

Part 2

The Universality of Categories and Their Functions



The Primacy and Fate of Predicativity in Tupi-Guarani

Francesc QUEIXALÓS¹

Introduction

Among the typologically interesting characteristics of the languages of the Tupi-Guarani family we find on the one hand the existence of a class of lexical roots that are relatively stable and homogeneous in terms of the elements it is comprised of, but not in terms of the functions it performs from one language to another, and on the other hand, the high frequency of a suffix whose role in the word cannot be easily identified in terms of inflectional or derivational morphology. In this study I attempt to show how a detour into diachrony clarifies facts in the light of recent proposals concerning lexical categories manifested by languages in other regions in the world, and how this detour leads to a unified interpretation of the two phenomena. To this end we must shift the emphasis from the celebrated but rather narrow noun/verb distinction in favour of the question of whether lexical entries are predicative or not.

A brief presentation of the distinction between lexical classes (Section 1) introduces the extent to which the syntactic functions of argument and predicate may be irrelevant in establishing lexical classes in various languages (1.1), and the reasons why this question is important in the Tupi-Guarani family

- 1 Centre d'Etudes des Langues Indigènes d'Amérique (IRD-CNRS) & Laboratório de Línguas Indígenas (Universidade de Brasília). This text is a considerably revised version of Queixalós (2001b). The contribution of Ana Suely Cabral and Aryon Rodrigues in terms of ideas and information is so pervasive in this study that it would be tedious to mention it on all the necessary occasions. To give one example, all the non-Emerillon linguistic data that are not referenced come from personal communications with A.S. Cabral or A. Rodrigues. I also received a great amount of help from Walkiria Praça. My deepest appreciation is extended to all three. I am also indebted to Mônica Borges, Yonne Leite, Dale Kinkade, and Gerald Taylor for additional comments and information. Any factual errors or conceptual shortcomings are entirely my own responsibility.

(1.2). The main body of the article consists of an illustration of the fact that in a certain type of grammatical systems the argument function is derived from the predicative, as is seen in several languages of the world (2), and within the Tupi-Guarani family (3). Kamayura (3.1) and Emerillon (3.2) introduce the study of the derivational process and of the use of morphology to this end in Tupi-Guarani. This morphological material is examined in terms of its nature (3.3) and its diachronic destiny, both from the phonological point of view (3.4.1) and the grammatical (3.4.2). The question of the copula and nominalising morphology as means of access to, respectively, the predicative and argument functions is addressed before the conclusion is reached on the main theme of the article, now from a diachronic point of view: the distinction between nouns and verbs (4).

1 Nouns and Verbs

Linguists try to distinguish between nouns and verbs in at least four different ways. On the *semantic* level, ideas centre around two prototypical poles: one of these refers to spatial configuration and temporal stability, the other to temporal configuration with no spatial dimension (Givón, 2001). This polarisation provides the bases for a contrast between entities and events.² On the *formal* level, it shows very generalised affinities between certain roots and certain categories that are expressed by means of explicit material, for example, gender-class on the one hand, time-aspect on the other. On the *pragmatic* level, some classes of roots are more apt than others for introducing manipulable participants into the discourse and others more apt for reporting events (Hopper & Thompson 1984). From a conceptually more homogenous point of view, a specialisation may be observed, sometimes concentrating on the expression of theme(s), sometimes on the expression of the rheme, which would eventually map onto the level of *function* by means of the distinction between argument and predicate. Mappings between these levels are common cross-

- 2 Chinese seems to be especially eloquent in terms of the correspondence firstly between noun and space and secondly, between verb and time (Paris 1984).

linguistically, not only between two given levels, but also between the four levels taken together (e.g. entity / marked for gender/theme/subject). These affinities constitute the basis of the distinction between *noun* and *verb*. Unfortunately for the convenience of our intellectual schemata, several linguistic systems show, when closely observed, that some affinities are not as natural as they may appear in the light of the presuppositions indicated above.

1.1 *Argument and Predicate*

Since Swadesh's article on Nootka, it has often been thought that a language that contains constructions such as

- (1) *mamo-k-ma* *qoʔas-ʔi*
 Work-3SingularIndicative *Man-Definite*³
 'the man works'
- (2) *qoʔas-ma* *mamo-k-ʔi*
 Man-3SingularIndicative *Work-Definite*
 'the worker is a man'

does not recognise the distinction that other languages make between nouns and verbs⁴ (Swadesh 1939:78). In the same way we find in Tongan (Tchekhoff 1984):

- (3) *na'e* *si'i* 'a e pepe
 Past *Smallness* *Absolutive* *Article* *Baby*
 'the baby was small'
- (4) *na'e* *kai* 'a e si'i
 Past *Food* *Absolutive* *Article* *smallness*
 'the boy was eating'

What we find here is the role that the functional predisposition of the lexeme plays in identifying classes. In languages of this type all the roots that belong

3 The segmentation of the definite is taken from Jacobsen (1979).

4 Boas (1911) had already drawn attention to a similar situation in Kwakiutl, from the same Salish family. See Davis & Saunders (1997:83) concerning an identical configuration in a language that is geographically close but which has no genetic bond, Bella Coola.

to an open inventory – ‘lexemes’ – are predicative in nature.⁵ What is less common, and what Nootka and Tongan demonstrate to a certain extent, is that these same roots also tend, at first sight, to have the role of argument.

Languages of the same class as Nootka and Tongan, such as Tagalog, Chercassian, Nahuatl, Cayuga and many others (cf. Lazard 1999), show that the predicative function cannot be taken as a serious criterion for distinguishing between nouns and verbs. Many of these languages also show that semantic categories with an explicit morphological expression, such as referentiality-definitude or aspect-tense, often actually constitute the core of criteria that allow the identification of two classes of roots in the lexicon.⁶

1.2 *Argument and Predicate in Tupi-Guarani*

The question is doubly essential within this South American family of languages.⁷

1.2.1 The Class of States

First, the question is essential because this family presents a class of lexical morphemes, referred to as ‘descriptives’ in the literature, with the following characteristics.

- 5 With the frequent exception of proper nouns which form a class of designations together with the deitic morphemes.
- 6 This is what Jacobsen (1979) claims concerning Nootka, in contrast to Swadesh’s opinion. It is also the position of Launey (1994).
- 7 Tupi-Guarani constitutes, in terms of number of languages – about 45 – the largest of the ten families branching out from the Tupi phylum. Predominantly located in the Amazon basin, it spreads from the north of Argentina to French Guyana, and from the western tributaries of the Orinoco to, formerly, the east coast of Brazil. On the basis of phonological and grammatical correspondences, its languages are gathered into eight groups, as follows (languages mentioned in this study are in italics): 1) *Old Guarani*, Kaiwa, Nandeva, *Paraguayan Guarani*, Mbya, Xeta, Tapiete, Chiriguano, Izoceño, Guayaki; 2) Guarayo, Siriono, Hora; 3) Tupi, São Paulo Lingua Geral, *Tupinamba*, *Amazonia Lingua Geral*; 4) *Tapirape*, *Tocantins Assurini*, Parakanã, *Surui*, *Ava-Canoeiro*, *Tembe*, *Guajajara*, Turiwara; 5) Arawete, Ararandewara-Amanaje, Anambe of Cairari river, Xingu Assurini; 6) *Kayabi*, Apiaka, Parintintin, Tupi-Kawahib, Juma; 7) *Kamayura*; 8) *Wayampi*, Wayampipuku, *Emerillon*, *Jo’e*, *Urubu-Ka’apor*, Ehrenreich Anambe, Guaja, Awre, Awra, Takunhape (Rodrigues & Cabral 2002).

At the semantic level it is quite homogeneous. A sample follows, taken from Kamayura (Seki, 2000):

- | | | |
|-----|------------|--------------------|
| (5) | atua'i | "low" |
| | je'ya | "high" |
| | huku | "long" |
| | akup | "hot" |
| | atã | "hard" |
| | pyw | "soft" |
| | jup | "yellow" |
| | tsiŋ | "white" |
| | pitsun | "black" |
| | oryp | "happy" |
| | koay | "annoyed" |
| | myrã | "old" |
| | pyau | "new" |
| | katu | "good" |
| | arõ | "pretty, pleasant" |
| | tykwara'ip | "rushed" |

Class (5) is stable from a diachronic point of view in the sense that all its meanings belong to a given formal class in a given language. However, in certain documented languages this class is considered to be made up of nouns (Rodrigues 1996a; Dietrich 2001), while in other languages it is seen as having come to make up a special class of intransitive verbs, the statives (Seki 1990; Leite 1990). Whether categorized as verbs or nouns, the elements of this class perform a predicative function. Emerillon (Couchili et al. 2002) is the only language in this family in which this class has split into two sub-classes that are also, internally, more or less semantically homogeneous: that which describes internal states that determine mental and/or physical states of being, such as

- | | | |
|-----|-----------|---------------|
| (6) | awu | "speech" |
| | baʔewal | "hunger" |
| | batsutsuk | "instability" |
| | bopil | "desire" |
| | kalai | "fever" |
| | kaneʔõ | "fatigue" |
| | katiʔi | "selfishness" |

nũʔāng	“cold”
oli	“pleasure”
uwedj	“thirst”

and that of permanent or temporary physical conditions such as:

(7)	tukuk	“short”
	atã	“hard”
	bik	“black”
	tawa	“yellow”
	tipi	“deed”
	puku	“long”
	tsikaĩ	“small”
	kuning	“on its side”
	tsĩng	“white”
	tsö	“fat”

The first of these sub-classes is similar but not identical to that of nouns. The second has no properties of nouns nor of what might be called an adjective class, without going as far as to merge with verbs: although this class appears to have most affinity with the predicative function, it needs additional grammatical material to perform this function.

1.2.2 The Primacy of Predicativity

The second reason for the non-criterial nature of the predicative function being central in Tupi-Guarani is the assumption that the members of this family are descendants of a language in which all lexical entries were predicative. Other contemporary and genetically independent (groups of) languages such as Salish (Kinkade 1983), the Philippine languages (Lemaréchal 1989, 1991) and Nahuatl (Launey 1994), have been described as showing this pattern of morphosyntactical organisation for which Launey coined the term ‘omni-predicativity’ (1986). The most important consequence of this pattern is that there is no lexical class in which the argument function is primary. This has to be derived.⁸

8 In 1968, Bach drew attention to the possibility of and the interest in describing English in this same way.

2 Derived Arguments

Languages of this type diverge in terms of the way in which what seems to be, at first sight, the acquisition of argument function is performed. A dichotomy appears depending on whether this process shows, by itself, a difference between the two classes, or not. Sikuni (Queixalós 2000) shows a notable asymmetry in the way in which a predicative root adopts the argument function.

- (8) pebi newüthü
 Man Jaguar
 'the man is a jaguar'
- (9) newüthü aitañibi
 Jaguar BeDrunk
 'the jaguar is drunk'
- (10) pe-aitañibi-nü yawahiba
 3-BeDrunk-Masculine Dance
 'the drunk man is dancing'

The roots of the class newüthü, 'jaguar', take on the argument function as in (9), with no addition of morphological material. The roots of the class aitañibi, 'to be drunk' take on the function of argument as in (10) by means of adding morphological material: a personal prefix and a gender suffix.

A more subtle means of showing the asymmetry between the two classes is seen in Nootka and Nahuatl. The Nootka roots of the class qoʔas, 'man', may appear as an argument without the definite -ʔi.

- (11) mamok-ma qoʔas
 Work-3SingularIndicative Man
 'a man works'

This means that the suffix is not the morphological trigger of function change. But, at the same time, the roots of the class mamok 'to work', have to join the -ʔi in order to perform the role of argument (Jacobsen 1979). In these roots, the typically nominal grammatical distinction of definitude is neutralised as an inflection. It becomes a mere marker of change in function. In the same way, Nahuatl (Launey 1994) has:

- (12) chōca in piltōntli
 Cry *Definite* Child
 ‘the child is crying’
- (13) piltōntli in chōca
 Child *Definite* Cry
 ‘the one who is crying is a child’

with the definite in indicating the argument function of the following term. However, as an argument, only the class of piltōntli ‘child’, to which the ‘grasshopper’ of (14) belongs, can dispense with the definite:

- (14) temōquê chapōltin
 GoDown Grasshoppers
 ‘grasshoppers landed’

The asymmetry between the two classes of elements in relation to a marker such as the definite shows that there does indeed exist a lexical difference between qoʔas and mamok in Nootka, and between piltōntli and chōca in Nahuatl, which comes not from their membership of a specific class of potential predicates or arguments – since both classes are primarily predicative – but from their affinity with semantic categories which, in many and different languages, differentiate between nouns and verbs by means of inflectional morphology (basically, extensionality for nouns and tense-aspect for verbs).

As for Tagalog, it applies the same treatment to all its predicative roots.⁹ (Lemaréchal 1989).

- (15) tumakbo iyon
 Run ThatOne
 ‘that one ran’
- (16) bahay ito
 House This
 ‘this is a house’

When a root of the sub-class of tumakbo, ‘to run’, appears as an argument, it is introduced by a proclitic ang, which performs the role of signalling the change of function.

9 From which deitics and pronouns are excluded.

- (17) pumatay ang nagnakaw
 Kill Proclitic Steal
 ‘the thief¹⁰ killed’

The same has to occur with a root taken from the sub-class bahay, ‘house’, such as bata, ‘child’:

- (18) tumakbo ang bata
 Run Proclitic Child
 ‘the child ran’

Several scholars relate omnipredicativity to one of the characteristics pointed out by Klimov (1974) for active-stative languages, and by Jelinek (1984) for non-configurational languages: pronominal affixes to the predicate constitute the expression of arguments. Similarly, phrases with a nominal lexical head that relate, referentially, to these affixes are adjuncts that bear no grammatical relation to the predicate. For Lemaréchal (1991), analysing the Palau language of the Philippines, these phrases are ‘expansions’ (a term inherited from Martinet). Kinkade (1983), along with other authors, observes that in the Salish languages, all lexical entries are predicates and considers the relationship between the adjunct – or descriptor – phrases¹¹ and the pronominal arguments to be appositive in nature. In a later study, Jelinek (1993) notes that another Salish language, Lummi, categorises almost all words as predicates. The presence of these predicates in the sentence in a non-predicative function occurs through generation of a determiner phrase introduced by a determiner *cə*:

- (19) k^wey'-Ø cə nə-ŋənə¹²
 Hungry-3Absolutive Determiner My-Child
 ‘he is hungry, (the one who is) my child’

10 I have simplified Lemaréchal’s translation: “the one who stole”.

11 The term ‘adjunct’ seems to collapse into a single function that of the lexically headed noun phrases, and that of adverbial or oblique phrases, on the basis that neither of the two phrase types bears a direct grammatical relationship to the predicate. However, the first of these functions deals only with the specification of the reference of core participants, while the second deals with the specification of circumstances, manner, space-time localisation or peripheral participants. In order to highlight this difference, I put forth the term *descriptor* as best suited to indicate the phrases that identify the reference of pronominal arguments.

12 I ignore the distinction between affixes and clitics.

- (20) nə-ŋənə-Ø cə k'wey'
 My-Child-3*Absolutive* *Determiner* Hungry
 'he is my child, the (one who is) hungry'¹³

Up to this point there is no difference here with what has been seen above in Tongan, Nootka or Nahuatl. The interesting point in Jelinek's proposal is the reaffirmation, within the context of non-configurationality, of an idea previously suggested by Thompson & Thompson in 1983, and also by Launey in 1986, namely, that the descriptor phrases are predicates subordinated to the main predicate, as the translations of the examples suggest. Elaborating on this idea of subordination, Launey goes further in identifying the referential link between descriptor phrases and pronominal arguments (1986, 1994). If the descriptor phrase is a subordinated predicate, it must contain the pronominal affixation required by any predicate. In

- (21) ka Ø-λa'toa in Ø-tīsiλ
 Assertion 3*Subject-Speak* *Determiner* 3*Subject-Doctor*
 'the doctor speaks'

and in contrast to what Jelinek (1984) claims, the co-referential relationship is not between the subject prefix of the verbal predicate 'speaks' and the descriptor phrase 'the doctor' taken as a whole, but between the subject prefix of the verbal predicate 'speaks' and the subject prefix of the nominal predicate heading the descriptor phrase, '(be) the doctor'. When we change the person, it becomes quite clear who co-refers with whom. The Sikuani example

- (22) newüthü-mü aitañibi-mü
 Jaguar-2*Subject* BeDrunk-2*Subject*
 'you, jaguar, are drunk'

clearly indicates that (9) should be read as

- (23) newüthü-Ø aitañibi-Ø
 Jaguar-*Subject*3 BeDrunk-*Subject*3
 'the jaguar is drunk'

A detail related by Kinkade (1983) shows that the idea that descriptor phrases are subordinate predicates has psychological reality. For a statement like:

13 The difference in the position of brackets is in the original text.

- (24) xes-ĩc'e? x^we c'i?
 Good-Flesh The Deer
 'venison is delicious'

a Cœur d'Alène native speaking writer gives the following English rendering: "[They are good to eat] [those which are deer]" (brackets added, FQ), which we may compare with Jelinek's translations of (19)–(20).

There are two consequences of this: 1) the main predicate, provided with its lexical head and pronominal arguments, is a morphosyntactically complete sentence; and 2) an utterance with more than one lexical word is in fact a complex sentence made up of various predicates.

3 Derived Arguments in Tupi-Guarani

The whole family has clear affinities with what we have just seen¹⁴. And these affinities, in my view, go a long way towards explaining several basic properties of the grammar of these languages and its history.

3.1 An Example from Kamayura

Let us look at the following data taken from Kamayura (Seki 2001: 56; 2000: 72), paying particular attention to the suffix -a, glossed as X:

- (25) kunu'um-a ka'i-a r-uwaj-a w-ekyj
 Boy-X Monkey-X Contiguity-Tail-X 3-Pull
 'the boy is pulling the monkey's tail'¹⁵

14 For greater ease of reading what follows, I shall not go into the question of Jelinek's argument type (see Vieira 1993; Leite 2001 for this discussion on Tupi-Guarani languages). This decision does not impinge on the issue at stake.

15 'Contiguity' and 'non-contiguity' refer to paradigms of prefixes whose function, according to Rodrigues (1990; 2001) and various other specialists, among them Cabral (1999), consists in indicating whether or not an internal complement is immediately adjacent to its head. Some authors, such as Seki, only recognise the existence of the paradigm of

- (26) ko-a wi e'ym a-jot
 Garden-X Ablative Negation 1 Singular-Come
 'I am not coming from the garden'

Note the suffix accompanying the nouns in the following functions: subject ('boy'), object ('monkey's tail'), adnominal complement ('monkey'), postposition complement ('garden'). Specialists in languages of the Tupi-Guarani family have found this suffix in almost all the languages, but its nature has given rise to different names: 'nominal, or rather, nominalising suffix' (Barbosa 1956), 'nominal index' (Rodrigues 1953, Barbosa 1956), 'argumentative', 'nominal', 'nominative' 'case' (Rodrigues 1996a; 2001), 'onomastic case' (Adelaar 1997), 'nominal function marker' (Seki 1990), 'nominal case' (Jensen 1999), 'nuclear case' (Seki 2000).

Now let us compare the distribution of the -a suffix in (25) and (26) to the distribution of the a marker, also glossed as X, placed before nouns in examples of the Palau language belonging to the same family as Tagalog (Lemaréchal 1991):

- (27) ng-remúrt a ngáleḵ
 3-Run X Child
 'the child runs'
- (28) ak-sileḵbek-íi a ngáleḵ
 1-Kick-3 X Child
 'I kicked the child'
- (29) blí-l a seḵeḵlí-k
 House -3 X Friend -1
 'my friend's house'
- (30) ak-milsúub ɛr a skúul
 1-Study Locative X School
 'I study in school'

The similarity between Kamayura -a and Palau a is striking: both morphemes are associated with

contiguity. I have made the glosses uniform for this latter paradigm, without this having any implication concerning my own position with regard to this question.

- a subject in (25) and (27);
- an object in (25) and (28);
- an adnominal complement in (25) and (29);
- an adposition complement in (26) and (30).

Lemaréchal attributes to the Palau a the same function as to the Tagalog *ang*, that is, to indicate that the lexical item associated with it no longer performs the predicate function, or, in the words of Thompson & Thompson, Launey or Jelinek, that the lexical element has become subordinate to another predicate. I propose the hypothesis that this is also the function of the *-a* in Tupi-Guarani. The Tupi-Guarani and the Philippine languages are original in that they have a specialised morpheme performing this function – it is the precise linguistic equivalent of the iota operator used in propositional calculus – while in Tongan, Nootka and Nahuatl it is a typical element in nominal morphology, the definites *e*, *-i*, and in respectively, that performs this role.

3.2 An Example from Emerillon

However, we may observe that alongside the pervasive *-a* in Kamayura – which I have come to identify with a term that I shall justify below: *referrer* – its cognate in Emerillon is much more limited in its distribution.¹⁶

(31) *ial* *oike*
 Canoe ItSinks
 ‘the canoe sank’

(32) *ial* *pepotat*
 Canoe YouWantIt
 ‘you want the canoe’

The subject and object functions are not associated to the referrer. However, it is present in the functions of adnominal complement and postposition complement:

¹⁶ Except where otherwise stated, the data for Emerillon in this article were taken from Maurel (1998), Couchili *et al.* (2001), or were supplied orally by Ti’iwan Couchili, Françoise Rose and Didier Maurel. My thanks to these three people. For /l/ at the end of a morpheme, I have unified the representation according to Rose’s (2000) phonological analysis.

- (33) ial-a kupatsi
 Canoe -*Referrer* Gunwale
 ‘canoe gunwale’
- (34) ailul enamiñ-a pe
 IBroughtItBack MyGrandfather-*Referrer* *Locative*
 ‘I brought it back to my grandfather’

As for the Wayampi language of the Oyapock region, it shows no trace of the referrer, neither in subject nor object (35), nor the adnominal complement (36), nor the postposition complement (37) (Grenand 1980: 75, 46, 62):

- (35) ε-mẽ kɔ w-aapɪ
 I-Husband Garden 3-Burn
 ‘my husband is burning the garden’
- (36) tayaʔu-ka ɔ-maʔẽ
 WildPig-Fat 3-Melt
 ‘the pig fat is melting’
- (37) ɔ-i-mɔ-kiʔa iwi le
 3-*Causative-Reflexive*-BeDirty Earth With
 ‘he is dirty with earth [on him]’

My hypothesis concerning the function of the -a in Tupi-Guarani has a diachronic corollary: Kamayura with its pervasive referrer, Emerillon with its referrer of limited distribution, and Wayampi in which the referrer is unknown, represent three successive stages of a single evolution that involves the whole family.

3.3 *The Nature of the Referrer*

Before going into details about this evolution, we need to observe minimal pairs such as those found in Kamayura (Seki 2000: 161–162):

- (38)a jetutyr-a morerekwat
 MyUncle-*Referrer* Chief
 ‘my uncle is a chief’

- b jetutyr-a morerekwar-á
 MyUncle-Referrer Chief-Referrer
 ‘my uncle is the chief’

Tagalog has the same minimal pair (Lemaréchal 1989, 1991; I extend the denomination “referrer” to ang):

- (39)a amerikano ang titser
 American Referrer Teacher
 ‘the teacher is an American’
 b ang aleman ang doktor
 Referrer German Referrer Doctor
 ‘the doctor is the German’

The absence or presence of the referrer *in the predicate* makes the difference between the two types of predication. The basic function of this marker, therefore, cannot be to change a predicate into an argument, or, to put it another way, to subordinate a predicate to another predicate. The difference between structures *a* and *b* lies in the properties of the predicate term with respect to reference. ‘Chief’ and ‘American’ in *a* do not refer, i.e. they do not hark back to any entity the speaker gives as existing. They only hark back to a class of elements defined, in comprehension, by a certain property, ‘being chief’, ‘being an American’. The relationship between subject and predicate in *a* is *inclusive*, which is the same as saying that the referent of the subject is identified as coinciding with any element included in the class defined by the expression functioning as predicate. As for [‘chief’-a] and [ang ‘German’] in *b*, they really do refer: they hark back to a class extensionally defined by its unique element which is given as an existing and individually identifiable entity. In *b* the relationship between subject and predicate is *equative*: the referent of the subject is identified with the *unique element* of the class included in the expression functioning as predicate.

In the light of the form of the predicates in examples (38) and (39), it is clear that the function of the referrer, far from consisting of deriving an argument from a predicative element, is limited to constructing a referent on a root which by itself is unable to refer because of its predicative nature. This fact is the reason for the frequency with which the referrer occurs in nominalisations – with or without a specifically nominalising morpheme – as in Tembe (Cabral 2001: 145):

- (40) izé-a a-kwáw karaí Ø-zeʔéŋ-a
I-Referrer I-Know WhiteMan Contiguity-Speech-Referrer
 ‘I know the white man’s language’
- (41) izé-a a-kwáw karaí Ø-zeʔéŋ-háw-a
I-Referrer I-Know WhiteMan Contiguity-Speech-Nominaliser-Referrer
 ‘I know the white man’s language’

This being the case, and in the same way that to a certain extent the difference between the two lexical categories noun and verb may be seen as a sub-product of a more fundamental distinction, that which exists between the two functions of argument and predicate, the distinction between argument and predicate should be seen as a sub-product of an even more fundamental difference, that which exists between the two basic constitutive operations of the act of speech: referring and predicating (Lyons, 1977: 429; cf. also the notions of indicativity and predicativity in Seiler’s typological model, for example, Broschart 1988). A basic dichotomy in the lexicon of languages is established between the class of elements whose function is to institute *predicates* and the class of elements whose function is to institute *referring expressions*. It is in the nature of omni-predicative languages for the second class to be extremely limited.

The characterisation of the referrer must be refined further. The following example from Tembe (Cabral 2001: 147)

- (42) kaʔí-a w-er-ekó huáz-a
Monkey-Referrer 3-CausativeComitative-BeInMovement Tail-Referrer
 ‘(the) monkey has a tail’

shows that the referrer is not building any reference, as I have just stated. The interpretation of this example could be ‘the monkey in question has a tail’, or ‘monkey is a class of beings whose characteristic is to have a tail’. In the second interpretation, neither ‘monkey’ nor ‘tail’ refer.

More clearly, we also have the referrer together with a non-referential ‘comb’ in Kamayura (Seki 2000: 309):

- (43) kunu'um-a kywaw-a o-'awyky-potat o-uw-a
 wite
Boy-Referrer Comb-Referrer 3-Make-Desiderative 3-Father-Referrer
Comparative
 ‘the boy wants to make a comb like his father’s’

and together with a non-referential ‘leaving’ in Tembe (Cabral 2001: 146):

- (44) izé-a a-putár né Ø-hém-haw-a
1-Referrer 1-Want 2 Contiguity-Leave-Nominaliser-Referrer
 ‘I want you to leave (lit.: ‘your leaving’)

We may say therefore that the referrer institutes an expression that is *capable of referring*, and not an expression *that in fact refers*. For this reason I find it preferable to substitute *referring expression* for the less literally explicit term *designation*, which is understood as an expression with the ability to refer. In noun+noun and noun+verb constructions it is this ability to refer that determines whether it is obligatory or impossible for the referrer to occur after the first element¹⁷. In Kamayura (Seki 2000: 376, 145, 365, 355) the first lexical element is not a designation in the compound

- (45) potap-yru
 Food-Recipient
 ‘stomach’

nor in the incorporation construction

- (46) na-je-'aju-wuku-ite
Negation-1Singular-Neck-BeLong-Negation
 ‘I do not have a long neck’

The first lexical element is a designation in genitive position

- (47) [...] o-'atywahaw-a pyr-im
3-Cousin-Referrer House-In
 ‘[...] in his cousin’s house’

whether it is referential as in (47), or not, as in (48):

- (48) jawar-a 'yhwāpē-her-a 'aŋ
Jaguar-Referrer Claw-Passado-Referrer This
 ‘this is a jaguar claw’

17 Here, the interpretation of the -a as a referrer is equivalent, in terms of explicative value, to the interpretation of the -a as a case marker.

Similarly the first lexical element is a designation in the position of a verb argument

- (49) je-'ajur-a n-i-huku-ite
 ISingular-Neck-Referrer Negation-3-BeLong-Negation
 ‘my neck is not long’

whether it is referential as in (49), or not, as in (43) – (44).

It is easy to understand the affinity of Tongan, Nootka and Nahuatl for the definite particle when the lexical element is no longer the predicate of the sentence: the definite particle performs in principle, the role of a reference generator. Equally, it is clear why in Tagalog, the only morphemes that are incompatible with *ang* are demonstratives and pronouns, which are already, by their nature, designations. Rose's two observations concerning Emerillon thus come to be particularly significant, linking the referrer *-a* to the demonstrative *aʔe*: 1) in certain constructions the referrer varies freely with a *aʔe* form (119); 2) in other constructions it is the demonstrative that varies freely with an *a* form (189). Likewise, Emerillon has a prefix with a pronominal deictic value occurring in a position typical of nouns – complement of a postposition – and whose phonological form is a (Maurel 2001).

- (50) a-l-ehe-te oho
 That-Contiguity-Sociative-Emphatic HeLeaves
 ‘that is why he left’

Clearly we need to know more about the etymology of the referrer.

3.4 *The Diachronic Scenario*

Within the omnipredicative pattern there is a notable cross-linguistic uniformity in the class of designations: this class is usually made up of deictically based elements and at times also of names. The greatest degree of omnipredicativity in a language is shown in the sparseness of this designation class. In Nahuatl just two deictics make up the class (Launey 1994: 282). In Nootka, the class is a little broader: it includes names, pronouns and some deictic adverbs (Swadesh 1939). In Apinaje (Cunha 2003) and in Palau (Lemaréchal 1991) names fall outside the class. In the Salish languages names and pronouns are outside this

class (Kinkade 1983). Tupi-Guarani has a strong degree of omnipredicativity comparable to that of Nahuatl: according to Cabral (2001: 131) its designation class is reduced to a few deitics. In Kayamura, for example, names are outside this class (Seki 2000: 187):

- (51) Sapaĩ-a a-enõj moĩ-a juka-tar-am
 Sapai-Referrer 1Singular-Call Snake-Referrer Kill-Nominaliser-Attributive Case
 ‘I called Sapai for him to kill the snake’

To be more precise, strong omnipredicativity cannot be attributed to Tupi-Guarani as a family because the documented languages present varying degrees of scarcity in the class of designations, as is shown in the distributions of the referrer in the Kamayura examples (25)–(26), Emerillon (31)–(34), and Wayampi (35)–(37). Let us look at this in more detail.

In his 1999 study, Jensen states that in proto-Tupi-Guarani the noun can function as a predicate, taking as the only clue of the derivation the absence of -a (149). In his 1998 work, Jensen states that in Kayabi the -a has spread to roots ending in a vowel (506). The same writer takes up again the spirit of this statement in his 1999 work, in which she mentions the extension of this morpheme to all the nouns in Tocantins Assurini (149).

I assume, on the contrary, that this suffix has never spread from one context to another – except in a very particular case, reported below –, but that it was omnipresent at a certain period and since then has simply receded until it has completely disappeared in some languages. I shall call this period, which only comparative study can place in a relative chronology within linguistic evolution, the *initial stage*. Cabral (2001: 135, 142) shows that the suffix was already present in proto-Tupi-Guarani. The same writer provides strong evidence that what I call the initial stage occurred beyond proto-Tupi-Guarani. Observations relating to pronunciation and mentioned by Cabral (slow or rapid articulation, generations co-existing in time, testimonies from tens or hundreds of years ago, and close varieties of a single language) are eloquent with respect to the sense of evolution and indicate that in certain cases, this evolution is still going on. For example, the Wayampi variety studied by Coudreau at the end of the 19th century still has the -a very much in evidence. Today it no longer exists.

We may think of the initial stage in the following way.

(52) *basic property:*

- 1) all or almost all the lexical roots are predicative¹⁸

corollaries of 1):

- 2) a referrer -a derives designations from these roots;
- 3) there exists a small lexical class of designations; it never combines with the referrer,

corollary of 2):

- 4) each constituent in the functions of
 equative predicate
 subject
 object
 adnominal complement
 adverbial complement

which does not belong to the class of designations, has to occur in combination with the referrer.

The subsequent story – which we are to a certain extent able to observe – is that of the erosion of this architecture. The documented languages – living ones such as Kamayura, or extinct ones such as Tupinamba – are testimony to the different stages that the disaggregation of the initial system has reached. Each amputation observed – each lacuna in the supposed initial distribution of the referrer – should be seen as symptomatic of one and the same process: the loss of omnipredicativity as a basic principle of the organisation of the grammatical system. It is important to point out two independent evolutionary factors which, to a certain extent, come together to produce the same result. The first is strictly grammatical: the loss of omnipredicativity causes the loss of the referrer's functionality, making it vanish completely or lose its status as a morpheme. The second goes beyond the boundaries of grammar. Because of its position within the word and its phonological make-up, the referrer is weak and its phonological erosion causes it to disappear easily. Naturally, it only ceases to exist phonologically because its grammatical role has become obsolete. For greater clarity in the synopsis I shall now propose, I shall show separately the effects of phonological evolution and grammatical evolution.

18 This does not exclude the possible existence at that stage of non functional differences between verbal roots and nominal roots (as in Nootka and Nahuatl).

With the same end in mind, I shall give a single example of a language for each stage identified, without having to conclude from this that the stage in question is only found in the language cited. Unless specifically indicated, the data for languages other than Emerillon come from Cabral (2001).

3.4.1 Phonology

The phonological history of the referrer is traced in detail in Cabral (2001). In brief: the more the segment preceding it resembles it, the more likely the referrer is to disappear. A variety of Tocantins Assurini represents the inheritance that keeps most of the initial stage, as it shows the referrer in all its phonic contexts. The following is an example in which, according to Harrison (1960) quoted in Cabral (2001), it occurs after a:

- (53) i-pepá-a
Non-Contiguity-Wing-Referrer
 'its wing'

The other variety of the same language illustrates the first stage in the disappearance of the referrer for apparently phonological reasons, diluting the suffix when the vowel a precedes it, but not when other vowels do (Cabral 1998, quoting Nicholson 1976):

- (54) arawasá-ø ereká h-ye'í-a ryrúramu
Sieve-Referrer YouAreMoving Non-Contiguity-Intestines-Referrer AsAReceptacle
 'you have a sieve as a receptacle for his intestines'¹⁹

The Tupinamba language (Rodrigues 1996a) accentuates this tendency and suppresses, after any vowel, the phoneme that carries the suffix:

- (55) kwesé pajé-ø maʔéasíβór-a suβáni
Yesterday Shaman-Referrer SickMan-Referrer Suck
 'yesterday the shaman sucked the sick man'

Wayampi goes further still: by eliminating final consonants, it allows the existence of only contexts contrary to the expression of the referrer. We have already seen examples of the Oyapock river dialect. In the example below, from

19 -ø represents, for some scholars, the phonologically null allomorph of the referrer.

the Jari river dialect (Jensen 1998), a final *r* has disappeared from the last two terms, in which the suffix should be found in the initial stage²⁰

- (56) amẽ jawĩ omimoĩ tapi'i rookwe
 Then Turtle HeCooks Tapir Meat
 'then the turtle cooks tapir meat'

* * *

Digression. As a sub-product of what has been stated above, we can now understand why most authorities consider the referrer to be a case marker.²¹ This idea is based on the distribution, and especially the non co-occurrence of the referrer with case markers. Let us look at these case markers in two languages (respectively, Rodrigues 2001 and Seki 2000). The allomorph depends on whether the case mark follows a consonant or a vowel (I simplify the list of allomorphs).

(57)	Tupinamba		Kamayura	
	...C-	...V-	...C-	...V-
translative	-amo	-ramo	attributive	-am -ram
locative 1	-ipe	-pe	locative	-ip -p
locative 2	-iβo	-βo		
locative 3 ²²	-i	-j		

I take Tupinanba as an example of what may have happened in the languages in which the referrer forms, distributionally, a paradigm with the markers of (57). Let us consider the possibility of a time in which these markers occurred

20 There is a possibility that the subsequent phonic context works as a factor favouring the phonological extinction of the suffix. Cabral (2001) refers to this in Jo'e and in Tembe in rapid speech. The same author suspects that it exists in Emerillon. Cf. similarly Seki (2001) for Kamayura .

21 Vieira (1993) is unusual in considering the suffix to be a nominaliser. The referrer hypothesis has of course more affinity with Vieira than with other authors.

22 The denominations of the different locatives in Rodrigues (2001) are, in the following order: punctual, diffuse and situational.

in the form of their base allomorph, presumably -ramo -ipe, -iβo, -i²³. Let us consider furthermore the possibility that at this time the referrer was fully active. We have to admit the strong possibility of the following sequences:

- (58) a ...C-a-ramo
 b ...V-a-ramo
 c ...C-a-ipe
 d ...V-a-ipe
 e ...C-a-iβo
 f ...V-a-iβo
 g ...C-a-i
 h ...V-a-i

The previous co-occurrence of the referrer with the markers of (57) is independently confirmed in the preservation of at least the reflex of the sequence ...C-a-(i)pe in two types of language. Emerillon represents the type in which the referrer has been partly lost. One of the contexts in which the sequence appears is found in (Rose 2003):

- (59) Kayen-a-pe ikem̃n za-iko-nam
 Cayenne-Referrer-Locative Previously IndeterminatePerson-Live-When
 ‘when, previously, we lived in Cayenne’

Paraguayan Guarani is an example of the type in which the -a has come to be part of the phonological make-up of the roots. The root ‘house’, as it occurs in the second word of the following example (Correa 1981, quoted in Velázquez-Castillo 2002), shows the reflex of the same sequence:

- (60) upe óga-pe-gua temi-mo-ngakuaa [...]
 That House-Locative-From NominalisingResultative-Causative-Big
 [...] i-kuñatai ramo-ramó-va²⁴
 Inactive-YoungWoman Just-Just-Relativizer
 ‘someone raised in that household who was just beginning to become
 a young woman’

23 The first two are in accordance with Seki’s options. The last two are my own assumptions.

24 In A. Rodrigues’ interpretation, the first ramo is a case suffix of the preceding noun.

The apparent incorporation of the referrer into the case paradigm is a mechanical consequence of the conjuncture of two facts: its functional weakening and the characteristics of the contexts (58)*b–h* that make it phonetically precarious.²⁵

Two more points must be made to complete this digression, which is important because the nature of the -a suffix is the cornerstone of the overall hypothesis. First, to provide an understanding of why the sequence of (58)*a* is not maintained following the example of the preservation of the -a in the contexts given in example (59) from Emerillon. I can give no satisfactory answer to this, but perhaps we should presume that the a-ramo sequence never occurred because ramo is not – or was not, during the period in question – a case marker. Rose (2003: 334–341) gives a detailed demonstration of how untypical its reflex -am is as a case marker in Emerillon. Second, we have to consider the ‘non-marked’ case that Seki identifies in Kamayura. It takes the form of -Ø. Its functions are to mark the noun as predicate,²⁶ as a dislocated constituent, as a vocative, or as a citational form. There is no common point between these functions except that the noun is simply not marked for case, either because it is the head of the predicate or by being located outside clause syntax. *End of digression.*

* * *

3.4.2 Grammar

In the grammatical history of the referrer, the conservative inheritance is also represented by Tocantins Assurini. The language has kept the referrer in the different functional contexts that we may assume to have existed in the initial stage, such as that of the subject and the adnominal complement (Cabral 1998):

- (61) sahý-a r-uwý-a utururú hehé
 Moon-*Referrer* Contiguity-Blood-*Referrer* ItDrops OnHim
 ‘the blood of the moon dropped on him’

25 In relation to the sequences *c*, *e*, *g*, see note (20) on the possible effect of the subsequent vowel context.

26 ... non-equative [FQ].

and equally in the equative predicate²⁷

- (62) Apetyráw-a txé r-ér-a
 Apetyráwa-Referrer 1 Contiguity-Name-Referrer
 ‘Apetyráwa is my name’

The fact of Tocantins Assurini being considered conservative in its phonology and its grammar does not mean that there is necessarily any superposition of the two evolutions in the same language. Cabral (2001) argues in favour of this disassociation and indicates the Guajajara, Ka’apor, Surui, Emerillon, Jo’e languages, and a variety of Tembe, as having experienced a total or partial loss of the referrer without any phonological conditions that would favour this process – basically the loss of final consonants – having occurred.

The vitality of the referrer in Tocantins Assurini is also clear through the regularity with which the suffix is capable, on its own, to nominalise verbs:

- (63) tʃé Ø-kér-a
 1 Contiguity- Sleep-Referrer
 ‘my sleep’

We have seen the nominalisations in Tembe where the referrer acts alone or in combination with a nominaliser (examples (40) and (41)). In other cases the referrer no longer has the ability to perform nominalisation on its own. The previous example:

- (64) izé-a a-putár né Ø-hém-haw-a
 1-Referrer 1-Want 2 Contiguity-Leave-Nominaliser-Referrer
 ‘I want you to leave (lit.: ‘your leaving’)

is not grammatical without the -haw nominaliser.

Kamayura goes further in losing the vitality of the referrer in nominalisations, without this fact, however, affecting other functions. The presence of the referrer in verb roots in a non-predicative function has serious restrictions (Seki, 2000: 121–123): it is not truly productive,²⁸ and of the eight nominaliser morphemes of Kamayura, none combines with -a.²⁹

27 A.S. Cabral, personal communication.

28 “*There are situations* (my emphasis, FQ) in which verb roots occur in noun positions, marked by the suffix [-a]”.

29 With the exception of the nominaliser -ama’e if this phonological sequence were to be analysed as I do in the example from Emerillon which follows.

The second stage is represented by Emerillon. The argument functions no longer show any trace of the referrer, as is seen in examples (31) and (32). Furthermore, it is no longer totally self-sufficient in nominalisations, in which it has to be associated with the nominaliser *-maʔe*:

- (65) o-wata-tat-a-maʔe
 3-Walk-Future-Referrer-Nominaliser
 ‘future walker’

The other contexts in which the referrer is still observable in Emerillon are the adnominal complement, example (33), and the postposition complement, example (34)³⁰. A possible reason for preserving the referrer in nominalisations, in the adnominal complement and in the postposition complement, will be suggested below.

The final stage of this development is represented by Guajajara (Harrison 1986). There is no longer any trace of the referrer but the disappearance of the latter cannot even superficially be attributed to phonological reasons. If the following example belonged either to the initial stage or to the stage where deletion after vowel occurs, the referrer would appear in the second and third words:

- (66) u-munyk t-azyr i-petym Ø-eraha i-zupe aʔe
 3-Light 3-Daughter 3-Tobacco 3Singular-Taking 3Singular-To 3°
 ‘his daughter lit his cigar and took it to him’

Under the same conditions, it would occur after the first morpheme in

- (67) mykur-rekuhaw-pe
 Possum-Dwelling-At
 ‘at possum’s place’

We return to Emerillon, recapitulating the occurrences of the referrer:

adnominal complement

- (68) ial-a kupatsi
 Canoe-Referrer Gunwale
 ‘gunwale of the canoe’

30 Together with a few more contexts mentioned in Couchili et al. (2001) and in Rose (2003: 116), one of which I shall deal with below.

postposition complement

- (69) ailul okal-a bitep
 IBroughtItBack VillageSquare-Referrer Locative
 'I brought it back to the village square'

nominalisation

- (70) o-wata-tal-a-maʔe
 3-Walk-Future-Referrer-Nominaliser
 'future walker'

and in the plural construction of the noun, introduced in the following example:

- (71) enamiñ-a-kom
 Grandfather-Referrer-Collective
 'grandparents'

Couchili *et al.* contend that in (71) kom should be considered as the head of the nominal phrase, giving the meaning 'group of' rather than contributing as a plural marker. On the other hand, the nominaliser maʔe could have nominal origin of generic meaning. In several languages of the family this morpheme means 'thing'. Thus, the constructions in (68), (70) and (71) are the same: a phrase made up of an adnominal complement followed by its head, 'gun-wale', 'thing' and 'group' respectively.³¹ This order within the noun phrase is fixed in Emerillon (Rose, 2003: 557). I suggest that the Emerillon referrer was preserved only in the contexts in which it occurs bound in strictly ordered sequences of morphemes, either in the nominal phrase or in the postpositional phrase. Other occurrences of the referrer also show this characteristic of insertion in frequently recurring and fixed sequences of morphemes³². Subject and object phrases do not always appear closely preceding the verb (Rose, 2003: 554), so they do not offer the same morphological or intra-phrasal 'protection' to the occurrence of the referrer.

If the reason for the fragmentary distribution of the referrer in this language is as has just been described – there are no functional reasons for preserving

31 The noun 'thing' and the nominaliser could have different origins. However, a nominal etymology for the nominaliser maʔe would explain why it follows the referrer instead of preceding it as it does in the other suffixing nominalisations, e.g. (44).

32 Basically before subordinators and some clitics (Rose 2003: 116).

the referrer where it occurs – we are led to draw far-reaching conclusions about the diachrony of omnipredicativity in Tupi-Guarani. It is possible that there still exist today quite omnipredicative languages within the family. Tocantins Assurini could be one of these. It is very probable that the family also has languages significantly less omnipredicative, such as Guajajara. The languages that are in an intermediary stage between Assurini and Guajajara, that is, those like Emerillon in which the presence of the referrer has already begun to diminish, are languages whose grammar has abandoned the characteristic of omnipredicativity. And because the referrer has lost its functionality, it disappears, revealing itself only where it is still protected, ‘encysted’ in the chain of morphemes. This explanation also accounts for the purported cases of weakening and loss due to the phonological environment. Based on a consensus in which the loss of the suffix was seen as a result of diachronic processes of weakening that affect phonemes at the end of words, Cabral (2001) insists on the pre-eminence of purely grammatical reasons for this development: the reduction “[of] the predicative properties of nouns and [of] the possibilities for verbs and descriptives to function as arguments”. In my terms, the loss of omnipredicativity is what in the final instance opens the way for the well documented phonological erosion of the final part of words, through the disappearance of a referrer that no longer has a function. Seen from this angle, the Emerillon -a is no more than a relic with no grammatical status, which has only been preserved where it prevents the occurrence of non-permitted phonemic sequences.

Languages like Guajajara which have simply stopped using the referrer without affecting the phonological constitution of the final section of words, show – against a background in which the process of the retreat of omnipredicativity may have taken a considerable number of centuries to be completed – how quickly linguistic change can occur, even in the field of grammatical architecture. The processes of phonological evolution are also rapid at times, although they can also be gradual. Surui might at first sight be considered representative of a conservative stage since in terms of the referrer this language is comparable to Tocantins Assurini. Ruth Monserrat’s field notes, quoted in Cabral (2001) give the following information – items within brackets are mine. [First phase] In the language of the oldest Surui speakers the suffix is audible after any segment except a. Among the young, however, various changes are found. In one and the same generation, I would emphasise. These are as follows. [Second phase] The suffix is audible only after consonantic segments. [Third

phase] After *r*, *-a* varies with *-ə*. [Fourth phase] After [other?] consonants *-a* varies with zero. It is clear that we are not able to reach a decision on whether the first phase, that of the living elders, constitutes the smallest observable sign that the process of losing omnipredicativity has been completed, or if the phase simply reflects an innocuous simplification of the phonic sequence ...*a*. But, in the first hypothesis the gradual nature of the phonological reduction of *-a* does not prevent the diachronic process from happening over a period of two or three generations; in the second hypothesis, it is the end of omnipredicativity as the basis of a grammatical system that has happened in such a short period. All of which does not mean that every grammatical property related to omnipredicativity has to disappear at the same rate. There is more in the history of modern Surui that can sustain the second hypothesis: Monserrat also observes that in the speech of the younger Surui, final *ə* extends to forms that were never associated with the *-a* suffix, such as finite verbs. In other words, the suffix, which is already grammatically opaque, is turning into a simple word final phoneme.

This fossilisation of the referrer is reminiscent of what has happened to three languages which have been most affected by historical events that have left them, to different degrees, genetically atypical: Paraguayan Guarani – a linguistic fusion of creole society and that of indians de-tribalised by the missions (Melià 1983: 48) –, *Lingua Geral* of Amazonia – a product of miscegenation between Indians and Portuguese spread by the missions of the Amazon and Rio Negro (Rodrigues 1996b) –, and Cocama or Omagua – a contact language imposed by the missions on the upper Amazon (Cabral 1995: 255). It is clear that in these languages, and due to the fact of their having been introduced into non-Tupi-Guarani populations, the loss of omnipredicative features has happened more quickly, as has the loss of other typically Tupi-Guarani grammatical characteristics. What makes Surui similar to these three languages is that the same phenomenon of what will constitute its fifth phase – the referrer, no longer having any grammatical function, will become encysted and come to be part of the lexical root – is found in them. Thus, they have the unsegmentable terms

Guarani (Rodrigues 1995)

- (72) ména ‘husband’
 óga ‘house’

Lingua Geral of Amazonia

- (73) ygára 'canoe'
tapiíra 'tapir'

and Cocama (Cabral 1995: 125)

- (74) uba 'father'
oka 'house'

It should be noted that these words, in the Tupi-Guarani language that gave rise to them (Old Guarani in the first case, Tupinamba in the second two), contained a referrer that was still productive. As in Surui, this feature of the initial stage has not passed on to new generations of speakers.

3.5 Questions of Chronology

One aspect that should be highlighted after this brief historical synopsis based on the hypothesis of an omnipredicative architecture in retreat, is the variation that can exist in the different rates at which different languages follow analogous paths of evolution. The rise of a *copula* in a few languages of this family will be taken as an example of this observation. One innovation resulting from the loss of omnipredicativity should be the generation of copulative elements in stages following the initial period, i.e. elements to be used to construct predicates with those items that can no longer do it for themselves.

It is well known that the copula is not a characteristic of the Tupi-Guarani family. However, Emerillon does have copulative morphemes with suppletive forms for third and first/second persons. Their use is obligatory in the subclass of absolute nouns – mainly the names of ethnic groups and animal species – and optional for other sub-classes of nouns.

- (75) dzawal o-tui tal
Dog 3-Copula Future
'it will be a dog'

- (76) i-dzēbulupa a-dju
3-Friend 1-Copula
'I am his friend'

We know that Emerillon represents an advanced stage in the loss of omni-predicativity. The presence of the two copulas – neither is reconstructed as such in proto-Tupi-Guarani (Rose 2003: 261) – is, therefore, coherent. More surprising is the existence in Kamayura, a relatively conservative language in terms of the referrer, of constructions with an element that Seki (2000: 158) identifies as a copula:

- (77) paje ere-ko
 Shaman 2-Copula
 ‘you are the shaman’

We note that, as is the case of the Emerillon copula *dju*, its use is restricted to the two first persons. This restriction in terms of persons could indicate a recent appearance of the copulative forms.³³ A third language with a copula is Ava-Canoeiro (Borges forthcoming). It seems to have reached a stage slightly more advanced than Emerillon, in that: 1) the same copulative form used for first and second persons – cognate with that of Kamayura – has extended to the third person,

- (78) ava a-ko
 AvaCanoeiro 1-Copula
 ‘I am an Ava-Canoeiro’
- (79) Tutaw i-ko Silene r-amyj
 Tutaw 3-Copula Silene Contiguity-Grandfather
 ‘Tutaw is Silene’s grandfather’

and 2) in some cases the referrer is still to be considered a suffix, while in others it has fossilised into a final phoneme of the root, as in, respectively

- (80) tukaN-a txi-txu
 AntSp.-Referrer 1Object-Sting
 ‘The ant stung me’
- (81) oka ‘house’
 k^wara ‘hole’

33 Lingua Geral of Amazonia (or Nhêngatu) could be a fourth language with a copula. But its copulative constructions seem to be of a different kind from that being studied here (Gerald Taylor, pers. com.).

If omnipredicativity had already begun to decline in pre-historic times³⁴ while there are still today living languages that do not have copulas, we have to recognise that systems do not respond at the same rate and in the same way to a given change in grammatical organisation. Cross-linguistically, the amount and the diversity of grammars that do not present a copula is notable (Benveniste, 1950). Launey (1994) suggests that nominal predication without copula is distributed along an axis: at the hard end is Nahuatl, which is omnipredicative, and at the soft end, non-omnipredicative Moroccan Arabic. The latter, like Russian, has the characteristic of generating a copula in tenses other than the present, an indication that in this tense we have the so-called ‘zero copula’.³⁵ This test is difficult to apply to Tupi-Guarani because the family has markers for past and future tenses which are exclusive to nouns. In Tapirape³⁶

- (82) ié ã-óp kyxé-kwér-ã
 1 1-Find Knife-Past-Referrer
 ‘I found what had been a knife’

The same markers indicate tense in nominal predication:

- (83) Ārāreme’í-Ø Kamoriwāgato’í-Ø r-atý-kwér-ã
 Ārāreme’í-Referrer Kamoriwāgato’í-Referrer Contiguity-Wife-Past-Referrer
 ‘Ārāreme’í was the wife of Kamoriwāgato’í’

The existence of markers specifically dedicated to expressing tense in relation to the noun may presumably have had a delaying effect on the creation of a copula.³⁷

What we have here is the question known in historical linguistics as the chronological mismatch between re-analysis and extension (Glidea 1998: 41), perhaps inherited from the theory of biological evolution (which in turn seems to have inherited it from tectonic plate theory): visible “changes occur by means of big leaps after a slow accumulation of tensions which a structure resists until it reaches its breaking point” (Gould 1985: 195). This breaking

34 That is, before 1500 A.D.

35 Nahuatl also has this characteristic, but the preservation of pronominal indexes on predicate nouns prevents comparison with Arabic or Russian.

36 Walkiria Praça (pers com.).

37 If, on the semantic level, the nominal tense of Tupi-Guarani were demonstrably different to the verbal tense, its incidence in relation to the creation of the copula should be revised.

point is different from one language to another. In languages such as Tocantins Assurini, according to Rut Monserrat's observations referred to above, the breaking point is virtually visible at present.

It is clear that the hypothesis defended in this paper should be extended and sustained by the localisation of the 'initial stage' in the relative diachrony of languages and families. There are indications that this stage is prior to proto-Tupi-Guarani. Cabral (2001) makes this explicit claim in recognising cognates of the referrer in various Tupi families in addition to Tupi-Guarani, in which they are either frozen as a final phoneme in words, or they occur – with quite restricted productivity in terms of distribution – in the nominalisation of verbs. Xipaya illustrates the two cases clearly:

	phoneme a	nominalisation (only with verbs ending in u)
(84)	aká 'house'	etúku 'to eat', etúka 'food'

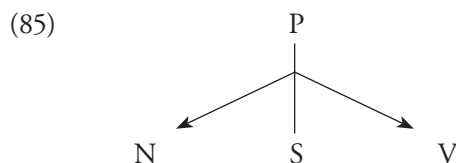
There is another independent clue to an initial stage older than proto-Tupi-Guarani. Along with a possible copula generation, another predictable consequence of the loss of omnipredicativity may be formulated as follows. Since sentences need to include designations with a certain frequency, the creation of the latter should be effected in the initial stage by means of light resources, say, by the simple addition of the referrer -a, as we may still see in various languages, among them Tocantins Assurini (example (63)). In subsequent stages, while a lexical class no longer needs the referrer to function as an argument – as we have illustrated plentifully above –, another class generates a complex and diversified system of forms for this same purpose: the *nominalisations*. In other words, nouns come to function as arguments in a *more* natural way, and verbs *less* so.

However, if the morphology of various nominalisations – circumstance, agent, patient co-occurring with agent, patient not co-occurring with agent – can be reconstructed for proto-Tupi-Guarani (Jensen, 1999)³⁸, this means that at this stage a class no longer had the ability to function as an argument with its former naturalness – i.e. with the referrer alone –, and that the relatively light forms such as that in (63), although still productive, are relics destined to disappear, which is presumably shown in the observations made concerning Tembe and illustrated in examples (40), (41) and (64).

38 Some may be even older, according to A. Rodrigues (pers. com.).

4 The Speciation of Nouns and Verbs

Looking at the comparative facts from the point of view of omnipredicativity in decline, the class of states referred to at the start of this paper should be seen as modern evidence of a period of *functional*³⁹ indistinction between lexical classes within a superclass of predicates, P. This period probably corresponded to the initial or omnipredicative stage. The argument function was, for all, the marked function. Members of this superclass suffered, with the loss of omnipredicativity, something like a (stronger) speciation, presumably driven by the semantics inherent in each member. Some became (more) nouns, N, others (more) verbs, V, by means of the appearance of an imbalance between the two classes in their respective capacities of accessing the marked function, the argumental (since overall, nouns continue predicating in the family). Other members of the superclass of predicates, precisely because of their hybrid semantics, remained functionally amorphic, generating by default the class of states, S. This split corresponds to the phase of the decline of omnipredicativity.

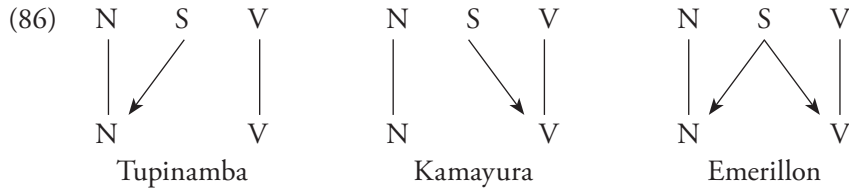


It was the conservation, in this later stage, of its functional indistinction that allowed the class of states to lean, in subsequent stages, either towards nouns, as in Guarani, Tupinamba, or Tocantins Assurini, or towards verbs, as in Kamayura, Tapirape⁴⁰ or Kayabi, without ever merging completely with neither nouns nor verbs. The originality of Emerillon in this context consists in a sort of repetition of history: the class of states undergoes in its turn a speciation – again driven by the meaning of each member – producing two subclasses that once again contrast with each other on the basis of their differential access to the

39 'Functional' since, as we have seen in Nootka or Nahuatl, nothing prevents lexical roots from being bound to either extensionality or tense-aspect and from having, thus, some kind of flexional morphology marking this distinction.

40 In Leite's (1990) interpretation. Walkiria Praça has a different point of view on this (pers. com.).

argument function: a class of quasi-nouns and a class of quasi-verbs (respectively, the nominoids and the attributives of Couchili et al. 2002),



The second line of (86) represents our empirical observational base – documented grammars. To discover whether the first and second lines of (85) represent respectively proto-Tupi and proto-Tupi-Guarani must be left to future studies which, crucially for the ideas advanced in this paper, must 1) resolve if and how the class of states is reconstructed into proto-Tupi-Guarani⁴¹ and 2) substantiate the assumption of semantics as a decisive factor in the two speciations.

5 Conclusion

The Tupi-Guarani family illustrates how an ancient omnipredicative pattern may be hypothesised through its manifestations observed in various stages of decay. An initial state, in which the basis of grammatical architecture is characterised by the primarily predicative nature of all (or almost all) of the lexical entries, comes to lose this characteristic over time and through the process of differentiation between languages. Remnants of the initial state appear in different members of the family – some as still active elements, others less so, others fossilised. Nouns and verbs grow apart from each other with respect to their functional affinities. The plausible formal consequences of this grammatical re-organisation – copulative elements, nominalising morphology – appear according to independent chronologies from one language to another, but the

⁴¹ The reconstruction of the class of states is being studied by A. S. Cabral and A. Rodrigues, of the University of Brasília.

family as a whole seems to be carried along by the same underlying drift. If comparative studies confirm the hypothesis advanced in this paper, we shall achieve an extremely wide-ranging and stimulating view concerning the way in which classes of words are shaped over time.

Bibliography

- Adelaar, W. 1997. O sufixo onomático e a preservação das consoantes do fim de raiz em Guaraní Paraguayo. *Boletim da Associação Brasileira de Linguística* 20, 11–19.
- Bach, E. 1968. Nouns and Noun Phrases, in E. Bach & R. T. Harms (eds), *Universals in Language*, New York, Holt & Rinehart, 89–122.
- Barbosa, Pe, A. L. 1956. *Curso de Tupi Antigo. Gramática, exercícios, textos* Rio de Janeiro, Livraria São José.
- Benveniste, E. 1950. La phrase nominale. *Bulletin de la Société de Linguistique de Paris*, 46, reprinted In *Problèmes de linguistique générale I*, 1966, 151–167 Gallimard, Paris.
- Boas, F. 1911. Kwakiutl. *Handbook of American Indian Languages*, BAE-B 40, Part 1, Washington, Smithsonian Institution.
- Borges, M. V. (forthcoming) Aspectos fonológicos e morfossintáticos do Avá-Canoeiro (Tupi-Guarani). Doctoral Dissertation, University of Campinas.
- Broschart, J. 1988. On the sequence of the Techniques on the Dimension of Participation. *AKUP* 76.
- Cabral, A. S. A. 1995. Contact-induced Language Change in the Western Amazon: The Non-Genetic Origin of the Kokama Language. PhD. Dissertation University of Pittsburgh.
- Cabral, A. S. A. 1998. Observações sobre caso no Asuriní do Tocantins. *Seminário Permanente de Línguas Indígenas*, Belém, UFPa.
- Cabral, A. S. A. 1999. Flexão relacional na família tupi-guarani. *Boletim ABRALIN*, 233–261.
- Cabral, A. S. A. 2001. Observações sobre a história do morfema *-a* da família Tupi-Guaraní, in F. Queixalós (ed), 133–162.
- Correa, J. 1981. Sandia Yvyvy, in A. Pecci (ed.), *Teatro breve del Paraguay*, Ediciones Napa.
- Couchili, T., Maurel, D. and Queixalós, F. 2002. Classes de lexèmes en émérillon. *Amerindia* 26/27, 173–208.
- Cunha de Oliveira, C. 2003. Lexical categories and the status of Descriptives in Apinaje. *International Journal of American Linguistics*, 69(3), 243–275.
- Davis, P. W. and R. Saunders. 1997. *A Grammar of Bella Coola*, UMOPL 13, Missoula, The University of Montana.
- Dietrich, W. 2001. Categorias lexicais nas línguas tupi-guarani (visão comparativa). In F. Queixalós (ed.), 21–37.
- Gildea, S. 1998. *On Reconstructing Grammar*. Oxford University Press, New York & Oxford.

- Givón, T. 2001. *Syntax I*. John Benjamins, Amsterdam.
- Gould, S. J. 1985. *El pulgar del panda*. Orbis, Barcelona.
- Grenand, F. 1980. *La langue wayâpi*. SELAF, Paris.
- Harrison, C. H. 1960. *Formulário dos vocabulários padrões para estudos comparativos preliminares nas línguas indígenas brasileiras: língua Asurini*, Anthropology Division, Linguistics Section, National Museum, Rio de Janeiro
- Harrison, C. H. 1986. Verb prominence, verb initialness, ergativity and typological disharmony in Guajajara, in D. C. Derbyshire and G. K. Pullum (eds), *Handbook of amazonian languages I*, 408–415. Mouton de Gruyter, Berlin.
- Hopper, P. J. and S. A. Thompson. 1984. The discourse basis for lexical categories in universal grammar. *Language*, 60(4), 703–752.
- Jacobsen W. H. Jr. 1979. Noun and Verb in Nootkan, in B. Efrat (ed.), *The Victoria conference on Northwestern languages*, 83–155. British Colombia Provincial Museum Heritage Record, no. 4, Victoria.
- Jelinek, E. 1984. Empty categories, case and configurationality. *Natural language and linguistic theory* 2, 39–76.
- Jelinek, E. 1993. Ergative ‘Splits’ and Argument Type. *MIT Working Papers in Linguistics* 18, 15–42.
- Jensen, C. 1998. Comparative Tupí-Guaraní Morphosyntax, in D. C. Derbyshire and G. K. Pullum (eds), *Handbook of Amazonian Languages* 4, 408–415. Mouton de Gruyter, Berlin and New York.
- Jensen, C. 1999. Tupi-Guarani, in R. Dixon and A. Aikhenvald (eds), *The Amazonian Languages*, 125–163. Cambridge University Press, Cambridge.
- Kinkade, M. D. 1983. Salish evidence against the Universality of ‘Noun’ and ‘Verb’. *Lingua* 60, 25–40.
- Klimov, G. A. 1974. On the character of languages of active typology. *Linguistics* 131, 11–25.
- Launey, M. 1986. Catégories et opérations dans la grammaire nahuatl. Thèse d’Etat, Paris, Université de Paris IV.
- Launey, M. 1994. *Une grammaire omniprédicative. Essai sur la morphosyntaxe du nahuatl classique*. CNRS Editions, Paris.
- Lazard, G. 1999. La question de la distinction entre nom et verbe en perspective typologique. *Folia Linguistica*, 33(3–4), 389–418.
- Leite, Y. 1990. Para uma tipologia ativa do tapirapé. *Cadernos de estudos linguísticos* 18, 37–56.
- Leite, Y. 2001. O estatuto dos sintagmas nominais de sujeito e objeto em Tapirapé, in F. Queixalós (ed.), 87–101.
- Lemaréchal, A. 1989. *Les parties du discours. Sémantique et syntaxe*. PUF, Paris.
- Lemaréchal, A. 1991. *Problèmes de sémantique et de syntaxe en palau*. CNRS Editions, Paris.
- Lyons, J. 1977. *Semantics 2*. Cambridge University Press, Cambridge.
- Maurel, D. 1998. *Éléments de grammaire émerillon*, Chantiers Amerindia, Supplément 1 au n° 23 d’Amerindia.
- Maurel, D. 2001. “Etre” et “avoir” en émerillon. ms.
- Melià, B. 1983. La lengua guaraní del Paraguay, in B. Pottier (ed.), *América latina en sus lenguas indígenas*. UNESCO/Monte Avila Editores, Caracas.

- Monserrat, R. 1985. Notas de campo.
- Nicholson, V. 1976. 6 textos na Língua Asuriní. *Arquivo Lingüístico* 016, Brasília, SIL.
- Paris, M. C. 1984. Nom et verbe en chinois: ressemblances et différences. *Modèles linguistiques* VI.1, 101–115.
- Praça, W. 1999. Nomes como predicados em Tapirapé. Master thesis, Universidade de Brasília.
- Queixalós, F. 2000. *Syntaxe sikuani*. Peeters, Louvain.
- Queixalós, F. (ed.). 2001a. *Noms et verbes en tupi-guarani: état de la question*, Studies in Native American Languages 37. Lincom Europa, Munich.
- Queixalós, F. 2001b. Le suffixe référentiel en émérillon, in F. Queixalós (ed.), 117–132.
- Reich, U. 2003. Mudança sintática e pragmática na Língua Geral Amazônica (LGA): marcação de caso e sistema pronominal, in J. R. Bessa Freire and M. C. Rosa (eds), *Línguas gerais. Política lingüística e catequese na América do Sul no período colonial*. UERJ, Rio de Janeiro.
- Rodrigues, A. D. 1953. Morfologia do verbo tupi. *Letras* 1, 121–152.
- Rodrigues, A. D. 1990. You and I = Neither You nor I: The Personal System of Tupinamba, in P. L. Payne (ed.), *Amazonian Linguistics. Studies in Lowland South American Languages*, 393–405. University of Texas Press, Austin.
- Rodrigues, A. D. 1995. Oposición conceptual y analogía fonológica en guarani: ‘casa’ y ‘chacra’. *Lingüística aborigen II*, Buenos Aires, UBA, 325–329.
- Rodrigues, A. D. 1996a. Argumento e predicado em Tupinambá. *Boletim da Associação Brasileira de Lingüística* 19, 57–66.
- Rodrigues, A. D. 1996b. As línguas gerais sul-americanas. *Papia. Revista de Crioulos de Base Ibérica* 4.2.
- Rodrigues, A. D. 2001. Sobre a natureza do caso argumentativo, in F. Queixalós (ed.), 103–114.
- Rodrigues, A. A. and A. S. A. C. Cabral. 2002. Revendo a classificação interna da família Tupi-Guaraní, in A. S. A. C. Cabral and A. A. Rodrigues (eds), *Línguas Indígenas Brasileiras. Fonologia, Gramática, História*, UFPa, Belém.
- Rose, F. 2000. Eléments de phonétique, phonologie et morphophonologie de l’émérillon (teko). D.E.A., Université de Lyon II.
- Rose, F. 2003. Morphosyntaxe de l’émérillon. Doctoral Dissertation, Lyon, Université Lumière Lyon II.
- Seki, L. 1990. Kamaiurá (Tupi-Guaraní) as an Active-Stativ Language, in E. Payne (ed.), *Amazonian Linguistics, Studies in Lowland South American Languages*, 367–391. Austin University of Texas Press.
- Seki, L. 2000. *Gramática do Kamaiurá*. Editora de Unicamp, São Paulo.
- Swadesh, M. 1939. Nootka internal syntax. *International Journal of American Linguistics* 9, 77–102.
- Tchekhoff, C. 1984. Une langue sans opposition verbo-nominale: le tongien. *Modèles linguistiques* VI.1, 125–132.
- Thompson L. C. and M. T. Thompson. *Thompson ms.*, quoted in Kinkade (1983), published as L. C. Thompson, M. T. Thompson and S. M. Egesdal. 1996. Sketch of Thompson, a Salishan language. In I. Goddard (ed.), *Handbook of North American Indians*, vol. 17, *Languages*, 609–643. Smithsonian Institution, Washington, D.C.

- Velazquez-Castillo, M. 2002. Guarani Causative Constructions, in M. Shibatani (ed.), *The grammar of causation and interpersonal manipulation*. Typological Studies in Language Series 48, 507–534. John Benjamins, Amsterdam.
- Vieira 1993. O problema da configuracionalidade na língua Asuriní: uma consequência da projeção dos argumentos do predicado verbal, Doctoral Dissertation, Campinas, University of Campinas.

