# The Shape of Arawakan Classifiers: An Internal Source for the Arawakan Classifier System(s)

Dunn, Saskia

# THE SHAPE OF ARAWAKAN CLASSIFIERS: AN INTERNAL SOURCE FOR THE ARAWAKAN CLASSIFIER SYSTEM(S)

Saskia Dunn (s2924994)

Leiden

08-07-2022

MA THESIS


The Shape of Arawakan Classifiers: An internal source for the Arawakan classifier system(s)


Saskia Dunn (s2924994)


Submitted for the partial satisfaction of the requirements for the degree of
MASTER OF ARTS
at
LEIDEN UNIVERSITY
in the programme
LINGUISTICS (RESEARCH)


Written under the supervision of:
Dr. Martine Bruil

Second reader:
Dr. Konrad Rybka

## Abstract

The Arawakan languages are renowned for their classifier systems, which allow classifiers in multiple positions in the clause and are underpinned by agreement. It is known that they expanded greatly within the subfamilies (Aikhenvald 2019), however what existed at the level of Proto-Arawakan is still unknown. In this thesis, I propose a Proto-Arawakan classifier system which allowed classifiers in the numeral, nominal, and verbal positions, and that at least twelve of the classifiers found today have roots in this system. This argument is supported by evidence in the morphosyntactic structures of the synchronic Arawakan classifier systems, cognates found between the classifiers relating to physical properties, and the semantics underlying the synchronic systems.

I expand on the proposal above by arguing that the Proto-Arawakan classifier system developed first in the numeral position and expanded to the nominal and verbal positions through the use of repeater classifiers. I argue that the set of forms which can be reconstructed to this system show signs of semantic bleaching and phonological erosion within this system, but that the systems only become fully grammaticalised within the subfamilies, where they expanded greatly.

# Contents

## List of Figures and Tables

**Figures**

**Tables**

## Acknowledgements

# Abbreviations and symbols

| | | | | |
|---|---|---|---|---|
| - | affix boundary | N | nasal |
| = | clitic boundary | NC1, NC2... | noun class 1, noun class 2 |
| 1,2,3 | first, second, third person | NEUT | neuter |
| 7,' | glottal stop | NF, NON.FEM | non-feminine |
| A | subject of transitive verb | NM | non-masculine |
| AB | absential | NOM | nominaliser |
| ADJ | adjective | NUM.CL | numeral classifier |
| AG | agreement marker | OBJ | object |
| ART | article | PL | plural |
| AUX | auxiliary | POS | possessor |
| BDR | bivalent direct | POSS | possessive marker |
| BR | bound root | PROG | progressive |
| CL, CLF | classifier | PST | past |
| CL.T, TCL | class term | REAL | realis |
| COMPL | complement | REDUP | reduplication |
| CON | connector | REP | repetitive |
| COP | copula | RP | repeater |
| DEF | definite article | S | subject of intransitive verb |
| DEM | demonstrative | SG | singular |
| DIM | diminutive | TH | thematic suffix |
| DIST | distal | TOPIC.ADVANCING | topic advancing voice |
| DL | dual | V.CL | verb classifier |
| DUB | dubitative | VERT | vertical |
| DUR | durative | VLBZ | verbaliser |
| F | feminine | | |
| hPL | human plural | | |
| IFV | imperfective | PA | Proto-Arawakan |
| IMP | impersonal | NA | North Arawakan |
| INAN | inanimate | NW, NWA | North West Amazon |
| INTER | interrogative | SSWA | South and South Western Arawakan |
| IRR | irrealis | | |
| LK | linker | SW, SWA | South West Amazon |
| M | masculine | | |

## Presentation of examples

This thesis uses five- and four-line examples. The first line gives the language and typological information (family, country or linguistic area as appropriate). The second line gives the text in the language's orthography in italics, if one is provided. The third line gives the text divided by morphemes. The fourth line gives a morpheme-by-morpheme gloss of the third line. The fifth line gives a free translation in English. The source is given following the example.

(1) BAURE (ARAWAKAN: GUAPORÉ-MAMORÉ)

*iyowkopan te kokononev*

iyowko-**po**-a-no te kokono-nev

various-CLF:tiny-LK-NOM1 DEM1.M caterpillar-PL

'various caterpillars'

(Danielsen 2007, 139)

Throughout, the key element (usually the classifier) is highlighted in bold.

## 1. Introduction

Nominal classification systems, systems by which languages categorise and organise nouns, and in particular systems referred to as 'classifier systems', have seen a great deal of discussion over the last fifty years (see for instance Allan 1977; Dixon 1986; Aikhenvald 2000; Kilarski 2013). In particular, the classifier systems of the North West Amazon (NWA) initially proved difficult to place within existing typologies, falling somewhere between gender and noun class systems and other nominal classification systems, as defined by Dixon (1986; see also Doris Payne 1987; Grinevald and Seifart 2004; Seifart and Payne 2007). [1] Although this question is not fully resolved, as new typologies have developed and understandings of nominal classification systems have expanded, the systems of the NWA have remained of interest, not least because they show particularly complex and parallel systems, often showing both gender like systems and classifier like systems, but also because the intensity of contact in the region has led to similarity among the systems of unrelated languages and obscured patterns of familial inheritance (see Aikhenvald 2012, 279–303 for an overview, and Epps and Michael 2017 for discussion of linguistic areas).

The Arawakan languages form a widespread language family, spoken far beyond the boundaries of the Amazon, in the east as far as the Andean foothills, in the south as far as the Chaco region, and in the north beyond even South America, believed to have been spoken in Central America prior to the European conquest (Aikhenvald 2012, 33). The Arawakan languages exhibit a gender system which can be reconstructed to Proto-Arawakan (PA) (Aikhenvald 1994a; 2019), but many also make use of classifier systems which show similarities in their function and morphosyntactic locus, but differ in their size, ranging from small systems of less than ten forms to expansive systems of over one hundred. Due to these differences, it has been argued that the classifier systems are a product of the subfamilies (Aikhenvald 1994a; 2019), and result from contact with similar systems in neighbouring languages. However, the presence of classifier systems across the family raises the question of how much similarity could be found between them, and whether there is evidence of any system at the level of PA which allowed the expansion of such systems later on. This question, which forms the primary motivation for this paper, is given below, along with the research objectives:

(1) Can the classifier systems seen in synchronic Arawakan languages be shown to share the same diachronic source?

      a. To establish to what degree are the Arawakan systems of the NWA unique to this area?

      b. To establish to what degree can cognates be identified in the synchronic systems which point to an early development of a classifier system?

      c. To establish to what degree can semantic evidence be identified in the synchronic systems which points to an early development of a classifier system?

      d. To establish to what degree do the morphosyntactic properties of the synchronic systems point to an early development of a classifier system?

This paper is structured as follows. **2** puts forward the necessary background on nominal classification systems and defines the terminology which will be used throughout this paper. **3** introduces the linguistics of South America and the Amazonian basin, with specific reference to classifiers, and further introduces the Arawakan family and some relevant linguistic features. The methodology is discussed in **4**. **5** gives an overview of the Arawakan classifier systems and their morphosyntactic properties, reviewing their grammatical functions and structure with comparison to each other and other nearby languages, and discusses the underlying semantics of the synchronic systems. **6** presents some possible cognate sets. **7** is a discussion of

---

[1] Both David and Doris Payne are cited in this paper. Based on frequency, David Payne will henceforth be referenced as Payne; Doris Payne will be referenced by first and last name.

the evidence presented in **5** and **6**, and proposes a developmental path which connects the morphosyntactic evidence with the evidence of cognates. Finally, a conclusion is given in **8**.

## 2. Theoretical background

This chapter will present the terminology which will be used throughout (2.1) and introduce the relevant theoretical background necessary to address the questions at the heart of this paper. The interaction between classifier locus, semantics, and function and how these characteristics can be used in identifying and defining different subtypes of classifiers and the cooccurrence of systems will be discussed in 2.2. Finally, 2.3 will present the key notions of grammaticalisation, and some developmental paths possible for classifier systems.

### 2.1. Terminology

Before moving on to the literature on nominal categorisation systems, it is necessary to define some terminology. The term **nominal categorisation system** (Senft 2000) is used to encompass any linguistic construction which can be considered to categorise a noun or noun phrase within a language: systems of **gender**, **noun class**, and **classifier** will be of importance for this paper.

**Gender** and **noun class** systems are often used interchangeably depending on the linguistic area in question: they are both closed grammatical systems in which a noun is assigned to (usually) one of between two and 20 classes and take part in grammatical agreement (Dixon 1986; Corbett 1991). Where a distinction may be made between them, gender systems are associated with smaller systems of 2-4 classes, while noun classes are associated with systems at the larger end of the spectrum; in this paper, noun class will be used to refer to both, with more specific reference made as appropriate. For both, grammatical agreement is a defining feature, and although a definition will be explored later in this section, it is enough to understand agreement as a formal realisation of a feature of a noun or verb on another element of the same phrase or clause. Both gender and noun class markers are considered to be largely semantically empty (Allan 1977). An example of a gender system is given in (2), and an example of a noun class system is given in (3).

(2) Gender: RUSSIAN (INDO-EUROPEAN)

    a.  Nov-**yj**   žurnal
        New-M   magazine
        'New magazine'

    b.  Nov-**aja** kniga
        New-F    book
        'New book'

    c.  Nov-**oe**        pis'mo
        New-NEUT     letter
        'New letter'

                               (Corbett 1991, 106)

(3) Noun class: SWAHILI (BANTU)

| **Ki**-kapu | **ki**-kubwa | **ki**-moja | **ki**-lianguka |
|---|---|---|---|
| NC7-basket | NC7-large | NC7-one | NC7-fell |

    'One large basket fell.'

                               (Welmers 1973, 171)

(2) demonstrates the three genders of the Russian gender system, in which gender is controlled by the head noun but marked overtly only on the modifier. (3) demonstrates a noun class system typical of a Bantu language: the head noun, 'basket', is from noun class 7 (of 18), and agreement is marked on two modifiers and incorporated into the verb as subject agreement.

**Classifiers** can be understood in opposition to noun class and gender systems, in that they prototypically consist of a large set of forms which are only marked once in a phrase and nouns may occur with a variety of

classifiers (Dixon 1986). The assignment of nouns to classifiers is highly semantically motivated, and the line between classifier and noun may be blurred. An example of classifiers in use is given in (4), where the noun classifiers *naj* 'man' and *no7* 'animal' are preposed to the nouns they occur with, and serve the function of identifying a core feature of the lexical noun.

(4)  Classifier: JAKALTEK (MAYAN: GUATEMALA).[2]

      xil     **naj**     xuwan   **no7**        lab'a

      saw    CL:man John   CL:animal    snake

      "(man) John saw the (animal) snake."

<div align="right">(Craig, 1986, as cited in Grinevald 2002, 65)</div>

The opposition between noun class and classifier systems described above was first explored by Dixon (1986), and can be summarised as in the table below, taken from Grinevald (2000).

<div align="center">TABLE 1</div>
<div align="center">NOUN CLASS VS GENDER SYSTEMS (GRINEVALD 2000, 62)</div>

| Noun class systems | Classifier systems |
|---|---|
| classify *all* nouns | do not classify all nouns |
| into a smallish number of classes | into a largish number |
| of a *closed* system | of an open system |
| fused with other grammatical categories (Def, Nb, Case) | independent constituent |
| can be marked on noun | not affixed to noun |
| realised in agreement patterns | marked once |
| N uniquely assigned to a class with no speaker variation | N possibly assigned to various classes at speaker's will |
| no variation in register | formal/informal uses |

Although this definition is not irrelevant, as more systems were described, examples were found which contradicted these tendencies. Examples are often given from the Amazonian region, where classifier systems range from less than ten forms to over a hundred, are very often bound forms, and the languages tend to be agglutinating rather than isolating. This is exemplified by systems such as the Miraña (Witotoan: NWA) system, which has a system of 72 class markers, 66 of which look more like classifiers, with shape-based semantics, and six of which look more like noun class markers, encoding animacy, natural gender, and number (Seifart 2005). Miraña is a language in the Colombian NWA; the characteristics that make the categorisation of this system difficult are found throughout the region.

Although later typologies have attempted to incorporate the evidence of recently described systems while maintaining a theoretical distinction between classifiers and noun class systems, Corbett and Fedden (2016; 2017) argue that a true theoretical distinction cannot be maintained.

Instead, Corbett and Fedden (2016) define an imagined ideal, based on a prototypical gender system, which they refer to as Canonical Gender. Canonical Gender follows the Canonical Gender Principle, given in (i), below, and three criteria, also given below (ii).

(i)  *Canonical Gender Principle* (*Corbett & Fedden, 2016*)

      In a canonical gender system, each noun has a single gender value.

---

[2] The example given here is of a noun classifier: Dixon's (1986) definition included noun classifiers, numeral classifiers, and verb classifiers ('classificatory verbs'), the distinction of which will be returned to in **2.2.1.**

(ii) *Canonical Gender – Criteria 1-3 (Corbett & Fedden, 2016)*
  a) Canonical gender values match agreement classes.
  b) In a canonical gender system the gender of a noun is constant across all domains in which a given language shows agreement.
  c) In a canonical gender assignment system, the gender of a noun can be read unambiguously off its lexical entry.

Both classifiers and noun classes may vary from this imagined ideal in realisation while still be categorised as nominal classification systems in theoretical approaches. By consequence, classifiers are not an opposing type to noun class systems, as presented by Dixon (1986), but a useful label for a variety of systems which diverge from the theoretical ideal of Canonical Gender (Corbett and Fedden 2017). This is the view which will be maintained in this paper.

This paper addresses languages which feature systems traditionally referred to as classifier systems alongside traditional gender systems. Although I will not attempt to reclassify any of these systems, something which may be left to later discussion, it is useful to bear in mind that where systems coexist, they are not different types of system, but coexisting systems of the same type. Whether or not it is indeed possible (or useful) to think of multiple systems coexisting in one language will be addressed in **2.2**.

### 2.2. Theoretical Approaches to Classifiers

Having discussed where classifiers fit in relation to other nominal classification systems, the following section will discuss different subtypes of classifiers, and how these can be identified and defined. **2.2.1** will discuss defining classifiers by morphosyntactic locus; **2.2.2** will discuss the relationship between morphosyntactic position and semantics; **2.2.3** will discuss the relationship between classifiers and function. The aim of these discussions is not to conclude the debates, but to provide the context necessary to understanding the systems discussed in this paper, and to justify the methodological approaches taken in later chapters.

#### 2.2.1. *Classifiers by locus*

Dixon (1986) focused his discussion on noun classifiers, with which he includes numeral classifiers, distinguishing these from verb classifiers (which he calls classificatory verbs). Grinevald (2002) distinguishes between genitive, numeral, noun, and verbal classifiers, but allows possible reference to demonstrative and article classifiers. Aikhenvald (2000) includes the same four as Grinevald (2002), but widens her scope to include locative and demonstrative, or deictic, classifiers as well as possessive classifiers which are distinct from the genitive type. These authors define classifiers by their morphosyntactic locus, referring to the element in the clause the classifier occurs with, by which each is considered a distinct type. That these are distinct types is supported by evidence that classifiers in different positions have different functions and semantic profiles, and that multiple types occur within the same language (Grinevald 2002). This has been used to argue for the existence of **multiple classifier systems**, languages which have one set of forms in multiple positions (Aikhenvald 2000). In this section I will argue that it is possible for classifiers of one system to occur with multiple loci, and that the existence of more than one system must be motivated by further evidence.

An example of a multiple classifier system is given below, from Tariana (Arawakan: NWA), which shows a demonstrative classifier, a nominaliser, a possessive classifier, a verbal classifier, and a noun class agreement marker all occurring with the same form *dapana* 'CL:house'.

(5)  TARIANA (ARAWAKAN: NWA)

|     |     |     |     |
|-----|-----|-----|-----|
| ha-**dapana** | pa-**dapana** | na-tape-**dapana** | na-ya-**dapana** |
| DEM:INAN-CL:house | one-CL:house | 3PL-medicine-CL:house | 3PL-POSS-CL:house |

hanu-**dapana**  heku
big-CL:house  wood

na-ni-ni-**dapana**-mahka
3PL-make-TOP.ADV-CL:HOUSE-RECENT.PAST.NON.VISUAL

'This one big hospital of theirs has been made of wood.'

(Aikhenvald 2000, 204)

Aikhenvald (2000) argues that maintaining a distinction between these forms allows a distinction to be drawn between noun class systems and classifiers, however it has already been argued in this paper that a strong distinction between these systems cannot be maintained. Further to this, Seifart argues that taking a locus-based approach obscures agreement patterns and jeopardizes the recognition of a coherent system (2005, 322). Corbett and Fedden (2017) have argued that the same criteria would not be applied to other grammatical features, such as the Indo-European gender systems. In the example below, from French, gender is considered a singular system, despite occurring not only on different loci, but in different forms on each locus. In (6a), masculine is marked with -ə on the definite article, but with no marking on the noun or adjective. In (6b), feminine is marked with -a on the definite article, no marking on the noun, and -t on the adjective.

(6)  FRENCH (INDO-EUROPEAN)

    a.  *Le garçon est content.*

|     |     |     |     |
|-----|-----|-----|-----|
| l-**ə** | gaʁsõ | ɛ | kõtã |
| DEF-M.SG | boy(M)[SG] | is | happy[M.SG] |

       'The boy is happy.'

    b.  *La femme est contente.*

|     |     |     |     |
|-----|-----|-----|-----|
| l-**a** | fam | ɛ | kõtã-**t** |
| DEF-F.SG | woman(F)[SG] | is | happy-F.SG |

       'The woman is happy.'

(Corbett and Fedden 2017, 25)

Further to this I argue that identifying classifiers by their locus obscures patterns of development of classifier systems, such as the spread of classifiers from one locus to another, which is argued for many languages: in Ngan'gityemerri (Australia: Southern Daly), classifiers spread from the nominal to modifier position (Sands 1995), and in Yagua (Peba-Yaguan: NWA), classifiers are argued to have spread from a nominal compound position to a numeral position (Doris Payne 2007).

Corbett and Fedden (2017) argue that multiple systems must be identified on the basis of distinctions in both semantics and form. This is seen in the two classifier systems of Akatek (Mayan: Guatemala, Mexico).[3] As can be seen in Table 2 below, the sets show a distinction both in phonological form and in locus, as there is no overlapping set of forms which can occur in both positions. They also differ in semantics, as the nominal set distinguishes the generic category of the referent, while the numeral set refers to the shape of the referent. That these are distinct systems can be evidenced three ways: in phonological form, in semantics, and in morphosyntactic locus.

---

[3] Akatek also has a paradigm of classificatory suffixes which show behaviour both of canonical classifiers and canonical noun classes, and two plural forms, which are not discussed here (Zavala 2000).

TABLE 2

NUMERAL AND NOUN CLASSIFIERS IN AKATEK (ZAVALA 2000)

| sortal numeral classifiers (numeral locus) | | noun classifiers (nominal locus) | |
|---|---|---|---|
| *wa'an*[4] | 'erect' | *naj* | 'man' |
| *k'itan* | 'separated' | *'ix* | 'woman' |
| *kupan* | 'bent, half-a-circle-shape' | *k'o* | 'honorific' |
| *xoyan* | 'circle-shaped and coiled-up' | *yab'* | 'familiar' |
| *patxan* | 'wide and flat' | *no'* | 'animal' |
| *jenan* | 'two-dimensional, extended' | *te'* | 'tree' |
| *k'olan* | 'spherical' | *ch'en* | 'rock' |
| *b'ilan* | 'spherical, small' | *('i)xim* | 'corn' |
| *pilan* | 'big spherical or oval' | *tx'an* | 'thread' |
| *xilan* | 'two- or three-dimensional round' | *tx'otx'* | 'soil/dirt' |
| *jilan* | 'three-dimensional, with longitudinal shape' | *a'* | 'water' |
| | | *ka'* | 'fire' |
| | | *tx'am* | 'salt' |
| | | *an* | 'vegetable' |

Recognising classifier systems which operate over multiple loci and those which exist alongside other, separate systems allows for a better understanding of how those systems operate as an entire system, and for a better identification of the developmental paths of those systems.

In this section, it has been shown how classifiers have been defined by their locus in previous literature, although it has also been argued that locus alone is not enough to motivate distinct systems, which will inform the rest of this discussion. The cross-linguistic overlap of semantics and locus will be considered in the next section.

### 2.2.2. *Semantics of classifiers*

In this section, it will be shown that the position of the classifier is relevant in discussions of semantics, and that this can have relevance in identifying the developmental path of a system.

That there is semantic organisation to classifiers was first discussed by Denny (1976), who identified three categories of organisation with specific reference to noun classifiers, and Allan (1977), who identified seven. Denny's (1976) categories were based on human interactions: physical interaction, functional interaction, and social interaction. Physical interaction referred to classifiers which describe the shape and strength of objects; functional interaction refers to classifiers which describe how an object is used; social interaction refers to any classifier which denotes a place in human hierarchies, such as animateness or social position (Denny 1976). Allan's (1977) seven categories were: material, shape, consistency, size, location, arrangement, and quanta. That these categories align with the position of the classifier in the sentence has also long been recognised. Adams and Conklin (1973) discussed not only that there is a strong tendency towards shape based semantics in numeral classifier systems, but that these systems can be broadly categorised into three basic shapes: long, flat, and round. These shapes are often found in combination with a characteristic relating to consistency or size, and most commonly have lexical origins in plant terms: long-rigid (tree or trunk), flat-flexible (leaf), and round (fruit) (Adams and Conklin 1973). Although much of the discussion of semantics in classifier systems has remained in the relationship between shape semantics and numeral classifiers, there are also tendencies between other semantic groups and other syntactic positions. Olness (1991, cited in Grinevald 2002) shows that functional classifiers correlate with the genitive position and material (or 'essence') classifiers correlate with the nominal position, as given in the table below.

---

[4] The suffix *-an* which occurs consistently through this set is unique to classifiers, although it is unclear what its meaning is; for more details on this see Zavala (1992, 130–34).

TABLE 3

CORRELATION OF FUNCTION AND POSITION (GRINEVALD 2002, 263)

| Morphosyntactic locus | Semantic profile | Example |
|---|---|---|
| numeral | physical | 'two-ROUND oranges' |
| genitive | functional | 'my-EDIBLE food' |
| noun | material/essence | 'an ANIMAL deer' |

These correlations are by no means absolute, but they are supported by a convincing cross-linguistic tendency: 63% of numeral classifiers have physical semantics, 78% of noun classifiers have material semantics, and 86% of genitive classifiers have functional semantics (Olness 1991, cited in Grinevald 2002). Craig (1992) argues that this can reflect the real world grammaticalisation environment: numeral classifiers are used in quantification, so are likely to develop in sales-based situations involving handling of objects; genitive classifiers develop from alienable constructions, involving nouns which are highly valued and sought for a particular function. Although these have yet to be evidenced in language, they reflect the correlations described.

These alignments of semantics and locus can also have consequences for understanding the development of classifier systems, in highlighting the grammaticalisation pathway in systems in which the classifiers have spread from one position to another. This will be of relevance in discussing the systems in this paper, which show classifiers with multiple loci, but which share underlying semantics. The following section will discuss the functions of classifiers, and whether or not these can also be aligned by locus.

### 2.2.3. Functions of classifiers

Classifiers are recognised for both their pragmatic and grammatical functions, which will be discussed here. This section will consider whether there can also be an alignment between locus and function, and will argue that although there is not a one to one match of function to locus, the functionality of classifiers is a motivating force behind their development. Greenberg (1972) and Croft (1994) recognise the main purpose of classification as individuation, particularly with reference to classifiers in numeral constructions, but also the individuation of referents in discourse. Bisang (2002) recognises a two-step process of individuation and identification, in which identification is a prerequisite for individuation. These align with the development of classifiers in different positions as laid out below, in Figure 1.

FIGURE 1

CLASSIFIERS AND THEIR FUNCTIONS (Bisang 2002, 303)[5]

| classification | → | identification | → | individuation |
|---|---|---|---|---|
| | | *discourse*      *deixis* | | |
| relational classifier | possessed classifier (possessor classifier) | noun classifier verbal classifier numeral classifier | deictive classifier locative classifier | numeral classifier |
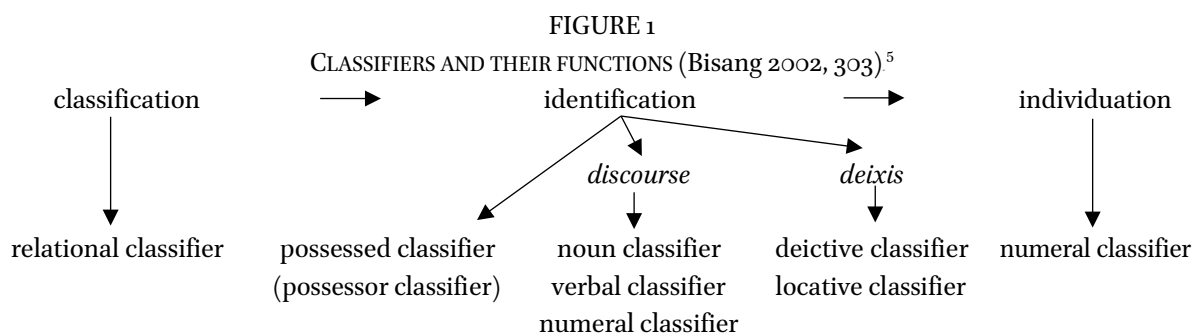
An entity must be identified before it can be used in discourse, thus identification allows for reference and is seen in the marking of definiteness and specificity, but also in the marking of relations in possessive or relative clauses (Bisang 2002). Identification can be further understood as identification in discourse or identification in a location, allowing for locative and deictic reference. A step beyond identification is individuation, which has long been recognised as the function of classifiers in languages in which nouns are inherently uncountable.

---

[5] For explanation of classifier types not discussed here, see Bisang (2002).

However, Contini-Morava and Kilarski (2013) discuss the functionality of classifiers from the perspective of semantic and discourse functions. Semantic functions include expansion of the lexicon (through derivation), differentiating referents, individuation, and ascribing properties to references. Discourse functions include referent identification, reference management, and re-presentation of referents. These functions are not necessarily tied to morphosyntactic locus, although again the use of individuation by numeral classifiers is noted.

Both of these approaches consider the functions of classifiers to be for manipulation and organisation of both discourse and the lexicon. Of relevance here are derivation, a semantic function, and anaphoric reference and agreement, both discourse functions. As this paper is not concerned with the analysis of these patterns but in the description of their realisations across languages, broad definitions are taken here. Derivation is understood in terms of Contini-Morava and Kilarski's (2013, 269–72) expansion of the lexicon, in reclassifying a given root by a new property, and anaphoric reference is understood as any use of a classifier in place of an overt noun. However, the definition of agreement requires a closer treatment.

At a broad level, agreement is "some systematic covariance between a semantic or formal property of one element and a formal property of another" (Steele 1978, cited in Corbett 1991). This is usually understood as a relationship between a controller, which in a nominal phrase would be the head noun, and a target, some kind of modifier. This definition allows for a variety of interpretations: although there must be some formal realisation outside the controller of agreement, on the target, there is no requirement for obligatory marking on the controller itself. The earlier Russian example (2a, repeated here as 7) demonstrates agreement controlled by a semantic element of the head noun, with no overt marking. The earlier Swahili example (3, repeated here as 8) demonstrates agreement controlled by a formal element of the head noun, with overt marking on both the noun and its modifiers.

(7)  RUSSIAN (INDO-EUROPEAN)

    Nov-**yj**   žurnal
    New-M   magazine
    'New magazine'

                                                               (Corbett 1991, 106)

(8)  SWAHILI (BANTU)

| **Ki**-kapu | **ki**-kubwa | **ki**-moja | **ki**-lianguka |
|---|---|---|---|
| NC7-basket | NC7-large | NC7-one | NC7-fell |

    'One large basket fell.'

                                                               (Welmers 1973, 171)

However, other definitions are more narrow; Siewierska (1999) defines grammatical agreement as the obligatory marking both on the target of agreement and on the controller. Corbett (2006) also sees obligatoriness and overt marking on both the target and controller (among other things) as necessary for Canonical Agreement,[6] although he allows that other realisations of agreement vary from this ideal. For this paper, I take the broadest possible position and return to the definition already given by Steele (1978, cited in Corbett 1991): agreement is the formal realisation of one element determined by a semantic or formal property of another.

This section has discussed the functions of classifiers in relation to discourse and the lexicon, as well as their possible alignment with classifier locus. Although there is not a one-to-one match of classifier position to function as in the case of semantics, it is clear that the motivating force behind classifiers and their use

---

[6] Canonical Agreement is a representation of agreement as an imagined ideal, following the same concept as Canonical Gender. See Corbett (2006) for more.

differs in different positions. It remains to be seen if and how this impacts the semantic patterns which develop in those positions, however both of these patterns may be of use in identifying how and where classifiers developed in a language, particularly in languages in which classifiers can occur with multiple loci in the sentence, as is the case in the languages discussed in this paper. The development of classifiers will be the focus of the next section.

### 2.3. Diachronic perspectives on classifiers

If grammar is "a system of more or less stable, regular, and productive form-meaning mappings" (Wiemer and Bisang 2004, 4), grammaticalisation is the instability by which it is produced. More specifically, grammaticalisation is a shift from a lexical function to a grammatical one, involving phonological, morphological, semantic, pragmatic, and syntactic changes (Passer 2016). Kouteva et al. (2019) define this as characterised by four processes: use in new contexts (extension), loss of meaning (desemanticisation), loss of characteristic lexical morphosyntactic properties (decategorialisation), and loss of phonetic substance (erosion). Semantic extension is the use of a term in a context beyond its literal meaning, as in the widely cited example of *pas* 'step' coming to be used in French negation with the meaning of reinforcing the statement 'not one step'. In the French example, this is followed by desemanticisation, in which the meaning of 'step' is lost in this context. Decategorialisation, or decategorisation, refers to the loss of categorial properties associated with the lexical form, and is often accompanied by phonological erosion, which can refer to any loss or reduction of phonological material.

Grammaticalisation is often presented as cyclical, in which a lexical element reduces to a grammatical element but is ultimately reduced to zero, and the grammatical function is filled again by a new lexical element undergoing grammaticalisation. However, an element does not have to go through the entire cycle; they may become stable at any point in the process, and so it is not always necessary for each of the four processes mentioned above to be found (Passer 2016). This section will discuss the role of grammaticalisation in the development of classifier systems and forms.

Grammaticalisation has been significant in the literature regarding classifiers, as for some time it defined the typology by which nominal classification systems were understood. Craig (1986) identifies classifiers as a midway point between purely lexical systems and purely grammatical systems, as they take part in lexical processes of derivation and compounding while also showing grammatical properties, in agreement. This system as presented by Grinevald (2002) is given in Figure 2, below.

FIGURE 2

NOMINAL CLASSIFICATION SYSTEMS AND GRAMMATICALISATION (GRINEVALD 2002)

| Lexical | | Grammatical |
|---|---|---|
| Class terms | CLASSIFIERS | Noun classes |
| Measure terms | | Gender |
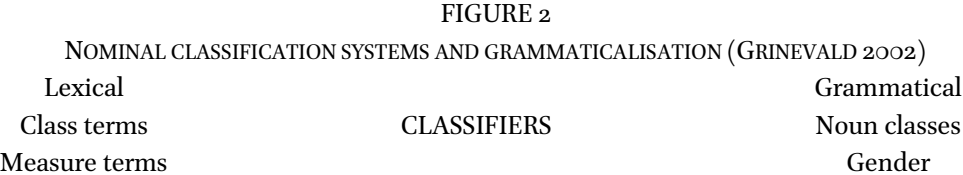
The two lexical types given in this continuum are both seen in English. Measure terms are phrases like *a glass of water*, or *a head of cattle* which allow mass nouns, usually considered uncountable, to be counted. Class terms have perhaps the most similarity with the classifiers which will be seen in this paper: they are lexical terms which allow the derivation of a noun into a category through compounding, as in English *straw**berry**, blue**berry**, rasp**berry***.

At the other end of the spectrum are 'purely' grammatical noun class systems. This typology is supported by evidence that classifiers are one source of noun classes: this is documented for some Australian languages (Sands 1995) and is proposed for some Bantu languages (Grinevald 2002). This typology is also used to categorise different types of classifiers, with classifiers occurring in numeral constructions considered less grammaticalised than classifiers occurring in genitive constructions (Craig 1992).

However, the usefulness of defining classifiers by locus has already been discussed in **2.2.1**, and it is grammaticalisation as a diachronic development path which is of importance here. Classifier forms are known to develop from lexical nouns (Grinevald 2000, 55–62), and can be shown to undergo changes in their morphology, semantics, phonology, and syntactic and pragmatic use. Of particular relevance is decategorialisation, as emerging classifiers lose properties associated with nouns, and loss, or changes, in meaning. As was discussed previously, a common shift is from nouns relating to plants to classifiers relating to shape and dimension. The classifiers considered in this study are bound forms, demonstrating phonological reduction of the lexical noun to become dependent on a new host. However, how classifiers undergo these changes (and the level to which they do) differs between languages and systems, and the grammaticalisation of a system differs from the grammaticalisation of a form. Below, some grammaticalisation pathways seen in classifier systems are discussed.

One source of classifier systems is in the grammaticalisation of a type of classifier referred to as 'repeaters' (Aikhenvald 2000). The term may be used with a variety of definitions: narrowly it refers to a classifier which is identical (or partly identical) in form and meaning to the noun it classifies (9), but it is also used to refer more broadly to a free noun used in a classifier position (10). In both uses of the term, repeaters are expected to show fewer signs of grammaticalisation than classifiers, as they are not expected to show signs of semantic bleaching or phonological erosion. However, it is in the grammaticalisation of these forms that classifier systems are argued to arise. Note that in (10), (10a) shows the use of the free noun *báhku* 'bone' in the position of a classifier, incorporated to a predicate, while (10b) shows the same noun in a position not available to classifiers, with a possessor prefix.

(9)  JAPANESE (JAPONIC)
   a.  **bin**        hito-**bin**
       bottle     one-bottle
       'a bottle'
   b.  **kyoku**           ik-**kyoku**
       composition      one-composition
       'a piece of music'

(Beckwith 2007, 100)

(10) MIRAŇA (WITOTOAN: NWA)
   a.  kátúɪ:βɛ-**báhku**
       fall-RP.bone
       'The bone fell.'
   b.  tᵼáj-**báhku**
       POS.1SG-bone
       'my bone'

(Seifart 2005, 81)

Repeaters are common in the classifier systems of South East Asia, and are proposed as the origin of the Sino-Tibetan classifier systems, including the Chinese system (Aikhenvald 2000, 361). This developmental path is also proposed for Kilivila (Austronesian: Papua New Guinea), where it is argued that 25 out of 88 classifiers are nouns which are used as classifiers but do not show any signs of grammaticalisation, that is, repeaters in the broad sense (Senft 1993). Senft (1993) argues that the original classifier system consisted only of repeaters, but that through use the majority underwent phonological reduction. Notably, Passer (2016) states that no system has been found which consists only of repeaters, however it is unclear if he is using the term in a narrow or broad sense.

An additional possible step in the development of classifiers is the development of classifiers via compounding systems, which is described for both Yagua (Peba-Yaguan, NWA: Doris Payne 2007) and the Witotoan languages (Seifart 2007), which are found in the NWA. In the Witotoan languages, classifiers are bound suffixes which can be used with numerals, demonstratives, and nouns, and they are also used in compound constructions. Seifart (2007) proposes that some classifiers may have grammaticalised from nouns in compound constructions, by the broadening of their syntactic use to appear in positions available to classifiers and consequently undergoing grammaticalisation in these positions. The same process is proposed for Yagua, where compounding is a productive process and a source of new classifiers (Doris Payne 2007).

A further step in the development of classifier systems is the spread of classifiers from one position to another. This process may be similar to the development of agreement, in which an element becomes redundant in its position, in a shift such as 'the man, he arrived' to 'the man he-arrived': in this imagined example, a pronoun becomes an agreement marker in the reanalysis of a topic construction (Corbett 2006, 265). This is proposed for Ngan'gityemerri (Southern Daly: Australia), where it is argued that classifiers originated in generic-specific noun pairings and the spread of the generic to other elements of the sentence. The example demonstrates three stages of this. In (11a), the generic and specific noun are followed by an unmarked adjective modifier. In (11b), the generic noun is used anaphorically, without a specific noun: this allows an association between the generic and the adjective. (11c) shows that when an overt specific noun is used, the generic must also be used with the adjective.

(11)  NGAN'GITYEMERRI (SOUTHERN DALY: AUSTRALIA).[7]

    a.  gagu    wamaŋgal    kerre    ŋeben-da
       animal  wallaby      big     1SGS/AUX-shoot
       'I shot a big wallaby.'

    b.  gagu   kerre   ŋeben-da
       animal  big    1SGS/AUX-shoot
       'I show a big wallaby.'

    c.  gagu    wamaŋgal    gagu    kerre    ŋeben-da
       animal  wallaby      animal  big    1SGS/AUX-shoot
       'I shot a big wallaby.'

<div align="right">(Reid 1994, cited in Sands 1995, 253–54)</div>

This section has discussed the relevance of grammaticalisation in the development of classifier systems, discussing both how they develop and how they spread, and demonstrating the importance of both repeaters and compound systems in this process.

---

[7] This construction is still used in the language, and is argued to be the origin of the agreement classes which have developed in parallel to this usage. In the example below, *wa=* 'MALE' has grammaticalised from a generic noun.

i.    NGAN'GITYEMERRI (SOUTHERN DALY: AUSTRALIA)

    wa=ŋurmumba    wa=ŋayi    daranj-fipal-njine
    MALE=youth       MALE=big    3SG/AUX-return-FOCUS
    'My initiand son has just returned.'

<div align="right">(Reid 1994, cited in Sands 1995, 254)</div>

### 2.4. Conclusion

This chapter has aimed to clarify the terminology and theoretical background necessary to understand the rest of this paper, in clarifying what a classifier is and how they can be understood in relation to other nominal classification systems, but also how different types can be seen within the category of classifiers and how these can be defined. It has been argued that although multiple systems can coexist in languages, these should be motivated by a difference both in semantics and phonological form, and that classifier locus alone is not enough to argue for a distinct system. However, it has also been argued that the relevance of locus is supported by a correlation with the semantic basis of the system and possibly with function; this may have consequences for identifying how a classifier system developed in a language, particularly when classifiers occur with multiple loci. Finally, some diachronic pathways of classifier development have been discussed, with examples from a number of languages. Of particular relevance are the use of repeaters in the development of classifier systems and the development of classifiers from compound systems, seen in the Witotoan languages and in Yagua. The following chapter will present the relevant linguistic features of the Arawakan family (3), before turning to the methodology used in this study (4). 5 and 6 will present the data, which will be discussed in 7 and concluded in 8.

# 3. Linguistic Background

Having presented the theoretical ideas which will be of relevance in this paper, this chapter will give background on the area in question: South America and the Amazonian Basin will be discussed initially in **3.1**, and the defined linguistics areas in **3.1.1**, before moving on to some relevant background on classifiers in this region (**3.1.2**). **3.2** will present the Arawakan family, including the literature and classifications (**3.2.1**) and previous research on Arawakan classifiers (**3.2.2**). **3.3** will present the linguistic features of the family, including descriptions of possession, classifiers (**3.3.1**), gender (**3.3.2**), and compounding (**3.3.3**).

## 3.1. South America and the Amazonian Basin

South America is a continent notable for its linguistic diversity: it is home to 108 language families, estimated to be around a quarter of the world's language families (Campbell 2012). It is estimated that in the past there were some 1500 spoken languages on the continent, although today the number is closer to 420, including 55 isolates with no known genetic relative (Campbell 2012). This diversity is particularly concentrated in Brazil, where there are an estimated 150–190 Indigenous languages (Campbell 2012), and in the Amazonian Basin. Despite this high number of language families, the continent is dominated by only a few large language families: Arawakan, Cariban, Chapacuran, Chibchan, Jê, Panoan(-Takanan), Nadahupan, Quechuan, Tukanoan, and Tupian (Campbell 2012; Michael 2021), of which Arawakan, the largest language family on the continent, is discussed here.

The Amazonian Basin, an area of 6,900,000 km², is home to 15 language families encompassing some 350 languages (Aikhenvald 2012). These languages range from those with only a handful of speakers to those with over ten thousand speakers, such as Guaraní, Garifuna, and Guajiro, although such large numbers are a rarity (Aikhenvald 2012). The area has seen a great amount of change since colonisation: pre-Colombian estimates of the population of the Greater Amazonian region are placed between four and five million, while mid-century estimates were around 200,000 Indigenous people, not all of whom spoke an Indigenous language (Aikhenvald 2012). Estimates of the date of the first population of South America have some variation, but are often placed at around 12,000 years ago, with the Andes and the Amazon being populated around the same time 9–11,000 years ago (Aikhenvald 2012). Evidence suggests that large scale societies dominated the river banks, while within the rainforest smaller communities existed on subsistence lifestyles (Aikhenvald 2012). The area is also home to a great deal of contact, which is not surprising given the large number of languages present. Trade routes connected the Amazon and the Andes, evidencing the possible range of contact between languages (Aikhenvald 2012). Many linguistic areas of shared features have been proposed within the Amazon, which will be discussed below.

### 3.1.1. Linguistic Areas in the Amazonian Basin

South America is home to some proposed linguistic areas, including the Vaupés region, the Caquetá-Putumayo region, both of which are in the NWA, and the Guaporé-Mamoré region in the South West Amazon (SWA), as well as the Upper Xingu region (Epps and Michael 2017). These regions have high levels of multilingualism, characterised by low levels of lexical borrowing but high levels of calquing and grammatical convergence (Epps and Michael 2017). The geographical zones of most relevance to this paper are the NWA (including both the Vaupés and the Caquetá-Putumayo regions) and the Guaporé-Mamoré region in the SWA. For the NWA, I have followed the boundaries of Eriksen (2011), who defines the region as the area between the Andes mountain range and the Río Negro and Orinoco River; the southern border is defined by the Putumayo River (Figure 3). The Guaporé-Mamoré region is a small region of two river basins, positioned south of the Madre de Dios and Madeira rivers and just east of the Andes (Figure 4: Crevels and Van der Voort 2008).

FIGURE 3
THE NORTH WEST AMAZON (ERIKSEN 2011, 170)
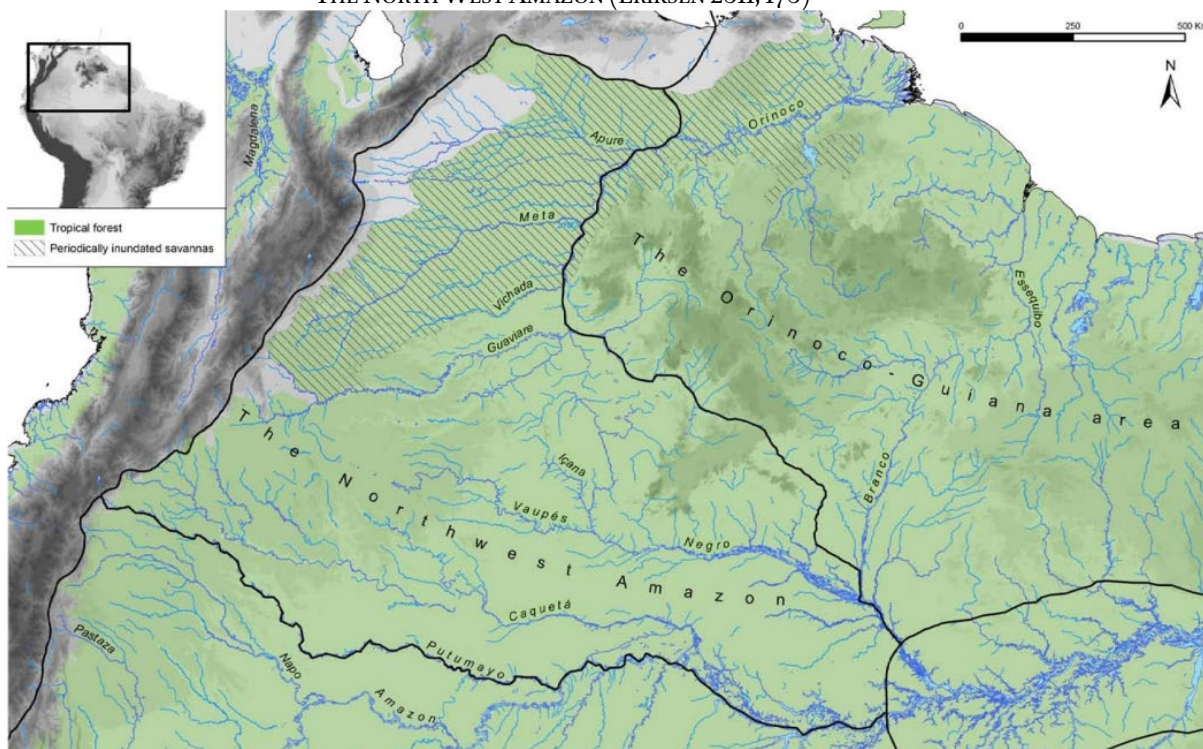


FIGURE 4
THE GUAPORÉ-MAMORÉ REGION (CREVELS AND VAN DER VOORT 2008, 157)
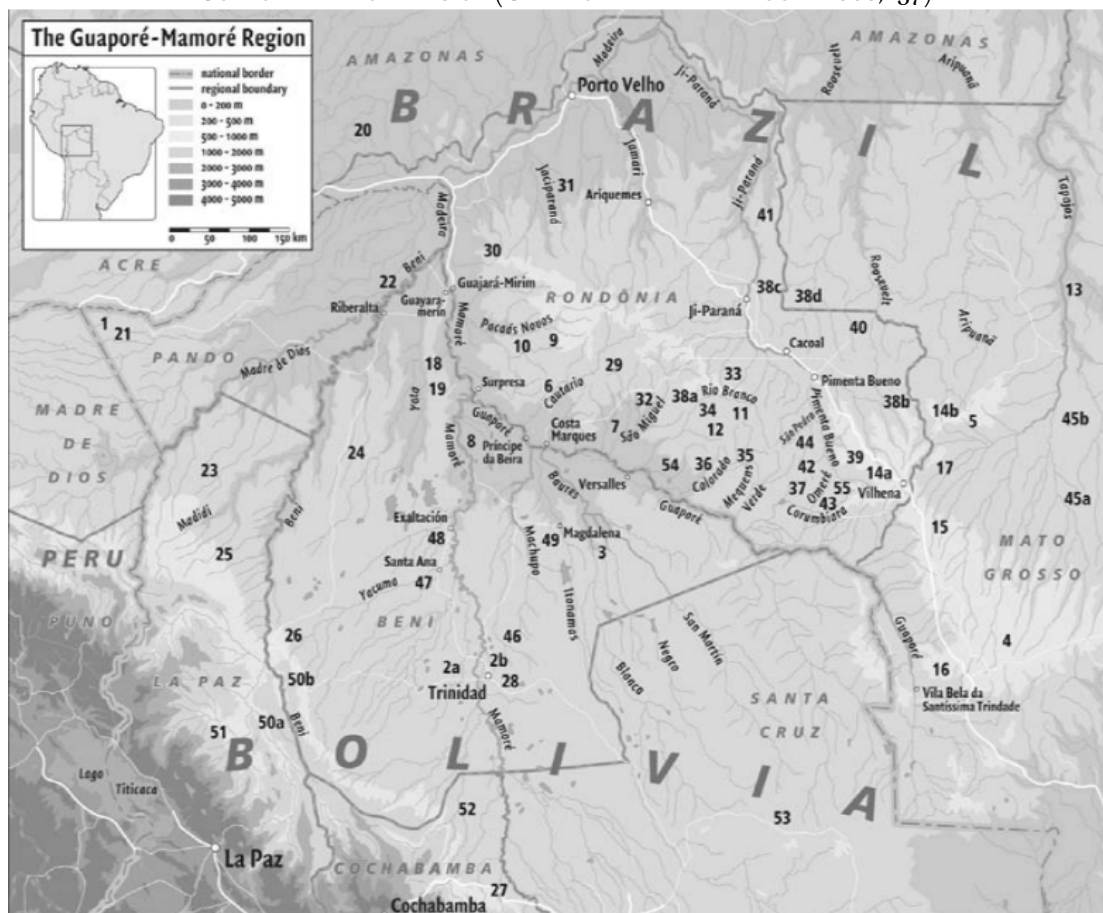
There have been attempts to argue for all of Amazonia as a macro linguistic area, which would include nominal classification systems as a widespread areal feature (see for instance Aikhenvald and Dixon 1999), however more recent work has focused on a west-east divide of the continent, which would divide the Amazonian region into the north and south Amazon in Eastern South America and the western Amazon in Western South America (Krasnoukhova 2012; Birchall 2014). The following section will discuss the realisation and areal patterns of classifiers across the Amazonian Basin.

### 3.1.2. *Classifiers in the Amazononian Basin*

Before moving on to discussing the Arawakan language family, the focus of this paper, it is necessary to give some background on classifiers in the Amazon as a whole, where they are considered an areal feature due to their appearance in many unrelated languages. These classifier systems are typified by their use of classifiers in numerous morphosyntactic loci and their multifunctionality: classifiers are used in derivational processes, in anaphoric reference, and, to varying degrees, in grammatical agreement. As such, Krasnoukhova (2012) refers to these systems as 'multifunctional classifier systems'. It has been argued that these unique features are particularly concentrated in the NWA, at the borders of Colombia, Brazil, and Peru, where languages from the Witotoan, Peba-Yaguan, Arawakan, and Eastern Tucanoan families all show overlapping similarities in their classifier systems (Seifart and Payne 2007).

However, the SWA, particularly the Guaporé and Mamoré river basins, also have many unrelated languages showing classifier pattens similar to those in the NWA. Van der Voort (2018) shows that many languages in this area have classifier systems which use multiple morphosyntactic loci for multiple functions, including anaphoric reference, agreement, and nominalisation, although the extent to which these functions are productive varies between languages. That the SWA has classifiers as an areal feature has also been recognised by Krasnoukhova (2012), who notes that both the NW and SWA are epicentres of multifunctional classification systems, but that the languages between these areas do not have classification systems.

Beyond the Amazonian region, classifiers also occur notably in the languages of the Chaco region, where a genitive classifier is used in constructions possessing domestic animals, which are otherwise unpossessible (Campbell and Grondona 2012). Classifiers are largely absent from the Andean region, although there are exceptions in the Barbacoan languages, in Tsafiki, Cuna (Chibchan), and Mochica (Isolate), and in the Southern Cone (Krasnoukhova 2012; Bellamy 2018).

### 3.2. The Arawakan languages

The remainder of this chapter will present the Arawakan languages: their history and classification will be discussed in **3.2.1** and the background literature on Arawakan classifiers will be given in **3.2.2**.

### 3.2.1. *Classification and history*

FIGURE 5

ARAWAKAN LANGUAGES[8]



The Arawakan[9] language family (Figure 5) consists of 40 languages spoken over eight countries in South America and four in Central America, and even includes a large immigrant community in North America; it is the most widespread Indigenous language family in South America (Aikhenvald 1999; 2012). The family was at one point much larger; the number of languages attested to have belonged to this family is around 80 (Michael 2021). However, a number of other languages have been at times proposed as being related and later discounted; these may still be referred to as Macro-Arawakan languages.[10]

The family was first recognised in 1783 by Father Gilij, an Italian priest, although as already noted, the internal subdivisions are still disputed (Aikhenvald 1999). Major family classifications have been proposed by Payne (1991), Kaufman (1994, cited in Michael 2021), Aikhenvald (1999), and most recently Ramirez (2020). Although Payne's (1991) classification is not widely used, it recognised the distinction between what are referred to now as Arawakan and Macro-Arawakan (Maipuran and Arawakan in his terminology).

---

[8] Map generated with Glottospace (Norder et al. 2022).

[9] There are some disputes as to the naming conventions: Maipure or Maipuran is often used to refer to a core of definitively related languages and contrasted with a wider, tenuous grouping of Arawakan (now Macro-Arawakan) languages (Payne 1991; Aikhenvald 2012), and Aikhenvald (2012) argues against the use of 'Arawakan' in favour of 'Arawak' due to the disputed history of the former term. I use 'Arawakan' here as it is the term in widest usage, following Michael (2021).

[10] Macro-Arawakan languages include Arawan, Guahiban, Timotean, Tiniguan, and Harakmbut languages, as well as some isolates (Michael 2021).

Aikhenvald's (1999) is the most influential classification, and is based on low level areal groupings as well as grammatical similarities (Figure 6). It recognises a major split between South (and South-Western) Arawakan (SSWA) and North Arawakan (NA) with the NA branch consisting primarily of the North Amazonian group; this is consistent with the classification of Kaufman (1994, cited in Michael, 2021). Figure 6 gives the family as presented by Aikhenvald (1999), with languages in italics and language groups in roman.

FIGURE 6
AIKHENVALD'S (1999) CLASSIFICATION OF ARAWAKAN

South and South-Western Arawakan

| | | | |
|---|---|---|---|
| South Arawakan | *Terêna, Bauré, Moxo (Ignaciano), Moxo (Trinitario), Salumã* | | |
| Paresi-Xingu | Xingu | *Waurá, Mehinaku, Yawalapiti* | |
| | Paresi-Saraveca | *Paresi (Haliti)* | |
| South-Western Arawakan | Piro-Apuriná | *Piro, Chontaquiro, Apurina/Ipurina, Cangiti, Mashko-Piro* | |
| Kampa | *Ashaninca, Asheninca, Caquinte, Machiguenga, Nomatsiguenga, Pajonal Kampa* | | |
| Amuesha | *Amuesha* | | |
| Chamicuro | *Chamicuro* | | |

North-Arawakan

| | | | |
|---|---|---|---|
| Rio Branco | *Wapishana* | | |
| | *Mawayana* | | |
| Palikur | *Palikur* | | |
| Caribbean (or Extreme North) | *Garifuna* | | |
| | TA-Arawak subgroup | *Lokono/Arawak, Guajiro/Wayyu, Añun/Parauhano* | |
| North Amazonian | Colombian | *Yucuna, Achagua, Piapoco, Cabiyari* | |
| | Upper Rio Negro | *Baniwa of Icana/Kurripako, Tariana, Guarequena* | |
| | Orinoco | *Bare, Baniwa of Guainia* | |
| | Middle Rio Negro | *Kaiʃana, Bahwana/Chiriana* | |

Attempts to refine this tree have used computational methods, the primary being that of Walker and Ribeiro (2011) which compared lists of 100 common vocabulary items across languages. From this, they argued against the high-level split of the North-South division, but argued for grouping most of the North Eastern languages together. More recently, Ramirez (2020) has presented a classification based on his own research. He recognises twelve subgroups, and also does not include a major split between North and South Arawakan. A major grouping within his classification is that of the Japurá-Colômbia subgroup, which includes many of the NWA languages. This classification is given below, in Figure 7.

FIGURE 7
RAMIREZ'S (2020, 1:43) CLASSIFICATION OF ARAWAKAN
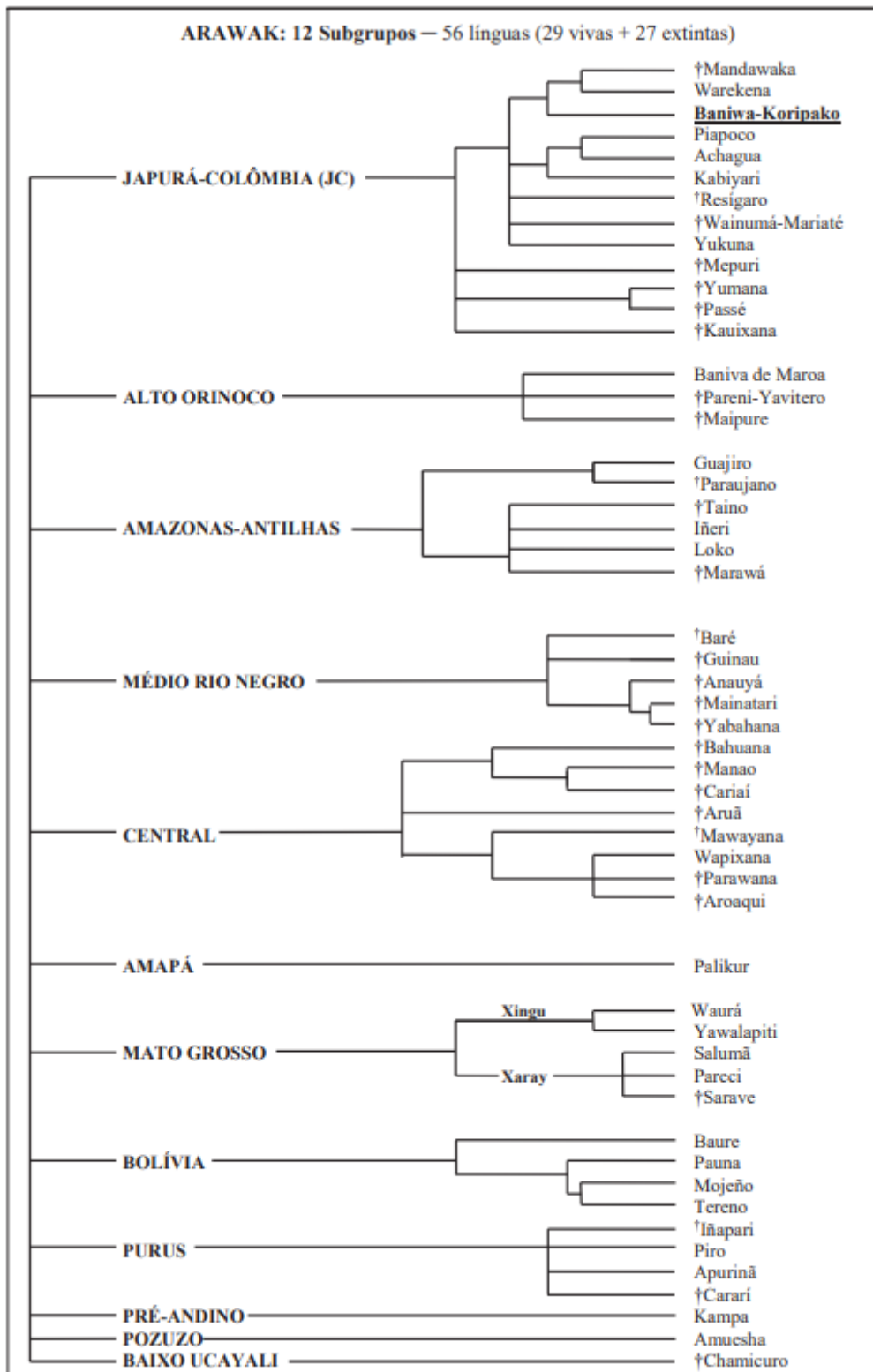
Given the lack of agreement about internal subdivisions, historical reconstructions have focused on low-level groupings (see Michael et al. 2010 on Proto-Kampa; Carvalho 2016b on Xinguan Arawak; 2021 on Proto-Purus; Jolkesky 2016 on Proto-Mamoré-Guaporé; Ramirez 2020 on Japurá-Colômbia). The only major attempt at a reconstruction of PA is that of Payne (1991). His reconstruction is not without criticism, namely that he

bases his evidence on only a small number of languages and considers lexical retentions rather than innovations; he has also proposed phoneme correspondences which are not supported by further evidence.[11] However it remains one of the few large-scale attempts to reconstruct PA.

Despite the widespread area that the Arawakan languages cover today, it is possible to tentatively identify a homeland between the Rio Negro and the Orinoco rivers, based on the high density of Arawakan languages in the area and on origin myths still known among speakers (Heckenberger 2002; Aikhenvald 1999; 2012). However, this is far from confirmed: Walker and Ribeiro (2011) use statistical methods to point to a homeland in western Amazonia. It has also been argued that the Arawakan languages form part of a cultural package which spread through a trade network referred to as the Arawakan matrix (Santos-Granero 2002; Eriksen and Danielsen 2014). This would have the Arawakan languages spread through diffusion, rather than population movement, thus indicating that the term Arawakan refers primarily to a shared linguistic heritage. However, this position points to a diversification point of 600AD, after the break-up of the trade network (Eriksen and Danielsen 2014); Michael (2021) argues instead for a time depth of around 5000 years, and questions whether there is enough evidence for including language in the cultural exchange. These questions remain relevant for reconstructing scenarios of language contact, and although more recent contact situations are known, a conclusion on the spread of early Arawakan has not yet been reached.

### 3.2.2. *Previous Research on Arawakan Classifiers*

As has been referenced, the Arawakan languages feature a range of classifier systems which show a great deal of variety in their forms, grammatical structure, and size. The linguistic properties of classifiers in Arawakan will be discussed in **3.3.1**; this section will introduce previous literature on the systems and proposals as to their origin and the origin of classifier systems in South America.

At this point, little work has been done on whether a classifier system could have existed in PA. Aikhenvald (see 1994a; 2006; 2019 among others) has conducted some studies on Arawakan classifiers of the NWA, and recognises a small number of cognates between systems: Tariana and Baniwa share 31, but all other languages share 5 or less.[12] Aikhenvald (2019, 115–16) argues that classifiers have developed at the subgroup level, and for the systems of the NWA proposes two possible paths of development, given here in (i) and (ii).

(i) Classifiers developed in the numeral position.
(ii) A set of bound inalienable nouns formed a proto set of classifiers, and agreement developed in other positions.

The proposal that classifiers have developed in the numeral position is based on the observation that this position is the most widespread for languages of the NWA. The second proposal, that classifiers originated as inalienable nouns, comes from the similarity in these systems, and the observation that forms with reference to shape or quantity were already bound forms in this position (Aikhenvald 2019, 115–16).

This second hypothesis is similar to that of Van der Voort (2018) with regards languages in the SWA, including Baure (Arawakan) and many unrelated languages, who suggests that classifiers are a result of noun grammaticalisation, most likely out of the bound nouns which appear in the region. In a number of the languages of this region, including the Arawakan languages, there is a closed or semi-closed set of bound nouns which appear in compounding processes, often in possessive constructions. As will be discussed in **3.3.3**, compound systems share many characteristics with the classifier sets in the same languages. However, as with the Arawakan family, the origins of the classifier systems in the Amazon more generally are unclear, and it is likely that contact played an important role. Within the NWA, lexical borrowing is generally avoided, however there are well documented cases of structural borrowing: it is likely that Kubeo (Tukanoan)

---

[11] For one such mismatch, see Carvalho (2016b).
[12] For a list of some shared forms see Aikhenvald (2019).

expanded the loci of its classifier system under influence from an Arawakan language, and that Tariana (Arawakan) did the same under influence from a Tukanoan language (Seifart 2007). Lexical borrowing is also seen, however: Resígaro has borrowed around 36 out of 56 forms from Bora (Witatoan), with which it has had intense contact (Seifart 2007).

The morphosyntactic similarities of the classifier systems across the region obscure possible origin points. In many families, classifier systems can only be reconstructed to a subfamily level rather than a proto family level, suggesting that classifiers as an areal feature are something of an innovation. In the Tukanoan family, Wiegertjes (2022) reconstructs a gender system to Proto-Tukanoan, which allowed the possibility of classifier development. In the Witotoan family, it is possible to reconstruct a system to the subfamily level, but it is so far not possible to reconstruct this back to Proto-Witotoan (Seifart 2007). At this point, much remains to be discovered for both the Arawakan classifier systems and the systems of the wider region and, although this paper focuses on family-internal developments, the impact of contact should not be ignored.

### 3.3.  Linguistic features of Arawakan

As discussed, the Arawakan languages form a widespread family with diversity in their linguistic structures. This section will introduce some of the primary linguistic features of the nominal system in order to contextualise the rest of the discussion. The Arawakan languages are polysynthetic, agglutinating, and head marking (Aikhenvald 1999). Although the languages are predominantly suffixing, there is a largely uniform set of cross-referencing prefixes which encode gender, often referred to as cross-referencing indexes, which will be discussed in detail in **3.3.2**. As well as classifiers, two nominal constructions are of relevance here: possession and compounding. A brief introduction to possession will be given here, and a longer treatment of compound systems will be given in **3.3.3**. The classifier system will also be introduced in **3.3.1**.

Within possessive constructions, an alienability distinction recognises three noun classes: alienable, inalienable, and non-possessable, with some further distinctions within these (Aikhenvald 1999). Across the family, body parts and kinship terms are usually treated as inalienable items, along with some other nouns. This distinction dates to PA; Payne (1987; 1990) reconstructs three markers for alienable classes (*-ni*, *-te*, *-re*), as well as a final vowel change, and zero marking in the inalienable class. In all of these, person is marked with a cross-referencing prefix. Examples of these are given below (12), (13), from Ashéninka (Kampa, Arawakan).

(12) Alienable possession: ASHÉNINKA (KAMPA, ARAWAKAN: PERU)

    a.   no-sari-**ne**
        1SG-macaw-POSS
        'my macaw'

    b.   no-toniro-**te**
        1SG-aguaje.palm-POSS
        'my aguaje palm'

    c.   n-aNpee-**re**
        1SG-cotton-POSS
        'my cotton'

    d.   no-yanire              (cf. kaniri 'manioc')
        1SG-manioc.POSS
        'my manioc'

                                             (Payne 1990, 80)

(13) Inalienable possession: ASHÉNINKA (KAMPA, ARAWAKAN: PERU)

no-čʰarine                    (cf. čʰarine 'grandfather')

1SG-grandfather

'my grandfather'

(Payne 1990, 81)

In the examples above, the order Possessor-Possessed is seen, and this is most common through the family both with full nouns and with indexes, although the Possessed-Possessor order is also found (Aikhenvald 1999). This construction is related to the compound construction, described in **3.3.3.**

### 3.3.1. Classifiers in Arawakan

Classifier systems in the Arawakan languages vary in the size and function, ranging from less than ten to over a hundred forms. These include grammaticalised monosyllabic forms with no known cognate and multisyllabic forms with cognates in the same and neighbouring languages. They can be used with a range of morphosyntactic loci, including the numeral, verb, noun, possessive constructions, demonstratives, and, in some marginal cases, locatives. In most of the languages presented here classifiers occur as bound suffixes (14), however they also occur as infixes in numbers in Baure (15), and as prefixes with the numeral 'two' in Warekena (16).

(14) YUKUNA (ARAWAKAN: NWA)

wéjiké-hilá            a'napitá

three-CL:long         arm

'three arms'

(Lemus Serrano 2020, 58)

(15) BAURE (ARAWAKAN: GUAPORÉ-MAMORÉ)

po-**po**-š

one-CLF:tiny-one

'one fish'

(Danielsen 2007, 139)

(16) WAREKENA (ARAWAKAN: NWA)

**e**-naba               '(two) human masculine'

**tuwa**-naba          '(two) human feminine'

**pamiña**-naba       '(two) animals'

(Aikhenvald 1998, 299)

Most of the systems surveyed here constitute closed systems, although Baure also makes use of repeaters (classifiers which share a form with a noun). As mentioned, classifiers can occur on a number of morphosyntactic loci and in most cases the same set of forms is used in all positions. However, some languages show restrictions in different positions, such as Alto Perené, which only allows five classifiers to occur on verbs. Palikur forms a unique case, as some forms show phonological differences when occurring with different loci, as in the case of -*tra* 'NUM. CL: linear' and -*buhku* 'LOC. CL: linear'.

These systems will be the focus of study in **5.**

### 3.3.2. Classifiers and gender in Arawakan

As already mentioned, there is a set of cross-referencing indexes which appear in a roughly uniform manner across the family. These are used with nouns and verbs: to mark the subject of a transitive or intransitive verb and to mark the possessor of a noun. As indicated in Table 4, these can be reconstructed to PA.

TABLE 4

PROTO-ARAWAKAN CROSS-REFERENCING AFFIXES (AIKHENVALD 2019)

| | Prefixes | | Suffixes | |
|---|---|---|---|---|
| | sg | pl | sg | pl |
| 1 | *nu- or ta- | *wa- | *-n or *-te | *-wa |
| 2 | *pi- | *(h)i | *-pi | *-hi |
| 3nf | *ri-, i- | *na- | *-ri, -i | *-na |
| 3f | *thu-, ru-, u- | *na- | *-thu, -ru, -u | *-na |

In the third person, the cross-reference indexes distinguish (usually) feminine and non-feminine gender. Gender is also marked as agreement markers on nominal modifiers. This is demonstrated in the examples below, which show feminine (17a) and non-feminine (17b) marked on the modifier and verb, as well as in the choice of demonstrative in Warekena.

(17) WAREKENA (NWA)

a. **ayuta** neyawa weduana-ɾi-**yawa** **yu**-tapa-pa
    DEM.DIST.FEM.SG woman good-ADJ-FEM 3sgfem-come-REDUP
    'That good woman is coming.'

b. **eta** enami weduana-**ɾi** **i**-tapa-pa
    DEM.DIST.NON.FEM.SG man good-ADJ.NON.FEM 3sgnf-come-REDUP
    'That good man is coming.'

(Aikhenvald 2012, 282–83)

In **2.2.1**, the distinction of concurrent systems was established as necessitating difference in both form and semantics. There is reason to maintain a distinction between gender and classifiers in the Arawakan languages (Aikhenvald 1994a), although there are some crossovers in both the semantics and the forms, which will be discussed here.

Particularly in the languages of the NWA, gendered classifiers, which reflect the distinctions of the wider gender system, are included within the classifier system, suggesting semantic overlap of the classifier system and the gender system. This is the case in Resígaro, Tariana, and Warekena, which are all included in this study. In each of these languages, the gendered classifiers show the same behaviour as other classifiers, rather than the behaviour of the wider gender system, and also usually differ in phonological form from the cross-referencing indexes. Outside of the NWA, only Yanesha' has a gendered classifier (although it does not have cross-reference indexes, see Duff-Trip (1997, 245)), although Wapishana also has a gendered compound noun. Both the cross-referencing affixes and the gender sensitive classifiers are given below in Table 5.

TABLE 5

GENDER SENSITIVE CLASSIFIERS AND CROSS-REFERENCING PREFIXES IN ARAWAKAN LANGUAGES

| | | Gender sensitive classifiers | Gender sensitive cross-referencing prefixes (sg) |
|---|---|---|---|
| Tariana (Aikhenvald 2003) | Feminine | *-ma* | *du-* |
| | Non-feminine | - | *di-* |
| Resígaro (Seifart 2012a) | Feminine | *-pɨ̃jé* | *do-* |
| | Non-feminine | *-gí* | *gi-* |
| Warekena[13] (Aikhenvald 1998) | Feminine | *tuwa*(*naba*)[14] | *yu-* |
| | Non-feminine/Masculine | *e*(*naba*) | *ø-*/*i-* |
| Yanesha' (Duff-Tripp 1997) | Feminine | *-apñor-* | - |
| | Non-feminine | - | - |
| Wapishana (Santos 2006) | Feminine | (*a*)*ba*[15] | *ɽu* |
| | Masculine | - | *ɽɨ* |

Of the other languages included in this study, none show gendered markers within the classifier system, and the gender and classifier systems show different morphosyntactic behaviour, having different target possibilities, and showing different agreement patterns, as is described here for Alto Perené. In (18), gender is marked twice on the verb, indicating agreement with the grammatical noun class of the subject and the object, and on the possessed noun indicating a non-masculine possessor, while a classifier is marked on the possessed noun indicating a physical characteristic of the possessed object. As well as a difference in possible targets and realisation, this also shows a difference in controller.

(18) ALTO PERENÉ (ARAWAKAN: PERU)

**ant**-atz-i-**ro**          **o**-shinki-**a**-te

3NM.A.make-PROG-REAL-3NM.O     3NM.POSS-maize-CL:liquid-POSS

'She was making her fermented corn drink'

(Mihas 2015, 416)

The data presented here and in previous studies of gender and classifiers in Arawakan (see Aikhenvald 1994a; 2019) do not suggest that there is any evidence to consider these as anything other than two distinct systems. However, the overlap in semantics of the two systems has primarily occurred in the NWA and not in other areas, and this development may prove interesting in future investigations of the influences of contact and unique changes in the languages of this area.

### 3.3.3. Classifiers and compound systems in Arawakan

As was mentioned previously, many Arawakan languages show a secondary nominal system of compounding, by which nouns are compounded with a closed or semi-closed set of forms to derive new lexical items. These sets are in many ways similar to the classifier sets, although with differences in distribution and semantic scope. This is of particular interest for the languages of the Guaporé-Mamoré region: compound systems are widespread across many unrelated languages in this region, and Van der Voort (2018) proposes that these are a likely source of the expanded classifier systems. In the Arawakan languages of the SWA, such systems are found in Baure and Paresi. Beyond this sphere, extensive systems are found in

---

[13] In Warekena, this distinction is made only in the classifier with 'two'; with 'one' both classifiers are *peya*.

[14] Classifiers in Warekena has phonologically dependent on numerals; the classifiers are *tuwa-* and *e-*.

[15] This form is a compound noun rather than a classifier; the distinction will be discussed in **3.3.3**.

Apurinã, Wapishana, and Alto Perené.[16] The languages of the NWA do not have such systems: in Yukuna, compound constructions are possible, but only minimal description is available and there appears to be limited similarity with the systems under consideration here;[17] in Warekena, there is a minimal system in which verbal roots and speech act participants can be compounded to derive a noun. As these do not show similarity with the systems in the rest of the family, they will not be further considered here. Given this, the systems most similar to each other are found in Baure, Paresi, Apurinã, Wapishana, and Alto Perené. The examples below demonstrate possible compounds in each language. As can be seen, many of these (20), (22) relate to plant parts, and others (19), (21) relate to both human and animal body parts. In each example, the bound noun is used for a derivational function, creating a new noun.

(19) BAURE (GUAPORÉ-MAMORÉ)

   *tiporekpo'e*
   tiporek-**po'e**
   chicken-head
   'chicken head'

(Danielsen 2007, 96)

(20) PARESI (ARAWAKAN: GUAPORÉ-MAMORÉ)

   atya-**hana**      'tree leaf'
   takola-**hana**    'bamboo leaf'
   zohitya-**hana**   'cashew leaf'

(Brandão 2014, 182)

(21) APURINÃ (ARAWAKAN: BRAZIL)

   potokoko-**piti**   'dove feather'
   pataru-**piti**     'chicken feather'
   irẽka-**piti**      '"mutum" (bird) feather'

(Facundes 2000, 168)

(22) ALTO PERENÉ (PERU)

   o-tzina-tai-t-aty-a                kipatsi-**pani**
   3NM.S-rise-NEW.DISC-EP-PROG-REAL soil-CL.T:powder.like
   'Dust particles went up into the air [after the vehicle passed].'

(Mihas 2015, 412)

(23) WAPISHANA (ARAWAKAN: BRAZIL, GUYANA)

   a.  atamɨn-**ak**
       àrvore-TCL:fruta
       'fruta da àrvore' (tr. 'tree fruit')
   b.  wabu-**ak**
       açaí-TCL:fruta
       'fruta do açaizeiro' (tr. 'açaí fruit')

(Santos 2006, 107)

It has been argued for some of these languages that the compound nouns evolved from inalienable nouns in the possessive construction already discussed in **3.3** (see Baure: Danielsen 2007, 134; Apurina: Facundes

---

[16] An interesting case is the system in Wapishana, where the number of bound compound forms far exceeds the number of classifiers, and all that distinguishes these forms is the possibilities of the morphosyntactic loci.

[17] There does not appear to be a closed set of bound forms which occur in compounds, and they do not appear to be used derivationally, but only coreferentially.

2000, 150; Paresi: Brandão 2014, 179). However, in the languages with established (and well-described) compounding systems, it is clear that the distinction between classifiers and bound nouns in compounds is somewhat blurry. In all languages except Wapishana, the classifiers have a wider semantic meaning than the bound nouns used in compounds, suggesting that grammaticalisation has caused semantic bleaching, the loss of some lexical material. However, in both Baure and Apurinã, the set of bound nouns can occur in more positions than just with another noun: in Baure, the set of bound nouns can function morphosyntactically as classifiers, but show distinctions both in semantics (being more specific) and in their use as independent lexical items (which is not possible for classifiers). In Apurinã, the distinction is primarily a semantic one, as classifiers do not show the same range as in other languages. A summary of the possibilities for bound nouns in relation to classifiers is given in the table below; languages of the Guaporé-Mamoré region are indicated in italics.

TABLE 6
BOUND NOUNS AND CLASSIFIERS POSSIBILITIES

|  | Classifiers have broader semantic meanings | Classifiers can occur in wider set of morphosyntactic environments | Bound nouns can occur both independently and as bound nouns |
|---|---|---|---|
| *Baure* | Y | N | Y |
| *Paresi* | Y | Y | N[18] |
| Apurinã | Y | N | N |
| Alto Perené | Y | Y | N |
| Wapishana | Y | Y | Y |

Throughout this paper I will maintain the distinctions drawn in the respective grammars, however as I will describe in 4, both classifiers and compound nouns will be used in this study. The discussion of bound nouns and their relationship with classifiers will be returned to in 7.

### 3.4. Conclusion
This section has given an overview of the geographical and historical relevance of South America and the Amazonian basin, as well as an introduction to the Arawakan language family and its linguistic features, and the work that exists on PA. The aim has been to highlight the importance of areal linguistics in understanding the position of the Arawakan family. 4 will discuss the methodology used in this study; 5 and 6 will present the data; 7 will discuss the findings, and the final conclusions will be presented in 8.

---

[18] This is not described for Paresi, however based on the details available bound nouns do not appear to be able to be used independently.

# 4. Methodology

This section will describe the methodology used to address the questions set out in 1, and will discuss how the sample and methods were chosen and applied. As laid out in 1, this study aims to identify a common diachronic source of the Arawakan classifier systems. To do this, both the grammatical structures and the classifier forms will be considered from languages across the family.

The data presented in this paper is based on a sample of 11 languages, due to both the scope of this paper and the lack of sufficient documentation for every language. As the aim here is to gain an insight into the PA nominal classification system, one language from each subfamily (based on Aikhenvald's (1999) classification) has been selected. Additionally, in order to gain some insight into the extent to which the Arawakan languages of the NWA diverge from those in the rest of the family, four languages from this region were selected (Resígaro, Yukuna, Tariana, Warekena). The languages chosen are given in Table 7, along with their ISO code, familial classification to two levels and the primary source used for this paper. The names given below are those used throughout, and where possible are based on the preferred names of the speaker community or that used in the primary grammar, although widely used alternatives are given in brackets. Ultimately, it was not possible to include a language from the Caribbean subfamily, as the languages either lack classifiers or the classifiers are known to come from a non-Arawakan source (for Garifuna, see Haurholm-Larsen 2016). In tables throughout, languages in the Guaporé-Mamoré region are highlighted in italics; languages in the NWA are highlighted in bold. A map of the sample is also included below the table.

TABLE 7
SAMPLE OF ARAWAKAN LANGUAGES

|  | Classification (Aikhenvald 1999) | Primary Source |
|---|---|---|
| *Baure (brg)* | South Arawak (SSWA) | (Danielsen 2007) |
| *Paresi (Parecís, Pareci: pab)* | Paresi-Xingu (SSWA) | (Brandão 2014) |
| Apurinã (Ipurinã: apu) [19] | South Western Arawak (SSWA) | (Facundes 2000) |
| Alto Perené (Ashéninka Perené: prq) [20] | Kampa (SSWA) | (Mihas 2015) |
| Yanesha' (Amuesha: ame) | Amuesha (SSWA) | (Duff-Tripp 1997) |
| Wapishana (wap) [21] | Rio Branco (NA) | (Santos 2006) |
| Palikur (plu) | Palikur (NA) | (Aikhenvald and Green 1998) |
| **Resígaro (rgr)** | Colombian (NA) | (Allin 1976) |
| **Yukuna (ycn)** | Colombian (NA) | (Lemus Serrano 2020) |
| **Tariana (tae)** | Upper Rio Negro (NA) | (Aikhenvald 2003) |
| **Warekena (Baniva de Maroa: gae)** | Orinoco (NA) | (Aikhenvald 1998) |

---

[19] Facundes (2000) states that Apurinã is the Portuguese term for both the language and the people, however it is used consistently in literature to refer to the language.

[20] Alto Perené is the variety of Ashéninka which is described in Mihas (2015), and is the preferred autonym of the speakers.

[21] I follow Gomes (2022) in using the anglicised autonym Wapishana rather than Wapichan, which is also used by speakers and more accurately reflects the pronunciation of the name.

## FIGURE 8
### SAMPLE OF ARAWAKAN LANGUAGES.[22]



Selecting such a broad sample allows for a broad study, but it does not allow for a deep one. Ultimately such choices must be made in constructing a study of this small size, and it is hoped that the results gained from this sample provide evidence that further in-depth study is both necessary and worthwhile.

The source material for this study consists of the classifier sets as presented in the grammars referenced above, as well as some closed sets of compound nouns where available.[23] These sets are provided in the appendices (I-XI). Not all reference grammars describe such sets with the same terminology; Facundes (2000) refers to two sets of 'classificatory nouns' for Apurinã, reasoning that these nouns have the semantic properties of prototypical classifiers, but the syntactic (and at times semantic) properties of noun class and gender systems. This is a terminological issue, and one which has already been addressed; the classificatory nouns show similar morphosyntactic behaviour as the classifiers in other languages, as they are used in derivational noun compounding, can be used anaphorically, and can be incorporated into the verb (Facundes 2000). For this reason, I consider the semantically extended set of classificatory nouns (Facundes' CN²s) equivalent to the classifiers of other languages, and the non-extended set, CN¹s, equivalent to compound nouns for the purposes of this study, and refer to them as such throughout. Palikur uses different classifier forms with different loci (Aikhenvald and Green 1998); these will all be included in this study and will be referred to by their locus.

Not all classifiers were included in this study: there is a great range in the size of classifier sets, and the only semantic basis shared among all languages is classifiers relating to shapes and physical properties. This was supported by an initial study of the complete cognate sets in Paresi, Apurinã, and Alto Perené, which have midsized classifier sets (20-60 forms) and Baure, which has over 100.[24] As such, due to constraints of

---

[22] Maps were generated using Glottospace (Norder et al. 2022).

[23] The compound nouns from Baure were not included as, although the construction is in productive use, they do not form a closed set.

[24] Only one cognate was found by this initial survey not relating to physical form; this is discussed in 5.

time and scope, it was decided to focus on these delimited classifier sets rather than the complete sets. The classifiers included are marked in the relevant appendices.

The analysis conducted in this study will be presented in two sections. In the first section, **5**, an overview of the morphosyntactic properties of the Arawakan classifier systems is presented, with particular reference to the classifier locus and the primary functions of the classifier systems. This explanation is based on the descriptions and examples provided in the grammars references in **7**, above; other materials are cited as necessary.

The second section, **6**, is based on a study of the classifier forms, in which cognates were identified using the comparative method: a systematic study of sound correspondences with the aim of uncovering forms derived from the same lexeme at an earlier stage of the language family. For this process, I made reference to forms with both corresponding phonological shape, and some semantic adjacency, as described below.

There has been little reconstructive work on PA, and as such the primary source for PA phonology has been Payne (1991). However, as discussed previously, this reconstruction is not without its issues, and where possible I have further supported the given phonological reconstructions with studies based on lower level subgroups, particularly the work of Michael et al. (2010), Jolkesky (2016), and Carvalho (2016a; 2018; 2021).

I have aimed to support the proposed cognate sets with evidence of semantic relatedness, particularly given the lack of phonological work on PA. Two issues must be borne in mind when considering the semantics of the classifiers presented here: that of semantic shift, and the use of analogical extension in classifier use. As discussed in **2.3**, classifiers are formed through the grammaticalisation of lexical nouns, and semantic changes often seen are shift, widening, or bleaching. Even with a common source, the changes may vary greatly between languages. Secondly, the use of classifiers has a pragmatic basis, and the semantics are often underpinned by analogical extension,[25] as was found in the study of the underlying semantics of shape based classifier systems by Adams and Conklin (1973), introduced in **2.2.2**. How these might be applied is likely to be language and society dependent.

It is intended that by taking into account both the grammatical structures and the classifier forms, it will be possible to identify what, if any, PA material allowed the evolution of such diverse systems seen today. In identifying this, and in considering the semantic basis of the classifiers, it is also hoped to shed light on the proposals put forward by Aikhenvald (2019) and Van der Voort (2018) discussed in **3.2.2**. These conclusions will be discussed in **7**, and concluded in **8**. The following chapters, **5** and **6**, will present the data.

---

[25] For an account of this in Apurinã, see Facundes (2000: 169-175).

# 5. Arawakan classifiers

Having presented the theoretical and linguistic background necessary to move forward, this chapter will consider the grammatical functions and syntactic structures of the Arawakan classifier systems included in this study. In 3, it was discussed that the languages of the NWA are considered to hold a unique position within South America, due to their use of classifiers with multiple loci. The aim here is to present evidence that the morphosyntactic characteristics which have been claimed to be unique to the NWA, which only comprises some of the Arawakan languages, are present throughout the family. The discussion will begin with a presentation of the syntactic structures in the various languages (5.1), before considering the functions of the systems (5.2), which are closely associated with the morphosyntactic loci. It will conclude with a discussion of the semantic basis of the various systems (5.3). Where possible, comparison will be made with other languages in the area.

## 5.1. Morphosyntactic locus

The issues and relevance of morphosyntactic locus was discussed in 2: in this section, the morphosyntactic locus will be used as a means to compare the different grammatical structures across languages.

Across the Arawakan family, classifiers can be found with numerals, nouns, modifiers, verbs, possessives, locative nouns, and interrogatives. Of these, the first four can be considered major positions, seen in at least five of the surveyed languages. This section will demonstrate how these positions are used with classifiers across the family and discuss the consequences of this for this study.

### 5.1.1. Numeral locus

The numeral position for classifiers is the most widespread position in the sample. In Warekena, this is the only position in which classifiers are seen (Aikhenvald 1998, 298–99), and in Yukuna, this is the only position in which classifiers are found in most recent descriptions (Lemus Serrano 2020, 57–58).[26] However, Apurinã does not have classifiers in this position (Facundes 2000).

With numerals, classifiers characterise an aspect of the head noun, as seen in example (24) below.

(24) Tariana (NWA)

| pa:-**kha** | kule-**kha** |
|---|---|
| one-CL:curved | fishing.tool-CL:curved |

'one fishing line'

(Aikhenvald 2003, 88)

It is often the case that classifiers occurring on numerals can also occur on quantifiers, as in example (25), below.

(25) Baure (Guaporé-Mamoré)

a. *roti' roenisa popoš him pitir piti'*

| roti' | rʋinisa | po-**po**-š | him | pitir | piti' |
|---|---|---|---|---|---|
| 3SGm | 3SGm=fish | one-CLF:tiny-one | fish | 2SG.POSS | 2SG |

'He caught one fish for you.'

---

[26] There are conflicting reports on usage in Yukuna. Although some reports indicate that classifiers can be used in nouns, adverbs, and adjectives in incantations, a much more restricted use was most recently reported (see Lemus Serrano 2020, 57–58 for discussion). This restricted usage in incantations points to a historic usage, and I have included the system as it is most broadly described. Yukuna is an endangered language, and this loss is likely a consequence of attrition.

b. *iyowkopan te kokononev*

iyowko-**po**-a-no             te              kokono-nev

various-CLF:tiny-LK-NOM1 DEM1.M        caterpillar-PL

'various caterpillars'

(Danielsen 2007, 139)

In some languages, as is the case in Baure, classifiers are obligatory with all numerals (Danielsen 2007, 138–40). In others, as in Alto Perené, classifiers are optional, and may not occur often in this position (Mihas 2015, 144).

As discussed in **3.2.2**, Aikhenvald (2019) proposes the numeral locus as the oldest of Arawakan systems in the NWA, as it is the only position consistent among those languages. This is also true of this sample, excluding Apurinã, suggesting not only that the numeral position may be the oldest in the wider family, but that it can be reconstructed to PA, and has later been lost in Apurinã. Furthermore, classifiers in this position can be shown to be a familial rather than areal trait: outside the Arawakan family, the numeral position is also seen in many languages of the NWA, including Yagua and the Witotoan languages (Seifart and Payne 2007). It is rarer in the SWA, but is seen at least in Kwaza and Itonama, which are both isolates (Van der Voort 2018).

### 5.1.2. *Nominal locus*

Both Van der Voort (2018), regarding languages of several families in the SWA, and Aikhenvald (2019), regarding NWA Arawakan languages, propose that classifiers could have originated in the nominal position, from a set of bound nouns. However, classifiers occur with nouns in only seven of the languages, as they do not occur here in Palikur (Aikhenvald and Green 1998) or Warekena (Aikhenvald 1998, 298–99), and only marginally in Wapishana and Yukuna.[27] If this position is shown to be the originating position of classifiers, it must be assumed that the classifier system in these languages has eroded, something which is already described for Yukuna (Lemus Serrano 2020, 58).

As was discussed in **2**, classifiers in the nominal position most commonly have a function of categorising a noun by its core features, as in many Australian languages. None of the Arawakan languages show this function, instead demonstrating a derivational function as in example (26), below.

(26) APURINÃ (BRAZIL)

mãko-**tsota**

mango-trunk.of

'mango tree trunk'

(Facundes 2000, 163)

In Paresi, classifiers can also have a nominalising function, when attached to stative verbs (Brandão 2014, 193), as shown in the example below. However, this function is not widespread throughout the family.

---

[27] In Wapishana, only the numeral, verb, and possessive positions are described as available for classifiers, however a number of classifier forms also occur as bound compound nouns (Santos 2006, 114–15). In Yukuna, classifiers are reported as occurring on nouns in incantations, and are not found in more recent descriptions (Lemus Serrano 2020, 57–58).

(27) PARESI (GUAPORÉ-MAMORÉ)

> *wiyeri*
> wiya-**li**
> be.sweet-CLF:round
> 'candy'

<div align="right">(Brandão 2014, 194)</div>

Classifiers are widely found in the nominal position throughout the Amazon, and as in Arawakan this is most often found with a derivational function (Seifart and Payne 2007; Van der Voort 2018). Although little can be drawn from this in considering familial or areal influence, the consistency of occurrence across the Arawakan family would indicate the possibility of the nominal position being inherited: it is found in all language branches, and does not show areal patterning.

### 5.1.3. Modifier locus

The term modifier here is used in a general sense to capture adjectives,[28] demonstratives, and other possible nominal modifiers. The modifier locus[29] is discussed here separately from the numeral locus as the classifiers in this position show a different distribution: only five languages allow classifiers with a nominal modifier: Baure (Danielsen 2007, 139), Paresi (Brandão 2014, 193–94), Resígaro (Allin 1976, 152), Alto Perené (Mihas 2015, 414), and Tariana (Aikhenvald 2003, 88). The distribution of classifiers in the modifier position varies across languages. In some languages, it is only possible with adjectives (Baure), or with demonstratives (Paresi), while in others it is possible to see classifiers in all modifier positions (Tariana). (28) shows a classifier with an adjective, while (29) shows a classifier with a demonstrative.

(28) ALTO PERENÉ (PERU)

| p-amen-i | antaro-**meni** | kotsiro |
|---|---|---|
| 2S-look-IRR | big-CL:flat.triangular | machete |

'Look for a machete with a big blade.'

<div align="right">(Mihas 2015, 423)</div>

(29) PARESI (GUAPORE-MAMORÉ)

> *ezetse naikoli kaweta*

| eze-**tse** | n=aikoli | ∅=kaweta |
|---|---|---|
| This-CLF:small | 1SG=tooth | 3SG=hurt-IFV |

'This tooth hurts.'

<div align="right">(Brandão 2014, 195)</div>

The nominal and modifier positions are closely related: in both Apuriña and Yanesha' a classifier can appear on the noun but not the modifier (Facundes 2000, 169–79; Lemus Serrano 2020, 57–58), but the reverse is not possible in any language.

It is possible that the modifier position spread either from the numeral position or from the nominal position. However, as it is not widely found in the family, it is not clear that this should be reconstructed as a grammatical possibility for PA. An alternative scenario would be that classifiers in the modifier position developed independently in subfamilies: this position is widely attested in the NWA, in the Witotoan languages and in Yagua (Seifart and Payne 2007), and it is possible that this influenced its development in the

---

[28] Many languages lack an adjective class; I have followed the description of the grammar in defining word classes, and also in identifying something as occurring in a modifying position.

[29] Classifiers with the modifier are considered part of the noun classifier type (Grinevald 2002). As the focus here is on behaviour rather than definition of systems, these loci are kept separate.

NWA Arawakan languages. However it is rarer in the SWA, appearing only in Karo (Tupian), where classifiers mark agreement on both the modifier and the noun (Van der Voort 2018), suggesting that contact is less likely to have been involved in the development of this position in Guaporé-Mamoré, where both Baure and Paresi are found. The development of the modifier position should be investigated at lower levels before a PA reconstruction is attempted.

### 5.1.4. *Verb locus*

The verbal position is used in eight of the languages, but it is not used in Resígaro (Allin 1976), Yukuna (Lemus Serrano 2020), or Warekena (Aikhenvald 1998), and it is only possible for a small number of forms in Alto Perené (Mihas 2015, 414). Across the family, this occurs as verbal incorporation of the classifier, which references the object of the verb. This can occur both with (30) and without (31) the overt lexical noun.

(30) BAURE (GUAPORÉ-MAMORÉ)

 *vehpaw to etip*

 vi=eh-**po**-a-wo    to  etip

 1PL=wash-CLF:tiny-LK-COP ART manioc.starch

 'We wash the manioc starch.'

(Danielsen 2007, 139)

(31) APURINÃ (BRAZIL)

 ata  yotika-**xiti**-ta   txa-ru

 1PL  burn-CL:earth.of-VLBZ AUX-3M.OBJ

 '... we set it (the field farm) on fire.'

(Facundes 2000, 165)

In most of the languages described here, incorporated classifiers are used, as in the examples, to modify or refer to the nominal referent. However, there is some variation in use: in Palikur, classifiers incorporated in stative verbs indicate the complete involvement of the subject, while in transitive verbs, it indicates direct physical contact with the object (Aikhenvald and Green 1998, 448–51).

Outside the Arawakan family, classifier systems show a different pattern to their use of the verbal locus than to the nominal and numeral loci. This position is more widespread in the SWA than in the NWA, as it is not found in the Witotoan languages or in Yagua (Seifart and Payne 2007). In the South West, they occur in all but one of Van der Voort's (2018) sample: as in Arawakan, they are found as incorporated nouns referring to an argument of the verb, but also as nominalisers (32) and as verbal modifiers (33). In (32), the classifier is used with an intransitive predicate which must be nominalised in order to be used as an attributive adjective. In (33), the classifier modifies the action of the verb, rather than referring to a property of the argument, as is reflected in the translation.

(32) KWAZA (ISOLATE: GUAPORÉ-MAMORÉ)

 (axy) haka-**xy**

 house old-CLF:house

 'old house'

(Van der Voort 2018, 218)

(33) MOVIMA (ISOLATE: GUAPORÉ-MAMORÉ)

| dan-a-**ba**=as | os | rey | dowe |
|---|---|---|---|
| chew-BDR-BR:round=NEUT.AB | ART.NEUT.PST | again | clothes |

'It [the cow] chewed my dress into a ball.'

(Van der Voort 2018, 217)

This difference in distribution between the NW and SWA is also reflected in the Arawakan sample considered here. Classifiers are found in the verbal position in every language outside the NWA, including the two which are located in the Guaporé-Mamoré region (Baure and Paresi). In the NWA, only Tariana has classifiers in the verbal position. That incorporation is possible in Tariana, against the pattern of areal influence, points to a family-internal development, and here it is tentatively proposed that incorporation of the classifier to the verb was possible at PA, but was later lost in the NA languages.[30]

### 5.1.5. *Locus: Conclusion*

In exploring the morphosyntactic possibilities of the Arawakan systems, I have aimed to show that more than reflecting the systems of the wider areas, the languages show a lot of consistency between each other. In the wider region, the nominal and numeral loci are both positions in use in languages throughout the region, although the numeral position is less widely found in the SWA. Both the nominal and numeral loci are widespread throughout the family, and the modifier position is available as part of the nominal locus in some languages. Within the family, the verbal locus is the only position which reflects the surrounding region, as it is found commonly in the SWA in both Arawakan and non-Arawakan languages, although sometimes with differing functions, but less so in the NWA. The occurrence of the four major classifier loci discussed above are summarised in the table below.

TABLE 8
CLASSIFIER LOCI IN ELEVEN ARAWAKAN LANGUAGES

| | Numeral | Noun | Modifier | Verb |
|---|---|---|---|---|
| *Baure* | Y | Y | Y | Y |
| *Paresi* | Y | Y | Y | Y |
| Apurinã | | Y | | Y |
| Alto Perené | Y | Y | Y | Y |
| Yanesha' | Y | Y | | Y |
| Wapishana[31] | Y | (Y) | | Y |
| Palikur | Y | | | Y |
| **Resígaro** | Y | Y | Y | |
| **Yukuna**[32] | Y | (Y) | (Y) | |
| **Tariana** | Y | Y | Y | Y |
| **Warekena** | Y | | | |

The morphosyntactic locus was defined on a positive basis: if a classifier can be shown in this position either through the grammar description or an example, it is considered to occur in that position.

---

[30] It is unclear to what extent this extended to noun incorporation as a whole as the relationship between classifier incorporation and noun incorporation has yet to be explored. In the absence of further evidence, I remain with the proposal above, with no comment on incorporation of the lexical noun.

[31] Only the numeral, verb, and possessive positions are described as available for classifiers, however the nominal position is available for compound nouns. Given the overlapping forms between compound and classifier systems in Wapishana, I have also indicated this position in brackets.

[32] Brackets indicate historic usage.

Based on the distribution, it is possible to propose that the numeral, nominal, and verbal positions were available in PA, although erosion of this full system is seen in some languages, in the numeral position in Apurinã, the nominal position in Palikur, and all positions but the numeral in Yukuna and Warekena. This begins to answer the questions put forward in 1: the features which were claimed to be unique to the NWA, namely the occurrence of classifiers in multiple positions, can be shown to be common throughout the family, and there is enough consistency in the grammatical structures of the family to point to a PA source of classifiers in the nominal, numeral, and verbal positions.

In the next section, the function of the Arawakan classifiers will be discussed, and it will be shown that these are very closely related to the classifier locus.

## 5.2.  The function of Arawakan classifiers

As highlighted by Krasnoukhova (2012), a core characteristic of the Amazonian classifiers is their multifunctionality: she particularly highlights the agreement and derivational functions, as well as the function of semantic categorisation. With reference to the NWA, Derbyshire and Payne (1990) also highlight their use in anaphoric reference, which is a common feature of classifiers cross-linguistically, while Seifart and Payne (2007) emphasise the derivation and agreement functions, but also mention the occurrence of anaphoric reference. With reference to the SWA, Van der Voort (2018) makes reference to these same three functions, but also notes a nominalising function which is found in three languages of the region. Although not all of these functions are found in every language surveyed here, there is no clear correlation of function to a particular area, including those the NWA. With reference to derivation, agreement, and anaphoric reference, only Baure, Paresi, Alto Perené, Palikur, and Tariana use classifiers for all three, although the majority of languages use classifiers for at least two, showing that the classifiers in Arawakan reflect the multifunctionality of classifiers generally in South America. However, how these functions occur across the languages here is not identical in every case. The following section will review these three functions in a handful of languages, showing the range with which they can appear and the spread across the Arawakan family. As derivation and agreement are widespread functions in both the NW and SWA, reference to areal distributions outside the Arawakan family will only be highlighted as necessary (Seifart and Payne 2007; Van der Voort 2018). The aim of discussing each function is comparative rather than analytical, and therefore broad definitions are taken as the starting point. Finally, there will be some discussion of whether these functions and their variation are important in considering whether the Arawakan classifiers constitute one diachronic system.

### 5.2.1.  Derivation

Derivation is a lexical process associated with the nominal position, which is reflected in this sample: in all languages in which the nominal position is found, it is found with a derivational function. Looking first at the classificatory nouns of Apurinã, classifier -*ku* can be used with a diminutive function which has moved beyond its lexical meaning of 'seed'.

(34) APURINÃ (BRAZIL)
     amarunu-**ku**
     boy-DIM
     'little boy'

(Facundes 2000, 179)

In the example, the diminutive classifier redefines the referent by the characteristic of smallness; the classifier is not added because the characteristic is already present in the referent, but in order to derive a new referent in which the characteristic is present. This is the most common occurrence of derivation in this sample.

Another example comes from Resígaro, as is seen in (35), where a new word is formed with the addition of the classifier. However, a different derivational function is also found in Resígaro. Noun stems are grammatically uncountable, and as such, the derivational function of the classifier is to derive a countable noun (Seifart 2012b). This is shown in (36).

(35) RESÍGARO (NWA)
    a.   hooní
        'water'
    b.   hooni-ʔ**éhú**
        water-CL:hole
        'a well'

(Allin 1976, 151, 161)

(36) RESÍGARO (NWA)
    mi-**mii**-kú           hiítá-**mii**-kú
    two-CL:canoe-DL     canoe-CL:canoe-DL
    'two canoes'

(Allin 1976, 201)

Although using classifiers to individualise a noun is common in South East Asian languages (Bisang 1999), Resígaro is the only example of individuation in this sample.

There are two languages in the sample, Warekena in the NWA, and Palikur, which are considered to have a derivational function but which do not show classifiers appearing on nouns. In Warekena, the derivational function is marginal and, from the available data, is likely fossilised. Based on the only example, (37), it is possible that this may also have had a nominalising function, allowing nominal morphology to occur on a numeral. Although it is unclear to what extent this is productive in the language, this may point to a previously productive process.

(37) WAREKENA (NWA)
    pa-**buya**-pe
    one-CL:time-PL
    'several times'

(Aikhenvald 1998, 299)

In Palikur, the derivational function of classifiers is again marginal, through the verbalisation of nouns to create stative verbs, as in (38), below.

(38) PALIKUR (BRAZIL; FRENCH GUIANA)
    ahamna-**boy**-ye
    leaf-V.CL:flat-DUR.NF
    'leaf-coloured' (lit. 'to be leaf-coloured')

(Aikhenvald and Green 1998, 448)

Derivation is found in nearly all languages in the sample, although it is more restricted in languages of the NWA, as it is fossilised in Warekena and not found in Yukuna. It is closely related to the nominal position, although there are differences in how it is used within each language. It is likely that the development of the nominal classifier locus was as a derivational position, and that diversification of this function has resulted from innovation in use and contact with non-Arawakan languages at the subfamily level.

### 5.2.2. Anaphoric reference

Anaphoric reference is a discourse function in which a classifier can be used in place of an overt noun. This appears in three ways in a number of the languages: through verbal incorporation, through the use of the classifier and numeral with no overt noun phrase, and through use of the classifier in isolation in place of the noun phrase. It is not described for Resígaro (Allin 1976), Warekena (Aikhenvald 1998), or Yukuna (Lemus Serrano 2020), but is available in some from in all other languages sampled.

The first type, anaphoric reference through verbal incorporation, is seen in Paresi, as in the example below:

(39) PARESI (GUAPORÉ-MAMORÉ)

*Hahanahaliya tyokanatseta*

| ha=hana | haliya | ∅=tyoka-**natse**-ta |
|---|---|---|
| 3SG=house | along | 3SG=sit-CLF:cylindrical-IFV |

'[The dog] is sitting close to the house'

(Brandão 2014, 194)

In (39), the dog is not given as an overt noun phrase, but through the incorporation of the classifier *-natse* 'CLF:cylindrical'. This is only possible when the referent is already established in discourse. This is the most common form of anaphoric reference in Arawakan, available in all languages which allow verb incorporation.

The second type of anaphoric reference is seen in the use of a classifier and numeral without an overt noun. This is exemplified with Palikur, below.

(40) PALIKUR (BRAZIL; FRENCH GUIANA)

| ba | pis | muwaka | ax | paha-**t** |
|---|---|---|---|---|
| INTER | 2SG | want | eat | one-NUM.CL:vertical |

'Do you want to eat one (banana fruit)?'

(Aikhenvald and Green 1998, 445)

The third type of anaphoric reference, of a classifier occurring independently without an overt noun, is similar to the structure seen above in (40), but rather than being bound to a numeral the classifier is bound to another element. Both Paresi and Tariana make use of a proclitic, giving a meaning along the lines of 'one thing of this type'. In the following example from Paresi, (41), the referent *kozeto* 'corn' is established, and is later referred to only by the use of the classifier *-tse* 'small' attached to a proclitic.

(41) PARESI (GUAPORÉ-MAMORÉ)

*kozeto tximate koni*

| **kozeto** | tximate | koni |
|---|---|---|
| corn | pile? | in.the.midle.of |

'A pile of corn'

[...]

*motehekoatya hoka nitsa, kala nita ene hatse hololo*

| ∅=mo-tya-hekoa-tya | hoka | ∅=nitsa | kala | ∅=nea-ita=ene |
|---|---|---|---|---|
| 3sg=put-TH-REP-TH | CON | 3sg=eat.meat | DUB | 3sg=say-IFV=PST |

| ha=**tse** | ∅=hololo |
|---|---|
| 3sg=CLF:small | 3sg=drop |

'He was crunching it and eating, and then he dropped a grain (of corn).'

(Brandão 2014, 197–98)

This type is more restricted in its distribution; some languages, such as Yanesha', do not report anaphoric reference outside of verb incorporation (Duff-Tripp 1997). Anaphoric reference of some sort is common throughout the languages; outside the NWA it is available to some extent in all languages, although it is restricted to Tariana within the NWA. Beyond the Arawakan family, this function appears to be widespread in the SWA but not found beyond the Arawakan family in the NWA (Seifart and Payne 2007; Van der Voort 2018). This reflects the spread of the verbal classifier position, to which its development can be tied. However, the history of the use of classifiers for anaphoric reference without the use of the verbal structure is harder to trace.

### 5.2.3. Agreement

As already discussed, the broad distinction made between prototypical noun class systems and prototypical classifier systems often depends on whether or not grammatical agreement is a main function of the system (Dixon 1986), and it is the appearance of agreement in the Amazonian systems which forced this distinction to be reconsidered (Doris Payne 1987; Derbyshire and Payne 1990; Grinevald and Seifart 2004). Within the Arawakan languages, it can present in three ways.

The first two ways agreement can appear is as would be expected of a traditional agreement system: with an agreement marker on a modifier, and with an agreement marker on both the modifier and the head. This is schematised in (42), below:

(42) MOD-CL HEAD(-CL)

Many of the languages surveyed allow both of these types. The first type, with the classifier given overtly both on a modifier and on the noun, is exemplified by Tariana in (43), below, although the order of elements can vary.

(43) TARIANA (NWA)

kule-**kha**          matʃ-**kha**
fishing.tool-CL:curved    good-CL:curved
'a good fishing line'

(Aikhenvald 2003, 88)

The second type, without an overt classifier on the controller, is exemplified with Baure in (44), below.

(44) BAURE (GUAPORÉ-MAMORÉ)

*to pari čino-mpe*

to      part    čino-**mpe**
ART     house   old-CLF:flat
'an/the old house'

(Danielsen 2007, 87)

A final potential agreement construction which can be considered here is agreement through incorporation into the verb. Verb-incorporated nouns would not usually be considered agreement markers, as they are not a copy of the noun or some information from it, but the only realisation of that noun (Derbyshire and Payne 1990), and as the marker is outside the local domain of the noun phrase (Corbett 2006). However, there are instances among these languages when the classifier is incorporated into the verb and an overt lexical noun occurs independently, as in Baure, exemplified in (45), below. In cases like this, verb-incorporated classifiers look very similar to other instances of agreement.

(45) BAURE (GUAPORÉ-MAMORÉ)

*vehpaw to etip*

| Vi=eh-**po**-a-wo | | to | etip |
|---|---|---|---|
| 1PL=wash-CLF:tiny-LK-COP | ART | | manioc.starch |

'We wash the manioc starch.'

<div align="right">(Danielsen 2007, 139)</div>

It is also possible for the classifier to occur both in the incorporated position, and on a modifier in the noun phrase, as is shown here in Tariana, in (46).

(46) TARIANA (NWA)

| pa-thi | dhema-**na** | diha | ha-**na** | heku-**na** |
|---|---|---|---|---|
| IMP-eye | 3sgnf+burn-CL:VERT | ART | this-CL:VERT | tree-CL:VERT |

'This tree is the one that burns one's eyes.'

<div align="right">(Aikhenvald 2003, 107)</div>

Agreement is seen in some form in all of the languages sampled here, showing no areal patterns, and is most commonly of the second or third type, with a classifier on the modifier or verb, but not marked on the head noun.

### 5.2.4. *Functions: Conclusion*

In surveying the range of features available in the Arawakan systems under discussion here, the aim has been to show that these systems share many properties across languages. Above all is the function of agreement, which occurs between the head and any modifying element, although there is variation in whether agreement is also realised on the noun. The derivational possibilities are also widespread, occurring in all languages which mark the classifier on a nominal locus. Anaphoric reference is only possible in languages which have expanded classifier system, and is not described for Resígaro, Yukuna, or Warekena, which are all located in the NWA. None of the features are restricted to one region of Arawakan, reflecting the argument of both Van der Voort (2018) and Krasnoukhova (2012) that this multifunctionality should not be considered a speciality of the NWA, but can be seen throughout the Amazonian region.

It is also clear that function is tied to position, particularly in the use of derivation and agreement; it is likely that these functions developed as part of the development of these positions. However, it is also true that there is a lot of diversity between languages in how they use classifiers: in Resígaro, classifiers have an individuation function which is not seen elsewhere, and Paresi widely uses classifiers to derive nouns from stative verbs, which is also not a function commonly found. This diversity points to the level of change within each language, and most likely the impact of contact on each language.

Having considered the grammatical structures and functions of the Arawakan systems, the next section will consider the semantics and size of the systems, and will argue that the systems share an underlying semantic base relating to physical properties.

## 5.3. The semantics of Arawakan classifiers

In 2, it was presented that classifier systems show underlyingly similar semantics, which tend to correlate to the position they develop in. This section will discuss the characteristics of the classifier systems in both size and semantics.

The classifier sets in the languages included in this study range from less than ten classifiers to over 100, as indicated in Table 9 below.

TABLE 9

NUMBER OF CLASSIFIERS PER LANGUAGE

|  | No: classifiers |
|---|---|
| *Baure* (Danielsen 2007) | 76 |
| *Paresi* (Brandão 2014) | 11 |
| Apurinã (Facundes 2000) | 14 |
| Alto Perené (Mihas 2015) | 28 |
| Yanesha' (Duff-Tripp 1997) | 121 |
| Wapishana (Santos 2006) | 10 |
| Palikur (Aikhenvald and Green 1998) | 55 |
| **Resígaro** (Allin 1976) | 56 |
| **Yukuna** (Lemus Serrano 2020) | 14 |
| **Tariana** (Aikhenvald 2003) | 81-91 [33] |
| **Warekena** (Aikhenvald 1998) | 6-12 |

Given this, the range of meanings included in each set can vary widely: the small set in Warekena includes two classifiers for masculine and feminine humans, and four classifiers which have very different semantics, including one shape classifier, one classifier for time, and two relating to animates – animals and fish. Alternatively, Yanesha' has one of the largest sets, and includes many classifiers relating to physical properties, but also to location, time, or function. Large and small sets are also geographically widespread: Tariana (81-91 classifiers) is located in the NWA, as is Warekena (6-12 classifiers); Baure (76 classifiers) is located in the SWA, as is Paresi (11); one of the largest sets, in Yanesha' (121 classifiers), is located nearby Alto Perené (28 classifiers). This section will review the semantic possibilities of the classifier sets according to five main categories: physical properties, plants, animates, body parts, and other.

### 5.3.1.  *Physical properties*

Classifiers relating to physical properties are available in every language considered here. These physical properties broadly relate to size or shape (large, round, long, etc), but also include texture (liquid, cloth, etc), and dimensions (one dimension, two dimensions, etc). In Yukuna, six of the 14 classifiers relate to size and shape; in Tariana and Yanesha', around 30 classifiers are related to physical properties. However, the category is not well represented in Warekena, which has the smallest number of classifiers and only one related to physical properties: *pau.lia.luni* 'curvilinear objects'. An example of a classifier relating to shape is given below, from Baure, where it is used derivationally to form a new noun.

(47) BAURE (GUAPORÉ-MAMORÉ)
   *kaharopi*
   kaharo-**pi**
   cotton-CLF:long.thin
   'cotton thread'

(Danielsen 2007, 88)

That classifiers relating to physical properties are so widespread in the Arawakan languages is of significance in understanding the development of these systems; returning to the correlations highlighted by Grinevald (2002), a semantic base of physical properties reinforces the proposal that the numeral locus is the oldest in the family.

---

[33] Where a range of numbers is given, this is due to questions of whether some forms consititue one or more classifiers, ie. due to changes in form with different hosts.

### 5.3.2. Plant-based classifiers

Baure, Apurinã, Yanesha', Palikur, Tariana, and Resígaro all include classifiers relating to plants, which broadly relate to generic plant parts: seed, plant, tree, leaf, etc. There are also more specific classifiers: banana (Baure), a new shoot (Resígaro), and low bush (Tariana). The language with the least number of plant-based classifiers is Palikur, which has only two, both of which are glossed as 'plant'. There is some overlap with the category of physical properties, as classifiers indicating 'trunk' or 'stick' may then also be used to refer to things of that shape. An example of a classifier referring to leaves (alongside a classifier referring to shape) is given below.

(48) APURINÃ (BRAZIL)
ãã-**muna**
plant-log.of
'tree'

(Facundes 2000, 209)

### 5.3.3. Body parts

Another major semantic grouping is that of body parts, which occur in Baure, Apurinã, Yanesha', Palikur, Tariana, and Resígaro. These broadly include classifiers for major body parts: leg, face, hand, etc; Yanesha' also includes classifiers for the back of the hand and the soul. Not all classifiers in this category relate to human body parts, as both Apurinã and Baure include classifiers for feathers (in Apurinã the same form is used for both feathers and hair), and in Yanesha' one classifier is used for both human and animal feet. An example of a body part classifier referring metaphorically to something 'leg like' is given below, from Tariana.

(49) TARIANA(NWA)[34]
sidu-kawa
arrow-CL:leg
'the long part of an arrow'

(Aikhenvald 2003, 91)

### 5.3.4. Animates

As already discussed, the division of languages which include classifiers relating to gender shows an interesting divide between the NWA and elsewhere, and although it is not a large semantic grouping in the family, it is included here for that reason. Languages outside of this area also show small numbers of classifiers relating to animate beings: Baure has a generic classifier for humans and another for animals, as well as a classifier for snails. The largest number is in Yanesha', which has classifiers for mice, rats, and worms, as well as generic terms for animals and people. An example of classifiers for animals and humans are given below, from Baure.

(50) BAURE (GUAPORÉ-MAMORÉ)
a.  ti      kove'   eton    aw      te      p-**a**          hir
    DEM1F   dog     woman   and.not DEM1M   other-CLF:animal man
    'This dog is female and the other one is male'

(Danielsen 2007, 388)

---

[34] The gloss for *sidu* is not given, but seems to come from *siduana* 'arrow', given in Aikhenvald (2003, 679).

41

b. *pon mapin etonanev*

| po-**no** | mapi-no | eton-anev |
|---|---|---|
| other-CLF.human | two-CLF:human | woman-hPL |

'The other two women.'

(Danielsen 2007, 171)

### *5.3.5. Other*

The semantic groups discussed above are the semantic groups seen widely throughout the family, however many languages include classifiers relating to other semantic fields, including food and tools, measurements, locations, and time. Classifiers relating to measurements are widely but thinly spread: in Baure and Apurinã there are only one or two classifiers relating to small measurements ('a handful', 'a spot of'), while in Tariana there are at least eight, which mostly relate to gatherings of objects: 'piles' or 'heaps'. Several languages also have classifiers relating to functional tools, often denoting things in everyday use, such as housing, transport, or cooking implements. The most restricted categories are time and location classifiers: these occur only in Yanesha', Palikur, Tariana, and Resígaro, where the location classifiers form the largest subgroup.

## 5.4. Conclusion

This chapter has presented the grammatical structures and functions of the sampled classifier systems, as well as shown the semantic areas underlying the systems. Based on distribution, the numeral, nominal, and verbal positions can be reconstructed to PA with the same functions they are used for today. The widespread occurrence of the numeral locus points to this position as the oldest classifier position, something further evidenced by the widespread use of classifiers relating to physical properties. However, it is also evident that the systems diversified after the divergence of PA, seen in the development of the modifier classifier position and in the varying realisations of the derivational, agreement, and anaphoric functions. This gives insight into the questions put forward in 1: not only can the characteristics of the NWA be shown to be present more widely throughout the family, but that these positions can be reconstructed points to a shared source at the level of PA. The realisation of this system will be discussed further in 7; whether any classifiers of the synchronic systems can be shown to have taken part in this system will be the focus of the next chapter.

# 6. Cognates

## 6.1. Introduction

The investigation into cognate sets among the Arawakan languages yielded 12 sets, discussed in sections **6.2-6.9**, with meanings related to shape, physical material, and time. These sets will be discussed under eight headings: long, round, flat (one-dimensional), containers, liquid, powder, cloth, and time. Of these, several are associated with multiple forms, while others have multiple distinct sets which may themselves be related. Of the 12 sets, eight are found spanning four or more languages, and only five include five or more languages.

An attempt has been made at reconstructing shared ancestral forms, by which I propose that the form was in use in the PA system discussed in the previous chapter. The number of languages supporting each set is small; in order to maximise evidence that reconstructed forms date from PA and are not independent parallel developments, reconstructions are supported with both shared phonology and meaning across at least three languages.

Each set gives information of genealogical relation within the family, language specific form, language specific gloss, how the form is categorised within the source material, and the reconstructed ancestral form.

## 6.2. Long: *pi, *kig, *na, *pa/ap (Sets 1-4)

The following forms are found with meanings related to length and large size: *pi, *kig, *na, *pa/ap.

### 6.2.1. Set 1: *pi

Set 1 consists of forms from five languages, with only one from NA. The meanings in each language are very similar, although Paresi differs slightly in not specifying characteristics of length and thinness, but of a natural element which has these characteristics; Alto Perené has been included as the meaning of the form 'rigid' relates to the meaning pairing 'long-rigid' given by Adams and Conklin (1973) as a commonly found meaning in classifiers.  Four languages have forms beginning with /p/ and one, Paresi, has a form beginning with /h/. These forms can be reconstructed as *pi; the change from */p/ > /h/ seen in Paresi is a regular sound change, discussed in Carvalho (2016a).

TABLE 10
SET 1: *pi 'LONG AND THIN'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| *Baure* | SSWA: South Arawak | -pi | Long and thin | classifier |
| Yanesha' | SSWA: Amuesha | -Vp̃, -p̃-, -p- | Thin and long things | classifier |
| **Tariana** | NA: North-Amazonian (Upper Rio Negro) | -pi | Long and thin | classifier |
| *Paresi* | SSWA: Paresi-Xingu | -hi | Vine-like | classifier, compound |
| Alto Perené | SSWA: Kampa | -pi | Rigid, hollow | classifier |

The Baure and Paresi forms here were first proposed as cognate by Payne (1991, 383–84), who gives them as reflexes of the form *apʰi 'snake'. If this lexical noun is cognate, it demonstrates grammaticalisation, showing semantic widening, decategorialisation, and phonological erosion. Aikhenvald (2019, 118) also notes cognacy with forms in Baniwa of Içana, Cabiyari, Waurá, Terêna, and Ignaciano, languages from both North and South Arawakan.

A morpheme of the shape *pi* is found throughout the languages surveyed here, in both Alto Perené and in Tariana. In Tariana, it occurs in –*napi* 'bone' and –*pida* 'crown and branches of tree'; in Alto Perené, it is in the form –*impiki* 'pointed'. It also occurs in Resígaro in the form -*ʔaapí* 'arm'. In other cases, the same shape occurs in forms where it is harder to make the connection with length: *api* 'under' in Baure, *depita* 'night' and *hipita* 'land' in Tariana. The occurrence of *pi* in -*ʔaapí* in Resígaro suggests that a classifier of this shape may have been in use in this language at an earlier stage, but has since been lost, demonstrating that it has a wider spread than in just the languages presented above.

### 6.2.2. Set 2: *kig

Set 2 includes three forms across languages of both SSWA and NA. There is a wider variation in meaning; in Alto Perené and Palikur, the meanings are given as 'pointed', while in Tariana the meaning relates to thin and curved objects. There is also variation in form, as although in all languages the form has an initial voiceless velar stop (which is aspirated in Tariana) and a high front vowel which varies in length, in Palikur it has a consonantal coda, *g*, and in some forms a second syllable, -*ya*. Despite this, it is possible to reconstruct an earlier form, *kig*. Evidence that the earlier form is *kig* rather than *ki* comes from work on the phonology of Proto-Kampa, of which Alto Perené is descended. Michael et al. (2010) argue that */g/ > */y/ in all environments in Ashéninka. Mihas (2015) further argues that palatal glides do not occur in final position in Alto Perené (part of the Ashéninka dialect continuum); it is possible that the representation given below is the result of a different analysis, which has analysed palatal glides in this position as vocalic rather than consonantal, or that the palatal glide has assimilated to the preceding high front vowel given the relative closeness in pronunciation. As this distinction does not greatly affect the reconstruction, it will not be discussed further here. There is no historical phonological work on Tariana, therefore it is unclear at this point if the loss of */g/ in this form is part of a regular process.

TABLE 11
SET 2: *kig* 'POINTED'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| Alto Perené | SSWA: Kampa | -*kii* | pointed | classifier |
| Palikur | NA: Palikur | -*kiya, -kig* | pointed | verbal classifier |
| **Tariana** | NA: North-Amazonian (Upper Rio Negro) | -*khi, -ki* | Thin, curved | classifier |

This form is also seen in another complex forms in Alto Perené, already discussed in **6.2.1**, -*impiki* 'pointed'. Although this cognate set only includes three languages, there is no clear areal or subfamily pattern to where it is found.

### 6.2.3. Set 3: *na

Set 3 comprises three classifiers which all have meanings relating to large size: in both Alto Perené and Tariana, the meanings relate to length, and in Yukuna the meaning is 'big'. Once again this set does not form an areal or subfamily pattern. There is no variation in form, therefore it is easily possible to reconstruct a proto form of the same shape and to propose that this form was in use in an earlier stage in the language.

A possible related form is also seen in Yanesha', -*n*- 'canoe', which may have developed from the meaning related to length. However, the form has further meanings 'animal, human' and ultimately this connection is too tenuous to include here.

TABLE 12
SET 3: *na 'LONG'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| Alto Perené | SSWA: Kampa | *-na* | 'long hollow' | classifier |
| **Yukuna** | NA: North-Amazonian (Colombian) | *-na* | 'big' | classifier |
| **Tariana** | NA: North-Amazonian (Upper Rio Negro) | *-na* | 'long vertical' | classifier |

Although Set 3 only consists of three members, it occurs in a number of forms in Paresi, Wapishana, and Apurinã, giving it a wider distribution than presented in this cognate set. Aikhenvald (2019, 118) also notes possible cognates in Baniwa of Içana, Achagua, Bahuwana, all in the NWA, and Yawalapiti. It is also found in further forms in Tariana and Alto Perené, suggesting that its development predates other classifiers in these sets. These complexities will be discussed here.

In Paresi, it occurs in the form *natse* 'long horizontally, cylindrical, three-dimension', where it is seen in combination with the Paresi classifier *tse* 'small', although how the meaning is arrived at is unclear.[35] In Apurinã, the form *muna* 'trunk; long; cylindrical' likely also includes this form, although it is unclear what the source of the initial segment is. In both Alto Perené and Tariana, it is found in combination with the Set 2 classifier *-ki*, in the form *naki*. In Alto Perené this has the meaning 'hollow and round', and in Tariana it has the meaning 'long handle'; due to the distance between these languages, there is no reason to think that these forms are themselves related, but more likely independent innovations from the same material. Finally, both Wapishana and Tariana have the form *puna*, meaning 'extension' in Wapishana and 'road' in Tariana. Although these languages are not closely related within the family, the semantics would suggest that these forms could be related, although again at this point it is unclear what the precise composition of the form is.

As in Set 2, the form *-na* occurs in only a small set of languages in its simple form, but with a greater distribution when complex forms are considered, demonstrating the possibility that this form existed at an earlier stage in these languages but was later lost.

### 6.2.4. Set 4: *pa/ap

Set 4 is discussed here as 4a and 4b due to the relatedness of their forms, which are *pa('ll)* and *ap*. I will consider first their semantic relationship before turning to their phonological characteristics.

Both sets include the meanings 'banana', 'extension', 'cylindrical' or 'concave', and 'long'. These meanings have overlapping features of length and curvature, and likely derive from the same source. For this reason, it is possible to reconstruct the meaning 'long and curved' to an earlier stage of the language family.

Returning to the phonological shape, the forms have been divided as CV or VC. Yanesha' includes both forms, although in one form it is given with a final palatized lateral consonant. It is well documented that metathesis is a widespread process in Yanesha' which obscures cognancy (Payne 1991, 430-431), although the process has not been shown to be more widespread in the Arawakan family, and the spread of languages does not indicate the change happening in one subfamily. It is also possible that this change happened in Yanesha' and spread through contact, although this is not supported by the areal distribution of the languages. Ultimately, the spread of this form will require further investigation. At this point, a classifier with the meaning 'long and curved' is tentatively proposed as being in use at an earlier stage in the language, although it is unclear if the form should be reconstructed as *pa* or *ap* . A map of the spread of these forms is given below the table.

---

[35] In Paresi, it also occurs in two compound forms, *mena* 'stalk of' and *hana* 'leaf of', which will not be discussed here.

TABLE 13

SET 4: *pa, *ap 'LONG AND CURVED'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| 4a | | | | |
| Yanesha' | SSWA: Amuesha | *-ap-* | long bone-shaped, banana | classifier |
| Wapishana | NA: Rio Branco | *ap* | extension | classifier |
| Palikur | NA: Palikur | *-ap* | concave | verbal classifier |
| 4b.[36] | | | | |
| Alto Perené | SSWA: Kampa | *-pa* | cylindrical | classifier |
| **Yanesha'**[37] | SSWA: Amuesha | *-pa'll-* /paʔʂ/ | banana shape | classifier |
| **Tariana** | NA: North-Amazonian (Upper Rio Negro) | *-pa* | largish and long | classifier |

FIGURE 9

AREAL DISTRIBUTION OF 'LONG AND CURVED' CLASSIFIER.[38]



---
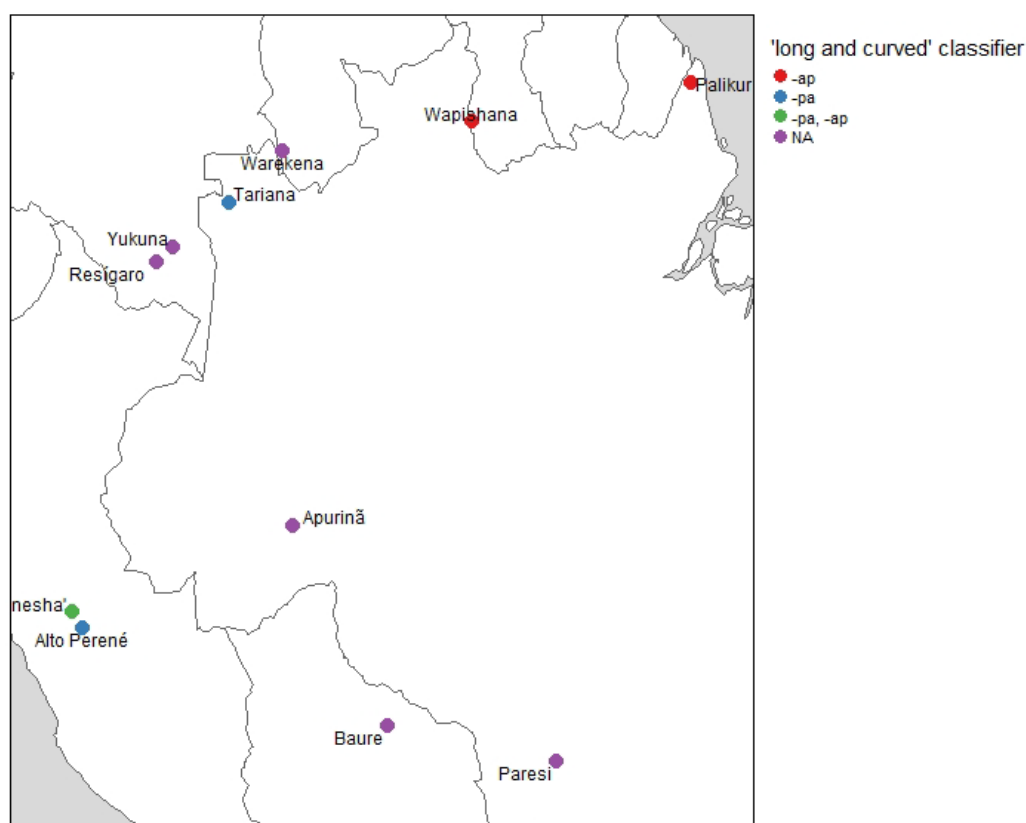
[36] This set originally included the Baure form *-poe* 'banana', however this is a repeater from the lexical noun *erapoe'* 'plantain' (Danielsen 2007, 446). Given the difference in shape, although it may be shown to derive from the same source as the other classifiers, it is most likely an independent development.

[37] Yanesha' includes a number of words for bananas which include this form: *empa'lla* 'plátano verde' (Wise 1998, 178), *parantapa'll* 'plátano chico muy verde' (Wise 1998, 276). It is unclear if these forms have incorporated the classifier, or

Although this form is not particularly widespread beyond these languages, an element of the same shape is seen in Tariana with a meaning related to curvature, in -*turapa* 'cone', and in Yukuna with a seemingly unrelated meaning: -*a'pa* 'part of whole'. [39]

## 6.3. Round: *sV* (Set 5)

Throughout the languages surveyed, many classifiers were found with meanings related to roundedness, however ultimately very few form coherent sets, either because too few forms were found (Tariana and Wapishana -*da* and -*q̇a* respectively, both meaning 'round'; Alto Perené -*ina* meaning 'spherical' and Tariana -*wina* meaning 'pile'), or because the meanings are inconsistent.

Set 5 includes classifiers from three languages, two from NA (Palikur, Resígaro) and one from SSWA (Baure). The forms show similarity in meaning, relating to shapes, and some similarity in phonological form, occurring with an initial /s/ in both Baure and Palikur, and /o/ in Palikur and Resígaro.

However, the set is thrown into question, as the Resígaro form is suggested by Seifart (2011) to be a loan from Bora, and indeed the shift from /s/ to /ʔ/ is not attested in Payne (1991). However, a shift of */s/ > /ʔ/ is also seen in another cognate set presented later in this section (Set 9), in which PA *sa* > Resígaro *ʔa*. Furthermore, the consistency in meaning between all three forms is enough to warrant inclusion here. An additional issue is the vowel split between SSWA (Baure) and NA (Palikur and Resígaro); although there has been study of the historical phonology of Baure at the subfamily level (Jolkesky 2016), it does not make reference to this vowel change.

TABLE 14

SET 5: *sV* 'OVAL'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| *Baure* | SSWA: South Arawak | -*se* | oval | classifier |
| Palikur | NA: Palikur | -*so* | equal dimensions | numeral classifier |
| Resígaro | NA: North-Amazonian (Colombian) | -*ʔo* | long, oval | classifier |

Given both of the issues discussed above, the classifier *sV* is tentatively reconstructed for PA, however further research will be required to see if these changes can ultimately be shown to be more widely spread.

## 6.4. 1D (flat): *ta, *kwa (Sets 6-7)

In this section, two sets are presented with meanings related to flatness and one dimensionality.

### 6.4.1. Set 6: *ta

Set 6 is largely uniform in both its form and meaning: it has a shared form in Alto Perené, Palikur, and Yukuna, and a similar form in Yanesha' and Wapishana. As there is very little variation in both form and meaning, it is possible to propose a reconstructed form *ta*. [40]

---

if the classifier is derived from these forms. Particularly in the case of the latter form, it is likely that *pa'll* is acting as a classifier here.

[38] Map generated with Glottospace (Norder et al. 2022).

[39] Aikhenvald (2019, 119) also connects this with a form in Apurinã, however no such form was given by Facundes (2000).

[40] Aikhenvald (2019, 118) connects the Yukuna form with a number of classifiers with the form -*da*, generally meaning 'round'; this is found in Tariana.

TABLE 15
SET 6: *ta '1D (FLAT)'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| Alto Perené | SSWA: Kampa | *-ta* | flat, vertical | classifier |
| **Yanesha'** | SSWA: Amuesha | *-tall-* /taɬ/ | flattened things, bark | classifier |
| Palikur | NA: Palikur | *-ta-* | vertical | numeral classifier |
| **Yukuna** | NA: North-Amazonian (Colombian) | *-ta* | flat | classifier |
| Wapishana | NA: Rio Branco | *t* | long | compound |

Two additional languages show the same form in other classifiers: Paresi has the form *taotse* 'flat surface, one dimensional', and Apurinã has the form *tãta* 'bark; shell; flat; thick'. It is also of significance that both Apurinã and Yanesha' are extended to include natural elements such as tree bark, a cross-linguistically common source of classifiers of this type (Adams and Conklin 1973). Tariana also shows use of this form, however it is unclear if its use is semantically related: it occurs in a number of forms relating to space, such as *patawai* 'gorge', *pusita* 'clearing' and *hipita* 'land'. As well as the languages presented here, this evidence suggests that it was once also in use in at least Paresi and Apurinã, but has since been lost.

### 6.4.2. Set 7: *kwa

Set 7 is a smaller set, but which also shows a lot of unity in meaning, and some unity in form. Although there is little historical reconstruction of these languages, the variation in form can be understood as a loss, in Palikur, of a medial glide which has been represented consonantally in Tariana and vocalically in Paresi. For this reason, a reconstruction can be proposed as *kwa*, with a meaning '1D (flat)'.

TABLE 16
SET 7: *kwa* '1D FLAT'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| *Paresi* | SSWA: Paresi-Xingu | *-koa* | one-dimensional | classifier |
| Palikur | NA: Palikur | *-ka* | flat | numeral classifier |
| **Tariana** | NA: North-Amazonian (Upper Rio Negro) | *-kwa* | flat surface | classifier |

Aikhenvald (2019, 118) gives cognates in Baniwa of Içana, Achagua, Cabiyari, Yawalapiti, and Waurá, and includes *-a'aku* 'concave' (Yukuna) and *-koa* 'round hollow container' (Apurinã), which were not included here, based on differing semantics.

### 6.5. Container: *Vko (Set 8)

Set 8 is discussed as two sets, 8a and 8b, based on six languages with classifiers of forms varying on *Vko*. In Baure, Paresi, Alto Perené, and Resígaro, languages from both SSWA and from NA, the form occurs with a mid to low initial vowel; in Apurinã and Palikur, which are again from both sides of the higher-level split in the family, the form occurs with a high initial vowel.

Despite the difference in the initial vowel, the same source can be shown for both: Payne (1991, 416) reconstructs the form *kopi(tʰi)* for 'Pot (Ceramic Cooking Pot)', and it is highly likely that the *-ko* element in Set 8 comes from this form. However, at this point it is unclear whether these variations were independent developments of the form *-ko*, or there was a shift from *-ako* to *-iko* (or vice versa). This form can be shown to result from grammaticalisation, showing semantic widening of cooking pot to the more general container,

and phonological reduction in the loss of the second syllable(s), although the classifier form has also gained an initial vowel.

TABLE 17
SET 8: *ako 'CONTAINER'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| **8a** | | | | |
| *Baure* | SSWA: South Arawak | *-eko* | jug | classifier |
| *Paresi* | SSWA: Paresi-Xingu | *-ako* | inside of a hollow, three-dimensional | classifier |
| Alto Perené | SSWA: Kampa | *-ako* | container | classifier |
| **Resígaro**.[41] | NA: North-Amazonian (Colombian) | *-ʔoohú* | metal or tin container | classifier |
| **8b** | | - | | |
| Apurinã | SSWA: South-Western Arawak | *riko* | hole | classifier |
| Palikur | NA: Palikur | *-iku* | within a periphery | numeral classifier |

Starting with the meanings given above, there are two broad meanings, which largely correspond to the divergence in the initial vowel. The forms in Set 8a with the initial mid to low vowel have meanings relating to containers, retaining the functional meaning of the referent (excepting Paresi), and the Set 8b forms with the initial *i-* (but including Paresi) have meanings relating to spatial dimensions. These meanings could be related: 'within a periphery' also has notions of containment, and this meaning is shared by Paresi in Set 8a, where the meaning also does not relate to a functional tool, but to spatial dimensions.[42]

There are some issues to be addressed in relation to the phonological form, particularly the Resígaro form. It is the first syllable, *-ʔo*, which is potentially cognate, however a change from /k/ > /ʔ/ is not attested in Payne (1991). The lengthening of the vowel, however, is expected: Payne (1991) cites Resígaro as having 35 forms with unexplained long vowel innovations. As already noted, the divergence between the two sets is in the initial vowel, although there are also some small differences within each set: the initial vowel in Baure is higher than in the other languages, suggesting a change of /a/ > /e/, and the form in Resígaro has lost the initial vowel. In Set 8b, Apurinã has an initial consonant *r-*, and there is variation in the final vowel of Palikur, which has *u*.

For Set 8a, *ako* can be reconstructed based on shared semantics and phonological similarity, although this would be supported with further reconstructive work on the historical phonology of Resígaro. However, it remains unclear how exactly the forms in Set 8b are related to the set above; further investigation may show whether the form is a further evolution of *ako* or an independent innovation from *kopi(tʰi)*. If these sets have developed independently from each other, the occurrence of reflexes of *kopi(tʰi)* throughout the family may be worthy of further study.

---

[41] A second Resígaro form, *-ʔíǰó* 'earthenware container for liquid' was considered for inclusion here, however this is a loan from Bora (Seifart 2011).

[42] The overlap between function and shape with particular relation to meanings such as 'container' is discussed by Grinevald (2015).

### 6.6. Liquid: *sa (Set 9)

Set 9 presents forms relating to liquid, found in seven languages and discussed as two subsets. It is based on a set proposed by Payne (1991, 425), which gives *isa[pʰa] as a form meaning 'wet' with reflexes in Chamicuro, Wapishana, Palikur, Achagua, Piapoco, and Yukuna, however, only in Wapishana does this form occur as a classifier.

The forms in this set show two paths of development: those which have kept an initial /s, z/ (Baure, Paresi, Yanesha', Wapishana), which themselves show a distinction in voicing, and those which haven't (Apurinã, Alto Perené, Resígaro). Again, there is no areal pattern in these two groups, as shown in the map below (Figure 10, below table). A note is also necessary on the form in Yanesha': as mentioned previously, metathesis of both CV segments and entire syllable sequences is reported to be very common in the language (Payne 1991, see also Wise, 1976), and for this reason when considering the form -Vs, I have considered it as patterning with Set 9a.

The form from Resígaro is included in Set 9b, however it includes an additional element after the root and an initial glottal stop. Although this is not reported in historical reconstructions of Resígaro (Carvalho 2018), it is also not reported to be a loan, and the same phonological shift of */s/ > /ʔ/ is seen in Set 4 (although that form may be a loan).

TABLE 18
SET 9: *sa 'LIQUID'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| 9a | | | | |
| *Baure* | SSWA: South Arawak | *-sa* | water | classifier |
| *Paresi* | SSWA: Paresi-Xingu | *-za* | liquid; speech | classifier |
| Yanesha' | SSWA: Amuesha | *-Vs* | water, river | classifier |
| Wapishana | NA: Rio Branco | *ʐa, (i)ʐ* | damp, liquid | compound, classifier |
| 9b | | | | |
| Apurinã | SSWA: South-Western Arawak | *-ã̃* | water, juice, tear, liquid | classifier |
| Alto Perené | SSWA: Kampa | *-aa* | liquid | classifier |
| **Resígaro** | NA: North-Amazonian (Colombian) | *- ʔaanú* | liquid | classifier |

FIGURE 10
AREAL DISTRIBUTION OF 'LIQUID' CLASSIFIER.[43]
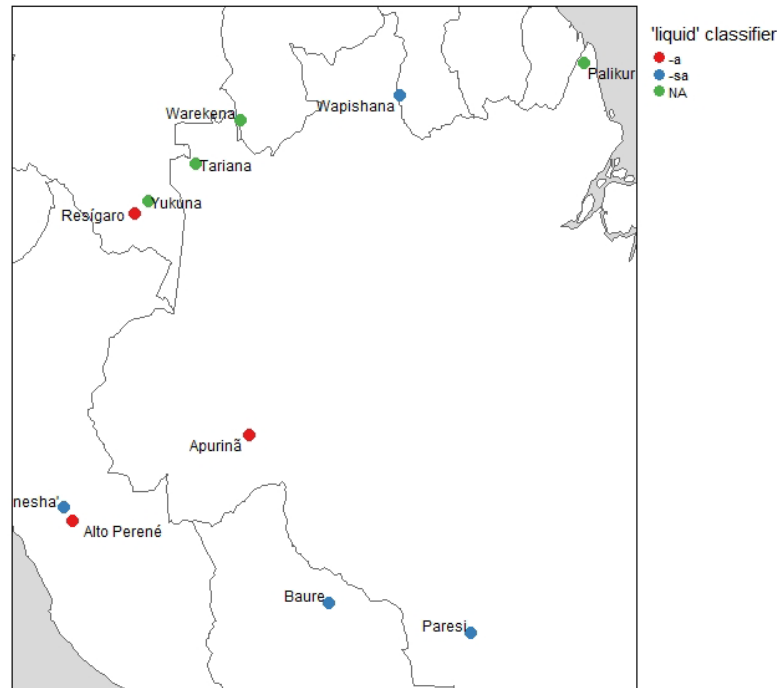


Regarding the voicing split seen in Set 9a, the decision to reconstruct the unvoiced form, *sa, is based on the inclusion of */s/ (but not */z/) in Payne's (1991) inventory of PA. The patterning of the loss of this phoneme does not show a clear areal pattern, and it is possible that this happened independently in a number of languages. Carvalho (2021) discusses the patterning of the Apurinã classifier with its cognate in Iñapari, a closely related language, and notes that it is likely that in Proto-Puru the form was *ha. Given this, data from a wider spread of languages may shed light on an early shift of */s/ > */h/ or */ø/, as well as the possible Resígaro shift of */s/ > /ʔ/.

As already mentioned, this form is likely derived from PA *isa[pʰa] 'wet', and therefore it shows the initial stages of grammaticalisation, in the loss of the initial vowel and the final syllable, although it does not show the same level of semantic widening which was seen in the other forms which can be connected to PA lexical nouns.

### 6.7.  Powder: *panhi (Set 10)

Set 10 presents classifiers from five languages relating to powder and dust. It is discussed as two subsets, based on a difference in phonological development. A map of this division is given below the table.

---

[43] Map generated using Glottospace (Norder et al. 2022).

TABLE 19

SET 10: *panhi 'POWDER'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|----------|-------------------|------|-------|------|
| 10a | | | | |
| Apurinã | SSWA: South-Western Arawak | *-panhi* | powder | classifier |
| Alto Perené | SSWA: Kampa | *-pani* | powder | compound |
| Yanesha' | SSWA: Amuesha | *-pueñ-/-Vpñ- /pʷenʲ/* | ash, dust | classifier |
| 10b.[44] | | | | |
| *Paresi* | SSWA: Paresi-Xingu | *he* | powder | classifier |
| **Tariana** | NA: North-Amazonian (Upper Rio Negro) | *-hwi* | particles | classifier |

FIGURE 11

AREAL DISTRIBUTION OF 'POWDER' CLASSIFIER [45]
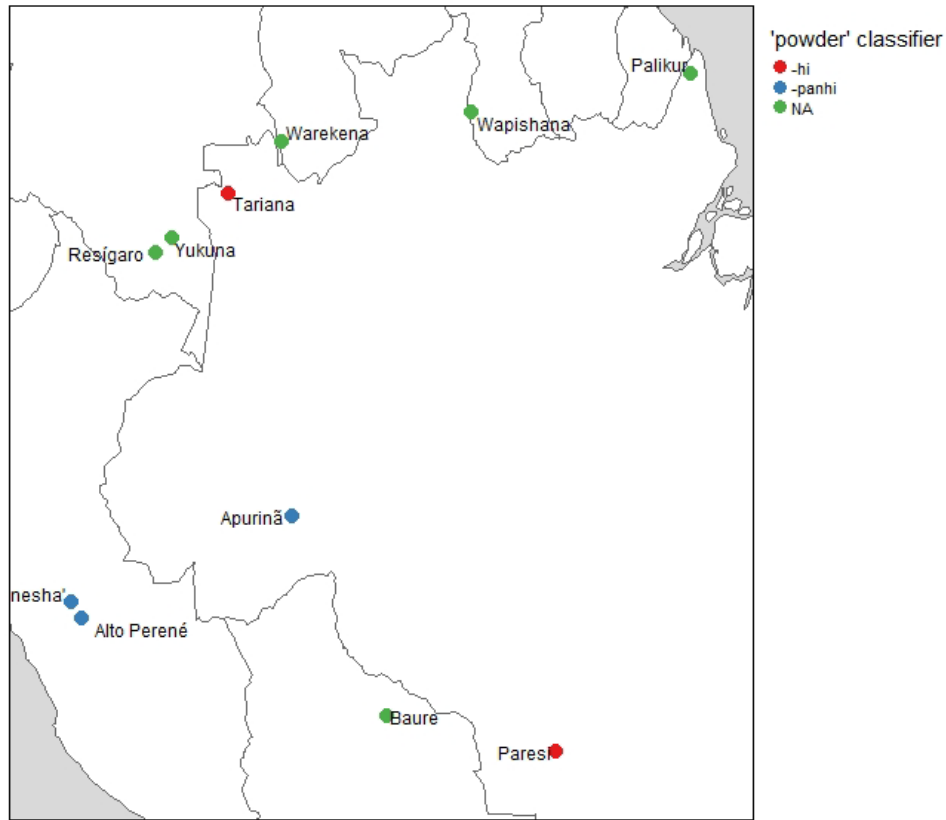


Set 10 is based on a PA form proposed by Payne (1991, 401): *pʰa/e[ne] 'dust, ash', which includes Yanesha', Chamicuro, Paresi, Waurá, Terena, Baure, Ignaciano, Piro, Apurinã, Machiguenga, Ashéninka, Palikur, Garifuna, and Yukuna, languages from across the family. Although this includes languages from this paper's sample, in some of these the reflex occurs as a lexical noun rather than a classifier. In Payne's reconstruction, all languages (except Paresi) include the initial segment *pV*, but only Yanesha', Piro, Apurinã, Machiguenga, Ashéninka, and Palikur also include a second syllable, which forms the basis of Set 10a.

Turning to the classifiers from this sample, the three languages in Set 10a come from the SSWA branch of Arawakan, demonstrating an areal pattern. Set 10b consists of one language from the SSWA branch, and one from NA. The two sets are given to be related based on a shared meaning and the overlap in shape of the

---

[44] Set 14b initially included -ǰíhú 'dust' (Resígaro), however this is a loan from Bora (Seifart 2011).

[45] Map generated using Glottospace (Norder et al. 2022).

syllable (*h*)*i*, but are distinguished on the basis of their phonological development; the languages in Set 10b have lost the initial segment *pan*. Based on the cognacy with the form *p$^h$a/e[ne]* discussed above, it is assumed that *panhi* is the older form, and the forms in Set 10b are innovations. It is likely that the forms in Paresi and Tariana are two independent innovations given their geographical and genetic distance. The remaining differences in form, such as the loss of aspiration in Alto Perené and in Yanesha', the development of labialisation in Tariana, and the development of */i/ > /e/ in Paresi would need to be explained with historical evidence from these languages. Despite these small variations in form, I have reconstructed the form *panhi* on the basis of the data from both sets.

This reconstructed form also shows some signs of grammaticalisation, although in its reconstructed form it does not show phonological erosion, as was argued earlier for *pi*, *ako*, and *sa*. However, it shows some semantic widening, in the shift from a specific to a general type of powder.

### 6.8.  Cloth: *ma* (Set 11)

Set 11 includes the forms *mo* and *ma*, which have meanings related to woven material. The vowel split shows an areal divide, as the /o/ form is found only in the SSWA languages, and the /a/ form is found only in the Northern languages. In Yukuna, the form is included but has the meaning 'whole', demonstrating semantic shift. It is possible that the form in Tariana is a complex classifier, including the form -*ka*, which was seen in Set 7, meaning 'flat'.

TABLE 20

SET 11: *ma* 'CLOTH'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|----------|-------------------|------|-------|------|
| *Baure* | SSWA: South Arawak | -*mo* | woven | classifier |
| Yanesha' | SSWA: Amuesha | -*amo*-/-*om*- | cloth | classifier |
| **Yukuna** | NA: North-Amazonian (Colombian) | -*a'ma* | whole | classifier |
| **Tariana** | NA: North-Amazonian (Upper Rio Negro) | -*máka* | extended cloth | classifier |

The form is also seen in two other languages: in Alto Perené it is part of the form -*mashi* 'flat open space', and in Palikur it is in the form -*madka* 'flat'. The Alto Perené form may also include the classifier -*shi* 'leaf-like'. In both cases, the extension of the meaning 'woven material' to the meaning 'flat' is not surprising. Given the appearance of /ma/ in languages beyond the Northern Arawakan subfamily, as in the forms in Alto Perené and Palikur, it is possible to tentatively propose that this vowel form is the older form. This is supported by two reconstructions in Payne (1991): both *idama* and *mata* are reconstructed for 'skin' (and are possibly the same form undergoing metathesis, see Payne (1991, 418)), and given the semantic link with flatness and one-dimensionality, it is possible that these are the source nouns for this classifier, or that these forms have incorporated the classifier. However, it would be necessary to support this with data from more languages.

### 6.9.  Time: *chi* (Set 12)

The set of classifiers presented in this set are connected by similarity in form, and meanings relating to the moon and time. No similar forms are given by Payne (1991), although possibilities of a lexical source are seen in some languages. In Paresi, *mitxini /mitʃini/* 'month' is found (Brandão 2014, 104), while in the NA languages, words are found with the morpheme *ke*: *keegi* 'month, moon' (Resígar: Allin 1976, 501), *keri* 'moon' (Yukuna: Lemus Serrano 2020, 69), *keeri* 'month' (Tariana: Aikhenvald 2003, 90), *queeri* 'month' (Piapoco: Klumpp 2019, 49). In Baure, the lexical noun 'moon' is *kiher* (Danielsen 2007, 28), and in Alto Perené it is *kashiri* (Mihas 2015, 48). The forms in Set 12 show differences in the vowel, however they also differ from the lexical nouns in the relevant languages, suggesting that they are not recent developments. A tentative reconstruction is proposed as *chi* 'moon'.

TABLE 21

SET 12: *chi 'MONTH, MOON'

| Language | Aikhenvald (1999) | Form | Gloss | Type |
|---|---|---|---|---|
| *Baure* | SSWA: South Arawak | *-či /chi/* | month | classifier |
| Alto Perené | SSWA: Kampa | *-chee* | crescent shaped | compound |
| Palikur | NA: Palikur | *-i/-i-* | series (time) | numeral classifier |

### 6.10. Reconstructions

Sections 6.2-6.9 presented 12 cognate sets for which PA reconstructions can be proposed, given below in Table 22. Although some of these forms are included in Payne's (1991) reconstructions alongside lexical nouns and Aikhenvald (2019, 118–19) has discussed the forms of some sets in relation to languages of the NWA, this is the first time that these cognates have been recognised between the classifier systems across the Arawakan family. These reconstructed forms are proposed as forms which were in use in a grammatical structure which predated the current classifier systems; it is likely that these forms were grammaticalised at the level of PA. This argument, and what this means for a PA classifier system, will be expanded in the following chapter. What is significant is that these forms have reflexes throughout the classifier systems of Arawakan, showing family-internal historical connections between the synchronic classifier systems.

TABLE 22

PROPOSED RECONSTRUCTIONS

|  | Proposed reconstructed form | Proposed reconstructed meaning | Language | # langs |
|---|---|---|---|---|
| Set 1 | *pi* | long and thin | *Baure*, Yanesha', **Tariana**, Paresi, Alto Perené | 5 |
| Set 2 | *kig* | pointed | Alto Perené, Palikur, **Tariana** | 3 |
| Set 3 | *na* | large, long | Alto Perené, **Yukuna**, **Tariana** | 3 |
| Set 4 | *pa/ap* | long and curved | Yanesha', Wapishana, Palikur, Alto Perené, **Tariana** | 5 |
| Set 5 | *sV* | oval | *Baure*, Palikur, **Resígaro** | 3 |
| Set 6 | *ta* | 1D (flat) | Alto Perené, Yanesha', Palikur, **Yukuna**, Wapishana | 5 |
| Set 7 | *kwa* | 1D (flat) | *Paresi*, Palikur, **Tariana** | 4 |
| Set 8a | *ako* | container | *Baure*, *Paresi*, Alto Perené, **Resígaro** | 4 |
| Set 9 | *sa* | liquid | *Baure*, *Paresi*, Yanesha', Wapishana, Apurinã, Alto Perené, **Resígaro** | 7 |
| Set 10 | *panhi* | powder | Apurinã, Alto Perené, Yanesha', *Paresi*, **Tariana** | 5 |
| Set 11 | *ma* | cloth | *Baure*, Yanesha', **Yukuna**, **Tariana** | 4 |
| Set 12 | *chi* | moon | **Baure**, Alto Perené, Palikur | 3 |

Although most of the cognates show a cross-family range, often being found in languages across a number of branches and across the high-level north-south division, most of them are supported by only a few languages. One classifier, *sa*, is found in seven languages, but only four are found in five languages, and all others are found in less. However, that these classifiers can be reconstructed across these branches with both

phonological and semantic evidences lends support to the argument that these forms developed before the divisions of these subfamilies.

Four of the forms can also be linked to PA lexical nouns, and these reconstructions show evidence of grammaticalisation, either in phonological erosion in the proposed reconstruction, or in the semantic widening of the form. These four are: *pi* 'long and thin'(*ap^hi* 'snake'), *ako* 'container' (*kopi(t^hi)* 'ceramic cooking pot'), *sa* 'liquid' (*isa[p^ha]* 'wet'), *panhi* 'powder' (*p^ha/e[ne]* 'dust, ash').[46]

The semantics of the forms found cover a range of meanings relating to shape (long, pointed, oval, curved), but also to texture, consistency, and material (liquid, powder, cloth). These last three forms are some of the more widespread forms, occurring in seven, five, and four languages respectively, while there is more variety in the forms relating to shape. It has already been discussed that semantics relating to physical form underpin the Arawakan classifier systems, and this is again supported by the results shown here. That there are a number of classifiers relating to the same shapes, particularly those three shapes which were argued to be so key to shape-based classifier systems (long-rigid, flat-flexible, round), is a particularly interesting finding, and may point to repeated processes of grammaticalisation.

Some notes are necessary on individual languages. None of the reconstructed forms in Table 22 have reflexes in Warekena, however the evidence of 5 showed that Warekena has a much-reduced system, and the lack of cognates does not prevent proposing that it came from a PA source. Apurinã, similarly, has a reduced system, having lost the numeral position, however it's appearance in two of the sets here (Set 9, *sa* and Set 10, *panhi*) evidences that it too evolved from the same PA system. Palikur, which has different classifier forms with different loci, has reflexes which occur both with a numeral locus and with a verbal locus: Sets 2 and 4 (*kig* 'pointed', *pa/ap* 'long and curved') have reflexes in Palikur which are verbal classifiers, and Sets 5, 6, and 7 (*sV* 'oval', *ta* 'flat', *kwa* 'flat') have reflexes in Palikur which are numeral classifiers. If the sets are now distinct systems, as argued by Aikhenvald and Green (1998), this would indicate that they began as one system and the division evolved later.

The reconstructed forms presented here were based on shared similarity in form and meaning across at least three languages; a number of other potential cognates were found for which a reconstruction ultimately could not be proposed due to a variety of reasons, such as lacking evidence of a shared origin either in meaning or form, or in the weight of evidence for an alternative argument of development. The classifier *tsa* 'long' is found in both Alto Perené and Apurinã, with a possible cognate *tse* 'small' in Paresi, however, the form in Paresi has a lexical cognate within the language which is a more likely source. Similarly, the forms -*eši* /eʃi/ 'piece' (Baure) and -(*i*)*tʃi* 'bundle' (Tariana) are likely cognate, and a reconstruction could have been proposed on the basis of a third form, -*ʔí* 'bunch' (Resígaro), however this is a loan from Bora (Witotoan: Seifart 2011). The forms *ku* 'rounded' (Apurinã), *mku* 'concave' (Palikur), -*ú* 'spherical' (Resígaro), and *a'ku* 'concave' (Yukuna) share overlapping semantics and a common element (*k*)*u*, however the Resígaro form is known to be a loan from Bora (Seifart 2011), and as yet the lack of phonological evidence does not support proposing these forms as cognate. These classifiers, as well as others for which a reconstruction could not be proposed, demonstrate that a wider study of more languages may well reveal more cognates.

This chapter has presented the evidence of cognate sets between the shape-based classifiers of the Arawakan classifier systems, and shown that a small number of them can be reconstructed to PA. The following chapter will discuss this in relation to the evidence which was presented in 5, and make a final proposal as to what the system looked like at the level of PA.

---

[46] Payne (1991) uses [] to indicate material which synchronically or diachronically has morphemic status, and () to indicate that a source indicates distinct morphemic status for this segment. See Payne (1991, 390) for more details.

## 7. Discussion

In 1, I set out the research question and four investigative aims, which are repeated below as (49).

(49) Can the classifier systems seen in synchronic Arawakan languages be shown to share the same diachronic source?

      a.    To what degree are the Arawakan systems of the NWA unique to this area?

      b.    To what degree can cognates be identified in the synchronic systems which point to an early development of a classifier system?

      c.    To what degree can semantic evidence be identified in the synchronic systems which points to an early development of a classifier system?

      d.    To what degree do the morphosyntactic properties of the synchronic systems point to an early development of a classifier system?

Based on these aims, I investigated the morphosyntactic similarities among 11 Arawakan classifier systems, and showed that not only are these systems similar in their functional use and in their structural possibilities, but that the features which have been claimed to be unique to the NWA are found throughout the Arawakan family, which lends credence to the possibility of a family internal source at the level of PA. This is supported by evidence of the cognacy between systems, which indicated that a small group of classifiers can be reconstructed to PA which do not show any areal patterns in their distribution, development, or survival. Although the cognates survive in only a handful of languages, they were supported by both semantic and phonological evidence, pointing to a shared origin over the possibility of a number of independent parallel developments. Furthermore, the study revealed a semantic basis to the system based on physical properties, including material, texture, and shape. Based on these results, it is possible to propose that the Arawakan classifier systems seen in the languages today developed from a single PA system, which, although different in many ways from its descendent systems, would have been recognisable as a classifier system. It is also possible to make some further claims about what this looked like and how it developed, which I set out below.

      i.    Proto-Arawakan allowed classifiers in the numeral position.

      ii.    At a later stage, Proto-Arawakan featured an open system which allowed lexical nouns to occur with the noun (through compounding), with the numeral (agreement), or with the verb (through incorporation).

      iii.    In Proto-Arawakan, a small set of nouns already showed signs of grammaticalisation, particularly semantic bleaching and phonological erosion.

      iv.    After diversification of the daughter languages, the system become a closed classifier system and the use of full nouns (and some grammaticalised forms) was lost.

Each of these points will be discussed in further detail below.

      i.    Proto-Arawakan allowed classifiers in the numeral position.

Prior to discussing other elements of the system I have proposed above, it is possible to make some comments on how this system evolved in PA. As already established, there is a cross linguistic tendency for classifiers in the numeral position to be selected for semantics based on shape and physicality, which is related to their use in operations of individuation and relating to object handling (Craig 1992; Bisang 2002; Grinevald 2002). This is opposed to classifiers which develop in the nominal position, which are selected for semantics relating to an objects category or 'essence' (Craig 1992). This supports the proposal that the use of classifiers in PA was first found with numerals, most likely as ungrammaticalised morphemes sourced from

lexical nouns. This system would have looked similar to the repeater constructions seen in Japanese and Miraña, in which lexical nouns can be used as classifiers. The examples from Miraña are repeated below, demonstrating the use of a lexical noun in a position which is usually taken by a classifier (50a), and which is not available to classifiers (50b).

(50) MIRAÑA (WITOTOAN: NWA)
    a.    kátúꞝ:βɛ-**báhku**
          fall-RP.bone
          'The bone fell.'
    b.    táj-**báhku**
          POS.1S-bone
          'my bone'

<div align="right">(Seifart 2005, 81)</div>

By using lexical nouns, this would have formed an open system, in which any noun in theory could be used in this position. However, it was established previously that the functions of the classifier system (agreement, derivation, anaphoric reference) are tied to the positions in which classifiers are used, and it is likely that this already acted as a limit on the selection of nouns, and impacted the meanings of the nouns when they occurred in this position.

That the numeral position for classifiers developed first is supported by the widespread occurrence of classifiers in the numeral position across the sample and throughout the family, and the semantics which underpin the synchronic systems.

    ii.    At a later stage, Proto-Arawakan featured an open system which allowed lexical nouns to occur with the noun (through compounding), with the numeral (agreement), or with the verb (through incorporation).

The spread of classifiers to nominal and verbal positions is proposed as a second stage of development in later PA. I argue that the system remained an open system which used lexical nouns, although as in the previous stage this would have been restricted by the developing functions of each position. Classifiers in a verbal position would have been a natural consequence of the spread to the nominal position, although it is unclear if it preceded or followed the development of incorporation of the noun more generally.

In proposing this wider system, an obligatory system in which classifiers occurred in every position is not expected or evidenced. Instead, as in the descendent languages, these positions developed as optional positions for classifiers, depending on the intended usage. That these positions developed later in the PA period is reflected in their less widespread usage; if these positions were less well established, they would have been more easily lost in later stages of development. This instability is reflected in the optionality of classifiers in a number of languages, as in the examples below from Tariana.

(51) TARIANA (NWA)
    a.    ha      papera  iri-**yami**
          DEM    paper   red-CL:piece
          'this piece of paper is red'
    b.    ha      papera-yami     iri-**yami**
          DEM    paper-CL:piece   red-CL:piece
          'this piece of paper is red'

<div align="right">(Aikhenvald 1994b, 439)</div>

iii.  In Proto-Arawakan, a small set of nouns already showed signs of grammaticalisation, particularly semantic bleaching and phonological erosion.

The twelve classifier forms presented in **6** date to PA, and were already in this form at that level, although a number have not survived in the daughter languages. This large-scale loss across languages points to the instability of the PA system, and may also have been aided by high levels of contact with non-Arawakan languages.

Although most of these reconstructed classifiers have not yet been connected with lexical nouns, such links may be made as further phonological and lexical reconstructions of the earlier stages of the languages are developed. However, evidence of grammaticalisation can be seen in four of the reconstructed forms (*pi, *ako, *sa, *panhi*), which show semantic bleaching and phonological erosion, although to different extents.

That these four forms already show signs of grammaticalisation at the PA stage suggests a level of grammaticalisation in the wider system which continued into the daughter languages with the reduction of the system.

iv.  After diversification of the daughter languages, the system become a closed classifier system and the use of full nouns (and some grammaticalised forms) was lost.

That the systems which have survived today show so much diversity in their grammatical function, position, system size, and classifier sets, points to a system which still showed fluctuation and did not stabalise as a grammatical system until after diversification. By fluctuation, it is intended that they system still allowed a great deal of freedom in which nouns could and would regularly be used, and that it was an optional rather than obligatory system. Such fluctuation is seen in the differences in how anaphoric reference has developed: in some languages, such as Palikur, it is an extension of the numeral position, while in others, such as Paresi and Tariana, a pronominal structure is used. Both of these structures are exemplified below.

(51) PALIKUR (BRAZIL; FRENCH GUIANA)

| ba | pis | muwaka | ax | paha-**t** |
|----|-----|--------|-----|-----------|
| INTER | 2SG | want | eat | one-NUM.CL:vertical |

'Do you want to eat one (banana fruit)?'

(Aikhenvald and Green 1998, 445)

(52) PARESI (GAUPORÉ-MAMORÉ)

*taika hanatse nikare*

| ø=taika | ha=**natse** | nikare |
|---------|-----------|--------|
| 3sg=break | 3sg=CLF:long | like.this |

'He broke (it) like this.'

(Brandão 2014, 265)

Above I have presented four stages of the PA classifier system, which I claim was established as a numeral classifier system and spread to the verbal and nominal positions sometime within the PA period. This system was a single system with some level of grammaticalisation, visible in the forms which have survived today, but was an open system available to free nouns, and did not become a closed system until after the diversification of the language family. Aikhenvald (2019) has previously argued that the numeral position may have been the earliest point of development of classifiers in languages of the NWA. Here I have argued not only that this is more likely than her alternative proposal that classifiers developed in the nominal position, but that the numeral position is the earliest point of development of classifiers across the whole family, not just in the NWA. However, how this classifier system relates to the compound systems remains an unresolved issue which requires further attention, and will be the focus of the rest of this chapter.

It was presented above that the numeral position for classifiers developed prior to the nominal and verbal positions in PA, on the basis of the spread of the numeral position throughout the family and the semantic basis of compounds. This position can be further supported by the relationships both between the classifier and compounds sets of different languages and within the same languages. In **6**, closed compound sets were included in the study of cognacy where they were available. However, only one cognate set included a form considered a compound (Set 12). This indicates that the cognacy is between the classifier sets, but not between the compound sets, suggesting that the classifiers predate the compounds. Furthermore, some compound forms show evidence of being composed of classifiers; this is particularly evident in Alto Perené. In the example below, form (53a) is a classifier, and the forms in (53b) are class terms, used in compound constructions. In each case, the class term includes the syllable *-ki*, and the meaning is related to smallness, indicating a semantic link with the classifier.

(53)  ALTO PERENÉ (PERU)

    a.   *-ki*                round, small
    b.   *-monki*         small-scale convexity, suspended
        *-naraki*        small, flat, elongated
        *-pentaki*      rectangular, thin, small
        *-petoki*        small piece, small sized
        *-tsintaki*     cluster of small, ball-like objects

<div align="right">(Mihas 2015, 410, 414)</div>

Although further investigation would be required to see if the class terms are morphologically complex forms, this strongly suggests that these terms developed after the formation of the classifier.

The evidence presented above points to both the classifier forms and the classifier system being older than the compound forms and the compound system. However, it has also been argued that the compound systems in a number of these languages are related to the possessive construction, as an evolution of the alienability distinction available in those languages, which dates to PA (see Baure: Danielsen 2007, 134; Apurina: Facundes 2000, 150; Paresi: Brandão 2014, 179). [47] If this were the case, it could be argued that the nominal position predated the numeral position, rather than the other way around. In this paper I maintain that the data points to an earlier numeral position, however, the connection between classifiers, compound nouns, and inalienable possession may yet prove a fruitful field of further research in the Arawakan languages. Furthermore, although Van der Voort (2018) considers compound systems an areal feature of the Guaporé-Mamoré region, the description here has highlighted that these systems are much more widespread in the Arawakan family than just this region. Whether the Arawakan languages are the source of this areal feature may also be worthy of further investigation.

---

[47] Facundes (2000) considers both classifiers (CN2s) and compound nouns (CN1s) types of inalienable noun.

## 8. Conclusion

This paper set out to consider the possibility of an internal source for the classifier systems seen in Arawakan languages today, given the evidence of subfamily innovations and areal influence from neighbouring languages. The research question and objectives given in **1** are repeated below.

(55) Can the classifier systems seen in synchronic Arawakan languages be shown to share the same diachronic source?

      a.   To establish to what degree are the Arawakan systems of the NWA unique to this area?

      b.   To establish to what degree can cognates be identified in the synchronic systems which point to an early development of a classifier system?

      c.   To establish to what degree can semantic evidence be identified in the synchronic systems which points to an early development of a classifier system?

      d.   To establish to what degree do the morphosyntactic properties of the synchronic systems point to an early development of a classifier system?

Two sources of evidence were considered: that of cognancy between the systems, and that of the structural similarities between the systems. Although a small set of cognates were found, they were thinly spread between the languages given here. In contrast, the languages showed a great deal of similarity in the structures of the systems, and in a number of cases languages show a stronger familial similarity than they do areal similarity. This speaks directly to point (a): although multiple classifier systems have been described as characteristic of the North West Amazon, systems in which classifiers can be used in a number of positions are characteristic of the Arawakan family as a whole, and should not be considered restricted to one area within the Amazon, as also argued by Van der Voort (2018).

The remaining points (b-d) motivated a cognate search, a comparison of the synchronic structural systems and consideration of the semantic basis of the systems. The evidence of both the cognate search and the comparison of structural systems points towards a family internal source of the underlying structure which was greatly expanded in individual subfamilies, leading to little similarity in the classifier forms themselves. This is supported by the evidence of the semantics: the systems show a strong bias towards semantics relating to shape and physical form. This is further evidence that the systems developed from one source, although the forms themselves have diverged greatly.

In arguing that the Arawakan classifier systems have developed from one structural source, I claimed the following four things, repeated here:

    i.    Proto-Arawakan allowed classifiers in the numeral position.

    ii.   At a later stage, Proto-Arawakan featured an open system which allowed lexical nouns to occur with the noun (through compounding), with the numeral (agreement), or with the verb (through incorporation).

    iii.  In Proto-Arawakan, a small set of nouns already showed signs of grammaticalisation, particularly semantic bleaching and phonological erosion.

    iv.  After diversification of the daughter languages, the system become a closed classifier system and the use of full nouns (and some grammaticalised forms) was lost.

The evidence presented in this paper points to a classifier system at the level of PA which supplied the structural material to allow divergent systems to develop in the daughter languages, and that much of the material which would become the classifier forms was already in use in such positions.

The evidence presented in this paper has come from a subset of the Arawakan languages, aiming to represent a broad picture of the family, and further study of more languages may yet reveal further cognate

forms and give greater detail on a Proto-Arawakan classifier system. Moreover, a number of features were discussed here which do not conform to areal patterns or which are more widely spread in the Arawakan family than an areal study would suggest: this was particularly the case with compounding structures. For this reason, further study into the position of the Arawakan family in relation to areal features may reveal interesting patterns as to the origin of those features.

# 9. Bibliography

Adams, Karen L., and Nancy Faires Conklin. 1973. 'Toward a Theory of Natural Classification'. In Papers from the Ninth Regional Meeting Chicago Linguistic Society, edited by Claudia Corum, T. Cedric Smith-Stark, and Ann Weiser. Chicago: Chicago Linguistic Society.

Aikhenvald, Alexandra Y. 1994a. 'Classe Nominal E Gênero Nas Línguas Aruák [Nominal Class and Gender in Arawakan Languages]'. Boletim Do Museo Paraense Emílio Goeldi 10: 137–259.

———. 1994b. 'Classifiers in Tariana'. Anthropological Linguistics 36 (4): 407–65.

———. 1998. 'Warekena'. In Handbook of Amazonian Languages, edited by Desmond C. Derbyshire and Geoffrey K. Pullum, 4:225–440. Berlin: Mouton de Gruyter.

———. 1999. 'The Arawak Language Family'. In The Amazonian Languages, edited by Alexandra Y. Aikhenvald and R. M. W Dixon, 65–106. Cambridge: Cambridge University Press.

———. 2000. Classifiers: A Typology of Noun Categorization Devices. Oxford: Oxford University Press.

———. 2003. A Grammar of Tariana, from Northwest Amazonia. Cambridge: Cambridge University Press.

———. 2006. 'Areal Diffusion, Genetic Inheritance, and Problems of Subgrouping: A North Arawak Case Study'. In Areal Diffusion and Genetic Inheritance: Problems in Comparative Linguistics, edited by Alexandra Y. Aikhenvald and R. M. W Dixon, 167–94. Oxford: Oxford University Press.

———. 2012. The Languages of the Amazon. New York: Oxford University Press.

———. 2019. 'A View from the North: Genders and Classifiers in Arawak Languages of North-West Amazonia'. In Gender and Classifiers: A Cross-Linguistic Typology, edited by Alexandra Y. Aikhenvald and Elena Mihas, 103–43. Oxford: Oxford University Press.

Aikhenvald, Alexandra Y., and R. M. W Dixon. 1999. 'Introduction'. In The Amazonian Languages, edited by Alexandra Y. Aikhenvald and R. M. W Dixon, 1–22. Cambridge: Cambridge University Press.

Aikhenvald, Alexandra Y., and Diana Green. 1998. 'Palikur and the Typology of Classifiers'. Anthropological Linguistics 40 (3): 429–80.

Allan, Keith. 1977. 'Classifiers'. Language 53 (2): 285–311.

Allin, Trevor R. 1976. 'A Grammar of Resígaro'. PhD Thesis, University of St Andrews.

Beckwith, Christopher. 2007. Phoronyms: Classifiers, Class Nouns, and the Pseudopartitive Construction.

Bellamy, Kate. 2018. 'Re-Evaluating the "numeral Classifier" System in Mochica'. LIAMES: Línguas Indígenas Americanas 18 (1): 81–98.

Birchall, Joshua. 2014. 'Argument Marking Patterns in South American Languages'. PhD Thesis, Nijmegen: Radboud Universiteit Nijmegen.

Bisang, Walter. 1999. 'Classifiers in East and Southeast Asian Languages: Counting and Beyond'. In Numeral Types and Changes Worldwide, edited by Jadranka Gvozdanovic, 118:113–86. Berlin: De Gruyter Mouton.

———. 2002. 'Classification and the Evolution of Grammatical Structures: A Universal Perspective'. STUF - Language Typology and Universals 55 (3): 289–308.

Brandão, Ana Paula Barros. 2014. 'A Reference Grammar of Paresi-Haliti (Arawak)'. PhD Thesis, The University of Texas at Austin.

Campbell, Lyle. 2012. 'Classification of the Indigenous Languages of South America'. In The Indigenous Languages of South America: A Comprehensive Guide, edited by Lyle Campbell and Verónica Grondona, 2:59–166. Berlin: Mouton de Gruyter.

Campbell, Lyle, and Verónica Grondona. 2012. 'Languages of the Chaco and Southern Cone'. In The Indigenous Languages of South America: A Comprehensive Guide, edited by Lyle Campbell and Verónica Grondona, 2:625–67. Berlin: Mouton de Gruyter.

Carvalho, Fernando O. de. 2016a. 'Internal and Comparative Reconstruction in Yawalapiti: Palatization and Rule Telescoping'. International Journal of American Linguistics 82 (3): 285–316.

———. 2016b. 'Obscure Cognates and Lexical Reconstruction: Notes on the Diachrony of the Xinguan Arawak Languages'. Boletim Do Museu Paraense Emílio Goeldi. Ciências Humanas 11 (1): 277–94.

———. 2018. 'Diachronic Split and Phoneme Borrowing in Resígaro (Arawakan)'. Canadian Journal of
     Linguistics/Revue Canadienne de Linguistique 63 (3): 339–58.

———. 2021. 'A Comparative Reconstruction of Proto-Purus (Arawakan) Segmental Phonology'.
     International Journal of American Linguistics 87 (1): 49–108.

Contini-Morava, Ellen, and Marcin Kilarski. 2013. 'Functions of Nominal Classification'. Language Sciences
     40: 263–99.

Corbett, Greville G. 1991. Gender. Cambridge: Cambridge University Press.

———. 2006. Agreement. Cambridge: Cambridge University Press.

Corbett, Greville G., and Sebastian Fedden. 2016. 'Canonical Gender'. Linguistics 52 (3): 495–531.

———. 2017. 'Gender and Classifiers in Concurrent Systems: Refining the Typology of Nominal
     Classification'. Glossa: A Journal of General Linguistics 2 (1): 1–47.

Craig, Colette G. 1986. 'Jacaltec Noun Classifiers: A Study in Grammaticalization'. Lingua 70 (4): 241–84.

———. 1992. 'Classifiers in a Functional Perspective'. In Layered Structure and Reference in a Functional
     Perspective: Papers From the Functional Grammar Conference in Copenhagen, 1990, edited by Lars
     Kristoffersen, Peter Harder, and Michael D. Fortescue, 277–302. Amsterdam: John Benjamins
     Publishing Company.

Crevels, Mily, and Hein Van der Voort. 2008. 'The Guaporé-Mamoré Region as a Linguistic Area'. In From
     Linguistic Areas to Areal Linguistics, edited by Pieter Muysken, 151–80. Amsterdam: John
     Benjamins Publishing Company.

Croft, William. 1994. 'Semantic Universals in Classifier Systems'. Word 45 (2): 145–71.

Danielsen, Swintha. 2007. 'Baure: An Arawak Language of Bolivia'. PhD Thesis, Radboud Universiteit
     Nijmegen.

Denny, J.P. 1976. 'What Are Noun Classifiers Good For?' In Papers from the Regional Meeting of the Chicago
     Linguistic Society 12, 122–32.

Derbyshire, D. C., and Doris L. Payne. 1990. 'Noun Classification Systems of Amazonian Languages'. In
     Amazonian Linguistics: Studies in Lowland South American Languages, edited by Doris L. Payne,
     243–71. Texas: University of Texas Press.

Dixon, R. M. W. 1986. 'Noun Classes and Noun Classification in Typological Perspective'. In Noun Classes
     and Categorization Proceedings of a Symposium on Categorization and Noun Classification,
     Eugene, Oregon, October 1983, edited by Colette Grinevald Craig, 105–12. Amsterdam: John
     Benjamins Publishing Company.

Duff-Tripp, Martha. 1997. Gramatica Del Idioma Yanesha' (Amuesha). Lima: Ministerio de Educación,
     Instituto Lingüístico de Verano.

Epps, Patience, and Lev Michael. 2017. 'The Areal Linguistics of Amazonia'. In The Cambridge Handbook of
     Areal Linguistics, 934–63. Cambridge: Cambridge University Press.

Eriksen, Love. 2011. 'Nature and Culture in Prehistoric Amazonia: Using G.I.S to Reconstruct Ancient
     Ethnogenetic Processes from Archaeology, Linguistics, Geography, and Ethnohistory'. PhD Thesis,
     Lund University.

Eriksen, Love, and Swintha Danielsen. 2014. 'The Arawakan Matrix'. In The Native Languages of South
     America: Origins, Development, Typology, edited by Loretta O'Connor and Pieter Muysken, 152–76.
     Cambridge: Cambridge University Press.

Facundes, Sidney da Silva. 2000. 'The Language of the Apurinã People of Brazil: (Maipure/Arawak)'. PhD
     Thesis, University of New York at Buffalo.

Gomes, Adrian Sydney. 2022. 'Introduction of a Wapishana-English Bilingual Education Programme: An
     Evaluation of the Early Stages'. PhD Thesis, Leiden University.

Greenberg, Joseph H. 1972. 'Numeral Classifiers and Substantival Number: Problems in the Genesis of a
     Linguistic Type'. Working Papers on Language Universals 9: 1–39.

Grinevald, Colette. 2000. 'A Morphosyntactic Typology of Classifiers'. In Systems of Nominal Classification,
     edited by Senft, Gunter, 50–92. Cambridge: Cambridge University Press.

———. 2002. 'Making Sense of Nominal Classification Systems'. In New Reflections on Grammaticalization, edited by Ilse Wischer and Gabriele Diewald, 259–76. Amsterdam: John Benjamins Publishing Company.

———. 2015. 'Classifiers, Linguistics Of'. In International Encyclopedia of the Social & Behavioral Sciences, edited by James D. Wright, 2nd ed., 811–18. Oxford: Elsevier.

Grinevald, Colette, and Frank Seifart. 2004. 'Noun Classes in African and Amazonian Languages: Towards a Comparison'. Linguistic Typology 8 (2): 243–85.

Haurholm-Larsen, Steffen. 2016. 'A Grammar of Garifuna'. PhD Thesis, Bern: University of Bern.

Heckenberger, Michael. 2002. 'Rethinking the Arawakan Diaspora: Hierarchy, Regionality, and the Amazonian Formative'. In Comparative Arawakan Histories : Rethinking Language Family and Culture Area in Amazonia, edited by Jonathan D. Hill and Fernando Santos-Granero, 99–122. Urbana: University of Illinois Press.

Jolkesky, Marcelo. 2016. 'Uma Reconstrução Do Proto-Mamoré-Guaporé (Família Arawák) [A Reconstruction of Proto-Mamoré-Guaporé (Arawakan)]'. LIAMES: Línguas Indígenas Americanas 16 (1): 7–37.

Kaufman, Terrence. 1994. 'The Native Languages of South America'. In Atlas of the World's Languages, 59–93. London: Routledge.

Kilarski, Marcin. 2013. Nominal Classification: A History of Its Study from the Classical Period to the Present. Amsterdam: John Benjamins Publishing Company.

Klumpp, Deloris A. 2019. A Grammar of Piapoco. Edited by Barbara E. Hollenbach. SIL International.

Kouteva, Tania, Bernd Heine, Bo Hong, Haiping Long, Heiko Narrog, and Seongha Rhee. 2019. 'Introduction'. In World Lexicon of Grammaticalization, 1–22. Cambridge: Cambridge University Press.

Krasnoukhova, Olga. 2012. 'The Noun Phrase in the Languages of South America'. PhD Thesis, Utrecht: Radboud Universiteit Nijmegen.

Lemus Serrano, Magdalena. 2020. 'Pervasive Nominalization in Yukuna: An Arawak Language of Colombian Amazonia'. PhD Thesis, Université Lumière Lyon 2: Sciences du Langage.

Michael, Lev. 2021. 'The Classification of South American Languages'. Annual Review of Linguistics 7: 329–49.

Michael, Lev, Robin Alcorn, Lisa Fann, Briana Van Epps, and Mona Zarka. 2010. 'Phonological Reconstruction of the Kampan Branch of Arawak'. Presented at the Workshop on American Indian Languages.

Mihas, Elena. 2015. A Grammar of Alto Perené (Arawak). Berlin: De Gruyter Mouton.

Norder, S. J. et al. 2022. Glottospace: R Package for the Geospatial Analysis of Linguistic and Cultural Data. https://github.com/SietzeN/glottospace.

Olness, G. 1991. 'Semantic Correlate of Three Morphosyntactic Classifier Types: A Cross-Linguistic Survey'. Ms., University of Oregon.

Passer, Matthias Benjamin. 2016. 'The Typology and Diachrony of Nominal Classification'. PhD Thesis, Universiteit van Amsterdam.

Payne, David. 1987. 'Some Morphological Elements of Maipuran Arawakan: Agreement Affixes and the Genitive Construction'. Language Sciences 9 (1): 57–75.

———. 1990. 'Some Widespread Grammatical Forms in South American Languages'. In Amazonian Linguistics: Studies in Lowland South American Languages, edited by Doris L. Payne, 75–88. Austin: University of Texas Press.

———. 1991. 'A Classification of Maipuran (Arawakan) Languages Based on Shared Lexical Retentions'. In Handbook of Amazonian Languages, edited by D. C. Derbyshire and Geoffrey K. Pullum, 3:355–499. Berlin: De Gruyter Mouton.

Payne, Doris L. 1987. 'Noun Classification in the Western Amazon'. Language Sciences 9 (1): 21–44.

———. 2007. 'Source of the Yagua Nominal Classification System'. International Journal of American Linguistics 73 (4): 447–74.

Ramirez, Henri. 2020. Enciclopédia Das Línguas Arawak: Acrescida de Seis Novas Línguas e Dois Bancos de Dados. Vol. 1. 4 vols. Curitiba, Brazil: Editora CRV.

Sands, Kristina. 1995. 'Nominal Classification in Australia'. Anthropological Linguistics 37 (3): 247–346.

Santos, Manoel Gomes dos. 2006. 'Uma Gramática Do Wapixana (Aruák) - Aspectos Da Fonologia, Da Morfologia e Da Sintaxe'. PhD Thesis, Instituto de Estudos da Linguagem.

Santos-Granero, Fernando. 2002. 'The Arawakan Matrix: Ethos, Language, and History in Native South America'. In Comparative Arawakan Histories: Rethinking Language Family and Culture Area in Amazonia, edited by Jonathan D. Hill and Fernando Santos-Granero, 25–50. Urbana: University of Illinois Press.

Seifart, Frank. 2005. 'The Structure and Use of Shape-Based Noun Classes in Miraña (North West Amazon)'. PhD Thesis, Wageningen: Radboud Universiteit Nijmegen.

———. 2007. 'The Prehistory of Nominal Classification in Witotoan Languages'. International Journal of American Linguistics 73 (4): 411–45.

———. 2011. Bora Loans in Resígaro: Massive Morphological and Little Lexical Borrowingi n a Moribund Arawakan Language. Max Planck Institute for Evolutionary Anthropology.

———. 2012a. 'Causative Marking in Resígaro (Arawakan): A Descriptive and Comparative Perspective'. International Journal of American Linguistics 78 (3): 369–84.

———. 2012b. 'The Principle of Morphosyntactic Subsystem Integrity in Language Contact: Evidence from Morphological Borrowing in Resígaro (Arawakan)'. Diachronica 29 (4): 471–504.

Seifart, Frank, and Doris L. Payne. 2007. 'Nominal Classification in the North West Amazon: Issues in Areal Diffusion and Typological Characterization'. International Journal of American Linguistics 73 (4): 381–87.

Senft, Gunter. 1993. 'A Grammaticalization Hypothesis on the Origin of Kilivila Classificatory Particles'. STUF - Language Typology and Universals 46 (2): 100–112.

———. 2000. 'What Do We Really Know about Nominal Classification Systems?' In Systems of Nominal Classification, edited by Senft, Gunter, 11–49. Cambridge: Cambridge University Press.

Siewierska, Anna. 1999. 'From Anaphoric Pronoun to Grammatical Agreement Marker: Why Objects Don't Make It.' Folia Linguistica 33 (1–2): 225–52.

Steele, S. 1978. 'Word Order Variation: A Typology Study'. In Universals of Human Language, edited by Joseph H. Greenberg, C. A. Ferguson, and E. A. Moravscik, 585–623. Syntax, IV. Stanford: Stanford University Press.

Van der Voort, Hein. 2018. 'Development and Diffusion of Classifier Systems in Southwestern Amazonia'. In The Diachrony of Classification Systems, edited by William McGregor and Søren Wichmann, 201–40. John Benjamins Publishing Company.

Walker, Robert S., and Lincoln Ribeiro. 2011. 'Bayesian Phylogeography of the Arawak Expansion in Lowland South America'. Proceedings of the Royal Society B: Biological Sciences 278 (1718): 2562–67.

Welmers, William E. 1973. African Language Structures. Berkeley: University of California Press.

Wiegertjes, Jurriaan. 2022. 'The Development of Classifiers in the Tukanoan Family'. MA Thesis, Leiden University.

Wiemer, Björn, and Walter Bisang. 2004. 'What Makes Grammaticalization? An Appraisal of Its Components and Its Fringes'. In What Makes Grammaticalization? A Look from Its Fringes and Its Components, edited by Walter Bisang, Nikolaus P. Himmelmann, and Björn Wiemer, 3–21. Berlin: Mouton de Gruyter.

Wise, Mary Ruth, ed. 1998. Diccionario Yanesha' (Amuesha) ~ Castellano. Lima: Ministerio de Educación, Instituto Lingüístico de Verano.

Zavala, Roberto. 1992. El Kanjobal De San Miguel Acatán. Mexico: Universidad Nacional Autónoma de México. Instituto de Investigaciones Filológicas.

———. 2000. 'Multiple Classifier Systems in Akatek (Mayan)'. In Systems of Nominal Classifications, edited by Gunter Senft, 114–46. Cambridge: Cambridge University Press.

## Appendix I.       Baure classifiers

All classifiers from Baure were considered for cognancy, including repeaters. Although compound nouns are found, there is no closed set of bound nouns used in compounds, so bound nouns were not included.

The following orthographic conventions are used:

š    /ʃ/
č    /tʃ, dʒ/
'    /ʔ/

Source: (Danielsen 2007, 444–47)

| Form | Gloss | Type |
| --- | --- | --- |
| *-no* | human | classifier |
| *-a* | animal | classifier |
| *-i* | fruit and bird | classifier |
| *-e* | unsweet | classifier |
| *-pa* | flat and round | classifier |
| *-pe* | cutlery | classifier |
| *-pi* | long and thing | classifier |
| *-po* | tiny | classifier |
| *-mo* | woven | classifier |
| *-mpe* | flat and round | classifier |
| *-sa* | water | classifier |
| *-se* | oval | classifier |
| *-si* | stick | classifier |
| *-iro* | round | classifier |
| *-aro* | liquid | classifier |
| *-či* | month | classifier |
| *-aha* | dish | classifier |
| *-ki* | contents | classifier |
| *-'ino* | tool | classifier |
| *-seki* | oval container contents | classifier |
| *-eko* | jug | classifier |
| *-iki* | net | classifier |
| *-ahaki* | contents of pots | classifier |
| *-koki* | inside | classifier |
| *-poki* | hammock | classifier |
| *-ake* | side3D | classifier |
| *-sare* | side2D | classifier |
| *-aso* | time | classifier |
| *-pi* | word | classifier |
| *-amok* | flat and raised | classifier |
| *-yok* | times | classifier |
| *-čipi* | roof | classifier |
| *-imir* | face | classifier |
| *-kis* | eye | classifier |
| *-waki* | handful | classifier |
| *-čipie* | pile | classifier |
| *-čow* | circle | classifier |

| Form | Gloss | Type |
|---|---|---|
| *-aki* | diameter | classifier |
| *-še* | trunc | classifier |
| *-poe* | down | classifier |
| *-api* | under | classifier |
| *ači* | other. Place | classifier |
| *-pon* | leaf | repeater |
| *-ewo* | raceme | repeater |
| *-soki* | seed | repeater |
| *-poe* | banana | repeater |
| *-moto* | plant | repeater |
| *-wok* | tree | repeater |
| *-čomomoe'* | flower | repeater |
| *-ni* | voice | repeater |
| *-wohis* | hand | repeater |
| *-ser* | tooth | repeater |
| *-pasiri* | nose | repeater |
| *-čokon* | ear | repeater |
| *-pes* | leg | repeater |
| *-imas* | genitals | repeater |
| *-hi* | horn | repeater |
| *-iši* | feather | repeater |
| *-ihis* | tail | repeater |
| *-powoki* | arm | repeater |
| *-tip* | nail | repeater |
| *-morekoe'* | year | repeater |
| *-sekone* | day | repeater |
| *-yitinoe* | night | repeater |
| *-eši* | piece | repeater |
| *-čon* | snail | repeater |
| *-čosoe* | fence | repeater |
| *-čoro* | frame | repeater |
| *-esonokoe* | door | repeater |
| *-torokoe* | corner | repeater |
| *-kori* | arrow | repeater |
| *-toto* | pebble | repeater |
| *-kor* | fishing net | repeater |
| *-si* | sweet potato | repeater |
| *-tono* | button | repeater |
| *-mes* | table | repeater |

## Appendix II.    Paresi classifiers

All classifiers and bound nouns were considered for cognancy.

Source: (Brandão 2014, 182, 185)

| Form | Gloss | Type |
| --- | --- | --- |
| *tse* | small, headwater | classifier |
| *hi* | long, slender, flexible (vine-like) | classifier |
| *li* | round | classifier |
| *he* | powder | classifier |
| *natse* | long horizontally, cylindrical, three-dimension | classifier |
| *katse* | thin, rigid, long vertically (stick-like) | classifier |
| *za* | liquid; speech | classifier |
| *hoko* | circled, three-dimension | classifier |
| *taotse* | piece, one-dimension | classifier |
| *koa* | flat surface, one-dimension | classifier |
| *ako* | inside of a hollow, three-dimension | classifier |
| *mena* | stalk of | compound |
| *tyahare* | root of | compound |
| *walahi* | vein of | compound |
| *hana* | leaf of | compound |
| *tane* | feather of | compound |
| *kino* | log | compound |
| *tyatya* | bark of | compound |
| *iye* | flower of | compound |
| *mili* | skin of | compound |
| *ri* | fruit of | compound |
| *tse* | seed of; tuber of | compound |
| *hi* | fiber of | compound |

# Appendix III.  Apurinã classifiers

All classifiers and bound nouns were considered for cognancy.

Source: (Facundes 2000, 176)

| Form | Gloss | Type |
|------|-------|------|
| *iri* | fruit of; bunch of | compound noun |
| *katu* | treetop of; thin branch of | compound noun |
| *kotsa* | root of | compound noun |
| *piti* | feather of; hair of | compound noun |
| *poru* | thick branch of | compound noun |
| *tako* | thin stalk of | compound noun |
| *tsota* | trunk of | compound noun |
| *xiti* | earth of | compound noun |
| *yoka* | spot of | compound noun |
| *ã* | water, juice; tear; liquid | classifier |
| *ke* | wood stick; long, thing | classifier |
| *ku* | rounded, small, hard | classifier |
| *mata* | skin; flat, soft | classifier |
| *muna* | trunk; long, cylindrical | classifier |
| *pẽ* | water, juice; liquid | classifier |
| *panhi* | powder | classifier |
| *pe* | mush; paste | classifier |
| *putsa* | liana; tripe; long, flexible | classifier |
| *riko* | hole | classifier |
| *tãta* | bark; shell ;flat, thick | classifier |
| *tsa* | liana; long, flexible | classifier |
| *tsopa* | wide leaf; flat, wide | classifier |
| *xike* | small leaf; flat, flexible | classifier |

## Appendix IV.    Alto Perené classifiers

All classifiers and class terms were considered for cognancy.

Source: (Mihas 2015, 410-411, 414)

| Form | Gloss | Type |
|------|-------|------|
| *-aa* | liquid | classifier |
| *-ako* | container | classifier |
| *-ampoiri* | tracks | classifier |
| *-chari* | grassy bottom of the river | classifier |
| *-impiki* | branching from one source, pointed | classifier |
| *-ina* | ball-like, spherical | classifier |
| *-ki* | round, small | classifier |
| *-ki* | social insect | classifier |
| *-kii* | pointed | classifier |
| *-mashi* | flat, broad, open space | classifier |
| *-meni* | flat, triangular, wedge-like | classifier |
| *-na* | long, hollow | classifier |
| *-naki* | hollow, round | classifier |
| *-ña* | stagnant water pool | classifier |
| *-ni* | running water feature | classifier |
| *-niro* | simulative | classifier |
| *-nki* | snake | classifier |
| *-nko* | branching from one source down | classifier |
| *-pa* | pod-like, cylindrical | classifier |
| *-paite* | period | classifier |
| *-pata* | flat, oblong, oar blade-like | classifier |
| *-peta* | devalued, old | classifier |
| *-pi* | rigid, hollow | classifier |
| *-to* | hollow, long, rigid | classifier |
| *-ta* | flat, vertical | classifier |
| *-tsa* | long, curvilinear | classifier |
| *-tsaranka* | big leaf-like, wide, flexible | classifier |
| *-tsori* | nuclear-like kin | classifier |
| *-chee* | crescent-shaped, sharp-ended | class term |
| *-chempoki* | naked | class term |
| *-ito* | sacred habitat area | class term |
| *-keni* | generic palm worm | class term |
| *-kitso* | hard, small, oval | class term |
| *-konaki* | gnarled, knotty, lumpy | class term |
| *-kota* | piece | class term |
| *-meriki* | slightly curved, rigid, thin | class term |
| *-meronta* | flat, smooth, rigid | class term |
| *-metara* | flat, circular | class term |
| *-moko* | spherical, solid | class term |
| *-monki* | small-scale convexity, suspended | class term |
| *-motsa* | soft, flowing, shapeless | class term |
| *-nampi* | side | class term |
| *-naraki* | small, flat, elongated | class term |

| Form | Gloss | Type |
|---|---|---|
| *-pana* | green leaf | class term |
| *-pani* | powder-like substance | class term |
| *-panki* | long, rigid | class term |
| *-panko* | sacred dwelling | class term |
| *-pashi* | dry leaf | class term |
| *-moko-/tero-/pompo-INA* | ball-like, spherical | class term |
| *-patsa* | jelly like substance | class term |
| *-pentaki* | rectangular, thin, small | class term |
| *-petoki* | small piece, small sized | class term |
| *-piro* | thin, flaccid | class term |
| *-pontso* | stump-like, with the blunt end | class term |
| *-poroki* | group | class term |
| *-sani* | generic wasp | class term |
| *-shi* | leaf-like, leafy substance | class term |
| *-tapo* | disc-like, curved | class term |
| *-taro* | thick, rigid | class term |
| *-tenka/-teni* | convex plane | class term |
| *-tonki* | long, thin, rigid | class term |
| *-tonoki* | fleshy, elongated, crescent-shaped | class term |
| *-tsintaki* | cluster of small, ball-like objects | class term |
| *-tsomonte* | bulge | class term |
| *-tsovaki* | rigid, thin, curved, pointed | class term |
| *-vatsa* | fibrous substance | class term |
| *-venki* | magic plant | class term |

## Appendix V.    Yanesha' classifiers

Only classifiers relating to physical form were checked for cognancy: these are indicated in the list below. There is no closed set of bound nouns used in compounds.

The following orthographic conventions are used (for more details see Wise (1998, 11–12)):

| | | | |
|---|---|---|---|
| b | /β/ | ll | /ʎ, ʎ̥/ |
| ch | /t͡ʂ/ | rr | /ɻ/ |
| g | /ɣ/ | sh | /ʃ/ |
| hu | /w/ | C̃ | /Cʲ/ |

Source: (Duff-Tripp 1997, 253-266)

| Form | Spanish definition | English translation | Type | Physical classifier |
|---|---|---|---|---|
| *-acr-/-tacr-* | rama de arbol o rio; horquillado | tree or river branch; bracketing | classifier | + |
| *-am-/-omo-* | tela, ropa | cloth, clothes | classifier | + |
| *-ama'-* | nido de pajaro, gallina, hormiga | bird's nest, hen, ant | classifier | |
| *-ame'-/'-m̃-* | huevo de pajaro, gallina, horminga, pez | bird egg, hen, ant, fish | classifier | |
| *-ap-/-p-/-a'p-* | largo en forma de hueso, platano | long bone-shaped, banana | classifier | + |
| *-a'puets-/-puets-* | lugar, sitio para hacer algo | place, place to do something | classifier | |
| *-ap̃nor-* | mujer | woman | classifier | |
| *-as-* | cosas pequenas; cosas de caza | small things; hunting stuff | classifier | + |
| *-as-/-a's-* | diente | tooth | classifier | |
| *-at̃-/-a't̃-/-pa't̃- /-Vpa't̃-* | estomago | stomach | classifier | |
| *-atsr-* | cosas en general | things in general | classifier | |
| *-com̃t-* | dorso de mano | back of hand | classifier | |
| *-con-* | tronco; puente | trunk; bridge | classifier | |
| *-cotom-/-ctom-/-cotm-/-Vctm-* | barriga | belly | classifier | |
| *-cheñ-* | choclo | corn | classifier | |
| *-che'ñ-* | espalda | back | classifier | |
| *-chep-* | muslo | thigh | classifier | |
| *-chets-* | carne, miembros del cuerpo | flesh, body parts | classifier | |
| *-chmap̃-* | ano | year | classifier | |
| *-chm̃-* | atado, manojo, paquete, racimo | tied, bunch, bundle, bunch | classifier | + |
| *-choy-/-chy-* | alma | soul | classifier | |
| *-čherret-* | mandibula | mandible | classifier | |
| *-čhn-* | cuello | neck | classifier | |
| *-čhočh-* | brote | outbreak | classifier | |

| Form | Spanish definition | English translation | Type | Physical classifier |
|---|---|---|---|---|
| *-čhoñ-* | producto de la chacra | product of the farm | classifier | |
| *-čho'ña* | | | classifier | |
| *-ñhoy* | cria (de animal, pajaro) | cria (of animal, bird) | classifier | |
| *-em̃-/-m̃-/-m-* | fruta y otras cosas pequenas y redondas | fruit and other small round things | classifier | + |
| *-em̃-/-m̃-/-m-* | mes; luna | month; Moon | classifier | |
| *-epue'-* | ala (de ave) | wing (of bird) | classifier | |
| *-eshll-* | piel pelada | peeled skin | classifier | + |
| *-etor-* | artifice | architect | classifier | |
| *-gorr-* | chucillo; anzuelo | chucillo; fish hook | classifier | |
| *-lle'ch-* | pedazos como palitos | pieces like sticks | classifier | + |
| *-llem̃-/-llm-* | mostacillas | beads | classifier | + |
| *-llom̃-/-llm-* | semilla, fruto con semillas y otras cosas redondas y pequenas incluso varios productos de la chacra; carachamas; pastillas; chaquira | seed, fruit with seeds and other round and small things including various products of the farm; carachamas; tablets; chaquira | classifier | + |
| *-ma'-* | arboleda, jardin, chacra | grove, garden, farm | classifier | |
| *-mat-/-muet-/-mt-* | dias | days | classifier | |
| *-me'r-/-mer-* | terron; manojo | lump; bunch | classifier | + |
| *-met-/-mt-* | cuchillo y otras cosas con superficie plana | knife and other things with flat surface | classifier | + |
| *-moc-* | cuerpo; animal | body; animal | classifier | |
| *-moch-/-mot̃-/-moc-* | gusano | worm | classifier | |
| *-mom-/-muen-/-muen-* | yuca | Yucca | classifier | |
| *-mo'r-* | follaje, ramas | foliage, branches | classifier | |
| *-mo'y-* | nido (hormiga); monton de algo incluyendo olas de agua | nest (ant); lot of something including water waves | classifier | |
| *-m̃-* | cascara dura (de zapallo, mani) | hard shell (pumpkin, peanut) | classifier | + |
| *-m̃-/-am̃-* | pelo | hair | classifier | |

| Form | Spanish definition | English translation | Type | Physical classifier |
|---|---|---|---|---|
| *-m̃ta'n-/-em̃ta'n-* | un lado de rio, barranco | one side of the river, ravine | classifier | |
| *-n-* | cuerpo animal, humano | animal body, human | classifier | |
| *-n-* | canoa | canoe | classifier | |
| *-nam-* | boca | mouth | classifier | |
| *-no'-* | en forma de cano, tubo | pipe-shaped, tube | classifier | + |
| *-non-* | en forma de cilindro grande | in the shape of a large cylinder | classifier | + |
| *-ñeñ* | lengua | tongue | classifier | |
| *-ñyap̃* | garganta | throat | classifier | |
| *-om-/-tom-* | tira, liber, forma de bejuco | tira, liber, liana form | classifier | + |
| *-ot* | mano | hand | classifier | |
| *-pa'll-* | forma de platano, pacay, vaina | banana shape, pacay, pod | classifier | + |
| *-pan-/-Vpn-* | hoja, hoja de papel | sheet, sheet of paper | classifier | + |
| *-par-/-Vpar-* | tratamiento medicinal; pusanga | medicinal treatment; pusanga | classifier | |
| *-pe'ch-* | flecha; palitos | arrow; sticks | classifier | + |
| *-po'ch-* | pie, pierna | foot leg | classifier | |
| *-pon-* | hueco; nalga, rabo | hole; buttock, tail | classifier | + |
| *-po'r-* | bulto; pluma | package; pen | classifier | |
| *-porr-* | nariz, pico, punta | nose, beak, tip | classifier | |
| *-pt-* | chala de maiz | corn husk | classifier | |
| *-puell-* | pluma | pen | classifier | |
| *-pueñ-/-Vpñ-* | ceniza, polvo; vapor | ash, dust; steam | classifier | + |
| *-p̃ar-/-Vp̃ar-* | tratamiento de una variedad de piripiri para encontrar animales en la caza, pusanga | treatment of a variety of piripiri to find animals on the hunt, pusanga | classifier | |
| *-que'-* | brazo | arm | classifier | |
| *-quet-/-Vquet-* | orega; ojo de aguja | pray; needle eye | classifier | |
| *-rëch-/-Včh-* | raiz, bejuco, soga, hilo; vena; intestino | root, vine, rope, thread; come to; intestine | classifier | + |
| *-rëch-* | cancion; nombre | song; Name | classifier | |

| Form | Spanish definition | English translation | Type | Physical classifier |
|---|---|---|---|---|
| *-roc-/-rec-/-rac-* | caracteristico | characteristic | classifier | |
| *-rrep̃-* | pie, base | foot, base | classifier | |
| *-rrog-/-rr-/-Vrr* | olla | cooking pot | classifier | + |
| *-rrom-/-rrm-* | piel o como piel, cuero; cascara | skin or like skin, leather; shell | classifier | + |
| *-rrorm̃-* | hueco redondo en la tierra | round hole in the ground | classifier | + |
| *-rrorr-* | humoso, nublado | smoky, cloudy | classifier | + |
| *-rrot-* | hombros | shoulders | classifier | |
| *-se'-/-soy-* | grasa; sebo; manteca | grease; tallow; butter | classifier | |
| *-se'-/-so'y-* | resina; savia; aceite vegetal; lagrima; gota | resin; sap; vegetable oil; tear; gout | classifier | |
| *-so'n-* | rata | rat | classifier | |
| *-sonap-* | pecho | chest | classifier | |
| *-sechep-* | una; garra | one; claw | classifier | |
| *-she'm-/-shem-/-sh-/-Vsh-/-e'm-* | malo | bad | classifier | |
| *-sheñ-* | persona | person | classifier | |
| *-shem̃t-* | un lado del techo | one side of the roof | classifier | + |
| *-shoñ-* | labios | lips | classifier | |
| *-tac-/-tec-/-tc-/-Vtc-* | pie, pata | foot, paw | classifier | |
| *-tag-/-t-/-tr-/-Vto'-* | cabeza | head | classifier | |
| *-tall-* | cosas aplanadas como escama, cascara, corteza | flattened things like scale, husk, bark | classifier | + |
| *-ta'r* | racimo, manojo, gajo | cluster, bunch, bunch | classifier | + |
| *-to'm-* | frente | forehead | classifier | |
| *-topeñ-/'-topñ-/-Vtopeñ-/-Vtopñ-* | limite, borde | limit, edge | classifier | + |
| *-tro'-/-tero'-* | conjunto | set | classifier | |
| *-t̃ap-* | pierna | leg | classifier | |
| *-t̃ep̃-/-ep̃-/-Vt̃ep-* | camino | way | classifier | |
| *-tsa'p-* | carrillo, mejilla | cheek, cheek | classifier | |
| *-Vcll-* | ojo; cara | eye; expensive | classifier | |
| *-Vch-* | arbol; tronco; tallo; palo | tree; trunk; stem; stick | classifier | + |
| *-Vchor-* | chaparro | shorty | classifier | + |

| Form | Spanish definition | English translation | Type | Physical classifier |
|------|--------------------|--------------------|------|---------------------|
| *-Včh-* | palabra, lengua | word, language | classifier | |
| *-Včhm-* | palos rectos; arboles | straight sticks; trees | classifier | + |
| *-Vll-/-ll-* | curvado; circular | bent; circular | classifier | + |
| *-Vmpañ-* | lugar, region | place, region | classifier | |
| *-Vp-* | semilla de zapallo, sandia | pumpkin seed, watermelon | classifier | |
| *-Vpen-/-pen-/-Vp̃n-/-p̃n-* | cerro, sierra; espalda, espinazo de pez o animal | hill, mountain range; back, backbone of fish or animal | classifier | |
| *-Vpe'n/-pe'n-* | alto como un cerro | high as a hill | classifier | + |
| *-Vpo-/-po-* | epoca, tiempo | time, time | classifier | |
| *-Vpo'-/-po'-/-pahu-* | casa | House | classifier | |
| *-Vp̃-/-p̃-/-p-* | cosas delgadas y largas (lapiz, culebra, aguja, palito) | thin and long things (pencil, snake, needle, stick) | classifier | + |
| *-Vque'-* | planta de semilla; renuevo de planta; semilla de adorno | seed plant; plant renewal; ornamental seed | classifier | |
| *-Vs-/-V's-/-s-* | agua, liquido, jugo; rio | water, liquid, juice; River | classifier | + |
| *-Vsme'll-* | cola, rabo | tail, tail | classifier | |
| *-Vso'n-* | raton | mouse | classifier | |
| *-Vt-* | superficie plana | flat surface | classifier | + |
| *-Vtn-/-n-* | tierra; piso | land; flat | classifier | + |
| *-Vt̃oll-/-t̃oll-* | pequeno | small | classifier | + |
| *-Vts-/-ts-* | pula, pure; hojarasca, hierba, plantas, ramas | pula, puree; leaf litter, grass, plants, branches | classifier | |
| *-V'tsore'-* | padrastro, madrastra | stepfather stepmother | classifier | |

## Appendix VI.    Wapishana classifiers

Although all classifiers were checked for cognnancy, only a subset of bound nouns were checked for cognnancy, as indicated below.

Source: (Santos 2006, 114-116)

| Form | Significado: Portugues | Meaning: English | Type | Physical classifier |
|---|---|---|---|---|
| *ap* | extensão | extension | classifier | + |
| *(a)ɽa* | falado | spoken | classifier | + |
| *b(iʔ)* | massa | pasta | classifier | + |
| *ɖa* | redondo | round | classifier | + |
| *ɖap* | habitação | housing | classifier | + |
| *iẓ, ɨ(ẓ)* | não discreto | not discreet | classifier | + |
| *puna* | extensão | extension | classifier | + |
| *ʃimak* | flexível | flexible | classifier | + |
| *ẓim* | fogo | fire | classifier | + |
| *ʔ(ɨ)* | partitivo, pare de | partitive, stop | classifier | + |
| *(a)ba* | feminino | feminine | compound noun | |
| *aib* | seiva | sap | compound noun | |
| *ak* | fruta | fruit | compound noun | |
| *(ana)b* | folha, palha | leaf, straw | compound noun | |
| *(a)p* | extensão | extension | compound noun | + |
| *b(a)* | parte de | part of | compound noun | + |
| *ba* | coletivo | collective | compound noun | |
| *baɽ* | superficie plana | flat surface | compound noun | + |
| *bi, b(iʔ)* | massa, pasta, particulas | dough, paste, particles | compound noun | + |
| *ɖa* | redondo | round | compound noun | + |
| *ɖakuɽi* | semelhante a galho, afluente (de rio) | branch-like, tributary (of river) | compound noun | + |
| *ɖaku* | boca, bico, extremidade | mouth, beak, end | compound noun | |
| *ɖani* | filho, pequeno | child, small | compound noun | + |
| *ɖap* | habitação | housing | compound noun | |
| *ɖaɽɨ* | pai, semelhante a pai, grande | father, father-like, great | compound noun | |
| *ɖin* | onça | jaguar | compound noun | |
| *ɖɨkɨu* | montanha | mountain | compound noun | |
| *ɖi, ʔi* | fino, flexíbel, linear | thin, flexible, linear | compound noun | + |
| *iɖ* | pena | feather | compound noun | |
| *iʃ(i)* | pêlo | fur | compound noun | |
| *(i)ẓ, (ɨ)ẓ* | líquido, em quantidade, cacho, não discreto | liquid, in quantity, bunch, not descrete | compound noun | + |
| *kin* | continente | continent | compound noun | |
| *kiɲa:* | instrumento | instrument | compound noun | |
| *kiẓi* | lugar | place | compound noun | |
| *maɖ* | couro, pele, casca | leather, skin, bark | compound noun | + |
| *suɖi* | pequeno | small | compound noun | + |
| *t* | longo | far away | compound noun | + |
| *tabaɨ* | perna, semelhante a perna, cabo | leg, leg-like, cable | compound noun | + |
| *uɽuɖ* | tronco, base | trunk, base | compound noun | |

| Form | Significado: Portugues | Meaning: English | Type | Physical classifier |
|------|------------------------|------------------|------|---------------------|
| *wauɖ* | ombro, semelhante a ombro, galho | shoulder, shoulder-like, branch | compound noun | |
| *ʃa* | fumaça, leve, esparso | smoke, light, sparse | compound noun | |
| *ʃiu* | duro, flor | hard, flower | compound noun | |
| *ɨːɖ* | semente, caroça | seed, pit | compound noun | |
| *ʐa* | úmido | damp | compound noun | |
| *ʐu* | oco | hollow | compound noun | + |
| *ʐuaɨ* | cabeça, topo | head, top | compound noun | + |

## Appendix VII.   Palikur classifiers

Only classifiers relating to physical form were checked for cognancy: these are indicated in the list below. There is no closed set of bound nouns used in compounds.

Source: (Aikhenvald and Green 1998, 462)

| Form | Gloss | Type | Physical classifier |
|------|-------|------|---------------------|
| *-p-ri* | animate masculine | numeral classifier | |
| *-p-ru* | animate feminine | numeral classifier | |
| *-tra/-tahr-* | linear: straight, curved, flexible: river, rope | numeral classifier | + |
| *-u/-so-* | equal dimensions: round/square: orange, box (circle) | numeral classifier | + |
| *-a/-sa-* | irregular dimensions: other shapes (oval, rectangular) and irregular: egg, basket, house, land | numeral classifier | + |
| *-mku/-muk-* | concave three dimensional: concave: bowl, canoe; metal objects: knife, ring, coin | numeral classifier | + |
| *-k/-ka-/-bu* | two dimensional: flat, nonextended: hammock, mat | numeral classifier | + |
| *-t/-ta-* | one dimensional: rigid, vertical: arrow, cigar | numeral classifier | + |
| *-iku/-rik-* | extended, with boundaries: perimeter, height, fire, field, hold, waterfall, pile | numeral classifier | + |
| *-kti/-kat-* | plants | numeral classifier | |
| *-imku/-say-* | wrapped | numeral classifier | |
| *-bru/-bohr-* | group | numeral classifier | |
| *-i/-i-* | series (spans of time) | numeral classifier | |
| *-twi/-tiw-* | cluster | numeral classifier | + |
| *-ki/-ki-* | tied bundle or strings | numeral classifier | + |
| *-ayh/-si-/-psi* | basketfuls | numeral classifier | + |
| *-uhri* | part | numeral classifier | + |
| *-bak/-bk-* | side | numeral classifier | + |
| *-biu/-biy-* | mouth, mouthful | numeral classifier | + |
| *-uku/-wok-* | hand, handful | numeral classifier | + |
| *-buka/'-buk* | linear shape | Verbal classifiers (stative/transitive) | + |
| *-pit/-pit* | irregular or round shape | Verbal classifiers (stative/transitive) | + |
| *-kiya/-kig* | pointed shape | Verbal classifiers (stative/transitive) | + |
| *-pewa/-peru* | branchlike shape | Verbal classifiers (stative/transitive) | + |
| *-apa/-ap* | concave, three dimensional shape | Verbal classifiers (stative/transitive) | + |
| *-boha/-bo* | two dimensions, flat | Verbal classifiers (stative/transitive) | + |
| *-min/-min* | one dimension, vertical | Verbal classifiers (stative/transitive) | + |

| Form | Gloss | Type | Physical classifier |
|---|---|---|---|
| *-muh/-muh* | side of object | Verbal classifiers (stative/transitive) | + |
| *-ekuh/-ik* | the inside part of the object | Verbal classifiers (stative/transitive) | + |
| *-kisa* | the edge of the object | Verbal classifiers (stative/transitive) | + |
| *-kat* | the trunk of the object | Verbal classifiers (stative/transitive) | + |
| *-duk/-duka* | chest | Verbal classifiers (body parts) | |
| *-kug/-kuga* | foot | Verbal classifiers (body parts) | |
| *-ok/-oka* | hand | Verbal classifiers (body parts) | |
| *-tiw/-tew* | head | Verbal classifiers (body parts) | |
| *-ot/-(h)ot(a)* | eye | Verbal classifiers (body parts) | |
| *-bi/-biya* | mouth | Verbal classifiers (body parts) | |
| *-tip/-tipa* | top (lid) | Verbal classifiers (body parts) | + |
| *-buhku(mna)* | linear shape | Locative classifiers | + |
| *-pit* | irregular or round shape | Locative classifiers | + |
| *-kigsa* | pointed shape | Locative classifiers | + |
| *-peru* | branchlike shape | Locative classifiers | + |
| *-kigbi(mna)* | sharp-edged shape | Locative classifiers | + |
| *-apa* | concave, three dimensional shape | Locative classifiers | + |
| *-madka* | two dimensions, flat | Locative classifiers | + |
| *-min* | one dimension, vertical | Locative classifiers | + |
| *-iku* | bounded: within a periphery, inside | Locative classifiers | + |
| *-bet* | unbounded: substances; otherwise unclassifiable items | Locative classifiers | + |
| *-hakwa* | water | Locative classifiers | + |
| *-vigku* | road, river | Locative classifiers | |
| *-pig* | 'pet' | Possessive classifiers | |
| *-mana* | 'food' | Possessive classifiers | |
| *-mutra* | 'plant' | Possessive classifiers | |

| Form | Gloss | Type | Physical classifier |
|------|-------|------|---------------------|
| *-win* | 'catch, animal caught to be eaten' | Possessive classifiers | |
| -kamkayh | 'child' | Possessive classifiers | |

# Appendix VIII. Resígaro classifiers

Only classifiers relating to physical form were checked for cognancy: these are indicated in the list below. There is no closed set of bound nouns used in compounds.

Source: (Allin 1976, 154-163)

| Form | Gloss | Type | Physical classifiers |
|------|-------|------|---------------------|
| -ʔootsí | "lungs" | classifier | |
| -ʔoohú | "metal or tin container" / "a room" | classifier | + |
| -ʔó | "longish and oval-shaped" | classifier | + |
| -ʔíǰó | "earthenware container for liquid" | classifier | + |
| -ʔiikó | "a new shoot" | classifier | |
| -ʔí | "bunch" | classifier | + |
| -ʔeté | "flower" | classifier | |
| -ʔéhú | "hole in the ground" | classifier | + |
| -ʔé | "trunk of a tree" | classifier | |
| -ʔásí | "central place" | classifier | |
| -ʔaaví | "side of" | classifier | + |
| -ʔaanú | "liquid" (uncountable) | classifier | + |
| -ʔaamí | "leaf-like" | classifier | |
| -vuudú | "log" | classifier | |
| -uuʔó | "rope-like" | classifier | |
| -ú | "spherical" | classifier | + |
| -píǰé | "human female" | classifier | |
| -pí | "liquid" (countable) | classifier | + |
| -pekó | "day" | classifier | |
| -pásí | "ring" | classifier | |
| -paahí | "hollow" | classifier | + |
| -mú | "tube-like" | classifier | + |
| -mokí | "dead" | classifier | |
| -mííʔo | "skin(-like)" | classifier | + |
| -ǰiiʔó | "stringy" | classifier | + |
| -ǰíhú | "dust" | classifier | + |
| -í | "stick-like" | classifier | + |
| -hú | "long and flat - horizontal" / "speech" | classifier | + |
| -hí | "round and flat" | classifier | + |
| -gú | "long and flat" | classifier | + |
| -gí | "human, male, and all non-human animates" | classifier | |
| -ʔúmí | "face" | classifier | |
| -ʔosí | "hand" | classifier | |
| -ʔípi | "machine" | classifier | |
| -ʔápo | "short cut" | classifier | |
| -ʔabaú | "shoulder" | classifier | |
| -ʔaapí | "arm" | classifier | |
| -uuʔá | "a part of" | classifier | + |
| -tuʔá | "foot" | classifier | |

| Form | Gloss | Type | Physical classifiers |
|---|---|---|---|
| *-tsiʔaaʔdó* | "shoe" | classifier | |
| *-pákó* | "honey" | classifier | |
| *-paǰí* | "field" | classifier | |
| *-mi* | "canoe" | classifier | |
| *-kuba* | "leg" | classifier | |
| *-kooʔú* | "broom" | classifier | |
| *-koomí* | "village" | classifier | |
| *-kó* | "a thick stick" | classifier | + |
| *-kaaʔdó* | "watering-place" | classifier | |
| *-ííʔo* | "long and oval-shaped" | classifier | + |
| *-hugí* | "path" | classifier | |
| *-gahú* | "above" | classifier | |
| *-gaaʔžó* | "raft, floating thing" | classifier | |
| *-dó* | "female" | classifier | |
| *-bú* | "made of liana cord" / "a felled tree" | classifier | |
| *-bahú* | "uninhabited part of the jungle" | classifier | |
| *-bábú* | "that which belongs to something" | classifier | |

# Appendix IX.  Yukuna classifiers

All classifiers from this set were considered in this study.

Source: (Lemus Serrano 2020, 57)

| Form | Gloss | Type |
|------|-------|------|
| *-a'jlu* | round recipient | classifier |
| *-a'ku* | concave | classifier |
| *-a'la* | round, whole | classifier |
| *-ja* | human | classifier |
| *-na* | big | classifier |
| *-ta* | flat | classifier |
| *-hila* | long | classifier |
| *-ipi* | pack | classifier |
| *-thá̱* | basket | classifier |
| *-i'lá* | string | classifier |
| *-a'pa* | part of a whole | classifier |
| *-i'jñe* | pack | classifier |
| *-a'ma* | whole | classifier |
| *-nachi* | animal separated from pack | classifier |

## Appendix X.    Tariana classifiers

Only classifiers relating to physical form were checked for cognacy.

Source: (Aikhenvald 2003, 89-92)

| Meaning | Noun cl sg | noun cl pl | num cl | poss cl | verbal cl |
|---|---|---|---|---|---|
| General animate | *-ite* | *-peni* | *-ita* | *-ite* | *-ita* |
| Human | *-ite* | *-peni* | *-hipa* | *-ite* | *-ita* |

| Form | Gloss | Type | Physical classifiers |
|---|---|---|---|
| *-ma* | feminine | Classifier | |
| *-da* | round objects | Classifier | + |
| *-hiwi* | thin long objects | Classifier | + |
| *-ipa* | big open space | Classifier | + |
| *-ipu* | long, hollow, bundle-like | Classifier | + |
| *-ita* | inanimate object, instrument | Classifier | |
| *-ku* | folded cloth | Classifier | + |
| *-kwa* | flat surface | Classifier | + |
| *-kwana* | plain | Classifier | |
| *-kwema* | flat and round | Classifier | + |
| *-kha* | curvilinear | Classifier | + |
| *-khi, -ki* | thin curved (rope, tape) | Classifier | + |
| *-maka* | extended cloth | Classifier | + |
| *-mapha* | completely covered | Classifier | + |
| *-na* | long vertical | Classifier | + |
| *-pa* | largish and long | Classifier | + |
| *-peku* | thin stretch | Classifier | + |
| *-pi* | long, thing, vertical; cycle of time | Classifier | + |
| *-pukwi* | round and hollow | Classifier | + |
| *-puna* | stretch of a road, road | Classifier | |
| *-phe* | leaflike | Classifier | |
| *-aphi* | smaller, hollow | Classifier | + |
| *-hwi* | particles, small thin objects | Classifier | + |
| *-yawa* | holes | Classifier | + |
| *-(i)tʃi* | bundle | Classifier | + |
| *-ima* | a paired object; one side of two | Classifier | + |
| *-iphina* | a quarter | Classifier | + |
| *-pada* | piece (evenly cut) | Classifier | + |
| *-piu, -pio, -pia* | time | Classifier | |
| *-sawa* | group | Classifier | |
| *-yami* | piece (unevenly cut or torn off) | Classifier | + |
| *-dapana* | habitation | Classifier | |
| *-whya* | canoe, transport | Classifier | |
| *-anhi* | line of a song | Classifier | |
| *-dawa* | corner, limited space | Classifier | + |

| Form | Gloss | Type | Physical classifiers |
|---|---|---|---|
| *-depita* | night | Classifier | |
| *-hipita* | land | Classifier | |
| *-ithi* | seed | Classifier | |
| *-iwai* | trap, wall | Classifier | |
| *-kada* | a day | Classifier | |
| *-kairathe* | heap of stones | Classifier | + |
| *-kawa* | leg, handle, anything leg-like | Classifier | + |
| *-kena* | branch | Classifier | + |
| *-kiyere, -kyere* | island | Classifier | |
| *-kuda* | trunk | Classifier | |
| *-kuda-mi* | cut-off trunk | Classifier | |
| *-kuma* | smoke-like; custom; level (generation; floor (of a house)) | Classifier | |
| *-kuya* | extended part | Classifier | + |
| *-kuthe* | manioc bread | Classifier | |
| *-kuru* | mist, smoke | Classifier | + |
| *-khay, khe* | thick jungle | Classifier | |
| *-lima* | place with certain type of plant | Classifier | |
| *-nai* | lake | Classifier | |
| *-naki* | long handle | Classifier | + |
| *-numa(na)* | word, language | Classifier | |
| *-ñapi* | bone | Classifier | |
| *-nhe* | verse, line | Classifier | |
| *-patawa* | gorge | Classifier | |
| *-peda* | low bush, small heap | Classifier | |
| *-pida* | crown and branches of a tree | Classifier | |
| *-pina* | swamp | Classifier | |
| *-pua* | river | Classifier | |
| *-purikuda* | mountain | Classifier | |
| *-pukuda* | bush | Classifier | |
| *-pukuipe* | turn | Classifier | |
| *-pusita* | clearing | Classifier | |
| *-taku* | endpoint | Classifier | |
| *-turapa* | cone | Classifier | |
| *-the* | knot | Classifier | |
| *-ithepu* | bow | Classifier | |
| *-thiwa* | ravine | Classifier | |
| *-wa* | heap of stones | Classifier | |
| *-wata* | -ful | Classifier | |
| *-wathe* | joint | Classifier | |
| *-wi* | trap, wall | Classifier | |
| *-wina* | pile | Classifier | + |
| *-yapi* | time span | Classifier | + |

## Appendix XI.     Warekena classifiers

All classifiers were considered for cognacy.

Source: (Aikhenvald 1998, 299)

| Gloss | with 'one' | with 'two' |
|---|---|---|
| human masculine | *peya* | *enaba* |
| human feminine | *peya* | *tuwanaba* |
| animals | *pamiña* | *pamiñanaba* |
| fish | *peɺeyaɺu* | *eɺenaba* |
| curvilinear objects | *pauɺiaɺuni* | *enaba* |
| periods of time | *babuya* | *bunaba* |