



# INTERNAL ARGUMENTS IN KATUKINA-KANAMARI AND BEYOND

Francesc Queixalós

## ► To cite this version:

Francesc Queixalós. INTERNAL ARGUMENTS IN KATUKINA-KANAMARI AND BEYOND. Cadernos de Etnolinguística, In press. halshs-03628963

**HAL Id: halshs-03628963**

**<https://shs.hal.science/halshs-03628963>**

Submitted on 3 Apr 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# INTERNAL ARGUMENTS IN KATUKINA-KANAMARI AND BEYOND<sup>1</sup>

Francesc Queixalós

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE, France

## 1 INTRODUCTION

### 1.1 THE LANGUAGE

### 1.2 TYPOLOGICAL OUTLINE

## 2 F-PREDICATES AND VALENCE

## 3 NOMINALIZATION

### 3.1 KANAMARI

### 3.2 BIA KATUKINA

## 4 NOMINALIZATION AND SUBORDINATION

## 5 CASE-MARKING

## 6 THE DIACHRONY OF NOMINALIZATION

### 6.1 BRINGING IN NEW TENSE-ASPECT MORPHOLOGY

### 6.2 SHAPING INDEPENDENT-CLAUSE ALIGNMENTS: THE ERGATIVITY BIAS

#### 6.2.1 ESKIMO

#### 6.2.2 MOVIMA AND TAGALOG

## 7 THE SYNCHRONIC PATTERNS OF KATUKINA-KANAMARI INDEPENDENT CLAUSE

## 8 INTERIM SUMMARY

## 9 A TAKE ON TUPI(-GUARANI) INTERNAL ARGUMENTS

### 9.1 PROVIDING NORMALIZED DATA

### 9.2 ACCOUNTING FOR THE DATA

#### 9.2.1 THE STRUCTURE OF PHRASES

#### 9.2.2 THE CASE FOR R-

## 10 RESULTS AND PROSPECTS

## REFERENCES

## ABSTRACT

My main point in this paper is to illustrate the way typology, as a (theoretically-informed) cross-linguistic comparison of facts, can help us unravel the grammatical make-up of a particular language / family of languages, and can — as a more ambitious offshoot — help give insights into issues of significance for our understanding of some important topics in grammar. For doing so, I take a cue from my analysis of Katukina-Kanamari's structure of complex phrases to join a persistent debate surrounding certain morphosyntactic aspects of Tupi / Tupi-Guarani languages. Between both ends, other languages — Movima, Eskimo, Tagalog — contribute toward further sharpening the specific concern addressed in the paper, internal argumenthood. The conspicuous parallelisms that can be spotted between the languages reviewed and Tupi(-Guarani) open the door for drawing some preliminary conclusions with respect to lingering uncertainties over how to account for intricate aspects of morphology in this latter family of languages.

**KEY-WORDS:** internal arguments, phrase constituency, nominalization, diachrony, alignment.

---

<sup>1</sup> Uploaded November 30 2021

## 1 INTRODUCTION

The notion of internal (*vs.* external) argument has been put forward and mainly used by various approaches within the generative tradition (Williams 1981), a direct consequence of the emphasis put on constituency as a cornerstone of the grammatical architecture proper to, mainly, well-studied languages. Taking a typological stand, I will here avail myself of the concepts of *constituency* and *internal argument* to bring to light non-trivial isomorphisms between complex phrases in a small sample of languages. Katukina-Kanamari will first contribute with a thorough instance of parallelisms involving constituency and case-marking that obtain between phrases headed by transitive verbs, relational nouns and postpositions. Particularly, the idea that nouns are endowed with valency properties will arise as we review some prominent characteristics of Katukina-Kanamari grammar. Through tentative internal reconstruction, we will gain from this virtually isolated language a preliminary appraisal of the impact of verb nominalization on the diachronic emergence of basic independent-clause patterns of alignment. The results thus attained will then be shown to echo to a significant extent cardinal features in the grammars of Eskimo, Movima and Tagalog. In order to illustrate how the hypotheses set out for Katukina-Kanamari fit in the overall synchronic morphosyntax of the language, an outline of independent-clause patterns will be supplied before turning to what might be seen as the fulfilment of the whole discussion: an attempt at unveiling the very grammatical nature of a Tupi(-Guarani)'s rather peculiar morpheme supposed to signalize "contiguity". The final part of the paper is devoted to delineate the issues yet to be addressed for a better understanding of Tupi(-Guarani) grammars, particularly the apparent conundrum emerging from strong constituency and non-configurationality as hypothesized for the family.

### 1.1 THE LANGUAGE

The people speaking Katukina-Kanamari are scattered in a large rainforest area of Western Brazil, from the middle Jurua up to the Javari. The language belongs to the Katukinan family, itself surrounded by Panoan languages to the West and South, Arawan and Arawakan to the East, and Tikuna plus the Tupi-Guarani Kokama to the North. Katukina-Kanamari is a single language split in two dialects. The Katukina dialect is circumscribed by the Bia river, a tributary of the Jutai. The Kanamari data here used were collected in the upper Itaquai (Vale do Javari). The other language in the family might be Katawixi, reputed to remain extant to the South-East, outside the mentioned area. Recent academic research on Katukina-Kanamari includes Adelaar (2000, 2007), dos Anjos (2005a, 2012, 2005b, 2011), and Ishy de Magalhães (2012, 2014). Since 1995, the present author has published fifteen articles on grammatical topics, plus a co-written Queixalós & dos Anjos (2007).

### 1.2 TYPOLOGICAL OUTLINE

Major morphosyntactic characteristics of Katukina-Kanamari include morphology of modest complexity, solid and far-reaching constituency structure, clause basic-order with initial predicate phrase, and head-final complex phrases (*i.e.* phrases containing some word-size dependent). Lexical classes include verbs, nouns, postpositions, adverbs and particles. TAM distinctions are almost exclusively coded by particles. Verbs, nouns and postpositions occur as heads of complex phrases, and also as clause predicates. Nouns in predicate position do not require any copulative element. The language lacks trivalent verbs (excepting, perhaps, 'say').

Transitive verb arguments are formally contrasted outright in several dimensions. No intransitive split between unergatives / unaccusatives has been so far detected besides noun incorporation ("possessor" ascension). Valence changing devices encompass a de-transitivizing voice as well as applicative, causative, and noun incorporation derived constructions. Above all, what makes Katukina-Kanamari a typologically and theoretically interesting language is the absolutive-ergative alignment that pervades its grammatical architecture. I will hereafter be broaching a few key issues related to this characteristic.

## 2 F-PREDICATES AND VALENCE

Katukina-Kanamari is emblematic as to extensive formal parallelisms between nouns and verbs. In part this stems from nouns being predicates in the two most basic meanings of the term *predicate* that we commonly ascribe to verbs (see a second factor in 4 NOMINALIZATION AND SUBORDINATION *in fine*). First, as the pivotal piece of information in the clause a predicate bears the Praguean *rheme*, itself a legacy of the Aristotelian dichotomy *onoma* / *rhema*. Second, by way of its embodiment in a lexical class of items endowed with argument structure, a predicate is a logical n-tuple operator (function) with places (operands, variables) requiring saturation by specific values (Frege 1891). I will henceforth speak — when useful — of A(ristotelian)-predicates /vs./ F(reguean)-predicates, respectively. Nouns — likewise verbs — are F-predicates, which entails featuring valence values. In Katukina-Kanamari this fact — along with their capacity to occur as A-predicates — is conspicuous: example (1) shows two monovalent F-predicates in A-predicate position, noun in (a) and verb in (b); in (2) the same parallelism obtains between divalent F-predicates as A-predicates, again noun in (a) and verb in (b).

- |                                                                              |                                                                       |
|------------------------------------------------------------------------------|-----------------------------------------------------------------------|
| (1) (a) <i>piya adu</i><br>man 1SINGULAR<br>'I'm a man.'                     | (b) <i>i:kao adu</i><br>cry 1SINGULAR<br>'I cried.'                   |
| (2) (a) <i>a-pama adu</i><br>3SINGULAR-father 1SINGULAR<br>'I'm his father.' | (b) <i>a-hak adu</i><br>3SINGULAR-spear 1SINGULAR<br>'He speared me.' |

Noticeably, the (b) examples show that the argument realizing the single participant of 'cry' aligns with the argument realizing the patient of 'spear': post-predicate position and identical first-person pronoun. As for both characteristics, in (a) examples the argument referring to the one who is 'father-of' aligns with the argument referring to the one who is 'man'. Thus, Katukina-Kanamari displays two subclasses of nouns along the same lines as verbs, monovalent in (1), divalent in (2).

The mono- / di-valent contrast in verbs is commonly dubbed *intransitive* /vs./ *transitive*. On the noun side, we have *alienably possessed* /vs./ *inalienably possessed*. As for nouns, more accurate (less rooted in intuitive semantics) labels are perhaps *relational* /vs./ *non-relational*. However, none of these three terminological dyads can be generalized over both nouns and verbs. We further note that the argument realizing the agent of the verb 'spear' in (2b) shares its formal properties with the genitive argument of the noun 'father' in (2a): a person prefix. This series of prefixes — supplied in (3) hereafter — is, primarily, the coding for the genitive dependent of divalent ("inalienably possessed") nouns.

Of course, semantic-role labels such as *patient* / *agent* are unworkable with respect to most nouns that are not verb-derived. As regards languages endowed with a valence-based bipartition of the class of nouns, I submit *intrinsic linkee* for the participant surfacing as the

divalent noun's genitive — *i.e.* the entity *inherently* coupled to the entity denoted by the head noun —, letting go by the name of *possessor* the participant that surfaces as the monovalent noun's genitive — *i.e.* the entity *transitorily* coupled to the entity denoted by the head noun (for further justification and illustrations see Queixalós 2017a). In Katukina-Kanamari noun phrases with a possessor participant exhibit a radically different structural template from those involving an intrinsic linkee.

As we will fully confirm below in dealing with complex noun and verb phrases (**5 CASE-MARKING**), a crucial consequence of the ongoing discussion is that the genitive dependent of a divalent noun and the argument coding the agent of a divalent verb are internal arguments of their respective heads, whereas the genitive dependent of a monovalent noun is a modifier.<sup>2</sup> Another no-less crucial consequence is, regarding verbs, an absolutive-ergative alignment driven by valence-and-constituency, (2a-b). In what follows I will use superscripts <sup>I/II</sup> to flag the valence of a given F-predicate, for instance <sup>I</sup>noun /vs./ <sup>II</sup>noun.

(3) shows free personal pronouns and the argument-indexing series on <sup>II</sup>F-predicates, respectively (slightly different forms between both dialects occur for third-person pronouns and the first-singular prefix):

(3)	singular	plural	singular	plural
	<i>adu</i>	<i>adik</i>	<i>yo-</i>	<i>tyo-</i>
	<i>idi:k</i>	<i>idi:ki</i>	<i>no-</i>	<i>na-</i>
	<i>anyan</i>	<i>anyan hinuk</i>	<i>(h)a-</i>	<i>ma-</i>

### 3 NOMINALIZATION

Nominalizing verbs is an exceedingly useful and massively used device in languages. The reason for this is two-fold. First, bringing a manner-of-existing together with its participant(s) into a non-verbal position within clause structure serves to describe an entity-type, build reference, and hierarchize A-predicates, three central mechanisms put to work by speakers in forging utterances. Second, nominalizations allow for doing so in a specially-compact make-up. In this section I will broach the synchronic mechanisms yielding oriented nominalizations ("participant nominalizations") in Katukina-Kanamari. Their form and function are somewhat different from one dialect to the other.

### 3.1 KANAMARI

Non-oriented nominalizations use a suffix *-nin*.

(4) *waok-dyi-nin* *hon*  
arrive-CENTRIPETAL-NOMINALIZER CAUSE

<sup>2</sup> As an attempt at terminological prophylaxy, I use *argument* for a linguistic expression — including zero — that bears a structural relationship to a predicate (*cf.* "syntactic argument"), along with representing one of the semantically obligatory participants. Arguments qualify as *nominative*, *accusative*, etc.; participants as *agent*, *patient*, etc. (and grammatical relations as *subject*, *object*, etc.). I further appeal to the notion of "constituency" adopting for *internal / external* a relatively theory-neutral definition, based on the observable syntagmatic groupings of clause components (*cf.* "surface constituent structure"; also Givón's "observable components of grammatical structure" 1995 ch. 5). All types of meanings conveyed by predicates — actions, events, processes, states, properties, qualities, inclusion (class-membership), identification, comparison, possession, etc. — will be here subsumed under the cover phrase *manner-of-existing*. Finally, an *adjunct* is an expression bearing no direct syntactic relation to the predicate and denoting either a non-obligatory participant or a backgrounded manner-of-existing (a third function for *adjunct* will be put forth below).

'because of (his) arrival here'

A deictic-based free form *nyan* related to a demonstrative *anyan* — see (3) — follows the <sup>II</sup>verb stem to produce patient-oriented nominalizations. As (5) shows, *nyan* nominalizations align absolutely. To get an agent-oriented nominalization out of this pattern a voice-changing prefix *wa-* preempts the person prefix slot (agent), thus deriving a <sup>I</sup>verb, (6) (antipassive; see (55) below for a finite clause).

- (5)(a) *a<sub>2</sub>-wahak*                      **nyan<sub>I</sub>**  
           3SINGULAR-<sup>II</sup>cook    DEICTIC  
           'her<sub>2</sub> cooked thing<sub>I</sub>'
- (b) *tyuku*    **nyan**  
           die    DEICTIC  
           'the dead one'
- (6)    **wa-wahak**    *nyan*  
           VOICE-cook    DEICTIC  
           'the cook'

### 3.2 BIA KATUKINA

The *nyan* device for nominalizing is unknown in this dialect, and the nominalizer *-nin* is more widely used as it serves for non-oriented, patient-oriented, and some otherwise-oriented nominalizations. Parallely to (5a-b), examples (7b-c) show an absolutive alignment (monovalent English 'cook' in (c)).

- (7)(a) *kirian-nin*                      (b) *tyo-tikok-nin*                      (c) *pia-nin*                      (d) *hoka-nin*  
           thunder-                      1PLURAL-know                      <sup>I</sup>cook-                      BeClean-  
           NOMINALIZER                      -NOMINALIZER                      NOMINALIZER                      NOMINALIZER  
           'thundering'                      'the one we know'                      'cooked thing'                      'dancing place'

In spite of having, likewise Kanamari in (6), a voice-derived agent nominalization, (8), Bia Katukina has innovated in creating a primary agent-nominalization, in which the verb is flanked on both sides by a discontinuous nominalizer *mo-...-hi*. (The second component *-hi* might be related to 1. a collective in noun morphology, (23), or 2. an unconfirmed habitual in verb morphology; presently, *mo-* remains opaque.) The person series of prefixes opens the nominalization sequence and now stands for the patient participant, (9). See *mo-...-hi* nominalizations as A-predicates in (10) and further below (26).

- (8) **wa-hak-nin**                      *daan-di*  
           VOICE-spear-NOMINALIZER    go-CENTRIPETAL  
           ' The fisherman came. '

(-di / -dyi are dialectal variants.)

- (9) [**no<sub>I</sub>-mo-ku-hi**]<sub>2</sub>  
           2SINGULAR-NOMINALIZER-bite-NOMINALIZER  
           'the one that bit you<sub>I</sub> (lit. [your<sub>I</sub> biter]<sub>2</sub>)'

- (10) *no<sub>I</sub>-mo-hak-hi*                      *adu<sub>2</sub>*  
           2SINGULAR-NOMINALIZER-spear-NOMINALIZER    1SINGULAR  
           'I<sub>2</sub> am the one who speared you<sub>I</sub> (lit. I<sub>2</sub> (am) [your<sub>I</sub> spearer]<sub>2</sub>) '

## 4 NOMINALIZATION AND SUBORDINATION

Cross-linguistically, subordinate clauses tend to specialize in form and syntagmatic position according to the functions mentioned in 3 **NOMINALIZATION**, all of them calling for a more or less radical loss of finiteness: relative clauses for identifying a referent; complement clauses for saturating an F-predicate's variable with a referent or manner-of-existing, and adverbial clauses for providing background information with respect to the manner-of-existing described by A-predicates in the main narrative line. Languages differ regarding the extent to which they make use of nominalization for subordinating (Givón 2015:665). (11) is taken from Sikuaní, a typical language as to vast-scale subordination through nominalization. Another such language is Ute, (12).

- (11) [[*pihawa*] *pe-xai-nae-nū*]<sub>1</sub> [*pe-n-ue-hawa*]<sub>2</sub> *ø<sub>2</sub>-kopata-ø<sub>1</sub>*  
 wife 3-own-IRREALIS-MASCULINE 3-cry-IRREALIS-NONANIMATE 3-GiveUp-3  
 'The husband stopped crying (lit. the [wife] owner)<sub>1</sub> stopped [his crying]<sub>2</sub>). '
- (12) '*áa'apachi-u* [ [*páa-y*] '*ivi-kya-paa-chi-mu*]<sub>1</sub> *ásti-kya-y*  
 BoySUBJECT- water-OBJECT drink-PLURAL-IRREALIS- want-PLURAL-IMMINENT  
 PLURAL NOMINALIZER-PLURAL  
 'The children want [to drink [water] ]<sub>1</sub>. '

UTE (Givón 2011:215; bracketing mine.)

See Givon (2015:666) for a sample of linguistic families behaving likewise.

Katukina-Kanamari resorts to nominalization in a similar way, as (13) shows for Bia Katukina, and (14) for Kanamari.

- (13) *yo*<sub>1</sub>-*tohi:k* [*ha*<sub>2</sub>-*ti-nin* *takara*<sub>3</sub>]  
 1SINGULAR-see 3SINGULAR-kill-NOMINALIZER hen  
 'I<sub>1</sub> saw him<sub>2</sub> kill the hen<sub>3</sub> (lit. I<sub>1</sub> saw [the hen<sub>3</sub> (being) his<sub>2</sub> killed one<sub>3</sub>]).'
- (14) *tona ityaropu*<sub>1</sub> [*yo*<sub>2</sub>-*wu-tu-nin*] [Ø<sub>1</sub>]  
 go girl 1SINGULAR-want-NEGATION-NOMINALIZER 3  
 'The girl<sub>1</sub> went away (because) I<sub>2</sub> didn't want her<sub>1</sub>  
 (lit. the girl<sub>1</sub> went away [my<sub>2</sub> not wanting her<sub>1</sub>]).'

In Bia-Katukina the agent-oriented nominalization also serves for subordinating.

Unsurprisingly, *-nin* as a subordinator does not occur in this pattern of nominalization.

- (15) *yo<sub>1</sub>-wu-tu* [ [*ha<sub>2</sub>-mo-koni-hi*] [Ø<sub>3</sub>] ]  
 1SINGULAR-want-NEGATION 3SINGULAR-NOMINALIZER-catch-NOMINALIZER 3  
 'I<sub>1</sub> don't want him<sub>3</sub> to catch (it)<sub>2</sub> (lit. I<sub>1</sub> don't want [he<sub>3</sub> (be) its<sub>2</sub> catcher<sub>3</sub>]).'

(See (26) for the syntagmatic position of the argument here represented by  $\emptyset$ .)

## 5 CASE-MARKING

Thus far we have met two types of F-predicates, monovalent and divalent, and two lexical classes embodying them, nouns and verbs. The two arguments encompassed by <sup>II</sup>F-predicates

contrast with each other in form, constituency and distribution. One argument surfaces as a noun phrase or a free pronoun, is external to the F-predicate phrase, and typically occurs after the A-predicate, *adu* in (1)-(2). It denotes: on nouns, the entity characterized / identified by the predicate; on verbs, the patient. The other argument is internal and surfaces as a prefix, *a-* in (1)-(2). On nouns it denotes the intrinsic linkee, on verbs the agent. The external argument can be zeroed in sufficiently strong pragmatic contexts (*pro*). The internal argument is mandatorily overt, and can only occur once, either as a person prefix or, as we will now see, as a noun phrase. This latter exponent of the internal argument occurs before the <sup>II</sup>F-predicate with an obligatory morpheme *na* intervening. Resuming (2):

- |         |                             |            |     |                            |            |
|---------|-----------------------------|------------|-----|----------------------------|------------|
| (16)(a) | [Kora... <b>na</b> ...pama] | <i>adu</i> | (b) | [Kora... <b>na</b> ...hak] | <i>adu</i> |
|         | Kora... <i>na</i> ...father | 1 SINGULAR |     | Kora... <i>na</i> ...spear | 1 SINGULAR |
|         | 'I'm Kora's father.'        |            |     | 'Kora speared me.'         |            |

Phonologically *na* is bound to the phrase head — noun, verb — on its right: /Kora#napama/, /Kora#nahak/ (see Queixalós 2010 and 2017b). Dos Anjos (2011:134) reports that speakers accept a pause between *na* and the preceding element, but consistently reject it between *na* and the subsequent element.

There are, however, strong reasons to assume that this morpheme started out diachronically as an allative-case suffix, to end up as a phrase-head proclitic after having undergone a re-grammaticalization pathway along the lines of

- (17)
1. allative, (20) from Bia Katukina;  
>
  2. recipient of 'give'-type verbs, (21) from Bia Katukina;  
>
  3. genitive in <sup>II</sup>noun phrases, (22) from Kanamari;  
>
  4. genitive (agent) in <sup>II</sup>verb nominalizations, (23) from Kanamari; (24) from Bia Katukina;  
>
  5. agent argument on <sup>II</sup>finite verbs, (25) from Bia Katukina;  
>
  6. internal argument head-proclitic on <sup>II</sup>F-predicates, *i.e.* nouns, verbs, and postpositions, (27) from Kanamari.

Heuristically, the semantic cogency of steps 1. to 5. in languages such as Katukina-Kanamari, deprived of any historical record and with little possibility of comparative inferences, is mainly inferrable from synchronic case-syncretism. Elsewhere, typology and historical records back up the soundness of the evolution / polysemy put forward in (17). Among many others:

- (18)
- |           |                                                                                                                                                                                                                                              |
|-----------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 1 > 2     | Atong (Breugel 2008:338)<br>Naukan Eskimo (Menovščikov 1975:79-100, cited in Kibrik 2003),<br>English                                                                                                                                        |
| 1 > 2 > 3 | French (as to 2 > 3: former stages of the language — and to-day's colloquial register — possibly <i>via</i> the dative external-possessor construction of the <i>mihi est</i> type, common in Ancient Indo-European languages, Luraghi 2020) |
| 3 > 5     | Lak (Kibrik 2003)                                                                                                                                                                                                                            |



	Aleut (Menovščikov 1968:391, 403, cited in Kibrik 2003).
1 > 2 > 5	Northern Carib (Gildea 2002a)
	Dargwa (Melčuk 2006:143)
1 > 3 > 5	Pano (Valenzuela 2003:323)
3 > 4 > 5	Eskimo (Mahieu 2004, 2009).

Other instances of similar syncretisms can be found in Palancar (2009).<sup>3</sup> Following this author the allative is, among spatial cases, the less prone to be used for marking the agent argument, and such association might be circumscribed to South America. Some of these putative semantic drifts — the initial ones — are consistent with well known grammaticalization pathways undergone by case-markers (Heine 2009), ultimately based on cognitive-metaphorical affinities (Luraghi 2009). Now, other drifts in (17) are more formally motivated:

1. the sequence genitive > agent marker has as a transitional step a <sup>II</sup>verb nominalization (and its subsequent re-finitization, see below **6.1 BRINGING IN NEW TENSE-ASPECT MORPHOLOGY**), on which the following structural constraint obtains: a <sup>II</sup>F-predicate — noun, verb — indexes uniformly its internal argument;
2. likewise, 6 is not as much an end point as the synoptic picture achieved by adding postpositions to steps 3 > 4 > 5 due to their inherent character of <sup>II</sup>F-predicates (on adpositions as a dyadic class, see Hale & Keyser nd.). A key factor in propounding 6 is the cross-linguistically common <sup>II</sup>nominal origin of postpositions (*e.g.* Givón 2021), a surmise, though, with yet little empirical footing in Katukina-Kanamari besides (*h*)*a-pori*, 'his back' > *hak por*i, 'behind the house' and, in the Bia dialect, the occurrence of the allative suffix on spatial postpositions, compare (19)-(20):

- (19) *hak iki-na da-dyoran*  
house INESSIVE-ALLATIVE DISLOCATIVE-go  
'He went into the house.'

The whole of the presumed diachronic pathway in (17) is only discernable in Bia Katukina, since Kanamari lacks 1. and marks the recipient role of 'give'-type verbs — syntactically an adjunct — with a postposition.

allative

- (20) *yo-dahu wankurun hak-na*  
1SINGULAR-carry pan house-ALLATIVE  
'I took the pan to the house.'

recipient

- (21) *Ayobi...na...nuhuk poako Kontan-na*  
Ayobi...na...give paddle Kontan-RECIPIENT  
'Ayobi gave a paddle to Kontan.'

intrinsic linkee

- (22) *opatyin...na...wadik*  
child...na...name

<sup>3</sup> At no moment I am suggesting an all-inclusive grammaticalization chain 1 through 5. See Givón (2015:305) for a dismissal of grammaticalization chains, much in the spirit of the global-rule rejection in early generative phonology.

'child's name'

agent in Kanamari *nyan* nominalization, patient-oriented, *cf.* (5a)

- (23) *paiko-hi= nuk...na...bahon-nyan*  
 elder-COLLECTIVE= group...na...replace-NOMINALIZER  
 'those replaced by the elders (lit. the elders' replacees)'

agent in Bia Katukina *-nin* nominalization, patient-oriented, *cf.* (7b)

- (24) *Hi:wuk...na...ka:k-nin*  
 Hi:wuk...na...break-NOMINALIZER  
 'that which Hi:wuk broke'

agent in finite <sup>II</sup>verb (both dialects)

- (25) *Ayobi...na...wapukni tawa*  
 Ayobi...na...DigOut manioc  
 'Ayobi dug out manioc tubers.'

From the systematic absence of the person prefix on verbs in (22)-(25) one can safely infer that the noun phrase followed by ...*na*... is indeed the expression of the internal argument.

Remarkably, the grammaticalization of ...*na*... has been pushed far enough to deprive it of any semantic-role correlate. It thus becomes apt, as a purely formal tool for marking internal arguments, to tag the patient in Bia Katukina *mo-...-hi* nominalization, agent-oriented, *cf.* (9).

- (26) *dyara...na...mo-wu-hi* *anya*  
 WhitePeople...na... NOMINALIZER-want-NOMINALIZER woman  
 'The woman likes Whites (lit. the woman (is) a Whites liker).'

(In the context: 'white men'.)

So as to visualize the synchronic mismatch that ...*na*... features regarding its grammatical /vs./ phonological status, I will henceforth represent it as *-na=* — grammatically a suffix, phonologically a proclitic — and write it bound to its grammatical host. Also, as the single overt grammatical ("structural") case-marker in the language, *-na=* will be merely glossed *grammatical case-marker*. (27) displays ...*na*...'s unified notation, as well as its uniform grammatical nature: an internal-argument case-marker in <sup>II</sup>F-predicate phrases (<sup>II</sup>noun, <sup>II</sup>verb, and postposition).

(27)

<sup>II</sup>noun

- (a) [ [*opatyin-na=*] *wadik*]  
 child-GRAMCASEMARK= name  
 'child's name'

<sup>II</sup>verb

- (b) [ [*Mayon<sub>1</sub>-na=*] *ikubalama*] [Ø<sub>2</sub>]  
 Mayon-GRAMCASEMARK= DigOut 3  
 'Mayon<sub>1</sub> dug it<sub>2</sub> out.'

postposition

- (c) [ [kotyia-**na**=]                      *katu*]  
       otter-GRAMCASEMARK=        COMITATIVE  
       'with the otters'

Let us observe, *en passant*, that *-na*= is subject to animacy-based differential marking with noun and postposition heads; compare (27a) with (28a), and (27c) with (4), resumed here as (28b). See, on the contrary, *-na*= occurring on a non-animate verb's dependent.

- (28) (a) [ [*ityowa ityonin*] *wadik*]  
           our        territory    name  
           'our territory's name '
- (b) [ [*waok-dyi-nin*]                      *hon*]  
       arrive-CENTRIPETAL-NOMINALIZER    CAUSE  
       'because of (his) arrival here'
- (c) *oman-na*=                      *ti tukuna*  
       tree- GRAMCASEMARK=        kill    people  
       'Trees kill people. '

Of special note is that, more generally, the language shows a clear propensity to a syntagmatic drift of bound / unstressed elements toward phrase heads. Thus, a collective suffix *-hi* cliticizes to the right located nominal — see it in (3) *anyan hinuk*, and (23) — and so does, to the right located auxiliary, the subordinating suffix *-nin*:

- (29) [ [*ikao-nin*=]                      *bak*]  
       cry-NOMINALIZER=    BeGood  
       'cry a lot (lit. be good at crying)'

pronounced /*ikao#ninbak*/ (Queixalós 2017b).

This phenomenon is well attested cross-linguistically, (30), and diversely dubbed (*head attraction* in Haig 2008:336; *headward-migrated dependent marking* in Nichols & Bickel 2005, among others).

- (30) English        *if a dog's **gotta** go then it's **gonna** go*  
       Gitksan        (Rigsby 1986)  
       Hausa        (Heine 2009b)  
       Movima        (Haude 2010a)  
       Nheengatu    (Cruz 2017)  
       Otomi        (Palancar 2016)  
       Panare        (Payne 1997:104)  
       Romance        (Vincent 2001)  
       Slavic        (Vincent 2001)  
       West Iranian (Haig 2008:336)

## 6 THE DIACHRONY OF NOMINALIZATION

I will now broaden the scope of nominalization issues beyond the above succinct description: first, by putting diachrony at the heart of the overall picture; then, by sketching the scenario along which nominalization feeds through to significant aspects of synchronic morphology and syntax; and finally yet importantly, by fleshing out the scenario with a choice of data from a few other, genetically unconnected, languages.

### 6.1 BRINGING IN NEW TENSE-ASPECT MORPHOLOGY

So far, I have presented the set of Katukina-Kanamari facts that evidence, on a synchronic basis, the tightness of the link between *-nin* as a nominalizer and *-nin* as a subordinator. A third function of this morpheme lies in marking durative aspect on finite verbs, (31). Significantly, *-nin* is the only verb-bound exponent of TAM categories (if we except the uncertain habitual *-hi*). Other TAM distinctions are coded by free standing particles (a plausible clue to the loss of finiteness exponents in verb morphology and the subsequent implementation of new tools for TAM), (32)-(33). Let us further note that neither *-nin* nor any particle occurs for conveying TAM semantics on nominalizations / subordinate clauses, pointing unambiguously to the non-finiteness of both.

- (31) *paiko-na=* *dyuru-nin* *ityaro*  
 AdultMan-GRAMCASEMARK= <sup>II</sup>HaveSex (lit. penetrate)-DURATIVE woman  
 'The man was having sex with the woman.'

- (32) *taanin Hayo-na=* *pu wa*  
 later Hayo-GRAMCASEMARK= eat FUTURE  
 'Hayo will eat (it) later.'

- (33) *Karon dyan* *dirin*  
 Karon GoHunting FRUSTRATIVE  
 'Karon went hunting (but didn't kill anything).'

The re-finitization of nominalizing-subordinating structures that ensues from the process of rising to independent-clause status (insubordination in Evans 2007, de-subordination in Givón 2015:661 *ff.*) is a well-documented fact in the diachrony of clause syntax, as is its effect on tense-aspect renovation. A thorough scrutiny of the issue in Arawakan, along with further references, can be found in Durand (to appear). Also, particularly akin to the Katukina-Kanamari shift from nominalizer *-nin* to durative *-nin* are various tense-aspect markers as reflexes of erstwhile nominalizers in Cariban (Gildea 1998:242 *ff.*; 2008) and Ute (Givón 2015:685).

### 6.2 SHAPING INDEPENDENT-CLAUSE ALIGNMENTS: THE ERGATIVITY BIAS

To begin with, I posit a fictitious scenario "accounting" for the syntactic evolution of a fictitious language. At a given stage, a <sup>II</sup>F-predicate *employ*(x, y) occurs in A-predicate position as (34a), and so does a <sup>I</sup>F-predicate *run*(x), (34b). At a different stage of evolution — obviating, for the sake of simplicity, the subordination / de-subordination pathway alluded to above —, the same <sup>II</sup>F-predicate eventually comes to be expressed as (35a), or perhaps as (35b). (The brackets stand for presence / absence of a copula in non-verbal predicates.)

- (34)(a) *Mary employs John* (b) *Elmer runs*

(35)(a) *Mary (is) John's employ-er*

(b) *John (is) Mary's employ-ee*

Let us further assume that at a subsequent time in diachrony, (35a) or its alternative option (35b) undergoes, as an independent clause, a process of re-finitization whereby *employ-* captures anew some of the typical properties of verbs. In the process which leads *employ* back to verbhood, *-er* and *-ee* concomitantly endure a gradual loss of their nominalizing function and a growing aptitude to serve other purposes (Gildea 1998:35-37 and 153-160; Givón 2015:685), commonly TAM finiteness-coding (see Gildea 1998:242 for the tense-aspect semantics that can be expected from each kind of oriented nominalization). Now, in such an event the genitive in *John's* and *Mary's* ends up as the case-marking of some verb's argument. And this is where the innovative alignment pattern emerges: the genitive of the agent-oriented nominalization, (35a), is now an accusative case-marker (prototypically patient), whereas the genitive of the patient-oriented nominalization, (35b), re-surfaces as an ergative case-marker (prototypically agent).

Ute (Givón 2011:99-101; 2015:669, 674) and 18th century Achagua (Durand, to appear) are cases in point as to a (35a)-like precursor for the new accusative-nominative pattern, whereas Eskimo (Mahieu 2004), Mayan (Sasse 1991) and Lak (Gilles Authier p.c. 2021; Friedman 2021) are, in turn, representative of the ergative-absolutive outcome of an erstwhile (35b)-like structure. For this latter case, Haude & Zúñiga (2016) speak of an "ergative bias".

### 6.2.1 ESKIMO

A closer look at Eskimo should shed some light on the role played by nominalization in the emergence of independent-clause structures. For this purpose I will mainly rely on Fortescue (1995, 2000) and Mahieu (2004, 2009). (Data from Nunavimmiutitut, Arctic Quebec, Mahieu 2004).

In noun phrases, the "possessor" role surfaces as a personal suffix (in fact a portmanteau morpheme standing for both participants, possessor and possessum), (36a). A noun phrase may further describe / specify the possessor's referent. The latter is then governed by the head noun, bears a genitive suffix, and occurs in internal position, (36b). A patient-oriented nominalization, (36c), is structurally identical to the possessive noun phrase (36b) regarding constituency, case-marking and indexation on the head. The pattern in (36c) occurs as the predicate of subordinate and, in a presumably ensuing point in time, independent<sup>II</sup> clauses, (37) (I replace Mahieu's gloss GENITIVE by ERGATIVE, assuming that the example features a finite-clause pattern).

(36)(a) *qaja(q)-nga*  
kayak-POSSESSIVE  
'his kayak'

(b) [ *Piita-up* ] *qaja(q)-nga*  
Peter-GENITIVE kayak-POSSESSIVE  
'Peter's kayak (lit. Peter's his-kayak)'

(c) [ *Piita-up* ] *sana-sima-ja(q)-nga*  
Peter-GENITIVE make-PERFECTIVE-PASSIVE NOMINALIZER-POSSESSIVE  
'(the) one made by Peter (more lit. (the) Peter's made-one)'

(37) [ *qajaq* ] [ *Piita-up* ] *sana-sima-ja(q)-nga*  
kayak Peter-ERGATIVE make-PERFECTIVE-PASSIVE NOMINALIZER-POSSESSIVE

' Peter made a kayak. '

In this variety of Eskimo, "topicalizing" the agent argument is carried out through left-dislocation, (38), which is tantamount to extracting it out of the predicate phrase.

- (38) [Piita-up]            [qajaq]    [sana-sima-ja(q)-nga]  
       Peter-ERGATIVE   kayak    make-PERFECTIVE-PASSIVE NOMINALIZER-POSSESSIVE  
       ' Peter made a kayak. '

And central to our present purposes (9.2.1 THE STRUCTURE OF PHRASES) is the functional competition that emerges between (37) and (38). Statistical evidence seems to show that the innovative pattern with initial-external agent argument, (38), is having the upper hand. In the more conservative Central Alaska Yupik, both orders / constituencies are yet on a par regarding frequency.

## 6.2.2 MOVIMA AND TAGALOG

Languages that afford their speakers different types of *grammatical* nominalizations — in my usage: rule-based, semantically compositional — should tend to eventually use both (35a) and (35b) to give birth to *two* independent-clause alignments that in no sense lend themselves to be viewed as one deriving synchronically from the other.<sup>4</sup>

This is what Movima, an isolate from Bolivia, seems to have achieved (Haude 2010a, 2012). This language displays minimal syntactic contrast between nouns and verbs.<sup>11</sup> Noun phrases and <sup>11</sup>verb phrases — head-initial — pattern identically: the head governs an internal argument. See the latter as pronominally indexed in both phrases:

- (39) *naye-~~le~~-na=us<sub>1</sub>*                      *kinos*                      *alwaj-a=us<sub>1</sub>*  
       marry-COPARTICIPANT-            ARTICLE FEMENINE    spouse=LINKING VOWEL  
       *na=3* MASCULINE ABSENTIAL    ABSENTIAL            *=3* MASCULINE ABSENTIAL  
       ' He<sub>1</sub> married his<sub>1</sub> wife. '

MOVIMA (Haude 2009a)

When this argument surfaces as a noun phrase, the constituency pattern [<sub>HEAD</sub> [...]NP]<sub>NP</sub> as in (40) and [<sub>HEAD</sub> [...]NP]<sub>VP</sub> as in (41) obtains. An article precedes the internal noun phrase. This is Movima's instance of head attraction announced in (30) above: the article cliticizes to the left-positioned phrase head. (For the sake of clarity in representing grammatical structure, and ease of comparison between languages, I will use a notation inspired in Katukina-Kanamari's *-na=*. The author uses --.)

- (40) [*as*                      *tobet-a*                      [=as                      *koʔ*]<sub>NP</sub>]<sub>NP</sub>  
       ARTICLE NEUTER    skin-LINKING VOWEL    =ARTICLE NEUTER    tree  
       'the bark of the tree'

MOVIMA (Haude 2006:253)

- (41) [*vel-na*                      [=ʔnes                      *senyo:ra*]<sub>NP</sub>]<sub>VP</sub>    [*kos*                      *asna*]<sub>NP</sub>

<sup>4</sup> This section expands — and sustains with some empirical data — a surmise put forth in Queixalós (2013:69 ff.) as to the parallels that can be drawn between Movima and Tagalog. My attention toward Movima was first attracted by inspiring exchanges of views with Katharina Haude during the years 2008-2009, a period of intense involvement in ergativity issues on my part.

LookAfter-*na* ARTICLEFEMENINE lady ARTICLENEUTER MyHome  
ABSENTIAL

'The lady looks after my house.'

MOVIMA (Haude 2009a; internal brackets and phrase labels mine.)

The external argument noun phrase follows the verb phrase. Since no left phrase-head is available, its article does not cliticize.

<sup>I</sup>Verbs take an external argument (as far as one can see, no difference exists between unergatives and unaccusatives in this respect). <sup>II</sup>Verbs select a suffix out of a set of two — *-na* / *-kay*, (41)-(42) — that orients the predicate towards the patient / agent external argument respectively. As Haude (2010b) states, even though the internal argument is formally identical to the expression of a "possessor", and "clauses with verbal predicates are analogous to equational clauses with predicate nominals", the lexical heads of (41)-(42) predicates are, on account of morphological and syntactic evidence — and in spite of the "subtle" character of such evidence —, synchronically verbs (also Haude 2008). Crucially, several formal properties single out the external argument as the privileged one on syntactic grounds, and the patient-oriented (41) is recognized as ergatively structured and pragmatically unmarked if compared to the accusative, agent-oriented, pattern (42). In text counts the former amounts to 90% of all transitive clauses with only third-person arguments (Haude & Zúñiga 2016).

(42) [*bijawni-kay-a* [=is *majniwa=yti*]<sub>NP</sub>]<sub>VP</sub> [*os baytim-wa:nas*]<sub>NP</sub>  
raise-*kay*- ARTICLEPLURAL ChildOf ARTICLE field-ABSTRACT  
LINKINGVOWEL NEUTER  
PLURAL

'The field raised our children.'

MOVIMA (Haude 2006:278; example slightly simplified; for the sake of the discussion I defer the use of the author's glosses for *-na* and *-kay*.)

I will now take a detour outside the American continent so as to, hopefully, improve our understanding of the Movima case. Philippine-type languages, especially Tagalog that I will use for illustration, have also created more than one oriented nominalization (Starosta *et al.* 1982; Himmelmann 1991). Their synchronic verbal reflexes select either the patient, (43), the agent, (44), or one among several other roles semantically peripheral, (45). A proclitic *ang=* flags the noun phrase picked out as the verb-orientation's target. The latter gets its semantic role from the reflexes — quite intricate synchronically — of the former nominalizing morphology. The verb phrase and the noun phrase — both head-initial — are isomorphic as for constituency and dependency marking, (46) (see *Excursus* below for a justification of the gloss REFERRER).

(43) [*binili* [*ng=lalake*]<sub>NP</sub>]<sub>VP</sub> **ang=***isda* *sa=tindahan*  
BuyPERFECTIVE GENITIVE=man REFERRER=fish DATIVE/LOCATIVE=store  
PATIENTORIENTED

'The man bought the fish at the store.'

(44) *bumili* **ang=***lalake* *ng=isda* *sa=tindahan*  
BuyPERFECTIVE REFERRER=man GENITIVE=fish DATIVE/LOCATIVE=store  
AGENTORIENTED

'The man bought fish at the store.'

(45) [*binilhan* [*ng=lalake*]<sub>NP</sub>]<sub>VP</sub> *ng=isda* **ang=***tindahan*

- BuyPERFECTIVE                      GENITIVE=man      GENITIVE=fish      REFERRER=store  
 DATIVE/LOCATIVEORIENTED  
 ' The man bought fish at the store. '
- (46) [pinutol                      [ng=magsasaka]<sub>NP</sub> ]<sub>VP</sub>      ang=[sungay      [ng=kalabaw]<sub>NP</sub> ]<sub>NP</sub>  
 CutPERFECTIVE                      GENITIVE=farmer                      REFERRER=horn                      GENITIVE=buffalo  
 PATIENTORIENTED  
 ' The farmer cut the buffalo's horn. '
- TAGALOG (Kroeger 1993:12, 32; free translations as in the original; tentative brackets and phrase labels mine.)

Worthy of note is the definite translation for 'fish' in (43), a fact to be correlated with several authors' opinion as to the basic status of the patient-oriented — ergative clause — pattern in these languages (Ceña 1977; Himmelmann 1991; De Guzmán 1992; Kroeger 1993:56; Aldridge 2012). Such basicness I take as a clue to the *diachronic* precedence of the patient-oriented pattern ("basicness" of the *ergative pattern*; this would not hold for whichever basic pattern). I will supply below some substantiation of this conjecture. Up to this point, the parallelism between Movima and Tagalog regarding constituency, predicate orientation, and noun-verb isomorphism should be clear. Another, less direct, parallelism rests on the primacy of agent- and patient-oriented nominalizations over the rest of participant nominalizations when it comes to bring about independent non-derived clause patterns. Just as Movima only features these two clause types, many Austronesian languages spoken in Malaysia and western Indonesia have reduced the Tagalog-like set of clause patterns to two: agent-oriented and patient-oriented (Adelaar 2005).

*Excursus.* Having anticipatively in mind the last section of this article, two remarks on Tagalog data might prove useful.

First, the brackets in examples are, as said, my contribution (only partially in Movima's examples). One may face a difficult task in tackling constituency matters while obviating both brackets and trees. I allow myself such an intrusion into someone else's data for obvious didactic reasons. Actually, the authors give in the text some clues to such a syntagmatic organization of constituents (*e.g.* Haude's "internal cliticization"; Kroeger 1993:111 on word order patterns; notwithstanding, the latter author — based on the lack of phrase clitics — denies the existence of complex-phrase level constituents, pp. 98-99, 229; in contrast, Manning & Sag 1999 provide evidence in favor of strong constituency in some Western Austronesian languages). All in all Tagalog is, reputedly, fairly free as to the mutual order of noun phrases (Schachter & Otnes 1972:83), with pragmatic, stylistic and role-semantics factors interfering. One instance of the latter ingredient is agent-first, (44), provided the co-present participants disallow ambiguity as to their involvement in the manner-of-existing described.

Second, in regard to analysis, *ang* has been, as a noun phrase marker, diversely interpreted and labelled, the commonest terms being "topic marker" (Schachter 1977) and "nominative case-marker" (Kaufman 1974; Kroeger 1993:13). An alternative view is that of Lemaréchal (1989). Nouns and verbs are, first and above all, predicates (in my *F-predicate* sense) — a logic-based version of the more common notion of precategoriality (Foley 1998). Cross-linguistically, noticeable exceptions to the F-predicate nature of lexical items are deixis-based expressions and, depending on the language, individual proper nouns. Referring is thus a capacity that has to be explicitly implemented as much for common nouns as for verbs. This task is carried out by some dedicated morphological tool — the equivalent of the iota operator in propositional logic — that for this type of languages is often (and, in my view,



misguidedly) regarded as an article. Significantly in Teko (Emerillon, Tupi-Guarani) a suffix with — I contend — identical function can arguably be traced back to a demonstrative (Rose 2003:118-123). (These ideas should sound familiar to those espousing Abney's 1987 proposal for a determiner phrase. Thus, *ang*-like elements, particularly those which capture semantic or pragmatic features such as number, gender and definiteness distinctions, *i.e.* so-called "articles" in Nahuatl — Launey 1994 — and Movima, head a constituent encompassing the noun phrase.) In Queixalós (2006) the label "referrer" ("référentiant" in Queixalós 2001) is put forward — *cf.* the glosses in (43)-(46) — for the marker that *enables* a nominal / verbal F-predicate to refer and, *hence*, to occur as argument. This explains its pervasiveness in clauses. The next two examples will show that Tagalog *ang*= neither marks any pragmatic function such as topic nor any case identity such as nominative.

(47) displays the "topicalization" — Kroeger's term — of the possessor in (46), 'buffalo'. The verb 'cut' retains the patient orientation having 'horn' as its target, *ang*= marked in (47) as it is in (46). (I have no hint as to the different spelling for 'cut'.) We are left, with (47) *ang*=horn and *ang*=buffalo, with two nominatives, or two topics, or a nominative and a topic mutually isomorphic. Obviously 'horn' is a verb's argument whereas 'buffalo' is not. According to my (*i.e.* Lemaréchal's) approach, both can share the same marker simply because *ang*= belongs not to the grammatical domain [verb / argument] but to the logical domain [predication / reference].

Now, we would expect that the function of allowing a noun to refer<sup>5</sup> should naturally be put to work in more than one place in the clause. (47) shows that. But how about the other occurrences of nouns in positions different from the orientation target *ang*=horn and the "topicalized" *ang*=buffalo in (47)? Actually, such occurrences are there in front of our eyes in every single Tagalog example adduced so far. Let us have a look at (48). The orientation target is consistently marked by the referrer, *ang*=letter. The genitive case surfaces not as *ng*=, displayed by (47) on *ng=magsasaka*, but as *ni*=. Lemaréchal's (1989:23) approach entails that there are no two genitive markers but one, *ni*=. And accordingly, *ng*= is nothing but the artifact that results from using as an analytical unit the conventional writing for phonological /*nang*/, which realizes the bimorphemic sequence {*ni-ang*=}. Thus, *ni-ang=magsasaka*, GENITIVE-REFERRER=farmer, *ni-Juan*, REFERRER GENITIVE-Juan, but \**ni-ang=Juan* owing to, as said, the non-F-predicative nature of proper nouns, which dispenses them from being implemented as able to refer.

The orientation-target noun phrase is indeed, as Kroeger says, a nominative, but as far as case is concerned, an *unmarked* one. More precisely: in the agent-oriented pattern, (44), the *ang*= noun phrase is an unmarked nominative, while in the patient-oriented pattern, (48), it is an unmarked absolutive. Accordingly, a more suitable representation for an example like (46) should be (49). Notice that I remain non-committal as to grammatical relations. For Kroeger (1993:21) the *ang*= noun phrase is the subject, much in the vein of Schachter (1977). Shibatani (2021) posits a split between *subject* and *topic* as distinct grammatical relations.

- (47) *ang=kalabaw*      *pinutul*      *ng=magsasaka*      *ang=sungay*  
 REFERRER=buffalo      CutPERFECTIVE      GENITIVE=farmer      REFERRER=horn  
                                          PATIENTORIENTED  
 'The farmer cut the buffalo's horn.'

- (48) *sinulat*      *ni=Juan*      *ang=liham*  
 WritePERFECTIVEPATIENTORIENTED      GENITIVE=Juan      REFERRER=letter

<sup>5</sup> I eschew the time-honoured "referring expression" label in order to avoid any ambiguity between an element being referential and an element being — what the referrer accomplishes — *able to refer*.

' Juan wrote the letter. '

- (49) [pinutol                                           [**ni-ang=magsasaka**]<sub>NP</sub> ]<sub>VP</sub>...  
CutPERFECTIVEPATIENTORIENTED GENITIVE-REFERRER=farmer  
'The farmer cut...'

... **ang**=[*sungay*                    **[ni-ang=kababaw]<sub>NP</sub>**]<sub>NP</sub>  
                                                  GENITIVE-REFERRER=**buffalo**  
 REFERRER=**horn**  
 '...the buffalo's horn. '

TAGALOG (Kroeger 1993:32, 111)

*End of excursus*

Up to this point, we have seen that one single oriented nominalization — patient — gave origin to contemporary <sup>ii</sup>clauses in Katukina-Kanamari and Eskimo, while two oriented nominalizations — patient and agent — did so in Movima. Tagalog went noticeably beyond, with its nominalizations encompassing other semantic participants beyond patient and agent.

Let us now get back to Movima. By Haude's account, *-na/-kay* — cf. (41)-(42) — stand for direct / inverse markers. Indeed, the direct / inverse analysis rests on a number of examples that clearly reveal the extent to which animacy (and in some way topicality) hierarchies prescribe which participant the verb has to be oriented to. At the functional level, however, the way these hierarchies map on grammatical structure starkly runs counter to the common and well documented direct / inverse systems elsewhere. As the author puts it (2009b; brackets and italics mine): "Transitive clauses are organized according to the relative position of the nominal referents on a referential hierarchy. *The lower-ranking participant is the privileged syntactic argument*, whereas *the higher-ranking participant is the one that is encoded like a possessor*." Also: "[...] contrary to what is stated for other inverse systems, the inverse [in Movima] would have an A-promoting and not an O-promoting function" (hence the present author's active translation of (42), instead of Haude's "Our children were raised by the field"). Said in a nutshell:

- (50)  $[\alpha \text{ animacy hierarchy}] \rightarrow [-\alpha \text{ grammatical hierarchy}]$

In an attempt at making sense of such a counter-intuitive grammatical architecture, I would now like to submit a sketchy and mostly tentative internal reconstruction along the lines that follow.<sup>6</sup>

1. Examples (41)-(42) above are the verbal finite-clause reflexes of two former oriented nominalizations, totally parallel to (35b) and (35a) respectively:

<sup>6</sup> I am grateful to a reviewer for pointing out that "the key elements of this reconstruction were already proposed and defended in some detail in Gildea & Haude's (2011) SLE presentation, from which Gildea & Zúñiga (2016: 510-519) extract both examples and the primary argumentation in published form." I am at present unacquainted with Gildea & Haude (2011) other than through Gildea & Haude (2021), which resumes co-authored (2011). As for Gildea & Zúñiga (2016), such "key elements" appear to consist in posing the oriented nominalizers as the diachronic precursors of the direction suffixes. Now, the reason why the patient-oriented nominalization is syntactically basic, and how this fact impinges upon the singular mapping between functional and formal hierarchies — the crux of the matter in accounting for the Movima case — does not seem to have been a real issue in either contribution, contrary to Queixalós's (2013: 69 *ff.*).

as patient-oriented in (41) ' My house (is) [the lady's looked-after one]. '

as agent-oriented in (42) ' The field (is) [our children's raiser]. '

2. Prior to (42), the patient-oriented nominalization (41) might have emerged *in-situ* as an independent clause for the mere purpose of *downplaying the agent participant*. This is a conceivable implementation in grammar of pragmatic-communicative strategies related to the tension existing between the cognitive primacy of the agent role on the one side, and the existence, in certain linguistic communities, of some sort of a culturally-molded association like volition-responsibility(-blame) on the other side (Givón 1981; Duranti 1994; Moyse-Faurie 2003, 2004). Here, attention must be paid to the oft-invoked "cleft" — (contrastive) focus in my terms — as a trigger for patient-oriented nominalizations, a motivation seemingly far less weighty for speakers than some kind of passivizing intention. To wit: text counts in English show, first, that "both subject and object cleft structures are exceedingly rare", object clefts being even rarer than subject clefts, while passives occur more frequently than subject / object clefts taken together (Roland *et al.* 2007), and, second, that the average of agentless passives roughly amounts to eighty per cent (Givón 1979:58; Shibatani 1985).

This conjecture suggests in itself an alternative to de-subordination as a diachronic mechanism able to generate new clause patterns out of nominalization (Queixalós 2013:40 *ff.*). An *in-situ* nominalization thus occurs in predicate position of independent clauses, with or without a copula. DeLancey on Tibeto-Burman: "The copula may be dropped over time [or reanalyzed as TAM morphology], and sometimes nominalizations are simply used as finite clauses". Such an *in-situ* nominalization is attested in Xavante for agent-downplaying and could well be an option as the precursor to ergative finite-clauses: an independent-clause <sup>II</sup>verb with a second person as agent allows for an honorific form consisting of either a patient-oriented or a non-oriented nominalization (Estevam 2011:233).

- (51) (a) [the fish] (is) [your cooked one]  
 (b) (there is) [your cooking [of the fish] ]

That said, nothing critical for the present discussion follows from adopting one or the other diachronic account, that is, either [nominalization-subordination > de-subordination > re-finitization], or *in-situ* [nominalization > re-finitization].

3. A key factor in my hypothesis is that, by finding itself demoted to internal argument — (41) and its twin example (51a) —, the expression of the agent hands over to that of the patient the whole array of formal properties attached to the syntactically privileged argument (Keenan 1976's behaviour and control; Givón 1994). This characterizes the Movima clause-type in (41), which only diverges from Katukina-Kanamari in that in Movima the agent noun phrase lacks any case-marking.
4. However, and due to the universal cognitive prominence of the agent role (Plank 1979:25; Foley & Van Valin 1984; Givón 2001:108, 199), such a grammatical architecture is counter-functional (see Ariel 2008 :147 for self-defeating form–function correlations in language and their diachronic fate). So the basic independent clause will, of necessity, evolve toward having the agent's expression back to the forefront by means of granting it the whole catalog of subject's syntactic properties, included, last but not least, the external position relative to the predicate phrase (see above Eskimo in 6.2 SHAPING

**INDEPENDENT-CLAUSE ALIGNMENTS**; Givón 2001:219; and the *reaccusativization pathway* in Queixalós 2013:67).

5. At this point, an admittedly unsettled assumption I would like to insist on is that some languages among those whose ergative clause stems diachronically from an erstwhile patient-oriented nominalization do things differently. They simply bestow the agent expression with the syntactically privileged status by creating an alternative nominalization, this time agent-oriented. The latter will eventually wind up as a re-finitized independent clause aligning accusatively at the morphological and syntactic levels. This would be the case of Movima and Tagalog.

The presumed time-sequentiality — first patient-nominalization then agent-nominalization — propounded in point 2. and here is not entirely speculative: the *innovative* agent-oriented nominalization entering the scene in the Bia dialect of Katukina-Kanamari (3.2 BIA KATUKINA) is clearly supportive in this respect. Of course, one does not expect any comparative evidence from Movima, but regarding Tagalog the vast array of languages genetically related to it should provide some hints as to the plausibility of the suggested relative chronology. Unfortunately, after a quick perusal of Austronesian literature it would seem that nobody has ever wondered about that surmise, not to say addressed it as an issue. For the sake of underpinning my conjecture I however provide hereafter a few clues to such possibility (contents of square brackets mine).

Starosta *et al.* (1982) readily posit patient-oriented nominalizers \*-en, \*ni-, \*-in in Proto-Austronesian, whereas agent-oriented nominalizers \*mu-, \*um appear as reconstructible albeit systematically subject to qualification: "possibly" / "probably" occur eight times in the text for these agent-oriented nominalizers; also "PAn [Proto-Austronesian] was a strongly noun-oriented language, with a high percentage of nominalization strategies. The affixes \*-en, \*ni-, \*-in [...], and [again] possibly \*mu-, -um- [sic] functioned to derive nouns from verbs".

Similarly, for Pawley & Reid (1999) "the use of verb stems plus non-Actor focus [other authors' *topic*] affixes as nouns is clearly PAn". They also find that "throughout the Austronesian family the person markers for the agents of passive verbs are the same as the genitive pronouns marking the possessors of underived nouns"; and "the nominalizations of POc [Proto-Oceanic] and many of its descendants show a contrast between actor-possessor and patient-possessor. Further, the particle marking actor-possessor is probably a reflex of PAn \*na [...] corresponding to the personal noun marker PAn \*ni 'actor/owner' "; in my rephrasing: reflexes of patient-oriented nominalizations, with the agent surfacing as genitive-marked internal argument, pervade the whole of the Austronesian family, a fair indication of ancientness.

Finally, Ross (2009) reconstructs Proto-Rukai (Formosan, Austronesian) verbal morphology as nominalizations for patient, location and instrument, leaving the agent-oriented nominalization (Table 7) as a set of three dots for, I presume, 'unknown form'.

It should be no surprise that the answer to a never-raised question be hard to come by. Though, the frequent hedging in contexts where the proto-forms for agent nominalization are expected, sharply contrasts with the positively reconstructed proto-forms for patient nominalization anytime the latter is mentioned.

Following up the assumption in 5., we may further forecast the possibility that the accusative pattern, functionally more congruous with the inherent prominence of agents, ends taking over the status of *basic* independent clause, not the case of Movima and

Tagalog ...yet, but indeed that of [East] Oceanic languages; for Pawley & Reid (1999): "the sequence of innovations leading from a Philippine-like ancestral system to the Oceanic type: [...] *subject became basically equated with actor*, and [...] *new passives* based on these active constructions *appeared* (possibly *independently*) in certain Indonesian and Oceanic languages [my italics]." Another language in case is Rukai (Formosan), which, according to Ross (2009), "has lost the undergoer voice and become an accusative language" with, again, an innovative passive construction. Outside Austronesian, Dargi (Tsezic, Eastern Caucasus), also featuring two mutually-underived "voices" (Authier 2016), might be experiencing the same reaccusativization drift.

6. As structures like (41)-(42) eventually shift from nominalized(-subordinate) status to finite(-independent) status, the formal apparatus of nominalization(-subordination) is renewed with different material (something that reminds the partial Katukina-Kanamari's reshaping of nominalization strategies that, what is more, appear as dialectally divergent). Synchronically, the current verbal status of Movima's former nominalizations has as a consequence calling into play an "article" / "determiner"<sup>7</sup> for nominalizing, much in the vein of the *referrer* introduced above for Tagalog and to be introduced below for Tupi(-Guarani). In the long run, the "direct / inverse" markers are the mere synchronic reflexes of the erstwhile nominalizers having retained their former orientating function, as the mentioned authors state.
7. Now, the ergativity bias understood as: re-finitized patient-oriented nominalizations yield ergative clause-patterns, is not *per se* a remarkable phenomenon. Many languages beyond Katukina-Kanamari, Eskimo, Tagalog and Movima have undergone this same historical process. Many other languages could as well be credited of an accusativity bias. What the ergativity bias contributes to make Movima outstanding in this context is a twofold mechanism. First, functional hierarchies interfere heavily with the remodelling of clause-patterns. Second and decisively — this is my understanding of Movima's ergativity bias —, the directionality make-up flows in a straight line from a primary clause-pattern aligned ergatively. I hence posit that the Movima's peculiar "direct / inverse" *effect* arises from three interconnected factors: A. the ergatively-aligned verb orientation inherited from a patient-oriented nominalization; B. the synchronically basic status of such alignment, resulting from its likely precedence in the diachronic line of time; and C. the inherent cognitive prominence of agents prompting an agent-oriented nominalization that leads to an accusatively-aligned verb orientation.

Thus seen, the way the ergativity bias informs the Movima grammatical patterns can be unfolded as follows. The patient-oriented clause sets in internal-argument position the expression of the agent participant (the basic pattern in Katukina-Kanamari, Tagalog, and former stages of Eskimo). Since agents are typically higher than patients in animacy hierarchies, the diachronic outcome of re-finitization will be an independent clause whose internal argument stands, regarding these hierarchies, as *mechanically* higher than the external argument. In stark illustration of the way in which the structurally distinct patterns can be conditioned by their genetic inheritance, the patient-oriented nominalization as basic forces the grammatical reworking into a mismatch between form /vs./ function that ends up as a frozen constraint upon unmarked <sup>II</sup>verb clauses. Although Haude & Zúñiga (2016) acknowledge nominalizing morphology as the diachronic precursor of "direct / inverse" marking in Movima, they do not engage in doing justice to what would be the ultimate factor behind the functionally-inverted "inverse" (in the same vein, see footnote 5 above).

<sup>7</sup> Thanks are due to Spike Gildea for bringing to my attention this capacity of the "determiner".

Voice might be a suitable candidate for accommodating the Movima facts. It has been put forward for Tagalog by several authors with different theoretical backgrounds (and as early as Blake, F. 1916 and Bloomfield 1917). Movima might lend itself to such an approach, as (52) shows.

- (52)(a) 

<i>[sal-na</i>	<i>[=as</i>	<i>pa:ko]</i> <sub>NP</sub>	<i>]VP</i>	<i>[kus</i>	<i>ona:cho]</i> <sub>NP</sub>
search-DIRECT	ARTICLE	NEUTER	dog	ARTICLE	MyGrandchild
				MASCULINE	
				ABSENTIAL	

' The dog looks for my grandson. '

- (b) 

<i>[sal-kay-a</i>	<i>[=kus</i>	<i>ona:cho]</i> <sub>NP</sub>	<i>]VP</i>	<i>[as</i>	<i>pa:ko]</i> <sub>NP</sub>
search-INVERSE-	ARTICLE	MyGrandchild	ARTICLE	NEUTER	dog
LINKINGVOWEL	MASCULINE				
	ABSENTIAL				

' The dog looks for my grandson. '

(Movima, Haude 2009a; internal brackets and phrase labels mine.)

In Haude's words, "[...] two speakers [were] offered independently [(52a)], where the lower ranking participant is encoded as [internal] and the verb is marked as direct. The inverse construction [(52b)] was accepted when I suggested it, but the speakers did explicitly say that it was no better than [(52a)]" (contents of square brackets mine). Nothing is reported as to the relative frequency of (52a, b) in natural production. The author considers — following Judy (1965) — the likelihood of the voice option, but rejects it based on the oddity of an inverse voice upgrading the centrality of the agent participant. (Note that the same oddity argument could be used to rule out the "direct / inverse" analysis.) There are, to be sure, several reasons to discard voice as the main function of the "direct / inverse" contrast in Movima. First, a secondary voice is expected to derive from a basic one. In this language no formal fact whatsoever registers such a *synchronic derivation* — neither in Tagalog. Second, the basic would-be voice is ergatively oriented, and a genuine antipassive exists as derived from it (Haude 2012) — contrary to Tagalog where the only candidate for the antipassive status would be the agent-oriented clause (Aldridge 2012). And third, voice commonly encodes strong pragmatic motivations (Givón 1994), but here the semantic motivation, *i.e.* animacy hierarchies ("rigidity" in Haude & Zúñiga 2016) is paramount, even though Haude (2008) mentions topicality ("flexibility" in Haude & Zúñiga 2016) among the hierarchies conditioning the switch between "direct" and "inverse". Now, and pace the precise meaning of *topicality* throughout the relevant literature, in Movima the internal argument is higher in topicality while in Tagalog and generally in Western Malayo-Polynesian languages it is the syntactically privileged argument — external in my terms — that is topical / definite / specific (Shibatani 2021), more in line with the cross-linguistically common mappings between function and syntax.

We have so far overviewed four languages — Katukina-Kanamari, Eskimo, Movima, and Tagalog — that, with perhaps some restrictions, make a good case for

1. the synchronic alignment of independent clauses as stemming from the re-finitization of nominalized clauses;
2. the formal parallelism between <sup>II</sup>F-predicates, nouns and verbs;

3. the existence of internal arguments identically marked and positioned in noun and verb phrases.

I have further suggested that in Movima and Tagalog the agent-oriented nominalization might have emerged, in diachrony times, later than the patient-oriented one as a functionally-based adjustment, in a syntactically ergative environment, toward reinstating the agent participant to the privileged-argument condition.

To round off this section by venturing a further conjecture on Movima and Tagalog, I would like to put forward *diathesis* as an overall term for what, in the foregoing discussion on independent clause patterns, I have been calling *orientation*, a term better suited to nominalization processes (and potentially mistakable, in other contexts, for spatial coordinates). "Diathesis" originates in Tesnière's work (1959:242) for the orientation of verbs toward a given argument which, that being so, qualifies as *subject*. Though, the examples the author used for illustration purposes were French active / passive clauses, a slight bias that led several readers to misguidedly equate *diathesis* and *voice* (e.g. Martinet 1965:221, and less neatly Witzlack-Makarevich 2010:197 ff.) and turn the term useless. Voice is a sub-type of diathesis: a secondary orientation set down on the verb's primary orientation by some overt morphosyntactic device. In the absence of such synchronic derivation, we do not have voice but merely (non-derivational) diathesis. This is the meaning of "symmetrical voice" (Foley 1998; Himmelmann 2005) as used for Philippine-type languages. Thus, in Standard Indonesian and Mori Bawah (Western Austronesian) *assymetrical* voices (Himmelmann's term) passive and antipassive are respectively superimposed, as synchronically-derived alternations, on symmetrical voices (diatheses). (The latter language even features a passive and an antipassive altogether, Mead 2005.) Similarly, Haudricourt (1979) carefully tells diathesis and voice apart on the basis that, in Paiwan (Formosan), (assymetrical) voice needs dedicated derivational morphology and that, on top of it, such morphology can co-occur with different diatheses.

Much in the vein of Lazard (1997) and the Typology Group of Saint Peterburg's work (e.g. Kulikov 2010), diathesis is here seen as the layer in the verb's lexical argument-structure that, beyond the stipulations on number of core participants (valence), their semantic properties (roles, so-called "phi-features") and their coding, singles out a given argument as prominent on syntactic grounds. Diathesis thus defined is by all means a pivotal factor in splitting the lexicon of verbs into formal classes. In this perspective, lability (A-lability as well as P-lability) is an instance of a class of verbs allowing for more than one diathesis, as are so-called locative / instrumental "voices" of, e.g., English *This **bed** sleeps three persons / **Money** can't buy me love.*

Movima and Katukina-Kanamari do have a secondary voice, the antipassive, brought into being by overt derivational morphology. Movima and Tagalog have more than one diathesis (symmetrical voices). In the latter two languages, the lack of any formal device giving rise to a synchronically-derived pattern stems from the fact that what stands as the diachronic source of the diathesis morphology is nothing but the erstwhile nominalizing morphology. Katukina-Kanamari also has more than one (insisting: non-derivational) diathesis. It, however, obeys different factors, as we will see hereafter.

## 7 THE SYNCHRONIC PATTERNS OF KATUKINA-KANAMARI INDEPENDENT CLAUSE

In this section I will survey the essentials of clause structure in order to display the overall picture which the particular issues so far addressed fit into. As said in section 2 **F-PREDICATES AND VALENCE** above, the external argument of <sup>II</sup>verbs — semantically patient — and the unique argument of <sup>I</sup>verbs align for constituency, lack of overt case-marking,

and basic post-verbal position, (1b)-(2b) renumbered (53a-b) respectively. Additionally this argument allows for *pro*: it may saturate a lexically required F-predicate's place while remaining syntactically unexpressed. The other argument — semantically agent — is internal to the verb phrase and mandatorily overt, either indexed to the verb, (53a), or as a case-marked noun phrase in fixed pre-verbal position, (54).

- (53)(a) [a-hak] [adu]  
           3 SINGULAR-spear 1 SINGULAR  
           ' He speared me. '
- (b) [i:kao] [adu]  
           cry 1 SINGULAR  
           ' I cried. '
- (54) [ [Tamakori-na=] buhuk] [tukuna]  
       Tamakori-GRAMCASEMARK= make Indians  
       ' Tamakori created the people. '

Such an alignment is thoroughly mirrored in syntax. Queixalós (2017b) supplies full-fledged formal evidence for the external argument's subjecthood (*e.g.* privileged access to coordination, focus, question, relativization, and control of coreference). As an instance of this, I will mention just one type of syntactic process, movement. The external argument noun phrase can be fronted with no other consequence on the clause structure, compare (56)-(54). (I will only provide examples for <sup>II</sup>verbs.) The internal argument noun phrase undergoes extraction from the verb phrase in a variety of syntactically constrained circumstances, all granting it access to the external argument's formal privileges. Thus, a de-transitivizing voice deprives the patient participant of its argument status — the antipassive mentioned in 6.2 SHAPING INDEPENDENT-CLAUSE ALIGNMENTS *in fine* — while promoting the agent participant to external argument along with crediting it with the differential properties attached to this position, such as *pro* in the following example (compare Ø in (27b)).

- (55) nuk<sub>1</sub> waokdyiwaokdyan-nin wa-pu Ø<sub>1</sub>  
       group ArrivePLURAL-NOMINALIZER ANTIPASSIVE-eat 3  
       ' As the mob<sub>1</sub> arrived (home), they<sub>1</sub> ate. '

Besides the syntactic motivation triggering the de-transitivization process just mentioned — access to subjecthood —, pragmatic forces can prompt the internal noun phrase (agent) to move off the verb phrase. A left-dislocated noun phrase loses its case-marker and, abiding by the constraint on overt internal-argument, leaves behind the person prefix on the verb, compare (57)-(53a). The noun phrase is, hence, no longer an argument but a coreferential adjunct.<sup>8</sup>

- (56) [tukuna] [ [Tamakori-na=] buhuk]  
       Indians Tamakori-GRAMCASEMARK= make  
       ' Tamakori created the people. '
- (57) [mapiri]<sub>1</sub> [ [a<sub>1</sub>-binik] [wa:pa]<sub>2</sub> ]  
       anaconda 3SINGULAR-swallow dog  
       ' (As for) the anaconda<sub>1</sub>, it<sub>1</sub> swallowed up the dog<sub>2</sub>. '

<sup>8</sup> In languages featuring argument-indexing systems *à la* Jelinek (1984) and Mithun (2003), two kinds of adjuncts ought to be distinguished functionally: those that contribute circumstantial information to the manner-of-existing denoted by the predicate, *adverbial adjuncts*, vs. those that help narrowing down the reference of the argument indexes, *coreferential adjuncts* (Queixalós 2016).



An alternative diathesis comes out in Katukina-Kanamari as what might be also called *split transitivity* (or, a bit less appropriately, "split ergativity"). Generic and non-count patient participants trigger an accusatively aligned pattern. These participants surface as an internal argument unmarked for case, whereas the expression of the agent participant takes over the external argument's properties described above, (58b). Speech-act participants and proper nouns do not occur in internal argument position, (59b). The verb lacks any argument indexing.

- (58)(a) [yo-wu]                      [oba-kon]                      (b) [ [oba]              wu]              [adu]  
           1 SINGULAR-want      tobacco-CYLINDER              tobacco      want      1 SINGULAR  
           ' I want a cigarette. '                      ' I want tobacco. '

- (59)(a) [ [mapiri-na=]                      duni]              [adu]  
           anaconda-GRAMCASEMARK=      catch      1 SINGULAR  
           ' The anaconda caught me. '

- (b) \*[ [adu] duni] [mapiri]

Although semantically marked (generic, non-count), the accusative pattern cannot be seen as *derived* from that seen in examples (56)-(57). Rather, in the diachronic scenario sketched above for Katukina-Kanamari (3 NOMINALIZATION and *ff.*), it might fit as the independent verbal <sup>II</sup> clause preceding the innovative ergative clause built on a patient-oriented nominalization. Bearing in mind that low topicality-referentiality makes an argument a poor candidate to subjecthood, one understands why the accusative clause-type survived, no matter how much residually, after such a radical turn as the advent of syntactic ergativity.

The basic premiss behind the scenario accounting for the emergence of the ergative pattern — *-nin* implementing a patient-oriented <sup>II</sup> nominalization — suggests there to be no synchronic durative *-nin* occurring either in the accusative pattern or on <sup>I</sup> verbs, since no patient-oriented nominalization might have been involved. (60)-(61) provide evidence to the contrary.

- (60) Tamakori    na              [ [a-tukuna]              makoniok-nin]  
           Tamakori    FOCUS    INDEFINITE-people    edify-DURATIVE  
           ' Tamakori was edifying the people. '

- (61) wa:pa    kitan-nin  
           dog      sleep-DURATIVE  
           ' The dog is sleeping. '

As a matter of fact, these data may be viewed as evidence for the reanalysis of nominalizer *-nin* as durative: the innovative function — aspect — extends (Harris & Campbell 1994:51; Gildea 1998:38) to environments where its erstwhile function — nominalizer — would not have applied.

Non-verbal A-predicates are nominal and adverbial. Nominal predicates can be monovalent (<sup>I</sup>nouns), (62a), or divalent (<sup>II</sup>nouns), (62b). Similarly, adverbial predicates can be either monovalent (adverbs), (63a), or divalent (postpositions), (63b). Besides a non-

obligatory existential copula *an*, (64), no copulative element occurs in non verbal predications.

- (62)(a) *tukuna Dyano*  
 Indian Dyano  
 'Dyano is an Indian. '
- (b) *Aro-na= tyo Ino*  
 Aro-GRAMCASEMARK= daughter Ino  
 'Ino is Aro's daughter. '
- (63)(a) *kodo kamudya*  
 UpThere MonkeySp.  
 'The monkey *Sp.* is up there. '
- (b) *itivan-na= katu Owi*  
 ThatOne-GRAMCASEMARK= COMITATIVE Owi  
 'Owi is with that one. '
- (64) *koni ki:dak an*  
 story OldThing COPULA  
 'There are old stories (going on). '

It should be noted, to conclude, that in spite of allowing nouns in A-predicate position without any copular element, and, importantly, in spite of the parallelisms — constituency, case marking — in noun and verb phrases hosting an internal argument, Katukina-Kanamari displays clear-cut formal differences between these lexical classes: 1. nouns need no implementation when occurring in argument positions, while verbs must be nominalized, and 2. <sup>II</sup>nouns do not feature different diatheses nor voices, neither do they permit movement / extraction of their internal argument.

## 8 INTERIM SUMMARY

I will here recapitulate the main results of the foregoing discussion so as to lay the background against which the next section will tackle some Tupi(-Guarani) issues. What the languages overviewed so far have showed us is that:

1. Nominalization processes — mainly those which single out a given participant as the orientation target — have far-reaching consequences on the synchronic structure of independent <sup>II</sup>clauses *via* subordination and the diachronically subsequent de-subordination, or (my *in-situ* nominalization regarding some syntactically ergative languages) by way of a diachronic shortcut aiming at downgrading, right in the independent clause, the agent participant.
2. In the process of re-finitization, the nominalizing morphology acquires other functions, commonly TAM distinctions, while new formal devices take over the nominalizing / subordinating tasks. (The reinstitution of innovative nominalizing morphology may, however, not happen overnight, as Bia Katukina shows, section 3.2.)
3. The verb phrase having inherited the structure of a noun phrase will have one of its core participants surface in the formal mould of a dependent *inside* a noun phrase — a genitive.
4. This dependent winds up as an ergative argument when the source orientation targets the patient, and as an accusative argument when the source orientation targets the agent.
5. The synchronic formal parallel between noun phrases and verb phrases may extend to adpositional phrases due to the frequent nominal origin of adpositions.
6. In a language contrasting <sup>I</sup>nouns and <sup>II</sup>nouns, the verb phrase that develops out of a nominalized <sup>II</sup>verb is isomorphic to the <sup>II</sup>noun phrase.

7. Unstressed elements bound to the dependent, internal, constituent and intermediating linearly between the latter and the phrase head, may find themselves cliticized to the head.
8. Verb-phrase internal arguments move off their original *locus* for pragmatic purposes (typically: uprating their informational import).
9. Their secondary sequential position can yield innovative constituency or word-order liable to grammaticalize as basic.
10. A language may generate, for a given lexical verb, more than one (non-derivational) diathesis out of more than one nominalization pattern.

## 9 A TAKE ON TUPI(-GUARANI) INTERNAL ARGUMENTS

The purpose of this section is to address the question of internal arguments in the Tupi family of languages. For doing so, I will not only present a brief appraisal of the existing knowledge on the topic itself, but also advance specific proposals as to the synchronic nature of the facts and the diachronic evolution that may have led to their current character.

This will call forth the need to assess whether several structural parallelisms — and asymmetries — between nouns, verbs and postpositions are due to anything other than chance. The negative answer is, of course, rooted in the F-predicate nature of such word classes and the effects of diachronic change. In all probability this will, as a bonus, set the conditions for casting some light on one lingering puzzle in Tupi morphology. Several of the hitherto reviewed phenomena will prove essential to the argument and, needless to say, the assumptions made in the present section will part company in fundamental ways from what is standardly taken for granted.

### 9.1 PROVIDING NORMALIZED DATA

For the present attempt at submitting an alternative analysis of the Tupi(-Guarani) facts I will first adduce some evidence of the parallelism between phrases in Guaja, a Tupi-Guarani language spoken in the Brazilian state of Maranhão and described by Marina Magalhães. To keep things simple and amenable to comparison, I will somehow regulate the data. Especially, constituency and morphology will be recorded — along with their attendant terminology and glosses — in the light of the analyses proposed in the preceding sections (with one exception: 2. hereafter). To wit:

1. The commonly used term "determiner" for noun phrases or personal bound forms is replaced by *argument* (often in reference to the internal argument; as for "intransitive" — unaccusative? — and transitive verbs, Cabral 2002 uses *internal argument*);
2. A pervasive "contiguity relational prefix" — an "adjacency marker" in Magalhães & Mattos (2014) — is dubbed R- in glosses as a convenient provisory label (from its more basic allomorph *r-*, that whenever possible I will pick out of a motley array of co-allomorphs; I suggest the reader take notice of it up front); it will be represented as in the original examples till we reach section 9.2.2 THE CASE FOR R-.
3. Another pervasive morpheme, the purported "nominal case" or "argumentative case" or "nuclear case" suffix *-a*, is glossed as *referrer* (see Queixalós 2001, 2006 for a

justification) since it signals the same function than Tagalog's proclitic *ang*= (section 6.2.2 MOVIMA AND TAGALOG *Excursus*): endowing an F-predicate — nouns as well as verbs — with the capacity for referring;

4. Some authors make constituency explicit while others do not. Constituency, however, may be pivotal for the issues at stake. I tentatively supply it wherever it seems relevant.
5. A morpheme often cast in terms of "non-contiguity relational prefix" is understood, along with several authors working on either Tupi-Guarani or Tupi languages (Jensen 1999; Corrêa da Silva 2002; Couchili, Maurel & Queixalós 2002; Galucio 2002; Rose 2003; Praça 2007; Copin 2012; Magalhães & Mattos 2014), as indexing a third person argument (a distributional basis for this choice will be supplied in due time below); the point of note here is Praça's quote of Apyãwa (Tapirape) school pupils describing this morpheme as "the mirror of whom/what people are or have been speaking about" (p. 38).
6. So as to make the data more transparent, morphophonologically deleted segments will be restored as (x).

The following examples, excerpted from Magalhães (2007), show <sup>II</sup>noun and postposition phrases with their internal argument surfacing as a bound pronominal form in examples (a), and as a noun phrase in examples (b) (/j/ is [y] and /y/ is [u]; constituency brackets mine).

<sup>II</sup>noun

(65)(a) [ [ha=r-imirikó-a]  
1=R-wife-REFERRER  
' my wife ' (p. 131)

(b) [ [xahú] r-awáj-a] a-xá  
peccary R-tail-REFERRER 1-see  
' I saw the peccary's tail. ' (p. 27)

postposition

(66)(a) [ [ha=]r-aké]  
1=R-ADESSIVE  
' next to me ' (p. 133)

(b) [ [pé] r-ipí]  
trail R-PERLATIVE  
' along the trail ' (p. 133)

Now turning to verbs, we face a twofold novelty if we compare with the languages overviewed above. First, <sup>II</sup>verb phrases do not parallel <sup>II</sup>noun / postposition phrases for constituency: the internal argument is there, but only under its indexed form, (67). The lexical noun phrase occurs externally, either before or after the verb (mostly before), (68). Second, unaccusative verbs display exactly this same behaviour, (69). (A person hierarchy obtains for the indexation of arguments on <sup>II</sup>verbs.)

<sup>II</sup>verb

(67) [ [ha<sub>1</sub>=]r-ixá] [a?iá]<sub>2</sub>

1=R-see                    3  
' He<sub>2</sub> saw me<sub>1</sub>. '

(p. 193)

(68)(a) [jakaré-a]<sub>1</sub>                    [pirá-(a)]<sub>2</sub>                    [u<sub>1</sub>-ʔú]  
alligator-REFERRER    fish-REFERRER    3-eat  
' The alligator<sub>1</sub> ate the fish<sub>2</sub>. '

(p. 142)

(b) [takÿn-a]<sub>1</sub>                    [u<sub>1</sub>-ʔu]                    [ha-miʔũ-a]<sub>2</sub>                    [i-kira=pa]  
toucan-REFERRER    3-eat                    3-food-REFERRER    3-BeFat-PURPOSE  
' The toucan<sub>1</sub> eats<sub>1</sub> his food<sub>2</sub> to get fat. '

(Marina Magalhães, p.c. February 2021)

<sup>I</sup>verb, unaccusative

(69)(a) [ [ha=]r-ahý ]  
1=R-BeSick  
' I am sick. '

(p. 52)

(b) [i-mymýr-a]<sub>1</sub>                    [h<sub>1</sub>-ahý]  
3-son-REFERRER    3-BeSick  
' Her son<sub>1</sub> is<sub>1</sub> sick. '

(p. 169)

(c) ... [h-ahy]                    [ha=ø-py-a]                    any  
3-BeSick    1-R-foot-REFERRER                    COORDINATION  
' ... and my feet were hurting. '

(Marina Magalhães, p.c. February 2021)

With little variation, analogous examples can be found in — among others —:

Apyãwa (Praça 2007:16, 34-36, 38, 73-76), Tocantins Assurini (Cabral 2007), Kamayura (Seki 2000) and Warázu (Pauserna, Guarasugwe; Ramirez & de França 2017), all of them Tupi-Guarani languages;

Sateré-Mawe (Silva 2010:176, 196), a language forming with Tupi-Guarani a sub-group within the Tupi family (Rodrigues & Dietrich 1997), and

Munduruku (Gomes 2007:22, 32, and 2017), a Tupi language outside the Tupi-Guarani-Aweti cluster. Gomes (2017) explicitly argues for the existence of such parallelisms between <sup>II</sup>noun, <sup>II</sup>verb, unaccusative, and postposition phrases.

## 9.2 ACCOUNTING FOR THE DATA

Here, constituency and morphology will be inspected so as to substantiate the mutual congruence of the data supplied.

### 9.2.1 THE STRUCTURE OF PHRASES

As anticipated at the outset of this section, the consistent patterns between these phrases is assumed to be a direct function of the lexical-relational structure of their heads, namely <sup>II</sup>F-predicates, which gives raise to contrasted argument coding and constituency. In particular,

one of the co-present arguments is directly case-governed by the head and occurs internally to the phrase thus generated. Moreover, the form of this argument should easily be traced back to the nominal morphology imposed by either verb nominalizations or the assumed <sup>II</sup>noun precursors of postpositions.

We saw above that this fact unifies <sup>II</sup>nouns, <sup>II</sup>verbs, and postpositions in Katukina-Kanamari (5 CASE-MARKING, (27), resumed / renumbered here for ease of reference). It does so, though partially, in the Tupi-Guarani data supplied hitherto.<sup>9</sup>

<sup>II</sup>noun, compare with (65)

- (70)(a) [ [*opatyin-na*=]                      *wadik*]  
           child-GRAMCASEMARK=            name  
           ' child's name '

<sup>II</sup>verb, compare with (67a)

- (b) [ [*Mayon-na*=]                      *ikubalama*]  
       Mayon-GRAMCASEMARK=          DigItOut  
       ' Mayon dug it out. '

postposition, compare with (66)

- (c) [ [*kotyia-na*=]                      *katu*]  
       otter-GRAMCASEMARK=          COMITATIVE  
       'with the otters'

We also know that Guaja matches this pattern uniquely with <sup>II</sup>noun and postpositional heads (9.1 PROVIDING NORMALIZED DATA). With <sup>II</sup>verbs, the pronominal exponent of the internal argument — semantically patient in Tupi(-Guarani) — also fits in here, (67), but the noun is clearly positioned outside the verb phrase, (68). A coherently tenable assumption is that, in the same vein than its correspondent in Katukina-Kanamari, the internal noun phrase migrates outside the verb phrase for pragmatic reasons, whichever these might be. Cabral (2002) points out that in the Tupi languages of Rondônia the occurrence of internal arguments as free pronouns outside the verb phrase signals emphasis, with the exception of the innovating Surui where the cognate forms occur internally. (In my opinion Surui should rather be held as the conservative language.)

An important difference between Guaja and Katukina-Kanamari arises from the fact that such migration is a synchronic process in the latter — left-dislocation *via* extraction —, but a diachronically-fixed change in the former. Subordinate clauses — more conservative than independent ones due to their lower exposure to pragmatically motivated forces — provide a significant clue to the originally internal position of this noun phrase, (71), from which the occurrence of the referrer is banned.

- (71) *jahá a-jahó kaʔá-pe*                      [ [*ni=r-ú*]    *r-aká=pa*]  
       1        1-go    jungle-LOCATIVE    2=R-father    R-LookFor-SIMULTANEITY  
       ' I went to the jungle looking for your father. '                      (p. 267)

- (72) *ha=r-imiriko-a*                      *o-inẽ*    *ha-ipa-pe...*  
       1=R-wife-REFERRER    3-stay    3-home-LOCATIVE  
       ' My wife stayed at home... '

<sup>9</sup> My first intuitions about the nature of (Tupi)-Guarani R- — needless to say prompted by the Katukina-Kanamari facts — were presented orally in Queixalós (2015).

...[ **[xahú]**      **r-aká**              *mehẽ jahá*  
          peccary    R-LookFor    when    1  
 ' ...when I went out to look for a peccary. '

(Marina Magalhães, p.c. November 2021)

Also in line with Katukina-Kanamari (see 5 CASE-MARKING), a constraint is here at work regarding the internal-argument: the latter must be realized overtly, either as a noun phrase or as a verbal index. In the interests of perspicuity, let us first resume the Katukina-Kanamari relevant examples (comparison is to be drawn between the items in bold type).

(73)(a) [ **[opatyin-na=]**                      *wadik*  
          child-GRAMCASEMARK=          name  
          'the child's name'

(b) [ **a-**                                      *wadik*  
       3SINGULAR-                          name  
       'her name'

(74)(a) [ **[kotyia-na=]**                      *katu*  
          otter-GRAMCASEMARK=          COMITATIVE  
          'with the otters'

(b) [ **a-**                                      *katu*  
       3SINGULAR-                          COMITATIVE  
       'with her'

(75)(a) [ **[Mayon-na=]**                      *ikubalama*  
          Mayon-GRAMCASEMARK=      DigOut  
          'Mayon dug (it) out. '

(b) [ **a-**                                      *hak*                      *adu*  
       3SINGULAR-                          3-spear                1SINGULAR  
       'He speared me. '

One can safely infer — and the mere observation of the data should confirm this — that to all analytical intents and purposes the Tupi-Guarani alleged "non-contiguity relational prefix" features, in nominal and postpositional phrases, the same distribution than the Katukina-Kanamari's third person *a-* — realizing the internal argument —, and undoubtedly quite different a distribution from *R-*, (76)-(77). It moreover does so in the erstwhile — today subordinate — verb phrases, (71)-(78), making up a substitution class not solely with *R-* but rather with the whole sequence [argument(=)R-]. (I only provide illustration with noun phrases, but an identical observation can be carried out with pronominal first and second person clitics, hence the equal sign in [argument(=)R-] above.) This yields clear evidence for its pronominal nature, as underscored by the Apyãwa pupils.

(76)(a) [ **[xahú]**      **r-**      *awáj-a*  
          peccary    R-    tail-REFERRER  
          'the peccary's tail '

(excerpted from (34b))

- (b) [i-                      *mymýr-a*]  
       3                      -son-REFERRER  
       'her son '                      (excerpted from (58))

- (77)(a) [ [pé]        r-        *ipĩ*]  
           trail        R-        PERLATIVE  
           'along the trail '                      (excerpted from (55))

- (b) [i-                      *pamẽ*]  
       3                      COMITATIVE  
       'with him '                      (excerpted from example 1033 p. 284)

- (78) *jahá<sub>1</sub> a<sub>1</sub>-xá [i<sub>2</sub>-ʔú] mehẽ*  
       1        1-see    3-eat    when  
       'I<sub>1</sub> saw him<sub>2</sub> eating (it) (lit. I saw when he ate it). '                      (p. 134)

Turning to unaccusative verbs, the unique-argument noun phrase also occurs externally to the verb phrase in independent clauses, (69b-c), but internally in subordinate ones, as in:

- (79) a-jaho    ta                      Santa Inês    (r-)pe                      [ [ha=Ø-mymy(r)]    r-ahy]                      *mehẽ*  
       1-go        FUTURE    Santa Inês    R-LOCATIVE        1-R-son                      R-BeSick    when  
       'I'll go to Santa Inês when my son is sick. '                      (Marina Magalhães, p.c. February 2021)

Constituency and morphology, thus, point to a clear parallel between <sup>II</sup>verbs and unaccusatives as to the realization of their "object" / single argument.

### 9.2.2 THE CASE FOR R-

We are now in a position to wrap up the discussion in terms of a coherent diachronic hypothesis for the *grammatical* nature and origin of the — at first glance — relic of something flagging an as yet rather enigmatic function. Given the tight parallel in the structure of phrases between Tupi(-Guarani) and Katukina-Kanamari, one may wonder what in the former — if any — might tally with the latter's *-na=*. We have a ready candidate: R-. In effect, as the reader will have pinpointed, Guaja R- displays the same distribution as the Katukina-Kanamari case-marker (except for unaccusative verbs, that Katukina-Kanamari discloses through other formal properties): bound to the immediately subsequent phrase-head (the purported "prefix") while remaining *grammatically* inseparable from the immediately preceding internal argument. I hence contend that Tupi(-Guarani) R- must be, in diachronic perspective, traced back to a case suffix that, due to a long and intimate cohabitation, in medial position, with the two components of a constituent [dependent + head], has migrated from left to right under the effect of head attraction.<sup>10</sup> In sum, R- is a case proclitic on governing heads. Since all other case-markers in these languages seem to be of the so-called semantic / inherent sort, I will henceforth gloss it *grammatical case-marker*. As an indication

<sup>10</sup> In also a perfect parallel to Gitksan's so-called "connector", were it not the reverse linear shift [head =CASE-noun] (Rigsby 1986:277, cited in Mithun 2008).



to a Tupi(-Guarani) leaning toward head attraction — recall that Katukina-Kanamari's *-na=* is not single of its ilk regarding that process — let us take note of the following record by Maurel (2019) concerning the referrer suffix *-a* in Teko: the grammatical sequence {*tapidz-a pope*}, house-REFERRER INESSIVE, 'in the house', was submitted to 86 speakers for segmentation into words. Almost all the interviewees returned /*tapidz#apope*/. What is more, Maurel (p.c. March 2021) reports a clear bias toward /*tapidza#pope*/ on the part of older speakers. As regards the referrer, head attraction is an ongoing process in Teko.

I resume below some of the Guaja examples in accordance with the notational convention and gloss used for Katukina-Kanamari's *-na=*: a case suffix surfacing as a head proclitic (examples renumbered; case-marked dependent in bold type):

<sup>II</sup><sub>noun</sub>

(80)(a) **ha-r=** *imirikó-a*  
1-GRAMCASEMARK= wife-REFERRER  
'my wife'

(b) **xahú-r=** *awáj-a*  
peccary-GRAMCASEMARK= tail-REFERRER  
'the peccary's tail'

postposition

(c) **ha-r=** *aké*  
1-GRAMCASEMARK= ADESSIVE  
'next to me'

(d) **pé-r=** *ipí*  
trail-GRAMCASEMARK= PERLATIVE  
'along the trail'

<sup>II</sup><sub>verb</sub>

(e) **ha-r=** *ixá* *aʔiá<sub>2</sub>*  
1-GRAMCASEMARK= see 3  
'He<sub>2</sub> saw me<sub>1</sub>. '

<sup>I</sup><sub>verb, unaccusative</sub>

(f) **ha-r=** *ahý*  
1-GRAMCASEMARK= BeSick  
'I am sick. '

Now, how long-lasting such a cohabitation must have been is a question that, for the time being, can only be answered in differential terms: certainly more in Tupi(-Guarani) than in Katukina-Kanamari, since in the latter the phonological form of the case-marker has remained untouched — that is, it has integrally preserved its allative-etymon make-up — while in Tupi(-Guarani) it (that is, the reconstructed *\*t-* / *\*tʔ* / *\*d*; see references in Meira & Drude 2013, and especially Cabral 2007 for *\*tʔ*) has endured extensive phonological bleaching such as dissimilation, lenition (Drude & Meira 2013), fusion with the head's initial segment (Gomes 2007:32; Ramirez & de França 2017), and even erosion down to zero. The profuse allomorphs of R- are, thus, the mere footprints of such a protracted (morpho-)phonotactic attrition.

At this point it is further of interest to wonder what *-r=* is, precisely, the case-marker of. The answer readily ensues from (80): taking *internal* / *external* in the understood sense of plainly observational syntactic positions, *-r=* case-marks the internal argument of governing heads.

Of course unaccusative verbs deserve special qualification in this context. Following a widely shared view initiated by Perlmutter (1978), in the syntactic derivation of unaccusative verbs the unique argument is "initially generated" as internal ("object"), then "raised" to external ("subject"), a conjecture unveiling a cogent commonality with genuine <sup>11</sup>verbs. This synchronically-"underlying" internal position of the sole argument faithfully mirrors its diachronically-prior internal position in the scenario put forward for Guaja unaccusatives.

The structure of phrases and the bound morphology therein are ultimately the contemporary reflexes of the way genitive noun phrases patterned at the time when

1. they shaped verb nominalizations, first used in subordinate clauses, then — coupled with re-finitization — as independent clauses (before pragmatic forces exert their effects on constituency hence word-order), and
2. some of them developed as postpositional phrases.

The Tupi family features ergative as well as accusative alignments in independent clauses, for instance ergative in the Guaporé / Aripuanã languages (Cabral 2002), but accusative in Apyãwa (Praça 2007) and Guaja (Magalhães 2007). Moreover, some languages, such as Juruna, are said to have switched from one alignment to the other (Cabral 2002). Given the assumptions put forth above (3 NOMINALIZATION, 4 NOMINALIZATION AND SUBORDINATION), such versatility at the family level adds weight to the view that the constant availability, to speakers, of oriented *grammatical* nominalizations may be a significant factor in fuelling alignment changes along the life-span of languages (see DeLancey 2011 for an instance thereof in Tibeto-Burman languages).

## 10 RESULTS AND PROSPECTS

This study was concerned with what the notion of internal argument can contribute to the understanding of phrase and clause patterns in languages where, plausibly, the structure of noun phrases has diachronically informed that of independent verb phrases *via* nominalization then (subordination then de-subordination then )re-finitization. Starting off with Katukina-Kanamari, then widening the empirical scope to a few other (families of) languages — Eskimo, Movima and Tagalog —, I piled up pieces of evidence that have been of service for facing the Tupi(-Guarani) facts through the prism of, mainly, Guaja. I recapitulate the results as a synopsis of the converging grammatical aspects featured by the languages discussed thus far. In doing so, I will content myself with taking into account *basic* independent-clause patterns exclusively.

<sup>11</sup>Verb phrases as A-predicates can be traced back to verb nominalizations. The nominalization orientation targets one argument, semantically patient (patient nominalization, ergative alignment) in all five reviewed languages but Guaja, in which the targetted argument is semantically agent (agent nominalization, accusative alignment).<sup>11</sup> Constituency and bound morphology — argument indexing, case-marking — yield a clear contrast between both arguments. In all languages the pronominal exponent of the syntactically non-privileged argument appears in the form and sequential position proper to

<sup>11</sup> Recall that, differently from the Kanamari dialect and Eskimo, Bia-Katukina, Movima and Tagalog have given raise to an alternatively-oriented nominalization targetting the agent participant.

a genitive. As a noun phrase it is also internal to the verb phrase in Katukina-Kanamari, Movima and — in a less obvious manner, at least synchronically — Tagalog (agent in the basic ergative pattern). In Eskimo and Guaja this noun phrase (agent in the former, patient in the latter) has migrated, at some point in the history of the languages, from inside to outside the verb phrase. Katukina-Kanamari might feature the incipient stage of such migration: the noun phrase *can* occur externally under pragmatically-driven motivations.

Now in Eskimo, Katukina-Kanamari and Tupi(-Guarani) we have in all likelihood a discrepancy as to the differential timing between the effects of head attraction and the noun phrase relocation off the verb phrase. Thus:

1. The Eskimo genitive suffix could not have clitized to the verb at the point where the noun phrase migrated since the latter took it along.
2. Katukina-Kanamari has initiated the migration process; though, since the case-marker is already procliticized, the noun phrase gets off its original locus as caseless.
3. Given the time-depth we can conjecture for the Tupi(-Guarani) case-suffix proclitization, and the ensuing phonological consequences, at no point would the noun phrase have done other than leave the verb phrase caseless.

Substantiating the claims made hitherto on Tupi(-Guarani) facts of course lies well outside the scope of the present paper. Rather, the set of proposals here submitted should be taken, above all else, as a program for future research. I will briefly sketch some of the open issues yet to be tackled. The heart of the matter is, doubtlessly, the origin of R- and its nature in synchronic perspective, not an accross-the-board grammatical phenomenon. Above I have put forward that R- is an internal argument case-suffix procliticized to the right adjacent head. My account is in no way competing with others', such as Jensen (1990), Rodrigues (1953, 1996) or Meira & Drude (2013), for the following reasons:

1. In the first-author's account, R- is a *linker*, which is tantamount to deliver a photograph of the facts at the more superficial level. (Hardly an enriching insight, one would think.)
2. The second-author's account is more complex. Cabral (1997), resuming previous work by Rodrigues, endows R- with a two-fold grammatical function: A. flagging contiguity; and B. establishing a *dependency* relation. The contiguity facet is equivalent to 1. under a less rudimentary phraseology. Now, if one understands "dependency" in Nichols (1986)'s sense, then we face two options. The first would strongly remind the Semitic *status constructus* (see Roberts 2001 for a basic description in Hebrew) as well as the Persian *ezāfe*, two amid the collection of Nichols' head-marking devices (the more common being pronominal indexing), whereby the head itself undergoes drastic phonological changes. The second option is to have on the head a dependency marker that, while missing in Nichols' inventory, fits exactly my hypothesized head attraction as exerted on a nominal case-marker.<sup>12</sup> Unfortunately, to date no principled implementation of what establishing such "dependency" might consist in seems to be available in the Tupi(-Guarani) literature.
3. Finally, Meira & Drude's analysis proves quite productive in motivating the synchronic phonological make-up of R- through cogent diachronic pathways — not a meagre

---

<sup>12</sup> Status constructus / *ezāfe* might ultimately qualify as the outcome of some utterly remote R-like case marker having been radically absorbed by the head.

affair. Such approach, although, is tangential to my purposes since my concern here is not R-'s phonological genealogy but its *grammatical nature*. The concluding lines of Meira & Drude's work, however, deserve to be paid special attention: an "intriguing possibility" is that the proteiform R- could result, in the final analysis, from contiguous morphosyntactic conditionings.

All said, the two tasks pending so as to settle the issue of Tupi(-Guarani) R- can be phrased, for now, as 1. identify the phonological form of the most remote precursor at, no doubt, the Proto-Tupi level; and 2. uncover a language in which a cognate of that putative precursor might still and undisputedly be a nominal case *suffix*. This would provide decisive support for the hypothesized nature of R- but, alas, also looks like a hopeless pursuit on the grounds given above: the vast time-depth presumed for the head attraction effects precludes the viability of unearthing any internal argument at a stage where no proclitization has yet taken place; therefore, no Eskimo-style case-marker should be expected on noun phrases located outside the verb phrase.

Notwithstanding that, I will indulge in a brief autobiographical tale that fits here as a possible roundabout route leading to such an achievement. When first studying the Kanamari dialect, I envisaged the case-marker nature of *-na=* as a cogent rationale behind such a mysterious *na-* "prefix" on heads. Uncovering it as a suffix case-marker soon looked like a definitely hopeless pursuit. Then I started documenting the Bia dialect. Eventually the case-marker suffix popped up on postpositions and on the recipient phrase of 'give' (see examples above): naturally, no right-located head to attach to was there available. In other words, opening the scope beyond arguments may be a serious option in tracking the origin of Tupi(-Guarani) R-. In a similar manner, it remains to see whether the plausible existence of an R-equivalent — let alone cognate — in Macro-Jê and Carib (Rodrigues 2009) could be conducive to any clue as regards the account proposed here.

Turning to the person prefix *i-* (and its allomorphs), one may wonder what basis there is for sustaining a pronominal function in spite of *i-*'s different distribution from first and second person clitics (as a reminder: *i-kirá*, 'he is fat' /vs./ *ha-r= ahý*, 'I am sick', Magalhães 2007:111 and 52 respectively; R- notation as *-r=* mine). The distribution of first and second person bound forms shows that they must be traced back to free pronouns, *i.e.* personal forms standing in noun phrase positions, clearly not the case for *i-*. The latter is plausibly better thought to have been nascent already when the first and second person free pronouns occurred internally and bound (for a detailed attempt at reconstruction, see Gildea 2002b). Its argument indexing function on <sup>II</sup>verbs and unaccusatives would result from the constraint whereby the internal argument must be overtly realized in its canonical position: likewise in Katukina-Kanamari, it stands for the corresponding noun phrase once the latter leaves the verb phrase.

The observant eye will have taken notice of an important gap in my account of <sup>I</sup>verbs constituency properties. In 9.2.1 THE STRUCTURE OF PHRASES *in fine* I argue that the unaccusative-verb's unique argument tightly parallels the <sup>II</sup>verb argument realizing the non-agent participant, both in independent and subordinate clauses. Though, such a property only could get full significance if found to contrast with the unergative-verb unique argument's properties, something not warranted for the time being: the evidence gathered through my inquiries (colleague scholars, published material) as to the noun phrase location — internal / external — in subordinate unergatives is anything but conclusive on either front: factors such as the language under consideration, subordinate-clause type, and even speakers' preferences, yield a picture far less clear than what I could reasonably cope with here. Notwithstanding, the topic remains of some interest to the question of whether, in subordinate clauses, the unaccusative unique argument is internal due to its "underlying"

syntactic derivation in synchrony (*cf.* the unaccusative hypothesis), or merely because in subordinates, mainly those eligible for a nominalization analysis, the unique argument of any <sup>I</sup>verb is internal. An indication as to a close parallelism, both in independent and subordinate clauses, between the <sup>II</sup>verb's internal argument and the unaccusative's unique argument on one hand, and <sup>II</sup>verb's external argument and the unergative's unique argument on the other hand — favoring, hence, the unaccusative hypothesis approach — might come from Nheengatu (Cruz 2011:492), were it not the language's peculiar history, encompassing significant divergences from the other Tupi(-Guarani) languages.

The last issue I will address is the bearing of my analysis of R- on the ongoing discussion of the omnipredicative hypothesis as applied to Tupi(-Guarani) languages (Queixalós 2001, 2006; Magalhães, Praça & Cruz 2019). The term *omnipredicativity* was coined by Launey (1986, 1994) more or less simultaneously to the accounts of Salishan languages by Thompson & Thompson ([prior to] 1983), Kinkade (1983) and Jelinek (1984, 1993), as well as Lemaréchal's proposal for two Philippine languages, Tagalog and Palawan (1989, 1991). The idea is that in some languages (almost) all the lexical classes are primarily F-predicates in the terms put forward above, something that entails a natural aptitude to occupy the A-predicate position (*cf.* 2 F-PREDICATES AND VALENCE). In many respects this notion of omnipredicativity and the more widespread one of non-configurationality (which also popped up in the eighties, Hale 1983 and three less-known previous papers by the same author) are the two faces of the same coin, one in logic perspective, the other in the more narrowly syntactic terms that prompted the "flat tree" metaphor: syntactic constituents do not cluster in larger constituents nor are they hierarchically ranked. The converging properties between these two approaches could seldom be a product of chance, as the following sample — selected for the present purposes — shows:

(81)

1. nouns are natural A-predicates; hence, no copula is needed;
2. the A-predicate is the only obligatory clause-constituent;
3. the A-predicate consists of a lexical item flanked by pronominal bound forms;
4. the pronominal bound forms are indeed the realization ("projection") of the predicate's arguments ("indexed arguments" in the sense of Haspelmath 2017) (see Jelinek 1984; outside the generative paradigm, Mithun 2003 and, earlier, Klimov 1974 for split-intransitive languages — which incidentally many Tupi-Guarani languages are said to be);
5. noun phrases are adjoined at the clause level (Jelinek 1984; Baker 1995) — in my terms coreferential adjuncts — or, equivalently, apposed to the argument-indexing morphology (*e.g.* Klimov 1974);
6. constituents, especially noun phrases, are syntagmatically discontinuous (Hale 1983);<sup>13</sup>
7. order of constituents is not set up by syntax but by pragmatic factors.

<sup>13</sup> In my view quite an ethnocentric appraisal of how languages tune up an entity-type and a referent: some build noun phrases, [*the three lost little children I told you about*]<sub>NP</sub>, while others scatter un-hierarchized full noun phrases all over the clause, [*child*]<sub>NP</sub> ... [*little-one*]<sub>NP</sub> ... [*mentioned-one*]<sub>NP</sub> ... [*threesome*]<sub>NP</sub> ... [*lost-one*]<sub>NP</sub>. See, for instance, Blake, B. (1983) on non-configurational Kalkatungu, and Luraghi (2010) on the "nouny" nature of "attributive adjectives" in non-configurational Ancient Indo-European languages.

For a retrospective of the different approaches, see Jelinek (2017). (Givón 2002 ch. 3 impugns the tenets of non-configurationality as a whole.) The non-configurationality "galaxy" allows for a large set of mixed-type "constellations". Of special interest in the present context is the co-existence in a single language of nouns as *primary* F-predicates, and strong constituency, as Movima exhibits and, to a significantly lesser extent, Katukina-Kanamari.

Unsurprisingly, at least two authors of generative obedience have, to my knowledge, addressed the question of non-configurationality in Tupi(-Guarani) languages, Leite (2001) on Apyãwa and Vieira (1993) on Trocara Assurini. The lack of mutual awareness between both perspectives — omnipredicativity / non-configurationality — is patent in how Leite (2001) deals with the referrer suffix: "Noun phrases [for agent and patient, FQ] take the same suffix *-a* on nouns with a final consonant." Finis.

As far as our immediate goals are concerned, there seem to be sufficient grounds for ascertaining that this family of languages lends itself to an omnipredicative / non-configurational account, however we dub it. Notwithstanding, the type of constituency discussed up to this point hits against the putative lack of internally-hierarchized complex constituents (an observation that has direct bearing on Movima as well).

I will start out calling to mind that the alleged omnipredicative nature of Tupi(-Guarani) languages is not asserted of present-day languages, but of a proto-X (the existence of the referrer suffix *-a* back to proto-Tupi is substantiated in Cabral 2001) whose currently observable offsprings show variably-downgraded stages of the omnipredicativity architecture. In such a perspective, the referrer's differential vitality in two given languages merely diagnoses the latter's different locations on the cline of the omnipredicativity recession. (Suggestively, the referrer is totally non-existent in contemporary Wayampi but appears as very much in evidence in Coudreau's 1892 record of the language.) The inescapably speculative scenario to be spelled out hereafter is based on two premises: first, Tupi(-Guarani) postpositions have their source in <sup>II</sup>nouns; second, nominalization prompts constituency.

1. The onset of the cycle features complex noun phrases internally hierarchized, a property shared with postpositional phrases. (Luraghi 2010 posits, for Ancient Indo-European languages, constituency starting from adpositional phrases then spreading to noun phrases then to verb phrases; so does Reinöhl 2016 for Indo-Aryan. In my view, for those adpositions that originate in <sup>II</sup>nouns, noun phrases should be temporally prior.)
2. Nominalization is used for complementation. Note that nominalizations can be reconstructed for proto-Tupi-Guarani (Jensen 1999). According to Aryon Rodrigues (p.c. 2005), some nominalizations *may be even older*.
3. As an effect of verb nominalizing, an argument is retrieved in the form of a genitive. The nominalization orientation targets the agent participant. Semantically the genitive-marked argument is, thus, a patient. This is where we get the internal-argument case-suffix R- altogether in noun phrases, postpositional phrases, and, once the nominalization is re-finitized as per its (innovative) independent status, verb phrases, a stage distinctively characterized by pervading strong constituency.
4. Communicative preferences eventually disfavor categorical utterances (theme / rheme, subject / A-predicate) while favoring thetic ones (all-rheme, compact formal make-up; Kuroda 1972-73; Sasse 1987; cf. also Lambrecht's 1987 "sentence focus").

Queixalós (2016) argues: first, that theticity is motivated by the communicative appeal of scoop effects and, second, that nominalizations deprived of nominally-expressed

arguments are a suitable tool for this kind of information packaging (here, **3 NOMINALIZATION**; see Moyse-Faurie 2011 for the use of nominalizations as exclamative clauses in Oceanic languages).

As this new pattern gains ground, the diachronically-derived verbal clause winds up as a verb flanked by the expression of its arguments through indexing morphology.

This leads to the emergence of omnipredicativity and its corollary emanation, the referrer suffix *-a* for endowing F-predicates with the capacity to refer (as mentioned above in dealing with Tagalog, Rose 2003:118-123 argues for a common origin in Teko between *-a* and a demonstrative *aʔe*).

The internal noun phrase has left the verb phrase, and the overall grammatical design meets all or most the properties listed in (81).

5. At this stage *any* noun phrase bears the referrer, internal arguments included. At least three languages unveil such a stage. In Teko, we can detect it in non-verbal F-predicates, that is, noun phrases and postpositional phrases, (82)-(Error! Reference source not found.) (only two postpositions take R-; Maurel 2015). See it also in Tocantins Assurini, (84). Notably, in Tocantins Assurini and Zo'e *-a* is extant in dependent verb phrases too, (85)-(86) (examples regulated).

- (82) [ [*tadz**at-a*]-*r=*                      *apat*  
chief-REFERRER-GRAMCASEMARK=*=*    weapon  
'the chief's weapon'

- (83) [ [*e-r=ait-a*]-*r=*                      *upi*                      *a-ho-tat*  
       1-GRAMCASEMARK=son-REFERRER-     COMITATIVE          1SINGULAR-go-FUTURE  
       GRAMCASEMARK=  
       ' I will go with my son. '

TEKO (Maurel p.c. March 2021; more examples in Rose 2003:125)

- (84) [sahý-a]-r=                      uwý-a]                  utururú    hehé  
moon-REFERRER-GRAMCASEMARK= blood-REFERRER ItDrops OnHim  
'The blood of the moon dropped on him.'

TOCANTINS ASSURINI (Cabral 1998)

- (85) *a-há* [ *[we-hy-a]-r=* *esák-a* ]  
 1-go 1COREFERENTIAL-mother-REFERRER-GRAMCASEMARK= see-GERUND  
 'I go see my mother.'

TOCANTINS ASSURINI (Cabral 2008)

- (86) *a-há* [ [*Xuam-a*]-**r=** *esak*]  
 1-go João-REFERRER-GRAMCASEMARK= see  
 'I go see João.'

Zo'E (Cabral 2008)

In other languages the loss of the referrer in the sequence [ [noun phrase-*a*]-*r*= head] could not be due to the morpheme's functional depletion in stages comparatively ahead on the omnipredicativity fading-cline: as we know, Guaja rules out the referrer *internally* while retaining it with great vitality elsewhere. One can neither call on some wide-ranging phonetic-based elision and leave it at that: 1. a quick inspection of the data does not abide

for generalizing on possible propitious environments; 2. critically, phonological factors may boost the omnipredicativity downfall but they are not the engine for it (see Cabral 2001 on languages having totally dispensed with *-a* while failing to meet any suitable phonological condition).

The case of Guaja demands further inspection. Admittedly, in an omnipredicative setup pronominal indexing conveys argumenthood while noun phrases stand as coreferential adjuncts. At the same time, no internal position could possibly host a coreferential adjunct. Featuring pronominally indexed arguments along with an internal noun phrase, this language presents itself as a hybrid on the time-line of diachrony: *in independent clauses*, [constituency / noun-phrase argumenthood] lingers into robustly settled [lack of constituency / noun-phrase "adjuncthood"]. In other words, Guaja exhibits vestiges of configurationality that overlap with dominant non-configurationality, a rather thought-provoking situation in the context of a family of languages undergoing, on the whole, an unmistakable decline of non-configurationality.

The above hypothesis of a drift in the history of Tupi(-Guarani) from configurationality (constituency) to non-configurationality (omnipredicativity) is backed by the following comparative facts regarding Uto-Aztecan (sources: Givón 2011; 2015 ch. 3, 18, 19, 20, 27; Tomas Givón p.c. April 2021; for Nahuatl, Launey 1994:83):

1. To the North we have conservative languages, with
  - overt case-marking in Yaqui and Ute;
  - rather rigid word-order in Yaqui with pre-verbal object, and high text-frequency of pre-verbal object in Ute;
  - nominalized subordinate clauses;
  - nouns in A-predicate position taking an overt copula in tense / aspect environments other than present / habitual (so-called zero copula).
2. Nahuatl, along with the Southern-most end of the family, is innovative: no case-marking, free word order, no nominalized subordinate clauses, and no copula whatsoever.
3. In the middle, a transitional area (*e.g.* Tarahumara) with a few typical-Northern features, mainly as relics in conservative environments (subordinate clauses).

In light of the notion of internal argument as embodied in a few genetically unconnected languages, I have sought to uncover the origin of the Tupi(-Guarani) *-R=* as a grammatical case-marker. I see head attraction — and hence *R=*'s "prefix" appearance, in fact a proclitic — as an epiphenomenon, and contend that *-R=* as a case-marker is yet synchronically functional behind its at first-sight fossilized-relic aspect (with the possible exception of a *stabilized* total fusion with the head's initial segment, *cf.* 9.2.2 THE CASE FOR *R=*). The hypothesized diachronic evolution has then, as an Ariadne's thread, led into a scale-shift embarking us on an ambitious course of inquiry concerning the abstract grammatical designs of these languages and their drifts along the line of time: from configurationality to non-configurationality / omnipredicativity and presumably back, with unavoidable intersecting phases (see Luraghi 2010 on an inverted Indo-European evolution: from non-configurationality to configurationality, then back to incipient non-configurationality in certain Romance languages). These results seem promising as the groundwork for future inquiries on Tupi(-Guarani). Conceivably, they might furnish the foundations for casting light on some of the still open concerns.



## REFERENCES

- Abney, S. (1987) *The English Noun Phrase and its Sentential Aspect*, PhD. Dissertation, Cambridge, M.I.T.
- Adelaar, A. (2005) "The Austronesian languages of Asia and Madagascar: a historical perspective" Adelaar, A. & Himmelmann, N. P. (eds) *The Austronesian Languages of Asia and Madagascar*, London Routledge, 1-42
- Aldridge, E. (2012) "Antipassive and ergativity in Tagalog" Laka, I. & Fernandez, B. (eds) *Accounting for Ergativity, Lingua* (special issue), 122.3, 192-203
- Anjos, dos, da Silva, Z. (2011) *Fonologia e Gramática Katukina-Kanamari*, PhD. Dissertation, Amsterdam Vrije Universit  t
- Ariel, M. (2008) *Pragmatics and Grammar*, Cambridge, Cambridge University Press
- Authier, G. (2016) "The antipassive construction in Dargi" *SLE Naples*, September 2, available at [https://www.academia.edu/28212766/The\\_antipassive\\_construction\\_in\\_Dargi](https://www.academia.edu/28212766/The_antipassive_construction_in_Dargi)
- Baker, M. (1995) *The Polysynthesis Parameter*, Oxford, Oxford University Press
- Blake, B. J. (1983) "Structure and word order in Kalkatungu: the anatomy of a flat language" *Australian Journal of Linguistics*, 3.2, 143-175
- Blake, F. R. (1916) "The Tagalog verb" *Journal of the American Oriental Society*, 36, 396-414
- Bloomfield, L. (1917) *Tagalog texts with grammatical analysis*, University of Illinois Studies in Language and Literature, 3
- Borella, C. de C  ssia (2000) *Aspectos morfossint  ticos da l  ngua Aweti (Tupi)*, Campinas, PhD. Dissertation, Universidade de Campinas
- Breugel, S. van (2008) *A Grammar of Atong*, Bundoora, PhD. Dissertation, La Trobe University
- Cabral, A. S. Arruda C  mara (1997) "Prefixos relacionais no Assurini do Tocantins" *Moara*, 8, 7-24
- Cabral, A. S. Arruda C  mara (2001) "Observa  es sobre a hist  ria do morfema -a da fam  lia Tup  -Guaran  " Queixal  s, F. (ed.) *Noms et verbes en tupi-guarani:   tat de la question*, Studies in Native American Languages 37, Munich, Lincom Europa, 133-162
- Cabral, A. S. Arruda C  mara (2002) "Natureza e dire   es das mudan  as de alinhamento ocorridas no tronco Tup  " Queixal  s, F. (ed.) *Ergatividade na amaz  nia I*, 85-87, available at <http://celia.cnrs.fr/FichExt/Documents%20de%20travail/Ergativite/1iCabral.htm>
- Cabral, A. S. Arruda C  mara (2007) "Fundamentando a Reconstru   o de Prefixos Relacionais para o Proto-Tup  " *Caderno de Resumos*, V Congresso Internacional da Associa   o Brasileira de Ling   stica, Belo Horizonte, 438-439
- Cabral, A. S. (2008) "Algunes consideracions sobre la llengua zo'  " *Marc de refer  ncies digital, revista catalana en format digital*, available at <http://www.referencies.cat/num32/art/art6.html>
- Ce  a, R. M. (1977) "Patient primacy in Tagalog" Paper presented at the LSA Annual Meeting, Chicago
- Copin, F. (2012) *Grammaire wayampi (famille tupi-guarani)*, Paris, PhD. Dissertation, Universit   de Paris 7
- Coudreau, H. A. (1892) *Vocabulaires m  thodiques des langues Ouayana, Aparai, Oyampi,   merillon*, Biblioth  que Linguistique Am  ricaine 15, Paris, Maisonneuve
- Cruz, A. da (2011) *Fonologia e gram  tica do Nheengat  *, Amsterdam, PhD. Dissertation, Vrije Universit  t

- Cruz, A. da (2017) "Towards an understanding of the origin of aspectual marks on nouns: evidence from Nheengatu and Tupinambá" Queixalós F. & Gomes, D. (eds) *O sintagma nominal em línguas amazônicas*, Brasília, Pontes, 45-73
- Couchili, T., Maurel, D. & Queixalós, F. (2002) "[Classes de lexèmes en émérillon](#)" *Amerindia* 26/27, 173-208
- De Guzmán, V. P. (1992) "Morphological Evidence for the Primacy of Patient as Subject in Tagalog" Ross, M. D. (ed.) *Papers in Austronesian Linguistics* 2, Canberra, Australian National University, 87-96
- DeLancey, S. (2011) "Finite structures from clausal nominalization in Tibeto-Burman" Yap, F.-H., Grunow-Harsta, K. & Wrona, J. (eds) *Nominalization in Asian Languages: Diachronic and Typological Perspectives*, Amsterdam / Philadelphia, John Benjamins, 243-362
- Durand, T. (to appear) "Insubordination and finitization in Arawakan languages" *International Journal of American Linguistics*
- Duranti, A. (1994) *From Grammar to Politics. Linguistic Anthropology in a Western Samoan Village*, Berkeley, University of California Press
- Evans, N. (2007) "Insubordination and its uses" Nikolaeva I. (ed.) *Finiteness: Theoretical and Empirical Foundations*, Oxford, Oxford University Press, 366-431
- Foley, W. A. (1998) "Symmetrical voice systems and precategoriality in Philippine languages" Butt, M. & King, T. H. (eds) *Proceedings of the Third Lexical Functional Grammar Conference, Workshop on Voice and Grammatical Functions in Austronesian Languages*, Brisbane, University of Queensland
- Foley, W. & Van Valin, R. (1984) *Functional syntax and universal grammar*, London / New York, Cambridge University Press
- Fortescue, M. (1995) "The historical source and typological position of ergativity in Eskimo languages" *Etudes/Inuit/Studies*, 19.2, 61-76
- Fortescue, M. (2000) "Parenté génétique des langues eskaléoutes" Tersis, N. & Therrien, M. (eds) *Les langues eskaléoutes*, Paris, CNRS Editions, 71-89
- Frege, G. (1891) *Function und Begriff. Vortrag gehalten in der Sitzung vom 9. Januar 1891 der Jenaischen Gesellschaft für Medizin und Naturwissenschaft (Function and Concept. An Address to the Jenaische Gesellschaft für Medizin und Naturwissenschaft on 9 January 1891)*, Verlag Hermann Pohle, Jena, 1891; translated as "Función y concepto", *Estudios sobre semántica*, Barcelona: Ediciones Ariel, 1973, 17-48
- Friedman, V. A. (2021) "Lak" Polinsky, M. (ed.) *The Oxford Handbook of Languages of the Caucasus*, Oxford, Oxford University Press, 201-242
- Galúcio, A. V. (2002) "O prefixo *i-* em Tupi: morfema antipassivo vs. marcador pronominal incorporado" *Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da Anpoll I*, Belém, UFPA, 274-287
- Gildea, S. (1998) *On Reconstructing Grammar: Comparative Cariban Morphosyntax*, Oxford, Oxford University Press.
- Gildea, S. (2002a) "Ergativity in the northern Cariban Languages" Queixalós, F. (ed.) *Ergatividade na Amazônia I*, Brasília, Universidade de Brasília, available at <http://celia.cnrs.fr/FichExt/Documents%20de%20travail/Ergativite/1kGildea.htm>
- Gildea, S. (2002b) "Pre-Proto-Tupí-Guaraní main clause person-marking" *Atas do I Encontro Internacional do Grupo de Trabalho sobre Línguas Indígenas da Anpoll*, 1, Cabral, A. S. & Rodrigues, A. (eds), Belém, Editora Universitária UFPA, 315-326
- Gildea, S. (2008) "Explaining similarities between independent clauses and nominalized phrases" Bruno, A. C., Pacheco, F., Queixalós, F. & Wetzels, L. (eds) *Amerindia* 32, 57-74

- Gildea, S. & Haude, K. (2021) *An Austronesian-type voice system in an Amazonian isolate*, University of Oregon Linguistics Department Colloquium, October 15
- Gildea, S. & Zúñiga, F. (2016) "Referential hierarchies: A new look at some historical and typological patterns" *Linguistics*, 54.3 483-529
- Givón, T. (1979) *On Understanding Grammar*, New York, Academic Press
- Givón, T. (1994) "The pragmatics of de-transitivizing voice" Givón, T. (ed.) *Voice and inversion*, Amsterdam / Philadelphia, John Benjamins, 3-42
- Givón, T. (1995) *Functionalism and grammar*, Amsterdam / Philadelphia, John Benjamins
- Givón, T. (1997) "Introduction" Givón, T. (ed.) *Grammatical Relations: a Functionalist Perspective*, Amsterdam / Philadelphia, John Benjamins, 1-84
- Givón, T. (2001) *Syntax : An introduction 1*, Amsterdam / Philadelphia, John Benjamins
- Givón, T. (2002) *Bio-linguistics: The Santa Barbara Lectures*, Amsterdam / Philadelphia, John Benjamins
- Givón, T. (2011) *Ute Reference Grammar*, Amsterdam / Philadelphia, John Benjamins
- Givón, T. (2015) *The Diachrony of Grammar*, Amsterdam / Philadelphia, John Benjamins
- Givón, T. (2018) *On Understanding Grammar*, revised edition, Amsterdam / Philadelphia, John Benjamins
- Givón, T. (2021) *The Life Cycle of Adpositions*, Amsterdam / Philadelphia, John Benjamins
- Gomes, D. (2017) "Sintagma nominal, verbal e posposicional em Mundurukú (Tupí): possível isomorfismo" Queixalós F. & Gomes, D. (eds) *O sintagma nominal em línguas amazônicas*, Brasília, Pontes, 263-290
- Haig, G. (2008) *Alignment change in Iranian languages. A Construction Grammar approach*, Berlin, Mouton De Gruyter
- Hale, K. (1983) "Warlpiri and the Grammar of Non-Configurational Languages" *Natural Language & Linguistic Theory* 1.1, 5-47
- Hale, K. & Keyser, J. (nd.) "The Basic Elements of Argument Structure" available at <http://lingphil.mit.edu/papers/hale/papers/hale012.pdf>
- Haspelmath, M. (2017) "Argument indexing: A conceptual framework for the syntactic status of bound person forms", available at <https://www.researchgate.net/publication/309109313>
- Haude, K. (2006) *A Grammar of Movima*, Nijmegen, PhD. Dissertation, Radboud Universiteit
- Haude, K. (2008) "Clause structure and alignment in Movima" Bruno, A. C., Pacheco, F., Queixalós, F. & Wetzels, L. (eds) *Amerindia* 32, 89-107
- Haude, K. (2009a) "Hierarchical Alignment in Movima" *International Journal of American Linguistics*, 75.4, 513-532
- Haude, K. (2009b) "Reference and predication in Movima" Epps, P. & Arkhipov, A. (eds) *New Challenges in Typology: Transcending the Borders and Refining the Distinctions*, Berlin, Mouton de Gruyter, 323-342
- Haude, K. (2010a) "The intransitive basis of Movima clause structure" Gildea, S. & Queixalós, F. (eds) *Ergativity in Amazonia*, Amsterdam / Philadelphia, John Benjamins, 285-315
- Haude, K. (2010b) "Referring to states and events: Subordination in Movima" Gijn, R. van, Haude, K. & Muysken, P. (eds) *Subordination in South American Languages*, 141-168
- Haude, K. (2012) "Undergoer orientation in Movima" Authier, G. & Haude, K. (eds) *Ergativity, valency and voice*, Berlin / Boston, de Gruyter Mouton, 259-287
- Haude, K. & Zúñiga, F. (2016) "Inverse and symmetrical voice: On languages with two transitive constructions" *Linguistics* 54.3, 443-481

- Haudricourt, A. G. (1979) "Importance de la relation équative en linguistique générale (sur des exemples de langues austronésiennes)" Paris, C. (ed.) *Relations prédicat-actants dans des langues de types divers*, Paris, SELAF, 11-14
- Heine, B. (2009a), "Grammaticalization of Cases" Malchukov, A. & Spencer, A. (eds) *The Oxford Handbook of Case*, Oxford, Oxford University Press, 458-469
- Heine, B. (2009b) "From nominal to clausal morphosyntax: complexity via expansion" Givón, T. & Shibatani, M. (eds) *Syntactic Complexity*, Amsterdam / Philadelphia, John Benjamins, 23-51
- Himmelman, N. P. (1991) "The Philippine challenge to universal grammar" *Arbeitspapier*, 15, Institut für Sprachwissenschaft, Cologne University
- Himmelman, N. P. (2005) "The Austronesian languages of Asia and Madagascar: typological characteristics" Adelaar, A. & Himmelman, N. P. (eds) *The Austronesian Languages of Asia and Madagascar*, London, Routledge, 110-181
- Ishy de Magalhães, P. H. (2012) *A phonological analysis of the Kanamari language (Katukina)*, Campinas, Master Thesis, Universidade de Campinas
- Ishy de Magalhães, P. H. (2018) *Kanamari do Juruá (família Katukina): aspectos fonológicos e morfosintáticos*, Campinas, PhD. Dissertation, Universidade de Campinas
- 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
- Jelinek, E. (2017) "Case and Configurationality" Martin-Callejo-Manandise, E. (ed.) *Coyote Papers: Working Papers in Linguistics from A-Z, Exploring Language: Linguistic Heresies from the Desert* 4 available at <https://repository.arizona.edu/handle/10150/226535>
- Jensen, C. (1990) "Cross-referencing changes in some Tupi-Guarani languages" Payne, D. (ed.) *Amazonian linguistics*, Austin, University of Texas Press, 117-158
- Jensen, C. (1999) "Tupi-Guarani" Dixon, R. & Aikhenvald, A. (eds) *The Amazonian Languages*, London / New York, Cambridge University Press, 125-163
- Judy, J. E. (1965) "Independent verbs in Movima" *Anthropological Linguistics* 7.7, 10-15
- Kaufman, T. (1974) "Meso-american Indian languages" *Encyclopaedia Britannica* 22, 767-774
- Keenan, E. (1976) "Towards a universal definition of subject" Li, C. (ed.) *Subject and topic*, New York, Academic Press, 305-333
- Kibrik, A. E. (2003) "Nominal inflection galore: Daghestanian, with side glances at Europe and the world" Plank, F. (ed.) *Noun Phrase Structure in the Languages of Europe* Mouton de Gruyter, Berlin New York, 37-112
- 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
- Kulikov, L. (2010) "Voice typology" Jae Jung Song (ed.) *The Oxford Handbook of Linguistic Typology*, Oxford, Oxford University Press, 368-398
- Kuroda, S. Y. (1972-73) "The Categorical and the Thetic Judgement. Evidence from Japanese Syntax" *Foundations of Language* 9.1, 153-185
- Lambrecht, K. (1987) "Aboutness as a Cognitive Category: The Thetic-Categorical Distinction Revisited" *Proceedings of the Thirteenth Annual Meeting of the Berkeley Linguistics Society*, 366-381
- Launey, M. (1986) *Catégories et opérations dans la grammaire nahuatl*, Thèse d'Etat, Paris, Université de Paris 7



- Launey, M. (1994) *Une grammaire omniprédicative. Essai sur la morphosyntaxe du nahuatl classique*, Paris, CNRS Editions
- Lazard, G. (1997) "Actance, diathèse : questions de définition" *Bulletin de la Société de linguistique de Paris* 92.1, 115-136
- Leite, Y. (2001) "O estatuto dos sintagmas nominais de sujeito e objeto em Tapirapé" Queixalós, F. (ed.) *Noms et verbes en tupi-guarani : état de la question*, Studies in Native American Languages 37, Munich, LincomEuropa, 89-101
- Lemaréchal, A. (1989) *Les parties du discours. Sémantique et syntaxe*, Paris, PUF
- Levin, B. (1983) *On the Nature of Ergativity*, PhD. Dissertation, Cambridge, Massachusetts Institute of Technology
- Luraghi, S. (2009) "Case in Cognitive Grammar" Malchukov, A. & Spencer, A. (eds) *The Oxford Handbook of case*, Oxford, Oxford University Press, 136-150
- Luraghi, S. (2010) "The rise (and possible downfall) of configurationality" Luraghi, S. & Bubenik, V. (eds) *The Continuum Companion to Historical Linguistics*. London / New York. Continuum, 212-229
- Luraghi, S. (2020) "External possessor constructions in Indo-European" Barðdal, J., Gildea, S. & Luján, E. R. (eds) *Reconstructing Syntax*, Leiden, Brill, 162-196
- Magalhães, M. (2007) *Sobre a morfologia e a sintaxe da língua Guajá (família Tupi-Guaraní)*, Brasília, PhD. Dissertation, Universidade de Brasília
- Magalhães, M. & Mattos, A. (2014) "Classes de palavras, tipos de predicados e sua relação com a intransitividade cindida em Guajá" *Via Litterae, Revista de Linguística e Teoria Literária*, 251-284
- Magalhães, M., Praça, W., & Cruz, A. da (2019) "Gradação da omnipredicatividade na família Tupi-Guaraní" *Forma y Función*, 32.2, 151-189, available at <http://dx.doi.org/10.15446/fyf.v32n2.80818>
- Mahieu, M.-A. (2004) "L'ergativité en nunavimmiutitut (Québec arctique)" Queixalós, F. (ed.) *Ergativity in Amazonia III*, Centre d'études des langues indigènes d'Amérique (CNRS, IRD) / Laboratório de Línguas Indígenas (UnB), 33-44, available at <http://celia.cnrs.fr/FichExt/Documents%20de%20travail/Ergativite/3fMahieu.htm>
- Mahieu, M.-A. (2009) "Objective conjugations in Eskaleout and Uralic: Evidence from Inuit and Mansi" Mahieu, M.-A. & Tersis, N. (eds) *Variations on Polysynthesis. The Eskaleout languages*, Amsterdam / Philadelphia, John Benjamins, 115-134
- Manning, Ch. & Sag, I. (1999) "Dissociations between argument structure and grammatical relations" Kathol et al. (eds), *Lexical and Constructional Aspects of Linguistic Explanation*, CSLI Publications, 236-258
- Maurel, D. (2019) *Le (vieux) problème : où coller, où séparer?* ms.
- Martinet, A. (1965) *La linguistique synchronique*, Paris, PUF
- Mead, D. (2005) "Mori Bawah" Adelaar, A. & Himmelmann, N. P. (eds) *The Austronesian Languages of Asia and Madagascar*, London Routledge 683-708
- Meira, S. & Drude, S. (2013) "Sobre a origem histórica dos "prefixos relacionais" das línguas Tupi-Guaraní" *Cadernos de Etnolinguística* 5.1
- Melčuk, I. (2006) *Aspects of the theory of morphology*, edited by David Beck, Berlin, Mouton De Gruyter
- Menovščikov, G. A. (1968) "Aleutskij jazyk [Aleut]" *Jazyki narodov SSSR*, 5, *Mongol'skie, tungusoman čžurskie i paleo-aziatskie jazyki*, Moscow, Nauka, 386-406
- Menovščikov, G. A. (1975) *Jazyk naukanskich eskimosov* [Naukan Eskimo], Leningrad, Nauka
- Mithun, M. (2003) "Pronouns and Agreement: The Information Status of Pronominal Affixes" *Transactions of the Philological Society* 101.2, 235-278
- Mithun, M. (2008) "The Extension of Dependency Beyond the Sentence" *Language*, 84.1, 69-119

- Moyse-Faurie, C. (2003) "The ergative features of Papuan and Austronesian languages" Queixalós, F. (ed.) *Ergativity in Amazonia II*, Centre d'études des langues indigènes d'Amérique (CNRS, IRD) / Laboratório de Línguas Indígenas (UnB), available at [http://celia.cnrs.fr/FichExt/Documents%20de%20travail/Ergativite/Introductions\\_ergativite.htm](http://celia.cnrs.fr/FichExt/Documents%20de%20travail/Ergativite/Introductions_ergativite.htm)
- Moyse-Faurie, C. (2004) "Politesse et violence verbale en Océanie" Angleviél, F. (ed.) *Violences océaniques*, L'Harmattan, 31-46
- Moyse-Faurie, C. (2011) "Nominalisation and exclamation in Oceanic languages" Moyse-Faurie, C. & Sabel, J. (eds), *Topics in Oceanic Morphosyntax*, Berlin, Mouton de Gruyter, 135-160
- Nichols, J. (1986) "Head-marking and dependent marking grammar" *Language* 62.1, 56-119
- Nichols, J. & Bickel, B. (2005) "Locus of marking" Haspelmath, M., Dryer, M., Gil, D. & Comrie, B. (eds) *The World Atlas of Language Structures*, 98-109, Oxford, Oxford University Press
- Palancar, E. (2016) "On the tightrope between infinitives and action nouns: The case of Otomi nominalizations" Chamoreau, C. & Estrada, Z. (eds), *Finiteness and Nominalization* Amsterdam / Philadelphia, John Benjamins, 141-170
- Palancar, E. (2009) "Varieties of ergative" Malchukov, A. & Spencer, A. (eds) *The Oxford Handbook of case*, Oxford, Oxford University Press
- Pawley, A. K. & Reid, L. A. (1979) "The evolution of transitive constructions in Austronesian" Naylor, B. (ed.) *Austronesian studies: Papers from the Second Eastern Conference of Austronesian Languages*, Ann Arbor, Center for South and Southeast Asian Studies, University of Michigan, 103-130
- Perlmutter, D. (1978) "Impersonal Passives and the Unaccusative Hypothesis", Jaeger, J. et al. (eds) *Proceedings of the 4th Annual Meeting of the Berkeley Linguistic Society*, 157-189
- Plank, F. (1979) "Ergativity, Syntactic Typology and Universal Grammar: Some past and present viewpoints" Plank, F. (ed.) *Ergativity*, London, Academic Press, 3-36
- Praça, W. (2007) *Morfossintaxe da língua Tapirapé (família Tupi-Guarani)*, PhD. Dissertation, Brasília, Universidade de Brasília
- Payne, T. E. (1997) *Describing morphosyntax. A guide for field linguists*, London / New York, The University of Cambridge Press
- Queixalós, F. (2001) "Le suffixe référentiel en émérillon" Queixalós, F. (ed.) *Noms et verbes en tupi-guarani : état de la question*, Studies in Native American Languages 37, Munich, LincomEuropa, 117-132
- Queixalós, F. (2006) "The Primacy and Fate of Predicativity in Tupi-Guarani" Lois, X. & Vapnarsky, V. (eds) *Root classes and lexical categories in Amerindian languages*, Viena, Peter Lang, 249-287
- Queixalós, F. (2010) "Grammatical Relations in Katukina-Kanamari", Gildea, S. & Queixalós, F. (eds) *Ergativity in Amazonia*, Amsterdam / Philadelphia, John Benjamins, 235-284
- Queixalós, F. (2013) *L'ergativité est-elle un oiseau bleu?* Munich, LincomEuropa, LSLT 26
- Queixalós, F. (2015) "Formas jazentes em Katukina-Kanamari e em Tupi-Guarani" ABRALIN, Cruz, A., Praça, W. & Magalhães, M. (orgs) *Workshop on Formas jazentes*, Belem, February 16
- Queixalós, F. (2016) "The role of nominalization in theticity: a Sikuani contribution" Chamoreau, C. & Estrada Fernández, Z. (eds) *Finiteness and Nominalization*, Amsterdam / Philadelphia, John Benjamins, 205-242

- Queixalós, F. (2017a) "Valence in Katukina-Kanamari noun phrases and the nature of genitive classifiers" Queixalós, F. & Gomes, D. *O sintagma nominal em línguas amazônicas*, Brasília, Pontes 141-176
- Queixalós, F. (2017b) "What being a Syntactically Ergative Language means for Katukina-Kanamari" Coon, J., Massam, D. & Demena Travis, L. (eds) *The Oxford Handbook of Ergativity* Oxford, Oxford University Press, 1035-1064
- Ramirez, H. & de França, M. C. V. (2017) "O warázu do Guaporé (Tupi-Guarani): primeira descrição linguística" *Liames* 17.2, 1-96
- Reinöhl, U. (2016) *Grammaticalization and the Rise of Configurationality in Indo-Aryan*, Oxford, Oxford University Press
- Rigsby, B. (1986) *Gitksan grammar*, Victoria, British Columbia Provincial Museum Linguistics Division ms.
- Roberts, I. (2001) "Head movement" Baltin, M. & Collins, C. (eds) *The Handbook of Contemporary Syntactic Theory*, 113-147
- Rodrigues, A. (1953) "Morfologia do verbo tupí" *Letras*, 1, 121-153
- Rodrigues, A. (1996) "Argumento e predicado em Tupinambá" *Boletim da Associação Brasileira de Linguística* 19, 6-18
- Rodrigues, A. (2009) "A Case of Affinity Among Tupí, Karíb, and Macro-Jê" *Revista Brasileira de Linguística Antropológica* 1.1, 137-161
- Rodrigues, A. & Dietrich, W. (1997) "On the linguistic relationship between Mawé and Tupí-Guaraní" *Diachronica* 14.2, 265-304
- Roland, D., Dick, F. & Elman, F. L. (2007) "Frequency of Basic English Grammatical Structures: A Corpus Analysis" *Journal of Memory and Language* 57.3, 348-379
- Sasse, H.-J. (1987) "The thetic / categorial distinction revisited" *Linguistics* 25, 511-580
- Sasse, H.-J. (1991) "Predication and Sentence Construction in Universal perspective" Zaefferer, D. (ed.) *Semantic Universals and Universal Semantics*, Berlin / New York, Foris Publications, 76-95
- Schachter, P. (1977) "Reference-related and role-related properties of subjects" Cole, P. & Sadock, J. M. (eds) *Grammatical Relations*, Syntax and Semantics 8, New York, Academic Press, 279-305
- Schachter, P. & Otnes, F. T. (1972) *Tagalog Reference Grammar*, Berkeley, University of California Press
- Shibatani, M. (1985) "Passives and related constructions : A prototype analysis" *Language* 61.4, 821-848
- Shibatani, M. (2021) "Syntactic typology" *Oxford Research Encyclopedia of Linguistics*, Oxford University Press, available at <https://doi.org/10.1093/acrefore/9780199384655.013.154>
- Silva, R. da (2010) *Estudo morfossintático da língua sateré-mawé*, Campinas, PhD. Dissertation, Universidade de Campinas
- Silva, B. Carreta Corrêa da (1997) *Urubú-Ka'apor, da Gramática à História: A Trajetória de um Povo*, Brasília, Master Thesis, Universidade de Brasília
- Starosta, S., Pawley, A. K. & Reid, L. A. (1982) "The Evolution of Focus in Austronesian" Amran, H., Carrington, L. & Wurm, S. A. (eds) *Papers from the third International Conference on Austronesian Linguistics 2, Tracking the travellers*, Canberra, Australian National University, 145-170
- Thompson, L. C. & Thompson, M. T. ([prior to] 1983) *Thompson*, ms. quoted in Kinkade (1983), published as Thompson, L. C., Thompson, M. T. & Egesdal, S. M. (1996) "Sketch of Thompson, a Salishan language " Goddard, I. (ed.), *Handbook of North American Indians* 17, *Languages*, Washington D.C., Smithsonian Institution 609-643

- Valenzuela, P. (2003) *Transitivity in Shipibo-Konibo*, PhD. Dissertation, Eugene, University of Oregon, and Ann Arbor, UMI Dissertation Services
- Vieira, M. D. (1993) *O problema da configuracionalidade na língua Assurini: uma consequência da projeção dos argumentos do predicado verbal*, Campinas, PhD. Dissertation, Universidade de Campinas
- Vincent, N. (2001) "LFG as a model of syntactic change" Butt, M. & Holloway King, T. (eds) *Time over matter. Diachronic perspectives on morphosyntax*, 1-42
- Williams, E. (1981) "Argument Structure and Morphology" *The Linguistic Review* 1, 81-114
- Witzlack-Makarevich, A. (2010) *Typological variation in grammatical relations*, PhD. Dissertation, Leipzig, Universität Leipzig



This work has benefited from valuable inputs by Gilles Authier, Aline Cruz, Alex François, Spike Gildea, Tomas Givón, Marina Magalhães, Didier Maurel, Claire Moyse, Walkíria Praça, Malcolm Ross and Masayoshi Shibatani. Also, since 2015 a continued exchange in epistolary form with Marina Magalhães significantly helped maturing my views on the topics at stake. It goes without saying that these generous people are all absolved of any misuse I might have made of their contributions.