Topics on the syntax of Kawahíva: A Tupí-Guaraní language from the Brazilian Amazon

by

Wesley Nascimento dos Santos

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Linguistics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Lev Michael, Co-chair
Professor Peter Jenks, Co-chair
Professor Line Mikkelsen

Summer 2024

Topics on the syntax of Kawahíva: A Tupí-Guaraní language from the Brazilian Amazon

Abstract

Topics on the syntax of Kawahíva: A Tupí-Guaraní language from the Brazilian Amazon

by

Wesley Nascimento dos Santos

Doctor of Philosophy in Linguistics

University of California, Berkeley

Professor Lev Michael, Co-chair

Professor Peter Jenks, Co-chair

This dissertation provides a description and analysis of the syntax of Kawahíva, a critically endangered Tupí-Guaraní language spoken by approximately 560 people in the Brazilian Amazon basin. The first part of this dissertation includes an overview of Kawahíva documentation and revitalization efforts, its phonology, and a comprehensive grammatical sketch of the language.

In the second half of this dissertation, I analyze two key topics of Kawahíva syntax: clause structure and relativization. In particular, I offer an analysis of VSO word order of Kawahíva clause structure. I argue that verb-initial word order in Kawahíva results from long head movement, a type of syntactic head movement (Harizanov and Gribanova 2019), similar to phrasal movement. This analysis is supported by the hallmarks of syntactic movement in Kawahíva verb movement, including interpretive semantic effects and nonlocality. I also show that two theoretical accounts of the V1 order – Remnant VP Movement and Head Movement – are insufficient to derive the V1 order in Kawahíva. This study is the first to demonstrate that a language can use long head movement as a general principle to create the verb-initial order.

The final chapter of the dissertation investigates relativization in Kawahíva. Several languages are documented with a relativization strategy where a clausal nominalization is used as an adnominal modifier. It follows from this that relativization in these languages is achieved by nominalization morphology itself, without any additional syntactic processes taking place (Comrie and Thompson 1985; Keenan 1985; Andrews 2007; Shibatani 2009). However, I argue that characterizing these structures as nominalizations in Kawahíva is insufficient and that there is additional clear evidence that the nominalized clause also involves a distinct operation, namely relativization. Evidence for this claim will come from showing that adnominal nominalizations exhibit the hallmark properties of relative clauses (and other extraction-based constructions) cross-linguistically: i) sensitivity to island effects and ii) the formation of a long-distance dependency between the gap in the nominalization and the filler (i.e., the modified noun).

À memória da minha mãe-avó, Joana Darthe da Silva Nascimento, e aos Juma gã e os Jupaú gã.

Contents

List of Figures

List of Tables

List of Abbreviations

1	first person
2	second person
3	third person
A	SET A, subject agreement
ABL	ablative
ACC	accusative
AF	distant
AG.NMLZ	agent nominalizer
ANT	anterior
APPL	applicative
ASP	aspect
ATEN	attenuative
ATR	attributive
AUG	augmentative
AV	active voice
B	SET B, object agreement
C	consonant
CAUS	causative
CIRC	circumstantial mood
CM	class marker
COMP	complementizer
CONT	continuative
COR	co-referent
CPL	completive
CS	case marker
CV	circumstantial voice
DAT	dative
DIM	diminutive
DIST.PST	distant past
DEF	definite

DEM	demonstrative
DIR.EVID	direct evidential
DIST	distal, distant
DUB	dubitative
ERG	ergative
EV	event nominalizer
EXCL	exclusive
FEM	feminine
FM	female speaker
FOC	focus
FREQ	frequentative
FRUST	frustrative
FS	end of sentence
FUT	future tense
GEN	genitive
HAB	habitual
HYP	hypothetical
ICP	incompletive
IMP	imperative
IRR	irrealis mood
INCL	inclusive
INDF	indefinite
INTR	intransitive
JUA	ISO code for Juma dialect
LK	linking morpheme
LOC	locative
LOC	locative voice
MASC	masculine
MED	medial
MF	male speaker
N	nuclear case
NEG	negation
NF	non-finite
NMLZ	nominalizer
NOM	nominative
NOM.FUT	nominal future tense
NOM.NEG	nominal negation
NOM.PST	nominal past tense
NON.SPEC.INDF	nonspecific indefinite
NS	nonsubject voice
NWIT	non-witnessed
OPT	optative mood
OV	object voice
P	patient

PFV	perfective
PIV	topic marking
PL	plural
POT	potential
POST	postposition
PRF	perfect
PRO	pronoun
PROGR	progressive
PROJ	projective aspect
PROX	proximal
PST	past tense
PV	patient voice
REAL	realis mood
RECP	reciprocal
RED	reduplicant
REFL	reflexive
REL	relativizer
REM	remote
RN	relational noun
S	subject
SA	the unique argument of active verbs
SG	singular
SO	the unique argument of stative verbs
SOC	sociative
TOP	topic
TRANS	transitive
URZ	ISO code for Jupaú (also known as Uru Eu Wau Wau)
V	vowel
WH.INTR.SUBJ	*wh*-agreement marker *-va'e* for intransitive subjects
WH.OBJ.I	*wh*-agreement marker *remb-* for transitive objects
WH.OBJ.II	*wh*-agreement marker *-pyr* for transitive objects
WH.OBL.I	*wh*-agreement marker *-hav* for a subset of obliques
WH.OBL.II	*wh*-agreement marker *-var* for a subset of obliques
WH.TRANS.SUBJ	*wh*-agreement marker *-har* for transitive subjects
WIT	witnessed

Acknowledgments

This dissertation would not have been possible without the Juma gã and Jupaú gã, who welcomed me into their houses, shared their food, and helped me navigate life in the Amazon. Their strength and dedication to unknown endeavors have always inspired me to continue working on this project, and their humor—much like that of Brazilian *Nordestinos* like myself—helped me feel I was at home surrounded by my close relatives during my time in the field.

Some Juma gã and Jupaú gã have contributed substantially to this dissertation, and I want to thank them by name, including my two friends, collaborators, and teachers, Mandeí Juma and Mandá Uru Eu Wau Wau (Jaikara village). Mandeí hẽa has been my primary teacher of the language and one of the most amazing individuals I have had the pleasure to meet. Her humor, intelligence, and bravery are unparalleled. Mandá hẽa taught me the Kawahiva language and culture in a nuanced and insightful way. For example, she drew my attention to a 'dialect of gender' in Jupaú: only female speech is transparent to long-distance nasal harmony at morpheme boundaries (see §3.2.3.2). She did so by pointing out that my pronunciation of Kawahíva was "too nasal", meaning that I sounded feminine for someone who appeared male.

Several other Kawahíva individuals have helped me several times by providing their time to check grammaticality judgments, translations, storytelling of a traditional narrative or personal story, or helped document the language by recording their parents and relatives. They should be thanked by name (with village name when needed for disambiguation): Ariva'ea (†2021), Bitaté ga, Boakara ga (Jaikara village), Boajupava'ea (623 village), Boreá hẽa (623 village), Boreá hẽa (Juma village), Boroap hẽa (Alto Jaru village), Boropó hẽa (Jaikara village), Kwari ga (Jamari village), Makana ga (623 village), Mandeí hẽa (Jaikara village), Maytá hẽa (Juma village), Puré ga (Jaikara village), Tangãi ga (Jaikara village), Tejuvi hẽa (Juma village), and Warina ga (Jamari village).

I would not have met the Kawahíva were it not for an invitation from esteemed colleagues at the Museu Paraense Emílio Goeldi to help survey the number of speakers of the Indigenous languages in the state of Rondônia, Brazil. Denny Moore, one of the big supporters of fieldwork and language documentation in South America, served as a mentor on my first field trips. Denny was also the one who initially encouraged me to pursue a PhD abroad and supported me in several ways, including with the fees for the US graduate school application and language proficiency exams. Together with Ana Vilacy Galucio and Hein van der Voort, they also provided feedback on my first attempts at the phonology and grammar of Kawahíva.

In Berkeley, I was closely mentored by three faculty who served on my dissertation committee and for whom I developed a strong respect and admiration. I could not have asked for a more knowledgeable fieldwork and language documentarian than my primary advisor Lev Michael. Among many remarkable skills, Lev's feedback on framing my work to a broader audience stands out. Peter Jenks took on the challenge of co-advising someone with little experience in formal syntax after I took his 2019 undergrad syntax course. His way of presenting the challenges, thinking, and responses to cross-linguistic phenomena in syntax helped me develop a strong connection with this field, and his unwavering support made me confident that I could do this type of work. I have learned so much from Line Mikkelsen, who personifies the language fieldworker,

theoretical linguist, and professor I aspire to be, and whose work demonstrates how one can be elegantly clear and concise. I also had the privilege of serving as a Graduate Student Instructor (GSI) for the three of them, and I will always remember students smiling at their excitement during lectures and laughing at silly jokes and examples.

I am also thankful for having faculty who shared feedback on other projects, including Amy Rose Deal and Hannah Sande, and the staff who made my life simpler as an international student and simplified my fieldwork logistics and accounting, Belén Flores, Paula Flores, Jhonny Arellano, Siti Keo, and Susan Luong.

Here, I was also fortunate to have met the best cohort: Maddie Bossi, Emily Drummond, Emily Grabowski, and Lupita Vasquez. I love you all! Your friendship and alliance have inspired and helped me get through the downs of graduate school (and life!) all these years.

My friends and colleagues at the Department of Linguistics deserve appraisal for their camaraderie – Myriam Lapierre, Zach O'Hagan, Allegra Robertson, Amber Galvano, Ana Lívia Agostinho, Anna Björklund, Anna Macnick, Becky Jarvis, Dakota Robinson, Edwin Ko, Franco Liu, Katie Russell, Julia Peck, Justin Royer, Maks Dąbkowski, Marta Schwarz, Måsi Santos, Shweta Akolkar, Tessa Scott, Tzintia Montaño, Tyler Lemon, Wendy Lopéz Márquez, and Zach Wellstood.

My non-linguist friends—Henryque Luna, Ben Goldstein, Maddie Wood, Alex Kushner, and Freddie Martinez—are the best people I know. Your friendship has helped me stay sane in this turmoil of life as a graduate student.

I owe just as much to my dear friend Az Hussain and his mom Ali. Time in California can be divided between before and after they entered my life. Thank you for the rides, food, priceless support, and everything in between over the last few years.

I am eternally grateful to my family for their endless support. Most of my childhood memories include my aunts taking me to preschool on a bike and helping me with homework. While only a few graduated from high school, they always understood that education was synonymous with transformative social justice. My 'mom' (and grandmother) Joana Nascimento (†2022) deserves a special mention. I have drawn so many times on memories of her strength and dedication to completing this work, including when she took me to her nurse technician evening course when I was seven, after an all-day shift at a hospital, and encouraged me to practice writing by copying the blackboard notes.

Finally, I would not have survived graduate school (and life in the US!) if it had not been for *jirembirikohẽa* Priscila. Thank you for holding up during my absences for months in the field, for always listening with interest to my excitement at new ideas, and for being the extrovert personality in our duo.

Chapter 1

Introduction

1.1 Overview

This dissertation advances the documentation and analysis of Kawahíva, a critically endangered Tupí-Guaraní Indigenous language spoken by around 560 people from eight Indigenous communities in the Brazilian Amazon basin. The Kawahíva are losing their culture and language, primarily because of social injustice and external influences, including systemic discrimination and prejudice, that have undermined the perceived value of the language (Dos Santos accepted). As a result, children have ceased acquiring the Indigenous language in almost all Kawahíva communities. Despite the existence of previous work on Kawahíva, which includes a grammar sketch (Pease 1968) and dictionary (Betts 2012),[1] these resources were not effective in supporting the communities' revitalization efforts.

A crucial ingredient in fighting language loss is the detailed documentation and analysis of the grammar of endangered languages, which crucially informs revitalization projects. This task is urgent in the Kawahíva case – if current trends continue, the present time represents the last opportunity to document the language fully. Chapter 3 of the dissertation addresses this issue by providing careful documentation of Kawahíva in the form of a sketch grammar. This resource can be used to develop educational materials for Kawahíva schools, such as lesson plans and teaching materials that cover the Kawahíva sound system and grammatical concepts.

Additionally, this dissertation provides the first in-depth description and analysis of Kawahíva syntax by investigating the topics of clause structure and relativization. In particular, I provide the first description and analysis of verb-initial clauses and relativization in Kawahíva. This work draws on high-quality data derived from the description and documentation of Kawahíva grammar as part of the Kawahíva Language Documentation Project. Additionally, the work presented here also strives to bring this documentary data into dialogue with modern syntactic theory, which not only enriches the depth of the descriptive and documentary work on Kawahíva but also fosters meaningful dialogue between documentation and theory.

The topics in syntax explored in this dissertation, clause structure and relativization, while

[1] See a list of previous documentation on Kawahíva in §2.2.

understudied in Kawahíva, are relatively important topics in the literature on Tupí-Guaraní, as several languages of the family have been the focus of descriptive studies for the past 50 years, which addressed these topics. With regard to word order in matrix clauses, the picture that emerges from the previous descriptive scholarship on Tupí-Guaraní can be summarized in (1) (Jensen 1998; Rodrigues and Cabral 2012a):

(1) *Word order in matrix clauses of Tupí-Guaraní languages*:
Most Tupí-Guaraní languages exhibit verb-final word order in matrix clauses.

While this suggests word order in matrix clauses is quite homogeneous throughout the family, not all languages conform to the above generalization. In at least one language, Tenetehára/Guajajára, matrix clauses exhibit a verb-first pattern (Harrison 1986; Duarte 2012), as defined in (2).

(2) *The verb-first pattern (V1 pattern)*:
A language has a verb-first pattern when the verb obligatorily precedes its arguments in declarative sentences with neutral information structure.

The documentation of the Kawahíva clause structure also reveals that the language exhibits the verb-first pattern in (2), since in informational neutral contexts (e.g., out-of-the-blue contexts), main clauses exhibit a strict VSO word order. I provide two examples of this pattern in (3).

(3) Verb-initial word order (VSO) in Kawahíva
a. A-hepiag-ipe ki ji gã ko.
1.SG.A-see-ALREADY PST 1.SG 3.PL REAL
'I already saw them the other day' (Juma: Text)
b. Ere-imbory-vepia-piang nde akaritajavuhua.
2.SG.A-yellow-RED-see 2.SG headdress
'You saw several yellow headdresses' (Juma: Text from a song)

I connect this verb-initial pattern (often abbreviated as V1) to the wider theoretical cross-linguistic literature, where V1 has been a long-standing topic of discussion in theoretical linguistics (see Carnie and Guilfoyle (2000); Carnie et al. (2005)). However, while it is estimated that 10% of the world's languages show this pattern, only a few language families have been considered in addressing how this word order is derived, mostly prominently Austronesian languages (Clemens and Polinsky 2017), Celtic languages (McCloskey 2006, 2017; Borsley and Kathol 2000; Jouitteau 2005), Mayan languages (Clemens and Coon 2018), and Berber languages (Ouhalla 1994). South American languages, which probably represent the greatest number of unrelated languages with the verb-initial pattern, have received little attention in this theoretical domain, with the exception of Tenetehára/Guajajára (Duarte 2012). This scenario is surely at odds with the assumption that answers to the question of how V1 is derived in natural language presuppose detailed formal syntactic accounts of clause structures across a typologically diverse set of languages.

Some fundamental questions about the pattern of verb-initial clauses include the following:

(4) a. What are the syntactic processes that underlie verb-initial orders?
b. Are those processes uniform across languages?

For some time now, the answer to the question in (4b) has been known to be negative. In current theoretical linguistics, it is understood that natural languages display V1 orders that are the product of one of two distinct processes, Head Movement ("Amalgamation") and Remnant VP Movement. In some languages, verb-initial order is due to head movement, wherein the verb locally head moves to some position above the subject (Eberhardt 1999; Macaulay 2005; Clemens and Coon 2018; Bossi and Diercks 2019; Ostrove 2020). However, other verb-initial languages require a different account, in which the verb phrase (VP) undergoes phrasal movement to some position above the subject, what is called Remnant VP movement (Massam 2001; Pearson 2001; Lee 2006; Medeiros 2013; Collins 2017; Adler et al. 2018; Van Urk 2022; Yuan 2022).

Kawahíva provides unique insights into the theories of word order variation. I demonstrate that the existing accounts of the V1 order cannot capture the Kawahíva V1 facts. Kawahíva word order is due to a previously unattested mechanism for deriving V1 order, namely Long Head movement. In this process, the verb undergoes nonlocal syntactic movement (Harizanov and Gribanova 2019).

The results of the documentation of Kawahíva clause structure are significant – they address the underrepresentation of South American languages in the theory of verb-initial word order formation. This underrepresentation perpetuates a skewed perspective within the discipline, often overlooking the Indigenous knowledge of this region. Additionally, this research fills in a real gap in the available information about Amazonian linguistic diversity (Epps and Michael 2023), whose overall diversity is only matched by New Guinea (Rodrigues 2000), and expands the typology of mechanisms involved in verb-initial clause formation in natural languages, which should include long head movement.

This dissertation also addresses relativization in Kawahíva for the first time. Along with verbal agreement, this is perhaps one of the most studied phenomena of the grammar of Tupí-Guaraní languages. Like many other languages, particularly languages in the Americas and Asia, Tupí-Guaraní languages generally employ a strategy for relativization that consists of using nominalizations adnominally, sometimes referred to as the 'nominalization strategy' (Thornes 2023:341). In other words, relativization in these languages is achieved by nominalization morphology itself, without any additional syntactic processes taking place (Comrie and Thompson 1985; Keenan 1985; Andrews 2007; Shibatani 2009). The pattern of relativization in Tupí-Guaraní and the 'nominalization strategy' are defined as in (5).

(5) a. *Relativization in Tupí-Guaraní languages*:
Tupí-Guaraní languages exhibit the 'nominalization strategy' for relativization, schematized in (5b), where a nominalization functions as a modifier to a noun.
b. *The nominalization strategy for relativization*:
Noun [... nominalization ...]

I provide an example of the Kawahíva adnominal nominalization in (6), which is enclosed within brackets, and an underscore indicates a gap position in the nominalization. Building on

previous work on adnominal nominalizations in sister languages to Kawahíva, I demonstrate that analogous constructions in Kawahíva are also nominalizations. In particular, I show that they exhibit the same nominal morphosyntax well-known from parallel structures in sister languages. For instance, they exhibit the nominal temporal marker *-ver*.

(6) **Tapy'ynha** [__ mõhãnga mõmbu-har-aver=a]=hẽa
non.indigenous medicine give-TR.SUBJ-PST.NOM=NMLZ=3.SG.FEM
'The female non-indigenous people that gave out medicine.' (Juma: Elicit)

In Chapter 5, I argue that simply analyzing these structures as nominalizations is insufficient. There is clear evidence that Kawahíva adnominal nominalizations also involve a distinct operation, namely relativization or 'extraction'. I demonstrate that the dependency between the gap and the modified noun in (6) displays the cross-linguistic hallmarks of 'extraction' (Chomsky 1977; Richards 2014).[2] For example, these dependencies are sensitive to island effects. Furthermore, the dependency between the gap in the nominalization and the filler (i.e., the modified noun) can occur at a long distance. These facts strongly suggest that the bracketed structure in (6) involves relativization, and that relativization is not simply a byproduct of an adnominal nominalization. I propose that Kawahíva adnominal nominalizations are best described as nominalized relative clauses.

The primary empirical focus of study in this dissertation is the syntax of Kawahíva, but this work is just one facet of a larger project to document and revitalize Kawahíva language and culture. More generally, this dissertation also aims to provide a broader description of the language, the context of its documentation, and the resources that we have created to support language revitalization. In the remainder of this chapter, I provide a brief overview of the Kawahíva language, followed by a more detailed outline of the structure of the dissertation.

1.2 The Kawahíva language

Kawahíva, which means 'Indigenous person', is an Amazonian language of the Tupí-Guaraní branch of the Tupían family from South America. It is spoken by eight ethnically distinct Indigenous communities: the Amondawa, Jiahui, Juma, Jupaú (Uru-Eu-Wau-Wau), Karipuna, Parintintin, Piripkura, and Tenharin (Pyri). Kawahíva is currently highly endangered and understudied, with around 560 speakers, mostly elders, and middle-aged adults, in a population of 1072-1376 individuals. Unfortunately, Kawahíva is the primary language of communication in only one out of eight communities. In all other communities, the Kawahíva are either passive bilinguals, with Portuguese as their primary language, or monolingual in Portuguese.

Previous documentation of the Kawahíva languages is limited, preliminary, and generally quite dated; a list of all known work that includes documentation of the Kawahíva language (and culture) is summarized in Chapter 2. The most extensive works are the Parintintin dictionary

[2] Alternative terms for extraction are 'A-bar' dependencies (Chomsky 1977; Richards 2014), 'filler-gap' dependencies, and 'unbounded' dependencies (Kroeger 2007).

(Betts 1981) and the pan-dialectal Kawahíva dictionary (Betts 2012). The latter was based on extensive fieldwork with the Parintintin and Tenharin and shorter research trips lasting two weeks each with the Amondawa, Jupaú, and Karipuna. The most extensive grammatical work is a Parintintin sketch grammar, which is 78 pages long (Pease 1968), which has not been updated since. Documentation of Kawahíva discourse only exists in the form of transcribed and translated texts, which, in a few cases, are also provided with glosses. However, this previous documentation did not produce any documentation of natural discourse in the form of accessible recordings, making the recordings on which this dissertation draws the only recordings of any Kawahíva variety available.

The material presented in this dissertation is the result of a seven-year collaborative project to document and revitalize Kawahíva. The methodology used to collect the data conjoins data from elicitation sessions and recorded and annotated oral texts. The elicitation data comes from three main consultants, Mandeí Juma, Tangãi Uru Eu Wau Wau, and Awip (Davi) Uru Eu Wau Wau. Additional examples are drawn from recorded and annotated spontaneous speech texts with several speakers. All data are deposited and acknowledged in the Kawahíva collection at the California Language Archive at UC Berkeley.[3]

1.3 Outline of the dissertation

This dissertation has two main parts. In the first part, which includes Chapters 2 and 3, I provide a descriptive overview of the Kawahíva language and its context. In the second part, Chapters 4 and 5, I investigate the structure of Kawahíva verb-initial clauses and its relativization strategy, respectively.

Chapter 2, *Kawahíva: a Tupí-Guaraní language from the Brazilian Amazon*, provides some information on the geographic and linguistic context of Kawahíva, the linguistic classification, ecology, and vitality. After that, I outline previous resources on the language, primarily created by the missionaries of the Summer Institute of Linguistics (SIL) and non-missionary scholars. Finally, I discuss the outcomes of the Kawahíva Language Documentation Project, created to aid in Kawahíva linguistic and cultural revitalization.

Chapter 3, *A Kawahíva Grammar Sketch*, provides a general overview of Kawahíva grammar, focusing on Kawahíva morphology and syntax. Most of these aspects of Kawahíva grammar have not been documented elsewhere and are only briefly mentioned in previous works. While this overview sets the scene for the rest of the dissertation, subsequent chapters are also intended to stand alone, with relevant background information provided in each.

Chapter 4, *Verb Initiality*, delves into the under-studied phenomenon of verb-initial clauses in Kawahíva based on careful elicitation and natural speech data. This phenomenon, present in 10% of all languages worldwide, has only been explored in one other language within the family (Harrison 1986; Duarte 2012). I argue the most explanatory account of verb-initial clauses requires long head movement (Harizanov and Gribanova 2019). This idea has not been popularly used as a general means of deriving V1 order as a general principle of clausal organization. However, if

[3] http://dx.doi.org/doi:10.7297/X2P26W9H

the theory predicts long head movement is available, we would expect to find more languages where it is used to create the V1 order; I propose Kawahíva is one such language.

Finally, Chapter 5, *Nominalization and relativization in Kawahíva*, investigates the Kawahíva strategy for relativization. We will see that Kawahíva adnominal nominalizations also display the syntax of relativization. This result contrasts with a sizable body of work on sister languages, where nominalization and relativization are understood as indistinguishable, with relativization as the byproduct of nominalization. I conclude that the best analysis of Kawahíva adnominal nominalizations is they are nominalized relative clauses.

Chapter 2

Kawahíva, an Amazonian Tupí-Guaraní language

The primary empirical data discussed in this dissertation are drawn from seven years of collaboration with speakers of Kawahíva, a Tupí-Guaraní language spoken in the Brazilian Amazon Basin. The sketch grammar and linguistic analyses presented here constitute just one facet of the Kawahíva Language Documentation Project, a joint effort between myself and community members to preserve the Kawahíva language and culture. This chapter aims to present Kawahíva in its linguistic and geographic context and describe the collaborative language documentation and revitalization work we have undertaken. In addition to creating new documentary materials, we have aimed to make existing documentary materials accessible to community members, develop community programming to support language vitality, and create digital and print resources for language and cultural education in the Kawahíva community.

The major challenges to language maintenance in the Kawahíva community stem from education — or, more accurately, the lack of educational instruction in the language — and the rapidly changing social context, including the introduction of alcohol in the villages. I return to these issues later in this chapter. Our work seeks to address these challenges by creating resources that meet the needs of Kawahíva community members and expanding the use of the language to new domain contexts, such as cell phones.

This chapter is laid out as follows. In §2.1, I provide information on the Kawahíva people and language, including the geographic and linguistic context, as well as linguistic classification, language use, and vitality and linguistic context of Kawahíva. In §2.2, I outline previous resources on the Kawahíva language and culture, mostly created by the Canadian missionaries Helen Pease and LaVera Betts. Finally, §2.3 provides a brief history of the Kawahíva Language Documentation Project, including a description of current and future projects that aid in Kawahíva linguistic and cultural revitalization.

2.1 Kawahíva: people and language

2.1.1 Kawahíva people

The Kawahíva consist of several Amazonian Indigenous communities living in the Southwest of the Amazon basin, whose lands are in the present-day Brazilian states of Amazonas, Mato Grosso, and Rondônia. These lands are located in the basin of the Madeira River in the Amazon rain forest, one of the largest rivers in South America and the biggest tributary of the Amazon River; the latter is the second longest river in the world and the largest by discharge. A map of the Madeira River watershed is provided in (2.1).

Figure 2.1: Map of the Madeira River. Picture taken from Wikipedia.

There are currently eight ethnic Kawahíva communities in contact with the non-indigenous society. We will also see that there are two uncontacted communities whose population is unknown. The names of the eight communities in contact are provided in Table (2.1).[1], along with

[1]Note from Table (2.1) the community name Tenharin is spelled out with a final <n>, but the land exhibits a final

the state where the community lives and the name of their land. The approximate locations of the Kawahíva are depicted in Figure (2.2). Some of the Kawahíva lands are in the top 50 biggest Indigenous reserves in Brazil, including the Terra Indígena Uru-Eu-Wau-Wau (about 1.9 million hectares) and Terra Indígena Tenharim-Marmelos (about 500.000 hectares).[2] For comparison, the former is almost the size of Slovenia (about 2 million hectares).

NAME	STATE	INDIGENOUS LAND
Amondawa	Rondônia	T.I Uru-Eu-Wau-Wau
Karipuna	Rondônia	T.I Karipuna
Jiahui	Amazonas	T.I Diahui
Juma	Amazonas	T.I Juma
Uru-Eu-Wau-Wau (‘Jupaú’)	Rondônia	T.I Uru-Eu-Wau-Wau
Parintintin	Amazonas	T.I Nove de Janeiro
		T.I Ipixuna
Piripkura	Mato Grosso and Rondônia	T.I Piripkura
		T.I Karipuna
Tenharin (‘Pyri’)	Amazonas	T.I Tenharim-Marmelos
		T.I Tenharim of Igarapé Preto
		T.I Sepoti

Table 2.1: Names, states, and lands of the Kawahíva communities. “T.I” stands for “Terra Indígena” (Indigenous Land).

<m>. The first choice is based on the current community practice of using <n> in their names on social media. This choice was further confirmed through direct communication with the chief of Tenharin Marmelos village, João Sena Tenharin, during an exchange in late July 2023 in Humaitá, Brazil.

[2] A ranking of the officially demarcated Indigenous reserves in Brazil by size is available online: https://terrasindigenas.org.br/en.

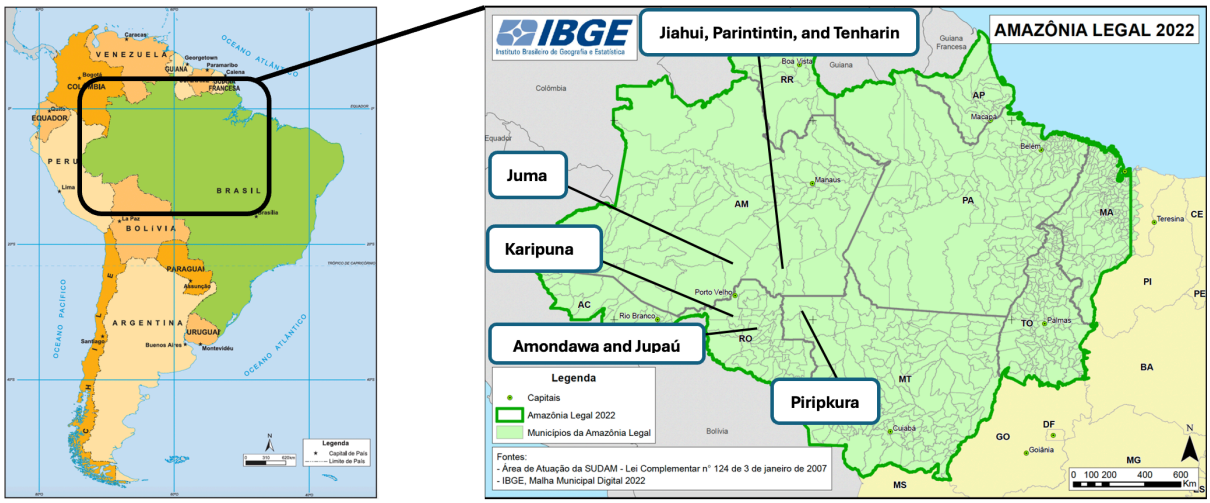


Figure 2.2: The approximate locations of the Kawahíva communities. Maps adapted and permitted for use from the "Map of the Legal Amazon" (Right) and "Map of South America" (Left) by the Brazilian Institute of Geography and Statistics (IBGE). The two-letter abbreviations of the Brazilian states where the Kawahíva are Amazonas (AM), Mato Grosso (MT), and Rondônia (RO).

In most cases among the eight ethnic communities, the community is concentrated in one or two villages, as shown in Table (2.2), which also displays the names of the villages. Except for the names of the villages of the Jiahui and Tenharin, the source for these names comes from personal knowledge (i.e., visiting the village or meeting a resident from that village in a nearby town). Alberto Tenharin (p.c., April 2024) provided the names of the Jiahui and Tenharin villages. A picture from a drone of the Juma village is provided in (2.3).

COMMUNITY	NO. OF VILLAGES	VILLAGE NAME
Amondawa	1	*Trincheira*
Karipuna	1	*Panorama* or *Karipuna*
Jiahui	2	*Ju'í* and *Kwaiari*
Juma	1	*Juma*
Uru-Eu-Wau-Wau ('Jupaú')	6	*Jaikara* (or 621), *623*, *Jamari*, *Alto Jamari*, *Aldeia Nova*, and *Alto Jaru*
Parintintin	3	*Poção*, *Pupunha*, and *Traíra*
Piripkura	2	*Panorama or Karipuna*
Tenharin ('Pyri')	11	Village name in Mato Grosso unknown *Marmelos*, *Vila Nova*, *Bela Vista*, *Tracoá*, *Campinho*, *Taboca*, *Mafuí*, *Castanheira* *Jacuí*, *Pakyri*, and *Caranaí*

Table 2.2: Number and names of Kawahíva villages.

The Kawahíva population totals between 1070 and 1380 individuals, depending on the source. The numerical breakdown by community is provided in Table (2.3).

NAME	POPULATION	SOURCE
Amondawa	130	Dos Santos (2017)
Jiahui	50	Moore et al. (2008)
Juma	12	Fieldwork
Uru-Eu-Wau-Wau ('Jupaú')	107	Dos Santos (2017)
Karipuna	29	Dos Santos (2017)
Parintintin	156	Moore et al. (2008)
	460	Thiago Parintintin (p.c)
Piripkura	3	Dos Santos (2017)
Tenharin ('Pyri')	585	Moore et al. (2008)
Total	1072-1376	

Table 2.3: Kawahíva population.

These numbers do not include the uncontacted Kawahíva people. To date, there are at least two Kawahíva communities that have had no contact with non-indigenous populations. One of these communities resides in the Terra Indígena Uru-Eu-Wau-Wau, which is also home to the Amondawa and Jupaú. Both of these groups refer to the neighboring isolated community as *Jurure'í*. Another group lives in the state of Mato Grosso, and they are referred to as the Kawahíva of the Pardo River, a designation based on their officially demarcated land (Azanha 2006:16). Additionally, in Mato Grosso, two of the three members of the remaining Piripkura community live in intermittent isolation in the Terra Indígena Piripkura.[3] They approach, at times, an official government post on the reserve to ask for fire to keep their torch lit.[4]

Historically, the Kawahíva derive from a common ancestor that once inhabited an area now corresponding to the state of Mato Grosso in Central Brazil. Both traditional narratives and ethnohistorical research suggest a path of migration from east to southwest along the Tapajós River, a tributary of the Amazonas River. This ancestral group subsequently split, likely due to internal fights and external conflicts with the Mundurukú, a non-Tupí-Guaraní Tupían people (Nimuendajú 1981; Menendéz 1989). Nimuendajú (1981)'s ethnohistorical map suggests that Proto-Kawahíva split into three groups. One group migrated to the region where the Amondawa, Jupaú, and Karipuna now live and may have reached their current location (the state of Rondônia) via tributaries of the Madeira River, including the Jacy-Paraná and Jamari rivers (Leonel 1995:33). Another group migrated to the region where the Parintintin and Tenharin currently inhabit, near the Marmelos River, in the current state of Amazonas. The third group is the Apiaká - whose language was reported to have one fluent speaker some years ago (Padua 2007:7). This group settled close to the Upper Tapajós, in what is now the state of Mato Grosso.

[3] The third Piripkura, Rita, is married to a Karipuna and lives among them in Rondônia.

[4] A 2017 documentary depicts these encounters with the two isolated Piripkura, an uncle and his nephew. The documentary is available online in Portuguese: https://www.youtube.com/watch?v=M4adxtjSWx4. The documentary trailer with English subtitles can also be found online: https://www.youtube.com/watch?v=ebWSK3Ptwj8.

Figure 2.3: Juma village. Courtesy photo of Puré Juma.

2.1.2 Kawahíva language

2.1.2.1 Language classification

Kawahíva belongs to the Tupí-Guaraní language family, one of the ten branches of the Tupían family (Rodrigues and Cabral 2002; Michael et al. 2015). The internal classification of the Tupían family is given in (2.4). Kawahíva is the sole Tupí-Guaraní language spoken in the likely homeland of Proto-Tupían, which is the present-day state of Rondônia (Rodrigues and Cabral 2012a:499). Seven out of the ten Tupían branches are exclusively spoken in this area. The region is also the most linguistically diverse state of Brazil, with 26 languages (Galucio et al. 2018). Moreover, Kawahíva is also spoken in two other states, Amazonas and Mato Grosso.

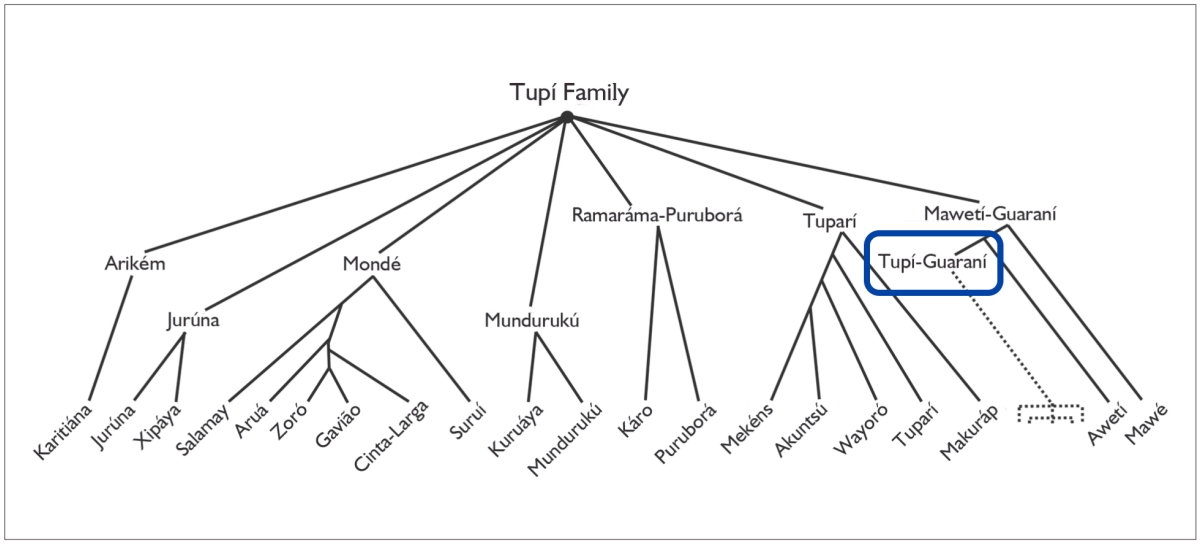


Figure 2.4: The language internal classification of the Tupían family. Adapted from Gijn et al. (2015:300).

Comparative and phylogenetic evidence suggests that Kawahíva falls into a Tupí-Guaraní subgroup which also includes Kawaiwete (also known as Kayabí) (Rodrigues and Cabral 2002; Michael et al. 2015). This subgrouping is depicted in Figure (2.5).

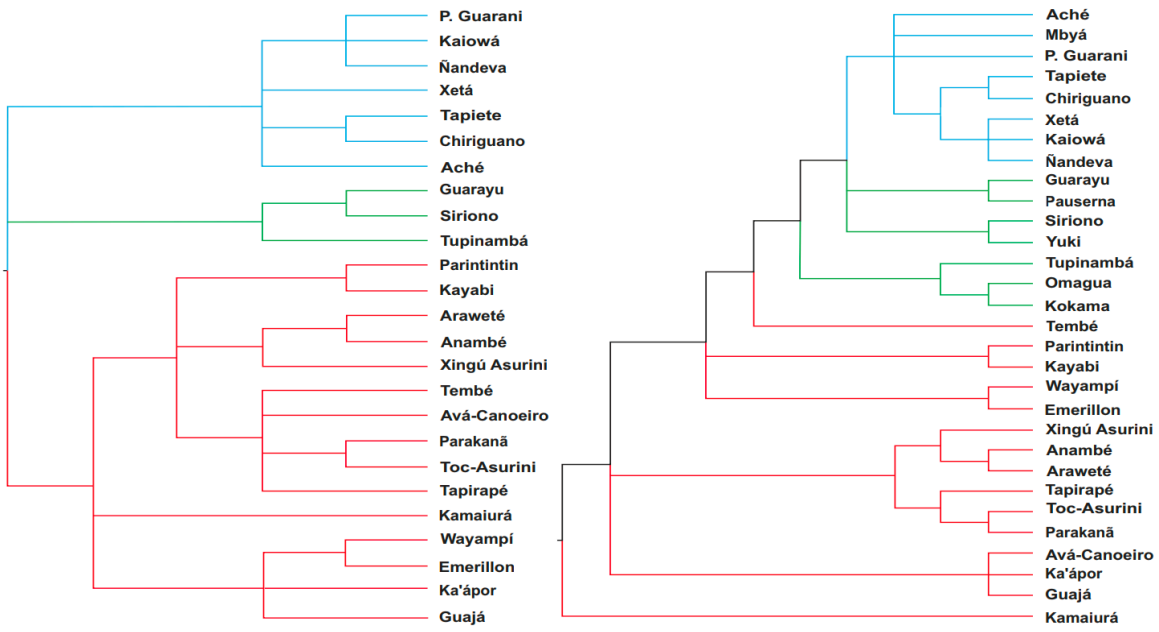


Figure 2.5: Tupí-Guaraní language internal classifications based on Rodrigues and Cabral (2002) (left), and Michael et al. (2015) (right). Image adapted from Michael et al. (2015:208).

There has also been an attempt at a Kawahiva-internal classification. Sampaio (2001) proposes two language-internal classifications based on phonological similarity and lexicostatistics. These classifications differ in some ways. For instance, the classification based on phonological similarity places Karipuna as closer to Jupaú (URU in Figure 2.6) and Amondawa (AMO). However, the lexicostatistical classification indicates that Karipuna is closer to Parintintin (PAH) and Tenharin (TEM). Additionally, the phonological classification suggests that Jiahui (DIA) is the most divergent dialect. Instead, the lexicostatistics classification indicates this is Juma (JUM).

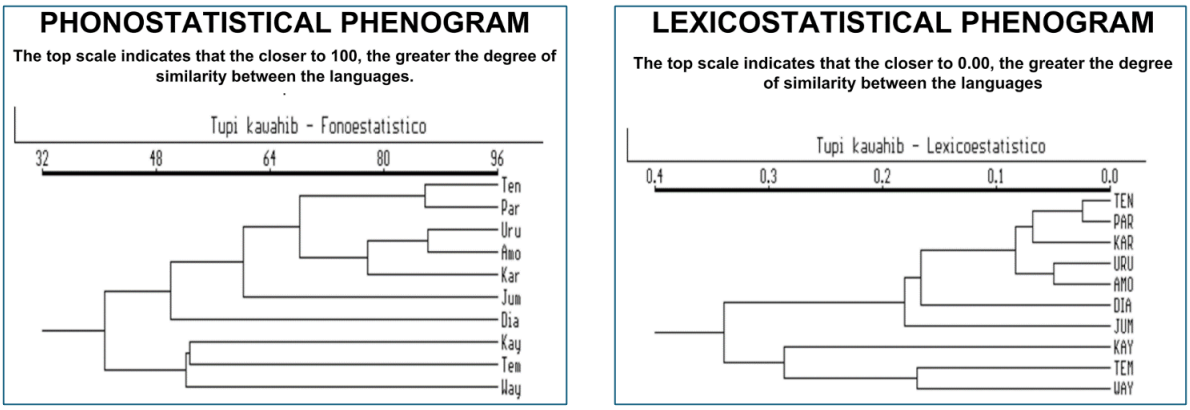


Figure 2.6: Kawahíva language internal classification based on Sampaio (2001:93-94).

Finally, while some previous works have adopted a distinction between northern and southern Kawahíva dialects (Aguilar 2015; Marçoli 2018), following Kracke (2007), this distinction is solely based on the geographical location of the speakers of these languages and is not based on linguistic criteria. It is clear that more systematic documentation across the pan-Kawahíva dialectal complex is needed to provide a better picture of the internal Kawahíva diversity.

2.1.2.2 Language use and vitality

Kawahíva consists of eight varieties associated with ethnic subgroups. Following a common practice among the communities, I refer to the name of a language variety by the name of its speech community. Sadly, all but the variety spoken by the Tenharin is severely endangered (Moore et al. 2008; Dos Santos 2017). The estimated number of speakers for each ethnic variety is displayed in Table (2.5), which also includes Glottolog and ISO codes. The existence of an ISO code for each ethnic variety suggests that they are distinct languages. However, these dialects are highly mutually intelligible. It is also the consensus among scholars who have previously conducted phonological and lexical comparative studies on the Kawahíva varieties that they comprise a single language (Sampaio 2001; Aguilar 2015; Marçoli 2018).

Other previously noted Kawahíva dialects have become dormant. This includes the most recent case of Capivari, whose last known speaker, Mr. Pitanga Capivari, passed away in 2022 in the city of Porto Velho (state of Rondônia) at an advanced age (Hanmin Kim, p.c., May 2023). Other varieties were spoken by groups that once inhabited areas near the Machado River, like the Paranawat and Wiraféd, whose existence is attested to by wordlists collected by Curt Nimuendajú between the years 1940-1950, and those varieties spoken by communities on the tributaries of the same river, closer to the Muqui River, like the Takwatip and the Ipotewát, as documented by Curt Nimuendajú and Lévi-Strauss around the same period. Lévi-Strauss also mentions people who were already almost extinct at that time, who lived near the Machado/Ji-Paraná River, like the Tucumanfét and the Jabotiféd, and the Mialat, who inhabited the Leitão River region.

NAME	POP.	SPEAKERS	PASS. BIL.	GLOTTOLOG	ISO	SOURCE
Amondawa	130	81	10	amun1246	adw	(Dos Santos 2017)
Jiahui	50	1	-	diah1239	pah	(Moore et al. 2008)
Juma	12	3	-	juma1249	jua	Fieldwork
Jupaú	107	64	17	urue1240	urz	(Dos Santos 2017)
Karipuna	29	10	0	kari1317	kuq	(Dos Santos 2017)
Parintintin	156	10	-	pari1258	pah	Moore et al. (2008)
	460	19				T. Parintintin (p.c)
Piripkura	3	3	-	N/A	N/A	Fieldwork
Tenharin	585	350	-	nucl1663	pah	(Moore et al. 2008)
Total	1072-1376	522-531	27			

Table 2.4: Speaker number of Kawahíva and Glottolog and ISO codes. "Pass. Bil." stands for passive bilinguals. "T. Parintintin" is Thiago Parintintin.

The Kawahíva are losing their culture and language. This is primarily due to social injustice, economic marginalization, and external influences. Systemic discrimination and prejudice have undermined the perceived value of the language. Expropriation of traditional lands and resources have deprived the communities of their traditional means of subsistence. Additionally, contact with outsiders has introduced detrimental societal changes, such as alcohol consumption. Particularly concerning is the rapid disappearance of many ethnic dialects within a single generation. Despite Kawahíva being the official language of elementary school instruction, the lack of materials, insufficient teacher training, and inadequate institutional support have hindered its effective usage and, consequently, its learning in schools. Sadly, among all communities but the Tenharin, children have ceased acquiring Kawahíva as their first language. Opportunities for learning the language are limited at the elementary school level, and the transition to non-indigenous educational systems for further schooling – if it occurs – effectively shuts down the chance for most children to continue learning it.

2.2 Previous documentation of Kawahíva

This section gives a brief overview of previous documentation of the Kawahíva culture and language. Earlier documentation primarily focused on word lists and phonological descriptions, though there is a brief grammatical sketch and glossed and translated verbal-art narratives. It also included ethnographic works that included material and verbal culture documentation. To my knowledge, the previous verbal art documentation is not available in the form of accessible recordings. All known work that includes documentation of the Kawahíva culture and language is summarized in tables 2.5-2.7. These materials are mostly inaccessible to community members, either because of inexperience in working with academic sources, lack of access to library resources, or language barriers.

The earliest documentation of a Kawahíva language comes from German-Brazilian ethnologist and anthropologist Curt Nimuendajú's on Parintintin. In the early 1920s, he was assigned

the task of forcing the Parintintin into contact by the *Serviço de Proteção ao Índio* (The Indian Protection Service office). His 78-page ethnographic sketch (Nimuendajú 1924), based on trips in 1921-23 to carry out the assigned task with the Parintintin, gives a short description of several aspects of Parintintin culture, including the language, history, religion, customs, behavior, and physical attributes of the Parintintin people. In the section on Parintintin, "Language", he notes the Parintintin language is "pure Tupían" and gives a wordlist of 328 items that he transliterated (see Figure 2.7), which mostly uses Portuguese orthographic conventions. This vocabulary includes words for body parts, cultural artifacts, kinship, animals, and simple sentences.

Following Nimuendajú's wordlist, most of the language documentation on Kawahíva was produced by Helen Pease and LaVera Betts between the 1960s-1980s, who were missionaries for the Summer Institute of Linguistics (SIL). This documentation consists of dictionaries, wordlists, short grammar sketches of some varieties, transcribed and translated narratives, and a translation of the New Testament. Their works on Parintintin are based on long-term fieldwork trips, while works on the other varieties are based on fieldwork that lasted for two weeks.

The most extensive works by these missionaries are the Parintintin-Portuguese dictionary (Betts 1981) and a pan-dialectal dictionary (Betts 2012), and several transcribed and translated texts (with a few of them glossed), most of them in Tenharin. The pan-dialectal dictionary also included some words from the Amondawa, Karipuna, Jupaú, and Tenharin varieties, although the dictionary is mostly based on Parintintin. Additionally, the 78-page Parintintin grammar sketch by Pease (Pease 1968) is a tagmemic-transformational-relational grammatical description of the major word classes, phrases, and sentence types of Parintintin. Despite being embedded in a particular framework, the grammar sketch is generally readable to non-practitioners of the framework employed. It is also the earliest grammatical description of Kawahíva.

Further documentation of the other aspects of Kawahíva culture was continued by a number of anthropologists. This includes the work by the American anthropologist Waud Kracke, who started doing fieldwork with the Parintintin in the 1960s. Among the numerous outcomes of his work is a PhD dissertation based on fieldwork trips that totaled ten months during 1966-1968 (Kracke 1973), which includes a 47-page ethnography of the Parintintin people, and a book chapter that presents a detailed description of food taboos among the Parintintin, the only documentation on this topic among the Kawahíva (Kracke 1981). Work by the Brazilian anthropologists Edmundo Peggion and Luciana França provided the first systematic ethnography of the kinship system of Kawahíva communities, the former based on fieldwork with the Tenharin and the latter with the Jupaú (Peggion 1996, 2005; França 2012). Additional cultural documentation includes an overview of the history and material culture of the Jupaú by Dresden Klaus-Peter (Klaus-Peter 2005), the edited volume on feathered artifacts by Ana Carla Bruno and Simone Gomes (Bruno and Gomes 2010), the documentation of Piripkura verbal art by João Paulo Denófrio (Denófrio 2012), and the photographic and descriptive documentation of several Amondawa artifacts, by João Paulo Denófrio and Tambura Amondawa (Denófrio 2012).

LANGUAGE	RESOURCE TYPE	TITLE	SOURCE
Parintintin	Article	The Parintintin of the Madeira River	(Nimuendajú 1924)
	Translated narrative	Bahira and his experiences	(Pereira 1940)
	Book chapter	The Cawahib, Parintintin, and their neighbors	(Nimuendajú 1948)
	Word list	Parintintin word list	(Lowe 1960)
	Transcribed and translated narratives	Orekwatijar	(Parintintin 1940)
	Word list	Parintintin word list	(Pease and Betts 1962)
	Transcribed and translated narratives	Moroğita: Parintintin myths	(Parintintin 1966)
	Transcribed, glossed, and translated narratives	Transcription of recordings of Parintintín texts	(Betts 1967)
Juma	Word list	Juma word list	(Abrahamson 1968)
	Grammar sketch	Filler of the verb complex slot	(Abrahamson n.d)
Parintintin	Grammar sketch	Parintintin grammar	(Pease 1968)
	Transcribed and translated narrative	Parintintin discourse	(Betts 2008[1969])
	Transcribed and translated narrative	Parintintin selected texts	(SIL 1972)
	PhD dissertation (w/ an ethnography)	Emotions and personality in Parintintin politics	(Kracke 1973)
	Article	Parintintin time sets, *po*, and *ko*	(Pease 1973)
Tenharin	Transcribed narratives	Plays of Margarida	(SIL 1976a)
	Transcribed, glossed, and translated narratives	Stories of Kairana	(SIL 1976b)
	Transcribed and translated narratives	Stories of Margarida	(SIL 1976c)
Juma and Parintintin	Word list	Juma-Parintintín similarities	(Pease 2009[1977])
	Transcribed and translated narratives	Ore nhomomirika ’ndika’ndirame	(Tenharim 1977a)
	Transcribed and translated narratives	Ahererohokava	(Tenharim 1977b)
	Transcribed and translated narratives	Ore horame cidade pe	(Tenharim 1977c)
	Transcribed and translated narratives	5° BEC	(Tenharim 1977d)

Table 2.5: Previous documentation of the Kawahíva language and culture.

LANGUAGE	RESOURCE TYPE	TITLE	SOURCE
Tenharin	Short anatomic descriptions	Ahera'oa	(Tenharim 1977e)
	Shorts texts about "malaria" and "cold" and their treatment	Ahe rehe i'ara ahe'mbyva ipohaĝa pavẽi	(Tenharim 1977f)
	Transcribed, glossed, and translated texts	A story of Manuel João	(SIL 1977)
	Transcribed and translated texts	Mbatera ga imombe'ui ji ve	(Tenharim 1977g)
	Transcribed and translated texts	The jaguar	(Tenharim 1977h)
	Transcribed and translated texts	Ore onga hepiapavi	(Tenharim 1977i)
	Transcribed and translated texts	Yvakytihava olaria	(Tenharim 1977j)
	Transcribed and translated texts	Dalvahẽa okwatijarypy	(Tenharim 1978a)
	Transcribed and translated texts	Gilenahẽa okwatija	(Tenharim 1978b)
	Transcribed and translated texts	Lurdehẽa okwatijara'ja	(Tenharim 1978c)
	Transcribed and translated texts	Lurdehẽa okwatijarypy	(Tenharim 1978d)
	Transcribed and translated texts	Mariahẽa okwatija	(Tenharim 1978e)
	Transcribed and translated texts	Porciano'ga okwatija	(Tenharim 1978f)
	Transcribed and translated texts	Zelito'ga okwatijarypy	(Tenharim 1978g)
	Transcribed and translated texts	Zelito'ga okwatijara'ja	(Tenharim 1978h)
Juma and Karipuna	Word list	Karipuna and Jumá comparative study	(Abrahamson 1980)
Parintintin	Book chapter	Don't let the piranha bite your liver	(Kracke 1981)
	Dictionary	Parintintin-Portuguese Portuguese-Parintintin dictionary	(Betts 1981)
Amondawa	Word list	Amondawa word list	(Pease and Betts 1991b)

Table 2.6: Previous documentation of the Kawahíva language and culture.

LANGUAGE	RESOURCE TYPE	TITLE	SOURCE
Amondawa	Transcribed and translated texts	Texts of an Amondawa	(Pease and Betts 1991d)
Jupaú	Word list Grammar sketch	Jupaú word list Comments on Uru-Eu-Wau-Wau	(Pease and Betts 1991c) (Pease and Betts 1991a)
Tenharin	Ethnography Bible	The Tenharim kinship system Tupana'ga nhi'iğa	(Peggion 1996) (Wycliffe 1996)
Amondawa	Narratives in Portuguese	Amondawa myths	(Silva et al. 2004)
Jupaú	Ethnography	Contributions to the history and culture of the Urueuwauwau	(Klaus-Peter 2005)
Tenharin	Ethnography	The dualist organization of the Kagwahiva people from the Amazon	(Peggion 2005)
Amondawa	Book chapter	The noun class: the Amondáwa onomastics	(Peggion 2007)
Tenharin, Jiahui, Parintintin	Ethnographic book	The feathered art of the Kagwahiva on the Transamazônica	(Bruno and Gomes 2010)
Jupaú	Ethnography	Crossed paths: kinship, difference, & movement among the Kagwahiva	(França 2012)
Amondawa, Karipuna, Jupaú, Parintintin, Tenharin	Dictionary	Kagwahiva dictionary	(Betts 2012)
Piripkura	Ethnography	Brief ethnographic contribution about the Kagwahiva: the Piripkura	(Denófrio 2012)
Amondawa	Ethnography	Object and person: Amondawa ethnographic fragments	(Amondawa and Denófrio 2015)

Table 2.7: Previous documentation of the Kawahíva language and culture.

1. **Vocabulario Parintintin.**

Levantado no Posto de Pacificação do Maicy mirim em dezembro de 1922 e janeiro de 1923.

Cabeça	*ae-akáñ*	Sovaco	*ae-pepowíd*
Olho	*ae-reakwád*	Mão	*ae-po*
Orelha	*ae-nambí*	Dedo da mão	*ae-pũ*
Furo da orelha	*nambikwád*	Unha da mão	*ae-pũpã*
Bocca	*a-yurú*	Coxa	*ae-ub*
Labio	*ae-rembé*	Pé	*ae-pị*
Nariz	*ae-tĩ*	Pelle	*ae-pid*
Lingua	*ae-ku*	Osso	*ae-kañ*
Dente	*ae-ráĩ*	minha veia	*ye-rayíg*
Cabello	*ae-ab*	Ferida da pelle	*mbirurú*
o cabello delle	*i-ab*	Dôr	*habị*
Sobrancelha	*ae-rapịkã*	Espectro?	*anyán*
Barba	*ae-rendịvá*	Mentira	*yapú* (?)
Cabello do corpo	*ae-rab*	Cubitus	*menó*
Pescoço	*ae-ayúd*	Tatuagem	*kwatihád*
Peito	*ae-potíd*	no braço sup.	*pirakáñga-rañgáb*
Bico do peito	*ae-kam*	no braço inf.	*yawára-rañgáb*
Barriga	*ae-rihé*	vossa imagem	*pe-nañgáb*
teu umbigo	*ne-tuá*	Sol	*kwará*
Nadegas	*ae-revĩ*	luz do sol	*kwarendị* (?)
Anus (delle)	*hevikwád*	calor do sol	*kwarahịahị* (?)
Ventosidade	*pinombó*	Lua	*yahị*
Membro delle	*hakwáĩ*	lua nova	*yahị ipị*

Figure 2.7: The earliest Parintintin wordlist, compiled by Nimuendajú (1924:262-66). The section title reads, "1. Parintintin vocabulary". The text below reads, "(The vocabulary was) collected at the Pacification post on the Maicy Mirim (River) in December 1922 and January 1923."

2.3 The Kawahíva Language Documentation Project

This section describes the present contributions to Kawahíva documentation and revitalization by the Kawahíva Language Documentation Project (which has provided the foundation for the research presented in this dissertation) and future endeavors.

My work with the Kawahíva community began in 2017 as part of a pilot language survey, the *Inventário Nacional de Diversidade Linguística* (INDL) (National Survey of Language Diversity).

The INDL survey had the following major goals (Galucio et al. 2018:218):

- Survey the number of speakers of Indigenous languages spoken in the state of Rondônia, the most linguistically diverse state of Brazil, with 26 languages (Galucio et al. 2018), to better inform institutional language policy and planning projects.
- Diagnose language vitality.
- Identify the number of people literate in each Indigenous language and the effectiveness of the writing systems available.
- Indicate the level of maintenance of verbal culture and the obstacles to its preservation.

In this capacity, I connected with the Amondawa, Jupaú (Uru-Eu-Wau-Wau), and Karipuna communities. The project was headed by linguists associated with the *Museu Paraense Emílio Goeldi*, a Brazilian Museum, and one of the most renowned institutions in South America for research on the Amazon. The principal investigators included Ana Vilacy Galucio, Dennis (Denny) Moore, and Hein van der Voort. The INDL survey allowed members of the survey team to conduct fieldwork in Rondônia. I was able to conduct a total of four months of fieldwork with the abovementioned communities during the months of February, June-July, and October 2017.

The INDL survey also aimed to document the demand for language documentation work among the communities surveyed. According to the survey results (Dos Santos 2017, 2019), community stakeholders ranked audiovisual language documentation, support for language learning and teaching materials, orthography reform, and training of community members in basic language documentation methods as the most urgent demands for language work.

Following the language survey in 2017, the Jupaú and Karipuna communities and I created the Kawahíva Language Documentation Project, a collaborative effort to address the communities' needs for language work and revitalization, as listed above. In 2019, the Juma community requested the project be extended to include their ethnic variety. Two other communities, the Parintintin and Tenharin, have also requested to be part of the documentation project: the Parintintin in Fall 2022 and the Tenharin in Spring 2024. Due to the usual time constraints imposed by a PhD, documentation of these two dialects had to be postponed. However, starting in Fall 2024, documentation of Parintintin will start as part of a 2-year postdoctoral fellowship funded by the Endangered Language Documentation Project (ELDP).[5]

Here, I describe aspects of the Kawahíva Language Documentation Project, including linguistic and cultural documentation, training of community members in language documentation, the creation of digital and printed resources, such as the Kawahíva multimedia dictionary for cellphones, a bilingual storybook, a letter tracing workbook, the organization of community workshops and panels.

[5]The collection's landing page of this project is available online: https://www.elararchive.org/dk0805.

2.3.1 Language documentation

The original goal of the project, and the foundation for this dissertation, is the documentation of the Kawahíva language, including the lexicon, grammar, and discourse. All documentary materials are archived in the open-access Kawahíva Language Documentation Project collection (2019-06) in the California Language Archive, which is housed at the University of California, Berkeley.[6] This documentation forms the basis for the digital and print resources described in §2.3.2-2.3.3.

Much of the linguistic documentation was created through recording and annotation of verbal art and elicitation interviews, and occasional WhatsApp chats. The primary methods used for data collection in this dissertation are summarized in (1). Throughout this dissertation, I indicate which method was used to collect a data point by writing "Text," "Elicit," or "WhatsApp" next to the translation of the example.

(1) a. Recording, transcription, and translation of Kawahíva spontaneous speech
b. Translation from Portuguese to Kawahiva
c. Well-formedness judgments of Kawahíva utterances
d. Comparison of the well-formedness of two or more Kawahíva utterances
e. Acceptability judgments of Kawahíva utterances in a constructed context
f. Description of visual stimuli in Kawahiva

The recordings made using the method in (1a) form the corpus of high-quality audiovisual, carefully annotated records of Kawahíva verbal art and conversational texts, including personal histories, traditional oral narratives, procedural descriptions, songs, and staged conversations. The annotation tasks in (1a) followed the workflow in (2). The text respeaking step in (2b) is an adaptation of the Basic Oral Language Documentation (BOLD) transcription method outlined in Reiman (2010).

(2) a. Text segmentation: Recorded texts are segmented into sentences and used to create a .eaf file using the function "auto segmenter" of the software program *SayMore*.
b. Text respeaking: Speakers play back each sentence on ELAN and are recorded carefully, respeaking the sentence in Kawahíva and translating it into Portuguese.
c. Text annotation: Speakers assist in transcribing and translating the recorded text on ELAN, or I used recorded, respoken texts in (2b) to annotate recordings.

[6] This collection is available online: http://dx.doi.org/doi:10.7297/X2P26W9H.

Figure 2.8: Oral transcription.

The elicitation interview tasks described in (1b-f) were primarily conducted with bilingual speakers in Kawahíva and Portuguese, although fluency in Portuguese varies; elderly speakers tend to be less proficient in Portuguese, which they started learning during adolescence. For Juma, interviews were conducted with Boreá and Mandeí Juma. For Jupaú, interviews were conducted with the Jupaú speakers Awip (Davi), Mandá, Puré, and Tãngai, both in person (2017-19, 2021-23) and over Zoom (2020).

Figure 2.9: Juma and Jupaú speakers. Clockwise: Mandeí, Boreá, Mandá, Puré, Tangãi, and Awip (Davi).

All interviews were conducted with one speaker at a time. Translation tasks described in (1b) were used to obtain lexical, syntactic, and semantic information about Kawahiva. The tasks in (1c-d) aimed to capture aspects of Kawahíva syntax. To indicate combinations that are not possible in the language, I marked sentences that were judged ill-formed in any context with an asterisk (*). The tasks in (1e-f) aimed to capture aspects of Kawahíva meaning, which places constraints on the acceptability of well-formed utterances in certain situations. Utterances that are well-formed but not felicitous in a particular context are presented with a pound sign (#). During elicitation interviews, utterance contexts were constructed and supplied in Portuguese or supplied visually using picture stimuli, followed by a Kawahíva utterance intended to be spoken in that context.

Figure 2.10: In-person elicitation.

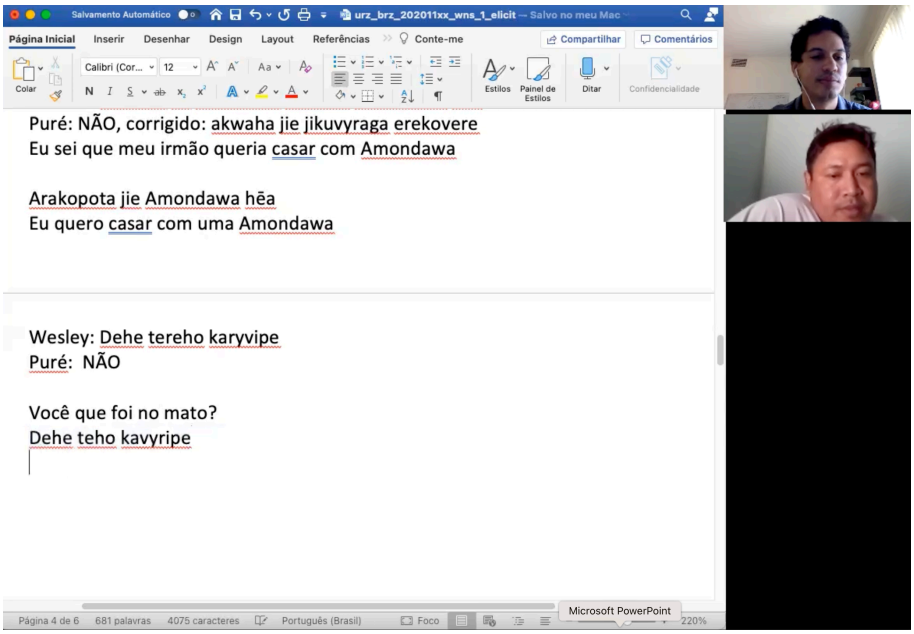


Figure 2.11: Zoom elicitation during the COVID-19 pandemic.

Finally, we also documented cultural knowledge, including traditional skills, through audio-visual recording and photos. For instance, we filmed and photographed the processes of basket weaving, flute, bracelet, and necklace making, arrow and headdress feathering, and honey extraction. These are key knowledge domains for documentation and revitalization, as only a few people among the Jupaú and Juma know them, some of whom are shown in Figure (2.12).

Figure 2.12: Documentation of cultural knowledge.

2.3.2 Digital resources

Digital resources for language documentation offer several advantages: they can be easily updated and changed, they may embed multimedia recordings, such as audio and video, and they can be easily accessed from anywhere in the world.[7] During this documentation project, we created two digital resources: a digital multimedia dictionary app for cell phones and customized keyboards for computers and cell phones. The development of the database that underpins the cell phone app was possible due to funding provided by a partnership between UNESCO and the Museu do Índio (FUNAI).[8,9]

[7]While cell coverage is not available for most Kawahíva, internet coverage has become quite widespread in the villages as the result of the expansion of satellite internet service in the Amazon. By Fall 2023, 90% of all municipalities in the region had at least one customer, according to data from Brazil's National Telecommunications Agency, *Anatel* (BBC 2024).

[8]The Kawahíva dictionary app is available for download through this link from Google Play.

[9]The Kawahíva keyboards are available for download through this link.

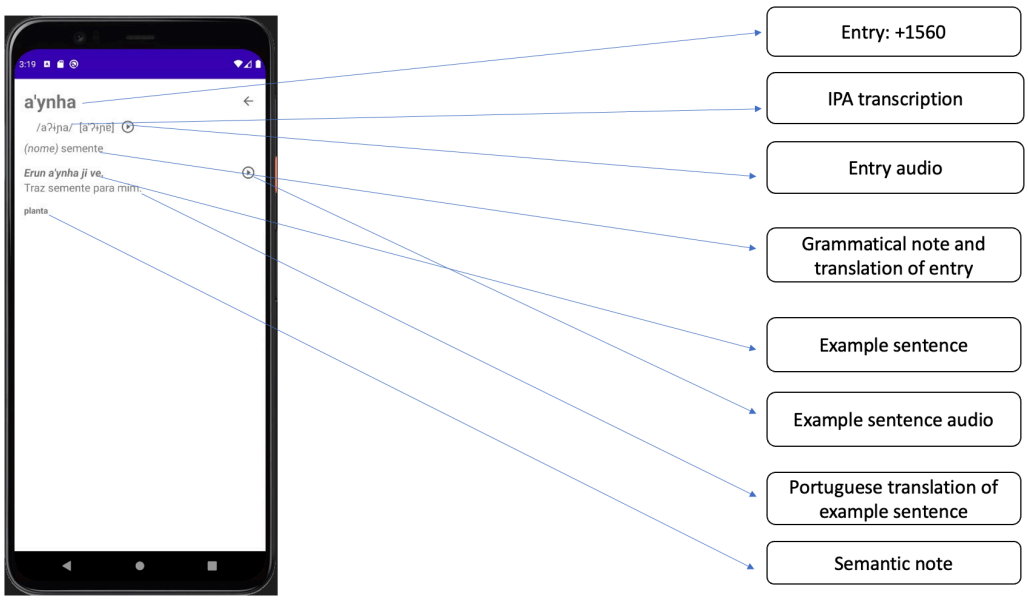


Figure 2.13: The Kawahíva multimedia dictionary app for cell phones.

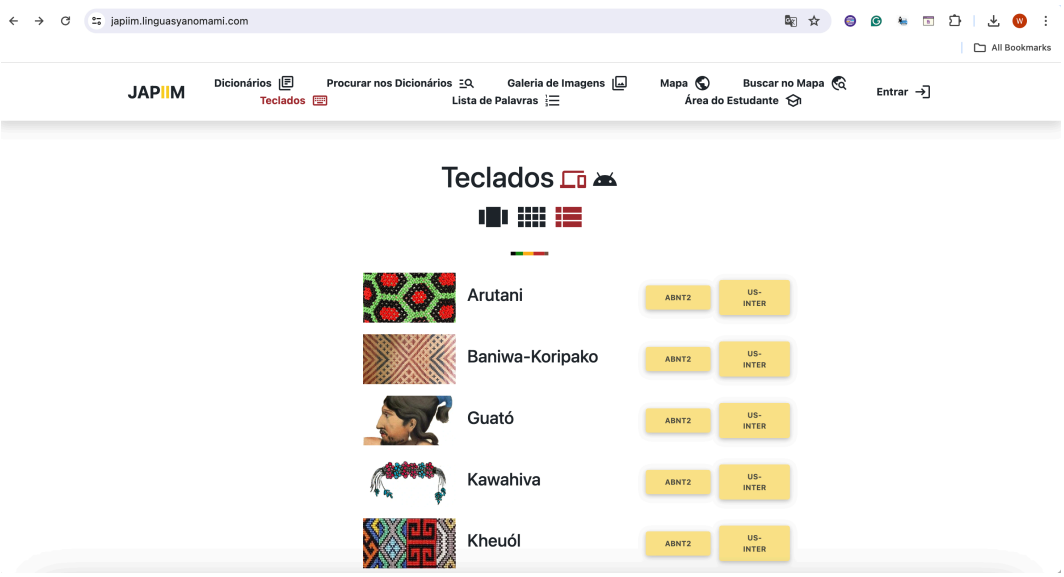
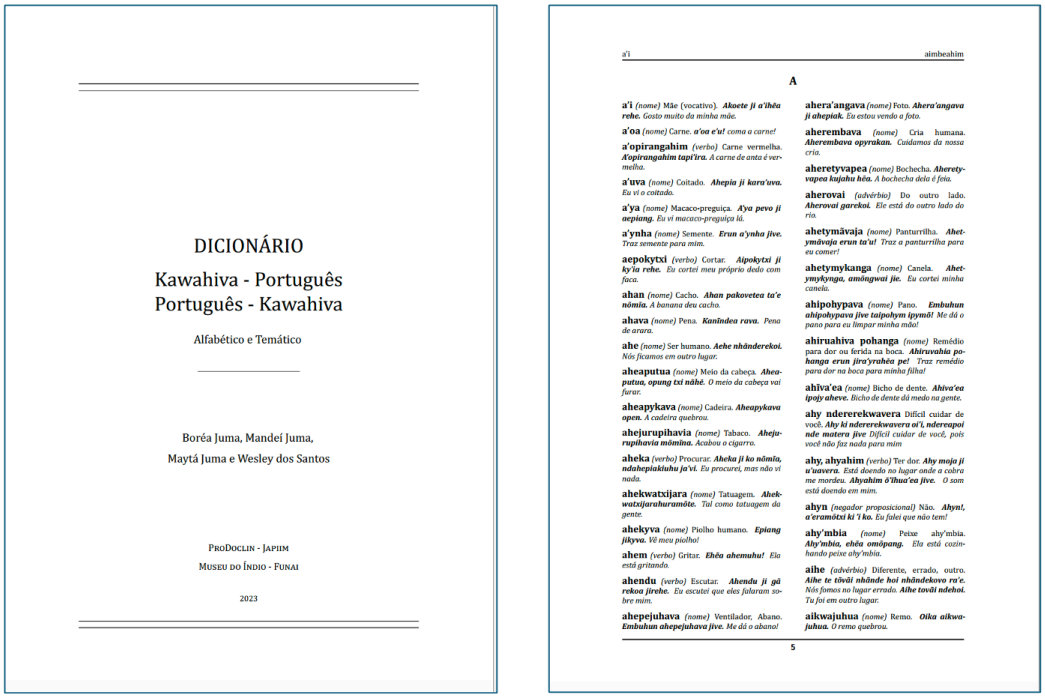


Figure 2.14: Customized keyboards for computers and cell phones for download.

2.3.3 Printed and forthcoming resources

The first printed resource produced was the thematic and encyclopedic Kawahiva-Portuguese, Portuguese-Kawahiva dictionary. This resource is the printed version of the digital app version

described in section (2.3.2). It includes all the lexical entries, example sentences, and grammatical notes found in its digital counterpart. A representative page of the dictionary is provided in (2.15).



DICIONÁRIO

Kawahiva - Português
Português - Kawahiva

Alfabético e Temático

Boréa Juma, Mandeí Juma,
Maytá Juma e Wesley dos Santos

ProDoclin - Japiim
Museu do Índio - Funai

2023

a'i aimbeahim

A

a'i *(nome)* Mãe (vocativo). ***Akoete ji a'ihèa rehe.*** *Gosto muito da minha mãe.*

a'oa *(nome)* Carne. ***a'oa e'u!*** *coma a carne!*

a'opirangahim *(verbo)* Carne vermelha. ***A'opirangahim tapi'ira.*** *A carne de anta é vermelha.*

a'uva *(nome)* Coitado. ***Ahepia ji kara'uva.*** *Eu vi o coitado.*

a'ya *(nome)* Macaco-preguiça. ***A'ya pevo ji aepiang.*** *Eu vi macaco-preguiça lá.*

a'ynha *(nome)* Semente. ***Erun a'ynha jive.*** *Traz semente para mim.*

aepokytxi *(verbo)* Cortar. ***Aipokytxi ji ky'ia rehe.*** *Eu cortei meu próprio dedo com faca.*

ahan *(nome)* Cacho. ***Ahan pakovetea ta'e nómia.*** *A banana deu cacho.*

ahava *(nome)* Pena. ***Kanîndea rava.*** *Pena de arara.*

ahe *(nome)* Ser humano. ***Aehe nhãnderekoi.*** *Nós ficamos em outro lugar.*

aheaputua *(nome)* Meio da cabeça. ***Aheaputua, opung txi nãhê.*** *O meio da cabeça vai furar.*

aheapykava *(nome)* Cadeira. ***Aheapykava open.*** *A cadeira quebrou.*

ahejurupihavia *(nome)* Tabaco. ***Ahejurupihavia mómina.*** *Acabou o cigarro.*

aheka *(verbo)* Procurar. ***Aheka ji ko nõmia, ndahepiakiuhu ja'ví.*** *Eu procurei, mas não vi nada.*

ahekwatxijara *(nome)* Tatuagem. ***Ahekwatxijarahuramõte.*** *Tal como tatuagem da gente.*

ahekyva *(nome)* Piolho humano. ***Epiang jikyva.*** *Vê meu piolho!*

ahem *(verbo)* Gritar. ***Ehêa ahemuhu!*** *Ela está gritando.*

ahendu *(verbo)* Escutar. ***Ahendu ji gã rekoa jirehe.*** *Eu escutei que eles falaram sobre mim.*

ahepejuhava *(nome)* Ventilador, Abano. ***Embuhun ahepejahava jive.*** *Me dá o abano!*

ahera'angava *(nome)* Foto. ***Ahera'angava ji ahepiak.*** *Eu estou vendo a foto.*

aherembava *(nome)* Cria humana. ***Aherembava opyrakan.*** *Cuidamos da nossa cria.*

aheretyvapea *(nome)* Bochecha. ***Aheretyvapea kujahu héa.*** *A bochecha dela é feia.*

aherovai *(advérbio)* Do outro lado. ***Aherovai garekoi.*** *Ele está do outro lado do rio.*

ahetymãvaja *(nome)* Panturrilha. ***Ahetymãvaja erun ta'u!*** *Traz a panturrilha para eu comer!*

ahetymykanga *(nome)* Canela. ***Ahetymykynga, amõngwai jie.*** *Eu cortei minha canela.*

ahipohypava *(nome)* Pano. ***Embuhun ahipohypava jive taipohym ipymã!*** *Me dá o pano para eu limpar minha mão!*

ahiruahiva pohanga *(nome)* Remédio para dor ou ferida na boca. ***Ahiruvahia pohanga erun jira'yruhéa pe!*** *Traz remédio para dor na boca para minha filha!*

ahîva'ea *(nome)* Bicho de dente. ***Ahiva'ea ipojy aheve.*** *Bicho de dente dá medo na gente.*

ahy ndererekwavera Difícil cuidar de você. ***Ahy ki ndererekwavera oi'i, ndereapoi nde matera jive*** *Difícil cuidar de você, pois você não faz nada para mim*

ahy, ahyahim *(verbo)* Ter dor. ***Ahy moja ji u'uavera.*** *Está doendo no lugar onde a cobra me mordeu.* ***Ahyahim ó'ihua'ea jive.*** *O som está doendo em mim.*

ahyn *(negador proposicional)* Não. ***Ahyn!, a'eramõtxi ki 'i ko.*** *Eu falei que não tem!*

ahy'mbia *(nome)* Peixe ahy'mbia. ***Ahy'mbia, ehêa omõpang.*** *Ela está cozinhando peixe ahy'mbia.*

aihe *(advérbio)* Diferente, errado, outro. ***Aihe te tõvãi nhãnde hoi nhãndekovo ra'e.*** *Nós fomos no lugar errado.* ***Aihe tavãi ndehoi.*** *Tu foi em outro lugar.*

aikwajuhua *(nome)* Remo. ***Oika aikwajuhua.*** *O remo quebrou.*

5

Figure 2.15: The Kawahíva print dictionary.

A second printed resource is a letter tracing workbook (2.16). This resource addresses one of the goals for language work on Kawahíva, i.e., to produce practical materials in the language, such as storybooks and other pedagogical materials, to increase the institutional visibility of the Kawahíva language and provide the Kawahiva, including children, with more ways to engage with the language. This workbook was made possible by a mini-grant fund from the Designated Emphasis on Language Revitalization in the Department of Linguistics at UC Berkeley.

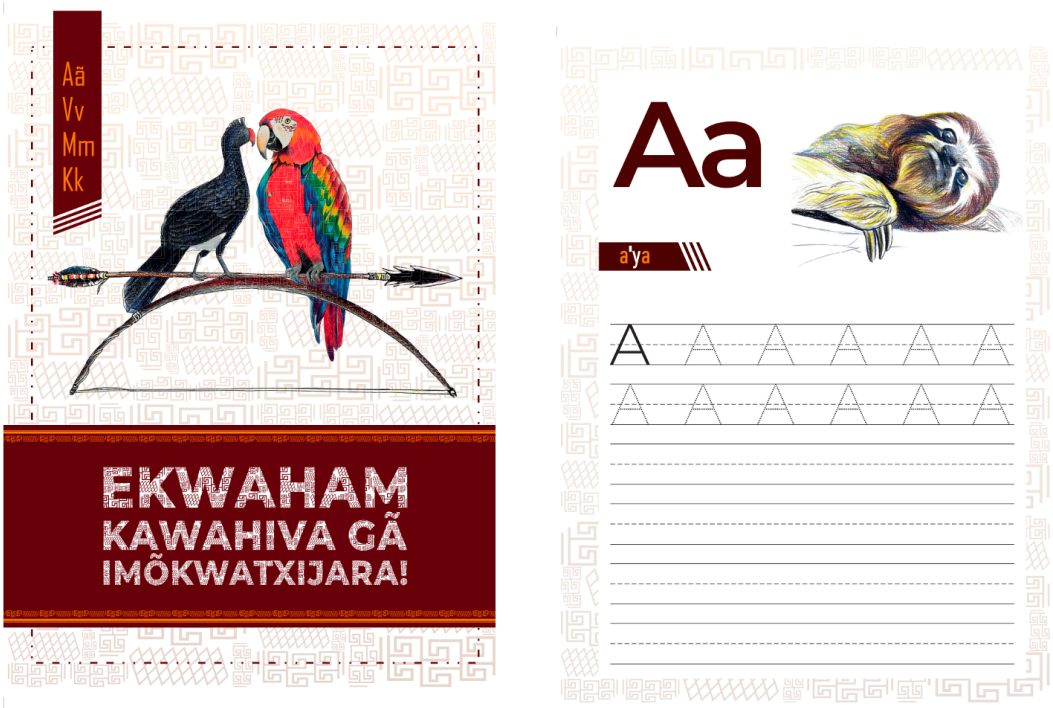
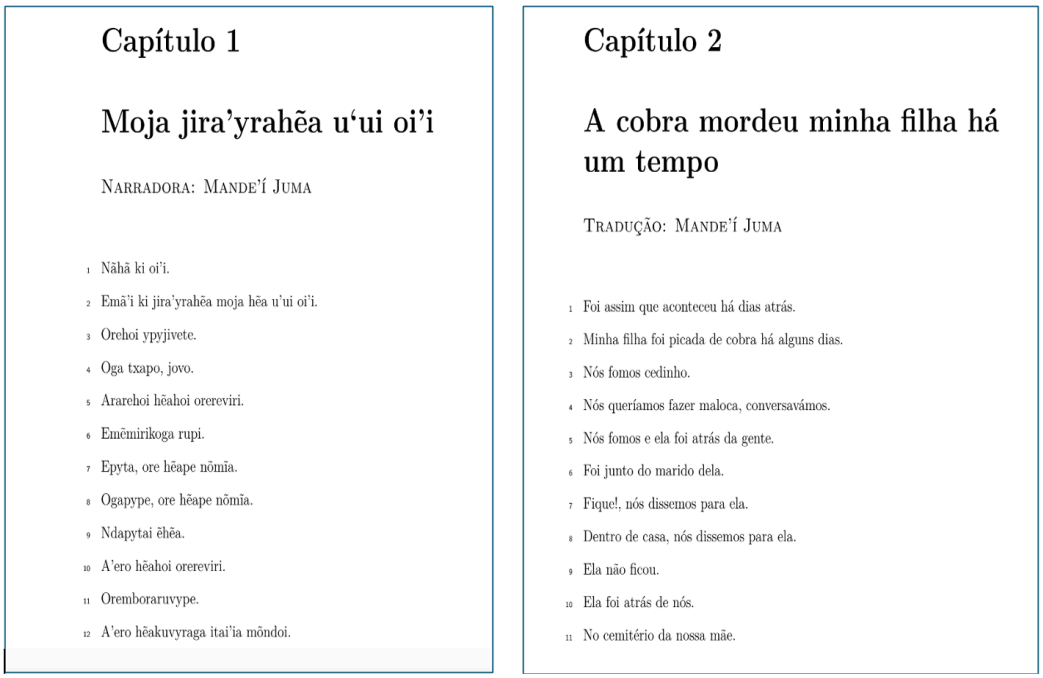


Figure 2.16: A Kawahíva letter tracing workbook. "Ekwaham Kawahiva Gã Imõkwatxijara!" means as "Learn the Kawahíva Writing!".

A third printed resource is a set of Kawahiva-Portuguese text collections, which will be delivered to the Juma and Jupaú communities within a semester. These resources were developed with the support of undergraduate research assistants at UC Berkeley: Matthew Ji, Maria Clara Castro, and Samantha Doyle-Jakobson. There are currently four sets of bilingual text collections for Juma, totaling over 1200 pages, and a 260-page text collection for Jupaú. These texts draw on the narratives and personal story recordings that were transcribed and translated by myself with the help of a native speaker using ELAN transcription software.



Capítulo 1

Moja jira'yrahẽa u'ui oi'i

NARRADORA: MANDE'Í JUMA

1 Nãhã ki oi'i.
2 Emã'i ki jira'yrahẽa moja hẽa u'ui oi'i.
3 Orehoi ypyjivete.
4 Oga txapo, jovo.
5 Ararehoi hẽahoi orereviri.
6 Emẽmirikoga rupi.
7 Epyta, ore hẽape nõmĩa.
8 Ogapype, ore hẽape nõmĩa.
9 Ndapytai ẽhẽa.
10 A'ero hẽahoi orereviri.
11 Oremboraruvype.
12 A'ero hẽakuvyraga itai'ia mõndoi.

Capítulo 2

A cobra mordeu minha filha há um tempo

TRADUÇÃO: MANDE'Í JUMA

1 Foi assim que aconteceu há dias atrás.
2 Minha filha foi picada de cobra há alguns dias.
3 Nós fomos cedinho.
4 Nós queríamos fazer maloca, conversavámos.
5 Nós fomos e ela foi atrás da gente.
6 Foi junto do marido dela.
7 Fique!, nós dissemos para ela.
8 Dentro de casa, nós dissemos para ela.
9 Ela não ficou.
10 Ela foi atrás de nós.
11 No cemitério da nossa mãe.

Figure 2.17: Kawahíva text collections.

Finally, we also plan to develop a children's illustrated storybook by the end of 2024. This book illustrates the story *Mbahira vapo kunhã* 'Mbahira creates the woman', on the genesis of the woman. This is one of the most traditional stories among the Kawahiva, and the first choice of storytellers when they are invited to narrate for recording.

2.3.4 Community workshops and panel

Another component of the Kawahíva Language Documentation Project involves organizing community workshops and panels with the Kawahíva. The workshops aim to address issues such as literacy resources and orthographic reform. Outside the villages, the panels help to raise awareness in the region about issues such as language endangerment and loss and language oppression and prejudice.

In conjunction with the Jupaú and Juma, we have held community workshops, which have been invaluable for expanding the project by creating new digital and print resources to increase literacy in the Indigenous language and assist schoolteachers with issues related to the practical writing system.

Figure 2.18: Community workshops among the Jupaú (top) and Juma.

Additionally, I organized a panel at a local public university, which took place in person and online.[10] The in-person venue was the Humaitá campus of the Universidade Federal do Amazonas, in the state of Amazonas, Brazil. Several Indigenous communities, both Kawahíva (Jiahui, Juma, Parintintin, and Tenharin) and non-Kawahiva (Munduruku, Mura, and Pirahã), come to this town to buy food, collect government social benefits, have medical examination and tests, or to live for work and/or to study at a local university. The audience at the panel included undergraduates from the School of Letters and the Department of Anthropology, members of the general community – parents and friends of undergraduates – and members of other Indigenous communities local to the area. Among other things, the event served to raise awareness about the Juma community and promote greater cultural understanding and sensitivity. Furthermore, the event was an opportunity for the Juma community to speak for themselves and share their story in person, which is a critical aspect of cultural preservation that helps to empower the community.

[10]This event was supported by a grant from the Center for Latin America and Caribbean Studies (CLACS) at UC Berkeley. The recording of the panel is available on the YouTube channel of CLACS: https://www.youtube.com/watch?v=mOcseVxdFmA.

Figure 2.19: The panel *The Juma in the South of the Amazon: memory, culture, identity, and language preservation through a multimedia dictionary.*

2.4 Conclusion

This chapter has provided the linguistic and sociohistorical background on the Kawahíva language and introduced the Kawahíva Language Documentation Project. The work I describe here is ongoing and aims to take a holistic view of language documentation and revitalization, incorporating the documentation of cultural knowledge.

Chapter 3

A Kawahíva grammar sketch

3.1 Introduction

This chapter provides an overview of Kawahíva grammar, focusing on the morphosyntactic properties of the language, which have not been thoroughly documented in previous descriptive work. While this overview sets the scene for the rest of the dissertation, subsequent chapters are also intended to stand alone, with relevant background information provided in each.

The chapter is structured as follows: First, I give a brief overview of Kawahíva phonology in §3.2, including the segmental phonology in §3.2.1, the orthography adopted in this dissertation in §3.2.2, and the major phonological processes in §3.2.3. Then, I turn to an overview of the morphosyntax and distribution of the major word classes in the language, including nouns in §3.3, verbs in §3.4, particles in §3.5, and postpositions in §3.6. Finally, I describe the clause structure and several clause-level phenomena in §3.7.

3.2 Phonology and writing system

This section presents my phonological analysis of Kawahíva, based on the Juma and Jupaú dialects. For alternative analyses of Kawahíva phonology, see Pease and Betts (1971) for the Parintintin dialect, Netto and Moraes (1993) for Amondawa, Sampaio (1997) for Jupaú, and Marçoli (2018) for Amondawa, Jiahui, and Tenharin.[1]

[1]The previous proposals generally overlap with mine, although some show greater similarity than others. For instance, the number of phonemic consonants proposed here, 13, matches the number of consonants proposed in Marçoli (2018). In contrast, Pease and Betts (1971) and Sampaio (1997) propose a total of 14 consonants. The difference of one extra phoneme in the latter authors' works is due to their including the nasal stop [ŋʷ] as a phoneme, which is never observed in more recent proposals, including the one presented in this chapter. Additionally, Netto and Moraes (1993) describe a total of 12 consonants; the smaller number of consonant phonemes in their work follows from their not including the affricate [tʃ]. A thorough comparison between these proposals is beyond the scope of this chapter, which I plan to address in later work.

3.2.1 Sound inventory, phonotactics and stress

Kawahíva's phone inventory includes 34 sounds: 22 consonants, provided in Table 3.1, and 12 vowels, provided in Tables 3.5-3.6. Phones in square brackets result from allophony.

	bilabial	labiodental	alveolar	postalveolar	palatal	velar	labiovelar	glottal
stop	p		t			k	k^w	ʔ
fricative		v			[ʒ]			h
affricate				tʃ [dʒ]				
nasal	m		n		[ɲ]	ŋ		
postoralized nasal	[mb]		[nd]		[ɲdʒ]	[ŋg]	[ŋg^w]	
flap			ɾ					
approximant					j		[w]	

Table 3.1: Kawahíva consonant phone inventory.

Kawahíva's phonemic inventory includes a total of 13 consonants, as shown in Table 3.2. Consonants contrast three manners of articulation: obstruent, oral sonorant, and nasal sonorant. Obstruents contrast six places of articulation: bilabial, labiodental, alveolar, postalveolar, velar, and labiovelar; while oral sonorants contrast only three places of articulation. Specifically, oral sonorants contrast alveolar, palatal, and glottal points of articulation, while nasal ones contrast bilabial, alveolar, palatal, and velar points of articulation. Nasalized sonorants are also observed as the result of phonological processes but are not phonemic.

	bilabial	labiodental	alveolar	postalveolar	palatal	velar	labiovelar	glottal
obstruent	p	v	t	tʃ		k	k^w	
oral sonorant			ɾ		j			ʔ h
nasal sonorant	m		n			ŋ		

Table 3.2: Kawahíva phonemic consonant inventory.

Tables 3.3-3.4 provide (near-)minimal pairs supporting the phonemic status of each consonant, as well as the manner and place of articulation contrasts shown in Table 3.2.

Consonant	Minimal and near-minimal pairs	
p : m	[pɛ̃mɛ̃] ‘to you (PL.)’	[mɛ̃mɛ̃] ‘equal’
p : m	[ɔpĩn] ‘s/he shut off’	[ɔmĩm] ‘s/he stole’
v : m	[ɨvɨa] ‘land, ground’	[ɨ̃mɨ̃] ‘be a long time ago’
v : ɾ	[-vaʔɛ] ‘WH.INTR.SUBJ’	[ɾaʔɛ] ‘today past’
t : ɾ	[ãtã] ‘be hard’	[aɾa] ‘day’
t : ɾ	[vata] ‘s/he walked, defecated’	[ɔvaɾa] ‘s/he mixed (it) up’
t : n	[ituʔã] ‘someone’s navel’	[inũa] ‘another one’
tʃ : j	[tʃuɾuguhua] ‘bread’	[ʒuɾua] ‘someone’s mouth’
tʃ : j	[ĩtʃĩa] ‘something’s beak’	[ĩɲĩʔã] ‘*mandi* fish’
k : ŋ	[ãkã] ‘branch’	[ʔãŋã] ‘DEM:PROX:ELONGATED’
kʷ : ŋ	[ãkã] ‘branch’	[ʔãŋã] ‘DEM:PROX:ELONGATED’
j : ʔ	[ɔdʒãm] ‘it hatched’	[ɔʔãn] ‘s/he fell (purposefully)’
j : h	[ɔdʒãm] ‘it hatched’	[ɔhãm] ‘s/he cut (it)’

Table 3.3: Consonant manner of articulation contrasts.

Consonant	Minimal pairs and near-minimal pairs	
p : v	[kapɨra] ‘forest’	[kavɨra] ‘leaf’
p : v	[pɨɾ̃ĩm] ‘be correct, well’	[vɨhɨa] ‘bleed’
p : k	[pɔɾɔ] ‘onomatopoeia for any noise’	[kɔɾɔ] ‘DEM:PROX:ROUND’
p : tʃ	[piʔua] ‘gnat’	[tʃiʔu] ‘let’s eat’
p : t	[pɛhɛa] ‘path’	[tɛhɛ] ‘idly’
t : v	[tata] ‘fire’	[vata] ‘s/he walked, defecated’
t : tʃ	[tuɾuʔũma] ‘mud’	[tʃuɾuguhua] ‘bread’
t : k	[tata] ‘fire’	[kava] ‘wasp’
k : kʷ	[kaɾa] ‘yam’	[kʷaɾa] ‘sun’
k : ʔ	[ɛki] ‘come in!’	[ɛʔi] ‘say (it)!’
j : ɾ	[ɔjãm] ‘it hatched’	[ɔɾãm] ‘s/he untied (it)’
ʔ : h	[aʔɛa] ‘this one’	[ahɛ] ‘person’
m : n	[mãhã] ‘when’	[nãhã] ‘thus’
m : n	[ɔpãm] ‘s/he finished’	[ɔpãn] ‘s/he lost the way’
m : ŋ	[mãhã] ‘when’	[ŋãhã] ‘3.PL.FOC’
m : ŋ	[ikãm] ‘be fat’	[ikãŋ] ‘be dry’
n : ŋ	[nãhã] ‘thus’	[ŋãhã] ‘3.PL.FOC’
n : ŋ	[ɔpɔ̃n] ‘s/he jumped’	[ɔpɔ̃ŋ] ‘s/he exploded’

Table 3.4: Consonant place of articulation contrasts.

Kawahíva’s vowel inventory is provided in Tables 3.5-3.6. Vowels in Kawahíva may be oral or nasal, but this contrast appears to be restricted to the root-final syllable, which coincides with the location of stress. Vowels are analyzed as contrasting two heights, high and low, and three values of backness, front, central, and back.

	Front	Central	Back
High	i	ɨ	u
Low	ɛ	a	ɔ

Table 3.5: Oral vowel phonemes.

	Front	Central	Back
High	ĩ	ɨ̃	ũ
Low	ɛ̃	ã	ɔ̃

Table 3.6: Nasal vowel phonemes.

Tables 3.7-3.9 provide minimal and near-minimal pairs supporting the phonemic status of each vowel. The minimal pairs in Table 3.7 illustrate a two-way height contrast for front vowels

/i, ɛ/, central vowels /ɨ, a/, and back vowels /u, ɔ/.

Vowel	Minimal and near-minimal pairs	
i : ɛ	[ɛpi] ‘shut off!’	[pɛ] ‘to, for’
ɨ : a	[ahɨ] ‘be in pain’	[ɛha] ‘have good vision’
u : ɔ	[aɾua] ‘*aɾua* frog’	[aʔɛɾɔ] ‘then’

Table 3.7: Vowel height contrasts.

Table 3.8 provides minimal and near minimal pairs supporting a three-way backness contrast for high vowels /i, ɨ, u/ and low vowels /ɛ, a, ɔ/.

Vowel	Minimal and near-minimal pairs	
i : ɨ	[ɛpi] ‘shut off!’	[ɛpɨ] ‘blow (it)!’
i : u	[ɛpi] ‘shut off!’	[ɛʔu] ‘bite (it)!’
u : ɨ	[idʒũn] ‘be rotten’	[idʒɨ̃n] ‘be hard’
ɛ : a	[ahɛ] ‘person’	[-aha] ‘unfortunately’
a : ɔ	[ɛha] ‘have good vision’	[ɛhɔ] ‘go!’

Table 3.8: Vowel backness contrasts.

Finally, Table 3.9 provides minimal and near minimal pairs as evidence of contrastive vowel nasality for six vowel qualities /ĩ, ɨ̃, ũ, ɛ̃, ɔ̃, ã/.

Vowel	Minimal and near-minimal pairs	
i : ĩ	[ɛviɾi] ‘behind something’	[ɔpɔkɨ̃ɾ̃ĩ] ‘s/he tickled someone’
ɨ : ɨ̃	[ɛʔɨi] ‘many (precise number)’	[vɛ̃ʔɨ̃ĩ] ‘s/he scratched’
u : ũ	[ikuʔa] ‘an animal’s hing legs’	[ikũa] ‘someone’s tongue
ɔ : ɔ̃	[pɛɾɔ] ‘LOC.DEM:MED’	[vãɾ̃õ] ‘s/he waited’
ɛ : ɛ̃	[pɛ] ‘to, for’	[pɛ̃] ‘2.PL’
a : ã	[ga] ‘3.SG.MASC’	[ŋã] ‘3.PL’

Table 3.9: Vowel nasality contrasts.

Vowel-vowel sequences, such as the ones listed in Table 3.10, are frequent in Kawahíva. I leave to future work whether these vowel sequences are single complex units (i.e., diphthongs) or sequences of vowels that occupy the nuclei of distinct syllables.

Vowel-vowel	Example	Gloss
ai	[aipɛʔi]	‘one, once’
au	[auhu]	‘maybe’
ia	[iɾipiakaɾa]	‘my soul’
ea	[ɛaviɨɾɨ]	‘be drunk’
ei	[ɛiɾuva]	‘bee sp.’
eu	[ivɛuvɛ]	‘be lazy’
oa	[ndipɔakaɾi]	‘weak’
oi	[kɔiʔi]	‘many’
oɨ	[ɔɨʔĩjãɾɔ]	‘hit’
ɨa	[ipɨahuɛtɛka]	‘young man’
ɨu	[ipɨu]	‘ripe’
ɨi	[ipɔhɨi]	‘heavy’
ui	[ɨvuivɛʔi]	‘small’

Table 3.10: Vowel-vowel sequences.

Kawahíva syllable structure is summarized as (C)V(C) where the CV syllable is the most common type. Any vowel may occupy the nucleus position of the syllable, and any consonant can be the onset of a syllable.

(1) [ã.tã] ‘be hard’ (V)

(2) [ga] ‘3.SG.MASC’ (CV)

Kawahíva also allows CVC syllables, but with some restrictions on the segments that can appear in the coda position. Codas most commonly appear word-finally, where only nasal stops (and their oralized allophones) and glottal stops are attested.[2]

(3) [i.pu.tu.pãm] ~ [i.pu.tu.pap] ‘be tired’

(4) [ɔ.hĩn] ~ [ɔ.hit] ‘s/he fell’

(5) [ɔ.pɨ.hĩŋ] ~ [ɔ.pɨ.hɨk] ‘s/he caught (it)’

(6) [mãʔ] ‘huh?, what’

[2]The alternation between the word-final oral and nasal stops in examples (3-5) occurs within the same speaker of both dialects when pronouncing these words in isolation. In most cases, the nasal stop alternant occurs in careful pronunciation, while the oral stop alternant occurs in fast pronunciation. A similar alternation is described in Awetí (Drude 2021), the closest relative to the Tupí-Guaraní language family. However, the Awetí pattern is not analyzed in these terms. Rather, Drude (2021) proposes the archiphonemes /P, T, K/, which surface as an oral stop or nasal stop depending on the oral or nasal status of the preceding vowel; if nasal, then the stop surfaces with the nasal allophone, but as an oral stop otherwise. This analysis does not carry over to Kawahíva as the vowel is underlyingly oral in these contexts. If it were nasal, the vowel would trigger a process of nasal harmony to its left (described in §3.2.3.2), which is not true, as also evidenced by the examples (3-5).

Word-internally, only glottal stops appear in the coda position. However, this surface distribution results from metathesis of the only glottal stop-initial morpheme, the diminutive suffix /-ʔi/. I discuss this process in §3.2.3.10. Examples of a glottal stop followed by a consonant also constitute the only heterosyllabic consonant sequences in the language; there are no instances of tautosyllabic consonant sequences. Additionally, I have not found examples of the VC syllable type.

Stress is assigned to the root-final syllable of a word. Examples of stress assignment to the root-final syllable are provided for verbs and nouns in (7-8) and (9-10), respectively. In nouns, this stress pattern is a bit obscured as they always surface with the extrametrical nominalizing enclitic *=a*. On the surface, then, stress falls on the penultimate syllable of nouns bearing the enclitic *=a*.

(7) /ɛ-pi/ → [ɛ.ˈpi]
2.SG.IMP-shut.off
'Shut off!'

(8) /ɛ-kɨtʃi/ → [ɛ.kɨ.ˈtʃi]
2.SG.IMP-cut
'Cut (it)!'

(9) /janu=a/ → [ɲã.ˈndu.a]
spider=NMLZ
'spider'

(10) /tapiʔiɾ=a/ → [ta.pi.ˈʔi.ɾa]
tapir=NMLZ
'tapir'

However, some suffixes trigger stress shift from the root-final syllable, both in verbs and nouns. For instance, stress remains on the final syllable of the verb root in (11) after circumfixation of the sentential negation *nd(a)-...-i*, but shifts in (12) after suffixation of *-ahim* 'VERY'. Another suffix that triggers stress shift is *-uhu* 'AUG' (13-14). Examples (15)-(16) provide evidence of stress shift due to suffixation in nouns.

(11) /n-aimɛ-i/ → [nd-aĩˈmbɛ-i]
NEG-sharp-NEG
'be dull'

(12) /aimɛ-ahim/ → [aĩmbɛ-aˈhĩm]
sharp-VERY
'be very sharp'

(13) /i-ipi/ → [j-iˈpi]
i-dark
'be dark'

(14) /i-ipi-uhu/ → [j-ipi-uˈhu]
i-be.dark-AUG
'be very dark'

(15) /vɨɾa=a/ → [vɨˈɾa]
bird=NMLZ
'bird'

(16) /vɨɾa-ʔi=a/ → [vɨɾaˈʔia]
bird-DIM=NMLZ
'small bird'

3.2.2 Writing system

In this dissertation, I adopt the practical writing system used by the Jupaú and Juma, which is provided in tables 3.11-3.12.

	bilabial	alveolar	postalveolar	palatal	velar	labiovelar	glottal
stop	p	t			k ~ g	kw	'
fricative	v						h
affricate			tx				
nasal	m	n		nh	ng	ngw	
flap		r					
approximant				j			

Table 3.11: Kawahíva orthographic consonant inventory.

	front	central	back
high	i, ĩ	y, ỹ	u, ũ
mid	e, ẽ		o õ
low		a, ã	

Table 3.12: Kawahíva orthographic vowel inventory.

Most graphemes represent the same sounds they represent in English, with a few orthographic conventions worth noting, including ⟨y⟩ = /ɨ/, the apostrophe ⟨'⟩ = /ʔ/, and ⟨tx⟩ = /ʧ/. The most significant orthographic idiosyncrasy of this system is the use of <g> to represent /ŋ/ in the sequence ⟨gã⟩.

This writing system is quite similar to the one used in the materials developed by the missionaries of the Summer Institute of Linguistics (SIL), LaVera Betts and Helen Pease, for the Parintintin

and Tenharin dialects, including the Parintintin dictionary (Betts 1981:7-8) and Kawahíva pan-dialectal dictionary (Betts 2012:2-3). Given the similarities between the writing systems adopted in the Juma and Jupaú villages, I suspect Betts and Pease's system was disseminated among the Jupaú and Juma, most likely via other missionaries who lived in these communities, as Betts and Pease never lived among the Juma and Jupaú.

3.2.3 Phonological and morphophological processes

I now discuss central phonological and morphophonological processes in Kawahíva based on the dialects spoken by the Jupaú and Juma.

3.2.3.1 Nasal post-oralization

Kawahíva nasal consonants show no contrast in stressed syllables. In stressed syllables, plain nasal consonants become post-oralized before oral vowels, which I schematize as a phonological rule in 17; 'N' stands for a plain nasal stop, 'ND' for a post-oralized nasal stop, and 'V' for an oral vowel. This process is illustrated in the examples in (18). I indicate primary stress in phonological rules and examples with the diacritic (ˈ).

(17) N → ND / ˈ_V

(18) a. /amɛ=a/ → [ãˈmbɛ=a] 'animal's belly'
b. /ɛnɨ/ → [ɛ̃ˈndɨ] 'be lit'

In unstressed syllables, plain and postoralized nasals are in free variation, as evidenced in the examples from Juma (19) and Jupaú (20).

(19) a. /mɛɾu=a/ → [mɛˈɾu=a] 'fly' (Juma)
b. /a-mɛka/ → [a-mɛˈka] 'I opened (it)' (Juma)
c. /mɛvɛ/ → [mbɛˈvɛ] 'be slow' (Juma)
d. /miaɾa=a/ → [mbiaˈɾa] 'meat' (Juma)

(20) a. /mɛɾu=a/ → [mbɛˈɾu=a] 'fly' (Jupaú)
b. /mutuk=a/ → [muˈtuk=a] 'horsefly' (Juma and Jupaú)

The analysis of the alternations between plain nasal stops and homorganic post-oralized nasal stops in stressed syllables is a major point of disagreement among analyses of Tupí-Guaraní languages. In some languages, the plain nasal stops are posited as the unconditioned allophones, and post-oralized nasals, as conditioned allophones, while other languages have received the opposite analysis (Lapierre and Michael 2017; Miranda 2018:for an overview). In previous phonological proposals for Kawahíva, the plain nasal stops have been unanimously posited as the unconditioned allophones, but without further discussion (Pease and Betts 1971; Sampaio 1997; Marçoli 2018). Perhaps the best piece of evidence in Kawahíva that the plain nasal stops are the unconditioned allophones is they occur in both onset and coda position, while post-oralized nasals are restricted to onset position.

3.2.3.2 Anticipatory nasal assimilation and long-distance nasal harmony

I turn now to the discussion of the patterns of nasality, in which the two dialects behave similarly. Both dialects exhibit a pattern of nasal assimilation and long-distance nasal harmony (Walker 2011:for a typological overview). In local nasal assimilation, a nasal stop (plain or postoralized) nasalizes an immediately preceding vowel.[3]

Plain nasal stops in word-final coda position nasalize an immediately preceding vowel. This process is schematized as a phonological rule in (21). Examples of this process are provided in (22); 'N' stands for a plain nasal stop. Likewise, postoralized nasal stops nasalize the immediately preceding adjacent vowel, as in (23). Nasalized vowels are not considered underlyingly nasal, as phonemic nasal vowels trigger a process of long-distance nasal harmony, as I discuss next. Crucially, this process does not arise in (21)-(23).

(21) $V \rightarrow \tilde{V}$ / _N

(22) a. /ɨɾup**ɛm**=a/ → [ɨɾup**ɛ̃m**=a] 'basket'
b. /v-ɔvavɨt**im**/ → [v-ɔvavɨt**ĩm**] 's/he corralled (it)'
c. /a-mɔ-havivu**n**/ → [ã-mbɔ-havivu**ũn**] 'I made (it) stuffed'
d. /ɛ-up**in**/ → [ɛ-up**ĩn**] 'lift (it)!'
e. /i-kaɾ**aŋ**/ → [i-kaɾ**ãŋ**] 'be handicapped'
f. /ɔ-kaɾ**uŋ**/ → [ɔ-kaɾ**ũŋ**] 's/he returned'

(23) a. /ak**im**ajuv=a/ → [ak**ĩmb**adʒuv=a] 'bamboo'
b. /ak^{w}**ɛm**aʔɛ=a/ → [ak^{w}**ɛ̃mb**aʔɛ=a] 'man'
c. /v-ah**in**ɔŋ/ → [v-ah**ĩnd**ɔŋ] 'it ripped'
d. /i-kupɛk**aŋ**=a/ → [i-kupɛk**ãŋ**g=a] 'someone's back'
e. /i-p**uŋ**a/ → [i-p**ũŋg**a] 'be swollen'

Kawahíva also exhibits long-distance nasal harmony, in which oral segments nasalize when a phonemic nasal vowel appears to their right. I adopt the standard terms *trigger*, *target*, *blocker* (or *opaque*) segment and *transparent* segment to describe the behavior of Kawahíva phonemes within this pattern of nasal harmony (Walker 2011). A *trigger* is a segment that causes nasal harmony. A *target* segment is one that undergoes nasalization. A *blocker* or *opaque* segment is one that blocks segments to its left from being nasalized by the trigger. Finally, a *transparent* segment is a segment that does not undergo nasalization but does not block nasal spreading. Table 3.13 summarizes the behavior of the Kawahíva segments under long-distance nasal harmony. The discussion below moves from the top to the bottom row.

[3]While the claims made here about the patterns of nasalization are impressionistic, preliminary observations of airflow data in words with long-distance nasal harmony seem to support these generalizations.

behavior	segment
trigger	/ĩ, ɪ̃, ũ, ɛ̃, ɔ̃, ã/
target	/v, j, ɾ, i, ɨ, u, ɛ, ɔ, a/
transparent	/p, v, t, tʃ, k, kʷ, m, n, ŋ, h/
blocker	/ʔ/

Table 3.13: Segment behavior under nasal harmony.

Phonemically nasal vowels are the only triggers of long-distance nasal harmony, which generally nasalizes segments to their left in both dialects. The segments affected in both dialects include sonorant segments (vowels, the tap, and /j/) and /v/. Examples of this process in Juma are provided in (24). Note that these examples also show that long-distance nasal harmony does not cross a morpheme boundary in Juma. I return to this point shortly.

(24) a. /ɛ-avĩ/ → [ɛ-ãṽĩ] ‘squeeze (it)!’
b. /ɔ-kutɛ̃i/ → [ɔ-kũtɛ̃i] ‘to be loose’
c. /akã/ → [ãkã] ‘branch’
d. /kujã=a/ → [kũɲã] ~ [kũj̃ã] ‘woman’
e. /v-ɛɾõi/ → [v-ɛ̃ɾ̃õi] ~ [v-ɛ̃nõi] ‘s/he remembered (it)’
f. /v-apatʃĩ/ → [v-ãpãtʃĩ] ‘s/he tied (it)’
g. /v-ɔvatʃĩ/ → [v-ɔ̃ṽãtʃĩ] ‘s/he met (someone)’
h. /j-atɨaɾɔ̃/ → [ʒ-ãtɪ̃ãɾ̃ɔ̃] ‘be ripe’

Nasal consonants never trigger long-distance nasal harmony, even when they occur in the root-final position, as demonstrated by the examples in (22) above. All obstruents (stops and fricatives), except for the glottal stop, are transparent to long-distance nasal harmony.

The glottal stop /ʔ/ blocks long-distance nasal harmony, evidenced in the examples in (25).

(25) a. /i-tuʔã=a/ → [i-tuʔã] ‘someone’s navel’
b. /jaʔĩ=a/ → [ʒaʔĩ=a] ~ [jaʔĩ=a] ‘catfish’
c. /tʃuʔĩ/ → [tʃuʔĩ] ‘be small’
d. /ɔ-kaʔɛ̃/ → [ɔ-kaʔɛ̃] ‘s/he got cured’
e. /ɔ-kɔʔɔ̃/ → [ɔ-kɔʔɔ̃] ‘s/he got burned’
f. /i-kuɾuʔã=a/ → [i-kuɾuʔã] ‘someone’s throat’
g. /ji-ɾɛɾɨpɨʔã=a/ → [ji-ɾɛɾɨpɨʔã] ‘my knee’

The two dialects exhibit different patterns of long-distance nasal harmony across morpheme boundaries. In Juma, morpheme boundaries are generally opaque to nasal harmony; this is also illustrated in the above examples in (24). The only instance of nasal harmony crossing a morpheme boundary in Juma occurs with the prefix *i-*, as in (26). I describe the distribution of *i-* in §3.7.15.

(26) a. /i-pɔ̃-uhu=a/ → [ĩ-pɔ̃-ũhũ=ã] ‘someone’s thumb’

In Jupaú, this process depends on the speaker's gender: in women's speech, but not men's speech, long-distance nasal harmony affects the morpheme immediately to the left of the root. To illustrate this pattern, I provide the following examples in (27). The first pair shows a verb inflected with the 3rd person subject agreement marker *ɔ-*, whereas the second pair of examples show the same verb inflected with the 1st person exclusive plural agreement marker *ore=*.

(27) a. /ɔ-tʃĩ/ → [ɔ̃-tʃĩ] (3.A-tie.up) 's/he tied (someone) up' (women's speech)
b. /ɔ-tʃĩ/ → [ɔ-tʃĩ] (3.A-tie.up) 's/he tied (someone) up' (men's speech)
c. /ɔɾɛ=tʃĩ/ → [ɔ̃ɾ̃ɛ̃=tʃĩ] (1.EXCL=tie.up) '(someone) tied us up' (women's speech)
d. /ɔɾɛ=tʃĩ/ → [ɔɾɛ=tʃĩ] (1.EXCL=tie.up) '(someone) tied us up' (men's speech)

In describing the pattern of long-distance nasal harmony at the morpheme boundary as tied to the speaker's gender, Jupaú can be characterized as having a 'genderlectal phonological rule' (or 'genderlects'). That is, the speaker's gender is indexed in the phonology (Dunn 2014; Rose 2015:for overviews).

3.2.3.3 Progressive nasal assimilation

Both dialects also exhibit a pattern of progressive nasal assimilation, in which a nasal vowel partially or completely nasalizes the immediately following segment(s). I schematize this process in (28), where 'T' stands for a voiceless oral stop. In particular, oral stops are partially nasalized, appearing as prenasalized stops, as shown in the examples (29), while oral vowels are nasalized, as evidenced in the examples in (30).

(28) /V, T/ → [Ṽ, ND] / Ṽ+_

(29) a. /mɔ̃-pɨta/ 'CAUS-stay/stop' → [mɔ̃-mbɨta] 'stop (something/someone)'
b. /mɔ̃-katu/ 'CAUS-be.pretty' → [mɔ̃-ngatu] 'dress up (someone)'
c. /mɔ̃-kwerai/ 'CAUS-be.tired' → [mɔ̃-ngwerai] 'make someone tired'
d. /ɔ-ikɔtʃĩ-pap/ → [ɔ-ĩkɔ̃tʃĩ-mbap] 'melt completely'
e. /ɔ-ji-mɔ̃ji-katu/ → [ɔ-ji-mɔ̃ɲĩ-ngatu] 'recover oneself well'

(30) a. /i-pɔ̃-uhu=a/ → [ĩ-pɔ̃-ũhũ=ã] 'someone's thumb'
b. /i-pɔ̃-ʔi=a/ → [ĩ-pɔ̃-ʔĩ=ã] 'someone's small finger'

I characterize progressive nasal assimilation as a different process from anticipatory nasal assimilation, which I described in §3.2.3.2. The former, but not the latter, creates a prenasalized oral stop. One question for future work is whether prenasalized stops [^{n}t] differ phonetically from the postoralized nasal stops [n^{t}], as it has been demonstrated for the Panãra language (Jê family, Brazil). In Panãra, there is a surface contrast between prenasalized oral stops [^{n}t] and postoralized nasal stops [n^{t}] – the two differ in the extent of the duration of nasal airflow (Lapierre 2021).

3.2.3.4 /i/-allomorphy

Both dialects exhibit a phonologically conditioned allomorphy of the prefix /i-/, which becomes [j-] ~ [dʒ-] ~ [ʒ-] when prefixed to non-nasal vowel-initial roots (31a), [ɲ-] when prefixed to nasal vowel-initial roots (31b), [ĩ-] when attached to a nasal vowel-initial root (31c), and [i-] elsewhere (31d). The grammatical distribution of the prefix /i-/ is presented in §3.7.15.

(31) a. /i-aku-ahĩm/ → [j-aku-ahĩm] ~ [dʒ-aku-ahĩm] ~ [ʒ-aku-ahĩm] (i-be.hot) 'be hot'
b. /i-uɾãm-ɛtɛ/ → [ɲ-ũr̃ãmɛ̃tɛ] (i-be.slow-REALLY) 'be slow'
c. /i-nẽm-ahim/ → [ĩ-nẽm-ahĩm] (i-be.rotten) 'be rotten'
d. /i-katu/ → [i-katu] (i-be.pretty) 'be pretty'

3.2.3.5 Root allomorphy

Juma and Jupaú exhibit a process of initial-consonant alternation, in which the initial consonant of some roots undergoes a systematic alternation with a phonetically distinct consonant. As an example of such a systematic alternation in the initial consonant of a word, consider the root pairs from Juma in (32). As seen in the pairs, roots alternate between an [ɾ]-initial form and one other form that depends on the root. This other form can be either one of the following: [t]-, [h]-initial, or a consonant-less form. I assume which roots exhibit which alternation to be lexically determined.[4]

(32) a. [taʔɨ] ~ [ɾaʔɨ] 'offspring'
b. [hɛpia] ~ [ɾɛpia] 'see'
c. [ɛkɔ] ~ [ɾɛkɔ] 'be, live, marry'
d. [ɛhɛ] ~ [ɾɛhɛ] 'at'

The grammatical context in which the word appears determines which of two forms occurs in any particular instance: when the root and the immediately preceding constituent (usually their syntactic complement) belong to the same phrase, the [ɾ]-form appears; if there is no immediately preceding constituent or the immediately preceding constituent is not part of the same phrase, the root appears with the other alternant. Examples in (33) through (35) illustrate provide evidence for this context trigger for nominal, verbal, and postpositional roots. In (33a), the possessor of the inalienable noun is overt, and immediately precedes the possessed noun. However, in (33b), the possessor is null. As a result, the root surfaces as [ɾ]-initial only in the first case and [t]-initial in the second one.

(33) a. nde=**ra'yr**=a=ga
2.SG=offspring=3.SG.MASC
'Your son' (Juma: Elicit)

[4]In general, both Jupaú and Juma exhibit the same pairs of alternations. However, in Jupaú, the alternating pair [ɾ]- versus [h]-initial form is very restricted. It has been found only with the postposition [hɛhɛ] ~ [ɾɛhɛ] 'at', and in one single instance, which comes from the traditional story on the genesis of women narrated by Mandá Uru Eu Wau Wau.

b. **ta'yr**=a
offspring=NMLZ
'(Someone's/something's) offspring' (Juma: Elicit)

In (34), the root alternation involves the verb 'be, live, marry'. In (34a), this verb appears in the [ɾ]-initial form, as it is preceded by the same phrase (i.e., the VP) noun *matera* 'food, thing'. In contrast, in (34b), this verb appears in its initial consonant-less form, as the preceding constituents, including the postpositional phrase *'ype* 'in the river', are not part of the VP with the verb. For instance, there is a pause between the postpositional phrase *'ype* 'in the river' and the verb (marked with commas), which I take as evidence that they do not belong to the same prosodic phrase, and therefore, they do not belong to the same syntactic phrase either, assuming that prosodic phrases match syntactic phrases (Selkirk 2011).[5]

(34) a. Ji **matera reko**-i j-apo-vo, 'i:te:ki:ko:ra.
1.SG food be-i i-make-VO DIR.EVID
'I was making food.' (Juma: Text)

b. Jakare=a, 'y-pe, **eko**-i.
alligator=NMLZ river-in live-i
'Alligators live in the river.' (Juma: Elicit)

In addition to nominal and verbal roots in (33) and (34), postpositional roots also exhibit root alternation conditioned by the presence of a constituent that belongs to the same phrase. Examples of these alternations are provided in (35).

(35) a. E-ho ga=**renõnde**.
2.SG.IMP-go 3.SG.MASC=ahead
'Go ahead of him!' (Juma: Elicit)

b. Gã=hã ki **tenõnde** uhu ko.
3.PL.FOC PST ahead come REAL
'THEY came in first.' (lit.: they came ahead (of us) (Juma: Text)

I characterize the alternations in (32) as root allomorphy, and each alternant in the pair is an allomorph. This is also summarized in Table (3.14).

context	allomorph
$[\text{YP X}]_{\text{XP}}$	[r]-form
(YP), $[\text{X}]_{\text{XP}}$	[t]-, [h]-initial, or a consonant-less form

Table 3.14: Distribution of the syntactically conditioned root allomorphy.

[5]The direct evidential in example (34a), *'i te ki ko*, described in §3.5.4, is a non-compositional expression formed by a verb, an adverbial suffix, and some particles. I parse the different morphemes using colons but gloss the combination as 'direct evidential'.

The root allomorphy account of the initial-consonant alternations in (32) contrasts with the usual approach to similar facts in sister languages to Kawahíva. Several scholars consider the alternating consonants as 'relational prefixes' (Rodrigues and Cabral 2012a). For this view, for instance, the tap [ɾ-] is not part of the root but a prefix. This morpheme indicates "in a stem that it forms a syntactic unit with its determiner, which is the immediately preceding expression" (Rodrigues and Cabral 2012a:511), while the other morphemes indicate the absence of the said 'determiner'. I interpret the term 'determiner' in this definition as referring to the 'complement' in a HEAD-COMPLEMENT syntactic dependency. A summary of the two competing accounts for root alternations is presented in Table 3.15.

	triggering context	alternating consonants
Root allomorphy	any/no same-phrase constituent	not morphemes
Relational prefix	presence/absence of same-phrase complement	morphemes

Table 3.15: Summary of approaches to initial-consonant alternations.

However, the relational prefix account and its conditioning of the alternations in (32) face challenges. To start with the latter, we will see that [ɾ]-roots surface regardless of whether the immediately preceding word is a complement. This can be seen, for example, in (36) with the verb 'sleep' and 'come'. These verbs show an [ɾ]-form despite the fact that *kavyripe* 'in the river' and *ore* '1.EXCL' are not their syntactic complement.

(36) a. Kavyr-ipe **reki**=a nhãnde reko-i hemõ nõmĩa.
jungle-in sleep=NMLZ 1.INCL be-i ORAL.TRAD.EVID FRUST
'Sleeping in the jungle, we were.' (Juma: Text)

b. Pevo ore **rur**-i irytxangwa'ea o-vo-vo, 'i:te:ki:ko.
there 1.EXCL come-i soda 3.COR-drink-VO DIR.EVID
'There, we came to drink soda.' (Juma: Text)

The [ɾ]-form of the verbs *reki* 'sleep' and *reko* 'be' is an expected result under the relational prefix account, which would assume that the constituent immediately preceding the verbs is their complements, contrary to the truth. However, their [ɾ]-form follows straightforwardly within the root allomorphy account – the verbs and the constituent immediately preceding them form part of the same phrase, which should trigger the [ɾ]-initial allomorph.

The relational prefix account also faces challenges for historical and parsimony reasons. On the first count, Meira and Drude (2013) calls into doubt the morphological analysis historically, who see the "'relational' morphemes not as prefixes historically, but root-initial consonants that developed differently in different phonological and syntactic environments (Meira and Drude 2013:23)." The prefixal account requires assuming that there has been a reanalysis or morphologization of the alternating segments, an assumption not shared with the root allomorphy proposal adopted here.

A cross-linguistic argument can also be used to assess the feasibility of the relational prefix account. The Kawahíva root alternations are strikingly similar to the well-known phenomenon of *consonant mutation* from other language families (Grijzenhout 2011:for an overview), like Celtic

languages, where they are treated as stem alternations (Hannahs 2011:for an overview).[6] If we assume the root allomorphy idea, these similar phenomena can be explained in a similar way in both Celtic and Tupí-Guaraní languages. However, there is no such parallel for the relational prefix account. In the interest of analyzing empirically similar phenomena in a similar way, the allomorphic analysis of consonant mutation is strong evidence in favor of the root allomorphy account in Kawahíva (and Tupí-Guaraní more generally).

In summary, the root allomorphy analysis is the most historically coherent and analytically parsimonious account of the root alternations in Kawahíva.

3.2.3.6 Lenition (softening)

Kawahíva exhibits lenition of stops. For example, /p/ becomes [v] in contexts where it occurs between two phonologically oral front vowels at a syntactic word boundary, as in (37).

(37) *Lenition of /p/ → [v] between front vowels*
   a. /ji # pɛ/ → [dʒivɛ] '1.SG to'
   b. /nɛ # pɛ/ → [ndɛvɛ] '2.SG to'
   c. /janɛ # pɛ/ → [ɲãndɛvɛ] '1.INCL to'
   d. /ɔɾɛ # pɛ/ → [ɔɾɛvɛ] '1.EXCL to'

However, the stop /p/ remains unchanged in intervocalic contexts involving nonidentical oral vowels in a syntactic word boundary.

(38) *No lenition of /p/ between nonidentical vowels*
   a. /ga # pɛ/ → [gapɛ] '3.SG.MASC to'
   b. /hɛ̃a # pɛ/ → [hɛ̃apɛ] '3.SG.FEM to'
   c. /ŋã # pɛ/ → [ŋãpɛ] '3.PL to'

Finally, /p/ becomes a homorganic nasal stop counterpart [m] if the preceding vowel is nasal, as in (39). Note the vowel nucleus of [p] also nasalizes.

(39) *No lenition of /p/ between nasal vowels*
   /pɛ̃ # pɛ/ → [pɛ̃mɛ̃] '2.PL to'

3.2.3.7 Vowel deletion

At syntactic word boundaries, vowel hiatus is resolved by deleting the first vowel, regardless of whether it is stressed or not. Examples of vowel deletion in vowel hiatus are provided in (40).

(40) *Vowel deletion in vowel hiatus*
   a. /ki # ɔɾɛ/ 'PST' + '1.EXCL' → [kɔˈɾɛ]

[6] This similarity was first noted in Meira and Drude (2013:10).

b. /pɔ # ahɛ/ ‘IRR + person’ → [paˈ̃hɛ]
c. /dʒi # uvi/ ‘1.SG + from’ → [dʒuˈvi]
d. /kʷaɾa=a # ipɛ/ ‘sun + in’ → [kʷaɾiˈ̃pɛ]
e. /ãmãn=a # ipɛ/ ‘rain + in’ → [ãmãniˈ̃pɛ]
f. /imiaɾ=a # ipɛ/ ‘dwell in’ → [ĩmbiaɾiˈ̃pɛ]
g. /kavɨɾ=a # ipɛ/ ‘village + in’ → [kavɨɾiˈ̃pɛ]
h. /ɔkaɾ=a # ipɛ/ ‘village + in’ → [ɔkaɾiˈ̃pɛ]

At morpheme boundaries, vowel deletion applies to repair vowel hiatus involving identical vowels. For instance, vowel hiatus of identical vowels is created when roots that end with an underlying [a] vowel are combined with the nominalizing enclitic [=a], and one of the vowels is deleted, as in (41). Note the two [a] vowels surface when not immediately adjacent, as when a suffix intervenes between the root and the enclitic, as in (42).

(41) /vɨɾa=a/ → [vɨˈ̃ɾa]
bird=NMLZ
‘Bird’

(42) /vɨɾa-ʔi=a/ → [vɨɾaˈʔia]
bird-DIM=NMLZ
‘Small bird’

However, in cases of vowel hiatus of non-identical vowels at a morpheme boundary, vowel deletion applies to the vowel of some morphemes but not others. In examples (43), this process targets the initial vowel of the negation suffix *-e’ym* in (43a) but not the initial vowel of aspect morphemes (43b-c), person agreement (43d), and the imperative (43e). It remains for future work to determine what conditioning factor underlies this vowel deletion process to resolve vowel hiatus at morpheme boundaries in some cases, but not others.

(43) *Vowel hiatus resolution of non-identical vowels at a morpheme boundary*
a. /uʔi-ɛʔɨm=a/ ‘flour-NOM.NEG=NMLZ’ → [uʔi-ˈʔɨ̃m=a]
b. /a-ʔu-ipɛ/ ‘1.SG.A-eat-ALREADY’ → [a-ʔu-iˈ̃pɛ]
c. /ɛhɔ̃i-uhu/ ‘be.big-AUG’ → [ɛhɔ̃i-ũˈhũ]
d. /ɔ-ɨɾɔ̃/ ‘3.A-wait’ → [ɔ-ɨ̃ˈɾɔ̃]
e. /ɛ-auʔi/ ‘2.SG.IMP-cover’ → [ɛ-auˈʔi]

3.2.3.8 /h/-deletion

The Juma dialect exhibits a process of root-initial /h/-deletion in unstressed syllables at an intervocalic morpheme boundary between two identical vowels (e.g., /ɛ/_/ɛ/). This process is schematized as a phonological rule in (44); the identity between the vowels in the conditioning environment

for /h/-deletion is represented using identical subscripts. Examples of /h/-deletion are provided in (45).[7]

(44) *The root-initial /h/-deletion process in Juma*
$h \rightarrow \varnothing$ / V_i_+V_i

(45) *Root-initial /h/-deletion in morpheme boundary between /ɛ/_/ɛ/ in unstressed syllables*
a. /a-hɛpiaŋ/ '1.SG.A-see' → [a-hɛˈpiãŋ]
b. /ɛɾɛ-hɛpiaŋ/ '2.SG.A-see' → [ɛɾɛ-ˈpiãŋ] (after further deletion of root vowel [ɛ])
c. /ɔɾɔ-hɛpiaŋ/ '1.EXCL.A-see' → [ɔɾɔ-hɛˈpiãŋ]
d. /tʃi-hɛpiaŋ/ '1.INCL.A-see' → [tʃi-hɛˈpiãŋ]
e. /pɛ-hɛpiaŋ/ '2.PL.A-see' → [pɛ-ˈpiãŋ] (after further deletion of root vowel [ɛ])
f. /ɛ-hɛpiaŋ/ '2.SG.IMP-see' → [ɛ-ˈpiãŋ]

The example in (46) shows that vowels other than /ɛ/ also trigger /h/-deletion.[8]

(46) *Root-initial /h/-deletion in morpheme boundary between /a/_/a/ in unstressed syllables*
/a-haimɛɾuŋ/ '1.SG.A-sharpen' → [a-imɛˈɾũŋ]

The /h/-deletion process is exclusive to unstressed syllables at morpheme boundaries. Example (47) shows that /h/ is not deleted when the onset of a stressed syllable between two identical vowels.

(47) *No /h/-deletion in morpheme boundary between identical vowels in stressed syllables*
a. /ɛ-hɛi/ '2.SG.IMP-clean' → [ɛ-ˈhɛi]
b. /ɔ-hɔ/ '3.A-go' → [ɔ-ˈhɔ]

Finally, root-initial /h/-deletion in Juma is exclusive to morpheme boundaries. Examples in (48) indicate that /h/ is not deleted in word- and morpheme-internal position, regardless of whether it appears between two identical vowels or not.

(48) a. [ɛ-pɛhɛˈʔɔ̃ŋ] '2.SG.IMP-cut (it) in half!'
b. [ɛhaiˈtɛ] 'be dubious'
c. [ɨˈhɛ=a] 'insides'
d. [aˈhɛ] 'person'
e. [ɨˈhɔk=a] 'caterpillar'
f. [vaˈhaj=ɐ]['tail'
g. [auˈhu=a] 'stomach'
h. [dʒ-ipi-uˈhu] 'i-be.black-AUG'
i. [i-dʒuv-aˈhĩm] 'i-be.yellow-VERY'

[7]There is one instance of deletion of /h/ that does not follow this rule: the /h/ in the verb *hun* 'come' is deleted when the verb combines with the theme applicative *rero-*, i.e., *rer-un* 'bring'. The final vowel of the applicative suffix, /ɔ/, is subject to the process of vowel deletion of non-identical vowels in morpheme boundary, which I described in §3.2.3.7.

[8]Other verbs subject to /h/-deletion include *hekan* 'look for' and *hendum* 'listen'.

3.2.3.9 /a/-epenthesis

Consonant clusters at a morpheme boundary generally trigger a process of /a/-epenthesis. An example of /a/-epenthesis is provided in (49). Epenthetic /a/ are indicated in bold in relevant examples.

(49) */a/-epenthesis in consonant clusters at morpheme boundaries*
/aviɾ-vɛɾ=a/ 'scratch-NOM.PST=NMLZ' → [aviɾ**a**-vɛɾ=a] 'scratch (noun)'

The exception is a consonant cluster that involves the glottal stop /ʔ/, as I described in the following section.

3.2.3.10 Glottal stop metathesis

The glottal stop in the diminutive suffix /-ʔi/ undergoes metathesis with the immediately preceding consonant of a consonant-final morpheme. I provide examples of this process in (50).

(50) a. /pikɨɾ-ʔi=a/ 'be.young-DIM=NMLZ' → [pikɨʔɾia] '*lambari* fish (small fish species)'
b. /akɨkɨ+putaŋg-ʔi=a/ → [a.kɨ.kɨ+pu.taʔ.**ŋg**i.a] 'cup'
c. /ajav-ʔi/ → [-a.jaʔ-vi] '(AGAIN-DIM)'

Other instances of glottal stop metathesis are found in words such as *akyky[ʔɾ]ia* 'Caica parrot', *juparapa[ʔɾ]ia* 'pineapple', *ayve[ʔɾ]ia* 'be half full', *embyɾe[ʔɾ]ia* 'half', and *ipo[ʔɾ]ia* 'be orphan'.

3.2.3.11 Portuguese borrowings

Kawahíva speakers adapt Portuguese borrowing to Kawahíva phonology. These adaptations include phonological adaptation and syllable reduction. For instance, the Portuguese alveolar fricative [s] is adapted to its homorganic stop counterpart [t] in Kawahíva, as in the examples below.

(51) Portuguese → Kawahíva
a. [sãndaliɐ] → [tãndaliɐ] 'flip-flop'
b. [sɔ] → [tɔ] 'only'

Other accommodations involve nasal sounds in Portuguese, which follow the Kawahíva rules for this natural class. For instance, speakers produce words containing oral vowels in Portuguese as nasalized due to the presence of a nasalized vowel interpreted as an underlying nasal. This pattern follows from the Kawahíva nasal harmony process described in section §3.2.3.2, where underlying nasal vowels trigger nasal spreading leftward.

(52) Portuguese → Kawahíva
[ɔRgãnizaR] → [ɔ̃gãniza] 'set up'

A further nasality-related adaptation is observed in Portuguese words with a nasal stop-vowel sequence, where the vowel is oral. Speakers interpret the vowel as underlyingly oral and post-oralize the nasal stop, as in (53).

(53) Portuguese → Kawahíva
[na] → [nda] 'in'

Finally, speakers delete word-internal coda consonants of Portuguese words that do not follow the Kawahíva restrictions to a word-internal coda, as I described in §3.2.3.10; recall that Kawahíva only allows a word-internal coda with a glottal stop [ʔ], but Portuguese lacks this sound.

(54) Portuguese → Kawahíva
a. [ɔspitaʊ] → [ɔpita-uhu-a] (hospital-AUG=NMLZ) 'hospital'
b. [wɛzlɛɪ] → [ɛliɪ] 'Wesley'

3.3 Nominals and the noun phrase

In this section, I provide an overview of nominals in Kawahíva. It covers nouns §3.3.1 and their properties, which include number expression §3.3.2, nominal temporal markers §3.3.3, nominal constituent negation §3.3.4, and the privative §3.3.5. Following that, I describe the elements of the noun phrase, including demonstratives §3.3.6, quantifiers §3.3.7, and possession §3.3.8. Finally, §3.3.9 describes the pronouns and definiteness §3.3.10.

3.3.1 Nouns

Kawahíva nouns consist of an acategorial root and the enclitic *=a*, as shown for a common noun in (55). Proper names, too, appear with the enclitic *=a*. Examples of proper names with the enclitic are provided in (56).[9] Proper names, however, do not bear *=a* when used as a vocative, as in (57).

(55) tajahu=a
peccary=NMLZ
'A/the peccary'

(56) a. **Aviv=a**=ga pamẽ ga ho-i, 'i:ko.
Avip=NMLZ=3.SG.MASC together 3.SG.MASC go-i DIR.EVID
'He was together with Avip.' (Juma: Text)
b. **Jaru=a** jãnhã=mẽ.
Jaru.river=NMLZ up=to
'Up the Jaru river.' (Juma: Text)

[9] Sometimes, speakers will drop the enclitic in Portuguese proper names.

c. **Alto Jaru=a** pupe ki nhãnde ruv-i, ‘i:te:ki nũmĩa.
Alto Jaru=NMLZ inside PST 1.INCL stay DIR.EVID FRUST
‘To the Alto Jaru (basin) we came.’ (Jupaú: Text)

d. **Vĩntʃiũ=a** uvi ore pira i-mõkã’ẽ-mbyr=a erur-i avo
21.village=NMLZ from 1.EXCL fish i-roast-WH.OBJ.I=NMLZ bring-i here
Jumã=ipe.
Jumã.village=in
‘(It was) from 21 (village) that we brought fish that was roasted here, to the Juma village.’ (Juma: Text)

(57) **Avip**, e-run yhya ore=ve.
Avip, 2.SG.IMP water 1.EXCL=for
‘Avip, bring water for us.’ (Jupaú: Common in everyday speech)

I characterize *=a* as a ‘syntactic’ nominalizer in Kawahíva. That is, it derives nouns from acategorial roots. In this capacity, it is different from the more usual ‘nomenclative’ function of nominalizers – to create names for things in the world by converting members of other word classes into nouns; see Kastovsky (1985:225) and (Toosarvandani 2010:91). Instead, *=a* in Kawahíva satisfies the syntactic requirement that acategorial roots must bear it to function as a noun. This requirement is exemplified in (58): the absence of *=a* on words in the syntactic position of a noun is ungrammatical, such as heads of possessive phrases and objects of a transitive clause.[10]

(58) a. javatxing=a ra’yr=*(a)
dog=NMLZ offspring=NMLZ
‘A/the dog’s puppy.’

b. A-hepiang ki jie javatxing=a ra’yr=*(a) ko.
1.SG.A-see PST 1.SG dog=NMLZ offspring=NMLZ REAL
‘I saw the a/the dog’s puppy.’

Evidence for *=a*’s nominal categorizing or ‘nominalizing’ derivation of acategorial roots comes from the fact that when the same word lacks *=a*, it behaves as an existential predicate. For example, in (59a), the root *kapyr* ‘leaf, money’ functions as a predicate ‘to have money’, and bears predicate negation, which verbs like *’u* ‘eat’ also bear, as in (59b). In contrast, the form *kapyr* can’t appear with sentential negation in combination with *=a*, as seen in (59c).

(59) a. Nd-a-kapyr-i jie.
NEG-1.SG.A-leaf-NEG 1.SG
‘I do not have money.’ (Elicit)

b. Nd-a-’u-i ki jie pira.
NEG-1.SG.A-eat-NEG PST 1.SG fish
‘I did not eat fish.’ (Elicit)

[10]For alternative analyses of *=a*, see Rodrigues (2001) for a case marker and Queixalós (2006) for a determiner account of this enclitic.

c. *Nd-a-kapyr=a-i jie.
NEG-1.SG.A-leaf=NMLZ-NEG 1.SG
'Intended: I do not have money.' (Elicit)

Nevertheless, nominalizing *=a* also exhibits a 'nomenclative' function, in that it forms nouns from postpositions, number words, verbs, and clauses (clausal nominalizations). This is demonstrated in examples in (60) for postpositions (60a) and number words (60b), which must bear the nominalizer to function as the complement of a postposition and the subject of an intransitive verb, respectively.

(60) a. **Pe=*(a)** uvi gã ho-i.
to=NMLZ from 3.PL go-i
'From there they went.' (Juma: Text)

b. **A'ipe=*(a)** i-nhãn-i o-vo-vo.
one=NMLZ i-run-i 3.COR-go-VO
'The other (one) ran away and went.' (Juma: Text)

The nominal status of deverbal nouns can be diagnosed based on their nominal morphosyntax when in combination with *=a*. The verb *karuvan* 'be sick' in (61a), for instance, can be used as the argument of a verb, in (61b), or the complement of a postposition, as shown in (61c), with *=a*.

(61) a. **Karuvan** ga 'up=a.
be.sick 3.SG.MASC be.laying.down=NMLZ
'He is sick, lying down.' (Juma: Elicit)

b. I-poju **karuvar=a** ore=ve.
i-be.scared be.sick=NMLZ 1.EXCL=to
'Diseases are scary to us.' (Juma: Text)

c. Oro-kyji ore jup=a **karuvar=a** uvi.
1.EXCL.A-be.afraid 1.EXCL be.laying.down=NMLZ be.sick=NMLZ from
'We are afraid of diseases, while lying down.' (Juma: Text)

Additionally, the enclitic *=a* nominalizes clauses, as in (62), which can serve as verbal complements for clause-embedding verbs, such as *hepiang* 'see' and *kwaham* 'know'.

(62) a. A-hepia ji [Puré=ga piauhua, i-pyhyg=a].
1.SG.A-see 1.SG Puré=3.SG.MASC *piauhua*.fish i-catch=NMLZ
'I saw Puré caught a *piauhua* fish.' (Juma: Elicit)

b. A-kwaham ji [Puré=ga piauhua, i-pyhyg=a].
1.SG.A-know 1.SG Puré=3.SG.MASC *piauhua*.fish i-catch=NMLZ
'I know Puré caught a *piauhua* fish.' (Juma: Elicit)

The analysis of *=a* as the nominalizer in Kawahíva stands in contrast to the set of morphemes typically analyzed as nominalizers in Tupí-Guaraní languages (Jensen 1999:159-160). In Kawahíva, the cognates of these morphemes are those in Table 3.16.

Juma	Jupaú	Gloss
-har		transitive subjects
-hav & *-var*		postpositional obliques
rembi- & *-pyr*	*rimbi-* & *-pyr*	objects
-va'e		intransitive subjects

Table 3.16: Kawahíva cognates of Tupí-Guaraní nominalizers.

Here, I do not extend the nominalizing analysis to the Kawahíva morphemes in Table 3.16. In Kawahíva, these morphemes always co-occur with the nominalizer *=a* in relative clauses, as I describe in §3.7.11. Therefore, it would be redundant to call them nominalizers. In work in progress, I propose that they are better understood as *wh*-agreement markers in Kawahíva (Dos Santos In prep.), i.e., markers of a special form of agreement cross-referencing an extracted phrase (Deal 2016:170).[11] In particular, they cross-reference relative clause heads in Kawahíva.

In addition to the nominalizer *=a*, nouns referring to living humans (and God) combine with the enclitics *ga* '3.SG.MASC', *hẽa* '3.SG.FEM', and *gã* '3.PL', as shown in (63). However, nouns referring to a deceased human referent cannot bear these enclitics, as shown by (63e). Instead, they must bear the suffix *-va'e*, as in (64).

(63) a. kwemba'e=a=ga
man=NMLZ=3.SG.MASC
'Man'

b. kunhã=a=hẽa
woman=NMLZ=3.SG.FEM
'Woman'

c. tapy'ynh=a=gã
non.indigenous=NMLZ=3.PL
'Non-indigenous (people)'

d. tupan=a=ga
god=NMLZ=3.SG.MASC
'God'

e. jajia=(*__hẽa__)
deceased.aunt=3.SG.FEM
'Deceased aunt.' (Juma: Elicit)

(64) Depois **Dudu-va'e=a** ho-i mãipo.
after.that Dudu-va'e=a go-i WIT.DISTANT.PST
'After that, the late Dudu went.' (Jupaú: Text)

[11] This type of agreement was first described for the Austronesian languages Chamorro and Palauan (Chung and Georgopoulos 1988; Chung 1994).

3.3.2 Number and numerals

Kawahíva nouns are unspecified for number. Speakers use numerals and quantifiers to express number and quantity distinctions with respect to nouns. Numerals and quantificational elements are provided in Table 3.17. I describe quantifiers in §3.3.7. Examples in (65) show that numerals may or may not appear immediately adjacent to the noun.

Table 3.17: Number words and quantifiers.

Morpheme	Gloss
aipe'i	'one, once'
mõkõi	'two, twice'
mboapyn (Juma only)	'three, three times'
mõkõi katu katu (Juma only)	'four, four times'
e'yi	'(exact) many'
koi'i	'(inexact) many'

(65) a. Aipe'i pira ji i-pyhyg-i.
one fish 1.SG i-catch-i
'I caught one fish.' (Juma: Elicit)

b. Tapi'ira ji a-juka aipe'i.
tapir 1.SG 1.SG.A-kill one
'I killed one tapir.' (Juma: Elicit)

3.3.3 Nominal temporal markers

Kawahíva exhibits a two-way temporal distinction in nouns, which distinguishes between nominal past and nominal future. The two exponents of these categories are provided in Table 3.18.[12] The description presented here is only preliminary.

Table 3.18: Nominal temporal morphology.

Juma	Jupaú	Gloss
-ver ~ -kwer	*-ver ~ -kwer*	nominal past
-ram	*-ham*	nominal future

The nominal past suffix is used when the time at which the property denoted by the nominal is true of the individual prior to the utterance time. Examples of the nominal past suffix are provided in (66), including a non-deverbal noun (66b) and a deverbal noun in (66c).[13]

[12]The allomorphs for the nominal past suffix, *-ver ~ -kwer*, are lexically conditioned.

[13]Other deverbal nouns with the nominal past morpheme include *jiu'uavera* 'bite (noun)' (a nominalization of *u'u* 'bite'), *kokweruhua* 'old chacra/field', *aviravera* 'scratch (noun)' (a nominalization of *avin* 'scratch (verb)'), *ayvera* 'stump', *'aravera* 'birthday' (a nominalization of *'an* 'fall naturally, be born'), *jatavera* 'feces' (a nominalization of *ata* 'walk, defecate'), *tatara'uvera* 'torch', *imõpuavera* 'a hole made by a bullet' (a nominalization of *mõpu* 'shoot'), *ipirogavera* 'peeled skin (as in snakes)' (a nominalization of *pirong* 'peel'), and *rupã-rupãvera* 'hit' (a nominalization of the reduplicated form *nupã* 'hit (verb)').

(66) a. Ga=rembiriko=a=hẽa.
3.SG.MASC=partner=NMLZ=3.SG.FEM
'His wife.' (Juma: Common in everyday speech)

b. Ga=rembiriko-**kwer**=a=hẽa jie a-reko-ajam.
3.SG.MASC=partner-NOM.FUT=NMLZ=3.SG.FEM 1.SG 1.SG.A-marry-AGAIN
'His ex-wife, I married (her).' (Juma: Elicit)

c. Ore=ho-**aver**=a rupi pe ho-i.
1.EXCL=go-NOM.FUT=NMLZ through 2.PL go-I
'Through our going, you (PL.) went.' (Jupaú: Elicit)

In contrast, the nominal future suffix is used when the time at which the property denoted by the nominal will be true of the individual after the utterance time. On nouns, I have only identified this suffix with *rimbiko* 'partner' thus far, as shown in (67).

(67) Ji=rimbiriko-**ram**=a ki txiro o-ko ko.
1.SG=partner-NOM.FUT=NMLZ PST DEM:MED:ROUND 3.A-be REAL
'That is my future partner.' (Juma: Elicit)

Most Tupí-Guaraní languages exhibit cognate forms of the Kawahíva temporal markers, including the best-studied languages for this phenomenon, Paraguayan Guaraní (Tonhauser 2006, 2007) and Nheengatu (Cruz 2016). However, the semantic analysis of the cognate morphemes *-kue* and *-rã* of the most-studied language of the family in this respect, Paraguayan Guaraní, is divided between two positions: one which considers the temporal markers to be tense (Nordlinger and Sadler 2004) and one which rejects the tense analysis and instead analyzes these morphemes belonging to a distinct category (Tonhauser 2008). The argument that nominal temporal morphemes are not tense markers in Paraguayan Guaraní is based on fact that they do not behave like verbal tense markers: for instance, these morphemes may co-occur, unlike tense markers cross-linguistically. In Kawahíva, too, these morphemes co-occur.[14] Further systematic investigation of these morphemes in Kawahíva is needed to understand the similarity of their semantics with their cognates in Paraguayan Guaraní.

(68) Ji=rimbiriko-**ram-akwer**=a=ga txiro o-ko.
1.SG=partner-NOM.FUT=NOM.PST=NMLZ=3.SG.MASC DEM:MED:ROUND 3.A-be
'That is my promised partner.' (Juma: Elicit)

Finally, nominal temporal morphemes are also used for the expression of tense in nominalized dependent clauses, including complement clauses §3.7.9, adjunct clauses §3.7.10, and relative clauses §3.7.11.

[14]The example in (68) is the only instance of co-occurrence found thus far. Interestingly, Tonhauser (2007:859) notes the only order allowed in co-occurrence between the cognate morphemes in Paraguayan Guaraní is *-kue* (…) *-rã*, the opposite morpheme order in (68).

3.3.4 Constituent negation

Kawahíva noun phrases can be negated with the enclitic *=rõi* (~ *=nõi*). The same morpheme is also used with other phrases, including pronouns, as in (70a), and postpositional phrases, as in (70b).

(69) a. O-kyn **amãn=a** gã=pe kavyr-ipe.
3.A-rain rain=NMLZ 3.PL=to jungle-in
'It rained rain on them in the jungle.' (Juma: Elicit)

b. A'ero **amãn=a=rõi** tũhũi, 'i:te:ki:ko.
then rain=NMLZ=NOM.NEG INDEED DIR.EVID
'Then it was not rain indeed.' (Juma: Text)

(70) a. **Ji=rõi**, irãmbihuteteva'ea.
1.SG=NEG Big.Ears
'(It was) not me, Big Ears.' (Juma: Text)

b. **Korupi=rõi** tũhũi te, ga rur-i auhu, 'i ki gã
through.here=NEG INDEED REALLY 3.SG.MASC come-i seem say PST 3.PL
ko=ra.
REAL=TODAY.PST
'It seems not through here that he came indeed, they said.' (Juma: Text)

3.3.5 Privative

Kawahíva also exhibits a privative morpheme *-e'ym* (in Juma, and *-i'im* in Jupaú), that derives a stem from a noun that expresses the lack or absence of the referent of the root (71).

(71) u'i-**'ym**=a
flour-NOM.NEG=NMLZ
'(A meal) without flour' (Juma: Text)

The privative *-e'ym* is also used for negation in dependent clauses, including complement clauses §3.7.9, adjunct clauses §3.7.10, and relative clauses §3.7.11

3.3.6 Demonstratives

Kawahíva demonstratives show a three-way deictic distinction, which distinguishes between proximal, medial, and distal. Each of these, in turn, shows a two- or three-way shape categorization of their referent, distinguishing between elongated, round, and flat. This system yields the rich paradigm presented in Table 3.19.

Table 3.19: Demonstratives.

	elongated	round	flat
proximal	*ʻãngã*	*koro*	*kohoa*
medial	*mĩhĩa*	*txiro*	*pehea*
distal	*mĩhĩa*	*txijia*	*txijia*

Adnominal demonstratives follow the noun. Examples of adnominal demonstratives with nouns referring to elongated referents are provided in (72). These demonstratives have also been found in combination with words for 'water barrel/tank'.

(72) a. 'yngu'a **'ãngã**
mortar DEM:PROX:ELONGATED
'This mortar'

b. 'yngu'a **mĩhĩa**
mortar DEM:MED:ELONGATED
'That (distant) mortar'

Examples of adnominal demonstratives with nouns referring to round referents are provided in (73). Other nouns that can combine with these demonstrative include *irãmutxinguhua* 'hen', *pira* 'fish', *tamĩapiruhua* 'cloth', *ĩjihava* 'hammock', *mitakwanha* 'bread', *mbiara* 'meat', and *jãipepoa* 'bowl'.

(73) a. Yvutua **koro**
ball DEM:PROX:ROUND
'This ball'

b. Yvutua **txiro**
ball DEM:MED:ROUND
'That ball'

c. Javatxinga **txijia**
puppy/dog DEM:DIST:ROUND/FLAT
'That (distant) puppy'

Examples of adnominal demonstratives with nouns referring to flat referents are given in (74). Other nouns that may combine with these demonstrative include *ytypejuhava* 'broom'.

(74) a. Yhyaruhavuhua **kohoa**
hose DEM:PROX:FLAT
'This hose'

b. Yhyaruhavuhua **pehea**
hose DEM:PROX:ROUND
'That hose'

c. Yhyaruhavuhua **txijia**
hose DEM:DIST:ROUND/FLAT
'That (distant) hose'

When they bear focus, demonstratives tend to occur in the left periphery of the clause. In such cases, they appear discontinuous from the noun, as in (75).

(75) **Txijia** ki **javatxinga** o-mõnhãn tapi'ira ko.
DEM:DIST:ROUND/FLAT PST dog 3.A-corral tapir REAL
'That dog corralled the tapir.'

Locative demonstrative pronouns also show a three-way deictic distinction, as presented in Table 3.20. Examples of adverbial demonstratives are provided in (76).

Table 3.20: Locative demonstratives.

	Gloss
avo	proximal
pero	medial
pevo	distal

(76) a. A-pota ji nde=reku=a **avo**.
1.SG.A-want 1.SG 2.SG=live=NMLZ LOC.DEM:PROX
'I wanted you to live here.' (Juma: Elicit)

b. Mãngãn rupava **pero** 'um?
who hammock LOC.DEM:MED be.horizontal
'Whose hammock (is) there, (positioned) horizontally?' (Juma: Elicit)

c. E-mbohaeko **pevo** oga=pype.
2.SG.IMP-hang LOC.DEM:DIST house=inside
'Hang (it) inside the house!' (Juma: Elicit)

Finally, the demonstratives in Table 3.19 are also used as pronouns (e.g., 'this', 'that').

(77) a. Mãngã apoa **txiro**?
who whatchamacallit DEM:MED:ROUND
'Whose thing is that?' (pointing at a spoon) (Juma: Elicit)

b. E-juka **kohoa**!
2.SG.IMP-kill DEM:PROX:FLAT
'Kill this!' (pointing at a hen) (Juma: Elicit)

3.3.7 Quantifiers

Quantificational semantics in Kawahíva (e.g., 'many', 'a few', 'all') is expressed via with prosodically independent words, provided in Table 3.21, and suffixes, described in §3.4.8.

Table 3.21: Quantifiers.

quantifier	gloss
e'yi	many (exact)
koi'i	many (inexact)

The quantifier *e'yi* 'exact many' is used when the speaker knows the exact quantity of items being referred to. It implies a specific and quantifiable count. Consider (78), where the speaker knew the exact number of cars that arrived in the village.

(78) A'ero, ki onhãruhua'ea **e'yi** vĩntʃiũa=mẽ tur-i oi'i.
then PST car MANY.EX 21=to come-i some.time.ago
'Then, many cars arrived in the 21 village.' (Juma: Elicit)

In contrast, the quantifier *koi'i* 'indefinite many' is used when speakers do not know the exact quantity. It implies a general sense of abundance without precise quantification. An example is the number of ticks on a tapir, as in example (79).

(79) Jatevuguhua **koi'i** tapi'ira rehe.
tick MANY.INEX tapir at
'(There are) many ticks on the tapir.' (Juma: Elicit)

The same noun can be combined with either quantifier. Consider the examples in (80) and (81). In (80a) and (81a), the speakers use *e'yi* 'exact many', since they know the number of fish they roasted and caught, while in (80b) and (81b), speakers use *koi'i* 'inexact many' since they do not know the number of fish during the summer, nor the number of *yryvihua* trees on the reserve.

The quantifiers can appear postnominally, as in (78-79), and prenominally, as in (80). They may also be discontinuous with the quantified noun, as in (81).

(80) a. A-mõka'ẽ ji **e'yi** pira.
1.SG.A-roast 1.SG MANY.EX fish
'I roasted many fish.' (Juma: Elicit)

b. Kwaripe **koi'i** pira.
summer MANY.INEX fish
'In the summer, (there are) many fishes.' (Juma: Elicit)

(81) a. **E'yi** ji pira, i-pyhyk-i.
MANY.EX 1.SG fish, i-catch-i
'It was many fish that I caught.' (Juma: Elicit)

b. Yryvihua pevo **koi'i**.
yryvihua.tree LOC.DEM:DIST MANY.INEX
'(There are) many *yryvihua* trees there.' (Juma: Elicit)

3.3.8 Possession

Possessive relations in Kawahíva are expressed by juxtaposing the possessor and possessed noun. Possessors occur to the left of the possessed noun. Examples of pronouns as possessors are provided in (82), while examples of nouns as possessors are provided in (83).

(82) a. E-piang **ji**=kyva!
2.SG.IMP-see 1.SG=lice
'See my lice!' (Juma: Elicit)

b. O-ji-ram **hẽa**=rupahama.
3.A-REFL-untie 3.SG.FEM=hammock
'Her hammock became untied.' (Juma: Elicit)

(83) a. **Tapi'ira rauhua** a-'u.
tapir insides 1.SG.A-eat
'(I'm) eating the tapir's insides.' (Juma: Elicit)

b. **Mayta=hẽa ra'yr=a=hẽa** ki o-ho ore=rupi ko.
Maytá=3.SG.FEM offspring=NMLZ=3.SG.FEM PST 3.A-go 1.EXCL=with REAL
'Maytá's daughter went with us.' (Juma: Text)

Kawahíva does not exhibit a canonical distinction between alienable and inalienable possession constructions (Aikhenvald 2012). However, there is evidence that nouns that we would expect, on the basis of cross-linguistic tendencies, to fall into alienable vs. alienable categories are treated differently in the grammar. In particular, inalienable nouns exhibit two root allomorphs, conditioned by the presence of an adjacent possessor, as in (84). In contrast, alienable nouns never show root allomorphy, as in (85).

(84) a. **ta'yr**=a
offspring=NMLZ
'(Someone's/something's) offspring' (Juma: Elicit)

b. nde=**ra'yr**=a=ga
2.SG=offspring=3.SG.MASC
'Your son' (Juma: Elicit)

(85) a. **tupahũ**=a
gun=NMLZ
'A gun' (Juma: Elicit)

b. nde=**tupahũ**=a
2.SG=gun=NMLZ
'Your gun' (Juma: Text)

I take this pattern to reflect the cross-linguistic distinction that possessors of inalienable nouns, but not possessors of alienable nouns, are syntactic complements of possessed nouns (Alexiadou 2003). This distinction immediately explains why only inalienable nouns undergo root allomorphy with an overt possessor. If root allomorphy is triggered only when the root and preceding word (usually its syntactic complement) are in the same phrase (as described in §3.2.3.5), then it follows that the possessor of the inalienable noun in (84) is a syntactic complement, but the possessor of the alienable noun in (85) is not.

3.3.9 Pronouns

Kawahíva has two sets of personal pronouns, non-focused and focused pronouns, which expresses person, number, clusivity, and gender. The first person plural pronouns draw a clusivity distinction, and the third person singular pronouns are further distinguished for gender. Table 3.22 shows the non-focused personal pronouns in Juma and Jupaú.

Table 3.22: Non-focused personal pronouns.

	Juma	Jupaú
1st singular	*ji*	
2nd singular	*nde*	
1st plural inclusive	*nhãnde*	
1st plural exclusive	*ore*	*ore ~ are*
2nd plural	*pẽ*	*pe*
3rd singular masc.	*ga*	*ka ~ ga*
3rd singular fem.	*hẽa*	
3rd plural	*gã*	
3rd mythical	*ju*	

Non-focused pronouns are not distinguished for case. In the examples (86-88), the 3rd person masculine pronoun form remains the same regardless of its grammatical function as a subject in (86), an object in (87), or the complement of a postposition in (88).

(86) V-epia ka moja ko.
3.A-see 3.SG.MASC snake REAL
'He saw a/the snake.'

(87) V-epia hẽa ka ko.
3.A-see 3.SG.FEM 3.SG.MASC REAL
'She saw him.'

(88) O-mõndo hẽa pira ka=pe ko.
3.A-give 3.SG.FEM fish 3.SG.MASC=to REAL
'She gave him fish.'

3rd person pronouns are generally used for living human referents; they cannot be used to refer to inanimate referents, such as *mbiara* 'meat' in (89a), nor animals, as *akara'ia* '*akara* fish' (89b).

(89) a. Aramẽ ki hẽa, ga='y=hẽa, mbiara rero-'ong-i,
after.that PST 3.SG.FEM 3.SG.MASC=mother=3.SG.FEM meat SOC.CAUS-take.out-i
(***ga/*hẽa**) i-mõndo-vo i-mõkup=a, 'i:te:ki:ko.
3.SG.MASC/3.SG.FEM i-put-VO i-boil=NMLZ DIR.EVID
'After that, she, his mother, took out the meat and put (it) to boil (it).' (Juma: Elicit)

b. A'ero ga akara'ia pyhyg-i. A'ero (***ga**/***hẽa**) i-hir-i 'yame'y ipe.
then 3.SG.MASC *akara*.fish catch-i then 3.SG.MASC/3.SG.FEM i-fall-i river.shore in
'Then he caught an *akara* fish. Then (it=fish) fell by the river shore.' (Juma: Text)

3rd person pronouns may be used to refer to non-living humans in personal stories and mythical entities in traditional stories, although these instances occur in limited contexts, including as part of a direct quotation and the fixed phrase *s/he said* after a direct quotation, as in (90a). Outside these contexts, these referents can be referred to with *a'ea* 'this one' (90b) or are null (90c). The latter two examples come from a speaker's recollection of the events the day before a Jupaú schoolteacher had been murdered.

(90) a. T-a-piang nde=kyva, apin. Po 'i!, 'i ky **ka** ko=ra.
OPT-1.SG.A-see 2.SG=lice father IRR say say PST 3.SG.MASC REAL=TODAY.PST
'Let me see your lice, Father. (She) speaks!, he said.' (Jupaú: Text)

b. **A'ea** te i-ho-i ipymbuhuva'ea rehe.
this.one REALLY i-go-i motorcycle at
'This one (=he) left on the motorcycle.' (Juma: Text)

c. Aramẽ ki ajiramẽ tur-i, 'i-te ki ko.
after.that PST later come-i say-REALLY PST REAL
'After that, later, (he) came back.' (Juma: Text)

Kawahíva also allows *pro*-drop of 1st and 2nd persons (91a-91b). *Pro*-drop of a 3rd person pronoun referring to a living human, however, is not permitted, as in (91c).

(91) a. Ere-'u po (nde)?
2.SG.A-eat IRR 2.SG
'Did you eat?' (Juma and Jupaú: Common in everyday speech)

b. A-'u-ipe (ji) ko.
1.SG.A-eat-ALREADY 1.SG REAL
'I already did.' (Juma and Jupaú: Common in everyday speech)

c. Nd-o-mõkã'e-i ki *(gã).
NEG-3.A-roast-NEG PST 3.PL
'(They) did not roast (it).' (Juma: Elicit)

Kawahíva has a pronoun *ju*, which is used for 3rd person referents. I have found this pronoun only in traditional stories, where it refers to mythical entities in a benefactive role, and bears the postposition *pe* 'to, for'.

(92) Kamĩnha po-i mikytã mõndo-vo **ju=pe**.
manioc/corn.beer make-i cake give-VO 3.mythical=to
'(He) made the manioc/corn beer and gave the cake to her.' (Jupaú: Text)

A phonologically similar set of pronouns is used when the referent of the pronoun is focused. Table 3.23 shows the complete list of the focused pronouns. This paradigm is generally formed based on the non-focused form (see Table 3.22) and the syllable *hV*, where the vowel V is either a copy of the immediately preceding vowel or shares its backness value. However, some pronouns have identical focused and non-focused forms, including the 1st person plural pronoun and the 3rd person feminine; the latter also bears an initial vowel /e/.

Table 3.23: Focused personal pronouns.

	Juma	Jupaú
1st singular	*jihe ~ jihi*	
2nd singular	*ndehe*	
1st plural inclusive	*nhãnde*	
1st plural exclusive	*ore*	*ore ~ are*
2nd plural	*põhẽ*	*pehe*
3rd singular masc.	*gaha*	*kaha ~ gaha*
3rd singular fem.	*ehẽa*	
3rd plural	*gãhã*	

The default position of focused pronouns is clause-initial, as in (93a), the same position occupied by interrogative phrases. Focused pronouns may also appear in non-initial clause position, as in exhaustive focus constructions, as in (93b), where it is followed by the particle *te'i* 'only', and in polar questions, as in (93c).

(93) a. **Jihe** ki tevir-ete a-ko ko.
1.SG.FOC PST behind-REALLY 1.SG.A-be REAL
'It was me who was really behind.' (Juma: Text)

b. A'ero po **gaha** te'i i-mõkwatxijar-i nãhẽ, a-'i ki ji nõmĩa.
then IRR 3.SG.MASC.FOC only i-write-i POT 1.SG.A-say PST 1.SG FRUST
'Then only he would write it (the Kawahíva language), I thought.' (Juma: Text)

c. Nde-pe-mõpu-i-uhu **põhẽ**?, 'i ki ka ko.
NEG-2.PL.A-shoot-NEG-AUG 2.PL.FOC say PST 3.SG.MASC REAL
'Didn't y'all shoot (it)? He asked.' (Juma: Text)

3.3.10 Definiteness

Determiners follow nouns and encode definiteness. Definite determiners are restricted to living human referents (and God). Definite determiners show a two-way number distinction (singular and plural) and a two-way gender distinction in the singular (feminine and masculine). An indefinite determiner is used for non-specific indefinite readings. Table 3.24 lists the determiners in Kawahíva.

Kawahíva also allows nouns to appear without a determiner (i.e., bare). Bare nouns in Kawahíva can be used for specific indefinite, generic, and kind readings.

Table 3.24: Determiners.

determiner	gloss
=*ga*	'SG.MASC definite determiner'
=*hẽa*	'SG.FEM definite determiner'
=*gã*	'PL definite determiner'
=*mõ*	'indefinite determiner'

3.3.10.1 Definite determiners

Definite determiners in Kawahíva draw on the paradigm of independent 3rd person pronouns. They make a two-way number distinction, and in the singular, a two-way gender distinction: =*ga* is the singular masculine form, and =*hẽa* is the feminine singular form.

(94) a. Ji i-mõndo-i, ga=ruva=**ga** pe, 'i:te:ki:ko.
1.SG i-give-i 3.SG.MASC=father=3.SG.FEM to DIR.EVID
'I gave (it) to his father.' (Juma: Text)

b. Ore pyta-i tapy'ynh=a=**hẽa**.
1.EXCL stop-i non.indigenous=3.SG.FEM
'We stopped to bring the female non-indigenous people.' (Juma: Text)

There is an additional plural determiner =*gã*, which is gender-invariant. This determiner may also combine with a living human proper noun and function as an associative plural 'X and the group associated with X'.

(95) a. E-'i-ramõ-te ki jimyjive=**gã** ipi.
i-say-LIKE-REALLY PST elder=3.PL always
'The elders used to say thus always.' (Juma: Text)

b. Davi=**gã** ho-i txikyty.
Davi=3.PL go-i ahead
'Davi's group went ahead.' (Juma: Text)

These enclitics have several properties that are characteristic of definite determiners. For instance, nominals marked with them cannot appear in existential contexts (Dawson and Jenks 2023).[15]

(96) a. Nd-o-ko-i akwemba'ea=(***ga**/***gã**) jahya pyri.
NEG-3.A-be-NEG man=3.SG.MASC/3.PL moon near
'There is no man on the moon.' (Juma: Elicit)

[15]Other diagnostics are not available in Kawahíva, given the restriction that these determiners combine with nouns whose referents are living humans. Therefore, it is not possible to use them to describe parts of previously-introduced wholes (e.g., I bought a bicycle for my son, but the tires were worn out).

b. N-o-ko-i 21=mẽ tapy'ynha=(*__gã__).
NEG-3.A-be-NEG 21.village=to non.indigenous=3.PL
'There are no non-indigenous people in the 21 (village).' (Juma: Elicit)

The enclitic determiner *=ga* '3rd singular masculine' is used to refer to a unique referent like *Tupan* 'God'.

(97) Mãrãn-ũhũ po txia txirove, Tupan=a=__ga__.
why-AUG IRR DEM now God=NMLZ=3.SG.MASC
'Why (is) this now, God.' (Lamenting about the COVID situation) (Juma: Text)

Nouns, in combination with the enclitic determiners, pass the 'test of consistency' or 'law of contradiction' (Löbner 2002). According to it, a noun referring to an individual cannot both hold and not hold a given property. This is unexpected, as definite articles create arguments that refer to individuals and such nouns obey the law of contradiction. For example, in (98), *kunhã* 'woman' must refer to the same individual in each instance. Therefore, the enclitic must be a definite determiner.

(98) *Kunhã=__hẽa__ i-katu. Kunhã=__hẽa__ nd-i-katu-i.
woman=3.SG.FEM i-be.pretty woman=3.SG.FEM NEG-i-be.pretty-NEG
'The woman is pretty, and the woman is not pretty.' (Juma: Elicit)

Finally, the enclitic determiners can be used for generic readings (99a). However, they are not required in these contexts (99b).

(99) a. Tapy'ynh=a=__gã__ a-kate'ym matera rehe.
non.indigenous=NMLZ=3.PL 3.A-be.miserly thing at
'Non-indigenous people are miserly about things.' (Juma: Elicit)

b. __Akwemba'e=a__ ki akwemba'e=a pomẽ o-ingwarai.
boy=NMLZ PST boy=NMLZ together 3.A-play
'Boys play with boys.' (Juma: Text)

3.3.10.2 Bare nouns

Most Kawahíva nouns appear without a determiner (i.e., bare). Bare nouns are not construed as definite. For instance, they do not pass the law of contradiction.

(100) a. __Tapi'ira__ ehe. __Tapi'ira__ nda-he-i.
tapir be.tasty tapir NEG-be.tasty-NEG
'There is tapir that is delicious, there is tapir that is not delicious.' (Juma: Elicit)

b. __Kunhã__ i-katu. __Kunhã__ nd-i-katu-i.
woman i-be.pretty woman NEG-i-be.pretty-NEG
'There is one woman who is pretty, and there is a woman who is not pretty.' (Juma: Elicit)

Bare nouns can be used in cases of specific indefinite reference, where the speaker has a particular referent in mind, as in (101a), as well as cases of non-specific reference, where the speaker does not have a particular referent in mind, as in (101b).

(101) a. Are reko-i **ehira** vo-vo. Yvahua, o nome dele.
1.EXCL be-i honey eat-VO *yvahua*.honey the name of.it
'We were eating honey. *Yvahua* honey, the name of it.' (Jupaú: Text)

b. Ore ho-i **iramburava** rek=a oro-ko-vo, 'i:te:ki:ko rõmĩa.
1.EXCL go-i agouti look.for=NMLZ 1.EXCL.COR-be-VO DIR.EVID FRUST
Nd-or-epiak-i ki ore ko.
NEG-1.EXCL-see-NEG PST 1.EXCL REAL
'We had looked for an agouti. We did not see (any).' (Juma: Text)

Bare nouns in Kawahíva can have a number of readings, including wide scope indefinite, generic, and kind readings. In wide scope readings, bare nouns must be interpreted before scope-sensitive elements in linear order, like negation (102) and *-ajam* 'again' (103). For example, the bare noun *javatxinga* 'dog' follows negation in (102a). However, the only available interpretation is one where the bare noun takes scope before the negation operator. The same wide scope pattern with negation is observed when the bare noun precedes it (102b). Additionally, bare nouns must also have a wide scope interpretation relative to the operator *-ajam* 'again' (103).

(102) a. Nd-o-nhãr-i **javatxinga**.
NEG-3.A-run-NEG dog
'✓Context 1: There is one dog that did not run.
'✗Context 2: No dog ran. (Juma: Elicit)

b. Aipe'i **javatxinga** nd-o-i'ing-i.
one dog NEG-3.A-bark-NEG
'✓Context 1: There is one dog that did not bark.
'✗Context 2: No dog barked. (Juma: Elicit)

(103) V-epiag-aja-(a)jam ki gã **yvujara** oi'i.
3.A-see-RED-AGAIN PST 3.PL *yvujara*.worm some.time.ago
'✓ Context 1: They saw the (same) *yvujara* worm.
'✗ Context 2: They each saw a (different) *yvujara* worm.' (Juma: Elicit)

Bare nouns obligatorily take narrow scope relative to propositional attitude verbs like *pota* 'want, desire' (104) and in conditional clauses (105). In other words, in narrow scope readings, the noun is interpreted within the scope of the scope-sensitive element in linear order.

(104) O-pota-ete ga **tapy'ynha** kwarupi.
3.A-want-REALLY 3.SG.MASC non.indigenous today
'✓ Context 1: He wants (to date) any non-indigenous person today.
'✗ Context 2: He wants (to date) a specific non-indigenous person today.' (Juma: Elicit)

(105) Agui=ga o-je'o [**javatxinga** ekir=a=mẽ] nãhẽ.
Agui=3.SG.MASC 3.A-cry dog enter=NMLZ=IF POT
'✓ Context 1: Agui will cry if any dog enters (the house).
'✗ Context 2: Agui will cry if a specific dog enters (the house).' (Juma: Elicit)

Finally, a bare noun can also be used for generic and kind readings.

(106) a. Nda-vy-i **kavarapea**.
NEG-blood-NEG cockroach
'Cockroaches do not have blood.' (Juma: Elicit)

b. Mbahira v-apo **kunhã**.
Mbahira 3.A-make woman
'Mbahira created women.' (Jupaú: Text)

3.3.10.3 Indefinite determiner

Kawahíva uses the determiner *=mõ* for nonspecific indefinite contexts where the speaker does not have a particular referent in mind.

(107) Nd-a-hepiag-i ki ji pira=**mõ**.
NEG-1.SG.A-see-NEG PST 1.SG fish=NON.SPEC.INDEF
'I did not see any fish.' (Juma: Text)

Nonspecific indefinites must take narrow scope with respect to operators like negation.

(108) Nd-o-nhãr-i javatxinga=**mõ**.
NEG-3.A-run/walk-NEG dog=NON.SPEC.INDF
'✓No dog ran (after the peccary).'
'✗(Among many dogs) One dog did not run (after the peccary).' (Juma: Elicit)

Nonspecific indefinites also obligatorily take narrow scope in regard to verbs like *pota* 'want, desire' and in conditional clauses; the latter suggests that nonspecific indefinites are sensitive to "scope islands".

(109) O-pota-ete ga tapy'ynha=**mõ** kwarupi.
3.A-want-REALLY 3.SG.MASC non.indigenous=NON.SPEC.INDEF today
'✓ Context 1: He wants (to date) any non-indigenous person today.
'✗ Context 2: He wants (to date) a specific non-indigenous person today.' (Juma: Elicit)

(110) Agui=ga o-je'o [**javatxinga=mõ** ekir=a=mẽ] nãhẽ.
Agui=3.SG.MASC 3.A-cry dog=NON.SPEC.INDF enter=NMLZ=IF POT
'✓ Context 1: Agui will cry if any dog enters (the house).
'✗ Context 2: Agui will cry if a specific dog enters (the house).' (Juma: Elicit)

Nouns combined with *=mõ* show a behavior similar to Negative Polarity Items (111). For instance, while they occur in negative polarity environments (111a), they can't occur in a positive polarity environment (111b). Additionally, they occur in polar questions (112a), and cannot occur in a content question like (112b), which presupposes a specific lighter.

(111) a. Nd-a-hepiag-i jie matera=**mõ**.
NEG-1.SG.A-see-NEG 1.SG thing=NON.SPEC.INDF
'I saw nothing.' (Juma: Elicit)

b. *A-hepia jie matera=**mõ**?
1.SG.A-see 1.SG thing=NON.SPEC.INDF
Intended: 'I saw nothing.' (Juma: Elicit)

(112) a. Pe=reko pe tata'ia=**mõ**?
2.PL.B-have 2.PL lighter=NON.SPEC.INDF
'Do you have a lighter?' (Juma: Elicit)

b. Mãhã tata'ia=(***mõ**)?
where lighter=NON.SPEC.INDF
Intended: 'Where is a lighter?' (Juma: Elicit)

Finally, *=mõ* may also combine with the 3rd person plural pronoun *gã*.

(113) Nda=ji=repiag-i **gã=mõ**.
NEG-1.SG.B=see-NEG 3.PL=NON.SPEC.INDF
'Nobody saw me.' (Juma: Elicit)

3.4 Verbs

In this section, I describe basic morphological properties of the Kawahíva verb.

The verb consists of morphemes that can be grouped into a large number of distinct categories by distributional tests. In this discussion, I distinguish between morphology that appears to the left and the right of the root.

On the left of the verb root, as depicted in (114), in order, are: negation (the prefix of the discontinuous standard negation morpheme *n(da)-...-i*), argument agreement, the reciprocal, the causative, the reflexive, and the reduplicant.

(114) NEG - {SUBJ AGR|OBJ AGR|IMP} - RECP - {CAUS} - REFL - RED - root

On the right side of the verb root are, as shown in order in (115), are: the transitive applicative, agreement (including imperative agreement), aspect, and negation (the suffix of the discontinuous standard negation morpheme *n(da)-...-i*).

(115) root - APPL - ASPECT - NEG - ASPECT

Setting aside the negation prefix, we can group the leftmost morphemes listed above into a category of argument markers and valency change markers. Excluding the applicative, we can group the rightmost morphemes into a category of inflectional affixes. This simplified picture of the verb morphology is shown in (116):

(116) argument marking - valency change - root - inflectional suffixes

The discussion below generally moves through the verbal morphology from left to right, structured by the breakdown into three zones in (116).

3.4.1 Agreement

The Juma and Jupaú dialects exhibit a pattern of head-marking of arguments in matrix clauses, in which matrix verbs bear markers that index the person, number, and clusivity (for the 1st person plural) of either the subject or object. Whether the subject or object properties are marked on the verb depends on a person hierarchy 1>2>3 in the Juma and Jupaú dialects (Dos Santos 2021a, 2023a). The grammatical dependency relation between argument and verb is expressed by two sets of verbal indexes, which are shown in Table 3.25.[16] Following Dos Santos (2023a), I refer to the subject markers as SET A, and the object markers as SET B;[17] I also treat SET B markers as pronominal agreement and SET A as grammatical agreement markers.

Table 3.25: Person/number indices.

	SET A	SET B
1st person singular	a-	ji=
2nd person singular	ere-	nde=
3rd person singular and plural	v- ∼ o-	–
1st person inclusive	txi-	nhãnde=
1st person exclusive	oro-	ore=
2nd person plural	pe-	pe=

Table 3.26 summarizes all the possible combinations of subjects and objects in a matrix clause with a transitive verb, which I discuss next.

[16] The alternation *v-* ∼ *o-* is phonologically-conditioned: *v-* is used with vowel-initial verbs, and *o-* is used with consonant-initial verbs.

[17] The two sets are also commonly referred to as Set I and II after Jensen (1990)'s comparative work.

Table 3.26: Possible combinations of arguments in transitive sentences and selected prefixes.

↓ S O →	1.sg	1.incl	1.excl	2.sg	2.pl	3.sg./pl
1st singular				a-	a-	a-
1st plural inclusive						txi-
1st plural exclusive				oro-	oro-	oro-
2nd singular	ji=		ore=			ere-
2nd plural	ji=		ore=			pe-
3rd singular/plural	ji=	nhãnde=	ore=	nde=	pe=	v- ~ o-

The SET A markers cross-reference the subject in transitive clauses, as shown in (117-125); these examples draw on the Juma dialect. The different surface forms of the verb 'see' (i.e., *hepia* ~ *epia*) are the result of regular morphophonological processes in Juma, including [h]-deletion (described in §3.2.3.8) and vowel deletion (described in §3.2.3.7).

(117) **A**hepia ki ji ga. (**1.SG** → 3)
a–hepia ki ji ga.
1.SG.A-see PST 1.SG 3.SG.MASC
'I saw him.'

(118) **Ere**pia ki nde ga. (**2.SG** → 3)
ere–hepia ki nde ga.
2.SG.A–see PST 2.SG 3.SG.MASC
'You (SG.) saw him.'

(119) **V**epia ki hẽa ga. (**3.MASC** → 3.FEM)
v–hepia ki hẽa ga.
3.A–see PST 3.SG.FEM 3.SG.MASC
'She saw him.'

(120) **V**epia ki ga hẽa. (**3.FEM** → 3.MASC)
v–hepia ki ga hẽa.
3.A–see PST 3.SG.MASC 3.SG.FEM
'He saw her.'

(121) **Pe**pia ki pe ga. (**2.PL** → 3)
pe–hepia ki pe ga.
2.PL.A–see PST 2.PL 3.SG.MASC
'You (PL.) saw him.'

(122) **Oro**hepia ki ore ga. (**1.EXCL** → 3)
oro–hepia ki ore ga.
1.EXCL.A–see PST 1.EXCL 3.SG.MASC
'We (EXCL.) saw him.'

(123) **Txi**pia ki nhãnde ga. (**1.INCL** → 3)
txi–hepia ki nhãnde ga.
1.INCL.A-see PST 1.INCL 3.SG.MASC
'We (INCL.) saw him.'

(124) **V**epia ki gã hẽa. (**3.PL** → 3.SG.)
v–hepia ki gã hẽa.
3.A–see PST 3.PL 3.SG.FEM
'They saw her.'

(125) **V**epia ki kwemba'ea javatxinga. (**NP** → NP)
v–hepia ki kwemba'ea javatxinga.
3.A–see PST man dog
'The man saw a/the dog.'

In contrast, SET B markers cross-reference the object in transitive clauses, which I illustrate in examples (126-130).

(126) **Ji**repia ki kwemba'eaga. (NP → **1.SG**)
ji=repia ki kwemba'ea=ga.
1.SG.B=see PST man=3.SG.MASC
'The man saw me.'

(127) **Nde**repia ki kwemba'eaga. (NP → **2.SG**)
nde=repia ki kwemba'ea=ga.
2.SG.B=see PST man=3.SG.MASC
'The man saw you (SG.).'

(128) **Pe**repia ki kwemba'eaga. (NP → **2.PL**)
pe=repia ki kwemba'ea=ga.
2.PL.B=see PST man=3.SG.MASC
'The man saw you (PL.).'

(129) **Ore**repia ki kwemba'eaga. (NP → **1.EXCL**)
ore=repia ki kwemba'ea=ga.
1.EXCL.B=see PST man=3.SG.MASC
'The man saw us (EXCL.).'

(130) **Nhãnde**repia ki kwemba'eaga. (NP → **1.INCL**)
nhãnde=repia ki kwemba'ea=ga.
1.INCL.B=see PST man=3.SG.MASC
'The man saw us (INCL.).'

Examples (131-137) are of transitive matrix clauses with local persons (i.e., 1st and 2nd). They support the claim that Juma and Jupaú exhibit a person hierarchy 1>2>3.

(131) **A**hepia ji nde. (**1.SG** → 2.SG)
a-hepia ji nde
1.SG.A=see 1.SG 2.SG
'I saw you (SG.).'

(132) **A**hepia ji pe. (**1.SG** → 2.PL)
a-hepia ji pe
1.SG.A=see 1.SG 2.PL
'I saw you (PL.).'

(133) **Ji**repia nde. (2.SG → **1.SG**)
ji=repia nde
1.SG.B=see 2.SG
'You (SG.) saw me.'

(134) **Ji**repia pe. (2.PL → **1.SG**)
ji=repia pe
1.SG.B=see 2.PL
'You (PL.) saw me.'

(135) **Ore**=repia nde. (2.SG → **1.EXCL**)
ore=repia nde
1.EXCL.B=see 2.SG
'You (SG.) saw us (EXCL.).'

(136) **Ore**=repia gã. (3.PL → **1.EXCL**)
ore=repia gã
1.EXCL.B=see 3.PL
'They saw us (EXCL.).'

(137) **Nhãnde**=repia gã. (3.PL → **1.INCL**)
nhãnde=repia gã
1.INCL.B=see 3.PL
'They saw us (INCL.).'

The Juma and Jupaú person hierarchy 1>2>3 contrasts with the pattern of person marking in local scenarios in the Karipuna and Parintintin dialects, which exhibit a 1/2>3 person hierarchy, a domain of dialectal variation previously addressed in Dos Santos (2021a). In particular, in the latter dialects, the verb bears markers that cross-reference both the subject and object in local person scenarios (e.g., 1 → 2) via the portmanteaux affixes *oro-* '1→2.SG' and *opo-* '1→2.PL' (Dos Santos 2021a). I illustrate these morphemes with the Karipuna examples in (138a-b).

(138) a. **Oro**–vepia.
1>2.SG–see
'I saw you (SG.).'

b. **Opo**-vepia.
1>2.PL–see
'I saw you (PL.).'

Intransitive verbs also take SET A and SET B markers. These verbs fall into two broad classes: the active class (e.g., *dance*, *jump*) and the stative class (e.g., *be pretty*, *be tired*). Examples of the active verb *kwam* 'dance' with SET A markers are provided in (139-146).

(139) **A**kwam ki ji.
a–kwam ki ji.
1.SG.A–dance PST 1.SG
'I danced.'

(140) **Ere**kwam ki nde.
ere–kwam ki nde.
2.SG.A–dance PST 2.SG
'You (SG.) danced.'

(141) **O**kwam ki kwemba'ea.
o–kwam ki kwemba'ea.
3.A–dance PST man
'A man danced.'

(142) **O**kwam ki hẽa.
o–kwam ki hẽa.
3.A–dance PST 3.SG.FEM
'She danced.'

(143) **Oro**kwam ki ore.
oro–kwam ki ore.
1.EXCL.A–dance PST 1.EXCL
'We (EXCL.) danced.'

(144) **Txi**kwam ki nhãnde.
txi–kwam ki nhãnde.
1.INCL.A–dance PST 1.INCL
'We (INCL.) danced.'

(145) **Pe**kwam ki pe.
pe–kwam ki pe.
2.PL.A–dance PST 2.PL
'You (PL.) danced.'

(146) **O**kwam ki kunhã=hẽa.
o–kwam ki kunhã=hẽa.
3.A–see PST woman=3.SG.FEM
'The woman danced.'

Intransitive stative verbs fall into two morphologically distinguishable subclasses: one subclass takes SET B markers (i.e., object agreement), and another appears with the prefix *i-* (~ *j-* [dʒ]).[18] The former subclass includes verbs like *tur ~ rur* 'come, arrive', *ten ~ ren* 'be seated', *tup ~ 'up* 'be in a horizontal position/be laid down', and *tuv ~ ruv* 'stay'. I provide one example of a stative verb with SET B marker in (147).

(147) **Ore**=rur–i tapy'ynha=gã pyri.
1.EXCL.B=come–i non.indigenous=3.PL by
'We came to town.' (Lit.: We came to the non-indigenous people) (Text: Juma)

Intransitive stative verbs marked with *i-* include most of the stative verbs, such as *katu* 'be pretty, well', *kwerai* 'be tired', *rovia* 'be happy', *rãite* 'be ugly', *ro'y* 'be cold', among others.

(148) **I**-kwerai ji.
i–be.tired 1.SG
'I am tired.' (Juma and Jupaú: Common in everyday speech)

3.4.2 Imperative

Imperatives are formed using a verb prefixed with a 2nd person agreement prefix. The agreement paradigm in imperatives is provided in Table 3.27; this paradigm is partially syncretic with the 2nd person subject marking of declarative clauses (see Table 3.25). Additionally, the implied 2nd person subject is null. An example of each imperative agreement marking in positive polarity contexts is provided in examples (149), and in negative polarity contexts, in examples (150).

Table 3.27: Verbal agreement paradigm in imperatives.

	singular	plural
positive polarity	e-	pe-
negative polarity	ere-	epe-

(149) a. **E**-mõmbe'u ji=ve e-piag-aver=a!
2.SG.IMP-tell 1.SG=to 2.SG.COR-see-NOM.PST=NMLZ
'Tell me (what) you (SG.) saw!' (Juma: Elicit)

b. Txuruguhua **pe**-apo kwarupi txi-'u!
bread 2.PL.IMP-make today 1.INCL.COR-eat
'You (PL.) make bread for us to eat!' (Juma: Elicit)

[18]The [i-] ~ *j-* [dʒ] allomorphy is phonologically-conditioned: [i-] appears with consonant-initial roots, whereas *j-* [dʒ] appears with vowel-initial roots.

(150) a. Nd-**ere**-'u-i karavaruhua, e-ja'yra rakuv=a=mẽ!
NEG-2.PL.IMP-eat-NEG agouti 2.SG.COR-offspring be.hot=NMLZ=WHEN
'Do you (SG.) not eat agouti when your child is feverish!'
(Juma: Elicit)

b. Nd-**epe**-'u-i karavaruhua, e-jakuv=a=mẽ!
NEG-2.PL.IMP-eat-NEG agouti 2.SG.COR-be.hot=NMLZ=WHEN
'Do you (PL.) not eat agouti when you are feverish!' (Juma: Elicit)

Imperatives are also required in some modal constructions, which I describe in §3.7.8.

3.4.3 Reciprocals

Kawahíva uses the prefix *jo-* as a reciprocal, which appears on verbs like *hepiang* 'see', *mõngyta* 'tell a story', and *mõi'ing* 'talk'. Two examples of the reciprocal prefix are provided in (151).

(151) a. O-**jo**-vepiang ki gã ko.
3.A-RECP-see PST 3.PL REAL
'They met.' (Lit.: they saw each other) (Juma: Elicit)

b. O-**jo**-mõngyta gã 'up=a.
3.A-RECP-tell.story 3.PL be.horizontally=NMLZ
'They are telling each other stories, (positioned) lying.' (Juma: Elicit)

The reciprocal *jo-* is also lexicalized on a number of verbs that no longer carry transparently reciprocal meaning (152-153), and others that may no longer be decomposed into their component parts (154-155). Additionally, in one instance, the lexicalized *jo-* verb occurs with a form that is used as a postposition elsewhere (156).

(152) *jo-* 'RECP' + *mbo'e* 'teach' → *jombo'e* 'laugh with each other'

(153) *jo-* 'RECP' + *vopian* 'fence' → *jovopian* 'shield eyes (protect one's eyes from the brightness of the sunlight)'

(154) *jo-* 'RECP' + *ajun* '???' → *joajun* 'have two partners'

(155) *jo-* 'RECP' + *vavun* '???' → *jovavun* 'look up'

(156) *jo-* 'RECP' + *koty* 'side of' → *jokoty* 'be lefthanded'

When verbs that would otherwise bear the reciprocal take postpositional objects, the reciprocal appears as the postpositional object, and is not prefixed to the verb. Verbs that take a postpositional object include the transitive *nupã* 'hit' and ditransitive *mõndo* 'send, give, put'. Additionally, the reciprocal is preceded by a coreferential marker that co-varies with the person and number feature of the subject (see §3.7.13 for a description of coreferential marking).

(157) a. Nd-o-mõndo-i ahe matera o-**jo**=pe.
NEG-3.A-give-NEG people thing 3.COR-RECP=to
'We do not give things to each other.' (Juma: Text)

b. Aramẽ ahe, ahe kwar-uhu-i o-**jo**=vehe.
after.that people people tie-AUG-i 3.COR-RECP-at
'After that, we tie each other.' (Juma: Text)

c. Ore nd-oro-mõndo-i ore oro-**jo**=pe.
1.EXCL NEG-1.EXCL.A-give-NEG 1.EXCL 1.EXCL.COR-RECP=to
'We do not give (things) to each other.' (Juma: Text)

3.4.4 Causative

Kawahíva has two causative prefixes, *mõ-* (~ *mbo-* /_(C)V) and sociative causative *rero-*. The former appears on intransitive verbs, lexical (158) or derived (159), to introduce an agent/causer subject, resulting in a derived transitive verb.

(158) a. O-hin ji=mẽn=a=ga.
3.A-fall 1.SG=partner=NMLZ=3.SG.MASC
'My husband fell.' (Juma: Elicit)

b. O-**mbo**-hin kurumĩ=a=gã nh-ãpepo=a.
3.A-CAUS-fall kid=NMLZ=3.PL i-pan=NMLZ
'The kids dropped someone's pan.' (Juma: Elicit)

(159) a. O-j-apo ka kavyr=ipe.
3.A-REFL-do 3.SG.MASC jungle=in
'He got hurt in the jungle.' (Juma: Elicit)

b. O-**mbo**-j-apo gã ji=ra'yr=a=hẽa.
3.A-CAUS-REFL-do 3.PL 1.SG=offspring=NMLZ=3.SG.FEM
'They hurt my daughter.' (Juma: Elicit)

The causative *mõ-* is also lexicalized on a number of verbs which no longer carry transparently causative meaning (160-161), and others which may no longer be decomposable into their component parts (162-163). Additionally, in one case, the lexicalized *mõ-* verb combines with an element that is found as an affix elsewhere in the language (164), and in another case, the lexicalized *mõ-* verb is an intransitive verb (165).

(160) *mõ-* 'CAUS' + *'e* 'say' → *mbo'e* 'teach'

(161) *mõ-* 'CAUS' + *ahy* 'be in pain' → *mboahy* 'miss someone'

(162) *mõ-* 'CAUS' + *mbe'u* '???' → *mõmbe'u* 'tell something'

(163) *mõ-* 'CAUS' + *te* '???' → *mbote* 'become, transform'

(164) *mõ-* 'CAUS' + *'i* 'DIM' → *mbo'i* 'grind'

(165) *mõ-* 'CAUS' + *katxing* 'be odorous' → *mõkatxing* 'be odorous (like a snake)'

Transitive verbs cannot be causativized using *mõ-* (166a). To express the meaning that a *mõ-* transitive verb would have, speakers volunteer a construction where the transitive verb is used in the imperative in a direct quotation (166b).

(166) a. *O-mõ-mĩ gã ga.
3.A-CAUS-steal 3.PL 3.SG.MASC
'We made him steal (something).' (Jupaú: Elicit)

b. E-mĩ!, ‘i ki gã ga=pe.
2.SG.IMP-steal say PST 3.PL 3.SG.MASC=to
'Steal (it)! they said to him.' (Jupaú: Elicit)

The sociative causative *rero-* also adds a causer to an intransitive verb, which also participates in the action. Two examples of the sociative causative are provided in (167) and (168). In Table 3.28, I have listed all the verbs that combine with this causative.

(167) a. Karuvaruhua o-hẽm yputurimõ.
agouti 3.A-leave in.evening
'AGOUTIS go out in the evening.' (Jupaú: Elicit)

b. Ore hẽa mõãtã-i hẽa **rero**-hẽm=a, ‘i:te:ki:ko:ra.
1.EXCL 3.SG.FEM pull-i 3.SG.FEM SOC.CAUS-leave-NMLZ DIR.EVID
'We pulled her to take her out.' (Jupaú: Elicit)

(168) a. Akwembe'ea o-ho kavyr-ipe.
man 3.A-go jungle-in
'The MAN went to the jungle.' (Jupaú: Elicit)

b. Mãngã pe=**rero**-ho nãhẽ?, ‘i ki gã ore=ve ko.
who 2.PL.B=SOC.CAUS-go POT say PST 3.PL 1.EXCL=to REAL
'Who will you all take? They asked us.' (Jupaú: Text)

Table 3.28: Verbs with the applicative *rero-*.

	morpheme breakdown	gloss
/ɾɛɾɔ-hɔ/	SOC.CAUS-go	'take'
/ɾɛɾɔ-hun/	SOC.CAUS-come	'bring'
/ɾɛɾɔ-ɛkɔ/	SOC.CAUS-be/live	'have'
/ɾɛɾɔ-hɛ̃m/	SOC.CAUS-leave	'take out'
/ɾɛɾɔ-gahap/	SOC.CAUS-cross	'cross with'
/ɾɛɾɔ-kwam/	SOC.CAUS-pass	'pass by with'

3.4.5 Reflexive

Kawahíva forms reflexive verbs with the prefix *ji-*, which marks the identity between subject and object. I provide examples of the reflexive prefix in (169).

(169) a. O-ji-ape ga 'oin=a.
3.A-REFL-illuminate 3.SG. be.standing=NMLZ
'He is illuminating himself.' (Jupaú: Elicit)

b. O-ji-pe'a javytxinga jawara uvi.
3.A-REFL-distance dog jaguar from
'The dog distanced itself from the jaguar.' (Jupaú: Elicit)

The reflexive prefix *ji-* is also lexicalized on a number of verbs that no longer carry transparently reflexive meaning (170-172), and others that may no longer be decomposed into their component parts (173-174).

(170) *ji-* 'REFL' + *apo* 'do/make' → *jiapo* 'hurt oneself'

(171) *ji-* 'REFL' + *'u* 'eat' → *ji'u* 'have sex with someone from the same moiety.'

(172) *ji-* 'REFL' + *kwaham* 'know' → *jikwaham* 'to have a presentiment'

(173) *ji-* 'REFL' + *ahung* '???' → *jiahung* 'wash'

(174) *ji-* 'REFL' + *apu'u* '???' → *jiapu'u* 'burn (intr.)'

When verbs that would otherwise bear the reflexive take postpositional objects, the reflexive appears as the postpositional object, and is not prefixed to the verb. Verbs that take a postpositional object include transitive verbs like *kwan* 'tie', as in (175). Additionally, the reflexive is preceded by a coreferential marker that co-varies with the person and number features of the subject (see §3.7.13 for a description of coreferential marking).

(175) Gaha tehe tũhũi o-kwan o-ji=he.
3.SG.MASC.FOC IDLY INDEED 3.A-tie 3.COR-REFL=at
'He idly tied himself indeed.' (Juma: Elicit)

3.4.6 Pluractionality

Kawahíva marks pluractionality via reduplication. Specifically, reduplication indicates multiple participants in an action, as in (176a), or iteration of an action, as in (176b).

(176) a. **Are=rety–are=retyk** ki gã ko.
1.EXCL.B=RED-pull.down PST 3.PL REAL
'They pulled us all down.' (Jupaú: Elicit)

b. Anhãnga o-**kutu-kutu**-pam j-urua.
ghost 3.A-RED-pierce-COMPLETELY i-mouth
'It was the ghost that pierced his mouth (many times).' (Jupaú: Text)

Reduplication copies the entire root into the reduplicant, plus the object marker and *wh*-agreement object marker if either of them is present, no matter how many segments or syllables. Additionally, the prefix *i-* is copied just in case the *wh*-agreement object marker is also present;

contrast the example in (177c) where these morphemes co-occur and reduplication applies with example (178c) above, where only the prefix *i-* is present and is not copied into the reduplicant.

(177) a. **[Oreren]**-ore=ren-i, ki ko.
RED-1.EXCL.B=be.seated-i PST REAL
'We all sat down.'

b. *Ore-**[re]**-ren-i, ki ko.
1.EXCL.B-[red]-be.seated-i PST REAL
'We all sat down.'

c. Gã **[rembity]-(e)mb-i-tyk**-aver=a.
3.PL RED-WH.OBJ.II-i-take.down-NOM.PST=NMLZ
'Those (trees) which they (the non-indigenous people) took down.' (Juma: Elicit)

There does not seem to be any restriction on the phonological size of the structure that can be copied into the reduplicant. Verbs are mostly mono- or disyllabic, but trisyllabic roots also exist.

(178) a. Gã, i-mbo-**[ʼu]-ʼu**-i, ʻi:te:ki:ko.
3.PL i-CAUS-RED-eat-i DIR.EVID
'They fueled it (=the car).' (Lit.: they made it drink multiple times) (Juma: Text)

b. Aramẽ ki amãna **[ndara]-ndarang**-i, ʻi:te:ki:ko.
after.that PST rain RED-thunder–i DIR.EVID
'After that, it thundered.' (Juma: Text)

c. Aramẽ ki Uka=ga, j-**[apyha]-(a)pyhar**-i, ʻi:te:ki:ko, ore=ve ko.
after.that PST Uka=MASC i-RED-tie-i DIR.EVID 1.EXCL=for REAL
'After that, Uka tied it (many times) for us.'

3.4.7 Applicative

Kawahíva has one applicative morpheme, *-okan* (~ *-ukan*). The suffix *-okan* adds an argument with a goal or benefactive role to a transitive verb, lexical or derived. Examples of the transitive applicative suffix are provided in (179) and (180). In Table 3.29, I have listed all the verbs I found in combination with the suffix *-okan*.

(179) a. O-mbo-hin kurumĩ=gã j-ãpepoa.
3.A-CAUS-fall kid=PL i-pan
'The kids dropped (someone's) pan.' (Juma: Elicit)

b. O-mbo-hir-**ukan** hẽa mbiara javatxinga pe.
3.A-CAUS-fall-APPL 3.SG.FEM meat dog for
'She dropped the meat for the dog.' (Jupaú: Elicit)

(180) a. O-kwaha-te-hu ka matera.
3.A-know-REALLY-AUG 3.SG.MASC thing
'He knows things.' (Juma: Elicit)

b. O-kwaha-**ukan** javatxinga tajahua ore=ve.
3.A-know-APPL dog peccary 1.EXCL=for
'The dog indicated the peccary for us.' (Jupaú: Elicit)

Table 3.29: Verbs with the applicative *-okan*.

	morpheme breakdown	gloss
/hepiang-okan/	see-APPL	'show, introduce'
/kwaham-okan/	kwaha-APPL	'indicate'
/mbo-hin-okan/	CAUS-fall-APPL	'drop for'
/mbo-hun-okan/	CAUS-come-APPL	'follow someone'

3.4.8 Adverbials

A variety of adverbial meanings are conveyed by verbal suffixes, as shown in Table 3.30.

Table 3.30: Kawahíva adverbial suffixes.

Suffixes	Gloss
-jipe ∼ *-ipe*	'already'
-katu	'well'
-ete	'really, truly'
-pav (∼ *-pam* ∼ *-pap*)	'completely'
-uhu	'a lot'
-'i	'a bit, a little, a few'
-jav (∼ *-jam* ∼ *-ajap*)	'again'
-hura	'apparently, seemingly'
-ta	'almost'
-ypy	'firstly'
-verev	'almost'
-aha	'unfortunately'
-ahy	'forcefully'
-ahiv	'much, very'
-je	'still'
-a'ang	'pretendingly, deceivingly'

The morpheme order among these suffixes is given in (181)[19] I determined the order of the morphemes based on the observed transitivity relation among them.[20] Some adverbial suffixes

[19]More systematic documentation of semantic scope differences resulting from affix order differences is needed, but impressionistically speaking, the order of some affixes contributes to their relative scope to others in the template.

[20]That is to say that, if *-ajav* was found preceding *-ahiv*, and *-ahiv* was found preceding *-katu*, then I concluded that *-ajav* also preceded *-katu* unless there was an example that contradicted this conclusion. However, I could not determine the relative order among the suffixes *-pav*, *-hiv*, and *-verev*, for the lack of data showing a combination between them.

have a fixed order, but others may occupy different positions with respect to other suffixes; I use asterisks to indicate that a morpheme is also found in other positions in this schema. Template (182) shows the same order, this time using the gloss for each morpheme.

(181) root - pav* - ipe - ramõ* - i* - ajav - ha - i* - {pav*/hiv*/verev} - ʻi - uhu - ramõ* - katu - ete - i* - hiv* - ramõ* - tete

(182) root - COMPLETELY* - ALREADY - LIKE* - NEG* - AGAIN - UNFORTUNATELY - NEG* - {COMPLETELY*/FORCEFULLY*/ALMOST} - DIM - AUG - LIKE* - WELL - REALLY - NEG* - FORCEFULLY* - LIKE* - IDLY

3.4.9 Negation

Standard negation is expressed using the circumfix morpheme *n(da)-..-i*. The prefixal part is the outermost morpheme among the verbal prefixes, occurring outside the imperative agreement, subject agreement, and object agreement markers. The suffix in *n(da)-..-i* can be followed by some agent-oriented adverbial suffixes (e.g., *-ahy* 'forcefully' and *tete* 'idly'), as well as the adverbial suffix *-ramõ* 'like'.

(183) a. Nd-epe-'u-i karuvaruhu=a, e-jakuv=a=mẽ!
NEG-2.PL.IMP-eat-NEG agouti=NMLZ 2.SG.COR-be.hot=NMLZ=WHEN
'Do (you PL.) not eat agouti when you are feverish!' (Juma: Elicit)

b. Nd-a-hepiag-i jie matera=mõ.
NEG-1.SG.A-see-NEG 1.SG thing=NON.SPEC.INDF
'I did not see a thing/anything.' (Juma: Elicit)

c. Nda=ji=repiag-i gã=mõ.
NEG-1.SG.B=see-NEG 3.PL=NON.SPEC.INDF
'Nobody saw me.' (Juma: Elicit)

There is also a negative particle *ahyn* used in negative existentials and as a negative response particle.

(184) **Ahyn** tapy'ynha avatxi=a também ji=ve.
NEG non.indigenous.person corn=NMLZ also 1.SG=for
'There is no non-indigenous people's corn for me either.' (Jupaú: Text)

(185) **Ahyn**, nd-a-ro-ko-i dinheiro-hu=a, ʻi ka=pe ko nõmĩa.
NEG NEG-1.SG.A-SOC.CAUS-be-NEG money-AUG=NMLZ say 3.SG.MASC=to REAL FRUST
'No, I don't have money.' (uttered as an answer to an invite to visit another village) (Jupaú: Text)

Finally, Kawahíva has a privative *e'ym* (in Juma, and *i'im* in Jupaú), described in §3.3.5, and illustrated again in (186). This morpheme is also used in dependent clauses, including complement clauses §3.7.9, adjunct clauses §3.7.10, and relative clauses §3.7.11

(186) u'i-**'ym**=a
flour-NOM.NEG=NMLZ
'(A meal) without flour' (Juma: Text)

3.4.10 Habitual

Kawahíva uses the verb *eko* 'be' as an auxiliary verb to express habitual actions. In this capacity, *eko* 'be' is prefixed with a coreferential marker (see §3.7.13 for a description of coreferential marking), and bears the suffix *-vo*, which is also used in adjunct purposive clauses and clause chaining clauses (see §3.7.10 and §3.7.12 for a description of these clauses).

(187) O-i'inguhu javatxinga o-ko-vo ypyturimõ.
3.A-bark dog 3.COR-be-VO in.evening
'The dog has been barking a lot in the evenings.' (Jupaú: Elicit)

The auxiliary follows the subject and verb, as in (187), as well as the object, as in (188).

(188) O-kwan hẽa kurumĩ=a=gã o-ko-vo.
3.A-tie 3.SG.FEM kid=NMLZ=3.PL 3.COR-be-VO
'She has tied up the kids.' (Jupaú: Elicit)

The auxiliary may follow or precede VP-adverbials, such as *kavyripe* 'in the jungle'.

(189) a. Ji=ropa ji it-eko-vo kavyr=ipe.
1.SG.B=lose.way 1.SG 1.SG.COR-be-VO jungle=in
'I always lose my way (walking) in the jungle.' (Jupaú: Elicit)

b. I-ty'a ji kavyr=ipe it-eko-vo.
i-be.hungry 1.SG jungle=in 1.SG.COR-be-VO
'I always feel hungry (walking) in the jungle.' (Jupaú: Elicit)

Finally, the coreferential marking on the habitual auxiliary co-varies with the person/number features of the subject (190a), and never that of the object (190b).

(190) a. Ji=repia ka o-ko-vo.
1.SG.B=see 3.SG.MASC 3.COR-be-VO
'He has watched me.' (Jupaú: Elicit)

b. *Ji=repia ka ji/it-eko-vo.
1.SG.B=see 3.SG.MASC 1.SG.B/1.SG.COR-be-VO
'He has watched me.' (Jupaú: Elicit)

3.5 Particles

Kawahíva has several particles that convey a range of meanings, including tense and a reality status (i.e., irrealis and realis). The surface order of these particles in a matrix clause is schematized

in (191), which includes the verb (the bolded **v**) for reference. Most particles have a fixed order, with a few exceptions, including tense, frustrative, and progressive particles; I use asterisks to indicate their variability on surface order.

(191) IRR - TENSE* - **V** - TENSE - POT - EVID - REAL - FRUST* - HYP - FRUST*

The discussion below moves from the leftmost particles to the rightmost particles.

3.5.1 Mood

Kawahíva has a reality status system distinction, which distinguishes irrealis and realis. The realis is used for clauses that describe "situations as actualized, as having occurred or actually occurring, knowable through direct perception. The irrealis portrays situations as purely within the realm of thought, knowable only through imagination" (Mithun 1995). Table 3.31 presents the two particles that realize the reality status distinction in Kawahíva.

Table 3.31: Kawahíva reality particles.

Particles	Gloss
po	irrealis
ko	realis

I characterize these particles as realizing a reality system based on several semantic parameters from typological studies (Mithun 1995; Elliott 2000; Michael 2014), including (i) temporal reference (non-future vs. future), (ii) polarity (positive vs. negative), (iii) speaker-oriented modality, and (iv) hypotheticality.

The reality system follows an expected distribution with respect to temporal reference. In particular, clauses with non-future temporal reference exhibit realis marking, as in (192), while clauses with future temporal reference have irrealis marking, as in (193).

(192) Oji ve-'i ki Takãi Bitate=ga rur-i avo ko.
some.time.ago for-DIM PST Takãi Bitate=3.SG.MASC come-i here REAL
'Some days ago, Takãi Bitate came here.' (Juma: Text)

(193) Po ga rur-ajav-i nãhẽ.
IRR 3.SG.MASC come-AGAIN-i POT
'(Let's see if) he comes again.' (Juma: Text)

Additionally, the example in (194) shows that the realis particle is not compatible with future temporal reference markers (e.g., future tense marker or future temporal adverbials) in the same sentence. However, as we see in (195a), the irrealis particle is compatible with the past tense marker. Nevertheless, the meaning of the sentence is slightly different from its counterpart without the irrealis particle (195b). The sentence includes a component of doubt as to whether the event described actually happened.

(194) *Oro-ho txi ore kavyr-ipe ko'emamẽ ko.
1.EXCL.A FUT 1.EXCL. jungle-in in.morning REAL
Intended: 'We will go to the jungle tomorrow.' (Juma: Elicit)

(195) a. A-hepiang po jie gã oi'i Uru.Eu=gã pyri.
1.SG.A-see IRR 1.SG 3.PL some.time.ago Uru.Eu=3.PL by
'I perhaps saw them on the Uru Eu (reserve).' (Juma: Elicit)
b. A-hepiang jie gã oi'i Uru.Eu=gã pyri.
1.SG.A-see 1.SG 3.PL some.time.ago Uru.Eu=3.PL by
'I saw them on the Uru Eu (reserve) (some time) ago.' (Juma: Elicit)

The realis/irrealis distinction also patterns as expected with respect to sentence polarity: positive polarity sentences with non-future temporal reference exhibit realis marking, as in (192), while their negative polarity counterparts exhibit irrealis marking, as in (196).

(196) "Nd-ahe-rer-eko-katu-i po gã," a-'i tũhũi ji nõmĩa.
NEG-person-SOC.CAUS-be-WELL-NEG IRR 3.PL 1.SG.A-say INDEED 1.SG FRUST
'"They don't take care of us well", I said indeed, in vain.' (Juma: Text)

The relation of the realis/irrealis distinction to modality, e.g., polite imperative, also accords with a notional definition of (ir)realisness. The Kawahíva polite imperative utterance employs special 2nd person pronouns, i.e., *ene* '2.SG' and *epe* '2.PL', as in (197). The irrealis *po* here follows from the unrealized event expressed by this sentence.

(197) "Ene po txi nde=nhãrukanga nãhẽ", 'i ki ka ko.
2.SG IRR FUT 2.SG=rib POT say PST 3.SG.MASC REAL
'"Be careful with your ribs", he said.' (Juma: Text)

One other domain where irrealis is commonly found is hypotheticals, e.g., conditional sentences. Irrealis marking is to be expected in this domain because no specific realized event has occurred. Accordingly, Kawahíva requires irrealis marking in conditionals, as in (198).

(198) Nde=he i-mõmbe'u-ramẽ *(po), ji kwahav-i ra'e amõ.
2.SG=FOC i-tell-IF IRR 1.SG know-i TODAY.PST SUBJ
'If you were to tell it (to me), I would have known.' (Juma: Text)

Finally, in addition to these parameters, irrealis marking is possible in interrogative sentences, whether it is a yes-no question, as in (199), or a content question, like (200), although not obligatory in all interrogatives. Irrealis marking in yes-no questions is unsurprising because the speaker implies a lack of knowledge of whether the event was realized (Chafe 1995:354), as in (199a). It might be surprising to see its absence in the same contexts. However, such cases become less mysterious when one observes that they refer to contexts where the speaker directly perceives the event or has normal expectations that led them to believe that the event accords with reality, as is true for the sentence described in (199b).

(199) a. O-'u po nde=ra'yra=ga pira?
3.A-eat IRR 2.SG=offspring=3.SG.MASC fish
'Does your son eat fish?' (Juma: Elicit)

b. Ere-'u nde?
2.SG.A-eat 2.SG
'Did you eat?' (Common in everyday speech upon seeing that someone finished eating lunch or dinner, or around that time)

The pattern of irrealis marking might also be surprising in content questions since the event is presupposed, and the speaker is simply seeking more information about the event, as is the case in (200a). However, speakers may use irrealis marking in content questions to indicate uncertainty about the event actually having happened. For instance, in (200b), irrealis marking is possible because the speaker is unsure whether she heard a gunshot, i.e., whether the shooting event happened.

(200) a. Gara nde ere-'u?
what 2.SG 2.SG.A-eat
'What did you eat?' (Common in everyday speech upon seeing that someone finished eating lunch or dinner, or around that time)

b. Mãngã po o-mõpu tupãhua?
who IRR 3.A-shoot gun
'Who shot the gun?' (said upon hearing a noise similar to a gunshot coming from the jungle, but it could not be confirmed whether was actually a gunshot) (Juma: Text)

3.5.2 Tense

Tense information is encoded by the particles shown in Table 3.32. Two of them encode general past and future semantics. The remaining particles distinguish multiple temporal distinctions of the past, which indicates that Kawahíva has a graded past tense system (Dahl and Viveka 2011). The general particles may be dropped, in which case the default temporal reading of the sentence is non-future. Additionally, the graded past tense system also distinguishes whether the event described was witnessed by the speaker or someone else.[21] I also describe the temporal adverbial *oi'i* 'some time ago', whose semantics include several degrees of past, between today past and several months ago.

Table 3.32: Tense particles.

[21]Other sister languages of Kawahíva have been described as having particles that convey temporal and evidential distinctions, including Guajajára, Kaiowá, Tocantins do Asuriní, and Wayampí (Jensen 1998:553-554).

Particles	Gloss
ki	general past
txi	general future
ra'e	hodiernal (today) past
rai'i	hesternal (yesterday) past
ra'ikwehe	non-witnessed (by the speaker) distant past (about 20 years ago)
ikwehe	witnessed distant past (about 20 years ago)
rimba'e	non-witnessed (by the speaker) remote past (about 40 years ago)
mãipo	witnessed remote past (about 40 years ago)

The general past tense particle *ki* indicates that the action occurs anytime before the utterance time. In that way, it is compatible with past temporal adverbials, such as *oi'i* (~ *oji'i*) 'some time ago'.

(201) a. **Ki** ore ho-i, karuk=a=mẽ ko.
PST 1.EXCL go-i afternoon=NMLZ=TO REAL
'We went in the afternoon.' (Juma: Text)

b. Gara **ki** nde ere-apo oi'i?
what PST 2.SG 2.SG.A-do some.time.ago
'What did you do yesterday?' (Juma: Elicit)

c. Humaitá uvi ore ju=a=rinẽ, **ki** ore ho-i kavyr=ipe ko.
Humaitá from 1.EXCL come=NMLZ=AFTER PST 1.EXCL go-i jungle=in REAL
'After we returned from Humaitá, we went into the jungle.' (Juma: Elicit)

The future tense particle *txi* indicates that the action will occur sometime after the utterance time. It is compatible with future temporal adverbials, including *koi'iramẽ* 'later', and adverbial clauses with a future reference time.

(202) a. Gara po **txi** nde ere-apo koi'iramẽ?
what IRR FUT 2.SG 2.SG.A-do later
'What will you do later?' (Juma: Elicit)

b. Wesley=ga ho=ramẽ ko'emamẽ, **txi** jie mitakwanha po-i nãhẽ.
Wesley=3.SG.MASC go=WHEN tomorrow/morning FUT 1.SG bread make-i
POT
'When Wesley leaves tomorrow, I will make bread.' (Juma: Elicit)

The hodiernal or today past particle *ra'e* expresses past temporal reference within the day containing the utterance time. This is confirmed by the scenario described in the context in (203), presented to speakers, and the back translation of the question, "Did Mandeí make bread today?" Using other tense particles, such as *rai'i* 'yesterday past', is infelicitous considering this scenario and marked with a pound sign (#).

(203) Context: *Mandeí made bread and invited me to come for breakfast early today. Later that day, during the evening, my partner wonders what Mandeí made for breakfast.*

a. V-apo Mandeí=hẽa mitakwãnha **ra'e**?
3.A-make Mandeí=3.SG.FEM bread TODAY.PST
'Did Mandeí make bread today? (Juma: Elicit)

b. #V-apo Mandeí=hẽa mitakwãnha **rai'i**?
3.A-make Mandeí=3.SG.FEM bread YESTERDAY.PST
'Did Mandeí make bread yesterday? (Juma: Elicit)

The hesternal or yesterday past tense particle *rai'i* expresses past temporal reference within the day preceding the day of the utterance time. In the scenario in (204), the speaker inquires about an event that happened on the day immediately before. The question is felicitous if it includes *rai'i*, as in (204a) but not felicitous if it includes a different past tense particle, like *ra'e* 'today past', as in (204b).

(204) Context: *Mandeí made bread yesterday and invited me to come over for breakfast early today. Later that day, during the evening, my partner inquired what Mandeí made for breakfast the day before.*

a. Gara nde ere-apo **rai'i**?
what 2.SG 2.SG.A-make YESTERDAY.PST
'What did you make yesterday? (Juma: Elicit)

b. #Gara nde ere-apo **ra'e**?
what 2.SG 2.SG.A-make TODAY.PST
'What did you make today? (Juma: Elicit)

The non-witnessed distant past particle *ra'ikwehe* is used to refer to past actions that happened around 20 years before the present time. For instance, if someone asks a question about an event that happened when their interlocutor, who is 20-25 years old, was a child, and the speaker did not witness, the question should include the *ra'ikwehe* particle, as in (205a). Speakers are very strict about the contexts where these particles can be used. They will say that the person asking the question must use *ra'ikwehe* because "the person does not know about what happened and, therefore, is asking the question". Using *ikwehe* 'witnessed distant past' is infelicitous, as in (205b); speakers comment that *ikwehe* is used as part of an answer to a question with *ra'ikwehe*.

(205) Context: *Imagine you are Bitaté or Kunhãvé (who are 20-25 years old). Imagine your father broke his finger when you were just a little kid. Someone who did not witness this event but heard about your father's experience asks you how it happened.*

a. Mãrãnuhu nde=ruv=a=ga ga=reko-i o-po'ia i-mõpẽ-i
how.exactly 2.SG=father=NMLZ=3.SG.MASC 3.SG=be-i 3.COR-little.finger i-break-i
ra'ikwehe?
NWIT.DIST.PST

'How did your father break his little finger?' (Lit.: how exactly was your father (to) break his own little finger?) (Juma: Elicit)

b. #Mãrãnuhu nde=ruv=a=ga ga=reko-i o-po'ia i-mõpẽ-i
how.exactly 2.SG=father=NMLZ=3.SG.MASC 3.SG=be-i 3.COR-little.finger i-break-i
ikwehe?
WIT.DIST.PST
'How did your father break his little finger?' (Lit.: how exactly was your father (to) break his own little finger?) (Juma: Elicit)

In contrast, the witnessed distant past particle *ikwehe* indicates an event approximately 20 years before the utterance time and that the speaker witnessed it. For instance, when asked to volunteer a sentence related to the scenario in (206), the speaker offered (206a) as an explanation for the context in which their father broke a finger. Their sentence includes the particle *ikwehe*. Speakers reject using *ra'ikwehe* 'non-witnessed distant past' in this scenario, as in (206b), and explain the choice by adding "it cannot be *ra'ikwehe* because I was there".

(206) Context: *Imagine you are Bitaté or Kunhãvé (who are in their 20-25 years old). You saw your father breaking his finger when you were just a child. Someone asks about your father's experience, and you explain it to them.*

a. Tapyj=a ki ji=ruv=a=ga v-apo **ikwehe**.
house=NMLZ PST 2.SG=father=NMLZ=3.SG.MASC 3.A-make WIT.DIST.PST

'My father was building a house.' (Juma: Elicit)

b. #Tapyj=a ki ji=ruv=a=ga v-apo **ra'ikwehe**.
house=NMLZ PST 2.SG=father=NMLZ=3.SG.MASC 3.A-make NWIT.DIST.PST
'My father was building a house.' (Juma: Elicit)

Kawahíva also exhibits two particles that encode remote past. One is the non-witnessed remote past particle *rimba'e*, which is used to indicate past actions that happened around 40 years ago. The consultant, around 40, uses this past tense particle to refer to situations when she was a child, as in (207a). The speaker does not accept the other graded past tense particles in the same context, including *mãipo* 'witnessed remote past', as in (207b).

(207) Context: *Imagine your father broke his finger when you were just a child. Someone who did not witness this event but heard about your father's experience asks you how it happened.*

a. Mãrãnuhu nde=ruv=a=ga ga=reko-i o-po'ia i-mõpẽ-i
how.exactly 2.SG=father=NMLZ=3.SG.MASC 3.SG=be-i 3.COR-little.finger i-break-i
rimba'e?
NWIT.DIST.PST

'How did your father break his little finger? (Lit.: how exactly was your father (to) break his own little finger?) (Juma: Elicit)

b. #Mãrãnuhu nde=ruv=a=ga ga=reko-i o-po'ia i-mõpẽ-i
how.exactly 2.SG=father=NMLZ=3.SG.MASC 3.SG=be-i 3.COR-little.finger i-break-i
mãipo?
WIT.DIST.PST
'How did your father break his little finger? (Lit.: how exactly was your father (to) break his own little finger?) (Juma: Elicit)

Other scenarios where the particle *rimba'e* is used include events that occurred around the time when both the Juma and Jupaú were forced into permanent contact, i.e., in the late 1980s. The examples in (208) draw from texts that describe events from that time. The particle also appears in traditional stories.

(208) a. Mãrãmẽ gã ero-ho-pav-i iputahuva'ea **rimba'e**?
where 3.PL SOC.CAUS-go-COMPLETELY-i cattle NWIT.REM.PST
'Where did they take the cattle?' (Jupaú: Text)

b. O-kahup=a, po i-ho-i **rimba'e** nõmĩa.
3.COR-hunt=NMLZ IRR i-go-i NWIT.REM.PST FRUST
'To hunt, he might have gone.' (Juma: Text)

The other member in the pair of remote past tense particles is the witnessed remote past morpheme *mãipo*. It indicates an event witnessed by the speaker that happened around 40 years before the present. The imagined scenario in (209) includes an event that occurred over 20 years ago. The speaker, who is 40, offers (209a) as part of the context when their father broke a finger. The same speaker finds (209b) incongruous given the context, as this example includes *rimba'e* 'non-witnessed remote past', while the scenario indicates the speaker was present at the event.

(209) Context: *You saw your father breaking his finger when you were just a child. Someone asks about your father's experience, and you explain it to them.*

a. Tapyja ki ji=ruva=ga v-apo **mãipo**.
house PST 2.SG=father=NMLZ=3.SG.MASC 3.A-make WIT.DIST.PST

'My father was building a house.' (Juma: Elicit)

b. #Tapyja ki ji=ruva=ga v-apo **rimba'e**.
house PST 2.SG=father=NMLZ=3.SG.MASC 3.A-make NWIT.DIST.PST
'My father was building a house.' (Juma: Elicit)

Additionally, like *rimba'e*, the particle *mãipo* is commonly used in sentences that refer to events that took place around when the Juma and Jupaú established permanent contact with the non-Kawahíva society (i.e., in the '80s), with the additional piece that the speaker witnessed the event described. Example (210) is uttered in the same text of the question in (208) above. However, unlike the destination of the illegal cattle on the Jupaú land, the speaker can attest to the event when the cattle were removed.

(210) Gã iputahuva'ea rero-ho-pav-i **mãipo**.
3.PL cattle SOC.CAUS-go-COMPLETELY-i WIT.REM.PST
'They took all the cattle.' (Jupaú: Text)

Finally, speakers may use the adverbial *oi'i* (~ *oji'i*) 'some time ago' to refer to several degrees of past.[22] For instance, it can refer to events within the 'today past' and 'yesterday past', as in the examples in (211).

(211) Context: *Mandeí made bread and invited me to come for breakfast early today. Later that day, during the evening, my partner wonders what Mandeí made for breakfast, to which the speaker replies the following.*

a. V-apo Mandeí=hẽa mitakwanha **oi'i**.
3.A-make Mandeí=3.SG.FEM bread some.time.ago
'Mandeí made bread today.' (Juma: Elicit)

Context: *Mandeí made bread yesterday and invited me to come over for breakfast early morning today. Later that day, during the evening, my partner wonders what Mandeí made for breakfast, to which the speaker replies the following.*

b. V-apo Mandeí=hẽa mitakwanha **oi'i**.
3.A-make Mandeí=3.SG.FEM bread some.time.ago
'Mandeí made bread yesterday.' (Juma: Elicit)

Additionally, the adverbial *oi'i* 'some time ago' may be used to refer to past events going back in one week or several months, as in (212). When asked to consider the context in (212), the speaker included the adverbial in the example volunteered below.

(212) Context: *I have been to the Jupaú village Jaikara (aka 621) for a week, having first gone to the Juma village. Mandá wonders about what the Juma made for me to eat. I describe some of the things they made.*

[22] The morpheme *oi'i* is described as adverbial, rather than a particle, for its syntactic behavior: it can be fronted like other adverbials, such as *ko'emamẽ* 'tomorrow, in the morning'. In contrast, particles, including the tense particles described in this section, have a fixed position and cannot be fronted.

a. Mandeí=hẽa v-apo mitakwãnha o-ji-vyr=ipe ji=ve
Mandeí=3.SG.FEM 3.A-make bread 3.COR-REFL-village=in 1.SG=for
oi'i.
some.time.ago
'Mandeí made bread for me in her village.' (Juma: Elicit)

Context: *I have been to the Juma village for six months after going to the Jupaú village. Mandeí wonders about what the Jupaú made for me to eat. I describe some of the things they made.*

b. V-apo gã mitakwãnha **oi'i**.
3.A-make 3.PL bread some.time.ago
'They made bread.' (Juma: Elicit)

The particle *oi'i* may co-occur with the general past tense particle *ki*, as in (213). This indicates that *oi'i* does not form a paradigm with the tense particles.

(213) Gara **ki** nde ere-apo **oi'i**?
what PST 2.SG 2.SG.A-make some.time.ago
'What were you doing?' (Juma: Elicit)

3.5.3 Potential

Kawahíva uses the particle *nãhẽ* to indicate that an event is likely to happen. As such, *nãhẽ* is often translated with a near future interpretation.

(214) A-ho Aldeia Nova pype nãhẽ, 'i ki ka ko.
1.SG.A-go Aldeia Nova inside POT say PST 3.SG.MASC REAL
'I'll go to Aldeia Nova (village) soon, he said.' (Juma: Text)

Additionally, *nãhẽ* is compatible with the future adverbials, the future particle *txi*, and adverbial clauses with a future reference time.

(215) a. Koi'iramẽ ji rur-ajav-i, pe=pyri nãhẽ, 'i ki ka ko.
later 1.SG come-AGAIN-i 2.PL=near POT say PST 3.SG.MASC REAL
'Later, I will come again soon, he said.' (Juma: Text)

b. Wesley=ga ho-inẽ ko'emamẽ, txi jie mitakwanha apo-i nãhẽ.
W.=3.SG.MASC go-AFTER tomorrow FUT 1.SG cake make-i POT
'After Wesley leaves tomorrow, I'll make cake soon.' (Juma: Elicit)

c. Wesley=ga ho-inẽ ko'emamẽ, jie mitakwanha apo-i nãhẽ.
W.=3.SG.MASC go-AFTER tomorrow 1.SG cake make-i POT
'After Wesley leaves tomorrow, I'll make cake soon.' (Juma: Elicit)

Its syntactic status as a particle, rather than an adverbial, is consistent with the behavior of other particles. Like them, *nãhẽ* is strictly clause-final and cannot be fronted.

3.5.4 Evidentiality

In the Juma dialect, speakers use *'ite ki ko*, with the variants *'i ki ko*, *'ite ki ko*, and *'ite ki kora* to mark that an event was directly witnessed, and *'i ke hemõ* to indicate the event described is part of oral tradition. I describe *'ite ki ko* and *'i ke hemõ* as specialized markings of the source of information. This characterization fits the semantic definition of evidential markers (Brugman and Macaulay 2015; Aikhenvald 2018). The distribution of these markers is summarized in Table 3.33. Besides these evidentiality exclusive markers, recall that some tense particles indicate evidential semantics, as described in §3.5.2.

Table 3.33: Evidentiality marking.

Marking	Gloss	Distribution
'ite ki ko	direct evidential	Firsthand descriptions
'i ke hemõ	oral tradition evidential	Traditional stories, imagined events

The direct evidential marker *'ite ki kora* is commonly used in firsthand descriptions of events, as in (216), which also includes the context in which this sentence is uttered. It never appears in myths and narratives about the distant past; instead, *'i ke hemõ* is used in these contexts. Additionally, *'ite ki ko* never appears in texts where the speaker describes imagined events. In the examples, note that I have used colons between the morphemes that make up the evidential markers (e.g., *'ite ki kora*) as a way to indicate its non-compositionality meaning. I come back to compositionality shortly.

(216) *Context: In a monologue about the aftermath of a day of hunting, the speaker describes who brought which parts of the game animal, a common topic in conversations about hunting.*

Jie erur-i, nh-ã'ihua, i-mbotyryry-vo, ‘i:te:ki:ko:ra.
1.SG bring-i i-heart i-fry-VO DIR.EVID

‘I brought the (animal's) heart to fry it.' (Juma: Text)

In contrast, the evidential *'i ke hemõ* is common in traditional stories, as in (217), and never used in firsthand descriptions.

(217) *Context: A myth about a deer that got lost after the non-indigenous people made a group of deer scatter, but managed to return to its group.*

A'eramẽ ki, o-ho-hu-te ‘yhua ji=rejar=a, ‘i:ke:hemõ.
after.that PST 3.A-go-AUG-REALLY deer 1.SG=leave=NMLZ ORAL.TRAD.EVID

‘After that, the deer went away and left me.' (Juma: Text)

Morphologically, the more complex variant *'ite ki kora* is a combination of a verb, a suffix, and three particles: the verb *'i* ‘say', suffixed with *-te* ‘really', plus the general past tense particle *ki*, followed by the realis marker *ko* and a clitic form of *ra'e* ‘today past', i.e. *ra*. Despite the fact that each component of this marker can be identified separately, the meaning of *'ite ki ko* is not

compositional. The morphological composition of *'i ke hemõ* is less straightforward; I take it to be the combination of the verb *'i* 'say', the general past tense particle *ki*, the noun *ahe* 'person', and the particle *amõ*, which I characterize as a subjunctive particle in §3.5.6.

3.5.5 Frustrative

The frustative semantics is encoded by the particle *nõmĩa* (~ *rõmĩa*). Generally speaking, frustratives are grammatical markers that express the non-realisation of some expected outcome implied by the proposition expressed in the marked clause (Overall 2017). I characterize this particle as frustrative based on the scenarios volunteered by speakers where the particle is present, such as the one for the example in (218).

(218) Speaker's comment: Viviane wanted to eat fish but couldn't because she choked.

T-a-'u pira hẽa nõmĩa.
OPT-1.SG.A-eat fish 3.SG.FEM FRUST

'She wanted to eat fish in vain.' (Jupaú: Elicit)

Sometimes, especially in spontaneous speech, it is not straightforward to tell what led a speaker to use the frustrative. For instance, the particle is consistently used to reply to a greeting without further elaboration, as in the common exchange in (219).

(219) a. Pyryete nde?
be.well 2.SG
'Are you good?' (Common in everyday speech)

b. Pyryete nõmĩa.
be.well FRUST
'It's all good.' (Common in everyday speech)

3.5.6 Hypothetical

Kawahíva exhibits the particle *amõ*, which serves to indicate hypotheticality. The characterization of this particle as a hypothetical marker is based on its perceived semantics and discourse contexts of use in spontaneous speech. For instance, *amõ* was used in 44 out of 53 sentences of a text in which a Juma speaker speculates what her first days back in her village will be like after having been in a different Kawahíva village for months due to the COVID pandemic. All other sentences without *amõ* describe the speaker's situation in the current village. These examples also show that *amõ* is compatible with the irrealis particle *po*. This is expected since the events described were speculations, and the irrealis particle is triggered in situations knowable only through imagination, as described in §3.5.1.

(220) a. Nãhã po ji ho-i ji=vyr=ipe ra'e amõ.
thus IRR 1.SG go-i 1.SG=village=in TODAY.PST HYP
'Thus I will go to my village.' (Juma: Text)

b. Aramẽ po ji matera apo-i ore=ve i-vo-vo ra'e amõ.
after.that IRR 1.SG food make-i 1.EXCL=for i-eat-VO TODAY.PST HYP
'After that, I will make food for us to eat.' (Juma: Text)

An identical cognate morpheme is described in sister languages to Kawahíva, including Kamaiurá and Guajajará, as an indefinite pronoun meaning 'other'. In Kawahíva, the particle *amõ* is also used as a pronoun, translated as 'other' when combined with the nominalizer enclitic *=a*. However, there does not seem to be any sister language with a cognate morpheme *amõ* showing similar semantics of hypothetical.

3.6 Postpositions

Kawahíva has several postpositions that primarily express a variety of spatial relationships. The postpositions and their functions are summarized in Table 3.34.

Table 3.34: Postposition forms and functions.

postposition	gloss
pe	to, for
rehe	at
upe	on
ipe	in
ruvi	from
rupi	through, along
pype	inside, into
pymõ	along with
pomẽ	with
koty	beside
irẽ	after
pyri	near
erõnde	ahead
reviri	behind, before
ʔarimõ	on top of
vyripe	below
pyteripe	in the middle of

Postpositions exhibit a cluster of properties that bring them closer to verbs than nouns. For instance, they can take a SET B marker, as in (221), which are bound pronouns (Dos Santos 2023a).

(221) O-mõpu ga **ji**=rehe.
3.A-shoot 3.SG.MASC 1.SG.B=at
'He shot at me.' (Juma: Elicit)

Postpositions also host voice morphology, such as the reflexive and the reciprocal, and co-referential markers, as shown in examples (222a-b).

(222) a. Ere-mõpu nde e-**ji**-ehe.
2.SG.A-shoot 2.SG 2.SG.COR-REFL=at
'You shot at yourself.' (Juma: Elicit)

b. T-a-hepiang nde=kyva, e-'i ki anhãnga o-**jo**=upe.
OPT-1.SG.A-see 2.SG=lice i-say PST spirit 3.COR-RECP=on
'"Let me see your lice", the spirits said to each other.' (Juma: Text)

Postpositions, like verbs, can combine with adverbial suffixes, like the diminutive *'i*, in example (223).

(223) Oji ve-**'i** ki Takãi Bitate=ga rur-i avo ko.
some.time.ago for-DIM PST Takãi Bitate=MASC come-i here REAL
'Just the other day, Takãi Bitaté came here.' (Juma: Text)

Finally, although the primary function of postpositions is to express spatial relations, there are a few cases where at least some of them are involved in the expression of temporal semantics. The most straightforward instance involves *pe* 'to, for', as in (224). Additionally, forms identical to *irẽ* 'after' and *eronde* 'ahead' appear in *after-* (225) and *before-* adjunct clauses (226).

(224) a. ko'em=a=**mẽ**
early=NMLZ=to
'In the morning, tomorrow' (Common in everyday speech)

b. karuk=a=**mẽ**
afternoon=NMLZ=to
'In the afternoon' (Common in everyday speech)

c. oi'i=**ve**
some.time.ago=for
'Some time ago' (Common in everyday speech)

(225) E-nduvun [e-võnhã mbovur-**irẽ**]!
2.SG.IMP-spit 2.SG.COR-tooth remove-AFTER
'Spit after you remove your tooth!' (Juma: Elicit)

(226) [Ji=ho=a=**rerõnde**=koty] ga 'e-jav-i.
1.SG=go=NMLZ=BEFORE=side 3.SG.MASC say-AGAIN-i
'Before I went, he said again.' (Juma: Text)

3.7 Clauses and clausal phenomena

3.7.1 Word order

Kawahíva constituent order is VSO in discourse-neutral scenarios in matrix clauses (e.g., new-event reporting and out-of-the-blue matrix clauses). I provide three examples below, uttered in a new-information reporting context, for a transitive sentence (227), a ditransitive sentence (228), and an intransitive sentence (229).

(227) V-ero-ho ki polícia=gã gã ko.
3.A-SOC.CAUS-go PST police=3.PL 3.PL REAL
'The police took them.' (Juma: Text)

(228) O-mõndo ki ga pira akwemba'ea=pe ko.
3.A-send PST 3.SG.MASC fish man=to REAL
'He gave fish to the man.' (Juma: Elicit)

(229) J-ape'um-ahi ki gã ko.
i-be.dirty-VERY PST 3.PL REAL
'They were filthy.' (Juma: Text)

Dependent clauses exhibit SOV order, as in the complement clause in (39). Note that any position other than clause-final for the verb is unacceptable. This verb-final order pattern also extends to adjunct and relative clauses (not shown below, but described in §3.7.10 and §3.7.11).

(230) a. A-hepiang ki jie [ji=ruva=ga pira **'u=a**].
1.SG.A-see PST 1.SG 1.SG=father=3.SG.MASC fish eat=NMLZ
'I saw [my father eating fish].' (Juma: Elicit)

b. *A-hepiang ki jie [ji=ruva=ga **'u=a** pira].
1.SG.A-see PST 1.SG 1.SG=father=3.SG.MASC eat=NMLZ fish
'I saw [my father eating fish].' (Juma: Elicit)

c. *A-hepiang ki jie [**'u=a** ji=ruva=ga pira].
1.SG.A-see PST 1.SG eat=NMLZ 1.SG=father=3.SG.MASC fish
'I saw [my father eating fish].' (Juma: Elicit)

In Chapter 4, I develop the first account of verb-initial clauses in Kawahíva, arguing that this pattern results from the long head movement (or syntactic head movement) of the verb to the clause-initial position. This is the first study to show that a language can use long head movement to derive the verb-initial word order.

SOV order is found in matrix clauses with clause-initial adverbials, including i) discourse particles, such as *a'ero* 'then, therefore', *aramẽ* 'after that', *a'ea rupi* 'with that', as in (231a). The SOV order is also found when the clause-initial position is occupied by ii) adverbial frame-setting

postpositional phrases, such as *oi'i ve* 'some time ago', *kavyripe* 'in the woods', and others, as in (231b).[23]

(231) a. A'ero hẽa=kuvyra=ga itai'ia mõndo–i.
then 3.SG.FEM=brother=3.SG.MASC fishhook throw-i
Then her young brother (was) throwing a fishhook.' (Juma: Text)

b. Oi'i ve-'i ji=ruva=ga uruvia pyhyk–i ko.
some.time.ago for-DIM 1.SG=father=3.SG.MASC catfish catch-i REAL
'Some time ago, my father caught catfish.' (Juma: Text)

OSV word order is found in object extraction contexts, including object questions (232), object focus (233), and topicalized objects (234).

(232) Mãngã po ji=ruvyra v-epiang rimba'e?
who IRR 1.SG=uncle 3.A-see DISTANT.PST
'Who did my uncle see back in the day?' (Juma: Text)

(233) Pirapetxinguhua ki hẽa o-pyhy-pyhy ko.
pirapetxinguhua.fish PST 3.SG.FEM 3.A-RED-catch REAL
'It was *pirapetxinguhua* fish that she caught.' (Juma: Text)

(234) Takãi Kajuvi ki, tapy'ynha o-purun ga=rehe.
Takãi Kajuvi PST non.indigenous 3.A-hit 3.SG.MASC=at
'(As for) Takãi Kajuvi, the non-indigenous person ran over him.' (Juma: Text)

Kawahíva exhibits SVO word order in matrix clauses with subject extraction, including subject questions (235), subject focus (236), and topicalized subjects (237).

(235) Mãngã ki v-epiang kwemba'ea?
who PST 3.A-see man
'Who saw the man?' (Juma: Elicit)

(236) Mbahira ki v-apo kunhã.
Mbahira PST 3.A-make woman
'It was Mbahira who made the woman.' (Jupaú: Text)

(237) Nhãnde=katu nde-kwahav-i nanungara.
1.INCL=WELL NEG-know-NEG thing
'As for us (INCL.), (we INCL.) don't know (how to) things.' (Jupaú: Text)

[23]In natural speech, speakers tend to drop the adverbial in clause-initial position in adjacent sentences after it has been repeated. Speakers see no difference when asked to judge pairwise comparisons of SOV sentences with and without the clause-initial adjunct.

Finally, dependent clauses show fixed SOV word order. I provide examples showing the requirement that the verb must be final with respect to its arguments in complement clauses (238), relative clauses (239), and adjunct clauses (240).

(238) a. A-hepiang ki jie [ji=ruva=ga pira **ʻu**=a].
1.SG.A-see PST 1.SG 1.SG=father=3.SG.MASC fish eat=NMLZ
'I saw [my father eating fish].' (Juma: Elicit)

b. *A-hepiang ki jie [ji=ruva=ga **ʻu**=a pira].
1.SG.A-see PST 1.SG 1.SG=father=3.SG.MASC eat=NMLZ fish
'I saw [my father eating fish].' (Juma: Elicit)

c. *A-hepiang ki jie [**ʻu**=a ji=ruva=ga pira].
1.SG.A-see PST 1.SG eat=NMLZ 1.SG=father=3.SG.MASC fish
'I saw [my father eating fish].' (Juma: Elicit)

(239) a. [Tapy'ynha mõhãnga mbu-**hu**-har=a]=gã te'i o–hun
non.indigenous medicine CAUS-come–WH.OBL.I=NMLZ=3.PL only 3.A–come
ore=pyri.
1.EXCL=by
'The non-indigenous people [who bring medicine] are the only ones that come to us.' (Juma: Text)

b. *[Tapy'ynha mbu-**hu**-har=a mõhãnga]=gã te'i o–hun
non.indigenous CAUS-come–WH.OBL.I=NMLZ medicine=3.PL only 3.A–come
ore=pyri.
1.EXCL=by
'The non-indigenous people [who bring medicine] are the only ones that come to us.' (Juma: Elicit)

c. *[Mbu-**hu**-har=a tapy'ynha mõhãnga]=gã te'i o–hun
CAUS-come–WH.OBL.I=NMLZ non.indigenous medicine=3.PL only 3.A–come
ore=pyri.
1.EXCL=by
'The non-indigenous people [who bring medicine] are the only ones that come to us.' (Juma: Elicit)

(240) a. [Ji kandambuhua **ʻu**-ramẽ], ji=reveka nda-katu-i.
1.SG papaya eat–WHEN 1.SG=stomach NEG-be.well-i
'[When I eat papaya], my stomach gets bad.' (Juma: Elicit)

b. *[Ji **ʻu**-ramẽ kandambuhua], ji=reveka nda-katu-i.
1.SG eat-WHEN papaya 1.SG=stomach NEG-be.well-i
'[When I eat papaya], my stomach gets bad.' (Juma: Elicit)

c. *[**‘u**-ramẽ ji kandambuhua], ji=reveka nda-katu-i.
eat-WHEN 1.SG papaya 1.SG=stomach NEG-be.well-i
‘[When I eat papaya], my stomach gets bad.’ (Juma: Elicit)

3.7.2 Non-verbal predicates

In addition to verbal predicates, Kawahíva allows noun phrases to act as predicates. Noun phrases may be used as predicates without any verbalizing morphology. These can be used in predicational, identificational, specificational, and equative constructions.

In predicational constructions, the nominal predicates a property of the subject: in (241), *tapy’ynha* predicates the non-indigenous property of the subject referent. Predicates are shown in bold.

(241) **Tapy’ynh=a** jie.
non.indigenous=NMLZ 1.SG
‘I am a non-indigenous person.’ (Juma: Elicit)

Identificational constructions, which relate a deictic element and a nominal, are formed using a demonstrative pronoun followed by a noun phrase, which appears with a nominal enclitic for living humans, as in example (242).

(242) Koro **ji=rembiriko=a=hẽa**.
DEM:PROX:ROUND 1.SG=partner=NMLZ=3.SG.FEM
‘This is my wife.’ (Juma: Elicit)

Equative constructions, which relate two referential expressions, involve two nominals, as in (243); the nominal predicate also appears with the enclitic for living humans.

(243) a. Ji=ruva=ga **tapy’ynh=a=ga**.
1.SG=father=3.SG.MASC non.indigenous=NMLZ=3.SG.MASC
‘My father is the non-indigenous person (in the village).’ (Juma: Elicit)

b. Ji=y=hẽa **i-mõkwatxijahar-uhu=a=hẽa**.
1.SG=mother=3.SG.FEM i-teacher-AUG=NMLZ=3.SG.FEM
‘My mother is the teacher (of the village).’ (Juma: Elicit)

In specificational constructions, which specify who or what a particular individual is, the nominal predicate appears without an enclitic for living humans: in (244), *Ĵuma gã tavijara* ‘The Juma chief’ specifies who Boreá is.

(244) Juma=gã **tavijar=a** Borea=hẽa.
Juma=3.PL chief=NMLZ B.=3.SG.FEM
‘The Juma chief is Borea.’ (Juma: Elicit)

In negation, the predicate is negated with the constituent negation marker *=rõi*, as in the predicational construction in (245a) and the equative construction in (245b).

(245) a. Ji=ruva=ga **kawahiv=a=rõi**.
1.SG=father=3.SG.MASC indigenous=NMLZ=NEG
'My father is not indigenous.' (Juma: Elicit)

b. Ji=ruva=ga **kawahiv=a=ga=rõi**.
1.SG=father=3.SG.MASC indigenous=NMLZ=3.SG.MASC=NEG
'My father is not the indigenous person.' (Juma: Elicit)

3.7.3 Question

3.7.3.1 Polar questions

Matrix polar questions have a clause-final falling intonation and the same VSO constituent order as declarative matrix clauses in discourse-neutral contexts (246).

(246) O-'u po nde=ra'yr=a=ga pira?
3.A-eat IRR 2.SG=offspring=NMLZ=3.SG.MASC fish
'Does your son eat fish?' (Juma: Elicit)

Kawahíva does not have embedded polar questions. These are conveyed using a direct quotation introduced by the verb *'i* 'say, ask', as in (247).

(247) "Nd-ahy-i-hu-te po nde=ve", 'i ki ahe=ruva ahe=ve.
NEG-hurt-NEG-AUG-REALLY IRR 2.SG=to say PST person=father person=to
'"Didn't it really hurt you?" Our father said to us.' (Juma: Text)

3.7.3.2 Content questions

Matrix content questions are formed by fronting the question word, so that the question word appears at the left edge of the clause. A list of question words in Kawahíva is provided in Table 3.35.

Table 3.35: Question words.

question word	gloss
mãngã	‘who’
mãngãngatu	‘who exactly’
gara	‘what’
garakatu	‘what exactly’
gara rehe	‘for what purpose’
mõmẽ	‘where’
mõmẽngatu	‘where exactly’
mãrã	‘how’
mãrãngatu	‘how exactly’
mãrĩmẽ	‘what time’
mãrãmẽ	‘when’
mãrupi	‘what way’
mãrãnũhũ	‘what reason’
mãrãmõmĩ	‘how many’
garamõ	‘for what’

Additionally, the question word appears before the irrealis and tense particles. Subject questions exhibit SVO word order, as in (248); object questions exhibit OSV word order, as in (249); and non-argument question constructions is SOV, as in (250).

(248) Mãngã po o-mõpu tupãhua?
who IRR 3.A-shoot gun
‘Who shot the gun?’ (Juma: Text)

(249) Mãngã po ji=ruvyra v-epiang rimba’e.
whom IRR 1.SG=deceased.uncle 3.A-see NWIT.DISTANT.PST
‘Whom did my uncle see in the old times?’ (Juma: Text)

(250) Mõmẽ ki Mandeí=hẽa Clebson=ga repiag-i?
where PST Mandeí=3.SG.FEM Clebson=3.SG.MASC see-i
‘Where did Mandeí see Clebson?’ (Juma: Elicit)

Possessors, under questioning, undergo fronting with the entire possessive phrase, as shown in (251).

(251) “**Mãngã pira** ga txiro o-mõ-ngi”, a-’i ki ji
who clothing 3.SG.MASC CONTINUOUSLY 3.A-CAUS-enter 1.SG.A-say PST 1.SG
ko=ra.
REAL=TODAY.PST
‘“Whose clothes is he using?”, I asked’ (Juma: Text)

Kawahíva allows multiple question words in a sentence, as in (252a). However, there is a requirement regarding the order of subject and object question words. The subject question word takes precedence over the object question word in appearing at the left edge of the clause (252b). Fronting both question words despite preserving that order requirement is also ungrammatical, as shown in (252c).

(252) a. Mãngã v-epiang gara?
who 3.A-see what
'Who saw what?' (Juma: Elicit)

b. *Gara mãngã v-epiang?
what who 3.A-see
'What did who see?' (Juma: Elicit)

c. *Mãngã gara v-epiang?
who what 3.A-see
'Who saw what?' (Juma: Elicit)

Kawahíva does not exhibit embedded questions. I provide two examples in (253) to show that *wh*-words are rejected by speakers in embedded clauses under *kwaham* 'know' and *hepiang* 'see'.

(253) a. *Nd-a-kwahav-i jie [mãngã=pe Wesley=ga u'ia
NEG-1.SG.A-know-NEG 1.SG who=to W.=3.SG.MASC manioc.flour
rero-ho=a].
SOC.CAUS-go=NMLZ
Intended: 'I don't know who Wesley took manioc flour to.' (Juma: Elicit)

b. *Nd-a-hepiak-i jie [mãngã=pe Wesley=ga u'ia rero-ho=a].
NEG-1.SG.A-see-NEG 1.SG who=to W.=3.SG.MASC manioc.flour SOC.CAUS-go=NMLZ
Intended: 'I didn't see who Wesley took manioc flour to.' (Juma: Elicit)

Speakers use two juxtaposed independent clauses to convey the meaning of an embedded question, as in (254), wherein one of the clauses is a simple content question.

(254) Mãngã=pe Wesley=ga u'ia rero-ho-i? Nd-a-kwahav-i jie.
who=to Wesley=3.SG.MASC manioc.flour SOC.CAUS-go-i NEG-1.SG.A-know-NEG 1.SG
'Who did Wesley take manioc flour to? I don't know.'

3.7.4 Focus

Kawahíva exhibits distinct constructions for information and contrastive focus, on the one hand, and exhaustive focus, on the other hand (Aissen 2023:for an overview). Information and contrastive focus involve the same fronting construction also employed for content interrogatives, while exhaustive focus involves the particle *te'i* 'only'. Further investigation is needed to clarify whether information and contrastive focus are distinguished by other means, including prosody.

In information focus constructions, the focused constituent appears clause-initially. The commonly used method for eliciting information focus constructions is questioning: the focused constituent is the part of the sentence corresponding to the answer to the question. In examples (255) and (256), the second member of the pair includes, respectively, a focused subject constituent and a focused object constituent in clause-initial position, which corresponds to the question word (i.e., *mãngã* 'who' and *gara* 'what') in the question. Note the constituent bearing information focus precedes irrealis mood and tense particles when these particles are present, just like the subject and object question words in (255) and (256). Since question words occur in clause-initial position by fronting, the analogous distribution of focused constituents indicates that they are also fronted to that position. Additionally, there is no pause following the focused constituent in these contexts. Focused constituents are bolded in the examples below, both in Kawahíva and free translation lines.

(255) a. **Mãngã** ki v-er-eko Amondawa=hẽa ko?
who PST 3.A-SOC.CAUS-be Amondawa=3.SG.FEM REAL
'Who married an Amondawa woman?' (Jupaú: Elicit)

b. **Ji=rikira=ga** ki v-er-eko Amondawa=hẽa ko.
1.SG=brother=3.SG.MASC PST 3.A-SOC.CAUS-be Amondawa=3.SG.FEM REAL
'**My brother** married an Amondawa woman.' (Jupaú: Elicit)

(256) a. **Gara** ki nde ere-apo oi'i?
what PST 2.SG 2.SG.A-make some.time.ago
'What did you make yesterday?' (Jupaú: Elicit)

b. **Mitakwanha** ki ji a-po oi'i.
bread PST 2.SG 2.SG.A-make some.time.ago
'I made **bread** yesterday.' (Jupaú: Elicit)

Contrastive focus constructions also involve the fronting of the focused constituent, as shown in the two most common subtypes of contrastive focus. In selective focus constructions, as in (257), the object focus appears clause-initially. Likewise, in corrective focus constructions, as does the corrective object focus in (258).

(257) a. (Person A) Who did you see, Wesley or his partner?

b. (Person B) **Wesley=ga** ki ji a-hepia ko.
W.=3.SG.MASC PST 1.SG 1.SG.A-see REAL
'It was Wesley who I saw.' (Juma: Elicit)

(258) a. (Person A) The 623 villagers planted beans.

b. (Person B) Ahyn! **Mãjioka** ki gã o-tym ko.
NEG manioc PST 3.PL 3.A-plant REAL
'No! (It was) manioc (that) they planted.' (Jupaú: Elicit)

Exhaustive focus constructions are formed using the particle *te'i* 'only',[24] placed to the right edge of the focused constituent. The focused constituent may or may not be fronted; the choice depends on whether another constituent has already filled the clause-initial position. Both subject and object constituents under exhaustive focus appear clause-initially in (259) but not in (260); note that clause-initial exhaustively focused constituents appear before the tense particle.

(259) a. **Tampinha=ga** te'i ki o-puta ko.
T.=3.SG.MASC only PST 3.A-stay REAL
'Only Tampinha stayed.' (Juma: Text)

b. **Pirahua** te'i ki ahe o-u'u.
jatuarana.fish only PST people 3.A-eat
'It is only *jatuarana* fish that we eat.' (Juma: Text)

(260) a. Aramẽ ki **ore** te'i ore reko-i.
after.that PST 1.EXCL only 1.EXCL live-i
'After that, only we stayed.' (Juma: Text)

b. Ore ki **pira-'i=a** te'i oro-verun ko.
1.EXCL PST fish-DIM=NMLZ only 1.EXCL.A-bring REAL
'We brought only small fish.' (Juma: Text)

I do not observe any difference between the syntactic strategies in focus constructions involving nouns versus pronouns. Information focus is formed with the special set of focusing-bearing pronouns (described in §3.3.9) in clause-initial position, as shown in (261), a pronominal subject focus construction. Note the subject pronoun *gaha* precedes the tense particle.

(261) **Gaha** ki ore=rer-u ko.
3.SG.MASC.FOC PST 1.EXCL.B=SOC.CAUS-come REAL
'(It was) he (who) brought us.' (Juma: Text)

Exhaustive focus constructions also exhibit the particle *te'i* 'only' at the right edge of the focusing-bearing pronoun, as in (262).

(262) A'ero po **gaha** **te'i** i-mõkwatxijar-i nãhẽ, a-'i ki ji nõmĩa.
then IRR 3.SG.MASC.FOC only i-write-i POT 1.SG.A-say PST 1.SG FRUST
'Then only he would write it (the Kawahíva language), I thought.' (Juma: Text)

The third and final strategy, contrastive focus, also exhibits the noun-focused strategy (i.e., fronting the focused pronoun), as demonstrated with an example of corrective focus in (263).

(263) a. (Speaker A) She went to the field.

[24] This particle is related to the verbal predicate *te'i* 'be alone'.

b. (Speaker B) Ahyn! **Gaha** ki o-ho karyr-ipe ko.
NEG 3.SG.MASC.FOC PST 3.A-go field-in REAL
'No! (It was) he (who) went to the field.' (Juma: Elicit)

3.7.5 Topic

Kawahíva distinguishes three strategies for the expression of topic, including i) the use of unfocused pronouns (for living human referents) and null argument (for non-living human and non-human referents), ii) left-dislocation followed by resumption, and iii) the use of *katu* 'well'. Each strategy can be associated with a different subtype of topic expression, including topic continuity, switch topic, and contrastive topic. This description follows Lambrecht (1994)'s theoretical framework on topic.

Topics are expressed using non-focused pronouns for living human referents, as in (264). In the first example, the first line of a personal experience narrative, the speaker mentions her youngest son *Kajuvi*, who got run over by a motorcycle, for the first time; she then describes what happened after she received the bad news about her son. In the second mention of *Kajuvi*, in (264b), the speaker uses the pronoun *ga* '3.SG.MASC'.

(264) a. Nãhã ki **Kajuvi**=ga apoa, ga=pyg=a=mẽ ko.
thus PST Kajuvi=3.SG.MASC whatchamacallit 3.SG.MASC=hit=NMLZ=WHEN REAL
'It was thus when the whatchamacallit hit Kajuvi.' (Juma: Text)

b. Ipyambu'iva'ea **ga**=pyg=a=mẽ ko.
motorcycle 3.SG.MASC=hit-NMLZ=WHEN REAL
'When the motorcycle hit him.' (Juma: Text)

In contrast, topics with a non-living human or a non-human referent show a different strategy. They usually appear as null arguments or may be referred to using the prosodically independent word *a'ea* 'this one'. For instance, example (265a) opens a story where the speaker describes how her late uncle got injured and killed by a jaguar. Another instance of this null strategy for topics is illustrated in (266). The speaker, describing how she went fishing, mentions the noun phrase *pira* 'fish' in (266a) twice but does not use it again in the subsequent sentence in (266b) that also contains the referent of this noun phrase.

(265) a. Nãhã ki **ore=ruvyra**.
thus pst 1.excl=late.uncle
'It was thus (with) our late uncle.' (Juma: Text)

b. Jawara i-pyhyk-i mãipo.
jaguar i-catch-i WIT.REM.PST
'The jaguar caught (him).' (Juma: Text)

(266) a. Karuka=mẽ ki ji ho-i **pira** reru=a oi'i. **Pira** rerur-i.
afternoon=to PST 1.SG go-i fish bring=NMLZ some.time.ago fish bring-i
'In the afternoon, I went to bring fish. (I) brought fish.' (Juma: Text)

b. Ko'ema=mẽ jie i-mõtyryry-i.
morning=to 1.SG i-fry-i
'Early morning, I fried (it).' (Juma: Text)

Switch topic is expressed by a construction in which the switch topic appears at the left edge of the clause, followed by a prosodic break, as in (267). The referent of *Takãi Kajuvi* is a topic argument in this example: it is first mentioned at the beginning of the text (sentences 1 and 2, the examples in 264 above); then it is only ever mentioned again in sentence 18 of the text, the example (267). Additionally, the proper noun *Takãi Kajuvi* and the determiner *ga* are non-contiguous.[25]

(267) **Takãi Kajuvi**, tapy'ynha o-purun ga=rehe.
Takãi Kajuvi non.indigenous 3.A-hit 3.SG.MASC=at
'(About) Takãi Kajuvi, the non-indigenous person ran over him.' (Juma: Text)

Contrastive topics are marked with *katu* 'well',[26] as exemplified in (268), which draws on a long conversation (a 561-sentence text) between a couple. I also provide the previous context for these sentences for reference. Among the various topics discussed, they talk about the impact of COVID within their communities and in the lives of non-indigenous people. At some point, they start contrasting hygienic habits between non-indigenous and indigenous people. The subject in the sentences, contrastive topic arguments, are marked with *=katu* 'well', in (268a) and (268b).

(268) Context: A: *The non-indigenous people are those who wash up their hands well, but the non-indigenous people are the ones who see diseases. It is them who see diseases, [but] they wash up their hands well. B: They say it is what it is. A: It is what it is, the non-indigenous people said to us. Yet, they see diseases. [And they] bring them to us. B: The non-indigenous people did not deserve to see this type of thing. They say they, the natives, don't know anything. The non-indigenous people say to us. A: They have everything to see exams. B: We will stay like that then. A: They talk about labs. Yet, they still see diseases. After that, they use gloves. They even use condoms, yet they feel itchy. They should not see these types of things [i.e., diseases]. B: We [natives] should see it. We didn't have these before. A: We are the ones who should feel itchy. A: We only had malaria, that's what we heard them [ancestors] saying. The elders used to talk about [malaria]. We didn't see them talking about the worst diseases. A: Their penises, they wash up. Their hands, they wash up. Some of them (even) make soap.*

a. **Nhãnde**=katu nd-o-vapo-i nanũngara.
1.INCL=WELL NEG-3.A-make-NEG thing
'We don't make (these) things.' (Jupaú: Text)

[25] As (267) suggests, Kawahíva does not exhibit pronominal resumption in topic constructions.

[26] Contrastive topics have also been studied in a sister language, Paraguayan Guaraní (PG) (Tonhauser 2015). She found that *=katu* 'contrast' is a contrastive topic marker in PG. She also notes that this discourse notion could be better described as the result of a confluence of factors, including prosody and word order, as well as *=katu*.

b. **Nhãnde**=katu nde-kwahav-i nanũngara.
1.INCL=WELL NEG-know-NEG thing
'We don't know (those) things.' (Jupaú: Text)

Future research is needed to show whether there are differences between argument topics vs. non-argument topics and subject topics vs. object topics for each type of topic expression. Additionally, it is necessary to ascertain whether factors like prosody contribute to the expression of topicality in Kawahíva.

3.7.6 Comparatives and superlatives

Kawahíva has two strategies to form comparative constructions. The first strategy involves a biclausal sentence with two independent clauses, in which the first clause contains the inexact quantifier *koi'i* followed by a verbal predicate, and the second clause contains a negated version of the predicate found in the first clause. This strategy is called the conjoined comparative in typological studies (Stassen 2006).

(269) Ji=rupava koi'i i-pypin, nde=rupava nd-i-pypir-i.
1.SG=hammock MANY.INEX i-be.large 2.SG=hammock NEG-i-be.large-NEG
'My hammock is larger than yours.' (lit. My hammock is very large; your hammock is not (large)) (Juma: Elicit)

The second strategy for forming comparative constructions uses the inexact quantifier *koi'i* followed by a postpositional phrase headed by *uvi* 'from' with the standard of comparison as its complement. This strategy is called the locational comparative in typological studies (Stassen 2006).

(270) Gaha koi'i yvovo nde=uvi.
3.SG.MASC.FOC MANY.INEX tall 2.SG=from
'He is taller than you.' (Juma: Elicit)

Superlatives are formed using the indefinite quantifier *koi'i* followed by a predicate.

(271) Ji=rapyja koi'i i-katu nhãnde=vyr=ipe.
1.SG=house MANY.INEX i-be.pretty 1.INCL=village=in
'My house is the prettiest in the village.' (Juma: Elicit)

3.7.7 Conjunction and disjunction

Kawahíva does not exhibit an overt grammatical mechanism for conjunction or disjunction. Rather, coordination and disjunction arise as an implicature when same-type constituents appear with a pause between them. This is the strategy that I have found to coordinate nouns. In elicitation, speakers produce two independent sentences when asked to translate conjoined verb

phrases from the target language, Portuguese. I suspect verb phrase conjunction is disallowed as they would be similar to clause chaining constructions (described in §3.7.12).

(272) a. **Ahe=hy=a**, **ahe=ruv=a** ki v-epiang ahe=ve nõmĩa.
person=mother=NMLZ person=father=NMLZ PST 3.A-see person=for FRUST
'One's mother and one's father seek out (someone) for us.' (Juma: Text)

In disjunction, like conjunction, there is a pause between the elements disjunctively conjoined. Like conjunction, I have identified this strategy for disjunction used with nouns only thus far (273).

(273) **Funai=gã**, **Cimi=gã**, j-auhu, 'i ki mãipo.
Funai=3.PL Cimi=3.PL i-seem say PST NWIT.REM.PST
'It seemed (to be) Funai or Cimi.' (Juma: Text)

Clauses can be disjunctively conjoined using the same pause strategy, as I demonstrate with the examples provided in (274). Notice the example in (274c) involves the Portuguese verb *não sei* 'I don't know'.

(274) a. Txi te gã o-juka miara, txi te gã nd-o-juka-i?
FUT REALLY 3.PL 3.A-kill meat FUT REALLY 3.PL NEG-3.A-kill-NEG
'Will they kill the game, or will they not?' (Juma: Text)

b. T-a-ho, ga 'e-i, nd-a-ho-i ji, ga 'e-i, 'i:te:ki:ko:ra.
OPT-1.SG.A-go 3.SG.MASC say-i NEG-1.SG.A-go-NEG 1.SG 3.SG.MASC say-i DIR.EVID
'Do I go, or do I not?, he said.' (Juma: Text)

c. Não.sei Kawahíva ra'em, não.sei jawara ra'em, 'i ki ahe
I.not.know indigenous TODAY.PST I.not.know jaguar TODAY.PST say PST people
mãipo nõmĩa.
WIT.REM.PST FRUST
'I don't know if (they were) indigenous people, or (they were) jaguars, we said.' (Juma: Text)

3.7.8 Modality

Modal categories and their different flavors are conveyed through various strategies, including the optative *t-*. Table 3.36 summarizes possible combinations of modal flavor (e.g., epistemic, circumstantial, and deontic) and modal force (e.g., necessity or possibility), indicating the Kawahíva strategies that are used in those contexts. Note that a verb in the optative *t-* is used in both deontic and circumstantial necessity modality contexts. Kawahíva does not seem to have any lexicalized modal verbs.

Table 3.36: Modal categories and their expression.

	necessity	possibility
deontic	Optative *t-*	
circumstantial	Optative *t-*, Desiderative *-pota*	Polarity disjunction with imperative
epistemic	Basic clause	Basic clause with *auhu* 'seem'

Deontic modality describes what is required or permitted given a set of laws, rules, or moral principles. If something is required by a law, rule, or principle (i.e., deontic necessity), of if it is merely possible or permitted according to a rule or principle (i.e., deontic possibility), it can be expressed in Kawahíva using the verb with the optative *t-*, as in (275).

(275) *Context*: *Mbotawa* is the main ceremony among the Kawahíva, when a girl reaches puberty, and is then ready to get married. During the *mbotawa* period, the men must go hunting to bring game meat for the party. One may comment about a man who did not observe his duties during the *mbotawa* period with the sentence above.

Gaha t-o-ho kavyr=ipe.
3.SG.MASC.FOC OPT-3.A-go jungle=in

'He should have gone to the jungle (to hunt).' (Juma: Elicit)

Circumstantial modality describes what is possible or necessary due to a particular set of circumstances. If something must occur due to the current conditions (i.e., circumstantial necessity), it is expressed using the optative prefix *ta-*, or the desiderative suffix *-pota*.

(276) a. Akwẽ nã'ẽ, t-a-nhãtxym a=jie.
wait first OPT-1.SG.A-sneeze ?=1.SG
'Hold on, I have to sneeze.' (Juma: Elicit)

b. I-kuru-pota-ruhu jie.
i-pee-want-AUG 1.SG
'I want to pee.' (Juma: Elicit)

If something is possible given a particular set of conditions (i.e., circumstantial possibility), it can be expressed using a polarity disjunction where the verb is in the imperative.

(277) *Context*: Every time I need to sleep in the town of Lábrea, I have two options for lodging: either sleep in the Catholic house or go to a Funai employer's home. At the Catholic house, the host tells me I can stay with them or go elsewhere.

"Avo-ete e-kin!", ehẽa ji=ve; "A'ero e-vo-vo-pota,
here-REALLY 2.SG.IMP-sleep 3.SG.FEM.FOC 1.SG=to then 2.SG.IMP-go-VO-want
e-ho a'ero!", ehẽa ji=ve.
2.SG.IMP-go then 3.SG.FEM 1.SG=to

'"Sleep right here!" She (said) to me; "or if you want to go, go then!", she (said) to me.' (Juma: Elicit)

Finally, epistemic modality describes what is possible or necessary due to one's particular knowledge or evidence. If something must be true given the available evidence and knowledge about the world (i.e., epistemic necessity), it is expressed using a basic clause without any tense or mood marking.

(278) *Context*: Mandeí is a punctual schoolteacher who must be at school at 7am. I stop by Mandeí's house one day past 7am, and Mandeí isn't there, so I say:

Imõkwatxijauhua Mandeí=hẽa o-mõndo kurumi=gã pe.
writing Mandeí=3.SG.FEM 3.A-give kid=3.PL to

'Mandeí is teaching the kids.' (Lit.: Mandeí is giving writings to the kids) (Juma: Elicit)

If something is possible given the evidence and knowledge about the world (i.e., epistemic possibility), it can be expressed by using a basic clause without tense and mood marking accompanied by the word *auhu* 'seem'.

(279) *Context*: The schoolteacher Mandeí regularly misses school, so the students never know whether she will be there on a given day.

U-hu hẽa auhu kwarupi i-mõndo-vo imõkwatxijauhua.
3.A-come 3.SG.FEM seem today i-give-VO writing

'It seems that she comes today to give class.' (Juma: Elicit)

3.7.9 Complement clauses

There are two main complementation strategies in Kawahíva: the first of these involves a clause that can include the arguments of the verb and most of the finiteness distinctions found in matrix clauses, and is embedded under verbs like *kwaham* 'know, remember'.[27] The second strategy involves a structurally smaller clause, consisting of just a verb phrase embedded under the verb *eko* 'be'. The former complement clause is nominalized with the invariant nominalizer *=a*. In contrast, the complement clause under *eko* 'be' may either be nominalized with the same *=a* or appear with suffix *-vo*, which also marks verbs in purposive clauses (described in §3.7.10) and clause chaining constructions (described in §3.7.12). Table 3.37 summarizes the differences between the two types of clausal complements, which I label Type I and Type II.

Table 3.37: Complement clause strategies.

	Nominalization	Tense	Agreement
Type I	✓	✓	Object agreement only
Type II	✓	✗	✗

[27] The verb *'i* 'say' does not embed complement clauses in Kawahíva as it appears with a direct quote.

3.7.9.1 Type I complement clauses

Type I complement clauses appear under verbs like *kwaham* 'know, remember', *hepia* 'see', *koite* 'like', *hendum* 'hear', and *rovia* 'believe'. Examples (280a-c) illustrate complement clauses under the first two verbs. I enclose the complement clause within brackets for ease of identification.

(280) a. A-**kuha** tuma=hẽa [ka ji=repiak=a].
3.A-know mother=3.SG.FEM 3.SG.MASC 1.SG=see=NMLZ
'Mother knows he saw me.' (Jupaú: Elicit)

b. Ere-**pia** ki [nde jie tapi'ira juka=a] ko.
2.SG.A-see PST 2.SG 1.SG tapir kill=NMLZ REAL
'You know I killed a tapir.' (Jupaú: Elicit)

c. A-**koite** jie [hẽa=rehe hẽa kwahav=a].
1.SG.A-like 1.SG 3.SG.FEM=at 3.SG.FEM dance=NMLZ
'I liked that she danced.' (Jupaú: Elicit)

An important morphosyntactic property of these structures is they are nominalized – they must appear with the nominalizer *=a*. The example in (281) forms a minimal pair with the example in (280a), but the former is ungrammatical due to the absence of the nominalizer *=a*.

(281) *A-**kuha** tuma=hẽa [ka ji=repiak].
3.A-know mother=3.SG.FEM 3.SG.MASC 1.SG=see
'Mother knows he saw me.' (Jupaú: Elicit)

The clausal complement of these predicates exhibits some but not all finiteness markers of the matrix clause. For instance, they exhibit agreement markers, although these are restricted to object agreement markers.

(282) a. O-kwaha ki ga [ji **nde**=repiak=a].
3.A-know PST 3.SG.MASC 1.SG 2.SG=see=NMLZ
'He knew I saw you.'

b. Ere-pia ki nde [jawara tapi'ira **(*o)**-juka=a] ko.
2.SG.A-see PST 2.SG jaguar tapir 3.A-kill=NMLZ REAL
'You saw the jaguar killing a tapir.'

Type I complement clauses exhibit a tense distinction but with important differences from the matrix clause. The temporal distinction is a simple two-way difference between 'past' with *-ver* (~ *-kwer*) versus 'future' with *-ham* in Jupaú, or *-ram* in Juma, as in examples (283); recall that matrix clauses exhibit a graded past tense distinction, as described in §3.5.2. Note the morphological exponents of tense are different: bound morphemes in complement clauses, but particles in matrix clauses. Example (284) shows the future tense particle *txi* 'future', which occurs in matrix clauses, is not allowed within a complement clause.

(283) a. A-epia ji [tapy'ynha=ga ji=rendyr=a=hẽa
1.SG.A-see 1.SG non.indigenous=3.SG.MASC 1.SG=sister=NMLZ=3.SG.FEM
potar-**aver**=a].
want-NOM.PST=NMLZ
'I saw the non-indigenous person dating my sister.' (Jupaú: Elicit)

b. A-kwaha ji [tapy'ynha=gã yva rety-**ham**=a].
1.SG.A-know 1.SG non.indigenous=3.PL tree take.down-NOM.FUT=NMLZ
'I know the non-indigenous people will take down trees.' (Jupaú: Elicit)

(284) *Arovia jie [Puré=ga **txi** tapi'ira juka nãhẽ].
1.SG.A-believe 1.SG P.=3.SG.MASC FUT tapir kill=NMLZ POT
'I believe Puré will kill a tapir soon.' (Juma: Elicit)

Type I complement clauses also allow negation but with a different morphological exponent. Instead of the matrix clause, circumfix *n-..-i*, these complement clauses use the suffix *-e'ym* (Juma) (*-i'im*, in Jupaú), which also marks negation on nouns (as described in §3.3.4).

(285) A-epia ji [kurumi=gã pira 'u-**'im**=a].
1.SG.A-see 1.SG child=3.PL fish eat-NOM.NEG=NMLZ
'I saw the children didn't eat fish.' (Jupaú: Elicit)

These complement clauses lack mood distinctions, as the examples in (286) demonstrate. While the baseline example in (286a) shows no mood marking a complement clause embedded under *rovia* 'believe', the examples (286b-c) show the irrealis particle is not allowed in the same clause.

(286) a. A-rovia jie [Puré=ga tapi'ira juka=a].
1.SG.A-believe 1.SG P.=3.SG.MASC tapir kill=NMLZ
'I believe Puré killed a tapir.' (Juma: Elicit)

b. *A-rovia jie [Puré=ga **po** tapi'ira juka=a].
1.SG.A-believe 1.SG P.=3.SG.MASC IRR tapir kill=NMLZ
'I believe Puré (might have) killed a tapir.' (Juma: Elicit)

c. *A-rovia jie [**po** Puré=ga tapi'ira juka=a].
1.SG.A-believe 1.SG IRR P.=3.SG.MASC tapir kill=NMLZ
'I believe Puré (might have) killed a tapir.' (Juma: Elicit)

Other clause-like properties of Type I complement clauses are the use of the transitive applicative *-okan*, as shown in example (287), and adverbial suffixes, including *-ipe* 'already' (288a) and *-vere* 'almost' (288b).

(287) A-kwaha jie ji=kuvyr=a=ga ga=rembiriko=hẽa
3.A-know 1.SG 1.SG=young.brother=NMLZ=3.SG.MASC 3.SG.MASC=spouse=3.SG.FEM
epia-**ok**=a apin=a=ga=pe.
see-APPL=NMLZ father=NMLZ=3.SG.MASC=to

'I know my young brother introduced her wife to my father.' (Lit.: I know my young brother showed his wife to my father) (Jupaú: Elicit)

(288) a. A-kwaha jie ji=kuvyr=a=ga imbiriko=hẽa
3.A-know 1.SG 1.SG=young.brother=NMLZ=3.SG.MASC spouse=3.SG.FEM
ero-ho-**ipe**=a epiak=a apin=a=ga=pe.
SOC.CAUS-go-ALREADY=NMLZ see=NMLZ father=NMLZ=3.SG.MASC=for
'I know my young brother already took his wife for my father to see.' (Jupaú: Elicit)

b. A-kwaha jie [ji=kuvyra=ga ereko-**vere**=a
3.A-know 1.SG 1.SG=young.brother=3.SG.MASC marry-ALMOST=NMLZ
Amondawa=hẽa].
amondawa=3.SG.FEM
'I know my young brother almost married an Amondawa woman.' (Jupaú: Elicit)

Finally, Type I complement clauses exhibit a strict SOV word order. I provide the examples in (289b-c) to show that word orders different from SOV are ungrammatical in these clauses, which must be SOV, as in (289a).

(289) a. A-hepia ki jie [ji=ruv=a=ga pira **ʻu=a**].
1.SG.A-see PST 1.SG 1.SG=father=NMLZ=3.SG.MASC fish eat=NMLZ
'I saw [my father eating fish].' (Juma: Elicit)

b. *A-hepia ki jie [**ʻu=a** ji=ruv=a=ga pira].
1.SG.A-see PST 1.SG eat=NMLZ 1.SG=father=NMLZ=3.SG.MASC fish
'I saw [my father eating fish].' (Juma: Elicit)

c. *A-hepia ki jie [ji=ruv=a=ga **ʻu=a** pira].
1.SG.A-see PST 1.SG 1.SG=father=NMLZ=3.SG.MASC eat=NMLZ fish
'I saw [my father eating fish].' (Juma: Elicit)

3.7.9.2 Type II complement clauses

Type II complement clauses are embedded under the verb *eko* 'be' and consist of just a verb phrase. Morphologically, the complement clause is marked with *=a* or *-vo* and allows reduplication and causative morphology. The difference between the two types is the object position in the complement contains a transitive verb.

In type II complement clauses marked with the nominalizer *=a*, the semantic object appears right next to the verb in the complement clause, as in (290), and is ungrammatical if it appears preceding *eko* 'be', as in (291).[28,29]

[28]I briefly address the lack of subject agreement in matrix clauses, including cases like (290), in §3.7.16.

[29]The progressive/imperfective translations provided to these and other examples in this section by speakers bear some similarities with the habitual semantics of constructions with the auxiliary *eko* 'be', as described in (§3.4.10).

(290) Ore reko-i, [**ore=rapyj=a** pehi=**a**].
1.EXCL be-i 1.EXCL=house=NMLZ sweep=NMLZ
'We were sweeping our house.' (Juma: Text)

(291) *Aramẽ ki ji **i-tapyja** reko-i [i-pehi=**a**].
after PST 1.SG 1.SG.COR-house be-i i-sweep=NMLZ
'After that I was sweeping my house.' (Juma: Elicit)

In contrast, when the complement marked is marked with *-vo*, the semantic object of the verb in the complement clause may or may not appear adjacent to it. The semantic object of the verb in the complement clause, *kamipiva* 'coffee', is adjacent to the embedded verb in (292), but not in (293), where the complement *matera* 'food, thing' precedes the matrix verb *eko* 'be'. It is much more common for the object to appear in the position where it precedes *eko* 'be'; in fact, the example (292) is the single spontaneous speech example in the corpus where the object is linearly adjacent to the embedded verb.

(292) Ji reko-i, [**kamipiva**, j-apo-vo], 'i:te:ki:kora.
1.SG be-i coffee i-make-VO DIR.EVID
'I was making coffee.' (Juma: Text)

(293) Ji **matera** reko-i [j-apo-**vo**], 'i:te:ki:ko:ra.
1.SG food be-i i-make-VO DIR.EVID
'I was making food.' (Juma: Text)

Additionally, Type II complement clauses allow reduplication morphology and the causative *mõ-*. An example of reduplication morphology and one of causative morphology in Type II complement clauses are provided in (294) and (295), respectively.

(294) Ore reko-i [i-**kytxi-kytxi-vo**].
1.EXCL be-i i-RED-cut-VO
'We were cutting (it).' (Juma: Text)

(295) Aramẽ ki Kunhavé=hẽa te, a'oa reko-i [i-**mõ**-tyryry-**vo**], 'i:te:ki:ko:ra.
after PST K.=3.SG.FEM REALLY meat be-i i-CAUS-drip-VO DIR.EVID
'After that, Kunhavé was frying the meat.' (Juma: Text)

Finally, note this syntactic behavior of *eko* 'be' is quite distinct from the one described when this form is used as an auxiliary to express the habitual in (3.4.10). In that capacity, *eko* 'be' is strictly in clause-final position and must bear the suffix *-vo* and a coreferential marker that covaries with the subject features. I argue these differences reflect distinct selective requirements of the same verb: as an auxiliary, *eko* 'be' requires a simple verb phrase as its complement; however, as a non-auxiliary, it takes a complement clause as the complement.

3.7.10 Adjunct clauses

Most adjunct clauses are distinguished by the different markings that surface on the verb of the adjunct clause. Table 3.38 summarizes the various types of adjunct clauses.

Adjunct clause	Marker
Temporal overlap, Reason, and Conditional	*-mẽ*
After-	*-irẽ*
Purpose	*=a* and *-vo*
Before-	*renõnde*

Table 3.38: Adverbial suffixes.

Temporal overlap, Reason, and Conditional adjunct clauses are marked with *-mẽ*.

(296) [Amãna kyr=a=mẽ], ki ahe nd-o-avyky-i matera.
rain rain=NMLZ=WHEN PST people NEG-3.A-move-NEG thing
'When it rains, we do not work out things.' (Juma: Elicit)

(297) Nd-o-poravyky-i ga [o-karuvar=a=mẽ]
NEG-3.A-work-NEG 3.SG.MASC 3.COR-sick-NMLZ=BECAUSE
'He does not work because he is sick.' (Juma: Elicit)

(298) A'ero [amãna kyr-e'ym=a=mẽ], ore ho-i pevo, kavyr-ipe,
then rain rain-NOM.NEG=NLMZ=IF 1.EXCL go-i there jungle-in
oro-ki=a.
1.EXCL.COR-sleep=NMLZ
'If it does not rain, we will go there, into the jungle, to sleep.' (Juma: Text)

After-clauses are marked with the suffix *-irẽ*.

(299) E-nduvun [e-võnhã mbovur-irẽ]!
2.SG.IMP-spit 2.SG.COR-tooth remove-AFTER
'Spit after you remove your tooth!' (Juma: Elicit)

Purpose clauses are marked either with the enclitic *=a* or the suffix *-vo*, as shown in examples (300), where they are bracketed. The choice depends on a semantic presupposition vs. entailment distinction: in purpose clauses marked with *=a*, the event is a presupposition, whereas, in purpose clauses marked with *-vo*, it is an entailment. In (300a), the speaker can only presuppose that the event of hearing a peccary happened, while in the scenario in (300b), this is confirmed.[30]

[30]Baranger (2022:1) describes a similar distinction in Mbya-Guaraní purposive clauses, which are marked with either *-aguã* or *-vy*: "[...] *Aguã*-marked clauses portray an intended hypothetical outcome of the event or state-of-affair, [...] while -vy clauses consistently trigger a result interpretation, entailing that the intended state-of-affairs was accomplished".

(300) a. O-ho ga [tajahua rendup=a].
2.SG.A-go 3.SG.MASC peccary hear=NLMZ
'He went to hear peccary.' (Elicit)
✓ Someone went hunting and has not returned yet.
✗ Someone came back from hunting and said they heard a peccary.

b. O-ho ga [tajahua rendup(a)-vo].
2.SG.A-go 3.SG.MASC peccary hear-VO
'He went to hear peccary.' (Elicit)
✓ Someone came back from hunting and said they heard a peccary.
✗ Someone went hunting and has not returned yet.

Before-clauses are marked with *renõnde*.

(301) [Ji=ho=a=rerõnde=koty] ga 'e-jav-i.
1.SG=go=NMLZ=BEFORE=side 3.SG.MASC say-AGAIN-i
'Before I went, he said again.' (Juma: Text)

Adjunct clauses may be initial, medial, or final within the matrix clause. The adjunct clauses marked with *-mẽ* in examples (302-304) illustrate these different positions of the adverbial clause.

(302) [Karuvara kwa-pav-ahiv=a=mẽ] tehe txi ji ho-i, 'i ki ga
disease pass-COMPLETELY-VERY=NMLZ=WHEN IDLY FUT 1.SG go-i say PST 3.SG.MASC
ko.
REAL
'Only when the coronavirus is gone completely, I will leave, he said.'
(Juma: Text)

(303) A'ea ki ahe [mater-e'ym=a=mẽ] o-'u nũmĩa.
this.one PST people thing-NOM.NEG=NMLZ=WHEN 3.A-eat FRUST
'This one (is) what we eat when there is nothing.' (Lit.: That is what we, when there is nothing, eat) (Juma: Text)

(304) Tapy'ynha=gã i-mbuhu-mbuhu-tehe ore=ve mãipo nõmia
non.indigenous=3.PL i-RED-give-IDLY 1.EXCL=to WIT.DIST.PST FRUST
[i-'u-e'ym=a=mẽ].
i-eat-NOM.NEG=NMLZ=WHEN
'The non-indigenous people (used to) give it to us in the old days when we didn't eat (them).'
(Jupaú: Text)

Stacking adjunct clauses in initial, medial, and final positions is also possible, as the *-mẽ* clauses examples demonstrate in (305-307).

(305) [Karuvara mõmĩn-ahiv=a=mẽ], [i-kwa-pa-hiv=a=mẽ], po ji
disease end-VERY=NMLZ=WHEN i-pass-COMPLETELY-pass-VERY=NMLZ=WHEN IRR 1.SG
ho-i nãhẽ.
go-i POT
'When the disease ends, completely gone, I will leave.' (Juma: Text)

(306) [Ahe i-kwahav=a=mẽ], [matera renduv=a=mẽ], aí o-ko matera ahe
people i-know=NMLZ=WHEN thing hear=NMLZ=WHEN then 3.A-be thing person
rehe.
at
'When we know it, when (we) understand things, then things happen to us.'
(Juma: Text)

(307) Marã po txi nãhẽ, [koi'ir=a=mẽ], [karuk=a=mẽ] nãhẽ.
how IRR FUT POT later=NMLZ=WHEN afternoon=NMLZ=WHEN POT
'(I don't know) how it will be later in the afternoon.' (Lit.: (I don't know) how it will be when it is later, in the afternoon) (Juma: Text)

Adjunct clauses allow few finiteness distinctions, including the causative *mõ-* and aspect-related adverbial suffixes, such as *-pam* 'completely'

(308) A-ho jie [televisão **mõ**-mbik=a] ji=rapyja pupe.
1.SG.A-go 1.SG TV CAUS-turn.off=NMLZ 1.SG=house inside
'I will into my house to turn off the TV.' (Jupaú: Elicit)

(309) [Pe j-u-**pa**-mẽ=hẽ], po txa-ho kotxi, e-'i-te ka.
2.PL i-eat-COMPLETELY=WHEN=FOC IRR LET'S-go later i-say-REALLY 3.SG.MASC
'When y'all eat (it) completely, let's go, he said.' (Juma: Text)

Negation is also allowed in adjunct clauses but with an important distinction. Most adjunct clauses mark negation with *-e'ym* [Juma] or *-i'im* [Jupaú], as shown with the *when*-clause in (310). This negation marker is also used in nominalized complement clauses and non-clausal nominals. In contrast, purpose clauses mark negation with *-imẽ*, as shown in the example with *-vo* in (311).[31]

(310) [Gã=vyr=ipe ji=ho-'im-amẽ], ki gã ho-i o-kahup=a oi'i.
3.PL=village-in 1.SG=go-NOM.NEG-WHEN PST 3.PL go-i 3.A-hunt=NMLZ some.time.ago
'When I didn't go to their village, they went hunting.' (Jupaú: Elicit)

[31] A prefix *ta-* is also marked on the negated verb in this example, as well as the non-negated form of the verb in the purposive clause (not shown). An identical morpheme surfaces in constructions with deontic ('I should eat fish') and necessity modal meanings (I would like to eat fish). More research is necessary to pin down the semantics of *ta-* in and outside purposive clauses.

(311) O-nhã ka [t(a)-'u-**imẽ** pira jo-vo].
3.A-run 3.SG.MASC OPT-eat-NOM.NEG fish ?-VO
'He ran to not eat fish.' (Jupaú: Elicit)

However, adjunct clauses lack several finiteness distinctions, including realis/irrealis and tense. For instance, the ungrammatical examples in (313-314) form minimal pairs with the grammatical baseline example in (312), which does not include the particle *po* 'irrealis'.

(312) [Humaitá pe ore matera 'u-ramẽ], po ore mãrãngatu ko.
Humaitá to 1.EXCL food eat-IF IRR 1.EXCL how.exactly REAL
'If we had eaten something in Humaitá, we (would be), 'what really happened?' (Jupaú: Elicit)

(313) *[**Po** Humaitá pe ore matera 'u-ramẽ], po ore mãrãngatu ko.
IRR Humaitá to 1.EXCL food eat-IF IRR 1.EXCL how.exactly REAL
'If we had eaten something in Humaitá, we (would be), 'what really happened?' (Jupaú: Elicit)

(314) *[Humaitá pe **po** ore matera 'u-ramẽ], po ore mãrãngatu ko.
Humaitá to IRR 1.EXCL food eat-IF IRR 1.EXCL how.exactly REAL
'If we had eaten something in Humaitá, we (would be), 'what really happened?' (Jupaú: Elicit)

Adjunct clauses exhibit nominal tense morphology with *-ver* 'nominal past' and *-(a)ham* 'nominal future'. For instance, while the baseline *when*-clause in (315) does not include *ki* 'past' and is grammatical, minimal pairs like (317) and (318) are ruled out. Consultants consistently offer the example without the tense particle in (315) as a solution.

(315) [Humaitá pe ore matera 'u-irẽ], ki ore mãrãngatu ko.
Humaitá to 1.EXCL food eat-AFTER PST 1.EXCL how.exactly REAL
'After we ate something in Humaitá, we (were), "What really happened?"'
Intended: After we ate something in Humaitá, we got stomachache (Jupaú: Elicit)

(316) *[**Ki** Humaitá pe ore matera 'u-irẽ], ki ore mãrãngatu ko.
PST Humaitá to 1.EXCL food eat-AFTER PST 1.EXCL how.exactly REAL
'After we ate something in Humaitá, we (were), "What really happened?"'
Intended: After we ate something in Humaitá, we got stomachache (Jupaú: Elicit)

(317) *[Humaitá pe **ki** ore matera 'u-irẽ], ki ore mãrãngatu ko.
Humaitá to PST 1.EXCL food eat-AFTER PST 1.EXCL how.exactly REAL
'After we ate something in Humaitá, we (were), "What really happened?"'
Intended: After we ate something in Humaitá, we got stomachache (Jupaú: Elicit)

Finally, the morphological past tense marker *-ver* is not accepted in adjunct clauses. The following example illustrates the presence of this marker in a *when*-clause, which makes the sentence bad. The consultant offered a fix, which consisted of removing this suffix.

(318) *Gã=vyr-ipe ji=ho-**ver**-amẽ, ki ore ho-i ore=kahup=a.
3.PL=village=in 1.SG=go-NOM.PST-WHEN PST 1.EXCL go-i 1.EXCL=hunt=NMLZ
'When I went to their village, we went hunting.' (Jupaú: Elicit)

In brief, I have shown that the following properties characterize adjunct clauses: they can be stacked; they exhibit fewer finiteness properties, allowing causative, negation, and aspect but disallowing mood and tense.

3.7.11 Relative clauses

Relative clauses in Kawahíva are post-nominal: the head precedes the relative clause, leaving a gap in the canonical position of the relative head, as illustrated by the object gap (marked with a slash) in the bracketed structure in (319a). There is no relative pronoun or complementizer. Additionally, relative clauses bear the nominalizing enclitic *=a*, as shown by the minimal pair in (319), making them nominalized relative clauses.

(319) a. A-epia ki jie **y'va** [tapy'ynha=gã __
1.SG.A-see PST 1.SG tree non-indigenous=PL
remb-i-tyk-aver=a].
WH.OBJ.I-i-take.down-NOM.PST=NMLZ
'I saw the trees which the non-indigenous people took down.'

b. *A-epia ki jie **y'va** [tapy'ynha=gã __ remb-i-tyk-aver].
1.SG.A-see PST 1.SG tree non-indigenous=PL WH.OBJ.I-i-take.down-NOM.PST
'I saw the trees which the non-indigenous people took down.'

A core morphosyntactic property of relative clauses is that verbs in relative clauses obligatorily bear one of the morphemes listed in Table 3.39.

RELATIVE HEAD	MARKER	RELATIVE HEAD	MARKER
trans. subjects	*-har*	postpositional obliques	*-hav* & *-var*
trans. objects	*remb-* & *-pyr*	intrans. subjects	*-va'e*

Table 3.39: Kawahíva relative clause morphemes.

The choice of marker from Table 3.39 depends on the grammatical function of the head within the relative clause. In particular, *-har* is used when the relative head (or the gapped position within the nominalized relative clause) is a transitive subject. The morphemes *remb-* and *-pyr* are used when the relative head is a transitive object. In turn, *-hav* and *-var* are used when the relative head is the complement of a postposition that heads an oblique PP. Finally, *-va'e* is

employed when the relative head is the subject of an intransitive verb. In cases where there are two morphemes for the same gap position (i.e., object position and complement of postpositions), the choice boils down to nuanced aspectual differences between *remb-* and *-pyr*, and in the case between *-hav* and *-var*, the choice depends on the postposition in the oblique phrase. In Chapter 5, I provide a thorough description of the distribution of these morphemes.

In addition to headed relative clauses, Kawahíva also exhibits headless relative clauses, an example of which is shown in (320).

(320) A-epia ki jie [tapy'ynha=gã remb-i-tyk-aver=a].
1.SG.A-see PST 1.SG non.indigenous=PL WH.OBJ.II-i-take.down-NOM.PST=NMLZ
'I saw the ones (= the trees) that the non-indigenous people took down.' (Jupaú: Elicit)

To the best of my knowledge, the choice between a headed and headless relative clause depends on the referent of the latter being *active* and/or *given* in the discourse (Chafe 1976; Lambrecht 1994), easily recovered by its previous mentioning in the discourse.

The description of bracketed structures like (319a) above as relative clauses contrasts with the common characterization of similar structures in the literature on sister languages to Kawahíva. In this literature, the mentioned bracketed structure is a nominalization; juxtaposed with a noun, nominalizations function as an adnominal structure. Relativization, then, is a byproduct of nominalizations. In Chapter 5, I argue that relativization is an independent process in Kawahíva and not a function of nominalizations, although nominalization plays a major role in the expression of relativization.

3.7.12 Clause chaining

Sequential actions are often conveyed using clause chaining constructions, which follow a clause with an intransitive or transitive verb with one or more adjuncts and mark the verb of the chained clauses with the suffix *-vo*. In example (321), the chained clause is within brackets and has an intransitive verb.[32]

(321) Ore rur-i, tapyja=upe, [oro-puta-vo].
1.EXCL come-i house=inside 1.EXCL.COR-stop-VO
'We came into the house and stopped.'

Several different verbs, intransitive or transitive, have been found in the first position in the clause-chained construction. As the table summarizes, all valency combinations between a matrix and a chained verb are allowed.

Table 3.40: Combinations of matrix and subsequent verbs.

	SUBSEQUENT INTRANSITIVE	SUBSEQUENT TRANSITIVE
MATRIX INTRANSITIVE	✓	✓
MATRIX TRANSITIVE	✓	✓

[32]I briefly address the phenomenon of lack of subject agreement in matrix clauses, as in (321), in §3.7.16.

Table 3.41 summarizes all the verbs that have been found in the matrix position in a clause chaining construction, whereas Table 3.42 summarizes the verbs that head the chained clause and follow the main clause.

Table 3.41: Types of matrix verbs in clause chaining constructions.

MATRIX INTRANSITIVE	MATRIX TRANSITIVE
ho 'go'	*kwav* 'pass by'
(t/r/nd)ur 'come'	*gyahav* 'cross'
vag 'return'	*upir* 'lift'
eko 'be'	*ereko* 'stay with'
kwaham 'know'	*erun* 'bring'
ja 'laugh (at)'	*erokwav* 'pass with'
nhãi 'run'	

Table 3.42: Types of V2 verbs.

V2 INTRANSITIVE	V2 TRANSITIVE
pyta 'stay'	*mõndo* 'put, give'
vo 'go'	*vo* 'eat'
eko 'be'	*ero* 'take'
jo 'come'	*potxi-potxi* 'tie repeatedly'
	apitxi 'kill'
	po'o 'peel (bird)'
	apo 'make'
	rerovo 'go with'
	kytxi 'cut'
	mbotyryry 'fry'
	ryrõ 'await'
	pyta 'stay'
	pyhyk 'catch'
	myty 'pull'
	mbojo 'make come'
	joverovovo 'one and the other go'

In clause chains with shared-subject, the subjects of chained adjunct intransitive clauses are null and marked with the set of coreferential markers, which co-vary with the features of the matrix subject, as the example in (322) illustrates.

(322) Txiro ore kwav-i, [oro-vo-vo], [oro-puta-vo],
CONTINUOUSLY 1.EXCL pass.by-I 1.EXCL.COR-go-VO 1.EXCL.COR-stop-VO
oro-jup=a.
1.EXCL.COR-be=NMLZ
'We passed by, went, stopped, to be (there).' (Juma: Text)

In contrast, the subjects of chained adjunct transitive clauses are null and not marked by the coreferential paradigm, as in (323).

(323) Aramẽ ki hẽa, ga=’y=hẽa, mbiara rero-’e-i
after.that PST 3.SG.FEM 3.SG.MASC=mother=3.SG.FEM meat SOC.CAUS-leave-i
[i-mõndo-vo] i-mõkup=a, ‘i:te:ki:ko.
i-put-VO i-heat-NMLZ DIR.EVID
‘After that, his mother took out the meat and put it to heat up.’ (Juma: Text)

3.7.13 Coreferential marking

Coreference between the verbal subject and the complement of a postposition (324), a possessor (325), and the auxiliary verb *eko* ‘be, live, stay’ (326) is indicated with the prefixes in Table 3.43.

Table 3.43: Coreferential markers.

	PREFIXES
1st person singular	i- ~ it-
2nd person singular	e-
1st person plural inclusive	nhãnde-
1st person plural exclusive	oro-
2nd person plural	pe-
3rd person singular/plural	v- ~ o-

(324) a. Ere-mõpu nde **e**-ji-ehe.
2.SG.A-shoot 2.SG 2.SG.COR-REFL=at
‘You shot at yourself.’ (Juma: Elicit)

b. T-a-hepiang nde=kyva, e-’i ki anhãnga **o**-jo=upe.
OPT-1.SG.A-see 2.SG=lice i-say PST spirit 3.COR-RECP=on
‘“Let me see your lice”, the spirits said to each other.’ (Juma: Text)

(325) O-mõmbe’u ga **v**-uva=ga repiag-aver=a.
3.A-tell 3.SG.MASC 3.COR-father=3.SG.MASC see-NOM.PST=NMLZ
‘He told (to us) he saw his own father.’ (Juma: Elicit)

(326) a. Ji=repia ka **o**-ko-vo.
1.SG.B=see 3.SG.MASC 3.COR-be-VO
‘He has watched me.’ (Jupaú: Elicit)

b. *Ji=repia ka **it**-eko-vo.
1.SG.B=see 3.SG.MASC 1.SG.COR-be-VO
‘He has watched me.’ (Jupaú: Elicit)

3.7.14 Cross-clausal coreference

Coreference between the subject of the matrix clause and the subject or object of a dependent clause is distinguished from coreference between the object of the matrix clause and the subject or object of a dependent clause. These combinations are summarized in Table 3.44. Coreference between the matrix clause subject and the subject or object argument in the dependent clause is indicated on the dependent verb with markers from the coreferential paradigm discussed in the previous section in §3.7.13. In contrast, coreference between the object of the matrix clause and the subject or object argument in the dependent clause is indicated on the dependent verb by the prefix *i-*. I provide examples of these combinations in (327); dependent clauses are shown in brackets.

Table 3.44: Cross-clausal coreference marking on dependent verbs.

↓ MATRIX DEPENDENT →	COREFERENT SUBJECT	COREFERENT OBJECT
SUBJECT	coreferential marking	coreferential marking
OBJECT	i-	i-

(327) a. Nd-o-juka-i ga tapi'ira [**o**-nhãr=a=mẽ].
NEG-3.A-kill-NEG 3.SG.MASC tapir 3.SG.COR-run=NMLZ=BECAUSE
'He did not kill the tapir because he ran (away).' (matrix subject = dependent subject)

b. [**O**-mõngyjip=a=inẽ], javatxinga j=u'u-i.
3.COR-scary=NMLZ=AFTER dog 1.SG.B=bite=i
'After I scared (it), the dog bit me.' (matrix subject = dependent object)

c. Nd-o-juka-i ga tapi'ira [**i**-nhãr=a=mẽ].
NEG-3.A-kill-NEG 3.SG.MASC tapir i-run=NMLZ=BECAUSE
'He did not kill the tapir because it ran (away).' (matrix object = dependent subject)

d. [Ji **i**-mõngyjip=a=inẽ], e-nupã javatxinga rehe!
1.SG i-scary=NMLZ=AFTER 2.SG.IMP-hit dog at
'After I scare (it), hit the dog!' (matrix object = dependent object)

In multiple coreference, there can be co-occurrence between coreferential markers and the prefix *i-*. However, this co-occurrence arises only in one combination, namely, when coreference is between arguments with the same grammatical role in the matrix and dependent clause. In (328), matrix and dependent subjects and matrix and dependent objects are coreferent. As a result, the dependent verb exhibits the coreferential marking to indicate the former, and *i-*, to indicate the latter coreferential pattern. In contrast, if the co-referring arguments have different grammatical roles in both clauses, only the coreferential marking is allowed, as in (329).

(328) E-mõmbon pira [**e**-imb-**i**-'u-e'ym=a]!
2.SG.IMP-throw fish 2.SG.COR-WH.OBJ.I-eat-NEG=NMLZ
'Throw (out) the fish that (you) did not eat!' (matrix subject = dependent subject; matrix object = dependent object) (Juma: Elicit)

(329) a. E-juka kava [**e**-pi-har=a]!
2.SG.IMP-kill wasp 2.SG.COR-bite-WH.TRANS.SUBJ=NMLZ
'Kill the wasp that bit (you)!' (matrix subject = dependent object; matrix object = dependent subject) (Juma: Elicit)

b. *E-juka kava [**e-i**-pi-har=a]!
2.SG.IMP-kill wasp 2.SG.COR-i-bite-WH.TRANS.SUBJ=NMLZ
'Kill the wasp that bit (you)!' (matrix subject = dependent object; matrix object = dependent subject) (Juma: Elicit)

Finally, co-referential markers do not co-occur with object agreement nor the prefix *i-*, as shown in (330). Dependent clauses do not exhibit subject agreement; therefore, co-occurrence between them and co-referential markers or the prefix *i-* is ruled out for this reason.

(330) a. A-y'u jie kaminha **(*ji)=i**-karuvar=a=mẽ.
1.SG.A-drink 1.SG manioc.beer 1.SG.B=1.SG.COR-be.sick=NMLZ=BECAUSE
'I drank *kaminha* because I was sick.' (matrix subject = matrix subject) (Juma: Elicit)

b. [Ji **(*nde)=i**-mõngyjip=a=inẽ], ga nupã-i nde=rehe.
1.SG 2.SG.B=i-scary=NMLZ=AFTER 3.SG.MASC hit-i 2.SG=at
'After I scared you, he hit you.' (Juma: Elicit)

3.7.15 The pattern of i-prefixation

In Kawahíva, there is a pattern of prefixation with *i-* on verbs in contexts of object extraction. The most straightforward instances of this pattern are cases where an object is displaced in sentences with a clause-final verb, including dependent clauses. Consider the examples in (331), which include a complement clause within brackets. In (331a), there is no object extraction out of the embedded clause. Extraction of the object is seen in (331b); as a result, the object surfaces in the initial position of the matrix clause. Additionally, it shows the prefix *i-* on the dependent verb. Omitting the prefix is ungrammatical, as shown in (331c).

(331) a. Ere-piang nde [apinaga pira 'u-a] rai'i.
2.SG.A-see 2.SG father fish eat-NMLZ YESTERDAY.PST
'You saw Father eating fish.'

b. Gara nde ere-piang [apinaga i-'u-a] rai'i?
what 2.SG 2.SG.A-see father i-eat-NMLZ YESTERDAY.PST
'What did you see father eating?'

c. *Gara nde ere-piang [apinaga 'u-a] rai'i?
what 2.SG 2.SG.A-see father eat-NMLZ YESTERDAY.PST
'What did you see father eating?'

For other contexts of object displacement that trigger the prefix *i-*, see Chapter 4.

3.7.16 The pattern of i-suffixation and loss of subject agreement in matrix clauses

Kawahíva matrix verbs exhibit two morphosyntactic phenomena related to the extraposition of adjuncts to clause-initial position. The first is the appearance of a verbal suffix *-i* in contexts when the clause-initial position is filled by a postpositional phrase (PP), as in (332), a discourse particle, as in (333), or an adjunct clause, as in (334).

(332) a. A-hepia ki jie gã **oi'i=ve**.
1.SG.A-see PST 1.SG 3.PL some.time.ago=to
'I saw them yesterday.' (Juma: Elicit)

b. **Oi'i=ve** ki jie gã=repiak**-i**.
some.time.ago=to PST 1.SG 3.PL=see-i
'Yesterday I saw them.' (Juma: Elicit)

(333) a. A-hepia ki jie Kanindé=gã **a'ero**.
1.SG.A-see PST 1.SG Kanindé=3.PL then
'I saw the Kanindé (group) then.' (Juma: Elicit)

b. **A'ero** ki jie Kanindé=gã repiak**-i**.
then PST 1.SG Kanindé=3.PL see-i
'Then I saw the Kanindé (group).' (Juma: Elicit)

(334) a. A-ko jie ji=rapyja pupe **amãna kyr=a=mẽ**.
1.SG.A 1.SG 1.SG=house inside rain rain=NMLZ=WHEN
'I stay inside my home when it rains.' (Juma: Elicit)

b. **Amãna kyr=a=mẽ**, jie ko**-i** ji=rapyja pupe.
rain rain=NMLZ=WHEN 1.SG be-i 1.SG=house inside
'When it rains, I stay inside my home.' (Juma: Elicit)

In natural speech, especially when narrating, speakers tend to repeat connectives like *a'ero* 'then' at the beginning of a new sentence as a strategy for text cohesion.[33] Notably, speakers commonly drop these clause-initial adjuncts after a few repetitions. However, the suffix *-i* is still required, despite the absence of an overt clause-initial adjunct. When presented with the version that includes the dropped constituent, speakers comment they are the same but that the version without the clause-initial adjunct is less repetitive, suggesting that the omitted element is recoverable from the context in connected speech.

This distribution of *-i* is also attested in sister languages of the Tupí-Guaraní family (Jensen 1999:156), where it is referred to as *oblique-topicalization/oblique-topicalized verb* (Jensen 1999; Vieira 2014).

[33] In this capacity, connectives like *a'ero* 'then' function analogously to words like *aí* 'then, so' in spoken Brazilian Portuguese, where they are used as cohesive devices.

These examples also show that matrix verbs lose subject agreement with subjects when adjuncts appear clause-initialy. Indeed, subject agreement in such cases is ungrammatical. Compare the (b) examples in (332)-(334) with the examples in (335), which exhibit agreement with the subject.

(335) a. *Oi'i=ve ki jie gã **a**-hepiak-i.
some.time.ago=to PST 1.SG 3.PL 1.SG.A-see-i
'Yesterday I saw them.' (Juma: Elicit)

b. *A'ero ki jie Kanindé=gã **a**-hepiak-i.
then PST 1.SG Kanindé=3.PL 1.SG.A-see-i
'Then I saw the Kanindé (group).' (Juma: Elicit)

c. *Amãna kyr=a=mẽ, jie **a**-ko-i ji=rapyja pupe.
rain rain=NMLZ=WHEN 1.SG 1.SG.A-be-i 1.SG=house inside
'When it rains, I stay inside my home.' (Juma: Elicit)

Chapter 4

Verb initiality

4.1 Introduction

10% of all languages exhibit a word order pattern where the verb obligatorily occupies the initial position (with respect to its arguments) in declarative sentences with neutral information structure (Dryer 2005). I refer to this pattern as the 'verb first pattern', as in (1).

(1) The verb first pattern (V1 pattern):
A language has a verb first pattern when the verb is obligatorily the first with respect to its arguments in a declarative sentence with neutral information structure.

In this chapter, I argue that verb-initial (VSO) word order in Kawahíva is a result of (long) syntactic head movement, akin to phrasal movement. This analysis is supported by the hallmarks of syntactic movement present in the head movement responsible for creating V1 order, including interpretive effects, nonlocality, and a specifier landing site. I also show that two alternative analyses to V1 order – Remnant VP Movement and Morphological Amalgamation – fail to show their predicted outcomes, i.e., the creation of a Remnant VP and a strictly local head movement.

Prominent examples of the verb first pattern can be found in Austronesian languages (Clemens and Polinsky 2017), Celtic languages (McCloskey 2006, 2017; Borsley and Kathol 2000; Jouitteau 2005), Mayan languages (Clemens and Coon 2018), and Berber languages (Ouhalla 1994).

In South America, we also encounter several unrelated language families and isolates that exhibit the V1 pattern. These include the Arawak languages Baure (Danielsen 2007), Caquinte (O'Hagan 2020), and Yanesha' (Duff-Tripp 1997); the Carib languages Ikpeng (Pachêco 2001) and Panare (Gildea 1993); the Chapacuran language Wari' (Everett and Kern 1997; Apontes 2015; Birchall To appear); the Macro-Jê language Kipéa do Karirí (Rodrigues 1999; Queiroz 2012); the Maku languages Nadëb and Kuyawi (Martins and Martins 1999); the Tupí-Guaraní language Tenetehará (Harrison 1986; Duarte 2012); the Tacanan language Reyesano (Guillaume 2012); the Peba-Yaguan language Yagua (Payne 1990); and the linguistic isolates Guató (Palácio 1984; Rodrigues 1999; Balykova 2019), Itonama (Mily Crevels, p.c for SAILS), Muniche (Michael et al. 2023), and Taushiro

(O'Hagan 2023).[1]

However, despite the fact that V1 has been a long-standing topic of discussion in theoretical linguistics (see Carnie and Guilfoyle (2000); Carnie et al. (2005)), the V1 pattern in South American languages has not received significant attention in this theoretical literature, with the exception of Duarte (2012)'s work on Tenetehára (Tupí-Guaraní).

Answers to the question of how V1 is derived in natural language presuppose detailed formal syntactic accounts of clause structures across a typologically diverse set of languages. This chapter contributes to the cross-linguistic and theoretical study of the V1 pattern, as described in (1), by examining the verb-initial (VSO) clauses of Kawahíva, based on the dialects spoken by the Juma and Jupaú (aka Uru-Eu-Wau-Wau (UEWW)). We see the Kawahíva V1 pattern below in (2).

(2) a. A-hepia ki jie anhãnga ko.
1.SG.A-see PST 1.SG ghost REAL
'I saw a ghost.' (Juma: Elicit)

b. Ere-'u-ipe po nde?
2.SG.A-eat-ALREADY IRR 2.SG
'Did you already eat?' (Juma: Elicit)

The V1 pattern also applies to sentences with a nonpronominal subject and object, as demonstrated by the elicited examples in (3a) and (3b), and the spontaneous speech example in (3c).

(3) a. V-epiang po txi ji=rembiriko=ga kaninde=gã nãhẽ.
3.A-see IRR FUT 1.SG=partner=3.SG.MASC Kanindé=PL POT
'My husband might see the Kanindé NGO group.' (Juma: Elicit)

b. O-i'inguhu ki javatxinga kohoa ko.
3.A-bark PST dog DEM:PROX:FLAT REAL
'That dog barked (at night).' (Juma: Elicit)

c. Nd-o-ko-potar-i tũhũi ki ji=rekira=hẽa Maytá=hẽa
NEG-3.A-live-DES-NEG INDEED PST 1.SG=older.sister=3.SG.FEM Maytá=3.SG.FEM
pevo oi'i nõmĩa.
LOC.DEM:DIST some.time.ago FRUSTR
'My sister Maytá did not want to live there anymore indeed.' (Juma: Text)

An important aspect of this pattern is a complementary distribution between clause-initial verbs and informationally prominent constituents in the clausal left periphery, as in the examples with a subject question and event report sentence in (4a-b), respectively.[2]

[1]Some of these languages and relevant sources were found through the *South American Indigenous Languages Structures* database (SAILS) (Krasnoukhova 2016).

[2]Following Lambrecht (1994), I assume that sentences like (4b) can be uttered out-of-the-blue and their only pragmatic requirement to be felicitous is to convey that something happened.

(4) a. Gara po txi nde ere-apo koi'iramẽ?
what IRR FUT 2.SG 2.SG.A-do later
'What will you do later?' (Juma: Elicit)

b. A-hepiang po txi jie UEWW=gã nãhẽ.
1.SG.A-see IRR FUT 1.SG UEWW=PL POT
'I might see the Uru-Eu-Wau-Wau (in the meeting I'm attending soon).' (Juma: Elicit)

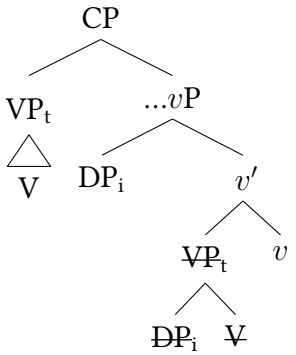
The subject question constituent in (4a) is initial and immediately followed by a mood and a tense particle, respectively. In (4b), the verb is initial and followed by the same particles. Examples (5a-b) demonstrate that only one of these elements may precede those particles at a time, regardless of their linear order before the particles.

(5) a. *V-epiang Kaninde=gã **po txi** ji=rembiriko=ga.
3.A-see Kanindé=PL IRR FUT 1.SG=partner=3.SG.MASC
'It is Kanindé that may see my husband.' (Juma: Elicit)

b. *Kaninde=gã v-epiang **po txi** ji=rembiriko=ga.
Kanindé=PL 3.A-see IRR FUT 1.SG=partner=3.SG.MASC
'It is Kanindé that may see my husband.' (Juma: Elicit)

A similar distributional pattern is also found in Niuean, where predicative phrases or verbs can occupy the initial position in the clause (Massam 2001). Massam (2001) develops a Remnant VP analysis to account for this pattern; she proposes that the elements involved are both phrasal, and the verb is embedded within a VP whose object is evacuated prior to the fronting of the VP. Consequently, the initial VP surfaces with only the verb head, although it is still a phrasal constituent, as represented in Tree (6). Similar facts are also reported for Tenetehára (Duarte 2012), which I will review later in this chapter.

(6)

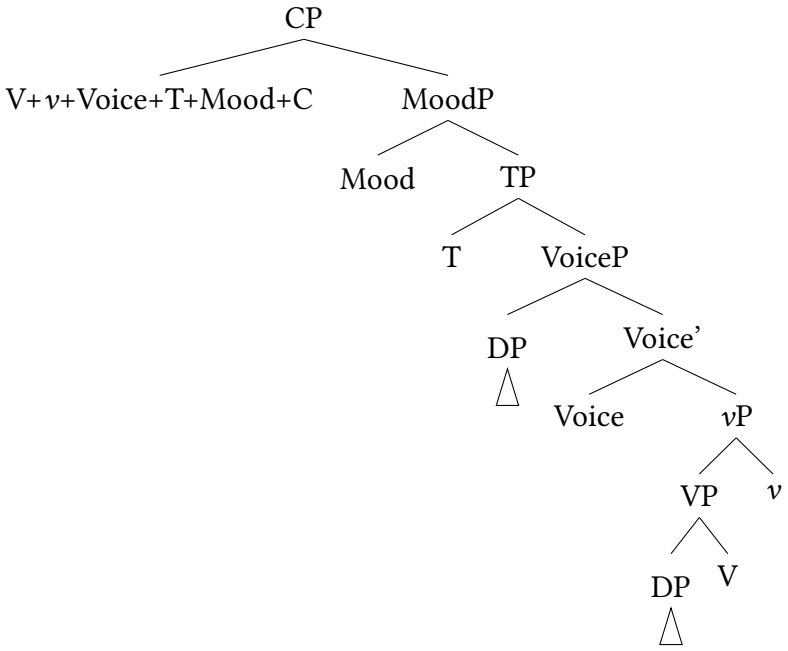


Notably, this account necessitates that object evacuation is independently attested in the language. In section 4.4, I argue this approach is not suitable for Kawahíva V1 on the basis of positive evidence showing that the object does not evacuate the VP in V1 clauses.

I explore a non-syntactic derivation of the Kawahíva V1 pattern through Morphological amalgamation (Harizanov and Gribanova 2019); however, I ultimately reject this alternative. A clear

proposal for a V1 pattern using this approach is Hammerly (2021) for the Algonquian language Ojibwe. In a morphological amalgamation account to V1, the verb moves through all intermediate between its base position and landing site. This account is schematized in (7):

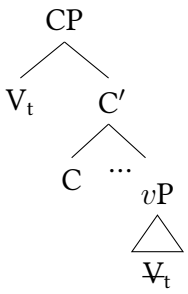
(7)



In section (4.5), I present several arguments that challenge this analysis for Kawahíva V1, all having to do with the absence of the hallmarks of morphological amalgamation, including locality (Harizanov and Gribanova 2019) – while morphological amalgamation abides by the Head Movement Constraint (Travis 1984) (i.e., a head may only move to the next head up), verb fronting in Kawahíva skips at least TP, thus strongly suggesting that it is not strictly local.

In section (4.6), I will argue that the correct analysis of Kawahíva V1 is syntactic head movement, illustrated in (8); intermediate projections are omitted.[3]

(8)



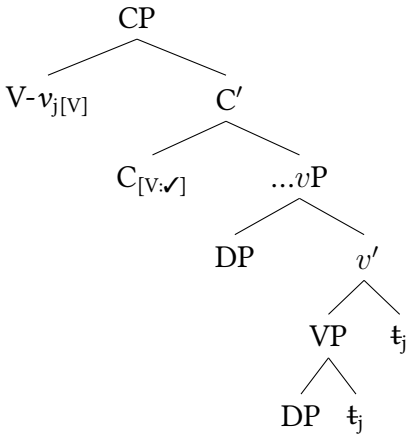
[3]In §4.7, we will see that Kawahíva V1 also involves a few steps of local movement in the lower domain of the clause, namely *v*P and VoiceP.

This idea has not been popular as a way of deriving V1 order as a general principle of clausal organization. To my knowledge, only Mayan (Clemens and Coon 2018), Otomanguean (Lee 2005; Macaulay 2005; Eberhardt 1999) and, under one proposal, the Polynesian language Niuean (Clemens 2019) have explicitly received this account. However, given that the theory predicts syntactic head movement is available, we would expect to find more languages where it is used in the creation of the V1 order. In this chapter, I argue that Kawahíva is one such language where this approach applies effectively to account for its V1 pattern.

Supporting evidence for this idea comes from properties usually associated with syntactic movement, including the presence of interpretive effects, nonlocality, and a specifier position as the landing site (Vicente 2009; Harizanov 2019). In addition, this approach effectively accounts for the complementary distribution between the verb and phrases of the CP domain (e.g., foci, topics, frame-setting adverbial PPs) as a competition for this specifier position.

More technically, I propose the Kawahíva syntactic movement is driven by the feature [V] with the EPP property on CP, following similar proposals for the feature trigger in other verb-initial languages (Massam 2020; Van Urk 2022). Additionally, I posit [V] is also present on *v*P. A consequence of the latter is the head V will acquire the feature [V] after head-moving onto *v*. The complex head V-*v* bearing [V] will move to Spec,CP to satisfy the EPP of [V] on C, as schematized in (9).

(9)



In the final section in (4.7), I extend this analysis to non-verb-initial clauses, which involve fronting of a PP or DP for extraction. I argue that when PPs, but not DPs, front due to A'-extraction, they check the [V]-feature trigger on C. In cases of PP fronting, the verb is not able to undergo fronting. Support for this idea comes from morphosyntactic changes observed on the verb when PPs are initial, but not when DPs are initial. In contrast, in DP fronting scenarios, the DP does not remove the feature trigger for verb movement, thus still allowing the verb to front to the same specifier position the DP lands in (i.e., Spec,CP). I suggest a restriction on spelling out the verb in this position applies such that the verb is spelled out lower than the DP, but higher than its base position.

This chapter is organized as follows. §4.2 reviews the key aspects of Kawahíva for this chapter, including verbal agreement, matrix clauses, and dependent clauses. In §5.3, I propose a phrase structure for matrix 4.3.1 and dependent clauses 4.3.2 to account for their different properties, as well as the complementary distribution between verbs and phrases in the initial position. Then I offer a brief account of verbal agreement that captures straightforwardly the absence of subject agreement in dependent clauses, in §4.3.3. After that, in §4.4, I consider and discard alternatives to derive V1 clauses, namely Remnant VP movement and morphological amalgamation, in §4.5. In §4.6, I present the proposed analysis of Kawahíva V1, the Syntactic Head Movement/Head-to-Spec triggered by [V] with the EPP property. Finally, in §4.7, I discuss how my analysis extends to non-verb-initial clauses, consider a feature account of the complementary distribution, and address the issue of XP vs. X movement.

4.2 Key aspects of Kawahíva

4.2.1 Verbal agreement

Recall from §3.4.1 that Kawahíva exhibits agreement between a verb and the [ϕ]-features (person, number, and Case) of one of its nominal arguments. I provide the complete set of person indexes in Table 4.1. Subject agreement is expressed via inflectional prefixes (i.e., SET A), whereas object agreement is realized with bound pronouns (i.e., SET B) (Dos Santos 2021a, 2023a). Note Kawahíva does not exhibit agreement with 3rd person objects.

	SET A	SET B
1ST PERSON SG.	a-	ji=
2ND PERSON SG.	ere-	nde=
1ST PERSON INCL.	txi-	nhãnde=
1ST PERSON EXCL.	oro-	ore=
2ND PERSON PL.	pe-	pe=
3RD PERSON	v- ∼ o-	–

Table 4.1: Kawahíva indexes.

In matrix clauses, transitive verbs index either the subject or object; the choice depends on the person hierarchy 1>2>3 (Dos Santos 2021a, 2023a). Examples in (10) demonstrate that, in 1->2 and 2->1 contexts, the 1st person pronoun is always marked on the verb regardless of the grammatical role.

(10) a. A-hepia ki ji nde ko.
1.SG.A-see PST 1.SG 2.SG REAL
'I saw you.'

b. Ji=repia ki nde ko.
1.SG.B=see PST 2.SG REAL
'You saw me.'

Intransitive verbs fall into three subgroups. Two of them are marked with one of the sets presented in Table 4.1, while a third subgroup is marked with the invariant prefix *i-*. I summarize the morphologically-distinguished intransitive verbs in Table 4.2.

	SET A	SET B (*n=4*)	PREFIX *i-*
1ST PERSON SG.	a-	ji=	i-
2ND PERSON SG.	ere-	nde=	i-
1ST PERSON INCL.	txi-	nhãnde=	i-
1ST PERSON EXCL.	oro-	ore=	i-
2ND PERSON PL.	pe-	pe=	i-
3RD PERSON	v- ∼ o-	–	i-

Table 4.2: Morphological-distinct subgroups of intransitive verbs.

The subgroup of intransitive verbs that exhibit set A marking includes verbs like *kwam* 'dance, jump', among others; an example is (11). I refer to this subgroup as unergative verbs.

(11) A-kwam ki ji.
1.SG.A-dance PST 1.SG
'I danced.'

The subgroup of intransitive verbs that take set B marking is small. It includes the verbs *tur ~ rur* 'come, arrive', *ten ~ ren* 'be seated', *tup ~ 'up* 'be in a horizontal position, be laid down', and *tuv ~ ruv* 'stay'. (12) provides an example of this subgroup. I refer to these verbs as unaccusative verbs.

(12) A'ero txiro, ji=rur-i carro-'i=a pype.
then CONSTINUOUSLY 1.EXCL.B=come-i car-DIM=NMLZ inside
'Then I was coming in the small car.'

Additionally, the subgroup of intransitive verbs that takes the invariant prefix *i-* includes verbs like *katu* 'be pretty, well,' *kwerai* 'be tired,' *rovia* 'be happy,' *rãite* 'be ugly,' and *ro'y* 'be cold', among many others. (13) serves as an example of this morphological subgroup.

(13) I-kwerai ji.
i–be.tired 1.SG
'I am tired.'

I treat the sole argument of *i*-marked intransitive verbs on a par with the SET B-marked ones syntactically – they are both objects. In other words, they are unaccusative verbs (Perlmutter 1978). Evidence for this syntactic object status in the former subgroup comes from Dos Santos (2022), where I discussed the language-internal diagnostic for unaccusativity in Kawahíva.[4] This

[4]Standard cross-linguistic tests for unaccusativity, such as passives, auxiliary selection, resultative predicates, and participle formation, are not applicable in Kawahíva (Dos Santos 2022).

argument draws on the distribution of the prefix *i-* with transitive verbs: *i-* is the morphological reflex of extraction of objects of transitive verbs in clause-final position, such as dependent clauses; this pattern is demonstrated in (14). In it, extraction of the object within a complement clause to the matrix clause triggers the prefix *i-* on the dependent verb. Examples in (15) show that subject extraction does not trigger *i-*.

(14) a. Ere-piang nde [apinaga pira 'u=a] rai'i.
2.SG.A-see 2.SG father fish eat=NMLZ YESTERDAY.PST
'You saw father eating fish yesterday.'

b. Gara nde ere-piang [apinaga __ i-'u=a] rai'i?
what 2.SG 2.SG.A-see father i-eat=NMLZ YESTERDAY.PST
'What did you see Father eating yesterday?'

c. *Gara nde ere-piang [apinaga __ 'u=a] rai'i?
what 2.SG 2.SG.A-see father eat=NMLZ YESTERDAY.PST
'What did you see Father eating yesterday?'

(15) a. Mãngãn nde ere-piang [__ pira 'u=a] rai'i?
who 2.SG 2.SG.A-see fish eat=NMLZ YESTERDAY.PST
'Who did you see eating fish yesterday?'

b. *Mãngãn nde ere-piang [__ pira i-'u=a] rai'i?
who 2.SG 2.SG.A-see fish eat=NMLZ YESTERDAY.PST
'Who did you see eating fish yesterday?'

As the sole argument of unaccusative verbs is an object, which becomes derived subjects by moving to the subject position, the sole argument of *i*-marked intransitive verbs must have moved in Kawahíva as well. I propose this object movement, like object extraction in (14b), triggers *i-* in Kawahíva. In brief, while unaccusative verbs in Kawahíva exhibit a different morphology – SET B (object agreement morphology) and *i*-prefixation – this morphology still shows their sole argument is an object (i.e., they have the syntax of unaccusatives).

4.2.2 Matrix clauses

Regular declarative matrix clauses exhibit a VSO word order in event reports, out-of-the-blue clauses, an example of which is (16a), and songs in Kawahíva, as in (16b).[5] We also see the VSO order in polar questions, as in (17).[6]

[5]The invariable VSO order in songs is particularly telling as songs are a good source to diagnose pragmatic unmarked word order in languages where word order is pragmatically determined (Turpin 2013).

[6]There is one reference to Kawahíva clause structure in previous work, including verb-initial clauses. Pease (1968), in her grammar sketch, writes: *When the object is a noun the preferred order of occurrence is S P ±O [SVO] in the declarative intransitive and demonstrative transitive clauses, and also in the declarative transitive type except when the subject is a pronoun in which case the preferred order is P S O [VSO].*

(16) a. A-hepiag-ipe ki ji gã ko.
1.SG.A-see-ALREADY PST 1.SG 3.PL REAL
'I already saw them the other day' (Juma: Text)

b. Ere-imbory-vepia-piang nde akaritajavuhua.
2.SG.A-yellow-RED-see 2.SG headdress
'You saw several yellow headdresses' (Juma: Text from a song)

(17) O-'u po nde=ra'yra=ga pira?
3.A-eat IRR 2.SG=offspring=3.SG.MASC fish
'Does your son eat fish?' (Juma: Elicit)

Matrix clauses in Kawahíva are further characterized by the presence of the mood particles *po* 'irrealis' and *ko* 'realis', as well as the tense particles *ki* 'past' and *txi* 'future'. In natural speech, these particles are sometimes dropped, with their meanings inferred from the context. A brief characterization of the surface position of these particles is relevant for determining the landing site of the verb.

The tense particles *ki* 'past' and *txi* 'future' are mutually exclusive, and their surface position depends on the presence of other elements within the clause. These particles occur in the initial position in the clause if no constituent or verb is fronted, and if *po* 'irrealis' is absent, as examples (18) and (19) demonstrate for each tense particle.

(18) **Ki** ore ho-i, karuka=mẽ ko.
PST 1.EXCL go-i afternoon=to REAL
'We went in the afternoon.' (Juma: Text)

(19) **Txi** te gã o-juka miara, txi te gã nd-o-juka-i.
FUT REALLY 3.PL 3.A-kill meat FUT REALLY 3.PL NEG-3.A-kill-NEG
'(I do not know) whether they will kill something or not (lit.: (I do not know) whether they will kill game meat, or they will not kill (it)).' (Juma: Text)

Tense particles, on the other hand, appear in the second position within the clause if a constituent is fronted, as seen in example (20) where the verb occupies the clause-initial position. The same position is observed when there is no fronting, but *po* 'irrealis' is present, as demonstrated in example (21).

(20) A-hendu-katu-ramõ **txi** jie nõmĩa, 'i ki ji ko=ra.
1.SG.A-listen-WELL-LIKE FUT 1.SG FRUSTR say PST 1.SG REAL=TODAY.PST
'I got it, I said.' (Lit.:'I understood it well, I said') (Juma: Text)

(21) Po **txi** (nãhẽ)!
IRR FUT POT
'Later!' (Common in everyday speech when one is asked to do something and they respond that they will perhaps do it later. This is most of the time uttered as a joke.)

Finally, these particles show up in the third position of the clause when there is both fronting of a constituent or the verb, and *po* 'irrealis' is overt, as shown in (22) with constituent fronting.

(22) Gara po **txi** nde ere-apo koi'iramẽ?
what IRR FUT 2.SG 2.SG.A-do later
'What will you do later?' (Juma: Elicit)

This is a similar distribution found with *ki* 'past', with the caveat that this particle does not occur in the third position; this would require co-occurrence with *po* 'irrealis', which is semantically incompatible.

The particles *po* and *ko* are mood particles with the meanings 'irrealis' and 'realis', respectively (Dos Santos 2021b). While these particles exhibit a paradigmatic relation, they have different surface positions. On the one hand, *po* 'irrealis' appears either in the first position when no constituent or the verb is fronted or in the second position when a constituent or the verb is fronted. On the other hand, *ko* 'realis' appears in sentence-final position.[7] In section (5.3), I revisit this surface distinction to propose that it ultimately boils down to the headedness parameter of the phrase where these particles reside. This phrase is head-final in realis-marked sentences but head-initial in irrealis-marked sentences. Due to the surface differences between the mood particles and the present focus on the left periphery of the clause, I will primarily consider *po* 'irrealis' in the discussion and examples.

As previously mentioned, determining the surface position of these particles is crucial for identifying the initial verb's position. In all the examples presented so far, the initial verbs precede the mood particle *po* 'irrealis' and the tense particles *ki* 'past' and *txi* 'future'. In fact, in pragmatically-neutral contexts, the verb must surface before each of these particles. We see this requirement in past tense-marked clauses with *ki* 'past' in (23). Any deviation from this word order results in unacceptable sentences.

(23) a. A-hepiang **ki** jie gã ko
1.SG.A-see PST 1.SG 3.PL REAL

b. *__Ki__ a-hepiang jie gã ko
PST 1.SG.A-see 1.SG 3.PL REAL

c. *__Ki__ jie a-hepiang gã ko
PST 1.SG 1.SG.A-see 3.PL REAL

d. *__Ki__ jie gã a-hepiang ko
PST 1.SG 3.PL 1.SG.A-see REAL

e. *__Ki__ jie gã ko a-hepiang
PST 1.SG 3.PL REAL 1.SG.A-see

'I saw them' (Juma: Elicit)

[7] The example in (20) initially appears to be a counterexample for this generalization. However, see footnote (10), where I revisit this data point in the context of the proposed clause structure for Kawahíva.

The same generalization holds for irrealis-marked sentences with *po* 'irrealis' — the verb must precede this particle, or the sentence becomes unacceptable:[8]

(24) a. *__Po__ o-'u nde=ra'yra=ga pira?
IRR 3.A-eat 2.SG=offspring=3.SG.MASC fish

b. *__Po__ nde=ra'yra=ga o-'u pira?
IRR 2.SG=offspring=3.SG.MASC 3.A-eat fish

c. *__Po__ nde=ra'yra=ga pira o-'u?
IRR 2.SG=offspring=3.SG.MASC fish 3.A-eat

'Does your son eat fish?' (Juma: Elicit)

A further property of initial verbs is agreement marking. In the clause-initial position, verbs are required to exhibit agreement marking, a subject person index (SET A) or an object person index (SET B). The choice is governed by the person hierarchy 1>2>3, as discussed in §4.2.1. This requirement is illustrated for subject agreement in (25).

(25) a. *Hepiang __ki__ jie gã ko.
see PST 1.SG 3.PL REAL

'I saw them.' (Juma: Elicit)

b. A-hepiang __ki__ jie gã ko.
1.SG.A-see PST 1.SG 3.PL REAL

'I saw them.' (Juma: Elicit)

However, if the verb is not in the initial position, it may or may not bear a subject index (SET A marker). This difference is correlated with whether the clause-initial constituent is an adverbial or

[8]It is worth noting that this requirement for the verb to precede the mood and tense particles is not the result of a language-wide prosodic constraint against these particles in clause-initial position. Examples like those in (1), which can be found in the corpora of Kawahíva spontaneous speech, demonstrate that each particle may occur clause-initially. Additionally, we already saw in (21) that the mood and tense particles can co-occur, where the mood particle is in the clause-initial position.

(1) a. __Ki__ ore ho-i, karuka=mẽ ko.
PST 1.EXCL go-i afternoon=to REAL

'We went in the afternoon.' (Juma: Text)

b. __Po__ ahe ho-i a'ero, 'i ki gã ko.
IRR people go-i then say PST 3.PL REAL

'Let's go then, they said.' (Juma: Text)

One could imagine the alternative that these particles exhibit prosodically weak variants that are used in clause-initial position. In this view, then, it would be the weak variants that are banned from the initial position. However, considering the complete analysis that will be developed, the reason these particles are prohibited from appearing at the beginning of a clause in examples (23)-(24) is due to the categorical requirement that the verb must move across them to reach the clause-initial position in pragmatically-neutral declarative matrix clauses.

an extracted DP, as I will discuss next, along with other accompanying morphosyntactic changes in non-verb-initial clauses. Firstly, I characterize the instances of non-verb-initial clauses.

Non-verb-initial clauses are found in two instances. In one, the initial position is filled in with an informationally prominent DP, such as foci (26)-(27), *wh*-words (28)-(29), and topics (30). Crucially, a SET A argument index on the verb can appear in all these cases; the choice depends on the person hierarchy.

(26) Mbahira ki v-apo kunhã.
Mbahira PST 3.A-make woman
'It was Mbahira who made the woman.' (Jupaú: Text)

(27) Pirapetxinguhua ki hẽa o-pyhy-pyhy ko.
pirapetxinguhua.fish PST 3.SG.FEM 3.A-RED-catch REAL
'It was *pirapetxinguhua* fish that she caught.' (Juma: Text)

(28) Mãngã ki v-epiang kwemba'ea?
who PST 3.A-see man
'Who saw the man?' (Juma: Elicit)

(29) Mãngã po ji=ruvyra v-epiang rimba'e?
who IRR 1.SG=uncle 3.A-see NWIT.REM.PST
'Who did my uncle see back in the day?' (Juma: Text)

(30) Takãi Kajuvi ki, tapy'ynha o-purun ga=rehe.
Takãi Kajuvi PST non-indigene 3.A-hit 3.SG.MASC=at
'(As for) Takãi Kajuvi, the non-indigene ran over him.' (Juma: Text)

A second instance of non-verb-initial clauses include cases where the clause-initial position is filled in with a postpositional phrase (PP) that serves as an adverbial frame-setting element, including adverbial *wh*-words like *maramẽ* 'when', as in (31), and a discourse particle, which is generally a PP. In contrast to DP-initial clauses, in PP-initial clauses, the verb cannot bear a SET A argument index, which (31) also serves to illustrate.

(31) Mãrã=mẽ po ga pira (*o-)'u-i?
where=to IRR 3.SG fish (3.A-)eat-i
'Where does he eat fish?' (Juma: Elicit)

In §4.7, I suggest the complementarity distribution between subject indexation and clause-initial PPs follows straightforwardly in my account of the V1 order in Kawahíva. I refer the reader to §4.7 for further details, as previewing them in this subsection would take us too far afield.

Informationally prominent constituents also interfere with the ability of a matrix verb to surface before the mood and tense particles when in clause-initial position, in that the verb can no longer appear before them, as the example in (32) shows.

(32) *Mbahira v-apo ki kunhã.
Mbahira 3.A-make PST woman
'It was Mbahira who made the woman.' (Jupaú: Elicit)

(32) is similar to (26). The difference is the verb precedes the past tense particle in the former but not in the latter example. This change in word order causes unacceptability. This sentence is unacceptable regardless of whether the verb precedes or follows the focused constituent, as shown in (33).

(33) *V-apo Mbahira ki kunhã.
3.A-make Mbahira PST woman
'It was Mbahira who made the woman.' (Jupaú: Elicit)

This data strongly suggests that the reason that the verb and informationally prominent constituents cannot precede the mood and/or tense particles at the same time is structural. They compete for the same clause-initial position.

Two final properties of non-verb-initial clauses are the surface position of the verb and specifically in the case of PP-initial clauses, the appearance of the suffix *-i*. Regarding verb position, in DP-initial clauses, as in (26), the verb occupies a medial position. This is evidenced in subject extraction since, in object-prominent clauses, the object is fronted. In PP-initial clauses, however, the verb is always final, as in (31). This and other differences discussed above are summarized in Table 4.3:

CLAUSE-INITIAL XP	verb position	agreement	*-i* suffixation
DPs	medial	sets A & B	✗
PPs	final	set B	✓

Table 4.3: Summary of clause-initial elements and morphosyntactic properties.

Note how these different orderings between the object and verb correlate with the availability of subject marking (i.e., a SET A marker). In particular, PP-initial clauses, but not DP-initial subjects, seem to interfere with the availability of subject marking on the non-initial verbs. I will return to this distribution in §4.5, where I demonstrate how my analysis accounts for these cases.

In brief, I have shown that the following properties characterize main clauses: the availability of two sets of agreement markers, the availability of verb fronting, the presence of reality marking particles, and the availability of a left peripheral focus and topic position, which competes with verb fronting. In the following section, we will see that dependent clauses lack all of these properties. This scenario, I argue, strongly suggests the same initial position in matrix clauses is missing in dependent clauses. I will propose in §5.3 that this is because the dependent clauses lack a CP.

4.2.3 Dependent clauses

In contrast to matrix clauses, dependent clauses exhibit a truncated structure. They lack an internal landing site for extraction, mood particles, verb fronting, and subject agreement. The latter will be discussed only after presenting a proposal for the clause structure in §4.3.3.

4.2.3.1 Absence of landing site for extraction within embedded clauses

The most telling piece of evidence that dependent clauses are truncated comes from embedded questions. In elicitation, the strategy of questioning a phrase within the embedded clause is always rejected, as the examples below demonstrate.

(34) a. *A-kwaham jie [mõrõmõ=pe ji=mitakwanha i-'u=a].
1.SG.A-know 1.SG which=2.PL bread 1.SG-eat=NMLZ
Intended: 'I know which one of you ate my bread.' (Juma: Elicit)

b. *Nd-a-kwahav-i jie [mãngã=pe Wesley=ga u'ia reroho=a].
NEG-1.SG.A-know-NEG 1.SG who=to W.=3.SG.MASC manioc.flour take=NMLZ
Intended: 'I don't know who Wesley took manioc flour to.' (Juma: Elicit)

One of the different strategies to convey an embedded question involves two independent clauses, as illustrated in the following sentence.

(35) Mãngã=pe Wesley=ga u'ia eroho-i? Nd-a-kwahav-i jie.
who=to W.=3.SG.MASC manioc.flour take-i NEG-1.SG.A-know-NEG 1.SG
'Who did Wesley take manioc flour to? I don't know.'

In one intricate instance of the two-sentence strategy to convey the embedded question meaning, the speaker uses a complex clause wherein the first part shows the matrix verb *kwahav* 'know' embedding a clause, but the embedded clause lacks the *wh*-word, as in (36a). The second part of the two-sentence structure is a simple interrogative clause. This example was given as the repair to example (34b) from above. The other strategy is pervasive in natural speech. In it, what is consistently translated as embedded questions is formally not an embedded question. Rather, the structure used is simply a matrix question with the irrealis particle *po*, as the examples (36b-c) illustrate.

(36) a. Nd-a-kwahav-i jie [Wesley=ga u'ia reroho=a]. Mãngã=pe po
NEG-1.SG.A-know-NEG 1.SG W.=3.SG.MASC manioc.flour take=NMLZ who=to IRR
ga eroho-i, u'ia?
3.SG.MASC take-i manioc.flour
'I don't know (who) Wesley took manioc flour to. To whom did he take it, the manioc flour?' (Backtranslation of "I don't know who Wesley took manioc flour to?")

b. Mãngã po ji=ruvyra v-epiang rimba'e?
who IRR 2.SG=late.uncle 3.A-see NWIT.REM.PST
'Who did my late uncle see long ago?' (Juma: Text)

c. Mãrãmẽ po ore=ho-i?
when IRR 1.SG.B=go-i
'When will we travel?' (Juma: Text)

Taken together, these facts suggest that Kawahíva lacks embedded questions, and their meanings are conveyed with other structures. I propose that this follows from embedded clauses lacking the relevant position that hosts questioned phrases (=embedded clauses are truncated). The absence of embedded questions would be surprising for Kawahíva if we viewed both matrix and dependent clauses with a uniform clausal structure. In the next subsections, I demonstrate that mood particles, verb fronting, and subject agreement are also unavailable in embedded clauses. After assembling these facts, it will become evident that the position responsible for the different morphosyntactic properties between matrix and dependent clauses is the clause-initial position. In §5.3, I will identify this position as C.

4.2.3.2 Absence of mood particles

A second piece of evidence indicating that the clause-initial position of matrix clauses is missing in dependent clauses is their lack of reality marking. For instance, the use of reality mood is unacceptable in complement clauses (37a) and relative clauses (37b).

(37) a. *A-rovia jie [Puré=ga po tapi'ira juka=a].
1.SG.A-believe 1.SG IRR P.=3.SG.MASC tapir kill=NMLZ
'I believe Puré perhaps killed a tapir.'

b. *E-mõtyryry irãmutxinguhua [Wesley=ga remb-i-mõndo=a ko]!
2.SG.IMP-fry chicken W.=3.SG.MASC WH.OBJ.I-i-send=NMLZ REAL
'Fry the chicken which Wesley sent!'

The absence of reality marking in a dependent clause might not be strong evidence in itself that dependent clauses are truncated.[9] However, conjoined with the result that embedded clauses also lack a position for extraction, it becomes natural to link the absence of the two to a single explanation, which is the absence of a specific structural domain in their structure.

4.2.3.3 Absence of verb fronting

The absence of embedded questions and mood particles in embedded clauses leads to another expectation: that verbs should not be allowed to surface at the edge of the embedded clause. This prediction arises from the fact that in matrix clauses, verbs and pragmatically-prominent phrases compete for the clausal-initial position, and both immediately appear before *po* 'irrealis'. If neither of these properties is observed in dependent clauses because these clauses are truncated, then it follows that verb fronting must also be lacking in embedded clauses. This expectation is confirmed as follows.

[9]One could think this is due to dependent clauses not being truth functions, as they are not assertions.

Dependent clauses are strictly verb-final (i.e., SOV order). The following examples provide positive evidence of this verb-final order in all types of dependent clauses present in the language, including complement clauses (38a), relative clauses (38b), and adverbial clauses (38c).

(38) a. A-hepiang ki jie [ji=ruva=ga pira ʻu=a].
1.SG.A-see PST 1.SG 1.SG=father=3.SG.MASC fish eat=NMLZ
'I saw [my father eating fish].' (Juma: Elicit)

b. [Tapy'ynha mõhanga mbuhu-har=a]=gã te'i o-hun ore=pyri.
non.indigenous medicine bring-WH.TRANS.SUBJ=NMLZ=PL only 3.A-come 1.EXCL=by
'The non-indigenous (people) [who bring medicine] are the only ones that come to us.' (Juma: Text)

c. [Ji kandambuhua ʻu-ramẽ], ji=reveka nda–katu–i.
1.SG papaya eat–WHEN 1.SG=stomach NEG–be.well–i
'[When I eat papaya], my stomach gets bad.' (Juma: Elicit)

Indeed, any other position for the verb within dependent clauses is unacceptable, as shown in (39), which includes an argument complement clause. This pattern extends to relative clauses and adjunct clauses likewise (not shown below).

(39) a. *A-hepiang ki jie [ʻu=a ji=ruva=ga pira].
1.SG.A-see PST 1.SG eat=NMLZ 1.SG=father=3.SG.MASC fish
'I saw [my father eating fish].' (Juma: Elicit)

b. *A-hepiang ki jie [ji=ruva=ga ʻu=a pira].
1.SG.A-see PST 1.SG 1.SG=father=3.SG.MASC eat=NMLZ fish
'I saw [my father eating fish].' (Juma: Elicit)

4.3 Clause structure and verbal agreement

I argue the differences between matrix and dependent clauses receive a natural account under the proposal that the clause-initial position is the CP domain and this domain is lacking in dependent clauses.

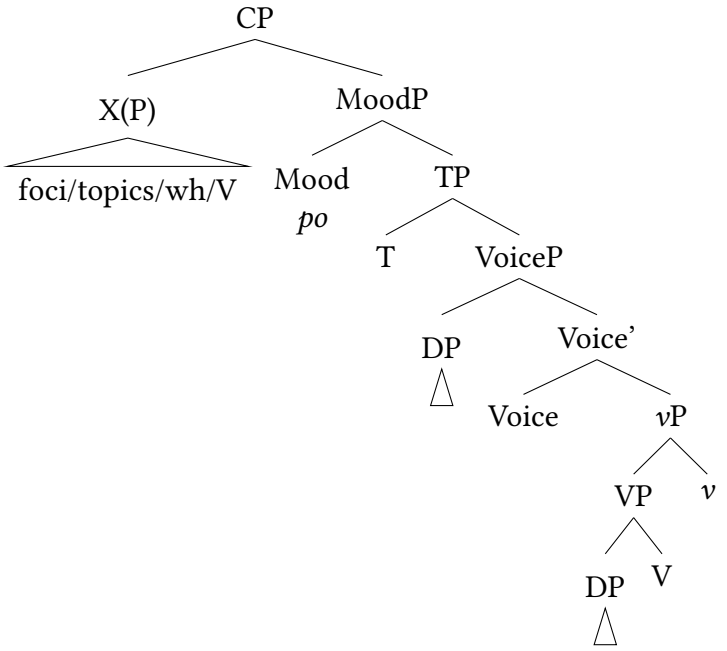
4.3.1 A proposal for the clause structure

I propose the matrix clause structure is as shown in (40).

(40) CP – MoodP – TP – VoiceP – *v*P – VP

The CP, MoodP, and TP domains are supported by the properties of matrix clauses discussed in the previous section. The other projections, VoiceP, *v*P and VP, are unexceptional on the layered approach to clausal structure arrived at in various schools of syntactic analysis (Kratzer 1996; Pylkkänen 2002; Harley 2013). The trees below illustrate that discourse-prominent constituents and the verb occupy the CP domain when clause-initially.

(41)



In this view, the surface difference between the mood particles *po* 'irrealis' and *ko* 'realis' boils down to the headedness parameter within MoodP, as I propose in this chapter. In particular, *po* 'irrealis' is left-headed (41), but *ko* 'realis' is right-headed, as in (42):[10]

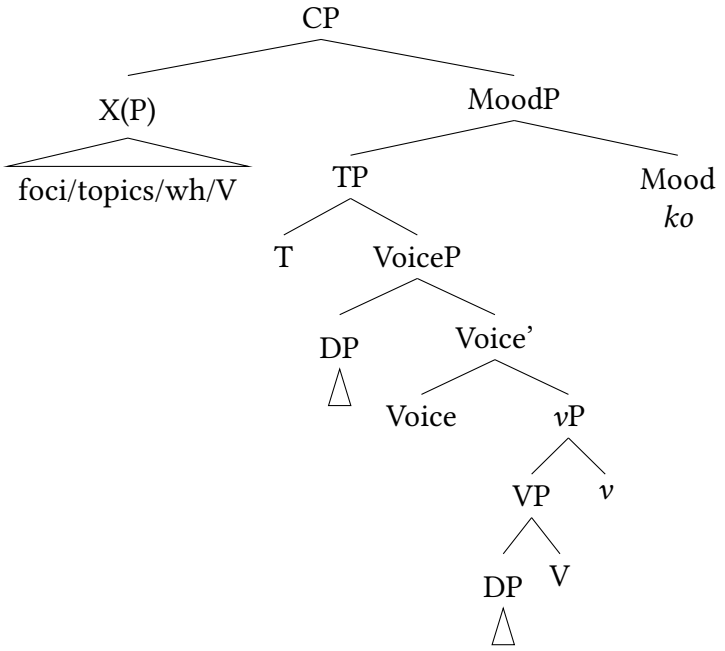
[10]The non-clause-final position of *ko* 'realis' in example (20) might, at first sight, contradict the clause structure proposed for the matrix clause in section (5.3). The realis particle realizes the highest right-headed functional projection of the clause spine, yet it is not the final element in the clause.

However, this particular order between the two particles arises only in cases where *ra'e* shows up in its reduced form and encliticizes to the right of *ko*. In the natural speech corpora, the instances wherein the full form of *ra'e* is used and it co-occurs with *ko* shows that the order is always *ra'e* followed by *ko*, as in the following example:

(1) Mã! i-pojita j-uvi ra'e ko.
Huh! i-be.afraid i-from TODAY.PST REAL
'Huh! One is afraid of it.' (Juma: Text)

Given that the two orders seem to be conditioned by the full phonological form of one of the particles, one way to understand these ordering differences is as the result of some postsyntactic rearrangement between the two, possibly due to a requirement imposed by the reduced form of *ra'e* that it needs a host to its left. Alternatively, the two particles might not be compositionally transparent at this point, in which case there is only one morpheme synchronically. Consequently, the issue of morpheme order becomes irrelevant. Notably, the glossing would have to change to reflect this descriptive analysis. Either way, the proposed clause structure in section (5.3) can still be maintained despite the surface order of the two particles.

(42)



A mixed-headedness parameter finds further support in a distantly related language in the Tupían stock. It has been independently argued for to explain the surface differences in the realization of members of the paradigm that belongs to T in Tuparí (Tuparí branch of Tupían). In particular, Singerman (2018, 2021) argues that the Tuparí tense domain (=TP) has head-initial and head-final properties.

Furthermore, the complementary distribution between discourse prominent phrases and the verb is a concomitant of both competing for the same landing site, Spec,CP.[11]

Finally, I propose that the VP is underlyingly head-final (i.e., OV). Verbs in all types of dependent clauses exhibit a strict SOV word order; this contrasts with the fact that SOV word order in matrix clauses is only possible in cases where the trigger for verb fronting is removed (i.e., sentences with an initial adverbial PP).[12] Additionally, this head-final property correlates with other properties generally found in languages that have been described as head-final. Particularly, it correlates with the strict head-final order of other phrases in the language, including the *noun-postposition* order, *possessor-possessee*, *noun-demonstrative*, and the *verb-auxiliary* order.

[11] In §4.7.2, I discuss an alternative account for this distribution wherein the competition between informationally prominent phrases and verbs for the clause-initial position is due to the feature trigger for fronting.

[12] This is analogous to the reasoning that German is underlyingly OV. We see OV in embedded clauses where V can't move but V2 in matrix clauses where V can move. However, the reason V can't move in embedded clauses is different in both languages: C is occupied (German) vs. C is absent (Kawahiva) (Vikner 1995).

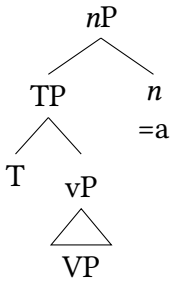
4.3.2 The structural difference between matrix and dependent clauses

The absence of embedded questions becomes quite straightforward given the matrix clause structure. We saw in section §4.2.3 that dependent clauses lack some of the morphological and structural properties of matrix clauses, including the absence of subject agreement, a host position for extraction and verb fronting, and the lack of mood particles.[13] Translated into the current proposal wherein these properties are associated with the CP and MoodP domains, this fact means that dependent clauses lack both. Consequently, I propose the dependent clause structure as represented in (43).

(43) TP – VoiceP – *v*P – VP

I also posit that dependent clauses are embedded under a syntactic nominalizing head to account for their nominal properties.[14] In Kawahíva, this head is spelled out as *=a* (see Dos Santos (2023b) for a nominalizer account of *=a*). Diagram (44) represents the proposed structure of dependent clauses, where their highest verbal functional projection, TP, is the complement of the nominalizing *=a*.[15]

(44)



To summarize, in Kawahíva, verbs and prominent constituents compete for the same matrix clause-initial position, as evidenced by their complementary distribution. I have proposed that both the verb and prominent constituents land in the specifier of CP. The position of Spec,CP as the landing site for verbs and prominent constituents is supported by their consistent placement before the mood and tense particles, which are the exponents of the heads of MoodP and TP respectively, when fronted. Therefore, given the conventional layered approach to clause structure (CP-TP-VoiceP-vP-VP), it is natural to propose that this position is CP.

[13]Other differences between matrix and dependent clauses include tense and nominalization status – dependent clauses, but not matrix clauses, exhibit nominal bound tense morphology and are nominalized.

[14]The idea that the nominal properties of a nominalization are contributed by a nominal projection is at the core of the Functional Nominalization Thesis (Kornfilt and Whitman 2011).

[15]In the Tupían family, a *n* selecting a functional complement in dependent clause structures is not unheard of. Vivanco (2018:87-130) argues that a cognate morpheme in the distantly related language Karitiana (Arikém family), embeds an Evidential Phrase.

4.3.3 Verbal agreement: a further argument for dependent clauses truncation

4.3.3.1 The pattern

Further evidence that dependent clauses are truncated can finally be appreciated, namely their lack of subject agreement. While matrix clause verbs can index either the subject or object due to the person hierarchy 1>2>3, dependent clause verbs index only the object (Dos Santos 2021a, 2023a), thus neutralizing the person hierarchy of matrix clauses. Therefore, while the 1>2 and 2>1 scenarios consistently show agreement with the more local person in (45), dependent clauses consistently show agreement with the object in (46).

(45) a. A-hepia ki ji nde ko.
1.SG.A-see PST 1.SG 2.SG REAL
'I saw you.'

b. Ji=repia ki nde ko.
1.SG.B=see PST 2.SG REAL
'You saw me.'

(46) O-kwaha ki ga [ji nde=repiak=a].
3.A-know PST 3.SG.MASC 1.SG 2.SG=see=NMLZ
'He knew I saw you.'

This pattern of object agreement only in dependent clauses becomes even more evident in non-local scenarios (i.e., 3>3). Recall from Table (4.1) that Kawahíva exhibits agreement with 3rd person subjects (marked with *o-*), but not 3rd person objects. If dependent clause verbs agree only with objects, then we expect the prefix *o-* to be ruled out from dependent clauses. This is true according to the ungrammaticality of (47).[16]

[16]It is important to note that in Kawahíva (and other TG languages) a prefix *o-* can appear on the verb of a dependent clause, but only when there is coreference between the matrix subject and either argument of the dependent clause (as described in §3.7.13 on 'coreferential marking'). This distribution is essential for characterizing the prefix *o-* in example (47) as part of the set of prefixes known as 'cross-referencing' prefixes of Tupí-Guaraní languages (Jensen 1998:for a family overview on this morphology), which are presented in Table 4.4, rather than being considered the 3rd person subject index (i.e., a SET A marker), the paradigm of which is also included in Table 4.4 for convenience.

PERSON/NUMBER	CROSS-REFERENCING MARKERS	SET A
1ST PERSON SG.	it–	a–
2ND PERSON SG.	e–	ere–
1ST PERSON PL. INCL.	nhãnde–	txi–
1ST PERSON PL. EXCL.	oro–	oro–
2ND PERSON PL.	pe–	pe–
3RD PERSON	v– ~ o–	v– ~ o–

Table 4.4: Kawahíva cross-reference agreement markers.

(47) *Ere-pia ki nde [jawara tapi'ira o-juka=a] ko.
2.SG.a-see PST 2.SG jaguar tapir 3.A-kill=NMLZ REAL
'You saw the jaguar killing a tapir.'

I take the absence of subject agreement in dependent clauses to suggest the existence of two distinct sources for verbal agreement in Kawahíva. This idea receives further support from verbal reduplication.

In reduplication, object markers, but not subject markers, can be copied into the reduplicant (Dos Santos 2023a). To illustrate, consider (48a), which presents a baseline example of object agreement without reduplication. In turn, (48b) presents an example of object agreement with reduplication. Notably, the object marker is copied into the reduplicant. Finally, (48c) reveals that object markers must be copied in reduplication.

(48) a. Are=rety ki gã ko.
1.EXCL.B=pull.down PST 3.PL REAL
'They pulled us down.'

b. [Arerety]–are=retyk ki gã ko.
[RED]–1.EXCL.B=pull.down PST 3.PL REAL
'They pulled us all down.'

c. *Are=[rety]–retyk ki gã ko.
1.EXCL.B=[RED]–pull.down PST 3.PL REAL
'They pulled us all down.'

Unlike the obligatory copying of object markers in reduplication, subject agreement markers do not have a similar requirement. Consider (49a), which is a baseline example of subject agreement in non-local scenarios (i.e., 3->3), where the verb is marked with *o-*; (49b) and (49c), respectively, demonstrate that the subject agreement marker is not copied and, and any attempt to copy it results in unacceptable sentences.

(49) a. Anhãnga o–kutu–pam j–urua.
ghost 3.A–pierce–COMPLETELY i–mouth
'It was the ghost that pierced all his mouth (the chief's mouth).'

b. Anhãnga o–[kutu]–kutu–pam j–urua.
ghost 3.A–[RED]–pierce–COMPLETELY i–mouth
'It was the ghost that pierced his mouth (the chief's mouth) multiple times.'

c. *Anhãnga [okutu]–o–kutu–pam j–urua.
ghost [RED]–3.A–pierce–COMPLETELY i–mouth
'It was the ghost that pierced his mouth (the chief's mouth) multiple times.'

The fact that the 3rd person morphemes in both paradigms are homophonous is a mere coincidence. As one can see from this table, some other members of the two paradigms, but not all, are homophonous. For instance, the 1st person (both singular and plural inclusive) and 2nd person singular forms are not homophonous.

Assuming that agreement dependencies arise from a functional head in the clausal spine that copies morphosyntactic features from a nominal (Chomsky 2001), resulting in the proper agreement markers, the absence of subject markers in dependent clauses and reduplication cannot be explained if the same functional head is involved in both subject and object agreement. However, this pattern becomes clear if we posit that subject and object agreement markers have distinct structural loci.[17] Next, I propose a derivation for verbal agreement in Kawahíva that captures this asymmetrical behavior between the two agreement sets. This pattern and the analysis also concur that dependent clauses are truncated.

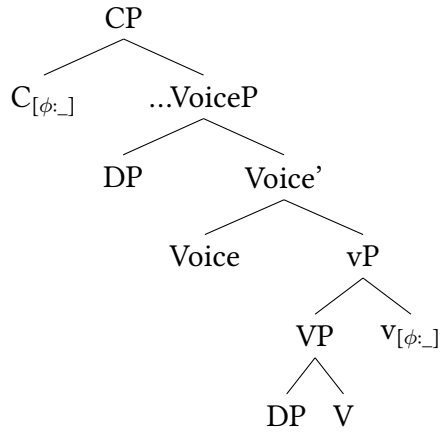
4.3.3.2 The derivation of verbal agreement

I treat agreement phenomena as dependencies that result from the syntactic operation *Agree* (Chomsky 2001). In particular, they involve a relation between a probe on a functional head and a nominal goal within its search domain.

Translating the Kawahíva system of agreement into this framework, I propose that each agreement set (i.e., SET A and SET B) is the result of Agree dependencies created through probes that are hosted on different functional heads of the clause structure. I also associate the lack of subject agreement with the absence of the CP within dependent clauses and propose that the CP domain is related to subject agreement. In addition, I propose that the *v*P head is implicated in object agreement.

More technically, matrix clauses come with two probes, one on the CP head and another on the *v*P head. Diagram (50) represents the clause structure for matrix clauses based on these ideas; a dash next to a feature in the probe specification means that the probe has not yet begun its search for that feature. In contrast, dependent clauses lack the CP domain and, consequently, the CP-probe (not shown here).

(50)



The rules of subject and object agreement are (51) and (52), respectively.

[17] The use of reduplication as a diagnostic tool in Avar by Rudnev (2020) served as an inspiration for the similar purpose of locating the structural source of agreement in Kawahíva.

(51) Subject agreement
Subject agreement is the overt spellout of an Agree relation between a probe on C and the features of an external argument.

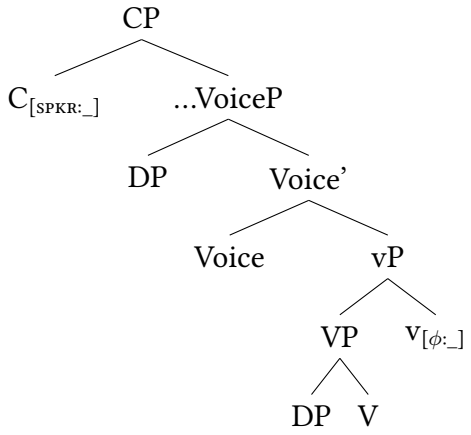
(52) Object agreement
Object agreement is the overt spellout of an Agree relation between a probe on v and the features of an internal argument.

Recall, however, that Kawahíva allows only one argument at a time to be indexed on the verb, and the choice is dependent on the person hierarchy 1>2>3. To account for these person hierarchy effects, however, the generalized [ϕ]-probes on CP and *v*P are not sufficient. Therefore, we need a more specific feature than [ϕ].

I propose the CP-probe is relativized to the feature [SPKR]. This is so because matrix clauses, but not dependent clauses, exhibit person hierarchy effects. Given the absence of CPs in dependent clauses, it follows that CPs carry the probe with the feature [SPKR] in matrix clauses. I propose the probe on *v* is a standard flat probe that can be satisfied by any [ϕ] feature-bearing goal.

The final representation of Kawahíva matrix clauses can be illustrated through diagram (53) after incorporating the agreement rules in (51) and (52).

(53)



We can now derive the verbal agreement pattern of matrix and dependent clauses. First, consider a local scenario 1>2 in a simple matrix clause in (54). In these cases, the CP-probe copies features of the external argument, and the *v*P-probe copies feature of the internal argument.

(54) A-hepia ki ji nde ko.
1.SG.A-see PST 1.SG 2.SG REAL
'I saw you.'

To ensure that the resulting agreement marker on the verb obeys the person hierarchy 1>2>3, and that *a-* is correctly chosen in (54), I assume this choice is calculated according to the *Subset Principle* (Keine 2010).

(55) Subset Principle, adapted from Keine (2010:8).
a) The morphosyntactic features of the agreement marker are a subset of the morphosyntactic features copied back by the probe.
b) The agreement marker has the highest number of morphosyntactic features that are a subset of the features copied by the probe.

Moreover, a preliminary version of the composite feature bundles of *a-* '1st singular' and *ere-* '2nd singular', assuming the feature-geometric representation proposed in (Harley and Ritter 2002), and their vocabulary insertion rules are provided in Table (4.5). These bundles will be revised shortly to include one more feature specification (i.e., Case).

[SPKR, PART, ϕ, SG, C]	$\longleftrightarrow$	a-
[PART, ϕ, SG, C]	$\longleftrightarrow$	ere-

Table 4.5: Vocabulary items for 1ST and 2ND SG. subject markers.

We can now derive the choice of agreement marker in different person scenarios within matrix clauses. Let's consider, for instance, the local scenario 1->2. In these cases, the features of 1st person are copied onto C, while the features of 2nd person are copied onto *v*. However, as there exists only a single verbal slot for verbal agreement, only a single feature bundle will be overtly realized. At Spellout, *a-* will be chosen as its features match the highest number of features with the subset of features on C, following the second clause of the definition of the Subset Principle in (55). The 2nd person subject agreement index *ere-* cannot be selected, as its features match a fewer number of the feature bundle on C.

Meanwhile, in dependent clauses, only the features of the object will be considered for overt realization, as these clauses lack the subject agreement probe. Since the object probe is always satisfied with the [ϕ]-features of any nominal in the internal argument position, the result is object agreement throughout all dependent clauses but relative clauses.

These assumptions are sufficient to correctly derive the person hierarchy effects. However, something must be said about the actual form of agreement marker chosen. Most combinations of person features exhibit two choices, one for subject agreement and another for object agreement. For example, in 1->2 scenarios, the agreement marker is *a-* '1st person singular subject', not *ji=* '1st person singular object'. I argue this choice is due to the first clause of the Subset Principle in (55) coupled with the Case feature of nominal arguments. First, I briefly lay out my assumptions about Case assignment.

I will assume that Case is assigned under a dependency between a nominal DP and a functional head. The heads involved in Case assignment are T and *v*. In matrix clauses, T assigns Case to its closest c-commanded DP, the external argument. In turn, *v* assigns Case to its closest c-commanded DP, the internal argument.[18]

[18]In intransitive clauses, T assigns Case to the sole argument of unergative and unaccusative verbs. The Case on *v* will not be discharged onto any DP; the sole argument of the unergative is in Spec,Voice, out of the c-command reach of *v*. The probe on *v* in unergatives will not find a goal in its search domain, and Agree fails, in the sense of Preminger (2011). The probe on C finds the unergative argument in Spec,Voice and copies its ϕ-features and Case

Putting together these assumptions about Case assignment and DP-internal features, Table (4.6) shows the composite feature bundles of the singular agreement markers in non-matrix clause scenarios. I indicate the Case feature on a nominal argument with K and a subscript indicates the functional head responsible for assigning Case (e.g., K_T means that Case is assigned by T).

1st person singular subject	[SPKR, PART, ϕ, SG, K_T, C]	$\longleftrightarrow$	a-
1st person singular object	[SPKR, PART, ϕ, SG, K_v, v]	$\longleftrightarrow$	ji=
2nd person singular subject	[PART, ϕ, SG, K_T, C]	$\longleftrightarrow$	ere-
2nd person singular object	[PART, ϕ, SG, K_v, v]	$\longleftrightarrow$	nde=
3rd person subject	[ϕ, K_T, C]	$\longleftrightarrow$	v-
3rd person object	[ϕ, K_v, v]	$\longleftrightarrow$	ø

Table 4.6: Feature composite and vocabulary insertion rules of a set of person markers

According to Table (4.6), in 1->2 local scenarios, a 1st person will be selected by the Subset Principle over a 2nd person for the verbal agreement slot, as per the previous discussion. Additionally, the agreement marker *a-* '1st person singular subject' wins over *ji=* '1st person singular object' because the feature bundle copied onto C is an exact match of the feature bundle of *a-*, including the Case feature. In contrast, the feature bundle of *ji=*, while including features that match most of the features on C, differs from it for having a distinct Case feature.

To summarize, in §4.2.3, I showed several pieces of evidence that dependent clauses are truncated as they lack the CP (e.g., they miss a landing site for extracted constituents). In this section, I discuss additional evidence for this claim based on verbal agreement. In particular, I showed that matrix clauses exhibit agreement with the subject or object. However, dependent clauses show agreement only with objects. A unified account of the differences between matrix and dependent clauses that also capture this asymmetry in agreement marking between both clauses is possible: the absence of subject agreement in dependent clauses follows from the lack of the structural position that hosts subject agreement in matrix clauses, the CP layer, independently supported by the pieces of evidence shown in §4.2.3.

Next, I discuss three analytical alternatives to derive V1 in Kawahíva, starting with the VP remnant movement. Then, I consider morphological amalgamation in section (5.3), and finally, syntactic head movement (or head-to-specifier movement) in section (4.4).

4.4 Remnant movement

One initially appealing analysis of Kawahíva verb movement would be as Remnant VP movement. This view has been advocated elsewhere, including for Zapotec (Lee 2000), Niuean (Massam

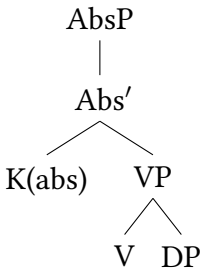
feature assigned by T. As for unaccusatives, the unaccusative v does not assign Case (The Unaccusativity Hypothesis). The unaccusative argument must evacuate the VP and move to Spec,Voice so as to receive Case from T. This movement is triggered by an EPP feature on v. The probe on unaccusative v copies the features of the unaccusative argument. However, as I propose in section (4.4.1.5), the EPP feature, not the ϕ-features of the unaccusative argument, is spelled out. Finally, the probe on C can no longer interact with the unaccusative DP because the latter had interacted with v already.

2000, 2001) Malagasy (Rackowski and Travis 2000), the related Tupí-Guaraní language Tenetehára (Duarte 2012), among other languages; see (Thiersch 2017) for a lengthy overview of this approach.

A remnant movement analysis predicts two steps of movement, one for the object and another one for the complement-free VP. The movement of the object is motivated, in some languages, by the need of a DP to check a case feature. This idea builds off the assumption that, while NPs do not have case features to be checked, DPs do (see Bittner and Hale (1996); Giorgi and Longobardi (1991)).

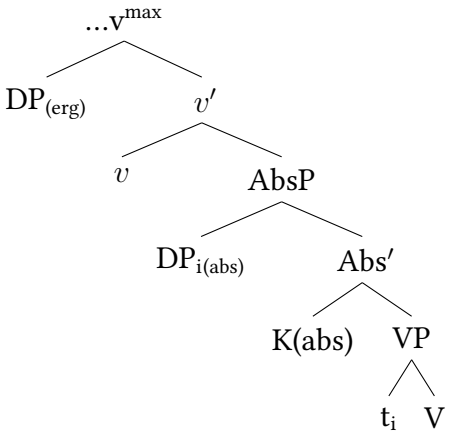
As said, the idea that VP remnants are involved in the creation of V1 clauses has been proposed for the VSO/VOS alternating orders of Niuean (Seiter 1979; Massam 2001), an ergative-absolutive language. In this language, the first step in the creation of VSO clauses is object DP movement to the specifier of Abs(olutive)P. A diagram showing the underlying structure of the VP and this functional projection is presented in (56); unnecessary details are omitted. At this point in the derivation, the DP object is VP-internal but it needs to check its case feature.

(56)



The DP object checks its case feature by moving to Spec,AbsP, thus evacuating the VP, which would contain only V in the simplest case.[19] The derivation proceeds with the merge of *v* and the ergative DP.

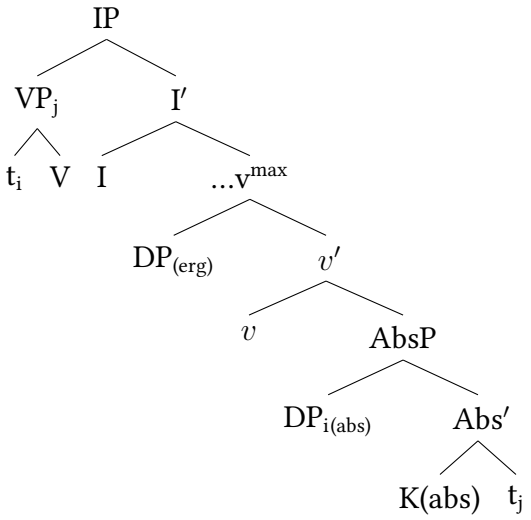
(57)



[19]If other VP-internal elements are present, evacuation of these VP-internal elements may be required.

After that, the head of IP is merged. Massam (2001) — building off previous works (Massam and Smallwood 1997; Massam 2000) — assumes the head of IP has a strong predicate feature [pred] with the EPP property for movement. This feature attracts the VP to the specifier of IP, namely the VP, yielding VSO in Niuean.

(58)



Tenetehára, a sister language to Kawahíva, with alternating VSO/VOS orders, has also been analyzed as undergoing remnant movement by Duarte (2012). In this account, VSO follows from the creation of a remnant VP that moves to Spec,CP – this landing site is proposed based on the surface order of the verb, which precedes the temporal adverbials adjoined to TP. Additionally, VOS is the result of the movement of the VP with an overt object. Duarte (2012) claims that the object evacuates the VP due to Case. Therefore, like Massam (2001), the author draws on the idea that only nominals with a DP layer need to be Case-assigned. Duarte (2012:372) draws his conclusion that DPs evacuate the VP in VSO, but not VOS, from positive evidence showing that D-level material (e.g., demonstratives) is present with objects in VSO clauses, but not objects in VOS clauses. In the latter cases, as the object does not evacuate the VP, it will be dragged along to the clause-initial position as the VP fronts.

There are two pieces of evidence against the remnant approach to V1 clauses in Kawahíva, including one piece that directly shows that these clauses do not involve object DP movement prior to the fronting of VP.

The first evidence against a VP-raising analysis is the absence of VOS constituent order in informationally neutral matrix clauses, prevalent in V1 languages that have received a Remnant VP account, including Niuean (Massam 2001), Malagasy (Pearson 2001), among others. Example (59) cannot mean that Clebson saw a jaguar, which the parsing of this sentence with a VOS order would lead us to expect, regardless of the definiteness status of ‘jaguar’.

(59) V-epia jawara Clebson=ga.
3.A-see jaguar Clebson=3.SG.MASC

‘A jaguar saw Clebson.’

‘*Clebson saw a/the jaguar.’ (Jupaú: Elicit)

One could maintain that the VSO order of Kawahíva is derived via object shift followed by movement of the remnant VP, even for those cases wherein the object is an NP (rather than a DP). This, however, would be against the very reason object shift was proposed to account for the verb-initial orders in VSO/VOS alternating languages, namely the need for Case licensing on objects with a DP layer. Therefore, I take the lack of informationally neutral VOS order in Kawahíva to be unexpected under the Remnant VP approach to verb fronting in the language.[20]

Having shown negative evidence that object shift is not involved in Kawahíva V1, I discuss a positive piece of evidence that these V1 clauses do not involve object evacuation from the VP next.

4.4.1 The argument from the pattern of *i-* prefixation

The second piece of evidence that V1 clauses in Kawahíva do not involve a step of object movement comes from a language-internal argument based on the distribution of the prefix *i-*. This prefix is triggered when the object vacates the VP.

Three instances of object movement out of the VP trigger the prefix *i-*. In the first case, a thematic object, but not the verb, vacates the VP. Secondly, a thematic object is dropped, and the verb does not leave the VP. The third and final case is when a thematic object vacates the VP prior to the movement of the verb. In brief, the appearance of *i-* is tied to the absence of an overt DP within a VP with an overt verb. Table (4.7) summarizes the structural context for the occurrence of *i-*.[21]

[20]Interestingly, Chung (2005) sets apart languages with strict VSO order in pragmatically neutral clauses in a discussion about whether all verb-initial orders should be derived by a Remnant VP approach.

[21]There is a different reflex than *i-* found on a subset of verbs when they meet the same conditions that lead to *i-*. The members of this subgroup instead undergo root allomorphy wherein their initial segment surface with either [h] or no initial-consonant under conditions a) through c) above, or with an initial [ɾ] otherwise.

	h-/r- verbs	*i*-marked verbs
VP-INTERNAL OBJECT DPS	[DP repia]$_{VP}$	[DP mõndo]$_{VP}$
VP-EXTERNAL OBJECT DPS	(DP), [hepia/epia]	DP, [i-mõndo]
	‘see DP’	‘throw away DP’

Table 4.7: The distribution of the *h-/r-* and i-marked verbs.

Table (4.7) illustrates the resulting form of the verb under the same conditions that lead to the *i*-prefixation alongside the parallel in verbs that surface with *i-* instead. Whether a verb will surface with one or the other pattern can only be lexically determined synchronically (see Meira and Drude (2013) for a historical development that led to this present-day two-way split). In the next subsections, I focus on the subset of verbs that show the pattern of *i*-prefixation because it includes a much larger number of verbs compared to the number of verbs that have the pattern of root allomorphy.

	i-marked verbs
VP-INTERNAL OBJECT DPS	[DP mõndo]$_{VP}$
VP-EXTERNAL OBJECT DPS	DP, [i-mõndo]
	'throw away DP'

Table 4.8: The distribution of the i-marked verbs.

The relevance of the *i*-prefixation to the present discussion is an expectation within the Remnant VP approach to V1 related to object movement prior to VP fronting. Given that *i-* is triggered under object movement out of the VP, then a clear prediction arises: if V1 clauses in Kawahíva involve a step of object movement prior to VP fronting, V1 clauses must surface with an initial verb marked with *i-*. However, this is only true in the case of unaccusative verbs. We will see that even in these cases, the use of *i-* follows from an independent reason, namely the derivational life of unaccusative verbs. The failed prediction of the Remnant VP approach that clause-initial verbs should bear the prefix *i-* argues against it as the means to derive V1 in Kawahíva.

It should be noted that the pattern of *i*-prefixation is not restricted to verbs, nor it is exclusive to Kawahíva. The prefix *i-* extends to all cases where the complement of a head is moved or dropped, which includes objects of postpositions and inalienable possessors. These other instances will be discussed in (4.4.1.4). The pattern of *i*-prefixation is also widespread across the Tupí-Guaraní family, wherein this prefix is referred to as a "relational prefix" and is generally described as a marking of syntactic adjacency between a complement and its selecting head (Rodrigues and Cabral 2012a). Rodrigues and Cabral (2012a:511) write that this prefix indicates "in a stem that it forms a syntactic unit with its determiner, which is the immediately preceding expression." Similar conditions for the appearance of a cognate morphological reflex are described for members of distant-related language families of the Tupían stock. These include the single member of the Arikém family Karitiana (see Storto (1999:194-204) and Vivanco (2018:133-177)) and the Tuparí language Mekens (Galucio 2001, 2002), wherein a complement-less VP leads to the morpheme *ti-* showing up on the verb.

4.4.1.1 Object movement in the absence of verb movement

One instance wherein object movement triggers the prefix *i-* involves dependent clauses and matrix clauses with a clause-initial PP. In both structures, the verb whose object is moved shows up final in regard to its arguments. We already saw that the verb is final in dependent clauses, a baseline example of which is (60), reproduced from Dos Santos (2022). I will show that PP-fronting competes with V-fronting in (4.7), thus making the verb also occur clause-finally.

(60) Ere-piang nde [apinaga pira 'u=a] rai'i.
2.SG.A-see 2.SG father fish eat=NMLZ YESTERDAY.PST
'You saw Father eating fish.'

I take the final position of the verb within the complement clause to mean the verb stays in its base-generated position, i.e., right-heading the big VP, given the proposed structure of a dependent clause in (44). Focusing first on dependent clauses, non-relative extraction of the object

triggers the prefix *i-* on the dependent verb, as demonstrated with example (61); otherwise, the sentence becomes unacceptable, as in (61b).[22]

(61) a. Gara nde ere-piang [apinaga i-'u=a] rai'i?
what 2.SG 2.SG.A-see father i-eat=NMLZ YESTERDAY.PST
'What did you see Father eating?'

b. *Gara nde ere-piang [apinaga 'u=a] rai'i?
what 2.SG 2.SG.A-see father eat=NMLZ YESTERDAY.PST
'What did you see Father eating?'

The second verb-final structure where *i*-prefixation occurs is in clauses where an adverbial PP/particle has been fronted and the object leaves the VP due to its pragmatic and/or semantic status, as illustrated in (62). Several independent observations support this claim, including an intonational break between the pre-verbal object and the verb – which is marked via a comma – the associated pragmatic and/or semantic status(es) of the object, and the presence of *i-*.

(62) A'ero jie te, 'yhea, i-mboapyg-i, 'i:te:ki:ko:ra.
then 1.SG REALLY insides i-cook-i DIR.EVID
'Then I cooked (*the) (peccary's) insides.' (Juma: text)

The factors that condition the intonational break between the object and verb, and the consequent marking with *i-* on the verb of Adverbial-SOV matrix clauses are those usually associated with pragmatic topics. In particular, the referents of these objects are *active* and *given* in the discourse (Chafe 1976; Lambrecht 1994). Other properties that correlate with the prosodic break setting the object apart from the verb are the semantic features *definiteness* and *specificity*. While objects might have all these properties at once, sometimes they have only one. For instance, only the definite status of *'yhea* 'bowel' conditions the prosodic break between the object and the verb in (62), given the context in which this particular example is uttered.[23]

The intonational break between the preverbal object and the verb, and the consequent triggering of *i-*, become highly suggestive that the object is VP-external when we consider that prosodic constituency can be a diagnostic for syntactic constituency. If we assume that prosodic

[22]Object relative clauses show a different morphological reflex as in Kawahíva they lead to *wh*-agreement for Case with the relative head and never use *i-* in object relative clauses. There are two exponents of *wh*-agreement with the Case feature of a relative object, *-pyr* and *remb-*; an example of the latter can be seen in (37b). The difference is determined by a nuanced contrast between completive vs. non-completive readings, as well as the distinction between wholly affected and partially affected readings (as described in §5.4.1).

[23]The sentence is extracted from a 362-sentence personal narrative self-recorded by the speaker, wherein she narrates a land patrol activity during which some people went hunting. In line 184, the speaker describes that one group member killed a peccary (wild pig). The next four sentences are the speaker's elaboration of the *postmortem* cleaning of the peccary by herself, which culminates with the sentence in (62). The referent of the peccary is given at this point, and so the peccary is not overtly mentioned. Its bowel, however, had not been mentioned, and so its first mentioning through the phrase *'yhea* 'bowel' is overtly realized. Although this is the first mention of *'yhea*, its possessor can be easily recovered from the context. This is so because the speaker has been talking about the peccary and its cleaning all along.

constituents generally match syntactic constituents, as Selkirk (2011) argues, then the object in example (62) is parsed into a prosodic constituent different from that where the verb belongs. Consequently, if prosody reads off syntactic structures, then the object and the verb belong to distinct syntactic constituents. The idea that a different prosodic constituency is indicative of a different syntactic constituent is further strengthened by the morphological reflex of *i-* on the verb in (62), which would meet the same structural condition of the *i-* that appears as the result of object extraction from an embedded clause to the clause-initial position of the matrix clause, as demonstrated based on the example in (61) from above. All in all, these facts support that preverbal objects are not in their thematic position in sentences like (62), that is, they occupy a position external to the VP.

The discussion developed above leads us to expect that VP-internal objects will have the opposite pragmatic/semantic status (nonspecific) compared to VP-external objects (e.g., specific), in addition to no prosodic break between it and the verb. As far as I can tell, based on my spontaneous speech corpora, all illustrative examples realize this prediction. In the following example, a clause-final verb within a purposive clause is preceded by its thematic object. Notably, no prosodic break between the two is observed, and the bare plural object *agoutis* of the translation correctly captures the specificity and givenness status of the object in the context where this example is uttered.[24]

(63) Davi=ga ho-i, karuvoruhua juka-vo, ’i:te:ki:ko.
Davi=3.SG.MASC go-i agouti kill-VO DIR.EVID
‘(Then) Davi went to kill agoutis.’ (Juma: text)

Now that we have seen the associated readings resulting from a VP-internal and VP-external object, we can tie the movement of the object outside the VP as the consequence of the pragmatic/semantic effects that would be associated if it stayed in its thematic position. Take specificity, for instance. It is a standard assumption that indefinite objects receive an obligatory specific reading if they move out of their thematic position (see Diesing (1992)), but can be ambiguous if they stay VP-internally[25].

4.4.1.2 Object drop in the absence of verb movement

Instances of object omission in verb-final clauses also lead to *i*-prefixation on the final verb. Consider (64) for an example.

[24]This sentence is drawn from the same text wherein (62) was uttered. In it, *karuvoruhua* ‘agouti(s)’ is mentioned only once – the sentence in (63) – right after the speaker describes a couple of events, including that it was already evening, she cooked, she and her land patrol team ate everything, and they all went to sleep, except her partner (sentence (63)), who went out to kill agoutis, and his nephew came along. To go after agoutis after having dinner is a common practice among the Juma and Jupaú. Despite the latter, one could go out to hunt many other things that are not agoutis. This might be one other reason that led the speaker to avoid dropping the object of *juka* ‘kill’ in (63).

[25]See Vicente (2009) for a rejection of the Remnant VP approach to derive predicate topicalization in Spanish based on the fact that the object in predicate topicalization can be read as nonspecific, contrary to the prediction expected under a Remnant VP approach.

(64) Aramẽ ki ka, i-mõmbo-i repiag=a.
after.that PST 3.SG.MASC i-throw-i see=NMLZ
'After that he threw (it = fishing hook) to see.' (Juma: Text)
Context: A speaker is narrating a fishing day. Five sentences earlier, she said that she and her partner prepared their *vỹrỹtxĩ'ĩa* 'fishing hook,' made a bet on who would first catch a fish, and then threw it the first time in vain. Then, her partner tried it again, which the speaker describes in the sentence above.

I argue these cases can be subsumed under the previous cases we saw where the pragmatic and/or semantic status of the object is the key element conditioning a VP-external position of a preverbal object. The difference seems to be in whether the referent of the object is given or not in the discourse. In particular, the givenness status of the object in (64) seems to be the important feature for its omission, which the context helps to identify. As it shows, the referent of the object has been introduced, and the speaker wants to refer back to it, but without overtly using linguistic material to do so (clearly because the speaker assumes that the interlocutor will be able to recover its referent from the context as it is still very active in the discourse given a recent mentioning). The object is omitted, and the verb is marked with *i-*.

The structure of sentences with object drop can easily be unified with the structure of sentences wherein the object is overt and VP-external, as discussed in the previous subsection for the example in (62). Consequently, the idea that the object is in a position external to the VP in (62) can also be extended to these cases with null objects like (64). I concur that the syntactic position of the object in (64) is outside the VP.

4.4.1.3 Unaccusative verbs

The third and final instance of *i*-prefixation is that on unaccusative clauses, as example of which is presented in (65) (see section 4.2.1 for further discussion).

(65) I-kari ga.
i-be.jealous 3.SG.MASC
'He is jealous.'

The use of *i-* on unaccusatives follows from the conditions that lead to *i-* and the well-known fact that the sole argument of these verbs vacates the VP for Case. The latter is a consequence of the inability of unaccusative verbs to license Case onto their sole argument. The appearance of *i-* on this subset of intransitive verbs is the strongest (and only) piece of evidence that these verbs are unaccusative, as proposed in Dos Santos (2022).

Dos Santos (2022) proposes an account of the *i*-prefixation that supports the analysis of these verbs as unaccusative in the absence of standard tests for unaccusativity in Kawahíva. This proposal conjoins the results of the use and distribution of *i*-prefixation beyond this subset of intransitive verbs – which was described in subsections (4.4.1.1)-(4.4.1.2) – and the derivational life of unaccusative verbs. In terms of the latter, the idea that these are unaccusative verbs draw on

the standard assumption that their sole argument, which is base-generated in the complement position of the verb (Perlmutter 1978), evacuates this position (possibly due to Case licensing as unaccusative verbs cannot license Case). Consequently, given how *i-* is triggered, namely, upon the absence of a thematic object within the VP with a stranded verb, and the sole argument of an unaccusative verb leaves the VP, then it follows that the condition for *i*-prefixation is also met during the derivation of unaccusative clauses.

Conjoining the Remnant VP movement approach and the independent morphological reflex of object movement out of the VP leads to the expectation that V1 clauses show up with the initial verb bearing the prefix *i-*. Let us see this in detail. A remnant movement of a VP involves the pre-movement of the object. The movement of objects out of a VP triggers the morphological reflex of *i-* on the verb that is left stranded inside that VP. Subsequently, the VP raises without the object. Consequently, V1 clauses would consistently show up with *i-* attached to the initial verb. The failure of this prediction is patently demonstrated by examples (66-67), which include clause-initial transitive and active intransitive verbs, respectively; only in those cases where the verb bears a verbal agreement marker, but not *i-*, is the sentence well-formed. with and without *i-*.

(66) a. A-hepia ki ji ga.
1.SG.A-see PST 1.SG 3.SG.MASC
'I saw him.'

b. *I-hepia ki ji ga.
i-see PST 1.SG 3.SG.MASC
'I saw him.'

(67) a. A-kwam ki ji.
1.SG.A–dance PST 1.SG
'I danced.'

b. *I-kwam ki ji.
i–dance PST 1.SG
'I danced.'

In brief, the overall distribution of *i*-prefixation extends to the subset of unaccusative verbs that surface with this marking and its account does not involve a Remnant VP movement.

4.4.1.4 The pattern of *i-* prefixation beyond verbs

The absence of an overt complement to a postposition also triggers the pattern of *i*-prefixation on the latter, thus showing this pattern extends beyond verbs:

(68) a. **Avo uvi** i-ho-i, kahua.
LOC.DEM:PROX from i-go-i car
'It will leave from here, the car.' (Juma: Text)

b. Pevo-'i po ahe ho-i o-vag=a **j-uvi**, a-'i ki ji
LOC.DEM:DIST-DIM IRR people go-i 3COR-return=NMLZ i-from 1.SG.A-say PST 1.SG
ko=ra.
REAL=TODAY.PST
'(Let's) go a bit further there to turn back from (there).' (Juma: Text)

Additionally, *i-* is also used to mark an inalienable noun (e.g., body parts) within a possessive phrase whose inalienable possessor is absent; I assume inalienable possessors are syntactic complements of inalienable nouns (Alexiadou 2003:among others).

(69) a. O-mbo-tyky-tyky hẽa **mbiara kava** rereko-vo.
3.A-CAUS-RED-drip 3.SG.FEM meat fat have-VO
'She had the meat's fat to drip.' (Juma: Text)

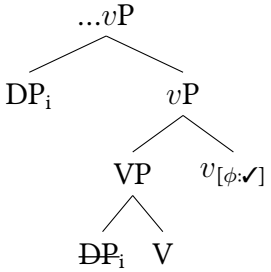
b. **I-kava** ko.
i-fat REAL
'(She was frying meat's) fat.' (Juma: Text)

The characterization of *i-* in the previous sections, as resulting from the absence of an overt verbal object within a VP containing an overt verb, or the "premature" movement of the object before the verb, can naturally be extended to these instances of *i*-prefixation. In this broader pattern, *i-* is marked on all heads whose complements meet these conditions.

4.4.1.5 A proposal and the relevance of the pattern of *i*-prefixation

I propose the *i*-prefixation pattern is the morphological reflex of a functional head bearing an edge feature with the EPP property. In particular, I propose that the [ϕ]-probe on *v*P can be optionally merged with the EPP property to facilitate movement of the DP to the edge of *v*P; in essence, a kind of object shift (Holmberg 1986, 1999). This idea is represented in (70).[26,27]

(70)



[26]I also posit that an edge feature is present in the instances of *i*-prefixation beyond VPs, that is, PPs and PossPs headed by inalienable nouns, as discussed in section (4.4.1.4). In these cases, the functional heads bearing the edge features are on the heads of *p*P and *n*P, respectively.

[27]A similar proposal that an edge feature is spelled out overtly on *v* is given to the tonal changes on a verb in instances of A'-movement in the Kwa language Asante Twi (Korsah and Murphy 2020).

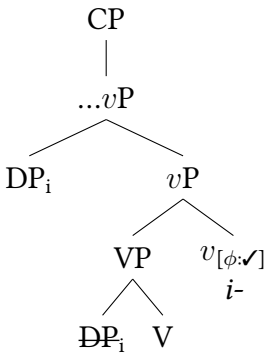
When the object leaves the VP with the verb stranded inside, or when the object evacuates the VP prior to the verb head – for Case reasons as in the scenario with unaccusatives – the EPP feature is merged with the [ϕ]-probe on *v*. In those cases, *v* spells out with *i-*. In contrast, if the object does not move, the EPP feature is not merged with the [ϕ]-probe. In those cases, *v* does not spell out with *i-*. More technically, the rules for vocabulary insertion at Spellout in these two scenarios are as follows:

[ϕ, EPP, *v*]	$\longleftrightarrow$	i-
[ϕ, *v*]	$\longleftrightarrow$	$\emptyset$

Table 4.9: Vocabulary items for *v*.

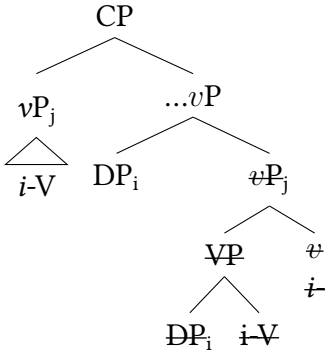
This pattern of prefixation with *i-* in Kawahíva gives us an empirical prediction about whether V1 clauses involve remnant movement. Under this view, V1 in Kawahíva should have resulted from evacuating the object outside the VP and then fronting the VP. The first step, which is illustrated in diagram (71), is object movement. It moves the object out of the VP prior to verb fronting, resulting in a VP whose head is internal but without a complement. This is the same configuration where *i-* is triggered elsewhere, as discussed thus far. Consequently, *i-* should be realized on *v* at Spellout.

(71)



The second step is Remnant VP fronting, as represented in (72). The result should be a clause-initial *i*-marked verb in all instances of clause-initial verbs.

(72)



This prediction is not correct, however, as example (73) demonstrates — as well as all other instances of clause-initial verbs except with unaccusative verbs, which bear *i-*.

(73) *I-'u ki jie pira ko.
i-eat PST 1.SG fish REAL
'I ate fish.'

4.4.2 Interim summary

This section presented evidence against the idea that the creation of a Remnant VP is involved in Kawahíva V1 clauses, given that there is a morphological reflex of object movement with the prefix *i-*. The prediction under the Remnant VP account that V1 clauses would surface with *i-* is not confirmed, however.

A note is in order about the pattern of *i-* prefixation across the Tupí-Guaraní family in regards to the formation of V1. We saw that at least the related language Tenetehára shows the V1 pattern (Duarte 2012), and the analysis proposed is the Remnant VP. Given that Tenetehára also shows the pattern of *i-* prefixation (Duarte 2007), the discussion in this section would suggest that Tenetehára V1 clauses should show *i-*. This is not confirmed, however, as the examples in Duarte (2012) demonstrate. However, as far as I can tell, an alternative analysis compatible with the facts presented in Duarte (2012) is possible. In it, the verb head undergoes syntactic head movement, and the bare NP object is prosodically reordered next to the verb in the postsyntactic component, thus resulting in the VOS order. If syntactic head movement is the actual means that derives V1 in Tenetehára, then the absence of *i-* prefixation in Tenetehára V1 is actually expected. A prosodic reordering analysis for alternating VSO/VOS languages is proposed for a subgroup of Mayan languages (Clemens and Coon 2018) and the Polynesian language Niuean (Clemens 2019).

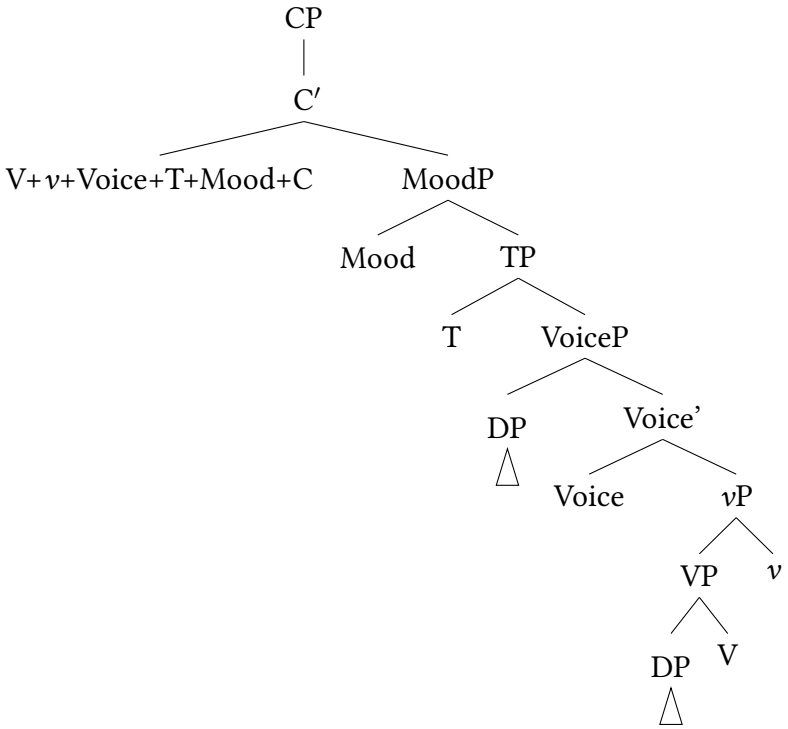
4.5 Morphological Amalgamation

Kawahíva V1 clauses might be the result of morphological amalgamation.[28] This view has been proposed in the account of the V1 pattern of several languages (see Carnie and Guilfoyle (2000); Carnie et al. (2005)). We will see here that this analysis is not only empirically inadequate, but verb fronting in Kawahíva does not show the crosslinguistic hallmarks of morphological amalgamation.

Under the idea that Kawahíva V1 is created via morphological amalgamation (Harizanov and Gribanova 2019), the verb moves locally through all heads on its way up to the clause-initial position, as diagram (74) illustrates, assuming the Kawahíva matrix clause structure proposed in section (5.3):

[28] Morphological amalgamation is the postsyntactic counterpart of syntactic head movement, as proposed in Harizanov and Gribanova (2019). I refer to this movement as 'morphological amalgamation,' rather than 'head movement,' to avoid a long-standing controversy around this term and its implications (cf. Dékány (2018) for a critical assessment of the different properties under the phenomena with the rubric 'head movement'.).

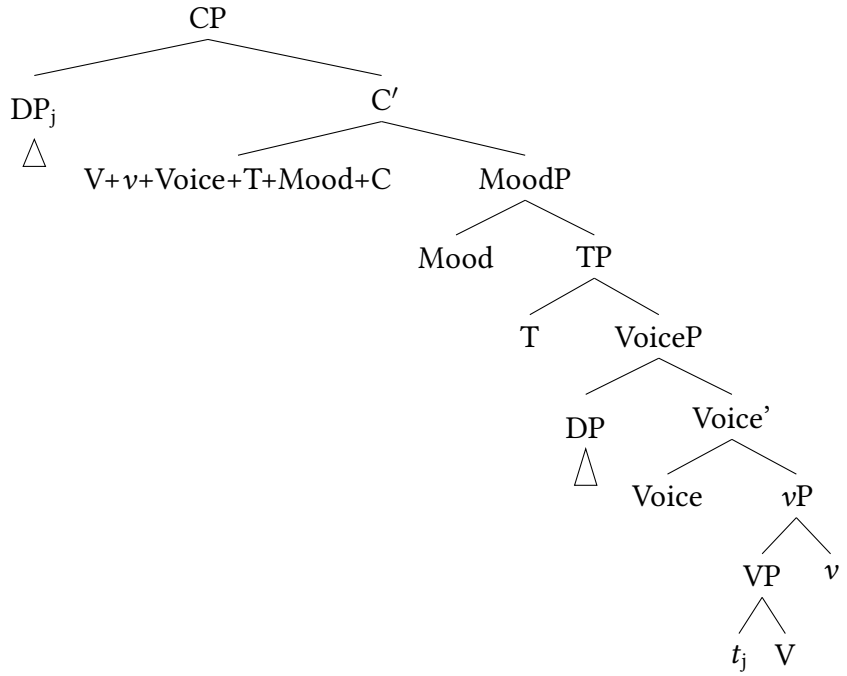
(74)



The first issue with the morphological amalgamation analysis is a failed prediction that a verb fronted to C should be able to co-occur with overt elements in Spec,CP. After all, in the head movement demonstrated in (74), the verb landing in C and the filling in of Spec,CP should have been compatible. However, we have seen this is not possible, and, in fact, informational prominent phrases and verbs are in complementary distribution for the clause-initial position.

One alternative compatible with these facts would be to assume this complementary distribution follows from a version of the Doubly-Filled Comp filter (Chomsky and Lasnik 1977), which rules out sentences like *I wonder who (*that) they saw* in English, for instance. To illustrate in a concrete fashion, this proposal says that Kawahíva does not allow matrix clause configurations like (75), where both Spec,CP and C are filled in – the former with an extracted object, and the latter, with the verb after verb movement.

(75)



Appealing to the Doubly-Filled Comp constraint as an alternative to the mentioned complementarity is, at best, a stipulation since its structural properties in English are not found in Kawahíva. So, while the Doubly-Filled Comp restriction in English is exclusive to dependent clauses and involves wh-words and complementizers, in contrast, it would be a root phenomenon in Kawahíva, would not involve only wh-words, and would never implicate complementizers.

Another strong piece of evidence against this analysis is the absence of the hallmarks of morphological amalgamation in Kawahíva verb fronting. The core properties of this movement are listed below, following Harizanov and Gribanova (2019):

- (i) Morphological amalgamation obeys the Head Movement Constraint (HMC) (Travis 1984).
- (ii) Morphological amalgamation directly reflects syntactic structure (Baker (1985)'s Mirror Principle).
- (iii) Morphological amalgamation creates a head that functions as a unit for the purposes of phonology and syntax (Harizanov and Gribanova 2019).[29]

[29]I do not discuss here a fourth aspect of morphological amalgamation, that morphological amalgamation is clause-bound. In (4.6.2), I discuss some data showing that embedded verbs do not front to the upstairs clause, which suggests this verb movement is clause-bound. I show this aspect can be accommodated in the syntactic head movement account I advocate for Kawahíva V1 in (4.6). Additionally, its clause-boundedness might also suggest that Kawahíva V-fronting is a type of A-movement – which is typically clause-bounded – which comes without surprise since it is not A'-movement, as discussed in (4.6.3). All, the clause-boundedness of Kawahíva V-fronting is unexpected under the idea that it is syntactic and non-A'-movement, as proposed in this chapter.

Note how the HMC – which states that a head may only move to the next head up, or in other words, a head cannot skip the next head up (Travis 1984) – contributes to morphological amalgamation to obey the "mirror generalization" – that is, the order of syntax structure is reflected in the affix ordering.

Firstly, Kawahíva verb fronting does not obey the HMC and does not yield mirror effects because it skips at least the T head; we will see shortly that it also skips Mood. This is one way to see (76), where the particle *po* 'irrealis' is closer to the verb stem than *txi* 'future'.

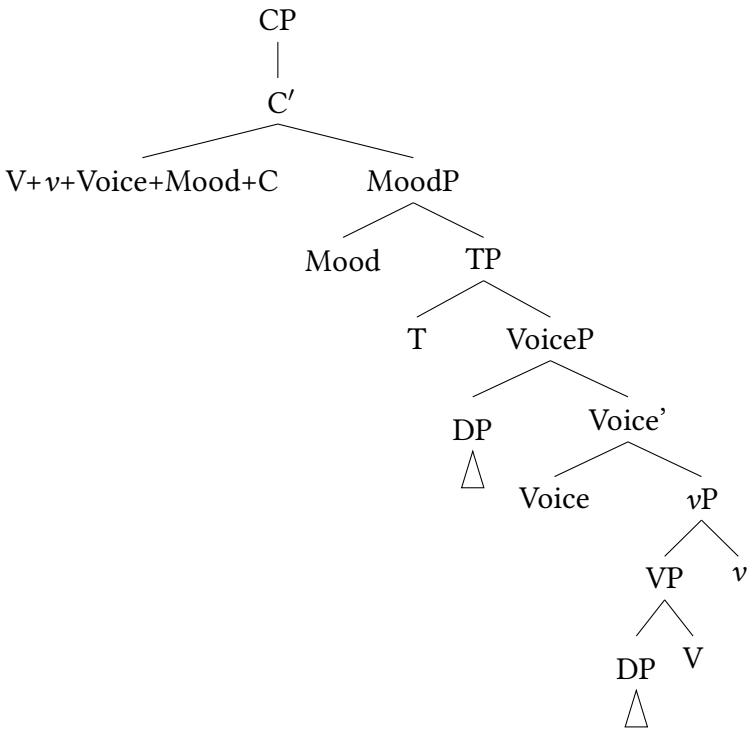
(76) A-hepiang **po** **txi** jie gã nãhẽ.
1.SG.A-see IRR FUT 1.SG 3.PL POT
'I perhaps see them.' (Juma: Elicit)

If Kawahíva verb fronting were HMC-compliant, which would also create mirror effects, the exponent of T, *txi* 'future,' would surface closer to the verb stem than (one of) the exponent(s) of Mood, *po* 'irrealis.' In fact, we would expect the following order to be grammatical:

(77) *A-hepiang **txi** **po** jie gã nãhẽ.
1.SG.A-see FUT IRR 1.SG 3.PL POT
'I perhaps see them.' (Juma: Elicit)

However, if we assumed the movement that derives V1 skips at least T, then (76) above would follow straightforwardly. To illustrate concretely the structure that underlies this example, we could schematically represent it with the tree in (78).

(78)



In sum, (76) and (77) strongly suggest that verb fronting in Kawahíva is not local (non-HMC compliant) and yields anti-mirror effects. These properties are unlike morphological amalgamation, which is strictly local and has a tendency to generate mirror effects.[30]

Finally, verb fronting in Kawahíva does not result in a syntactic or prosodic unit. To see that the clause-initial verb is not a syntactic unit, consider (79), where the adverbial *tūhūi* 'really, even' appears between the clause-initial verb and irrealis particle.

(79) V-eruri **tũhũi** po ka o-pira, 'i ki ji ko=ra.
3.A-bring INDEED IRR 3.SG.MASC 3.COR-cloth say PST 1.SG REAL=TODAY.PST
'He indeed brought his own clothes, I said.' (Juma: Text)

Assuming morphological amalgamation creates a head that is a complex syntactic unit (Harizanov and Gribanova 2019), it is a surprise that the adverbial can interrupt the sequence with the verb and irrealis particle, both of which are on C, following the idea that V1 is the result of amalgamation (74).

Furthermore, the initial-clause verb fails to show the expected prosodic behavior of heads created via morphological amalgamation, as not only does the verb bear stress in clauses where it is initial but so does the irrealis particle. If prosodic constituents generally match syntactic constituents (Selkirk 2011), the sequence verb and particle would map into one single prosodic constituent, contrary to fact.

Additionally, the breakup between the initial verb and the irrealis particle, both syntactically and prosodically, strongly suggests they are not on the same head (i.e., C). I take this to mean that the verb also skips MoodP on its way to C, lending further support that Kawahíva V1 is not strictly local.

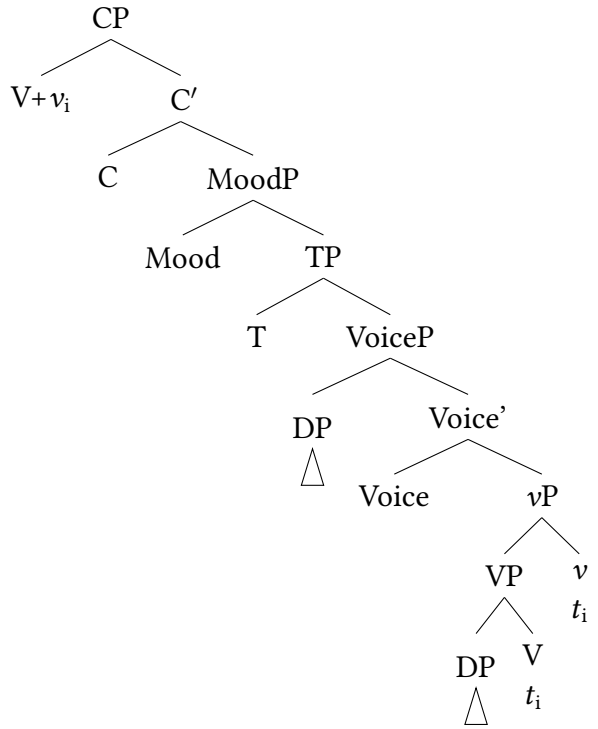
In this section, I have shown that Kawahíva verb fronting does not show the properties generally associated with morphological amalgamation crosslinguistically, as it does not comply with the Head Movement Constraint (Travis 1984), nor the Mirror Principle (Baker 1985), and does not show the syntactic or prosodic hallmarks of heads created via this type of movement (Harizanov and Gribanova 2019). Therefore, I conclude that Kawahíva V1 clauses do not involve morphological amalgamation.

4.6 Syntactic head movement (Head-to-Specifer movement)

I consider syntactic head movement as an ultimate means to derive VSO in Kawahíva. In particular, VSO results from verb fronting to Spec,CP, as illustrated in (80):

[30]There are certainly cases where anti-mirror effects arise, as in Kîîtharaka (Muriungi 2008), but these are not as frequent as one might expect.

(80)



A core piece of this proposal is that the landing site of V-fronting is a specifier position. In essence, this is equivalent to long head movement. Precursors to this line of research date back to Koopman (1984), but it has been argued for elsewhere (Fukui and Takano 1998; Toyoshima 2001; Matushansky 2006; Vicente 2007, 2009:among others), including in more recent work by Harizanov (2019) and Preminger (2019).

This type of movement had once been considered incongruent with a general condition on syntactic movement, the Chain Uniformity Condition (CUC) — "a chain is uniform with regards to phrase structure status" (Chomsky 1995:253). Together with a distinction between minimal (i.e., a terminal node) and maximal (i.e., a maximal projection) syntactic objects, the CUC rules that links of a movement chain will have the same status, as a head or a phrase. In practice, the CUC ensures that phrases move to specifiers as their landing sites and heads adjoin to a head position.

Several scholars, however, have pointed out that the CUC is, at least, theoretically problematic (Fukui and Takano 1998; Toyoshima 2001; Matushansky 2006; Vicente 2007, 2009). For one, the CUC is redundant: the uniformity condition on the shape of a movement chain can be derived via another independently motivated principle, the Extension Condition (Chomsky 1995). This condition ensures that an XP lands in a specifier position since Merge has to target the root of the syntactic tree (i.e., the position where a syntactic object is not dominated by another syntactic object). There are, therefore, reasons to drop this condition on the shape of movement. Doing so would leave us with a scenario where no restrictions are imposed on the input of Merge. As

a result of that, Merge should apply to minimal and maximal syntactic objects alike. Once these assumptions are in place, head-to-specifier movement is ruled in.

The theoretical possibility of a head-to-specifier movement has prompted scholars to analyze some head-fronting phenomena as such. For example, some scholars, like Vicente (2009) and Harizanov (2019), argued that a head-to-specifier movement is empirically attested in Spanish predicate cleft (81a) and Bulgarian participle fronting (81b), respectively.

(81) a. **Leer**, Juan ha leído un libro.
read.INF Juan has read a book
'As for reading, Juan has read a book.' (Vicente 2009:159)

b. **Pročel** bjah knigata.
read be.1.SG.PST the.book
'I had read the book.' (Harizanov 2019:5)

These scholars argue that, in either case above, a head – a bare verb (81a) or a participle (81b) – fronts and lands in the specifier of a higher functional projection – Spec,TopicP or Spec,AuxP – by virtue of an A' feature with the EPP property for movement.

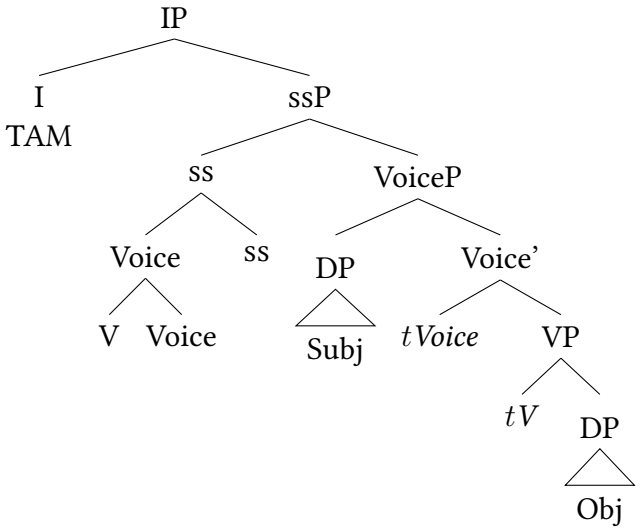
Spanish and Bulgarian represent a growing list of languages where a head-fronting phenomenon has been successfully analyzed as the result of a head-to-specifier phenomenon; others include predicate fronting/cleft in Brazilian Portuguese (Bastos-Gee 2009), Yiddish (Bleaman 2022), Hebrew (Landau 2006; Harbour 1999), Finnish (Brattico 2022); see Harizanov (2019:29) for a larger list of phenomena that are potentially amenable to this approach.

Yet, this approach has not been widely proposed as a way of deriving V1 order as a general principle of clausal structure. To my knowledge, only Mayan (Clemens and Coon 2018), Otomanguean (Lee 2005; Macaulay 2005; Eberhardt 1999) and, under one proposal, the Polynesian language Niuean (Clemens 2019) have received explicit accounts that verb-initial orders result from syntactic head movement.[31]

Similar to Kawahíva, in informationally neutral clauses, Mayan languages show the V1 pattern. However, while some Mayan languages show VSO order in this context, another group of Mayan languages alternate VSO/VOS orders. According to Clemens and Coon (2018), the verb head moves through Voice to a position above the subject position in Spec,Voice, as demonstrated in Tree (82). This straightforwardly captures the VSO languages of the family.

[31]The Nilotic language Kipsigis also shows a pattern of V1 analyzed as the result of head movement of the verb to a position higher than its arguments, as proposed by Bossi and Diercks (2019). However, since the authors explicitly write that "syntactic and postsyntactic analyses of head movement both derive the appropriate word order patterns in the language" (Bossi and Diercks 2019:17 (footnote 17)), I will not consider the V1 pattern in Kipsigis on a par with the pattern of V1 in Mayan nor Kawahíva.

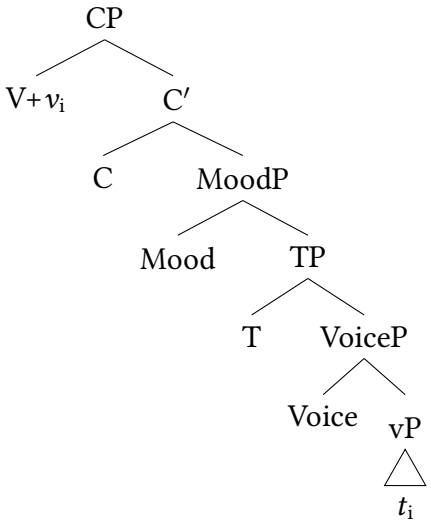
(82)



Additionally, the authors posit that this analysis can also account for the alternating VOS order in VSO/VOS languages by arguing that the surface position of the object in VOS order follows from three other phenomena, which are i) rightward subject topics, ii) heavy-NP shift of phonologically heavy subjects, and iii) prosodic reordering of bare NP objects; see Clemens and Coon (2018) for a detailed discussion about how the effects of these phenomena lead to VOS.

I argue that a similar analysis can derive Kawahíva V1. Unlike the head landing site of the verb in (82), the landing site of Kawahíva V1 is Spec,CP, as illustrated in (83) below.

(83)

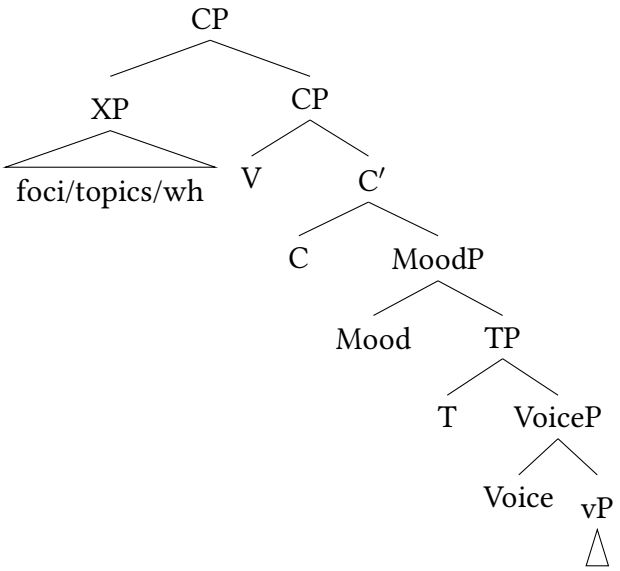


The exact feature triggering the movement of the verb will be discussed in section (4.6.3).

This analysis would immediately explain the complementary distribution between verbs and phrases in the initial position of the clause. Simply put, this distribution is due to a competition

for Spec,CP, as illustrated in the diagram below.[32]

(84)



In the remainder of this section, I discuss other properties of this verb movement that strongly support the idea that it is syntactic.

4.6.1 The hallmarks of syntactic head movement

The discussion about the properties of syntactic head movement in Kawahíva V1 assumes the type of movement proposed in Harizanov and Gribanova (2019). To foreshadow these properties, they are briefly described below:

(85) a. Syntactic head movement leads to interpretive semantic effects (Harizanov and Gribanova 2019).
b. Syntactic head movement does not obey the Head Movement Constraint (HMC) (Travis 1984).
c. Syntactic head movement does not need to directly reflect syntactic structure (Baker (1985)'s Mirror Principle 1985).
d. Syntactic head movement does not create a head that functions as a unit for the purposes of phonology and syntax (Harizanov and Gribanova 2019).

As one can see, and as also discussed by those authors, the properties of syntactic head movement contrast with the properties of its post-syntactic kin, (morphological) amalgamation, which we had discussed in section (4.5).

[32]One might wonder why Kawahíva does not allow fronted verbs and informationally prominent constituents to co-occur in multiple specifiers of CP. I discuss this constraint in §4.7.1.2.

4.6.1.1 Semantic effects

The most telling piece of evidence that verb fronting in Kawahíva is syntactic head movement is the semantic import associated with it. To see this, consider examples with universal quantification, which is expressed via the adverbial suffix *-pam* 'completely,' [33] and its scope relative to a focused bare noun in subject position:

(86) Javatxinga ki o-mõnhã-pam mbarakaja'ia ko.
dog PST 3.A-corral-COMPLETELY cat REAL
'It was a/the dog that corralled all the cats.' (∃ > ∀)
'#It was all the dogs that corralled a/the cat.' (#∀ > ∃) (Juma: Elicit)

In this SVO clause wherein the clause-initial subject bears (contrastive) focus and the clause-medial verb is suffixed with *-pam* 'completely,' the latter cannot scope over the bare noun in subject position; the resulting reading is infelicitous. Instead, the only element over which *-pam* 'completely' can have scope is the bare noun in object position.

However, if the verb marked with *-pam* 'completely' is in clause-initial position, thus resulting in a VSO clause, then the reading where universal quantification scopes above the bare noun in subject position becomes felicitous. Interestingly, the bare noun in subject position cannot outscope *-pam* 'completely.'

(87) O-mõnhã-pam ki javatxinga mbarakaja'ia ko.
3.A-corral-COMPLETELY PST dog cat REAL
'All the dogs corralled the cat.' (∀ > ∃)
'#A/the dog corralled all the cats.' (#∃ > ∀) (Juma: Elicit)

In the analysis developed in this chapter, the clause-initial position of a pragmatically prominent constituent, like the focused phrase in (86), as well as the clause-initial position of the verb in (87), is achieved through syntactic movement to Spec,CP. The distinct scope readings resulting from the varied positions of the verb suffixed with *-pam* 'completely' relative to the subject nominal provide evidence for the syntactic nature of the verb's movement. Given that only movement within narrow syntax, therefore movement prior to LF can yield these semantic distinctions, it follows that the movement of Kawahíva verbs to the clause-initial position in VSO clauses must be syntactic in nature.

The result of this section – verb head movement in Kawahíva is syntactic because it leads to semantic effects – also accords with the results in a growing body of literature on syntactic

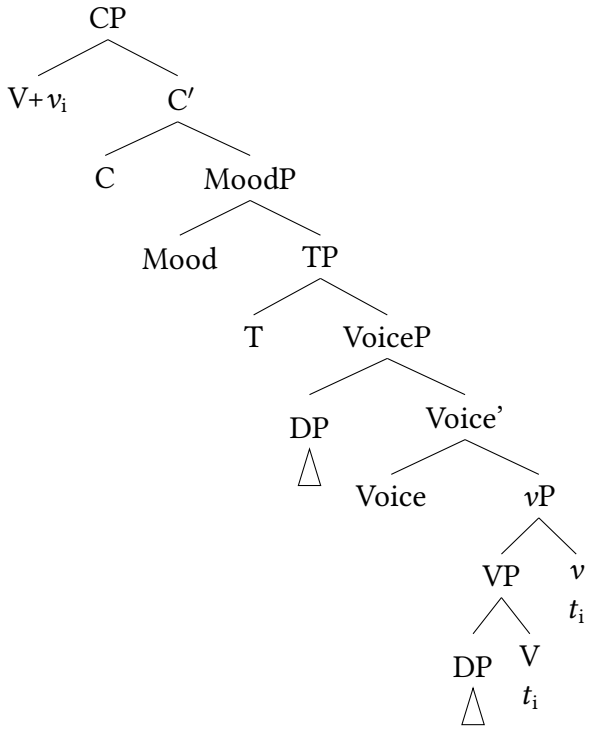
[33] Several adverbial meanings, including (some types of) aspect, evaluative mood, speech act mood, among others, are realized through adverbial suffixes (e.g., *-pam* 'completely,' *-aha* 'unfortunately,' *-ete* 'really'). Furthermore, a cursory exploration of their ordering on the verb complex reveals that almost all of them have a fixed order, with a few being able to occupy different positions in the verbal stem complex. Negation is also realized through bound morphemes, specifically the circumfix *n(da)-…-i*. I assume these suffixes can be optionally merged onto the verb, which accounts for the fact that they are not required, and their almost fixed ordering is the result of certain combinations being ruled out because they produce semantically ill-formed structures, following the Scopal Theory approach to the syntax of adverbials (Ernst 2020).

head movement that has shown similar semantic effects of movement of quantificational heads (Lechner 2006; Lee Ms.).

4.6.1.2 Non-locality

We expect syntactic head movement to be non-local in the sense of being able to skip heads and, therefore, disobey the Head Movement Constraint (HMC) (Travis 1984). Consequently, we would expect the fronted verb to skip an intervening head. Tree (88) demonstrates the conclusion from section (4.5) that the moving head skips T and Mood.

(88)



The ability to skip T and Mood shows that Kawahíva V1 fronting is non-local and it does not abide by the Head Movement Constraint (HMC) (Travis 1984), another trait of this type of movement (Harizanov and Gribanova 2019).[34]

[34]The non-local aspect of verb movement in Kawahíva brings it close to the so-called Long Head Movement (LHM) discussed in several works by Lema and Rivero (1990:among many others publications by Rivero) based mostly on Slavic languages. However, the cases of LHM discussed in their work seem to differ from the present LHM in Kawahíva in at least two aspects. One is the landing site: as opposed to landing in a specifier position, the moving head, in their case, is a head position. The other is nonlocality: Kawahíva LHM has some intermediate steps of local movement, such as V-to-v, while LHM in Slavic does not seem so. Curiously, in a proposed typology of V(P)-movement, Travis and Massam (2021) identify two types of LHM, one of which they exemplify with the LHM of Slavic languages. The other one is exemplified by pronominal cliticization in Romance languages. I leave a full comparison and discussion of Kawahíva LHM and the LHMs identified by Travis and Massam (2021) for future work.

4.6.1.3 Absence of mirror effects

A further consequence of the syntactic movement of the verb in (88) is that it will not pick up the content on T and Mood, thus in an apparent violation of Baker (1985)'s Mirror Principle and rendering anti-mirror effects. This is so because, for instance, given the representation in the diagram in (88), and assuming an instance where only Mood and T are overtly exponed, we would see the morpheme order V-*v*-Voice followed by Mood and T. As suggested, these anti-mirror effects are only apparent. This resulting morpheme order is simply a consequence of the complex V-*v* not moving through these heads since the moving head skips them.

4.6.1.4 Absence of properties of a phonological/syntactic head

Finally, the resulting head of a head-to-specifier movement does need to form a syntactic unit with the head of the surrounding projection. Given the tree in (88), the C head would not have been part of V-*v* at Spellout in the post-syntactic component. This fact would predict, for instance, that the V-*v* and C heads may not form a prosodic unit because each of them, for instance, bears their own stress. This prediction is also fulfilled in Kawahíva, as it has been shown in section (4.5) that the V-*v* does not form a prosodic constituent with the C head because they each are stressed.

4.6.2 Kawahíva V fronting is not A'-movement

It remains to be discussed the identity of the featural trigger of V fronting, given that movement is feature-driven (Chomsky 1995, 2001). I first consider the option where this feature is an A'-feature in this section. I demonstrate this alternative overgenerates. Then, in the following section, I propose the relevant feature is the lexical feature [V]. I show this option does not suffer from the overgeneration issue raised in the A'-feature competing alternative and can also be extended to cases where the clause-initial element is not a verb.

Analyses of predicate fronting where the verb lands in a specifier position have proposed the trigger is an A' feature; this is true for predicate fronting/cleft in Brazilian Portuguese (Bastos-Gee 2009), Spanish (Vicente 2007, 2009), Yiddish (Bleaman 2022), Hebrew (Landau 2006; Harbour 1999), Finnish (Brattico (2022), as well as participle fronting in Bulgarian (Harizanov 2019), among others. The major reason for an A'-feature as the movement-inducing feature has been that some of its properties can be associated with the properties of A'-movement, which are laid out in the following table for convenience based on Richards (2014)'s overview.

In particular, this A'-movement of a head is non-local in the sense of non-local A'-movement, wherein the moving element can cross finite clausal boundaries in long-distance extraction. Additionally, it can be associated with pragmatic effects such as focus, as in Brazilian Portuguese predicate cleft (Bastos-Gee 2009), or topic, as in the Spanish infinitive predicate topicalization (Vicente 2007, 2009). I discuss the locality of this movement shortly.

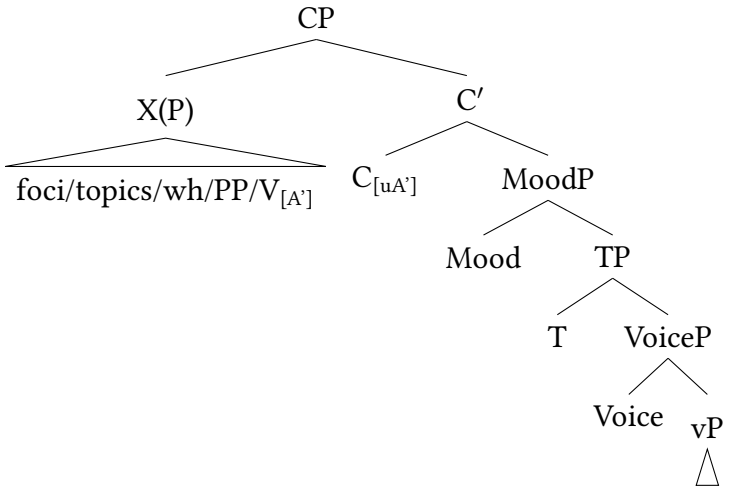
As for the association of this movement with pragmatic effects, this idea might initially receive some support from the fact that verbs compete for the clause-initial position with A'-feature-driven elements (e.g., topics, foci, and questions). Furthermore, this view would offer a unified

PROPERTIES	A-PROPERTIES	A'-PROPERTIES
Local	✓	✗
Restricted to nominals	✓	✗
No reconstruction for Condition C	✗	✓
No Weak Crossover	✓	✗
New antecedents for anaphors	✓	✗
Parasitic gap licensing	✗	✓

Table 4.10: Summary of the A/A' distinctive properties.

account of V-fronting and A'-movement in the matrix clause, in that movement to Spec,CP would be triggered by a generalized A'-feature on C – we can call this feature as [A']. (89) demonstrates this alternative in a diagram:

(89)



However, the parallel between Kawahíva V fronting and instances of A'-movement of a predicate starts to break down when we consider locality, in particular, long-distance extraction – while long-distance movement is available and required of embedded XPs, it is not available for the embedded verb.

Long-distance movement of constituents out of (nominalized) dependent clauses is allowed (see section 4.2.3). Contrast the baseline example without extraction in (90a) to that with extraction in (90b).

(90) a. Ere-piang nde [akwemba'ea=ga pira pyhyk=a].
2.SG.A-see 2.SG man=3.SG.MASC fish catch=NMLZ
'You saw the man catching fish.' (Juma: Elicit)

b. Gara nde ere-piang [akwemba'ea=ga i-pyhyk=a]?
what 2.SG 2.SG.A-see man=3.SG.MASC i-catch=NMLZ
'What did you see the man catching?' (Juma: Elicit)

I assume the driving force for the object extraction in the (b) example is an A'-feature on the matrix C head. It would only be natural to assume that this feature also drives short phrasal A'-movement within a matrix clause. Consequently, if it is also involved in verb fronting in matrix clauses, as one would propose under the account of V1 as A'-movement, the long-distance extraction over an embedded verb should also be possible. This prediction is not true, however.

Long-distance fronting of the dependent verb to the matrix clause is not allowed, as the example in (91) demonstrates. In particular, in B's response in (91b), the dependent verb is contrastively focused, and fronted to the matrix clause. The result is unacceptable. Therefore, the A'-driven movement view overgenerates.

(91) Context: *You were walking in the forest and saw your father kneeling next to a tapir lying on the ground who was squealing. You came back home and told me, your brother, that father was helping a tapir. I know that father was killing the tapir because he told me when he came back. My brother (person A) says, "I saw Father healing a tapir," to which I (person B) reply, "Nope! You saw Father killing a tapir".*

a. A's comment

A-hepiang jie [apina=ga tapi'ira mõmberav=a].
1.SG.A-see 1.SG father=3.SG.MASC tapir heal=NMLZ

'I saw Father healing a tapir.' (Juma: Elicit)

b. B's response with long-distance verb fronting.

Ahan! *Juka=a nde ere-piang [apina=ga tapi'ira].
Nope! kill=NMLZ 2.SG 2.SG.A-see father=3.SG.MASC tapir

'Nope! You saw Father KILLING a tapir.' (Juma: Elicit)

The repair strategy used by the speaker to convey corrective focus over the embedded verb in this scenario is demonstrated in example (92). This repair employs two independent clauses. The first negates the healing event, and the second uses an adverbial purposive clause (formally a nominalized verb with the corrected verb *juka* 'kill') in clause-initial position to assert the correct event seen by the interlocutor according to the speaker.

(92) A'ea=hãi nde repiak-i. I-juka=a te nde=repiak-i.
this.one=NEG 2.SG see-i i-kill=NMLZ REALLY 2.SG=see-i

'It was not this one you saw. To kill it, you did.' (Juma: Elicit)

An argument from the formal typological literature related to verb fronting (be it head fronting or verbal phrase fronting) can also be adduced, which is consistent with Kawahíva verb fronting not involving an A'-feature. It has been shown that languages with verb head fronting or verb phrase fronting derived via A'-movement show two gap avoidance strategies: i) a copy of the displaced head or ii) a semantically vacuous dummy verb (Hein 2017, 2018). None of these effects are found in Kawahíva verb fronting. Consequently, this accords with the idea that A'-movement of the verb is not involved in the creation of V1 clauses in the language.

In the next section, I propose that the driving force of V fronting is a lexical [V] feature. Importantly, the issue of overgeneration does not arise in the analysis of V1 clauses as driven by this feature.

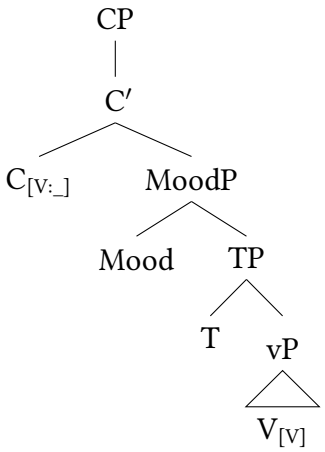
4.6.3 The feature trigger of verb fronting: a [V] feature

I propose the feature [V] is responsible for verb fronting in Kawahíva; this idea follows a similar identity for the feature responsible for predicate fronting in the Polynesian languages Niuean (Massam 2020) and Imere (Van Urk 2022). This feature is hosted on the C head of regular declarative matrix clauses and comes with the EPP property, as defined in (93).

(93) **The [V] feature on C**:
A [V]-bearing element must check off [V] on C by moving to Spec,CP.

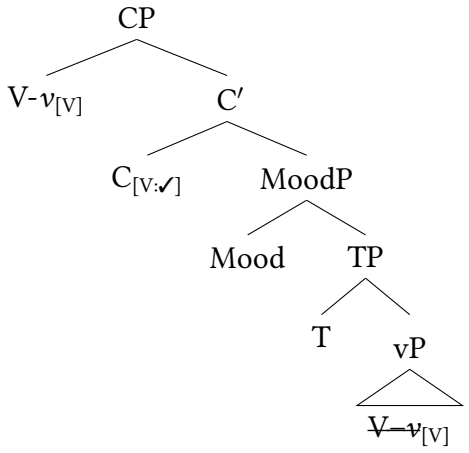
I also posit that *v* comes with the [V] feature that can satisfy the EPP of [V] on C, and so will the verb head once it moves onto *v*. The relevant structure pre-movement of the verb is demonstrated in (94).

(94)



Once the C head endowed with the EPP-bearing [V] feature is merged, it will trigger the movement of a [V]-bearing element to its specifier. In Kawahíva, this is the complex head V-*v*, as shown in (95). In section (4.7.3), I flesh out this analysis to account for a separate issue that it raises, namely, that the complex head carrying the [V] feature is targeted for movement, but not the more local target phrase *v*P that also carries [V]. For now, I ask the reader to bear with the idea that the target for movement to Spec,CP is V-*v*.

(95)



Additionally, I posit that this complex head M-merger with C (Matushansky 2006) in the PF component. This accounts for the realization of [Φ]-agreement morphology hosted on C onto the verb.

The proposal with [V] avoids the main issue raised by the A'-analysis of verb fronting. In that analysis, there was no reason why phrases within a dependent clause, but not an embedded verb, could be extracted to the matrix clause. To see this again, consider the next examples.

In (96), the dependent verb is *in situ*, and the matrix verb is initial.

(96) A-hepiang jie [apina=ga tapi'ira mõmberav=a].
1.SG.A-see 1.SG father=3.SG.MASC tapir heal=NMLZ
'I saw Father healing a tapir.' (Juma: Elicit)

In (97), however, the dependent verb appears initially. This configuration is unacceptable, as previously known from the discussion of the A'-analysis.

(97) *Ahan! juka=a nde ere-piang [apina=ga tapi'ira].
Nope! kill=NMLZ 2.SG 2.SG.A-see father=3.SG.MASC tapir
'Nope! You saw Father KILLING a tapir.' (Juma: Elicit)

In the present account with the [V] feature, however, this result is expected: verbs in dependent clauses do not front to the matrix clause because there will always be a closer match for [V] to move (i.e., the matrix verb) than moving the dependent verb. Moving the dependent verb, as in (97), incurs a violation of a property of syntactic movement, the Relativized Minimality, by which movement targets the closest goal to the probe in terms of c-command (Rizzi 1900; Chomsky 1995).

The analysis with a [V] endowed with the EPP property might receive further support from the fact that Kawahíva shares properties with other languages that have been claimed to have a predicate fronting through [pred] endowed with the EPP property. In a recent typology of EPP features, Doner (2017, 2019) proposes there are two major types of EPP, predicate- and nominal-sensitive EPPs. The first type is found in Celtic (except Breton) (Biberauer 2010), Inuktitut (Doner

2017), Niuean (Massam and Smallwood 1997), a.o., while the nominal-sensitive EPP is found in English, for instance. Among the properties of languages with a predicate EPP, which are found in Kawahíva as well, are i) the raised constituent is not nominal, ii) non-verbal predicates also front (but non-functional verbal elements, i.e., auxiliaries, light verbs, and copulas, do not front) and iii) the existence of a defective D.[35] Although the current account does not propose that the feature trigger is [pred], I subsume the feature [V] to be a subtype of a [pred] feature. If this reasoning is accepted, then the Kawahíva properties discussed in the next paragraph are expected, given the properties found in predicate-sensitive EPP languages.

Properties i) and ii) are illustrated in examples (98) and (99), respectively.[36] (99) also serves to show the lack of a systematic contrast of definiteness, which would support the existence of a defective D head in the language since there is no definiteness marking on the object *mbiara* 'meat'; consequently, the object can receive either a definite or indefinite reading.

(98) Tapy'ynh=a jie.
non.indigenous=NMLZ 1.SG
'I am a non-indigenous (person).' (Juma: Elicit)

(99) a. O-'u ve-'e auhu ki hẽa mbiara.
3.A-eat to-DIM seem PST 3.SG.FEM meat

b. *Auhu ki hẽa o-'u (ve-'e) mbiara.
seem PST 3.SG.FEM 3.A-eat to-DIM meat
'It seems that she ate a/the meat.' (Juma: Elicit)

In short, the proposal that a [V] on C is involved in the creation of Kawahíva V1 clauses not only solves the main issue raised in the alternative analysis with an A'-feature but also fits in a typology of languages that have received an account with a feature which could be regarded as similar to [V].[37]

The remainder of the chapter demonstrates how this proposal extends to non-verbal initial clauses and addresses a connected debate about X (head) versus XP (phrasal) movement given the assumption within Bare Phrase Structure that both contain the same label and the latter should be a closer target for movement.

[35] In Doner's words (2019: 183), "a defective D […] (corresponds to) the lack of a systematic definiteness contrast."

[36] Verbs like *auhu* 'seem' are referred to here 'light' verbs in the sense that they are functional elements; another verb that falls under this category is *rãmbuve* 'begin.' They do not behave as prototypical verbs in Kawahíva, given that they do not show some core verbal properties, including the ability to bear agreement inflection. I conclude these 'light' verbs cannot undergo fronting because of their condition as functional elements, which preclude them from being combined with the [V] feature.

[37] Kawahíva also accords to Alexiadou and Anagnostopoulou (1998)'s typology of movement to check an EPP feature, whereby languages that move a head to check EPP do so by V-raising.

4.7 Extension and consequences

4.7.1 Non-verb-initial clauses

We have seen in section (4.2.2) cases where the verb is not the first element in the clause. In matrix clauses, in particular, they can also be medial and final depending on whether the initial element is an extracted PP or DP. These word order changes are also accompanied by different morphological changes as summarized in Table (4.11).

CLAUSE-INITIAL X(P)	verb position	agreement	*-i* suffixation
Verb	–	sets A & B	✗
Extracted DPs	medial	sets A *&* B	✗
Extracted PPs	final	set B	✓

Table 4.11: Summary of clause-initial elements and morphosyntactic properties.

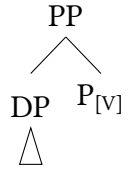
Descriptively, when the clause-initial position is occupied by an extracted PP, the main verb i) is strictly in the final position, ii) is voided of subject marking, and iii) receives a special suffix *-i*.

In what follows, I demonstrate that the [V] feature with the EPP on C also explains the differences in Table (4.11). I argue these correlate with the ability of the initial element other than a verb to check off [V] on C – if it can, then it bleeds the movement of the verb, and if it cannot, then the movement of the verb should still be possible. I discuss clause-initial PP sentences to illustrate the first scenario, where verbs do not undergo movement at all. Then, I discuss (non-predicative) clause-initial DP sentences to illustrate the second scenario, wherein DPs are not able to check the [V] on C, and verbs still move to Spec,CP; however, verbs are not pronounced in Spec,CP due to a language-internal PF constraint to avoid spellout of multiple elements in Spec,CP. Consequently, a linearization algorithm results in the spell out of the verb in an intermediate position.

4.7.1.1 PP-initial clauses

Starting with the instances where extracted PPs fill in the clause-initial, I propose PPs can check the [V] feature on C. Consequently, clause-initial PPs can bleed the movement of the verb as the former removes the feature trigger for verb fronting. I propose they do so because PP heads can merge with the [V] feature as well. In essence, this means that postpositions and verbs form a natural class.

(100)



The idea that postpositions and verbs form a natural class is not surprising when considering that postpositions share several morphosyntactic properties with verbs in the language. In particular, postpositions are able to host agreement markers (101), voice morphology (102), aspect morphology (103), and the *i-* prefix marking (104).[38]

(101) O-mõpu ga **ji**=rehe.
3.A-shoot 3.SG.MASC 1.SG.B=at
'He shot at me.'

(102) Ere-mõpu nde e-**ji**=ehe.
2.SG.A-shoot 2.SG 2.SG.COR-REFL=at
'You shot at yourself.'

(103) Oji ve-**'i** ki Takãi Bitate=ga rur-i avo ko.
some.time.ago to-DIM PST Takãi Bitate=3.SG.MASC come-i LOC.DEM:PROX REAL
'Just in the other day Takãi Bitate came here.' (Juma: Text)

(104) a. Nd-a-jor-i **j**-uvi.
NEG-1.SG.A-come-NEG i-from
'I didn't come from (there).'

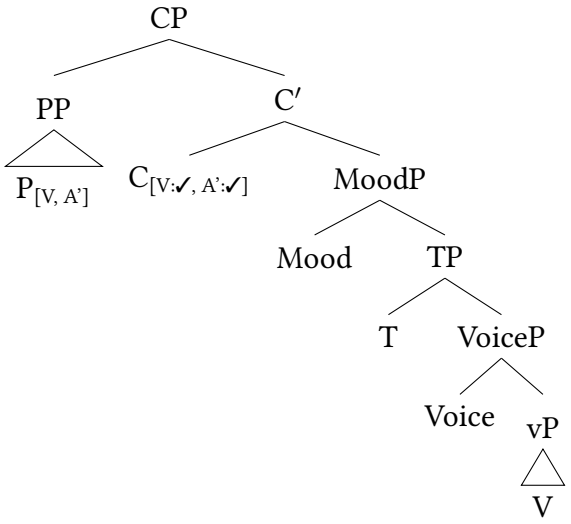
b. Nd-o-ko-i i-pupe.
NEG-3.A-be-NEG i-inside
'It's not inside (it).'

I posit the verbal behavior of postpositions is straightforwardly explained under the idea that the postposition categorizing functional head also contains a [V] feature and that the affixes above all are subcategorized to only occur on [V]-bearing heads.

In addition, PPs also bear an A'-feature when they are extracted. Given that C heads can optionally bear the A'-feature, extracted PPs will be able to satisfy both features on C at once. Consequently, when PPs are initial, they remove the trigger for verb fronting, as demonstrated below:

[38]Postpositions as predicates are attested in non-related languages (Hagège 2010:245-249). In the Amazon, an analogous case is known from the Carib language family (Meira 2006).

(105)



As a result of that, the verb stays in its base position within the VP — which equates with the sentence-final position on the surface — and it cannot expone subject agreement morphology – which is the spellout of a probe on C (see section (4.3.3) for how verbal agreement is derived).

A final aspect of PP-initial clauses is the pattern of *-i* suffixation. This suffix is only and always triggered in the context of clause-initial PPs in matrix clauses.[39] I suggest this verbal morphology realizes the [V] feature on *v* in cases where the verb is trapped within the *v*P.[40] More technically, I propose the following rules for Vocabulary Insertion.

[V, *v*]/	⟷	-i

Table 4.12: Vocabulary items for V+*v*.

[39]The adverbial PP is optionally dropped in natural speech. Therefore, examples in this chapter where the the suffix *-i* is observed without an overt PP should be interpreted as having a covert PP clause-initially.

[40]The pattern of *-i* suffixation and its structural condition is also attested in other members of the Tupí-Guaraní family (for a language family overview, see Jensen (1999:156)). However, I am not aware of any analyses for this pattern in any language within the Tupí-Guaraní family or related language families. Some labels used by scholars to refer to *-i*, such as *oblique-topicalization/oblique-topicalized verb* (Jensen 1999; Vieira 2014), suggest an approach wherein the suffix *-i* is a morphological reflex of oblique extraction (i.e., *wh*-agreement morphology). However, it seems unlikely that there would be distinctive agreement morphology to cross-reference oblique topics but not argument topics. This is significant in Kawahíva, which exhibits *wh*-agreement morphology for Case in relative clauses, and this paradigm includes agreement markers for both arguments and obliques, none of which are *-i* (Dos Santos In prep.).

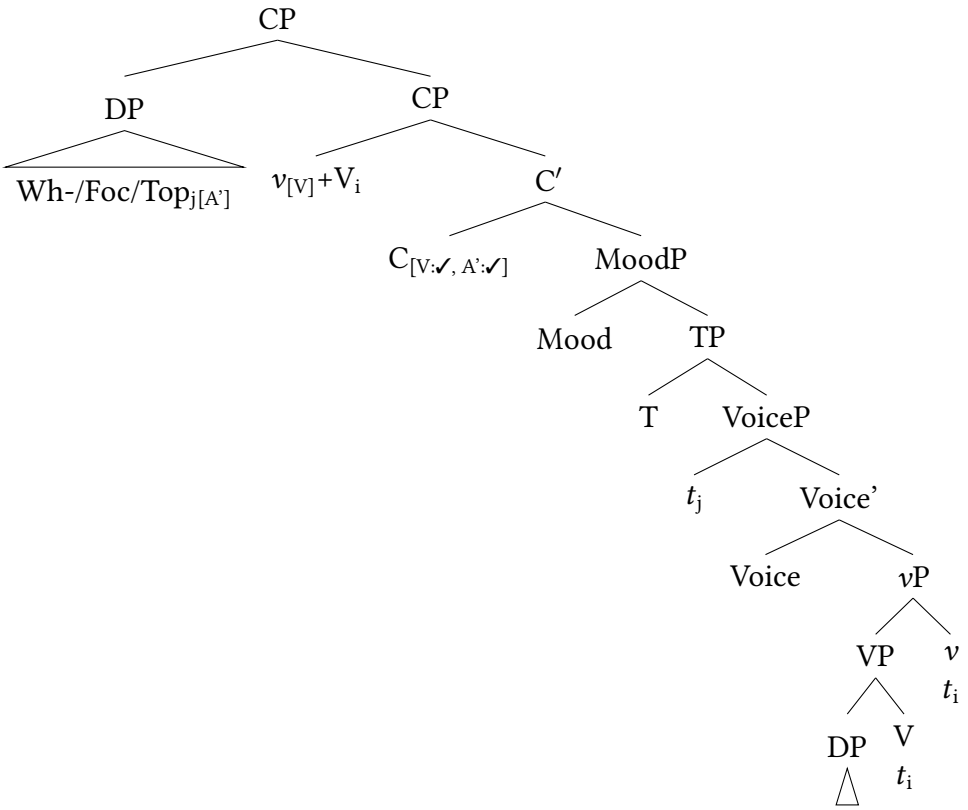
Additionally, there has been an attempt to reanalyze the structures with *-i* as a nominalized clause where *-i* serves as the nominalizer (Praça et al. 2017). The supporting argument for this proposal is the absence of subject agreement markers in these structures, just like other dependent clauses, consistently analyzed as nominalized in the literature of this family. However, the absence of subject agreement in structures with *-i* would be the only shared property with truly nominalized clauses. This parallel breaks down when other morphological properties are considered, such as the exponent of negation and availability of coreferential markers – the structures with *-i* suffixation use the same negative exponent as matrix clauses, which is unlike dependent clauses. Furthermore, coreferential markers are not available in structures with *-i* suffixation, the opposite pattern found with truly nominalized clauses.

This idea should capture the fact that this suffix is only employed in cases where the verb does not leave the *v*P, and therefore, the verb will move as far as to *v*.

4.7.1.2 DP-initial clauses

In contrast to PP-initial clauses, extracted DPs cannot check off the [V] on C. In these cases, I propose that despite its inability to occur in an initial position, the verb undergoes movement to Spec,CP to check off [V] on C, as demonstrated in (106). This claim is supported by the fact that the verb can bear subject agreement morphology (107), which is located on C (as discussed in §4.3.3).

(106)



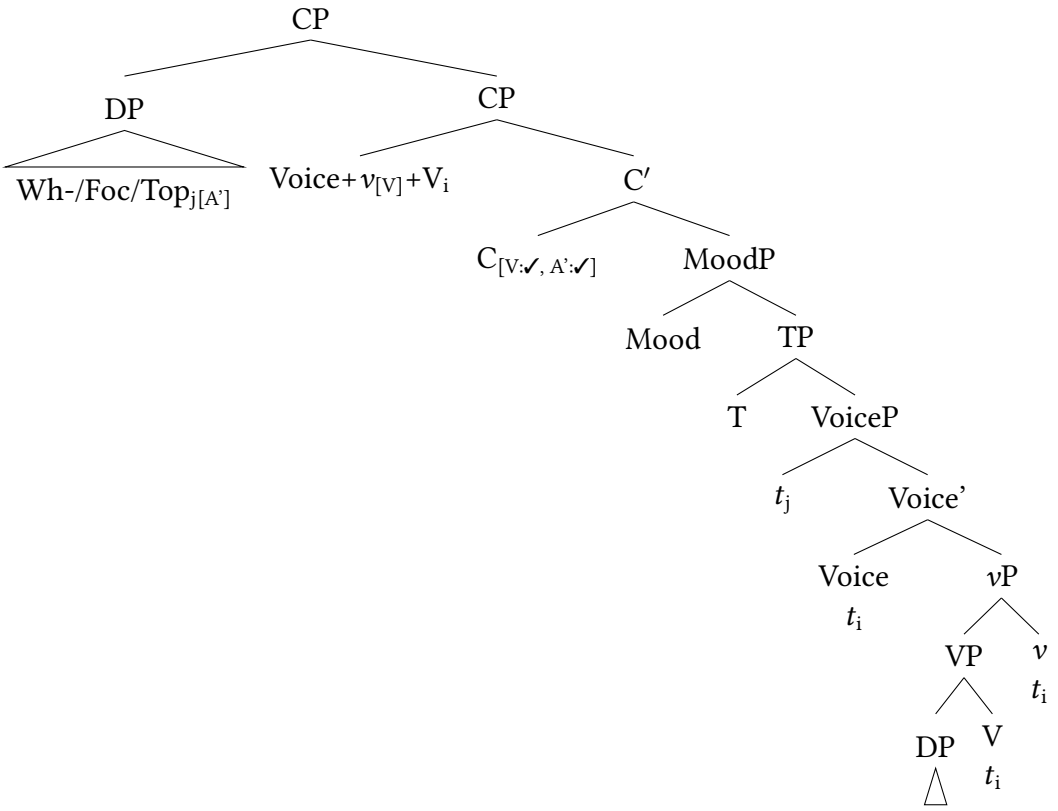
Notably, however, the verb is not realized before T and Mood, contrary to what the representation in (106) would presuppose. This has already been shown with a DP subject question in (28) and is also seen with a DP object focus in (107). In these cases, the verb surfaces below the subject, which is in Spec,Voice according to the phrase structure proposed for Kawahíva.

(107) a. Pirapetxinguhua **ki** hẽa o-pyhy-pyhy ko.
pirapetxinguhua.fish PST 3.SG.FEM 3.A-RED-catch REAL
'It was *pirapetxinguhua* fish that she caught.' (Juma: Text)

b. Mãngã **po** ji=ruvyra v-epiang rimba'e?
who IRR 1.SG=uncle 3.A-see NWIT.REM.PST
'Who did my uncle see back in the day?' (Juma: Text)

To account for the surface position of the verb in clause-initial (non-predicative) DP sentences, I posit that what surfaces is an intermediate copy of the verb. Given that this intermediate copy spells out below the reality particles (=MoodP), tense particles (=T), and subject (=Spec,VoiceP), but above the object (=*v*P/VP), I propose this is the Voice head, as illustrated below:

(108)



Furthermore, I posit that there are two lexical Voice heads. One is selected for cases wherein the verb is spelled out in an intermediate position and it has the ability to trigger head movement of the head *v*+V onto it. In contrast, a different Voice head will be merged for those cases where the verb is spelled out lower, following the object (i.e., PP-initial clauses). This head does not attract the head *v*+V.

This idea captures the fact that, in DP-initial clauses, the verb undergoes movement through Voice on its way up to Spec,CP, while in PP-initial clauses, it does not, which is evidenced by its strict final position. Additionally, a PF constraint, to be elaborated shortly, precludes spellout of

both the verb and extracted DPs in Spec,CP. Consequently, the intermediate copy of the verb on Voice is pronounced at PF.[41]

With these ideas in mind, I propose extracted DPs and verbs cannot co-occur in Spec,CP due to a language-internal PF constraint that forbids spellout of multiple specifiers in Spec,CP, as defined below:[42]

(109) **The Kawahíva Doubly-Filled Spec,CP PF Constraint**: Spec,CP cannot spell out more than one specifier.

The PF component then must spell out a lower copy of the verb. I propose this is just the highest copy of the complex head V+*v* before reaching Spec,CP, that is, the copy on Voice.[43,44]

A question that arises concerns the choice to realize the verb in a lower position: why is it the verb, and not the extracted DP, that must be realized lower? I suggest the alternative choice where the extracted DP surfaces lower is ungrammatical because it violates the linearization ordering established between the extracted DP and the verb in earlier spell-out domains. This idea builds on Fox and Pesetsky (2005)'s linearization algorithm for the syntax-phonology mapping. This algorithm prohibits contradictions between the orderings of words fixed in an earlier spell-out domain. For our purpose here, it will be assumed that the spell-out domains in Kawahíva correspond to the complement of the phasal heads C and *v*.

Consider the different orderings established in scenarios of co-occurring DP extraction and verb fronting, starting with subject extraction, which is illustrated in (110). The expression "A < B" should be read as "A precedes B":

(110) **Spell-out applies to CP**
$[_{CP}\ DP_{subj} < V < Mood < T < \cancel{DP}_{subj} < \cancel{V}\ [_{vP} < DP_{obj} < \cancel{V}\ ]\ ]$

In subject extraction, the relevant spell-out domain is the CP; the relative ordering within the VP is non-relevant because the subject argument is not merged at the point when Spellout applies to the VP.

[41] One might raise a concern that this account amounts to a conspiracy – the verb moves to Voice only when it might end up being spelled out on this head. However, this issue is not exclusive to Kawahíva. Harizanov and Gribanova (2019:500) also noted a reviewer's reaction to their alternative approach to the account of German V2 that maintains V-to-T. As the authors pointed out, the traditional account of German V2 also leads to a conspiracy, wherein *v* attracts V, and then, T attracts *v*. The upshot is that even for well-studied phenomena like German V2, the mainstream account exhibits a flavor of conspiracy.

[42] Additional evidence that Spec,CP does not allow multiple specifiers comes from the existence of multiple wh-words in a sentence, as in "who saw what?", in which the subject, but not the object, always fronts.

[43] A similar pattern of spelling out a lower copy of a fronted verb just in case the original landing site has been taken is found in Classical Hebrew predicate clefts (Harbour 1999). Harbour shows that verbs move to T in non-predicate cleft constructions. In predicate cleft constructions, the verb moves to C. However, if C is taken by a head that must occur in this position – such as a frame-setting adverbial like 'in the beginning' – then the verb cannot raise to C anymore and spells out lower.

[44] The creation of copies as the verb moves up is also consistent with the current hypothesis that Kawahíva V1 results from syntactic movement, as syntactic movement, but not morphological amalgamation, can create copies (Harizanov and Gribanova 2019:487).

Within the CP are the DP subject and verb, in this order – I assume the verb tucks in beneath the verb (Richards 2014), thus keeping the relative order between the two prior to movement to Spec,CP unchanged. However, the PF constraint in (109) rules out spelling out of both of the initial constituents in this algorithm. Spelling out V lower rather than the subject DP allows the relative ordering of the subject DP and V to be maintained, so only the initial V is deleted.

Consider now object extraction. In these cases, both CP and VP count for the evaluation of relative ordering by the linearization algorithm. First of all, the relative ordering of the object relative to the verb is such that the object precedes the verb in the VP, as demonstrated below.

(111) **Spell-out applies to VP**
$DP_{obj} < V$

Subsequently, the object is moved to Spec,CP. The verb moves to Voice and then to Spec,CP. The final order between the two at the moment when Spell-out applies to CP is represented in (112).

(112) **Spell-out applies to CP**
$[_{CP}\ DP_{obj} < V < Mood < T < DP_{subj} < \sout{V}\ [_{vP} < \sout{DP}_{obj} < \sout{V}\ ]\ ]$

As it can be seen, realizing the verb in Spec,CP instead of the DP object would contradict the relative order established when Spell-out applied to VP. Spelling out a lower copy of V instead of the initial copy maintains the relative order between the extracted object and verb established within the VP.

In sum, choosing to realize the extracted DP rather than the verb in Spec,CP is a consequence of the mapping from syntax to phonology, in particular, an effect of the order preservation monotonic property of Spell-out proposed by Fox and Pesetsky (2005).

4.7.2 An alternative to the complementary distribution between phrases and verbs

In footnote (11), an alternative approach to the competition between informationally prominent phrases and verbs for the clause-initial position was raised. In it, this distribution is due to different C-heads hosting different feature triggers.

This alternative would also have to take into account the morphosyntactic differences that arise depending on the syntactic category of the initial element when it is not the verb. Specifically, we would have a C-head that is endowed with the feature that leads to extraction of DPs, one for PPs, and yet a different one for verb fronting. Then, it could still be maintained that phrases land in the usual position for phrases, namely a specifier. In Kawahíva, this would be Spec,CP. In contrast, the verb lands in the usual position for a head. This position would be the C head in cases of verb fronting. Under this view, the complementary distribution between these elements, therefore, boils down into there being a different trigger on the different C-heads for each type of fronting, rather than a competition for the same position.

The discussion in the preceding section has shown that a single explanation is possible if we posit that the feature triggers can actually co-exist on C. Positing that they co-occur on C when extraction is involved in matrix clauses straightforwardly accounts for the morphosyntactic differences observed on the verb depending on the syntactic category of the element in clause-initial position (i.e., DPs vs. PPs). In the competing account, however, we would miss the generalization that extracted PPs, but not DPs, remove the ability of the verb to undergo fronting.

4.7.3 The XP vs. X movement debate

The proposed analysis of verb movement for Kawahíva V1 circumvents an issue that arises in other case studies of verb fronting wherein the head X of XP fronts despite the fact that XP is the maximal projection of X and should be a closer target for fronting. This condition of a closer target follows from the standard assumption in Bare Phrase Structure (BPS) that a phrase and its head share the same label. If V-*v* and the *v*P that it is the head of bear the same features, they both should be eligible goals for fronting triggered by [V] on C. That is, if *v* has the [V] feature, so does *v*P. However, if that is the case, and without introducing additional assumptions, the system would appear to predict — contrary to fact — that *v*P should undergo movement rather than V-*v* because *v*P is closer to the probe than V-*v*.

This issue does not arise in the creation of Kawahíva V1, according to the present proposal. This is so because once V-*v* moves to Voice, they become a closer target for movement to Spec,CP than *v*P. Additionally, in instances where V-*v* does not front, the trigger for this fronting has been removed by a clause-initial PP. Consequently, *v*P cannot be eligible for fronting either.

4.8 Conclusion

I hope to have presented a compelling case that the creation of verb-initial (V1) clauses in informationally neutral matrix clauses in Kawahíva (Tupí-Guaraní; Brazil) involves (long) syntactic head movement of the verb to a position above its arguments. Arguments for this idea came from the properties associated with syntactic movement, including semantic effects, nonlocality, and a specifier landing site.

Competing alternatives, such as the Remnant VP movement and Morphological amalgamation, were considered but ultimately set aside due to their core predictions not being realized. Notably, these include the movement of the object prior to the movement of the remnant VP under the Remnant approach and a strictly local movement under the morphological amalgamation account. However, positive evidence that Kawahíva V1 does not involve object movement prior to verb movement strongly argues against the first alternative. In addition, the nonlocal property of the movement that leads to V1 discourages the amalgamation analysis.

This analysis contributes to the growing list of languages where syntactic head movement is a general mechanism for deriving verb-initial order. Such languages include Mayan (Clemens and Coon 2018), Otomanguean (Lee 2005; Macaulay 2005; Eberhardt 1999) and, under one proposal, the Polynesian language Niuean (Clemens 2019).

Chapter 5

Nominalization and relativization

5.1 Introduction

Several languages employ a strategy for relativization that consists of using clausal nominalization as an adnominal modifier. It follows from this that relativization in these languages is achieved by nominalization morphology itself, without any additional syntactic processes taking place (Comrie and Thompson 1985; Keenan 1985; Andrews 2007; Shibatani 2009). This strategy, which I shall call the 'nominalization strategy' after Thornes (2023:341), is schematized in (1). Throughout this chapter, I enclose the adnominal nominalization within brackets and present the modified noun in bold.

(1) *The nominalization strategy for relativization*:
Noun [… nominalization …]

This pattern of relativization is highly pervasive in Asia and the Americas (Yap and Grunow-Hårsta 2010; Yap et al. 2011; Comrie and Estrada-Fernández 2012; Zariquiey et al. 2019). We see it in (2) from the Tibeto-Burman language Chantyal, a language family in which relative clauses have been almost exclusively studied under the umbrella of nominalization (Matisoff 1972; DeLancey 2002; Noonan 2008; Genetti 2011). In this example, the modified noun *gayye sya* 'beef' is preceded by the adnominal nominalization *mənchisə casiwa* 'that the person ate'. In languages of the Tibeto-Burman family, it is maintained that relative clauses are "best viewed as NPs, juxtaposed to the NP they are modifying, the two NPs constituting, therefore, a sort of appositional structure" (Noonan 2008:226).

(2) [Mənchi-sə ca-si-wa] **gay-ye sya**
person-ERG eat-ANT-NMLZ cow-GEN meat
'The beef that the person ate.' (Noonan 2008:223 for Chantyal)

The 'nominalization strategy' for relativization is particularly prevalent in the languages of the Americas (Comrie and Estrada-Fernández 2012; Zariquiey et al. 2019; Thornes 2023); in South America only, the *South American Indigenous Languages Structures* database (SAILS) (Krasnoukhova

2016) lists 50 languages belonging to 23 language families with this pattern. Among the well-known South American language families with this strategy are Quechua (Cole and Hermon 2011) and Tupían (Rodrigues and Cabral 2012b:538). It is a well-documented pattern in languages of the Tupí-Guaraní (TG) branch of the Tupían family, according to several individual studies of different languages of this family, including Anambé (Julião 2005:120-125), Araweté (Solano 2009:310-323), Asuriní (Tocantins) (Vieira 1993:175-194), Asuriní (Xingu) (Pereira 2009:211-216; 311-319) and (Pereira 2010), Guajá (Magalhães 2007:208-223), Guajajara-Tenetehára (Camargos 2017:185-236), Kamaiurá (Seki 2000:121-185), Guaraní Mbyá (Martins 2003:124-137), Paraguayan Guaraní (Estigarribia 2020:71-82), Parakanã (Silva 1999:43-49) and Tapirapé (Praça 2007:67-71).[1] Outside the Tupí-Guaraní branch, the 'nominalization strategy' is also described in members of two distantly related branches, in Sakurabiat (Mekens) of the Tuparí branch (Galúcio 2006, 2011) and Gavião of Rondônia of the Mondé branch (Moore 1989, 2006, 2012).

In Kamaiurá, for instance, Seki (2000:177) writes, "the strategy for relativization in the language is the nominalization of the embedded clause [...]". Consider the Kamaiurá example in (3). According to Seki, "the NP object *wyrapya* is modified (restricted) by the clause *kunu'uma pyhykarera* 'that caught the boy', which in turn has the verb *pyhyk* 'catch, hold' nominalized by the suffix *-tat* 'agentive'" (Seki 2000:178). In other words, the bracketed structure in (3) is a nominalization that can function as a nominal modifier.

(3) Akwama'e-a o-juka **wyrapy-a** [kunu'um-a pyhyk-ar-er-a].
man-N 3-kill hawk-N boy-N catch-NMLZ-PST-N
'The man killed the hawk that caught the boy.' (Seki 2000:178)

Against this backdrop, this chapter investigates the relationship between nominalization and relativization in Kawahíva, previously undocumented in the language. An example of the Kawahíva adnominal structure to be examined is provided within brackets in (4). In it, there is a gap in the nominalization, which I indicate with a dash. This position is associated with the modified noun (in bold).[2]

(4) **Tapy'ynha** [__ mõhãnga mõmbu-har-aver=a]=hẽa
non.indigenous medicine give-TR.SUBJ-PST.NOM=NMLZ=3.SG.FEM
'The female non-indigenous person that gave out medicine.' (Juma: Elicit)

In this chapter, I show that adnominal structures like (4) are nominalizations in Kawahíva. This is evidenced by their nominal morphosyntax, well-known from its sister languages. For instance, they exhibit the nominal temporal marker *-ver*. However, I argue that characterizing

[1] This list contains at least one member of each of the branches proposed in language-internal classifications of the Tupí-Guaraní branch (Rodrigues and Cabral 2002; Michael et al. 2015).

[2] I do not indicate the gap position when reproducing data with adnominal nominalizations from other languages as the source does not include them in the examples. However, some sources indicate that the modified noun is associated with a position within the nominalization while discussing the examples. For example, Seki (2000:178) states, "the noun phrase relativized is omitted from the relative clause, and its syntactic-semantic role is recovered by the nominalizing affix [...]".

these structures as nominalizations on its own is insufficient and that there is additional clear evidence that the nominalized clause also involves a distinct operation, namely relativization. I argue this operation involves 'extraction' or a 'filler-gap' dependency, a hallmark of relative clauses (and other extraction-based constructions) cross-linguistically. Evidence for this proposal will come from the fact that adnominal nominalizations i) exhibit sensitivity to island effects and ii) allow a long-distance dependency between the gap in the nominalization and the filler (i.e., the modified noun); the latter is the long-distance dependency property of extraction.

Furthermore, we will also see that adnominal nominalizations cannot undergo extraposition, a trait of a particular type of relative clause formation in several languages (i.e., the 'raising' strategy for relative clause formation) (Bhatt 2002; Hulsey and Sauerland 2006). I tie the absence of extraposition of adnominal nominalizations to my proposal that such structures also involve relativization in Kawahíva. If extraposition is impossible, then it is likely that the adnominal nominalization involves a 'raising' strategy for relativization, which does not admit extraposition. These facts support the proposal that adnominal nominalizations are best described as nominalized relative clauses, at least in Kawahíva.

This chapter is structured as follows. In section §5.2, I discuss my assumptions about nominalization and relativization. Then, in section §5.3, I present the structure described as an adnominal nominalization in the Tupí-Guaraní language family. In section §5.4, I show these structures exhibit the hallmarks of both nominalization and relativization. Based on these results, I propose that adnominal nominalizations are nominalized relative clauses, at least in Kawahíva.

5.2 Assumptions about nominalization and relativization

In this section, I review the phenomena of nominalization and relativization and their hallmark of syntactic properties. This will form the basis for analyzing adnominal structures in Kawahíva in §5.4.

5.2.1 Nominalization

Nominalization, the derivational process of forming a nominal from some other word class (mostly commonly a verb), is highly pervasive cross-linguistically (Crystal 1980:328). Typologically, nominalizations are commonly distinguished between lexical nominalization and clausal nominalization (Comrie and Thompson 1985; Baker and Vinokurova 2009). English agent-denoting nominalizer *-er* and clausal nominalizer with *-ing*, as exemplified in (5), are two of the most well-known nominalizations. In (5a), the *-er* nominalizer creates the individual-denoting noun *dancer* out of the verb *dance*. In (5b), the verb of the complement clause of *see* is nominalized with *-ing*.

(5) *English lexical and clausal nominalizations*
   a. Danc-**er**
   b. Max saw [the volunteers register-**ing** voters] (adapted from Toosarvandani (2010)).

Lexical nominalization (also referred to as 'entity', 'participant', 'argument', or 'nomenclative' nominalization) creates names for things in the world (Comrie and Thompson 1985; Kastovsky 1985; Toosarvandani 2014). It has become a common view of lexical nominalizations that they avoid a clause-like syntax (Baker and Vinokurova 2009): Lexical nominalizations, but not clausal nominalizations, "[bear] only morphological and (often unpredictable and idiosyncratic) semantic relations to the associated verb" (Comrie and Thompson 1985:349). For instance, in the Turkic language Sakha, lexical nominalizations (created with the nominalizer *-AAccY*), but not clausal nominalizations, lack i) adverbs, ii) aspect, and iii) negation.[3]

(6) *Sakha agent nominalization disallow clausal syntax*

a. *Tünnüg-ü **sorujan** aldjat-aaccy kel-le
window-ACC intentionally break-AG.NMLZ come-PST
'The one who intentionally broke the window.'

b. *Bar-**ytalaa**-ccy kel-le.
go-FREQ-AG.NMLZ come-PST.3.SG.S
'The frequent goer came.'

c. *Suruj-**um**-aaccy kel-le.
write-NEG-AG.NMLZ come-PST.3.SG.S
'The one who does not write came; the nonwriter came.'

Clausal nominalizations (also referred to as 'event' or 'grammatical' nominalizations) are derived nominals with a verbal core. However, in contrast to lexical nominalizations, they "retain certain properties of the verbs […] they are related to". For instance, Sakha clausal-nominalizations (created with the nominalizer *-YY*) allow i) adverbs, ii) aspect, and iii) negation (Baker and Vinokurova 2009).

(7) *Sakha clausal nominalization allow clausal syntax*

a. Misha tünnüg-ü **sorujan** aldjat-yy-ta miigin kyyhyr(t)-ta.
Misha window-ACC intentionally break-EV.NMLZ-3.SG.P me.ACC anger-PST.3.SG.S
'Misha' s intentionally breaking the window angered me.' (Baker and Vinokurova 2009:523)

b. Misha onno-manna bar-**ytalaa**-hyn-a miigin kyyhyr-ta.
Misha here-there go-FREQ-EV.NMLZ-3.SG.P me.ACC anger-PST.3.SG.S
'Misha' s repeated goings hither and thither angered me.' (Baker and Vinokurova 2009:525)

c. Bil-**im**-ii
know-NEG-EV.NMLZ
'Ignorance' (Baker and Vinokurova 2009:526)

[3]The absence of clausal structure is also true for well-known lexical nominalizations, like the English agent nominalization with *-er*. For example, this nominalization can't be modified with an adverb (e.g., *I gave a reward to [the finder of the wallet so quickly], from Baker and Vinokurova (2009:519)).

In brief, nominalization is a process that creates a noun out of other word classes. Additionally, clausal nominalizations, but not lexical nominalizations, exhibit a clause-like syntax.

5.2.2 Relativization

Relative clauses are part of a family of constructions in natural language that also includes topicalization, *wh*-question, and focus constructions. These constructions are referred to as 'extraction' throughout several theoretical frameworks – formal (Chomsky 1977; Richards 2014) and typologically-oriented (Van Valin and Lapolla 1997; Croft 2001; Goldberg 2006; Kroeger 2007) - but also referred to as 'A-bar'-dependencies (Chomsky 1977; Richards 2014), 'filler-gap' dependencies or 'unbounded' dependencies (Kroeger 2007); I adopt the term 'extraction' throughout this paper as it is the most commonly used across different theoretical frameworks. In these constructions, an element moves from its initial or basic position (the 'base-generated' or 'gap' position) into its final, visible position (the 'landing site' or 'filler'). To illustrate concretely, consider the example in (8), an object *wh*-question in English: the object *wh*-word moves from its postverbal position to the clause-initial position, above the subject.

(8) **What** are you looking at __?

We see the same displacement process in an object relative clause like (9), where the object appears preceding the subject and not in the usual postverbal position for objects in English.

(9) Show me [**the book** which you were reading __]!

In this section, I review two general syntactic properties of relative clauses, nonlocality and island effects, which are also typical of constructions involving extraction. Additionally, in §5.4.3, I draw on the availability of extraposition as a diagnostic for relativization. I aim to establish that if extraposition is not possible for an adnominal nominalization, it indicates that the nominalization involves the syntax of relativization. This discussion will serve as a framework for examining the presence of a syntax of relativization in Kawahíva adnominal nominalizations.

Movement for extraction may be nonlocal (i.e., unbounded), in that there is no principled bound on the length of the distance between the base position of the moved element and its landing site (Kroeger 2007; Richards 2014). In (10-11), an object *wh*-question and object relative from English, the moved object crosses a clause boundary, enclosed within brackets.

(10) **What** did you say [you were looking at __]?

(11) Show me [**the book** which you said [you were reading __]]!

Non-locality is typical of extraction-based constructions cross-linguistically. It is found in extraction constructions of languages such as the Nilotic language Dinka (Van Urk and Richards 2015) and the Austronesian language Tagalog (Rackowski and Richards 2005). We can see the pattern of non-locality in *wh*-questions (12a-b) and relative clauses (13a-b) in both languages.[4]

[4]In the example (12b), the adjunct *kailan* 'when' modifies the event in the complement clause with *uwi* 'go home', not the 'saying' event of the intermediate verb. This is evidenced in the agreement registered on the latter, which indicates agreement with the object complement clause, not adjuncts.

(12) *Unboundedness in wh-questions in Dinka and Tagalog*

a. **Yeŋà** yḛ̱ Yâa̱r tâak, [ké c̱íi Bôl luéel, [yè cṵkkiṳ̀ __ tîŋ]]?
who HAB.NS Yaar.GEN think COMP PRF.NS Bol.GEN say COMP PRF.1.PL see
'Who does Yaar think that Bol said that we saw?' (Van Urk and Richards 2015:126)

b. **Kailan** [sa-sabih-in ng sundalo [na Ø-u-uwi ang pangulo __]]?
when ASP-say-ACC CS soldier that NOM-ASP-go.home ANG president
'When will the soldier say that the president will go home?' (Rackowski and Richards 2005:586)

(13) *Unboundedness in relative clauses in Dinka and Tagalog*

a. Ye ŋà [yí̱i tâak, [c̱íi Bôl __ tṭ̱iŋ]]?
be **who** HAB.2.SG think.NF PRF.OV Bol.GEN see.NF
'Who do you think Bol has seen?' (Van Urk 2015:97)

b. Ang **kalabaw** [na s<in>abi-∅ ng guro [na bi-bigy-an ng lalaki ng
ang water.buffalo that -ASP-say-ACC CS teacher that ASP-give-DAT CS man CS
bulaklak __]]
flower
'The water buffalo that the teacher said that the man would give a flower to.' (Rackowski and Richards 2005:586-587)

Another hallmark of extraction is a sensitivity to syntactic islands, a pattern well-known since Ross (1967): in extraction-based constructions, like relative clauses, the moving element cannot escape syntactic islands, such as adverbial clauses. They exhibit 'island effects'. We see this pattern in the English examples in (14-15), an object *wh*-question and an object relative. Crucially, the moved element is separated from its original position by a syntactic island, the *after*-adverbial clause.

(14) ??**What** did you buy [$_{island}$ after looking at __ for hours]]?

(15) ??Show me [**the book** which you bought [$_{island}$ after you read __ at the library]]!

Island effects are also found in Dinka topicalization (16a) and relativization out of an adverbial clause island (16b), as well as in Tagalog, both in *wh*-questions (17a) and relativization from the adverbial island (17b).

(16) *Island effects in wh-questions in Dinka*

a. *Tòony à-c̱íi Ádìt jà̱a̱l [$_{island}$ wu̱í̱n c̱íi Máyèn __ kuêem].
pot 3.S-PRF.OV Adit.GEN leave.NF when PRF.OV Mayen.GEN break.NF
'(Lit.) The pot, Adit left because Mayen broke' (Van Urk 2015:98-99)

b. *Yè **ŋó̱** c̱íi Ádìt jàal [$_{island}$ wu̱í̱n c̱íi Máyèn __ kuêem]?
be what PRF.OV Adit.GEN leave.NF when PRF.OV Mayen.GEN break.NF
'(Lit.) What did Adit leave [because Mayen broke _]?' (Van Urk 2015:98-99)

(17) *Island effects in relative clauses in Tagalog*

a. *Ano ang s<in>abi mo [$_{island}$ kung kailan b<in>ili ni Vic
what.NOM NOM <PFV>say[PV] 2.SG.GEN COMP when <PFV>buy[PV] GEN Vic
__]?

'What did you say when Vic bought?' (Hsieh 2019:555)

b. *Ang **libro=ng** s<in>abi mo [$_{island}$ kung kailan b<in>ili ni Vic __]
nom book=LK <PFV>say[PV] 2.SG.GEN COMP when <PFV>buy[PV] GEN Vic
'The book that you said when Vic bought.' (Hsieh 2019:555)

In §5.4, I show that Kawahíva adnominal nominalizations exhibit the two general hallmarks of extraction and cannot be extraposed. I contend this evidence strongly indicates that Kawahíva nominalizations are nominalized relative clauses. Before that, I review the 'nominalization strategy' of Tupí-Guaraní languages.

5.3 The 'nominalization strategy' for relativization in Tupí-Guaraní

In descriptive studies of several languages of the Tupí-Guaraní language family, relativization is described as another function of nominalization, i.e., they exhibit the 'nominalization strategy' for relativization, as schematized in (18).

(18) *The nominalization strategy for relativization*:
Noun [… nominalization …]

In this section, I review the adnominal nominalization construction of Tupí-Guaraní languages and the supporting arguments of it being a (clausal) nominalization. This discussion draws mainly on the Kamaiurá language, which is perhaps the best-described language in this respect as it has the most extensive discussion of nominalizations available within the family, with a 32-page chapter dedicated to the topic (Seki 2000:170-202). In addition to Kamaiurá, several individual studies on sister languages describe a similar strategy for relativization. The list includes descriptive studies by Julião (2005:120-125, for Anambé), Solano (2009:310-323, for Araweté), Vieira (1993:175-194, for Asuriní (Tocantins)), Pereira (2009:211-216; 311-319, for Asuriní (Xingu)) and Pereira (2010), Magalhães (2007:208-223, for Guajá), Camargos and Castro (2013); Camargos (2017:185-236, for Guajajara-Tenetehára), Martins (2003:124-137, for Guaraní Mbyá), Estigarribia (2020:71-82, for Paraguayan Guaraní), Silva (1999:43-49, for Parakanã), and Praça (2007:67-71, for Tapirapé).

5.3.1 The adnominal nominalization in Tupí-Guaraní

The relevant structure examined is provided in (19), an example from Kamaiurá. Seki (2000) describes the Kamaiurá structure within brackets in (19) as follows: "the NP object *wyrapya* is

modified (restricted) by the clause *kunu'uma pyhykarera* 'that caught the boy', which in turn has the verb *pyhyk* 'catch, hold' nominalized by the suffix *-tat* 'agentive'" (Seki 2000:178). In other words, the bracketed structure in (19) is a clausal nominalization that can function as a modifier to a noun.

(19) Akwama'e-a o-juka **wyrapy-a** [kunu'um-a pyhyk-ar-er-a].
man-N 3-kill hawk-N boy-N catch-NMLZ-PST-N
'The man killed the hawk that caught the boy.' (Seki 2000:178)

It is worth mentioning that Seki (2000) also refers to bracketed structures like (19) as relative clauses. However, it is clear that she does not assume that such structures involve the syntax of a relative clause. For example, she points out, "In Kamaiurá, there are no lexical items corresponding to relative pronouns. The strategy for relativization in the language is the nominalization of the embedded clause with nominalizing affixes" (Seki 2000:177-178).

Seki (2000:184) also describes similar structures, as in (20), as adnominal nominalizations but without an overt modified noun. She refers to them as 'headless relatives'. As mentioned in the previous paragraph, Seki (2000) does not suggest that these structures involve relativization, despite the label.

(20) 'aŋ a'iweru i-'am-ĩ [je=reraha-tar-er-a] ko=ra'e apa.
here ATEN 3-be-CIRC 1.SG=take-NMLZ-NOM.PST-N FS=AF/FM father
'Here is the one who took me, Father.' (Seki 2000:184)

Tupí-Guaraní scholars have examined several morphosyntactic properties of the adnominal structures such as (19) to describe them as nominalizations, including i) temporal markers, ii) negation, iii) agreement, and iv) syntactic complement status. I review these properties, mainly drawing on data from the Kamaiurá language.

One of the nominal properties of the adnominal nominalizations in Kamaiurá is their ability to bear nominal temporal morphology. In the example in (19), the adnominal nominalization appears with the nominal suffix *-er*. This temporal marker also appears on lexical nouns, as in (21).[5,6] To state this argument based on nominal temporal morphology more clearly, if structures like (19) can bear temporal markers such as *-er*, as do lexical nominals, it is because they are nominal.[7]

[5]In other languages, cognate morphemes are labeled in a way that also suggests the nominal past tense semantics, as in (Magalhães 2007:161)'s *sufixo de atualização nominal retrospectiva* (*the retrospective nominal suffix*)

[6]This temporal-related morphology associated with nouns has also been extensively discussed in Tonhauser (2006) based on the TG sister language Paraguayan Guaraní and analyzed as having an aspect/modality semantics. The ultimate analysis of these morphemes is irrelevant to the main observation that this morphology is associated with nouns, a morphosyntactic aspect of this word category not disputed by Tonhauser (2006).

[7]One additional argument one could make for the nominal status of the bracketed structure in (19) would draw on the morpheme ordering between the morpheme *-tar*, described as a 'nominalizer', and the nominal temporal suffix *-er*. That is, the 'nominalizer' inside the temporal marker follows the 'Mirror Principle' (Baker 1985), which states that the order of morphemes reflects the syntactic structure. As the temporal marker attaches to nominals, it is expected to appear after the nominalizer by that principle. However, this argument would not be available in Kawahíva as I do

(21) Pe-etsak=ane=te [i-kyaw-**er**-awa] ko=kwãj
2.PL.IMP-see=CONT=FOC 3-hammock-PST-PL FS=MS
'See their hammocks!' (Seki 2000:90)

Negation is another supporting evidence for the nominalization of the adnominal structure in Tupí-Guaraní languages. Seki (2000:123) describes that Kamaiurá adnominal structures exhibit *-e'ym* for negation. This morpheme is a negation marker for nouns, as in (22). While Seki (2000)'s grammar does not contain any example of the morpheme *-e'ym* in adnominal nominalizations, we still find examples of the cognate negation morpheme *-y'ým* in adnominal structures of Guajá (Magalhães 2007), as shown in (23).[8]

(22) Anite paku-**e'ym**-a rak tapi'ir-a te rake wa.
not agouti-NEG-N ATT tapir-N F ATT MS
'No, it was not agouti. It was tapir.' (Seki 2000:213)

(23) Arapahá [ika-pyr-**y'ỹm**-a] a-jká tá.
deer kill-NMLZ-NEG-N 1-kill PROJ
'I will kill the deer that was not killed.' (Magalhães 2007:285)

The argument marking morphology (or set B morphology, as described in this dissertation) is also drawn upon to support further that adnominal modifiers are nominalizations in Tupí-Guaraní languages. In these, local persons are marked similarly to possessors in possessive phrases. (24) provides an example of pronominal agreement marking in the Kamaiurá adnominal construction. In turn, (25) shows the same morpheme marking the possessor in a possessive phrase.

(24) Kye'i-a [**ne**=remi-ekar-er-a].
knife-N 2.SG=NMLZ-look.for-NOM.PST-N
'(My brother saw) the knife that you were looking for.' (Seki 2000:180)

(25) Mame **ne**=ra'yr-a reko-w?
where 2.SG-offspring-N be-CIRC
'Where is your son?' (Seki 2000:164)

not view the cognate morpheme to Kamaiurá *-tar*, which in Kawahíva is *-har*, as a nominalizer. In work in progress, I propose this morpheme is part of a paradigm of *wh*-agreement markers (Dos Santos In prep.). In section §3.3.1, I described the nominal morpheme *=a* as a syntactic nominalizer, both in nouns and adnominal nominalizations. Given my analyses of these morphemes, the morpheme ordering facts between the temporal marker and nominalizer do not follow the 'Mirror Principle'. However, my proposal that adnominal nominalizations exhibit the syntax of relativization does not hinge upon the analysis of the said morphology. I leave it for future work to reconcile the morpheme ordering facts and the present proposal for adnominal nominalizations.

[8]One could also draw on the morpheme order between the nominal negation marker *-y'ým* and the Guajá 'nominalizer' *-pyr* to argue this order also follows the Mirror Principle and, then, how this fact supports the nominalization status of adnominal nominalizations. While this argument might be available for Guajá, it does not carry over to Kawahíva as the cognate of Guajá *-pyr* is best understood as a *wh*-agreement marker (Dos Santos In prep.). I refer the reader to footnote 7 for a similar issue with this argument in Kawahíva involving nominal temporal markers.

Using an adnominal nominalization as the complement of a postposition provides additional evidence for its nominal syntactic status. In the Kamaiurá example in (26), the adnominal structure is followed by the postposition *wi* 'ablative'.

(26) A'e-a n=o-kwahaw-ite [je=o-taw-er-a] **wi**.
this-N NEG=3-know-NEG 1.SG=go-NMLZ-NOM.PST-N ABL
'(He) does not know (the place) I came from.' (Seki 2000:185)

A final argument for the nominal status of the adnominal structures in TG can be adduced based on the distribution of the morpheme *-a*. A core morphosyntactic property of nouns in Kamaiurá (Seki 2000:107-109) and several other TG languages is they occur with *-a*; for instance, this morpheme is present on subjects (27a), objects of verbs (27b), and objects of postpositions (27c).

(27) a. Kunu'um-**a** tete rak o-ho ko'yt.
boy-N only ATT 3-go FS
'Only the boy went.' (Seki 2000:107)

b. Kunu'um-a huwaj-**a** w-ekyj.
boy-N tail-N 3-pull
'The boy is pulling (its) tail.' (Seki 2000:107)

c. Je=ruw-**a** nite
1.SG=father-N with
'With my father' (Seki 2000:107)

Notably, *-a* also appears in the adnominal structures. Although Seki (2000) does not draw on the use of *-a* in these structures as supporting evidence for their nominalization, this is nevertheless explicitly done elsewhere. For instance, in a section that describes cognate structures in the sister language Guajá, Magalhães (2007:213) says that "the resulting nominalizations, like any other noun, receives the nominal suffix *-a* [...]". The analysis of the morpheme *-a* in the TG family varies; it has been suggested it is a case marker (Rodrigues 1996, 2001) and a determiner (Queixalós 2006); I describe it as a syntactic nominalizer in (see §3.3.1). The point of the matter is that nouns, but not other word categories, must appear with *-a*. Given that this requirement is also in place for adnominal structures, I take it as further that these structures are nominal and comptible with the idea that they are nominalizations.

In sum, the following properties characterize the adnominal clausal nominalization in Tupí-Guaraní: nominal morphology, including nominal temporal markers, nominal negation, pronominal agreement morphology, and nominal syntax, as they are used as syntactic complements.

5.4 Distinguishing nominalization from relativization in Kawahíva

In this section, I argue that in addition to the typical Tupí-Guaraní nominal morphosyntax, Kawahíva adnominal nominalizations exhibit A-bar movement, the typical syntax of relativization. I start by presenting further details about the Kawahíva constructions described as adnominal nominalizations in sister languages to Kawahíva. I propose these structures are best understood as nominalized relative clauses. On the first count, they are nominalizations due to their undeniably nominal morphosyntax, the most well-studied aspect of these constructions in the Tupí-Guaraní family. However, I show that these structures also involve relativization as they exhibit syntactic properties of relative clauses, including extraction (Ross 1967; Kroeger 2007; Richards 2014).

An instance of the adnominal structure examined in Kawahíva is provided in (28). Adnominal constructions occur immediately to the right of the modified noun unless the modified noun is fronted for pragmatic reasons (e.g., focus) and appears in clause-initial position.

(28) A-epia ki jie **y'va** [tapy'ynha=gã __ remb-i-tyk-aver=a].
1.SG.A-see PST 1.SG tree non-indigenous=PL WH.OBJ.I-i-take.down-NOM.PST=NMLZ
'I saw the trees that the non-indigenous people took down.'

In addition to the bracketed adnominal nominalization in (28), Kawahíva also exhibits structures like (29), in which there is no overt head that is associated with the gap in the nominalization.

(29) A-epia ki jie [tapy'ynha=gã __ remb-i-tyk-aver=a].
1.SG.A-see PST 1.SG non-indigenous=PL WH.OBJ.I-i-take.down-NOM.PST=NMLZ
'I saw the ones (= the trees) that the non-indigenous people took down.'

I assume cases like (29) involve the same nominalization in (28), except that the overt noun in the latter is dropped. In line with this idea and the proposal in this chapter that such Kawahíva structures are nominalized relative clauses, I consider (29) a nominalized headless or free relative (see de Vries (2018); Caponigro (2021) for an overview on this type of relative clause). I treat headed and headless nominalized relative clauses on par here.[9]

An important morphosyntactic property of the adnominal nominalizations in Kawahíva is they exhibit the morphological paradigm in Table 5.1. This paradigm is similar to other Tupí-Guaraní languages, not shown here.

GAP POSITION	MARKER	GAP POSITION	MARKER
trans. subject	*-har*	oblique	*-hav & -var*
trans. object	*remb- & -pyr*	intrans. subject	*-va'e*

Table 5.1: Co-varying morphology of Kawahíva adnominal nominalizations.

[9]To the best of my knowledge, the choice between an overt and a null noun depends on the recoverability of its referent. Further systematic investigation of this distribution should clarify whether 'recoverability' is the correct condition for either choice.

The choice of a particular marker from Table 5.1 depends on the grammatical function associated with the gap position within the adnominal structure. Descriptively, *-har* is used when the gap corresponds to a transitive subject. The morphemes *remb-* and *-pyr* are used when the gapped position is a transitive object. In turn, *-hav* and *-var* are used when the gap position is associated with an oblique. Finally, *-va'e* is used for the gap of the subject of an intransitive verb. We will see that, in cases where there are two morphemes for the same gap position (i.e., object position and complement of postpositions), the choice boils down to nuanced aspectual differences between *remb-* and *-pyr*, and in the case between *-hav* and *-var*, the choice depends on the postposition in the oblique phrase. When referring to a particular adnominal nominalization, I include the label for the grammatical role associated with the gap position in the adnominal nominalization. Thus, adnominal nominalizations with *-har* are referred to as 'transitive subject adnominal nominalizations'; those with *-va'e* as 'intransitive subject adnominal nominalizations', and so on.

While cognates of these morphemes are described as nominalizers in sister languages to Kawahíva (Jensen 1998, 1999; Rodrigues 2001), they are best understood as *wh*-agreement markers in Kawahíva (Dos Santos In prep.), a type of agreement first described for the Austronesian languages Chamorro and Palauan (Chung and Georgopoulos 1988; Chung 1994). *Wh*-agreement is a special form of agreement indexing an extracted phrase (Deal 2016:170). Note, however, that the proposal presented in this chapter does not depend on any analysis of this morphological paradigm in Kawahíva.[10]

In the remainder of this section, I establish that Kawahíva adnominal structures (like the bracketed structure in (28) above) exhibit both nominal-like morphosyntax (based on their use of nominal temporal and negation morphology) and clausal-like morphosyntax (based on the use of adverbials). These facts provide strong evidence that the adnominal structures are clausal nominalizations.

5.4.1 Arguments for clausal nominalization

The first evidence that adnominal structures are clausal nominalizations comes from the enclitic *=a*. This nominalizer (described in §3.3.1) is required in these constructions. Examples (30) provide a minimal pair with (30a) and without (30b) the enclitic. Only in the absence of *=a* is the adnominal structure ungrammatical.

(30) a. A-epia ki jie **y'va** [tapy'ynha=gã __
1.SG.A-see PST 1.SG tree non-indigenous=PL
remb-i-tyk-aver=a].
WH.OBJ.I-i-take.down-NOM.PST=NMLZ

[10]However, a *wh*-agreement account of this morphology would count as a piece of morphosyntactic evidence for my proposal that adnominal nominalizations involve relativization. This is because if there is relativization in said structures, we could imagine a language where overt agreement with relative heads is available. One case where erstwhile adnominal nominalizations are analyzed as involving relativization and having *wh*-agreement with the relative head is the Algonquian language Ojibwe (Lochbihler and Mathieu 2013).

‘I saw the trees which the non-indigenous people took down.’

b. *A-epia ki jie **y’va** [tapy’ynha=gã __ remb-i-tyk-aver].
1.SG.A-see PST 1.SG tree non-indigenous=PL WH.OBJ.I-i-take.down-NOM.PST
‘I saw the trees which the non-indigenous people took down.’

Transitive subject adnominal nominalizations (or nominalizations with a gap in the transitive subject position) exhibit the suffix *-har* (31). The grammatical role must be that of a transitive subject. Attempts to use *-har* in a nominalization with a gap in the intransitive subject position are not acceptable (32). They are revised by the speaker with a repair like (33). In this example, the morpheme that appears is *-va’e*, used in nominalizations with a gap in the intransitive subject position, as we will see shortly.

(31) **Tapy’ynha** [__ mõhãnga mõmbu-**har**=a]=hẽa
non.indigenous medicine give-WH.TRANS.SUBJ=NMLZ=3.SG.FEM
‘The female non-indigenous person who gives out medicine’ (Juma: Elicit)

(32) *A-pyhy jie **irãmutxinguhua** [__ o-hẽr-**ahar**=a].
1.SG.A-catch 1.SG chicken 3.COR-leave-WH.TRANS.SUBJ=NMLZ
‘I caught the chicken that left.’ (Juma: Elicit)

(33) A-pyhy jie **irãmutxinguhua** [__ o-hẽr-**ava’e**=a].
1.SG.A-catch 1SG chicken 3.COR-leave-WH.INTR.SUBJ=NMLZ
‘I caught the chicken that left.’ (Juma: Elicit)

Transitive subject adnominal nominalizations exhibit the nominal past and future temporal markers *-ver* and *-ham* in (34a-b) and the nominal negation morpheme *-i’im* in (34c). These adnominal structures also exhibit clausal-like morphosyntax in that adverbials appear with the verb in the adnominal nominalization, including *-katu* ‘well, prettily’ (35a), *-hym* ‘forcefully’ (35b), and *-tehe* ‘in vain, idly’ (35c).

(34) a. T-a-juka jie **piranhuhua** [__
OPT-1.SG.A-kill 1.SG piranha
nde=po-u’u-har-**aver**=a].
2.SG=hand-bite-WH.TRANS.SUBJ-NOM.PST=NMLZ
‘I’d like to kill the piranha that bit your hand.’ (Juma: Elicit)

b. A-epia ji **koemba’ea** [__ v-a’yra
1.SG.A-see 1.SG man 3.COR-offspring
mbo’e-har-**aham**=a]=ga.
teach-WH.TRANS.SUBJ-NOM.FUT=NMLZ=3.SG.MASC
‘I saw the man who will teach his son.’ (Jupaú: Elicit)

c. A-epia ji **koemba'ea** [__ v-a'yra
1SG.A-see 1.SG man 3.COR-offspring
mbo'e-har-**i'im**=a]=ga.
teach-WH.TRANS.SUBJ-NOM.NEG=NMLZ=3.SG.MASC
'I saw the man who does not teach his own offspring.' (Jupaú: Elicit)

(35) a. T-a-piang **kunhã** [__ i-kwa-har-**akatu**=a].
OPT-1.SG.A-see woman i-know-WH.TRANS.SUBJ-WELL=NMLZ
'I would like to see the girl who knows (it) well.' (Juma: Elicit)

b. A-epia ji **tapy'ynha** [__ y'va
1.SG.A-see 1.SG non.indigenous tree
retyk-ahar-**ahym**=a]=ga.
take.down-WH.TRANS.SUBJ-FORCEFULLY=NMLZ=3.SG.MASC
'I saw the non-indigenous person who forcefully took down a tree.' (Juma: Elicit)

c. A-epia ji **tapy'ynha** [__ y'va
1.SG.A-see 1.SG non.indigenous tree
retyk-ahar-**atehe**=a]=ga.
take.down-WH.TRANS.SUBJ-IN.VAIN=NMLZ=MASC
'I saw the non-indigenous person who took down a tree for no reason.' (Juma: Elicit)

Transitive object adnominal nominalizations (or nominalizations with a gap in the transitive object position) exhibit one of two morphemes. They are *remb-* and *-pyr*. The choice depends on a nuanced contrast of two aspect distinctions, one between 'completive' vs. 'non-completive' readings and another between 'wholly affected' and 'partially affected' readings. Examples (36), including the speaker's comments, illustrate the distribution of these morphemes based on the completive vs. non-completive distinction.

(36) a. E-piokan **karovaruhua** ji=ve [Wesley=ga __ i-mõpu-**pyr**=a]!
2.SG.IMP-show agouti 1.SG=to Wesley=3.SG.MASC i-shoot-WH.OBJ.II=NMLZ
'Show me the agoutis which Wesley shot dead!' (Juma: Elicit)
Speaker's comment: The agoutis are already dead.

b. E-piokan **karovaruhua** ji=ve [Wesley=ga __ **rem**-i-mõpu=a]!
2.SG.IMP-show agouti 1.SG=to Wesley=3.SG.MASC WH.OBJ.I-i-shoot=NMLZ
'Show me the agoutis which Wesley shot!' (Juma: Elicit)
Speaker's comment: this can be uttered in a scenario where there are two agoutis shot but not dead yet, and I want to know which one Wesley shot.

Examples like (37) also support that the completive vs. non-completive contrast underlies the choice between one of the two morphemes. An example like (37a) is ill-formed as the choice of *-pyr* in the adnominal nominalization, which implies the fish-eating event was total, clashes with the following sentence, which implies that the event was partial. On the other hand, an example

like (37b) is well-formed as the choice of *remb-* does not imply that the fish-eating event described in the nominalization was not total, which is further confirmed by the speaker's comment.

(37) a. *E-mõmbon [ga __ i-'u-pyr=a], **pira**!
2.SG.IMP-throw.out 3.SG.MASC i-eat-WH.OBJ.II=NMLZ fish
Nd-o-'u-pav-i ga.
NEG-3.A-eat-COMPLETELY-NEG 3.SG.MASC
'Throw out (that) which he ate, the fish! He did not eat it completely.' (Juma: Elicit)

b. E-mõmbon [ga __ remb-i-'u=a], **pira**!
2.SG.IMP-throw.out 3.SG.MASC WH.OBJ.I-i-eat=NMLZ fish
'Throw out (that) which he ate, the fish!' (Juma: Elicit)
Speaker's comment: He ate half of the fish.

Object adnominal nominalizations also show a nominal-like and clausal-like morphosyntax. The former is manifested in the use of nominal temporal (38) and nominal negation morphology (39) in the nominalization. Its clausal morphosyntax is shown through the use of adverbials, including *-tehe* 'in vain, idly' (40), *-katu* 'well, prettily' (41), and *-hym* 'forcefully' (42).

(38) a. A-'u ji **mbiara** [__ i-mõtyryry-pyr-**aver**=a].
1.SG.A-eat 1.SG meat i-fry-WH.OBJ.II-NOM.PST=NMLZ
'I ate the meat that (someone) fried.' (Juma: Elicit)

b. **'yva** ki [tapy'ynha=ga __ remb-i-tyk-**aver**=a].
tree PST non.indigenous=PL WH.OBJ.I-i-take.down-NOM.PST=NMLZ
'It was the tree that the non-indigenous person took down.' (Juma: Elicit)

(39) a. Oro-'u ore **mbiara** [__ i-mõtyryry-pyr-**e'ym**=a].
1.EXCL.A-eat 1.EXCL meat i-fry-WH.OBJ.II-NOM.NEG=NMLZ
'We ate meat that is not fried.' (Juma: Elicit)

b. A-'u jie **matera=mõ** [ga __
1.EXCL.A-eat 1.EXCL whatchamacallit=INDF 3.SG.MASC
remb-i-mboapyk-**e'ym**=a].
WH.OBJ.I-i-cook-NOM.NEG=NMLZ
'I ate something that he did not cook.' (Juma: Elicit)

(40) a. E-mõmbon **pira** [Wesley=ga __ i-mõtyryry-pyr-**atehe**=a]!
2.SG.IMP-throw fish Wesley=3.SG.MASC i-fry-WH.OBJ.II-IN.VAIN=NMLZ
'Throw away the fish that Wesley (completely) fried in vain (because he won't eat it)!' (Juma: Elicit)

b. E-mõmbon **pira** [Wesley=ga __ remb-i-mbotyryry-**tehe**=a]!
2.SG.IMP-throw fish Wesley=3.SG.MASC WH.OBJ.I-i-fry-IN.VAIN=NMLZ
'Throw away the fish that Wesley (partially) fried in vain (because he won't eat it)!' (Juma: Elicit)

(41) a. Oro-'u ore **mbiara** [Wesley=ga __ i-mõtyryry-**katu**-pyr=a].
1.EXCL.A-eat 1.EXCL meat Wesley=3.SG.MASC i-fry-WH.OBJ.II-WELL=NMLZ
'I ate the meat that Wesley fried well.' (Juma: Elicit)

b. A-'u jie **matera=mõ** [Wesley=ga __
1.EXCL.A-eat 1.EXCL whatchamacallit=INDF W.=3.SG.MASC
remb-i-mõtyryry-**katu**-e'ym=a].
WH.OBJ.I-i-fry-WELL-NOM.NEG=NMLZ
'I ate something that Wesley did not fry well.' (Juma: Elicit)

(42) a. Ere-piang po nde **inihãma** [Wesley=ga __ i-mõmbor-ipyr-**ahym**=a]?
2.SG.A-see IRR 2.SG rope W.=3.SG.MASC i-throw-WH.OBJ.II-FORCEFULLY=NMLZ
'Did you see the rope that Wesley threw forcefully' (Juma: Elicit)

b. Ere-piang po nde **inihãma** [Wesley=ga __ rem-i-mõmbor-**ahym**=a]?
2.SG.A-see IRR 2.SG rope W.=3.SG.MASC WH.OBJ.I-i-throw-FORCEFULLY=NMLZ
'Did you see the rope that Wesley partially threw forcefully' (Juma: Elicit)

Kawahíva has two oblique adnominal nominalizations (or nominalizations with a gap associated with the complement of a postpositional phrase in oblique function). They are distinguished by the morphemes *-hav* and *-var*. The choice between the two depends on the postpositional phrase in oblique function; there is no property of a noun or pronoun complement to the postposition that correlates with this distinction. Table 5.2 provides a subset of postpositions that appear in the adnominal nominalization marked with *-hav* and *-var*.[11]

-hav	*-var*
pymõ 'with'	*ipe* 'in'
pe 'to, for'	*eviri* 'behind'
	enõnde 'in front'
	pype 'inside'
	pyri 'near'

Table 5.2: The distribution of *-hav* and *-var* in oblique PP relatives.

Descriptively, a gap in the complement position of the postpositional phrase headed by *pymõ* 'with' triggers *-hav* in the nominalization. First, consider (43a), an example of a transitive subject nominalization that also contains the oblique PP headed by *pymõ* 'with'. Notably, this construction is marked with *-har*. In contrast, (43b) includes an adnominal nominalization where the gap is associated with the oblique phrase headed by *pymõ* 'with'. As a result, the nominalization exhibits *-hav*. In turn, the example (43c) shows that the nominalization with a gap in the complement of *pymõ* 'with' cannot appear with a different gap-tracking morpheme than *-hav*, such as *-var*.

[11]Note how the postpositions that trigger *-var*, but not those that appear with *-hav*, are semantically related to spatial deixis. One could argue that this distinction is also responsible for the distribution of the two types of oblique adnominal nominalizations.

(43) a. **Tapy'ynha** [__ ʻyvakytxihauhua=**pymõ** ʻyva
non.indigenous chainsaw=with tree
rety-har=a]=ga
take.down-WH.TRANS.SUBJ=NMLZ=3.SG.MASC
'(I saw) the non-indigenous male who took down (a) tree(s) with a chainsaw.' (Juma: Elicit)

b. **Paratxia** [__ ahe mbiara kytxi-**hav**=a].
knife people meat cut-WH.OBL.I=NMLZ
'The knife [(with) which we cut meat].' (Juma: Elicit)

c. *__Paratxia__ [__ ahe mbiara kytxi-**var**=a].
knife people meat cut-WH.OBL.II=NMLZ
'The knife [with which we cut meat].' (Juma: Elicit)

In contrast, a gap in the postpositional phrase headed by a postposition like *ipe* 'in' triggers the suffix *-var* in the adnominal construction. First, consider the example in (44a), which contains an oblique postpositional phrase with *ipe* 'in'. Example (44b) shows the adnominal structure marked with *-var*. In turn, example (44c) demonstrates that an adnominal structure marked with *-var* is not compatible with cases where the gap is within a postpositional phrase headed by *pymõ* 'with'.

(44) a. Hẽa=ra'uva, jie a-hepiang okar=**ipe**.
3.SG.FEM=spirit 1.SG 1.SG.A-see terrace=in
'Her spirit, I saw (it) outside.'

b. **Ipymbu'iva'ea** [__ okar=ipe-**var**=a]
motorcycle terrace=in-WH.OBL.II=NMLZ
'The motorcycle which is outside' (Juma: Elicit)

c. *Oga=**pymõ**-var=a
house=with-WH.OBL.II=NMLZ
'The one with the house' (Juma: Elicit)

Importantly, oblique adnominal nominalizations exhibit nominal morphosyntax; they use nominal temporal markers (45) and nominal negation (46).

(45) a. **Paratxia** [__ ahe mbiara kytxi-hav-**aver**=a]
knife people meat cut-WH.OBL.I-NOM.PST=NMLZ
'The knife [with which we cut meat].' (Juma: Elicit)

b. E-run **yngu'a** [__ nde=pyri-var-**aver**=a]!
2.SG.IMP-bring mortar 2.SG=near-WH.OBL.II-NOM.PST=NMLZ
'Bring the mortar that was close to you!' (Juma: Elicit)

(46) a. A-mõkanhym ji **paratxia** [__ ahe mbiara i-kytxi-hav-**e'ym**=a].
1.SG.A-forget 1.SG knife people meat i-cut-WH.OBL.I-NOM.NEG=NMLZ
'I forgot the knife with which we did not cut the meat.' (Juma: Elicit)

b. E-run **yngu'a** [__ okar=ipe-var-**e'ym**=a]!
2.SG.IMP-bring mortar terrace=in-WH.OBL.II-NOM.NEG=NMLZ
'Bring the mortar that is not outside!' (Juma: Elicit)

Oblique adnominal nominalizations also appear with adverbials, showing a clause-like syntax. This is shown for both *hav*-marked nominalizations, in (47), and *var*-marked nominalizations, in (48).

(47) a. A-mõkanhym ji **paratxia** [__ ahe mbiara i-kytxi-hav-**akatu**=a].
1.SG.A-forget 1.SG knife people meat i-cut-WH.OBL.I-WELL=NMLZ
'I forgot the knife with which we cut the meat well.' (Juma: Elicit)

b. A-mõkanhym ji **paratxia** [__ ahe mbiara i-kytxi-hav-**ahym**=a].
1.SG.A-forget 1.SG knife people meat i-cut-WH.OBL.I-FORCEFULLY=NMLZ
'I forgot the knife with which we cut the meat forcefully.' (Juma: Elicit)

c. A-mõkanhym ji **paratxia** [__ ahe mbiara i-kytxi-hav-**atehe**=a].
1.SG.A-forget 1.SG knife people meat i-cut-WH.OBL.I-IN.VAIN=NMLZ
'I forgot the knife with which we cut the meat for no reason.' (Juma: Elicit)

(48) a. E-run **yngu'a** [__ okar=ipe-var-**akatu**=a]!
2.SG.IMP-bring mortar terrace=in-WH.OBL.II-WELL=NNMLZ
'Bring the mortar that is just outside!' (Juma: Elicit)

b. E-run **yngu'a** [__ okar=ipe-var-**atehe**=a]!
2.SG.IMP-bring mortar terrace=in-WH.OBL.II-IN.VAIN=NMLZ
'Bring the mortar that is outside for no reason!' (Juma: Elicit)

Finally, intransitive subject adnominal nominalizations (or nominalizations with a gap in the intransitive subject position) are marked with the suffix *-va'e*.[12] In the examples in (49), the suffix appears in nominalizations whose gap is connected to the subject of an active/unergative verb (49a) and also to the subject of a descriptive/unaccusative verb (49b). Nominalizations of derived intransitive verbs and verbs that can be transitive and intransitive (aka *labile* verbs) also exhibit this suffix. For example, contrast the ungrammatical combination of the plain transitive verb *ka* 'break something open' with *-va'e* to the grammatical combination of the same verb marked with the reflexive morpheme *ji-* and *-va'e* in (50a-b).

(49) a. E-rupã **javatxinga** [__ o-kuru-va'e=a oga=pyri]!
2.SG.IMP-hit dog 3.COR-urinate-WH.INTR.SUBJ=NMLZ house=near
'Hit the dog that urinated near the house!' (Juma: Elicit)

b. Ahe nd-o-hepiag-i-ve **mãjioka** [__ ymy-hu-**va'e**=a].
people NEG-3.A-see-NEG-YET yuca be.old-AUG-WH.INTR.SUBJ=NMLZ

[12]The three-way distinction between transitive subjects, objects, and intransitive subjects in this *wh*-agreement pattern could be described as a tripartite alignment system.

'We did not see yuca that is from the olden times.' (talking about yuca that was common during the times before the contact with the non-indigenous people) (Juma: Text)

(50) a. *O-ka-**va'e**=a
3.COR-break.open-WH.INTR.SUBJ=NMLZ
'That which someone broke' (Juma: Elicit)

b. O-ji-ka-**va'e**=a
3.COR-REFL-break.open-WH.INTR.SUBJ=NMLZ
'That which broke open' (Juma: Elicit)

Like all other adnominal nominalizations, intransitive subject adnominal nominalizations also exhibit nominal morphosyntax, including nominal temporal (51a) and nominal negation (51b). These adnominal structures also appear with adverbs (52).

(51) a. A-pyhy jie **irãmutxinguhua** [__ o-hẽr-ava'e-**ver**=a].
1.SG.A-catch 1.SG chicken 3.COR-leave-WH.INTR.SUBJ-NOM.PST=NMLZ
'I caught the chicken that left.' (Juma: Elicit)

b. A-pyhy jie **irãmutxinguhua** [__ o-hẽr-ava'e-**'ym**=a].
1.SG.A-catch 1.SG chicken 3.COR-leave-WH.INTR.SUBJ-NOM.NEG=NMLZ
'I caught the chicken that did not leave (the chicken house).' (Juma: Elicit)

(52) a. **Okyta'ietea** [__ txũ'i-va'e-**katu**=a]
dry.wood be.small-WH.INTR.SUBJ-WELL=NMLZ
'The dry wood which was good' (Juma: Elicit)

b. **Okyta'ietea** [__ txũ'i-**tete**-va'e=a]
dry.wood be.small-REALLY-WH.INTR.SUBJ=NMLZ
'The dry wood which was really small' (Juma: Elicit)

In brief, I have shown that the following properties characterize the adnominal clauses in Kawahíva as clausal nominalizations: they appear with the nominalizer *=a*, nominal temporal markers, nominal negation, and adverbs. We have also seen that adnominal nominalizations exhibit a four-way morphological paradigm that co-varies with the grammatical role of the internal gap position.

In the following section, we will see that these adnominal clausal nominalizations also exhibit the syntax of relativization.

5.4.2 Arguments for relativization

Recall from §5.2.2 that relative clauses are part of a family of constructions that involve 'extraction'. Cross-linguistically, extraction i) exhibits nonlocality and ii) is sensitive to syntactic islands (i.e., they exhibit island effects) (Richards 2014). Additionally, in some languages, relative clauses can't be extraposed.

In this section, I establish that adnominal nominalizations in Kawahíva involve extraction and can't be extraposed. I argue that these facts strongly suggest that adnominal structures in Kawahíva are not simply clausal nominalizations in adnominal function. I propose they are nominalized relative clauses.

5.4.2.1 Nonlocality

Nonlocality provides evidence that adnominal nominalizations in Kawahíva involve extraction. This pattern is seen in long-distance extraction examples like (53). No previous work has described long-distance extraction in Tupí-Guaraní, to my knowledge. In (53), the adnominal nominalized structure also contains a complement clause with a gap; the modified noun is associated with the gap within the complement clause. I use subscripts to indicate the gap position is associated with the modified noun.

(53) **Mboja**$_{i}$ [Kajubi=ga remb-i-piak=a [_$_{i}$ javatxinga
snake Kajubi=3.SG.MASC WH.OBJ.I-i-see=NMLZ dog
u'u-har=a]]
bite-WH.TRANS.SUBJ=NMLZ
'(I killed) the snake [which Kajubi saw [that bit the dog]]' (Juma: Elicit)

In (53), the transitive subject in the innermost bracketed clause, the complement clause of *hepiang* 'see', is associated with the noun *mboja* 'snake'; the latter has undergone long-distance extraction.

While the example in (53) shows an instance of long-distance extraction for transitive subject adnominal nominalizations, this is also true for all different adnominal nominalizations discussed in the previous section, namely object (54), oblique (55), and intransitive subject (57) adnominal nominalizations. They can all exhibit a gap separated from the modified noun by an intermediate clause boundary.

(54) **Tupãhua** [Kajuvi=ga rem-i-mõmbe'u=a [Davi=ga __ remb-i-piak=a]]
gun K.=3.SG.MASC WH.OBJ.I-i-tell=NMLZ D.=3.SG.MASC WH.OBJ.I-i-see=NMLZ
'(I brought) the gun which Kajuvi told (me) that Davi saw' (Juma: Elicit)

(55) **Karovaruhua** [nde rem-i-mõmbe'u=a [Wesley=ga __ i-mõpu-pyr=a]]
agouti 2.SG WH.OBJ.I-i-tell=NMLZ W.=3.SG.MASC i-shoot-WH.OBJ.II=NMLZ
'(Show me) the agouti which you told (me) that Wesley shot (dead)' (Juma: Elicit)

(56) **Y'vakytxihauhua** [Davi=ga remb-i-piak=a [__ tapy'ynha=gã
chainsaw D.=3.SG.MASC WH.OBJ.I-i-see=NMLZ non.indigenous=3.SG.MASC
y'va kytxi-hav=a]]
tree cut.down-WH.OBL.I=NMLZ
'(I threw away) the chainsaw which Davi saw that the non-indigenous people used to cut down trees' (Juma: Elicit)

(57) **Javatxinga** [Kajubi=ga rem-ĩ-mõmbe'u=a [__ o-mõnõ-va'e=a]]
dog K.=3.SG.MASC WH.OBJ.I-i-tell=NMLZ 3.COR-die-WH.INTR.SUBJ=NMLZ
'(Threw away) the dog which Kajuvi told (me) that died' (Juma: Elicit)

The next section demonstrates that the dependency between the modified noun and gap position in the adnominal nominalization cannot be established over a syntactic island.

5.4.2.2 Island effects

Further evidence that Kawahíva adnominal nominalizations involve extraction is they exhibit island effects. Consider the example in (58). The innermost bracketed structure is an adjunct clause, with a subject gap. The outermost bracketed structure is the adnominal nominalization, which modifies the noun *javatxinga* 'dog'.

(58) *E-juka **javatxinga** [Kajubi=ga remb-i-kwar=a [__ Davi=ga
2.SG.IMP-kill dog Kajubi=3.SG.MASC WH.OBJ.I-i-tie=NMLZ Davi=3.SG.MASC
u'ur=a=mẽ]]!
bite=NMLZ=BECAUSE
'Kill the dog that Kajubi tied up because (it) bit Davi!' (Juma: Elicit)

Notably, this example shows that an attempt to modify *javatxinga* 'dog' with a juxtaposed nominalization, which is associated with the subject gap position of the adjunct clause, is not well-formed (note that the same effects are observed in the English translation). Other adnominal nominalizations—transitive subject (59a), intransitive subject (59b), and oblique adnominal nominalization (59c)—also exhibit the same effects.

(59) a. *E-piokan ji=ve **tapy'ynha** [__ mõhãnga
2.SG.IMP-show 1.SG=to non.indigenous medicine
mõmbu-har-aver=a [__ rur=a=mẽ]=hẽa]!
give-WH.TRANS.SUBJ-NOM.PST=NMLZ come=NMLZ=WHEN=3.SG.FEM
'Show me the non-indigenous female person who gave out medicine when she came in!' (Juma: Elicit)

b. *E-pyhyng **irãmutxinguhua** [__ o-hẽr-ava'e=a [__ moja
2.SG.IMP-catch chicken 3.COR-leave-WH.INTR.SUBJ=NMLZ snake
repiak=a=mẽ]]!
see=NMLZ=WHEN
'Catch the chicken that left when (it) saw a snake!' (Juma: Elicit)

c. *E-piokan ji=ve **paratxia** [__ nde mbiara kytxi-hav=a [__
2.SG.IMP-show 1.SGto knife 2.SG meat cut-WH.OBL.I=NMLZ
'ar=a=mẽ]]!
fall=NMLZ=WHEN
'Show me the knife you cut meat with when (it) fell!' (Juma: Elicit)

I tie the ungrammaticality of these examples to the adjunct clause: adjunct clauses are syntactic islands, and extraction from it results in ungrammaticality, as widely attested in relative clauses cross-linguistically.

Put together, the pattern of island effects and nonlocality exhibited by adnominal nominalizations strongly indicate that these structures involve extraction, a hallmark of relative clauses.

5.4.3 An argument for relativization from extraposition

The claim that adnominal nominalizations in Kawahíva involve relativization is further supported by extraposition, which involves the placement of a constituent in a right-peripheral position, as shown in (60) with an extraposed relative clause.

(60) A woman came in [__ who was wearing a blazer].

Specifically, I argue that a ban on extraposing an adnominal nominalization can be tied to its internal syntax of relativization. In short, the argument runs as follows: if an adnominal nominalization cannot be extraposed, it involves the syntax of a relative clause that does not admit extraposition, known as 'raising' relatives. I also show that the ban on extraposition does not stem from the nominalization as non-adnominal nominalizations can be extraposed. The overall conclusion is that adnominal nominalizations involve relativization. However, the same diagnostic has nothing to say about the derivation of adnominal nominalizations that are extraposable. I elaborate on the argument from extraposition in the remainder of this section and show how it applies to Kawahíva.

In theoretical linguistics, it has become more or less consensus that natural languages employ a few choices in forming a relative clause structure, including the 'raising' and 'non-raising' relatives (Bhatt 2002; Hulsey and Sauerland 2006); the latter is also called 'head-external' relatives. Glossing over non-relevant technical details, in a raising relative clause, the relative head originates within the relative clause and is then extracted to its surface position. This structure is schematized in (61), showing the relative head initially within the relative clause in (61a) but moved out from it in a later step of the derivation (61b). In contrast, in a non-raising relative or head-external relative, the relative head starts outside the relative clause, as in (62).[13]

(61) *The raising relative structure*

a. $_{\text{NP}}[\ [\ldots\ \textbf{[relative head]}\ \ldots]_{\textit{relative clause}}]$

b. $_{\text{NP}}[\ \textbf{[relative head]}\ [\ldots\ \textbf{[}\sout{\textbf{relative head}}\textbf{]}\ \ldots]_{\textit{relative clause}}]$

(62) *The non-raising, or head-external, relative structure*

$_{\text{NP}}[\ \textbf{[relative head]}\ [\ldots\ \ \ldots]_{\textit{relative clause}}]$

[13] One other type of relative clause is the so-called 'matching' relative (Chomsky 1965; Sauerland 2003). Matching relative clauses involve the structure of a 'raising' and 'non-raising' relative, either of which can be "forced" in the proper environment. Therefore, outside these proper environments, one expects matching relatives to allow extraposition of an adnominal nominalization that involves relativization. As the availability of extraposition of the adnominal nominalization is not informative about the presence of internal relativization, I set aside matching relatives from this discussion.

The crucial distinction between the two types of relatives concerns the constituency status of the relative head and relative clause: in 'raising' relatives, the relative head is initially embedded within the relative clause, forming a constituent with it (61). This configuration allows the relative head to be interpreted in its original position before being moved to the surface position. In contrast, non-raising relatives or head-external relatives have the relative head outside the relative clause from the start (62). This separation means that the relative head and the relative clause do not form a constituent at the surface structure.

The idea that there is a tight connection between the relative head and relative clause parts in one type of relative but not the other plays a key role in the account of various syntactic phenomena related to constituency, including extraposition, which I discuss now. Other phenomena include idiom interpretability and binding effects (see (Schenner 2019:for an overview)).

A growing body of crosslinguistic data has shown that relative clause extraposition is available for languages that exhibit non-raising relative clauses of (62) but not for languages with the raising relatives of (61). Languages that exhibit the 'raising' strategy include the Mixe-Zoquean language Nuntaɨiyi, according to Lopéz Márquez (2024). Consider the examples in (63).[14] In (63a), the relative clause is placed right after the relative head *je?m ʔitumin* 'her money', and the adverbial *joymɨ* 'tomorrow' occurs before the modified noun. In contrast, in (63b), the relative clause appears extraposed over the adverbial *joymɨ* 'tomorrow'. As a result, this structure is ungrammatical in Nuntaɨiyi.

(63) *Extraposition of relative clauses in Nuntaɨiyi is ungrammatical*

a. ʔan=nuʔm-pa joy=mɨ **je?m ʔi=tumin** [Maliyaj=pɨʔk ʔi=neʔk-wɨ
1.ERG=steal-ICP tomorrow=POST DEF 3.POSS=money Maliyaj=REL 3.ERG=hide-CPL
ʔaani-tek-pok=joj=mɨ].
tortilla-throw.gourd=RN:inside=POST

'I will steal tomorrow her money that Maliyaj put in the tortilla basket.'
(Lopéz Márquez 2024:10)

b. *ʔan=nuʔm-pa **je?m ʔi=tumin** joy=mɨ [Maliyaj=pɨʔk ʔi=neʔk-wɨ
1.ERG=steal-ICP DEF 3.POSS=money tomorrow=POST Maliyaj=REL 3.ERG=hide-CPL
ʔaani-tek-pok=joj=mɨ].
tortilla-throw.gourd=RN:inside=POST

'Intended: I will steal her money tomorrow that Maliyaj put in the tortilla basket.'
(Lopéz Márquez 2024:10)

In contrast, in languages like Spanish, where restrictive relative clauses are analyzed as employing a non-raising strategy (Eguren 2017), the extraposition of a relative clause is freely available. This is shown in the examples in (64). In (64a), the relative clause within brackets immediately follows the relative head *un señor* 'an elderly man'. In (64b), the same relative clause appears extraposed over the adverbial *ayer* 'yesterday', and the result is grammatical.

[14] More particularly, extraposition is barred with noun phrase relatives in the language but freely available with other phrases (e.g., adverbials). This distinction is absent in Kawahíva.

(64) *Extraposition of relative clauses in Spanish is freely available*

a. Vino **un señor** [que no sabía usar el teléfono].
I.saw a MASC.elder that not knows use the.MASC cellphone
'(I) saw an elderly man who did not know to use the cellphone.'
(Gutiérrez Bravo 2015:100)

b. Vino **un señor** ayer [que no sabía usar el teléfono].
I.saw a MASC.elder yesterday that not knows use the.MASC cellphone
'(I) saw an elderly man yesterday who did not know to use the cellphone.'
(Gutiérrez Bravo 2015:100)

As shown, the availability or absence of extraposition for relative clauses in these languages is accounted for by the type of relative clause. Extraposition is not available if the language has only 'raising' relatives (e.g., Nuntaɨiyi). However, if the language exhibits a 'non-raising' relative, extraposition is freely available (e.g., Spanish).

I propose this correlation involving relative clause extraposition can be used as a diagnostic of relativization in adnominal nominalizations. Specifically, if extraposing an adnominal nominalization is ungrammatical, then the nominalization involves relativization. I also discuss independent evidence that non-adnominal nominalizations are extraposable in Kawahíva. This confirms that their adnominal counterparts are banned from extraposition for another reason other than the nominalization status. I argue that they involve a 'raising' relative clause, which does not allow extraposition, as in Nuntaɨiyi (Lopéz Márquez 2024). However, it is important to note that this correlation is only useful in one direction: the ban on the extraposition of adnominal nominalizations can indicate that the nominalization has the internal syntax of relative clauses, but the ability to extrapose an adnominal nominalization is inconclusive as to whether it exhibits relativization. There are two possible explanations for the availability of extraposition for the adnominal nominalization: either the internal syntax of the nominalization involves a 'non-raising' strategy for relativization, allowing extraposition, or the adnominal nominalization does not involve relativization, and extraposition exists for other reasons, such as to satisfy a prosodic constraint (Féry 2015; Potsdam 2022).

In Kawahíva, extraposition of the adnominal nominalization yields ungrammaticality. Consider the examples in (65). (65a) is the non-extraposed version. The modified noun and the adnominal nominalization (an intransitive subject nominalization marked with *-va'e*) appear before the postposition *rehe* 'at'. In contrast, (65b) shows the nominalization extraposed over the postposition. The result is ungrammatical. Notably, the ungrammaticality of these cases is not due to a general ban on postposition stranding in the language. As (65c) shows, Kawahíva allows stranded postpositions (e.g., as the result of topicalization of their complement).

(65) a. T-a-mõpu **mytua** [__ o-veve-va'e=a] rehe.
OPT-1.SG.A-shoot curassow 3.COR-fly-WH.INTR.SUBJ=NMLZ at
'I wanted to shoot at the curassow that flew.' (Juma: Elicit)

b. *T-a-mõpu **mytua** rehe [__ o-veve-va'e=a].
OPT-1.SG.A-shoot curassow at 3.COR-fly-WH.INTR.SUBJ=NMLZ

'I wanted to shoot at the curassow that flew.' (Juma: Elicit)

c. Nh=ãpynha, koi'i pirahuva ehe.
1.SG=nose INEX.MANY acne at
'In my nose, there is much acne.' (Lit.: My nose, there is much acne at)

It may be that adnominal nominalizations are not extraposable in Kawahíva due to a specific ban on extraposition of nominalizations rather than the breakup of the modified noun and adnominal nominalization. We can rule out this alternative explanation for the ban on extraposition by looking at extraposition in complement clauses, which are also nominalized (as described in §3.7.9). Consider the example in (66), with a nominalized complement clause embedded under *kwaham* 'know, remember'. Notably, a nominalized complement clause can be extraposed over an adverbial, such as *ko'emamẽ* 'tomorrow'. This data supports the idea that adnominal nominalizations are not banned from extraposition due to nominalization. If they were, then nominalized complement clauses could not be extraposable, contrary to truth.[15,16]

(66) A-kwaham ko'emamẽ [Puré=ga i-mõpu=a taitetua rehe oi'i].
1.SG.A-remember morning P.=3.SG.MASC i-shoot=NMLZ peccary at SOME.TIME.AGO
'I remembered this morning that Puré shot at the peccary some time ago.' (Juma: Elicit)

Under the view that the adnominal nominalization involves relativization, non-extraposition of the nominalized relative clause is expected. The relativization strategy used is the 'raising' relative. In this strategy, the relative head and relative clauses form a constituent. Extraposing the relative clause would break apart this constituency and, therefore, ungrammaticality.[17]

5.5 Conclusion

In this chapter, I have provided the first investigation of relativization in Kawahíva. Descriptive studies of languages of the Tupí-Guaraní family have proposed that relativization is accomplished by using a nominalization in adnominal function, the 'nominalization strategy' (Thornes 2023). This strategy is widespread cross-linguistically, being common to languages of the Americas and Asia.

[15]One other piece of evidence that (66) follows from extraposition of the complement clause over the adverbial *ko'emamẽ* 'tomorrow' – rather than a process that dislocates the adverbial to its surface position – is adverbials appear in the clause-final position by default. When they appear in a different position, they are separated from the other parts of the sentence by a prosodic break, which is not seen in (66).

[16]Note that I demonstrate the acceptability of extraposition of the nominalized complement clause over an adverbial, rather than a postposition, as I did for adnominal nominalizations. This is so because Kawahíva does not have a verb that takes a clausal complement introduced by a postposition.

[17]Two other cross-linguistics diagnostics typically used to diagnose the relative clause formation strategy available in a language are NPI licensing and idioms. The former is also available in Kawahíva, and it also shows that the language uses a raising strategy, as the gap position in the adnominal nominalizations can be associated with an NPI outside the nominalization. The diagnostic from idioms cannot be applied to the only idiom I have identified in the language thus far, i.e., 'X person is standing up', which can mean 'X is dead'. The idiom requires using a term for a deceased person as the relative head, which is highly culturally sensitive.

I have argued that such structures in Kawahíva are not simply nominalizations in adnominal function. Additional clear evidence shows that they involve the process of relativization. On one count, the adnominal nominalizations are indeed clausal nominalizations as they exhibit nominal morphosyntax. Evidence shows they appear with the nominalizer *=a*, nominal temporal markers, and nominal negation. They are clausal-like in that they also show adverbs. The adnominal nominalizations also exhibit the syntax of relativization, including sensitivity to island effects, nonlocality, and cannot be extraposed. These facts strongly suggest that adnominal nominalizations are best understood as nominalized relative clauses.

It is possible that the additional process of relativization in adnominal nominalizations is a distinctive property of Kawahíva. However, I do not know of any studies on adnominal nominalizations in related languages that tried to rule out that they also involve relativization. I hope the present case study will inspire new research in this area.

Bibliography

Abrahamson. n.d. *Filler of the verb complex slots*.

Abrahamson, Arne. 1968. *Júma word list*.

Abrahamson, Arne/Joyce. 1980. Karipuna and Júma comparative study.

Adler, Jeff, Steven Foley, Jed Pizarro-Guevara, Kelsey Sasaki, and Maziar Toosarvandani. 2018. The derivation of verb initiality in Santiago Laxopa Zapotec. In *A reasonable way to proceed: Essays in honor of Jim McCloskey*, ed. J. Merchant, L. Mikkelsen, D. Rudin, and K. Sasaki, 31–49. Santa Cruz, Berkeley, and Chicaco: University of California's eScholarship Repository.

Aguilar, Ana Maria Gouveia Cavalcanti. 2015. Contribuições para os estudos histórico comparativos sobre a diversificação do sub-ramo VI da família linguística Tupí-Guaraní. Doctoral Dissertation, Universidade de Brasília, Brasília.

Aikhenvald, Alexandra Y. 2012. Possession and ownership: a cross-linguistic perspective. In *Possession and ownership*, ed. Alexandra Y. Aikhenvald and R. M. W. Dixon, 1–64. Oxford: Oxford Academic.

Aikhenvald, Alexandra Y. 2018. Evidentiality: the framework. In *The Oxford Handbook of Evidentiality*, ed. Alexandra Y. Aikhenvald, 1–44. Oxford: Oxford University Press.

Aissen, Judith. 2023. Documenting topic and focus. In *Key topics in language documentation and description*, ed. Peter Jenks and Michael Lev, 11–58. Honolulu: University of Hawai'i Press.

Alexiadou, Artemis. 2003. Some notes on the structure of alienable and inalienable possessors. from np to dp. In *The Expression of Possession in Noun phrases*, ed. Martine Coene and Yves d'Hulst, 167–188. Philadelphia: John Benjamins.

Alexiadou, Artemis, and Elena Anagnostopoulou. 1998. Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language & Linguistic Theory* 16:491–539.

Amondawa, Tambura, and João Paulo Marra Denófrio. 2015. *Objeto e Pessoa: Fragmentos etnográficos amondawa*. FUNAI-CGIIRC.

Andrews, Avery D. 2007. Relative clauses. In *Language typology and syntactic description*, ed. Timothy Shopen, volume 3, 206–236. Cambridge: Cambridge University Press.

Apontes, Selmo Azevedo. 2015. Descrição gramatical do Oro Waram, variante Wari' norte (Pakaa Nova, Txapakura): fonologia, morfologia e sintaxe [Grammatical description of Oro Waram, variety of Northern Wari' (Pakaa Nova, Txapakura): phonology, morphology and syntax. Doctoral Dissertation, Universidade Federal de Minas Gerais/Belo Horizonte.

Azanha, Gilberto. 2006. Relatório antropológico circunstanciado de identificação da Terra Indí-

gena Rio Pardo (Kawahíva do Mato Grosso) [Anthropological report for the identification of the Indigenous Land Kawahíva do Rio Pardo (Kawahíva from Mato Grosso)].

Baker, Mark. 1985. The mirror principle and morphosyntactic explanation. *Linguistic Inquiry* 16:373–415.

Baker, Mark, and Nadya Vinokurova. 2009. On agent nominalizations and why they are not like event nominalizations. *Language* 85:517–556.

Balykova, Kristina. 2019. Expressão de propriedades no Guató e Wa'ikhana [The expression of properties in Guató and Wa'ikhana]. Rio de Janeiro: Universidade Federal do Rio de Janeiro. Master's dissertation.

Baranger, Estefanía. 2022. On syntactic integration and semantico-pragmatic distribution of Mbya Guarani purpose coding strategies. *Languages* 7(52):1–20.

Bastos-Gee, Ana C. 2009. Topicalization of verbal projections in Brazilian Portuguese. In *Minimalist Essays on Brazilian Portuguese syntax*, ed. J. Nunes, 161–190. Amsterdam/Philadelphia: John Benjamins Publishing Company.

BBC. 2024. Elon musk domina internet por satélite na Amazônia com antenas em 90% das cidades [Elon Musk dominates satellite-internet service in the Amazon with antennas in 90% of towns]. *BBC News Portuguese* URL `https://www.bbc.com/portuguese/articles/cv2edkw84zmo`.

Betts, LaVera. 1967. *Transcriptions of Recordings of Parintintín Texts*. SIL International.

Betts, LaVera. 1981. *Dicionário Parintintin-Português, Português-Parintintin*. SIL.

Betts, LaVera. 2008[1969]. *Parintintin discourse*. Anápolis (Goiás): SIL.

Betts, LaVera. 2012. Kagwahiva dictionary. Available at `https://www.sil.org/resources/archives/74512`.

Bhatt, Rajesh. 2002. The raising analysis of relative clauses: evidence from adjectival modification. *Natural Language Semantics* 43–90.

Biberauer, Theresa. 2010. Semi null-subject languages, expletives, and expletive *pro* reconsidered. In *Parametric variation: Null subjects in Minimalist Theory*, ed. Theresa Biberauer, Anders Holmberg, Ian Roberts, and Michelle Sheehan, 35–67. Amsterdam: John Benjamins.

Birchall, Joshua. To appear. The Chapacuran language family. In *Amazonian Languages: An International Handbook*, ed. Patience Epps and Lev Michael. Berlin/Boston: De Gruyter Mouton.

Bittner, Maria, and Ken Hale. 1996. The structural determination of case and agreement. *Linguistic Inquiry* 27:219–251.

Bleaman, Isaac L. 2022. Predicate fronting in Yiddish and conditions on multiple copy Spell-Out. *NLLT* 393–424.

Borsley, Robert, and Andreas Kathol. 2000. Breton as a V2 language. *Linguistics* 38:665–710.

Bossi, Madeline, and Michael Diercks. 2019. V1 in Kipsigis: Head movement and discourse-based scrambling. *Glossa* 4(1):1–43.

Brattico, Pauli. 2022. Predicate clefing and long head movement in Finnish. *Linguistic Inquiry* 4:663–705.

Brugman, Claudia M., and Monica Macaulay. 2015. Characterizing evidentiality. *Linguistic Ty-*

pology 19:201–237.

Bruno, Ana Carla, and Simone do Carmo Gomes. 2010. *A arte plumária dos povos Kagwahíva da Transamazônica: etnias Tenharim, Parintintin e Diahoi*. Editora INPA.

Camargos, Quesler Fagundes. 2017. Aplicativização, causativização e nominalização: uma análise unificada de estruturas argumentais em Tenetehára-Guajajára (Família Tupí-Guaraní) [Applicativization, causativization, and nominalization: a unified analysis of argumental structures in Tenetehára-Guajajára (Família Tupí-Guaraní)]. Doctoral Dissertation, Universidade Federal de Minas Gerais (UFMG).

Camargos, Quesler Fagundes, and Ricardo Campos Castro. 2013. Paralelismo entre DP e CP a partir de nominalizações na língua Tenetehára. *Revista da Anpoll* 285–303.

Caponigro, Ivano. 2021. Introducing Headless Relative Clauses and the Findings from Mesoamerican Languages. In *Headless Relative Clauses in Mesoamerican Languages*, ed. Ivano; Caponigro, Harold Torrence, and Roberto Zavala Maldonado, 1–57. Oxford: Oxford University Press. URL https://doi.org/10.1093/oso/9780197518373.001.0001.

Carnie, Andrew, and Eithne Guilfoyle. 2000. *The syntax of verb initial languages*. New York: Oxford University Press.

Carnie, Andrew, Heidi Harley, and Sheila Ann Dooley. 2005. *Verb first: On the syntax of verb initial languages*. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Chafe, W. 1976. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In *Subject and topic: A new typology of language*, ed. Charles N Li, 25–55. New York: Academic Press.

Chafe, W. 1995. The realis-irrealis distinction in caddo, the northern iroquoian languages, and english. In *Modality in grammar and discourse*, ed. J. Bybee and S. Fleischman, 349–365. Amsterdam: John Benjamins.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1977. On *wh*-movement. In *Formal syntax*, ed. Thomas Wasow Peter W. Culicover and Adrian Akmajian, 71–131. New York: Academic Press.

Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In *Ken Hale. A life in language*, ed. Michael Kenstowicz, 1–52. Cambridge: MIT Press.

Chomsky, Noam, and Howard Lasnik. 1977. Filters and control. *Linguistic Inquiry* 8:425–504.

Chung, Sandra. 1994. Wh-agreement and "referentiality" in Chamorro. *Linguistic Inquiry* 25 (1):1–44.

Chung, Sandra. 2005. What fronts? On the VP-raising account of verb-initial order. In *Verb First: On the syntax of verb initial languages*, ed. A. Carnie, H. Harley, and S. A Dooley, 9–30. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Chung, Sandra, and Carol Georgopoulos. 1988. Agreement with Gaps in Chamorro and Palauan. In *Agreement in Natural Language: Approaches, Theories, Descriptions*, ed. Charles A. Barlow, Michael; Ferguson, 109–137. Center for the Study of Language and Information.

Clemens, Laura Eby, and Maria Polinsky. 2017. Verb-initial word orders, primarily in Austrone-

sian and Mayan languages. In *The Wiley Blackwell Companion to Syntax (Second Edition)*, ed. Martin Everaert and Henk C. van Riemsdijk. John Wiley and Sons, Inc.

Clemens, Lauren. 2019. Prosodic noun incorporation: The relationship between prosody and argument structure in Niuean. *Syntax* 22(4):337–377.

Clemens, Lauren, and Jessica Coon. 2018. Deriving verb-initial word order in Mayan. *Language* 94:237–280.

Cole, Peter, and Gabriella Hermon. 2011. Nominalization and case assignment in Quechua. *Lingua* 1225–1251.

Collins, James N. 2017. Samoan predicate initial word order and object positions. *Natural Language and Linguistic Theory* 35:1–59.

Comrie, Bernard, and Zarina Estrada-Fernández. 2012. *Relative clauses in languages of the Americas: A typological overview*. Amsterdam/Philadelphia: John Benjamins.

Comrie, Bernard, and Sandra A. Thompson. 1985. Lexical nominalization. In *Language Typology and Syntactic Description, Vol. III: Grammatical Categories and the Lexicon*, ed. Timothy Shopen, 349–398. Cambridge University Press.

Croft, William. 2001. *Radical construction grammar: Syntactic theory in typological perspective*. Oxford: Oxford University Press.

Cruz, Aline da. 2016. Towards an understanding of the origin of aspectual marks on nouns: evidence from Nheengatu and Tupinambá. In *O sintagma nominal em línguas amazônicas [The noun phrase in Amazonian languages]*, ed. Francesc Queixalós and Dioney M. Gomes, 45–73. Campinas: Pontes.

Crystal, David. 1980. *A dictionary of linguistics and phonetics*. Oxford: Wiley-Blackwell.

Dahl, Östen, and Velupilai Viveka. 2011. The past tense. In *The world atlas of language structures online*, ed. Matthew S. Dryer and Martin Haspelmath. Munich: Max Planck Digital Library. Available online at `http://wals.info/chapter/66`, Accessed on 2024-06-15.

Danielsen, Swintha. 2007. *Baure: An Arawak language of Bolivia*, volume 6 of *Indigenous Languages of Latin America*. Leiden: Research School of Asian, African and Amerindian Studies (CNWS), Universiteit Leiden. PhD 2007 Radboud Universiteit Nijmegen.

Dawson, Virginia, and Peter Jenks. 2023. Describing definites and indefinites. In *Key topics in language documentation and description*, ed. Peter Jenks and Lev Michael, volume 26, 99–174. Honolulu: University of Hawai'i Press.

Deal, Amy Rose. 2016. Syntactic ergativity: Analysis and identification. *Annual Review of Linguistics* 2:165–185.

DeLancey, Scott. 2002. Relativization and nominalization in bodic. In *Proceedings of the twenty-eighth annual meeting of the Berkeley Linguistics Society: Special session on Tibeto-Burman and southeast Asian linguistics*, 55–72. Berkeley, CA: Cascadilla Proceedings Project.

Denófrio, João Paulo Marra. 2012. *Breve contribuição etnográfica sobre os Kagwahiva – o coletivo Piripkura. Relatório – Etnologia e Antropologia Social*. FUNAI-CGIIRC.

Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

Doner, Julianne. 2017. Predicate-Sensitive EPP. In *Proceedings of the 2017 annual conference of the*

Canadian Linguistic Association, 1–16.

Doner, Julianne. 2019. The EPP across Languages. Doctoral Dissertation, University of Toronto.

Dos Santos, Wesley. 2017. Inventário Nacional de Diversidade Linguística: Kawahíva [National Survey of Linguistic Diversity: Kawahíva]. Official unpublished report on the Kawahíva varieties in Rondônia and their vitality.

Dos Santos, Wesley. 2019. Kawahiva: do diagnóstico de vulnerabilidade à documentação [kawahiva: from the vulnerability assessment to the documentation]. *Revista Brasileira de Línguas Indígenas - RBLI* 2(1):105–120. Available at https://core.ac.uk/download/pdf/327113794.pdf.

Dos Santos, Wesley. 2021a. Alignment, agreement, and person hierarchy in Kawahíva. Online talk given at Amazonicas VIII. Recording available at https://www.youtube.com/watch?v=NIuzhl6bAHI&feature=youtu.be.

Dos Santos, Wesley. 2021b. For a realis and irrealis account of *ko* and *po* in Kawahíva. Online talk given at the Anual Meeting of the Society for the Study of the Indigenous Languages of the Americas (SSILA).

Dos Santos, Wesley. 2022. Diagnosing unaccusativity in Kawahíva. In *Proceedings of the 2022 Annual Meeting of The Linguistic Society of America*, 1–8. Linguistic Society of America (LSA). DOI: https://doi.org/10.3765/plsa.v7i1.5262.

Dos Santos, Wesley. 2023a. Asymmetries among person indexes in Kawahíva. *International Journal of American Linguistics* 89(4):427–455. DOI: https://doi.org/10.1086/726149.

Dos Santos, Wesley. 2023b. A nominalizer acccount of *=a* in Kawahíva (Tupí-Guaraní). Talk given at the 2023 Annual Meeting of The Linguistic Society of America.

Dos Santos, Wesley. accepted. Kawahíva (Brazil) - Language Snapshot. *Language Documentation and Description* .

Dos Santos, Wesley. In prep. Morphological reflexes of relativization in Kawahíva.

Drude, Sebastian. 2021. Alterações 'morfo-fonológicas' em Awetí: em defesa de fonemas abstratos [Morphophonological alternations in Awetí: in defense of abstract phonemes]. Talk presented at the Amazonicas VIII, available at https://www.youtube.com/watch?app=desktop&v=RoWwGXq23rw.

Dryer, Matthew. 2005. Order of subject, object and verb. In *World Atlas of Language Structures*, ed. Matthew; Dryer, Martin; Haspelmath, David; Gil, and Bernard Comrie, 123–152. Oxford: Oxford University Press.

Duarte, Fábio Bonfim. 2007. *Estudos de Morfossintaxe Tenetehára [Morphosyntactic Studies on Tenetehára]*. Belo Horizonte: Faculdade de Letras da UFMG.

Duarte, Fábio Bonfim. 2012. Tenetehára: A predicate-fronting language. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 57:359–386.

Duff-Tripp, Martha. 1997. *Gramática del idioma yanesha' (amuesha)*, volume 43 of *Serie Lingüística Peruana*. Lima: Lima: Ministerio de Educación and Instituto Lingüístico de Verano. URL http://www.sil.org/acpub/repository/slp43.pdf,http://

www.sil.org/acpub/repository/slp43-sample.pdf.

Dunn, Michael. 2014. Gender determined dialect variation. In *The expression of gender*, ed. Greville Corbett, 39–68. Berlin: De Gruyter.

Dékány, Éva. 2018. Approaches to head movement: A critical assessment. *Glossa* 1:1–43.

Eberhardt, Roy. 1999. Questions and Inversion in Ocotepec Mixtec. In *Work Papers of the Summer Institute of Linguistics*, volume 43, 1–21. North Dakota: Work Papers of the Summer Institute of Linguistics.

Eguren, Luis. 2017. Possessives and relative clauses in Spanish. *Probus* 41–72.

Elliott, Jennifer R. 2000. Realis and irrealis: Forms of grammaticalisation of reality. *Linguistic Typology* 5:55–90.

Epps, Patience, and Lev Michael. 2023. Introduction: The languages of Amazonia. In *Volume 1 Language Isolates I: Aikanã to Kandozi-Shapra: An International Handbook*, ed. Patience Epps and Lev Michael, xvii–lxii. Berlin/Boston: De Gruyter Mouton.

Ernst, Thomas. 2020. The Syntax of Adverbials. *Annual Review of Linguistics* 6:89–109.

Estigarribia, Bruno. 2020. *A grammar of Paraguayan-Guarani*. London: UCL Press.

Everett, Daniel L., and Barbara Kern. 1997. *Wari': The Pacaas Novos language of Western Brazil*. Descriptive Grammars Series. London/New York: Routledge.

Fox, Danny;, and David Pesetsky. 2005. Cyclic Linearization of Syntactic Structure. *Theoretical Linguistics* 31:1–45.

França, Luciana Barroso Costa. 2012. Caminhos cruzados: parentesco, diferença e movimento entre os Kagwahiva. Doctoral Dissertation, Universidade Federal do Rio de Janeiro.

Fukui, Naoki, and Yuji Takano. 1998. Symmetry in syntax: Merge and demerge. *Journal of East Asian Linguistics* 1:27–86.

Féry, Caroline. 2015. Extraposition and Prosodic Monsters in German. In *Explicit and implicit prosody in sentence processing. studies in theoretical psycholinguistics*, ed. Lyn Frazier and Edward Gibson. Springer.

Galucio, Ana Vilacy. 2001. The Morphosyntax of Mekens. Doctoral Dissertation, University of Chicago.

Galucio, Ana Vilacy. 2002. O prefixo i- em tupi: morfema antipassivo vs. marcador pronominal incorporado. In *Línguas indígenas brasileiras : Fonologia, gramática e história: Atas do i encontro internacional do grupo de trabalho sobre línguas indígenas da anpoll*, ed. Ana Suely A. C. Cabral and Aryon dall'Igna Rodrigues, 274–287. Belém: EDUFPA.

Galucio, Ana Vilacy, Denny Moore, and Hein Voort. 2018. O patrimônio linguístico do Brasil: Novas perspectivas e abordagens no planejamento e gestão de uma política da diversidade linguística. *Revista do Patrimônio Histórico e Artístico Nacional* 38:194–219. Available at: Galucio_Moore_van_der_voort_2018_O_patrimonio_linguistico_do_Brasil.pdf.

Galúcio, Ana Vilacy. 2006. Relativização na língua Sakurabiat (Mekens) [Relativization in Sakurabiat (Mekens)]. *Boletim do Museu Paraense Emílio Goeldi* 49–59.

Galúcio, Ana Vilacy. 2011. Nominalization in the Mekens language. *Amerindia* 237–260.

Genetti, Carol. 2011. Nominalization in Tibeto-Burman languages of the Himalayan area: A typological perspective. In *Nominalization in Asian languages: Diachronic and typological perspectives (typological studies in language*, ed. Foong Ha Yap, Karen Grunow-Hårsta, and Janick Wrona, 163–193. Amsterdam/Philadelphia: John Benjamins.

Gijn, Rik van, Ana Vilacy Galucio, and Antonia Fernanda Nogueira. 2015. Subordination strategies in Tupian languages. *Boletim do Museu Paraense Emílio Goeldi* 10:297–324.

Gildea, Spike. 1993. The development of tense markers from demonstrative pronouns in Panare (Cariban). *Studies in Language* 17:53–74.

Giorgi, Alessandra, and Giuseppe Longobardi. 1991. *The syntax of noun phrases.* Cambridge: Cambridge University Press.

Goldberg, Adele E. 2006. *Constructions at work: the nature of generalization in language.* New York: Oxford University Press.

Grijzenhout, Janet. 2011. *Consonant mutation*, chapter 65, 1–22. John Wiley Sons, Ltd. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9781444335262.wbctp0065.

Guillaume, Antoine. 2012. Maropa (Reyesano). In *Ambito Andino*, ed. Mily Crevels and Pieter Muysken, volume 2 of *Lenguas de Bolivia*, 191–232. La Paz: Plural Editores.

Gutiérrez Bravo, Rodrigo. 2015. *Las cláusulas relativas en maya yucateco.* Mexico: El Colegio de México.

Hagège, Claude. 2010. *Adpositions.* Oxford: Oxford University Press.

Hammerly, Christopher. 2021. A verb raising analysis of the Ojibwe VOS/VSO alternation: Lessons for feature copying and movement. *https://ling.auf.net/lingbuzz/004786* .

Hannahs, S. J. 2011. *Celtic mutations*, chapter 117, 1–24. John Wiley Sons, Ltd. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/9781444335262.wbctp0117.

Harbour, Daniel. 1999. The two types of predicate clefts: Classical Hebrew and beyond. In *Papers on morphology and syntax, cycle two*, ed. Benjamin Bruening Vivian Lin, Cornelia Krause and Karlos Arregi, 159–175. MIT Working Papers in Linguistics.

Harizanov, Boris. 2019. Head movement to specifier positions. *Glossa* 4(1):1–36.

Harizanov, Boris, and Vera Gribanova. 2019. Whither head movement? *Natural Language and Linguistic Theory* 37:461–522.

Harley, Heidi. 2013. External arguments and the Mirror Principle: On the distinctness of voice and v. *Lingua* 125:34–57.

Harley, Heidi, and Elisabeth Ritter. 2002. Person and number in pronouns: A feature-geometric analysis. *Language* 78(3):482–526.

Harrison, Carl H. 1986. Verb prominence, verb initialness, ergativity and typological disharmony in Guajajara. In *Handbook of Amazonian Languages*, ed. Desmond C. Derbyshire and Geoffrey K. Pullum, volume 1, 407–439. Mouton de Gruyter.

Hein, Johannes. 2017. Deriving the typology of verbal fronting. In *Proceedings of the Forty-Eight Annual Meeting of the North East Linguistics Society*, ed. Max Nelson and Sherry Hucklebridge,

volume 2, 29–38. Amherst: GLSA Amherst.

Hein, Johannes. 2018. Verbal fronting: Typology and Theory. Doctoral Dissertation, Leipzig University, Leipzig.

Holmberg, Anders. 1986. Word Order and Syntactic Features. Doctoral Dissertation, Stockholm University, Stockholm.

Holmberg, Anders. 1999. Remarks on Holmberg's Generalization. *Studia Linguistica* 53:1–39.

Hsieh, Henrison. 2019. Distinguishing nouns and verbs: A Tagalog case study. *NLLT* 523–569.

Hulsey, Sarah, and Uli Sauerland. 2006. Sorting out relative clauses. *Natural Language Semantics* 14(2):111–137.

Jensen, Cheryl. 1990. Cross-referencing changes in some Tupí-Guaraní languages. In *Amazonian Linguistics: Studies in Lowland South American Languages*, ed. Doris L. Payne, 117–158. Austin: University of Texas Press.

Jensen, Cheryl. 1998. Comparative Tupí-Guaraní morphosyntax. In *Handbook of Amazonian languages*, ed. Desmond C. Derbyshire and Geoffrey K. Pullum, volume 4, 489–618. Berlin: Mouton.

Jensen, Cheryl. 1999. Tupi-Guarani. In *The Amazonian Languages*, ed. R. M. W. Dixon and Alexandra Aikhenvald, 125–164. Cambridge: Cambridge University Press.

Jouitteau, Mélanie. 2005. Nominal properties of *v*Ps in Breton, a hypothesis for the typology of VSO languages. In *Verb first: On the syntax of verb initial languages*, ed. H; Dooley S. A. Carnie, A; Harley, 265–280. John Benjamins Publishing Company.

Julião, Maria Risolêta Silva. 2005. Aspects Morphosyntaxiques de L'Anambe. Doctoral Dissertation, L'Universite de Toulouse (Le Mirail), Toulouse.

Kastovsky, Dieter. 1985. Deverbal nouns in Old and Modern English: From stem-formation to word-formation. In *Historical Semantics Historical Word-formation*, ed. Jacek Fisiak, 221–261. Berlin: Mouton.

Keenan, Edward. 1985. Relative clauses. In *Language Typology and Syntactic Description*, ed. Timothy Shopen, 141–170. Cambridge: Cambridge University Press.

Keine, Stephan. 2010. *Case and agreement from fringe to core: A minimalist approach*, volume 536 of Linguistische Arbeiten. Germany: De Gruyter.

Klaus-Peter, Dresden Kästner. 2005. Beitrage zur Geschichte und Kultur der Urueuwauwau-Stamme (Rondonia, Brasilien) [Contributions to the history and culture of the Urueuwauwau tribes (Rondonia, Brazil)]. In *Abhandlungen und Berichte der Staatlichen Ethnographischen Sammlungen Sachsen*, ed. Claus Deimel, 91–133. VWB - Verlag fur Wissenschaft und Bildung.

Kornfilt, Jaklin, and John Whitman. 2011. Afterword: Nominalizations in syntactic theory. *Lingua* 1297–1313.

Korsah, Sampson, and Andrew Murphy. 2020. Tonal reflexes of movement in Asante Twi. *NLLT* 38:827–885.

Kracke, Waud H. 1973. Emotions and personality in Parintintin politics: a study of political structure and commotive leadership in an Amazon Indian tribe. Doctoral Dissertation, Universidade of Chicaco, Chicago (Illinois).

Kracke, Waud H. 1981. Don't let the piranha bite your liver. In *Working papers on South American Indians: Food taboos in Lowland South America*, ed. Kenneth M. Kesinger and Waud H. Kracke, 91–142. Bennington College.

Kracke, Waud H. 2007. A posição histórica dos parintintín na evolução das culturas Tupí-Guarani. In *Línguas e culturas Tupí*, ed. Aryon D. Rodrigues and Ana S. A. C. Cabral, 23–35. Editora Curt Nimuendajú.

Krasnoukhova, Olga. 2016. Is nominalization the main strategy for attributive (relative) clause formation? In *South American indigenous language structures (SAILS) online*, ed. Pieter Muysken, Harald Hammarström, Olga Krasnoukhova, Neele Müller, Joshua Birchall, Simon van de Kerke, Loretta O'Connor, Swintha Danielsen, Rik van Gijn, and George Saad. John Benjamins. https://sails.clld.org/parameters/NP10#5/1.746/289.565.

Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase structure and the lexicon*, ed. L. Rooryck, J.; Zaring, volume 33, 109–137. Springer.

Kroeger, Paul R. 2007. Filler–gap dependencies and relativization. In *Analyzing syntax: A lexical-functional approach*, ed. Paul R. Kroeger, 165–191. Cambridge University Press.

Lambrecht, Knud. 1994. *Information structure and sentence form: topic, focus, and the mental representations of discourse referents.* Cambridge: Cambridge University Press.

Landau, Ian. 2006. Chain resolution in Hebrew V(P) fronting. *Syntax* 9:32–66.

Lapierre, Myriam. 2021. Towards a theory of subsegmental and subfeatural representations: The phonology and typology of nasality. Doctoral Dissertation, University of California, Berkeley, Berkeley, CA.

Lapierre, Myriam, and Lev Michael. 2017. Nasal segments in Tupi-Guaraní: A comparative synthesis. Paper presented at the 8th Conference on Indigenous Languages of Latin America (CILLA VIII), Austin, Texas, USA, October 26–28.

Lechner, Winfried. 2006. An interpretive effect of head movement. In *Phases of Interpretation*, ed. Mara Frascarelli, 45–70. Berlin: Mouton.

Lee, Felicia. 2000. VP remnant movement and VSO in Quiavini Zapotec. In *The syntax of verb initial languages*, ed. Andrew Carnie and Eithne Guilfoyle, Oxford Studies in Comparative Syntax, 143–162. Oxford: Oxford University Press.

Lee, Felicia. 2005. Force first: Clause-fronting and clause typing in San Lucas Quiaviní Zapotec. In *Verb First: On the syntax of verb initial languages*, ed. A. Carnie, H. Harley, and S. A Dooley, 91–106. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Lee, Felicia. 2006. *Remnant raising and VSO clausal architecture.* Dordrecht: Springer.

Lee, Tommy Tsz-Ming. Ms. Movement of quantificational heads and its scope effects. *Manuscript* 1–45.

Lema, José, and María-Luisa Rivero. 1990. Long Head Movement: ECP vs. HMC. In *Proceedings of the North East Linguistics Society (20)*, ed. N/A, 1–16. Cornell University, Ithaca, NY: North East Linguistics Society (NELS).

Leonel, Mauro. 1995. *Etnodiceia uréu-au-au.* São Paulo: Edusp.

Lochbihler, Bethany, and Eric Mathieu. 2013. Wh-agreement in Ojibwe relative clauses: Evidence

for CP structure. *Canadian Journal of Linguistics/Revue canadienne de linguistique* 58:293–318.
Lopéz Márquez, Wendy. 2024. The structure of Nuntaɨiyi DPs and non-DP relative clauses. Talk given at Syntax and Semantics Circle at UC Berkeley.
Lowe, Ivan. 1960. *Questionnaire of the standard vocabularies for preliminary comparative studies on the Brazilian Indigenous languages - Parintintins language.*
Löbner, Sebastian. 2002. *Understanding semantics.* London: Arnold Publishers.
Macaulay, Monica. 2005. The syntax of Chalcatongo Mixtec: Preverbal and postverbal. In *Verb First: On the syntax of verb initial languages*, ed. A. Carnie, H. Harley, and S. A Dooley, 341–366. Amsterdam/Philadelphia: John Benjamins Publishing Company.
Magalhães, Marina Maria Silva. 2007. Sobre a morfologia e a sintaxe da língua Guajá (família Tupí-Guaraní). Doctoral Dissertation, Universidade Federal de Brasília, Brasília.
Martins, Marci Fileti. 2003. Descrição e análise de aspectos da gramática do Guarani Mbyá. Doctoral Dissertation, Universidade Estadual de Campinas, Campinas.
Martins, Silvana, and Valteir Martins. 1999. Makú. In *The Amazonian Languages. Cambridge Language Surveys*, ed. Alexandra Y. Dixon, R. M. W.; Aikhenvald, 251–268. Cambridge: Cambridge University Press.
Marçoli, Osmar. 2018. Estudo comparativo dos dialetos da língua Kawahib (Tupi-Guarani): Tenharim, Jiahui e Amondawa. Master's dissertation.
Massam, Diane. 2000. VSO versus VOS: Aspects of Niuean word order. In *The syntax of verb initial languages*, ed. Andrew Carnie and Eithne Guilfoyle, Oxford Studies in Comparative Syntax, 97–116. Oxford: Oxford University Press.
Massam, Diane. 2001. Pseudo Noun Incorportation in Nieuan. *Natural Language and Linguistics Theory* 19:153–197.
Massam, Diane. 2020. *Niuean predicates and arguments in an isolating language.* Oxford: Oxford University Press.
Massam, Diane, and Carolyn Smallwood. 1997. Essential Features of Predication in English and Niuean. In *North East Linguistics Society*, ed. K. Kusumoto, 263–272. UMass: ScholarWorks@UMass Amherst.
Matisoff, James A. 1972. Lahu nominalization, relativization, and genetivization. In *Syntax and semantics volume 1*, ed. J. P. Kimball, 237–257. New York/London: Seminar Press.
Matushansky, Ora. 2006. Head movement in linguistic theory. *Linguistic Inquiry* 1:27–86.
McCloskey, James. 2006. On the scope of verb movement in Irish. *Natural Language and Linguistic Theory* 14:47–104.
McCloskey, James. 2017. Ellipsis, polarity, and the cartography of verb-initial orders in Irish. In *Elements of comparative syntax*, ed. Enoch Aboh, Eric Haeberli, Genoveva Puskás, and Manuela Schönenberger, volume 127 in the series Studies in Generative Grammar [SGG]. John Wiley and Sons, Inc.
Medeiros, David J. 2013. Hawaiian VP-remnant movement: A cyclic linearization approach. *Lingua* 127:72–97.
Meira, Sérgio. 2006. Mental state postpositions in Tiriyó and other Cariban languages. *Linguistic*

Typology 8:213–241.

Meira, Sérgio, and Sebastian Drude. 2013. Sobre a origem histórica dos "prefixos relacionais" das línguas Tupí-Guaraní. *Cadernos de Etnolingüística* 5(1):1–30.

Menendéz, Miguel A. 1989. Os Kawahiwa: uma contribuição para o estudo dos Tupí Centrais. Doctoral Dissertation, Universidade de São Paulo.

Michael, Lev. 2014. The Nanti reality status system: Implication for the typological validity of the realis/irrealis contrast. *Linguistic Typology* 18(5):251–288.

Michael, Lev, Natalia Chousou-Polydouri, Keith Bartolomei, Erin Donnelly, Vivian Wauters, Sérgio Meira, and Zachary O'Hagan. 2015. A Bayesian Phylogenetic Classification of Tupí-Guaraní. *Liames* 2:193–221.

Michael, Lev, Stephanie Farmer, Gregory Finley, Karina S. Acosta, Christine Beier, Alexandra C. Icahuate, Donalia I. Baneo, and Melchor S. Saita. 2023. Muniche. In *Amazonian Languages: Language Isolates ii: Kanoé to Yurakaré*, ed. Patience Epps and Lev Michael, 851–892. Berlin/Boston: De Gruyter Mouton.

Miranda, Camille Cardoso. 2018. Estudo comparativo do fenômeno de nasalização em línguas da família Tupí-Guaraní (Tronco Tupí). Doctoral Dissertation, Universidade Federal do Pará, Belém/Pará.

Mithun, Marianne. 1995. On the relativity of irreality. In *Modality in Grammar and Discourse*, ed. Joan L. Bybee and Suzanne Fleischman, 367–387. Amsterdam: John Benjamins.

Moore, Denny. 1989. Gavião nominalizations as relative clause and sentential complement equivalents. *International Journal of American Linguistics* 309–325.

Moore, Denny. 2006. Cláusulas relativas em Gavião de Rondônia [Relative clauses in Gavião of Rondônia]. *Boletim do Museu Paraense Emílio Goeldi* 135–143.

Moore, Denny. 2012. Relative clauses in Gavião of Rondônia. In *Relative clauses in languages of the americas: a typological overview*, ed. Bernard Comrie and Zarina Estrada-Fernández, 243–252. John Benjamins.

Moore, Denny, Vilacy Galucio, and Nilson Gabas Jr. 2008. O desafio de documentar e preservar as línguas amazônicas. *Scientific American (Brasil)* 3:36–43.

Muriungi, Peter K. 2008. Phrasal movement inside Bantu verbs. deriving affix scope and order in Kîîtharaka. Doctoral Dissertation, University of Tromsø.

Netto, Waldemar F., and Jorge M Moraes. 1993. Descrição fonológica preliminar da língua dos uru eu wau wau: subgrupo mondáwa.

Nimuendajú, Curt. 1924. Os índios Parintintin do rio Madeira. *Journal de la société des Américanistes* 16:201–278.

Nimuendajú, Curt. 1948. The Cawahíb, Parintintin, and their neighbors. In *Handbook of South American Indians*, ed. Julian H. Steward, 283–297. US Government Printing Office.

Nimuendajú, Curt. 1981. Mapa etno-histórico do Brasil e regiões adjacentes. Instituto Brasileiro de Geografia e Estatística.

Noonan, Michael. 2008. Nominalizations in Bodic languages. In *Rethinking grammaticalization: New perspectives*, ed. Maria José López-Couso and Elena Seoane, 219–237. Berlin: John Ben-

jamins.

Nordlinger, Rachel, and Louisa Sadler. 2004. Nominal tense in crosslinguistic perspective. *Language* 80(4):325–331.

O'Hagan, Zachary. 2020. Focus in Caquinte. Doctoral Dissertation, University of California, Berkeley, Berkeley, CA.

O'Hagan, Zachary. 2023. Taushiro. In *Amazonian Languages: Language Isolates ii: Kanoé to Yurakaré*, ed. Patience Epps and Lev Michael, 995–1028. Berlin/Boston: De Gruyter Mouton.

Ostrove, Jason. 2020. When φ-agreement targets topics: The view from San Martín Peras Mixtec. Doctoral Dissertation, University of California, Santa Cruz, Santa Cruz, CA.

Ouhalla, Jamal. 1994. Verb movement and word order in Arabic. In *Verb movement*, ed. David Lightfoot and Norbert Hornstein, 41–72. Cambridge: Cambridge University Press.

Overall, Simon. 2017. A Typology of Frustrative Marking in Amazonian Languages. In *The Cambridge Handbook of Linguistic Typology*, ed. Aikhenvald A. Y. and Dixon R. M. W., 477–512. Cambridge: Cambridge University Press.

Pachêco, Frantomé Bezerra. 2001. Morfossintaxe do verbo Ikpeng (Karíb) [The morphosyntax of the Ikpeng verb (Karíb)]. Doctoral Dissertation, Universidade Estadual de Campinas, Campinas.

Padua, Alexandre J. 2007. Contribuição para a fonologia da língua Apiaká (Tupí-Guaraní). Doctoral Dissertation, Universidade de Brasília.

Palácio, Adair Pimentel. 1984. Guató: A língua dos índios canoeiros do Rio Paraguai [Guató: The language of the canoeiros of Paraguay River]. Doctoral Dissertation, Universidade Estadual de Campinas, Campinas.

Parintintin. 1940. *Orekwatijar*. Rio de Janeiro: Summer Institute of Linguistics.

Parintintin. 1966. *Morogĩta: lendas Parintintin*. Rio de Janeiro: Summer Institute of Linguistics.

Payne, D. L. 1990. *The pragmatics of word order: Typological dimensions of verb initial languages. empirical approaches to language typology*. Berlin/New York: De Gruyter Mouton.

Pearson, Matthew. 2001. Clause structure of malagasy: a minimalist approach. Doctoral Dissertation, University of California, Los Angeles, Los Angeles, CA.

Pease, Helen. 1968. Parintintin Grammar. Manuscript: `https://doi.org/10.3765/plsa.v7i1.5262`.

Pease, Helen. 1973. Parintintin time sets, *po*, and *ko*.

Pease, Helen. 2009[1977]. *Juma-parintintín similarities*. SIL.

Pease, Helen, and LaVera Betts. 1962. Questionnaire of the standard vocabularies for preliminary comparative studies on the Brazilian Indigenous languages - Parintintin language.

Pease, Helen, and LaVera Betts. 1971. Parintintin phonology. In *Tupi Studies I*, ed. David Bendor-Samuel, 1–14. Norman, Oklahoma: Summer Institute of Linguistics.

Pease, Helen, and LaVera Betts. 1991a. *Comments on Uru-Eu-Wau-Wau*.

Pease, Helen, and LaVera Betts. 1991b. *Questionnaire of the standard vocabularies for preliminary comparative studies on the Brazilian Indigenous languages - Amundava language*.

Pease, Helen, and LaVera Betts. 1991c. *Questionnaire of the standard vocabularies for preliminary*

comparative studies on the Brazilian Indigenous languages - Uru-Eu-Wau-Wau language.

Pease, Helen, and LaVera Betts. 1991d. *Textos de um Amundava.* SIL.

Peggion, Edmundo Antonio. 1996. Forma e função: uma etnografia do sistema de parentesco Tenharim (Kagwahí, AM). Doctoral Dissertation, Universidade Estadual de Campinas.

Peggion, Edmundo Antonio. 2005. Relações em perpétuo desequilíbrio: a organização dualista dos povos Kagwahiva da Amazônia. Doctoral Dissertation, Universidade de São Paulo (USP).

Peggion, Edmundo Antonio. 2007. A classe dos nomes: a onomástica Amondáwa (Kagwahiva/RO). In *Línguas e culturas Tupí*, ed. Aryon D. Rodrigues and Ana S. A. C. Cabral, 123–130. Editora Curt Nimuendajú.

Pereira, Antônia Alves. 2009. Estudo morfossintático do Asurini do Xingu. Doctoral Dissertation, Universidade Estadual de Campinas, Campinas.

Pereira, Antônia Alves. 2010. A nominalização e as orações relativas no Asurini do Xingu (Tupi-Guarani). *Liames* 101–113.

Pereira, Nunes. 1940. *Bahira e suas experiências: etnologia amazonica.* Belém (Pará): Albano H. Martins and Ca.

Perlmutter, David M. 1978. Impersonal passives and unaccusative hypothesis. In *Proceedings of the Annual Meeting of the Berkeley Linguistics Society 38*, 157–189. Berkeley Linguistic Society (BLS)/Linguistic Society of America (LSA).

Potsdam, Eric. 2022. Malagasy extraposition: Evidence for PF movement. *Natural Language and Linguistic Theory* 195–237.

Praça, Walkíria Neiva. 2007. Morfossintaxe da língua Tapirapé [The Morphosyntax of Tapirapé]. Doctoral Dissertation, Universidade de Brasília (UnB), Brasília.

Praça, Walkíria Neiva, Marina Maria Silva Magalhães, and Aline Cruz. 2017. Indicativo II da família Tupí-Guaraní: uma questão de modo? *Liames* 17(1):39–58.

Preminger, Omer. 2011. Agreement as a Fallible Operation. Doctoral Dissertation, MIT, Cambridge, MA.

Preminger, Omer. 2019. What the PCC tells us about "abstract" agreement, head movement, and locality. *Glossa: a journal of general linguistics* 4(1):1–42.

Pylkkänen, Liina. 2002. *Introducing arguments.* Cambridge, MA: MIT Press.

Queiroz, José Márcio Correia de. 2012. Um estudo gramatical da língua Dzubukuá, família Karirí [A grammatical study of the Dzubukuá language, Karirí family]. Doctoral Dissertation, Universidade Federal da Paraíba, João Pessoa.

Queixalós, Francesc. 2006. The primacy and fate of predicativity in Tupi-Guarani. In *Lexical categories and root classes in Ameridian languages*, ed. Valentina Vapnarsky and Ximena Louis, 249–288. Bern: Peter Lang.

Rackowski, Andrea, and Norvin Richards. 2005. Phase Edge and Extraction: A Tagalog Case Study. *Linguistic Inquiry* 36:565–599.

Rackowski, Andrea, and Lisa Travis. 2000. V-initial languages: X or XP movement and adverbial placement. In *The syntax of verb initial languages*, ed. Andrew Carnie and Eithne Guilfoyle, Oxford Studies in Comparative Syntax, 117–142. Oxford: Oxford University Press.

Reiman, D. Will. 2010. Basic oral language documentation. *Language Documentation and Conservation* 4:254–268. Available at http://hdl.handle.net/10125/4479.

Richards, Norvin. 2014. A-bar movement. In *The Routledge Handbook of Syntax*, ed. Yosuke; Siddiqi Daniel Carnie, Andrew; Sato, 167–191. London/New York: Routledge Taylor and Francis Group.

Rizzi, Luigi. 1900. *Relativized minimality*. MIT Press: Cambridge.

Rodrigues, Aryon D. 1996. Argumento e predicado em Tupinambá. *Boletim da ABRALIN* 57–66.

Rodrigues, Aryon D. 1999. Macro-jê. In *The Amazonian Languages. Cambridge Language Surveys*, ed. Alexandra Y. Dixon, R. M. W.; Aikhenvald, 165–206. Cambridge: Cambridge University Press.

Rodrigues, Aryon D. 2000. Panorama das línguas indígenas da Amazônia: as línguas amazônicas hoje [An overview of the Indigenous languages in the Amazon: The Indigenous languages today]. Available at https://www.comciencia.br/dossies-1-72/reportagens/amazonia/amaz5.htm.

Rodrigues, Aryon D. 2001. Sobre a natureza do caso argumentativo. In *Des noms et des verbes en tupi-guarani*, ed. Francesc Queixalós, 103–114. Lincom Europa.

Rodrigues, Aryon D., and Ana Suelly A. C. Cabral. 2002. Revendo a classificação interna da família tupí-guaraní. In *Línguas indígenas brasileiras: Fonologia, gramática e história*, ed. Aryon D. Rodrigues and Ana Suelly A. C. Cabral, 327–337. Editoria Universitária UFPA.

Rodrigues, Aryon D, and Ana Suelly A. C. Cabral. 2012a. Tupían. In *The Indigenous Languages of South America: A Comprehensive Guide*, ed. Vera Campbell, Lyle; Grondona, 495–574. De Gruyter Mouton.

Rodrigues, Aryon D., and Ana Suelly Arruda Câmara Cabral. 2012b. Tupían. In *The Indigenous languages of South America: A comprehensive guide*, ed. Lyle Campbell and Vera Grondona, 495–574. Berlin/Boston: De Gruyter Mouton.

Rose, Françoise. 2015. On male and female speech and more: Categorical gender indexicality in indigenous south american languages. *International Journal of American Linguistics* 81(4):457–608. DOI: https://doi.org/10.1086/683158.

Ross, John Robert. 1967. Constraints on variable in syntax. Doctoral Dissertation, MIT.

Rudnev, Pavel. 2020. The anaphor agreement effect is not about featural deficiency: Evidence from avar. *Glossa* 5(1):1–19.

Sampaio, Wany B. A. 1997. Estudo comparativo sincrônico entre o Parintintin (Tenharin) e o Uru-Eu-Uau-Uau (Amondava): Contribuições para uma revisão na classificação das línguas Tupi-Kawahiva.

Sampaio, Wany B. A. 2001. As línguas Tupi-Kawahib: um estudo sistemático filogenético. Doctoral Dissertation, Fundação Universidade Federal de Rondônia, Porto Velho.

Sauerland, Uli. 2003. Unpronounced heads in relative clauses. In *The Interfaces: Deriving and Interpreting Omitted Structures*, ed. K. Schwabe and S. Winkler. John Benjamins.

Schenner, Mathias. 2019. An introduction to reconstruction effects in relative clauses,. In *Reconstruction Effects in Relative Clauses*, ed. Manfred Krifka and Mathias Schenner. De Gruyter.

Seiter, William J. 1979. Studies in Niuean Syntax. Doctoral Dissertation, University of California, San Diego, San Diego, CA.

Seki, Lucy. 2000. *Gramática do Kamaiurá: Língua Tupi-Guarani do Alto Xingu*. Campinas: Editora da Unicamp.

Selkirk, Elisabeth. 2011. The syntax-phonology interface. In *The handbook of phonological theory*, ed. Jason Riggle John Goldsmith and Alan Yu, 435–484. Oxford: Blackwell.

Shibatani, Masayoshi. 2009. Syntactic complexity: Diachrony, acquisition, neuro-cognition, evolution. In *Elements of complex structures, where recursion isn't: The case of relativization*, ed. T. Givón and Masayoshi Shibatani, 163–198. Amsterdam/Philadelphia: John Benjamins.

SIL. 1972. *Seleções Parintintín*. SIL.

SIL. 1976a. *Diversões de Margarida*. SIL.

SIL. 1976b. *Histórias de Kairana*. SIL.

SIL. 1976c. *Histórias de Margarida*. SIL.

SIL. 1977. *História de manuel joão*. SIL.

Silva, Auristéa Caetana Souza. 1999. Aspectos da referência alternada em Parakanã. Doctoral Dissertation, Universidade Federal do Pará, Belém, PA.

Silva, Vera da, Wany Sampaio, and Valdemir Miotello. 2004. *Mitos Amondawa*, volume 1. Porto Velho: Edufro.

Singerman, Adam Roth. 2018. The morphosyntax of Tuparí, a Tupian langauge of the Brazilian Amazon. Doctoral Dissertation, University of Chicago, Chicago.

Singerman, Adam Roth. 2021. On finite embedded clauses in Tuparí: their synchrony, diachrony, and typology. *Glossa: a journal of general linguistics* 6(1):1–46.

Solano, Eliete de Jesus Bararuá. 2009. Descrição gramatical da língua Araweté. Doctoral Dissertation, Universidade de Brasília, Brasília.

Stassen, Leon. 2006. Comparative constructions. In *Encyclopedia of language and linguistics*, ed. Keith Brown, 686–690. Oxford: Elsevier Ltd.

Storto, Luciana R. 1999. Aspects of a Karitiana Grammar. Doctoral Dissertation, MIT.

Tenharim, Dalva. 1978a. *Dalvahẽa okwatijarypy*. SIL.

Tenharim, Gilena. 1978b. *Gilenahẽa okwatija*. SIL.

Tenharim, João Bosco da Silva. 1977a. *Ore nhomomirika 'ndika'ndirame*. SIL.

Tenharim, João Bosco da Silva. 1977b. *Ahererohokava*. SIL.

Tenharim, João Bosco da Silva. 1977c. *Ore horame cidade pe*. SIL.

Tenharim, João Bosco da Silva. 1977d. *5° bec*. SIL.

Tenharim, João Bosco da Silva. 1977e. *Ahera'oa*. SIL.

Tenharim, João Bosco da Silva. 1977f. *Ahe rehe i'ara ahe'mbyva ipohaã pavẽi*. SIL.

Tenharim, João Bosco da Silva. 1977g. *Mbatera ga imombe'ui ji ve*. SIL.

Tenharim, João Bosco da Silva. 1977h. *A onça*. SIL.

Tenharim, João Bosco da Silva. 1977i. *Ore onga hepiapavi*. SIL.

Tenharim, João Bosco da Silva. 1977j. *Yvakytihava olaria*. SIL.

Tenharim, Lurde. 1978c. *Lurdehẽa okwatijara'ja*. SIL.

Tenharim, Lurde. 1978d. *Lurdehẽa okwatijarypy*. SIL.

Tenharim, Maria. 1978e. *Mariahẽa okwatija*. SIL.

Tenharim, Porciano. 1978f. *Porciano'ga okwatija*. SIL.

Tenharim, Zelito. 1978g. *Zelito'ga okwatijarypy*. SIL.

Tenharim, Zelito. 1978h. *Zelito'ga okwatijarypy*. SIL.

Thiersch, Craig. 2017. Remnant movement. In *Wiley online companion to syntax*, 435–484. New York: Wiley and Sons.

Thornes, Tim. 2023. Clause combining: relative clauses. In *The languages and linguistics of Indigenous North America*, ed. Carmen Dagostino, Marianne Mithun, and Keren Rice, volume 1, 323–343. Berlin/Boston: De Gruyter Mouton.

Tonhauser, Judith. 2006. The Temporal Semantics of Noun Phrases: Evidence from Guaraní. Doctoral Dissertation, Standford University, Standford, CA.

Tonhauser, Judith. 2007. Nominal tense? the meaning of Guaraní nominal temporal markers. *Language* 83(4):831–869.

Tonhauser, Judith. 2008. Defining crosslinguistic categories: The case of nominal tense (Reply to Nordlinger and Sadler). *Language* 84(2):332–342.

Tonhauser, Judith. 2015. Contrastive topics in Paraguayan Guaraní discourse. In *Proceedings from Semantics and Linguistic Theory*.

Toosarvandani, Maziar. 2010. Patterns of Nominalization in Numic. *International Journal of American Linguistics* 76:71–100.

Toosarvandani, Maziar. 2014. Two types of deverbal nominalization in northern paiute. *Language* 90:786–833.

Toyoshima, Takashi. 2001. Head-to-spec movement. In *The minimalist parameter*, ed. Galina M. Alexandrova and OlgaArnaudova, 435–484. Amsterdam and Philadelphia,PA: John Benjamins.

Travis, Lisa DeMena. 1984. Parameters and effects of word order variation. Doctoral Dissertation, MIT.

Travis, Lisa DeMena, and Diane Massam. 2021. VP-fronting within a movement typology. In *Parameters of Predicate Fronting*, ed. Vera; Lee-Schoenfeld and Dennis Ott, 33–62. Oxford: Oxford University Press.

Turpin, Myfany. 2013. Verb-final word order in the poetry of non-configurational languages. In *Proceedings of the 44th Annual Meeting of the Australian Linguistics Society*. Melbourne: Melbourne University.

Van Urk, Coppe. 2015. A uniform syntax for phrasal movement: A case study of Dinka Bor. Doctoral Dissertation, MIT, Boston.

Van Urk, Coppe. 2022. Constraining Head Movement. *Linguistic Inquiry* 1–45.

Van Urk, Coppe, and Norvin Richards. 2015. Two components of long-distance extraction: Successive cyclicity in dinka. *Linguistic Inquiry* 46:113–155.

Van Valin, Robert, and Randy J. Lapolla. 1997. *Syntax: structure, meaning, and function*. Cambridge: Cambrdige University Press.

Vicente, Luis. 2007. The syntax of heads and phrases: A study of verb (phrase) fronting. Doctoral

Dissertation, Leiden University, Leiden, Netherlands.

Vicente, Luis. 2009. An alternative to remnant movement for partial predicate fronting. *Syntax* 12(2):158–191.

Vieira, Marcia Maria Damaso. 1993. O fenômeno da não-configuracionalidade na língua Asuriní do Trocará: um problema derivado da projeção dos argumentos verbais. Doctoral Dissertation, Universidade Estadual de Campinas, Campinas (São Paulo).

Vieira, Marcia Maria Damaso. 2014. The manifestation of Topic and Focus in Tupi-Guarani Languages. *Delta* 30 (Special Issue):659–683.

Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. New York/Oxford: Oxford University Press.

Vivanco, Karin. 2018. Perguntas qu-, orações subordinadas e ordem de palavras em Karitiana [wh questions, subordinate clauses and word order in Karitiana]. Doctoral Dissertation, University of São Paulo.

de Vries, Mark. 2018. Relative Clauses in Syntax. Oxford Research Encyclopedia of Linguistics.

Walker, Rachel. 2011. Nasal harmony. In *The Blackwell Companion to Phonology*, ed. Keren; Rice, Colin J.; Ewen, Elizabeth Hume, and Marc Oostendorp, volume 3. Oxford, UK: John Wiley Sons, Ltd.

Wycliffe. 1996. *Tupana'ga nhi'iĝa*.

Yap, Foong Ha, and Karen Grunow-Hårsta. 2010. Non-referential uses of nominalization constructions: Asian perspectives. *Language and Linguistics Compass* 1154–1175.

Yap, Foong Ha, Karen Grunow-Hårsta, and Janick Wrona. 2011. *Nominalization in Asian Languages: Diachronic and typological perspectives*. Amsterdam/Philadelphia: John Benjamins.

Yuan, Michelle. 2022. Deriving VSO in San Juan Piñas Mixtec (and some puzzles along the way). Handout from talk given at the MIT Syntax Square.

Zariquiey, Roberto, Masayoshi Shibatani, and David W. Fleck. 2019. *Nominalization in the languages of the Americas*. Amsterdam/Philadelphia: John Benjamins.