

THE UNIVERSITY OF CHICAGO

A GRAMMAR OF KARAJÁ

A DISSERTATION SUBMITTED TO
THE FACULTY OF THE DIVISION OF THE HUMANITIES
IN CANDIDACY FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

DEPARTMENT OF LINGUISTICS

BY

EDUARDO RIVAIL RIBEIRO

CHICAGO, ILLINOIS

JUNE 2012

UMI Number: 3517263

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



UMI 3517263

Copyright 2012 by ProQuest LLC.

All rights reserved. This edition of the work is protected against unauthorized copying under Title 17, United States Code.



ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 - 1346

COPYRIGHT © 2012 by Eduardo Rivail Ribeiro

All rights reserved.

I dedicate this dissertation
to my mother-in-law,
Dr. Kalavati J. Dave,
and to my parents,
José Ilário Ribeiro and
Terezinha Pereira Ribeiro.

CONTENTS

Acknowledgements	v
Abbreviations	viii
List of tables	ix
Chapter 1: Introduction	1
Chapter 2: Phonology	73
Chapter 3: Female and male speech	127
Chapter 4: Morphology	172
Chapter 5: Subordinate clauses	222
Chapter 6: Karajá as a Macro-Jê language	262
Bibliography	288

ACKNOWLEDGEMENTS

This dissertation would not have been possible without the support of several individuals and institutions, whose contributions I would like to acknowledge. My studies at the University of Chicago were supported by grants from CNPq (Conselho Nacional de Desenvolvimento Científico e Tecnológico, Brazil) and the Wenner-Gren Foundation (under the sponsorship of Michael Silverstein), as well as a Century Scholarship from the University's Division of Humanities. Field work in Brazil was made possible, on several occasions (1999-2003), by Tinker Field Research Grants provided by the University's Center for Latin American Studies.

Throughout the long years leading to this dissertation, I was fortunate to rely on the unwavering support of my families—the one I was born into, the one I was welcomed into, and the one I would later build. To my parents, José Ilário and Terezinha, my siblings (and their spouses), nephew, and nieces, and my mother-in-law, Kalavati Dave, my sincerest thanks for all the love and encouragement. My wife, Neha, and my daughters Uma (a.k.a. “Doc”) and Roma (a.k.a. “Post-Doc”), were always an inspiration, without which life—academics and dissertation included—would have seemed purposeless. My little girls' insistence, at bedtime, to hear Karajá stories over and over again showed me that my two ‘universes’, continents apart, were both inextricably weaved into our lives. And when academics became, at times, a rather secondary pursuit, friends such as Valerie Brown would always remind me of what I was capable of achieving.

I have learned Karajá from too many native speakers, always so generous in sharing their knowledge of their language and culture. Naming each one of them would be impossible, since every interaction with an *Iny*, in their villages or elsewhere, was a potential learning opportunity. I will limit myself to acknowledging by name my main Karajá teachers, who have also become dear friends. I owe much of what I know about the language to the late Ijeseberi Karajá, a true scholar who, although educated in the city and well-versed in Brazilian society, remained deeply rooted in his native culture. Sinvaldo Wahùka and Luiz Kurikala welcomed me as a pupil, as a guest, and as a friend. Their insights into their own language's complexities expanded not only my understanding of Karajá, but also of language in general and its relationship to culture.

Thanks to female speakers such as Xirikeru Karajá and Violeta Karajá (one of the main guardians of the Xambioá culture), I gained valuable insight into the characteristics of female speech, often neglected in the previous literature on the language. Many talented storytellers (among them, the late Ereheni, Mauare, Ijyraru, Alice Kuabiru, Madalena Areraki, and Arumani) kindly introduced me to the fascinating Karajá oral tradition, narrating traditional tales and historical accounts which richly illustrate complex grammatical phenomena and poetic resources. These individuals, as well as their fellow tribesmen, illustrate the true meaning of hospitality, welcoming me not only into their villages and homes, but also into their culture and everyday life pursuits. That experience, no less than the countless data collection and elicitation sessions, was invaluable for my learning of Karajá. This dissertation is a modest tribute to my Karajá teachers' patience and generosity and to their society's endurance.

Among my fellow South Americanists, especial thanks are due to Denny Moore, to whom I owe the fateful suggestion—and constant encouragement—to pursue my studies at the University of Chicago, and Mônica Veloso Borges, who, in addition to being a confidant in matters of linguistic analysis, has been a loyal friend to my family and I since my undergraduate years. I would also like to thank Walkíria Neiva Praça, for sharing her knowledge of Tapirapé with me, and Andrés Salanova, for kindly providing me with data and insights on Mebengokre, which have been instrumental for my understanding of Proto-Jê and its relationship with Karajá. Hein van der Voort, in particular, has become a dear friend, a trusted companion in the pursuit of learning, and a constant source of guidance, inspiration, and support in both difficult and happy moments.

Lastly, I would like to thank my professors at the University of Chicago, to whom I owe any merits this dissertation may have. Many of the ideas in this dissertation grew out of classes I took with Jerry Sadock, Amy Dahlstrom, Michael Silverstein, Salikoko Mufwene, John Goldsmith, Kostas Kazazis, Anthony Buccini, and Gunnar Hansson, among others. Chicago provided me with a deeper understanding of the value of theoretical diversity and with a set of analytic tools which I will forever cherish in my career as a linguist. I was particularly fortunate to have in my dissertation committee linguists who, while having solid reputations as theoretical linguists, equally value data-oriented linguistic analysis: Amy Dahlstrom, Jerry Sadock, and Michael Silverstein. I am particularly indebted to Amy Dahlstrom, not only for her contributions to my education as a linguist, but also for her kind support and encouragement, without which this dissertation would not have come to light.

ABBREVIATIONS

ALL	Allative
ANTI	Antipassive
ASSERT	Assertive
ATR	Advanced tongue root
CONV	Converb
CTFG	Centrifugal direction
CTPT	Centripetal
EMPH	Emphatic
EXC	Excitement
FUT	Future
IMPERF	Imperfective
INSTR	Instrumental
INTER	Interrogative
INTR	Intransitive
LOC	Locative
NEG	Negation
NOM	Nominalizer
PERF	Perfective
PL	Plural
POT	Potential
PROGR	Progressive
REL	Relational, or linking, prefix
STRESS	Subordinating accent
TEMP	Temporal
TRANS	Transitive
VERB	Verbalizer

LIST OF TABLES

Table 1.1. Língua Geral loans in Karajá	7
Table 1.2. Some Tapirapé loans in Karajá	11
Table 1.3. Some Karajá loans in Tapirapé	11
Table 1.4 . Some Karajá (Xambioá) loans in Kaiapó (esp. Xikrin)	13
Table 1.5. Karajá dialects, regular phonological differences: schwa	21
Table 1.6. Karajá dialects, regular phonological differences: CCV vs. CVCV	21
Table 1.7. Karajá dialects: a few irregular differences	21
Table 1.8. Consonantal inventory of Karajá	26
Table 1.9. Vocalic inventory of Karajá	28
Table 1.10. Vowels according to their behavior in terms of vowel harmony	33
Table 1.11. Possessive prefixes in Karajá (Ribeiro 1996)	42
Table 1.12. Subject agreement markers in Karajá	47
Table 1.13. Tense/aspect clitics in Karajá	59
Table 1.14. Karajá postpositions	62
Table 2.1. Consonantal inventory of Karajá	83
Table 2.2. Vocalic inventory of Karajá	86
Table 2.3. Vowels according to their behavior in terms of vowel harmony	107
Table 3.1. Historical development of the form ‘jaguar’	132
Table 3.2. Historical development of the form ‘river’	133
Table 3.3. Class <i>a</i> : absolute exceptions to the rule of <i>k</i> -dropping	143
Table 3.4. Class <i>b</i> : words for which the rule of <i>k</i> -dropping is optional	144
Table 3.5. Class <i>c</i> : words for which the rule of <i>k</i> -dropping is obligatory	144
Table 3.6. Class <i>c</i> male speech forms occurring in early word lists	147
Table 3.7. Class <i>b</i> words occurring in early word lists	148
Table 3.8. Female speech in Javaé as opposed to Karajá	157
Table 3.9. Proto-Jê consonantal inventory (Ribeiro 2005)	165
Table 4.1. Subject agreement markers in Karajá	174
Table 4.2. Possessive prefixes in Karajá (Ribeiro 1996)	205
Table 5.1. Karajá postpositions	224
Table 5.2. Paradigms for some Karajá postpositions	225
Table 6.1. The Macro-Jê Hypothesis	264
Table 6.2. Some lexical cognates between Proto-Jê and Karajá (apud Davis 1968)	269

Table 6.3. Additional lexical cognates	270
Table 6.4. Derivational suffixes in Jê and Karajá	275
Table 6.5. Consonant replacement in Karajá: k > r	278
Table 6.6. Proto-Jê consonant inventory (Ribeiro 2005)	279
Table 6.7. Karajá consonant inventory (Ribeiro 2007)	279
Table 6.8. Cluster simplification in Pre-Proto-Karajá	280
Table 6.9. “Pseudo metathesis”	282
Table 6.10. From suffix to “replacive infix”	282
Table 6.11. Consonant replacement in Karajá: h > r	

Chapter 1

INTRODUCTION

1. The Iny and their language

The main purpose of this dissertation is to provide an introductory grammar of Karajá, a Macro-Jê language spoken along the Araguaia River, in Central Brazil. Karajá has three mutually-intelligible dialects—Karajá proper, Javaé, and Xambioá. Karajá proper can be further divided into Northern and Southern Karajá. Combined, the four dialects have a total of approximately three thousand speakers (Silva 2009), in the Brazilian states of Goiás, Mato Grosso, Tocantins, and Pará. The dissertation takes into account data of all four dialects, collected by the author in several field trips. Except for the Xambioá dialect, Karajá is quite vigorous in sociolinguistic terms, being the first language to be learned by the children and the language used in everyday life in nearly all of the villages.

Speakers of the four different dialects refer to their people as *Iny* [i'nɔ̃] ‘people, human’, and to their language as *Iny Rybe* [i'nɔ̃ ri'be]. Ethnographic naming practices follow in part the point of view of the Karajá proper, who call their northernmost kin *ijɔ̃bɪkɔ̃wa* ‘companion people’ (*ijɔ̃* ‘people’ + *bɪkɔ̃wa* ‘friend, companion’), whence the Portuguese word *Xambioá*. Despite their obvious linguistic and cultural communion, the Karajá refer to the Javaé as *ijɔ̃dʒu*, a term generally reserved for non-Karajá Indians such as the Xavánte. As we will see (Chapter 2), such attitude has less to do with linguistic differences than with cultural and historical ones. The Javaé, on the other hand, call the remaining Karajá *berohokɔ̃ mahãdu* ‘the people of the big river [i.e. the Araguaia]’. Non-Indians are called

dori or, less-commonly, *weku* or *wəɫə* (words whose etymologies are unknown).

Bilingualism is very common among the men, who frequently visit the neighboring towns in order to trade fish and other commodities; some (including entire families) travel even farther, seeking education or medical treatment.

In this introduction, I provide background information on the language and its speakers: Section 2 describes the geographic location of the Karajá-speaking tribes; Section 3 discusses the language's genetic affiliation; Section 4 describes contacts with speakers of different languages; Section 5 offers a brief appraisal of previous scholarship on Karajá language and culture; Section 6 describes basic dialectal differences; Section 7 briefly discusses the circumstances surrounding the collection and elicitation of the data on which the present dissertation is based. The introduction concludes with a brief grammatical sketch of the language (Section 8) and a narrative table of contents (Section 9).

2. Location and subsistence

The Karajá, Javaé, and Xambioá inhabit the margins of the Araguaia River, a tributary of the Tocantins (which, in turn, is a southern tributary of the Amazon). It is a transition area between the savannahs (Portuguese *cerrados*) of the Central Plateau (*Planalto Central*), to the south, east, and west, and the Amazon rainforest, to the north and northwest. The heart of their territory, where three of the four dialects are spoken, is the Bananal Island, the world's largest fluvial island, which is a mosaic of cerrado, forest, and swamp vegetation. The Karajá proper and Xambioá occupy the main course of the Araguaia River, while the Javaé live by its smaller branch, aptly named Javaé River (although historically they lived in

the interior of the island, until at least the first half of the last century). Although they concentrate largely in the island itself, there are Karajá villages on both sides of the Araguaia, between the states of Goiás, Tocantins (right banks), Mato Grosso, and Pará (left banks). The Xambioá, the northernmost Karajá-speaking group, live in a more Amazonian setting, roughly across from the mouth of the Maria River, a western tributary of the Araguaia.

It is likely that the Karajá were already occupying the Araguaia valley centuries before the arrival of the first European sailors to the Brazilian coast in 1500. Archaeological evidence indicates a continuum between the ceramics unearthed in the area, dated between 1190 AD and 1260 AD, and the ceramics historically associated with the Karajá (Lima Filho 1994:29). It could have been that the ancestors of the current-day Karajá, Javaé, and Xambioá simply adopted the local ceramic tradition upon migrating to the area. However, as far as Karajá oral history is concerned, there is no indication of recent migrations from another area. In fact, all the places mentioned in their myths of origin are very precise geographic locations within their traditional territory in or around the Bananal Island. Therefore, Toral's (1992) claim that "all Karajá hail from the north" (that is, from near the point where the Araguaia and Tocantins Rivers merge) also seems to lack ethnographic and linguistic basis. Linguistically, the Bananal Island is a more likely center of dispersal, since it is where most dialectal diversity is found. Plants and animals mentioned in the origin myths are also typical cerrado species, not rainforest ones (which would be expected were the Karajá of a northern, more Amazonian origin).

The Karajá-speaking groups are masterfully adapted to the Araguaia, from whose waters they take most of their sustenance. As excellent fisherman, the Karajá are in contrast

with the Jê tribes to the east (who used to rely more heavily on hunting, which to the Karajá has only secondary importance), and the Tupí-Guaraní-speaking Tapirapé to the west, traditional agriculturalists. Seasonal gathering (turtle eggs, nuts for the production of oils, palm leaves for basketry, honey, beeswax, fruits and berries, etc.) also plays an important role, both economically and symbolically. They are also well-established agriculturalists, planting maize, manioc (both sweet and bitter), yam, cotton, squash, urucum, chilly peppers, and introduced plants such as rice and sugarcane. The Karajá are also increasingly integrated with the local economies, often visiting the nearby towns to sell fish and crafts (Karajá traditional clay dolls, ♀ *ritfoko* (♂ *ritfoo*) are particularly appreciated among tourists). Some Karajá are also employed by the government as teachers, nurses, boat pilots, or in bureaucratic positions with the Brazilian Indian affairs agency, FUNAI (Fundação Nacional do Índio). That, and retirement pensions, are increasingly-important sources of income among Brazilian Indians in general.

3. Genetic classification

A relationship between Karajá and the Jê family (which forms the core of the Macro-Jê stock) was first proposed by Karl von den Steinen (1886), but the evidence presented then was far from convincing, as criticized by Ehrenreich (1894). Later classifications would consider Karajá as an isolate (Loukotka 1968, Mason 1950, McQuown 1955), until Irvine Davis (1968), using the standard historical-comparative method, compared Karajá and Maxakalí data with his own reconstructed Proto-Jê forms (Davis 1966), detecting a number of phonological and lexical correspondences. All subsequent classifications of South

American languages (Rodrigues 1970, 1986, 1999; Greenberg 1987, Kaufman 1994) agree in including Karajá into the Macro-Jê stock. As I have shown elsewhere (Ribeiro 2005), Davis' Proto-Jê requires a thorough revision. However, as I discuss in Chapter 6, most of the correspondences between Jê and Karajá he detected are being further corroborated—and often refined--by additional evidence.

4. Contacts

The penetration of Europeans (or their Brazilian descendents) into Karajá territory took place mainly in two different fronts: Jesuit missionaries from Pará, to the north, and *bandeirantes*, explorers from São Paulo, to the south (Palacin 1972). There are written reports of contacts between Jesuits and the Karajá as early as 1658 (Leite 1943:338), although such contacts seem to have been rather sporadic. Several bandeirante parties, in search of gold and Indian slaves, seemed to have reached the Araguaia as well, throughout the 1600s. No lasting contacts were made, however. Penetration into the Karajá territory would intensify with the founding (1727) of Sant'Anna (later Vila Boa, currently Cidade de Goiás), a bandeirante town by the Vermelho River, an eastern tributary of the Araguaia. Although their territory was mostly seen as a passageway to more profitable enterprises (commerce with Belém, to the north; search for mostly-mythical gold mines, to the north and west; contact with the town of Cuiabá, in Mato Grosso, to the west), the Karajá were, nonetheless, victims of merciless attacks by the conquering forces.

One of the most memorable attacks by the bandeirantes against the Karajá was perpetrated by Antônio Pires de Campos, in the mid-1700s, while marching from Cuiabá (the

current capital of Mato Grosso state) on his way to support Goiás colonists against the Southern Kaiapó (ancestors of the current-day Panará, a Northern Jê tribe). Backed by 500 Boróro warriors, who had become allies of the bandeirantes in raids against other tribes, Pires de Campos attacked by surprise a Southern Karajá village, killing many and taking many others as prisoners to be sold as slaves, some of which would later escape and return to the Araguaia. Decades later, when government troops under the command of José Pinto da Fonseca established the first peaceful contacts with the Southern Karajá, Pires de Campos' attack was still fresh in their memory.

Fonseca's report to the governor of Goiás (Fonseca 1846[1775]) is one of the first trustworthy sources on the tribe. To serve as an interpreter, he had brought back a Karajá woman, one of the captives taken by Pires de Campos decades earlier. Fonseca also mentions the presence of a Boróro "slave," belonging to the Karajá chief, presumably taken as prisoner during Pires de Campos' raid. Pieces of information such as these are not just historical curiosities, but may help understand early language contacts between the Karajá and the colonizers. For much of the colonial period (1500-1808), *Língua Geral*, a *lingua franca* based on Old Tupí or Tupinambá (the Tupí-Guaraní language which was spoken along most of the Brazilian coast by the time of the arrival of the Portuguese), was a major colonial language—more important in some places, even, than Portuguese (Rodrigues 1996). The bandeirantes spoke the southern variety of the language, *Língua Geral Paulista* (based on the Tupí dialect of São Vicente, today São Paulo state), and spread it as new territories were conquered and colonized. Pires de Campos and his Boróro allies certainly spoke *Língua Geral Paulista*. More than half a century after the attack, descendants of the Boróro, settled by Pires de Campos in a village along the route between Goiás and São Paulo to serve as a

permanent outpost against the Southern Kaiapó, were visited by the French Botanist Auguste de Saint-Hilaire, who collected a vocabulary of their language (Saint-Hilaire 1847); they still spoke Língua Geral, which had by then largely disappeared from other areas. Língua Geral loans in Karajá (Table 1), referring to elements introduced with the European colonization, probably date from this early period of contact. The other variety of Língua Geral, Língua Geral Amazônica, which survives to this day (under the name Nheengatú), enjoyed even more widespread use in the Amazon region, initially as the language favored by the Jesuits in their missions. At least one loanword, restricted to the Xambioá dialect, may trace back to Língua Geral Amazônica: *bãbera* [mabe'ra] ‘paper’ (< Língua Geral *papéra* < Portuguese *papel*).

Table 1.1. Língua Geral loans in Karajá

Karajá	Old Tupí	
<i>dzikira</i>	<i>Jucúra</i>	‘salt’
<i>bãkawa</i>	<i>Mocába</i>	‘firearm’
<i>burure ~ brure</i>	<i>Pururé</i>	‘hoe’
<i>kobãda</i>	<i>Comandá</i>	‘beans’
<i>bãbera</i>	<i>Papéra</i>	‘paper’ (Xambioá only)
<i>ĩdadzuwa</i>	<i>Itajúba</i>	‘money’ (no longer in use)

The ethnonym "Karajá" itself may be particularly revealing of these earlier contacts with the bandeirantes. Although most authors agree on its Tupí-Guaraní origin—in Guaraní, it refers to a large monkey species, called *guariba* in Portuguese--, its precise provenance is rarely discussed. It so happens that a cognate of the Guaraní word is not found in any of the Amazonian Tupí-Guaraní languages with which the Karajá were likely in contact. Its occurrence is restricted to southern Tupí-Guaraní languages such as Guaraní and Sirionó (Mello 2000). Baldus (1948) suggests that such an epithet is also due to the bandeirantes,

speakers of Língua Geral Paulista. Although a similar word was not documented for Old Tupí, it is not unlikely that it may have occurred in Língua Geral Paulista, since the Tupí of São Vicente was described as having been influenced by Guaraní (maybe forming a link in a continuum between Guaraní and Tupinambá). Furthermore, one of the main economic occupations of the bandeirantes was raiding the Jesuit missions among the Guaraní in the south, in current-day Paraguay, where they could very well have adopted the word. Also of uncertain origin, the term *Javaé* may have come from a Tupí-Guaraní language as well (cf. Kaapór *sawa'e* 'man'; Kakumasu & Kakumasu 2007:50).

Fonseca's visit marks a turning point in the relationship between the government and the Karajá. The governors of Goiás were now interested in bringing the Indian tribes under the watchful eye of the government, through the creation of official settlements (*aldeamentos*) gathering different ethnic groups, generally away from their traditional territories (such as São José dos Mossâmedes, located south of Vila Boa). Hundreds of Javaé and Southern Karajá would agree to join the *aldeamentos*, only to later abandon them because of widespread death by disease and mistreatments. Closer to the indigenous territories, the government would create *presídios*, military outposts which would be often attacked and destroyed by the Karajá and Javaé (allied with other indigenous tribes such as the Xerénte). Despite such opposition, *aldeamentos* and *presídios* would gradually contribute to the weakening of native resistance. In 1850, the *presídio* of Leopoldina was founded near the mouth of the Vermelho River. Though destroyed and rebuilt more than once, the *presídio* signals a permanent official presence in Southern Karajá territory. With the consolidation of Leopoldina, later a town (today's Aruanã), the Karajá were now in closer contact with the national, Portuguese-speaking population. An increased Brazilian presence in the Araguaia,

both in Leopoldina and in Southern Pará (near the territories of the Xambioá and Northern Karajá) also meant that it was safer for naturalists, such as Castelnau, and ethnographers, such as Ehrenreich, to travel up and down the river, contacting the Karajá and collecting information on their language and culture. In most of its extension, however, the Araguaia was still free from permanent Brazilian settlements, a situation which would last well into the 20th century. Despite initial attempts to increase the commercial use of the Araguaia, with the creation of steamboat lines in the second half of the 19th century, such initiatives were short-lived. Although limiting the influx of large waves of colonists, the lack of economic prosperity did not mean an end to the Karajá's troubles. The Xambioá, once described as the most prosperous and numerous among the Karajá-speaking groups (Ehrenreich 1891), would come close to extinction, victims of military actions and epidemics (Toral 1992). In the beginning of the 20th century, there were only a few scattered Xambioá families, which would later be gathered by the SPI (Serviço de Proteção ao Índio), a forerunner of FUNAI. These families are the ancestors of the current-day Xambioá. Due to such drastic population decline and their geographic distance from the other Karajá villages, the Xambioá married local Brazilians. Karajá language and culture would gradually lose ground, in such a way that the Xambioá dialect is now nearing extinction, with only a handful of elderly speakers.

Starting in the 1930s, a new political climate brought a renewed interest in colonizing the region west of the Araguaia. This "March to the West" (*Marcha para o Oeste*) would result in an ever-growing presence of cattle ranches in indigenous territory, and the founding of towns such as Luciara (near Northern Karajá villages in the north of the Bananal Island), Santa Terezinha do Araguaia (near the Southern Karajá village of São Domingos, or Krehãwa), and São Félix do Araguaia (near the Southern Karajá village of Santa Isabel do

Morro, or Hãwalò). Roads and a brand new city—Goiânia—would be built, attracting even more newcomers to Central Brazil. Airplanes would bring adventurers and journalists from Rio and São Paulo, who fascinated their readers with stories of untamed nature and men. Even two of Brazil's most enterprising and popular presidents, Getúlio Vargas (1940) and Juscelino Kubitschek (1960), would visit the Karajá village of Santa Isabel do Morro. The rush of activity in the region, this time, was more than a mere phase, bringing lasting changes and additional challenges for the Karajá, who were now in permanent contact with an ever-growing Brazilian population. But, after centuries of dealing with the newcomers, the Karajá had learned to survive their intrusion, managing to preserve their own culture—as well as much of their traditional land—against all odds.

Both before and after the arrival of the European colonizer, the Karajá were also in contact with neighboring tribes, or tribes who would eventually make incursions into the Araguaia. Of these, contacts with the Tapirapé (♀ *woku*, ♂ *wou*), a Tupí-Guaraní tribe, were certainly the most important, for having resulted in mutual, lasting cultural influences. Traditional inhabitants of the upper course of the Tapirapé River (a western tributary of the Araguaia), the Tapirapé would seasonally frequent the region at the mouth of the Tapirapé River, where they would trade (and occasionally fight) with the Northern Karajá. Occupying the main course of a major river, the Karajá would act as intermediates between the Brazilians and the Tapirapé, exchanging old iron tools for Tapirapé goods. As Baldus points out, the Karajá traditionally acted as purveyors of “chants, tools, and epidemics” to the Tapirapé (Baldus 1970:65). The Tapirapé, on the other hand, would provide the Karajá with agricultural products and tamed macaws (whose feathers are prized for their ornamental value). The loans exchanged between both languages reflect not only the economic expertise

of each tribe (Tables 1.2 and 1.3), but also the exchange of ritual elements. A few Karajá loans, such as the words for ‘banana’ and ‘White man’, occur not only in Tapirapé, but also in Asuriní of Tocantins and Parakanã, closely-related Tupí-Guaraní languages whose speakers live near the lower Tocantins River, providing clues for determining a relative time frame and approximate migratory route for the Tupí-Guaraní tribes in the region.

Table 1.2. Some Tapirapé loans in Karajá

Karajá	Tapirapé	
<i>behira</i>	<i>pe?ira</i>	‘carrying basket’
<i>kɔbɛdawira</i>	<i>komana?iwira</i>	‘ <i>andu</i> beans’
<i>hārara</i>	<i>ārārā</i>	‘macaw (sp.)’
<i>darawe</i>	<i>tārāwe</i>	‘parakeet (sp.)’
<i>tʃakohi</i>	<i>tʃākoʔi</i>	‘Txakohi ceremonial mask’

Table 1.3. Some Karajá loans in Tapirapé

Karajá	Tapirapé	
<i>idʒada</i>	<i>tʃātā</i>	‘banana’
<i>dori</i>	<i>tori</i>	‘White man’
<i>bɔrɔɔ</i>	<i>marara</i>	‘turtle stew’
<i>irabure</i>	<i>irāwore</i>	‘Irabure ceremonial mask’

Some Tapirapé groups roamed even further east, settling in the interior of the Bananal Island (their presence being mentioned by early reports on the Karajá, such as Fonseca’s). There, according to the oral histories of both tribes, they would have mixed with the ancestors of present-day Javaé, contributing to set this Karajá-speaking group apart—culturally and linguistically—from the remaining of the tribe (see Chapter 3). In addition to Tapirapé loans such as the ones listed in Table 2, which are common to all Karajá dialects, a few are restricted to the Javaé dialect. That is the case, for instance, of the word for ‘garbage’, Javaé *hidĩ* (< Tapirapé *?itĩ*), and proper nouns such as *Kujamōkō* (cf. Ribeiro 2001), all of which suggest a more intimate contact than a merely commercial one. There are

also several cultural features shared by the Tapirapé and the Javaé, but not the other Karajá groups.

There is an epilogue in the history of contacts between the Karajá and the Tapirapé. In the 1940s, following a drastic population drop resulting from attacks by the Kaiapó (Northern Jê), the surviving Tapirapé came to live near a Northern Karajá village at the mouth of the Tapirapé River. From relatively sporadic, contact between both tribes came to become permanent, resulting in many interethnic marriages characterized by a situation of domestic bilingualism with relative equilibrium between both languages. Recently, after having recovered their traditional lands in the west, most Tapirapé left the Araguaia, taking along the many additional Karajá influences acquired in the past half a century. Most mixed families, however, remained behind. Among these, the effects of language contact can be felt not only in the lexicon, but also in the phonology and syntax; such situation is a fascinating topic for future studies.

Another tribe with whom the Karajá were in close contact were the Kaiapó (*karalahu*), speakers of a Northern Jê language. At the beginning of the 19th century, the Kaiapó lived in the area between the lower Araguaia and the Tocantins, according to early historians (Cunha Mattos (1874-1875[1824]), Silva e Souza (1842[1812])). Due to the invasion of their territory by cattle ranchers, they begin a gradual migration towards west, across the Araguaia and into Pará. When Coudreau and Krause visited the area, between 1890 and 1910, all Kaiapó-speaking tribes were already settled west of the Araguaia River. The first wave of Kaiapó migrants were the ancestors of the present-day Xikrin, who would have migrated in the first decades of the 19th century (Vidal 1977). Marching west, they would establish peaceful contacts with the northernmost Karajá-speaking group, the

Xambioá. As a result of such contacts, the Kaiapó borrowed a number of cultural elements (with their terminology) from Xambioá. The linguistic loans include not only terms referring to items of material and ceremonial culture, but possibly even a relationship term, *bikwa* ‘relative’ (< Karajá *bikɔwa* ‘friend, companion’, ♂ *biɔwa*). Additional waves of Kaiapó would progressively migrate west, this time mostly through Northern Karajá territory, and far from peacefully (I have recorded, among the Northern Karajá of Santana do Araguaia, narratives of such conflicts). While some of these late migrants would end up being catechized by the missionaries and absorbed into the regional population, disappearing as an ethnic group, most continued their migration towards west, ending up as far away as the Xingu River. Currently, no Kaiapó-speaking groups live near the Araguaia.

Table 1.4 . Some Karajá (Xambioá) loans in Kaiapó (esp. Xikrin)¹

Kaiapó	Karajá	
<i>wiwi</i>	<i>wii</i>	‘song, chant’
<i>warikɔkɔ</i>	<i>warikɔkɔ</i>	‘tobacco pipe’
<i>rara</i>	<i>lala</i>	‘basket (type of)’
<i>bikwa</i> ‘relative’	<i>bikɔwa</i>	‘friend, companion’
<i>berɔ</i>	<i>berɔ</i>	‘puba (manioc flour)’

At least one Karajá flora word may have been borrowed from Kaiapó, *kabiri* ‘bacaba’ (a palm-tree berry closely related to açai), called *kamere* in Kaiapó (and closely related Jê languages). But it may simply be a case of common inheritance from Proto-Macro-Jê, given

¹ Besides its importance in providing evidence of past contacts between the Xambioá and the Kaiapó, the identification of such loans also helps explain apparent exceptions to diachronic phonological rules undergone by the latter. As I have shown elsewhere (Ribeiro 2005), Proto-Jê *w is reflected as /b/ in Kaiapó in most environments. Examples such as *warikɔkɔ* and *wiwi* would remain puzzling mysteries were it not for their identification as Karajá loans. This also helps provide a relative chronology for the transformation of Proto-Jê *w into Kaiapó /b/ (a major departure from its closest relatives, Apinajé and Timbira), which would certainly have happened before the loans were introduced—therefore, no later than the first half of the 19th century (Vidal 1977)

that their similarities seem to be matched by regular phonological correspondences between Karajá and Jê. At least one proper male noun in Karajá—Krumare [kru'marɛ]—was probably borrowed from Kaiapó,² either as a result of peaceful relationships, or inherited from captured Kaiapó women.³

Finally, a few loanwords seem to belong to an older layer of contacts, with tribes with whom the Karajá are no longer in touch. That is the case of the word for ‘maize’, *bãki* [ma'ki], which was probably borrowed (directly or indirectly) from one of the Arawák-speaking tribes of the Xingu region (cf. Ribeiro 2009), and the word for ‘poisonous arrow’, *dʒ-uwaθa*, probably borrowed from an unknown Tupí-Guaraní tribe (cf. Tupinambá *r-ubasy*). A sui generis kind of loan is the word *dʒakare* ‘alligator’ (obviously of Tupí-Guaraní origin), which occurs only as a formulaic expression in a myth; versions of the same myth, not found among the Tapirapé, are found among Xingu tribes (Ribeiro 2001a). The Karajá, in fact, share a number of cultural features with the Xingu, including a particular form of wrestling and the ceremonial use of the atlatl (Toral 1992, Lima Filho 1994). The territory between the Araguaia and Xingu rivers was likely inhabited by tribes who have migrated or become extinct early on. As Baldus (1938) suggests, such tribes could have served as an “ethnographic bridge” between the Araguaia and the Xingu. Furthermore, as traditional inhabitants of a major waterway, the Karajá were certainly in contact with several tribes migrating east or west of their territory (as illustrated by the Tapirapé and Kaiapó cases). Loanwords of uncertain origin suggest a much more diverse ethnographic landscape than the

² *Krôma-re* was the name, for instance, of a prestigious Kaiapó chief, often mentioned by Verswijver (1992); the suffix *-re* is very common in Kaiapó proper nouns.

³ The presence of captured Kaiapó and Tapirapé women among the Karajá is often mentioned in the early literature. Captured women would become full-fledged members of Karajá society.

one found at the time of the first documented contacts, with cultural networks which are only beginning to be understood.

Contacts with other tribes of the Araguaia-Tocantins area, such as the Apinajé (whom the Karajá call *wɛbɪdõle* [wɛbɪmõ'le]) and the Xavánte (*kɪθa*, ♂ *ɛɪθa*), seem to have been rather sporadic. Although the Xavánte were the most formidable enemies of the Southern Karajá until very recently (again, I have collected many stories of skirmishes between both tribes), their interaction did not result in any apparent exchange of cultural or linguistic loans. That is a fact important to emphasize, vis-à-vis Rodrigues' (1999, 2004) hypothesis that the differences between male and female speech in Karajá would be a result of Xavánte influence (see Chapter 2). Especially in the case of Jê tribes such as the Xavánte and the Kaiapó, forced to migrate west by increasing colonial encroachment upon their original territories, much of the Karajá contacts with tribes to the east were triggered by the arrival of the Portuguese and their descendents. That is not to say that there were no contacts with eastern tribes in pre-colonial times; but, if there were such contacts, they seem to have had less cultural importance than contacts with tribes to the west.

5. Previous sources

The Karajá language has been mentioned since the mid-19th century, mainly in works by naturalists, ethnographers, and voyagers, who collected vocabularies (Castelnau 1851, Coudreau 1897, Ehrenreich 1891, Sócrates 1892, Krause 1911, Brito Machado 1947, Sekelj 1948). There is also a grammatical essay published by a Dominican missionary, Luiz Palha (1942). Most of the descriptive work on Karajá, however, was done by David and Gretchen

Fortune, from the Summer Institute of Linguistics (SIL), who have been working on the Northern Karajá dialect since 1958, with the main goal of translating the Bible (the New Testament has already been published (1983)). Their descriptive works are restricted to a few articles (most of which are unpublished), including a phonological description (Fortune and Fortune 1963), two grammatical essays (Fortune and Fortune 1964; Fortune 1975), and a study of the differences between men and women's speech (Fortune and Fortune 1975). More recent works include Maia's (1986[1998]) description of typological aspects of Javaé (based on Fortune's description of the Northern Karajá dialect), Cavalcante's description of the phonology of Southern Karajá (1992), Viana's MA thesis on the expression of the attribute (1995), my own MA thesis on verb morphology (Ribeiro 1996), and Borges' (1997) study on the differences between male and female speech. While preparing this dissertation, I presented preliminary findings in several conferences; some of such talks resulted in publications (Ribeiro 2001b, 2002, 2003, 2004)

Thus, the history of Karajá language documentation can be divided between two clearly-distinct phases. The early phase, spanning roughly a century (between Castelnau and Sekelj), consists of data collected by a naturalist (Castelnau), a geographer (Coudreau), ethnographers (Ehrenreich, Krause), a missionary (Palha), a journalist (Sekelj), and government employees (Sócrates, Britto Machado), in impressionistic, pre-IPA transcriptions. The second phase, starting in the late 1950s with the Fortunes, is being conducted by professional, academically-trained linguists, either on a descriptive fashion or following theoretical trends of the day.

Of the early accounts, Ehrenreich (1894) is the first one who tries—with a relative degree of success—to provide an analysis of the data, comparing previous vocabularies with

his own (collected from Xambioá speakers), segmenting and identifying grammatical morphemes, hinting at the conservative nature of female speech, mentioning possible lexical exchanges with Xingu tribes, and discussing von den Steinen's inclusion of Karajá in his *Tapuya-Stämme*. Krause, the first researcher to visit the Javaé, provides a vocabulary of the latter, in addition to a more extensive Southern Karajá vocabulary. He also collected a text, which he attempts to segment and translate—largely unsuccessfully. By the time of Krause's visit, Dominican missionaries were already firmly established in the town of Conceição do Araguaia (which grew out of a mission founded in 1888), near the Northern Karajá and the Xambioá, where they managed to pacify local Kaiapó groups. A few decades later, one of such missionaries, Friar Luiz Palha, would produce a short essay on the Karajá language (Palha 1942), which, given all the time he spent visiting Northern Karajá and Xambioá villages, is disappointingly superficial.⁴

As one would expect, such works, based on data transcribed using the orthographies with which the authors were familiar (Portuguese, French, and German), fail to capture the phonological minutiae of the language. Furthermore, mistranslations were not uncommon, considering that the data were collected at a time when few—if any—Karajá were fully bilingual. Although their value as language documentation may be negligible, such works have important historical and diagnostic value. A few lexical items recorded in these wordlists have since fallen into disuse. For instance, the word *ĩdadzuwa* 'money', a likely

⁴ One of the Xambioá speakers I interviewed, Madalena Areraki (which was in her 80s at the time of my last visit to the Xambioá (1998)), still remembered Friar Luiz and the missionary school in Conceição do Araguaia, to where she was taken as a child. Many Xambioá and Northern Karajá children would be taken to the missionary school, an institution that further contributed to the cultural downfall of the region's Indians. There they would be offered a Christian education to wean them out of their original culture and turn them into "useful citizens"; many would end up as servants in Brazilian homes. Madalena's father, noticing how miserable she was at the school (she especially missed tobacco, obviously forbidden by the missionaries), took her back to her people (apparently, without the knowledge of the missionaries). Madalena would become the matriarch of a significant portion of the Xambioá population; her daughter, now in her 60s, is a fluent speaker of Xambioá and a strong advocate for cultural revitalization among her tribesmen.

loan from *Língua Geral* (cf. Tupinambá *itajúba*), is only documented in Socrates' (1892) vocabulary (printed as *intadiná*, for *intadiuá*); it would later be replaced by *đieru* [nie'ru], a Portuguese loan (< *dinheiro*). A Portuguese loan for 'suit, jacket', [mari'dɔ] (< Portuguese *paletó*), documented by both Machado and Sekelj, illustrates the full adaptation—not only in terms of phonological inventory, but of phonotactics as well—of a loan containing a bilabial stop before /a/; as the Karajá become more familiar with Portuguese, such adaptations tend to be less thorough (Ribeiro 2001a).⁵ Furthermore, as it will be suggested in Chapter 2 ('Female versus male speech'), Castelnau's wordlist may help put to rest a theory, set forth by Rodrigues (1999, 2004), according to which the differences between male and female speech were due to Xavánte influence. Old wordlists are also useful in discussing the use of *k*-preserving forms by male speakers (see Chapter 3) and in evaluating whether the use of non-possessed body-part terms as citation forms would have been a result of Portuguese influence (see Chapter 4).

Among the authors from the second phase, few bring additional contributions to the descriptive knowledge of the language, beyond what has been produced by the Fortunes. Maia's MA thesis, *Aspectos Tipológicos da Língua Javaé* (1986), later published as a book (1998), is a case in point. Although apparently devoted to the Javaé dialect, its title is misleading, since none of the morphological characteristics particular to Javaé are mentioned. His analysis is essentially dependent on the description of the Northern Karajá dialect provided by the Fortunes, failing to notice important grammatical phenomena such as noun incorporation and voice inflection. His transcription of the data, relying on the

⁵ As we will see (Chapter 1), Karajá does not have a voiceless bilabial stop /p/, for which /b/ is substituted in loans; however, the oral allophone of /b/ does not occur before /a/ in the native lexicon of the language. Another loan illustrating the same constraint is Xambioá [mabe'ra] 'paper' (< *Língua Geral papéira*), mentioned above (Table 1).

orthography designed by the Fortunes, also leads to analytical problems: since the current orthography does not represent stress, the use of stress shift to signal subordination went unnoticed (see Chapter 5).

As mentioned in some of the following chapters, Maia's works often illustrate the risks of doing theoretical linguistics without a solid descriptive grasp of a language. In contrast, the present dissertation has an essentially descriptive goal. However, although not subordinated to any theoretical framework in particular, it intends to contribute to theoretical linguistics by describing phenomena—previously undocumented for Karajá and other South American languages—which may help shed light into broader theoretical issues.

While there is no comprehensive grammar or dictionary of the Karajá language, there are a number of fairly-detailed accounts of Karajá culture (Ehrenreich 1891, Krause 1911, Donahue 1982, Petesch 2001, Toral 1992, Lima Filho 1994, Rodrigues 2006). Most of such anthropological monographs are very useful sources for understanding the social context in which Karajá is spoken. The differences between male and female speech, for instance, are part of a wider set of well-differentiated gender roles, whose details can be found in the anthropological literature. They also describe cultural contacts between the Karajá and neighboring tribes, largely corroborated by the loanwords described here, and offer more or less detailed accounts of the history of the Karajá-speaking peoples. Works on neighboring tribes, such as the Kaiapó (Vidal 1977, Verswijger 1992) and, especially, the Tapirapé (Baldus 1970, Wagley 1977), are also very useful.

An indirect source on the Karajá and their language are the many semi-fictitious jungle-adventure novels which became very popular in the mid-1900s (some of which even include Karajá vocabularies) by popular authors such as Willy Aureli (1940, 1963, 1964,

1966), José Mauro de Vasconcelos (1979), Hermano Ribeiro da Silva (1935), and Tibor Sekelj—who, in addition to a children’s book about the adventures of a Karajá boy, published originally in Esperanto (1979), also produced a journalistic article on the Karajá (1948), mentioned above. As one may suppose, such books are more useful as windows into national perception of the Indians than as reliable information sources on Karajá language and culture. Unlike the Xavánte (which by then were still largely ‘unpacified’), generally portrayed as treacherous savages, the Karajá are portrayed in a benign—albeit condescending—light. In part because of such popular literature, the Karajá occupy, more than other tribes of the region, a place in the national imaginary. They were even the theme of a *samba enredo* of a major Rio de Janeiro samba school, Estácio de Sá (1979), narrating the exploits of the trickster Kynyxiwe.⁶

6. Dialectal differences

In phonological terms, the Karajá dialects can be divided into two groups, the schwa-dialects (Southern and Northern Karajá) and the non-schwa dialects (Xambioá and Javaé). As shown in Table 5 below, corresponding to the schwa in Southern and Northern Karajá, Xambioá and Javaé generally present a vowel identical with the vowel in the following syllable. The schwa-dialects also present CrV syllables where Xambioá and Javaé have CVrV syllables (Table 6). Both facts—the occurrence of the schwa and of CCV syllables—are probably interconnected, as CCV syllables can be analyzed as the result of schwa syncope (see Chapter 2).

⁶ The song, entitled *Das trevas à luz do Sol: Uma odisséia Karajá*, was probably based on the anthropological literature on Karajá mythology (cf. Baldus 1951, for instance).

Table 1.5. Karajá dialects, regular phonological differences: schwa

S. Karajá	N. Karajá	Javaé	Xambioá	
<i>bədi</i>	<i>bədi</i>	<i>bidi</i>	<i>bidi</i>	‘honey’
<i>-dəkə</i>	<i>-dəkə</i>	<i>-dəkə</i>	<i>-dəkə</i>	‘causative suf.’
<i>dəka</i>	<i>dəka</i>	<i>daka</i>	<i>daka</i>	‘to tie’
<i>kəbə</i>	<i>kəbə</i>	<i>kəbə</i>	<i>kəbə</i>	‘wave (river)’

Table 1.6. Karajá dialects, regular phonological differences: CCV vs. CVCV

S. Karajá	N. Karajá	Javaé	Xambioá	
<i>krəbi</i>	<i>krəbi</i>	<i>krəbi</i>	<i>krəbi</i>	‘monkey’
<i>kre</i>	<i>kre</i>	<i>kerē</i>	<i>kerē</i>	‘kingfisher’
<i>brure</i>	<i>brure</i>	<i>bururē</i>	<i>bururē</i>	‘to tie’
<i>ibru</i>	<i>ibru</i>	<i>iburū</i>	<i>iburū</i>	‘cry, wailing’

Besides these regular phonological similarities, which could have resulted from independent developments in Javaé and Xambioá, there are also shared lexical similarities, which, for not being regular, are more suggestive of a closer relationship. Javaé and Xambioá also share neologisms coined to describe cultural items introduced by the Portuguese-speaking colonizers, such as ‘sugarcane’. Thus, it is very likely that both Javaé and Xambioá form a subgroup among the dialects. That would suggest a shared period of unity (possibly, in the Bananal Island), before the Xambioá headed to their current location. That split would certainly have taken place before the intense interethnic contact between Javaé and Tapirapé (cf. Section 4 above; see Chapter 3).

Table 1.7. Karajá dialects: a few irregular differences

S. Karajá	N. Karajá	Javaé	Xambioá	
<i>kuritʃi</i>	<i>kōritʃi</i>	<i>nawakɪ</i>	<i>nawɪkɪ</i>	‘mutum bird’
<i>lei</i>	<i>lei</i>	<i>rei</i>	<i>rei</i>	‘anaconda’
<i>dōθō</i>	<i>dōθō</i>	<i>θōθō</i>	<i>θōθō</i>	‘woodpecker’
<i>bākidɪ</i>	<i>bākidɪ</i>	<i>bɪdɪdɪ</i>	<i>bɪdɪdɪ</i>	‘sugarcane’

Contact between speakers of the four dialects was probably common—and it remains so. That is certainly the case of Javaé and Southern Karajá, a fact documented as early as 1775 by Fonseca, who reported the presence of Javaé among the Karajá he contacted in the southern tip of the Bananal Island. The presence of Javaé families among Southern Karajá populations is a tendency which persists to this day, being still common, for instance, in Aruanã (Buridina), the southernmost Karajá village. As we will see (Chapter 5), the fact that both dialects share an innovation—the further grammaticalization of a habitual marker—may be a result of contact. Contacts between the Southern Karajá and the Northern Karajá tribes of the Bananal Island were also frequent, and the same situation was probably true between the Northern Karajá from Pará and the Xambioá (at least before the latter’s population was drastically reduced). At any rate, the shallow degree of differentiation among the dialects suggests a very recent split—maybe even post-Columbian. This dissertation describes the language from a (Southern and Northern) Karajá perspective, but dialectal differences will be described whenever relevant.

7. Data

The data on which this dissertation is based were collected by me from male and female native speakers of all four dialects, mostly during field trips conducted between 1996 and 2003. The bulk of the materials consists of texts of different genres, including traditional narratives and historical and personal accounts. In an effort to document traditional manufacturing techniques (particularly endangered among the Xambioá), I have also recorded texts of a more descriptive nature, explaining how to extract nut oils, how to spin

and weave, etc.⁷ For comparative purposes, I collected standardized lexical lists, in addition to morphological paradigms and sentences. The data were transcribed and elicited with the assistance of native speakers—in particular, Ijeseberi Karajá (Southern Karajá), Sinvaldo Wahùka (Northern Karajá and Javaé), and Luiz Kurikala (Northern Karajá and Xambioá). Their particular sets of life experiences were extremely useful for the task of transcribing and analyzing the texts. The late Ijeseberi, possessor of an amazing, intuitive perceptiveness to linguistic detail, had already experience working with the SIL missionaries and with Marcus Maia; a “modern” young man, Ijeseberi was however deeply rooted in his native culture, helping me document and analyze traditional songs and explaining in detail the collected myths. Sinvaldo Wahùka, the son of a Javaé couple raised in a Northern Karajá village, is a Karajá-language teacher, a poet in his native language, and is very aware of the shortcomings of the official orthography, which we often discussed. Finally, Luiz Kurikala, a Northern Karajá teacher currently living among the Xambioá, is very familiar with both dialects and very keen to the small details that differentiate them. In general, I listened to the recordings side by side with them, making sure to note phonological details which are not captured by the current orthography. My knowledge of Karajá has also greatly benefitted from my participation, as a consultant, in several training programs for indigenous teachers (discussing, among other things, the orthography and assisting in the production of indigenous literacy materials). Thanks to their interest and generosity, every interaction with Karajá speakers has been a unique learning opportunity.

8. A brief grammatical sketch

⁷ Among the Xambioá, I interviewed all of the fluent speakers (a total of 8, by the time of my last visit (1998)), to check for the possibility of dialectal variation (since their population is a result of the gathering of different local groups) and whether there were signs of language obsolescence.

As an introduction to the language's main aspects, this section presents a sketch of Karajá, succinctly describing phenomena which will receive an in-depth treatment in the following chapters and offering an overview of grammatical aspects which will only be dealt with briefly in this dissertation (such as negation and interrogation).

8.1 Female *versus* male speech

Karajá shows differences between female and male speech to a degree that is not found in other Brazilian languages. These differences, first mentioned by Ehrenreich (1891, 1894) and studied more recently by Fortune & Fortune (1975) and Borges (1994, 1997), can generally be accounted for by regular phonological rules. As in Koasati (Haas 1964), female speech can be considered as more conservative, male speech being characterized, in general, by the deletion of a velar stop occurring in the corresponding female speech form:

(1)		Female	Male	
	a.	<i>kɔwɔɾɔ</i>	<i>ɔwɔɾɔ</i>	‘wood’
	b.	<i>kɔhã</i>	<i>ɔhã</i>	‘armadillo’
	c.	<i>ɖɪkarɔ̃</i>	<i>ɖɪarɔ̃</i>	‘I’

The deletion of the velar stop can make possible the fusion between vowels:

(2)	a.	<i>hãlɔkɔɛ</i>	<i>hãlɔɛ</i>	‘jaguar’
	b.	<i>ɾuku</i>	<i>ɾu</i>	‘night’
	c.	<i>beraku</i>	<i>bero</i>	‘river’

Although there are a few exceptions (such as, for example, the locative postposition *kɪ* and the attitude marker *ka* ‘resolution, certainty’), the deletion of the velar stop is a very productive process, applying even to borrowings from Portuguese and other indigenous languages:

(3)	a.	<i>kubɛɖã</i>	<i>ubɛɖã</i>	‘blanket’ (from Portuguese <i>coberta</i>)
-----	----	---------------	--------------	---

- b. *kabɛ* *abɛ* ‘coffee’ (from Portuguese *café*)
 c. *bãkawa* *bãawa* ‘firearm’ (from Língua Geral *mukáwa*)

Although the differences between female and male speech in Karajá have been previously studied, interesting phenomena have been traditionally overlooked. For example, the processes of vowel fusion illustrated above (♀ *beraku* :: ♂ *bero*) can apply across morpheme boundaries, a fact that can render the morphological segmentation less obvious in the male speech. This occurs, for example, with intransitive verbs, which are generally marked by the prefix *a-*. In the male speech, this prefix can be fused with the first vowel of the verb stem under the same conditions shown above:

- | | | | | |
|-----|---|-------------------------------|---|-------------------------------|
| (4) | ♀ | <i>rakubədǽreri</i> | ♂ | <i>robədǽreri</i> |
| | | Ø-r-a-kubədǽ=r-eri | | Ø-r-a-ubədǽ=r-eri |
| | | 3-CTFG-INTR-spread=CTFG-PROGR | | 3-CTFG-INTR-spread=CTFG-PROGR |
| | | ‘They are spreading.’ | | ‘They are spreading.’ |
| | | | | |
| (5) | ♀ | <i>rakədúkǽreri</i> | ♂ | <i>rədúǽreri</i> |
| | | Ø-r-a-kədúkǽ=r-eri | | Ø-r-a-ədúǽ=r-eri |
| | | 3-CTFG-INTR-go.up=CTFG-PROGR | | 3-CTFG-INTR-go.up=CTFG-PROGR |
| | | ‘S/he is going up.’ | | ‘S/he is going up.’ |

Furthermore, one of the main shortcomings of Fortune & Fortune’s (1975) and Borges’s (1994, 1997) works is their lack of a comparative, pan-dialectal perspective. Although differences between male and female speech are present throughout all the dialects (and, therefore, probably predate their diversification), all the previous studies limit themselves to one single dialect. Thus, while Ehrenreich’s pioneering remarks are likely based mainly on data of the Xambioá dialect, Fortune and Fortune’s account is based on Northern Karajá, and Borges’s on Southern Karajá.

Thus, data of Javaé, the most divergent of the four dialects, were never seriously considered. According to Fortune & Fortune (1963), the differences between male and

female speech would not occur in Javaé. Both males and females would speak what corresponds to the male speech in [Northern] Karajá. However, this information, based likely on the opinion of speakers of Northern Karajá, is not completely accurate. In fact, differences between male and female speech *do* occur in Javaé, although to a lesser extent. Interestingly, in Javaé the distinctions between female and male speech are much less remarkable than in the other three dialects, in the sense that many forms used exclusively by males in Xambioá, Northern Karajá, and Southern Karajá, are also used by females in Javaé.

Another important aspect of male vs. female speech distinctions in Karajá is the fact that there are different degrees of obligatoriness as to the deletion of /k/ in male speech. The three personal pronouns provide a useful illustration: while k-deletion is obligatory with the first-person pronoun (♀ *dɪkarə̃*, ♂ *dɪarə̃*), it does not happen with the second-person pronoun (*kai*), and it is optional with the third-person pronoun (*dəki*, ♂ *dɪi*).

Gender-related speech differences such as the ones found in Karajá (and Koasati (1964), Chiquitano (Falkinger 2002), etc.) can be seen as a matter of *gender deixis*, in the sense that, as Folley (1997) points out, “some actual linguistic elements are indexicals of some fact about gender [...]. The choice of a set of forms over the other can be said to point directly to the speaker’s sex.” Throughout this dissertation, the symbols ♀ ‘female’ and ♂ ‘male’ will be used to indicate such deictic value of a morpheme. Since forms such as *kai* and *dəki* can be used by both male and female speakers, they are left unmarked.

Chapter 3 undertakes a thorough description of the differences between female and male speech in Karajá, taking into consideration for the first time data from the four different

dialects and approaching facts that were not mentioned in previous studies. Possible scenarios for the diachronic origin of such distinctions will also be discussed.

8.2 Phonology

Besides an ‘asymmetrical’ inventory of consonants (which lacks, for example, the voiceless dental stop *t* and the voiceless bilabial *p*), the language presents a fairly large number of vowels, a characteristic of Macro-Jê languages.

Table 1.8. Consonantal inventory of Karajá

		(tʃ)	k
b	d	(dʒ)	
	f		
	θ	(ʃ)	h
	l		
w	r		

Contrasting with the previous phonological descriptions of the language (Fortune & Fortune 1963, Cavalcante 1992), my own analysis points to a larger system of vowels, in which the feature [ATR] ‘advanced tongue root’ plays a major role (Ribeiro 2000). The main difference is that I recognize a phonemic opposition between the high [+ATR] vowels /i/, /ɨ/, and /u/ and their [-ATR] counterparts /ɪ/, /ɨ̃/, and /ʊ/, a distinction not mentioned in the previous works. This distinction has pervasive consequences for the understanding of Karajá morphophonemics. Besides the existence of minimal pairs (for example, *lahi* ‘to curse’ vs. *lahɪ* ‘grandmother’, and *-uka* ‘to split’ vs. *-ʊka* ‘to cook’), the phonemic character of the opposition is also shown by the fact that [-ATR] and [+ATR] high vowels have exactly inverse behaviors in the processes of vowel harmony (2.4) and palatalization (2.5).

Table 1.9. Vocalic inventory of Karajá

oral			nasal	
i	ĩ	u	ĩ	
ɪ	ɨ	ʊ		
e	ə	o	ẽ	õ
ɛ	(ə)	ɔ		
	a		ã	

In addition to differences in the vocalic inventory, I propose a different analysis of the consonantal inventory as well. Palatal consonants, considered as independent phonemes by Fortune & Fortune (1963) and Cavalcante (1992), are very likely allophones of their non-palatal counterparts (2.5). Also questionable is the phonemic status of the schwa, which occurs only in two of the dialects, but is probably reconstructible for Proto-Karajá (2.3). Furthermore, there is no consensus on how many nasal vowels there are: Fortune & Fortune (1963) describe four (as in Table 2 above) and Cavalcante (1992) describes three (treating the nasal low vowel as an allophone of its oral counterpart), while Rodrigues (1999), in a recent survey of the Macro-Jê languages, states that Karajá “has only two nasal vowels, /ẽ/ and /õ/.” As for the nasal low vowel, Rodrigues considers it “an automatic realization of the phoneme /a/ when it either stands at the beginning of a word or is preceded by /h/ or by a voiced stop [...].” Although such a description accounts for the majority of examples containing this vowel (6), there are a number of counterexamples to it. As shown by the examples in (7) below, a number of words fitting the structural description provided by Rodrigues present an oral low vowel instead. In addition, there are a number of examples in which the nasal low vowel occurs in onsetless syllables word-medially (8), a position in which oral low vowels are also attested (9):

- | | | |
|-----|--|-------------------------------------|
| (6) | <i>hãbu</i> ‘man’ | <i>ãθu</i> ‘embaúba (type of tree)’ |
| | <i>kɔhã</i> ‘armadillo’ | <i>ã-ra</i> ‘your nephew’ |
| | <i>hãwa</i> ‘place’ | <i>ãθi</i> ‘grass’ |
| | <i>dakidã</i> ‘star’ [daki’na] | <i>wobã</i> ‘axe’ [wo’ma] |
| (7) | <i>aθara</i> ‘calango (type of lizard)’ | |
| | <i>haka</i> ‘buritirana (type of palm tree)’ | |
| | <i>waha</i> ‘my father’ | |
| (8) | <i>dohokuã</i> ‘newborn baby’ | |
| | <i>d-εãθõ</i> ‘nose’ | |
| (9) | <i>l-uahi</i> ‘medicine’ | |

Rodrigues also ignores the nasal front vowel /i/, which occurs in very few examples, such as *ãhĩ* ‘mosquito’,⁸ *hĩ* ‘(a woman’s) older brother’ (cf. *ã-hĩ* ‘your older brother’)--which forms a ‘minimal trio’ with *hi* ‘cry’ (cf. *ãhĩ* ‘your cry’), and *hi* ‘to drive away’--, and the extremely productive derivational suffix *-dĩ* [nĩ] ‘similar to’.⁹ Note that the environments in which /i/ occurs are almost the same as the ones in which /ã/ occurs: that is, onsetless syllables, after the glottal fricative /h/, and after voiced stops. Although the environments in which both /ã/ and /i/ occur are quite restricted, their distribution does not seem to be totally predictable, a fact that grants them phonemic status. This issue will be further discussed in Chapter 1.

8.2.1 Syllable and bimoraic-minimality constraint

The canonical syllabic pattern in the language is (C)V. However, in Southern and Northern Karajá, surface CCV syllables can appear as the result of the deletion of a schwa

⁸ Although the occurrence of the low-frequency vowel /i/ with *ãhĩ* could in principle be attributed to the likely onomatopoeic origin of the word, the same obviously cannot be said of its other attestations.

⁹ An exhaustive list of the occurrences of this n

occurring between a stop and the alveolar flap /r/: *kəɾɔ* ‘frog’ [kɾɔ], *kəɾɔbɪ* ‘monkey’ [kɾɔ'bɪ]. The minimal phonological word in Karajá must be bimoraic. In order to satisfy such constraint, an underlyingly monosyllabic root must, when pronounced by itself, duplicate its vowel: *bɛ* ‘water’ [bɛ'ɛ], *wa* ‘pé’ [wa'a], *tʃu* ‘sun’ [tʃu'u]. CCV-words behave, as one would expect, as bimoraic. As we will see (Chapter 2), CCV syllables also behave phonologically as bimoraic units in processes such as reduplication and hypocoristic shortening.

8.2.2 Stress

The position of the stress is predictable: in general, it falls on the last syllable of the (isolated) word (10). However, ‘minimal pairs’ for stress can appear in a phrasal level, a fact which is due to the contrast between intrinsically stressed and unstressed words (11). Furthermore, there is at least one circumstance in which stress shift is grammatically meaningful: subordination (12; see Chapter 5).

- (10) *hãbu* [ha'bu] ‘man’ *hãdĩke* [hãni'ke] ‘chicken’
kədõra [kənõ'ra] ‘sand’ *biθikərəda* [biθikre'da] ‘bike’
 (from Portuguese *bicicleta*)

- (11) a. *hãbu kɔ* [habu'kɔ] b. *hãbu=kɔ* [hã'bukɔ]
 man face man=AL
 ‘man’s face’ ‘to the man’

- (12) a. *hãbu Ø-r-Ø-urɔ=r-a* [rɔ'rɔra]
 man 3-CTFG-INTR-die=CTFG-PERF
 ‘The man died.’
- b. *hãbu Ø-r-Ø-urɔ=r-a=Ø* [rɔrɔ'ra]
 man 3-CTFG-INTR-die=CTFG-PERF=SUBORD
 ‘The man who died.’

8.2.3 Schwa

As mentioned above (Section 7), the main phonological difference among the four dialects consists in the occurrence in Southern and Northern Karajá of a schwa [ə] in unstressed positions, corresponding to environments in which Xambioá and Javaé generally present a vowel identical to the one occurring in the following syllable:

(13)	Karajá	Javaé, Xambioá	
	<i>kədɔ</i>	<i>kɔdɔ</i>	‘termite’
	<i>bədɪ</i>	<i>bɪdɪ</i>	‘honey’
	<i>-dəkə̃</i>	<i>-dəkə̃</i>	‘causative suffix’
	<i>rəkʊ</i>	<i>rʊkʊ</i>	‘gourd’

As I mentioned above, another difference is the occurrence, in the Northern and Southern Karajá dialects, of surface CCV syllables, resulting from a process of syncope of a schwa occurring between a stop and the alveolar approximant /r/ (in male speech, this schwa may surface due to the rule of *k*-dropping: δ *əɾɔ* ‘frog’, *əɾɛ* ‘kingfisher’). As expected, Xambioá and Javaé present a vowel identical to the vowel in the following syllable and no syncope takes place:

(14)	Karajá	Javaé, Xambioá	
	<i>kəɾɔ</i> [krɔ]	<i>kɔɾɔ</i> [kɔ'rɔ]	‘frog’
	<i>kəɾɛ</i> [krɛ]	<i>kɛɾɛ</i> [kɛ're]	‘kingfisher’
	<i>bəɾɔ</i> [brɔ]	<i>bɔɾɔ</i> [bɔ'rɔ]	‘back’

There are reasons to believe that Southern and Northern Karajá are, with respect to the existence of the schwa, the most conservative dialects.¹⁰ The alternative would be to consider the scenario occurring in Javaé and Xambioá as the most conservative, postulating for Southern and Northern Karajá a rule of *lenition* of an unstressed vowel when followed by a syllable containing an identical vowel. However, this hypothesis seems to be ruled out by the existence, in Southern and Northern Karajá, of minimal pairs such as ♀ *daka* ‘to take off’ versus *dəka* ‘to tie’ (corresponding to a homophonous pair in Xambioá), as well as *bɔrɔ* [bɔːrɔ] ‘stingray’ versus *bəɔrɔ* [brɔ] ‘back’ (corresponding to a homophonous pair in both Javaé and Xambioá). However, the fact that the schwa occurs only in unstressed positions makes its phonemic status problematic. This topic will be thoroughly investigated in Chapter 1.

8.2.4 Vowel harmony

Karajá presents an intricate system of vowel harmony in terms of the feature [ATR]—apparently, the first documented case of [ATR] vowel harmony in a South American language (Ribeiro 2000, 2002). Vowel harmony in Karajá can be roughly described as a process of regressive spreading of the feature value [+ATR] to [-ATR] vowels. According to their behavior in triggering, undergoing, or blocking vowel harmony, the vowels of Karajá can be grouped as in the table below:

¹⁰ That is probably a different conclusion from the one reached by the Fortunes. Although they do not discuss dialectal differences such as the ones illustrated above, their decision concerning Karajá orthography suggests that they consider Javaé and Xambioá to be more conservative. The schwa is not represented in the common orthography adopted for the four dialects. Thus, morphemes such as *dəka* ‘to tie’ and *daka* ‘to take off’ (a minimal pair in Karajá, but a homophonous pair in Xambioá) are not distinguished orthographically in any of the dialects.

Table 1.10. Vowels according to their behavior in terms of vowel harmony

Oral

[+ATR]		opaque		[-ATR]
i	ɨ	u		ɪ
e	ə	o		ɛ
		a		ɔ

Nasal

[+ATR]		opaque
ĩ		ẽ ã
		ã

As the examples in (15) below demonstrate, all combinations of vowels in a phonological word are possible, except [-ATR] vowels preceding [+ATR] vowels. This is the circumstance under which vowel harmony takes place. Thus, in (15d) the stems *dɔɾɛ* ‘parrot’ and *ɛbɔ* ‘hand’ undergo vowel harmony when followed by the stems *d-e* ‘wing’ and ♀ *kube* ‘palm’.

- | | | | | | |
|------|----|---------------------------|----------|--------------------|--------------------------|
| (15) | a. | [+ATR] [+ATR] | | b. | [+ATR] [-ATR] |
| | | <i>kube</i> [ku'be] | ‘palm’ | | <i>tʃuʃɔ</i> [tʃu'ʃɔ] |
| | | <i>kərotʃu</i> [kro' tʃu] | ‘tamal’ | | <i>ritʃɔɾɛ</i> [ritʃɔɾɛ] |
| | | | | | ‘offspring’ |
| | c. | [-ATR] [-ATR] | | | |
| | | <i>dɔɾɛ</i> [dɔ'ɾɛ] | ‘parrot’ | | |
| | | <i>dɛbɔ</i> [dɛ'bɔ] | ‘hand’ | | |
| | d. | *[-ATR] [+ATR] | | | |
| | | <i>dɔɾɛ d-e</i> [dore'de] | | <i>d-ɛbɔ</i> | <i>kube</i> [deboku'be] |
| | | parrot REL-wing | | REL-hand | palm |
| | | ‘parrot’s wing’ | | ‘palm of the hand’ | |

Vowel harmony occurs either morpheme-internally or across word boundaries, with the phonological word (characterized by a single primary stress) being its apparent domain. Any morpheme containing a [+ATR] vowel can trigger vowel harmony, regardless of its morphological status. Vowel harmony can be triggered by noun or verb roots (15d), clitics (16), or suffixes (17).

(16) ♀ *hãlɔkɔɛ=ɬi* [hãloko'etʃi]
 jaguar=LOC
 'at the jaguar'

(17) ♀ *hãlɔkɔɛ-dĩ* [hãloko'e'ni]
 jaguar-similar.to
 'domestic cat'

The fact that [-ATR] vowels can follow, but not precede [+ATR] vowels clearly shows that vowel harmony in Karajá is strictly a right-to-left process. This is further illustrated by the examples in (18) below, involving the verb stem *ɬuhɔ* 'to curse'. Since this stem contains both a [+ATR] vowel and a [-ATR] vowel, it can either trigger or undergo vowel harmony. In (18a), the vowel in the first syllable of the stem triggers vowel harmony on the antipassive prefix *ɔ-*; however, the vowel in the second syllable of the stem, as well as the vowels of the progressive auxiliary clitic remain intact. In (18b), on the other hand, the vowel in the last syllable of the stem undergoes vowel harmony, under the effect of the [+ATR] vowel of the imperfective auxiliary.

(18) a. *r-ɔ-ɬuhɔ=rɛrɪ* [rotʃu'hɔrɛrɪ] b. *r-ɔ-ɬuhɔ=r-e* [rotʃu'hɔrɛ]
 CTFG-ANTI-curse=CTFG-PROGR CTFG-ANTI-curse=CTFG-IMPERF
 'He is cursing.' 'He cursed.'

The straightforward relevance of directionality for the description of vowel harmony in Karajá may have interesting implications for theories that argue against directionality as an independent parameter of assimilation (Beckman 1995, 1997, 1998; Lombardi 1996, 1999; Bakovič 2000, etc.). Such theories treat directionality as an epiphenomenon dependent on the morphological structure of the language. Examples such as the ones presented above, in which directionality is clearly at play, are, according to Bakovič (*op. cit.* 6-8), ‘unattested’. Thus, the Karajá data are a strong counterexample to such claims, suggesting that such theories are inadequate as a universal characterization of vowel harmony phenomena (Ribeiro 2002).

8.2.5 Palatalization

As I mentioned above, one of the consequences of the analysis of the vowels proposed here is that it also makes possible to review the phonemic character of the complete series of palatals, /tʃ/, dʒ/, and ʃ/, considered independent phonemes by Fortune & Fortune (1963) and Cavalcante (1992). A careful examination of the distribution of palatal consonants in Karajá reveals that they occur in very restricted environments—generally in contiguity to high [+ATR] vowels. Thus, as the examples below show, the interdental fricative and the palatal fricative are in complementary distribution: [ʃ] occurs after the high [+ATR] vowels /u/ and /i/, whereas [θ] occurs elsewhere.

- (20) *biθa* ‘macaw’ *-kiθe* ‘scratch’ *-uθa* ‘forget’
ifa ‘bowl’ *kufe* ‘fish flour’ *rufa* ‘raw’

Therefore, the distinction between [+ATR] and [-ATR] high vowels proves to be crucial also for the analysis of the consonantal system. The allophonic nature of the variations shown above is missed if the oppositions between /i, ɨ, u/ and /ɪ, ɨ, ʊ/ are not recognized.

The study of some morphophonemic alternations also corroborates the analysis suggested above. Thus, the nominalizer suffix *-θV* (where V stands for a vowel identical to the last vowel in the verb root) is palatalized when attached to a verb root ending in the high front [+ATR] vowel /i/, but not after its [-ATR] counterpart /ɪ/:

(21)	Verb	Noun	
	<i>-aha</i>	<i>-aha-θa</i>	‘find’
	<i>-ʊɔ</i>	<i>-ʊɔ-θɔ</i>	‘fly’
	<i>-lədɪ</i>	<i>-lədɪ-θɪ</i>	‘put’
	<i>-obi</i>	<i>-obi-ʃi</i>	‘see’

The same can be postulated for the remaining palatal consonants, the affricates /tʃ/ and /dʒ/. As the examples below show, they also occur generally in contiguity to high [+ATR] vowels.¹¹ Notice the contrast with the [-ATR] high vowels /ʊ/, /ɨ/, and /ɪ/, which occur with non-palatal consonants:

(22)	<i>butʃi</i>	‘clay pot’	<i>kɔdɪ</i>	‘tobacco’
	<i>tʃi</i>	‘loc. postposition’	<i>dɪ</i>	‘bone, leg’
	<i>tʃɨ</i>	‘(a) squeeze’	<i>dɨ</i>	‘vagina’
	<i>tʃu</i>	‘sun’	<i>dʊ</i>	‘loin cloth’
	<i>ãdʒikura</i>	‘manioc’	<i>-ãdɪ</i>	‘mother’
	<i>kɨdʒɨ</i>	‘fish scale; dead skin’	<i>dɨ</i>	‘to carry (animate objects)’
	<i>hodʒu</i>	‘pole’	<i>-dʊ</i>	‘nominal suffix’

¹¹ I am aware of only one apparent exception to this rule, the word *tʃakohi*, the name of a ceremonial mask, which is very likely a Tupí-Guaraní loanword (Ribeiro 2001).

One of the sources of *tʃ* in Southern and Northern Karajá is the palatalization of the velar stop /k/ after the [high, +ATR] vowel /i/, a process that does not occur in Javaé and Xambioá (23). Again, the high front [-ATR] vowel /ɪ/ does not trigger palatalization: *bikɔwa* ‘friend’ [bikɔʷwa].

(23)	Javaé, Xambioá	Karajá	
	<i>rikɔɾɛ</i>	<i>ritʃɔɾɛ</i>	‘offspring’
	<i>ikɔɾɔ</i>	<i>itʃɔɾɔ</i>	‘fox’

A note on the representation of the palatal consonants needs to be added. Despite their lack of phonemic status--their occurrence being restricted, as I have shown, to environments contiguous to [high, +ATR] vowels--, palatal consonants will be transcribed as such in the present dissertation, for the sake of concreteness. Since a single palatal consonant may trace back to different sources, determining an underlying form may at times be an arbitrary decision, with little analytical gain and potentially detrimental consequences, both descriptively and comparatively. As it will be further discussed in Chapter 1, palatalization is not only useful as an indicator of dialectal differences, but also provides clues on the degree of morphological integration (distinguishing between compounds and phrases and between productive and fossilized morphology).

8.3 Morphosyntax

Karajá is an SOV, head-marking language. Core NP arguments—that is, subject and direct object—are not morphologically marked. Pronominal subjects are expressed by a series of free pronouns, such as ♀ *dɪkarɔ̃* ‘I’ in (28) below, while pronominal objects are

expressed by a series of bound morphemes, such as *wa-* ‘1st person’ in (29). In contrast to a fairly simple nominal morphology, Karajá presents a complex verb morphology, being traditionally described as having a very irregular fused set of prefixes indicating person, aspect, object, and direction (Fortune & Fortune 1964, Wiesemann 1986, Maia 1998). However, a more careful analysis reveals a rather regular, mostly agglutinating morphology, with separate prefixes indicating *person* (and cumulatively, *mood*), *direction*, and *valence* (28). In addition, pronominal direct objects are obligatorily incorporated into the verb (29).

(28) ♀ *dɪkarə̃ waha* *ka-r-i-rakɔ=kəre*
 I my.father¹² *I-CTFG-TRANS-wait=FUT*
 ‘I will wait for my father.’

(29) ♀ *waha* *Ø-r-i-wa-rakɔ=kəre*
 my.father 3-CTFG-TRANS-*I*-wait=FUT
 ‘My father will wait for me.’

Nominals functioning as subjects and direct objects are not morphologically marked, but there is a clear-cut distinction between subject pronouns and non-subject pronouns.

Pronominal subjects are expressed by the independent pronouns ♀ *dɪkarə̃* (♂ *dɪarə̃*) ‘I’, *kai* ‘you’, and *dəkɪ* (♂ *dɪn*) ‘he/she/it’, all of which can be pluralized by the morpheme *-boho* (*dɪkarə̃-boho* ‘we (exclusive)’, *kai-boho* ‘you (plural)’, and *dəkɪ-boho* ‘they’). Non-subject pronominal NPs (such as direct objects, objects of postpositions, and possessors) are expressed by a series of pronominal affixes (Table 11). The noun *idə̃* ‘human, people’ followed by the pluralizer *-boho* works as a first person plural inclusive pronoun. A major

¹² Although the word *waha* would appear to contain the first-person prefix *wa-*, it cannot be segmented. There are suppletive forms for both the second- (*bv* ‘your father’) and third-persons (*dəbi* ‘his father’).

difference between nouns such as *idə* and real pronouns is in the fact that independent pronouns such as *dikarə* cannot occur with postpositions or as possessors.

As the examples above suggest, Karajá is predominantly agglutinative, fusion between morphemes being fairly rare. However, under certain circumstances, person and voice/valence markers can be merged phonologically, making the morphological segmentation less obvious. Thus, in the examples below, the second person prefix *da-* fuses with the transitive marker *-* (30b) and with the intransitive marker *a-* (31):

- (30) a. *dadiwɨde*
 da-d-I-wi=d-e
 2-CTPT-TRANS-carry=2-IMPERF
 ‘You brought it.’
- b. *dewɨde*
 da-∅-I-wi=d-e
 2-CTFG-TRANS-carry=2-IMPERF
 ‘You took it away.’
- (31) ♂ *daraθɨda*
 da-∅-a-ra-θi=da
 2-CTFG-INTR-head-shave=2-PERF
 ‘Have you shaved your head?’
- aɔbo*
 aɔbo
 INTER

The distribution of the inflectional prefixes is illustrated in (32) below with the complete paradigm for the verb *wi* ‘to carry’, both in the realis and the irrealis mood.

(32) verb *wi* ‘to carry’

realis

- a. *rewɨre*
 r-a-∅-I-∅-wi=r-e
- b. *nadiwɨde*
 d-a-d-I-∅-wi=d-e¹³

¹³ The double marking of direction in the 1st person centripetal of the realis (32b) is restricted to the Southern Karajá dialect. In Javaé, Xambioá, and Northern Karajá, direction is marked only once in such cases:

- a. **Javaé, Xambioá, Northern Karajá**
ãdiwɨde
 ã-d-I-∅-wi=d-e
 1-CTPT-TRANS-3-carry=CTPT-IMPERF
 ‘I brought it.’

- | | | | |
|-----------------|---|----|--|
| | <i>CTFG-1-CTFG-TRNS-3-carry=CTFG-IMPRF</i>
'I took it away.' | | <i>CTPT-1-CTPT-TRNS-3-carry=CTPT-IMPEF</i>
'I brought it.' |
| c. | <i>dewɨde</i>
d̥a-∅-I-∅-wi=d̥-e
2-CTFG-TRANS-3-carry=2-IMPERF
'You took it away.' | d. | <i>d̥adiwɨde</i>
d̥a-d̥-I-∅-wi=d̥-e
2-CTPT-TRANS-3-carry=2-IMPERF
'You brought it.' |
| e. | <i>riwɨre</i>
∅-r-I-∅-wi=r-e
3-CTFG-TRANS-3-carry=CTFG-IMPERF
'S/he took it away.' | f. | <i>diwɨde</i>
∅-d̥-I-∅-wi=d̥-e
3-CTPT-TRANS-3-carry=CTPT-IMPERF
'S/he brought it.' |
| <i>irrealis</i> | | | |
| g. | <i>kariwɨkre</i>
ka-r-I-∅-wi=kəre
1-CTFG-TRANS-3-carry=FUT
'I will take it away.' | h. | <i>kadiwɨkre</i>
ka-d̥-I-∅-wi=kəre
1-CTPT-TRANS-3-carry=FUT
'I will bring it.' |
| i. | <i>biwɨkre</i>
b-∅-I-∅-wi=kəre
2-CTFG-TRANS-3-carry=FUT
'You will take it away.' | j. | <i>b̥adiwɨkre</i>
b̥-d̥-I-∅-wi=kəre
2-CTPT-TRANS-3-carry=FUT
'You will bring it.' |
| k. | <i>riwɨkre</i>
∅-r-I-∅-wi=kəre
3-CTFG-TRANS-3-carry=FUT
'S/he took it away.' | l. | <i>k̥adiwɨkre</i>
k̥-d̥-I-∅-wi=kəre
3-CTPT-TRANS-3-carry=FUT
'S/he will bring it.' |

Karajá presents straightforward morphological criteria for subjecthood. Thus, in the example (28) above, *d̥ikar̥ɨ* 'I' can be identified as the subject for the circumstances that (i) it

The interlinear gloss provided for the 1st person centrifugal (32a) is somewhat abstract. It is reconstructed internally, taking the 1st person centripetal as the model. Forms such as (32a) occur in Southern Karajá, Northern Karajá, and Xambioá. In Javaé, on the other hand, 1st person centrifugal forms completely parallel the centripetal form, as it can be seen in the examples below. In this sense, the Javaé dialect presents a more regular paradigm. Javaé may reflect more closely what may have existed in Proto-Karajá, or such regularity may be an innovation by analogy.

- | | |
|--------------|---|
| Javaé | |
| b. | <i>ãriwɨre</i>
ã-r-I-∅-wi=r-e
1-CTFG-TRANS-3-carry=CTFG-IMPERF
'I took it.' |
| c. | <i>ãdiwɨde</i>
ã-d̥-I-∅-wi=d̥-e
1-CTFG-TRANS-3-carry=CTFG-IMPERF
'I brought it.' |

is part of the subject pronominal series and (ii) it triggers verb agreement. Direct objects, on the other hand, can be identified by their position and by the circumstance that, when pronominal, they are obligatorily incorporated into the verb (29). However, when both subject and object are nouns, they can only be distinguished by the word order, canonically SOV:

- (33) a. ♀ hābu hālɔkɔɛ Ø-r-ɪ-rəbʉdǎ=r-a
 man jaguar 3-CTFG-TRANS-kill=CTFG-PERF
 ‘The man killed the jaguar.’
- b. ♀ hālɔkɔɛ hābu Ø-r-ɪ-rəbʉdǎ=r-a
 jaguar man 3-CTFG-TRANS-kill=CTFG-PERF
 ‘The jaguar killed the man.’

8.3.1 Possession and lexical classes

Possessed nouns are preceded by their possessors (34a, 35a). Pronominal possession is indicated by prefixes, which coincide in part with those occurring as direct objects (incorporated into the verb) or objects of postpositions. Most noun and verb stems can be divided into two lexical classes, which I will label *i-class* and *d-class*, after the third-person prefixes taken by nouns in each class. As illustrated below by the paradigms for *kɔɾʉ* ‘forehead’ (34) and *ɛbɔ* ‘hand’ (35), the differences are not limited to the third person. While *i-class* paradigms are more straightforward in terms of segmentation, *d-class* paradigms present a ‘linking prefix’ when the possessor is a noun or a first person pronominal prefix, and a zero-prefix for the second person:

- (34) a. *hābu kɔɾʉ* ‘man’s forehead’

- b. *wa-kɔɾʊ* ‘my forehead’
 c. *a-kɔɾʊ* ‘your forehead’
 d. *i-kɔɾʊ* ‘his/her/its forehead’
 e. *dʌ-kɔɾʊ* ‘his/her/its own forehead’
- (35) a. *hãbu d-ɛbɔ* ‘man’s hand’
 b. *wa-d-ɛbɔ* ‘my hand’
 c. *∅-ɛbɔ* ‘your hand’
 d. *d-ɛbɔ* ‘his hand’ or ‘his own hand’

Table 1.11. Possessive prefixes in Karajá (Ribeiro 1996)

Person	i-class	d-class
1 st	<i>wa-</i>	<i>wa-</i>
2 nd	<i>a-</i>	<i>∅-</i>
3 rd	<i>i-</i>	<i>d-</i>
3 rd REFL	<i>dʌ-</i>	

Whereas the *i*-class prefix series distinguishes a reflexive third person (*dʌ-*) from a non-reflexive one (*i-*), the *d*-class series has only one third person prefix (*d-*), which covers the range of meanings of both reflexive and non-reflexive third persons. Furthermore, *d*-class stems present a linking prefix in the first person (*wa-d-ɛbɔ* ‘my hand’), when preceded by a nominal possessor (*hãbu d-ɛbɔ* ‘the man’s hand’), or in their citation forms (*d-ɛbɔ* ‘hand’). The linking prefix has two allomorphs whose choice does not seem to be conditioned synchronically: *d-* (*d-ɛbɔ* ‘hand’, *d-õhõ* ‘pet’, etc.) and *l-* (*l-ɔdɪ* ‘throat’, *l-ɔhɪ* ‘medicine’, etc.). The function of this prefix is synchronically fairly opaque, but its distribution resembles that of the so-called *relational prefixes*, linking morphemes described for other Macro-Jê languages (Rodrigues 2000), as well as languages from the Tupí stock and the Karíb family (Rodrigues 2009). As discussed in Chapter 5, the occurrence of such linking prefixes in Karajá may provide further evidence for its inclusion in the Macro-Jê stock. Throughout this dissertation, *d*-class noun stems will be cited preceded by the linking prefix.

Although most nouns can occur, in their citation form, without a possessor, some *d*-stems may take a generic possessor prefix, *hã-*, as illustrated below by the stem *l-əbu* ‘blood’. As we will see (Chapter 6), this prefix has likely cognates in Jê and other Macro-Jê families.

- (36) a. *hãbu l-əbu* ‘the man’s blood’
 b. *hã-l-əbu* ‘blood’
 c. *wa-l-əbu* ‘my blood’
 d. *Ø-əbu* ‘your blood’
 e. *d-əbu* ‘his blood’

The distinction between both lexical classes pervades the entire morphology of the language. With verbs, the main difference between both classes is in the fact that *d*-class intransitive verbs, such as *udədə* ‘to become cold’, are marked by a zero prefix, whereas *i*-class intransitive verbs, such as *dɔdəkɛ* ‘to become hot’, are marked by the prefix *a-*:

- (37) a. *bəde* *Ø-r-Ø-udədə=r-a*
 weather 3-CTFG-INTR-become.cold=CTFG-PERF
 ‘The weather got cold.’
 b. *bəde* *Ø-r-a-dɔdəkɛ=r-a*
 weather 3-CTFG-INTR-become.hot=CTFG-PERF
 ‘The weather got hot.’

d-class transitive stems, such as *vahidə* ‘to treat’¹⁴, take the prefix *d-* when not immediately preceded by a pronominal direct object or by an incorporated noun (38a). When immediately preceded by a pronominal direct object or an incorporated noun, the relational prefix *l-* is used (38b):

¹⁴ This is a denominal verb derived from *-vahu* ‘medicine’. It is tempting to consider the prefix *d-* occurring with transitive verbs as simply a marker of 3rd person object. However, this prefix also occurs with *antipassive* constructions (54), which do not allow explicit direct objects.

- (38) a. həri wa-ritfərə Ø-r-*t*-uahi- dǎ=r-e
 shaman 1-child 3-CTFG-TRANS-3-medicine-VERB=CTFG-IMPERF
 ‘The shaman treated my child.’
- b. həri Ø-r-*t*-wa-*t*-uahi-dǎ=r-e
 shaman 3-CTFG-TRANS-1-REL-medicine-VERB=CTFG-IMPERF
 ‘The shaman treated me.’

8.3.2 Possessive predicates

There are no independent possessive pronouns (such as ‘mine’, ‘yours’, etc.) in Karajá, this function being played by the generic nouns *hōrō* ‘thing’ (39) and *d-ōhō* ‘pet’ (40), which are used in possessive statements and questions (41, 42):

- (39) kia wɛriri wa-***hōrō*** Ø-r-Ø-a=r-e
 this basket 1-***thing*** 3-CTFG-INTR-go=CTFG-IMPERF
 ‘This basket is mine.’ (Lit. ‘This basket is my thing.’)
- (40) kia idzɔrɔθa wa-d-***ōhō*** Ø-r-Ø-a=r-e
 this basket 1-REL-***pet*** 3-CTFG-INTR-go=CTFG-IMPERF
 ‘This dog is mine.’ (Lit. ‘This dog is my pet.’)
- (41) bǎ=d-***ōhō***=bo kia idzɔrɔθa Ø-r-Ø-a=r-e?
 human-REL-pet=INTERthis dog 3-CTFG-INTR-go=CTFG-IMPERF
 ‘Whose dog is this?’
- (42) bǎ=***hōrō***=bo kia wɛriri Ø-r-Ø-a=r-e?
 human=thing=INTER this basket 3-CTFG-INTR-go=CTFG-IMPERF
 ‘Whose basket is this?’

A common characteristic of lowland South American languages is the existence, with nouns, of different degrees of grammatical possession: there are obligatorily-possessed nouns (e.g. body-part terms, for instance), optionally-possessed nouns (e.g. cultural items), and nouns which cannot be directly possessed (e.g. items of nature, such as rocks, water, animals, etc.). That is a well-documented phenomenon in several Macro-Jê languages, including those of the Boróro, Jê, and Karirí families (Rodrigues 1999:190-192). In the latter case (that is, items which cannot be directly possessed), generic nouns such as the ones described above are used as possessive classifiers in indirect possessive constructions, such as 43 below, from Boróro (Crowell 1977:178). In Karajá, however, that is not the case, as any noun can in principle be directly possessed (*wa-idʒɔrɔθa* ‘my dog’, etc.).

- (43) a. i-n-ago kogariga
 1-REL-pet chicken
 ‘my chicken’
- b. i-n-o tori
 1-REL-thing stone
 ‘my stone’

Corresponding to possessive verbs such as English *have*, Karajá presents a construction involving the instrumental postposition =*dr* attached to the indefinite pronoun *idʒõ* ‘some, other’ (44a) or to the possessed noun (44b):

- (44) a. ɔ̃ dɪarɔ̃ wa-dɪeru ***idʒõ=dr=r-e***
 I 1-money some=INSTR=CTFG-IMPERF
 ‘I have money.’
- b. ɔ̃ dɪarɔ̃ wa-***dɪeru=dr=r-e***

Table 1.12. Subject agreement markers in Karajá

Person	Realis	Irrealis
1 st	<i>a-</i>	♀ <i>ka-l</i> ♂ <i>a-</i>
2 nd	<i>da-</i>	<i>bə-l/b-</i>
3 rd	<i>∅-</i>	<i>∅-</i> ; <i>kə-l</i> ♂ <i>ə</i>

- (46) a. ♀ *kraritfakre* b. ♀ *kariθuhokre*
ka-r-a-rika=kəre *ka-r-i-∅-θuhə=kəre*
1-CTFG-INTR-walk=FUT *1-CTFG-TRANS-3-wash=FUT*
‘I will walk.’ ‘I will wash it.’

According to Fortune & Fortune (1964) and Maia (1998), Karajá verbs would be divided into two different classes, active and stative, with the latter consisting essentially of predicates denoting adjectival meanings, the so-called ‘descriptive verbs.’ Descriptive predicates take exactly the same series of person markers taken by nouns (Table 11), as shown by the examples below:

- (47) a. *wa-d-εbure=r-e* b. *∅-εbure=d-e*
1-REL-get.angry=CTFG-IMPERF *2-get.angry=2-IMPERF*
‘I am angry.’ ‘You are angry.’
- c. *d-εbure=r-e*
3-get.angry=CTFG-IMPERF
‘He is angry.’
- (48) a. *wa-itʃəde=r-e* b. *a-itʃəde=d-e*
1-get.crazy=CTFG-IMPERF *2-get.crazy=2-IMPERF*
‘I am crazy.’ ‘You are crazy.’
- c. *i-itʃəde=r-e*
3-get.crazy=CTFG-IMPERF
‘S/he/it is crazy.’

As we have seen, pronominal direct objects are obligatorily incorporated into the verb. The series of direct object prefixes is partially the same that occurs with nouns and descriptive predicates:

- (49) a. dəkɪ Ø-r-ɪ-wa-rakɔ=kəre
 s/he 3-CTFG-TRANS-1-wait=FUT
 ‘S/he will wait for me.’
- b. dəkɪ a-r-a-rakɔ=kəre
 s/he 2-CTFG-2-wait=FUT
 ‘S/he will wait for you.’
- c. dəkɪ Ø-r-ɪ-Ø-rakɔ=kəre
 s/he 3-CTFG-TRANS-3-wait=FUT
 ‘S/he will wait for her/him.’

8.3.4 Direction

The verb also inflects for *direction*, according to the speaker’s point of view.

Centrifugal direction (‘thither’), marked by the prefix *r-* or by its zero-allomorph, indicates that the process is seen as occurring away from the current location of the speaker (50a).

Centripetal direction (‘hither’), marked by the prefix *d-*, indicates that the process is seen as occurring towards the current location of the speaker (50b). Centrifugal direction is the unmarked member of the opposition. All verbs are marked for direction, including those that apparently do not indicate a motion whatsoever, such as *urɔ* ‘to die’. Notice that the clitic aspectual auxiliaries also inflect for direction, and, in the 2nd person, also for person (32).

- (50) a. ♀ ka-r-ɪ-Ø-wi=kəre b. ♀ ka-d-ɪ-Ø-wi=kəre
 1-CTFG-TRANS-3-carry=FUT 1-CTPT-TRANS-3-carry=FUT
 ‘I will take it.’ ‘I will bring it.’

The system of directional markers in Karajá presents characteristics that traditionally define an inflectional category, such as obligatoriness, semantic and formal regularity, and productivity (Anderson 1985:163; Bauer 1988:73-87; Bybee 1985:11), in spite of predictions according to which direction would not occur inflectionally (Bybee 1985, Talmy 1985). A possible explanation for the *sui generis* nature of the phenomenon may be found in the wide range of pragmatic and discourse functions that may be played by directional marking in Karajá: as described in Chapter 3, besides its strictly directional use, directional inflection can be used with evidential purposes, as well as to code empathy relationships between the participants of the speech act and between narrator and characters in a narrative text.

8.3.5 Valence and voice

Karajá verbs are lexically either transitive or intransitive. Intransitive verbs (8.3.5.1) may have their valence increased through causativization or through oblique promotion, which are derivational processes. Transitive verbs (8.3.5.2), on the other hand, may have their valence decreased through reflexivization, passivization, and antipassivization.

8.3.5.1 Intransitive verbs

Intransitive verbs are those that do not take a direct object as one of their arguments, such as *-udədə* ‘to become cold’ (38a), *dədəkε* (*ɔ̃ dədεε*) ‘to become hot’ (38b), and ♀ *ritfa* (*ɔ̃ ritfa*) ‘to walk’ (40a). As we have seen above, *i*-class intransitive verbs are generally

marked by the prefix *a-*, while *d*-class intransitive verbs are marked by a zero allomorph. The overwhelming majority of verb stems can be assigned to either class; a few stems, however, are marked by less-common prefixes. The class of intransitive verbs includes not only one-place verbs such as *rika* ‘to walk’ and *-udədə* ‘to become cold’, but also verbs such as *-obi* ‘to see’, whose arguments are oblique NPs—in this case, a locative, marked by the postposition *bə* ‘diffuse locative’ (51). Although notionally transitive, such verbs behave as intransitive for all purposes. For example, they cannot be made passive or antipassive, and their arguments cannot be incorporated.

- (51) ♀ *dɪkarə halokɔemə rabire*
 dɪkarə halokɔe=bə *r-a-Ø-obi=r-e*
 I jaguar=LOC CTFG-*I*-INTR-see=CTFG-IMPERF
 ‘I saw the jaguar.’

Most intransitive verbs can be transitivized, either through causativization or through the promotion of an oblique to direct object. The transitivized stem is formed by the nominal form of the verb plus the verbalizer suffix *-də*. This is illustrated in the example (52b) below, where the intransitive verb *uka* II ‘to be cooked’ is transitivized:¹⁶

- (52) a. *iweru rukareri*
 iweru Ø-r-Ø-uka=r-eri
 calugi 3-CTFG-INTR-be.cooked=CTFG-PROGR
 ‘The *calugi* (a kind of porridge) is cooking.’

¹⁶ This example illustrates a very common process for deriving nouns from verb roots, namely *consonantal replacement*, which consists in replacing a velar stop or a glottal fricative occurring in the last syllable of the verb root with an alveolar flap: *rika* I ‘to walk’ > *rira* ‘the action of walking’, *rira-dā* ‘walking place’, *rira-du* ‘the one who walks’; *θohɔ* I ‘to wash’ > *θorɔ* ‘the action of washing’, *θorɔ-dā* ‘washing material’, *θorɔ-du* ‘the one who washes’ (Ribeiro 1996). Thus, the transitive stem in (46b) above is constructed with the nominal form of the verb *uka* ‘to be cooked’, *-ora* ‘the action of cooking’, followed by the verbalizer suffix *-də*.

- b. *ahāwəki iweru rɪdʊranəkre*
 a-hāwəki iweru Ø-r-ɪ-d-ɪra-də̃=kəre
 2-woman calugi 3-CTFG-TRANS-3/REL-be.cooked-VERB=FUT
 ‘Your wife will cook the *calugi*.’

8.3.5.1.1 Causativization

Causative stems derived from unergative verbs, such as *rika* ‘to walk’, are formed with the causativizer suffix *-dəkə̃* plus the verbalizer suffix *-də̃* (53). However, the causative suffix does not occur in causative stems derived from unaccusative verbs, such as *-vka* ‘to be cooked’ in (52b) above.

- (53) *hābu kuladʊ Ø-r-ɪ-rira-dəkə̃-də̃=r-ɛrɪ*
 man child 3-CTFG-TRANS-walk-CAUS-VERB=CTFG-PROGR
 ‘The man is making the child walk.’

8.3.5.1.2 Oblique promotion

With a few semi-transitive verbs (i.e., those which take non-canonically marked objects, such as an allative or dative argument), such as *-ə̃kə̃rafi* ‘to ask’, transitivity results in the promotion of the former oblique argument to direct object (examples from the Xambioá dialect):

- (54) a. *hawiki dʌ-rikərə=kɔ Ø-r-Ø-ə̃kə̃rafi=r-e*
 woman 3REFL-offspring=AL 3-CTFG-INTR-ask=CTFG-IMPERF
 ‘The woman asked her son.’
- b. *hawiki dʌ-rikərə Ø-r-ɪ-d-ə̃kə̃raθi-də̃=r-e*
 woman 3REFL-offspring 3-CTFG-TRANS-d-ask-VERB=CTFG-IMPERF
 ‘The woman questioned her son.’

8.3.5.2 Transitive verbs

Transitive verbs are those that take a direct object as one of their arguments. In Karajá, transitive verbs are always marked by the prefix *r-*. As I mentioned above, both transitive and intransitive valence prefixes may fuse with the preceding personal prefix under certain circumstances, such as in the 2nd person in the centrifugal direction of the realis mood (55a). Notice that there is no fusion in the centripetal direction (55b).

- (55) a. *dɛrakodɛ*
dʌ-∅-r-∅-rakɔ=d-e
 2-CTFG-TRANS-3-wait=2-IMPERF
 ‘You waited for him (thither).’
- b. *dʌdrakodɛ*
dʌ-d-r-∅-rakɔ=d-e
 2-CTPT-TRANS-3-wait=2-IMPERF
 ‘You waited for him (hither).’

8.3.5.2.1 Reflexive

There are two allomorphs of the reflexive morpheme, *ɛʃi-* and *iʃi-*. The former is incorporated into the verb, when the NP co-referential with the subject is a direct object (56a). The latter is attached to postpositions, when the coreferential NP is an oblique (56b).

- (56) a. ♀ *dɪkarɔ̃ ka-r-ɛʃi-θuɬɔ=kərə*
 I 1-CTFG-REFL-wash=FUT
 ‘I will wash myself.’
- b. *habu iʃi=bɔ̃ ∅-r-∅-obi=r-e*
 man REFL=LOC 3-CTFG-INTR-see=CTFG-IMPERF
 ‘The man saw himself.’

8.3.5.2.2 Passive

Passive verbs are marked by the prefix *a-*, with class I stems such as *θuhɔ* ‘to wash’ (57b), or its zero allomorph, with class II stems, such as *-uka* ‘to split’ (57b). Notice that this is apparently the same prefix that occurs with basic intransitive verbs such as *rika* I ‘to walk’ and *-obi* ‘to see’. With transitive roots, however, this prefix will always convey a passive or anticausative meaning.

- (57) a. d-ãdɪ wa-dəki Ø-r-ɪ-θuhɔ=r-ɛɪɪ
REL-mother 1-clothes 3-CTFG-TRANS-wash=CTFG-PROGR
‘My mother is washing my clothes.’
- b. wa-dəki Ø-r-a-θuhɔ=r-ɛɪɪ
1-clothes 3-CTFG-PASS-wash=CTFG-PROGR
‘My clothes are being washed.’
- (58) a. ♀ kɔdʊ hãlɔkɔɛkɔɾʊ Ø-r-ɪ-tʃ-uka=r-e
 turtle jaguar forehead 3-CTFG-TRANS-tʃ-split=CTFG-IMPERF
 ‘The turtle split the jaguar’s forehead.’
- b. ♀ hãlɔkɔɛ kɔɾʊ Ø-r-Ø-uka=r-e
 jaguar forehead 3-CTFG-PASS-split=CTFG-IMPERF
 ‘The jaguar’s forehead was split.’

In the passive construction, the original O becomes the subject, as in languages such as English, for example. However, unlike English, where the agent in a passive construction can be expressed as an oblique (‘by-phrase’), in Karajá the agent, although sometimes implicit, cannot be expressed at all. Thus, passives in Karajá are both a *backgrounding* construction, functioning to delete unknown or irrelevant subjects, and a *foregrounding* construction, since it results in the promotion of the original O to subject position (Foley and Van Valin 1985).

8.3.5.2.3 Antipassive

Antipassive is a phenomenon typical of ergative languages, corresponding functionally to a ‘mirror image’ of the passive construction in nominative-accusative languages (Silverstein 1976). In a syntactically ergative language, “while the A and the O in an ergative clause are marked as ergative and absolutive respectively, the A in an antipassive is typically coded as an absolutive NP, and the O (if present) appears in a case other than the absolutive” (Cooreman 1994: 50).¹⁷ Although some authors, such as Cooreman, limit the discussion of antipassive constructions to ergative languages, nominative-accusative languages may also present *backgrounding* antipassives, which “function to demote the undergoer to peripheral status” (Foley and Van Valin 1985: 338). This is what occurs in Karajá, where antipassive, marked by the prefix ɔ-, results in the deletion of an unknown or irrelevant direct object:

(59) d-ãdi Ø-r-ɔ-θuhɔ=rɛri
REL-mother 3-CTFG-ANTI-wash=CTFG-PROGR
‘My mother is washing (something).’

(60) hãbu Ø-r-ɔ-tʃ-uka=r-ɛri
man 3-CTFG-ANTI-d-split=CTFG-PROGR
‘The man is splitting (something).’

As these examples show, antipassive in Karajá is not promotional (or foregrounding), in the sense that the A remains in the same syntactic relation it occupies in the corresponding active, transitive voice. Furthermore, the antipassive construction in Karajá does not allow

¹⁷ I will follow Cooreman in adopting Dixon’s (1979) use of the labels A and O to refer to the two participants in a two-participant clause—prototypically, the agent and the patient, respectively.

the expression of the demoted O whatsoever, which is an interesting parallel with what happens to the agent in the passive construction.

8.3.6. Noun incorporation

Noun incorporation in Karajá is a process by which the head of the absolutive noun phrase is inserted into the verb, thereby forming a compound. The more productive pattern of noun incorporation involves only body-part terms, which are in general inherently possessed nouns. Since only the head of the absolutive noun phrase is incorporated, the valence of the resulting noun-verb compound remains unaltered, as the possessor is promoted to subject with intransitive, unaccusative verbs such as *bɔhɔ* ‘to break’ (61), or to object with transitive verbs such as *dɔka* I ‘to tie’ (62):¹⁸

- (61) a. *idə we rɪkɪ ɖai Ø-r-a-bɔhɔ=r-e*
 people belly NARR 3.LOC 3-CTFG-INTR-break=CTFG-IMPERF
 ‘The people’s bellies were broken there, it is said.’
- b. *idə rɪkɪ ɖai Ø-r-a-we-bɔhɔ=r-e*
 people NARR 3.LOC 3-CTFG-INTR-belly-break=CTFG-IMPERF
 ‘The people’s bellies were broken there, it is said.’

¹⁸ Examples (55a) and (55b) are from the Javaé dialect. Although Maia (*op. cit.*: 63) claims that object incorporation does not occur in Javaé, noun incorporation seems to be as common in Javaé as it is in the other three dialects. The example below, involving the incorporation of the noun *dɪkɔhɔ* I ‘knee’ to the transitive verb *we* I ‘to penetrate’, occurs in the same text from which the examples above were obtained:

- (a) *rɪdɪkɔhɔwɛrɛ*, *ɪdɪkɔhɔ* *riwɛrɛ*
 Ø-r-I-Ø-dɪkɔhɔ-wɛ=r-e *ɪ-dɪkɔhɔ* Ø-r-I-wɛ=r-e
 3-CTFG-TRANS-3-knee-penetrate=CTFG-IMPERF 3-knee 3-CTFG-TRANS-penetrate=CTFG-IMPERF
 ‘[He] stabbed him in the knee, he stabbed his knee.’

- (62) a. kədǽjiwɛ kuθehewe dɪ Ø-r-ɪ-dʔkɑ=r-e
 K. rhea leg 3-CTFG-TRANS-tie=CTFG-IMPERF
 ‘Kynyxiwè tied the legs of the rhea.’
- b. kədǽjiwɛ kuθehewe Ø-r-ɪ-dɪ-dʔkɑ=r-e
 Kynyxiwè rhea 3-CTFG-TRANS-leg-tie=CTFG-IMPERF
 ‘Kynyxiwè tied the legs of the rhea.’

A morphologically similar construction involves the combination of body-part terms with predicates in a fashion similar to what Mithun (1984: 863) labels ‘classificatory noun incorporation’ (57b). In such cases, the noun stems are body-part terms that ordinarily function as *measure terms* (57a), such as *ra* ‘head’ (measure term for potatoes and yams), *ru* ‘eye’ (measure term for grains), and *wɛ* ‘belly’ (measure term for round fruits):

- (63) a. adǽdǽ i-wɛ θohodi
 pineapple 3-belly one
 ‘one pineapple’
- b. adǽdǽ Ø-r-ɪ-wɛ-kʔkɑ=r-ɛɪɪ
 pineapple 3-CTFG-TRANS-belly-grate=CTFG-PROGR
 ‘She is grating pineapple.’

In addition to body-part terms, the reciprocal morpheme *wi* and the semantically-empty noun *bǽdɛ* can also be incorporated, with specific sets of verbs. Noun incorporation will be more thoroughly discussed in Chapter 3.

8.3.6.1 Interactions between voice and noun incorporation

As we have seen in above, noun incorporation in Karajá is generally a valence-preserving process. Therefore, since a transitive verb remains transitive after having incorporated a noun, it can still be made passive (64b) or antipassive (64c):

- (64) a. ♀ d-ãdɪ wa-ritʃɔɾɛ Ø-r-ɾ-radɛ-kəɾɔ=r-ɛɪɪ
REL-mother 1-offspring 3-CTFG-TRANS-hair-cut=CTFG-PROGR
‘My mother is cutting my child’s hair.’
[Lit. ‘My mother is hair-cutting my child.’]
- b. ♀ wa-ritʃɔɾɛ Ø-r-a-radɛ-kəɾɔ=r-ɛɪɪ
1-offspring 3-CTFG-PASS-hair-cut=CTFG-PROGR
‘My child’s hair is being cut.’
[Lit. ‘My child is being hair-cut.’]
- c. d-ãdɪ Ø-r-ɔ-radɛ-kəɾɔ=r-ɛɪɪ
REL-mother 3-CTFG-ANTI-hair-cut=CTFG-PROGR
‘My mother is cutting hair.’
[Lit. ‘My mother is hair-cutting (*someone*).’]

Examples such as (58c), in which antipassive markers can co-occur with an incorporated noun, may have interesting implications for theories that treat antipassive as a special kind of noun incorporation, such as seen in Baker’s approach (1988). As I have suggested elsewhere (Ribeiro 2001a), if antipassive is to be treated as a matter of noun incorporation, the interaction between antipassive and noun incorporation in Karajá provides a strong counterexample to Baker’s claim against the occurrence of multiple incorporations.

8.7 Attitude markers

Karajá presents a number of discourse-oriented particles, indicating mostly the attitude of the speaker in relation to what he or she is uttering—examples of which are =kɔɾɪ

‘admiration, surprise’, =*θd̥* ‘boredom, repetitiveness’, =*θɔ* ‘excitement’, =*ləh ə* ‘compassion’, and =*kɔθəd̥* ‘doubt’. These *attitude markers* (Sadock & Zwicky 1985: 161) are, in general, unstressed, and appear cliticized to the first syntactic constituent of the sentence. Their distribution is exemplified by the particle =*θɔ* in the example below:

- (65) a. ♀ *bikuθɔ* *kədɔkakra*
 biku=*θɔ* kə-d-Ø-ɔka=*kərə*
 rain=*EXCIT* 3-CTPT-INTR-rain=*FUT*
 ‘It will rain (and that’s great).’

8.8 Tense

Tense does not occur as an inflectional category in Karajá. Instead, tense distinctions are encoded by temporal-aspectual auxiliaries and particles that cliticize to the main verb. As with main verbs, the auxiliaries *-ɛrɪ* ‘progressive’, *-a* ‘perfective’, and *-e* ‘imperfective’ inflect for direction (and, in the second person, also for person), as shown by the examples above. The particles =*kərə* ‘future’, =*kɛ* ‘potential’, and =*hed̥* ‘admonitory’, restricted to the irrealis mood, are invariable. A finite verb in Karajá is typically inflected for all categories (*person, direction, and voice/valence*) and followed by a temporal-aspectual auxiliary or particle.

Table 1.13. Tense/aspect clitics in Karajá

Auxiliaries	Particles
= <i>(r)e</i> ‘imperfective’	= <i>kərə</i> ‘future’
= <i>(r)a</i> ‘perfective’	= <i>kɛ</i> ‘potential’
= <i>(r)ɛrɪ</i> ‘progressive’	= <i>hed̥</i> ‘admonitory’

These tense and aspect markers have traditionally been considered as inflectional affixes (Fortune & Fortune 1964, Fortune 1973; Maia 1998). However, a more recent analysis has shown that they are, in fact, clitics (Ribeiro 1996). As such, they can attach to any element occurring as a predicate, including postpositional phrases (66), pronouns (67), and nouns (68):

- (66) ♂ wad̥au d̥əbõra=u=r-e
 Watau youth=TEMP=CTFG-IMPERF
 ‘It was during Watau’s youth.’
- (67) ♀ dɪkarõ=kəre
 I=FUT
 ‘It will be me.’
- (68) ♂ dzuhu=rəkɪ hʊrəθõ=rəkɪ idõ hɛɔd̥i=r-ɛdõ=r-e
 before=NARR lightning.bugs people fire=CTFG-PL=CTFG-IMPERF
 ‘It is said that, in the old times, lightning bugs were the fire of mankind.’

Auxiliaries and particles belong to clearly-distinct grammatical categories, based on semantic, morphological, and distributional criteria. Semantically, particles are more properly described as modals (future, potential), while auxiliaries are more properly described as aspectual markers (progressive, perfective, imperfective). Morphologically, particles are invariable, while auxiliaries are inflected, forming a defective sub-class of verbs. Furthermore, auxiliaries and particles can co-occur (a fact which was not mentioned in any previous description of Karajá), demonstrating that they both occupy different positional slots in the complex verbal word. Although the differences between =r-a and =r-e are traditionally described as one of remoteness (‘recent past’ vs. ‘remote past’, respectively) by both Fortune and Maia, that is certainly an inadequate characterization, given that =r-e

occurs mainly in descriptive constructions, narratives (as a sort of historical present), and even to denote future events (signaling habituality, versus punctuality; cf. (69d)). Chapter 4 further discusses the semantics of the temporal-aspectual morphemes. A habitual morpheme, *=bãhã*, resulting from the grammaticalization of a converb construction, will be described in Chapter 5.

- (69) a. *ɗəkɪ* *kəɗura* \emptyset -r-i-rɔ=**r-a**
 he fish 3-CTFG-TRANS-eat=**CTFG-PERF**
 ‘He’s eaten (the) fish.’
- b. *ɗəkɪ* *kəɗura* \emptyset -r-i-rɔ=**r-e**
 he fish 3-CTFG-TRANS-eat=**CTFG-IMPERF**
 ‘He ate the fish.’
 ‘He eats fish.’
- c. *ɗəkɪ* *kəɗura* \emptyset -r-i-rɔ=**kre**
 he fish 3-CTFG-TRANS-eat=**FUT**
 ‘He will eat the fish.’
- d. *ɗəkɪ* *kəɗura* \emptyset -r-i-rɔ=**r-e=kre**
 he fish 3-CTFG-TRANS-eat=**CTFG-IMPERF=FUT**
 ‘He is going to be eating fish.’
 (i.e. ‘He will live on fish.’)

8.9 Number

Number is not an inflectional category in Karajá, being optional for both nouns and most verbs. In nouns, plurality is indicated by at least three different devices: reduplication, the use of the noun *bãhãɗv* ‘group, crowd’, and the attachment of the pluralizer *=boho*. Plurality in the verb is marked by the auxiliary *-edã*, which presents the same inflectional properties of the temporal-aspectual auxiliaries mentioned above. Although it agrees only with the subject, the auxiliary *-edã* can express the plurality of the subject, the object, or

both, as the different translations for the example (70) below show. This morpheme occurs also with oblique pronominal objects (71).

- (70) itʃɔrɔθa Ø-r-I-wa-rɔ=r-ɛdǎ=r-e
 dog 3-CTFG-TRANS-1-bite=CTFG-PLURAL=CTFG-IMPERF
 ‘The dogs bit me/The dog bit *us*/The dogs bit *us*.’
- (71) hǎbu wa-bǎ=r-ɛdǎ Ø-r-Ø-obi=r-e
 man 1-LOC=CTFG-PL 3-CTFG-INTR-see=CTFG-IMPERF
 ‘The man saw *us*.’

8.10 Postpositions

There are at least eleven postpositions in Karajá (Table 14). Except for the reflexive (marked by the prefix *ij/i-*), postpositions generally take the same series of prefixes as i-class stems (listed in Table 4 above). The only exception is the postposition *bǎ*, which takes the prefix *u-* in the second person (the only occurrence of this prefix whatsoever). Some postpositions present suppletive pronominal forms corresponding to the third person. Most of the postpositions are unstressed, with the exception of the dative *dǎkɛ* and the the evitative *laku*.

Table 1.14. Karajá postpositions

<i>Postposition</i>	<i>Approximate meaning</i>	<i>3rd person suppletive pronouns</i>
= <i>kɪ</i>	‘stationary locative’	<i>dai</i>
= <i>tʃi</i>	‘dynamic locative’	
= <i>bǎ</i>	‘diffuse locative’	<i>dʊ</i>
♀ = <i>ku</i> (♂ = <i>u</i>)	‘temporal’	
= <i>rɔbɪ</i>	‘ablative’	
= <i>dɪ</i>	‘instrumental’	
= <i>dǎkɛ</i> (♂	‘dative’	<i>dabǎ</i>

=dεε)		
=kɔ (ɔ̂ =ɔ)	‘allative’	dabã
♀ =laku (ɔ̂	‘evitative’	
=lau)		
=wadã	‘comitative’	
=wəθe	‘comparative’	

In most languages where there are distinctions such as the ones between i-class stems and d-class stems, postpositions are also divided between both classes. That may also have been the case in Karajá. Notice that all suppletive third-person forms begin with /d/, suggesting that such ‘pronouns’ are relics of d-class postpositional paradigms. Notice also that, if this theory is correct, a second-person form corresponding to *dʊ* would have been **∅-ʊ*; that may be the origin of the second-person prefix *ʊ-* mentioned above.

Karajá does not seem to have adverbs as an independent part-of-speech, adverbial functions being performed by postpositional phrases. Consequently, corresponding to adverbial clauses, Karajá presents nominalized clauses occurring as objects of postpositions. Example (72) below illustrates the use of the comitative postposition *=wadã* to form conditional clauses (example from the Xambioá dialect). Given the importance of postpositions in forming adverbial subordinate clauses, their semantics will be further described in Chapter 5.

- (72) *idã ∅-r-a-ri=r-e=wadã*
 people 3-CTFG-INTR-leave=CTFG-IMPERF=COM
- idã=deke=ka ki-d-∅-iri=kere*
 people=DAT=ASSERT 3-CTPT-INTR-call.out=FUT
 ‘If anyone was left, do call out to us!’

8.11 Negation

There is one basic negative morpheme in Karajá, *-kõ*; suffixed to the verb, it negates the entire sentence (73). It can also occur with the pro-form ♀ *adõ* ‘thing’ (♂ *aõ*), forming a negative particle which negates individual constituents of the sentence (74) and also occurs in answers to yes/no questions (75b). The negative morpheme, followed by the imperfective clitic, can also be used in such circumstances (75c). An emphatic morpheme, *hiki* (♂ *hiki*), occurs with *adõkõ* or a verb to convey a strong negation (75d).

- (73) ♂ dəkɪ ɔhã Ø-ɪ-ɪ-rɔ=**kõ**=ɪ-e
 he armadillo 3-CTFG-TRANS-eat=**NEG**=CTFG-IMPERF
 ‘He doesn’t eat armadillo.’
- (74) ♂ dɪarẽ **aõkõ**, kai=dɪa
 I **NEG** you=ASSERT
 ‘Not me, but you instead.’
- (75) a. ♂ *kaiaõbo* *ɔhã* *d'erode*
 kai=aõbo ɔhã dɪa-Ø-ɪ-rɔ=d-e
 you=INTER armadillo 2-CTFG-TRANS-eat=2-IMPERF
 ‘Do you eat armadillo?’
- b. ♂ aõkõ
 NEG
 ‘No.’
- c. ♂ kõ=r-e
 NEG=CTFG-IMPERF
 ‘No.’
- d. ♂ aõkõ-hiki
 NEG-EMPH
 ‘Not at all.’

- c. ♀ *bõ=ritʃɔɾɛ=bo* ♂ *bõ=riɔɾɛ=bo*
 ANIM=child=INTER
 ‘Whose child?’
- d. ♀ *bõ=ritʃɔɾɛ=dəkɛ=bo* ♂ *bõ=riɔɾɛ=dɛɛ=bo*
 ANIM=child=DAT=INTER
 ‘To whom’s child?’

The interrogative word ♀ *adõ=bo* (♂ *aõ=bo*) occurs as a second-position clitic to form polar (“yes/no”) questions (79). The interrogative morpheme can also occur by itself, as a second-position clitic, to situate the focus of the question on specific constituents of the sentence (80).

- (79) ♀ *kaianõbo inã ribe dekeride*
*kai=**adõbo** idã ribe* *ɔa-Ø-i-keri=d-e*
 you=**INTER** people speech 2-CTFG-TRANS-know=2-IMPERF
 ‘Do you know Karajá?’
- (80) ♂ *kaiwe inã ribe dekeride*
*kai=**we** idã ribe* *ɔa-Ø-i-keri=d-e*
 you=**INTER** people speech 2-CTFG-TRANS-know=2-IMPERF
 ‘Are **you** the one who knows Karajá?’

8.13 Numerals

Karajá has numeral words (or constructions) for the concepts ‘one’ to ‘twenty’. For numerals ‘one’ to ‘five’, there are single words (81). Some of these numerals seem to be made up of more than one identifiable, albeit fossilized, morpheme, including third person prefix *i-*, *bikɔwa* ‘companion’, *=dõ* ‘indefinite article’ (as suggested by stress and its behavior in male speech), and *kire* ‘half’. The use of words such as ‘companion’ with the numeral ‘four’ is also found in other lowland South American languages (Green 1997), and could be an areal phenomenon. The numeral ‘five’ could arguably (and rather speculatively)

be segmented as *i-ru-kire* [3-eye-half] ‘the half (of both hands?)’, a descriptive construction including *ru* ‘eye’, the classifier for ‘round, small things’ such as *beans*, *beads*, and *coins*.

(81) θ ohodʒi
‘one’¹⁹

(82) idātʃi
‘two’

(83) ♀ idāʃadō ♂ idāʃaō
‘three’

(84) ♀ idākubikowa ♂ idāubiowa
‘four’

(85) ♀ irukire ♂ iruire
‘five’

Higher numerals involve increasingly complex constructions—in fact, entire sentences including the verb form ♀ *rekurə* ‘(it/one) crossed’ (♂ *reurə*). Thus, ‘six’ is literally ‘(it) crosses to one (at the other) hand’; ‘sixteen’ is ‘(it) crosses to one at the foot’, and so forth. The data for the higher numerals are given here in a very careful—didactic, even--form. In casual speech, and especially in itemized counting, both the allative postposition =*kə* and the verb can be suppressed. Depending on factors such as speed and informality, constructions such as these can be pronounced as a single phonological word (in which case vowel harmony may take place) or as two separate words: e.g. [dɛboʔoʔoʔdʒi] ~ [dɛʔbo ʔoʔoʔdʒi] ‘six’. Currently, the overwhelming tendency is to have such constructions

¹⁹ In Javaé, the numeral ‘one’ is lohodʒi; the correspondence Javaé /l/ :: Karajá /θ/ is not a regular one.

replaced by Portuguese numerals. Numeral words generally precede the quantified noun
(101).

- | | | | | |
|------|---|------------------------|-------------|-------------------|
| (86) | ♀ | d-εbɔ=kɔ | θohodzi | ∅-r-∅-ekuro |
| | ♂ | d-εbɔ=ɔ | θohodzi | ∅-r-∅-euro |
| | | REL-hand=AL
'six' | one | 3-CTFG-INTR-cross |
| (87) | ♀ | dεbɔ=kɔ | idātʃi | ∅-r-∅-ekuro |
| | ♂ | dεbɔ=ɔ | idātʃi | ∅-r-∅-euro |
| | | REL-hand=AL
'seven' | two | 3-CTFG-INTR-cross |
| (88) | ♀ | d-εbɔ=kɔ | idāɖadõ | ∅-r-∅-ekuro |
| | ♂ | d-εbɔ=ɔ | idāɖaõ | ∅-r-∅-euro |
| | | REL-hand=AL
'eight' | three | 3-CTFG-INTR-cross |
| (89) | ♀ | d-εbɔ=kɔ | idākubikɔwa | ∅-r-∅-ekuro |
| | | d-εbɔ=ɔ | idāubɔwa | ∅-r-∅-euro |
| | | REL-hand=AL
'nine' | four | 3-CTFG-INTR-cross |
| (90) | | d-εbɔ | i-dʊε | |
| | | REL-hand | 3-finish | |
| | | 'ten' | | |
| (91) | ♀ | wa=kɔ | θohodzi | ∅-r-∅-ekuro |
| | ♂ | wa=ɔ | θohodzi | ∅-r-∅-euro |
| | | foot-AL
'eleven' | one | 3-CTFG-INTR-cross |
| (92) | ♀ | wa=kɔ | idātʃi | ∅-r-∅-ekuro |
| | ♂ | wa=ɔ | idātʃi | ∅-r-∅-euro |
| | | foot=AL
'twelve' | two | 3-CTFG-INTR-cross |
| (93) | ♀ | wa=kɔ | idāɖadõ | ∅-r-∅-ekuro |

	♂	wa=ɔ	idāɗaō	Ø-r-Ø-euro	
		foot-AL	three	3-CTFG-INTR-cross	
		'thirteen'			
(94)	♀	wa=kɔ	idākɔbikɔwa	Ø-r-Ø-ekuro	
	♂	wa=ɔ	idāɔbrɔwa	Ø-r-Ø-euro	
		foot=AL	four	3-CTFG-INTR-cross	
		'fourteen'			
(95)	♀	wa=kɔ	irukire	Ø-r-Ø-ekuro	
		wa=ɔ	iruire	Ø-r-Ø-euro	
		foot=AL	five	3-CTFG-INTR-cross	
		'fifteen'			
(96)	♀	wa	wi=kɔ	θohodzi	Ø-r-Ø-ekuro
	♂	wa	wi=ɔ	θohodzi	Ø-r-Ø-euro
		foot	both-ALL	one	3-CTFG-INTR-cross
		'sixteen'			
(97)	♀	wa	wi=kɔ	idātʃi	Ø-r-Ø-ekuro
	♂	wa	wi=ɔ	idātʃi	Ø-r-Ø-euro
		foot	both	two	3-CTFG-INTR-cross
		'seventeen'			
(98)	♀	wa	wi=kɔ	idāɗadō	Ø-r-Ø-ekuro
		wa	wi=ɔ	idāɗaō	Ø-r-Ø-euro
		foot	both=AL	three	3-CTFG-INTR-cross
		'eighteen'			
(99)	♀	wa	wi=kɔ	idākɔbikɔwa	Ø-r-Ø-ekuro
	♂	wa	wi=ɔ	idāɔbrɔwa	Ø-r-Ø-euro
		foot	both=AL	four	3-CTFG-INTR-cross
		'nineteen'			
(100)		wa	idʊe		
		foot	3-finish		
		'twenty'			
(101)	♂	idāɔbrɔwa	əriθa	udī	

four Xavánte spirit
'four Xavánte spirits'

9. Table of contents

As I intended to show in this introduction, many aspects of Karajá phonology, morphology, and syntax (as well as its peculiar system of gender deixis) remain to be thoroughly understood. However, even a succinct overview of the language reveals a number of phenomena rich in theoretical and descriptive implications. It is my hope that this dissertation will provide us with a better understanding not only of the Karajá language, but also of the general analytical questions its study raises. The following chapters provide an in depth description of phenomena which are central to the understanding of Karajá.

Chapter 1 provides a thorough account of the language's phonological inventory, syllable and stress patterns, and morphophonology. A major difference between the account given here and those offered in previous sources lies in the vocalic inventory: here it is shown that the feature [ATR] 'advanced tongue root' plays a central role in Karajá phonology, as evidenced by a pervasive process of vowel harmony--the first of this kind ever described for a South American language. [ATR]-ness is also relevant in describing the second-most common morphophonemic process in the language, palatalization. Reduplication, maybe the most phonologically-oriented of all word-formation processes, will also be treated in this chapter.

Chapter 2 deals with an oft-mentioned—yet poorly understood—peculiarity of Karajá: the existence of systematic differences between male and female speech. As suggested early on by Ehrenreich (1894), female speech is more conservative, male speech

being generally characterized by the deletion of the velar stop /k/. However, although such ‘genderlectal’ differences are often described as being the result of regular phonological processes, they can have morphological consequences as well—a fact that went unnoticed in previous descriptions. The chapter also includes an account of male vs. female speech distinctions in Javaé, where genderlectal variation is much less prominent (Javaé women speaking what seems to be a relaxed version of the male speech in the other dialects). A possible explanation for such a situation, based on a well-documented case of language contact involving the Javaé, is proposed. Hypotheses on the possible diachronic origins of the male vs. female speech distinctions are also discussed. Contrary to previous suggestions, according to which they would be contact-induced (Rodrigues 1999, 2004), the facts seem to favor an explanation in terms of the language’s own internal historical tendencies.

Chapter 3 describes the language’s word-formation strategies, including, in addition to morphology proper (that is, the combination of roots and affixes), the use of particles and clitics. Verb morphology is where most of Karajá’s grammatical complexity lies: the verb inflects for person (with portmanteau morphemes which also indicate mood), direction (which also play evidential functions), and voice (transitive, passive, and antipassive), besides incorporating object pronouns and nouns (sometimes with classificatory purposes). Tense and aspect are conveyed by clitics and particles, respectively. As in many other lowland South American languages, nouns, postpositions, and (to a lesser extent) verbs share person-marking paradigms. Comparative evidence plays an important role in explaining less-than-transparent morphological devices, such as linking prefixes (which, as it turns out, are cognates with similar prefixes in other Macro-Jê languages), synchronically invariable roots with fossilized prefixes, and a non-concatenative mechanism to derive nouns from

verbs. The latter constitute an unusual case of replacive infixation, which, thanks to comparative evidence, can be traced back to the affixation of a nominalizing suffix.

Chapter 4 deals with complex sentences. Subordinate clauses in Karajá are always the result of sentential nominalizations, not only in the cases of relative and complement clauses, but of adverbial clauses as well. It is here claimed that adverbs in Karajá do not exist as an independent part-of-speech (or, at least, constitute a marginal part-of-speech with very few members), adverbial functions being played by postpositional phrases. Consequently, adverbial clauses are postpositional clauses, marked by the same postpositions which mark nouns. Their use as subordination markers mirrors their use with nouns, even in the case of converbs—by far, the most common type of subordinate clauses--, which are marked by a locative postposition. Considering the role they play with adverbial clauses, postpositions are described in this chapter. A case of discourse (rather than grammatical) subordination, the *=dokuri clause*, will also be described in this chapter.

The last chapter, Chapter 5, provides an appraisal of the evidence for the inclusion of Karajá in the Macro-Jê stock, taking into consideration additional evidence uncovered by the present study, informed by recent improvements in the reconstruction of Proto-Jê (Ribeiro 2005). Although a relationship between Karajá and the Jê family had been suggested early by von den Steinen (1884), Karajá was one of the last languages to be included in the Macro-Jê stock (Davis 1968). As its late inclusion in the stock suggests, the similarities between Karajá and other Macro-Jê languages are far from obvious. The evidence presented by Davis, however, is being further corroborated by additional data, as this chapter hopefully demonstrates. Especially compelling for the hypothesis of genetic relationship are those instances in which comparative evidence helps explain synchronic irregularities of Karajá, as

in the case of the replacive infix. The chapter demonstrates the interplay between descriptive and historical linguistics: on the one hand, it shows how a better descriptive knowledge of a language may contribute to uncover additional pieces of comparative evidence; on the other, it shows how a better comparative knowledge may contribute to improve our understanding of synchronic phenomena.

CHAPTER 2

Phonology

1. Introduction

This chapter describes the phonology of Karajá, providing a thorough account of the language's phonological inventory, syllable and stress patterns, and phonological processes. Whenever relevant, differences between the present account and those proposed in previous descriptions of the language¹ will be outlined, and possible diachronic scenarios for the development of certain aspects of Karajá phonology will be discussed. Phonemic inventory is not taken here as a static notion, as diachronic processes and language contact often result in 'irregularities' (semi-productive processes, low-frequency phonemes, and changes in phonotactic patterns) that can only be explained when the language is seen as a dynamic entity. This point is particularly illustrated by the study of loanwords, described throughout this chapter.

Sections 2, 3 and 4 describe the minimal shape of a phonological word, syllabic patterns, and stress placement rules, respectively. Section 5 deals with the phonemic inventory of Karajá, including information on phonotactic restrictions. Consonants are described in Section 4.1 and vowels in Section 5.2, with special attention to sounds whose phonological status may be problematic: palatal consonants, the nasal vowels /ĩ/ and /ã/, and the schwa. A major difference between the account given here and those offered in previous sources lies in the vocalic inventory: here it is shown that the feature [ATR] 'advanced

¹ Pioneering attempts to analyze the sound system of Karajá were made by Hugo Künike (1916, 1919), based on vocabularies collected by Castelnau, Ehrenreich, Coudreau, and Krause. Based, as they were, on poorly transcribed data, such works will not be discussed here.

tongue root’ plays a central role in Karajá phonology. This has profound consequences for the understanding of the language’s phonology, prompting a reanalysis of its consonantal inventory and revealing a pervasive process of vowel harmony--the first of this kind ever described for a South American language. Section 6 describes phonological processes (including morphophonological ones): Section 6.1 describes vowel harmony, Section 6.2 deals with palatalization, and Section 6.3 describes reduplication.

2. The minimal word

The minimal phonological word in Karajá must contain at least two moras. In order to obey this constraint, underlying monomoraic stems must duplicate their vowel (1a), when occurring by themselves in a phrase, or they may cliticize to a following word in the same phrase (1b, 1c). This is illustrated below by the noun stem *bɛ* ‘water’ (1) and by the pronoun *dʊ* ‘3.LOC’ (2):

- (1) a. [kɔla'dʊ bɛ'ɛ ritʃi'wira]
 kɔladʊ bɛ Ø-r-ɪ-tʃiwi=r-a
 kid water 3-CTFG-TRANS-pour/spill=CTFG-PERF
 ‘The kid spilled water.’
- b. [bedʒudʑdʑə]
 bɛ dʒ-udʑdʑə
 water REL-cold
 ‘cold water’
- c. [bɛ'dɪ]
 bɛ=dɪ
 water=INSTR
 ‘with water’

- (2) [dʊʊ ,dado'bidε] ~ [dʊ,dado'bidε]
 dʊ dʌ-d-Ø-obi=d-ε
 3.LOC 2-CTPT-INTR-see=2-IMPERF
 'You saw it.'

Although some (d-class) roots are constituted of a single vowel (*u* 'tooth', *ɔ* 'cluster, bunch; stem'), no root is constituted by a consonant alone. With the exception of the causative suffix *-dəkə̃* (*ɔ̃ dəkə̃*), affixes (some of which are constituted of a consonant alone; e.g. *r-* 'centrifugal direction', *d-* 'centripetal direction', *d-* '3rd person') are all monosyllabic. Verb and noun stems can be monosyllabic or polysyllabic. Most postpositions are monosyllabic, but a few are disyllabic (= *wədā* 'comitative', = *rəbɪ* 'ablative', = *wəθε* 'comparative', = *dəkε* 'dative', = *laku* 'evitative').

3. Syllable

The canonical syllabic pattern in the language is (C)V. However, in Southern and Northern Karajá, surface CCV syllables can appear as the result of the deletion of a schwa occurring between a stop (/b/, /k/, and, less commonly, /d/) and the alveolar flap /r/: *kəɔ* 'toad' [kɔɔ], *kəɔbɪ* 'monkey' [kɔɔ'bɪ], *kəɛ* 'martin-pescador (bird sp.)' [kɛɛ]. As we will see (Chapter 3), the schwa surfaces when a velar stop is deleted in male speech (*ɔ̃ əɔɔ*, *ɔ̃ əɔbɪ*, *ɔ̃ əɛɛ*).

In loanwords containing unacceptable consonant clusters (such as (C)VC.CV), the canonical pattern is maintained through the insertion of a schwa: *bārəso* [marə'so] 'March

(the month of)’ (< Portuguese *março* [ˈmarsu].² In the non-schwa dialects, a vowel identical with the one in the following syllable is inserted: *bãroso* [maroˈso]. CCV syllables are clearly bimoraic, as shown by their behavior in processes such as hypochoristic formation and reduplication. As one would expect, surface monosyllabic words such as *krɔ* ‘toad’ are phonologically self-standing—that is, they can occur by themselves with no need to cliticize or undergo vowel reduplication (3). Notice that, unlike a monosyllabic stem, which has no stress when cliticized to a postposition (cf. (1c) above), CCV stems maintain their accent under such circumstances (3b):

- (3) a. ♂ [ɔlaˈdu ˈkrɔ rɪˈmõre]
 ɔladu krɔ Ø-r-r-bã=r-e
 child toad 3-CTFG-TRANS-catch=CTFG-IMPERF
 ‘The child caught the toad.’
- b. ♂ [ɔlaˈdu ˈkrɔmã roˈbire]
 ɔladu krɔ=bã Ø-r-Ø-obi=r-e
 child toad=DIFF 3-CTFG-INTR-see=CTFG-IMPERF
 ‘The child saw the toad.’

4. Stress

Phonetically, stressed syllables are characterized by having a higher pitch than unstressed ones. The position of the stress is predictable: in general, it falls on the last syllable of the (isolated) word (4). As these examples show, stress follows an iambic rhythmic pattern, being assigned in a right-to-left fashion (4e-g). By default, primary stress

² The Karajá name for their village in Aruanã (Goiás), Buridina [bʊrɛdɪˈna], probably derived from Portuguese *Leopoldina* (the old name of the town), also illustrates this use of the schwa.

is assigned to the rightmost syllable; secondary stress is assigned to every right syllable of a binary iambic foot:

- | | | | | |
|-----|----|-------------------|----------------|--|
| (4) | a. | <i>bʊ</i> | [bʊʊ] | ‘your father’ |
| | b. | <i>hãbu</i> | [hã'bu] | ‘man’ |
| | c. | <i>kəbɔ</i> | [kə'bɔ] | ‘ <i>banzeiro</i> (wind-caused waves)’ |
| | d. | <i>bɔrɔ</i> | [bɔ'rɔ] | ‘stingray’ |
| | e. | <i>hãləkʊ</i> | [hãləkʊ] | ‘hole’ |
| | f. | <i>kɔdʊbɔdã</i> | [kɔdʊbɔ'na] | ‘tortoise’ |
| | g. | <i>wafiwahadɛ</i> | [waʃi,waha'dɛ] | ‘bow (for arrows)’ |

Loanwords whose stress do not fall on the last syllable in the language of origin undergo adaptation to conform to the stress pattern of Karajá: *beradʊ* [bera'dʊ] ‘treacle’ (< Portuguese *melado* [mɛ'lado]), *dakəʃi* [dʌkəʃi] ‘taxi’ (< Portuguese *táxi* [taksi]), *bifikərədɔ* [bifikrɛ'dɔ] ‘bicycle’ (< Portuguese *bicicleta* [bisi'kleta]).

The iambic nature of stress assignment in Karajá is especially clear with morphologically complex words. Notice that, in compounds, the original stress pattern of the individual stems is ‘rearranged’ in order to follow the iambic pattern. One can assume that, although stems can be said to carry stress, the position of such stress is not underlyingly predetermined. Although the stress is, by default, assigned to the last syllable of a stem, its ultimate position will depend on the morphological or syntactic position of the stem. That is illustrated below by a compound (5a), a case of derivational suffixation (6b), and by a vocative phrase involving the vocative interjection, which is stressed (7b). The behavior of the stress is further illustrated in (8) and (9), which show stress (re-)assignment under suffixation (8b) and compounding (9b). As shown in (9c), a possessive phrase such as *brɔrɛ-dĩ hukã-θĩ* ‘cow’s milk’ can be pronounced as two words (in which case both preserve their primary stress) or as a single phonological word.

- (5) a. ♂ *wa-riɔɾɛ* [wa,riɔ'ɾɛ]
1-child
'my child'
- ♂ *wa-riɔɾɛ-riɔɾɛ* [wari,ɔɾeriɔ'ɾɛ]
1-child-child
'my grandchild'
- (6) a. *kɔlɔbɔɾɛ* [kɔ,lɔbɔ'ɾɛ] b. *kɔlɔbɔɾɛ-dʔi* [kɔlɔ,bɔɾɛ'ni]
red.ant red.ant-SIMILAR
'red ant' '(type of) red ant'
- (7) a. *wa-ladzirɔ* [wə,ladzɪ'rɔ] b. *wa-ladzirɔ wɨ* [wəla,dzɪrɔ'wɨ]
1-aunt VOC³
'my maternal aunt' 'oh, my maternal aunt!'
- (8) a. *brɔɾɛ* [brɔ'ɾɛ]
'deer'
- b. *brɔɾɛ-dʔi* [,brɔɾɛ'ni]
deer-similar
'cow'
- (9) a. *hɔkɔ̃* [hɔ'kɔ̃]
'breast'
- b. *hɔkɔ̃-θi* [,hɔkɔ̃'θi]
breast-liquid
'milk'
- c. *brɔɾɛ-dʔi hɔkɔ̃-θi* [brɔɾɛ'ni ,hɔkɔ̃'θi] ~ [brɔ,ɾɛni,hɔkɔ̃'θi]
deer-similar breast-liquid
'cow milk'

Although stress-placement is basically predictable, 'minimal pairs' for stress can appear in a phrasal level, a fact which is due to the contrast between intrinsically stressed vs. unstressed morphemes: while noun and verb stems and derivational suffixes are stressed, a

³ The vocative interjection *wɨ* is used in female speech, corresponding to *kɨ* in male speech.

‘The White man arrived.’

- b. *d'ori* Ø-d-Ø-ɔrɔ=d-e=Ø [doro'de]
White 3-CTPT-INTR-arrive=CTPT-IMPRF=SUBORD
‘The White man who arrived.’
- (13) a. *hābu* Ø-r-Ø-urɔ=r-a [rɔ'rɔra]
man 3-CTFG-INTR-die=CTFG-PERF
‘The man died.’
- b. *hābu* Ø-r-Ø-urɔ=r-a=Ø [rɔrɔ'ra]
man 3-CTFG-INTR-die=CTFG-PERF=SUBORD
‘The man who died.’

4.1 Proper nouns

Although the stress falls regularly on the last syllable of noun stems, proper nouns are frequently exceptions. While some proper nouns are preferentially pronounced with stress in the last syllable (14), some are preferentially pronounced with penultimate stress (15):

(14) **female names, final stress**

- a. M̃yixa [mɔ̃i'ʃa]
b. Areraki [arɛra'ki]
c. Jijuke [dʒidʒu'kɛ]
d. Ixahakaru [iʃahaka'rɔ]

male names, final stress

- d. Maluare [malua're]
e. Watau [wada'u]
f. Maurehi [maure'hi]
g. Ijeseberi [idʒeθebe'ri]

(15) **female names, penultimate stress**

- a. Xikireru [ʃiri'kerɔ]
b. Koaxiru [koa'ʃirɔ]
c. Koabiru [koa'birɔ]

male names, penultimate stress

- c. Ijyrraru [idʒi'raru]

- (17) a. *bãdũ!* [maɖũ] (< *bãdũkari* [maɖũkari] ‘old man’)
 ‘hey, old man!’
- b. *θɛdã!* [θɛna] (< *θɛdãdu* [θɛna'du] ‘old woman’)
 ‘hey, old lady!’

4.3 Enumerations

Another instance in which penultimate stress is favored is when enumerating items, such as in (18) below:

- (18) [iɾə 'maki ã'dəhə]
 iɾə, bãki, ãdəhə
 manioc corn yam
 ‘[He planted many things, such as] manioc, corn, yam...’

5. Phonological inventory

This section describes the segmental inventory of Karajá; Section 4.1 deals with consonantal phonemes, whereas Section 4.2 deals with vowels. In Karajá, consonants and vowels form two clearly distinct classes; only vowels can be the nucleus of a syllable, and a syllable can be constituted of a vowel alone (e.g. *kai* ‘you’ [ka'i]). Consonants, on the other hand, only occur as the onset of a syllable. There are no ambiguous segments; the glide /w/ behaves clearly as consonant, not a semivowel. There are no long vowels or diphthongs; sequences of vowels are always tautosyllabic.

5.1 Consonants

According to my analysis, there are nine phonemic consonants in Karajá (Table 1). A major departure from previous descriptions of Karajá phonology (Fortune & Fortune, Cavalcante) is the fact that I do not treat the palatal consonants [j], [tʃ], and [dʒ] as independent phonemes (as indicated by the parentheses in Table 1). They generally are, as it will be shown, allophones of non-palatal consonants in contiguity to [high, +ATR] vowels. A remarkable fact regarding the consonantal inventory is the rarity of voiceless stops: the language lacks both a bilabial voiceless stop /p/ and an alveolar voiceless stop /t/. As further discussed in Chapters 3 and 6, the lack of cross-linguistically common phonemes such as /p/ and /t/ is probably the result of a historical drift towards the lenition of voiceless stops. The fact that the only voiceless stop in the language, /k/, is ‘unstable’--being, as it is, the consonant involved in the process of deletion which characterizes male speech (Chapter 3)--is in itself instructive of such tendency.

Table 2.1. Consonantal inventory of Karajá

		(tʃ)	k
b	d	(dʒ)	
	ɗ		
	θ	(j)	h
	l		
w	r		

Most of the phonemes listed in Table 1 above are represented with their standard IPA values and do not require additional explanation as to their pronunciation, except for the voiced stops /d/ and /b/ and the fricative [ʃ]. As in most Macro-Jê languages (such as those of the Jê, Maxakalí, and Jabutí families; Ribeiro & van der Voort 2010), in Karajá the voiced stops /b/ and /d/ do no contrast phonologically with their nasal counterparts. They are

pronounced as fully oral consonants before oral vowels and fully nasal consonants before nasal vowels, as illustrated below by the prefixes *d-* ‘centripetal direction’ and *b-* ‘2nd person (irrealis)’. Loanwords that do not fit such pattern tend to be adapted to Karajá phonotactics, as in *beradu* [bera'du] ‘treacle’ (< Portuguese *melado* [me'ladu]), where Portuguese /m/ is adapted to /b/, since there is no nasal counterpart to the middle front vowel /ɛ/ in Karajá.

- (19) a. *b-obi=kre* [b'ɔbikre]
 2-see=FUT
 ‘you will see’
- b. *b-õhõ=kre* [mõ'hõkre]
 2-bathe=FUT
 ‘you will bathe’
- (20) a. *d-obi=d-e* [d'ɔbide]
 CTPT-see=CTPT-IMPERF
 ‘he saw (it) (hither)’
- b. *d-õhõ=de* [nõ'hõde]
 CTPT-bathe=CTPT-IMPERF
 ‘he bathed (hither)’

The pronunciation of the sound transcribed as [ʃ] throughout this dissertation varies between [s] and [ʃ]. Portuguese /s/, borrowed as [s] in Karajá, is perceived as being a variation of [ʃ], being therefore written as <x> in Karajá orthography. A commonly-used Portuguese loanword is <*xoba*> [so'ba] ‘to soak’ (< Portuguese *ensopar*), which preserves two violations to Karajá phonotactics (the occurrence of [s] in contiguity to a non-high vowel, and of [b] before /a/). The loan is, however, fully adapted grammatically (21a), and, as expected, it triggers vowel harmony (21b):

- (21) a. wa-dəki Ø-r-a-soba=r-a

1-clothes 3-CTFG-INTR-soak=CTFG-PERF
 ‘My clothes got soaked.’

- b. kədǝde soba [kənǝdeso'ba]
 flour soak
 ‘soaked flour’

Aside the phonemic processes described below (palatalization and vowel harmony) and the distributional restrictions concerning the nasal vowels /ã/ (Section 4.2.1) and /ɨ/ (4.2.2), there does not seem to be any restrictions as to the co-occurrence of consonants and vowels. Unattested combinations are /w/ before a homorganic vowel /u/ or /ʊ/, or before nasal vowels, and the lateral /l/ before /ɔ/, but those gaps could very well be an artifact of the data (which do not, of course, exhaust the whole lexicon of the language). Hiatuses are common (22)--and not only as the result of k-dropping in male speech (Chapter 3)--, but a bilabial glide tends to be inserted between a back vowel and an onsetless vowel (23):

- (22) a. ♀ hãlɔkɔɛ ♂ hãlɔɛ ‘jaguar’
 b. riu ‘hunting’
 c. kai ‘you’

- (23) a. hɨlɔi ‘to vomit’ [hɨlɔ'i] ~ [hɨlɔ'wi]
 b. θɔɛ ‘many, much’ [θɔ'ɛ] ~ [θɔ'wɛ]
 c. -ʊɔ ‘to fly’ [ʊ'ɔ] ~ [ʊ'wɔ]
 d. -uaθa ‘poisonous arrow’ [ua'θa] ~ [uwa'θa]

5.2 Vowels

Contrasting with the previous phonological descriptions of the language (Fortune & Fortune 1963, Cavalcante 1992), my own analysis points to a larger system of vowels, in which the feature [ATR] ‘advanced tongue root’ plays a major role (Ribeiro 2000).

According to my description, there are seventeen vowels in Karajá: thirteen oral and four nasal (Table 2). Compared to Fortune & Fortune (1963), the main difference is that I recognize a phonemic opposition between the high [+ATR] vowels /i/, /ĩ/, and /u/ and their [-ATR] counterparts /ɪ/, /i/, and /ʊ/, a distinction not mentioned in the previous works.⁶ Besides the existence of minimal pairs (for example, *-lahi* ‘to curse’ vs. *lahi* ‘grandmother’, *kĩ* ‘inside’ vs. *kɪ* ‘tree bark fiber’, and *-uka* ‘to split’ vs. *-ʊka* ‘to cook’), the phonemic character of the opposition is also shown by the fact that [-ATR] and [+ATR] high vowels have exactly inverse behaviors in the processes of vowel harmony (Section 5.1) and palatalization (Section 5.2).

Table 2.2. Vocalic inventory of Karajá⁷

oral			nasal	
i	ĩ	u	ĩ	
ɪ	i	ʊ		
e	ə	o	ẽ	õ
ɛ	(ə)	ɔ		
	a		ã	

Although both Fortune & Fortune (1963) and Cavalcante (1992) arrive at similar consonantal inventories,⁸ the complete vocalic inventory of Karajá remained, to a certain

⁶ Actually, Fortune & Fortune (1963) described the phonemic contrast between /u/ and /ʊ/ and even represented both separately in an early version of the Karajá orthography (with the letters <u> and <ù>, respectively; cf. Fortune & Fortune 1972). However, in their latest version of the Karajá orthography (as in Alford, Fortune & Fortune 1983), such distinction is no longer represented. As we will see, Fortune & Fortune (1963) briefly mentioned vowel harmony as a matter of height assimilation, without noticing the opposite behaviors of /u/ (which triggers vowel harmony) and /ʊ/ (which undergoes it), despite both being high vowels. The use of a grave accent with <ù> parallels its use with the back mid vowels, where <ò> represents the open-mid vowel /ɔ/ and <o> represents the close-mid vowel /o/. The decision to graphically mark the [-ATR] vowels, instead of the [+ATR] ones, is in itself problematic, since [-ATR] vowels (as the unmarked members of the opposition) are much more common. That is, the system of diacritics devised to differentiate [-ATR] and [+ATR] vowels is functionally backwards.

⁷ I am using the symbol ə to represent the mid-close central [+ATR] vowel which occurs in such words as *irə* ‘sweet manioc.’ This phoneme is transcribed by Fortune & Fortune as *ĩ* and by Cavalcante as *ə*.

extent, an open question, especially regarding nasal vowels. Fortune & Fortune (1963) describe four (as in Table 2 above) and Cavalcante (1992) describes three nasal vowels (treating the nasal low vowel as an allophone of its oral counterpart), while Rodrigues (1999), in a recent survey of the Macro-Jê languages, states that Karajá “has only two nasal vowels, /ã/ and /õ/.” The apparently questionable phonemic status of both /ĩ/ and /ã/ is due to their low frequency. But, as it will be further discussed in Sections 4.2.1 and 4.2.2, their distribution is not predictable and there is fairly clear evidence of contrast between them and their oral counterparts. The schwa, whose phonemic status is also problematic, will be thoroughly discussed in Section 4.2.3.

5.2.1 The nasal low vowel /ã/

As I mentioned above, the nasal low vowel is considered as an independent phoneme by Fortune & Fortune (1963), but not by Cavalcante (1992), who treats it as an allophone of its oral counterpart, /a/. Rodrigues considers it “an automatic realization of the phoneme /a/ when it either stands at the beginning of a word or is preceded by /h/ or by a voiced stop [...]” Although such a description accounts for the majority of examples containing this vowel (24, 26), there are a number of counterexamples to it. As shown by the examples in (25, 27) below, a number of words fitting the structural description provided by Rodrigues present an oral low vowel instead. In addition, there are examples in which the nasal low vowel occurs in onsetless syllables word-medially (28), a position in which oral low vowels are also attested (29):

⁸ The only difference between both analyses of the consonantal inventory is in the fact that Cavalcante (1992) considers [m] and [n] as independent phonemes, whereas Fortune & Fortune (1969) treat them as allophones of their oral counterparts, /b/ and /d/, respectively. My analysis agrees with the Fortunes’ in this particular aspect.

- | | | | | | | | |
|------|----|----------------|----------------------|------|----|---------------|---------------------------------|
| (24) | a. | <i>hābu</i> | ‘man’ | (25) | a. | <i>haka</i> | ‘buritirana (tree sp.)’ |
| | b. | <i>kɔhā</i> | ‘armadillo’ | | b. | <i>waha</i> | ‘my father’ |
| | c. | <i>hāwa</i> | ‘place’ | | c. | <i>āθi</i> | ‘grass’ |
| (26) | a. | <i>āθi</i> | ‘grass’ | (27) | a. | <i>aθara</i> | ‘calango (lizard sp.)’ |
| | b. | <i>ā-ra</i> | ‘your nephew’ | | b. | <i>adī</i> | ‘to bring (s.t.) down; to fell’ |
| | c. | <i>āθu</i> | ‘embaúba (tree sp.)’ | | | | |
| (28) | | <i>dohokuā</i> | ‘newborn’ | (29) | | <i>kuadzī</i> | ‘rainbow’ |

Further complicating matters, the oral allophone of the voiced stops /b/ and /d/ do not occur before /a/ in the native lexicon (30). That is further illustrated, again, by the prefixes *b-* 2nd person (irrealis) and *d-* ‘centripetal direction’. Notice that, even though the vowel of the intransitive prefix *a-* is apparently oral, it seems to trigger the nasalization of the prefixes (31). That suggests that, historically, /a/ was intrinsically nasal; that nasality would later be lost in most environments, except after /h/, in onsetless positions, and, as a historical relic, with previously nasalized consonants. As noticed by Fortune & Fortune (1963), the nasality of /ā/ after [m] and [n] is much less pronounced than with other nasal vowels. Old loans (32) tend to adapt to the native phonotactic restrictions.

- | | | | | |
|------|----|----------------|-------------------------|------------|
| (30) | a. | <i>dʼakidā</i> | ‘star’ | [dʼakiˈna] |
| | b. | <i>ladā</i> | ‘uncle’ | [laˈna] |
| | c. | <i>wobā</i> | ‘axe’ | [woˈma] |
| | d. | <i>webā</i> | ‘matamatá (turtle sp.)’ | [weˈma] |

- | | | | |
|------|----|----------------------|--------------|
| (31) | a. | <i>b-a-ritʃa=kε</i> | [mariˈtʃakε] |
| | | 2-INTR-walk=POT | |
| | | ‘Walk!’ | |
| | b. | <i>d-a-ritʃa=d-e</i> | [nariˈtʃade] |
| | | 3-CTPT-INTR-walk=POT | |
| | | ‘he walked (hither)’ | |

- (32) a. *bãri'dɔ* [marɪ'dɔ] 'coat, suit' (< Portuguese *paletó*)
 b. *bãbera* [mabɛ'ra] 'paper' (< Língua Geral *papéra*)

Corroborating the hypothesis that there was a 'wholesale' nasalization of /a/ diachronically (at least in certain positions) is the fact that the male speech form for the first-person prefix in the irrealis mood, ♀ *ka-*, is nasalized: ♂ *ã-*.⁹ The same does not happen with recent loans such as *kara* 'yam' (< Portuguese *cará*, ultimately from Tupinambá *kará*) and *kabe* 'coffee' (< Portuguese *café*), whose low vowels do not get nasalized upon the dropping of /k/ in male speech. The introduction of such loanwords creates minimal pairs for the contrast between /a/ and /ã/ (33):

- (33) a. *ã-ra* 'your nephew' vs. ♂ *ara* 'cará' (< *kara*)
 b. *ã-be* 'your water' vs. ♂ *abe* 'coffee' (< *kabe*)

Even if minimal pairs such as the ones above are considered 'spurious' as evidence for the contrast between /a/ and /ã/, given the fact that they are based on recent loanwords, they do illustrate the evolving nature of the language's phonology, under a situation of ever-increasing bilingualism. Recent loanwords, such as *babai* [baba'i] 'daddy (vocative)' (<

⁹ Another example in which male speech /ã/ corresponds to female speech /ka/ is the 1st-person pronoun ♀ *džikarɛ* 'I', ♂ *džiãrɛ* (notice that this is the word used in all dialects except Southern Karajá, where the form of the pronoun is ♀ *dɪkarɛ* (♂ *dɪarɛ*); the difference in the [ATR] quality of the first vowel is not a systematic correspondence among the dialects; other sporadic cases of similar mismatches are discussed below (see Section 5.2, 'Palatalization')). There seems to be a correlation between [+ATR]-ness and nasalization, as suggested by the contrast between *dɪarɛ* and *džiãrɛ* (see also *ɪvahɪ* 'medicine' vs. *dohokuã* 'newborn'). A similar example is the loan [siã'ru] 'cigarette' (< Portuguese *cigarro*), where the nasality of the vowel cannot be attributed to any regular phonological rule. This is a matter for future investigations.

Portuguese *papai*), do not show the same degree of adaptation as the one seen with the examples in (32).¹⁰ At any rate, the phonemic contrast between /a/ and /ã/ does not need to rely on minimal pairs such as the ones in (33), since their distribution is unpredictable even within the native lexicon.

5.2.2 The nasal high front vowel /i/

Another vowel ignored by Rodrigues (1999), but whose phonemic status was recognized by Fortune and Fortune (1963) and Cavalcante (1992), is the nasal front vowel /ĩ/, which occurs in very few examples, such as *hĩ* ‘(a woman’s) older brother’ (which forms a ‘minimal trio’ with *hi* ‘cry’ and *hi* ‘to drive away’), *ãhĩ* ‘mosquito’¹¹, *kudĩ* [ku’ni] ‘spirit, ghost’, and the extremely productive derivational suffix *-dĩ* [ni] ‘similar to’. The contrast between /ĩ/ and its oral counterpart is demonstrated by the minimal pair in (34) below.

Another minimal pair, *raĩ=ra* ‘I stood up’ (35b) vs. *rai=ra* ‘I laid down’ (36b), can be explained morphologically as the result of the deletion of the nasal vowel in the root *-ãi* ‘to stand up’ and the transference of the nasality to the stressed vowel:¹²

¹⁰ With some loans, different forms may coexist, varying as to the degree of adaptation. That is the case of the Portuguese loan for ‘bread’ (Portuguese *pão*). The form [ma’õ] is said to be used only by those with little or no knowledge of Portuguese; the less adapted forms [ba’õ] or even [pa’õ] are more commonly used.

¹¹ The word for ‘mosquito’, *ãhĩ*, is probably of onomatopoeic origin. Another onomatopoeic word containing /ĩ/ is [sĩ’ka], the name of a bad-omen bird.

¹² Another example of nasal stability under vowel deletion, also involving the nasal mid-central vowel /õ/, occurs with the root *hõ* ‘to be’, when it is prefixed with the 1st person prefix *wa-* (cf. (a) below; for additional details on this root, which becomes grammaticalized as a habitual marker, see Chapter 5). Thus, there is yet another (superficial) minimal pair for the contrast between /a/ and /ã/: *waha=r-e* ‘(he) is my father’ [wa’hare] vs. *wa-hã=r-e* ‘I am’ [wa’hãre].

(a)	wa-hã=r-e 1-be=CTFG-IMPERF ‘I am.’	(b)	ã-hã=ɸ-e 2-be=2-IMPERF ‘You are.’	(c)	i-hõ=r-e 3-be=CTFG-IMPERF ‘He is.’
-----	--	-----	---	-----	--

- (34) *ãhĩ* ‘mosquito’ vs. *ãhi* ‘your cry, your crying’
ã-hĩ ‘your older brother’
- (35) a. \emptyset -r- \emptyset - $\tilde{a}i=r-a$ [rõ'ira]
 3-CTFG-INTR-stand.up=CTFG-PERF
 ‘He stood up.’
- b. *r-a- \emptyset - \emptyset - $\tilde{a}i=ra$* [ra'ira]
 CTFG-1-CTFG-INTR-stand.up=CTFG-PERF
 ‘I stood up.’
- (36) a. \emptyset -r- \emptyset -oi=r-a [ro'ira]
 3-CTFG-INTR-stand.up=CTFG-PERF
 ‘He laid down.’
- b. *r-a- \emptyset - \emptyset -oi=ra* [ra'ira]
 CTFG-1-CTFG-INTR-stand.up=CTFG-PERF
 ‘I laid down.’

Note that the environments in which *hĩ*/ occurs are the same as the ones in which *ãhĩ*/ occurs: that is, onsetless syllables, after the glottal fricative /h/ and after voiced stops. Another similarity is the fact that the nasality of *hĩ*/ is much less pronounced after [n].¹³ As with *ãhĩ*/, although the environments in which *hĩ*/ occurs are quite restricted, its distribution does not seem to be totally predictable, a fact that—in addition to the existence of unquestionable minimal pairs (34)--grants it phonemic status.

5.2.3 The schwa

The schwa in Karajá corresponds to the traditional definitions of the term, in both its phonetic and phonological acceptations. Phonetically, it is an unstressed mid-central vowel, as indicated by its transcription with the IPA symbol [ə]. Phonologically, it is the default vowel

¹³ I have not yet found any instance of [m] before *hĩ*/, except for the interjection ♀ *bĩ* [mĩ] ‘surprise’.

par excellence, occurring to break up disallowed consonantal sequences (37). As the French schwa in Anderson’s (1982) analysis, the schwa in Karajá can be described as a featureless vowel; its lack of ‘phonemic personality’ is particularly demonstrated by the fact that it tends to completely assimilate all the features of a contiguous vowel (*total harmony*), a process that seems to be obligatory morpheme-internally (38). The schwa generally occurs after consonants (at least in the female speech, which is more conservative; see Chapter 3); so far, I have found only two examples of it occurring in onsetless syllables, əθɛ ‘fart’ (39) and əfi ‘soft’.

- (37) a. *bɔsəka* *bɔsaa* ‘matches’ (< Portuguese *fósforo*¹⁴)
 b. *dakəsi* *daasi* ‘taxi’ (< Portuguese *táxi*)

- (38) a. ♂ *hãlɔv* ‘hole’ (< *hãlɔkɔ*)
 b. ♂ *dʌa* ‘to tie’ (< *dʌka*)
 c. ♂ *-dãã* ‘causative’¹⁵ (< *dəkã*)

- (39) a. Ø-*r-a-əθɛ-dã=ra* [raə'θɛnãra] ~ [raa'θɛnãra] ~ [ra:'θɛnãra]¹⁶
 3-CTFG-INTR-fart-VERB=CTFG-PERF
 ‘He farted.’

As mentioned in Chapter 1, the main phonological difference among the four dialects consists in the occurrence in Southern and Northern Karajá of a schwa [ə] in unstressed positions, corresponding to environments in which Xambioá and Javaé present a vowel identical to the one occurring in the following syllable:

¹⁴ Probably via regional Portuguese *fosco* [fɔskɔ].

¹⁵ Notice that derived nasal vowels (such as the one resulting from total harmony in this example) do not trigger the nasalization of the previous consonant.

¹⁶ Notice that, in unstressed positions, an [əV] sequence may be optionally pronounced as [V:].

(40)	Karajá	Javaé, Xambioá	
	<i>kədɔ</i>	<i>kɔdɔ</i>	‘termite’
	<i>bədɪ</i>	<i>bɪdɪ</i>	‘honey’
	<i>-dəkǽ</i>	<i>-dǽkǽ</i>	‘causative suffix’
	<i>rəkʊ</i>	<i>rʊkʊ</i>	‘gourd’

Another difference is the occurrence, in the Northern and Southern Karajá dialects, of surface CCV syllables, resulting from a process of syncope of a schwa occurring between a stop and the alveolar approximant /r/. CCV syllables can, in such cases, be analyzed as being the result of schwa syncope in CərV environments (where C is either /k/, /b/, or /d/). That becomes obvious in male speech: with the suppression of /k/, the schwa surfaces (41a, 41b). As expected, Xambioá and Javaé present a vowel identical to the one in the following syllable and no syncope takes place:

(41)	Karajá	Javaé, Xambioá	
a.	<i>krɔ, ɔ əɔ</i>	<i>kɔɔɔ, ɔ əɔ</i>	‘toad’
b.	<i>krɛ, ɔ əɛ</i>	<i>kɛɛɛ, ɔ ɛɛ</i>	‘martim pescador (bird sp.)’
c.	<i>brɔ</i>	<i>bɔɔ</i>	‘back’

There are reasons to believe that Southern and Northern Karajá are, with respect to the existence of the schwa, the most conservative dialects.¹⁷ The alternative would be to consider the scenario occurring in Javaé and Xambioá as the most conservative, postulating for Southern and Northern Karajá a rule of *lenition* of an unstressed vowel when followed by a syllable containing an identical vowel. However, this hypothesis is ruled out by the

¹⁷ That is probably a different conclusion from the one reached by the Fortunes. Although they do not discuss dialectal differences such as the ones illustrated above, their decision concerning Karajá orthography suggests that they consider Javaé and Xambioá to be more conservative. The schwa is not represented in the common orthography adopted for the four dialects. Thus, morphemes such as *dəka* ‘to tie’ and *daka* ‘to take off’ (a minimal pair in Karajá, but a homophonous pair in Xambioá) are not distinguished orthographically in any of the dialects.

existence, in Southern and Northern Karajá, of minimal pairs such as ♀ *daka* ‘to take off’ versus *dəka* ‘to tie’ (corresponding to a homophonous pair in Xambioá), as well as *bəɾɔ* [bɔˈrɔ] ‘stingray’ versus *bəɾɔ* [brɔ] ‘back’ (corresponding to a homophonous pair in both Javaé and Xambioá).

However, the fact that the schwa occurs only in unstressed positions makes its phonemic status problematic. Furthermore, there is plenty of independent evidence demonstrating that Karajá proper presents, indeed, a tendency towards dissimilation. I will argue that such tendency is actually inherited from Proto-Karajá (and, therefore, that the Xambioá and Javaé dialects have innovated). The following sections discuss arguments against and in favor of considering the schwa as a conservative feature. Based on morphophonemic data, as well as on comparative evidence and the behavior of loanwords, Section 5.2.3.1 describes a strong tendency towards dissimilation in Karajá proper, a phenomenon which is at the root of the origin of the schwa. Section 4.2.3.2 shows that, on the other hand, deviations in the general correspondence rules set above (e.g. *bəɾɪ :: bɪɾɪ* ‘honey’), as well as the existence of multiple reflexes, in Javaé and Xambioá, of a single morpheme in Karajá, strongly suggest that there was indeed a schwa in Proto-Karajá. Considerations of predictability are especially important in considering the schwa as a conservative feature: while it is possible to predict the shape of a morpheme in Xambioá and Javaé from their shape in Karajá, the opposite is not true.

5.2.3.1 The schwa as innovation

There are several arguments to consider the schwa as an innovation, a result of dissimilation whenever there is a sequence of identical vowels—a process reminiscent of those described, for tones and other features, in terms of the Obligatory Contour Principle (OCP). In the few cases for which comparative evidence is available, it seems that the schwa occurs in a position where originally there was a full vowel:¹⁸

(42)	Jê¹⁹	Karajá	Javaé, Xambioá	
a.	* <i>prãm</i>	[rə'ma]	[ra'ma]	‘hunger’
b.	* <i>j-am</i>	[l-ə'ma]	[l-a'ma]	‘to stand up’
c.	<i>mɛŋ</i> (Apinajé)	[bə'dɪ]	[b'r'dɪ]	‘honey’

A common process in Macro-Jê languages, such as Ofayé and those of the Jê family, is the insertion of *echo vowels* (see, for instance, Oliveira 2005, for Apinajé) at the end of consonant-final stems. That may have been the origin of the sequences of syllables containing identical vowels in Karajá, as in the examples above. Internal evidence from Karajá morphophonemics also seems to reveal a similar situation. Karajá has two nominalizing suffixes, *-θV* and *-dV*, whose shapes are determined by the final vowel of verb root (see Chapter 4). In Karajá proper, the original final vowel of the verb is dissimilated upon suffixing:

(43)	Verb	Noun		
		Karajá	Javaé & Xambioá	
a.	<i>hʊ</i>	<i>hə-dʊ</i>	<i>hʊ-dʊ</i>	‘to finish’
b.	<i>wɛ</i>	<i>wə-dɛ</i>	<i>wɛ-dɛ</i>	‘to penetrate’
b.	<i>hɪ</i>	<i>rə-dɪ</i>	<i>rɪ-dɪ</i>	‘to drive away’ ²⁰

¹⁸ Such correspondences will play an important role in discussing the likely origins of *replacement infixation* of the nominalizer *-r-* in Karajá (Chapter 6).

¹⁹ Forms preceded by an asterisk are Proto-Jê forms reconstructed by me (Ribeiro 2005).

²⁰ The derivation *hɪ* > *rɪ-dɪ* involves both suffixation and *consonantal replacement*, a process that will be further described in Chapter 4 (and whose likely diachronic origins will be discussed in Chapter 6).

c. -*obi* -*obə-fi* *obi-fi* ‘to see’

Cases in which the schwa occur after a nasal consonant also presuppose dissimilation, since, as we have seen, nasal consonants only occur before nasal vowels. Thus, the presence of a nasal consonant at the beginning of the word for ‘rock, stone’, pronounced as [mə'na] in Karajá proper and [ma'na] in Javaé and Xambioá, can only be explained if one assumes that the consonant is followed by an underlyingly nasal vowel, /bādã/. A rule of dissimilation in Karajá is, thus, more plausible. This point is further illustrated by the prefix *bV-* ‘2nd person (irrealis)’ in the examples below. As we have seen, /b/ is pronounced as an oral consonant before an oral vowel and as a nasal consonant before nasal vowels. The examples below, in which /b/ surfaces as a nasal consonant before the schwa also suggests an analysis in terms of dissimilation:

- (41) a. *bV-d-Ø-ɔrɔ=kɛ* [bədɔ'rɔkɛ]
 2-CTPT-INTR-arrive=POT
 ‘Come (hither)!’
- b. *bV-d-Ø-ðhð=kɛ* [mənð'hðkɛ]
 2-CTPT-INTR-bathe=POT
 ‘Bathe (hither)!’

Another clear case of dissimilation in Karajá proper involves the first-person prefix *wa-* when attached to a number of kinship terms (all of which starting with /la/):

- (42) a. *wa-lahɪ* [wəla'hɪ] ‘my grandmother’
 b. *wa-labɪkɛ* [wəlabɪ'kɛ] ‘my grandfather’
 c. *wa-ladzɪrə* [wəladzɪ'rə] ‘my aunt’

Finally, there are a number of examples demonstrating that CCV syllables in Karajá proper may also be the result of dissimilation. This is illustrated below by an example involving the prefix ♀ *ka-* ‘1st. person irrealis’ followed by *r-* ‘centrifugal direction’ and the marker of intransitive verbs *a-* (42). Additional examples include the nominal form of the verb *daka*, formed by consonantal replacement and suffixation (*dara-θa* [dra'θa]), and the loan *brɔre* ‘hoe’ (< Língua Geral *pururé*).

- (43) ♀ *ka-r-a-ritfa=kre* [krari'tʃakre]
 1-CTFG-INTR-walk=FUT
 ‘I will walk (thither).’

5.2.3.2 The schwa as a conservative feature

As we have seen, the schwa in Karajá proper generally corresponds to a synharmonic vowel in the non-schwa dialects. While this correspondence rule accounts for the majority of examples, there is a subset of examples in Javaé that require a different explanation. When the following vowel is /i/, the Karajá schwa will correspond to /e/ in Javaé (44). Again, Javaé and Xambioá will have homophonous pairs where Karajá presents minimal pairs. For example, corresponding to the Karajá stem *əfi* ‘soft’, Xambioá has *ifi* (homophonous with the reflexive morpheme *ifi*) and Javaé has *efi* (homophonous with *efi* ‘younger brother’).

- | | | | | |
|------|---------------|--------------|----------------|---------------------|
| (44) | Karajá | Javaé | Xambioá | |
| a. | <i>dəbie</i> | <i>dəbie</i> | <i>tʃibie</i> | ‘to raise, to feed’ |
| b. | <i>hətʃi</i> | <i>hetʃi</i> | <i>hitʃi</i> | ‘bottom; buttocks’ |
| c. | <i>əfi</i> | <i>efi</i> | <i>ifi</i> | ‘soft’ |

d. *dədɪ* *dədʒɪ*²¹ *dɪdɪ* ‘to put’

The root *hətʃi* ‘bottom’ (44b) is particularly instructive. The glottal fricative /h/ behaves as a ‘transparent’ consonant, allowing total harmony to take place just as with onsetless syllables. Therefore, when *hətʃi* follows another root in a compound, the schwa assimilates the features of the last vowel in the preceding root (45). Interestingly enough, compounds such as *berahətʃi* ‘river bottom’ are common to all dialects, demonstrating that a ‘harmonizing’ form (that is, a form containing a schwa) must have already been present in the proto-language.²²

(45) *bera-hətʃi* [berahaʔʃi]
 water-bottom
 ‘water bottom’

²¹ The variation in the [ATR] value of the high vowel in this stem is not a systematic difference between Javaé and the other dialects. Other examples of sporadic differences of this type are *ɖɪkarə* ‘I’ (in Southern Karajá) versus *dʒɪkarə*, *kɪla* ‘small’ versus *kɪdʒa*.

²² That is probably the origin of the interesting pattern of morphophonemic alternations found with the ‘augmentative’ morpheme, whose first vowel is always a copy of the last vowel in the preceding root (see below). The difference is that the ‘augmentative’ morpheme is a bound root, occurring always as the second element in a compound. Therefore, an underlying, schwa-preserving form is not found in surface representation. Stems such as *hətʃi*, on the other hand, although obligatorily possessed, are morphologically independent, and may be found without any preceding morphological element.

♀	a.	<i>beraku-hukə</i> [bera,kuhu'kə]	b.	<i>dakidā-həkə</i> [daki,nahə'kə]
		river-big ‘the Araguaia River’		star-big ‘big star (Venus)’
	c.	<i>ibəru-hukə</i> [i,bɾuhʊ'kə]	d.	<i>i-θɔ-həkə</i> [iθɔhɔ'kə]
		wail-big ‘long ritual wailing’		3-red-big ‘very red’

A similar example involves the root *hədɛ* ‘hit’, which generally falls into the overall pattern of correspondence. However, while its basic form in Xambioá is, expectedly, *hədɛ* (46a), it also occurs as *hadɛ* when preceded by the incorporated object pronouns of first and second persons, *wa-* (46) and *a-*. That is, corresponding to a single form in Karajá proper, Xambioá has two forms: *hədɛ* and *hadɛ*. Again, the postulation of a schwa in the proto-language provides a more straightforward explanation for the existence of multiple reflexes in Xambioá.

- (46) a. \emptyset -*r*-*r*- \emptyset -*hədɛ*-*dǝ=r-a* [rihɛ'dɛnǝra]
 3-CTFG-TRANS-3-hit-VERB=CTFG-PERF
 ‘he hit him’
- b. \emptyset -*r*-*r*-*wa*-*hadɛ*-*dǝ=r-a* [ruaha'dɛdǝra]
 3-CTFG-TRANS-1-hit-VERB=CTFG-PERF
 ‘he hit me’

Finally, examples involving the nasalization of /b/ and /d/ also seem to corroborate the need for postulating the existence of a schwa in the proto-language. As we have seen, a form such as *-dǝkǝ* [dǝkǝ] ‘causative’ (Xambioá and Javaé *-dǝkǝ* [nǝkǝ]) may occur without the velar stop in male speech. As a consequence of k-deletion, the schwa assimilates all the features of the following vowel (*-dǝǝ* > *-dǝǝ*), but the consonant remains oral: [dǝǝ].

Examples such as this demonstrate the need to differentiate cases such as *bǎnǎ* [mǎna] ‘stone’, where the schwa results from the dissimilation of a nasal vowel, from *-dǝkǝ* ~ *dǝǝ* [dǝkǝ ~ dǝǝ] ‘causative’, where the schwa is originally an oral vowel.

²³ It is interesting to notice the contrast between forms such as (46b) and those containing the intransitive marker *a-*; although the phonological environment would be presumably the same, the intransitive form is *hədɛ*. This suggests that object markers have a more ‘intimate’ relationship with the verb stem than inflectional prefixes such as the intransitive marker (even if, on the surface, they seem to be occurring in the ‘same’ position). A similar state of affairs—the fact that palatalization is restricted to certain prefixes (equivalent to Lexical Phonology’s *level 1* affixes)—is discussed below.

5.2.3.3 Schwa vs. /ə/

Considering that its occurrence is limited to unstressed positions, could the schwa be treated as an allophone of a full vowel phoneme? The vowel phoneme whose articulatory characteristics are closer to the schwa is the mid-central [+ATR] vowel /ə/. Considering that this vowel occurs mostly in stressed position, it seems to be in complementary distribution with the schwa.²⁴ There are, however, a few examples in which /ə/ occurs in unstressed position; these examples, however, would not pose a problem to an analysis in terms of complementary distribution, since the schwa does not seem to occur in such positions (47c-e). At any rate, treating the schwa as an allophone of /ə/ or any other phoneme would be a rather arbitrary solution, given its origin as the result of dissimilation of any vowel in the inventory.

- (47) a. *riə* 'hammock'
b. *irə* 'sweet manioc'
c. *hələ* 'peccari'
d. *ãdəhə* 'yam (sp.)'
e. *həri* 'shaman'

6. Phonemic processes

²⁴ Initially, when the schwa was represented in Karajá orthography, it was written with the same symbol used to represent /ə/, <à>. In orthographic terms, such decision makes sense, since /ə/ tends to occur mostly in stressed position.

6.1 Vowel harmony

Vowel harmony is certainly the most remarkable feature of Karajá phonology. In both previous descriptions of Karajá phonology, Fortune & Fortune (1963) and Cavalcante (1992), vowel harmony is briefly mentioned as a process by which a high or close-mid vowel “closes” an open-mid vowel in a preceding syllable.²⁵ The Fortunes’ examples are reproduced below (48), with my own transcription and morphological segmentation:

- | | | | | |
|---------|---|--|------------|----------------|
| (48) a. | ♂ | /d-εbɔ-ube/
REL-hand-palm | [debou'be] | ‘palm of hand’ |
| b. | | /r-ɪ-dɔ=r-e/
CTFG-TRANS-eat=CTFG-IMPERF
‘S/he ate (it).’ | [ri'dore] | |
| c. | | /budɔε-dĩ/
deer-similar.to | [budoe'ni] | ‘sheep’ |

Although both accounts are limited to a handful of examples, without any attempt at further generalization, it is clear that they consider vowel harmony to be a case of height assimilation. Thus, discussing the example [debou'be] ‘palm of hand’, Fortune & Fortune suggest that “the high /u/ of *ube* seems to have influenced all the preceding vowels to the higher vowel position.” However, this formulation would not account for a number of cases

²⁵ Interestingly, Fortune & Fortune label as a case of vowel harmony only the example reproduced as (48a) above, naming the process illustrated by (48b) and (48c) as ‘vowel assimilation across consonants.’ However, the examples illustrate obviously one and the same process. Notice that example (48a) is [debokube] in female speech—and therefore a case of “assimilation across consonants” as well. Cavalcante’s description is limited to a rule of ‘raising of /ε/ and /ɔ/’ in unstressed positions, although all examples she mentions involve open-mid vowels followed by [+ATR] vowels.

in which a high vowel would ‘fail’ to trigger vowel harmony. For example, the postpositions *tʃi* ‘locative’ and *dɪ* ‘instrumental’, which would contain the same vowel in Fortune & Fortune’s and Cavalcante’s transcriptions, have totally inverse behaviors in regard to vowel harmony (49). Similarly, while the root *-u* ‘tooth’ triggers vowel harmony, the root *rʊ* ‘thigh’ does not (50):

- (49) a. /hãwəkɔ=tʃi/ [hawəkɔtʃi] ‘in the canoe’
 canoe=LOC
- b. /hãwəkɔ=dɪ/ [hawəkɔdɪ] ‘with the canoe’
 canoe=INSTR
- (50) a. ♀ /wa-ritʃɔɾɛ dʒ-u/ [waritʃɔɾɛ'dʒu]
 1-offspring REL-tooth
 ‘my child’s tooth’
- b. ♀ /wa-ritʃɔɾɛ rʊ/ [waritʃɔɾɛ'rʊ]
 1-offspring thigh
 ‘my child’s thigh’

Examples such as the ones above show that vowel harmony in Karajá cannot be described as a matter of height assimilation, since some morphemes containing high vowels, such as *tʃi*, would trigger vowel harmony, whereas others, such as *dɪ*, would fail to do it under the same circumstances. I claim that what underlies the differences in phonological behavior between morphemes such as *tʃi* and *dɪ* is the feature [ATR]. Consequently, I propose that vowel harmony in Karajá is better described as the regressive spreading of the feature value [+ATR] to [-ATR] vowels, a well-documented phenomenon in West African languages, but apparently rare in Brazilian languages.

This analysis is contingent on a revision of the phonemic inventory of Karajá. Contrasting with the previous phonological descriptions of the language, this analysis points to a larger inventory of vowels in which the feature [ATR] plays a major role (Table 1). The main difference is that in this analysis, a phonemic opposition is recognized between the high ‘tense’ vowels /i/, /ĩ/, and /u/ and their ‘lax’ counterparts /ɪ/, /ɨ/, and /ʊ/²⁶, as shown by the minimal pairs given below (4). This distinction was not considered in previous works.²⁷

- | | | | | | | |
|------|----|-------------|------------|----|-------------|----------------|
| (51) | a. | <i>lahi</i> | ‘to curse’ | b. | <i>lahɪ</i> | ‘grandmother’ |
| | c. | <i>bɨ</i> | ‘silence’ | d. | <i>bɨ</i> | ‘answer’ |
| | e. | <i>-uka</i> | ‘to split’ | f. | <i>-ʊka</i> | ‘to be cooked’ |

Besides the existence of minimal pairs, the phonemic character of the opposition is also shown by the fact that lax and tense vowels have quite different behaviors in the processes of vowel harmony and palatalization. As the examples below demonstrate, the high tense vowels /i/, /ĩ/, and /u/ and the close-mid vowels /e/, /ɛ/, and /o/ are dominant²⁸, triggering vowel harmony, whereas the high lax vowels /ɪ/, /ɨ/, and /ʊ/ do not trigger vowel harmony. This fact would be difficult to explain through an analysis in terms of the feature [high]. The fact that some hig

²⁶ I am using the terms [+ATR]/[-ATR] and tense/lax more or less interchangeably here. Noske (1995) finds “the [+ATR] and [-ATR] vowels of Turkana auditorily quite close to the tense and lax vowels of English and German.” I would say the same at least with respect to the front vowels /i/ and /ɪ/ of Karajá, which sound to me similar to the vowels in *beat* and *bit*, respectively, but, as Noske noted for Turkana, without the length distinction occurring in the English examples. The mid vowels resemble the ones of Portuguese. An acoustic analysis of the Karajá vowels is still to be done, though.

²⁷ In fact, Fortune & Fortune (1963) include /ʊ/ as a phoneme in their initial description of Karajá phonology. However, more recent works, such as Fortune (1973), do not mention it. Although this phoneme was initially represented in Fortune’s Karajá orthography, it was not used in more recent published works, such as the latest version of the New Testament (1983) and literacy materials. After I started discussing the distinction between [-ATR] and [+ATR] vowels with the Karajá speakers, they decided to use of the letter <ù> to represent /ʊ/ (the same letter initially adopted by the Fortunes).

²⁸ For the use of the terms ‘dominant’ and ‘recessive’ in the sense adopted here, see Rigsby & Silverstein (1969).

(52)	a.	/hãḏike chicken	ritʃɔre/ offspring	[hãnikeritʃɔre]	‘chick’
	b.	/hadike chicken	θɪ/ egg	[hãnike'θɪ]	‘chicken’s egg’
	c.	/hadike chicken	d-e/ REL-wing	[hãnike'de]	‘chicken’s wing’
(53)	a.	/budɔe deer	dʒ-u/ REL-tooth	[budɔe'dʒu]	‘deer’s tooth’
	b.	/budɔe deer	ru/ thigh	[budɔe'ru]	‘deer’s thigh’
	c.	♀ deer	/budɔe woku/ deer stomach	[budoewo'ku]	‘deer’s stomach’

This distinction has turned out to be crucial for the study of vowel harmony in Karajá in particular, and of Karajá phonology in general. By recognizing the distinction between lax and tense (or [-ATR] and [+ATR]) vowels, one can account for a number of otherwise unexplainable ‘exceptions’ in which a high front or back vowel would ‘fail’ to trigger vowel harmony. In addition, one is able to provide a consistent account of palatalization, as we will see.

Vowel harmony languages are characterized by constraints on which vowels may co-occur within a given phonological domain (typically, the phonological word). In Karajá, the parameter governing these constraints is, as claimed in this work, the feature [ATR]. The examples below show the patterns of vowel combinations in the phonological word in Karajá, according to the feature [ATR].

- (54) [+ATR] [+ATR]
- | | | | |
|----|-----------|-----------|----------------------------------|
| a. | /hedə/ | [he'də] | 'smoke (noun)' |
| b. | /kubə/ | [ku'be] | 'palm' |
| c. | /kuʃe/ | [ku'ʃe] | 'fish flour' |
| d. | /-urə/ | [u'rə] | 'tip' |
| e. | /kərotʃu/ | [kro'tʃu] | 'pamonha (a type of corn bread)' |
- (55) [+ATR] [-ATR]
- | | | | |
|----|-----------|------------|----------------------------|
| a. | /tʃuʃɔ/ | [tʃu'ʃɔ] | 'quati (a type of mammal)' |
| b. | /ritʃɔrɛ/ | [ritʃɔ'rɛ] | 'offspring' |
| c. | /hãd̥ikɛ/ | [hãni'kɛ] | 'chicken' |
| d. | /itʃərə/ | [itʃərə] | 'fried' |
- (56) [-ATR] [-ATR]
- | | | | |
|----|---------|----------|---------------------------------|
| a. | /rɔrɪ/ | [rɔ'rɪ] | '(a type of) basket' |
| b. | /dɔrɛ/ | [dɔ'rɛ] | 'parrot' |
| c. | /bɛrɔ/ | [bɛ'rɔ] | 'puba (a type of manioc flour)' |
| d. | /hɛd̥i/ | [hɛ'd̥i] | 'blanket' |
- (57) *[-ATR] [+ATR]
- | | | | |
|----|-------------------------------|---------------|---------------------------|
| a. | /həlɔkɔɛ-d̥i/ | [hãlɔkɔe'ni] | '(wild or domestic) cat' |
| | jaguar-similar.to | | |
| b. | /r-a-ruɛ-d̥ə=r-a/ | [rarue'd̥ərə] | 'he/she/it became blind.' |
| | CTFG-INTR-eye-close=CTFG-PERF | | |
| c. | /dɔrɛ d-e/ | [dore'de] | 'parrot's wing' |
| | parrot REL-wing | | |

As the examples above show, all combinations of vowels in a phonological word are possible, except [-ATR] vowels preceding [+ATR] vowels. This is the circumstance under which vowel harmony takes place (10). Thus, for example, the morpheme *dɔrɛ* 'parrot' undergoes [+ATR] vowel harmony when followed by the morpheme *-e* 'wing' (10c). Conversely, a [+ATR] vowel remains unchanged when followed by a [-ATR] vowel, such as in *ritʃɔrɛ* 'offspring' (8b). Therefore, vowel harmony in Karajá can be defined as a process of regressive spreading of the feature value [+ATR] to [-ATR] vowels.

Karajá presents an asymmetrical kind of vowel harmony, since only the feature specification [+ATR] seems to be phonologically active, triggering vowel harmony. In symmetrical vowel harmony systems—that is, in systems where both feature values are phonologically active—, a given feature value is generally considered to be a property of the entire morpheme. In Karajá that is certainly not the case. As shown above, [+ATR] vowels can be followed by [-ATR] vowels in the same morpheme. Therefore, a word such as *ritʃɔɾe* ‘offspring’ may either trigger (a) or undergo (b) vowel harmony:

- (58) a. ♀ /wa-θɛ-ritʃɔɾe/ [waθɛritʃɔɾɛ] ‘my sibling’
 1-mother-offspring
- b. ♀ /wa-ritʃɔɾɛ boho/ [waritʃɔɾɛboʰo] ‘my children’
 1-offspring PLURAL

The existence of disharmonic roots such as ♀ *ritʃɔɾe* ‘offspring’, which are rather common in the language’s lexicon, demonstrates that vowel harmony in Karajá is a strictly directional, right-to-left process.

According to their behavior in triggering, undergoing, or blocking vowel harmony, the vowels of Karajá can be grouped as in Table 3 below. The fact both groups of vowels, dominant versus recessive, include both high and mid vowels demonstrate that vowel harmony in Karajá is of the *cross-height* type, a terminology which fell into disuse once the mechanism underlying this type of vowel harmony was found to be not height, but tongue root position.

Table 2.3. Vowels according to their behavior in terms of vowel harmony

Oral

[+ATR]	opaque	[-ATR] ²⁹
i ɨ u		ɪ ɨ ʊ
e ə o		ɛ ɔ
	a	

Nasal

[+ATR]	opaque
ĩ	õ ã
	ã

Any morpheme containing a [+ATR] vowel can trigger vowel harmony, regardless of its morphological or stress status. Vowel harmony can be triggered by noun or verb roots, clitics (such as the locative postposition *tʃi* and the imperfective auxiliary (*r*)*e*), and suffixes. Vowel harmony occurs either morpheme-internally or across word-boundaries, with the phonological word (characterized by a single primary stress) being its apparent domain. I will exemplify the behavior of each vowel in relation to vowel harmony by taking verb forms involving the clitic auxiliary (*r*)*e* ‘imperfective.’

The mid open vowels /ɛ/ and /ɔ/ undergo [+ATR] assimilation in an iterative manner. As shown by examples (11) and (12), vowel harmony is not restricted to the verb root, but also affects the vowel of the prefix.

²⁹ By ‘[-ATR] vowels’ I refer to those vowels which will surface as [-ATR] unless they undergo [+ATR] vowel harmony. I do not refer to their underlying feature specification. In fact, the Karajá data likely support an analysis in which the [-ATR] feature value is not present underlyingly for the recessive vowel series, being rather introduced by a redundancy rule. Therefore, [+ATR] spreading would probably be a feature-filling rule. I shall also explain the use of the term ‘opaque’ in this paper. In the description of some languages, such as Akan, this term is used to describe vowels that not only block [+ATR] spreading, but also trigger [-ATR] vowel harmony (Kenstowicz 1994, 351). In Karajá, the opaque vowels /a, ã, õ, and ɔ/ simply block vowel harmony. As a first approximation, it may be hypothesized that opaque vowels are underlyingly specified as [-ATR] and, therefore, do not undergo [+ATR] vowel harmony, which would be a feature-filling process.

- (59) /r-ε-rɔ=r-e/ [re'rore]
 CTFG-1+TRANS-eat.solids=CTFG-IMPERF
 'I ate (it).'
- (60) /r-ε-he=r-e/ [re'here]
 CTFG-1+TRANS-scratch=CTFG-IMPERF
 'I scratched (it).'

As occurs with the open mid vowels, the high lax vowels /ɪ/, /ɨ/, and /ʊ/ also undergo vowel harmony. However, the further spread of vowel harmony seems to be optional:

- (61) /r-ε-hɪ=r-e/ [re'hire] ~ [re'hire]
 CTFG-1+TRANS-drive.away=CTFG-IMPERF
 'I drove (it) away.'
- (62) /r-ε-hʊ=r-e/ [re'hure] ~ [re'hure]
 CTFG-1+TRANS-finish=CTFG-IMPERF
 'I finished (it).'
- (63) /r-ε-hʊkɔdε=r-e/ [rehuko'dere] ~ [rehuko'dere]
 CTFG-1+TRANS-lend=CTFG-IMPERF
 'I lent (it).'
- (64) /r-ε-ki=r-e/ [re'kire] ~ [re'kire]
 CTFG-1+TRANS-eat.grains=CTFG-IMPERF
 'I ate (it).'

Therefore, high and mid [-ATR] vowels differ in the way they undergo vowel harmony. As illustrated by the examples (11)-(15) above, the open-mid vowels /ɛ/ and /ɔ/ undergo vowel harmony iteratively, while with the high lax vowels /ɪ/, /ɨ/, and /ʊ/, vowel harmony may optionally occur only locally. There is actually a great deal of variation in the extent to which vowel harmony can take place with high [-ATR] vowels, not only from

speaker to speaker, but also within the speech of a single individual.³⁰ Such variations need to be further investigated.

The examples below, involving the suffix *-dĩ* ‘similar to’, further illustrate the differences between mid and high [-ATR] vowels. While harmonization of [-ATR] high vowels may optionally take place iteratively when triggered by aspectual clitics such as *=r-e* ‘imperfective’, it seems to be strictly local when triggered by the derivational suffix *-dĩ*.

- (65) a. /krɔbɪ-dĩ/ [krɔbɪni] ‘a type of monkey’
 monkey-similar.to
- b. /kɔɖʊ-dĩ/ [kɔɖʊni] ‘a type of turtle’
 turtle-similar.to
- c. /brɔrɛ-dĩ/ [brɔrɛni] ‘cow’
 deer-similar.to
- d. /bedɔ-dĩ/ [bedɔni] ‘a type of fish’
 fish (sp.)-similar.to

Finally, the vowels /a/, /ɔ̃/, and /ɔ̄/ block vowel harmony, as illustrated by examples (18), (19), and (20) below:

- (66) /r-ɛ-ka=r-e/ [rɛ'kare]
 CTFG-1+TRANS-dig=CTFG-IMPERF
 ‘I dug (it).’
- (67) r-ɛ-hãdɛ=r-e [rɛhã'dere]
 CTFG-1+TRANS-hit=CTFG-IMPERF
 ‘I hit (it).’
- (68) /r-ɛ-bɔ̃=r-e/ [rɛ'mɔ̃re]
 CTFG-1+TRANS-take=CTFG-IMPERF
 ‘I took (it).’

³⁰ Differences in behavior between high and mid [-ATR] vowels are common in other languages as well.

- (69) /r-ε-ō=r-e/ [rε'ōre]
 CTFG-1+TRANS-give=CTFG-IMPERF
 'I gave (it).'

The example *beraku* 'river' illustrates well the way the three types of vowels—'dominant', 'recessive', and 'blocking'—interact. In this example, the spreading of the [+ATR] feature value of the vowel /u/ in the last syllable to the [-ATR] vowel /ε/ in the first syllable is blocked by the presence of the opaque vowel /a/ in the medial syllable. However, a quite different situation results in the male speech. In this case, the velar stop is dropped, making possible the fusion between the vowels /a/ and /u/, resulting in the mid-close vowel /o/, which then triggers vowel harmony in the first vowel:³¹

- | | | |
|------|----------------------|----------|
| (70) | | /beraku/ |
| | k-dropping | bɛrau |
| | vowel fusion | bɛro |
| | vowel harmony | bɛro |
| | ♂ | [bɛ'ro] |

6.1.1 Directionality

As we have seen, the fact that [-ATR] vowels can follow, but not precede [+ATR] vowels clearly shows that vowel harmony in Karajā is strictly a right-to-left process. This is further illustrated by the examples below, involving the stems *budε* 'few, little', *rikɔre* 'offspring', *tʃuhɔ* 'to curse', and *kɔfε* 'grassy'. Since these stems contain both dominant and

³¹ Except for cases such as this, involving vowel fusion, there are no differences between male and female speech concerning vowel harmony.

recessive vowels, they can either trigger (a) or undergo (b) vowel harmony. Not surprisingly, stems such as ♀ *ritʃɔɾɛ* (♂ *riɔɾɛ*) can trigger vowel harmony ‘on themselves’ (75).

- (71) a. *bəde-bude* [bədebu'dɛ]
land-few
‘island’
- b. *i-bude=r-e* [ibu'dere]
3-few=CTFG-IMPERF
‘It is little.’
- (72) a. *wa-θɛ-ritʃɔɾɛ* [waθɛritʃɔɾɛ]
1-mother-offspring
‘my sibling’
- b. *wa-ritʃɔɾɛ boho* [waritʃɔɾɛbo'ho]
1-offspring PLURAL
‘my children’
- (73) a. *Ø-r-ɔ-tʃuhɔ=rɛrɪ* [rɔtʃu'hɔɾɛrɪ]
3-CTFG-ANTI-curse=CTFG-PROGR
‘He is cursing.’
- b. *Ø-r-ɔ-tʃuhɔ=r-e* [rɔtʃu'hɔɾɛ]
3-CTFG-ANTI-curse=CTFG-IMPERF
‘He cursed.’
- (74) a. *bəde kɪʃɛ* [bədekɪʃɛ]
land grassy
‘grassy land’
- b. *i-kɪʃɛ=r-e* [ikɪʃɛrɛ]
3-grassy=CTFG-IMPERF
‘It is grassy.’
- (75) ♂ *wa-riɔɾɛ-riɔɾɛ* [wariorɛriɔɾɛ]
1-child-child
‘my grandchild’

As mentioned above, examples such as these, in which directionality is clearly at play, pose an interesting challenge to theories which reject directionality as an independent parameter of assimilation, such as the one proposed by Bakovič (2000). Bakovič claims that “agreement constraints are *left-right symmetrical*” (p. 6), and that directionality is in fact an epiphenomenon derived from morphological considerations. This claim seems to be rather plausible in the cases of languages presenting stem-controlled vowel harmony. As he states, the majority of languages with vowel harmony (such as Turkish and Hungarian) are strictly suffixing, and present *stem-controlled* vowel harmony. Thus, despite the appearances that vowel harmony in these languages is unidirectional, left-to-right, this directionality would be merely a consequence of the morphological structure of the language (p. 7). In other vowel harmony languages, such as Yoruba, “morphology is strictly prefixal; the apparent right-to-left directionality of [ATR] harmony is thus a reflection of stem control” (p. 61).

As for dominant-recessive harmony systems, Bakovič’s proposal seems to be based on the assumption, tacitly or explicitly stated in the literature on vowel harmony, that dominant-recessive harmony systems are inherently bidirectional. Examples such as *budε* ‘few’ and *rikɔrε* ‘offspring’, presented above, in which directionality is clearly at play, are, according to Bakovič, ‘unattested’:

“If dominant-recessive harmony could in principle be unidirectional, then we would expect to find a language in which the recessive vowels on one side of a dominant vowel are affected by harmony, while those on the other side remain unaffected. *Such a pattern is entirely unattested.*” (Bakovič 2000: 8; italics added)

The Karajá data, as we have seen, demonstrate that this is definitely not the case. Such a pattern is actually rather common in Karajá, not only in polymorphemic constructions, but in

monomorphemic words as well (55). Thus, Karajá provides a strong counterexample to such claims, showing that strict directionality can also be found in dominant-recessive vowel harmony systems, constituting in such cases an independent parameter of assimilation.

6.1.2 Domain

As the data presented above suggest, vowel harmony in Karajá is extremely pervasive, applying not only in and across compounds, such as those formed by noun incorporation (76), but also across word boundaries (77):

- (76) a. *Ø-r-a-rue-də=r-a* [rarue'dəra]
 3-CTFG-INTR-eye-close=CTFG-PERF
 'He/she became blind.'
- b. *Ø-r-a-we-bəhə=r-e* [rawebo'hore]
 3-CTFG-INTR-belly-break=CTFG-IMPERF
 '[They] had their bellies broken.'
- (77) a. ♀ *hāləkəe* *kɨdʒa* [hāləkə,ekɨdʒa]
 jaguar small
 'small jaguar'
- b. ♀ *wa-ritʃore* *dʒ-u* [waritʃore'dʒu]
 1-offspring REL-tooth
 'my child's tooth'

The domain of vowel harmony seems to be the phonological word, characterized by a single primary stress. Although stress seems to be useful in determining the domain of vowel harmony, it is irrelevant in characterizing triggers, since, as we have seen, vowel harmony

can be triggered not only by stems and derivational suffixes (which are intrinsically tonic), but by clitics (which are intrinsically unstressed) as well.

Since Karajá is an SOV language and the object tends to form a prosodic unit with the verb, vowel harmony can take place between the verb and the object. Considerations of frequency seems to play a role in predicting whether vowel harmony will take place. That seems to be especially common with idiomatic and formulaic constructions (such as curses). A case in point is the noun *bəde* ‘time, ground, world’, which occurs in a number of idiomatic expressions as an ‘empty’ object (*bəde* + *iri* ‘remember’ > ‘to become sad’; *bəde* + *buka* ‘disagree; not fit’ > ‘to be unlucky, when fishing, hunting, etc.’; *bəde* + *keri* ‘to know’ > ‘to realize’). This word also occurs in non-idiomatic constructions, as the subject of ‘atmospheric’ verbs (‘to cool down’, ‘to become silent’, ‘to become dark’, etc.) or as an object. In its more idiomatic uses, *bəde* can cliticize to the verb or even be incorporated into it; as the examples below demonstrate, vowel harmony triggered by the first vowel in the verb root *buka* extends all the way to the object *bəde* (78). Vowel harmony between the verb and its object is further illustrated by (79), a curse, where the clitic *=r-e* ‘imperfective’ triggers vowel harmony throughout the verb, reaching the object *rue* ‘eye’.

- (78) a. *waha* *bəde* \emptyset -*r-I-buka=ra* [bədəribu'kara]
my.father world 3-CTFG-TRANS-disagree=CTFG-PERF
‘My father got unlucky.’
- b. *waha* \emptyset -*r-a-bəde-bukara* [rabədəbu'kara]
my.father 3-CTFG-INTR-world-disagree=CTFG-PERF
‘My father got unlucky.’
- (79) *kofidabāru* *a-rue* \emptyset -*r-I-dɔ=kre* [arueri'dokre]
hawk 2-eye 3-CTFG-TRANS-suck=FUT

‘May the hawk eat your eyes.’³²
 (lit. ‘The hawk will eat your eyes.’)

6.1.3 Vowel harmony and grammatical function

As metaphony in Portuguese (which serves to distinguish nouns from verbs, and secondarily reinforces gender and number distinctions), vowel harmony in Karajá (or its lack thereof) may play a role in grammatical distinctions. Although a finite verb in Karajá generally ends with an aspectual-modal clitic, such a clitic may be suppressed in casual speech. As we have seen, two of the aspectual-modal clitics, *=kre* ‘future’ and *=r-e* ‘imperfective’, can trigger vowel harmony. The latter, by far the most frequent aspectual-modal clitic, can be suppressed while still triggering vowel harmony (80)—a case of assimilation *in absentia*, as I mentioned elsewhere (Ribeiro 2005:103). Vowel harmony may be, in such cases, the sole mark of aspect. On the other hand, the lack of vowel harmony is, to a certain extent, a mark of aspect as well, in examples such as (81) and (82) below; the fact that such examples do not display vowel harmony is in itself an indication, by exclusion, that such verbs are *not* imperfective (marked by the harmony-triggering clitic *=r-e*) or future (marked by *=kre*).

(80) *bɛra=tʃi* \emptyset -r-ɛ-θɛ= \emptyset [re'θe]
 water=LOC CTFG-INTR-fall=IMPERF
 ‘[He] jumped into the water.’

(81) *kədʒiwe* *dalahɪ* \emptyset -r-ɪ-wɛ
 K. 3REFL-grandmother 3-CTFG-TRANS-penetrate
 ‘Kynyxiwe has had sex with his grandmother.’³³

³² This curse is said as a response to the singing of the *sĩka* bird, which is considered a bad omen.

³³ This sentence is a verse in a song sung by a character from a traditional myth.

- (82) idʒõ b-ɪ-rɔ
some 2-TRANS-eat
'Eat some!'

6.1.4 Metalanguage: 'heavy' versus 'light'

While discussing issues related to Karajá orthography with native teachers, I used the terms 'strong' and 'weak' to refer to the high vowels /u/ and /ʊ/, respectively. The terms were translated by the Karajá teachers as *ikutʃie* 'heavy' and *iwɛdʌri* 'light', which provide a rather useful (albeit impressionistic) description of the oppositions between [+ATR] and [-ATR] vowels in Karajá.³⁴ The adoption of such native metalinguistic terminology by the Karajá has facilitated the discussion of [ATR] distinctions in the language, as native speakers' intuitions (and my non-native impressions) on whether a vowel is *heavy* or *light* tend to match its (morpho)phonemic behavior concerning vowel harmony and palatalization.

6.2 Palatalization

Considering the reanalysis of the vocalic inventory proposed here, it is possible to question the phonemic character of the complete series of palatals, /tʃ, dʒ, and ʃ/. A careful examination of the distribution of palatal consonants in Karajá reveals that they occur in very restricted environments—generally in contiguity to high [+ATR] vowels. Thus, as the examples below show, the interdental fricative and the palatal fricative are in complementary distribution: [ʃ] occurs after the high [+ATR] vowels /u/ and /i/, whereas [θ] occurs elsewhere.

³⁴ The Karajá terminology is reminiscent of the distinction between *big* and *thin* in Kujamutay (Greenberg & Sapir 1978).

- (83) *bɪθa* ‘macaw’ *-kiθe* ‘scratch’ *-ʊθa* ‘forget’
ifa ‘bowl’ *kufe* ‘fish flour’ *rufa* ‘raw’

Therefore, the distinction between [+ATR] and [-ATR] high vowels proves to be crucial also for the analysis of the consonantal system. The allophonic nature of the variations shown above is missed if the oppositions between /i, i, u/ and /ɪ, ɪ, ʊ/ are not recognized. It can be reasonably concluded that the interdental fricative and the palatal fricative are allophones of the same phoneme.

The study of some morphophonemic alternations also corroborates the analysis suggested above. Thus, the nominalizer suffix *-θV* (where V stands for a vowel identical to the last vowel in the verb root) is palatalized when attached to a verb root ending in the high front [+ATR] vowel /i/, but not after its [-ATR] counterpart /ɪ/:

- | | | | |
|------|--------------|-----------------|--------|
| (84) | Verb | Noun | |
| | <i>-aha</i> | <i>-aha-θa</i> | ‘find’ |
| | <i>-ʊɔ</i> | <i>-ʊɔ-θɔ</i> | ‘fly’ |
| | <i>-lədɪ</i> | <i>-lədɪ-θɪ</i> | ‘put’ |
| | <i>-obi</i> | <i>-obi-ʃi</i> | ‘see’ |

The same can be postulated for the remaining palatal consonants, the affricates /tʃ/ and /dʒ/. As the examples below show, they also occur generally in contiguity to high [+ATR] vowels. Notice the contrast with the [-ATR] high vowels /ʊ/ and /ɪ/, which occur with non-palatal consonants:

- | | | | | |
|------|--------------|------------|-------------|--------------|
| (85) | <i>tʃu</i> | ‘sun’ | <i>dʊ</i> | ‘loin cloth’ |
| | <i>tʃi</i> | ‘locative’ | <i>dɪ</i> | ‘bone, leg’ |
| | <i>butʃi</i> | ‘pot’ | <i>kɔdɪ</i> | ‘tobacco’ |

<i>adzikura</i>	‘manioc’	<i>-ãdɪ</i>	‘mother’
<i>hodzu</i>	‘pole’	<i>-dʊ</i>	‘nominal suffix’

One of the sources of *tf* in Southern and Northern Karajá is the palatalization of the velar stop /k/ after the [high, +ATR] vowel /i/, a process that does not occur in Javaé and Xambioá (86). Again, the high front [-ATR] vowel /ɪ/ does not trigger palatalization: *bɪkɔwa* ‘friend’ [bɪkɔwa].

(86)	Javaé, Xambioá	Karajá	
	<i>rikɔɛ</i>	<i>ritʃɔɛ</i>	‘offspring’
	<i>ikɔɔ</i>	<i>itʃɔɔ</i>	‘fox’

Another source of *tf* in all the four dialects is the palatalization of an initial alveolar implosive /d/ in a verb root when preceded by the prefix *i-*, a fossilized prefix that appears with deverbal nouns (87a-b). The same prefix triggers palatalization with roots beginning with /θ/ (87c), /d/ (87d) and /l/ (87e).

(87)	Verb	Noun	
a.	<i>-dɛhɛ</i>	<i>i-tʃɛɛ</i>	‘look’ ³⁵
b.	<i>-dɔki</i>	<i>i-tʃɔki</i>	‘carry’
c.	<i>-θɛ</i>	<i>i-ʃɛ</i>	‘dance’
d.	<i>-dɔ</i>	<i>i-dʒɔ</i>	‘fight (cursing)’
e.	<i>-ləki</i>	<i>i-dʒɔki</i>	‘tell’

The same alternations are found with personal and relational prefixes.³⁶ The 3rd-person prefix *d-* is pronounced *tf-* when attached to roots beginning with /u/, /ɥ/, or /i/:

³⁵ This example presents *consonantal replacement*, a common derivational process to create nouns from verb stems. It consists in the replacement of a velar stop or a glottal fricative occurring in the last syllable of the verb root with an alveolar flap in the corresponding noun form. (29c) and (29d) also show examples of consonantal replacement. This process may appear combined with affixation.

(88)	<i>d-adɪ</i>	‘his mother’	<i>tʃ-u</i>	‘his tooth’
	<i>d-ɔdɪ</i>	‘his throat’	<i>tʃ-urə</i>	‘its tip’
	<i>d-uahɪ</i>	‘his medicine’	<i>tʃ-iru</i>	‘its branches’
	<i>d-e</i>	‘its wing’	<i>tʃ-ira</i>	‘its pit’
	<i>d-ira</i>	‘its cleanness’	<i>tʃ-ɪdʒa</i>	‘its chills’

Similarly, the ‘relational’ prefix *d- ~ l-* is pronounced as *dʒ-* before /u/ and /ɪ/:³⁷

(89)	<i>waha d-ɛbɔ</i>	‘my father’s hand’	<i>waha dʒ-u</i>	‘my father’s tooth’
	<i>waha l-uahɪ</i>	‘my father’s medicine’	<i>wihɪ dʒ-urə</i>	‘arrow tip’
	<i>d-ɛbɔ l-ira</i>	‘clean hands’	<i>wa-dʒ-ɪdʒa</i>	‘my chills’

Finally, the nominalizer suffix *-dV* is subject to the same palatalization rule shown in

(90) above for the suffix *-θV*:

(90)	Verb	Noun	
a.	<i>-hʊ</i>	<i>-hʊ-dʊ</i>	‘finish’
b.	<i>-bɔ̃</i>	<i>-bɔ̃-dɔ̃</i>	‘take’
c.	<i>-hɛ</i>	<i>-rɛ-dɛ</i>	‘scarify’
d.	<i>-hɪ</i>	<i>-rɪ-dɪ</i>	‘drive away’
e.	<i>-lahɪ</i>	<i>-lahɪ-dʒɪ</i> ³⁸	‘curse’

Therefore, the evidence presented here demonstrates that the palatal consonants *tʃ*, *dʒ*, and *f* trace back to non-palatal consonants occurring in contiguity to high [+ATR] vowels, a

³⁶ For more on relational prefixes, see Chapters 4 and 5.

³⁷ Although one would expect the relational prefix with i-initial d-class stems to be *dʒ-*, that is not the case. The only two stems I am aware of, *-iru* and *-ira*, occur with no linking prefix: cf. *kɔwɔru iru* ‘tree branches’. The relational prefix is reconstructed as *j- for Proto-Jê, a form which may be close to the one present in Proto-Macro-Jê. If the form of the relational prefix was something similar to *j- in Pre-Proto-Karajá, the prefix could have coalesced with the initial vowel of i-initial d-stems.

³⁸ Besides palatalization of the suffix consonant (*lahɪ-dɪ > lahidʒɪ*), the deverbal noun *lahadʒɪ* ‘the action of cursing’ exemplifies two other phonological processes, dissimilation (> *lahidʒɪ > lahɔdʒɪ*) and total harmony across /h/ (> *lahɔdʒɪ > lahadʒɪ*).

fact that can only be captured when the phonological opposition between the high [+ATR] vowels /i/, /ī/ and /u/ and their [-ATR] counterparts /ɪ/, /ɪ̄/, and /ʊ/ is recognized.

Palatalization is clearly a lexical, word-internal process. Therefore, although it can be triggered between affixes and roots and between elements in a compound, it does not occur across word boundaries (between nouns and adpositions, for example, or between different words in an NP). As the examples below show, palatalization can be a useful criterion in distinguishing compounds and phrases. In (91a), a compound, the last vowel of *bāki* ‘corn’ triggers palatalization on the initial consonant of *θəbō* ‘small’; on the other hand, palatalization does not occur in (91b), a phrase:³⁹

- (91) ♀ a. *bāki-θəbō* [ma,kiθəmə] b. *bāki θəbō* [ma,kiθəmə]
 corn-small corn small
 ‘rice’ ‘small (type of) corn’

In addition, inflectional prefixes *d-* ‘centripetal direction’ and *d-* ‘2nd person’ do not undergo palatalization when attached to stems beginning with a high [+ATR] vowel, thus contrasting sharply with *d-* ‘3rd person’ and *d- ~* ‘relational’ which, as we have seen, undergo palatalization under the same circumstances.

- (92) ♂ a. *∅-d-∅-ua=d-e* [du'ade]
 3-CTPT-INTR-split=CTPT-IMPERF
 ‘It split (hither).’
 b. *∅-r-r-kʊrʊ-dʒ-ua=r-e*
 3-CTFG-TRANS-forehead-REL-split=CTFG-IMPERF
 ‘He split his forehead.’

³⁹ Additional examples involving palatalization of the initial consonant of *θəbō* in compounds: ♀ *wekiri-θəbō* [wekiriθəmə] ‘small boy’; *ru-θəbō* [ruθəmə] ‘small EYE (said of small, round things, such as beans and coins’, as in *kəbēda ruθəbō* ‘small beans’).

- (93) a. *d-∅-uhu=d-e* [d'uɦude]
 2-CTPT-INTR-stop=2-IMPERF
 'You stopped.'
- b. *∅-r-I-tf-uhu-dǎ=r-e*
 3-CTFG-TRANS-3-stop-VERB=CTFG-IMPERF
 'He stopped it.'

Palatalization is, thus, clearly a word-internal, lexical rule. It is a fairly useful indicator of morphological constituency, distinguishing compounds from syntactic phrases (91), and distinguishing “old” morphology from presumably more recent constructions. Notice that the prefixes *d-* and *I- ~ d-*, which undergo palatalization, have clear cognates in Proto-Jê and in other families of the Macro-Jê stock (see Chapter 6), a fact which attests to their antiquity.⁴⁰

6.2.1 Exceptions

⁴⁰ In Lexical Phonology terms, prefixes such as *d-* ‘3rd person’ and *I- ~ d-* ‘relational’, as well as fossilized *i-*, are reminiscent of “Level 1” affixes in English, such as *-ity* (in *electric-ity*), while prefixes such as *d-* ‘centripetal direction’ and productive *i-* ‘3rd person’ would belong to shallower morphological levels. The fossilized prefix *i-* probably has the same origin as the 3rd person marker. Both clearly contrast synchronically, though, as illustrated by the fact that they can cooccur (cf. *idzəkí* ‘story’ vs. *i-idzəkí* ‘his story’). The contrast between both prefixes is further illustrated below by the root *rədǎ* ‘feces’ (*idzədǎ* ‘feces’ vs. *i-rədǎ* ‘his feces’).

- | | | | | | | | |
|----|------------------------------|----|-------------------------------|----|------------------------------|----|------------------------|
| a. | <i>idzədǎ</i>
‘feces’ | b. | <i>i-rədǎ</i>
‘his feces’ | c. | <i>wa-rədǎ</i>
‘my feces’ | d. | <i>a-rədǎ</i>
‘your |
| c. | <i>wa-rədǎ</i>
‘my feces’ | d. | <i>a-rədǎ</i>
‘your feces’ | | | | |

As the discussion above has shown, palatal consonants in Karajá only occur around high [+ATR] vowels. There are a few apparent exceptions to this claim, but all of them seem to find a straightforward explanation. One of the exceptions is the word *tfakohi*, the name of a ritual mask, which is clearly a Tupi-Guarani loanword (likely from Tapirapé *tfakuʔi*; cf. Ribeiro 2002). Another apparent exception is the emphatic particle $=f\varepsilon$, which presents a palatal consonant despite lacking a high [+ATR] vowel. Another phonological peculiarity involving $=f\varepsilon$ is the fact that it triggers vowel harmony, although it does not contain a [+ATR] vowel superficially. Combined, both apparent anomalies shed light on this morpheme's apparent exceptionality: $=f\varepsilon$ would contain a high [+ATR] vowel underlyingly, $*/\theta i\varepsilon/$, which would account for both facts.

- (94) *bɔ̃=wɛ=f\varepsilon* [mɔ̃'wɛf\varepsilon]
 HUM=INTER=EMPH
 'Who (the heck)?'

Another apparent exception is the verb root *-uri* 'to extinguish (fire)'. Although it begins with a [high, +ATR] vowel, it takes the non-palatal prefixes *d-* '3rd person' and *l-*. Notice that the root triggers vowel harmony in the preceding [-ATR] vowels.

- (95) a. *r-ɛ-d-uri-dɔ̃=r-a* [redʉrinɔ̃ra]
 CTFG-1+TRANS-3/REL-extinguish-VERB=CTFG-PERF
 'I've extinguished (it).'
- b. ♂ *wɔ-l-uri-fi-dɔ̃-dv* [wɔluriʃidɔ̃'dv]
 fire-REL-extinguish-NOM-CAUS-SUBJ
 'firefighter' (lit. 'the one who extinguishes fire')

A likely explanation would be that *uri* underlyingly contains a [-ATR] vowel in its first syllable: /uri/. The surface [+ATR] quality of its first vowel would be a consequence of vowel harmony triggered by the [+ATR] vowel occurring in its last syllable. The rules of vowel harmony and palatalization would be in a counter-feeding order. The initial vowel may very well have been a fossilized intransitive marker, *u-*, attested with other stems (see Chapter 4).

6.3 Interactions between vowel harmony and palatalization

As we have seen, palatalization is clearly a lexical process in Karajá, whereas vowel harmony is likely a post-lexical one, occurring across word-boundaries; vowel harmony and palatalization would be, in traditional terms, in *counter-feeding* order. There are, however, at least two cases in which vowel harmony clearly feeds palatalization. These can be seen as old compounds, which are not synchronically perceived as such. One is the kinship term *ãdzikura* ‘mother’s older sister’, which results probably from the combination of the morphemes *ãdi* ‘mother’ and *kura* ‘white’⁴¹ (compare it with *waha* ‘my father’ *versus* *wahakura* ‘my father’s older brother’). In this case, the high [+ATR] vowel /u/ triggers vowel harmony in the [-ATR] vowel of the preceding root, which then triggers palatalization on the voiced stop /d/, as shown by the derivation given below:

(96)	input	/adɪ-kura/
	vowel harmony	adikura
	palatalization	adzikura
	output	[adziku'ra]

⁴¹ For a discussion of the use of color terms in Karajá kinship terminology, see Donahue (1982:158).

Another example in which vowel harmony seems to have fed palatalization is the particle *tʃibo* ‘approximately’, which is probably the combination of the pro-form *dɪ* and the interrogative particle *=bo* (*dɪbo* [dʲibo] ‘how?’ > *tʃibo*):⁴²

(97)	input	/dɪ-bo/
	vowel harmony	dʲibo
	palatalization	tʃibo
	output	[tʃibo]

5.4 Reduplication

Reduplication is very commonly used with verb stems, signaling iteration or repetition. The entire verb root is reduplicated (98). Again, a bimoraic minimality constraint is at play: when the root is monosyllabic, reduplication includes all the verb prefixes:

- (98) a. b-∅-ɪ-krɔ=kre [bi'krokre]
 2-CTFG-TRANS-cut=FUT
 ‘You will cut it.’
- b. b-∅-ɪ-krɔ-krɔ=kre [bikro'krokre]
 2-CTFG-TRANS-cut-REDUP=FUT
 ‘You will chop it.’
- (99) a. idʒɔrɔθa irɔdʊ dɪ ∅-r-ɪ-rɔ=r-a
 dog animal bone 3-CTFG-TRANS-eat=CTFG-PERF
 ‘The dog ate the bone.’

⁴² The semantic development involved here is akin to what is found in Spanish, in which interrogative *cómo?* is used to signal imprecision (as *like* in English): *Ella tiene como veinte años* ‘She is, like, twenty years old.’

- b. idʒəɾəθa irədʊ dɪ Ø-r-ɪ-rə=rɪrə=r-ɛɪ
 dog animal bone 3-CTFG-TRNS-eat=REDUP=CTFG-PROG
 ‘The dog is gnawing the bone.’

With nouns, reduplication indicates plurality:

- (100) a. irədʊ
 animal
 ‘animal’
- b. irədʊ irədʊ
 animal animal
 ‘animals’

6. Concluding remarks

This chapter provided an account of Karajá phonology. Compared with previous analyses of the language, the analysis presented here proposes a larger vowel inventory; a major contribution of the present analysis lies in demonstrating the relevance of the feature [ATR] ‘advanced tongue root’, as evidenced by a pervasive process of vowel harmony. The reanalysis of the vocalic inventory has substantial consequences for the analysis of the consonantal inventory as well: palatal consonants, considered as independent phonemes in previous analysis, are the result of a process of palatalization around [high, +ATR] vowels. Although the distribution of palatal consonants is fairly predictable from a morphologically-informed point of view, a process of phonogenesis seems to be taking place: in addition to the adoption of loanwords which violate native phonotactic constraints, there are even minimal pairs (cf. *idʒədɔ̃* ‘feces’ vs *i-rədɔ̃* ‘his feces’), albeit morphologically conditioned. Such factors result in a fluid situation in which palatal consonants may end up acquiring phonological status.

Vowel harmony is among the most remarkable characteristics of Karajá phonology. Common in African languages, [ATR] vowel harmony is exceedingly rare outside of Africa; Karajá presents the first documented case of this type of vowel harmony in South America. Since most languages presenting [ATR] vowel harmony are genetically or areally related, the Karajá case provides a different perspective for the typological understanding of the phenomenon (Ribeiro 2002). Once again, the chapter does not provide the final word on the subject, indicating instead additional topics for investigation; for instance, a thorough acoustic analysis of the vowel inventory is yet to be conducted, in order to investigate whether the mechanisms underlying [ATR] distinctions in Karajá are similar to those found in better-known languages.

CHAPTER 3

Female and Male Speech

1. Introduction

This chapter describes the differences between female and male speech in Karajá, a peculiarity of this language which drew attention of ethnographers and linguists early on, but which remains largely misunderstood.¹ Ehrenreich, the first to describe this phenomenon, does it in the following terms:

The most notable peculiarity [of the Karajá language] is the existence of a special language for the men and another one for the women, more or less like **it** happens among the Guaikuru and Chikitano. However, only a few words are totally different; in the majority of them, only slight modification in the form is noticed. For example, in those cases where, in the male dialect, two vowels follow each other, a *k* is inserted between them in the female dialect. Thus,

‘rain’ ♂ *biū*, ♀ *bikū* ‘corn’ ♂ *mahĩ*, ♀ *makĩ*

At times the female word simply has one additional final syllable, etc. *It is likely that the women have simply preserved an older form of the idiom.* [Ehrenreich 1948:29, my translation; emphasis added]

The differences between the female and male ‘dialects’ in Karajá have been further studied by Kunike (1916, 1919), Fortune and Fortune (1975) and, more extensively, by Borges (1994, 1997). Corroborating

¹ The existence of distinctions between the speech of males and females has been documented for a number of other Lowland South American languages since the beginning of the colonization of the continent, but most of them seem to be limited to sporadic lexical differences, involving mostly interjections and a few grammatical morphemes (see, for example, Mamiani (1877[1699]:96-97), for Kariri, an extinct Macro-Jê language from northeast Brazil, and Rodrigues (1952:68) for Tupinambá, a Tupí-Guaraní language once spoken along the Brazilian coast). Besides Karajá, the only other Brazilian language in which male and female speech are also characterized by systematic [and pervasive] phonological differences seems to be Pirahã (Mura-Pirahã family; Everett 1986).

Ehrenreich's intuitions, both Fortune and Fortune (1975) and Borges (1994) consider female speech as the more conservative, postulating more or less exhaustive phonological rules to derive male speech forms from the corresponding female speech forms. However, interesting phenomena have been traditionally overlooked. For example, the rules deriving male speech forms from female speech can dramatically alter the morphological shape of the words, a fact not even mentioned in the previous studies.

Furthermore, one of the main shortcomings of Fortune & Fortune's (1975) and Borges's (1994, 1997) works is their lack of a comparative, pan-dialectal perspective. Although differences between male and female speech are present throughout all the dialects (and, therefore, probably predate their diversification), all the previous studies limit themselves to one single dialect. Thus, while Ehrenreich's pioneering remarks are probably based mainly on data of Northern Karajá and Xambioá dialects, Fortune and Fortune's account is based on Northern Karajá, and Borges's on Southern Karajá.

Thus, data of the Javaé dialect were never seriously considered. According to Fortune & Fortune (1963), the differences between male and female speech would not occur in Javaé. Both males and females would speak what corresponds to the male speech in [Northern] Karajá. However, this information, based likely on the opinion of speakers of Northern Karajá, is not completely accurate. In fact, differences between male and female speech do occur in Javaé, although to a lesser extent. Interestingly, in Javaé the distinctions between female and male speech are much less remarkable than in the other three dialects, in the sense that many forms used exclusively by males in Xambioá, Northern Karajá, and Southern Karajá, are also used by females in Javaé.

Therefore, the main purpose of this chapter is to describe the differences between female and male speech in Karajá, taking into consideration for the first time data from the four different dialects and approaching facts that were not mentioned in previous studies. In addition, I investigate the nature of the remaining male versus female speech distinctions that still occur in Javaé. As I intend to demonstrate, these distinctions were not preserved at

random. Rather, they display a very clear pattern, revealed by the comparative analysis of the dialects. Finally, I challenge Rodrigues' (1999, 2002) hypothesis on the origins of the female versus male speech distinctions, according to which such a phenomenon would be a result of language contact. I propose, instead, that a likely explanation shall be sought in internal factors.

2. Back to phonology: some dialectal differences

As we have seen in the previous chapters, the four dialects share essentially the same phonological inventory, although there are slight differences with respect to the occurrence of certain phonological processes. In this section, I summarize some of those differences which will be relevant for the topic of the present chapter.

The main phonological difference consists in the occurrence in Southern and Northern Karajá of a schwa [ə] in unstressed positions, corresponding to environments in which Xambioá and Javaé present a vowel identical to the one occurring in the following syllable:

(1)	Karajá	Javaé, Xambioá	
	<i>kədɔ</i>	<i>kɔdɔ</i>	‘termite’
	<i>bədɪ</i>	<i>bɪdɪ</i>	‘honey’
	<i>-dəkə</i>	<i>-dəkə</i>	‘causative suffix’
	<i>rəkʊ</i>	<i>rʊkʊ</i>	‘gourd’

Another difference is the occurrence, in the Karajá dialects, of surface CCV syllables, resulting from a process of syncope of a schwa occurring between a stop and the alveolar approximant /r/. As expected, Xambioá and Javaé present a vowel identical to the vowel in the following syllable and no epenthesis takes place:

(2)	Karajá	Javaé, Xambioá	
	<i>kəɾɔ</i> [krɔ]	<i>kɔɾɔ</i> [kɔ'ɾɔ]	'frog'
	<i>kəɾe</i> [kre]	<i>keɾe</i> [ke're]	'martim-pescador (bird sp.)'
	<i>bəɾɔ</i> [brɔ]	<i>bɔɾɔ</i> [bɔ'ɾɔ]	'back'

As I have suggested in the previous chapter, there are reasons to believe that Southern and Northern Karajá are, with respect to the existence of the schwa, the most conservative dialects. The alternative would be to consider the scenario occurring in Javaé and Xambioá as the most conservative, postulating for Southern and Northern Karajá a rule of *lenition* of an unstressed vowel when followed by a syllable containing an identical vowel. However, this hypothesis is ruled out by the existence, in Southern and Northern Karajá, of minimal pairs such as *daka* [d'a'ka] 'to take off' versus *dəka* [də'ka] 'to tie' (corresponding to a homophonous pair in Xambioá), as well as *bɔɾɔ* [bɔ'ɾɔ] 'stingray' versus *bəɾɔ* [brɔ] 'back' (corresponding to a homophonous pair in both Javaé and Xambioá). As will be seen, the facts of the male-female speech distinction also argue in favor of the hypothesis that the schwa can be reconstructed for Proto-Karajá.

Another phonological process that differentiates Javaé and Xambioá from Southern and Northern Karajá is the palatalization, in the latter, of the velar stop /k/ when following the [high, front, +ATR] vowel /i/.² As will be shown, the resulting affricate, [tʃ], will still behave as /k/ in relation to the process of *k*-dropping.

(3)	Karajá	Javaé, Xambioá	
	<i>itʃɔɾɔ</i>	<i>ikɔɾɔ</i>	'fox'
	<i>ritʃoko</i>	<i>rikoko</i>	'doll'

3. Female versus male speech

This section describes the differences between female and male speech forms in Karajá, considering the scenario occurring in the Southern Karajá, Northern Karajá, and

² In Xambioá, the phoneme /k/ is heavily palatalized in this environment: *ikɔɾɔ* 'fox' [içɔ'ɾɔ], etc.

Xambioá dialects as the prototype. As will be shown later, it is rather reasonable to claim that Javaé, where most of the distinctions are now obliterated, evolved from a situation similar to the one still present in the other dialects.

3.1 *k*-dropping

Ehrenreich's brief description summarizes the most common differences between female and male speech in Karajá. The main difference, as he points out, is the absence, in the male speech, of a velar stop occurring in the corresponding female speech form (*kɔwɔrv* 'tree, wood', ♂ *ɔwɔrv*). The deletion of the velar stop can make possible the fusion between vowels, the result being that the male forms may have a smaller number of syllables (♀ *hãlɔkɔɛ* 'jaguar', ♂ *hãlɔɛ*), a fact also noticed by Ehrenreich.

The postulation of the female speech as the more conservative one seems to be easily arguable. The main consideration is *predictability*: that is, whereas it is generally possible to predict the male form vis-à-vis the female form, the inverse does not occur. There are several examples in which sequences of vowels occur both in the female and the male speech, such as ♀ ♂ *riu* 'hunt' and ♀ ♂ *vahr* 'medicine', which makes an alternative rule of *k*-insertion unlikely. Considerations of predictability are even more evident in cases where there is vowel fusion: in comparing ♀ *beraku* 'river' with ♂ *bero*, it is rather clear that the latter can be derived from the former, but not vice-versa. Loanwords also provide arguments for considering the female speech as the less innovative, since borrowings containing a velar

stop will have this consonant systematically suppressed in the male speech: *kawaru* ‘horse’, ♂ *awaru* (from Portuguese *cavalo*); *kubeda* ‘blanket’, ♂ *ubeda* (Portuguese *coberta*); *bãka* ‘mango’, ♂ *bãa* (Portuguese *manga*). On the other hand, there is no insertion of a consonant in loanwords containing sequences of vowels: *dïeru* [nie'ru] ‘money’ (from Portuguese *dinheiro*). Morphological considerations also support the postulation of female speech as more conservative (section 3.2). Finally, in the few cases for which comparative evidence is available, the velar stop can be shown to be a retention from the proto-language: Proto-Jê **ko* ‘wood’ :: *kɔ*; Proto-Jê **ku* ‘to eat’ :: *kɨ*.

The fact that female speech is historically more conservative obviously does not imply that male speech is derived from it. Instead, both gender dialects are probably derived from a common source, which the female speech happens to represent more faithfully, since it did not undergo the phonological processes which affected the male speech. Thus, although female speech forms will coincide with the ones postulated as the historical input for the male speech forms, this simply reflects the fact that female speech did not undergo processes such as *k*-deletion and vowel fusion:

Table 3.1. Historical development of the form ‘jaguar’

	♀	♂
Pre-Proto-Karajá	* <i>hãlɔkɔɛ</i>	* <i>hãlɔkɔɛ</i>
k-deletion	<i>does not apply</i>	<i>hãlɔɔɛ</i>
vowel fusion	<i>does not apply</i>	<i>hãlɔɛ</i>
output	<i>hãlɔkɔɛ</i>	<i>hãlɔɛ</i>

Table 3.2. Historical development of the form ‘river’

	♀	♂
Pre-Proto-Karajá	* <i>beraku</i>	* <i>beraku</i>
k-deletion	<i>does not apply</i>	<i>berau</i>
vowel fusion	<i>does not apply</i>	<i>bero</i>
vowel harmony ³	<i>does not apply</i>	<i>bero</i>
Output	<i>beraku</i>	<i>bero</i>

3.1.1. Vowel fusion

Examples such as ♀ *biku* ‘rain’ (♂ *biu*) and ♀ *bāki* ‘corn’ (♂ *bāi*), where the velar stop is dropped without any further phonological consequences, illustrate the simplest instance of the processes that differentiate female from male speech. However, there are several circumstances in which the deletion of the velar stop creates the conditions under which different vocalic change processes can take place. If the *k*-dropping rule results in two identical contiguous vowels, they undergo a process of crasis:⁴

(4)	♀ <i>daka</i>	♂ <i>d'a</i>	‘to take off’
	♀ <i>hālɔkɔɛ</i>	♂ <i>hālɔɛ</i>	‘jaguar’
	♀ <i>kɔhɔkɔɛ</i>	♂ <i>ɔhɔɛ</i>	‘to cross’
	♀ <i>ruku</i>	♂ <i>ru</i>	‘night’
	♀ <i>aka</i>	♂ <i>a</i>	‘to run’
	♀ <i>rukudɔ</i>	♂ <i>rudɔ</i>	‘to thrust’

³ In this example, vowel fusion feeds vowel harmony. As we have seen (Chapter 2), the low vowel /a/ blocks vowel harmony, and that is why the first vowel in ♀ *beraku* ‘river’ is not affected by the [+ATR] vowel /u/ in the last syllable. In the male speech, the deletion of the velar stop makes possible the fusion between /a/ and /u/, resulting in the [+ATR] mid-back vowel /o/, which then triggers vowel harmony in the first vowel (see Chapter 2, ‘Phonology’, for a thorough description of vowel harmony).

⁴ Since there is a constraint against surface monosyllabic words (see Chapter 1, ‘Phonology’), the word for ‘night’ is lengthened in the citation form in male speech: ♂ *ru* > *ruu* ‘night’, but ♂ *ru=bɔ* [ru'mɔ] ‘at night’, ♂ *ruku=bɔ* [ru'kumɔ].

The remaining coalescence processes take place when the low vowel /a/ is followed by a back vowel. When the deletion of /k/ results in a sequence of /a/ plus the high back vowel /u/, both vowels are fused, resulting in the mid back vowel /o/. If the *k*-dropping rule results in a sequence of /a/ plus the mid back vowel /ɔ/, the vowel /a/ is dropped:

- | | | | |
|-----|-------------------|-----------------|------------------|
| (5) | ♀ <i>beraku</i> | ♂ <i>bero</i> | ‘river’ |
| | ♀ <i>rakufi</i> | ♂ <i>rofi</i> | ‘to eat’ |
| (6) | ♀ <i>wakɔɾɛ</i> | ♂ <i>wɔɾɛ</i> | ‘a type of bird’ |
| | ♀ <i>rahakɔɾɛ</i> | ♂ <i>rahɔɾɛ</i> | ‘to turn over’ |

If /k/ is deleted between the mid back vowel /o/ and the high back vowel /u/, the latter is dropped:

- | | | | |
|-----|--------------------|------------------|-----------------------|
| (7) | ♀ <i>hedoku</i> | ♂ <i>hedo</i> | ‘house’ |
| | ♀ <i>kowoku</i> | ♂ <i>owo</i> | ‘mortar’ |
| | ♀ <i>ohokudira</i> | ♂ <i>ohodira</i> | ‘to shoot at’ |
| | ♀ <i>woku</i> | ♂ <i>wo</i> | ‘inside’ ⁵ |

A few examples may suggest that, with some lexical items, a syllable *ku* might have indeed been *inserted* in the female speech forms. For example, corresponding to Southern Karajá ♀ *wiku* ‘song’, ♂ *wiu*, the remaining dialects have ♀ *wiku*, ♂ *wi*.⁶

⁵ Notice that, with ♀ *woku* ‘war; Tapirapé’, a homophone of ♀ *woku* ‘inside’, vowel fusion does not occur in male speech: ♂ *wou*. The difference in behavior may have to do with differences in their morphological status. While ‘inside’ always never occurs by itself, ‘Tapirapé, war’ may occur freely.

3.1.2. Schwa assimilation

In cases in which the deletion of the velar stop results in a sequence of a schwa followed by another vowel, the schwa assimilates the features of the following vowel. The examples below are restricted to Southern and Northern Karajá (7), the ‘schwa-dialects’ (cf. Chapter 1). This process also applies to loanwords, such as *dakəfi* ‘taxi’ (♂ *daafi*).

(7)	♀ <i>dəka</i>	♀ <i>daa</i>	‘to tie’
	♀ <i>rəkʊ</i>	♂ <i>ruʊ</i>	‘gourd’
	♀ <i>dɔdəkɛ</i>	♂ <i>dɔdɛɛ</i>	‘to become hot’
	♀ <i>dəkɛ</i>	♂ <i>dɛɛ</i>	‘dative postposition’
	♀ <i>əkʊ</i>	♂ <i>oo</i>	‘to be burnt’

As mentioned above, the Xambioá and Javaé dialects lack the schwa. In these dialects, the cognates of the words above present a vowel identical to the vowel in the following syllable also in female speech: *rukʊ* ‘gourd’, *dɔdɛkɛ* ‘to become hot’, etc. However, these words still behave as they do in Karajá. Thus, the cognates of the Karajá verbs *dəka* ‘to tie’ and *daka* ‘to take off’ are homophonous in female speech in Xambioá, but not in male speech (they are not homophonous in Javaé for reasons that will be discussed below, in Section 4). This constitutes a further piece of evidence to consider the ‘schwa-dialects’ as more conservative.

⁶ Another example, somewhat more complicated, is ♀ *bədokudzəkɛ* ‘pirarucu fish’, ♂ *bəɔləkɛ ~ bəɔləɛ* (see 2.3 below).

(8) **Verbs ‘to take’ and ‘to tie’ in Xambioá**

♀ <i>daka</i>	♂ <i>da</i>	‘to take’
♀ <i>daka</i>	♂ <i>daa</i>	‘to tie’

These examples also show that the differentiation between male and female speech cannot be thought of as a synchronic process by which the male speech forms are derived from the corresponding female speech forms. If this were the case, one would expect that, in Xambioá, the roots for ‘to take’ and ‘to tie’ would have identical shapes in the male speech, as it happens in the female speech. Facts such as this allow us to postulate a relative dating for some phonological processes. It is likely that, when the process of schwa assimilation (**dəka* > *daka*) took place in Xambioá, the male-female speech distinctions were already established, corroborating the assumption that they were already present before the separation of the dialects. Furthermore, the fact that Tupí-Guaraní loanwords such as *hākuri* ‘agouti’ and *tʃakohi*, the name of a ceremonial mask, do not undergo the process of vowel fusion in the male speech (♂ *hāuri*, ♂ *tʃaohi*) suggests that these words entered the Karajá lexicon when this phonological process was no longer productive.⁷

3.1.3. Epenthesis

⁷ Cf. also the contrasting behaviors between native ♂ *θaku* ‘inside’ (♂ *θo*) and the loan *saku* ‘bag’ (♂ *sau*) (< Portuguese *saco*).

Finally, when the velar stop is preceded by the high front [+ATR] vowel /i/ occurring in an onsetless syllable, a voiced affricate is epenthesized between the vowels after the deletion of the velar stop in male speech:⁸

- | | | | | | |
|-----|---|----------------------|---|---------------|--------------|
| (9) | ♀ | <i>ikəɾɔ, itfəɾɔ</i> | ♂ | <i>idʒəɾɔ</i> | ‘fox’ |
| | ♀ | <i>ikəɾɛ, itfəɾɛ</i> | ♂ | <i>idʒəɾɛ</i> | ‘roasted’ |
| | ♀ | <i>ikoi, itfoi</i> | ♂ | <i>idʒoi</i> | ‘male group’ |

3.1.4. Morphological consequences

The rules of vowel fusion described above can apply across morpheme boundaries, a fact that can render the morphological segmentation less obvious in male speech. This occurs, for example, with intransitive verbs, which are generally marked by the prefix a-. In the male speech, this prefix can be fused with the first vowel of the verb stem under the same conditions described above:

- | | | | | |
|------|---|---------------------|---|-------------------|
| (10) | ♀ | <i>rakubədǎɾɛɾɪ</i> | ♂ | <i>robədǎɾɛɾɪ</i> |
| | | Ø-r-a-kubədǎ=r-ɛɾɪ | | Ø-r-a-ubədǎ=r-ɛɾɪ |

⁸ Note that the epenthetic consonant is only inserted when the high front [+ATR] vowel /i/ is in an onsetless syllable. It does not occur with the verbs ‘to walk’ or ‘to invite’, for example: ♀ *rika, ritfa* ‘to walk’, ♂ *ria*; ♀ *rarikəwɨ, raritəwɨ* ‘to invite’ > ♀ *rariəwɨ*. Taking into consideration only data from the Southern and Northern Karajá dialects, Borges (1994: 335) postulates a rule of voicing to account for pairs of the kind ♀ *itfəɾɔ* ‘fox’, ♂ *idʒəɾɔ*. However, as the examples above demonstrate, the voiceless affricate in these positions traces back to a velar stop diachronically. The same process of epenthesis occurs when a vowel-initial deverbal noun is preceded by the fossilized prefix *i-*, such as in *idʒɛθo* ‘the action of wrestling’ (*i-* + *ɛθo* ‘to wrestle’) and *idʒara* ‘the action of running’ (*i-* + *ara* ‘the action of running’, from ♀ *aka* ‘to run’).

		3-CTFG-INTR-spread=CTFG-PROGR 'They are spreading.'		3-CTFG-INTR-spread=CTFG-PROGR 'They are spreading.'
(11)	♀	<i>rakɔdʊkɔrɛri</i> Ø-r-a-kɔdʊkɔ=r-ɛri 3-CTFG-INTR-go.up=CTFG-PROGR 'S/he is going up.'	♂	<i>rɔdʊɔrɛri</i> Ø-r-a-ɔdʊɔ=r-ɛri 3-CTFG-INTR-go.up=CTFG-PROGR 'S/he is going up.'

These facts show that important aspects of the language's morphology may be overlooked if the morphological analysis is based only on data from male speech, which has traditionally been the case. One of the consequences of this 'male-centric' approach is the fact that a rather common device to derive nouns from verbs, *consonantal replacement* (Ribeiro 1996), was completely ignored in the previous analyses (Fortune & Fortune 1964; Fortune 1973; Maia 1998[1986]). This device consists in the substitution of an alveolar flap /r/ for a velar stop occurring in the last syllable of the verb root. Deverbal nouns thus formed are common to both female and male speech, suggesting that this derivational process dates back to a period when male-female speech distinctions were not yet in place.

(12)	Verb	Noun	
	<i>kɪ</i>	<i>rɪ</i>	'eat (grains)'
	<i>ka</i>	<i>ra</i>	'dig'
	<i>ritʃa, rika</i>	<i>rira</i>	'walk'
	<i>kukɔ</i>	<i>kurɔ</i>	'lift'
	<i>-ɔkʊ</i>	<i>ɔrʊ</i>	'ripen'
	<i>-ʊkɛ</i>	<i>ʊrɛ</i>	'become dry'

Since the velar stop is generally suppressed in male speech, the formal identity between the verb root and the deverbal noun can be rather obfuscated. For example, although the verb roots ♀ *aka* 'to run' and ♀ *ka* 'to dig' may become homophonous in male

speech (δ *a*), their deverbal nouns will still differ: *ara* ‘the action of running’ *versus* *ra* ‘the action of digging’. Therefore, a regular derivational rule can only be identified if the female speech is taken into consideration.⁹

3.2. *d*-dropping

This process consists in the deletion in the male speech of an alveolar nasal consonant occurring in the corresponding female speech form. It seems to involve only two morphemes

⁹ A misunderstanding of the male-female speech distinctions may also lead to erroneous ethnographic statements. Thus, Donahue (1982), appraising the antiquity of Karajá agriculture, suggests that the Karajá word for ‘corn’ entered the Karajá lexicon through a European language, and that, therefore, the cultivation of this product was a recent introduction into Karajá culture:

The Karajá word for corn is *mai*, which happens also to be the French word for corn, both surely derived from maize, a word of Carib [sic] origin. It could well be that corn was introduced to the Karajá by Europeans after contact, an interesting case of circuitous diffusion. If such an American staple of corn were introduced by Europeans, it might indicate that the Karajá have come to horticulture quite a bit later than many other central Brazilian tribes. [Donahue 1982: 83]

Donahue would probably have reached a different conclusion if the female speech form, *bāki* [ma'ki] (which, as Ehrenreich suggests, represents “an older form of the idiom”), were taken into consideration. Another example is the analysis of Javaé sexuality suggested by anthropologist Patrícia Rodrigues (1993), based on the mistaken assumption that the words for ‘vagina’ and ‘skin’ were homophonous:

“[...] the fact that *tyky* means both ‘skin’, ‘clothes’, and one of the words for ‘body’, and also ‘vagina’, could mean that, contrary to western conception, which restrict sexuality to the genital organs, the Javaé would understand the whole female body as an extension of the vagina, an eroticized whole. The skin that covers the entire body would be as eroticized and permeated with sexuality as the genital organ, a notion that transforms the woman (as wife/sexual partner) into a *sexual being* [...], as a whole, and not only as a possessor of sexual anatomic parts.” [Rodrigues 1993: 312]

Interesting as it may sound, this hypothesis does not find any support in the linguistic data. In fact, the words for ‘skin’ and ‘vagina’ may become homophonous in the male speech, but only when pronounced in isolation. Since there is a constraint against surface monosyllabic words, the word for ‘vagina’, *dĩ*, will be pronounced as *dĩĩ* when occurring by itself, and, given the phonological processes described above, the word for ‘skin, clothes’, *dəki* (*dĩki*, in Javaé and Xambioá) may be optionally pronounced as δ *dĩĩ*. Homophony between these two words is limited to these cases. They never become homophonous in the female speech. Even in male speech, they will never become homophonous when combined with other morphemes: *i-dĩ* ‘her vagina’, *i-dĩĩ* ‘her/his/its skin, clothes’.

(which may be etymologically related): the clitic indefinite article ♀ =*dõ* [nõ] (♂ *õ*), and the pro-form ♀ *adõ* [a'nõ] ‘something’ (♂ *aõ*). The indefinite article attaches to entire noun phrases (13). The pro-form ♀ *adõ* generally occurs in combination with emphatic, interrogative, and discourse particles (14). It also occurs in combination with the nominalizer suffix *-dã* (15)¹⁰.

- | | | | | |
|------|---|--|---|------------------------|
| (13) | ♀ | <i>hirari rufera=dõ</i>
girl beautiful=INDEF
‘a beautiful girl’ | ♂ | <i>hirari rufera=õ</i> |
| (14) | ♀ | <i>adõ=hε=r-a=bo</i>
something=EMPH=CTFG-PERF=INTER
‘What’s up?’ | ♂ | <i>aõherabo</i> |
| (15) | ♀ | <i>adõ-dã</i>
something-NOM
‘thing’ | ♂ | <i>aõdã</i> |

Therefore, this process occurs only with a closed set of words, all of them containing the morphemes *adõ* and =*dõ*. An apparent exception is the word ♀ *adõdã* [anõ'na] ‘pineapple’, which is probably a loanword of Tupí-Guaraní origin, either from Tapirapé or from a Tupí-Guaraní-based *lingua franca* spoken in Brazil in the first centuries of the colonial period (cf. Tupinambá *ananá*). This word also undergoes the process of *d*-dropping

¹⁰ The indefinite article is probably also present in the numeral ♀ *idãdadõ* [ina'danõ] ‘three’, ♂ *idãdaõ* [ina'daõ], as suggested by the fact that the stress in this word occurs in the penultimate syllable. In Karajá, the stress generally falls on the last syllable of the word (as shown in Chapter 2), the only exception being clitics, such as the indefinite article =*dõ*, which are intrinsically unstressed.

in the male speech, becoming ♂ *aōdā* [aō'na], probably due to analogy with its native homonym, *adōdā* 'thing'.¹¹

3.3. Lexical differences

In some cases, the forms occurring in the male and female speech may be totally different words. That is the case of the vocative interjections (♀ *wu*, ♂ *kɨ*) and of the interjections expressing surprise or admiration (♀ *bɨ*, ♂ *bɛbɛ*). Examples such as these--involving interjections and other "emotive words"--are rather common in other languages. Fortune and Fortune (1975: 115-6) also mention two other cases in which "the men's speech form is totally different from the women's"—the verbs 'to cry' and (in my translation) 'to become angry'. This is clearly a mistake. The roots *obu* and *hi* 'to cry' and *θira* and *ɛbɔɛ* 'to become angry' are used in both female and male speech to refer to processes happening to female and male subjects, respectively—a difference having no relation with the female *versus* male speech distinctions with which this chapter is concerned.¹² Similar examples are found in Yana, according to Sapir's (1949) description:

¹¹ In Javaé, as we will see, only the male forms of the words containing the morphemes *adō* and *dō* were preserved: *aōda* 'thing', *hirari=ō* 'a girl' etc. However, the loanword for 'pineapple' in Javaé maintains the original [n] in both male and female speech: *adōdā*. This example suggests that the drift towards male speech (if there was indeed one; see Section 6) was already consolidated by the time this loanword entered the language (cf. section 4).

¹² Similar mistakes were made by Ehrenreich (1894), in a time when knowledge of Portuguese among the Karajá was probably very rare, making communication between outsiders and the Karajá a challenging endeavor. He listed as male-female speech variants words that actually have different meanings, and are common to both varieties. That is the case of the words *watjiwi* 'cooking pot' and *bəθɛ* 'plate', listed by Ehrenreich as the male- and female-speech forms for 'pot', respectively (Ehrenreich 1894:23).

(...) there is a small number of verb stems which apply exclusively to activity carried on by a male or by a female; e.g., *ni-*, *nĩ-* “a male goes” but ‘*a-* “a female goes,” *bu-ri-*, *bu-rĩ-* “a man dances” but *dja-ri*, *dja-rĩ-*, “a woman dances.” In the latter case the difference of verb probably reflects an actual difference in the style of dancing. [1949:206]

Following Sapir, I shall suggest that, for ‘to cry’ and ‘to get angry’, the differences between the verbs in Karajá would probably also reflect actual differences in the processes of crying and getting angry. Notice that the verb *obu* ‘(a woman) to cry’ is also used to refer to the sounds made by birds and to the ritual wailing of deceased relatives, performed exclusively by women.

Fortune & Fortune (1963, 16) also include in the same category of ‘totally different’ words the verb ♀ *ohokudza* ‘to listen’ (♂ *ohola*). However, in spite of the somewhat striking surface differences between the form used in female speech and the one used in male speech, the latter can be regularly derived from the former by the processes described above: *k*-dropping and vowel deletion. The differences between the consonants occurring in the last syllable of both forms, *dʒ* versus *l*, can also be explained as the result of a regular phonological process. As seen in Chapter 2, the lateral /l/ becomes palatalized when contiguous to the [high, +ATR] vowels /i/, /ĩ/, and /u/. In the aforementioned example, palatalization does not occur in the male form because the vowel /u/, which would trigger it, is not present. A similar example is ♀ *bədokudʒəke* ‘pirarucu fish’, ♂ *bədɔləke* ~ *bədɔləe*.

3.4. Exceptions, obligatoriness, and variation

Although it is a very productive process, applying even to loanwords, there are a few exceptions to the *k*-dropping rule. These are mainly grammatical words, such as the personal pronoun *kai* ‘you’, the demonstratives *ka* ‘this’, *kua* ‘that (far from both the speaker and the addressee)’, and *kia* ‘that (close to the addressee)’, the locative postposition =*kɪ*, the verb particles =*kəre* ‘future marker’ and =*kɛ* ‘potential marker’, and discourse markers, such as =*rəkɪ* ‘narrative’. I am not aware of any noun or verb roots constituting absolute exceptions to the *k*-dropping rule.

Although the previous studies do mention the existence of such exceptions, they fail to notice a very remarkable fact. Even among the words for which the *k*-dropping rule is possible there are differences concerning the *degree of obligatoriness* with which the differentiation takes place. That is, while for some words *k*-dropping is obligatory, for others it is optional. According to this criterion, words presenting a velar stop may be classed into three different categories: (a) *absolute exceptions*, that is, words for which the deletion of /k/ does not occur at all; (b) words for which the *k*-dropping rule is optional; (c) words for which the *k*-dropping rule is obligatory.¹³ Tables 3, 4, and 5 below show examples of members of each class:

Table 3.3. Class a: absolute exceptions to the rule of *k*-dropping

FEMALE	MALE	
<i>kai</i>	<i>kai</i>	‘you’
<i>kɛ</i>	<i>kɛ</i>	‘potential marker’
<i>kɪ</i>	<i>kɪ</i>	‘locative postposition’
<i>kəre</i>	<i>kəre</i>	‘future marker’
<i>rəkɪ</i>	<i>rəkɪ</i>	‘narrative particle’

¹³ The differentiation between female and male forms is also obligatory for all the cases of d-dropping described in section 3.2.

Table 3.4. Class *b*: words for which the rule of *k*-dropping is optional

FEMALE	MALE	
<i>dəka</i>	<i>dəka, dəa</i>	‘to tie’
<i>dəkɪ</i>	<i>dəkɪ, dɪɪ</i>	‘he, she, it’
<i>kədǎde</i>	<i>kədǎde, ədǎde</i>	‘flour’
<i>kəhəde</i>	<i>kəhəde, əhəde</i>	‘club’
<i>dəki</i>	<i>dəki, dɪi</i>	‘skin, bark, cloth’
<i>kəɾə</i>	<i>kəɾə, əɾə</i>	‘to break’
<i>kəɾəbɪ</i>	<i>kəɾəbɪ, əɾəbɪ</i>	‘monkey’
<i>dəke</i>	<i>dəke, dəe</i>	‘dative postposition’
<i>kə-</i>	<i>kə-, ə-</i>	‘3 rd person verb prefix’

Table 3.5. Class *c*: words for which the rule of *k*-dropping is obligatory

FEMALE	MALE	
<i>dɪkarǎ</i>	<i>dɪarǎ</i>	‘I’
<i>ka-</i>	<i>a-</i>	‘1 st person verb prefix’
<i>rakufi</i>	<i>rofi</i>	‘to eat’
<i>beraku</i>	<i>bero</i>	‘river’
<i>hedoku</i>	<i>hedo</i>	‘house’
<i>daka</i>	<i>da</i>	‘to take off’
<i>ruku</i>	<i>ru</i>	‘night’
<i>rukudǎ</i>	<i>rudǎ</i>	‘to thrust’

That these differences do not relate to criteria such as morphological status or part-of-speech is shown by the fact that each of the three subject pronouns behaves in a different way. For the first person pronoun, *k*-dropping is obligatory: ♀ *dɪkarǎ*, ♂ *dɪarǎ*. The second person pronoun, *kai*, belongs to the small class of absolute exceptions. Finally, for the third person pronoun, *dəkɪ*, *k*-dropping is optional. The same assumption is also valid for the postpositions: while the locative postposition =*kɪ* belongs to Class *a*, the dative postposition *dəke* belongs to Class *b*, and the temporal postposition =*ku* to Class *c*. Finally, the same differences in behavior occur with affixes: although *k*-dropping is obligatory with the first person irrealis prefix *ka-*, it is optional with the third person irrealis prefix *kə-*.

Furthermore, the *k*-dropping rule does not apply equally to all the occurrences of the velar stop in a given word. If two adjacent syllables in the same word contain a /k/, the

tendency is that only one of its occurrences will be deleted, generally the first one: ♀ *kɔka* ‘to shred’ (♂ *ɔka*); ♀ *kukɔ* ‘to lift’ (♂ *ukɔ*). It is plausible that at least some of these restrictions are due to phonological constraints, such as to limit the number of vowels in a row. Also, at least some cases of obligatory male-female speech differentiation can be stated in phonological terms. Thus, the velar stop seems to be always deleted when occurring before /u/. Thus, in the male speech form corresponding to ♀ *kowoku* ‘mortar’, *k*-dropping is obligatory in the last syllable, but optional in the first (♂ *kowo*, *owo*).¹⁴ The deletion of the velar stop seems to be generally optional in the cases described in section 2.1.2, which involve schwa assimilation. On the other hand, differentiation is always obligatory in the cases described in section 2.1.1 above, in which vowel fusion takes place (♀ *beraku* ‘river’, ♂ *bero*), and in section 2.1.3, involving epenthesis (♀ *ikɔrɔ* ‘fox’, ♂ *idʒɔrɔ*). Thus, the differentiation between male and female speech seems to be obligatory exactly in those cases in which the differences between the words in both gender dialects are more salient.

Some morphemes seem to be on the borderline between Class *a* and Class *b*. That is the case of the ‘augmentative’ morpheme *-hVkɛ̃*, which generally preserves the velar stop in the male speech, such as in ♀ *berohokɛ̃* [river-big] ‘great river (the Araguaia)’, ♂ *berakuhukɛ̃*. Although this morpheme may occasionally occur without the velar stop in the speech of some individuals, this is very sporadic, and considered wrong by the majority of the speakers. Therefore, although the classification suggested here accounts for the behavior

¹⁴ Therefore, examples such as *kowoku* ‘mortar’ belong simultaneously to classes *b* and *c*.

- b. ♂ *dɪarǝ* *adɪdʌkʌkre*
dɪarǝ a-d-I-Ø-dʌkʌ=kʌre
I 1-CTPT-TRANS-3-take=FUT
‘I will take it off.’
- c. *dʌkɪ* *kʌdɪdʌkʌkre*
dʌkɪ kʌ-d-I-Ø-dʌkʌ=kʌre
he/she 3-CTPT-TRANS-3-tie=FUT
‘She/he will tie it.’

3.5 Early data and borrowing

This analysis of the variation in the use of Class *b* words in male speech is corroborated by the early vocabularies collected by Castelnau (1851), Ehrenreich (1894), Coudreau (1897), and Krause (1911). Given the fact that women traditionally had scarce contact with the outside world, most of these early linguistic data were obtained from male speakers. Thus, while Class *c* words are consistently given in their male speech form (Table 6), Class *b* words tend to occur in their neutral (that is, *k*-preserving) form (Table 7).

Table 3.6. Class *c* male speech forms occurring in early word lists¹⁵

CASTELNAU (1851)	EHRENREICH (1894)	COUDREAU (1897)	KRAUSE (1911)	IPA transcription	
<i>Aeto</i>	<i>heto</i>	<i>Ééto</i>	<i>Häto</i>	[he'do]	‘house’
<i>Bero</i>	<i>bero</i>	<i>bero[oco]</i>	<i>Bäro</i>	[be'ro]	‘river’
<i>Loosi</i>	<i>Roši</i>	<i>Rochi</i>	<i>Dosi</i>	[ro'ʃi]	‘to eat’
<i>Awo</i>	<i>auno</i>	<i>aoun'o</i>	<i>(h)awo</i>	[hã'wɔ], [ã'wɔ]	‘canoe’
<i>Bi-ou</i>	<i>Biu</i>	<i>Biou</i>	<i>Biu</i>	[bi'u]	‘rain’
<i>Roou</i>	<i>Ruu(rere)</i>	<i>rou ou</i>	<i>(d)luu</i>	[ru'u]	‘night’
	<i>mahi</i>	<i>Mai</i>	<i>Mai</i>	[ma'i]	‘corn’
<i>Badero</i>		<i>Bédéro</i>	<i>Bedelo</i>	[bɛde'ro]	‘savanna’
<i>Avoai</i>	<i>an ʋaua</i>	<i>Anolé</i>	<i>anloä,</i> <i>andoä</i>	[hãlɔ'ɛ], [ãlɔ'ɛ]	‘jaguar’

¹⁵ Diacritics, which are especially abundant in Krause’s transcriptions of the Karajá vowels, were suppressed in the tables above.

Table 3.7. Class b words occurring in early word lists

CASTELNAU (1851)	EHRENREICH (1894)	COUDREAU (1897)	KRAUSE (1911)	IPA transcription	
<i>Takeu,</i> <i>tacou</i>	<i>Teke</i>	<i>Tacou</i>	<i>Deke</i>	[dʰəki]	‘skin, cloth, clothes’
<i>Cooté</i>	<i>Koti, kote</i>	<i>Cooti</i>	<i>Koti</i>	[kəʔdɪ]	‘tobacco’
<i>Comota</i>		<i>Comata</i>		[kəməʔdʰa]	‘beans’
<i>Craobi</i>	<i>kraobi</i>		<i>Klaobi</i>	[kərəʔbɪ], [krəʔbɪ]	‘monkey’
<i>Coonri</i>	<i>kaonri</i>		<i>oli, kōli</i>	[kōʔri]	‘tapir’
<i>Mokawa</i>	<i>makaua</i>	<i>Macaoua</i>	<i>Mahaua</i>	[mʰəkaʔwa]	‘firearm’
<i>Kanara</i>	<i>kenara</i>	<i>can oura,</i> <i>cao uara</i>	<i>kanula,</i> <i>kinōla,</i> <i>kenona</i>	[kənəʔra]	‘sand’
<i>matokari</i>	<i>matokare</i>	<i>Matoucari</i>		[madʰəkəri]	‘old man’
<i>Awkeu</i>	<i>hanökö(e)</i>	<i>Anoucou</i>	<i>hau(e)ke</i>	[həwəkɪ]	‘woman’
	<i>katu</i>	<i>Cootou</i>	<i>Kodu</i>	[kəʔdʰu]	‘turtle (sp.)’
	<i>katora,</i> <i>katura</i>	<i>Catoura</i>	<i>Kadola</i>	[kədʰura]	‘fish’
	<i>kašüära</i>	<i>cachi’ouéra</i>	<i>Kašiwera</i>	[kaʃiwɛʔra]	‘pepper’
	<i>kanande</i>	<i>Canandé</i>	<i>kenondä,</i> <i>kenōde</i>	[kənəʔdɛ]	‘(manioc) flour’

The same tendency can be observed with the few loanwords borrowed into Portuguese. These are mainly proper nouns, such as *Xambioá*, the name of a town in Tocantins (♂ *ifəbikəwa* [ifəbikəʔwa] ‘companion people’) and *Tainá*, a female proper noun (♀ *dakidã* [dʰakina] ‘star’, ♂ *dardã*). As expected, Class *b* loanwords tend to be *k*-preserving, such as in *Crixás*, the name of a town in Goiás (*kəriθa* [kriθa] ‘Xavante’, ♂ *əriθa*).

Karajá borrowings to other indigenous languages show the same tendency. Thus, a certain type of tobacco pipe, called *werikəkw* in Karajá (♂ *werikəkw*, *werikəkw*) was borrowed by the Kayapó (a neighboring tribe who was in contact with the Northern Karajá and

Xambioá), who preserved its original name. As Krause (1942[1911]:160) points out, the loanword, although likely taught to the Kayapó by Karajá men, is *k*-preserving: “It is curious to notice that the Kayapó employ the name *walikokó* [for the pipe], which is a term from the women’s speech. [...] The use of the term from the women’s speech can be explained by the fact that the [Karajá] men habitually use it when talking to strangers.” (Krause 1942[1911]:160, my translation). Of course, contrary to what Krause states, *wɛɾɪkəkɔ* is not exclusively a term from the women’s speech, being simply one of the examples for which the differentiation between both speeches is optional.

3.6 Use of female speech by men, and vice-versa

Examples in which the *k*-dropping rule may or may not apply cannot be confused with cases in which a male speaker deliberately uses female speech forms. In fact, as in Koasati (Haas 1964) and Yana (Sapir 1949), female and male speakers can use each other’s speech forms when ‘impersonating’ a character of the opposite sex, such as when narrating a story or baby-sitting. Thus, if a male speaker is, for example, quoting a female character or baby-talking to his daughter, he will use plain female speech in doing so.¹⁶ Inversely, women will use male speech forms when quoting a male character or baby-talking to a baby boy. The use of female speech by a male speaker is further illustrated by the sentence below,

¹⁶ A very interesting question to be investigated is when and how are the distinctions between female and male speech learned by the child. According to Fortune & Fortune (1975:115), “at about age three, mothers start insisting in the Karajá tribe that boys use only men’s speech. This is the beginning of the socialization process which continues until the adult male is admitted to the men’s house where no young initiate would ever use women’s speech.” However, the observance of the female-male speech differences in baby-talk suggests that these distinctions start being acquired much earlier.

heard from a father addressing his baby daughter in baby-talk. Notice that ♀ *dzikarõ* ‘I’, ♀

ka- ‘first person prefix (irrealis)’, and ♀ *rakufi* ‘to eat’ are exclusively female forms.

- | | | | |
|------|-------------------------|----------------|--------------------------------------|
| (18) | <i>wáha,</i> | <i>dzikarõ</i> | <i>karirakufikemõrerri!</i> |
| | waha, | <u>dzikarõ</u> | <u>ka-r-i-rakufi</u> =kε=bõ=r-εrɪ |
| | my.father-VOC | I | 1-CTFG-TRANS-eat=POT=CONV=CTFG-PROGR |
| | ‘Daddy, I want to eat!’ | | |

3.7 Hypercorrection

When one is imitating or quoting a different dialect or a closely-related language, analogical mistakes are common, due to the over-application of phonological correspondence rules.¹⁷ Similar mistakes may happen in Karajá when a female speaker is quoting a male speaker, or vice-versa. Although most evidence I have for such hypercorrection is anecdotal,¹⁸ I managed to find at least one example on print. In a first, experimental edition of the gospel according to St. John, in the famous passage in which Jesus talks to the Samaritan woman by the well (John 4:7-9), a [tʃ] was inserted in the word *Samaria* in the quotation of

¹⁷ For instance, a Portuguese speaker may erroneously assume that the Spanish form corresponding to Portuguese *corte* ‘a cut’ is **cuerte*, by analogy with several cognate pairs displaying regular phonological correspondences (Port. *sorte* :: Sp. *suerte*; Port. *morte* :: Sp. *muerte*; Port. *forte* :: Sp. *fuerte*; etc.).

¹⁸ Although both varieties are mutually intelligible, one must be careful in trying to elicit data from the female speech with a male consultant, or vice-versa. As speakers of different dialects tend to do, traits considered as being ‘typical’ of the other dialect may be exaggerated. While working on a project of lexical compilation, a number of Karajá male teachers were discussing the differences between homophony and polysemy. A pair of homophonous words was then chosen as example: *riu* ‘to hunt’ and *riu* ‘he/she defecated’. When asked whether these words would also be identical in the female speech, they promptly said that both would occur with the velar stop, only to be immediately corrected by the only Karajá female in the room. As it turns out, *riu* ‘to hunt’ and *riku* ‘he/she defecated’ are not homophonous in the female speech.

the Samaritan woman's speech (Fortune & Fortune 1972:14), by analogy with forms such as

♀ *ritfa* (♂ *ria*) 'to walk':¹⁹

7 Tahe **Samaría** Mahādu hāwyy-ō dobehede [...].

'A woman from Samaria came down [to draw water].'

[...]

9 Tahe hāwyy rarybere: kai Judeu Mahādu tate, Jikarÿ **Samaritxa** Mahādu

rare. [...]

'Then the woman said: you are a Jew, I am a Samaritan.'

3.8 Social correlates: *idzoi* vs. *ifš*

As Fortune & Fortune (1975) point out, male vs. female speech distinctions are part of a wider set of social rules concerning gender differentiations, very important in Karajá society. Although both males and females traditionally received, upon puberty, the circular facial tattoos characteristic of the tribe, body painting patterns and ornaments are clearly distinct with relation to gender. Besides common roles traditionally associated with each gender (women as ceramists, men as wood carvers; women as manufacturers of delicately woven baskets, men as manufacturers of rough and sturdy carrying baskets; etc.), social space is clearly delineated geographically as well: the village, seen as the woman's place par excellence in a matrilocal society, is called *ifš hāwa* 'the place of *ifš*', a noun which

¹⁹ The translation work seems to be done in cooperation between native consultants and the missionaries. As far as I know, all the native consultants were males. Considering that the Fortunes' analysis of the use of forms such as ♀ *ritfa* vs. ♂ *ria* is essentially correct (female speech is considered as more conservative, and [tʃ] is shown to be the result of the palatalization of /k/; cf. Fortune & Fortune 1977), a hypercorrection mistake such as this could only have been made by a native speaker. The mistake was duly corrected in the current edition of the New Testament (1983).

designates a category including women and children of both genders--as opposed to ♀ *itfoi* (♂ *idzoi*), generally translated in Portuguese as *a rapaziada* ‘the guys’, which refers to the totality of initiated men.²⁰

The exclusively-male ritual spaces, where the main festivals are planned and performed, are called *idzoi-da* ‘the place of the guys’. The *idzoi-da*--which includes not only the ritual plaza (*kube*) and the house where ritual masks are made and stored (*hetokre*), but also the paths leading to them—is strictly forbidden to women (Toral 1992). Although the women also have their own ritual space, the *hirari-da* (*hirari* ‘young girl’), such space is not exclusively for them, as men are free to wander around. While the male ritual space is actively defined as the one where rituals take place and women are forbidden, the female ritual space is passively defined: it is the place where they witness, mostly as mere expectators, the rituals taking place. As Toral (1992:68) points out, “strictly speaking, the only space that truly belongs to the women is the interior of the houses.” The parallels between such social restrictions and language use are transparent: while male speech is characterized by a higher degree of variation and flexibility (since many k-preserving forms can be optionally used by men), female speech is characterized by a higher degree of rigidity, reflecting women’s general lack of social mobility in Karajá society.

Satirical songs--including the *weru wiu* ‘maraca songs’ (Conrad 1997; cf. Chapter 5), generally used to mock atypical social behaviors--often criticize an individual whose transgressions blur the traditional gender lines within Karajá society, frowning upon, for instance, a man who serves his own food (therefore invading a typically female social space):

²⁰ For the ongoing grammaticalization of *idzoi* into a 1st person pronoun, see Chapter 4.

criticize their exaggerations. An extreme, oft-mentioned example is the radical adaptation of Portuguese *Coca-Cola* into *ɔ̃ ɔa ɔla* [ɔ'a ɔ'la].

As with other markers of social belonging, such as body-painting and ornaments, the use of language as an emblem of masculinity seems to fade as one enters mature age, retiring from the public sphere. A number of well-respected chiefs and shamans (such as the late Watau, from Santa Isabel do Morro, and Maurehi, from Aruanã) are said to have indulged in adopting female speech forms, in their older years, without incurring any censorship from their communities.

The contrast in linguistic behavior between the young men (*weriribɔ*) and the elders (*badokari*) reflects rather fittingly the different stages they occupy in Karajá social life. As Toral (1992:122; my translation) puts it, “at the vigor of their age, ornamented and always ready for ceremonial activities”, the *weriribɔ* are both “the pride” of the village and “the soul” of the *hedokre*, and, despite the recentness of their admittance, “exercise their rights as effective members of the *idzoi* with much more emphasis than the other age groups.” On the other hand, although older men are still revered in matters of ritual concern, they tend to be portrayed, in traditional mythology at least, as little more than domestic fixtures, sitting by the kitchen fire all covered in ashes.²¹

An even more striking parallel between linguistic behavior and other marks of social belonging concerns the use of labrets, a habit now largely abandoned by the Karajá.²² As described by Donahue (1982:132-3; italics added), upon their initiation (“usually around twelve or thirteen”), during the *hed'ohokɔ̃* ‘big house’ festival,

²¹ *radidɔ̃hɔ̃kɔ̃=bɔ̃ rvdɔ̃=bɔ̃hɔ̃* ‘sitting all covered in ashes’ occurs as a semi-formulaic phrase in many traditional tales, describing an old man’s place in the household.

²² According to Donahue (1982:133), “[A]t present the lip plugs are only used for ceremonial purposes—only a few older Karajá still use them in daily life.”

“the boys’ lower lip is pierced and a lip plug is placed in the hole. For the young boys the lip plug is a small stone; as they reach maturity, the plug is longer and longer and is made of wood. As a man reaches his physical prime, the lip plug may reach down to his chest. As he gets older, he uses a smaller and smaller plug, until for old men there is nothing left but a small wood plug that does not even come out of the hole. *The Karajá lip plug seems to be a graphic symbol of the rise and fall of potency.*”

4. Female versus male speech in Javaé

In this section, I will describe the female *versus* male speech distinctions occurring in Javaé as in comparison with the other three dialects. According to Fortune & Fortune (1963), such differences would not occur in Javaé, as both males and females would speak what corresponds to the male speech in the remaining dialects. In fact, that is the opinion of most of the speakers of the Karajá and Xambioá dialects, for whom “Javaé women speak like men.” As I intend to show in this section, although such an affirmation may be sensible from the perspective of the Karajá and Xambioá speakers, it is not completely accurate.

In fact, Javaé preserves the distinctions between female and male speech in a number of words. For example, while a Javaé woman says *hãlɔkɔɛ* ‘jaguar’, a man says *hãlɔɛ*. However, most of the words for which the distinctions are maintained are those for which the deletion of the velar stop /k/ in the male speech is optional (‘Class b’ words). For the great majority of the words for which the *k*-dropping is obligatory in the other dialects, Javaé women adopted the corresponding male form: *dʒiãrɛ̃* ‘I’, *bero* ‘river’, *riɔrɛ* ‘offspring’. That is why, from the viewpoint of Karajá and Xambioá speakers, there is nothing exclusively ‘feminine’ in the speech of the Javaé women. The text fragment below, from a story told by a Javaé woman from the village of Boto Velho, illustrates this point. Notice that *aõ* ‘thing’,

dɪarɔ̃ ‘I’, *a-* ‘first person irrealis prefix’, and *ko* ‘to’ are exclusively male forms in the other three dialects. The only morpheme corresponding to a female form in the other dialects is *likɪ* ‘to tell’, for which the *k*-dropping rule is optional:

(20) *aɔ̃ma* *dɪarɔ̃ araribekere,* *inɔ̃wɛbɔ̃hɔ̃narɪkɪ,*
aɔ̃-bã dɪarɔ̃ a-r-a-rɪbɛ=kere, idɔ̃-wɛ-bɔ̃hɔ̃-dã=rɪkɪ,
 thing-PHAT I 1-CTFG-INTR-speak=FUT people-belly-break-NOM=NARR

inɔ̃rɪkɪ *dai rawebohore.*
 idɔ̃=rɪkɪ dai Ø-r-a-wɛ-bɔ̃hɔ̃=r-e.
 people=NARR 3.LOC 3-CTFG-INTR-belly-break=CTFG-IMPERF
 ‘I will talk about [a place called] Inywebòhòna. There, it is said, the people’s bellies were broken.’

aɔ̃marerɪkɪ *dʒɪrɪrɪkɪ* *dadiko*
aɔ̃-bã=r-e=rɪkɪ dʒɪrɛ=rɪkɪ d-adɪ=ko
 thing-PHAT=CTFG-IMPERF=NARR boy=NARR 3-mother=to

relɪkɪrɛ.
 Ø-r-e-likɪ=r-e.
 3-CTFG-INTR-tell=CTFG-IMPERF
 ‘It is said that a boy told his mother [the secrets of the men’s ceremonial house].’

Table 8 below lists a number of examples from the speech of Javaé females. As mentioned before, the majority of the cases for which *k*-preserving forms were maintained in Javaé involve Class *b* words, such as *dɪkɪ* ‘he, she, it’. For the majority of Class *c* words, however, the Javaé women adopted male forms. This is true for most of the words that present vowel fusion in the male form of the other dialects, such as *rofi* ‘to eat’, *bero* ‘river’, and *da* ‘to take off’. The few exceptions include ♀ *hãlɔ̃kɔ̃ɛ* (♂ *hãlɔ̃ɛ*) ‘jaguar’ and ♀ *ikɔ̃rɔ̃* (♂ *idɔ̃rɔ̃*) ‘fox’, which can be seen as relics from a time when the distinctions between male and female speech in Javaé were still strongly observed.

Table 3.8. Female speech in Javaé as opposed to Karajá

	JAVAÉ ♀			KARAJÁ ♀ ²³			CLAS S
	SPOKEN WOMAN”	“AS	A	SPOKEN MAN”	“AS	A	
1. (s)he	<i>diki</i>						<i>dəki</i> b
2. 1 st person irrealis				<i>a-</i>			<i>ka-</i> c
3. arm				<i>afiə</i>			<i>afikiə</i> c
4. bird	<i>dāwaki</i>						<i>dāwəki</i> b
5. bitter manioc				<i>adziura</i>			<i>adzikura</i> c
6. canoe				<i>hāwə</i>			<i>hāwəkə</i> c
7. club	<i>kəhəde</i>						<i>kəhəde</i> b
8. dative postposition	<i>deke</i>						<i>dəke</i> b
9. doll	<i>rikoko</i>						<i>ritfoko</i> b
10. to eat				<i>rofi</i>			<i>rakufi</i> c
11. fan	<i>kəri</i>						<i>kəri</i> b
12. firearm	<i>bəkawa</i>						<i>bəkawa</i> b
13. flour	<i>kədəde</i>						<i>kədəde</i> b
14. forehead	<i>kəru</i>						<i>kəru</i> b
15. friend				<i>bəwa</i>			<i>bikəwa</i> c
16. grandfather				<i>labie</i>			<i>labike</i> c
17. harbor	<i>bekə</i>						<i>bekə</i> b
18. house				<i>hedə</i>			<i>hedoku</i> c
19. I				<i>dīarə</i>			<i>dīkarə</i> c
20. jatobá (tree sp.)	<i>kīwa</i>						<i>kīwa</i> b
21. knife	<i>baki</i>						<i>baki</i> b
22. little boy				<i>oladu</i>			<i>koladu</i> c
23. maize				<i>mai</i>			<i>maki</i> c
24. male group				<i>idzoi</i>			<i>itfoi</i> c
25. monkey	<i>kəɔbri</i>						<i>kəɔbri</i> b
26. mortar				<i>kowo</i>			<i>kowoku</i> b, c
27. offspring				<i>riɔre</i>			<i>ritfɔre</i> c
28. pepper	<i>kafiwera</i>						<i>kafiwera</i> b
29. pestle	<i>hākə</i>						<i>hākə</i> b???
30. river				<i>berə</i>			<i>bəroku</i> c
31. sand, beach	<i>kədəra</i>						<i>kədəra</i> b
32. skin, cloth	<i>diki</i>						<i>dəki</i> b
33. sky, rain				<i>biu</i>			<i>biku</i> c
34. straw mat				<i>bire</i>			<i>bikire</i> c
35. tapir	<i>kōri</i>						<i>kōri</i> b
36. three				<i>idadaə</i>			<i>idadadə</i> c
37. to				<i>ko</i>			<i>koku</i> c
38. to be born				<i>ua</i>			<i>uka</i> c
39. to break, to cut	<i>kəɔ</i>						<i>kəɔ</i> b

²³ The female speech forms in Xambioá are the same as in Karajá, except for the phonological rules described in section 3 above.

40. to burn	<i>oko</i>		<i>əko</i>	<i>b</i>
41. to dig	<i>ka</i>		<i>ka</i>	<i>b</i>
42. to eat (grains)	<i>kɨ</i>		<i>kɨ</i>	?
43. to fetch		<i>ɔri</i>	<i>kɔri</i>	<i>c</i>
44. to know		<i>eri</i>	<i>keri</i>	<i>c</i>
45. to make holes		<i>ɔa</i>	<i>ɔka</i>	<i>c</i>
46. to run		<i>a</i>	<i>aka</i>	<i>c</i>
47. to scratch, to rub	<i>kɨθe</i>		<i>kɨθe</i>	<i>b</i>
48. to take off		<i>da</i>	<i>daka</i>	<i>c</i>
49. to tell	<i>liki</i>		<i>ləki</i>	<i>b</i>
50. to thrust		<i>rudə</i>	<i>rukudə</i>	<i>c</i>
51. to tie	<i>daka</i>		<i>daka</i>	<i>b</i>
52. to weed		<i>lafikure</i>	<i>lafikure</i>	<i>c</i>
53. tree, wood	<i>kɔwɔɾu</i>		<i>kɔwɔɾu</i>	<i>b</i>
54. trunk, tree	<i>kɔ</i>		<i>kɔ</i>	<i>b</i>
55. turtle	<i>kɔdʊ</i>		<i>kɔdʊ</i>	<i>b</i>
56. woman	<i>hawiki</i>		<i>hawəki</i>	<i>b</i>

As expected, the female speech word for ‘mortar’ in Javaé is *kowo*, a form also available in the male speech (♂ *kowo*, *owo*), not only in Karajá and Xambioá, but in Javaé as well. Therefore, male speech in Javaé presents the same kind of variation involving Class *b* words that is found in the other dialects. This is illustrated by the sentences below, both of which were obtained from a text narrated by a Javaé male from the village of Boto Velho. Likewise what happens in Karajá and Xambioá, the dative postposition =*deke* may appear with (21a) or without (21b) the velar stop. In addition, the pronoun *dɨki* ‘he, she, it’ occurs in its *k*-preserving form:

- (21) a. *dɔriwana* *wideke* *raribereri*
dɔri=wadã wi=deke r-a-ribe=r-eri
White=COM REC=DAT CTFG-INTR-speak=CTFG-PROGR
‘[He] is talking with the White fellow.’
- b. *dɨkiθe* *wadeerenõ* *raribera,* *wad̥auhe*
dɨki=he wa-deε=redõ r-a-ribe=r-a wad̥au=he
(s)he=EMPH 1-DAT=PL CTFG-INTR-speak=CTFG-PERF W.=EMPH
‘He told us himself, Watau.’

Therefore, it is reasonable to suppose that Javaé evolved historically from a situation similar to the one still present in the other dialects to a situation in which most of the characteristics of male speech were adopted by female speakers. Besides its obvious historical-comparative implications, such a historical change carries considerable morphological consequences. For example, the process of *consonantal replacement*, which derives nouns from verbs, is in most cases obscured in Javaé, since the identity between the primitive verb and its derived noun cannot be stated in synchronic terms: *a* ‘to run’ > *ara* ‘the action of running’; *ria* ‘to walk’ > *rira* ‘the action of walking’; *da* ‘to take off’ > *daraθa* ‘the action of taking off’. The following sections are devoted to discussing the likely origins of the male-female speech distinctions in Karajá, as well as suggesting a possible scenario in which the Javaé drift towards male speech may have taken place.

5. Language contact and the ‘Xavánte hypothesis’

The development of male and female speech distinctions such as the ones occurring in Karajá may, in several cases, be the result of situations of language contact.²⁴ The classical example of such a situation occurs in the so-called Island-Carib language (Arawak), mentioned as early as the 17th century. According to Jespersen (1964:237), “[T]he first to mention their distinct sex dialects was the Dominican Breton, who, in his *Dictionnaire*

²⁴ Although phenomena such as pronoun borrowing and male vs. female speech distinctions are generally discussed in reference to ‘exotic’ languages, both can be illustrated by a case of borrowing in Portuguese. In [Brazilian] Portuguese, French *moi* is commonly used informally, mainly in a sort of tongue-in-cheek register. My impression as a native speaker (corroborated by internet searches and consultations with other native speakers) is that such usage of *moi* is essentially restricted to female speech, illustrating how males and females may be more or less receptive to language innovations.

Carai-be-français (1664), says that the Caribbean chief had exterminated all the natives except the women, who had retained part of their ancient language.” The women preserved their Arawakan language, adopted later by the Kariban conquerors, who retained elements of their original language in their speech—giving origin to what came to be called the ‘men’s language.’²⁵ Jespersen views this theory with suspicion, and tries to explain such differences as a result of language taboo. However, in spite of Jespersen’s skepticism, more recent studies have given further support to the traditional explanation. According to Taylor & Hoff, “it is impossible to explain the available linguistic data unless one accepts the native tradition that indeed there has been such an invasion”:

The language that is called the women’s language by Breton, but which actually was and is the native language of both sexes, is a typical Arawakan language [...]. The lexicon is Arawakan too, except that it contains a rather large proportion of borrowed Carib (Karina) lexemes: about twenty-two percent of the basic vocabulary. Over and above these loanwords adopted by the common tongue, the so-called men’s language contained a large number of lexemes and some grammatical morphemes of Carib (Karina) provenance, the latter being insufficient to formulate most simple sentences without admixture of inherited Arawakan grammatical morphemes. [Taylor and Hoff 1980:301]

Another case in which language contact may have played a role in the genesis of male versus female speech distinctions is Kokama, a language spoken in lowland Peru. Kokama is traditionally classified as a Tupí-Guaraní language. However, Cabral (1995) argues that “Kokama cannot even be classified genetically because its linguistic subsystems do not find correlates in a single language or linguistic family source.” Through a comparison with Tupinambá (Tupí-Guaraní), a language assumed to have been the lexifier of Kokama, Cabral shows that “except for lexical similarities and a number of sound correspondences, Kokama differs greatly from the Tupinambá language in its phonology, morphology and syntax,

²⁵ According to Taylor & Hoff (*op. cit.*:302), the Island-Carib male speech is “virtually extinct today.”

though the latter is the source of a significant portion of the vocabulary of the former.” Cabral suggests, as a likely historical scenario for the development of Kokama, the Jesuit missionary villages, where speakers of languages of different families (including Arawakan and Panoan) had to learn Kokama, the official language of the missions, “in this process failing to learn the Tupí-Guaraní language as a whole.” Among the non-Tupian features found in Kokama, Cabral (*op. cit.*:285) mentions the existence of differences between male and female speech in the pronominal system, which “is neither a Tupian feature nor a feature characteristic of the other native languages that we know the Kokama speakers have had some contact with during the last 400 years.” Interestingly, the female pronominal forms are mostly of Tupian origin, whereas the male forms are mostly non-Tupian. The explanation Cabral suggests for the origin of such differentiations lies in the situation of intense language contact:

This gender distinction found in Kokama could be either a feature created in a contact situation where a social distinction based on sex and ethnicity was called for, or it could have developed as an extension of a feature found in the non-Tupian language(s) participating in the development of Kokama. This extension could well have been caused by the contact situation itself, where mixing of Indian groups with significant proportions of males from one ethnicity and females from another could be a favorable situation for the development of speaker sex distinctions in linguistic forms. [Cabral 1995:287-88]

In Island-Carib and Kokama, the role of language contact in the development of male versus female speech distinctions seems to be quite straightforward, since males and females use lexemes of distinct genetic sources. That is a very different situation from what happens in Karajá, where, as we have seen, males and females share the same lexical items (with the exception of a few interjections), the differences between both speeches being simply a result of fairly regular phonological processes. However, Rodrigues (1999) suggests for Karajá a scenario very similar to what may have happened in Island-Carib:

Perhaps in the past the women of one dialectal group of Karajá could have been subjugated by warriors speaking another language, say one lacking velar stops but having glottal stops. Such warriors could have killed all Karajá male adults, taken their place as husbands and learned the Karajá language from their new wives, but substituted their glottal stops for the velar stops of the women. This bad pronunciation by the new masters of the group would then have been maintained through the following generations and spread to other dialectal groups. In the course of time, the articulation of the glottal stop would have weakened and finally disappeared, giving place to vowel sequences and contractions. [...] It happens that the neighboring language Xavánte has historically undergone the systematic change of velar consonants into glottal stops. [...] This language, which does not show a similar difference correlated with the sex of the speakers, may well be the source of the difference between Karajá men's and women's speech. [Rodrigues 1999:177-8]²⁶

This imaginative hypothesis runs into a series of obstacles, the main one being that it presupposes a rather intimate situation of language contact, which would presumably have had a more pervasive impact on the Karajá language—including a number of borrowings, such as observed in Island-Carib and Kokama. However, a lexical comparison between Karajá and Xavánte has of yet not revealed a single clear case of Xavánte loanwords in Karajá (or vice-versa, for that matter).²⁷

A more damaging piece of evidence against the Xavante hypothesis is provided by the earliest vocabularies of both languages, collected in the first half of the 19th century. While the first Karajá vocabulary, collected by Castelnau (1850), already show k-dropping forms, the first Xavante vocabularies, collected by Pohl (1832) and by Castelnau (1850)

²⁶ For a more elaborate version of the 'Xavánte hypothesis', see Rodrigues (2002).

²⁷ I compared an extensive Karajá vocabulary with the Xavánte data published by Hall, McLeod & Mitchell (1987).

demonstrate that the rule transforming *k into a glottal stop had not taken place yet.

Rodrigues (2004) comes up with yet another ad hoc explanation, suggesting that the present-day Xavante do not descend from the ones whose language was documented by Pohl and Castelnau, but from another, closely-related group. Such hypothesis, however, has no support in the history of Xavante migrations, which is fairly well-documented by the early colonial sources on Goiás.

6. Internal factors: a hypothesis on the origins of male v. female speech distinctions

As we have seen, there is no convincing evidence supporting the hypothesis that the female-male speech distinctions in Karajá are a result of external factors, such as suggested by Rodrigues (1999, 2002). Instead, a rather more plausible hypothesis is that such differences were the result of internal factors. It is not uncommon for phonological processes to be treated differently according to the gender of the speaker (Labov 2001). It may very well have been the case that a process of velar weakening took place in the past, affecting a certain area of the lexicon. Women would have been less receptive to this innovation and, with time, velar weakening was reinterpreted as a sign of masculinity (or, on the other hand, the preservation of the velar stop was seen as a sign of femininity). The *k*-dropping rule would have then extended to other areas of the vocabulary, no longer as a purely phonological rule, but as a *socially motivated* phonological rule—that is, a phonological rule with a ‘sociolinguistic twist’.²⁸ This would explain the contrast between Class *c* forms (which would belong to the earlier layer of the lexicon to be affected by the *k*-dropping rule;

²⁸ A similar explanation is suggested by Dunn (2000), for the male-female speech varieties in Tangoa and Chukchi.

notice that this class includes all the words which undergo more ‘radical’ phonological processes, such as vowel fusion) and Class *b* words (which would constitute more recent innovations, resulting from analogy; notice that this class includes loanwords, for which *k*-dropping seems to be generally optional).

This scenario seems to be corroborated by comparative evidence. When one compares the consonantal inventory of Karajá with that of Proto-Jê (Table 9 below), an eye-catching difference is the gap, in Karajá, in the inventory of voiceless stops: as we have seen, Karajá lacks both a /p/ and a /t/. Comparisons with different families within Macro-Jê (Davis 1968, Ribeiro & van der Voort 2010) attest to the conservative nature of Proto-Jê, especially concerning its inventory of voiceless stops, which was very likely inherited as such from Proto-Macro-Jê. Lexical comparisons (Davis 1968, Ribeiro 2005), indicate that Karajá underwent a systematic process of voiceless-stop lenition. Before /r/, the process of lenition would have resulted in the complete deletion of the consonant (21); before a vowel, *p and *t would have become the approximants /w/ (Proto-Jê **par* ‘foot’ :: Karajá *wa*) and /r/, respectively. Curiously enough, the only voiceless stop in Karajá, /k/, happens to be the one which is often deleted in male speech—hardly a coincidence, given the apparent overall tendency towards voiceless stop lenition just described. If a similar process of lenition had affected *k before vowels to the same extent in which it affected both *p and *t, a likely reflex would have been the velar approximant *ɥ, a rare and diachronically unstable phoneme which could have been further weakened, giving rise the alternation between /k/ and zero found in Karajá today. Thus, there seems to be rather compelling (albeit circumstantial) evidence for an internal (rather than contact-induced) origin for the

differentiation between male and female speech in Karajá, although the lack of comparative evidence may never allow for a conclusive explanation.

Table 3.9. Proto-Jê consonantal inventory (Ribeiro 2005)

p	t		k
m	n	j	ŋ
	s		
w	r		

(22)	Proto-Jê	Karajá	
a.	* <i>krã</i>	<i>ra</i>	‘head’
b.	* <i>kra</i> ‘offspring’	<i>ra</i>	‘nephew’
b.	* <i>prĩ</i>	<i>rĩ</i>	‘path’
c.	* <i>prãm</i>	<i>rəbã</i>	‘hunger’

The process of k-dropping in Karajá would seem at first to make a strong case for lexical diffusion, rather than for an exceptionless, regular phonological process (Labov 1994: 421-439), since words presenting apparently the same phonological environments undergo the rule to different degrees of obligatoriness. However, it could very well be the case that, when such phonological process first originated, such words did not present the same ‘structural description’ required for the process to take place.²⁹ Given the lack of comparative Macro-Jê data so far, any defense of one or the other position would be highly speculative. The investigation of this and any other hypothesis on the origins of the female-male speech distinctions in Karajá depends on a better knowledge of the languages of the Macro-Jê stock. However, even the little comparative evidence available strongly suggests a

²⁹ That is, k-dropping might have been at first a regular phonological process. In a possible scenario, there would have been two different velar consonants—say, *k and *k^h. Initially, only *k would undergo lenition (or deletion), which would be a regular process in the male speech. In a later stage, the contrast between *k and *k^h would have disappeared; by analogy, reflexes of old *k^h would then also undergo k-dropping. It could be that even the cases of n-dropping would trace back to such a scenario: the consonant [n] would be, in the cases in which it is dropped in the male speech, a reflex of *k in nasal environments (**akō* > **aŋō* > ♀ *adō*, ♂ *aō*).

historical drift towards voiceless stop lenition in Karajá which could be at the root of the k-dropping rule which characterizes male speech.

At any rate, although the regularity of phonological processes remains a major tenet of historical linguistics, there are well-documented cases in which semantic or functional factors may interfere with such regularity. A rather illustrative example is provided by Goddard (1991:60), in discussing Cheyenne vis-à-vis Proto-Algonquian:

“Cheyenne [...] shows one intriguing development that is not yet explained: *p and *k are both dropped, but only, it would seem, optionally; where doublets can be compared, the form with retention has a diminutive meaning beside the form with loss [...]. Such doublets appear to point to the **negative semantic conditioning of a sound law**, surely not a very common type of sound change.”

Likewise, the preservation of /k/ in Karajá female speech could be seen as the result of a **negative sociolinguistic conditioning of a sound law**—again, “not a very common type of sound change”. The synchronic variations in the degree to which k-deletion takes place, coupled with its productive application even to recent loanwords, are in themselves very enlightening, demonstrating how a phonological rule gains a life of its own once it is invested with a sociolinguistic function.

7. Language contact in Javaé

If, as it seems, there was indeed a drift towards the male speech in the Javaé dialect (with women adopting male speech forms), this could have possibly been triggered by

external factors.³⁰ The Javaé data suggest that, at a certain point in the history of this dialect, the male speech was reanalyzed as the basic one. A similar historical change would also have happened in Koasati (as well as in other languages of the eastern branch of the Muskogean family). In Koasati, according to Haas (1964: 4), “at the present time only middle-aged and elderly women use the women’s forms, while younger women are now using the forms characteristic of men’s speech.”³¹ However, different from what happened in Koasati, the differentiation between male and female speech is still productive in Javaé, applying even to loanwords, such as *kabɪɖaõ* ‘chief’, ♂ *abɪɖaõ* (from Portuguese *capitão* ‘captain’). That is, the phonological rule of *k*-dropping that differentiates male and female speech is synchronically the same in Javaé as it is in the remaining dialects.

A possible scenario in which the drift towards male speech in Javaé took place may have involved a great number of foreign women having to learn the Karajá language from their Karajá husbands. The oral tradition of the Karajá-speaking groups, as well as reports by early travelers, mention a custom which seems to have been quite common in the past of this people: the capture of women from other tribes. Their favorite victims were certainly the peaceful Tapirapé (Tupí-Guaraní family), their closest neighbors. As Lipkind (1948:188) points out, the only captives taken by the Karajá were women and small children, “treated as full members of the group.” A number of Karajá families are said to be descendants of Tapirapé women captured by the Karajá (Toral 1992: 5-6).

³⁰ The change was certainly not motivated by the contact with the Brazilian population, which in the case of the Javaé is much more recent than for the other Karajá-speaking groups. While Karajá and Xambioá speakers are in contact with Portuguese-speaking colonizers since mid-1700s, the Javaé only came to establish permanent contact in the 1900s.

³¹ Haas’s findings have been recently challenged by Kimball (1987; see also Saville-Troike 1988 and Kimball 1990).

The tradition of marrying Tapirapé women may have been even more common among the ancestors of the present-day Javaé, since, according to oral tradition, the Tapirapé lived for a time in the interior of the Bananal Island with the Javaé (Wagley & Galvão 1948:167; Lipkind 1948:180; Baldus 1970:35-38; Toral 1992). The contact between the Javaé and Tapirapé apparently had a lasting effect on both tribes. There are a number of Karajá cultural elements imported by the Tapirapé, probably from their former Javaé neighbors. These include ceremonial songs and masks, as well as items of material culture. In addition, there are a number of cultural traits that the Javaé share with the Tapirapé, but not with the other Karajá-speaking groups. Lima Filho (1994:17), for instance, reports that the ceremonial men's house of the Javaé "is totally different from the one of the Karajá", being "apparently very similar in shape to the men's house of the Tapirapé."

On the other hand, Lipkind, in a letter written to Herbert Baldus, points out the existence of physical resemblances between the Javaé and the Tapirapé: "You may be interested to know that in the innermost Javahe villages there are some distinctly Tapirape types, short stature, light coloring, face cast" (Baldus 1970, 1936). According to Baldus (*op. cit.* 38), "the existence of 'Tapirapé types' among the Javahé may be also explained by the great number of Tapirapé women 'imported' by the Karajá." These physical characteristics are pointed out by the other Karajá-speaking groups as a distinctive feature of the Javaé, whom they call *ifãdzu*—the same term used to refer to the other, non-Karajá speaking indigenous groups. The intimate contact between the ancestors of the present-day Javaé and Tapirapé is recorded in the oral traditions of both tribes (for a published version of the story in both Karajá and Portuguese, in which the mixed Karajá-Tapirapé origin of the Javaé is suggested, see Silva & Rocha (org., 2006:139-145).

As we have seen, there are a number of Tupí-Guaraní loanwords in Karajá, including items such as *kobãda* ‘beans’, *bãkawa* ‘firearm’ and *brure* ‘hoe’, borrowed from Língua Geral, *kobãdãwira* ‘andu (a type of pea)’, *hãrara* ‘blue macaw’ and *dãrawe* ‘curica (a small type of macaw)’, from Tapirapé, as well as a number of words from unidentified Tupí-Guaraní sources, such as *hãkuri* ‘agouti’ e *-uwaθa* ‘poisoned arrow’. Such loanwords, common to all four dialects, provide interesting clues on the nature of past contacts between the Karajá and Tupí-Guaraní-speaking groups (Ribeiro 2001/2002). The nature of such loanwords suggests that these contacts were mainly of a commercial nature, since most of them refer to utensils, domesticated plants, and semi-domesticated birds, items traditionally exchanged among Brazilian indigenous groups. There are, however, some loanwords, traditionally associated with the Javaé dialect, which suggest a more intimate cultural contact between the latter and a Tupí-Guaraní-speaking tribe. These include terms such as *hidĩ* ‘garbage’ and proper nouns such as *Kujamõkõ* [kudzamõ’kõ], a female personal name.

Therefore, a reasonable amount of evidence suggests that a situation of intense contact with another tribe (presumably, the Tapirapé) may have contributed to set the Javaé apart from the other Karajá groups—physically, culturally, and linguistically. Since foreign women would have learned the Karajá language from their husbands (and not the traditional way, during childhood, in a female-dominated household), they would have learnt it imperfectly, without the sociolinguistic details which are so important in Karajá-speaking society. Incidentally, since marriages between Karajá men and non-Karajá women are common, similar cases of ‘imperfect learning’ of Karajá by foreign women are not unheard of even today; I have heard several anecdotal reports of non-Karajá women learning to speak ‘like a man’.

8. Concluding remarks

The main purpose of this chapter was to describe the female *versus* male speech distinctions in Karajá and to outline some of the questions such a phenomenon may pose to historical linguistics and sociolinguistics, as well as the implications of such a phenomenon for the analysis of Karajá phonology and morphology. Although male *versus* female speech distinctions are commonly found in the languages of the world, very little is known on how such distinctions are brought about historically, how they correlate with other social institutions, and how and why they cease to exist. In this sense, the example of the Karajá dialects is particularly interesting, for it provides both a situation in which the distinctions are vigorously preserved and a case in which they seem to have faded away diachronically.

The phenomenon of gender dialects in Karajá illustrates a *sui generis* case of language change, whose social and historical motivations constitute a rich theme to be further explored in future studies. Rather than providing the final word on such a complex and fascinating topic, I hope the main merit of this chapter is in revealing additional lines of inquiry. The data suggest that the genesis of such distinctions may be related to an overall tendency towards voiceless stop lenition, a hypothesis which needs to be further tested as our comparative knowledge of Macro-Jê improves. There seems to be a strong correlation between age and the extent to which genderlectal distinctions are carried out by individual speakers, which begs for additional investigations into how and when such distinctions are acquired. The possibility of imperfect learning of the peculiarities of each speech within mixed families—a factor which, as I suggest, may have been at the heart of the weakening of

genderlectal distinctions in Javaé—also deserves further investigation, beyond the anecdotal (albeit compelling) reports by native speakers. An in-depth investigation into these issues must be preceded by an accurate description of the data, and it is my hope that this chapter may represent a substantial contribution towards that goal.

CHAPTER 4

Morphology

1. Introduction

This chapter describes Karajá's word-formation strategies, including, in addition to morphology proper (that is, the combination of roots and affixes), the use of particles and clitics. Most of the chapter will be devoted to verb morphology, which is where most of Karajá's grammatical complexity lies: the verb inflects for person (with portmanteau morphemes which also indicate mood), direction (which also play evidential functions), and voice (transitive, passive, and antipassive), besides incorporating object pronouns and nouns (sometimes with classificatory purposes). Tense and aspect are conveyed by clitics and particles, respectively. As in many other lowland South American languages, nouns, postpositions, and (to a lesser extent) verbs share person-marking paradigms. Since postpositions have already been described in Chapter 1, and their semantic and syntactic characteristics will be further discussed in Chapter 5, they will not be dealt with in this chapter. Section 2 describes the inflectional properties of verbs. Section 3 deals with the inflectional properties of nouns and briefly discusses the small class of personal pronouns, outlining their main differences in relation to nouns. Section 4 discusses derivational morphology, describing processes used to create new noun (Section 4.1) and verb (4.2) stems.

Karajá, like most (if not all) Macro-Jê languages, lacks adjectives as an independent part of speech. In languages lacking adjectives, their functions tend to be expressed by nouns or verbs. The tendency, in Macro-Jê studies, has been to describe 'descriptives' (i.e., words expressing adjectival functions) as verbs, an interpretation which has also been proposed for

Karajá (Fortune 1973, Maia 1986). Section 5 describes the morphological and syntactic properties of such descriptive words, demonstrating that, contrary to previous claims, they are indeed nouns.

2. Verb morphology

Contrasting with a fairly simple noun morphology, Karajá presents a rather complex verb morphology. Stem-formation processes such as compounding, noun incorporation, and reduplication are commonly used. In addition, the verb inflects for *person* (and, cumulatively, *mood*), *direction*, and *voice/valence*. Pronominal direct objects are obligatorily incorporated into the verb. The example below illustrates the distribution of morphemes within the verb word:

- (1) ♀ *kai* *bəɖɪwarakokre*
 kai bə-d-I-wa-rakɔ=kəre
 you 2-CTPT-TRANS-1-wait.for=FUT
 ‘You will wait for me.’

The following sections provide a detailed description of the morphological categories which are expressed by the verb in Karajá: person (Section 2.1), direction (Section 2.2), voice and valence (Section 2.3), and object (Section 2.4). Section 2.5 describes noun incorporation, a process which has interesting implications for the understanding of argument structure. Section 2.6 describes the verbalization of noun-noun compounds, which share some surface properties with noun incorporation.

2.1 Subject agreement

Person agreement displays a strictly nominative pattern, with the verb always agreeing with the subject, be it intransitive (2a) or transitive (2b). Person agreement markers are distributed into two different sets, one occurring in the realis (present and past tenses) and the other in the irrealis (future, potential, and admonitory). These prefixes are listed in Table 1 below.³²

Table 4.1. Subject agreement markers in Karajá

Person	Realis	Irrealis
1 st	<i>a-</i>	♀ <i>ka-lδ a-</i>
2 nd	<i>da-</i>	<i>bə-lb-</i>
3 rd	<i>Ø-</i>	<i>Ø-; kə-lδ ə</i>

- (2) ♀ a. *kraritfakre*
ka-r-a-rika=kəre
I-CTFG-INTR-walk=FUT
 ‘I will walk.’
- b. *kariθuhokre*
ka-r-I-Ø-θuhə=kəre
I-CTFG-TRANS-3-wash=FUT
 ‘I will wash it.’

2.2 Direction

An interesting difference between Karajá on the one hand, and languages such as English or Portuguese, on the other, is that in Karajá there are no lexical opposites for direction, such as English *come* versus *go* and *bring* versus *take*. In Karajá, such a distinction is completely dependent on the morphological mechanisms that are made available by the language. Thus, in the examples below, the prefix *r-* indicates that the event is seen as occurring away from the current location of the speaker (centrifugal direction),

³² The same set of prefixes is used for singular and plural. There is also a distinction between a first person plural exclusive (marked by the same set of prefixes used for first person singular) and a first person plural inclusive (inflected for third person). The prefix *kə-* ‘3rd person’ is restricted to the centripetal direction of the irrealis mood.

whereas the prefix *d-* indicates that the event is seen as occurring towards the current location of the speaker (centripetal direction):

- | | | | | | |
|-----|---|----|------------------------|----|------------------------|
| (3) | ♀ | a. | <i>krakre</i> | b. | <i>kanakre</i> |
| | | | ka- <i>r</i> -∅-a=køre | | ka- <i>d</i> -∅-a=køre |
| | | | 1-CTFG-INTR-move=FUT | | 1-CTPT-INTR-move=FUT |
| | | | ‘I will go.’ | | ‘I will come.’ |

The system of directional markers in Karajá presents characteristics that traditionally define an inflectional category, such as obligatoriness, semantic and formal regularity, and productivity (Anderson 1985:163; Bauer 1988:73-87; Bybee 1985:11). The existence of direction as an inflectional category seems to be a fairly rare phenomenon. Talmy (1985:135), for example, in a survey of grammatical categories, states categorically that direction “is not marked inflectionally.” According to Bybee, this would be due to its lack of *lexical generality*:

By definition, an inflectional category must be applicable to all stems of the appropriate semantic and syntactic category and must obligatorily occur in the appropriate syntactic context. In order for a morphological process to be so general, it must have only minimal semantic content. If a semantic element has high content, i.e. is very specific, it simply will not be applicable to a large number of stems. [Bybee 1985:16-17]

That is, since the category of direction would in principle be relevant only for motion verbs, it would not apply to a sufficient number of verbs to constitute an inflectional category. Thus, besides being commonly expressed lexically such as in Portuguese *levar* ‘to take away’ versus *trazer* ‘to bring’ and English *come* versus *go*, and by clitics, direction is also found expressed derivationally, such as in Latin *eō* ‘I go’, *ex-eō* ‘I go out’, *trans-eō* ‘I go across.’ However, as Bybee (*op. cit.*, 17) observes, “each of these prefixes has a limited

lexical applicability, for they are only appropriately added to verbs indicating motion of some sort. Their semantic content prevents them from meeting one of the criteria for inflectional status.”

Bybee’s predictions apply well to languages such as Georgian, for example, where the opposition between the particles *mo* ‘hither’ and *mi* ‘thither’ seems to be productive only with lexemes denoting some sort of “notional direction towards a goal” (Manning 1996:250).³³ However, in Karajá all verbs inflect for direction, including those that apparently do not indicate a motion at all, such as *uru* ‘to die’ (2) and *ekiwəθe* ‘to get tired’ (3), as shown by the examples below:

- | | | | | |
|-----|----|---|----|--|
| (4) | a. | <i>rurure</i>
Ø-r-Ø-uru=r-e
3-CTFG-INTR-die=CTFG-IMP
‘He died (thither).’ | b. | <i>durude</i>
Ø-d-Ø-uru=d-e
3-CTPT-die=CTPT-IMP
‘He died (hither).’ |
| (5) | a. | <i>rekɨwəθere</i>
Ø-r-Ø-ekiwəθe=r-e
3-CTFG-INTR-get.tired=CTFG-IMP
‘He got tired (thither).’ | b. | <i>dekɨwəθede</i>
Ø-d-Ø-ekiwəθe=d-e
3-CTPT-INTR-get.tired=CTPT-IMP
‘He got tired (hither).’ |

This suggests that, besides its basic, strictly directional use, directional inflection is also used for other purposes. Analyzing the use of the directional markers in narrative texts, this section explores the hypothesis that directional inflection can be used to show empathy relationships between the participants of the speech act and between narrator and characters

³³ The grammatical category expressed by the opposition between the particles *mo* and *mi* is called ‘orientation’ in Georgian linguistics, whereas the term ‘direction’ is used to refer to another grammatical category, which “consists of a multi-term system of preverbs, arranged in opposed doublets such as ‘in’ versus ‘out,’ ‘up’ versus ‘down,’ etc.” (Manning, *op. cit.*, 250) The use of the term ‘direction’ in this paper coincides with Talmy’s (*op. cit.*, 135) definition: “[Direction] refers to whether the Figure in a Motion event is moving toward or away from the speaker.”

in a narrative text, playing a role similar to what in other languages may be performed by obviation systems, evidentials, or attitude markers.

Centrifugal direction ('thither') is clearly the unmarked member of the centrifugal/centripetal opposition. Not only it can, under certain circumstances, be marked by a zero prefix, but it is also overwhelmingly more common in narrative and descriptive texts and everyday verbal interactions. Descriptive predicates, marked with auxiliary clitics which present directional markers, are always marked for centrifugal direction. Furthermore, verbs in the imperfective (marked by the clitic *=r-a*) *cannot* inflect for centripetal direction (an exception which remains to be explained) and there seem to be pragmatic restrictions as to the occurrence of centripetal marking with verbs in the progressive (marked with the auxiliary clitic *=r-err*). As we will see (Section 2.2.2), such restrictions demonstrate that direction can clearly be used with evidential purposes.

2.2.1 Distribution of directional marking

As mentioned above, all verbs in Karajá inflect for direction (*centrifugal* or *centripetal*, depending on the speaker's viewpoint). Centripetal direction, marked by the prefix *d-* (realized as [n] before nasals and /a/), indicates that the process occurs towards the speaker. Centrifugal direction, marked by the prefix *r-* (or by its zero-allomorph), indicates that the process occurs in the direction away from the speaker. Thus, both uses of the verb *lo* 'to enter' in (6) below convey basically the same meaning, the difference being in the location of the deictic center, the speaker. In the first, marked for centrifugal direction, the

speaker is out of the house; in the second, marked for centripetal direction, the speaker is in the house:

- (6) a. *malɔke* b. *mənalɔke*
 b-∅-a-lɔ=ke bə-d-a-lɔ=ke
 2-CTFG-INTR-enter=POT 2-CTPT-INTR-enter=POT
 ‘Enter!’ ‘Enter!’

The distribution of the directional prefixes is illustrated below with the complete paradigm for the verb *wí* ‘to carry’, both in the realis (7) and the irrealis (8) mood.³⁴ As mentioned above, centrifugal direction is marked, in the 1st and 2nd persons of the realis, by a zero prefix, allowing for the fusion between the person-agreement marker and the voice/valence marker—or, with vowel-initial stems such as *-obi* ‘to see’ and *-eburɛ* ‘to get angry’, with the initial vowel of the stem: when the initial vowel of the stem is a back vowel, the personal prefix prevails (9); when the initial vowel of stem is a front vowel, the personal prefix is deleted (10). As the paradigms below illustrate, morphological segmentation is much more straightforward in the centripetal direction, particularly in the irrealis mood.

³⁴ The double marking of direction in the 1st person centripetal of the realis (7b) is restricted to the Southern Karajá dialect. In Javaé, Xambioá, and Northern Karajá, direction is marked only once in such cases:

- a. **Javaé, Xambioá, Northern Karajá**
adiwide
 a-d-I-∅-wi=d-e
 1-CTPT-TRANS-3-carry=CTPT-IMP
 ‘I brought it.’

The interlinear gloss provided for the 1st person centrifugal (7a) is somewhat abstract. It is reconstructed internally, taking the 1st person centripetal as the model. Forms such as (7a) occur in Southern Karajá, Northern Karajá, and Xambioá. In Javaé, on the other hand, 1st person centrifugal forms completely parallel the centripetal form, as can be seen in the examples below. In this sense, the Javaé dialect presents a more regular paradigm, and may reflect more closely what may have existed in Proto-Karajá (or, alternatively, may have undergone analogical regularization):

- b. **Javaé**
ariwire
 a-r-I-∅-wi=r-e
 1-CTFG-TRANS-3-carry=CTFG-IMP
 ‘I took it.’
- c. *adiwide*
 a-d-I-∅-wi=d-e
 1-CTFG-TRANS-3-carry=CTFG-IMP
 ‘I brought it.’

- (7) **realis**
- a. *rewjire*
r-a-∅-I-∅-wi=r-e
CTFG-1-CTFG-TRNS-3-carry=CTFG-IMPCTPT-1-CTPT-TRANS-3-carry=CTPT-IMP
 ‘I took it away.’
- b. *nadiwjde*
d-a-d-I-∅-wi=d-e
CTFG-1-CTFG-TRNS-3-carry=CTFG-IMPCTPT-1-CTPT-TRANS-3-carry=CTPT-IMP
 ‘I brought it.’
- c. *dewjde*
d-a-∅-I-∅-wi=d-e
2-CTFG-TRANS-3-carry=2-IMP
 ‘You took it away.’
- d. *dadiwjde*
d-a-d-I-∅-wi=d-e
2-CTPT-TRANS-3-carry=2-IMP
 ‘You brought it.’
- e. *riwjire*
∅-r-I-∅-wi=r-e
3-CTFG-TRANS-3-carry=CTFG-IMP
 ‘S/he took it away.’
- f. *diwjde*
∅-d-I-∅-wi=d-e
3-CTPT-TRANS-3-carry=CTPT-IMP
 ‘S/he brought it.’
- (8) **irrealis**
- a. *kariwjkre*
ka-r-I-∅-wi=kəre
1-CTFG-TRANS-3-carry=FUT
 ‘I will take it away.’
- b. *kadiwjkre*
ka-d-I-∅-wi=kəre
1-CTPT-TRANS-3-carry=FUT
 ‘I will bring it.’
- c. *biwjkre*
b-∅-I-∅-wi=kəre
2-CTFG-TRANS-3-carry=FUT
 ‘You will take it away.’
- d. *bədiwjkre*
bə-d-I-∅-wi=kəre
2-CTPT-TRANS-3-carry=FUT
 ‘You will bring it.’
- e. *riwjkre*
∅-r-I-∅-wi=kəre
3-CTFG-TRANS-3-carry=FUT
 ‘S/he took it away.’
- f. *kədiwjkre*
kə-d-I-∅-wi=kəre
3-CTPT-TRANS-3-carry=FUT
 ‘S/he will bring it.’
- (9) a. *rabire*
r-a-∅-∅-obi=r-e
CTFG-1-CTFG-INTR-see=CTFG-IMP
 ‘I saw (it) (thither).’
- b. *nadobide*
d-a-d-∅-obi=d-e
CTPT-1-CTPT-INTR-see=CTPT-IMP
 ‘I saw (it) (hither).’
- (10) a. *reburere*

r-a-Ø-Ø-εbure=r-e
CTFG-1-CTFG-INTR-get.angry=CTFG-IMP
 ‘I got angry (thither).’

- b. *nadeburede*
d-a-d-Ø-εbure=d-e
CTPT-1-CTPT-INTR-see=CTPT-IMP
 ‘I got angry (hither).’

As the examples above show, the clitic auxiliary *=(r-)e* ‘perfective’ agrees in person (when in the 2nd person) or direction (when in the 1st and 3rd persons) with the main verb, the same happening to the auxiliaries *=(r-)eri* ‘progressive’ and *=(r-)a* ‘perfective’. The pluralizer morpheme *(r-)edā* also agrees in person and/or direction with the main verb:

- (11) a. *biwibenākre*
b-Ø-I-Ø-wi=b-Ø-εdā=kāre
2-CTFG-TRANS-3-carry=2-CTFG-PL=FUT
 ‘You (plural) will take it away.’
- b. *bədiwibədenākre*
bə-d-I-Ø-wi=bə-d-εdā=kāre
2-CTPT-TRANS-3-carry=2-CTPT-PL=FUT
 ‘You (plural) will bring it.’

2.2.2 Empathy and discourse strategies

This section briefly analyzes the use of directional inflection in narrative texts, as well as in everyday verbal interactions. As mentioned above, the hypothesis suggested by the data is that, besides its ‘literal,’ strictly directional use, directional inflection may be used to signal empathy relationships between participants of the speech act and, in a narrative text, between characters and narrator. The rationale that underlies this hypothesis is straightforward. In any deictic system, the speaker is canonically the deictic center. Directionals can then be

used to establish relations of (*physical, emotional, ideological, etc.*) approximation or distancing from such a deictic center.

The use of directional inflection to signal empathy relationships between participants of the speech act is particularly clear with non-motion verbs, for which both centrifugal and centripetal marking options are available without great semantic constraints. In such cases, the use of a centripetal marker seems to imply that the speaker is somewhat more involved with the process described by the verb. That is why centripetal marking seems to be preferred in situations where advice is being given, such as in the constructions (12) and (13) below, marked by the admonitory particle =*hedǎ*. On the other hand, the use of centrifugal marking in such constructions would fail to convey the concern of the speaker with the fate of the addressee.

- (12) *kǎwǎrukǎ makǎwǎnǎkekǎ bǎdeθehenǎ*
 kǎwǎrukǎ=kǎ b-Ø-a-kǎwǎdǎ=kǎkǎ bǎ-d-ε-θε=*hedǎ*
 tree=AL 2-CTFG-INTR-climb=COND 2-CTPT-INTR-move.down=ADM
 ‘If you climb the tree, you may fall.’

- (13) *bǎdurǎhenǎ!*
 bǎ-d-Ø-urǎ=*hedǎ*!
 2-CTPT-INTR-die=ADM
 ‘[Be careful,] you may die!’

Furthermore, in imperative constructions, marked by the potential particle =*ke*, the use of centripetal markers seems to convey a higher degree of camaraderie (14b, 15b), while the use of centrifugal marking would suggest a more distant relationship (14a, 15a).

- (14) a. *idzǎki wakoku bǎlǎkǎkǎ*
 idzǎki wa-koku b-Ø-ε-lǎkǎ=*ke*
 story 1-to 2-CTFG-INTR-tell=POT

‘Tell me a story.’

- b. *idzəki wakoku bədeləkike*
idzəki wa-koku bə-d-ε-ləki=kε
story 1-to 2-CTPT-INTR-tell=POT
‘Tell me a story.’
- (15) a. *kakə bunəke*
ka=kə b-∅-∅-udə=kε
this=AL 2-CTFG-INTR-sit.down=POT
‘Sit down here.’
- b. *kakə bəduṇəke*
ka=kə bə-d-∅-udə=kε
this=AL 2-CTPT-INTR-sit.down=POT
‘Sit down here.’

In narrative texts, especially those narrated mostly in 3rd person, directional inflection is frequently used to signal which character the speaker chooses to be more relevant for the story, by assigning to him or her the role of deictic center. This function of the directional markers in Karajá is similar to the role played by obviation systems, such as the one occurring in Algonquian languages. In these languages, according to Dahlstrom (1999:36), “if more than one third person is mentioned within a certain syntactic domain, then the third person most central to the discourse is referred to by *proximate* forms (of nouns, pronouns, or verb agreement) and the more peripheral third persons are referred to by *obviative* forms.” The factors involved on assigning proximate or obviative status are mostly semantic or discursive. Thus, “if one of the two third persons is a human and the other is (notionally) inanimate, the human will always be proximate and the inanimate third person will always be obviative.” When both third persons are humans, “the third person chosen as proximate is often the one the speaker feels closest to, such as a relative of the speaker as opposed to a nonrelative, a Mesquakie as opposed to an Indian of another tribe, or an Indian as opposed to

a non-Indian.” However, Dahlstrom remarks that these are tendencies, “not hard and fast rules” (Dahlstrom *op. cit.*, 44-45).

In Karajá, quite interesting is the fact that, in choosing the character to whom to assign higher discourse prominence, objective factors such as physical closeness to the speaker can be overcome by factors such as the place where crucial actions are taking place. Thus, in the short text presented here, ‘The Hawk and the Snake’ (16), the hawk is initially chosen as the deictic center, as shown by the opposition between the verb forms *robi=re* ‘he/she saw (thither),’ in Line D, and *dobi=de* ‘he/she saw (hither),’ in Line E. Notice that at this point the hawk is flying in the sky. The choice of ‘sky’ as opposed to ‘ground’ is not exactly what one would expect if a purely anthropocentric perspective were adopted.

(16)

THE HAWK AND THE SNAKE:
A KARAJÁ TEXT³⁵

- A. *nawiihikã* *heka* *hemãlalawana*
 dawiihikã *heka* *hebãlala=wãda*
 hawk DESCR snake=COM
- wimã* *radãnmãhãre.*
 wi=bã Ø-r-a-dãdã=bãhã=r-e
 both=LOC 3-CTFG-INT-fight=CONT=CTFG-IMP
 ‘The hawk and the snake usually fight with each other.’
- B. *ðahe* *nawiihikã* *ronãmhãhã*
 ða=he *dawiihikã* Ø-r-Ø-udã=bãhã
 and=EMPH hawk 3-CTFG-INTR-sit.down=CONT
- õwõrudĩreki* *rãma* *rĩthamã.*
 õwõru dĩre=ki rãba Ø-r-i-tha=bã
 tree on=LOC hunger 3-CTFG-TRANS-hurt=CONV
 ‘Then the hawk sat down on a tree, hungry.’
- C. *ðahe* *ruore,* *nawiihikã,*

³⁵ This text was collected from a Southern Karajá male speaker from the village of Hawalò, in June of 1993.

- d̥a=he* *∅-r-∅-uɔ=r-e* *dawiihik̄*
 and=EMPH 3-CTFG-INTR-fly=CTFG-IMP hawk
- rɔɔm̄* *rat̄jim̄* *raraθem̄.*
∅-r-∅-uɔ=b̄ *∅-r-a-t̄ji=b̄* *rara* *uθe=b̄*
 3-CTFG-INTR-fly=CONV 3-CTFG-INTR-be=CONV vulture like=LOC
 ‘Then the hawk flew, and kept flying like a vulture.’
- D. *d̥arək̄i* *rɔɔm̄* *rat̄jim̄h̄*
d̥a=rək̄i *∅-r-∅-uɔ=b̄* *∅-r-a-t̄ji=b̄h̄*
 and=NAR 3-CTFG-INTR-fly=CONV 3-CTFG-INTR-be=CONT
- d̥ahe* *hem̄lalam̄* *robire.*
d̥a=he *heb̄lala=b̄* *∅-r-∅-obi=r-e*
 and=EMPH snake=LOC 3-CTFG-INTR-see=CTFG-IMP
 ‘And, while flying, he saw the snake.’
- E. *d̥ahe* *nawiihik̄* *irabim̄* *ream̄h̄,*
d̥a=he *dawiihik̄* *i-rabi=b̄* *∅-r-e-a=b̄h̄*
 and=EMPH hawk 3-over=LOC 3-CTFG-INTR-move=CONT
- idr* *d̥aθ̄* *hem̄lala* *d̥u* *dobide.*
i-d̄i *d̥a=θ̄* *heb̄lala* *d̄u* *∅-d̄-∅-obi=d̄-e*
 3-INSTR and=REPET snake 3.LOC 3-CTPT-INTR-see=CTPT-IMP
 ‘The hawk kept flying over it [the snake], and then the snake saw it [the hawk] too.’
- F. *nawiihik̄* *d̄am̄* *deθede*
dawiihik̄ *d̄ab̄* *∅-d̄-e-θe=d̄-e*
 hawk 3.AL 3-CTPT-INTR-move.down=CTPT-IMP
- d̄am̄* *raaōbinan̄krem̄.*
d̄ab̄ *∅-r-a-aōbidad̄=k̄re=b̄*
 3.AL 3-CTFG-INTR-fight=FUT=CONV
 ‘Then the hawk came down, in order to fight with it [the snake]’
- G. *d̥ahe* *wiwəna* *raaōbinan̄m̄h̄ren̄.*
d̥a=he *wi=wəda* *∅-r-a-aōbidad̄=b̄h̄=r-ed̄*
 and=EMPH both=COM 3-CTFG-INTR-fight=CONT=CTFG-PL
 ‘They fought each other.’
- H. *idr* *d̥ahe* *nawiihik̄* *hem̄lala* *rrəbun̄re.*
i-d̄i *d̥a=he* *dawiihik̄* *heb̄lala* *∅-r-i-rəbud̄=r-e.*
 3-INSTR and=EMPH hawk snake 3-CTFG-TRANS-kill=CTFG-IMP
- idr* *rrwim̄* *ruore.*
i-d̄i *∅-r-i-∅-wi=b̄* *∅-r-∅-uɔ=r-e*

3-INSTR 3-CTFG-TRANS-3-carry=CONV 3-CTFG-INTR-fly=CTFG-IMP
 ‘Then the hawk killed the snake and flew away, taking it.’

I. *kiehe* *idʒĩĩ.*
 kie=he *idʒĩĩ*
 that=EMPH story
 ‘That’s it, the story.’

Thus, there is a strong tendency to assign the role of deictic center to the most salient character or location, and not necessarily to the character or location which is physically closer to the speaker. The same tendency is manifested in other, longer texts, such as in the text ‘Krysa-my Ijyy,’ which tells the story of a war occurred between the Karajá and the Xavante, their traditional enemies. The text was narrated by a Karajá man in the same village where some of the narrated events took place. Most of the story is about an expedition of Karajá warriors to Xavante territory in order to revenge the killing of a Karajá youth. The Karajá village is initially presented as the deictic center. At a certain point, the Karajá men ask for the help of Bandeira, a White man, who lends them firearms. This fragment of the text is reproduced below (17). Notice that the narrator adopts Bandeira’s viewpoint (*dɔɪdɛdã=de* ‘they came’) and not the viewpoint of the Karajá men. This reflects the more central role played by Bandeira, the man who has the firearms, at this point of the narrative.

(17) **Fragment of a Karajá text (Ribeiro 1999)**

- a. *dãmãle* *dɔɪdenãde* *duidzĩmã.*
dãbã=le Ø-d-Ø-ɔɪ=d-ɛdã=d-e dɔ=ɪdzĩ=ãbã
3.AL=EMPH 3-CTPT-move(plural)=CTPT-PLURAL=CTPT-IMP 3.LOC=story=LOC
'They *came* to him to tell the story.'
- b. *dahe* *mãawãmãawãmã* *residɔenãrenãre.*
dã=he bãawa-bãawa=bã Ø-r-ɛ-sidɔɛdã=r-ɛdã=r-e.
and=EMPH firearm-REDUP=LOC 3-CTFG-INTR-ask.for=CTFG-PLURAL=CTFG-IMP
'And they asked him for firearms.'

Notice that while (17a) is marked for centripetal direction, the verb in the following sentence, (17b), is marked for centrifugal direction, a fact rather common in narrative texts. This suggests that the choice of a centripetal verb is made only to signal a shift in the deictic center. Once the identity of the new deictic center is stated, the narrator takes a neutral position. The same tendency can be noticed in the text 'The Hawk and the Snake' above. In Line F, the verb *deθe=de* 'he/she moved down (hither)' signals a shift in the deictic center, reflecting the fact that now the crucial actions are going to take place on the ground. However, the verb in the following sentence, *raaθbidãdã=kre=bã* 'in order to fight,' presents centrifugal marking.

The data presented here suggest that directional marking is commonly used as a tool to confer dramaticity and dynamism to the narrative. In narrating an event in which an action can be contemplated from different angles, the speaker rarely adopts a static point of view. Thus, even the "enemy" can be chosen as the deictic center, if the more intense actions are taking place in the enemy's field. Thus, in the text 'Krysa-my Ijyy,' the deictic center is switched to a Xavante man who is shot (18).

(18) **Fragment of a Karajá text**

- a. *darəkthe imadialemō diwede,*
d̥a=rəkɪ=hɛ i-ba-d̥i̯a=lɛ=b̥ō Ø-d-ɪ-wɛ=d-e
and=NAR=EMPH3-liver-middle=EMPH=LOC 3-CTPT-TRANS-penetrate=CTPT-IMP
'Then [the bullet] penetrated him right in his liver.'
- b. *rəkiləhə rahinōhiki reθere.*
rəkɪ=ləhə Ø-r-a-hid̥ō=hiki Ø-r-ɛ-θɛ=r-e
NAR=COMP 3-CTFG-INT-cry=EMPH 3-CTFG-INTR-fall=CTFG-IMP
'Crying a lot, the poor thing fell.'

Note that the second sentence in the fragment above presents the particle *=ləhə*, which indicates 'compassion.' This reinforces the interpretation given here of centripetal marking as a mechanism to signal empathy towards a given character. As we have seen in Chapter 1, Karajá presents a number of *attitude markers*, discourse-oriented particles indicating the attitude of the speaker in relation to what he or she is uttering—examples of which are *=kɔɪɪ* 'admiration, surprise,' *=ka* 'certainty', *=θɔ* 'excitement', *=ləhə* 'compassion', *=d̥ō* 'sympathy', and *=kɔθəd̥ō* 'doubt.' The interaction between directional marking and such particles constitutes a rich theme to be explored in future studies.

2.2.3 Pragmatic restrictions and evidentiality: the speaker as an observer

Despite the pervasiveness of directional marking in Karajá, a preliminary study (Ribeiro 2002) revealed what seemed to be a striking gap in the directional paradigm: the occurrence of centripetal markers with progressive verbs seemed to be rather limited.³⁶

³⁶ Another unexplained gap remains to be further investigated: the absence of centripetal marking with perfective verbs, marked by the auxiliary *=r-a*. In addition, as I mentioned above, non-verbal predicates ('descriptives,' postpositional phrases, etc.) do not occur with the centripetal marker, although they take the

As it turns out, this apparent exception is due to somewhat unusual pragmatic constraints, rather than to a morphological idiosyncrasy. Although a few occurrences of 3rd person verbs in the centripetal direction could be found in the text corpus (19), there are no examples of centripetal progressive verbs in the 1st or 2nd persons. The example below is from a traditional Karajá story:

- (19) *idïï dïbimã robire, d'aina dïï dïbimã,*
 i-dïï dïbi=b ã Ø-**r**-Ø-obi=**r**-e d'aidã dïï dïbi=bã
 3-skin old=LOC 3-CTFG-INTR-see=CTFG-IMP Tainá skin OLD=LOC
- d'ai d'ahε naderi*
 d'ai d'a=hε Ø-**d**-Ø-ã=**d**-εri
 3.LOC ASSERT=EMPH 3-CTPT-INTR-go=CTPT-PROGR
 'She saw his old skin, Tainá's old skin, it was there [visible to her].'

As it turns out, unlike its counterparts with imperfective, future, or potential verbs, a progressive verb inflected for centripetal direction conveys an extra piece of evidential information, implying that the narrated event is being witnessed by the speaker (or by whomever is assigned the role of deictic center; see (19) above): *dãribe=dεri* 'he is talking [and I'm witnessing him talk],'
duru=dεri 'he is dying (and I'm witnessing it),' etc. In the majority of the examples, such constructions are said to be appropriate only in indexical (=pointing) use:

- (20) a. *rariareri* b. *nariaderi*
 Ø-**r**-ã-ria=**r**-εri Ø-**d**-ã-ria=**d**-εri
 3-CTFG-INTR-walk=CTFG-PROGR 3-CTPT-INTR-walk=CTPT-PROGR
 'He is walking.' 'He is walking [pointing].'

same series of tense-aspect morphemes which occur with 'true' verbs. Directional inflection is an exclusively verbal category in Karajá.

Such evidential use restrains the range of pragmatic situations in which centripetal progressive verbs can occur, helping to explain the rarity of such verbs in narrative texts, for example. Why then are 1st and 2nd person centripetal progressive verbs even harder to come by? Oppositions between sentences such as ‘I am coming’ vs. ‘I am going’, and ‘I am bringing it’ versus ‘I am taking it away’ seem to be neutralized, as suggested by examples such as (21a, b) and (22a) below, in which a progressive verb marked with the centrifugal prefix *r-* may be interpreted as denoting either a centrifugal or a centripetal motion:

- (21) a. *dad̄eri* *anōbo?*
 d̄a-∅-a=d̄-eri adōbo
 2-CTFG-go=2-PROGR QUEST
 ‘Are you going/coming?’
- b. *r̄ari!*
 r-a-∅-a=r-eri
 CTFG-1-CTFG-INTR-go=CTFG-PROGR
 ‘I am going/coming.’
- (22) a. *d̄ikar̄* *ar̄əθ̄na* *r̄ew̄iri*
 d̄ikar̄ a-r̄əθ̄dā r̄-a-∅-i-wi=r-eri
 I 2-food CTFG-1-CTFG-TRANS-carry=CTFG-PROGR
 ‘I’m bringing/taking your food.’

At first, the same progressive verb seemed to never occur with centripetal markers (b), either in the text corpus or in elicitation attempts:

- b. *?d̄ikar̄* *ar̄əθ̄na* *nad̄iw̄iri*
 d̄ikar̄ a-r̄əθ̄dā d̄-a-d̄-i-wi=d̄-eri
 I 2-food CTPT-1-CTPT-TRANS-carry=CTPT-PROGR
 ‘I’m bringing your food.’

However, as it turns out, constructions such as (22b) above are indeed possible, but in very rare contexts. Constructions such as (22b), (23b), and (24b) are acceptable if the events reported are being seen indirectly, through a video recording (!) or a mirror, for example.

- (23) a. *dariaderi*
 da-Ø-ã-ria=dɛɾɪ
 2-CTFG-INTR-walk=2-PROGR
 ‘You are walking.’
- b. *danariaderi*
 da-**d**-ã-ria=d-ɛɾɪ
 2-CTPT-INTR-walk=2-PROGR
 ‘You are walking [in the video].’
- (24) a. *rariareri*
 r-a-Ø-a-ria=rɛɾɪ
 CTFG-1-CTFG-INTR-walk=CTFG-PROGR
 ‘I am walking.’
- b. *nanariaderi*
d-ã-**d**-ã-ria=**d**-ɛɾɪ
 CTPT-1-CTPT-INTR-walk=CTPT-PROGR
 ‘I am walking [in the video].’

A plausible generalization, to be further investigated, is that centripetal direction in progressive verbs is acceptable when describing a narrated event, but not as a part of the speech event itself.

2.3 Valence and voice

Karajá verbs are lexically either transitive or intransitive.³⁷ Intransitive verbs may have their valence increased through causativization (2.3.1.1) or through oblique promotion

³⁷ Maia (1998, 79) mentions the existence of ‘diffuse verbs’, that is, verbs that can be used either transitively or intransitively without any morphological alteration. In our data, however, the only verb he mentions as being ‘diffuse,’ -*okã* ‘to dry’, has exactly the same behavior of other intransitive verbs, such as *oka* ‘to be cooked’. As shown in the example (b) below, this verb presents transitive morphology when used transitively. Notice that the transitive stem is a denominal verb formed by the deverbal noun *orã* ‘the action of drying’ followed by the verbalizer suffix -*dã*:

- (a) *bee rokãre*
 be Ø-r-Ø-ukã=r-e
 water 3-CTFG-INTR-dry=CTFG-IMPERF
 ‘The water dried.’
- (b) *tfuu bee rɪdʊrãñãkre*
 du be Ø-r-ɪ-d-*orã*-dã=kre
 sun water 3-CTFG-TRANS-3/REL-dry-VERB-FUT
 ‘The sun will dry the water.’

(2.3.1.2). Transitive verbs, on the other hand, may have their valence decreased through reflexivization (2.3.2.1), passivization (2.3.2.2), and antipassivization (2.3.2.3).

2.3.1. Intransitive verbs

Intransitive verbs are those that do not take a direct object as one of their arguments, such as *-udədə* ‘to become cold’, *dədəke* ‘to become hot’, and ♀ *rika* ~ *ritfa* (♂ *ria*) ‘to walk’. As we have seen (Chapter 1), i-class intransitive verbs are generally marked by the prefix *a-*, while f-class intransitive verbs are marked by a zero allomorph. In addition, a few intransitive verbs, such as *θe* ‘to fall’, are marked by the prefix *ε-*. The class of intransitive verbs includes not only one-place verbs such as *rika* ‘to walk’ and *udədə* ‘to become cold’, but also *extended intransitive* verbs such as *-obi* ‘to see’, whose arguments are oblique NPs—in this case, a locative, marked by the postposition *bə* ‘diffuse locative’ (25). Although notionally transitive, such verbs behave as intransitive for all purposes. For example, they cannot be made passive or antipassive, and their arguments cannot be incorporated.

- (25) *dɪkarə haləkəmə rabire*
dɪkarə haləkəε=bə r-a-Ø-obi=r-e
 I jaguar=LOC CTFG-I-INTR-see=CTFG-IMPERF
 ‘I saw the jaguar.’

Most intransitive verbs can be transitivized, either through causativization or through the promotion of an oblique to direct object. The transitivized stem is formed by the nominal

form of the verb plus the verbalizer suffix *-dǎ*. This is illustrated in the example (26b) below, where the intransitive verb *-uka* ‘to be cooked’ is transitivized.³⁸

- (26) a. *iweru røkareri*
iweru Ø-r-Ø-uka=r-eri
 calugi 3-CTFG-INTR-be.cooked=CTFG-PROGR
 ‘The *calugi* (a kind of drink) is cooking.’
- b. *ahāwəki iweru rɪdʊranǎkre*
a-hāwəki iweru Ø-r-ɪ-d-ura-dǎ=kəre
 2-woman calugi 3-CTFG-TRANS-3/REL-be.cooked-VERB=FUT
 ‘Your wife will cook the *calugi*.’

2.3.1.1. Causativization

Causative stems derived from unergative verbs, such as *rika* ‘to walk’, are formed with the causativizer suffix *-dəkǎ* plus the verbalizer suffix *-dǎ* (27). However, the causative suffix does not occur in causative stems derived from unaccusative verbs, such as *-uka* ‘to be cooked’ in (26) above.

- (27) *habu kuladu riradəkǎnǎeri*
habu kuladu Ø-r-ɪ-rira-dəkǎ-dǎ=r-eri
 man child 3-CTFG-TRANS-walk-CAUS-VERB=CTFG-PROGR
 ‘The man is making the child walk.’

2.3.1.2. Oblique promotion

³⁸ This example illustrates a very common process for deriving nouns from verb roots, namely *consonantal replacement*, which consists in replacing a velar stop or a glottal fricative occurring in the last syllable of the verb root with an alveolar flap: *rika* ‘to walk’ > *rira* ‘the action of walking’, *rira-dǎ* ‘walking place’, *rira-du* ‘the one who walks’; *θohɔ* ‘to wash’ > *θurɔ* ‘the action of washing’, *θurɔ-dǎ* ‘washing material’, *θurɔ-du* ‘the one who washes’ (see Section 4.1). Thus, the transitive stem in (12b) above is constructed with the nominal form of the verb *uka* ‘to be cooked’, *ura* ‘the action of cooking’, followed by the verbalizer suffix *-dǎ*.

With a few *extended intransitive* verbs which take an allative or dative argument, such as *ɛkɛrafi* ‘to ask’, transitivization results in the promotion of the former oblique argument to direct object (examples from the Xambioá dialect):

- (28) a. *hawiki* *ɔarikɔɛkɔ* *rɛkɛrafire*
hawiki ɔa-rikɔɛ=kɔ Ø-r-Ø-ɛkɛraθi=r-e
woman 3REFL-child=AL 3-CTFG-INTR-ask=CTFG-IMPERF
‘The woman asked her son.’
- b. *hawiki* *ɔarikɔɛ* *ɾɔɛkɛrafinɔɛ*
hawiki ɔa-rikɔɛ Ø-r-ɾ-ɔkɛraθi-ɔɛ=r-e
woman 3REFL-child 3-CTFG-TRANS-3/REL-ask-VERB=CTFG-IMPERF
‘The woman questioned her son.’

2.3.2 Transitive verbs

Transitive verbs are those that take a direct object as one of their arguments. In Karajá, transitive verbs are always marked by the prefix *ɾ-*, as shown by the paradigms given in (7) above. As mentioned above (Section 2.2), transitive and intransitive valence prefixes may fuse with the preceding personal prefix under certain circumstances, such as in the 2nd person in the centrifugal direction of the realis mood (29a). Notice that there is no fusion in the centripetal direction (29b).

- (29) a. *ɔerakode* b. *ɔadɾakode*
ɔa-Ø-ɾ-Ø-rakɔ=d-e ɔa-d-ɾ-Ø-rakɔ=d-e
2-CTFG-TRANS-3-wait=2-IMPERF 2-CTPT-TRANS-3-wait=2-IMPERF
‘You waited for him (thither).’ ‘You waited for him (hither).’

2.3.2.1 Reflexive

There are two allomorphs of the reflexive morpheme, *efi-* and *ifi-*. The former is incorporated into the verb, when the NP coreferential with the subject is a direct object (30a). The latter is attached to postpositions, when the coreferential NP is an oblique (30b).

- (30) a. *dɪkarə̃ kareʃiθuhokre*
 dɪkarə̃ ka-r-eθi-θuhɔ=kəre
 I 1-CTFG-REFL-wash=FUT
 ‘I will wash myself.’
- b. *habu ifimə̃ robire*
 habu iθi=bə̃ Ø-r-Ø-obi=r-e
 man REFL=LOC 3-CTFG-INTR-see=CTFG-IMPERF
 ‘The man saw himself.’

2.3.2.2. Passive

Passive verbs are marked by the prefix *a-*, with i-class stems such as *θuhɔ* ‘to wash’ (31b), or its zero allomorph, with d-class stems, such as *uka* ‘to split’ (32b). Notice that this is apparently the same prefix that occurs with basic intransitive verbs such as *rika* ‘to walk’ and *-obi* ‘to see’. With transitive roots, however, this prefix will always convey a passive or anticausative meaning.

- (31) a. *nadɪ wadəki rɪθuhɔreri*
 d-ädi wa-dəki Ø-r-ɪ-θuhɔ=r-eri
 REL-mother 1-clothes 3-CTFG-TRANS-wash=CTFG-PROGR
 ‘My mother is washing my clothes.’
- b. *wadəki raθuhɔreri*
 wa-dəki Ø-r-a-θuhɔ=r-eri
 1-clothes 3-CTFG-PASS-wash=CTFG-PROGR
 ‘My clothes are being washed.’

- (32) a. *kɔdʊ hãlɔkɔɛ kɔɾʊ ritfukare*
 kɔdʊ hãlɔkɔɛ kɔɾʊ Ø-r-ɾ-d-uka=r-e
 turtle jaguar forehead 3-CTFG-TRANS-3/REL-split=CTFG-IMPERF
 ‘The turtle split the jaguar’s forehead.’
- b. *hãlɔkɔɛ kɔɾʊ rukare*
 hãlɔkɔɛ kɔɾʊ Ø-r-Ø-uka=r-e
 jaguar forehead 3-CTFG-PASS-split=CTFG-IMPERF
 ‘The jaguar’s forehead was split.’

In the passive construction, the original O becomes the subject, as happens in languages such as English, for example. However, unlike English, where the agent in a passive construction can be expressed as an oblique (*‘by-phrase’*), in Karajá the agent, although sometimes implicit, cannot be expressed at all. Thus, passives in Karajá are both a *backgrounding* construction, functioning to delete unknown or irrelevant subjects, and a *foregrounding* construction, since they result in the promotion of the original O to subject position (Foley and Van Valin 1985).

2.3.2.3. Antipassive

Antipassive is a phenomenon typical of ergative languages, corresponding functionally to a ‘mirror image’ of the passive construction in nominative-accusative languages (Silverstein 1976). In a syntactically ergative language, “while the A and the O in an ergative clause are marked as ergative and absolutive respectively, the A in an antipassive is typically coded as an absolutive NP, and the O (if present) appears in a case other than the absolutive” (Cooreman 1994, 50).³⁹ Although some authors, such as Cooreman, limit the discussion of antipassive constructions to ergative languages, nominative-accusative

³⁹ I will follow Cooreman in adopting Dixon’s (1979) use of the labels A and O to refer to the two participants in a two-participant clause—prototypically, the agent and the patient, respectively.

languages may also present *backgrounding* antipassives, which “function to demote the undergoer to peripheral status” (Foley and Van Valin 1985: 338). This is what occurs in Karajá, where antipassive, marked by the prefix *ɔ-*, results in the deletion of an unknown or irrelevant direct object:

- (33) *nadi rɔθuhɔɾɛɾɪ*
 d-ãdi Ø-r-ɔ-θuhɔ=ɾɛɾɪ
 REL-mother 3-CTFG-ANTI-wash=CTFG-PROGR
 ‘My mother is washing (something).’
- (34) *hãbu rotfukarɛɾɪ*
 hãbu Ø-r-ɔ-d-uka=r-ɛɾɪ
 man 3-CTFG-ANTI-3/REL-split=CTFG-PROGR
 ‘The man is splitting (something).’

As these examples show, antipassive in Karajá is not promotional (or foregrounding), in the sense that the A remains in the same syntactic relation it occupies in the corresponding active, transitive voice. Furthermore, the antipassive construction in Karajá does not allow the expression of the demoted O whatsoever, which is an interesting parallel to what occurs with the agent in the passive construction.

Antipassives in Karajá are used to suppress an irrelevant, unknown, or undetermined object. Common examples involve typical activities such as ‘to clear (a garden)’:

- (35) *idzoi woba=di Ø-r-ɔ-lafiuɾ=kɾe*
 guys ax=INSTR 3-CTFG-ANTI-clear=FUT
 ‘The guys will clear [the garden] with axes.’

At least one intransitive stem, *-ɔbrɔ*, seems to have resulted from reanalysis of a formerly transitive stem including the antipassive marker, as suggested by the existence of an allomorph without the initial vowel (*brɔ*; cf. also the nominal form, *a-l-ɔbrɔ-da* ‘plant’). As

the example below shows, *-ɔbrɔ* ‘to plant’ is an intransitive, intrinsically antipassive verb, which needs to be transitivized in order for it to occur with objects:⁴⁰

- (36) *warikɔrɛ* *wadeke* *rikɔrɔmɔ́dʌ*
 wa-rikɔrɛ wa-dɛkɛ Ø-r-ɪ-kɔrɔ=bɔ́=dʌ
 1-child 1-DAT 3-CTFG-TRANS-cut=CONV=ASSERT

dʊrɔbɔrɔwahãre,
 dʊ=r-a-Ø-ɔbɔrɔ=wahã=r-e,
 3.LOC=CTFG-1-INTR-plant=1.HABIT=CTFG-IMPERF

adzikura *rebɔrɔdɔ́wahãre,*
 adzikura r-a-Ø-ɪ-bɔrɔ-dɔ́=wahã=r-e
 manioc CTFG-1-TRANS-plant-VERB=1.HABIT=CTFG-IMPERF

karalibi *rebɔrɔdɔ́wahãre*
 kara-libi r-a-Ø-ɪ-bɔrɔ-dɔ́=wahã=r-e
 yam-black CTFG-1-TRANS-plant-VERB=1.HABIT=CTFG-IMPERF
 ‘Once my son cuts [i.e. clears] it, I plant in it. I plant manioc, black yam...’

Although the antipassive was not described by previous authors, it is commonly found in both narrative and descriptive texts, commonly being used in contexts such as the one above, in which an antipassive form of the verb is later followed by a transitive one, specifying the object. Another example, from a traditional tale, is provided below. The sentences are part of a passage in which two young men, returning to their village, hear someone pounding something with a mortar. At first, the verb *θi* is given in an antipassive form, later followed by a transitive one:

- (37) *idɔ́* *idatʃi=rɛkɪ* *Ø-r-ɔ-θi=r-ɛrɪ,*
 Karajá two=QUOT 3-CTFG-ANTI-pound=CTFG-PROG

berɔ *Ø-r-ɪ-θi=r-ɛdɔ́=r-ɛrɪ*

⁴⁰ Notice that the examples above are from the Xambioá (that is, a non-schwa) dialect.

manioc.flour 3-CTFG-*TRANS-pound*=CTFG-PROG
'Two people are pounding. They are pounding manioc flour.'

2.5 Pronominal objects

Pronominal direct objects are obligatorily incorporated into the verb (38). The series of direct object prefixes coincides partially with the series occurring with nouns and descriptive predicates. Pronominal subjects of transitive and intransitive predicates are expressed by free pronouns (*dɪkarə* 'I', *kai* 'you', *dəkɪ* 'he, she, it').

- (38) a. *dəkɪ* Ø-r-ɪ-wa-rakɔ=kərə
s/he 3-CTFG-TRANS-1-wait=FUT
'S/he will wait for me.'
- b. *dəkɪ* a-r-a-rakɔ=kərə
s/he 2-CTFG-2-wait=FUT
'S/he will wait for you.'
- c. *dəkɪ* Ø-r-ɪ-Ø-rakɔ=kərə
s/he 3-CTFG-TRANS-3-wait=FUT
'S/he will wait for her/him.'

2.6. Noun incorporation

Noun incorporation in Karajá is a process by which the head of the absolutive noun phrase is inserted into the verb, thereby forming a compound. The more productive pattern of noun incorporation involves only possessed nouns—that is, productive noun incorporation in Karajá is a possessor-raising strategy. Since only the head of the absolutive noun phrase is incorporated, the valence of the resulting noun-verb compound remains unaltered, as the

possessor is promoted to subject with intransitive, unaccusative verbs such as *bɔhɔ* ‘to break’ (38), or to object with transitive verbs such as *dɔka* I ‘to tie’ (39).⁴¹

(38) a. idǎ wɛ=ɾɪkɪ ɖai Ø-r-a-bɔhɔ=r-e
 people belly=NARR 3.LOC 3-CTFG-INTR-break=CTFG-IMPERF
 ‘The people’s bellies were broken there, it is said.’

b. idǎ=ɾɪkɪ ɖai Ø-r-a-wɛ-bɔhɔ=r-e
 people NARR 3.LOC 3-CTFG-INTR-belly-break=CTFG-IMPERF
 ‘The people’s bellies were broken there, it is said.’

(39) a. kǎdǎjɪwɛ kuθehewe dɾi Ø-r-ɪ-dɔka=r-e
 K. rhea leg 3-CTFG-TRANS-tie=CTFG-IMPERF
 ‘Kynyxiwe tied the legs of the rhea.’

b. kǎdǎjɪwɛ kuθehewe Ø-r-ɪ-dɾi-dɔka=r-e
 K. rhea 3-CTFG-TRANS-leg-tie=CTFG-IMPERF
 ‘Kynyxiwe tied the legs of the rhea.’

Since, as we have seen above, noun incorporation is a valence-preserving process, an incorporating transitive verb can still be made passive or antipassive.

⁴¹ Examples (21a) and (21b) are from the Javaé dialect. Although Maia (*op. cit.*: 63) claims that object incorporation does not occur in Javaé, noun incorporation seems to be as common in Javaé as it is in the other three dialects. The example below, involving the incorporation of the noun *dikɔhu* ‘knee’ to the transitive verb *wɛ* ‘to penetrate’, occurs in the same text from which the examples above were obtained:

(a) *ɾɪdikɔhuwere*, *ɪdikɔhu* *ɾiwere*
 Ø-r-ɪ-Ø-dikɔhu-wɛ=r-e *i-dikɔhu* Ø-r-ɪ-wɛ=r-e
 3-CTFG-TRANS-3-knee-penetrate=CTFG-IMPERF 3-knee 3-CTFG-TRANS-penetrate=CTFG-IMPERF
 ‘[He] stabbed him in the knee, he stabbed his knee.’

- (40) a. *dori wa-riɔɾɛ Ø-r-ɪ-ra-θi=kre*
 White 1-child 3-CTFG-TRANS-head-pluck=FUT
 ‘The White man will shave my child’s head.’
- b. *wa-riɔɾɛ Ø-r-a-ra-θi=kre*
 1-child 3-CTFG-head-pluck=FUT
 ‘My child will have his head shaved.’
- c. *dori Ø-r-ɔ-ra-θi=kre*
 White 3-CTFG-ANTI-head-pluck=FUT
 ‘The White man will shave [someone’s] head.’

Noun incorporation of body-part terms is generally used to form compound verb stems with a more specific meaning, referring to ‘nameworthy’ activities, such as *dɪ-dʌka* in example (39) above: while other animals can be tied in different ways, a bird is more commonly tied by the legs. A similar example, overheard from a man instructing his son on how to kill a chicken, is given below (41). Additional examples are *kɔ-krɔ* [wood-cut] ‘to clear (a garden)’, *kɔ-θuɦɔ* [face-wash] ‘to wash [someone’s] face’, and *-ɔdɪ-krɔ -ɔdɪ-krɔ* [neck-cut] ‘to cut [someone’s] neck’. Although every incorporating construction is said to be replaceable with a more analytical, non-incorporating version, there is a tendency to use incorporation for more idiomatic compounds: *kɔ-θuɦɔ* ‘to face-wash’, for instance, refers to a ritual in which someone washes the face of a masked performer (cf. Toral 1992), although it can also be used literally. The same can be said about *-ɔdɪ-krɔ*, which has an idiomatic use (‘to become downcast (Portuguese *cabisbaixo*)’), in addition to a literal one, ‘to cut [one’s] neck’.

- (41) b- \emptyset -r-*θira-θi*=kε,
2-CTFG-TRANS-feather-pluck=POT

b- \emptyset -r-*d-ɔdi-θira*-θi=kε
2-CTFG-TRANS-3-neck-feather-pluck=POT
'Pluck its feathers, pluck its neck's feathers!'⁴²

2.6.1 Incorporation of *bəde* 'world, time'

In addition to the incorporation of body-part terms, which does not alter a verb's valence, a less common type involves the incorporation of the noun *bəde* 'world, time', which occurs as a direct object in idiomatic constructions such as the ones in (42) and (43) below. Examples such as these are the only case in which noun incorporation results in valence change, turning transitive verbs into intransitive ones.

- (42) a. *badʊkari bəde* \emptyset -r-r-*kəri=r-a*
old.man world 3-CTFG-TRANS-know=CTFG-PERF
'The old man found out.'
- b. *badʊkari* \emptyset -r-a-*bəde-kəri=r-a*
old.man 3-CTFG-INTR-world-know=CTFG-PERF
'The old man found out.'
- (43) a. *hābu bəde* \emptyset -r-r-*riri=r-a*
man world 3-CTFG-TRANS-remember=CTFG-PERF
'The man became sad.'
- a. *hābu* \emptyset -r-a-*bəderiri=r-a*

⁴² These examples may suggest, at the first sight, the existence of double incorporation (first of *θira*, followed by the incorporation of *ɔdi*). It is more likely, however, that *ɔdi-θira* is a compound, in which case incorporation would have taken place only once.

man 3-CTFG-INTR-worldremember=CTFG-PERF
 ‘The man became sad.’

Although the noun *bəde* also occurs as the subject in ‘impersonal’ constructions—particularly verbs referring to weather-related phenomena, such as in *bəde rudʔəfə=reri* ‘the weather is getting cold’, it cannot be incorporated under such circumstances.

2.6.2 Incorporation of *wi* ‘reciprocal’

In addition to the incorporation of body-part nouns and the semantically empty noun *bəde*, the morpheme *wi* ‘both; RECIPROCAL’ can also be incorporated, with intransitive verbs such as *kudʔ* ‘to gather’ and *rareə* ‘to separate’--verbs which are semantically intrinsically reciprocal. Once again, both the incorporating form and its corresponding analytical form are considered as synonymous by the native speakers.

(44) ♂ a. dawii *wii* Ø-r-a-*kudʔ*=r-a
 bird **REC** 3-CTFG-INTR-*gather*=CTFG-PERF
 ‘The birds gathered.’

a. dawii Ø-r-a-*wi-kudʔ*=r-a
 bird 3-CTFG-INTR-**REC**-*gather*=CTFG-PERF
 ‘The birds gathered.’

(45) ♂ a. dawii *wii* Ø-r-a-*rareə*=r-a
 bird **REC** 3-CTFG-INTR-*separate*=CTFG-PERF
 ‘The birds scattered.’

a. dawii Ø-r-a-*wi-rareə*=r-a

bird 3-CTFG-INTR-*REC-separate*=CTFG-PERF
 ‘The birds scattered.’

2.6.3 Verbalization of noun-noun compounds

A number of verbs present compound stems which superficially resemble noun incorporation constructions, but which are actually the result of the verbalization of noun-noun compounds. In some cases, the second element in the noun-noun compound is a deverbal noun, creating constructions which may look, at first sight, as the result of valence-changing noun incorporation, as in the example below:

- (46) ♂ a. biu=ki Ø-r-a-*iwa-dɔ*-dã=bã
 top=LOC 3-CTFG-INTR-*jatobá-eat*-VERB=CONV
- Ø-r-Ø-udã=bãhã
 3-CTFG-INTR-sit=HABIT
 ‘He sat at the top eating jatobá fruit.’

That *dɔ* ‘to eat (sweets), suck’ in (46) is in its nominal form (derived by zero-affixation or conversion) is shown by the fact that it requires the verbalizing suffix *-dã*. That is particularly clear when the deverbal noun is derived by consonantal replacement, as in (47) below, where *ri* is the nominal form of the verb *kri* ‘to eat (grains, bread, etc.)’:

- (47) ♂ a-r-a-*bolo-ri*-dã=kre
 a-CTFG-INTR-*cake-eat*-VERB=FUT
 ‘I will do some cake-eating.’

2.6.4 Derivation as incorporation: causative *-dekã*

All the incorporation constructions described above fit well into the traditional, Sapirian definition of noun incorporation, since they involve independent noun and verb stems which can also occur in non-incorporating constructions. I suggest, however, that Karajá also has a less usual type of incorporation, similar to what has been described by Sadock for West Greenlandic (Sadock 1980), involving the causativizer *-dəkə̃*. As we have seen, *-dəkə̃* selects only nominal stems, in such a way that, in order to be causativized, a verb root must first be nominalized. I suggest that *-dəkə̃*, although morphologically dependent, is a noun-incorporating stem.

Such an analysis is particularly appealing if one takes into consideration *extended intransitive verbs*, which, as we have seen, take non-canonically marked objects. That is the case of the verb *-ehu* ‘to throw’, whose objects are marked with the instrumental postposition *=dɪ* (48a). While, with transitive verbs, the original object cannot be expressed under causativization, with extended intransitive verbs the original object is preserved with its original marking, and the original subject (the causée) is expressed as a direct object (48b):

- (48) a. *kolado mənadɪ* *rehure* *bədɔɪdɪ*
kolado bə̃dā=dɪ Ø-r-Ø-ehu=r-e bədɔɪ=dɪ
child rock=INSTR 3-CTFG-INTR-throw=CTFG-IMPERF sling=INSTR
‘The child threw the rock with a sling.’
- b. *hābu kolado* *mənadɪ*
hābu kolado bə̃dādɪ
man child rock=INSTR
- ridəhudəkə̃nə̃re* *bədɔɪdɪ*
Ø-r-ɪ-d-ehu-dəkə̃-də̃=r-e bədɔɪ=dɪ
3-CTFG-TRANS-3?-throw-CAUS-VERB=CTFG-IMPERF sling=INSTR
The man made the child throw the rock with a sling.’

The main argument for treating causativization as a case of noun incorporation à la Sadock (i.e., a case of syntactic word-formation) is the fact that the original verb root seems to preserve some syntactic saliency, still assigning case to its original, non-canonically marked object.

3. Noun morphology

As we have seen (Chapter 1), the only inflectional morphology occurring with nouns are the pronominal markers of possessors, listed again in Table 2 below. Most noun stems can be divided into two lexical classes, depending on the personal prefixes they take, arbitrarily labeled i-class and d-class, after the 3rd person marker with which they occur. The main difference between both stems of both classes is in the series of personal prefixes they take, as illustrated below by the paradigms for *kəɾʊ* ‘forehead’ (49) and *ɛbɔ* ‘hand’ (50). The series of personal prefixes occurring with both classes are summarized in Table 2 below.

(49)	<i>i-class</i>		(50)	<i>d-class</i>	
	<i>habu kəɾʊ</i>	‘man’s forehead’		<i>habu d-ɛbɔ</i>	‘man’s hand’
	<i>wa-kəɾʊ</i>	‘my forehead’		<i>wa-d-ɛbɔ</i>	‘my hand’
	<i>a-kəɾʊ</i>	‘your forehead’		<i>∅-ɛbɔ</i>	‘your hand’
	<i>i-kəɾʊ</i>	‘his/her/its forehead’		<i>d-ɛbɔ</i>	‘his/her/its hand’, or
	<i>d’a-kəɾʊ</i>	‘his/her/its own forehead’			‘his/her/its own hand’

Table 4.2. Possessive prefixes in Karajá (Ribeiro 1996)⁴³

Person	i-class	d-class
1 st	<i>wa-</i>	<i>wa-d-</i>
2 nd	<i>a-</i>	<i>∅-</i>
3 rd	<i>i-</i>	<i>d-</i>
3 rd REFL	<i>d’a-</i>	

⁴³ Although all d-class stems are vowel-initial and most of i-class stems are consonant-initial, the distinction cannot be reduced to phonological terms, since the i-class also includes some vowel-initial stems, such as *aʃikɔ* ‘arm’ (*ʃ aʃikɔ*), *ɛθɔ* ‘cotton’, *ari* ‘to gather’, *ɛʃi* ‘younger brother’, etc.

Whereas the i-class prefix series distinguishes a reflexive third person (*dʼa-*) from a non-reflexive one (*i-*), the d-class series has only one third person prefix (*d-*), which covers the range of meanings of both reflexive and non-reflexive third persons. Furthermore, the d-class stem *-ɛbɔ* ‘hand’ presents a prefix *d-* in the first person and when preceded by a nominal possessor. The function of this prefix is synchronically fairly opaque, but its distribution resembles that of *relational prefixes*, linking prefixes which have been described for several other Macro-Jê families. Relational prefixes were first described as a grammatical peculiarity of Tupí-Guaraní languages, and their occurrence in languages of Karíb and Macro-Jê stocks, as well as in languages of other branches within the Tupí stock, has been pointed out as evidence for a genetic relationship among these three groups (Rodrigues 1994). Such linking prefixes are very likely cognates with similar morphemes in Jê and other families within Macro-Jê (Ribeiro 2004, 2005b, 2011) and provide further corroboration for the inclusion of Karajá in the Macro-Jê stock (see Chapter 6).

3.1 Personal pronouns

There are three independent personal pronouns in Karajá: ♀ *dɪkarẽ* (♂ *dɪarẽ*) ‘I’, *kai* ‘you’, and *dɔkɪ* (♂ *dɪɪ*) ‘he, she, it’. All of them can be pluralized by the pluralizer *boho*, which also occurs with nouns. The plural form of the 1st person pronoun has an exclusive reading (51). For an inclusive 1st person plural reading, the noun *idẽ* ‘human, Karajá’ is used with the pluralizer *boho*. The verb, in such cases, is generally inflected for third person,

except in the centrifugal direction of the irrealis mood, when the verb occurs with the prefix *rək-* (52):

(51) *dɪrə̃ boho kədura a-r-r-rɔ=rɛdɔ̃=kre*
 I PL fish 1-CTFG-TRANS-eat=CTFG-PL=FUT
 ‘We (excl.) will eat fish.’

(52) *idɔ̃ boho kədura rək-r-rɔ=r-ɛdɔ̃=kre*
 Karajá PLURAL fish 1PL.INCL.-TRANS-eat=POT
 ‘We (incl.) will eat fish.’

Personal pronouns are clearly distinct from nouns since, unlike nouns, they cannot occur as (direct or postpositional) objects or possessors (personal prefixes of the possessive series being used instead).

3.1.1 Grammaticalization of *idɔi* ‘guys’

As with *idɔ̃* ‘human, Karajá,’ the noun *idɔi* ‘guys’ seems to be undergoing a process of grammaticalization towards pronounness, being often used with first person (singular or plural) connotations. So, a sentence such as (53) below can have a literal reading (a), an exclusive (b) or inclusive (c) 1st person plural reading, a 1st person singular reading (d), or a soft imperative or exhortatory reading (e). In its non-literal uses, *idɔi* is used mostly in departing or arriving greetings (‘I’m out of here’, ‘Here we are’, etc.).

(53) *idɔi Ø-r-Ø-oi=kre*
 guys 3-CTFG-INTR-go.PL=FUT
 a. ‘The guys will go.’
 b. ‘We (incl.) will go.’

- c. 'We (excl.) will go.'
- d. 'I will go.'
- e. 'Let's go.'

The text fragment in (54), from a traditional tale (of which another, shorter version was published as Ribeiro 2005a), brilliantly illustrates how the ambiguity of *id3oi* can be exploited for comical effects. In the story, the monkey (*krɔbi*) kindly offers to take the turtle (*kɔdudĩ*) up a jatobá tree so that the latter could eat some fruit. The monkey then nonchalantly decides to leave the turtle behind, up on the tree. The monkey's departing words are (54a), to which the turtle replies with (54b); while the monkey clearly has a first-person singular meaning in mind, the turtle demands a first-person plural inclusive interpretation.⁴⁴

- (54) a. *id3oi* Ø-*r-εθε=kre* [re'θekre]
 guys 3-CTFG-INTR-fall=FUT
 'I'm out of here.' (Lit. 'The guys are going down.')
- b. *id3oi=he* Ø-*r-ε-θε* [re'θe],
guys=EMPH 3-CTFG-INTR-go.down
- d'a* *b-Ø-I-wa-adi=kre*
 then 2-CTFG-TRANS-1-fall=FUT
 '[You're damn right] *the guys* are going down, so take me down!'

The grammaticalization of *idɔ̃* and *id3oi* parallels that of Brazilian Portuguese *a gente* 'the people', which is mostly used with 1st person plural connotations (even though, as in Karajá, the verb displays third-person agreement). Another parallel to the grammaticalization of *id3oi* is the evolution of English *guys* as a 2nd person pronoun

⁴⁴ Notice that, in (10b), the verb illustrates a case of assimilation *in absentia* (cf. Chapter 2): the future marker =*kre*, although not present, triggers vowel harmony in the preceding [-ATR] vowels.

pluralizer. A major difference is that the gender of the original noun, semantically bleached in English (since *you guys* can be used to refer to both males and females), was not bleached in Karajá. As far as my data are concerned, the non-literal, 1st person uses of *idzoi* are restricted to male speech.

4. Derivational morphology

Karajá presents clear-cut morphological devices to derive nouns from verbs and vice-versa. Interestingly enough, while all inflectional affixes in Karajá are prefixes, all derivational affixes are suffixes. Furthermore, contrasting with its complex inflectional morphology, verbs apparently do not present any productive derivational morphology: as we will see, while all derivational morphemes attaching to noun stems (including the causativizer, as seen above, and the suffixes *-dʊ* and *-da*, which create subject and instrument nouns) have a regular and productive distribution, nominalizing morphology is largely irregular and lexically-determined.

4.1 Verbalization

Any noun can be turned into a verb by means of the verbalizer suffix *-dǎ*. Although the meaning of the derived verb may at times be idiosyncratic, verbs formed in this manner can generally be translated as ‘to provide OBJ with NOUN’ (when transitive) or ‘to be provided with NOUN’ (when intransitive).

- (51) ♀ dɪkarǽ ka-r-ɛʃi-wa-dǽ=kǽre
 I 1-CTFG-REFL-*foot/shoe*-VERB=FUT
 ‘I will put shoes on.’
- (52) hǽri wa-rikǽrɛ Ø-r-ɪ-d-*vahɪ*-dǽ=r-a
 shaman 1-child 3-CTFG-TRANS-3/REL-*medicine*-VERB=CTFG-PERF
 ‘The shaman gave medicine to my child.’
- (53) d-ǽdɪ Ø-r-a-dɛ-dǽ=r-e
 REL-mother 3-CTFG-INTR-*flesh*-VERB=CTFG-IMPERF
 ‘My mother got fat.’

4.2 Nominalization

There are at least three devices to create action nouns from verb stems: zero-affixation or ‘conversion’ (54), consonantal replacement (55), and suffixation (56).

- (54) *deverbal nouns derived by zero-affixation (‘conversion’)*

<i>verb</i>		<i>noun</i>
<i>ribe</i> I	‘to speak’	<i>ribe</i>
<i>dɔ</i> I	‘to eat (soft stuff)’	<i>dɔ</i>
<i>rɔ</i> I	‘to eat (hard stuff)’	<i>rɔ</i>
<i>era</i> II	‘to copulate’	<i>era</i>

- (55) *deverbal nouns derived by consonantal replacement*

<i>verb</i>		<i>noun</i>
<i>θu<u>h</u></i> I	‘to wash’	<i>θu<u>r</u></i>
<i>ri<u>k</u>a</i> I	‘to walk’	<i>ri<u>r</u>a</i>
<i>kɔ<u>k</u>a</i> I	‘to shred’	<i><u>k</u>ɔra</i>
<i>u<u>k</u>ǽ</i> II	‘to get dry’	<i>u<u>r</u>ǽ</i>
<i><u>k</u>i</i> I	‘to eat (grains)’	<i><u>r</u>i</i>

There are two nominalizing suffixes, *-dV* and *-θV* (where ‘V’ represents a copy of the last vowel of the root), whose distribution seems to be lexically determined:

(56) *deverbal nouns derived by suffixation*

<i>verb</i>		<i>noun</i>
<i>wɛ</i> I	‘to penetrate’	<i>wɛ-dɛ</i>
<i>bɔ̃</i> I	‘to catch’	<i>bɔ̃-dɔ̃</i>
<i>hʊ</i> I	‘to finish’	<i>hʊ-du</i>
<i>obi</i> II	‘to see’	<i>obi-θi</i>
<i>aha</i> II	‘to find’	<i>aha-θa</i>

In some cases, consonantal replacement may co-occur with suffixation:

(57) *deverbal nouns derived by both consonantal replacement and suffixation*

<i>verb</i>		<i>noun</i>
<i>d̥aka</i>	‘to take off’	<i>d̥ara-θa</i>
<i>bɛhɛ</i>	‘to go down’	<i>bɛrɛ-θɛ</i>

Action nouns function as obligatorily possessed nouns, taking the same series of prefixes described in Table 1 above. The possessor will correspond to the absolutive argument of the original verb (that is, the subject of an intransitive verb or the object of a transitive verb): *wa-r̥ibɛ* ‘my speech’, *wa-rira* ‘my walk’, *bɛ l-orɔ̃* ‘the water drought’, *d̥ak̥i θurɔ̃* ‘the washing of clothes’, *b̥adi d̥ɔ̃* ‘the eating of honey’, *ãdikura k̥ɔra* ‘manioc grating’, etc.

4.2.1 ‘Instrument’ and ‘subject’ nouns

Action nouns can take additional derivational morphology to form subject and instrument nouns: *rira-du* ‘the one who walks’, *rira-d̥ã* ‘the place where one walks, the instrument with which one walks’; *d̥ak̥i θurɔ̃-du* ‘the one who washes clothes’, *d̥ak̥i θurɔ̃-d̥ã* ‘instrument to wash clothes; soap’; *k̥ɔra-du* ‘the one who grates’, *k̥ɔra-d̥ã* ‘grater’, etc.

Although the suffixes *-da* and *-du* are traditionally described as ‘nominalizers’, they attach to any noun, including, as we have seen, *previously nominalized* verbs.

4.2.2 The suffix *-dĩ*

The suffix *-dĩ* ‘similar to’ is among the most productive derivational morphemes in Karajá, being used to derive, for instance, names of recently introduced animals (*brɔɾe-dĩ* ‘cow’, from *brɔɾe* ‘deer’, etc.). It may be etymologically related to the noun *dĩ* ‘noun’ and the transitive verb *dĩdĩ* ‘to name, to call’.

5. Descriptive words

A number of Macro-Jê languages—including Karajá—are traditionally described as presenting an active-stative verb agreement system to some degree. According to Fortune & Fortune (1964) and Maia (1998), Karajá verbs are divided into two different classes, active and stative, with the latter consisting essentially of predicates which denote adjectival meanings, the so-called ‘descriptive verbs.’ These predicates take a series of person markers which partially coincides with the pronominal object markers in transitive verbs, and which happens to be the same series of possessive prefixes occurring with nouns. A more careful examination, however, demonstrates that such ‘verbs’ are, in fact, nouns, and that transitivity, rather than stativity, plays a central role in Karajá grammar.

The case of Karajá (as well as a brief observation of other languages of the stock) suggests that one must be cautious in analyzing descriptive predicates as verbs in Macro-Jê

languages. In cases in which there are clear criteria to distinguish nouns and verbs, descriptive predicates tend to behave mostly like nouns. For example, in Karajá a verb must first be nominalized in order to appear in a descriptive predicate. Consequently, one must also be cautious in relying on the behavior of descriptive predicates as a sole criterion to propose active-stative systems.

Predicative and attributive adjectives are among the lesser known aspects of Macro-Jê languages. According to the descriptions available, in all Macro-Jê languages the ‘adjective’ would follow the noun (Rodrigues 1999: 193). Since most Macro-Jê languages are SOV, the position of the adjective in these languages would constitute a systematic counter-example to Greenberg’s implicational universals, according to which adjectives would tend to precede the modified noun in SOV languages (Greenberg 1966). However, since most such descriptions do not give clear evidence for the existence of adjectives as an independent part of speech, the apparently exceptional nature of adjectives in Macro-Jê may well be just another undesirable result of what one might call ‘translation-based linguistic analysis.’

That is certainly the case of Karajá, also traditionally described as having postposed adjectives (Maia 1998: 32). However, a careful analysis of the Karajá data clearly demonstrates that words denoting adjectival meanings in this language are abstract nouns whose behavior is identical to that of obligatorily possessed nouns—i.e., corresponding to attributive constructions such as *the angry man*, Karajá presents a genitive construction, *the man’s anger*. This implies that, instead of occurring as a modifier, the “adjectival” word would be, in fact, the head of the noun phrase, a phenomenon documented in languages such as Aleut (Sadock 2000), but commonly overlooked when considering Lowland South American languages.

This section examines the hypothesis that lexemes denoting adjectival meanings in Karajá (henceforth ‘descriptives’) are, in fact, nouns (generally, abstract nouns), be they in predicative or attributive use. That is, words corresponding to adjectives in languages such as English or Portuguese, such as ‘angry,’ ‘red,’ ‘short,’ would be better translated as abstract nouns—‘anger,’ ‘redness,’ and ‘shortness,’ respectively.

Descriptive predicates take exactly the same series of prefixes that occur with nouns (Table 2):

- (59) a. *wa-d-εbure=r-e*
1-REL-get.angry=CTFG-IMPERF
 ‘I am angry.’
- b. *Ø-εbure=d-e*
2-get.angry=2-IMPERF
 ‘You are angry.’
- d. *d-εbure=r-e*
3-get.angry=CTFG-IMPERF
 ‘He is angry.’
- (60) a. *wa-itʃǎdε=r-e*
1-get.crazy=CTFG-IMPERF
 ‘I am crazy.’
- b. *a-itʃǎdε=d-e*
2-get.crazy=2-IMPERF
 ‘You are crazy.’
- c. *i-itʃǎdε=r-e*
3-get.crazy=CTFG-IMPERF
 ‘S/he/it is crazy.’

The only property that descriptives share with verbs is the fact that both can occur as predicates and, therefore, appear with tense/aspect markers. As we have seen, these markers have been traditionally considered inflectional affixes. According to this view, any lexeme

occurring with these ‘affixes’ would be by definition a verb. However, a more careful analysis has shown that the tense/aspect markers are, in fact, clitics, and can therefore attach to any element occurring as predicate, including postpositional phrases, pronouns, and nouns (Ribeiro 1996). As such, they can attach to any element occurring as a predicate, including postpositional phrases (61), pronouns (62), and nouns (63).

- (61) wad̥au d̥əb̥ɔ̃ra=k̥u=r-e
 W. young=TEMP=CTFG-IMPERF
 ‘It was during Watau’s youth.’
- (62) d̥ikar̥ɔ̃=k̥əre
 I=FUT
 ‘It will be me.’
- (63) d̥zuhu=r̥ək̥ɪ h̥ur̥əθ̥ɔ̃=r̥ək̥ɪ id̥ɔ̃ h̥ɛɔ̃d̥ĩ=r̥-ɛd̥ɔ̃=r-e
 before=NARR lightning.bugs=NAR people fire=CTFG-PL=CTFG-IMPERF
 ‘It is said that, in the old times, lightning bugs were the fire of mankind.’

As the examples below further show, any noun can occur as a descriptive predicate, including not only more abstract nouns such as -*εθε* ‘pain’ and -*əd̥ɔ̃re* ‘fragrance, smell’, but also concrete nouns such as *dε* ‘flesh’, -*εδε* ‘thorn’, and *kub̥ɔ̃* ‘body’:

- (64) a. d̥z-u d̥-εθε
 REL-tooth REL-pain
 ‘tooth ache’
- b. wa-d̥z-u d̥-εθε=r̥-εr̥ɪ
 1-REL-tooth 3-pain=CTFG-PROGR
 ‘My tooth is aching.’
- (65) a. h̥ãd̥ike d̥ε
 chicken flesh
 ‘chicken’s meat’

- b. hãdike i-dɛ=r-e
 chicken 3-flesh=CTFG-IMPERF
 ‘The chicken is fat.’

Claiming that descriptives are, in fact, nouns, does not imply that they are always originally nouns. Indeed, a number of descriptives are derived from change-of-state verbs, such as *itfãdɛ* ‘to become crazy’ (66), *ɛbure* ‘to get angry’ (67), and *-udədə* ‘to become cold’ (68):

- (66) a. hãbu Ø-r-a-itfãdɛ=r-e
 man 3-CTFG-INTR-become.crazy=CTFG-IMPERF
 ‘The man got crazy.’
- b. hãbu itfãdɛ
 man crazy
 ‘crazy man; the man’s craze’
- (67) a. habu Ø-r-Ø-ɛbure=r-e
 man 3-CTFG-INTR-get.angry=CTFG-IMPERF
 ‘The man got angry.’
- b. habu d-ɛbure
 man REL-angry
 ‘angry man; man’s anger’
- (68) a. bədə Ø-r-Ø-udədə=r-a
 weather 3-CTFG-INTR-get.cold=CTFG-PERF
 ‘The weather got cold.’
- b. bədə tʃ-udədə=r-e
 weather 3-get.cold=CTFG-IMPERF
 ‘The weather is cold.’

Notice that the basic meaning of these verbs is not descriptive or stative, but processual: they denote the inchoative aspect of what in other languages such as English

would correspond to an adjective. Other descriptives, such as *kutfiɛ* ‘heavy’, are originally nouns (69).

- (69) a. wa-bɛhira Ø-r-a-kutfiɛ-dǝ=r-e
 1-basket 3-CTFG-INTR-heavy-VERB=CTFG-IMPERF
 ‘My basket got heavy.’
- b. wa-bɛhira i-kutfiɛ=r-e
 1-basket 3-heavy=CTFG-IMPERF
 ‘My basket is heavy.’

As we have seen, ‘descriptives’ behave morphologically as nouns. I propose that descriptives derived from verbs, such as in (66)-(68) above, are also nouns. Although this is not obvious in the examples above, since the primitive verbs and their noun forms happen to be homonymous, it becomes clear with verbs that are formally differentiated from their derived nouns. Such is the case with verbs such as *ɔka* ‘to make holes’, *kika* ‘to tear’, and *ukǝ* ‘to dry’, whose corresponding noun forms are derived by *consonantal replacement* (as we have seen, a common mechanism to derive nouns from verbs):

- (70) a. wa-bɛ-d-ɛkɔ Ø-r-Ø-ɔka=r-a
 1-water-REL-utensil 3-CTFG-INTR-make.holes=CTFG-PERF
 ‘My cup got a hole in it.’
- b. wa-bɛ-d-ɛkɔ d-ɔra=r-e
 1-water-REL-utensil 3-make.holes=CTFG-IMPERF
 ‘My cup has a hole in it.’
- (71) a. wa-d-õrɛ Ø-r-a-kika=r-a
 1-REL-shirt 3-CTFG-INTR-tear=CTFG-PERF
 ‘My shirt tore.’
- b. wa-d-õrɛ i-kira=r-e
 1-REL-shirt 3-tear=CTFG-IMPERF
 ‘My shirt is torn.’

- (72) a. bε Ø-r-Ø-ukǎ=r-εrɪ
 water 3-CTFG-INTR-dry=CTFG-PROGR
 ‘The water is drying.’
- b. ãhʊ d-urǎ=r-e
 lake 3-dry=CTFG-IMPERF
 ‘The lake is dry.’

Descriptives derived from verbs are generally translated as participles—that is, verbal adjectives. Notwithstanding this, I consider them to be verbal *nouns*, which is a corollary of the lack of distinction between nouns and adjectives in Karajá. As Haspelmath (1994, 152) states, “it is quite clear that not all languages have participles. The definition of participle (‘verb-derived adjective within a verbal paradigm’) apparently presupposes the existence of adjectives in a language. Since there are many languages that lack primary adjectives, at least these languages will also lack participles.”

5.1 ‘Semantic types’

As the examples above illustrate, concepts expressed as adjectives in other languages can be lexicalized either as nouns or as change-of-state verbs in Karajá. The assignment of a given concept to one or another part of speech seems to be semantically fairly consistent. Opposite pairs (‘antonyms’) generally fall into the same part of speech, what is probably related to the fact pointed out by Dixon (1977, 27) that “each semantic type has basic or ‘norm’ connection with a single part of speech.” Thus, while *dɔdǎkε* ‘to become hot’ and *udǎdǎ* ‘to become cold’ are both verbs, *-urʊ* ‘darkness’ and *-ufa* ‘clearness’ are both nouns. Similarly, all basic-color concepts are basically lexicalized as verbs (*θɔ* ‘to be made red’,

kura ‘to be made white’, *lobi* ‘to be made black’). The distribution of the different semantic types between nouns and verbs constitutes an interesting question to be investigated in the future.

5.2 Headship

The data analyzed here clearly demonstrate that words denoting adjectival concepts in Karajá are nouns—especially abstract nouns. These nouns present a morphological behavior identical to obligatorily possessed nouns—i.e., corresponding to attributive constructions such as the angry man, Karajá presents a possessive construction, the man’s anger. This would imply that, instead of occurring as a modifier, the “adjectival” word would be, in fact, the head of the noun phrase. Uncommon as this may seem, such a situation is reminiscent of what happens in Tupinambá and Aleut, for example.

In Tupinambá, a Tupí-Guaraní language once spoken along the Brazilian coast, Rodrigues (1996) also treats descriptive words in Tupinambá as nouns; “

Muitos nomes designam qualidades ou estados, tais como *oríβ* ‘alegria’, *-así* ‘dor’, *-ún* ‘pretume, escuridão’, *piráŋ* ‘vermelhidão’, *poráŋ* ‘beleza’, *kaneʔõ* ‘cansaço’, *maʔenwár* ‘lembrança’, etc. Como esses nomes são mais freqüentemente usados como predicados, eles têm sido muitas vezes considerado adjetivos ou verbos descritivos. Entretanto, morfológica e sintaticamente eles não diferem dos nomes possuíveis [...]. [Rodrigues 1996: 62-63]

(73) **Tupinambá** (Rodrigues 1996: 63)

- a. *né* *r-oríβ*
 you CNT-happiness
 ‘You are happy’ or ‘you have happiness.’
- b. *né* *r-uʔúβ*
 you CNT-arrow
 ‘You have arrows.’
- c. *né* *r-oríβ-a* *o-páβ*
 you CNT-happiness-ARG 3SUBJ-finish

‘Your happiness is finished.’

- d. *né* *r-uʔúβ-a* *o-páβ*
you CNT-arrow-ARG 3SUBJ-finish
‘Your arrows are finished.’

A similar situation is presented by the ‘relative-anaphoric’ construction in Aleut (Sadock 2000). Sadock describes it as a “highly favored structure in Aleut,” which is “deployed in a surprising array of formally and semantically distinct cases”, including *genitivization* (74), *nominal modification* (75), and *relativization* (76).⁴⁵

Aleut (Sadock 2000)

- (74) *tayaġu-m* *ula-a*
man-3/rel/s house-3/A/s
‘the man’s house’
- (75) *ula-m* *tagada-a*
house-3/rel/s be.new-3/A/s
‘the/a new house’
- (76) *hla-m* *aygag-na-a*
boy-3/rel/s walk-PART-3/A/s
‘a walking boy’

However, according to Sadock (*op. cit.*), in (41) and (42) “the notional modifier is the formal head of the phrase. [...] Despite the formal headship facts, the relative case NP (the counterpart of a possessor) is the semantic head.” Thus, the relative-anaphoric construction in Aleut presents a striking similarity with genitive-like constructions in Karajá (and in Jê, and in Tupí-Guaraní). The distribution of the relative case marker in Aleut in (39)-(41) is extremely similar to the distribution of relational prefixes in Macro-Jê.

⁴⁵ Relative case marks a syntactically dependent form; absolutive case marks an independent form.

6. Final remarks

This chapter described the morphology of Karajá, focusing especially on the inflectional and derivational properties of nouns and verbs. A thorough description of verb and noun morphology makes it possible to determine the morphosyntactic position of ‘descriptives’, i.e. lexemes that denote adjectival meanings. Contrary to previous analyses (Fortune 1973, Maia 1986), the description provided here demonstrates that descriptives are in fact nouns. A conclusive proof of their nominal nature is the fact that, in order to be used as a descriptive word, a verb must be first nominalized, which is particularly clear in cases for which nominalization is indicated by consonantal replacement.

Subordinate Clauses

1. Introduction

This chapter describes subordinate clauses in Karajá, including relative clauses, complement clauses, and (the functional equivalents of) adverbial clauses. Although this study is mainly concerned with subordinate clauses proper (that is, complements or modifiers which present the internal structure of a full clause), it will also briefly discuss nominalized complements, very common functional equivalents of subordinate clauses in Karajá.

The basic mechanism to signal subordination, as illustrated by relative and complement clauses in subject and direct object positions, is stress shift, an interesting example of non-concatenative morphology. As we have seen, a typical finite verb in Karajá presents tense-aspect clitics, which are unstressed. In independent clauses, the stress falls ordinarily on the last syllable of the verb stem (1). In relative clauses, however, the stress shifts to the tense-aspect clitic (2). The entire subordinate clause behaves as a noun phrase, and, as such, may take nominal clitics such as the indefinite article (3) ♀ =*dō* (♂ =*ō*) (and, as we will see, postpositions):

- (1) dori Ø-d-Ø-ɔɔ=d-e [dɔ'rode]
 White 3-CTPT-INTR-go.ashore=CTPT-IMPRF
 'The white man came ashore.'
- (2) ♀ dori Ø-d-Ø-ɔɔ=d-é [dɔrɔ'de]
 White 3-CTPT-INTR-go.ashore=CTPT-IMPRF+SUBORD

2). Exceptions are the postposition *=bã*, which takes the prefix *ʌ-* in the second person (the only occurrence of this prefix whatsoever), and the allative postposition *♀ =kɔ* (*♂ =ɔ*), which does not seem to occur with prefixes.¹ In addition, *=bã* ‘diffuse locative’, *=kɪ* ‘stationary locative’, and *=dəkɛ* ‘dative’ present suppletive pronominal forms for the third person (*dʊ*, *dai*, and *dabã*, respectively). Most postpositions are unstressed, with the exception of the dative *=dəkɛ* and the the evitative *♀ =laku* (*♂ =lau*).

Table 5.1. Karajá postpositions

<i>Postposition</i>	<i>Approximate meaning</i>	<i>3rd person suppletive pronouns</i>
<i>=kɪ</i>	‘stationary locative’	<i>dai</i>
<i>=tʃi</i>	‘dynamic locative’	
<i>=bã</i>	‘diffuse locative’	<i>dʊ</i>
<i>♀ =ku</i> (<i>♂ =u</i>)	‘temporal’	
<i>=rɔbɪ</i>	‘ablative’	
<i>=dɪ</i>	‘instrumental’	
<i>=dəkɛ</i> (<i>♂</i>	‘dative’	<i>dabã</i>
<i>=dɛɛ</i>)		
<i>=kɔ</i> (<i>♂ =ɔ</i>)	‘allative’	<i>dabã</i>
<i>♀ =laku</i> (<i>♂</i>	‘evitative’	
<i>=lau</i>)		
<i>=wɔdã</i>	‘comitative’	
<i>=wɔθɛ</i>	‘comparative’	

¹ Both the allative postposition *=kɔ* and the dative postposition *=dəkɛ* mark the recipient or the direction of a given process. Although both can occur with nouns, *=dəkɛ* seems to occur only with [+animate] nouns, while *=kɔ* can occur with both [+animate] and [-animate] nouns. Because of this semantic constraint, *=dəkɛ* occurs more frequently in functions that are traditionally associated with that of datives. As mentioned above, *kɔ* does not take prefixes. Thus, corresponding to both postpositions, there is only one set of inflected forms: *wa-dəkɛ* ‘to me’, *ã-dəkɛ* ‘to you’, *dabã* ‘to him’, *iθi-dəkɛ* ‘to himself’.

Table 5.2. Paradigms for some Karajá postpositions

	‘instrumental’	‘ablative’	‘diffuse locative’	‘stationary locative’	‘dative’
‘man’	<i>hābu=dɪ</i>	<i>hābu=rəbɪ</i>	<i>hābu=bə̃</i>	<i>hābu=kɪ</i>	<i>hābu=dəke</i>
1 st	<i>wa-dɪ</i>	<i>wa-rəbɪ</i>	<i>wa-bə̃</i>	<i>wa-kɪ</i>	<i>wa-dəke</i>
2 nd	<i>ã-dɪ</i>	<i>ã-rəbɪ</i>	<i>u-bə̃</i>	<i>ã-kɪ</i>	<i>ã-dəke</i>
3 rd	<i>i-dɪ</i>	<i>i-rəbɪ</i>	<i>du</i>	<i>dai</i>	<i>dabə̃</i>
REFL	<i>ifi-dɪ</i>	<i>ifi-rəbɪ</i>	<i>ifi-bə̃</i>	<i>ifi-kɪ</i>	<i>ifi-dəke</i>

Considering the semantic idiosyncrasies that generally characterize adpositions, naming them is only a first attempt towards defining their semantic properties.² Although the translation of some of the postpositions listed above is rather straightforward (vis-à-vis their approximate translation given in Table 1), a number of them require further explanation (either because of their highly polysemic nature, or because of the semantic subtleties that differentiate them). A brief explanation of the meanings most commonly associated with some of them shall be helpful in understanding their use with subordinate clauses.

The *evitative* postposition indicates what is to be avoided or feared, marking the oblique complements of verbs such as *-uberu* ‘to be frightened’ (4) and *wə* ‘to hide’.³ An example of the *temporal* postposition (‘when’) is given in (5) below. Besides indicating source (‘from’), the *ablative* postposition marks objects of comparison (6).

- (4) ♂ wa-riore dori=***lau*** Ø-r-Ø-uberu=r-eri
 1-child White=***EVIT*** 3-CTFG-INTR-be.frightened=CTFG=PROGR
 ‘My child is afraid of the White man.’

² A good example of the problem posed by the translation of adpositions is the preposition *by* in English, which has instrumental (*he went by train*), locative (*he is by the fountain*), and temporal (*I’ll be back by midnight*) uses, among others. The criteria to determine the ‘basic’ meaning of postpositions may be fairly arbitrary. In labeling the Karajá postpositions, I took into consideration what seems to be their more general, productive, and non-idiomatic uses (as far as possible).

³ Kariri, a Macro-Jê language remotely related to Karajá, also has an evitative postposition (Mamiani 1877:77). According to Blake (1994: 156), evitative (also called ‘aversive’) case markers are also common in Australian languages.

- (5) ♂ *wira=u* wa-hedo a-r-i-wi-dõ=kre
dry.season=*TEMP* 1-house 1-CTFG-TRANS-make-VERB=*FUT*
‘I will build my house *in the dry season.*’
- (6) waha i-rarie=r-e bu=*rõbi*
my.father 3-height=*CTFG-IMPRF* 2.father=*ABL*
‘My father is taller *than your father.*’

The *dynamic locative =tʃi* is used with figures in motion (7), while the *stationary locative =kɪ* occurs with static figures (8). While both *=tʃi* and *=kɪ* denote more precise paths, the *diffuse locative =bõ* denotes more widespread, less ‘pinpointable’ paths (10, 11).⁴

- (7) ♂ urĩthĩ bera=*tʃi* Ø-r-e-θe=r-e
Urisiri water=*LOC* 3-CTFG-INTR-fall=*CTFG-IMPRF*
‘Urisiri jumped [or fell] *into the water.*’
- (8) ♂ dækɪ θohodzi iwa=*kɪ* Ø-d-a-rɪ=d-e
he one jatobá=*LOC* 3-CTPT-INTR-leave=*CTPT-IMPRF*
‘He was left alone *on the jatobá tree.*’
- (9) ♀ dõre baderaku=*bõ* Ø-r-Ø-õi=bõhõ=r-e
parrot wetland=*LOC* 3-CTFG-stand.up=*HAB=CTFG-IMPRF*
‘The parrot lives *in the wetlands.*’
- (10) ♂ ɔwõru ruru=*bõ* bõdi Ø-r-i-ε=r-e
tree branch=*LOC* honey 3-CTFG-TRANS-look.for=*CTFG-IMPRF*
‘He searched for honey *in the tree’s branches.*’

Besides its more clearly locative uses, *=bõ* can be used to indicate destination (11), time (12) and manner (13, 14).⁵ The oblique complements of verbs such as *ask, say, call,*

⁴ These semantic characterizations are approximations. For the concepts of *figure* and *path* adopted here, see Talmy (1985:129), who defines figure as “the salient moving or stationary object in a motion event” and path as a category which “refers to the variety of paths followed, or sites occupied, by the Figure object.”

⁵ Although the range of meanings displayed by some of the Karajá pospositions may at first seem unusual, it is very easy to find parallels in other, well-known languages such as English or Portuguese. The use of locative postpositions with temporal purposes (as a metaphorical extension from *location in space* to *location in time*),

name, think and tell (that is, the complements denoting that which is said, heard, thought, asked, or told) are also marked by the postposition =*bǎ* (15). As we will see, this semantic diversity accounts for the use of this postposition to mark both purpose clauses and converbs, as well as sentential complements of verbs such as *say, tell, etc.*⁶

- (11) ♂ wa-oworu=*bǎ* a-r-Ø-a=kre
 1-garden=*LOC* 1-CTFG-INTR-go=FUT
 ‘I will go *to my garden.*’
- (12) ♂ waha biuraθɔ=*bǎ* kə-d-Ø-a=kre
 my.father tomorrow=*LOC* 3-CTPT-INTR-go=FUT
 ‘My father will come *tomorrow.*’
- (13) wi=*bǎ* Ø-r-Ø-obu-dǎ=r-a
 speed=*LOC* 3-CTFG-INTR-swim-VERB=CTFG-PERF
 ‘He swam *quickly.*’
- (14) ɖa-dɪ=le=*bǎ* Ø-r-Ø-a=r-ɛɪ
 3REFL-leg=EMPH=*LOC* 3-CTFG-INTR-go=CTFG-PROGR
 ‘He is coming *on foot* (lit. on his legs alone).’
- (15) ♂ wa-bɪɔwa boho kawidǎ=*bǎ*
 1-friend PL Kawina=*LOC*
 Ø-r-ɪ-wa-dɪdɪ=bǎhǎ=r-e
 3-CTFG-TR-1-name=HAB=CTFG-IMPRF
 ‘My friends call me *Kawina.*’

for example, is a rather common phenomenon. Even more idiosyncratic uses of postpositions (as non-canonical markers of what would be translated as direct objects in English) may also reflect cross-linguistic tendencies. As Jespersen (1965:159) points out, verbs for ‘to throw’ take objects marked as instrumentals in many languages, including Old Norse. The use of a locative postposition to mark the object of a verb such as ♀ *rakufi* ‘to eat’ (♂ *rofɪ*) is paralleled in English by the use of *on* in constructions such as *deer feed on grass* or *the castaway lived on coconuts for weeks*.

⁶ This apparent diversity of meanings could be due not to polysemy, but to homophony (that is, instead of these being instances of different meanings associated with a single postposition, it could be the case that there would be different postpositions with the same phonological shape). However, the fact that a postposition such as =*bǎ* presents the same irregular allomorphy in all the uses mentioned above seems to rule out the hypothesis of accidental similarity.

In addition to the paradigmatic peculiarities and wide range of semantic uses, the ‘diffuse locative’ postposition, *=bǎ*, presents some distributional peculiarities. As illustrated by (16) below, in which *=bǎ* occurs as a purpose marker, it can follow other postpositions (in this case, the ‘static locative’ *=ki*). [The example also illustrates the use of the dative postposition, *=dǎkε ~ =dεε*, with benefactive functions.]

- (16) wa-dεε a-ru=*ki*=*bǎ*
 1-DAT 2-sight=**LOC**=**LOC**
 ‘Watch out for me!’ (lit. ‘[It is] for [it to be] in your eyes for me.’)

Another characteristic which sets *=bǎ* apart from all the remaining postpositions is its position with relation to the contrastive morpheme *=le* ‘only’: while *=le* follows all the other postpositions, as illustrated below by the dative *=dǎkε* (17a), it precedes *=bǎ* (17b):

- (17) a. hǎbu=*dǎkε*=*le* Ø-r-r-ǒ=r-e
 man=**DAT**=*only* 3-CTFG-TRANS-give=CTFG-IMPERF
 ‘He gave only to the man.’
- b. hǎbu=*le*=*bǎ* Ø-r-Ø-obi=r-e
 man=*only*=**LOC** 3-CTFG-INTR-see=CTFG-IMPERF
 ‘He saw only the man.’

2.1 Extended intransitive verbs

A number of verbs which would be translated as transitives in English or Portuguese have objects marked by oblique postpositions in Karajá. For example, the verb *-obi* ‘to see’ takes objects marked by the locative postposition *=bǎ* (18); *-ehu* ‘to throw’ takes

complements marked by the instrumental postposition *=dɪ* (19); *-aha* ‘to find’ takes complements marked by the locative postpositions *=tʃi* or *=kɪ* (20).⁷ The verb ♀ *rakufi* ‘to eat’ (♂ *roʃi*), although marked by the prefix *r-*, which generally occurs with transitive verbs, takes objects marked by the postposition *=bɔ̃* (21).⁸

- (18) kai wa-**bɔ̃** dʌ-d-obi=d-e
 you 1-**LOC** 2-CTPT-INTR-see=2-IMPRF
 ‘You saw me.’
- (19) ♂ dʌriɔre=**dɪ** biu=rɔbi=hɪkɔ̃ Ø-r-Ø-ehu=r-e
 3.REFL-child=**INSTR** high=ABL=big 3-CTFG-INTR-throw=CTFG-IMPRF
 ‘He threw his child from a very high altitude.’
- (20) dəkɪ hāwa witʃira=**tʃi** Ø-r-Ø-aha=r-a
 he place different=**LOC** 3-CTFG-INTR-find=CTFG-PERF
 ‘He found a different place.’
- (21) ♂ ɔhā hɛka iʃɔdɔ=**bɔ̃** Ø-r-ɪ-roʃi=bɔ̃hɔ̃=r-e
 armadillo ASSRT worm=**LOC** 3-CTFG-TRANS-eat=HAB=CTFG-IMPRF
 ‘The armadillo eats worms.’

Although they occur with an ‘object,’ all these verbs behave as intransitives. As such, they do not inflect for voice and cannot incorporate their objects.

2.2 Temporal constructions

⁷ As in English or Portuguese, the verb *aha* ‘to find’ is also used with an ‘evaluative’ meaning (as in *I found the play annoying*). In this case, the evaluative phrase is marked by the postposition *=bɔ̃*.

- (i) hābu i-bɪda=**bɔ̃** dʌi Ø-r-Ø-aha=r-e
 man 3-bad=**LOC** 3.LOC 3-CTFG-INTR-find=CTFG-IMPRF
 ‘The man found it bad.’

⁸ Another verb which presents a similar irregular behavior is ♂ ‘to drink’.

The semantics of some postpositions can be further illustrated by their use in temporal constructions. In addition to the temporal postposition proper, *=ku*, locative *=bǝ* and instrumental *=dɪ* were also attested in temporal uses: *uku=bǝ* ‘at night’, *uku=dɪ* ‘at dawn’; *bikuraθɔ=ku* ‘in the morning’, *bikuraθɔ=bǝ* ‘tomorrow’; etc. The use of *=dɪ* ‘instrumental’ in *uku=dɪ* ‘at dawn’ (literally, ‘after night’) mirrors its use with subordinate clauses such as (22) below:

- (22) *kə-d-∅-ehɛbǝdǝ=kre=dɪ* *a-r-∅-a=kre*
 3-CTPT-INTR-arrive=FUT=*INSTR* 1-CTFG-go=FUT
 ‘After he arrives, I will go.’

3. Relative clauses

For the sake of terminological clarity, I will adopt here the basic terminology employed by Keenan (1985). Like Keenan, I regard relative clauses to be full NPs (a fact that is rather clear in Karajá, where relative clauses can take articles, as we have seen, and postpositions, as it will be shown below). As Keenan points out, a relative clause typically consists of a common noun (or a pronoun), the *domain noun*, optionally accompanied by determiners, and a restrictive clause: “Semantically the common noun determines a class of objects, which we shall call the domain of relativization, and the restrictive clause identifies a subset of the domain, those elements which satisfy the conditions given by the restrictive clause” (Keenan 1985: 142). Thus, in *dʊri dɔrɔ=dé* ‘a White man who arrived’, *dʊri* ‘White man’ is the domain noun, and *dɔrɔ=dé* ‘who arrived’ is the restrictive clause.

As we have seen, the only surface difference between the independent clause in (1) and its subordinate counterpart in (2) is in the location of the stress: [*dʊri dɔrɔde*] ‘the White

man came ashore’ versus [d'o'ri doro'de] ‘the White man *who* came ashore’. In fact, stress shift is the basic mechanism to mark the subordinate status of relative clauses that modify core NPs (that is, subjects and direct objects). That the subordinate clause occupies the position of an NP is further demonstrated by the fact that it can be followed by the enclitic indefinite article ♀ =*dō*, ♂ =*ō* (24), which attaches to the rightmost element of an NP (23). In addition, an NP modified by a subordinate clause can still act as a possessor, and, in these cases, the possessed noun will occur *after* the subordinate clause (25).⁹ Notice that the relative clause is postnominal, occurring after the modified noun. This fact, as we will see (Section 6), reflects the general distribution of modifiers in attributive constructions.

(23) ♀ d'ori=*dō* Ø-d-Ø-ɔrɔ=d-e
 White=*INDEF* 3-CTPT-INTR-go.ashore=CTPT-IMPRF
 ‘A white man came ashore.’

(24) ♀ d'ori Ø-d-Ø-ɔrɔ=d-e=*dō*
 White 3-CTPT-INTR-go.ashore=CTPT-IMPRF=*INDEF*

 wa-ritʃoko Ø-r-r-kōbra-dě=r-a
 1-clay.doll 3-CTFG-TRANS-buy-VERB=CTFG-PERF
 ‘A White man who came ashore bought my clay dolls’

(25) ♀ *d'ori* Ø-d-Ø-ɔrɔ=d-é ritʃɔɛ
 White 3-CTPT-INTR-go.ashore=CTPT-IMPRF+SUBRD child

⁹ Notice that example (37) is only grammatical if *riɔɛ* ‘child’, an obligatorily-possessed noun, is interpreted as being the head of the preceding construction.

wa-ritfoko Ø-r-ɪ-kōbra-dǝ=r-a
 1-clay.doll 3-CTFG-TRANS-buy-VERB=CTFG-PERF
 ‘The White man who came ashore’s child bought my clay dolls.’

Examples (2), (24), and (25) above illustrate restrictive clauses which modify domain nouns in subject position. The examples below illustrate relative clauses modifying the direct objects of the transitive verbs ♂ *urihi* ‘to try’ (♀ *kurihi*) and *rəθǝ* ‘to eat’:

(26) ♂ *bādvuri* *kva* Ø-d-Ø-udǝ=d-é
 old.man that¹⁰ 3-CTPT-INTR-sit=CTPT-IMPRF+SUBORD

b-Ø-ɪ-urihi=b-ɛdǝ=ke
 2-CTFG-TRANS-try=2-PLURAL=POT
 ‘Try that old man who is sitting there.’

(27) kai *rəθǝ-dǎ* *ǎ-d-ɪ-wi=d-é*
 you eat-INSTR 1-CTPT-TRANS-carry=CTFG-IMPRF+SUBORD

dǎ-d-ɪ-rəθǝ=d-e
 2-CTPT-TRANS-eat=2-IMPRF
 ‘You ate the food that I brought.’

In all the examples above, the domain noun occupies the position of core arguments of the verb (subject or direct object), which are not morphologically marked in Karajá. However, when the domain noun is an oblique NP, the postposition that marks such NPs will occur after the restrictive clause, cliticized to the verb (a further argument for the nominal status of the relative clause). Thus, while with core arguments stress shift may be the only surface indicator of subordination, in examples such as (28) the postposition can also be seen as an indicator of the subordinate nature of the clause. Notice that stress shift also occurs in

¹⁰ *kva* ‘that (distant from both the speaker and the addressee)’ is a demonstrative (not a relative pronoun, as the translation of the example may suggest). There are two other demonstratives in Karajá, *ka* ‘this’ and *kia* ‘that (close to the addressee)’ (see example 40).

those cases in which the subordinate verb is followed by an article or a postposition. In such cases, however, stress shift may be analyzed as being triggered by the article or the postposition (see 3.3 below). Therefore, stress shift will only be indicated in this paper in cases in which it is the only surface marker of subordination.

In (28) below, the relative clause modifies the object of the verb ♀ *rakufi* ‘to eat’ (♂ *rofi*), which is marked by the postposition =*bã*, as we have seen above (21). The postposition marking the object will occur after the relative clause. This is further illustrated by examples of relative clauses modifying the objects of the postpositions =*bã*, =*laku*, and =*kr*.

- (28) ♂ *kia=redã* Ø-r-Ø-ãi=r-eri=le=bã
 that=PL 3-CTFG-INTR-stand.up=CTFG-PROGR=only=LOC
- b-Ø-i-rofi=bã b-Ø-Ø-ã=b-e=køre
 2-CTFG-TRANS?-eat=LOC 2-CTFG-INTR-go=2-IMPERF=FUT
 ‘You will be eating *just those [animals] standing there.*’

- (29) kai hãbu Ø-r-a-idzøra-dã=r-eri=bã
 you man 3-CTFG-INTR-run=VERB=CTFG-PROGR=LOC

d-a-Ø-bi=d-a
 2-CTFG-INTR-see¹¹=2-PERF
 ‘You saw the man running.’

- (30) ♂ wa-riøre dori d-Ø-øro=d-e=lau
 1-child White 3-CTPT-INTR-go.ashore=CTFG-IMPRF=EVIT
- Ø-r-Ø-uberu=r-eri
 3-CTFG-INTR-be.freightened=CTFG-PROGR
 ‘My child is afraid of the White man who came ashore.’

¹¹ Intransitive class II verb stems such as *obi* ‘to see’ lose their initial vowels in the first and second persons of the realis mood: in the example above, the underlying form of the verb is *da-obi=d-a* [d’a’bida].

- (31) ♂ kai d̥a-d-a-hāwa-d̥ɔ̃=b̥h̥ɔ̃=d̥-e
 you 2-CTPT-INTR-place-VERB=HAB=2-IMPRF
 hed̥o r-ɛ-wi-d̥ɔ̃=r-e=ki
 house CTFG-1+TR-make-VERB=CTFG-IMPRF=LOC
 ‘You live in the house that I built.’

3.1 Internal versus external relative clauses

Another important distinction is between *internal* and *external* relative clauses, according to the location of the domain noun (whether inside or outside the relative clause, respectively; Keenan 1985:143). In all the examples given so far, the domain nouns are inside the relative clauses. However, examples in which the domain noun is outside the relative clause are also easily attested (32).

- (32) kai rəθɔ̃d̥ã
 you food
 d̥a-d-i-rəθɔ̃=d̥-e ã-d-i-wi=d̥-é
 2-CTPT-TR-eat=2-IMPRF 1-CTPT-TRANS-carry=CTPT-IMPRF+SUBRD
 ‘You ate the food that I brought.’

When an oblique domain noun is separated from the relative clause, notice that the postposition marking the oblique NP occurs twice—with both the domain noun and the subordinate clause. That is, the stranded relative clause agrees in case with the domain noun:

- (33) iwero=b̥ɔ̃ r-ɛ-õ=r-a Ø-ãd̥i
 porridge=LOC CTFG-1+TR=DRINK=CTFG-PERF 2-mother
 Ø-r-i-d̥-ora-d̥ɔ̃=r-e=b̥ɔ̃
 3-CTFG-TRANS-d̥-cook-VERB=CTFG-IMPRF=LOC
 ‘I drank the porridge that your mother cooked.’

3.2 Pronoun retaining strategies

In the previous sections, we concentrate on the syntactic properties of the domain nouns. In this section, attention will be given to the relativized position—that is, the position in the relative clause which is co-referential with the domain noun. In all the examples of relative clauses presented so far, the relativized position corresponds to a core argument of the subordinate clause—either the subject, as in (32), or the direct object, as in (33). As these examples show, there is no surface indication of which position is relativized, which is indicated by gapping—that is, the relativized NP is simply absent from the relative clause. This lack of explicit indication of which position is relativized may lead to ambiguity when one has a third-person transitive verb with a third-person object. Thus, in both examples below, the relativized position may be interpreted as being either the subject or the direct object of the subordinate clause:

- (34) ♂ hāwii=bō r-a-Ø-bi=r-a
 woman=LOC CTFG-1-INTR-see=CTFG-PERF

hābu Ø-r-ɪ-həɖɛ-dō=r-a=bō
 man 3-CTFG-TRANS-hit-VERB=CTFG-PERF=LOC
 ‘I saw the woman who hit the man’
 or ‘I saw the woman who the man hit.’

- (35) ♂ hāwii hābu Ø-r-ɪ-həɖɛ-dō=r-á
 woman man 3-CTFG-TRANS-hit-VERB=CTFG-PERF+SUBORD

Ø-r-Ø-ɛhəbō-dō=r-ɛɪ
 3-CTFG-INTR-arrive-VERB=CTFG-PROGR
 ‘The woman who hit the man is arriving’
 or ‘the man who hit the woman is arriving.’

Subject and direct objects are the positions which tend to be more easily relativizable cross-linguistically, when compared with other syntactic positions. The hierarchy of which positions are more likely to allow relativization is thus summarized by Keenan and Comrie (1977):

- (36) **Keenan and Comrie's (1977) Accessibility Hierarchy:**
 Subject > Direct Object > Indirect Object > Object of Adposition > Possessor

While subject and direct object relativized positions are not marked in Karajá, relativized positions occupying a lower position in the accessibility hierarchy are marked by a third-person pronominal form. That is, relativization of any syntactic position besides subject and direct object is accomplished through *pronoun retaining* strategies. This is the case for objects of all the postpositions (including the dative, which behaves as any other oblique NP in the language, and objects of comparison), as well as possessors. Below are examples of relativized locative (37), ablative (38), and genitive (39) positions:

- (37) kai hāwa=rəbɪ d-∅-∅-ɛhɛbǎ-dǎ=d-ɛrɪ
 you place=ABL 2-CTFG-INTR-arrive-VERB=2-PROGR

dai r-a-hāwa-dǎ=bǎhǎ=r-e=rəbɪ
 3.LOC CTFG-1+INTR-place-VERB=HAB=CTFG-IMPRF=ABL
 'You are arriving from the village *where I live.*'
 [Lit. 'You are arriving from the village [*that*] I live in it.']

- (38) dɪarǎ hāwa=kɪ r-a-hāwa-dǎ=bǎhǎ=r-e
 I place=LOC CTFG-1+INTR-place-VERB=HAB=CTFG-IMPRF

i-rəbɪ d-∅-∅-ɛhɛbǎ-dǎ=d-ɛrɪ=kɪ
 3-ABL 2-CTFG-INTR-arrive-VERB=2-PROGR=LOC
 'I live in the village *where you are coming from.*'

[Lit. ‘I live in the village *[that]* you are coming from it.’]

- (39) *dɪarə* *hãwii=bə* *r-a-∅-bi=r-a*
 I woman=LOC CTFG-1-INTR-see=CTFG-IMPRF

i-hãbu *∅-r-∅-urɔ=r-a=bə*
 3-man 3-CTFG-INTR-see=CTFG-PERF=LOC
 ‘I saw the woman *whose husband died.*’
 [Lit. ‘I saw the woman *[that]* her husband died.’]

The example below further illustrates a pronoun-retaining relative clause, involving the object of the verb *-ehu* ‘to throw, to shoot’, which is marked with the instrumental posposition *=di*. Notice that the clitic *=le* ‘only’, which generally attaches to the right edge of a noun phrase, occurs after the subordinate verb:

- (40) *biura=ɔ* *i-di* *∅-r-∅-ehu=r-a=le*
 east=AL 3-INSTR 3-CTFG-INTR-throw=CTFG-PERF=*only*

∅-r-a-bədzidə-dihɪ=bə *∅-r-∅-oi=r-e*
 3-CTFG-INTR-melt-strong=CONV 3-CTFG-INTR-lie.down=CTFG-IMPERF
 ‘Only the one [the wax-tipped arrow] he had shot towards east was melted.’

3.3 Possible origins of the ‘subordinating accent’

As we have seen, stress shift plays an important role in subordination in Karajá. This device—which we may call ‘subordinating accent’—is reminiscent of the ‘definite accent’ in Tonga, described by Anderson (1992: 212) as an example of non-affixal clitic. In this language, according to Anderson, “the normal location of stress is on (the syllable containing) the penultimate mora. A sort of definiteness is marked by a stress shift to the final mora of the entire NP.” In this case, the stress shift is explained as being the reflex of a demonstrative morpheme **a* which would have existed diachronically.

A similar state of affairs may very well have been the diachronic source of the ‘subordinating accent’ in Karajá. The shift could have been originally triggered by a subordinating morpheme of some sort,¹² which would have later disappeared, leaving stress shift as evidence of its diachronic existence.

Another possible origin would be a more general, delimiting use of stress (common in languages where stress is predictable), to mark the boundaries of the noun phrase. However, given the lack of comparative evidence, one can only speculate on the origins of such phenomenon.¹³

3.4 Previous descriptions

In spite of its extreme pervasiveness, the phenomenon of the ‘subordinating accent’ was not mentioned in previous accounts of Karajá grammar (Fortune & Fortune 1964, Fortune 1970, Fortune 1973, Maia 1998). In his typological study, Maia states that he was able to find “hypothetical relative constructions” only through direct elicitation, “similar structures not being detected in the texts analyzed” (Maia 1998:33).¹⁴ The “hypothetical” here relates to the fact that he could not identify any relative pronoun (or any subordinating morpheme, for that matter) in the constructions considered. Notwithstanding this, the examples of “hypothetical relative constructions” that he provides (reproduced below with

¹² This ‘subordinating’ morpheme could have been a nominal clitic, such as an article (which would thus pair with the indefinite article ♀ =*dõ*).

¹³ Synchronically, one could postulate the existence of a ‘zero clitic’ to account for the stress shift. Since this more abstract analysis does not necessarily present any analytical advantage, I will adopt a more concrete analysis in this paper, indicating the stress shift whenever necessary.

¹⁴ Although Maia’s statement seems to suggest that relative clauses are less common in texts than in elicited materials, that is certainly not the case. Most examples in this paper were indeed taken from texts, where subordination is extremely common. In fact, all the subordinate clause types here described are highly favored constructions in Karajá grammar.

my own transcription and morphological segmentation) very likely presented the stress shift illustrated above. Since Maia relies on the rather inaccurate orthography used in Karajá schools, the stress shift went unnoticed in his analysis.¹⁵

Maia's examples of "hypothetical relative constructions" reanalyzed
(after Maia 1998:33)¹⁶

- (41) weriri dɔlɔra Ø-d-ɪ-widɔ̃=d-é
basket Dolora 3-CTPT-TRANS-make=CTPT-IMPRF+SUBORD

ka=u awi=r-e
PROX=TEMP good=CTFG-IMPRF
'The basket Dolora made yesterday is beautiful.'

- (42) hābu wihɪ=dɪ Ø-r-Ø-ehu=r-á
man arrow=INSTR 3-CTFG-INTR-throw=CTFG-PERF+SUBORD

aθɔ̃dɪ-hĩkɔ̃=bɔ̃ Ø-r-Ø-obi=r-a
guariba-big=LOC 3-CTFG-INTR-see=CTFG-PERF
'The man who threw the arrow saw the big guariba monkey.'

Although stress shift as a subordinating device is not mentioned in the work of the SIL missionaries, it is indicated in their translation of the New Testament by the use of an accent on the aspectual clitic (as mentioned in Chapter 1). This procedure, however, is not adopted in the common orthography taught at the Karajá schools.

4. Complementation

¹⁵ The fact that Maia does not signal any pause between both sentences corroborates my analysis of these examples as cases of subordination, not juxtaposition of two independent clauses ('the man threw the arrow; he saw the guariba monkey'), which would be a rather odd construction in Karajá, given the lack of particles marking any type of relation between the last clause and the preceding one.

¹⁶ Maia's original transcriptions of these examples are (i) *weriri Dolora dewinade kau awire* and (ii) *Hābu whyh rehura asynihiky-my robina*. In addition to the overall shortcomings of Karajá orthography, Maia's transcription also presents some inaccuracies of its own. For example, the verb *ehu* 'to throw' takes objects marked with the instrumental postposition =*dɪ*, never a 'bare' object as Maia's transcription suggests. This and other inaccuracies are corrected in (41) and (42) above.

Complement clauses are those which occupy the position of a noun phrase, occurring as arguments of the verb or as objects of postpositions. Like relative clauses, complement clauses in Karajá are also characterized by stress shift. In (43) below, the subordinate clause is the subject of the main clause, whereas in (44), the subordinate clause is the object of the verb ♀ *keri*, ♂ *eri* ‘to know’.

- (43) a-riβe=r-e *idʒaʔθo* *kə-d-∅-ɔhɔdɔ=kəré*
 2-speak.NOM=CTFG-IMPERF Aruanã 3-CTPT-INTR-exit=FUT+SUBORD
 ‘You said (lit. it is your word) *that the Aruanãs are coming out.*’¹⁷

- (44) *dibõbo* *∅-r-a-kɔraru=r-é*
 how 3-CTFG-INTR-start=CTFG-IMPERF+SUBORD

r-ε-keri=kõ=r-e
 CTFG-1+TRANS-know=NEG=CTFG-IMPERF
 ‘I don’t know *how it starts.*’

With predicates of saying and thinking, such as *ask, say, call, name, think, and tell*, the complements denoting that which is said, told, heard, or thought are marked by the postposition =*bã*. Therefore, sentential complements of these verbs will also be marked by this postposition. The same happens to complements of the verb *obi* ‘to see’.

- (45) *hε* *kə-d-r-wi=d-εdɔ=kəre=bã*
 firewood 3-CTPT-TRANS-carry=CTPT-PL=FUT=LOC

∅-r-a-riβe=r-a
 3-CTFG-INTR-speak=CTFG-PERF
 ‘She said *that she would bring firewood.*’

- (46) i-θε *∅-r-r-d-əãraθi-dã=r-e*
 3-mother 3-CTFG-TRANS-3-ask.NOM-VERB=CTFG-IMPERF

¹⁷ *Aruanãs* are masked dancers who represent forest and underwater spirits. Aruanã festivals are among the most highly anticipated events in Karajá social life.

aðhebo *Ø-r-Ø-udəð-də=r-a=bə*
 what 3-CTFG-INTR-happen=CTFG-PERF=LOC
 ‘His mother asked him *what had happened.*’

- (47) *ifə=dəke* *Ø-r-Ø-iri=bəhə* *hāwa widira=tʃi*
 people=DAT 3-CTFG-INTR-yell=HABIT place different=LOC

Ø-r-Ø-aha=r-a=bə
 3-CTFG-INTR-find=CTFG-PERF=LOC
 ‘He yelled to the people *that he had found a different place.*’

- (48) *buhā* *ikoi* *kɨ=rəbɪ* *wekiribə*
 dolphin guys amidst=ABL young.man

Ø-r-ɪ-di=r-e=bə *ka-r-e-ləki=kəre*
 3-CTFG-TR-carry.ANIM=LOC 1-CTFG-INTR-tell=FUT
 ‘I will tell you [a story about] *the dolphin taking a young man away from his group.*’

- (49) *kai* *hābu* *Ø-r-a-idzəra-də=r-erɪ=bə*
 you man 3-CTFG-INTR-run=VERB=CTFG-PROGR

də-bi=d-a
 2-CTFG-INTR-see¹⁸=2-PERF
 ‘You saw the man running.’

A complement clause marked with =bə can be replaced by the third-person suppletive form *dʊ*:

- (50) *dʊ* *Ø-r-a-ribe=r-a*
 3.LOC 3-CTFG-INTR-speak=CTFG-PERF
 ‘He said thus.’

¹⁸ Notice that, in this example, the complement clause could also be translated as a relative clause headed by *hābu* ‘man’ (‘You saw the man who is/was running’).

As we will see in the following section, functional equivalents of adverbial clauses in Karajá can also be seen as complement clauses occupying the position of NPs in postpositional phrases.

5. Adverbial clauses

Karajá apparently has few, if any, true adverbial clauses, a fact that mirrors the overall lack of adverbs in the language.¹⁹ Corresponding to locative adverbs ('here', 'there', 'hither', 'thither') and temporal adverbs ('yesterday', 'tomorrow'), Karajá has demonstrative pronouns (*ka* 'proximal', *kəḁã* 'distal') followed by postpositions (*ka=ki* 'here'; ♀ *ka=ku*, ♂ *ka=u* 'yesterday'; ♀ *kəḁã=ku*, ♂ *kəḁã=u* 'the day before yesterday; the day after tomorrow'). Corresponding to manner adverbs ('quickly', 'swiftly,' etc.), Karajá has nouns followed by the postposition *=bə* (*wi=bə* 'quickly,' etc.). Thus, just as most adverbial functions are played by postpositional phrases, functions traditionally associated with adverbial clauses are played by 'postpositional clauses' in Karajá. Below are examples of allative, locative, and temporal clauses:

- (51) ♀ i-rəbɪ a-he=bə ã-d-Ø-ã=d-e=kɔ=le
 3-ABL 2-behind=LOC 1-CTPT-INTR-ir=CTPT-IMPRF=AL=EMPH
- b-Ø-ε-aka=bə b-Ø-ã=b-e=kre
 2-CTFG-INTR-ir=LOC 2-INTR-ir=2-IMPRF=FUT
 'You may keep on going *to where I followed you from.*'

¹⁹ One candidate for adverbhood would be *widzi* 'today, now'. The locative morphemes *dú*, *dabə*, and *dai* (which, as we have seen, are suppletive third-person forms of postpositions) could also be alternatively analyzed as adverbs. Even if these morphemes are considered as adverbs, this would be a closed class in Karajá.

- (52) ♀ dʌbǎ Ø-r-Ø-a=bǎhǎ=r-e=kɔ
 3.AL 3-CTFG-INTR-go=HAB=CTFG-IMPRF=AL
- dʌ=rǎkɪ Ø-r-ɪ-d-ɛhɛbǎ-dǎ=r-e
 so=QUOT 3-CTFG-TRANS?-3?-arrive-VERB=CTFG-IMPRF
 ‘They arrived, it is said, *where he frequently went.*’
- (53) ♂ ididɪ=tʃi Ø-r-ɔdʊɔ=r-a=kɪ dʌ=rǎkɪ
 ground=LOC 3-CTFG-go.up=CTFG-PERF=LOC so=QUOT
- ifǎ dʊ Ø-r-Ø-ǎi=r-e
 peccary 3.LOC 3-CTFG-INTR-stand.up=CTFG-IMPRF
 ‘*Where they came up to the surface*, it is said that there were peccaries there.’
- (54) ♂ d-ǎdɪ kǎ-d-ɛhɛbǎ-dǎ=kre=u
 REL-mother 3-CTPT-INTR-arrive-VERB=FUT=TEMP
- a-r-Ø-a=kre
 1-CTFG-INTR-go=FUT
 ‘I will go *when my mother arrives.*’

Notice that the adverbial-like subordinate clauses above are structurally identical to relative clauses which modify oblique domain nouns (as seen in Section 3 above), except for the lack of a domain noun. That is, the locative clauses above can be analyzed as headless relative clauses. This structural identity between relative clauses and adverbial clauses is not unusual. As Thompson & Longacre (1985:179) point out, adverbial clauses expressing time, location, and manner can commonly be paraphrased, in many languages, “with a relative clause with a generic and relatively semantically empty head noun: *time*, *place*, and *way/manner*, respectively;” such clauses have also in common the fact that they can be replaced by a single word, an adverb, such as *today* (time), *here* (location), and *quickly* (manner) in English and other languages with a clear class of adverbs.

In Karajá, however, not only temporal, locative, and manner adverbial notions are expressed through postpositional clauses. As we will see below, purpose, conditional, and reason adverbial clauses, as well as converbs (subordinate clauses denoting a wide range of circumstantial meanings), are also constructed with postpositions. Thus, (notional) adverbial clauses in Karajá can be seen as complement clauses occupying the NP node in a postpositional phrase, being structurally identical with both relative and complement clauses.

5.1 Purpose.

Purpose clauses are constructed with the postposition *=bǝ* attached to verbs in the future or potential. This is probably an extension of the use of this postposition to indicate destination (see 2.6 above).²⁰

- (55) i-hãwii=ɔ idʒoi a-r-a-di=kəre
 3-woman=AL guys 2-CTFG-2-carry.ANIM=FUT
- ihãwii=wədã b-∅-∅-oi=kəre=*bǝ*
 3-woman=COMIT 2-CTFG-INTR-lie.down=FUT=LOC
 ‘We will take you to his wife in order for you to marry her.’

5.1.1. Nominal imperatives

A common type of negative imperative is formed with an action noun followed by the postposition *=bǝ* and, optionally, either the future marker *=kre* or the potential marker *=ke*.²¹

²⁰ The use of destination adpositions to mark purpose clauses is commonly attested, as in English (*I came to see you*) and Portuguese (*eu vim para vê-lo*).

²¹ Notice that the future and potential markers are also optional in verbal imperatives.

The action noun may be an original noun, such as *hi* ‘cry’ (56), or a deverbal noun, such as *ad̄iθi* ‘(the action of) bringing down’ (57) and *rira* ‘walk’ (58):

(56) *hi=kō=bō=ke*
 cry=NEG=LOC=POT
 ‘Don’t cry!’

(57) ♂ *i-ad̄iθi=ō=bō=kre*
 3-bring.down.N=NEG=LOC=FUT
 ‘Don’t bring it down!’

(58) ♂ *kia=bō* *rira=ō=bō*
 there=LOC walk.NOM=NEG=LOC

bād̄id̄i *Ø-r-a-ru-hukō=kre=bō*
 peanut 3-CTFG-INTR-EYE-grow=FUT=LOC
 ‘Don’t walk over there, in order for the peanut plants grow.’

5.2 Reason

Subordinate clauses indicating the cause of the event denoted by the main clause are formed with the locative postposition *=ki*.

(59) ♀ *wa-eθoru* *wa-rikōre* *Ø-r-ɪ-həd̄ε-d̄ō=r-a=ki*
 1-younger.sister 1-child 3-TR-hit-VERB=CTFG-PERF=LOC

ā-d̄ake *i-di* *Ø-r-Ø-a=r-εri*
 2-DAT 3-INSTR 3-CTFG-INTR-go=CTFG-PROGR
 ‘Because my younger sister hit my son, I’m bringing him to you.’

As with NPs marked with the locative postposition *=ki*, a reason clause can be replaced with the third-person suppletive form the the postposition, *d̄ai*:

- (60) *i-wiu=rəbɪ wa-d-ðhōdɪ Ø-r-ʊθa=r-a,*
 3-song=ABL 1-REL-ear 3-CTFG-INTR-forget=CTFG-PERF
- daj=le dʊʊ ako a-r-a-wiu-də=ð=kəre*
 3.LOC=EMPH 3.LOC 2-to 1-CTFG-INTR-song-VERB=NEG=FUT
 ‘I forgot their song, that’s why I won’t sing it to you.’

5.3 Conditional

There are two ways of forming conditional clauses: with the locative postposition =*kɪ* attached to potential verbs (61) or with the comitative postposition =*wədā* (62, 63).

- (61) *b-Ø-efi-ɛlehi-də=kɛ=kɪ b-efi-ɛlehi-də=kɛ*
 2-CTFG-REFL-rest-VERB=POT=LOC 2-CTFG-REFL-rest-VERB=POT
 ‘If you want to rest, rest.’

- (62) *idə Ø-r-a-ɾɪ=r-e=wədā hɛka*
 people 3-CTFG-INTR-leave=CTFG-IMPRF=COM ASSRT
- idə=dəkɛ=ka kə-d-a-ribe=kəre*
 people=DAT=ASSRT 3-CTPT-INTR-speak=FUT
 ‘If anyone was left, do call for us.’

- (63) *ð l-ʊahɪ=bə b-Ø-ɪ-ð=kð=kre=wədā*
 REL-medicine=LOC 2-CTFG-TRANS-drink=NEG=FUT=COM
- dɪɛ=bə bə-d-Ø-ʊɾʊ=b-e=kre*
 always=LOC 2-CTPT-INTR-die=2-IMPRF=FUT
 ‘If you don’t take medicine, you’ll always be feverish.’

The difference between conditional constructions with =*kɛ=kɪ* and constructions with =*wədā* is not fully understood yet. An obvious difference, however, is in the fact that =*kɛ=kɪ* constructions are restricted to the irrealis mood, and are thus commonly used to signal

hypothetical conditionals such as (64) and (65), while *=wədã* constructions are common with habitual conditionals such as (66):

- (64) wira=**ke=kr**=d'a=he
 dry.season=**POT=LOC=ASSERT=EMPH**
 θõεbõ ədʊra a-r-ɪ-bõ=ke
 much fish 1-CTFG-TRANS-catch=**POT**
 'If it were the dry season, I would certainly catch lots of fish.'
- (65) dziarõ a-hãbu=**ke=kr** a-r-a-d'εɔ-dõ-õ-d'ihɪ=ke
 I 2-man=**POT=LOC** 1-CTFG-2-send-VERB-NEG-strong=**POT**
 'If I were your husband, I wouldn't boss you around.'
- (66) wari heka i-ruɖa Ø-r-Ø-a=r-e,
 maguari ASSERT 3-shy 3-CTFG-INTR-go=**CTFG-IMPERF**
 idã=bõ i-rehe=ɔ Ø-r-Ø-obi=bõhõ=r-e=**wədã**
 people 3-long=**AL** 3-CTFG-INTR-see=**HABIT=CTFG-IMPERF=COM**
 urile idã=lau Ø-r-Ø-uwɔ=bõhõ=r-e
 simply people=**EVIT** 3-CTFG-INTR-fly=**HABIT=CTFG-IMPERF**
 'The *maguari* bird is shy. If (or whenever) it sees a person far away, it flies away.'

5.4 Converbs

Converbs are by far the most common subordinate clauses in Karajá. Marked by the postposition *=bõ*, converbs indicate that the event denoted by the subordinate clause is somehow related to the event coded by the main clause, signaling a range of circumstantial meanings, including *manner* (67, 68), *temporal sequence* (69), *simultaneity* (70, 71), and relationships of *cause and effect* (72).

- (67) ♂ bãdʊari-hikõ Ø-r-a-rabu-dõ=bõ
 old.man-big 3-CTFG-INTR-bend-VERB=**LOC**

Ø-r-εʃi-ukɔ=r-e
 3-CTFG-REFL-lift=CTFG-IMPRF
 ‘The very old man got up [lit. lift himself up], *bending [his back]*.’

- (68) ♂ *bε=ɗɪ* *Ø-r-a-hilɔɪ=bɔ̃*
 water=INSTR 3-CTFG-INTR-vomit=CONV

irɔɗu *hεɔɗi* *Ø-r-ɪ-ɗ-ɔri-ɗɔ̃=r-e*
 animal fire 3-CTFG-TRANS-3?-extinguish-VERB=CTFG-IMPRF
 ‘*Vomiting water*, she extinguished the animal’s fire.’

- (69) ♂ *bāi-d-εθε* *Ø-r-ɪ-ɗa=bɔ̃* *bāi-d-εθε=ɗɪ*
 knife-REL-sharp 3-CTFG-TRANS=LOC knife-REL-sharp=INSTR

ɗai *Ø-r-Ø-ɔɗzi=r-e=rəkɪ=hi=ɗɔ̃*
 3.LOC 3-CTFG-INTR-stab=CTFG-IMPRF=QUOT=EMPH=SIMPATY
 ‘*Taking a dagger*, it is said that he stabbed him, the poor thing.’

- (70) ♂ *bɔ̃awa=ɗɪ* *Ø-r-Ø-ehu=bɔ̃=rəkɪ=hε*
 firearm=INSTR 3-CTFG-INTR-throw=LOC=QUOT=EMPH

idɔ̃ *bāhāɗu* *ɗabɔ̃* *Ø-r-a-ɔɾu-ɗɔ̃=r-e*
 people group 3.AL 3-CTFG-INTR-run-VERB=CTFG-IMPRF
 ‘*Firing their guns*, the Karajā ran after them, it is said.’

- (71) ♂ *ɗabɔ̃* *Ø-r-Ø-εhεbɔ̃-ɗɔ̃=bɔ̃*
 3.AL 3-CTFG-INTR-arrive-VERB=LOC

idɔ̃ *boho* *Ø-d-ɪ-ɔri=d-εɗɔ̃=kəre*
 people PL 3-CTPT-TRANS-fetch=CTPT=PL=FUT
 ‘*As soon as he gets there*, they will come for us.’²²

Haspelmath (1995:3) defines converb as “*a nonfinite verb form whose main function is to mark adverbial subordination*” [his emphasis] “Another way of putting it”, he adds, “is that converbs are verbal adverbs, just like participles are verbal adjectives.”²³ At first, from a

²² This is one of the few examples of a third-person centripetal verb in the realis that does not occur with the prefix *kə-*. Possible semantic motivations for the absence of the prefix are being investigated.

²³ Similar constructions have been called ‘absolute clauses’ by some authors. According to Thompson & Longacre (1985:200-203), absolute “is a cover term for a subordinate clause type in which the following conditions hold: (i) the clause is marked in some way as being subordinate; (ii) there is no explicit signal of the

- (74) hãwɔ=rəkɪ Ø-r-ɪ-wɪ-dǎ=r-ɛdǎ=r-e,
 canoe=QUOT 3-CTFG-TRANS-make-VERB=CTFG-PL=CTFG-IMPRF
- hãwɔ Ø-r-ɪ-wɪ-dǎ=r-ɛdǎ=bǎ d'a=rəkɪ
 canoe 3-CTFG-TRANS-make-VERB=CTFG-PL=LOC so=QUOT
- idʒǎ iraru=ɔ Ø-r-ɛ-a=r-e
 some north=ALL 3-CTFG-INTR-run=CTFG-IMPRF
 'They made canoes, it is said. *Having made canoes*, some went north.'

5.4.1 Auxiliary-like converb constructions

Converbs commonly occur in auxiliary-like constructions in which the final, tense-marked verb is a positional verb, such as *-oi* 'to lie down', *-odǎ* 'to sit down', and *-ǎi* 'to stand up'. As the example below illustrate, the result is a stative meaning (75a), as opposed to a processual meaning (75b):

- (75) a. d̥i̯ Ø-r-a-dʊu=bǎ
 clothes 3-CTFG-INTR-get.wet=CONV
- Ø-r-Ø-oi=r-ɛrɪ
 3-CTFG-INTR-lie.down=CTFG-PROGR
 'The clothes are wet.'
- b. d̥i̯ Ø-radʊurɪ
 clothes 3-CTFG-INTR-get.wet=CTFG-PROGR
 'The clothes are getting wet.'

Although such auxiliary-like constructions are, in principle, simply another case of converb subordination, there seems to be arguments to consider them as slightly more grammaticalized than the converb constructions illustrated above. Since both the subordinate and the main verbs share the same subject, they tend to form a contiguous, uninterrupted

construction. In addition, there is at least one clear case in which an auxiliary-like construction underwent further grammaticalization, resulting in a aspectual marker, as described in the following section.

5.4.2 A case of grammaticalization

In addition to the posture verbs listed above, another common auxiliary-like construction involves the stem *hã*, which can be translated as ‘to be’, which takes the nominal series of prefixes. Notice that the stem undergoes total harmony (or vowel copy) in the first (*wa-hã=r-e* ‘I am’) and second persons (*ã-hã=d-e* ‘you are’), but not in the third person (*i-hã=r-e* ‘he is’). As the examples below illustrated, the construction signals a contrastive present (‘is VERBing now, unlike before’)

- | | | | |
|------|----|---|--|
| (76) | a. | <i>rɛrɔmã</i>
r-a-Ø-rɔ=bã
CTFG-1-CTFG-TRANS-eat=CONV
‘I (now) eat it.’ | <i>wahãre</i>
wa-hã=r-e
1-be=CTFG-IMPERF |
| | b. | <i>dɛrɔmã</i>
dã-Ø-rɔ=bã
2-CTFG-TRANS-eat=CONV
‘You (now) eat it.’ | <i>ãhãde</i>
ã-hã=d-e
2-be=2-IMPERF |
| | c. | <i>rɪrɔmã</i>
Ø-r-rɔ=bã
3-CTFG-TRANS-eat=CONV
‘He (now) eats it.’ | <i>ihãre</i>
i-hã=r-e
3-be=CTFG-IMPERF |

Constructions such as (78) above seem to have been the source for the rather productive habitual construction illustrated by examples (79) below. Both constructions

involve exactly the same morphemes (a combination of a converb, marked by *=bã*, and the inflected root *hã*), but present different degrees of phonological integration: while the constructions above form two distinct phonological words, the ones below form one single phonological word. Furthermore, the constructions below present a higher degree of morphological fusion, as the combination of the converb marker *=bã* and the inflected verb results in the loss of phonological material. The results are *=wahã* ‘1st person’ (in which the converb marker is dropped), *=bãhã* ‘2nd person’ (in which the converb marker fuses with the inflection prefix of auxiliary-like root), and *=bãhã* ‘3rd person’ (in which the 3rd person marker of the auxiliary-like root is dropped):

- (77) a. *rɛrɔwahãre*
 r-a-Ø-I-rɔ=**wahã**=r-e
 CTFG-1-CTFG-TRANS-eat=**1.HABIT**=CTFG-IMPERF
 ‘I eat it (habitually).’
- b. *dɛrɔmãhãde*
 dɛa-Ø-I-rɔ=**bãhã**=d-e
 2-CTFG-TRANS-eat=**2.HABIT**=2-IMPERF
 ‘You eat it (habitually).’
- c. *rɪrɔmãhãre*
 Ø-r-I-rɔ=**bãhã**=r-e
 3-CTFG-TRANS-eat=**3.HABIT**=CTFG-IMPERF
 ‘He eats it (habitually).’

The constructions above are common to both the Northern Karajá and Xambioá dialects. In Southern Karajá (78) and Javaé,²⁵ on the other hand, grammaticalization went a step further, with the generalization of the third-person form, *=bãhã*, to all persons:

²⁵ The contrast in the degree of grammaticalization of the habitual marker demonstrates the descriptive usefulness of distinguishing between Southern and Northern Karajá, despite the fact that they are traditionally seen as one single dialect. Considering the traditional situation of contact between the Javaé and Southern

- (78) a. *rɛrɔwahāre*
 r-a-∅-I-rɔ=**bəhə**=r-e
 CTFG-1-CTFG-TRANS-eat=**HABIT**=CTFG-IMPERF
 ‘I eat it (habitually).’
- b. *dɛrɔmāhāde*
 d-a-∅-I-rɔ=**bəhə**=d-e
 2-CTFG-TRANS-eat=**HABIT**=2-IMPERF
 ‘You eat it (habitually).’
- c. *rɪrɔmāhāre*
 ∅-r-I-rɔ=**bəhə**=r-e
 3-CTFG-TRANS-eat=**HABIT**=CTFG-IMPERF
 ‘He eats it (habitually).’

The development of the habitual marker in Karajá provides an interesting case-study in which every step in the grammaticalization process can still be witnessed synchronically, both in terms of internal reconstruction and interdialectal comparison. It is a compelling example of the usefulness of a pan-dialectal analysis, even in cases in which the dialectal differences may, at first sight, be considered too shallow to yield any descriptive insight.

5.4.3 Converbs in traditional songs

While converb constructions generally require one finite verb, exceptions are commonly found in poetic language. In a particular poetic genre, the *werɔ wiu* ‘maraca songs’ (*werɔ* ‘maraca’, *wiu* ‘song’), entire songs may not contain a single finite verb, as illustrated by the example below:

Karajá speakers (cf. Chapter 1), the sharing of this peculiarity may be due to dialectal influences from Southern Karajá onto Javaé.

- (79) *d-ō* *rɔdɛ* *Ø-r-ɪ-ɛɪ=kɛ=bõ* *aõbo*
REL-penis insert 3-CTFG-TRANS-know=POT=**CONV** INTER
- idzõdzu* *itʃɛɛ-da=dɪ*
foreigner look-INSTR=INSTR
- idzi=dəke* *Ø-r-ɛ-dɛhɛ=bõ*
REFL=DAT Ø-CTFG-INTR-look=**CONV**
- ɔwɔɾu kre=dɪ* *Ø-r-edzi-we=bõ*
wood piece=INSTR Ø-CTFG-REFL-penetrate=**CONV**

‘Does she want to know [how] penis penetration [is]?’ With a foreigner’s looking device [mirror] she watches herself (and) with a piece of wood penetrates herself.’

Maraca songs are performed by masked male dancers during the traditional *Aruanã* festivals; therefore, the songs are always in male speech. In addition to the common use of non-finite sentences, another linguistic peculiarity of the *wɛɾu wiu* is the substitution of [dʒ] for everyday-language’s [j], as in *idzõdzu* for *ifõdzu* ‘foreigner’ and *idzi* for *ifi* ‘reflexive’. As mentioned before (cf. Chapter 3), maraca songs are used mostly as a tool for social criticism against behaviors seen as “un-Karajá”—such as, in the example above, female masturbation. The fact that *k*-preserving forms (*kre* ‘piece’, *ɔ ɛɛ*; *dəke* ‘dative’, *ɔ dɛɛ*) are found even in such a strictly-male form of speech further corroborates the analysis given in Chapter 3 concerning the behavior of class b words.

5.5 Postpositions as discourse connectors

Reflecting their use with adverbial clauses, third-person forms of adpositions are used as discourse connectors (corresponding to conjunctions in languages such as English and Portuguese): *dai* ‘therefore, because of that’, *idr* ‘then, after that’, etc.

6. *Dokuri*-clause

Besides the ‘postpositional clauses’ described in this chapter, a few other constructions can also be probably treated as functional equivalents of adverbial clauses. The most common among them is the *dokuri*-clause, characterized by the presence of the second-position clitic ♀ =*dokuri* (♂ =*dori*). Such clauses have an explicative function and can be translated using English *for* (80) or *because*.

- (80) ♂ *dio=he* *b-Ø-Ø-udã*,
 PERM=EMPH 2-CTFG-INTR-sit.down
- i-rehe=bã=dori* *Ø-r-Ø-a=bãhã=r-e*
 3-long=LOC=*dori* 3-CTFG-INTR-go=HAB=CTFG-IMPRF
 ‘You better stay, for I’m going far.’

Maia (2002:155) analyzes ♂ =*dori* as a ‘focus marker,’ which is probably a mistake: being a second-position clitic, it is not surprising that this morpheme occurs after fronted elements, but that is also the case of most discourse-oriented particles in the language (see Chapter 4). Fortune (1970) translates it (rightly, I think) as ‘because.’ As a second-position clitic, =*dokuri* attaches to the first constituent of the clause, regardless of its syntactic status. In (79) above, =*dokuri* attaches to a postpositional phrase. The examples below illustrate its occurrence with noun phrases (81) and a predicate (82):

- (81) ♂ *bə-d-a-lə=ke*, *biu=dori* *kə-d-Ø-ɔa=kəre*

2-CTPT-INTR-enter=POT rain=*dori* 3-CTPT-INTR-rain=FUT
 ‘Enter, because it is going to rain.’

- (82) ♂ *a-r-Ø-a=õ=kre,* *i-ɔhãru=r-e=dori*
 1-CTFG-INTR-go=NEG=FUT 3-danger=CTFG-IMPERF=*dori*
 ‘I won’t go, because it is dangerous.’

In fact, its distribution and phonological behavior suggest that *=dokuri* belongs in the large class of *speech-act particles*, that inform on the attitude of the speaker with relation to the utterance (indicating certainty, doubt, excitement, boredom, sympathy, surprise, assertiveness, resolve, etc.). As other speech act clitics, *=dokuri* is unstressed; notice that, unlike adpositions, *=dokuri* does not change the stress of the verb phrase to which it attaches; thus, *i-ɔhãru=r-e=dori* ‘because it is dangerous’ is pronounced [iɔhãruredori].

Analyzing *=dokuri* as a speech act clitic helps us to understand the differences between *dokuri*-clauses and the reason clauses described above: *dokuri*-clauses are probably examples of *speech act* adverbial clauses, whose function, as defined by Thompson & Longacre (1985:203) “is not to modify or qualify the main clause in any way, but to modify or qualify, as it were, the speech act which the speaker is performing in uttering the main clause.” In fact, most attitude markers can be seen as speech-act subordinators to a certain extent, since they anchor the utterance to a previous one; in an example such as (83) below, the attitude markers *=θɔɛ* ‘enthusiasm’ and *=θɔ̃* ‘repetition’ can only be interpreted in reference to a previous speech act (it was jokingly said as a warning to a child who insisted in climbing a tree against a parent’s repeated admonition):

- (83) bə-d-Ø-uru=kre=θɔɛ=θɔ̃
 2-CTPT-INTR-die=FUT=ENTH=REPET
 ‘[As I have warned you before,] you will die [deservedly].’

7. Nominalized complements

Corresponding roughly to the subordinate clauses described above, one commonly finds nominalized forms of the verbs playing the role of complements. Deverbal nouns may be preceded by a grammatical possessor which corresponds to the original absolutive argument of the verb—that is, to the subject of an intransitive verb, such as in *wa-rira* ‘my walking’ (from ♀ *rika* ~ *ritfa* ‘to walk’), or to the object of a transitive verb, such as in *kɔdʊdɪ bɔ̃dɔ̃* ‘turtle catching.’ (from *bɔ̃* ‘to catch’). These constructions occur in any of the functions performed by subordinate clauses (or by any NP, for that matter)—that is, as subjects (84), direct objects (85) or objects of postpositions (86-91):

(84) ♀ *irə dʌrəθa* Ø-r-ɪ-wa-d-ɛkiwəθɛ-dɔ̃=r-a
 maniocake.NOM 3-CTFG-TRANS-1-REL-get.tired-VERB=CTFG-PERF
 ‘Pulling manioc made me tired.’

(85) ♀ *dɪkarɔ̃ bɛ l-urɔ̃* ka-r-ɪ-rakɔ̃=kərə
 I water REL-dry.NOM 1-CTFG-TRANS-wait=FUT
 ‘I will wait until the water dries.’

(86) ♂ *dɪarɔ̃ ɔdʊdɪ bɔ̃dɔ̃=bɔ̃* a-r-Ø-a=kərə
 I turtle catch.NOM=LOC 1-CTFG-INTR-go=FUT
 ‘I’ll go catch turtles.’

(87) ♂ *idʒɔ̃rɛ l-urə=u* a-r-ɪ-dɔ̃=kərə
 porridge REL-cook.NOM=TEMP 1-CTFG-TRANS-eat=FUT
 ‘When the porridge is cooked, I’ll eat it.’

(88) ♂ *kədθəiwe ɔdʊrə rəhɔ̃rɛθɛ=lɛ=bɔ̃*
 Kɔ̃nyxiwe fish turn.NOM=EMPH=LOC

Ø-r-Ø-øhōdɪ-dɔ̃=r-e
 3-CTFG-INTR-ear-VERB=CTFG-IMPERF

‘Kÿnyxiwe just thought about transforming the fish.’

- (89) ♂ idʒoi *bɔla=dr* *d-ehu=dr*
 guys ball=INSTR REL-throw.NOM=INSTR

Ø-r-ε-dεhe=kəre
 3-CTFG-INTR-watch=FUT
 ‘The guys will watch the ball game.’²⁶

- (90) ♂ wi Ø-r-ɪ-dɔɛdɔ̃=r-e
 REC 3-CTFG-TRANS-prohibit=CTFG-IMPRF

wi *wihɪ=dr* *d-ehu=lau*
 REC arrow=INSTR REL-throw=EVIT
 ‘They prohibited each other from shooting arrows.’

- (91) ♂ *i-θira=bɔ̃* *d’a=hε*
 3-difficult=LOC ASSERT=EMPH

udʒaudʒa=bɔ̃ *l-obiʃi* Ø-r-Ø-a=r-e
 sariema=LOC REL-see.NOM 3-CTFG-INTR-go=CTFG-IMPERF
 ‘Seeing the sariema bird is difficult.’

In addition, the suffix *-dɔ* forms subject nouns (that is, nouns referring to the subject of the original verb), which can be translated as relative clauses (92). The suffix *-dɔ̃*, on the other hand, forms nouns which name instruments, places, or other circumstances related to the action denoted by the verb (*rəθɔ̃-dɔ̃* ‘food’ < *rəθɔ̃* ‘to eat’; *kɔra-dɔ̃* ‘shredder’ < *kɔka* ‘to shred’; *θurɔ-dɔ̃* ‘soap (or any washing device)’ < *θuhɔ* ‘to wash’).

- (92) *rəθɔ̃-dɔ̃* *widi-dɔ* Ø-r-Ø-ɔrɔ=r-εrɪ
 eat-INSTR carry.NOM-SUBJ 3-CTFG-INTR-go.ashore=CTFG-PROGR
 ‘The one who brings food is coming ashore.’

²⁶ Both *dεhe* ‘to watch, to look at’ and *ehu* ‘to throw’ take objects marked by the instrumental postposition *=dr* (corresponding to what would be direct objects in English or Portuguese). The ‘ball game’ [lit. ‘the throwing of the ball’] in the example above refers, naturally, to soccer.

Verb nouns can also be used as (notional) modifiers of other nouns, both as predicates (93a) or attributes (93b). Notice that constructions such as *dəki θurɔ* can be translated as ‘washed clothes’ or ‘the clothes’ washing’, being structurally identical to genitive phrases. Notice that the location of the relative clause mirrors the location of the modifier in simple attributive constructions.

- (93) a. *wa-dəki* *i-θurɔ=ɾ-e*
 1-clothes 3-wash.NOM=CTFG-IMPRF
 ‘My clothes are washed.’
- b. *dəki* *θurɔ=dɪ* *a-r-eθi-dəki-də=kəre*
 clothes wash.NOM=INSTR 1-CTFG-REFL-clothes-VERB=FUT
 ‘I will wear clean clothes.’

The differences between subordinate clauses proper and constructions involving nominalized verbs are rather straightforward. With subordinate clauses, the verb remains intact, preserving the same inflectional properties and argument structure as in independent clauses. The deverbal nouns discussed in this section, on the other hand, are typical nouns, in that only one argument may be present (as the possessor) and verbal inflectional categories (such as direction and voice) are absent.²⁷

That is, while subordination is clearly a syntactic operation, the creation of deverbal nouns is a matter of derivational morphology. As such, it is subject to the irregularities which generally characterize derivational processes. The examples above illustrate different devices to derive nouns from verbs (cf. Chapter 4): (i) *ura* ‘cooking’ (from *uka* ‘to be cooked’), *l-urə* ‘drying’ (from *ukə* ‘to dry’), and *θurɔ* ‘washing’ (from *θuhɔ* ‘to wash’) are

²⁷ For example, while the verb *wi* ‘to carry’ can occur in both centripetal and centrifugal directions (*dɪwi=de* ‘he brought it’ versus *ɾɪwi=re* ‘he took it away’), this distinction is neutralized in the nominal form, in such a way that *ɾəθəɔdā wɪdɪ-du* ‘the one who carries food’ can be translated as either ‘the one who brings food’ or ‘the one who takes food away’.

derived through *consonantal replacement*, a process by which a velar stop or a glottal fricative in the last syllable of the verb root is replaced with the alveolar flap /r/ in the corresponding nominal form; (ii) *b̃d̃* ‘catching’ (from *b̃* ‘to catch’) is derived by the suffixation of *-dV*; (iii) ♂ *rah̃r̃ε̃θ̃ε̃*, ♀ *rah̃ak̃r̃ε̃θ̃ε̃* ‘turning, transforming’ (from ♀ *rah̃āk̃r̃ε̃*, ♂ *rah̃r̃ε̃* ‘to turn; to transform’ ‘to turn, to transform’) is derived by the suffixation of *-θV*; *d̃arãθ̃a* ‘taking’ (from ♀ *d̃aka*, ♂ *d̃a* ‘to take’) illustrates both consonantal replacement and the suffixation of *-θV*;²⁸ (iv) *d̃-ehu* ‘throwing’ (from *ehu* ‘to throw’) is derived by conversion (or zero-affixation). In addition, a number of verbs have suppletive noun counterparts (cf. *ud̃* ‘to sit down’, *r̃i* ‘sitting’; *ōr̃ō* ‘to sleep’, *d̃āhi* ‘sleep’; *oi* ‘to lie down’, *heri* ‘lying down’; etc.).

8. Final remarks

Karajá presents a pervasive and straightforward mechanism to create subordinate clauses, shared by relative, complement, and ‘adverbial’ clauses alike. This device is *clause nominalization*: an entire clause is nominalized and, as such, can occur in all the positions traditionally occupied by noun phrases (including subject and direct objects, objects of postpositions, and noun modifiers). Clause nominalization (a syntactic process) differs from *lexical nominalization* (a morphological process) in a number of ways, the main one being that the former preserves all the inflectional properties of the verb. With subordinate clauses occurring in the position of a bare nominal (that is, a noun phrase not followed by

²⁸ Notice that in both nominalizing suffixes, *-dV* and *-θV*, ‘V’ is a copy of the last vowel in the verb stem.

postpositions or an article), the subordinate status of the clause is indicated by *stress shift*, an interesting case of non-concatenative morphology which went unnoticed in previous descriptions of Karajá grammar.

CHAPTER 6

Karajá as a Macro-Jê language

1. Introduction

This chapter provides an appraisal of the evidence for the inclusion of Karajá in the Macro-Jê stock, taking into consideration additional evidence uncovered by the present study, informed by recent improvements in the reconstruction of Proto-Jê (Ribeiro 2005). A relationship between Karajá and the Jê family (which forms the core of the Macro-Jê stock) was first proposed by Karl von den Steinen (1886), but the evidence presented then was far from convincing, as criticized by Ehrenreich (1894). Later classifications would consider Karajá as an isolate (Loukotka 1968, Mason 1950, McQuown 1955), until Irvine Davis (1968), using the standard historical-comparative method, compared Karajá and Maxakalí data with his own reconstructed Proto-Jê forms (Davis 1966), detecting a number of phonological and lexical correspondences. All subsequent classifications of South American languages (Rodrigues 1970, 1986, 1999; Greenberg 1987, Kaufman 1994) agree in including Karajá into the Macro-Jê stock. As its late inclusion in the stock suggests, the similarities between Karajá and other Macro-Jê languages are far from obvious. The evidence presented by Davis, however, is being further corroborated by additional data, as this chapter hopefully demonstrates.

2. The ‘Macro-Jê hypothesis’

The Macro-Jê stock comprises the Jê family and a number of possibly related language families, most of which are located in Brazil (the exception being Chiquitano, which is also spoken in Bolivia). Macro-Jê is arguably one of the lesser-known language

groups of South America, its very existence as a genetic unit being still “a working hypothesis” (Rodrigues 1999: 165). According to Rodrigues (1986, 1999), whose classification has been the most widely accepted among researchers working on Brazilian languages, the ‘Macro-Jê hypothesis’ comprises 12 different language families: Jê, Kamakã, Maxakalí, Krenák, Purí, Karirí, Yatê, Karajá, Ofayé, Boróro, Guató, and Rikbaktsa. The existence of Jê as a language family has been recognized since early classifications of South American languages (Martius, 1867). ‘Jê’ is a Portuguese spelling for a Northern Jê collective morpheme ([je] in Apinajé, for instance) that occurs in the names of several Jê-speaking peoples. The term ‘Macro-Jê’ was coined by Mason (1950), replacing earlier labels, such as ‘Tapuya’ and ‘Tapuya-Je.’

Recent classifications (Rodrigues, 1986; Greenberg, 1987; Kaufman, 1994) differ as to the precise scope of Macro-Jê, although there is agreement on the inclusion of most of the families (Table 1). Except for Karirí (included only by Rodrigues), Greenberg and Kaufman included all the families listed above. In addition, Greenberg included Chiquitano (also included by Kaufman), Jabutí, and Otí. Given the lack of comprehensive comparative studies, the Macro-Jê status of some of these families is still an open question. Although Guató is included in the stock by all of the aforementioned classifications, there is no convincing evidence for its inclusion. While the evidence presented by Greenberg (1987) and Rodrigues (1986, 1999) was rather superficial and inconclusive, a recent study by one of Rodrigues’ students (Martins 2011), purporting to present additional evidence based on a perusal of the entire corpus available of Guató, fails to provide any convincing new arguments. On the other hand, recent studies have revealed compelling additional evidence for the inclusion of Chiquitano and the Jabutí family into the Macro-Jê stock. Thus, based on

such recent developments, my own classification includes all the families included by Greenberg, except for Guató and Otí, a poorly documented language once spoken in southern Brazil, whose meager available data do not support its inclusion in the Macro-Jê stock (or any family, for that matter). Like Rodrigues, I also include the Karirí family, based especially on particularly-suggestive grammatical evidence (Rodrigues 1992, Ribeiro 2002).

Table 6.1. The Macro-Jê Hypothesis

1.	Jê Amazonian Jê †Jeikó Northern Jê: Panará, Suyá, Kayapó, Apinajé, Timbira (Parkatêjê, Pykobjê, Krahô, etc.) Central Jê: Xavánte, Xerénte, †Akroá-Mirim, †Xakriabá Southern Jê Kaingáng, Xokléng, †Ingáin
2.	Kamakã †Kamakã, †Mongoyó, †Menién, †Kotoxó, †Masakarã
3.	Maxakalí Maxakalí, †Pataxó, †Kapoxó, †Monoxó, †Makoní, †Malalí
4.	Borum (Krenák, Botocudo) Borum
5.	Purí (Coroado) †Purí, †Coroado, †Koropó
6.	Ofayé Ofayé
7.	Rikbaktsá Rikbaktsá
8.	Boróro Boróro, †Umutína, †Otuquê
9.	Karajá Karajá
10.	Jabutí Djeoromitxí, Arikapú
11.	Chiquitano Chiquitano (Bésiro)
12.	Karirí †Kipeá, †Dzubukuá, †Pedra Branca, †Sabuyá
13.	Yatê Yatê
14.	Guató Guató

	(although the inclusion of Guató is agreed upon by all major classifications, there is simply no compelling evidence of genetic relationship with members of the Macro-Jê stock)
14	Otí †Otí (Eo-Xavánte) (the inclusion of Otí, proposed only by Greenberg, is not substantiated by the available data)
* Extinct languages are indicated by †. Based on Greenberg (1987), Rodrigues (1970, 1986, 1999), Kaufman (1994), Adelaar (2005), Ribeiro & van der Voort (2010), and Ribeiro (2011).	

There are three family-level reconstructions available: Davis (1966) and Ribeiro (2005), for Proto-Jê, and van der Voort (2009), for Jabutí. So far, lexical comparative evidence supporting the inclusion of individual families in the Macro-Jê stock has been presented for Kamakã (Loukotka 1932), Maxakalí (Loukotka 1931, 1939; Davis 1968), Purí (Loukotka 1937), Boróro (Guérios 1939), Krenák (Loukotka 1955; Seki 2002), Karajá (Davis 1968), Ofayé (Gudschinsky 1971), Rikbaktsá (Boswood 1973), Chiquitano (Adelaar 2005) and Jabutí (van der Voort and Ribeiro 2009). In addition, some studies have shown very suggestive cases of morphological idiosyncrasies shared by Jê, Boróro, Maxakalí, Karirí, Karajá, and Ofayé (Rodrigues, 1992, 2000b); some of such idiosyncrasies, namely the existence of ‘linking prefixes’, are also found in Jabutí and Chiquitano (Ribeiro 2011). Thus, although the inclusion of many of the families into the Macro-Jê stock is being further corroborated by additional research, for others (namely Iatê and Boróro) the hypothesis has yet to be systematically tested. The precise relationship among the suggested members of the stock also remains to be worked out. Some of the families may constitute subgroups within the stock. For instance, Borum (Krenák or Botocudo) and the Maxakalí family are closely-related, and both (plus, possibly, the Purí and Kamakã families) are more obviously related to Jê than Karajá is.

Besides the scarcity of data on some of its now-extinct members, such as the languages belonging to the Purí and Kamakã families, Macro-Jê comparative studies run into problems that typically plague long-range comparativists: roots are generally isolating and monosyllabic, and syllabic patterns are rather simple. Since most likely cognates tend to be of the CV type, there is a greater risk of accidental similarities. That risk is particularly increased by an apparent neglect to abide by the comparative method when proposing cognates. A case in point is the work of Aryon Rodrigues, a major Brazilian scholar, on comparative Macro-Jê linguistics. Besides adding little to the cognate sets already identified by previous authors (Guérios, Loukotka), Rodrigues' cognate inventories (Rodrigues 1999, Rodrigues & Cabral 2007) illustrate primary methodological mistakes, including an obvious case of borrowing (the word for 'maize', which is likely of Arawák origin; cf. Ribeiro 2010) and assigning cognate words to wrong cognate sets--for instance, Kaingáng *krẽ* 'egg' is considered a cognate of the word for 'egg' in other Jê languages, which trace back to Proto-Jê **ɲrɛ*, even though (as Davis had already demonstrated) Kaingáng preserved Proto-Jê **ɲ* as /*ɲ*/; Kaingáng *krẽ* traces back to Proto-Jê *kra* 'offspring' instead. As examples such as this illustrate, little attention is given to phonological regularity, even in the case of a fairly well-established family as Jê. Rodrigues' main contribution to comparative Macro-Jê studies was the detection of shared morphological similarities which are likely due to common inheritance (Rodrigues 1992, 2000b; Ribeiro 2011). Particularly relevant for Karajá are the so-called "relational prefixes", linking morphemes discussed in Section 3.2 below (cf. Chapter 4 as well).

While Macro-Jê comparative linguistics has been marked by an overall lack of methodological soundness, resulting in questionable proposals of genetic relationship (as

illustrated by the inclusion of Guató), it has attracted, on the other hand, unfounded pessimism. A case in point is Dixon & Aikhenvald's (1999:18) claim that the similarities among the several families within Macro-Jê would be due to areal, rather than genetic, factors:

“Rather than all the Macro-Je families being related in a higher-level tree, it seems to us that they could constitute a long-term linguistic area; this would account for their considerable typological similarities. (There may, of course, be genetic links between *some* of the established families, within the linguistic area.)”

Such suggestion reveals a lack of familiarity with the evidence put forth by authors such as Guérios, Loukotka, Davis, Gudschinsky, Boswood, and Rodrigues—mostly items of basic vocabulary, such as body-part terms, and grammatical morphemes. Although not as robust as one would wish (due in part to the temporal depth involved), such evidence points to genetic, rather than areal, relationship.²⁹ Furthermore, the typological similarities alluded to by Dixon & Aikhenvald are common in lowland South America, and thus do not set Macro-Jê languages apart as a group.

Macro-Jê is, to a certain extent, reminiscent of North America's Algic, in the sense that, in both cases, there is a close-knit, well-established family (Algoquian; Jê, in our case) and a number of distantly-related families or isolates (Wiyot and Yurok; Karajá, Ofayé, and several others, in our case). A major difference is that there is even more internal diversity in Macro-Jê, suggesting that it is even older than Algic. If, as Goddard (1991:64) states, “Algonquian, Wiyot, and Yurok form a genetic grouping that is at, or very close to, the maximum depth at which it is possible to reconstruct features of the protolanguage that retain any appreciable degree of resolution,” Macro-Jê may provide an even fuzzier picture. For

²⁹ Similar objections against Dixon & Aikhenvald's claims are raised by Voort (2004:213).

most of the stock, however, regular phonological and grammatical correspondences are still being uncovered, but one has to keep in mind the limitations mentioned by Goddard (*op. cit.*) for Algic:

“[...] it is methodologically instructive to note that this is a case in which the comparative method produces a proof of genetic relationship without there being a reconstructed phonology or phonological history beyond what is implied by a handful of equations of identity, or near identity.”

3. Karajá as Macro-Jê: from von den Steinen (1886) to Davis (1967)

The first author to suggest that Karajá was related to the Jê languages was Karl von den Steinen (1886), who classified Karajá as part of his “Tapuya-Stämme” (in addition to Northern and Central Jê languages and languages which are nowadays grouped as part of the Borum (Botocudo or Krenák), Kamakã, and Maxakalí families). Ehrenreich (1894) takes issue especially with von den Steinen’s reliance on pronominal similarities, but some of the lexical similarities pointed out by von den Steinen turn out to be actual cognates, including the first-person prefix *wa-* and two body-part terms, ‘tooth’ and ‘tongue’. Nonetheless, Karajá remained as an unclassified language for most of the 20th century, until Irvine Davis (1968), based on data collected by the Fortunes, undertook a comparison among Karajá, Maxakalí, and his own reconstruction of Proto-Jê (Davis 1966), detecting a number of phonological correspondences in basic vocabulary (mostly body-part terms and verbs, but also two pronominal morphemes; Table 2). Davis relies on the diagnostic value of Swadesh’s list to propose the inclusion of Karajá into the Macro-Jê stock:

“Maxakalí and Karajá are included in the same stock with the Jê languages on the basis of the fact that regular sound correspondences are detectable in a relatively small corpus of data and on the basis of lexical similarity. Lexicostatistical comparisons based on the Swadesh 100–word list show about 25% shared cognates between Maxakalí or Karajá and individual Jê languages.”

3.1 Lexical correspondences

Although Davis’ reconstruction requires, as I have shown elsewhere (Ribeiro 2005), a thorough revision, most of the phonological correspondences he detected between Jê and Karajá (Table 2) are being further corroborated by additional evidence (Table 3).

Table 6.2. Some lexical cognates between Proto-Jê and Karajá (apud Davis 1968)

Proto-Jê	Karajá	
* <i>a-</i>	<i>a-</i>	‘2 nd person’
* <i>j-ua</i>	<i>dʒ-u</i>	‘tooth’
* <i>so</i>	<i>dɔ</i>	‘to suck’
* <i>krã</i>	<i>ra</i>	‘head’
* <i>ku</i>	<i>kɨ</i>	‘to eat’
* <i>ma</i>	<i>ba</i>	‘liver’
* <i>j-ĩ</i>	<i>dɛ</i>	‘meat’
* <i>j-ĩja</i>	<i>d-ɛa(θɔ̃)</i> <i>d-ɛadɪ</i> ‘nasal bridge’	‘nose’
* <i>j-õtɔ</i>	<i>d-ɔrɔ(dɔ)</i>	‘tongue’
* <i>j-ã</i>	<i>(u)dɔ̃</i>	‘to sit’
* <i>ɲɛ</i>	<i>θɪ</i>	‘egg’
* <i>par</i>	<i>wa</i>	‘pé’
* <i>wa</i> ~ * <i>pa</i>	<i>wa-</i>	‘1 st person’
* <i>pɪ</i>	<i>ɾɨ</i>	‘path’
* <i>sɨ</i>	<i>dɪ</i>	‘bone’
* <i>sɨ</i>	<i>dɨ</i>	‘seed’

Proto-Jê	Karajá	
* <i>mrɔ(k)</i>	<i>bɪbɪ</i>	‘ashes’
Table 6.3. Additional lexical cognates		
Proto-Jê	Karajá	
* <i>rɛ</i>	<i>rɪ</i>	‘to leave (trans.)’
* <i>sɪ</i>	<i>dʃɪ</i>	‘to weave’
* <i>ŋrɛ</i>	<i>ʊ-θɪ</i>	‘to dance’
* <i>kra</i>	<i>ra</i> ‘nephew’	‘child, offspring’
* <i>ko</i>	<i>kɔ</i>	‘wood’
* <i>j-are</i>	<i>l-adzɪkura</i> ‘manioc’ (cf. <i>kura</i> ‘white’)	‘root’
* <i>mã</i>	<i>bɔ̃</i>	‘allative, dative, locative postposition’
* <i>we</i>	<i>obi</i>	‘to see’
* <i>wẽ</i>	<i>rɪbɛ</i>	‘to speak’
* <i>prãm</i>	<i>rəma</i>	‘hunger’
* <i>j-am</i>	<i>l-əba</i>	‘to stand up’
* <i>s-um</i>	<i>dəbɪ</i>	‘his father’
* <i>j-ɪ(j)</i>	<i>l-ədɪ</i>	‘to lay (s.t.) down’
* <i>tɪk</i>	<i>rʊ ~ rəbʊ</i>	‘to die’
* <i>kok</i>	<i>kəbɔ</i>	‘wind’

Some phonological correspondences are pretty well-attested, providing the basis for the detection, on solid grounds, of grammatical correspondences as well (see Section 3.2):

- (1) **Proto-Jê *m :: Karajá /b/**
- a. **ma* *ba* ‘liver’
 - b. **mrɔ(k)* *bɪbɪ* ‘ashes’
 - c. **mã* *bɔ̃* ‘locative, allative postposition’
 - d. **prãm* *rəba* ‘hunger’
 - e. **j-am* *l-əba* ‘to stand up’
 - f. **j-um* *dəbɪ* ‘father’

- (2) **Proto-Jê *k :: Karajá /k/**
- | | | | |
|----|------------|-----------|--------------|
| a. | <i>*ku</i> | <i>kĩ</i> | ‘to eat’ |
| b. | <i>*ko</i> | <i>kɔ</i> | ‘wood; horn’ |
- (3) **Proto-Jê *s :: Karajá /d/**
- | | | | |
|----|------------|-----------|------------|
| a. | <i>*sĩ</i> | <i>dĩ</i> | ‘bone’ |
| b. | <i>*so</i> | <i>dɔ</i> | ‘to suck’ |
| c. | <i>*sĩ</i> | <i>dĩ</i> | ‘seed’ |
| d. | <i>*sĩ</i> | <i>dĩ</i> | ‘to weave’ |
- (4) **Proto-Jê *j :: Karajá /d/**
- | | | | |
|----|---------------|------------------|----------------|
| a. | <i>*jĩ</i> | <i>dɛ</i> | ‘meat, flesh’ |
| b. | <i>*j-am</i> | <i>l-əba</i> | ‘to stand up’ |
| c. | <i>*j-ɔtɔ</i> | <i>d-ɔrɔ(dɔ)</i> | ‘tongue’ |
| d. | <i>*j-ĩja</i> | <i>d-ɛa(θə)</i> | ‘nose’ |
| | | <i>d-ɛadĩ</i> | ‘nasal bridge’ |
- (5) **Proto-Jê *r :: Karajá /r/**
- | | | | |
|----|----------------|--------------|---------------------|
| a. | <i>*rɛ</i> | <i>rĩ</i> | ‘to leave (trans.)’ |
| b. | <i>*prĩ</i> | <i>rĩ</i> | ‘path’ |
| c. | <i>*krã</i> | <i>ra</i> | ‘head’ |
| d. | <i>*kra</i> | <i>ra</i> | ‘child’ |
| e. | <i>*prãm</i> | <i>rəba</i> | ‘hunger’ |
| f. | <i>*mrɔ(k)</i> | <i>brĩbĩ</i> | ‘ash’ |

Overall correspondences shown by Davis’ data are also being further corroborated.

Karajá has only one voiceless stop, /k/, which corresponds regularly to *k in Proto-Jê. The lack of /p/ and /t/ in Karajá was likely the result of a systematic process of lenition (cf. Proto-Jê **par* ‘foot’ :: Karajá *wa*; Proto-Jê **tik* ‘to die’ :: Karajá *rəbɔ*). Furthermore, consonantal clusters involving a voiceless stop were systematically simplified by the complete elimination of the stop:

- (6) **cluster simplification in Karajá**
- | | | | |
|----|---------------|-------------|----------|
| a. | * <i>prĩ</i> | <i>rĩ</i> | ‘path’ |
| b. | * <i>krã</i> | <i>ra</i> | ‘head’ |
| c. | * <i>kra</i> | <i>ra</i> | ‘child’ |
| d. | * <i>prãm</i> | <i>rəba</i> | ‘hunger’ |

Both processes—lenition of the voiceless stops *p and *t and cluster simplification (6)—seem to suggest an overall tendency towards voiceless stop lenition, which may help explain the genesis of male vs. female speech differentiations (Chapter 3) and may also play an important role in explaining the origins of the derivational process of consonantal replacement described in Chapter 4 (cf. Section 3.4 below).

Advances in both descriptive and comparative Macro-Jê linguistics are contributing to improve not only the *quantity* of cognates, but—more importantly—their *quality*. Consider, for instance, the words for ‘nose’ in both Karajá and Proto-Jê, which Davis had already identified as likely cognates. Fossilized phonological material which was initially unaccounted for can now find an explanation, thanks to a better knowledge of the Karajá lexicon: the final syllable in the Karajá word is synchronically unanalyzable, but a related word, *d-ɛadĩ* ‘nasal bridge’ (cf. *dĩ* ‘bone’), clearly suggests that the last syllable in *d-ɛaθ̃* was once a separate morpheme.³⁰

In addition to corroborating Davis’ correspondences, additional research is contributing to refine them. For instance, the final syllable in the word for ‘ash’ in Karajá

³⁰ The word reconstructed by Davis for ‘nose’ was **nĩjakrɛ*, identical with the Apinajé word. The final syllable, present in most Northern Jê languages, is probably a reflex of Proto-Jê **krɛ* ‘hole’; therefore, Apinajé *nĩjakrɛ* (and related forms in other Northern Jê languages) probably meant ‘nostrils’ originally. My reconstruction takes into consideration the Southern Jê word (Kaingáng *nĩjê*), which does not include the final syllable; further corroboration for my reconstruction comes from Ofayé: *j-ĩjɛ* ‘nose’. The final syllable in Karajá *d-ɛaθ̃* may be ultimately related to the final syllable in the Apinajé form.

remained unaccounted for in Davis' comparison. Additional lexical evidence, however, suggests a regular correspondence between Proto-Jê final *k and Karajá /b/:

- (7) **Proto-Jê final *k :: Karajá /b/**
- | | | | |
|----|-----------------|--------------|-------------------------|
| a. | * <i>mrɔ(k)</i> | <i>brɪbɪ</i> | 'ash', ³¹ |
| b. | * <i>kok</i> | <i>kəbɔ</i> | 'wind' |
| c. | * <i>tɪk</i> | <i>rəbʊ</i> | 'to die', ³² |

Phonological correspondences within the Jê family are well established. Therefore, one can make well-informed decisions on likely Jê/Karajá cognates even in cases for which a Proto-Jê form cannot be reconstructed. For instance, although the Central Jê allative postposition (Xerénte *ku*, Xavante *ʔu*) cannot, so far, be reconstructed for Proto-Jê, it is still safe to compare it with the allative postposition *kɔ* in Karajá. In both Karajá and Central Jê, the allative postposition is homophonous with the word for 'wood' (Karajá *kɔ*, Proto-Jê **ko*, Xerénte *ku*). Also, the locative postposition *ki* in Kaingáng, so far attested only for the Southern Branch of the Jê family, can be suggested as a likely cognate of the locative postposition *ki* in Karajá.

- (8) a. Xerénte (Central Jê) *ku* 'allative postposition' :: Karajá *kɔ*
 b. Kaingáng (Southern Jê) *ki* 'locative postposition' :: Karajá *ki*

3.2 Personal and 'relational' prefixes

³¹ The final consonant does not occur in Northern and Central Jê languages, but its reconstruction is suggested by the Southern Jê forms (cf. Kaingáng *mrɛj*) and further corroborated by Maxakalí putok 'ashes'. Notice that Proto-Jê final *k occurs as /j/ in Kaingáng: **pek* 'fart' > Kaingáng *pej*.

³² Karajá *rəbʊ* 'death' is actually a nominal form corresponding to the verb *ɾʊ* 'to die'. That is reminiscent of Jê languages, where a deverbal noun is generally characterized by the dropping of a final consonant (cf. Mebengrokre *tuuk* ~ *tu*).

Davis (1968) had already listed two personal prefixes as likely cognates: **wa-* ~ *pa-* ‘1st person :: Karajá *wa-*; **a-* ‘2nd person’ :: Karajá *a-*. We can now add a few more prefixes to the list of cognates. As it turns out, Jê languages also have two 3rd person markers whose distribution is similar to the distribution of Karajá *d-* and *i-*. In Suyá (Northern Jê), for instance, stems such as *wa* ‘tooth’ take a consonantal 3rd person prefix, *s-*, which alternates with a ‘linking prefix’ *t-*, as in the examples below (9). The parallel with Karajá (10) is remarkable:

- (9) **Suyá**
- | | | | |
|----|---------------|----|-------------|
| a. | <i>i-t-wa</i> | b. | <i>s-wa</i> |
| | 1-REL-tooth | | 3-tooth |
| | ‘my tooth’ | | ‘his tooth’ |
-
- (10) a. *wa-dʒ-u*
- | | | | |
|--|-------------|----|--------------|
| | 1-REL-tooth | b. | <i>tf-uu</i> |
| | ‘my tooth’ | | 3-tooth |

The initial consonants in examples such as Proto-Jê **j-wa* ‘tooth’ and Karajá *dʒ-u* were treated by Davis as part of the stem, but we now know that, in both cases, the initial consonant is a linking prefix, another likely cognate between Karajá and Jê (Ribeiro 2004). Furthermore, Suyá *s-* (Xoklém *ð-*, Timbira *h-*, Xerente *s-*, etc.) can be reconstructed as **s-* for Proto-Jê, a consonant that, as we have seen in (3), corresponds systematically with Karajá /*d/*. Therefore, the cognacy between, say, Karajá *d-εaθõ* and Proto-Jê **j-ĩja* is based not only on phonological correspondences, but on their morphological behavior as well: both belong to a class of stems that take the ‘linking prefix’ (Proto-Jê **j-*, Karajá *d-* ~ *l-* ~ *dʒ-*), which

alternates with 3rd person markers which are, in all likelihood, also cognates (Proto-Jê *s- and Karajá *d-*).

The detection of these additional cognates also help explain cases that seem to contain a fossilized prefix in Karajá. One of them is the noun *dəbɨ* ‘father; his father’, which cannot synchronically take a possessive prefix. As it turns out, its likely cognate in Jê, **j-um*, belongs to the class of stems that take the consonantal 3rd person marker *s, a cognate of Karajá *d-*. The same can be said about the replacive third-person form for the locative postposition =*bə̃*, *dʊ*, which also probably contains a fossilized third-person prefix; revealingly, a likely cognate of *dʊ* in Jê languages present alternating forms: Xoklêng (Southern Jê) *ð-o ~ j-o* ‘to’ and Xerénte (Central Jê) *s-o ~ z-o* ‘to’ (Ribeiro 2004:96).

3.3 Derivational suffixes

Most Jê languages present three derivational suffixes which have likely cognates in Karajá. (Table 4). For the sake of clarity, the discussion here presented will be based on data from Mebengokre (Kaiapó), which can be considered, for our purposes, as a typical Northern Jê language.

Table 6.4. Derivational suffixes in Jê and Karajá

	Karajá	Jê family		
		Mebengokre	Xerente	
	<i>-dʊ</i>	<i>dʒwɔŋ</i>	<i>-kwa</i>	‘AGENT, SUBJ’, ³³

³³ Although the phonological differences between the cognates in this series (cf. KRJ *-dʊ* :: MEB *dʒwɔŋ* :: XER *-kwa*) may at first seem disconcerting, the correspondences they illustrate are indeed regular. MEB /dʒ/ traces back to Proto-Jê *j, which corresponds regularly to KRJ /d/; although XER /z/ is the most common reflex of Proto-Jê *j before oral vowels, the Proto-Jê sequence *[jw] occurs systematically in XER as /kw/, as further

	Karajá	Jê family		
		Mebengokre	Xerente	
	<i>-da</i>	<i>dʒʌ</i>	<i>-zɛ</i>	'INSTR'
	<i>-ɾ-</i>	<i>-ɾ</i>	<i>-ɾ</i>	'NOMLZ'

Considering their semantic and phonological characteristics, a plausible hypothesis is that the Jê suffixes that derive nouns of agent (Mebengokre *dʒwɣŋ*) and of instrument (Mebengokre *dʒʌ*) are cognates of the Karajá suffixes *-dʊ* and *-da*. Notice that Mebengokre traces back to Proto-Jê *j-, which, as we have seen, corresponds regularly to Karajá /d/. The cognates in both languages have exactly the same function. In Mebengokre, the morpheme *dʒʌ* derives nouns of instruments, places, objects. In Karajá, the suffix *-da* presents the same wide range of meanings as its likely cognates in Jê, deriving nouns of instruments (*kɔra-da* 'grinder'), places (*rira-da* 'the place where one walks'), objects (*rəθə̃-da* 'food'), etc.

Besides their obvious semantic affinities, the suffixes present identical grammatical behaviors in both Jê and Karajá. Although previous descriptions of Karajá consider *-da* and *-dʊ* as nominalizers, they are not nominalizers at all, attaching not only to *previously nominalized* verbs (*ritʃa* 'to walk', *rira* 'the action of walking', *rira-da* 'a place (or instrument) for walking', *rira-dʊ* 'the one who walks'), but also to nouns in general (cf. Chapter 4). Likewise, their likely Jê cognates are generally treated as nominalizers in Jê grammatical tradition (by, among others, Oliveira (2005), who translates the Apinajé equivalents of Mebengokre *-dʒʌ* and *dʒwɣŋ* as 'nominalizer of agent' and 'nominalizer of place and/or instrument', respectively). But, as Salanova points out, "what the "nominalizers" attach to is already nominal", a fact that is particularly clear in the

illustrated by the nearly-homophonous cognate series for 'tooth': KRJ *dʒ-u* :: MEB *dʒ-wa*, XER *kwa* (< Proto-Jê *j-ua [jwa]).

Mebengokre example below (11), where the verb *ku* ‘to eat’ occurs with the nominalizer suffix *-r*, which is then followed by the ‘instrument-noun’ suffix (Salanova 2007:88). The same is true for Central Jê languages, such as Xerênte, where the ‘nominalizers’ attach to already-nominalized verbs (12):

(11) **Mebengokre** (Salanova 2007)

i-dʒʌ-ku-r-dʒʌ

1-anti-eat.n-container

‘My eating utensils’, but also: ‘my eating place’, ‘my food’, etc.

(12) **Xerente** (Krieger & Krieger)

a. *du* ‘to carry’

b. *du-r-kwa* ‘the one who carries’

c. *du-r-zɛ* ‘carrying utensil’

Notice again that, in Northern Jê languages, both morphemes begin with the same consonant (/dʒ/ in Mebengokre, for instance), which traces back to Proto-Jê *j (Ribeiro 2005)--which, in turn, corresponds regularly with Karajá *d-*. Another cognate pair corroborates the phonological correspondences between Karajá *-du* and Mebengokre *dʒwɔɔ*: Karajá *ku* ‘to defecate’ :: Mebengokre *kwɔ*. Therefore, the Karajá nominal suffixes *-du* and *-da* are perfect matches with their likely Jê cognates, semantically, grammatically, and—last but not least—phonologically.

Considering that Karajá /r/ corresponds systematically to Proto-Jê *r (5), the Jê nominalizer *-r* would in principle be a good candidate as a cognate of “replacive morph” *-r-* that occurs with Karajá deverbal nouns. The challenge, however, lies in explaining the peculiar distribution of the latter—as a replacive infix, instead of a suffix. As it turns out, consonant replacement in Karajá can indeed be traced back to the suffixation of a

nominalizing morpheme *-r, its non-linear location being a consequence of a succession of regular phonological processes. That is what the following section describes.

3.4 ‘Trapped in amber’: from suffix to infix

As in Jê, the distribution of the morphemes in Karajá tends to follow a rather linear, agglutinative pattern, with one major exception—Karajá presents, among its nominalizing strategies, a process of *consonant replacement*, by which a velar stop in the last syllable of a verb root is replaced with /r/ in the corresponding nominal form: *ka* ‘to dig’ > *ra* ‘the action of digging’, etc. (Table 5).

Table 6.5. Consonant replacement in Karajá: k > r

	Verb	Noun	
1.	<i>ka</i>	<i>ra</i>	‘dig’
2.	<i>kɔka</i>	<i>kɔra</i>	‘grind’
3.	<i>rika</i>	<i>rira</i>	‘walk’
4.	<i>kĩ</i>	<i>rĩ</i>	‘eat’
5.	<i>kukɔ</i>	<i>kurɔ</i>	‘lift’
6.	<i>-uka</i>	<i>-ura</i>	‘split’
7.	<i>-ukɔ̃</i>	<i>-urɔ̃</i>	‘dry’ ³⁴
8.	<i>-uku</i>	<i>-uru</i>	‘ripen’

This section provides a description of the phonological processes underlying the diachronic evolution of a suffix, *-r, into a synchronic pattern of consonant alternation or (replacive) infixation in Karajá. First, a brief comparison between Karajá and Proto-Jê consonant inventories is necessary. Comparing both phonological inventories (Tables 6 and 7), the most remarkable difference is the lack, in Karajá, of the stops /p/ and /t/; /k/ is the only

³⁴ A likely cognate of this stem in the Jê family is Mebengokre *kukɔ̃* (nominal form *kukɔ̃j*). The initial syllable in Mebengokre is probably a fossilized (?) prefix; the same can be said of the initial vowel in the Karajá example.

voiceless stop in Karajá. Both *p and *t are reflected in Karajá as /w/ and /r/, respectively (Proto-Jê **par* ‘foot’, Karajá *wa*; Proto-Jê *j-ðtɔ* ‘tongue’; Karajá *d-ɔrəd’ɔ*). Since the comparative study of Macro-Jê phonology (including the comparison between Proto-Jê and another reconstructed language, Proto-Jabuti; Ribeiro & van der Voort 2005) strongly suggests that the consonant inventory was originally closer to the one found in Proto-Jê, it is safe to assume that both *p and *t underwent a process of lenition in (Pre-)Proto-Karajá.

Table 6.6
Proto-Jê consonant inventory (Ribeiro 2005)

p	t	k	(ʔ)
m	n	ŋ	
	s		
w	r	j	

Table 6.7
Karajá consonant inventory (Ribeiro 2007)³⁵

		(tʃ)	k
b	d	(dʒ)	
	dʃ		
	θ	(f)	h
	l		
w	r		

Since, as will be shown, processes of syllabic rearrangement played a crucial role in the development of consonantal replacement in Karajá, a brief description of the syllabic patterns of both Proto-Jê and Karajá is also necessary. In Macro-Jê, syllabic patterns are rather simple. As summarized by Rodrigues (1999), clusters “are limited to the combination of grave (labial and velar) stops followed by a central or a lateral approximant.” This state of affairs can be reconstructed for Proto-Jê, and is maintained in most Jê languages. Karajá also presents the same type of clusters (cf. *krɔ* ‘toad’, *brɔ* ‘back’). However, as shown by a

³⁵ As we have seen in Chapter 2, the series of palatal consonants in Karajá is an innovation (resulting from palatalization of velar, alveolar, and dental consonants around [high, +ATR] vowels). Their status as phonemes is still questionable.

comparison with Proto-Jê (and other better-known Macro-Jê languages such as Maxakalí), old clusters with an initial voiceless stop were systematically simplified in Karajá (Table 8), losing the initial stop. The process of cluster simplification seems to have affected only voiceless stops (voiced stops were preserved, as illustrated by likely cognate pairs such as Proto-Jê **mrɔk* ‘ash’, Karajá *bribi*). Clusters with an initial voiceless stop were later reintroduced in Karajá through a process of syncope (**kəɾɔ* > *krɔ* ‘frog’).³⁶

Table 6.8. Cluster simplification in Pre-Proto-Karajá

	(Proto-)Jê	Karajá		Maxakalí
9.	* <i>pri</i>	<i>ri</i>	‘path’	<i>pita</i> ~ <i>pta</i>
10.	* <i>prām</i>	<i>rəba</i>	‘hunger’	<i>pitip</i> ~ <i>ptip</i>
11.	* <i>kra</i>	<i>ra</i> ‘nephew’	‘offspring’	
12.	* <i>krã</i>	<i>ra</i>	‘head’	

Correspondences between Proto-Jê and Karajá consonants are fairly well-established, although a few gaps remain.³⁷ However, for the purposes of the present study, the facts are clear. Concerning the inventory, the crucial facts are those related to the aforementioned generalizations about consonant clusters: (a) Karajá /k/, the only voiceless stop in this language, corresponds to Proto-Jê *k (Proto-Jê **ko* ‘wood, horn’ :: Karajá *kɔ*; Proto-Jê **ku* ‘to eat’ :: Karajá *kĩ*; also ex. 12 and 13); (b) the voiced stop /b/ in Karajá corresponds to Proto-Jê *m (Proto-Jê **ma* ‘liver’ :: Karajá *ba*; Proto-Jê **mã* ‘dative’, Karajá *bã*; also ex. 14

³⁶ Such syllables can still be considered as underlyingly dissyllabic (/kə.ɾɔ/), as demonstrated by their behavior in processes such as reduplication (which, in Karajá, is sensitive to the number of moras).

³⁷ For instance, correspondences involving Proto-Jê *ŋ are still unclear. The two likely cognates involving Proto-Jê *ŋ we have so far happen to be a homophonous pair: Proto-Jê **ŋɾe* ‘egg’ :: Karajá *θɾ*, Proto-Jê **ŋɾe* ‘to dance, sing’ :: Karajá *θ-θɾ* ‘to dance’. Here, Karajá /θ/ corresponds to Proto-Jê **ŋɾ*, and not only *r, contrary to what Davis suggests (again, Proto-Jê *r corresponds to /r/ in Karajá: Proto-Jê **ɾe* ‘to leave behind’ :: Karajá *ɾ*). So, both *kr and *ŋr are reflected in Karajá as /θ/. This suggests that both *k and *ŋ merged as /k/ in Karajá, although the data do not allow a definite answer at this time, since no clear examples illustrating correspondences with *ŋ in other positions have been identified so far.

and 15); (c) Proto-Jê *r corresponds to Karajá /r/ (Proto-Jê *rɛ ‘to leave behind’ :: Karajá *rr*, also ex. 10-13 and 19).

Unlike Karajá, Jê languages allow consonant-final syllables (cf. Proto-Jê *pek ‘fart’, *tət ‘hard’, *par ‘foot’, *prām ‘hunger’, etc.). Corresponding to a CVC sequence in Jê, Karajá has two syllables, the cognate vowel showing up at the last one (CəCV). This apparent change in the position of the vowel is here termed “pseudo metathesis”—“pseudo” because, as will be shown, the vowel did not change its position at all. The key to this process is a rather common phenomenon in Macro-Jê--the insertion of so-called “echo vowels,” which has been documented in at least three families (Jê, Ofayé, and now Karajá; see the discussion of the schwa as a likely case of dissimilation, in Chapter 2). Although languages vary as to the inventory of consonants involved (in Ofayé, it only takes place after /r/: *ɸar* ‘foot’ [ɸʷaraʔ]; in Kaingáng, only after approximants; in Suyá, after all final consonants), they have one thing in common: a vowel identical with the one in the last syllable is inserted after the final consonant. In Karajá, the original vowel was then weakened to a schwa, a process that finds parallels in other Macro-Jê languages.³⁸ Notice, however, that if the final consonant is *r, the original vowel is deleted, the result being a CCV cluster (ex. 18):

³⁸ A very similar succession of processes happened in Xerente (Central Jê), as shown by a comparison with Proto-Jê and Xavante (the other surviving Central Jê language):

Syncope in Xerente

	(Proto-)Jê	Xavante	Xerente	
1.	*par	para	pra	‘foot’
2.	*pur	buru	bru	‘garden, field’
3.	*tep	tebe	tbe	‘fish’
4.	*met	pese	pse	‘good’
5.	*pok	puʔu	pku	‘to pierce’
6.	*kēn	ʔene	kne	‘stone’

Table 6.9. “Pseudo metathesis”

	(Proto-)Jê	Karajá	
13.	* <i>prãm</i>	<i>rəma</i>	‘hunger’
14.	* <i>j-am</i>	<i>l-əma</i>	‘to stand up’
15.	* <i>z-um</i>	<i>dəbi</i>	‘his father’
16.	* <i>j-í(j)</i>	<i>l-ədɪ</i>	‘to lay down’
17.	Apinajé <i>mɛɲ</i>	<i>bədɪ</i>	‘honey’
18.	Apinajé <i>por</i> ‘lower back’ (< * <i>wor</i>)	<i>brɔ</i>	‘back’

We are now in a position to propose a diachronic explanation for the emergence of the consonant replacement pattern shown in Table 1. As the correspondences illustrated above show, it is rather plausible to consider (on phonological and semantic grounds) the “replacive infix” –r- in Karajá as a cognate of the Jê nominalizer *-r. Its peculiar distribution can be accounted for by a succession of independently-demonstrable diachronic processes. This is illustrated below (Table 10) by the Karajá root *kɪ* ‘to eat’ (nominal form *rɪ*), a likely cognate of Proto-Jê **ku* (nominal form **ku-r*)

Table 6.10. From suffix to “replacive infix”

(a) suffixation	* <i>kɪ-r</i>
(b) echo-vowel insertion (Table 5)	* <i>kɪrɪ</i>
(c) vowel weakening/dissimilation (Table 5)	* <i>kəɾɪ</i>
(d) cluster formation (Table 5, ex. 19)	* <i>krɪ</i>
(e) cluster simplification (Table 5)	<i>rɪ</i>

If the description provided here is correct, we should be able to find elsewhere (besides those examples involving suffixation of the nominalizer) additional cases in which final *r came to be in prevocalic position. That is indeed the case: the Karajá cognate of the Jê verb root **kɪr* ‘to call, to bellow’³⁹ is *rɪ*, displaying exactly the same correspondences

³⁹ In this case, the final /r/ is indeed part of the stem, as suggested by Southern Jê (Xoklég *kɪl* ‘to shout out, to cry’; Urban 1985:177), where the nominalizer *-r does not occur (cf. also Maxakalí *cat* ‘to call’; Pereira

described above (see also Table 9, ex. 18). Examples such as these show that the infixal location of the nominalizer is a byproduct of regular phonological processes which affected *r-final stems in general.

Also, if consonant clusters involving voiced stops were indeed preserved in Karajá, one should be able to find cases in which the nominalizer has been preserved after /b/. That, in fact, also happens: the nominal form corresponding to the Karajá verb *-bu* ‘to cry’ is *-bru* (cf. Meb. *muw* ‘to cry’, nominal form *muw-r*; also compare example 18 in Table 9). Cases such as **par* ‘foot’ :: Karajá *wa*, where lenition seems to have prevented the creation of an environment for cluster formation, provide clues on the relative chronology of the phonological processes postulated herein (*p-lenition must have preceded cluster formation).⁴⁰

3.4.1 Residual examples: h-replacement

There are a few verb stems for which, instead of /k/, it is /h/ that is replaced by /r/ in the nominal form (Table 11). Except for *ho* ‘to wash’,⁴¹ all of these stems present, after the glottal fricative, a front vowel. Thus, it is possible that such cases represent original instances in which *k preceded a front vowel. The velar stop would have undergone a

1992:95). The final consonant was apparently reinterpreted as a suffix in the other branches of the Jê family, Central and Northern Jê, where the nominalizer *-r* is still productive.

⁴⁰ One should not assume, however, that lenition took place simultaneously for *p and *t. As the survival of *k demonstrates, different consonants (although belonging to the same “natural class”) may behave differently in processes such as this (and the same applies to cluster simplification). Lenition of *p and *t and cluster simplification can both be seen as instances of the same tendency (a Sapirian “drift,” if we will) to weaken voiceless stops (fully consolidated in pre-consonantal position, but only half-way so in pre-vocalic position). As for the “preservation” of *k in Karajá, that is only half the story, as we have seen in Chapter 3.

⁴¹ Clear correspondences in Jê for Karajá /h/ have not yet been found. /h/ could have derived historically from a stop (for instance, *ʔ); that would provide an explanation for its behavior, in terms of consonantal replacement, with the stem *ho* ‘to wash’.

process of spirantization after the alternation between *k and *r was already in place (e.g. *kɛ ‘to scratch’ ~ nominal form *rɛ > hɛ ~ rɛ). In fact, there are no cases of r-infixation involving the velar stop before a front vowel; thus, it is reasonable to suggest that, in the small corpus of consonant-replacing verbs, /k/ and /h/ are in complementary distribution (again, with the exception of hɔ ‘to wash’). Likely cognates in Jê suggest that this interpretation is on the right path (cf. Meb -ke ~ -kep ‘to scratch’ :: Karajá hɛ; Meb. kɔkje ~ kɔkjer ‘to cut with a cutting instrument’ :: Karajá kɔhɛ).

Table 6.11. Consonant replacement in Karajá: h > r

	Verb	noun	
19.	<i>hɔ</i>	<i>rɔ</i>	‘wash’
20.	<i>hɪ</i>	<i>rɔdɪ</i>	‘to chase’
21.	<i>hɛ</i>	<i>rɔdɛ</i>	‘scratch’
22.	<i>hɛ</i>	<i>rɔdɛ</i>	‘pluck, pull off, remove’
23.	<i>bɛhɛ</i>	<i>bɛrɛθɛ</i>	‘go down’
24.	<i>kɔhɛ</i>	<i>kɔrɛ</i>	‘cut’

3.4.2 An alternative analysis?

In principle, an alternative, more abstract explanation for the evolution of the suffix *-r into an infix in Karajá could also be proposed, similar to the one suggested by Milizia for the Proto-Indo-European nasal infixation rule (Milizia 2004). Simply put, the suffix *-r would have undergone metathesis as a strategy to resolve a constraint against syllable-final consonants (*kɪr > *krɪ). This would provide a more straightforward explanation, eliminating the need to postulate steps (b), (c), and (d) (Table 10), which could be replaced by one single rule of metathesis.

The explanation offered above (Table 10), however, has at least two major advantages. First, it provides a unified treatment of final consonants (including cases that led to schwa-formation, described in Chapter 2). Second, it is further confirmed by independent internal evidence. Cluster formation resulting from syncope is attested, for instance, in examples such as (12) below, where the prefix combination *ka-r-a-* is reduced to [kra]:

- (12) ka-r-a-rika=kre [kraritʃakre]
 1-CTFG-INTR-walk=FUT
 ‘I will walk.’

3.4.3 Karajá and the typology of infixation

In a recent study on infixation, Yu (2007) describes four different paths which may lead to the creation of infixes: (1) *metathesis*, (2) *entrapment* (a “scenario in which a morpheme is stranded between a fossilized composite of an affix and a root”), (3) *reduplication mutation*, and (4) *morphological excrescence and prosodic stem association*. The rather unusual development of the Karajá infix –r- does not seem to fit into any of such likely scenarios (although, as I mention above, metathesis may be seen as a possible alternative explanation). The development of –r-, however, seems to be closer to what Yu describes as entrapment—except that, in our case, the morpheme was trapped by fossilized phonological, rather than morphological, material. If the scenario here described is correct, the notion of entrapment could be extended to deal with cases of phonological entrapment as well, in addition to cases of morphological entrapment.

4. Final remarks

As this chapter demonstrates, despite the limited number of likely lexical cognates, it is still necessary--and rather rewarding--to always take into consideration phonological evidence when proposing grammatical cognates. Thus, for instance, the Karajá third-person *d-* was here shown to be a cognate of Proto-Jê third-person prefix **s-* based not only on superficial similarities, but on distributional (both Karajá *d-* and Proto-Jê **s-* occur with vowel-initial stems and alternate with a ‘linking prefix’) and phonological grounds (Karajá /d/ corresponds systematically to Proto-Jê **s*). The regularity of the correspondence is further illustrated by the fact that a homophonous pair in Karajá (the words for ‘seed’ and ‘to weave’) corresponds to a homophonous pair in Jê (cf. 3c and 3d).

A more dramatic example of the interplay between descriptive and historical linguistics is the case of the ‘replacive infix’ *-r-* in Karajá. Despite its pervasiveness, the replacive infix was not even mentioned by previous authors (Fortune, Maia). While I described this phenomenon in my MA thesis (Ribeiro 1996), its origin would have remained a mystery were it not for the comparative evidence that is gradually being amassed, thanks to recent improvements in the descriptive knowledge of several Jê and Macro-Jê languages. By demonstrating that such a phenomenon traces back to a suffix with a clear cognate in Jê, an additional piece of evidence is added to the comparative pool. On the other hand, comparative evidence allows for a better understanding of how this phenomenon came about diachronically, providing additional data for the typological understanding of infixation, for which data from South American languages are conspicuously missing. The Karajá data here provided may help fill such a gap.

Unlike its cognates in most Jê languages, *-r-* is not a productive morpheme in Karajá,

surviving only as a relic in examples involving the infixation pattern described in this study. By “encasing” the original suffix in the root, the phonological processes here described preserved a further piece of evidence for the genetic relationship between Karajá and the Jê family, as part of the Macro-Jê stock. As this study briefly illustrates, regular phonological processes can substantially alter the shape of stems, making the identification of cognates between Karajá and Jê even more difficult (that is particularly the case with examples involving cluster formation followed by cluster simplification, such as Proto-Jê **kír* ‘to call’: Karajá *ri*). The lesson the data teach us, however, is one of optimism: as we gain a better understanding of the individual languages that constitute the Macro-Jê stock, a better picture of the phonological and morphological correspondences among them starts to emerge; that, in turn, provides explanations for puzzling synchronic issues.

BIBLIOGRAPHY

- Anderson, Stephen. 1985. Inflectional morphology. In Shopen, Timothy (ed.), *Language Typology and Syntactic Description*, vol. III, 150-201. Cambridge University Press.
- Baker, Mark C. 1988. *Incorporation: A Theory of Grammatical Function Changing*. Chicago: University of Chicago Press.
- Bauer, Laurie. 1988. *Introducing Linguistic Morphology*. Edinburgh University Press.
- Bybee, Joan. 1985. *Morphology: a Study of the Relation between Meaning and Form*. Amsterdam: John Benjamins.
- Bakovic, Eric. 2000. Harmony, dominance, and control. Doctoral dissertation. Rutgers University.
- Beckman, Jill. 1995. Shona height harmony: markedness and positional identity. *UMOP* 18, 53-76.
- Beckman, Jill. 1997. Positional faithfulness, positional neutralization, and Shona vowel harmony. *Phonology* 14, 1-46.
- Beckman, Jill. 1998. Positional faithfulness. Doctoral dissertation, University of Massachusetts, Amherst.
- Borges, Mônica Veloso. 1994. Diferenças entre a fala feminina e a masculina na língua Karajá (Brasil). In *II Jornadas de Lingüística Aborígen* (Proceedings). Buenos Aires: Universidad de Buenos Aires.
- Borges, Mônica Veloso. 1997. As falas masculina e feminina em Karajá. Master's thesis. Goiânia: Universidade Federal de Goiás.
- Brito Machado, Otton Xavier de. 1947. *Os Carajás (Inan-Son-Uéra)*. Rio de Janeiro: Imprensa Nacional.
- Castelnau, Francis. 1851. *Expédition dans les Parties Centrales de l'Amérique du Sud V*. Paris.
- Cavalcante, Marita Pôrto. 1992. Fonologia do Karajá. *Revista do Museu Antropológico* 1. Goiânia: Editora da UFG.
- Cooreman, Ann. 1994. A functional typology of antipassives. In Fox, Barbara and Paul J Hopper (eds.) *Voice: Form and Function*, 49-87. Amsterdam: John Benjamins.
- Coudreau, Henri. 1897. *Voyage au Tocantins-Araguaya*. Paris.

- Davis, Irvine. 1966. Comparative Jê phonology. *Estudos Lingüísticos: Revista Brasileira de Lingüística Teórica e Aplicada*, 1:2.10-24.
- Davis, Irvine. 1968. Some Macro-Jê relationships. *IJAL*, 34.42-7.
- Dixon, Robert M. W. 1979. Ergativity. *Language* 55: 59-138.
- Ehrenreich, Paul. 1891. *Beiträge zur Völkerkunde Brasiliens*. Berlin. [Portuguese translation: 1948. Contribuições para a Etnologia do Brasil. *Revista do Museu Paulista*, Vol. II. São Paulo.]
- Ehrenreich, Paul. 1894. Materialien zur Sprachenkunde Brasiliens, I: die Sprache der Caraya (Goyaz), *Zeitschrift für Ethnologie*, 26.20-37, 49-60.
- Foley, William, and Robert Van Valin, Jr. 1985. Information packaging in the clause. In Shopen, Timothy (ed.), *Language Typology and Syntactic Description I*, 282-364. Cambridge University Press.
- Fonseca, José Pinto da. 1867 [1775]. Cópia da carta que o alferes José Pinto da Fonseca escreveu ao Exmo. General de Goyazes, dando-lhe conta do descobrimento de duas nações de índios, dirigida do sítio onde portou. *Revista Trimensal de Historia e Geographia*, Tomo VIII, pp. 376-390.
- Fortune, David & Fortune, Gretchen. 1963. The phonemes of the Karajá language (manuscript). Rio de Janeiro: Arquivo Lingüístico do Museu Nacional.
- _____. 1964. Karajá grammar (unpublished manuscript). Rio de Janeiro: Arquivo Lingüístico do Museu Nacional.
- _____. 1975. Karajá men's women's speech differences with social correlates. *Arquivos de Anatomia e Antropologia* 1. Rio de Janeiro: Instituto de Antropologia Prof. Souza Marques.
- Fortune, David. 1973. Gramática Karajá: um estudo preliminar em forma transformacional. *Série Lingüística* 1, 101-161. Brasília: SIL.
- Gildea, Spike. 1998. *On Reconstructing Grammar: Comparative Cariban Morphosyntax*. New York/Oxford: Oxford University Press.
- Haas, Mary. 1964. Men's and women's speech in Koasati. In Hymes, D. (editor) *Language in Culture and Society: a Reader in Linguistics and Anthropology*, 228-233. New York: Harper and Row.

- Hulst, Harry van der & Jeroen van der Weijer. 1995. Vowel harmony. *Handbook of Phonological Theory*, ed. by John Goldsmith, pp. 495-534. Cambridge, MA: Blackwell.
- Krause, Fritz. 1911. *In den Wildnissen Brasiliens*. Bericht und Ergebnisse der Leipziger Araguaya-Expedition 1908. Leipzig: R. Voigtländers Verlag. [Portuguese translation: 1940-1944. Nos sertões do Brasil. *Revista do Arquivo Municipal*, vol. 66-95.]
- Kunike, H. 1916. Beiträge zur Phonetik der Karajá-Sprache (Brasilien). *Internationales Archiv für Ethnographie* 23.147-82.
- Kunike, H. 1919. Die Phonetik der Karaiá-Sprache nach linguistischen Principien. *Journal de la Société des Américanistes*, n.s., 11.139-81.
- Lipkind, William. 1948. The Carajá. In Steward, J. H., *Handbook of South American Indians*, v. 3, 179-191. Washington: Smithsonian Institution, Bureau of American Ethnology.
- Lombardi, Linda. 1996. Restrictions on direction of voicing assimilation: an OT account. *University of Maryland Working Papers in Linguistics* 4, 84-102.
- Lombardi, Linda. 1999. Positional faithfulness and voicing assimilation in Optimality Theory. *NLLT* 17, 267-302.
- Loukotka, Chestmir. 1968. *Classification of South American Indian Languages*. Los Angeles: Latin American Center, UCLA.
- Maia, Marcus. 1998. *Aspectos Tipológicos da Língua Javaé*. München: Lincom Europa. [Publication of his 1986 master's thesis; Universidade Federal do Rio de Janeiro.]
- Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60.847-93.
- Mithun, Marianne. 1986. On the nature of noun incorporation. *Language* 62.32-37.
- Noske, Manuela. 1995. *The Ternary Use of Distinctive Features: in Defense of the Number Three*. PhD dissertation. Chicago: University of Chicago.
- Palha, Frei Luiz. 1942. *Ensaio de Gramática e Vocabulário da Língua Karajá, Falada pelos Índios Remeiros do Rio Araguaia*. Rio de Janeiro.
- Ribeiro, Eduardo R. 1996. Morfologia do verbo Karajá. Master's thesis. Goiânia: Universidade Federal de Goiás.

- _____ 2000. [ATR] vowel harmony and palatalization in Karajá. *Santa Barbara Papers in Linguistics, Vol. 10: Proceedings of WAIL 2000*, p. 80-92. Santa Barbara: University of California.
- _____ 2000a. Direction in Karajá. Paper presented at the VI Encuentro de Lingüística en el Noroeste. Sonora, México: Universidad de Sonora.
- _____ 2001. Empréstimos Tupí-Guaraní em Karajá. To appear in *Revista do Museu Antropológico*. Universidade Federal de Goiás, Brazil.
- _____ 2001a. Valence, voice, and noun incorporation in Karajá. Paper presented at the 27th BLS Meeting. Berkeley: University of California.
- _____ 2002. Directionality in vowel harmony: the case of Karajá. Paper presented at the 28th BLS Meeting. Berkeley: University of California.
- _____ 2002a. Relational prefixes and the ‘Macro-Jê hypothesis.’ Paper presented at the 7th Workshop on Structure and Constituency in the Languages of the Americas. University of Alberta at Edmonton, Canada.
- _____ 2002b. On verbs, nouns, and ‘adjectives’ in Karajá. Paper presented at the 38th Meeting of the Chicago Linguistic Society. Chicago, April 2002.
- Rigsby, Bruce & Michael Silverstein. 1969. Nez Perce vowels and Proto-Sahaptian vowel harmony. *Language* 45.1: 45-59.
- Rodrigues, Aryon. 1985. “Evidence for Tupi-Carib relationships.” In Harriet Klein and Louisa Stark (eds.), *South American Indian Languages, Retrospect and Prospect*, 371-404. Austin: University of Texas Press.
- _____ 1985a. Relações internas na família lingüística Tupí-Guaraní. *Revista de Antropologia* 27/28, 33.53.
- _____ 1986. *Línguas Brasileiras: para o Conhecimento das Línguas Indígenas*. São Paulo: Edições Loyola.
- _____ 1994. Grammatical affinities among Tupí, Carib, and Macro-Jê. Unpublished manuscript, Universidade de Brasília.
- _____ 1999. Macro-Jê. In Aikhenvald & Dixon (editors), *The Amazonian Languages*. Cambridge: Cambridge University Press.
- _____ 2000. ‘Ge-Pano-Carib’ versus ‘Jê-Tupí-Karib’: sobre relaciones lingüísticas prehistóricas en Sudamérica. In Miranda. Luis (editor), *Actas del I Congreso de Lenguas Indígenas de Sudamérica*, Tomo I, 95-104. Lima: Universidad Ricardo Palma.

- Sadock, Jerrold and Arnold M. Zwicky. 1985. Speech act distinctions in syntax. In Shopen, Timothy (ed.), *Language Typology and Syntactic Description* vol. I, 155-95. Cambridge: Cambridge University Press.
- Sadock, Jerrold. 1980. Noun incorporation in Greenlandic: a case of syntactic word-formation. *Language* 56.300-319.
- _____. 1986. Some notes on noun incorporation. *Language* 62.19-31.
- _____. 1991. *Autolexical Syntax*. Chicago: University of Chicago Press.
- Sapir, Edward. 1911. The problem of noun incorporation in American languages. *American Anthropologist* 13.250-282.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Dixon, R. (editor), *Grammatical Categories in Australian Languages*, 112-71. Canberra: Australian Institute of Aboriginal Studies.
- Talmy, Leonard. 1985. Lexicalization patterns: semantic structure in lexical forms. In Shopen (ed.), *Language Typology and Syntactic Description*, vol. III, 57-149. Cambridge University Press.
- Viana, Adriana. 1995. A expressão do atributo na língua Karajá. Master's thesis. Brasília: Universidade de Brasília.
- Wiesemann, Ursula. 1986. The pronoun systems of some Jê and Macro-Jê languages. In Wiesemann, U. (ed.), *Pronominal Systems*, 359-80. Tübingen: Gunther Narr Verlag.