A grammar of Baniwa classifiers

Cronhamn, Sandra

2025

Document Version:
Publisher's PDF, also known as Version of record

A grammar of Baniwa classifiers

SANDRA CRONHAMN
CENTRE FOR LANGUAGES AND LITERATURE | LUND UNIVERSITY

A grammar of Baniwa classifiers

A grammar of Baniwa classifiers

by Sandra Cronhamn

LUND
UNIVERSITY

Thesis for the degree of Doctor of Philosophy
Thesis advisors: Prof. Niclas Burenhult, Dr. Karolin Obert,
Prof. Gerd Carling
Faculty opponent: Assoc. Prof. Katharina Haude

To be presented, with the permission of The Joint Faculties of Humanities and Theology of Lund University, for public criticism in SOL:H104 on Friday, the 31st of January 2025 at 10:15.

Organization
LUND UNIVERSITY

Centre for Languages and Literature
Box 201
SE–221 00 LUND
Sweden

Document name
DOCTORAL DISSERTATION

Date of disputation
2025-01-31

Sponsoring organization
–

Author(s)
Sandra Cronhamn

Title and subtitle
A grammar of Baniwa classifiers

Abstract
Nominal classification systems provide a unique window into the intersection of grammar, semantics, and cognition. Found in more than half of the world's languages, these systems possess both universal and language-specific properties. Nominal classification systems of a specific type, featuring classifiers marked in multiple morphosyntactic loci, are found in many languages in Northwestern Amazonia. These systems are of particular typological interest as they share properties with several other types of nominal classification systems, yet few of them have been described in detail. A classifier system of this type is found in the Arawakan language Baniwa, spoken by a few thousand people in Northwestern Brazil.

This thesis provides a detailed analysis of the Baniwa classifier system from phonological, morphological, syntactic, functional, semantic, typological, historical and contact perspectives, based on first-hand data from field work and a combination of descriptive and experimental approaches. The analysis outlines a flexible and versatile system encoding semantic distinctions of animacy, shape, and part–whole relations. Classifiers have several central functions in the grammar of Baniwa, including derivation, inflection, and referent tracking in discourse. The system has developed over the course of several millennia, and continues to develop in the face of ongoing cultural change and language contact. It shares many of its properties with classifier systems in other Arawakan languages, as well as with classifier systems in unrelated languages in the area.

This account of Baniwa classifiers contributes to the understanding of nominal classification systems more widely, in particular those of Arawakan and Northwestern Amazonian languages, and illuminates the structure, development and maintenance of such systems. The analysis sheds light on a number of commonly posited dichotomies in linguistic theory, such as the distinction between lexical and grammatical forms, and between inflectional and derivational processes. It is also a contribution to our knowledge of lesser-known languages.

Key words
Baniwa, Arawakan languages, Northwestern Amazonian languages, Amazonian languages, South American languages, classifiers, nominal classification

Classification system and/or index terms (if any)
–

Supplementary bibliographical information
–

Language
English

ISSN and key title
–

ISBN
978-91-89874-69-5 (print)
978-91-89874-70-1 (pdf)

Recipient's notes
–

Number of pages
325

Price
Free

Security classification
–

A grammar of Baniwa classifiers

by Sandra Cronhamn

LUND
UNIVERSITY

A doctoral thesis at a university in Sweden takes either the form of a single, cohesive research study (monograph) or a summary of research papers (compilation thesis), which the doctoral student has written alone or together with one or several other author(s). This thesis is a monograph.

Cover photo by Sandra Cronhamn

The Joint Faculties of Humanities and Theology, Centre for Languages and Literature

ISBN: 978-91-89874-69-5 (print)
ISBN: 978-91-89874-70-1 (pdf)

Printed in Sweden by Media-Tryck, Lund University, Lund 2025

Media-Tryck is a Nordic Swan Ecolabel certified provider of printed material. Read more about our environmental work at www.mediatryck.lu.se

Printed matter 3041 0903

MADE IN SWEDEN

Le langage est une traduction de la réalité, une transposition où les objets figurent déjà généralisés et classifiés par le travail de la pensée

Bréal (1897: 275)

Contents

List of Figures

List of Tables

Acknowledgements

While there is only one name listed as the author of this thesis, there are many people without whom this work would not have been possible.

First and foremost, I want to express my deepest gratitude to all the speakers of Baniwa that I have had the pleasure to work with: Valkíria Apolinário, Franklin da Silva, Horácio da Silva, Frankito Apolinário, Bette Olegário da Silva, Alberto Apolinário, Mônica Apolinário, Valentim Lisbão, Francinaldo Baniwa, Laureano Rodrigues, Graceleano Maudre Pinheiro, Fernado José, Alino Candido da Silva Souza dos Santos, Hermes Ernesto Vitorino Maudú, Rosiane de Lima, Dona Alda, Santolina Alexandre, Tereza da Silva Torres, Adelina Alexandre Marcelino, André Pedro, Armindo Hermes dos Santos, Valdis José de Souza Brazão, Cláudia Garrido Ricardo, and José Apolinário. Thank you for welcoming me to your homes and for being so generous with your knowledge, hospitality, time, and patience when teaching me your wonderful language. I am especially grateful to Valkíria, Frank, and Horácio for their invaluable contribution to this thesis, as well as for their great company.

I would also like to thank my supervisors, for a lot of things: the knowledge, the support, the enthusiasm, and the intellectual challenges that they have poured into this project. In particular, thank you, Niclas and Karol, for all the hours you have put in during this past year, as well as for the encouragement that helped me stay sane during the process. And thank you, Niclas and Gerd, for taking me on as a young student and introducing me to the world of linguistics.

A number of people have kindly given feedback on parts of my drafts and analyses. I especially want to thank Pattie Epps for providing feedback and perspectives that greatly helped me improve this thesis. I also want to thank Françoise Rose for our discussions in the spring 2022, which gave me lots of new ideas, Colette Grinevald for taking the time to discuss Baniwa classifiers with me, Saskia Dunn for valuable comments on the historical analysis, Mechtild Tronnier for helpful feedback on the phonological analyses, and Niklas Erben Johansson for all his input over the years.

Thank you to all my colleagues at the Linguistics department in Lund for the intellectually stimulating research environment, but also for the silliness around the kitchen table. In particular, I have enjoyed being part of the PhD student body with its collegial atmosphere. A special thanks to Niklas Erben Johansson, Filip Larsson, Olof Lundgren, Victor Bogren Svensson, Eline Visser, Jo Yager, Felix Ahlner, Kajsa Gullberg, and Anna Hjortdal—it has been a pleasure to share parts of the PhD journey with you.

Apart from Lund University, me and my work have also benefited greatly from shorter research stays at other departments. I want to thank Ana Suelly Arruda Câmara Cabral and the rest of the people at LALLI for welcoming me in Brasília. I also want to thank Françoise Rose and the team at DDL for hosting me in Lyon.

I am grateful to FOIRN, especially the late Isaias Pereira Fontes, for the encouragement and interest in my project. I also thank the Birgit Rausing Language Programme and the Joint Faculties of Humanities and Theology at Lund University for funding field trips and recording equipment. I would also like to thank the Lund University Humanities Lab, especially Jens Larsson, for helping with archiving.

Things do not always go as planned, especially field work. I am grateful to the Swedish Consulate in Manaus, Falck Global Assistance, Kammarkollegiet, Maria Persson, and Arthur Holmer for their assistance and support when I found myself stuck in Amazonia in the middle of a global pandemic. I would also like to thank my maternity care midwife Hanna for not simply brushing me off when I announced that I had to go and do field work in remote Amazonia at five months pregnant, instead going out of her way to help me sort out all the challenges. Thank you also to Sanja Wihk Monção and Guilherme Reis Monção for the medical advice, as well as for the good times in São Gabriel.

Christian—what would I have done without you? I would certainly not have finished this thesis. Thank you, from the bottom of my heart, for all your support and understanding, for your unbelievable patience, for putting up with my absences (both geographical and mental) during the course of this work, for bringing me snacks, and for sacrificing so much of your own time so that I was able to focus on writing. I also want to thank my parents, my sister Sofia, my in-laws, and my friends for all their support and help, and for always wanting to hear about my research.

Last but not least: thank you, Edith, for being such a ray of sunshine.

Abstract

Nominal classification systems provide a unique window into the intersection of grammar, semantics, and cognition. Found in more than half of the world's languages, these systems possess both universal and language-specific properties. Nominal classification systems of a specific type, featuring classifiers marked in multiple morphosyntactic loci, are found in many languages in Northwestern Amazonia. These systems are of particular typological interest as they share properties with several other types of nominal classification systems, yet few of them have been described in detail. A classifier system of this type is found in the Arawakan language Baniwa, spoken by a few thousand people in Northwestern Brazil.

This thesis provides a detailed analysis of the Baniwa classifier system from phonological, morphological, syntactic, functional, semantic, typological, historical and contact perspectives, based on first-hand data from field work and a combination of descriptive and experimental approaches. The analysis outlines a flexible and versatile system encoding semantic distinctions of animacy, shape, and part–whole relations. Classifiers have several central functions in the grammar of Baniwa, including derivation, inflection, and referent tracking in discourse. The system has developed over the course of several millennia, and continues to develop in the face of ongoing cultural change and language contact. It shares many of its properties with classifier systems in other Arawakan languages, as well as with classifier systems in unrelated languages in the area.

This account of Baniwa classifiers contributes to the understanding of nominal classification systems more widely, in particular those of Arawakan and Northwestern Amazonian languages, and illuminates the structure, development and maintenance of such systems. The analysis sheds light on a number of commonly posited dichotomies in linguistic theory, such as the distinction between lexical and grammatical forms, and between inflectional and derivational processes. It is also a contribution to our knowledge of lesser-known languages.

Abbreviations

Languages

PA	Proto-Arawak
PJC	Proto-Japurá-Colombia
Port.	Portuguese

Phonology

C	consonant
N	nasal
V	vowel

Grammar

A	subject of transitive verb
Adj	adjective
CLF^s	simplex classifier form
CLF^c	complex classifier form
N	noun
NUM	numeral
NC	nominal classification
NP	noun phrase
O	object of transitive verb
POSS	possessor
RC	relative clause
S_a	subject of active intransitive verb
S_o	subject of adjective/stative intransitive verb
V	verb

Glossing in examples and tables

ACCUM	accumulative ‘increasingly’
ADJZ	adjectivizer
ADD	additional ‘also’
ADM	admonitive ‘warning’
ADVS	adversative
AL	alienator
ALL	allative
AUG	augmentative
AUX	auxiliary
BEN	benefactive
CAUS	causative
CEXP	counterexpectative
CLF	classifier
CLF.CURV	curvilinear classifier (full gloss: CLF.CURVILINEAR)
CLF.EXCR	excrement classifier (full gloss: CLF.EXCREMENT)
CLF.GNR	generic classifier (full gloss: CLF.GENERIC)
COMPR	comparative
COND	conditional
CONN	connective
CONST	constant ‘always’
CNSQ	consequential
CNTR	contrastive
DEM	demonstrative
DISJ	disjunctive
DUR	durative
EVD	evidentive ‘of course’
F	feminine
FRUS	frustrative
FUT	future
GER	gerund
GNO	gnomic
IMM	immediative
IMPL	implicative
IMPRS	impersonal
INAL	inalienator
INGR	ingressive
INSTR	instrumental
INTS	intensive
IRR	irrealis

LOC.in	locative ‘inside’
LOC.on	locative ‘on top of’
MID	middle voice
N	neuter
NEG	negative
NF	non-feminine
NMLZ	nominalizer
NMLZ.AG	agent nominalizer
NMLZ.INSTR	instrument nominalizer
NMLZ.LOC	locative nominalizer
NMLZ.PAT	patient nominalizer
OPT	optative
PERS	persistive
PL	plural
POS	positive
PQ	polar question particle
PROG	progressive
PRF	perfect
PRIV	privative
PROG.INCH	progressive inchoative ‘ever since’
PRP	purposive
PST	past
PST.INCH	past inchoative ‘already’
Q	interrogative root
REFL	reflexive
REL	relative
REP	repetitive ‘again’
RESTR	restrictive
RPT	reportative
SEP	separative
SG	singular
SMBL	semblative
SEQ	sequential
SIM.DIST	simultaneous distal ‘at the time when’
SIM.PROG	simultaneous progressive ‘during’
SIM.PUNC	simultaneous punctual ‘when’
SIM.SEP	simultaneous separative ‘during (unnoticed)’
SUB	subordinator
SUSP	suspensive

Part I

Preliminaries

Chapter 1

Introduction

The question of how humans categorize the world through language has long fascinated linguists, philosophers, psychologists, and cognitive scientists, as it presumably tells us something about the inner workings of our minds (Aristotle, ca 350 BCE/2014; Rosch, 1975; Senft, 2000, etc.). Nominal classification is a grammatical phenomenon whereby entities referred to by nouns are explicitly divided into categories, thus offering an excellent opportunity for the study of categorization. Nominal classification systems are found in more than half of the world's languages (Allassonnière-Tang et al., 2021) and bear striking similarities in their semantic underpinnings (Seifart, 2010). They typically show great complexity which takes time for the language learner to master (Aikhenvald, 2000: 413–421), and the fact that they exist despite this suggests that they are cognitively and communicatively helpful for language users. Their cross-linguistic prevalence and relative semantic cohesion also suggest that they encode categories of universal cognitive importance. At the same time, each nominal classification system is unique in its particular division of entities. The study of nominal classification systems thus has the potential to broaden our understanding of both universal and language-specific properties of human communication systems.

This thesis focuses on one particular nominal classification system, namely the classifier system of the Arawakan language Baniwa spoken in Northwestern Amazonia. The nominal classification systems in languages of this region are highly interesting from a typological point of view for several reasons. First, they typically combine properties previously thought to be incompatible, such as large inventories and flexible assignment on the one hand, and bound forms and syntactic agreement on the other (Dixon, 1986). Their scientific discovery therefore inspired a revision of existing typologies (Seifart & Payne, 2007). Sec-

ond, they display a high degree of structural similarity across language families (Krasnoukhova, 2012), begging the question of what features of nominal classification, or indeed language in general, are prone to areal diffusion. Northwestern Amazonian languages are therefore significant for our understanding of nominal classification as a whole. However, many of the languages in question remain un- or underdescribed (Seifart & Payne, 2007), and the careful description of individual systems is a prerequisite for further theorization and comparison. In this thesis, I provide a detailed description and analysis of the classifier system in Baniwa, as spoken in an urban regional center in the middle of the Brazilian Amazon. Apart from constituting a study in its own right, it is also a contribution to the study of Arawakan and Northwestern Amazonian nominal classification systems, as well as to the study of nominal classification in general.

In this chapter, I present the language and its speakers (§ 1.1), before going on to present the study (§ 1.2) in terms of its aims, methods, and data. § 1.3 introduces the conventions I adopt, and § 1.4 outlines the structure of the thesis.

1.1 Baniwa: the language and its speakers

Baniwa-Koripako is a continuum of mutually intelligible language varieties[1] spoken in the Upper Rio Negro region, in the borderlands of Brazil, Colombia, and Venezuela. The speakers of Baniwa-Koripako mainly live along the Içana river and its tributaries, as well as along the Guainia river. There is also a substantial population in and around the Brazilian town of São Gabriel da Cachoeira (see § 1.1.2.1), as well as in some urban areas of Colombia and Venezuela (Ramirez, 2020a: 37).

Ramirez (2001a: 2–3; 2001b: 21 ff.; 2020a: 37–40, 43–46) recognizes three main dialectal varieties of Baniwa-Koripako, called *Koripako* ("Northern"), *Baniwa* ("Central"), and *Karotana* ("Southern"), the geographical labels referring to their historical relative locations[2]. The same tripartite division was recognized by Nimuen-

[1] In Glottolog (Hammarström et al., 2024), Baniwa-Koripako is associated with the glottocode [bani1259], further subdivided into "Baniwa do Icana" [bani1255] and "Curripaco" [curr1243]. In Ethnologue (Eberhard, Simons & Fennig, 2024), it is divided into "Baniwa" and "Curripaco", associated with the respective ISO 639-codes [bwi] and [kpc]. Wikipedia has an entry about the language under the name "Karu" (Wikipedia, 2024a). Further alternative names for the language include Baniba, Baniva, Maniba (do Içana/del Isana), Kurripako, Walimanai, and Wakuenai. The language described in this thesis should not be confused with Baniva (of Maroa/Guainia) [guar1293] of Venezuela, another Arawakan language.

[2] Ramirez (2020a: 37–40) additionally treats Tariana ("Tariano") as a fourth dialect of the same language—see further in § 1.1.1.

dajú already in 1927 (Nimuendajú, 1950: 169), and is relatively well accepted among linguists[3], even if the dialects also contain a certain amount of internal variation (Gonçalves, 2018). Koripako is spoken mainly in Colombia, on the upper Içana river, the Guainia river (one of the headstreams of Rio Negro), and the headwaters of the Cuiari (a tributary of the Içana river). Baniwa is spoken in Brazil, on the middle Içana river between Assunção mission and Matapi, as well as on the Aiari and Cuiari tributaries. Karotana used to be spoken on the lower Içana river south of the Assunção mission until the mid-20th century. Some of the groups who spoke this dialect have shifted to Nheengatu/Língua Geral (a Tupí-Guaraní language functioning as a local lingua franca) or to Portuguese. One group migrated in the 19th century to Victorino on the Guainia river in Venezuela, the only place where the Southern dialect is still spoken today. A map of the geographical extension of each dialect is shown in Figure 1.1. It is the linguistic variety known as Baniwa that will be the focus of this thesis.

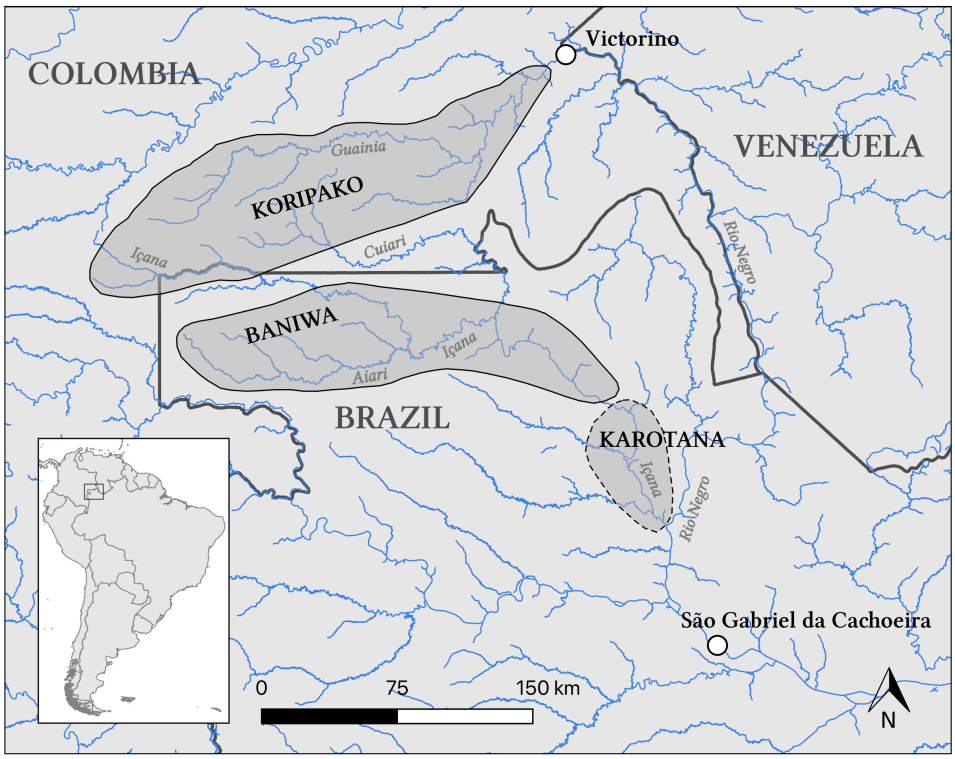

Figure 1.1: Geographical extension of Baniwa-Koripako dialects (based on Ramirez, 2020a: 39)

[3] An alternative dialectal division is based on the isoglosses for different variants of the words for 'yes' (*aha, ehe, oho*) and 'no' (*kori, khenim, karo, ñame*) (Granadillo, 2006). The name *Koripako* itself is an example of this: *kori pa-áako* /NEG IMPRS-to.say/, lit. 'those who say *kori* (for 'no')'.

According to the territorial and environmental management plan of the Baniwa-Koripako people (PGTA Nadzoeri, 2021), Baniwa-Koripako speakers use the self-denomination *Medzeníakonai*, which refers to them as speakers of the original language. Wright & Andrello (2002) further note that Baniwa-Koripako speakers commonly use the name of their phratry as a collective self-denomination. However, the speakers I worked with in São Gabriel da Cachoeira (see § 1.1.2.1, § 1.2.3) typically present themselves as speakers of Baniwa (often in contrast to speakers of Koripako), in line with Wright & Andrello's (2002) statement that the name Baniwa[4] is used "in multiethnic contexts or to the non-indigenous world" (see also Garnelo, 2007: 209).

All dialects of the Baniwa-Koripako continuum are mutually intelligible and share around 90% of their vocabulary (Aikhenvald, 2007: 476; Ramirez, 2020a: 46). The dialectal situation is complicated by the fact that marriages between speakers of different dialects are common. Traditionally, marriages have taken place according to clanic exogamy, that is, between people of different phratries. The phratries consist of different sibs which in turn are associated at least to some degree with different dialects, meaning that there is constant contact between dialectal varieties, with the possible effect of diffusion of dialectal features (Wright, 1981: 28 ff.; Ramirez, 2001a: 2).

Baniwa is a relatively large language in the Amazonian context. Estimates of the number of speakers vary, and are partly complicated by several factors. One is the difficulty of distinguishing between ethnic population and speaker population in some reports. Another is the difficulty of knowing which dialects are included in the estimate. A third complicating factor, related to the second one, is the fact that the Baniwa-Koripako continuum is spoken in three different countries. In any case, the number of speakers appears to be in the thousands. Aikhenvald (2007: 476) estimates a total of 3,000–4,000 speakers of Baniwa-Koripako. Ethnologue lists 6,420 speakers of Baniwa [bwi] and 12,070 of Koripako [kpc], amounting to 18,490 in total (Eberhard, Simons & Fennig, 2024). Crevels (2012: 182, 195, 220) estimates 16,554 speakers, whereof 5,811 in Brazil. The publication *Povos Indígenas no Brasil* reports an ethnic population of 7,145 in Brazil based on 2014 statistics from the Secretariat for Indigenous Health (SESAI), but does not mention anything about the number of speakers of the language (Ricardo, Klein & Santos, 2023).

Baniwa enjoys co-official status with Tucano (East Tucanoan), Nheengatu (Tupí-Guaraní), and Portuguese in the municipality of São Gabriel da Cachoeira (see

[4]The name *Baniwa* itself is an exonym that has been used for several Arawakan-speaking groups in the same region, which may go back to the Nheengatu term *maniyua* 'bitter manioc' (Stradelli, 1929: 513; Aikhenvald, 2012: 21; but see Floyd, 2013: 287).

further in § 1.1.2.1). Despite this, Crevels (2012: 182, 195, 220) classifies Baniwa as "potentially endangered". Ethnologue classifies Baniwa [bwi] as "in trouble" and Koripako [kpc] as "developing" in the Language Cloud (Eberhard, Simons & Fennig, 2024). It is unclear to what extent the language is being transmitted to children across the speech community (but see § 1.1.2.1 for the situation in the city of São Gabriel da Cachoeira). Speakers generally seem to be of the opinion that the Baniwa use Portuguese more if they are 1) younger, and 2) live in the city. Indeed, the dialectal variety whose speakers have largely shifted to Portuguese or Nheengatu is the one traditionally spoken closest to the city (the Southern dialect/Karotana, see above).

The speakers of Baniwa traditionally live close to the major rivers and practice small-scale agriculture with bitter manioc (*Manihot esculenta*) as the staple crop. Manioc tubers are prepared for consumption through a process of fermentation including scraping, washing, grating, squeezing, and cooking, resulting in *farinha* (roasted manioc flour) or *beiju* (large, circular flatbreads made of manioc). Other subsistence activities include gathering, fishing, and hunting.

Baniwa speakers are part of a regional exchange system in the Upper Rio Negro region, where the social interaction and exchange are built around language, marriage patterns, and subsistence practices (Epps, 2020: 276–280). Many groups in this region, in particular in the Vaupés, practice linguistic exogamy (see further in § 2.7.2)—Baniwa-speakers have only been part of this practice to a marginal extent through the intermarrying of some groups with speakers of Kubeo (East Tucanoan, Gomez-Imbert, 1996: 445–447), as the predominant marriage pattern of Baniwa-Koripako speakers has been one of clanic exogamy (see above). The exchange system also features a division of labor, where each group occupies its own specialized niche: for example, the predominantly foraging Naduhup- and Kakua-Nɨkak-speaking groups provide hunted meat and forest products to the riverine, agricultarist Tucanoan- and Arawakan-speaking groups in exchange for agricultural products. The division of labor also extends to crafted goods, which are typically connected to specific groups: Tuyuka speakers make canoes, Tucano speakers make benches, Naduhup speakers make *aturás* (large baskets for carrying manioc), and Baniwa speakers make *ralos*, or manioc graters: flat boards with small, sharp pieces of quartz stones (Epps & Stenzel, 2013: 21; Epps, 2020: 277–278). As (Epps, 2020: 277) states, "the distinctions among groups facilitate their functioning together as a system of interdependent, complementary parts".

In the 20th century, many Baniwa speakers worked in the extraction of forest resources like rubber and Brazil nuts, often under slavery-like conditions. They also began producing baskets for commercial purposes, and are still well-known

for their basketry skills. Today, many work in the fields of education and health care, and the local economy is becoming more and more monetarized (Wright, 1981: 15–28; da Silva, 2021: 224).

1.1.1 Genealogical context

Baniwa is part of the Arawakan language family (also known as Maipurean), which constitutes the largest language family in South America by number of languages. There are around 40 living Arawakan languages, spoken across a vast area comprising both South and Central America (historically, Arawakan languages were also spoken in the Carribean). The distribution of the Arawakan family in South America is shown in Figure 1.2. The geographical spread of the Arawakan languages is associated with a number of salient cultural features often referred to as the "Arawakan matrix", including farming techniques, ceramic traditions, ritual practices involving sacred musical instruments, a social organization based on consanguinity, and the suppression of warfare with linguistically related groups, which has been adopted by neighboring non-Arawakan groups to varying extents and likely contributed to the expansion of the language family (Santos-Granero, 2002; Hornborg, 2005; Eriksen, 2011). The Arawakan homeland is hypothesized to be in Northwestern Amazonia (Zucchi, 2002; Hornborg, 2005).

Figure 1.2: Geographical distribution of Arawakan languages in South America (map created by Davius, released to the public domain)

The linguistic relatedness of the Arawakan languages was recognized by Father Gilij as early as 1783, predating Sir William Jones's recognition of the relatedness of the Indo-European languages by three years (Aikhenvald, 1999: 73). Despite the long history of comparative and historical study, there is much disagreement on the internal classification of the Arawakan language family. This is due in part to an insufficient level of description for individual languages, but also to problems of teasing apart genealogical and areal effects (Aikhenvald, 1999; Dunn, 2022). Aikhenvald (1999) proposes a primary split into a Northern and a Southern branch. Ramirez (2020b; 2020c) presents a different classification where Proto-Arawak splits into 12 different subgroups (Figure 1.3[5]). Ramirez's classification has been accepted in a number of publications and databases by other researchers (Michael, 2021; Cayón & Chacon, 2022; Hammarström et al., 2024), and will be adopted in this study as well. In Ramirez's classification, Baniwa belongs to the Japurá-Colombia subgroup together with 12 other languages[6].

The closest relative of Baniwa is Tariana, a nearly extinct language reported in Ramirez (2001b) to be so closely related that the languages are mutually intelligible, and in Ramirez (2020a) considered a dialect of the Baniwa-Koripako continuum. This is contested by Aikhenvald (2007: 476), who claims that there is no mutual intelligibility. In any case, there is a close genealogical relationship between the two varieties, although Tariana has been heavily restructured due to contact with East Tucanoan languages (see § 2.7.2).

In terms of their general typological profile, Arawakan languages can be characterized as highly synthetic, predominantly agglutinating (mostly suffixing) and mostly head-marking. Nominal morphology is fairly uniform throughout the family, whereas verbal morphology displays a higher degree of variation and can be very complex (Aikhenvald, 1999: 65, 80–81).

1.1.2 Geographical context

Baniwa is spoken in the Upper Rio Negro region, a tropical rainforest environment known for its linguistic diversity (Epps & Stenzel, 2013). Languages of five different families (Naduhup, East Tucanoan, Arawakan, Kakua-Nɨkak, and Carib)

[5] In the family tree, extinct languages are markes with a dagger (†) and nearly extinct languages are marked with a superscript dagger ([†]).

[6] These are Warekena [ware1258?], †Mandahuaca [mand1448], Piapoco [piap1246], Achagua [acha1250], Cabiyarí [cabi1241], Yucuna [yucu1253], †Kaishana [kais1242], [†]Resígaro [resi1247], †Wainumá-Mariaté [uain1239], †Jumana [juma1250], †Passe [pass1250], and †Mepuri [mepu1234] (Ramirez, 2020b: 17). Recall that Ramirez (Ramirez, 2020a treats Tariana as a dialect of Baniwa (§ 1.1).

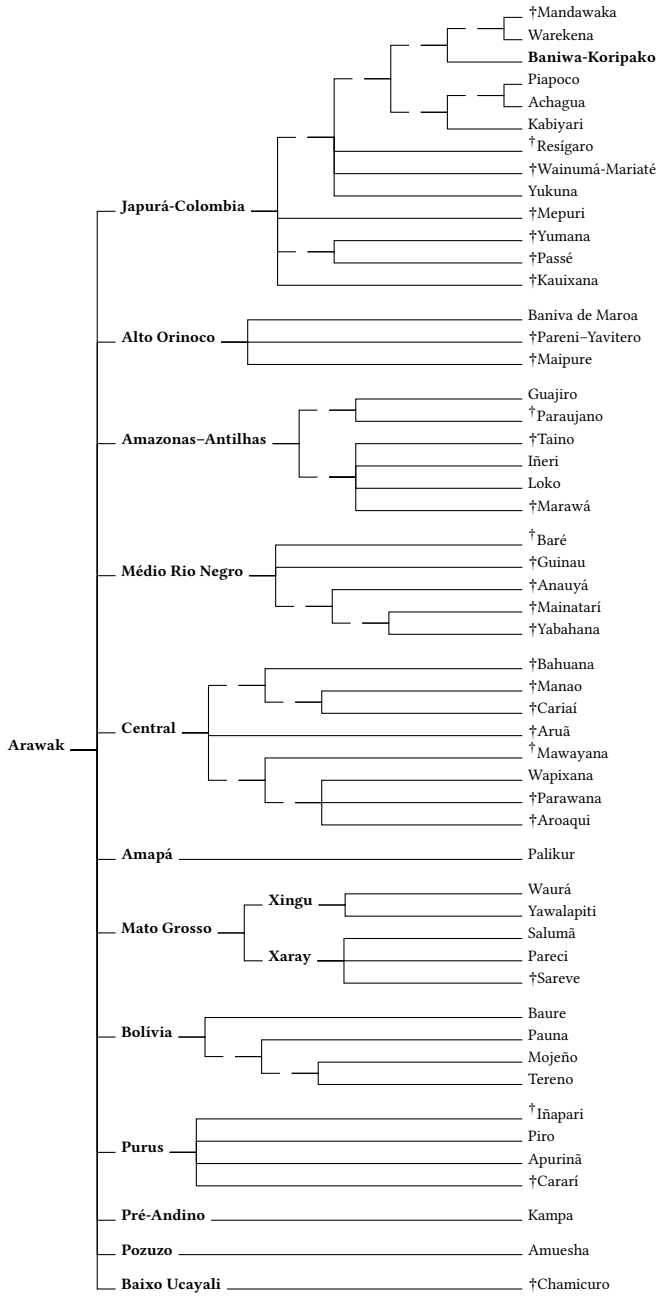

Figure 1.3: Classification of the Arawakan languages (adapted from Ramirez, 2020a: 43)

have a long co-presence in the region. In addition, the local lingua franca Nheengatu (Tupí-Guaraní) has gained ground in the past centuries, as have Portuguese

(on the Brazilian side) and Spanish (on the Colombian side). The Upper Rio Negro region, like many other multi-lingual regions, is characterized by a seemingly paradoxical pattern: simultaneous diversity and similarity. Some aspects of language and culture have diffused throughout the region, while others are kept strictly apart, in a dynamic that has intrigued many researchers in the fields of linguistics and anthropology (Sorensen Jr., 1967; Hugh-Jones, 1979; Aikhenvald, 1996a; Stenzel, 2005; Epps, 2018).

The Upper Rio Negro region, in turn, is part of a larger area loosely referred to as Northwestern Amazonia, as seen in Figure 1.4, where the Upper Rio Negro region is outlined in dark gray (Cayón & Chacon, 2022: 2).

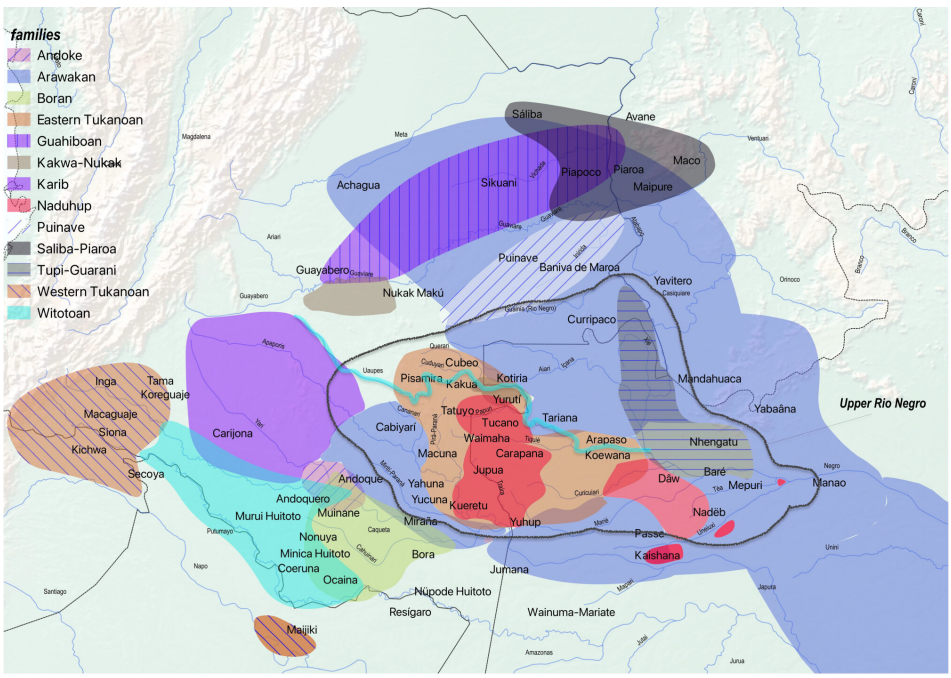

Figure 1.4: Language families in Northwestern Amazonia (map from Cayón & Chacon, 2022: 2)

The region of Northwestern Amazonia stretches over adjacent areas in Brazil, Colombia, Ecuador, Peru, and Venezuela. In addition to the language families present in the Upper Rio Negro region, Northwestern Amazonia is typically understood to comprise languages from the Andoke, Boran, Guahiboan, Peba-Yaguan, Puinave, Saliba-Piaroa, West Tucanoan, and Witotoan lineages. The speakers of these languages have been in contact with each other in various constellations over long periods of time, and share a number of cultural traits relating to, e.g., ritual practices and social organization (Cayón & Chacon, 2022). A widespread linguistic phenomenon in Northwestern Amazonia is the existence of

nominal classification systems of the *multifunctional* kind (Krasnoukhova, 2012). These are described in more detail in § 2.7.

1.1.2.1 Setting of study: São Gabriel da Cachoeira

The data for this thesis were collected in the small town of São Gabriel da Cachoeira, which was established in 1668 by jesuits and has a population of about 17,000 people, 85% of whom are of Indigenous descent[7] (Shulist, 2018a). The town is therefore frequently referred to as "the Indigenous capital". The municipality with the same name is the third largest in Brazil in terms of geographical area but very sparsely populated. It is known as *Cabeça do Cachorro* 'the dog's head' for the shape outlined by the national borders to Colombia and Venezuela.

São Gabriel da Cachoeira is located on the northern shore of the Rio Negro, just south of the equator, and marks the farthest point on the river that can be accessed by regular boat traffic; areas upstream are only accessible by canoe or motorboat due to the many rapids. The town also hosts a small airport with flights to and from Manaus (the state capital of Amazonas, situated at about 850 km of distance from São Gabriel da Cachoeira). Due to its remote location in the middle of the Amazon rainforest, the town cannot be accessed via road.

São Gabriel da Cachoeira is a true melting pot: the Indigenous people are of 23 different ethnicities, speaking more than 20 different languages of the Arawakan, Tucanoan, Naduhup, and Yanomami language families (Cabalzar & Ricardo, 2006). In addition, the Tupí-Guaraní language Nheengatu is widely spoken throughout the region, as is Portuguese. In 2003, three Indigenous languages of São Gabriel da Cachoeira were granted co-official status with Portuguese: Baniwa, Nheengatu, and Tucano. This was the first promotion of any Indigenous language to official municipal status in Brazil (Crevels, 2012: 181).

While São Gabriel da Cachoeira can be considered remote in a nation-wide context, both in terms of geographical location and economic and political power and influence, it functions as a regional urban center, and is perceived as standing in contrast to the more rural communities within the municipality (Shulist, 2018b: 22–29). In terms of language, rural communities in the area are generally characterized by a stronger vitality of Indigenous languages as well as a higher degree of linguistic homogeneity compared to the city (but see § 2.7.2 for a brief depiction of the situation in the Vaupés). The associations of particular languages with particular territories in the rural areas stand in opposition to the Portuguese-

[7] For comparison, the Indigenous population in Brazil as a whole is 0.83%, according to the 2022 census (Instituto Brasileiro de Geografia e Estatística, 2023).

dominant city, which "has been marked as non-Indigenous (Portuguese) space" (Shulist, 2018b: 28; see also Epps, 2018).

Nevertheless, compared to many other Indigenous groups in the city, the Baniwa-speaking population of São Gabriel appears to be retaining and transmitting their language to children to a larger extent. Shulist (2018b: 47–48, 158–159) mentions two possible reasons for this. First, they are among the more recent migrants to the city area. Second, evangelism has had different effects on different Indigenous groups in the region depending on the Christian orientation of the organization. The evangelical protestant missionaries among the Baniwa-speaking groups on the Içana and the Catholic missionaries among Tucanoan-speaking groups on the Vaupés both largely prohibited the use of many cultural and religious practices, but a key difference was that the evangelical missionaries did not forbid the use of the local language, instead translating the New Testament into Baniwa. One result of this is that Baniwa is the only local language that has a relatively accepted standardization (see § 3.1.4). Another result, according to Shulist, is that Baniwa-Koripako speakers feel more comfortable using their language in formal contexts, e.g., in the church and schools.

1.1.3 Previous work on Baniwa

The earliest sources on Baniwa-Koripako consist of a number of word lists from the second half of the 19th century and the first half of the 20th century, of which the most notable is that of Nimuendajú (1932). In the second half of the 20th century, a number of partial language descriptions appeared: Scheibe (1957) on the phonology of Baniwa, Boley & Boley (1979) and Taylor (1991) on the grammar of Baniwa, and Valadares (1993) on the phonology and morphology of Koripako. Ramirez has conducted the most comprehensive work on the language, resulting in a grammar (2001b)[8] and dictionary (2001a) of Baniwa. Since then, a number of further publications have appeared, including Bezerra (2005) on the grammar of Koripako, Granadillo (2006) featuring an ethnographic account of Koripako language documentation, Aikhenvald (2007) on Baniwa-Koripako classifiers (see further below), and de Souza (2012) on the phonology of Baniwa-Koripako. For a more detailed account of the general history of study of Baniwa, see Ramirez (2020a: 46–48).

A number of anthropologists have conducted ethnographic work on Baniwa-Koripako-speaking groups. Robin M. Wright has worked primarily with Baniwa

[8]The grammar was edited and republished as part of a four-volume encyclopedia on the Arawakan languages (Ramirez 2020a; 2020b; 2020c; 2020d), which also includes historical and comparative work on the family.

speakers on the Aiari river, and has published extensively on topics relating to history, religion, cosmology, shamanism, and medicine (Wright 1981; 1992a; 1992b; 1998, etc.). Jonathan D. Hill worked mainly with Koripako speakers in Venezuela and Colombia, his publications spanning topics such as mythology, musicology, ritual, and ethnicity (Hill 1984; 1993; 2009a, etc.). Both Wright and Hill have large data collections archived in AILLA (The Archive of the Indigenous Languages of Latin America) (Wright, 2020; Hill, n.d.).

The classifier system of Baniwa has been the subject of some previous publications, most notably those by Aikhenvald (1996b; 2007) and Ramirez (2001b; 2020a). Aikhenvald has described the classifier system of Baniwa-Koripako in two articles (1996b; 2007) as well as in a few other publications with a comparative perspective (e.g., 2019). In her (1996b) paper, Aikhenvald analyzes Baniwa as having separate classifier and noun class systems, although they consist of almost the same forms. She later unites them into a single system marked in different morphosyntactic contexts, and proposes a division of the classifiers into four sets, based on a mixture of semantic and morphosyntactic criteria (2007). The most extensive description to date of the Baniwa classifier system is found in the grammar by Ramirez (2001b), later re-published in an edited version (Ramirez, 2020a). Some complementary information on classifiers can be found in Ramirez's dictionary (2001a). Ramirez's study provides a relatively thorough account of the classifier system, bringing up aspects of morphosyntax, semantics, and assignment.

Apart from the studies by Aikhenvald and Ramirez, a number of other publications deal with aspects of the Baniwa-Koripako classifier system. Hill (1988) analyzes the classifier system and its connection to ritual and myth in a number of dialects spoken in Venezuela (see further in § 7.3.3.3). Taylor (1991: 39–45, 121–133) presents a non-exhaustive inventory of 40 classifiers and identifies some of the morphosyntactic contexts in which they appear (such as on numerals and adjectives). Other publications include those by Gonzalez-Ñáñez (1985), Baltar (1995), Gomez-Imbert (Gomez-Imbert, 1996) and Melgueiro (2009).

The previous analyses differ in several respects: for example, in their classifier inventories, in the phonological forms of classifiers, in the morphosyntactic contexts they consider classifiers to appear in, in their semantic descriptions, and in the distribution of allomorphs across the different morphosyntactic contexts. An example of differing analyses is shown in Table 1.1, where the inventories of classifier forms in three publications are compared (spelling unaltered): Aikhenvald (2007), Ramirez's grammar (2001b), and Ramirez's dictionary (2001a)—which presents different inventories in the table on page 21 and in the

actual dictionary entries[9]. A total of 58 different morphemes are identified as classifiers across these publications, the individual inventories varying in size. As can be seen for some classifiers in the list, the analyses also differ in what they count as allomorphs of the same classifier.

1.2 This study

1.2.1 Aims and scope

The previous literature (summarized in § 1.1.3) forms a valuable basis for the study of classifiers in Baniwa. Especially Ramirez (2001b; 2020a) and Aikhenvald (2007) have been instrumental in paving the way for the current study. Certain aspects of the system, such as its morphosyntactic and semantic properties, have been described in earlier studies (Aikhenvald, 2007; Ramirez 2001b; 2020a), although the analyses differ in several respects. Other aspects have received less attention. These include, for example, the functions of the classifier system, the historical origins and age-depth of classifiers[10], language contact phenomena, as well as in-depth analysis of the semantic organization. No study so far has provided a comprehensive account of the classifier system and its various properties. This thesis aims to fill that gap.

Thus, as the title suggests, the main aim of this thesis is to provide a comprehensive descriptive account of the properties and workings of the classifier system in Baniwa: a "grammar" of classifiers, speaking to the holistic approach that I am adopting. The classifier system is explored from a number of different angles, corresponding to individual chapters: formal properties (including phonology and morphosyntax) (§ 5), functions (§ 6), semantic properties (§ 7), historical development (§ 8), and language contact (§ 9).

A further aim of this thesis is to put Baniwa classifiers in a typological perspective, and to contribute to the theoretical and typological understanding of nominal classification systems. Given the geographical and genealogical position of

[9] Among the classifiers that are lacking in Aikhenvald's (2007) inventory, three may be due to the fact that she does not appear to distinguish vowel length in the language. The three classifiers in question differ from other classifiers only in vowel length in certain contexts, such as on the numeral *apa-* 'one': *-ma*^{1} (lacking) vs. *-ma*^{2} (*apáma* vs. *aapáma*), *-daa* (lacking) vs. *-da* (*apádaa* vs. *apáda*), and *-aana* (lacking) vs. *-na* (*apáana* vs. *aapána*).

[10] Ramirez (2020b) has conducted extensive historical work on the Arawakan language family, including on the Japurá-Colombia subgroup which Baniwa belongs to, but this has been presented from the point of view of the subgroup as a whole rather than from the perspective of present-day Baniwa.

Table 1.1: Comparison of classifier inventories in previous descriptions

Aikhenvald (2007)	Ramirez (2001b)	Ramirez (2001a: 21)	Ramirez (2001a: 31–337)
-Ø, -a, -a.ɻi	–	–	–
-Ø , -ya.ɻi	*-Ø(-.Ii)*	*-Ø(-.Ii)*	–
–	*-áana*	–	*-áana*
-ahna(y)	*-ánhaa, -áanhai*	*-ánhaa, -áanhai*	*-áanhaa*
-apa(-.ɻi/-ʒu)	*-aápa(-.Ii)*	*-aápa(-.Ii)*	*-aápa*
-aphi	*-áaphi(-i)*	*-áaphi(-i)*	*-áaphi*
-api	*-áapi(-i)*	*-áapi(-i)*	*-áapi*
-apu	*-áapo*	*-áapo*	*-áapo*
-da(-.ɻi)	*-da(-.Ii)*	*-da(-.Ii)*	*-da*
–	*-daa, -da-.Ii*	–	*-daa*
-dapana	*-dápana*	*-dápana*	*-dápana*
-i	*-e/-i(-.Ii)*	*-e/-i*	*-i*
–	*-éekhe*	–	–
-ima(.ɻi)	*-eéma(-.Ii)*	*-eéma(-.Ii)*	*-eéma*
-ihwe	*-éewhe, -da-.Ii*	–	–
-hiku	*-híko, híki*	*-híko, híki*	*-híko*
-hipa, -ite, -da-.ɻi	*-hípa, -íite, -da-.Ii*	–	*-hípa*
-hipada	*-hipáda*	*-hipáda*	*-hipáda*
-hipani	*-hípani*	–	*-hípani*
–	*-híwa(-.Ii)*	*-híwa(-.Ii)*	*-híwa*
-hiwi	*-híwi(-i)*	*-híwi(-i)*	–
-ida	*-íida*	–	*-íida*
-iʃi	*-íijhi, -da-.Ii*	–	–
–	–	–	*-íiniri*
-ita, -ite, -da-.ɻi	*-iíta, -íite*	*-iíta, -íite*	*-iíta*
-itʃia	–	–	–
–	*-íiwi*	*-íiwi*	–
-ya(.ɻi)	*-ja*	*-ja*	–
-yawa	*-jawa*	–	–
-ʃa(.ɻi)	*-(j)haa*	*-jhaa*	–
–	*-kénaa*	*-kénaa*	–
-kha, -khay	*-khaa(-i)*	*-khaa(-i)*	*-khaa*
-ku,-ki	*-ko(-i)*	*-ko(-i)*	*-ko*
-kwa, -kwe	*-koa(-i)*	*-koa(-i)*	*-koa*
–	*-kódzoa*	–	–
–	*-ma*[1]	–	*-ma*
-ma, -da-ʒu	*-ma*[2]*; -da-.ɻo*	–	*-ma*
-maka, -make	*-máka, -máki*	*-máka, -máki*	*-máka*
-na, -ne, -nay	*-na(-i)*	*-na(-i)*	*-na*
-naku, -naki	*-náko(-i)*	*-náko(-i)*	*-náko*
-pa(.ɻi)	*-pa(-.Ii)*	*-pa(-.Ii)*	–
-pawa(ni)	*-pawa*	–	–
-peku, -peki	*-péko(-i)*	*-péko(-i)*	*-péko*
-peni	–	–	–
-peʒi	–	–	–
-phe, -phay	*-phe, -phai*	–	–
-pi	*-pi*	–	*-pi*
-puku, -puki	*-póko(-i)*	*-póko(-i)*	*-póko*
–	*-pokóda*	*-pokóda*	*-pokóda*
-tawah.ɻe	–	–	–
-tuwhya	*–ttówhia*	*-tówhia*	–
–	*-tsoi*	–	–
-wa	*-wa(-.Ii)*	*-wa(-.Ii)*	*-wa*
-wa.ɻi	*-wá.Ii*	–	–
–	*-wá.Ihia*	*-wá.Ihia*	*-wá.Ihia*
-wana, -wane	*-wána*	*-wána*	*-wána*
-wata, -wate	*-wáta*	*-wáta*	–
–	*-wáthe*	*-wáthe*	–
45 classifiers	52 classifiers	35 classifiers	34 classifiers

Baniwa, the current description feeds into the body of literature on nominal classification systems in Northwestern Amazonian languages (§ 2.7) on the one hand, and Arawakan languages (§ 2.8) on the other. It also informs the study of nominal classification more generally. An additional dimension of the present study is that it is situated in an urban context, allowing for the exploration of questions relating to language contact, cultural change, and majority language influence, and how these are reflected in classifier systems.

Last but not least, an additional aim of this thesis is to contribute to the description, documentation, and linguistic analysis of the Baniwa language.

This thesis is based on first-hand data from original field work; mostly data from targeted elicitation and speaker interviews, but also on naturalistic texts. The methodological choice to collect first-hand materials was made for several reasons. First, the descriptive nature of the work required first-hand contact with speakers, spending time and studying the language in the environments where it is spoken. Second, the data collection was necessary due to the scarcity of available data.

In some cases, second-hand data produced by other authors have also been taken into account. In Chapter 8 (Diachrony), I have relied on previous historical work on Arawakan languages by other scholars when it comes to reconstructions at earlier stages. Chapter 3 (Grammatical preliminaries) draws heavily on the linguistic analysis by Ramirez (2020a). Throughout the work with this thesis, I have had the privilege to rely of Ramirez's comprehensive dictionary (2001a) and grammar (2020a) as reference works .

1.2.2 Data collection

The first-hand data for this thesis were collected on three occasions: during a first pilot trip to Brasília and two later field trips to São Gabriel da Cachoeira (§ 1.1.2.1), as summarized in Table 1.2.

Table 1.2: Occasions for data collection

Location	Time
Brasília	March–April 2019
São Gabriel da Cachoeira	February–March 2020
São Gabriel da Cachoeira	September–October 2022

During the first trip to Brasília, I came into contact with colleagues who were studying in the linguistics and anthropology graduate programs at the University

of Brasília at the time, and who were native speakers of Baniwa. During this trip, some preliminary data were collected, primarily in the form of elicited materials and semi-structured picture-book narratives. During the first trip to São Gabriel da Cachoeira, the most extensive data collection was carried out. This included the collection of the Free listing and Noun list data sets, the recording of both semi-structured and unstructured narratives, and plenty of additional elicitation. The final field trip to São Gabriel da Cachoeira was mostly devoted to following up on the previous data collection, and mainly comprised structured interviews and elicitation.

An account of the individual data types and the methods used for collecting them is given in § 1.2.4–1.2.5. The sessions were all recorded in audio, most of them also in video. All data, both recordings and additional files, are archived online at the Lund University Humanities Lab Archive. The data collection is available on request at https://hdl.handle.net/10050/84e2a81b-501a-41a9-a59d-9c62becbc6a1.

1.2.3 Participants

This thesis is based on linguistic materials from 12 native speakers of Baniwa. Table 1.3 presents some basic information about the them, as well as the types of data each speaker provided: naturalistic texts (NT), Free listing (FL), Noun list (NL), and general elicitation (GE).

Table 1.3: Overview of participants

Speaker information						Data types			
ID	**Age**	**Sex**	**Languages**	**Place of birth**	**Occupation**	**NT**	**FL**	**NL**	**GE**
A	74	M	B (P, Kr, N)	Tunuí Cachoeira	agriculture	–	✓	–	✓
B	45	F	B (Kb, P, S)	Canadá	agriculture	✓	–	–	–
C	70	F	B (P (Kr))	Ambaúba	agriculture	–	–	–	✓
D	34	F	B (P, N)	São Gabriel	agriculture	–	–	–	✓
E	25	M	P (B)	São Gabriel	education	✓	✓	–	✓
F	45	M	B (P, S, Kr)	Tukumã Rupitã	education	–	✓	✓	✓
G	41	M	B (P, Kr, S (N))	Tukumã Rupitã	education	–	✓	–	✓
H	36	M	B (P, Kr (N, S))	Tunuí Cachoeira	education	–	✓	–	–
I	43	F	B (P (N))	São Gabriel	education	–	–	–	✓
J	80	F	B (P, N)	Nazaré	agriculture	✓	–	–	–
K	45	F	B (P, Kr (N))	Tunuí Cachoeira	education	–	✓	✓	✓
L	45	M	B (P)	Aracu	agriculture	–	–	–	✓

Out of the speakers, six were female and six male, their ages ranging between (roughly) 25 and 80 at the time of data collection. Almost all the speakers I worked with resided permanently in different neighborhoods in São Gabriel da

Cachoeira or in smaller communities in the outskirts of town, and had done so for 10 years or more, typically in households with three generations. While Shulist (2018b: 48) writes that Baniwa speakers are "more likely than nearly any other group [in the city] to live in close proximity to their own", none of the families I worked with lived within walking distance of each other. Two speakers (H and L) lived in communities on traditional Baniwa territory (Tunuí Cachoeira and Aracu, respectively), and were interviewed during temporary stays in town. The participants all work either in agriculture (small-scale cultivation in *roça* manioc gardens, processing of manioc products such as *beiju* (manioc bread) and *farinha* (roasted manioc flour), and/or selling agricultural products in markets or along rivers) or in education (including school teachers, school administrators, and students in higher education). Even in town, many families have their own manioc gardens, although some family members may also have other jobs.

The speakers were asked to list their strongest language, any additional languages they speak (given in parentheses), and any additional languages they understand to some degree (given in double parentheses). All speakers were at least bilingual in Baniwa and Portuguese (at least in spoken form). They all listed Baniwa (B) as their strongest language except the youngest speaker, who listed Portuguese (P). Besides these two languages, some participants also speak or understand Nheengatu (N, Tupí-Guaraní, seven speakers), Koripako (Kr, six speakers), Spanish (S, four speakers), and Kubeo (Kb, East Tucanoan, one speaker)[11]. All participants lived in households that are primarily Baniwa-speaking, although not exclusively in all cases; some of the households also consist partly of speakers of East Tucanoan languages (Tucano, Kubeo) by marriage. In such cases, it is common to rely on Portuguese as a common language in the interaction with non-Baniwa-speaking family members. All participants also stated using both Baniwa and Portuguese in their daily life outside their home, to varying degrees. Despite the fact that those who work in the educational sector are involved in Indigenous educational programs, they use Portuguese (both written and spoken) to a larger extent than those who work in the agricultural sector, due to Portuguese being the language of administration. Those who work

[11]It should be noted that the languages listed here are a matter of self-assessment. Different speakers may have different views on what it means to be able to speak or understand a language, and thus differ in which languages they choose to mention when being asked in this way. The self-perceived level of proficiency probably also depends on what they use the different languages for. In addition, speakers may differ in what they think of as a "language"; as they all speak Baniwa, presumably more than six of the speakers are in fact able to communicate with Koripako speakers, but perhaps they did not all think of listing it as a separate language (see § 1.1). They same may hold for Spanish to some extent; as all speakers are proficient in Portuguese, it is likely the case that more than four of them understand some Spanish, but perhaps some of them think of this competence as closely linked to their competence in Portuguese.

in the agricultural sector mainly use Portuguese, in spoken form, when selling their agricultural products.

Speakers were recruited based on interest and availability, and the body of participants grew organically through the addition of family members, friends, neighbors, acquaintances, and colleagues of the speakers I had talked to previously, beginning with the speakers I had met in Brasília in 2019 (see § 1.2.2). During field work, I interviewed and collected various types of data from a total of 25 Baniwa speakers. Not all of the data I collected ended up being analyzed for this thesis, which is why the additional 13 speakers are not represented in Table 1.3.

Prior to engaging any participant, informed consent was obtained in oral and written form. I explained my research objectives (in Portuguese), stated what kind of data I was collecting and why, described the means by which the data would be collected (audio and video recording), and how I would like to use and store the materials (in scientific publications and in an online archive). In some cases, one of my main collaborators (see below) assisted me in clarifying some of this information in Baniwa for the participant. The participants all agreed to being interviewed, to being recorded during the time, and for the data to be published and used for research purposes. All recorded speakers also completed a questionnaire to collect comparable information about their sociolinguistic backgrounds (some of which is summarized in Table 1.3). This featured questions about their age, sex, place of birth, the language(s) of their parents, the languages they know besides Baniwa, and their language use in everyday situations. All speakers were compensated for their time and for sharing their knowledge, either with money or goods, according to agreement.

Dialectal differences are not the topic of this thesis, and it also does not focus on any particular dialectal variety of Baniwa. All speakers consulted spoke some version of the variety known as Baniwa (that is, the "Central" dialect of the Baniwa-Koripako language continuum, see § 1.1). The place of birth of each speaker is plotted on a map in Figure 1.5 overlaid with the areas associated with the main Baniwa-Koripako dialects. All speakers except four were born in locations within the Central dialect area. Of these four, three were born in the city of São Gabriel da Cachoeira, which is outside traditional Baniwa territory. Finally, one was born just on the border to the Koripako (Northern) dialect area, but lived in Tunuí Cachoeira within the Baniwa dialect area for a long time and self-identifies as speaking Baniwa.

Three speakers (F, G, and K) functioned as my language teachers and main collaborators during various stages of the work, and were thus involved in the work in a more profound way than the other speakers. The collaborators performed

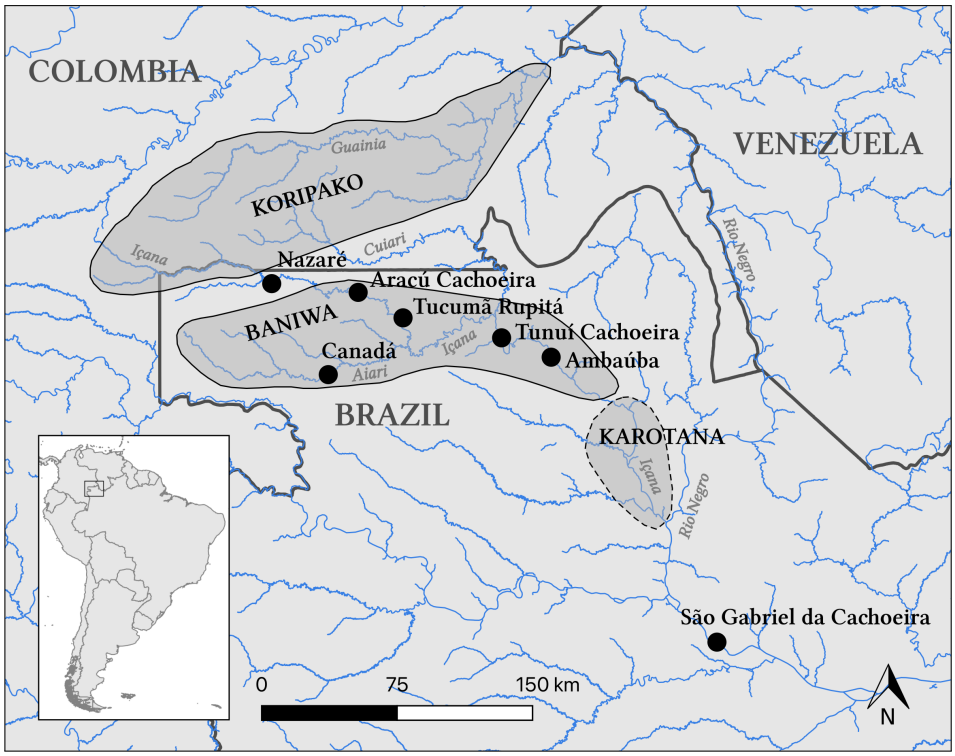

Figure 1.5: Birth locations of participants

tasks such as transcription and translation, and participated in more comprehensive elicitation. They also helped me with many other tasks, such as recruiting speakers, getting around town, and clarifying my research objectives and various instructions in Baniwa whenever necessary.

To summarize, the speaker sample reflects a speech community living in an urban, multilingual environment where Portuguese is used to a much larger extent compared to more traditional settlements (see further in § 1.1.2.1).

1.2.4 Elicited data

This category includes several different types of elicited data, summarized in Table 1.4 along with the approximate number of hours of recordings of each data type[12]. Throughout elicitation, Portuguese has been used as lingua franca.

[12] All data are archived at https://hdl.handle.net/10050/84e2a81b-501a-41a9-a59d-9c62becbc6a1, and is available on request.

Table 1.4: Elicited data types

Data type	Duration of recordings
General elicitation	24 hours
Free listing data set	7 hours
Noun list data set	25 hours

The general elicitation category contains structured interviews using a diverse range of methods, including translation and back-translation, paradigm elicitation, grammaticality judgments, and data manipulation (Bowern, 2015: 90), as well as interviews with speakers about any kinds of questions that come up during the course of field work. These data types have typically been collected in tandem and are intermixed in the same files (for example, grammaticality judgments are often interspersed with interviews about speaker intuitions on usage and semantics).

The Noun list and Free listing data sets are specific data types that warrant more detailed methodological description in terms of elicitation, coding, and analysis. They are described below.

1.2.4.1 Free listing data set

The Free listing data set consists of spontaneously produced lists of nouns for 49 classifiers[13]. This data set takes classifiers as the basis and investigates the internal structure of the categories they represent through the items they combine with. The method of collection and analysis was developed by Franjieh (2012) for possessive classifiers in North Ambrym (Oceanic, Vanuatu) within the framework of prototype theory (see § 2.4).

The data were collected in individual elicitation sessions with six native speakers of Baniwa (see further below). In the experiment, the participant heard a stimulus phrase consisting of the numeral *apa-* 'one' suffixed with a classifier (e.g., *apa-khaa* 'one-CLF.CURVILINEAR') and was asked to list as many items as they could think of that could be counted with that particular classifier-inflected form of the

[13]In the experiment, 48 of the 53 classifiers (see § 4.2) were tested. The four classifiers *-éekhe* 'CLF.SMALL.SEED', *-kódzoa* 'CLF.BEND', *-tsoi* 'CLF.PILE', and *-wálhia* 'CLF.YEAR' are missing, as they had not yet been recognized as part of the inventory at the time of the data collection. In addition, the two classifiers *-ma*^{1} and *-ma*^{2} were accidentally conflated during the preparation of the stimuli, and were presented only once to the participants, some of which interpreted it as the former, some as the latter (only one participant listed examples of both groups, as separate categories). The results were subsequently split into two separate groups, leading to a final data set comprising 49 classifiers.

numeral. The procedure is illustrated below. The participants were also asked to translate each listed item into Portuguese. When a participant was done with the list for one classifier, the procedure was repeated with the next classifier.

- Stimulus: 'one'-CLF
- Response: item[1], item[2], item[3]...

The stimulus was presented audially to the participants[14]. The ordering of the list of classifiers was randomized using an online tool (Haarh, 2020), so that different orders of classifiers were generated for each participant[15]. The responses to the stimuli were given orally by the participants, and recorded in writing by me. All elicitation sessions were recorded in audio, most of them also in video. The notes taken during the sessions were later double-checked against the recordings to establish the lists for each participant and classifier. The running of the experiment took anywhere from 28 minutes to 1 hour and 40 minutes.

The numeral 'one' was chosen as the carrier for the classifier for a number of reasons. First, low numerals (1–3) are one of the contexts in which classifiers are obligatory (§ 5.3), which means that every time *apa-* 'one' is used, a classifier must be selected. Second, in Baniwa noun phrases, numerals precede the noun (§ 5.3.3), which was found to be an advantage for a prompt. Third, numerals were considered to impose less of a semantic restriction on the possible responses than, e.g., adjectives, as adjectives typically carry more meaning and thus run the risk of leading participants to list items with a certain property (for example, using the adjective *maka-* 'big' as a carrier would likely skew the sample towards words denoting big things).

The study was conducted with six participants, presented in Table 1.5 in order of participation in the study. The six participants are all native speakers of Baniwa. Five of the participants are male and one female, their ages ranged from 25–74

[14]In the first three trials of the experiment, I (a non-native, non-fluent speaker) pronounced the stimulus phrases (*apa*-CLF 'one') for each classifier category (version A). This method proved ineffective, as the consultants did not always recognize the targeted classifier. Therefore, for the last three trials of the experiment, sound clips of a native speaker pronouncing the stimulus phrases were played for the consultants instead (version B).

[15]This was done to control for two possible kinds of biases. If the classifiers would have been presented in the same order to each participant, it might have led to a disproportionately high number of items being listed for the classifiers presented first, due to the participants' gradually declining energy levels. The nature of the reported items for a certain classifier category might also be influenced by the classifiers mentioned prior to the classifier in question, through priming effects involving either similarity or contrast. Presenting the classifiers in the same order to all participants could strengthen such effects, and thereby significantly influence the composition of the lists for individual classifiers.

years old at the time of data collection, and they come from different parts of the geographical area where Baniwa is spoken (§ 1.1)[16]. The Stimulus column lists whether the participant was subjected to stimulus version A or B (see above)[17]. The Average column lists the average number of items that the participant listed per classifier. The Blanks column lists how many blank responses the participant gave, i.e., the number of classifiers for which the participant did not list any items at all.

Table 1.5: Participants in the Free listing experiment

ID	Age	Sex	Place of birth	Stimulus	Average	Blanks
K	45	F	Tunuí Cachoeira	A	2.9	8
E	25	M	São Gabriel da Cachoeira	A	0.46	28
F	45	M	Tukumã Rupitã	A	4.54	2
G	41	M	Tukumã Rupitã	B	3.96	0
H	36	M	Tunuí Cachoeira	B	2.14	3
A	74	M	Tunuí Cachoeira	B	1.02	8

The average number of items listed per classifier varies from 0.46 to 4.54. The number of blank responses varies from 0 to 28. There can be a number of reasons for a participant to give a blank response: perhaps the participant could not think of any examples at that moment, or perhaps the participant is unfamiliar with the classifier in question.[18]

The data collection resulted in participant-specific lists of items for each classifier. All lists were ranked individually according to Franjieh's (2012) model, where the ten first items in each individual list received points in descending order (starting at 10 points). Thus, the earlier an item is listed, the higher its score. In Table 1.6, the individual lists for the classifier *-híwi* 'CLF.POINTED' are shown along with the score of each item.

[16] § 1.2.3 contains more information about the sociolinguistic background of the participants.

[17] In most cases, all six participants were subjected to the individual stimulus, but for a few classifiers, not everyone was. This is due to one of the following reasons: a) the accidental conflation of the two classifiers *-ma*^{1} 'CLF.PAIR' and *-ma*^{2} CLF.FEMALE' (explained above), or b) failure by me to pronounce the stimulus phrase well enough for it to be intelligible to the participant (for participants K, E, and F)—because of this, a new method was adopted for the remaining three participants (see above). Six classifiers had less than six participants. These are *-na* 'CLF.TRUNK', *-íixi* 'CLF.SEED', *-éewhe* 'CLF.EGG', *-daa* 'CLF.DAY' (five participants), *-ma*^{2} 'CLF.FEMALE' (four participants), and *-ma*^{1} 'CLF.PAIR' (three participants).

[18] One participant stands out from the rest when it comes to the number of blank responses: the youngest one, E, who gave 28 blank responses (57% of the cases). A possible reason for this discrepancy is that this participant does not have the full inventory of 53 classifiers (see § 4.2) in active use, but instead uses a smaller set. This is supported by the fact that in an entire recorded picture book narrative [1904095_BOY_E], the same speaker uses only the generic classifier *-da*.

Table 1.6: Individual lists for *-híwi* 'CLF.POINTED' by each participant

***-híwi* 'CLF.POINTED'**			
K	**E**	**F**	**Points**
áawi 'needle'	*áawi* 'needle'	*iítsa* 'fish hook'	10
-eétsha 'tooth'		*áawi* 'needle'	9
-íidzo 'strand of fur/hair'		*haikhíwi* 'twig'	8
háiko 'stick'		*kaapáwi* 'arrow'	7
		-eétsha 'tooth'	6
		haikóapo 'stick'	5
		-íidzo 'strand of fur/hair'	4
G	**H**	**A**	**Points**
[*caneta*] 'pen'	*haikhíwi* 'stick'	*áawi* 'needle'	10
[*lápis*] 'pencil'	[*caneta*] 'pen'	*iitsáapo* 'fishing rod'	9
-íiwi 'thorn, spur'	*tápoa* 'nail'		8
áawi 'needle'	*iítsa* 'fish hook'		7

After the score of each listed item had been established, the individual lists were compiled into collective lists with what is assumed to be the most prototypical meanings of each classifier. The use of collective lists across several participants (as opposed to lists based on single speakers) is taken to aid in evening out idiolectal and chance effects. The collective list for *-híwi* is shown in Table 1.7, where *áawi* 'needle' is the most prototypical item (46 points), followed by [*caneta*] 'pen' (19 points), etc.

Table 1.7: Collective list for *-híwi* 'CLF.POINTED'

***-híwi* 'CLF.POINTED'**	
Listed items	**Points**
áawi 'needle'	46
[*caneta*] 'pen'	19
haikhíwi 'stick, twig'	18
iítsa 'fish hook'	17
-eétsha 'tooth'	15
-íidzo 'strand of fur/hair'	12
[*lápis*] 'pencil'	9
iitsáapo 'fishing rod'	9
tápoa '(metal) nail'	8
-íiwi 'thorn, spur'	8
háiko 'stick'	7
kaapáwi 'arrow'	7
haikóapo 'stick'	5

The collective lists for each classifier can be found in Appendix B, along with

a description of the response types as well as detailed description of how the compilation was conducted. All material relating to this data set is available on request at https://hdl.handle.net/10050/84e2a81b-501a-41a9-a59d-9c62becbc6a1 (including the recordings and the individual lists).

1.2.4.2 Noun list data set

The Noun list data set consists of a list of 653 nouns and the classifiers that they are able to combine with. It takes the individual noun as its starting point, and explores classifier–noun combinability from that perspective. 39 of the 53 classifiers appear in this data set.

Each noun in the Noun list data set contains information on the classifiers that the lexeme in question can be used with. A simplified example to illustrate the structure of the data set is given in Table 1.8, where *dáapa* 'paca (rodent sp.)' is shown to be compatible with *-da* 'CLF.GENERIC', *dzáawi* 'jaguar' with *-na* 'CLF.TRUNK', and *áatti* 'chili pepper' with both.

Table 1.8: Simplified example of lexeme–classifier combinability coding

Lexeme	**Meaning**	***-da***	***-na***
dáapa	'paca'	✓	
dzáawi	'jaguar'		✓
áatti	'chili pepper'	✓	✓

The Noun list data set was collected as part of a broader word list elicitation also including other parts of speech. The word list was compiled of five previously established word lists from different sources: The Intercontinental Dictionary Series (Key & Comrie, 2015a), The World Loanword Database (Haspelmath & Tadmor, 2009), the Swadesh 207 list (Swadesh, 1950), and two versions of the South America culture list from The Diachronic Atlas of Comparative Linguistics (Carling, 2024). These were translated from English to Portuguese, except for the Intercontinental Dictionary Series list, which was already available in a Portuguese version (Key & Comrie, 2015b). The word lists were selected in order to cover a broad spectrum of semantic domains, including both basic and culture-specific vocabulary.

Translations of the meanings were elicited orally in recorded sessions. Two native speakers took turns participating in the sessions. Multiple translations per meaning were accepted, as were entries corresponding to more than one meaning, as well as other lexical items that came up in connection with the items in the list. This resulted in an initial list of ca 1800 lexical items in Baniwa. All

entries were coded for word class with reference to Ramirez (2001a; 2001b), and only free nouns, bound nouns, and compounds (§ 3.2.5) were retained, resulting in the final 653-item Noun list. The list contains 110 loanwords, mostly from Portuguese. The distribution of noun types is shown in Table 1.9.

Table 1.9: Distribution of noun types in the Noun list data set

Word class	Number of instances
Free nouns	406
Inherently free nouns	382
Converted from bound nouns	24
Bound nouns	213
Compounds	34
Total	**653**

The participants were asked to produce examples of how the nouns could be used in contexts where classifiers are obligatory, most often with lower numerals (§ 5.3.3) but also with, e.g., adjectives (§ 5.3.4). They were encouraged to produce as many alternative versions as they could come up with, using different classifiers.

The elicitation resulted in 25 hours of recordings. Notes were taken during the sessions and were later checked against the recordings. The nouns, their meanings, and the information on which classifiers each lexeme can combine with were concatenated into a spreadsheet (see Table 1.8 above for a very simplified example of the structure).

The main purpose of this data collection was to investigate noun–classifier combinability. The method of using a word list was selected for its capability of targeting the combinatory potential of classifiers. In addition, it is a relatively efficient way of assembling large amounts of comparable lexical data. The results of the Noun list data set have been used throughout this thesis, especially for the semantic property analyses (Chapter 7). Both the recordings and the spreadsheet can be found among the archived materials, available on request at https://hdl.handle.net/10050/84e2a81b-501a-41a9-a59d-9c62becbc6a1.

1.2.5 Naturalistic texts

In addition to the elicited data (§ 1.2.4), this work has also been informed by a few naturalistic and semi-structured texts, summarized in Table 1.10. The three naturalistic texts comprise a total of 320 utterances. One of the texts is a picture book narrative (Mayer, 1967), and the other two are free narratives. During the

recording of both free narratives, other people were present and occasionally chimed in with questions or comments to the narrative, which can be heard in the recordings. Otherwise, the texts are basically monologic in their format. The texts were annotated using the ELAN software (Sloetjes & Wittenburg, 2008; Max Planck Institute for Psycholinguistics, The Language Archive, 2023). They were first transcribed and translated into Portuguese with the help of native speakers, and later provided with interlinearized glosses and English translation by me.

Table 1.10: Naturalistic texts

Title	Type	Speaker	Utterances
A boy, a dog, and a frog (Mayer, 1967)	Picture book narrative	E	103
How to make tapioca	Free narrative	B	162
The story of the lizards	Free narrative	J	55
Total			**320**

Naturalistic data provide information about how language is used in free discourse (Bowern, 2015: 131 ff.). The naturalistic texts have informed the analysis of classifiers throughout this thesis. They have played a particularly important role in the identification of morphosyntactic loci (§ 5.3.1) and functions in discourse (§ 6.3). The interlinearized texts and recordings are available on request at https://hdl.handle.net/10050/84e2a81b-501a-41a9-a59d-9c62becbc6a1.

1.3 Conventions

All examples from my own data are marked with the source file ID. The data collection is hosted at the Lund University Humanities Lab Archive, available on request at https://hdl.handle.net/10050/84e2a81b-501a-41a9-a59d-9c62becbc6a1.

For Baniwa, I use the orthography developed by Ramirez (§ 3.1.4). Examples from other sources have been adapted to these conventions unless indicated otherwise (e.g., Table 1.1 where the classifier sets found in previous sources are compared). In cases of Portuguese code switching, Portuguese elements are enclosed in square brackets to indicate that Portuguese orthographic rules are followed instead: [*lápis*] 'pencil'. However, if a Portuguese loanword has been adapted to Baniwa pronunciation, square brackets are not used: *garáapha* 'bottle'.

For the glossing of examples, the Leipzig glossing rules (Bickel, Comrie & Haspelmath, 2008) have been followed as far as possible. For abbreviations of grammatical functions not listed in the Leipzig glossing rules, the Wikipedia (2024b) List of glossing abbreviations has been consulted. A list of all abbreviations used in

glosses is found in the frontmatter of this thesis. In accordance with convention, an asterisk (*) is used to mark ungrammaticality. In the diachronic contexts, asterisks are also used to mark reconstructed forms: a single asterisk (*) for the younger Proto-Japurá-Colombia reconstructions, and a double asterisk (**) for the older Proto-Arawak reconstructions.

Translations from Portuguese to English have been made by me. This primarily concerns cases of reproduced example sentences from Portuguese-language literature (in particular that of Ramirez), and applies both to glossing and free translation. Reproduced data from other sources are marked as "adapted from" the original source if they have in any way been altered (in terms of translation, glossing, orthographic representation, etc.). Some regional Portuguese terms are given in italics, sometimes followed by a clarification in parentheses; a list of these regional terms is also found in the the glossary, with descriptions (Appendix A). Scientific names of flora and fauna that are not common knowledge are also found in the glossary (Appendix A).

When necessary, the languages mentioned in this work will be specified by means of language family, country/area, and/or glottocode (Hammarström et al., 2024).

1.4 Outline of the thesis

This thesis is divided into three parts: Part: I Preliminaries, Part II: The classifier system, and Part III: Concluding remarks.

Part I contains Chapters 1–3: Introduction, Nominal classification: theoretical background, and Grammatical preliminaries. These chapters constitute the background for the thesis, and present information prerequisite for the ensuing description and analysis.

Part II contains Chapters 4–9: Definition and delimitation, Formal properties, Functions, Semantic properties, Diachrony, and Contact. These chapters constitute the main part of the thesis, each focusing of a certain aspect of the Baniwa classifier system.

Part III contains Chapter 10: Discussion and conclusion, concluding the thesis by summarizing and discussing its results and contributions.

Chapter 2

Nominal classification: theoretical background

Classifier systems, such as that of Baniwa, fall under the more general category of nominal classification. This chapter presents the theoretical background on nominal classification that is of relevance to the thesis, as well as the terminology that will be employed. § 2.1-2.6 are concerned with nominal classification systems in general. § 2.7 describes nominal classification in Northwestern Amazonia. § 2.8 describes nominal classification in the Arawakan language family.

2.1 Typological overview

Nominal classification systems are defined by Seifart (2010: 719) according to the following four criteria:

i. "Nouns collocate in well-defined grammatical environments with classificatory elements (these may be free forms, clitics, affixes, etc., and these may also occur elsewhere)."

ii. "The number of classificatory elements is larger than 1 but significantly smaller than the number of nouns."

iii. "Classificatory elements show different patterns of collocation with nouns, i.e. they impose a classification (some overlap is allowed; prototypically, there is a relatively equal division of the nominal lexicon by classificatory elements)."

iv. "At least a substantial subpart of nouns are classified in this way."

This definition encompasses a diverse range of systems that typically go under different names in the literature, such as gender (e.g., in Indo-European languages), noun classes (e.g., in Bantu languages), and numeral classifiers (e.g., in many East and South East Asian languages), to name a few well-known examples. Roughly half of the world's languages (53.4%) have a nominal classification system of some kind (Allassonnière-Tang et al., 2021). The common denominator of nominal classification systems is that they categorize nouns (or referents of nouns, see § 2.3), but this is not their only function[19]: they also serve to individuate and differentiate referents, to expand the lexicon, and to manage reference in discourse, among other things (Contini-Morava & Kilarski, 2013). The categorization is primarily semantic in the sense that all nominal classification systems have at least a partial semantic basis, although some systems may additionally operate on formal (phonological and/or morphological) criteria (Corbett, 1991: 8; Aikhenvald, 2000: 22 ff.).

A common way to divide nominal classification systems into types is by first making a distinction between the systems that display agreement, i.e., gender and noun class systems, and those that do not, i.e., classifier systems (see, e.g., Dixon, 1986; Aikhenvald, 2000; Grinevald, 2000; Senft, 2007; Seifart, 2010; Kilarski, 2013; Allassonnière-Tang et al., 2021). Classifier systems are then typically divided into types based on their morphosyntactic locus, e.g., numeral, noun, and verbal classifiers. Figure 2.1 is a synthesis of the typologies of various scholars.

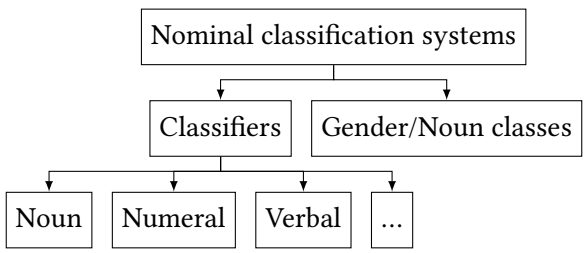

Figure 2.1: Typology of nominal classification (based on Dixon, 1986; Aikhenvald, 2000; Grinevald, 2000; Seifart, 2010; Kilarski, 2013)

Gender/noun class and classifier systems have sometimes been treated as opposing categories. This approach was taken by Dixon (1982), who set up a number of criteria for differentiating between the two (labelled "noun classes" and "classifiers"), summarized in a slightly simplified manner in Table 2.1. In addition, Dixon (1982: 218–219) argued that the type of nominal classification system that

[19]Some would go so far as to say that nominal classification systems do not have a classifying function at all: François (1999) argues that nominal classification systems may have a *classifying effect* ("effet classifiant"), but that this is merely a consequence of some other, syntactic function of the system, and never a function in itself.

a language may have is conditioned by its general morphological type: noun class systems are found in synthetic languages, and classifier systems in analytic languages.

Table 2.1: Summary of Dixon's (1982; 1986) criteria for distinction between noun classes and classifiers (adapted from Aikhenvald, 2000: 6; Grinevald, 2000: 62)

Noun classes	**Classifiers**
all nouns are classified	not all nouns are classified
smallish number of classes	largish number of classes
closed system	open system
fusion with other categories	independent constituent
can be marked on noun	not marked on noun
displays agreement	marked once
rigid assignment	flexible assignment

A subsequent increase in descriptions of nominal classification systems brought about the need for more nuanced typological approaches (Grinevald, 2000: 81; Singer, 2018: 126–127). Corbett & Fedden (2016: 496) cite three main reasons for this. The first is the increased awareness of the diversity of classifier systems, recognized in typological overviews such as Aikhenvald (2000), Grinevald (2000), and Kilarski (2013). The second is the discovery of some languages possessing more than one type of system (Derbyshire & Payne, 1990; Fedden & Corbett, 2017). The mere existence of such languages challenges Dixon's (1982: 218–219) claim of a connection between nominal classification system type and morphological type. Finally, descriptions have emerged of systems that cannot easily be sorted into any particular type. One notable example is Seifart's (2005) description of the nominal classification system in Miraña (Bora-Witotoan, Colombia), which has some properties of the gender/noun class type (e.g., agreement) and others of the classifier type (e.g., a large inventory).

Craig/Grinevald[20] (1992; 2000) and Aikhenvald (2000) are examples of influential typological works that recognize nominal classification as some sort of continuum, while still distinguishing types. Both take a prototype approach (see § 2.4) implying "a gradient rather than categorical treatment" of the properties of nominal classification systems, which are regarded as "focal points on various continua" (Aikhenvald, 2000: 8).

Corbett & Fedden (2016) take an even more unifying approach, treating nominal classification as a continuum without clearly identifiable dividing lines (see also Singer, 2018). In their canonical gender approach, a "canonical ideal" is identified,

[20]These names refer to the same person, Colette Grinevald, who has also published under her former name Colette Craig.

against which actual nominal classification systems are compared according to a number of criteria. This line of research makes no distinctions into types—all kinds of nominal classification systems are called "gender" systems. Corbett & Fedden (2016: 498) clarify that the canonical ideal of their appoach is different from a prototype, as there is "no requirement to produce a canonical exemplar". However, a shared trait between the canonical approach and prototype theory is the gradient nature of category membership.

Whether or not types are distinguished to some degree, most of the relatively recent approaches thus agree that nominal classification systems can be placed on some sort of continuum without definitive borders between the types (Craig, 1992; Aikhenvald, 2000; Grinevald, 2000; Seifart, 2010; Kilarski, 2013; Corbett & Fedden, 2016). This is illustrated in Figure 2.2, where the labels "classifiers" and "gender/noun classes" are used for convenience[21]. The continuum approaches all build at least partially on ideas from prototype (§ 2.4) and grammaticalization (§ 2.2) theory, whether explicitly or implicitly.

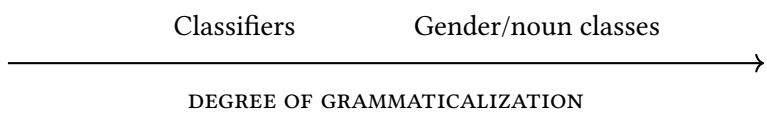

Figure 2.2: Nominal classification continuum (adapted from Kilarski, 2013: 9)

2.1.1 Terminological note: nominal classification

The terms *gender* and *noun class* are used inconsistently in the literature—they have traditionally been used for systems of the Indo-European and the Bantu types, respectively, and continue to be used according to these traditions. Many scholars treat them as the same kind of system, but differ in which label they use as a cover term: Corbett (1991) uses "gender"; Aikhenvald (2000) and Contini-Morava & Kilarski (2013) use "noun class". The term *classifier* appears to be applied more consistently to the less grammaticalized types of nominal classification systems (although this category of systems in itself contains a lot of diversity). However, in some cases it is used in contradicting ways: for instance, Aikhenvald (2000: 1 ff.) sometimes uses "classifier" as a cover term for all devices of nominal classification, and sometimes in reference to a less grammaticalized subgroup in opposition to gender and noun class systems.

[21] Note that the arrow representing the increasing degree of grammaticalization does not entail that all nominal classification systems are necessarily moving from the less grammaticalized end to the more grammaticalized end of the continuum (see further in § 2.2).

The use of the term *nominal classification* is rather uncontroversial; it is used by many authors in roughly the same sense, that is, as a cover term for all kinds of systems that categorize nouns[22] (e.g., Senft, 2007; Seifart, 2010; Kilarski, 2013). In reference to the morphemes that constitute a nominal classification system, no widely accepted cover term has been established. Haspelmath (2021: 45–46) has suggested "nomifier", but this term does not seem to have gained any significant ground yet. Aikhenvald (2000) uses "noun categorization device", a term that has been adopted in a number of publications by other scholars (e.g., Vittrant, 2005). In the same publication, Aikhenvald (2000: 1 ff.) also used "classifier" as a cover term (although inconsistently, as mentioned in the previous paragraph). In Corbett & Fedden (2016), the term "gender" is used for any nominal classification system. Contini-Morava & Kilarski (2013: 266) use "classification marker" in the same sense.

In this work, the term *nominal classification* will be used to refer to the phenomenon as a whole, that is, for anything that qualifies as a nominal classification system according to Seifart's (2010) definition above. When necessary, the abbreviation NC will be employed.

The distinctions of subtypes of nominal classification systems is not a major concern in this thesis, but it will sometimes be convenient to contrast the less grammaticalized and more grammaticalized system types. *Classifier* will be used for the former type, and *gender* or *gender/noun class* for the latter. In particular, the term *classifier* will be used for the grammatical system in Baniwa that is the focus of this thesis, as well as for similar systems. For the bipartite system found in Baniwa and other Arawakan languages (as well as for similar, highly grammaticalized systems with few distinctions), *gender* will be used. Otherwise, *gender/noun class* will be used in reference to the more grammaticalized types of agreement systems.

2.2 Grammaticalization

Grammaticalization is the study of "the evolution from lexical to grammatical forms and from grammatical to even more grammatical forms" (Heine & Kuteva, 2002: 377). It is a theory that offers an explanation to how and why grammatical forms come about based on the underlying assumption that there is a motivation on the language users' part to communicate successfully. To do so, linguistic forms for concrete meanings (*source*) are frequently employed to also

[22] But see Contini-Morava & Kilarski (2013: 265) for a problematization of the theoretical implications of the terminology.

express less concrete meanings (*target*), giving rise to the process of grammaticalization. Heine (2004: 578–579) identifies four key mechanisms of grammaticalization, which are interrelated in the sense that they form part of the same general grammaticalization process:

i. Desemanticization (bleaching): loss in meaning content
ii. Extension (context generalization): use in new contexts
iii. Decategorialization: loss in morphosyntactic properties, including independent word status (cliticization, affixation)
iv. Erosion (phonetic reduction): loss in phonetic substance

According to Heine (2004: 579), the grammaticalization process takes place in an overlap model featuring three stages. In the first stage, there is a linguistic expression (A). In the second stage, A acquires a second pattern of usage (B), and there is now ambiguity between A and B. In the third and final stage, A is lost and there is only B left (although not all grammaticalization processes reach the third stage). The overlap model can be illustrated as follows:

i. A
ii. A, B
iii. B

A central claim of grammaticalization theory is that grammaticalization is essentially unidirectional, i.e., that grammatical forms normally do not become lexical forms (but see Heine (2004: 582–584) for a discussion of some counterexamples).

As pointed out in § 2.1 (see, e.g., Figure 2.2), NC systems can differ in their degree of grammaticalization. The systems traditionally referred to as classifiers are closer to the lexical end of the grammaticalization cline, whereas gender and noun class systems are more grammaticalized and thus closer to the grammatical end (§ 2.1). Some properties that have been argued to be indicative of less vs. more grammaticalized nominal classification systems are summarized in Table 2.2 (see also Dixon, 1986: 106–107; 63 ff. Seifart, 2005: 102–103; Lehmann, 2015; Passer, 2016: 103).

Grammaticalization explains where NC markers come from etymologically. The unidirectionality principle of grammaticalization theory predicts that the origins of grammatical markers are ultimately to be found in the lexicon. For less grammaticalized systems, whose origins are the easiest to localize, this is indeed true:

Table 2.2: Properties indicative of less grammaticalized vs. more grammaticalized nominal classification systems (adapted from Erben Johansson & Cronhamn, 2023: 286)

Less grammaticalized	More grammaticalized
free forms	bound forms
more categories	fewer categories
semantically transparent	semantically opaque/bleached
longer forms	shorter forms
less phonetically eroded	more phonetically eroded
optional	obligatory
flexibly assigned	rigidly assigned
marked once	multiple marking (agreement)
inside NP	inside and outside NP
paradigmatic regularity	paradigmatic irregularity

their most common source, by far, is lexical nouns (Mithun, 1986: 395; Aikhenvald, 2000: 353 ff.; Seifart, 2010: 727). The NC markers of more grammaticalized systems appear to arise from a more diverse array of sources, such as lexical nouns, gender-sensitive pronouns, and certain types of classifiers (Aikhenvald, 2016; Fedden & Corbett, 2017).

If a system of any type is old enough, the etymological origins of NC markers may have become obscured (as seems to be the case in the Bantu languages, see Grinevald & Seifart, 2004: 256). Several different processes may be at play in such cases: phonological erosion and the subsequent merger of homophonous NC markers, reanalysis of other linguistic material as NC markers, recategorization of nouns, and re-semanticization (i.e., regularization) of the system (Seifart, 2010: 729; Aikhenvald, 2016: 83 ff.).

A phenomenon that may help understand the pathway from open-class noun to a marker of nominal classification is that of repeaters, that is, nouns that are either fully or partially repeated in an obligatory slot for a nominal classification marker in a language (Aikhenvald, 2000). This process provides a very clear illustration of the overlap model: a linguistic form is extended into a new context where it becomes re-interpreted, triggering a process of further grammaticalization (Heine, 2004: 578–580). Repeaters are most commonly found in languages with less grammaticalized nominal classification systems, e.g., Truquese (Austronesioan, Micronesia), but are not unheard of in more grammaticalized ones, e.g., Baïnouk (Atlantic-Congo, Senegal/Guinea-Bissau) (Aikhenvald, 2000: 61, 361).

An issue that has been debated in the nominal classification literature concerns whether the grammaticalization cline in Figure 2.1 should simply be seen as a representational tool for the relative degrees of grammaticalization of different systems, or whether the cline also implies that less grammaticalized NC systems

may in fact develop into more grammaticalized ones. In other words, is there any evidence that systems possessing the properties in the left-side column of Table 2.2 over time may come to acquire the properties on the right-hand side?

It has sometimes been hypothesized that classifiers may grammaticalize into noun classes. Grinevald (2002: 264) cites studies on African, Australian, and Amazonian languages in support of an evolutionary path from nouns to classifiers to noun class markers. Grinevald & Seifart (2004) suggested that noun class systems in Niger-Congo languages may have developed along pathways similar to those of many Amazonian NC systems. Franjieh, Corbett & Grandison (2020) made a similar claim based on cognate NC markers in six Oceanic languages that display varying degrees of grammaticalization, ranging from typical possessive classifiers to more gender/noun class-like markers. The best argument for the ability of NC systems to move from the less to the more grammaticalized types is perhaps the synchronic existence of NC markers displaying differing degrees of grammaticalization within one and the same system. This was shown by Seifart (2005) for Miraña. In Miraña, the NC markers can be divided into several subsets that differ in terms of, e.g., frequency, bondedness, semantic generality, and phonological complexity. The subsets can be mapped onto a grammaticalization cline in such a way that their properties correlate just as predicted by grammaticalization theory (Seifart, 2005: 102).

On the other hand, Passer (2016) conducted an in-depth typological study investigating the hypothesis that classifier systems can develop into agreement-like systems, and found that there is no evidence in the form of a documented case of such a shift.

2.3 Semantics

Nominal classification is the grammatical encoding of a semantic categorization imposed on the nominal lexicon of a language. The cross-linguistic semantic variation in nominal classification systems can be studied from a typological perspective. As Evans (2010: 504–505) puts it, linguistic typology "is concerned with exploring the deep regularities which underlie the incredible diversity in how particular languages work". Indeed, although the particular categorization is language-specific, typological research shows that there is a remarkable unity in the kinds of distinctions that are made in nominal classification systems (Seifart, 2010). Aikhenvald (2006: 468) mentions a set of "core parameters", namely a) animacy, b) physical properties, c) functional properties, and d) arrangement (see also Denny, 1976; Allan, 1977). The universality of this set of parameters has

been attributed to their high *cue validity*, i.e., their ability to serve as good predictors of other features, as evidenced by research in experimental psychology (Rosch, Mervis et al., 1976; Seifart, 2010: 725). Physical shape, in particular, "allows a perceiver to predict more facts about an object than any other property" (Palmer, 1999: 363). By contrast, properties with low cue validity (e.g., color, sound, feel, taste, and smell) have not been found to play a role in any nominal classification system described so far (Seifart, 2010: 726).

Animacy-related properties include distinctions such as animate/inanimate and human/non-human. Human referents can be further classified according to sex, social status, kinship relations, etc. For inanimate referents, time-stable physical properties such as shape, dimensionality, direction, size, consistency, and material come into play. Distinctions in linear dimensions (1D, 2D, 3D) are particularly common. Functional properties typically relate to how people interact with the referents in question: whether they can be consumed, domesticated, used for transportation, etc. Finally, arrangement properties relate to configuration and other time-sensitive properties of referents (Aikhenvald 2000: 271–274; 2006: 468; Seifart, 2010: 725–726).

It has sometimes been suggested, most notably by Croft (1994), that certain types of nominal classification systems favor certain semantic parameters—for example, shape properties are identified as especially important to numeral classifier systems, and functional properties such as edibility to possessive classifier systems. However, it should be noted that some of Croft's proposals appear not to hold up in the light of more recently available data (Seifart, 2010: 726).

A fundamental question in nominal classification is whether it is the *nouns* as such (that is, the lexemes), or the *referents* of nouns (that is, the real-world entities) that are categorized (Lucy, 2000: 328–331; Seifart, 2010: 725; see § 2.3.1 below). The classification of referents is more closely connected to the discourse and context, whereas the classification of lexemes is indicative of a certain level of linguistic conventionalization. Referent classification is typically a precursor to lexeme classification, the development occurring through a grammaticalization process (Seifart, 2010: 725). Both levels of classification may, however, exist simultaneously within a system, even applying to the same classifying device.

2.3.1 Terminological note: sense, reference, and extension

This section contains an account of the terminology used to describe semantic relationships in this thesis.

As noted by Saussure (1959), the linguistic sign derives its meaning from two sources: the world that it describes, and the language that it is part of (that is, its relationship with the other linguistic signs within that language). The relationship between a linguistic sign and the real world (that is, between the *signifier* and the *signified*) is one of *reference* or *denotation*, while the relationship between a linguistic sign and its mental representation is one of *sense* (Saeed, 2009: 12, 23 ff.).

The term reference is used for the action on the part of speakers of pointing out specific entities in the world: when speakers use expressions such as *she* or *the dog*, they refer to specific real-world entities, or referents. In contrast, the extension of a linguistic expression is its range of possible referents: for *dog*, it is the entire set of dogs, that is, a particular class of items. The denotation of a linguistic sign is its relationship to its extension. Thus, the denotation is a stable property of the word itself, while reference is context-dependent. Consequently, the extension is the category of entities, while a referent is a particular entity (Lyons, 1977: 174 ff.; Saeed, 2009: 12, 23–27).

With his famous example of *the morning star* and *the evening star*, Frege (1980) argued for a distinction between the sense (German *Sinn*) and the reference (German *Bedeutung*) of linguistic expressions. The morning star and the evening star have the same reference (the planet Venus), but they arguably have different meanings. A perhaps even clearer example is expressions like *the President of Brazil* and *Lula da Silva*: at the time of writing this, they refer to the same individual, but the fact that they do not always do so indicates that there must be more to meaning than reference. This is because these expressions differ in sense, that is, in the mental representation that the respective expressions evoke. According to Frege, sense is primary because it allows for reference: it is because we know the sense of *the President of Brazil* that we can use it to refer to Lula da Silva (or Jair Bolsonaro, or Dilma Rousseff, depending on the time) (Frege, 1980; Saeed, 2009: 12, 30 ff.).

2.4 Prototype theory

Prototype theory is a cognitive model in which categories are construed around a *prototype*, rather than defined by a set of criteria. According to this theory, categories have gradient membership, and some members of a category are more prototypical than others. A commonly cited example features the category of birds. The members of this category typically share a number of attributes: they have wings, can fly, lay eggs, have feathers, etc. This set of criteria succeeds in in-

cluding some category members, such as robins, but fails to include others, such as penguins. Prototype theory offers a model where both robins and penguins are members of the bird category, but robins are more prototypical members than penguins—that is, they are better examples of birds (Rosch, 1973; Lakoff, 1986; Geeraerts, 2006). In other words, according to prototype theory, concepts and categories have an *analog* structure rather than a *digital* one, as often assumed by earlier theories—such assumptions go all the way back to Aristotle, but can also be found in more recent semantic approaches within, e.g., the structuralist and generativist movements (Lakoff, 1986; Geeraerts, 2006: 142).

Prototype theory was developed by the psychologist Eleanor Rosch in the 1970s, through a series of empirical experiments. Rosch's theory of categorization rests on two basic principles, relating to cognitive economy and perceived world structure:

> "The first has to do with the function of category systems and asserts that the task of category systems is to provide *maximum information with the least cognitive effort*; the second has to do with the structure of the information so provided and asserts that the perceived world comes as *structured information rather than as arbitrary or unpredictable attributes*. Thus maximum information with least cognitive effort is achieved if categories map the perceived world structure as closely as possible." (Rosch, 1978:28, emphasis added)

Geeraerts (2006: 146) states the following four characteristics of prototypical categories:

i. "Prototypical categories cannot be defined by means of a single set of criterial (necessary and sufficient) attributes"
ii. "Prototypical categories exhibit a family resemblance structure, or more generally, their semantic structure takes the form of a radial set of clustered and overlapping meanings"
iii. "Prototypical categories exhibit degrees of category membership; not every member is equally representative for a category"
iv. "Prototypical categories are blurred at the edges"

Identifying the most central members of a category is important for the understanding of the category as a whole, and for the understanding of the role and structure of categories in cognition. Berlin and Kay (1969) showed in their seminal study on color terms that a category is defined by its central members rather

than by its full meaning range. Prototypical members of a category possess certain cognitive characteristics that set them apart, known as *prototype effects*, such as faster recognition, earlier acquisition and higher frequency in use. Prototypical members also seem to be used as reference points that people generalize from (Lakoff, 1986: 32), and are more frequently listed by subjects when primed with a category label, probably because of faster retrieval from memory (Rosch, Simpson & Miller, 1976; Rosch, 1978).

Prototype theory can be used to analyze any cognitive category. As nominal classification systems consist of categories, they easily lend themselves to prototype-theoretical analyses. Such studies have been made in the past, perhaps most famously of the noun classes of Dyirbal (Pama-Nyungan, Northeastern Australia; Dixon, 1982; Lakoff, 1987).

Rosch's (1978: 28) principle of cognitive economy may explain the very existence of nominal classification systems: they must aid in providing important information, otherwise it would not be economical to maintain them. In addition, it may be hypothesized that it is the matching of categorization and perceived world structure that is responsible for the striking typological similarities of classifier systems worldwide (see § 2.3), since to a large extent, humans perceive the world in similar ways (see, e.g., Berlin & Kay, 1969). However, Rosch's definition of the second principle also leaves room for cross-cultural variation, and the individual differences in nominal classification systems may indeed be caused by differences in the living environments of speakers of different languages, and the kinds of things they typically come across in their everyday lives.

When it comes to the categorization underlying nominal classification systems, a prototype-theoretical approach would imply a model where a nominal classification marker's applicability to nouns is centered around a prototypical core and extended radially by a family resemblance structure (also known as *chaining*; Lakoff, 1986: 17), rather than determined by binary traits. According to this model, some nouns (or rather, meanings) would be more representative members of a class than others, and there may not be strict lines of inclusion vs. exclusion between classes. In Chapter 7, the classifier system of Baniwa will be studied from a prototype-theoretical perspective.

2.5 Types of classifier markers

26.5% of the world's languages have some kind of non-agreeing nominal classification system, commonly known as a classifier system (Allassonnière-Tang et

al., 2021: 2). As mentioned in § 2.1, classifier systems are often divided into types based on their morphosyntactic locus. Grinevald (2000) recognizes numeral, noun, possessive (genitive) and verbal classifiers. Aikhenvald (2000) adds locative and deictic classifiers to the typology[23].

The typological literature also distinguishes types of classifiers (that is, types of individual markers) according to the different functions they have within a system. A few different kinds of classifier types that will be of relevance for this thesis are brought up in this section: sortal and mensural classifiers (§ 2.5.1), generic classifiers (§ 2.5.2), and specific classifiers (§ 2.5.3).

2.5.1 Sortal and mensural classifiers

A sortal classifier is defined by Lyons (1977: 463) as "one which individuates whatever it refers to in terms of the kind of entity that it is", whereas a mensural classifier is defined as "one which individuates in terms of quantity". The choice of a sortal classifier is conditioned by properties of the referent, while the choice of a mensural classifier is conditioned both by the quantity and the properties of the referent (Aikhenvald, 2000: 115). Below is an example from Vietnamese showing a sortal (1a) and a mensural (1b) classifier.

(1) Vietnamese (Löbel, 2000: 261)

a. *mót* ***con*** *cá*
one **CLF.ANIMAL** fish
'a fish'

b. *mót* ***cân*** *cá*
one **CLF.POUND** fish
'a pound of fish'

Several arguments for the mensural/sortal distinction have been put forth in the literature (Lyons, 1977: 463; Aikhenvald, 2000: 114 ff.; Grinevald, 2004: 1020; Kilarski, 2013: 35; Her, Hammarström & Allassonnière-Tang, 2022). Some of the most common ones are summarized in Table 2.3. However, this distinction has also been criticized for not being a valid distinction language-internally (see, e.g., Lucy, 2000). The issue of mensural and sortal classifiers will be revisited in § 7.6, where it will be argued that such a distinction is not present in Baniwa.

[23] Aikhenvald (2000: 204 ff.) also recognizes so-called "multiple classifier systems", i.e., systems that use the same set of classifiers in different environments. This system type will be treated in § 2.7 and § 2.8 below.

Table 2.3: Differences between sortal and mensural classifiers (based on Lyons, 1977: 463; Aikhenvald, 2000: 114 ff.; Grinevald, 2004: 1020; Kilarski, 2013: 35)

Sortal classifiers	Mensural classifiers
Individuate in terms of kind	Individuate in terms of quantity
Only used with count nouns	Used with both mass and count nouns
Refer to inherent properties of nouns	Refer to temporary properties of nouns
May appear semantically redundant	Contribute meaning
Less freedom of choice	More freedom of choice

2.5.2 Generic classifiers

Languages with nominal classification systems differ in how they handle referents that do not fit neatly into the existing categories of the system. For instance, in Vietnamese and Burmese, abstract nouns may appear unclassified, and in Tai languages, abstract nouns tend to take repeaters (Aikhenvald, 2000: 336). A third strategy for solving this problem is found in languages with general or generic classifiers. The concept of generic classifiers may appear straightforward, but Aikhenvald (2000: 335–337) shows that generic classifiers actually represent at least three distinct functions: the residue function, the default function, and the unspecified referent function (see also Zubin & Shimojo, 1993).

Residue classifiers function as categories for referents outside the domains covered by the nominal classification system, that is, they are "everything else" categories. Default classifiers can replace other classifiers under certain conditions, that is, they are over-arching categories. Unspecified referent classifiers, finally, are used for unknown referents (Aikhenvald, 2000: 335–337).

In some languages, the same classifier may be employed for all three generic functions (e.g., *-da* in Baniwa, see § 7.3.3.1). In other languages, the burden may be divided between different classifiers—for instance, Navajo uses one classificatory verb in the residue function and another in the unspecified referent function (Aikhenvald, 2000: 336).

Generic classifiers often have a more specific core meaning. Greenberg (1972: 34–35) notes that generic classifiers often develop from classifiers with shape semantics, most commonly from classifiers for round objects. Aikhenvald (2000: 336–337) also recognizes original roundness connotations for generic classifiers in some languages, such as Ponapean (Austronesian, Micronesia) and Burmese (Sino-Tibetan)[24]. According to Denny (1979: 99–100), it is typically classifiers for

[24]However, Aikhenvald (2000: 336) also gives some examples to the contrary: in Palikur (Arawakan, Brazil/French Guyana) the numeral classifier for vertical objects is used as a generic

three-dimensional objects that develop into generic classifiers (e.g., in Koyukon (Athabaskan, Alaska)).

2.5.3 Specific classifiers

The previous section was concerned with generic classifiers, that is, classifiers with a broad classificatory scope. This section deals with so-called specific classifiers, that is, classifiers that are used with a single noun, or a very limited set. Such classifiers typically relate to either social structures or to objects with specific functions. They appear in less grammaticalized nominal classification systems, and are usually relatively infrequent in use compared to other classifier types. Specific classifiers that are only used with a single noun are called unique classifiers (Grinevald, 2015: 815; Aikhenvald, 2021: 234).

§ 2.3 brought up the fact that nominal classification systems cross-linguistically show notable similarities when it comes to the semantic distinctions they make, and that these distinctions are in line with cognitively useful categories. This raises the question why specific classifiers exist at all, as they do not appear to classify anything, nor to comply with the basic principles of human cognition.

Aikhenvald (2021) argues that specific classifiers reflect salient aspects of the society and culture that the language in question is used in, and that these distinctions are, consequently, highly sensitive to cultural and societal change. Khmer (Austroasiatic, Cambodia) used to have specific classifiers for clergy, royalty, and the image of Buddha, but they appear to be going out of use since the Khmer Rouge revolution. Speakers of Nivkh (unclassified, Eastern Russia) used to depend on fishing for their livelihood, and had many specific classifiers for objects used in the fishing enterprise, none of which are used or even recognized by Nivkh speakers today.

In line with the general principles of grammaticalization, Aikhenvald suggests that specific classifiers may represent an evolutionary stepping stone between lexical nouns and more general classifiers bridged by the repeater strategy (§ 2.2). She argues that they "constitute the pool of potential forms which can give rise to further, more generalised, categorization devices" (Aikhenvald, 2021: 247).

While language structures may certainly reflect societal and cultural structures, as specific classifiers appear to do according to Aikhenvald (2021), there is so far no explanation as to why a certain specific classifier arises and not another.

classifier, and in Mandarin Chinese, the generic classifier developed from a noun meaning 'bamboo stalk'.

There are arguably central aspects of any society that are not reflected in their nominal classification system. Likewise, as Aikhenvald mentions, specific classifiers do not necessarily reflect salient cultural aspects (for instance, Indonesian has a specific classifier for umbrellas; Aikhenvald, 2021: 243). More research is needed for a better understanding of the function of specific classifiers. In any case, specific classifiers abound in classifier systems—they are also found in Baniwa, and are described in § 7.3.3.2.

2.6 Language contact

Allassonnière-Tang et al. (2021) demonstrated, based on a global sample, that less grammaticalized nominal classification systems are more likely to spread by diffusion than more highly grammaticalized nominal classification systems, which are more often inherited. This is in line with the general tendency of items associated with a lower degree of grammaticalization to be more readily borrowable, e.g., content > function words, and free forms > bound forms (Thomason & Kaufman, 1988; Matras, 2009; Tadmor, 2009).

Nominal classification systems can be affected by language contact in different ways. For example, they can undergo semantic and/or morphosyntactic restructuring, as in the cases of Baniwa's neighbors Tariana (Arawakan; Aikhenvald, 2002) and Kubeo (Tucanoan; Gomez-Imbert, 1996), or they can borrow nominal classification markers, as Resígaro (Arawakan, Peru; Aikhenvald, 2001: 186 ff.; Seifart, 2012) (see § 2.7.2 below).

Another contact-related phenomenon that affects classifier systems is the issue of how new, contact-induced phenomena are incorporated into the classification (whether referred to by loanwords or not). In some languages, there are generic or residue classifiers that take care of this problem (see § 2.5.2). In others, such items may escape classification (e.g., objects made of plastic in Jacaltec (Mayan, Mexico/Guatemala; Craig, 1986: 274)). Nominal classification systems may also adapt to accomodate new members, such as the possessive classifier system of Iaai (Oceanic, New Caledonia), where a classifier traditionally encoding possession of items to sit on now also marks possession of any kind of vehicle for transportation (Dotte, 2017).

Language contact can also lead to the reduction and/or loss of nominal classification systems, typically in couple with a process of language attrition (Aikhenvald, 2000). Dyirbal (Pama-Nyungan, Northeastern Australia) used to have four noun classes with intricate principles of assignment closely connected to mythological

themes (Dixon, 1982: 178 ff.), but this system was modified into a semantically transparent three-way system, probably due to influence from the English pronouns *he*, *she*, and *it* (Schmidt, 1985: 151 ff.; see further in § 7.3.3.3). Warekena (Arawakan, Brazil), one of Baniwa's closest relatives, was described in Aikhenvald (1998: 298–299) as being in a process of losing its classifier system due to language shift to Nheengatu (Tupí-Guaraní) and Portuguese. Today, Warekena is barely spoken anymore (Ramirez, 2020b: 96–97).

Chapter § 9 deals with the issue of classifiers and language contact in Baniwa, including the classification of more recently introduced items and the adaptability of the classifier system. This chapter also addresses a question that has received very little attention in the previous literature, namely what happens phonologically when classifiers are attached to borrowed host words.

2.7 Nominal classification in Northwestern Amazonia

Nominal classification is a common feature in the languages of Northwestern Amazonia, a region known for its dense linguistic diversity and areal diffusion (Grinevald & Seifart, 2004; Seifart & Payne, 2007; Epps & Michael, 2017). Not only is it present in many languages in the region; the systems are also highly similar across language families. Several Amazonian NC systems were described already in the 1960-70s, but largely escaped the attention of nominal classification typologists, perhaps partly due to the use of different terminology (Seifart & Payne, 2007). In the 1980s, detailed descriptions of Amazonian NC systems using more generally accepted terminology began to emerge (Seifart & Payne, 2007)—notable early examples are Gomez-Imbert's (1982) description of Tatuyo (East Tucanoan, Colombia) and Payne's (1986) description of Yagua (Peba-Yaguan, Peru). The earliest overviews of nominal classification in the area were provided by Payne (1987) and Derbyshire & Payne (1990).

The Amazonian nominal classification systems began to attract the attention of typologists, since they typically display characteristics of both less grammaticalized (classifier) systems and more grammaticalized (gender/noun class) systems, thereby challenging the prevailing dichotomy (Aikhenvald, 2000: 6; Grinevald, 2000; Seifart & Payne, 2007; see also § 2.1, § 2.2). For instance, Amazonian systems often have flexible classifier assignment, a typical property of classifier systems, and may at the same time display agreement, a typical property of gender/noun class systems (Krasnoukhova, 2012: 205). Having been largely absent from earlier typological works (e.g., Denny, 1976; Allan, 1977; Dixon, 1986; Corbett, 1991), efforts to include Amazonian systems into the cross-linguistic typol-

ogy of nominal classification started to be made from the 1990s on (Craig, 1992; Aikhenvald, 2000; Grinevald, 2000).

As a consequence of the earlier typological distinctions, Northwestern Amazonian NC systems have often been characterizes as "mixed" (Payne, 1987; Derbyshire & Payne, 1990), as systems of "multiple classifiers" (Aikhenvald, 2000), or as transitional in nature—that is, as noun class system in the early stages of grammaticalization (Grinevald, 2000; Grinevald & Seifart, 2004). As Seifart & Payne (2007: 384) point out, however, "there is nothing particularly 'mixed' about their behaviour, nor do they necessarily constitute 'multiple' systems when viewed from a language-internal perspective".

A more recent areal–typological survey is found in Krasnoukhova (2012: 193 ff.), who uses the term *multifunctional classifier systems*, a term that better accounts for the systems' language-internal cohesion. Descriptions of nominal classification systems in Northwestern Amazonia have continued to surface, although many systems in the region are still un- or underdescribed (Seifart & Payne, 2007). Table 2.4 contains a non-exhaustive list of some of the most important works on the topic.

Table 2.4: Selected publications on nominal classification in Northwestern Amazonian languages

Language(s)	**Publication(s)**
Arawakan languages	Aikhenvald (2019)
Baniwa	Ramirez (2001b; 2020a); Aikhenvald (2007)
Tariana	Aikhenvald (2003)
Bora-Witotoan languages	Seifart (2007)
Bora	Weber (2010)
Miraña	Seifart, 2005
Murui	Wojtylak (2016)
Naduhup languages	Epps & Obert (2022)
Hup	Epps (2007)
Peba-Yaguan languages	
Yagua	Payne (1986; 2007)
Tucanoan languages	
Kubeo	Chacon (2022)
Máíhĩki	Farmer (2015)
Tatuyo	Gomez-Imbert (1982; 2007)
Tuyuca	Barnes (1990)
Wa'ikhana	Balykova (2019)
Areal overviews	Payne (1987); Derbyshire & Payne (1990); Krasnoukhova (2012)
Contact phenomena	
Arawakan/Tucanoan	Gomez-Imbert (1996); Aikhenvald (2002)

2.7.1 General characteristics

A prominent feature of classifiers in Northwestern Amazonian languages is their ability to to be used for both inflection and derivation. Payne (1985) was the first to draw attention to this dual capacity in her description of Yagua (Peba-Yaguan, Peru). The same pattern has since been described for many other languages in in the region (Aikhenvald, 2000; Seifart & Payne, 2007) from different language families: Arawakan (Tariana; Aikhenvald, 2003), Bora-Witotoan (Miraña; Seifart, 2005), and Tucanoan (Tatuyo; Gomez-Imbert, 2007), among others.

Krasnoukhova (2012: 193 ff.) enriched the typological description by identifying three main functions of nominal classification systems in Amazonia: semantic categorization, derivation, and agreement. She notes that the functions are manifested to different degrees in the individual languages, but that "this particular combination of properties is what makes a multifunctional classifier system different from the more prototypical categories of nominal classification" (Krasnoukhova, 2012: 193).

Table 2.5 displays the properties of multifunctional classifier systems that Krasnoukhova (2012: 206) identifies. The properties are categorized under the three main functions. The table also indicates whether the property in question was analyzed as typical of gender/noun class systems (G/NC) and/or classifier systems (CLF) according to Dixon (1982; 1986). As is clear from the table, the multifunctional classifier systems indeed bear characteristics of both of Dixon's types. They have more classifier-like properties relating to their semantic function, and more gender/noun class-like properties relating to the agreement function, while the derivational function is largely absent from Dixon's typology. In what follows, the three functions will be described briefly in terms of Krasnoukhova's properties, as well as some properties noted by other researchers.

In terms of semantics, Amazonian systems typically have large inventories of classes that can be assigned flexibly, meaning that most nouns can take more than one NC marker. The systems are typically open to new members, primarily recruited from lexical nouns (in particular bound nouns), often through the use of repeaters. The properties in this category are all typical of less grammaticalized NC systems (Krasnoukhova, 2012: 207–209). The systems commonly encode shape, but also consistency, size, function, and sex. They often include a fundamental animate/inanimate distinction, and there also tends to be generic classifiers (Seifart & Payne, 2007: 382–383; Epps & Obert, 2022: 6).

Among the derivational properties, we find the ability to derive noun stems. Such stems can be derived either from other noun stems or roots, or from verbal stems

Table 2.5: Properties of multifunctional classifier systems (adapted from Krasnoukhova, 2012: 206)

Properties of a multifunctional classifier system		(G/NC)	(CLF)
Associated with semantic function			
1	Nouns can be assigned to various classes freely	–	+
2	Form largish number of classes	–	+
3	Constitute an open system	–	+
Associated with derivational function			
4	Can derive noun stems		
4a	Either from noun stems or roots	–	–
4b	Or from verbal stems or roots	–	–
5	Can form a full NP when occurring on a modifier	–	–/+
Associated with agreement function			
6	Can occur on predicates to mark core arguments		
6a	Either on *any* predicate	+	–/+
6b	Or only on a subclass of predicates		
7	Can participate in agreement within the NP	+	–
8	Classify all nouns	+	–

or roots. Another typical property is the ability of classifying elements to form full noun phrases together with the modifiers on which they occur, meaning that there need not be an overt noun. This also frequently happens when classifiers are used anaphorically. The properties relating to the derivational function are generally not identified as characteristic of either gender/noun class or classifier systems (Krasnoukhova, 2012: 209–212).

The agreement properties generally align with typical properties of gender/noun class systems. NC markers can typically appear on predicates to mark core arguments, either on any predicate, or on a subclass such as nominal predicates or stative verbs. They can also participate in agreement patterns on constituents within the noun phrase, although Krasnoukhova notes that in many of the languages in her sample, agreement is only realized optionally. Finally, many systems obligatorily classify all nouns, but Krasnoukhova flags this feature as somewhat problematic, due to the existence of neutral classifying elements. In addition, NC markers in Amazonian systems are often realized as bound suffixes (Seifart & Payne, 2007: 382–383).

2.7.1.1 Terminological note: inflection and derivation

In this thesis, the two basic morphological relationships of *inflection* and *derivation* are distinguished (Haspelmath & Sims, 2010: 19–22, 81 ff.). Both derivation and inflection are generally expressed by affixes, that is, morphemes that attach

to words or main parts of words. Affixes often have abstract meaning and grammatical functions, are generally unanalyzable, and cannot occur on their own. Bases or stems are the parts of words that affixes attach to. They may be morphologically complex, and may or may not be able to occur on their own; stems that cannot are called bound stems. Bases or stems that cannot be analyzed into constituents are called roots. Inflection is lexeme-internal, and affects the relationship between different forms of the same lexeme. Derivation is lexeme-external, and creates new lexemes. Derivation is thus a kind of word formation strategy, alongside compounding. The difference between compounding and derivation is that compounding combines more than one root into a new lexeme, whereas derivation only involves one root. The morphological relationships are summarized in Figure 2.3.

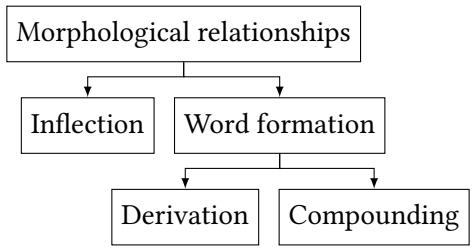

Figure 2.3: Types of morphological relationships (adapted from Haspelmath & Sims, 2010: 18–9)

Haspelmath & Sims (2010: 81) point out that while the conceptual distinction between inflection and derivation is "quite basic to most morphological theorizing and terminology", the distinction between the two is not always straightforward. There are two very different approaches to this: on the one hand, there is the dichotomy approach (e.g., Scalise, 1988; Bickel & Nichols, 2007), which argues that there is sufficient evidence for a formal distinction between inflection and derivation, and on the other hand, there is the continuum approach (e.g., Bybee, 1985), which argues that these are related phenomena which are best characterized on a continuum with "canonical inflection at one extreme, and canonical derivation at the other, but many intermediary types" (Haspelmath & Sims, 2010: 89–91).

To determine whether or not a certain form is inflectional or derivational, the following criteria (Haspelmath & Sims, 2010: 90 ff.) may be used:

i. Syntactic relevance: in terms of syntactic agreement, syntactic government and other syntactic rules, inflection is relevant to the syntax, whereas derivation is not.

ii. Obligatoriness: inflectional features are obligatorily expressed on all applicable word forms, whereas derivational meanings can be omitted.

iii. Limitations on application: inflectional values can be applied to their base without arbitrary limitations (i.e., inflectional paradigms are usually complete), whereas derivational formations may be limited in arbitrary ways.

iv. Same concept as base: canonical inflected word forms express the same concept as the base, whereas canonical derived lexemes express a new concept.

v. Abstractness: inflectional values express relatively abstract meanings, whereas derivational meanings are relatively concrete.

vi. Meaning compositionality: canonical inflected word forms have compositional meaning, whereas canonical derived lexemes have non-compositional meaning.

vii. Position relative to base: canonical inflection is expressed at the periphery of words, whereas canonical derivation is expressed closer to the root.

viii. Base allomorphy: inflection induces less base allomorphy, whereas derivation induces more base allomorphy.

ix. Word class change: canonical inflection does not change the word class of the base, whereas derivational affixes may do so.

x. Cumulative expression: inflectional values may be expressed cumulatively, whereas derivational meanings are not.

xi. Iterability: inflectional values cannot be iterated, whereas derivational values can.

The dichotomy and continuum approaches generally agree on the criteria as such, but disagree about which criteria are the most important (Haspelmath & Sims, 2010: 98 ff.). Proponents of the dichotomy approach consider relevance to the syntax, obligatoriness and limitations on applicability the most important criteria. Within the continuum approach, however, one generally wants to avoid making such, as it is referred to, "arbitrary" choices. Haspelmath & Sims (2010: 99) point out that "[i]f all these criteria are taken seriously, then the continuum approach is almost inevitable, because different criteria may point to different conclusions".

A modification of the dichotomy approach suggests a tripartition, by subdividing inflection into two types: inherent and contextual, where the former type includes values that are not required by the syntactic context but contribute independent semantic information, such as tense and nominal number, and the latter encompasses values that are dictated by the syntactic context but semantically largely redundant, e.g., structural cases and grammatical person marking on verbs (Booij, 1996). This distinction takes better account of the phenomena than the dichotomy approach, since inherent inflection usually has some properties in common with derivation that are lacking in contextual inflection, especially when it comes to meaning compositionality, position relative to base, and base allomorphy (Haspelmath & Sims, 2010: 100 ff.).

The issue of inflectional and derivational functions in Baniwa classifiers is analyzed in § 6.1. As we will see, the distinction is not always so clear-cut; see § 6.4.

2.7.2 Areal diffusion

Nominal classification systems can be affected by language contact, especially those of the less grammaticalized, non-agreeing type (Allassonnière-Tang et al., 2021; see also § 2.6). The general similarity across family boundaries of Amazonian classifier systems has often been attributed to areal diffusion of structural features (Seifart & Payne, 2007). In addition, Epps & Michael (2017) note that while lexical borrowing in Amazonia is sparse, when it comes to nominal classification, even bound morphology is sometimes borrowed. In what follows, a few specific examples of contact situations that have affected nominal classification systems in various ways will be described.

A well-described case of structural convergence between unrelated languages comes from the Vaupés river basin, which is part of the larger Upper Rio Negro region of Brazil and Colombia. In this region, speakers of Tariana (Arawakan) and various East Tucanoan languages participate in a strict intermarriage pattern (linguistic exogamy), which results in bilingual children and "obligatory societal multilingualism" (Aikhenvald, 2012: 76). Due to a strong taboo against language mixing, the languages do not borrow words from each other. Instead, there is heavy structural convergence: apart from a few lingering Arawakan traits, Tariana has radically reshaped its grammar in accordance with that of the East Tucanoan languages. The classifier system of Tariana has developed some traits that are found in East Tucanoan languages but are absent from Baniwa (the closest relative of Tariana), such as obligatory classifier marking on demonstratives, and the use of repeaters (see Aikhenvald, 2002: 86 ff. for more examples). Tariana has also adapted semantically to East Tucanoan languages, for example by classifying animals according to their animacy rather than their shape, as is the case in Baniwa (see § 7.2.1) Another language of the region, Hup (Naduhup), is spoken by people who do not practice linguistic exogamy, but have continuous trade relations with East Tucanoan speakers. As a result, Hup also shares some features with East Tucanoan languages (although much less so than Tariana) that are absent from its relatives spoken outside the Vaupés, including an incipient nominal classification system (Epps, 2007; Epps & Obert, 2022).

Another case from the Upper Rio Negro region displays influence in the opposite direction, from an Arawakan to a Tucanoan language, namely from Baniwa to Kubeo (Gomez-Imbert, 1996). The classifier systems of East Tucanoan languages

generally make a strict division between animate and inanimate referents, where shape classifiers are only used for inanimates. In Kubeo, however, animals are classified according to their shape, just as in Baniwa (see § 7.2.1). Gomez-Imbert argues that this is due to extensive contact with Baniwa speakers resulting from linguistic exogamy.

The contact scenarios brought up so far have resulted in structural changes in nominal classification systems, but there are also examples of lexical borrowing. One such case is Resígaro (Arawakan, Peru), whose speakers have long been in close contact with speakers of the neighboring Bora language (Bora-Witotoan). While it is likely, based on comparison with sister languages, that Resígaro had a nominal classification system in place already before coming into contact with Bora, not much remains of it, as Resígaro has essentially borrowed its entire paradigm of classifier suffixes from Bora (Aikhenvald, 2001: 186 ff.; Seifart, 2012).

2.8 Nominal classification in Arawakan languages

Several types of nominal classification systems are found in Arawakan languages. Most languages in this family have a gender system that distinguishes masculine and feminine, where feminine is usually the marked category. Such systems are stable across the family, both in terms of the markers themselves and the morphosyntactic contexts that they occur in, which include person affixes, demonstratives, and kinship terms (Aikhenvald, 2020).

The majority of Arawakan languages also have classifier systems marked in multiple morphosyntactic contexts, labelled *multiple classifier systems* (Aikhenvald, 2000), *multifunctional classifier systems* (Krasnoukhova, 2012), or *multilocus classifier systems* (Dunn & Rose, forthcoming). The four most common contexts are numerals, nouns, nominal modifiers[25], and verbs. Across the family, the four main loci are about equally common hosts for classifiers. Classifiers in Arawakan languages typically make distinctions related to physical properties, but also to animacy and function (Dunn & Rose, forthcoming).

Many Arawakan languages have more than one nominal classification system, a feature that is relatively rare cross-linguistically (Aikhenvald, 2019: 103; see also Fedden & Corbett, 2017). For example, several North Arawakan languages have both a small, sex- and animacy-based gender system, and a larger, predominantly shape-based classifier system (Aikhenvald, 2019: 123).

[25] As such, Dunn & Rose include at least adjectives and demonstratives (but not numerals).

Classifiers are found in eight of the 15 Arawakan languages that are spoken in the Orinoco-Rio Negro area, as well as in two closely related languages outside this area, Resígaro and Yucuna (Aikhenvald, 2007: 495–498). Of these eight languages, only Baniwa, Tariana and Resígaro mark classifiers in multiple morphosyntactic contexts—the others only have numeral classifiers. Aikhenvald argues that the emergence of multiple classifier contexts is an innovation of the Baniwa-Tariana subgroup[26]. It is unclear what made this system arise, but the East Tucanoan languages, which belong to a different language family altogether but are spoken within the Vaupés-Içana area, share structural similarities with Baniwa and Tariana—among them classifiers in multiple contexts. For this reason, Aikhenvald (2007: 497–498) suggests that the Proto-Baniwa-Tariana innovation of classifiers in multiple contexts "could be indicative of old contact-induced changes in an area which goes beyond the Vaupés into the basin of the Içana and its tributaries" (see § 2.7.2 for more on the much heavier structural convergence in the Vaupés).

2.8.1 Classifiers in Proto-Arawak and Proto-Japurá-Colombia

As Aikhenvald (2019: 103–104) notes, grammatical gender is strikingly uniform and stable across the Arawakan family. The markers go back to Proto-Arawak (PA) ***(-)(r)u(-)* 'feminine' and ***(-)(r)i(-)* 'masculine' (Aikhenvald, 1999; Aikhenvald, 2020). Although classifier systems are found in many Arawakan languages, they do not show the same level of family-wide unity, neither in terms of their forms, meanings, or the contexts in which they are used (Aikhenvald, 1999: 84). In linguistically diverse Northwestern Amazonia, for instance, similar systems are found across family boundaries (§ 2.7), and the region in general is known for its structural convergence (§ 1.1.2). There are several documented cases of languages whose nominal classification systems have been influenced by another language (§ 2.7.2). These differences between gender and classifiers in Arawakan languages are in line with the cross-linguistic finding that classifier systems, being less grammaticalized, are more prone to spread by language contact (Allassonnière-Tang et al., 2021; see also Epps, van Gijn & Emlen, forthcoming).

Despite their relative non-uniformity compared to the gender system, classifier systems as such are not recent innovations in Arawakan languages. Some kind of classifier system appears to have been present already in PA, tentatively dated to ca 4,000–3,500 years BP (Noble, 1965; Zucchi, 2002), although it has undergone

[26] In Resígaro, it can likely be attributed to diffusion from Bora (Aikhenvald, 2001: 186; Seifart, 2012: 483 ff.; see also § 2.7.2).

significant development in the individual branches.

Aikhenvald (2019: 117–120) lists 11 classifiers that have cognates in at least two Arawakan languages of Northwestern Amazonia[27]. All 11 are present in Baniwa. However, this does not appear to be a proposal for a subgroup; Aikhenvald does not present any reconstructions or historical analysis, so it is unclear whether the forms are in fact regular reflexes of common ancestors or not. Ramirez (2020d: 96–97) reconstructs 15 classifiers to Proto-Japurá-Colombia (PJC; see § 1.1.1), which is dated to ca 2,000 years BP (Ramirez, 2020b: 205). All 15 classifiers in question are present in Baniwa[28]. Ramirez's work is explicitly historically oriented, positing a subgroup and making use of the comparative method for reconstructions. PJC will therefore be considered here as an intermediate stage between PA and Baniwa. Table 2.6 lists the 15 reconstructed PJC classifiers along with their Baniwa reflexes.

Table 2.6: Reconstructed classifiers in Proto-Japurá-Colombia and their reflexes in Baniwa (adapted from Ramirez, 2020d: 96–97)

PJC reconstruction		**Baniwa classifier**	
**-iita*	'human'	*-iíta*	CLF.HUMAN
**-aʔa-la*	'round'	*-da*	CLF.GENERIC
**-kʊa*	'limited area'	*-koa*	CLF.SURFACE
**-kaSa*	'threadlike (snake, liana)'	*-khaa*	CLF.CURVILINEAR
**-na*	'trunk, mammal'	*-na*	CLF.TRUNK
**-maka*	'fabric, cloth'	*-máka*	CLF.FABRIC
**-aa*	'water'	*-áanhaa*	CLF.LIQUID
**-eema*	'side'	*-eéma*	CLF.SIDE
**-Sikʊ*	'tube'	*-híko*	CLF.LONG
**-pɨ*	'long tube, rope, liana'	*-pi*	CLF.TUBE
**-wa*	'hole'	*-wa*	CLF.HOLE
**-aapi*	'container'	*-áapi*	CLF.HOLLOW
**-aapʊ*	'long and flexible (pole, path)'	*-áapo*	CLF.STICK
**-pʊkʊ(i)*	'circular'	*-póko*	CLF.CIRCLE
**-Siwa*	'manioc bread'	*-híwa*	CLF.BEIJU

It is not known exactly how the 15 morphemes were used in PJC. According to Ramirez (2020b: 193–194), they are synchronically used as classifiers only in some of the daughter languages, including Baniwa—others merely use them as derivational morphemes on nominals. He categorizes the morphemes as gram-

[27] As such, Aikhenvald (2019: 124–127) counts Piapoco [piap1246], Warekena [ware1258?], Baré [bare1276], Warekena of Xié [guar1293?], Baniva of Guainia [guar1293?], Bahuana [bahu1238], Achagua [acha1250], Yucuna [yucu1253], Baniwa-Curripaco [bani1259], and Tariana [tari1259].

[28] This includes nine of the 11 classifiers listed by Aikhenvald (2019). The two Baniwa classifiers that Aikhenvald lists, but that are absent from Ramirez's (2020d) list of reconstructions, are *-aápa* 'CLF.OBLONG' and *-ko* 'CLF.HAMMOCK'.

matical affixes rather than lexical roots (Ramirez, 2020d: 96–97), thus distinguishing them from bound nouns.

Two studies have attempted to reconstruct classifiers for PA: Payne (1991) and Dunn (2022). Payne reconstructs 2 classifiers at this level, while Dunn reconstructs 12, including the two by Payne. Only Dunn's list of reconstructed classifiers in PA will therefore be considered in what follows.

Baniwa is not part of the 11-language sample that Dunn (2022: 27) based her reconstructions on (the Upper Rio Negro branch in Aikhenvald's (1999) classification is represented by Tariana). However, several of the reconstructed roots have probable reflexes in Baniwa. Four Baniwa classifiers correspond closely to the Tariana reflexes listed by Dunn in both form and meaning, and can be counted as reflexes with a fair amount of certainty: *-pi*, *-na*, *-koa*, and *-máka*. These four classifiers are all reconstructed to PJC too. Another classifier *-áanhaa* 'CLF.LIQUID' might be a reflex of PA ***sa* 'liquid'. There does not seem to be any Tariana equivalent, but the Resígaro reflex *-ʔaanú* 'liquid' (Dunn, 2022: 50) is similar in both form and meaning. The classifier *-áanhaa* is also reconstructed to PJC as **-aa*, which brings it closer in form to the PA reconstruction[29]. Finally, *-aápa* 'CLF.OBLONG', is a possible cognate to Tariana *-pa* 'largish, long', but is not treated as such in this work due to lack of certainty (see § 8.3). Table 2.7 shows the PA reconstructions, the Tariana reflexes, and the (possible) Baniwa reflexes.

Table 2.7: Reconstructed classifiers in Proto-Arawak and their (possible) reflexes in Baniwa and Tariana (based on Dunn, 2022: 54)

PA reconstruction		**Baniwa reflex**		**Tariana reflex**	
***kwa*	'flat'	*-koa*	CLF.SURFACE	*-kwa*	'flat surface'
***ma*	'cloth'	*-máka*	CLF.FABRIC	*-máka*	'extended cloth'
***na*	'large, long'	*-na*	CLF.TRUNK	*-na*	'long, vertical'
***pi*	'long, thin'	*-pi*	CLF.TUBE	*-pi*	'long,thin'
***sa*	'liquid'	*-áanhaa*	CLF.LIQUID	–	–
***pa/ap*	'long, curved'	(? *-aápa*	CLF.OBLONG)	*-pa*	'largish, long'
***kig*	'pointed'	–	–	*-khi, ki*	'thin, curved'
***panhi*	'powder'	–	–	*-hwi*	'particles'
***ako*	'container'	–	–	–	–
***chi*	'moon'	–	–	–	–
***sV*	'oval'	–	–	–	–
***ta*	'flat'	–	–	–	–

Dunn (2022: 56 ff.) proposes, based on occurrence patterns in her cross-family sample of 11 languages, that PA first allowed classifiers on numerals, and at a later stage also in compounding and verb incorporation. She argues that the

[29] Several of the other reflexes of PA ***sa* have also dropped the initial consonant.

classifier system showed signs of grammaticalization already in PA, but that the class remained open for new members from the noun lexicon. The fact that classifier systems look relatively different across the Arawakan family is taken to reflect the fluidity and instability that is argued to have been present in the system at the protolanguage stage.

A different scenario for the development of Arawakan classifiers is sketched in Dunn & Rose (forthcoming: 35), where it is proposed that classifiers may instead have grammaticalized in several loci independently, mutually reinforcing the grammaticalization process by analogy between the contexts. The possibility is evoked by the finding that in their sample of 35 languages, numerals no longer stand out as the most common classifier locus (contra Dunn, 2022). Instead, the sample shows a relatively equal distribution across the four "major loci": numerals, nouns, nominal modifiers (e.g., adjectives and demonstratives), and verbs. This finding is linked to the fact that in Mojeño Trinitario (Arawakan, Bolivia), each of the classifier loci can be linked to a corresponding source construction with a bound noun (Rose & Van linden, 2023; see also Rose, 2024b for evidence of similar constructions in three other Arawakan languages). However, Dunn & Rose (forthcoming: 35) conclude that for the time being, there is not enough evidence to determine whether parallel grammaticalization may have happened in PA or independently in its sub-branches.

Thus, a separate class of morphemes with at least some of the functions and characteristics of the present-day classifiers appears to have been present in the lineage of Baniwa ever since PA. Over time, the system has evolved grammatically and been enriched with more classifiers, some of which are not even present in Baniwa's closest relative Tariana[30]. The diachrony of the Baniwa classifier system is the subject of Chapter § 8.

[30] According to Aikhenvald (2019: 117), 31 of the ca 45 classifiers she recognizes for Baniwa are shared with Tariana, although she does not provide a full list.

Chapter 3

Grammatical preliminaries

This chapter presents a phonological (§ 3.1) and morphosyntactic (§ 3.2) profile of Baniwa, with a focus on properties relevant to the ensuing analysis of its classifiers. The account is based on a combination of previous sources (especially Ramirez, 2020a) and my own analysis.

3.1 Phonology

This section presents the basic properties of Baniwa phonology, drawing largely on the analysis of Ramirez (2020a: 49 ff.). § 3.1.1 focuses on segmental phonology and presents the consonant and vowel inventories. § 3.1.2 is concerned with suprasegmental phonology, and introduces syllabic and moraic structure, stress placement, and tone. § 3.1.3 describes phonological and morphophonological processes of /Ch/ clusters, /h/-metathesis, vowel nasalization, and vowel fusion. In § 3.1.4, the practical orthography used in this thesis is outlined

3.1.1 Segmental phonology

3.1.1.1 Consonant inventory

The 17 consonantal phonemes that Ramirez (2020a: 49) recognizes are shown in Table 3.1.

There are six manners of articulation (nasal, stop, fricative, affricate, flap, and approximant), and eight places of articulation (bilabial, dental, alveolar, postalve-

Table 3.1: Consonant inventory in IPA (adapted from Ramirez, 2020a: 49)

	Bilab.	Dent.	Alv.	Postalv.	Retr.	Pal.	Vel.	Glott.	Unspec.
Nasal	m		n			ɲ			N
Stop	p b	t̪	t d				k		
Fric.								h	
Affr.				ts dz					
Flap			ɾ						
Approx.	w				ɻ	j			

olar, retroflex, palatal, velar, and glottal). One phoneme, the nasal /N/, is unspecified for place of articulation, as it is phonetically realized as a nasalization of an adjacent vowel (§ 3.1.3.3). Phonemic contrasts in voicing[31] are present in the stop (/p/ vs. /b/, /t/ vs. /d/) and affricate (/ts/ vs. /dz/) series. There is a phonemic contrast between voiceless dental /t̪/ and alveolar /t/ stops, but when it comes to voiced stops, the contrast is neutralized, [d̪] and [d] being allophones in free variation of the same phoneme /d/ (Ramirez, 2020a: 53).

Some of the phonemes possess allophonic variants in complementary distribution[32]. These are displayed in Table 3.2.

Table 3.2: Allophonic variation in consonant phonemes (adapted from Ramirez, 2020a: 51 ff., 88 ff.)

Phoneme	Allophone	Context	Example
/b/	[b]	V_V	*no**b**ittiméta* [no**b**iṭiméta] ‘I fragment’
	[ᵐb]	#_	***b**ittíme* [ᵐ**b**iṭíme] ‘fragmented’
/d/	[d]	V_V	*no**d**áka* [no**d**áka] ‘my urine’
	[ⁿd]	#_	***d**áapa* [ⁿ**d**á:pa] ‘paca (rodent sp.)’
/ts/	[t͡ʃ]	_i	***ts**íi**ts**i* [**t͡ʃ**í:**t͡ʃ**i] ‘monkey sp.’
	[t͡s]	_V (V≠i)	***ts**óome* [**t͡s**ó:me] ‘close, near’
/dz/	[ⁿd͡ʒ]	#_i	***dz**íiro* [ⁿ**d͡ʒ**í:ɻo] ‘grasshopper’
	[d͡ʒ]	_i	*a**dz**írhi* [a**d͡ʒ**íɾi] ‘hither’
	[ⁿd͡z]	#_V (V≠i)	***dz**éema* [ⁿ**d͡z**é:ma] ‘tobacco’
	[d͡z]	_V (V≠i)	*íi**dz**a* [í:**d͡z**a] ‘rain’
/w/	[ɥ]	_i	*a**w**iɲa* [a**ɥ**iɲa] ‘tree sp.’
	[w]	_V (V≠i)	*a**w**akáda* [a**w**akáda] ‘forest’
/ɻ/	[ɻ]	i,e_	*pánttiriko* [pá̋ⁿṭiɻiko] ‘in the house’
	[ɾ]	V_ (V≠i,e)	*awakádaliko* [awakádaɾiko] ‘in the forest’

Voiced stops and affricates (/b, d, dz/) are prenasalized ([ᵐb, ⁿd, ⁿdz]) in word-

[31] Ramirez does not count /g/ as a phoneme in Baniwa, although it is present in a handful of loanwords (e.g., *garáapha* ‘bottle’).

[32] Allophones in free variation are disregarded here.

initial position. Some phonemes (/ts, dz, w/) undergo palatalization ([ʧ, ʤ, ɥ]) when they are followed by /i/. In the case of /ɪ/, it is instead the preceding vowel that determines the phonetic realization: after front vowels (/i, e/), /ɪ/ is realized as a retroflex [ɻ][33] (Ramirez, 2020a: 51 ff., 88 ff.).

3.1.1.2 Vowel inventory

The four vowel qualities in Baniwa are presented in Table 3.3 (Ramirez, 2020a: 49).

Table 3.3: Vowel inventory (based on Ramirez, 2020a: 49)

	Front	**Central**	**Back**
High	i		
Medial	e		o
Low		a	

All four vowels in Baniwa have short and long versions, as shown in Table 3.4 (Ramirez, 2020a: 80 ff.). As the table also shows, vowels can contrast in length in any syllable in the phonological word (final /a/, penultimate /i/, antepenultimate /e/ and preantepenultimate /o/), including in more than one syllable per word (as in, e.g., [ká:na:] 'watery'). Vowel length is independent of stress (§ 3.1.2.2) and tone (§ 3.1.2.3).

Table 3.4: Vowel length contrasts (adapted from Ramirez, 2020a: 80)

Vowel	**Short version**	**Long version**
/a/	[ká:n**a**] 'corn'	[ká:n**a:**] 'watery'
/i/	[wa**ɻí**me] 'newlywed'	[wa**ɻí:**me] 'moving slowly'
/e/	[h**é**maɻi] 'abiu fruit'	[h**é:**mali] 'fish sp.'
/o/	[n**o**panápia] 'I planted'	[n**o:**panápia] 'my former house'

Baniwa also has 10 diphthongs (/ai, ao, ea, ei, ia, ie, io, oa, oe, oi/) (Ramirez, 2020a: 82). Short vowels are monomoraic, and long vowels and dipthongs are bimoraic (more on this in § 3.1.2.1). Each vowel can be nasalized through one of the two nasalization processes described in § 3.1.3.3.

The vowel /i/ has an allophonic variant [j], which occurs word-initially when /i/ is immediately followed by a vowel other than /i/ (Table 3.5). The realization of /i/ as [j] also results in lengthening of the subsequent vowel, in order to retain

[33] Note that, while [ɻ] is an allophone of /ɪ/, /ɻ/ and /ɪ/ are also separate phonemes under other phonological conditions.

the mora count as /i/ loses its syllabicity (Ramirez, 2020a: 78)—see § 3.1.2.1 for more on the moraic structure of Baniwa.

Table 3.5: Allophonic variation in vowel phonemes (adapted from Ramirez, 2020a: 78)

Phoneme	Allophone	Context	Example
/i/	[j]	_V (V≠i)	***í**akawa* [**j**á:kawa] 'you(PL) go'
	[i]	elsewhere	***iní**po* [**i**ní**i**po] 'path'

3.1.2 Suprasegmental phonology

3.1.2.1 Syllabic and moraic structure

Syllables (σ) in Baniwa are always open. They can be either light or heavy, consisting of one or two moras (μ), respectively (Ramirez, 2020a: 50). The two syllable types are schematized in Table 3.6, where C stands for any consonant other than /h/. The phonemic sequence /Ch/ is the only consonant cluster permitted in Baniwa—see § 3.1.3.1 for its phonetic realizations.

Table 3.6: Syllable types (adapted from Ramirez, 2020a: 50)

	Moraic structure	Segmental structure
σ light	1μ	(C) (h) V
σ heavy	2μ	(C) (h) V_1 V_2

The rhyme of a light syllable contains one mora, which corresponds to a short vowel. The rhyme of a heavy syllable contains two moras, which corresponds to either a long vowel or a diphthong[34]. Phonologically, both long vowels and diphthongs can be said to consist of two short vowels (and this is indeed how they are represented orthographically, see § 3.1.4). As for the syllable onset, it may consist of Ø, C, h or Ch [Ch, Ç] (Ramirez, 2020a: 50).

Baniwa syllables abide by two basic principles: moraic contraction and trimoraic avoidance (Ramirez, 2020a: 85). These principles, which are at the root of the Baniwa vowel fusion rules, are explained in more detail in § 3.1.3.4.

[34]The bimoraic interpretation is supported by comparative work within the Arawakan family, which reveals that deletion of intervocalic consonants (VwV, VdV, V.IV) has given rise to many of the long vowels in Baniwa (Ramirez, 2020a: 81).

3.1.2.2 Stress

Baniwa has mixed stress placement rules. On the one hand, it has fixed stress placement: the primary stress always falls on the penultimate metric syllable (post-lexical stress). On the other hand, suffixes in Baniwa can be metric or extrametric—this is an inherent lexical property of the suffix (lexical stress). Metric suffixes are those that count with respect to stress placement, and extrametric suffixes are those that do not count in this respect (Ramirez, 2020a: 51, 93 ff.). In accordance with Ramirez, stressed syllables are marked with an acute accent [ˊ] throughout this thesis.

The contrast between metric and extrametric suffixes explains how, despite the fixed stress placement, stress is still a distinctive feature, as evidenced by the minimal pair *limáka* and *límaka* in (2). The suffix *-ka* (underlined) in (2b) is extrametric, which means that the stress placement rule holds by falling on the second-to-last metric syllable. Extrametric suffixes are only invisible to stress placement as long as they are at the right periphery of a phonological word. Should they be followed by a metric suffix, they behave just like metric suffixes (Ramirez, 2020a: 93–94).

(2) a. *li-máka*
3NFSG-to.leave
'he leaves' (adapted from Ramirez, 2020a: 93)

b. *lí-ma-ka*
3NFSG-to.fish.with.poison-SUB
'he is fishing with fish poison' (adapted from Ramirez, 2020a: 93)

The phonetic realization of stressed syllables in Baniwa does not have to do with either intensity or duration (recall that vowel length is independent of stress, as mentioned in § 3.1.1.2). Instead, stress is marked by pitch, i.e., the stressed syllable in a phonological word is the syllable with the highest pitch (Ramirez, 2020a: 92).

3.1.2.3 Tone

As mentioned in § 3.1.1.2 and § 3.1.2.2, word-level stress and vowel length are independent features in Baniwa. When stress falls on a light syllable (i.e., a short vowel), it always receives high tone (Ramirez, 2020a: 81). When stress falls on a heavy syllable, that is, a bimoraic one (a long vowel or a diphthong), the high

tone can dock on either mora. This is distinguished phonetically as a high tone (/V́V/, [ˊ]) and a rising contour tone (/VV́/, [ˇ]), as shown in the minimal pair *kéetto* [ké:t̪o] 'capybara' and *keétto* [kě:t̪o] 'ant sp.' (Ramirez, 2020a: 95). Tone is an inherent lexical property.

Another distinction comes into play when we consider metric suffixes and how they affect the tone. As established in § 3.1.2.2, a metric suffix will have the stress fall on the penultimate syllable. When this syllable is heavy, metric suffixes can be further divided into centripetal and centrifugal suffixes, where the former has stress fall on the second mora of the preceding syllable (/VV́/, rising tone), and the latter has stress fall on the first mora of the preceding syllable (/V́V/, high tone). Centripetal suffixes greatly outnumber the centrifugal ones (Ramirez, 2020a: 96–98). An example of each is shown in Table 3.7, where extrametric suffixes are underlined.

Table 3.7: Types of metric suffixes (adapted from Ramirez, 2020a: 97)

	Suffix	Example
Centripetal	*-karo* 'in order to'	*waá**karo**wa* /wa-aa-**ka.ɾo**-wa/ 'for us to go'
Centrifugal	*-nhi* 'stative'	*wáa**nhi**kawa* /wa-aa-**nhi**-ka-wa/ 'we are going'

As centripetal and centrifugal suffixes may affect the tone of their preceding syllables differently—including, as in Table 3.7, the same syllable of the same morpheme—this may, at least in part, offer a historical explanation for the tone distinction (Ramirez, 2020a: 96–98).

3.1.3 Phonological and morphophonological processes

3.1.3.1 /Ch/ clusters

The sequence /Ch/ (where C stands for any consonant other than /h/) is the only consonant cluster permitted in Baniwa. /Ch/ is realized as an aspirated consonant [C^{h}] if C is an obstruent, and as a voiceless consonant [C̥] if C is a sonorant (Ramirez, 2020a: 50). This is illustrated in the following pair of rules:

- /Ch/ → [C^{h}] / C = obstruent
- /Ch/ → [C̥] / C = sonorant

The phonetic realizations of /Ch/ are shown in Table 3.8. As can be seen, voicing contrasts in obstruents are neutralized when aspirated: /th/ and /d/ are both real-

ized as [tʰ], and /ts/ and /dz/ are both realized as [t͡sʰ]∼[t͡ʃʰ] (Ramirez, 2020a: 69)[35]. The devoicing processes are regular except for /jh/ → [ç]∼[ʃ][36].

Table 3.8: Phonetic realizations of the /Ch/ cluster (adapted from Ramirez, 2020a: 67 ff.)

/Ch/	Realization	Example			
Obstruents → aspirated					
/ph/	[pʰ]	*nó**ph**ia*	/nó**ph**ia/	[nó**p**ʰia]	'I blow'
/th/	[tʰ]	*nó**th**i*	/nó**th**i/	[nó**t**ʰi]	'my eye'
/dh/		*notó**dh**eni*	/notó**dh**eni/	[notó**t**ʰeni]	'I punched'
/t̠h/	[t̠]	*pán**tth**ii*	/paNV**t̠h**ii/	[pa̋ː^{n}**t̠**ʰiː]	'pillar'
/kh/	[kʰ]	***kh**éetti*	/**kh**éeṭi/	[**k**ʰéːṭi]	'sprouting'
/t͡sh/	[t͡sʰ]~[t͡ʃʰ]	*lí**tsh**awa*	/lí**t͡sh**awa/	[lí**t͡s**ʰawa]	'he rips'
/d͡zh/		*nói**dzh**eni*	/nói**d͡zh**eni/	[nói**t͡s**ʰeni]	'I cried'
Sonorants → devoiced					
/mh/	[m̥]	***mh**óokoli*	/**mh**óokoɭi/	[**m̥**óokoɭi]	'catfish sp.'
/nh/	[n̥]	***nh**óa*	/**nh**óa/	[**n̥**óa]	'1SG'
/ɲh/	[ɲ̥]	*píi**ñh**a*	/píi**ɲh**a/	[píː**ɲ̥**a]	'you eat'
/ɭh/	[ɭ̥]	***lh**íma*	/**ɭh**íma/	[**ɭ̥**íma]	'he hears'
/ɻh/	[ɻ̥]	***rh**óa*	/**ɻh**óa/	[**ɻ̥**óa]	'she'
/wh/	[w̥]~[ɥ̥]	***wh**áa*	/**wh**áa/	[**w̥**aː]	'we'
/jh/	[ç]~[ʃ]	*li**x**ápi*	/li**jh**ápi/	[liçápi]~[liʃápi]	'his intestines'

3.1.3.2 /h/-metathesis

A phenomenon related to /Ch/ clusters is that of metathesis involving /h/, illustrated here by means of the two verb conjugation series in Table 3.9.

Table 3.9: Phonetic realizations of some verbal conjugations to illustrate the process of /h/-metathesis (adapted from Ramirez, 2020a: 73–74)

	***-kápa* 'to see'**	***-haaméeta* 'to win'**
1SG	[nokápa]	[n̥oamé:ta]
2SG	[pikápa]	[pʰiamé:ta]
1PL	[wakápa]	[w̥a:mé:ta]
2PL	[ikápa]	[j̥a:mé:ta]

The first series in Table 3.9 suggests the following, quite straightforward, morpheme analysis: *no-* '1SG', *pi-* '2SG', *wa-* '1PL', *i-* '2PL', *-kapa* 'to see'. The second series is less transparently analyzable, but taking into account the information

[35]The consonant cluster /bh/ is unattested—/b/ itself is a rare, borrowed phoneme mainly occurring in onomatopoeia (Ramirez, 2020a: 52, 69).

[36]Ramirez (2020a: 70–72) argues convincingly for its inclusion in this group, but due to lack of space, the discussion will not be repeated here.

about /Ch/ clusters summarized in Table 3.8, the voiceless and aspirated consonants can be analyzed into their underlying phonological /Ch/ clusters. By postulating a metathesis at the morpheme boundary, the forms can be explained as regular derivations involving the same person prefixes and the /h/-initial verb *-haaméeta* 'to win', as shown in Table 3.10[37] (Ramirez, 2020a: 73).

Table 3.10: Underlying forms of the second series in Table 3.9 (adapted from Ramirez, 2020a: 73–74)

Underlying form	Post-metathesis	Phonetic realization
/**n**o-**h**aaméeta/	→ /**nh**oaméeta/	→ [n̥oamé:ta]
/**p**i-**h**aaméeta/	→ /**ph**iaméeta/	→ [p^{h}iamé:ta]
/**w**a-**h**aaméeta/	→ /**wh**aaméeta/	→ [w̥a:mé:ta]
/**i**-**h**aaméeta/	→ /**hi**améeta/	→ [j̥a:mé:ta]

3.1.3.3 Vowel nasalization

As mentioned in § 3.1.1.2, vowels have oral and nasal counterparts. According to Ramirez (2020a: 59 ff.), there are two kinds of nasal vowels: one kind arises as a consequence of coarticulation when a vowel follows a nasal consonant (/m, n, ŋ/), and is thus phonetic rather than phonological. The other nasalization process is, in Ramirez's analysis, due to an underlying nasal consonant /N/, which is unspecified for place of articulation, and whose surface representation is visible only in the nasalization of the following vowel. /N/ has two different phonetic outcomes, depending on the environment:

- /VNV/ → [Ṽ:]: ***pántti*** /pa**NV́**n̥ti/ [pa̰**:**^{n}n̥ti] 'house'
- /VNhV/ → [VhṼ]: ***óohõ*** /**óoNho**/ [**ó:hõ**] 'yes'

In 1, /N/ stands between two vowels: /VNV/. In a first step, /N/ assimilates with the following vowel and nasalizes it: /VṼ/. The two (now adjacent) vowels then assimilate into a long vowel /Ṽ:/, to which the first vowel lends its vowel quality (regressive assimilation) and the second its newly acquired nasality (progressive assimilation). In 2, again, /N/ stands between two vowels, but is now immediately followed by /h/: /VNhV/. The first step is the same as above, the nasalization of the following vowel: /VhṼ/, but the vowel assimilation is blocked by the presence of /h/ (Ramirez, 2020a: 60 ff.). The presence of an underlying nasal consonant that surfaces as nasalization of vowels will become relevant for the analysis of classifiers on Portuguese numerals in § 9.2.1.

[37] Recall the allophonic variation of /i/ (3.1.1.2), where the rule states that /#i/ → [j] / _V (V≠/i/). In analogy with this rule, (Ramirez, 2020a: 75) postulates the rule /#ih/ → /#hi/ → [j̥] / _V (V≠/i/), explaining the form [j̥a:mé:ta] on the final row of Table 3.10.

3.1.3.4 Vowel fusion

As was shown in § 3.1.2.1, vowels sometimes fuse in order to maintain the language's preferred moraic structure. Vowel fusion is triggered when vowels meet on opposite sides of a morpheme boundary within a phonological word. Table 3.11 displays the fusion rules for a number of short vowels, long vowels, and diphthongs (Ramirez, 2020a: 85 ff.).

Table 3.11: Vowel fusion rules (adapted from Ramirez, 2020a: 86)

↓ + →	**a**	**e**	**i**	**o**	**aa**	**ee**	**ii**	**oo**
a	a		e		aa	ee	ai/ee	aa
e			e		ea/ia	ee	ee/ei	
i	a/i		i		ia	ie	ee/ei	
o	o	e	i/o		oa	oe	oi	oo
aa			ee		aa	ee	ai/ee	
ee			ee		aa	ee	ai/ee	
ii					ia	ie		
oo					oa	oe		
ai					aijaa	aijee		
ia			ie		ia		ie	
ie						ie		
io							ioi	
oa			oe		oa	oe	oe	
oe						oe		
oi					owia	oe		

The vowel fusion rules follow some general patterns, based on the principles of moraic contraction and trimoraic avoidance (Ramirez, 2020a: 85):

- **Moraic contraction**: 1μ+1μ → **1μ**
- **Trimoraic avoidance**: 1μ+2μ / 2μ+1μ / 2μ+2μ → **2μ**

Thus, when two monomoraic syllables meet, the resulting syllable is monomoraic. When two syllables meet and at least one of them is bimoraic, the resulting syllable is bimoraic. Examples of two vowel fusion rules that are grounded in the principles of moraic contraction and trimoraic avoidance, respectively, are shown in Table 3.12. Note that /h/-initial morphemes behave as their vowel-initial counterparts—see § 3.1.3.2 on /h/-metathesis.

A few exceptions to the two principles are found in cases where a final diphthong meets an initial long vowel. In three cases (*ai* + *aa* → *aijaa*, *ai* + *ee* → *eijee*, *oi* + *aa* → *owia*), an epenthetic glide is inserted between the two syllables. In one

Table 3.12: Examples of moraic contraction and trimoraic avoidance in vowel fusion rules (adapted from Ramirez, 2020a: 86–87)

Principle	Rule	Example
1μ+1μ → **1μ**	*a* + (*h*)*i* → (*h*)*e*	***na-i****maa* (3PL-sleep) → ***né****maa* 'they sleep' ***na-hi****ma* (3PL-hear) → ***nhé****ma* 'they hear'
1μ+2μ → **2μ**	*a* + (*h*)*oo* → (*h*)*aa*	***na-oó****ma* (3PL-to.want) → ***naá****ma* 'they want' ***na-hóo****rhi* (3PL-pair.of.knees) → ***nháa****rhi* 'their knees'

case (*io* + *ii* → *ioi*), the result is a triphthong (Ramirez, 2020a: 85–88). Table 3.13 shows an example of each type of exception.

Table 3.13: Vowel fusion rules that constitute exceptions from the moraic principles (adapted from Ramirez, 2020a: 86–88)

Rule	Example
oi + *aa* → *owia*	*kód****oi-aa****li* (cujubim-river) → *Kod****owía****li* 'Rio-do-Cujubim (toponym)'
io + *ii* → *ioi*	*li-lh****io=ii****hĩ* (3SG-for=DEM1.NFSG) → *Jíɭh****ioi****hĩ* 'for this'

3.1.4 Orthography

The practical orthography used in this thesis follows the conventions used in Ramirez's dictionary (2001a: 4–8). This is close to the orthography taught in Baniwa schools (Ramirez, 2020a: 99), the most widely accepted orthography that exists for the language. Table 3.14 shows the correlation between Baniwa phonemes and the graphemes used in this thesis, with examples of their use in the practical orthography, phonetic transcription, and sometimes also phonological transcription.

The four vowel phonemes /i, e, o, a/ are represented by graphemes corresponding to their IPA symbols <**i, e, o, a**>. Long vowels (§ 3.1.1.2) are orthographically represented by double letters <**ii, ee, oo, aa**>. The stressed syllable (§ 3.1.2.2) is always marked with an accent <´>. On bimoraic stressed syllables, the tone distinction (§ 3.1.2.3) is indicated through the placement of the accent: <**V́V**> for high tone and <**VV́**> for rising tone.

The consonantal phonemes /m, n, p, b, t, d, k, g, h, w/ are also represented by their IPA symbols: <**m, n, p, b, t, d, k, g, h, w**>. Note that <**g**> does not represent a phoneme in Baniwa according to Ramirez's analysis (2020a), but it is present in a number of loanwords. The orthographic sequences <**ts**> and <**dz**> both represent single phonemes in Baniwa, namely /ts/ and /dz/.

Other consonantal phonemes are orthographically represented by something

Table 3.14: Grapheme–phoneme correlation (Ramirez 2001a: 4–8; 2020a: 99)

Orth.	Phon.	Examples
i	/i/	***íi**ta* [**í:**ta] ‘canoe’
e	/e/	*h**é**ema* [h**é:**ma] ‘tapir’
o	/o/	*k**o**mána* [k**o**mána] ‘bean’
a	/a/	*íit**a*** [í:t**a**] ‘canoe’
m	/m/	***m**anákhe* [**m**anákʰe] ‘açaí (palm sp.)’
n	/n/	***n**éeri* [**n**é:ɻi] ‘deer’
p	/p/	***p**eéthe* [**p**ĕ:tʰe] ‘*beiju*’
b	/b/	*no**b**ittiméta* [no**b**iṯiméta] ‘I fragment’
t	/t/	*íi**t**a* [í:**t**a] ‘canoe’
d	/d/	*nhó**d**oa* [n̪ó**d**oa] ‘my mother’
k	/k/	***k**éeri* [**k**é:ɻi] ‘moon’
g	–	**g**aráapha [**g**aɻá:pʰa] ‘bottle’
h	/h/	***h**éema* [**h**e:ma] ‘tapir’, *manák**h**e* [manak**h**e] ‘açaí (palm sp.)’ , *n**h**óa* [n̪óa] ‘1SG’
w	/w/	*a**w**akáda* [a**w**akáda] ‘forest’
ts	/ts/	***ts**íi**ts**i* [**ʧ**í:**ʧ**i] ‘monkey sp.’, ***ts**óome* [**ʦ**ó:me] ‘close, near’
dz	/dz/	*a**dz**írhi* [a**ʤ**íɻi] ‘hither’, *íi**dz**a* [í:**ʤ**a] ‘rain’
ñ	/ɲ/	***ñ**áme* [**ɲ**áme] ‘no’
tt	/ṯ/	*íi**tt**a* [í:**ṯ**a] ‘smoke’
l	/ɹ/	*kaaro**l**iko* [ká:ɻo**ɹ**iko] ‘in the car’, *pántti**r**iko* [pã́:ⁿṯi**ɻ**iko] ‘in the house’
r	/ɻ/	*née**r**i* [né:**ɻ**i] ‘deer’
y	/j/	*lí**y**a* [lí**j**a] ‘his/its skin/bark’
n	/N/	*pá**n**tti* [pã̌:ⁿṯi] /pa**N**V́ṯi/ ‘house’, *áah**ã*** [á:h**ã**] /áa**N**hV/ ‘here’
x	/j/+/h/	*lí**x**a* [lí:**ʃ**a] /lí**jh**a/ ‘his/its excrement’

other than their IPA symbol, but straightforwardly so: <**ñ**> for /ɲ/, <**tt**> for /ṯ/, <**l**> for /ɹ/, <**r**> for /ɻ/, and <**y**> for /j/.

The orthographic sequence <**Ch**> is pronounced either as an aspirated obstruent (*manák**h**e* [manak**h**e] ‘açaí (palm sp.)’) or a devoiced sonorant (***nh**óa* [n̪óa] ‘1SG’)—see § 3.1.3.1.

The underlying nasal consonant /N/ is phonetically realized as nasalization on an adjacent vowel (§ 3.1.3.3). It is orthographically represented either as <**n**> (*pá**n**tti* [pã̌:ⁿṯi] /pa**N**V́ṯi/ ‘house’) or <**Ṽ**> (*áah**ã*** [á:h**ã**] /áa**N**hV/ ‘here’).

The grapheme <**x**> is special in that it does not correspond to a phoneme in Baniwa, but to the phoneme sequence /j/ + /h/, which is pronounced as [ʃ] (*lí**x**a* [lí:**ʃ**a] /lí**jh**a/ ‘his/its excrement’)—see § 3.1.3.1.

3.2 Morphology and syntax

This section introduces the main characteristics of Baniwa morphology and syntax, based primarily on Ramirez's (2020a) description. The account pays particular attention to themes relevant to the ensuing analysis of classifiers. § 3.2.1 gives an overview of the main characteristics and parts of speech. § 3.2.2 describes the morphosyntactic alignment and the basic constituent order. § 3.2.3 introduces the person marking affixes. The properties of verbs and adjectives are described in § 3.2.4. § 3.2.5 deals with free and bound nouns, including related processes and categories such as possession, conversion, compounding, number, and gender. § 3.2.6 presents the structure of the noun phrase. In § 3.2.7, relational nouns and postpositions are described. § 3.2.8 gives an account of verb nominalization and relative clauses. § 3.2.9 describes quantifiers, including numerals. § 3.2.10 presents the demonstrative system. Finally, § 3.2.11 relates interrogative clauses.

3.2.1 Overview

Typologically, Baniwa (like most Arawakan languages) is highly synthetic, predominantly agglutinating, and head-marking (Aikhenvald, 1999: 80). Affixation is achieved primarily through suffixing, although some prefixes exist (e.g., the person prefixes described in § 3.2.3).

Baniwa has four main parts of speech: free nouns, bound nouns, adjectives and verbs (Ramirez, 2020a: 102 ff.). The distinction of these four parts of speech is based on two dimensions of morphosyntactic criteria, as seen in Table 3.15: free and bound nouns take the same set of nominal suffixes, and adjectives and verbs take the same set of verbal suffixes. Free nouns and adjectives are free morphemes, while bound nouns and verbs are bound morphemes.

Table 3.15: Main parts of speech (adapted from Ramirez, 2020a: 103)

	Nominal suffixes	**Verbal suffixes**
Free morpheme	Free nouns	Adjectives
Bound morpheme	Bound nouns	Verbs

Syntactically, verbs and adjectives typically function as predicates and free and bound nouns as arguments, but nouns can also function as predicates under certain circumstances (see § 3.2.5).

3.2.2 Alignment and basic constituent order

Baniwa has active–stative (or split–ergative) alignment (Aikhenvald, 1999: 87 ff.; Ramirez, 2020a: 108 ff.). The subject of a transitive verb (A) and the subject of an active intransitive verb (S_a) are both marked with person prefixes (3a, 3b), while the object of a transitive verb (O) and the subject of an adjective, also called a stative intransitive verb[38] (S_o), both take person suffixes (3a, 3c).

(3) a. *líiñhakani*
li-íiñha-ka-**ni**
3NFSG-to.eat-SUB-**3NFSG**
A-V-**O**

‘**He** eats **it**’ (adapted from Ramirez, 2020a: 110)

b. *lhiékokawa*
li-heéko-ka-wa
3NFSG-to.run-SUB-MID
S_a-V

‘**He** is running’ (adapted from Ramirez, 2020a: 110)

c. *iinónakani*
iinónaa-ka-**ni**
sad-SUB-**3NFSG**
Adj-**S_o**

‘**He** is sad’ (adapted from Ramirez, 2020a: 110)

The alignment is mirrored in the basic constituent order: A/S_a V/Adj O/S_o. The examples in (4a–4c) are equivalent of those in (3a–3c), but with full noun phrases as arguments. In (4a) and (4b), the connective prefix *i-* (see § 3.16) cross-references the preceding noun phrase as the subject. Indirect objects typically precede direct objects, and oblique arguments typically follow all core arguments (Ramirez, 2020a: 109–111).

[38] Here, I will follow Ramirez (Ramirez, 2020a: 104) in using the term *adjective* (§ 3.2.4.2) rather than *stative verb*. This is a practical decision: in this thesis, this part of speech will primarily figure in their attributive function (which is where they take classifiers, see § 3.2.4.2, § 5.3.4), and the term “adjective” was considered to convey this in the clearest way.

(4) a. *pédoro* *íiñhaka* *palána*
pédoro i-íiñha-ka **palána**
Pedro CONN-to.eat-SUB **banana**
A V **O**

‘**Pedro** eats **banana**’ (adapted from Ramirez, 2020a: 110)

b. *pédoro* *hiékokawa*
pédoro i-heéko-ka-wa
Pedro CONN-to.run-SUB-MID
S_a V

‘**Pedro** is running’ (adapted from Ramirez, 2020a: 110)

c. *iinónaka* *pédoro*
iinónaa-ka **pédoro**
sad-SUB **Pedro**
Adj **S_o**

‘**Pedro** is sad’ (adapted from Ramirez, 2020a: 110)

3.2.3 Person marking

In Baniwa, there are several series of pronominal forms that mark the grammatical person: one of prefixes, one of suffixes, and one of free pronouns (Ramirez, 2020a: 104 ff.). They are treated here because they apply to many different parts of speech. Table 3.16 shows the different series.

Table 3.16: Person affixes (adapted from Ramirez, 2020a: 105)

	Prefix	Suffix	Free pronoun
1SG	*no-*	*-nhoa*	*nhóa*
2SG	*pi-*	*-phia*	*phía*
3NFSG	*li-*	*-ni, -lhia*	*lhía*
1FSG	*ro-*	*-no, -rhoa*	*rhóa*
1PL	*wa-*	*-whaa*	*wháa*
2PL	*i-*	*-hia*	*hía*
3PL	*na-*	*-na, -nhaa*	*nháa*
IMPRS	*pa-*	*-phaa*	*pháa*
CONN	*i-*	–	–

The pronominal paradigm distinguishes singular and plural number. In the third person singular, feminine and non-feminine are distinguished, but this distinction collapses in the plural. Apart from 1–3 person singular and plural forms,

the paradigm also contains an impersonal form *pa-/-phaa*, whose closest English equivalents are the impersonal pronoun 'one' and the indefinite pronouns 'someone'. There is also a connective prefix *i-*. (Ramirez, 2020a: 104 ff.).

The person prefixes encode the subject of a verb (transitive or intransitive), the possessor of a bound noun, or the complement of a relational noun (Ramirez, 2020a: 105). The former two are exemplified in Table 3.17.

Table 3.17: Person prefixes on verbs and bound nouns (adapted from Ramirez, 2020a: 105)

	Subject of verb		**Possessor of bound noun**	
1SG	***nó****-dia*	'I return'	***no****-náapa*	'my arms'
2SG	***pí****-dia*	'you(SG) return'	***pi****-náapa*	'your(SG) arms'
3NFSG	***lí****-dia*	'he/it returns'	***li****-náapa*	'his arms'
1FSG	***ró****-dia*	'she returns'	***ro****-náapa*	'her arms'
1PL	***wá****-dia*	'we return'	***wa****-náapa*	'our arms'
2PL	***í****-dia*	'you(PL) return'	***i****-náapa*	'your(PL) arms'
3PL	***ná****-dia*	'they return'	***na****-náapa*	'their arms'
IMPRS	***pá****-dia*	'(some)one returns'	***pa****-náapa*	'(some)one's arms'
CONN	*Pédoro* ***í****-dia*	'Pedro returns'	*Pédoro* ***i****-náapa*	'Pedro's arms'

The last row in Table 3.17 clearly illustrates the function of the connective prefix *i-*: it serves to connect the preceding noun phrase with the verbal, nominal or relational root that it attaches to, in order to mark it as the non-pronominal subject (*Pédoro* ***í****-dia* 'Pedro returns') or the possessor (*Pédoro* ***i****-náapa* 'Pedro's arms'). This prefix is obligatorily used instead of a person marker whenever there is a non-pronominal subject/possessor/complement, and is therefore counted as part of the paradigm (Ramirez, 2020a: 108). The connective prefix is identical to the second person plural prefix. Both impersonal *pa-* and connective *i-* are frequently used as default possessors.

The person suffixes encode the object of a transitive verb or the subject of an adjective (Ramirez, 2020a: 105). Both are exemplified in Table 3.18.

Table 3.18: Person suffixes on verbs and adjectives (adapted from Ramirez, 2020a: 105)

	Object of transitive verb		**Subject of adjective**	
1SG	*pi-kápa-****nhoa***	'you(SG) see me'	*iinónaa-****nhoa***	'I am sad'
2SG	*no-kápa-****phia***	'I see you(SG)'	*iinónaa-****phia***	'you(SG) are sad'
3NFSG	*no-kápa-****ni***	'I see him'	*iinónaa-****ni***	'he is sad'
1FSG	*no-kápa-****no***	'I see her'	*iinónaa-****no***	'she is sad'
1PL	*pi-kápa-****whaa***	'you(SG) see us'	*iinónaa-****whaa***	'we are sad'
2PL	*no-kápa-****hia***	'I see you(PL)'	*iinónaa-****hia***	'you(PL) are sad'
3PL	*no-kápa-****na***	'I see them'	*iinónaa-****na***	'they are sad'
IMPRS	–	–	*iinónaa-****phaa***	'(some)one is sad'

The free pronoun series is used primarily for emphasis, when there is no root for a person affix to attach to (5) (Ramirez, 2020a: 307–310).

(5) a. *kóaka íiñhali?*
koa-ka i-íiñha-li
Q-SUB CONN-to.eat-REL

'Who is eating?'

b. ***nhóa!***
nhóa
1SG

'Me!' (adapted from Ramirez, 2020a: 307)

3.2.4 Verbs and adjectives

Verbs and adjectives take the same morphological marking, namely the large set of verbal suffixes that appear in the 26 suffix positions in Table 3.19[39] (Ramirez, 2020a: 176–177). The verbal suffixes encode categories such as aspect, tense, mood, and subordination. None of the positions are obligatorily filled; they merely indicate the fixed order in which the suffixes appear if there are several (some variation exists, e.g., in positions 14 and 25). It is rare for the individual instance of a verb or adjective to contain more than four or five suffixes (Ramirez, 2020a: 176).

The verbal suffix *-ka* is very frequent and covers a range of different meanings: locating a verbal action in time, marking complement and subordinate clauses, marking nominal predicates, expressing progressive aspect, expressing contrastive focus, etc. (Ramirez, 2020a: 207–208, 304–305). In this thesis, *-ka* will be glossed consistently as 'SUB' (subordinator) regardless of its function in the particular utterance.[40]

(6) shows verbal suffixes (augmentative *-phaa*, subordinator *-ka*, past inchoative *-iina*, and the third person singular S_o/O marker *-ni*) attached to the verb *-áapoa* 'to sing' (6a) and to the adjective *kadaawáka* 'dark' (6b).

[39] When nouns occur as predicates, e.g., in existential clauses, they may also take verbal suffixes. A few of the suffixes listed here can also be used on nouns even when they are not predicates, e.g., the restrictive suffix *-tsa* (Ramirez, 2020a: 165 ff., 178).

[40] Ramirez (2020a: 304–306) speculates that the homonymous agent nominalizer *-ka* (§ 5.3.6.1) may also be an instantiation of the same morpheme, although the nominalizer is metric and the subordinator extrametric (§ 3.1.2.2). Due to this formal difference, I considered them separate morphemes.

Table 3.19: Order of verbal suffixes (adapted from Ramirez, 2020a: 176–177)

Position	Suffix(es)
1	*-(íi)ta* ‘CAUS’
2	*-áaka* ‘REFL’
3	*-nhi* ‘PROG’
4	*-karo* ‘PRP’ / *-kádaa* ‘COND’ / *-kadáana* ‘SIM.PROG’ / *-kádanako* ‘SIM.PUNC’ / *-kawálhi* ‘SIM.DIST’ / ‘*-kápemi* ‘SIM.SEP’ / *-kádzaa-mi* ‘SEQ’ / *-kápoa* ‘CNSQ’ / *-káde* ‘GNO’ / *-kápidzo* ‘COMPR’ / *-khaa* ‘OPT’
5	*-kánhe* ‘INTS’ / *-panali* ‘SMBL’
6	*-phaa* ‘AUG’
7	*-káadzawa* ‘ACCUM’
8	*-pia* ‘PST’
9	*-khe* ‘GER’
10	*-hini* ‘PRF’
11	*-de* ‘DUR’ / *-dali* ‘CONST’ / *-dawatsa* ‘ADM’
12	*-dzo* ‘COMPR’
13	*-me* ‘CNTR’ (in combination with *-tsa* ‘RESTR’, slot #16)
14	*-li* ‘REL’ *-ka* ‘SUB’ / *-khétta* ‘IMPL’ / *-kháapani* ‘IMM’
15	*-ttoa* ‘PERS’ / *-iitsa* ‘PROG.INCH’
16	*-tsa* ‘RESTR’
17	*-dekha(a)* ‘ADVS’ / *-poni* ‘EVD’
18	*-iina* ‘PST.INCH’
19	*-tshaa* ‘CEXP’
20	*-tsakha* ‘ADD’ / *-tseenakha* ‘REP’
21	*-tha* ‘FRUS’
22	*-pida* ‘RPT’
23	*-mitha* ‘IRR’
24	*-wa(tsa)* ‘FUT’
25	*-iika* ‘INGR’ *-wa* ‘MID’ / person suffix (§ 3.16)
26	*-aaka* ‘SEP’

(6) a. *kaláka* *íapoakeena*
kaláka i-áapoa-**ka-iina**
hen/rooster CONN-to.sing-**SUB-PST.INCH**

'The rooster is already singing' (adapted from Ramirez, 2020a: 176)

b. *kadaawakáphaakeenani*
kadaawaka-phaa-**ka-iina-ni**
dark-**AUG-SUB-PST.INCH-3NFSG**

'It is already darker' (adapted from Ramirez, 2020a: 176)

3.2.4.1 Verbs

Verbs in Baniwa are bound forms that express actions, processes, and in some cases states (*-aa* 'to go', *-íiñha* 'to eat', *-óowhaa* 'to sit', *-éema* 'to stand' etc.). They always function as predicates, and take the set of verbal suffixes described above. Verbs, unlike adjectives, are bound forms (Ramirez, 2020a: 103).

Verbs are either transitive or intransitive (Ramirez, 2020a: 175–176). Intransitive verbs include *-aa* 'to go', *-áara* 'to fly', *-heéko* 'to run', and *-híwa* 'to fall'. Transitive verbs include *-áanhee* 'to know', *-dzeekáta* 'to make', *-íiñha* 'to eat', and *-kápa* 'to see'.

All verbs obligatorily take a person prefix (§ 3.2.3) indicating the subject (***nó****-dia* 'I return', ***lí****-dia* 'he returns'). Transitive verbs also take a suffix to indicate pronominal objects (*pi-kápa-**nhoa*** 'you see me', *no-kápa-**phia*** 'I see you') (Ramirez, 2020a: 105, 175).

As will be shown in § 3.2.5, Baniwa lacks a copula verb.

3.2.4.2 Adjectives

Adjectives express properties, states, and processes (*hamíña* 'heavy', *iinónaa* 'sad', *íitta* 'black', *kattíima* 'happy', *mhedzáako* 'weak', *tshíome* 'tearing', etc.). They take the same suffixes as verbs (§ 3.2.4.1), but unlike verbs, they are free forms. The subject of an adjective aligns with the object of a transitive verb by taking the same person suffixes: *iinónaa-ka-**nhoa*** /sad-SUB-**1SG**/ 'I am sad' (Ramirez, 2020a: 265).

The primary function of adjectives is predicative, but they can also be used at-

tributively, i.e., modifying a head noun in a noun phrase. The attributive use of adjectives is described in more detail in § 5.3.4. Attributive adjectives obligatorily take classifier marking. A small class of adjectives show an irregular pattern in that they require a classifier even when used as a predicate: *maka-**dápana**-ka pántti* /big-**CLF.HOUSE**-SUB house/ 'the house is big' (Ramirez, 2020a: 267). The bound adjectives include *maka-* 'big', *opi-* 'old', and *waali-* 'new, young'.

Apart from the class of words that are inherent adjectives, Baniwa also has several productive adjectivizing affixes that derive adjectives from other parts of speech. The positive and privative prefixes *ka-* (7a) and *ma-* (7b) derive adjectives from bound nouns, and the adjectivizing suffix *-me* (8a) derives adjectives from onomatopoetic expressions (8) (Ramirez, 2020a: 268–277).

(7) *-kóada* 'value' (adapted from Ramirez, 2020a: 269)

a. *kakóada*
ka-kóada
POS-value
'expensive'

b. *makóada*
ma-kóada
PRIV-value
'cheap'

(8) *tshio* 'sound of tearing' (adapted from Ramirez, 2001a: 295)

a. *tshíome*
tshío-**me**
sound.of.tearing-**ADJZ**
'tearing'

3.2.5 Free and bound nouns

Like other Arawakan languages, Baniwa makes a formal distinction between free and bound nouns, which constitute separate parts of speech[41] (Aikhenvald, 1999: 82; Ramirez, 2020a: 102–104). They are primarily distinguished by their

[41] In the typological literature, the terms *alienable* and *inalienable* are commonly used (Nichols, 1988). In the literature on Arawakan languages, free nouns are sometimes also referred to as *absolute*, *autonomous*, or *independent* nouns, and bound nouns as *relative*, *possessed*, or *dependent* nouns (Ramirez, 2020a: 104, my translation). Ramirez uses the terms *independent noun* ("nome independente") and *dependent noun* ("nome dependente").

status as free vs. bound morphemes and by their behavior in possession. However, either type can be productively converted into the other, as will be shown in § 3.2.5.1.

Free nouns are free nominal roots that typically denote human referents (*atsíanli* 'man'), animals (*dzáawi* 'jaguar'), plants (*káini* 'manioc'), natural phenomena (*íidza* 'rain'), and objects (*malíye* 'knife') (Ramirez, 2020a: 103). Proper names belong to this noun type, and most borrowed nouns are found in this category (see § 9.1).

Bound nouns are bound nominal roots that require an obligatory person marking prefix, i.e., they are obligatorily possessed. They include kinship terms (*-hádoa* 'mother'), parts of, e.g., bodies (*-náapa* 'arm'), objects (*-imána* 'tip, point'), plants (*-ke* 'tree branch') and spaces (*-tawána* 'corner'), and other things that are typically in some kind of close relationship with their possessor (*-iipítana* 'name') (Ramirez, 2020a: 103).

The two classes of nouns are defined morphologically by their ability to take the same set of nominal suffixes, which include ablative, allative, plural, collective, augmentative, diminutive, and pejorative suffixes, as well as the relational nouns which are described in more detail in § 3.2.7 (Ramirez, 2020a: 103, 153 ff.). The examples below show the allative case suffix *-lhe* being used on a free (9a) and a bound (9b) noun.

(9) a. *líakawa* *híipanakolhe*
li-aa-ka-wa híipanako-**lhe**
3NFSG-to.go-SUB-MID São.Gabriel-**ALL**

'He goes to São Gabriel.' (adapted from Ramirez, 2020a: 153)

b. *noómakeena* *nódiakawa* *nodzákalerhe*
no-oóma-ka-iina no-dia-ka-wa no-dzakále-**lhe**
1SG-to.want-SUB-already 1SG-to.return-SUB-MID 1SG-village-**ALL**

'I already want to return to my village.'
(adapted from Ramirez, 2020a: 153)

Free and bound nouns display the same syntactic behavior. They typically constitute the head of noun phrases functioning as clausal arguments, but can also function as predicates in existential clauses, as in (10) where the proper name *Pédoro* is the predicate. This is due to the fact that Baniwa lacks a copula (Ramirez, 2020a: 120 ff.).

(10) *noipítana Pédoro*
no-iipítana **Pédoro**
1SG-name **Pedro**

'My name (is) Pedro.' (adapted from Ramirez, 2020a: 120–121)

3.2.5.1 Possession and conversion

Bound nouns are obligatorily possessed. In Table 3.17 in § 3.2.3, the full set of obligatory possessive person markers were shown on the bound noun *-náapa* 'arm(s)': ***no****-náapa* 'my arms', ***pi****-náapa* 'your(SG) arms', ***li****-náapa* 'his arms', etc.

Free nouns cannot be directly possessed by person marking prefixes (**no-tsíino*, intended: 'my dog', Ramirez, 2020a: 124), unless they are accompanied by one of a set of inalienator suffixes which converts them into bound nouns. This operation is regular (with a few exceptions), and the noun keeps its original meaning. Free nouns divide into three alienable possession classes based on which inalienator suffix they take: *-ni*, *-te*, or *-le*[42]. Table 3.20 illustrates the conversion for the three groups.

Table 3.20: Conversion from free to bound nouns (adapted from Ramirez, 2020a: 128)

Free N		**Possessed inalienated N**			
tsíino	→	*notsínoni*	*(no-tsíino-**ni**)*	/1SG-dog-**INAL**/	'my dog'
oñái	→	*nooñáite*	*(no-oñái-**te**)*	/1SG-harbour-**INAL**/	'my harbour'
keníke	→	*nokeníkere*	*(no-keníke-**le**)*	/1SG-manioc.garden-**INAL**/	'my manioc garden'

The *-ni* class is the most common one, comprising 66% of the free nouns. This class includes most nouns denoting animals (especially mammals), some humans, food, tools, physiological states, as well as the majority of borrowed nouns. The *-te* class comprises 22% of the free nouns. These include most nouns denoting humans, but also some animals (especially fish, reptiles, and invertebrates), most humans, fruit, plants, and natural phenomena. Finally, the *-le* class comprises 11% of the free nouns, primarily denoting instruments and artefacts, but also a few plants and animals (Ramirez, 2020a: 129–130).

The alienable possession classes display some irregularities. A few free nouns can be possessed without an overt suffix (e.g., *dzéema* 'tobacco' → *no-dzéma-**Ø*** 'my tobacco'). There is also a fourth inalienator suffix *-de* that only applies to a handful of nouns (*téwa* '*piaba* (fish sp.)' → *no-tewá-**de*** 'my *piaba*'). There

[42]These could also be interpreted as a system of possessive classifiers (Ramirez, 2020a: 132 ff.).

is a great deal of variation in the choice of inalienator suffix: there are both dialectal and generational differences, and in some cases the same individual can use either *-ni*, *-te*, or *-le* interchangeably for the same noun without any apparent change in meaning. In combination with other suffixes, the three-way distinction is neutralized and *-le* is used across the board (Ramirez, 2020a: 131–134).

Bound nouns can also be converted to free nouns, allowing them to stand without a possessor. This is done by means of an alienator construction, a combination of the prefix *i-* and the suffix *-tti* (Ramirez, 2020a: 127–128), illustrated in (11). As these suffixes always appear together, they may alternatively be analyzed as a circumfix. Again, the noun keeps its original meaning through the conversion.

(11) a. *nokáale*
no-káale
1SG-heart

'my heart' [200310_ELIC_K_02]

b. *dzamáda* *ikaalétti*
dzáma-da **i**-káale-**tti**
two-CLF.GNR **AL**-heart-**AL**

'two hearts' [200310_ELIC_K_02]

An exception to the alienation strategy can be seen in kinship terms, which generally cannot be converted in this way[43]. To achieve a non-possessive reading of a bound kinship noun, it can be prefixed with either the impersonal *pa-* or the third person plural *na-*. A few bound nouns have an irregular conversion pattern, e.g., *-pana* '(someone's) house' → *pán-tti* 'house' (Ramirez, 2020a: 128).

Baniwa also has an alienable possession strategy formed with the bound noun *-dzaa* 'belongings'. This construction involves classifiers and will be described in more detail in § 5.3.5.1.

3.2.5.2 Compounding

Compound nouns are common in Baniwa. The first part generally consists of a free noun and the second of a bound noun. The bound noun can either attach directly to the first component, or stand freely, in which case it obligatorily takes

[43] A notable exception being *-eenípe* 'son, offspring', which can be regularly converted into *ienipétti* 'child (non-adult human)', but with the additional, atypical meaning change.

the connective prefix *i-*. Thus, there are the two alternative formats N_1 CONN-N_2 (12a) and N_1N_2 (12b) (N_1 = free noun, N_2 = bound noun). There does not appear to be any difference in meaning between the two constructions.

(12) a. *háiko íiwi*
háiko i-íiwi
tree CONN-flower

'tree flower' [200321_FREE_A_01]

b. *haikóiwi*
háiko-íiwi
tree-flower

'tree flower' [200319_FREE_H_01]

A similar construction involves nominalized forms of verbs (§ 3.2.8). A nominalized verb (which functions as a bound noun) can form the first part of a compound-like construction, where it is juxtaposed with a second noun that represents the object of the (now nominalized) verb. The format of this construction is PREFIX-N_1 (PREFIX)-N_2, where N_1 is a nominalized verb, and N_2 is either a free or bound noun. An example with a free noun as the second part is (13), which uses the connective *i-* as a default possessor, allowing it to function basically as a free noun.

(13) *iphiakáda* *káwaale*
i-phia-ka-da káwaale
CONN-to.blow-NMLZ.AG-CLF.GNR wind

'ventilator' (something that blows wind) [200322_ELIC_F_01]

Examples with a nominalized verb as the first part and a bound noun as the second part are (14a) and (14b). In (14b), impersonal *pa-* functions as a default possessor. In both examples, the two obligatorily possessed compound parts harmonize in person. In (14b), the compound is the head of a noun phrase that is modified by a numeral, showing that these constructions indeed function as nominal elements syntactically.

(14) a. *roparaaxoópa* *rotsíkolee*
ro-páraa-xoópa ro-tsíkolee
3FSG-to.tie-NMLZ.INSTR 3FSG-hair

'her hair tie' (something that is used to tie (up) hair)
[200302_ELIC_F_01]

b. *apaíta* *patakaaxoópa* *patsínoma*
apa-iíta pa-tákaa-xoópa pa-tsínoma
one-CLF.HUMAN IMPRS-to.cut-NMLZ.INSTR IMPRS-beard

'one razor' (something that is used to cut beard) [200302_ELIC_F_01]

The meaning of the compound depends on the type of nominalizer used. Both the agent nominalizer (13) and the instrument nominalizer (14a, 14b) are attested in this construction.

3.2.5.3 Number and countability

The two plural suffixes *-pe* and *-nai* attach to nouns. The *-pe* suffix is rarely used directly on nominal roots, but can be suffixed to nominalized verbs (15). The *-nai* suffix is mostly used with nouns for animate beings, especially humans (*newíki* 'person' → *newikí-**nai*** 'persons'). Certain kinship terms even take a combination of the suffixes (*-iito* 'daughter' → *iito-**nái-pe*** 'daughters') (Ramirez, 2020a: 154 ff.).

(15) *paiñhakarodápe*
pa-íira-karo-da-**pe**
IMPRS-to.eat-NMLZ.LOC-CLF.GNR-**PL**

'kitchen utensils (generic term)' [200315_ELIC_F_01]

However, in general Baniwa disfavors nominal plural marking[44]. For example, in expressions with numerals (16a) and quantifiers (16b) denoting multiple referents, the noun stem is unmarked for number[45].

[44]It appears to be uncommon for languages with obligatory numeral classifiers to have obligatory plural marking on nouns (Greenberg, 1972; Sanches & Slobin, 1973; Tang & Her, 2019).

[45]Note, however, the fossilized plural marker in the quantifier *manópe* 'many' (§ 3.2.9).

(16) a. *manópe* *háiko*
many tree

‘many trees’ [200326_ELIC_F_01]

b. *dzamána* *háiko*
dzama-na háiko
two-CLF.TRUNK tree

‘two trees’ [221001_ELIC_K_01]

To express plurality, a common strategy is to use a periphrastic construction with the third person plural free pronoun/demonstrative *nhaa* (§ 3.2.3, § 3.2.10): ***nhaa*** *tsíino* ‘the/these dogs’ (in contrast to singular ***lhie*** *tsíino* ‘the/this dog’). In these cases too, the noun stem remains unmarked for number, suggesting that it is underspecified for this category rather than inherently singular.

The underspecification of number on nouns is interesting from a countability perspective. In a language like English, only mass nouns are underspecified for number (e.g., *water*), whereas count nouns have distinct singular and plural forms (e.g., *tree* vs. *trees*). In Baniwa, there is no such formal distinction on the noun itself. Whether countability is distinguished depends on the category of the modifying form: for example, some quantifiers are sensitive to countability (§ 3.2.9), but numerals are not (§ 3.2.9.1). Numerals obligatorily take a classifier regardless of the noun’s countability, and there is no distinction between mensural and sortal classifiers (as will be shown in § 7.6).

The underspecification of number can be seen in the analysis in § 7.2.4.2, where it is shown that nouns denoting items that typically occur in pairs (such as shoes) get singular vs. dual readings by combining with the side and pair classifiers, respectively. This is contrasted with paired body part terms, which are an exception to this pattern, and in fact have an inherently dual reading.

Regarding the interaction of classifiers and plurals, see § 4.3.2.

3.2.5.4 Gender marking

Baniwa, like almost all Arawakan languages, has a binary grammatical gender system (that is, an additional nominal classification system besides the classifier system, see Seifart, 2010: 719). Gender is obligatorily marked in several morphosyntactic contexts, including on person marking affixes (17) and demonstra-

tives in the third person singular. Gender suffixes are also used in derivation, both on nouns and as nominalization suffixes on verbal roots.

(17) a. *líekoa*
li-éekoa
3NFSG-face

'his face' [200310_ELIC_K_01]

b. *róekoa*
ro-éekoa
3FSG-face

'her face' [200310_ELIC_K_01]

The two genders denote feminine and non-feminine referents, and the system is almost always semantically transparent: *-ro* 'feminine' is used for all female referents, and *-li* 'non-feminine' is used for everything else. Humans are typically referred to according to their biological sex, but in cases where this is unknown or unspecified, the non-feminine gender is used. Animals are referred to with the non-feminine gender, unless they are expressly female—this typically only happens with domestic animals like pet dogs or chickens, where the sex may be known and of relevance. Thus, (18a) is used for male humans and animals in general, whereas (18b) is typically only used for female humans, but can be used for female animals if the sex-distinction is relevant.

(18) a. *líiñhaka*
li-íiñha-ka
3NFSG-to.eat-SUB

'he/it is eating' (of male humans, animals in general)
[221005_ELIC_G_01]

b. *róiñhaka*
ro-íiñha-ka
3FSG-to.eat-SUB

'she is eating' (of female humans, expressly female animals)
[221005_ELIC_G_01]

3.2.6 Noun phrase structure

The basic structure of the Baniwa noun phrase can be depicted as follows (adapted from Ramirez, 2020a: 258):

[DEM + NUM + NOUN + ADJ + RC]

The head (underlined) of the noun phrase must be either a free noun (19a, 19b) or a possessed bound noun (19c). The noun phrase may also include modifiers. Demonstratives and numerals precede the head: (19a) features a demonstrative, (19b) features a numeral, and (19c) features both. Adjectives (19b) and relative clauses (19d) follow the noun.

(19) a. *lhíehẽ* *ienipétti*
lhíe(hẽ) ienipétti
DEM1.NFSG child

'this child' [190405_BOY_E_01]

b. *apadápana* *pántti* *makadápana*
apa-dápana pántti maka-dápana
one-CLF.HOUSE house big-CLF.HOUSE

'one big house' [200305_ELIC_F_01]

c. *lhiénahã* *aphépa* *nakitsíndatsa*
lhiéna(hã) apa-hípa na-kitsínda-tsa
DEM3.NFSG one-CLF.MALE 3PL-companion-RESTR

'that one (male) relative of theirs' [200308_NARR_J_01]

d. *máapa* *pitakhaanhíko*
máapa pi-tákhaa-ni-híko
bee/honey/sugar.cane 2SG-to.cut-NMLZ.PAT-CLF.LONG

'(the) sugar cane that you cut' [200930_ELIC_K_01]

The noun phrase as a domain is highly relevant for classifier use in Baniwa. Numerals and attributive adjectives always carry classifier marking (see § 5.3.3 and § 5.3.4, respectively), and relative clauses often do, too (§ 5.3.6). In addition, elements that typically function as modifiers can also head noun phrases under some conditions. These processes are described and discussed in detail in § 6.1 and § 6.6.

3.2.7 Relational nouns and postpositions

Relational nouns are a subclass of bound nouns. Their separate status allows them to introduce non-core arguments in clauses (e.g., indirect objects and locatives), which sets them apart from other bound nouns (Ramirez, 2020a: 136 ff.). One of their most common functions is to create locative expressions equivalent to adpositional phrases.

Relational nouns are bound forms that take person prefixes. In (20a), *li-péedza* is equivalent to an adpositional phrase 'in front of him', where the complement is expressed solely by the person prefix. In (20b), there is a noun phrase representing the complement: *dzakálee i-íikaa* 'above the village', and the connective prefix is used to link the relational noun to the complement (Ramirez, 2020a: 141).

(20) a. *nóema* *lipéedza*
no-éema li-**péedza**
1SG-to.stand 3NFSG-**in.front.of**

'I am standing **in front of** him.' (adapted from Ramirez, 2020a: 141)

b. *wáadzoli* *íarakawa* *dzakálee* *íikaa*
wáadzoli i-áara-ka-wa dzakálee i-**íikaa**
vulture CONN-to.fly-SUB-MID village CONN-**above**

'A vulture is flying **above** the village.'
(adapted from Ramirez, 2020a: 141)

Two common relational nouns are *-liko* 'internal part; inside' (21a) and *-nako* 'surface; on (top of)' (21b). In contrast to other relational nouns, they often attach directly to the complement noun, essentially functioning as postpositions (Ramirez, 2020a: 138), which suggests that they are more grammaticalized.

(21) a. *áyaahã* *óoniriko* *nakápa* *híiparo*
áyaahã óoni-**riko** na-kápa híiparo
DEM1.LOC water-**LOC.in** 3PL-to.see frog

'Here **in** the water, they see a frog.' [190405_BOY_E_01]

b. *lhíra* *nóada* *háikonako*
li-híra no-aawáada háiko-**nako**
3FSG-to.climb 1SG-to.think tree-**LOC.on**

'He climbed **up** the tree, I think.' [190405_BOY_E_01]

3.2.8 Nominalization of verbs

There are several different nominalizing constructions in Baniwa, which use different nominalizing suffixes and have different functions. An example of a nominalized verb derived with the instrument nominalizer *-xoópa* is shown in (22). Here, the nominalized verb functions as the head of a noun phrase. Nominalized verbs also commonly function as relative clauses (see § 3.2.8.1).

(22)

líoma	*lhípakani*	*líyo*	*lhiénahã*
li-óoma	li-hípa-ka-ni	li-yo	lhiénahã
3NFSG-to.want	3NFSG-to.catch-SUB-3NFSG	3NFSG-INSTR	DEM3.NFSG

phapaxoópa
pa-hípa-xoópa
IMPRS-to.catch-NMLZ.INSTR

'He wants to catch it with that **hand net** (lit. 'instrument for catching')' [190405_BOY_E_01]

Table 3.21 displays an overview of all the nominalizers found among nominalizations in the Noun list data, ordered by number of occurrences[46]. More nominalizing suffixes are listed in Ramirez (2020a: 279 ff.).

Table 3.21: Nominalizers found in the Noun list data set

Nominalizer	Function	Number of items	Classifier?
-karo	location nominalizer	27	✓
-ka	agent nominalizer	21	✓
-xoópa	instrument nominalizer	18	–
-khe	action nominalizer	17	–
-ni	patient nominalizer	15	✓
-kali	state nominalizer	6	–
-kaáwa	location nominalizer	2	–
-xoóda	instrument nominalizer	1	–

When a verb is nominalized, it takes the shape of a bound noun, meaning that it requires a person prefix indicating its possessor. Many of the nominalized forms

[46]Some of the nominalizers have similar semantics. According to Ramirez (2020a: 282–283), the only difference between the location nominalizers *-karo* and *-kaáwa* is that the former takes a classifier and thus allows for the specification of the shape of the location. The two instrument nominalizers *-xoóda* and *-xoópa* appear to be synonymous (Ramirez, 2020a: 291)—it is not known what conditions the choice of one over the other. In the material at hand, *-xoópa* is far more common.

have become lexicalized either with the connective prefix *i-* or the impersonal *pa-* (see § 3.2.5.2). Both *i-* and *pa-* are part of the person prefix paradigm (§ 3.2.3).

Three of the nominalizers in Table 3.21 take classifiers in all attested cases: the agent nominalizer *-ka*, the patient nominalizer *-ni*, and the location nominalizer *-karo*. Nominalization with classifiers is described in more detail in § 5.3.6.

3.2.8.1 Relative clauses

Baniwa forms relative clauses through the nominalization of verbs (Ramirez, 2020a: 298 ff.; see § 3.2.8), which is a very common strategy in South American languages (van Gijn, Haude & Muysken, 2011: 10–11). The verb is nominalized either with a nominalizing suffix (23a) or the relative marker *-li* (23b). The relative clause functions as a modifier in the noun phrase (§ 3.2.6) and always follows the head noun.

(23) a. *atsíanlika idzaamikaíta*
atsíanli-ka **i-dzáami-ka-iíta**
man-SUB **CONN-to.be.ill-NMLZ.AG-CLF.HUMAN**

‘the man **who is ill**’ (adapted from Ramirez, 2020a: 300)

b. *atsíanlika idzáamiri*
atsíanli-ka **i-dzáami-ri**
man-SUB **CONN-to.be.ill-REL**

‘the man **who is ill**’ (adapted from Ramirez, 2020a: 300)

The choice of nominalizer depends on what function the relativized noun has in the relative clause: subject (23), object (24a), location (24b), instrument (24c), etc. In many situations, more than one construction is possible (for example, the two in (23) above), whereas in others, different constructions come with differences in meaning. Some constructions include a relational noun that follows the nominalized verb (e.g., 24b, 24c). For a more detailed account of Baniwa relative clauses, see Ramirez (2020a: 298–304).

(24) a. *tápee nóirali*
tápee **no-íira-li**
medicine **1SG-to.drink-REL**

‘the medicine **that I drank**’ (adapted from Ramirez, 2020a: 298)

b. *méedzaka limaaníkoa ináko*
méedza-ka **li-ímaa-ni-koa** **i-náko**
table-SUB **3NFSG-to.sleep-NMLZ.PAT-CLF.SURFACE** **CONN-LOC.on**

'the table **on top of which he sleeps**'
(adapted from Ramirez, 2020a: 303)

c. *pía nólhio malíye natákhaali iyo kóphe*
pi-aa no-lhio malíye **na-tákhaa-li** **i-yo** **kóphe**
2SG-to.give 1SG-BEN knife **3PL-to.cut-REL** **CONN-INSTR** **fish**

'Give me the knife **with which they cut the fish**!'
(adapted from Ramirez, 2020a: 301)

3.2.9 Quantifiers

Baniwa has a number of non-numeral quantifiers, like *hóre* 'much' (25a), *manópe* 'many' (25b), *tsóo(-tsa)* 'little' (25c), and *menakóda* 'few' (25d) (see also Ramirez 2020a: 250–251 for a non-exhaustive list).

(25) a. *phiwánhika hóre óoni*
pi-hiwánhi-ka **hóre** óoni
2SG-to.transport-SUB **much** water

'You carry a lot of water' [200308_NARR_B_01]

b. *manópe newíki*
manópe newíki
many person

'many people' [200326_ELIC_F_01]

c. *tsóotsa matsóka*
tsoo-tsa matsóka
little-RESTR farinha

'(a) little *farinha* (roasted manioc flour)' [200326_ELIC_F_01]

d. *menakóda tsíino*
menakóda tsíino
few dog

'few dogs' (adapted from Ramirez, 2020a: 251)

The use of quantifiers reflect a mass/count distinction. *Hóre* 'much' appears to be used primarily with masses (25a), and *manópe* 'many' and *menakóda* 'few' with countable items[47]. An example of this can be seen below with *íidza* 'rain': *hóre íidza* refers to a large quantity of rain, whereas *manópe íidza* is interpreted as referring to a large number of separate rainfalls in different locations. *Tsóo-* can mean both 'little' and 'few'[48], but the reading 'few' requires a plural marker.

(26) a. *hóre* *íidza*
much rain

'a lot of rain' [200326_ELIC_F_01]

b. *manópe* *íidza*
many rain

'many rains (separated, in different places)'
*'a lot of rain' [200326_ELIC_F_01]

3.2.9.1 Numerals

Baniwa's lowest numerals *apa-* 'one' *dzama-* 'two', and *madali-* 'three' are unanalyzable. The numeral *likoa-…-áaka* 'four' is clearly morphologically complex, but barely analyzable[49]. One–three inflect for classifiers obligatorily. Four has traditionally taken an obligatory classifier infix and may still do so for some speakers, but for others (as a testament to its intransparency) it has fossilized into the form *likoadáaka* that only takes a classifier suffix optionally (see further in § 5.3.3).

From five and on, numerals consist of descriptive expressions involving hands, fingers, feet, and toes, producing very long strings of words, as exemplified by the numeral 'six' in (27). These analyzable, higher numerals do not inflect for classifiers (even optionally). Table 3.22 shows the native numerals 1–20.

[47] As a further indication, the latter two possibly contain some fossilized suffixes hinting at the countability of their referents: plural *-pe* in *manope*, and generic classifier *-da* in *menakóda*).

[48] *Tsoo* is also used as an adjective with the meaning 'small'.

[49] Analyses of this numeral form differ: Ramirez (2020a: 247) tentatively interprets it as *li-koa-…-áaka* /3SGNF-opposite-…-REFL/, meaning something like 'it is opposite of itself'. Aikhenvald (1996a: 99; 2002: 107–108) analyzes it as *li-koa-…-ka* /3SGNF-be.enough-…-DECL/ 'it is enough/the one that is enough'.

(27) *apeéma* *pakáapi* *aphéwi* *apeémanakhitte*
apa-eéma pa-káapi apa-híwi apa-eéma-náko-hítte
one-CLF.SIDE IMPRS-pair.of.hands one-CLF.POINTED one-CLF.SIDE

'six' (lit. 'one side of a pair of hands and one (finger) from the other side')
[220921_ELIC_F_01]

Table 3.22: Native numeral system

	Numeral form	Translation
1	*apa*-CLF	–
2	*dzama*-CLF	–
3	*madali*-CLF	–
4	*likoa*-CLF-*áaka* / *likoadaaka*(-CLF)	–
5	*apeéma pakáapi*	'1 hand'
6	*apeéma pakáapi aphéwi apeémanakhitte*	'1 hand + 1 from other side'
7	*apeéma pakáapi dzamhéwi apeémanakhitte*	'1 hand + 2 from other side'
8	*apeéma pakáapi madalhíwi apeémanakhitte*	'1 hand + 3 from other side'
9	*apeéma pakáapi likoadaakhéwi apeémanakhitte*	'1 hand + 4 from other side'
10	*dzameéma pakáapi*	'2 hands'
11	*dzameéma pakáapi aphéwi pháipanakhitte*	'2 hands + 1 from foot'
12	*dzameéma pakáapi dzamhéwi pháipanakhitte*	'2 hands + 2 from foot'
13	*dzameéma pakáapi madalhíwi pháipanakhitte*	'2 hands + 3 from foot'
14	*dzameéma pakáapi likoadaakhéwi pháipanakhitte*	'2 hands + 4 from foot'
15	*dzameéma pakáapi apeéma pháipa*	'2 hands + 1 foot'
16	*dzameéma pakáapi apeéma pháipa aphéwi apeémanakhitte*	'2 hands + 1 foot + 1 from other side'
17	*dzameéma pakáapi apeéma pháipa dzamhéwi apeémanakhitte*	'2 hands + 1 foot + 2 from other side'
18	*dzameéma pakáapi apeéma pháipa madalhíwi apeémanakhitte*	'2 hands + 1 foot + 3 from other side'
19	*dzameéma pakáapi apeéma pháipa likoadaakhéwi apeémanakhitte*	'2 hands + 1 foot + 4 from other side'
20	*apaíta nawíki*	'1 person (2 hands + 2 feet)'

Borrowed numerals from Portuguese are common in everyday use, especially for higher numerals (see Dixon 2012: 71 ff.). Portuguese numerals may optionally inflect for classifiers (see further in § 9.2.1).

The numeral *apa-* 'one' is also used in the sense 'other' (28).

(28) *apáaphi* *wakinikíre*
apa-áaphi wa-kiníki-re
one-CLF.AREA 1PL-manioc.garden-INAN

'our other manioc garden' [200311_FREE_E_01]

Numerals and (non-numeral) quantifiers show some interesting differences. First, numerals obligatorily take classifiers, which quantifiers never do. Second, a mass/count distinction can be seen in quantifier use (§ 3.2.9), but is not reflected

in the use of numerals; in particular, no distinction can be made between mensural and sortal classifiers, as is discussed at length in § 7.6. A property that both quantifier and numeral expressions have in common is that the noun stem remains unmarked for number (§ 3.2.5.3).

3.2.10 Demonstratives

The demonstrative system of Baniwa is presented in Table 3.23. The table includes both nominal demonstratives (equivalent to 'this' and 'that') and adverbial demonstratives (equivalent to 'here' and 'there'), as these belong to the same paradigm. There are four distinctions in distance referred to as zones 1–4, where zone 1 is the closest to the speaker and zone 4 the farthest away (Ramirez, 2020a: 310 ff.).

Table 3.23: Demonstrative system (adapted from Ramirez, 2020a: 310)

		Distance from speaker →			
		Zone 1	**Zone 2**	**Zone 3**	**Zone 4**
Nominal	**NFSG**	*lhíe(hẽ)* *=ii(hĩ)* DEM1.NFSG	*lhíera(hã)* *=éera(hã)* DEM2.NFSG	*lhiéna(hã)* *=eéna(hã)* DEM3.NFSG	*lhiéta(hã)* *=eéta(hã)* DEM4.NFSG
	FSG	*rhóa(hã)* DEM1.FSG	*rhóara(hã)* DEM2.FSG	*rhoána(hã)* DEM3.FSG	*rhoáta(hã)* DEM4.FSG
	PL	*nháa(hã)* DEM1.PL	*nháara(hã)* DEM2.PL	*nhaána(hã)* DEM3.PL	*nhaáta(hã)* DEM4.PL
Adverbial		*-áa(hã)* *=áa(hã)* DEM1.LOC	*-áara(hã)* *=áara(hã)* DEM2.LOC	*-aána(hã)* *=aána(hã)* DEM3.LOC	*-aáta(hã)* *=aáta(hã)* DEM4.LOC

Nominal demonstratives are marked for number (singular/plural) and, in the singular, gender (feminine/non-feminine), resulting in three series of forms. The points of departure for these series are the independent third person pronouns *lhía* '3NFSG', *rhóa* '3FSG', and *nháa* '3PL' (see § 3.2.3). The adverbial demonstrative series is built on the root *-áa*. The zone 1 forms are unmarked[50], and in zones 2–4, the pronominal roots receive the suffixes *-ra*, *-na*, and *-ta*, respectively. Each form has an optional suffix *-hṼ*, where Ṽ = a nasalized version of the last vowel before the suffix. The non-feminine singular (*lhía-*) and spatial (*áa-*) series have parallel enclitic forms that attach to the previous word in the sentence under certain circumstances (Ramirez, 2020a: 310–313).

[50]The zone 1 forms are identical to the independent pronouns, with the exception of the non-feminine form which has an irregular vowel: *lhie(hẽ)* ≠ *lhia*.

Nominal demonstratives are typically used as modifiers in noun phrases, where they precede the head noun. They can also form independent noun phrases (without formal changes) if the reference is clear from the context. Demonstratives can be deictic in both space and discourse. In other words, they can be used with referents at various physical distances, and also anaphorically for previously mentioned entities (Ramirez, 2020a: 311–319).

3.2.11 Interrogative clauses

There are two ways to form questions in Baniwa (Ramirez, 2020a: 331 ff.). Polar questions typically start with the interrogative particle *káphaa?* (29a). This is not obligatory in negative polar questions like (29b), which are distinguished from declarative clauses only by intonation.

(29) a. *káphaa piiñhakéera?*
káphaa pi-íiñha-ka=íira
PQ 2SG-to.eat-SUB=DEM2.NFSG

'Are you eating?' (adapted from Ramirez, 2020a: 331)

b. *ñámetshaa píiñhaka?*
ñáme-tshaa pi-íiñha-ka
NEG-CEXP 2SG-to.eat-SUB

'Didn't you eat?' (adapted from Ramirez, 2020a: 331)

Wh-questions are formed with a variety of interrogative pronouns, most of which are based on the interrogative root *kóa?* 'what/who?'. Some examples are given below with *kóa*[51] (30a, 30b), *kóame?* 'how?' (30c), *koawáda?* 'why?' (30d), and *kóadzo?* 'how many?' (30e). There are also some interrogative pronouns that are not based on *koa-*, e.g., *kadali-?* 'how many?' (30f). Interrogative clauses tend to be wh-initial (Ramirez, 2020a: 331 ff.).

(30) a. *kóakatshaa pidéenhiri?*
kóa-ka-tshaa pi-déenhi-li
Q-SUB-CEXP 2SG-to.do-REL

'**What** are you doing?' (lit. 'What is it that you are doing?')
(adapted from Ramirez, 2020a: 331)

[51] The subordinating suffix *-ka* is more or less obligatory on the form *kóa?*, and if the entity asked about is the subject or the object of the predicate, it is marked with relative *-li* (Ramirez, 2020a: 332).

b. *kóaka íokaliwa?*
kóa-ka i-óoka-li-wa
Q-SUB CONN-to.arrive-REL-MID

‘**Who** (is it that) is coming?’ (adapted from Ramirez, 2020a: 332)

c. *kóame pidzeekátaka íita?*
kóame pi-dzeekata-ka íita
how 2SG-to.make-SUB canoe

‘**How** do you make canoes? (adapted from Ramirez, 2020a: 334)

d. *koawádatshaa pidzeekáta kádzoahã?*
koawáda-tshaa pi-dzeekata kádzo=aahã
why-CEXP 2SG-to.make like.this=DEM1.LOC

‘**Why** are you making this? (adapted from Ramirez, 2020a: 334)

e. *kóadzo kaláka lírhioka?*
kóadzo kaláka li-lhio-ka
how.many hen/rooster 1NFSG-to.have-SUB

‘**How many** hens does he have? (adapted from Ramirez, 2020a: 336)

f. *kadaliíta malíye piómaka?*
kadali-iíta malíye pi-oóma-ka
how.many-CLF.HUMAN knife 2SG-to.want-SUB

‘**How many** knives do you want?
(adapted from Ramirez, 2020a: 336)

Some interrogatives take classifier suffixes; these are analyzed in § 5.3.8 and § 6.1.1.5.

Part II

The classifier system

Chapter 4

Definition and delimitation

This chapter introduces the classifier system in Baniwa and serves to define and delimit the classifier set. In § 4.1, a definition of classifiers in Baniwa is provided. In § 4.2, the full inventory of 53 classifiers, determined on the basis of the definition, is presented. § 4.3 underscores the boundaries of the classifier class by discussing it in relation to other morphemes with which it shares some properties. Finally, § 4.4 describes the general principles of classifier assignment.

4.1 Definition

Classifiers in Baniwa are suffixes that occur on a number of elements within the noun phrase, denoting some semantic property of the real-world referent of the noun phrase head. In order to delimit classifiers as a class and to separate it from other parts of speech, I define as classifiers only the set of mutually exclusive suffixes that are capable of:

i. attaching directly to the root of the numerals 1–3,

ii. allowing the creation of a noun phrase consisting of a numeral 1–3 (with a classifier) and a noun: NUM-CLF N,

iii. producing a noun phrase expressing the quantity of the noun in question: 'one/two/three N'.

In the defining construction, N may be either a free or a bound noun (but in case it is a bound noun, it appears with an obligatory possessor prefix; see § 3.2.3).

An example of a morpheme that can be identified as a classifier based on the criteria above is the generic classifier *-da*. (31a) shows *-da* attached to a numeral that forms a noun phrase with a free noun (NUM-CLF N), and (31b) shows the same classifier and numeral being used with a bound noun with a possessor prefix (NUM-CLF POSS-N).

(31) a. *apáda* *hiipáda*
apa-**da** hiipáda
one-**CLF.GNR** stone

'one stone' [200228_ELIC_F_01]

b. *apáda* *lidákeeri*
apa-**da** li-dakee-ri
one-**CLF.GNR** 3NFSG-grandchild-NF

'one grandson of his' [200311_ELIC_K_06]

In § 1.1.3, it was shown that previous descriptions of the Baniwa classifier system differ considerably in terms of the classifier inventories they identify (see Table 1.1). A clear definition is useful, as it provides concrete criteria for the inclusion or exclusion of a morpheme in the classifier class[52]. This is particularly important, as Baniwa classifiers share some properties with other parts of speech, which makes it difficult to separate them in some contexts (see further in § 4.3).

The reason for using lower numerals as the definitional context is a practical one. Numerals are bound forms that never occur without one of a limited set of suffixes—the ones that I have chosen to label classifiers[53]. In particular, bound nouns cannot attach to numerals (*NUM-N). On the other hand, classifiers and bound nouns can both occur as the second part of complex nominal word formations on free nouns (N-CLF, N-N), in many cases both formally and semantically indistinguishable from each other (see § 4.3.1, § 8.2). This makes free nouns less suitable as a definitional context. Numerals also do not require any other suffixes than classifiers, in contrast to, e.g., attributive adjectives, where most

[52]Neither Aikhenvald nor Ramirez provides a precise definition of classifiers: Aikhenvald (2007: 479) declares that "[c]lassifiers in Baniwa/Kurripako [...] occur in various morphosyntactic environments", which are specified later in the text. Ramirez (2020a:233, my translation) states that "in Baniwa-Koripako, there is a series of morphemes that are automatically suffixed to the main determiners of the noun phrase (adjectives in attributive function, the equivalent of relative clauses, numerals, etc.)". Both of these descriptions are broad enough to also encompass other morphemes, such as various nominalizers, which are nevertheless absent from their inventories.

[53]To my knowledge, the only exception is the morpheme *-wali* 'times', which I exclude from the classifier set for other reasons—see § 4.3.3.

classifiers only occur in combination with another nominalizing suffix (referred to as their "complex form" in this thesis, see § 5.3.4). Numerals thus allow for a more straightforward identification of the underlying classifier forms, which is difficult to determine with certainty on adjectives due to the complex morphophonological phenomena at morpheme boundaries (see § 3.1.3.2, § 3.1.3.3).

4.2 Inventory

Based on the definition in § 4.1, 53 classifiers can be identified. These are presented in alphabetical order in Table 4.1, along with their label (used as a gloss), and examples of the referents they associate with.

The inventory is divided into individual classifiers according to the principle one form = one morpheme = one label. In other words, no homonymous classifiers are recognized. For example, the semantic extension of *-iíta* is difficult to subsume under a single label; it can be used to refer to (among other things) humans, some (but not all) kinds of fish, insects, and some objects which have a flat part (for example, knives, paddles, and cell phones). Nevertheless, *-iíta* is treated as a single classifier, rather than several homonymous ones. The alternative way of dividing up the inventory—to allow for homonymous morphemes in order to account for widely differing semantics (e.g., *-iíta*^{1} 'CLF.HUMAN', *-iíta*^{2} 'CLF.FISH', *-iíta*^{3} 'CLF.INSECT', *-iíta*^{4} 'CLF.FLAT.OBJECT')—was opted out of due to the difficulty of drawing a line between homonymy and polysemy in a non-arbitrary way.

The two classifiers *-ma*^{1} 'CLF.PAIR' and *-ma*^{2} 'CLF.FEMALE' look identical in their citation forms, but are underlyingly different, since *-ma*^{2} regularly causes lengthening of the initial vowel in *apa-* 'one' (32a), a morphophonological effect that it shares with a few other classifiers (see § 5.2.2). The female classifier *-ma*^{1} does not cause initial lengthening (32b), meaning that there is a formal difference that justifies the treatment of *-ma*^{1} and *-ma*^{2} as separate classifiers.

(32) a. ***aa****páma* *íinaro*
apa-ma^{2} íinaro
one-CLF.FEMALE woman

'one woman' [200319_FREE_H_01]

b. ***a****páma* *tshapáto*
apa-ma^{1} tshapáto
one-CLF.PAIR shoe

'one pair of shoes' [200319_FREE_H_01]

Table 4.1: Classifier inventory

Classifier	Label	Examples
-Ø	CLF.CANOE	Canoes, calabash bowls
-áana	CLF.GROUP	Groups (e.g., of people, animals, objects)
-áanhaa	CLF.LIQUID	Liquids (e.g., water, juices, drinks)
-aápa	CLF.OBLONG	Oblong fruits (e.g., banana, corn, manioc), some birds
-áaphi	CLF.AREA	Manioc gardens, lakes, villages
-áapi	CLF.HOLLOW	Plates, bowls, pots, vessels, bones
-áapo	CLF.STICK	Sticks, fishing rods, trees, paths, creeks
-da	CLF.GENERIC	Almost anything, round objects (e.g., stones, fruits)
-daa	CLF.DAY	Days, nights
-dápana	CLF.HOUSE	Houses, buildings, camps
-eéma	CLF.SIDE	Sides, halves, one of a pair (e.g., animals, body parts)
-éekhe	CLF.SMALL.SEED	Small seeds (esp. of chili peppers)
-éewhe	CLF.EGG	Eggs
-híko	CLF.LONG	Trees, sugar canes, bottles
-hípa	CLF.MALE	Male humans
-hipáda	CLF.PIECE	Pieces (e.g., meat, fruit, wood)
-hípani	CLF.RAPIDS	Rapids
-híwa	CLF.BEIJU	*Beiju* (manioc bread)
-híwi	CLF.POINTED	Long, thin, pointed objects (e.g., needles, hooks, arrows)
-i	CLF.BASKET	Baskets, bunches of palm tree fruit)
-íida	CLF.HALF	Halves, sections (e.g., animals, houses, land)
-iíta	CLF.HUMAN	Humans, some fish, insects, knives, paddles
-iítsia	CLF.BUNCH	Bunches of palm tree fruit (without the fruit)
-íiwi	CLF.FLOWER	Flowers, thorns
-íixi	CLF.SEED	Seeds, kernels
-kénaa	CLF.BRANCH	Tree branches
-khaa	CLF.CURVILINEAR	Long, thin, flexible things (e.g., snakes, rivers, lianas)
-ko	CLF.HAMMOCK	Hammocks
-koa	CLF.SURFACE	Flat surfaces (e.g., beaches, mirrors, tables)
-kódzoa	CLF.BEND	River bends
-ma[1]	CLF.PAIR	Pairs (e.g., of shoes, certain flutes)
-ma[2]	CLF.FEMALE	Female humans
-máka	CLF.FABRIC	Fabrics (e.g., cloths, towels)
-na	CLF.TRUNK	Trees, some mammals (e.g., dogs, jaguars, tapirs), some fish
-náko	CLF.BUNDLE	Bundles (e.g., of sticks, firewood, sugar cane)
-pa	CLF.PACKAGE	Packages, parcels, leaf-wraps (e.g., of coffee, salt, fish)
-páwa	CLF.RIVER	Rivers, tributaries, creeks
-péko	CLF.PATH	Paths, rivers
-phe	CLF.LEAF	Leaves, feathers, sheets (e.g., of paper, metal), fabrics
-pi	CLF.TUBE	Palm trees, rifles, blowguns, *tipitís* (manioc squeezers)
-póko	CLF.CIRCLE	Circular objects, bunches (e.g., Amazon grape), quarters of animals
-pokóda	CLF.STUMP	Tree stumps, tree trunks
-tsoi	CLF.PILE	Piles, heaps (e.g., of stones, soil, fruit)
-ttáwalhe	CLF.CUT	Cut off pieces (e.g., of wood, meat, land, sugar cane, soap)
-ttówhia	CLF.ROOM	Rooms
-wa	CLF.SPACE	Holes, doors, villages
-wálhia	CLF.YEAR	Years
-wána	CLF.SLICE	Slices, pieces, shards (e.g., of *beiju*, land, glass, ceramics)
-wáta	CLF.BRACT	*Curuatás* (leaves around flower bases), containers
-wáthe	CLF.NODE	Knots, joints of sugar canes, elbows
-xaa	CLF.EXCREMENT	Excrement
-ya	CLF.SKIN	Bark, skin
-yáwa	CLF.HOLE	Holes, doors, wounds

One classifier is analyzed as a zero morpheme: *-Ø* 'CLF.CANOE' (33a–33b). This analysis is justified by the fact that the complex form of this classifier is *-Ø-li* (33b), in analogy with other classifiers that take the *-li* suffix in this context, such as *-híwa* 'CLF.BEIJU' (34a–34b). Complex forms are described in more detail in § 5.3.2.

(33) a. *ápa*
apa-**Ø**
one-**CLF.CANOE**

'one (of a canoe)' [200323_ELIC_F_01]

b. *makáli*
maka-**Ø**-li
big-**CLF.CANOE**-NF

'big (of a canoe)' [220918_ELIC_F_01]

(34) a. *aphéwa*
apa-**híwa**
one-**CLF.BEIJU**

'one (of a *beiju*)' [200315_ELIC_F_01]

b. *makhéwali*
maka-**híwa**-li
big-**CLF.BEIJU**-NF

'big (of a *beiju*)' [220918_ELIC_F_01]

4.3 Non-classifier categories with related properties

The classifier class is delimited by the definition in § 4.1, but there are several morphemes that do not qualify as classifiers according to the definition, but show similar characteristics due to common historical origins, or to grammatical or semantic semblance.

First, many classifiers can be shown to originate from bound nouns (see further in § 8.2), and therefore these two parts of speech share many characteristics. Several classifiers have bound noun counterparts that are highly similar or identical in both form and meaning, such as the classifier *-éewhe* 'CLF.EGG' and the

bound noun *-éewhe* 'egg'. The relationship between classifiers and bound nouns is treated in § 4.3.1.

Second, the plural suffix *-pe* (§ 3.2.5.3) competes for the same morphosyntactic slot as the classifiers in some contexts, for example on some nominalized verb forms. It is formally excluded from the classifier inventory on the basis of criterion i. (§ 4.1), as it cannot be attached directly to numerals. The relationship between classifiers and plural suffixes is treated in more detail in § 4.3.2.

Third, there is a morpheme *-wali* 'times' which attaches to numerals in the same slot as classifiers do. Most commonly, *-wali* is used to denote the number of times an event occurs. It is excluded from the classifier set on the basis of a combination of the definitional criteria ii. and iii. (§ 4.1), explained in more detail in § 4.3.3.

Finally, Baniwa has a bipartite grammatical gender system which shares some characteristics with the classifier system, as they are both nominal classification systems (see Seifart, 2010: 719). The two gender markers are easily excluded from the classifier set, as they occur in different morphosyntactic contexts—for instance, gender is not marked on numerals, which is the definitional context of classifiers (§ 4.1). The relationship and interaction between the classifier and gender systems is treated in § 4.3.4.

4.3.1 Bound nouns

As described in § 3.2.5, Baniwa has separate classes of free (alienable) and bound (inalienable) nouns. Most classifier morphemes can be shown to have developed from the class of bound nouns, a diachronic relationship described in detail in § 8.2. As the shift from one word class to another is a gradual process, it is unsurprising that bound nouns and classifiers share some properties.

For example, bound nouns and classifiers sometimes occur in the same locus: on free nouns, classifiers can occur as suffixes and bound nouns can be attached as the second part of compounds. Despite the similar surface structure, in many cases, it is clear if this second element is a classifier (35a) or a bound noun (35b), simply because no homonymous morpheme exists in the other category. However, 16 of the classifiers have an identical form that is synchronically used as a bound noun[54]. These include *-éewhe*, which means that the second element in (35c) can be analyzed as the classifier 'CLF.EGG' (in analogy with (35a)) or as the bound noun *-éewhe* 'egg' (in analogy with (35b))—both analyses are equally

[54]The test for bound nounhood used here is whether or not the form can be used with a person prefix in a possessive construction (see further in § 8.1).

plausible, from both a formal and a semantic perspective[55].

(35) a. *tapeedápana*
tápee-**dápana**
medicine-**CLF.HOUSE**

'hospital' [200328_ELIC_F_04]

b. *panttinóma*
pántti-**nóma**
house-**mouth**

'door' [200305_ELIC_F_01]

c. *kalakéewhe*
kaláka-**éewhe**
hen/rooster-**CLF.EGG** / **egg** ?

'hen's egg' [200319_FREE_H_01]

4.3.2 Plural marker *-pe*

In some contexts in Baniwa where classifiers are used, classifier distinctions that are made in the singular are neutralized in the plural, because the plural marker *-pe* (§ 3.2.5.3) competes for the same slot as classifiers. One such case is the alienable possession construction (36), described further in § 5.3.5.1.

(36) a. *lhíehẽ* *pántti* *nodzaarodápana*
lhíehẽ pántti no-dzaa-ro-**dápana**
DEM1.NFSG house 1SG-belongings-F-**CLF.HOUSE**

'this house of mine' [220919_ELIC_F_01]

b. *nháahã* *pántti* *nodzaarópe*
nháahã pántti no-dzaa-ro-**pe**
DEM1.NFSG house 1SG-belongings-F-**PL**

'these houses of mine' [221006_ELIC_G_01]

[55]The full list of the 16 classifiers that have an identical corresponding bound noun is presented later in this thesis, in Table 8.3 (§ 8.1).

Another example of a construction where the plural marker *-pe* replaces classifiers is on verbs that are nominalized with the agent nominalizer *-ka* (§ 5.3.6.1), as shown in (37) where the nominalized verbs act as relative clauses.

(37) a. *haikóapo* *hiwakáapo*
háiko-áapo i-híwa-ka-**áapo**
tree/wood-CLF.STICK CONN-to.fall-NMLZ.AG-**CLF.STICK**

'(the) stick that is falling' [220929_ELIC_G_01]

b. *haikóapo* *hiwakápe*
háiko-áapo i-híwa-ka-**pe**
tree/wood-CLF.STICK CONN-to.fall-NMLZ.AG-**PL**

'(the) sticks that are falling' [221006_ELIC_G_01]

However, in nominalizations with the location nominalizer *-karo* (§ 5.3.6.3), the plural marker follows the classifier instead of replacing it (38). This clearly shows that the plural marker does not belong to the set of classifiers.

(38) a. *pherakaróda*
pa-híra-karo-**da**
IMPRS-to.climb-NMLZ.LOC-**CLF.GNR**

'ladder' [200305_ELIC_F_01]

b. *paiñhakarodápe*
pa-íiñha-karo-**da-pe**
IMPRS-to.eat-NMLZ.LOC-**CLF.GNR-PL**

'kitchen utensils' (collective term) [200315_ELIC_F_01]

Most importantly, from a definitional perspective, the plural marker does not attach directly to numeral roots, which excludes them from the classifier set on the basis of criterion i. (§ 4.1). In ordinary numeral constructions, no plural suffix occurs even on numbers higher than 'one' (neither on the numeral nor the noun, see § 3.2.5.3). However, Ramirez (2020a: 245) notes that plural *-pe* may follow the classifier on a numeral in expressions like 'one by one', 'two by two', etc. (39).

(39) *dóomali* *híwa* *apadápe*
dóomali i-híwa apa-**da-pe**
umari CONN-to.fall one-**CLF.GNR-PL**

'The *umari* fruits fall one by one.' (adapted from Ramirez, 2020a: 245)

4.3.3 The morpheme *-wali* 'times'

The morpheme *-wali* 'times' constitutes a borderline case for inclusion in the classifier set. Most commonly, it attaches to a numeral and is used without any overt noun to denote the number of times an event occurs, as in (40).

(40) *apáwali* *nomheéta* *nonóma*
apa-**wali** no-mheéta no-nóma
one-**times** 1SG-to.open 1SG-mouth

'I open my mouth once.' [200310_ELIC_K_02]

This morpheme is a borderline case since it fulfills criterion i. in the definition of classifiers presented in § 4.1—that is, it attaches directly to numeral roots in the same slot that classifiers do. It can even be said to fulfill criterion ii., as it does occur with an overt noun (*heekóapi* 'day') in at least one construction. However, it is excluded from the classifier set on the basis of criterion iii., as the only construction in which it has been recorded with an overt noun (41) does not refer to a quantity of the noun in question, but instead is a conventionalized expression used in the beginning of narratives. This has also been noted by Ramirez (2020a: 237), who nevertheless analyzes *-wali* as a classifier.

(41) *apáwali* *heekóapi…*
apa-**wali** heekóapi
one-**times** day

'One day/Once upon a time…' [200313_FREE_F_01]

In order to refer to a quantity of the noun *heekóapi* 'day', such as 'one day' or 'two days', etc., the classifier *-daa* 'CLF.DAY' is used (42).

(42) *lhíatsheenahã* *apádaa* *heekóapikatsa*
lhíatsha-iina=áa(hã) apa-**daa** heekóapi-ka-tsa
however-already=DEM1.LOC one-**CLF.DAY** day-SUB-RESTR

pamettátakatsani *dorómeri…* *óohõ*
pa-mettáta-ka-tsa-ni doróme-li óohõ
IMPRS-to.dry-SUB-RESTR-3NFSG time.consuming-REL yes

'But still, this one takes a whole day just to dry it, which is time-consuming… yes.' [200308_NARR_B_01]

The morpheme *-wali* is also used on the numeral *apa-* 'one' in the conventionalized expression 'sometimes', together with the plural suffix *-pe* (43).

(43) *apawalípe* *rhoána* *hipakádani…*
apa-**wali-pe** rhoána i-hípa-ka-da-ni
one-**times-PL** DEM3.FSG CONN-to.wash-NMLZ.AG-CLF.GNR-3NFSG

pikitsindátakano
pi-kitsindáta-ka-no
2SG-to.help-SUB-3FSG

'Sometimes, she who washes it… you have to help her.'
[200308_NARR_B_01]

4.3.4 Gender

Besides its classifiers, Baniwa has another nominal classification system: a grammatical gender system with two distinctions, feminine *-ro* and non-feminine *-li* (§ 3.2.5.4). Gender and classifiers occur in different grammatical environments and do not compete with each other for the same slots. Thus, it is easy to exclude the two gender markers from the classifier set, simply because gender is not marked on numerals, which is the definitional context of classifiers (§ 4.1). However, there are certain contexts where there is interaction between the two systems.

One such context is attributive adjectives, where the classifier is accompanied by an additional nominalizing suffix in most cases (see further in § 5.3.4). In this environment, either *-ro* or *-li* is used as the nominalizer with a subset of the classifiers (Group I, see further in § 5.3.2). The choice between *-ro* and *-li* depends on the whether the referent is feminine (*tsoo-**dá-ro*** /small-**CLF.GNR-F**/ 'small (about a woman)') or non-feminine (*tsoo-**dá-li*** /small-**CLF.GNR-NF**/ 'small (about, e.g., a stone)').

Classifiers and gender also interact on demonstratives. Demonstratives only take classifiers when they are used contrastively (see § 5.3.7). Demonstratives themselves, however, have different forms in the singular depending on gender (see § 3.2.10): for example, the proximal demonstratives *lhia* and *rhoa* are used for non-feminine and feminine referents, respectively. When classifiers are used in this construction, it means that they choose a suitable demonstrative form to attach to depending on whether the referent is feminine (*rhóa-ma*^{2} /DEM1.FSG-CLF.FEMALE/ 'this (one)' (contrastive, about a woman)) or non-feminine (*lhía-da* /DEM1.NFSG-CLF.GENERIC/ 'this (one)' (contrastive, about, e.g., a stone)).

To my knowledge, the only example of a gender marker used in a conventionalized, i.e., non-transparent way, is in the alienable possession construction described in § 5.3.5.1. In this construction, the bound noun *-dzaa* 'belongings' takes two suffixes: the feminine marker *-ro* and a classifier (44). In this construction, there is no alternation between *-li* and *-ro* depending on the gender of the referent.

(44) *lhíehẽ* *papéra* *nodzaaróphe*
lhíehẽ papéra no-dzaa-**ro-phe**
DEM1.NFSG paper 1SG-belongings-**F-CLF.LEAF**

'this paper of mine' [220918_ELIC_F_02]

4.4 The nature of classifier assignment

In any nominal classification system, it is, by definition, nouns that are the grammatical elements being classified. In Baniwa, the classification applies to all nouns, both free and bound ones. Free nouns can stand on their own, while bound nouns require an explicit possessor. Each can be converted to the other category (§ 3.2.5.1). Free nouns can be classified as they are (45), while bound nouns either have to overtly mark possession via a person prefix (46a), or be converted to a free noun by the alieanator construction consisting of the prefix *i-* and the suffix *-tti* (46b).

(45) *apáapi* *paráto*
apa-áapi **paráto**
one-CLF.HOLLOW **plate**

'one plate' [200321_FREE_A_01]

(46) a. *apakénaa* *líke*
apa-kénaa li-**ke**
one-CLF.BRANCH 3NFSG-**branch**

'one branch (of his/its)' [200312_ELIC_K_01]

b. *apáda* *ixadátti*
apa-da i-**xáda**-tti
one-CLF.GNR AL-**belly**-AL

'one belly' [200310_ELIC_K_02]

Baniwa classifiers encode semantic properties such as shape, animacy, and part–whole relations (see further in Chapter 7). The assignment is flexible, in the sense that one and the same noun can generally occur with more than one classifier[56], as exemplified in (47) with *malíitsi* which has the basic meaning 'grass'. This demonstrates that the actual entities that are being classified are not the nouns themselves, e.g., *malíitsi*, but the real-world referents, e.g., a certain field of grass, or a certain blade of grass (see § 2.3).

(47) *malíitsi* 'grass' [200312_ELIC_K_01]

a. *apáaphi* *malíitsi*
apa-**áaphi** malíitsi
one-**CLF.AREA** grass
'one field of grass'

b. *aphéwi* *malíitsi*
apa-**híwi** malíitsi
one-**CLF.POINTED** grass
'one blade of grass'

Even the same real-world referent can be assigned different classifiers, as long as there is more than one classifier that is semantically compatible with the referent in question. The choice can also depend on which feature of the referent the speaker chooses to emphasize. (48) shows that at least three different classifiers can be used more or less interchangeably to refer to rivers, in combination with the noun *óoni*.

(48) *óoni* 'water, river' [200326_ELIC_F_01]

a. *apapáwa* *óoni*
apa-**páwa** óoni
one-**CLF.RIVER** water
'one river'

b. *apapéko* *óoni*
apa-**péko** óoni
one-**CLF.PATH** water
'one river'

c. *apákhaa* *óoni*
apa-**khaa** óoni
one-**CLF.CURV** water
'one river'

Classifier assignment in Baniwa thus operates on a semantic basis. This is typical of less grammaticalized nominal classification systems, and stands in contrast to

[56] In the Noun list data set (§ 1.2.4), speakers listed up to seven different classifiers as compatible with a single noun (average 1.89).

more grammaticalized systems where the assignment may, in addition, be based on phonological or morphological criteria (see § 2.1). However, phonological and morphological criteria presuppose that the system operates on lexeme-level, as real-world referents do not have phonological or morphological properties, which excludes this possibility in the Baniwa case (see Corbett, 1991: 33 ff.).

Classifier assignment in Baniwa is, in most cases, relatively semantically transparent (see Chapter 7). For example, objects with large, flat surfaces such as tables, mirrors, and manioc ovens typically associate with the surface classifier *-koa*. Likewise, long, thin and pointed objects such as needles and pencils tend to associate with the classifier *-híwi*. However, there is also some degree of opacity in the system, at least to an outside observer. For instance, the classifier *-khaa* is a shape classifier normally used for long, thin, and flexible objects like lianas and snakes, but it is also used for ideas, languages, and songs (see § 7.2.2.7). In this case, it is the shift from concrete to abstract that turns the assignment more opaque. Another example is the classifier *-iíta* which is used for both humans, objects with flat parts, and some fish, among other things (§ 7.1.22). In this case, it is the complex semantic structure of the classifier that makes the assignment principles opaque.

4.5 Chapter summary

This chapter has served to establish and delimit what constitutes the classifier system in Baniwa, a prerequisite for the ensuing analysis in the subsequent chapters. I have provided a definition of Baniwa classifiers that allows for the identification of 53 formally distinct classifier morphemes, and at the same time excludes other morphemes (such as bound nouns, plural and gender markers, and the morpheme *-wali* 'times') from being treated as classifiers. The analysis in the following chapters is based on the inventory identified here.

This chapter has also introduced the basic assignment principles of the classifier system, which are central to understanding the nature of the system. Classifiers associate with referents rather than nouns, and the assignment operates on a semantic basis that allows for a high degree of flexibility.

Chapter 5

Formal properties

This chapter describes the formal properties of classifiers, that is, their phonological, morphophonological, and morphological behavior, as well as their morphosyntactic distribution. § 5.1 introduces the basic phonological characteristics of the classifier morphemes, such as syllable structure and stress placement. § 5.2 describes how classifiers are involved in morphophonological processes across morpheme boundaries when attached to host words: vowel fusion, /h/-metathesis, and initial vowel lengthening. § 5.3 describes the morphosyntactic distribution of classifiers. The simplex and complex forms of classifiers, which occur in different loci, are presented (§ 5.3.2). The morphological characteristics of classifier the their various loci are then outlined: numerals (§ 5.3.3), adjectives (§ 5.3.4), nouns (§ 5.3.5), nominalized verbs (§ 5.3.6), demonstratives (§ 5.3.7), and interrogatives (§ 5.3.8).

5.1 Phonology

Classifier morphemes in Baniwa contain up to three syllables (σ). The distribution is shown in Figure 5.1. Disyllabic classifiers are most common (31 classifiers, 58%), followed by monosyllabic (16 classifiers, 30%) and trisyllabic ones (5 classifiers, 9%). One classifier, *-Ø* 'CLF.CANOE', is analyzed as a zero morpheme (see § 4.2), thus in theory consisting of zero syllables. In Arawakan languages, affixes are mostly monosyllabic, and those that are disyllabic typically go back to grammaticalized free forms (Aikhenvald, 2020: 5). This is the case for many of the classifiers, as we will see in Chapter 8. Most of the trisyllabic classifiers are etymologically analyzable.

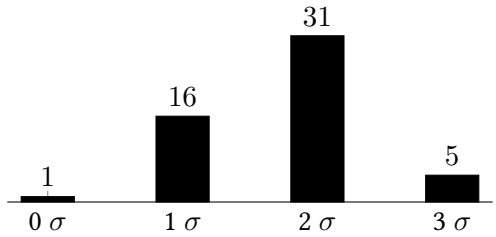

Figure 5.1: Classifiers by number of syllables

Lexical stress in Baniwa always falls on the penultimate metric syllable (see § 3.1.2.2). Most classifier morphemes are fully metric, meaning that they count with respect to stress placement. An example of a typical metric classifier is *-ya* 'CLF.SKIN' (49a). The stress in this case falls on the penultimate syllable of the word. However, there is one classifier, *-wa* 'CLF.SPACE', which is entirely extrametric (Ramirez, 2001a: 315), as can be seen in example (49b), where the stress falls on the antepenultimate syllable. Apart from *-wa*, three of the trisyllabic classifiers (*-dápana*, *-hípani*, and *-ttáwalhe*) end in an extrametric (underlined) syllable, but their first two syllables are metric, resulting in antepenultimate stress. Extrametric syllables are not exclusive to classifiers, but are also observed elsewhere in the lexicon (see § 3.1.2).

(49) a. *apáya*
apa-ya
one-CLF.SKIN

'one (skin)' [200302_ELIC_F_01]

b. *ápawa*
apa-wa
one-CLF.SPACE

'one (hole)' [200325_ELIC_F_02]

5.2 Morphophonology

5.2.1 Vowel fusion and /h/-metathesis

Vowel-initial classifiers behave like other vowel-initial bound morphemes when suffixed to a host word, i.e., they trigger vowel fusion (see § 3.1.3.4). This is illustrated in (50), where *a + ii* → *ee*.

(50) *apéexi* *líixi*
ap**a-íi**xi li-íixi
one-CLF.SEED 3NFSG-seed

'one seed (of him/it)' [200325_ELIC_F_02]

Classifiers beginning in /h/, like other /h/-initial morphemes, trigger both vowel fusion (see § 3.1.3.4) and /h/-metathesis (see § 3.1.3.2). This is illustrated in (51), where *a + (h)i → (h)e*.

(51) *a**phé**wi* *áawi*
a**pa-hí**wi áawi
one-CLF.POINTED needle

'one needle' [200302_ELIC_F_01]

5.2.2 Lengthening of initial vowel

When classifiers are suffixed to the numeral *apa-* 'one', the majority of them—e.g., *-da* 'CLF.GENERIC' in (52a)—leave the initial short vowel of *apa-* unaffected. However, a few classifiers cause lengthening of this vowel ([apa] → [a:pa]), as in (52b). The classifiers that do this are *-ma²* 'CLF.FEMALE', *-na* 'CLF.TRUNK', and *-pi* 'CLF.TUBE'. One classifier, *-ko* 'CLF.HAMMOCK' shows inter-speaker variation, and can yield either [apako] or [a:pako].

(52) a. ***a**páda* *itsída*
apa-da itsída
one-CLF.GNR tortoise

'one tortoise' [200311_ELIC_K_06]

b. ***aa**páma* *íinaro*
apa-ma² íinaro
one-CLF.FEMALE woman

'one woman' [200311_ELIC_K_04]

Table 5.1 lists the classifiers that cause initial lengthening in *apa-* along with their forms on the numerals one, two, and three for comparison.

Table 5.1: Classifiers that cause initial vowel lengthening in *apa-* 'one'

Classifier	Gloss	***apa-* 'one'**	***dzama-* 'two'**	***madali-* 'three'**
-ma²	CLF.FEMALE	*aapáma*	*dzamáma*	*madalíma*
-na	CLF.TRUNK	*aapána*	*dzamána*	*madalína*
-pi	CLF.TUBE	*aapápi*	*dzamápi*	*madalípi*
(*-ko*)	CLF.HAMMOCK	*aapáko*∼*apáko*	*dzamáko*	*madalíko*

Apa- appears to be the only host word which is affected in this way by these classifiers. No other numeral displays this behavior, and neither does any other host word in any other morphosyntactic context where classifiers appear, to my knowledge. For comparison, (53) shows the numeral *dzama-* 'two' suffixed with the same two classifiers as in (52). In contrast to *apa-*, *dzama-* shows no variation in initial vowel length. Table 5.1 shows that this holds true across the four classifiers for both *dzama-* 'two' and *madali-* 'three'.

(53) a. *dz**a**máda*
dz**a**ma-da
two-CLF.GNR

'two' [220918_ELIC_F_01]

b. *dz**a**máma*
dz**a**ma-ma²
two-CLF.FEMALE

'two (women)' [220918_ELIC_F_01]

It is not clear why *apa-* is the only word which undergoes this process. Among the numerals that take classifiers (§ 5.3.3), it is the only one beginning in a vowel, which may be of relevance. In addition, *apa-* is a frequent morpheme, and frequently used values tend to be more differentiated than rare ones due to their "greater memory strength" (Haspelmath & Sims, 2010: 268–270, 272–273). This may explain the presence of more differentiation in *apa-* in relation to the other numerals, and perhaps also in relation to other (less frequent) morphemes which can act as hosts for classifiers.

It is also not clear why only the four classifiers in Table 5.1 cause lengthening of the initial vowel in *apa-*. Phonologically, they are all monosyllabic and metric, meaning that the stress falls on the same syllable (the penultimate one) in the host word. They begin with an obstruent (either a nasal or a plosive), which in three cases is labial (/m, n, p/) and in one case velar (/k/). However, there are also classifiers that share these properties but do not cause initial lengthening: most

notably *-ma*[1] 'CLF.PAIR' (which is homonymous with *-ma*[2] 'CLF.FEMALE' except for in this very construction), but also *-pa* 'CLF.PACKAGE'.

In the case of *-ma*[2] 'CLF.FEMALE' (which renders *aapáma*), part of the explanation could be homonymy avoidance with *-ma*[1] 'CLF.PAIR' (which renders *apáma*). A problem with this explanation is that while homonymy is avoided in one very frequent construction, the two classifiers still produce identical forms in all other morphosyntactic contexts (see § 5.3). Another problem is that it does not explain why *-na*, *-pi*, and *-ko* trigger initial vowel lengthening, as none of these have homonymous classifiers.

5.3 Morphology and syntax

5.3.1 Morphosyntactic distribution

Table 5.2 lists all the loci in which classifiers occur. The Form column indicates which set of forms, labelled *simplex* and *complex*, is used in each context (see § 5.3.2).

Table 5.2: Morphosyntactic loci of classifiers

Locus	Form	Section
Numerals	simplex	(§ 5.3.3)
Adjectives (attributive)	complex	(§ 5.3.4)
Nouns	simplex	(§ 5.3.5)
Verbs (nominalized)	simplex	(§ 5.3.6)
Demonstratives (contrastive)	simplex	(§ 5.3.7)
Interrogatives	simplex, complex	(§ 5.3.8)

Classifiers in Baniwa are primarily used within the noun phrase. They are found on nouns and on nominal modifiers such as numerals and attributive adjectives. When classifiers are used on verbs, they are always nominalized and occur either as nouns in their own right of as relative clauses in the noun phrase. Demonstratives are not marked with classifiers, except in contrastive situations. Classifiers are also found in some interrogative constructions. Classifiers are always suffixed to their host, but see § 5.3.3.2 for a (now obsolete) context where classifiers were used as infixes.

For Arawakan languages in general, Dunn & Rose (forthcoming) identify nouns, numerals, modifiers (e.g., adjectives and demonstratives), and verbs as the four major loci. In this sense, the distribution of classifiers in Baniwa is relatively representative of the family as a whole, except for the fact that classifiers are

not marked on verbal predicates (only on nominalized forms of verbs). However, Dunn & Rose also state that many Arawakan languages in Northwestern Amazonia tend to lack classifiers on verbs. An unusual feature within the family appears to be the marking of classifiers on interrogatives in Baniwa (Dunn & Rose, forthcoming: 25).

Classifiers can generally occur both with and without an overt noun. This is discussed in more detail in Chapter 6, which deals with the functions of classifiers.

Across the naturalistic texts (see § 1.2.5), classifiers occur 59 times. The total number of morpheme tokens in the texts is 2962, meaning that classifiers account for 1.99% of the morpheme tokens in Baniwa, or that roughly every 50th morpheme is a classifier. Table 5.3 lists all the morphosyntactic loci where classifiers have been found in the naturalistic texts. The numeral locus seems to be the most common one, accounting for more than a third of all occurrences.

Table 5.3: Frequency of classifier loci in naturalistic texts

Locus	Count	Share of clf tokens
Numeral	22	37.3%
Nominalized verb	17	28.8%
Noun	12	20.3%
Adjective	3	5.1%
Interrogative	3	5.1%
Demonstrative	2	3.4%

5.3.2 Simplex and complex classifier forms

There are two sets of classifier forms in Baniwa, labelled *simplex* and *complex* with reference to their morphological composition. The simplex form is the bare, monomorphemic form, which is also used as the citation form for the classifiers. The complex form is composed of a classifier and an additional suffix. In what follows, whenever they need to be differentiated, the simplex form will be referred to as CLF^{s}, and the complex form as CLF^{c}.

Each morphosyntactic locus where classifiers occur takes either the simplex or the complex set of classifiers (as specified in Table 5.2 above). For example, numerals always take the simplex form, and adjectives always take the complex form, as shown in (54). The additional suffix in the complex form is a nominalizer. This speaks to a functional difference between classifiers in different contexts, which will be further explored in § 6.1. In connection to this, the Baniwa classifier system cannot be said to exhibit agreement between constituents (although phrases such as the one in (54) may invite such an interpretation)—this

will be discussed in § 6.5.

(54) *aapá**na*** *háiko* *mhatsamé**ne***
apa-**na** háiko mhatsa-me-**na-i**
one-**CLF.TRUNK** tree/wood snap-ADJZ-**CLF.TRUNK-N**

'one broken (snapped) piece of wood' [200322_ELIC_F_01]

Despite the functional differences, it is useful to think of the forms as constituting two classifier sets, because the choice of nominalizer in the complex form is conditioned by the classifier. The classifiers can be divided into four groups based on which of the four nominalizers *-i* 'N' (neuter), *-li* 'NF' (non-feminine), *-ro* 'F' (feminine), or *-ni* '?' they combine with (Table 5.4). Group I takes either *-li* or *-ro*, group II takes *-i*, and group III takes *-ni*. Group IV does not take any suffix, but uses the same (simplex) form in both sets.

5.3.2.1 Group I: *-li* 'NF' or *-ro* 'F'

Group I includes the 12 classifiers that take the non-feminine *-li* and/or feminine *-ro* suffixes on adjectives (23% of the classifier set)[57]. As mentioned in § 5.3.2, these two morphemes likely share an origin with the homonymous derivational suffixes *-li* and *-ro*, also used for marking gender elsewhere in the language (§ 3.2.5.4). Gender is systematically encoded through the choice of either *-li* or *-ro* whenever the complex form of classifiers is used, which is why the classifiers taking these morphemes are lumped together here.

Gender assignment in Baniwa is semantic (§ 3.2.5.4): all feminine referents are covered by *-ro*, and everything else is covered by *-li*. Most of the classifiers in Group II are only compatible with *-li* 'NF', due to the fact that their referents all belong to the non-feminine category. The opposite is true for the classifier *-ma²* 'CLF.FEMALE', whose referents are all in the feminine category. The only classifier whose possible referents span these two categories, generic *-da*, accordingly is compatible with both suffixes, depending on its referent. This interaction of classifiers and gender on adjectives is cross-linguistically unusual, as pointed out by Aikhenvald (2019; 2020). (55) shows two examples of *-da* on adjectives, one with a feminine referent (55a) and one with a non-feminine referent (55b).

[57] The complex form of one classifier in this group, *-xaa* 'CLF.EXCREMENT', is irregular: *-(íi)xali*.

Table 5.4: Classifiers grouped according to their complex form

	Gloss		Simplex	Complex	(Analysis)
Group I	CLF.CANOE		*-Ø*	*-Ø-li*	
	CLF.OBLONG		*-aápa*	*-aápa-li*	
	CLF.GENERIC		*-da*	*-da-li/-da-ro*	
	CLF.SIDE		*-eéma*	*-eéma-li*	
	CLF.MALE		*-hípa*	*-hípa-li*	
	CLF.BEIJU		*-híwa*	*-híwa-li*	
	CLF.BASKET		*-i*	*-iri* (*-i-li*)	
	CLF.PAIR		*-ma*^{1}	*-ma*^{1}*-li*	
	CLF.FEMALE		*-ma*^{2}	*-ma*^{2}*-ro*	
	CLF.PACKAGE		*-pa*	*-pa-li*	
	CLF.EXCREMENT		*-xaa*	*-(íi)xali*	*-(íi)-xa-li*
	CLF.SKIN		*-ya*	*-ya-li*	
Group II	CLF.GROUP		*-áana*	*-áane*	*-áana-i*
	CLF.LIQUID		*-áanhaa*	*-áanhai*	*-áanhaa-i*
	CLF.AREA	?	*-áaphi*	*-áaphi*	*-áaphi-i*?
	CLF.HOLLOW	?	*-áapi*	*-áapi*	*-áapi-i*?
	CLF.STICK	?	*-áapo*	*-áapo*	*-áapo-i*?
	CLF.LONG		*-híko*	*-híki*	*-híko-i*
	CLF.PIECE		*-hipáda*	*-hipáde*	*-hipáda-i*
	CLF.RAPIDS	?	*-hípani*	*-hípani*	*-hípani-i*?
	CLF.POINTED	?	*-híwi*	*-híwi*	*-híwi-i*?
	CLF.HUMAN		*-iíta*	*-iíte*	*-iíta-i*
	CLF.FLOWER	?	*-íiwi*	*-íiwi*	*-íiwi-i*?
	CLF.SEED	?	*-íixi*	*-íixi*	*-íixi-i*?
	CLF.CURVILINEAR		*-khaa*	*-khai*	*-khaa-i*
	CLF.HAMMOCK		*-ko*	*-ko*∼*-ki*	*-ko-i*
	CLF.SURFACE		*-koa*	*-koe*	*-koa-i*
	CLF.FABRIC		*-máka*	*-máke*	*-máka-i*
	CLF.TRUNK		*-na*	*-ne*	*-na-i*
	CLF.BUNDLE		*-náko*	*-náko*∼*-náki*	*-náko-i*
	CLF.PATH		*-péko*	*-péki*	*-péko-i*
	CLF.LEAF		*-phe*	*-phai**	*-phe-i*
	CLF.TUBE	?	*-pi*	*-pi*	*-pi-i*?
	CLF.CIRCLE		*-póko*	*-póko*∼*-póki*	*-póko-i*
	CLF.STUMP		*-pokóda*	*-pokóde*	*-pokóda-i*
	CLF.PILE	?	*-tsoi*	*-tsoi*	*-tsoi-i*?
	CLF.ROOM		*-ttówhia*	*-ttówhie*	*-ttówhia-i*
	CLF.SLICE		*-wána*	*-wáne*	*-wána-i*
	CLF.BRACT		*-wáta*	*-wáte*	*-wáta-i*
Group III	CLF.RIVER		*-páwa*	*-páwa-ni*	
	CLF.SPACE		*-wa*	*-wa-ni*	
	CLF.HOLE		*-yáwa*	*-yáwa-ni*	
Group IV	CLF.DAY		*-daa*	*-daa*	
	CLF.HOUSE		*-dápana*	*-dápana*	
	CLF.SMALL.SEED		*-éekhe*	*-éekhe*	
	CLF.EGG		*-éewhe*	*-éewhe*	
	CLF.BUNCH		*-iítsia*	*-iítsia*	
	CLF.SIDE		*-íida*	*-íida*	
	CLF.BRANCH		*-kénaa*	*-kénaa*	
	CLF.BEND		*-kódzoa*	*-kódzoa*	
	CLF.CUT		*-ttáwalhe*	*-ttáwalhe*	
	CLF.YEAR		*-wálhia*	*-wálhia*	
	CLF.NODE		*-wáthe*	*-wáthe*	

(55) a. *íinaro madoádaro*
íinaro mádoa-**da-ro**
woman short-**CLF.GNR-F**

'short woman' [200325_ELIC_F_02]

b. *hiipáda iittádali*
hiipáda íitta-**da-li**
stone black-**CLF.GNR-NF**

'black stone' [200228_ELIC_F_01]

5.3.2.2 Group II: *-i* 'N'

Group II contains the 27 classifiers that take *-i* on adjectives (51% of the classifier set). According to Ramirez (2020a: 371–375), *-i* may have developed from an old neuter marker that used to exist alongside non-feminine *-li* and feminine *-ro*, but the morpheme does not appear to carry any meaning today.

Nine classifiers are tentatively included in this group (Table 5.5). They are compatible with an analysis featuring an *-i* suffix, but may also not contain any suffix at all—the vowel fusion rules (§ 3.1.3.4) do not permit a conclusive analysis[58].

Table 5.5: Classifiers tentatively included in Group II

Gloss	Simplex	Complex
CLF.AREA	*-áaphi*	*-áaphi* (*-áaphi-i*?)
CLF.HOLLOW	*-áapi*	*-áapi* (*-áapi-i*?)
CLF.STICK	*-áapo*	*-áapo* (*-áapo-i*?)
CLF.RAPIDS	*-hípani*	*-hípani* (*-hípani-i*?)
CLF.POINTED	*-híwi*	*-híwi* (*-híwi-i*?)
CLF.FLOWER	*-íiwi*	*-íiwi* (*-íiwi-i*?)
CLF.SEED	*-íixi*	*-íixi* (*-íixi-i*?)
CLF.TUBE	*-pi*	*-pi* (*-pi-i*?)
CLF.PILE	*-tsoi*	*-tsoi* (*-tsoi-i*?)

The classifiers that end in *-o* display some variation in their complex forms: some end in *-o* (e.g., *-áapo*), some end in *-i* (e.g., *-híki*), and some show variation between the two (e.g., *-póko*∼*-póki*). Variants ending in *-o* and *-i* are both regular outputs of the vowel fusion rules of Baniwa (§ 3.1.1.2), where *o* + *i* → *i, o*.

[58] Note that all the tentatively included suffixes end in /i/, /o/, or /oi/, as these are the post-fusion vowel qualities that cannot be analyzed unambiguously (cf. Group IV, § 5.3.2.4).

The same cannot be said for *-phe* 'CLF.LEAF', whose complex form is *-phai*. It is included in this group anyway for lack of better options, although we would expect the form *-phe* also on the adjective (*e* + *i* → *e*)[59].

5.3.2.3 Group III: *-ni* '?'

Group III contains the 3 classifiers that take the suffix *-ni* on adjectives (6% of the classifier set). The etymology of this suffix is uncertain (see further in § 5.3.2.5 below).

The oldest classifier in this group is *-wa* 'CLF.SPACE', which can be reconstructed to Proto-Japurá-Colombia (see § 8.1). The two other classifiers are possibly complex forms carrying this suffix as their second part: *-páwa* 'CLF.RIVER' and *-yáwa* 'CLF.HOLE' (Ramirez, 2020a: 164–165). This group may thus be united on morphological grounds[60].

5.3.2.4 Group IV: no suffix

Group IV contains all classifiers whose "complex" form is identical to their simplex form, i.e., they do not take any additional suffix[61]. There are 11 such classifiers (21% of the classifier set). They all end in /a/, /a:/, /e/, /ia/, or /oa/. Most classifiers in Group IV belong to the classifiers that have a synchronically identical corresponding bound noun (see § 8.1).

5.3.2.5 Historical note

The origin of the suffixes *-i*, *-li*, *-ro*, and *-ni* is not entirely clear. In the case of *-li*, *-ro*, and *-i*, the suffixes are identical to the derivational suffixes *-li* 'masculine', *-ro* 'feminine', and *-i* 'neuter', which Ramirez (2020a: 371–375) hypothesizes to be remnants of an old tripartite gender marking system. In the case of *-li* and *-ro*, they still distinguish the gender of the referent when they are combined with classifiers in the complex set of forms. Apart from the few cases where *-li* and *-ro* alternate on the generic classifier depending on the referent, the suffixes do not

[59] This analysis is also tentatively suggested by Ramirez (2020a: 253).

[60] The classifier *-híwa* 'CLF.BEIJU' (Group I) also ends in *-wa*, but this classifier goes back to a Proto-Japurá-Colombia *beiju* classifier **-Siwa* which does not seem to be related to *-wa* 'CLF.SPACE'.

[61] Note that the nine classifiers that are tentatively included in Group II (Table 5.5) might also be analyzed as lacking a suffix.

appear to carry much meaning today: they simply have a nominalizing function, and have each become associated with particular classifiers by convention.

Ramirez (2020a: 251–255) does not recognize *-ni* as part of the group of suffixes that can follow classifiers on adjectives. Aikhenvald (2007: 484) recognizes one classifier with *-ni* (*-páwa* 'CLF.RIVER') and suggests that *-ni* might be related to the homonymous third person non-feminine singular suffix. Another homonymous morpheme is the *-ni* suffix that marks one of the alienable possession classes (§ 3.2.5.1). Due to the uncertain meaning, origin and function of *-ni*, it is glossed as '?'.

Ramirez's hypothesis about the complex forms being remnants of an older system is supported by the historical analysis in § 8.1, as all of the classifiers that can be reconstructed to Proto-Japurá-Colombia and Proto-Arawak take some kind of suffix. The use of the simplex form in all environments appears to be a relatively recent innovation, which likely did not take place until the suffixes used in the complex form had started to lose their meaning.

5.3.3 Classifiers on numerals

The numeral system of Baniwa was described in § 3.2.9.1. This section focuses on how classifiers are used within the Baniwa numeral system. Classifiers occur as obligatory suffixes on numerals 1–3 (§ 5.3.3.1). Historically, the numeral 4 has taken classifiers as obligatory infixes, but this system is falling out of use. Today, the same numeral may take suffixes instead, but only optionally (§ 5.3.3.2). Most commonly, however, numerals from four and up are borrowed from Portuguese. Classifier use on borrowed numerals is not treated here, but described in detail in § 9.2.1.

5.3.3.1 Classifiers on numerals 1–3

Classifiers are obligatory on the numerals *apa-* 'one', *dzama-* 'two', and *madali-* 'three'. The full inflectional paradigm on numerals 1–3 is given in Table 5.6, where irregular forms[62] are marked with asterisks (*). Examples of noun phrases featuring each of the three numerals are shown in (56). In these constructions, the noun follows the numeral and remains unmarked for number. The functions of classifiers on lower numerals is described in more detail in § 6.1.1.1.

[62]Irregular forms will not be discussed further here, but will be referenced in the analysis of classifiers on Portuguese numerals in § 9.2.

Table 5.6: Classifier paradigm on numerals 1–3

Classifier	Gloss	*apa-* 'one'	*dzama-* 'two'	*madali-* 'three'
-Ø	CLF.CANOE	*ápa*	*dzáma*	*madála**
-áana	CLF.GROUP	*apáana*	*dzamáana*	*madalíana*
-áanhaa	CLF.LIQUID	*apáanhaa*	*dzamáanhaa*	*madalíanhaa*
-aápa	CLF.OBLONG	*apaápa*	*dzamaápa*	*madaliápa*
-áaphi	CLF.AREA	*apáaphi*	*dzamáaphi*	*madalíaphi*
-áapi	CLF.HOLLOW	*apáapi*	*dzamáapi*	*madalíapi*
-áapo	CLF.STICK	*apáapo*	*dzamáapo*	*madalíapo*
-da	CLF.GENERIC	*apáda*	*dzamáda*	*madalída*
-daa	CLF.DAY	*apádaa*	*dzamádaa*	*madalídaa*
-dápana	CLF.HOUSE	*apadápana*	*dzamadápana*	*madalidápana*
-eéma	CLF.SIDE	*apeéma*	*dzameéma*	*madaliéma*
-éekhe	CLF.SMALL.SEED	*apéekhe*	*dzaméekhe*	*madalíekhe*
-éewhe	CLF.EGG	*apéewhe*	*dzaméewhe*	*madalíewhe*
-híko	CLF.LONG	*aphéko*	*dzamhéko*	*madalhíko*
-hípa	CLF.MALE	*aphépa*	*dzamhépa*	*madalhípa*
-hipáda	CLF.PIECE	*aphepáda*	*dzamhepáda*	*madalhipáda*
-hípani	CLF.RAPIDS	*aphépani*	*dzamhépani*	*madalhípani*
-híwa	CLF.BEIJU	*aphéwa*	*dzamhéwa*	*madalhíwa*
-híwi	CLF.POINTED	*aphéwi*	*dzamhéwi*	*madalhíwi*
-i	CLF.BASKET	*ápe*	*dzáme*	*madalíye**
-íida	CLF.HALF	*apáida*	*dzamáida*	*madalíeda**
-iíta	CLF.HUMAN	*apaíta*	*dzamaíta*	*madaliíta*
-iítsia	CLF.BUNCH	*apeétsia*	*dzameétsia*	*madaliítsia*
-íiwi	CLF.FLOWER	*apéewi*	*dzaméewi*	*madalíiwi*
-íixi	CLF.SEED	*apéexi*	*dzaméexi*	*madalíixi*
-kénaa	CLF.BRANCH	*apakénaa*	*dzamakénaa*	*madalikénaa*
-khaa	CLF.CURVILINEAR	*apákhaa*	*dzamákhaa*	*madalíkhaa*
-ko	CLF.HAMMOCK	*apáko∼aapáko**	*dzamáko*	*madalíko*
-koa	CLF.SURFACE	*apákoa*	*dzamákoa*	*madalíkoa*
-kódzoa	CLF.BEND	*apakódzoa*	*dzamakódzoa*	*madalikódzoa*
-ma[1]	CLF.PAIR	*apáma*	*dzamáma*	*madalíma*
-ma[2]	CLF.FEMALE	*aapáma**	*dzamáma*	*madalíma*
-máka	CLF.FABRIC	*apamáka*	*dzamamáka*	*madalimáka*
-na	CLF.TRUNK	*aapána**	*dzamána*	*madalína*
-náko	CLF.BUNDLE	*apanáko*	*dzamanáko*	*madalináko*
-pa	CLF.PACKAGE	*apápa*	*dzamápa*	*madalípa*
-páwa	CLF.RIVER	*apapáwa*	*dzamapáwa*	*madalipáwa*
-péko	CLF.PATH	*apapéko*	*dzamapéko*	*madalipéko*
-phe	CLF.LEAF	*apáphe*	*dzamáphe*	*madalíphe*
-pi	CLF.TUBE	*aapápi**	*dzamápi*	*madalípi*
-póko	CLF.CIRCLE	*apapóko*	*dzamapóko*	*madalipóko*
-pokóda	CLF.STUMP	*apapokóda*	*dzamapokóda*	*madalipokóda*
-tsoi	CLF.PILE	*apátsoi*	*dzamátsoi*	*madalítsoi*
-ttáwalhe	CLF.CUT	*apattáwalhe*	*dzamattáwalhe*	*madalittáwalhe*
-ttówhia	CLF.ROOM	*apattówhia*	*dzamattówhia*	*madalittówhia*
-wa	CLF.SPACE	*apáwa*	*dzamáwa*	*madaláwa*∼madaliyáwa**
-wálhia	CLF.YEAR	*apawálhia*	*dzamawálhia*	*madaliwálhia*
-wána	CLF.SLICE	*apawána*	*dzamawána*	*madaliwána*
-wáta	CLF.BRACT	*apawáta*	*dzamawáta*	*madaliwáta*
-wáthe	CLF.NODE	*apawáthe*	*dzamawáthe*	*madaliwáthe*
-xaa	CLF.EXCREMENT	*apáxaa*	*dzamáxaa*	*madalíxaa*
-ya	CLF.SKIN	*apáya*	*dzamáya*	*madalíya*
-yáwa	CLF.HOLE	*apayáwa*	*dzamayáwa*	*madaliyáwa*

(56) a. *apáda* *ienipétti*
apa-**da** ienipétti
one-**CLF.GNR** child

‘one child’ [190405_BOY_E_01]

b. *dzamáda* *kóphe*
dzama-**da** kóphe
two-**CLF.GNR** fish

‘two fish’ [200311_ELIC_K_06]

c. *madála* *íitta*
madali-**Ø** íitta
three-**CLF.CANOE** canoe

‘three canoes’ [220918_ELIC_F_01]

5.3.3.2 Classifiers on the numeral 4

As described in § 3.2.9.1, *likoa-…-áaka* ‘four’ is a complex expression where the empty slot in the middle has traditionally taken an obligatory classifier infix (Aikhenvald, 1996a: 99; Ramirez, 2020a: 248). The fact that the classifier shows up as an infix in this context (but as a suffix in all other contexts) is due to the classifier historically being a suffix followed by another suffix, namely the reflexive *-áaka*, which has become an invariable part of the conventionalized expression. Three examples of *likoa-…-áaka* with classifier infixes are shown in (57).

(57) Infixing pattern on 4 (adapted from Ramirez, 2020a: 248)

-áapa	‘CLF.OBLONG’	*likoapáaka*	*li-koa<**áapa**>áaka*
-da	‘CLF.GENERIC’	*likoadáaka*	*li-koa<**da**>áaka*
-híwa	‘CLF.BEIJU’	*likhoewáaka*	*li-koa<**híwa**>áaka*

At least for some speakers, *likoa-…-áaka* appears to have undergone a fossilization process with the infixed generic classifier *-da*, resulting in the non-inflecting form *likoadáaka* being used with all referents.

In a further developmental step, the fossilized form *likoadáaka* has undergone analogical levelling with the rest of the lower numerals (1–3, see § 5.3.3.1), and can now (optionally) take the full classifier paradigm as suffixes. Three examples—

parallel to those above—of the fossilized *likoadáaka* with classifier suffixes are shown in (58).

(58) Suffixing pattern on 4 [220918_ELIC_F_01]

-áapa	'CLF.OBLONG'	*likoadaakáapa*	*likoadaaka-**áapa***
-da	'CLF.GENERIC'	*likoadaakáda*	*likoadaaka-**da***
-híwa	'CLF.BEIJU'	*likoadaakhéwa*	*likoadaaka-**hiwa***

The fact that classifiers may now optionally be attached as suffixes to *likoadáaka* suggests that the form is not analyzable to speakers anymore, and that the infixed <*da*> probably is not recognized as a classifier. The fact that many of the classifier distinctions collapse due to morphophonological processes like vowel fusion and /h/-metathesis may also have contributed to the functional demise of the infixing system. For example, both <*áapi*> 'CLF.HOLLOW' and <*pi*> 'CLF.TUBE' result in the form *likoapíaka* by completely regular processes of vowel fusion (see § 3.1.3.4).

The traditional infixing system appears to be in exceedingly rare use today. Many speakers do not recognize the forms when prompted with them. As for the suffixing system, it is unclear to what extent it is used. The suffixing system has not been described in any of the previous sources. The most common strategy seems to be the use of the Portuguese numeral *quatro* 'four', i.e., avoiding the Baniwa numeral altogether. Classifier use on Portuguese numerals is described in § 9.2.1.

5.3.4 Classifiers on adjectives

Classifiers are obligatorily marked on attributive adjectives, i.e., adjectives that modify heads of noun phrases. An example is given in (59). The complex form (see § 5.3.2) of the classifier is used on adjectives, as a result of the fact that adjectives need to be nominalized in order to be used attributively. This and other functions of classifiers on adjectives are described in more detail in § 6.1.1.2.

(59) *pántti makadápana*
pántti maka-**dápana**
house big-**CLF.HOUSE**

'big house' [200305_ELIC_F_01]

As mentioned in § 3.2.4.2, adjectives typically do not take classifier marking when they are used as predicates, but there is a handful of bound adjectives like *maka-* 'big' that obligatorily take a classifier in all contexts (Ramirez, 2020a: 267). An example is given in (60).

(60) *makadápanaka* *pántti*
maka-**dápana**-ka pántti
big-**CLF.HOUSE**-SUB house
'The house is big' (adapted from Ramirez, 2020a: 267)

5.3.5 Classifiers on nominal roots

Classifiers can be suffixed to nouns—both free (61) and bound (62) ones—to modify their meaning or to derive new nouns (see further in § 6.1.1.6).

(61) *tsipála* 'metal (material)'

a. *tsipaláapi*
tsipála-**áapi**
metal-**CLF.HOLLOW**
'pan' [200315_ELIC_F_01]

b. *tsipaláda*
tsipála-**da**
metal-**CLF.GNR**
'firearm bullet' (or other round metal object) [200322_ELIC_F_01]

(62) *-nóma* 'mouth'

a. *nonóma*
no-nóma
1SG-mouth
'my mouth' [200310_ELIC_K_02]

b. *nonomáya*
no-nóma-**ya**
1SG-mouth-**CLF.SKIN**
'my lip' (lit. 'my mouth-skin') [200310_ELIC_K_02]

The classifier serves to assign properties to the referent, typically related to its shape, as in (61a) where the word for 'pan' is derived with *-áapi*, denoting hollow objects, and in (61b) where the word for 'firearm bullet' is derived with *-da*, here denoting something small and roundish. Classifiers can also be used for specifying a part–whole relation, as in (62b), where the classifier *-ya* highlights the 'skin' part of the mouth when referring to the lip.

Over time, as these processes form lexicalized expressions, classifiers regularly become incorporated into new nominal roots. In some cases, such roots can be further specified with another classifier. Two examples of such diachronic classifier stacking are given in (63b) and (64b). The form *-thimakáya* 'eyelids' (63b) appears to be possible only because *-thimáka* (63a)—itself a derivation attested in Ramirez (2001a: 286–287)—had already been lexicalized. Likewise, *-hiwidáapi* 'skull' (64b) is derived from *-hiwída* 'head' (64a), which in turn was created via derivation with a classifier. There is no evidence to suggest that Baniwa allows synchronic classifier stacking.

(63) a. *nothimáka*
no-thi-**maka**
1SG-eye-**CLF.FABRIC**

'my eyelids' (adapted from Ramirez, 2001a: 286–287)

b. *apeéma* *lithimakáya*
apa-eema li-thimáka-**ya**
one-CLF.SIDE 3NFSG-eyelids-**CLF.SKIN**

'one of his eyelids' [200310_ELIC_K_02]

(64) a. *rhowída*
ro-hiwi-**da**
3NFSG-hook/arrow-**CLF.GNR**

'her head' [200310_ELIC_K_01]

b. *lhiwidáapi*
li-hiwída-**áapi**
3NFSG-head-**CLF.HOLLOW**

'his skull' [200310_ELIC_K_01]

5.3.5.1 Classifiers in alienable possession constructions

Baniwa has an alienable possession construction formed with the bound noun *-dzaa* 'belongings'. The construction is POSS-*dzaa-ro*-CLF, where the obligatory person prefix represents the possessor. The construction also obligatorily contains the feminine nominalizer *-ro*, as well as an obligatory classifier suffix in the simplex form (65a–65b). The generic classifier *-da* forms an irregular variant of this construction by omitting *-ro*: POSS-*dzaa-da* (65c).

(65) a. *lhíehẽ* *ttirolípi* *nodzaarópi*
lhíehẽ ttirolípi no-dzaa-**ro-pi**
DEM1.NFSG *tipití* 1SG-belongings-**F-CLF.TUBE**

'this *tipití* of mine' [220918_ELIC_F_02]

b. *lhíehẽ* *piéta* *nodzaaróko*
lhíehẽ piéta no-dzaa-**ro-ko**
DEM1.NFSG hammock 1SG-belongings-**F-CLF.HAMMOCK**

'this hammock of mine' [220918_ELIC_F_02]

c. *lhíehẽ* *dáapa* *nodzaáda*
lhíehẽ dáapa no-dzaa-**da**
DEM1.NFSG paca 1SG-belongings-**CLF.GNR**

'this paca of mine' [220918_ELIC_F_02]

Gender marking is typically semantically transparent in Baniwa (see § 3.2.5.4). The alienable possession construction is, to my knowledge, the only context where gender marking does not alternate between feminine *-ro* and non-feminine *-li*: here, the feminine marker is used irrespective of the referent.

5.3.6 Classifiers on verbal roots

Classifiers in Baniwa are primarily marked within the noun phrase. Classifiers do not appear on verbs when they are used as predicates, but they are used in certain nominalization constructions that permit verbal roots to occur in the noun phrase, either as nouns or as modifiers in the form of relative clauses (see further in § 6.1.1.3).

Nominalization of verbs is common in Baniwa, and there are several different constructions with different nominalizers, depending on semantic role (§ 3.2.8).

Three of the nominalization constructions take classifiers: the ones derived with the agent nominalizer *-ka*, the patient nominalizer *-ni*, and the location nominalizer *-karo*[63] (Table 5.7). They all take the simplex form of the classifier. Nominalization of verbs results in bound nouns, which obligatorily take a person prefix (§ 3.2.3).

Table 5.7: Nominalization constructions with classifiers

Nominalizer	Function	Construction	Result
-ka	agent nominalizer	V-*ka*-CLF	→ bound noun
-ni	patient nominalizer	V-*ni*-CLF	→ bound noun
-karo	location nominalizer	V-*karo*-CLF	→ bound noun

(66) shows three different nominalizations of the verb *-kadzeekáta* 'to teach', derived with the agent nominalizer *-ka* (66a, 66b) and the place nominalizer *-karo* (66c). These expressions, which all function as nouns, have become lexicalized with either connective *i-* or impersonal *pa-*.

(66) *-kadzeekáta* 'to teach' [200327_ELIC_F_01]

a. *ikadzeekatakaíta*
i-kadzeekáta-**ka-iíta**
CONN-to.teach-**NMLZ.AG-CLF.HUMAN**

'teacher' (lit. '(the one) who teaches')

b. *ikadzeekataakakaíta*
i-kadzeekáta-aaka-**ka-iíta**
CONN-to.teach-REFL-**NMLZ.AG-CLF.HUMAN**

'student' (lit. '(the one) who teaches him-/herself')

c. *pakadzeekataakakaróda*
pa-kadzeekáta-aaka-**karo-da**
IMPRS-to.teach-REFL-**NMLZ.LOC-CLF.GNR**

'school' (lit. '(the place) where one teaches oneself')

Nominalized verbs can also function as relative clauses. This is exemplified in (67a) with the agent nominalizer *-ka*, which refers to a subject, and in (67b) with

[63] For the instrument nominalizers *-xoóda* and *-xoópa*, Ramirez (2020a: 291) suggests that these may contain fossilized versions of the two classifiers *-da* 'generic' and *-aápa* 'oblong', respectively. This may explain why they do not take classifiers synchronically.

the patient nominalizer *-ni*, which refers to an object. The functions of classifiers on verbal roots are described in more detail in § 6.1.1.3.

(67) a. *áawi* *hiwakhéwi*
áawi i-hiwa-**ka-híwi**
needle CONN-to.fall-**NMLZ.AG-CLF.POINTED**

'(the) needle which is falling' [220929_ELIC_G_01]

b. *áawi* *pikapanhíwi*
áawi pi-kapa-**ni-híwi**
needle 2SG-to.see-**NMLZ.PAT-CLF.POINTED**

'(the) needle which you saw' [220930_ELIC_K_01]

5.3.6.1 Agent nominalization with *-ka* + CLF

The agent nominalizer *-ka* constructs nominalizations referring to the agent of the action expressed by the verb. Agents are typically human, and the most common classifiers in this construction are *-iíta* 'CLF.HUMAN' (68a), *-da* 'CLF.GENERIC', and *-ma*^{2} 'CLF.FEMALE' (68b)—three classifiers typically used with human referents (§ 7.2.1). However, the generic classifier *-da* can also be used in this construction with non-human agents (68c). The classifier in nominalizations with *-ka* can also be exchanged for the plural suffix *-pe*, as shown in (68d), where *iarakápe* refers to a collective of birds or flying animals.

(68) a. *ideenhikaíta*
i-déenhi-ka-**iíta**
CONN-to.work-NMLZ.AG-**CLF.HUMAN**

'worker' (someone (human) who works) [200322_ELIC_F_01]

b. *iitsaletakáma*
i-iitsaléta-ka-**ma**
CONN-to.fish-NMLZ.AG-**CLF.FEMALE**

'fisherwoman' (someone (female) who fishes) [200327_ELIC_F_04]

c. *iphiakáda* *káwaale*
i-phia-ka-**da** káwaale
CONN-to.blow-NMLZ.AG-**CLF.GNR** wind

'ventilator' (something that blows wind) [200322_ELIC_F_01]

d. *iarakápe*
i-áara-ka-**pe**
CONN-to.fly-NMLZ.AG-**PL**

'birds, flying animals' (something (plural/collective) that flies)
[200311_ELIC_K_06]

(68c) is interesting in that the transitive verb *-phia* 'to blow' takes the overt object *káwaale* 'wind', which is incorporated into the nominalization in a way that allows the creation of a kind of compound (by juxtaposition). Normally in Baniwa compounds, the second part consists of a bound noun, but free nouns are also allowed in this position in cases where the first part in the compound is a nominalized transitive verb, of which the free noun in the second part is the direct object. This construction (as well as compounding in general) is described in § 3.2.5.2.

5.3.6.2 Patient nominalization with *-ni* + CLF

The patient nominalizer *-ni* creates nominalizations referring to the undergoer of the action expressed by a transitive verb. This role is compatible with both human and non-human referents, and thus, as expected, the classifiers used in this nominalization construction belong both to those used with human referents (69a), and to those used with non-human referents (69b). As with the agent nominalizer *-ka*, the classifier in the nominalization construction formed with *-ni* can be exchanged for a plural suffix for a collective reference. (69c) is an example of this, as *beiju* dough is what results from manioc that has been grated and strained.

(69) a. *painoaniíta*
pa-íinoa-ni-**iíta**
IMPRS-to.kill-NMLZ.PAT-**CLF.HUMAN**

'victim' (someone (human) who is killed) [200328_ELIC_F_01]

b. *paphianída*
pa-phia-ni-**da**
IMPRS-to.blow-NMLZ.PAT-**CLF.GNR**

'flute' (something that is blown) [200327_ELIC_F_02]

c. *rowidzonípe*
ro-wídzo-ni-**pe**
3FSG-to.strain-NMLZ.PAT-**PL**

'her *beiju* dough' (something (plural/collective) that is strained)
[200315_ELIC_F_01]

5.3.6.3 Location nominalization *-karo* + CLF

The location nominalizer *-karo* creates nominalizations referring to a location where the action denoted by the verb takes place. The location in question can vary in size, from an area associated with an action (70a), to a container associated with an action (70b, 70c). The nominalizer commonly occurs with the generic classifier *-da*, but also with classifiers referring to spatial entities, such as *-áaphi* 'CLF.AREA'. (70a) is another example of incorporation of the object, in the form of the borrowed noun [*bola*] 'soccer' (cf. (68c)).

(70) a. *natopikakaróaphi* [*bola*]
na-topika-karo-**áaphi** [bola]
1PL-to.play-NMLZ.LOC-**CLF.AREA** ball/soccer

'our soccer field' (somewhere where playing of soccer takes place)
[200316_ELIC_F_01]

b. *padzeeneetakaróda*
pa-dzeeteéta-karo-**da**
IMPRS-to.make.pass-NMLZ.LOC-**CLF.GNR**

'sieve, strainer' (somewhere where straining takes place)
[200315_ELIC_F_01]

c. *pairakarodápe*
pa-íira-karo-**da-pe**
IMPRS-to.drink-NMLZ.LOC-**CLF.GNR-PL**

'vessels for drinking' (somewhere where drinking takes place)
[200315_ELIC_F_01]

An interesting feature that sets *-karo* apart from the agent and patient nominalizers *-ka* and *-ni* is that it does not seem to be possible for the plural suffix *-pe* to fill the classifier slot in this environment. The plural suffix can be added to the construction, but following the classifier, as in (70c).

5.3.7 Classifiers on demonstratives

Demonstratives in Baniwa usually do not take classifiers (see § 3.2.10). An example of a demonstrative without classifier is shown in (71), where the non-feminine singular proximal demonstrative *lhíehẽ* forms a noun phrase with the head noun *híiparo* 'frog sp.'.

(71)

nheekokháapani	*nóada*	*líikaalhe*	***lhíehẽ***	*híiparo*
na-heéko-kháapani	no-aawáada	li-íikaa-lhe	**lhíe(hẽ)**	híiparo
3PL-to.run-IMM	1SG-to.think	3NFSG-above-ALL	**DEM1.NFSG**	frog

'They ran, I think, towards **this** frog.' [190405_BOY_E_01]

However, a classifier may be suffixed to a demonstrative in order to express focus or contrast. The simplex form of the classifier then attaches to the pronominal root that forms the basis for the demonstrative system (see § 3.2.10). In (72a) and (72b), the classifiers *-aápa* 'CLF.OBLONG' and *-ko* 'CLF.HAMMOCK' are attached to the same demonstrative/pronominal root *lhía*, resulting in a contrastive reading—i.e., that a certain referent is singled out among several alternative ones. In both of these elicited examples, a noun is overtly expressed. The noun in this construction obligatorily takes the subordinator *-ka*, which turns it into the predicate of an existential clause (although best translated as a relative clause, cf. Ramirez, 2020a: 208): *palanáka* 'which is a banana' (72a), *piétaka* 'which is a hammock' (72b). The noun is effectively pushed out of the noun phrase, which is headed by the demonstrative. This is in contrast to (71), where *híiparo* stands as the head of the noun phrase and therefore does not take the subordinator *-ka*. (§ 6.1.1.4) describes the functions of classifiers on demonstratives in more detail.

(72) a.

lhiápa	*palanáka*
lhia-**aápa**	palana-ka
DEM1.NFSG-**CLF.OBLONG**	banana-SUB

'this banana' (pointing out one among several alternatives)
(lit. 'this (oblong) one, which is a banana') [220919_ELIC_F_01]

b.

lhíako	*piétaka*
lhia-**ko**	pieta-ka
DEM1.NFSG-**CLF.HAMMOCK**	hammock-SUB

'this hammock' (pointing out one among several alternatives)
(lit. 'this (hammock) one, which is a hammock') [220919_ELIC_F_01]

5.3.8 Classifiers on interrogatives

A general introduction to interrogative clauses in Baniwa was given in § 3.2.11. Wh-questions in Baniwa are formed with interrogative pronouns, most of which are based on the interrogative root *kóa* 'what?'. The interrogative constructions that take classifiers are listed in Table 5.8. They involve the bases *kóame* 'how?', *kóaka* 'what/who?', *kóadzo* 'how many?', and *kadali-* 'how many?', which are described in turn in § 5.3.8.1–5.3.8.4 below.

Table 5.8: Interrogative expressions with classifiers

Base	Meaning	Construction	Meaning	Clf form
kadali-	'how many?'	*kadali-*CLF^{s}	'how many?'	simplex
kóaka	'what/who?'	*kóaka-*CLF^{s}	'what kind of?'	simplex
kóadzo	'how many?'	*kóadzo-*(CLF^{s}) *kóadzo-*CLF^{c}	'how many?' 'what size?'	simplex complex
kóame	'how?'	*kóame-*CLF^{c}	'which?'	complex

The first three constructions in Table 5.8 consistently take the simplex classifier form, and the last two take the complex one. The choice of form depends on whether the classifier has an inflectional (simplex), qualificational (simplex), or derivational (complex) function in the expression; in other words, what relationship it has to the interrogative base. This is discussed in more detail in § 6.1.1.5.

5.3.8.1 *Kadali-* 'how many?'

The interrogative morpheme *kadali-* 'how many?' obligatorily takes a classifier (73a). The typical expected answer to a question with *kadali-* consists of a classifier-inflected numeral (73b), often with the restrictive suffix *-tsa*.

(73) a. *kadaliíta* *kóphe?*
kadali-**iíta** kóphe
how.many-**CLF.HUMAN** fish

'How many fish?' [220927_ELIC_G_01]

b. *apaítatsa*
apa-**iíta**-tsa
one-**CLF.HUMAN**-RESTR

'Just one (fish).' [220927_ELIC_G_01]

5.3.8.2 *Kóaka* 'what/who?'

The interrogative *kóaka* is typically used without a classifier to express the meaning 'what?'[64], as in (74). *Kóaka* consists of the interrogative morpheme *kóa* and the suffix *-ka* (§ 3.2.4), which makes it predicative (Ramirez, 2020a: 332–333).

(74) *heee* *kóaka* *lhíerapida* *liakoána*
heee **kóa-ka** lhíera(hã)-pida li-áako=aána(hã)
[screaming] **Q-SUB** DEM2.NFSG-RPT 3NFSG-to.say=DEM3.LOC
'*Heee! What (is) that?*, he said (reported)' [220927_ELIC_G_01]

Kóaka can also take a classifier suffix to form the construction *kóaka*-CLF 'what kind of?' (75). The simplex form of the classifier is used in this construction.

(75) *koakáda* *haikóthe* *píiñhali?*
kóa-ka-**da** haikóthe pi-íiñha-li
Q-SUB-**CLF.GNR** fruit 2SG-to.eat-REL
'What kind of fruit (is it) that you ate?' [220928_ELIC_G_01]

5.3.8.3 *Kóadzo* 'how many?'

The interrogative morpheme *kóadzo*, with the basic meaning 'how many?' (see § 3.2.11), is used in two separate interrogative constructions with classifiers.

The construction *kóadzo*-(CLF^{s}) 'how many?' may take a classifier in the simplex form (76a). In this construction, the classifier is optional—it can also be left out (76b). Some speakers alternate between (76a) and (76b), whereas for others, the use of a classifier on *kóadzo* is simply ungrammatical. These speakers either use the bare form *kóadzo?* 'how many?' for quantity-questioning, or the synonymous construction *kadali*-CLF 'how many?' (§ 5.3.8.1), which allows them to specify the referent by means of a classifier. The choice likely depends on the level of specificity required in the specific situation. A question formed with *kóadzo* expects the same kind of answer as described above for *kadali-* (73b).

[64] *Kóaka* can also mean 'who?', see § 3.2.11.

(76) a. *koadzóphe?*
kóadzo-**phe**
how.many-**CLF.LEAF**

'How many?' (of leaves) [220927_ELIC_G_01]

b. *kóadzo?*
kóadzo
how.many

'How many?' [220930_ELIC_K_01]

By contrast, a classifier is needed in the construction *kóadzo*-CLF^{c} 'what size?' (77a). In this construction, the complex form of the classifier is used. A question like the one in (77a) would typically be answered with a construction with *kádzo* 'like this' (77b). *Kádzo* is usually used without a classifier (77c), but to form the construction *kadzo*-CLF^{c}-*aahã* 'of this size', a classifier is obligatory (77b).

(77) a. *koadzóphai?*
kóadzo-**phe-i**
how.many-**CLF.LEAF-N**

'What size?' (of a leaf) [220927_ELIC_G_01]

b. *kadzókhaiyaahã*
kádzo-**khaa-i**=aahã
like.this-**CLF.CURV-N**=DEM1.LOC

'of this size (of thread, etc.)' [220930_ELIC_K_01]

c. *kádzo nokaitekápidzo pírhio!*
kádzo no-kaíte-kápidzo pi-lhio
like.this 1SG-to.tell-COMPR 2SG-BEN

'Like I told you!' [200308_NARR_B_01]

5.3.8.4 *Kóame* 'how?'

Classifiers are used on the interrogative morpheme *kóame*, which has the basic meaning 'how?' (§ 3.2.11), to form the interrogative construction *kóame*-CLF 'which?', that is, to form a question serving to point out a referent from a limited set of possible ones (78a). In response to such a question, the root *kóame* can

also be employed in the indefinite construction *kóame*-CLF-*katsa* 'any' (78b). The noun (in this case *haikóthe* 'fruit') is often omitted if the referent is clear from the context.

(78) a. *koamédali* *haikóthe* *píiñhali?*
koame-**da-li** haikóthe pi-iiñha-li
how-**CLF.GNR-NF** fruit 2SG-to.eat-REL
'Which fruit did you eat?' [220927_ELIC_G_01]

b. *koamédalikatsa* *(haikóthe)*
koame-**da-li**-ka-tsa (haikóthe)
how-**CLF.GNR-NF**-SUB-RESTR (fruit)
'any (fruit)' [220928_ELIC_G_01]

5.4 Chapter summary

This chapter has outlined the formal characteristics of the classifier system.

In terms of phonology, classifier morphemes consist of up to three syllables, and most commonly of only metric syllables (with a few exceptions). They abide by the general morphophonological rules of Baniwa, meaning that vowel-initial classifiers trigger vowel fusion and /h/-initial classifiers trigger /h/-metathesis. When suffixed to *apa-* 'one', some classifiers trigger irregular lengthening of the initial vowel of the host word.

Morphosyntactically, classifiers are distributed across several loci within the noun phrase. They can occur on heads (that is, on nouns), as well as on modifiers: low numerals (1–4), attribute adjectives, a possessive construction formed with the bound noun *-dzaa*, relative clauses formed with *-ka* and *-ni*, demonstratives when used contrastively, as well as a number of interrogative constructions. In many of these loci, they are obligatory (for example, a low numeral or an attributive adjective can never occur without a classifier), and in others they are optional (e.g., on demonstratives).

Most classifiers have two forms: a simplex and a complex one. The complex form contains an additional nominalizer, either *-li*, *-ro*, *-i*, or *-ni*. It is the morphosyntactic context that determines which of the two forms is used. A few classifiers use the same form in all contexts. Classifiers always appear as suffixes; the only previously infixing pattern (on the numeral 4) has been regularized into an optional suffixing pattern.

Chapter 6

Functions

This chapter explores the different functions that classifiers exhibit in Baniwa. In § 6.1, three functions that apply on word level are analyzed together: derivation, inflection, and qualification. § 6.2 is concerned with the function of referent specification, and § 6.3 deals with the use of classifiers in discourse.

The remaining sections discuss aspects of the functions that are of particular significance to the analysis. § 6.4 considers the distinction between inflection and derivation. § 6.5 shows that classifiers are not agreement markers in Baniwa. Finally, § 6.6 takes a second look at the structure of the noun phrase, in light of the functions that have been presented earlier in the chapter.

6.1 Word-level functions: derivation, inflection, and qualification

This section is concerned with the relationship between classifiers and the hosts that they attach to. It treats the three interrelated functions of derivation, inflection, and qualification, the distinguishing properties of which are summarized in Table 6.1.

Table 6.1: Properties of derivation, inflection, and qualification

	Lexeme-external	Obligatory
Derivation	✓	–
Inflection	–	✓
Qualification	–	–

Derivation is a lexeme-external process (i.e., it creates new lexemes), and in this sense it is distinct from both inflection and qualification. Derivation can be evidenced by change of word class (e.g., nominalization of verbs), which grants the derived word access to new morphosyntactic contexts. However, derivation does not always entail word class change: for example, nouns can be derived from other nouns. In such cases, a semantic change is necessary, such that the derived word "denotes an entity (or type) different from that denoted by the nominal root" (Rose & Van linden, 2023: 262).

Inflection is a lexeme-internal process, that is, it does not derive new lexemes. Obligatoriness is often taken to be one of the most important criteria for inflection (see Bybee, 1985: 81–82), and here it will be used to distinguish inflection from qualification. As inflectional morphemes, classifiers can be said to exhibit a kind of agreement with the head noun (or rather, its referent). An important point is that the obligatoriness of the classifier in an inflectional position does not entail any particular restriction on the inflectional *value*—that is, a classifier is obligatory, but the classifier choice is quite free (see § 4.4).

Qualification (also called property assignment; Dunn & Rose, forthcoming: 26) is also a lexeme-internal process, but is always optional, setting it apart from inflection. I follow Rose & Van linden (2023: 262) in distinguishing qualification and derivation on a semantic basis: qualification results in an instance of the type denoted by the root. Note that qualification is used here for a process that modifies the meaning of the *host* that it attaches to (hence its inclusion under word-level functions), not of the *head noun*, in case the classifier is marked on a modifier. However, in the case of classifier marking on nouns themselves, this is of course indistinguishable.

6.1.1 Distribution by morphosyntactic host

Table 6.2 shows the distribution of the three functions across the main loci that classifiers occur in. In what follows, the functions are outlined and discussed.

Table 6.2: Distribution of word-level functions by morphosyntactic host

Locus	Derivation	Inflection	Qualification
Numeral	✓	✓	
Adjective	✓		
Verb	✓		
Demonstrative	✓		
Interrogative	✓	✓	✓
Noun	✓		✓

6.1.1.1 On numerals 1–3

The clearest cases of inflection are when classifiers are used on lower numerals (1–3) that modify the head noun (underlined) in a noun phrase (NP) In (79a), the trunk classifier *-na* is suffixed to the numeral *apa-* 'one', which modifies the free noun *háiko* 'tree'. In (79b), the generic classifier *-da* is suffixed to the numeral *apa-* 'one', which modifies the free noun *yamakátti* 'cloth'. Here, the whole NP *apáda yamakátti* 'one/a cloth' functions as the object of the preceding predicate.

(79) a.

aapána	*háiko*
apa-**na**	háiko
one-**CLF.TRUNK**	tree

'one tree' [200228_ELIC_F_01]

b.

pakádaa	*apáda*	*yamakátti*	*mátsia*	*náani*	*dopítsi*
pa-kádaa	apá-**da**	yamakátti	mátsia	náani	dopítsi
IMPRS-to.leave	one-**CLF.GNR**	cloth	good	SUSP	sieve

iikáanaami
i-iikáanaami
CONN-on.top.of

'One leaves a cloth well on top of a sieve.' [200308_NARR_B_01]

Classifiers are obligatory on lower numerals; a lower numeral can never appear without a classifier[65]. This contrasts them with verbs and adjectives, whose primary function is predicative (see § 3.2.4). For verbs and adjectives to function as modifiers of noun phrase heads, they must be derived (nominalized) with the appropriate morphology, which in both cases includes an obligatory classifier, but also some additional suffix. Numerals always appear inside the noun phrase, always with a classifier but never with a nominalizing suffix.

Numerals can also function as heads of noun phrases in the absence of nouns, where the classifier can be seen as a derivational suffix deriving an independent noun phrase from a numeral, as required for a noun phrase head. The marking (i.e., the simplex form of the classifier) is the same in both the inflectional (80a) and the derivational (80b) function.

[65]The only exception is when numerals appear with the morpheme *-wali* 'times', which was discussed in § 4.3.3.

(80) a. *apéewhe* *ainíewhe*
apa-**éewhe** aini-éewhe
one-**CLF.EGG** wasp-egg

'one wasp's egg' [200321_FREE_A_01]

b. *apéewhe*
apa-**éewhe**
one-**CLF.EGG**

'one (egg)' [200309_FREE_K_02]

Sometimes a derived numeral takes additional morphology that confirms that it functions as the NP head, such as the relational suffix *-nako* in (81).

(81) … *dzamádanako* *pándza* *pakéñoa* *paároka*
… dzama-da-**nako** pándza pa-kéñoa pa-aáro-ka
… two-CLF.GNR-**LOC.on** now IMPRS-to.begin IMPRS-to.pour-SUB
líawa *lhiéna* *óoni*
li-aa-wa lhiéna(hã) óoni
3NFSG-to.go-MID DEM3.NFSG water

'… in two (hours), one starts to pour out that water.' [200308_NARR_B_01]

The likelihood of a classifier to be used in the inflectional vs. the derivational function on numerals (that is, of occurring with or without a head noun) depends to some extent on its specificity. Table 6.3 shows some examples of classifiers that typically appear without an overt noun. They are all highly specific and thus restricted in their range of possible referents, which seems to explain the tendency.

Table 6.3: Examples of classifiers typically used without overt nouns

Classifier	Gloss	Example phrase	Meaning
-wálhia	CLF.YEAR	*apawálhia*	'one year'
-hípani	CLF.RAPIDS	*aphépani*	'one rapid'
-ttówhia	CLF.ROOM	*apattówhia*	'one room'
-kódzoa	CLF.BEND	*apakódzoa*	'one bend'

To conclude, classifiers on numerals are either inflectional (if the numeral modifies the NP head) or derivational (allowing the numeral to head the NP). Numerals often occur as derived NPs for anaphoric reference in discourse (see § 6.3).

6.1.1.2 On adjectives

The primary function of adjectives in Baniwa is predicative (82a), but they can also be used attributively, modifying the head in a noun phrase (82b). Attributive use requires a classifier suffix in the complex form, which consists of a classifier and (in most cases) an additional nominalizing suffix, the choice of which is determined by the classifier (see further in § 5.3.2). Thus, adjectives in Baniwa are nominalized in order to function as modifiers in the noun phrase. They typically follow the head noun.

(82) a. *hanipáka óoni*
hanípa-ka óoni
huge-SUB water/river

'The river is flooding (lit. 'being huge')' [200325_ELIC_F_02]

b. *awakáda hanipádali*
awakáda hanípa-**da-li**
forest huge-**CLF.GNR-NF**

'huge forest' [200325_ELIC_F_02]

The same derivational construction that derives attributive adjectives can also derive independent NPs from adjectives. For example, the noun *poottidzádali* 'candy' (83) is derived from the adjective *poottídza* 'sweet' (cf. the English noun *sweets*). This supports the analysis of attributive adjectives as nominalized expressions.

(83) *poottidzádali*
poottidza-**da-li**
sweet-**CLF.GNR-NF**

'candy' [200326_ELIC_F_03]

Similarly, the noun *kaxadádali* denotes a certain jug-shaped basket made by the Baniwa (known as *jarro* in Portuguese; see Ricardo, 2001: 12–13). It is derived from the adjective *kaxáda* which means 'big-bellied' (which can also be used about people with large stomachs), and refers to the appearance of the basket. (84a) shows that it can head a NP where it is modified by a numeral, in a structure analogous to that in (84b), where the inherent free noun *tshéeto* '*aturá* basket' is the head.

(84) a. *aapána* *kaxadádali*
apa-na kaxáda-da-li
one-CLF.TRUNK big.bellied-CLF.GNR-NF

'one *jarro* basket' (lit. 'one big-bellied one') [200322_ELIC_F_01]

b. *ápe* *tshéeto*
apa-i tshéeto
one-CLF.BASKET *aturá*

'one *aturá* basket' [200322_ELIC_F_01]

To conclude, classifiers on adjectives clearly have a derivational function, as shown by the word class change they induce. This is the case both when they stand as independent NPs and when they modify a head noun.

6.1.1.3 On verbs

Verbs in Baniwa can be nominalized with any of a number of nominalizing suffixes, the choice of which depends on which argument function is being nominalized (see § 3.2.8). Some of the nominalization constructions require an additional classifier (see § 5.3.6), like the location nominalizer *-karo.* In (85), the noun *pawéntakaroda* 'shop, store, market' is derived from the verb *-wénta* 'to buy'.

(85) *pawéntakaroda*
pa-wénta-**karo-da**
IMPRS-to.buy-**NMLZ.LOC-CLF.GNR**

'shop/store/market (lit. 'buying place')' [200324_ELIC_F_01]

The nominalized verb functions as any bound noun syntactically, as demonstrated by the comparison between the nominalized verbs *-ímaakaroda* 'house (sleeping place)' (86a) and *-pítakaroda* 'bathtub (bathing place)' (86b), and the inherent bound noun *-pana* 'house' (86c): both kinds can take the relational suffix *-liko* 'LOC.in', which attaches to nouns (see § 3.2.7).

(86) a. *némaakarodaliko*
na-ímaa-karo-da-**liko**
3PL-to.sleep-NMLZ.LOC-CLF.GNR-**LOC.in**

'In their house (lit. 'sleeping place').' [190405_BOY_E_01]

b. *napítakarodaliko* *nhaána* *yalánawinai*
na-píta-karo-da-**liko** nhaána yalánawi-nai
3PL-to.bathe-NMLZ.LOC-CLF.GNR-**LOC.in** DEM3.PL white.person-PL
'In the bathtub (lit. 'bathing place') of those white people.' [190405_BOY_E_01]

c. *piówhaa* *áatsa* *yáahã* *wáapanaliko* *…*
pi-oówhaa aa-tsa yáahã wa-pana-**liko** …
2SG-to.sit AUX-RESTR DEM1.LOC 1PL-house-**LOC.in** …
'Stay right here in our house […]' [200308_NARR_J_01]

Similarly to adjectives, the nominalized status of verbs also permits them to function as modifiers of the NP head, in the form of relative clauses (see § 3.2.8.1, § 5.3.6), as shown in (87). The connective prefix *i-* links the relative clause *iwátshaakada* to the preceding head noun *litsínoni*.

(87) *lhíe* *litsínoni* *iwátshaakada*
lhíe(hẽ) li-tsíino-ni i-wátshaa-**ka-da**
DEM1.NFSG 3NFSG-dog-INAL CONN-to.jump-**NMLZ.AG-CLF.GNR**
'This dog of his that jumped' [190405_BOY_E_01]

The function of classifiers on verbs is very similar to their function on adjectives: both adjectives and verbs are normally employed as predicates, but they can be nominalized in order to function both as independent NPs and as modifiers of NP heads. As evidenced by this change of word class, classifiers always have a derivational function on verbal roots.

6.1.1.4 On demonstratives

Demonstratives usually occur without classifiers (see § 3.2.10, § 5.3.7). They can either function as modifiers of NP heads (88a) or constitute a full NP in their own right (88b).

(88) a. *náamaka* *nhépaka* *rhoaátahã* *ienipétti*
na-óoma-ka na-hípa-ka **rhoáta(hã)** ienipétti
3PL-to.want-SUB 3PL-to.catch-SUB **DEM4.FSG** child
'They want to catch that girl.' [200308_NARR_J_01]

b. *nhéette likápa nháahã*
nhéette li-kápa **nhá(hã)**
then 3NFSG-to.see **DEM1.PL**

'Then, he saw these.' [190405_BOY_E_01]

Classifiers on demonstratives are optional, and are only used to add a focused or contrastive reading, singling out a referent among several potential ones. The demonstrative constitutes an independent NP in both examples in (89); in (89b), the noun *tsíino* 'dog' is added to specify that the referent is a dog, but this noun receives the *-ka* suffix (see § 3.2.4) that makes it function as the predicate.

(89) a. *lhiadápana*
lhia-**dápana**
DEM1.NFSG-**CLF.HOUSE**

'that (house) one' (pointing out one among several alternatives)
[220919_ELIC_F_01]

b. *lhíana tsiinóka*
lhia-**na** tsíino-ka
DEM1.NFSG-**CLF.TRUNK** dog-SUB

'this dog' (pointing out one among several alternatives)
(lit. 'this (trunk-shaped) one, (which) is a dog') [220919_ELIC_F_01]

In other words, when classifiers are used on demonstratives, the demonstratives always function as NP heads rather than modifiers. This, as well as their optional status, shows that they cannot be inflectional suffixes. Instead, they derive an independent NP, as evidenced by the inability of the derived demonstratives to function as modifiers—which they can without classifiers, cf. (88a).

6.1.1.5 On interrogatives

The different functions of classifiers on interrogatives (§ 5.3.8) are particularly illuminating. On interrogatives, classifiers are used in all three functions, and the function is strictly connected to the interrogative construction in question. Table 6.4 shows the distribution of functions by construction. Each construction is described in more detail below.

Table 6.4: Functions of classifiers in interrogative constructions

Base		Construction		Clf form	Der.	Infl.	Qual.
kóame	'how?'	*kóame*-CLF^{c}	'which?'	complex	✓		
kóadzo	'how many?'	*kóadzo*-CLF^{c}	'what size?'	complex	✓		
		kóadzo-(CLF^{s})	'how many?'	simplex			✓
kóaka	'what/who?'	*kóaka*-CLF^{s}	'what kind of?'	simplex			✓
kadali-	'how many?'	*kadali*-CLF^{s}	'how many?'	simplex		✓	

The construction *kóame*-CLF^{c} 'which?' is derivational, as the complex form of the classifier derives a new meaning from the interrogative base *kóame* 'how?' that it attaches to. The derivation results in a nominalized form (analogous to the one on adjectives, § 6.1.1.2) that allows it to be used as a modifier of the NP head (see also Ramirez, 2020a: 334–335). (90a) shows an example of the use of *kóame* 'how?' without a classifier, where the speaker uses the expression *kóame páako*, literally translating to 'how does one say?' when she cannot think of the name of a certain type of fuel that she uses for the manioc oven. (90b) shows the use of the derived construction *kóame*-CLF^{c} 'which?', which modifies the noun *íinaro* 'woman'.

(90) a. *kóame* *páako* *háiko* *ñáme* *háapee*
kóame pa-áako háiko ñáme háapee
how IMPRS-to.say tree/wood NEG really

'What's-its-name (lit. 'how does one say?'), wood, no, not really…' [200308_NARR_B_01]

b. *koamédaro* *íinaro* *piokéeta?*
kóame-da-ro íinaro pi-ookéeta
how-CLF.GNR-FSG woman 2SG-to.meet

'Which woman did you meet?' [220928_ELIC_G_01]

The interrogative base *kóadzo* means 'how many?', and can be used either without (91a) or with (91b) a classifier to ask about the quantity of something. A noun can be either expressed or omitted in both cases, depending on the level of specificity required. The optional addition of the classifier in the construction *kóadzo*-(CLF^{s}) 'how many?' does not change the meaning of the interrogative, but merely narrows down the possible reference to referents of a certain kind, e.g., packages in (91b). Therefore, this is analyzed as an example of the qualifying function.

(91) a. *nhéette* *kóadzo* *tsheetoéna* *idéenhikaro*
nhéette **kóadzo** tsheeto=eéna(hã) i-déenhi-karo
then **how.many** aturá=DEM3.NFSG CONN-to.work-in.order.to
aapána [*lata*]*?*
apa-na [lata]
one-CLF.TRUNK can

'So how many *aturá* baskets (of manioc) does it take to make one can (of *tapioca*)?' [200308_NARR_B_01]

b. *koadzópa?*
kóadzo-pa
how.many-CLF.PACKAGE

'How many?' (of packages) [220927_ELIC_G_01]

Classifiers also appear in a different construction on *kóadzo* 'how many?', namely *kóadzo*-CLF^{c} 'what size?' (92). This use differs from the one in (91b): it derives a very different meaning, and it uses a different form of the classifier (the complex one, which acts as a nominalizer). These two properties demonstrate that the classifier here is used in the derivational function.

(92) *koadzópali?*
kóadzo-**pa-li**
how.many-**CLF.PACKAGE-NF**

'What size?' (of a package) [220927_ELIC_G_01]

The interrogative base *kóaka* 'what/who?' is composed of the interrogative morpheme *kóa* and the suffix *-ka* (see § 3.2.4), which shows that it is used as a predicate (§ 5.3.8.2). In (93a), it is used without a classifier. In (93b), the classifier *-da* is added, narrowing down the meaning from 'what?' to 'what kind/type of?', at the same time specifying that the referent asked about is expected to be compatible with the classifier. This is analyzed as a qualifying use of the classifier, as it does not radically change the meaning of the interrogative base, nor its syntactic properties (it still functions as the predicate). Accordingly, the construction takes the simplex classifier form.

(93) a. *kóakatshaa pidéenhiri?*
kóa-ka-tshaa pi-déenhi-ri
Q-SUB-CEXP 2SG-to.do-REL

'What (is it) that you are doing?' (adapted from Ramirez, 2020a: 332)

b. *koakáda haikóthe píiñhali?*
kóa-ka-**da** haikóthe pi-íiñha-li
Q-SUB-**CLF.GNR** fruit 2SG-to.eat-REL

'What kind of fruit (is it) that you ate?' [220928_ELIC_G_01]

The use of classifiers on *kóaka* 'what/who?' is noteworthy, as *kóaka* functions as a predicate. Apart from this construction and the handful of adjectives that take classifiers also in their predicative use (§ 5.3.4), classifiers only occur on elements within the noun phrase.

Finally, the interrogative *kadali-* 'how many?' (94) differs from the interrogatives discussed above in that it displays inflectional classifier use. This interrogative, in contrast to the others, is a bound morpheme that never appears without a classifier, much like the lower numerals 1–3 (§ 5.3.3.1, § 6.1.1.1) with which it has an obvious link (see, for example, the question–answer pair in (73)).

(94) *kadalhípa atsíanli?*
kadali-**hípa** atsíanli
how.many-**CLF.MALE** man

'How many men?' [220927_ELIC_G_01]

Thus, on interrogatives, classifiers are attested in all three functions. The classifier is derivational in *kóame*-CLF^{c} 'which?' and *kóadzo*-CLF^{c} 'what size?', as it leads to a salient change in meaning (and, at least in the former construction, a change in syntactic properties). In *kóadzo*-(CLF^{s}) 'how many?' and *kóaka* 'what/who?', the classifier is used in the qualification function, as it is optional and does not change the meaning of the expression significantly. In *kadali*-CLF^{s} 'how many?', the classifier is inflectional since it is obligatory.

6.1.1.6 On nouns

On nouns, the simplex form of the classifier attaches directly to the root. A classifier can derive new nouns from both free and bound nouns, and the resulting

noun remains in the same noun category. In (95a), the free noun *iitsaákhaa* 'fishing line' is derived from the free noun *iítsa* 'fish hook' with the curvilinear classifier *-khaa*. In (95b), the bound noun *-thiíta* 'glasses, spectacles' is derived from the bound noun *-thi* 'eyes' with the human classifier *-iíta*[66]. These examples constitute cases of derivation (Rose & Van linden, 2023: 263): a fishing line is not a type of fish hook, and glasses are not a type of eyes.

(95) a. *iitsaákhaa*
iítsa-**khaa**
fish.hook-**CLF.CURV**

'fishing line' [200327_ELIC_F_04]

b. *pathiíta*
pa-thi-**iíta**
IMPRS-eyes-**CLF.HUMAN**

'glasses, spectacles' [200328_ELIC_F_04]

Classifiers attached to nouns can also be instances of qualification, if the N-CLF construction "designates an instance of the type denoted by the nominal root" (Rose & Van linden, 2023: 263; see also Pepper, 2020: 203). The qualification function is also applicable to both free and bound nouns. In (96a), the surface classifier *-koa* is used to specify that the stone is flat. In (96b), the generic classifier *-da*, which also has roundness connotations (see § 7.1.8, § 7.3.3.1) is used to emphasize the roundness of the knot.

(96) a. *hipákoa*
híipa-**koa**
stone-**CLF.SURFACE**

'flat stone' [200318_FREE_G_01]

b. *piwathéda*
pi-wáthe-**da**
2SG-knot-**CLF.GNR**

'your (round) knot' [200309_FREE_K_01]

[66]Here, *-iíta* is used in its 'object with a flat part' sense (see § 7.1.22).

6.1.2 Derivation: summary

The derivational function of classifiers shows up on all six morphosyntactic hosts, and is therefore the most widespread morphological function in the classifier system. Derivation is lexeme-external in the sense that it gives rise to new lexemes, which in the case of Baniwa classifiers always results in nouns. In this way, classifiers can be said to have a nominalizing function.

In some cases, the derivational function can be evidenced by the morphological form of the derived word. This is the case for adjectives and verbs, which are primarily predicative but can be nominalized in order to be used as modifiers in NPs. In both of these cases, there is also a more dedicated nominalizer. The two interrogatives where classifiers are used derivationally pattern with adjectives in taking the same complex form consisting of a classifier and a nominalizer. In all of these cases, the nominalization also permits access to the NP domain to a part of speech that is usually found outside the NP.

In other cases, the derivational function cannot be discerned merely from the form of the derived word. On numerals, for instance, the morphological form is identical regardless of whether it is used inflectionally or derivationally. Here, the derivational function is evidenced instead by syntactic criteria alone: the ability of the derived form to head the NP.

Finally, on nouns, the derivational function can only be evidenced by semantic criteria; more precisely, that the classifier derives a denotation that is not a type of the denotation of the nominal root.

6.1.3 Inflection: summary

Inflection is a marginal function of classifiers in Baniwa: it only shows up on lower numerals and on the interrogative *kadali-* 'how many?'. The main characteristic of these roots is the fact that they never occur without classifiers. This sets them apart from other classifier hosts. They also take the simplex classifier form, in contrast to, e.g., adjectives and derived interrogatives.

The inflectional function of classifiers is an example of inherent rather than contextual inflection, as it contributes some independent information, and the inflectional value is not determined by the syntactic context in the same way that, e.g., core case marking is (Booij 1996, Haspelmath & Sims 2010: 100; see also § 2.7.1.1). The relationship between inflection and derivation in Baniwa classifiers is elaborated in § 6.4–6.6 below.

6.1.4 Qualification: summary

Qualification (property assignment) occurs on two types of hosts: nouns and certain interrogatives. The qualified noun (N-CLF) is formally identical to the derived noun, but can be distinguished semantically, as the qualified expression denotes an entity of the same kind as the root. Likewise for the interrogatives where classifiers are used for qualification, the meaning of the root is narrowed down or specified. Similarly to the inflectional function, only the simplex classifier form is used for qualification purposes.

6.2 Referent specification

The flexible nature of classifier assignment (§ 4.4) allows classifiers to play a central role when it comes to identifying referents. As seen in § 6.1.1.6, classifiers can be used on nouns to derive new lexemes. Some further examples of this function are shown in (97), where three different lexemes are derived from the noun *haíko* 'tree, wood'[67].

(97) *háiko* 'tree' [200313_FREE_F_01]

a. *haikóapo*
háiko-**áapo**
tree/wood-**CLF.STICK**
'stick'

b. *haikoíta*
háiko-**iíta**
tree/wood-**CLF.HUMAN**
'board'

c. *haikhíwi*
háiko-**híwi**
tree/wood-**CLF.POINTED**
'twig'

However, a similar way of referent specification can also be achieved through the use of classifiers on modifiers. In many cases, the noun and the classifier both contribute to establishing the meaning of an expression, in a way that is semantically reminiscent of the derivational examples in (97) above. The phrases in (98) are minimal triplets differing only in classifier choice on the numeral,

[67] In (97b), as in (95b) above, *-iíta* is used in its 'object with a flat part' sense (see § 7.1.22).

which results in quite different readings. Likewise, in (99), four different shape classifiers are used to establish different referents of the noun *kamárai* 'light'.

(98) *óowi* 'war' [200327_ELIC_F_04]

a. *apáda* *óowi*
apa-**da** óowi
one-**CLF.GNR** war
'one war'

b. *apaíta* *óowi*
apa-**iíta** óowi
one-**CLF.HUMAN** war
'one soldier'

c. *apáana* *óowi*
apa-**áana** óowi
one-**CLF.GROUP** war
'one army/group of soldiers'

(99) *kamárai* 'light' [200228_ELIC_F_01], [200305_ELIC_F_01]

a. *apáda* *kamárai*
apa-**da** kamárai
one-**CLF.GNR** light
'one lamp'

b. *apákhaa* *kamárai*
apa-**khaa** kamárai
one-**CLF.CURV** light
'one line of string lights'

c. *aapána* *kamárai*
apa-**na** kamárai
one-**CLF.TRUNK** light
'one flashlight'

d. *aphéwi* *kamárai*
apa-**híwi** kamárai
one-**CLF.POINTED** light
'one candle'

Crucially, the classifier modifies the meaning of the head noun rather than the meaning of the host. For this reason, they are not treated as examples of derivation here.

6.3 Functions in discourse

Cross-linguistically, classifiers typically have some functions on discourse level, such as identifying, establishing, disambiguating, and tracking referents, or altering their status in terms of definiteness, specificity, topicality, thematic salience,

etc. (Contini-Morava & Kilarski, 2013: 279–291). A common discourse-level function of classifiers is their anaphoric use, that is, their ability to refer back to a referent that has been established by an antecedent earlier in discourse (Lyons, 1977: 660; Contini-Morava & Kilarski, 2013: 280).

In Baniwa too, classifiers can be used anaphorically. When used in this function, the classifier typically occurs without the noun as it is redundant. (100) shows an utterance from a story, where the storyteller describes how to make *tapioca* (manioc starch). The noun phrase *apáda* [*hora*] 'one hour' is mentioned in full in the beginning of the utterance. Subsequently, the numeral–classifier combination *dzamáda* 'two (hours)' is enough to understand that the storyteller is still talking about hours.

(100)

apáda	[*hora*]	*pándza*	*pakádaa*	*dzamádanako*	*pándza*
apa-da	[hora]	pándza	pa-kádaa	**dzama-da**-nako	pándza
one-CLF.GNR	hour	now	IMPRS-to.leave	**two-CLF.GNR**-LOC.on	now

pakéñoa	*paároka*	*líawa*	*lhiéna*	*óoni*
pa-kéñoa	pa-aáro-ka	li-aa-wa	lhiéna(hã)	óoni
IMPRS-to.begin	IMPRS-to.pour-SUB	3NFSG-to.go-MID	DEM3.NFSG	water

'One leaves it one hour, in two (hours), one starts to pour out that water.'
[200308_NARR_B_01]

Anaphora can also appear in consecutive utterances. In (101), the overt noun *heekóapi* 'day' is mentioned in the first utterance. In the next utterance, the duration of a day is referred back to twice: once with the numeral *apa-* and once with the demonstrative *lhia-*. Both take the day classifier *-daa*, and no overt noun is expressed in either case.

(101) a.

lhíatsheenahã	*apádaa*	*heekóapikatsa*
lhíatsha-iina=áa(hã)	**apa-daa**	heekóapi-ka-tsa
however-already=DEM1.LOC	**one-CLF.DAY**	day-SUB-RESTR

pamettátakatsani	*dorómeri*	*óohõ*
pa-mettáta-ka-tsa-ni	doróme-li	óohõ
IMPRS-to.dry-SUB-RESTR-3NFSG	time.consuming-REL	yes

'But still, [it takes] a whole day just to dry it, which is time-consuming… that's right.'

b. *apádaakatsa* *pakótshokani*
apa-daa-ka-tsa pa-kótsho-ka-ni
one-CLF.DAY-SUB-RESTR IMPRS-to.wash-SUB-3NFSG

lhíadaatsatha *lhiéna* *pimeérani* *óohõ*
lhía-daa-tsa-tha lhiéna(hã) pi-meéra-ni óohõ
3NFSG-CLF.DAY-RESTR-FRUS DEM.3NFSG 2SG-to.toast-3NFSG yes

'[It takes] one day to wash it, that very same one (day) that you toast it, yes.' [200308_NARR_B_01]

However, in the case of the demonstrative *lhia-*, the absence of an overt noun can hardly be taken as evidence for anaphoric use of the classifier, as classifiers on demonstratives do not form NPs with overt nouns at all. Instead, this is an example of when classifiers are used for focus or contrast, a function that is associated with demonstratives specifically (see § 5.3.7, § 6.1.1.4).

Further examples of anaphoric reference can also be seen in answers to questions. In (102a), it is established that women are the referents whose quantity is in question. In the answer in (102b), then, the anaphoric use of the classifier conveys enough information, and no overt noun is expressed.

(102) a. *kadalíma* *íinaro?*
kadali-**ma** íinaro
how.many-**CLF.FEMALE** woman

'How many women?' [220927_ELIC_G_01]

b. *aapámatsa*
apa-**ma**-tsa
one-**CLF.FEMALE**-RESTR

'Just one (woman).' [220927_ELIC_G_01]

A similar example from discourse is shown in (103). In (103b), a classifier on the numeral *dzama-* 'two' is enough as a reference to 'two baskets', since the baskets were established as referents by the antecedent *tshéeto* '*aturá* basket' in the preceding question in (103a).

(103) a. *nhéette* *kóadzo* *tsheetoéna* *idéenhikaro*
nhéette kóadzo tsheeto=eéna(hã) i-déenhi-karo
then how.many aturá=DEM3.NFSG CONN-to.work-in.order.to
aapána [*lata*]*?*
apa-na [lata]
one-CLF.TRUNK can

'So how many *aturá* baskets (of manioc) does it take to make one can (of *tapioca*)?'

b. *dzáme* *tsoo-…* *náani* *tsoopekádaatsa* *káini*
dzama-i tsoo-… náani tsoo-pe-kádaa-tsa káini
two-CLF.BASKET little-… SUSP little-PL-COND-RESTR manioc
dzáme
dzama-i
two-CLF.BASKET

'Two… well, if the maniocs are few, then two (baskets).'
[200308_NARR_B_01]

The anaphoric use of classifiers has been identified in other Amazonian languages (Krasnoukhova, 2012: 212), including in the Arawakan family (Dunn & Rose, forthcoming: 30–32). Classifiers can also have other discourse-level functions in Arawakan languages: in Mojeño Trinitario (Bolivia), when a classifier occurs on a verb without an associated NP, the referent is backgrounded, whereas an overt NP in combination with a classifier on the verb foregrounds the referent (Dunn & Rose, forthcoming: 30–32). This is interesting in light of the fact that in Baniwa, classifiers are used on demonstratives without overt nouns to express focus or contrast; thus, classifiers in Arawakan languages appear to be capable of expressing both increased and decreased prominence. A systematic study on a larger corpus of naturalistic texts is needed in order to examine the discourse-level functions of Baniwa classifiers in greater detail.

6.4 The distinction between inflection and derivation

In § 6.1.1.2, it was shown that classifiers can derive nouns from adjectives. For instance, the adjective *poottídza* 'sweet' gives rise to the noun *poottidzádali* 'candy' (104a). However, the form *poottidzádali* can also appear as an attributive adjective modifying the nominal head of the NP, as in (104b) where it is used in the adjectival sense 'sweet' about wild fruit. In this case, the classifier could be seen

as both derivational (as it participates in the nominalization of the adjective)[68] and inflectional (as it appears on a modifier, refers to the head, and is obligatory in this context).

(104) a. *poottidzádali*
poottidza-**da-li**
sweet-**CLF.GNR-NF**

'candy' [200326_ELIC_F_03]

b. *haikóthe poottidzádali*
haikóthe poottidza-**da-li**
wild.fruit sweet-**CLF.GNR-NF**

'sweet fruit' [200326_ELIC_F_03]

Although the semantics ('candy' vs. 'sweet') and NP structure (independent vs. modifying) do highlight some differences between the two uses of *poottidzádali*, the fact that they are formally identical complicates the analysis. The question is at what point the suffix sequence *-da-li* stops being derivational and becomes inflectional.

Nominalized verbs (§ 6.1.1.3) can be analyzed in a similar way: when they are used as relative clauses modifying NP heads, the classifier can be viewed as either derivational or inflectional. The case for the inflectional use of classifiers in relative clauses is weaker, however, as only some of the nominalizers require an additional classifier (see § 3.2.8). Thus, as classifiers can be shown to be part of the derivational template of particular nominalization constructions, and as relative clauses with some nominalizers lack classifiers altogether, it is difficult to argue for the inflectional status of classifiers in this context.

Numerals exhibit the inverse situation. As shown in § 6.1.1.1, numerals obligatorily take classifiers, and a numeral's primary position is as a modifier in a noun phrase. This makes numerals different from attributive adjectives and relative clauses, which can also modify NP heads but have to be nominalized first, as the primary function of the roots (adjectival vs. verbal) is predicative. Consequently, classifier use on numerals is considered here as inflectional.

However, noun phrases require heads in the form of nouns (§ 3.2.6), and classifier-inflected numerals frequently appear without the head noun without any addi-

[68] The fact that the classifier is expressed closer to the root than the nominalizer *-li*, arguably a derivational morpheme, also supports the derivational analysis of the classifier (cf. criterion 9, position relative to the base, in § 2.7.1.1).

tional morphology. The classifier can then arguably also serve a derivational function on numerals: it derives independent NPs.

§ 2.7.1 contained a theoretical discussion on whether or not inflection and derivation can be distinguished systematically. Some scholars, such as Bickel & Nichols (2007), argue that they are distinct; others, such as Bybee (1985), argue that they are end points on a continuum; yet others, such as Booij (1996), argue for a more nuanced understanding of the inflectional category. As shown above, Booij's model works well for Baniwa classifiers, as they can be accounted for as instances of inherent inflection in some cases. Examples such as the ones discussed in this section, where the same instance of a classifier can be analyzed as both inflectional and derivational, also lend some support to the continuum approach (see also Payne's (1985) discussion on the inflectional and derivation functions of classifiers in Yagua (Peba-Yaguan, Peru)).

6.5 Why classifiers are not agreement devices

Agreement is defined by Steele (1978: 610) as "some systematic covariance between a semantic or formal property of one element and a formal property of another". The question of agreement in nominal classification systems is of particular relevance since, as Seifart (2005: 157) notes, "[t]he presence of agreement is a crucial criterion in the typology of systems of nominal classification, where it is taken as the main definitional characteristic of noun class systems as opposed to classifier systems" (see also § 2.1). Baniwa classifiers are not used as agreement devices: modifiers in the same NP may be marked with the same classifier (105a), but they may also be marked with different classifiers (105b, 105c).

(105) a. *aphéwi* *háiko* *mhatsamhéwi*
apa-**híwi** háiko mhatsa-me-**hiwi**
one-**CLF.POINTED** tree/wood snap-ADJZ-**CLF.POINTED**

'one broken (snapped) stick/twig' [200322_ELIC_F_01]

b. *aapána* *palána* *kakoadádali*
apa-**na** palána kakóada-**da-li**
one-**CLF.TRUNK** banana expensive-**CLF.GNR-NF**

'one expensive banana' [200324_ELIC_F_01]

c. *apákhaa* *óoni* *tsoopéki*
apa-**khaa** óoni tsoo-**péko-i**
one-**CLF.CURV** water small-**CLF.PATH-N**

'one small river' [200228_ELIC_F_01]

Speakers even accept phrases such as the one in (106). The three classifiers *-khaa* 'CLF.CURVILINEAR', *-péko* 'CLF.PATH', and *-páwa* 'CLF.RIVER' are all regularly used for creeks, and the semantic compatibility allows them to be marked on one modifier each (a numeral and two adjectives) in the same NP.

(106) *apákhaa* *ñawápo* *makapéki* *iittapáwani*
apa-**khaa** ñawápo maka-**péko-i** íitta-**páwa-ni**
one-**CLF.CURV** creek big-**CLF.PATH-N** black-**CLF.RIVER-?**

'one big black creek' [221005_ELIC_G_01]

The existence of non-agreeing phrases is unsurprising, given that agreement is a syntactic phenomenon that essentially presupposes inflectional morphology (Dahl, 2004: 196). As shown above (§ 6.1), classifiers in Baniwa are derivational rather than inflectional on all NP modifiers except numerals. This explains why modifiers within the same NP can take different classifier marking: the function of classifiers on, e.g., attributive adjectives is not to show agreement (and thereby constituency) with the head noun, but primarily to grant the adjective access to the modifier position in the first place. The multiple marking of classifiers in the Baniwa NP is thus a by-product of its constraints on various modifiers (i.e., that they have to be nominalized, and that the nominalization constructions contain classifiers), rather than a device for marking agreement within the NP. The requirement on various NP modifiers to be nominalized, with classifiers or other nominalizing morphemes, appears to be relatively widespread in South America (Krasnoukhova, 2012: 266). For instance, in Kwaza (unclassified, Brazil), demonstratives, numerals and attributive roots must be nominalized in order to occur as NP modifiers (van der Voort, 2004: 180–181).

However, the Baniwa NP may shed some light on the possible development of agreement of the gender/noun class type (see also Greenberg, 1978; Corbett, 1991: 137 ff.). Baniwa classifiers are part of derivational templates that are required for certain parts of speech to function as NP modifiers. Thus, for an adjective to function attributively, it necessarily carries classifier marking. As classifiers (whether used derivationally or inflectionally) reflect some semantic

property of the nominal referent, these derivational classifier suffixes, which *de facto* express some kind of semantic agreement with the head noun, could in time become interpreted as inflectional suffixes that have to agree across the NP. However, it should be emphasized that systems like the one in Baniwa are functional in their own right, and agreement marking should not be seen as an evolutionary goal.

From a genealogical perspective, classifier agreement is rare in Arawakan multilocus classifier systems due to the flexible assignment principles and optional status in many of the languages (Dunn & Rose, forthcoming: 27–29)[69]. Classifier systems in some unrelated Northwestern Amazonian languages, notably Miraña (Bora-Witoto) and Tatuyo (Tucanoan), have been described as exhibiting systematic agreement, although it operates via a generic–specific division, where general class markers can replace specific class markers in agreement marking without affecting co-referentiality (Seifart, 2005: 80; Gomez-Imbert, 2007).

6.6 Noun phrase structure revisited

In § 3.2.6, the basic NP structure was described as follows (adapted from Ramirez, 2020a: 258):

[DEM + NUM + NOUN + ADJ + RC]

The head of the noun phrase must consist of a noun. If a noun phrase contains a non-derived noun, it functions as the head, as it cannot function as a modifier. However, as has been shown repeatedly above, nouns are frequently absent from the noun phrase, in which case an element that typically functions as a modifier can take over the position as head: numerals (81), nominalized adjectives (84a), nominalized verbs (86a), and demonstratives (89b) can all head NPs. In addition, with the exception of demonstratives[70], there are no morphological clues to the status of the element in question as a modifier or head, as it does not require any additional marking. Sometimes, as in (83), a word may have different meanings when used as an adjective ('sweet') and as a noun ('candy'), but not all modifiers have such lexicalized counterparts. In other cases, the analysis is truly ambiguous, as in (107) from Ramirez (2020a: 255).

[69] But see Aikhenvald (2019) for a different view.

[70] When demonstratives take classifiers, they form full NPs on their own—in case the speaker wishes to specify the referent by means of a noun, this noun is pushed out into an existential clause. Therefore, demonstratives will be excluded below in the discussion of ambiguous cases of headedness relations within the NP.

(107) *noóma* *apáda* *kamítsha* *iittádali* *nheétte*
no-oóma apa-da kamítsha íitta-da-li nheétte
1SG-to.want one-CLF.GNR shirt black-CLF.GNR-NF then

apáda *iiráidali*
apa-da **íirai-da-li**
one-CLF.GNR **red-CLF.GNR-NF**

'I want one black shirt and **one red (shirt)**'
(adapted from Ramirez, 2020a: 255)

The clause above features two coordinated noun phrases (*apáda kamítsha iittádali* 'one black shirt' and *apáda iiráidali* 'one red (shirt)'). The first NP features a nominal head (*kamítsha* 'shirt'), which is omitted in the second one. As we have seen above, both numerals and nominalized adjectives can function as both modifiers and heads of NPs. In *apáda iiráidali*, the morphological marking is identical to the counterparts in the preceding *apáda kamítsha iittádali* for both numeral and adjective. Thus, the NP is in theory compatible with two analyses: one where the numeral is the syntactic head (a), and one where the adjective is the syntactic head (b).

a. *apáda iiráidali*

b. *apáda iiráidali*

A similar example is shown in the isolated phrase (108), which is compatible with the headedness analyses in both (a) and (b)[71].

(108) *apaíta* *manaanidálitsa*
apa-iíta ma-náani-da-li-tsa
one-CLF.HUMAN PRIV-clothes-CLF.GNR-NF-RESTR

'one naked person' [200311_ELIC_K_04]

a. *apaíta manaanidálitsa*

b. *apaíta manaanidálitsa*

[71]The restrictive suffix *-tsa* does not facilitate the analysis, as it is used here in a lexical sense, accompanying the privative adjectivizer *ma-* (Ramirez, 2020a: 269–270, 283–284), and therefore does not have any syntactic relevance.

Yet another example is shown in (109), but with a numeral and a nominalized verb (with an accompanying adpositional phrase). The numeral can be interpreted as the head with the nominalized verb as a modifying relative clause: "one who guards underneath it", (a), or the nominalized verb can be interpreted as the head, with the numeral as a modifier: "one underneath-it-guarder", (b).

(109)

apaíta	*iwápakada*	*líaphitte*
apa-iíta	i-wápa-ka-da	li-áaphitte
one-CLF.HUMAN	CONN-to.wait/guard-NMLZ.AG-CLF.GNR	3NFSG-beneath

'one who guards underneath it' / 'one underneath-it-guarder'
(said about a person who keeps watch over the fire in a manioc oven during the process of roasting *tapioca*) [200308_NARR_B_01]

a. *apaíta iwápakada líaphitte*

b. *apaíta iwápakada líaphitte*

The ambiguity stems from the fact that all elements in the NP are able to form NPs on their own—that is, they are all essentially nouns, as illustrated in the modified NP structure below.

$[[\text{DEM}]_{\text{NP}}\ [\text{NUM-CLF}]_{\text{NP}}\ \underline{[\text{N}]_{\text{NP}}}\ [\text{ADJ-CLF(-NMLZ)}]_{\text{NP}}\ [\text{POSS-V-NMLZ-(CLF)}]_{\text{NP}}]_{\text{NP}}$

The NP structure is reminiscent of a construction involving the apposition of separate NPs (cf. Epps, 2008: 278–279 for Hup (Naduhup, Brazil)), that is, a construction where co-referential modifiers "are not part of the (integral) phrase containing the head noun" (Rijkhoff, 2002: 22). The lack of agreement between modifiers (§ 6.5) also suggests the possibility of an appositional relationship. However, the NP word order is relatively fixed, which is indicative of NP constituency, as otherwise a more free constituent order might be expected. Another argument in favor of the modifiers actually forming a larger NP unit comes from intonation; appositional phrases tend to be separated by pauses (Acuña-Fariña, 1999), which is not the case in Baniwa NPs.

6.7 Chapter summary

The analysis of the functions of classifiers highlights the versatility of the system. On word level, classifiers are used for inflection, derivation, and qualification. Different morphosyntactic hosts have different functions available to them.

Derivation is the most widespread function in terms of the number of loci, being a possible function in all of them. Inflection is almost exclusively used on numerals, and qualification on nouns. Interrogatives are interesting because they display a split, where different constructions are connected to the three different functions.

The distinction between inflection and derivation is not always straightforward; certain constructions can be analyzed in different ways. This also complicates the analysis of the noun phrase structure, as it is not always clear what is a head and what is a modifier.

Classifiers also serve purposes above word level. They play an important role in specifying referents of nouns by assigning properties to them. In discourse, classifiers can be used to refer back to previously mentioned referents, as well as to express focus or contrast.

Chapter 7

Semantic properties

This chapter explores the semantic properties of the classifier system. § 7.1 provides individual semantic descriptions of all classifiers. § 7.2 examines the semantic parameters along which classification is structured. § 7.3 explores the semantic core and extension of classifiers and describes their structure as categories of referents. Finally, a number of additional themes relating to semantics are brought up: § 7.4 discusses the variation in classifier use between speakers, § 7.5 describes how multiple referents associated with different classifiers are classified, and § 7.6 argues against a distinction between mensural and sortal classifiers in Baniwa.

7.1 Semantic descriptions of individual classifiers

The purpose of this section is to provide detailed descriptions of the semantic properties of each classifier, before going into generalizations in the subsequent sections. The descriptions in § 7.1.1–7.1.53 comprise the full set of documented classifiers, and are presented in alphabetical order.

The descriptions in this section are based on the entire material at hand, especially the Noun list and Free listing data sets (§ 1.2.4). A concise table of the classifier inventory and some examples of the referents each classifiers typically associates with can be found in Table 4.1 (§ 4.2).

7.1.1 *-Ø* 'CLF.CANOE'

The zero classifier *-Ø* is associated with containment, and is used for objects that have a hollow shape as a result of having been carved. This comprises canoes (*íita*), as well as different types of calabash bowls (*kóoya*, *átta*[72]). Extending from the use for canoes, *-Ø* can also be used for other container-like vehicles such as airplanes, buses, and cars (but not bicycles; see § 9.1).

For some speakers, it is possible to use *-Ø* with skin and bark (referents that are more commonly associated with *-ya* 'CLF.SKIN'). There is a possible etymological link between *-Ø* and *-ya* that is described in § 8.3.1.

For non-carved hollow objects, see *-áapi* 'CLF.HOLLOW' (§ 7.1.6).

7.1.2 *-áana* 'CLF.GROUP'

The group classifier *-áana* is used for multitudes of entities of all kinds, spanning groups of humans (e.g., *nawíki* 'person', *-kitsíenaa* 'relatives'), animals (e.g., *pówe* 'monkey (generic)', *aapídza* 'white-lipped peccary', *doólo* 'insect') and inanimate entities (e.g., *-náani* 'belongings', *paiñhakarodápe* 'kitchen utensils').

7.1.3 *-áanhaa* 'CLF.LIQUID'

The liquid classifier *-áanhaa* is associated with fluidity, and is used for water (*óoni*), for liquids of fruits and vegetables (e.g., *manákhea* 'açaí juice', *maawíroa* 'pineapple juice', *káaphe* 'coffee', *ttíape* '*tucupí* (sauce made from manioc liquid)', *káapi* '*caapi* (hallucinogenic beverage)', *patshíakaa* '*chibé* (drink made of water and manioc flour)', *yalákhi* 'fermented drink'), for bodily fluids (e.g., *-aanómhaa* 'saliva'), and for any other types of liquids (e.g., [*gasolina*] 'gasoline', *tápee* '(liquid) medicine', *-daataxoópa* 'paint').

This classifier is also used for certain units of time, most commonly *míitsha* 'week' (§ 7.2.2.7). Some speakers also use it for *hámoli* 'dry season; year'.

[72] *Kóoya* bowls are typically teardrop shaped, and are made from a fruit that grows on a vine. They are used both as bowls and as ladles. *Átta* bowls are rounder and made from the fruit of a tree.

7.1.4 *-aápa* 'CLF.OBLONG'

The classifier *-aápa* is associated with distinctly elongated, three-dimensional shapes, and is used for both animate beings and inanimate objects.

It is used for a wide array of different animals: most birds (e.g., *kepíreeni* 'small bird (generic)', *kóitsi* 'curassow', *kaláka* 'hen/rooster'), other flying animals (*píittiri* 'bat'), a few fish species (*kettínali* '*jacundá*', *táali* '*aracu*'), some small mammals (e.g., *dzáawi* 'ocelot', *phíitsi* 'agouti (rodent sp.)', *póotto* 'small agouti sp.'), as well as some insects (*kawípe* 'ant sp.') and crustaceans (*dzáaka* 'shrimp').

It is commonly used with some fruits and vegetables of oblong shape, referring to the individual fruit (e.g., *palána* 'banana', *hawádza* '*ingá* sp.', *káana* 'ear of corn', *káini* 'manioc tuber', *korhitáko* 'large variety of *malagueta* pepper'). Less commonly, it can refer to other inanimate objects of an oblong shape, e.g., [*pão*] '(loaf of) bread', *tsháako* 'bag'.

7.1.5 *-áaphi* 'CLF.AREA'

The classifier *-áaphi* is used for delimited areas and places, e.g., *kiníki* 'manioc garden', *káida* 'beach', *kalítta* 'lake', *dzakálee* 'village, town', *-hinaáphia(mi)* 'shadow', and *midzáka* 'sacred place'.

It is commonly used to derive nouns from verbs together with the location nominalizer *-karo* (§ 5.3.6.3), resulting in nouns referring to locations where certain actions take place, e.g., *-kadzeekataakakaróaphi* 'classroom (lit. place of learning)', *-deenhikaróaphi* 'workplace (lit. place of working)', *-topikakaróaphi* [*bola*] 'soccer field (lit. place of playing)'.

This classifier can also be used for holes and openings of different kinds, especially when emphasizing their location, for example *-nómawa* 'hole', *panttinóma* 'door', *moólo* 'back door; hole of armadillo or paca' (see also *-péko* 'CLF.PATH' (§ 7.1.38), *-wa* 'CLF.SPACE' (§ 7.1.46), and *-yáwa* 'CLF.HOLE' (§ 7.1.53)).

7.1.6 *-áapi* 'CLF.HOLLOW'

The classifier *-áapi* is used for hollow containers such as *paráto* 'plate', *tsipaláapi* 'pot, pan', *aakhéepa* 'clay pot', and [*bacia*] 'bowl', as well as for bones, which are also hollow (e.g., *iñápi* 'bone', *-peréma* 'rib', *-ttamáapi* 'dorsal spine'). It can also be used with liquids, e.g., *manákhea* 'açaí juice' and *óoni* 'water', to refer to a

container full of the liquid in question.

For carved containers like canoes and calabash bowls, see *-Ø* 'CLF.CANOE' (§ 7.1.1).

7.1.7 *-áapo* 'CLF.STICK'

The classifier *-áapo* is associated with thin objects with a salient one-dimensional extension. It is used for sticks and stick-like objects (e.g., *haikóapo* 'stick', *iitsáapo* 'fishing rod', *idzawithiapótti* 'bow'), as well as for other things that are long in shape, such as trees (e.g., *máawi* 'palm sp.'). It is also used for passages of various kinds, which may not be physical objects but which are still perceived as having a dimension of length (e.g., *inípo* 'path', *ñawápo* 'creek', *-hitakóliko* 'nostrils').

7.1.8 *-da* 'CLF.GENERIC'

The classifier *-da* originated as a classifier for round objects (see § 8.1), and has subsequently developed into a generic classifier.

Although *-da* can be used with almost any referent under the right circumstances (see § 7.3.3.1), it is most canonically used with round(ish), three-dimensional referents like stones (*hiipáda*), eggs (*-éehwe*), tortoises (*itsída*), piranhas (*ómai*), and round fruits (*-iináka* 'fruit (from a tree)', *manákhe* 'açaí (individual fruit)'). This is also the classifier used for preadolescent human children (*ienipétti*). It is sometimes used for referents that do not fit neatly into any other category, as is the case for the Portuguese loanwords [*computador*] 'computer' and [*caderno*] 'notebook', but even referents that are closely associated with another classifier can typically be used with *-da* as well (e.g., *heekóapi* 'day' (*-daa* 'CLF.DAY'), *íita* 'canoe' (*-Ø* 'CLF.CANOE'), *tshéeto* '*aturá* basket' (*-i* 'CLF.BASKET'), *tsipaláapi* 'pot, pan' (*-áapi* 'CLF.HOLLOW'), and *pántti* 'house' (*-dápana* 'CLF.HOUSE')).

The generic function of *-da* and its place within the classifier system is described in more detail in § 7.3.3.1.

7.1.9 *-daa* 'CLF.DAY'

The classifier *-daa* is used for units of solar time. It occurs only with two nouns: *heekóapi* 'day' and *déepi* 'night'.

7.1.10 *-dápana* 'CLF.HOUSE'

The classifier *-dápana* is associated with human-made structures for human occupancy. It is used for permanent buildings (e.g., *pántti* 'house', *tapeedápana* 'hospital') as well as for temporary habitations (e.g., *pamákali* 'tent'), but not with animal habitations. It can also be used for the afterworld (*iarodátti*).

7.1.11 *-eéma* 'CLF.SIDE'

The classifier *-eéma* is used to refer to one side of an entity, or to one of a pair of entities. A common function of *-eéma* is its use with body parts. There, it refers to one of a pair of entities, e.g., *-híipa* 'foot (one of a pair)', *-káapi* 'hand (one of a pair)', *-kawa* 'leg (one of a pair)', *-héeni* 'ear (one of a pair)', *-thi* 'eye (one of a pair)', *-náapa* 'arm (one of a pair)'. The same principle applies to other paired entities, e.g., *-heenittáda* 'earring (one of a pair)'. Such overt marking of a singular item in a pair is called antidual marking (Michaelis et al., 2013a: 104–105). Antidual marking in Baniwa is described in more detail in § 7.2.4.2.

This classifier is also commonly used for sides or halves of animal bodies (e.g., *dáapa* 'paca (rodent sp.)', *néeri* 'deer', *kóphe* 'fish', *phíitsi* 'agouti (rodent sp.)', *héema* 'tapir'), or of other kinds of entities (e.g., *máawiro* 'pineapple'). The entity must be cut lengthwise, so that the halves are symmetrical. For halves of entities cut widthwise, see *-íida* 'CLF.HALF' (§ 7.1.21) and *-hipáda* 'CLF.PIECE' (§ 7.1.16).

7.1.12 *-éekhe* 'CLF.SMALL.SEED'

The classifier *-éekhe* is used only with small seeds, in particular those of chili peppers. It appears to be going out of use, as it is only recognized by some Baniwa speakers. Seeds in general are usually referred to by *-íixi* 'CLF.SEED' (§ 7.1.25).

7.1.13 *-éewhe* 'CLF.EGG'

The classifier *-éewhe* is used exclusively for eggs of different animals (e.g., *kalakéewhe* 'hen's egg', *iikolíewhe* 'turtle's egg', *ainíewhe* 'wasp's egg'). Eggs are also commonly referred to by the generic classifier *-da* (§ 7.1.8).

7.1.14 *-híko* 'CLF.LONG'

The classifier *-híko* is used for long and mostly rigid things, such as trees (especially palm trees, e.g., *ponáma* 'patawa', *íitewi* 'moriche palm', *máawi* '*jupati*', *píipiri* 'peach palm', and *manákhe* 'açaí', but also other kind of trees, e.g., *akáyo* 'cashew' and *dóomali* 'umari', as well as the generic *háiko* 'tree'), other long and rigid plants (*póapoa* '*arumã* (plant used for basketry)', *máapa* 'sugar cane'), and bottles (*garáapha*).

This classifier can be used with a few mammals (*néeri* 'deer', *tsíino* 'dog') and fish species (*íiniri* '*traíra*', *táali* '*aracu*').

7.1.15 *-hípa* 'CLF.MALE'

The classifier *-hípa* is used exclusively with male human referents, e.g., *atsíanli* 'man', *-eenípe* 'son', *walhípali* 'young man', and *pedália* '(male) elder'. Any referent of *-hípa* can also be referred to by *-iíta* 'CLF.HUMAN', although *-iíta* also includes other meanings besides male humans.

This classifier is not used for male animals, which are classified by shape rather than sex within the classifier system (cf. § 3.2.5.4 on grammatical gender).

7.1.16 *-hipáda* 'CLF.PIECE'

The classifier *-hipáda* is used for detached pieces of any size of all kinds of entities, including animals (*kóphe* 'fish', *dáapa* 'paca (rodent sp.)'), plants (*háiko* 'stick/piece of wood', *máapa* 'sugar cane'), food (*palána* 'banana', *peéthe* '*beiju*'), and inanimate objects (*hinoorótti* 'hammock rope', *ttáawaali* 'thread', [*sabão*] 'soap').

This classifier is very close in meaning to *-iída* 'CLF.HALF' (§ 7.1.21) and *-ttáwalhe* 'CLF.CUT' (§ 7.1.44), and to some extent also to *-wána* 'CLF.SLICE' (§ 7.1.48).

7.1.17 *-hípani* 'CLF.RAPIDS'

This is a highly specific classifier, used only for rapids (*híipa*).

7.1.18 *-híwa* 'CLF.BEIJU'

The classifier *-híwa* is only used with a single noun, *peéthe* '*beiju* (manioc bread)'.

7.1.19 *-híwi* 'CLF.POINTED'

The classifier *-híwi* refers to pointed, long, and thin objects, e.g., *áawi* 'needle', *iítsa* 'fish hook', [*lápis*] 'pencil', [*caneta*] 'pen', *kaapáwi* 'arrow', *kapitsíri* 'blow-pipe arrow', *haikhíwi* 'stick, twig', and *-íidza* 'strand of fur/hair'.

7.1.20 *-i* 'CLF.BASKET'

The classifier *-i* can be used for certain kinds of baskets, especially *tshéeto* '*aturá*', a large basket for carrying manioc, but also *mokóto* '*panacu*', a simple container made of braided leaves. It is also used for bunches of fruit from certain palm trees where the fruits grow in clusters, in particular *póoperi* '*bacaba*', *ponáma* 'patawa', *píipiri* 'peach palm', and *manákhe* 'açaí'. The semantic association may stem from fruit bunches being carried in baskets (see *-póko* 'CLF.CIRCLE' for a similar situation).

For bunches of palm trees without the fruit on, see *-iítsia* 'CLF.BUNCH' (§ 7.1.23).

7.1.21 *-íida* 'CLF.HALF'

The classifier *-iída* often refers to halves of animal bodies, especially when they are cut widthwise (e.g., *dáapa* 'paca (rodent sp.)', *kóphe* 'fish'). It can also refer to pieces or sections that are not halves, e.g., *iitsaákhaa* 'piece of a fishing line', *háiko* 'piece of wood', *-kiníki* 'part of a manioc garden', *pántti* 'part of a house (e.g., a room)'.

This classifier is highly similar to *-hipáda* 'CLF.PIECE' (§ 7.1.16) in meaning. For some speakers, the two classifiers appear to be used almost synonymously and interchangeably. For halves of animal bodies cut lengthwise, see *-eéma* 'CLF.SIDE' (§ 7.1.11).

7.1.22 *-iíta* 'CLF.HUMAN'

The classifier *-iíta* is used for humans, certain animals, and some inanimate objects with flat parts.

With human beings, it is used for males, like *atsíanli* 'man', or for persons not specified for gender, like *nawíki* 'person' (see § 7.5; see also the feminine vs. non-feminine split in the Baniwa gender system, § 3.2.5.4). With expressly female human referents, *-ma*^{2} 'CLF.FEMALE' (§ 7.1.32) is preferred. The classifier *-hípa* 'CLF.MALE' (§ 7.1.15) is even more strictly connected to male human referents.

With animals, it is used for monkeys (*pówe* 'monkey (generic)', *íitsi* 'howler monkey', *káaparo* 'woolly monkey'), for certain fish (*dzáapa* 'peacock bass', *wéemai* '*peixe-cachorro*', *ñámaro* 'ray', *ttéephe* '*piaba*'), for insects and spiders (*ttowída* 'louse', *kóopali* 'tick', *éeni* 'spider (generic)'), and for some other small animals (*páitsi* 'frog sp.', *íikoli* 'turtle').

It is also used for inanimate objects and instruments that are either flat or have a salient flat part (*ttiíwe* 'paddle', *matshéeta* 'machete', *peéthe* '*beiju*', *haikoíta* 'wooden board', *malíye* 'knife', *-thiíta* 'spectacles, glasses', [*celular*] 'cell phone').

The meaning of *-iíta* is difficult to subsume under a single label, and it is not known how the disparate meanings have developed. It goes back to a reconstructed human classifier **-iita* in Proto-Japurá-Colombia (Ramirez (2020d: 97); see also § 8.1), so the use with human referents appears to be the oldest. Hill (1988: 68; 2009a: 22, 117) suggests that the placement of most fish species in this category is rooted in myths where fish appear as phallic symbols, thus strongly connected to male humans (see § 7.3.3.3).

7.1.23 *-iítsia* 'CLF.BUNCH'

The classifier *-iítsia* is used to refer to bunches, especially from palm trees, where the fruit has already been picked from the cluster. The species for which this classifier can be used are, e.g., *ponáma* 'patawa', *manákhe* 'açaí', and *kamhéro* 'Amazon grape'. They are largely the same as those used with the classifier *-i* 'CLF.BASKET' (§ 7.1.20) which, in one of its functions, refers to the same kinds of bunches, but with the fruit still on them. It appears that some speakers can use *-iítsia* for bunches both with and without fruit (see § 7.4).

7.1.24 *-íiwi* 'CLF.FLOWER'

The classifier *-íiwi* is used primarily for flowers (*haikóiwi*), but also for thorns (*dóowiri*).

7.1.25 *-íixi* 'CLF.SEED'

The classifier *-íixi* is used for seeds and kernels of various kinds (e.g., *manakhéexi* 'açaí seed', *piiridzéexi* 'avocado seed'), and by extension for small pieces of stone (*hiipáda*). For small seeds, especially chili pepper seeds, *-éekhe* 'CLF.SMALL.SEED' (§ 7.1.12) may be used instead.

7.1.26 *-kénaa* 'CLF.BRANCH'

The classifier *-kénaa* is used only for tree branches, either generically (*haikóke* 'tree branch'), or of specific trees (e.g., *irímawa* 'lemon').

7.1.27 *-khaa* 'CLF.CURVILINEAR'

The classifier *-khaa* is used for all kinds of long, thin, and flexible referents with a salient extension in one dimension. This includes animals (e.g., *áapi* 'snake', *oomápi* 'worm', *mhóokoli* 'catfish sp.', *kolíri* 'catfish sp.', *iitsítsi* 'eel sp.', *katshíri* 'caiman'), plant parts (e.g., *ádapi* 'liana, vine', *-pali* 'root'), body parts (e.g., *-iittípi* 'tail', *-tarapíte* 'vein, artery', *-tsíkole* 'strand of hair') and inanimate objects (e.g., *ttáawaali* 'thread', *iitsaákhaa* 'fishing line', *hinoorótti* 'hammock rope'). It can also be used with paths, rivers and other kinds of passages that are curvilinear in form (e.g., *inípo* 'path', *óoni* 'river', *ñawápo* 'creek').

This classifier has abstract manifestations (partly shared with *-péko* 'CLF.PATH', § 7.1.38) where it is used for concepts relating to language, cognition, and emotions, possibly through a perceived notion of linearity. For example, it can be used with *iakótti* 'voice, language, word', *-iipítana* 'name', *-nakoapanínaa* 'myth, story', *-wapiñeetákhe* 'thought, idea', *iaanheekhétti* 'feeling, emotion; knowledge', [*canção*] 'song', *hiepakhétti* 'faith, belief, religion', *ikaakhétti* 'laughter, smile', *ixáatti* 'lie', and *iñapakátti* 'prayer' (see further in § 7.2.2.7). Another abstract sense of *-khaa* is that of lineages, e.g., *-newikíka* 'tribe, clan'.

7.1.28 *-ko* 'CLF.HAMMOCK'

This is a highly specific classifier used only with the noun *piéta* 'hammock'.

7.1.29 *-koa* 'CLF.SURFACE'

The classifier *-koa* is used for flat, two-dimensional entities at various scales. This includes objects (e.g., *póali* '*forno* (flat oven used for making *farinha* and *beiju*)', *méedza* 'table', *kánaali* 'mirror', *áada* 'manioc grater', *yóora* '*jirau* (type of shelf)', *-rhoakaróda* 'bed') as well as locations and landscape features (e.g., *dzakálee* 'village, town', *káida* 'beach', *hipákoa* 'flat rock').

7.1.30 *-kódzoa* 'CLF.BEND'

The classifier *-kódzoa* is used for bends, in particular river bends, but also bends of, e.g., a rope.

7.1.31 *-ma[1]* 'CLF.PAIR'

This classifier is typically used for entities that come in pairs, such as *-heenittáda* 'earring' and *tshapáto* 'shoe'. In extension, it can be used for the palm tree *máawi* '*jupati*', possibly because it is used to craft blowpipes, which consist of pairs of shafts (Ramirez, 2001a: 202), as well as ceremonial flutes that consist of pairs of tubes (Hill, 2009a: 96–97, 174).

For many paired objects, *-ma[1]* and *-eéma* 'CLF.SIDE' (§ 7.1.11) can be used in reference to the pair (110a) vs. one of the pair (110b)[73].

(110) a. *apáma* [*sandália*]
apa-**ma[1]** [sandália]
one-**CLF.PAIR** sandal

'one pair of sandals' [200318_FREE_G_01]

b. *apeéma* [*sandália*]
apa-**eéma** [sandália]
one-**CLF.SIDE** sandal

'one sandal (i.e., one half of a pair of sandals)' [200318_FREE_G_01]

[73] This use of *-ma[1]* does not apply to paired body parts—see § 7.2.4.2.

In addition to the pair sense, *-ma*[1] can also be used for entities that consist of several similar objects in parallel rows, such as combs of bananas (*palána*) and strips of some material lined up for braiding. It is also used for some objects that contain saliently parallel parts, e.g., *pheelóma* 'type of pan flute' and *yóora* '*jirau* (type of shelf with a horizontal plane consisting of parallel pieces of wood)'.

7.1.32 *-ma*[2] 'CLF.FEMALE'

This classifier is used exclusively with expressly female human referents such as *íinaro* 'woman', *ideenhikáma* '(female) worker', and *iokokáma* 'seamstress'. The complex form of this classifier is rarely used, being largely replaced by the generic/feminine combination *-dá-ro*. For human referents not specified for gender, *-iíta* 'CLF.HUMAN' (§ 7.1.22) is preferred.

The classifier *-ma*[2] is not used for female animals, as these are classified by shape rather than sex within the classifier system (cf. § 3.2.5.4 on grammatical gender).

7.1.33 *-máka* 'CLF.FABRIC'

The classifier *-máka* is used with all kinds of textiles and other woven, flexible, two-dimensional objects, such as *yamakátti* 'cloth', *-palhéeda* 'cover', *waxímaka* 'hand net (for fishing)', and *koyáma* 'type of straw mat'. Its use covers many Portuguese loanwords, e.g., [*toalha*] 'towel', [*malhadeira*] 'fishnet', and [*lençol*] 'sheet'. In a non-fabric sense, *-máka* is also part of the derived body part noun *-thi-máka* /eye-CLF.FABRIC/ 'eyelids'.

7.1.34 *-na* 'CLF.TRUNK'

The classifier *-na* refers to solid, three-dimensional, cylindrical referents. It is used for both animals and inanimate objects.

This classifier is used for trees (e.g., *háiko* 'tree', *-nánaa* 'tree trunk'), some mammals (e.g., *tsíino* 'dog', *héema* 'tapir', *dzáawi* 'jaguar', *pitsána* 'cat', *néeri* 'deer', *pówe* 'monkey (generic)', *wáamo* 'sloth', *aapídza* 'white-lipped peccary', *kéetto* 'capybara'), some fish species (e.g., *táali* '*aracu*', *íiniri* '*traíra*', *kettínali* '*jacundá*', *dzáapa* 'peacock bass'), and trunk-shaped, upright objects (e.g., [*lata*] 'can for measuring manioc flour', *garáapha* 'bottle', *kaxadádali* 'type of curved basket', *wepóne* 'quiver', *pántthii* 'vertical pillar of a house', *tsháako* 'bag', *korótsha* 'cross').

7.1.35 *-náko* 'CLF.BUNDLE'

The classifier *-náko* refers to bundles of long and thin objects, such as *máapa* 'sugar cane', *haikóapo* 'sticks', *hawádza* '*ingá* pod', *ttídzee* 'firewood', and *tsikólee* 'hair'.

7.1.36 *-pa* 'CLF.PACKAGE'

The classifier *-pa* is used for packages of various kinds, such as *hiemakalepátti* 'fish wrapped in a leaf for cooking', *mokóto* '*panacu*', and *tsháako* 'bag'. It can also be used with nouns referring to the contents of packages (e.g., *yokíra* 'salt', *káaphe* 'coffee').

7.1.37 *-páwa* 'CLF.RIVER'

The classifier *-páwa* is highly specific, used only for watercourses such as *óoni* 'river', *ñawápo* 'creek', and *-ke* 'tributary'.

7.1.38 *-péko* 'CLF.PATH'

The classifier *-péko* is used for passages and courses (e.g., *inípo* 'path', *óoni* 'river', *ñawápo* 'creek'), as well as openings (e.g., *panttinóma* 'door', *-heewakaropéki* 'entrance', *iniphiwída* 'harbour', *-nóma* 'mouth', *-nómawa* 'hole, cave').

This classifier also covers abstract senses relating to language, cognition and emotions (e.g., *iakótti* 'voice, language, word', *kattimákali* 'happiness, joy', *ixáatti* 'lie'), in a similar fashion to the classifier *-khaa* 'CLF.CURVILINEAR' (see § 7.2.2.7).

7.1.39 *-phe* 'CLF.LEAF'

The classifier *-phe* is associated with thin, flat, two-dimensional entities. It is used for leaves (e.g., *panáphe* 'leaf (of a tree)', *manakhéphe* 'açaí leaf', *palanáphe* 'banana leaf'), for feathers (*-phe*), and for various kinds of entities that come in sheets (*papéra* 'paper', *yamakátti* 'cloth', *paláata* 'money (here: banknote)').

7.1.40 *-pi* 'CLF.TUBE'

The classifier *-pi* refers to long, thin, cylindrical, hollow, and tube-shaped objects such as *mókawa* 'rifle, gun', *mawípi* 'blowpipe', and *ttirolípi* '*tipití* (manioc squeezer)'. It can also refer to stems of trees (*háiko* 'tree stem') especially those of palm trees, e.g., *manakhépi* 'açaí stem', *pooperípi* '*bacaba* stem', *ponamápi* 'patawa stem'. This classifier is also used to refer to the time unit of a month, *kéeri* (§ 7.2.2).

7.1.41 *-póko* 'CLF.CIRCLE'

The classifier *-póko* is used with several different types of referents. The core meaning of *-póko* is that of circular, flat, two-dimensional shapes. In this sense, it is used for rings (*-kaapíiro*), earrings (*-heenittáda*), bracelets (*-naapápi*), and hair ties (*-paraaxoópa itsíkole*), for wheels (*róoda*), and by extension for bicycles ([*bicicleta*]). This classifier is also used for bird's nests (*-éewhe iaróda*), for hand nets for fishing (*waxímaka*), and for flat woven sieves of the types called *dopítsi* and *waláya* (both used in manioc processing), which carry both circular and container semantics.

In another sense, *-póko* is used for bundles of some fruits (*kamhéro* 'Amazon grape' and *palána* 'banana') or small game animals (e.g., *kóphe* 'fish', *dáapa* 'paca (rodent sp.)', and *phíitsi* 'agouti (rodent sp.)'). Some speakers use it for groups of animate referents that are not in bundles, e.g., schools of fish (*kóphe*) or groups of humans (*nawíki*).

This classifier also has a function where it is used for parts of entities. In this sense, it is used for quarters of bodies of relatively small game animals, e.g., pacas (*dáapa*), agoutis (*phíitsi*), and deer (*néeri*).

The classifier *-póko* is etymologically related to *-pokóda* 'CLF.STUMP' (§ 7.1.42). This connection is visible in the way some speakers can use *-póko* to refer to tree stumps. The semantic connection is likely in the flat, circular surface of a tree stump.

7.1.42 *-pokóda* 'CLF.STUMP'

The classifier *-pokóda* refers to tree stumps and trunks, e.g., *haikopokóda* 'tree trunk' and *-tanhída* 'tree stump'.

7.1.43 *-tsoi* 'CLF.PILE'

The classifier *-tsoi* refers to piles of objects, e.g., *káini* 'manioc' and *hiipáda* 'stone'.

7.1.44 *-ttáwalhe* 'CLF.CUT'

The classifier *-ttáwalhe* refers to cut off or separated pieces of various entities. It is used for pieces of plants (e.g., *háiko* 'tree/wood', *máapa* 'sugar cane'), pieces of animal bodies (e.g., *kóphe* 'fish', *omáwali* 'anaconda', *dáapa* 'paca (rodent sp.)'), pieces of landscape entities (e.g., *híipai* 'land', *óoni* 'river'), or pieces of objects (e.g., [*sabão*] 'soap'). Some speakers use it interchangeably with *-ttówhia* 'CLF.ROOM' (§ 7.1.45) to refer to rooms of houses (e.g., *-imaakaróda* 'bedroom', [*quarto*] 'room').

Semantically, *-ttáwalhe* is highly similar to *-hipáda* 'CLF.PIECE' (§ 7.1.16), and to some extent to *-iída* 'CLF.HALF' (§ 7.1.21) and *-wána* 'CLF.SLICE' (§ 7.1.48).

7.1.45 *-ttówhia* 'CLF.ROOM'

The classifier *-ttówhia* is used only for rooms of houses. While rooms as such are not a very old concept in the region, a *maloca* (longhouse), the traditional dwelling, is typically divided into compartments, and this may have been the original use of this classifier in the Indigenous context (as speculated by Aikhenvald 2000: 344). This classifier often occurs without an overt noun, but it can also associate with the noun [*quarto*] 'room', borrowed from Portuguese.

7.1.46 *-wa* 'CLF.SPACE'

The classifier *-wa* is used with referents that have either a convex, concave, or flat shape. It refers to holes and openings (e.g., *-nómawa* 'hole', *haláaphi* 'buraco', *-nóma* 'mouth', *-eeríko* 'anus', *dóopo inómawa* 'entrance of a lizard's nest', *okéna* 'door'), to certain locations, areas and spaces (e.g., *-heepídaa* 'slope', *dzakálee* 'village, town'), and to piles of objects (e.g., *híiniri* '*ucuquí* fruit', *dóomali* 'umari fruit', *líxo* 'rubbish').

This classifier is quite similar to *-yáwa* 'CLF.HOLE' (§ 7.1.53) in meaning. The two classifiers appear to be used almost interchangeably in many contexts, and may share an etymological link (see § 8.3.10).

7.1.47 *-wálhia* 'CLF.YEAR'

The classifier *-wálhia* is used only for yearly cycles (*hámoli* 'dry season; year').

7.1.48 *-wána* 'CLF.SLICE'

The classifier *-wána* is used for slices of food items (e.g., *peéthe* '*beiju*', [*pizza*] 'pizza', *míitsi* 'barbecued meat', [*bolo*] 'cake'), for shards of broken, hard objects (e.g., *kánaali* 'mirror', *paráto* 'plate', *-eétsha* 'tooth', *déekai* 'ceramic object'), for sections of land (e.g., *híipai* 'land', *kiníki* 'manioc garden'), and for pieces of animal bodies (e.g., *dáapa* 'paca (rodent sp.)', *phíitsi* 'agouti (rodent sp.)', *héema* 'tapir').

7.1.49 *-wáta* 'CLF.BRACT'

The classifier *-wáta* is used for bracts (known as *curuatá* in local Portuguese), a botanical term for a certain type of curled leaf that grows around flower bases of palm trees like *manákhe* 'açaí', *ponáma* 'patawa', and *póoperi* '*bacaba*'.

In an extended sense, *-wáta* can be used for thin, curled up objects (e.g., *haikóya* 'tree bark (here: fallen and rolled up)', *papéra* 'paper (here: folded into the shape of a funnel)'), as well as for different types of calabash bowls (*kóoya*, *átta*).

7.1.50 *-wáthe* 'CLF.NODE'

The classifier *-wáthe* is used for various kinds of nodes, knots, and joints, e.g., *máapa iwáthe* 'joint of a sugar cane', *iitshaákhaa iwáthe* 'knot on a fishing line', *ttáawaali iwáthe* 'knot on a thread', and *-nawathére* 'elbow'.

7.1.51 *-xaa* 'CLF.EXCREMENT'

The classifier *-xaa* is used exclusively for faeces, particularly that of various animals (e.g., *tsiinóxaa* 'dog excrement', *heemáxaa* 'tapir excrement'), but also of humans.

7.1.52 *-ya* 'CLF.SKIN'

The classifier *-ya* is used mainly for tree bark and animal skin (*-ya* 'bark; skin'). It can also be used for *curuatás*, e.g., *manakhewáta* '*curuatá* of açaí', which bear resemblance to bark in being thin, flat plant parts (see § 7.1.49).

Some speakers can use *-ya* to refer to canoes (*íita*), while for others this use is impossible (see § 7.4)—in fact, somewhat surprisingly, *íita* 'canoe' emerges as the most prototypical member of this classifier category in the Free listing data set (see § 7.3.1; Appendix B). There appears to be an etymological link between *-ya* and *-Ø* which is described in § 8.3.1.

7.1.53 *-yáwa* 'CLF.HOLE'

The classifier *-yáwa* is used for holes and openings (e.g., *-nomáwa* 'hole', *pant-tinóma* 'door', *kawawalikoáaphi* 'hole in the ground', *-dzáanaa* 'wound', *-aapówa* 'burrow, den (of an animal)', *kawípe inómawa* 'entrance of an ant's nest').

Some speakers can use *-yáwa* to refer to skin and bark (*-ya*), but this appears to be uncommon. However, it does raise the question of whether *-yáwa* could etymologically be a compound classifier, consisting of *-ya* 'CLF.SKIN' and *-wa* 'CLF.SPACE' (see § 8.3.10; also discussed by Ramirez 2020a: 165). The classifiers *-yáwa* and *-wa* appear to be used almost interchangeably in many contexts.

7.2 Semantic parameters

The previous section (§ 7.1) described in detail the semantic properties of each individual classifier. This section aims to generalize the semantic organization of the classifier system by describing the semantic parameters that are fundamental to the workings of the classifier system: animacy (§ 7.2.1), physical properties (§ 7.2.2), collections (§ 7.2.3) and parts (§ 7.2.4).

7.2.1 Animacy

As is common in multifunctional classifier systems of Amazonia (§ 2.7.1), animacy is a fundamental parameter in the classifier system of Baniwa. Table 7.1

shows the distribution of the three referent types humans, animals[74], and inanimates among the classifiers of Baniwa.

As table 7.1 shows, there are no examples of classifiers that can be used with human and inanimate referents, but not with animals. Thus, it seems to be the case that animal referents occupy an intermediate position between humans and inanimates in the Baniwa classifier system. The sequence humans—animals—inanimates follows the sequence in Corbett's (2000) Animacy Hierachy, although the predictions made by the hierachy are different and have to do with grammatical number.

7.2.1.1 Human referents

Five classifiers can be used with human referents (Table 7.1). Two of the classifiers are used exclusively for humans: *-hípa* 'CLF.MALE' for males and *-ma²* 'CLF.FEMALE' for females. The remaining three can be used for both humans, animals and inanimates. The classifier *-iíta* 'CLF.HUMAN' is used for any human referent unless it is expressly female, as the referent will typically be interpreted as male. This classifier is also used for certain animals (see § 7.2.1.2 below) and for some objects with flat parts. The generic classifier *-da* is used for preadolescent children, but it functions primarily as a generic classifier (§ 7.3.3.1) and also carries connotations of roundness. The group classifier *-áana* is used for groups of anything: people, animals, or inanimate objects.

In other words, within the classifier system, humans are differentiated by age (adult vs. child), and adult humans are further differentiated by sex (female vs. male). When referring to groups of people, all of these distinctions collapse.

Within the Arawakan family, classifiers expressing biological sex are almost exclusively found in the languages spoken in Northwestern Amazonia (Dunn & Rose, forthcoming: 17). This trait is also found in the neighboring East Tucanoan languages, although there it applies to animates referents in general, and not just humans (Barnes, 1990; Chacon, 2007; Gomez-Imbert, 2007; Stenzel, 2013).

[74]The animal category here is restricted to living animals, not to (parts or collections of) dead animals.

Table 7.1: Classifier use according to animacy

Classifier	Gloss	Humans	Animals	Inanimates
-hípa	CLF.MALE	✓	–	–
-ma²	CLF.FEMALE	✓	–	–
-áana	CLF.GROUP	✓	✓	✓
-da	CLF.GENERIC	✓	✓	✓
-iíta	CLF.HUMAN	✓	✓	✓
-aápa	CLF.OBLONG	–	✓	✓
-híko	CLF.LONG	–	✓	✓
-khaa	CLF.CURVILINEAR	–	✓	✓
-na	CLF.TRUNK	–	✓	✓
-phe	CLF.LEAF	–	✓	✓
-Ø	CLF.CANOE	–	–	✓
-áanhaa	CLF.LIQUID	–	–	✓
-áaphi	CLF.AREA	–	–	✓
-áapi	CLF.HOLLOW	–	–	✓
-áapo	CLF.STICK	–	–	✓
-daa	CLF.DAY	–	–	✓
-dápana	CLF.HOUSE	–	–	✓
-eéma	CLF.SIDE	–	–	✓
-éekhe	CLF.SMALL.SEED	–	–	✓
-éewhe	CLF.EGG	–	–	✓
-hipáda	CLF.PIECE	–	–	✓
-hípani	CLF.RAPIDS	–	–	✓
-híwa	CLF.BEIJU	–	–	✓
-híwi	CLF.POINTED	–	–	✓
-i	CLF.BASKET	–	–	✓
-íida	CLF.HALF	–	–	✓
-iítsia	CLF.BUNCH	–	–	✓
-íiwi	CLF.FLOWER	–	–	✓
-íixi	CLF.SEED	–	–	✓
-kénaa	CLF.BRANCH	–	–	✓
-ko	CLF.HAMMOCK	–	–	✓
-koa	CLF.SURFACE	–	–	✓
-kódzoa	CLF.BEND	–	–	✓
-ma¹	CLF.PAIR	–	–	✓
-máka	CLF.FABRIC	–	–	✓
-náko	CLF.BUNDLE	–	–	✓
-pa	CLF.PACKAGE	–	–	✓
-páwa	CLF.RIVER	–	–	✓
-péko	CLF.PATH	–	–	✓
-pi	CLF.TUBE	–	–	✓
-póko	CLF.CIRCLE	–	–	✓
-pokóda	CLF.STUMP	–	–	✓
-tsoi	CLF.PILE	–	–	✓
-ttáwalhe	CLF.CUT	–	–	✓
-ttówhia	CLF.ROOM	–	–	✓
-wa	CLF.SPACE	–	–	✓
-wálhia	CLF.YEAR	–	–	✓
-wána	CLF.SLICE	–	–	✓
-wáta	CLF.BRACT	–	–	✓
-wáthe	CLF.NODE	–	–	✓
-xaa	CLF.EXCREMENT	–	–	✓
-ya	CLF.SKIN	–	–	✓
-yáwa	CLF.HOLE	–	–	✓

7.2.1.2 Animal referents

In Baniwa, (non-human) animals are classified by shape[75], as noted by Gomez-Imbert (1996) and Ramirez (2020a: 233–234). In the singular, living individual animals are primarily referred by one of the classifiers *-aápa* 'CLF.OBLONG', *-da* 'CLF.GENERIC', *-iíta* 'CLF.HUMAN', *-na* 'CLF.TRUNK', or *-khaa* 'CLF.CURVILINEAR'[76]. Three of these classifiers (*-áana*, *-da*, and *-iíta*) are also used with both humans and inanimates. The other three (*-aápa*, *-khaa*, and *-na*) can be used with animals and inanimates, but not with humans.

The classifier *-da* is used for some mammals (*dáapa* 'paca (rodent sp.)'), some fish (*ómai* 'piranha', *tsíipa* '*pacu*'), some reptiles (*itsída* 'tortoise'), some amphibians (*híiparo* 'frog sp.'), and some birds (*wáaro* 'parrot'). The animals are all round(ish) and relatively small.

The classifier *-iíta* refers to a number of disparate types of animals, namely monkeys (*pówe* 'monkey (generic)', *íitsi* 'howler monkey', *káaparo* 'woolly monkey'), insects and spiders (*poléta* 'cockroach', *éeni* 'spider (generic)'), some fish species (*dzáapa* 'peacock bass', *ñámaro* 'ray', *ttéephe* '*piaba*'), some reptiles (*páitsi* 'frog sp.') and some amphibians (*íikoli* 'turtle'). Most of the animals in this group may be perceived as somewhat flat, although in the case of monkeys, it may also be their similarity to humans that places them in this group (*-iíta* is also used for humans, see § 7.2.1.1).

The classifier *-na* is used with animals that are relatively large and trunk-like in shape. It is mostly used with mammals (*tsíino* 'dog', *héema* 'tapir', *dzáawi* 'jaguar', *pitsána* 'cat', *néeri* 'deer', *pówe* 'monkey (generic)', *aapídza* 'white-lipped peccary', *kéetto* 'capybara', *wáamo* 'sloth'), but also with some fish species (*táali* '*aracu*', *íiniri* '*traíra*', *kettínali* '*jacundá*', *dzáapa* 'peacock bass').

The classifier *-khaa* is used for any animal that is long, relatively thin and curvilinear in shape. This includes snakes and other reptiles (*áapi* 'snake (generic)', *katshíri* 'caiman', *dóopo* 'lizard'), worms and centipedes (*oomápi* 'worm', *khápi* 'centipede'), as well as some fish (*mhóokoli* 'catfish sp.', *kolíri* 'catfish sp.', *iitsítsi* 'eel sp.').

The classifier *-aápa* is used with most birds (*kepíreeni* 'small bird (generic)', *kóitsi* 'curassow', *kaláka* 'hen/rooster'), but also with some mammals (*dzáawi* 'ocelot'),

[75] Animals are not sex-differentiated within the classifier system, but they can be through the grammatical gender system (§ 3.2.5.4), which also surfaces in the complex form of the generic classifier (§ 5.3.2).

[76] Classification of individual animals also occurs with *-híko* 'CLF.LONG' and *-phe* 'CLF.LEAF', but only for one or very few species.

including flying ones (*píittiri* 'bat'), some fish species (*kettínali* '*jacundá*', *táali* '*aracu*'), some insects (*kawípe* 'ant sp.'), and some crustaceans (*dzáaka* 'shrimp'). The referents all have a somewhat oblong shape.

The five main classifiers that are used for animals all have a broader reference that operates at least partly on shape distinctions (see § 7.2.2 below) and includes both animal and non-animal referents. There are no indications that animals would be categorized according to biological taxonomy (mammals are found in the *-da*, *-iíta*, *-na*, and *-aápa* groups), habitat (aquatic animals are found in all five groups), edibility (edible species are found in all five groups), or harmfulness (animals that can cause harm are found in all groups). In addition, there are plenty of examples of animal species that can be referred to by more than one classifier, e.g., *íikoli* 'turtle' (*-iíta* or *-da*) and *dzáapa* 'fish sp.' (*-íita* or *-na*). This variation may depend on both inter-speaker preferences and the appearance of individual animals.

Classification of animals by shape is common across the Arawakan language family (Dunn & Rose, forthcoming: 17). By contrast, in the neighboring East Tucanoan languages, animals are classified as animates and not by shape (Tariana has also adopted this practice; Aikhenvald, 2002: 88–92). An exception is Kubeo, which has aligned its classification of animals with that of Baniwa (Gomez-Imbert, 1996). Both of these contact situations are described in § 2.7.2.

7.2.1.3 Inanimate referents

Fifty-one classifiers, all but *-hípa* 'CLF.MALE' and *-ma²* 'CLF.FEMALE', can be used with inanimate referents. Five of these can also be used with animals, and an additional three with animals and humans (see Table 7.1 above). However, the great majority of Baniwa classifiers are associated exclusively with inanimate referents, which are typically classified according to their physical properties (§ 7.2.2), or whether they occur as a part or a collection of something (§ 7.2.3).

7.2.2 Physical properties

Physical properties, in particular shape, are encoded in many Baniwa classifiers. Both inanimate objects and animals are classified by their physical shape. The most relevant physical parameters are summarized in Table 7.2 and described below: dimensionality (§ 7.2.2.1), flexibility (§ 7.2.2.2), curvature (§ 7.2.2.3), hollowness (§ 7.2.2.4), pointedness (§ 7.2.2.5), and texture (§ 7.2.2.6).

Note that the classifier descriptions in this section may not cover the entire ex-

Table 7.2: Examples of classifiers denoting physical properties

Classifier	Gloss	Associated shape	*Flexibility*	*Curvature*	*Hollowness*	*Pointedness*	*Texture*
1D referents							
-áapo	CLF.STICK	Stick-shaped referents					
-híko	CLF.LONG	Long & thin referents					
-híwi	CLF.POINTED	Long, thin & pointed referents				✓	
-khaa	CLF.CURVILINEAR	Long, thin & flexible referents	✓	(✓)			
-péko	CLF.PATH	Paths & passages		✓			
-pi	CLF.TUBE	Tubes			✓		
2D referents							
-áaphi	CLF.AREA	Areas & locations					
-iíta	CLF.HUMAN	Referents with flat parts					
-koa	CLF.SURFACE	Surfaces					
-máka	CLF.FABRIC	Fabrics	✓	✓			✓
-phe	CLF.LEAF	Flat referents	(✓)				
-póko	CLF.CIRCLE	Flat & round referents		✓			
-wa	CLF.SPACE	Holes, openings & piles		✓			
-ya	CLF.SKIN	Skin and bark	(✓)				
-yáwa	CLF.HOLE	Holes & openings		✓			
3D referents							
-Ø	CLF.CANOE	Hollow referents			✓		
-aápa	CLF.OBLONG	Oblong referents					
-áapi	CLF.HOLLOW	Hollow referents			✓		
-da	CLF.GENERIC	Round referents					
-na	CLF.TRUNK	Trunk-shaped referents					
Not saliently dimensional							
-áanhaa	CLF.LIQUID	Liquid referents					✓

tension of the individual classifier; many of them denote both physical properties and something else (see § 7.2.1, § 7.2.3, § 7.2.4).

7.2.2.1 Dimensionality

One of the most central physical properties encoded in the classifier system is dimensionality. Many classifiers associate with referents with a saliently one-dimensional, two-dimensional, or three-dimensional extension.

The 1D category comprises classifiers that refer to stick-like objects (*-áapo*); long and thin things (*-híko*); long, thin and pointed things (*-híwi*); long, thin and flex-

ible things (*-khaa*); and tubes (*-pi*). The path classifier *-péko* is also placed in this category as it only emphasizes extendedness in a single dimension.

The 2D category contains classifiers referring to surfaces (*-koa*); bounded areas (*-áaphi*); flat objects (*-phe*): flat and round objects (*-póko*); fabrics (*-máka*); skin and bark (*-ya*); objects with saliently flat parts (*-iíta*); concave objects (*-yáwa*); and concave and convex objects alike (*-wa*).

The 3D category features classifiers that are used with oblong objects (*-aápa*); trunk-shaped objects (*-na*); round objects (*-da*); round and small objects (*-íixi*); and hollow objects (*-Ø, -áapi*).

Within the three dimensions (1D, 2D, and 3D), a number of other parameters play a part in making further distinctions. They are described below.

7.2.2.2 Flexibility

Flexible 1D referents associate with the curvilinear classifier *-khaa*. This stands in contrast to most of the other 1D classifiers, which usually denote straight and rigid objects, e.g., *-híko* 'CLF.LONG'.

For 2D objects, it is primarily the fabric classifier *-máka* that denotes flexibility (although this may also happen with some referents of the leaf classifier *-phe* and the skin and bark classifier *-ya*). In contrast, the surface classifier *-koa* generally denotes rigid 2D referents.

7.2.2.3 Curvature

Some 2D referents are distinguished by curvature. For example, *-póko* 'CLF.CIRCLE' is used for 2D referents with a curved edge (i.e., flat and round), like wheels. Flat referents without a curved edge (like tables and mirrors) are typically assigned the surface classifier *-koa*[77].

2D objects can also be distinguished by curvature in another sense, namely by concavity and convexity. The hole classifier *-yáwa* is generally used for concave shapes, while the space classifier *-wa* can be used for both concave and convex

[77] This classifier also encompasses some referents with curved edges: *póali* '*forno*', a large, flat and round oven for roasting manioc produce, is generally used with *-koa*. By contrast, *-póko* is never used for referents that are flat and square, so the the presence of curvature appears to be a more marked feature.

shapes. By contrast, the surface classifier *-koa* generally denotes flat shapes without any curvature.

7.2.2.4 Hollowness

The tube classifier *-pi* associates with hollow 1D referents, while solid 1D referents take, e.g., *-híko* 'CLF.LONG'.

For 3D referents, hollowness in the three dimensions distinguishes hollow from solid objects. While *-aápa* 'CLF.OBLONG', *-da* 'CLF.GENERIC', and *-na* 'CLF.TRUNK' generally are used for solid referents, hollow objects are referred to by either *-Ø* 'CLF.CANOE' or *-áapi* 'CLF.HOLLOW'.

7.2.2.5 Pointedness

The classifier *-híwi* is used for long, thin, and pointed objects, in contrast to, e.g., *-híko*, which does not have pointed connotations.

7.2.2.6 Texture

Texture is a minor parameter in the classifier system of Baniwa. Only two classifiers denote textures: the liquid classifier *-áanhaa* and the fabric classifier *-máka*.

7.2.2.7 Abstract referents

As seen above, the classifier system in Baniwa makes extensive use of distinctions based on the physical shape of referents, but some referents are abstract and thus do not have a physical shape. In most cases, abstract referents will take the generic classifier *-da* (§ 7.3.3.1), e.g., *tápoli* 'dream' and *idzaamikátti* 'disease', but there are two notable exceptions: concepts relating to language and cognition, as well as time units.

The classifiers *-khaa* 'CLF.CURVILINEAR' and *-péko* 'CLF.PATH' are frequently used for abstract concepts relating to language and cognition, such as ideas, words, languages, songs, and prayers. Two examples are shown in (111). This semantic extension hints at a linear conception of linguistic and cognitive processes. Alternatively, as both *-khaa* and *-péko* are regularly used for rivers, it may be the case that language and cognition are viewed as streams or flows.

(111) a. *kadalíkhaa* *pianhéeka* *iákotti?*
kadali-**khaa** pi-áanhee-ka i-áako-tti
how.many-**CLF.CURV** 2SG-to.know-SUB AL-speech-AL

'How many languages do you know?' [221006_ELIC_G_01]

b. *apapéko* *ixáatti*
apa-**péko** i-xaa-tti
one-**CLF.PATH** AL-lie-AL

'one lie' [200326_ELIC_F_04]

Strikingly similar semantic extensions occur elsewhere in the Arawakan language family. In Mojeño Trinitario (Bolivia), the classifier *-pi* typically denotes long, thin, narrow, and flexible objects, but can also be used for speech, words, languages, songs, and music (Rose, 2024a: 56, 65–66). Rose suggests that this is a metaphorical extension to a linear process in the sound domain. Baure (Bolivia) also has the form *-pi*, analyzed by Danielsen (2007: 444–445) as two separate classifiers: one for long and thin objects, and another for words. Baniwa's closest relative Tariana, however, does not appear to have the same extension—instead, it has a dedicated classifier *-numa(na)* for words and language, which has developed from the bound noun *-numa(na)* 'mouth' (Aikhenvald, 2003: 92, 99).

Another subset of abstract referents is that of time units. Time units in Baniwa are associated with a number of different classifiers, as summarized in Table 7.3. The classifier *-daa* 'CLF.DAY' is used for days (*heekóapi*) and nights (*déepi*), *-áanhaa* 'CLF.LIQUID' is used for weeks (*míitsha*), *-pi* 'CLF.TUBE' is used for months (*kéeri*), and *-wálhia* 'CLF.YEAR' is used for years (*hámoli*)[78].

Table 7.3: Classifiers associated with time units

Classifier	Gloss	Associated time unit	Other use
-daa	CLF.DAY	Days, nights	–
-áanhaa	CLF.LIQUID	Weeks	Liquids
-pi	CLF.TUBE	Months	Tubes
-wálhia	CLF.YEAR	Years (dry seasons)	–

The classifier *-daa* is used for days and nights only. It is probably related to the Tariana classifier *-kada*, which is also used for both days and nights (Aikhenvald, 1994: 423–424, 455). Another specific classifier is *-wálhia*, used only for yearly

[78] Yearly cycles are generally counted in dry seasons: the noun *hámoli* means both 'dry season' and 'year'. The dry season extends roughly from September to March. *Oníyaa* 'rainy season' extends roughly from April to August.

cycles. This classifier does not appear to have any cognate classifier in Tariana, which uses a different classifier *-yapi* for years and possibly for other time-spans (Aikhenvald, 1994: 457).

The classifier *-pi* is primarily a shape classifier used for tubes, which has come to be associated with months. It has a cognate *-pi* in Tariana with a very similar extension: it is primarily used for long and thin things (e.g., *tipitís*, trees), but also for months (Aikhenvald, 2003: 90). It is thus likely that the temporal use goes back at least to Proto-Baniwa-Tariana. The classifier itself appears to be very old; it can be reconstructed to Proto-Arawak *-pi* as a classifier for long and thin objects (Dunn, 2022: 43–44; see also § 2.8.1)[79], and cognates abound in both closely and distantly related languages (Aikhenvald, 2019: 118; Mihas, 2019: 61; Ramirez, 2020d: 97). In Mojeño Trinitario and Baure, this classifier is used for abstract referents in the language and cognition domains (as described above), but also in the time domain: in Baure it is used for weeks (Danielsen, 2007: 13), while in Mojeño it associates with a more general temporal extension (Rose, 2024a: 65–66). For most of the other languages that have a reflex of this classifier, there is no information about any use with time units.

The classifier *-áanhaa* is primarily used for liquids, but is also used for weeks. This particular semantic extension is likely relatively recent, as the week is a concept introduced through Christianity[80]. One older speaker in this study also uses it with *hámoli* 'dry season; year', and according to Ramirez (2001a: 240), *oníyaa* 'rainy season' may take the liquid classifier[81]. It is possible that the use of *-áanhaa* for time units grew out of the strong association between the rainy season and water, through heavy rains and flooding of rivers (the word *oníyaa* is even derived from *óoni* 'water').

7.2.3 Collections

Some classifiers are used for groups of objects (e.g., groups, pairs), quantities of objects (e.g., bottlefuls, potfuls), or objects in certain kinds of collections (e.g., bundles, piles). A non-exhaustive list is given in Table 7.4.

Most of the classifiers in Table 7.4 are used for collections of inanimate objects. An exception is the group classifier *-áana*, which can be used for groups of inan-

[79]Dunn (2022: 53–54) also reconstructs a classifier ***chi* for months and moons for Proto-Arawak.

[80]The word for 'Sunday; week' in Baniwa is *míitsha*, borrowed from Portuguese *missa* 'mass' (Ramirez, 2001a: 220).

[81]In my data, *oníyaa* 'rainy season' is only used with the generic classifier *-da*.

Table 7.4: Examples of classifiers denoting collections

Classifier	Gloss	Associated collection
-áana	CLF.GROUP	Groups
-áapi	CLF.HOLLOW	Potfuls, bowlfuls
-i	CLF.BASKET	Bunches of fruit
-ma[1]	CLF.PAIR	Pairs
-na	CLF.TRUNK	Bottlefuls
-náko	CLF.BUNDLE	Bundles of long & thin objects
-pa	CLF.PACKAGE	Packages
-tsoi	CLF.PILE	Piles

imate objects (112a)[82], animals (112b), and humans (112c).

(112) a. *apáana* *nonáani*
apa-**áana** no-náani
one-**CLF.GROUP** 1SG-belongings

'my other things' [200324_ELIC_F_01]

b. *apáana* *pówe*
apa-**áana** pówe
one-**CLF.GROUP** monkey

'one group of monkeys' [200311_ELIC_K_06]

c. *apáana* *newíki*
apa-**áana** newíki
one-**CLF.GROUP** person

'one group of people; one clan' [200319_FREE_H_01]

Some of the classifiers in Table 7.4 are only used to denote collections, e.g., *-náko* 'CLF.BUNDLE' (113a) and *-tsoi* 'CLF.PILE' (113b).

(113) a. *apanáko* *ttidzéena*
apa-**náko** ttidzée-na
one-**CLF.BUNDLE** firewood-CLF.TRUNK

'one bundle of firewood' [200319_FREE_H_01]

[82] In this example, *apa-* 'one' is used in the sense 'other' (see § 3.2.9.1).

b. *apátsoi* *hiipáda* *makátsoi*
apa-**tsoi** hiipáda maka-**tsoi**
one-**CLF.PILE** stone big-**CLF.PILE**

'one big pile of stones' [220930_ELIC_K_01]

Other classifiers in Table 7.4 do not only denote collections. For example, *-na* is primarily a shape classifier denoting trunk-like shapes, but can be extended to denote bottlefuls (114a), due to the trunk-like shape of bottles. A similar extension occurs for *-áapi*, a shape classifier for hollow objects, which is also used for bowlfuls (114b).

(114) a. *aapána* *áatti*
apa-**na** áatti
one-**CLF.TRUNK** jiquitaia

'one bottleful of *jiquitaia* (chili pepper powder)' [200315_ELIC_F_01]

b. *apáapi* *manákhea*
apa-**áapi** manákhe-aa
one-**CLF.HOLLOW** açaí-liquid

'one bowlful of açaí juice' [200309_FREE_K_02]

7.2.4 Parts

Classifiers are sometimes used to indicate parts of objects. A non-exhaustive list can be found in Table 7.5.

Table 7.5: Examples of classifiers denoting parts

Classifier	Gloss	Associated part
-eéma	CLF.SIDE	Sides, halves (cut lengthwise), one of a pair
-hipáda	CLF.PIECE	Pieces
-íida	CLF.HALF	Halves (cut widthwise)
-kódzoa	CLF.BEND	Bends (of rivers)
-phe	CLF.LEAF	Leaves (of trees)
-ttáwalhe	CLF.CUT	Cut off or separated pieces
-wána	CLF.SLICE	Slices/shards of flat objects
-wáthe	CLF.NODE	Nodes (of, e.g., sugar canes)
-ya	CLF.SKIN	Skins (of animals), bark (of trees)

In many cases, part classifiers denote some generalized kind of piece, e.g., *-hipáda* 'CLF.PIECE' (115a), or *-wána* 'CLF.SLICE' (115b).

(115) a. *aphepáda* *palána*
apa-**hipáda** palána
one-**CLF.PIECE** banana

'one piece of a banana' [200309_FREE_K_02]

b. *apawána* *kánaali*
apa-**wána** kánaali
one-**CLF.SLICE** mirror

'one shard of a (broken) mirror' [200319_FREE_H_01]

For cut off halves of hunted animals, a distinction is made in the cutting direction: *-eéma* 'CLF.SIDE' is used for symmetrical halves of animal bodies cut from head to tail, while *-íida* 'CLF.HALF' is used for halves of animal bodies cut along the waist.

In other cases, a part classifier denotes something that is in an intrinsic part–whole relationship with something else. For instance, *-kódzoa* 'CLF.BEND' is used for bends of rivers, and *-phe* 'CLF.LEAF' is used for leaves of trees. Some such classifiers may also be employed as shape classifiers, for instance when *-phe* is used for sheets of paper or bank notes. Many classifiers of this kind denote parts of plants—these are described in more detail in § 7.2.4.1 below.

7.2.4.1 Plants and their parts

A wide array of classifiers are used for plants and their various parts. They are very commonly used as suffixes directly on nouns to modify the meaning and indicate the part that is referred to. An overview of these is presented in Table 7.6. Plant nouns can be said to refer to the plant as a whole, with classifiers providing the necessary degree of detail in order to identify specific parts.

Trees are typically divided into palm trees and non-palm trees in terms of classifier use: *-pi* 'CLF.TUBE' is used for palm trees (*manakhépi* 'açaí palm') and *-na* 'CLF.TRUNK' is used for other trees (*doomalína* 'umari tree').

Many botanical parts are associated with certain classifiers. Some examples are *-íiwi* for flowers (*haikóiwi* 'flower of a tree'), *-wáta* for *curuatás* (*manakhewáta*

Table 7.6: Classifiers associated with plants and their parts

Classifier	Gloss	Associated plant (part)
Trees		
-na	CLF.TRUNK	trees, tree trunks (mainly non-palm trees)
-pi	CLF.TUBE	palm trees (stems)
Botanical parts		
-áapo	CLF.STICK	sticks
-híwi	CLF.POINTED	sticks, twigs
-íiwi	CLF.FLOWER	flowers, thorns
-kénaa	CLF.BRANCH	branches
-khaa	CLF.CURVILINEAR	lianas
-phe	CLF.LEAF	leaves
-pokóda	CLF.STUMP	tree stumps
-ttáwalhe	CLF.CUT	piece between two joints (sugar cane, bamboo)
-wáta	CLF.BRACT	*curuatás* (bracts)
-wáthe	CLF.NODE	wood knots
-ya	CLF.SKIN	tree bark, fruit peel
Fruit and bunches of fruit		
-áanhaa	CLF.LIQUID	fruit juices
-aápa	CLF.OBLONG	single oblong fruits (banana, ingá, corn, manioc)
-da	CLF.GENERIC	single roundish fruits; some bunches
-i	CLF.BASKET	bunches of fruit
-iítsia	CLF.BUNCH	bunches of fruit (bare, without the fruit)
-ma[1]	CLF.PAIR	hands of banana
-póko	CLF.CIRCLE	bunches of Amazon grape, hands of banana
Seeds		
-éekhe	CLF.SMALL.SEED	small seeds (esp. chili peppers)
-íixi	CLF.SEED	seeds

'*curuatá* of açaí'), *-phe* for leaves (*doomalíphe* 'umari leaf'), and *-ttáwalhe* for pieces between two joints (*maapattáwalhe* 'piece of sugar cane').

Single fruits typically take either *-da* 'CLF.GENERIC' or *-aápa* 'CLF.OBLONG', depending on the shape of the fruit: *-da* is used for round fruits (e.g., *póoperi* '*bacaba*') and *-aápa* for oblong fruits (e.g., *palána* 'banana'). For bunches, there are several different classifiers depending on the type: *-i* is used for bunches of fruit from palm trees, while *-iítsia* is used for bare bunches from the same trees, that is, without the fruit—for instance, if the fruit has already been harvested from it (*manakhéetsia* 'bunch of açaí (without fruit)'). Bunches of Amazon grape and hands of banana associate with *-póko* 'CLF.CIRCLE'. The latter can also be used with *-ma*[1] 'CLF.PAIR'.

Finally, there are two classifiers for seeds: *-íixi*, which refers to any seed (*manákhe* 'açaí seed'), and *-éekhe*, which refers to small seeds, especially of chili peppers

(*aattíekhe*).

The high level of differentiation regarding plant parts reflects intimate knowledge of the plants in question. For example, palm tree species like açaí, *bacaba*, and patawa are native to the area, and are important to Baniwa culture and subsistence (Hill, 2009b). Accordingly, palm trees typically take a different classifier than other trees, and specific parts of palm trees, like the *curuatá*, have their own associated classifiers. Bunches are even differentiated according to whether or not they still have the fruits left on them. Another case concerns the culturally important chili peppers (Wright, 2009), the seeds of which have their own specific classifier.

Bound nouns denoting plant parts is a common source for Baniwa classifiers to develop from, which explains the fact that there are so many different classifiers in the plant domain. This is further discussed in § 8.2.2.

7.2.4.2 Body parts

The names for many body parts seem to contain classifier suffixes. Some examples are given in Table 7.7[83].

Table 7.7: Body part terms comprising classifiers

Body part	Meaning	Segmentation	Gloss
-eetshaápa	'jawbone, gums'	*-eétsha-aápa*	tooth-CLF.OBLONG
-hiwída	'head'	*-hiwí-da*	hook/arrow-CLF.GENERIC
-hiwidáapi	'skull'	*-hiwída-áapi*	head-CLF.HOLLOW
-kaaléewhe	'testicles'	*-káale-éewhe*	heart-CLF.EGG
-kakoápa	'jaw'	*-káko-aápa*	cheek-CLF.OBLONG
-kodáapi	'collarbone'	*-kóda-áapi*	chest-CLF.HOLLOW
-nomaya	'lips'	*-nóma-ya*	mouth-CLF.SKIN
-tarapiténa	'sinew, tendon'	*-tarapíte-na*	vein-CLF.TRUNK
-thimáka	'eyelids'	*-thi-máka*	eyes-CLF.FABRIC
-thimakáya	'eyelids'	*-thimáka-ya*	eyelids-CLF.SKIN
-ttamáapi	'spine'	*-ttáma-áapi*	back-CLF.HOLLOW
-ttipalékhaa	'urethra'	*-ttípale-khaa*	urine-CLF.CURVILINEAR

For body parts that come in pairs (e.g., eyes and hands), the unmarked form refers to the pair rather than the singular item. Thus, in (116a), *nóthi* refers to 'my (pair of) eyes'. In order to refer to only one eye, a classifier construction like *apeéma nóthi*, literally 'one side of my (pair of) eyes', in (116b) is necessary. To really

[83]Two of the body parts in Table 7.7, *-hiwidáapi* 'skull' and *-thimákaya* 'eyelids', involve diachronic classifier stacking (see § 5.3.5).

emphasize that one is talking about both eyes, a construction like that in (116c) can be used.

(116) a. *nóthi*
no-thi
1SG-pair.of.eyes

'my (pair of) eyes' [220919_ELIC_F_01]

b. *apeéma* *nóthi*
apa-**eéma** no-thi
one-**CLF.SIDE** 1SG-pair.of.eyes

'my one eye' (lit. 'one **side** of my pair of eyes') [220919_ELIC_F_01]

c. *nhaa* *dzameéma* *nóthi*
nhaa dzama-**eéma** no-thi
3PL two-**CLF.SIDE** 1SG-pair.of.eyes

'both of my eyes' (lit. 'two/both **sides** of my pair of eyes')
[220919_ELIC_F_01]

All paired body parts in Baniwa follow the same pattern as *-thi* '(pair of) eyes': the unmarked form of the noun refers to the pair, while the classifier *-eéma* 'CLF.SIDE' is used as an antidual marker to refer to the singular item[84]. Some other examples are *-híipa* '(pair of) feet', *-káapi* '(pair of) hands', *-kawa* '(pair of) legs', *-héeni* '(pair of) ears', and *-náapa* '(pair of) arms'. This pattern is also visible in the traditional counting system (§ 3.2.9.1), as shown in (117a–117b).

(117) a. *apeéma* *pakáapi*
apa-**eéma** pa-káapi
one-**CLF.SIDE** IMPRS-pair.of.hands

'five' (lit. 'one **side** of a pair of hands') [220921_ELIC_F_01]

b. *dzameéma* *pakáapi*
dzama-**eéma** pa-káapi
two-**CLF.SIDE** IMPRS-pair.of.hands

'ten' (lit. 'two/both **sides** of a pair of hands') [220921_ELIC_F_01]

[84]In some cases, the generic classifier *-da* can also be used: *apa-da no-thi* /one-CLF.GNR 1SG-pair.of.eyes/ 'my one eye' (lit. 'one of my pair of eyes')—cf. (116b).

Other paired entities, e.g., shoes and earrings, follow a slightly different pattern. As for body parts, *-eéma* 'CLF.SIDE' can be used for the singular item (118a), but when referring to the pair, the pair classifier *-ma[1]* is used (118b).

(118) a. *apeéma* *tshapáto*
apa-**eéma** tshapáto
one-**CLF.SIDE** shoe

'one shoe (i.e., one **side** of a pair of shoes)' [200318_FREE_G_01]

b. *apáma* *tshapáto*
apa-**ma[1]** tshapáto
one-**CLF.PAIR** shoe

'one **pair** of shoes' [200318_FREE_G_01]

The pair classifier *-ma[1]*, however, does not combine with paired body parts, as illustrated in (119b) ((119a) features the singular reading with *-eéma* 'CLF.SIDE' for reference). As seen above (cf. (116a)), to refer to both of one's ears, *nhóeni* 'my (pair of) ears' is enough. To really emphasize the reference to both ears, a construction analogous to the one in (116c) above can be used.

(119) a. *apeéma* *nhóeni*
apa-**eéma** no-héeni
one-**CLF.SIDE** 1SG-pair.of.ears

'my one ear' (lit. 'one **side** of my pair of ears') [220921_ELIC_F_01]

b. * *apáma* *nhóeni*
apa-**ma[1]** no-héeni
one-**CLF.PAIR** 1SG-pair.of.ears

Intended: 'my (**pair** of) ears' [220921_ELIC_F_01]

This structural difference between paired body parts and other paired items indicates a difference in their semantics: while the pair is the underlying quantity of the body parts, the nouns of other paired items are rather underspecified for number, which is why the pair classifier is necessary for the latter but ungrammatical with the former.

The overt marking of singular instances of paired body part terms is similar to singulative marking, i.e., the use of a marked singular form in opposition to an

unmarked collective form (e.g., Welsh *moch* 'pigs' vs. *mochyn* 'a pig', Nurmio, 2023: 156). However, the Baniwa construction differs from singulatives in important ways. First, the term *singulative* typically refers to a derived form of the noun (Nurmio, 2023: 157). In Baniwa, the noun itself remains unmarked in both (116a) and (116b), and the singulative-like marking is in the form of a numeral with a classifier suffix. Second, singulatives typically stand in opposition to nouns with collective meaning in general, but in Baniwa this construction concerns paired body part terms specifically.

Michaelis et al. (2013a: 104–105) use the term *antidual* for singulative-like marking referring to singular forms in opposition to non-marked dual forms. The phenomenon appears to be rare, but a number of languages have been described as using antidual marking for body part terms, including Hungarian (Moravcsik, 2003: 148) and six creole languages in the APiCS database (Michaelis et al., 2013a: 104–105). None of these use the dual–antidual as their only strategy, however. It is also not clear whether the antidual construction is part of a larger collective–singulative pattern in these languages or not. In Baniwa, however, the antidual construction appears to be the only option for referring to single instances of body parts that come in pairs. In addition, as described above, it is specific to paired body part terms—other unmarked pairs behave slightly differently.

As for the distribution of antidual marking in the APiCS database, Michaelis et al. (2013a: 105) note that all three French-based Indian Ocean creoles show the construction. The origin of the feature in these languages is unknown, as it can neither be attributed to French nor to either of the two major substratal sources (Malagasy and Eastern Bantu languages). The other three languages showing antidual marking are Haitian Creole, Creolese, and Nengee, all of which are spoken in the Caribbean, the former French-based and the latter two English-based. Similarly, the origin of the feature in these languages remains unknown according to Michaelis et al. It is interesting to note, however, that at least two of the Caribbean antidual-marking languages have Arawakan languages among their substratal sources—Haitian Creole (Taíno substrate; Michaelis et al., 2013c: 195) and Creolese (Lokono substrate; Michaelis et al., 2013b: 49).

7.2.5 Summary of semantic parameters

This section has attempted to generalize the semantic properties associated with individual classifiers (§ 7.1) by sorting them under the parameters animacy, physical properties, collections, and parts, in order to analyze the semantic organization of the system.

In some cases, a single semantic property (or the combination of a few properties) does a good job of defining the extension of a classifier. For example, *-ma*^{2} 'CLF.FEMALE' is used for human females and nothing else. Likewise, *-híwi* 'CLF.POINTED' is well defined by a combination of the properties [+LONG], [+THIN], and [+POINTED], which applies to all of its referents. The classifier *-wána* is used for slices of flat objects, thereby denoting both a certain physical shape and a part–whole relationship.

In other cases, a certain semantic property (or combination thereof) cannot alone define the necessary and sufficient conditions for the combinability between a referent and a classifier. Classifiers are often characterized by overlapping meanings, as illustrated in Figure 7.1. The two classifiers *-Ø* and *-áapi* share connotations of hollowness. Both *-Ø* and *-áapi* can be used for bowls, but only *-áapi* is used for bones and pots, and only *-Ø* is used for canoes and airplanes. Thus, the semantic trait [+HOLLOW] does not correspond neatly to a single classifier.

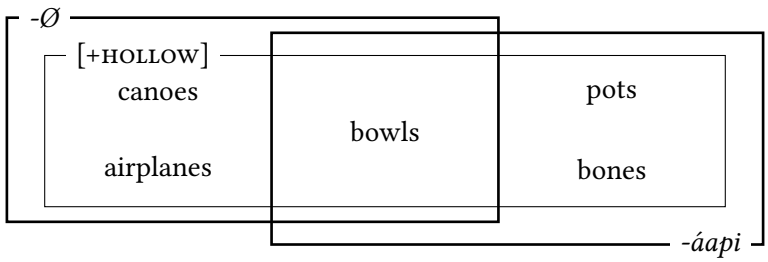

Figure 7.1: Semantic overlap of *-Ø* 'CLF.CANOE' and *-áapi* 'CLF.HOLLOW'

Another example is *-iíta* 'CLF.HUMAN' (Figure 7.2), a classifier used for human males, insects, some fish species, and some objects with flat parts (among other things). There is no single property (or set of properties) that defines all the referents that *-iíta* can be used with. In addition, some of the subsets of referents of this classifier are united by properties that also cover referents associated with other classifiers (e.g., [+FISH] and [+FLAT]).

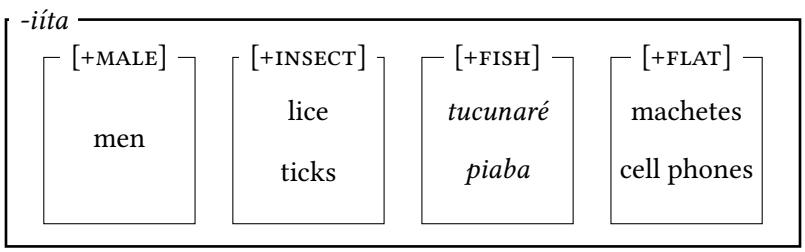

Figure 7.2: Semantic structure of *-iíta* 'CLF.HUMAN'

In order to reach beyond what semantic properties can tell us, the next section (§ 7.3) explores the notion of prototypicality by analyzing the core and extension of classifier categories.

7.3 The structure of classifier categories

In this section, classifiers are viewed as categories, and their internal structure is analyzed in terms of two intersecting dimensions: the strength of their *semantic core* (that is, how well speakers agree on the most central members), and the range of their *semantic extension* (that is, how widely applicable the classifier is). The analysis presented here is based on the framework of prototype theory (§ 2.3).

In an attempt to quantify the notions of semantic core and semantic extension, the Free listing experiment (described in detail in § 1.2.4.1; see also Franjieh, 2012) was conducted. In the experiment, participants were asked individually to list members of each classifier category—that is, nouns compatible with particular classifiers. The lists of all participants were then aggregated. It was hypothesized that when presented with a category and asked to list its members, more prototypical members would be listed earlier and by a greater number of participants than less prototypical members. This is based on findings by Rosch et al. (1976), where subjects listed prototypical members more frequently when primed with a category label. The most mentioned item of each classifier category was taken to represent its semantic core, and the number of listed items across participants was taken to represent the range of its extension. The full collective lists of the Free listing data set, as well as a description of the compilation principles, can be found in Appendix B.

Table 7.8 shows the general results of the Free listing experiment. Under 'Number of items', 'Total' shows the number of items that were listed across participants[85], and 'Total/participant' shows the total number of items divided by the number of participants who left a non-blank response. The number in 'Total/participant' is used as a measure of the semantic extension of the classifier, which will be the subject of § 7.3.2 below. Under 'Core consensus', 'Score' shows the score of the highest-ranking item for each classifier (see § 1.2.4.1 for a description of the scoring system), and '% of max' shows the score as a percentage of the maximal

[85]In types, not tokens—in other words, an item was only counted once even if several participants listed it (e.g., *áawi* 'needle', which was mentioned by five participants for *-híwi* 'CLF.POINTED').

score possible for that particular item[86]. The number in '% of max' will be used as a measure of the semantic core consensus, which will be the subject of § 7.3.1 below.

7.3.1 Semantic core

Identifying the most central members of a category is important for the understanding of the category as a whole (Berlin & Kay, 1969; see § 2.4). Table 7.9 shows the highest-ranking item of each classifier category along with its core consensus measure as a percentage of the maximum (see Table 7.8). The highest-ranking item is taken to represent the semantic core of the classifier.

Three classifier categories have two items tied for first place: *-da* 'CLF.GENERIC' (*hiipáda* 'stone' and *dáapa* 'paca', 18 points), *-wáta* 'CLF.BRACT' (*-wáta* '*curuatá*' and *kóoya* 'calabash bowl', 19 points), and *-xaa* 'CLF.EXCREMENT' (*tsíino íixa* 'dog excrement' and *héema íixa* 'tapir/cow excrement', 19 points).

For two classifier categories, *-éewhe* 'CLF.EGG' and *-ttówhia* 'CLF.ROOM', the most commonly listed item was the numeral–classifier combination itself without an overt noun, rendered as "Ø". In these cases, the numeral–classifier combination is sufficient for conveying the intended meaning ('egg' and 'room', respectively), and that this is the most common way to use these classifiers. Thus, their most prototypical "items" are the meanings conveyed through the lack of an overt noun. For *-máka* 'CLF.FABRIC', the most prototypical member is a Portuguese noun ([*toalha*] 'towel').

In some cases, one and the same lexeme can be the most prototypical member of more than one classifier category. One example is *atsíanli* 'man', which is the most prototypical member of both *-hípa* 'CLF.MALE' and *-iíta* 'CLF.HUMAN'. In other cases, the same lexeme may be the most prototypical member of two different categories, but with different meanings. This is the case for *dáapa* 'paca', which is listed for both *-da* 'CLF.GENERIC', *-eéma* 'CLF.SIDE' and *-íida* 'CLF.HALF'. When used with *-da*, *dáapa* refers to the animal in general. When used with *-eéma* and *-íida*, *dáapa* refers to a piece of a paca that has been cut off.

The most prototypical members are spread out over various semantic domains. Some denote human referents ('woman', 'man'). Some denote animals ('paca',

[86]This was calculated by comparing the score of the highest-ranking item to the maximal score any item could receive in that category, given the number of participants that left a non-blank response. The maximum score of any given classifier category is equal to the number of participants x10. 10 is the score given to the the item that is mentioned first in each individual list, and the maximum score can only be reached if the lexeme is the first mention by every participant.

Table 7.8: General statistics of the Free listing data set

Classifier	Gloss	Number of items		Core consensus	
		Total	Total/participant	Score	% of max
-Ø	CLF.CANOE	4	1	38	95%
-áana	CLF.GROUP	6	1.5	36	90%
-áanhaa	CLF.LIQUID	8	2.67	17	56.7%
-aápa	CLF.OBLONG	15	2.5	45	75%
-áaphi	CLF.AREA	12	2	40	66.7%
-áapi	CLF.HOLLOW	9	1.8	45	90%
-áapo	CLF.STICK	5	1	39	78%
-da	CLF.GENERIC	40	6.67	18	30%
-daa	CLF.DAY	2	0.4	49	98%
-dápana	CLF.HOUSE	3	0.5	60	100%
-eéma	CLF.SIDE	17	3.4	42	84%
-éewhe	CLF.EGG	8	1.6	20	40%
-híko	CLF.LONG	14	2.33	38	63.3%
-hípa	CLF.MALE	5	0.83	60	100%
-hipáda	CLF.PIECE	14	3.5	25	62.5%
-hípani	CLF.RAPIDS	1	0.5	20	100%
-híwa	CLF.BEIJU	1	0.25	40	100%
-híwi	CLF.POINTED	13	2.17	46	76.7%
-i	CLF.BASKET	7	1.75	26	65%
-íida	CLF.HALF	9	2.25	20	50%
-iíta	CLF.HUMAN	27	4.5	40	66.7%
-iítsia	CLF.BUNCH	3	1.5	20	100%
-íiwi	CLF.FLOWER	3	0.6	30	60%
-íixi	CLF.SEED	6	2	18	60%
-kénaa	CLF.BRANCH	5	1	30	60%
-khaa	CLF.CURVILINEAR	21	3.5	38	63.3%
-ko	CLF.HAMMOCK	1	0.2	50	100%
-koa	CLF.SURFACE	11	1.83	44	73.3%
-ma[1]	CLF.PAIR	6	2	30	100%
-ma[2]	CLF.FEMALE	2	0.5	40	100%
-máka	CLF.FABRIC	11	2.2	27	54%
-na	CLF.TRUNK	24	4.8	37	74%
-náko	CLF.BUNDLE	9	1.8	30	60%
-pa	CLF.PACKAGE	11	2.2	27	54%
-páwa	CLF.RIVER	3	0.6	48	96%
-péko	CLF.PATH	8	1.33	40	66.7%
-phe	CLF.LEAF	17	3.4	33	66%
-pi	CLF.TUBE	13	2.17	27	45%
-póko	CLF.CIRCLE	13	2.17	34	56.7%
-pokóda	CLF.STUMP	10	2	17	34%
-ttáwalhe	CLF.CUT	11	2.2	20	40%
-ttówhia	CLF.ROOM	2	0.5	30	75%
-wa	CLF.SPACE	13	2.6	19	38%
-wána	CLF.SLICE	15	2.5	47	78.3%
-wáta	CLF.BRACT	16	2.67	19	31.7%
-wáthe	CLF.NODE	9	2.25	17	42.5%
-xaa	CLF.EXCREMENT	3	1	19	63.3%
-ya	CLF.SKIN	7	1.4	28	56%
-yáwa	CLF.HOLE	9	2.25	27	67.5%

Table 7.9: Semantic core of classifiers

Classifier	Gloss	Most prototypical member(s)	Core consensus
-dápana	CLF.HOUSE	*pántti* 'house'	100 %
-hípa	CLF.MALE	*atsíanli* 'man'	100 %
-hípani	CLF.RAPIDS	*híipa* 'rapids'	100 %
-híwa	CLF.BEIJU	*peéthe* '(piece of) *beiju*'	100 %
-iítsia	CLF.BUNCH	*manákhe* 'branch of açaí fruit'	100 %
-ko	CLF.HAMMOCK	*piéta* 'hammock'	100 %
-ma[1]	CLF.PAIR	*tshapáto* 'shoe/pair of shoes'	100 %
-ma[2]	CLF.FEMALE	*íinaro* 'woman'	100 %
-daa	CLF.DAY	*heekóapi* 'day'	98 %
-páwa	CLF.RIVER	*óoni* 'river, creek, tributary'	96 %
-Ø	CLF.CANOE	*íita* 'canoe, boat'	95 %
-áana	CLF.GROUP	*newíki* 'group of people, clan'	90 %
-áapi	CLF.HOLLOW	*paráto* 'plate'	90 %
-eéma	CLF.SIDE	*dáapa* 'half/side of a paca'	84 %
-wána	CLF.SLICE	*peéthe* 'piece/slice/half of a *beiju*'	78.3 %
-áapo	CLF.STICK	*iitsáapo* 'fishing rod'	78 %
-híwi	CLF.POINTED	*áawi* 'needle'	76.7 %
-aápa	CLF.OBLONG	*palána* 'banana (individual fruit)'	75 %
-ttówhia	CLF.ROOM	Ø 'room'	75 %
-na	CLF.TRUNK	*háiko* 'tree'	74 %
-koa	CLF.SURFACE	*póali* 'manioc oven'	73.3 %
-yáwa	CLF.HOLE	*-nómawa* 'hole'	67.5 %
-áaphi	CLF.AREA	*kiníki* 'manioc garden'	66.7 %
-iíta	CLF.HUMAN	*atsíanli* 'man'	66.7 %
-péko	CLF.PATH	*inípo* 'path'	66.7 %
-phe	CLF.LEAF	*papéra* 'sheet of paper'	66 %
-i	CLF.BASKET	*póoperi* 'bunch of *bacaba* fruit'	65 %
-híko	CLF.LONG	*háiko* 'tree, piece of wood'	63.3 %
-khaa	CLF.CURVILINEAR	*áapi* 'snake'	63.3 %
-xaa	CLF.EXCREMENT	*tsíino íixa* 'dog excrement'; *héema íixa* 'tapir/cow excrement'	63.3 %
-hipáda	CLF.PIECE	*kóphe* 'piece/half of a fish'	62.5 %
-íiwi	CLF.FLOWER	*-íiwi* 'flower'	60 %
-íixi	CLF.SEED	*-íixi(mi)* 'seed'	60 %
-kénaa	CLF.BRANCH	*-ke* 'branch'	60 %
-náko	CLF.BUNDLE	*kiníkhii* 'bundle of *maniva*'	60 %
-áanhaa	CLF.LIQUID	*manákhe(a)* 'açaí juice'	56.7 %
-póko	CLF.CIRCLE	*kamhéro* 'bunch of Amazon grape'	56.7 %
-ya	CLF.SKIN	*íita* 'canoe'	56 %
-máka	CLF.FABRIC	[*toalha*] 'towel'	54 %
-pa	CLF.PACKAGE	*mokóto* '*panacu*'	54 %
-íida	CLF.HALF	*dáapa* 'half/piece of a paca'	50 %
-pi	CLF.TUBE	*mókawa* 'rifle, gun'	45 %
-wáthe	CLF.NODE	*-wáthe* 'knot, tree knot'	42.5 %
-éewhe	CLF.EGG	Ø 'egg'	40 %
-ttáwalhe	CLF.CUT	*háiko* 'piece of wood'	40 %
-wa	CLF.SPACE	*-nómawa* 'hole'	38 %
-pokóda	CLF.STUMP	*palána* '(group of) banana tree(s)'	34 %
-wáta	CLF.BRACT	*-wáta* '*curuatá*'; *kóoya* 'calabash bowl'	31.7 %
-da	CLF.GENERIC	*hiipáda* 'stone; *dáapa* 'paca'	30 %

'snake') or parts of their dead bodies ('piece of a paca', 'piece of a fish'). Quite a large number denote parts of plants ('tree branch', 'flower'). Several relate to the cultivation and production of manioc produce ('*beiju*', 'bundle of *maniva* (manioc stems)', 'manioc garden', 'manioc oven'). Some relate to water ('rapids', 'river') or land ('stone', 'hole', 'path'). Others relate to habitation ('house', 'hammock'), transportation ('canoe'), or hunting, fishing, and gathering ('fishing rod', 'rifle', '*panacu*').

7.3.1.1 Semantic core consensus

The Free listing data set shows that classifiers differ in the degree to which participants agree on their semantic core (see Table 7.9). It was hypothesized that the degree of consensus between participants on the highest-ranking item in a particular category is indicative of the degree of prototypicality of that item. Categories with a high degree of core consensus were taken to be more clearly oriented around a prototypical core.

Eight classifier categories have a semantic core consensus equal to 100% of the maximum score (that is, there is maximal agreement among the participants on the most prototypical member). These are *-dápana* 'CLF.HOUSE', *-hípa* 'CLF.MALE', *-hípani* 'CLF.RAPIDS', *-híwa* 'CLF.BEIJU', *-iítsia* 'CLF.BUNCH', *-ko* 'CLF.HAMMOCK', *-ma*^{1} 'CLF.PAIR', and *-ma*^{2} 'CLF.FEMALE'.

The classifier with the lowest degree of consensus on its most prototypical member was the generic classifier *-da*.

7.3.2 Semantic extension

The classifiers differ greatly in how widely applicable they are. Table 7.10 shows the semantic extension of each classifier, quantified as the number of items (types) listed for the classifier per participant (see Table 7.8).

The highest number of items per participant was listed for *-da* 'CLF.GENERIC' (6.67). Six other classifiers had more than three items listed per participant: *-na* 'CLF.TRUNK', *-iíta* 'CLF.HUMAN', *-khaa* 'CLF.CURVILINEAR', *-hipáda* 'CLF.PIECE', *-eéma* 'CLF.SIDE', and *-phe* 'CLF.LEAF'.

The 10 classifiers *-ko* 'CLF.HAMMOCK', *-híwa* 'CLF.BEIJU', *-daa* 'CLF.DAY', *-hípani* 'CLF.RAPIDS', *-ttówhia* 'CLF.ROOM', *-ma*^{2} 'CLF.FEMALE', *-dápana* 'CLF.HOUSE', *-páwa* 'CLF.RIVER, *-íiwi* 'CLF.FLOWER', and *-hípa* 'CLF.MALE' all had less than one item per

Table 7.10: Semantic extension of classifiers

Classifier	Gloss	Semantic extension
-da	CLF.GENERIC	6.67
-na	CLF.TRUNK	4.8
-iíta	CLF.HUMAN	4.5
-hipáda	CLF.PIECE	3.5
-khaa	CLF.CURVILINEAR	3.5
-eéma	CLF.SIDE	3.4
-phe	CLF.LEAF	3.4
-áanhaa	CLF.LIQUID	2.67
-wáta	CLF.BRACT	2.67
-wa	CLF.SPACE	2.6
-aápa	CLF.OBLONG	2.5
-wána	CLF.SLICE	2.5
-híko	CLF.LONG	2.33
-íida	CLF.HALF	2.25
-wáthe	CLF.NODE	2.25
-yáwa	CLF.HOLE	2.25
-máka	CLF.FABRIC	2.2
-pa	CLF.PACKAGE	2.2
-ttáwalhe	CLF.CUT	2.2
-híwi	CLF.POINTED	2.17
-pi	CLF.TUBE	2.17
-póko	CLF.CIRCLE	2.17
-áaphi	CLF.AREA	2
-íixi	CLF.SEED	2
-ma[1]	CLF.PAIR	2
-pokóda	CLF.STUMP	2
-koa	CLF.SURFACE	1.83
-áapi	CLF.HOLLOW	1.8
-náko	CLF.BUNDLE	1.8
-i	CLF.BASKET	1.75
-éewhe	CLF.EGG	1.6
-áana	CLF.GROUP	1.5
-iítsia	CLF.BUNCH	1.5
-ya	CLF.SKIN	1.4
-péko	CLF.PATH	1.33
-Ø	CLF.CANOE	1
-áapo	CLF.STICK	1
-kénaa	CLF.BRANCH	1
-xaa	CLF.EXCREMENT	1
-hípa	CLF.MALE	0.83
-íiwi	CLF.FLOWER	0.6
-páwa	CLF.RIVER	0.6
-dápana	CLF.HOUSE	0.5
-hípani	CLF.RAPIDS	0.5
-ma[2]	CLF.FEMALE	0.5
-ttówhia	CLF.ROOM	0.5
-daa	CLF.DAY	0.4
-híwa	CLF.BEIJU	0.25
-ko	CLF.HAMMOCK	0.2

participant listed (that is, some participants did not list any items at all for these classifiers, decreasing the average number to less than one).

7.3.3 Category structure

The scatter plot in Figure 7.3 combines the measures of semantic core consensus[87] (*y* axis) and semantic extension[88] (*x* axis), and each classifier is plotted in this two-dimensional space. This plot will be used as a starting point for discussing the category structure of classifiers.

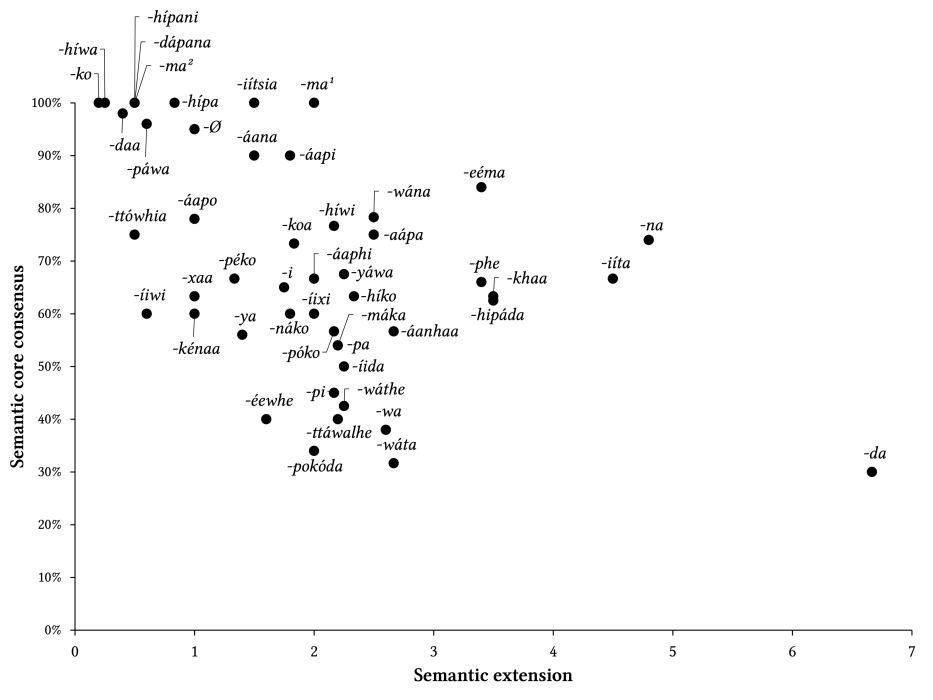

Figure 7.3: Semantic core consensus and semantic extension of classifiers

In general terms, a wide extension is associated with a lower degree of core consensus, and a narrower extension is associated with a higher degree of core consensus. The majority of the classifiers position themselves to the middle–left of the plot, forming a loose cluster of categories with a medium core consensus coupled with a relatively narrow extension. The following sections contain a

[87] Measured as the score of the highest-ranking member, as a percentage of the maximum score possible (see § 7.3.1).

[88] Measured as the number of listed items per participant (see § 7.3.2).

more general discussion of the two extremes in Figure 7.3: the generic classifier (§ 7.3.3.1) and the specific classifiers (§ 7.3.3.2).

7.3.3.1 The generic classifier

The most extreme outlier in Figure 7.3 is *-da*, which functions as a generic classifier. This classifier has the lowest degree of core consensus among the classifiers, combined a much wider semantic extension than any other classifier.

The generic classifier can be used in reference to most objects, depending on the circumstances. The use of *-da* covers all three possible functions of generic classifiers identified by Aikhenvald (2000: 335–337): the residue function, the default function, and the unspecified referent function (see § 2.5.2).

Due to its status as a generic classifier, *-da* can be used with almost any referent, and is by far the most commonly used classifier in Baniwa. It is the classifier for which the highest number of unique items were listed (40, compared to the average of 9.7). In the Noun list data set (§ 1.2.4.2), *-da* is compatible with at least 70% of the sample. In the naturalistic texts (§ 1.2.5), it accounts for 62% of the classifier occurrences. The youngest of the speakers consulted for this thesis, a 25-year-old man, uses no other classifier than *-da* throughout an entire picture book narrative (Mayer, 1967). Whether there is a generational shift going on, where the generic classifier is gaining ground at the expense of more specific classifiers among the young, urbanized speakers of Baniwa remains a question for future research.

Historically, *-da* can be reconstructed to PJC **-aʔa-la* 'round' (Ramirez, 2020d: 96; see § 2.5.2). According to Greenberg (1972: 34–35), generic classifiers typically develop from classifiers for round things. This classifier still retains a strong connotation of roundness for Baniwa speakers, as they state upon being asked about its use. Its roundness connotation is also clear from many of the most commonly listed items: these include roundish objects and beings, like stones, eggs, pacas, piranhas, and fruits (see Appendix B for a more detailed list).

In some cases, *-da* can be used to add a connotation of roundness to a referent, in particular if the referent is typically referred to by another classifier. For example, *aapáma íinaro* /one.CLF.FEMALE woman/ is the typical way of referring to 'one woman'. *Apáda íinaro* /one.CLF.GENERIC woman/ is likely to be interpreted as 'one short and chubby woman'.

The generic meaning and high frequency of *-da* explains its positioning in Figure 7.3: it is widely applicable but semantically rather general, and thus there is

comparatively little agreement on a prototypical member (it even has a tie for the first place between *hiipáda* 'stone' and *dáapa* 'paca', 18 points). As Lakoff (1986: 17) points out, so-called "everything else" categories have neither central members nor chaining. Interestingly, however, the only four items listed by two speakers (*hiipáda* 'stone', *dáapa* 'paca', *-iináka* 'fruit', *itsída* 'tortoise') are all examples of round(ish) things. It thus seems like *-da* retains a core connotation of roundness, despite having become semantically bleached.

On the other hand, many of the referents used with *-da*—including referents that score high in the Free listing toplist—lack a round(ish) shape. This includes shapeless referents such as *déepi* 'night' and *-wapiñeetákhe* 'thought, idea', as well as more recent Portuguese loanwords for concepts that may not fit neatly into any other classifier category, e.g., [*computador*] 'computer', [*caderno*] 'notebook' and [*relógio*] 'clock, watch'.

Many referents that are typically used with another classifier can alternatively be used with *-da*. These include *déepi* 'night' and *heekóapi* 'day' (*-daa* 'CLF.DAY'), *íita* 'canoe' and *kóoya* 'calabash bowl' (*-Ø* 'CLF.CANOE'), *tsipaláapi* 'pot, pan' and *-peréma* 'rib' (*-áapi* 'CLF.HOLLOW'), *tshéeto* '*aturá* basket' (*-i* 'CLF.BASKET'), [*lápis*] 'pencil' (*-híwi* 'CLF.POINTED'), and *pántti* 'house' (*-dápana* 'CLF.HOUSE'). While it is true that many referents may be compatible with several classifiers in the flexible system of Baniwa, the compatibility with *-da* stands out in that it does not seem to have any obvious semantic limits.

However, the generality of *-da* does not mean that it is always compatible with all referents under all circumstances. For example, it appears to be somewhat less favoured as an alternative for distinctly non-round entities, e.g., long objects. The appropriateness of a certain classifier is also highly context-dependent and determined by, e.g., the level of specificity that the speaker needs or wants to convey. For example, the noun *máapa* refers to both bees, honey and sugar canes. To refer to a sugar cane, the classifier *-híko* 'CLF.LONG' is typically used, and while the generic classifier *-da* may not be theoretically incompatible with sugar canes, it is more likely to instead convey the meaning 'bee'.

In expressions referring to multiple referents that normally take different classifiers, such as a number of objects of different shapes, or a group of people consisting of both men and women, *-da* is the classifier of choice (see § 7.5).

In some languages, items that do not fit into the categorization may simply remain unclassified (see § 2.5.2). This is not the case in Baniwa. Baniwa does have a zero classifier *-Ø*, but it is reserved for canoes and hollow objects (see § 7.1)—zero marking cannot be used as a default classifier, or as a way of "avoiding" classification. Instead, the generic classifier is used to accommodate such items.

7.3.3.2 Specific classifiers

Specific classifiers have a narrow extension coupled with a very high prototype consensus, and are found in the upper left corner of Figure 7.3. These classifiers are typically centered around a prototypical referent rather than determined by certain semantic properties. Some examples of specific classifiers are listed in Table 7.11.

Table 7.11: Examples of specific classifiers

Classifier	Gloss	Use
-Ø	CLF.CANOE	Canoes (+bowls, contained vehicles)
-daa	CLF.DAY	Days, nights
-dápana	CLF.HOUSE	Houses (+buildings, tents)
-hípa	CLF.MALE	Men
-hípani	CLF.RAPIDS	Rapids
-híwa	CLF.BEIJU	*Beiju*
-ko	CLF.HAMMOCK	Hammocks
-ma²	CLF.FEMALE	Women
-páwa	CLF.RIVER	Rivers (+tributaries, creeks)
-wálhia	CLF.YEAR	Years

Four of the classifiers in Table 7.11 are so-called "unique classifiers" (Grinevald, 2015: 815; Aikhenvald, 2021: 234; see § 2.5.3) which only ever combine with a single noun: *-ko* 'CLF.HAMMOCK' (with *piéta* 'hammock'), *-híwa* 'CLF.BEIJU' (with *peéthe* '*beiju*'), *-hípani* 'CLF.RAPIDS' (with *híipa* 'rapids'), and *-wálhia* 'CLF.YEAR' (with *hámoli* 'dry season; year').

Other classifiers in this group have a very specific meaning, but are not unique classifiers in the sense that they only combine with a single *noun*. Rather, they combine with referents of a very specific type, but the referents may be linguistically represented by different nouns. These include *-daa* 'CLF.DAY' (for days and nights), *-ma²* 'CLF.FEMALE' (for women), *-hípa* 'CLF.MALE' (for men), and *-páwa* 'CLF.RIVER' (for rivers and similar watercourses).

Yet other classifiers in Table 7.11 are centered on a specific referent, but can be extended to cover others—that is, they have a slightly wider extension. An example of one such classifier is *-dápana* 'CLF.HOUSE': it is centered on houses, but can be used for any buildings and human habitations (including tents) by extension. Another is *-Ø* 'CLF.CANOE', which is centered on canoes, but can also be applied to some other hollow, carved referents (e.g., calabash bowls), as well as other contained vehicles (e.g., airplanes).

Table 7.11 contains several specific classifiers that are used with items of great

importance in Baniwa culture. These classifiers relate to the domains of habitation (*-dápana* for houses; *-ko* for hammocks), transportation (*-Ø* for canoes, the main vehicle; *-páwa* for rivers, the main watercourses; *-hípani* for dangerous rapids), and nutrition (*-híwa* for the staple *beiju* (manioc bread)). Habitation, transportation and nutrition are basic human activities relevant to all people, but the Baniwa classifiers in question reflect the specific traditional sociocultural environment of Baniwa speakers.

It has been suggested that specific classifiers cross-linguistically tend to reflect items of cultural salience (Grinevald, 2015: 815–816; Aikhenvald, 2021; see § 2.5.3). However, there are other salient items in Baniwa culture that do not have specific classifiers (e.g., manioc tubers and *farinha*, another manioc-based staple), so the cultural status of the items mentioned above cannot alone account for the existence of the corresponding specific classifiers. Very little attention has been devoted to specific classifiers in the typological literature, and more research is needed for a better understanding of the link between culturally important items and specific classifiers.

7.3.3.3 The role of mythology and ritual in the classifier system

Dixon's (1982) influential work on Dyirbal (Pama-Nyungan, Northeastern Australia) showed that the assignment of referents into the language's four noun classes is based on a complex combination of taxonomic principles and mythical associations (see also Lakoff, 1987). For Baniwa-Koripako varieties spoken in Venezuela, Hill (1988) provides an account of how social, ritual, mythological, and symbolic associations may play a role in the semantic organization of the nominal classification system. He argues that the linguistic classifier system is intimately connected to categorizations made in the ritual chant-language (a highly specialized form of discourse mastered only by shamans, or "chant-owners"), and that a basic, tripartite distinction of masculine, feminine, and presexual underlies the classifier categories. For example, he argues that the placement of male humans and most fish species in the same category (*-iíta* 'CLF.HUMAN') is due to a mythological association, since fish appear as symbols of the phallus in several myths (see, e.g., Hill (1988: 68; 2009a: 22, 117), Wright (1993/1994), and Albuquerque & Garnelo (2018)).

From a semantic perspective, it can be argued that some of Hill's provided explanations for the grouping of referents can be explained by other principles. For instance, the fact that *surubim* catfish, vines, and snakes are categorized together (*-khaa* 'CLF.CURVILINEAR') is rooted in ceremonial and mythical associations of these referents according to Hill (1988: 68). An alternative explanation

is the similarity in physical shape of these and other referents in the category (see § 7.2.2)[89]. This explanation is also in line with the general tendency for languages in Northwestern Amazonia to have shape-based nominal classification systems (see § 2.7.1). On the other hand, it is possible that physical similarities and linguistic associations may lead items to be associated in ritual in the first place, which may strengthen the associations further.

Another relevant point concerns the difference between the speakers Hill worked with (ritual specialists in traditional settings) and the speakers I worked with (an urbanized population without specialized ritual knowledge). To the extent that mythical associations ever played a role in the Baniwa classifier system, it is likely that they are less salient for the speakers in my study than the ones Hill worked with four decades ago, as the daily interaction with the mythical realm differs significantly for individuals in these two groups. Due to the rapid pace of cultural change resulting from contact with the majority culture (Shulist, 2018b), it is possible that the understanding of the underlying organizational principles have been reinterpreted to some extent. In any case, the speakers I worked with have all expressed very clear ideas about the fact that referents are grouped together based on similarities in shape.

A possible parallel can be drawn to Dyirbal. Dixon (1972; 1982) began his documentation in the 1960s, when the language was still spoken in a relatively traditional setting. In the following decades, influence from Australian majority culture and the English language led to a rapid and drastic restructuring of Dyirbal, including the noun class system. As Schmidt (1985) demonstrated, Dyirbal speakers in the 1980s could be divided into three generational strata: the older generation who spoke traditional Dyirbal and mastered the full, traditional, quadripartite noun class system, the intermediate generation who used a slightly simplified form of the system, and the younger generation who used a very simplified, tripartite system (see also Lakoff, 1987: 97–98). Among the Baniwa speakers I have worked with, the youngest one appears to be using a significantly reduced system compared to the rest of the speakers (see § 1.2.4.1, 7.3.3.1).

7.3.4 Reflections on prototypicality

Language is shaped by the cultural context in which it is used, and nominal classification systems are no exception. As will be shown in more detail in Chapter 9 on language contact, the Baniwa classifier system has adapted to the influence

[89] Hill (1988: 69) does mention their similarity in shape, but analyzes the shape association as secondary to the mythical association.

from Brazilian majority culture in a number of different ways.

The results of the Free listing experiment are a case in point, where the influence of modern life is obvious: rooms, towels, and rifles are all non-Indigenous concepts that were nevertheless mentioned most often in their respective classifier categories (see Table 7.9) by the particular set of speakers I worked with. While the prototypicality of certain members in classifier categories may largely be generalizable across the urbanized Baniwa-speaking population in São Gabriel da Cachoeira today, it is likely that the same study would have yielded at least partially different results among Baniwa speakers living in more traditional settings, not to mention among past generations, where some of the objects that are commonly encountered in town might be rare or even virtually unheard of. Conversely, there are many things that are more present in traditional village life (see the discussion in § 7.3.3.3 of variation in access to mythical knowledge for Baniwa speakers in different social settings). This goes to show how the notion of prototypicality can shift quite rapidly, as an effect of cultural changes in terms of what the speakers of a particular linguistic variety tend to encounter in their daily lives. This is discussed in more detail in § 10.2.

7.4 Inter-speaker variation

Apart from the variation in classifier assignment that is due to the flexibility of the system itself (described in § 4.4), there is some inter-speaker variation when it comes to classifier choice.

As described in § 1.1, the Baniwa-Koripako continuum comprises several closely related dialects. Speakers from various parts of the Central dialect continuum (the variety known as *Baniwa*) live in São Gabriel da Cachoeira where the data for this thesis were collected, and are represented among the participants (see § 1.2.3). The basic classifier assignment principles are essentially the same regardless of dialect (see Aikhenvald, 2007: 498). Although no systematic study is made here of the dialectal variation relating to classifier assignment, it is clear that some dialectal differences do exist, that speakers are highly aware of. When asking speakers about the use of classifiers, they commonly reply that while they may not combine a certain classifier with a certain noun, they have heard people from some other place or speakers of some other dialect do so.

One example of something that may be a dialectal feature is the classifier assignment to the noun *íita* ‘canoe’, which is most canonically assigned the dedicated canoe classifier *-Ø* (§ 7.1.1). However, some people may also use the skin and

bark classifier *-ya* (§ 7.1.52) for canoes. For other speakers this is impossible, and they claim that it is only "people from Aiari" who speak like that. Another dialectal feature attributed (by others) to speakers who live on the Aiari river is the use of the classifier *-iítsia* 'CLF.BUNCH' with bunches of palm fruits where the fruit is still left on the bunch. Other speakers use this classifier only for bunches where the fruit has already been picked (for bunches with the fruit left, *-i* 'CLF.BASKET' is used).

Another likely source of inter-speaker variation comes from differences in living environment, in particular between speakers residing in remote, traditional settlements and the Baniwa-speaking population in the regional urban center São Gabriel da Cachoeira (see § 1.1.2.1). As almost all of the data collected for this thesis comes from speakers who live in town, comparison with speakers from other localities is outside the scope of this study, but one likely effect of the town-dwelling participant sample can be seen in the prototype effects of some classifiers (see § 7.3), which was discussed in § 7.3.4. Another effect of the urban environment is the increased level of engagement with Portuguese in daily life and, consequently, differences between speakers in linguistic competence in Baniwa. Among the speakers consulted for this study, one states that Portuguese is his strongest language (for all others, it is Baniwa, see § 1.2.3). This person is also the youngest in the sample, which suggests a possible generational difference.

Finally, there seems to be some idiolectal variation involved in classifier use, in the sense that it is sometimes a matter of personal preference. However, such variation is difficult to distinguish from the more structural kinds of variation described above without systematic study. In any case, even siblings sometimes have different intuitions about how classifiers are assigned and used.

7.5 Classifier use with multiple referents

When talking about multiple referents that are normally assigned different classifiers, speakers resort to the generic classifier *-da* (§ 7.3.3.1). This is the case both for inanimate objects and human referents. When referring to an egg, a thread, and a leaf with the numeral *madali*-CLF 'three', *-da* seems to be the only grammatical classifier option (120). The curvilinear classifier *-khaa* and the leaf classifier *-phe* (used for threads and leaves, respectively) are deemed ungrammatical when referring to this constellation of objects[90].

[90] However, *madalíkhaa* can be used for three threads, *madalíphe* for three leaves, etc.

(120) a. *madalída*
madali-**da**
three-**CLF.GNR**

‘three (of egg, thread, and leaf)’ [221005_ELIC_K_01]

b. * *madalíkhaa*
madali-**khaa**
three-**CLF.CURV**

Intended: ‘three (of egg, thread, and leaf)’ [221005_ELIC_K_01]

c. * *madalíphe*
madali-**phe**
three-**CLF.LEAF**

Intended: ‘three (of egg, thread, and leaf)’ [221005_ELIC_K_01]

Likewise, when using the phrase *madali*-CLF *newíki* ‘three people’ in reference to two men and one woman, *-da* is the only classifier alternative judged felicitous by participants (121), as the female classifier *-ma²* cannot be used with male referents, and the human classifier *-iíta* cannot be used with expressly female referents.

(121) a. *madalída* *newíki*
madali-**da** newíki
one-**CLF.GNR** person

‘three (of two men and a woman)’ [221005_ELIC_K_01]

b. * *madaliíta* *newíki*
madali-**iíta** newíki
one-**CLF.HUMAN** person

Intended: ‘three (of two men and a woman)’ [221005_ELIC_K_01]

c. * *madalíma* *newíki*
madali-**ma²** newíki
one-**CLF.FEMALE** person

Intended: ‘three (of two men and a woman)’ [221005_ELIC_K_01]

7.6 The issue of sortal and mensural classifiers

As mentioned in § 2.7.1, the typological literature on numeral classifiers generally distinguishes two different types: sortal classifiers that classify in terms of what entity is in question, and mensural classifiers that classify in terms of the quantity of the entity (Lyons, 1977: 463; Aikhenvald, 2000: 114 ff.; Kilarski, 2013: 35; Her, Hammarström & Allassonnière-Tang, 2022). This distinction has been debated; for instance, Senft (2000: 22–23) argues that it may hold for individual languages, but only if there is some formal evidence for it (consistent with an *emic* rather than an *etic* point of view, see Pike, 1954: 8 ff.; Evans, 2010). The Baniwa system contains classifiers whose meanings correspond to those typically expressed by sortal and mensural classifiers, respectively, but there is little language-internal evidence for a mensural/sortal distinction.

First, all classifiers in Baniwa are marked in the same morphosyntactic environments, where they compete for the same slots (§ 5.3.1). It is not possible to distinguish a mensural and a sortal set based on their morphosyntactic behavior or their formal characteristics—as is the case in, e.g., Ojibwe (Algonquian, United States/Canada), where sortal and mensural classifiers are used with different forms of the numeral 'one' (Meyer, 2020: 34 ff.).

Second, all classifiers in Baniwa have a specifying function (§ 6.2)—there is no reason to impose a distinction into "quantities" and "kinds" (Lyons, 1977: 463; Aikhenvald, 2000: 114–115). (122) shows two phrases whose semantics correspond to typical examples of mensural (122a) and sortal (122b) classification. In both cases, the classifier specifies the shape of the water-related referent.

(122) a. *aapána* *óoni*
apa-**na** óoni
one-**CLF.TRUNK** water

'one bottle of water' (lit. 'one trunk-shaped unit of water')
[200315_ELIC_F_01]

b. *apákhaa* *óoni*
apa-**khaa** óoni
one-**CLF.CURV** water

'one river' (lit. 'one curvilinear unit of water') [200326_ELIC_F_01]

Another example is shown in (123) with the noun *peéthe* '*beiju* (manioc bread)'. In both examples, the classifier specifies a unit of a certain kind—a slice unit

in (123a), and a *beiju* unit in (123b). There is no language-internal reason to analyze (123a) as an example of "quantity" classification and (123b) as "kind" classification.

(123) a. *apawána* *peéthe*
apa-**wána** peéthe
one-**CLF.SLICE** beiju

'one slice of *beiju*' (one slice unit of *beiju*) [200309_FREE_K_01]

b. *aphéwa* *peéthe*
apa-**híwa** peéthe
one-**CLF.BEIJU** beiju

'one *beiju*' (one *beiju* unit of *beiju*) [200309_FREE_K_02]

Third, in the literature, it has been argued that mensural classifiers can be used with both mass and count nouns while sortal classifiers can be used only with count nouns (Grinevald, 2004: 1020; Kilarski, 2013: 35). As discussed in § 3.2.5, nouns in Baniwa are formally unmarked for countability and number in most cases, so this is not a very reliable diagnostic. Some quantifiers do appear to be sensitive to the countability of nouns (§ 3.2.9), but this is not the case for numerals (§ 3.2.9.1): all nouns behave similarly in the sense that for any noun, the referent needs to be specified by means of a classifier if it is used in a morphosyntactic context that requires it (122–123).

Fourth, it is sometimes argued that mensural classifiers clearly contribute meaning while sortal classifiers may seem redundant (Grinevald, 2004: 1020; Kilarski, 2013: 35), which is an argument that stems mainly from the way we choose to translate them into non-classifier languages, as pointed out by Lucy (2000: 331–332). As shown throughout this thesis, it is certainly not the case that classifiers in Baniwa are redundant. In (122b), the curvilinear classifier *-khaa* clearly contributes to establishing the 'river' meaning, especially if we consider the fact that the noun *óoni* simply means 'water'. Even in a case like (123b), where the *beiju* classifier *-híwa* is used to specify a unit of the noun *peéthe* '*beiju*', the classifier is not redundant, as shown by the contrasting example in (123a) which specifies a different unit of *beiju*.

Fifth, mensural classifiers are sometimes characterized as denoting temporary properties, and sortal classifiers as denoting inherent properties of nouns (Aikhenvald, 2000: 115; Kilarski, 2013: 35). Considering the examples with *óoni* 'water' in (122), we might conclude that the trunk-shaped quality of a bottle of water is

an example of a temporary property of water, but can the curvilinear quality of a river really be said to be a inherent property of water? This highlights the important difference between nouns and referents, e.g., 'water' vs. 'river' in (122b), and is also related to the previous point about the crucial semantic contribution of classifiers to expressions.

Sixth, Aikhenvald (2000: 115) describes mensural classifiers as allowing more freedom of choice than sortal classifiers. The general situation in Baniwa is that almost any noun allows a high degree of freedom when it comes to classifier choice of any kind (see § 4.4)—this it not restricted to mensural uses.

To conclude, nouns and classifiers in Baniwa take part in an intricate interplay when it comes to establishing reference. As shown above, there is little evidence that the system makes any straightforward distinctions between "kind" and "quantity" classification, "redundant" and "necessary" classifiers, or "temporary" and "inherent" properties. I have found no formal or semantic properties that warrant a distinction between mensural and sortal classifiers in Baniwa.

7.7 Chapter summary

The classifier system can be described through a combination of semantic properties and semantic prototypes, both of which play a role for the extension of individual classifiers. Among the semantic parameters relevant to the classifier system, animacy plays an important role as it largely sets human and non-human referents apart. For non-human referents, including animals, physical shape is the most salient parameter. Shape classifiers are also metaphorically extended to abstract referents, like time units and concepts relating to language and cognition. Many classifiers can also be used to specify parts or collections of things. The same classifier may carry connotations of several different parameters, such as a shape property and a part–whole relationship. The commonly made distinction between mensural and sortal classifiers does not apply in Baniwa.

Some classifiers are better defined by their most prototypical members than by semantic properties. The classifiers differ in how strongly speakers agree on their most prototypical members, as well as in how wide a range of referents they can associate with. Highly specific classifiers are only used with a very narrow type of referent (e.g., *-ko*, used exclusively for hammocks), while others can be extended to other referents by perceived similarity with the prototype (e.g., *-Ø*, which is centered on canoes but can be used for other hollow objects and contained vehicles). One classifier, *-da*, functions as a generic classifier.

Chapter 8

Diachrony

This chapter is concerned with the historical origins of the classifier system. It starts by proposing a number of diachronic strata for the classifiers based on previous historical work on the Arawakan language family[91] (§ 8.1). It then goes on to describe the possible pathway from bound nouns to classifiers (§ 8.2), discussing the relationship between classifiers and bound nouns and how they each contribute to the enrichment of the other class over time (§ 8.2.1), as well as the possibility that the semantic distinctions present in the classifier system are partly inherited from the ones present in the class of bound nouns (§ 8.2.2). Finally, the etymological origins of individual classifiers are examined (§ 8.3).

8.1 Diachronic strata of Baniwa classifiers

In § 2.8.1, it was shown that the classifier system of Baniwa has roots dating as far back as Proto-Arawak (PA)[92] (Dunn, 2022). Five present-day classifiers are likely to be reflexes of classifiers from that earliest reconstructable stage, and are taken to represent the oldest stratum of classifiers in Baniwa. At the Proto-Japurá-Colombia (PJC)[93] level, 15 of the classifiers in Baniwa can be reconstructed (Ramirez, 2020d), including the five from PA. The 10 classifiers that are not also reconstructable to PA are taken to represent the second oldest stratum of classifiers in Baniwa.

[91] In this chapter, I do not propose any reconstructions—the Proto-Arawak forms are reconstructed by Dunn (2022), and the Proto-Japurá-Colombia forms by Ramirez (2020d).

[92] It is not known when PA was spoken, but some sources estimate a time-depth of ca 4,000–3,500 years BP (Noble, 1965; Zucchi, 2002).

[93] PJC is estimated by Ramirez (2020d) to have been spoken around 2,000 years BP.

The reconstructable classifiers are shown in Table 8.1. The oldest stratum contains the five classifiers that are (likely) reflexes of classifiers that Dunn (2022: 54) has reconstructed for PA[94]. The second oldest stratum contains the 10 classifiers that date from PJC (Ramirez, 2020d: 96–97).

Table 8.1: Reconstructable classifiers

Clf	Gloss	PA (Dunn, 2022: 54)	PJC (Ramirez, 2020d: 96–97)
Oldest stratum: classifiers with reconstructed root in PA			
-áanhaa	CLF.LIQUID	***sa* 'liquid'	**-aa* 'water'
-koa	CLF.SURFACE	***kwa* 'flat'	**-kʊa* 'limited area'
-máka	CLF.FABRIC	***ma* 'cloth'	**-maka* 'fabric, cloth'
-na	CLF.TRUNK	***na* 'large, long'	**-na* 'trunk, mammal'
-pi	CLF.TUBE	***pi* 'long and thin'	**-pɨ* 'long tube, rope, liana'
Second oldest stratum: classifiers with reconstructed root in PJC			
-áapi	CLF.HOLLOW	–	**-aapi* 'container'
-áapo	CLF.STICK	–	**-aapʊ* 'long and flexible'
-da	CLF.GNR	–	**-aʔa-la* 'round'
-eéma	CLF.SIDE	–	**-eema* 'side'
-híko	CLF.LONG	–	**-Sikʊ* 'tube'
-híwa	CLF.BEIJU	–	**-Siwa* 'manioc bread'
-iíta	CLF.HUMAN	–	**-iita* 'human'
-khaa	CLF.CURV	–	**-kaSa* 'threadlike'
-póko	CLF.CIRCLE	–	**-pʊkʊ(i)* 'circular'
-wa	CLF.SPACE	–	**-wa* 'hole'

Subsequently, individual classifiers appear to have emerged at different points in time, via grammaticalization from bound nouns (see § 8.2). It is difficult to reconstruct a precise timeline, for several reasons: the lack of historical records (even relatively recent ones), the lack of consensus on the internal branching within the family, and the tendency of classifier systems to be affected by language contact (§ 2.7.2). However, it is possible to make an internal comparison of classifiers and bound nouns in order to identify the forms that belong to both categories simultaneously (see Heine's (2004: 579) overlap model, § 2.2). This is illustrated in Table 8.2.

Table 8.2: Criteria for the status as classifier vs. bound noun

		Can take person prefix	Can attach to numeral
-héeni	'ear'	✓	–
-ya	'skin'; 'CLF.SKIN'	✓	✓
-da	'CLF.GENERIC'	–	✓

[94] The 'long and thin' classifier ***pi* was also reconstructed by Payne (1991) as a classifier, albeit in a slightly different form: ***-apʰi* 'long, slender, snake-like'.

Classifiers were defined in § 4.1 by their ability to attach to numerals (e.g., *-da*, *-ya*). Bound nouns (§ 3.2.5) obligatorily take a person prefix (e.g., *-ya*, *-héeni*), which classifiers cannot do, as they are suffixes and not roots (e.g., *-da*). The ability of a classifier form to take a person prefix was therefore used as a diagnostic for its dual status as classifier and bound noun (e.g., *-ya*). This was tested by asking native speakers to judge the grammaticality of each identified classifier morpheme prefixed with a person prefix, and to provide a translation if the combination was deemed grammatical. 16 classifiers were found to have a formally identical bound noun, which was similar or identical in meaning. These are shown in Table 8.3. Grammaticalization typically entails both semantic bleaching and phonetic erosion (Heine, 2004; see § 2.2), which none of the 16 classifiers in question have undergone much of. It may therefore be hypothesized that these classifiers represent the most recent stratum of classifiers in Baniwa.

Table 8.3: Classifiers with identical corresponding bound nouns

Classifier		**Bound noun** (with 3SGNF prefix)	
-wána	CLF.SLICE	*li-wána*	'his/its slice/piece'
-wáta	CLF.BRACT	*li-wáta*	'his/its *curuatá*'
-ya	CLF.SKIN	*lí-ya*	'his/its skin/bark'
-íixi	CLF.SEED	*líixi* (*li-íixi*)	'his/its seed/kernel'
-phe	CLF.LEAF	*lí-phe*	'his/its leaf'
-hipáda	CLF.PIECE	*lhipáda* (*li-hipáda*)	'his/its piece'
-ttáwalhe	CLF.CUT	*li-ttáwalhe*	'his/its piece'
-éewhe	CLF.EGG	*líewhe* (*li-éewhe*)	'his/its egg'
-ttówhia	CLF.ROOM	*li-ttówhia*	'his/its room'
-iítsia	CLF.BUNCH	*liítsia* (*li-iítsia*)	'his/its bunch'
-íiwi	CLF.FLOWER	*líiwi* (*li-íiwi*)	'his/its flower/thorn'
-kénaa	CLF.BRANCH	*lí-ke(naa)*	'his/its branch'
-pokóda	CLF.STUMP	*li-pokóda*	'his/its tree strump'
-wáthe	CLF.NODE	*li-wáthe*	'his/its knot'
-éekhe	CLF.SMALL.SEED	*líekhe* (*li-éekhe*)	'his/its seed'
-kódzoa	CLF.BEND	*li-kódzoa*	'his/its curve'

Note that the overlap between classifiers and bound nouns is restricted to a small set of forms. For this reason, I do not consider Baniwa to make use of repeaters (§ 2.2), as these typically imply that the classifier system is an open class (Aikhenvald, 2000: 103).

There is no overlap between the classifiers that are reconstructable to PA/PJC (the hypothesized oldest stratum, § 8.1) and those that can be used as bound nouns synchronically (the hypothesized youngest stratum, § 8.3)—that is, none of the reconstructable classifiers is able to operate as a bound noun in present-day Baniwa. The fact that they do not overlap permits the hypothesis that they represent opposite ends on a time-depth continuum.

There are 22 classifiers that can neither be reconstructed nor used as bound nouns synchronically. These are shown in Table 8.4, and are referred to here as being of unknown age. In particular when it comes to historical reconstruction, it cannot be ruled out that some of the yet uncategorized classifiers could have been present at some earlier point in history, but there is currently no evidence to support it. However, it is likely that many of the classifiers in this group would position themselves somewhere between the oldest and the youngest stratum on the time-depth continuum. Semantically, their meanings range from quite specific (*-wálhia* 'CLF.YEAR', *-hípani* 'CLF.RAPIDS') to quite general (*-aápa* 'CLF.OBLONG', *-áaphi* 'CLF.AREA'). Their etymological origins will be investigated in § 8.3.

Table 8.4: Classifiers of unknown age

Clf	Gloss	Clf	Gloss
-Ø	CLF.CANOE	*-ko*	CLF.HAMMOCK
-áana	CLF.GROUP	*-ma*^{1}	CLF.PAIR
-aápa	CLF.OBLONG	*-ma*^{2}	CLF.FEMALE
-áaphi	CLF.AREA	*-náko*	CLF.BUNDLE
-daa	CLF.DAY	*-pa*	CLF.PACKAGE
-dápana	CLF.HOUSE	*-páwa*	CLF.RIVER
-hípa	CLF.MALE	*-péko*	CLF.PATH
-hípani	CLF.RAPIDS	*-tsoi*	CLF.PILE
-híwi	CLF.POINTED	*-wálhia*	CLF.YEAR
-i	CLF.BASKET	*-xaa*	CLF.EXCREMENT
-íida	CLF.HALF	*-yáwa*	CLF.HOLE

The classifier strata display some interesting differences when it comes to phonology, semantics, and morphology, summarized in Table 8.5.

Table 8.5: Properties of the oldest and youngest classifier strata

	Oldest strata	Youngest stratum
Phonology	shorter forms	longer forms
Semantics	more general	more specific
Morphology	all have complex form	not all have complex form

In terms of phonology, more grammaticalized forms tend to be more phonetically eroded (§ 2.2). The two reconstructable strata have an average number of syllables of 1.6, and the youngest stratum of 2.06. Classifiers in the reconstructable strata are maximally disyllabic; all the trisyllabic classifiers (§ 5.1) are found either in the youngest stratum or among the classifiers of unknown age. Most of the trisyllabic classifiers are analyzable etymologically (see § 8.3).

The youngest stratum includes many classifiers for parts of objects, in particular plants (e.g., *-éekhe* 'CLF.SMALL.SEED', *-íiwi* 'CLF.FLOWER', and *-kénaa* 'CLF.BRANCH';

see further in § 8.2.2), whereas the two oldest strata primarily consist of shape classifiers and classifiers with quite general semantics (e.g., *-híko* 'CLF.LONG', *-da* 'CLF.GENERIC', and *-khaa* 'CLF.CURVILINEAR')—the notable exception being the highly specific *beiju* classifier *-híwa*.

As shown in § 5.3.2, most classifiers have a complex form with an additional nominalizing suffix used in certain contexts. These suffixes are hypothesized by Ramirez (2020a: 371–375) to be remnants of an older tripartite gender marking system. A small set of classifiers never take an additional suffix in these contexts; these classifiers are all found either in the youngest stratum or among the classifiers of unknown age. The ability of classifiers to use the same form in all contexts appears to be an innovation, and the classifiers that do not take a suffix in the complex form appear to have developed into classifiers after the suffixes used in the complex form had already started to lose their meaning.

8.2 From bound nouns to classifiers

There is broad consensus on the view that classifiers develop from lexical sources, primarily bound nouns, and come to function as classifiers through a process of grammaticalization; this has been proposed for Baniwa (Ramirez, 2020a: 239–241), for the Arawakan languages more generally (Aikhenvald, 2019: 117–129; Dunn & Rose, forthcoming: 34), as well as for other unrelated languages in Northwestern Amazonia (Payne, 1986; Seifart, 2005; Stenzel, 2013: 128 ff.; Epps & Obert, 2022). This is in line with the general typological tendencies, as nouns are the primary source of classifiers cross-linguistically (Aikhenvald, 2000: 353 ff.). The fact that bound nouns seem particularly prone to develop into classifiers is likely related to their morphosyntactic properties, as we will see below.

As shown in § 2.8.1, a classifier system (marked on numerals and used in compounding and verb incorporation) is likely to have existed since PA (Dunn, 2022). It is also clear that new classifiers have developed over time in Baniwa, and a probable bridging context from bound nouns to classifiers is that of compounds. As shown in § 3.2.5.2, Baniwa compounds typically consist of a free noun (first component) and a bound noun (second component), where the bound noun is either prefixed with connecting *i-* or attaches directly to the preceding noun, as in (124)[95].

[95] See Example (12) in § 3.2.5.2 for an illustration of the two parallel constructions.

(124) *iniphiwída*
inípo-**hiwída**
path-**head**

'harbour' (lit. 'head (=end) of a path') [200323_ELIC_F_01]

In some cases, similar constructions are compatible with two analyses: they can be compounds consisting of a free noun and a bound noun (125b), or they can be derived nouns consisting of a free noun with a classifier suffix (125a). The ambiguity stems from the fact that the morpheme *-ya* is used both as a bound noun, as evidenced by its ability to occur with a person prefix, and as a classifier, as evidenced by its ability to occur on a numeral (see Table 8.2). In this case, neither the form nor the meaning permits us to decide conclusively on either of the analyses.

(125) *haikóya* 'tree bark' [221006_ELIC_G_01]

a. *háiko-**ya***
tree/wood-**skin/bark**

b. *háiko-**ya***
tree/wood-**CLF.SKIN**

As both bound nouns and classifiers can occur in the same position in such constructions[96], it is easy to see how bound nouns could grammaticalize into classifiers in this context. Regarding the other contexts where classifiers appear in Baniwa, the possible paths of grammaticalization are not as clear, as they do not have parallel constructions with bound nouns that could serve as source constructions, at least not synchronically. For example, numerals, adjectives, and verbs do not combine with nouns through compounding or noun incorporation processes. Ramirez (2020a: 239) proposes a scenario for the development of new numeral classifiers in Baniwa through a repeater mechanism, where the bound noun is copied onto a modifying numeral (see below for a similar proposal by Aikhenvald, 2019, but for the North Arawakan languages).

For Arawakan languages in general, Rose (2024b) explores the possibility that they may have grammaticalized in parallel in different loci (on numerals, nouns, adjectives, and verbs), based on evidence from four different languages which have corresponding source constructions with nouns to support such a hypothesis (see also Rose & Van linden, 2023; Dunn & Rose, forthcoming). It is difficult to assess the plausibility of parallel development in Baniwa due to the synchronic

[96] See Example (35) in § 4.3.1 for more examples of similar constructions.

lack of such source constructions for other contexts than nouns. However, this does not necessarily exclude it as an explanation for the historical development of the classifier loci, as bound nouns may have been used in these positions at earlier stages. Occasional lexicalized forms invite such interpretations; for example, the noun *dzamápali* '*zagaia*', designating a spear used for fishing, appears to be etymologically comprised of the numeral *dzama-* 'two' and the bound noun *-pali* 'root' (lit. 'two roots'), possibly because the spear is constructed by splitting it at one end in order to accomodate the spearhead. The composition of this noun may reflect an earlier stage where it was possible more broadly for numerals and bound nouns to form compound constructions.

Aikhenvald (2019: 133 ff.) proposes a different historical scenario for the North Arawakan languages, where bound nouns in compound constructions similar to the ones in (124) and (125) give rise to agreement markers first on numerals, and then on adjectives and other targets. She also proposes an alternative scenario where numerals was the first classifier context for these languages. Both alternatives are based on the fact that numerals are the most common classifier locus in the North Arawakan languages in her sample. Dunn (Dunn, 2022) draws similar conclusion on family level. However, as the more recent family-wide and more comprehensive comparison by Dunn & Rose (forthcoming) suggests that classifiers may have been present in multiple loci already in PA, the diachronic primacy of the numeral locus in North Arawakan languages is called into question.

8.2.1 The coevolution of classifiers and nouns

Classifiers and bound nouns coexist in Baniwa, and appear to have done so for a long period of time: both classifiers and bound nouns seem to have been present in Proto-Arawak (Payne, 1991; Aikhenvald, 1999; Dunn, 2022). Over time, nouns and classifiers feed off each other: bound nouns always provide a source for classifiers to grammaticalize from, and classifiers are in turn regularly suffixed to nouns to derive new nouns. For example, as illustrated in (126), the free noun *hiipáda* is probably derived from a combination of the free noun *híipa* 'rapids' (i.e., rocky parts of rivers) and the round/generic classifier *-da*. In turn, as illustrated in (127), *hiipáda* 'stone' gave rise to the classifier *-hipáda* 'CLF.PIECE' (Ramirez, 2020a: 243) via its bound noun version *-hipáda* 'piece' (see § 8.1)[97].

[97] *Hiipáda* is one of the few free nouns whose conversion into a bound noun does not involve any of the alienable possession suffixes (§ 3.2.5.1; see also Ramirez, 2020a: 131). Note the change in meaning from *hiipáda* 'stone' to *-hipáda* 'piece'.

(126) Classifier deriving new noun

híipa ‘rapids’ + *-da* ‘CLF.GENERIC’ → *hiipáda* ‘stone’

(127) Noun developing into classifier

hiipáda ‘stone’ → *-hipáda* ‘piece’ → *-hipáda* ‘CLF.PIECE’

8.2.2 Classifiers as a product of their origins

The fact that bound nouns are the main source of classifiers likely has implications for the semantic distinctions found within the classifier set. Bound nouns (§ 3.2.5) in Baniwa typically include meanings relating to specific parts of things, like plants or the body. Table 8.6 lists the classifiers that originate (or are hypothesized to originate) in either plant or body parts terms. Plant parts in particular appear to be very prone to develop into classifiers (see further about plant part classification in § 7.2.4.1); body parts to a slightly lesser extent. Sometimes the origin is ambiguous between plant and body part, as in the case of *-phe* ‘leaf; feather’ and *-ya* ‘skin; bark’.

On the other hand, kinship relations also constitute a significant proportion of the bound noun set, and yet there are no known examples of kinship terms that have grammaticalized into classifiers in Baniwa.

It also seems possible for free nouns to develop into classifiers as long as they are first converted into bound nouns, which any free noun can be (see § 3.2.5). Thus, also meanings typically associated with free nouns can grammaticalize into classifiers, at least theoretically, although only three such example have been found so far in Baniwa (see *-pi*, *-hipáda*, and *-hípani* in § 8.3). Thus, free noun semantics do not affect the semantic distinctions of classifiers on system-level the way bound noun semantics do.

Terms denoting both plant and body parts are a common source of classifiers (see, e.g., Aikhenvald, 2000: 355). They occur across the Arawakan family, and there is a tendency for them to extend to cover other objects of similar shapes (Dunn & Rose, forthcoming: 14–15, 34). Baniwa’s closest relative Tariana shares some of the classifiers in Table 8.6 (e.g., *-phe* ‘leaf-like’; Aikhenvald, 1994: 454), but also appears to have developed additional body part classifiers that are not found in Baniwa (Aikhenvald, 2003: 99).

Table 8.6: Classifiers originating in plant and body part terms

Classifier	Gloss	Origin
Classifier < plant part		
-éekhe	CLF.SMALL.SEED	< *-éekhe* 'seed'
-iítsia	CLF.BUNCH	< *-iítsia* 'bunch'
-íiwi	CLF.FLOWER	< *-íiwi* 'flower/thorn'
-íixi	CLF.SEED	< *-íixi* 'seed/kernel'
-kénaa	CLF.BRANCH	< *-ke/-kénaa* 'branch'
-pokóda	CLF.STUMP	< *-pokóda* 'tree strump'
-wáta	CLF.BRACT	< *-wáta* '*curuatá*'
-wáthe	CLF.NODE	< *-wáthe* 'node'
Ambiguous/combination		
-phe	CLF.LEAF	< *-phe* 'leaf, feather'
-ya	CLF.SKIN	< *-ya* 'skin, bark'
-yáwa	CLF.HOLE	< ? *-ya* 'skin, bark' + < ? PJC **-wawa* 'heart, abdomen'
Classifier < body part/bodily excretion		
-áapi	CLF.HOLLOW	< ? PJC **inapi/-aapi* 'bone'
-éewhe	CLF.EGG	< *-éewhe* 'egg'
-híwi	CLF.POINTED	< *-hiwí(da)* 'head'
-wa	CLF.SPACE	< ? PJC **-wawa* 'heart, abdomen'
-xaa	CLF.EXCREMENT	< *-íixa* 'excrement'

Similar patterns can be found in unrelated languages in the Upper Rio Negro region. Tucanoan languages have nominal suffixes for both plant and body parts that appear to be mid-way between bound nouns and classifiers (for example, Barnes (1990) analyzes them as classifiers, but Ramirez (1997) as nouns). Hup (Naduhup) has an incipient system of classifiers that have developed primarily from bound nouns denoting plant parts (Epps, 2007).

Semantic similarities in the classifier systems of Amazonian languages may thus, at least in part, be due to the widespread existence of a class of bound nouns (Krasnoukhova, 2012: 57), which commonly encompasses body parts and other part–whole relations (Nichols, 1992: 119–120; Ortmann, 2018: 102), and which in turn tend to develop into classifiers.

8.3 Etymological origins of individual classifiers

Table 8.7 presents an overview of the proposed and tentative etymological origins of some of the Baniwa classifiers, most of which are compiled from other sources (Payne, 1991; Aikhenvald, 2019; Ramirez 2020a; 2020d; Dunn, 2022). The etymologies are discussed in more detail below.

Table 8.7: Etymological origins of classifiers (based on Payne, 1991; Aikhenvald, 2019; Ramirez 2020a; 2020d; Dunn, 2022)

Classifier	Gloss	Origin
Classifiers reconstructable as classifiers in PJC/PA		
-áanhaa	CLF.LIQUID	< PJC **-aa* 'water' (CLF) < PA ***sa* 'liquid' (CLF) < ***isa-pha* 'wet' (ADJ?)
-koa	CLF.SURFACE	< PJC **-kʊa* 'limited area' (CLF) < PA ***kwa* 'flat' (CLF)
-máka	CLF.FABRIC	< PJC **-maka* 'fabric, cloth' (CLF) < PA ***ma* 'cloth' (CLF)
-na	CLF.TRUNK	< PJC **-na* 'trunk, mammal' (CLF) < PA ***na* 'large, long' (CLF)
-pi	CLF.TUBE	< PJC **-pɨ* 'long tube, rope, liana' (CLF) < PA ***pi* 'long, thin' (CLF) < PA ***apʰi* 'snake' (FREE N)
-áapi	CLF.HOLLOW	< PJC **-aapi* 'container' (CLF) (< ? PJC **inapi/-aapi* 'bone')
-áapo	CLF.STICK	< PJC **-aapʊ* 'long, flexible' (CLF)
-da	CLF.GNR	< PJC **-aʔa-la* 'round' (CLF)
-eéma	CLF.SIDE	< PJC **-eema* 'side' (CLF)
-híko	CLF.LONG	< PJC **-Sikʊ* 'tube' (CLF)
-híwa	CLF.BEIJU	< PJC **-Siwa* 'manioc bread' (CLF)
-iíta	CLF.HUMAN	< PJC **-iita* 'human' (CLF)
-khaa	CLF.CURV	< PJC **-kaSa* 'threadlike' (CLF)
-póko	CLF.CIRCLE	< PJC **-pʊkʊ(i)* 'circular' (CLF)
-wa	CLF.SPACE	< PJC **-wa* 'hole' (CLF) (< ? PJC **-wawa* 'heart, abdomen')
Classifiers with synchronically identical bound nouns		
-éekhe	CLF.SMALL.SEED	< *-éekhe* 'seed' (BOUND N)
-éewhe	CLF.EGG	< *-éewhe* 'egg' (BOUND N)
-hipáda	CLF.PIECE	< *-hipáda* 'piece' (BOUND N) < *hiipáda* 'stone' (FREE N)
-iítsia	CLF.BUNCH	< *-iítsia* 'bunch' (BOUND N)
-íiwi	CLF.FLOWER	< *-íiwi* 'flower/thorn' (BOUND N)
-íixi	CLF.SEED	< *-íixi* 'seed/kernel' (BOUND N)
-kénaa	CLF.BRANCH	< *-ke/-kénaa* 'branch' (BOUND N)
-kódzoa	CLF.BEND	< *-kódzoa* 'curve' (BOUND N)
-phe	CLF.LEAF	< *-phe* 'leaf, feather' (BOUND N)
-pokóda	CLF.STUMP	< *-pokóda* 'tree strump' (BOUND N) < ? *-póko* 'CLF.CIRCLE' (CLF) + *-da* 'CLF.GNR' (CLF)
-ttáwalhe	CLF.CUT	< *-ttáwalhe* 'piece' (BOUND N)
-ttówhia	CLF.ROOM	< *-ttówhia* 'room' (BOUND N)
-wána	CLF.SLICE	< *-wána* 'slice' (BOUND N)
-wáta	CLF.BRACT	< *-wáta* '*curuatá*' (BOUND N) < ? Nheengatu *carauatai* 'id.'
-wáthe	CLF.NODE	< *-wáthe* 'knot' (BOUND N)
-ya	CLF.SKIN	< *-ya* 'skin, bark' (BOUND N)
Classifiers with a tentative etymology		
-Ø	CLF.CANOE	< ? Dial. *-hwya* 'concave objects' (CLF) / ? *-ya* 'CLF.SKIN'
-aápa	CLF.OBLONG	< ? PA ***pa/ap* 'long, curved' (CLF)
-áaphi	CLF.AREA	< *-aaphítte* 'part below' (BOUND N)
-dápana	CLF.HOUSE	< ? *-da* 'CLF.GNR' + *-pana* 'house' (BOUND N)
-hípani	CLF.RAPIDS	< *híipa* 'rapids' (FREE N) + *-ni* '?'
-híwi	CLF.POINTED	< *-hiwí(da)* 'head' (BOUND N) < PJC **-Siw-* 'head' (BOUND N)
-ko	CLF.HAMMOCK	< ? *-koa* 'to lie (in a hammock)' (V)
-wálhia	CLF.YEAR	< ? *waali-* 'new, young' (< PJC **waari* 'new')
-xaa	CLF.EXCR	< *-íixa* 'excrement' (BOUND N)
-yáwa	CLF.HOLE	< ? *-ya* 'CLF.SKIN' + *-wa* 'CLF.SPACE'

Fifteen classifiers are reconstructable as classifiers in PJC. Five of these are also reconstructable to PA, again *as* classifiers (see § 8.1). Thus, if these classifiers have developed from lexemes belonging to other word classes (such as bound nouns, see § 8.2), this development must have taken place at some earlier stage in history. For two of the five Baniwa classifiers that are reconstructed to PA, a lexical origin has been proposed. Payne (1991: 383) posits the form ***ap^{h}i* 'snake' as the source of the long and thin classifier ***pi* (Dunn's form), and Dunn (2022: 55) posits the the form ***isa-pha* 'wet' (Payne's form) as the source of the liquid classifier ***sa*. It is not clear exactly what word class status these items had in PA, but ***ap^{h}i* was presumably a free noun and ***isa-pha* may have been an adjective, based on their semantics. For two of the PJC classifiers, Ramirez (2020d: 96–97) notes a connection to lexical items for body parts: the container classifier **-aapi* 'container' + the free noun **inapi* 'bone' (which also has the irregular bound version **-aapi*), and the hole classifier **-wa* + the bound noun **-wawa* 'heart, abdomen'. The fabric classifier **-máka* is clearly related to the the widespread Arawakan noun *hamaka* 'hammock'[98] (Aikhenvald, 2019: 120; Ramirez, 2020d: 97), but it is unclear whether the noun developed into the classifier or the other way round (see § 8.2.1). These forms may also be related to a verb ***imaka* 'to sleep' Payne (1991: 418).

Sixteen classifiers are identical in form to bound nouns in current use with identical or highly similar meanings. It is assumed here that the bound nouns in question have given rise to the corresponding classifiers (see § 8.2). Therefore, this set of classifiers will not be discussed further below, but a brief deviation will be made here to propose a speculative, but possibly noteworthy etymology of *-wáta* 'CLF.BRACT': this may be a reduced form of a borrowing from Nheengatú *carauatai* (Stradelli, 1929: 400); in fact, the same word that gave rise to local Portuguese *curuatá*. If this is the case, it would be the only known classifier in Baniwa to have developed from a loanword.

The remaining classifiers are those which have neither a reconstructable origin as classifiers, nor a synchronic identical counterpart in a bound noun. Ten of these may be tentatively connected to some other item in the lexicon of Baniwa, a related language, or a reconstruction. They are discussed below.

8.3.1 *-Ø* 'CLF.CANOE'

As zero morphemes contain no phonological material, it is in principle impossible to establish an etymology for *-Ø* 'CLF.CANOE'. However, based on the in-

[98] English *hammock* is also of Arawakan origin, borrowed via Spanish from Taíno.

terchangeability of *-Ø* with the skin/bark classifier *-ya* (see § 7.4, § 7.1.1, and § 7.1.52), it is possible that it is related to the form *-whia* [ʍ̥ia], a classifier for concave objects appearing in the Kumandene and Ayanene varieties (Valadares, 1993: 127; Aikhenvald, 2007: 498). Tariana has the related form *-hwya*, which functions as a classifier for canoes (Aikhenvald, 2007: 498).

Apart from canoes, *-Ø* is also used for calabash bowls of the *kóoya* and *átta* types, made from hollowed out fruits. A connection between these objects and skin and bark might stem from their common 'thin outer layer' sense. This is perhaps most clearly visible in the bowls, where only the peel of the fruit is left, but dugout canoes are made according to the same basic principle.

The connection between canoes and bark is further strengthened by the existence of a canoe type in the Upper Rio Negro region that is made from a large sheet of bark that is stripped off a tree[99]. Bark canoes in this region appear to be associated primarily with non-riverine groups like speakers of Dâw and Nadëb (Naduhup), where they show up in plenty of stories (Assis, 2001; P. Epps and K. Obert, personal communication, October 7–9, 2024), but Wright (1981: 102–103) reports that the Baniwa-speaking Hohodene sib state that their ancestors did not have dugout canoes, but instead used long strips of bark for shorter journeys on water. Interestingly, the use of the skin/bark classifier *-ya* for canoes is primarily associated with the dialectal varieties spoken on the Aiari river (see Figure 1.1 in § 1.1), which is exactly where the Hohodene live.

8.3.2 *-aápa* 'CLF.OBLONG'

The classifier *-aápa* 'CLF.OBLONG' may be a reflex of the PA classifier ***pa/ap* 'long, curved', as reconstructed by Dunn (2022: 45), although it is not included in the oldest classifier stratum in for lack of certainty (see § 8.1). The oblong meaning of Baniwa *-aápa* speaks in favour of it as a reflex, especially its use with bananas, which is also reported for the cognate set. Aikhenvald (2019: 119, 137) possibly connects the Baniwa and Tariana classifiers, but she sometimes reports the form *-pa* and sometimes *-apa* for Tariana (both with the meaning 'largish long'). The reason to be cautious about positing a PA origin for this classifier is twofold. First, the classifier is not reconstructed to the intermediate stage PJC (Ramirez, 2020d: 96–97). Second, there is some doubt on Dunn's part about the phonological form of the reconstruction, as the reflexes are either VC (***pa*) or CV (***ap*). Perhaps the VCV form in Baniwa suggests a reconstruction ***apa* which could be a common origin of both sets.

[99] Bark canoes are also found in other parts of South America and the world (Arnold, 2017).

8.3.3 *-áaphi* 'CLF.AREA'

Ramirez (2020a: 241) posits that *-áaphi* 'CLF.AREA' stems from *-aaphítte*, a composite form consisting of the relational noun *-aapi* 'below' and the ablative suffix *-hitte*.

8.3.4 *-dápana* 'CLF.HOUSE'

Ramirez (2020a: 241) also suggests that *-dápana* 'CLF.HOUSE' stems from *-Vpana* 'house'[100]. The preceding *-da* may be the generic classifier (as suggested by Aikhenvald, 2019: 117), which would entail that etymologically, one classifier is stacked on another (see § 5.3.5).

8.3.5 *-hípani* 'CLF.RAPIDS'

This classifier quite likely stems from the free noun *híipa* 'rapids' + the inalienator suffix *-ni* (Ramirez, 2020a: 241), which converts free nouns into bound ones (§ 3.2.5.1). The form *-hípani* also occurs as the second part in compound toponyms (128).

(128) *Dzatthépani*
dzáatte-hípani
toucan-**CLF.RAPIDS**

'Cachoeira-do-Tucano' (toponym) (adapted from Ramirez, 2020a: 87)

8.3.6 *-híwi* 'CLF.POINTED'

The classifier *-híwi* 'CLF.POINTED' seems to go back to PJC **-Siw-* 'head' (Ramirez, 2020d: 74), probably via the extended sense 'end' (see, e.g., (124)). The present-day Baniwa form *-hiwída* 'head' likely contains the generic classifier *-da* as a fossilized suffix (Ramirez, 2020a: 241).

[100] V in *-Vpana* represents an underlying vowel which takes on the quality of the final vowel of the host word and lengthens it, e.g., *nóopana* < *no-Vpana* /1SG-house/ 'my house' (Ramirez, 2020a: 89)

8.3.7 *-ko* 'CLF.HAMMOCK'

Ramirez (2001a: 180) lists the verb *-koa* 'to lie (in a hammock)' (< PJC **-kʊʔa* 'id.', Ramirez, 2020d: 46), which may be related to the classifier *-ko* 'CLF.HAMMOCK' (although he does not make this connection). As verbs do not generally appear to give rise to classifiers in Baniwa, it is possible that the classifier developed from some related nominal form that is no longer in use[101] (cf. the verb–noun connection discussed in relation to *-máka* 'CLF.FABRIC', § 8.3). Tariana has a similar classifier *-ku* 'folded cloth' (Aikhenvald, 2019: 137).

8.3.8 *-wálhia* 'CLF.YEAR'

The year classifier *-wálhia* may have its origins in the adjective *waali-* 'new, young' (< PJC **waari* 'new', Ramirez, 2020d: 85), which shares a temporal meaning. This hypothesis is strengthened by the existence of other formally similar items with time-related reference, like the morpheme *-wali* 'times' (§ 4.3.3) and the simultaneous distal verbal suffix *-kawálhi* 'at the time when'. The form of *-wálhia* is consistent with the analysis *waali-* + the suffix *-haa*, which attaches to adjectives and modifies their meaning, usually translatable as 'a little': *íitta* 'black' vs. *ítthaa* 'a little black'; *hámo* 'hot' vs. *hámhoa* 'lukewarm' (Ramirez, 2020a: 267–268). However, in some adjectives it produces a different kind of meaning change, e.g., *iirai* 'red' vs. *iráixaa* 'reddened'. It is possible to imagine an analogous situation along the lines of *waali-* 'new' vs. *-wálhia* '(re)newed', which could fit with the cyclic sense of the classifier.

8.3.9 *-xaa* 'CLF.EXCREMENT'

The classifier *-xaa* 'CLF.EXCREMENT' has developed from the bound noun *-íixa* 'excrement'. The complex form of this classifier still retains the short [a], and sometimes also the initial vowel of the bound noun (e.g., *makéexali* < *maka-**íixa**-li* /big-CLF.EXCR-NF/ 'big excrement', see § 5.3.2).

8.3.10 *-yáwa* 'CLF.HOLE'

The classifier *-yáwa* 'CLF.HOLE' is suggested by Ramirez (2020a: 165) to possibly consist of two historically stacked classifiers: *-ya* 'CLF.SKIN' and *-wa* 'CLF.SPACE'.

[101] But see Mihas (2019: 50–51) for two classifiers with suggested verbal origins in the Kampan branch of the Arawakan family.

The connection between *-yáwa* and *-wa* is clear, as they are semantically close and overlap to a great extent in terms of the nouns they are used with (see § 7.1.46, § 7.1.53). The connection between *-yáwa* and *-ya* is less clear, but perhaps strengthened by the seemingly parallel connection between *-Ø* and *-ya* (§ 8.3.1), as both *-Ø* and *-yáwa* refer to hollow referents.

8.3.11 Classifiers of unknown origin

The remaining 12 classifiers[102] are treated here as classifiers of unknown origin, although some of them may have connections to other items.

For *-i* 'CLF.BASKET', a classifier which is also used for bunches of fruit, Ramirez (2020a: 243) suggests that it may come from *-iítsia* 'bunch' (thus sharing an origin with the classifier *-iítsia* 'CLF.BUNCH'). There is too little phonological material in *-i* to determine the validity of this etymology.

Ramirez (2020a: 242, 247) mentions in passing a possible connection between the pair classifier *-ma[1]* and the numeral *dzama-* 'two', seemingly suggesting that this numeral comprises *-ma[1]* as a derivational suffix. As *dzama-* goes back to PJC **(pʊi)jama* and has cognates in other Arawakan languages that also feature the *-ma* element (Ramirez, 2020d: 98), this derivation must have happened a long time ago. However, *-ma[1]* is not one of the classifiers reconstructed to PJC. Thus, the origin of this classifier is still unresolved.

The classifiers *-páwa* 'CLF.RIVER', *-yáwa* 'CLF.HOLE', and *-tsoi* 'CLF.PILE' will be treated together. Ramirez (2020a: 164–165) lists these as nominal suffixes that can attach to nouns in compound-like constructions, but cannot be used as bound nouns independently (for instance, with just a possessor prefix). The form *-páwa(ni)* 'creek' appears as the second part in some compound toponyms (e.g., *Koomadanaípawa*) and on numerals and adjectives. The form *-tsoi* 'CLF.PILE' can also appear as a second part in compounds (e.g., *hipadátsoi*, 'pile of stones'). The form *-yáwa* 'CLF.HOLE' is simply reported to be suffixed to numerals and adjectives. It is not clear how these "nominal suffixes" differ from classifiers, as classifiers can be used in all of these environments. Thus, these suffixes cannot be treated as etymological origins, as they simply refer to the classifier forms.

Bound nouns are indeed the most common lexical source for classifiers to grammaticalize from. Naturally, this is true for the entire set of classifiers that have

[102] These are *-áana* 'CLF.GROUP', *-daa* 'CLF.DAY', *-hípa* 'CLF.MALE', *-i* 'CLF.BASKET', *-íida* 'CLF.HALF', *-ma[1]* 'CLF.PAIR', *-ma[2]* 'CLF.FEMALE', *-náko* 'CLF.BUNDLE', *-pa* 'CLF.PACKAGE', *-páwa* 'CLF.RIVER', *-péko* 'CLF.PATH', and *-tsoi* 'CLF.PILE'.

identical bound noun forms synchronically. It is also (likely) the case for many of the classifiers that are given tentative etymologies, such as *-áaphi* 'CLF.AREA', *-dápana* 'CLF.HOUSE', *-híwi* 'CLF.POINTED', and *-xaa* 'CLF.EXCREMENT'. One classifier, *-hípani* 'CLF.RAPIDS', is ultimately connected to a free noun. A free noun (***ap^{h}i* 'snake') also appears to be the source of *-pi* 'CLF.TUBE', but it grammaticalized into a classifier already in PA. Two classifiers (one of them in PA) are tentatively linked to adjectives: *-áanhaa* 'CLF.LIQUID' < PA ***sa* 'liquid classifier' < ***isa-pha* 'wet', and *-wálhia* 'CLF.YEAR' < ? *waali-* 'new, young'. Another set of classifiers can only be connected to other classifier morphemes, either in Baniwa (*-yáwa* 'CLF.HOLE'), in related dialects/languages (*-Ø* 'CLF.CANOE'), or in reconstructed stages (*-aápa* 'CLF.OBLONG'). Finally, one classifier (*-ko* 'CLF.HAMMOCK') can be connected to a verb (*-koa* 'to lie (in a hammock)').

8.4 Chapter summary

The classifier system of Baniwa has developed over a long period of time; it has both ancient roots and members of much more recent origin. Five classifiers can be reconstructed all the way to Proto-Arawak, and an additional 10 to Proto-Japurá-Colombia. Thus, about a third of the classifiers are several thousand years old. The rest of the classifiers appear to have developed into classifiers more recently. In particular, there is a set of 16 classifiers whose lexical origins are extremely transparent, to the point that there is an identical form in use as a bound noun simultaneously. The different age strata of classifiers tend to correlate with other properties: reconstructable classifiers have shorter forms and more general semantics than the classifiers in the youngest age layer, which are also used as bound nouns synchronically. These differences are in line with the general principles of grammaticalization (§ 2.2). In addition, all reconstructable classifiers take one of the additional nominalizing suffixes in their complex form, whereas only some of the classifiers in the youngest stratum do so. This lends support to Ramirez's (2020a: 371–375) hypothesis that these nominalizers are remnants of an old gender marking system which is no longer productive.

New classifiers have developed from bound nouns over time through a pathway of compounding, and possibly through other pathways in the past. The dominance of bound noun origins is clear, but some other lexical sources, such as free nouns, adjectives, and verbs, have been proposed for a minority of classifiers.

Chapter 9

Contact

The Baniwa speakers living in São Gabriel da Cachoeira are generally bilingual in Portuguese (see § 1.1.2.1), which is also the case for all participants in this study (see § 1.2.3). Code-switching between the two languages is common, as appears to be the case for other Indigenous languages spoken in the region (Epps, 2018: 164–166). This chapter explores how the classifier system of Baniwa is affected by the lexical influence from Portuguese, the majority language in the urban environment of São Gabriel da Cachoeira. Lexical borrowing challenges the classifier system in two ways. First, borrowed nouns, typically for recently introduced items and concepts, need to be handled by the classification system in some way (§ 9.1). Second, classifiers may need to attach to borrowed hosts (§ 9.2). These patterns are described in this chapter.

9.1 Classification of borrowed nouns

The Noun list data set (§ 1.2.4) contains more than 100 borrowed nouns (all free nouns). The majority of these are recent Portuguese loans which have not been integrated phonologically, e.g., [*avião*] 'airplane' and [*polícia*] 'police' (given in brackets, so as to underline that they do not follow Baniwa orthography). There are also examples of Portuguese loanwords that show phonological integration, e.g., *kamítsha* 'shirt' < *camisa* 'id.' and *garáapha* 'bottle' < *garrafa* 'id.'. There are also a handful of loanwords from Spanish, e.g., *paláata* 'money' (< *plata* 'silver; money') and Nheengatu, e.g., *áawi* 'needle' (< *awi* 'id.') (Ramirez, 2001a).

When Portuguese nouns are counted in Baniwa, they behave as Baniwa nouns in that they are not marked for number (see § 3.2.5.3). An example is shown in

(129) using the singular form *filme*, not the plural **filmes*.

(129) *dzamáda* [***filme***] *liwadzákakawa*
dzama-da [**filme**] li-wadzáka-ka-wa
two-CLF.GNR **film** 3NFSG-to.finish-SUB-MID

'He finished (watching) two movies.' [200326_ELIC_F_02]

When it comes to semantics, physical shape is the most important parameter for classifier assignment to non-human referents (see § 7.2). Loanwords are no exception to this principle: they behave just as native words in that they are classified by shape, as shown by the examples in (130). A cell phone takes the human classifier *-iíta*, which is also regularly used for objects with flat parts (130a), a bank note takes the leaf classifier *-phe* (130b), and a key takes the pointed classifier *-híwi* (130c).

(130) a. *apáita* [*celular*]
apa-**iíta** [celular]
one-**CLF.HUMAN** cell.phone

'one cell phone' [200228_ELIC_F_01]

b. *apáphe* *paláata*
apa-**phe** paláata
one-**CLF.LEAF** money

'one bank note' [200324_ELIC_F_01]

c. *aphéwi* [*chave*]
apa-**híwi** [chave]
one-**CLF.POINTED** key

'one key' [200305_ELIC_F_01]

The semantic domain of vehicles offers some interesting clues to the semantics of the canoe classifier *-Ø*. With the exception of *íita* 'canoe' (the only traditional vehicle), these are referred to by loanwords in Baniwa, whether phonologically integrated (*káaro* 'car') or not ([*avião*] 'airplane'). As Table 9.1 shows, the vehicles that take the canoe classifier *-Ø* are all container-like in shape: the only two vehicles that take different classifiers are bicycles and motorcycles. Thus, the extension of *-Ø* to include these particular vehicles is not only based on function, but also on shape.

Table 9.1: Vehicles and their associated classifiers

Form	Meaning	Classifier	Gloss
íita	'canoe'	*-Ø*	CLF.CANOE
káaro	'car'	*-Ø*	CLF.CANOE
[*avião*]	'airplane'	*-Ø*	CLF.CANOE
[*ônibus*]	'bus'	*-Ø*	CLF.CANOE
[*trêm*]	'train'	*-Ø*	CLF.CANOE
[*lancha*]	'boat'	*-Ø*	CLF.CANOE
[*bicicleta*]	'bicycle'	*-póko*	CLF.CIRCLE
[*moto*]	'motorcycle'	*-da*	CLF.GENERIC

An intriguing question is how extensions like these affect the internal structure of *-Ø* as a category. This category traditionally consisted of carved, hollowed-out items (§ 7.1.1), reflecting an intimate awareness of how artefacts are manufactured (see also Costa Oliveira, 2015: 316). In the case of motorized vehicles like airplanes, there is no relationship to the manufacturing process, eliminating the salience of this particular aspect of objects. Instead, many vehicles are assigned to this category by a function + shape association with the canoe as a container-like vehicle. It is possible that the inclusion of vehicles in this category will shift the focus away from the items' manufacturing process to their functional properties (even if the shape dimension appears to be largely intact).

In Tariana, the canoe classifier *-hwya* is also used with more recently introduced vehicles, but as a derivational device to form nouns for, e.g., airplanes (131a) and cars (131b) (Aikhenvald, 2019: 121–122). This supports the analysis that it is indeed the canoe meaning of the zero morpheme *-Ø* that is extended in Baniwa, and not just absence of classifier marking.

(131) Tariana (adapted from Aikhenvald, 2019: 121–122)

a. *kara-ka-**whya***
REL.fly-TH-**CLF.CANOE**

'airplane'

b. *ka-kolo-ka-**whya***
REL-roll-TH-**CLF.CANOE**

'car'

Another interesting case is that of plants and their parts. As shown in § 7.2.4.1, Baniwa has a very rich system for differentiating parts of plants with the use of classifiers, built on the intimate knowledge of and dependence on the plants in

question. This high level of differentiation is, unsurprisingly, lost when it comes to imported produce that is bought in the market rather than harvested. An example is oranges, which are categorized as round and whose juice is categorized as a liquid, but the plant itself is irrelevant. Another example is peanuts; the edible seeds are also categorized as round, but as one speaker remarked, "nobody knows what the plant looks like".

The data presented in this section shows that loanwords (even the recent ones that have not undergone phonological adaptation) are well integrated in the classifier system in the sense that they behave just as native nouns. They are classified by the same semantic principles, they can take classifiers as suffixes, and they do not show plural marking when combining with numerals higher than 1.

9.2 Classifiers on borrowed hosts

The previous section dealt with the semantic question of how borrowed nouns, and by extension recently introduced concepts, are classified. This section explores the suffixation of classifiers to borrowed host words, which is primarily a formal issue.

Classifiers attach to borrowed words belonging to the word classes that can act as classifier hosts (see § 5.3.1), as shown in (132). (132a) features the leaf classifier *-phe* on the noun [*dopará*] (< Portuguese *açaí do Pará*'). (132b) shows the complex form of the generic classifier *-da* on the borrowed adjective *ríiko* (< Portuguese *rico* 'rich'). In (132c), the borrowed verb *-kánta* (< Portuguese *cantar* 'to sing') is nominalized with the location nominalizer and the generic classifier into a noun referring to the church.

(132) a. [*dopará*]*phe*
[dopara]-**phe**
açaí.do.Pará-**CLF.LEAF**

'*açaí* leaf' [200308_NARR_B_01]

b. *apáda* *dzakálee* *riikódali*
apa-da dzakálee ríiko-**da-li**
one-CLF.GNR village/town rich-**CLF.GNR-M**

'one rich town' [200324_ELIC_F_01]

c. *pakántakaroda*
pa-kánta-**karo-da**
IMPRS-to.sing-**NMLZ.LOC-CLF.GNR**
'church (lit. 'singing place')' [200328_ELIC_F_02]

The rest of this section will be devoted to classifier use on one specific kind of borrowed host, namely Portuguese numerals. The phonological structure of these unintegrated loanwords make for an interesting case study on how morphophonological processes are affected by lexical borrowing.

9.2.1 Classifiers on Portuguese numerals

In everyday use, the Baniwa numeral system (§ 3.2.9.1) is commonly replaced by numerals borrowed from Portuguese from 4 and up. The example in (133) illustrates this split within the same utterance: Baniwa numerals are used up to 'three' (*dzama-* 'two', *madali-* 'three'), and Portuguese numerals are used for higher numbers ([*oito*] 'eight', [*nove*] 'nine').

(133) *hanipakádaa pidéenhikani* [*até*] **[*oito*]** [*hora*] **[*nove*]** [*hora*]
hanipa-kádaa pi-déenhi-ka-ni [até] **[oito]** [hora] **[nove]** [hora]
much-COND 2SG-to.work-SUB-3NFSG until **eight** hour **nine** hour
pittáita *piwatshéetaka* ***dzamána*** [*lata*] [*ou*]
pi-ttáita pi-wátshaa-íita-ka **dzama-na** [lata] [ou]
2SG-to.be.able.to 2SG-to.jump-CAUS-SUB **two-CLF.TRUNK** can DISJ
madalína [*láta*]*kani*
madali-na [lata]-kani
three-CLF.TRUNK can-?

'If you are making a lot of it, up to eight or nine hours you pop them, two or three cans…' [200308_NARR_B_01]

As noted by Ramirez (2020a: 249), whenever a Portuguese numeral is used, an appropriate classifier may optionally be attached to it—unlike the Baniwa numerals 1–3 (§ 5.3.3.1), which obligatorily take a classifier. In (133), both [*oito*] 'eight' and [*nove*] 'nine' appear without classifier marking. (134) gives two synonymous examples showing that the phrase in question is grammatical both with and without a classifier on the Portuguese numeral *cinco* 'five'.

(134) a. [*cinco*] *háiko*
[cinco] háiko
five tree

'five trees' [220919_ELIC_F_01]

b. [*cinco*]***na*** *háiko*
[cinco]-**na** háiko
five-**CLF.TRUNK** tree

'five trees' [220919_ELIC_F_01]

When using Portuguese numerals, Portuguese pronunciation is used[103]. This leads to the use of some phonetic segments not found in the Baniwa phoneme inventory, as well as syllable types that violate the Baniwa syllable structure rules. When classifiers are suffixed to Portuguese numerals, the morphophonological rules of Baniwa are challenged.

In order to investigate how the morphophonological processes would play out under these conditions, an elicitation task was conducted with a native speaker. Two types of data were elicited:

- all 53 classifiers suffixed to the Portuguese numerals *cinco* 'five' and *seis* 'six'
- the classifiers *-da* 'CLF.GENERIC' and *-híwi* 'CLF.POINTED' suffixed to a selection of Portuguese numerals (4–23, 30, 40, 50, 60, 70, 80, 90, 100, 1,000)

The numerals *cinco* 'five' and *seis* 'six' were selected for phonological reasons. *Cinco* [sĩː.ku] obeys Baniwa syllable structure, and all sounds in the last syllable ([k], [u]) are native to Baniwa[104]. *Seis* [seis], on the other hand, violates Baniwa syllable stucture in that it is consonant-final. In addition, [s] is not found in the Baniwa phoneme inventory (§ 3.1.1.1). Thus, *cinco* and *seis* provide good conditions for testing the morphophonological processes on foreign phonological material, both when it obeys and disobeys the phonological structure. *Cinco* is taken to represent all V-final Portuguese numerals and *seis* all C-final ones.

The classifiers *-da* 'CLF.GENERIC' and *-híwi* 'CLF.POINTED' were selected on similar grounds. While *-da* does not trigger any morphophonological processed across morpheme boundaries, *-híwi* triggers both /h/-metathesis (§ 3.1.3.2) and vowel

[103]A similar situation is described for Tariana–Portuguese code switching by Aikhenvald (2010: 185–186).

[104]The Baniwa phoneme /o/ has [u] as one of its phonetic realizations (Ramirez, 2020a: 79).

fusion (§ 3.1.3.4). Here, *-da* is taken to represent all classifiers that do not trigger morphophonological processes, and *-híwi* 'CLF.POINTED' all those that do (both /h/-metathesis and vowel fusion).

Below, the classifiers are divided into four phonological types: monomoraic, V-initial, C-initial (C ≠ /h/), and /h/-initial classifiers. Each type will be described in terms of how it behaves when confronted with V- vs. C-final roots. Reference is frequently made to the forms on the Baniwa numerals 1–3 (§ 5.3.3.1), as well as the suffixing forms on the Baniwa numeral 4 (§ 5.3.3.2). As some speech sounds in the data are not present in the Baniwa phonological system, and thus do not have any conventionalized orthographic representation, all elicited forms in this section are presented in IPA.

9.2.1.1 Monomoraic classifiers

Only one classifier is monomoraic, that is, consists only of a short vowel: *-i* 'CLF.BASKET'. Table 9.2 summarizes its morphophonological behavior with both Baniwa and Portuguese numerals.

Table 9.2: Morphophonology of monomoraic classifiers with numerals

Numeral	Gloss	Clf	Gloss	Form	(Expected)
Regular: [a] + [i] → [e]					
apa-	'1'	*-i*	'CLF.BASKET'	[á.p**e**]	
dzama-	'2'	*-i*	'CLF.BASKET'	[ʤá.m**e**]	
Irregular: [j]-epenthesis; [i] → [e]					
madali-	'3'	*-i*	'CLF.BASKET'	[ma.da.ɽí.**je**]	*[ma.dá.ɽi]
likoadaaka-	'4'	*-i*	'CLF.BASKET'	[ɽi.kua.da:.ká.**je**]	*[ɽi.kua.dá:.ke]
cinco	'5'	*-i*	'CLF.BASKET'	[sĩ:.kú.**je**]	*[sí̃:.ki]∼[sí̃:.ku]
Irregular: [j]-epenthesis; [i] → [a]					
seis	'6'	*-i*	'CLF.BASKET'	[séis.**ja**]	*[séis.i]

On the native numerals 1 and 2, *-i* follows the regular vowel fusion rule [a] + [i] → [e]. When it come to Portuguese numerals, both V-final (*cinco*) and C-final (*seis*) ones trigger the irregular insertion of an epenthetic glide [j] between the root and the suffix. V-final *cinco* triggers the irregular vowel change [i] → [e], aligning itself with the native numerals 3 and 4 (also V-final) which show the same pattern. C-final *seis* instead triggers [i] → [a].

9.2.1.2 V-initial classifiers

Fourteen classifiers are vowel-initial (not counting monomoraic-*i* 'CLF.BASKET', see § 9.2.1.1). They all begin in the long vowels [a:] (e.g., *-áanhaa*, 'CLF.LIQUID'), [e:] (e.g., *-éewhe* 'CLF.EGG'), or [i:] (e.g., *-íida* 'CLF.HALF'). Table 9.3 summarizes their morphophonological behavior with both Baniwa and Portuguese numerals.

Table 9.3: Morphophonology of V-initial classifiers with V-final numerals

Numeral	Gloss	Clf	Gloss	Form	(Expected)
Regular [a:]-initial: [a] + [a:] → [a:], [i] + [a:] → [ia], [u] + [a:] → [ua]					
apa-	'1'	*-áanhaa*	'CLF.LIQUID'	[a.p**á:**.n̥a:]	
madali-	'3'	*-áanhaa*	'CLF.LIQUID'	[ma.da.ɽ**ía**.n̥a:]	
cinco	'5'	*-áanhaa*	'CLF.LIQUID'	[sĩ:.k**úa**.n̥a:]	
Regular [e:]-initial: [a] + [e:] → [e:], [i] + [e:] → [ie], [u] + [e:] → [ue]					
apa-	'1'	*-éewhe*	'CLF.EGG'	[a.p**é:**.w̥e]	
madali-	'3'	*-éewhe*	'CLF.EGG'	[ma.da.ɽ**íe**.w̥e]	
cinco	'5'	*-éewhe*	'CLF.EGG'	[sĩ:.k**úe**.w̥e]	
Regular [i:]-initial: [a] + [i:] → [ai]∼[e:], [i] + [i:] → [i:], [u] + [i:] → [ui]					
apa-	'1'	*-iíta*	'CLF.HUMAN'	[a.p**ǎi**.ta]	
madali-	'3'	*-iíta*	'CLF.HUMAN'	[ma.da.ɽ**ĭ:**.ta]	
cinco	'5'	*-iíta*	'CLF.HUMAN'	[sĩ:.k**ŭi**.ta]	
Irregular [i:]-initial: [a] + [i:] → [ai], [i] + [i:] → [ie], [u] + [i:] → [ue]					
apa-	'1'	*-íida*	'CLF.HALF'	[a.p**ái**.da]	
madali-	'3'	*-íida*	'CLF.HALF'	[ma.da.ɽ**íe**.da]	*[ma.da.ɽ**í:**.da]
cinco	'5'	*-íida*	'CLF.HALF'	[sĩ:.k**úe**.da]	*[sĩ:.k**úi**.da]

As native roots in Baniwa are always vowel-final, these classifiers trigger regular vowel fusion rules when used on native numerals. When suffixed to Portuguese vowel-final numerals—which follow the syllable structure rules of Baniwa—the same vowel fusions are triggered. The examples all end in [u], which is a speech sound in Baniwa (an allophone of the phoneme /o/), meaning that there is a vowel fusion template (that of *o*-final morphemes) for the classifiers to latch onto. In 13 out of 14 cases, regular forms are produced.

The only exception is *-iída*, which behaves in this context as if the underlying form were *-éeda*. It also behaves as underlying *-éeda* on the native numeral 3. On the native numeral 1, it produces a regular form (i.e., behaves as *-iída*).

When vowel-initial classifiers are suffixed to Portuguese consonant-final numerals, which violate the syllable structure rules of Baniwa and therefore do not have counterparts in the native vocabulary, the root-final consonant blocks the vowel fusion that would normally take place. This blocking apparently triggers various other morphophonological phenomena, such as epenthesis, elision, and

vowel quality changes. A few different outcomes are recorded, depending partly on the initial vowel quality of the classifier. The forms are shown in Table 9.4. As they all involve a non-native, consonant-final syllable structure, they cannot be said to be either regular or irregular compared to the behavior of native vocabulary.

Table 9.4: Morphophonology of V-initial classifiers with C-final numerals

Numeral	Gloss	Clf	Gloss	Form
[a:]-initial: [j]-epenthesis				
seis	'6'	*-aápa*	'CLF.OBLONG'	[seis.**jă:**.pa]
seis	'6'	*-áanhaa*	'CLF.LIQUID'	[seis.**já:**.n̪a:]
seis	'6'	*-áaphi*	'CLF.AREA'	[seis.**já:**.pʰi]
[i:C[-VOICE]]-initial: elision [i:] → Ø				
seis	'6'	*-iíta*	'CLF.HUMAN'	[séis.ta]
seis	'6'	*-iítsia*	'CLF.BUNCH'	[séis.t͡ʃia]
seis	'6'	*-íixi*	'CLF.SEED'	[séis.ʃi]∼[séi.ʃi]
[i:C[+VOICE]]-initial: [j]-epenthesis, [i:] → [e:]				
seis	'6'	*-íida*	'CLF.HALF'	[seis.**jé:**.da]
seis	'6'	*-íiwi*	'CLF.FLOWER'	[seis.**jé:**.wi]
[e:-initial: elision [e:] → Ø				
seis	'6'	*-éewhe*	'CLF.EGG'	[séis.e̯we]
[e:]-initial: [j]-epenthesis				
seis	'6'	*-eéma*	'CLF.HUMAN'	[seis.**jĕ:**.ma]
seis	'6'	*-éekhe*	'CLF.SMALL.SEED'	[seis.**jé:**.kʰe]

For the [a:]-initial classifiers, the clash with a C-final root triggers the insertion of an epenthetic glide [j] between the numeral and the classifier. This process is regular across the six [a:]-initial classifiers (three of which are shown in Table 9.4).

The [i:]-initial classifiers show two different patterns: in three out of five cases, [i:] is elided, and in two out of five, an epenthetic [j] is inserted which appears to cause a vowel change [i:] → [e:]. The split is possibly conditioned by the voicing of the consonant following immediately upon initial [i:], so that voiceless consonants (C[-VOICE]) trigger elision, and voiced consonants (C[+VOICE]) trigger [j]-epenthesis. Stress placement, tone, number of syllables, as well as place or manner of articulation of the following consonant can all be ruled out as conditioning factors.

When the initial vowel of *-íixi* has been elided, this classifier can optionally undergo a second elision process where [s] → Ø, yielding the parallel forms [séis.ʃi]∼[séi.ʃi]. That this elision is a form of assimilation triggered by the ap-

pearance of suffix-initial [ʃ] is supported by the fact that the [ʃ]-initial classifier *-xaa* 'CLF.EXCREMENT' triggers the same assimilation (see § 9.2.1.3).

The [e:]-initial classifiers show the same two patterns as the [i:]-initial ones: in one case, [j]-epenthesis is triggered, and in the other two cases, [e:] is elided. However, the explanation given for the [i:]-initial classifiers, that the split is conditioned by the voicing of the immediately following consonant, does not hold up for these examples. The pattern featuring vowel breaking contains both a classifier with a voiced consonant and a classifier with a voiceless consonant.

9.2.1.3 C-initial classifiers (C ≠ /h/)

Thirty-one classifiers begin with a consonant other that /h/ (/h/-initial classifiers are treated separately in the following section, § 9.2.1.4). In the phonological system of Baniwa, consonant-initial forms do not trigger any morpheme boundary phenomena, but produce predictable, agglutinating forms. When C-final classifiers are suffixed to V-final Portuguese numerals (which follow the syllable structure rules of Baniwa), they behave as they do on native roots. Fifty-one such examples were recorded, all of which are regular. Table 9.5 summarizes the use of C-initial classifiers on V-final roots through a selection of examples with the generic classifier *-da*.

Table 9.5: Morphophonology of C-initial classifiers with V-final numerals

Numeral	Gloss	Clf	Gloss	Form
Regular: agglutinating				
apa-	'1'	*-da*	'CLF.GENERIC'	[a.pá.da]
madali-	'3'	*-da*	'CLF.GENERIC'	[ma.da.ɽí.da]
cinco	'5'	*-da*	'CLF.GENERIC'	[sĩ:.kú.da]
nove	'9'	*-da*	'CLF.GENERIC'	[nɔ:.ví.da]
setenta	'70'	*-da*	'CLF.GENERIC'	[se.tẽ:.tá.da]

The non-native, consonant-final syllable structure of some Portuguese roots does not cause any major problems for consonant-initial classifiers. In 34 out of 36 examples, no morpheme boundary phenomena are triggered, and predictable, agglutinating forms are produced. In one case (*-xaa* 'CLF.EXCREMENT'), suffix-initial [ʃ] causes elision or assimilation of root-final [s] (the same thing happens with *-íixa* 'CLF.SEED' after the loss of its initial vowel, see § 9.2.1.2). One form is completely irregular: *-pa* 'CLF.PACKAGE' triggers epenthesis of a syllable [ja:], for no apparent reason.

Two examples were recorded of C-initial classifiers suffixed to Portuguese nu-

merals ending in a nasal vowel [Ṽ] (*vinte e um* [vĩ:.t͡ʃi.ũ:] '21', *cem* [sẽĩ] '100'). The resulting forms were agglutinating, with no morphophonological processes taking place across the morpheme boundary, as is the case for the regular C-final roots. Ṽ-final roots are treated as C-final ones due to the assumption that they have an underlying final nasal (Ṽ = VN, as proposed by Ramirez, 2020a: 59 ff.). The alignment of Ṽ-final roots with C-final ones, rather than V-final ones, only becomes apparent when they receive /h/-initial suffixes (see § 9.2.1.4). Table 9.6 summarized the forms produced by C-initial classifiers on C-final (including Ṽ-final) numerals.

Table 9.6: Morphophonology of C-initial classifiers with C-final numerals

Numeral	Gloss	Clf	Gloss	Form	(Expected)
No morphophonological processes					
seis	'6'	*-wáta*	'CLF.BRACT'	[seis.wá.ta]	
vinte e um	'21'	*-da*	'CLF.GENERIC'	[vĩ:.t͡ʃi.ű:.da]	
cem	'100'	*-da*	'CLF.GENERIC'	[sẽĩ.da]	
Elision: [s] → Ø / _[s]					
seis	'6'	*-xaa*	'CLF.EXCREMENT'	[séi.ʃa:]	*[séis.ʃa:]
Epenthetic syllable (?)					
seis	'6'	*-pa*	'CLF.PACKAGE'	[seis.já:.pa]	*[séis.pa]

9.2.1.4 /h/-initial classifiers

Six classifiers are /h/-initial. As described in § 3.1, /h/ does not behave like other consonants in Baniwa: it is the only consonant that can follow other consonants in the syllable onset, and when it occurs morpheme-initially—as in the classifiers in question—it triggers /h/-metathesis (§ 3.1.3.2). The process of /h/-metathesis affects the last consonant of the root (i.e., the consonant in the onset of the last syllable): obstruents become aspirated and sonorants become devoiced. After metathesis has taken place, the root-final vowel and the (now) suffix-initial vowel fuse according to the regular vowel fusion rules (§ 3.1.3.4).

As long as the Portuguese root is vowel-final (i.e., consistent with the syllable structure of Baniwa), the morphophonological processes operate just as they do on native roots—that is, /h/-metathesis triggers aspiration of obstruents and devoicing of sonorants, followed by fusion of the resulting vowel hiatus according to regular processes. Twenty-six such examples are found on Portuguese numerals, all of which follow these principles regularly. Table 9.7 shows a selection of examples of the dual processes on both Baniwa and Portuguese numerals.

Table 9.7: Morphophonology of /h/-initial classifiers with V-final numerals

Numeral	Gloss	Clf	Gloss	Form
Aspiration of preceding obstruent, vowel fusion				
apa-	'1'	*-híwi*	'CLF.POINTED'	[a.**pʰé**.wi]
noventa [nɔ.vẽ:.ta]	'90'	*-híwi*	'CLF.POINTED'	[nɔ.vẽ:.**tʰé**.wi]
sete [se:.ʧi]	'7'	*-híwi*	'CLF.POINTED'	[se:.**ʧʰí**.wi]
cinco[sĩ:.ku]	'5'	*-híwi*	'CLF.POINTED'	[sĩ:.**kʰú**.wi]
onze [ɔ̃:.zi]	'11'	*-híwi*	'CLF.POINTED'	[ɔ̃:.**sʰí**.wi]
Devoicing of preceding sonorant, vowel fusion				
dzama-	'2'	*-híwi*	'CLF.POINTED'	[ʤa.**m̥é**.wi]
madali-	'3'	*-híwi*	'CLF.POINTED'	[ma.da.**ɾ̥í**.wi]
nove [nɔ:.vi]	'9'	*-híwi*	'CLF.POINTED'	[nɔ:.**fí**.wi]
quatro [kuat.ɾu]	'4'	*-híwi*	'CLF.POINTED'	[kuat.**ɾ̥ú**.wi]

One of the most surprising findings is that /h/-initial classifiers trigger aspiration of obstruents and devoicing of sonorants even on non-native phonemes, as seen in the examples with numerals [v], [z], and [ɾ] are the last consonants, which turn into [f], [sʰ], and [ɾ̥], respectively. The change [z] → [sʰ] involves both devoicing ([z] → [s]) and aspiration ([s] → [sʰ]), but is categorized as aspiration[105]. When it comes to vowels, the vowel-final Portuguese numerals in the sample only end in vowels that are present in Baniwa, and the vowel fusion rules operate as they would on native vocabulary.

If the Portuguese numeral is consonant-final (violating Baniwa syllable structure), an even more surprising process takes place. In these cases, the suffix-initial /h/ is blocked from operating on the preceding consonant, i.e., regressively. Instead, surprisingly, the /h/-metathesis operates on the following consonant in the suffix, i.e., progressively. Progressive /h/-metathesis is not seen elsewhere in the language, but must be triggered by the normally non-permitted syllable structures that /h/-initial classifiers are faced with when latching onto Portuguese numerals. The pressure on /h/ to be realized through the aspiration or devoicing of another consonant is apparently strong enough to redirect the metathesis process to the following consonant instead of the preceding one. In the cases at hand, no vowels collide at the morpheme boundary, and therefore there is no vowel fusion. Instead, the initial vowel of the classifier is elided. This pattern is regular across all 12 examples.

Interestingly, a root-final nasal vowel [Ṽ] also blocks regressive /h/-metathesis. It behaves like C-final roots and not V-final ones in that it triggers progressive

[105] This is in analogy with the native Baniwa phonemes [ts] and [dz], whose voicing contrast is neutralized when they face /h/-metathesis before [i], so that [ts], [dz] → [tsʰ] (Ramirez, 2020a: 69).

/h/-metathesis as well as elision of the first vowel in the classifier suffix. This is regular across the two recorded examples (*vinte e um* '21', *cem* '100'), and suggests that the nasal vowel has an underlying final nasal consonant (Ṽ = VN; as suggested by Ramirez, 2020a: 59 ff.). An analysis of Ṽ as aligning with C-final roots is also compatible with the behavior of Ṽ-final roots with C-initial suffixes (§ 9.2.1.3). There are no examples with Ṽ-final roots and V-initial classifiers in the material at hand. A selection of examples of /h/-initial classifiers with C- or Ṽ-final numerals are shown in Table 9.8.

Table 9.8: Morphophonology of /h/-initial classifiers with C-final numerals

Numeral	Gloss	Clf	Gloss	Form
Aspiration of following obstruent, vowel elision				
seis	'6'	*-hipáda*	'CLF.PIECE'	[seis.**pʰá**.da]
seis	'6'	*-híko*	'CLF.LONG'	[séis.**kʰu**]
Devoicing of following sonorant, vowel elision				
seis	'6'	*-híwi*	'CLF.POINTED'	[séis.**̥wi**]
vinte e um	'21'	*-híwi*	'CLF.POINTED'	[vĩ:.t͡ʃi.ũ:.**̥wi**]
cem	'100'	*-híwi*	'CLF.POINTED'	[sẽ̃ĩ.**̥wi**]

9.2.1.5 Summary of classifier use on Portuguese numerals

The (optional) suffixation of Baniwa classifiers to Portuguese higher numerals forces speakers to juggle Baniwa-internal morphophonological rules on the one hand and foreign phonemes and syllable structures on the other. Some of the morphophonological processes are productive in this new environment. For example, vowel fusion rules at morpheme boundaries generally operate in the same way as they do on Baniwa morphemes (although all root-final vowel qualities encountered so far are also present in Baniwa). In many cases, the process of /h/-metathesis has been found to operate according to the same principles even on phonemes not found in Baniwa, both through devoicing of sonorants and aspiration of obstruents. In other cases, innovative processes are created, such as /h/-metathesis operating progressively instead of regressively when blocked by C-final roots.

Vowel-final roots generally posed less problems than consonant-final ones, which is unsurprising given that Baniwa only allows V-final syllables. The V-final roots were primarily represented by Portuguese *cinco* 'five'. With V-initial classifiers, the outcome was mostly regular according to the vowel fusion rules, and with C-initial classifiers—including /h/-initial ones—the outcome was entirely regular, agglutinating forms. With /h/-initial classifiers, the principles of /h/-metathesis

were applied so regularly that they affected non-native consonants just as they would native ones: sonorants were devoiced ([v] → [f], [ɾ] → [ɾ̥]) and obstruents aspirated ([z] → [sʰ]).

The Portuguese numeral *seis* 'six' served as the main consonant-final example in the elicitation task. *Seis* was indeed the root that caused the most problems for the morphophonological processes, as C-final roots do not exist in Baniwa, and morpheme boundary phenomena such as /h/-metathesis and vowel fusion require vowel-final roots to operate regularly. V-initial classifiers handled the C-final root through a number of strategies, such as epenthesis of [j], elision and breaking. C-initial classifiers, for the most part, produced agglutinating forms without any morphophonological processes taking place, even if it led to the creation of consonant clusters normally not permitted in Baniwa. An exception was [ʃ]-initial classifiers, which caused the final [s] of the root to elide, or possibly to be co-articulated with [ʃ]. The combination of C-final roots and /h/-initial classifiers was the most difficult for the Baniwa phonological system to handle, as both the /h/-metathesis and the vowel fusion rules that normally take place with /h/-initial suffixes were blocked by the final consonant of the root. This leads the /h/-metathesis to operate progressively instead of regressively, affecting the following consonant of the classifier suffix as it would the preceding under normal circumstances: devoicing in case of sonorants ([w] → [w̥]), and aspiration in case of obstruents ([p] → [pʰ], [k] → [kʰ]). Ṽ-final classifiers behave just like C-final ones when suffixed with /h/-initial classifiers, supporting Ramirez's (2020a: 59 ff.) analysis that Ṽ is underlyingly VN.

9.3 Chapter summary

This chapter has explored classifier use in relation to lexical borrowing from Portuguese, from two perspectives: how borrowed nouns are classified, and how classifiers attach to borrowed host words.

Regarding the classification of borrowed nouns, and by extension recently introduced items, they behave as native nouns and follow the basic semantic principles of assignment: for example, inanimate objects are generally categorized according to their shape. But the accommodation of new members may also lead to changes in classifier structure. For example, if a category receives a large influx of new referents, it may lead to shifts in the balance between members of the category, which may in turn lead to reanalysis of the classification principles. This may be happening in the case of motorized vehicles and the canoe classifier *-Ø*. Both the case of vehicles and the case of non-native fruits shed light

on the fact classification systems are affected by the speakers' experiences with the objects in question. For associations between objects that are manufactured in similar ways, knowledge of the manufacturing process is required. Similarly, a highly detailed degree of differentiation of plant parts likely only arises as a result of close familiarity with the plants in question.

Borrowed words can also function as the hosts to which classifiers attach. A case study on classifier suffixation to Portuguese numerals illustrates the specific ways in which borrowed material behaves in a highly agglutinating language with extensive morphophonological processes taking place at morpheme boundaries. When classifiers attach to Portuguese numerals, those processes must adapt to the borrowed host words and their foreign phonetic material and syllable structures. Vowel fusion rules generally operate regularly. When it comes to /h/-metathesis, it sometimes extends productively to foreign phonemes (aspirating obstruents and devoicing sonorants), and sometimes innovates new processes (e.g., changing directionality when blocked by consonant-final roots).

Part III

Concluding remarks

Chapter 10

Discussion and conclusion

The main aim of this thesis has been to provide a comprehensive descriptive account of the classifier system in Baniwa. I have addressed this by analyzing the system from various angles: its phonological and morphosyntactic properties, its functions, its semantic organization, its historical development, and its adaptation in the face of language contact. Another aim has been to put the description in a typological light, and to contribute to the understanding of nominal classification systems from a typological perspective. I have done so by bringing Baniwa into the typological picture, as well as by making reference to other languages, especially Arawakan and Northwestern Amazonian ones, throughout the thesis.

The classifier system of Baniwa has been the subject of linguistic descriptions before, and the findings of previous scholars have laid the groundwork for this particular thesis by providing a rich source of previous materials to draw on. Some aspects of the classifier system had been described previously, but were reanalyzed and put into a broader context. These include the morphosyntactic behavior, semantics, and some of the functions of the classifiers. Other aspects that I have brought up in this thesis were previously undescribed, including the suffixing pattern on the numeral 4, the antidual marking on paired body-part terms, and the morphophonological behavior of classifier suffixes on Portuguese host words. An additional dimension was the casting of the study in the urban environment of São Gabriel da Cachoeira, which provided many illuminating perspectives. Considering the rapid cultural change associated with this living environment, the description provided in this thesis offers a unique snapshot documenting the use of Baniwa in a particular time and place. This feeds into the additional aim of this thesis to contribute to the description, documentation, and linguistic analysis of Baniwa.

In this final chapter, I discuss the findings and contributions and present some general conclusions and avenues for future research, arranged around the major themes that have come up during the course of the thesis.

10.1 Summary of the main findings

Baniwa has a highly versatile system of 53 classifiers (§ 4.2) that are marked in several different loci and perform a number of different functions. Classifiers are flexibly assigned, and associate with real-world referents rather than nouns (§ 4.4).

Classifiers contain up to three syllables (§ 5.1) and are marked as suffixes in various loci within the noun phrase (§ 5.3), both on heads and modifiers (§ 6.1). A classifier-marked modifier can in many cases constitute an independent noun phrase; in such cases, there may be no overt head noun (§ 6.1.1.1, § 6.3). The loci include lower numerals, attributive adjectives, nouns (including a possessive construction), nominalized verbs, and some interrogatives. Most classifiers have a complex form used in certain contexts that includes one of four additional suffixes—the suffix choice is conditioned by the classifier and appears to be a remnant of an older gender-marking system (§ 5.3.2).

Classifiers are used as inflectional suffixes on lower numerals, but in almost all other contexts they function as derivational suffixes—for instance, on adjectives and verbs, they are used in derivational processes in combination with other nominalizers, granting these (primarily predicative) word classes access to the noun phrase (§ 6.1). This analysis suggests that classifiers are not agreement devices in Baniwa, which is also confirmed by the data (§ 6.5). Classifiers are also used for qualification, for example on nouns (§ 6.1). In many cases, these functions are difficult to tease apart, and certain instances are compatible with more than one analysis (§ 6.4).

Other functions of classifiers include referent specification, most obviously in cases with semantically general nouns (§ 6.2), as well as discourse functions such as anaphoric reference and the expression of focus and contrast (§ 6.3).

In terms of semantic properties, the Baniwa classifier system operates along similar parameters as classifier systems cross-linguistically, encoding distinctions primarily in shape, animacy, and part–whole relations (§ 7.2). Humans are differentiated according to sex and age (§ 7.2.1.1), while animals are categorized by their shape, with a subset of the classifiers used for inanimate referents (§ 7.2.1.2). Inanimate referents, in turn, are categorized according to a wider variety of shape

properties, with distinctions in dimensionality, flexibility, curvature, hollowness, pointedness, and texture (§ 7.2.2).

Classifiers differ greatly in their internal semantic structure (§ 7.3). Some of them have a very clear prototypical referent that speakers agree on; this is generally correlated with their degree of semantic specificity. The most extreme cases are, on the one hand, the classifiers that are only ever used with a very specific type of referent (§ 7.3.3.2), and on the other hand, the generic classifier (§ 7.3.3.1).

Most classifiers have grammaticalized from bound nouns, but a few may have a different lexical source, such as a free noun, an adjective, or a verb (§ 8.3). The pathway from bound noun to classifier is likely nominal compound constructions (§ 8.2).

The classifier system has developed over a long time span: for example, some can be reconstructed to Proto-Japurá-Colombia or even to Proto-Arawak, whereas others display such a low degree of grammaticalization that they can be used as bound nouns synchronically (§ 8.1). These three age strata explain some of the variation seen in the classifier system: the reconstructable strata have shorter forms and more general semantics than the youngest stratum.

As classifiers are an integral part of Baniwa grammar, their study also brings light to many other domains in the language. One such domain is the semantics of the nominal lexicon; it was shown in § 6.2 that expressions often derive their meaning from a combination of noun and classifier semantics. A very specific case is the antidual marking of paired body part terms (§ 7.2.4.2), which becomes visible through the scrutiny of classifier use. Another is the use of classifier suffixes on Portuguese numerals (§ 9.2.1), which sheds light on morphophonological processes like vowel fusion and /h/-metathesis and their stability.

10.2 The malleability of the classifier system

A major theme throughout this thesis has been the use of an Indigenous South American language in an urban setting, and how the data in this study reflects the ways in which the classifier system of Baniwa adapts to the rapid cultural and linguistic change associated with this relatively new living environment. The effects on classifier use are many, and indicative both of a great degree of malleability and a possible emergent restructuring of the system. They are potentially informative about the way nominal classification systems in general may adapt in similar situations.

The data from the Free listing experiment (§ 7.3, Appendix § B) showcases plenty of examples of how non-Indigenous items are grouped according to the general semantic parameters that the classifier system is organized around (§ 7.2). For example, pens and pencils are grouped with other pointed objects such as needles, arrows, and sticks (*-híwi* 'CLF.POINTED'), beds are grouped with other objects with flat surfaces such as manioc graters and manioc ovens (*-koa* 'CLF.SURFACE'), industrial packages of coffee, biscuits, and milk are grouped with leaf-wrapped fish packages and *panacu* containers (*-pa* 'CLF.PACKAGE'), and slices of pizza are grouped with slices of *beiju* manioc flatbread (*-wána* 'CLF.SLICE'). The same principles show up in the way loanwords are categorized more generally (§ 9.1); for example, container-like motorized vehicles are categorized with canoes (*-Ø* 'CLF.CANOE'). These examples demonstrate a high degree of semantic robustness of the system. The semantic parameters in the Baniwa classifier system, which feature distinctions relating to, e.g., animacy and shape, are in line with how nominal classification systems are organized cross-linguistically. The fact that nominal classification systems resemble each other in systematic ways has been taken as evidence for such systems to be reflecting distinctions of relevance to humans in general, not just to speakers of particular languages in particular cultural contexts (see § 2.3). The Baniwa data is partly in line with this view, as it shows that a semantic organization that was developed in an Indigenous context also appears to be functional in an urban environment.

At the same time, the Free listing data is also indicative of how the prototypicality status of members within categories may shift, probably as a result of changes in the frequency of interaction between different types of objects. As discussed in § 7.3.4, some of the items that show up as the most prototypical ones within their respective categories are relatively recently introduced objects and concepts for speakers of Baniwa. One example is the rifle, which in my data is the most widely agreed upon member of the tube classifier category (*-pi*). The classifier *-pi* goes back to a classifier in Proto-Arawak, spoken long before rifles entered the South American continent. Therefore, the rifle cannot possibly have been the most prototypical member of this classifier category for the entire time period of its existence, showing that categories are malleable also in the sense of their prototypicality, not just in the members they include at their peripheries.

The classifier system also displays signs of adaptability in a formal sense. In particular, the way and extent to which numerals are used by Baniwa speakers in the urban setting is likely to differ significantly from how numerals were used in the past (see Epps & Salanova, 2013: 5–6 for the limited importance of numerals in Amazonian languages). Urban environments with monetarized economies make exact counting a necessity, which in Baniwa has led to the borrowing of

Portuguese numerals from four and up. These borrowed numerals take classifiers only optionally, but when they do, they are integrated into the linguistic structure to the extent that the complex morphophonological rules apply even to non-native speech sounds (§ 9.2.1); again demonstrating a certain linguistic robustness, namely of the morphophonological kind. The use of Portuguese numerals has also largely replaced the native numeral 4, which used to take classifiers as infixes. To the extent that this form is used at all, it has fossilize in its generic form, optionally taking a classifier as a suffix instead, in analogy with the lower numerals but in a slot where classifiers were never used originally (§ 5.3.3.2).

Most of the speakers I worked with were in their forties or older. The only speaker who was in his twenties at the time of data collection uses classifiers in a way that differs significantly from the other speakers. While the language use of a single person is not enough to draw any far-reaching conclusions about inter-generational differences in Baniwa classifier use, it suggests an interesting avenue for future research. In the Free listing experiment, when speakers were asked to list examples of items that could go into each classifier category, this speaker gave many more blank responses than the rest of the speakers (28, compared to 0–8 among the other speakers). This suggests that the classifier system that this speaker has active command of consists of a smaller set compared to those of the older speakers. In addition, in spontaneous speech, the young speaker uses a classifier in all the contexts where they obligatorily occur, but appears to use the generic classifier *-da* more or less exclusively. Thus, while adhering to the grammatical rules concerning classifier occurrence, the speaker appears not to be relying on classifiers as a strategy for identifying and differentiating referents, as is otherwise characteristic of the system (see § 6.2), presumably compensating for this function of classifiers by some other linguistic strategy. This situation brings to mind the restructuring of noun classes described for Young People's Dyirbal (Schmidt, 1985; see also § 7.3.3.3).

10.3 Baniwa classifiers in a typological perspective

One of the goals of this thesis was to put the description of Baniwa classifiers in a typological perspective, especially in light of its genealogical and geographical position, and to contribute to the theoretical and typological understanding of nominal classification systems. In this section, I summarize the findings from a typological viewpoint, and in § 10.3.1, I discuss their theoretical implications.

The classifier system in Baniwa is relatively typical of the kinds of Amazonian nominal classification systems that have been called *multiple* (Aikhenvald, 2000),

multifunctional (Krasnoukhova, 2012), or *multilocus* (Dunn & Rose, forthcoming) classifier systems. In Table 10.1, the properties of Baniwa are compared to those identified by Krasnoukhova (2012) as typical for such systems.

Table 10.1: Properties of Baniwa classifiers compared to multifunctional classifier systems (adapted from Krasnoukhova, 2012: 206)

	Properties of a multifunctional classifier system	**Baniwa**
Associated with semantic function		
1	Nouns can be assigned to various classes freely	✓
2	Form largish number of classes	✓
3	Constitute an open system	-
Associated with derivational function		
4	Can derive noun stems	✓
4a	Either from noun stems or roots	(✓)
4b	Or from verbal stems or roots	(✓)
5	Can form a full NP when occurring on a modifier	✓
Associated with agreement function		
6	Can occur on predicates to mark core arguments	-
6a	Either on *any* predicate	-
6b	Or only on a subclass of predicates	-
7	Can participate in agreement within the NP	-
8	Classify all nouns	✓

Among the properties associated with semantics (Krasnoukhova, 2012: 207–209), Baniwa possesses two out of three. Flexible assignment is characteristic of the Baniwa classifier system, in line with most of the languages in Krasnoukhova's sample. In terms of inventory size, Baniwa positions itself somewhere between the extremes (some Nambikwaran languages with seven–eight; Kubeo with more than 150). While Baniwa has some classifiers that can also function as nouns, this is restricted to a limited set, which is why I do not consider the system to be open. It is difficult to assess whether this behavior of Baniwa is similar to or different from the languages in Krasnoukhova's sample, as the notion of an "open system" is somewhat open to interpretation.

The derivational properties are an important function that sets the Amazonian type apart from other kinds of classifier systems (Krasnoukhova, 2012: 209–212). Baniwa classifiers can derive noun stems from both nominal and verbal roots, and in many cases form independent noun phrases with their host words without overt nouns (including for anaphoric use).

Most of the agreement properties (Krasnoukhova, 2012: 212–217) are absent in Baniwa, although this does not necessarily make Baniwa very unusual in comparison to other multifunctional classifier systems, as Krasnoukhova also notes that

some of these properties are problematic. The only property in this category that Baniwa possesses is a system that applies to all nouns—however, Krasnoukhova questions the relevance of this feature in light of the existence of neutral classifying elements. The rest of the properties are absent from Baniwa. Baniwa classifiers are not agreement devices, but this is in fact quite typical of the languages in Krasnoukhova's sample, too; many of them display some optionality in the realization of agreement (the same is noted for Arawakan languages by Dunn & Rose, forthcoming: 27–29). The main point where Baniwa differs from Krasnoukhova's sample is when it comes to classifier marking on predicates. Baniwa does not mark classifiers on predicates[106], while such marking is "quite typical" in other Amazonian languages (Krasnoukhova, 2012: 212).

It is worth noting that the absence of classifier marking on predicates is also the main point where Baniwa deviates from the Arawakan languages. Verbs, understood as predicates, are identified by Dunn & Rose (forthcoming) as one of the four major classifier loci in the family, although they do note that this locus is uncommon in the Northwestern Arawakan languages. The other three loci—nouns, numerals, modifiers, which are all within the noun phrase—are shared by Baniwa. Another relatively unusual property of Baniwa, at least from an Arawakan perspective, seems to be the use of classifiers on interrogatives.

The semantic organization of the Baniwa classifier system also has many properties in common with those of other Amazonian languages (see, e.g., Epps & Obert, 2022: 6). The system features distinctions primarily in physical properties, and animacy and sex are also salient. The existence of a generic classifier is also common, as well as a number of highly specific classifiers that overlap with nouns.

In some cases, Baniwa shape classifiers are extended to cover abstract referents. In particular, classifiers denoting various kinds of long and thin objects appear to take on abstract senses (e.g., the curvilinear classifier *-khaa* for concepts relating to language and cognition, and the tube classifier *-pi* for months). Very similar extensions have been noted elsewhere in the Arawakan language family, although far from all sources provide information about classifier use with abstract referents.

Classifiers in Baniwa are similar to many other Arawakan and Amazonian systems in that they have bound nouns as their primary lexical source. In Baniwa, the distribution of classifiers and bound nouns only overlaps on nouns, whereas

[106]The only exceptions are the use of classifiers on one predicative interrogative (§ 6.1.1.5) and a handful of irregular adjectives that obligatorily take classifiers also when used as predicates (§ 5.3.4), which I do not consider enough to constitute a subclass of predicates.

in some other Arawakan languages, it overlaps in several contexts. Plant and body part terms are frequently mentioned as the most common sources for classifiers; in Baniwa, bound nouns denoting plant parts seem particularly prone to develop inte classifiers.

10.3.1 Theoretical implications

The analysis of Baniwa classifiers presented in this thesis highlights a number of issues in classifier research that have implications for the theoretical understanding of such systems, as well as of language more broadly. A few specific areas are mentioned here.

The analysis of classifier functions shows that in most of the loci, classifiers have a derivational rather than inflectional function. This explains why the system does not display agreement. Other authors have pointed to the problems with positing agreement for Amazonian languages, typically in light of the optional use of classifiers in some languages, as well as the lack of a rigid association between classifiers and nouns (Krasnoukhova, 2012: 212–217; Dunn & Rose, forthcoming: 27–29). The Baniwa data illustrate a third reason why systems may not be displaying agreement: because their function is in many cases derivational and not inflectional, due to the requirement of noun phrase modifiers to be nominalized. To what extent this analysis applies to classifier systems in other languages, in Amazonia or elsewhere, is a question for future research. Krasnoukhova's (2012) sample indicates that the nominalization of noun phrase modifiers is a fairly widespread phenomenon in South America, although not all of the surveyed cases employ classifiers as nominalizers.

The derivation–inflection question also has other potentially significant implications, such as pointing to a possible pathway for the development of inflectional morphology and agreement. A Baniwa adjective or verb needs to be nominalized in order to occur as an modifier in a noun phrase, and the nominalization construction in most cases requires a classifier, which must be semantically compatible with the head noun. Such situations could, over time, give rise to a reanalysis of classifiers as inflectional markers, and perhaps result in agreement constraints within the noun phrase.

10.4 A categorization, or something more?

In the opening sentence of this thesis, I declared the fascination for how humans categorize the world through language as a point of departure for this whole endeavor. But as we have seen along the way, nominal classification is about more than just a rigid categorization.

The classifiers in Baniwa can indeed be construed as categories, as seen in § 7.3. This, together with the analysis of the semantic parameters that underpin them § 7.2, provide useful tools for understanding the basic semantic structure of the system, in terms of properties, associated members, prototype effects, and metaphorical extensions.

However, for being a system of categories, the classifier system is characterized by a general lack of strict boundaries. This applies on several levels. Nouns are typically compatible with several different classifiers, which is perhaps unsurprising, given that it is not nouns, but referents, that classifiers associate with. But even the same referent can generally be used with different classifiers, depending on the context. There are also no strict boundaries between the semantic parameters, as the same classifier may make use of combinations.

Thus, if we stop at categorization, we risk missing a lot of what makes classifiers interesting. Rather than imposing a strict categorization, classifiers serve to highlight things that people pay attention to in particular contexts. As such, category membership is constantly in flux. The flexibility of classifier assignment is a rich source of linguistic creativity, that can be exploited at the speaker's will.

10.5 Outlook

This thesis has focused on describing the classifier system of Baniwa, outlining the inventory and providing detailed descriptions of the formal, functional, and semantic properties of classifiers, with further reference to diachronic and contact-related aspects of the system. It is my hope that this description can pave the way for future studies on areas that are outside the scope of this thesis.

A complementary study with a greater focus on how Baniwa classifiers are used in context would be necessary in order to advance the general understanding of the phenomenon. Such a study would have the potential of exploring classifier functions at the discourse level in greater detail, such as their involvement in the management of reference and information structure. It would also be able to

better incorporate the notions of frequency and speaker variation.

Nouns in Baniwa are formally divided into categories by several separate grammatical subsystems. Besides the classifier system, which has been the focus of this thesis, there is a bipartite division according to gender, a bipartite division into free and bound nouns, and, among bound nouns, a tripartite division into alienable possession classes. An interesting area for future research is the intersection of these subsystems: what functions each system has, how they cross-cut each other, and how the systems interact.

The understanding of Baniwa classifiers would greatly benefit from studies from sociolinguistic and language acquisition perspectives. A study on how children learn to use classifiers would not only further the understanding of the acquisition patterns, but would also have the potential to enlighten the analysis of the system itself, in terms of possible acquisitional differences between classifiers of different semantic types, and between different morphosyntactic loci and functions. Likewise, a study on sociolinguistic variation in classifier use, especially on the possible intergenerational differences, would have the potential to broaden our understanding of how classifier systems are affected in processes of language attrition and system restructuring.

Appendices

Appendix A

Glossaries

These glossaries, especially the scientific names for plant and animal species, build partly on Epps (2008: 935–937), Ramirez (2001a; 2020a: 15–17) and Obert (2019: 308–309). In some cases, it was not possible to provide a species name or an English translation, as the exact species or genus was not identifiable.

Plant and animal species

Baniwa

aapídza	white-lipped peccary (*Tayassu pecari*; Port. *queixada*)
dáapa	paca, rodent sp. (*Cuniculus paca*; Port. *paca*)
dóomali	umari, tree sp. with edible fruit (*Poraqueiba serica*; Port. *umari*)
dzáapa	peacock bass, fish sp. (*Cichla monoculus*; Port. *tucunaré*)
hawádza	tree sp. with beanlike edible fruits (*Inga sp.*; Port. *ingá*)
hémali	abiu, tree sp. with edible fruit (*Sapotaceae* sp.; Port. *abiú*)
héemali	fish sp., larger variety of peacock bass (*Cichla temensis*; Port. *tucunaré-açu*)
híiniri	tree sp. with edible fruit (*Pouteria ucuqui*; Port. *ucuqui*)
íiniri	fish sp. (*Hoplias malabaricus*; Port. *traíra*)
íitewi	moriche palm, palm sp. with edible fruit (*Mauritia flexuosa*; Port. *buriti*)
íitsi	howler monkey (*Alouata sp*; Port. *guariba*)
iitsítsi	eel sp. (*Synbranchus marmoratus*; Port. *muçum*)
káaparo	woolly monkey (*Lagothrix sp*; Port. *macaco-barrigudo*)
káini	bitter manioc (*Manihot esculenta*; Port *mandioca*)

kamhéro Amazon grape (*Pourouma cecropiifolia*; Port. *cucura*)
kawiápali tree sp. with beanlike edible fruits (*Inga sp.*; Port. *ingá*)
kettínali fish species (*Crenicichla lenticulata*; Port. *jacundá*)
kóitsi curassow, large edible bird sp. (*Cracidae sp.*; Port. *mutum*)
kolíri catfish sp. (*Pseudoplatystoma fasciatum*; Port. *surubim*)
kóowhe leaf-cutter ant (*Atta sp.*; Port. *saúva*)
kópiro plant with edible fruit (*Solanum sessiliflorum*; Port. *cúbio*)
máaliphe fish sp. (*Hydrolycus sp.*; Port. *peixe-cachorro/pirandira* variety)
máawi palm sp. used for flutes and blowpipes (*Iriartella setigera*; Port. *jupati/pari*)
manákhe açaí, palm sp. with edible fruit (*Euterpe precatoria*; Port. *açaí*)
máre guan bird, large edible bird sp. (*Cracidae sp.*; Port. *jacú*)
mhóokoli large catfish sp. (*Brachyplatystoma vaillantii*; Port. *piraíba*)
onítholo catfish sp. (*Asterophysus batrachus*; Port. *mamaiacu*)
ówhii fish sp. (*Sternopygus macrurus*; Port. *sarapó* variety)
phíitsi agouti, rodent sp. (*Dasyprocta sp.*; Port. *cutia*)
píipiri peach palm, palm sp. with edible fruit (*Bactris gasipaes*; Port. *pupunha*)
póapoa plant whose fibres are used for basketry (*Ischnosiphon sp.*; Port. *arumã*)
ponáma patawa, palm sp. (*Jessenia bataua*; Port. *patauá*)
póoperi palm sp. with edible fruit (*Oenocarpus bacaba*; Port. *bacaba*)
póotto small agouti sp. (*Myoprocta pratti*)
porámo palm sp., variety of *açaí* (*Euterpe catinga*; Port. *açaí-chumbo*)
táali fish sp. (*Anostomoides laticeps*; Port. *aracu*)
téwa fish sp. (Port. *piaba* variety)
tsíipa fish sp. (*Myleus pacu*; Port. *pacu*)
tsíitsi monkey sp. (*Cacajao* sp.; Port. *uacari*)
ttéephe fish sp. (*Characidae sp.*; Port. *piaba*)
ttíiña palm sp. (*Mauritiella armata*; Port. *caranã*)
wéemai fish sp. (*Hydrolycus sp.*; Port. *peixe-cachorro/pirandira* variety)

Regional Portuguese terms

aturá	large basket with forehead strap used for carrying harvest, primarily manioc tubers
beiju	flat, round bread made of bitter manioc; staple food
caapi	ayahuasca, hallucinogenic drink made from a vine (*Banisteriopsis caapi*)
caxiri	fermented alcoholic drink made from manioc or fruit
chibé	drink made of *farinha* and water
curuatá	bract (botanical term for a particular leaf that grows around the inflorescence of certain palm trees)
farinha	roasted bitter manioc flour; staple food
forno	manioc roasting oven, used for making *beiju*, *farinha*, and *tapioca*
jarro	type of curved basket commercially produced by the Baniwa
jiquitaia	dried spice made of ground chili peppers and salt
jirau	wooden frame similar to a stand with a horizontal plane consisting of parallel pieces, used for drying or storing
maçoca	certain type of *farinha*
maloca	traditional longhouse
maniva	the stems of manioc that grow above ground
panacu	disposable container made of braided leaves, carried on forehead
ralo	manioc grater made of paxiuba wood (*Socratea exorrhiza*) and small, sharp pieces of quartz stones, traditionally manufactured by the Baniwa
roça	manioc garden
tapioca	manioc starch
tipití	woven tube used to squeeze the liquid out of grated manioc
tucum	fibre from a palm leaf (*Astrocaryum tucuma*) used to make hammocks
tucupí	spicy sauce made from squeezed-out manioc liquid
zagaia	fishing spear

Appendix B

Free listing data

This appendix contains the collective lists produced in the Free listing experiment, where participants listed items associated with individual classifiers (see further in § 1.2.4.1). The items that the participants listed in response to the stimulus prompt were of a few different types:

- Free nouns, such as *áatti* 'chili pepper'.
- Bound nouns, which are always used with a possessor by the participants, but cited here in their bare form, e.g., *-nawathére* 'elbow'.
- Compounds, both with connector (*íikoli íewhe* 'turtle's egg') and without (*iikolíewhe* 'turtle's egg'), see § 3.2.5.2.
- Unintegrated loanwords from Portuguese, such as [*sabão*] 'soap'. These are rendered with their Portuguese spelling, and put in square brackets to indicate that the spelling conventions for Baniwa otherwise used in this thesis (§ 3.1.4) do not apply.
- Classifiers are sometimes used without an overt noun (see, e.g., § 6.1.1.1, § 6.3). Consequently, participants occasionally listed the numeral-classifier combination as an entity of its own. These instances were coded as *Ø* in the data, to represent the lack of a noun. For example, *Ø* 'egg' occurs in the list of *-éewhe* 'CLF.EGG'. This means that the participant has given the example *apéewhe* (*Ø*) and given the translation 'egg'.

A possible bias in the lists of items is that they are likely influenced by objects that were within the participants' eyesight at the time and place of the elicitation sessions. For example, one participant listed the Portuguese noun [*computador*] 'computer' for the category *-da* 'CLF.GENERIC'. It is more likely that this item was

listed for being in the vicinity of the participant than for representing a typical member of the classifier category in question.

When compiling the collective lists, a number of principles were followed:

- Variants of the same lexeme were counted as instances of the same item (e.g., *atsíanli*∼*tsíanli*∼*áatsia* 'man'; Ramirez, 2001a: 50).
- Derived variants of the same lexeme base were counted as separate items, even in cases where the meaning was (largely) identical (e.g., *háiko* 'stick' ≠ *haikóapo* 'stick').
- Integrated and unintegrated loanwords from the same source word were counted as instances of the same item (e.g., *méedza* = [*mesa*] 'table' < Portuguese *mesa* 'id.').
- Compounds in their full and contracted versions (§ 3.2.5.2) were treated as instances of the same item (e.g., *íikoli íewhe* = *iikolíewhe* 'turtle's egg').
- Bound nouns and their converted free versions (§ 3.2.5) were counted as instances of the same item (e.g., *-hinóoro* = *hinoorótti* 'hammock rope').

The data are presented in Table B.1 in the form of the compiled collective lists of each of the 48 classifiers tested, presented in alphabetical order. 'Order' specifies the order of the items in each list. The items themselves are listed under 'Item', and their meaning under 'Meaning'[107]. The column 'Points' shows the score of each item, and 'Tokens' shows the number of times the item in question was mentioned (i.e., the number of participants who mentioned it).

[107]Note that the 'Meaning' column here specifies the meaning of the classifier-lexeme combination, rather than the basic meaning of the lexeme itself. An example is *newíki*, which means 'person' in its basic form, but 'group of people, clan' when used with *-áana* 'CLF.GROUP'.

Table B.1: Collective lists from the Free listing experiment

Order	Item	Meaning	Points	Tokens
***-Ø* 'CLF.CANOE'**				
1	*íita*	'canoe, boat'	38	4
2	*kóoya*	'type of calabash bowl'	19	2
3	*átta*	'type of calabash bowl'	17	2
4	[*concha*]	'ladle'	7	1
***-áana* 'CLF.GROUP'**				
1	*newíki*	'group of people, clan'	36	4
2	*aapídza*	'group of white-lipped peccaries'	10	1
3	*pówe*	'group of monkeys'	9	1
3	*itopikakápe*	'football team'	9	1
5	*íitsi*	'group of howler monkeys'	8	1
6	*káaparo*	'group of woolly monkeys'	7	1
***-áanhaa* 'CLF.LIQUID'**				
1	*manákhe(a)*	'açaí juice'	17	2
2	*óoni*	'water'	10	1
2	*hámoli*	'year'	10	1
4	*líana*	'backwater'	9	1
4	*maawíroa*	'pineapple juice'	9	1
5	*ponáma*	'patawa juice'	8	1
6	*patshíakaa*	'drink made of water and manioc flour'	6	1
7	*tápee*	'(liquid) medicine'	5	1
***-aápa* 'CLF.OBLONG'**				
1	*palána*	'banana (individual fruit)'	45	5
2	*kóitsi*	'curassow'	28	3
3	*kaláka*	'hen/rooster'	18	2
4	*hawádza*	'*ingá* sp. (individual fruit)'	15	2
4	*kepíreeni*	'small bird (generic)'	15	2
6	*máali*	'heron'	11	2
7	*máre*	'guan bird'	9	1
8	*wáaro*	'parrot'	8	1
9	*káana*	'ear of corn'	7	2
9	*kawiápali*	'*ingá* sp.'	7	1
11	*kettínali*	'fish sp.'	6	1
12	*dzáawi*	'ocelot'	5	1
13	*phíitsi*	'agouti (rodent sp.)'	4	1
14	*póotto*	'small agouti sp.'	3	1
	káini	'manioc tuber'		1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
***-áaphi* 'CLF.AREA'**				
1	*kiníki*	'manioc garden'	40	4
2	[*quarto*]	'room'	16	2
3	*-nómawa*	'hole'	14	2
4	*halhaméaphi*	'fold (of, e.g., cloth)'	10	1
4	*kalítta*	'lake'	10	1
6	*panttinóma*	'door'	9	1
6	*óoni*	'lake'	9	1
6	*-aaróaphi*	'place for lying down'	9	1
9	*dzakálee*	'community, town'	8	1
9	*-hinaáphia(mi)*	'shadow'	8	1
9	*amáka*	'cleared and burnt area'	8	1
12	*-hikaniyáaphi*	'hole'	6	1
***-áapi* 'CLF.HOLLOW'**				
1	*paráto*	'plate'	45	5
2	*iñápi*	'bone'	22	3
3	[*bacia*]	'bowl'	15	2
4	*tsipaláapi*	'pot, pan'	10	1
5	*aakhéepa*	'clay pot'	9	1
5	[*panela*]	'pot'	9	1
7	*manákhea*	'açaí juice (in a dish or pot)'	8	1
8	*óoni*	'water (in a dish or pot)'	7	1
8	*-aaróapi*	'vessel'	7	1
***-áapo* 'CLF.STICK'**				
1	*iitsáapo*	'fishing rod'	39	4
1	*haikóapo*	'stick'	37	4
1	*háiko*	'tree'	10	1
1	*inípo*	'path'	8	1
1	*ñawápo*	'creek'	7	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
***-da* 'CLF.GENERIC'**				
1	*hiipáda*	'stone'	18	2
1	*dáapa*	'paca (rodent sp.)'	18	2
3	*-iináka (háiko)*	'fruit (from a tree)'	17	2
4	*déepi*	'night'	10	1
4	*íita*	'canoe'	10	1
4	*irímawa*	'lemon'	10	1
4	*-éehwe*	'egg'	10	1
8	*itsída*	'tortoise'	9	2
8	*naláya*	'orange (fruit)'	9	1
8	*heekóapi*	'day'	9	1
8	*kaláka íewhe*	'hen's egg'	9	1
8	*manákhe*	'a single açaí fruit'	9	1
13	*-kóloni*	'sweeling (on skin)'	8	1
14	*tshéeto*	'*aturá* basket'	7	1
14	*tsipaláapi*	'pot, pan'	7	1
16	*mólhoi*	'*maçoca*'	6	1
16	*aakhéepa*	'clay pot'	6	1
16	*kópiro*	'*cúbio* fruit'	6	1
19	[*laranja*]	'orange (fruit)'	5	1
19	*áatti*	'chili pepper'	5	1
19	*ómai*	'piranha'	5	1
22	[*caneco*]	'cup'	4	1
22	*tsíipa*	'fish sp.'	4	1
22	*hémali*	'abiu fruit'	4	1
25	[*computador*]	'computer'	3	1
25	*íikoli*	'turtle'	3	1
25	*papéra*	'book'	3	1
28	*padanakaróda*	'notebook'	2	1
28	[*lápis*]	'pencil'	2	1
30	[*relógio*]	'clock, watch'	1	1
30	*híiparo*	'frog sp.'	1	1
30	[*caderno*]	'notebook'	1	1
	-thi	'eye'		1
	kóoya	'type of calabash bowl'		1
	onítholo	'catfish sp.'		1
	wáaro	'parrot'		1
	pántti	'house'		1
	-áako	'word, language'		1
	[*bóla*]	'ball'		1
	[*balão*]	'balloon'		1
***-daa* 'CLF.DAY'**				
1	*heekóapi*	'day'	49	5
2	*déepi*	'night'	19	2

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
-dápana **'CLF.HOUSE'**				
1	*pántti*	'house'	60	6
2	*pamákali*	'tent'	9	1
2	*iarodátti*	'afterworld'	9	1
-eéma **'CLF.SIDE'**				
1	*dáapa*	'half/side of a paca (lengthwise)'	42	5
2	*-híipa*	'foot (one of a pair)'	19	2
3	*-káapi*	'hand (one of a pair)'	17	2
4	*-kawa*	'leg (one of a pair)'	13	2
5	*-heenittáda*	'earring (one of a pair)'	10	1
5	*peéthe*	'slice of *beiju*'	10	1
7	*néeri*	'half of a deer (lengthwise)'	9	1
8	*-héeni*	'ear (one of a pair)'	8	1
8	*héema*	'half of a tapir (lengthwise)'	8	1
8	*máawiro*	'half of a pineapple (sliced)'	8	1
11	[*melancia*]	'half/slice of a watermelon'	7	1
12	*kóphe*	'half of a fish (sliced)'	6	1
12	*-thi*	'eye (one of a pair)'	6	1
12	*pántti*	'side of a house'	6	1
15	*-náapa*	'arm (one of a pair)'	5	1
15	*íitsiri*	'side of an animal'	5	1
17	*phíitsi*	'side of an agouti'	3	1
-éewhe **'CLF.EGG'**				
1	Ø	'egg'	20	2
2	*kaláka íewhe*	'hen's egg'	19	2
3	*íikoli íewhe*	'turtle's egg'	16	2
4	*-éewhe*	'egg'	10	1
4	*ainíewhe*	'wasp's egg'	10	1
6	*ttowída*	'louse egg'	9	1
6	[*pato*] *íewhe*	'duck's egg'	9	1
8	*kóphe íewhe*	'fish egg'	7	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
-híko **'CLF.LONG'**				
1	*háiko*	'tree, piece of wood'	38	4
2	*máapa*	'sugar cane'	20	2
3	*tsíino*	'dog'	16	2
3	*garáapha*	'bottle'	16	2
5	[*lata*] *matsóka*	'can of manioc flour'	10	1
6	*áatti*	'chili pepper'	9	1
7	*ponáma*	'patawa palm'	8	1
7	*íiniri*	'fish sp.'	8	1
9	*íitewi*	'moriche palm'	7	1
10	[*lanterna*]	'flashlight'	6	1
11	*néeri*	'deer'	5	1
11	*kamárai*	'flashlight'	5	1
13	*manákhe*	'açaí palm'	4	1
13	*yalákhi*	'bottle of *caxiri*'	4	1
-hípa **'CLF.MALE'**				
1	*atsíanli*	'man'	60	6
2	*ienipétti*	'child'	17	2
2	*walhípali*	'young man'	17	2
4	*pedália*	'elder'	9	1
4	*newíki*	'person'	9	1
-hipáda **'CLF.PIECE'**				
1	*kóphe*	'piece/half of a fish (widthwise)'	25	3
2	*háiko*	'half of a stick/piece of wood'	16	2
3	*palána*	'piece of a banana'	10	1
3	*máapa*	'piece of sugar cane'	10	1
3	*dáapa*	'half of a paca (sliced)'	10	1
6	*pántti*	'half of a house'	9	1
7	[*sabão*]	'piece of soap'	8	1
7	[*bolo*]	'piece of a cake'	8	1
7	[*tábua*]	'one half of a board'	8	1
10	*peéthe*	'piece of a *beiju*'	7	1
10	[*pilha*]	'battery'	7	1
12	*hinoorótti*	'piece of hammock rope'	5	1
13	*ttáawaali*	'piece of thread'	4	1
-hípani **'CLF.RAPIDS'**				
1	*híipa*	'rapids'	20	2

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
***-híwa* 'CLF.BEIJU'**				
1	*peéthe*	'(piece of) *beiju*'	40	4
***-híwi* 'CLF.POINTED'**				
1	*áawi*	'needle'	46	5
2	[*caneta*]	'pen'	19	2
3	*haikhíwi*	'stick, twig'	18	2
4	*iítsa*	'fish hook'	17	2
5	*-eétsha*	'tooth'	15	2
6	*-íidzo*	'strand of fur/hair'	12	2
7	[*lápis*]	'pencil'	9	1
7	*iitsáapo*	'fishing rod'	9	1
9	*tápoa*	'nail'	8	1
9	*-íiwi*	'thorn (of plant), spur (of fish/ant)'	8	1
11	*háiko*	'stick'	7	1
11	*kaapáwi*	'arrow'	7	1
13	*haikóapo*	'stick'	5	1
***-i* 'CLF.BASKET'**				
1	*póoperi*	'bunch of *bacaba* fruit'	26	3
2	*ponáma*	'bunch of patawa fruit'	24	3
3	*tshéeto*	'*aturá* basket'	20	2
4	*manákhe*	'bunch of açaí fruit'	19	2
5	*mokóto*	'*panacu* container'	9	1
6	*píipiri*	'bunch of peach palm fruit'	6	1
7	*kamhéro*	'bunch of Amazon grapes'	5	1
***-íida* 'CLF.HALF'**				
1	*dáapa*	'half/piece of a paca (widthwise)'	20	2
2	*kóphe*	'half/piece of a fish (widthwise)'	18	2
3	*-pána*	'part of a house'	10	1
3	*híipai*	'piece of land'	10	1
5	*iitsaákhaa*	'piece of a fishing line'	9	1
5	*kiníki*	'part of a manioc garden'	9	1
7	*pántti*	'part of house (e.g., room)'	8	1
8	*háiko*	'piece of wood'	7	1
9	*irímawa*	'piece of lemon'	6	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
-iíta **'CLF.HUMAN'**				
1	*atsíanli*	'man'	40	4
2	*nawíki*	'person'	30	4
3	*kóphe*	'fish'	25	3
4	*peéthe*	'*beiju*'	24	3
5	*ttiíwe*	'paddle'	22	3
6	[*colher*]	'spoon'	12	2
7	*pówe*	'monkey (generic)'	11	2
8	*matshéeta*	'machete'	8	2
8	*dzáapa*	'fish sp.'	8	1
10	*wéemai*	'fish sp.'	7	1
11	*ñámaro*	'ray (fish)'	5	2
11	*-híipa*	'foot'	5	1
13	[*concha*]	'shell'	3	1
13	*íikoli*	'turtle'	3	1
15	*-eétsha*	'tooth'	2	1
16	*ttowída*	'louse'	1	1
16	*íitsi*	'howler monkey'	1	1
	-héeni	'ear'		1
	-tsóta	'fingernail, toenail'		1
	haikoíta	'board'		1
	ttéephe	'fish sp.'		1
	iitsaákhaa	'spool of fishing line'		1
	káaparo	'woolly monkey'		1
	kóopali	'tick'		1
	malíye	'knife'		1
	-iitshére	'spot (on skin)'		1
	páitsi	'frog sp.'		1
-iítsia **'CLF.BUNCH'**				
1	*manákhe*	'branch of açaí fruit'	20	2
2	*ponáma*	'branch of patawa fruit'	9	1
3	*porámo*	'branch of *açaí-chumbo* fruit'	8	1
-íiwi **'CLF.FLOWER'**				
1	*-íiwi*	'flower'	30	3
2	*háiko íiwi*	'flower'	20	2
3	*dóowiri*	'thorn'	9	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
***-íixi* 'CLF.SEED'**				
1	*-íixi(mi)*	'seed'	18	2
2	*manakhéexi*	'açaí seed'	10	1
2	*manákhe*	'açaí (seed, individual fruit)'	10	1
4	*hiipáda*	'small piece of stone'	9	1
4	*piiridzéexi*	'avocado seed'	9	1
4	*ponáma*	'patawa (seed, individual fruit)'	9	1
***-kénaa* 'CLF.BRANCH'**				
1	*-ke*	'branch'	30	3
2	*haikóke*	'tree branch'	10	1
2	*irímawa*	'branch of lemon tree'	10	1
4	*háiko*	'tree branch'	9	1
4	*-íiwi*	'branch with flowers'	9	1
***-khaa* 'CLF.CURVILINEAR'**				
1	*áapi*	'snake'	38	4
2	*iitsaákhaa*	'fishing line'	28	4
3	*ádapi/dápi*	'liana, vine'	21	4
4	*inípo*	'path'	15	2
5	*ówhii*	'fish sp.'	11	2
6	*okáana*	'liana sp.'	10	3
6	*-hinóoro/hinoorótti*	'hammock rope'	10	2
6	Ø	'thread'	10	1
9	*óoni*	'river'	8	1
9	*iakótti*	'language'	8	1
11	*oomápi*	'worm'	7	1
11	*ñawápo*	'creek'	7	1
11	*keewiíte*	'catfish sp.'	7	1
14	*ttáawaali*	'thread'	6	1
15	*komália*	'*tucum* fibre'	4	1
15	*omáwali*	'anaconda'	4	1
15	*mhóokoli*	'catfish sp.'	4	1
18	*kolíri*	'catfish sp.'	3	1
18	*tsikólee*	'strand of hair'	3	1
20	*iitsítsi*	'eel sp.'	2	1
21	*dóopo*	'lizard'	1	1
***-ko* 'CLF.HAMMOCK'**				
1	*piéta*	'hammock'	50	5

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
-koa **'CLF.SURFACE'**				
1	*póali*	'oven'	44	5
2	*dzakálee*	'community, village, town'	29	3
3	*méedza*~[*mesa*]	'table'	18	2
4	*káida*	'praia'	13	2
5	*kánaali*	'mirror'	10	1
5	*áada*	'traditional grater'	10	1
7	*-éekoa*	'forehead, surface'	8	1
7	*yóora*	'*jirau* shelf'	8	1
9	*hipákoa*	'flat stone'	7	1
10	*meeranípe*	'tray of manioc flour being roasted'	6	1
11	[*cama*]	'bed'	5	1
***-ma*^{1} 'CLF.PAIR'**				
1	*tshapáto*	'shoe/pair of shoes'	30	3
2	[*sandália*]	'pair of sandals'	18	2
3	*máawi*	'palm sp.'	17	2
4	*pheelóma*	'type of flute'	8	1
5	*yóora*	'shelf'	7	1
5	(inaudible)	'?'	7	1
***-ma*^{2} 'CLF.FEMALE'**				
1	*íinaro*	'woman'	40	4
2	*pedália*	'old woman'	9	1
-máka **'CLF.FABRIC'**				
1	[*toalha*]	'towel'	27	3
2	[*malhadeira*]	'fishnet'	19	2
3	*yamakátti*	'cloth'	18	2
4	*-palhéeda*	'cover'	10	1
4	*waxímaka*	'hand net (for fishing)'	10	1
4	[*lençol*]	'sheet'	10	1
7	*máawi*	'palm sp.'	8	1
7	[*lona*]	'canvas'	8	1
7	*-pálhee*	'sheet'	8	1
10	*iaromakátti*	'blanket'	7	1
10	*koyáma*	'carpet'	7	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
-na **'CLF.TRUNK'**				
1	*háiko*	'tree'	37	4
2	*tsíino*	'dog'	22	3
3	*héema*	'tapir, ox'	21	3
4	*palána*	'bunch of bananas'	15	2
5	*mokóto*	'large *panacu* container'	11	2
6	[*lata*]	'can (for measuring manioc flour)'	10	1
6	*míitsha*	'week'	10	1
8	*tsháako*	'bag'	9	1
9	*dzáawi*	'jaguar'	8	2
9	*káini*	'manioc'	8	1
11	*keníkhee*	'manioc stem'	7	1
12	*phíitsi*	'agouti (rodent sp.)'	6	2
12	[*gato*]	'cat'	6	1
12	*nawíki*	'person'	6	1
12	*máapa*	'sugar cane'	6	1
16	*pówe*	'monkey (generic)'	5	1
16	*ttíiña*	'*caraná* palm'	5	1
18	*kóitsi*	'curassow'	4	1
18	*káaparo*	'woolly monkey'	4	1
20	*néeri*	'deer'	2	1
	ónoli	'heron sp.'		1
	táali	'fish sp.'		1
	íiniri	'fish sp.'		1
	kettínali	'fish sp.'		1
-náko **'CLF.BUNDLE'**				
1	*kiníkhii*	'bundle of *maniva*'	30	3
2	*ttíiña*	'bundle of *caraná* branches'	17	2
2	*haikóapo*	'bundle of sticks'	17	2
4	*máapa*	'bundle of sugar cane'	10	1
4	*ttidzéena*	'bundle of firewood'	10	1
6	*kóona*	'bundle of a certain type of liana'	9	1
6	*tsikólee*	'tied up hair'	9	1
8	*ttídzee*	'bundle of firewood'	8	1
8	*hawádza*	'bundle of *ingá* pods'	8	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
-pa **'CLF.PACKAGE'**				
1	*mokóto*	'*panacu* container'	27	3
2	*hiemakalepátti*	'fish wrapped in a leaf'	17	2
3	*tsháako*	'bag'	10	1
3	*-hiríko*	'young leaf (of palm tree)'	10	1
3	*ttéephe*	'*piaba* fish wrapped in a leaf'	10	1
6	*póapoa*	'flat reed of *arumã* (for basketry)'	9	1
6	*yokíra*	'small bag/package of salt'	9	1
8	[*bolacha*]	'package of bisquits'	8	1
9	[*café*]	'package of coffee'	7	1
10	*-iítsia*	'branch, bunch (of fruit tree)'	6	1
10	[*leite*]	'package of milk'	6	1
-páwa **'CLF.RIVER'**				
1	*óoni*	'river, creek, tributary'	48	5
2	*ñawápo*	'creek'	28	3
3	*-ke*	'tributary'	9	1
-péko **'CLF.PATH'**				
1	*inípo*	'path'	40	4
2	*óoni*	'river'	26	3
3	*ñawápo*	'creek'	17	2
4	*pántti inóma*	'door'	15	2
5	*iakótti*	'idea, thought, language, word'	14	2
6	*tshéeto*	'*aturá* basket'	10	1
7	*líapo*	'type of linear basketry pattern'	6	1
8	*tsoopéki*	'window'	4	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
-phe **'CLF.LEAF'**				
1	*papéra*	'sheet of paper'	33	4
2	*panáphe*	'leaf (of a tree)'	20	2
3	*yamakátti*	'cloth'	16	2
4	*-phe*	'leaf, feather'	13	2
5	*tshaayápe*	'dress'	10	1
6	[*caderno*]	'notebook'	9	1
6	[*zinco*]	'sheet of zinc'	9	1
8	[*alumínio*]	'sheet of aluminium'	8	1
8	*manakhéphe*	'açaí leaf'	8	1
10	[*laranja*] *íphe*	'orange leaf'	7	1
11	*ipalheedátti*	'blanket'	6	1
11	*deríphe*	'leaf of *sororoca* (wild banana sp.)'	6	1
13	*palanáphe*	'banana leaf'	5	1
14	[*plástico*]	'tarpaulin'	4	1
15	[*lona*]	'canvas'	3	1
15	*máaliphe*	'fish sp.'	3	1
17	*píittiri makáphe/náphe*	'umbrella' (lit. bat's wing)	2	1
-pi **'CLF.TUBE'**				
1	*mókawa*	'rifle, gun'	27	3
2	*kéeri*	'month'	26	3
3	*manákhe*	'açaí palm'	24	3
4	*ttirolípi*	'*tipití*'	19	2
5	*póoperi∼pooperípi*	'*bacaba* stem'	18	2
6	*mawípi*	'blowpipe'	17	2
7	*-kodzópi*	'long path, stretch, river'	9	1
8	*ponáma*	'patawa palm'	8	1
9	*manakhépi*	'açaí stem'	7	1
10	*ponamápi*	'patawa stem'	6	1
11	*háiko*	'tree stem'	5	1
11	*-kódzoa*	'river bend'	5	1
12	*-yapíka*	'river bend'	4	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
***-póko* 'CLF.CIRCLE'**				
1	*kamhéro*	'bunch of Amazon grape'	34	4
2	*palána*	'comb of banana'	23	3
3	*dopítsi*	'sieve'	16	2
4	*kóphe*	'bundle of fish, school of fish'	14	2
5	*áaxi*	'yam'	10	1
5	*tshéeto*	'*aturá* basket'	10	1
7	*dáapa*	'bundle of paca'	9	1
8	*newíki*	'group of people'	8	1
8	*íitsiri*	'bundle of animals'	8	1
10	*néeri*	'bundle of deer'	7	1
11	*phíitsi*	'bundle of agouti'	6	1
11	*waláya*	'type of flat basket'	6	1
13	*waxímaka*	'hand net (for fishing)'	5	1
***-pokóda* 'CLF.STUMP'**				
1	*palána*	'banana tree; collection of bananas'	17	2
2	*haikopokóda*	'tree trunk'	10	1
2	*háiko*	'tree stump'	10	1
2	*-tanhída*	'tree stump'	10	1
2	*áaxi*	'yam'	10	1
2	*-koróda*	'tree stump'	10	1
7	*káini*	'collection of maniocs'	9	1
7	*haikópali*	'tree stump'	9	1
9	*manákhe*	'açaí seedling'	8	1
10	[*limão*]	'lemon tree'	7	1
***-ttáwalhe* 'CLF.CUT'**				
1	*háiko*	'piece of wood'	20	2
2	*kóphe*	'piece of fish'	17	2
3	Ø	'room'	10	1
3	*omáwali*	'piece of anaconda'	10	1
5	[*sabão*]	'piece of soap'	9	1
5	[*quarto*]	'room'	9	1
5	*dáapa*	'piece of paca'	9	1
8	*óoni*	'microregion/part of a river'	8	1
8	*-imaakaróda*	'room'	8	1
8	[*pilha*]	'battery'	8	1
11	*máapa*	'piece of sugar cane'	6	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
***-ttówhia* 'CLF.ROOM'**				
1	Ø	'room'	30	3
2	[*quarto*]	'room'	10	1
***-wa* 'CLF.SPACE'**				
1	*-nómawa*	'hole'	19	2
2	*pántti inóma*	'door'	10	1
2	[*lixo*]	'a pile of rubbish'	10	1
2	*haláaphi*	'buraco'	10	1
2	*dzakálee*	'community, town'	10	1
6	*-dzáanaa*	'wound'	9	1
6	*híiniri*	'pile of *ucuquí* fruit'	9	1
8	*dóomali*	'pile of umari fruit'	8	1
8	*ñawápo*	'creek'	8	1
8	*kawípe inómawa*	'entrance of ant's nest'	8	1
11	*kóowhe inómawa*	'entrance of leaf-cutter ant's nest'	7	1
11	*-eeríko*	'anus'	7	1
13	*dóopo inómawa*	'entrance of lizard's nest'	6	1
***-wána* 'CLF.SLICE'**				
1	*peéthe*	'piece/slice/half of *beiju*'	47	5
2	[*pizza*]	'pizza slice'	18	2
3	*dáapa*	'piece/half of paca'	16	2
4	*híipai*	'piece of land'	10	1
4	*kánaali*	'piece of (broken) mirror'	10	1
6	[*vidro*]	'piece of (broken) window'	9	1
6	*kiníki*	'part of manioc garden'	9	1
8	*paráto*	'piece of (broken) plate'	8	1
8	*-eétsha*	'piece of broken tooth'	8	1
8	*míitsi*	'piece of barbecued meat'	8	1
11	*phíitsi*	'piece of agouti'	7	1
11	*ttiíwe*	'piece of a broken paddle'	7	1
11	*déekai*	'piece of (broken) ceramic'	7	1
11	[*bolo*]	'slice of cake'	7	1
15	*héema*	'piece of tapir'	5	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
-wáta **'CLF.BRACT'**				
1	*-wáta*	'*curuatá*'	19	2
1	*kóoya*	'type of calabash bowl'	19	2
3	*manákhe íphe*	'*acaí* leaf'	10	1
3	*Ø*	'*curuatá*'	10	1
3	[*colher*]	'spoon(ful), ladle(ful)'	10	1
3	*meettadádali*	'empty box'	10	1
7	*ponáma íphe*	'patawa leaf'	9	1
7	*-íiwi*	'*curuatá*'	9	1
9	*póoperi íphe*	'*bacaba* leaf'	8	1
9	*átta*	'type of calabash bowl'	8	1
9	*panáphe*	'leaf'	8	1
9	*ponáma íiwi*	'*curuatá* of patawa'	8	1
13	*háiko íya*	'tree bark (fallen and rolled up)'	7	1
13	*píipiri íiwi*	'*curuatá* of peach palm'	7	1
13	*papéra*	'folded paper (funnel shape)'	7	1
16	*manákhe íiwi*	'*curuatá* of açaí'	6	1
-wáthe **'CLF.NODE'**				
1	*-wáthe*	'knot, tree knot'	17	2
2	*-wathéda*	'knot'	10	1
2	*máapa iwáthe*	'joint of sugar cane'	10	1
2	*máapa*	'joint of sugar cane'	10	1
5	*-nawathére*	'elbow'	9	1
5	*háiko*	'tree knot'	9	1
7	*ádapi*	'knot of a liana'	8	1
8	*iitshaákhaa iwáthe*	'knot on a fishing line'	6	1
9	*ttáawaali iwáthe*	'knot on a thread'	5	1
-xaa **'CLF.EXCREMENT'**				
1	*tsíino íixa*	'dog excrement'	19	2
1	*héema íixa*	'tapir/cow excrement'	19	2
3	*-íixa*	'excrement'	10	1
-ya **'CLF.SKIN'**				
1	*íita*	'canoe'	28	3
2	*háiko íya*	'tree bark'	25	3
3	*manákhe iwáta*	'*curuatá* of açaí'	20	2
4	*-ya*	'bark, skin, hide'	19	2
5	*taapáya*	'bark of a dead tree'	16	2
6	*áatti*	'chili pepper'	9	1
7	*ponáma iwáta*	'*curuatá* of patawa'	8	1

Continued on next page

Table B.1 (continued)

Order	Item	Meaning	Points	Tokens
***-yáwa* 'CLF.HOLE'**				
1	*-nómawa*	'hole'	27	3
2	*-dzáanaa*	'wound'	20	2
3	*kawawalikóaphi*	'hole in the ground'	10	1
3	*pántti inóma*	'door'	10	1
5	*kawípe inómawa*	'entrance of ant's nest'	8	1
5	*-ya*	'skin, bark'	8	1
5	*-aapówa*	'burrow, den (of an animal)'	8	1
8	*kóowhe inómawa*	'entrance of leaf-cutter ant's nest'	7	1
9	*dóopo inómawa*	'entrance of lizard's nest'	6	1

References

Acuña-Fariña, Juan Carlos (1999). "On apposition". In: *English Language and Linguistics* 3, 59–81.

Aikhenvald, Alexandra Y. (1994). "Classifiers in Tariana". In: *Anthropological Linguistics* 36, 405–465.

— (1996a). "Areal diffusion in Northwest Amazonia: The case of Tariana". In: *Anthropological Linguistics* 38, 73–116.

— (1996b). "Noun classes, classifiers and agreement in Baniwa (Arawak)". In: *Moscow Journal of Linguistics* 3, 7–33.

— (1998). "Warekena". In: *Handbook of Amazonian Languages. Vol. 3*. Ed. by Desmond C. Derbyshire & Geoffrey K. Pullum. Berlin: Mouton de Gruyter, 215–439.

— (1999). "The Arawak language family". In: *The Amazonian Languages*. Ed. by Robert M. W. Dixon & Alexandra Y. Aikhenvald. Cambridge: Cambridge University Press, 65–106.

— (2000). *Classifiers: A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.

— (2001). "Areal diffusion, genetic inheritance and problems of subgrouping: A North Arawak case study". In: *Areal Diffusion and Genetic Inheritance*. Ed. by Robert M. W. Dixon & Alexandra Y. Aikhenvald. Oxford: Oxford University Press, 167–194.

— (2002). *Language contact in Amazonia*. Oxford: Oxford University Press.

— (2003). *A grammar of Tariana, from northwest Amazonia*. Cambridge: Cambridge University Press.

— (2006). "Classifiers and noun classes, semantics". In: *Encyclopedia of language and linguistics, Volume 1*. Ed. by Keith Brown. Oxford: Elsevier, 463–470.

— (2007). "Classifiers in multiple environments. Baniwa of Içana/Kurripako—A North Arawak perspective". In: *International Journal of American Linguistics* 73.4, 475–500.

— (2010). *Language contact in Amazonia*. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y. (2012). *The Languages of the Amazon*. Oxford: Oxford University Press.
— (2016). *How Gender Shapes the World*. Oxford: Oxford University Press.
— (2019). "A view from the north: Genders and classifiers in Arawak languages of north-west Amazonia". In: *Genders and classifiers: A cross-linguistic typology*. Ed. by Alexandra Y. Aikhenvald & Elena I. Mihas. Oxford: Oxford University Press, 103–143.
— (2020). "Morhology in Arawak languages". In: *Oxford Research Encyclopedia of Linguistics*.
— (2021). "One of a Kind. On the Utility of Specific Classifiers". In: *Cognitive Semantics* 7, 232–257.

Albuquerque, Gabriel & Luiza Garnelo (2018). "Entre mundos: homens, serpentes e peixes em dois mitos baniwa". In: *In Estudos de Literatura Brasileira Contemporânea* 53, 129–147.

Allan, Keith (1977). "Classifiers". In: *Language* 53, 285–311.

Allassonnière-Tang, Marc, Olof Lundgren, Maja Robbers, Sandra Cronhamn, Filip Larsson, One-Soon Her, Harald Hammarström & Gerd Carling (2021). "Expansion by migration and diffusion by contact is a source to the global diversity of linguistic nominal categorization systems". In: *Humanities and Social Sciences Communications* 8.1, 1–6.

Aristotle (ca 350 BCE/2014). "Categories". In: *The Complete Works of Aristotle: The Revised Oxford Translation, One-Volume Digital Edition*. Ed. by Jonathan Barnes. Princeton: Princeton University Press, 25–70.

Arnold, Béat (2017). *Les canoës en écorce d'Amérique du Sud : de l'Amazonie à la Terre de Feu*. Le tour du monde en 80 pirogues, fascicule 3. Le Locle: Editions G d'Encre.

Assis, Elias (2001). "Patrões e fregueses no alto Rio Negro: As relações de dominação no discurso do povo Dâw". Bachelor's thesis. Manaus: Universidade Federal do Amazonas.

Baltar, Marcos Antonio Rocha (1995). "Sistema de classificação do Baníwa do Içana-Hohodene, Língua Aruak do Norte – Rio Negro". Master's thesis. Florianópolis: Universidade Federal de Santa Catarina.

Balykova, Kristina (2019). "What do oranges and hammers have in common? The classifier 'round' in Wa'ikhana and other East Tukano languages". In: *LIAMES: Línguas Indígenas Americanas* 19, 1–24.

Barnes, Janet (1990). "Classifiers in Tuyuca". In: *Amazonian Linguistics: Studies in Lowland South American Languages*. Ed. by Doris L. Payne. Austin: University of Texas Press, 273–292.

Berlin, Brent & Paul Kay (1969). *Basic color terms: their universality and evolution*. Berkeley: University of California Press.

Bezerra, Zenilson Agostinho (2005). *Gramática Koripako*. São Gabriel da Cachoeira: Missão Novas Tribos do Brasil.

Bickel, Balthasar, Bernard Comrie & Martin Haspelmath (2008). *Leipzig Glossing Rules: Conventions for interlinear morpheme-bymorpheme glosses*. Available online at https://www.eva.mpg.de/lingua/resources/glossing-rules.php.

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

Boley, Frederick & Paula Boley (1979). *Descrições preliminares da gramática da língua Baniua*. Manuscript. New Tribes of Brasil.

Booij, Geert (1996). "Inherent versus contextual inflection and the split morphology hypothesis". In: *Yearbook of morphology 1995*. Ed. by Geert Booij & Jaap van Marle. Dordrecht: Kluwer, 1–16.

Bowern, Claire (2015). *Linguistic fieldwork: a practical guide*. 2nd ed. Basingstoke: Palgrave Macmillan.

Bréal, Michel (1897). *Essai de sémantiqueue: science des significations*. Paris: Hachette.

Bybee, Joan L. (1985). *Morphology: a study of the relation between meaning and form*. Amsterdam: John Benjamins.

Cabalzar, Alonsio & Carlos Alberto Ricardo (2006). *Povos Indígenas do Rio Negro: uma introdução à socioambiental do noroeste da Amazônia brasileira*. 3rd ed. São Gabriel da Cachoeira, AM: FOIRN – Federação das Organizações Indígenas do Rio Negro/ISA – Instituto Socioambiental.

Carling, Gerd, ed. (2024). *Diachronic Atlas of Comparative Linguistics Online*. Frankfurt am Main. URL: https://diacl.uni-frankfurt.de (visited on 03/02/2023).

Cayón, Luis & Thiago Costa Chacon (2022). "Diversity, multilingualism and interethnic relations in the long-term history of the Upper Rio Negro region of the Amazon". In: *Interface Focus* 13.

Chacon, Thiago Costa (2007). "O sistema de classificação nominal do Tukáno". In: *Revista de Estudos e Pesquisas (FUNAI)* 4.2, 147–197.
— (2022). "Gender and Noun Class Markers in Kubeo (Tukanoan)". In: *Faits de Langues* 52.2, 11–49.

Contini-Morava, Ellen & Marcin Kilarski (2013). "Functions of Nominal Classification". In: *Language Sciences* 40, 263–299.

Corbett, Greville G. (1991). *Gender*. Cambridge: Cambridge University Press.
— (2000). *Number*. Cambridge: Cambridge University Press.

Corbett, Greville G. & Sebastian Fedden (2016). "Canonical Gender". In: *Journal of Linguistics* 52, 495–531.

Costa Oliveira, Thiago Lopes da (2015). "Os Baniwa, os Artefatos e a Cultura Material no Alto Rio Negro". Doctoral dissertation. Museu Nacional, Universidade Federal do Rio de Janeiro.

Craig, Colette (1986). "Jacaltec noun classifiers: a study in language and culture". In: *Noun Classes and Categorization*. Ed. by Colette Craig. Amsterdam: Benjamins, 263–293.
— (1992). "Classifiers in a Functional Perspective". In: *Layered Structure and Reference in a Functional Perspective: Papers from the Functional Grammar Conference, Copenhagen, 1990*. Ed. by Michael Fortescue, Peter Harder & Lars Kristofferson. Amsterdam & Philadelphia: John Benjamins, 277–301.

Crevels, Mily (2012). "Language endangerment in South America: The clock is ticking". In: *The indigenous languages of South America: A comprehensive guide*. Ed. by Lyle Campbell & Verónica Grondona. Volume 2 of *The World of Linguistics*. Berlin: Mouton de Gruyter, 167–234.

Croft, William (1994). "Semantic universals in classifier systems". In: *Word* 45, 145–171.

da Silva, Franklin Paulo Eduardo (2021). "Pobres e Miseráveis? Preconceitos contra Povos Indígenas na Região do Alto Rio Negro (AM)". In: *Mediações* 26.1, 219–233.

Dahl, Östen (2004). *The growth and maintenance of linguistic complexity*. Amsterdam: Benjamins.

Danielsen, Swintha (2007). *Baure: An Arawak language of Bolivia*. Indigenous Languages of Latin America 6. Leiden: CNWS.

de Saussure, Ferdinand (1959). *Course in General Linguistics*. Ed. by Charles Bally & Albert Sechehaye. Trans. by Wade Baskin. (First published in 1915 as *Cours de Linguistique Générale*. Paris: Pyot.) New York: Philosophical Library.

de Souza, Erick Marcelo Lima (2012). "Estudo Fonológico da língua Baniwa-Kuripako". Master's thesis. Campinas: Universidade Estadual de Campinas.

Denny, J. Peter (1976). "What are noun classifiers good for?" In: *Papers from the twelfth regional meeting of the Chicago Linguistic Society*. Ed. by Salikoko S. Mufwene, Carol A. Walker & Sanford B. Steever. Chicago: Chicago Linguistic Society, 122–32.

— (1979). "The 'Extendedness' Variable in Classifier Semantics: Universal Semantic Features and Cultural Variation". In: *Ethnology: Boas, Sapir and Whorf Revisited*. Ed. by Madeleine Mathiot. The Hague: Mouton, 97–119.

Derbyshire, Desmond & Doris L. Payne (1990). "Noun classification systems of Amazonian languages". In: *Amazonian Linguistics: Studies in Lowland South American Languages*. Ed. by Doris L. Payne. Austin: University of Texas Press, 243–72.

Dixon, Robert M. W. (1972). *The Dyirbal Language of North Queensland*. Cambridge: Cambridge University Press.

— (1982). *Where have all the adjectives gone? and other essays in semantics and syntax*. Berlin: Mouton.

— (1986). *Noun classes and categorization*. Amsterdam/Philadelphia: John Benjamins.

— (2012). *Basic linguistic theory. Vol. 3: Further grammatical topics*. Oxford: Oxford University Press.

Dotte, Anne-Laure (2017). "Dynamism and Change in the Possessive Classifier System of Iaai". In: *Oceanic Linguistics* 56.2, 339–363.

Dunn, Saskia (2022). "The Shape of Arawakan Classifiers: An Internal Source for the Arawakan Classifier System(s)". Master's thesis. Leiden University.

Dunn, Saskia & Françoise Rose (forthcoming). "Arawak classifiers: a South American multilocus classifier system". In: *The Oxford Handbook of Gender and Classifiers*. Ed. by Niels Schiller & Tanja Kupisch. Oxford: Oxford University Press.

Eberhard, David M., Gary F. Simons & Charles D. Fennig (2024). *Ethnologue: Languages of the World*. URL: http://www.ethnologue.com (visited on 05/27/2024).

Epps, Patience (2007). "Birth of a noun classification system: The case of Hup". In: *Language Endangerment and Endangered Languages: Linguistic and Anthro-*

pological Studies with Special Emphasis on the Languages and Cultures of the Andean-Amazonian Border Area. Ed. by Leo Wetzels. Leiden: Publications of the Research School of Asian, African, & Amerindian Studies (CNWS), 107–128.

Epps, Patience (2008). *A grammar of Hup*. Berlin, New York: Mouton.

— (2018). "Contrasting linguistic ecologies: Indigenous and colonially mediated language contact in northwest Amazonia". In: *Language & Communication* 62.1, 156–169.

— (2020). "Amazonian linguistic diversity and its sociocultural correlates". In: *Language dispersal, diversification, and contact: A global perspective*. Ed. by Mily Crevels & Pieter Muysken. Oxford: Oxford University Press, 275–290.

Epps, Patience & Lev Michael (2017). "The Areal Linguistics of Amazonia". In: *The Cambridge Handbook of Areal Linguistics*. Ed. by Raymond Hickey. Cambridge: Cambridge University Press, 934–963.

Epps, Patience & Karolin Obert (2022). "Naduhup Languages and the Typology of Nominal Classification". In: *Cadernos de Etnolingüística* 10.1, 1–25. ISSN: 1946-7095.

Epps, Patience & Andrés Pablo Salanova (2013). "The Languages of Amazonia". In: *Tipití: Journal of the Society for the Anthropology of Lowland South America* 11.1, 1–28.

Epps, Patience & Kristine Stenzel (2013). *Upper Rio Negro: Cultural and Linguistic Interaction in Northwestern Amazonia*. Rio de Janeiro: Museu do Índio-FUNAI. URL: http://www.museunacional.ufrj.br/ppgas/livros_ele.html.

Epps, Patience, Rik van Gijn & Nicholas Q. Emlen (forthcoming). "South America". In: *The Oxford Handbook of Historical Morphology*. Ed. by Peter Arkadiev & Franz Rainer. Advance online publication. Oxford: Oxford University Press.

Erben Johansson, Niklas & Sandra Cronhamn (2023). "Vocal iconicity in nominal classification". In: *Language and Cognition* 15.2, 266–291. ISSN: 1946-7095.

Eriksen, Love (2011). "Nature and Culture in Prehistoric Amazonia: Using G.I.S. to reconstruct ancient ethnogenetic processes from archaeology, linguistics, geography, and ethnohistory". Doctoral dissertation. Lund University, Human Ecology Division.

Evans, Nicholas (2010). "Semantic typology". In: *The Oxford handbook of linguistic typology*. Ed. by Jae Jung Song. Oxford: Oxford University Press, 504–533.

Farmer, Stephanie (2015). "Establishing reference in Máíhĩki". Doctoral dissertation. UC Berkeley.

Fedden, Sebastian & Greville G. Corbett (2017). "Gender and Classifiers in Concurrent Systems: Refining the Typology of Nominal Classification". In: *Glossa: A Journal of General Linguistics* 2.1, 1–47. ISSN: 1946-7095.

Floyd, Simeon (2013). "Semantic transparency and cultural calquing in the north west Amazon". In: *Upper Rio Negro: Cultural and Linguistic Interaction in Northwestern Amazonia*. Ed. by Patience Epps & Kristine Stenzel. Rio de Janeiro: Museu do Índio–FUNAI, 271–306.

François, Alexandre (1999). "L'illusion des classificateurs". In: *Faits de langues* 7.14, 165–175.

Franjieh, Michael (2012). "Possessive classifiers in North Ambrym, a language of Vanuatu: explorations in semantic classification". Doctoral dissertation. SOAS, University of London.

Franjieh, Michael, Greville G. Corbett & Alexandra Grandison (2020). *How classifiers become gender in Oceania: An experimental approach*. Paper presented at the 53rd annual meeting of the Societas Linguistica Europaea, Bucharest, Romania, August 26-29, 2020.

Frege, Gottlob (1980). *Translations from the Philosophical Writings of Gottlob Frege*. Ed. by Peter Geach & Max Black. Oxford: Blackwell.

Garnelo, Luiza (2007). "Cosmology, environment, and health: Baniwa food myths and rituals". In: *História, Ciências, Saúde — Manguinhos* 14, supplement, 191–212.

Geeraerts, Dirk (2006). "Prototype theory: Prospects and problems of Prototype Theory". In: *Cognitive linguistics: Basic readings*. Ed. by Dirk Geeraerts. Berlin & New York: Mouton de Gruyter, 141–166.

Gomez-Imbert, Elsa (1982). "De la forme et du sens dans la classification nominale en tatuyo (langue Tukano Orientale d'Amazonie Colombienne)". Doctoral dissertation. Université de Paris-Sorbonne.

— (1996). "When animals become 'rounded' and 'feminine': Conceptual categories and linguistic classification in a multilingual setting". In: *Rethinking linguistic relativity*. Ed. by John J. Gumpertz & Stephen C. Levinson. Cambridge: Cambridge University Press.

— (2007). "Tukanoan Nominal Classification. The Tatuyo System". In: *Language Endangerment and Endangered Languages. Linguistic and anthropological studies*. Ed. by W. Leo Wetzels. Leiden: CNWS Publications, 402–428.

Gonçalves, Artur Garcia (2018). “Para uma dialetologia baniwa-koripako do rio Içana”. Master’s thesis. Universidade de Brasília.

González-Ñáñez, Omar (1985). “Los numerales en un dialecto Curripaco”. In: *Boletín de Lingüística* 5, 15–29.

Granadillo, Tania (2006). “An Ethnographic Account of Language Documentation among the Kurripako of Venezuela”. Doctoral dissertation. University of Arizona.

Greenberg, Joseph H. (1972). “Numeral classifiers and substantival number: problems in the genesis of a linguistic type”. In: *Working Papers on Language Universals*, 1–39.

— (1978). “How does a language acquire gender markers?” In: *Universals of Human Language, III: Word Structure*. Ed. by Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik. Stanford: Stanford University Press, 47–82.

Grinevald, Colette (2000). “A Morphosyntactic Typology of Classifiers”. In: *Systems of Nominal Classification*. Ed. by Gunter Senft. Cambridge: Cambridge University Press, 50–92.

— (2002). “Making sense of nominal classification systems. Noun classifiers and the grammaticalization variable”. In: *New Reflections on Grammaticalization*. Ed. by Ilse Wischer & Gabriele Diewald. Amsterdam/Philadelphia: John Benjamins, 259–275.

— (2004). “Classifiers”. In: *Morphologie / Morphology; Ein internationales Handbuch zur Flexion und Wortbildung / An international handbook on inflection and word-formation*. Ed. by Geert E. Booij, Christian Lehmann, Joachim Mugdan & Stavros Skopeteas. Vol. II. Berlin: Walter de Gruyter, 1016–1031.

— (2015). “Classifiers, linguistics of”. In: *International Encyclopedia of the Social & Behavioral Sciences*, 811–818.

Grinevald, Colette & Frank Seifart (2004). “Noun classes in African and Amazonian languages: Towards a comparison”. In: *Linguistic Typology* 8, 243–285.

Haarh, Mads (2020). *RANDOM.ORG: True Random Number Service*. URL: https://www.random.org (visited on 03/09/2020).

Hammarström, Harald, Robert Forkel, Martin Haspelmath & Sebastian Bank (2024). *Glottolog 5.0*. Available online at http://glottolog.org. URL: https://doi.org/10.5281/zenodo.10804357 (visited on 05/27/2024).

Haspelmath, Martin (2021). “Towards standardization of morphosyntactic terminology for general linguistics”. In: *Linguistic categories, language description and linguistic typology*. Ed. by Luca Alfieri, Giorgio Francesco Arcodia & Paolo Ramat. Amsterdam & Philadelphia: John Benjamins, 35–57.

Haspelmath, Martin & Andrea Sims (2010). *Understanding Morphology*. 2nd edition. London: Hodder Education.

Haspelmath, Martin & Uri Tadmor (2009). *World Loanword Database*. Leipzig. URL: http://wold.clld.org.

Heine, Bernd (2004). "Grammaticalization". In: *The handbook of historical linguistics*. Ed. by Brian D. Joseph & Richard D. Janda. Oxford: Blackwell, 575–601.

Heine, Bernd & Tania Kuteva (2002). "On the Evolution of Grammatical Forms". In: *The Transition to Language*. Ed. by Alison Wray. Oxford: Oxford University Press.

Her, One-Soon, Harald Hammarström & Marc Allassonnière-Tang (2022). "Defining numeral classifiers and identifying classifier languages of the world". In: *Linguistics Vanguard* 8.1, 151–164.

Hill, Jonathan D. (1984). "Social equality and ritual hierarchy: The Arawakan Wakuenai of Venezuela". In: *American Ethnologist* 11, 528–544.

— (1988). "The Soft and the Stiff: Ritual Power and Mythic Meaning in a Northern Arawakan Classifier System". In: *Antropólogica* 69, 55–77.

— (1993). *Keepers of the Sacred Chants: The Poetics of Ritual Power in an Amazonian Society*. University of Arizona Press.

— (2009a). *Made-from-Bone: Trickster Myths, Music, and History from the Amazon*. University of Illinois Press.

— (2009b). "The celestial umbilical cord: wild palm trees, adult male bodies, and sacred wind instruments among the Wakuénai of Venezuela". In: *Journal for the Study of Religion, Nature and Culture* 3.1, 99–125.

— (n.d.). *Curripaco Collection of Jonathan Hill*. PID Collection: 983. Accessed 9 October 2024. The Archive of the Indigenous Languages of Latin America (AILLA), ailla.utexas.org.

Hornborg, Alf (2005). "Ethnogenesis, regional interaction, and ecology in prehistoric Amazonia: towards a systemic perspective". In: *Current Anthropology* 46.4, 589–620.

Hugh-Jones, Christine (1979). *From the Milk River: Spatial and temporal processes in Northwest Amazonia*. Cambridge: Cambridge University Press.

Instituto Brasileiro de Geografia e Estatística (2023). *Censo Demográfico 2022: indígenas: primeiros resultados do universo*. Rio de Janeiro.

Key, Mary Ritchie & Bernard Comrie (2015a). *The Intercontinental Dictionary Series*. Leipzig. URL: http://ids.clld.org/.

Kilarski, Marcin (2013). *Nominal classification: A history of its study from the classical period to the present*. Amsterdam: John Benjamins.

Krasnoukhova, Olga (2012). "The Noun Phrase in the Languages of South America". Doctoral dissertation. Radboud University Nijmegen.

Lakoff, George (1986). "Classifiers as a reflection of mind". In: *Noun classes and categorization*. Ed. by Colette Craig. Amsterdam: John Benjamins, 13–51.

— (1987). *Women, Fire, and Dangerous Things: What Categories Reveal about the Mind*. Chicago: University of Chicago Press.

Lehmann, Christian (2015). *Thoughts on grammaticalization*. Language Science Press.

Löbel, Elisabeth (2000). "Classifiers versus genders and noun classes: A case study in Vietnamese". In: *Gender in Grammar and Cognition*. Ed. by Barbara Unterbeck, Matti Rissanen, Terttu Nevalainen & Mirja Saari. Berlin, New York: De Gruyter Mouton, 259–320.

Lucy, John A. (2000). "Systems of nominal classification: a concluding discussion". In: *Systems of nominal classification*. Ed. by Gunter Senft. New York: Cambridge University Press, 326–341.

Lyons, John (1977). *Semantics. 2 vols*. Cambridge: Cambridge University Press.

Matras, Yaron (2009). "The Borrowability of Structural Categories". In: *Grammatical Borrowing in Cross-linguistic Perspective*. Ed. by Yaron Matras & Jeanette Sakel. Berlin: Mouton de Gruyter, 31–73.

Max Planck Institute for Psycholinguistics, The Language Archive (2023). *ELAN*. Version 6.7. Nijmegen. URL: https://archive.mpi.nl/tla/elan.

Mayer, Mercer (1967). *A boy, a dog, and a frog*. New York: Dial Books for Young Readers.

Melgueiro, Edilson Martins (2009). "Sobre a natureza, expressão formal e escopo da classificação lingüística das entidades na concepção do mundo dos Baníwa". Master's thesis. Universidade de Brasília.

Meyer, Cherry Lynn (2020). "Noun Categorization in Ojibwe: gender and classifiers". Doctoral dissertation. University of Chicago.

Michael, Lev (2021). "The classification of South American languages". In: *Annual Review of Linguistics* 7, 329–349.

Michaelis, Susanna Maria, Philippe Maurer, Martin Haspelmath & Magnus Huber (2013a). *The Atlas of Pidgin and Creole Language Structures*. Oxford: Oxford University Press.
— (2013b). *The Survey of Pidgin and Creole Languages. Vol. 1, English-based and Dutch-based languages*. Oxford: Oxford University Press.
— (2013c). *The Survey of Pidgin and Creole Languages. Vol. 2, Portuguese-based, Spanish-based, and French-based languages*. Oxford: Oxford University Press.

Mihas, Elena I. (2019). "Genders and classifiers in Kampa (Arawak) languages of Peru". In: *Genders and classifiers: A cross-linguistic typology*. Ed. by Alexandra Y. Aikhenvald & Elena I. Mihas. Oxford: Oxford University Press, 30–66.

Mithun, Marianne (1986). "The convergence of noun classification systems". In: *Noun Classes and Categorization*. Ed. by Colette Craig. Amsterdam & Philadelphia: Benjamins, 379–397.

Moravcsik, Edith (2003). "Inflectional morphology in the Hungarian noun: A typological assessment". In: *Noun Phrase Structure in the Languages of Europe*. Ed. by Frans Plank. Berlin & New York: Mouton de Gruyter, 113–252.

Nichols, Johanna (1988). "On alienable and inalienable possession". In: *In honor of Mary Haas*. Ed. by William Shipley. Berlin: Mouton de Gruyter, 475–521.
— (1992). *Linguistic diversity in space and time*. Chicago: The University of Chicago Press.

Nimuendajú, Curt (1932). "Idiomas indígenas del Brasil". In: *Revista del Instituto de Etnología de La Universidad Nacional de Tucumán* 2.2, 590–618.
— (1950). "Reconhecimento dos rios Içana, Ayari e Uaupés. Relatório apresentado ao Serviço de Proteção aos Indios do Amazonas e Acre, 1927". In: *Journal de la Société des Américanistes* 39, 125–182.

Noble, G. Kingsley (1965). "Proto-Arawakan and its descendants". In: *International Journal of American Linguistics* 31.3, 1–129.

Nurmio, Silva (2023). "Towards a typology of singulatives: Definition and overview of markers". In: *Number categories: Dynamics, Contact, Typology*. Ed. by Deborah Arbes. Berlin: de Gruyter, 155–182. URL: https://doi.org/10.1515/9783110986600-005.

Obert, Karolin (2019). "The linguistic encoding of space in Dâw". Doctoral dissertation. Universidade de São Paulo.

Ortmann, Albert (2018). "Connecting the typology and semantics of nominal possession: Alienability splits and the morphology–semantics interface". In: *Morphology* 28.1, 99–144.

Palmer, Stephen E. (1999). *Vision science: Photons to phenomenology*. Cambridge, MA: MIT Press.

Passer, Matthias (2016). "The typology and diachrony of nominal classification". Doctoral dissertation. University of Amsterdam.

Payne, David L. (1991). "A classification of Maipuran (Arawakan) languages based on shared lexical retentions". In: *Handbook of Amazonian Languages. Vol. 3*. Ed. by Desmond C. Derbyshire & Geoffrey K. Pullum. Berlin: Mouton de Gruyter, 355–499.

Payne, Doris L. (1985). "Inflection versus derivation: Is there a difference?" In: *Proceedings of the First Annual Pacific Linguistics Conference*. Ed. by Scott DeLancey & Russel S. Tomlin. Eugene: University of Oregon, 247–260.
— (1986). "Noun classification in Yagua". In: *Noun classes and categorization*. Ed. by Colette Craig. Amsterdam: John Benjamins, 113–131.
— (1987). "Noun classification in the Western Amazon". In: *Language Sciences* 9.1, 22–44.
— (2007). "Source of the Yagua Nominal Classification System". In: *International Journal of American Linguistics* 73.4, 447–474.

Pepper, Steve (2020). "The typology and semantics of binominal lexemes. Noun-noun compounds and their functional equivalents". Doctoral dissertation. University of Oslo.

PGTA Nadzoeri: Plano de Gestão Territorial e Ambiental da Organização Baniwa e Koripako (2021). São Gabriel da Cachoeira: FOIRN/ISA.

Pike, Kenneth Lee (1954). *Language in relation to a unified theory of the structure of human behavior: Part I*. Glendale, California: Summer Institute of Linguistics.

"Portuguese Dictionary" (2015b). In: *The Intercontinental Dictionary Series*. Ed. by Mary Ritchie Key & Bernard Comrie. Leipzig: Max Planck Institute. URL: http://ids.clld.org/contributions/178.

Ramirez, Henri (1997). *A fala Tukano dos Ye'pâ-Masa. Tomo I: Gramática*. Manaus: Inspetoria Salesiana Missionária da Amazônia, CEDEM.
— (2001a). *Dicionário da Língua Baniwa*. Manaus: Editora da Universidade do Amazonas.
— (2001b). "Uma Gramática do Baniwa do Içana". Unpublished manuscript.
— (2020a). *Enciclopédia das línguas Arawak: Acrescida de seis novas línguas e dois bancos de dados. Vol. I: Um gramática do Baniwa do Içana*. Curitiba, Brazil: Editora CRV.

— (2020b). *Enciclopédia das línguas Arawak: Acrescida de seis novas línguas e dois bancos de dados. Vol. II: O subgrupo Japurá-Colômbia (JC).* Curitiba, Brazil: Editora CRV.
— (2020c). *Enciclopédia das línguas Arawak: Acrescida de seis novas línguas e dois bancos de dados. Vol. III: Os outros subgrupos Arawak.* Curitiba, Brazil: Editora CRV.
— (2020d). *Enciclopédia das línguas Arawak: Acrescida de seis novas línguas e dois bancos de dados. Vol. IV: Bancos de dados.* Curitiba, Brazil: Editora CRV.

Ricardo, Beto (2001). *Arte Baniwa: cestaria de arumã.* 3rd ed. São Gabriel da Cachoeira: FOIRN/ISA/OIBI.

Ricardo, Fany, Tatiane Klein & Tiago Moreira dos Santos (2023). *Povos indígenas no Brasil: 2017/2022.* 2nd ed. São Paulo: Instituto Socioambiental.

Rijkhoff, Jan (2002). *The noun phrase.* Oxford: Oxford University Press.

Rosch, Eleanor (1973). "Natural categories". In: *Cognitive psychology* 4.3, 328–350.
— (1975). "Cognitive representations of semantic categories". In: *Journal of experimental psychology: General* 104.3, 192–233.
— (1978). "Principles of categorization". In: *Cognition and Categorization.* Ed. by Eleanor Rosch & Barbara B. Lloyd. Hillsdale, NJ: Lawrence Erlbaum, 27–48.

Rosch, Eleanor, Carolyn B. Mervis, Wayne D. Gray, David M. Johnson & Penny Boyes-Braem (1976). "Basic objects in natural categories". In: *Cognitive Psychology* 8, 383–439.

Rosch, Eleanor, Carol Simpson & R. Scott Miller (1976). "Structural bases of typicality effects". In: *Journal of Experimental Psychology: Human perception and performance* 2.4, 491–502.

Rose, Françoise (2024a). *Mojeño Trinitario classifiers: a multilocus and multifunctional system.* Cadernos de Etnolingüística, Série Monografias 7.
— (2024b). *Thoughts on the development of multilocus classifier systems.* Paper presented at Classifiers in the Arawakan languages (ClAra), Lyon, France, September 9, 2024.

Rose, Françoise & An Van linden (2023). "The derivational use of classifiers in Western Amazonia". In: *Binominal Lexemes in Cross-Linguistic Perspective: Towards a Typology of Complex Lexemes.* Ed. by Steve Pepper & Francesca Masini & Simone Mattiola. Berlin, Boston: De Gruyter Mouton, 237–276. URL: https://doi.org/10.1515/9783110673494-008.

Saeed, John I. (2009). *Semantics.* 3rd ed. Malden, MA: Wiley-Blackwell.

Sanches, Mary & Linda Slobin (1973). "Numeral classifiers and plural marking: an implicational universal". In: *Working Papers in Language Universals* 11, 1–22.

Santos-Granero, Fernando (2002). "The Arawakan Matrix: Ethos, Langage and History in Native South America". In: *Comparative Arawakan Histories. Rethinking Language Family and Culture Area in Amazonia*. Ed. by Jonathan D. Hill & Fernando Santos-Granero. Urbana & Chicago: University of Illinois Press, 25–50.

Scalise, Sergio (1988). "Inflection and derivation". In: *Linguistics* 26.4, 561–582.

Scheibe, Paul (1957). "Phonemic analysis of Baniua". Manuscript. Museu Nacional, Rio de Janeiro.

Schmidt, Annette (1985). *Young people's Dyirbal: an example of language death from Australia*. Cambridge: Cambridge University Press.

Seifart, Frank (2005). "The structure and use of shape-based noun classes in Miraña (North West Amazon)". Doctoral dissertation. Radboud University, Nijmegen.
— (2007). "The prehistory of nominal classification in Witotoan languages". In: *International Journal of American Linguistics* 73.4, 411–445.
— (2010). "Nominal classification". In: *Language and Linguistics Compass* 4, 719–736.
— (2012). "The principle of morphosyntactic subsystem integrity in language contact: Evidence from morphological borrowing in Resígaro (Arawakan)". In: *Diachronica* 29.4, 471–504.

Seifart, Frank & Doris L. Payne (2007). "Nominal classification in the North West Amazon: Issues in areal diffusion and typological characterization". In: *International Journal of American Linguistics* 73.4, 381–387.

Senft, Gunter (2000). "What do we really know about nominal classification systems?" In: *Systems of nominal classification*. Ed. by Gunter Senft. Cambridge: Cambridge University Press, 11–49.
— (2007). "Nominal classification". In: *The Oxford handbook of cognitive linguistics*. Ed. by Dirk Geeraerts & Hubert Cuyckens. Oxford: Oxford University Press, 676–696.

Shulist, Sarah (2018a). "Signs of Status: Language Policy, Revitalization, and Visibility in Urban Amazonia". In: *Language Policy* 17.1, 523–543.
— (2018b). *Transforming Indigeneity: Urbanization and Language Revitalization in the Brazilian Amazon*. Toronto: University of Toronto Press.

Singer, Ruth (2018). "Beyond the classifier/gender dichotomy. The role of flexibility in a more integrated typology of nominal classification". In: *Non-Canonical Gender Systems*. Ed. by Sebastian Fedden, Jenny Audring & Greville G. Corbett. Oxford: Oxford University Press, 100–128.

Sloetjes, Han & Peter Wittenburg (2008). "Annotation by category - ELAN and ISO DCR". In: *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*.

Sorensen Jr., Arthur P. (1967). "Multilingualism in the Northwest Amazon". In: *American Anthropologist* 69, 670–684.

Steele, Susan (1978). "Word order variation. A typological study". In: *Universals of human language. Vol. 4: Syntax*. Ed. by Joseph H. Greenberg, Charles A. Ferguson & Edith A. Moravcsik. Stanford: Stanford University Press, 585–623.

Stenzel, Kristine (2005). "Multilingualism in the Northwest Amazon, revisited". In: *Memorias del Congreso de Idiomas Indígenas de Latinoamérica–II*. Ed. by I. I. L. Martín. University of Texas at Austin, 1–28.

— (2013). *A reference grammar of Kotiria (Wanano)*. Lincoln: University of Nebraska Press.

Stradelli, Ermanno (1929). *Vocabulários da Língua Geral Portuguez-Nheêngatú e Nheêngatú-Portuguez*. Rio de Janeiro: Revista do Instituto Historico-Geografico.

Swadesh, Morris (1950). "Salish Internal Relationships". In: *International Journal of American Linguistics* 16.4, 157–167.

Tadmor, Uri (2009). "Loanwords in the World's Languages: Findings and results". In: *Loanwords in the World's Languages: A Comparative Handbook*. Ed. by Martin Haspelmath & Uri Tadmor. Berlin: De Gruyter Mouton, 55–75.

Tang, Marc & One-Soon Her (2019). "Insights on the Greenberg-Sanches-Slobin generalization: Quantitative typological data on classifiers and plural markers". In: *Folia Linguistica* 53.2, 297–331.

Taylor, Gerald (1991). *Introdução à língua Baniwa do Içana*. Campinas: Editora da Unicamp.

Thomason, Sarah Grey & Terrence Kaufman (1988). *Language contact, creolization and genetic linguistics*. Berkeley & Los Angeles: University of California Press.

Valadares, Simoni M. B. (1993). "Aspectos fonológicos e morfológicos da língua Kurripáku (Kumandáminanai e Ayáneni)". Master's thesis. Universidade Federal da Santa Catarina.

van der Voort, Hein (2004). *A grammar of Kwaza*. Berlin: Mouton de Gruyter.

van Gijn, Rik, Katharina Haude & Pieter Muysken (2011). "Subordination in South America: An overview". In: *Subordination in native South American languages*. Ed. by Rik van Gijn, Katharina Haude & Pieter Muysken. Amsterdam: John Benjamins Publishing Company, 1–23.

Vittrant, Alice (2005). "Classifier systems and noun categorization devices in Burmese". In: *Proceedings of the twenty-eighth Annual Meeting of the Berkeley Linguistics Society*, 129–148.

Weber, David J. (2010). "The structural status of Bora classifiers". In: *Work Papers of the Summer Institute of Linguistics, University of North Dakota Session* 46.

Wikipedia (2024a). *Karu language — Wikipedia, The Free Encyclopedia*. URL: http://en.wikipedia.org/w/index.php?title=Karu%5C%20language&oldid=1225728197 (visited on 05/27/2024).

— (2024b). *List of glossing abbreviations — Wikipedia, The Free Encyclopedia*. URL: https://en.wikipedia.org/wiki/List_of_glossing_abbreviations (visited on 07/16/2024).

Wojtylak, Katarzyna (2016). "Classifiers as Derivational Markers in Murui (Northwest Amazonia)". In: *Word-Formation Across Languages*. Ed. by Lívia Körtvélyessy, Pavol Štekauer & Salvador Valera. Newcastle-upon-Tyne: Cambridge Scholars Publishing, 393–425.

Wright, Robin M. (1981). "History and Religion of the Baniwa Peoples of the Upper Rio Negro Valley (2 volumes)". Doctoral dissertation. Stanford University.

— (1992a). "Guardians of The Cosmos: Baniwa Shamans and Prophets (Part I)". In: *History of Religions* 32.1, 32–58.

— (1992b). "Guardians of The Cosmos: Baniwa Shamans and Prophets (Part II)". In: *History of Religions* 32.2, 126–145.

— (1998). *Cosmos, Self and History in Baniwa Religion. For Those Unborn*. Austin, TX: University of Texas Press.

— (2009). "The Fruit of Knowledge and the Bodies of the Gods: Religious Meanings of Plants among the Baniwa". In: *Journal for the Study of Religion, Nature, and Culture* 3.1, 126–153.

— (2020). *Baniwa of the Aiary and Içana Collection of Robin M. Wright Finding Aid*. PID Set: 14920. Accessed 9 October 2024. The Archive of the Indigenous Languages of Latin America (AILLA), ailla.utexas.org.

— (1993/1994). “Umawali. Hohodene myths of the Anaconda, father of fish”. In: *Bulletin de la Société Suisse des Américanistes* 57/58, 37–48.

Wright, Robin M. & Geraldo Andrello (2002). *Baniwa*. Povos Indígenas no Brasil. URL: https://pib.socioambiental.org/ (visited on 11/03/2024).

Zubin, David A. & Mitsuaki Shimojo (1993). “How “General” Are General Classifiers? With Special Reference to *ko* and *tsu* in Japanese”. In: *Proceedings of the Nineteenth Annual Meeting of the Berkeley Linguistics Society: General Session and Parasession on Semantic Typology and Semantic Universals*. Ed. by Joshua S. Guenter, Barbara A. Kaiser & Cheryl C. Zoll, 490–502.

Zucchi, Alberta (2002). “A New Model of the Northern Arawakan Expansion”. In: *Comparative Arawakan Histories: Rethinking language family and culture area in Amazonia*. Ed. by Jonathan D. Hill & Fernando Santos-Granero. Urbana & Chicago: University of Illinois Press, 199–222.

A grammar of Baniwa classifiers

Nominal classification systems provide a unique window into the intersection of grammar, semantics, and cognition. Found in more than half of the world's languages, these systems possess both universal and language-specific properties. Nominal classification systems of a specific type, featuring classifiers marked in multiple morphosyntactic loci, are found in many languages in Northwestern Amazonia. These systems are of particular typological interest as they share properties with several other types of nominal classification systems, yet few of them have been described in detail. A classifier system of this type is found in the Arawakan language Baniwa, spoken by a few thousand people in Northwestern Brazil.

This thesis provides a detailed analysis of the Baniwa classifier system from phonological, morphological, syntactic, functional, semantic, typological, historical and contact perspectives, based on first-hand data from field work and a combination of descriptive and experimental approaches. The analysis outlines a flexible and versatile system encoding semantic distinctions of animacy, shape, and part–whole relations. Classifiers have several central functions in the grammar of Baniwa, including derivation, inflection, and referent tracking in discourse. The system has developed over the course of several millennia, and continues to develop in the face of ongoing cultural change and language contact. It shares many of its properties with classifier systems in other Arawakan languages, as well as with classifier systems in unrelated languages in the area.

This account of Baniwa classifiers contributes to the understanding of nominal classification systems more widely, in particular those of Arawakan and Northwestern Amazonian languages, and illuminates the structure, development and maintenance of such systems. The analysis sheds light on a number of commonly posited dichotomies in linguistic theory, such as the distinction between lexical and grammatical forms, and between inflectional and derivational processes. It is also a contribution to our knowledge of lesser-known languages.

Printed by Media-Tryck, Lund 2025 NORDIC SWAN ECOLABEL 3041 0903

Joint Faculties of Humanities and Theology
Centre for Languages and Literature

ISBN 978-91-89874-69-5