A typological overview of Emerillon, a Tupi-Guarani language from French Guiana

F. Rose

Language Profile

A typological overview of Emerillon, a Tupí-Guaraní language from French Guiana

FRANÇOISE ROSE

Abstract

This article offers a typological overview of the Emerillon language, a Tupí-Guaraní language spoken by a small community in French Guiana. General information is provided on various aspects of the grammar, within the domains of phonology, morphology, and syntax. Special attention is given to a few features of the language that are rather rare and/or poorly discussed in the typological literature, namely morphemic nasal harmony, a hierarchical person indexation system, a rare type of nominal predication, and the existence of a specific marker for sociative causation. These features are all typical of the Tupí-Guaraní family.

Keywords: *alignment, Amazonia, causative, cross-reference, Emerillon, nasal harmony, nominal predication, person hierarchy, Tupí-Guaraní, valency*

1. Introduction

The Emerillon community consists of about 400 people living in two areas in French Guiana: next to the Maroni river (on the border with Suriname) in the western part and at the Oyapock-Camopi confluence (on the border with Brazil) in the eastern part. The community (auto-denominated Teko) is the result of the aggregation of surviving members of different small ethnic groups, mainly of Tupí-Guaraní origin (Navet 1994). The Emerillon language is still actively being used and passed on to children as a native language. Even though it is in close contact with several languages – essentially Wayampí and Brazilian Portuguese in the eastern part, Wayana and Eastern Maroon Creoles in the western part, French Guiana Creole and French, the language used in school, in both areas – the influence of European and Creole languages on Emerillon is

Linguistic Typology 12 (2008), 431–460
DOI 10.1515/LITY.2008.044
1430–0532/2008/012-0431

very limited and concerns mainly the lexicon (Rose & Renault-Lescure 2008). It must, nevertheless, be considered endangered.

Emerillon belongs to the Tupí-Guaraní family (consisting of over forty languages) of the Tupí stock (Rodrigues 1984–1985). The name Tupí-Guaraní stems from Tupinambá, the now extinct language spoken along the coast in the early colonization times, and Guaraní, which made it through the centuries to become one of the official languages of Paraguay, spoken by around five million people. This group is probably the best studied one in Amazonia, with a long tradition of linguistic investigation since Anchieta (1595). Nevertheless, only few recent and comprehensive grammars like Seki's Kamaiurá grammar (Seki 2000) are available. However, comparative study within the family (Jensen 1998) is facilitated since "Tupí-Guaraní is noted for a high degree of lexical and morphological similarity among its member languages in spite of their extensive geographical separation" (Jensen 1999: 128). These languages are indeed spoken throughout Brazil, in northern Argentina, Paraguay, Bolivia, and French Guyana. Internal classification of the Tupí-Guaraní family was established by Rodrigues on lexical and phonological grounds (Rodrigues 1984–1985) and then revised (Rodrigues & Cabral 2002). Emerillon belongs to subgroup VIII,[1] along Urubu-Ka'apor, Anambé de Ehrenreich, Guajá, Awré e Awrá, Takunhapé, and its closest relatives Wayampí, Wayampípukú, and Jo'é. Other classifications do not differ much as far as Emerillon is concerned (Dietrich 1990, Mello 2002).

The Emerillon language had been very little described before (Couchili et al. 2002, Maurel 1998, Queixalós 2001b). My own work (Rose 2003b, to appear) is based on texts and elicited items collected by myself in the field from various speakers, mainly from the eastern part, between 1999 and 2004.

This article aims at providing an overview of the typological characteristics of the language. I will try to stick to a synchronic description of the language within a typological perspective, referring when necessary to typological stud-

1. The eight subgroups are constituted as follows (Rodrigues & Cabral 2002):

   I: Old Guaraní; Kaiwá, Ñandeva, Paraguayan Guaraní; Mbyá-Guaraní; Xetá; Tapiete, Chiriguano, Izoceño; Guayakí

   II: Guarayo/Guarayú; Siriono, Horá; Yuki

   III: Tupí, Língua Geral Paulista; Língua Geral Amazônica, Tupinambá

   IV: Tapirapé; Asuriní do Tocantins, Parakanã; Suruí; Avá-Canoeiro; Tembé, Guajajára, Turiwára

   V: Araweté, Ararandewára-Amanajé, Anambé do Cairari; Asuriní do Xingu

   VI: Kayabí, Apiaká; Parintintín, Tupí-Kawahíb; Juma

   VII: Kamayurá

   VIII: Wayampí of French Guiana, Wayampípukú, EMERILLON, Jo'é; Urubu-Ka'apor, Anambé de Ehrenreich; Guajá; Awré e Awrá; Takoapé

ies, but leaving aside diachronic hypotheses and comparative remarks.[2] The major points developed here are quite representative of the whole Tupí-Guaraní family.

The Emerillon phonological system is presented in Section 2, with a special focus on nasal harmony. Section 3 will describe Emerillon verbal morphology, namely a very interesting indexation system based on both person and semantic role hierarchies. Section 4 will then go on with clause syntax issues such as word order, different classes of predicates, and alignment. Finally, Section 5 will give an overview of valency-changing derivations, with special emphasis on the sociative causative. For each typologically remarkable feature of Emerillon, I will also discuss how they fit in with the existing typology.

2. Phonology

2.1. *Phoneme inventory*

The inventory of consonants and vowels is given in Tables 1 and 2. In cases of allophony, phonetic realizations are specified in brackets using the IPA notation. The system is in several ways quite typical of Amazonian phonology (Aikhenvald 1996, Dixon & Aikhenvald 1999), for example as having one liquid, two glides, the closed central vowel. It is "richer" than the average systems of Amazonian languages, displaying a voice distinction within stops, and the *o*/*u* distinction. Like the rest of the system, /k^{w}/ and /ʔ/ are found throughout the Tupí-Guaraní family.

The various cases of allophony are conditioned by different factors:

(i) dialectal variation: /s/ and /z/[3] can be phonetically realized as either fricatives [s] and [z] or affricates [ts] and [dz];
(ii) the vocalic environment: [h] ∼ [ɣ], [w] ∼ [β];
(iii) the syllabic structure: mid vowels tend to be more open in closed syllables;
(iv) the position within the morpheme/word/prosodic phrase domain: in morpheme-final position before a prosodic break, the non-continuants are unreleased [p˺], [t˺], [ʧ˺], and [k˺]. In absence of release, voicing is irrelevant.

2. Emerillon offers some features that are interesting in a diachronic perspective, like the general restructuring of dependent clauses, with a shift in the indexation system, the emergence of serial verb constructions out of dependent constructions, and the replacement of ancient nominalization patterns by subordinators syncretic with adpositions (Rose 2006). It also raises many comparative or areal issues, like the descriptive word category, the so-called relational marker, the question of polyfunctionality or omnipredicativity (see, for example, Queixalós 2001a, b) that concern a great number of languages in the Amazonian area.
3. Even though /s/ and /z/ have to be considered distinct phonemes, it is noteworthy that /s/ is most often found before high vowels and /z/ essentially before mid or low vowels. /z/ and /ʤ/ are both reflexes of proto-Tupí-Guaraní †j. Emerillon /j/ probably emerged as a transition.

Table 1. *Emerillon consonant system*

	labial	alveolar	palatal	velar	labio-velar	glottal
voiceless non-continuant	p [p, pˀ]	t [t, tˀ]	ʧ [ʧ, ʧˀ]	k [k, kˀ]	k^w	ʔ
voiced non-continuant	b [b, mb, m]	d [d, nd, n]	ʤ [ʤ, ɲ]	g [g, ŋ]		
voiceless continuant		s [s, ts]				
voiced continuant		z [z, dz]				
non-obstruent	w [w, β, w̃]	ɾ [ɺ, ɾ, n]	j [j, j̃]			h [h, ɣ, h̃]

Table 2. *Emerillon vowel system*

	front	central	back
closed	i, ĩ	ɨ, ɨ̃	u, ũ
mid	e, ẽ [e, ɛ]	ə	o, õ [o, ɔ]
open		a, ã	

However, distribution facts suggest that the unreleased consonants are allophones of voiceless non-continuants,[4] since voiced non-continuants, as will be explained later, are nasalized in final position. Yet the voiceless non-continuants in this same morpheme-final position undergo intervocalic voicing (and are realized [b], [ɾ], [ʤ], and [g]), but only within the domain of the prosodic phrase.

(1) a. /kɨʤap/ [kɨʤapˀ] papaya ‘papaya’

b. /kɨʤap-a-kom/ [kɨʤabakom] papaya-REF-PL ‘papayas’

In the absence of nasality further to the right within the same morpheme, the voiced non-continuants /b/ and /d/ are generally realized as prenasals (or post-oralized nasals as we will see later) in intervocalic position only.

4. In the writing system, this final consonant is written with the voiced consonant symbol, to ensure a unique root form whatever the morphophonological context.

Prenasalization is systematic morpheme-internally (2a) and optional morpheme-initially when preceded by a vowel within the word domain (2b).[5]

(2)	a.	/sibo/	b.	/o-bo-ʔi/
		[sĩmbo]		[oboʔi ∼ omboʔi]
		rope		3.I-CAUS-small
		'rope'		'he cuts it into pieces'

(v) nasal harmony: in our analysis, nasality is not always contrastive at the segmental level (hence the absence of nasal consonants in the inventory), but is a feature spreading over morphemes, typically roots, nasalizing them totally or partially. This phenomenon will be described in greater depth in Section 2.4.

2.2. *Syllabic and word structure*

Attested syllabic structures are quite simple: CV (by far the most frequent), V, CVC, and VC. Vowel sequences are common, each vowel being the nucleus of a separate syllable. Closed syllables are only found in morpheme-final position.

The canonical pattern of an Emerillon word is a sequence of open syllables, possibly with a final closed syllable: (C)V…(C)V.(C)V(C). A major role of morphophonemics is to prevent the creation of consonant sequences within a word, at morpheme boundaries.

In word-initial position, all consonants but /g/, /j/, and /ɾ/ are found. In word-internal position, all consonants are found (though /g/ is restricted to morpheme-final position). In word-final position, only non-continuants are found and are either unreleased, therefore neutralizing the opposition between voiced and unvoiced, or nasal (see Section 2.4).

2.3. *Stress*

The domain of stress is the prosodic phrase. Primary stress (ˊ) usually falls on the penultimate syllable of the phrase. Secondary stress (ˋ) alternates on every second syllable counting backward from the primary stress: *pànanàɾupì omàʔẽnè nõdèɾehéo* '(when we are) on the sea, he still sees us'. Optional stress can also be found on the initial syllable of the phrase. Besides, stress is attracted by heavy syllables, and is often avoided on high vowels in onsetless syllables. The major acoustic correlates of stress are duration and intensity (Gordon & Rose 2006). Comparative Tupí-Guaraní studies on stress, based on a small number of impressionistic descriptions, traditionally reconstruct a delimitative

5. This general rule is yet subject to inter- and intra-speaker variation. Morpheme-finally, the voiceless continuants that are voiced when followed by a vowel within the prosodic phrase domain are never prenasalized.

stress on the final syllable, although Wayampí, Siriono, and Chiriguano are presented as historical deviations due to their penultimate stress patterns (Dietrich 1990: 16–17, Jensen 1999: 133).

2.4. *Nasal harmony*

In this section, I will show that nasality is a feature that spreads within morphemes and to adjacent morphemes, forming a type of nasal harmony that is not very common typologically.

2.4.1. *Nasals as allophonic variants of voiced consonants.* Three features of the language point to the fact that nasal consonants are allophones of their oral voiced corresponding sounds.

First, voiced consonants and nasal consonants are in complementary distribution, based on the oral vs. nasal environment of the segment concerned. No minimal pair can be found where the nasality of just one segment would be the distinctive feature: [baʔe] 'thing' vs. [mãʔẽ] 'REL'; [owa] 'face' vs. [õw̃ã] 'a little'; [paku] 'fish species' vs. [pẽkũ] 'bird species'.

Second, some affixes have two surface forms, characterized by absence vs. presence of nasality. In each case, the selection of one or the other form is determined by the presence vs. absence of nasality of the root to which they are affixed. Example (3) illustrates this phenomenon for the negation suffix [ʤi] ∼ [ɲĩ].

(3) a. [d-a-ʤapɨaka-ʤi]
NEG-1SG.I-think-NEG
'I don't think so.'

b. [d-o-mãnõ-ɲĩ]
NEG-3.I-die-NEG
'He didn't die.'

Third, the integration of loan words originally including some nasal segments shows either nasalization (Guianese Creole [dipẽ] becomes [nĩpẽ] 'bread' in Emerillon) or denasalization (Wayana [malija] becomes [baliʤa] 'knife' in Emerillon).

These three characteristics of Emerillon indicate an allophonic distribution between oral voiced and nasal segments in nasal contexts. Crosslinguistically, languages where nasal consonants are not distinctive are rare. Very often, this is explained by nasal consonants being allophones of oral consonants in nasal contexts, as found especially in West African and South American languages (Clements & Osu 2005). The next section focuses on the precise distribution of these allophones in Emerillon.

2.4.2. *The distribution of nasality within lexical roots.* Nasality in Emerillon affects strings of segments. Some lexical roots are completely affected by nasality, and their final segment is necessarily nasal, e.g., [ãmãn] 'rain', [w̃ãĩw̃ĩ] 'woman', [ãh̃ãɲ] 'tooth', [nãsĩʔõ] 'mosquito', [tɬ̃ãkã] 'river'. A smaller number of roots are disharmonic. They are affected by nasality only partially, more specifically to the left of a consonant [mb] or [nd], e.g., [nĩmbo] 'rope', [mɬ̃ndɨʤu] 'cotton', [kãndetatˀ] 'crown', [tãmbe] 'flat'.[6] No trace of nasality can be found to the right of that segment.

The nasal string can yet contain voiceless elements, both in totally ([nãsĩʔõ], [tɬ̃ãkã]) and partially ([kãndetatˀ], [tãmbe]) nasal roots, but absolutely all voiced segments within this string are nasal.

2.4.3. *Analysis.* My analysis is that nasality is lexically assigned in Emerillon, and some nasal harmony is going on, within the strict domain of the morpheme. Section 2.4.4 will discuss the propagation of nasality to adjacent morphemes. Within the scope of nasality, all voiced segments (apart from /z/)[7] are realized nasally. Voiceless phonemes are transparent to nasalization, but not opaque: they cannot be nasalized but yet do not block the spreading.

In the analysis of nasal harmony, once the domain of nasalization is delimited and the segments are classified as underspecified for nasality, transparent, and/or opaque, the major challenge is to define the trigger and the direction of spreading, as Walker recognizes for Tuyuca: "because nasality spreads to all nasalizable segments in a nasal morpheme, it is impossible to unambiguously pinpoint the segment from which spreading originates" (Walker 2003).

My hypothesis is that nasality in Emerillon is to be analyzed as a feature spreading leftward from the final segment of a root (the nasal allomorphs of /b/, /d/, /ʤ/ or /g/) or from its rightmost [mb] and [nd]. A first piece of evidence is the distribution of nasal elements within roots. There is always only one span of nasality per root. In roots totally affected by nasality, the final segment is always a nasalizable element, which is therefore realized as a nasal vowel or consonant, while their initial segment is not necessarily nasalizable (see [tɬ̃ãkã], [kãndetatˀ], and [tãmbe] from Section 2.4.2). There is absolutely no morpheme totally affected by nasalization ending with a non-nasal element. In roots partially affected by nasality, no /b/ or /d/ phonemes can be found to the right of the prenasalized element.

---

6. In some occurrences of such words, especially in frequent words, when the vowel immediately preceding this consonant is strongly nasalized, the nasal part of the consonant is not realized.
7. In my corpus, there are six roots affected by nasality within which /z/ (realized [dz]) is found. However, three of them are loan words, and the three remaining ones may be considered as part of the Amazonian substrate: [madziʔokˀ, dzakami, dzapakani]. There is, however, no evidence that /z/ has a nasal counterpart.

A second piece of evidence for the anchorage of nasality to the right end of a morpheme can be found in loan words:

> Although nativization strategies cannot always be equated with actual constraints that account for the phonotactic structure of the borrowing language, one would expect, on the other hand, that the constraints of the borrowing language are always active when words from another language are adapted. (Wetzels 2007)

In all the languages from which Emerillon has borrowed words (i.e., essentially French, Creoles, and other Amerindian languages), nasality is a segmental feature. In the process of adaptation to the target phonological system, these words are integrated as either oral or (partially or totally) nasal morphemes, whereby the lexeme undergoes either nasalization or denasalization so that the resulting form always fits with what is expected according to the native system. The selection of complete nasalization vs. denasalization seems to be induced by the nasality of its final segment in the source language. Words with a nasal final segment in the source language are completely nasalized in Emerillon (4a–c). Those with an oral final segment (even though some initial segments may be nasal) are denasalized (4d, e). Moreover, example (4f) shows that a [m] towards the end of the source lexeme may be realized as post-oralized and triggers partial nasalization.

(4) Nasalization

a. [dipẽ] (Guianese Creole) → [nĩpẽ] 'bread' (Emerillon)
b. [savɔ̃] (French) → [sãɨ̃wã] 'soap' (Emerillon)
c. [faʁin] (French or Guianese Creole)[8] → [pãnĩn] 'flour' (Emerillon)

Denasalization

d. [malija] (Wayana) → [balicʒa] 'knife' (Emerillon)
e. [mɔ̃pɛʁ] (French) → [bopeˀ] 'priest' (Emerillon)

Partial nasalization

f. [pomad] (French or Guianese Creole) → [põmbatˀ] 'gel' (Emerillon)

The question left to be discussed now is: What is the trigger of nasal harmony? Here are three alternative hypotheses.

8. [ʁ] of the source language is usually substituted with [ɾ], cf. [sitʁuj] 'pumpkin' (French) → [siturucʒ] (Emerillon).

On the first hypothesis, nasality is a suprasegmental feature, a phonological feature of morphemes as a whole, rather than of particular phonemes. It is mapped to the rightmost element of the morpheme and spreads leftward. Nasality is then assigned lexically. This explains the cases of totally nasalized roots. Then partially nasalized roots must be explained by phonetic spreading of nasality from prenasalized consonants, a pattern that is typologically uncommon (Leo Wetzels, personal communication). This analysis is proposed by Piggott for Guaraní, another Tupí-Guaraní language with a similar realization of nasality:

> [i]n the case of nasal morphemes, a floating nasal autosegment must be present. The other source of nasality must be segmental, since nasality can spread from a point internal to a morpheme. [...] Since it is the latter segment [a prenasalized stop – FR] that obligatorily appears between nasal and oral sequences, the most likely source of the nasality that spreads leftward is an underlying nasal consonant, which is phonetically realized as a prenasalized stop. (Piggott 1992: 57)

Tupi-Guaranists like Cabral follow a comparable analysis. The autosegmental feature is associated with the opposition between oral and nasal stresses (Cabral 2000, Grannier Rodrigues 1990). In a previous publication, our first analysis of Emerillon nasality was autosegmental (Rose 2002a).

Under the second hypothesis, nasality is always phonemic, and total and partial nasalization constitute a unique phenomenon. One would have to postulate underlying nasal vowels and nasal consonants, triggering nasal harmony leftward. To explain partially nasalized roots, consonants in intervocalic position should also be seen as post-oral variants of /m/ and /n/. This analysis is not satisfactory because it does not account for the facts that only one contrast per word is possible, that no oral voiced segment can be found to the right of a nasal segment (except when a prenasalized stop intervenes), and that voiced non-continuants are not found word-finally (except as a result of intervocalic voicing).

A third alternative satisfactorily explains both nasalization and final consonant distribution. A lenition rule according to which voiced non-continuants in final position are systematically nasalized must be postulated, accounting for morpheme-final consonants being either unreleased (e.g., /tapɨʧ/ [tapɨʧ̚] 'house') or nasal (e.g., /tabɨʤ/ [tãm̃ɲ] 'grandfather'). The same lenition rule applies exclusively to /b/ and /d/ in intervocalic position, nasalizing them. However, in intervocalic position, the nasalized consonant will be post-oralized (e.g., /aduʤa/ [ãnduʤa] 'mouse'). As a result, the distinction between oral stops, nasals, and prenasals is neutralized. The consonants that are phonetically realized as nasals or prenasals, whether in final or intervocalic position, will trigger nasal harmony to their left. However, this analysis still makes it necessary to assume underlying nasal vowels. To predict that morphemes have

at most one nasal span, nasal vowels must be restricted to morpheme-final position.

2.4.4. *The propagation of nasality to adjacent morphemes.* Affixes are either nasalizable or they are not (only enclitics can be lexically nasal). Indeed, nasal harmony can propagate within the word to some affixes (prefixes and suffixes) that are adjacent to a totally nasalized root or to the nasal part of a partially nasalized root. In the following examples, the relational *ɾ-* (5a), and the negation *-ʠi* (5d) are affected by nasal harmony on a totally nasal root as in (5b, e). Other prefixes are always oral, like *e-* in (5a–c), while a few are very rarely nasalized, like *d-* and *o-* (5d–f). Finally, some enclitics are always nasal (but do not trigger nasality propagation themselves, like *-(o)ŋ* in (5f)).

(5) a. [e-ɾ-apɨʧ̃]
1SG.II-RELN-house
'my house'

b. [e-n-ãmɨ̃ɲɪ]
1SG.II-RELN-grandfather
'my grandfather'

c. [e-mẽmbɨt̃]
1SG.II-son
'my son'

d. [d-a-ʠapɨaka-ʠi]
NEG-1SG.I-think-NEG
'I don't think so.'

e. [d-o-mãnõ-ɲĩ] ∼ [n-õ-mãnõ-ɲĩ]
NEG-3.I-die-NEG
'he didn't die'

f. [o-pɨhɨg-õŋ]
3.I-catch-PL.S
'they caught it'

It is worthwhile noting that while spreading within morphemes is strictly leftward, propagation to adjacent morphemes is attested both leftward and rightward.

2.4.5. *Emerillon within the typology of nasal harmony.* Recent research on nasal harmony, generally based on secondary data, distinguishes two types of nasal harmony. In the phonetic or segmental type of nasal harmony (Type A in Piggott's terms), nasality spreads from a nasal segment until it is blocked by an opaque segment (Piggott 1992). For example, in Sundanese, non-nasal, supralaryngeal consonants block the long distance spreading of nasality (Cohn 1990). In the prosodic or suprasegmental type, defined as Type B by Piggott,

nasal harmony is operating through morpheme-level or word-level specifications for [+nasal] or [−nasal], where most segments surface differently in nasal and non nasal morphemes. It is correlated with the absence of opaque segments: even plosives are transparent. In other words, there is no adjacency requirement: the rule applies no matter what intervenes between the trigger and the target (Odden 1994). Type B is exemplified by a very few languages in the literature, and essentially these are languages from South America, among which Guaraní and the Tucanoan languages (see, for example, Piggott 1992, Gomez-Imbert 1997). However, some authors suggest that suprasegmental nasality could always, in principle, be reanalyzed as based on segmental nasality.

Nasal harmony in Emerillon basically displays the characteristics of Piggott's Type B. However, it has been shown that nasal harmony in Emerillon can be seen as triggered by nasal allophones of oral segments. This is accounted for in Piggott's typology, where "a Type A language must have underlying nasal consonants, but similar segments may all be derived in a Type B language" (Piggott 1992: 62). However, Emerillon constitutes a very particular illustration of that last case, in that nasality does not have to be derived from a suprasegmental feature, but can be seen as the result of allophony.

Moreover, the phonological rule nasalizing voiced non-continuants in final position (and for some in intervocalic position) is not common, although it is plausible that voiced obstruents are phonetically hard to maintain in final position. Comparable processes of lenition of postvocalic stops into nasals are found in Japanese (Shibatani 1990, Tsujimura 1996). Rodrigues (2003) puts forward a phonetic hypothesis explaining how silence (i.e., final position) can cause nasalization, due to desynchronization of the velum movements.

Finally, the Emerillon data also challenge the question of transparency. Walker's (2003) scale, ranking segments according to their compatibility with nasalization, predicts that (i) if a segment blocks nasalization, all segments less compatible with nasality will also block it, and (ii) if a segment is permeable (nasalized or transparent), all segments more compatible by the nasalization hierarchy will also be permeable.

(6) Vowels > Glides > Liquids > Fricatives > Obstruent stops

Yet the hierarchy does not predict, within the permeable segments, which will be targets and which will be transparents. In Emerillon, voiced obstruent stops undergo nasalization while voiced fricatives resist nasalization (as well as all voiceless segments).

The Emerillon phonological system is particularly interesting in displaying a type of nasal harmony system at morpheme-level that has come to be discussed for South American languages (Peng 2000, Rodrigues 2003). Moreover, it chal-

lenges typologies of nasal harmony on several points, such as the trigger and transparency.[9]

3. Morphology

3.1. *Overview*

Morphological typology categorizes languages according to two parameters: fusion, the degree to which morphemes are easily segmentable, and synthesis, the number of morphemes per word (Comrie 1981). On these scales, Emerillon is an agglutinative language, leaning towards polysynthesis.

(7) *eɪe-mo-zaug-a-õwã-zepe-ʔe-po* *mamã,* *ʤasoɪ?*
2SG.I-CAUS-bathe-REF-little-CONCES-INTENS-INTER Mum Djasot
'But did you really wash Mum properly, Djasot?'

Emerillon employs predominantly suffixes or enclitics, with only limited prefixation. Prefixes are all person markers, some voice markers, one negative marker, and one subordinator. Emerillon is a head-marking language. There is greater morphological complexity on the predicate, which carries most of the grammatical morphemes, and also undergoes reduplication (Rose 2005, 2007).

3.2. *A hierarchical indexation system*

Within the domain of verb morphology, the Emerillon indexation system is the most challenging feature. In most Tupí-Guaraní languages, indexation differs in dependent and independent clauses.[10] In Emerillon, however, dependent and independent clauses follow the same indexation pattern. The system presented below is comparable to the indexation system in other Tupí-Guaraní independent clauses, but nowadays also applies to dependent clauses in Emerillon.

Person markers are divided into two main sets, called Set I and Set II, as shown in Table 3.

We will now look at their specific distribution on verbs. On intransitive verbs, only Set I is found, to index S:

(8) *a-ʔiɨta* *ʔi-pope.*
1SG.I-swim river-in
'I swim in the river.'

On transitive verbs, Set I is used for A and Set II for P.

9. In the transcription system used in the rest of the article, I note both nasal vowels and nasal consonants. It is therefore not a pure phonological writing, but aims at greater ease of processing.
10. In Tupí-Guaraní languages, indexation in dependent clauses is usually described as following an absolutive system with some coreference pattern (see Jensen 1998: 526–532 for further details).

Table 3. *Emerillon person indexes*

	Set I	Set II
1SG	*a-*	*e-*
2SG	*eɾe-*	*de-*
1INCL	*si-*	*nõde-/kõde-*
1EXCL	*oɾo-*	*oɾe-/oɾone-*
2PL	*pe-*	*pe-/pene-*
3	*o-*	*i-*
indeterminate	*za-*	*zo-/poɾo-*[a]

a. *zo-* is used on nouns, *poɾo-* on verbs to mark a generic human object 'people'.

Table 4. *Distribution of person indexes on transitive verbs*

	A	P
Set I	×	
Set II		×

The interesting point here is that there is only one slot for person indexes on transitive verbs. Whether A or P will be marked depends on different hierarchies. First, the person hierarchy (9) plays a role whenever a speech act participant and a 3rd person interact. Speech act participants are higher than the 3rd person on this scale.

(9) Person hierarchy: 1/2 > 3

The participant higher on the hierarchy is marked on the verb, whether A or P. In (10a, b), a 1st person and a 3rd person are involved. The 1st person being higher in the hierarchy, it is in both cases the one to fill the person index slot, with a Set I index when it is A (10a), with a Set II index when it is P (10b).

(10) a. *a-nupã.*
1SG.I-hit
'I hit him.'

b. *zawaɾ e-suʔu.*
dog 1SG.II-bite
'A dog bit me.'

Set II indexes do not only mark P on the verb. They are also used to refer to the possessor of nouns (11a) and the object of postpositions (11b).

(11) a. *mɨn i-men o-manõ.*
long.ago 3.II-husband 3.I-die
'Long ago, her husband died.'

b. *o-apɨg i-ʔaɾ.*
3.I-sit 3.II-on
'She sat on it.'

The second hierarchy is the semantic role hierarchy (12), at work whenever the person hierarchy is not relevant, i.e., between two 3rd persons or two speech act participants.

(12) Semantic roles hierarchy: A > P

When a 3rd person acts on another 3rd person, it is the A that is systematically marked (with a Set I index), whatever the arguments refer to.

(13) a. *patu-pope o-iɲuŋ.*
pot-in 3.I-put
'She puts them (the sweet potatoes) in the pot.'
b. *aɾakapusa-uhu o-mõduɾ-oŋ bal.*
gun-big 3.I-send-PL.S bullet
'Guns were shooting bullets.'
c. *o-peɾo-peɾog eiba Ø-owa.*
3.I-RED-lick 3.II.pet 3.II-face
'His dog licks his face.'

The local configurations (i.e., when the two speech-act participants are involved) basically also follow the A > P hierarchy, with residues of the Proto-Tupí-Guaraní 1 > 2 > 3 hierarchy. When a 2nd person acts on a 1st person, the A is marked on the verb, with a Set I index. To specify the person of the P, an extra independent marker is used.

(14) a. *eɾe-nupã eɾeɲ.*
2SG.I-hit 2SG+ɲ
'You (SG) hit me.'
b. *eɾe-nupã oɾone-kom.*
2SG.I-hit 1EXCL-PL
'You (SG) hit us (EXCL).'
c. *pe-nupã peɲ.*
2PL.I-hit 2PL+ɲ
'You (PL) hit me.'
d. *pe-nupã oɾone-kom.*
2PL.I-hit 1EXCL-PL
'You (PL) hit us (EXCL).'

The presence of the independent pronominal form is necessary to disambiguate the 2 → 1 configuration from the configuration where a 2nd person A acts on a 3rd person P (15).

(15) *eɾe-nupã.*
2SG.I-hit
'You (SG) hit him/her/it/them.'

In the four cases presented in (14), the Set I index on the verb refers to the 2nd person A. In (14b, d), *oɾonekom*, the normal free pronoun for 1st person exclusive, refers explicitly to P, as expected. Surprisingly, in (14a, c), the independent forms, used only in this specific configuration, are linked to 2nd person pronominal forms (*eɾe-* and *pe-* are the Set I indexes for 2nd person singular and plural, the origin of *ɲ* is unexplained) rather than to 1st persons. The system seems *a priori* aberrant in synchrony: in order to refer to a 2nd person A a 1st person P, two markers for 2nd person are used. This peculiar use of the independent person marker can probably be explained as a hypothetical residue of a Proto-Tupí-Guaraní marker for A. The hierarchy in Proto-Tupí-Guaraní is supposed to have been 1 > 2 > 3 (Jensen 1998, Montserrat & Soares 1983). Therefore, in the same configuration, the 1st person P, being higher in the hierarchy, is indexed before the verb and the 2nd person A is then expressed with an independent marker following the verb, as illustrated in (16) from Tupinambá (Jensen 1998).

(16) *sjé* *r-epjak* *pejepé.*
1SG.II RELN-see 2PL.PRO
'You (PL) all see me.'

My hypothesis is that the 2nd person marker remained in the same position in this configuration in Emerillon, even though the 1 > 2 hierarchy was neutralized between the speech act participants. The semantic role hierarchy came into play to compensate this loss. It is not surprising that a language would not hierarchically distinguish between the speech act participants. 1st and 2nd person are independent within the hierarchy, their relative order fluctuates from one language to the other (Silverstein 1976).

When a 1st person acts on a 2nd person singular, Emerillon uses the Set I marker for 1st person exclusive (17).

(17) *oɾo-nupã.*
1EXCL.I-hit
'I/we hit you.' (but also: 'We hit it/her/him/them.')

When a 1st person (singular or plural) acts on a 2nd person plural, A is marked with a Set I marker for 1st person singular, and P with an index referring to generic human object (*poɾo-*).

(18) *a-poɾo-nupã.*
1SG.I-INDET.II-hit
'I/we hit you all.' (lit. 'I hit people.')

The reorganization of the person hierarchy as well as the substitution patterns illustrated above for some local configurations can be explained by po-

liteness.[11] It is tightly correlated with the fact that languages disfavor transparent marking of 1st and 2nd person combinations (Heath 1998). This assertion is confirmed by the fact that both in Emerillon (1/2 > 3) and in the hypothetical Proto-Tupí-Guaraní (1 > 2 > 3), the marking is very clear and systematic when only one speech act participant is involved, but less so when both 1st and 2nd person are involved.

3.3. *Emerillon within the typology of indexation systems*

To give a brief summary of the Emerillon person indexation system on the verbs, intransitive verbs take a Set I index, while transitive verbs allow only one person index, from Set I for A or from Set II for P. The correct index is selected according to the relative position of the two arguments on both the person hierarchy 1/2 > 3 and the semantic role hierarchy A > P. Two hierarchies are involved, and are ordered as follows:

(19) Person hierarchy (1/2 > 3) > Semantic role hierarchy (A > P)

The terms "person hierarchy" and "semantic role hierarchy" used here correspond roughly to other designations that emerged after Silverstein's pioneer work on hierarchies of features (Silverstein 1976): "animacy hierarchy" (Comrie 1981), "referential or inherent topicality hierarchy" (Givón 1994), "saliency hierarchy" (Klaiman 1991), "empathy hierarchy" (DeLancey 1981). A cover term that surfaced recently is "indexability hierarchy" (Bickel & Nichols 2007). Both hierarchies used in Emerillon could logically be justified in terms of saliency, the most salient participant being put forward. It is, nevertheless, important to assert that this system is completely grammaticalized: whatever the characteristics of the participants are, what counts in the system is the grammatical persons and the semantic roles.

The semantic role hierarchy becomes relevant when, and only when, the person hierarchy is not. I argue that this type of indexation system should be classified as hierarchical, since the primary organizing pattern is the notion of hierarchy (Cf. also Rose 2001, 2003a).

D. Payne formulated the hypothesis that Tupí-Guaraní languages could be described as having an inverse system (D. Payne 1994). On the basis of Givón's

---

11. It is a general fact about communication that the situations involving both speech act participants create a confrontation between the speaker and the addressee who is in a lower position. Languages often use devices like pluralization or substitution of a person for another to soften this confrontation (Brown & Levinson 1987). Thus, the plural form for a singular in (17) may be a means of weakening the 1st person subject, thus making the relation less "threatening" for the addressee. In (18), the use of a singular form for the subject and a generic form for the 2nd person creates a distance that softens the confrontation.

definition of the inverse, she explains that in these languages, there is an "inherent topicality hierarchy" $1 > 2 > 3$ (D. Payne 1994). If a 1st person acts on a 2nd or a 3rd person, or if a 2nd person acts on a 3rd, the action flows in the natural direction (A is more topical than P). This is considered to be direct. If the action flows the other direction (when P is more topical than A), that is considered to be inverse. Furthermore, a canonical inverse language is a language that expresses the inverse direction with morphosyntactic devices in a transitive construction. Examples from Fox, an Algonquian language, are repeated below (from Comrie 1981: 122), where *-aa* is a direct marker (20a) and *-ek* an inverse one (20b).

(20) a. *ne-waapam-aa-wa.*
1SG-see-DIRECT-3
'I see him.'

b. *ne-waapam-ek-wa.*
1SG-see-INVERSE-3
'He sees me.'

Applying this definition to Tupí-Guaraní languages, direct situations are those where Set I is used, i.e., when A is 1st or 2nd person, and P is 3rd person, but also when two 3rd persons are involved. Inverse situations are those where Set II is used, i.e., when a 3rd person A acts on a 1st or 2nd person P, and also in the local configuration where a 2nd person acts on a 1st person (as in (16) from Tupinambá). This analysis assumes that the two speech act participants follow the hierarchy $1 > 2$. The other local configuration where a 1st person acts on a 2nd person does not participate in the system: Payne analyses the forms as portmanteau. Finally, she proposes that the relational *r-* prefix, found in co-occurrence with Set II, be analyzed as an inverse marker in Tupí-Guaraní languages.

This hypothesis could be applied as it is to Emerillon, since the indexation system on independent verbs diverges from the rest of the family only in local configurations. However, a canonical inverse language is a language that expresses the inverse direction with an explicit morphosyntactic device. The relational morpheme that according to Payne is an inverse marker in Tupí-Guaraní languages (the *r-* relational) is found throughout the family on a lexically defined sub-class of transitive roots, but also on some intransitive roots, some nouns and postpositions when preceded with a Set II index. This distribution does not fit in with the function of the inverse category. No particular morpheme, therefore, can be interpreted as an inverse marker in Tupí-Guaraní languages. As a consequence, for the Emerillon indexation system, one would have be willing to consider only the mere distribution of person markers into two sets as an inverse system.[12] It seems to me that positing person and semantic role hierarchies is sufficient to explain the cross-referencing system.

12. This is actually what T. Payne (1997) calls "special verb agreement markers for inverse situations", citing data from Wayampí, a close relative of Emerillon.

This argument boils down to a confrontation between two different approaches. In the functional approach, inverse is to be found whenever P is more topical than A, but A is still topical (Givón (ed.) 1994). In the syntactic approach, hierarchical alignment and direction (inverse) are logically independent features that can, but need not, co-occur (DeLancey 2001, Zúñiga 2002). The so-called inverse systems are then considered a particular case of hierarchical systems (Nichols 1992, Siewierska 2004), the two being different possible expressions of deictic orientation (DeLancey 1981, no date). In line with Heath's argument (Heath 1998), I advise against the extension of the "inverse" terminology to forms lacking an inverse marker, which I see as undermining the usefulness of such a term. The Emerillon indexation system on verbs is then a plain hierarchical system (for further detail see Rose 2001, 2003a).

4. Syntax: Predicates and alignment

4.1. *Word order*

Emerillon can best be described as a verb-final language. In a pragmatically neutral sentence, the subject, if expressed with a full NP, is normally placed somewhere before the main predicate. Noun modifiers precede the head of the noun phrase, and adverbs and subordinate clauses are clause-initial. Oblique arguments and subordinate clauses are marked with postposed elements (21).

(21) [*pitaŋ o-kiʤe-ɾ-ehe*], *takuɾu-ũwĩ-ɾ-ehe o-wuɾ.*
child 3.I-be.scared-RELN-because rock-DIM-RELN-on 3.I-go.up
'Since the child is scared, he goes up on a small rock.'

The position of the object is more flexible: it may precede or follow the predicate (see, for example, (33c) and (21)). Topicality counts for preverbal and postverbal objects did not reveal a clear correlation between topicality and position of the object. In the rare situation where both subject and object are expressed with full noun phrases, the more frequent and basic order is SOV.

(22) *mama baʔeɾupiʔa o-bo-siu-siu.*
Mum egg 3.I-CAUS-RED-fry
'Mum used to fry eggs.'

Any kind of constituent may be focused by dislocating it to the initial position, where it carries second position particles with various discursive functions (like focus *-te* and interrogative/exclamative *-sipo* in the following example).

(23) *e-men-a-te-sipo iʤe a-ʔu!*
1SG.II-husband-REF-FOC-INTER/EXCL PRO.1SG 1SG.I-eat
'I ate my husband!'

The predicate is the only obligatory constituent. Besides verbal predication, another noticeable type of predication is based on a nominal root.

4.2. *Possessive nominal predicates*

Items that are unambiguously described as nouns can function as intransitive predicates and constitute complete clauses (24).

(24) *oɾone-kaɾakuɾi-nam pe-ɾ-upi oɾo-ho-taɾ-uwe.*
1EXCL.II-money-when way-RELN-on 1EXCL.I-go-FUT-too
'When we have money, we'll go that way too.'

The following example illustrates the fact that a noun, once preceded by a Set II marker (usually marking the possessor on a nominal head), can function either as an argument or as a predicate. The first occurrence of *iʤaɾiʤ* is a predicate ('she has a grandmother'), while the second occurrence is an argument ('her grandmother').

(25) *i-ʤaɾiʤ-a-nam, i-ʤaɾiʤ-a-te o-maʔẽ*
3.II-grandmother-REF-when 3.II-grandmother-REF-FOC 3.I-watch
Ø-ehe.
3.II-POSTP
'If she has a grandmother, then it is her grandmother that watches her.'

This type of nominal predicates shares the properties of verbal predicates. The following pairs of sentences show that nominal predicates (in (a)) and verbal predicates (in (b)) equally carry the plural subject marker (26), sentence negation (27), TMA (28), and voice markers (29).

(26) a. *e-iba-ŋ, kuʤatãĩkiɾ-a-kom.*
3.II-pet-PL.S teenage.girl-REF-PL
'They have a pet, the girls.'
b. *e-potaɾ-oŋ.*
1SG.II-like-PL.S
'They like me.'

(27) a. *d-e-sapato-ʤi-ãhã iʤe t-a-kʷa-ɲ.*
NEG-1SG.II-shoe-NEG-only PRO.1SG PURP-1SG.I-PASS-CONT
'I don't have any shoes for me to go.'
b. *d-o-ʔu-ʤi sautu.*
NEG-3.I-eat-NEG salt
'She does not eat salt.'

(28) a. *e-ɾ-apiʤ-taɾ.*
1SG.II-RELN-house-FUT
'I will have a house. '

b. *wane-maʔẽ za-ɨkɨʤ-taɾ apam-a-wi.*
good-REL INDET.I-take-FUT foreigner-REF-ABL
'We will take the good things to the foreigners.'

(29) a. *o-ze-mo-kasi-ne.*
3.I-REFL-CAUS-strength-CONTRAST
'He made himself strong.'

b. *zawaɾ-a-ɾ-ehe o-ze-mõ-boɾɨ.*
dog-REL-RELN-with 3.I-REFL-CAUS-make.happy
'He is having fun with the dog.'

Although "the predicative function cannot be taken as a serious criterion for distinguishing between nouns and verbs" (Queixalós 2006: 252), yet the distinction between noun and verb categories is clear: while nouns can function as predicates with the appropriate morphology, verbs need to function as the predicate of a relative clause to be used as arguments (cf. (30a, b)). A collective book (Queixalós (ed.) 2001), as well as a paper in historical linguistics (Queixalós 2006), have been devoted to the topic of parts of speech and their function in Tupí-Guaraní languages.

(30) a. *Polo o-manõ.*
Paulo 3.I-die
'Paulo died.'

b. *o-kuwa-pa o-manõ-maʔẽ.*
3.I-know-COMPL 3.I-die-REL
'He (God) knows all the dead ones.'

In the typological literature, "nominal predication" usually refers to cases where a noun has a predicative function, without necessarily having the same morphology as verbs. Thus, in many cases of so-called nominal predication, copulas (may) appear. Nominal predicates[13] are not completely comparable to verbal predicates. In this sense, nominal predicates typically convey the meaning of inclusion, equation, attribution, location, and existence (T. Payne 1997).

Nominal predicates in Emerillon substantially deviate from this characterization (Rose 2002b). First, nominal predicates display exactly the same morphology as verbs (plural subject marker, negation, voice, TAM, cf. (26)–(29)), which is therefore called "predicate morphology". They only differ in the indexation of their unique argument. Nominal predicates encode it with a Set II index and intransitive verbs with a Set I index. Second, the meaning of these predications is possessive. They contrast with other devices for 'be' functions, such as the use of a verb *ʤu* 'be' for inclusion (31a), or the copula *kob* for existence (31b).

13. "Predicate nominals" in standard terminology.

(31) a. *teko-a-te* *eɾe-ʤu!*
Emerillon-REF-FOC 2SG.I-be
'Emerillon you are!'
b. *kob-we* *i-bote.*
COP-too 3.II-motor
'There is also his motor.' [in a description task]

This type of nominal predication in Tupí-Guaraní languages, where nouns behave as predicates sharing most of their morphology with verbal predicates, and express possession, does not seem to be attested as such elsewhere.

4.3. *Descriptive words*

An on-going debate among specialists of Tupí-Guaraní and related languages concerns the issue of which part of speech "descriptive" words belong to (Dietrich 2000, Queixalós (ed.) 2001, Queixalós 2006, Meira 2006, *inter alia*). These items refer to qualities or states, take Set II person markers, and are typically used as predicates, as illustrated by the Emerillon examples below:

(32) a. *e-kaneʔõ.*
1SG.II-fatigue
'I am tired.'
b. *i-ɲuʔaŋ-oŋ.*
3.II-cold-PL.S
'They are cold.'
c. *d-e-baʔewaɾ-i.*
NEG-1SG.II-hunger-NEG
'I'm not hungry.'

This class of words has often been described by Tupi-Guaranists as stative verbs (Kakumasu 1986, Leite 1990, Jensen 1998, Seki 2000, Franceschini 2002). However, some authors analyze them as nouns (Dietrich 1977, Rodrigues 1996, Couchili et al. 2002, Praça 1999, DaSilva 2000, Meira 2006). It is noticeable that some of the authors working on the same language are on opposite sides of the debate (Kakumasu and DaSilva for Ka'apor, Franceschini and Meira for Mawe), so that we are basically facing a problem of analysis. However, one should keep in mind that there are also real differences in the behaviour of those terms from one language to the other. In Emerillon, descriptive words split into two distinct categories. The one discussed in this section is comparable to the descriptive category in the other Tupí-Guaraní languages. The other displays completely different behaviour: these roots cannot take any person markers and need additional morphology to function as predicates, modifiers, or heads of nominal phrases. The Emerillon situation, with two categories of descriptive words, shows that the "descriptive" domain is an unstable area within the family. This section will focus on the descriptive category that is comparable to the one present throughout the family and already illustrated in (32).

It is noteworthy that the split between these two categories is semantically based (Queixalós 2001a), respecting moreover the semantic classes defined by Dixon (Dixon 1977). The descriptive category discussed in this section expresses "human propensity" and "physical property", while the one diverging from the Tupí-Guaraní descriptive category expresses "size, color, value".

The two alternative analyses of descriptive words as nouns or verbs stem from the fact that descriptive words are usually used as predicates (like verbs), but are only once preceded by a Set II marker, just like the nouns as we observed in Section 4.2. Since descriptive predicates (32) are parallel to other nominal predicates (cf. Section 4.2), they are basically nouns. In contrast to verbs, they do not need to be relativized to function as arguments. The differences they may show from other nouns (i.e., being rarely used as NP heads, with plural or copulas) may easily be explained by semantic factors. These words basically refer to human qualities, physical sensations, and mental phenomena.

4.4. *The Emerillon alignment system*

Now that both verbal and nominal predication have been presented, the issue can be addressed of how the single argument of intransitive predicates aligns with the arguments of transitive predicates. Since arguments are not marked for case and constituent order is quite flexible, the argumentation will be based on the indexation pattern described above and on syntactic phenomena.

In Section 3.2, it was shown that the single argument of an intransitive verb aligns with the A argument of a transitive verb: both are indexed with a Set I person marker, while P is indexed with a Set II marker. This forms a nominative-accusative system. Moreover, number agreement can be used to identify the subject relation: the plural clitic *(o)ŋ* in (33) agrees exclusively with the unique argument of a verbal (33a) or nominal (33b) intransitive predicate and with the A argument of a transitive verb (33c), thus defining the subject relation.

(33) a. *o-paʔam-oŋ* *ikiʔi.*
3.I-get.up-PL.S then
'Then they got up.'

b. *i-awu-ŋ.*
3.II-word-PL.S
'They speak.'

c. *baipuɾi* *o-pɨhig-oŋ.*
tapir 3.I-catch-PL.S
'They caught a tapir.'

d. **a-ɨkɨʤ-oɪ̨.*
1SG.I-catch-PL.S
'I caught them.'

Another criterion is widely used by Tupi-Guaranists to define the subject grammatical relation. The use of an *o-* index for 3rd person possessors or objects of postpositions is specifically triggered by coreference with the subject, be it of the unique argument of a verbal (34a) or nominal (34b) intransitive predicate and the A argument of a transitive verb (34c).

(34) a. *o-ho i-kotɨ o-wɨ-kotɨ.*
3.I-go 3.II-to 3.COREF-mother-to
'He is going to her, to his mother.'
b. *Bopea t-aʔɨr-pɨrɨ o-kuɲa-wi.*
Monpera 3.II-son-more 3.COREF-sister-ABL
'Monpera has more kids than his (own) sister does.'
c. *bokal-a-pe o-akaɪ̨ o-mõde.*
jar-REF-in 3.COREF-head 3.I-put
'He put his (own) head in the jar.'

However, if we take into account nominal predicates, the alignment system can also be interpreted as one of split intransitivity, since the unique argument of intransitive predicates is marked differently. Set I is found on verbal predicates (whether unaccusative or unergative) and Set II on nominal predicates.

(35) a. *a-ʔɨta ʔɨ-pope.*
1SG.I-swim river-in
'I swim in the river.'
b. *e-kaneʔõ.*
1SG.II-fatigue
'I am tired.'

Most languages displaying split-intransitivity morphologically distinguish intransitive predicates on the basis of either their aspect (active or stative) or the agentivity of the unique argument (Mithun 1991). In line with the analysis of descriptive words as nouns (see Section 4.3), the Emerillon split-intransitivity system is untypical, as it does not operate within verbs, but between intransitive verbs and nominal predicates. These latter are specialized in expressing a subset of stative events, namely possessive predication, physical properties, and human propensities.

In conclusion, the Emerillon alignment system can be said to be nominative-accusative for verbs, where S and A are marked with Set I and P with Set II (cf. Section 3.2), but with an intransitivity split on predicates in general, including

nominal predicates (S and A are marked with Set I on verbs, while P and the unique argument of nominal predicates are marked with Set II).

5. Valency issues

As far as valency issues are concerned, Emerillon is representative of the Tupí-Guaraní family. A sociative causative is one of the features shared by the whole family which was not accounted for previously in the typological literature.

5.1. *Valency and valency-changing derivations*

Underived predicates are clearly either intransitive or transitive in Emerillon. This language has neither ambitransitive nor ditransitive predicates.[14] Two predicates sharing their arguments can combine in a single clause by means of a serial verb construction (as elsewhere, see Durie 2000, Pawley & Lane 1998, Schiller 1990). Serial verb constructions in Emerillon express direction (36), motion, or sequential action.

(36) *kaʔi* *o-weʤu* *o-ʔu.*
monkey.species 3.I-go.down 3.I-come
'The monkey is coming down.'

Reflexive-reciprocal is the only valency-reducing derivation of Emerillon. There is no passive.

(37) a. *o-ze-mim.*
3.I-REFL/RECIP-hide
'He hid himself.'
b. *si-ze-aɨhɨ-kom!*
PURP.1INCL.I-REFL/RECIP-love-PL.S
'Let us love each other!'

Emerillon also exhibits a mechanism for reorganizing valency, namely noun incorporation. Only dependent nouns can be incorporated, i.e., nouns that are obligatorily modified by a possessor when they are heads of a nominal phrase: *ɨru* 'clothes' is such a noun (37b), whose possessor must always be specified, except when incorporated as in (38a). Their incorporation is triggered by the assignment of an affected human patient to the object position, like 'mother' in (38a) (Rose 2008).

14. There is a subclass of intransitive verbs that mark a second participant with the postposition *ehe* 'to'. These verbs express cognition or perception events, like *maʔẽ* 'to see', *ʤapɨaka* 'to think'.

(38) a. *o-ɨɾu-a-nẽ-ʔe* *o-iɲuŋ-oŋ.*
3.COREF-clothes-REF-CONTRAST-INTENS 3.I-put-PL.S
'They were wearing clothes.'
b. *o-ɨ* *o-ɨɾu-mõde.*
3.COREF-mother 3.I-clothes-put.on
'He dressed his mother.' (lit. 'He clothes-put his mother.')

Finally, the three valency-increasing operations of Emerillon are all different varieties of causatives. The first is specialized for intransitive predicates and expresses direct causation (typically with physical manipulation of a patientive causee, as in (39a)). The second is specialized for transitive predicates and expresses indirect causation (typically with an agentive causee obeying to oral direction as in (39b)). The third is specialized for sociative causation on intransitive stems, as illustrated in (39c); it is described in greater detail in the following section.

(39) a. *wane iʤe* *a-mo-zaug.*
well PRO.1SG 1SG.I-CAUS-bathe
'I wash her well.' (lit. 'I make her bathe well.')
b. *iʤe-a-te* *lekol* *a-iɲuŋ-okaɾ.*
PRO1SG-REF-FOC school 1SG.I-put-CAUS
'I had the school settle here.' (lit. 'I made them put the school.')
c. *pe-ɾo-poɾahaʤ-kom.*
2PL.I-CAUS.SOC-dance-PL
'You made them dance with you.'

5.2. *Sociative causative*

Tupí-Guaraní languages, and among them Emerillon, have a specific marker for sociative causation, a supposedly rare feature. Sociative causation is a special type of causation, where the causer not only makes the causee do an action, but also participates in it. On Emerillon intransitive predicates, the sociative causative marker *eɾo-* (40b) contrasts with the regular causative marker *bo-* ∼ *mõ-* (40a); further instances of *eɾo-* are given in (40c–e).

(40) a. *wãĩwĩ* *o-mõ-keɾ* *ʔimaʔẽ.*
woman 3.I-CAUS-sleep child
'The woman is putting the child to sleep.'
b. *o-eɾ-aho* *o-eɾo-keɾ.*
3.I-CAUS.SOC-go 3.I-CAUS.SOC-sleep
'(The husband) carries (his new wife, who is drunk) and makes her sleep with him.'

c. *o-eɾo-kwa* *baʔezaʔu.*
3.I-CAUS.SOC-pass food
'She distributed food' (lit. 'She passed by and had food pass with her.')

d. *zawaɾ-enam* *ka* *o-(w)eɾo-nan.*
dog-TOP.SWITCH wasp 3.I-CAUS.SOC-run
'The dog runs with the wasps.' (The dog makes the wasps "run" with him.)

e. *zawaɾ* *o-eɾo-ʔaɾ* *tupawəɾ.*
dog 3.I-CAUS.SOC-fall box
'The dog makes the box fall (when falling himself).'

Shibatani & Pardeshi (2002) distinguish three nuances of sociative causation, depending on the extent to which the action of the causer is identical to that of the causee: (i) joint action, when both causer and causee perform the same action; (ii) assistive, when the causer helps the causee without performing exactly the same action; (iii) supervision, when the causer supervises the action performed by the causee. Within the domain of specific markers for sociative causative, the Emerillon case is restricted to intransitive predicates, and to the semantic nuance of sociative causation called joint-action, where both causer and causee perform the same action.

Sociative causation is usually seen as a possible meaning extension of regular causative markers, as discussed by Kulikov (2001) and Shibatani & Pardeshi (2002). This is fairly frequent crosslinguistically. In contrast, specific sociative causative markers are hardly ever discussed in general typology, with the exception of Dixon's typology of causative that posits a special type of causative called "causative of involvement" (Dixon 2000). Specific markers for sociative causative were probably first identified in the Tupí-Guaraní family, under the name of comitative causation (Adam 1896, Rodrigues 1953). The existence of such a supposedly rare type of causative both in Emerillon and in another Amazonian language, Cavineña (Guillaume 2008), prompted us to carry out a crosslinguistic survey of sociative causation (Rose & Guillaume 2007). This survey demonstrates that in actual fact specific sociative causative markers are not that rare. They are above all widespread in South American languages, being attested for instance in the Tupí, Tacanan, and Quechuan families.

6. Conclusion

In this Language Profile, I have tried to give a brief overview of Emerillon grammar by presenting in a succinct way its phonological, morphological, and syntactic components. Special attention was given to characteristics that are both central to the language and representative of the Tupí-Guaraní family as a whole. The main focus has been on specific features that are little represented in

typological surveys: nasal harmony, hierarchical indexation system, possessive nominal predication, and sociative causative. These features may be recognized as less unusual in the future, if – as we expect – they are to be found in some other Amazonian languages that still lack a comprehensive description.

Received: 10 September 2007
Revised: 19 August 2008

Centre National de la Recherche Scientifique
Université Lumière Lyon 2

Correspondence address: Laboratoire Dynamique du Langage, CNRS, Institut des Sciences de l'Homme, 14 avenue Berthelot, 69363 Lyon Cedex 07, France; e-mail: francoise.rose@univ-lyon2.fr

Acknowledgements: This paper was initially presented in a seminar at the Research Center for Linguistic Typology, La Trobe University, on 31 January 2007. I wish to warmly thank Antoine Guillaume, Katharina Haude, and Alexandre François for their valuable comments on an initial version of this paper, and Nick Clements as well as Elsa Gomez-Imbert and Leo Wetzels for stimulating discussions on nasality. I would also like to express gratitude to three anonymous reviewers. Denis Creissels and Colette Grinevald have always deeply inspired my work.

Abbreviations: 1/2/3 1st/2nd/3rd person; I/II Set I/II person index; A agent of transitive verb; ABL ablative; CAUS causative; CAUS.SOC sociative causative; COMPL completive; CONCES concessive; CONT continuous; CONTRAST contrastive; COP copula; COREF coreferential; DIM diminutive; EXCL exclusive; FOC focus; FUT future; INCL inclusive; INDET indeterminate; INTENS intensifier; INTER/EXCL interrogation/exclamation particle; NEG negation; P patient of transitive verb; PL plural; POSTP postposition; PRO pronoun; PURP purposive; RED reduplication; REF referential; REFL/RECIP reflexive-reciprocal; REL relativizer; RELN relational marker; S argument of intransitive verb; SG singular; TOP.SWITCH topic switch.

References

Adam, Lucien. 1896. *Matériaux pour servir à l'établissement d'une grammaire comparée des dialectes de la famille Tupi*. Paris: Maisonneuve.

Aikhenvald, Alexandra. 1996. Amazonian languages. Reading packet. (The 3rd Australian Linguistic Institute. Australian National University). Unpublished manuscript.

Anchieta, Joseph de. 1595. *Arte de gramatica da lingua mais usada na costa do Brasil*. Coimbra: Mariz.

Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description*, Vol. 3: *Grammatical categories and the lexicon*, 169–240. Cambridge: Cambridge University Press.

Brown, Penelope & Stephen Levinson. 1987. *Politeness*. Cambridge: Cambridge University Press.

Cabral, Ana Suelly. 2000. Fonologia da língua Jo'é. *Universa* 8(3). 571–596.

Cabral, Ana Suelly & Aryon Rodrigues (eds.). 2002. *Línguas indígenas Brasileiras: Fonologia, gramática e história. Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da ANPOLL*. Belem: Editoria Universitária U.F.P.A.

Clements, Nick & Sylvester Osu. 2005. Nasal harmony in Ikwere, a language with no phonemic nasal consonants. *Journal of African Languages and Linguistics* 26. 201–211.

Cohn, Abigail. 1990. Phonetic and phonological rules of nasalization. *UCLA Working Papers in Phonetics* 76. 1–224.

Comrie, Bernard. 1981. *Language universals and linguistic typology*. Oxford: Blackwell.

Couchili, Ti'iwan, Didier Maurel & Francesc Queixalós. 2002. Classes de lexèmes en émérillon. *Amerindia* 26/27. 173–208.

DaSilva, Beatriz Corrêa. 2000. Considerações sobre classes de palavras em Ka'apor. *Universa* 8(3). 597–607.

DeLancey, Scott. 1981. An interpretation of split ergativity and related patterns. *Language* 57. 626–657.

DeLancey, Scott. 2001. Lectures on functional syntax. Class given at the LSA Summer Institute, University of California at Santa Barbara. http://www.uoregon.edu/~delancey/sb/fs.html

DeLancey, Scott (no date). Deixis, topicality and the inverse. Unpublished manuscript.

Dietrich, Wolf. 1977. Las categorías verbales (partes de la oración) en tupí-guaraní). *Indiana* 4. 245–261.

Dietrich, Wolf. 1990. More evidence for an internal classification of Tupi-Guarani languages (*Indiana* Supplement 12). Berlin: Mann.

Dietrich, Wolf. 2000. Problema de la categoría del adjetivo en las lenguas tupí-guaraníes. In Hein van der Voort & Simon van de Kerke (eds.), *Indigenous languages of Lowland South America*, 255–263. Leiden: CNWS, Universiteit Leiden.

Dixon, R. M. W. 1977. Where have all the adjectives gone? *Studies in Language* 1. 19–80.

Dixon, R. M. W. 2000. A typology of causatives: Form, syntax and meaning. In R. M. W. Dixon & Alexandra Aikhenvald (eds.), *Changing valency*, 30–83. Cambridge: Cambridge University Press.

Dixon, R. M. W. & Alexandra Aikhenvald. 1999. Introduction. In Dixon & Aikhenvald (eds.) 1999, 1–21.

Dixon, R. M. W. & Alexandra Aikhenvald (eds.). 1999. *The Amazonian languages*. Cambridge: Cambridge University Press.

Durie, Mark. 2000. Grammatical structures in verb serialization. In Alex Alsina, Joan Bresnan & Peter Sells (eds.), *Complex predicates*, 289–354. Stanford, CA: CSLI.

Franceschini, Dulce. 2002. A voz inversa em Sateré-Mawé (Tupi). In Cabral & Rodrigues (eds.) 2002, 222–233.

Givón, Talmy. 1994. The pragmatics of de-transitive voice: Functional and typological aspects of inversion. In Givón (ed.) 1994, 3–44.

Givón, Talmy (ed.). 1994. *Voice and inversion*. Amsterdam: Benjamins.

Gomez-Imbert, Elsa. 1997. Morphologie et phonologie Barasana: Approche non-linéaire. Université de Paris 8 doctorat d'état thesis.

Gordon, Matthew & Françoise Rose. 2006. Emérillon stress: A phonetic and phonological study. *Anthropological Linguistics* 48. 132–168.

Grannier Rodrigues, Daniele. 1990. *Fonologia do guarani antigo*. Campinas: Unicamp.

Guillaume, Antoine. 2008. *A grammar of Cavineña*. Berlin: Mouton de Gruyter.

Heath, Jeffrey. 1998. Pragmatic skewing in 1–2 pronominal combinations in native American Languages. *International Journal of American Linguistics* 64. 83–104.

Jensen, Cheryl. 1998. Comparative Tupí-Guaraní morpho-syntax. In Desmond Derbyshire & Geoffrey Pullum (eds.), *Handbook of Amazonian languages*, Vol. 4, 490–603. Berlin: Mouton de Gruyter.

Jensen, Cheryl. 1999. Tupí-Guaraní. In Dixon & Aikhenvald (eds.) 1999, 125–163.

Kakumasu, James. 1986. Urubu-Kaapor. In Desmond Derbyshire & Geoffrey Pullum (eds.), *Handbook of Amazonian languages*, Vol. 1, 326–403. Berlin: Mouton de Gruyter.

Klaiman, M. H. 1991. *Grammatical voice*. Cambridge: Cambridge University Press.

Kulikov, Leonid. 2001. Causatives. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language typology and universals*, Vol. 2, 886–898. Berlin: de Gruyter.

Leite, Yonne. 1990. Para uma tipologia ativa do Tapirapé: Os clíticos referenciais de pessoa. *Cadernos de estudios linguisticos* 18. 37–56.

Maurel, Didier. 1998. *Eléments de grammaire émérillon* (Chantiers Amerindia 23(1)). Paris: Association d'Ethnolinguistique Amérindienne.

Meira, Sergio. 2006. Mawe stative verbs and predicate possession. In Wolf Dietrich & Haralambos Symeonidis (eds.), *Guaraní y "Mawetí-Tupí-Guaraní": Estudios históricos y descriptivos sobre una familia lingüística de América del Sur*, 47–68. Berlin: LIT.

Mello, Antônio. 2002. Evidências fonológicas e lexicais para a sub-agrupamento interno Tupi-Guarani. In Cabral & Rodrigues (eds.) 2002, 338–342.

Mithun, Marianne. 1991. Active/agentive case marking and its motivations. *Language* 67. 510–546.

Montserrat, Ruth & Marília Faco Soares. 1983. Hierarquia referencial em línguas Tupi. *Ensaios de linguïstica* 9. 164–187.

Navet, Eric. 1994. Introduction. In Ti'iwan Couchili & Didier Maurel (eds.), *Contes des indiens émérillon*, 1–11. Paris: Conseil International de la Langue Française.

Nichols, Johanna. 1992. *Linguistic diversity in space and time*. Chicago: University of Chicago Press.

Odden, David. 1994. Adjacency parameters in phonology. *Language* 70. 289–330.

Pawley, Andrew & Jonathan Lane. 1998. From event sequence to grammar: Serial verb constructions in Kalam. In Anna Siewierska & Jae Jung Song (eds.), *Case, typology and grammar: In honor of Barry J. Blake*, 202–227. Amsterdam: Benjamins.

Payne, Doris. 1994. The Tupi-Guarani inverse. In Barbara Fox & Paul Hopper (eds.), *Voice: Form and function*, 313–340. Amsterdam: Benjamins.

Payne, Thomas. 1997. *Describing morpho-syntax: A guide for field linguists*. Cambridge: Cambridge University Press.

Peng, Long. 2000. Nasal harmony in three South American languages. *International Journal of American Linguistics* 66. 76–97.

Piggott, Glyne. 1992. Variability in feature dependency: The case of nasality. *Natural Language & Linguistic Theory* 10. 33–78.

Praça, Walkíria. 1999. Nomes como predicados na língua Tapirapé. Brasília: Universidade de Brasília master's thesis.

Queixalós, Francesco. 2001a. Le tupi-guarani en chantier. In Queixalós (ed.) 2001, 1–20.

Queixalós, Francesco. 2001b. Le suffixe référentiant en émérillon. In Queixalós (ed.) 2001, 115–132.

Queixalós, Francesco. 2006. The primacy and fate of predicativity in Tupi-Guarani. In Ximena Lois & Valentina Vapnarsky (eds.), *Lexical categories and root classes in Amerindian languages*, 249–287. Bern: Lang.

Queixalós, Francesc (ed.). 2001. *Des noms et des verbes en tupi-guarani: Etat de la question*. München: Lincom Europa.

Rodrigues, Aryon. 1953. Morfologia do verbo Tupi. *Letras* Separata n°1. 121–152.

Rodrigues, Aryon. 1984–1985. Relações internas na família linguïstica Tupí-Guaraní. *Revista de Antropologia* 27/28. 33–53.

Rodrigues, Aryon. 1996. Argumento e predicado em Tupinambá. *Boletim da Associação Brasileira de Lingüistica* 19. 57–66.

Rodrigues, Aryon. 2003. Silêncio, nasalidade e laringalidade em línguas indígenas brasileiras. *Letras de Hoje* 38(4). 11–24.

Rodrigues, Aryon & Ana Suelly Cabral. 2002. Revendo a classificação interna da família Tupí-Guaraní. In Cabral & Rodrigues (eds.) 2002, 327–337.

Rose, Françoise. 2001. Cross-referencing in Emérillon (Tupi-Guarani): A hierarchical agreement system. In Jeanie Castillo (ed.), *Proceedings from the Fourth Workshop on American Indigenous Languages* (Santa Barbara Papers in Linguistics 11), 71–84. Santa Barbara, CA: Department of Linguistics, University of California, Santa Barbara.

Rose, Françoise. 2002a. Le problème de la nasalité dans l'inventaire phonologique de l'émérillon. *Amérindia* 26/27. 147–172.

Rose, Françoise. 2002b. My hammock = I have a hammock: Possessed nouns constituting possessive clauses in Emérillon (Tupi-Guarani). In Cabral & Rodrigues (eds.) 2002, 392–402.
Rose, Françoise. 2003a. Le marquage des personnes en émérillon (tupi-guarani): Un système d'accord hiérarchique. *Faits de Langues* 21(2). 107–120.
Rose, Françoise. 2003b. Morphosyntaxe de l'émérillon: Une langue tupi-guarani de Guyane française. Lyon: Université Lumière Lyon 2 doctoral dissertation.
Rose, Françoise. 2005. Reduplication in Tupi-Guarani languages: Going into opposite directions. In Bernhard Hurch (ed.), *Studies on reduplication*, 351–368. Berlin: Mouton de Gruyter.
Rose, Françoise. 2006. Le syncrétisme adpositions/subordonnants: Proposition de typologie syntaxique. *Faits de Langues* 28. 205–216.
Rose, Françoise. 2007. Action répétitive et action répétée: Aspect et pluralité verbale dans la réduplication en émérillon. *Faits de Langues* 29. 125–143.
Rose, Françoise. 2008. L'incorporation nominale en émérillon: Une approche lexicale et discursive. *Amerindia* 31. 87–112.
Rose, Françoise. To appear. *Grammaire de l'émérillon (teko), langue tupi-guarani de Guyane française* (Langues et Sociétés d'Amérique traditionnelle). Leuven: Peeters.
Rose, Françoise & Antoine Guillaume. 2007. Sociative causative markers in South-American languages: A possible areal feature. Paper presented at SSILA Annual Meeting, Anaheim.
Rose, Françoise & Odile Renault-Lescure. 2008. Contact-induced changes in Amerindian Languages of French Guiana. In Thomas Stolz, Rosa Salas Palomo & Dik Bakker (eds.), *Aspects of language contact: New theoretical, methodological and empirical findings with special focus on romanicisation processes*, 349–376. Berlin: Mouton de Gruyter.
Schiller, Eric. 1990. The typology of serial verb constructions. *Chicago Linguistic Society* 26. 393–406.
Seki, Lucy. 2000. *Gramática do Kamaiurá*. Campinas: Unicamp.
Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge: Cambridge University Press.
Shibatani, Masayoshi & Prashant Pardeshi. 2002. The causative continuum. In Masayoshi Shibatani (ed.), *The grammar of causation and interpersonal manipulation*, 85–126. Amsterdam: Benjamins.
Siewierska, Anna. 2004. *Person*. Cambridge: Cambridge University Press.
Silverstein, Michael. 1976. Hierarchy of features and ergativity. In R. M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aboriginal Studies.
Tsujimura, Natsuko. 1996. *An introduction to Japanese linguistics*. Cambridge, MA: Blackwell.
Walker, Rachel. 2003. Reinterpreting transparency in nasal harmony. In Jeroen van de Weijer, Vincent van Heuven & Harry van der Hulst (eds.), *The phonological spectrum, Part I: Segmental structure*, 37–72. Amsterdam: Benjamins.
Wetzels, Leo. 2007. Sobre a representação da nasalidade em Maxacali: Evidências de empréstimos do Português. In Ataliba Teixeira de Castilho, Maria Aparecida Torres de Morais, Sonia Lazzarini Cyrino & Ruth Vasconcellos Lopes (eds.), *Descrição, história e aquisição do Português Brasileiro: Estudos dedicados à Mary Kato*, 217–240. Campinas: Pontes.
Zúñiga, Fernando. 2002. Inverse systems in indigenous languages of the Americas. Zürich: Universität Zürich doctoral dissertation.