

Liuro VI. Cap. XIII.

opiápe catû aipo é:jara oimopor aipô
éra. Anhangá ratápe coír oicôbaé,aépe

Fabrcio Ferraz Gerardi

A Role and Reference Grammar Description of Tupinambá



robáque nde rureímêbê nde jucáeimi
è oyepíca oangaturamamo ê, nden hé-
a rarômo ê. Emonanamamo Tûpá nheé-
goêra cuápa, nde remimôbeu poêra,
urai agoérabé, opabinhe ymoacipira
pira çocé, ymoáciabo, ceroîromo, eneí

**A ROLE AND REFERENCE GRAMMAR
DESCRIPTION OF TUPINAMBÁ**

Fabrício Ferraz Gerardi

A ROLE AND REFERENCE GRAMMAR DESCRIPTION OF TUPINAMBÁ

EBERHARD KARLS
UNIVERSITÄT
TÜBINGEN



TÜBINGEN
LIBRARY PUBLISHING

Bibliografische Information der Deutschen Nationalbibliothek

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie, detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.



Dieses Werk ist lizenziert unter einer Creative Commons Namensnennung - Keine Bearbeitungen 4.0 International Lizenz. Um eine Kopie dieser Lizenz einzusehen, konsultieren Sie <https://creativecommons.org/licenses/by-nd/4.0/legalcode> oder wenden Sie sich brieflich an Creative Commons, Postfach 1866, Mountain View, California, 94042, USA.

Die Online-Version dieser Publikation ist auf dem Repositorium der Universität Tübingen frei verfügbar (Open Access):

<http://hdl.handle.net/10900/142263>

<http://nbn-resolving.de/urn:nbn:de:bsz:21-dspace-1422636>

<http://dx.doi.org/10.15496/publikation-83610>

Tübingen Library Publishing 2023
Universitätsbibliothek Tübingen
Wilhelmstraße 32
72074 Tübingen
druckdienste@ub.uni-tuebingen.de
<https://tlp.uni-tuebingen.de>

ISBN (Hardcover): 978-3-946552-84-0

ISBN (PDF): 978-3-946552-85-7

Umschlaggestaltung: Sandra Binder, Universität Tübingen

Coverabbildung: Theodore de Bry, Public domain, via Wikimedia Commons

Satz: Fabrício Ferraz Gerardi

Druck und Bindung: Druckhaus Sportflieger in der Medialis Offset GmbH

Printed in Germany

Abstract

Tupinambá is the first attested language of the Tupí-Guaraní family and it has a story that one can follow from its first attestation to the present through its descendants. Making use of RRG, a linguistic theory that is informed by cross-linguistic diversity, I present the first typologically adequate description of Tupinambá. This description introduces the main aspects of Tupinambá grammar, including phonology, morphology, syntax, and information structure, accounting for the interface between syntax, semantics, and pragmatics.

Acknowledgements

I would like to thank those who, in different ways, have participated in the process of this dissertation. Gerhard Jäger, for allowing me to work on it and on other projects related to the languages of South America; Robert Van Valin Jr., for the generous offer of his time to comment on my use of his theory, for the friendly meetings; Wolf Dietrich, for his time, comments, friendly meetings, and for his invaluable contribution to the study of Tupí-Guaraní languages. I stand on the shoulders of you three.

I thank my friend Stanislav Reichert, for his friendship, generous support, and help in making many of the sources available to me.

To my friend Thiago Blumenthal: *você nos deixou tão cedo; saudade, lágrimas, e sorrisos.*

Lali and Carolina Aragon, since I cannot thank you enough, I dedicate this work to you both.

Contents

- 1 Introduction** **7**
- 1.1 The People 10
- 1.2 The Language 18
- 1.3 Typological Profile of TUP 23
- 1.4 Linguistic-Ethnographic remarks 25
- 1.5 Primary sources and previous work on Tupinambá 29

- 2 Phonology** **41**
- 2.1 Segmental phonology 42
- 2.1.1 Consonants 43
- 2.1.2 Phonetic realizations and phonological processes . . . 46
 - 2.1.2.1 Stops 47
 - 2.1.2.2 Fricatives 48
 - 2.1.2.3 Nasals 49
 - 2.1.2.4 Flaps 51
 - 2.1.2.5 Approximants 52
 - 2.1.2.6 Morphophonological processes 53
- 2.1.3 Vowels 54
- 2.1.4 Phonetic realizations and morphophonological processes 56
- 2.2 Syllable structure 56

2.3	Prosody	58
3	RRG	59
3.1	Introduction	59
3.1.1	The syntactic representation: the layered structure of the clause (LSC)	65
3.1.2	The constituent projection	67
3.1.3	The operator projection	70
3.1.4	The focus projection	72
3.2	The Layered Structure of the Reference Phrase	77
3.3	The semantic representation	80
3.3.1	Lexical representation of nominals	88
3.3.2	Semantic Macroroles and Lexical Entries for Verbs	90
3.3.3	Syntactic functions, case and preposition assignment	91
3.4	Linking algorithms	93
4	Word classes	97
4.1	Introduction	97
4.2	Word classes in Tupinambá	100
4.3	Noun classes	104
4.3.1	Argument indexes and pronouns	104
4.3.2	Relational Markers	111
4.3.3	Adverbs	119
5	Basic Clause Patterns	123
5.1	M-intransitive Verbs	124
5.2	M-transitive Verbs	128
5.3	Ditransitive Verbs	133
5.4	Privileged Syntactic Argument	138
5.5	Nonprototypical ('non-verbal') predication	140
5.6	Word order	145
5.7	Valency changing	148
5.7.1	Causatives	149
5.7.1.1	Causative of M-intransitive predicates	149
5.7.1.2	Causative of M-transitive predicates	154

5.7.1.3	Sociative causative	159
5.7.2	Incorporation	160
5.7.3	Antipassives: <i>poro</i> and <i>mbaʔe</i>	168
5.7.4	Reflexive, middle, and reciprocal voice	170
5.7.4.1	Reflexive	171
5.7.4.2	Middle	173
5.7.5	Reciprocal verb constructions	177
5.7.6	Is there a passive voice in TUP?	179
6	The Layered Structure of the Clause	183
6.1	PrDP	183
6.2	Pre-core slot (PrCS)	185
6.3	Extra-core slot (ECS)	188
6.4	The periphery	191
6.4.1	Modifier Phrase	194
6.5	Operators	195
6.5.1	Clause-level operators	196
6.5.1.1	Illocutionary force	197
6.5.1.2	Evidentials	207
6.5.1.3	Status	209
6.5.1.4	External negation	212
6.5.1.5	Tense	215
6.5.2	Core-level operators	219
6.5.2.1	Modality	219
6.5.2.2	Hortative modality	220
6.5.2.3	Core negation	222
6.5.3	Nuclear-level operators	223
6.5.3.1	Aspect	223
6.5.3.2	Nuclear negation	226
7	Lexical categories	229
7.1	Predicate semantic classes	229
7.1.1	State predicates	230
7.1.2	Activity predicates	232
7.1.3	Achievement predicates	233

7.1.4	Semelfactive	234
7.1.5	Accomplishments	234
7.2	Morphological predicate types	235
7.3	Nominal Categories	236
7.3.1	Nominal tense	236
7.3.2	Diminutive and augmentative	241
7.3.3	Nominal number	243
7.3.4	Alienability and Inalienability	245
7.4	Postpositions	246
7.4.1	Postposition <i>swi</i>	248
7.4.2	Postposition <i>ese</i>	248
7.4.3	The case of <i>irũ</i>	249
7.4.4	Some special cases	250
7.4.5	Other postpositions	252
7.5	Postposition assignment	254
7.6	Postpositional Phrase	256
7.6.1	Adjunct postpositional phrases	256
7.6.2	Argument-marking postpositional phrases	257
7.7	Particles	258
7.7.1	Aspect particles	259
7.7.2	Core modifying particles	259
7.7.3	Discourse particles	260
7.7.4	Illocutionary particles	262
7.8	Case	263
7.8.1	Locative-Dative cases	263
7.8.2	Perlative case	266
7.8.3	Translative case	267
8	Reference Phrase	269
8.1	Minimal RPs	269
8.2	RP operators	272
8.2.0.1	Adnominal quantifiers	273
8.2.0.2	Nominal negation	280
8.2.1	RP operators	282
8.2.1.1	Demonstratives	282

8.3	Nominalizers	285
8.4	Modifier Phrase	292
	8.4.0.1 Adverbial demonstratives	297
8.5	Attributive possession	299
9	Information Structure	305
9.1	Introduction	305
9.2	Predicate focus	307
9.3	Sentence focus	307
9.4	Focus positions in the syntax	309
	9.4.1 The Pre-Detached Position	309
	9.4.2 Pre-core slot	310
9.5	Narrow focus	313
9.6	Indicative II or oblique-topicalized?	314
	9.6.1 Topic	319
9.7	Nonfocal argument index	321
10	Complex sentences	325
10.1	Levels of juncture	325
	10.1.1 Nuclear junctures	326
	10.1.2 Core junctures	327
	10.1.3 Clause junctures	328
	10.1.4 Sentential junctures	329
10.2	Nexus relations	329
	10.2.1 Coordination	331
	10.2.2 Subordination	334
	10.2.3 Cosubordination	340
10.3	Complex RPs	361
	10.3.1 Coordination	361
	10.3.2 Relative clauses	366
	10.3.2.1 Restrictive relative clauses	368
11	Conclusion	375

List of Figures

1.1	Map showing possible Tupinambá presence along the coast in 1500 represented by the green shade	10
1.2	Representations of TUP villages with houses, patios, and fences	38
1.3	The Tupí-Guaraní languages (in green) along with the distribution of the TG archaeological record (black dots). TUP is represented by the dot on the coast (north-east). From Ferraz Gerardi et al. (2023)	39
2.1	Syllable structure	57
3.1	The architecture of RRG	61
3.2	Syntactic templates for a TUP WH-question. The dashed line indicates the focus domain and the triangle marks the narrow focus	64
3.3	WH-question with precore slot	65
3.4	Simple constituent projection	66
3.5	LSC with periphery at the core-level	68
3.6	PrCS with WH-word	69
3.7	The Layered Structure of the Clause	69
3.8	ECS with RPs coreferential with the bound argument indexes in the core	70

3.9	Some syntactic templates for English from Van Valin Jr (2022, 42)	71
3.10	Constituent and operator projections	73
3.11	Alternative operator projection	74
3.12	The focus projection	75
3.13	RPIP with definiteness and deixis operators	78
3.14	The layered structure of the RP	79
3.15	Complex RP	80
3.16	Thematic relations continuum along the actor-undergoer hierarchy. From Van Valin Jr (2001a)	85
3.17	Increasing generalization of semantic contrasts (from Van Valin Jr 2005, 54)	87
3.18	The actor-undergoer hierarchy. The arrows indicate the increasing markedness of the realization of an argument as macrorole	90
3.19	The linking algorithm, from Van Valin Jr (2005, 129)	94
4.1	Five coexpression types of function indicators from Haspelmath (2021b)	100
4.2	Independent pronoun in actor function with portmanteau argument index	109
5.1	A core template with three core arguments for TUP	124
5.2	Embedded possessive construction	133
5.3	Incorporation with possessor stranding	137
5.4	Representation of (188)	151
5.5	Representation of (189a), a causative construction with three core arguments	152
5.6	Factive construction with nuclear juncture (complex nucleus) in French	157
5.7	Causative clause with peripheral argument	158
5.8	Nominal incorporation	162
5.9	The layered structure of the word	164
5.10	Layered structure of a word with noun incorporation	164
5.11	Incorporation of possessed noun	167

5.12	Reflexive linking in TUP	172
6.1	PrDP template	184
6.2	Sentence with pre-detached position	184
6.3	Nuclear peripheral modifier	192
6.4	Core peripheral modifier	193
6.5	Clausal peripheral modifier	193
6.6	Modifier phrase	194
6.7	Nominal modification	195
6.8	RRG operator projection	196
6.9	Example of an imperative sentence	205
6.10	The tense operator	218
7.1	Object word in predicate function	230
7.2	The nominal tense RP operator	239
7.3	Structure of a predicative PP	256
7.4	Predicative postposition	257
7.5	PP as core argument	258
8.1	LSPR of a minimal RP	270
8.2	Demonstrative in argument function	271
8.3	Negation as a core-level operator of the RP	281
8.4	A deixis operator	283
8.5	Modification by composition in TUP	293
8.6	Recursive modification by composition in TUP	293
8.7	RP possessor in the RPIP	296
8.8	Embedded possessor RP	297
8.9	Nominal tense operator	299
9.1	Predicate focus in TUP	308
10.1	English nuclear juncture	327
10.2	Nuclear juncture	327
10.3	English core juncture	328
10.4	Nexus relations, from Van Valin Jr (2005, 188)	329

10.5	Types of connection in complex constructions, from Pavey (2010, 226)	330
10.6	Parataxis. Clausal juncture without a clause linkage marker . .	333
10.7	Example of (object) complementation (subordination) in English	335
10.8	Peripheral ad-clausal subordination	337
10.9	Ad-clausal subordination	338
10.10	Ad-clausal subordination with a clause linkage marker	339
10.11	Subordination at the Sentence-level	340
10.12	Nuclear cosubordination	343
10.13	Core cosubordination	345
10.14	Nuclear and core cosubordination in a sentence	348
10.15	Complex core with periphery in cosubordination	354
10.16	RP coordination	362
10.17	RP subordination	366
10.18	RP subordination	367
10.19	Relative clause or RP subordination	368
10.20	Restrictive relative clause	369

List of Tables

1.1	Prefixes and suffixes in TUP	26
1.2	Sources for the Tupinambá language	30
1.3	Some references to descriptions of TG languages	37
2.1	Tupinambá consonant phonemes	43
2.2	Tupinambá vocalic phonemes	43
2.3	Words with [h]	44
3.1	Constructional schema for TUP WH-questions	64
3.2	Semantic units underlying the syntactic units of the layered structure of the clause	66
3.3	Some of the most common operators and their respective levels in the LSC	72
3.4	RP operators	79
3.5	Characterization of verbal <i>Aktionsart</i> based on Van Valin Jr (2022)	80
3.6	<i>Aktionsart</i> and their logical structures, from Van Valin Jr (2005, 45)	82
3.7	Decomposition of state predicates and thematic relations	85
3.8	Decomposition of activity predicates and thematic relations	86

4.1	Information packaging and semantic root classes combined in English. The elements in the diagonal are those exhibiting “default” behavior	99
4.2	Semantic classes combined with propositional act functions and overtly marked structural coding for POS in Tupinambá .	101
4.3	Tupinambá person markers	105
4.4	Possessor markers in Mawé, Awetí, and coreferential markers in Tupinambá (Set III)	108
4.5	Tupinambá relational markers: R ₁ marks contiguity, R ₂ marks non-contiguity, R ₃ indicates coreference, and R ₄ indicates that the possessor is human	112
4.6	TUP noun classes. Examples in first person singular and third person	112
5.1	Some active and stative roots in Tupinambá. Colored pairs are antonyms that require different markers for active and stative forms	127
5.2	Example of an M-transitive verb paradigm	130
5.3	Causativized intransitive roots	153
6.1	Some WH-words in Tupinambá	187
6.2	Tupinambá operators	196
6.3	Categories of Illocutionary Force	197
6.4	Tupinambá WH-words	203
6.5	Imperative argument indexes	203
7.1	Some TUP postpositions formed out of spatial nouns or body-parts	254
8.1	RP levels and their operators	273
8.2	Demonstratives in TUP	283
8.3	Lexical roots with <i>emi-</i> and <i>-pir</i>	290
9.1	WH-words in Tupinambá	311
9.2	Non-focal constructions with fronted adverbials in TG languages	318

10.1 Frequency of *kojpo* and *konipo* 364

List of Abbreviations

1	first person
2	second person
3	third person
A	agent-like argument of canonical transitive verb
ABL	ablative
ACT	actor
ADV	adverb(ial)
AFD	actual focus domain
ALL	allative
ANTIP	antipassive
ARG(s)	argument(s)
ASP	aspect
AUG	augmentative

AUH acto-undergoer hierarchy
CAUS causative
CLM clause linkage marker
CONJ conjunction
DAT dative
DCA direct core argument
DELIB deliberative
DEM demonstrative
DEV_{PASS} deverbaler (passive)
DUR durative
ECS extra-core slot
EMPH emphatic
EPEN epenthesis
EV_{fh} evidential first hand
EXCL exclusive
FAC factive
FOC focus
FUT future
GER gerund
HAB.AG habitual agent
HOR hortative
hort permissive
IF illocutionary force

IMP	imperative
INCL	inclusive
INGR	ingressive
INTJ	interjection
INTS	intensive
IRR	irrealis
ITER	iterative
IU	information unit
LOC	locative
LS	logical structure
LSC	layered structure of the clause
LSPP	layered structure of the postpositional phrase
LSRP	layered structure of the reference phrase
LSW	layered structure of the word
MP	modifier phrase
NEG	negation, negative
NFOC	nonfocal
NMLZ	nominalizer/nominalization
NMLZ _{AG}	nominalizer (agentive)
NMLZ _{CIRC}	nominalizer (circumfix)
NMLZ _{PAT}	nominalizer (patient)
NMLZ _{REL}	nominalizer (relativizer)
NPST	nonpast

NUC	nucleus
OBJ	object
OBL	oblique
OCA	oblique core argument
OP	operator projection
OPT	optative
P	patient-like argument of canonical transitive verb
PERI	periphery
PFD	potential focus domain
PL	plural
PoCS	postcore slot
PoDP	post-detached position
POS	part-of-speech
POSP	postposition
POSS	possessive
PP	postpositional phrase
PRCL	particle
PRCL _F	particle used only by women
PRCL _M	particle used only by men
PrCS	precore slot
PrDP	pre-detached position
PROC	process
PRON	pronoun

PSA	privileged syntactic argument
PST	past
Q	question
R	relational morpheme
R ₁	relational morpheme of contiguity
R ₂	relational morpheme of non-contiguity
R ₃	relational morpheme of co-reference
R ₄	relational morpheme of human dependent
RECP	reciprocal
RED	reduplication
RED _D	reduplication disyllabic
RED _M	reduplication monosyllabic
REF	referential
RFLX	reflexive
RP	reference phrase
RPIP	reference phrase initial position
S	single argument of canonical intransitive verb
SCAU	sociative causation
SEMEL	semelfactive
SG	singular
SR	switch reference
TG	Tupí-Guaraní
TNS	tense

TOP topic

TRSL translative

TUP Tupinambá

UND undergoer

VOC vocative

Primary sources

AA Anchieta Arte (Anchieta 1595)

Araújo Araújo Catecismo (Araújo 1618b)

CC Cartas Camarões (Navarro 2022)

DC I Anchieta Doutrina Cristã I (Anchieta 1618a)

DC II Anchieta Doutrina Cristã II (Anchieta 1618b)

FA Figueira Arte (Figueira 1687)

Poemas Anchieta Poemas (Anchieta 1997)

Teatro Anchieta Teatro (Anchieta 2006)

Chapter 1

Introduction

Nenhuma língua primitiva do mundo, nem mesmo o sânscrito, ocupou tão grande extensão geográfica como o tupi e seus dialetos; com efeito, desde o Amapá até o Rio da Prata [...] desde o Cabo de São Roque até o Javari, [...] estão, nos nomes dos lugares, das plantas, dos rios e das tribos indígenas [...] os imperecedores vestígios dessa língua.
Magalhães, 1876, p. 28

This dissertation aims to describe Tupinambá (TUP) or, more precisely, the language attested almost exclusively in Jesuit religious texts, a native South American language spoken along the coast of Brazil at the time of the Brazilian ‘discovery’ in 1500. This language belongs to the Tupían family in which it is a member of the Tupí-Guaraní (TG) branch. The Tupían family is one of the largest language families in South America (see Section 1.2).

This language has been dead for about three hundred years, a fact that, combined with its relatively small corpus, poses a challenge for its description. Furthermore, although the term Tupinambá is controversial (see Section 1.2),

I employ it throughout this work to refer to whatever variant of the language was spoken on the coast, attested from Staden to Bettendorff (see Section 1.5). I also avoid the name Tupí, commonly used in Brazil, since “Tupí” has established itself as the name of the language family.

As expressed by Payne (1997, 2),

‘when a language does become extinct, . . . the linguistic description and other materials remain as a central part of the cultural heritage of descendants of the language’s speakers, as well as of all humanity. Without this documentation, the language, along with the cultural traditions and wisdom embodied in it, is lost forever.’

It is in this spirit that this study of TUP is presented. One of its goals is to contribute to typological research, to Tupían studies, and to human knowledge, since language is a fundamental aspect of the human species (see Everett 2012, 2017). According to Song (2018, 78), ‘10% of the world’s languages may have decent descriptions (read: adequate for typological research)’ – although this percentage may be higher now. This description of TUP will hopefully contribute to increasing this number.

The description of the language is carried out within the framework of Role and Reference Grammar (RRG) (see Section 3). The choice of the framework is straightforward. RRG captures and explains the interactions between syntax, semantics, and pragmatics through tools and principles which are motivated by typology and, in particular, the need to account for the cross-linguistic diversity exhibited by different grammatical systems. Thus, I consider it the best option available to describe these interactions, especially because Tupinambá has never been comprehensively described within a modern linguistic framework (see Section 1.5).

A grammar of TUP can play a significant role for historical linguistics in the diachronic study of Tupí-Guaraní (TG) languages. From the early texts in the sixteenth and seventeenth centuries, one can continuously follow its development through its descendants, the different forms of ‘Língua Geral’ (see Rodrigues 1996b), up to the modern Nheengatu, spoken nowadays by about

ten thousand individuals in northwestern Amazonia (Cruz 2011, 16-18). The study of TG languages has profited from Tupinambá and has contributed to our understanding of South America in many ways, including migrations, contact patterns, archaeology, plant ecology, and genetics (Noelli et al. 2018; Balee 2001; Castro e Silva et al. 2020; Silva et al. 2022; Castro e Silva et al. 2022; Ferraz Gerardi and Reichert 2021; Ferraz Gerardi et al. 2023). It has also been important for anthropology and for the study of other Tupí-Guaraní languages and cultures, as can be seen in works such as: Wagley and Galvão (1949); De Castro (1992); Ribeiro (1996); Cormier (2003).

This dissertation is organized as follows: in this chapter, an introduction to the Tupinambá people is given in Section 1.1, and their language is presented in Section 1.2. Section 1.3 briefly summarizes the typology of TUP. The main primary sources for the language are introduced in Section 1.5, and previous work on the language is discussed in Section (??). Chapter 2 discusses the phonology of TUP. The theoretical framework guiding the description, Role and Reference Grammar (RRG), is presented in Chapter 3. Chapter 4 discusses word classes, noun classes, relational markers and adverbs. In Chapter 5, basic clause patterns are presented followed by Chapter 6 on the layered structure of the clause in TUP. Lexical categories are presented in Chapter 7. The reference phrase is presented in Chapter 8. Chapter 9 deals with information structure, and Chapter 10 discusses complex sentence types. Finally, the last chapter provides some concluding remarks.

The orthography used in this study is based on the phonological inventory laid out for the language in Chapter 2 . I consider this to be an important issue because there is little agreement among published sources regarding how to write TUP, but all sources seem to prefer a Portuguese-based orthography. All the examples used throughout the work are from the original sources, thus avoiding the risk of presenting something which is not attested or which could not have existed. As for the translations of the examples into English, I have tried to keep the structure of the TUP constructions as close as I could to the original translations. This should account for the fact that some translations, although grammatically correct, may sound unnatural to native English speakers.

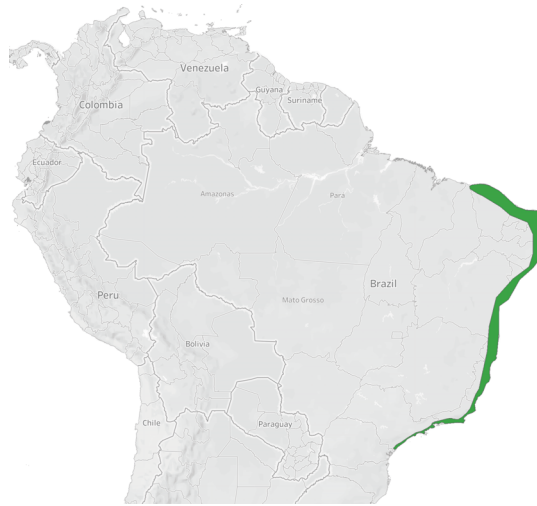


Figure 1.1: Map showing possible Tupinambá presence along the coast in 1500 represented by the green shade

1.1 The People

On April 22, 1500, the men from Cabral's fleet, on the Brazilian coast where nowadays lies the city of Porto Seguro in the state of Bahia, first encountered Brazilian indigenous people (Hemming 1978)¹. The language spoken by those people on the beach was unknown to the eight interpreters on board (see Bueno 2016, 36, 136), and it would reveal itself to be spread along the extensive Brazilian coastline, initially from the coast of the state of Ceará, in the northeast, to Cananeia in the south² (Métraux 1948, Anchieta (1933), Cardim 2009). The distribution of TUP groups along the coast (excluding the Guaraní groups) is shown in Figure 1.1.

Sixteenth and seventeenth century chroniclers registered the names of the

¹See the vivid description of this encounter in Bueno (2016).

²From Cananeia southwards, the Guaraní occupied the coast as far as Lagoas dos Patos and the Paraná-Paraguay basin. The language of the Guaraní, based on its first attestations, was certainly intelligible to the coastal Tupinambá speakers (see Anchieta 1997, xii-xiii, 78, 197, 210).

coastal groups, among which the following were found: *Ararape* (Cardim 2009, 197); *Uiatã* (Cardim 2009, 195); *Guaracaio* or *Itati* (Cardim 2009, 197); *Potiguara* on the coast between the Parnaíba and Paraíba rivers³; *Kaeté* on the coast between the mouth of the Paraíba and São Francisco; *Tupinambá* from the São Francisco river to Camamu or Ilhéus; *Tupinikin* from Camamu to the São Mateus or Cricaré river; *Temiminõ* in the southern state of Espírito Santo and on the lower Paraíba; *Tupinambá (Tamūja)* from Cape São Tomé to Angra dos Reis, but also in the hinterlands; *Tupinakin* from Angra dos Reis to Cananéia⁴; *Carijó* south of São Vicente on the coast and hinterlands until Paraguay; *Tupina* (see Métraux 1928a, 17-18) west of *Kaeté*; *Amoipira* on the left of the São Francisco in the hinterlands of Bahia (Sousa 1851, 44). There are also other groups whose locations cannot be precisely determined.⁵ In addition to these coastal tribes, different ethnic groups who spoke languages unrelated to TUP were found intermittently along the coast or at least not far from it, such as *Tobajar*, formerly at the Serra de Ibiapaba and which later migrated to the upper Mearim river in Maranhão (see Métraux 1928a, 16).

Whether the names of the groups indicate some kind of relations among them, e.g., that *Tamūja* ‘grandfather’ and *Temiminõ* ‘grandchild’ were so named because they were seen as early and later inhabitants of the coast, respectively, is but a conjecture. *Potiguara* ‘shrimp eater’, according to Edelweiss (1947, 33-55), is probably an epithet given by another group. The name *Tobajara* ‘enemy’ (see Staden 1557, chap. XIII) also means ‘brother-in-law’ (see Araújo 1618b, 116v and Anonymous 1952a, 87), and it was also applied to hostile groups who spoke unintelligible languages.⁶

The extensive territory occupied by TUP speaking groups given their linguistic⁷ and cultural similarity (Cardim 2009) is compatible with their pos-

³Métraux (1928a, 13) notes that the French chroniclers place the group they call *cannibals* in an area corresponding to that of the Potiguar.

⁴Edelweiss (1947, 39) notes that every time Anchieta mentions the *Tupinakin*, he places them in Porto Seguro and nearby.

⁵For the names of tribes and their locations, see Cardim (1881, 2009); Gândavo (1576); Sousa (1851); de Vasconcelos (1865).

⁶The meaning of the root *toβajar* ‘opposite, opponent, oppose’ is related to the meanings ‘enemy’ and ‘brother-in-law’ because it indicates ‘those from the other side’.

⁷In one letter of 1584, Anchieta, who had been to Bahia, Espírito Santo, Rio de Janeiro and

sibly recent arrival at the coast (Métraux 1928a, 12-19, Métraux 1927, Hemming 1978, 24, and Ferraz Gerardi et al. 2023)⁸. Upon arrival at the coast, Tupían speaking groups met with speakers of different ethnic groups known to have inhabited the coast (see Métraux 1948, 97): the *Guitaca* (see e.g. De Léry 1972, 354 and Anchieta 2006, 64) at the mouth of the Paraíba; the *Aimoré* between Espírito Santo and Bahia; and the *Tremembé* (see e.g. d'Evreux 2014, 178-180) between Ceará and Maranhão, among others. These non-Tupían groups were referred to as *Tapuja* (*tapi?ija* 'foreigner, barbarian') (see Anchieta 2006, 10,14,16 and Cardim 1881, 54-60; see also Ribeiro 2009).

Witness to other non-Tupían groups not far from the coast are the grammars or catechisms composed by religious missionaries: the Kiriri catechism published in 1698 (Mamiani della Rovere 1698), followed by a grammar of this language one year later (Mamiani della Rovere 1699). In 1709 the Capuchin Bernardo Nantes published his catechism (Bernardo et al. 1709) in the *Dzubukua* language, closely related to the *Kiriri* language described by Mamiani, *Kippea*. Other works have been lost, such as those by Manuel Viegas, published in the language of the *Maromomín* in 1585, the catechism by Father Manoel Nunes in the language of the *Nheengaiba*⁹, and the works of Bettendorff, who, besides a doctrine written in *Língua Geral*, composed catechisms in two other languages, *Tapajó* and *Urucuçu*, both now lost (Lee 2005, 141-143).

Returning to the Tupían groups on the coast, the linguistic and cultural similar-

São Paulo, wrote: *Desde o rio do Maranhão, que está além de Pernambuco para o Norte, até a terra dos Carijós, que se estende para o Sul desde a alagôa dos Patos até perto do rio que chamam de Martim Afonso, em que pode haver 800 léguas de costa, em todo o sertão dela que se estenderá com 200 ou 300 léguas, tirando o dos Carijós, que é muito maior e chega até ás serras do Peru, ha uma só língua.* (Anchieta 1933, 328) (From the Maranhão River, which is beyond Pernambuco to the North, to the land of Carijós, which extends to the South from the Patos floodplain to near the river they call Martim Afonso, where there may be 800 leagues of coastline, throughout the hinterland that will extend 200 or 300 leagues, except for Carijós, which is much larger and reaches the mountains of Peru, there is only one language.)

⁸Schmidt-Riese (1998) posits the hypothesis that the migration of Tupían groups towards the coast constitutes a suspension of the *equilibrium*, in the sense of Dixon et al. (1997), in which case the similarity of the variants along the coast would be a product of diversification and not of convergence.

⁹The term means 'bad speech' in Tupinambá, but it is not known which language it refers to.

ities among them based on the information handed down by the early sources suggest that these groups did not form discrete social units (Fausto 1998). This obscure scenario suggests the numerous disagreements regarding the ethnonyms and their locations. Were the Tupinikin¹⁰ the allies of the Portuguese in São Vicente (Cardim 2009, 197,274) or the allies of the French in Rio de Janeiro¹¹ (Thevet 1953, 296 and De Léry 1972, 29)? Staden (1557) asserts that in São Vicente the people called themselves *Tupinakiya* and were called *Tobajar* by their enemies, the *Tupinambá*. Anchieta, on the contrary, says that the natives from São Vicente were *Tupí* (see Edelweiss 1947, 44). Such disagreements abound among sixteenth and seventeenth century sources.

The coastal groups that spoke TUP in the sixteenth century had a sophisticated economy¹² and an almost amorphous social system, with nothing between the family and the tribe (De Castro 1992, 24)¹³. They lived in villages (*taβ*) consisting of a few large common rectangular houses (*ok*) whose sizes varied according to how many people lived in them. This type of house was common to all Tupí-Guaraní (TG) populations (Noelli 2022); the ceilings were made with palm leaves (*pinoβ*) (genus *Attalea*) or *patí* (*Syagrus botryophora* (Mart.) Mart. or *Syagrus romanzoffiana* (Cham.) Glassman), or with leaves of (*kapar*) (unknown species)¹⁴, on a wooden frame which lays on a column (*okita*) stuck in the ground. The spaces between these pillars were called *koti* ‘living room, corner, room’.¹⁵ On each side of the house, there was a door (*oken*). Each house hosted about thirty families. The building of the *ok^wam* ‘future house’ was the result of the *potirō* ‘collective effort’ (see Noelli 2022, 207). Three representations of TUP houses are given in Figure 1.2a. Villages

¹⁰Rodrigues (2010a, 27-28) says that the Tupinambá used ‘Tupinikin’ to refer the Tupí of São Vicente as well as those groups’ inhabitants of Espírito Santo and Southern Bahia.

¹¹For the names of some local groups in Rio de Janeiro, see Fernandes (1970, 60-61).

¹²As Fernandes (1949, 22) observes, the fundamental principle of the Tupinambá economy was the production of what was strictly necessary for immediate consumption. Accumulation of utilities for rationalizing was unknown to them.

¹³d’Abbeville (1614) writes that the chief of the Tupinambá ‘has no authority other than giving advice, especially when they are in their assembly or *carbet* which they hold every evening in the open space where their houses are.’ (apud MacCormack 1999, 120).

¹⁴See Sousa (1851, 222)

¹⁵Métraux (1928a, 47-48) and Fernandes (1963, 70) estimated, based on textual data, the *koti* as being 4×6.6 meters.

could eventually be surrounded by fences (*iβirá*), pointed platforms stuck in the ground, as shown in Figures 1.2a and 1.2.

The houses could last up to four years (Fernandes 1949, 35), and this was also the necessary amount of time for the soil of their slashes (*ko*) to be exhausted. Once the soil was exhausted and the roof could no longer contain the rain (*amān*), the group would migrate to a new area nearby (Staden 1557, 155). In order to be protected from incursions, the village could be surrounded by fences (*iβir*)¹⁶. Inside the house, there were no dividing walls, and each family occupied the space between two columns. Hammocks (*ini*) made mainly of cotton (*aminiju*) (Cardim 1881, 6-7) were to be found in the houses, along with other furniture pieces such as a wooden stool (*apikaβ, jur*), on which they kept their goods: gourds (*kuj*), pots (*kamusi, igasaβ*), baskets (*panaku*), sieves (*urupem*), weapons (*popeswar*), groceries, etc. Each family kept, in their division, a fire lit day and night (Cardim 1881, 9), which, during the night, protected against the cold and against mosquitoes (*jati?ũ*).

Common to all houses, the patio (*okar*) was the place where social and religious life evolved, and where rituals (*pepir*) and dances (*porasej*) were performed¹⁷ along with chants (*je?engar*). The villages were located preferably on hills, exposed to the wind, close to waterways and arable land, where they practiced horticulture. In their slashes (*ko*) they planted corn (*aβati*) and roots (*apo*), such as manioc (*mani?ok*), yam (*kara*) (Family: Dioscoreaceae), peanut (*manuβi*) (*Arachis hypogaea* L.), and sweet potato (*jetik*). They also cultivated pepper (*ki?ĩja*) and cashew (*akaju*), which was used to keep track of the months and year (d'Abbeville 1614, chap. 51). They used only two horticultural instruments, the digging stick (*sira*) and the stone axe (*ji*) (see d'Abbeville 1614, 226). They did practice exchange – even with non-Tupían groups (see De Léry 1972, 71 and d'Evreux 2014, 95, 184) – although not intensively.

¹⁶The villages protected by fences can be seen, e.g., in the drawings by Staden (1557) and van Groesen and Tise (2019).

¹⁷As told by d'Abbeville (1614) in a few passages, the dance had a special religious meaning for the Tupinambá, so that the Jesuits saw in it a sinful act (Anchieta 2006, 16, 32, 172, 193-194, 202, 204, 208). It had magic powers and was associated with the reincarnation of the shaman (*paje*) (d'Abbeville 1614, 209-209, 252-253).

The Tupinambá used many ornaments, such as hats with feathers (*akaŋaaβ*) of different colors, diadems (*akangatar*) made with red heron (*war*) or macaw (*kanine*) feathers, necklaces, bracelets, and leg ornaments (*poʔir*).¹⁸ Remarkable red heron feather cloaks (*waraaβusu*) were the most common feather ornaments. Mainly chiefs (*moruβisaβ*, *tuβisaβ*) and important men had shell necklaces, some up to nine meters long, which had to be wound several times around the neck; the women also wore long necklaces with loops that covered their chests. They also made bracelets with shells (*jaʔã*, *minõ*) or feathers (*awan*). All men, and only men, from the age of five or six, carried ‘tembetas’ (*temetar* or *metara*) of stones of different colors on their lower lip, especially those of green (*metaroβi*). The men also made a hole or two in the wings of their noses, into which they stuck long, thin pieces of wood or small white bones. Both men and women, but most commonly the latter, pierced the lobes of the ears, to introduce an ornament made of monkey bone (*namipaj*) or a wooden roller tangled with cotton thread. They removed all the hair on the body, including the eyelashes and eyebrows. Men shaved their hair from their foreheads up to their ears, using a bamboo (*kise*) or quartz knife (*itakise*). The women, however, wore their hair long and loose over their backs and, for work, tied it on top of their heads. Both men and women tattooed themselves, the latter only when they reached puberty, the former whenever they killed an enemy. They were also painted on all festive occasions, especially in black, with the juice of genipa (*janipaβ*), and in red, using the juice of annatto (*uruku*). Their conception of time was mainly based on the cycles of the moon (*jasì*), but apparently some stars (*jasitata*) or constellations also allowed them to keep track of time (Thevet 1953, 314). The name of a star (or constellation), *sejsu* [sej.ʃu], is mentioned in d’Abbeville (1614, chapter 51)¹⁹.

Hunting (*jeporakar*, *kaʔamomirõ*) and fishing (*ekij*, *pinaejtik*) were fundamental activities for the subsistence of the groups, and they were carried out mostly by men²⁰. Men chased deer (*swasu*), peccaries (*tajtetu*, *tajasu*), monkeys (*kaʔi*), agoutis (*akuti*), armadillos (*tatu*), and caymans (*jakare*). Jaguars

¹⁸ *Poʔir* actually means ‘bead(s)’ but also referred to necklaces and bracelets.

¹⁹ Also mentioned in Araújo (1618b, 17, 179v). See Lima and Moreira (2005).

²⁰ As Hemming (1978, 25) remarks, ‘[h]unting forced people to live in small groups, with enough men to hunt in packs, but never too many to exhaust an area’s fish or game.’

(*jawar*) and tapirs (*tapiʔir*) were caught in concealed pit falls (*mukuiʔiri*). Their hunting weapons were bows (*iβirapar*) and arrows (*uʔuβ*). As for fishing, they killed fish by poisoning calm waters with (*timo*) (*Dahlstedtia pinnata*) and (*tingi*) (*Magonia pubescens* A. St.Hil. or *Paullinia trigona* Vell.). They also used hooks made of thorns (*ju*, *juatĩ*) and fishing lines made with the fibers of *tukuma*, *tukũ* (*Bactris setosa*). This lines disappeared soon after European contact. They used canoes (*igar*) that could take up to thirty individuals (Anchieta 1933, 203). Other activities carried out only by the men included preparing the field for plantation, i.e., falling and burning, and building canoes (*igar*), bows (*iβirapar*) and arrows (*uʔuβ*), and clubs (*iβirapem*) (see Staden 1557, 177) and their adornments. Men also built houses, from cutting the wood with their stone axes (*ji*) to finishing the roofs, and were responsible for obtaining fire (*tata*).

The women maintained the slashes and collected roots, fruits, and cotton. They helped catch fish and oysters, and had to clean the canoes. An important task carried out by women, more precisely by pre-teenagers (*kujãtaiĩ*), was the preparation of an alcoholic fermented beverage (*kawiĩ*), and the fabrication of pottery (*ejaʔẽ*, *ejaʔẽpepo*, *kamu*). Only women took care of the house, cooked, kept the fire on, and made sure water was always at hand (see Fernandes 1949, 55-57). Women also prepared flour (*uʔi*) from different roots, as well as making manioc porridge (*miʔaʔu*). They roasted meat (*toʔo*) on the grill (*moka*) (see Léry 1578, 124-125 and Staden 1557, 15). As attested by written and visual sources, they had domestic animals (*mimaβ*) such as monkeys (*kaʔi*), armadillos (*tatu*), and parrots (*ajuru*).

War (*maran*, *warini*²¹) played an important role in the TUP society (see Fernandes 1970 and De Castro 2020), and it was intrinsically connected to the ancestor cult and to anthropophagic rituals.²² Exocannibalism, along with name bestowal and affinity (De Castro 1992, 155-163), are the main features that characterize TG peoples beyond their linguistic identity and behind their apparent morphosociological diversity (see De Castro 2020, 81-116). War was related to the preservation of the territory, demographic growth, and the conquest of

²¹From this word the name 'Guarani' is derived.

²²In fact these wars had nothing to do with territorial disputes. Their purpose was to maintain an infinite revenge cycle related to the killing of group members by an enemy group.

new territories in order to secure additional natural resources (Fernandes 1949, 43). The war was so important for a man that he dropped his childhood name only after having killed an enemy, after which he could marry, have legitimate children, and drink beer (*kawi*) (see De Castro 1992, 151). The drinking of beer *ka?u* was associated with leisure, celebrations, and singing *je?eŋ*, in opposition to the consumption of food, which was done in silence, as recorded by the early chroniclers (see De Castro 1992, 353 footnote 8).

The Tupinambá believed in the existence of supernatural entities²³ which inhabited the jungle. Many names are known from the extant texts: *Ajãŋ*²⁴, *Jurupari*, *Kurupir*, *Ma?etatá*, and *Wajupja*. The main supernatural entity was the spirit of thunder (*Tupãn*)²⁵, responsible for the rain, lightning (*tupãβeraβ*), and thunder (*tupãsunuŋ*). The connection between the physical world and that of the supernatural entities was made through the shamans (*pajé* or *karaiβ*), who possessed the knowledge of healing, either through spells (*jekaraimojãŋ*) or potions (*posaiŋ*).

After a few decades of peacefully trading brazilwood for metal tools, the rivalry between the Portuguese and the French spread among the indigenous coastal groups, with some groups taking the side of the French and other groups the side of the Portuguese. Subsequently, the enslavement of indigenous peoples allowed by royal patents (see Hemming 1978, 37-38) would cause many displeased natives to flee the Portuguese yoke. The arrival of the Jesuits in 1549 would give the natives more reasons to flee contact with the Europeans: the forced conversion and the various diseases (see Hemming 1978, 140-145). In his book published in 1576, Gândavo (2004, 32) writes that ‘there were many of these groups throughout the coast in the captaincies. They were everywhere when the Portuguese began to settle the land; but because these same Indians opposed the Portuguese, often betraying them, the governors and captains of the land slowly annihilated them, killing many of them. Others fled to the hinterland, leaving the coast free from natives throughout the captaincies’ (my translation).²⁶

²³For the religion of the Tupinambá, see Métraux (1928b).

²⁴This was the name given to the devil in the Jesuit texts.

²⁵This was the name that Jesuits used to refer to the Christian God in Tupí.

²⁶In the original: ‘Havia muitos destes índios pela Costa junto das capitánias, tudo enfim estava

It is not an exaggeration to assert that by the year 1700 the Tupinambá culture was already modified in its entirety, practically lost, perhaps except for some small groups that, having fled, found refuge in isolated areas away from the coast. Presently, there are about twenty-six thousand individuals who recognize themselves as Tupinambá, Tupinikin, or Potiguara²⁷. Since their ethnography lies beyond the scope of this short historical introduction, they will not be discussed.

1.2 The Language

Tupinambá belongs to the Tupí-Guaraní (TG) language family, a branch of Tupían established in 1958 (Rodrigues 1958a), though many of the internal relations now accepted were already known much earlier (see Hervás y Panduro 1805; von Martius 1867, Brinton 2009, 231-237). Recently, the classification has been refined and different proposals have been formulated thanks to the addition of more data and the inclusion of previously unstudied TG languages (Rodrigues 1985; Dietrich 1990b; Rodrigues and Cabral 2002; Ferraz Gerardi and Reichert 2021; Ferraz Gerardi et al. 2023).

Most of the sub-groups of Tupían are found in West Brazil, in the state of Rondônia (Galucio et al. 2015). Since this region contains the greatest diversity of sub-groups, it is considered to be the homeland of the family (see Rodrigues and Cabral 2012; Eriksen and Galucio 2014; Galucio et al. 2015). The largest sub-group, Tupí-Guaraní (TG), is also the most widely spread language family (sub-group) geographically in South America, with more than thirty living languages in Brazil, Peru, Argentina, Bolivia, French Guiana, and Paraguay. Based on the same criteria, its homeland is posited to be situated on the Xingu-Tocantins interfluve. As well as having the greatest diversity, it is also associated with a type of ceramics that may have spread from that area (see Ferraz Gerardi et al. 2023). Connections between many TG groups can

cheio deles quando começaram os portugueses a povoar a terra; mas porque os mesmos índios se levantaram contra eles e faziam-lhes muitas traições, os governadores e capitães da terra destruíram-nos pouco a pouco e mataram muitos deles, outros fugiram para o sertão, e assim ficou a Costa despovoada de gentio ao longo das capitánias’.

²⁷See <https://terrasindigenas.org.br>. Accessed on 01 September 2022.

also be inferred on cultural bases, such as the existence of agriculture and a strong tendency to sedentarism (Noelli 1996), a minimal repertoire of domesticates, and associated patterns of plant nomenclature and classification (Balée and Moore 1994). Additionally, parallels in the vocabulary of these languages are found in certain shared ritual and mythological complexes that are relevant to ethnozoology.

The Tupían groups inhabiting the coast spoke a single language, Tupinambá (see Sousa 1851, 57, Cardim 1881, 49,194-195) with minor differences in pronunciation and grammar (Anchieta 1595, 1v). These differences could hardly account for different languages, as suggested by e.g. Cabral (2011) (see Rodrigues 2011b).²⁸

While Anchieta (1595) acknowledges differences in pronunciation throughout the coast, Figueira (1687) is silent in this regard, perhaps because he was aware of the Jesuit's efforts to 'standardize and render closely-related Tupí-Guaraní speech forms considered the "general language of the coast" into one uniform language' (Lee 2005, 127) (see also Barros 2004). He is probably describing one already standardized and uniform language, for pedagogical and practical purposes (Zwartjes 2011, 165). His grammar was published about seventy years after Anchieta's draft had been sent for publication, when Tupinambá was already being used a lingua franca among the populations on and near the coast and in the Jesuitic mission (see Lee 2005, 50-51).

From the year of the 'discovery' (1500) to the end of the 1530s, a trade jargon (see Lee 2005, 46-49) began to develop. A prominent role was played by those foreigners who had become part of local native societies (see Léry 1578, chap. XVI), serving as mediators between the indigenous people and the Europeans. Since this scenario was common throughout the coast, there is no doubt that these men and women, whether exiled, survivors of shipwrecks, or sent by the crown, were crucial in the creation of a lingua franca (Lee 2005, 31-40).

²⁸It is worth mentioning that during the war between Paraguay and Brazil (1864-1870) speakers of Guaraní and Nheengatu could (still) understand each other (see Freire 2011, 102). This could hardly be the case if the southern and the northern language of the coast were two different languages three hundred and fifty years earlier.

Migration from Portugal intensified after 1530 in order to settle the land (Hemming 1978, 34-44). As soon as settlements were established, the practice of polygamy marked the alliances between the Portuguese and the TUP (Monteiro 2001, 34). Africans, Europeans, and natives of different tongues from the hinterland reached the settlements and thus the TUP spoken throughout the coast began to undergo changes for the sake of more efficient communication. Meanwhile, Portugal was sending orphans to the colony to learn the language and serve as interpreters (see de Almeida 1910a, 260 and de Almeida 1910b, II,280-281), and the French were doing the same (see de Léry 1957, 27). These are perhaps the bilinguals mentioned in Staden (1557, 55).

When the Jesuits arrived in Brazil for the first time in 1549, they chose Tupinambá as the common language (*língua geral*) (Altman 2003), and they founded mission villages through which, by the end of the century, they ‘controlled virtually all Indians under Portuguese rule’ in about thirty mission villages (Hemming 1978, 98,179). In these missions, children were separated from the adults in order to be instructed without the direct influence of their parents, and they spoke their native tongue, which was the language of the missions. At first, Jesuits used interpreters, the *línguas* (see Leite 1940; Barros 2002), which ‘were selected from colonists of Portuguese birth living in Brazil before the arrival of the Jesuits’ (Lee 2005, 132). With the arrival of Portuguese settlers (Gândavo 2004, 33) (see Rodrigues 2010a, 37, Schaden (1954), and Lee 2005, 156-162), mestizo populations were rapidly formed, and the bilingual children would later be used by the Jesuits. The bilingual generation instructed in a standardized language, the Jesuitic Tupinambá, in the missions (see Leite 1950, 40 and Barros 2004) shaped a kind of creole, the ‘Língua Geral’²⁹

²⁹*Língua geral* ‘general/common language’, as used by the Jesuits, first applied to languages that were spread through large communities in South America in the sixteenth century. This term must be distinct from ‘Língua Geral’ as the name of two language varieties that spread in the south (Língua Geral Paulista) and in the north of Brazil (Língua Geral Amazônica) as Tupinambá based creoles (Freire 2004, 93). For these varieties, see Rodrigues (1996b, 2010a), Schmidt-Riese (1998), and Muller et al. (2019, 19-22, 72-79). The term ‘Língua Geral’ was most probably associated with a standardized form used by the church with the intention of unifying the language of the villages established by missionaries. The term is used in Mexico, Peru, and Brazil (Pottier 1983, 21) with apparently the same meaning, that of lingua franca, i.e., the variety which had the highest degree of intelligibility, to refer to the coastal language in all its varieties (see Edelweiss 1947, 27-31). Nonetheless, in Brazil,

(see Muller et al. 2019, 19-22, 72-79). This language would later spread south and northwards. In the south, this language, *Língua Geral Paulista*, survived until the nineteenth century (see Rodrigues 1996a; Leite 2013) and the northern variety would reach Amazonia, with the name *Língua Geral Amazônica*, and become a *lingua franca* used by many ethnic groups (see Freire 2004), surviving to our days as *Nheengatu* (see Cruz 2011).

One strategy of the standardization process was the avoidance of grammatical patterns which did not have a parallel in Portuguese or which were perceived as complex. This is clear from the fact that some of the grammatical constructions presented in Anchieta (1595) and in Figueira (1687) are rarely found in other texts, even in other texts by Anchieta himself, such as object-incorporation with a stranded modifier. Moreover, in the prologue (*Aprovaçam*) to the grammar by (Figueira 1687), the editor writes that this description was less confusing than the previous one by Anchieta (1595) – not to mention the fact that Figueira (1687) seems to have abandoned some of the grammatical features present in Anchieta (1595).

The missions were a multilingual environment (see Cardim 2009, 288 and Lee 2005, chap. 2), where natives of different ethnic groups lived with TUP-speaking individuals. A complex language would be an obstacle to a faster learning process. Therefore, the language described by the Jesuits, far from being the real language as spoken by the natives, was ‘a simplified and poor idealization of it’ (de Freitas Leite 2005).³⁰ The variety developed in the missions which was used by non-native speakers, such as natives of other ethnic groups like Europeans and Africans, ended up establishing itself as an inter-ethnic means of communication, which, I suppose, made it quite different from the Tupinambá spoken before the arrival of the Europeans (Freire 2004, 66-81).

The foundation of the city of Belém, in 1616, brought the first settlers and missionaries to the Amazon. They had brought with them TUP-speaking in-

the ‘mestiçagem’ gives the whole thing a completely different character than in Hispano-America, where the ‘Lengua general’ never became the mother tongue of a large population of settlers and native women.

³⁰Against such opinions, others maintain that descriptions like Anchieta’s faithfully correspond to the language spoken as it was spoken by the natives Rodrigues et al. (1997).

dividuals from the coast, consequently putting them in contact with other TG languages, such as Guajajara (see Bettendorff 1698, 94, 303-307, 344). The language then occupied a larger part of the coast between the Pará (Tocantins) and Parnaíba rivers (and later also at the Pindaré, Mearim, and Itapicuru rivers) (see Sousa 1851; d'Abbeville 1614; Bettendorff 1698; Wagley and Galvão 1949; Gomes 2002; Wagley and Galvão (1949); Gomes (2002)). This *lingua franca* which was being used by different populations, especially in the missions, gradually diverged from the language of the coast, which was slowly disappearing, because natives from the coast increasingly fled from Europeans. By the end of the eighteenth century, Tupinambá was already an extinct language (Borges and Nunes 1998), and by the mid-eighteenth century, the language described by Anchieta (1595), Figueira (1687), and Araújo (1618b) was no longer understood by those in the north (Daniel 1975, 225). Bettendorff (1681) was written due to the difficulty the natives had with understanding the earlier written doctrines (see also Lee 2005, 185 and Freire 2004, 171), but it seems that the process of language change was not immediately captured by religious texts in use (see Daniel 1975, II,227; Barros 2003). The anonymous grammar of the *Língua Geral Amazônica*, a 1750 manuscript (Anonymous 1750), is an important piece of evidence attesting the changes that had taken place by the time of its publication. The language it describes differs in many aspects from the language described by Anchieta (1595) and Figueira (1687) (see Zwartjes 2011, 168-175).

Language manuals had already been circulating for some time before Anchieta concluded his grammar, which must have been before 1556, because Manoel da Nóbrega took one of its drafts to Salvador in this year (Schmidt-Riese 2016) (see also de Almeida 1910b, I,301). When it was published in 1595, it was the second grammar of a native American language to be published – the first was a grammar of Quechua by Domingo de Santo Tomás published in 1560. In 1574, a Christian doctrine had already been written in the language of the coast, by father Leonardo do Vale (see Barros 2004), but this text has been lost (Anchieta 1618a, 36). The forty years between Anchieta's final draft and its publication is certainly enough time for the 'standardization' process to show its initial effects beyond the natural changes that could have taken place in this time span. The same applies to the grammar by Figueira.

During the colonial period of Brazil (from the 16th to the 18th century), the name *Língua Brasileira*³¹ was often used in addition to ‘Língua Geral’ (see Zwartjes 2011). It was only in the nineteenth century that the name Tupí³² spread in Brazil, initially replacing the name *Língua Brasileira* in scholarly circles (Rodrigues 1958b, 5-6). The term Tupí originally referred to a group in São Vicente (see Anchieta 1595, 1v)³³

Among ethnologists, nowadays, Tupí is also the general name of the peoples considered to be related to the Tupinambá, and therefore designates some languages that are certainly related to the old Língua Geral, but not at all identical. As a result, the name Tupí has become an unsuitable term for describing either what is here called TUP or the variety of TUP spoken in the south. While the term ‘Tupí’ only started being used to refer to the language in the nineteenth century, the name Tupinambá appears in the eighteenth century, referring to the language of those Tupinambá from Pará, in order to distinguish this language from the language spoken there by the population of mixed origin, which was already different from it (Rodrigues 1986).

1.3 Typological Profile of TUP

Tupinambá has an average vowel inventory (see Maddieson 2013b) and a moderately small consonant inventory (see Maddieson 2013a). The syllabic structure is relatively simple (see Sec. 2.2). There is only one liquid phoneme /r/. There is a high central vowel /i/ and contrastive nasalization of vowels, both of which are distinct characteristics of lowland Amazonian languages (Aikhenvald 2012).

TUP is head-final and head-marking. This means that core arguments (A, S,

³¹The name *Língua Brasileira* appears in the titles of all works produced by the Jesuits up to the seventeenth century.

³²The name Tupí is of obscure origin. According to de Vasconcelos (1865, I, 109-110), it stems from a mythical character. This is a plausible etymology, since it appears to be related to the word for thunder in many Tupí-Guaraní languages (see the entry ‘thunder’ in Gerardi et al.), and it is associated with a divinity (Métraux 1928b).

³³In Anchieta (1595, 1v) ‘Tupí’ only refers to those natives south of the Tamoyos, in Rio de Janeiro, who called themselves Tupinambá. Hence the association of Tupí with the southern varieties and of Tupinambá with the northern varieties of speech (see Rodrigues 2010a).

and O) are expressed on the predicate with bound indexes in the SOV order. The alignment is exclusively nominative-accusative, contrary to what has been suggested previously, e.g., in Jensen (1998a, 565).

TUP is mildly agglutinating – not elaborately agglutinating as suggested by Rodrigues and Cabral (2012) – and combines suffixes and prefixes (see Table 1.1). It thus exhibits a weak degree of synthesis, although ‘the boundary between a synthetic, a highly synthetic, and a polysynthetic language is moot’ (Aikhenvald 2012, 129).

Some examples of agglutinating structures are given in (1):

- (1) a. Neporoamotareʔimawera
 ne=∅-poro-amotar-eʔim-wer-a
 2SG=R₁-ANTIP_{HUM}-care-PRIV-PST-REF
 ‘Your past disregard for people.’ (DC, II, 79)
- b. Eporapitiumẽ!
 e-poro-apiti-umẽ
 2SG.IMP-ANTIP-kill-NEG.IMP
 ‘Do not kill (people)!’ (DC, I, 143)

Polysynthesis is very reduced in TUP, being confined to cases of object incorporation, which when joined by an argument index form a whole sentence, as in (2):

- (2) a. Ojepoej
 o-je-po-ej
 3-RFLX-hand-wash
 ‘He washed his own hands.’ (Araújo, 61)
- b. Erejemoaipupukipe?
 ere-je-mo-aʔir-puk-puk-i=pe
 2SG-RFLX-CAUS-seed-bust-RED-EPEN=Q
 ‘Did you cause yourself to ejaculate?’ (DC, II, 90)

Gender and number are not categories of the TUP noun. However, TUP does have (nominal) tense, which is also a widespread feature in South America

(see Aikhenvald 2012, 59,159-162). In TUP, either a noun or a reference phrase can be specified for tense. TUP nouns are divided into possessed and non-possessed, with no morphological distinction between alienable and inalienable possession. The system of evidentials is very simple, with only one morpheme. TUP has no core cases, but it has locative(s), dative, translative, and perlative cases. Other oblique suffixes may be analyzed as cases. Arguments are bound to the predicate in S(O)V order. The past is the unmarked tense, while future is overtly marked. A dependent verbal form (gerund) appears in complex constructions with another, fully inflected verb. TUP has noun incorporation and two types of reduplication, mono- and disyllabic, each related to a different aspectual notion. Reduplication can also indicate plurality. There are many discourse particles. Demonstratives in TUP include the following categories: proximal- distal, and visible-invisible.

The first person plural distinguishes inclusive (1 person and 2 person) from exclusive (1 person and 3 person) and there is a generic index. A relational morpheme is used to mark the contiguity between a head and its dependent in many constructions.

Like many Amazonian languages (Aikhenvald 2012, 385), TUP lacks a possessive verb ‘have’, using verbless constructions instead.

Tupi-Guarani languages tend to exhibit more prefixes than do most western Amazonian families (Payne 1990), but the number of prefixes and suffixes in the language is nearly the same, as shown in Table 1.1. Derivational morphology is predominantly suffixal (see e.g., Dietrich 1990a).

Valency changing devices encompass causatives, one for transitive and another for intransitive verbs, and a sociative causative, antipassive, and incorporation. There is no passive derivation. Nominalization is very frequent in TUP and is used for complement and relative clauses.

1.4 Linguistic-Ethnographic remarks

There is not much that can be said about the sociolinguistics of Tupinambá, because no native texts of any kind have survived. There is, nonetheless, one

Feature	Prefix	Suffix	Section
Argument indexes	✓		4.3.1
Possessor indexes	✓		4.3.1
Oblique markers		✓	7.4
Causative intransitive	✓		5.7.1.1
Permissive	✓		6.5.1.1
Causative transitive		✓	5.7.1.2
Sociative causative intransitive	✓		5.7.1.3
Antipassive	✓		5.7.3
Reflexive	✓		5.7.4.1
Reciprocal	✓		5.7.5
Verbal Tense		✓	6.5.1.5
Nominal Tense		✓	7.3.1
Nominalizers	✓	✓	8.3
Gerund		✓	10.2.3

Table 1.1: Prefixes and suffixes in TUP

aspect that is attested by the Jesuit texts, which concerns male and female speech, a topic on which there are few cross-linguistic studies (see Aikhenvald 2016, chapter 9 and Aikhenvald 2012, 374-378).

An interesting sociolinguistic phenomenon is the fact that the gender of one speech act participant (speaker or hearer) determines the phonology, the lexicon, or the morphology of a language³⁴. This type of phenomenon, although it does occur on other continents, is more common in the Americas, especially in South-America³⁵ (Fortune and Fortune 1975; Borges 2004; Ribeiro 2006; Fleming 2012; Rose 2015a). Many Tupían languages have gender-specific lexical items. This chapter does not include cases relating the gender of the

³⁴Such a phenomenon was first described by Adam (1879). Cardim (2009) was the first to write that the language spoken on the coast had words only used by men and words only used by women.

³⁵The survey in Rose (2015a) found forty-one South-American languages belonging to thirteen stocks.

speaker and the gender of the ego of a kinship term, e.g., *pikiʔr* ‘older sister of woman’ and *endir* ‘sister of man’. The lack of attestation of the use of most of the elements presented in this chapter does not allow for any conjecture regarding the statistical indexing of gender, where some forms could be used by both women and men but are more commonly used by one or the other gender.

While cross-linguistically the locus of gender indexicality can be phonological, lexical, morphological, or pragmatic (discourse markers), TUP only has interjections/particles³⁶ and some kinship terms indexing gender. This is in line with the observation in Rose (2015a) that at the lexical level, distinctions tend to be limited to a few items. Although common, gender indexicality in discourse markers is restricted to some ten or less items. It is not known whether and to what extent gender indexicality can be reconstructed for Tupían or TG languages.

In Tupinambá, as far as it is attested, gender indexicality is limited to discourse markers and particles³⁷. Throughout this section, subscripts _f and _m mark female and male speech respectively.

The discourse particles *reʔĩ* and *reʔa* are used by women and men, respectively. Both express expectation.

- (3) a. *femenwera* *ipo reʔĩ*
 *f*e=∅-men-p^wer-a *ipo reʔĩ*
 1SG=R₂-husband-PST-REF ADV PRCL_F
 ‘It ought to be my ex-husband.’ (AT, 10)
- b. *Oso ipo reʔĩ*
 o-so *ipo reʔĩ*
 3-go ADV PRCL_F
 ‘It is expected that he/she/they/it went.’ (VLB, I, 102)

³⁶Rose (2015b) notes that the fuzzy boundary between discourse markers and other categories pose a problem for categorizing elements such as the particles/interjections presented in this section for Tupinambá.

³⁷Based on the definition proposed by Dingemanse (2021) I opt not to consider the elements in question as interjections. The items in question seem not to agree with one of the formal characteristics proposed, namely that they may function as stand-alone utterances.

- c. Oso ipo reʔa
 o-so ipo **reʔa**
 3-go ADV **PRCL_M**
 ‘It is expected that he/she/they/it went.’ (VLB, I, 102)
- d. Oimojaŋ ipo kori milagre amō maʔe
 o-i-mojaŋ ipo kori milagre amō maʔe-∅
 3-R₂-do certainly today miracle some thing-REF
 iaβaiβaʔe moaβaiβeʔima
 i-aβaiβ-βaʔe mo-aβaiβ-eʔim-a
 R₂-difficulty-NMLZ_{REL} CAUS-difficulty-PRIV-REF
 ʃeroβakéne reʔa
 ʃe=r-oβake=ne **reʔa**
 1SG=R₂-in.front.of=FUT **PRCL_M**
 ‘Hopefully he will do some miracle for me, discomplicating things.’
 (Araújo, 58v)

These particles, *reʔĩ* and *reʔa*, may also follow the negative *aani*:

- (4) a. Aani **reʔĩ**
 no **PRCL_F**
 ‘(It is) not like this.’ (FA, 127)
- b. Aani **reʔa**
 no **PRCL_M**
 ‘(It is) not like this.’ (FA, 127)

eʔĩ and *reʔa* are also used following the adverbs *serã* (see VLB, I, 87) and *apẽ* (see Figueira, 127).

The particles *ju / jo* and *we* are used by women and by men, respectively. They are also used following vocatives, as in 5:

- (5) a. ʃesi ju!
 ʃe=∅-si **ju**
 1SG=R₁-mother **PRCL_F**
 ‘Oh, my mother!’ (FA, 19)

- b. ʃeruβ we!
 ʃe=r-uβ we
 1SG=R₁-father PRCL_M
 ‘Oh, my father!’ (FA, 18)

See also other particles in (382).

As far as the lexicon is concerned, the vocative terms for sister, when used by a woman, are *kiʔĩ*, *kinaʔĩ* (VLB, II, 30; (AA, 14v), *nai* (AA, 14v), and *toʔi* (VLB, II, 30). The vocative terms for brother, as used by women, are *taj* and *tapiʔa* (VLB, II, 31). A vocative term for an older woman is *tape* (not given in the VLB) for women, and *tawpe* for men (AA, 14v). The latter is not said to be gender indexing in the VLB (II, 116).

1.5 Primary sources and previous work on Tupinambá

This section briefly introduces the sources containing written material in TUP. Primary sources are summarized in chronological order in Table 1.2.

The first significant attestation of words and sentences in TUP comes from the German gunner Hans Staden, in his *Wahrhaftige Historia* (Staden 1557)³⁸, in which he describes the nine months he spent as a prisoner of the Tupinambá during his second trip to Brazil. Hans Staden, besides recording aspects of Tupinambá culture, also wrote down sentences and words, some of which offer unique attestations.

The best known passage of his book is found in chapter XXIX (given below), where he describes the killing of the prisoner who will be eaten³⁹:

³⁸For an English translation of Staden’s book, see Staden (2008). For an overview of Staden’s account and drawings, see Duffy and Metcalf (2012).

³⁹While some authors have taken the various accounts of aspects of Tupinambá culture by the first sixteenth century authors as factual, due to their similarity, such as Métraux (1928a, 1948, 1979); Fernandes (1949, 1970), others have questioned this similarity, in particular the description of the anthropophagic ritual (see Ziebell-Wendt 1991).

Year	Author	Reference
1557	Staden	Staden (1557)
1558	Thevet	Thevet (1558)
1578	de Léry	Léry (1578); De Léry (1972)
15??	Anchieta	Anchieta (1988, 1618a,b)
1583	Anchieta	Martins (1941); Anchieta (1948, 2006)
1589-1594	Anchieta	Martins (1945); Anchieta (1997)
1595	Anchieta	Anchieta (1595)
1614	d'Abbeville	d'Abbeville 1614
1615	d'Evreux	d'Evreux (2014)
1616	Figueira	Figueira (1687)
1618	Araújo	Araújo (1618b)
1621	anonymous	Anonymous (1938, 1952a,b)
1645	Camarões, P. Poti	Navarro (2022)
1686	Araújo	Araújo (1618a)
1687	Bettendorff	Bettendorff (1681)

Table 1.2: Sources for the Tupinambá language

‘When all those [guests] who come from outside have now gathered together, the chief of the hut bids them welcome and says: Now come and help to eat your enemy. The day before they begin to drink, they tie the cord *Mussurana*⁴⁰ about the captives neck; on this day, they also paint the club called *Iwera Pemme* [*Ibira-pema*]⁴¹ with which they want to kill him.’ (Staden 2008, 396-397).

He provides not only important information on cultural aspects of the people, as in the quoted passage, but also many linguistic attestations, such as the following passage from chapter XX:

⁴⁰ *musu-rana* ‘rope-false’. See Métraux (1928a, 80-83) and Métraux (1979, 123-125).

⁴¹ *iβira-pema* ‘wood-angle’. Staden provides a drawing of this club.

Frenchman, which is sometimes interrupted by remarks of linguistic or moral character. Léry's register of TUP includes not only grammatical notes regarding pronouns and verbs, names of fauna and flora items, and cultural objects, but also examples taken from conversations as part of daily life.

During his years (1577-1587) as the head of the Jesuit order in Brazil, José de Anchieta wrote two theater plays: *Auto de São Lourenço* in Spanish, Portuguese, and Tupinambá, and *Na Aldeia de Guaraparim* in Tupinambá (Anchieta 2006). These were intended to be used as catechisms by the natives and the colonists. Both plays are wellsprings of valuable information on TUP culture. Anchieta wrote these plays in verses, with rhymes, as exemplified in (7) with each verse of the stanza glossed separately:

- (7) a. Ikawĩwasupipo?
 i-kawĩ-wasu=pe ipo
 R₂-beer-big=Q DEM
 'Does he in fact have a lot of beer (?)' (AT, 62: 698)
- b. Jeramũja Jawaruna?
 je=r-amũj-a Jawar-un-a?
 1SG=R₁-grandfather-REF jaguar-black-REF
 'My grandfather Black-Jaguar?' (AT, 62: 699)
- c. Eneĩ! Tasaβeipo!
 eneĩ t-a-s-aβeipor!
 INTJ HORT-1SG-get.drunk
 'Aha, may I get drunk!' (AT, 62: 700)
- d. Eri, awjete pako, ajewak
 eri awjete pako a-je-wak
 INTJ ADV PRCL 1SG-RFLX-embellish
 wijemouna
 wi-je-mo-un-a
 1SG_{CORF}-RFLX-CAUS-black-GER
 'Ah, I shall certainly adorn myself, painting myself black.' (AT, 62: 701-702)

(see, e.g., Gaffarel 1877; Cesar 2016).

In the above verses, the devil talks about beer consumed during the anthropogagic ritual of killing a prisoner, about the painting of ones body with *Genipa americana*, likewise on the occasion of the ritual killing of a prisoner, and even provides us with a proper name. While attestations like this are very rare, they permeate all the plays.

Anchieta's grammar (Anchieta 1595), published in 1595, clearly hints at a thorough reflection on how to present the content, since his description is concise and objective but nonetheless dense. He devotes fifteen pages to phonetics, provides detailed treatment of constructions and notes on the word order variation. Anchieta also wrote catechisms (Anchieta 1618a,b) and poems (Anchieta 1997).

Another Jesuitic text attesting the *Língua Brasileira* appeared in 1607. Some Christian prayers, occupying three and a half pages, appear in the *Rituale seu manuale peruanum* (de Oré 1607, 415-418). The prayers contained in the *Rituale* are already similar to those in Araújo (1618b), who had to reconsider the translations of some Christian concepts which turned out to be somewhat artificial – and apparently difficult to understand for the natives. Many of these were substituted by Portuguese words (compare the *Pater noster* version in Thevet and in Araújo given in Lee 2005, 136).

In 1614, the French Franciscan Claude d'Abbeville, who worked as a missionary with the Tupinamba in Maranhao, published *Histoire de la mission des pères Capucins en l'isle de Maragnan et terres circonvoisines* (d'Abbeville 1614). One year later, in 1615, the Franciscan Yves d'Evreux published his *Voyage au nord du Brésil (1615)* (d'Evreux 2014). Both were part of a French Catholic mission to the Tupinambá on the Maranhão island in the mouth of the Amazon, and each described different groups of the Tupinambá after having acquired some knowledge of their language. Léry's and Evreux's decriptions, along with those by Abbeville and Staden, are the most important ethnographic and lexical sources about the Tupinambá, especially their religion (see e.g. MacCormack 1999, 115-116).

The works of the Frenchman play an important role as they provide unique information about the culture of the Tupinambá, which often complement knowledge from Jesuitic texts. In Anchieta (1618b, 83), the priest asks a native if

(s)he believed in the dance of *Wajup'ia*:

- (8) Ereroβiáripe paje porapiti moʔaŋaʔuβa,
 Ere-eroβiár=pe paje-∅ poro-apiti moʔaŋaʔuβ-a
 2SG-believe=Q shaman-REF ANTIP-kill pretend-GER
 jekaraimojaŋa, moranjiwana pitanjeʔeŋa, Wajup'ia
 jekaraimojaŋ-a, morangiwan-a ∅-pitang-jeʔeŋ-a, Wajup'ia
 spell-∅ omen-REF R₁-child-speech-REF Wajup'ia
 moraseja, maraka poraseja, mosawsuβa?
 m.oraseja-∅, maraka-∅ porasej-a mosawsuβ-a
 R₃-dance-∅ rattle-REF dance-REF dream-REF
 ‘Do you believe in the shaman pretending to kill people, mysticism,
 children’s omen, Wajupia dance, rattle dance, and dreams?’ (DC II,
 83)

Nothing about this entity or place referred to as *Wajupja* is known through the Jesuits, in spite of Anchieta’s register. It is d’Abbeville (1614, 323) who talks about it as the place where the souls of the dead go after the death of the body. It is located beyond the mountains, inhabited by their ancestors⁴³. In d’Evreux (2014, 281), *Wajupja* are evil spirits or devils.

The catechism by the Jesuit Antonio de Araújo was published in 1618⁴⁴, before its publication was requested in 1592. It contains, in the initial pages, four short poems composed by another Jesuit, Antonio Valente (see Ayrosa 1941). Araújo’s *Catecismo* is precious because it contains the longest sentences known in the language. A new edition came out in 1686 which, as indicated in its prologue, was published because some of the vocabulary in

⁴³The same belief is found among the Tembé, Guarayo, and Apapocuva (see Métraux 1979, 35,51,54-55, 112). This Brazilian edition of Métraux’s essay is cited when notes accompanying the text are referred to.

⁴⁴In its original title: *Catecismo na Lingoa Brasilica, no qual se contem a svmma da doctrina christã. Com tudo o que pertence aos Mysterios de nossa sancta Fé & Bõs custumes. Composto a modo de Dialogos por Padres Doctos, & bons lingoas da Companhia de IESV. Agora nouamente concertado, ordenado, & acrescentado pello Padre Antonio d’Araújo Theologo, & lingoa da mesma Companhia.* The *imprimatur* of the catechism mentions that the text had been in use in the missions for about forty years

the 1618 edition had become obsolete, and because subtle changes in the doctrine had to be undertaken (see Araújo 1618a). It also contains a few pages listing kinship terminology (Araújo 1618b, 113v-117). Araújo's work can be seen as part of the standardization process of the language⁴⁵. Araújo's *Catecismo* is linguistically the most important text in TUP because of the many long sentences it contains.

In order to provide a more practical and straightforward description of the language that would allow for faster learning, since Anchieta's description was considered difficult, the *Arte da Língua Brasileira* by Luis Figueira was published in 1621, with a fourth edition in (Figueira 1687). The author 'sought out rural Indians and great missionary linguists born and raised among the Indians to consult' for his grammar (Lee 2005, 138). Figueira's *Arte* was intended to and probably did replace Anchieta's as a learning manual.

A manuscript from 1621⁴⁶ containing the *Vocabulário na língua brasileira* was first published in 1938 (Anonymous 1938) and subsequently in an augmented version in Anonymous (1952a). This text of unknown authorship is noteworthy not only due to its length, unique words, and many (short) sentences exemplifying their use, but also because, although it was copied in São Paulo, its lexemes are those of the northern variety.

The last texts in the *Língua Brasileira* is the *Compêndio da Doutrina Christã na Língua Portuguesa e Brasileira* by Bettendorff, published in 1681 (Bettendorff 1681). Although the language of these texts clearly underwent changes in comparison to the language of Anchieta – since it lacks some of the constructions and particles described in the work of Anchieta – they may be considered part of the corpus. The language of Bettendorff's catechism is morphologically and syntactically less complex than the language of previous texts. In this work, I do not discuss examples from Bettendorff (1681).

It is known that the natives also produced texts (see Lee 2005, 147-151), but only six letters written in 1645 by Tupinambás from Paraíba have survived.

⁴⁵Barros (2004) sees the standardization of Tupí in its religious context as a particular case of diglossia, where linguistic varieties acquire different uses, functions, and social values within a community, i.e., a linguistic variety becomes standard use in specific contexts.

⁴⁶For this manuscript, see de Paula Martins (1945, 1949).

Only two of these letters had been transcribed, translated, and published before 2022, see e.g. Sampaio (1906); Souto Maior (1912); Cerno and Obermeier (2013); Monserrat et al. (2020). In October 2022, all six letters were published, transcribed and translated in Navarro (2022).

If the goal of the standardization process was to ‘render closely-related Tupi-Guarani speech forms considered “the general language of the coast,” into one uniform language’ (Lee 2005, 127), it was a successful endeavor, as far as one can tell from the texts from Anchieta to Araújo.

The amount of work devoted exclusively to TUP is modest. Besides two grammars of pedagogical character, Barbosa (1956) and de Almeida Navarro (2004), there is no grammar adequate for typological research.

Barbosa’s grammar has many interesting insights, but one of its drawbacks is the many made-up, non-attested examples. The same problem is found in the grammar by de Almeida Navarro (2004). Neither grammar take the phonological system into account, confusing phones, phonemes, and graphemes. It is much more reasonable for a description of a dead language to contain exclusively attested examples. Barbosa (1956) is linguistically a better resource, especially in the treatment of syntax.

Aryon Rodrigues is the scholar who most prolifically wrote on Tupinambá. He published many articles describing aspects of Tupinambá grammar, mainly morphology, e.g., Rodrigues (1951a, 1952, 1953, 1996a, 1999a, 2004, 2009, 2010b).

Very important for the study of TUP are descriptions of other TG languages and comparative work within the TG family. While a comprehensive list of all the important contributions would be out of place here, one author, Wolf Dietrich, deserves special mention due to the quality and extent of his contributions, out of which the following deserve mention: Dietrich (1986, 1990b, 1994, 2000, 2001, 2006, 2009, 2014, 2017a,c,b, 2023).

For other TG languages, Table 1.3 shows what I consider to be the descriptions of TG languages (although the list is non-exhaustive) which are most important for the study of TUP, because they allow for a diachronic overview of aspects of TG languages.

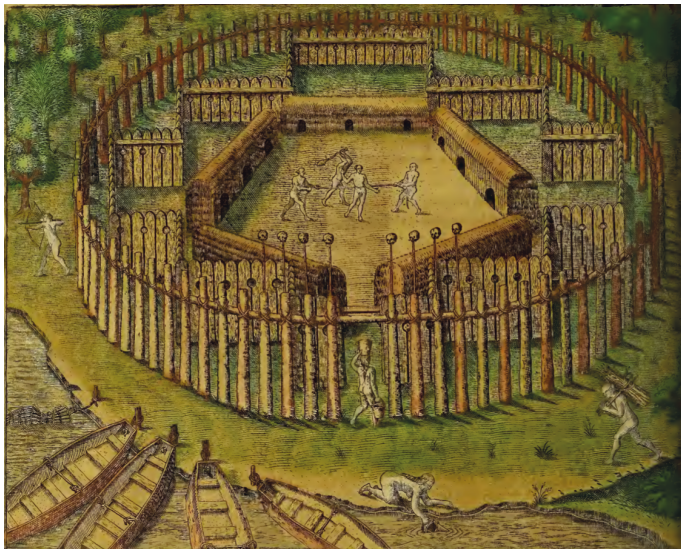
Language	Reference
Araweté	Solano (2010)
Avá Canoeiro	Borges et al. (2006), Silva (2020)
Chiriguano	Dietrich (1986)
Guajá	Magalhães (2010)
Guaraní	Estigarribia (2020)
Ka'apor	Kakumasu et al. (1986), Correa da Silva (1997)
Kamajurá	Seki (1990, 2000)
Kokama	Vallejos (2016)
Nheengatu	Cruz (2011)
Mbyá	Dooley (2015)
Old Guaraní	Montoya (1876); Restivo (1724)
Omagua	Michael and O'Hagan (2016)
Tapiete	González Vergara (2006)
Tekó	Rose (2011)
Wayampi	Grenand (1980); Jensen (1990b); Copin (2012)
Yuki	Villafañe (2004)

Table 1.3: Some references to descriptions of TG languages

Additionally, it is important to mention the chapter on Tupí-Guaraní by Jensen (1998a), and the more recent overview of Tupían languages in Rodrigues (1999b), Jensen (1999), and Rodrigues and Cabral (2012).



(a) Induction of the captive in a TUP village with four houses from Staden (1557)



38

(b) The killing of a prisoner in a Tupinambá village by de Bry (1592, 106)

Figure 1.2: Representations of TUP villages with houses, patios, and fences

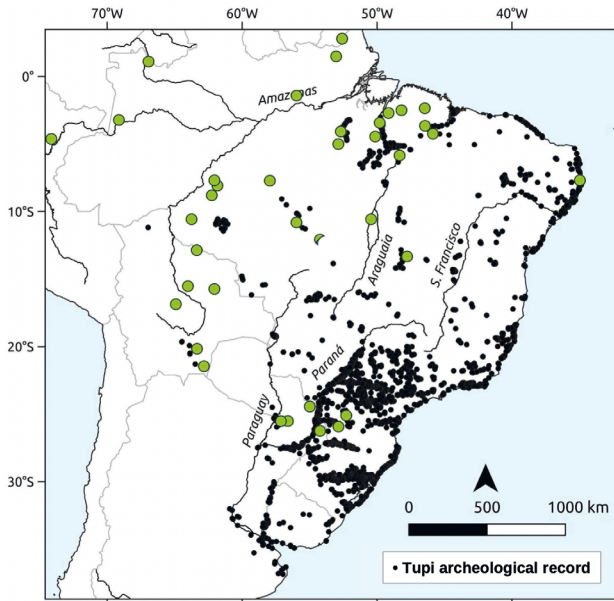


Figure 1.3: The Tupí-Guaraní languages (in green) along with the distribution of the TG archaeological record (black dots). TUP is represented by the dot on the coast (north-east). From Ferraz Gerardi et al. (2023)

Chapter 2

Phonology

Tupinambá phonology was first described by Rodrigues (1958b), who relied not only on the first sources, but also took into account the native languages of the authors of these sources in order to better grasp the phonology of TUP. The native language of an author is expected to reflect his/her orthography or transcription of TUP. By examining words as supplied by Portuguese, French, and German authors, Rodrigues was able to infer phenomena that would have remained otherwise unknown. Decades after Rodrigues' work, descriptions of other TG languages became available, which allowed for more solid tentative reconstructions of Proto-Tupi-Guaraní (Rodrigues 1958b; Jensen 1984, 1999; Rodrigues and Cabral 2012; Meira and Drude 2015) These reconstructions (Rodrigues and Dietrich 1997; Schleicher 1998; Mello et al. 2000; Meira and Drude 2015) and descriptions of other TG languages, have contributed to the understanding of TUP phonology (Jensen 1984).

Most of the TUP texts were written by Portuguese speakers with little variation among them, due to Jesuit standardization. If more variation were to be found, one would be able to inquire what these differences could reveal. While Figueira (1687) is almost tacit regarding the phonology of the language, Anchieta (1595) offers some information regarding variation in pronun-

ciation.¹ Araújo (1618b, 1-3) in the beginning of his catechism also provides some notes on pronunciation when presenting his orthography, but he is silent regarding variation.

Reconstructing the phonology of TUP is a difficult task since one must rely exclusively on documents from the fifteenth and sixteenth century, and on descriptions of other TG languages. An attempt at reconstructing TUP phonology is beyond the scope of this chapter and of this work, but nonetheless, I do think that the orthography here proposed, based on phonology, has advantages from a linguistic standpoint because it offers a more coherent view of the language's phonological system; but it will, at first, look uncommon to those who are familiar with TUP texts. One should not be surprised by eventual incoherencies. This chapter, beyond the presentation of the phonology, intends to facilitate the reading of the words and sentences in the pages that will follow.

2.1 Segmental phonology

The phonology of Tupinambá consists of fifteen consonants and six oral vowels with nasal counterparts, a small and an average size inventory respectively according to (Maddieson 2013a) and (Maddieson 2013b). Consonant and vowel phonemes are given in Tables 2.1 and 2.2:

The stress mostly falls on the last syllable with some exceptions described in Section 2.3. The basic syllable pattern is (C₁)V(C₂). The following sections introduce TUP phonology in a more detailed manner.

¹In this regard, Rodrigues writes: *Na exposição dos fenômenos fonéticos, detém-se Anchieta em quinze páginas, registrando, com meticulosidade rara mesmo em trabalhos modernos, mudanças e equivalências fonéticas, variantes individuais e diversidades dialetais.* (In the exposition of phonetic phenomena, Anchieta uses fifteen pages, recording, with rare precision, even in modern works, phonetic changes and equivalences, individual variants, and dialectal diversities.) (Rodrigues 1951b)

Consonants	Bilabial	Alveolar	Palatal	Velar	Glottal
Voiceless stops	p p ^w p ^j	t		k k ^w	ʔ
Fricatives	β	s			
Nasals	m	n		ŋ	
Flap		r			
Approximants	w		j		

Table 2.1: Tupinambá consonant phonemes

	Front		Central		Back	
	oral	nasal	oral	nasal	oral	nasal
High	i	ĩ	i	ĩ	u	ũ
Mid	e	ẽ			o	õ
Low			a	ã		

Table 2.2: Tupinambá vocalic phonemes

2.1.1 Consonants

There are fifteen consonant phonemes in TUP, as shown in Table 2.1. The glottal stop is included as a phoneme in this work, because minimal pairs seem to be attested, in spite of its absence in Rodrigues (1958b), who does not mention it at all, but includes it in many of his subsequent works (ex. Rodrigues 1996a, 2013).

The consonant [h] has not been included in the phonemic inventory. It is attested only in the items given in Table 2.3:

Since the phonemic consonants are established based on contrasts (Ladefoged and Maddieson 1996, 2), we proceed to present the most relevant oppositions for the consonant phonemes through minimal or near minimal pairs .

- (9) /p/ vs. /β/

Word	Word class	Example/Meaning	Source
haj	INTJ	‘Oh’ (expressing pity)	FA, 138
he		‘Eh...’	FA, 138; Léry, 344
hewe	INTJ (of man)	Eẽ hewe! ‘Oh, yes’	VLB, II, 117
hewi	INTJ (of woman)	Eẽ hewi! ‘Oh, yes’	VLB, II, 54, 117
hẽhẽ	particle	‘yes’	FA, 127

Table 2.3: Words with [h]

/pe/ ‘path’ /βe/ ‘also’
 /ipi/ ‘beginning’ /iβi/ ‘earth’

(10) /p/ vs. /m/

/pan/ ‘sideslip’ /man/ ‘bundle’
 /apo/ ‘root’ /amõ/ ‘someone’

(11) /p/ vs. /ʔ²/

/pok/ ‘pop’ /ʔok/ ‘cut’
 /ape/ ‘shell’ /aʔe/ ‘this’

(12) /p/ vs. /p^j/³

/pa/ ‘yes’ /p^ja/ ‘deviate’
 /epak/ ‘wake up! (imp)’ /ep^jak/ ‘see’

(13) /p/ vs. /p^w/⁴

/pan/ ‘wash’ /p^wã/ ‘finger’

²Minimal pairs with the glottal stop are not found abundantly (see 22), and are restricted to initial and medial position, since there is no evidence for the glottal stop in final position (see Jensen 1984, 53).

³/p^j/ is not abundantly attested. In word-initial position I have not found more than 10 items, including some compounds.

⁴/p^w/ is not abundantly attested.

- (14) /β/ vs. /m/
 /βoja/ ‘servant, disciple’ /moja/ ‘snake’
 /saβa/ ‘his feather(s)’ /sama/ ‘rope’
- (15) /β/ vs. /w/
 /βe/ ‘also’ /we/ ‘vocative marker’
 /aβa/ ‘person’ /awa/ ‘bumpiness’
- (16) /m/ vs. /w/
 /mã/ ‘ah (intj.)’ /wã/ ‘already’
 /aman/ ‘rain’ /awan/ ‘bracelet’
- (17) /t/ vs. /s/
 /tete/ ‘human body’ /sete/ ‘his body’
 /pita/ ‘heel’ /pisa/ ‘fishnet’
- (18) /t/ vs. /r/
 /ti/ ‘oh (interjection)’ /ri/ ‘because’
 /arara/ ‘macaw’ /atara/ ‘hiker’
- (19) /t/ vs. /n/
 /tupã/ ‘thunder’ /nupã/ ‘hit’
 /ini/ ‘hammock’ /iti/ ‘his nose’
- (20) /n/ vs. /r/
 /ne/ ‘future marker’ /re/ ‘after’
 /manaka/ ‘Brunfelsia hopeana Benth’ /maraka/ ‘rattle’
- (21) /k/ vs. /ŋ/
 /ʔaka/ ‘horn’ /ʔaŋa/ ‘soul’
 /puka/ ‘heavy’ /puŋa/ ‘swelling’
 /pak/ ‘awaken’ /mopapaŋ/ ‘do something slovenly’

(22) /k/ vs. /ʔ/

/kaβ/ ‘fat’ /ʔaβ/ ‘cut open’
/ake/ ‘this’ /aʔe/ ‘this’

(23) /k/ vs. /k^w/

/ka/ ‘break’ /k^wa/ ‘bay’
/kaβ/ ‘fat’ /k^waβ/ ‘pass’

(24) /n/ vs. /ŋ/

/men/ ‘husband’ /jeʔeŋ/ ‘speak’

(25) /m/ vs. /n/

/mã/ ‘oh’ /nã/ ‘thus’

(26) /w/ vs. /j/

/wara/ ‘eater’ /jara/ ‘owner’

2.1.2 Phonetic realizations and phonological processes

The phonemic consonants presented in Table 2.1 have different phonetic realizations with predictable distributions. These are summarized below:

/p/	→	[p], [ᵐb]
/pʲ/	→	[pʲ]
/pʷ/	→	[pʷ]
/t/	→	[t], [ᵐd]
/k/	→	[k], [ŋ]
/kʷ/	→	[kʷ]
/ʔ/	→	[ʔ]
/β/	→	[β], [p̣]
/s/	→	[s], [ʃ]
/m/	→	[m], [ᵐb]
/n/	→	[n], [ᵐd]
/ŋ/	→	[ŋ], [ᵑg]
/r/	→	[r], [ṛ]
/j/	→	[j], [ɲ], [dʒ]
/w/	→	[w]

In what follows, the most common distributions of consonant realizations are described along with phonological processes.

2.1.2.1 Stops

The voiceless bilabial stop /p/ only occurs in initial and medial position realized as [p] or [ᵐb]. Similar to /p/, /pʲ/, and /pʷ/ occur in syllable-initial position. The phoneme /t/ occurs in both initial and medial position. As for the voiceless velar stop /k/, it occurs in all positions, while the labio-velar /kʷ/ occurs only syllable-initially. The glottal stop seems to occur in root-initial and root-medial position, always followed by a stressed vowel.⁵

There are no voiced counterparts of the voiceless stops /p/, /pʲ/, /pʷ/, /t/, /k/, and /kʷ/. However, when preceded by a nasal vowel, the voiceless consonants /p/, /t/, and /k/ are nasalized whilst maintaining their place of articulation, as in Example (27).

⁵The primary sources do not graph or mention the glottal stop. Perhaps an indication of it can be found in Araújo (see Araújo 1618b, 3), who inconsistently marks the second vowel of a sequence with two dots, indicating that they do not form a diphthong, e.g. *mbaeü*, which could represent [ᵐbaʔe'ʔu] (see Araújo 1618b, 17v).

- (27) /p, t, k/ → [ᵐb, ᵐd, ŋ] / $\tilde{V}+_{-}$
- /jũ/ + /piteripe/ → [jũ.ᵐbi.te.ri.'pe] 'in the middle of the field'
 - /emi/ + /tipirõ/ → [ẽ.mi.ᵐdi.pi.'rõ] 'stew' (Arte, 13v)
 - /kujã/ + /katu/ → [ku.jã.ŋa.'tu] 'good woman' (Poemas, 86)

An exception to the above rule seems to occur, for instance in the examples shown in (28). This could be related to the orthography of the original sources, or to some phonological rule that can not be predicted from these sources.

- (28) a. /kunumĩ/ + /kaneʔõ/ → [ku.nu.mĩ.ka.ne.'ʔõ] 'tired boy'
- b. /kujã/ + /piatã/ → [ku.jã.pi.a.'tã] 'courageous woman'
- c. /kawĩ/ + /tata/ → [ka.wĩ.ta.'ta] 'strong spirit (beverage)'

2.1.2.2 Fricatives

The voiceless bilabial fricative /β/ is found in word-initial position, albeit infrequently. It is more common in root-medial position. The bilabial stops encountered in final position, [b] and [pʰ], are allophones of /β/, as shown in (29):

- (29) /β/ → [b] ~ [pʰ] / _#
- /seruβ/ → [ʃe.'rub] ~ [ʃe.'rupʰ] 'oh my father'
 - /kuwaβ/ → [ku.'wab] ~ [ku.'wapʰ] 'know'
 - /soβ/ → [sɔb] ~ [sɔpʰ] 'he has leaves'

The reason why [pʰ] is here taken to be an allophone of /β/ and not of /b/ can be explained as follows: Anchieta (1595, 4) says that [b] is pronounced as it is in Spanish, i.e., *aua* ([aβa]) instead of *aba* ([aba]). No other source provides information regarding the pronunciation, but the earliest sources do provide some clues regarding [p] and [b] in final position. French sources (De Léry 1972; d'Abbeville 1614) have [p] almost exclusively in final position. In Staden (1557), there is alternation of [b] and [p] in final position, occurring even in different attestations of the same word. This allophonic process is also suggested by Rodrigues (1958b) (see also Rodrigues 1999a).

The alveolar fricative [s] occurs in initial and medial position, e.g., ['swi] 'from', [si'rik] 'slide, leak out', [eβuru'su] 'big, large', [ja'suk] 'wash'. The allophone [ʃ] never follows [i] or [i], where its allophone [ʃ] occurs instead:

- (30) /s/ → [ʃ] / i, i _
- a. /isi/ → [i.'ʃi] 'his mother'.
 - b. /isupe/ → [i.'ʃu.'pe] 'in him'.
 - c. /papisuara/ → [pa.pi.'ʃwa.rə] 'what is on the wrist'.

I have not found minimal pairs of [s] and [ʃ], except for a near minimal pair:

- (31) /posij/ 'heavy' /po.ʃi/ 'ugly'

The post-alveolar fricative [ʃ] is found word-initially apparently only in onomatopoeic words or in words of foreign origin (e.g. ex.32a,b). Word initially it is found as reduced form of /ise/ 'I (e.g. ex.32c,d,e), or in suffixes that only occur followed by /i/, e.g., *swe* 'non-indicative future' and *so* 'non-indicative future'. This process where the alveolar becomes post-alveolar, only occurs before the deletion of the high vowel.

- (32) a. [ʃo.ro.'ro] 'tinamou (Family tinamidea)'
 b. [ʃe.ru.'ru] 'type of mollusk'
 c. [ʃe.'βe] 'me (dative)'
 d. [ʃe.'βo], 'me (dative)'
 e. [ʃe] 'my'

2.1.2.3 Nasals

The bilabial nasal /m/ occurs in initial, medial, and final position. It has two phonetic realizations [m] and [ᵐb]. The phone [ᵐb] is restricted to syllable-initial position and oral environments (33 and 34), while [m] is found in all other contexts (35). As illustrated by the examples below, nasality is triggered

by a nasal vowel or a nasal consonant regressively. In (34a), the [o] is nasalized due to the regressive spread by [ᵐb]. Otherwise, one would expect [ᵐboᵐbe'ʔu], which is not the case.

- (33) /m/ → [ᵐb] / #_V
- a. /míte/ → [ᵐbi.'te] 'still'
 - b. /mewe/ → [ᵐbe.'we] 'slow(ly)'
 - c. /maʔe/ → [ᵐba.'ʔe] 'thing'
- (34) /m/ → [ᵐb] / V_V
- a. /momeʔu/ → [mõ.ᵐbe.'ʔu] 'tell, accuse, confess'
 - b. /nami/ → [nã.ᵐbi] 'ear'
 - c. /pema/ → [pẽ.ᵐbə] 'angle'
 - d. /kami/ → [kã.ᵐbi] 'milk'
- (35) /m/ → [m] / \$_Ṽ
- a. /manõ/ → [mã.'nõ] 'die'⁶
 - b. /maʔẽ/ → [mã.'ʔẽ] 'look at'.
 - c. /memuã/ → [mẽ.'mũ.ã] 'evil act'.
 - d. /koʔẽma/ → [kõ.'ʔẽ.mã] 'morning'.

The alveolar nasal /n/ occurs in initial, medial, and final position. Parallel to /m/, it has two phonetic realizations [n] and [ᵐd]. The phone [ᵐd] is restricted to syllable initial position and oral environments (36) and (37), while [n] is found in all other contexts (38). As illustrated by the examples below, nasality is triggered by a nasal vowel or a nasal consonant regressively.

- (36) /n/ → [ᵐd] / #_V
- a. /ne/ → [ᵐde] 'your'
 - b. /na/ → [ᵐda] 'negative particle'

⁶*Manõ* 'die', like many other words, never appears with a nasalized [ã] in the sources. If the [a] is not nasalized one would expect [ᵐb], which never occurs.

c. /niβe/ → [ˈni.βe] ‘with’

(37) /n/ → [ˈd] / \$ _V

a. /eni/ → [e.ˈni] ‘flame’

b. /mena/ → [me.ˈna] ‘marry’

c. /anira/ → [a.ˈni.ra] ‘bat’

(38) /n/ → [n] / ˜V _

a. /marakanã/ → [ma.ra.kã.ˈnã] ‘bird sp.’

b. /aman/ → [a.ˈmã] ‘surround’

c. /marana/ → [ma.ˈrã.nə] ‘war’

The nasal velar /ŋ/ is realized as [ŋ] in syllable final position as in (39), and as [ŋg] syllable-initially, as in (40).⁷

(39) a. /kaŋwerĩ/ → [kãŋ.we.ˈrĩ] ‘bone (dim.)’

b. /aʔaŋaβa/ → [a.ʔãŋ.ˈa.βa] ‘image, model’

c. /piraŋ/ → [pi.ˈrãŋ] ‘red’

(40) a. /ajãŋa/ → [a.ˈɲã.ŋgə] ‘devil’

b. /posãŋa/ → [po.ˈsã.ŋgə] ‘medicine’

c. /jẽʔeŋaβa/ → [jẽ.ʔẽ.ˈŋga.βə] ‘stutterer’

2.1.2.4 Flaps

The flap /r/ occurs in initial, medial, and word-final position:

(41) a. [ra.ˈko] ‘actually’

b. [pi.ˈra] ‘fish’

c. [e.ˈkar] ‘seek’

⁷Rodrigues (1958b, 80) mentions that the realization of /ŋ/ as [ŋg] could also occur, but he acknowledges the impossibility of determining such realizations from the extant texts. See also Rodrigues (1974, 18-19).

In final position, /r/ seems to occur in free variation with [r] or [t], most often with the unreleased [t̚].

- (42) /r/ → [r] ~ [t̚] / _#
 a. /taʔir/ → [ra.ʔir] ~ [ta.ʔit̚] ‘daughter of men’
 b. /esaʔir/ → [e.sa.ʔir] ~ [e.sa.ʔit̚] ‘pupil’

2.1.2.5 Approximants

The palatal approximant /j/ occurs in initial, medial, and final position. The phone [j] occurs in oral environments and is in free variation with [dʒ] in syllable-initial position (43). When preceded or followed by a nasal syllable, /j/ is realized as [ɲ], as in (44).

- (43) /j/ → [j] ~ [dʒ] / \$_
 a. /ja/ → [ja] ~ [dʒa] ‘as’
 b. /akaju/ → [a.ka.ʔju] ~ [a.ka.ʔdʒu] ‘year’
- (44) /j/ → [ɲ] / _\$_[+nasal] or \$_[+nasal]_
 a. /jetɪŋ/ → [ɲɛ.ʔtɪŋ] ‘fly species’
 b. /jõ/ → [ɲõ] ‘only’
 c. /jeran/ → [ɲɛ.ʔrãŋ] ‘attack, get irritated’
 d. /atõja/ → [a.ʔtõ.ɲã] ‘touch’
 e. /manõja/ → [mã.ʔnõ.ɲã] ‘place’

The labial approximant /w/ occurs in word-initial, medial and final position, as shown below:

- (45) a. /waj/ → [waj] ‘tail’
 b. /kawĩ/ → [ka.ʔwĩ] ‘manioc beer’
 c. /jukaw/ → [ju.ʔkaw] ‘kill (non focal)’

2.1.2.6 Morphophonological processes

There is one process involving glottal stop deletion at the morpheme boundary (in composition or suffixation), which is independent of the phonetic characteristics of the final consonant of the first element, e.g.:

- (46) /ʔ/ → ∅ / +₋
- a. /pitaŋ/ + /-ʔi/ → [pi.taŋ.i] ‘baby (dim.)’ (Poemas, 86)
 - b. /petim/ + /ʔu/ → [pe.ti.^mbu] ‘to smoke’ (Teatro, 8)
 - c. /jukir/ + /ʔi/ → [ju.ki.^ri] ‘salt water’ (d’Abbeville, 306v)

Different consonants, with the exception of the glottal stop, may not co-occur across morpheme boundaries, in which case an epenthetic [i] or [i] is inserted.

- (47) ∅ → [i, i] / [+ consonant_i] +₋ [+ consonant_j]
- a. /ok/ + /-pe/ → [o.^ki.pe] ‘in the house’
 - b. /asep^jak/ + /≠ne/ → [a.se.p^ja.^ki.nẽ] ‘I will see it’
 - c. /ajaβaβ/ + /-swer/ → [a.ja.βa.^βi.swer] ‘I’m a runaway’ (VLB, II, 11)
 - d. /oker/ + /≠ne/ → [o.^ke.ri.ne] ‘he will sleep’
 - e. /ojeran/ + /-βaʔe/ → [o.ɲẽ.rã.ⁿi.βa.ʔe] ‘the one who attacks’
 - f. /oimõjãŋ/ + /≠pe/ → [oj.mo.^ɲã.ɲi.pe] ‘did he do it?’

There is consonant deletion at the morpheme boundary when both surface stop consonants share the same place and manner of articulation.

- (48) C_[stop] → ∅ / C_[stop_i] +₋ C_[stop_i]
- a. /taβ/⁸ + /-pe/ → [‘ta.pe] ‘in the village’
 - b. /ep^jak/ + /katu/ → [e.p^ja.ka.^tu] ‘see well, observe’

⁸See Section 2.1.2.1.

When followed or preceded by a nasal vowel at a morpheme boundary, [r] becomes [n].⁹

- (49) /r/ → [n] / $\tilde{V}+_-$ or $_- \tilde{V}+$
 a. /tupã/ + /-ramo/ → [tu.pã.'nã.mo] ‘as a God’
 b. /ero-/ + /jãñ/ → [e.nõ.'jãñ] ‘cause to run with oneself’

2.1.3 Vowels

Tupinambá has six oral and six nasal phonemic vowels, which are shown in Table 2.2. Minimal or near-minimal pairs are given below:

Oral vowels

- (50) /i/ vs. /e/
 /supi/ ‘truly’ /supe/ ‘to’
 /aʔi/ ‘mother (vocative)’ /aʔe/ ‘this’
- (51) /i/ vs. /o/
 /ipo/ ‘his hand’ /opo/ ‘his own hand’
 /pira/ ‘skin’ /pora/ ‘content’
- (52) /i/ vs. /u/
 /uʔi/ ‘spear’ /uʔu/ ‘cough’
 /mina/ ‘pointy thing’ /muna/ ‘spit’
- (53) /i/ vs. /i/
 /ita/ ‘stone’ /ita/ ‘scaffold’
 /piʔa/ ‘my son (voc.)’ /piʔa/ ‘liver’
- (54) /e/ vs. /a/
 /tete/ ‘body of a person’ /tata/ ‘fire’
 /meʔẽ/ ‘give’ /maʔẽ/ ‘look at’

⁹The phonological rule in (49) only applies to morpheme boundaries. However, it is not unreasonable to suppose that it might have occurred in syllable boundaries as well, as in /kurumĩ/ → [kũ.nũ.'mĩ] ‘boy’. The reason for this seems to be the fact that sonorants in TUP were affected by regressive nasality. The word *kunumĩ*, for example, is never attested with the flap. Its first attestation with the flap comes from 1739 (see Edelweiss 1969, 134-137).

- (55) /e/ vs. /o/
 /pe/ ‘bark’ /po/ ‘hand’
 /aʔe/ ‘actually’ /aʔo/ ‘revile’
- (56) /e/ vs. /u/
 /pe/ ‘bark’ /pu/ ‘noise’
 /en/ ‘spill’ /un/ ‘black’
- (57) /e/ vs. /i/
 /asem/ ‘I go out’ /asim/ ‘I slip’
 /ene/ ‘you’ /eni/ ‘saliva’
- (58) /o/ vs. /a/
 /soβa/ ‘his leaves’ /saβa/ ‘its feathers’
 /ʔuaβo/ ‘eating’ /ʔuaβa/ ‘eating place/instrument/occasion’
- (59) /o/ vs. /u/
 /soʔo/ ‘his flesh’ /suʔu/ ‘his bite’
 /oβ/ ‘leaf’ /uβ/ ‘father’
- (60) /i/ vs. /a/
 /pik/ ‘quit, press’ /pak/ ‘wake’
 /iti/ ‘dirt’ /ita/ ‘column’
- (61) /i/ vs. /o/
 /iβa/ ‘fruit’ /oβa/ ‘face’
 /pik/ ‘quit, press’ /pok/ ‘snap, crackle’
- (62) /i/ vs. /u/
 /pik/ ‘quit, press’ /puk/ ‘rift’
 /ʔi/ ‘water’ /ʔu/ ‘ingest’

Nasal vowels

- (63) /i/ vs. /ĩ/
 /aʔi/ ‘oh my brother (voc.)’ /aʔĩ/ ‘oh my mother (voc.)’
 /ʔi/ ‘be old’ /ʔĩ/ ‘frustrative’

- (64) /e/ vs. /ẽ/
 /ape/ ‘shell’ /apẽ/ ‘crook’
 /ʔe/ ‘say’ /ʔẽ/ ‘pour’
- (65) /a/ vs. /ã/
 /ita/ ‘stone’ /itã/ ‘mussel shell’
 /piʔa/ ‘oh my son (voc.)’ /piʔã/ ‘skin disease’
- (66) /o/ vs. /õ/
 /ro/ ‘blind’ /rõ/ ‘thus, then’
- (67) /i/ vs. /ĩ/
 /eʔij/ ‘crowd, multitude’ /eʔĩj/ ‘scratch’
- (68) /u/ vs. /ũ/
 /ju/ ‘thorn’ /jũ/ ‘field’

2.1.4 Phonetic realizations and morphophonological processes

All TUP vowels occur in initial, medial, and final position. They become nasalized in nasal environments, as below:

- (69) a. /anũ/ → [ã.'nũ] ‘bird (sp.)’
 b. /meʔeɲ/ → [mẽ.'ʔẽɲ] ‘give’
 c. /pitun/ → [pĩ.'tũn] ‘dark, night’
 d. /inĩ/ → [ĩ.'nĩ] ‘hammock’
 e. /emonanĩ/ → [ẽ.mõ.'nã.nĩ] ‘continuously’

2.2 Syllable structure

The syllable structure of Tupinambá is (C)V(C), with the following possible types: V, CV, VC, CVC. All consonants occur in the onset (C₁) except /ɲ/. In the coda, (C₂) which does not allow fricatives (/β/ is phonetically realized as [b] in coda position) and glottal consonants, although as far as the

glottal stop is concerned, this is no more than an assumption. The nucleus contains a single vowel and heavy syllables have a (C)VC structure (as in /tiŋ/ ‘white’).

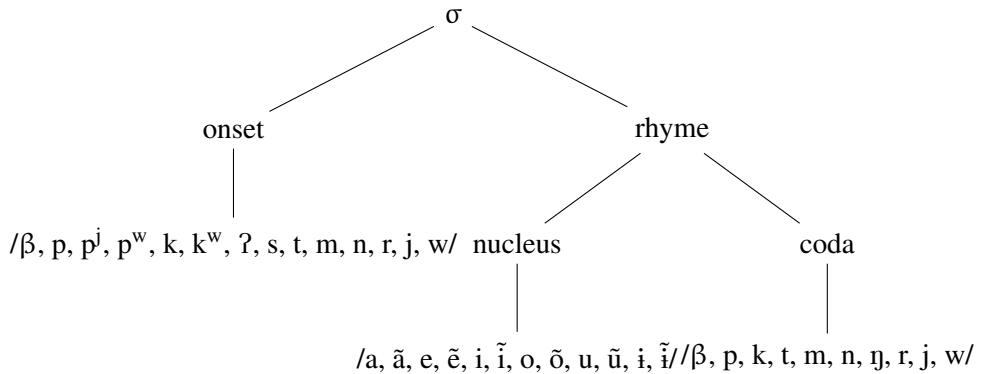


Figure 2.1: Syllable structure

- (70) V /ǎ/ ‘this’
 CV /si/ ‘mother’
 VC /ok/ ‘house’
 CVC /peβ/ ‘flat’

Sequences of two vowels are attested in TUP in all positions (initial, medial, and final). Due to the lack of native speaker intuitions, which have proven to be essential for the understanding of vocalic-like sequences within a syllable, it is impossible to predict vocalic sequences phonemically or to posit other rules by looking for phonetic cues. Another reason not to posit the presence of phonemic vocalic sequences is that whenever vowel clusters are found within a syllable, they are actually formed by the approximants /w/ or /j/ and a vowel (71) or the vowel is assigned to a different syllable (71c).

- (71) a. /ǎwa/ → [‘ǎ.wa] ‘that one.’ (VLB, I, 109)
 b. /awajmiri/ → [a.waj.mi.‘ri] ‘plant sp.’ (Piso, II, 175)
 c. /ae/ → [a.‘e] ‘emphatic particle’ (VLB, II, 36)

2.3 Prosody

The stress is fixed, occurring on the last syllable (72), but it might fall on the penultimate (73) or antepenultimate (74) under specific morphophonological conditions.

- (72) a. [ka.'ruk] 'urine'
 b. [ko.'ʔem] 'afternoon'
 c. [i.'ʃe] 'I'
 d. [ja.ka.'re] 'caiman'

As can be seen from the examples in (73), the stress falls on the penultimate syllable of derived words, either through suffixation (a,b) or through cliticization (c) and composition (d,e).

- (73) a. /juka/ + /-saβa/ → [ju.ka.'sa.βa] 'way of killing'
 b. /peʔa/ + /-pira/ → [pe.ʔa.'pi.ra] 'the one who has been removed'
 c. /a-/ + /so/ + /=ne/ → [a.'so.ne] 'I shall go'
 d. /tata/ + /eni/ → [ta.ta.e.'ni] 'flame'
 e. /ajuru/ + /juβ/ → [a.ju.ru.'juβ] 'blond (lit. yellow parrot)'

Some examples of stress on the anti-penultimate syllable are given in (74).

- (74) a. /mono/ + /reme/ → [mo.'ndo.re.me] 'send (irrealis)'
 b. /si/ + /-ramo/ → [ʃi.ra.mo] 'as a mother'
 c. /owata/ + /βaʔe/ → [o.wa.'ta.βa.ʔe] 'the one who walks'

Contrastive stress can only be seen in morphologically derived words (75).

- (75) a. /aβ/ + /-a/ → [ʃa.βa] 'feather'
 b. /aβa/ → [a'βa] 'man'

Chapter 3

RRG

3.1 Introduction

This chapter offers an introductory, non-exhaustive overview of Role and Reference Grammar (RRG). Its sole purpose is to outline the theory in order to facilitate the reading of the subsequent chapters, thus making them more accessible. Some aspects of RRG presented in this chapter will receive more attention in the subsequent chapters.

RRG appeared in the 1970s as a framework for describing language structure departing from languages other than English, such as Lakhota, Tagalog, and Dyirbal¹. It attempts to provide a model of syntax applicable to all languages, accounting for the variation of typological parameters, such as the presence or absence of verb phrases (see Van Valin Jr 2005, 80-88), syntactic or morphological expression of predicate-argument structure, grammatical relations (see Van Valin Jr 2005, 89-94), and serial verb constructions or chaining con-

¹Haspelmath (2008) criticizes RRG on the basis the same argument used by RRG to criticize the Chomskyan framework, saying that *Lakhotacentric or Tagalocentric frameworks are in no way better than Anglocentric frameworks*. However, as Jackendoff (2002, 75) writes, RRG has developed a syntactic machinery *more explicitly designed to speak to the varieties of syntactic phenomena in the languages of the world*.

structions. RRG attempts to capture this diversity and explain the interaction of syntax, semantics, and pragmatics in different grammatical systems (Van Valin Jr and LaPolla 1997, 14-15, Van Valin Jr 2005, 1).² In line with functional approaches to language study (Butler 2003b,a; Dik 1991; Givón 2001; Van Valin Jr 1991; Newmeyer 1991), RRG does not consider language to be an infinite set of structural descriptions, but views it as a system of human communication (Foley and Van Valin Jr 1984, 7, Van Valin Jr and LaPolla 1997, 11-15), thus attempting to characterize not only syntactic but also communicative competence. This characterization requires an analysis of the interaction between morphosyntactic form and communicative function (see Foley and Van Valin Jr 1984, 11-16). The emphasis in RRG has predominantly been on the exploration of language systems themselves from the perspective of their use in communication³, but as Bohmeyer and Van Valin Jr (2017, 144) observe, the typology of form-to-meaning mapping has also always been a theme in RRG research (see also Van Valin Jr 2009). Based on the classification of linguistic theories in Jackendoff (2002), RRG would fall into the *parallel architecture theory* type, in which syntax, semantics, and pragmatics are represented independently but interact directly with one another (see Jackendoff and Audring 2020, 5-9). RRG also takes a constructional approach, rejecting the principles-and-parameters approach (see Van Valin Jr 2022).

While syntax is said to constitute a system in the structuralist⁴ sense (see Van Valin Jr 1993, 2), the autonomy of syntax is rejected since form is assumed to be motivated by function (see Van Valin Jr 2003). The fundamental role of function requires grammatical structure to be understood and explained with reference to its semantic and communicative functions (pragmatics). Syntax is not arbitrary, but *relatively motivated by semantic, pragmatic and cognitive concerns* (Van Valin Jr 1991, 9). These assumptions place RRG in the middle of a functionalist continuum between Systemic Functional Grammar

²The main descriptions of the theory are found in Foley and Van Valin Jr (1984); Van Valin Jr (1991, 1993); Van Valin Jr and LaPolla (1997); Van Valin Jr (2001b, 2005, 2008a, 2022)

³“Languages are *systems* and not random collections of grammatical constructions. When explored from the perspective of how they achieve a certain communicative end, we see their systematic nature most clearly” (Foley and Van Valin Jr 1984, 374).

⁴For a description of Structuralism and its background, see Goldsmith and Laks (2019).

(Halliday 1985; Halliday and Matthiessen 2013)⁵ and Cognitive Grammar (Langacker 1987)⁶ (see Butler 2003b).

RRG looks at language structure from four perspectives: the surface forms, the underlying semantic structure, the modifying grammatical elements, and the pragmatic information structure. It offers mechanisms for discovering and describing how they interact with one another without referring to constituent structure or abstract levels such as Logical Form⁷ (see Van Valin Jr and LaPolla 1997, 317-319). This interaction is accounted for by typologically motivated principles which are captured by three main representations: syntactic (form of the utterance), semantic (meaning of the utterance), and information or focus structure (pragmatics). The linking algorithm, the core component of RRG, connects the syntactic and semantic representations with pragmatics playing a role in the linking process. A visual representation of the linking algorithm is given in Figure (3.1) (Van Valin Jr 2005, 131).

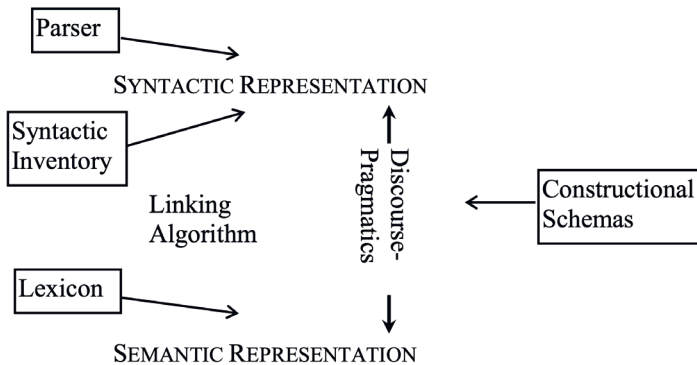


Figure 3.1: The architecture of RRG

⁵SFG takes a radical discourse-pragmatic view, starting with discourse and working ‘down’ to lower levels of grammatical structure.

⁶Langacker recognizes only the semantic, phonological and symbolic aspects of linguistic structure, and rejects the distinction between semantics and pragmatics, which he considers to be artificial and arbitrary (see Langacker 1990).

⁷This is one important characteristic of RRG that is shared with Constructional Grammar (see e.g. Goldberg 2003, 219).

The linking between syntax and semantics is bi-directional (see Van Valin Jr and LaPolla 1997, chap.7, Van Valin Jr 2005, chap. 5) and models the production process, starting with the formulation of a message, mapping it onto the appropriate morphosyntactic form, and uttering it. It also models the comprehension process, with an analysis of the utterance followed by mapping it onto a representation of its meaning⁸. Linking semantics to syntax begins in the lexicon, where a semantic representation is built. It takes a sentence as input, applies a syntactic parser and represents the morphosyntactic properties of this sentence by the layered structure of the clause (LSC) (see Sections 3.1.1 and 3.1.2). Explicit mapping rules link the syntactic representation to the semantic representation, providing an interpretation. Simple examples of the linking in both directions are explained below for a simple sentence based on (76)^{9, 10}.

- (76) Ojuka
 o-i-juka
 3-R₂-kill
 ‘She kills him.’

The semantics-to-syntax linking consists of five steps. Step one builds a semantic representation of the utterance from the logical structures associated with the lexical units (see Section 3.3). In step two, macroroles are assigned to the arguments of the predicator (see Section 3.3.2). In the third step, the morphosyntactic encoding of the arguments is determined (Privileged Syntactic Argument (PSA), case markers, adpositions, agreement). In step four, the syntactic templates are selected from the syntactic inventory. Finally, in

⁸The modeling of the communicative process, what the hearer and the speaker do, makes RRG suitable for computational implementation (see Kallmeyer et al. 2013 and Nolan 2004) and also applicable in psycho- and neurolinguistic studies (see Van Valin Jr 2006a) and cognitive science (see Van Valin Jr and LaPolla 1997, 640-649).

⁹The glossing of this example will become clear in the next chapters. For now, it suffices to say that ‘3’ refers to the subject, and ‘R₂’ refers to the object.

¹⁰The explanation of the linking algorithm provided here is a superficial one. For more details see: Van Valin Jr and LaPolla (1997, chap.7), Van Valin Jr (2005, chap. 5), Van Valin Jr (2006b), and Osswald and Kallmeyer (2018).

step five, the arguments are anchored to their position in the syntactic representation.

In the syntax-to-semantic linking, the parser outputs a syntactic structure (step one)¹¹, and morphosyntactic forms (verb-forms, voice, etc.) are retrieved, the PSA is determined (step two). In the active voice, the actor is the PSA (the core initial reference phrase (RP)) *o-*. The logical structure (LS) of the predicate is retrieved from the lexicon and macroroles are assigned. The last step establishes that *o-* is the actor and *i-* the undergoer.

Language-specific features, i.e., grammatical constructions, are captured in RRG by *constructional schemas* (Van Valin Jr 2005, 131-132)¹². But because constructional schemas reference general principles, they not only capture cross-linguistic generalizations, but at the same time express language-particular properties of grammars (see Jiménez-Briones and Luzondo-Oyón 2013). Constructional schemas contain morphological, syntactic, semantic, and pragmatic information, some as general principles and some as language-specific constraints¹³

Table 3.1 below provides a constructional schema for TUP WH-questions.

The template for the pre-core slot (PrCS) mentioned in Table 3.1 is given in Figure 3.2, and an example of a WH-question is given in (77), with its syntactic representation given in Figure 3.3.

- (77) Maʔepe pesekar?
 maʔe=pe pe-s-ekar
 thing=Q 2PL-R₂-seek
 ‘What do you seek?’ (Araújo, 54)

Constructional schemas contain language-specific morphosyntactic features as well as semantic and pragmatic information of the given construction (see

¹¹See Van Valin Jr and Usón (2014, footnote 2).

¹²These were called *constructional templates* in Van Valin Jr and LaPolla (1997).

¹³For some examples of constructional schemas, see: Van Valin Jr and LaPolla (1997, 433-436), Van Valin Jr (2005, 132-134), Jiménez-Briones and Luzondo-Oyón (2013), González Vergara (2006), Nolan (2013), Diedrichsen (2010).

Construction	Tupinambá WH-question
	Template(s): PrCS
Syntax:	PSA: None Linking: WH-XP to PrCS
Morphology:	Default
Semantics:	Contains an open proposition with a variable α , WH-XP = α
Pragmatics:	Illocutionary force: interrogative Focus structure: narrow focus on PrCS

Table 3.1: Constructional schema for TUP WH-questions

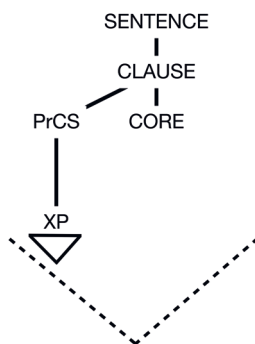


Figure 3.2: Syntactic templates for a TUP WH-question. The dashed line indicates the focus domain and the triangle marks the narrow focus

examples in Van Valin Jr 2005, 131-135, 148, 258, 265, 267). From Figure (3.1), it can be seen that discourse-pragmatics runs parallel to the linking algorithm, indicating the interactive role of discourse-pragmatics in both directions of the linking. The lexicon interacts with the semantic representation before the linking process takes place.

More recently, a series of extensions have been proposed for RRG (see e.g. Kailuweit 2018). Among these is a formalization of the theory, which can

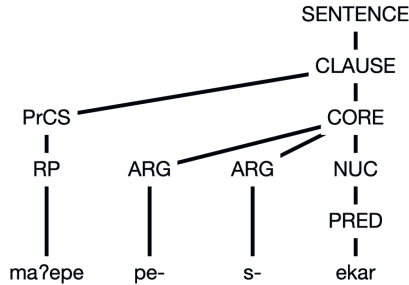


Figure 3.3: WH-question with precore slot

serve as a basis for computational implementations (Osswald and Kallmeyer 2018). A volume which is expected to be published in 2022 by Cambridge University Press will also contain many novelties regarding the theory and its applications.

The next sections introduce the main tools of RRG for investigating and describing the structure of a language. The syntactic representation is presented in Section 3.1.1, followed by the semantic representation in 3.3 and information structure in Section 3.3.

3.1.1 The syntactic representation: the layered structure of the clause (LSC)

RRG represents the hierarchical organization of sentences and clauses; that is, non-relational aspects of clause structure are represented through semantically motivated syntactic units, as shown in Table 3.2. The LSC is based on contrasts known to be found in all languages, namely the contrasts between expressions of semantic predicates, their arguments, and the modifiers of their projection (Van Valin Jr 1990, 1993; Van Valin Jr and LaPolla 1997). These contrasts are a consequence of the nature of language as a system of communication, which requires predication and reference in order to talk about states of affairs in the world. The LSC also accounts for aspects of clause structure common to languages of different types, such as word order, flat syntax, and head marking (see Van Valin Jr and LaPolla 1997, 22). These principles are

typologically grounded¹⁴, thus accounting for cross-linguistic diversity. Divergent features should be represented by comparable structures whilst pointing out the differences and similarities between them.

Semantic element(s)	Syntactic unit
Predicate	Nucleus
Predicate + arguments	Core
Predicate, arguments and non-arguments	Clause

Table 3.2: Semantic units underlying the syntactic units of the layered structure of the clause

The syntactic organization in RRG is linear and layered, without underlying derivations or multiple representations. Its internal structure consists of the following layers: *sentence*, *clause*, *core*, and *nucleus*. The layers in this organization are represented as labeled trees, as displayed in Figure (3.4). As Osswald and Kallmeyer (2018, 359) observe, “trees provide the most natural way to analyze syntactic structures since they build on the basic relations of immediate dominance and linear precedence.”

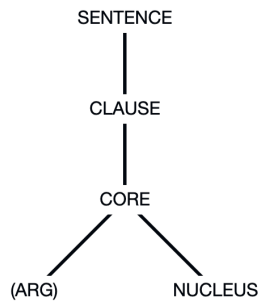


Figure 3.4: Simple constituent projection

The LSC has three independent but unified projections: the constituent, the

¹⁴The typological adequacy to which RRG subscribes was laid out in Dik (1991). See Van Valin Jr and LaPolla (1997, 14-15).

operator, and the focus projection. The constituent projection represents the syntactic categories, the operator projection represents grammatical categories called operators, and the focus projection represents the information structure. These are briefly discussed in the next sections.

3.1.2 The constituent projection

The nucleus (NUC), Core, and periphery (PERI) are the primary components of the LSC, which are semantically motivated and universal. The nucleus is the predicate, and since syntactic categories in RRG are non-endocentric (see Chapter 4), it is not connected to any particular lexical category (see Croft 2022a) and its head can be phrasal (see Everett 2008) or even an RP; the core contains the nucleus and its arguments, while the peripheries host non-arguments. The core periphery hosts, for example, non-argument adjuncts and temporal and locative modifiers. The core and its periphery make up the clause. All known human languages make a distinction between the core and the periphery, just like all languages distinguish between predicating and non-predicating elements as well as between arguments and non-arguments (adjuncts) (Van Valin Jr and LaPolla 1997; Van Valin Jr 2005). An LSC with core and clausal periphery is represented in Figure 3.5. The peripheries are, following Osswald and Kallmeyer (2018), not separated nodes, as in e.g. Van Valin Jr and LaPolla (1997); Van Valin Jr (2005), but marked by a feature [PERI+].

The non-universal components of the constituent projection are the pre-detached position (PrDP) and post-detached position (PoDP), and the PrCS and the post-core slot (PoCS), all of which are pragmatically motivated. The detached positions are for dislocated constituents, which are normally but not always separated from the main clause by an intonation break, or represented by a resumptive pronoun in the core. Detached units are outside the clause but within the sentence node. The precore and postcore slots are inside the clause. Unlike detached elements, units in the precore or postcore are not intonationally separated from the clause, and they are not represented by resumptive arguments in the core. They usually host focal elements. Example 78 shows an element in the PrDP.

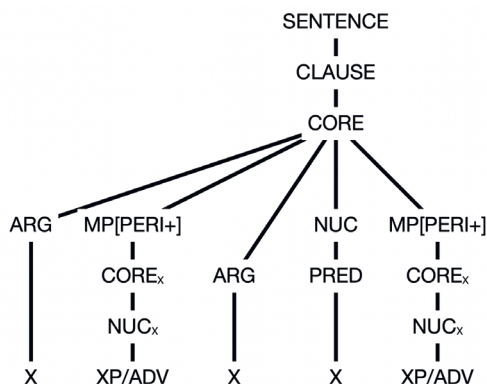


Figure 3.5: LSC with periphery at the core-level

(78) [[My sister]_{PRDP}, [[I have not seen her]_{CORE}]_{CLAUSE}]_{SENTENCE}

In languages like Tupinambá and English, with *ex situ* WH-questions, the PrCS becomes the position for the WH-words, as in (79) and its representation in 3.6.

(79) Ma?etepe peseka ko feretama pupe?
 ma?e-te=pe pe-s-eka ko fe=r-etam-a Ø-pupe
 WH-FOC=Q 2SG-R₂-seek here 1SG=R₁-country-REF R₁-POSP
 ‘What do you seek here in my land?’ (Teatro, 30)

The PoCS is found in languages in which WH-questions occur in a post-core position, e.g. in SOV languages such as Japanese. WH- and non-WH constituents can also occur in the PrCS and PoCS as focused or displaced constituents.

All the layers and constituents described above are summarized in Fig. 3.7

As will be shown in Chapter 10, the central components of the LSC also turn out to be the three fundamental building blocks of complex sentences in language, because the construction of complex sentences involves the linking of

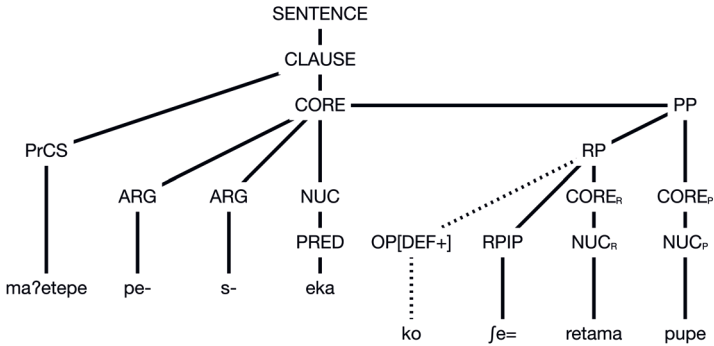


Figure 3.6: PrCS with WH-word

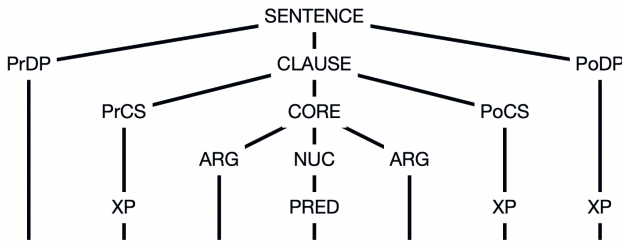


Figure 3.7: The Layered Structure of the Clause

these units: nucleus with nucleus, core with core, clause with clause, or sentence with sentence.

For head-marking languages¹⁵, there is an ECS (Van Valin Jr 2013) where independent lexical nominals, such as reference phrases coreferential with a bound argument index in the core, are located. In dependent-marking languages such as English, lexical nouns are the core arguments of the core, but

¹⁵The term ‘head-marking’ has a slightly different meaning in RRG than originally proposed in Nichols (1986), according to whom any head-dependent relation can be morphologically encoded on the dependent, the head, both, or neither. This definition requires one to consider, e.g., subject-verb agreement in English as head-marked. In RRG, the term implies, as in Bloomfield (1935), a head to which one or more bound morphemes are attached, filling the head’s semantic argument positions.

in head-marking languages, core arguments may or may not be coreferential with an overt lexical RP. This is illustrated in (80), which is represented in Figure 3.8. The arguments of the predicate are *o-* and *-s-*. The RPs, *Pedro* and *sword*, are semantically related to the arguments but are not arguments themselves. Subscripts indicating coreference help to identify the arguments to which the RPs are semantically related (see Section 6.3).

- (80) Pedro itangapema osekij
 Pedro_i i-itangapem-a_j o_i-s_j-ekij
 Pedro R₂-sword-REF 3-R₂-pull
 ‘Pedro pulled the sword.’ (see Araújo, 54v)

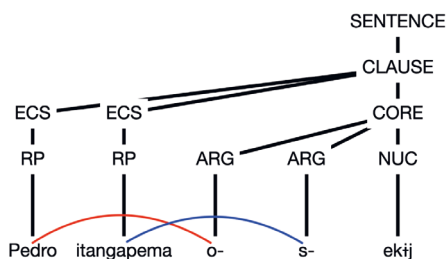


Figure 3.8: ECS with RPs coreferential with the bound argument indexes in the core

Van Valin Jr (2013) highlights important differences between the ECS and the PrCS (see also the discussion in (Kihara 2017, 61-66)).

The building blocks of the syntactic representations are the syntactic templates, which are stored in the syntactic inventory. Figure (3.9) shows some examples of syntactic templates.

3.1.3 The operator projection

The operator projection hosts grammatical categories such as aspect, negation, tense, directionality, event quantification, status, tense, evidential, and illocutionary force. Initially, the constituent and operator projections were homomorphic mirror images of each other (see Figure 3.10). This projection is

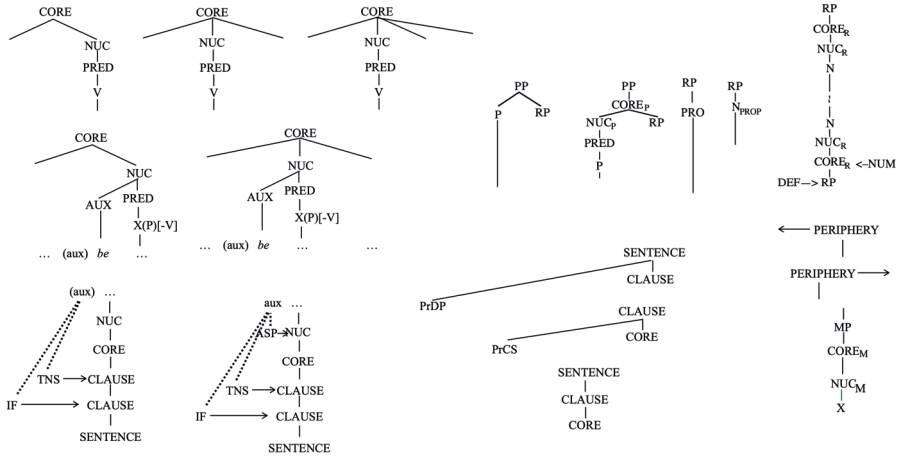


Figure 3.9: Some syntactic templates for English from Van Valin Jr (2022, 42)

necessary because operators are subject to different ordering constraints from the predicates, arguments, and adjunct modifiers (see Foley and Van Valin Jr 1985, 233, Van Valin Jr and LaPolla 1997, 46-52 and Van Valin Jr 2005, 9-11). Furthermore, the operator projection permits an accurate expression of the scope of operators in complex sentences (see Van Valin Jr 2005, 213-219).

All the layers, with the exception of the sentential layer, have their own operators, and each layer may be modified by more than one operator at a time. Operators play an important role in the determination of clause linkage types, since clauses that share an operator constitute a special type of clause linkage specific to RRG called cosubordination (see Chapter 10).

The scopes of an operator is an individual layer. The nucleus, for instance, is modified by aspect and directionals. Nuclear operators do not involve participants since they are outside of the nucleus. The core operators are negation, event quantification, and deontic modality. Epistemic modality, instead, is subsumed under status, a clausal operator together with the related notions of realis-irrealis. Tense modifies the clausal layer as well as evidentials, which are used to indicate the source of the information expressed by a speaker in

a proposition. Some operators, such as negation, may be found in more than one layer. The combined schema for the constituent and operator projections is given in Fig. 3.10 and summarized in Table 3.3.

Nuclear operators:	Aspect
	Negation
Core operators:	Directionals (only those modifying orientation of action or event without reference to participants)
	Directionals (only those expressing orientation or motion of one participant with reference to other participant or the speaker)
	Event quantification
Clausal operators:	Modality (root modals, e.g., ability, permission, obligation)
	Internal (narrow negation) negation
	Status (epistemic modals, external negation)
Clausal operators:	Tense
	Evidentials
	Illocutionary force

Table 3.3: Some of the most common operators and their respective levels in the LSC

Figure 3.10 shows the syntactic representation and the operator projection, whose separate representation was proposed by Johnson (1987).

In the formalization of RRG proposed by Osswald and Kallmeyer (2018, 360), the operator projection “assumes a single syntactic structure in which operator components are distinguished by the feature [OP]”, as exemplified in Figure (3.11), adapted from Van Valin Jr (2005, 14). This approach to marking operators is adopted throughout this work, with dashed lines connecting the operators to the nodes.

3.1.4 The focus projection

The distribution of information in sentences has morphosyntactic consequences. The information or focus structure interacts with the other projections (con-

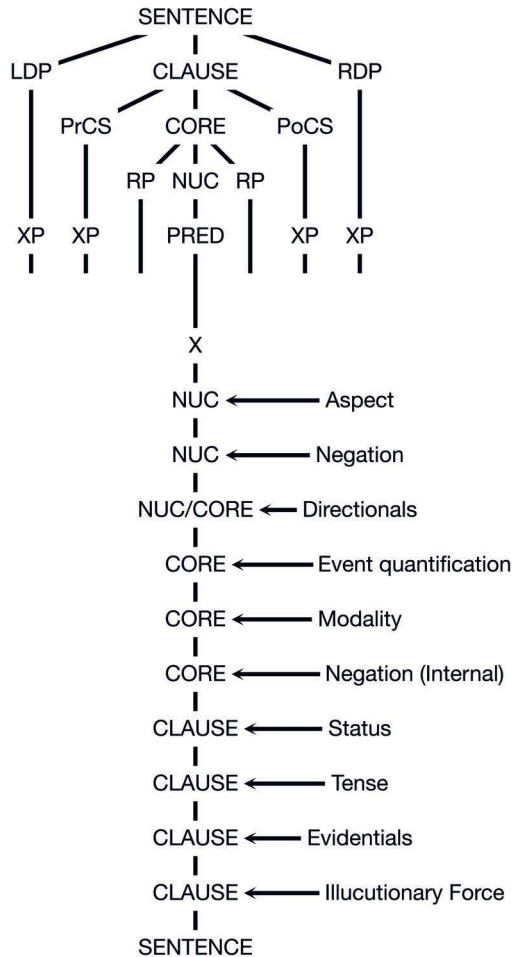


Figure 3.10: Constituent and operator projections

stituent and operator) in the articulation of statements. The components of the focus projection are the information units (IUs), which correspond to the amount of information contained in a simple WH-expression (Lambrecht 1986). The basis of the proposition of the focal structure is found in the notions of “pragmatic presupposition” and “pragmatic assertion”, expounded on

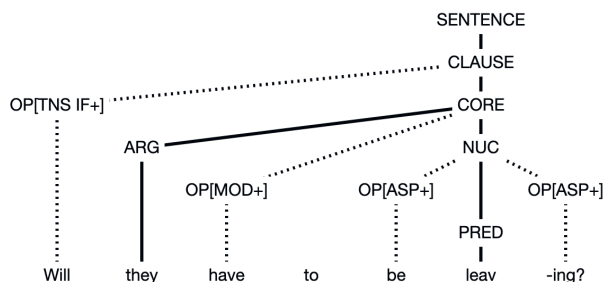


Figure 3.11: Alternative operator projection

in Lambrecht (1994, 52):

PRAGMATIC PRESUPPOSITION: The set of propositions lexico-grammatically evoked in a sentence which the speaker assumes the hearer already knows or is ready to take for granted at the time the sentence is uttered.

PRAGMATIC ASSERTION: The proposition expressed by a sentence which the hearer is expected to know or take for granted as a result of hearing the sentence uttered.

From the notions of pragmatic presupposition and pragmatic assertion, the concepts of “topic” and “focus” are derived. The former corresponds to the information which is interpreted pragmatically as something (the referent of the proposition) assumed or already known; it is therefore part of the pragmatic presupposition. The latter is the part of the propositional content that establishes the difference between the pragmatic presupposition and the pragmatic assertion, i.e., the asserted part of a declarative statement or what is asked in an interrogative statement. The information contained in the presupposition and how it relates to the new information (the focus) is the basis of the informative structure of the sentence. RRG calls the grammatical system that conventionally associates the distribution of information with a given sentence form, and which indicates the scope of the assertion in a sentence in a way that contrasts with the pragmatic presupposition, the “focal structure”. Two main sections can be distinguished in the focal structure: (a) the potential

focus domain (PFD), which corresponds to the syntactic domain in which the sentence focus can be found; and (b) the actual focus domain (AFD), which indicates the section of the statement that is effectively focused.

The marking of focus depends on the type of speech act; that is, it depends on the illocutionary force (IF) operator, because the potential focus domain must fall within the IF operator. The focus domain can be any of the units in the core or the peripheral PPs. The focus structure projection must represent both the potential focus domain and the actual focus domain. The speech act node, which is related to the illocutionary force operator, anchors the focus structure projection, and the potential and actual focus domains are represented within its scope. This information is captured in the RRG representation of the focal structure, as can be seen in Figure 3.12, where the two types of focus domain are outlined: potential (dashed line) and actual (triangle). The nodes called IU make up the basic units, while the “speech act” node, which is directly related to the IF operator, is the anchor point of the focal structure projection (see Van Valin Jr 1999b, 2014).

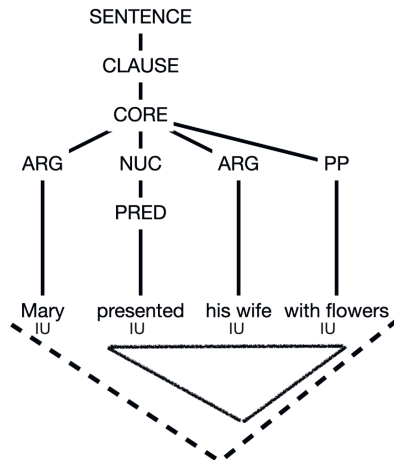


Figure 3.12: The focus projection

RRG classifies focal structures according to the following categories:

PREDICATE FOCUS STRUCTURE is found in sentences that express a constituent in a topic position, and in which the predicative phrase provides new information about it. It is the unmarked type of focus, and in it, the real focus domain is the predicative phrase. In the statement that expresses the answer in (81a), for example, the potential domain of focus considers the entire sentence, the topic corresponds to the noun phrase “my bicycle”, and the actual focus domain is equivalent to the comment “it is broken”.

- (81) a. What happened to your bicycle?
 b. [_{PFD}[My bicycle_{topic} [_{AFD}broke down_{focus}]]]

SENTENCE FOCUS. These structures do not have a topic. In them, both the potential and real focal domains are equivalent to the entire sentence. These sentences are generally used to introduce new referents in the discourse. In example (82a), it can be seen how all of the sentence constituents are in the domain of real focus and the sentence does not have a topic.

- (82) a. My bicycle broke down
 b. [_{PFD}[My bicycle [_{AFD}broke down]]]

NARROW FOCUS. In sentences with this type of structure, the real domain of focus corresponds to a single constituent. In the answer in (83b), for example, while the potential focus domain encompasses the entire sentence, the actual focus domain is concentrated in the constituent “MY CAR” (capital letters denote contrastive intonation).

- (83) a. I heard your skateboard broke?
 b. [_{AFD} My CAR [_{PFD}broke]]]

The potential focus domain in the simple sentence seems to correspond to the clause, so any constituent found in the nucleus, the core, or the peripheries can be focalized, while the elements that appear in the detached positions are topical by default and are therefore outside the potential domain of focus (see Sections 6.1 and 6.2).

The RRG theory of information structure was adapted from Lambrecht (1986, 1987, 1994, 2000), whose theory of information structure posits three types of foci: narrow, predicate, and sentence focus. These types indicate the focused constituents in a proposition. Lambrecht's theory was further enhanced in RRG with the introduction of the concepts of PFD potential focus domain and AFD. The former refers to the possible domain which can be focused, and the latter to the specific position of a focused element. RRG also adapted the Discourse Representation Theory of (Von Stechow 1999) in order to formally represent the interaction of presuppositions and assertions (Van Valin Jr 2005). Based on (O'Connor 2008), a representation of prosody has also been proposed.

3.2 The Layered Structure of the Reference Phrase

Similarities in the structure of the RP and the LSC become especially pronounced when RPs are complex derived nominals (see Van Valin Jr 2005, 24-30). The primary correspondences between RPs and clauses are between their layered structures and their operator projections. In the layered structure of the RP, there is a nominal nucleus (NUC_R) and a nominal core ($CORE_R$) consisting of the nucleus and the arguments of a complex derived nominal, e.g. 3.15, but there is only one level corresponding to the clause and sentence levels. The layers of the RP may also be modified by peripheral elements.

The layered structure of reference phrase (LSRP) contains a reference phrase initial position (RPIP), a daughter of the RP node. This position is occupied by a variety of elements, depending on the language: WH-words, demonstratives, possessive pronouns, articles, or possessor phrases (Van Valin Jr 2005, 26). The RPIP is a core-initial position that subsumes the functions of the PrDP and the PrCS in the LSC. This follows from the fact that unlike sentential units with four layers, a complex RP has only three.

The nuclear periphery is occupied by adjunct restrictive modifiers such as adjectives, nominal modifiers, and restrictive relative clauses. The core periphery is occupied by adjunct PPs and adverbials, and the RP periphery is occupied by non-modifiers such as non-restrictive relative clauses. Figure (3.13)

shows an example of an RPIP for the English RP *this book*:

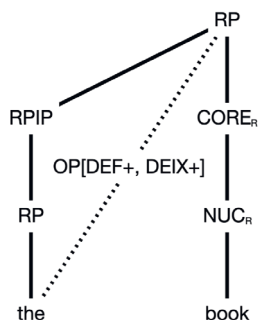


Figure 3.13: RPIP with definiteness and deixis operators

Table 3.4 shows the operators of the RP at each level. Nominal aspect, a nuclear (NUC_R) operator, involves the count-mass distinction in parallel with the telic/atelic distinction in verbs (see Jackendoff 1992, 29), as well as distinguishing whether the referent is an individual, part of an individual, or a set of individuals. Core ($CORE_R$) operators are about quantity and negation. Quantification is expressed through the grammatical category of number as well as lexical expressions like numerals and quantifiers. Negation may be expressed through a special negative form for RPs, such as *no* in English, special determiners which interact with sentential negation, such as English *any* as in *Mary didn't buy any books*, and nouns and pronouns with an inherently negative meaning, such as German *nichts*, Czech *nic*, and French *rien* 'nothing'. RP-level operators, encoding definiteness and deixis, are analogous to the function of some of the clause-level operators, such as tense and illocutionary force. They are primarily concerned with expressing the location of the referent with respect to a reference point, usually the interlocutors (deictics), and with indicating the speaker's assumption about the identifiability of the referent by the hearer (definiteness). The usual formal expression of these operators is in the form of determiners; in particular, articles and demonstratives. Operators in the RP follow the same iconic ordering constraint as operators within the clause (Rijkhoff 1990, 2002).

A preliminary general schema for the layered structure of the RP is given in

Level	Operator
Nuclear _N	Nominal aspect (count-mass distinction, classifiers) Number
Core _N	Quantification (quantifiers) Negation
RP	Definiteness Deixis

Table 3.4: RP operators

Figure 3.14 as a homomorphic mirror image (for ease of presentation). Syntactic templates for RPs and PPs would be stored in the syntactic inventory along with the other templates discussed in the previous section.

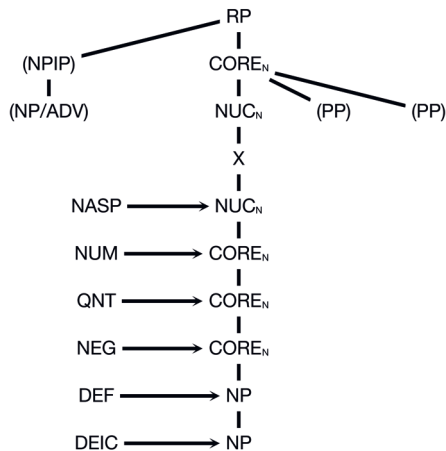


Figure 3.14: The layered structure of the RP

In Figure 3.15, the initial definite article *the* is attached to the RP node in the constituent projection by a dotted line due to its status as one of the RP operators.

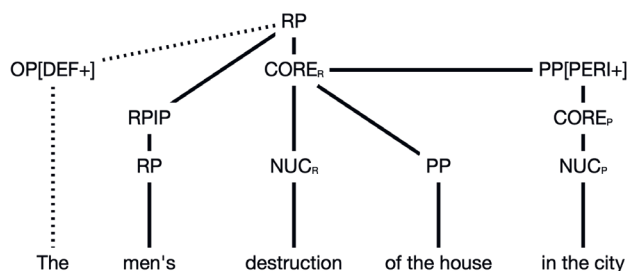


Figure 3.15: Complex RP

3.3 The semantic representation

As previously mentioned, the interplay of syntax and semantics lies at the heart of RRG, so that syntax no longer enjoys a dominant position. It is of little or no use to grasp the syntactic representation (Section 3.1.1) without meaning assigned to it. The semantic representation of a sentence is based on the lexical decomposition of the predicate in the nucleus, which falls under one of the six *Aktionsart* types¹⁶. The *Aktionsart* types and their defining features are given in Table 3.5.

	static	dynamic	telic	punctual
State	+	–	–	–
Activity	–	+	–	–
Achievement	–	–	+	+
Semelfactive	–	±	–	+
Process	–	–	–	–
Accomplishment	–	–	+	–
Active accomplishment	–	+	+	–

Table 3.5: Characterization of verbal *Aktionsart* based on Van Valin Jr (2022)

According to static parameters, the classes of states and activities make up the

¹⁶These distinctions were proposed in Vendler (1967), formalized in Dowty (1979), and extended by RRG (see Van Valin Jr and LaPolla 1997, 82-128, Van Valin Jr 2022)

most basic distinction. The dynamic parameter refers to any non-static event, telicity implies an inherent endpoint, and punctuality refers to instantaneity. Furthermore, the state-activity distinction is fundamental because they are the only classes to have argument positions that define thematic relations. The other classes are compositionally derived from one of these two classes, as can be seen in Table 3.6. Examples of English predicates of each class are given in (84) below:

- (84) a. **States** *be sick, be short, be dead, know, love*, etc.
 b. **Activities** *look at, walk, eat* (intransitive or transitive with non-referential object RP)¹⁷, *run* (without a complement), etc.
 c. **Achievements** *pop, shatter, explode* (intransitive), etc.
 d. **Process** *melt, freeze, grow*, etc.
 e. **Semelfactives** *sneeze, flash, blink, cough*, etc.
 f. **Accomplishments** *dry, dissolve* (intransitive), etc.
 g. **Active Accomplishments**¹⁸ *run* (+ goal PP), *eat* (+ a referential RP), *build* (+ RP), etc.

Furthermore each of these classes has a causative counterpart, as in (85) below:

- (85) a. **Causative state** *The jaguar frightens/scares the boy.*
 b. **Causative achievement** *The cat popped the balloon.*
 c. **Causative semelfactive** *The teacher tapped the pencil on the table.*
 d. **Causative accomplishment** *The hot water melted the ice.*
 e. **Causative activity** *The girl bounced the ball around the room.*
 f. **Causative active accomplishment** *The sergeant marched the soldiers to the park.*

¹⁷See Section 5.7.2.

¹⁸The difference between activities and active accomplishments is the telic use of activity verbs. This general pattern relates activity verbs of motion (e.g., *run*), consumption (e.g., *eat*), and creation (e.g., *paint*) to the corresponding active accomplishment verbs, (e.g., *run to the park*, *eat the cake*, and *paint the picture*), respectively.

The logical structure of each *Aktionsart* class is given in Table 3.6. Predicates (**pred'**) are represented in bold followed by a prime symbol. These are part of the vocabulary of the semantic metalanguage used in the decomposition not words in a particular language (e.g. English *hear* or German *hören*). Arguments are written inside the parenthesis, and variables 'x', 'y', and 'z' are used when no referents are specified.

Aktionsart class	Logical structure
State	pred' (x) or (x, y)
Activity	do' (x, [pred' (x), or, (x, y)])
Achievement	INGR pred' (x) or (x, y), or INGR do' (x, [pred' (x) or (x, y)])
Process	PROC being.consumed' (x) or (x, y) PROC becoming.higher/lower.on.[α]scale' (x) PROC moving(.direction)' (x)
Semelfactive	SEML pred' (x) or (x, y) SEML do' (x, [pred' (x) or (x, y)])
Accomplishment	BECOME pred' (x) or (x, y), or BECOME do' (x, [pred' (x) or (x, y)])
Activity accomplishment	do' (x, [pred' ₁ ' (x, (y))]) & INGR pred' ₂ ' (z, x) or (y)
Causative	α CAUSE β (α, β are logical structures of any type)

Table 3.6: *Aktionsart* and their logical structures, from Van Valin Jr (2005, 45)

As can be seen in Table 3.6, the activity, achievement, and semelfactive classes all have an activity component (**do'**[x,...]), which lexicalizes agency in the LS of the predicate and which is absent from state predicates. The remaining classes are related to a stative or inactive component in the LS: **predicate'** (x) or **predicate'** (x, y). Examples of each class with their respective logical decomposition are given in (86)-(92) below:

(86) STATES

- Igor is a fool **be'** (Igor, **fool'**)
 The window is shattered. **shattered'** (window)
 Fred is at the house. **be-at'** (house, Fred)
 John saw the picture. **see'** (John, picture)
- (87) ACTIVITIES
 The children cried. **do'** (children, [**cry'** (children)])
 The wheel squeaks. **do'** (wheel, [**squeak'** (wheel)])
 Carl ate snails. **do'** (Carl, [**eat'** (Carl, snails)])
- (88) ACHIEVEMENTS
 The window shattered. INGR **shattered'** (window)
 The balloon popped. INGR **popped'** (balloon)
- (89) SEMELFACTIVE
 John coughed. SEML **do'** (Mary, [**cough'** (Mary)])
 Mark glimpsed the image. SEML **do'** (Mark, [**glimpse'** (image)])
- (90) ACCOMPLISHMENTS
 The snow melted. BECOME **melted'** (snow)
 The sky reddened. BECOME **red'** (sky)
 Mary learned French. BECOME **know'** (Mary, French)
- (91) ACTIVITY ACCOMPLISHMENT
 Carl ate the snail
do' (Carl, [**eat'**] (Carl, snail)) \wedge PROC **being.consumed'** (snail) \wedge
 FIN **consumed'** (snail)
 Paul walked one mile to the store.
do' (Paul, [**walk'**] (Paul)) \wedge PROC **covering.path.distance'** (Paul, one
 mile) \wedge FIN **be-at'** (store, Paul)
- (92) CAUSATIVES
 The jaguar scared the boy.
 [**do'** (jaguar, \emptyset)] CAUSE [**feel'** (boy, [**afraid'**])]
 Max broke the window.
 [**do'** (Max, \emptyset)] CAUSE [BECOME **broken'** (window)]

The cat popped the balloon.

[**do'** (cat, ∅)] CAUSE [INGR **popped'** (balloon)]

Felix bounced the ball.

[**do'** (Felix, ∅)] CAUSE [**do'** (ball, [**bounce'** (ball)])]

The linguistically relevant distinctions of *Aktionsart* classes can be tested according to language-specific criteria (see Kroeger 2018, 381-386). In English, for example, situations describing states cannot be an answer to the question *What happened?* and cannot be used with the progressive form (be V-ing)¹⁹. Given the parameters that define the *Aktionsart* classes, such as duration, telicity, etc. (see table 3.5), the co-occurrence of an action with certain types of adverbs may serve as a reliable test. Activities and activity accomplishments may occur with adverbs that code dynamic action, as long as the adverbs do not require a controlling PSA (subject), e.g., *deliberately*, *carefully*, because they are incompatible with activity verbs that have PSAs referring to non-agentive participants, e.g., *the dog shivered violently*/**deliberately*. Accomplishments, such as *die* (BECOME **dead'** (x)) in languages such as English, German, and Portuguese, may co-occur with an adverb like *slowly*, which is impossible for achievements. In Mandarin, for example, the verb *die* is an achievement (INGR **dead'** (x)), so that **tā sǐ de kuài* 'he died quickly' is ungrammatical. For examples of other tests, see Van Valin Jr and LaPolla (1997, 94-102) and Kroeger (2018, 381-386).

Some languages may mark *Aktionsart* classes of predicates morphologically. In Tupinambá, most of the verbs which have the operator CAUSE in their logical structures will take the prefix *mo-*, indicating that they are causatives – compare Tupinambá *mboʔe* 'teach' (< *mo* + *ʔe*, literally 'cause to say') and English *teach*, where the causative is not morphologically marked. Both have the logical structure [**do'** (x, ∅)] CAUSE [BECOME **know'** (y_{COGNIZER}, z_{CONTENT})].

There are only five thematic relations, one for each position in the logical structure. This can be seen in the 'thematic relations continuum' along the

¹⁹Regarding the use of states with progressive forms, there are marked interpretations, such as *Henry is loving the game*.

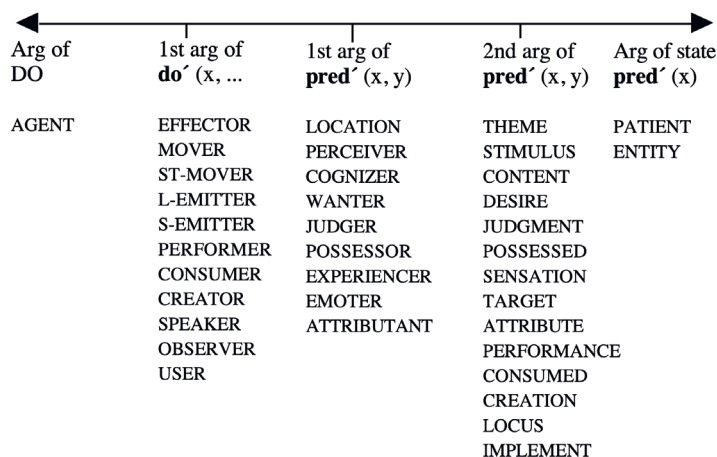


Figure 3.16: Thematic relations continuum along the actor-undergoer hierarchy. From Van Valin Jr (2001a)

actor-undergoer hierarchy shown in Figure 3.16.

The semantic interpretation of an argument in the logical decomposition is a function of its position in the LS of the predicate: the leftmost argument, in terms of the actor-undergoer hierarchy (AUH), is the actor, while the rightmost argument is the undergoer. Some examples are given in Table (3.7) for state predicates and in Table (3.8) for activity predicates (from Van Valin Jr 2005, 55).

State			
1 arg:	State/condition	broken' (x)	x = PATIENT
	Existence	exist' (x)	x = ENTITY
2 args:	Pure location	be-loc' (x, y)	x = LOCATION, y = THEME
	Perception	hear' (x, y)	x = PERCEIVER, y = STIMULUS
	Cognition	know' (x, y)	x = COGNIZER, y = CONTENT
	Identificational	be' (x, [pred'])	x = IDENTIFIED, y = IDENTITY

Table 3.7: Decomposition of state predicates and thematic relations

Activity			
1 arg	Unspecified action	do' (x, \emptyset)	x = EFFECTOR
	Motion	do' (x, [walk' (x)])	x = MOVER
	Static motion	do' (x, [spin' (x)])	x = ST-MOVER
	Light emission	do' (x, [shine' (x)])	x = L-EMITTER
2 args	Performance	do' (x, [sing' (x, (y))])	x = PERFORMER, y = PERFORMANCE
	Consumption	do' (x, [eat' (x, (y))])	x = CONSUMER, y = CONSUMED
	Repetitive action	do' (x, [tap' (x, (y))])	x = EFFECTOR, y = LOCUS
	Directed perception	do' (x, [see' (x, (y))])	x = OBSERVER, y = STIMULUS

Table 3.8: Decomposition of activity predicates and thematic relations

The arguments of predicates in each class share certain characteristics and semantic roles. Since many verb-specific semantic roles can be generalized, (e.g., giver, runner, killer, and speaker are all agents), semantic roles can be generalized through thematic relations,²⁰ which in turn are generalized by two semantic macroroles, actor and undergoer (see Section 3.3.2). Actor is a generalization across agent, experiencer, recipient, and other roles, while undergoer is a generalization subsuming stimulus, theme, recipient, and other roles²¹. Agent is the prototype for actor, and patient is the prototype for undergoer. Figure (3.17), from Van Valin Jr (2005, 54), summarizes these increasing generalizations:

Due to the characteristics that arguments of predicates share with semantic roles, it is possible to associate thematic relations with particular predicate classes or, more specifically, with particular positions in semantic representations. For example, with a one-argument state predicate, this argument will have the role of patient (**pred'** (x)). In the case of a two-argument state predicate, the leftmost argument in the hierarchy (Figure 3.18) will be the actor and the rightmost will be the undergoer. Each argument in the syntactic representation of a simple clause must be linked to an argument position in the LS of

²⁰Thematic relations do not play a role in the theory; they are mere mnemonics for the LS argument positions, e.g., ‘perceiver’ is the mnemonic for the first position (x) in a two-place perception LS like **hear'** (x, y).

²¹It is not a contradiction that the thematic relations of experiencer and recipient may be either actor or undergoer.

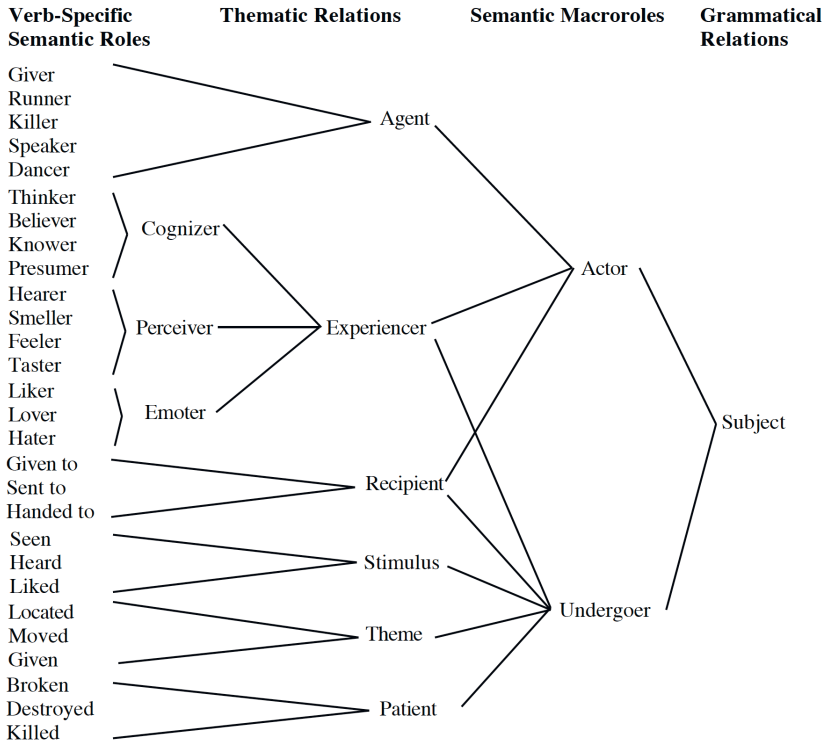


Figure 3.17: Increasing generalization of semantic contrasts (from Van Valin Jr 2005, 54)

the predicate.

In the continuum in Figure 3.16, it can be seen that with the exception of ‘agent’, each of the thematic relations listed under a particular argument position represents a distinct subclass of state or activity verb, and accordingly, each is a label for an argument position in the LS of a particular type of verb. For example, ‘cognizer’ means ‘first argument of a two-argument state predicate of experience’, and ‘content’ means ‘second argument of a two-argument state predicate of perception’. These thematic relations cannot be predicted, and therefore need to be made explicit: **know**' ($x_{\text{COGNIZER}}, y_{\text{CONTENT}}$).

3.3.1 Lexical representation of nominals

RRG bases the semantic representation of nominals on the theory of Nominal Qualia proposed by Pustejovsky (1995)²² and represents them in terms of the decomposition system used in RRG (see Van Valin Jr 2005, 50-3). The Generative Lexicon, proposed by Pustejovsky (1995), is an approach relying on a richly structured lexicon. This approach to lexical representation focuses on semantic phenomena such as coercion and systematic polysemy. Lexical entries include, in addition to argument structure, an “event structure” and a “Qualia structure”, both of which play a fundamental role in GL accounts of semantic composition.

The Qualia theory postulates that the meaning of nouns can be captured by four Qualia relations or roles, which together constitute the Qualia Structure of the word in question. These are the Formal, the Constitutive, the Telic, and the Agentive Quale. These can be understood as four different perspectives on defining sense, as given in (3.3.1).

(93) Qualia theory

- a. **Constitutive role:** accounts for the relation between an object and its constituents
 - 1. material
 - 2. weight
 - 3. parts and component elements
- b. **Formal role:** that which distinguishes the object within a large domain
 - 1. orientation
 - 2. magnitude
 - 3. shape
 - 4. dimensionality
 - 5. colour
 - 6. position
- c. **Telic role:** purpose and function of the object

²²See also Pustejovsky and Jezek (2016).

1. purpose that an agent has in performing an act
 2. built-in function or aim that specifies certain activities
- d. **Agentive role:** factors involved in the origin or ‘bringing about’ of an object
1. creator
 2. artifact
 3. natural kind
 4. causal chain

The Qualia structure of the TUP noun *miape* ‘bread’ is represented in the attribute value matrix in (94):

$$(94) \left[\begin{array}{l} \text{MIAPE} \\ \text{QUALIA} \left[\begin{array}{l} Q_F \textit{ food} \\ Q_C \textit{ flour, water, yeast, salt} \\ Q_T \textit{ eat} \\ Q_A \textit{ bake} \end{array} \right] \end{array} \right]$$

Referential identification can be specified as shown in (95). This says that *r* refers to an entity (or set of entities) in a domain of reference which is an element in the extension of the set of qualia properties $\{Q_C, Q_F, Q_T, Q_A\}$ in the domain of reference. For the reference of a particular noun, the qualia properties in (95) are coindexed with the qualia properties specified for that noun, as in (94).

$$(95) \text{ Referential identification: } \{Q_C, Q_F, Q_T, Q_A\} (r)$$

Nominal modifiers are also represented as predicates in the semantic representation. Attributive predication is represented by ‘**be**’ ($x, [\text{pred}']$), as in example (96) and (97), predicative and attributive modification respectively.

$$(96) \text{ The car is white} \\ \text{be}'(\text{car}, [\text{white}'])$$

$$(97) \text{ I see a white car} \\ \text{see}'(I, [\text{be}'(\text{car}, [\text{white}'])])$$

3.3.2 Semantic Macroroles and Lexical Entries for Verbs

The idea of semantic macroroles is unique to RRG (see Van Valin Jr 1999a, 2001a). They are the “primary interface between the LS and syntactic representations” (Van Valin Jr 2006b, 287)

The two semantic macroroles, actor and undergoer²³, are the two primary arguments of a transitive predication, either one of which may be the single argument of an intransitive verb²⁴. The correlation between the semantic position of the argument and how likely it is to be the actor or undergoer is captured by the AUH, shown in Figure 3.16:

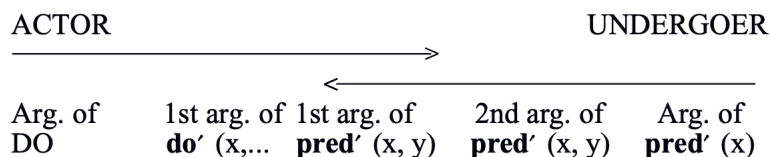


Figure 3.18: The actor-undergoer hierarchy. The arrows indicate the increasing markedness of the realization of an argument as macrorole

(98) Default Macrorole Assignment Principles

- a. Number: the number of macroroles a verb takes, which is less than or equal to the number of arguments in its logical structure
 1. If a verb has two or more arguments in its LS, it will take two macroroles.
 2. If a verb has one argument in its LS, it will take one macrorole.
- b. Nature: for verbs which take one macrorole
 1. If the verb has an activity predicate in its LS, the macrorole is the actor.

²³Van Valin Jr (2006b, 78-81) presents convincing evidence for positing only two macroroles. See also Van Valin Jr (1999a).

²⁴They correspond to the pre-theoretical notions of ‘logical subject’ and ‘logical object’ or, alternatively, to the general notions of ‘agent’ and ‘patient’ (Van Valin Jr 2022, 88). RRG does not use these labels because they are normally used to refer to syntactic rather than semantic relations (Van Valin Jr 2005, 60).

2. If the verb has no activity predicate in its LS, the macrorole is the undergoer.

Transitivity in RRG is defined semantically in terms of ‘M-transitivity’ (macrorole-transitivity), which corresponds to the number of macrorole arguments a predicate takes. This is in contrast to ‘S-transitivity’ (syntactic transitivity), which refers to the number of syntactic arguments a predicate takes. The M-transitivity of a verb can be predicted using the principles in (98a). If these principles are apparently violated, then the source of irregularity comes from the fact that a verb has fewer macroroles than the principles in (98a) would predict, and this irregular M-transitivity is marked in their lexical entries (see Van Valin Jr 2004), as e.g., the verb *helfen* ‘help’ and *gefallen* ‘please’ in German, which are syntactically transitive, but their M-transitivity is irregular. The three M-transitivity possibilities are: transitive (2 macroroles), intransitive (1 macrorole), and atransitive (0 macroroles).²⁵ Three-place predicates take three core arguments, but there can be no more than two macroroles (Van Valin Jr 2005, 64). The third argument in a ditransitive construction is not a macrorole, but rather a “non-macrorole core argument”.

3.3.3 Syntactic functions, case and preposition assignment

The linking of semantics and syntax has two phases: first, the determination of semantic macroroles based on the LS of the predicate; and second, the mapping of the macroroles and other arguments onto syntactic functions. RRG does not assume traditional grammatical relations like subject and object, positing instead a single, construction-specific grammatical relation, which is termed the privileged syntactic argument (PSA) of the construction (see Section 5.4). PSAs are associated with the notion of controller or pivot (Van Valin Jr and LaPolla 1997, chap. 6 and Van Valin Jr 2005, 94-101). The non-PSA syntactic arguments in the clause are referred to as the direct core argument (DCA) and oblique core argument (OCA). Languages have selection hierarchies to determine the PSA; the main ones are given in (99) and (100), from Van Valin Jr (2005, 100).

²⁵For the principles determining the m-transitivity of verbs, see Van Valin Jr (2005, 63-66). M-transitivity and S-transitivity may coincide, but this is not always the case.

- (99) Privileged Syntactic Argument Selection Hierarchy:
 arg of DO > 1st arg of **do'** > 1st arg of **pred'** (x, y) > 2nd arg of **pred'** (x, y) > arg of **pred'** (x)
- (100) Accessibility to Privileged Syntactic Argument Principles:
- a. Accusative constructions: Highest ranking direct core argument in terms of (99) [default]
 - b. Ergative constructions: Lowest ranking direct core argument in terms of (99) [default]
 - c. Restrictions on PSA in terms of macrorole status:
 1. Languages in which only macrorole arguments can be PSA: German, Italian, Dyirbal, Jacaltec, Sama, ...
 2. Languages in which non-macrorole direct core arguments can be PSA: Icelandic, Georgian, Japanese, Korean, Kinyarwanda, ...

The PSA selection hierarchy in (99) (from Van Valin Jr 2005, 100) is the actor part of the AUH. For a language like TUP, (100a) the actor is the PSA, but there is significant cross-linguistic variation regarding PSAs (Van Valin Jr and LaPolla 1997, chapter 6). Because Tupinambá has no passive voice, it is not possible for the undergoer of a transitive verb to function as the privileged argument; there is a restricted neutralization of semantic contrasts because either the actor or the undergoer can function as the PSA with intransitive verbs.

Case assignment rules are formulated with reference to the linking system. The basic rules for direct core arguments in accusative languages are given in (101).

- (101) Case marking rules for accusative languages:
- a. Highest ranking core macrorole takes nominative case.
 - b. Other core macrorole takes accusative case.
 - c. Non-macrorole direct core arguments take dative as their default case.

In a language like English, without RP case marking, there are rules for preposition assignment. The rules for ‘*to*’, ‘*from*’ and ‘*with*’ are given in (102).

- (102) Preposition assignment rules for English
- a. Assign ‘*to*’ to non-MR x argument in LS segment: ... BECOME/INGR **pred'** (x, y)
 - b. Assign ‘*from*’ to non-MR x argument in LS segment: ... BECOME/INGR NOT **pred'** (x, y)
 - c. Assign ‘*with*’ to non-MR y argument if, given two arguments, x and y, in a logical structure, with x lower than or equal to y on the Actor-Undergoer Hierarchy, y is not selected as a macrorole.

The rules in (102b,c) do not cover all of the uses of ‘*from*’ and ‘*with*’, and they are presented for illustrative purposes only. For more information on the assignment of adpositions, see Van Valin Jr and LaPolla (1997, 376-384).

3.4 Linking algorithms

RRG is a parallel architecture theory in terms of Jackendoff (2002) because syntax, semantics, and discourse-pragmatics have independent representations that may interact with each other. The previous sections have introduced the components that describe grammatical structure, i.e., the LSC, the lexical representation with semantic roles, syntactic functions, and focus structure. This is a significant difference from Transformational Grammar, in which an expression can have a number of constituent structures; that is, it is not possible for an expression to occupy one position at one level of structure and another position at another level. Similarly distinct from Minimalism, RRG does not assume exclusively binary branching, leading to simpler flat structures. Another difference from minimalism is the fact that phonologically empty elements are not part of the theory, so that different sentence structures in languages may correspond to the same semantic structure, as in both sentences in (103), in English and Czech respectively. Note how Czech lacks an article and does not need the independent pronoun.

- (103) a. I read a/the book.
 b. Čtu knih-u
 read.IPFV.1SG book-ACC.SG
 ‘I read a/the book.’

The linking system relating semantic and syntactic representations is summarized in Figure (3.19). Syntactic functions like the PSA and direct core arguments represent the syntactic pole of the system. These are structurally instantiated in the LSC. The logical structure represents the semantic pole.

The technical details of the linking algorithm are developed in Van Valin Jr and LaPolla (1997) and are not discussed here (see also Van Valin Jr 2022 for a recent summary).

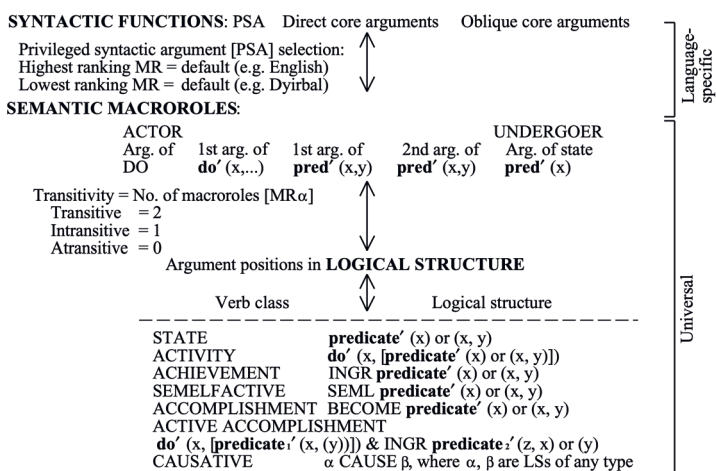


Figure 3.19: The linking algorithm, from Van Valin Jr (2005, 129)

The relation between logical structure and macroroles is mediated by the actorundergoer hierarchy, shown in Figure 3.18. The relation between macroroles (and non-macrorole arguments of the verb) and morphosyntactic functions is subject to extensive cross-linguistic variation and is affected by the PSA selection hierarchy in Example 99 and selection principles in 100, as

well as by the extent to which focus structure is grammaticalized in clause-internal relational syntax (see Van Valin Jr 2005, 101-107).

Chapter 4

Word classes

4.1 Introduction

Von Humboldt's (1836) assumption of language-specific word classes was later adopted by the American Structuralist School (see Boas 1911a) under the view that lexical categories do not display exactly the same grammatical properties cross-linguistically (see Boas 1911a; Sapir 1921; Anward et al. 1997, Croft 2001, 63-83, Foley 2017, 182 Haspelmath 2007, 2020).

In line with recent typological research, word classes¹ are here considered neither as inherent properties of lexical roots, nor as atomic, primitive units of grammatical analysis and structure (Dryer 1997a; Haspelmath 2007; Croft 2001; Song 2018; Haspelmath 2021b; Croft 2022a,b). Word classes are taken to be a comparative concept, and in this sense, the question of whether all languages have some or all of the same categories becomes meaningless (Dixon 1982; Anward et al. 1997; Dixon 2004; Croft 2001; Rijkhoff 2007; Rijkhoff and van Lier 2013; Croft 2022b). As Croft (2022a, 11) observes, *the facts that are supposedly about a word class in a language description are really*

¹For the disambiguation of terms such as parts of speech, lexical categories, syntactic categories, and word classes, see Rauh (2010).

facts about the construction(s) used to define that word class. Comparative concepts (Haspelmath 2010) involving form and function would serve as an appropriate basis for cross-linguistic comparison (Croft 1990, 2003)². Formal properties must be defined in a cross-linguistically valid fashion, i.e., not in terms of language-specific distributional patterns (see Croft 2001, 75-83 and Croft 2022b). Word classes must be defined in terms of their roles in constructions (Croft 2022b).

A common sense ontology, based on the denotational or contentive meaning of lexical roots (see Beck 2002, 12), reflects how entities populating the universe are perceived and conceived by human cognition (Jackendoff 1983; Braine 1992, 1994; Kemmerer 2019). This “universal” ontology or semantic categorization coincides with a linguistic categorization that classifies lexical items according to language-specific criteria. When used in a general sense, ‘noun’, ‘verb’, and ‘adjective’³ are useful terms for describing languages, but these concepts seldom correspond to or display any overlapping properties of these categories from one language to another (Anward et al. 1997; Haspelmath 2007; Evans 2000). *Whiteness* denotes a property but is classified as a noun in English, just as *destruction* denotes an action but is classified as a noun. Tupinambá, as well as other TG languages, has a unique form for *white*, *whiteness*, and *be white*, respectively (see Dietrich 2001), which can only be distinguished on the basis of constructions.

Beginning with semantics, object-, property-, and action-denoting roots can be categorized according to their occurrence in three propositional act functions or information packaging (Searle 1969; Croft 1991, 2001; Haspelmath 2021b; Croft 2022a): reference, modification, and predication. Some combinations of semantic classes and information packaging tend – though this is not a necessary condition – to exhibit a “default behavior” across languages, with each root type being more or less marked when associated with one of the

²Croft (2022b) prefers to avoid the term ‘comparative concept’.

³Following best practice in typology, for language-specific (descriptive) categories I capitalize the initial letters of the name of the category, while comparative concepts are not written with the first letter capitalized. So, for example “Adjective” refers to the language-specific lexical category manifesting a modifying function in the language under discussion, while “adjective” refers to the comparative concept as found in different languages.

information packaging types (see Croft 1991, 2001, 2022a). Treating information packaging and semantic classes as parameters that are non-independent, the default behavior or privileged combinations are the following: reference to object = the prototypical “noun”, modification by property = prototypical “adjective”, and predication of action = prototypical “verb”. These prototypes are based on the cross-linguistic distribution of language-internal distribution patterns, revealing a tendency of overt coding (markedness) and inflection to avoid prototypes (Haspelmath 2021b; Croft 2022a).

In Table 4.1, the “privileged combinations” are the values in the diagonal, i.e. <object, reference>, <property, modification>, and <action, predication>. Function indicators (additional coding), in the case of English, are given in bold.

	reference	modification	predication
objects	–	genitive flag <i>the rent of the house</i>	object-word copula <i>is a student</i>
properties	substantivizer <i>(the new one)</i>	–	property copula <i>is big</i>
actions	nominalizer <i>the open-ing</i>	relativizer <i>the work that they did</i>	–

Table 4.1: Information packaging and semantic root classes combined in English. The elements in the diagonal are those exhibiting “default” behavior

Based on the behavior displayed by these combinations, (2021b) posits five types which account for different indicator coexpression patterns for different aspects of coding. English, for example, in predicative function, is of the nominalis coexpression type, since it requires a copula for object- and property-roots, but not for the action-root (see Figure 4.1).

RRG, as a typologically oriented theory, also considers categories to be language-specific. RRG assumes functionally motivated non-endocentric syntactic categories such as the nucleus containing the predicate, (potentially) referential phrases, modifying phrases, or even clauses. These syntactic slots can be realized by whatever lexical category is employed in a given language’s specific

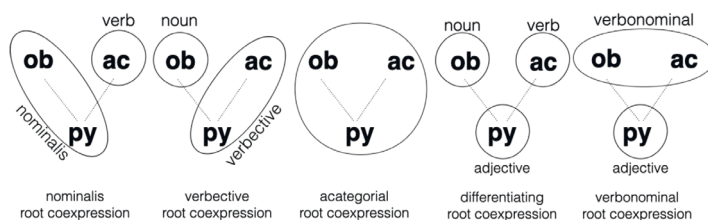


Figure 4.1: Five coexpression types of function indicators from Haspelmath (2021b)

syntactic templates. Categorical specifications of lexical items and the syntactic slots into which they are inserted must not match, since such a specification does not form part of the syntactic structure (see Van Valin Jr 2008b). As Van Valin Jr (2022, 15) observes, the status of lexical categories in RRG is related to the semantic distinction between referential expressions and predicates, a distinction borne of the fundamental opposition motivating the LSC, i.e., that between predicating and non-predicating elements. Thus, RRG also acknowledges the combination of lexical roots and information packaging types. Van Valin Jr (2016) also suggests that at the most basic level, lexical items fall into one of two classes: they are either referring expressions or predicates and they are like grammatical relations: language specific but with a universal semantic foundation. Table 4.2 shows that this is the case in TUP, where the basic distinction is between predication and non-predication. Modification is either found in predication or in referential function.

4.2 Word classes in Tupinambá

Word classes have been a recurrent topic in the studies of Tupían languages (Rodrigues 1996a; Dietrich 2001; Seki 1990, 2000; Queixalós 2001; Rose 2002; Meira 2006; Dietrich 2017c)⁴, and this is not surprising. Already in

⁴The topic has been discussed since the beginning of last century, regarding different language families: Boas (1911b); Swadesh (1938); Frachtenberg (1922); Davis and Saunders (1997); Hébert (1983). See also Mithun (2001, 56-67), Dixon (2010a, 37-61), Davis and Matthewson (2009), Lazard (1999); Broschart (1997); Evans and Osada (2005); Peterson (2007).

1595, Anchieta noted in his grammar that it was not possible to talk about word classes in a comparative way (see Anchieta 1595, 44v-45.)

Based on the types defined by Haspelmath (2021b) presented in the previous section, Tupinambá belongs to the acategorial type, since none of the root semantic classes require a copula in predicate function. TUP lexical roots are existential predicates and require additional coding (function indicating morphology) in order to be used in modification or reference. The three semantic classes, namely objects, properties, and actions (Croft 2022a), combined with propositional act functions and overtly marked structural coding for POS in Tupinambá, are exemplified in Table (4.2).

	reference	modification	predication
objects	<i>ok-a</i> ‘house’	<i>r-ok(-a)</i> ‘(of the) house’	<i>ok</i> ‘be a house’
properties	<i>poraj-a</i> ‘beauty’	<i>poraj</i> ‘beautiful’	<i>poraj</i> ‘be beautiful’
actions	<i>kutuk-a</i> ‘poking’	<i>o-kutuk-βaʔe</i> ‘one that pokes’	<i>kutuk</i> ‘poke’

Table 4.2: Semantic classes combined with propositional act functions and overtly marked structural coding for POS in Tupinambá

The fundamental distinction shown in Table 4.2 is that between predication and reference (see Meira 2006, 212), which is manifested through markedness⁵ and captured by a mapping between semantic class and pragmatic function.

While predication is characterized by the absence of overt coding, two types of predication are found in TUP and they are differentiated by constructions. The “verbal” type requires indexes from Set II and IV (see Table 4.3), while PREDPOSSESSIVE predication⁶ requires bound indexes from Set I (see Table 4.3) .

The propositional speech act function reference is illustrated below with all

⁵Markedness, as explained in Beck (2002, 21-24), should not be understood exclusively in terms of additional markedness (more marked). There are instances in languages of less marked where the less marked form, e.g., one that has undergone the loss of a morpheme (decategorization), suggests markedness.

⁶This terminology is taken from Haspelmath (2022).

three semantic root classes. All semantic classes require the referential suffix *-a* ~ \emptyset for this function⁷:

(104) Object-word in reference function

Nerera renupa aβe
 ne=r-er-**a** r-enup-a aβe
 2SG=R₁-name-REF R₁-hear-GER ADV
 ‘Also hearing your name.’ (Poemas, 174)

(105) Property-word in reference function

K^warasi sose oporaja k^waβeʔeŋa
 k^warasi- \emptyset \emptyset -sose o-poraŋ-**a** k^waβeʔeŋ-a
 sun-REF R₁-POSP CORF-beauty-REF show-GER
 ‘Showing her (own) beauty (which is) more than the sun.’ (Araújo, 4v)

(106) Action-word in reference function

Maʔeete kaʔuwasu
 maʔe-ete kaʔu-wasu- \emptyset
 thing-INTS drink-big-REF
 ‘A big drunkenness is a good thing.’ (Teatro, 8)

The propositional speech act function modification is illustrated with all three semantic root classes. Modifications will be either arguments or predicates and additional coding will vary depending on the type of modification: genitive flag as in (107), juxtaposition as in (108), and relativization as in (109).

(107) Object-word in modification function

Oka rerekoara
oka- \emptyset r-erekoar-a
house-REF R₁-guardian-REF
 ‘Guardian of the house.’ (VLB, I, 6)

(108) Property-word in modification function

⁷The allomorph *-a* is used after consonants, while \emptyset is used after vowels (see Table 4.5).

Nerekoporaja
 ne=r-eko-**poraj**-a
 2SG=R₁-be-**beauty**-REF
 ‘Your beautiful way (of being).’ (Teatro, 122)

- (109) Action-word in modification function

Owatáβaʔe
 o-**wata**-βaʔe
 3-**walk**-REL
 ‘The one going / the going one.’ (DC, II, 79)

The propositional speech act function predication is exemplified with object-word (110), property-word (111), and action-word (112) (see Section 5.5).

- (110) Object-word in predicate function (predpossessive/existential predication)

ʃerer
 ʃe=r-**er**
 1SG=R₁-**name**
 ‘I have a name / there is my name.’ (VLB, II, 50)

- (111) Property-word in predicate function (predpossessive/existential predication)

Iporaj
 i-**poraj**
 R₂-**beauty**
 ‘It was beautiful.’ (Poemas, 152)

- (112) Action-word in predicate function (verbal predication)

Paranã rupi awata
 paranã-∅ r-upi a-**wata**
 sea-REF R₁-through 1SG-**walk**
 ‘I walked through the sea.’ (VLB, II, 48)

Launey (1994, 2002, 2004) suggested the OMNIPREDICATIVE character of Classical Nahuatl and consequently described a language type in which members of all major open word classes may function equally and without derivation as predicates, and in which the predicative use is primary, while the referential is syntactically derived⁸. TUP and some TG languages have been analyzed as being of the omnipredicative type (Queixalós 2006; Magalhães et al. 2019)⁹. The system described by Launey is simply one of the root expression types found cross-linguistically, listed in Haspelmath (2021b), and since the terminology in Haspelmath (2021b) better connects all coexpression types, relating them to each other in a consistent manner, I will henceforth avoid the term omnipredicative.

4.3 Noun classes

Before presenting the noun classes, it is necessary to present the sets of person indexes (possessor and cross-index markers) and the so called relational morphemes, since they often occur together (4.3.1) and (4.3.2).

4.3.1 Argument indexes and pronouns

Tupinambá indexes arguments through bound INDEXES (See Section 6.3). Even though the distinction between affixes and clitics is gradient (Zwicky and Pullum 1983; Haspelmath 2002, 2011), verbal argument indexes are here considered to be prefixes, while possessor markers are considered to be clitics, because they combine not only with simple nominals but also with RPs (see examples (125a) and (126a))¹⁰.

⁸Omnipredicative languages in his terms roughly correspond to the nonconfigurational type defined by (Hale 1983; Jelinek 1984; Baker 2001).

⁹Other languages that have been said to be of the omnipredicative type are: Salish (Kinkade 1983), some Philippine languages (Lemaréchal 1989, 1991; Himmelmann 2008), Yucatec Maya (Vapnarsky 2013), Khoekhoe (Hahn 2014), Sikuani (Queixalós 2000).

¹⁰The definition of clitics and affixes is here taken to be a straightforward one: affixes are class-selective, while clitics are indiscriminate, combining with any word-class (see Haspelmath 2021b).

Like most Tupían languages, TUP uses the same set of argument indexes¹¹ in the form of prefixes for marking A/Sa (Set II) on verbs, and clitic person indexes for indexing possessors and the complement of postpositions (Set I, see Rodrigues and Cabral 2012, 543-552). Argument indexes and free person forms are given in 4.3.

	Set I	Set II	Set III	Set IV	Free Forms
1SG (1)	je=	a-	wi-		ise
2SG (2)	ne=	ere-	e-	oro-	ene
3SG/PL (3)	(see Section 4.3.2)	o-	o-		(aʔe)
1EXCL.PL (13)	ore=	oro-	oro-		ore
1INCL.PL (12)	jane=	ja-	ja-		jane
2PL (23)	pe=	pe-	peje-	opo-	peʔẽ
Generic (123)	(ase)				ase

Table 4.3: Tupinambá person markers

Table 4.3 shows that TUP has a system of six persons in cross-referencing prefixes and personal pronouns. Two numbers are distinguished for two persons (first and second), but not for the third. The first person plural distinguishes between two forms, which include or exclude the hearer. Personal cross-referencing on prefixed nouns is used to indicate possessor (Set I).

Members of Set I are proclitic markers which attach to lexical roots (possessed roots only), postpositions, and RPs (see examples (125b) and (126a,b)). Contrary to what has been said in the literature (Jensen 1998a; Rodrigues 2010b), they are not absolute markers. Rather, these clitics are either indexes on postpositions (see Section 7.4) or possessive indexes (see Dietrich 2001, 2017b). As clitics, they cannot be stressed, they cannot be focused, and require a host to attach to (see Spencer and Luís 2012).

- (113) a. Pemaʔenwar
 pe=∅-maʔenwar
 2PL=R₁-remember

¹¹The term ‘argument index’ is taken from Haspelmath (2013).

- ‘You remember / there is your remembrance.’ (Anch., Arte, 20v)
- b. *ʃeaiβ*
ʃe=∅-aiβ
 1SG=R₁-impaired
 ‘I am impaired / there is my impairment.’ (VLB, I, 83)
- c. *Nekoʔema*
ne=∅-koʔem-a
 2SG=R₁-morning-REF
 ‘Your morning.’ (see Cantigas, IV)
- d. *Neruβ*
ne=r-uβ
 2SG=R₁-father
 ‘You have a father / There is your father.’ (FA, 39)

While members of Set I clearly have their origin in the independent pronouns (see Table 4.3) – they are reduced forms of the independent pronouns – third person markers constitute an exception. It might well be possible that these were part or related to a series of person markers that have been lost (see Gildea 2002). These morphemes will be dealt with in Section 4.3.2. For now, it will be enough to call attention for the fact that while examples (113) with first or second person are possessive RPs, formed by three morphemes: [possessor + relational + possessed], examples (114) are formed by two: [relational + possessor]. This indicates that the relational is not a third person index but indicates its absence within the constituent.

- (114) a. *Imaʔenwar*
i-maʔenwar
R₂-remember
 ‘He remembers.’ (AA, 20v)
- b. *Tuβa*
t-uβ-a
R₂-father-REF
 ‘His father.’ (Araújo, 4)

Set II markers are used to instantiate the core arguments, actor (ACT) or undergoer (UND), of transitive or intransitive verbs.

- (115) a. Jũ rupi awata
 jũ-∅ r-upi a-wata
 field-REF R₁-through 1SG-walk
 ‘I walked through the fields.’ (FA, 123)
- b. Nojanduj moropotára
 na-**o**-i-anduβ-i moropotar-a
 NEG-3-R₂-feel-NEG lust-REF
 ‘She did not feel lust.’ (Poemas, 182)
- c. Aβeβuj
 a-βeβuj
 1SG-float
 ‘I float.’ (VLB, II 21)
- d. São Pedro itañapema osekij
 São Pedro itañapem-a **o**-s-ekij
 Saint Peter sword-REF 3-R₂-pull.out
 ‘Saint Peter pulled out the sword.’ (Araújo, 54v)

Set III markers are coreferential indexes used in core-junctures with the gerund (see Chapter 10) and should not be associated with switch-reference (see Van Valin Jr and LaPolla 1997, 287-294, Hammond 2015, van Gijn and Hammond 2016) (see Section 10.1.2). They are nominal in origin and must still be analyzed as such. As evidence for the nominal origin of Set III indexes, one may take their etymological connection with possessor markers in languages such as Mawé and Awetí (cf. Meira and Drude (2013, 4-5) and Jensen (1998b)). Table 4.4 shows possessor markers in Mawé (Silva et al. 2010) and Awetí (Reiter 2012), and coreferential markers in Tupinambá.

Set IV are portmanteau indexes¹², encoding features of two arguments of a verbal predicate through a single morpheme (Cysouw 2003, 18-19 and Trom-

¹²For an overview regarding portmanteau indexes in TG languages, see Rose (2009, 2015b). For a discussion of the possible origin of these morphemes, see Cabral (2001a).

	Mawé Poss.	Awetí Poss.	Tup. Corref. (Set III)
1SG	u(j)-	i-	wi-
2SG	e-	e-	e-
1EXCL.PL	uru-	ozo-	oro-
1INCL.PL	aj-	kaj-	jere-
2PL	ej-	eʔi-	peje-
3CRF	to-	o-	o-

Table 4.4: Possessor markers in Mawé, Awetí, and coreferential markers in Tupinambá (Set III)

mer 2007). TUP has two portmanteau markers: *oro-* 1(SG/PL) → 2SG (116)¹³ and *opo-* 1(SG/PL) → 2PL (117). As shown by Anchieta (1595, 12,37), the independent pronoun was used to disambiguate between the singular (117a) and plural (117b) of the first person actor:

- (116) Oropisiβ umã jandikaraiβa pupe
oro-pisiβ umã jandi-karaiβ-a Ø-pupe
 1.ACT.2.UND.SG-anoint already oil-holy-REF R₁-POSP
 ‘I have already anointed you with blessed oil.’ (Ar., 141)

- (117) a. Ise opojuka
 ise **opo-juka**
 I 1.ACT.2UND.PL-kill
 ‘I kill you.’ (AA, 12)
- b. Ore opojuka
 ore **opo-juka**
 We.EXCL 1.ACT.2.UND.PL-kill
 ‘We kill you.’ (AA, 12)

Independent pronouns alone constitute RPs which can be coindexed with the

¹³The edition consulted for the text of Araújo has *pytub* (*pituβ*) instead of *pisiβ*. Whether this is dialectal variation or a printing error is not known.

arguments, as in (118) (see section 6.3). Their use is often pragmatically motivated, e.g., for topicalization or contrastive focus (see section 9). Note that only the RP coindexed with the actor can be coded as an independent pronoun when portmanteau indexes (1 → 2) are used.

- (118) Ise orojuka
 ise **oro**-juka
 I 1.ACT.2.UND.SG-kill
 ‘(It is) I (who) kill you.’ (FA, 9)

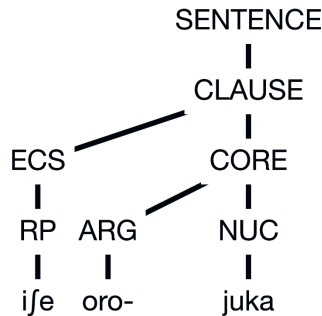


Figure 4.2: Independent pronoun in actor function with portmanteau argument index

There is no independent pronoun for third person; *a?e* (see Table 4.3) is a demonstrative (see Section 8.2.1.1), but there is a free form which requires the third person index *o-*, *ase*. Rodrigues (1990) treats *ase* in terms of a generic marker, including first, second and third persons. In (119), *ase* is given as a free form (referent of the third person pronoun), as in (119a). It is given as an argument of a postposition in (119b), and as a possessor index in (119c).

- (119) a. Tupã omanõ, memetipo ase omanõβo
 Tupã o-manõ memetipo **ase** o-manõ-βo
 God 3-die even.more **we.all** 3-die-GER
 ‘(If) God died, even more we are to die.’ (FA, 163)

- b. Aβa aβape ase rese Tupã moŋetasáramo
 aβa aβa=pe **ase** r-ese Tupã moŋeta-sar-ramo
 person perso=Q **we.all** R₁-because God pray-NMZL_{AG}-TRSL
 sekow?
 s-eko-w
 R₂-be-NFOC
 ‘Who prays to God because of us? / Who are those who pray to
 God for us? (Araújo, 23v)
- c. Marã e?ipe ase ruβa ase si ase
 marã e-?i=pe **ase** r-uβ-a **ase** Ø-si-Ø **ase**
 how 3-say-Q **our** R₁-father-REF **our** R₁-mother-REF **our**
 rerokara supe?
 r-erok-sar-a supe
 R₁-baptize-NMLZ_{AG}-REF to
 ‘How does our father, our mother say to our.’ (Araújo, 82)

Ase is often interpreted as an impersonal form, as in (120)

- (120) a. Oimoeteβe ase amõamõ ?ara, ipupe
 o-i-mo-ete-βe **ase** amõ-amõ ?ar-a i-pupe
 3-R₂-much-also **we.all** some-RED day-REF R₂-in
 oporaβiβkie?ima
 oporaβiβki-e?im-a
 work-PRIV-GER
 ‘Do we / Does one honor other days by not working (on these
 days)?’ (Araújo, 12v)
- b. Nane rako ase jeupiri iβakipene,
 nane rako **ase** je-upiri iβak-ipe=ne
 thus EV_{FH} **we.all** RFLX-elevate sky-EPEN-LOC=FUT
 oposijusu rejtikire
 o-posij-usu r-ejtik rire
 CORF-weight-big R₁-throw after
 ‘Thus, indeed, one/we all will rise to heaven after throwing away
 his/our burden.’ (Araújo, 169v)

Nonetheless, Anchieta (1595, 36v) explicitly says that the impersonal construction requires the first person plural inclusive. Examples are given in (121).

- (121) a. Jajuka
 ja-i-juka
 1PL.INCL-R₂-kill
 ‘One kills (it/him/her/them)’ (AA, 36v)
- b. Nomenari: emonã tekoarwera
 n-o-menar-i emonã t-eko-ar-wer-a
 NEG-3-marry-NEG thus R₂-be-NMLZ_{AG}-PST-REF
 jaipe?a
 ja-i-pe?a
 1PL.INCL-R₂-separate
 ‘He did not marry: (may) one separate the one who acted this way
 (marry against his/her own will)’ (Araújo, 128)

4.3.2 Relational Markers

In Tupían studies, the morphemes referred to by the name of RELATIONAL are a distinct feature of five of the ten branches of the family (Rodrigues and Cabral 2012, 496-499)¹⁴. Besides Tupían languages, they are found in some Jê languages (Cabral and da Costa 2004; Ribeiro 2004)¹⁵, in Jabuti and Chiquitano (Ribeiro 2011), and in Cariban languages (Rodrigues 2009). These morphemes define, according to their allomorphs, the noun classes in Tupinambá¹⁶. Possessed roots, i.e., those that combine with possessor indexes, are members of either of Classes I or II. Class III is that of unpossessed roots, whose members cannot be combined with possessor indexes. Table (4.5)¹⁷ illustrates all three root classes:

¹⁴If one considers their original status as root-initial segmental alternation, then it would be a feature of six branches (see Drude and Meira 2019).

¹⁵See Salanova (2009) for an alternative view.

¹⁶See Jensen (1998a) for a reconstruction of noun classes in PTG. See also Meira and Drude (2013).

¹⁷The abbreviation with the numbers (R₁, R₂, etc.) follows Rodrigues (1996a); Cabral (2000).

Class	R ₁	R ₂	R ₄	R ₃	Examples
Ia	∅-	i-	∅-	o-	<i>akaŋ</i> ‘head’, <i>si</i> ‘mother’, <i>sem</i> ‘leave’, <i>ker</i> ‘sleep’
Ib	∅-	i-	m-	o-	<i>po</i> ‘hand’, <i>posaj</i> ‘medicine’, <i>pita</i> ‘stay’
IIa	r-	s-	t-	o-	<i>esa</i> ‘eye’, <i>asem</i> ‘yell’, <i>enone</i> ‘in front of’, <i>e?ō</i> ‘die’
IIb	r-	t-	t-	o-	<i>uβ</i> ‘father’, <i>a?ir</i> ‘son’, <i>ur</i> ‘come’, <i>uβ</i> ‘lie down’
IIc	r-	s-	∅-	o-	<i>ok</i> ‘house’, <i>u?uβ</i> ‘arrow’
IId	r-	s-	(V → ∅-)	o-	<i>ape</i> ‘path’, <i>(e)kuj</i> ‘gourd’, <i>(e)pinō</i> ‘fart’
III	–	–	–	–	<i>kwarasi</i> ‘sun’, <i>iβirá</i> ‘tree’, <i>tapi?ir</i> ‘tapir’

Table 4.5: Tupinambá relational markers: R₁ marks contiguity, R₂ marks non-contiguity, R₃ indicates coreference, and R₄ indicates that the possessor is human

Some roots belonging to class IId have an alternate form with an initial *e*, which is a trace of a prefix of alienable possession, still found in Mundurukú (Gomes 2006), Mawé and Awetí (Meira and Drude 2013), among others.

	1SG	3
Ia	ʃe=∅-akaŋ ‘my head’	i-akaŋ ‘his head’
Ib	ʃe=∅-pó ‘my hand’	i-pó ‘his hand’
IIa	ʃe=r-esá ‘my eye’	s-esá ‘his eye’
IIb	ʃe=r-uβ ‘my father’	t-uβ ‘his father’
IIc	ʃe=r-ók ‘my house’	s-ok ‘his house’
IId	ʃe=r-ekuj ‘my gourd’	s-ekuj ‘his gourd’

Table 4.6: TUP noun classes. Examples in first person singular and third person

Functionally, relational markers mark the contiguity (R₁) or non-contiguity (R₂) of a head – any possessed root or a postposition – and its dependent (Rodrigues 1996a; Cabral 2000)¹⁸. The relational of contiguity (R₁) thus has

¹⁸For the problem of the origins of the relational morphemes and their development see Jensen (1998a); Cabral (2000); Gildea (2002); Meira and Drude (2013).

a twofold function: the flagging of contiguity and creating a dependency relation, while the relational of non-contiguity signalizes the absence of a dependent in the syntagma.

Many authors working on TG languages have treated the relational of non-contiguity (R_2) as an index of third person argument (Jensen 1999; Couchili et al. 2002; Neiva Praça 2007; Rose 2011; Copin 2012; Magalhaes and de Matos 2014). I consider this view to be wrong. The distribution of first and second person bound indexes clearly shows that they must be traced back to free pronouns, i.e., personal forms in reference phrase positions (Queixalós 2022). However, this does not apply to *i-* (R_2), since it is plausible that it already existed when the first and second person free indexes occurred internally and bound (see Gildea 2002). See examples (125) and (126).

The contiguity and non-contiguity of a postposition and its dependent are illustrated in (122) with R_1 and R_2 respectively.

- (122) a. Nerese
 ne=**r**-ese
 2SG= R_1 -because
 ‘Because of you.’ (see Figueira, 124)
- b. Sese
 s-ese
 R_2 -because
 ‘In him.’ (see Araújo, 60)

The contiguity or non-contiguity of possessor and possessed roots is illustrated in (123). If the possessor (dependent) is not the preceding element, in that it is outside the constituent or absent, then R_2 is used:

- (123) a. Isi
 i-si-∅
 R_2 -mother-REF
 ‘His mother.’ (Poemas, 184)

- b. Tuβa
 t-uβ-a
 R₂-father-REF
 ‘His father.’ (see Araújo, 5)
- c. fesi
 fe=∅-si-∅
 1SG=R₁-mother-REF
 ‘My mother.’ (Araújo, 33v)
- d. Soka
 s-ok-a
 R₂-house-REF
 ‘His house.’ (FA, 78)

With a contiguous nonpronominal possessor, an RP, R₁ is employed (124):

- (124) a. Anas roka
 Anas r-ok-a
 Anas R₁-house-REF
 ‘Anas’ house.’ (see Araújo, 55)
- b. K^warasi semaβa koti
 [k^warasi-∅ ∅-[sem-aβ-a ∅-koti-∅]]
 sun-REF R₁-exit-NMLZ-REF R₂-side-REF
 ‘The side of the rising of the sun.’ (Araújo, 3)
- c. Pero rekoaβa
 Pero r-eko-aβ-a
 Pero R₂-act-NMLZ-REF
 ‘Pero’s job.’ (see Araújo, 5)

The examples in (125), (126), and (127) are illustrative¹⁹. In the a) examples, the prefix of contiguity (*r-*) indicates that the dependent is adjacent (immediately to the left) to the head, while the b) examples have the non-contiguous

¹⁹The examples below, involving a possessive construction with a non-contiguous relational and first or second person, are rarely attested. I suppose this construction was avoided due to its complexity, which explains its mention in Anchieta’s grammar but its scarcity in the texts.

marker (*s-*, *i-*) which indicates that the dependents are not adjacent to the head, thus implying a constituent discontinuity. The square brackets indicate a syntagma. Note that in (125b) and (126a,b), the clitic markers attach to the RP.

- (125) a. jePedro rawsume
 [$\text{je=Pedro r-awsu}\beta\text{-me}$]
 1SG=Pedro \mathbf{R}_1 -love-CLM
 ‘Because I love Pedro.’ (Arte, 37)
- b. Pedro jesausume
 [$\text{Pedro je=s-awsu}\beta\text{-me}$]
 Pedro 1SG= \mathbf{R}_2 -love-CLM
 ‘Because I love Pedro.’ (Arte, 37)
- (126) a. jePedro jukáreme
 [$\text{je=Pedro } \emptyset\text{-juka-reme}$]
 1SG=Pedro \mathbf{R}_1 -kill-CLM
 ‘If I kill Pedro.’ (Arte, 37)
- b. Pedro fejjukáreme
 [$\text{Pedro je=i-juka-reme}$]
 Pedro 1SG= \mathbf{R}_2 -kill-CLM
 ‘If I kill Pedro.’ (Arte, 37)
- (127) a. $\text{K}^w\text{ese ka?a rupi owatá}\beta\text{o Pedro}$
 $\text{k}^w\text{ese ka?a-}\emptyset\text{ r-upi o-wata-}\beta\text{o [Pedro]}$
 yesterday forest-REF \mathbf{R}_1 -POSP 3-walk-GER Pedro
 ropári
 \mathbf{R}_1 -opar-i]
 \mathbf{R}_1 -get.lost-NFOC
 ‘Yesterday, Pedro got lost walking through the forest.’ (FA, 95)
- b. $\text{Kwese ka?a rupi Pedro owatá}\beta\text{o}$
 $\text{k}^w\text{ese ka?a-}\emptyset\text{ r-upi [Pedro] o-wata-}\beta\text{o}$
 yesterday forest-REF \mathbf{R}_1 -POSP Pedro 3-walssk-GER

sopari

[s-opár-i]

R₂-get.lost-NFOC

‘Yesterday, Pedro got lost walking through the forest.’ (FA, 95)

In (128a), the relational of contiguity \emptyset - signalsizes that *feruβa* is the dependent adjacent to the head *mongetaw*. In (128b), the dependent of the head *mongetaw* is not adjacent, signalized by the relational of non-contiguity *i*. Its dependent, *feruβa*, is in the first position. The head is given underlined, while the brackets mark the syntagma:

- (128) a. Koriteʔĩ Pedro [feruβa mongetaw]
 koriteʔĩ Pedro je=r-uβ-a \emptyset -mongeta-w
 now Pedro 1SG=R₁-father-REF **R**₁-talk-NFOC
 ‘It was now, that Pedro spoke to my father.’ (FA, 96)
- b. feruβa koriteʔĩ Pedro [imongetaw]
 je=r-uβ-a koriteʔĩ Pedro **i**-mongeta-w
 1SG=R₁-father-REF now Pedro **R**₂-talk-NFOC
 ‘Pedro has now spoken to my father.’ (FA, 96)

Of syntactic relevance is the fact that contiguity implies that dependent and head are inside a syntagma (129), while the absolute relational (**R**₄) implies the opposite, so that in (130), *t-aʔira* functions as an apposition.

- (129) Tupã raʔira
 [Tupã **r**-aʔir-a]
 God **R**₁-son-REF
 ‘Son of God.’ (DC, 166)

- (130) Tupã taʔira
 [Tupã] [**t**-aʔir-a]
 God **R**₄-son-REF
 ‘God (the) son.’ (DC, 131)

In certain syntactic environments, it becomes impossible to identify the possessor, as there is no referent, but there is a dependency relation with human beings in general expressed syntactically (Rodrigues 1996a, 96). This is signaled by (R⁴). Compare the pairs in (131) and (132).

(131) a. Moʔir
 m-poʔir
 R₄-necklace
 ‘Necklace (of a person).’ (see Léry, 346)

b. Poʔir
 poʔir
 necklace
 ‘Necklace.’ (VLB, II, 14)

(132) a. Tete
 t-ete
 R₄-body
 ‘Body (of a person).’ (see Léry, 346)

b. Sete
 s-ete
 R₂-body
 ‘Its body (of an animal).’ (VLB, II, 14)

The relational R₃ indicates that the dependent of a head is coreferential with the subject of the main clause, as in (133). Compare (133b) and (134) for coreferential and non-coreferential possession.

(133) a. Oatiʔiβa ri krusa osupi
 o-atiʔiβa ∅-ri krusa-∅ o-s-upir
 R₃-shoulder R₁-POSP cross-REF 3-R₂-lift
 ‘He lifts the cross on his own shoulder.’ (Poemas, 122)

b. Oaʔira
 o-aʔir-a
 R₃-son-REF

‘His own son.’ (see Araújo, 25v)

- (134) Taʔira
 t-aʔir-a
 R₂-son-REF
 ‘His (someone else’s) son.’ (see Araújo, 14v)

Another use of the relational marker of non-contiguity is to index the under-goer third person argument of a transitive verb, as in (135a):

- (135) a. Aik^waβ
 a-i-k^waβ
 1SG-R₂-thank
 ‘I thanked him.’ (VLB, I, 23)
- b. Eresekar
 ere-s-ekar
 2SG-R₂-seek
 ‘You seek him.’ (see D’Evreux, 144)

Tupinambá lacks third person possessor markers and pronouns²⁰. The relational of non-contiguity (R₂) is not a third person marker, in spite of its possible origin as a third person index (see Rodrigues and Cabral 2012). The independent pronoun for third person is actually a demonstrative (see Section 8.1). One more argument for the relational analysis is the occurrence of relational morphemes in languages of the Jê and Cariban families, where these are best viewed as relationals rather than third person indexes (see Rodrigues 2009)²¹.

The possessor character of R₂ has been suggested by Rose (2018), and as noted in Cabral (2001b), alternative analyses to the relational hypothesis do not agree among themselves as far as its status is concerned. Meira and Drude (2013) also argue in favor of the non-pronominal character of the relational

²⁰Regarding the lack of third person in languages, see Benveniste (1971); Bhat (2005); Siewierska (2005, 2009).

²¹For arguments against the presence of relationals in the Jê family, see Salanova (2004, 2009).

of non-contiguity, arguing that *i* and *o* are not clitics like members of Set I, which can be etymologically associated with the independent pronoun (see Set I and the independent pronouns in Table 4.5).²² An analysis by Gildea (2002) suggests that Proto-Tupí-Guaraní displays marking patterns stemming from competing pronominal systems, the oldest of which are to be seen in the *i* (R^2) and *o* (R^3) morphemes, indicating that the coreferential and non-coreferential opposition was already present in the language at an early stage (see Meira and Drude 2013, 5, note 4). He (Gildea) believes that this (original) system has been lost, leaving us unable to reconstruct it, and the relational morpheme(s) would be traces of this partially lost system (see also Jensen 1990a; Schleicher 1998).

4.3.3 Adverbs

Adverbs are one-place predicates which take a logical structure or part of a logical structure as their argument. The sentence *yesterday John gave Patty a flower in the garden* would be represented as in (136):

- (136) *give* [**do'** (w, \emptyset)] CAUSE [BECOME **have'** (x, y)]
in be-in' (x, y)
yesterday **yesterday'** (x)

$\langle_{\text{IF DEC}} \langle_{\text{TNS PST}} \langle \text{yesterday' (be-in' (garden, [do' (John, } \emptyset)] \text{ CAUSE [BECOME have' (Patty, flower)]})}) \rangle \rangle \rangle \rangle$

Although this chapter intends to explicate the question of lexical categories, treating mainly what is traditionally called ‘noun’, ‘adjective’, and ‘verb’, it is important to look at ‘adverbs’ because of their importance as modifiers at different levels of the LSC, and because of their lexical manifestations in TUP.

²²Gomes (2006, 39-40) also defends the relational hypothesis for Mundurukú, asserting that its relational morphemes have a complex allomorphy and are capable of explaining syntactic phenomena which would be impossible to explain if they were analyzed as third-person prefixes. One such explanation is given in Meira and Drude (2013, 19), where it is shown that *-iát*, but not *-at*, may nominalize a clause.

Example (138) shows one adverb in the periphery of the clause and one adverbial PP in the periphery of the core.

- (138) K^warasi nipo oβera putunusu k^waβire
 k^warasi-∅ nipo o-βeraβ putun-usu-∅ k^waβ-rire
 sol-REF certainly 3-shine night-big-REF pass-POSP

‘The sun certainly shines after the great night passes.’ (Poemas, 142)

Adverbs may modify different layers of the LSC. Clausal modifiers include speech act modifiers (*honestly*), evidential or epistemic modifiers (*evidently*, *allegedly*), and judgment (*appallingly*, *unfortunately*). Core modifiers include temporal modifiers (*yesterday*, *tomorrow*) and manner modifiers (*quickly*, *slowly*, *deliberately*, *carefully*, *violently*). Nuclear modifiers include aspectual modifiers (*completely*, *continuously*).

Since adverbs modify different levels of the LSC, they tend to appear closer to the layer they modify (see Van Valin Jr 2005, 19-21).²³

²³This seems to be connected with the tendency to minimize dependency length (see Liu 2008 and Futrell et al. 2015).

Chapter 5

Basic Clause Patterns

Tupinambá is a head-final and head-marking language, meaning that the core arguments are cross-referenced on the arguments by bound indexes which are indexed to the predicate in the SOV word order¹ (see Van Valin Jr 1987, 2013). The subject of intransitive predicates (S) and the subject of transitive predicates (A) are indexed by the same set of person indexes in independent clauses, therefore exhibiting a nominative-accusative alignment pattern. Independent pronouns or RPs semantically related to the arguments can seemingly – because not all possible orders are attested – appear in any order in relation to the core, a fact already noted by Anchieta (1595, 37) (see Section 5.6). Adjuncts tend to follow the core, but this is not obligatory.

RRG defines transitivity according to the number of macrorole arguments a predicate has, not the number of syntactic arguments (Van Valin Jr 2005, 60-67). Predicates with one core argument are termed M-intransitive, and those with two or three (ditransitive) are M-transitive (see Section 7.1).² The head-

¹The order of noun-postposition and genitive-noun is typologically consistent with OV order, while the order of noun-adjective is only slightly inconsistent with the other parameters (see Dryer 1997b).

²Some languages have atransitive predicates, i.e., predicates with a semantic and syntactic value of 0 (zero) (see Van Valin Jr 2005, 64).

final and head-marking character of TUP is illustrated in Figure 5.1, where the core, containing the nucleus and the core arguments, exhibits the fixed order of these elements.

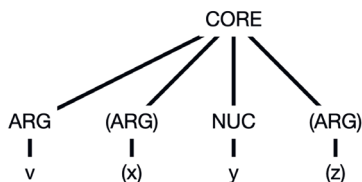


Figure 5.1: A core template with three core arguments for TUP

In the template in Figure 5.1, the core contains three arguments. The first argument (v) corresponds to the sole argument of an intransitive verb, while the second argument (x) corresponds to the undergoer of a transitive verb. Both v and x are pre-nuclear. The third argument (z) corresponds to the non-macrorole core argument (see Section 5.3), which normally occurs post-nuclear.

5.1 M-intransitive Verbs

Each semantic verb class described in Section 3.3 has M-intransitive examples. For activity verbs, the single macrorole is the actor (A). Stative verbal predication is not common in Tupinambá. Intransitive states are expressed mostly by non-verbal predication, as illustrated by the nominal predicates in (139) (see Section 5.5).

(139) State: undergoer PSA

- a. ʃekane?õ
 ʃe=∅-kane?õ
 1SG=R₁-tiredness
 ‘I am tired.’ (VLB, I, 65)
exist' ([have.as.part'(1SG, kane?õ)])

- b. ʃeakim
 $\text{ʃe}=\emptyset\text{-akim}$
 1SG=R₁-wet
 ‘I am wet / there is my being wet.’ (VLB, II, 40)
exist’ ([**have.as.part’**(1SG, akim)])

Examples of activity verbs are given in (140); examples of achievement are given in (141). Examples of accomplishment are given in (143). A semelfactive active verb is shown in (142).

(140) Activity verbs: actor PSA

- a. Ajākatune
 a-jā-katu=ne
 1SG-run-INTS=FUT
 ‘I will run a lot.’ (AT, 25)
do’ (1SG, [**run’**(I)])
- b. Jū rupa awata
 $\text{jū-}\emptyset \text{ r-upi a-wata}$
 field-REF R₁-through 1SG-walk
 ‘I walked through the field.’ (FA, 123)
do’ (1SG, [**walk’**(I)])
- c. $\text{Eraso koβaʔe neruβape}$
 $\text{e-era-so koβaʔe ne=r-uβ=pe}$
 2SG.IMP-SCAU-go DEM 2SG=R₁-father=LOC
 ‘Take this to your father.’ (FA, 121)
 [**do’**(2SG, \emptyset)] CAUSE [**do’**(koβaʔe, [so’(koβaʔe)])] & INGR **be-toward’**(koβaʔe, neruβa) & INGR **be-toward’**(2SG, neruβa)

(141) Achievement verbs: undergoer PSA

- a. iβira opuruk
 $\text{iβira-}\emptyset \text{ o-puruk}$
 tree-REF 3-snap
 ‘The tree snapped.’ (see VLB, I, 127)
 INGR **snapped’**(3sg [the tree])

- b. Oʔar mune
 o-ʔar mune
 3-fall trap
 ‘The trap fell.’ (VLB, I, 63)
 INGR **fall**’ (3sg [mune])
- c. Open iβira
 o-pen iβira-∅
 3-break stick-REF
 ‘The stick broke.’ (see VLB, II, 92)
 INGR **do**’ (3[iβira], [**break**’ (3)[iβirá]])

(142) Semelfactive verb: actor PSA

Ajemoesaβik
 a-je-mo-esa-βik
 1SG-RFLX-CAUS-eye-blink
 ‘I blinked.’ (VLB, I, 79)
 SEML **do**’ (1SG, [**blink**’ (I)])

(143) Accomplishment verb: undergoer PSA

- a. Aβa omanõ
 aβa-∅ o-manõ
 man-REF 3-die
 ‘A man died.’ (Fig., Arte, 69)
 BECOME **dead**’ (3sg [man])
- b. Atujuk
 a-tujuk
 1SG-rotten
 ‘I rotted.’ (VLB, I, 38)
 BECOME **rotten**’ (1SG)

Stative predicates in TUP can be expressed by two different constructions: verbal predication or existential predication (see Dixon 1979, 1994). Table (5.1) illustrates pairs of antonyms that are expressed by different constructions. Note that *uʔu* ‘cough’ and *pitu* ‘breathe’ are not seen as activities in TUP (see

Holisky 1987), which are always expressed with Set I indexes (see Table 4.3). While *posij* ‘be heavy’ is expressed through nominal syntax, *βεβuj* ‘fluctuate, be light’ is expressed through Set II markers (verbal syntax).

		Set II markers	Set I markers
heavy	posij		x
light	βεβuj	x	
dry	tiniŋ	x	
wet	akim		x
hard	atã		x
soft	memek	x	
remember	maʔenwar		x
cough	uʔu		x
steal	mona	x	
vomit	weʔen	x	
breathe	pitu		x

Table 5.1: Some active and stative roots in Tupinambá. Colored pairs are antonyms that require different markers for active and stative forms

Examples (144) in TUP and (145) in English are illustrative of the language-specific character of the semantics of lexical roots. The former requires a non-prototypical predication (see Section 5.5), while the latter is expressed through verbal predication. The lexical decomposition of these predicates is clearly different (see Section 3.3).

- (144) ʃeuʔu
 ʃe=∅-uʔu
 1SG=R₁-cough
 ‘I cough (there is my cough).’ (VLB, I, 62)
be’ (I, [cough’])

- (145) English I cough
exist’ ([have.as.part’(1SG,cough)])

Tupinambá and other Tupí-Guaraní languages have been classified as languages of the split-intransitive type (see Schmidt-Riese 1998; Jensen 1990a; Rose 2009) or active-stative (see Seki 1990; Jensen 1998a; Cabral 2009) type. According to this view, intransitive roots receive different argument indexes depending on the *Aktionsart*, active or stative (see Seki 1990). Note that in Table 5.1 some of the roots that are perceived as activities, e.g., in English, such as *remember*, *cough*, *breathe*, are states in TUP because they are perceived as independent of volition or control by the subject (see Dixon 1994, 78-83).

A special case concerns the verb *iko*, which can be translated into various meanings in English, such as ‘be, live, act, behave, happen’. While these meanings are semantically stative, *iko* – glossed as ‘be’ – is always used with Set I indexes, as in the examples below:

- (146) a. Aβátepe ojko jeoja...?
 aβa-te=pe o-**iko** je=∅-oja
 person-FOC=Q 3-**be** 1SG=R₁-similar
 ‘Who is like me ...?’ (Teatro, 20)
- b. Enépe ereiko?
 ene=pe ere-**iko**
 you=Q 2SG-**be**
 ‘How are you? (lit. you? you are/live)’ (VLB, II, 113)
- c. Peikoete peroβajarape
 pe-**iko**-ete pe=r-oβajar-a-pe
 2PL-**be**-INTS 2PL=R₁-enemy-REF-LOC
 ‘Be strong towards your enemies.’ (Araújo, 89)

5.2 M-transitive Verbs

The semantic verb classes presented above can all be M-transitive. In each case, the verb has an actor macrorole and an undergoer macrorole. Examples of stative verbs are given in (147), and activity verbs are given in (149). An example of an M-transitive (causative) achievement verb is given in (151),

and an example of an M-transitive (causative) accomplishment verb is given in (152). Examples of activity verbs are given in (149).

- (147) State verb: actor PSA, undergoer DCA
 Peipousuβ imē
 pe-i-pousuβ imē
 2PL-R₂-fear NEG
 ‘Fear it not.’ (Araújo, 4)
 NOT **fear'** (2sg, 3sg)
- (148) Naik^waβi aʔe aβa
 n-a-i-k^waβ-i aʔe aβa-∅
 NEG-1SG-R₂-know-NEG DEM man-REF
 ‘I do not know this man.’ (Araújo, 57)
 NOT **know'** (1SG, 3[aʔe aβa])
- (149) Activity verb: actor PSA, undergoer DCA
 Asasa
 a-s-asaβ
 1SG-R₂-cross
 ‘I cross(ed) it.’ (see VLB, II, 67)
do' (1SG, [**cross'** (1SG, it)])
- (150) Eresaʔangipe aβamemuã?
 ere-s-aʔang-i=pe aβa-memuã-∅
 2SG-R₂-imitate-EPEN=Q man-evil-REF
 ‘Did you imitate the evil men?’ (DC, II, 100)
do' (2SG, [**imitate'** (1SG, 3[aβamemuã])])
- (151) Causative achievement verb: actor PSA, undergoer DCA
 Asapi jū
 a-s-api jū-∅
 1SG-R₂-burn field-REF
 ‘I burned the field.’ (VLB, I, 140)
 [**do'** (1sg, ∅)] CAUSE [INGR **burnt'** (1sg [jū])]

- (152) Causative accomplishment verb: actor PSA, undergoer DCA
 Aimoiku
 a-i-mo-iku
 1SG-R₂-CAUS-liquid
 ‘I melted it.’ (VLB, I, 95)
 [dø’ (1sg, Ø)] CAUSE [BECOME **melted**’ (i-)]

The complete paradigm of a transitive verb is given in Table (5.2):

	Verb Conjugation	Translation
1SG → 3	a-i-akaβ	I fight/fought him/her/it/them
2SG → 3	ere-i-akaβ	you fight/fought him/her/it/them
3 → 3	o-i-akab	he/she/it/they fight/fought him/her/it/them
1PL.INCL → 3	ja-i-akaβ	we fight/fought him/her/it/them
1PL.EXCL → 3	ore-i-akaβ	we fight/fought him/her/it/them
2PL → 3	pe-i-akaβ	you fight/fought him/her/it/them
1 → 2SG	oro-akaβ	I/we fight you.SG
1 → 2PL	opo-akaβ	I/we fight you.PL

Table 5.2: Example of an M-transitive verb paradigm

M-transitive verbs cannot have a first person as an undergoer, or a second person in the case of a non-first person actor.³ These cases require a different construction, namely existential predication, as given in (153). An example such as (153a), literally ‘there is my calling’, would be ambiguous regarding who the caller could be, as it could be either the second or third person: ‘there is my killing by you/him’. In order to disambiguate the meaning, the oblique *jepe* (153b) is used in the case of a second person singular, and *pejepe* is used in the case of second person plural (153c). The fact that the second person plural takes *pe=*, which is associated with the second person plural (see Table 4.3), suggests that *jepe* (and *pejepe*) could be oblique markers/pronouns, whose ending could be associated with a locative case ending (see Section 7.8).

³For a similar pattern in other language families and discussion of the phenomenon, see Bresnan et al. (2001); Aissen (1999); Jelinek and Demers (1994, 1983).

- (153) a. \int erenõj
 \int e=r-enõj
 1SG=R₁-call
 ‘He calls me.’ (see VLB, II, 50)
- b. \int ejuka jepe
 \int e=Ø-juka **jepe**
 1SG=R₁-kill **OBL**
 ‘You kill me.’ (Teatro, 78)
- c. \int ejuka pejepe
 \int e=Ø-juka pe-jepe
 1SG=R₁-kill 2PL-OBL
 ‘You kill me.’ (Arte, 37)

A hierarchy, $1 > 2 > 3$, has been postulated for TG languages (see Jensen 1990a; Monserrat and Soares 1983; Magalhães 2010; Seki 1990; Rose 2009, 2015b) whereby the relative ranking of A and U determines which arguments are indexed on the predicate. The hierarchy predicts that if $A > U^4$, both arguments are cross-referenced, as in (154), where $1 > 3$.

- (154) Asawsuβ
 a-s-awsuβ
 1SG-R₂-love
 ‘I love him.’ (see Poemas, 102)

In the case of A being lower than U, only the highest argument is indexed, as in (155).

- (155) \int erep^jak
 \int e=r-ep^jak
 1SG=R₁-see
 ‘He/you see(s) me / (there) is sight of me / there is my sight.’ (Arte, 37v)

⁴See the 3.18 in Section 3.3.2.

It is possible to simplify the system by considering that TUP transitive constructions only exist with third person undergoers⁵ or second person undergoers in the case of a first person actor (portmanteau indexes, i.e., Set IV indexes in Table 4.3). All other cases are expressed with one nominal (possessor) argument only, through nonverbal predication (see Section 5.5). This is a case of Occam's Razor, because there is no reason to postulate two functions for the markers of Set I: that of possessor and that of absolutive markers. Similarly, there is no reason why the same construction must have two functions, that of existential predication and that of a transitive construction with a suppressed subject due to a hierarchical constraint. Note how (156) is a genitive construction, or more precisely, two genitives: 'your killing' and 'my lord' (literally 'my lord's killing of you'). This interpretation was first suggested by Dietrich (2001, 2017a) for other TG languages (see Rodrigues 2011a for Tupinambá).

- (156) Nejuka fejara
 ne=∅-juka-∅ je=∅-jar-a
 2SG=R₁-kill-REF 1SG=R₁-lord-REF
 'My lord kills you (lit. my lord's killing of you).' (Arte, 12v)

Another argument against the intransitive split due to verbal *Aktionsart* is the fact that many stative predicates are cross-referenced with markers of Set II indexes (see table 4.3). The examples in (157) show stative verbs with 'active markers' because specific stative meanings are expressed by active verbs. These are rare, but there are a few examples.

- (157) a. Aβeβuj
 a-βeβuj
 1SG-be.light
 'I am light / I float.' (VLB, II, 21)

⁵There seems to be an intransitive bias in OV languages, i.e., languages in which the object precedes the verb (see Hawkins 2014, 158,180), Nichols et al. 2004. See also (Progovac 2015; Queixalós 2010). Although the topic is not discussed in this work, it could be related, diachronically, to the intransitive bias found in TUP, and somehow related to relational morphemes. This is something scholars of Tupí-Guaraní should consider as a research topic.

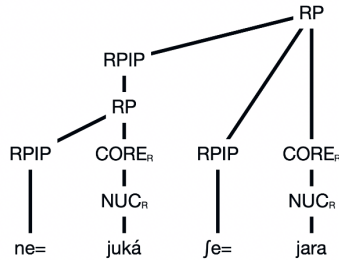


Figure 5.2: Embedded possessive construction

- b. Ain
a-in
1SG-be.still
'I am seated.' (FA, 58)
- c. Ajuß
a-juß
1SG-lie
'I am lying.' (FA, 57)
- d. Aiko
a-iko
1SG-be
'I am / I exist / I act.' (FA, 59)

5.3 Ditransitive Verbs

Some M-transitive verbs have a semantic valency of three, and are thus ditransitive verbs, but only two of the three arguments are macroroles (see Van Valin Jr and LaPolla 1997, 145-154, Van Valin Jr 2001a and Van Valin Jr 2005, 60-67). The third argument of ditransitive verbs in Tupinambá is a non-macrorole indirect core argument because it takes the dative case, as in (158), or a non-macrorole oblique core argument, as in (159), because it is adpositionally marked. Pronouns can receive the dative case in TUP, but RPs require a postposition (see Section 7.8.1).

- (158) a. Maʔepe Tupã ojmeʔeŋ aseβe iβakipene?
 maʔe=pe Tupã-∅ o-i-meʔeŋ **ase-βe** iβak-pe=ne
 thing=Q God-REF 3-R₂-give **we-DAT** heaven-POSP=FUT
 ‘What will God give us in heaven?’ (CA, 27)
 [do’ (Tupã, ∅)] CAUSE [BECOME **have’** (ase,i)]
- b. Tameʔene pira ruβa enéβo
 t-a-meʔeŋ=ne pira-∅ r-uβ-a **ene-βo**
 HORT-1SG-give=FUT fish-REF R₁-egg-∅ **you-DAT**
 ‘I shall give you fish eggs.’ (Teatro, 46)

Note that the only possibility of a non-macrorole direct core argument as an RP in TUP is with pronouns, because they can receive the dative case. Non-pronominal non-macrorole core arguments are always indirect, because these are marked by unstressed suffixes, as in (158) (see Section 7.4).

- (159) a. Apek^waβeʔeŋ kunumĩ supe
 a-pe-k^waβeʔeŋ kunumĩ-∅ **supe**
 1SG-path-show child-REF **to**
 ‘I show the way to the children.’ (see VLB, I, 152)
 [do’ (I, ∅)] CAUSE [BECOME **see’** (kununĩ, pe)]⁶
- b. Aimoin uʔuβa supe
 a-i-mo-in uʔuβ-a **supe**
 1SG-R₂-CAUS-place arrow-REF **with**
 ‘I point to him with the arrow (lit. I put the arrow towards him).’
 (VLB, I, 39)
- c. Aik^waβeʔeŋ X supe
 a-i-k^waβ-meʔeŋ X **supe**
 1SG-R₂-know-give X **to**
 ‘I offered it to X.’ (VLB, II, 54 modified)⁷
- d. Ereimomeʔupe aβa aŋaipaŋjemim-a
 ere-i-momeʔu=pe aβa-∅ ∅-aŋaipaβ-jemim-a
 2SG-R₂-tell=Q man-REF R₁-evil-hide-REF

⁶Operators not included.

⁷The source does not specify a recipient, using an abbreviation instead.

ik^wapare?ima supe?
 i-k^waβ-ar-e?im-a **supe**
 R₂-know-NMLZ-PRIV-GER **about**

‘Did you tell someone who did not know it about one’s hidden evil deeds?’ (Araújo, 108)

The closest to a dative shift alternation (see Haspelmath 2015, Van Valin Jr 2005, 60-62, 112-115) found in TUP occurs when the undergoer is not an argument index but an RP which must be a possessed root. The possessed root is incorporated with the relational of non-contiguity (R₂) because its possessor is not adjacent. The relational of non-contiguity (R₂) will be that of its noun class membership (see Section 4.3), in addition to *i-* or *s-* used to cross-reference the undergoer in transitive verbs. This is clear from example (161a), where the (R₂) is *t-*, indicating that the possessor is not the preceding element.

In (160), the recipient is coded by a PP (dative), but in (161), the lexical root is incorporated and the undergoer (recipient) is neither marked by case nor is it coded by a postposition. In other words, this type of incorporation advances an oblique argument into the case position vacated by the incorporation (see Mithun 1984). This difference points to a choice (marked or unmarked undergoer) regarding the undergoer, but there does not seem to be a semantic difference (see Van Valin Jr 2001a). This type of incorporation is only attested with ditransitive verbs that represent transfer of possession⁸.

- (160) Aime?ej aβa supe
 a-i-me?ej aβa-∅ supe
 1SG-R₂-give man-REF to
 ‘I gave it to the men.’ (Teatro, 48)
 [**do**’ (1SG, ∅)] CAUSE [BECOME **have**’ (i, aβa)]

In the examples in (161), the undergoer argument can be incorporated because it is a possessed noun, which is the reason why the predicate is transitive (two core arguments).

⁸The same has been observed in Oneida, an Iroquian language (see Michelson 1991).

- (161) a. Ataʔimeʔeŋ Pedro
 a-t-aʔir-meʔeŋ Pedro
 1SG-R₂-son-give Pedro
 ‘I give Pedro a son.’ (AA, 50v)
 [doʔ (1SG, ∅)] CAUSE [BECOME haveʔ (Pedro, t-aʔir)]
- b. Aiaomeʔeŋ Pedro
 a-i-aoβ-meʔeŋ Pedro
 1SG-R₂-clothes-give Pedro
 ‘I give clothes to Pedro / I clothe Pedro.’ (AA, 50v)

Comparing (161) with (162), the difference between them is the presence of another relational marker in the latter examples which indexes a new argument, namely the undergoer. The first relational marker indexes a macrorole core argument (the undergoer), while the second relational marks the non-contiguity of the possessed root and its possessor (the non-macrorole core argument).⁹

- (162) a. Aitaʔimeʔeŋ Pedro
 a-i-[t-aʔir]-meʔeŋ Pedro
 1SG-R₂-R₂-son-give Pedro
 ‘I give him Pedro’s son.’ (AA, 50v)
 [doʔ (1SG, ∅)] CAUSE [BECOME haveʔ (i, [Pedro t-aʔir])]
- b. Aijaomeʔeŋ Pedro
 a-i-[i-aoβ]-meʔeŋ Pedro
 1SG-R₂-R₂-clothes-give Pedro
 ‘I give him Pedro’s clothes.’ (AA, 50v)

The representation of (162a) is given in Figure 5.3.

⁹This type of *possessor-stranding* (Baker 1988, 96-105) seems to be found in other TG languages such as Araweté (Solano 2010, 330), where it is similar to the TUP construction, and in Tenetehara, with a somewhat different construction (see Castro and Camargos 2021). In other languages, e.g., Kamajura (Seki 1990, 143-144), Guajá (Magalhães 2010, 198), Tekó (Rose 2011, 266-269), and Wayampi (Copin 2012, 343-344), it is less clear if there is possessor-stranding, because the possessed element is not marked by a non-contiguity marker, as in TUP.

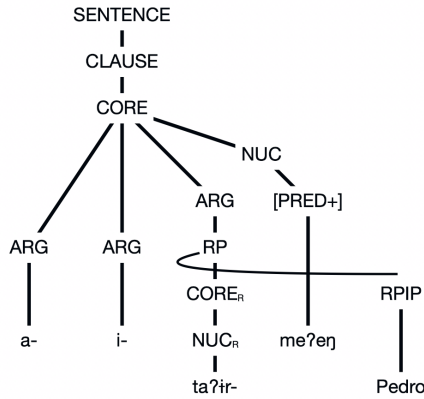


Figure 5.3: Incorporation with possessor stranding

Note that the recipient is also an unmarked choice of non-macrorole core argument, since it is cross-referenced by the relation (R_2) *i-*.

Without the incorporation, the construction above would be as in (163) below, with the possessive RP *Pedro r-a?ira* ‘Pedro’s son’ in the extra-core slot (ECS), coreferential with the *i* in the core:

- (163) Aime?eɲ Pedro ra?ira
 a-i-me?eɲ Pedro r-a?ir-a
 1SG-R₂-give Pedro R₁-son-REF
 ‘I give away Pedro’s son.’ (non-attested)

In (161), the recipient (Pedro) is not morphologically marked as such – it is not known if it could appear in a different order, e.g., pre-core.

Since examples of the type of incorporation displayed in (162) are not frequently attested, it is not possible to know if the dative shift involved in such constructions entailed semantic differences, exhibiting some kind of fluidity in terms of semantic role/syntactic function correspondences (see Dixon 2011).

5.4 Privileged Syntactic Argument

The treatment of grammatical relations in diverse languages based on the relations of subject, direct object, and indirect object has revealed itself to be problematic (see, e.g., Foley and Van Valin Jr 1984; Dryer 1997a). RRG posits a single construction-specific grammatical relation, called the PSA, of which ‘subject’ is a generalization. Thus, one can speak of the ‘Subject’ in German, English, Malagasy, etc., but not the *‘PSA in German’, for example. Conversely, one can speak of ‘the PSA of a raising construction’, but not *‘the subject of a raising construction.’

The notion of the PSA is justified by the fact that seemingly all languages have syntactic constructions in which there are restrictions on the RPs and PPs (arguments and non-arguments) that can be involved in them. These restrictions, the privileges given to a constituent, define a privileged syntagmatic function with respect to that construction (Van Valin Jr and LaPolla 1997, 251,277). Thus, acknowledging that grammatical relations do exist but are not necessarily universal, RRG assumes that grammatical relations are not only language-specific but also construction-specific (see LaPolla *ming*). The PSA is associated with the pivot or controller of a construction. Voice modulation may alter which of these is required in a construction, as when the passive requires the Undergoer to be the pivot or controller. TUP does not have a passive voice, and it also lacks constructions in which an argument of a linked construction is missing.

In TUP, the same argument indexes cross-reference the subject of transitive and intransitive verbs, whether this is actor or undergoer, i.e., there is a neutralization of semantic roles, as shown in (164):

- (164) a. Intransitive verb Subject is undergoer.
 Amanõ
 a-manõ
 1SG-die
 ‘I die.’ (VLB, II, 42)
- b. Intransitive verb. Subject is actor.

Ajān

a-jān

1SG-run

‘He runs.’ (cf. VLB, I, 82)

- c. Transitive verb. Subject is undergoer.

Asawsuβ

a-s-awsuβ

1SG-R₂-love

‘I love him.’ (VLB, I, 33)

- d. Transitive verb. Subject is actor.

Asupir

a-s-upir

1SG-R₂-lift

‘I lift it.’ (VLB, I, 121)

There is no split in TUP, as mentioned in Section 5.1. Many stative roots require the same active markers exemplified in (164): *k^waβ* ‘know’, *ikoβe* ‘live, be’, *ʔi / ʔe* ‘be of age, be late’.

Referring to the PSA as construction-specific means that there will often be conventionalized patterns, such as the position of a referring expression in the clause with some semantic role or macrorole, marking on nouns or pronouns with particular semantic roles, or reference to a referent in two clauses (see LaPolla *ming*), where the construction limits the possible interpretations of the role of a particular participant in the action described in the clause (Van Valin Jr and LaPolla 1997, 242-316).

The PSA in TUP is always the subject [A or S], which is a generalization across the PSAs of particular constructions. This is a default choice because, in accusative languages, the highest-ranking macrorole is the default (not the unique) choice for PSA. In TUP cosubordination, for example, there is a restriction on the interpretation of the argument in the nominalized core (see Section 10.1.2), which must be coreferential with the S in the main core. Coreference requires a different construction if the argument in the main core is the A, and it is impossible with an O argument.

The semantic representation (logical structure) of a grammatical clause is the first step in constructing a clause, as described in Section (3.3), where each *Aktionsart* has a unique logical structure that includes the salient argument positions. Based on the position of the arguments in the semantic representation, macroroles are assigned according to the AUH (see Section 3.3), and one of these is chosen to bear the privileged relation to the predicate PSA for the specific construction(s). This relation is privileged syntactically in that it is signaled by coding properties (e.g., agreement) and behavioral properties (e.g., the role of the RP), a distinction suggested by Keenan (1976).

5.5 Nonprototypical (‘non-verbal’) predication

Nonprototypical predication refers to the predication of concepts other than action concepts (see Chapter 4). It includes the predication of object concepts and property concepts. Nonprototypical predication also includes the predication of location (in which case it is existential) and the predication of possession (where it is predpossessive). Within the Tupían family, there are interesting types and variations of these (see Dietrich 2023).

Most types of nonprototypical predication (see Haspelmath 2022; Croft 2022a) in TUP require a unique construction, as exemplified in (165). Here, the lexical root functions as the predicate and is preceded by a relational marker (165a), which may be preceded by a possessor index (171) or, alternatively, preceded or followed by an RP, as in (165c) and (165d).

- (165) a. Attributinal construction
 Ikatu beʔĩ
 i-katu beʔĩ
 R₂-goo a.little.bit
 ‘It is a little better.’ (VLB, I, 31)
- b. Classificational construction
 ʃeporomoʔesar
 ʃe=poro-moʔe-sar
 1SG=ANTIP-teach-NMLZ_{AG}
 ‘I am a teacher.’ (cf. VLB, II, 62)

- c. Equational construction
 Iporeseʔõ ʃerera
 Iporeseʔõ ʃe=r-er-a
 Iporeseõ 1SG=R₁-name-REF
 ‘My name is Iporeseõ.’ (Poemas, 154)
- d. Equational construction
 ʃerera Kururupeβa
 ʃe=r-er-a Kururupeβa
 1SG=R₁-name-REF Kururupeba
 ‘My name is Kururupeba (flat frog).’ (Teatro, 92)
equate’ (ʃerera,Kururupeβa)
- e. Attributional construction
 Kunumĩ turusu
 kunumĩ-∅ t-urusu
 boy-REF R₂-big
 ‘The boy is big.’ (FA, 75)
big’(kunumĩ)
- f. Predpossessive construction
 ʃeko
 ʃe=∅-ko
 1SG=R₁-slash
 ‘I have a slash.’ (FA, 67)
have’(1SG,ko)

The existential type may also be expressed by the construction in (165), but often the (full) verb *ikoβe* ‘exist, be’¹⁰ (see Figueira 1687, 66) is employed in the texts.

- (166) a. Oikoβe ʃetaiaiβa
 o-ikoβe ʃe=∅-taiiaiβ-a
 3-exist 1SG=R₂-courage-REF
 ‘My courage exists.’ (Teatro, 24)
exist’(my courage)

¹⁰Formed by the root *iko* ‘be’ + the adverbial particle *βe* ‘still, also’.

- b. Oikoβepe amō aβa sekoβjaramo?
 o-ikoβe=pe amō aβa-∅ s-ekoβjar-ramo
 3-exist=Q other man-REF R₂-substitute-TRSL
 ‘Does another man exist as his successor?’ (Araújo, 50v)

All attestations of *ikoβe* have the verb preceding the RP in the ECS.¹¹

For equational predication, another construction is available which uses the nominalizer used for relativization (see Section 8.3). This construction is not particularly different since no copular element stands between both arguments. Two examples are given in (167).

- (167) a. Pedro osóβa?e
 Pedro o-so-βa?e
 Pedro 3-go-NMLZ_{REL}
 ‘Pedro is the one who goes / the going one.’ (AA, 30v)
equate’(Pedro, osóβa?e)
- b. Niporaŋiβa?e ruã a?e tata
 n-i-poraŋ-i-βa?e ruã a?e t-ata-∅
 NEG-R₂-beauty-EPEN-NMLZ_{REL} NEG this R₄-fire-REF
 ‘That fire is not the one which is beautiful.’ (Araújo, 163v)

The case of the predlocative construction is curious, because while all other non-verbal predicative constructions in TUP do not require a copula or a verb, the predlocative construction is almost exclusively attested with the verbs *iko* ‘be, act’ and *ikoβe* ‘exist, be’. Also puzzling is the fact that in many other TG languages, including Nheengatu, the direct descendant of TUP (see Magalhães et al. 2019, 179 and Cruz 2011, 471-477), the copula is not required.¹²

¹¹Araújo (1618b, 44v,46,50v), Anchieta (2006, 10,42,62,92,138,156). An exception is found in Anchieta (1997, 148).

¹²The case of TUP could be due to the fact that the authors of the texts were speakers of Portuguese, to whom the lack of a copula was perceived as something odd.

- (168) a. Pedro okope sekow
 Pedro o-ko-pe s-iko-w
 Pedro CORF-slash-LOC R₂-be-NFOC
 ‘Pedro is in his own slash.’ (FA, 81)
- b. iβakipe oiko jaβepe sekow?
 iβak-pe o-iko jaβe=pe s-eko-w
 sky-LOC 3-be as=Q R₂-be-NFOC
 ‘Is he in heaven as he is in it (in the wafer)?’ (DC, I, 215)

Two of the few attestations of a predlocative construction without a verb are given below:

- (169) Maʔepe calix pupe?
 maʔe=pe calix pupe
 thing=Q chalice in
 ‘What is in the chalice?’ (DC, I, 216)
- (170) Umāpe Tatapitera?
 umā=pe Tatapitera
 where=Q Tatapitera
 ‘Where is Tatapitera?’ (Teatro, 130)
- (171) a. Neruβ
 ne=r-uβ
 2SG=R₁-father
 ‘You have a father.’ (FA, 39)
- b. ʃerasi
 ʃe=r-asi
 1SG=R₁-pain
 ‘I have pain.’ (Teatro, 48)

In the case of a non-possessed root, there appears to be no need for the relational of non-contiguity, as in (172)¹³.

¹³*Maʔe* ‘thing’ is attested with the relational of non-contiguity and with the relational of coreference, as well as with possessor indexes.

- (172) Maʔeete Tupã repʔaka?
 maʔe-ete Tupã r-epʔak-a
 thing-EMPH God R₁-see-REF
 ‘Is it a good thing to see God? (lit. the vision of God)’ (DC, I, 173)

With a possessor index:

- (173) a. ʃeruβ
 ʃe=r-uβ
 1SG=R₁-father
 ‘I have a father / There is my father.’ (FA, 38)
- b. Neruβ
 ne=r-uβ
 2SG=R₁-father
 ‘You have a father / There is your father.’ (FA, 38)

Since TUP lacks a third person possessor index, whenever the possessor is expressed by an RP, the RP follows the predicate, which carries the relational of non-contiguity because its dependent does not precede the head.

- (174) a. Iporaŋ ko Tupãoka
 i-poraŋ ko Tupã-ok-a
 R₂-beauty this God-house-REF
 ‘This church is beautiful.’ (FA, 38)
- b. Ipor nerige
 i-por ne=r-ige
 R₂-full 2SG=R₁-womb
 ‘Your womb is full.’ (Poemas, 116)
- c. Ipoʃi
 i-poʃi
 R₂-ugly
 ‘He is ugly.’ (Araújo, 163v)

If the RP is fronted, preceding the predicate, then it is in a topical position (PrDP).

Predlocation (predication of location) is a curious case in TUP, not only because it differs from other nonprototypical predication constructions but because it also differs from other TG languages. Predlocation in TUP, as attested in the texts, always requires the verb *iko* ‘be, act, live, happen’, which takes person indexes from Set I (see Section 4.3.1).

5.6 Word order

In TUP, the order of bound elements in the core corresponds to SOV. Nonetheless, the word order of the RPs coreferential with the arguments is not fixed. Frequency is not the most prominent criterion in identifying word order (see Siewierska 1988), but it is an important one. The recently published treebank of Tupinambá in the Universal Dependencies (Gerardi 2020) project will be a useful tool for quantitative analyses as its coverage improves, and will allow for a clearer picture regarding not only word order, but other aspects of the language as well.

The RPs in ECS coreferential with the core arguments have no overt marking to indicate their grammatical functions, and their order seems to be flexible in relation to the core. The first grammarians, as expected, did not present a clear picture, but Anchieta (1595, 16v), for instance, does mention that in a sentence such as (175), ‘his (own) father’ refers to Joanne, which suggests that the RP coreferent with the actor more commonly preceded the undergoer-related RP (OSV), even though later texts seem to prefer SOV.

- (175) Joanne Pedro osawsuβ oguβa rawsume
 Joanne Pedro o-s-awsuβ o-uβ-a r-awsuβ-me
 Joanne Pedro 3-R₂-love R₄-father-REF R₁-love-CLM
 ‘Pedro loves Joanne because he loves his own father / Joanne loves Pedro because he loves his own father.’ (AA, 16v)

Anchieta (1595, 36v) also mentions a simpler case in which an animate entity acts on an inanimate object. In this case, the animate RP is taken to be the

actor independent of the order in which it appears:

- (176) a. OSV
 Miape Pedro oʔu
 miape-∅ Pedro o-ʔu
 bread-REF Pedro 3-eat
 ‘Pedro eats bread.’ (AA, 36v)
- b. SOV
 Pedro miape oʔu
 Pedro miape-∅ o-ʔu
 Pedro bread-REF 3-eat
 ‘Pedro eats bread.’ (AA, 36v)
- c. VSO
 Oʔu Pedro miape.
 o-ʔu Pedro miape-∅
 3-eat Pedro bread-REF
 ‘Pedro eats bread.’ (AA, 36v)

When both arguments are animate, Anchieta states (Anchieta 1595, 36v) the meaning is ambiguous. This seems to imply that any order would have been possible¹⁴, such that in the case of (177), the only way to avoid ambiguity would be to use ‘participles’, a term used by Anchieta to refer to nominalizers (see Section 8.3), such as those employed in (178).

- (177) Pedro Joanne ojuka
 Pedro Joanne o-i-juka
 Pedro Joanne 3-R₂-kill
 ‘Pedro kills Joanne / Joanne kills Pedro.’ (AA, 36v)
- (178) a. Pedro ijukasara
 Pedro i-juka-**sar**-a
 Pedro R₂-kill-NMLZ_{AG}-REF
 ‘Pedro was his killer.’ (AA, 36v)

¹⁴The same problem was noted by Montoya (1876, 35) for Old Guaraní.

- b. Pedro ijukapira
 Pedro i-juka-**pir**-a
 Pedro R₂-kill-NMLZ_{PAT}-REF
 ‘Pedro was the killed (one).’ (AA, 36v)

Besides the orders in examples above, the following orders are attested:

- (179) The RP related to the actor follows the core
 I have not found such a case in the texts.
- (180) The RP related to the undergoer precedes the core
 Neakaŋa juka ajpota korine
 ne=∅-akaŋ-a ∅-juka-∅ a-i-pota kori=ne
 [2SG=R₁-head-REF R₁-kill-REF]_j 1SG_i-R_{2j}-want today-FUT
 ‘I shall want to break your head later on today.’ (Staden, 156)
- (181) The RP related to the undergoer follows the core
 Oimome?u umã karaiβeβe ipiki?irape
 o_i-i_j-mome?u umã [karaiβeβe-∅]_i i-piki?ir-pe
 3-R₂-announce already angel-∅ R₂-younger.cousin-DAT
 ipuru?aramo seko
 i-puru?a-ramo s-eko
 R₂-pregnant-TRSL R₂-be
 ‘The angel had already told her cousin of her pregnancy (of her being pregnant).’ (Araújo, 6v)
- (182) Both RPs precede the core with the RP related to the DCA preceding the RP related to the ‘subject’
 Tupã ase ojmoete...
 [Tupã]_j [ase]_i o_i-i_j-mo-ete
 God we 3-R₂-CAUS-good
 ‘We honor God / one honors God [...].’ (Araújo, 101)

Noun modifiers (see Section 8.4) follow the head noun, as in (183), a property that correlates with OV order more commonly in South America (and Australia-New Guinea) than in other parts of the world (Dryer 1992b, 95).

- (183) Aβaporaja
 aβa-poraj-a
 man-beauty-REF
 ‘Beautiful man.’ (see DC, II, 97)

Determiners (see Section 8.2.1.1) precede head nouns (184), as well as genitives (185).

- (184) Iko aβa
 iko aβa-∅
 DEM man-REF
 ‘This man.’ (Araújo, 60v)

- (185) Tupana jeʔeɲa
 Tupana ∅-jeʔeɲ-a
 God R₁-word-REF
 ‘Word of God.’ (AT, 146)

SOV languages with postpositions, noun-modifier, and possessor-possessed order are neither the most common type of SOV, nor the least common (see Greenberg 1963; Hawkins 1983).

5.7 Valency changing

The notion of valency has a wide range of effects on the morphosyntax of Tupinambá. All valency changing morphemes in TUP, with one exception, are prefixes occurring closer to the predicate nucleus than other prefixes, such as person markers. This is a common characteristic of Tupián languages which is also common in Amazonia (cf, Dixon and Aikhenvald 1999, 9).

Verbal valence morphology reflects the influence of semantic valency on TUP morphosyntax. The fundamental valency distinction in this language is that between semantically monovalent and semantically polyvalent verbs. This distinction plays a prominent role in phenomena such as the formation of imperatives, interrogatives, and negative clauses. Not surprisingly, it also plays a prominent role in the establishment of grammatical relations.

5.7.1 Causatives

Causatives increase the semantic valency of predicates by introducing a new agent to their argument structure (Zúñiga 2020, 15). TUP has two types of causatives: lexical, of which there are no more than a dozen, and morphological¹⁵. Examples of lexical causatives, i.e., causatives which do not contain any formal differentiation between the causal predicate and the affected predicate, include: *poj* ‘feed’, *samok* ‘untie’, *upir* ‘lift, raise’, *juka* ‘kill’, and *apĩĩ* ‘tie’. A possible source of these lexical causatives lies in their colexifications: *juka* < *ajur* ‘neck’ + *ka* ‘break’, *samok* < *sam* ‘rope’ + (?)*ok* ‘cut’. There is also a pair of verbs which seems to stem from a single root displaying vowel alternation between /e/ and /u/ for the intransitive-causative opposition: *jeka* (intr.) ‘break’ and *juka* (tr.) ‘break’¹⁶. Possibly, there were other verbs exhibiting such an alternation that are not attested in the TUP corpus but which have survived in other languages, as in Paraguayan Guaraní (see Velázquez-Castillo 2002, 512).

5.7.1.1 Causative of M-intransitive predicates

There are two types of morphological causatives, depending on the M-transitivity of the predicates. M-intransitive predicates are causativized by the prefix *mo-*. The following examples (186) show causativized predicates:

- (186) a. *Aimojirõ* *Tupã fejopupe*
 a-i-mo-jirõ *Tupã je=jo-pupe*
 1SG-R₂-CAUS-calm God 1SG=RFLX-POSP

¹⁵These two types of causatives could be expressions of two different levels of lexical complexity, as suggested by Nichols et al. (2006); Nichols (2009).

¹⁶These verbs are found in Old Guaraní Montoya (1876, 199-199v).

- ‘I appease God for me.’ (FA, 81)
 [do’ (1SG, ∅)]CAUSE [feel’ (i(Tupã), [calm’])]
- b. Opa iβaka erejmopo, paranã iβi aβe
 opaβ iβak-a ere-i-mo-por paranã-∅ iβi-∅ aβe
 all sky-REF 2SG-R₂-CAUS-contain sea-REF earth-REF also
 ‘You fill all the skies, the sea and the earth.’ (Poemas, 128)

While obligatory in the northern variety described by Figueira (1687), the relational morpheme (R₂) indexing the undergoer of causativized predicates as in (186) was not required in the southern variety described by Anchieta, as exemplified in, (187)¹⁷.

- (187) a. Amoramwe
 a-mo-ramwe
 1SG-CAUS-frustrate
 ‘I frustrated them.’ (AT, 14)
 [do’ (1SG, ∅)]CAUSE [feel’ (∅, [frustrated’])]
- b. A-monij korinone
 a-mo-nij kori-no=ne
 1SG-CAUS-tremble today-PRCL=FUT
 ‘I will scare them today too.’ (AT, 20)

The absence of (R₂) indicates that the undergoer is zero realized and coreferential with an RP in the ECS, as in (188).

¹⁷The absence of an indexed undergoer in the cores of a predicate causativized by *mo-* is the most common situation in all TG languages, except the northern variety of TG. It seems more reasonable to assume the loss of R₂ before *mo-* in the TG languages, including the southern variety of TUP, than an innovation of the northern variety. The evidence suggesting the retention of R₂ is the fact that Mundurukú (Crofts 2004; Gomes 2006) and Mawé (Silva et al. 2010), two branches that split before TG (see Galucio et al. 2015; Rodrigues and Dietrich 1997), have the relational obligatorily followed by the causative morpheme in a causative construction. Awetí, the last Tupían split before TG, does not show it because only vestiges of relationals are found in this language (see Rodrigues and Cabral 2012, 514-515).

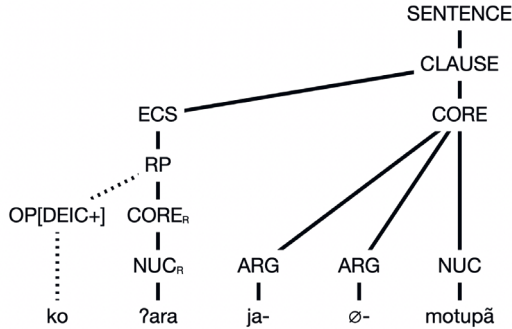


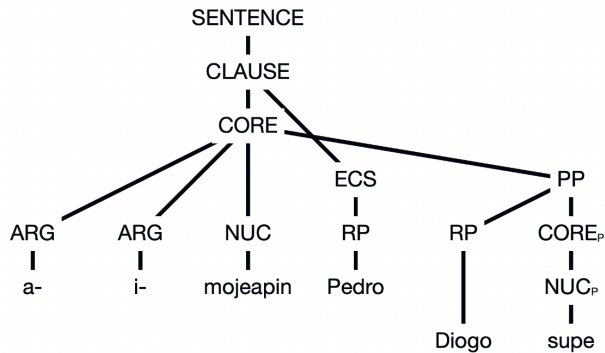
Figure 5.4: Representation of (188)

- (188) Ko ?ara jamotupã
 [ko ?ar-a]_i ja-∅_i-mo-tupã
 this day-REF 1PL.INCL-R₂-CAUS-God
 ‘We sanctify this day.’ (Araújo, 4v)

Other examples are shown in (189), with their syntactic representation given in Figure 5.5:

- (189) a. Aimojeapin Pedro Diogo supe
 a-i-mo-jeapin Pedro Diogo supe
 1SG-R₂-CAUS-shave Pedro Diogo POSP
 ‘I make Diogo shave Pedro.’ (FA, 90)
- b. Ko santo omoŋgetasara ojmojekosuβ
 ko santo o-moŋeta-sar-a o-i-mo-je-kosuβ
 this saint CORF-pray-NMLZ_{AG}-REF 3-R₂-CAUS-RFLX-rejoice
 ima?e ikajemira kojpo
 i-ma?e-∅ i-kajem-pir-a kojpo
 R₂-thing-∅ R₂-disappear-NMLZ-REF or
 semiawsujaβaβa supe
 s-emi-awsuβ-jaβaβ-a supe
 R₂-RES-slave-flee-REF to

‘This saint helps the one who prays to him to recover his lost things or his runaway slave.’ (Araújo, 6)



[do' (I, ∅)]CAUSE [do' (Diogo, [shave' (Diogo, i[Pedro])])]

Figure 5.5: Representation of (189a), a causative construction with three core arguments

The example (189a) is an interesting one, since *a-* is selected as the actor macrorole (see Figure 3.18) and Pedro as the undergoer macrorole. Diogo, despite being an actor, cannot be a macrorole, and is therefore marked by a postposition according to (101).

The prefix *mo-* displays considerable flexibility regarding the types of stem with which it combines. Some examples are given in (190)-(193):

(190) Nominal root

Aimoaβare Pedro

a-i-**mo**-aβare Pedro

1SG-R₂-CAUS-priest Pedro

‘I ordain Pedro / cause Pedro to be(come) a priest.’ (Arte, 48v)

(191) Intransitive state predicate

Oporomotekok^waβe?ima

o-poro-**mo**-tekok^waβ-e?im-a

3-ANTIP-CAUS-ignorant-PRIV-GER

‘Causing people to be ignorant.’ (Araújo, 83)

(192) Intransitive active predicate

Aimosem Ajanga fejoswi
a-i-**mo**-sem Ajaŋa fe=jo-swi
1SG-R₂-CAUS-leave devil 1SG=RFLX-POSP

‘I expel the Devil from myself.’ (DC, I 163)

(193) Numeral

Momosapir
mo-mosapir
CAUS-three

‘Cause to be the third (time).’ (VLB,II,115)

Table 5.3 has examples of causativized predicates. Note that all predicates are originally intransitive¹⁸.

Form	Translation	Causativized form	Translation
aβare	priest	mo -aβare	cause to be a priest / ordain (VLB, II, 58)
tiniŋ	dry (intr.)	mo -tiniŋ	cause to dry /dry (tr.) (VLB, II, 1 14)
so	go	mo -no	cause to go / send (FA 84)
akuβ	hot	mo -akuβ	heat up (DC, I, 221)
sii	tremble	mo -nii	scare /cause to tremble (AT 20)

Table 5.3: Causativized intransitive roots

Predicates derived with *mo-* have all the morphological possibilities of a regular transitive predicate, i.e., they may combine with the reflexive *je-* (194), the nominalizing resultative prefix *t-embí-*¹⁹ (195), and the nominalizing agentivizer *-sar* (196):

¹⁸The transitive verb *itarō* ‘satisfy’ is indeed attested with *-mo*, but without a change in meaning.

¹⁹The *t-* is the R₄, which indicates a non-specific possessor.

- (194) Ajemooriβusu, neroβake
 a-**je-mo**-oriβ-usu ne=r-oβake-∅
 1SG=**RFLX-CAUS**-happy-AUG 2SG=R₁-face-REF
 witu
 wit-u-∅
 1SG_{CORF}-come-GER
 ‘I rejoice greatly about you (because your face is coming to me).’ (Abbeville, 342)
- (195) feremimono
 fe=r-**emi-mo**-so-∅
 1SG=R₁-**RES-CAUS**-go-REF
 ‘My sent thing (thing I caused to go).’ (FA, 70)
- (196) Ajanga mosemara
 Ajanga-∅ **mo-sem-sar**-a
 devil-REF **CAUS**-leave-**NMLZ**_{AG}-REF
 ‘One who casts out the devil (who causes the devil to leave).’ (Cantigas, III)

In (197), the incorporation with the non-contiguous marker (R₂) makes the verb intransitive. Thus, the complex nucleus can combine with the causative *mo-* (see Section 5.7.1.2).

- (197) Aimotekok^waβ
 a-i-**mo**-t-eko-k^waβ
 1SG-R₂-**CAUS**-R₂-fact-know
 ‘I teach him (cause him to know facts).’ (VLB, II, 12)
 [do’ (I)] CAUSE [BECOME **know**’ (i,tekokuwaβ)]

5.7.1.2 Causative of M-transitive predicates

It is not uncommon cross-linguistically for predicates of different valencies to be causativized by different strategies (Zúñiga 2020, 31). M-transitive predic-

ates are causativized by *-ukar* suffixed to the predicate²⁰. Contrary to Navarro (2011, 2013), who considers *ukar* to be a verb because it may combine with nominalizers such as *-sar* and *-saβ*, it is here considered to be a bound lexical morpheme. Since most of the affixes that combine with the predicate nucleus are prefixes, it is indeed exceptional that *-ukar* is a suffix. The following example is offered by Navarro (2011):

- (198) Omena kojpo wemireko jukasara kojpo
 o-men-a kojpo o-emireko-∅ juka-sar-a kojpo
 3_{CORF}-husband-REF or 3_{CORF}-wife-REF kill-NMLZ_{AG}-REF or
 ijukaukasara (...) i-juka-**ukar-sar**-a
 R₂-kill-FAC-NMLZ_{AG}-REF
 ‘The one who kills his own husband or his own wife, or the one who causes their killing (makes them get killed).’ (Araújo [1686], 279)

In the example above, the nominalizer *sar* attaches to *ukar* when this is already combined with a lexical element.²¹

In favor of *-ukar* as a factive marker, there is the fact that it is never attested alone combined with argument indexes; (Figueira 1687, 146) states that it alone has no meaning. Furthermore, it is probably diachronically related to a transitivizer found in other Tupían languages, such as *ka* in Akuntsu (see e.g. Aragon 2014, 213-217). The sentences in (199) illustrate the use of *-ukar*.

²⁰In order to differentiate the causatives of intransitive from causatives of transitive, the latter are glossed as FAC ‘factive’.

²¹Out of fifteen TG languages, including Old Guaraní, where *ukar* is also used as a causative with transitive verbs, none has it as a lexical root: Apiaká (Sousa 2017), Asuriní Xingu (Pereira 2009), Asuriní Tocantins (Cabral et al. 2011), Chiriguano (Dietrich 1986), Tekó (Rose 2000), Guajá (Silva Magalhães 2002), Guajajara (Harrison and Harrison 2013), Guaraní (Estigarribia 2020), Guarayo (Höller 1932), Kamajurá (Seki 2000), Mbyá (Dooley 1998), Old Guaraní (Montoya 1876), Parakanã (Ferreira da Silva 2003), Tapirapé (Praça 2007). The only language where a cognate of *ukar* has a function other than a causative is Zo’e, where it appears to have a modal meaning (de Castro et al. 2020).

- (199) a. Herodes pitaja [...] mokōj roʔi
 Herodes pitaj-a [...] mokōj roʔi-∅
 Herodes child-REF [...] two year-REF
 omoawjeβaʔe, mopaβukarawera
 o-mo-awje-βaʔe **mo-paβ-ukar-saβ-wer-a**
 3-CAUS-terminate-REL CAUS-all-FAC-NMLZ-PST-REF
 ‘Herodes caused the annihilation of the children that completed
 two-years.’ (Araújo, 10)
- b. Judeus supe sepʔakuka
 judeus-∅ supe s-epʔak-**ukar-a**
 jews DAT R₁-see-FAC-GER
 ‘Showing it to the Jews (causing the Jews to see it.’ (Araújo, 60v)
- c. Aʔe omena supe ʔiβa ʔuukaraʔuβi
 aʔe o-men-a supe ʔiβa-∅ ʔu-**ukar-a**ʔuβ-i
 PRCL COREF-husband-REF POSP fruit-REF eat-FAC-false-NFOC
 ‘She wrongly made her husband eat the fruit.’ (Poemas, 178)
- d. Esepʔakukar oreβe
 e-s-epʔak-**ukar** ore-βe
 2SG.IMP-R₂-see-FAC 1PL.EXCL-DAT
 ‘Make us see him.’ (Araújo, 14v)
 [do’ (e)] CAUSE [see’ (ore,3)]
- e. Santa Helena, Constantino rei si,
 Santa Helena Constantino rei ∅-si-∅
 Saint Helena Constantino king R₁-mother-REF
 osekarukar
 o-s-ekar-ukar
 3-R₂-search-FAC
 ‘Saint Helena, mother of the king Constantine, caused (commanded)
 to seek it.’ (Araújo, 4v)

The interpretation of *ukar* as a verb has implications for the analysis of causative/factive constructions. In the case of a lexical verb with a causative function, such as the verb *faire* ‘do, make’ in French, the core contains a complex

nucleus formed by the junction of two nuclei . This is exemplified in (200), with its representation given in Figure 5.6 (see Section 10.1.1):

(200) Je fais manger les gâteaux à Fabrice.

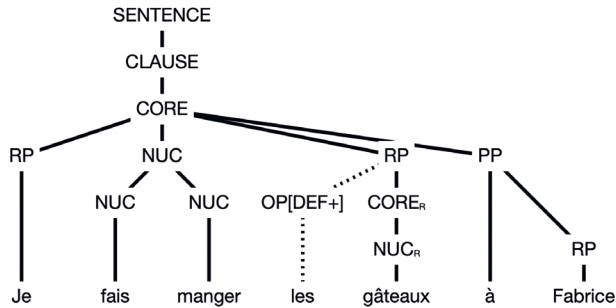


Figure 5.6: Factive construction with nuclear juncture (complex nucleus) in French

In the analysis proposed here, there is no nuclear juncture since *ukar* is a causative marker, not a lexical item. Thus, the structure of (201) can be seen in its semantic representation:

(201) Munepora mojepe pejmosemukar iséβe
 munepor-∅ mojepe pe-i-mo-sem-ukar ise-βe
 prisoner-REF one 2PL-R₂-CAUS-leave-CAUS I-DAT
 ‘You make me liberate one prisoner.’ (Araújo, 59v)
 [do' (pe, ∅)] CAUSE [do' (ise, ∅) CAUSE [BECOME be.free'(i[munepora])]]

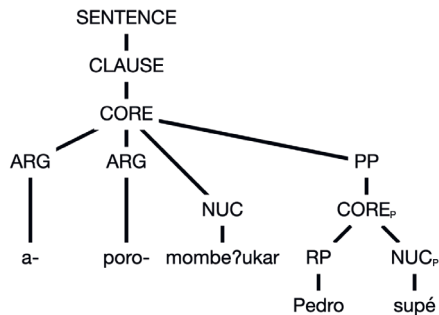
In (201), the causative predicate *mo-sem* ‘cause to leave’ is causativized by the causative *ukar*, which requires an additional argument. As the semantic representation shows, the initial effector of the causal chain is the actor macrorole (Van Valin Jr and LaPolla 1997, 84-85, 145-147, 377-382). The undergoer macrorole is *i*, with which the RP *munepora* ‘thief’ is coreferential, and the choice is clear: according to the AUH (3.18) it outranks *ise* ‘I’, which is also an argument, but an oblique non-macrorole argument. Since it is also

an effector (actor), it has to appear in the dative, the default case for non-macrorole arguments (Van Valin Jr 2005, 110-115).

As a causativizer of M-transitive predicates, *ukar* may combine with causative predicates (causativized by *mo*). An example was given in 201 and another follows in 202.

- (202) Aporomo?eukar Pedro supe
 a-poro-**mo**-?e-**ukar** Pedro supe
 1SG-ANTIP-CAUS-say-CAUS Pedro POSP
 ‘I make Pedro teach people.’ (FA, 146)

Here, the PSA is the causer agent. The causee, also an agent, bears the dative case and appears in the periphery as an adjunct (not as a core argument).



[**do'** (I, ∅)] CAUSE [**teach'** (Pedro, people)]

Figure 5.7: Causative clause with peripheral argument

The result of causativizing a causative predicate is a two-argument clause in which the causee bears dative case, as in (202), represented in (5.7).

5.7.1.3 Sociative causative

Causation in Tupinambá, as in many South American languages (Guillaume and Rose 2010), is semantically sensitive to a specific aspect of event structure which, according to Shibatani and Pardeshi (2002), lies between direct and indirect causation on a continuum, referred to as sociative causation²². Sociative causation is marked by the prefix (*e*)*ro*-²³, indicating that the causer makes the causee perform the action and takes part in it. A distinct marker for sociative causation is apparently an uncommon typological feature, although it is more commonly found in South America (see Guillaume and Rose 2010). Example (203) shows an example of sociative causation with its logical structure:

- (203) Aroβeβéne
 a-**ero**-βeβe=ne
 1SG-SCAU-fly-FUT
 ‘I will make them fly with me.’
 [**do**’ (1SG, ∅)] CAUSE [**do**’ [(they) **fly**’ ([they])] ∧ [**do**’ (1SG, ∅), **fly**’ (1SG)]

The sociative causative prefix only combines with intransitive roots, as in (204), where the semantics of *ero*- shows that it can have a sociative causative meaning (204a) as well as a comitative applicative meaning (204b). There is a difference, as illustrated in (204), between *I make them dance and dance with them* (sociative causative), and *I dance with them* (comitative applicative) (see Guillaume and Rose 2010).

- (204) a. Aimire, jaraso muru, tawje,
 Aimire ja-**ro**-so muru tawje
 Aimbere 1INCL.PL-SCAU-go unblessed soon

²²In the vast majority of TG literature, this has been referred to as causative-comitative voice (Rodrigues 2010b; Rose 2011; Seki 2000; Magalhães 2010).

²³Note that ‘an asymmetry in degree of control between causer and causee is necessary for causation to eventuate, an active causee is normally seen as a cooperating participant, albeit not always willing’ (Velázquez-Castillo 2002, 521). In this sense, there is no distinction regarding the volition of the causee expressed by the morpheme *ero*-.

janeroipira moesāja
 jane=r-oipir-a mo-esāj-a
 1INCL.PL=R₁-substitute-REF CAUS-happy-GER

‘Aimbire, let’s take (cause to go with us) the unblest, soon, in order to make our substitutes happy.’ (Teatro, 42)

- b. jeanameta aroporasej seru
 je=Ø-anam-eta a-ero-porasej s-er-u
 1SG=R₁-relative-PL 1SG-SCAU-dance R₂-SCAU-come.GER
 ‘Bringing my parents, make them dance with me.’ (AP, 138)²⁴

5.7.2 Incorporation

One type of incorporation, namely that of possessed roots and their possessors, has already been discussed in Section 5.3. This section discusses a different type of incorporation.

Nominal incorporation is a type of composition that inserts a lexical root into a verbal stem (Sapir 1921; Mithun 1984). At the same time, it can be a valency changing device, because the incorporation of a non-possessed noun with the undergoer function (205) reduces by one the number of independent syntactic arguments in the core. This turns the predicate into an M-intransitive predicate because incorporated lexical roots have low categoriality (Hopper and Thompson 1984, 711-714)²⁵. The incorporation of unpossessed nouns results in a non-referential reading, as the translations of the examples in (205) indicate²⁶ (1984, 856).’, and being non-referential, it cannot refer to a specific

²⁴The root *ur* ‘come’ with the sociative comitative prefix *ero-* forms the verb *erur* ‘bring’, e.g., *I go and cause X to go with me.*

²⁵This statement applies to word-class theories, e.g., as in Givón (2001), but following Velázquez-Castillo (1995, 677) this is taken to mean that incorporated nouns ‘lack many or all of the morphological trappings characteristic of nouns, such as gender, number, or definiteness marking.’ TUP does allow an incorporated noun to be modified by a possessor, providing an exception to a characteristic intrinsically related to transitivity, according to Hopper and Thompson (1984).

²⁶In this regard, Mithun observes that structures with noun incorporation ‘tend to be used in contexts without specific, individuated patients. They may be generic statements; or descriptions of on-going activities, in which a patient has been incompletely affected; or habitual activities, in which the specific patient may change; or projected activities, in which the spe-

‘affected’ participant, which is what the undergoer has to be. Therefore, this second argument of the activity predicate is a syntactic argument of the core and a semantic argument in the semantic representation, but it is not a macrorole. As such, there is only one macrorole in this type of activity predicate. The difference between *I drink beer* (non-referential) and *I drink a/the beer* (referential) is that the former is an activity and the latter an active accomplishment (see Section 3.3). The second argument of an activity predicate characterizes the action rather than describing a participant, and as such, it does not function as an undergoer macrorole.

The following examples are illustrative.

- (205) a. Aiβiraʔaβ
 a-iβira-ʔaβ
 1SG-tree-cut
 ‘I cut trees.’ (FA, 145)
- b. Aʔiasaβ
 a-ʔi-asaβ
 1SG-river-cross
 ‘I cross rivers.’ (VLB, II 67)

The structure of (205a) is given in Figure (5.8). The nucleus of the predicate becomes more complex after the incorporation; it can be considered a new compound word (see Van Valin Jr 2013 and Ullrich and Van Valin Jr 2007), and the predicate has one argument.

Because incorporation reduces the valency of transitive predicates, incorporated structures are expected to combine with the causative of intransitive verbs (206), but not with the causative of transitive verbs (see Section 5.7.1). This is in fact the case, as shown in (206).

- (206) a. Moʔiʔu
mo-ʔi-ʔu
 CAUS-water-ingest

cific patient is not yet identifiable; or joint activities, where an individual agent incompletely affects a particular patient; or activities directed at an unspecified portion of a mass.

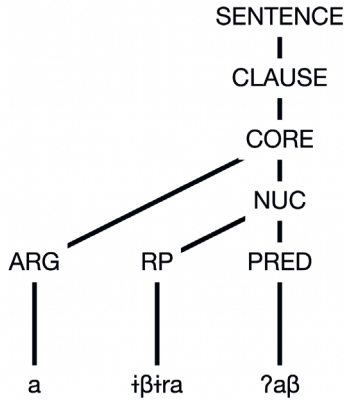


Figure 5.8: Nominal incorporation

‘Make drink water / give to drink.’ (VLB, I, 53)

- b. Oimoʔiʔupe wa?
 o-i-**mo**-ʔi-ʔu=pe wa
 3-R₂-CAUS-water-ingest=Q PRCL

‘Has he been given water? (Did he make him drink?)’ (Araújo, 63)

- c. Aβa mongawaβo
 aβa-∅ **mo**-kawĩ-ʔu-aβo
 person-∅ CAUS-beer-ingest-GER

‘Making people drink beer.’ (Araújo, 78)

As mentioned in Section 5.3 (see Examples (161) and (162)), RPs related to the arguments may be incorporated into the core, replacing the argument indexed by a bound index. Contrary to what is suggested by Zúñiga (2020, 59-60) (quoting Mithun 1984), this is not a type of applicative. As exemplified in (207), the R₂ is not indexing an additional argument; it belongs with the incorporated root, indicating the non-contiguity with the possessor (see Section 4.3.2).

- (207) a. Ajatipetek
 a-[**i-atiβ**]-petek
 1SG-[R₂-temple]-hit
 ‘I slapped his temples.’ (VLB, I, 56)
 [**do**’(1SG[slap’(1SG,i-atiβa)])]
- b. Moruβisaβa βoja amō osoβapetek
 Moruβisaβ-a ∅-βoja-∅ amō o-[**s-oβa**]-petek
 master-REF R₁-servant-REF some 3-[R₂-face-hit]
 ‘Some of the chief’s servants hit his face.’ (Araújo, 55v)
- c. Atujuka Francisco
 a-[**t-uβ**]-juka Francisco
 1SG-[R₂-father-kill] Francisco
 ‘I killed Francisco’s father.’ (FA, 88)

In order to explain the examples in (207), it is necessary to resort to the layered structure of the word (LSW), proposed in Van Valin Jr (2013) for head-marking languages.²⁷ The layered structure of the word has a nucleus (NUC_W), which can be internally complex, and a core_W. Inflectional affixes are considered formatives (FRM) and are assigned to the core_W (word core). Derivation, therefore, occurs at the NUC_W level and inflection at the core_W level. Clitics are formatives that attach to words in detached positions analogous to those in the sentence. Head-marked argument indexes are assigned as formatives to the core_W. A template for the layered structure of the word, adapted from Van Valin Jr (2013), is provided in Figure 5.9.

Thus, the representation of (207a) would be as in Figure 5.10.

Through the LSW, inflectional properties of words are accessible to syntax, as suggested by Van Valin Jr (2013). The inflectional affixes in the core of the word (core_W) also instantiate the core arguments in the core of the clause – both cores are coextensive, the nucleus of the word is also the nucleus of the clause, and the argument-signaling indexes in the core_W are the core arguments.

²⁷Van Valin Jr (2013) is motivated by the status of the independent RPs in head-marking languages (see Section 6.3), and by the targeting of elements inside the word by syntax.

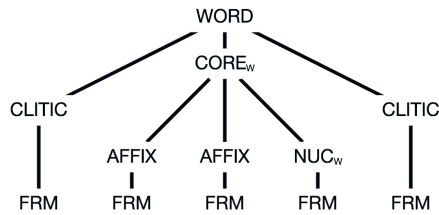


Figure 5.9: The layered structure of the word

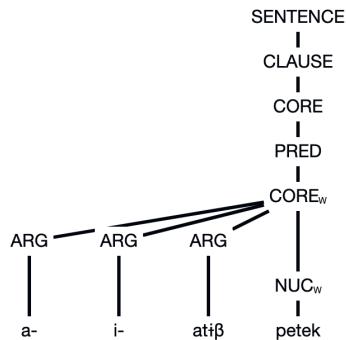


Figure 5.10: Layered structure of a word with noun incorporation

The example above can be compared with (208), where *s-eβira* is related to the argument but is not the core argument, since *i* is the core argument.

- (208) Seβira a_jpetek
 [s-eβir-a]_j a_i-i_j-petek
 R₂-buttocks-REF 1SG-R₂-hit
 ‘I slapped his buttocks.’ (VLB, II, 135)
 [do'(1SG[slap'(1SG,i[seβira]))]]

Another possibility is the incorporation of a whole possessed phrase, as in (209). The incorporated object and the predicate build a complex nucleus, and the RP is fully referential. The argument marked by the postposition is an oblique core argument. Examples such as (209) are only attested a couple of

times.

- (209) Naʃeraʔipotari nerese
 N-a-**ʃe=r-aʔir**-potar-i ne=r-ese
 NEG-1SG-1SG=R₁-son-want-NEG 2SG=R₁-POSP
 ‘I do not want you as a son (lit. I do not want my son as you).’ (FA,
 124)

Another possibility is the incorporation of the possessed RP without the possessor. This case is similar to (207a), but with a possessor RP in the ECS²⁸ of the possessor and not the possessed incorporated noun.

- (210) Aikomojaŋ ʃeruβa
 a-[i-ko]-mojaŋ ʃe=r-uβ-a
 1SG-[R₂-slash]-make 1SG=R₁-father-REF
 ‘I do my father’s slash.’ (FA, 87)

Other examples similar to those in (207a) are given in (211):

- (211) a. Asakamiʔok
 a-s-akami-ʔok
 1SG-R₂-fork-rip.out
 ‘I rip out the fork (of a branch).’ (VLB, I, 142)
- b. Aiakaŋeki-ekij
 a-i-akaŋ-eki-ekij
 1SG-R₂-head-RED-pull
 ‘I repeatedly pulled his head.’ (VLB, I, 142)

²⁸Due to the nature and function of the relational morpheme in Tupinambá, the analysis proposed here must be different from the analysis of the same phenomenon in Guaraní proposed by Velázquez-Castillo (1995, 682-685, 695). The main reason for this is the fact that the TUP relational of non-contiguity (R₂) is not viewed as such in Modern Guaraní, as the glossing of Examples (38) in Velázquez-Castillo (1995, 695) indicates. See also Estigarribia (2020, 133-135) and Rose (2011, 382-384).

- c. Atuparuŋ aβati
 a-t-upaβ-ruŋ aβati-∅
 1SG-R₂-lay-establish corn-REF

‘I established (laid out) a corn plantation.’ (VLB, II, 81)²⁹

The cases above are somewhat similar to reflexives, although reflexives display an intermediate status between one and two-argument predicates (Hopper and Thompson 1980), as shown in (212).

- (212) Ojepoej teʔija remiep^lakamo
 o-je-po-ej t-eʔij-a r-emi-ep^lak-amo
 3-RFLX-hand-wash R₃-crowd-REF R₁-RES-see-TRSL

‘He washed his (own) hands being seen by the crowd.’ (Araújo, 61)

The most interesting case of incorporation concerns ditransitive predicates with partial incorporation of the undergoer argument. This is the case of an undergoer possessed RP where only the possessed root is incorporated and the recipient is not an RP but an argument index. In this case, the undergoer is the *first* argument of the state predicate **have**' (x,y), rather than the second, as one would expect based on the AUH (see Figure 3.18). This is a case of an argument which has a thematic relation other than patient serving as undergoer (see Van Valin Jr 2005, 61). This is illustrated in (213). Figure (5.11) represents the structure of (213b).

- (213) a. Aitaʔi-meʔeŋ Pedro
 a-i-[t-aʔir]-meʔeŋ Pedro
 1SG-R₂-R₂-son-give Pedro
 ‘I give him/her/them Pedro’s son.’ (AA, 51)
- b. Aijaomeʔeŋ Pedro
 a-i-[i-aoβ]-meʔeŋ Pedro
 1SG-R₂-[R₂-cloth]-give Pedro
 ‘I give him/her/them Pedro’s clothes.’ (AA, 51)

²⁹The verb *ruŋ* ‘put, establish, arrange’, not frequently attested, is only attested with incorporated nouns, except for its nominal forms.

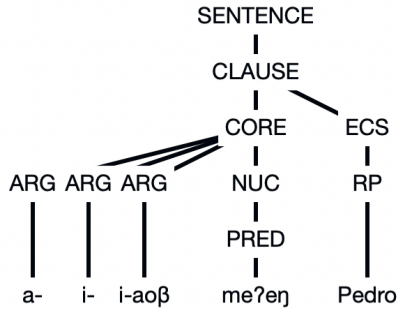


Figure 5.11: Incorporation of possessed noun

Examples like those in (213) are only attested in Anchieta (1595), while examples such as those in (162) are more frequently attested, even in a later source such as Bettendorff (1681). Further examples are given in (214).

- (214) a. Aikoruŋ ʃeruβa
 a-i-ko-ruŋ ʃe=r-uβ-a
 1SG-R₂-slash-prepare 1SG=R₁-father-REF
 ‘I prepare my father’s slash.’ (FA, 145)
- b. Asʔĩjok
 a-s-ʔĩj-ok
 1SG-R₂-seed-RIP.OUT
 ‘I ripped out his seeds.’ (VLB, I, 123)
- c. Eresausupotaretépe Tupã?
 ere-s-awsuβ-potar-ete=pe Tupã
 2SG-R₂-love-want-truth=Q God
 ‘Do you really want to love God?’ (Bettendorff, 125)
- d. Atuβajuka Francisco
 a-t-uβ-a-juka Francisco
 1SG-R₂-father-REF-kill Francisco
 ‘I killed Francisco’s father.’ (FA, 88)

5.7.3 Antipassives: *poro* and *maʔe*

Following Janic and Witzlack-Makarevich (2021), the antipassive is here treated as an intransitive construction in which: (i) the same verb with the same lexical meaning is also found in a transitive construction; (ii) the actor in the transitive construction is encoded as the sole argument of the intransitive construction in the corresponding antipassive construction; and (iii) the undergoer in the transitive construction is either encoded as an oblique or left unexpressed in the corresponding antipassive construction. Through this demotion of the patient argument, the antipassive construction increases the relative topicality of the actor and consequently decreases the relative topicality of the patient argument (Givón 1994). The antipassive markers in TUP are *poro* ‘human’ and *maʔe* ‘non-human, thing’. These meanings are associated with a widespread feature in the Americas: their grammaticalization into antipassives (Say 2021)³⁰. They are also associated with another cross-linguistically common feature, namely that antipassives are commonly found when the object is non-specific or indefinite (Hopper and Thompson 1980; Foley and Van Valin Jr 1985). The presence of two antipassives distinguishing between human and non-human markers is also cross-linguistically common (Janic and Witzlack-Makarevich 2021, 10).

An important characteristic of the antipassive construction, as the examples in this section will show, is the fact that they tend to express habitual, incomplete or non-punctual events (see Cooreman 1994, 57), with the demoted undergoer interpreted as non-referential, indefinite or generic in nature (Janic and Witzlack-Makarevich 2021, 3).

The example in (215) illustrates the difference between an active and an antipassive construction: (215a) is an example of a transitive construction, unmarked in the active voice, with the bivalent verb *suʔu* ‘bite’. The verb

³⁰The case of *maʔe* ‘thing’ is a clear case of grammaticalization of a nominal stem. The etymology of *poro* is not known, so it is not possible to establish its path of grammaticalization or even assert that it is a case of it. These lexemes have often been treated as cases of incorporation in descriptions of TG languages (Dietrich 1994). While *poro* is a case of a dedicated antipassive marker, as it has no other function, *maʔe* is a case of a syncretic antipassive, because this marker does have another function. Both types belong to a well-known cross-linguistic distinction (see Janic and Witzlack-Makarevich 2021, 11-14).

suʔu ‘bite’ is used with the same lexical meaning in the intransitive construction in (215b), with a single argument and *poro* ‘human’ as the antipassive marker.

- (215) a. Aisuʔu
 a-i-suʔu
 1SG-R₂-bite
 ‘I bite it.’ (D’Evreux, Viagem, 158)
- b. Moja oporosuʔu
 moj-a o-**poro**-suʔu
 snake-REF 3-**ANTIP**_{antip}-bite
 ‘The snake bites people.’ (FA, 6)

Following Rose (2011, 265-266), I do not consider the prefix *poro* to be a lexical root since it is not attested as such; it never combines with possessor indexes, nor does it modify other nouns. Rather, *poro* is a grammaticalized morpheme with the unique function of indicating a human indeterminate participant. That a predicate with *poro* is intransitive is corroborated by its co-occurrence with the causative *mo-*, as in (216) (see Section 5.7.1.1):

- (216) Imoporoamotareʔima
 i-**mo-poro**-amotar-eʔim-a
 R₂-CAUS-**ANTIP**-love-PRIV-GER
 ‘Causing them to hate people.’ (DC II, 103)
- (217) a. Aporojaj
 a-**poro**-jaj
 1SG-**ANTIP**-make.fun
 ‘I make fun of people.’ (VLB I, 123)
- b. Oporomoʔeaʔu Tupā jeʔeja raʔaja
 o-**poro**-moʔe-aʔu Tupā ∅-jeʔeja-a r-aʔaj-a
 3-**ANTIP**-teach-false God R₁-speech-REF R₁-experience-REF
 ‘Falsely teaches people to experience the word of God.’ (AT, 136)

- c. Oporomoiŋoβémoʔanga
 o-**poro**-mo-ikoβe-moʔaŋ-a
 3-ANTIP-CAUS-live-pretend-GER
 ‘Pretending to make people live.’ (Araújo, 160)

The non-human counterpart of *poro* is *maʔe* which, besides indicating an indeterminate non-human participant, is a regular lexical root meaning ‘thing’. The predicate that incorporates *maʔe* becomes intransitive (218) and, as such, may be causativized and have its valency increased (218b):

- (218) a. Amaʔeʔu neswi
 a-**maʔe**-ʔu ne=∅-swi
 1SG-ANTIP-eat 2SG=R₁-from
 ‘I eat (something/things) without you.’ (AA, 43)
- b. Aβa nojmomaʔeʔuj
 aβa-∅ na-o-i-**mo-maʔe**-ʔu-j
 person-REF NEG-3-R₂-CAUS-thing-eat-NEG
 ‘The man does not feed it.’ (Araújo, 11)

Although synchronically *poro* and *maʔe* seem to have a similar function, only differing according to the human/non-human parameter, they clearly have a different origin. While *poro-* is a prefix, *maʔe* ‘thing’ is a lexical item which, as such, can be incorporated. In (217c), the causative follows *poro-*, while in (218b), it precedes *maʔe*.

5.7.4 Reflexive, middle, and reciprocal voice

This section describes reflexive and middle constructions. Both types of constructions are in the same section for two reasons: (i) as observed by Zúñiga (2020, 151), many authors within the functional-typological tradition have treated them in a somewhat related way; and (ii) the same morpheme marks both functions in TUP, as in other TG languages.

5.7.4.1 Reflexive

A reflexive construction is a grammatical construction with two coreferential participants which are marked by a special form, a reflexivizer, which signals this coreference (Haspelmath 2021a)³¹. These constructions have one argument but two semantic roles. In RRG terms, this means that for the construction to be grammatical, the reflexive pronoun must not be higher in the AUH (see Fig. 3.18) than its antecedent. This means that RRG treats reflexivity semantically, i.e., with actors binding undergoers or agents binding patients, but not the other way around.³²

Tupinambá uses the reflexive voice marker prefix as its reflexivization strategy (see Haspelmath 2021a and Van Valin Jr and LaPolla 1997, 392-417). This takes the form of a verbal prefix, *je-*, which indicates the coreference of two participants of a verb, i.e., the actor and undergoer are linked to the same argument index. This morpheme is always bound to the predicate, occupying the slot reserved for the undergoer as in (219). This prefix reduces the M-transitivity of the predicate, since it combines with morphemes exclusively used with intransitive verbs, such as the causative in (221). Some examples of the reflexive construction are given in (219). Note that these are examples of a complete reflexive construction. Thus, they cannot be analyzed as involving coreference between two distinct referring expressions. Rather, this construction may be best analyzed as involving the linking of actor and undergoer to the same argument marker, as the representation of (219a) given in Figure 5.12 indicates.

³¹This section will not treat coreference within the clause that can be expressed in other ways than with a reflexive marker (see Haspelmath 2021a).

³²That the relationship between the antecedent and the reflexive is semantically motivated seems to be corroborated by the fact that there are no cases of objects binding subjects or patients binding agents. According to RRG, the more semantically motivated a feature, the less cross-linguistic variation it will show. Regarding how far away the reflexive can be from the antecedent, this is a syntactically motivated issue, and therefore more cross-linguistic variation is not only expected but is indeed what one observes (see Van Valin Jr and LaPolla 1997, 389-418, 604-615).

- (219) a. Ajeka
a-**je**-ka
1SG-**RFLX**-break
'I broke myself.' (VLB II, 92)
- b. Ajeāj
a-**je**-āj
1SG-**RFLX**-wrinkle
'I wrinkle myself (I frown).' (VLB I, 117)

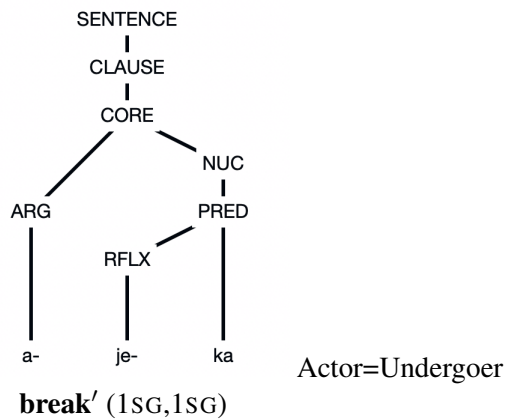


Figure 5.12: Reflexive linking in TUP

In the case of a possessive predication, the reflexive may be preceded by a relational of non-contiguity, as in (220).

- (220) iβíramo ijemojaŋine
iβir-ramo i-**je**-mojaŋ=ne
earth-TRSL R₂-**RFLX**-transform=FUT
'He will transform himself (in something) like the earth.' (DC I, 161)

- (221) a. Ajeran sese
a-**je**-ran s-ese
1SG-**RFLX**-rude R₁-against

‘I get angry at him.’ (VLB, II, 103)

b. Mojeran

mo-**je**-ran

CAUS-**RFLX**-rude

‘Cause to irritate oneself.’ (VLB, II, 89)

It is not uncommon for coreferential constructions not to be treated as reflexive constructions Haspelmath (2004). The relational morpheme (R_3) (see Section 4.3.2) could well be considered an anaphoric adpossessor modifying the object and be interpreted as coreferential with the subject.

(222) a. Pedro ojuka oguβa

Pedro o-i-juka o-uβ-a

Pedro 3- R_2 -kill R_3 -father-REF

‘Pedro killed his own father.’ (AA, 16)

b. Otupānamo taferereko

o-tupã-ramo ta-fe=r-ereko

R_3 -God-TRSL HORT-1SG= R_1 -treat

‘May I be as their own God.’ (Araújo, 160)

5.7.4.2 Middle

The middle voice is more difficult to define than the reflexive. The first reason is that different grammatical traditions define it by different criteria; however, the main reason is that what is referred to by the term ‘middle voice’ in the literature has a wide range of meanings (see Zúñiga 2020, 171). Thus, there seems to be no agreement on what counts as a middle marker cross-linguistically (Inglese 2021).

As with reflexives, middle constructions have a unique referent but two semantic roles. The ACT is simultaneously the causer and UND (patient or goal), but as Kemmer (1994, 181) points out, there is a semantic property which subsumes the notion of subject-affectedness that is crucial to the nature of the middle. This semantic property, which she terms ‘relative elaboration of events’, ‘is the parameter along which the reflexive and the middle can

be situated as semantic categories intermediate in transitivity between one-participant and two-participant events, and which, in addition, differentiates reflexive and middle from one another’.

Here, the ten situation types³³ or pragmatic contexts from Kemmer (1994) are employed to categorize the middle voice in Tupinambá. This combination differentiates the middle from the reflexive construction.³⁴ Although the middle marker and the reflexive voice marker are the same, *je-*, the marker in middle function is here glossed as MID.³⁵

(223) Grooming or body care

ʃeakaŋa mojewaka
 je=∅-akaŋ-a mo-**je**-wak-a
 1SG=R₁-head-REF CAUS-MID-embellish-GER
 ‘Adorning my head.’ (Poemas, 152)

(224) Nontranslational motion

- a. Ajereβjereβ
 a-**je**-reβ-jereβ
 1SG-MID-turn.over-RED
 ‘I keep on turning over (and over).’ (VLB, I, 127)
- b. Ojeaiβik
 o-**je**-aiβik
 3-MID-lower.the.head
 ‘He lowered his head.’ (Araújo, 63v)

³³‘The fact that these situation types recurrently cluster together in the languages of the world, i.e., are expressed by the same marker in a given language across many languages, suggests that the middle is a linguistic category with the potential for grammatical instantiation.’ (Kemmer 1994, 183).

³⁴The middle-reflexive opposition here is similar to the opposition of reflexives and false reflexives in Dixon (1972, 89-94).

³⁵For the relationship between middle and reflexive markers, see Inglese (2022). One interesting finding of Inglese (2021) is that middle markers ‘are most conspicuously associated with anticausative/spontaneous events and with verbs of translational motion, and less so with grooming and non-translational motion situations, which in Kemmers (1993) view represent the semantic middle prototype’.

(225) Change in body posture

- a. Erejeapik
 ere-**je**-apik
 2SG-**MID**-sit
 ‘You sit (down).’ (DC, II, 92)
- b. Ajepiso witupa
 a-**je**-piso wit-uβ-a
 1SG-**MID**-stretch 1SG_{CORF}-lie-GER
 ‘I lie stretched out.’ (VLB, I, 129)

(226) Translational motion

Jeupir
je-upir
MID-go.up
 ‘Rise / go up.’ (VLB, II, 119)

(227) Indirect middle

- a. Ajerure neβe ʃeremiʔurama
 a-**je**-rure ne=βe ʃe=r-emi-ʔu-ram-a
 1SG-**MID**-request 2SG=DAT 1SG=R₁-RES-ingest-FUT-REF
 rese
 r-ese
 R₁-for
 ‘I ask you for my food.’ (D’Evreux, Viagem, 144)
- b. Tekorama ri jeapisaka
 t-eko-ram-a Ø-ri **je**-apisaka
 R₄-be-FUT-REF R₁-POSP **MID**-pay.attention
 ‘Pay attention to future deeds.’ (Araújo, 19v)

(228) Emotion middle

Ajemoirō
 a-**je**-mo-irō
 1SG-**MID**-CAUS-angry

‘I got angry.’ (Teatro, 44)

(229) Emotive speech actions

a. Ejapirō

e-**je**-apirō

2SG.IMP-MID-complain

‘Complain.’ (Teatro, 44)

b. Tijeroβjar

apo aβa

ri

t-ja-**je**-roβjar

apo aβa-∅

ri

HORT-1PL.INCL-MID-believe this man-REF OBL

‘May we trust these men.’ (Léry, Histoire, 354)

(230) Cognition middle

Ojeanerekóβo

oajajpawera

rese

o-**je**-anerekó-βo

o-ajajpaβ-wer-a

r-ese

3-MID-thinking-GER CORF-wickedness-PST-REF R₁-about

‘Thinking about your wickedness.’ (Araújo, 74v-75)

(231) Spontaneous events

Jekij!

je-kij

MID-grow

‘To grow! (person, animal, tree).’ (VLB, I, 85)

Once the whole Tupinambá corpus is fully available in searchable format it will be possible to provide a full account of verbs that can take the middle voice marker based on the definition requiring that the construction has with the following characteristic, from Inglese (2021):

- i. it occurs with bivalent (or more) verbs to encode one or more of the following valency changing operations: passive, anticausative, reflexive, reciprocal, antipassive;
- ii. the same construction is also obligatory with some (at least monovalent) verbs that cannot occur without middle markers;

- iii. the semantics of (at least some of) the verbs in (i) does not match that of those in (ii) or vice versa.

5.7.5 Reciprocal verb constructions

Adding *jo-* to an M-transitive predicate in the undergoer slot creates a predicate in whose logical structure, predicate'(x,y), the x and y arguments are simultaneously reciprocal. The prefix *jo-* allows both actors to be merged into one macrorole, leaving the undergoers of the action implicit. Example (232) shows a reciprocal predicate with its logical structure.

- (232) Pejojuka
 pe-**jo**-juka
 2PL-RECP-kill
 'You kill each other.' (FA, 80)
do' (2PL, ∅ CAUSE [BECOME **dead'** (3)] ∧ **do'** (3, ∅ CAUSE [BECOME **dead'** (2PL)]

- (233) a. Aβa mojoamotare?imuka
 aβa-∅ mo-**jo**-amotar-e?im-uka-∅
 person-REF CAUS-RECP-like-NEG-CAUS-GER
 'Causing people to hate one another.' (Diálogo, 215)
- b. Orojoapiapi
 oro-**jo**-api-api
 1PL.EXCL-RECP-hit-RED
 'We keep on hitting each other.' (VLB, II, 32)

When postpositions combine with the reflexive, the reciprocal *jo-* may alternate with the reflexive *je-* having reflexive function, as in 234.

- (234) a. Atupãmoŋeta jejoese
 a-tupã-moŋeta je=**jo**-ese
 1SG-God-talk 1SG=**RFLX**-POSP
 'I pray for myself.' (FA, 82)

- b. Pedro toimoŋeta ojoese
 Pedro t-o-i-moŋeta o-**jo**-ese
 Pedro HORT-3-R₂-pray 3-**RFLX**-POSP
 ‘May Pedro pray for himself.’ (FA, 82)
- c. Ojeswi imoʔëuka ase rese
 o-**je**-swi i-mo-ʔë-uka ase r-ese
 3-**RFLX**-from R₂-CAUS-leak.out-CAUS.GER we R₁-POSP
 ‘The outpouring of himself for us.’ (Araújo, 43)

Speaking about reciprocals, Haspelmath (2021a) observes that there is a universal of reflexive constructions according to which ‘if a language has a reflexive voice marker, it also has a voice marker for reciprocal constructions’ (see also Dixon 2010b, 141). In TUP, the reciprocal voice marker is *jo-*. The controller in a reciprocal construction must have a plural reference, as in (235).

- (235) Orojoapi
 oro-**jo**-api
 1PL.EXCL-**RECP**-hit
 ‘We hit each other.’ (VLB, II, 32)
do' (1PL.EXCL, [**hit'** (1PL.EXCL, 2SG)]) \wedge **do'** (2SG, [**hit'** (2SG, 1SG)])

In (235), the actor is talking to someone other than the undergoer, since the form *oro-* excludes the hearer, in contrast to (236). Nonetheless, both cases remain ambiguous regarding the grammatical number of the undergoer, as it could be plural or singular.

- (236) Mewue jajomoŋeta
 mewue ja-**jo**-moŋeta
 low.volume 1PL.INCL-**RECP**-talk
 ‘We talk to each other quietly.’ (Teatro, 148)

Dixon (2010b, 147-151) notes that reciprocal constructions are also possible when the number of participants is greater than two, whether they are specified

or not. Such a case would allow for different interpretations. Consider (237), where the topic of the discourse is the inhabitants of a village. At one point, Aimberé talks about some of them as in (237):

- (237) Ojoapisapisapa
 o-**jo**-apisaβ-pisaβ-a
 3-**RECP**-wound-RED-GER
 ‘(They are) wounding each other continuously.’ (Teatro, 36)

Here it is not necessarily implied that all of them wound each other (full reciprocal), but that some – not all – of them wound each other.

5.7.6 Is there a passive voice in TUP?

Both authors of the grammars in Anchieta (1595, 35-35v) and Figueira (1687, 86, 90-91) agree that the prefix *je-* ‘reflexive marker’ can also be used to indicate the passive, as in (238):

- (238) a. Ojenuβ
 o-je-enuβ
 3-RFLX-hear
 ‘It is heard.’ (Anchieta, 35)
- b. Ajemojāj
 a-je-mojāj
 1SG-RFLX-make
 ‘I am made.’ (Anchieta, 35)
- c. Aje?u
 a-je-?u
 1SG-RFLX-eat
 ‘I eat myself / I am eaten.’ (FA, 90)
- d. Ajejuka
 a-je-juka
 1SG-RFLX-kill
 ‘I kill myself / I am killed.’ (FA, 86)

Nonetheless, there is no attestation (in either Anchieta or Figueira) of a reflexive construction with the optional oblique argument (the effector), i.e., an agentless passive. The example (238a) is the only example attested which is not third person in any TUP text.

In Araújo (1618b), some occurrences of *je-* with no reflexive meaning are actually instances of an impersonal construction.

- (239) Tojemojãŋ neremimotara
 t-o-je-mojãŋ ne=r-emi-potar-a
 HORT-3-RFLX-do 2SG=R₁-DEV_{PASS}-want-REF
 ‘May your will be done / may one do your will.’ (Araújo, 13v)

In the recently transcribed TUP letters (Navarro 2022), the following example is found. Navarro (2022) adds a footnote saying that the reflexive marker became a passive marker in Colonial Tupi, but the evidence supporting this claim is scarce.

- (240) Nojeme?eŋi jeí ãwa supe quartel
 n-o-je-me?eŋ-i jeí ãwa supe quartel
 NEG-3-RFLX-give-NEG today.past DEM to quartel
 ‘Their lives were not spared today.’ (CC, 1)³⁶

The lack of examples with an effector expressed by an oblique constituent, along with the fact that *je-* is rarely attested with a passive meaning, and when it does occur, this can be interpreted as an impersonal, may be taken as evidence that TUP did not have a passive construction. Either the Jesuits misunderstood the matter, influenced by the impersonal construction in Portuguese, which uses the same marker as the passive construction (*-se*), or the lack of examples is just a coincidence, though an improbable one. The lack of a passive is also attested for Old Guaraní, where *je-* has either a reflexive, middle, or impersonal reading (see Restivo 1724, 63). The examples of passive constructions provided in this section probably exhaust the examples of passive

³⁶‘To give quarters/barrack’ was synonym to ‘spare one’s life’.

constructions in the whole TUP corpus. The passive voice should not be considered part of the grammatical inventory of TUP. The evidence from the texts and the comparison with other TG languages support this fact. The examples in (238) are difficult to explain and could be a case of imperfect learning or a Jesuit attempt to shape the language.

Chapter 6

The Layered Structure of the Clause

Following the general discussion of the LSC in Section (3.1.1), this section presents the LSC and syntactic templates for TUP beyond the basic clause patterns presented in Chapter 5.

6.1 PrDP

The PrDP hosts elements set off by a pause, such as adverbials or topical information. There is often coreference of an argument in the clause with an element in the PrDP, such as a resumptive pronoun, as in (241b), where the free pronoun *ene* ‘you’ marks the topic. The syntactic template for the PrDP is given in Figure 6.1. The representation of (241a) is given in Figure (6.2).

- (241) a. Kwese, karáíβari ipokoki
k^wese karáíβ-a-ri i-pokok-i
yesterday non-indian-REF-POSP R₂-attack-NFOC

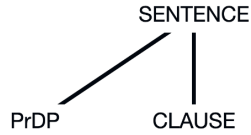


Figure 6.1: PrDP template

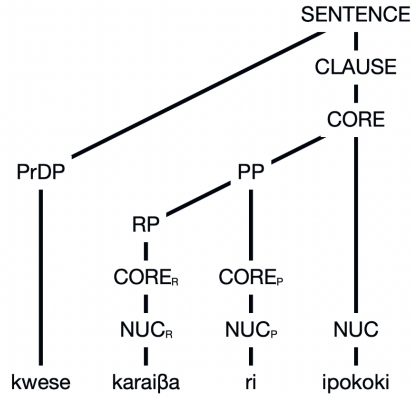


Figure 6.2: Sentence with pre-detached position

‘Yesterday, the white men were attacked.’ (AT, 30)

- b. [[Ene]_{PrDP}, [[nejiβápe]_{PERIPHERY} [Jesu]_{ECS}
 ene ne=∅-jiβa-pe Jesu
 you 2SG-REF-arm-LOC Jesus
 [eresupi]_{CORE}]_{CLAUSE}]_{SENTENCE}
 ere-s-upi
 2SG-R₂-lift

‘You, you carried Jesus in your arms.’ (Poemas, 118)

Another example of a topical RP in the PrDP is given in (242):

- (242) Tupinamba Parawasupenarwera, itupã
 Tupinamba Parawasu-pe-sar-wer-a i-tupã
 Tupinambá Paraguasu-LOC-NMLZ-PST-REF R₂-God
 osiβaʔep^wera opakatu jamopa
 o-sik-βaʔe-p^wer-a opa-katu ja-mo-paβ
 3-rub-REL-PST-REF all-INTS 1PL.INCL-CAUS-finish
 ‘The Tupinambá who were in Paraguasu, who rubbed (the statue of)
 their Gods, we exterminated them.’ (Teatro, 16)

There can be multiple units in the PrDP, as in 243.¹

- (243) Jemoirõ, morapiti, joʔu, tapuja
 je-moirõ-∅ poro-apiti-∅ jo-ʔu-∅ tapuj-a
 RFLX-anger-REF ANTIP-slay-REF RECP-eat-REF foreigner-REF
 rara, awasá, moropotara, majana,
 r-(j)ar-a awasá-∅ poro-potar-a majan-a
 R₁-capture-REF concubinage-REF ANTIP-want-REF pimping-REF
 siwaraji, najpotari aβá
 siwaraji-∅ n-a-i-potar-i aβá-∅
 prostitution-REF NEG-1SG-R₂-want-NEGman-REF R₂-abandon-REF
 sejara
 s-ejar-a

‘Getting angry, slaughtering people, eating each other, capturing tapuias,
 concubinage, sensual desire, covetousness, prostitution, I don’t want
 anyone to abandon these.’ (Teatro, 10)

6.2 Pre-core slot (PrCS)

The PrCS is restricted to main clauses due to its association with contrastive focus (244c), since the scope of illocutionary force does not extend to out-

¹One could argue that all the pre-core elements in this example are in the ECS, but since both editions of the *Teatro* have a dash instead of a comma, I take this to be a sign of a longer pause, and thus the PrDP is an interpretation that better fits the pragmatics of the sentence. Moreover, the ECS does not host multiple elements (phrases), whereas the PrDP does.

side the clause level. It is the position for focal (narrow-focus) elements and WH-words in languages in which these occur *ex situ*, such as TUP. It can be occupied by focal arguments and adjuncts, as in (244), in which case the adjuncts are not set off by a pause, as when they occur in the PrDP (see Section 6.1). The PrCS also hosts question words, as in (244e).

- (244) a. Awje kunumĩwasu oekoaiβete ojomim
 awje kunumĩ-wasu-∅ o-eko-aiβ-ete o-jo-mim
 finally boy-big-REF 3CORF-deed-evil-INTENS 3-R₂-hide
 ‘Finally the boys hide their evil deeds.’ (Teatro, 40).
- b. Emonã kori aikóne
 emonã kori a-iko=ne
 thus today 1-be=FUT
 ‘Today I shall act this way.’ (Araújo, 99v)
- c. Peróte toso
 Pero=te t-o-so
 Pedro-FOC HORT-3-go
 ‘May Pedro (not someone else) go.’ (VLB, I, 36)
- d. Sekóte ipofiete
 s-eko=te i-pofi-ete
 R₂-life-FOC R₂-bad-INTS
 ‘His life (not something else) is very bad.’ (Teatro, 30)
- e. Mba?epe ke kanineoβi jaswara?
 mba?e=pe ke kanine-oβi-∅ jaswar-a
 thing=Q here macaw-blue-REF similar-REF
 ‘What is here similar to a blue-yellow macaw?’ (Teatro, 64)
- f. Aβaβépe Tupã noimoetei?
 aβa-βe=pe Tupã n-o-i-mo-ete-i
 person-also-Q God NEG-3-R₂-CAU-INTENS-NEG
 ‘Who else does not honor God.’ (Araújo, 66)
- g. Mamõpe a?e iβoja sow a?erire?
 mamõ=pe a?e i-βoja-∅ so-w a?e-r-ire
 where=Q DEM R₂-disciple-REF go-NFOC this-R₁-after

‘Where did these disciples of his go afterwards?’ (DC, I, 170)

WH-words in TUP always occur *ex situ*, in the PrCS. Some of these words are shown in Table 6.1. All words are given with the question clitic =pe. The words without the question clitic and their meanings are given in the two rightmost columns.

WH-word	Meaning	Lexeme	Meaning
Ma?e=pe	what, which	ma?e	thing
Marã=pe	which, how	marã	
Marãmarã=pe	wich (plural)		
Ma?ema?epe	what, which (plural)		
Aβa=pe	who	aβa	person
Aβaaβa=pe	who (plural)	aβa	person
Marãjatu=pe	how		
Mamõ=pe	where	mamõ	where
Umã=pe	where	umã	where
Moβipe	how many		
Marãramo=pe	Why	marã + ramo	anything + translativo

Table 6.1: Some WH-words in Tupinambá

- (245) a. Ma?epe erejpotar?
ma?e=pe ere-i-potar
thing=Q 2SG-R₂-want
 ‘What do you want?’ (Léry, 347)
- b. Ma?epe amõ?
ma?e=pe amõ
thing=Q other
 ‘Which other? / What else?’ (Léry, 343)
- c. Marãpe peroβajara rera?
marã=pe pe=r-oβajar-a r-er-a
What=Q 2PL=R₂-enemy-REF R₂-name-REF
 ‘What is the name of your enemies?’ (Léry, 354)

- d. Moβipe tuβisakatu kiβð?
moβi=pe t-uβisáβ-katu-∅ kiβð
how.many=Q R₄-chief-good-REF around.here
 ‘How may great chiefs are there around here?’ (Léry, 350)

6.3 Extra-core slot (ECS)

In head-marking languages² such as TUP, bound argument indexes saturate the valency requirements of the predicate (Van Valin Jr 1977, 1985, 2013) (see Section 5.7.2). Syntactically optional RPs coreferential with the bound argument indexes are not core arguments because the core arguments are bound to the head. Following Haspelmath (2013), I use the term *conominal* to refer to these RPs. Van Valin Jr (2013) places these RPs in the ECS, a position only found in head-marking languages. This position is structurally analogous to the PrCS or PoCS because it is also a daughter of the clause node. These RPs must be instantiations of the core arguments with no fixed order in relation to the core (see Section 5.6), as in (246):

- (246) a. [[Ajmomaʔeete]_{CORE} [ne-r-oka]_{ECS}]_{CLAUSE}
 a-i-mo-maʔe-ete ne=r-ok-a
 1SG-R₂-CAUS-thing-good 2SG=R₁-house-REF
 ‘I honour your house.’ (Poemas, 170)
- b. [[Nerokangaturamwama]_{ECS} [orojmoĩ]_{CORE}]_{CLAUSE}
 ne=r-ok-angaturamwam-a oro-i-mo-in
 2SG=R₁-house-holy-FUT-REF 1PL.EXCL-R₂-CAUS-be.still
 ‘We build your holy house.’ (Poemas, 146)
- c. [[Osap^jape]_{CORE} [ase ijeʔenga]_{ECS}]_{CLAUSE} ...?
 o-s-ap^ja=pe ase i-jeʔeŋ-a
 3-R₂-obey=Q PRON R₂-speech-REF
 ‘Will one obey our words?’ (DC, I, 224)

²It is worth mentioning, as observed by Bohmeyer et al. (2016, 182), that *head-marking* has a more restricted meaning in RRG than it originally had in Nichols (1986).

- d. Noik^waβipe taʔa kawaramo
 n-o-i-k^waβ-i=pe taʔa kaʔu-ar-amo
 NEG-3-R₂-know-NEG=Q sir beer.drink-NMLZ_{AG}-TRSL
 ʃereko?
 ʃe=r-eko-∅
 1SG=R₁-be-REF
 ‘Doesn’t the master know that I am a drinker?’ (AT, 136)
- e. Nojnupäiswétepe aβa oaʔira
 N-o-i-nupã-i-swe-te=pe aβa-∅ o-aʔir-a
 NEG-3-R₂-hit-NEG-NEG.FUT-FOC=Q man-REF CORF-son-REF
 oemiawsuβane?
 o-emiawsuβ-a=ne
 CORF-slave-REF=FUT
 ‘But won’t the man punish his own son and his own slave?’ (Araújo,
 69v)

It has been claimed by Haspelmath (2013) that Van Valin Jr and LaPolla (1997) and Bresnan and Mchombo (1987) support the ‘dual-nature view’ of bound arguments, according to which the presence of a lexical RP makes the bound arguments agreement markers, and in their absence, the bound arguments are the arguments (see Bohnemeyer et al. 2016). In fact, RRG does not subscribe to the analysis in Bresnan and Mchombo (1987) on RPs functioning as subject and agreement preferring the term ‘coreference’ to ‘agreement’ since it considers the latter notion more Eurocentric than universal. Additionally, Van Valin Jr (2013) considers co-indexed RPs as being pragmatically unrestricted, since they can be topical or focal – but not the subject as in Bresnan and Mchombo (1987). In RRG, bound argument indexes are the core arguments regardless of the presence of co-indexed RPs. Van Valin Jr (2013) argues that bound argument markers are pronominal anaphors, capable of being locally bound or independently referential. RRG also distinguishes between clause-internal topics, located in the ECS, and clause-external topics, located in the detached positions, a distinction not made in Bresnan and Mchombo (1987). In this sense, the analysis by Siewierska (2001) is not accepted here because it is unclear regarding the status of bound indexes in head-marking

languages.

The ECS differs from the PrCS and PoCS in important ways (Van Valin Jr 2013). Firstly, the elements in the ECS are not associated with a specific pragmatic or discourse function, i.e., they can be focal or topical. When focal, these RPs are usually morphologically marked, as in (247a) (repeated from 244c). Secondly, they are not positionally restricted (see Section 5.6), although RPs marked by the focal *-te* do not appear postcore, because TUP has no PoCS, thus the ungrammaticality of (247b). Third, the PrCS/PoCS may instantiate arguments and adjuncts, while the RPs in the ECS must be instantiations of arguments.

(247) a. Perote toso.
 Pero_i-te t-o_i-so
 Pero-FOC HORT-3-go
 ‘May Pedro (not someone else) go.’ (VLB, I, 36)

b. *Toso Perote.
 t-o_i-so Pero_i-te
 HORT-3-go Pero-FOC
 ‘?’

(248) [Oso]_{CORE} [[βe]_{PERIPHERY} [amõ maranaritekoara]_{ECS}]_{CLAUSE}.
 o_i-so βe [amõ maranaritekoar-a]_i
 3-go again other soldier-REF
 ‘Other soldiers went also.’ (Araújo, 64)

In summary, the ECS is licensed by the cross-reference markers on the nucleus and therefore occurs exclusively in head-marking constructions restricted to cross-referenced RPs, whereas the PrCS and PoCS can accommodate other syntactic categories. A clause has exactly as many ECSs as its nucleus or nuclei carry cross-reference markers (whereas every clause has exactly one PrCS and PoCS).

6.4 The periphery

Section 3.1.1 showed that the core hosts the predicate and its arguments. The periphery is the place where non-arguments are hosted, which can be of two types: phrasal adjuncts such as PPs, and non-phrasal adjuncts such as adverbs. The distinction between the core and the periphery thus corresponds to the distinction between arguments and non-arguments. There is a periphery for each of the following levels: nucleus, core, and clause, because adjuncts have scope over specific levels.

The nuclear periphery contains aspectual adverbs such as *completely* and *continuously*. Example (249), with its syntactic representation in Figure 6.3, shows the adverb *paβ*³ ‘terminate, completely’ modifying the nucleus.

- (249) Aimaʔepisirōpaβ
 a-i-maʔe-pisirō-paβ
 1SG-R₂-thing-appropriate-completely
 ‘I took his things completely.’ (VLB, I, 100)

- (250) Oʔu api ahē maʔe
 o-ʔu api ahē maʔe-∅
 3-eat completely INTJ thing-REF
 ‘He completely eats things.’ (VLB, II, 52)

Adjuncts like temporal adverbs (e.g. *tomorrow*, *yesterday*) and manner adverbs (e.g. *quickly*, *carefully*, *violently*) modify the core when they express locational or temporal features of the state of affairs coded by the core. Examples of with temporal adverbs are given in (251), with the syntactic representation of (251a) given in 6.4. Manner adverbs are shown in (252), where they relate to pace and performance.

³*Paβ* is a lexical root meaning ‘terminate, all, finish’. One could well argue that it has grammaticalized as a completive aspectual marker, but since it continues to be used as a lexical root, I prefer to view this as a lexical root which, like others, can be used as a modifier.

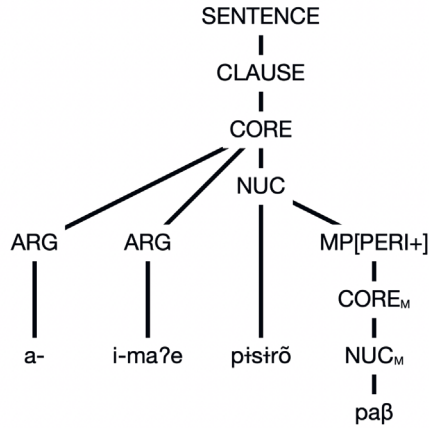


Figure 6.3: Nuclear peripheral modifier

- (251) a. Aseja k^wese ʃeroka
 a-s-eja(r) k^wese ʃe=r-oka
 1SG-R₂-abandon yesterday 1SG=R₁-house-REF
 ‘I left my house yesterday.’ (Poemas, 112)
- b. Tasep^ʃak tawje
 ta-s-ep^ʃak tawje
 HORT-R²-see soon
 ‘May I see them soon.’ (Léry, Histoire, 345)
- (252) a. Oporomoʔe aʔu
 o-poro-mo-ʔe aʔu
 3-ANTIP-CAUS-say false
 ‘He teaches people erroneously.’ (AT, 128)
- b. Aʔe umani^ĩ maʔe mojaŋa
 aʔe umani^ĩ maʔe-∅ mojaŋ-a
 DEM slowly thing-REF make-GER
 ‘(Me) doing things slowly.’ (AA, 56v)

The clausal periphery contains epistemic adverbs like *probably* and evidentials

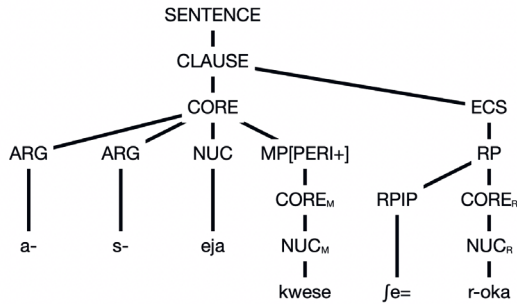


Figure 6.4: Core peripheral modifier

like *evidently*, speech act modifiers like *honestly*, and speaker attitude/judgement adverbs like *unfortunately*. A clausal peripheral modifier can be seen in (253), represented in Figure 6.5.

- (253) Ja omanõβo
 ja o-manõ-βo
 luckily 3-die-GER
 ‘Luckily he dies.’ (FA, 163)

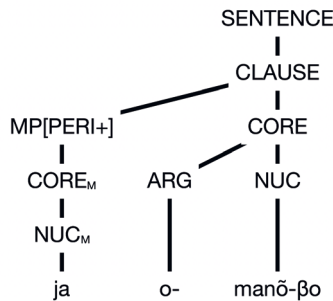


Figure 6.5: Clausal peripheral modifier

Sentences with multiple adverbs are not frequently attested, so it is not possible to verify, as suggested by Van Valin Jr (2005, 20), if in a sentence with more than one adverb, the adverbs are constrained by the LSC, i.e., whether

adverbs related to outer operators occur further from the predicate and vice-versa.

6.4.1 Modifier Phrase

Van Valin Jr (2008b) proposed the notion of a modifier phrase (MP) to accommodate attributive modifiers of all types, as well as adverbial modifiers and adpositional phrase modifiers. Thus, the primary syntactic categories in RRG are the RP, clause, adpositional phrase, and MP. These are related to the three speech-act functions mentioned in Chapter 4: reference, modification, and predication. MPs occur in the peripheries of the element modified, and they are not involved in predicative uses of lexical roots or adverbials. MPs, like PPs and RPs, also have a layered structure, a $core_M$ and a nuc_M , as illustrated in Figure (6.6), which represents the syntactic structure of Example (254) with a peripheral modifier at the nuclear level (degree modifiers).

- (254) Kunumĩporanga
 kunumĩ-poraŋ-a
 boy-beauty-REF
 ‘Beautiful boy.’ (Poemas, 164)

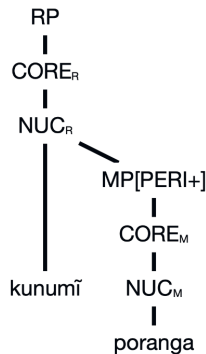


Figure 6.6: Modifier phrase

Manner adverbial modifiers are at the core level. Example (255) and its rep-

resentation in Figure (6.7) show an MP in the nuclear periphery. Note that degree modifiers like *ete* do not head phrases, and therefore cannot be in an MP.

- (255) Tupã siporaŋete
 Tupã Ø-si-poraŋ-ete
 God R₁-mother-beauty-INTS
 ‘Very beautiful mother of God.’ (Poemas, 82)

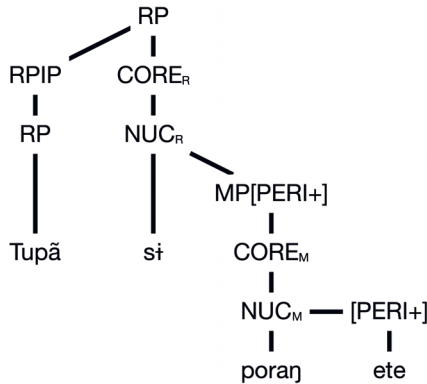


Figure 6.7: Nominal modification

More examples of MPs will be discussed in Chapter (8) on RPs.

6.5 Operators

In RRG, elements that are in a whole domain of their own because they represent grammatical categories that are qualitatively different from predicates and their arguments are called operators. They are organized according to the range of their scope, i.e., according to which level they modify: the whole clause, the core, or the nucleus, as shown in Figure (6.8).

The next sections present Tupinambá operators according to their scope. Operators modifying the clause are presented first, followed by operators modifying the core and subsequently followed by operators modifying the nucleus.

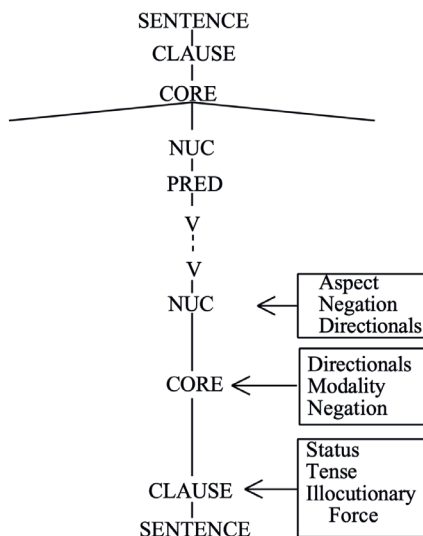


Figure 6.8: RRG operator projection

<i>DEC</i>	<i>IMP</i>	<i>INT</i>	<i>R</i>	<i>ATS</i>	<i>NAT</i>	<i>NFUT</i>	<i>SG</i>	<i>DUR</i>
⟨ _{IF}	⟨ _{STA}	⟨ _{IR}	⟨ _{EVID}	⟨ _{TNS}	⟨ _{EVG}	⟨ _{PL}	⟨ _{ASP}	⟨ _{NEG}
<i>HOR</i>	<i>IR</i>	<i>NAT</i>	<i>AT3</i>	<i>FUT</i>	<i>PL</i>	<i>ITER</i>	<i>LS</i>	⟩⟩⟩⟩⟩⟩⟩⟩⟩⟩⟩⟩⟩⟩⟩⟩
<i>OPT</i>						<i>PUNC</i>		

Table 6.2: Tupinambá operators

It is worth stressing that it is a difficult task to work out the operators from written texts alone without access to native speakers.

6.5.1 Clause-level operators

Clause-level operators (IF, evidentials, status, and tense) show a binary grouping, with one group containing tense (TNS) and status, (ST), and the other evidentials and illocutionary force (IF). Tense and status situate the proposition expressed by the clause within temporal and realis-irrealis continua (see

Comrie et al. 1985; Hornstein 1993). Evidentials indicate the epistemological basis of the state of affairs expressed, i.e., they indicate how the speaker came to be aware of the information uttered (Aikhenvald 2004), while illocutionary force specifies the type of speech act.

6.5.1.1 Illocutionary force

The illocutionary force (IF) operator modifies the clause, not just one of its constituent clauses. It occurs only in main clauses, i.e., clauses immediately dominated by the sentence node. Languages typically have three basic sentence types corresponding to the three types of illocutionary force⁴: declarative, imperative, and interrogative sentences. Bybee (1985, 22) defines illocutionary force ('mood') as an indication of 'what the speaker wants to do with the proposition' in a particular discourse context. In other words, IF is a grammatical reflection of the speaker's purpose in speaking. It would appear that every language has the means to express the major types of illocutionary force, and many can also express minor types, such as those given in Table (6.3). These categories of illocutionary force will comprise a system within a language and will be mutually exclusive, since it is impossible to mark a sentence as both declarative and imperative, for instance.

	Speech act	Sentence Type
Major	Assertion	Declarative
	Command	Imperative
	Question	Interrogative
Minor	Exhortation	Hortative
	Wish / Hope for	Optative
	Assert not true	Subjunctive

Table 6.3: Categories of Illocutionary Force

⁴Van Valin Jr and LaPolla (1997, 41) state that 'modality, status, and illocutionary force are all conflated in traditional grammar under the term "mood". Following this approach, the term 'mood' is not used in this work.'

The examples in (256) show some matches between speech act and sentence. The example in (256a) was already perceived by Anchieta as a mismatch.

- (256) a. Asótepe ise?
 a-so-te=pe ise
 1SG-go-FOC=Q I
 ‘Did I go (by the way)?/I didn’t go!’ (AA, 36)
- b. Assertion in declarative form
 A π a β a karai β a atuasa β a kori ojko
 a π a β -a karai β -a atua-sa β -a kori o-iko
 natives-REF christian-REF companion-NMLZ-REF today 3-be
 ‘The natives and the Christians are friends today.’ (Abbeville, 342)
- c. Question in interrogative form
 Ma ?e pe ereru nekaramemuã pupe?
 ma ?e =pe ere-ero-u ne= \emptyset -karamemuã- \emptyset \emptyset -pupe
 thing=Q 2SG-CAUS.SOC-come 2SG=R₁-box-REF R₁-LOC
 ‘What did you bring in your box?’ (Léry, 343)
- d. Command in imperative form
 Eje ?e ko ?i r!
 e-je ?e ko ?i r
 2SG.IMP-speak now
 ‘Speak now!’ (Staden, 154)
- e. Command in interrogative form
- f. Marãpe nerejemimi?
 marã=pe n-ere-je-mim-i
 why=Q NEG-2SG-RFLX-hide-NEG
 ‘Why don’t you hide yourself?’ (AT, 34)

The association of speech act with sentence type allows for cases of indirect speech acts, in which one illocutionary act is performed indirectly by way of performing another (Searle 1975). Some of these mismatches between IF and sentence type are illustrated in (257) below. A very common type of indirect speech act is the rhetorical question, where the interrogative form is employed

for some purpose other than to ask a question (257a). On the other hand, (257) and (257c) are not mismatches since the speech act in each corresponds to the sentence type. Since these mismatches require a change in intonation, it is difficult to find similar examples from written sources.

- (257) a. Why don't you just be quiet? [command in interrogative form]
 b. Don't tell me you lost it! [question in imperative form]
 c. Who cares? [assertion in interrogative form]
 d. I don't suppose you'd like to buy this from me? [question in declarative form,
 with modified intonation]

Currently, RRG does not have a compatible framework for handling skewing between the form of the utterance's illocutionary act and the intention of its illocutionary act, as illustrated by the English examples in (257).

Declarative

The assertion speech act makes a statement, and the default form of the sentence is declarative and is unmarked as such. An affirmative and a negative statement are given in (258); both are declarative IF.

- (258) Assertion expressed with a declarative sentence:
- a. Iporangete ã teko janéβe
 i-poraŋ-ete ã t-eko-∅ jane=βe
 R₂-beautiful-INTS DEM R₄-thing 1PL.INCL=DAT
 'These things are very beautiful to us.' (Léry, 355)
 ⟨_{IF} DEC **feel'**(jane, [(**be'** (teko, [**beautiful'**]))]) ⟩
- b. Naferoriβi
 na-je=r-oriβ-i
 NEG-1SG=R₁-happy-NEG
 'I am not happy.' (Anch., Arte, 34v)

- c. Ojeaiβik oaseasemamo omanōngatuaβo
 o-je-aiβik o-ase-ase-amo o-manō-katu-aβo
 3-RFLX-lower.the.head 3_{corf}-yell-yell-GER 3_{corf}-die-good-GER
 koite
 koite
 finally
 ‘He lowered his head, yelled repeatedly and finally really died.’
 (Araújo, 92)
 ⟨_{IF} DEC **do**’ (3 [lower.the.head’ (3)]) & SEML **do**’ (3 [yell’ (3)])
 & BECOME **dead**’ (3)⟩

Rhetorical questions are given in (259). In both cases, the focal clitic =*te* seems to signalize the mismatch between sentence type and speech act, i.e., despite being assertions, (259a) and (259b) both have ⟨_{IF} INT⟩ in their logical structures, because in RRG, the logical structure has to match the syntactic structure.

An assertion can also be expressed by a different sentence type, as in (259):

(259) Assertion expressed as an interrogative sentence. Example (259a) is repeated from (256a).

- a. Asotepe ise?
 a-so-**te**=pe ise
 1SG-go-FOC=Q I
 ‘Did I by the way go? (I didn’t go!)’ (AA, 36)
 ⟨_{IF} INT ⟨_{TNS} NFUT **do**’ (1SG, [move.away.from.ref.point’ (1SG)])
 ⟩⟩
- b. Nase retama ruãtepe iko iβi ase
 na-ase r-etama-∅ ruã-**te**=pe iko iβi-∅ ase
 NEG-our R₁-country-REF NEG-FOC=Q DEM land-REF our
 rekoaβa?
 r-eko-aβ-a
 R₁-live-NMLZ-REF
 ‘Isn’t this land where we live, by the way, our country?’ (Araújo, 23)

- c. Nererojeβiriβepotáripe nerekop^wera?
na-ere-ro-jeβiri-βe-potár-i=pe ne=r-eko-p^wer-a
NEG-2SG-SCAU-return-also-want-NEG=Q 2SG=R₁-life-PST-REF
‘Don’t you also wish to return to your past actions?’ (DC, II, 6)

Interrogative

All interrogative sentences, yes-no questions, information questions (question words), and alternative questions in TUP require the clitic =pe⁵. The interrogative clitic =pe has narrow focus; it can attach to any constituent. The examples in (260) show interrogative sentences with WH-words.

(260) WH-word

- a. Aβape ajpo-βaʔe ojmomaran?
aβa=pe ajpo-βaʔe o-i-momaran
person=Q DEM-NMLZ 3-R₂-obey
‘Who obeys that one?’ (Araújo, 67)
- b. Mamōpe Tupã rekow?
marmō=pe Tupã r-eko-w
where=Q God R₁-be-NFOC
‘Where is God?’ (Araújo, 26)
⟨ IF INT ⟨ TNS NFUT **be-at**’ (Tupã,marãpe), ⟩ ⟩
- c. Marãpe ase mojangi?
marã=pe ase mojaŋ-i
how=Q PRON make-NFOC
‘How did he make us? (lit. how was his making of us?)’ (Araújo, 25)
- d. Marãpe nerera?
marã=pe ne=r-er-a
what=Q 2SG=R₁-name-REF

⁵Bickel and Nichols (2005) distinguish two kinds of clitics: phonologically bound words, and those that do not select the category of the host they attach to, i.e., clitics that may attach to any type of word, or even affixed to constituents or clauses. The interrogative =pe and the future marker =ne are of the second type.

‘What is your name?’ (Léry, Histoire, 341)
 ⟨ IF INT ⟨ TNS NFUT **have**’ (2SG[ne],era), ⟩⟩

The following examples are cases of narrow focus in which different constituents are questioned.

(261) PSA questioned

fɛruβape oso?
 fɛ=r-uβ-a=**pe** o-so
 1SG=R₁-father-REF=Q 3-go

‘Did my father go?’ (AA, 36)

⟨ IF INT ⟨ TNS NFUT **do**’ (3[fɛruβ], [**move.away.from.ref.point**’ (3[fɛruβ]))
 ⟩⟩

(262) Predicate questioned

Asope isene?
 a-so=**pe** ise=ne
 1SG-go=Q I=FUT

‘Will I go?’ (FA, 166)

⟨ IF INT ⟨ TNS FUT **do**’ (1SG[ise], [**move.away.from.ref.point**’ (1SG[ise]))
 ⟩⟩

(263) Possessor questioned

Aβa raʔirape ne?
 [aβa r-aʔir-a]=**pe** ne
 person R₁-son-REF=Q you

‘Whose son are you?’ (VLB, I, 87)

Tupinambá WH-words are given in Table (6.4). Some are not attested in the texts, but appear in Anonymous (1952a) .

Command/imperative IF

Imperative sentences can be positive (command) or negative (prohibition). In each case, the sentence is marked as imperative by imperative person markers – given in Table 6.5 – on the verb.

WH-word	Translation	Attestation example
aβápe	who	Teatro, 46
erimaʔe(pe)	when	Araújo, 30v
ke	what size	VLB, II, 91
mamō	where	Araújo, 52v
manōj	whence	VLB, I, 106
marã(pe)	why, how	FA, 98
marãbaʔe	what kind of	Léry, Histoire, 363
marãeteʔi	how	Araújo, 156v
marã(na)mo(pe)	why	VLE, II, 82
marangatueté	how	VLB, I, 77
marangoti(pe)	in what direction	Araújo, 47
mbaʔe(pe)	what, which	Araújo, 43v
mbaʔereme(pe)	in what circumstance	Araújo, 90v
mboβi	how many	Araújo, 107
mojrã(pe)	when (in the future)	Araújo, 46
monomo	how many	VLB, II, 91
nãβo	how many	VLB, II, 91
nãmo / nōmo	what size	VLB, II, 91
umã	where	Teatro, 130
umãβaʔe	which one	DC, I, 212
umãme	where	FA, 127

Table 6.4: Tupinambá WH-words

Person	Index
2SG	e-
2PL	pe-

Table 6.5: Imperative argument indexes

An positive imperative clause is illustrated in (264), while (265c) illustrates a negative imperative, with the obligatory negative imperative marker *umẽ*:

(264) Imperative expressions:

Pejori, peraso muru
 pe-jori pe-era-so muru-∅
 2PL.IMP-come 2PL.IMP-SCAU-go darned-REF

‘Come, bring the darned ones.’ (AT, 92)

⟨_{IF} *IMP do'* (2PL, [**move.towards.ref.point'** (2PL)]) & [**do'** (x, ∅)]
 CAUSE [**do'** (muru, [**move.away.from.ref.point'** (muru)])]⟩

The imperative IF has its own negator, *umẽ*. Examples are given in (265):

(265) a. Eporapiti umẽ!
 e-poro-apiti **umẽ**
 2SG.IMP-ANTIP-slaughter NEG
 ‘Do not slaughter people.’ (Araújo, 69v)

b. Jerenõj umẽ jepe!
 je=r-enõj **umẽ** jepe
 1SG=R₁-call NEG PRON
 ‘Do not invoke my name!’ (Teatro, 32)

c. Ejemoririj umẽ
 e-je-moririj **umẽ**
 2SG.IMP-MID-worry NEG.IMP
 ‘Do not worry (yourself)!’ (AT, 32)
 ⟨_{IF} *IMP* ⟨_{NEG} [**do'** (∅,∅)] CAUSE[**be'** (2SG, [**worried'**])]⟩⟩

Optative

The optative IF expresses an unfulfilled wish/desire. It is realized by *mo* or *temõ* following the predicate and the sentence-final particle *mã* (267). If the event or wish is in the past and thus cannot be fulfilled anymore, it is realized by *meĩĩ* or *meĩĩmo*. With a non-verbal predicate, only the particle *mã*⁶ is used (266c).

⁶See section (7.7).

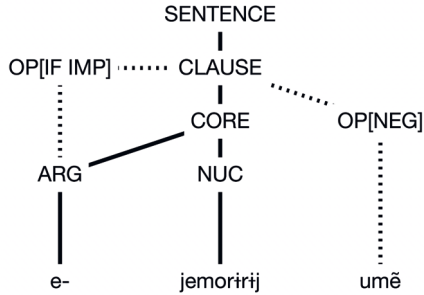


Figure 6.9: Example of an imperative sentence

- (266) a. Aso temō iβakipe mā!
 a-so **temō** iβak-Ø-pe **mā**
 1SG OPT sky-REF-LOC PRCL
 ‘If I could go to heaven!’ (AA, 24)
 ⟨_{IF} OPT [do’ (1SG, [move.away.from.ref.point’ (1SG[a)])] & INGR
 [be-at’ (iβak, 1SG[a])])⟩
- b. Ajukamo mā
 a-i-juka-mo **mā**
 1-SG-R₂-kill-OPT PRCL
 ‘I wish I could kill him.’ (AA, 18)
 ⟨_{IF} OPT [do’ (1SG, [move.away.from.ref.point’ (1SG)])] & INGR
 [be-at’ (iβak, 1SG)])⟩
- c. Akwej ko mā!
 akwej ko **mā**
 DEM here PRCL
 ‘I wish/if only that one were here!’ (DC, 93)

Note that while *mā* is a sentence-final particle, *temō*, *mo*, and *meʔĩ(mo)* always follow the predicate (second position).

- (267) a. Aso meʔĩmo iβak-i-pe mā
 a-so **meʔĩmo** iβak-EPEN-pe **mā**
 1-SG-go OPT sky-REF-LOC PRCL

‘I wish I had gone to heaven.’ (AA, 24)

- b. Ajuka meʔĩ mā
 a-∅-juka **meʔĩ mā**
 1SG-R₂-kill **OPT PRCL**

‘I wish I had killed him.’ (AA, 18)

Hortative

Descriptions of TG languages mention the existence of a permissive or exhortative mood^{7,8}. The hortative category expresses hues of a wish, request, desire, deliberation, intention, or obligation through the morpheme *ta*, perhaps cognate with the verb *potar* ‘want’, since a link between such markers and lexical sources meaning ‘desire, want’ are attested (see Bybee et al. 1994).

- (268) a. Tour e Jurupari!
 t-o-ur e Jurupari
HORT-3-come PRCL Jurupari

‘May Jurupari come!’ (D’Abbeville, Histoire, 357)

⟨_{IF} **OPT**⟨_{TNS} *NFUT*⟨_{MOD} *HORT* [**do**’ (Jurupari, [**move.away.from.ref.point**’ (Jurupari))] & INGR [**be-at**’ (Jurupari,?)])⟩⟩

- b. Tafeapipik e moinisemawera ka!
 t-je=∅-apipik e mo-inisem-wer-a ka
HORT-1SG=R₁-mistreat PRCL CAUS-full-PST-REF PRCL

‘May it harm me indeed, my drunkenness!’ (DC, II, 103)

Similarly to the imperative, the hortative is negated with *umē ~ imē*.

⁷For example: Dietrich (1986, 110-111), Rose (2011, 275-276), Seki (1990, 129-130), Villaña (2004, 210), Pease (2007, 53-57).

⁸The use of the term hortative seems more appropriate, as the term “permissive”, frequent in descriptions of Tupi-Guaraní languages. According to the examples found in the TUP corpus, its most frequent usage is in blessings and curses (as optative), although it is also used as non-second person commands (jussive and cohortative). However, as I did not find any clear example of its usage in permissive meaning (see van der Auwera et al. 2013), I regard “hortative” as a more appropriate label.

- (269) Toimoʔaŋimě aβa emonã oikoβaʔe
t-o-i-moʔaŋ-imě aβa-∅ emonã o-iko-βaʔe
HORT-3-R₂-fake-NEG person-REF thus 3-be-NMLZ_{REL}
‘May the Indians who acted this way not pretend.’ (CC, 1, 17)

6.5.1.2 Evidentials

Evidentiality is the grammaticalized marking of information source Aikhenvald (2004). In other words, it is a way of indicating the speaker’s assessment of the evidence for his or her statement.

Tupinambá has a relatively simple system of evidentials compared to other languages of the TG family (see Cabral 2007, 289), marking a three-way distinction between information directly attested by the speaker, information obtained by third-party attestation, and information heard, but not from direct testimony (see Willett 1988).

Information directly attested by the speaker is conveyed by *-rako*, as (270) illustrates:

- (270) a. Ak^wejme rako pira asekijmarajatu
ak^wejme **rako** pira-∅ a-s-ekij-marŋgatu
formerly EV_{FH} fish-REF 1SG-R₂-fish-favorably
‘In the old days, as a matter of fact, I used to fish favourably.’ (AP, 152)
⟨_{IF} DEC ⟨_{EV} ATT **do**’ (1SG, [**fish**’ (1SG)]) ⟩ ⟩
- b. Emonã rako sekow neswi
emonã **rako** s-eko-w ne=∅-swi
thus EV_{FH} R₂-act-NFOC 2SG-R₂₁-POSP
‘(Being) far from you, this is how he acted.’ (Araújo, 74 1686)
- c. Aʔe rako iangajpa
aʔe **rako** i-angajpa
They EV_{FH} R₁-evil
‘They, I know it, are evil.’ (AT, 16)

If the information being conveyed is attested not by the speaker, but by a third party, then *raʔe* is employed:

- (271) a. Oso raʔe
 o-so **raʔe**
 3-go EV_{NFH}
 ‘He went, it is said.’ (VLB, I, 104)
- b. Maria kujāŋatu opuruʔaramo, raʔe,
 Maria kujā-katu-∅ o-puruʔa-ramo **raʔe**
 Maria woman-good-REF CORF-pregnant-TRSL EV_{NFH}
 tekopofji ojmopuru
 t-eko-pofji o-i-mo-puru
 R₁-habit-bad-∅ 3-R₂-CAUS-damn
 ‘Maria, good woman, becoming pregnant, it is said, attempted against the sinns.’ (AP, 184)

Another morpheme indicating that the information has been indirectly obtained by having heard it from a third party is *-je*. As shown by (273a), it is possible to combine two evidential morphemes.⁹

- (272) Guarayo, Tupí-Guaraní, Bolívia
 Oso je rae
 3-go EV EV
 ‘It is said that he went.’ (Hoeller 1932, 216)

Another piece of evidence for two evidentials occurring adjacently is the double source for evidential morphemes in other TG languages (see Cabral 2007)

- (273) a. Emonã je raʔe
 emonã **je** **raʔe**
 thus EV_{NFH} EV_{NFH}
 ‘Thus it happened, it is said.’ (VLB, I, 104)

⁹A similar example, with the same evidentials in (273b) is found in Guarayo (my gloss and translation; the original orthography is maintained).

- b. Emonã je aβa rekow raʔe
 emonã **je** aβa-∅ r-eko-w **raʔe**
 thus EV_{NFH} man-REF R₁-act-NFOC formelry
 ‘It is said that the man formerly acted this way.’ (DC, II, 100)

6.5.1.3 Status

In Tupinambá, the realis status¹⁰ is unmarked, while irrealis status is expressed through an oblique suffix, originally a postposition meaning ‘on the occasion of, because’, which later grammaticalized as a translative case marker before losing its status as such (see Cabral and Rodrigues 2005). It attaches to the nominalized constituent and does not require the relational marker. As status is a clausal operator, it is unusual for it to be indicated on an RP rather than on a part of nucleus, as in TUP.

The irrealis (IRR) is marked on the RP by *-reme* ~ *-me*¹¹ ‘because of, on the occasion of, if, when’, which was originally an oblique suffix or postposition that later grammaticalized into an irrealis marker expressing simultaneity (274a), condition (274b), causality (274c), or temporality (274d).

- (274) a. Ajeʔej nesoreme
 a-jeʔej ne=R₁-so-**reme**
 1SG-speak 2SG=∅-go-**IRR**
 ‘I speak while/when you go (lit. I speak on the occasion of your going).’ (AA (1595), 29v)
 ⟨_{IF} DEC ⟨_{STA} IRR ⟨_{TNS} PRS **do'** (1SG [**speak'** (1SG)]) ∧ **do'** (2SG, [**move.away.from.ref.point'** (2SG)]))⟩⟩⟩
- b. Tupã ipotareʔĩme, najpotari
 Tupã i-potar-eʔĩ-**me** n-a-i-potar-i
 God R₁-want-not-**IRR** NEG-1SG=R₁-want-NEG

¹⁰The term ‘realis’ is controversial in linguistics (see von Prince et al. (2022)). I follow Haspelmath (2010), according to whom comparative concepts cannot be right or wrong, but more or less productive. In this sense, it is not controversial that what I here consider to be a marker of irrealis status is elsewhere considered a marker of subjunctive mood (e.g. Seki 2000, 130) or another grammatical category.

¹¹For the structure of clauses with *reme*, see Section 10.2.2.

‘I do not want it, if/when God does not want it (lit. because of God’s not wanting it I do not want it).’ (D’Abbeville, 351v)

- c. Pedro oso omonóreme

Pedro o-so **o-mo-so-reme**

Pedro 3-go 3_{conf}-CAUS-go-IRR

‘Pedro_i goes because/when/if he_i is sent.’ (FA, 84)

- d. Ojerokipe ase Jesus ?ereme?

o-je-roki=pe ase Jesus ?e-**reme**

3-RFLX-bow=Q we Jesus say-IRR

‘Do we bow when we say Jesus?’ (Araújo, 23)

⟨_{IF} int ⟨_{STA} irr ⟨_{TNS} PRS do’ (ase [duck (asé)]) ∧ do’ (2SG, [move.away.from.re]) (2SG)]⟩⟩⟩⟩

It is common to combine the irrealis marker with the optative/hypothetical marker.

- (275) a. Pedro jawar-a Ø-juká-reme

Pedro jawar-a Ø-juka-**reme**

Pedro jaguar-REF R₂-kill-IRR

‘If/when Pedro killed the jaguar.’ (FA, 155)

- b. femonorememo asómo

ʃe=Ø-mo-so-**reme-mo** a-so-**mo**

1SG=R₁-CAUS-go-IRR-OPT 1SG-go-OPT

‘If I were sent, I would go. (if they sent me, I’d go)’ (AA, 25)

The morpheme *nipo* is an alethic modality marker. It indicates the speaker’s estimation of the probability of the proposition expressed by his utterance. In this case the speaker considers it possible:

- (276) a. Oso nipo?

o-so **nipo**

3-go **DUB**

‘Is he perhaps going? / would he go?’ (VLB, I, 82)

- b. Ojepe-ĩomβe, nipo, iangajpaβ amōme e
 ojepe-ĩomβe **nipo** i-angajpaβ amōme e
 one-the.other DUB R₁-evil some.times PRCL
 ‘Maybe, one or the other was evil sometimes.’ (AT, 38)

Another category is the frustrative, which encodes epistemic modality indicating the ‘nonrealization of some expected outcome implied by the proposition expressed in the marked clause’ (Overall 2017) – because the category frustrative always implies two propositions, even though the second proposition often remains implicit. The frustrative is a common category of the verb in Amazonian languages, especially TG (see Aikhenvald 2012, 185 and Dietrich 2006).

In Tupinambá, this unrealized expectation is encoded by the particle *βiã*¹².

- (277) a. Aso βiã
 a-so **βiã**
 1SG-go **FRUST**
 ‘I went (in vain).’ (AA, 21v)
- b. Asopotar iβakipe, eʔi, βiã
 a-so-potar iβak-i-pe e-ʔi **βiã**
 1SG-go-want sky-LOC 3-say **FRUST**
 ‘I want to go to heaven, they say, in vain.’ (Araújo, 112)

Two other particles, *jõte*¹³, *jě*, and *(ʔ)ĩ*, which is also used for the diminutive (see Section 7.3.2), indicate other nuances of the frustrative modality: concessive or lusive, indicating that the goal of the action was not accomplished.

- (278) a. Aso jõe
 a-so **jõe**
 1SG-go **FRUST**

¹²For some cognates of TUP *βiã* in other TG languages and their meaning, see Jensen (1998a, 538-539).

¹³The lusive marker *jõe* probably should be analyzed as the combination of the particle/adverb *jõ* ‘only,just’ and the focal particle *te*.

- c. Nerejeʔēmotaripe nerapisara supeʔ
n-ere-jeʔeŋ-potar-i=pe ne=r-apisar-a supe
 NEG-2-SG-speak-want-NEG=Q 2SG=R₁-colleague-REF POSP
 ‘Didn’t you want to talk to your colleague?’ (Araújo, 102)
- d. Nereimojirōj Tupã nejoupe
n-ere-i-mo-jirō-i Tupã ne=∅-joupe
 NEG-2-R₁-CAUS-forgive-NEG God 2SG=R₁-RECP_{POSP}
 ‘You did not make God forgive you.’ (AC, 97)

Examples of *na ... i* negating nominal predicates are given in (281):

- (281) a. Nitiβi jeaβaetep^wera
n-i-tiβ-i je=∅-aβaete-p^wer-a
 NEG-R₂-existence-NEG 1SG=R₁-courage-PST-REF
 ‘I do not have courage. (there isn’t my old courage)’ (Teatro, 50)
- b. Naferāj
na-fe=r-āj-i
 NEG-1SG=R₁-tooth-NEG
 ‘I do not have teeth.’ (VLB, I, 97)
- c. Naferoriβi
na-fe=r-oriβ-i
 NEG-1SG=R₁-happy-NEG
 ‘I am not happy.’ (Anch., Arte, 34v)

In the future tense, negation is also discontinuous, but with the addition of *swe/so* (*na ... i swe / so*)¹⁴. Examples of future negation are given in (282a).

- (282) a. Nasawsuβejēswe Ajāṅane
n-a-s-awsuβ-βejē-i-swe Ajāṅa=ne
 NEG-1SG=R₁-love-again-NEG-NEG Devil=FUT
 ‘I shall not love the Devil again.’ (Araújo, 86)

¹⁴Most probably, *swe* and *swe / so* are the result of a palatalization due to a preceding *i-* (see example 30).

- b. Noromomeʔuisóne
n-oro-momeʔu-i-so=ne
 NEG-1PL.EXCL-denounce-NEG-NEG=FUT
 ‘I will not denounce you.’ (Teatro, 34)

There is a special form for negating the future, as in other Tupían languages (see Dietrich 2017b). This form is similar to the standard core negation (see Section 6.5.2.3) in that it uses a discontinuous morpheme *n(a)-...-i* with the addition of the suffix *swe* – of unknown etymology – which can optionally be followed by the future clitic marker *=ne* (*swe=ne*):

- (283) a. Nimaʔenwariswéne
n-i-maʔenwar-i **swe=ne**
 NEG-R₁-remember-NEG PRCL=FUT
 ‘They will not remember.’ (FA, 40)
- b. Naferekopofij *swe*
na-fe=r-eko-pofi-i **swe**
 NEG-1SG=R₁-life-evil-NEG NEG
 ‘I will not have sins.’ (AC, 106)
- c. Najukaj swéne
na-a-i-juka-i **swe=ne**
 NEG 1SG=R₁-kill-NEG NEG=FUT
 ‘I do not refrain from killing him (I will not not kill him).’ (FA, 34)
- d. Ase ʔaʔa jō nopaβi swéne, awjeramaje
 ase ʔaʔa-a jō na-opaβ-i **swe=ne** awjeramaje
 we soul-REF ADV NEG-terminate-NEG **NEG=FUT** eternally
 omanōβaʔerameʔima sekóreme
 o-manō-βaʔe-ram-eʔim-a s-eko-reme
 3-die-NMLZ-FUT-PRIV-REF R₂-be-POSP
 ‘Only our soul will not end, because it is what never dies.’ (Bettendorff, Compêndio, 58)

Double negation is mentioned by Anchieta (1595) and Figueira (1687) but

rarely seen in the texts. Double negation – privative (see Section 8.2.0.2) + negation – has affirmative meaning as in (284)¹⁵:

- (284) a. Najukaeʔimi
na-a-i-juka-**eʔim-i**
NEG-1SG-R₂-kill-**PRIV-NEG**
‘I kill him.’ (FA, 34)
- b. Naipotareʔimi swéne
na-a-i-potar-**eʔim-i** swe=ne
NEG-1SG-R₂-want-**PRIV-NEG** NEG=FUT
‘I will not not kill him.’ (Anchieta, 34v)

There is also an instance of the privative with the non-predicative negator *ruã* as in (285):

- (285) Nasekasaβa k^waβeʔima ruã
na-s-eka-saβ-a k^waβ-**eʔim-a** **ruã**
NEG-R₂-search-NMLZ-REF know-**PRIV-GER** NEG
‘Not ignoring (not knowing) what they were looking for (lit. knowing the thing searched for).’ (Araújo, 54)

6.5.1.5 Tense

Tense is a category which expresses a temporal relationship between the time of the described event and some reference time (S) which, in the unmarked case, is the moment of speech. In the simplest case, tense indicates the temporal relationship between the time of the event (E) and the time of the utterance (S) describing the event (Hornstein 1993; Comrie et al. 1985). In TUP, as in most TG languages, the time of the utterance mostly does not coincide with the time of the event, so unmarked predicates express the past. Special constructions are used for the present tense.

¹⁵These examples are translated in the original sources as ‘not refrain from (doing) X’. This nuance is still seen in double negation construction in Mbyá Guaraní (Dooley 2015, 88), where the construction translates to ‘not refrain from doing X/not refuse (to do) X.’

Tupinambá only has absolute tenses, and a future and non-future distinction. We can formally distinguish this binary opposition in terms of ‘*E* (moment of event) relative to *S* (moment of speech)’ (see Comrie et al. 1985, chap. 6)¹⁶. Tupinambá is a non-future language (‘*E* not-after *S*’), meaning that it has default [– future] (see Müller 2013, 38). Thus, the verb form not marked for tense can be either present or past (see Anchieta 1595, 21v), with the past being far more common.

Even though TUP is a non-future language, there is a particle, *jã*¹⁷, that indicates present tense, excluding the past-tense reading.

- (286) Asójã
 a-so-jã
 1SG-go-NPST
 ‘I go.’ (AA, 21v)

Future

The only marked tense in Tupinambá is the future, defined as E after S. It is always marked by the clitic =*ne*¹⁸, which tends to be placed at the end of the clause¹⁹ unless there is a sentence-final particle, as can be seen from the examples in (287).

- (287) a. Asóne
 a-so=**ne**
 1SG-go=**FUT**
 ‘I will go.’ (AA, 22)
 ⟨_{IF} DEC ⟨_{STA} IR ⟨_{TNS} FUT **do**’ (1SG [**move.away.from.ref.point**’ (1SG)])⟩⟩⟩

¹⁶For a similar analysis using three parameters but somewhat different terminology, see: Givón 2001; Bhat 1999.

¹⁷Anchieta 1595, 21v says that *a*, *nja*, and *iko* have the same meaning as *jã*. He provides examples like (286) with all these forms.

¹⁸I do not agree with the terminology in Rodrigues (1953, 129), which refers to the clitic =*ne* as intentional, because many occurrences of it exclude an intentional reading. Additionally, since intention relates to a participant, it would be a core and not a clausal operator.

¹⁹See note 5.

- b. Aso koríne
 a-so kori=**ne**
 1SG-go today_{fut}=FUT
 ‘I will go today.’ (AA, 22)
- c. Aso kori okipe neruriréne
 a-so kori oki=pe ne=r-ur-ríre=**ne**
 1SG-go today_{fut} house-LOC 2SG=R₁-come-after=FUT
 ‘I will go home today after you come.’ (AA, 22)
- d. ferejtik koríne mā!
 fe=r-ejtik kori=**ne** mā
 1SG=R₁-defeat today=FUT PRCL
 ‘Oh, I will be defeated today (lit. they will defeat me today).’ (AT, 28)
- e. Aʔepe miawsuβa nosap^lariswe
 aʔe=pe miawsuβ-a n-o-s-ap^lar-i-swe
 ?=Q slave-REF NEG-3-R₁-obey-NEG-NEG.FUT
 ojara jeʔeɲáne?
 o-jara-∅ jeʔeɲ-a=**ne**
 COREF-lord-REF word-REF=FUT
 ‘Won’t the slave obey the words of his own master?’ (Araújo, 69)

Figure 6.10 shows the representation of 287:

It also often appears attached to the predicate (288).

- (288) a. fepinaporangete topinaitikine
 fe=∅-pina-poraɲ-ete t-o-pina-itiki=**ne**
 1SG=R₁-fishhook-beauty-INTS HORT-3-fishhook-throw=FUT
 enéβo
 ene-βo
 2SG-DAT
 ‘May my very beautiful fish hook fish for you.’ (Poemas, 152)

ive aspect. It is similar in meaning to the so-called imperfect (conflated tense and aspect) of Portuguese or Spanish (see Comrie et al. 1985, 6-7). *Raka?e* is an adverb, not a tense or aspect marker, and is similar to English infinitives preceded by *used to* in that it expresses something that cannot be expressed otherwise in TUP. Its use in the texts clearly translates into the Portuguese imperfect tense.

- (290) a. Ise **raka?e**
 I **formerly**
 ‘It was I / It used to be me.’ (VLB, I, 121)
- b. Orojo?u raka?e
 oro-jo-?u **raka?e**
 1PL.EXCL-RECP-eat **formerly**
 ‘We used to eat each other / we were once eating each other.’
 (D’Abbeville, Histoire, 341v)

6.5.2 Core-level operators

Core operators modify the relation between a core argument, normally the actor, and the nucleus. This is especially true of core directionals and modality.

Modality operators in RRG refer to the deontic sense of modal verbs. This category includes such things as strong obligation (must or have to), ability (can or be able to), permission (may) and weak obligation (ought or should) (Van Valin Jr and LaPolla 1997, 41). Modality thus concerns the relationship between the referent of the subject RP and the action.

6.5.2.1 Modality

Ability, permission, and obligation are expressed in different ways (see Bybee et al. 1994, 177). Ability and permission are expressed through a lexical verb, *(e)?ikatu* ‘be able, can, be allowed’²⁰, which requires a complement.

²⁰Bybee et al. (1994, 181) notes that this type of modality is rarely expressed by inflectional affixes.

- (291) a. Teʔikatu nek^wapa ʃeruβa Tupinamβa!
 t-eʔikatu ne=k^waβ-a ʃe=r-uβ-a Tupinamβa!
 HORT-be.able 2SG=R₁-know 1SG=R₁-father-REF Tupinambá
 ‘May my Tupinambá father (get to) know you!’ (Poemas, 114)
 ⟨ IF DEC ⟨ MOD PER **know**’(uβa, 2SG) ⟩ ⟩
- b. Aʔekatu sep^jaka
 a-ʔekatu s-ep^jak-a
 1SG-be.able R₂-see-GER
 ‘I can / am able to see him.’ (AA, 56)
 ⟨ IF DEC ⟨ MOD PER/AB **see**’(1SG, 3) ⟩ ⟩

Obligation is expressed through the particle *mone*²¹. *Mone* expresses both nuances of obligation, i.e., strong obligation, here translated as ‘must’, and weak obligation, which is translated as ‘should’ (see Bybee et al. 1994, 177).

- (292) a. Kori mone aso
 kori **mone** a-so
 today MOD 1SG-go
 ‘I should go today.’ (AA, 25)
 ⟨ IF DEC ⟨ MOD OBL **do**’ 1GO [(**move.away.from.ref.point**’ (1GO)] ⟩ ⟩
- b. Ahě rajē temonemo!
 ahě rajē te-**mone**-mo
 that.one first FOC-MOD-IRR
 ‘He should be first (not someone else)!’ (VLB, II, 64)

6.5.2.2 Hortative modality

Hortative modality expresses a wish or an allowance and it is formed using the prefix *ta*, which combines with active (293) or stative (294) predicates. It also functions as an hortative marker, as in (293b).

²¹This particle is poorly attested. It is often associated with the irrealis marker *mo*. *Mone* is attested in the following sources: Araújo (1618b, 156v, 165v), Anchieta (1595, 25), Anonymus (1952b, 53, 59, 64).

- (293) a. Tomanõ
t-o-manõ
HORT-3-die
 ‘May he die.’ (Araújo, 56v)
- b. Tajajuka jemena
t-ja-∅-juka je=∅-men-a
HORT-1PL.INCL-R₁-kill 1SG=R₁-husband-REF
 ‘Let us/may we kill my husband.’ (Araújo, 279)
- (294) a. Tafejuka Pedro
ta-fe=∅-juka Pedro
HORT-1SG=R₁-kill Pedro
 ‘May Pedro kill me.’ (FA, 152)
- b. Tafema?enwar
ta-fe=∅-ma?enwar
HORT-1SG=R₁-remember(ance)
 ‘May I remember.’ (AF, 44)

The hortative mood also occurs in the future and is more frequently attested with the first person singular (295) and the first person plural inclusive (see Anchieta 1595, 23). In contrast, it is rarely attested with second or third (296) persons. Often, the combination of the hortative marker with the future marker expresses epistemic future (see Giannakidou and Mari 2018).

- (295) a. Eru pira ta?une
e-ero-ur pira-∅ **t-a-?u=ne**
2.IMP.SG-SCAU-bring fish-REF HORT-1SG-ingest-FUT
 ‘Bring fish, that I may eat it / I shall eat it.’ (Anch., Arte, 23)
- b. Tasóne, wi! Taka?une!
t-a-so=ne wi! **t-a-ka(wĩ)-?u=ne**
HORT-1SG-go=FUT INTJ HORT-1SG-beer-ingest=FUT
 ‘I shall go, I shall drink beer!’ (AT, 12)

- (296) a. Nemajanamo tojkóne!
 ne=∅-maja-namo **t-oiko=ne**
 2SG=R₁-spy-TRSL **HORT-3-be=FUT**
 ‘He might be your spy/spy on you!’ (AT, 34)
- b. Tasepi nemonawera
ta-s-epi ne=∅-mona-wer-a
HORT-R₁-payment 2SG=R₁-theft-PST-REF
 ‘May there be a price for your theft.’ (AT, 48)
- c. Kori e toromodone
 kori e **t-oro-mo-so=ne**
 today PRCL **HORT-1PL.EXCL-CAUS-go-FUT**
 ‘Today, I may turn you away.’ (AT, 34)

6.5.2.3 Core negation

Core-level negation in TUP negates a core argument. The core negator morpheme is *n-*, which precedes the term being negated and the focal *ruã* which follows it. This type is common to other TG languages (Dietrich 2017b, 22-23). Examples are given in 297. The element in the scope of *n ... ruã* is always focal.

- (297) a. Naβare ruã ise
n-aβare-∅ **ruã** ise
 NEG-priest-REF NEG I
 ‘I am no priest.’ (AA, 46v)
- b. Naseruβa ruãtepe ase rete ojmojaŋ?
na-ase-r-uβ-a **ruã-te=pe** ase r-ete o-i-mojaŋ
 NEG-our-R₁-father-REF NEG-FOC=Q we R₁-body 3-R₂-make
 ‘Wasn’t it our father who made our body?’ (Araújo, 25)
- c. Naferuβa supe ruã ajmeʔeŋ
na-fe=r-uβ-a supe **ruã** a-i-meʔeŋ
 NEG-1SG=R₁-father-REF to NEG 1SG-R₂-give
 ‘Not to my father I gave it.’ (AA, 47v)

- (298) a. Nerejeʔēmotaripe nerapisara supeʔ
 n-ere-jeʔeŋ-potar-i=pe ne=r-apisar-a supe
 NEG-2-SG-speak-want-NEG=Q 2SG=R₁-colleague-REF POSP
 ‘Didn’t you want to talk to your colleague?’ (Araújo, 102)
- b. Nereimójirōj Tupã nejoupe
 n-ere-i-mo-jirō-i Tupã ne=∅-joupe
 NEG-2-R₁-CAUS-forgiveness-NEG God 2SG=R₁-RECP_{POSP}
 ‘You did not make God forgive you.’ (Araújo, 97)
- c. Naferoriβi
 na-fe=r-oriβ-i
 NEG-1SG=R₁-happy-NEG
 ‘I am not happy.’ (AA, 34v)

6.5.3 Nuclear-level operators

Nuclear-level operators modify, as the name suggests, only the nucleus. Elements in other layers are not affected by them.

6.5.3.1 Aspect

Aspect, in spite of not relating to tense, i.e., the relation between event and time of utterance, concerns the internal temporal structure of the event itself (see Comrie 1976). Aspect is manifested in lexical semantics, since predicates have different *Aktionsart*, as well as in grammatical semantics through various grammatical constructions (see Croft 2012, 4). Another notion that is related to verbal aspect is that of causal or force dynamic structure (Croft 2012, 4). They are related in the sense that both are relevant to event structure. Aspectual markers in TUP always follow the predicate. Otherwise, they are adverbs.

The completive ‘aspect’ is marked by the adverb *pa* (from *paβ*), meaning ‘all, total(ly), complete(ly)’, and indicates that an action has been fully performed. As an adverb, it cannot be an operator because operators are closed class grammatical items.

In the examples in (299), it should be clear that the scope of the adverb is over the predicate only; the action is perceived as completed.

- (299) a. Kunumi mokongapa
 kunumi- \emptyset mokoŋ-a-**pa**
 boy-REF swallow-GER-**completely**
 ‘Swallowing the boy completely.’ (Poemas, 166)
- b. Taʔupa Jakarewasu pepira
 t-a-ʔu-**pa** Jakarewasu pepir-a
 HORT-1SG-ingest-**completely** Jakarewasu feast-REF
 ‘I shall eat up Jakarewasu’s feast.’ (AT, 64)
- c. Ereroirōpape seko?
 ere-eroirō-**pa**=pe s-eko
 2SG-hate-**completely**=Q R₁-deed
 ‘Do you completely hate his deeds?’ (Araújo, 114v)

Predicates may be modified by the lexical root *aiβ* ‘bad’, which is in opposition to the meaning expressed by *pa*, as an incomplete action marker, indicating that the action is partial or incomplete, as in (300):

- (300) a. Asenuβaiβ nejeʔeja
 a-s-enuβ-aiβ ne= \emptyset -jeʔeja
 1SG-R₂-listen-**bad** 2SG=R₁-speech-REF
 ‘I hear your words (but not all).’ (VLB, I, 119)
- b. Erejukaípe menareʔima imoposi
 ere-i-juka-aiβ=pe men-sar-eʔim-a i-mo-pofɨ
 2SG-R₂-kill-**bad**=Q husband-NMLZ_{AG}-PRIV-REF R₂-CAUS-evil
 janone kojpo moposipota?
 janone kojpo i-mo-posipota
 before or R₂-CAUS-evil-want.GER
 ‘Did you hit (without killing) a maiden before doing her evil or wanting to cause her evil?’ (Araújo, 103v)

- c. Amanōaiβ amiasi swi
 a-manō-aiβ ami-asi ∅-swi
 1SG-die-INCMP abdomen-pain R₁-from
 ‘I almost die of hunger.’ (VLB, II, 73)

For a particle signaling the imperfective aspect and past tense, see (6.5.1.5).

Lexical roots that combine with Set II indexes may combine with indexes from Set I (possessor) indexes in order to express the habitual or frequentative aspect, indicating that the actor frequently performs the action (see Anchieta 1595, 51).

- (301) a. Aʔitaβ
 a-ʔitaβ
 1SG-swim
 ‘I swim.’ (AA, 51)
- b. jeʔitaβ
 je=∅-ʔitaβ
 1SG=R₁-swim
 ‘I often swim / I am a swimmer / I can swim.’ (AA, 51)

Some nuances regarding the realization of the verbal process, e.g., repetition and customary activity (Sapir 1921, IV:24), are expressed through reduplication (see Inkelas and Downing 2015; Downing and Inkelas 2015), which is a common feature of Tupían languages (Rose 2005; Dietrich 2014). Tupinambá has two different types of reduplication, and they differ according to the number of syllables reduplicated. There are also verbs that only exist in reduplicated forms because the iterative aspect is inherent to the nature of the process they describe, such as: *papar* ‘count, numerate’, *βεβε* ‘fly’, *pupur* ‘boil’, *bubur* ‘gush’, etc.²²

Monosyllabic reduplication expresses event-internal repetition, as shown in (302). When reduplicated, transitive verbs indicate that the action is performed on different objects one at a time (302a). Meanwhile, in intransitive

²²These verbs seem to be onomatopoeic, imitating, for example, the noise of boiling water or the flapping of wings.

verbs, reduplication indicates that the action is performed by the actors successively or simultaneously (302b). This is the iterative aspect.

- (302) a. Aimokõkõŋ
 a-i-**mokõŋ-kõŋ**
 1SG-R₂-**swallow.RED**
 ‘I swallow one after the other.’ (VLB, I, 116)
- b. Osisik
 o-**sik-sik**
 3-**arrive.RED**
 ‘(S)He arrives again and again.’ (AA, 53v)

Disyllabic reduplication indicates iterative or frequentative meaning, i.e., the repetition of the verbal process (303a). If the verbal root is monosyllabic, it is possible to reduplicate two syllables by including the person index, as in (303b).

- (303) a. Wijemojewawaka
 wi-je-mo-**jewak-jewak-a**
 1SG_{CORF}-RFLX-CAUS-embellish-**RED**-GER
 ‘I keep embellishing myself.’ (Poemas, 110)
- b. Nesunesupa
 ne=∅-suβ-ne=∅-su-pa
 2SG=R₁-visit-2SG=R₁-visit-GER
 ‘Visiting you again and again.’ (Teatro, 84)

6.5.3.2 Nuclear negation

Nuclear-level negation is characterized by the privative suffix *-eʔim*²³. Its privative meaning is clear from examples such as those in (304).

²³For the cognates of the privative *-eʔim* in other Tupí-Guaraní languages and their use, see Dietrich (2017b).

- (304) a. Sieʔima
 si-**eʔim**-a
 mother-**PRIV**-REF
 ‘Orphan / motherless.’ (VLB, II, 59)
- b. Kereʔima
 ker-**eʔim**-a
 sleep-**PRIV**-REF
 ‘Sleeplessness.’ (see Teatro, 34; AC, 53)

In (305a) the privative is used as an argument, contrasting with its function in (305b), where the nominal with the privative suffix is the predicate:

- (305) a. Nojpotaripe Tupã jereʔõeʔima
 Na-o-i-potar=pe Tupã je=r-eʔõ-**eʔim**-a
 NEG-3-R₂-want=Q God 1SG=R₁-death-**PRIV**-REF
 jeretãme wisóβo?
 je=r-etam-pe wi-so-βo
 1SG=R₁-country-LOC 1SG_{CORF}-go-GER
 ‘Doesn’t God want that I do not die (my death-less-ness) to go to my country?’ (D’Abbeville, Histoire, 351v)
- b. Marãpe peruβisaβetaeʔim?
 marã=pe pe=r-uβisaβ-eta-**eʔim**
 why-Q 2PL=R₁-chief-many-**PRIV**
 ‘Why don’t you have many rulers?’ (Léry, 362)
- c. Serekokatupireʔimetémo
 s-ereko-katu-pir-**eʔim**-ete-mo
 R₂-treat-GOOD-DEV.PASS-**PRIV**-INTS-IRR
 ‘He would be not very well treated.’ (Léry, 353)

Although it is clearly more frequently used with nominals, the privative *-eʔim* is also found with verbal predicates:

- (306) Ajukaeʔim
 a-i-juka-**eʔim**
 1SG-R₂-kill-**PRIV**

'I don't kill (it/him/her/them).' (AA, 20)

Chapter 7

Lexical categories

Chapter 4 has shown that word classes are defined according to the combination of semantic categories and the speech act function they perform, and that there is a morphological difference between possessive and non-possessive predication. This section further discusses word classes, lexical categories, and some predicate types.

7.1 Predicate semantic classes

Aktionsart predicate classes were presented in Section 3.3. These, as expected, account for TUP-specific morphosyntactic generalizations or distinctions. *Aktionsart* predicates are described in terms of their LS, which includes the minimum number of semantic arguments that each predicate may require. This section discusses the representation of these logical structures for predicates in Tupinambá.

7.1.1 State predicates

States describe situations that do not change over time¹. They are also atelic, i.e., they do not have an endpoint.

The single argument (x) is an entity being identified by the predicate, in the case of identificational state as in (307) with its representation in Figure (7.1), but an entity bearing the specific individual-level property denoted in the root in the case of an attribute predicate, as in (308).

- (307) Tupã raʔira iko aβa
 Tupã-∅ r-aʔir-a iko aβa-∅
 God-REF R₁-son-REF DEM man-REF
 ‘This man is the son of God.’ (Ar, 64 modified)
equate’ (Tupã taʔira,iko aβa)

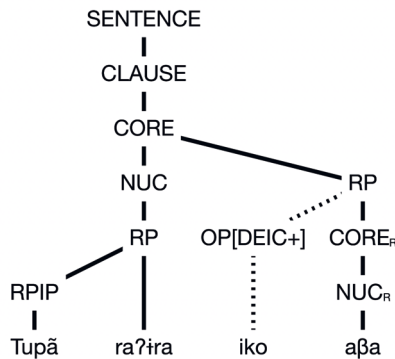


Figure 7.1: Object word in predicate function

- (308) ʃeroβi
 ʃe=r-oβi
 1SG=R₁-blue(ness)

¹Other types of states are discussed in Vendler (1967) which are not relevant to the morphosyntax of TUP (see Croft 2012, 77-83).

'I am blue.' (VLB, I, 49)
 ([**have.as.part'**(1SG,blueness)]

Physical, emotional, or mental experiences are temporary, stage-level states that have come about for an EXPERIENCER argument. State experiences do not denote cognitive attention or direction. The experiencer is not a controller of the state of affairs. Experiential states may be used as stative modifiers in a reference phrase, and their logical structure is **feel'** (x, [**root'**]). The latter will be used in order to differentiate experience states from other stative predicates. Examples are given in (309).

(309) Undergoer = experiencer

- a. \int eputupa β
 \int e= \emptyset -putupa β
 1SG=R₁-amaze
 'I am amazed.' (Léry, 353)
feel' (I, [**amazement'**])
- b. \int ero \int i
 \int e=r-o \int i
 1SG=R₁-cold
 'I am/feel cold.' (Léry, 367)
feel' (I, [**cold'**])

M-transitive state verbs have actors as cognizers, emoters, judges, and other semantic roles as actors, as in (310):

- (310) a. Actor = cognizer, undergoer = content
 Naik^wa β i a \int e a β a
 na-a-i-k^wa β -i a \int e a β a
 NEG-1SG=R₁-know-NEG DEM man- \emptyset
 'I do not know this man.' (Araújo, 57)
know'(1SG, a \int e a β a)
- b. Actor = emoter, undergoer = target

Asausu kujākaraiβa
 a-s-awsuβ kujã-karaiβ-a
 1-SG=R₁-love woman-non.indian-REF
 ‘I love a white woman.’ (D’Evreux, Viagem, 252)
love’(1SG, kujākaraiβa)

- c. Actor = judger, undergoer = judgement
 Naimoʔaŋi neso
 na-a-i-mo-ʔaŋ-i ne=∅-so-∅
 NEG-1SG-R₂-CAUS-idea-NEG 2SG=R₁-go-REF
 ‘I do not understand you going.’ (VLB, II, 110)

7.1.2 Activity predicates

Activity predicates are dynamic and temporally unbounded. Their logical structure is of the type **do’** (x, [**predicate’** (x) or (x, y)]).

- (311) a. Actor = mover
 Ereso
 ere-so
 2SG-go
 ‘You go/went.’
do’ (2SG, [**move.away.from.ref.point’** (2SG)])
- b. Actor = mover
 Ajākatune
 a-jān-katu=ne
 1SG-run-much=FUT
 ‘I will run a lot.’ (AT, 25)
do’ (1SG, [**run’** (I)])
- c. Actor = light emitter
 K^warasioβeraβ
 k^warasi-∅ o-βeraβ
 sun-REF3-shine
 ‘The sun certainly shines.’ (see Poemas, 142)
do’ (kwarasi, [**shine’** (kwarasi)])

- (312) Actor = user, undergoer=implement
 Ejporu nejemoʔeawera
 e-i-poru ne=∅-je-mo-ʔe-aβ-wer-a
 2SG.IMP-R₂-use 2SG=R₁-RFLX-CAUS-say-NMLZ-PST-REF
 ‘Use what you learned.’ (VLB, I, 131)
do'(e, [**use'**(e, nejemoʔeawera)])

7.1.3 Achievement predicates

Achievement verbs denote a punctual change of state which achieves an end point. They can be achievement INGR **predicate'**(x) or (x, y) or causative achievement [**do'**(x, ∅)] CAUSE [INGR **predicate'**(y)]. Examples are given below. Verbs such as *βok* ‘blast off’, *ʔar* ‘fall’, and *sok* ‘break’ only have an achievement sense and are M-intransitive.

- (313) a. Oβok nejiʔã
 o-βok ne=∅-jiʔã-REF
 3-blast.off 2SG=R₁-heart-REF
 ‘Your heart blasted off.’ (AP, 120)
 INGR **blast.off'**(3[nejiʔã])
- b. Oʔar iβipe
 o-ʔar iβi-∅-pe
 3-fall meat-REF earth-REF-LOC
 ‘It fell on the floor.’ (VLB, I, 72)
 INGR **fall'**(3)
- c. Osok
 o-sok
 3-break
 ‘It breaks.’ (AA, 53v) INGR **break'**(3)

Some verbs have an achievement sense and a causative achievement counterpart:

- (314) a. Ita ojeka
 ita-∅ o-je-ka
 stone-REF 3-RFLX-break
 ‘The stones break.’ (see Araújo 1618b, 64)
 INGR **break'** (3[ita])
- b. Erejoka
 ere-jo-ka
 2SG-R₂-break
 ‘You break them.’ (see Anchieta 2006, 48)
 [**do'** (2SG, ∅)] CAUSE [INGR **broken'** jo]

7.1.4 Semelfactive

- (315) Semelfactive verb: actor PSA
 Ajemoesaβik
 a-je-mo-esa-βik
 1SG-RFLX-CAUS-eye-touch
 ‘I blinked.’ (VLB, I, 79)
 SEML **do'** (1SG, [**blink'** (I)])

7.1.5 Accomplishments

Accomplishments are processes with endpoints.

- (316) a. Accomplishment verb: undergoer PSA
 Aβa omanõ
 aβa-∅ o-manõ
 man-REF 3-die
 ‘A man died.’ (Fig., Arte, 69)
 BECOME **dead'** (3sg [man])
- b. Active accomplishment
 Jasi maʔe oʔu
 jasi-∅ maʔe-∅ o-ʔu
 moon-REF thing-REF 3-eat

‘A thing ate the moon.’ (VLB, I, 108)

do' (3[maʔe], [**eat'** (3[maʔe], jasi)]) \wedge PROC **being.consumed'**
(jasi) \wedge FIN **consumed'** (3[jasi])

7.2 Morphological predicate types

Lexical roots functioning as predicates can be divided into two types: nominal existential predicates that combine with Set I indexes and ‘verbal’ predicates that combine with indexes from Set II. Both types share morphology associated with some grammatical categories, as the examples below attest.

Both may receive the future tense marker:

- (317) a. Ajukane
a-i-juka=**ne**
1SG-R₂-kill=**FUT**
‘I will kill him.’ (FA, 7)
- b. Neʔākitijokine
ne=∅-ʔāṅ-kitijok=**ne**
2SG=R₁-soul-clean.rubbing=**FUT**
‘You will have a clean soul.’ (DC, II, 113)

Both receive the same negation:

- (318) a. Namaʔēi
n-a-maʔē-i
NEG-1SG-see-NEG
‘I do not look at it.’ (VLB, I, 70)
- b. Naferoriβi
n-fe=r-oriβ-i
NEG-1SG=R₁-happy-NEG
‘I am not happy.’ (AA, 34v)

Both types of predicates combine with the irrealis marker:

- (319) a. Asómo
 a-so-**mo**
 1SG-go-**IRR**
 ‘If I went.’ (FA, 142)
- b. jesuʔumo mariwi
 je=∅-suʔu-**mo** mariwi
 1SG=R₁-bite-**IRR** mariwi
 ‘A bug would sting me.’ (Teatro, 64)
- (320) a. Ojeʔu
 o-je-ʔu
 3-RFLX-eat
 ‘He eats himself.’ (FA, 142)
- b. fejejok
 je=∅-je-jok
 1SG=R₁-RFLX-sob
 ‘I sob.’ (Teatro, 64)

Other types of predicates, which are associated with complex sentences, are discussed in Chapter 10.

7.3 Nominal Categories

This section deals with nominal categories in Tupinambá. These are: tense/aspect, degree (diminutive, augmentative, and intensity), nominal number, and alienability.

7.3.1 Nominal tense

The nominal tense in Tupinambá is indicated by the suffixes *-p^wer* ~ *-wer* ‘nominal past’ or *-ram* ~ *-wam* ‘nominal future’. Cognates of these morphemes are common in Tupí-Guaraní languages and have been treated as nominal tense in other language descriptions (Seki 2000, 54, Neiva Praça 2007,

19, Harrison and Harrison 2013, 20,98), and also as nominal aspect markers, in spite of terminological differences (Rose 2011, 236). This distinction, however, is irrelevant, as observed by Bertinetto (2020), because any tense by definition conveys a range of values in each TAM-component, besides the temporal reference.²

Examples of past tense are given in (321) and examples of future tense are given in (322):

- (321) a. Ok^wera
 ok-**p^wer**-a
 house-**PST**-REF
 ‘A degraded version of a house / former house.’ (AA, 33v)
- b. Nitiβi jeaβaetep^wera
 n-i-tiβ-i je=∅-aβaete-**p^wer**-a
 NEG-R₂-exist-NEG 1SG=R₁-courage-**PST**-REF
 ‘My old courage is not there (there isn’t my old courage).’ (Teatro, 50)
- (322) a. Tupã sirama ∅-ri
 Tupã ∅-si-**ram**-a ∅-ri
 God R₁-mother-**FUT**-REF R₁-POSP
 imojañipira
 i-mojañ-i-pir-a
 R₂-make-EPNT-NMLZ_{pass}-REF
 ‘She was made to become the (future) mother of God.’ (Poemas, 88)

²Although this idea is not developed here, I do consider the possibility of these markers being modifiers of the qualia (see Section 3.3.1) In this case, the past form of a house would indicate that one or more qualia has/have been modified, constitutive role, formal role, telic role, or agentive role. A house without a roof for example, would have its formal role altered, because a part of it is missing as well as its telic role, since it could not fulfill its goal of sheltering people. A house in construction, for example, would have some of its qualia different from a finished house, in the sense that its parts would not be the same, the same for its telic role, since people could not live inside it yet. these are basic examples, but it seems reasonable, I think, to pursue this possibility further.

- b. Ajune ise, peremi?urama!
 a-jur=ne ise pe=r-emi-?u-**ram**-a
 1SG-come=FUT I 2PL=R₁-RES-eat-FUT-REF
 ‘I am coming, I, your future meal!’ (Staden, 67)

The nominal past, as exemplified in (321a), denotes not only ‘a degraded version of a house’, e.g., a house without a roof, or the ruins of a house. It also denotes the time at which the noun property or the possessive relation holds, so that *okwera* could also mean ‘the former house (of someone)’. The same applies to the future tense, as in (323), which may indicate a house that already exists, but will belong to someone who will live there, or it may indicate a house that is being constructed but is not yet finished.

- (323) Okwama
 ok-ram-a
 house-FUT_N-REF
 ‘Future house.’ (cf. VLB, I, 108)

An interesting example is (324), which was the Jesuit choice for referring to the communion wafer after it becomes the body of Christ.

- (324) Miapep^wera
 miape-p^wer-a
 bread-PST_N-REF
 ‘What was bread (the communion wafer).’ (cf. AC, 87)

The representation of (325) is given in Figure 7.2, showing the nominal tense operator.

- (325) ferok^wama aimojaŋ
 fe=r-ok-**wam**-a a-i-mojaŋ
 1SG=R₁-house-FUT_N-REF 1SG-R₂-make
 ‘I am making my future house.’ (VLB, I, 108)

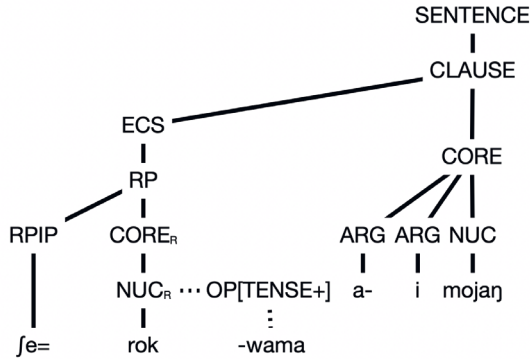


Figure 7.2: The nominal tense RP operator

- (326) Ajune ise, peremi?urama!
 a-jur=ne ise, pe=r-emi-?u-**ram**-a
 1SG-come=FUT I 2SG=R₁-RES-eat-FUT_N-REF
 ‘I am coming, I, your future meal!’ (Staden, 67)

It is possible to combine both past and future morphemes to form *ramwer* < *ram* + *p^wer*, and *p^weram* < *p^wer* + *ram*³. Their meaning is illustrated in the examples in (327) and (328):

- (327) a. Mijukaramwera
 (e)mi-juka-**ram-p^wer**-a
 RES-kill-FUT_N-PAS_N-REF
 ‘What would have been killed (but wasn’t).’ (AA, 19v)
- b. Ajaña ratápe soramwera
 Ajaña r-ata=pe so-**ram-p^wer**-a
 Devil REL₁-fire-LOC go-FUT_N-PST_N-REF
 ‘Should have gone to Devil’s fire.’ (DC, II, 77)

³The combination of *p^wer* with *ram*, *p^weram* lexicalizes with the meaning ‘frustration, to frustrate oneself’ (see Example (331)): see Araújo (1618b, 53,84,161) and Anchieta (1595, 34). It is not attested in Old Guaraní (see Montoya 1876, 29 and Restivo 1724, 51(43)), while *ramp^wer* is more frequently attested.

As Bertinetto (2020) observes, the combination of the past with the future marker – retrospective and prospective in his terminology – gives the modal meaning of counterfactuality:

- (328) a. Itĩmwerama
 i-tĩm-**p^wer-ram-a**
 R₂-bury-PST_N-FUT_N-REF
 ‘What will have been his burial.’ (AC, 56v)

P^w and *ram* also appear as the head of an RP, as in (329), or as nominal modifiers (see examples above). This does not imply that these are lexical roots only. They are both functional and lexical roots. They were, at the time of the first descriptions, already in the process of grammaticalization, since in spite of functioning as RPs, they still exhibit an allomorphy similar to that of suffixes (see Rodrigues (2010b) and Cruz (2016, 64)).

In (329), *p^wer* is used as the head of a nominal predicate; in (330), *ram* is the head of an RP, as it also is in 331:

- (329) Ip^wer tekoaiβa
 i-**p^wer** t-eko-aiβ-a
 R₂-PST_N R₄-state-bad-REF
 ‘His/her/their affliction has passed.’ (AA, 33v)

- (330) Neram
 ne=∅-**ram**
 2SG=R₁-FUT_N
 ‘You will be/there is your future.’ (AA, 33v)

- (331) ƚerampwer
 ƚe=∅-**ram-pwer**
 1SG=R₁-FUT_N-PST_N
 ‘I got frustrated.’ (VLB, II, 10)

7.3.2 Diminutive and augmentative

The diminutive form of a noun is (?)*ĩ* and it expresses the small size, small quantity of a referent, or even affection. This is the same suffix used to indicate the lusive aspect of processes (see Section 6.5.3.1).

- (332) a. Pitangĩnamo ereiko
 pitaj-*ĩ*-ramo ere-iko
 child-DIM-TRSL 2SG-be
 ‘You are like a little baby.’ (AP, 100)
- b. Asawsuβ nemembiri
 a-s-awsuβ ne=∅-membir-*ĩ*
 1SG-R₂-love 2SG=R₁-son-DIM
 ‘I love your little child.’ (AP, 102)
- c. Ipeŋ?õp^werĩ jaβiõ sekow?
 i-peŋ?õ-p^wer-*ĩ* jaβiõ seko-w
 R₂-be.in.pieces-DIM each R₂-be-NFOC
 ‘Is it in each tiny piece?’ (DC, I, 216)

The diminutive suffix may also express a small quantity:

- (333) Tojmoja?ok nememira tekokatu?ĩ amõ
 t-o-i-mo-ja?ok ne=∅-memir-a t-eko-katu-*ĩ* amõ
 HORT-3-R₂-CAUS-separate 2SG=R₁-son-REF R₂-be-good-DIM some
 oreβe!
 ore=βe
 1PL.EXCL-DAT
 ‘May your son share some of his virtue with us!’ (Araújo, 32v)

It is also used in many adverbials referring to a short period of time, brevity, closeness, or something that is imminent:

- (334) a. Ko?ĩ ‘very near’
 b. Korite?ĩ ‘soon, quickly’

- c. k^weʔĩ ‘near the hearer’
- d. Meweʔĩ ‘very slow, slowly’
- e. Rameʔĩ ‘similar’

The augmentative form of a noun is marked by the addition of *usu* ~ *wasu* and it expresses the large size of a referent and the positive quality or essence of something or someone.

- (335) a. Okusu
 ok-**usu**-∅
 house-AUG-REF
 ‘Big house.’ (AA, 13v)
- b. Guaisara maranusu
 Guaisara maran-**usu**-∅
 Guaisara battle-AUG-REF
 ‘The big Guaixara battle.’ (Teatro, 20)

Just like the diminutive suffix, albeit its opposite in meaning, the augmentative suffix *-usu* may also express a large quantity, as in (336).

- (336) Arurusu
 a-er-ur-**usu**
 1SG-SCAU-come-AUG
 ‘I brought many/ a large quantity.’ (AA, 13v)

In (337a), *katu* ‘goodness’ functions as a nominal modifier, while in (337b) it functions as a verbal modifier (adverb). However, *katu* can also function as the nucleus of an RP, as in (337c). Besides the referential marker (REF) in (337c) there is no morphology differentiating the functions of *katu* in the examples (337).

- (337) a. Tuβisakatu
 t-uβisaβ-**katu**
 R₄-chief-**good**
 ‘Good chiefs.’ (Poemas, 104)

- b. Ajemĩngatu
 a-je-mim-**katu**
 1SG-R₂-hide-**good**
 ‘I hide myself properly.’ (AT, 34)
- c. Aʔe aʔe koβaʔe katu meʔengara reʔa
 aʔe aʔe koβaʔe **katu**-∅ meʔeng-ar-a reʔa
 DEM PRCL DEM **goodness**-REF give-NMLZ-REF perhaps_m
 ‘It is perhaps this one, the giver of goodness.’ (Araújo, 66v)

7.3.3 Nominal number

Number is not a grammatical category of TUP nouns, which are optionally marked, not by a suffix but by the lexical root *eta* ‘great number, multitude, many), which functions as a nominal modifier (see Anchieta 1595, 8v). The use of *eta* as a plural marker can be seen in (338).

- (338) Jewakaβeta
 je-wak-aβ-eta
 RFLX-adorn-NMLZ-many
 ‘Many ornaments.’ (Poemas, 112)

The usage of *eta* as a lexical root can be understood from examples such as those in (339).

- (339) a. Oré reta
 ore r-eta-∅
 1PL.EXCL R₁-many-REF
 ‘We are many.’ (VLB, II, 44)
- b. Erupe itá kisé amõ?
 ere-ero-ur=pe itá-∅ ∅-kise amõ
 2SG-SCAU-come=Q stone-REF R₁-knife some
 Arureta
 a-ero-ur-eta
 1SG-SCAU-come-many

- ‘Did you bring some iron knives? Yes, I brought many.’ (Léry, Histoire, 346)
- c. Oretupãoketa
ore= \emptyset -tupã-ok-eta
1PL.EXCL=R₁-God-house-PL
‘We have many churches.’ (Poemas, 114)
- d. Jeanameta aroporasej seru
je= \emptyset -anam-eta a-ero-porasej s-er-ur
1SG=R₁-relative-PL 1SG-SCAU-dance R₂-SCAU-come.GER
‘Bringing my (many) relatives, I make them dance with me.’ (Poemas, 138)
- e. Kunumĩeta
kunumĩ-eta- \emptyset
boy-many-REF
‘(Many) boys.’ (AT, 26)
- f. Tataendieta osepiak
tata-endi-eta- \emptyset o-s-epiak
fire-flame-many-REF 3-R₂-see
‘They saw (many) flames of fire.’ (see Araújo, 45)

Another way of indicating plurality or collectives other than *eta* is through reduplication. Although poorly attested, it is reasonable to suspect that this strategy was somewhat productive because it is present in many TG languages. In TUP one finds: *mirĩ* ‘small thing’ (Anonymous 1952a, I, 78) and *mirĩmirĩ* ‘small things’ (Anonymous 1952a, II, 39); with the root *tiŋ*, there are formations such as *ti-tiŋ-a*: *?a-titia* ‘white stains on the hair’ (Anonymous 1952a, II, 29), *?api-titia* ‘white stains on the head’, *titiŋa* ‘white stains on the skin, yeasts’.⁴

⁴Anonymous (1952a, II, 10) translates: ‘to have yeast’ *je= \emptyset -ti-tiŋa* and if there are many syllables, then the last one is repeated, as in *je= \emptyset -ti-ti(ŋ)-tiŋ-a*. The formation is also seen in the names of some white-spotted animals: *pikitia* = *piktitiŋa* (Marcgrave and Piso 1648, 159), *jabutitiŋa* (Marcgrave and Piso 1648, 241) and Anonymous (1952a, I, 62). In these examples, it is not the modifying element that is pluralized. The head noun and its modifier become one word, and reduplication affects the last syllable, so that it is natural that only

When accompanied by numerals, nouns remain unchanged and cannot combine with *eta*:

- (340) Mokōj apiaβa
 mokōj apiaβa-∅
 two man-REF
 ‘Two men.’ (AA, 9v)

The collective can be expressed by adjoining roots to the root *tiβ* ‘gathering, set, great number of’.

- (341) a. Takwarasutiβa
 takwara-su-tiβ-a
 bamboo-big-abundance-REF
 ‘Place of many bamboos / canebrake.’ (Staden, 116)
- b. Paʔiwasu irūndiβa
 paʔi-wasu-∅ irū-tiβ-a
 master-big-REF companion-set-REF
 ‘The companions of the master.’ (Poemas, 114)

7.3.4 Alienability and Inalienability

Tupinambá has one class of possessed and one of non-possessed nouns. The former consists of utensils, kinship, and body parts terms, and the latter of natural elements such as trees, animals, celestial bodies, and etc. A alienable and inalienable distinctions seems to have been lost, since a mark of alienability *-e-* that still exists in some languages such as Satere-Mawé (Silva et al. 2010), Awetí (Reiter 2012), and Mundurukú (Gomes 2006), has left traces in a Tupinambá. It can be seen in some nouns of class IId (see Table 4.5). Nouns like *(e)kuj* ‘gourd’, *(e)pinõ* ‘fart’, *(e)poti* ‘feces, defecate’, *(e)panaku* ‘basket’, and

the modifying element is reduplicated. A similar phenomenon is seen in verbal reduplication, which sometimes reduplicates not only the verbal root but also the bound index, because some roots are monosyllabic, as in *eresiresik* ‘you frequently arrive’ (*ere-* + *sik*). Further examples are: *iβitiβitira* ‘mountain range’ Anonymous (1952a, II, 60), *mitamita* Anonymous (1952a, II, 132). See also (413).

(*e*)*nimo* ‘thick thread’ still maintain the initial *e* in some form of the possessed paradigm, as in the example (342), where the possessed form has the *e*:

- (342) a. Wajβĩ rekuja
 wajβĩ-∅ r-ekuj-a
 elderly.woman-REF R₁-gourd-REF
 ‘Old women’s gourds.’ (AT, 30)
- b. Kuja jẽ itĩŋa-tĩŋáβo
 kuj-a jẽ i-tĩ-ka-tĩ-ka-βo
 gourd-REF PRCL R₂-point-break-point-break-GER
 ‘Continuously breaking the tips of the gourds.’ (AT, 170)

7.4 Postpositions

Postpositions form an important minor closed class in Tupinambá. They express temporal or spatial relations between parts of a sentence: its object or complement and the predicate or a non-predicative noun (Hagège 2010, 1). When the complement of a postposition is a person index, only Set I indexes (see Table 4.3) are used.

As already mentioned (see Section 4.3.2), postpositions take relational markers, signaling their contiguity with their dependents (objects), possibly because these are grammaticalized lexical roots. The following sections provide an overview of some postpositions and their meanings and uses.

In Section 4.3.2, it was stated that postpositions combine (obligatorily) with relational markers, as in (343):

- (343) a. Aʔar nepupe
 a-ʔar ne=∅-**pupe**
 1SG-fall 2SG=R₁-**POSP**
 ‘I board with you.’ (Anch., Arte, 40v)
- b. Oso jerenone
 o-so je=r-**enone**
 3-go 1SG=R₁-**ahead.of**

‘He goes/went ahead of me.’ (FA, 122)

- c. Isupe e
i-DAT e
R₂-to PRCL

‘To him indeed.’ (Anch., Arte, 54)

- d. Sese oroso
s-ese oro-so
R₂-with 1PL.EXCL-go

‘I go with him / We go together.’ (AA, 44v)⁵

Nonetheless, there are instances of postpositions that seem to be in the process of grammaticalization, displaying an intermediary stage between, e.g., postpositions and case markers (see Section 7.8). These cases are well known cross-linguistically and have been discussed in the literature (see e.g., Hopper and Traugott 2003; Hagège 2010; Lehmann 2015; Kuteva et al. 2019). Here they are referred to as unstressed suffixes which are not preceded by relationals. The most frequent unstressed suffix is *supe* ‘to, against’ (see Section 5.3).

- (344) a. Ise supe⁶
ise **supe**
I to

‘To me.’ (VLB, II 64)

- b. Eresikijpe Ajãja, Tawaiβa, Kurupira, Jurupari kojpo
ere-s-ikij=pe Ajãja Tawaiβa Kurupira Jurupari kojpo
2SG-R₂-invoke=Q Anhanga Taguaiba Kurupira Jurupari or
teʔõ aβa supe?
t-eʔõ-∅ aβa-∅ **supe**
R₄-death-REF person-REF to

‘Did you invoke the devil, Taguaiba, Kurupira, Jurupari or the death upon someone?’ (Araujo, 102v)

⁵The postposition *ese* requires a plural subject index with some verbs (Anchieta 1595, 44v).

⁶While *supe* can follow an free pronoun, it cannot attach to a bound index.

7.4.1 Postposition *swi*

Swi is an ablative postposition and thus indicates motion away from something.

- (345) a. Ajur fekoswi
 ajur fe=∅-ko-∅ ∅-**swi**
 1SG-come 1SG=R₁-slash-REF R₁-**from**
 ‘I came from my slash.’ (FA, 9)
- b. Emonã rako sekow neswi
 emonã rako s-eko-w ne=∅-**swi**
 this.way ADV R-act-NFOC 2SG=R-**from**
 ‘Thus she reacted, actually, in your absence (being away from you).’ (Araújo, 74)
- c. iβaka swi erejur
 iβak-a **swi** ere-jur
 sky-REF **from** 2SG-come
 ‘You came from the sky/heaven.’ (Poemas, 100)

7.4.2 Postposition *ese*

Ese combines instrumental and comitative meanings, as in (346).

- (346) Nerese memẽ orojkó
 ne=r-**ese** memẽ oroiko
 2SG=R₁-**with** always 1PL.EXCL-be
 ‘We are always with you.’ (Poemas, 84)
- (347) Nerejk^waβipe koʔir teʔõ nerese
 n-ere-i-k^waβ-i=pe koʔir t-eʔõ-∅ ne=r-**ese**
 NEG-2SG-R₂-know-NEG=Q now R₄-death-REF 2SG=R₁-**with**
 seko?
 s-eko
 R₂-be
 ‘Do you not know that death is with you?’ (D’Abbeville, Histoire, 350)

Besides the comitative meaning, *ese* expresses direction: ‘towards’.

Similar to the postposition *ese* is the postposition *i*⁷, which cannot be used with the non-contiguous marker (R₂).

- (348) a. Janereko pupe p^lã jaiko tekoaiβa
 jane=r-eko-∅ pupe p^lã ja-iko t-eko-aiβ-a
 1PL.IN=R₁-act-REF with PRCL 1PL.IN-be R₄-action-evil-REF
 rine?
 r-**i**=ne
 R₂-**in**=FUT
 ‘Through our actions, by the way, will we be in sin.’ (CC, 6, 46)
- b. feputupaβ jẽ neri
 fe-∅-putupaβ jẽ ne=r-**i**
 1SG=R₁-surprise PRCL 2SG=R₁-**with**
 ‘You surprised me! (I am surprised because of/with you)’ (Léry, Histoire, 353)

7.4.3 The case of *irũ*

Irũ is not a comitative postposition. It is a lexical root meaning ‘companion’, as in (349a). It often receives the translative case (see Section 7.8.3) to take the form *irũramo* (*irũ* + *ramo*) ‘as/in the quality of a companion’ in order to express a comitative meaning, as in (350).

- (349) a. Tajko neirũ
 t-a-iko ne=∅-**irũ**
 HORT-1SG-be 2SG=R₁-**companion**
 ‘May I be your companion.’ (Léry, Histoire, 372)
- b. Tereiko paʔi Nikora irũ
 t-ere-iko paʔi Nikora ∅-**irũ**
 HORT-2SG-be master Nicolau R₁-**companion**
 ‘May you live as a companion of master Nicolau (with Nicolau).’
 (Léry, 352)

⁷In Navarro (2013), the form *ri* is given.

- (350) a. Neirūramo janejara rekow
 ne=∅-**irũ-ramo** jane=∅-jar-a r-eko-w
 2SG=R₁-**companion-TRSL** IPL.INCL-lord-REF R₁-be-NFOC
 ‘Our Lord is with you (is as your companion).’ (Araújo, 31v)
- b. Aβape irūramo turine?
 aβa=pe **irũ-ramo** t-ur-i=ne
 person=Q **companion-TRSL** R₂-come-NFOC=FUT
 ‘Who will come with him/as his companion?’ (Araújo, 46v)

One possible source of confusion is the fact that *irũ* has an initial *i* and therefore, when preceded by the relational of non-contiguity (R₂), no additional *i* is written down. This has led to confusion where *irũ* was often understood as a postposition, a fact already noted by Anchieta (1595, 6). Compare (351a) with (351b).

- (351) a. Irũ
 irũ
 ‘(A/The) Companion(s).’ (AA, 6)
- b. Irũ
 i-irũ
 R₂-companion
 ‘With him / He has (a) companion(s).’ (AA, 6)

7.4.4 Some special cases

Some postpositions cover a wide range of meanings. This can be illustrated with examples containing the postposition *ese*, which has locational or directional meaning:

- (352) a. Teʔõ nerese seko
 t-eʔõ-∅ ne=r-**ese** s-eko
 R₃-death-REF 2SG=R₁-**with** R₂-be
 ‘Death is with/at you.’ (D’Abbeville, Histoire, 350)

- b. Tupana rese ajko
 Tupana r-**ese** a-iko
 God R₁-**in** 1SG-be
 ‘I live in contact with God.’ (FA, 166)
- c. Ne emoŋeta neTupã tok^waβ ?e
 ne e-moŋeta ne=∅-Tupã t-o-k^waβ ?e
 You 2SG.IMP-pray 2SG=R₁-God HORT-3-pass say
 amanusu janemomaranē?ima
 aman-usu-∅ jane=∅-mo-maran-e?im-a
 rain-AUG-REF 1PL.INCL-R₁-CAUS-destroy-PRIV-REF
 rese
 r-ese
 R₁-in-order.to
 ‘Pray your God for the storm to be over, for us not to be destroyed.’
 (Staden, 66)

Just as one postposition expresses different meanings, one and the same meaning may be expressed by different postpositions, as in (353), where the comitative meaning is expressed by different roots:

- (353) a. Tisa?aŋ apiaβa marã
 t-ja-s-a?aŋ apiaβa-∅ ∅-marã-∅
 HORT-1PL.INCL-R₂-try man-REF R₁-power-REF
 janeirũ
 jane=∅-**irũ**
 1PL.INCL=R₁-**companion**
 ‘May we experience the strength of the men, our companions.’
 (Léry, Histoire, 357)
- b. Ma?epi?awp^lara kawĩaiβasi rese
 ma?e-pi?a-ip^lar-a kawĩ-aiβ-asi-∅ r-**ese**
 thing-liver-enemy-REF beer-bad-pain-REF R₁-**with**
 imonani, ipupe se?ima
 i-monan-i i-pupe s-e?im-a
 R₂-mix-NFOC R₂-to R₂-serve.drink-GER

‘Giving him to drink a gall-like substance mixed with vinegar.’
(Araújo, 63v)

- c. Oroso Pedro ni
oro-so Pedro \emptyset -**ni**
1PL.INCL-go Pedro R₁-**with**
‘I go with Pedro.’ (AA, 44)

7.4.5 Other postpositions

This section offers a list of postpositions and postpositional clitics. Each is presented with their corresponding translations and an attested example.

- (354) feraje ise mi
 fe=r-aje i-sem-i
1SG=R₁-**crosswise** R₂-exit-NFOC
‘It came out of me crosswise.’ (VLB, I, 102)
- (355) Lisboa akweakoti
Lisboa \emptyset -**akwe-a-koti**
Lisbon R₁-**other-NMLZ-towards**
‘To the other side of Lisbon.’ (VLB, I, 48)
- (356) Sakip^weri aso
s-**akip^wer-i** a-so
R₂-**footsteps**-LOC 1SG-go
‘I go behind him.’ (VLB, II, 135)
- (357) $\text{i}\beta\text{itira}$ $\text{amo}\eta\text{oti}$
 $\text{i}\beta\text{itir-a}$ \emptyset -**amõ-koti**
mountain-REF R₁-**other-side**
‘Beyond the mountains.’ (VLB, I, 31)
- (358) Arasó jeapiri
a-era-só $\text{je}=\emptyset$ -**apiri**
1SG-CC-go 1SG=R₁-**along**

‘I take it with me.’ (VLB, I, 35)

- (359) iβitinga ʔariβo
iβitinga ʔariβo
cloud R₁-**above**

‘Above the clouds.’ (Araújo, 56v)

- (360) ʃeatuaj turi
ʃe=∅-**atua-j** t-ur-i
1SG=R₁-**back**-LOC R₂-come-NFOC

‘He came from behind me (His coming was in my back).’ (AA, 41v)

- (361) K^weseβé maʔe naʔuj
k^wese-βé maʔe-∅ n-a-ʔu-i
yesterday-**more** thing-REF NEG-1SG-eat-NEG

‘I haven’t eaten since yesterday.’ (Poemas, 150)

- (362) Kaʔáβo
kaʔá-βo
wood-**PERL**

‘Through the woods.’ (VLB, II, 81)

- (363) Osó ʃerenone
o-so ʃe=r-**enone**
3-go 1SG-R₂-**in.front.of**

‘He went in front of me.’ (FA, 122)

- (364) Asawsuβ Pedro taʔira reseβe
a-s-awsuβ Pedro t-aʔir-a r-**eséβe**
1SG-R₂-love Pedro R₂-son-REF R₂-**together.with**

‘I love Pedro and his son.’ (AA, 44)

- (365) Etupãmojeta orerese
e-tupã-mojeta ore=r-**ese**
2SG.IMP-God-pray 1PL.EXCL-R₁-**for**

‘Pray (God) for us.’ (DC, I, 148)

TUP shows a process which is reconstructed for Proto-Tupí-Guaraní, in which oblique (case) markers combined with spatial nouns or body-parts have been grammaticized as postpositions or adverbials (see Jensen 1999, 149). Some examples are given in Table 7.1.

Lexeme	Meaning	Case	Lexeme	Meaning
wiri	under	pe (locative)	wiripe	below
ʔar	top	βo (perlative)	ʔariβo	above
ʔar	top	i (locative body-part)	ʔari	above, on top of
p̄i	foot	pe (locative)	p̄ipe	near
iβir	margin	i (locative body-part)	iβiri	along
atua	neck	i (locative body-part)	atuj	behind, after
akip ^w er	back part	i (locative body-part)	akip ^w eri	on the trail of

Table 7.1: Some TUP postpositions formed out of spatial nouns or body-parts

7.5 Postposition assignment

Non-macrorole core arguments of ditransitive verbs or of intransitive verbs often require an indirect complement (see Sec. 5.3). These non-macrorole arguments can be seen under the BECOME or INGR operators in the logical structure. The example below with the verb *meʔej* ‘give’ illustrates the argument marked by the postposition:

- (366) Aβa supe imaʔe meʔeɲa
aβa-∅ **supe** i-maʔe-∅ meʔeng-a
man-REF DAT R₂-thing-REF give-GER
 ‘Giving men their things (justice).’ (DC, 153)
 [do’ (x, ∅)] CAUSE [BECOME **have**’ (aβa, imaʔe)]

The assignment of postpositions to non-macrorole core arguments follows strict rules that are language-specific (see Van Valin Jr and LaPolla 1997, 352-

383). In the case of *supe*, for example, the following rule can be formulated (367):

- (367) *Rule assigning supe in TUP*
 Assign *supe* to the non-macrorole *x* argument, if it is third person, in the logical structure segment: ... BECOME/INGR **pred'** (x,y)

In case the non-macrorole core argument is first or second person, the rule must be formulated in terms of (368).

- (368) *Case assignment for first- and second-person arguments in TUP*
 Assign *-βε / -βο* to the non-macrorole *x* argument in the logical structure segment: ... BECOME/INGR **pred'** (x,y)

First and second person combine with the dative case *-βε / -βο*, but not with *supe*, as shown in (369):

- (369) a. Tasenōj neβε
 t-a-s-enōj **ne-βε**
 HORT-1SG-R₂-name **you-DAT**
 ‘May I name it for you.’ (cf. Léry, 360)
- b. Peimoŋeta Tupã janejara iseβε
 pe-i-moŋeta Tupã jane=∅-jar-a **ise-βε**
 SG-R₂-talk God 1PL.INCL=∅-lord-REF **I-DAT**
 ‘Pray you to God for me.’ (DC, I, 190)

Another example is given for the postposition *swi* ‘from’. The rule for its assignment is given in (370). An example is given in (371) with the verb *?ok* ‘cut, remove, rip out’:

- (370) *Rule assigning swi in TUP*
 Assign *swi* to the non-macrorole *x* argument in the logical structure segment: ... BECOME/INGR NOT **pred'** (x,y)

- (371) Tojeʔok ise swi feresaporopotara
 t-o-i-eʔok ise ∅-swi je=r-esa-poro-potar-a
 HORT-3-R₂-cut I R₁-from 1SG=R₁-exe-ANTIP-want-REF
 ‘May my lustful eyes be pulled out of me.’ (AP, 146)
 [do’ (3, ∅)] CAUSE [BECOME NOT have’ (ise, esaporopotara)]

7.6 Postpositional Phrase

The layered structure of the postposition (LSPP) is paralleled by the structure of the clause and LSRP (see Section 3.2) and by the structure of the RP (see Chapter 8). Its structure is given in Figure (7.3). The object of the postposition is a core argument and thus is inside the core_p.

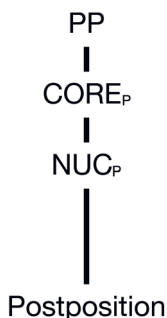


Figure 7.3: Structure of a predicative PP

Postpositional phrases can be predicative or non-predicative, and there are also postpositional phrases that function as arguments of the main predicate. Non-predicative PPs lack a layered structure. In this case, the postposition is not a semantic predicate, and the object is not a semantic argument of the adposition.

7.6.1 Adjunct postpositional phrases

Predicative postpositions are, as the name indicates, predicates. They provide semantic information about the clause in which they occur, both in terms of

their own meaning and in terms of the meaning of the RP that occurs with them (their argument). They are therefore adjuncts (or adverbials), elements that modify in some way the event or situation described by the main predicate. This function is reflected in their semantic and syntactic representation. As shown in (372), represented in Figure 7.4, the predicative PP takes the whole of the core as its second argument, and the adjunct postpositional phrase appears in the syntactic periphery.

- (372) Oimoasipe aʔe rire aʔe iβaʔuawera?
 o-i-mo-asi=pe aʔe ∅-rire aʔe iβa-ʔu-p^wer-a
 3-R₂-CAUS-pain=Q this R₁-after this fruit-eat-PST-REF
 ‘Did he regret it after eating that fruit? (Did he regret after it the past eating of this fruit?)’ (DC, I, 163)
be-after’ [**do’**(3) [**eat’**(3,)]] <_{IF} INT [**regret’**(3,3)] >

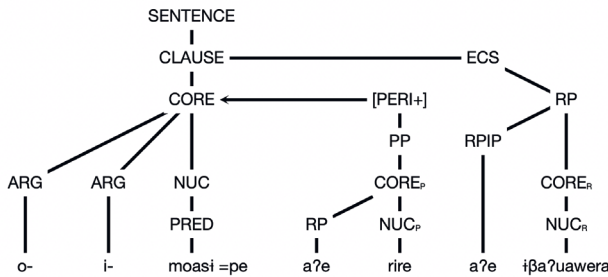


Figure 7.4: Predicative postposition

7.6.2 Argument-marking postpositional phrases

Contrary to adjunct postpositional phrases, argument-marking postpositions are non-predicative, because they are arguments of a predicate. They mark OCA of the predicate with a postposition (in case an oblique case is not used). The difference is clear when one compares example (372) with example (373), where the oblique core argument is marked by the dative case, or with example

(374), where the PP is an argument adjunct and therefore appears inside the core. The syntactic representation of (374) is given in Figure 7.5.

- (373) Eimeʔej pina iseβe
 e-i-meʔej pina-∅ ise-βe
 2SG.IMP-R₂-give fishhook-REF I-DAT
 ‘Give fishhooks to me.’ (AA, 34)
 ⟨_{IF} IMP [**do**’ (2SG, ∅)] CAUSE [BECOME **have**’ (ise, pina)] ⟩

- (374) Aotinga onoŋ ase rese
 aoβ-tiŋ-a o-noŋ ase r-ese
 clothes-white-REF 3-put we R₁-POSP
 ‘Put white garments on us.’ (CA, 81v)
 ⟨ [**do**’ (3, ∅)] CAUSE [BECOME **be-LOC**’ (aotiŋ, ase)] ⟩

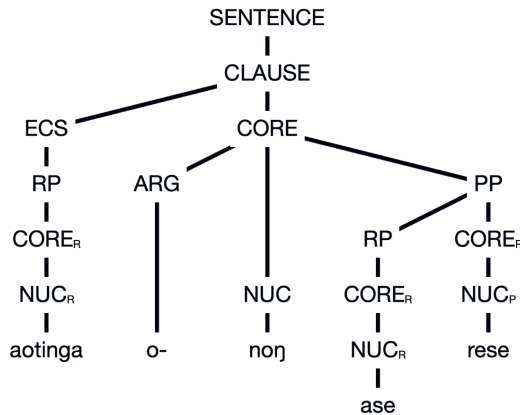


Figure 7.5: PP as core argument

7.7 Particles

Particles are morphological entities realized as a phonologically free unit (see Bickel and Nichols 2005). Tupinambá, like other Tupían languages, is rich

in particles, which confer nuances of many kinds to the sentence. They are operators, since they express grammatical features of lexical words such as temporal, modal, aspectual, discursive (e.g., discourse relations between propositional units), and illocutionary information (e.g., evidentiality). Although particles constitute a closed class in TUP, they are numerous. Some of these are illustrated in the following sub-sections. Each particle often expresses nuances of meanings, or even different meanings. In order to provide a clearer picture, a study devoted exclusively to TUP particles would be necessary, but such a study would face the fact that real discourse texts, such as legends, myths, etc. by native speakers are not attested, so it is impossible to know exactly what the meanings of certain particles are.

7.7.1 Aspect particles

The particle *mã* is often used to describe an action that is about to be initiated (inceptive aspect). It often accompanies the optative illocutionary force (see section 6.5.1.1), but it may stand alone. It always appears in clause final position:

- (375) a. *ƒerejtik korine mã!*
 ƒe=r-ejtik kori=ne mã
 1SG=R₁-defeat today=FUT PRCL
 ‘I will be defeated today / there will be my defeat today!’ (AT, 28)
- b. *Naijukaiswe temõ mã*
 n-a-i-juka-i-swe temõ mã
 NEG-1SG-R₂-kill-NEG-PRCL PRCL_{opt} PRCL
 ‘I hope I do not kill him / oh, if I only didn’t kill him.’ (Fig., Arte, 27)

7.7.2 Core modifying particles

Desideratives, intentionals, and deliberative necessarily refer to one of the core arguments, hence they must be core-level modifiers.

Temõ indicates the desiderative aspect. It appears seldom without the particle *mã*.

- (376) Iariβe temõ jeswi mã
 i-ariβe temõ je-swi mã
 R₂-cease PRCL_{opt} 1SG-ABL INTJ
 ‘Ah, if it only would cease.’ (Araújo, 165)

The particle *ka* expresses intention:

- (377) a. Aso ka
 a-so **ka**
 1SG-go PRCL
 ‘I intend to go.’ (FA, 139)
- b. Asap¹akatupe aḡire ka
 a-s-ap¹a-katu-pe aḡire **ka**
 1SG-R₂-obey-well-PRCL henceforth PRCL
 ‘I intend to obey him well henceforth.’ (Araújo, 77)

7.7.3 Discourse particles

The particle *e* emphasizes a constituent, sometimes as contrastive focus:

- (378) Ene e ajpo ere, eʔi
 ene e ajpo ere-ʔe e-ʔi
 you PRCL DEM 2SG-say, 3-say
 ‘They say that YOU are saying it.’ (Araújo, 56)

Another emphatic particle is *jẽ*:

- (379) Ajpo rese jẽ, koʔi asausu
 ajpo r-esé jẽ, koʔi a-s-awsuβ
 DEM R₁-because PRCL now 1SG-R₂-love
 ‘Because of this, in effect, I now love him.’ (Poemas, 108)

The particle *te* is focal (see Anchieta 1595, 36:

- (380) a. Maʔeasiβorate ajekatu iʔuu
 maʔe-asi-βor-a-te ajekatu i-ʔu-u
 thing-ill-HAB.AG-REF-FOC well R₂-ingest-NFOC
 ‘The ill certainly eat well.’ (Araújo, 77v)
- b. ʃemĩte jepé iswí
 ʃe=∅-mim-te jepé i-swi
 1SG=R₁-hide-FOC PRON R₂-from
 ‘Do hide me, from him.’ (Teatro, 34)

The particle *ri* has dubitative meaning.

- (381) Asópe isene ri?
 a-so=pe ise=ne ri
 1SG-go=Q I=FUT PRCL
 ‘Will I go?’ (VLB, II, 58)

The clitic *pe*, usually accompanied by sentence final particles *ka* used by men, or *ki* used by women, have a deliberative sense, i.e., they express the intention to undertake an action. Both, *ka* and *ki* may be used without *pe*. The future marker =*ne* may also be used with *ka* or *ki*.

- (382) a. Asóne ki
 a-so=**ne** **ki**
 1SG-go=FUT **DELIB_F**
 ‘I have to go (I intend to go and have decided I will).’ (FA, 139)
- b. Aso ka
 a-so **ka**
 1SG-go=FUT **DELIB_M**
 ‘I have to go (I intend to go and have decided I will).’ (FA, 139)
- c. Aso umēpe ki
 a-so umē=**pe** **ki**
 1SG NEG=**DELIB PRCL_F**
 ‘I have not to go (I intend not to go and have decided I will not).’
 (AA, 23)

- (384) a. Neʔĩ sekija koʔite!
neʔĩ s-ekij-a koʔite
 EXCL R₂-pull-GER finally
 ‘Pull him! (Ah, may there finally be his pulling)’ (VLB, II, 58)
- b. Neʔi toso!
neʔi t-o-so
 EXCL HORT-3-go
 ‘May he go!’ (AA, 56v)

7.8 Case

‘Cases mark dependent nouns for the type of relationship they bear to their heads. Traditionally, the term refers to inflectional marking’ (Blake 2004, 1). TUP has case endings for expressing some grammatical functions, which are given below.

7.8.1 Locative-Dative cases

The locative is expressed by *pe*, which also expresses motion towards the referent (allative), and also marks the recipient or beneficiary (dative). Locative examples are given in (385). All variants are here glossed as LOC.

- (385) a. ʃepope
 ʃe=po-**pe**
 1SG=hand-LOC
 ‘In my hand(s).’ (AT, 48)
- b. Nejurarawaj tápe
 ne=∅-jurarawaj taβ-**pe**
 2SG=R₁-lie village-LOC
 ‘You lie in the village.’ (DC, II, 84)
- c. Maʔeramaripe ase típe oeni
 maʔe-rama-ri=pe ase ∅-tĩ-**pe** o-eni-∅
 thing-FUT-PRCL=Q our R₁-nose-LOC CORF-saliva-REF

moini?

mo-in-i

CAUS-lay-NFOC

‘Why do you put your saliva on our nose?’ (Araújo, 81v)

Body parts have their own locative case marker, *-i*.

- (386) a. Oajuri serekóβo
 o-ajur-i s-ereko-βo
 CORF-neck-LOC R₂-be-GER
 ‘Having them on the neck.’ (Araújo, 12v)
- b. Pítaj
 pita-i
 heel-LOC
 ‘On the heel.’ (AA, 41v)

This case ending also appears in many postpositions and adverbials with the ending already grammaticalized, i.e., as part of the word and no longer analyzable as a case suffix, as in *wiri* ‘below’ (AA, 41v), *píri* ‘near, close to’ (FA, 126), *pukuj* ‘along (time and place)’ (VLB, II, 130), and *iβiri* ‘along’ (VLB, I, 106).

Examples of *-pe* expressing motion towards a referent (allative) are given in (387).

- (387) a. Aso okipe
 a-so ok-**pe**
 1SG-go house-LOC
 ‘I go home.’ (Anch., Arte, 40)
- b. Asopotar iβakipe
 a-so-potar iβak-i-**pe**
 1SG-go-want sky-EPEN-LOC
 ‘I want to go to heaven.’ (Araújo, 248)

- c. Eike kori feji?ãpe
 e-ike kori fe=Ø-ji?ã-**pe**
 2SG.IMP today 1SG=R-heart-POSP
 ‘Enter in my heart today.’ (Poemas, 92)

Examples of *-pe* marking the beneficiary or the recipient are shown in (388).

- (388) a. Aime?eŋ feruβape
 a-i-me?eŋ fe=r-uβ-a-**pe**
 1SG-R₂-give 1SG=R₁-father-REF-DAT
 ‘I gave it to my father.’ (AA, 42)
- b. Aime?eŋ aβa supe
 a-i-me?eŋ aβá-Ø **supe**
 1SG-R₂-give person-REF **POSP**
 ‘I gave it to the Indians.’ (Teatro, 48)
- c. Orejara ahẽpe
 ore=Ø-jar-a ahẽ-**pe**
 1PL.EXCL-R₁-carrier-REF DEM-DAT
 ‘We are carriers of goods to him.’ (Léry, 362)

Since the non-contiguous marker cannot receive case markers, a special form, *supe*, is used, which cannot be used with first or second person indexes. It is also used with RPs.

- (389) a. Erépe amõ kujã supe...?
 ere-?i=pe amõ kujã-Ø **supe**
 2SG-say=Q some woman-REF **to**
 ‘Did you say to any woman...?’ (Araújo, 104)
- b. Pitaŋ-ĩ supe ou
 pitaŋ-ĩ **supe** o-ur
 child-DIM **to** 3-come
 ‘They came to the little baby.’ (Poemas, 194)

- c. Moruβisaβa tuiβaʔe ojeʔeŋ memē isupe
 moruβisaβ-a tuiβaʔe o-jeʔeŋ memē **i-supe**
 chief-REF old 3-speak always R₂-**to**
 ‘The old chiefs always speak to him.’ (Teatro, 36)
- d. Tekoβe meʔeŋa isupe
 t-ekoβe-∅ meʔeŋ-a **i-supe**
 R₄-life-REF give-GER R₂-**to**
 ‘Giving life to him.’ (Araújo, 39)

Free pronouns have a special dative form which consists of the pronominal morpheme and the suffix *-βe* or *-βo*:

- (390) a. Oʔa janéβo kori
 oʔa jane-βo kori
 3-born 1PL.INCL-**DAT** today
 ‘He was born to us today.’ (Poemas, 94)
- b. Ijirō ipo kori iseβene
 i-jirō ipo kori ise-βe=ne
 R₂-forgiveness certainly today I-**DAT**=FUT
 ‘He will certainly forgive me today.’ (Araújo, 92v)
- c. Ajpo aβa jara janéβe
 ajpo aβa-∅ jar-a **jane-βe**
 DEM man-REF carrier-REF 1PL.INCL-**DAT**
 ‘These men are carrier of goods to us.’ (Léry, Histoire, 354)
- d. Ise aʔe ã aʔe umwã nako peñe
 ise aʔe ã a-ʔe umwã nako peñ-**me**
 I PRCL PRCL 1SG-say already PRCL 2PL-**DAT**
 ‘Behold, it is I, I already said this to you.’ (Araújo, 54v)

7.8.2 Perlative case

The perlative case, which is often called the *diffuse locative* in the TG literature (see e.g. Rodrigues 1996a; Jensen 1998a; Seki 2000) indicates a movement which goes ‘through, across, along’.

- (391) Kóβo
 ko-βo
 slash-PERL
 ‘Through the slash(es).’ (AA, 42)

The perlative marker *-βo* lends a plural reading to the word it attaches to⁸. This is illustrated in (392):

- (392) a. Ipotasaβokatu
 i-pota-sa-βo-katu
 R₂-want-NMLZ-PERL-truly
 ‘Purely through his wish.’ (Araújo, 53)
- b. Kaʔaβo ajko
 kaʔa-βo a-iko
 forest-PERL 1SG-be.in.movement
 ‘I go through the woods.’ (VLB, II, 41)
- c. Aso okiβo
 a-so ok-βo
 1SG-go house-PERL
 ‘I go through the houses.’ (FA, 7)

7.8.3 Translative case

The translative case indicates a change in state, which may be temporary.

- (393) a. Pitaŋamo seni Maria jiβape
 pitaŋ-amo s-en-i Maria jiβa-pe
 child-TRSL R₂-sit-NFOC Maria arm-LOC
 ‘As a child he is in Maria’s arms.’ (Poemas, 106)
- b. Marã oikóβotepe ase Aŋaŋa remiawsuβamo
 marã o-eko-βo-te=pe ase Aŋaŋa r-emiawsuβ-amo
 how 3-be-GER-FOC=Q we devil R₁-friend-TRSL

⁸This is not the case when the morpheme *-βo* has the meaning of ‘according to’.

sekow?

s-eko-w

R₂-be-NFOC

‘How are we/do we act like friends of the devil?’ (Araújo, 26)

Reference Phrase

Section 3.2 briefly introduced the LSRP. This section discusses the RP in TUP in detail, considering its operators and different types of modifiers.

8.1 Minimal RPs

A minimal RP in TUP may consist of a single lexical root, a proper noun, or a pronoun. A minimal RP consisting of only a noun is given in (394), with the LSRP of (394a) given in Figure 8.1.

- (394) a. Tuβa
t-uβ-a
R₄-father-REF
'(A/the) Father(s).'
- b. Jawara
jawar-a
jaguar-REF
'(A/the) jaguar(s).'



Figure 8.1: LSPR of a minimal RP

The presence of the relational morpheme preceding the lexical root in (394a) indicates that the root is possessed, even though a possessor is not specified in this case (see Section 4.3). When the root is unpossessed, the relational marker is absent, as in (394b). An RP does not require the referential morpheme, which only functions as an indicator that the lexical root is not predicative.

A minimal RP consisting of just a pronoun has the same structure as the examples in (394). In TUP, indefinite pronouns are RPs on their own, as in (395a), which is given as an answer to the question *Will everything burn?*

- (395) a. Paβ
 everything
 ‘All / Everything.’ (VLB, II, 130)
- b. Ise
 I
 ‘I.’ (cf. Teatro, 8)

Demonstrative pronouns

There is often a synchronic or diachronic relationship between demonstratives and third person pronouns (Rijkhoff 2002, 174). Among TUP demonstratives, there is one, *aʔe*, which often functions as a third person (independent) pronoun. In general, TUP demonstratives may function pronominally as arguments if combined either with the referential suffix (REF) or the nominalizer suffix *-βaʔe*.

- (396) a. Tupã aʔeβaʔe rejtika tatape
 tupã-∅ aʔe-βaʔe r-ejtik-a t-ata=pe
 God-REF DEM-NMLZ R₂-throw-GER R-fire-POSP
 ‘(...) God throwing those in the fire.’ (DC, I, 193)

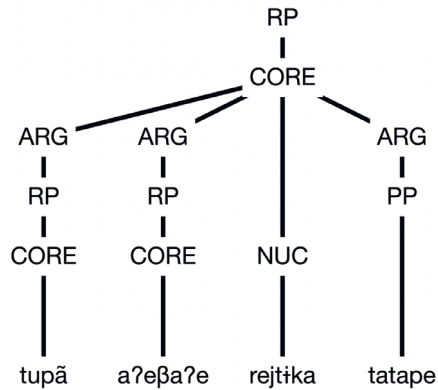


Figure 8.2: Demonstrative in argument function

- (397) a. Taβusupe wĩ?
 taβ-usu=pe wĩ-∅
 village-AUG=Q DEM-REF
 ‘Is this a city?’ (Léry, *Hist.*, 361)
- b. Jesus βojá ã iko
 Jesus ∅-βoja-∅ ã iko
 Jesus R-disciple-REF PTCL DEM
 ‘Behold, this is Jesus’ disciple.’ (Ar., *Cat.*, 79)
- c. Eβok^we nememira, kujã we!
 eβok^we ne=∅-memir-a kujã we
 DEM 2SG=R₁-son-REF woman VOC
 ‘Oh, woman, this is your son!’ (Ar., *Cat.*, 63)
- d. Ak^wej komã! Emonã rako jeeβokweja rerekow
 ak^wej komã emonã rako je=eβokwej-a r-ereko-w
 DEM here INTJ in.fact 1SG=**this-REF** R₁-treat-NFOC

‘If only that (one) were here! Thus, in fact, I treat this (one).’
(Anch, *Dout. II*, 93)

- e. Ak^wej temō our fepose mǎ!
ak^wej temō o-ur fe=∅-pose mǎ
 DEM PTCL_{opt} 3-come 1SG=R₁-towards VOC
 ‘If only that (one) came to me!’ (Anch, *Dout.*, 96)

- f. ?anga jape peroka?
?aŋ-a ja=pe pe=r-ok-a
 DEM-REF alike=Q 2PL=R₁-house-REF
 ‘Are your houses like these?’ (Léry, *Hist.*, 363)

- g. Iaja pa?i tupã nojpotari
Iaŋ-a pa?i tupã na-o-i-potar-i
this-REF father God NEG-3-R₂-want-R₂
 ‘God the father does not want this.’ (Ar., *Cat.*, 102v)

- h. Aβa ra?irape wĩ?
 Aβa r-a?ira=pe wĩ-∅
 person R₁-son=Q **this-REF**
 ‘Whose son is this?’ (Teatro, 48)

- i. Aβápe ajpoβa?e ojmomaran?
 Aβa=pe **ajpo-βa?e** o-i-momaran
 person=Q DEM-NMLZ 3-R₂-disobey
 ‘Who disobeys that one?’ (Araújo, 67)

- j. Ikoβa?e te!
Iko-βa?e te
 DEM-NMLZ INT
 ‘This one (not the other)!’ (VLB, I, 130)

8.2 RP operators

The operators for each layer are associated with a specific semantic domain: the nuclear_R operators express qualitative features of the referent, the core_R operators express quantitative characteristics of the referent, and the RP-level operators locate the referent within the immediate common ground, which

includes the discourse context and the physical environment (Van Valin Jr 2022, 36).

The operators of the RP in TUP are shown in Table (8.1).

Level	Operator type
Nuclear _R	Nominal tense
Core _R	Number, negation
RP	Deixis

Table 8.1: RP levels and their operators

In Section 3.2, it was mentioned that nominal aspect involves the count-mass distinction, which parallels the telic/atelic distinction in verbs (see Jackendoff 1992, 29), that is, whether the referent is an individual, part of an individual, or a set of individuals. TUP does not have classifiers, and the count-mass distinction is not marked morphologically, nor does it have any morphosyntactic implications.

The only nuclear operator (NUC_R) of the TUP RP is nominal tense, which was discussed in Section 7.3.1.

Since grammatical number is not a category of TUP nouns (see Section 7.3.3), the only RP operator at the core level is negation.

8.2.0.1 Adnominal quantifiers

Previous versions of RRG treated quantifiers as operators. Currently, they are considered to be peripheral modifiers of the core. In TUP, some behave as lexical items as well as modifiers, e.g. *jaβi?õ* ‘each’, which can function as the head of an RP (398a) and as a modifier (398b).

- (398) a. *Pejaβi?õ pa?i Tupã karaiβeβe moikow*
pe=∅-jaβi?õ pa?i Tupã karaiβeβe-∅ mo-iko-w
 2PL=R₁-each lord God angel-REF CAUS-be-NFOC
 ‘God the lord assigned each of you an angel.’ (Teatro, 52)

- b. ʔaretewasu jaβiʔõ
 ʔar-ete-wasu **jaβiʔõ**
 day-INTS-AUG **each**
 ‘Each Easter.’ (Araújo, 59v)

RPs may be quantified by overt numerals or general quantifiers in lexical expressions such as *three books, many dogs, few particles, every woman*.

Ancient sources agree on the fact that TUP could count to five (Thevet 1953, 239, Staden 1557, 185, De Léry 1972, 251, d’Evreux 2014, 121), but there were numerals only for one, two, three and four, making it a nearly anumeric language (see Everett 2013, ch. 6). For the number five, the word *po* ‘hand’ was used (see 409b), and the possibility that this was introduced by the Portuguese cannot be excluded¹.

- (399) a. Amõ mokõj mosanga
 amõ **mokõj** p.osaŋ-a
 other **two** R₃-medicine-REF
 ‘The other two medicines.’ (DC, I, 223)
- b. Mosapir tekokatu
mosapir t-eko-katu-∅
three R₃-life-good-REF
 ‘The three virtues.’ (DC, I, 153)

When employed without adjacent nouns, numerals do not take the relational marker:

- (400) a. Ererureta serã? Aani, mosapi jõ
 ere-rur-eta serã aani **mosapi** jõ
 2SG-bring-many PRCL no **three** only
 ‘Did you bring many, by the way? No, only three.’ (AT, 46-48)

TUP numerals may precede or follow the noun they modify (401). This is a rare feature cross-linguistically, as shown by Dryer (2013). Of a sample

¹This is conjecture based on a comparison with numerals in other Tupían languages.

consisting of 1154 languages, only 65 lack a dominant noun-numeral/numeral-noun order. Quantified RPs never combine with *eta*.

- (401) a. Ojepe kujã
ojepe kujã-Ø
one woman-REF
 ‘One woman.’ (AA, 9v)
- b. Kujã ojepe
 kujã-Ø **ojepe**
 woman-REF **one**
 ‘One woman.’ (AA, 9v)
- c. Mokõj ap^jaβa
mokõj ap^jaβ-a
two male-REF
 ‘Two men.’ (AG, 9v)
- d. Ap^jaβa mokõj
 ap^jaβ-a **mokõj**
 male-REF **two**
 ‘Two men.’ (AG, 9v)

When marked by REF, cardinals become ordinals or adverbs:

- (402) a. ?ara mosapira pupe
 ?ar-a **mosapir-a** Ø-pupe
 day-REF **three-REF** R₁-in
 ‘On the third day.’ (Araújo, 15)
- b. Imokõja
 i-**mokõj-a**
 R₂-**two-REF**
 ‘His second time.’ (VLB, II, 115)

When negated, numerals, despite not carrying the referential marker (REF), receive the non-predicative negation (see Section 6.5.2.3):

- (403) Namosapir ruã te tupã!
 na-mosapir ruã te tupã-Ø
 NEG-three NEG FOC God-REF
 ‘Not three Gods, instead!’ (ADC, I, 193)

While one, two, and three are consistent within the TG family, ‘four’ seems less stable. There are different forms for ‘four’ attested in Tupinambá:

- (404) Mokōmokōjsik
 mokō-mokōj-sik
 two-two-in.total
 ‘Four.’ (VLB, I, 154)
- (405) Ojoirunik
 four
 ‘Four.’ (Araújo, 77) ²
- (406) Ojoirūrū
 ojo-irū-irū
 RFLX-companion-companion
 ‘A pair of pairs.’ (VLB, I, 154) ³
- (407) Mojerunik
 four
 ‘Four.’ (FA 14, Bettendorff, 48)⁴

Araújo (1618b) uses two forms of the word as well as the Portuguese word:

- (408) a. Ojoirunik tekosikaβa
ojoriunik t-eko-sikaβ-a
four R₃-fact-last-REF
 ‘The last facts are four.’ (Araújo, 154v)

²Also spelled *ojeirundik*.

³*Ojoirū* means ‘companions of each-other’.

⁴Also spelled *mojerundik*.

- b. Quatro tekokatuitá
Quatro t-ekokatu-ita
four R₃-virtue-column
 ‘The cardinal virtues are four.’ (Araújo, 10)

Numbers other than four require the word for *hand*, *foot* to be expressed. I do agree with Wolf Dietrich (personal communication) that this could well be a Jesuit invention, but since there are many different rare types of numeral systems cross-linguistically, it is more prudent to be categorical (see Hammarström 2010).

- (409) a. Opá ko po mosapir misã ?ara sikeme
 Opa ko po-∅ mosapir misã-∅ ?ara-∅ sik-eme
 all this hand-REF three toe-REF day-REF arrive-POSP
 ‘When the thirteenth day came.’ (Araújo, 3)⁵
- b. jepo, jepi, aβa po, ipi
 je=∅-po je=∅-pi aβa-∅ ∅-po i-pi
 1SG=R₁-**hand** 1SG=R₁-**foot** man-REF R₁-**hand**-REF R₂-foot
 ?ara omemirawera k^waβire...
 ?ar-a o-memira-wer-a k^waβ-ire
 day-REF CORF-son-PST-REF pass-after
 ‘Forty days after the birth of her son had passed...’ (Araújo, 3v)⁶
- c. Opa ko po jabi?õ Tupã supe
 opa ko po-∅ jabi?õ Tupã supe
 all this hand-REF each God DAT
 ‘One (for) each ten to God.’ (Araújo, 78)

Some quantifiers may precede (410a) or follow (410c) the noun or the RP, like *amõ* ‘some, any, a certain, someone, other’ (410) or *opa(β)* ‘all, every’ (411):

⁵All these hands and three toes = 10 + 3.

⁶My hands, my feet, someone’s hands, his feet = 10 + 10 + 10 + 10.

- (410) a. Amõ aβa
amõ aβa
other man
 ‘Other men.’ (Ar, 128)
- b. Tupã amõ kujãkatu mojaŋi
 Tupã-∅ **amõ** kujã-katu mojaŋ-i
 God-REF some woman-good make-NFOC
 ‘God made a certain good woman. (God’s making of a certain good woman)’ (Poemas, 86)
- c. i-ŋupé oapisara amõ meʔeŋa
 i-ŋupé o-apisar-a **amõ** meʔeŋ-a
 R₂-POSP 3.COREF-similar-REF **some** give-GER
 ‘Giving him someone similar to himself.’ (Ar, 72)
- (411) a. Opa aβa jukaw
opa aβa-∅ juka-w
loc man-REF kill-NFOC
 ‘Killed all the men.’ (AG, 54v)
- b. ŋetekokuwaβa opa amokajem
 ŋe=t-ekokuwaβ-a **opa** a-mo-kajem
 1SG=R₂-knowledge-REF complete 1SG-CAUS-disappear
 ‘I made all my understanding disappear.’ (Poemas, 106)

With negation, the meaning of *amõ* is translated by ‘no, none, any’:

- (412) a. Naaruri amõ parati
n-a-rur-i **amõ** parati
 NEG-1SG-bring-NEG **any** parati
 ‘I have not brought any parati (species of fish).’ (Poemas, 154)

For a plural reading, *amõ* may be reduplicated:

- (413) a. Amõamõ santos
amõ-amõ santos
some-RED saints

‘Some saints.’ (Ar, 139 [1686])

- b. Karaiβa amōamō iangajpa
 Karaiβ-a **amō-amō** i-angajpa
 white.man-REF **some-RED** R₂-sin
 ‘Many white men are sinners.’ (Poesias, 55)

Some quantifiers can only follow the noun or pronoun:

- (414) a. ?ara jaβi?ō
 ?ara **jaβi?ō**
 day **each**
 ‘Each day.’ (VLB, I, 62)
- b. Pejaβi?ō pa?i Tupā karajβεβε moikow
 pe=**jaβi?ō** pa?i Tupā-∅ karajβεβε-∅ mo-eko-w
 2PL=**each** holy.man God-REF angel-REF CAUS-be-NFOC
 ‘For each one of you God the father delegated an angel.’ (AT, 50)
- c. Oporandupe Herodes ma?e tetiruã rese isupe?
 o-porandu=pe Herodes ma?e **tetiruã** r-ese i-supe
 3-ask-INT Herodes thing **any** R₂-about R-DAT
 ‘Did Herodes ask him about anything?’ (Araújo, 59)

Amō can also be used pronominally and as the head of an RP, as in (415):

- (415) Mokōj monaβora, i?ekatuaβa koti **amō**, a?e amō
 Mokōj mona-βor-a i-?ekatuaβa ∅-koti **amō**, a?e amō
 two steal-HAB.AG-REF R₂-right.side R₂-POSP PRO DEM other
 iasu koti
 i-asu ∅-koti
 R₂-left.side R₁-POSP
 ‘Two thieves, one on his right side and that other on his left.’ (Araújo, 62v)

Some other quantifiers, like *moβirjō* ‘some, few, not many’ only precede the noun:

- (416) Moβirjõ ipo erimaʔe kunumĩ kajemi
moβirjõ ipo erimaʔe kunumĩ-∅ kajem-i
few ADV once boy-REF disappear-NFOC
 ‘Once, certainly, only a few boys died.’ (AR, 157v)

8.2.0.2 Nominal negation

Negation in the RP denotes the absence of a referent, so it is no different from a quantifier which has a quantity of zero. TUP has a privative morpheme, *-eʔim*, which is the RP negation operator at the core level. This is illustrated in (417):

- (417) a. Poropotareʔima
 poro-potar-**eʔim**-a
 ANTIP-want-PRIV-REF
 ‘Lustlessness (lit. absence of desire for a person).’ (Poemas, 132)
- b. Tureʔimawama
 t-ur-**eʔim**-wam-a
 R₂-come-PRIV-FUT-REF
 ‘Their future not-coming.’ (Teatro, 14)
- c. Sieʔima
 si-**eʔim**-a
 mother-PRIV-REF
 ‘Orphan (lit. motherless).’ (VLB, II, 59)

Figure (8.3) shows the representation of (417c):

When the privative is negated, the meaning is non-negative, as in (418).

- (418) a. Najukaeʔimi
 n-a-i-juka-**eʔim**-i
 NEG-1SG-R₂-kill-**PRIV**-NEG
 ‘I do not not kill him.’ (FA, 34)

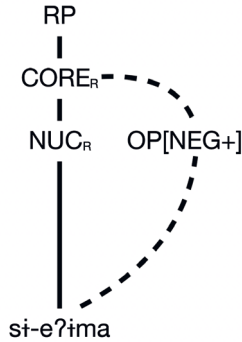


Figure 8.3: Negation as a core-level operator of the RP

- b. Naipotare?imi
 n-a-i-potar-**e?im-i**
 NEG-1SG-R₂-want-PRIV-NEG
 ‘It’s not the case that I do not want it/him.’ (AA, 34v)
- c. Napeamotare?imipe oreruβisaβa?
 n-pe-amotar-**e?im-i**=pe ore=r-uβisaβ-a
 NEG-2PL-hate-PRIV-NEG=Q 1PL.EXCL=R₁-chief-REF
 ‘Don’t you love (not not hate) your chief?’ (Léry, 353)

Thus, the double negation or double privative construction has a positive meaning. In (419), the lexical root *ekate?im* ‘avarice’⁷ is negated by the privative, deriving *ek(o)ate?ime?im* ‘lack of avarice, freedom’. The only attested instance of the double privative is found in Bettendorff (1681, 62), given in (419), whose language is already distinct from the language described by Anchieta (1595) and Figueira (1687).

⁷The lexical root *ekate?im* is possibly a compound of *ekar* ‘seek’ + *e?im* ‘privative’. Even if the etymology of the first element of the compound is uncertain, there is no question regarding the presence of the privative morpheme.

- (419) Tupã miatãeteete, sekoate?ime?imeteete
 tupã miatã-ete-ete s-ekoate?im-e?im-ete-ete
 God strength-INTS-INTS R₁-avarice-PRIV-INTS-INTS
 ‘The great power of God, his great freedom.’ (Bettendorff, 62)

8.2.1 RP operators

Operators that modify the whole RP ground the referent in the ‘real world’. These are related to locality and are similar to clause-level operators in the sentence. They mark the RP for deixis.

8.2.1.1 Demonstratives

Demonstratives are ‘deictic expressions which are used to orient and focus the hearer’s attention on objects or locations in the speech situation’ (Diessel 1999, 2). TUP has three types of demonstratives according to the syntactic context (*ibidem*): (i) adnominal demonstratives (used as modifiers of nouns), (ii) pronominal demonstratives (used as independent pronouns, i.e., as arguments of verbs and adpositions, which are full RPs on their own) (see Section 8.1), (iii) adverbial demonstratives (verb modifiers which are used for the specification of location). This section will only deal with adnominal demonstratives.

TUP adnominal demonstratives encode the following semantic features: distance contrast (proximal, distal), person-orientation contrast (near the speaker, near the hearer, away from the speaker) (see Diessel 1999, 39), and visibility contrast (in sight or not in sight). A list of TUP demonstratives is given in Table 8.2.

From the table above, it is possible to postulate an old, non-analyzable prefix indicating non-visibility, *a-*, and a prefix that indicates proximity to the listener, *eβo-*.

Demonstratives are always free roots that do not require derivational morphology. When followed by a noun, they form a tight constituent (RP), with the demonstrative modifying the noun directly. TUP demonstratives always precede the noun (see Dryer 1992b, 108).

	Near the speaker	Near the listener	Far from the speaker	Far from both
Visible	(i)ko	(eβo)k ^w e(j), (eβo)wĩ(ŋ), e(βo)wĩ, emonā	kwe(j), mō, erik ak ^w e(j), amō,	ajpo, eβapo,
Non-visible	ā, aŋ		awā, apo, anōj	ʔe, aʔe ako

Table 8.2: Demonstratives in TUP

Structurally, demonstratives are hosted in the RPIP (see Van Valin Jr 2005, 26-27), as shown in 420 and its representation in Figure 8.4:

- (420) Ikoʔara
iko ʔar-a
this world-REF
 ‘This world.’ (DC, I, 159)

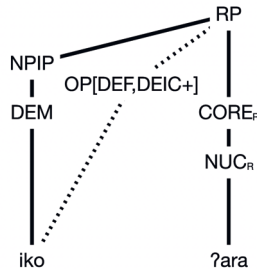


Figure 8.4: A deixis operator

- (421) a. Ikotaβa apamonana
iko taβ-a apamonan-a
this village confuse-GER
 ‘Confusing this village.’ (AT, 42)

- b. Moβipe ase iko mosanga rarine?
 moβi=pe ase **iko m.osang-a** r-ar-i=ne
 how.many=Q PRON this R₄.medicine-REF R-take-NFOC=FUT
 ‘How many times does one take this medicine?’ (AD I, 208)
- c. Ko aʔe iβaka janeremiep^jak^wama oimojaŋ
 ko **aʔe** iβak-a jane=r-emi-ep^jak-wam-a o-i-mojaŋ
 this **that** sky-REF 1PL.INCL=R₁-RES-see-FUT-REF 3-R₂-make
 ‘This one made that sky we will see.’ (Araújo, 86)
- d. Marātepe aŋ maʔekatupaβẽ orowerekóne?
 marã-te=pe **aŋ** maʔe-katu-pabẽ oro-wereko=ne
 how-FOC=Q DEM thing-INTS-all 1PL.EXCL-have=FUT
 ‘But how do we do with these many riches?’ (Araújo, 7)

When something is out of sight, either because it is far away or because it is an abstract entity, the demonstratives encoding non-visibility are used.

- (422) a. Taso ajpo jeʔenga mopo
 t-a-so **ajpo** jeʔeŋ-a mo-po(r)
 HORT-1SG-go **DEM** word-REF CAUS-happen
 ‘May I go fulfill these words.’ (AT, 62)
- b. Taso nepiri kori, ajpo tuβisaβa
 t-a-so ne=∅-piri kori **ajpo** t-uβisaβ-a
 HORT-1SG-go 2SG=R₁-near today **DEM** R³-chief-REF
 waβo
 w-aβo
 3_{CORF}.eat-GER
 ‘May I go to you, today, in order to eat those leaders.’ (AT, 68)
- c. Ajpo jōpipo nerera?
ajpo jō-pe-ipo ne=r-era
DEM only-Q-certainly 2SG=R₂-name
 ‘Is your name only this indeed?’

8.3 Nominalizers

TUP is the only TG language with nine nominalizers, all of which ‘have cognates in at least some TG languages’ (Schleicher 1998, 136). Their functions often depend on the transitivity of the root they combine with or on the semantics of the participants involved. Nominalized lexical roots, like any RP, may combine with casual suffixes, postpositions, or tense markers, and often require that their arguments, actor and undergoer, be encoded as the possessor.

Relativizer

The nominalizer *-βaʔe*, besides being used in equative predication (see Section 5.5), also nominalizes clauses, mostly with intransitive predicates⁸. An example of a nominalized clause as a restrictive modifier is given in (423).

- (423) *iβi opa iβitiŋa iβaka swi*
iβi-∅ opa iβitiŋ-a iβak-a ∅-swi
 earth-REF all clouds-REF sky-REF R₁-from
oʔariβaʔe iasoʔiune
o-ʔar-i-βaʔe iasoʔi-u=ne
 3-fall-EPEN-NMLZ_{REL} cover-NFOC=FUT
 ‘The earth, all the clouds that fall from the sky will cover it.’ (Araújo, 7)

The clause nominalized by the relativizer, when it follows an RP in a detached position (PrDP), i.e., with a pause, this RP is the undergoer of the action in the nominalized predicate, as in (424), with subscripted indices. This RP must not be an independent pronoun of either the first or second person. This type, however, is uncommon in the texts, because *-βaʔe* is far more common with intransitive verbs.

⁸The nominalizing suffix *-βaʔe* also combines with deictics (see Section 8.1) and lexical roots of different semantic categories, such as ‘one’ in *o-jepe-βaʔe* 3-one-NMLZ_{REL} ‘the one who is unique’ (DC, I, 141).

- (424) Pedro, ojúkáβaʔe
 Pedro_i o_j-i_i-júka-βaʔe
 Pedro 3-R₂-kill-NMLZ_{REL}
 ‘As for Pedro_i, he_j is the one who kills him_i.’ (AA, 30v)

-βaʔe can appear in the ECS, either related to a core argument (425) or as the argument of a postposition (426):

- (425) Nojaβiaŋajpe omendariβaʔe Tupā reko
 n-o-i-aβi-aŋajpaβ=pe o-mensar-βaʔe Tupā r-eko
 NEG-3-R₂-infringe-evil=Q 3-marry-NMLZ_{REL} God R₁-law
 ojopotá?
 o-jo-pota
 3-RECP-want.GER
 ‘Don’t those who are married commit sins in desiring each other?’
 (DC, I, 228)

-βaʔe combines with nominal tense, as in (426):

- (426) a. Jerepiramo [omanōβaʔep^wera ri]
 je=r-epi-ramo o-manō-βaʔe-p^wer-a ∅-ri
 1SG=R₁-price-TRSL 3-die-NMLZ_{REL}-PST-REF R₁-POSP
 jeakireʔimamo
 je=∅-akir-eʔim-aβo
 1SG=R₁-soften-PRIV-GER
 ‘Having compassion for the one who died as my saviour.’ (AC., 86)
- b. Omanōβaʔep^wera swi
 o-manō-βaʔe-p^wer-a ∅-swi
 3-die-NMLZ_{REL}-PST-REF R₁-from
 ‘From those who have died.’ (DC, I, 141)
- c. Ako omanōβaʔerameʔima βerameʔi
 ako o-manō-βaʔe-ram-eʔim-a βerameʔi
 that 3-die-NMLZ_{REL}-FUT-NEG-REF seem

Patient nominalizer

Transitive predicates can also be nominalized by *-pir* and designate the under-goer of an event.

- (429) a. Ijukapira
 i-juka-**pir**-a
 R₂-kill-NMLZ_{PAT}-REF
 ‘(The) one who is/must be killed.’ (AA, 19v)
- b. Tapeso pejekosupa
 t-pe-so pe-je-ekosuβ-a
 HORT-2SG.PL-go 2SG.PL-RFLX-be-delight-GER
 ipotaripira ri
 i-potar-**pir**-a Ø-ri
 R₂-want-NMLZ_{PAT}-REF R₁-POSP
 ‘May you go, rejoicing with what is desired.’ (Teatro, 58)
- c. Aβamona morapitjawera
 aβa-mona-Ø mor-apiti-sar-wer-a
 person-thief-REF ANTIP-slaughter-NMLZ-PST-REF
 repiramo muneokipe imoneβipirwera
 r-epi-ramo mune-ok-pe i-moneβ-**pir**-wer-a
 R₁-pay-TRSL prison-house-LOC R₂-arrest-NMLZ_{PAT}-PST-REF
 ‘A thief (who was) put in prison as payment for men’s slaughter.’
 (Araújo, 59v)

Resultative nominalizer

The prefix *emi-* – the only prefix among nominalizers – is functionally similar to the nominalizer *pir*. It ‘derives from a transitive predicate a noun which is the undergoer of the action from which it is derived.’ (Schleicher 1998, 136). It requires the undergoer of the predicate to be interpreted as a possessed noun: the literal translation of (431a) would be ‘killed thing of/by the jaguar’. Some nominals derived through *emi-* probably had been lexicalized so that the prefix in question was not perceived as a derivation, such as the examples in (430). Other instances, such as those in (431), are derived though prefixation.

- (430) a. Temi?u
t-**emi**-?u
R₄-**RES**-ingest
‘Food (lit. ingested thing).’ (VLB, I, 77)
- b. Temireko
t-**emi**-r-eko
R₄-**RES**-R₁-be
‘Wife (lit. made be with).’ (VLB, II, 40)
- c. Mika?u
(e)mi-ka?u
RES-make.soggy
‘Porridge (lit. (what is) made soggy).’ (Staden, 143)
- (431) a. Jawara remijukap^wera
javar-a r-**emi**-juka-p^wer-a
jaguar-REF R₁-**RES**-kill-PST-REF
‘What the jaguar killed / the killed by the jaguar.’ (Araújo, 107v)
- b. Tojemojãjneremimotara
t-o-je-mojãjne=r-emi-potar-a
HORT-3-RFLX-make 2SG=R₁-**RES**-want-REF
‘May your will (what is desired) be made.’ (Araújo, 13v)
- c. ?iβa Tupã remipisirõ
?iβa-∅ Tupã r-**emi**-pisirõ-∅
fruit-REF God R₁-**RES**-prohibit-REF
‘God’s prohibited fruit (that God prohibited).’ (Araújo, 84)

Although some of the examples containing *emi*- have been translated into English using the passive voice, it is not a syntactic passive. Its addition to a stem is a lexical operation.

There seems to be no functional difference between *mi*- and *-pir*, although this matter requires further investigation. While the Tupinambá treebank available on UD (Ferraz Gerardi 2022) is yet to reach a number of sentences that would be necessary for a quantitative analysis of the Tupinambá corpus, it is

possible to note that many lexical roots in fact combine with both of these nominalizers without a difference in meaning, even though certain nuances of meaning seem to be recognizable as Jesuits attempted to translate Christian ideas. Table (8.3) shows some lexical roots attested in combination with both forms, *emi-* and *-pir*.

Root	<i>emi-</i>	Meaning	Attestation	<i>-pir</i>	Attestation
<i>juka</i>	kill	<i>emi-juka</i>	Ar., 107v	<i>juka-pir</i>	Figueira, 8, 32
<i>moete</i>	honor	<i>emi-moete</i>	VLB, II, 87	<i>moete-pir</i>	VLB, II, 87
<i>iko</i>	be, act	<i>emi-(r)eko</i>	Araújo, 96	<i>serekopir</i>	Figueira, 107
<i>ep^jak</i>	see	<i>ep^jak-(i)pir</i>	Léry, 346	<i>emi-ep^jak</i>	Ar., 61
<i>moja?ok</i>	divide	<i>emi.moja?ok</i>	Ar., 162	<i>moja?ok-(i)-pir</i>	Ar., 78v

Table 8.3: Lexical roots with *emi-* and *-pir*

General nominalizer

The general nominalizer *-saβ* combines with active roots, indicating how the event happens, the instrument through which the event is accomplished, the goal of the event, or even the circumstances under which the event takes place. Examples are given in (432).

- (432) a. Ma?e rese ima?enwasaβa
 ma?e r-ese i-ma?enwar-**saβ**-a
 thing R₁-with R₁-remember-**NMLZ**-REF
 ‘The memory of things (lit. way of remembering things).’ (DC, I, 152)
- b. Ima?e potasaβa
 i-ma?e potar-**saβ**-a
 R₂-thing want-**NMLZ**-REF
 ‘Desire of its things.’ (DC, I, 152)
- c. Aseka jepe mitasaβa amō witekoβo
 a-s-eka jepe m.itar-**saβ**-a amō wi-t-eko-βo
 1SG-R₂-search one stay.R₃-**NMLZ**-REF other 1SG-EPEN-be-GER

‘I am looking for an inn (lit. place of staying).’ (Teatro, 128)

d. $\text{je nemoiko}\beta\text{osa}\beta\text{a}$

$\text{je ne}=\emptyset\text{-mo-iko-}\beta\text{o-}\mathbf{sa}\beta\text{-a}$

I 2SG=R₁-CAUS-be-GER-NMLZ-REF

‘I am the cause of your action.’ (Teatro, 176)

e. $\text{Seta j}\ddot{\text{e}} \text{ i}\text{?asa}\beta\text{usu}$

s-eta $\text{j}\ddot{\text{e}} \text{ i}\text{?a-}\mathbf{sa}\beta\text{-usu-REF}$

R₂-many already take.water-NMLZ-AUG-REF

‘The bows are many already (lit. instrument of taking water).’
(Teatro, 26)

Habitual agent

The habitual agent is expressed by *-bor*, as exemplified in (433). This suffix can express a habit (433a), or a constant state (433b).

(433) a. $\text{A}\eta\text{a ja a}\eta\text{ajpa}\beta\text{ora ajuka}$
 $\text{a}\eta\text{a ja a}\eta\text{ajpa}\beta\text{-}\mathbf{\beta\text{or-a}} \text{ a-i-juka}$
 this like sin-HAB.AG-REF 1SG-R₂-kill

‘As (with) these, I kill the sinners.’ (Poemas, 94)

b. $\text{Eremomuejra mara}\text{?a}\beta\text{ora}$
 $\text{ere-mo-muejra}\beta \text{ mara}\text{?a-}\mathbf{\beta\text{or-a}}$
 2SG-CAU-heal sick-HAB.AG-REF

‘You healed the sick.’ (Poemas, 122)

Gerund

For the gerund, see Section 10.2.3.

8.4 Modifier Phrase

Lexical roots, demonstratives, and nominalizers may modify an RP as well as a sentence with a PrCS. Like the RP and the predicative PP, the MP has a nucleus and a core, but no operators. MPs are peripheral modifiers at the nuclear level (degree modifiers) and the core level (manner adverbial modifiers).

The simplest type of modification is composition (Rodrigues 1951a). The morphophonemic processes present in composition (see 436) are clearly distinct from that of apposition/juxtaposition (see Section 8.5). While the juxtaposition of lexemes for a possessive construction requires the mediation of a relational morpheme (R_1) to signal dependency, a compound does not require a relational and the REF appears only once, with its scope over the entire compound, as in (434 and (254). As shown in Figure (8.5), the lexical modifier is in the nuclear periphery of the NUC_R . Figure (8.6) represents (434d), where a modified lexical root modifies another lexical root.

- (434) a. Aoβuna
 aoβ-un-a
 clothes-black-REF
 ‘Black clothe(s).’ (VLB, II, 86)
- b. ʔaŋaporaja
 ʔaŋa-poraŋ-a
 soul-beauty-REF
 ‘Beautiful soul.’ (Poemas, 140)
- c. Piraakāpuku
 pira-aka(ŋ)-puku-∅
 fish-head-long
 ‘Long headed fish.’ (VLB, I, 50) (*Bagre pinnimaculatus*)
- d. Pirajurumemeka
 pira-juru-memek-a
 fish-mouth-soft-REF
 ‘Jamaica weakfish.’ (Marcgrave, 149)

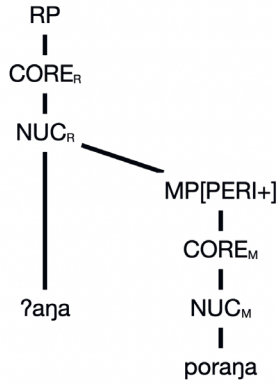


Figure 8.5: Modification by composition in TUP

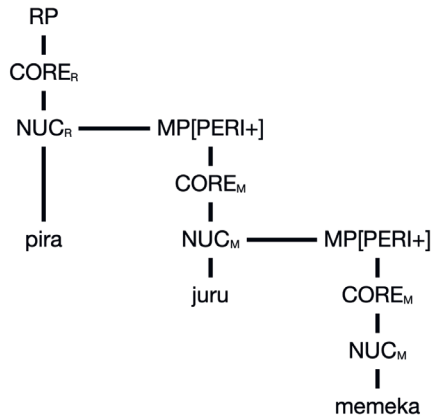


Figure 8.6: Recursive modification by composition in TUP

The following example has an action-word as modifier.

- (435) a. Piraβeβe
 piraβeβe-∅
 fish-fly-REF
 ‘Flying fish.’ (Marcgrave, 162; VLB, II, 70; 147)

- b. Tejujan
 teju-jan-a
 lizzard-run-REF
 ‘Running lizard.’ (Marcgrave, 238)

In compound modification, as in (434), the semantic head of the new lexeme is always the leftmost element, even if the composition is made up of more than two elements, as in (434c).

- (436) Pisaipeβa
 pisa-i(βa)-peβ-a
 fish.net-grip-flat-REF
 ‘Flat-grip-fishnet.’ (VLB, II, 99)

Most compositions are like (434a) and (436) in that the head of the compound is the leftmost element. Nonetheless, there are ‘compounds’ that seem to be headed by the rightmost element, as in (437). These are probably cases of grammaticized possessive expressions which lost the contiguity marker. Such cases are called ‘determinative compounds’ by Rodrigues (1951a). Note that the contiguity marker (R_1) is missing, for example, in (437a) (*r-up^jara*). The translation given to (437c) is ‘fish oil’. This translation would be correct if the contiguity marker were present: pira-REF R_1 -jani-REF. One would expect its meaning to be ‘oily fish’ if no grammaticalization process were involved.

- (437) a. Ajurujuβup^jar
 ajuru-juβ-up^jar
 parrot-yellow-enemy
 ‘Yellow parrot enemy (Frenchman).’ (cf. AT, 44)⁹
- b. Ataeniuru
 at-eni-uru
 fire-light-container

⁹‘Yellow parrot’ was the expression used to refer to the French (see VLB, I, 143) or other people of European descent (see also Marcgrave and Piso 1648, 268).

‘Fire light container (lamp).’ (VLB, I, 65)

- c. Pirajani
 pira-jani
 fish-oil

‘Fish oil / oil of fish.’ (VLB, I, 49)

Further examples are given in 438:

(438) a. Uʔuβuru

uʔuβ-uru

arrow-case

‘Arrow case.’ (VLB, I, 49)

b. Wirajeʔeηatu

wira-jeʔeη-katu

bird-sing-beautiful

‘Saffron finch (*Sicalis flaveola*).’ (Marcgrave, 211)

c. Meni

Men-si-∅

husband-mother-REF

‘Mother-in-law.’¹⁰ (Araújo, 115)

RPs containing a dependent noun which modifies the head noun may have different semantic functions besides that of possessor. (439) exemplifies this fact with agent (a), theme (b), patient (c), but other roles may also be encoded. Since these are possessive constructions, the possessor is in the RPIP, as shown in Figure 8.7 representing (439b).

¹⁰This example may be compared with the following example:

- (1) Mena si
 Men-a ∅-si-∅
 husband-REF R-mother-REF
 ‘Mother of the husband.’

which is not a compound, because there is a syntactic relation.

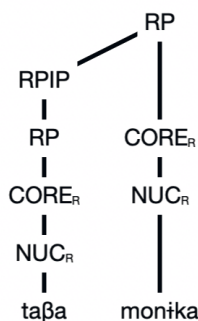


Figure 8.7: RP possessor in the RPIP

- (439) a. Sumarã puʔama
 sumarã-∅ ∅-puʔama-REF
 enemy-REF R-assault-∅
 ‘The enemy’s assault.’ (Poemas, 184)
- b. Taβa monika
 taβ-a ∅-monik-a
 village-REF R-destroy-REF
 ‘The village’s destruction.’ (Teatro, 12)
- c. Christo rawsuβa
 Christo r-awsuβ-a
 Christ R₁-love-REF
 ‘The love of Christ (one’s love of Christ)’ (Araújo, 161v)

Depending on the transitivity of a root, it is possible to embed an RP containing an RPIP in another RP with its own RPIP. The verb ‘to love’ takes two arguments, so the nominalized phrase ‘João’s love of Pedro’ *Joao Pedro rawsuβa* consists of two genitives [[the love [of João]] [of Pedro]]. In this case, the noun closest to the rightmost NUC_R is always the argument which is lower on the actor-undergoer hierarchy (see Figure 3.16), in parallel to the SOV orders of cores in clauses. The representation of (440) is given in Figure 8.8.

- (440) João Pedro sawsuβ
 João Pedro s-awsuβ
 João Pedro R-love-REF
 ‘João loves Pedro (lit. João’s love of/for Pedro).’ (AA, 16v)

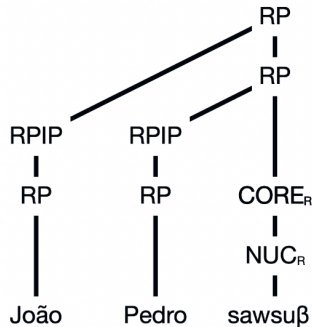


Figure 8.8: Embedded possessor RP

8.4.0.1 Adverbial demonstratives

Demonstratives may function as adverbs, as in (441).

- (441) a. Aso iko!
 a-so **iko**
 1sg-go DEM
 ‘Behold, I go!’ (Anch., *Arte*, 21v)
- b. Iesus Nazareno iko orosekar
 Iesus Nazareno **iko** oro=s-ekar
 Jesus of.Nazareth DEM 1PL_{EXCL}=R-search
 ‘Behold, we are searching Jesus of Nazareth.’ (Ar., *Cat.*, 54v)
- c. Ko s-eko-w ko
ko s-eko-w **ko**
 DEM R-be-NFOC DEM
 ‘Behold, it/(s)he is here.’ (VLB, I, 109)

- d. Eβok^we r-upi e-kuwaβ!
eβok^we r-upi e-kuwaβ
 DEM REL-POSP 2sg.imp-pass
 ‘Pass through here!’ (VLB, II, 81)
- e. **Ajpo** jê!
 DEM PRCL
 ‘(T)here it is!’ (Léry, *Hist.*, 353)

A noun can also be formed through the derivation of more than one morpheme. There are many such morphemes, which often become lexicalized. The following examples show the suffix *-usu* ‘big’ (442) and *-p^wera* ‘nominal past’ (324):

- (442) igarusu
 igar-usu
 canoe-big
 ‘Ship (big canoe).’ (see ADC, I, 212)

In the case of lexicalization, the composition forms a new noun, so that there is no modifier, but nominal tenses are operators functioning at the RP level, as can be seen from (443) represented in Figure 8.9:

- (443) Rekop^wera
 r-eko-p^wer-a
 R₁-law-PAST_N-REF
 ‘The old law.’ (Poemas, 104)

- (444) a. Aik^waβ maʔe
 a-i-k^waβ maʔe-∅
 1SG-R₂-know thing-REF
 ‘I know things.’ (Fig., 122)

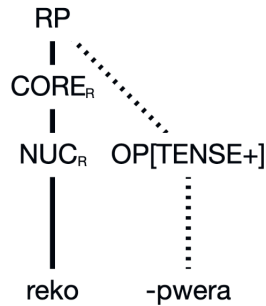


Figure 8.9: Nominal tense operator

8.5 Attributive possession

Possessive relations involve two entities: the possessor (dependent) and the possessed entity (head). Although there are three types of expression of possession (McGregor 2009, 2), this section will only deal with *attributive possession*, in which the possessed and the possessor form an RP, as in *Dave's car* or *his car* (Chappell and McGregor 1996).

Tupinambá nouns are divided into non-possessed and possessed. Non-possessed nouns include, for example, animals, trees, and non-cultivated plants. Possessed nouns include, for example, parts of a whole, attributes and members of a system of relations, tools, and cultivated plants. Possessed nouns are further divided into two categories (Chappell and McGregor 1996). The first category is that of *inalienable nouns*, or obligatorily possessed nouns, which require an overt statement since one of the elements, the head, is semantically incomplete, because it is relational (see Lehmann 1985). In the expression *John's father*, for example, 'father' would obligatorily require an adnominal possessor in TUP. The second category is that of *alienable nouns*, which are not obligatorily possessed nouns and may stand on their own without the specification of a possessor, such as 'shoe' in *Paul's shoe* (see Chappell and McGregor 1996; Velázquez-Castillo 1996). Non-possessed nouns cannot occur in possessive constructions, mainly because the cultural reality reflected in the

grammar¹¹ does not allow this.

Inalienable and alienable possession show no structural differences. They are formed by a possessor either expressed by an independently coded noun (445a,b) or by a pronominal proclitic (Set II in Table 4.3) (445c,d) always followed by the possessed noun. In both cases, the head is obligatorily marked by the relational morpheme(R). In the following examples (445a, b and c), the heads are inalienably possessed, while in (445d) the head is alienably possessed. This is atypical in terms of the parameters in Nichols (1988), according to which inalienability is associated with head marking or non-marking, whilst alienability is typically associated with dependent marking.

- (445) a. Wira raβa
 wira-∅ r-aβ-a
 bird-REF R₁-feather-REF
 ‘Feather of bird / bird’s feather.’ (FA, 71)
- b. Tupã raira
 Tupã r-air-a
 God R-son-REF
 ‘Son of God.’ (AT, 242)
- c. Neruβa
 ne=r-uβ-a
 2SG=R-father-REF
 ‘Your father.’ (Ar., Cat., 100v)
- d. Jepina
 Je=∅-pina-∅
 1SG=R-fish.hook-REF
 ‘My fishhook.’ (Anch., P, 152)

Compare the examples in 446: ‘God the son (one of the persons of the trinity)’ is a *classification* in the terminology of Chappell and McGregor (1989)¹². In

¹¹For the relationship between grammar and culture, see Everett (2012).

¹²Although this kind of apposition looks like modification by property, it is different from ‘son of God’, a possessive construction

(446a), the R *t*- indicates that the dependent is generic and human, while in (446b) *r*- indicates the contiguity of the dependent and the head noun (see Cabral and da Costa 2004, 8). The order of constituents in (446) is also relevant. In (446a) the element on the left side is the head, while in (446b) it is the element on the right:

- (446) a. Tupã taira
 Tupã t-air-a
 God R-son-REF
 ‘God the son.’ (ADC I, 134)
- b. Tupã raira
 Tupã r-air-a
 God R-son-REF
 ‘Son of God.’ (AT, 242).

Semantically, nouns in a possessive relation are referential, while nouns used to classify refer to a type or class in a classificatory construction. The referent noun in this construction is the head noun. One may also note that while the possessor in a possessive relation may be realized as a pronoun, this is impossible for classifying nouns because they lack referentiality.

The examples in (445) show that a possessive construction is either of the type

([N-REF $\overbrace{\text{R-N-REF}}^{\text{head}}$] or [proclitic= $\overbrace{\text{R-N-REF}}^{\text{head}}$]). It is also possible to embed an RP inside another RP. The structure of (447b) is $\underbrace{[[\text{N-REF R-N-REF}]_{\text{RP}} \overbrace{\text{R-N-REF}}^{\text{head}}]}$.

The same expansion can be used to generate (447c). The TUP corpus, however, does not attest to an embedding higher than depth 3, similar to (447b) (see Verhoeven and Lehmann 2018), but apparently there could be, though unnatural, a longer RP of this kind.

- (447) a. Aβá raira
 aβa-REF r-air-a
 man-REF R₁-son-REF
 ‘The man’s son.’ (Teatro, 50)

- b. Aβá raira rura
 aβa-REF R₁-air-a R₁-ur-a
 man-REF R-son-REF R-come-REF
 ‘The man’s son’s arrival.’ (not attested)
- c. Aβá raira rura ?ara
 aβa-REF r-air-a r-ur-a Ø-?ar-a
 man-REF R₁-son-REF R₁-come-REF R₁-day-REF
 ‘The day of the man’s son’s arrival.’ (not attested)

Mbyá Guaraní, a language related to TUP which has the same structure of possessive constructions, attests an example (Vieira 2018, 182) which, if translated into TUP, would yield (448) (I see no reason why this would be ungrammatical):

(448) Mbyá Guaraní (Tupí-Guaraní)

Ara ruβa irũ raira roka
 Ara r-uβa Ø-irũ- r-aira r-oka
 Ara R-father R₁-friend R₁-son R₁-house

‘Ara’s father’s friend’s son’s house.’ (Vieira 2018, 182)

As is common in many languages (Bickel and Nichols 2013), non-possessed nouns can be made grammatically (not semantically) possessed through the apposition of an abstract generic noun which is (can be) marked for possession. In TUP, animal names cannot take the usual head-marked possessive inflection, i.e., one cannot possess a cow unless the word *ejmβaβ* ‘domestic animal, pet, breeding animal’ is employed:

- (449) a. *ʃetapi?ira
 ʃe=tapi?ir-a
 1SG=cow-REF
 ‘?’
- b. ʃerejmβaβa tapi?ira
 ʃe=r-ejmβaβa tapi?ir-a
 1SG=R-pet-REF cow-REF

'My (domestic animal) cow.' (AG, 14v)

Chapter 9

Information Structure

9.1 Introduction

For ease of exposition, some typeface conventions are employed to represent properties of information structures. If a word (or phrase) bears the accent/intonation responsible for conveying focus, it is marked in SMALL CAPS; topic is signaled by **boldface**. For example, in the English Q/A pair in (450), *dog* and *Kim* bear the A and B accents (Jackendoff 1972), respectively, and the focus *the dog* (with the A-accent) conveys is projected to *chased the dog*.

- (450) Q: What about Kim? What did Kim do?
A: **Kim** chased the DOG.

In (450), the presupposition is that *Kim* is available as a topic for comment, i.e., the assertion that Kim *chased the dog* (see Lambrecht 1994, 226).

Information structure can be formally manifested in aspects of prosody, morphology (in the form of special grammatical markers), syntax (in particular nominal), word or constituent order (through displacement), clefting, through

the use of complex grammatical constructions, and in certain choices between related lexical items. The types of texts which make up the TUP corpus limit the extent to which information structure can be described.

Prosodic features cannot, for obvious reasons, be recovered. The homogeneity of the texts, mostly consisting of indoctrination, poses a challenge for understanding information structure through word order in TUP. For instance, Anchieta's theater plays and poems need to be approached carefully since they are rather artificial in their word order, and most of what is left are texts of catechetical nature. For this reason, information structure in TUP is mostly perceived through morphology and, to a limited extent, through word order variation.¹

The analysis proposed in this chapter is based on the theory of information structure developed by Lambrecht (1986, 1987, 1994, 2000). In this approach, three categories are fundamental: (i) PRESUPPOSITION and ASSERTION, relating to the structuring of propositions into portions which a speaker assumes an addressee already knows or does not yet know (see Lambrecht 1994, 52); (ii) IDENTIFIABILITY and ACTIVATION, relating to a speaker's assumptions about the status of the mental representations of discourse referents in the addressee's mind at the time of an utterance (see Van Valin Jr and LaPolla 1997, 199-201); and (iii) TOPIC and FOCUS, relating to a speaker's assessment of the relative predictability vs. unpredictability of the relations between propositions and their elements in given discourse situations (see Lambrecht 1994, 6).

RRG posits two basic types of focus structure:

- **Predicate focus structure:** *Sentence construction expressing a pragmatically structured proposition in which the PSA is a topic (hence within the presupposition) and in which the predicate expresses new information about this topic. The focus domain is the predicate phrase*

¹The comparison of TUP with other TG languages could be helpful in perhaps recovering aspects of information structure in TUP. Nonetheless, as far as I am concerned, the gap in the treatment of information structure in TG languages in general or in specific languages urgently needs to be filled. Descriptions of TG languages in general do not devote space to this subject.

(or part of it).²

- **Sentence focus structure:** *Sentence construction formally marked as expressing a pragmatically structured proposition in which both the PSA and the predicate are in focus. The focus domain is the sentence, minus any topical non-PSA arguments* (Lambrecht 2000, 617)

9.2 Predicate focus

Some examples of predicate focus are given below. Example (451) is an answer to the question ‘What did he do before dying?’ The referent of ‘he’ is available as a topic for comment, which is the new information (focus) provided by the assertion ‘ate with his disciples’. The focus domain is the predicate plus the remaining post-verbal core constituents. Figure 9.1 represents 451, showing the focus projection, where the triangle represents the actual focus domain, the part of the sentence that is actually in focus, and the dotted line represents the potential focus domain, i.e., the syntactic domain in which focus elements may occur.

- (451) Omaʔeʔu oemimoʔeeta piʔi
 o-maʔe-ʔu o-emi-moʔe-eta-∅ piʔi
 3-thing-ingest 3_{CORF}-NMLZ-disciple-many-REF together
 ‘He ate together with his disciples.’ (Araújo, 52)

9.3 Sentence focus

Sentence focus constructions have an entire sentence as a focus domain. Unlike in predicate focus, there is no presupposed topic; that is, the subject or referent in a sentence is not the topic, as new participants are introduced and the proposition expressed is, therefore, not a comment on the introduced topic (see Lambrecht 2000). Put another way, the referent and the proposition do not have a topic-comment relationship, since the utterance is not about the

²This definition is from Lambrecht (2000, 616), except for PSA which is used in the place of ‘subject’.

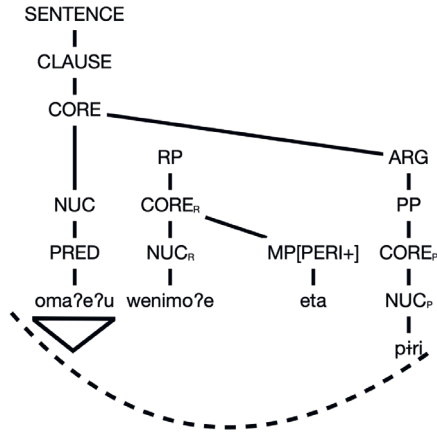


Figure 9.1: Predicate focus in TUP

coded referent. The referents introduced are all new information, just like anything else that may follow. Sentence focus constructions further lack pragmatic presuppositions, except for the non-distinctive presuppositions common in all focus types (Van Valin Jr and LaPolla 1997, 207).

Sentence focus is mainly found in presentational constructions such as narrative openings like *Once upon a time ...* and *There came a man ...*, and most often occurs with intransitive verbs (Lambrecht 2000, 617).

Tupinambá does not seem to have a special presentational construction, as far as attested by the texts, so we cannot know which forms were typically used to begin a story, but we can imagine a context in which sentence focus would occur using concrete examples. If example (452) occurred in a dialogue as an answer to the question *What is happening there?*, there would be no pragmatic presupposition in the above information structure as the assertion and focus are identical and the focus domain is on the whole clause.

- (452) O?ar so?o munépe
 o-?ar s-o?o-∅ mune-pe
 3-fall R₄-flesh-REF trap-LOC

‘Hunt fell into the trap/bag.’ (VLB, I, 63)

Example 452 does not have a topic. Its information structure is shown below:

Sentence: oʔar soʔo munépe

Presupposition: none

Assertion: oʔar soʔo munépe

Focus: oʔar soʔo munépe

Focus domain: CLAUSE

Other examples in which topics are not available are so-calledthetic/existential sentences, such as *There is somebody at the door*. This sentence carries all new information and thus there is no topic, as in (453).

- (453) Noikoj aβa nejaβe
 na-o-i-ko-i aβa-∅ ne=∅-jaβe
 NEG-3-R₂-NEG person-REF 2SG=R₁-like
 ‘There isn’t a person like you.’ (Poemas, 140)

9.4 Focus positions in the syntax

In this section we look at the PrDP, the PrCS and the pre-verbal slot as positions of focus or not.

9.4.1 The Pre-Detached Position

The PrDP (see Section 6.1) is a position outside the potential focus domain. In (454), for example, the RP *ise* ‘I’ is a clause-external topic in the PrDP and *a-* ‘1SG’ is the resumptive index in the clause. The potential focus domain is the clause *aporomoiño kaʔu rese jepi* ‘I make people be in drunkenness always’ and the actual focus domain in this case coincides with the predication *poromoiño kaʔu rese jepi*³. The dislocated topicalized pronoun would otherwise be in the ECS.

³In the poems by Anchieta (Anchieta 1997), it is often the case that he moves constituents around, changing the order of the sentences in order to maintain a specific meter and to obtain

- (454) Ise ko, kaʔu rese aporomoijo jepi
ise ko kaʔu-∅ r-ese a-poro-mo-iko jepi
 I behold drunkenness-REF R₁-POSP ISG-ANTIP-CAUS-be always
 ‘As for me, I make people permanently live in drunkenness.’ (Teatro, 136)

Another example of a (dislocated) topic in the PrDP is given in (455).

- (455) a. Aʔe rako, iaŋajpa
 aʔe rako i-aŋajpaβ
 this EV_{FH} R₂-evil
 ‘They (these old slave women), they are evil.’ (Teatro, 16)
- b. Ene, Tupã roripápe awjerama jẽ erejko
 ene Tupã r-oriβ-saβ-pe awjerama jẽ ere-iko
 you God R₁-happiness-NMLZ-LOC for.ever PRCL 2SG-be
 ‘You, you already are in the glory of God for eternity.’ (Teatro, 124)

9.4.2 Pre-core slot

For Van Valin Jr and LaPolla (1997, 228), the default interpretation of elements in the PrCS is focal, and this is the obligatory interpretation if they are WH-words. WH-words always comprise a particular type of argument-focus construction. Before proceeding to the examples, Table 9.1 shows WH-words in TUP. Some words can be reduplicated to indicate plural number e.g. (456a), and most, if not all of them may occur with and without the interrogative clitic =*pe*.

WH-words in Tupinambá are always focal in a WH-question and always show up in the PrCS. Some examples are given in (456):

the desired rhymes. These are probably cases of artificial constructions, since they are infrequent in the other texts – except for his theater plays (Anchieta 2006). The sentence in 454 is odd in that the PP *kaʔu rese* ‘in drunkenness’ appears pre-core. The normal, expected order would have the PP post-core.

WH-word	Meaning	Colexification	Attested in
Maʔe	what, which	thing	DC, I, 133
Aβa	who	person	
Eri	when		
Mamō	where		
Manōj	where from		VLB, I, 106
Marā	how, why		DC, I, 133
Marāmo	why	how + translative	
Marānamo	why	how + translative	
Marāŋatu	how + intensivizer		
Marāŋatuete	why	how + intensivizer + good	
Maʔete	what	what + focus	
Mobi	how many	some, few	DC, I, 133
Mojrā	when		
Monomo	how many		
Nāβo / (na)nāmo		how many	
Umā	where		Teatro, 130
Umāβaʔe	which		DC, I, 158
Umāme	where	DC, I, 180	
amāme	where		VLB, II, 57
Marāeteʔĩ	how	how + good	

Table 9.1: WH-words in Tupinambá

- (456) a. Maʔemaʔepe ajaŋa ojpotar?
 maʔe-maʔe=pe ajaŋ-a o-i-potar
 WH-WH=Q devil-REF 3-R₂-want
 ‘Which things does the devil want?’ (Araújo, 27v)
- b. Erimaʔepe erejur?
 erimaʔe=pe ere-jur
 WH=Q 2SG-come
 ‘When did you come?’ (FA, 166)
- c. Umāmepe seków?
 umāme=pe s-eko-w
 WH-Q R₂-be-NFOC

- ‘Where does he live?’ (DC, I, 180)
- d. Marāmarāpe santíssima trindade rera?
 marā-marā=pe santíssima trindade r-er-a
 WH-WH=Q holy trinity R₁-name-REF
 ‘What are the names of the Holy Trinity?’ (DC, I, 157)
- e. Ma?etepe peseka ko feretama pupe?
 ma?e-te=pe pe-s-eka ko fe=r-etam-a pupe
 WH-FOC=Q 2PL-R₂-search DEM 1SG=R₁-country-REF POSP
 ‘What then are you looking for in this country of mine?’ (Teatro, 28)
- f. Ma?epe ereru nekaramemuã pupe?
 ma?e=pe ere-er-ur ne=∅-karamemuã-∅ pupe
 WH=Q 2SG-SCAU-come 2SG=R₁-box-REF LOC
 ‘What have you brought in your box?’ (Léry, 342-343)
- g. Ma?epe ojonon iakaŋa ?ariβo?
 ma?e=pe o-jo-non i-akaŋ-a ?ar-βo
 WH=Q 3-R₂-put R₂-head-REF superior.part-PERL
 ‘What did they place around his head?’ (AC, 60v)

The interrogative clitic =pe may attach to the focalized element (narrow focus) instead of attaching to the WH-word:

- (457) a. Ma?e apiaβapajpo?
 ma?e apiaβ-a=pe ajpo
 WHAT nativeR₁=Q that
 ‘Which Indians are those?’ (Teatro, 142, 2006)
- b. Moβi ma?e resepe ase jerurew?
 moβi ma?e-∅ r-ese=pe ase jerure-w
 how.many thing-REF R₁-POSP=Q we ask-NFOC
 ‘For how many things do we ask?’ (Araújo, 26)

Sometimes =pe is omitted.

- (458) Aβa serā oeru?
 aβa serā o-er-ur
 WH by.the.way 3-CAUS.SOC-come
 ‘Who by the way has brought it?’ (Teatro, 6)

9.5 Narrow focus

The question clitic =*pe* (interrogative illocutionary force) follows narrow focus,. Thus, it can follow any fronted constituent (see Lambrecht 1994, 221-238). It follows the verbal predicate in (459a) and the nominal predicate in (459b). In (459c), it follows the pronoun in the PrDP⁴ and in (459d), it is placed after the time adverb:

- (459) a. Osap^jápe Pilatos ije?eηa a?éreme ko?ite?
 o-s-ap^jak=pe Pilatos i-je?eη-a a?ereme ko?ite
 3-R₂-obey-Q Pilatos R₂-speech-REF afterwards finally
 ‘Did Pilatos then finally obey their words?’ (Araújo, 61)
- b. Turusukatupe a?e cruz erima?e?
 t-urusu-katu=**pe** a?e cruz erima?e
 R₂-big-INTS=Q DEM cross ADV
 ‘Was his cross REALLY BIG?’ (Araújo, 61v)
- c. **jepe** asóne?
 je=pe a-so=ne
 1SG=Q 1SG-go=FUT
 ‘Is it I who will go?’ (FA, 166)
- d. Oiaοβok rajēpe iβia
 o-i-aoβ-ok rajē=pe iβia
 3-R₂-cloth-take.off ADV=Q PART
 ‘BEFORE IT his clothes were removed.’ (Araújo, 61v)

It seems that there is a restriction on question formation according to which the element questioned (the question word in a simple, direct WH-question

⁴Due to its topical status, i.e., its association with a specific discourse function, the RP coindexed with the actor argument cannot be in the ECS.

or the focal NP in a simple, direct yes-no question) must function in a clause which is within the potential focus domain of the sentence (Van Valin Jr. et al. 1996).

9.6 Indicative II or oblique-topicalized?

Many TG languages have a construction in which fronted adverbials (adverbs or postpositional phrases) trigger a change in the verbal morphology. Rodrigues (1953); Almeida et al. (1983); Praça (2001) called this construction ‘indicative II’. Rodrigues later abandoned this terminology and called the construction ‘circumstantial mood’ Rodrigues (1996a); Praça (1999); Seki (2000). Jensen (1999); Harrison (1986); Vieira (2014) refer to this construction as ‘oblique-topicalized’. I consider the former term inappropriate because the construction has nothing to do with mood (see Bybee et al. 1994, 176-181). Regarding topic, the construction indeed extracts the adverbial constituent into a more discourse prominent position, but the function of the extracted unit is that of focus. The focal status of the fronted adverbial constituents had already been suggested by Dobson 2005 and Vieira 2014. This construction is thus a type of extraction, because extraction is normally restricted to the informational focus of the proposition, i.e., extracted phrases correspond to the informational focus of the utterance (see Van Valin Jr. 1986; Van Valin Jr. et al. 1996; Goldberg 2013), and the adverbial occurs in a position different from its canonical position in a declarative sentence.

The presence of fronted adverbials in TUP in the PrCS triggers the nominalization of the main predicate, indicated by possessor indexes (Set II in Table 4.3) with the addition of the suffix *-i* following a consonant \sim *-w* following a vowel. In Tupinambá this construction is possible only with first and third person⁵, as in (460). Note that the gloss of the nominalized predicate is NFOC because the focus is on the fronted adverbial expression⁶. The translation

⁵The phenomenon is unevenly distributed among TG languages. In Guajá, Tapirapé, and Kamajurá the construction is only possible with third person, while in Parintintin, as in Tupinambá, only with first and third person. In Mbyá, the fronted adverbial is usually followed by a tense/aspect or modal particle (see Dooley 2015, 66). The construction has been lost in other languages, such as Tekó (Emérillon).

⁶The same interpretation of this construction, i.e., that the fronted adverbial expression as focal,

implies a cleft-like semantic structure with narrow contrastive focus on the fronted adverbials.⁷

- (460) a. Kori ijukaw
 kori i-juka-**w**
 yesterday R₁-kill-**NFOC**
 ‘It was yesterday that he killed it (his killing of it).’ (Arte, 39v)
- b. Koromō jekajemi
 Koromō je=∅-kajem-**i**
 soon 1SG=R₁-flee-**NFOC**
 ‘It is soon that I (shall) run away.’ (AA, 39v)
- c. Janekajemire Tupā amō kujāṅatu
 jane=kajem-ire Tupā amō kujā-ṅatu-∅
 1PL.INCL=loose.oneself-POSP God certain woman-good-REF
 mojaṅi
 mojaṅ-**i**
 make-**NFOC**
 ‘It was after we lost ourselves that God made a good woman.’ (AP, 86)
- d. K^wese Pedro nerese imaṅenwari
 k^wese Pedro ne=r-ese i-maṅenwar-**i**
 yesterday Pedro 2SG=R₁-POSP R₁-remember-**NFOC**
 ‘It was yesterday that Pedro remembered you.’ (Fig., 94)
- e. K^wese nerese Pedro imaṅenwari
 k^wese ne=r-ese Pedro ∅-maṅenwar-**i**
 yesterday 2SG=R₁-POSP Pedro R₁-remember-**NFOC**
 ‘It was yesterday that Pedro remembered you.’ (Fig., 94)
- f. Pedro okope sekow
 Pedro o-ko=pe s-eko-**w**
 Pedro 3_{CORF}-slash-POSP R₂-be-**NFOC**

can be found in Magalhães and Alves (2022).

⁷This type of contrastive focus expressed by fronted adverbials, adjuncts, or oblique arguments has been discussed in RRG terms for Tagalog in Latrouite and Van Valin Jr (2021).

‘It is in his own slash that Pedro is.’ (Fig., 84)

The association of this construction with focus is supported by the fact that questioning an adverbial expression triggers the nominalization and the non-focal suffix, i.e., a WH-word or a constituent in the scope of the focal clitic =*pe*, which are always focal and block the predicate from being the focalized constituent. Some examples are given in 461:

- (461) a. Mamõpe isow oma?e?upaßire?
 mamõ=pe i-so-w o-ma?e-?u-paß-rire
 where=Q R₂-go-NFOC 3_{CORF}-thing-ingest-finish-after
 ‘Where did he go after he finished eating?’ (Araújo, 52v)
- b. Marãtepe ase rekow oeõriréne?
 marã-te=pe ase r-eko-w o-eõ-rire=ne
 how-FOC=Q we R₁-be-NFOC 3_{CORF}-death-after=FUT
 ‘How will we be after his death?’ (DC, I, 161)
- c. Mamõpe Tupã rekow?
 mamõ=pe Tupã r-eko-w
 where=Q God R₁-be-NFOC
 ‘Where is God?’ (DC, I, 158)
- d. Opakatúpe Tupã ase pi?apenwara tirwã repiáki?
 opa-katu=pe Tupã ase pi?a-penwara tirwã r-epiak-**i**
 all-good=Q God we heart-NMZL_{CIRC} even R₁-see-NFOC
 ‘God sees everything, even what is in our hearts?’ (DC, I, 158)
- e. Marãmo ahẽ rekow oma?ekatúramo
 marãmo ahẽ r-eko-w o-ma?e-katu-ramo
 why DEM R₁-be-NFOC 3_{CORF}-thing-good-TRSL
 jeswi?
 je=Ø-swi
 1SG=R₁-from
 ‘Why does this one have more good things than me?’ (Araújo, 109v)

The nominalization of the verb is reflected by the person markers of Set I (see

Table 4.3), creating a genitive construction ([N R₁-N]). Examples (460c) and (460d) are illustrative in this regard: in (460c) *i-maʔenwar-i*, the head takes the non-contiguity marker *-i* signaling a constituent discontinuity, because Pedro (the dependent) is not contiguous to it. In (460d), the marker of contiguity signals that Pedro, the dependent, and the head *maʔenwar-i* form a genitive construction without constituent discontinuity.

Praça et al. (2017, 52) suggest⁸ that this construction requires nominalization of the main predicate because the adverbial becomes the main predicate when fronted and, as a consequence, the nominalized predicate becomes an argument of the adverbial predicate without the necessity of a copula. According to this view, the literal translation of (460a) and (460d) would be something like ‘It is tomorrow, Pedro’s remembrance of you’. They suggest this parallels the fact that in some TG languages, adverbial expressions may function as predicates without a copula⁹, a construction that the authors conjecture might have existed in TUP, despite not being attested, not even once. This construction is exemplified in (462) in Tapirapé¹⁰

(462) (Tapirapé, TG)

Tapi’ir-a ka’a=pe
Tapir-RFR forest-POSP

‘The Tapir is in the forest.’ (Praça et al. 2017, 48)

It is difficult to think of a reason why such a simple construction is not attested in the Tupinambá corpus, if it existed (it is also not attested in Old Guaraní). Furthermore, the relation between the constructions mentioned in Praça et al. (2017) is not so obvious. The distribution of both constructions in the TG languages is also not informative, as shown in Table 9.2.

⁸The authors neither write about the focal status of the fronted adverbials, nor about the non-focal status of the nominalized predicate.

⁹This construction exists, for example, in Guajá and Nheengatu (Praça et al. 2017), Kamajurá (Seki 2000), Avá Canoeiro (Borges et al. 2006), Tekó (Rose 2003, 185-187), Anambé (Julião 2005, ex. 133), Warazu (Ramirez et al. 2017, 489)

¹⁰My glossing, original orthography maintained.

Language	Presence
Asuriní Tocantins	1
Parakanã	1
Suruí	1
Tapirapé	1
Tembé	1
Guajajara	1
Parintintin	?
Kayabi	1
Asuriní Xingu	1
Arawete	0
Kamayurá	1
Emerillon	0
Guajá	1
Wayampi	0
Ka'apor	0
Anambé	1
Ava-canoeiro	1
Tupinambá	1
Nheengatu	0
Guarayo	?
Sirinono	?
Yuki	0
Warazu	0
Mbya	0
Guaraní	0
Chiriguano	0
Old Guarani	?
Kaiwá	0
Tapiete	0

Table 9.2: Non-focal constructions with fronted adverbials in TG languages

Undoubtedly, there is a discourse-pragmatic base for this construction – be it the type of predicate suggested by Praça et al. (2017) or not. Evidence for this includes not only the fact that topicalized elements are usually fronted to a detached position (see Lambrecht 1994), but also a clue given by the first Tupi grammars. Anchieta (1595, 39v) says that verbs can lose their person indexes (Set II in 4.3) if preceded by an adverb, preposition, gerund (...), or a phrase answering to another one¹¹; Figueira (1687, 93) is more precise when he writes of this construction, explaining that it can occur if it is preceded by some adverb, preposition, or gerund, or if one talks about something that has already been spoken about, pertaining to that verb, i.e., the topic (...)¹².

The nonfocal suffixes *-i* and *u* were not used in the southern variety described by Anchieta (1595, 40). In their place, the translative case marker is used.

- (463) Koromõ feroriβamo
 koromõ je=r-oriβ-amo
 soon 1SG=R₁-happy-TRSL
 ‘Soon I will be happy.’ (AA, 40)

9.6.1 Topic

In Section 4.3.2, it was shown that it is possible to have two different sentences with the same semantic interpretation. In example (125) each sentence has a different word order and hence a different contiguity marker (relational). One reason for these choices be topicalization. The fronting of the intransitive subject or the fronting of the DCA of a transitive verb implies topicalization of these elements. In (125) it seems – the sentences are given without a context – that the difference lies in the fact that the ‘object’ *Pedro* is fronted in (125b) as a marked topic.

It is possible to front the subject *pitaŋ* of (464a) by placing it in the PrDP, thus

¹¹*Mas tendo aduerbio, preposição, gerundio, supino, alguma oração antes, a que ha de responder outra* Anchieta (1595, 39v).

¹²*[q]uando antes dellas fica algum advérbio, ou preposição, ou gerundio; ou se relatamos a cousa, de que já fallamos pertencendo ao tal verbo* (Figueira 1687, 93-94)

topicalizing the subject. This is seen in (464b) where the resumptive argument index is co-indexed with a subscript.

- (464) a. Osi swi pitaja ʔareme...
 o-si-∅ swi pitaj-a ʔar-reme
 3_{CORF}-mother-REF from child-REF be.born-SUBJ
 ‘When a child is born from his/her mother...’ (Araújo, 8)
- b. Pitaja_i osi swi i_iʔareme...
 pitaj-a o-si-∅ ∅-swi i-ʔar-reme
 child-REF 3_{CORF}-mother-REF R₁-from R₂-be.born-SUBJ
 ‘A child, when he/she is born from his/her mother ...’ (Unat-
 tested)

The same contrast observed in the examples in (464a) can be observed in (465):

- (465) a. Koriteʔĩ Pedro ʃeruβa moŋetaw
 koriteʔĩ Pedro ʃe=r-uβ-a moŋeta-w
 now Pedro 1SG=R₁-fatherREF talk-NFOC
 ‘Now, Pedro talks to my father.’ (FA, 96)
- b. Koriteʔĩ ʃeruβa Pedro imoŋetaw
 koriteʔĩ ʃe=r-uβ-a Pedro i-moŋeta-w
 now 1SG=R₁-father-REF Pedro R₂-talk-NFOC
 ‘Now, with my father, Pedro speaks.’ (FA, 96)

In (464), as in (465), the b examples are not attested examples. Nonetheless, such contrasts are indeed attested; both constructions are found in TUP. This is seen in (466), with a topicalization in (466b), indicated by the fronting of the ‘object’ (*maʔeasiβora* ‘one who is ill’) to the PrCS, with a resumptive ‘index’ in the relational of non-contiguity (*i-*) with *suβan*, indicating that its dependent is not contiguous:

- (466) a. K^wese maʔeasiβora paje isuβani
 k^wese maʔeasi-βora paje-∅ i-suβan-i
 yesterday illness-NMZL_{CIRC}-REF shaman-REF R₂-suck-NFOC

‘Yesterday, the shaman sucked (healed) the one who was ill.’ (FA, 96)

- b. K^wese paje maʔeasiβora suβani
 k^wese paje-∅ maʔeasi-βor-a suβan-i
 yesterday shaman-REF illness-NMZL_{CIRC}-REF suck-NFOC
 ‘Yesterday, the one who was ill, the shaman sucked (healed) him.’
 (FA, 96)

9.7 Nonfocal argument index

One manifestation of topicality is seen in the non-canonical indexing of arguments when both arguments are third person. Both Anchieta (1595) and Figueira (1687) acknowledge the phenomenon (see Rodrigues 1990).

In the case of two third person core arguments, Rodrigues (1990, 398) observes: ‘If the subject, that is, the agent, is in focus, it is marked on the verb by *o-*; if conversely, the object, that is, the patient, is in focus, the subject is marked by *ya*’. This contrast is illustrated in 467.

- (467) a. Pedro moja ojuka
 Pedro moj-a o-i-juka
 Pedro snake-REF 3-R₂-kill
 ‘Pedro killed the snake.’ (FA, 99)
- b. Pedro moja jajuka
 Pedro moj-a ja-i-juka
 Pedro snake-REF 1PL.INCL-R₂-kill
 ‘Pedro killed THE SNAKE.’ (FA, 99)

Other examples provided by the first TUP grammar are given in (468).

- (468) a. Moja Pedro jaisuʔu
 moj-a Pedro ja-i-fuʔu
 snake-REF Pedro 1PL.INCL-R₂-bite
 ‘THE SNAKE bit Pedro.’ (AA, 36v)

- b. $\text{feru}\beta\text{a}$ $\text{to}\beta\text{ajara}$ $\text{ja}\text{?u}$
 $\text{fe=r-u}\beta\text{-a}$ $\text{to}\beta\text{ajar-a}$ **ja-?u**
 1SG=R₁-father-REF enemy-REF 1PL.INCL-eat
 ‘THE ENEMIES ate my father.’ (AA, 36v)
- c. Pedro $\text{ta}\text{?ira}$ $\text{jainup}\bar{\text{a}}$
 Pedro $\text{t-a}\text{?ir-a}$ $\text{ja-i-nup}\bar{\text{a}}$
 Pedro R₂-son-REF 1PL.INCL-R₂-hit
 ‘HIS SON hit Pedro.’ (AA, 36v)
- d. $\text{Moru}\beta\text{isa}\beta\text{a}$ mona jainamiokukar
 $\text{moru}\beta\text{isa}\beta\text{-a}$ $\text{mona-}\emptyset$ **ja-i-nami-ok-ukar**
 judge-REF thief-REF 1PL.INCL-R₂-ear-cut-FAC
 ‘The judge had the THIEF’S EAR cut off.’ (AA, 36v)
- e. $\text{Japop}^{\text{w}}\text{arat}\bar{\text{a}}$, $\text{imoa}\eta\text{aipapa}$. Suwi
 $\text{ja-po-p}^{\text{w}}\text{ar-at}\bar{\text{a}}$ $\text{i-mo-}\eta\text{aipap}\beta\text{-a}$ $\text{s-uw}\bar{\text{i}}-\emptyset$
 1PL.INCL-hand-tie-strong R₂-CAUS-evil-GER R₂-blood-REF
 momukapa , $\text{jainup}\bar{\text{a}}\text{nup}\bar{\text{a}}$
 $\text{mopukap-}\beta\text{a}$ $\text{ja-i-nup}\bar{\text{a}}.\text{nup}\bar{\text{a}}$
 CAUS-spill-GER 1PL.INCL-R₂-hit.ITER
 ‘They tied his hands, making him bleed and hitting him.’ (Poemas, 120)¹³

Outside the grammars, the nonfocal agent or focal-undergoer is further attested many times, indicating it was a common resource of the language, as (469) exemplifies. Examples such as the following are important because they are inserted in a discursive context, allowing for a better understanding of the phenomenon.

- (469) a. Moraseja $\text{rero}\beta\text{jara}$ $\text{ipi}\text{?a}$ jaiporaka
 p.orasej-a $\text{r-ero}\beta\text{jar-a}$ $\text{i-pi}\text{?a-}\emptyset$ **ja-i-poraka**
 R₃.dance-REF R₁-belief-REF R₂-heart-REF 1PL.INCL-R₂-fill
 ‘The belief in the dance fills their hearts.’ (Teatro, 32)

¹³*Amarraram suas mãos fortemente, fazendo-lhe mal. Derramando seu sangue, ficaram a açoítá-lo.* (Poemas, 120)

- b. Nomenari emonã tekoarwera
na-o-menar-i emonã t-eko-ar-wer-a
NEG-3-marry-NEG thus R₃-be-NMLZ_{AG}-PST-REF
jaipeʔa
ja-i-peʔa
1PL.INCL-R₂-separate
‘He/she did not marry. Thus, having been (married), he/she divorces her/him.’ (Araújo, 128)

The anonymous vocabulary in Anonymous (1952a) gives important information regarding word order. In the entry for the ‘lunar eclipse’ (eclipsarse a lua) (vol. I, 108), the example that translates to ‘something eats the moon’¹⁴ is given in two different orders, which are shown in (470).

- (470) a. Jasi maʔe jaʔu
jasi-∅ maʔe-∅ **ja**-ʔu
moon-REF thing-REF **1PL.INCL**-ingest
‘A THING eats the moon.’ (VLB, I, 108)
- b. Maʔe jasi jaʔu
maʔe-∅ jasi-∅ **ja**-ʔu
thing-REF moon-REF **1PL.INCL**-R₂-ingest
‘A THING eats the moon.’ (VLB, I, 108)

Just like in a direct-inverse system, this non-canonical marking in TUP contrasts with the opposition of active/passive/antipassive and obviative/proximate (see Givón 1994), and the basic function of these syntactic devices is to rank participant RPs along a certain dimension or hierarchy, which might be called the animacy hierarchy, and which displays great variety across languages (see Croft 2003, 128-157 and references in Oshima 2007, 733-734).

The non-canonical marking of third person with *ja-* avoids ambiguity since, as suggested by the comments of Anchieta (1595, 36v), Figueira (1687, 99),

¹⁴Anonymous (1952a, 108) says that the ‘Tupinambá’ (northern groups) thought that a jaguar ate the moon, but the ‘Tupis’ (southern groups) thought it was a snake. See d’Abbeville (1614, chap. 51).

and Anonymous (1952a, 108), there should have been a hierarchy of the type human > non-human for marking the lowest argument in the hierarchy as more salient, somewhat like an obviative-proximate distinction¹⁵, i.e., a non-salient or less topical (obviative) third-person referent and a more salient or more topical (proximate) third-person referent in a given discourse context. This is a distinction that, due to the character of the texts, cannot be recovered with precision, since in many examples both arguments are human (e.g. 468b, 468d, 468e, 469a).

¹⁵For obviative and proximate, see Dryer (1992a), Aissen (1997), Oshima (2007), Martin (2011), Kiparsky (2015).

Chapter 10

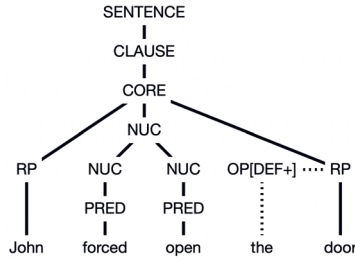
Complex sentences

This chapter presents aspects of complex sentences, of which RRG has a distinctive theory consisting of three main components: the theory of juncture, i.e., the units involved in building complex sentences; the theory of nexus, which deals with the relationship between units involved in the linking (Van Valin Jr and LaPolla 1997, 441); and finally, the theory of interclausal semantic relations, which deals with the semantic relationship between the units in the juncture.

The next sections deal with complex sentences in TUP based on the levels of juncture in Section 10.1 and the types of relations in Section 10.2. Complex RPs are dealt with in Section 10.3.

10.1 Levels of juncture

The units involved in complex constructions are those of the LSC: nucleus, core, and clause (see Section 3). The juncture of these levels causes the following patterns to emerge:



[do' (John, [force' (John, door)])] CAUSE [BECOME open' (door)]

Figure 10.1: English nuclear juncture

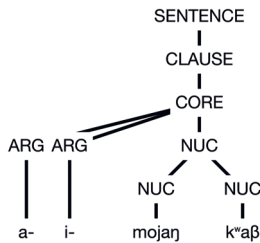


Figure 10.2: Nuclear juncture
[know' (I, [do'(1SG, 3)])]

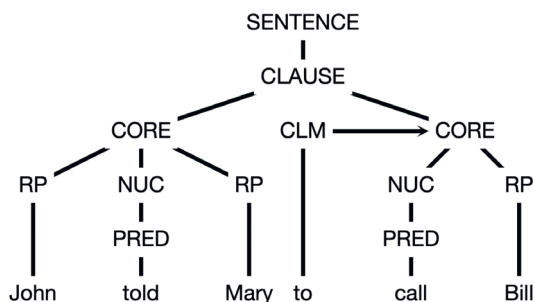
10.1.2 Core junctures

Core junctures are made up of multiple cores, each with its own nucleus and some (or all) of its arguments. In this type of juncture, one of the core arguments functions semantically as an argument of both predicates, as in (474), with its structure given in Figure 10.3¹:

(474) John saw Mary calling Bill
see' (John, do' (Mary) [call'(Mary, Bill)])

Both cores in Figure (10.3) have their own nuclei and arguments, and one argument appears in the semantic representation of both predicates, but only

¹For the logical structure of 'verbs of saying', see Van Valin Jr and LaPolla (1997, 116-118).



do'(John[**express**(α).**to**.(β).**in.language**.(γ ')(John, Mary)]) CAUSE
do'(Mary)[**call'**(Mary,Bill)]

Figure 10.3: English core juncture

once in the syntactic representation. In (474), *Mary* is the shared argument. It is the undergoer of *see* and the actor of *call*. In other words, the linked core is an argument of the matrix verb semantically but not syntactically. This is an example of core subordination whereby a core unit is an argument of a matrix core. The subordinate nature of the linked core is indicated by the gerund in the dependent core².

The only type of core juncture in TUP will be discussed in Section 10.2.3.

10.1.3 Clause junctures

Clause juncture constructions contain two (or more) independent clauses, all of which have their own arguments (see 475).

(475) Dana jogged through the park, and Kim waved to him

In the above example, *Dana jogged through the park* and *Kim waved to him* are distinct clauses, and each is linked independently of the other, just as if each were a simple sentence on its own. The fact that there is a pronoun in the

²See Van Valin Jr and LaPolla (1997, 444-447) for restrictions and structural differences between English core and nuclear junctures.

second clause referring (possibly) to Dana in the first clause does not affect the linking.

10.1.4 Sentential junctures

Sentential junctures are complex constructions comprising two or more sentences. Such junctures may be a sentence with several clauses or a clause with a core juncture (see Van Valin Jr and LaPolla (1997, 469) and Van Valin Jr (2005, 192)).

10.2 Nexus relations

In RRG, nexus relations consist of the common clause linkage relations, coordination and subordination, as well as a relation unique to RRG, cosubordination. These are divided along the features of embedding and dependence³. They are schematically shown in Figure (10.4) from Van Valin Jr (2005, 188).

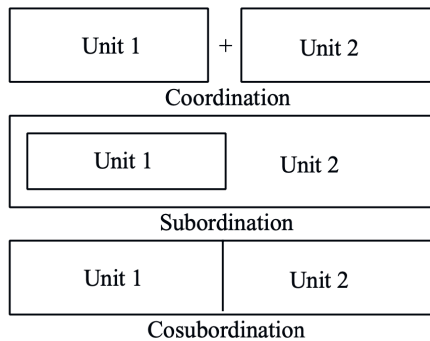


Figure 10.4: Nexus relations, from Van Valin Jr (2005, 188)

The schema in Figure (10.4) captures the specific feature of each nexus relation. The nexi are divided along the features of embedding, dependence, and independence. At the clause level, coordination shows two (or more)

³Independence, the ability to occur as an independent utterance, applies only to clausal junctures and therefore is not a general feature of nexus relations.

linked independent units. The linked units are self-sufficient as far as grammatical categories are concerned, as in (10.5). In a subordinate nexus relation, the schema indicates the embedding feature associated with subordination, wherein a matrix unit contains an embedded subordinate unit. The structurally embedded units can function as core arguments (complement clauses) or modifiers (adverbial clauses and relative clauses).

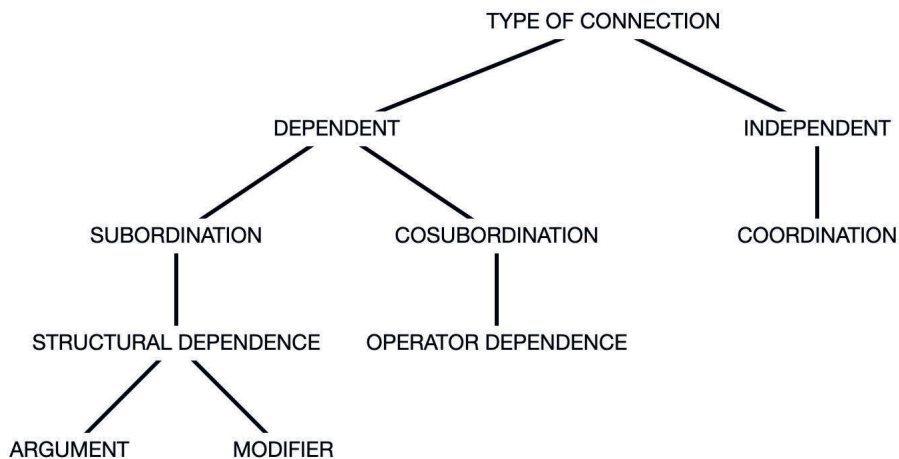


Figure 10.5: Types of connection in complex constructions, from Pavey (2010, 226)

Cosubordinate units are not independent; there is obligatory operator sharing at the level of juncture. They do not belong together with subordination, since subordination is defined in terms of embedding. The linked units appear to be independent, as in coordination, since none is embedded, but they practically belong together, as in subordination. Although cosubordinate units may or may not be conjunctively linked, some relations require them to belong together for semantic and syntactic reasons. Such a relation follows if one or two grammatical categories are shared between the linked units. The unit with all the grammatical operators may stand on its own, but not without the necessary operator for a specific layer. Figure 10.5 shows the nexus types.

Consequently, the notion of ‘dependence’ is either structural dependence or

operator dependence. Subordination and cosubordination share the ‘dependence’ feature, but not the type of dependence. The type of dependence in cosubordination is operator dependence, while the type in subordination is structural dependence through embedding.

Each nexus type can in principle occur at each level of juncture, generating nine possible juncture-nexus types. Hierarchically, they may be ranked according to the tightness of the linkage, yielding the clause linkage hierarchy in (476), from Van Valin Jr (2001b).

(476) Clause linkage hierarchy

[Tightest] Nuclear cosubordination > nuclear subordination > nuclear coordination > core cosubordination > core subordination > core coordination > clausal cosubordination > clausal subordination > clausal coordination [Weakest]

10.2.1 Coordination

The coordinated elements are of the same syntactic layer (nucleus, core, and clause) and have operator independence at the level of juncture.

Juxtaposition⁴ is very common in TUP; clauses are juxtaposed without any clause linkage markers (CLM), as in (477). For example, in (477c), each clause has its own operators: the IF of the first is imperative and the second is permissive, as shown by its representation in Figure 10.6.

- (477) a. [Tupã ajõ maʔeete], [aʔe janemojaŋáramo],
 Tupã ajõ maʔe-ete aʔe jane=∅-mojaŋ-ar-amo
 God only thing-INTS DEM 1PL.INCL=R₁-do-NMLZ_{AG}-TRSL
 [maʔe tetirwã mojangáramo sekow]
 maʔe-∅ tetirwã mojaŋ-ar-amo s-eko-w
 thing-REF all do-NMLZ_{AG}-TRSL R₂-be-NFOC

⁴It has been suggested that the juxtaposition of clauses is a linking type with prosodic bases (see Jiménez 2021). While there is no operator sharing or dependence among juxtaposed clauses, it is true that juxtaposed clauses tend to have the same IF and tense.

‘Only God is something great, being our creator, being the creator of all things.’ (DC, I, 131)

- b. [Ejori saʔaŋa rō], [totupājeʔeŋaβi],
 e-jori s-aʔaŋ-a rō t-o-tupā-jeʔeŋ-aβi
 2SG.IMP-come R₂-tempt-GER then HORT-3-God-word-fail
 [tokaʔu tomondarō], [toporepejan
 t-o-kaʔu t-o-mondarō t-o-poro-epejan
 HORT-3.beer.drink HORT-3-steal HORT-3-ANTIP-attack
 ojkoβo], [tojpuru tekopofi], [toso ko
 o-eko-βo t-o-i-puru t-eko-pofi-∅ t-o-so ko
 3-be-GER HORT-3-R₂-use R₂-law-evil-REF HORT-3-go this
 taβa swi]
 taβ-a swi
 village-REF from

‘Come, then, to tempt them, so that they violate God’s word, so that they drink beer, so that they steal, so that they attack people, so that they act sinfully, so that they go away from this village.’ (Teatro, 18)

- c. Ejasoʔjaβok nekaramemuã
 e-i-asoʔjaβ-ok ne=∅-karamemuã-∅
 2SG.IMP-R₂-cover-take.off 2SG=R₁-box-REF
 tasepʔak nemaʔe
 t-a-s-epʔak ne=∅-maʔe-∅
 HORT-1SG-R₂-see 2SG=R₁-thing-REF
 ‘Uncover your box (and) I may see your things.’ (Léry, 346)

Another example of parataxis is given in (478).

- (478) Erejojajpe, erejaʔope, erejangaʔope
 ere-i-jaj=pe ere-i-aʔo=pe ere-i-angaʔo=pe
 2SG-R₂-mock=Q 2SG-R₂-insult=Q 2SG-R₂-threaten=Q
 neruβa, nesi, neramĩja,
 ne=r-uβ-a ne=∅-si-∅ ne=r-amĩj-a
 2SG=R₁-father-REF 2SG=R₁-mother-REF 2SG=R₁-grandfather-REF

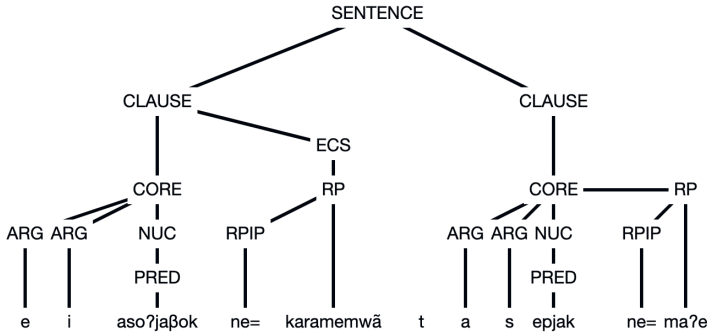


Figure 10.6: Parataxis. Clausal juncture without a clause linkage marker

nearija?
 ne=∅-arij-a
 2SG=R₁-grandmother-REF
 ‘Did you mock, insult, threaten your father, mother, grandfather, grand-
 mother?’ (Araújo, 100v)

In (479), each clause has its own tense marker (the future clitic =*ne*) and they are linked by the adversative CLM *konipo*.

(479) Ajpo e?i jōte isupéne konipo aβare supe imome?uw
 ajpo e-ʔi jōte i-supe=ne **konipo** aβare-∅ supe i-mome?u-w
 DEM 3-say only R₂-DAT=FUT **or** priest-REF to R₂-tell-NFOC
 iakakapawama resene
 i-akakaβ-wam-a r-ese=ne
 R₂-reprehend-FUT-REF R₁-BECAUSE-FUT
 ‘You will say this to him only or you’ll tell the priest so that he repre-
 hends him (lit. for his reprehension).’ (DC, I, 228)

Adversative coordination may involve expectedness in the form of ‘p but not q’, or presuppose that ‘normally, p and not q’ (Croft 2022a, 437). Adversative coordination is expressed through parataxis with the adversative particle *a?e*

in final position ⁵.

- (480) a. Naferemiawsuβa ruã, feremireko a?e
 na-fe=r-emiawsuβ-a ruã fe=r-emirekó-∅ **a?e**
 NEG-1SG=R₁-slave-REF NEG 1SG=R₁-wife-REF this
 ‘It is not my slave, but my wife.’ (Araújo, 95)
- b. Na Pero ruã, tiβira a?e
 na Pero ruã t-iβir-a **a?e**
 NEG Pero NEG R₂-brother-REF this
 ‘It was not Pero (who was going), but his brother.’ (VLB, II, 115)
- c. Karaiβa nasetaj, São Sebastião a?e
 karaiβ-a na-s-eta-i São Sebastião **a?e**
 christians-REF NEG-R₂-many-NEG Saint Sebastian this
 omondik tata sese
 o-mondik t-ata-∅ s-ese
 3-light R₃-fire-REF R₂-POSP
 ‘The Christians were not many, but Saint Sebastian ignited fire on them.’ (Teatro, 22)

10.2.2 Subordination

The RRG theory of clause linkage distinguishes between two types of subordination: daughter subordination, in which the subordinate junct is a daughter of a higher node, and peripheral subordination, in which the subordinate junct functions as a peripheral modifier of one of the layers, just like adverbials and adjuncts, as seen in Section 3.1.2 (Van Valin Jr 2007; Matic’ et al. 2014).

A common type of daughter subordination is complementation (see Van Valin Jr and LaPolla 1997, 492-504), whereby a larger unit is linked to a smaller unit. This is exemplified in (481), with its representation in Figure 10.7, where *believe* takes ‘that *pets are allowed in the airplane*’ as its complement. In other words, the clause ‘*that pets are allowed in the airplane*’ is an argument

⁵In this sense, Navarro’s interpretation of *a?e* as a conjunction Navarro (2013, 11) (*a?e*³) seems to be a misinterpretation of structures such as (480).

of *believe*. Note that in Figure (10.7), the embedded clause is a daughter of the core node.

(481) Sue believes that pets are allowed in the airplane

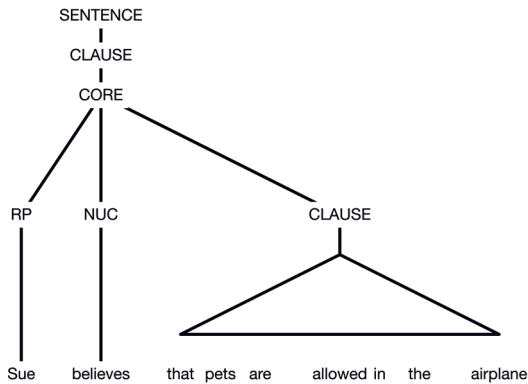


Figure 10.7: Example of (object) complementation (subordination) in English

In TUP, a clause cannot function as a core argument due to the head-marking character of the core, but it can be hosted in the ECS (see Section 6.3) as long as it is nominalized. In order to be in the ECS, it must be nominalized since the ECS in TUP cannot host a finite clause. The sentence in (481) would literally be translated into TUP as ‘Sue believes it, the allowance of pets in the plane’. In (482a), a possessive RP semantically related to the undergoer argument of *enuβ* ‘hear’, marked with *s-*, is in the ECS (see Section 6.3). Literally, in (482a), one must say, ‘May they hear it, that saying of yours’.

- (482) a. Tosenu ajpo ne?e
t-o_i-s_j-enuβ [ajpo ne=∅-?e-∅]_j
HORT-3-R₂-hear DEM 2SG=R₁-say-REF
‘May they hear that saying of yours.’ (Teatro, 186)
- b. Neakaŋa juka ajpota koríne
ne=∅-akaŋ-a ∅-juka-∅ a-i-pota kori=ne
2SG=R₁-head-REF -R₁-break-REF 1SG-R₂-want today=FUT

‘I shall want to break your head today.’ (Staden, 156)

Peripheral or adverbial subordination involves a clause appearing as a peripheral modifier, and because all three layers may be modified, there is ad-nuclear, ad-core, and ad-clausal subordination. Ad-nuclear subordination is not found in TUP. In ad-core subordination the core is modified by a peripheral adverbial providing information about time, space, manner, or pace, as in (483)⁶, where the adverbial expression *jeporupi* modifies the core *ejotĩ*.

- (483) Ejotĩ nekesaβa jepo rupi
 e-jotĩ ne=∅-kesaβa je=∅-po r-upi
 2SG.IMP-tie 2SG=R₁-sleeping.mat-REF 2SG=R₁-hand R₁-at
 ‘Tie your hammock next to me.’ (AA, 44)

Ad-clausal subordination is found, for example, when joined together by a causal link marker such as *because*. An example is given in (484), with its representation in Figure 10.8.

- (484) Pedro oso omonóreme
 Pedro o-so o-mo-so-reme
 Pedro 3-go CORF-CAUS-go-because
 ‘Pedro goes because/when/if he is sent.’ (FA, 84)

The suffix *-reme* is here characterized as a CLM, but diachronically, it is a postposition which, based on the extant texts, appears to have been going through a process of grammaticalization (see Section 6.5.1.3) towards the loss of its postpositional status, becoming a subordinating morpheme in Nheengatu (see Cruz 2011, 390-391)⁷.

Other examples of ad-clausal subordination are given in (485).⁸

⁶The postposition *porupi* ‘parallel to’ is formed by *po* ‘hand’ and *upi* ‘through, along, according to’. One may also regard it as an adverbial expression, as in *i-po r-upi* ‘parallel to one’s hand’. There is no way of knowing which is the correct analyses.

⁷The same process is attested in Tekó (Rose 2011, 303-307, 336-338).

⁸The subjunctive suffix *-reme* is sometimes glossed as CLM for a better grasping the structure of complex sentences, specially when the tree-like representation is given.

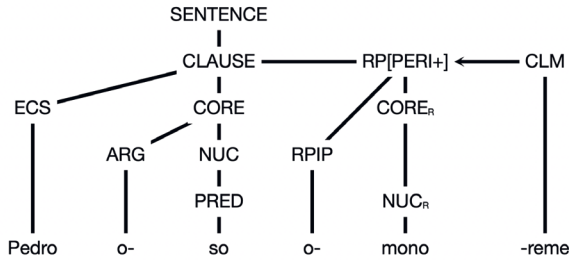


Figure 10.8: Peripheral ad-clausal subordination

- (485) a. Ereʔu memẽ soʔo ʔaretéreme
 ere-ʔu memẽ s-oʔo-∅ ʔar-ete-reme
 2SG-ingest always R₂-meat-REF day-INTS-CLM
 ‘You always ate meat when it was a holiday.’ (Teatro, 168)
- b. Marãpe Tupã serekow emonã sekóreme?
 Marã=pe Tupã s-ereko-w emonã s-eko-reme
 how=Q God R₂-treat-NFOC thus R₂-be-CLM
 ‘How did God treat them when/after they acted this way?’ (DC, I, 160)
- c. ʃerureme, asoβajtĩ ʃeremierekop^wera
 ʃe=r-ur-eme a-s-oβajtĩ ʃe=r-emi-ereko-p^wer-a
 1SG=R₁-come-CLM 1SG-R₂-meet 1SG=R₁-RES-keep-PST-REF
 ‘When I came, I found what I had been keeping.’ (Léry, 375)

The main clausal subordination construction type involves a clause as a peripheral modifier (ad-clausal subordination). A subordinate unit expresses, in a sense, a secondary event within the main event, so it may also have its own arguments and operators. In (486), the dependent clause in both sentences is linked to the main clause through a CLM. The LSC of (486a) is represented in Figure 10.9:

- (486) a. iβakipe Cristo janejara jeupirire ko sancto
 iβak-pe Cristo jane=∅-jar-a je-upir-rire ko sancto
 sky-LOC Christ 1SG=R₂-lord-REF RFLX-elevate-after this saint

clause linkage markers such as *jepe* ‘even if, despite, even though’. The example in (488) shows two clauses, each with its own tense operator, linked by *jepe*. Its representation is given in (10.10).

- (488) Ereipisirō jepene, neposwi
 ere-i-pisirō jepe=ne ne=∅-po-swi
 2SG-R₂-set.free even.if=FUT 2SG=R₁-hand-ABL
 arosēne
 a-ero-sē=ne
 2SG-SCAU-exit=FUT
 ‘Even if you release them, I will take them out of your hand.’ (Teatro, 42)

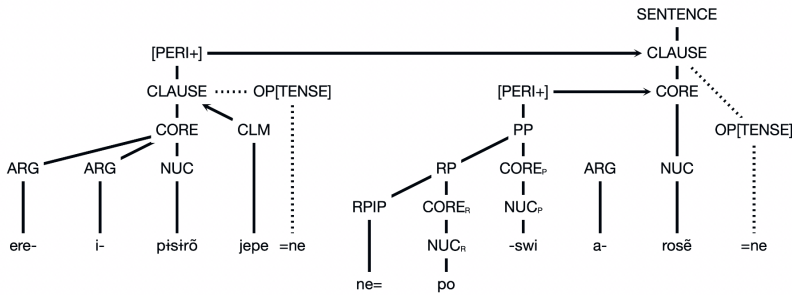


Figure 10.10: Ad-clausal subordination with a clause linkage marker

Sentential subordination is also possible. In (489a), the fronted subordinate (adverbial) clause *ajpo ojoupe ʔe aβe* is in the PrDP of the sentence *ojara repip^wera rejitiki Tupāokipe*. A simplified representation of it is given in Figure (10.11).

- (489) a. Ajpo ojoupe ʔe aβe, ojara
 ajpo o-jou-pe ʔe aβe o-jar-a
 DEM CORF-himself-DAT say CONJ CORF-master-REF
 repip^wera rejitiki Tupāokipe
 r-epi-p^wer-a r-ejitik-i Tupā-ok-pe
 R₁-payment-PST-REF R₁-throw-NFOC God-house-LOC

‘After saying this to himself, he threw the payment for his own master in the temple.’ (Ar., Cat., 57v)

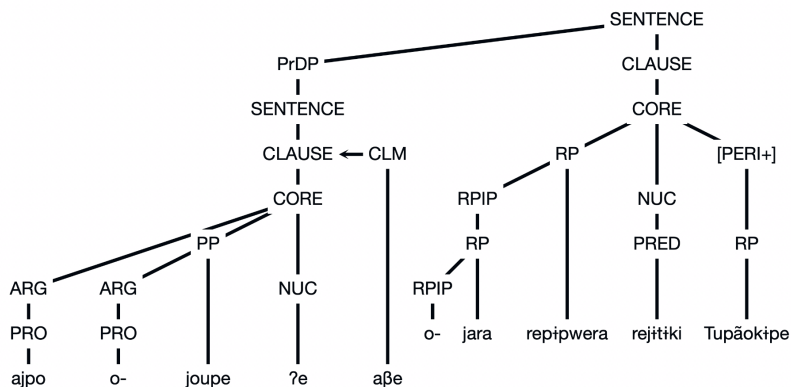


Figure 10.11: Subordination at the Sentence-level

A sentence may contain both, a daughter subordinate and an adverbial subordinate, as in (490):

- (490) Oipotarepe judeus ojuka iswi
 o-i-potar-e=pe judeus o-juka-∅ i-swi
 3-R₂-want-PRCL=Q jews CORF-kill-REF R₂-from
 ojepisirō?eima?
 o-je-pisirō-βo-?eim-a
 3-RFLX-free-GER-PRIV-REF

‘Did he indeed want his (own) killing by the Jews without getting rid of them?’ (Bettendorf, 46)

10.2.3 Cosubordination

The existence of cosubordination as an intermediate between coordination and subordination was first proposed in Foley and Van Valin Jr (1984) and has overcome significant theoretical criticisms regarding its validity (see Bickel

2010; Foley 2010). Recently, Van Valin Jr (2021) has put an end to the discussion by showing it must indeed be treated as a distinct nexus type.

Cosubordination resembles subordination in that it is structurally asymmetrical, comprising an independent clause and a cosubordinate clause which cannot stand alone as an independent unit. Thus, both cosubordinate and subordinate clauses are dependent. However, cosubordination resembles coordination in that there is no embedding. The dependence is exclusively at the level of operators: the linked unit must share at least one operator at the level of juncture with the licensing unit. Cosubordination cannot occur at the sentence level, because there are no sentence-level operators.

In (491), the examples of complex nuclei would seem to be instances of nuclear cosubordination, because both *awsuβ* and *peʔa* would share a nuclear operator.

- (491) a. Tupã osawsupeʔa
 Tupã o-s-**awsuβ-peʔa**
 God 3-R₂-love-quit
 ‘God stopped loving them.’ (Teatro, 30)
 [stop’(God, [love’(God, 3)]]
- b. Toiʔusejkatu Tupã reko
 t-o-i-ʔ**u-sej**-katu Tupã r-eko-∅
 HORT-3-R₂-**ingest-want**-INTS God R₁-law-REF
 ‘May he really want (to ingest) God’s law.’ (Araújo, 81v)

Nonetheless, there are cases of nuclear negation (see Section 6.5.3.2), which are only found in nominalized clauses with a complex nucleus. This means that nuclear cosubordination may have existed despite not being attested in the texts, because the nominalizations in (492) would not be possible if their non-nominalized counterparts did not exist.

- (492) a. Osopotareʔimaʔe
 o-**so-potar-eʔim**-βaʔe
 3-**go-want-PRIV**-NMLZ_{REL}
 ‘The one who does not want to go.’ (Araújo, 70v)

The translation of (495a) could be misleading, since it seems to imply that ‘aimlessly’ is a feature of the argument’s volition or intention, which would make the peripheral adverbial a core modifier, but in fact, it is only the action (in this case, walking) that is affected.

- (495) a. Awatawata tejẽ
 a-**wata-RED_D** tejẽ
 1SG-**walk-walk** in.vain
 ‘I walk / keep on walking around aimlessly.’ (VLB, II, 140)
- b. Kipe ajenupãnopã
 kipe a-je-**nupã-RED_D**
 for.a.long.time 1SG-RFLX-**hit-hit**
 ‘I kept punishing myself for a long time.’ (Teatro, 174)
- c. Anosesem
 a-**nosem-RED_M**
 1SG-**remove-RED_M**
 ‘I remove one after the other.’ (VLB, II, 129)

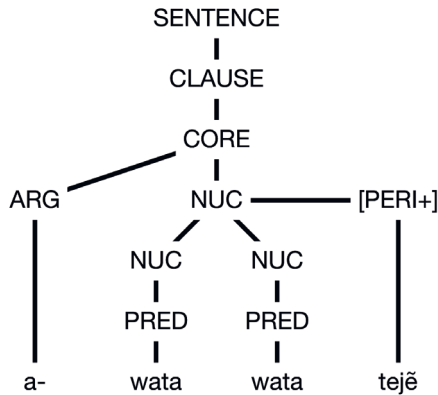


Figure 10.12: Nuclear cosubordination

Besides incorporation and reduplication, another instance of cosubordination commonly found in TUP is core cosubordination. In core cosubordination,

a gerund (nominalized core) combines with a finite core in a core juncture. Cosubordinate cores must show traits of grammatical dependence in two senses: they must depend, at least, on one of the core operators, and they must share, in TUP, a core argument (see Foley and Van Valin Jr 1984, 261,304 and Cerda 2021). The cosubordinate core must be non-finite and cannot have a subject or controller that is not an argument of the main predicate.

Examples of core cosubordination with a gerund are given in (496) and (497). Note that *o-* ‘third person’ in (496a) and the *wi-* in (496b) are the shared arguments, which are overtly marked in intransitive gerunds.

- (496) a. Oropaß oromanõmo
 oro-paß **oro-manõ-ßo**
 1PL.EXCL-terminate **1PL.EXCL-die-GER**
 ‘We come to an end (by) dying.’ (Poemas, 82)
- b. Asawsu sese wijemoririja
 a-s-awsuß s-ese wi-je-moririj-a
 1SG-R₂-love R₂-POSP 1SG_{CORF}-RFLX-care-**GER**
 ‘I love her (while) taking care/and take care of her.’ (Poemas, 182)

The representation of (496a) is given in Figure 10.13. Note that the non-finite core is a peripheral core modifier of the finite core, and the semantic relation between both cores (Van Valin Jr 2022) is that of a simultaneous event.

There is no coreference in core junctures, only argument sharing (see Van Valin Jr 2022, 135-141), because coreference is a property of clausal linkage and thus distinct from argument sharing. In the examples below, with a transitive verb in the second core, there is no overt marking, but still, in (497), *Pedro* is the shared argument between ‘go’ and ‘kill’.

- (497) Oso Pedro jawara jukaßo
 o-so Pedro jawar-a juka-ßo
 3-go Pedro jaguar-REF kill-**GER**
 ‘Pedro goes to kill the jaguar / Pedro goes killing the jaguar.’ (FA, 155)

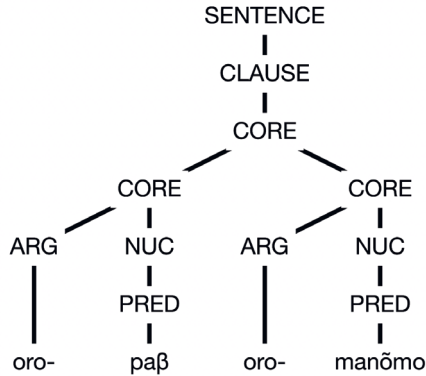


Figure 10.13: Core cosubordination

The translations in (496a), (496b), and (497), contain three different interclausal semantic relations. One could argue that TUP violates the predicted relationship between the interclausal syntactic relations hierarchy and interclausal semantic relations due to the fact that different types of actions are expressed through the same nexus relations and same construction – in this case, core cosubordination. These include purposive actions (intention), simultaneous or sequential (multiple actions), and manner and position actions (modifying sub-actions) (see Van Valin Jr 2022, 65-67).⁹

In (497), *jawara jukáβo* ‘killing the jaguar/to kill the jaguar’ is linked at the core level through cosubordination with *so* ‘go’ forming a complex core. Example (497) shows formal asymmetry between the linked predicative units, since *so* ‘go’ is a verb covertly marked for tense (non-future), and *jukáβo* is, on the other hand, a non-finite form. This asymmetry is not, however, relevant in the model for the theory of nexus relations and juncture, since it does not contradict the requirement of symmetry between layers or strata (of a functional nature): *so* ‘go’ and *jukáβo* ‘killing, to kill’ are joined at the core

⁹Some well-known examples after Chomsky (1965) are cases of the same sequence of words having different meanings. This is the Tupinambá equivalent of ambiguous English sentences, such as *flying planes can be dangerous*, *visiting anthropologists can be dangerous*, or *I saw the man with the telescope*. Obviously, they are ambiguous in a certain context, but they must have different structures which correspond to each possible interpretation.

level.

When a finite core with an M-transitive verb is joined by a gerund (nominalized core), the non-finite core, if M-intransitive, receives Set III indexes (see Table 4.3) signaling coreference of the subjects (498).

Example (498b) is repeated from (496b).

- (498) a. Nasopotari mamō nepiri witekoβo
 n-a-so-potar-i mamō ne=∅-piri **wit**-eko-βo
 NEG-1SG-go-want-NEG far 2SG=R₁-near **1SG**_{CORF}-be-GER
 jě
 jě
 PRCL
 ‘I do not want to go far away in order to be near you.’ (Poemas, 100)
- b. Asausu sese wijemoririja
 a-s-awsu [s-ese **wi**-je-moririj-a]
 1SG-R₂-love R₂-POSP **1SG**_{CORF}-care-GER
 ‘I love her (while) taking care/and take care of her.’ (Poemas, 182)

If the nominalized core is M-transitive, then, as already mentioned, there is argument sharing with the subject of the finite core. The direct, indirect, or oblique core argument obligatorily precedes the predicate of the nominalized core, either as a relational of non-contiguity (499a) or as an RP, as in (499b)¹⁰ and (499c). If the dependent predicate has three arguments, both direct and oblique or indirect core arguments precede it as in (499d).

- (499) a. Seroβjasare?ima potirōw
 s-eroβjar-sar-e?im-a potirō-w
 R₂-believe-NMLZ_{AG}-REF work.together-NFOC
 ijīβōjīβōmo, ijukáβo
 [**i**-jīβō-jīβō-βo] [**i**-juka-βo]
 R₂-shoot-ITER-GER R₂-kill-GER

¹⁰The verb *moŋeta* ‘speak’ is transitive, i.e., the addressee is a direct undergoer core argument.

‘Those who did not believe him worked together shooting arrows at him, killing him.’ (Araújo, 3v)

- b. Marã e?ipe ase karaiβeβe warõana
 Marã e-?i=pe ase [karaiβeβe-∅ o-arõ-ar-a
 what 3-say=Q we angel-REF CORF-guard-NMLZ_{AG}-REF
 moŋetaβo?
 moŋeta-βo]
 speak-GER

‘What do we say (when) praying to the guardian angel?’ (Araújo, 23v)

- c. ferarõkату jepe, nepi?a pupe
 je=r-arõ-kату jepe [ne=∅-pi?a-∅ pupe
 1SG=R₁-watch.over-INTS PRON 2SG=R₁-heart POSP
 femima
 je=∅-mim-a]
 1SG=R₁-hide-GER

‘Watch over me while hiding me in your heart.’ (Poemas, 133)

- d. Ojaβieteβepe kujã Tupã je?eŋa
 o-i-aβi-ete-βe=pe kujã-∅ Tupã ∅-je?eŋ-a
 3-R₂-transgress-INTS-also=Q woman-REF God R₁-word-REF
 ... isupe oapisara amõ me?eŋa
 ... i-supe o-apisar-a amõ me?eŋ-a
 ... R₂-DAT CORF-friend-REF some give-GER

‘Does a woman also transgress God’s word (by) giving him some of her friends.’ (Araújo, 72)

In (500), there is nuclear cosubordination and core cosubordination, showing that different levels of juncture may occur within a sentence. Its representation is given in Figure (10.14).

- (500) ?ara paβire imoingoβejeβiri opiri
 ?ar-a paβ-ire i-mo-ikoβe-jeβiri o-piri
 world-REF finish-after R₂-CAUS-live-return CORF-near

serasóβo awjeramajēne
 s-era-so-βo awieramajē=ne
 R₂-SCAU-go-GER eternally=FUT

‘After the end of the world, he will cause them to return to life, bringing them to himself for eternity.’ (Araújo, 27)

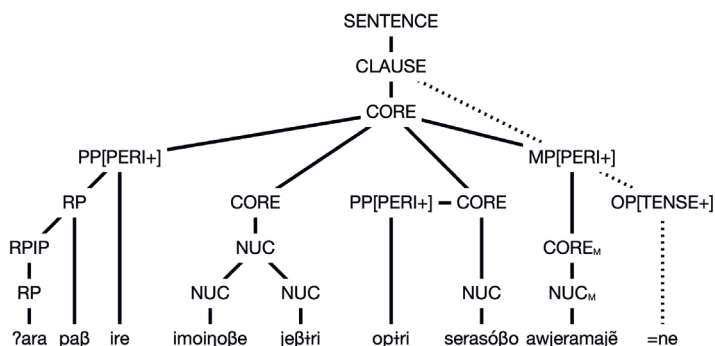


Figure 10.14: Nuclear and core cosubordination in a sentence

Core cosubordination has often been associated with switch-reference (SR) in some descriptions of TG languages (e.g. Cabral and Rodrigues 2005; Rodrigues and Cabral 2012; Silva and Cabral 2013; Dooley 2015)¹¹. Switch-reference is exclusively a property of clausal junctures (see Foley and Van Valin Jr 1984, 257-258, 276-277, Jacobsen Jr 1992, Roberts 1988, Kihara 2017, 150-151, and Hammond 2015).¹² In a switch-reference system, a particular syntactic or semantic function (see Van Valin Jr and LaPolla 1997, 287-288 and Van Valin Jr 2005, 101-107) is monitored (usually the pivot), and verbal affixes signal whether the RP which has that function in a particular clause is coreferential

¹¹Rose (2011, 425) says that this construction resembles switch-reference systems, but she was not categorical about it.

¹²According to Foley and Van Valin Jr (1984, 322-360) switch-reference is one of four types of morphosyntactic systems for signaling the reference relations of RP arguments in discourse, the others being switch-function, gender system, and zero anaphora. For Comrie (1989), there are five different types of reference-tracking system: gender/class indexing, reflexive pronouns, switch-function, switch-reference, and obviation.

or not with the RP which has that same function in a syntactically related controlling clause. Switch-reference is found mainly in predicate-final languages and it is canonically defined as an ‘inflectional category of the verb, which indicates whether or not its subject is identical with the subject of some other verb’ (Haiman and Munro 1983, ix). It is more common for SR to occur with finite verb forms, but there are cases of non-finite forms in SR, as in (van Gijn and Hammond 2016, 45), where the SR construction with a verb form reduced in inflectional potential encodes adverbial function, similarly to a converb (see also Nichols 1983, 245). The point to be made is that a variety of constructions have been referred to as SR (Haiman and Munro 1983; van Gijn and Hammond 2016).

It has already been said that the same construction, besides expressing purpose (a subject’s intention to proceed with a course of action), also expresses a simultaneous action. Some examples of purposive clauses are given in (501), and examples of simultaneous clauses are given in (502).

- (501) a. Asopota neretāme neporajatu
 a-so-pota ne=r-etāma=pe ne=∅-poraj-katu
 1SG-go-want 2SG=R₁-land-POSP 2SG=R₁-beauty-INTS
 rep^jaka
 r-ep^jak-a
 R₁-see-GER
 ‘I want to go to your land in order to/and see your great beauty.’
 (Poemas, 92)
- b. Tereju iβate fererasoβo
 t-ere-ju iβate je=r-era-so-βo
 HORT-2SG-come height 1SG=R₁-CAUS.SOC-go-GER
 ‘May you come in order to/and take me (with you) to heaven.’
 (Poemas, 102)
- c. Aike witupa
 a-ike wi-tup-a
 1SG-enter 1SG_{CORF}-lay.down-GER
 ‘I entered in order to/and laid down.’ (VLB, I, 18)

- (502) a. Pitangĩramo ereiko, Tupãnamo eikoβoβé
 pitang-ĩ-ramo ereiko Tupã-ramo e-iko-βo-βé
 child-DIM-TRSL 2SG-be God-TRSL 2SG_{CORF}-be-GER-also
 ‘You are a little child being also God.’ (Poemas, 100)
- b. Ise oromojasuk Tuβa, Taʔira,
 ise oro-mo-jasuk t-uβ-a, t-aʔir-a,
 I 1SG.2SG-CAUS-wash R₁-father-REF R₁-son-REF
 Espírito-Santo rera pupe wijáβo
 Espírito-Santo r-er-a pupe wi-ja-βo
 holy-ghost R₁-name-REF POSP 1SG_{CORF}-say-GER
 ‘I baptize you (while) saying: in the name of the father, of the son,
 and of the holy spirit.’ (DC, I, 200)
- c. Neporangatu rawsupa tekoaʔiβ
 ne=∅-poraŋ-katu r-awsuβ-a t-eko-aʔiβ
 2SG=R₁-beauty-INTS R₁-love-GER R₁-life-bad
 oromopo
 oro-mo-por
 1PL.EXCL-CAUS-jump
 ‘We throw away the bad life loving your great beauty.’ (Poemas,
 84)
- d. Oiporaraβépe maʔe amō aʔepe
 o-i-porara-βe=pe maʔe-∅ amō aʔe-pe
 3-R₂-suffer-also=Q thing-REF other DEM-LOC
 ojkóβone?
 o-eko-βo=ne
 3-R₂-live-GER=FUT
 ‘Will they suffer anything living there?’ (Araújo, 47)

It is not common cross-linguistically for purposive clauses and modifying sub-
 actions to be expressed by the same construction¹³. While some sentences are

¹³In his study of complement clauses, Schmidtke-Bode (2015, 69) mentions that in Matsés, for
 example, ‘apart from purpose, converbs may also be associated with other adverbial domains’.
 TUP is not alone in the Tupí-Guaraní family in its coding of purpose and simultaneity by the

in fact ambiguous, allowing both interpretations, such as (502c), (497), (498a), others clearly only allow one interpretation (502b), (502a), (501a).

RRG assumes that there is an iconic relation governing the interaction of syntax and semantics in clause linkage (see Van Valin Jr 2005, 205-213 and Van Valin Jr 2022). This relation is captured by the Interclausal Relations Hierarchy (IRH) (see Van Valin Jr 2022, 63-69), capturing the thematic relations of clause linkage. The coding of different types of interclausal semantic relations by the same construction and the same juncture nexus does not seem to be common. The construction with the gerund in a core-juncture described above codes the following semantic relations in Van Valin Jr (2022, 69): single actions (manner, position) (see e.g. 518), multiple actions (simultaneous, sequential), and intentions (purposive).

- a) Single actions
 - 1) Modifying sub-actions: manner
 - 2) Modifying sub-actions: position
- b) Multiple actions
 - 1) Simultaneous
 - 2) Sequential
- c) Intentions
 - 1) Purposive

TUP is very unusual in that it expresses different interclausal semantic relations through the same construction (nexus type).

It is also possible to have a core junction with more than one linked core, as in (503).

same converb construction. It is also the case in Old Guaraní (Montoya 1876, 16), Restivo (1724, 76,82,84,101); Kamajura (Seki 2000, 195-201); and Teko (Rose 2011, 359).

- (503) a. Ojeaiβik owasēasemamo, omanōḡatwaβo
 o-je-aiβik o-asē-REDUP-amo o-manō-katu-aβo
 3-RFLX-low.the.head 3_{CORF}-cry-cry-GER 3_{CORF}-die-INTS-GER
 koʔite
 koʔite
 finally
 ‘He lowered his head crying loud, finally dying.’ (Araújo, 63v)
- b. Angaipáβora ajuka ʃeratápe
 angaipaβ-βor-a a-∅-juka ʃe=ɾ-ata-∅=pe
 sin-HAB.AG-REF 1SG-R₂-kill 1SG=R₁-fire-R₁-LOC
 seroʔane serasóβo, iʔwaβo páne
 s-ero-ʔar=ne s-era-so-βo, i-ʔu-aβo paβ=ne
 R₂-SCAU-fall-FUT R₂-SCAU-go-GER R₂-eat-GER all=FUT
 ‘I will kill the sinners causing them to fall with me into my fire,
 leading them, eating them.’ (Teatro, 94)

The non-finite core found in core subordination is a gerund. The gerund is defined by Haspelmath (1995) as a ‘nonfinite verb form whose main function is to mark adverbial subordination’ (Haspelmath 1995, 3). According to Haspelmath (1995), the gerund (‘converb’ in his terminology) has the following characteristics (see also Tikkanen 2001):

1. It is inherently subordinate, i.e., involved in subordinate constructions
2. Non-finite: it lacks specifications for tense, aspect, and mood
3. Adverbial nature: it mainly functions as a modifier, not an argument, and modifies clauses, not RPs
4. It is marked by an affix
5. It is often used in constructions that are coreferential with the subject of another clause
6. It may be the focus of a polar question

The TUP converb is marked by the suffix $-(a)\beta o \sim -a^{14}$ or by the loss of a

¹⁴These forms are cognates, in spite of the apparent difference, see Schleicher (1998, 144), in

final *r*. This form has also been referred to as a gerund in the literature (see Jensen 1998a, Rodrigues 2011a and Aikhenvald 2012, 312-314).

Bohnmeyer and Van Valin Jr (2017) suggest that the Macro-Event-Property (MEP), which is a form-meaning mapping property, constrains the compatibility of event descriptions with time-positional modifiers. According to the MEP, the verbal core is the macro-event phrase. Simple cores have the MEP by default, while complex cores have it only in cosubordinate linkage. I have not found a clear example in TUP, parallel to the one in (504) from Bohnmeyer and Van Valin Jr (2017, 167), of the temporal peripheral modifier with scope over the complex core. There are examples of it preceding the complex core, but it seems that an intonational break is involved, indicating that the adverb is topicalized and thus belongs in the PrDP. The most frequent attestation places the time adverb between the linked cores, as in (505a). The representation of (505b) is given in Figure 10.15.

(504) [[[[[Chris went]CORE to [see Pat]CORE today]CORE]PERIPHERY]CLAUSE]SENTENCE

- (505) a. *ʃepeʔa maʔeaíβa swi kori Tupã*
ʃe=∅-peʔa maʔe-aiβ-a ∅-swi kori Tupã
 1SG=R₁-push.away thing-evil-REF R₂-from **today** God
remimotara rupi ʃemoiŋóβo
r-emi-potar-a r-upi ʃe=∅-mo-iko-βo
 R₁-RES-want-REF R₁-according 1SG=R₁-CAUS-be-GER
 ‘Push me away from evil things today in order to make me be according to the will of God.’ (DC, I, 190)
- b. *Jaso kori imomewaβo*
ja-so kori i-mo-meʔu-aβo
 1PL.INCL-go **today** R₂-CAUS-announce-GER
 ‘We go to announce him today.’ (Poemas, 110)

line with a proposal in Haspelmath (1995, 17). See also Jensen (1998a, 529-530), and Cabral and Rodrigues (2005).

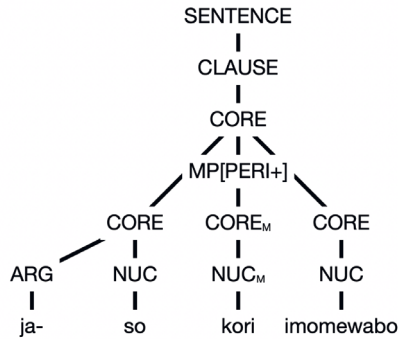


Figure 10.15: Complex core with periphery in cosubordination

Cosubordination with the gerund is commonly found with the permissive mood, as in (506a):

- (506) a. *Tasaβeipóne* *?ara* *mokajema...*
 t-a-saβeipo=ne ?ar-a mo-kajem-a
 HORT-1SG-get.drunk=FUT understanding-REF CAUS-hide-GER
 ‘May I get drunk in order to lose my understanding.’ (DC, II, 103)
- b. *Tafemaran* *umě iwaβo*
 t-fe=∅-maran umě i-?u-aβo
 HORT-1SG=R₁-ill NEG R₂-ingest-GER
 ‘May I not fall ill eating it.’ (Araújo, 21v)

The core juncture with the gerund is also attested expressing cause, although less frequently, as in (507), which is given as an answer to ‘Why is the cross the sign of the Christians?’

- (507) a. *Ipupe omanōmo janejara* *Jesus Cristo*
 i-pupe o-manō-βo jane=∅-jar-a Jesus Cristo
 R₂-in 3-die-GER 1PL.INCL=R₁-lord-REF Jesus Christ
 ‘(Because) Our Lord Jesus Christ died on it.’ (DC, I, 186)

- b. Jesus Cristo omenare?ima?ep^wera reko jaße
 Jesus Cristo o-menar-e?im-βa?e-p^wer-a r-eko jaße
 Jesus Christ 3-marry-PRIV-REL-PST-REF R₁-be like
 oikopota
 o-iko-pota
 3-be-want.GER
 ‘For wanting to live like Jesus Christ who also wasn’t married.’
 (DC, I, 224)

Although the finite core more often precedes the non-finite core in the texts, the opposite order, as in (508), is also possible and does not seem to be marked. Based on the texts and supported by a comparison with other TG languages, especially Old Guaraní, there is no doubt that the non-finite core more often followed the finite core¹⁵.

- (508) a. Wisoβo aso
 wi-so-βo a-so
 1SG_{CORF}-go-GER 1SG-go
 ‘I go in order to stay.’ (VLB, II, 41)
- b. Witu ajur
 wi-tu a-jur
 1SG_{CORF}-come.GER 1SG-come
 ‘I came to stay / I came and I stayed.’ (VLB, II, 41)

In (509), the comma seems to indicate a pause following the fronted core and the discourse particle, which would place the dislocated syntagma in the detached position. This fronting is not like the one in (508), but one to a more topical position (PrDP). Whether this (dislocation with a pause) is a calc of a similar construction in Portuguese cannot be asserted with certainty.

¹⁵In the verses by Anchieta (Anchieta 1997, 2006), the non-finite core more often precedes the matrix clause than in all other texts due to the poetic character of these texts, e.g., the search for rhymes. Another issue is the use of commas, which seems to follow its use in Portuguese, blurring the structure of the texts.

- (509) Nerep^laka pota jě, jaju kuwépe swi
 ne=r-ep^lak-a pota-∅ jě, ja-ju kuwepe ∅-swi
 2SG=R₁-see-REF want-GER PRCL 1PL.INCL-come far R₁-from
 ‘Wanting to see you, we came from afar.’ (Poemas, 96)

The presence of the discourse particle is not necessary for placing a constituent in the PrDP as long as it is set off by a pause, as in (510).

- (510) a. Tupana kuwapa, koʔi asausu fejara
 Tupana kuwaß-a, koʔi a-s-awsu je=∅-jar-a
 God know-GER now 1SG-R₂-love 1SG=R₁-master-REF
 Jesu
 Jesu
 Jesus
 ‘Knowing God, I now love my lord Jesus.’ (Poemas, 106)
- b. Tapujpepoʔi moripa, tupotareʔimi
 tapuj-pe-poʔi mo-oriß-a t-u-potar-eʔim-i
 slave-DAT-evil CAUS-amuse-GER R₂-come-want-PRIV-NFOC
 ike
 ike
 here
 ‘Amusing themselves with slaves, they did not want to come here.’
 (Teatro, 16)

In (511), the non-finite core in each example is in the PrCS as the focus of a polar question, and thus within the scope of the IF operator.

- (511) a. Ejemomewáβope erejur feraʔit?
 e-je-momeʔu-aβo=pe ere-jur je=r-aʔir
 2SG_{CORF}-RFLX-confess-GER=Q 2SG-come 2SG=R₁-son.VOC
 ‘Is it in order to confess that you came, my son?’ (DC, II, 77)
- b. Marā ojáβope irajtítataendi meʔéʔi ase
 Marā o-ʔi-aβo=pe irajti-tata-endi-∅ meʔéʔ-i ase
 what 3-say-GER=Q wax-fire-shine-REF give-NFOC PRON

pópe?
 pó-∅-pe
 hand-REF-POSP

‘By saying what he puts a candle in our hands?’ (DC, I, 204)

Since core junctures are within the scope of the clausal operators, as with tense, there can be only one IF operator. In (512a), the IF is interrogative, and it is imperative in (512b):

- (512) a. Erépe nerekoposip^we momoraŋa?
 er-ʔe=**pe** ne=r-eko-pofji-p^wer mo-poraŋ-a
 2SG-say-Q 2SG=R₁-be-evil-PST CAUS-beauty-GER
 ‘Did you say it while celebrating your evil behavior?’ (cf. DC, II, 93)
- b. Ejori orerese nememira moŋetaβo
 e-jori ore=r-ese ne=∅-memir-a moŋeta-βo
2SG.IMP-come 1SG.EXCL R₁-son-REF talk-GER
 ‘Come, in order to talk to your son about us.’ (Poemas, 82)

Since the linked cores are inside the clause, they depend on the clause for tense. In (513), the clause is overtly marked for tense:

- (513) Tameʔēne pira ruβa enéβo,
 t-a-meʔēŋ=**ne** pira-∅ r-uβ-a ene=βo,
 HORT-1SG-give=**FUT** fish-REF R₁-egg-REF 2SG=DAT
 wijepimeʔeŋa
 wi-je-epi-meʔeŋ-a
 1SG_{CORF}-RFLX-pay-give-GER
 ‘May I give you fish eggs to repay (you).’ (Teatro, 46)

Tense can be marked on the finite verb, as in the examples above. In (514), it appears attached to both verbs. This is a very rare case and should not be taken as evidence that the non-finite core may be independently marked for case. It seems, based on the translation of the example, that the second future simply emphasizes the tense already marked on the main verb.

- (514) Aʔereme amõ ajukáne, ʃejusanime
 aʔe-reme amõ a-i-juka=**ne** ʃe=∅-jusan-pe
 this-after other 1SG-R₂-kill-FUT 1SG=R₁-lasso-LOC
 imoʔa, enéβo imeʔeŋáβone
 i-mo-ʔa ene=βo i-meʔeŋ-áβo=**ne**
 R₂-CAUS-fall.GER 2SG-DAT R₂-give-GER-FUT
 ‘I shall kill some afterwards, making them fall in my lasso, giving them to you.’ (Poemas, 157)

The future marker appears most frequently attached only to the gerund, in sentence-final position, as in (515). Only a discourse particle may follow the tense marker in sentence-final position, as in (515b).

- (515) a. Ere sep^jakane
 ere-ʔe s-ep^jak-a=**ne**
 2SG-say R₂-see-GER=**FUT**
 ‘You shall see and believe (lit. you will say, seeing it).’ (FA, 159)
- b. Aʔe ipo maʔe repiramo ojmoariβeukar
 aʔe ipo maʔe-∅ r-epi-ramo o-i-mo-ariβe-ukar
 PRCL ADV thing-REF R₁-payment-TRSL 3-R₂CAUS-quiet-CAUS
 ʃeswi, ʃemop^werápane rea
 ʃe=∅-swi ʃe=∅-mo-p^weraβ-a=**ne** rea
 1SG=R₁-from 1SG=R₁-CAUS-heal-GER=**FUT** PRCL
 ‘Certainly, he will spare me as a reward of something, healing me.’
 (Araújo, 96v)

There are some apparent cases where a nominalized core takes the future clitic without the presence of a finite verb, but this is due to the presence of particles that require the gerund, like *memetipo* in (516a) and *ka* in (516b):

- (516) a. Memetipo ase isupe ojerokiaβone?
 memetipo ase i-supe o-je-roki-**aβo=ne**
 even.more we R₂-to 3-RFLX-bow-**GER=FUT**
 ‘Will we indeed bow to her?’ (Araújo, 31)

- b. Teʔijñene ojkóβo ka!
 Teʔijñe=**ne** o-eko-βo ka
 PRCL=**FUT** 3-be-GER PRCL
 ‘Let him be!’ (VLB, I, 92)

The only possible instance of clausal cosubordination I found is given in (517), where the tense operator seems to be shared by both clauses.

- (517) Anojan, aroβeβene
 a-**ero**-jan a-ro-βeβe=**ne**
 1SG-SC-run 1SG-SC-fly=**FUT**
 ‘I will make them run with me, will make them fly with me.’ (AT, 42)

Core-subordination is often used to indicate the position of actors while performing an action, i.e., whether they are standing, lying down, or sitting. Such constructions have often been said to involve positional auxiliaries, but they are no different from any other gerund nominalized core involved in core subordination. TUP does not have grammaticalized auxiliaries, as all positional auxiliaries are lexical roots. The lexical roots that may indicate the position of actors in TUP are (*j*)uβ ‘lay/lie’, (518), *iko* ‘be, be moving’ (519), *in* ‘sit down’ (520), *kuβ*¹⁶ ‘be’ (without reference to posture) (521).

- (518) a. Asasaβ pe witupa
 a-s-asaβ pe-∅ wi-tup-**a**
 1SG-R₂-cross path-REF 1SG_{CORF}-lay.down-GER
 ‘I occupy the path (lying down).’ (VLB, I, 47)
- b. Ajepiso witupa
 a-je-piso wi-tup-**a**
 1SG-RFLX-stretch 1SG_{CORF}-lay.down-GER
 ‘I am stretched (while laying down).’ (VLB, I, 129)
- c. Erekepipo ejupa?
 ere-ker=pe-ipo e-juβ-a
 2SG-sleep=Q-really 2SG_{CORF}-lie.down-GER

¹⁶*Kuβ* is only used with plural subjects.

‘Were you really sleeping?’ (Teatro, 12)

- (519) a. *feresaraj e witekoβo*
ʃe=r-esaraj e wit-eko-βo
 1SG=R₁-forget PRCL 1SG_{CORF}-be-GER
 ‘I in fact forget it (while being/doing something).’ (AP, 182)

- b. *Ajeroβjapirib witekoβo*
a-je-roβja-pirib wit-iko-βo
 1SG-RFLX-arrogant-somewhat 1SG_{CORF}-be-GER
 ‘I am (being) somewhat arrogant.’ (VLB, I, 33)

- (520) *Ajemopeβ witena*
a-je-mo-peβ wit-in-a
 1SG-RFLX-CAUS-flat 1SG_{CORF}-sit-GER
 ‘I am squatting.’ (VLB, I, 23)

- (521) a. *Nepo wiriβo paβẽ torojenoŋ*
ne=∅-po-∅ wiriβo paβẽ t-oro-je-noŋ
 2SG=R₁-hand-REF POSP all HORT-1PL.EXCL-RFLX-put
orojupa nememiramo
oro-juβ-a ne=∅-memir-amo
 1PL.EXCL-lay.down-GER 2SG=R₁-child-TRSL
orokupa
oro-kup-a
 1PL.EXCL-be-GER
 ‘May we (laying down) put ourselves under your hands, being as your children.’ (Poemas, 148)

- b. *Tjaso ke jawapika*
t-ja-so ke ja-wapik-a
 HORT-1PL.INCL.CORF-go PRCL 1PL.INCL.CORF-sit-GER
jakupa
ja-kuβ-a
 1PL.INCL.CORF-be-GER
 ‘Let’s remain seated (lit. let us go sitting being).’ (Teatro, 146)

10.3 Complex RPs

Complex RPs may contain complex modifiers such as genitives, possessives, and relative clauses as part of modification (Dryer 2007, 151; Van Valin Jr and LaPolla 1997, 492). These complex modifiers co-occur with simple modifiers in RPs, as the previous sections have shown.

The structural similarities between the LSC and the LSRP are strengthened by the application of the theory of clause linkage, juncture, and nexus (see Chap. 10) to the analysis of complex RPs. This allows the analysis of complex RPs to reflect that of complex sentences, although there are fewer RP junctures because the RP has fewer layers than the clause. The RP level is the maximal layer, followed by the $CORE_R$ and the NUC_R which make up the RP junctures. The three nexus relations – coordination, subordination, and cosubordination – can be applied to the analysis of complex RPs. The RP layer is compatible with all three juncture-nexus types.

10.3.1 Coordination

Coordination involves the linking of two or more coordinands, which are independent units that may be joined by coordinators (syndetic) or by simple juxtaposition (asyndetic) (Haspelmath 2004). Distinctions are usually made between three semantic types of coordinate constructions: conjunctive coordination (additive), disjunctive coordination, and adversative coordination.

Syndetic conjunctive coordination is expressed with the additive conjunction/adverb *aβé* ‘and, also’¹⁷, which occurs after the coordinands. The position of this coordinator, [A] [B co], where *A* and *B* stand for two coordinands and *co* stands for the coordinator, is the least frequent type cross-linguistically according to Haspelmath (2004, 6). *Aβé* only connects RPs, never clauses.

The examples in (522) are cases of RP-level coordination, since no operator

¹⁷Mithun (1988) suggests that additives are often grammaticalized to produce conjunction, claiming that conjunctions also frequently develop from an adverbial particle meaning ‘also, too, as well’. Further evidence is provided by Haspelmath (2004), according to whom coordinators often have other meanings/functions besides the function of marking coordinating constructions.

dependence is involved:

- (522) a. [S. Pedro]_{coordinand}, [São João]_{coordinand} [aβé]_{coordinator}
 S(aint) Pedro Saint João **aβé**
 S(aint) Peter Saint John **CONJ**
 ‘Saint Peter and Saint John.’ (AC, 55)
- b. Iawasa-βaʔe omenasaβeʔima
 [i-awasa-βaʔe]_{coordinand} [o-mena-saβ-eʔim-a
 R₁-concubinage-NMLZ COREF-spouse-NMLZ-NEG-REF
 rese ojkoβaʔe aβe
 r-ese o-jko-βaʔe]_{coordinand} [aβé]_{coordinator}
 R₁-POSP 3-be.with-NMLZ **CONJ**
 ‘The one in concubinage and the one who lives with whom is not his/her spouse.’ (Araújo, 71)
- c. Enéte, neresemõ [arijama]_{coordinand},
 ene-te ne=r-esemõ-∅ arijama
 2SG-FOC 2SG=R₁-left.over-REF bird
 [tajasu]_{coordinand}
 tajasu
 peccary
 ‘To you, though, there remain birds and peccaries.’ (AP 152)

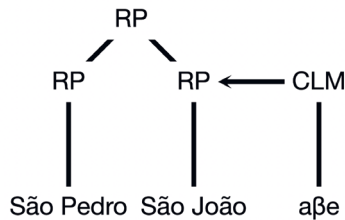


Figure 10.16: RP coordination

A reduced form of *aβe*, *-βe*, can be suffixed to a word. This was possibly used as a prosodical resource. The syntactic representation is given in Fig. 10.16.

- (523) a. Tosarō paʔi Jesu ʃeretama, neaβé
 t-o-sarō paʔi Jesu ʃe=r-etam-a, ne aβé
 HORT-3-keep lord Jesus 1SG=R₁-land-REF 2SG CONJ
 ‘May lord Jesus watch over my land and you as well.’ (Poemas, 112)

It is also common not to use a conjunction, i.e., simple juxtaposition, as in (525):

- (524) [Moropotara]_{coordinand}, [tesajnana]_{coordinand}, [marãʔe]_{coordinand},
 moro-potar-a t-esajnan-a marã-ʔe
 ANTIP-desire-REF R₂-lust-REF evil-say
 [mosarō]_{coordinand} [moʔema]_{coordinand} βe]_{coordinator}
 monarō-∅ moʔem-a βe
 theft-REF lie-REF CONJ
 ‘Sexual desire, lustful malediction, theft, and lies.’ (AT, 150)

- (525) Kaʔu, awasanemp^wera, temoʔema, marã
 kaʔu-∅ awasa-nem-p^wer-a t-emoʔem-a marã
 beer.drink-REF concubinage-stench-PST-REF R₂-lie-REF evil
 ʔe, joapisaβa, maranwera
 ʔe-∅ jo-api-saβ-a maran-p^wer-a
 say-REF REC-hit-NMLZ-REF war-PST-REF
 ‘Beer drinking, the old fetid concubinage, the lies, the ill saying, the mutual wounds, the old wars.’ (AT, 190)

- (526) Tosarō paʔi Jesu ʃeretama, neaβe
 t-o-sarō paʔi Jesu ʃe=r-etam-a ne-aβe
 HORT-3-keep lord Jesus 1SG=R₁-land-REF 2SG-CONJ
 ‘May lord Jesus watch over my land and you as well.’ (Poemas, 112)

The adverb/conjunction *aβe* never links clauses, only RPs or PPs. An examples of its linking to two PPs is given below:

- (527) Oka ripijawáma rese ajaɲa
 ok-a r-ipija-wám-a r-ese ajaɲ-a
 house-REF R₁-sprinkle-FUT-REF R₁-POSP devil-REF
 mojewasemawama reseβé
 mo-jewasema-wam-a r-ese-βé
 CAUS-flee-FUT-REF R₁-POSP-CONJ

‘For sprinkling the house and for expelling the devil.’ (DC, I 222)

Disjunctive coordination presents the elements in the construction as being alternatives to each other. Disjunctive coordination in TUP is either expressed using *kojpó* or *konipo*¹⁸. Disjunctive coordination appears to be similar to alternative questions (Croft 2022a, 436), so they are often co-expressed. Table 10.1 shows the frequency of both forms in the texts¹⁹. An interesting pattern emerges: while *kojpó* is more frequent in most texts, it is significantly less frequent in DC I, with *konipo* 85%, and *kojpó* 14%.

Source	Kojpo	konipo
Poemas	0	0
Teatro	1	1
DC I	1	6
DC II	9	0
Araújo	102	2
Camarões (1, 2, 3)	2	0
Betendorf	19	0
Total	134	9

Table 10.1: Frequency of *kojpo* and *konipo*

¹⁸The VLB (II, 60) also gives *komonipo*, which is not attested in any text.

¹⁹Nheengatu has substituted the TUP form(s) with *u* ‘or’ (from Portuguese *ou*) (Cruz 2011). *konipo* – like *komonipo* – occurs less frequently in the texts, perhaps because it is longer, with the exception of Anchieta (1618a) (see Table 10.1). Língua Geral does not know *konipo*, which fell into disuse; it only attests *kojpó* (e.g. Muller et al. 2019, 27,211).

- (528) a. Ene konipo ise?
 ene **konipo** ise
 you **or** I
 ‘You or I?’ (VLB, II, 60)
- b. Anira ruāpe e, panama kojpo wajkuika?
 anira-∅ ruā=pe e panama-∅ **kojpo** wajkuika-REF
 bat-REF PRCL=Q PRCL butterfly-REF **or** opossum-REF
 ‘Is it, actually, a bat, a butterfly, or an opossum?’ (AT, 44)
- c. Mamō serā fesōwne, konipo iβakipe,
 mamō serā je=∅-so-u=ne **konipo** iβaki-pe
 where by.the.way 1SG=R₁-go-NFOC=FUT **or** sky-LOC
 konipo Ajanga r-ata-pe-no?
konipo Ajanga rata-pe-no
or Devil R₁-fire-LOC-PRCL
 ‘Where, by the way, shall I go? To heaven or to the devil’s fire?’
 (DC, I, 221)
- d. Eresuñápe nerige nememira
 Ere-s-uñá=pe ne=r-ge-∅ ne=∅-memir-a
 2SG=R₁-touch=Q 2SG=R₁-womβ-REF 2SG=R₁-son-REF
 jukáβo ijukapota? Kojpo ere?upe ma?e
 juka-βo i-juka-pota **Kojpo** ere-?u=pe ma?e-∅
 kill-GER R₂-kill-want.GER **or** 2SG=ingest=Q thing-R₁ other
 amō tomanō jeswi ejaβo?
 amō t-o-manō je=-∅-swi ?e-aβo
 HORT-3-die 1SG=R₁-POSP say-GER
 ‘Did you touch your belly killing your son or did you ingest some-
 thing saying may he die coming out of me?’ (AC, 102)

In (529), the lexical root *apiatā* modifies *kujā* with a non-restrictive nominalization (RP periphery) – assuming the hearer knows who is being referred to – in the nuclear periphery, which modifies the RP as a whole since it, like non-restrictive clauses, has independent IF (see Van Valin Jr 2022, 41). The representation of (529) is given in Figure (10.17). This is a subordinate RP construction.

- (529) Kujāpiatā iβakipenwara
 kujā-piatā-∅ iβak-pe-nwar-a
 woman-strong-REF sky-LOC-NMLZ-REF
 ‘A strong woman who is in heaven.’ (Poemas, 126)

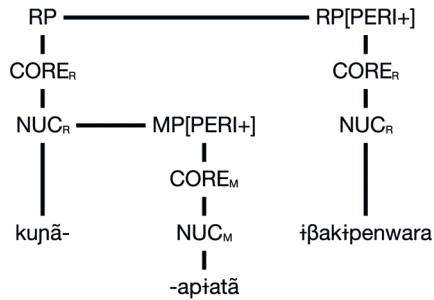


Figure 10.17: RP subordination

Example (530) shows an RP with a proper noun, which has no layered structure (cf. Van Valin Jr 2005, 222). It has an RPIP and two peripheries, and is modified by an RP which contains an RPIP with conjoined periphery units, resulting in a coordinate RP juncture-nexus type. Its syntactic representation is given in Figure 10.18. The RP modifying the RP with the proper noun is similar to a non-restrictive relative clause because it adds information about the head noun. This is a case of RP subordination because the RP *paʔi Tupã rawsupara* functions as an adjunct at the RP level.

- (530) Ako Ana wajβĩ rainha, paʔi Tupã rausupara
 ako Ana wajβĩ rainha paʔi Tupã r-awsuβ-ar-a
 DEM Ana old queen lord God R₁-love-NMLZ_{SG}-REF
 ‘That old queen Ana, who loves God / lover of God.’ (Teatro, 168)

10.3.2 Relative clauses

A typical example of RP subordination is the restrictive relative clause, in which a clause is used as a restrictive modifier of an RP. The modifying clause

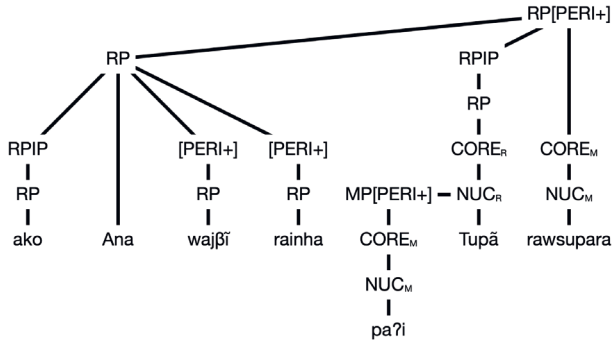


Figure 10.18: RP subordination

is hosted in the periphery, since it is an optional modifier. In TUP, since the head noun is outside the relative clause, the relative clause is externally-headed (see De Vries 2002).

As seen in Section 8.3, the nominalizer *-βaʔe* is often used as an RP modifier (relativizer) which restricts the interpretation of the RP head, as in (531). The nominalized clause *oporomonijβaʔe* is a peripheral modifier of the RP *t-eko*, which is an argument of the predicate *s-eta*. The representation of (531) is shown in Figure (10.19).

- (531) Seta teko oporomonijβaʔene
 s-eta t-eko-∅ o-poro-monij-βaʔe=ne
 R₂-many R₂-be-REF 3-ANTIP-frighten-NMLZ_{REL}=FUT
 ‘There are be many things that will frighten us.’ (Araújo, 159v)

The scope of the clitic =ne in (531) is over the main predicate only. If the nominalized clause were to be marked for tense, it would receive the nominal future-tense marker *-ram* (see Section 7.3.1), as in (532).

- (532) Aroβjar tekoβe opaβaʔerameʔima
 a-roβjar t-ekoβe-∅ opa-βaʔe-ram-eʔim-a
 1SG-believe R₂-life-REF finish-NMLZ_{REL}-FUT-PRIV-REF
 ‘I believe in the life that will not end.’ (DC, I, 142)

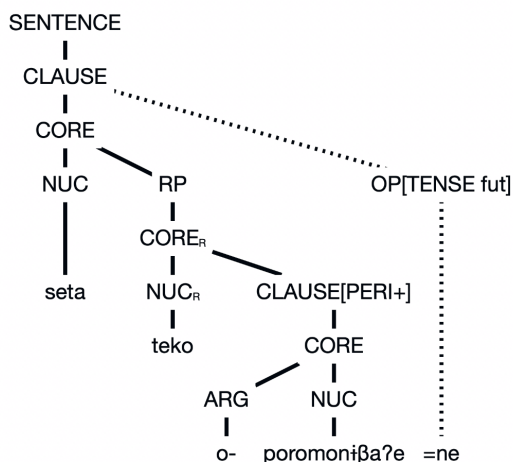


Figure 10.19: Relative clause or RP subordination

10.3.2.1 Restrictive relative clauses

A relative clause²⁰ is a clause inside an RP that provides more information about the referent of the head of the RP (restrictive). The head of an RP is coreferent with an argument of the predicate within the relative clause (henceforth RC), but that argument is ‘missing’ from the RC and the head noun controls its identity. This type is not available in Tupinambá, since the modifier of the RP-head is always a nominalized clause (also an RP).

Restrictive relative clauses, as the name suggests, restrict the interpretation of the referent of the matrix clause. They are non-argument, peripheral modifiers of the nominal and thus a case of nuclear_N subordination. The modifier in this case is a clause nominalized by *-βa?e*²¹. This is illustrated in (533) with its

²⁰This section discusses only restrictive relative clauses, which is the more common type; the term ‘relative clause’ will be taken to mean ‘restrictive relative clause’ unless specified otherwise. Non-restrictive relative clauses give extra information that is not necessarily needed to identify the referent, and may follow a pause: e.g., *my brother, who lives in Michigan, is older* (I have only one brother). This contrasts with *I texted my sister who lives in Canada*. (I have other sister(s), but I am talking about the one who lives in Canada).

²¹The similarity of the clause nominalizer *-βa?e* with the word *-ma?e* ‘thing’ allows for a

syntactic representation given in (10.20).

- (533) Pitaŋa mokōj roʔi omoawieβaʔe
 pitaŋ-a mokōj roʔi-∅ o-mo-awie-βaʔe
 child-REF two year-REF 3-CAUS-complete-NMLZ
 ‘The children who complete two years (of age).’ (Araújo, 10v)

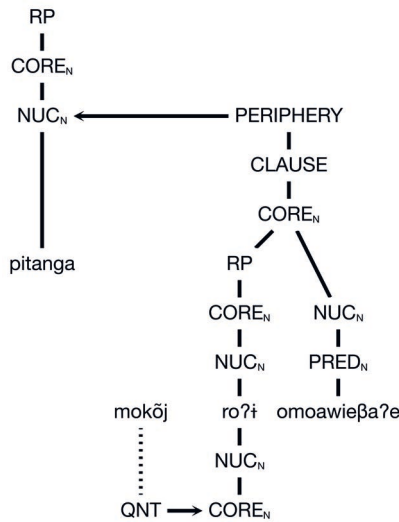


Figure 10.20: Restrictive relative clause

The relativization strategy employed in Tupinambá is nominalization, a strategy not accounted for either by formalists (De Vries 2002; Keenan and Comrie 1977, 1979; Comrie and Keenan 1979) or by functionalists (Cristofaro 2005; Song 2014) (see Lehmann 1984, 149-153).

conjecture regarding the origin of the nominalizer. The nominalization path from ‘thing’ to nominalizer is known from other languages (see Kuteva et al. 2019, 433-434). The use of nominalization as a strategy for subordination ‘is significantly more pervasive in South America than would be predicted on the basis of global patterns’ (Van Gijn 2014, 274)

- (534) Omemira Tupā apiaβamo gwigépe
 O-memir-a Tupā apiaβ-amo o-ge=pe
 3CORF-son-REF God man-TRSL 3CORF-womβ-POSP
 ojemojaŋjiβaʔe ʔarama osepiʔakaʔuβ.
 o-je-mojaŋ-βaʔe ʔa-ram-a o-s-epiʔak-aʔuβ
 3-RFLX-make-NMLZ_{REL} birth-FUT-REF 3-R₂-see-satisfy
 ‘(Maria) wishes earnestly to see the birth of her own son who generates himself as a human being in her womb.’ (Araújo, 9-9v)
- (535) iβakipe karaiβeβe maraŋatuβaʔe
 iβak=pe karai-βeβe-∅ maraŋatu-βaʔe
 sky-POSP white.person-fly-REF kindness-NMLZ_{REL}
 opitaβaʔep^wera ruβisaβa
 o-pit-βaʔe-p^wer-a r-uβisaβ-a
 3-stay-NMLZ_{REL}-NPST-REF R₁-chief-REF
 ‘Chief of the angels that are good, that remained in heaven.’ (Araújo, 8v)
- (536) Aŋwera aʔepe turrama osarōβaʔe
 aŋ-wer-a aʔepe t-ur-ram-a o-s-arō-βaʔe
 soul-PST-REF there R₂-come-FUT-REF 3-R₂-wait-NMLZ
 ‘The souls that waited there for his future return.’ (DC, I, 150)

Regarding the accessibility hierarchy indicating which RP positions can be relativized, (Keenan and Comrie 1977; Comrie 1981), some examples are given below which are illustrative of the possibilities in TUP.

- (537) a. Subject relativized
- | | | |
|---|-------------|---------------------------|
| Apiaβa kujā | rese | oeko |
| apiaβa kujā-∅ | r-ese | o-eko |
| man | woman-REF | R ₁ -WITH 3-be |
| osaʔaŋjepéβaʔe | | neʔikatui omena |
| o-s-aʔaŋ-jepe-βaʔe | | n-eʔi-katu-i o-mena |
| 3-R ₂ -attempt-in.vain-NMLZ _{REL} | NEG-can-NEG | CORF-marry |

‘The man who tries in vain to have sexual intercourse with a woman cannot get married.’ (AC, 131v)

- b. Aokerejua k^warasi sose oβeraβaʔe
 ao(β)-kerejua k^warasi-∅ -∅-sose o-βeraβ-βaʔe
 clothes banded.cotinga sun-REF R₁-above
 nuɟara
 nuɟara
 3-shine-NMLZ_{REL} similar

‘Similar to a garment made of a banded cotinga feathers²² that shines more than the sun.’ (AC, 37v)

(538) Apposition

- a. iβi apiteripe tataamõrore
 iβi-∅ apiteri-pe tata-amõrore-∅
 earth-REF middle-LOC fire-eternity-REF
 oweβaʔerameʔima mojanga
 o-weβ-βaʔe-ram-eʔim-a mojaɟ-a
 3-extinguish-NMLZ-FUT-PRIV-REF make-GER

‘In the middle of the earth making an eternal fire which won’t be extinguished.’ (AC, 38)

- b. iβitira Oliuete seriβaʔe apira
 iβitir-a Oliuete s-er-i-βaʔe apir-a
 mount-REF Olivet R₂-name-EPEN-NMLZ_{REL} ridge-REF
 ʔariβo osi oβoja
 ʔar-i-βo o-si-∅ o-βoja-∅
 over-EPEN-TRSL CORF-mother-REF CORF-disciple-REF
 rerasow
 r-era-so-w
 R₂-SCAU-go-NFOC

‘He took his mother and his disciples above the peak of the mountain called Oliuete.’ (AC, 4v)

(539) Direct Object relativized

²²Cotinga maculata is a species of bird from southeastern Brazil whose feathers shine intensely.

- a. Amō ʔiβa wemitima piteripe oʔambaʔe
 amō ʔiβ-a o-emitim-a piteri-pe o-ʔam-baʔe
 other tree-REF CORF-garden-REF middle-LOC 3-stand-NMLZ
 kuaβeʔenga
 kuaβeʔeŋ-a
 show-GER
 ‘Showing him a certain tree that stood in the middle of his garden.’
 (AC, 39v-40)
- b. Aroβjar Tupã Tuβa, opakatumbaʔe tetiruã
 a-roβjar Tupã T-uβ-a, opa-katu-mbaʔe-∅ tetirwã
 1SG-believe God R₄-father-REF all-INTS-thing-REF any
 mojaŋa eʔikatuβaʔe
 mojaŋ-a e-ʔi-katu-βaʔe
 make-GER 3-say-INTS-NMLZ_{REL}
 ‘I believe in God the father, the one who can do all and anything.’
 (Ar., Cat., 14v)

- (540) Niporangibəʔe ruã aʔe tata. sun,
 n-i-poraŋ-i-βaʔe ruã aʔe tata-∅ s-un,
 NEG-R₂-beauty-EPEN-NMLZ_{REL} NEG DEM fire-REF R₂-dark
 ipofɪ, oporoapieteβaʔe
 i-pofɪ o-poro-api-ete-βaʔe
 R₂-ugly 3-ANTIP-burn-EMPH-NMLZ_{REL}
 ‘That fire isn’t beautiful. It is dark, it is ugly, it is the one which
 intensively burns people.’ (AC, 163v)
- (541) Orojerure βe nerese tojemeʔeŋ
 oro-je-rure βe ne=r-ese t-o-je-meʔeŋ
 1PL.EXCL-ask ADV 2SG=R₁-because HORT-3-MID-give
 apiaβaŋaturama oreretama pora ri,
 apiaβ-aŋaturam-a ore-r-etam-a por-a ri,
 man-good-REF 1PL.EXCL-R₁-land-REF inhabitant-REF POSP
 paʔijemoʔesaβa βe Tupã rese iʔekatuβaʔe
 paʔi-jemoʔesaβ-a βe Tupã r-ese i-ʔekatu-βaʔe
 priest-wise-REF also God R₁-POSP R₂-worth-NMLZ_{REL}

‘We also ask you that they give us good men as inhabitants of our land, and wise priests who know God’s law.’ (D’Abbeville, Histoire, 342)

Chapter 11

Conclusion

The previous chapters have attempted to describe the grammar of TUP in a typologically adequate manner while making use of a modern linguistic framework. I have tried to avoid what I consider to be a frequent mistake in the past, where authors describing TG languages seem to have used previous descriptions of TG languages as templates.

In the introduction is a short *prolegomena* to the study of TUP, paving the way for anyone interested in the topic. It contains information on the people and the language, including a short typological profile, as well as details of the primary sources used and a short summary of previous work on the language. It also contains some comments of a socio-linguistic and ethnographic nature regarding the gender-exclusive distinctions in male and female dialects.

The chapter on phonology was a first attempt to review Rodrigues' work on the topic, which was not accessible to many because it was written in German. Furthermore, I have tried to improve the phonological description based on my knowledge of the language, which resulted in a phonological inventory somewhat different than that found in previous works, e.g., with the treatment of the phones [ʃ] and [ɲ] as surface representations. I have also attempted to provide an overview of syllabic and stress patterns by reviewing Rodrigues

(1958b).

In Chapter 3, I briefly introduced the framework employed to describe the language, RRG, as a syntax-semantics-pragmatics interface. While languages of different language families have been described using RRG, this is the first South American language to be described within this framework.

The morphological and syntactic discussions are spread from chapters 4 to 10 and shed some light on the morphosyntax and typology of TUP. Some topics were not discussed in depth because they would require a change in the course of the work, leading to a discussion of theoretical aspects of language or typology, which would be less focused on the language itself. This should not be interpreted as an excuse; rather, it is a consequence of the somewhat small corpus of the language, which is restricted to religious texts.

Particularly important is the treatment of word classes in Chapter 4, in which, instead of arguing for the existence or lack of a certain word class, I adopt a typological approach based on the idea of comparative concepts, which allows for an approach to word classes that is language-specific but still captures cross-linguistic generalizations. This approach depends on the idea of speech act functions combined with semantic classes, thus allowing language-specific constructions to define word classes. This is in line with current typological and constructional approaches, although the idea is not a recent one.

Another significant morphosyntactic aspect of this work lies in the treatment of bound indexes exclusively analyzed as possessor indexes, never as absolutive markers, as has been suggested for Tupí-Guaraní languages. As a consequence of the non-existence of absolutive markers, no person hierarchy must be assumed.

In Chapter 5, I laid out the basics of clause patterns and valency operations affecting it. A more in-depth analysis of the clause is presented in Chapter 6. Chapter 7 considers lexical categories and further morphosyntactic topics. The reference phrase is discussed in Chapter 8 and is followed by a chapter on information structure, discussing the pragmatics of some constructions. In this chapter, I argue against the so-called indicative-II or oblique-topicalized interpretation of a characteristic TG construction, that of fronted adverbials.

Here, it is analyzed as a construction in which the main predicate loses its focal status as an adverbial expression and is fronted to a topical position, thus becoming the main predicate. Finally, Chapter 10 presents the types and levels of clause juncture in TUP, a topic that is rarely treated in grammars of TG languages. The discussion of complex sentences has shown, among other things, that a particular type of clause linkage, cosubordination, is very common in TUP, especially at the core level, with nominalized cores (gerund constructions).

I certainly do not regard this description as complete. I would like to have discussed some aspects that require a digitally available corpus, enabling a quantitative analysis of phenomena such as word order, information structure, and variations in patterns. In this sense, I wait for the completion of the online Tupinambá treebank, available through the Universal Dependencies Project (UD), which would greatly contribute to this research. The treebank would also allow for the inclusion of TUP not only in comparative studies within the Tupían family, but also within other language families on UD.

Bibliography

- Adam, L. (1879). *Du Parler des hommes et du parler des femmes dans la langue Caraïbe, par Lucien Adam...* Maisonneuve. 26
- Aikhenvald, A. Y. (2004). *Evidentiality*. Oxford University Press. 197, 207
- Aikhenvald, A. Y. (2012). *Languages of the Amazon*. Oxford University Press. 23, 24, 25, 26, 211, 353
- Aikhenvald, A. Y. (2016). *How gender shapes the world*. Oxford: Oxford University Press. 26
- Aissen, J. (1997). On the syntax of obviation. *Language* 73(4), 705–750. 324
- Aissen, J. (1999). Markedness and subject choice in optimality theory. *Natural Language & Linguistic Theory* 17(4), 673–711. 130
- Almeida, A., I. de Jesus, and L. G. de Paula (1983). *A língua Tapirapé*. XEROX do Brasil. 314
- Altman, M. C. F. S. (2003). As línguas gerais sul-americanas e a empresa missionária: linguagem e representação nos séculos XVI e XVII. In J. R. Bessa Freire and M. C. Rosa (Eds.), *Línguas gerais: política lingüística e catequese na América do Sul no período colonial*, pp. 57–84. EdUERJ. 20

- Anchieta, J. d. (1933). *Cartas, informações, fragmentos históricos e sermões*. Civilização Brasileira. 10, 12, 16
- Anchieta, J. d. ((1933)[1595]). *Arte de gramática da língua mais usada na costa do Brasil*. Imprensa Nacional. 6, 19, 21, 22, 23, 30, 33, 41, 48, 101, 108, 111, 123, 145, 146, 167, 179, 214, 216, 220, 221, 225, 239, 243, 247, 250, 260, 281, 319, 321, 323
- Anchieta, J. d. (1948). *Auto representado na festa de São Lourenço*. Museu Paulista. Peça trilingue do séc. XVI, transcrita, comentada e traduzida, na parte tupi, por M. de L. de Paula Martins. 30
- Anchieta, J. d. (1988). *Diálogo da Fé*. São Paulo: Edições Loyola. Organização, tradução e notas do Pe. Armando Cardoso. 30
- Anchieta, J. d. (1992[1618]a). *Doutrina cristã: catecismo brasílico. Introdução e notas do padre Armando Cardoso, SJ*, Volume 1. Edições Loyola. 6, 22, 30, 33, 364
- Anchieta, J. d. (1992[1618]b). *Doutrina cristã: doutrina autógrafa e confessionalário*, Volume 2. Edições Loyola. 6, 30, 33
- Anchieta, J. d. (1997). *Poemas: lírica portuguesa e tupi*. Martins Fontes. 6, 10, 30, 33, 142, 309, 355
- Anchieta, J. d. (2006). *Teatro: Auto de São Lourenço-auto Na Aldeia de Guaraparim*. Martins Fontes. 6, 12, 14, 30, 32, 142, 234, 310, 355
- Anonymous (1750). Gramática geral da língua do brasil, com um dicionário dos vocábulos mais usuais para a inteligência da dita língua (gramatica da língua geral do brazil, com hum dicionario dos vocabulos mais uzuaes para a intelligencia da dita lingua). Manuscript. University of Coimbra (UCBG Ms. 69). 22
- Anonymous (1938). *Vocabulário da língua brasílica*. Departamento de Cultura. edited by Plinio Ayrosa. 30, 35
- Anonymous (1952a). *Vocabulário da língua brasílica*. (second ed.). Universidade de São Paulo. edited by Carlos Drumond. 11, 30, 35, 202, 244, 245, 323, 324

- Anonymous (1952b). *Vocabulário da língua brasílica. 2º vol. (IZ). Revista e conferida com os MS. Fg. 3144 da Bibl. Nacional de Lisboa por Carlos Drumond*. Universidade de São Paulo. 30, 220
- Anward, J., E. Moravcsik, and L. Stassen (1997). Parts of speech: A challenge for typology. *Linguistic Typology 1*, 167–184. 97, 98
- Aragon, C. C. (2014). *A grammar of Akuntsú, a Tupían language*. Ph. D. thesis, University of Hawai i,, at Mānoa. unpublished PhD thesis. 155
- Araújo, A. d. (1898[1618]a). *Catecismo brasílico da doutrina cristã, publicado por Julius Platzmann*. Teubner. 30, 35
- Araújo, A. d. (1952[1618]b). *Catecismo na língua brasílica*. Pontifícia Universidade Católica do Rio de Janeiro. 6, 11, 15, 22, 30, 33, 35, 42, 47, 142, 180, 220, 234, 239, 276
- Ayrosa, P. (1941). *Poemas brasílicos do Pe. Cristóvão Valente, SJ: notas e tradução*. Universidade de São Paulo. 34
- Baker, M. C. (1988). *Incorporation: A theory of grammatical function changing*. University of Chicago Press. 136
- Baker, M. C. (2001). The natures of nonconfigurationality. In C. Baltin, Mark & Collins (Ed.), *The Handbook of Contemporary Syntactic Theory*, pp. 407–438. Blackwell. 104
- Balee, W. (2001). Environment, culture, and Sirionó plant names. In L. Maffi (Ed.), *On Biocultral Diversity*, pp. 298–310. Smithsonian. 9
- Balée, W. and D. Moore (1994). Language, culture, and environment: Tupí-Guaraní plant names over time. In A. Roosevelt (Ed.), *Amazonian Indians from prehistory to the present: anthropological perspectives*, pp. 363–380. Tucson: University of Arizona Press. 19
- Barbosa, A. L. (1956). *Curso de tupi antigo: gramática, exercícios, textos*. Livraria São José. 36

- Barros, M. C. D. M. (2002). A origem intelectual das orações em tupi de andré thevet e yves devreux (séculos xvi-xvii): algumas hipóteses. *REVISTA DE ESTUDOS DA LINGUAGEM* 10(1), 139–187. 20
- Barros, M. C. D. M. (2003). A relação entre manuscritos e impressos em tupi como forma de estudo da política lingüística jesuítica no século xviii na amazônia. *Revista Letras* 61, 125–152. 22
- Barros, M. C. D. M. (2004). A política de estandardização da língua geral na amazônia no período colonial (séculos xvii-xviii). *Revista Internacional de Lingüística Iberoamericana* 2(1 (3), 23–54. 19, 20, 22, 35
- Beck, D. (2002). *The typology of parts of speech systems: The markedness of adjectives*. Routledge. 98, 101
- Benveniste, E. (1971). *Problems in general linguistics*. University of Miami Press. 118
- Bernardo, d. N. et al. (1709). *Katecismo Indico da Lingva Kariris, acrescentado de varias Praticas doutrinaes, & moraes, adaptadas ao genio, & capacidade dos Indios do Brasil, pelo Padre Fr. Bernardo de Nantes, Capuchinho, Pregador, & Missionario Apostolico; oferecido ao muy alto, e muy poderoso Rey de Portugal Dom Joaõ V. S.N. que Deus guarde*. Lisboa: Officina de Valentim da Costa Deslandes. 12
- Bertinetto, P. M. (2020). On nominal tense. *Linguistic Typology* 24(2), 311–352. 237, 240
- Bettendorff, J. F. (1800[1681]). *Compendio da doutrina christã na lingua portugueza e brasilica. Reimpresso por José Mariano, da ConceiçãoVELOZO*. Lisboa, Offic. de Simão Thaddeo Ferreira. 22, 30, 35, 167, 281
- Bettendorff, J. F. (2010[1698]). *Chronica da missão dos padres da Companhia de Jesus no Estado do Maranhão*. Brasília: Edições do Senado Federal. 22
- Bhat, D. N. S. (2005). *Third person pronouns and demonstratives*. Cambridge: Cambridge university Press. 118

- Bhat, D. S. (1999). *The prominence of tense, aspect and mood*. Amsterdam: John Benjamins Publishing. 216
- Bickel, B. (2010). Capturing particulars and universals in clause linkage. In I. Brill (Ed.), *Clause Linking and Clause Hierarchy: Syntax and Pragmatics*, pp. 51–101. Amsterdam: John Benjamins Publishing. 340
- Bickel, B. and J. Nichols (2005). Inflectional morphology. In T. Shopen (Ed.), *Language typology and syntactic description*, Volume 3, pp. 169–240. Cambridge: Cambridge University Press. 201, 258
- Bickel, B. and J. Nichols (2013). Obligatory Possessive Inflection. In M. S. Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. 302
- Blake, B. J. (2004). *Case* (2 ed.). Cambridge: Cambridge University Press. 263
- Bloomfield, L. (1935). *Language*. Georg Allen & Unwin Ltd. 69
- Boas, F. (1911a). Introduction. In F. Boas (Ed.), *Handbook of American Indian Languages*, Volume 1, pp. 5–83. Washington: Government Printing Office. 97
- Boas, F. (1911b). Kwakiutl. In F. Boas (Ed.), *Handbook of American Indian Languages. Part 1*. Washington: Government Printing Office. 100
- Bohnemeyer, J., L. K. Butler, and T. F. Jaeger (2016). Head-marking and agreement: Evidence from yucatec maya. In J. Fleischhauer, A. Latrouite, and R. Osswald (Eds.), *Explorations of the syntax-semantics interface*, pp. 169–208. Düsseldorf university press. 188, 189
- Bohnemeyer, J. and R. D. Van Valin Jr (2017). The macro-event property and the layered structure of the clause. *Studies in Language. International Journal sponsored by the Foundation Foundations of Language* 41(1), 142–197. 60, 353
- Borges, L. C. and J. H. Nunes (1998). La langue générale et la grammatisation des langues indiennes. *Langages* 2(130), 52–67. 22

- Borges, M. V. (2004). Diferenças entre as falas feminina e masculina no karajá e em outras línguas brasileiras: aspectos tipológicos. *LIAMES: Línguas Indígenas Americanas* 4(1), 103–113. 26
- Borges, M. V. et al. (2006). *Aspectos fonológicos e morfossintáticos da língua Avá-Canoeiro (Tupi-Guarani)*. Ph. D. thesis, Universidade de Campinas. Unpublished PhD thesis. 37, 317
- Braine, M. D. (1992). What sort of innate structure is needed to bootstrap into syntax? *Cognition* 45(1), 77–100. 98
- Braine, M. D. (1994). Mental logic and how to discover it. In G. R. J. Macnamara (Ed.), *The logical foundations of cognition*, pp. 241–263. Oxford University Press Oxford. 98
- Bresnan, J., S. Dingare, and C. D. Manning (2001). Soft constraints mirror hard constraints: Voice and person in english and lummi. In *Proceedings of the LFG01 Conference*, pp. 13–32. Stanford: CSLI Publications. 130
- Bresnan, J. and S. A. Mchombo (1987). Topic, pronoun, and agreement in chichewa. *Language* 63, 741–782. 189
- Brinton, D. G. (2009). *The American race: a linguistic classification and ethnographic description of the native tribes of North and South America*. Cambridge: Cambridge University Press. 18
- Broschart, J. (1997). Why tongan does it differently: Categorical distinctions in a language without nouns and verbs. *Linguistic Typology* 1(2), 123–166. 100
- Bueno, E. (2016). *A viagem do descobrimento: um olhar sobre a expedição de Cabral*. Estação Brasil. 10
- Butler, C. S. (2003a). *Structure and Function - A Guide to Three Major Structural-Functional Theories: Part 1: Approaches to the simplex clause*. Amsterdam: John Benjamins Publishing. 60
- Butler, C. S. (2003b). *Structure and Function - A Guide to Three Major Structural-Functional Theories: Part 2: From clause to discourse and beyond*. Amsterdam: John Benjamins Publishing. 60, 61

- Bybee, J. L. (1985). *Morphology: A study of the relation between meaning and form*, Volume 9. Amsterdam: John Benjamins Publishing. 197
- Bybee, J. L., R. D. Perkins, W. Pagliuca, et al. (1994). *The evolution of grammar: Tense, aspect, and modality in the languages of the world*. University of Chicago Press Chicago. 206, 219, 220, 314
- Cabral, A. (2000). Flexão relacional na família tupí-guaraní. *ABRALIN, Boletim da Associação Brasileira de Linguística* 25, 233–262. 111, 112
- Cabral, A. (2001a). O desenvolvimento da marca de objeto de segunda pessoa plural em tupí-guaraní. In A. Cabral and A. D. Rodrigues (Eds.), *Estudos sobre Línguas Indígenas*, Volume 1. Belém: Universidade Federal do Pará. 107
- Cabral, A. S. and L. da Costa (2004). Xikrín e línguas tupí-guaraní: marcas relacionais. *LIAMES: Línguas Indígenas Americanas* 4(1), 7–19. 111, 301
- Cabral, A. S. A. and A. d. Rodrigues (2005). O desenvolvimento do gerúndio e do subjuntivo em Tupí-Guaraní. In A. d. Rodrigues and A. S. A. Cabral (Eds.), *Novos estudos sobre línguas indígenas*, pp. 47–58. Editora da UnB Brasília. 209, 348, 353
- Cabral, A. S. A. C. (2001b). Observações sobre a história do morfema -a da família Tupí-Guaraní. In F. Queixalós (Ed.), *Des noms et des verbes en tupi-guarani: état de la question*, Volume 37, pp. 134–162. Lincom. 118
- Cabral, A. S. A. C. (2009). Caracterização do sistema de alinhamento do zoé e os fatores condicionadores de suas múltiplas cisões. In *Anais do VI Congresso Internacional da Abralín*, Volume 2, pp. 3145–3153. Abralín. 128
- Cabral, A. S. A. C. (2011). Different histories, different results: The origin and development of two amazonian languages. *PAPIA-Revista Brasileira de Estudos do Contato Linguístico* 21(1), 9–22. 19
- Cabral, A. S. A. C., A. P. do Couto e Silva, D. V. A. Martins, J. D. Lopes, J. P. F. de Carvalho, and S. A. Sousa (2011). Contribuições para o inventário

- da língua asuriní do tocantins: projeto piloto para a metodologia geral do inventário nacional da diversidade linguística. 155
- Cabral, A. S. d. C. (2007). L'expression des notions de l'épistémique et de l'alethique dans la famille Tupi-Guarani. In J. Guentchéva, Zlatka & Landaburu (Ed.), *L'enonciation médiatisée: Le Traitement Épistémologique de L'information: Illustrations Amérindiennes Et Caucasiennes*, pp. 267–292. Peeters. 207, 208
- Campbell, A. T. (1989). *To Square with Genesis: Causal Statements and Shamanic Ideas in Wayäap'i*. Edinburgh University Press. 31
- Cardim, F. (1881). *Do principio e origem dos indios do Brazil*, Volume 1. Rio de Janeiro:Typografia da Gazeta e Noticias. 11, 12, 14, 19
- Cardim, F. (2009). *Tratados da terra e gente do Brasil*. Hedra. 10, 11, 13, 21, 26
- Castro, R. C. d. and Q. F. Camargos (2021). The syntactic status of noun incorporation in the tenetehára language (tupí-guaraní family). *Revista Brasileira de Linguística Antropológica* 13(01), 299325. 136
- Castro e Silva, M. A., T. Ferraz, C. M. Couto-Silva, R. B. Lemes, K. Nunes, D. Comas, and T. Hünemeier (2022). Population histories and genomic diversity of south american natives. *Molecular biology and evolution* 39(1), msab339. 9
- Castro e Silva, M. A., K. Nunes, R. B. Lemes, À. Mas-Sandoval, C. E. Guerra Amorim, J. E. Krieger, J. G. Mill, F. M. Salzano, M. C. Bortolini, A. d. C. Pereira, D. Comas, and T. Hünemeier (2020). Genomic insight into the origins and dispersal of the brazilian coastal natives. *Proceedings of the National Academy of Sciences* 117(5), 2372–2377. 9
- Cerda, S. I. (2021). Complejidad intraclausal y cosubordinación en español El caso de la predicación secundaria. In C. C. Jiménez (Ed.), *Cosubordinación en español*, pp. 51–100. Berlin: Peter Lang. 344
- Cerno, L. and F. Obermeier (2013). Cartas de indígenas potiguaras de las guerras holandesas en el brasil. *Corpus* 3(1), 1–7. 36

- Cesar, J. T. G. (2016). Aryon rodrigues: restauração e tradução para o conhecimento da língua brasílica. Master's thesis, Universidade de Brasília. Unpublished master's thesis. 32
- Chappell, H. and W. McGregor (1989). Alienability, inalienability and nominal classification. *Annual Meeting of the Berkeley Linguistics Society* 15, 24–36. 300
- Chappell, H. and W. McGregor (1996). Prolegomena to a theory of inalienability. In H. Chappell and W. McGregor (Eds.), *The grammar of inalienability: A typological perspective on body part terms and part-whole relation*, pp. 3–30. Berlin: Walter de Gruyter. 299
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. MIT press. 345
- Comrie, B. (1976). *Aspect: An introduction to the study of verbal aspect and related problems*, Volume 2. Cambridge: Cambridge university press. 223
- Comrie, B. (1981). *Language universals and linguistic typology*. Blackwell. 370
- Comrie, B. (1989). Some general properties of reference-tracking systems. In D. Arnold and J. Durand (Eds.), *Essays on grammatical theory and Universal Grammar*, pp. 37–51. na. 348
- Comrie, B. et al. (1985). *Tense*. Cambridge: Cambridge university press. 197, 215, 216, 219
- Comrie, B. and E. L. Keenan (1979). Noun phrase accessibility revisited. *Language* 55(3), 649–664. 369
- Cooreman, A. (1994). A functional typology of antipassives. In B. A. Fox and P. J. Hopper (Eds.), *Voice: Form and function*, pp. 49–88. Amsterdam: John Benjamins Publishing. 168
- Copin, F. (2012). *Grammaire wayampi*. Ph. D. thesis, Univeristé Paris 7 Denis Diderot. Unpublished PhD thesis. 37, 113, 136
- Cormier, L. (2003). *Kinship with monkeys*. Columbia University Press. 9

- Correa da Silva, B. C. (1997). Urubú-ka'apor. da gramática à história, a trajetória de um povo. Master's thesis, universidade de Brasília. Unpublished master's thesis. 37
- Couchili, T., D. Maurel, and F. Queixalós (2002). Classes de lexèmes en émérillon. *Amerindia* 26(27), 173–208. 113
- Cristofaro, S. (2005). *Subordination*. Oxford University Press. 369
- Croft, W. (1990). A conceptual framework for grammatical categories (or: a taxonomy of propositional acts). *Journal of Semantics* 7(3), 245–279. 98
- Croft, W. (1991). *Syntactic categories and grammatical relations: The cognitive organization of information*. University of Chicago Press. 98, 99
- Croft, W. (2001). *Radical construction grammar: Syntactic theory in typological perspective*. Oxford University Press on Demand. 97, 98, 99
- Croft, W. (2003). *Typology and universals*. Cambridge: Cambridge University Press. 98, 323
- Croft, W. (2012). *Verbs: Aspect and causal structure*. Oxford University Press. 223, 230
- Croft, W. (2022a). *Morphosyntax: Constructions of the World's Languages*. Cambridge: Cambridge University Press. Draft version of 2021. 67, 97, 98, 99, 101, 120, 140, 333, 364
- Croft, W. (2022b). Word classes in Radical Construction Grammar. In E. van Lier (Ed.), *Oxford handbook of word classes*. Oxford University Press. 97, 98
- Crofts, M. (2004). *Aspectos da língua Mundurukú*. SIL. 150
- Cruz, A. (2011). *Fonologia e gramática do Nheengatu: a língua geral falada pelos povos Baré, Warekena e Baniwa. 2011*. Ph. D. thesis, Universiteit, Utrecht, The Netherlands. unpublished PhD thesis. 9, 21, 37, 142, 336, 364

- Cruz, A. d. (2016). Towards an understanding of the origin of aspectual marks on nouns: evidence from Nheengatu and Tupinambá. In . D. G. F. Queixalós (Ed.), *O Sintagma Nominal em Línguas Amazônicas*, pp. 45–74. Campinas: Pontes. 240
- Cysouw, M. (2003). *The paradigmatic structure of person marking*. Oxford: Oxford University Press. 107
- d'Abbeville, C. (1614). *Histoire de la mission des Peres Capucins en l'Isle de Maragnan et terres circonuoisines, où est traicté des singularitez admirables et des moeurs merueilleuses des Indiens habitans de ce pays, avec les missives et advis qui ont esté envoyez de nouveau par le RP Claude d'Abbeville*. François Huby, Paris. 13, 14, 15, 22, 30, 33, 34, 48, 323
- Daniel, J. (1975). *Tesouro descoberto no rio Amazonas*. 2 vols. Biblioteca Nacional. 22
- Davis, H. and L. Matthewson (2009). Issues in salish syntax and semantics. *Language and Linguistics Compass* 3(4), 1097–1166. 100
- Davis, P. W. and R. Saunders (1997). *A grammar of Bella Coola*, Volume 13. University of Montana. 100
- de Almeida, F. (1910a). *Cartas dos primeiros Jesuítas do Brasil (1538-1553; 1553-1558)*, Volume 2. Comissão do IV Centenário da Cidade de São Paulo. 20
- de Almeida, F. (1910b). *História da igreja em Portugal*. 2 volumes. Imprensa académica. 20, 22
- de Almeida Navarro, E. (2004). *Método moderno de tupi antigo: a língua do Brasil dos primeiros séculos*. Global. 36
- de Bry, T. (1592). *Americae tertia pars*. Officina Theodori de Bry. 38
- De Castro, E. V. (1992). *From the Enemy's Point of View*. University of Chicago Press. 9, 13, 16, 17
- De Castro, E. V. (2020). *From the Enemy's Point of View*. University of Chicago Press. 16

- de Castro, O. M., C. A. d. L. Carvalho, and E. G. da Luz (2020). Descrição morfológica da língua zo'ê. <https://www.recantodasletras.com.br/linguistica/6954065>. 155
- de Freitas Leite, Y. (2005). Arte de gramática da língua mais usada na costa do brasil: A criterion for evaluation. In *Missionary Linguistics II/Linguística misionera II*, pp. 191–204. Amsterdam: John Benjamins Publishing. 21
- De Léry, J. (1578[1972]). *Histoire d'un voyage fait en la terre de Brésil*. Geneva: Librairie Droz. 12, 13, 14, 30, 31, 48, 274
- de Léry, J. (1957). *Journal de bord de Jean de Lery en la terre de Brésil 1557*. Editions de Paris. 20
- de Oré, L. J. (1607). *Rituale, seu Manuale Peruanum, et forma brevis administrandi apud Indos sacrosancta Baptismi, Poenitentiae, Eucharistiae, Matrimonij, & Extremae vnctionis sacramenta*. apud Io. Iacobum Carlinum, & Costantinum Vitalem. 33
- de Paula Martins, M. d. L. (1945). O vocabulário brasileiro-portugues e o manuscrito 11.481 da biblioteca nacional. *Boletim bibliográfico da biblioteca pública municipa- de Saõ Paulo* 6, 59–93. 35
- de Paula Martins, M. d. L. (1949). Vocabulários tupis – o problema do vlb (1). *Boletim bibliográfico da biblioteca pública municipa- de Saõ Paulo* 13, 59–93. 35
- de Vasconcelos, S. (1865). *Chronica da Companhia de Jesu do estado do Brasil*. AJ Fernandes Lopes. 11, 23
- De Vries, M. (2002). *The syntax of relativization*. LOT, Utrecht. 367, 369
- d'Evreux, Y. (2014). *Voyage au nord du Brésil : (1615)*, Volume 4 of *Fontes Americanae*. Kiel: Westensee-Verl. 12, 14, 30, 33, 34, 274
- Diedrichsen, E. (2010). Towards a reconsideration of constructional schemas in rrg:: Are all constructions driven by constructions ? In W. Nakamura (Ed.), *International Conference on Role and Reference Grammar*, pp. 52–79. Unpublished conference proceedings. 63

- Diessel, H. (1999). *Demonstratives: Form, function and grammaticalization*. Amsterdam: John Benjamins Publishing. 282
- Dietrich, W. (1986). *El idioma chiriguano: gramática, textos, vocabulario*. Ediciones Cultura Hispánica, Instituto de Cooperación Iberoamericana. 36, 37, 155, 206
- Dietrich, W. (1990a). Chiriguano and Guarayo word formation. In D. L. Payne (Ed.), *Amazonian linguistics: studies in lowland South American languages*, pp. 293–321. University of Texas Press. 25
- Dietrich, W. (1990b). *More evidence for an internal classification of Tupi-Guarani languages*. Gebr. Mann Verlag Berlin. 18, 36
- Dietrich, W. (1994). Word formation, syntax, or noun classification? tupi-guarani mba'e-between lexicon and grammar. *Revista latinoamericana de estudios etnolingüísticos* 8, 113–128. 36, 168
- Dietrich, W. (2000). Problema de la categoría del adjetivo en las lenguas tupí-guaraníes. In H. van der Voort and S. van de Kerke (Eds.), *Indigenous languages of lowland South America*, pp. 255–263. Leiden: Research School CNWS, Leiden University. 36
- Dietrich, W. (2001). Categorias lexicais nas línguas tupi-guarani (visão comparativa). In F. Queixalós (Ed.), *Des noms et des verbes en tupi-guarani: état de la question*, pp. 23–37. Lincom. 36, 98, 100, 105, 132
- Dietrich, W. (2006). La categoría del frustrativo en las lenguas tupí-guaraníes. In W. Dietrich and H. Symeonidis (Eds.), *Guaraní y Mawetí-Tupí-Guaraní: Estudios históricos y descriptivos sobre una familia lingüística de América del Sur*, pp. 71–85. Berlin: Lit Verlag. 36, 211
- Dietrich, W. (2009). Cambio del orden de palabras en lenguas tupí-guaraníes. In *Ponencia presentada en el 53 Congreso Internacional de Americanistas. México*, Volume 1. Biblioteca Digital Curt Nimuendajú. 36
- Dietrich, W. (2014). Forms and Functions of Reduplication in Tupian Languages. In G. G. G. . H. van der Voort (Ed.), *Reduplication in Indigenous Languages of South America*, pp. 273–311. Brill. 36, 225

- Dietrich, W. (2017a). Sintaxis pronominal en las lenguas mawetí-guaraní. *Lingüística* 33(2), 69–95. 36, 132
- Dietrich, W. (2017b). Tipologia morfossintática da negação nas línguas do tronco tupi. *LIAMES: Línguas Indígenas Americanas* 17(1), 7–38. 36, 105, 212, 214, 222, 226
- Dietrich, W. (2017c). Word classes and word class switching in guaraní syntax. In B. Estigarribia and J. Pinta (Eds.), *Guarani Linguistics in the 21st Century*, pp. 158–. Brill. 36, 100
- Dietrich, W. (2023). Non-verbal predication in Tupian, especially Tupí-Guaraní languages. In P. M. Bertinetto, L. Ciucci, and D. Creissels (Eds.), *Non-verbal predication in the world's languages*. Berlin: De Gruyter. 36, 140
- Dik, S. C. (1991). Functional Grammar. In F. Droste and J. Joseph (Eds.), *Linguistic theory and grammatical description*, pp. 247–74. Amsterdam: John Benjamins Publishing. 60, 66
- Dingemanse, M. (2021, Jun). Interjections (oxford handbook of word classes). To appear in *Oxford Handbook of Word Classes*. 27
- Dixon, R. (2011). The Semantic Basis for a Typology. In A. Y. Aikhenvald and R. M. W. Dixon (Eds.), *Language at Large*, pp. 205–220. Brill. 137
- Dixon, R. M. (1972). *The Dyirbal language of north Queensland*. Cambridge: Cambridge University Press. 174
- Dixon, R. M. (1979). Ergativity. *Language* 55, 59–138. 126
- Dixon, R. M. (1982). *Where have all the adjectives gone?: and other essays in semantics and syntax*, Volume 107. Berlin: Walter de Gruyter. 97
- Dixon, R. M. (1994). *Ergativity*. Cambridge: Cambridge University Press. 126, 128
- Dixon, R. M. (2004). Adjective classes in typological perspective. In R. M. Dixon and A. Aikhenvald (Eds.), *Adjective classes: A cross-linguistic typology*, pp. 1–49. Oxford University Press. 97

- Dixon, R. M. (2010a). *Basic linguistic theory volume 2: Grammatical topics*, Volume 2. Oxford University Press. 100
- Dixon, R. M. (2010b). *Further grammatical topics*, Volume 3. Oxford University Press. 178
- Dixon, R. M. and A. Y. Aikhenvald (1999). Introduction. In R. M. Dixon and A. Aikhenvald (Eds.), *The Amazonian Languages*, pp. 1–21. Cambridge: Cambridge University Press. 148
- Dixon, R. M., R. M. Dixon, and D. Robert Malcolm Ward (1997). *The rise and fall of languages*. Cambridge: Cambridge University Press. 12
- Dobson, R. M. (2005). *Aspectos da língua kayabi*. SIL. 314
- Dooley, Rober, A. (2015). *Léxico Guarani, dialeto mbyá, introdução*. SIL International. 37, 215, 314, 348
- Dooley, R. A. (1998). *Léxico Guarani, dialeto mbya: versão para fins acadêmicos*. Sociedade Internacional de Linguística, Porto Velho. 155
- Downing, L. J. and S. Inkelas (2015). What is reduplication? typology and analysis part 2/2: The analysis of reduplication. *Language and Linguistics Compass* 9(12), 516–528. 225
- Dowty, D. R. (1979). *Word meaning and Montague grammar: The semantics of verbs and times in generative semantics and in Montague's PTQ*. D. Reidel. 80
- Drude, S. and S. Meira (2019, 11). Os assim-chamados 'prefixos relacionais' em línguas tupi: Uma análise alternativa. Unpublished presentation slides. I Simpósio de pesquisa em línguas indígenas Região norte (SIPLI-NORTE). 111
- Dryer, M. S. (1992a). A comparison of the obviation systems of kutenai and algonquian. In *Papers from the twenty-third annual Algonquian Conference*, Volume 23, pp. 119–163. Carleton University. 324
- Dryer, M. S. (1992b). The greenbergian word order correlations. *Language* 68(1), 81–138. 148, 282

- Dryer, M. S. (1997a). Are grammatical relations universal. In J. Bybee, J. Haiman, and S. A. Thompson (Eds.), *Essays on language function and language type*, pp. 115–143. Amsterdam: John Benjamins Publishing. 97, 138
- Dryer, M. S. (1997b). On the six-way word order typology. *Studies in Language. International Journal sponsored by the Foundation Foundations of Language* 21(1), 69–103. 123
- Dryer, M. S. (2007). Noun phrase structure. In T. Shopen (Ed.), *Language Typology and Syntactic Description*, Volume 2, pp. 151–205. Cambridge: Cambridge University Press. 361
- Dryer, M. S. (2013). Order of Numeral and Noun. In M. S. Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. 274
- Duffy, E. M. and A. C. Metcalf (2012). *The return of Hans Staden: a go-between in the Atlantic world*. JHU Press. 29
- Edelweiss, F. (1947). *Tupís e Guaranís: estudos de etnonímia e linguística*. Secretaria de Educação e Saúde. 11, 13, 20
- Edelweiss, F. (1969). *Estudos Tupís e Tupi-Guaranís: confrontos e revisões*. Livraria Brasileira. 54
- Eriksen, L. and A. V. Galucio (2014). The Tupian expansion. In L. O’connor and P. Muysken (Eds.), *The native languages of South America: Origins, development, typology*, pp. 177–199. Cambridge: Cambridge University Press. 18
- Estigarribia, B. (2020). *A Grammar of Paraguayan Guarani*. UCLPress. 37, 155, 165
- Evans, N. (2000). Kinship verbs. In P. M. V. . B. Comrie (Ed.), *Approaches to the typology of word classes*, pp. 103–172. Berlin: Mouton de Gruyter. 98
- Evans, N. and T. Osada (2005). Mundari: The myth of a language without word classes. *Linguistic Typology* 9(3), 351–390. 100

- Everett, C. (2013). *Linguistic relativity*. Berlin: De Gruyter Mouton. 274
- Everett, D. (2008). Wariintentional state constructions. In R. D. V. V. Jr (Ed.), *Investigations of the syntax-semantics-pragmatics interface*, pp. 381–409. Amsterdam: John Benjamins Publishing. 67
- Everett, D. (2017). *How language began: The story of humanity's greatest invention*. Profile Books. 8
- Everett, D. L. (2012). *Language: The cultural tool*. Vintage. 8, 300
- Fausto, C. (1998). Fragmentos de história e cultura tupinambá: da etnologia como instrumento crítico de conhecimento etno-histórico. In M. C. da Cunha (Ed.), *História dos índios no Brasil*. 2 ed., pp. 381–396. Editora Schwarcz. 13
- Fernandes, F. (1949). *A economia tupinambá: ensaio de interpretação sociológica do sistema econômico de uma sociedade tribal*. Prefeitura do Município de São Paulo. 13, 14, 16, 17, 29
- Fernandes, F. (1963). *Organização social dos Tupinambá*. Difusão Européia do Livro. 13
- Fernandes, F. (1970). *A função social da guerra na sociedade tupinambá* (3 ed.). São Paulo: Biblioteca Azul. 13, 16, 29
- Ferraz Gerardi, F. (2022). Tupinambá treebank (ud_tupinamba_tudet). online. Available online at <http://universaldependencies.org>. 289
- Ferraz Gerardi, F. and S. Reichert (2021). The tupí-guaraní language family: A phylogenetic classification. *Diachronica* 38(2), 151–188. 9, 18
- Ferraz Gerardi, F., T. Tresoldi, C. Aragon, S. Reichert, G. de Souza, Jonas, and F. S. Noelli (2023). Lexical phylogenetics of the tupí-guaraní family: Language, archaeology, and the problem of chronology. *PlosOne* xx(xx), xx–xx. xi, 9, 12, 18, 39
- Ferreira da Silva, G. (2003). Construindo um dicionário parakanã-português. Master's thesis, Universidade de Belém. Unpublished master's thesis. 155

- Figueira, L. (1878[1687]). *Grammatica da lingua do Brasil*. Leipzig: B. G. Teubner. 6, 19, 21, 22, 30, 35, 41, 141, 150, 155, 179, 214, 281, 319, 321, 323
- Fleming, L. (2012). Gender indexicality in the native americas: Contributions to the typology of social indexicality. *Language in Society* 41(3), 295–320. 26
- Foley, w. (2010). Clause linkage and nexus in Papuan languages. In I. Brill (Ed.), *Clause Linking and Clause Hierarchy: Syntax and Pragmatics*, pp. 27–50. Amsterdam: John Benjamins Publishing. 341
- Foley, W. and R. D. Van Valin Jr (1984). *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press. 60, 138, 340, 344, 348
- Foley, W. A. (2017). Structural and semantic dependencies in word class. In N. J. Enfield (Ed.), *Dependencies in language*, pp. 179–194. Language Science Press. 97
- Foley, W. A. and R. D. Van Valin Jr (1985). Information packaging in the clause. *Language typology and syntactic description 1*, 282–364. 71, 168
- Fortune, D. and G. Fortune (1975). Karajá mens-womens speech differences with social correlates. *Arquivos de Anatomia e Antropologia 1*, 111–124. 26
- Frachtenberg, L. J., R. M. (1922). Coos. In F. Boas (Ed.), *Handbook of American Indian Languages vol. II*. Washington Government Printing Office. 100
- Freire, J. R. B. (2004). *Rio Babel: a história das línguas na Amazônia*. Eduerj. 20, 21, 22
- Freire, J. R. B. (2011). *Rio Babel: a história das línguas na Amazônia* (2 ed.). Eduerj. 19
- Futrell, R., K. Mahowald, and E. Gibson (2015). Large-scale evidence of dependency length minimization in 37 languages. *Proceedings of the National Academy of Sciences 112*(33), 10336–10341. 121

- Gaffarel, P. (1877). *Jean de Léry: La langue Tupi*. Maisonneuve et Cie, Libraries-Éditeurs. 31, 32
- Galucio, A. V., S. Meira, J. Birchall, D. Moore, N. Gabas, S. Drude, L. Storto, G. Picanço, and C. R. Rodrigues (2015). Genealogical relations and lexical distances within the tupian linguistic family. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas* 10, 229–274. 18, 150
- Gerardi, F. F. (2020). UD_Tupinamba-TuDeT. https://github.com/UniversalDependencies/UD_Tupinamba-TuDeT. UD Treebank. 145
- Gerardi, F. F., S. Reichert, C. Aragon, J.-M. List, and T. Wientzek (2021, Mar). Tuled: Tupían lexical database. 23
- Giannakidou, A. and A. Mari (2018). A unified analysis of the future as epistemic modality. *Natural Language & Linguistic Theory* 36(1), 85–129. 221
- Gildea, S. (2002). Pre-*proto-tupí-guaraní* main clause person-marking. *Línguas Indígenas Brasileiras. Fonologia, Gramática e História. Atas do I Encontro Internacional do GTLI da ANPOLL 1*, 315–326. 106, 112, 113, 119
- Givón, T. (1994). The pragmatics of de-transitive voice: Functional and typological aspects of inversion. In T. Givón (Ed.), *Voice and inversion*, pp. 3–44. Amsterdam: John Benjamins Publishing. 168, 323
- Givón, T. (2001). *Syntax, vol. I*. Amsterdam: John Benjamins Publishing. 60, 160, 216
- Goldberg, A. E. (2003). Constructions: A new theoretical approach to language. *Trends in cognitive sciences* 7(5), 219–224. 61
- Goldberg, A. E. (2013). Backgrounded constituents cannot be extracted. *Experimental syntax and island effects* 221, 221–238. 314
- Goldsmith, J. A. and B. Laks (2019). *Battle in the mind fields*. University of Chicago Press. 60

- Gomes, D. M. (2006). *Estudo morfológico e sintático da língua Mundurukú (Tupí)*. Ph. D. thesis, Universidade de Brasília. unpublished phd thesis. 112, 119, 150, 245
- Gomes, M. P. (2002). *O índio na história: o povo Tenetehara em busca da liberdade*. Petrópolis: Vozes. 22
- González Vergara, C. E. (2006). La gramática del papel y la referencia: una aproximación al modelo. *Onomázein* 14(2), 101–140. 37, 63
- Greenberg, J. H. (1963). Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In J. H. Greenberg (Ed.), *Universals of Human Language*, pp. 73–113. Cambridge: MIT Press. 148
- Grenand, F. (1980). *La langue Wayãpi (Guyane française): Phonologie et grammaire*, Volume 41. Peeters Publishers. 37
- Guillaume, A. and F. Rose (2010). Sociative causative markers in South-American languages: a possible areal feature. In F. Franck (Ed.), *Essais de Typologie et de linguistique générale. Mélanges offerts à Denis Creissels.*, pp. 383–402. ENS Éditions. 159
- Gândavo, P. d. M. d. (1576). *Historia da província de sãcta Cruz a que vulgarmente chamamos Brasil*. Officina de Antonio Gonsalvez. 11
- Gândavo, P. d. M. d. (2004). *A primeira história do Brasil: história da província Santa Cruz a que vulgarmente chamamos Brasil*. Zahar. 17, 20
- Hagège, C. (2010). *Adpositions*. Oxford: Oxford University Press. 246, 247
- Hahn, M. (2014). Predication and NP Structure in an Omnipredicative Language: The Case of Khoekhoe. In S. Müller (Ed.), *Proceedings of the 21st International Conference on Head-Driven Phrase Structure Grammar*, pp. 238–258. CSLI Publications. 104
- Haiman, J. and P. Munro (1983). *Switch Reference and Universal Grammar: Proceedings of a symposium on switch reference and universal grammar, Winnipeg, May 1981*, Volume 2. Amsterdam: John Benjamins Publishing. 349

- Hale, K. (1983). Warlpiri and the grammar of non-configurational languages. *Natural Language & Linguistic Theory* 1(1), 5–47. 104
- Halliday, M. A. (1985). *Introduction to functional grammar*. University Park Pres. 61
- Halliday, M. A. K. and C. M. Matthiessen (2013). *Halliday's introduction to functional grammar*. Routledge. 61
- Hallonsten Halling, P. (2018). *Adverbs: A typological study of a disputed category*. Ph. D. thesis, Stockholm University. 120
- Hammarström, H. (2010). Rarities in numeral systems. In *Rethinking universals: How rarities affect linguistic theory*, Volume 45, pp. 11–53. Berlin: Mouton de Gruyter. 277
- Hammond, J. (2015). *Switch reference in Whitesands. Theoretical issues and experimental evidence*. Ph. D. thesis, Radboud Universiteit Nijmegen. 107, 348
- Harrison, C. (1986). Verb prominence, verb initialness, ergativity and typological disharmony in Guajajara. In D. Derbyshire and G. Pullum (Eds.), *Handbook of Amazonian languages*, Volume 1, pp. 407–439. Berlin: Walter de Gruyter. 314
- Harrison, C. and C. Harrison (2013). *Dicionário guajajara-português*. Anápolis: International Linguistic Association (SIL). 155, 237
- Haspelmath, M. (1995). The converb as a cross-linguistically valid category. In *Converbs in cross-linguistic perspective: structure and meaning of adverbial verb forms-adverbial participles, gerunds*, Volume 13, pp. 1–55. Berlin: de Gruyter. 352, 353
- Haspelmath, M. (2002). *Understanding morphology: Arnold London*. Routledge. 104
- Haspelmath, M. (2004). Coordinating constructions: An overview. In M. Haspelmath (Ed.), *Coordinating constructions*, Volume 58, pp. 3–40. Amsterdam: John Benjamins Publishing. 173, 361

- Haspelmath, M. (2007). Pre-established categories don't exist: Consequences for language description and typology. *Linguistic typology* 11(1), 119–132. 97, 98
- Haspelmath, M. (2008). Framework-free grammatical theory. In B. Heine and H. Narrog (Eds.), *The Oxford Handbook of Linguistic Analysis*. Oxford: Oxford University Press. 59
- Haspelmath, M. (2010). Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3), 663–687. 98, 209
- Haspelmath, M. (2011). The indeterminacy of word segmentation and the nature of morphology and syntax. *Folia linguistica* 45(1), 31–80. 104
- Haspelmath, M. (2013). Argument indexing: A conceptual framework for the syntactic status of bound person forms. *Languages across boundaries: Studies in memory of Anna Siewierska* 197, 226. 105, 188, 189
- Haspelmath, M. (2015). Ditransitive constructions. *Annu. Rev. Linguist.* 1(1), 19–41. 135
- Haspelmath, M. (2020). The structural uniqueness of languages and the value of comparison for language description. *Asian Languages and Linguistics* 1(2), 346–366. 97
- Haspelmath, M. (2021a). Comparing reflexive constructions in the world's languages. In K. Janic, N. Puddu, and M. Haspelmath (Eds.), *Reflexive constructions in the world's languages*, pp. 331–359. Berlin: LangSci Press. 171, 178
- Haspelmath, M. (2021b). Word class universals and language-particular analysis. In Fulano (Ed.), *Oxford handbook of word classes*. Oxford University Press. xii, 97, 98, 99, 100, 101, 104
- Haspelmath, M. (2022, 2). Nonverbal clause constructions. Online. Unpublished draft. 101, 140
- Hawkins, J. A. (1983). *Word Order Universals*. New York: Academic Press. 148

- Hawkins, J. A. (2014). *Cross-linguistic variation and efficiency*. Oxford University Press. 132
- Hébert, Y. M. (1983). Noun and verb in a salishan language. *Kansas working papers in linguistics* 8(2), 31–82. 100
- Hemming, J. (1978). *Red gold: the conquest of the Brazilian indians*. Harvard University Press. 10, 12, 15, 17, 20, 31
- Hervás y Panduro, L. (1800–1805). *Catálogo De Las Lenguas De Las Naciones Conocidas, Y Numeracion, Division, Y Clases De Estas Segun La Diversidad De Sus Idiomas Y Dialectos: Lenguas y naciones europeas; parte II, Naciones Europeas primitivas: sus lenguas matrices, y dialectos de estas*. Madrid: Administración del Real Arbitrio Beneficencia. 18
- Himmelman, N. P. (2008). Lexical categories and voice in Tagalog. In P. A. S. Musgrave (Ed.), *Voice and grammatical relations in Austronesian languages*, pp. 247–293. CSLI Stanford. 104
- Hoeller, A. (1932). *Guarayo-Deutsches Wörterbuch*. Verlag der Missionssprokura der pp Franziskaner, Hall in Tirol. 208
- Holisky, D. A. (1987). The case of the intransitive subject in tsova-tush (batsbi). *Lingua* 71(1-4), 103–132. 127
- Holmberg, A. H. (1950). *Nomads of the long bow: The Siriono of Eastern Bolivia*. Forgotten Books. 31
- Hopper, P. J. and S. A. Thompson (1980). Transitivity in grammar and discourse. *language* 56(2), 251–299. 166, 168
- Hopper, P. J. and S. A. Thompson (1984). The discourse basis for lexical categories in universal grammar. *Language* 4(60), 703–752. 160
- Hopper, P. J. and E. C. Traugott (2003). *Grammaticalization* (2 ed.). Cambridge: Cambridge University Press. 247
- Hornstein, N. (1993). *As time goes by: Tense and universal grammar*. MIT Press. 197, 215

- Höller, A. (1932). *Grammatik der Guarayo-Sprache*. Verlag d. Missionsprokura d. PP Franziskaner. 155
- Inglese, G. (2021). Towards a typology of middle voice systems. *Linguistic Typology* 26(3), 489–531. 173, 174, 176
- Inglese, G. (2022). How do middle voice markers and valency reducing constructions interact? typological tendencies and diachronic considerations. *Folia Linguistica* 56(2), 239–271. 174
- Inkelas, S. and L. J. Downing (2015). What is reduplication? typology and analysis part 1/2: The typology of reduplication. *Language and linguistics compass* 9(12), 502–515. 225
- Jackendoff, R. (1983). *Semantics and cognition*. MIT press. 98
- Jackendoff, R. (1992). *Semantic structures*, Volume 18. MIT press. 78, 273
- Jackendoff, R. and J. Audring (2020). *The texture of the lexicon: relational morphology and the parallel architecture*. Oxford University Press, USA. 60
- Jackendoff, R. S. (2002). *Foundations of language: Brain, meaning, grammar, evolution*. New York: Oxford University Press. 59, 60, 93
- Jacobsen Jr, W. H. (1992). Subordination and cosubordination in nootka: Clause combining in a polysynthetic verb-initial language. In R. D. Van Valin Jr (Ed.), *Advances in role and reference grammar*, pp. 235–274. Amsterdam: John Benjamins Publishing. 348
- Janic, K. and A. Witzlack-Makarevich (2021). *Antipassive: typology, diachrony, and related constructions*. John Benjamin. 168
- Jelinek, E. (1984). Empty categories and non-configurational languages. *Natural language and linguistic theory* 2, 39–76. 104
- Jelinek, E. and R. A. Demers (1983). The agent hierarchy and voice in some coast salish languages. *International Journal of American Linguistics* 49(2), 167–185. 130

- Jelinek, E. and R. A. Demers (1994). Predicates and pronominal arguments in straits salish. *Language*, 697–736. 130
- Jensen, C. (1990a). Cross-referencing changes in some Tupí-Guaraní languages. In *Amazonian linguistics: Studies in lowland South American languages*, pp. 117–158. Austin: University of Texas Press. 119, 128, 131
- Jensen, C. (1998a). Comparative tupí-guaraní morphosyntax. In B. S. Efrat (Ed.), *Handbook of Amazonian languages*, Volume 4, pp. 489–618. Berlin: Walter de Gruyter. 24, 37, 105, 111, 112, 128, 211, 266, 353
- Jensen, C. (1998b). The use of coreferential and reflexive markers in tupi-guarani languages. *Journal of Amazonian languages 1, 2*. 107
- Jensen, C. (1999). Tupí-Guaraní. In R. W. Dixon and A. Y. Aikhenvald (Eds.), *The Amazonian Languages*, pp. 125–163. Cambridge: Cambridge University Press. 37, 41, 113, 254, 314
- Jensen, C. J. (1984). O desenvolvimento histórico da língua wayampí. Master's thesis, Universidade de Campinas. unpublished master Thesis. 41, 44
- Jensen, C. J. S. (1990b). O desenvolvimento histórico da língua wayampi. Master's thesis, universidade de Campinas. Unpublished master's thesis. 37
- Jiménez-Briones, R. and A. Luzondo-Oyón (2013). Constructions in role and reference grammar. In B. Nolan and E. Diedrichsen (Eds.), *Linking constructions into functional linguistics: The role of constructions in grammar*, pp. 179–204. Amsterdam: John Benjamins Publishing. 63
- Jiménez, C. C. (2021). Cosubordinación clausal en español. In C. C. Jiménez (Ed.), *Cosubordinación en español*, pp. 179–204. Amsterdam: John Benjamins Publishing. 331
- Johnson, M. (1987). A new approach to clause structure in role and reference grammar. *Davis Working Papers in Linguistics 2*, 55–59. 72

- Julião, R. M. S. (2005). *Aspects morphosyntaxiques de l'anambé*. Ph. D. thesis, Toulouse 2. Unpublished PhD thesis. 317
- Kailuweit, R. (2018). Activity Hierarchy and Argument Realization in (R)RRG. In E. S. Rolf Kailuweit, Lisann Künkel (Ed.), *Applying and Expanding Role and Reference Grammar*, pp. 189–211. Freiburg: FRIAS. 64
- Kakumasu, J. et al. (1986). Urubu-Kaapor. In D. C. Derbyshire and G. K. Pullum (Eds.), *Handbook of Amazonian languages*, Volume 1, pp. 326–403. New York: Mouton de Gruyete. 37
- Kallmeyer, L., R. Osswald, and R. D. Van Valin Jr (2013). Tree wrapping for role and reference grammar. In G. Morrill and M.-J. Nederhof (Eds.), *Formal Grammar*, pp. 175–190. Berlin: Springer. 62
- Keenan, E. L. (1976). Towards a Universal Definition of "Subject" in Subject and Topic. In C. N. Li (Ed.), *Subject and topic*. Academic Press. 140
- Keenan, E. L. and B. Comrie (1977). Noun phrase accessibility and universal grammar. *Linguistic inquiry* 8(1), 63–99. 369, 370
- Keenan, E. L. and B. Comrie (1979). Data on the noun phrase accessibility hierarchy. *Language* 55(2), 333–351. 369
- Kemmer, S. (1994). Middle Voice, Transitivity. In B. A. Fox and P. J. Hopper (Eds.), *Voice: Form and function*, pp. 179–230. Amsterdam: John Benjamins Publishing. 173, 174
- Kemmerer, D. (2019). *Concepts in the brain: The view from cross-linguistic diversity*. Oxford University Press. 98
- Kihara, C. P. (2017). *Aspects of Gikūyū (Kikuyu) Complex Sentences: A Role and Reference Grammar Analysis*. Ph. D. thesis, Heinrich-Heine Universität–Düsseldorf. Unpublished PhD thesis. 70, 348
- Kinkade, M. D. (1983). Salish evidence against the universality of noun and verb. *Lingua* 60(1), 25–39. 104

- Kiparsky, P. (2015). Disjoint reference and the typology of pronouns. In I. K. . B. Stiebels (Ed.), *More than Words: A Festschrift for Dieter Wunderlich*, pp. 179–225. Akademie Verlag. 324
- Kroeger, P. (2018). *Analyzing meaning: An introduction to semantics and pragmatics*. Language Science Press. 84
- Kuteva, T., B. Heine, B. Hong, H. Long, H. Narrog, and S. Rhee (2019). *World lexicon of grammaticalization*. Cambridge: Cambridge University Press. 247, 369
- Ladefoged, P. and I. Maddieson (1996). *the sounds of the world's languages*. Blackwell. 43
- Lambrecht, K. (1986). *Topic, focus, and the grammar of spoken French*. Ph. D. thesis, University of California. Unpublished PhD thesis. 73, 77, 306
- Lambrecht, K. (1987). Sentence focus, information structure, and the thematic-categorical distinction. In *Proceedings of the 13th annual meeting of the Berkeley Linguistics Society*, Volume 13, pp. 366–382. Berkeley Linguistics Society. 77, 306
- Lambrecht, K. (1994). *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge university press. 74, 77, 305, 306, 313, 319
- Lambrecht, K. (2000). When subjects behave like objects: An analysis of the merging of s and o in sentence-focus constructions across languages. *Studies in Language. International Journal sponsored by the Foundation Foundations of Language* 24(3), 611–682. 77, 306, 307, 308
- Langacker, R. (1990). Settings, participants, and grammatical relations. In S. L. Tsohatzidis (Ed.), *Meaning and prototypes: studies in linguistic categorization*, pp. 213–238. Routledge. 61
- Langacker, R. W. (1987). *Foundations of cognitive grammar: Theoretical prerequisites*, Volume 1. Stanford university press. 61

- LaPolla, R. J. (Forthcoming). Grammatical Relations. In R. M. Robert D. Van Valin Jr, Delia Bentley and W. Nakamura (Eds.), *The Cambridge Handbook of Role and Reference Grammar*, pp. –. Cambridge University Press. 138, 139
- Latrouite, A. and R. Van Valin Jr (2021). A role and reference grammar account of aspects of the information structure-syntax interface in tagalog. In R. D. Van Valin, Jr (Ed.), *Challenges at the Syntax-Semantics- Pragmatics Interface*, pp. 258–283. 315
- Launey, M. (1994). *Une grammaire omniprédicative: Essai sur la morpho-syntaxe du nahuatl classique*. CNRS éditions. 104
- Launey, M. (2002). Compound nouns vs. incorporation in classical nahuatl. *Studia typologica* 4, 113–134. 104
- Launey, M. (2004). The features of omnipredicativity in classical nahuatl. *STUF-Language Typology and Universals* 57(1), 49–69. 104
- Lazard, G. (1999). La question de la distinction entre nom et verbe en perspective typologique. *Folia linguistica* 33(3-4), 389–418. 100
- Lee, M. K. (2005). *Conversing in colony: The Brasília and the Vulgar in Portuguese America, 1500–1759*. Ph. D. thesis, The Johns Hopkins University. Unpublished PhD thesis. 12, 19, 20, 21, 22, 33, 35, 36
- Lehmann, C. (1984). *Der relativsatz*. Gunter Narr Verlag. 369
- Lehmann, C. (1985). On grammatical relationality. *Folia Linguistica* 19(1-2), 67–110. 299
- Lehmann, C. (2015). *Thoughts on grammaticalization* (3 ed.). Language Science Press. 247
- Leite, F. R. (2013). A língua geral paulista e o “vocabulário elementar da língua geral brasílica”. Master’s thesis, Universidade de Campinas. Unpublished master’s thesis. 21
- Leite, S. (1938-1950). *História da Companhia de Jesus no Brasil. X vols., 1938-1950*. Lisbon: Livraria Portugália. 20

- Leite, S. (1940). *Novas cartas jesuíticas: de Nóbrega a Vieira*. Companhia Editora Nacional. 20
- Lemaréchal, A. (1989). *Les parties du discours: sémantique et syntaxe*. Paris: Presses Universitaires de France-PUF. 104
- Lemaréchal, A. (1991). *Problèmes de sémantique et de syntaxe en Palau*. Paris: Editions du Centre national de la recherche scientifique. 104
- Léry, J. d. (1578). *Histoire d'un voyage fait en la Terre du Bresil autrement dite Amerique*. Pour Antoine Chuppin. 16, 19, 30
- Lima, F. P. and I. d. C. Moreira (2005). Tradições astronômicas tupinambás na visão de Claude L'Abbeville. *Revista da Sociedade Brasileira de História da Ciência* 3(1), 4–19. 15
- Liu, H. (2008). Dependency distance as a metric of language comprehension difficulty. *Journal of Cognitive Science* 9(2), 159–191. 121
- MacCormack, S. (1999). Ethnography in South America: The first two hundred years. In *The Cambridge history of the native peoples of the Americas. Volume III, part 1*, pp. 96–187. Cambridge: Cambridge University Press. 13, 33
- Maddieson, I. (2013a). Consonant inventories. In M. S. Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. 23, 42
- Maddieson, I. (2013b). Vowel quality inventories. In M. S. Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. 23, 42
- Magalhães, M., W. Praça, and A. da Cruz (2019). Gradação da omnipredicatividade na família tupi-guarani. *Forma y Función* 32(2), 151–189. 104, 142
- Magalhães, M. M. S. (2010). *Sobre a morfologia e a sintaxe da língua Guajá (Família Tupí-Guaraní)*. Ph. D. thesis, Universidade de Brasília. Unpublished PhD thesis. 37, 131, 136, 159

- Magalhaes, M. M. S. and A. C. R. de Mattos (2014). Classes de palavras, tipos de predicados e sua relação com a intransitividade cindida em guajá. *Via Litterae (ISSN 2176-6800): Revista de Linguística e Teoria Literária* 6(2), 251–284. 113
- Magalhães, M. M. S. and D. O. S. Alves (2022). Alteração morfosintática de predicados no guajá pela anteposição de expressões adverbiais. *Diachronica* 17(1), 1–26. 315
- Mamiani della Rovere, L. V. (1698). *Catecismo da Doutrina Christãa. Na lingua brasilica da nação Kiriri composto pelo P. Luis Vincencio Mamiani, da Companhia de JESUS, Missionario da Provincia do Brasil*. Miguel Deslandes. 12
- Mamiani della Rovere, L. V. (1877[1699]). *Arte de grammatica da lingua brazilica da nação kiriri*. Typ. Central de Brown & Evaristo. 12
- Marcgrave, G. and W. Piso (1648). *Historia Naturalis Brasiliae ... in qua non tantum plantae et animalia, sed et indigenarum morbi, ingenia et mores describuntur et iconibus supra quingentas illustrantur*. Amstelodami: Lud. Elzevirium. 244, 294
- Martin, Jack, B. (2011). *A Grammar of Creek (Moskogee)*. University of Nebraska Press. 324
- Martins, M. d. L. d. P. (1941). *Contribuição para o estudo do teatro tupi de Anchieta*. Universidade de São Paulo. Boletim XXIV. 30
- Martins, M. d. L. d. P. (1945). *Poesias tupis (século XVI)*, Volume 51. Universidade de S. Paulo, Faculdade de Filosofia, Ciência e Letras. 30
- Matić, D., R. Van Gijn, and R. D. Van Valin Jr (2014). Information structure and reference tracking in complex sentences: an overview. In D. Matić, R. Van Gijn, and R. D. Van Valin Jr (Eds.), *Information structure and reference tracking in complex sentences*, pp. 1–43. Amsterdam: John Benjamins Publishing. 334
- McGregor, W. (2009). *The expression of possession*. Berlin: Walter de Gruyter. 299

- Meira, S. (2006). Stative verbs vs. nouns in sateré-mawé and the tupian family. *LOT Occasional Series 5*, 189–214. 100, 101
- Meira, S. and S. Drude (2013). Sobre a origem histórica dos prefixos relacionais das línguas tupí-guaraní. *Cadernos de Etnolingüística 5*(1), 1–30. 107, 111, 112, 118, 119
- Meira, S. and S. Drude (2015). A summary reconstruction of proto-maweti-guarani segmental phonology. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas 10*(2), 275–296. 41
- Mello, A. A. S. et al. (2000). *Estudo histórico da família linguística Tupí-Guaraní: aspectos fonológicos e lexicais*. Ph. D. thesis, Universidade Federal de Santa Catarina. Unpublished PhD thesis. 41
- Métraux, A. (1927). Migrations historiques des tupi-guarani. *Journal de la Société des Américanistes 19*, 1–45. 12
- Métraux, A. (1928a). *La civilisations matérielle des tribus Tuzpi-Guarani*. Librairie Orientaliste Paul Geuthner. 11, 12, 13, 29, 30
- Métraux, A. (1948). The Tupinamba. In J. H. Steward (Ed.), *Handbook of South American Indians: The tropical forest tribes*, Volume 3, pp. 95–134. US Government Printing Office. 10, 12, 29
- Métraux, A. (1979). *A religião dos Tupinambás* (second ed.). São Paulo: Companhia Editora Nacional. 29, 30, 31, 34
- Métraux, A. (2014[1928]b). *La religion des Tupinamba et ses rapports avec celle des autres tribus Tupi-Guarani*. Paris: PUF. 17, 23
- Michael, L. and Z. O'Hagan (2016). A linguistic analysis of old omagua ecclesiastical texts. *Cadernos de Etnolingüística, Série Monografias 4*, 1–176. 37
- Michelson, K. (1991). Possessor stranding in oneida. *Linguistic inquiry 22*(4), 756–761. 135
- Mithun, M. (1984). The evolution of noun incorporation. *Language 60*(4), 847–894. 135, 160, 162

- Mithun, M. (1988). The grammaticization of coordination. In J. Haiman and S. A. Thompson (Eds.), *Clause combining in grammar and discourse*, pp. 331–359. Amsterdam: John Benjamins Publishing. 361
- Mithun, M. (2001). *The languages of native North America*. Cambridge: Cambridge University Press. 100
- Monserrat, R., C. Barros, and B. F. Barbosa (2020). Um escrito tupi do capitão simão soares parayba (1645). *Corpus. Archivos virtuales de la alteridad americana* 10(2), 1–16. 36
- Monserrat, R. and M. F. Soares (1983). Hierarquia referencial em línguas tupi. *Ensaio de Linguística: Cadernos de Linguística e teoria da Liratura* 5(9), 164–187. 131
- Monteiro, J. M. (2001). Tupis, tapuias e historiadores. Unpublshed manuscript. Tese de livre docência, UNICAMP. 20
- Montoya, A. R. d. (1876). *Arte de la lengua guarani o mas bien tupi*. Faesy & Frick. 37, 146, 149, 155, 239, 351
- Muller, J. C., W. Dietrich, and R. Monserrat (2019). *Dicionário de Língua Geral Amazônica*. Potsdam: Universitätsverlag Potsdam Museu Paraense Emilio Goeldi. 20, 21, 364
- Müller, N. J. (2013). *Tense, aspect, modality, and evidentiality marking in South American indigenous languages*. Ph. D. thesis, Radboud Universiteit Nijmegen. 216
- Navarro, E. d. A. (2011). Recensão crítica: Fonologia e gramática do nheengatú. Unpublished review. Available online. 155
- Navarro, E. d. A. (2013). *Dicionário tupi antigo: a língua indígena clássica do Brasil*. Global Ed. 155, 249, 334
- Navarro, E. d. A. (2022). Transcrição e tradução integral anotada das cartas dos índios camarões, escritas em 1645 em tupi antigo. *Boletim do Museu. Paraense Emilio Goeldi. Ciências Humanas* 17(3), 1–49. 6, 30, 36, 180

- Neiva Praça, W. (2007). *Morfossintaxe da Língua Tapirapé (Família Tupí-Guaraní)*. Ph. D. thesis, Universidade de Brasília, Brasília. 113, 236
- Newmeyer, F. J. (1991). Functional explanation in linguistics and the origins of language. *Language & Communication* 11(1-2), 3–28. 60
- Nichols, J. (1983). Switch-reference in the Northeast Caucasus. In *Switch-reference and universal grammar*, pp. 245–265. Amsterdam: John Benjamins Publishing. 349
- Nichols, J. (1986). Head-marking and dependent-marking grammar. *Language* 62(1), 56–119. 69, 188
- Nichols, J. (1988). On alienable and inalienable possession. In W. Shipley (Ed.), *In honor of Mary Haas: from the Haas Festival Conference on Native American Linguistics*, pp. 557–609. Berlin: Mouton de Gruyter. 300
- Nichols, J. (2009). Linguistic complexity: a comprehensive definition and survey. In G. Sampson, D. Gil, and P. Trudgill (Eds.), *Language complexity as an evolving variable*, pp. 110–125. Oxford University Press. 149
- Nichols, J., J. Barnes, and D. A. Peterson (2006). The robust bell curve of morphological complexity. *Linguistic typology* 10(1), 96–106. 149
- Nichols, J., Peterson, D. A, and J. Barnes (2004). Transitivity and detransitivizing languages. *Linguistic Typology* 8(2), 149–211. 132
- Noelli, F. S. (1996). As hipóteses sobre o centro de origem e rotas de expansão dos tupi. *Revista de antropologia* 39(2), 7–53. 19
- Noelli, F. S. (2022). Che rog pypia aje katu! A linguagem da casa-longa Guaraní no século XVII. *Tellus* 48, 193–226. 13
- Noelli, F. S., J. P. Brochado, et al. (2018). A linguagem da cerâmica guaraní: sobre a persistência das práticas e materialidade (parte 1). *Revista Brasileira de Linguística Antropológica* 10(2), 167–200. 9

- Nolan, B. (2004). First steps toward a computational rrg. In B. Nolan (Ed.), *Linguistic theory and practice: description, implementation and processing: 2004 International Role and Reference Grammar Conference*, pp. 196–223. Citeseer. 62
- Nolan, B. (2013). Constructions as grammatical objects. In B. Nolan and E. Diedrichsen (Eds.), *Linking Constructions into Functional Linguistics: The role of constructions in grammar*, pp. 143–178. Amsterdam: John Benjamins Publishing. 63
- O'Connor, R. (2008). A prosodic projection for role and reference grammar. In R. D. J. Van Valin Jr (Ed.), *Investigations of the SyntaxSemanticsPragmatics Interface*. Amsterdam: John Benjamins Publishing. 77
- Oshima, D. Y. (2007). Syntactic direction and obviation as empathy-based phenomena: A typological approach. *Linguistics* 45(4), 727–763. 323, 324
- Osswald, R. and L. Kallmeyer (2018). Towards a formalization of Role and Reference Grammar. In *Proceedings of the 2013 Conference on Role and Reference Grammar*. Unpublished conference proceedings. 62, 65, 66, 67, 72
- Overall, S. (2017). A typology of frustrative marking in Amazonian languages. In A. Y. Aikhenvald and R. M. W. Dixon (Eds.), *The Cambridge handbook of linguistic typology*, pp. 477–512. Cambridge: Cambridge University Press. 211
- Pavey, E. L. (2010). *The structure of language: An introduction to grammatical analysis*. Cambridge: Cambridge University Press. xiv, 330
- Payne, D. L. (1990). Morphological characteristics of lowland South American languages. In D. L. Payne (Ed.), *Amazonian linguistics: studies in lowland South American languages*, pp. 213–241. University of Texas Press. 25
- Payne, T. E. (1997). *Describing morphosyntax: A guide for field linguists*. Cambridge: Cambridge University Press. 8

- Pease, H. (2007). *Parintintin grammar*. Associação Internacional de Linguística. 206
- Pereira, A. A. (2009). *Estudo morfossintático do Asuriní do Xingu*. Ph. D. thesis, Universidade Estadual de Campinas. Unpublished PhD thesis. 155
- Peterson, J. (2007). Languages without nouns and verbs? An alternative to lexical classes in Kharia. In *Old and new perspectives on South Asian languages. Grammar and semantics. Papers growing out of the Fifth International Conference on South Asian Linguistics (ICOSAL-5), held at Moscow, Russia in July*, pp. 274–303. Delhi : Motilal Banarsidass Publishers. 100
- Pottier, B. (1983). *América Latina en sus lenguas indígenas*. UNESCO, Monte Avila Editores, CA. 20
- Praça, W. (1999). Nomes como predicados na língua tapirapé. Master's thesis, Universidade de Brasília. unpublished master Thesis. 314
- Praça, W. N. (2001). Sobre o indicativo ii no tapirapé. *Estudos sobre línguas indígenas 1*, 167–176. 314
- Praça, W. N., M. M. S. Magalhaes, and A. da Cruz (2017). Indicativo ii da família tupi guarani: uma questão de modo? *LIAMES: Línguas Indígenas Americanas 17*(1), 39–58. 317, 319
- Praça, W. N. (2007). Morfossintaxe da língua tapirapé. Master's thesis, Universidade de Brasília. Unpublished master's thesis. 155
- Progovac, L. (2015). *Evolutionary Syntax*. Oxford University Press. 132
- Pustejovsky, J. (1995). *The Generative Lexicon*. Cambridge: The MIT Press. 88
- Pustejovsky, J. and E. Jezek (2016). *Integrating Generative Lexicon and Lexical Semantic Resources*. LREC 2016 Tutorial. 88
- Queixalós, F. (2000). *Syntaxe Sikuani*. Louvain: Peeters. 104

- Queixalós, F. (2001). *Des noms et des verbes en tupi-guarani: état de la question*. Number 37 in *Lincom Studies in Native American Linguistics*. Lincom Europa. 100
- Queixalós, F. (2006). The primacy and fate of predicativity in Tupi-Guarani. In X. Lois and V. Vapnarsky (Eds.), *Lexical categories and root classes in Amerindian languages*, pp. 249–287. Peter Lang. 104
- Queixalós, F. (2010). Grammatical relations in katukina-kanamari. In S. Gildea and F. Queixalós (Eds.), *Ergativity in Amazonia*, Amsterdam, pp. 235–283. John Benjamins Publishing. 132
- Queixalós, F. (2022). Internal arguments in katukina-kanamari and beyond. *Cadernos de Etnolinguística* 10(1), 1–44. 113
- Ramirez, H., V. Vegini, and M. C. V. de França (2017). O warázu do guaporé (tupi-guarani): primeira descrição linguística. *LIAMES: Línguas Indígenas Americanas* 17(2), 411–506. 317
- Rauh, G. (2010). *Syntactic categories: Their identification and description in linguistic theories*. Oxford University Press. 97
- Reiter, S. (2012). *Ideophones in Awetí*. Ph. D. thesis, Christian-Albrechts Universität Kiel. Unpublished PhD thesis. 107, 245
- Restivo, P. ((2010 [1724])). *Arte de la lengua guaraní*. versión digital con un facsímile de la edición de 1724, por silvio luzzi. Talk at the Harvard-Leipzig Workshop on Workshop on Morphology and Argument Encoding. Available online: http://celia.cnrs.fr/FichExtPaleographiesGuaraniGuarani_intro.htm. 37, 180, 239, 351
- Ribeiro, D. (1996). *Diários índios. Os Urubus-Kaapor*. São Paulo: Companhia das Letras. 9
- Ribeiro, E. R. (2004). Prefixos relacionais em jê e karajá: um estudo histórico-comparativo. *LIAMES: Línguas Indígenas Americanas* 4(1), 91–101. 111

- Ribeiro, E. R. (2006). Subordinate clauses in karajá. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas* 1, 17–47. 26
- Ribeiro, E. R. (2009). Tapuya connections: language contact in eastern Brazil. *LIAMES: Línguas Indígenas Americanas* 9(1), 61–76. 12
- Ribeiro, E. R. (2011). Prefixos relacionais como evidência histórico-comparativa: os casos chiquitano e jabutí. *Línguas e Culturas Macro-Jê* 2, 105–120. 111
- Rijkhoff, J. (1990). Explaining word order in the noun phrase. *Linguistics* 28, 5–42. 78
- Rijkhoff, J. (2002). Verbs and nouns from a cross-linguistic perspective. *Rivista di linguistica* 14(1), 115–147. 78, 270
- Rijkhoff, J. (2007). Word classes. *Language and linguistics compass* 1(6), 709–726. 97
- Rijkhoff, J. and E. van Lier (2013). *Flexible word classes: Typological studies of underspecified parts of speech*. OUP Oxford. 97
- Roberts, J. R. (1988). Amele switch-reference and the theory of grammar. *Linguistic Inquiry* 19(1), 45–63. 348
- Rodrigues, A. (2010a). Tupi, tupinambá, línguas gerais e português do Brasil. In V. Noll and W. Dietrich (Eds.), *O português e o tupi no Brasil. São Paulo: Contexto*, pp. 27–47. *Revista Brasileira de Linguística Antropológica*. 13, 20, 23
- Rodrigues, A. D. (1951a). A composição em tupi. *Revista Brasileira de Linguística Antropológica* 3(1), 23–29. 36, 292, 294
- Rodrigues, A. D. (1952). Análise morfológica de um texto tupi. *Logos* 15, 56–77. 36
- Rodrigues, A. D. (1953). Morfologia do verbo tupi. *Letras* 1, 121–152. 36, 216, 314

- Rodrigues, A. D. (1958a). Die klassifikation des Tupi-sprachstammes. In *Thirty-second international congress of americanists*. Copenhagen: Munksgaard. 18
- Rodrigues, A. d. (1958b). *Phonologie der Tupinambá-Sprache*. Ph. D. thesis, Universität Hamburg. Unpublished PhD thesis. 23, 41, 43, 48, 51, 375
- Rodrigues, A. D. (1985). Relações internas na família lingüística tupí-guaraní. *Revista de antropologia* 27/28, 33–53. 18
- Rodrigues, A. D. (1986). *Línguas brasileiras: para o conhecimento das línguas indígenas*, Volume 11. Edições Loyola. 23
- Rodrigues, A. D. (1990). You and I = Neither You nor I. The personal system of Tupinambá. In D. Payne (Ed.), *Amazonian Linguistics: studies in lowland South America*, pp. 393–405. Austin: University of Texas Press. 109, 321
- Rodrigues, A. D. (1996a). Argumento e predicado em tupinambá. *Boletim da associação brasileira de lingüística* 19, 6–18. 21, 36, 43, 100, 111, 112, 117, 266, 314
- Rodrigues, A. D. (1996b). As línguas gerais sul-americanas. *PAPIA-Revista Brasileira de Estudos do Contato Linguístico* 4(2), 6–18. 8, 20
- Rodrigues, A. D. (1999a). Macro-Jê. In R. M. W. Dixon and A. Aikhenvald (Eds.), *The Amazonian Languages*, pp. 165–206. Cambridge: Cambridge University Press. 36, 48
- Rodrigues, A. D. (2004). Tupi. In R. M. W. Dixon and A. Y. Aikhenvald (Eds.), *The Amazonian languages*, Chapter 4, pp. 107–124. Cambridge: Cambridge University Press. 36
- Rodrigues, A. D. (2009). A case of affinity among tupí, karífb, and macro-jê. *Revista Brasileira de Linguística Antropológica* 1(1), 137–162. 36, 111, 118
- Rodrigues, A. D. (2010b). Estrutura do tupinambá. In A. S. A. C. Cabral and A. D. Rodrigues (Eds.), *Língua e culturas Tupi*, pp. 11–42. LALI e Nimuendajú. 36, 105, 159, 240

- Rodrigues, A. D. (2011a). Argumento e predicado em tupinambá. *Revista Brasileira de Linguística Antropológica* 3(1), 93–102. 132, 353
- Rodrigues, A. D. (2011b). Diferenças fonéticas entre o tupi e o guaraní. *Revista Brasileira de Linguística Antropológica* 3(2), 135–152. 19
- Rodrigues, A. D. (2011 [1951]b). Esboço de uma introdução ao estudo da língua tupí. *Revista Brasileira de Linguística Antropológica* 3(1), 31–44. 42
- Rodrigues, A. D. (2013). Alguns problemas em torno da categoria lexical verbo em línguas tupí-guaraní. *Revista Brasileira de Linguística Antropológica* 3(1), 103–114. 43
- Rodrigues, A. D. et al. (1997). Descripción del tupinambá en el período colonial: el arte de José de Anchieta. In *La descripción de las lenguas amerindias en la época colonial*, Volume 63, pp. 371–400. Vervuert. 21
- Rodrigues, A. D. and A. Cabral (2012). Tupían. In L. Campbell and V. Grondona (Eds.), *The Indigenous Languages of South America*, pp. 495–574. Berlin: Mouton de Gruyter. 18, 24, 37, 41, 105, 111, 118, 150, 348
- Rodrigues, A. D. and A. S. A. C. Cabral (2002). Revendo a classificação interna da família Tupí-Guaraní. *Línguas Indígenas Brasileiras. Fonologia, Gramática e História*, Atas do I Encontro Internacional do GTLI da AN-POLL. 18
- Rodrigues, A. D. and W. Dietrich (1997). On the linguistic relationship between mawé and tupí-guaraní. *Diachronica* 14(2), 265–304. 41, 150
- Rodrigues, D. M. G. (1974). Fonologia do guaraní antigo. Master's thesis, Universidade de Campinas. Unpublished master's thesis. 51
- Rodrigues, Aryon, R. (1999b). Tupí-Guaraní. In R. W. Dixon and A. Y. Aikhenvald (Eds.), *The Amazonian Languages*, pp. 107–122. Cambridge: Cambridge University Press. 37
- Rose, F. (2000). *Eléments de phonétique, phonologie et morphonologie de l'Émerillon (teko)*. *Mémoire de DEA en Sciences du Langage*. Université de Lyon II. 155

- Rose, F. (2002). My hammock = I have a hammock. In A. S. C. A. Rodrigues (Ed.), *Línguas Indígenas Brasileiras. Fonologia, Gramática e História. Atas do I Encontro Internacional do GTLI da ANPOLL*, pp. 325–332. CNPQ & Universidade federal do Para. 100
- Rose, F. (2003). *Morphosyntaxe de l'émérillon. Une langue tupi-guarani de Guyane française*. Ph. D. thesis, Université Lumière-Lyon II. 317
- Rose, F. (2005). Reduplication in Tupi-Guarani languages: going into opposite directions. In B. H. . V. Mattes (Ed.), *Studies on Reduplication*, pp. 351–368. Berlin: Mouton de Gruyter. 225
- Rose, F. (2009). A hierarchical indexation system: the example of Emerillon (Teko). In A. A. Patience Epps (Ed.), *New challenges in typology: Transcending the borders and refining the distinctions*, Volume 2, pp. 63–84. Berlin: Mouton de Gruyter. 107, 128, 131
- Rose, F. (2011). *Grammaire del L'Émérillon Teko, une langue Tupi-Guarani de Guyane Française*. Peeters. 37, 113, 136, 159, 165, 169, 206, 237, 336, 348, 351
- Rose, F. (2015a). On male and female speech and more: categorical gender indexicality in indigenous south american languages. *International Journal of American Linguistics* 81(4), 495–537. 26, 27
- Rose, F. (2015b). When "you" and "i" mess around with the hierarchy: a comparative study of tupi-guarani hierarchical indexing systems. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas* 10(2), 347–369. 27, 107, 131
- Rose, F. (2018). Are the tupi-guarani hierarchical indexing systems really motivated by the person hierarchy? *Typological Hierarchies in Synchrony and Diachrony* 121, 289. 118
- Salanova, A. P. (2004). Subtractive morphology in meben-gokre. Manuscript. University of Ottawa, available online at <http://aix1.uottawa.ca/asalanov/Docs/initial-truncation.pdf>. 118

- Salanova, A. P. (2009). Não existem prefixos relacionais nas línguas Jê. In S. L. B. Braggio and S. M. d. Sousa Filho (Eds.), *Línguas e culturas macro-jê*, pp. 259–271. UNICAMP/IEL. 111, 118
- Sampaio, T. (1906). Cartas tupis dos camarões. *Revista do Instituto Archeológico e Geographico Pernambucano* 12(68), 201–305. 36
- Sapir, E. (1921). *Language: An introduction to the study of speech*. New York: Harcourt, Brace & World. 97, 160, 225
- Say, S. (2021). Antipassive and the lexical meaning of verbs. In *Antipassive. Typology, diachrony, and related constructions*, pp. 177–212. Amsterdam: John Benjamins Publishing. 168
- Schaden, E. (1954). Os primitivos habitantes do território paulista. *Revista de História* 8(18), 385–406. 20
- Schleicher, C. O. (1998). *Comparative and internal reconstruction of Proto-Tupi-Guarani*. Ph. D. thesis, Ph. D. dissertation, University of Wisconsin–Madison. 41, 119, 285, 288, 352
- Schmidt-Riese, R. (1998). Perspectivas diacrônicas brasileiras: o rastro das línguas gerais. *Romanistisches Jahrbuch* 49(1), 27–55. 12, 20, 128
- Schmidt-Riese, R. (2016). Anchieta 1595 e Figueira 1621. representações da gramática do Tupinambá. In V. Noll and W. Dietrich (Eds.), *O Português e o Tupi no Brasil*, pp. 49–59. Editora Contexto. 22
- Schmidtke-Bode, K. (2015). *Complement clauses and complementation systems: A cross-linguistic study of grammatical organization*. Ph. D. thesis, Friedrich-Schiller-Universität Jena. unpublished PhD thesis. 350
- Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge university press. 98
- Searle, J. R. (1975). Indirect speech acts. In P. Cole and J. L. Morgan (Eds.), *Syntax and semantics: speech acts*, Volume 3, pp. 59–82. Brill. 198

- Seki, L. (1990). Kamaiurá (Tupí-Guaraní) as an active-stative language. In D. L. Payne (Ed.), *Amazonian linguistics: studies in Lowland South American languages*, pp. 367–392. Austin: University of Texas Press. 37, 100, 128, 131, 136, 206
- Seki, L. (2000). *Gramática do kamaiurá: língua tupi-guarani do alto Xingu*. Editora da UNICAMP. 37, 100, 155, 159, 209, 236, 266, 314, 317, 351
- Shibatani, M. and P. Pardeshi (2002). The causative continuum. *Typological studies in language* 48, 85–126. 159
- Siewierska, A. (1988). *Word order rules*. London: Croom Helm. 145
- Siewierska, A. (2001, April). On the argument status of cross-referencing forms: Some problems. *Revista Canaria de Estudios Ingleses* 42, 215–236. 189
- Siewierska, A. (2005). Alignment of Verbal Person Marking. In M. S. Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. 118
- Siewierska, A. (2009). Person asymmetries in zero expression and grammatical function. In F. F. . D. C. Anne Abeillé (Ed.), *Essais de Linguistique Generale et de Typologie Linguistique offerts au Professeur Denis Creisels à l'occasion de ses 65 ans.*, pp. 425–438. Presses de LEcole Normale Supérieure. 118
- Silva, A. P. d. C. (2020). *Contribuições para o Conhecimento da História da Língua e da Cultura Avá-Canoeiro*. Phd thesis, Universidade de Brasília. 37
- Silva, A. P. d. C. and A. S. A. C. Cabral (2013, 1). Expressões de correferencialidade na história de línguas tupí-guaraní. 64a Reunião Anual da SBPC. Unpubslihed manuscript. 348
- Silva, M. A. C., T. Ferraz, and T. Hünemeier (2022). A genomic perspective on south american human history. *Genetics and Molecular Biology* 3(45), 1–15. 9

- Silva, R. G. P. d. et al. (2010). *Estudo morfossintático da língua Sateré-Mawé*. Ph. D. thesis, Universidade Estadual de Campinas. Unpublished PhD thesis. 107, 150, 245
- Silva Magalhães, M. M. (2002). Aspectos fonológicos e morfosintáticos da língua guajá. Master's thesis, University of Brasília. Unpublished master's thesis. 155
- Solano, E. d. J. B. (2010). *Descrição gramatical da língua araweté*. Ph. D. thesis, Universidade de Brasília. Unpublished PhD thesis. 37, 136
- Song, J. J. (2014). *Linguistic typology: Morphology and syntax*. London: Routledge. 369
- Song, J. J. (2018). *Linguistic typology*. Oxford University Press. 8, 97
- Sousa, G. S. d. (1851). *Tratado descritivo do Brasil em 1587*. Rio de Janeiro: Typographia Universal de Laemmert. 11, 13, 19, 22
- Sousa, S. A. (2017). *Elementos do léxico e da gramática Apiaká (subramo VI da família linguística Tupí-Guaraní)*. Ph. D. thesis, Universidade de Brasília. unpublished PhD thesis. 155
- Souto Maior, P. (1912). Dous índios notáveis e parentes próximos. *Revista do Instituto Histórico Arqueológico Pernambucano* 15, 61–71. 36
- Spencer, A. and A. R. Luís (2012). *Clitics: an introduction*. Cambridge: Cambridge University Press. 105
- Staden, H. (1978[1557]). *Wahrhaftige Historia und Beschreibung einer Landschaft der wilden, nackten, grimmigen Menschenfresser, in der Neuen Welt Amerika gelegen*. Kassel: Thiele & Schwarz. 11, 13, 14, 16, 20, 29, 30, 31, 38, 48, 274
- Staden, H. (2008). *Hans Staden's True History: An Account of Cannibal Captivity in Brazil*. Duke University Press. 29, 30
- Swadesh, M. (1938). Nootka internal syntax. *International Journal of American Linguistics* 9(2/4), 77–102. 100

- Thevet, A. (1558). *Les singularitez de la France antarctique, autrement nommée Amerique: & de plusieurs terres & isles decouvertes de nostre temps. Par F. André Thevet, natif d'Angoulesme*. Chez les heritiers de Maurice de la Porte, au Clos Bruneau, à l'enseigne S. Claude. 30, 31
- Thevet, A. (1953). *Les Français en Amérique pendant la deuxième moitié du XVI siècle: Le Brésil et les Brésiliens, par André Thevet*, Volume 1. Paris: Presses universitaires de France. 13, 15, 274
- Tikkanen, B. (2001). Converbs. In W. O. Martin Haspelmath, Ekkehard König and W. Raible (Eds.), *Language typology and language universals: An international handbook*, pp. 1112–1123. Berlin: De Gruyter Mouton. 352
- Trommer, J. (2007). On portmanteau agreement. Talk at the Harvard-Leipzig Workshop on Workshop on Morphology and Argument Encoding. Available online: <https://home.uni-leipzig.de/jtrommer/Harvard/trommer.pdf>. 107
- Ullrich, J. and R. D. Van Valin Jr (2007, 11). Nuclear structures: Passive and noun incorporation in lakhota. Unpublished. RRG Conference - Tokyo. 161
- Vallejos, R. (2016). *A Grammar of Kukama-Kukamiria: A language from the Amazon*. Brill. 37
- van der Auwera, J., N. Dobrushina, and V. Goussev (2013). Imperative-hortative systems (v2020.3). In M. S. Dryer and M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Zenodo. 206
- Van Gijn, R. (2014). Subordination strategies in South America: nominalization. In *The native languages of South America: Origins, development, typology*, pp. 274–296. Cambridge: Cambridge University Press. 369
- van Gijn, R. and J. Hammond (2016). *Switch Reference 2.0*, Volume 114. Amsterdam: John Benjamins Publishing. 107, 349
- van Groesen, M. and L. E. Tise (2019). *Theodor de Bry. America*. Cologne: Taschen. 14

- Van Valin Jr, Robert, D. (2021). Cosubordination. In R. D. Van Valin Jr (Ed.), *Challenges at the Syntax-Semantics-Pragmatics Interface*, pp. 241–256. Cambridge: Cambridge scholars publishing. 341
- Van Valin Jr, Robert, D. (2022). Principles of role and reference grammar. In R. D. Van Valin Jr (Ed.), *Cambridge Handbook of Role and Reference Grammar*, pp. 1–140. Cambridge: Cambridge University Press. xii, xv, 60, 71, 80, 90, 94, 100, 273, 344, 345, 351, 365
- Van Valin Jr, R. (2001a). Semantic macroroles in rrg [unpublished manuscript]. University of Buffalo, New York. xii, 85, 90, 133, 135
- Van Valin Jr, R. (2016). Van Valin Jr, r. d. on the status of lexical categories in rrg. Slides of presentationURL: <https://ubwp.buffalo.edu/rrg2019/wp-content/uploads/sites/101/2019/09/Van-Valin.pdf>. 100
- Van Valin Jr, R. D. (1977). *Aspects of Lakhota syntax*. Ph. D. thesis, University of California, Berkeley. Unpublished PhD thesis. 188
- Van Valin Jr, R. D. (1985). Case marking and the structure of the lakhota clause. In A. C. Nichols, Johanna Woodbury (Ed.), *Grammar inside and outside the clause*, pp. 363–413. Cambridge: Cambridge University Press. 188
- Van Valin Jr., R. D. (1986). Pragmatics, island phenomena, and linguistic competence. In A. M. Farley, P. T. Farley, and K.-E. McCullough (Eds.), *Papers from the 22nd Regional Meeting of the Chicago Linguistic Society*, Volume 22, pp. 223–233. Chicago Linguistic Society. 314
- Van Valin Jr, R. D. (1987). The role of government in the grammar of head-marking languages. *International Journal of American Linguistics* 53(4), 371–397. 123
- Van Valin Jr, R. D. (1990). Layered syntax in role and reference grammar. In J. Nuyts, A. M. Bolkestein, and C. Vet (Eds.), *Layers and levels of representation in language theory*, pp. 193–231. Amsterdam: John Benjamins Publishing. 65

- Van Valin Jr, R. D. (1991). Functionalist linguistic theory and language acquisition. *First Language* 11(31), 7–40. 60
- Van Valin Jr, R. D. (1993). A synopsis of Role and Reference Grammar. In R. D. Van Valin Jr (Ed.), *Advances in role and reference grammar*, pp. 1–164. Amsterdam: John Benjamins Publishing. 60, 65
- Van Valin Jr, R. D. (1999a). Generalized semantic roles and the syntax-semantics interface. In I. F. Corblin, C. Dobrovie-Sorin, and J.-M. Marandin (Eds.), *Empirical issues in formal syntax and semantics*, Volume 2, pp. 373–389. The Hague: Thesus Holland Academic Graphics. 90
- Van Valin Jr, R. D. (1999b). A typology of the interaction of focus structure and syntax. In E. Raxilina and J. Testelec (Eds.), *Typology and the theory of language: From description to explanation*, pp. 511–524. Languages of Russian Culture. 75
- Van Valin Jr, R. D. (2001b). The acquisition of complex sentences: a case study in the role of theory in the study of language development. *Chicago Linguistic Society Parasession* 36, 511–531. 60, 331
- Van Valin Jr, R. D. (2003). Functional linguistics. In M. Aronoff and J. Resnik (Eds.), *The handbook of linguistics*, pp. 319–336. Wiley Online Library. 60
- Van Valin Jr, R. D. (2004). Semantic macroroles in Role and Reference Grammar. In R. Kailuweit (Ed.), *Semantische Rollen*, pp. 62–82. Gunter Narr Verlag. 91
- Van Valin Jr, R. D. (2005). *Exploring the syntax-semantics interface*. Cambridge: Cambridge University Press. xii, xiii, xv, 59, 60, 61, 62, 63, 64, 67, 71, 72, 77, 82, 85, 86, 87, 88, 90, 91, 92, 94, 95, 121, 123, 133, 135, 158, 166, 193, 283, 329, 348, 351, 366
- Van Valin Jr, R. D. (2006a). On the relationship between syntactic theory and models of language processing. In B. C. u. A. D. F. Ina Bornkessel,

- Matthias Schlesewsky (Ed.), *Semantic role universals and argument linking: Theoretical, typological and psycho-neurolinguistic perspectives*, pp. 263–302. Berlin: Mouton de Gruyter. 62
- Van Valin Jr, R. D. (2006b). Semantic macroroles and language processing. In I. Bornkessel, M. Schlesewsky, B. Comrie, and A. D. Friederici (Eds.), *Semantic role universals and argument linking: Theoretical, typological, and psycholinguistic perspectives*, Volume 165, pp. 263–302. Berlin: Walter de Gruyter. 62, 90
- Van Valin Jr, R. D. (2007). Recent developments in the role and reference grammar theory of clause linkage. *Language and Linguistics* 8(1), 71–93. 334
- Van Valin Jr, R. D. (2008a). *Investigations of the Syntax Semantics Pragmatics Interface*. Amsterdam: John Benjamins Publishing. 60
- Van Valin Jr, R. D. (2008b). RPs and the nature of lexical and syntactic categories in Role and Reference Grammar. In R. D. Van Valin Jr (Ed.), *Investigations of the syntax-semantics-pragmatics interface*, pp. 161–78. Amsterdam: John Benjamins Publishing. 100, 194
- Van Valin Jr, R. D. (2009). Linguistics past and present: A view from the rhine. In Z. E. Fernández, A. Álvarez González, and M. B. Carpio (Eds.), *Ser lingüista: un oficio diverso y polifacético. Diez años de una Maestría en Lingüística*, pp. 155–162. Hermosillo: Editorial Universidad de Sonora. 60
- Van Valin Jr, R. D. (2013). Head-marking languages and linguistic theory. In B. Bickel, L. A. Grenoble, D. A. Peterson, and A. Timberlake (Eds.), *Language Typology and Historical Contingency: In honor of Johanna Nichols*, pp. 91–123. Amsterdam: John Benjamins Publishing. 69, 70, 123, 161, 163, 188, 189, 190
- Van Valin Jr, R. D. (2014). On the place of information structure in a grammar. unpublished. 75

- Van Valin Jr., R. D., B. Devriendt, L. Goossens, and J. van der Auwera (1996). Toward a functionalist account of so-called extraction constraints. In B. Devriendt, L. Goossens, and J. van der Auwera (Eds.), *Complex Structures: A Functionalist Perspective*, pp. 29–60. Berlin: Walter de Gruyter. 314
- Van Valin Jr, R. D. and R. J. LaPolla (1997). *Syntax: Structure, meaning, and function*. Cambridge: Cambridge University Press. 60, 61, 62, 63, 65, 66, 67, 71, 80, 84, 91, 92, 93, 94, 107, 133, 138, 139, 157, 171, 189, 197, 219, 254, 306, 308, 310, 325, 326, 327, 328, 329, 334, 348, 361
- Van Valin Jr, R. D. and M. R. Usón (2014). Interfacing the lexicon and an ontology in a linking system. In M. d. l. . G. González, F. J. R. d. M. Ibáñez, and F. González-García (Eds.), *Theory and practice in functional-cognitive space*, pp. 205–228. Amsterdam: John Benjamins Publishing. 63
- Vapnarsky, V. (2013). Is yucatec maya an omnipredicative language? predication, the copula and focus constructions. *STUF-Language Typology and Universals Sprachtypologie und Universalienforschung* 66(1), 40–86. 104
- Velázquez-Castillo, M. (1995). Noun incorporation in guaraní: a functional analysis. *Linguistics* 33, 673–709. 160, 165
- Velázquez-Castillo, M. (1996). *The grammar of possession: Inalienability, incorporation and possessor ascension in Guaraní*. Amsterdam: John Benjamins Publishing. 299
- Velázquez-Castillo, M. (2002). Guaraní causative constructions. In *The grammar of causation and interpersonal manipulation*, pp. 507–534. Amsterdam: John Benjamins Publishing. 149, 159
- Vendler, Z. (1967). *Linguistics in philosophy*. Cornell University Press, London. 80, 230
- Verhoeven, E. and N. Lehmann (2018). Self-embedding and complexity in oral registers. *Glossa : a journal of general linguistics* 3(1), 1–30. 301

- Vieira, M. M. D. (2014). A manifestação de tópico e foco em línguas da família tupi-guarani. *DELTA: Documentação e Estudos em Linguística Teórica e Aplicada* 30(3), 659–683. 314
- Vieira, M. M. D. (2018). Recursion in Tupi-Guarani Languages: The Cases of Tupinambá and Guarani. In L. Amaral, M. Maia, A. Nevins, and T. Roeper (Eds.), *Recursion across Domains*, pp. 166184. Cambridge: Cambridge University Press. 302
- Villafañe, L. (2004). *Gramática Yuki lengua Tupi-Guarani de Bolivia*. Ph. D. thesis, Katholieke Universiteit Nijmegen,. Unpublished PhD thesis. 37, 206
- Von Heusinger, K. (1999). *Intonation and information structure*. Ph. D. thesis, Citeseer. 77
- Von Humboldt, W. (2009 [1836]). *Über die Verschiedenheiten des menschlichen Sprachbaues*. Cambridge: Cambridge. 97
- von Martius, C. F. (1867). *Beiträge zur Ethnographie und Sprachenkunde Amerika s zumal Brasiliens*. Leipzig: Friedrich Fleischer. 18
- von Prince, K., A. Krajinović, and M. Krifka (2022). Irrealis is real. *Language* 98(2), 221–249. 209
- Wagley, C. and E. Galvão (1949). *The Tenetehara Indians of Brazil*. New York: Columbia University Press. 9, 22
- Willett, T. (1988). A cross-linguistic survey of the grammaticization of evidentiality. *Studies in Language. International Journal sponsored by the Foundation Foundations of Language* 12(1), 51–97. 207
- Ziebell-Wendt, Z. (1991). Die Tupinamba: Realität und Fiktion in den Berichten des 16. Jahrhunderts. In K. Kohut, J. Bär, E. G. Valdes, S. H. Groenwold, and H. Pietschmann (Eds.), *Der eroberte Kontinent: Historische Realität, Rechtfertigung und literarische Darstellung der Kolonisation Amerikas*, pp. 227–245. Frankfurt: Vervuert. 29

- Zúñiga, Fernando & Kittilä, S. (2020). *Grammatical voice*. Cambridge: Cambridge University Press. 149, 154, 162, 170, 173
- Zwartjes, O. (2011). *Portuguese missionary grammars in Asia, Africa and Brazil, 1550-1800*, Volume 117. Amsterdam: John Benjamins Publishing. 19, 22, 23
- Zwicky, A. M. and G. K. Pullum (1983). Cliticization vs. inflection: English n't. *Language* 59(3), 502–513. 104

Subject Index

- Accomplishments, 234
- Actor-undergoer hierarchy,
 - 85
- Adjunct, 257
- Adjuncts, 67
- Adverb, 119, 219
- Adverbial
 - fronted, 314
- Adverbs, 119
 - aspectual, 191
- Adversative coordination,
 - 333
- Agentive nominalizer, 287
- Agglutinative language, 24
- Aktionsart, 229
- Alethic modality, 210
- Alienability, 245
- Alignment pattern, 123
- Allative, 264
- Anambé, 318
- Antipassive, 168
 - human, 169
- Apiaká, 155
- Applicative), 162
- Arawete, 318
- Argument indexes, 105
- Aryon Rodrigues, 36
- Aspect, 223
 - completive, 223
 - frequentative, 225
 - habitual, 225
 - imperfective, 219
 - inceptive, 259
 - iterative, 226
- Asuriní Tocantins, 155, 318
- Asuriní Xingu, 318
- Asurní Xingu, 155
- Attributive modifiers, 194
- Augmentative, 242
- Auxiliaries, 359
- Ava-canoeiro, 318
- Awetí, 112, 150

- Basic Clause Patterns, 123
- Case
- allative, 265
 - dativative, 263, 265
 - locative, 263, 265
 - locative body parts, 264
 - perlative, 266
 - translative, 267
- Causative, 149
- comitative, 159
 - lexical, 149
 - morphological, 149
 - sociative, 159
 - transitive verbs, 154
- Causatives
- morphological, 149
- Chiriguano, 155, 318
- Circumstantial mood, 314
- Clausal subordination, 337
- Clause juncture, 328
- Clause linkage marker, 331, 332
- Clitic
- interrogative, 310, 313
 - question, 313
- Cohortative, 206
- Collective, 245
- Complex RP, 361
- Complex sentences, 325
- Compound, 292
- Conominal, 188
- Consonants, 42, 43
- Constituent projection, 67
- Construction Grammar, 60, 61
- Constructional schemas, 63
- Coordination, 331, 361
- adversative, 333
- Core, 67
- negation, 222
 - nominal, 77
- Core argument
- non-macrorole, 255
- Core cosubordination, 343, 346
- Coreferential, 349
- Cosubordination, 329, 340, 341
- core, 348
- Dative shift, 135, 137
- Deixis, 282
- Deliberative, 259
- Demonstrative, 282
- adnominal, 282
 - nonvisible, 284
 - pronominal, 270
- Demonstratives, 270, 282
- Desiderative, 259
- Detached position, 319
- Detached positions, 67
- Diminutive, 241
- Disjunctive coordination, 364
- Ditransitive verbs, 133, 135
- ECS, 335
- Emerillon, 318
- Evidential, 207, 208

- Exhortative, 206
 Experiencer, 231
 extra-core slot, 69, 70
 Extraction, 314
- Factive, 155
 Female speech, 26
 Finite clause, 335
 Focus, 74, 190, 314
 domain, 307
 predicate, 307
 Focus domain
 actual, 77
 potential, 77
 Focus projection, 67, 72, 75
 Free pronouns
 dative, 266
 free pronouns, 105
 Frustrative, 211
 concessive, 211
 lusive, 211
 Future
 epistemic, 221
- Generic index, 109
 Gerund, 107, 262, 291, 342, 344,
 352
 causal, 354
 Grammatical relations, 138
 Guajajara, 155
 Guajajara, 318
 Guajá, 37, 155, 314, 317,
 318
 Guaraní, 19, 37, 149, 155,
 318
 Guarayo, 155, 208, 318
- Habitual agent, 291
 Hans Staden, 29
 Head-marking, 69, 123,
 188
 Hortative, 197, 206, 220, 221,
 262
- Illocutionary Force
 declarative, 199
 hortative, 206
 Illocutionary force, 197
 exclamative, 262
 Imperative, 202
 negative, 203
 Impersonal, 180
 impersonal, 110, 111
 Incomplete action, 224
 Incorporation, 135, 160, 161,
 343
 Independent pronouns, 106
 Indicative-II, 314
 Information packaging, 98
 Information Structure, 305
 Intensional, 259
 Interclausal relations hierarchy,
 351
 Interclausal semantic relation,
 351
 Interclausal semantic relations,
 345, 351
 Interrogative, 201
 Intransitive bias, 132
 Irrealis, 209
- Juncture
 nuclear, 157

- Jussive, 206
- Ka'apor, 318
- Kaiwá, 318
- Kamajurá, 37, 155, 314, 317
- Kamayurá, 318
- Kayabi, 318
- Layered Structure of the Clause, 183
- Layered structure of the clause, 65, 68
- Layered structure of the RP, 77
- Layered structure of the word, 163
- Left-detached position, 183
- Lexical categories, 229
- Linking algorithm, 62
- Lusive, 241
- M-transitive verbs, 130
- M-transitivity, 149
- Macro-Event-Property, 353
- Macrorole, 177
- Macroroles, 90, 133
- Male speech, 26
- Markedness, 99
- Mawé, 112, 150
- Mbya, 314, 318
- Mbyá, 302
- Middle voice, 173
- Minimal pairs, 43
- Modality, 219
 - alethic, 210
- Modification
 - compound, 292
- Modifier phrase, 194, 292
- Modifiers, 67
 - locative, 67
 - temporal, 67
- Morpheme, 280
- Mundurukú, 112, 119, 150, 245
- Narrow focus, 313
- Negation
 - double, 214, 215, 280, 281
- Nexus relations, 329
- Nexus types, 329
- Nheengatu, 19, 37, 142, 317, 318, 336, 364
- Nominal Qualia, 88
- Nominal tense, 236
- Nominalizer, 285
 - general, 290
 - patient, 288
 - relative, 285
 - resultative, 288
- Non-arguments, 191
- Non-macrorole argument, 133
- Nonfocal, 319
- Noun classes, 104
- Nucleus, 67
 - complex, 341, 342
 - nominal, 77
- Numeral, 274, 275
- Oblique core argument,

- 257
- Oblique-topicalized, 314
- Old Guaraní, 318
- Old Guaraní, 37, 146, 149, 155,
180, 317
- Operator projection, 67
- operator projection, 70
- Operators, 67, 195
RP, 282
- Optative, 204, 259
- Orthography, 9
- OV languages, 132
- Parakanã, 155, 318
- Parataxis, 332
- Parintintin, 314, 318
- Particle
deliberative, 259
desiderative, 259
dubitative, 261
inceptive, 259
intentional, 259
- Particles, 259
aspect, 259
core modifier, 259
discourse, 260
illocutionary force,
262
- Passive voice, 179
- Periphery, 67, 191
nuclear, 77
- Perlative, 267
- Permissive, 206
- Person hierarchy, 131
- Person index
focal, 321
nonfocal, 321
- Person markers, 104
- Phonemic consonants, 43
- Phonology, 41, 42
- Phrasal adjuncts, 191
- Polysynthesis, 24
- Portmanteau indexes, 107
- Positional auxiliaries, *see*
Auxiliaries
- Possessor-stranding, 136
- Post core slot, 68
- Post-detached position, 67
- Postcore slot, 67
- Postpositional clitics, 252
- Postpositional Phrase, 256
- Postpositional phrase, 257
adjunct, 257
- Postpositions, 246, 252
- Potential focus domain,
309
- Pre-detached position, 67,
309
- Precore slot, 67
- Predicate focus, 76
- Predication, 101
classificational, 140
equational, 141, 142
existential, 101, 126,
141
nonverbal, 140
predlocative, 142
predpossessive, 101,
141
verbal, 101, 126

- Predicative postposition, 256
 Predlocation, 145
 Prefixes, 148
 Privative, 280, 281
 double, 281
 negation, 280
 Proper noun, 366
 Propositional act function, 98
 Prosody, 58
 PSA, 138

 Quantifiers, 273, 274

 Realis, 209
 Reciprocal, 177, 178
 Reduplication, 225, 244, 342, 343
 disyllabic, 226
 monosyllabic, 225
 plural, 244, 310
 Reference phrase, 269
 Reflexive, 166
 Relational markers, 111
 Relational morphemes, 111
 Relative clauses, 366
 Relativizer, 367
 Restrictive relative clauses, 368
 RP, 269
 RPIP, 283, 295
 RRG, 59

 Semantic macroroles, 86

 Semantic representation of nominals, 88
 Semantic roles, 86
 Sirinono, 318
 Sirionó, 31
 Sociative applicative, 159
 Sociative causative, 159
 Speech act function
 modification, 102
 reference, 101
 Split intransitive, 128
 Staden, 31
 State-predicate, 230
 Status, 209, 210
 Stress, 42, 58
 Subordination, 334
 daughter, 334
 peripheral, 334
 Suruí, 318
 Switch-reference, 348
 Syllable, 42, 56
 Syllable structure, 56

 Tapiete, 37, 318
 Tapirapé, 155, 314, 318
 Tekó, 37, 155, 314, 317, 336
 Tembé, 318
 Tense, 215
 Thematic relations, 86
 Topic, 74, 190, 309, 319
 Transitivity, 123
 Translative, 209, 319
 Tupinambá, 318

 Unstressed suffixes, 134,

- 247
- Valency, 149
- Valency change, 148
- Voiceless stops, 47
- Vowels, 42, 54
- Warazu, 317, 318
- Wayampi, 31, 318
- WH-questions, 68
- WH-words, 202, 310
- Wolf Dietrich Rodrigues,
36
- Word classes, 97
- Word order, 132, 145
- Yuki, 318
- Zo'e, 155

A Role and Reference Grammar Description of Tupinambá

Tupinambá is the first attested language of the Tupí-Guaraní family and it has a story that one can follow from its first attestation to the present through its descendants. Making use of RRG, a linguistic theory that is informed by cross-linguistic diversity, I present the first typologically adequate description of Tupinambá. This description introduces the main aspects of Tupinambá grammar, including phonology, morphology, syntax, and information structure, accounting for the interface between syntax, semantics, and pragmatics.

»This is the first comprehensive description of Tupinambá. The language was spoken on great parts of the Atlantic coast of Brazil when Jesuits, since the middle of the 16th century, began to use it for mission. In the forming of Brazil as an independent nation, Tupinambá became a national symbol: Nearly all the endemic Brazilian fauna and flora has Tupinambá names. It is present until now in lots of names of streets, place-names and geographical names, meals, fruits, first names of persons, and popular expressions in Brazilian Portuguese. In spite of being a natural element in Brazilian life and although there have been partial descriptions in early and also in modern times, Tupinambá never has been the object of a comprehensive syntax, morphology, and phonology.

Fabrício Ferraz Gerardi presents us with a convincing modern linguistic analysis on the basis of a corpus of texts written by missionaries of the 16th and 17th centuries. All we know about Tupinambá language is from Jesuit missionaries (catechisms, grammars, theater, poetry). This book stands for a significant advance not only in the analysis of Tupinambá syntactic structure but even in our understanding of how function the numerous languages of the whole linguistic family called Tupi-Guarani. This is a very important contribution to the linguistics of the languages of Brazil.«

Prof. Dr. Wolf Dietrich

