

## Morphology in Arawak languages

Alexandra Y. Aikhenvald

### Summary

The Arawak language family is the largest in South America in terms of its geographical spread, from Central America (Belize, Honduras, Guatemala, and Nicaragua) to as far south as Bolivia (and formerly Argentina and Paraguay). Within South America, Arawak languages are spoken in Lowland Amazonia and adjacent regions, covering Guyana, French Guiana, Surinam, Venezuela, Colombia, Peru and Brazil, in at least ten locations north of the River Amazon, and at least ten to the south of it. There are over forty extant languages and a few dozen extinct ones. The genetic unity of Arawak languages was first recognized by Father Filippo Salvatore Gilij as early as 1783, based on a comparison of pronominal prefixes in Maipure, a now extinct language from the Orinoco Valley, and in Moxo from Bolivia. The limits of the family were established by the early twentieth century. Proposals to include Arawak languages in spurious macro-groupings such as 'Arawakan' or 'Macro-Equatorial' have proved spurious and unsubstantiated. The heritage of Arawak languages survives in such common words as *hammock*, *hurricane*, *barbecue*, *guava* and *tobacco*.

Arawak languages are synthetic, predominantly head-marking and suffixing, with a closed and historically stable set of prefixes — bound pronouns on verbs, the relativising prefix *ka-* and its negative counterpart *ma-*. Personal prefixes distinguish first, second, and third person, and also impersonal and indefinite forms. Prefixes mark the subject of a transitive verb and of an intransitive active verb, and the possessor on nouns. In at least two thirds of the languages, personal suffixes or enclitics express the object of a transitive verb (O), and the subject of stative verbs (So) and/or the subject of non-verbal predicates. A few highly synthetic languages (including those from the Kampa subgroup in Peru) employ suffixes or enclitics to cross-reference the object and also the recipient or an oblique. There is typically a number of locative cases which can be stacked in one word. The majority of Arawak languages do not employ cases for marking core grammatical relations. The only exception is Tariana, from the multilingual Vaupés River Basin Linguistic area. Here, core cases were developed under the influence of the neighbouring Tucanoan languages. Inclusive-exclusive distinctions were developed in Resígaro and Palikur as a result of language contact.

Open classes are verbs and nouns; adjectives tend to form an open class, and share some features with nouns, and some with verbs. Verbal roots tend to be exclusively monosyllabic. Noun roots can contain more than one syllable. Derivational processes include affixation, compounding and various kinds of reduplication. Just a few languages have single-word serial verb constructions. The order of suffixes within a word can be variable, reflecting the scope of the morphemes.

Nouns divide into obligatorily, or inalienably, possessed and optionally, or alienably, possessed. Obligatorily possessed nouns are body parts, kinship terms and a few important possessions, e.g. 'name' and 'house'. If the possessor is not specified, these nouns take an unpossessed form marked with a suffix also used as a nominalizer on verbs. Alienably possessed nouns take a possessive prefix and an additional suffix (chosen based on the meaning of the noun). Most languages distinguish masculine and feminine genders in third person singular personal pronouns, demonstratives, nominalizations, and also as agreement markers on adjectives. More than half the languages have complex systems of classifiers on number words, and also on verbs, in possessive constructions, and on nouns themselves. They categorise the noun in terms of its shape, consistency, and animacy. Singular and plural numbers are fairly uniform across the family; dual has developed in Resígaro, as a consequence of language contact with the unrelated Bora. Other nominal grammatical categories include nominal tense, augmentative, diminutive, and approximative.

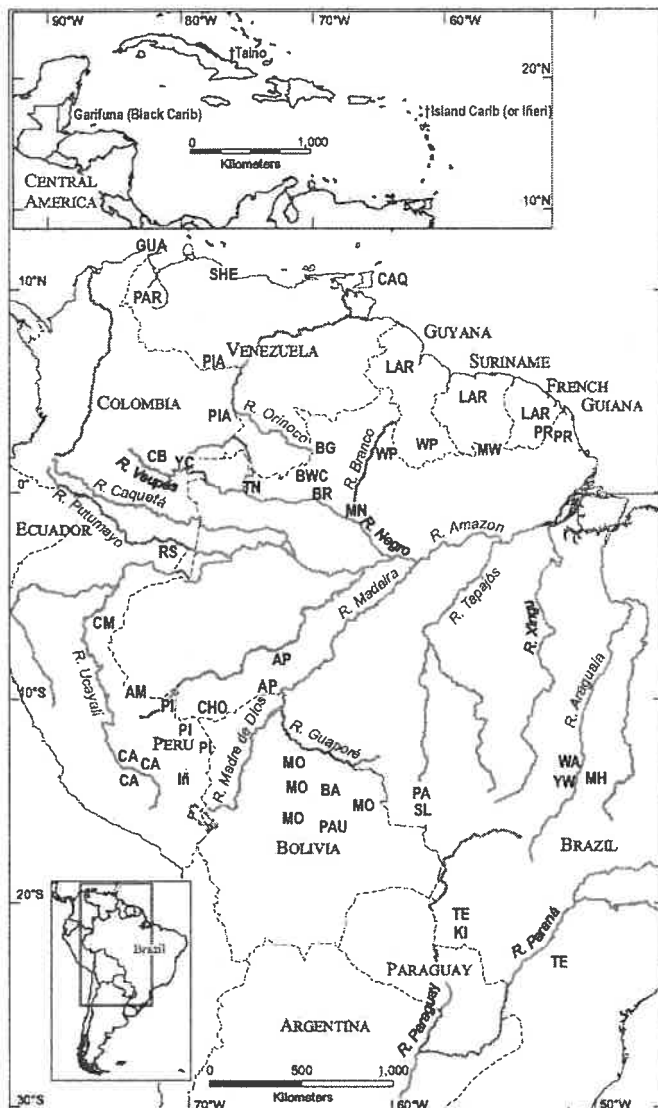
The verb is the most complicated part of the grammar of every Arawak language, and the only obligatory constituent in a clause. Typical verbal categories include tense, aspect, evidentiality, numerous modalities (including a frustrative meaning 'do in vain'), and valency changing derivations — passives, reflexives, reciprocals, causatives, and applicatives. Some Kampa languages have up to six applicative derivations, including comitative, benefactive, goal, presential, separative, and instrumental. Highly synthetic languages, such as Kampa and Palikur, have patterns of noun incorporation.

Many Arawak languages are located next to speakers of languages from other families. They take on their features, in grammar and sometimes also in lexicon. Tariana, the only Arawak language spoken in the multilingual Vaupés River Basin area surrounded by Tucanoan languages, has a distinct Tucanoan flavour to its grammar. Mawayana, Garifuna and Palikur, in contact with Carib languages, have acquired a few Carib features. Resígaro has been affected by Bora, and Amuesha bears traces of contact with Quechua and other, now hard-to-identify languages. The interaction of genetic inheritance, language contact, and independent innovations makes Arawak languages dauntingly diverse.

**Keywords:** contact linguistics, historical linguistics, synthetic languages, applicative, possession, Vaupés River Basin area, Tariana, Resígaro, Palikur, Kampa, Amuesha

**Arawak languages at a glance**

The Arawak language family is the largest in South America in terms of its geographical spread, with over forty extant languages and a few dozen extinct ones. Arawak languages are spoken in at least ten locations north of the River Amazon, and at least ten to the south of it (see Map). Arawak languages are spoken in Lowland Amazonia and beyond it, covering Guyana, French Guiana, Suriname, Venezuela, Colombia, Peru, Brazil and Bolivia, and formerly Paraguay and Argentina. Guajiro (or Wayyu-naiki) spoken around the Guajiro peninsula in Venezuela and Colombia is the largest language of the family. Garifuna is the only Arawak language spoken in Belize, Honduras, Nicaragua and Guatemala in Central America. Groups of Arawak speakers must have migrated from the Caribbean coast to the Antilles a few hundred years before the European conquest. The first native American peoples encountered by Columbus on the island of Hispaniola (the coast of modern Haiti) in 1492 were the Taino, speakers of an Arawak language, which became extinct within a hundred years of the invasion. Spanish and many other European languages inherited quite a few loans from the Arawak languages of the Caribbean, including Taino — their legacy survives in many common English words, including *hammock*, *hurricane*, *barbecue*, *iguana*, *maize*, *papaya*, *savanna*, and *guava*.



Map. Extant Arawak languages (Aikhenvald, 2012: 34-35; @Alexandra Aikhenvald)  
**Key to the Map:**

AM - Amuesha (or Yanasha')	MN - †Manao
AP - Apurina	MO - Mojo (Ignaciano, Trinitário)
BA - Baure	MW - Mawayana
BG - Baniva of Guainia (with closely related †Yavitero and Warekena of Xié)	PA - Paresi (or Haliti)
BR - †Bare	PAR - Parauhano (or Añun)
BWC - Baniwa of Içana/Kurripaco	PAU - Paiconeca, Paunaca
CA - Kampa languages (Machiguenga, Nanti, Nomatsiguenga, Ashaninca, Asheninca, Pichis. Alto Perené, Pajonal)	PI - Piro (or Yine)
CAQ - Caquetio	PIA - Piapoco (or Dzase)
CB - Cabiari (or Kawyari)	PR - Palikur
CHO - Chontaquiro	RS - Resígaro
CM - Chamicuro	SHE - Shebayo
GUA - Guajiro (or Wayyu-naiki)	SL - Salumã (or Enawenê-Nawê)
Iñ - Iñapari	TE - Terêna
KI - Kinikinao	TN - Tariana
LAR - Lokono Arawak	WA - Waura or Waujá (Xinguan Arawak)
MH - Mehinaku (Xinguan Arawak)	WP - Wapishana
	YC - Yucuna
	YW - Yawalapiti (Xinguan Arawak)

The highest concentration of recorded Arawak languages is found in the region between the Rio Negro and the Orinoco. This is potentially a strong linguistic argument in favour of the Arawak proto-home having been located there. This hypothesis is corroborated by the few mythical traditions of northern origin of Arawak-speaking peoples south of the Amazon. However, the diversity of Arawak languages south of the Amazon in central Peru and around the Rivers Purús and Madeira in Brazil must have been greater in the past than it is now. Due to mass extinction of languages this is hard to appreciate (see Lathrap, 1970, and a summary with references in Aikhenvald, 2013). The distribution of extant languages is shown in the Map.<sup>1</sup>

Arawak languages are synthetic, predominantly head-marking and suffixing, with a small and historically stable set of prefixes. All Arawak languages are predominantly agglutinating, with some fusion. Arawak languages spoken south of the Amazon — such as Kampa languages, the languages of Xingu, and South Arawak languages — tend to be highly synthetic, with intricate patterns of noun incorporation and long strings of verbal affixes, including markers of changing valency, aspect, and modality. Arawak languages spoken to the north of the Amazon including those of the Upper Rio Negro and the Orinoco basins tend to be less synthetic, with less complicated verb structure.

## 1 Word structure and word formation

### 1.1 Word structure

A grammatical word in Arawak languages contains at least one root morpheme. Obligatorily possessed nouns, and transitive and active intransitive verbs take a prefix indicating the person of the possessor or the subject, A/Sa. The majority of prefixes and suffixes across Arawak languages are monosyllabic. Disyllabic suffixes often go back to grammaticalised free forms, e.g. Baniwa of Içana, Tariana *-maka* 'classifier for stretchable extended objects, cloth' (Aikhenvald, 2002: 307-8), Apurinã *-maka* 'class term for soft, flat items', *māka* 'clothes' (Facundes, 2000), Paresi-Haliti *maka* 'hammock' (Brandão, 2014: 176), Bahuana *himaka* 'hammock' (Ramirez, 1992). Roots can consist of one or of two syllables; many verbal roots which go back to the proto-language are monosyllabic.

## 1.2 Morphological processes

Affixation is the most common process of word formation. Arawak languages have a small set of prefixes. These include bound pronouns on verbs (shown in Table 1 in §2.1), the attributive prefix *ka-* and its negative counterpart *ma-*. Suffixes of all kinds are less stable across the family; there is considerable diversity even across closely related languages. A few languages display circumfixation, infixation, and apophony. Suppletion is limited. Suffixes may display variable order, depending on their scope.

### 1.2.1 Prefixes and prefix positions

In many Arawak languages, the attributive-relative prefix *ka-* and its negative counterpart *ma-* derive an adjective from a noun, as in Piro *ka-yhi* (ATTR-tooth) 'having teeth', *ma-yhi* (NEG-tooth) 'toothless' (Matteson, 1965: 119), or a noun from a noun, e.g. Bare *ka-witi-w* (ATTR-eye-FEM) 'a woman with good eyes', *ma-witi-w* 'a woman with bad eyes; blind'. Prefix *ka-* often derives stative verbs with possessive meaning from obligatorily possessed nouns, e.g. Paresi-Haliti *iyane* 'husband', *ka-iyane* (ATTR-husband) 'get married (of a woman)', Tariana *-sa-niri* (spouse-masc) 'husband', *ka-sa-niri* (ATTR-spouse-masc) 'get married (of a woman)' (Brandão, 2014: 177, 329-330; Aikhenvald, 2003: 410-411).

A personal prefix may precede *ka-* or *ma-* in its derivational function, e.g. Paresi-Haliti *wi-ma-maha-za* (1pl-NEG-honey-POSSESSED) 'we had no honey'. In Tariana this is only possible if the stative verb is transitivised with the causative suffix, as in the following derivational chain: *-weni* 'price' > *ka-weni* (ATTR-price) 'pricey, expensive' > *wa-ka-weni-ta* (1pl-ATTR-price-CAUS) 'we pay'. In a few languages, the prefix *ka-* marks the predicate of relative clauses, e.g. Tariana *ka-inu* (ATTR-kill) 'the one who kills; hunter'.

In some languages, such as Resígaro, a moribund language spoken in the Putumayo basin in Colombia, the reflexes of *ka-* occur fossilized with just a few verb roots, e.g. *kemu* 'talk' (cognate to Warekena *-ma*, Piapoco *-ima* 'talk'); or *kainee* 'die, kill' (cognate to Baniwa and Tariana *-inu(a)*, Piapoco *-nua*, Achagua *-inua*, Wapishana *ian*, Iñapari *-ino* 'kill'). The prefix has been lost in Baniva of Guainia, Yavitero and Warekena of Xié. Its negative counterpart *ma-* tends to be productive in most languages. In a few, it is all but gone. For instance, in Warekena of Xié the negative prefix *ma-* occurs in a few intransitive verbs with inherently negative meaning, *ma-kare* (NEG-breath) 'be out of breath, breathless' and *ma-numa* (NEG-mouth) 'to be silent, mouthless'.

Causative and applicative prefixes on verbs are a feature of highly synthetic languages south of the Amazon — Paresi-Haliti, Xinguan, the languages of Peru (especially Kampa), and South Arawak, and also Palikur and Garifuna north of the Amazon. Examples include Paresi-Haliti, Piro, Apurinã, Baure and Ignaciano *ka-* 'causative, transitiviser', Palikur, Garifuna, Paresi-Haliti, Waujá, Bauiré, Ignaciano *a-/e-/i-* 'causative, transitiviser', Terêna *i-* 'causative'. Causative prefixes in the languages from the Kampa subgroup include Machiguenga *ogi-/og-/o-*, *omi(n)*, Asheninka Kampa *omi(n)-*, *oi-*, *o-*, *ow-*, Alto Perené *oi-*, *i-*, *v-*; Alto Perené has an additional malefactive causative prefix *mi-* (Mihás, 2015: 302-3, Payne, 2002, Wise, 1990a, b). The prefixes *him-/hi-* in Piro and *mi-* in Apurina mark the comitative on verbs (Wise, 1990a, b; Facundes, 2000).<sup>2</sup>

South Arawak languages, Kampa, Paresi-Haliti and Xinguan Arawak allow two prefix positions in a verbal grammatical word. Three prefix positions are a feature of a number of Kampa languages. For instance, in Nanti the person prefix can be followed by the marker of irrealis and a causative prefix, in this order. In other Kampa languages, including Alto Perené, an irrealis and a causative prefix are mutually exclusive, reducing the number of prefix positions to just two (Mihás, 2017: 795). Valency-changing prefixes are always placed close to the verbal root. The person prefix occupies the left-most position in the grammatical word.



palatalization in Alto Perené expresses diminution, e.g. *no-shinto-paye* (1sg-daughter-PL) 'my daughters', *nu-shintyo* (1sg-daughter) 'my little daughter' (Mihas, 2015: 437; 83).

Suppletion is rare and is generally restricted to possessed forms of some nouns, e.g. Terêna *cavâne* 'garden'; *issâne* 'his garden'; Mehinaku *amaka* 'hammock', *a-tunumala* 'his hammock', Baure *mon* 'sister-in-law' (non-possessed), *-tor* 'sister-in-law' (possessed), *wotoki* 'hammock' (non-possessed), *-imok* 'hammock' (possessed), Paresi-Haliti *irika-ti* (fire-NPOSS) 'fire', *in-itima* 'his fire', *hare* 'son', *i-tyani* 'his son' (Butler and Ekdahl, 1979; Medeiros, 1990; Danielsen, 2007: 122; Brandão, 2014: 167). In Asheninka and Ashaninka (both Kampa) possessed forms of some nouns display partial suppletion, e.g. Ashaninka *no-ovi-te* (1sg-pot-POSS) 'my pot', versus *koviti* 'pot' (non-possessed) (see Mihas, 2017: 788; 2015: 91-92). Vocative forms of kinship terms in Tariana are suppletive, e.g. *haniri* 'father', *paika* 'father: vocative'. In Tariana, the size adjective *hanu-* 'big' has a suppletive plural form: the stem *male* 'thick, heavy' is used as its plural counterpart, e.g. *hanu-da* (big-CL:ROUND) 'big round thing', *male-da-pe* (thick-CL:ROUND-pl) 'big round things'. The form *\*hanu-da-pe* is ungrammatical (if the plural marker *-pe* attaches directly to the adjective stem *hanu-*, the resulting form is a quantifier *hanu-pe* meaning 'many': Aikhenvald, 2003: 173).

In one case partial suppletion of a possessed form may go back to proto-Arawak. The common Arawak term for 'house, home' typically displays an alternation *pelpan*, cf. Terêna *péti*, possessed *pêno* 'house'; Ignaciano *peti*, possessed *pena*; Baniwa *pan-tti*, possessed form *-pana*; Palikur *payt*, possessed *-pi-*; Bahwana *φani-ci*, possessed *-ana*; Paresi *hati*, possessed *-hana* 'house', Baure *hari*, possessed *-wer* (Aikhenvald, 1999: 81; Brandão, 2014: 167; Danielsen, 2007: 122).

### 1.2.3 Variable ordering of suffixes

The order of suffixes within a grammatical word may be variable.<sup>3</sup> The semantic effect will then reflect the scope of an affix. Variable position of the causative suffix *-aka(g)* in Ashéninka, from the Kampa subgroup, correlates with the scope of causation. In (2), the habitual marker *-apiint* follows the causative. The causative, and not the caused action, is within the scope of the habitual. The meaning is 'constantly make cut'.

- (2) pi-n-chek-[aka-apiint]-e-ri Asheninka  
 2person-IRR-cut-CAUS-HAB-IRR-3mascO  
 'You should constantly be making him cut'

In (3), the habitual marker precedes the causative. The caused action ('cut') rather than the causation is 'habitual': the meaning of the form is 'make constantly cut' (Payne, 2002: 492).

- (3) pi-n-[chek-apiint-aka]-e-ri Asheninka  
 2person-IRR-cut-HAB-CAUS-IRR-3mascO  
 'You should make him constantly cut'

In Alto Perené, another Kampa language, a suffix may have different meanings depending on its place within the verbal word. The durative suffix *-vai* occurs close to the verbal root when it has intensifying function; in its aspectual durative meaning, it follows other suffixes, such as the frustrative *-ve-* 'do in vain' (Mihas, 2015: 104).<sup>4</sup>

Classifiers employed as derivational suffixes on Tariana nouns can occur in different order, with a semantic effect. The word *nu-kapi* (1sg-hand) means 'my hand' in general. The 'side of my hand' is *nu-kapi-ma* (1sg-hand-CL.SIDE.OF). The round part of the palm itself is *nu-kapi-ma-da* (1sg-hand-CL.SIDE.OF-ROUND) 'one palm of my hand'. The word for 'finger' is *nu-kapi-da* (1sg-hand-CL:ROUND.SMALL), literally, the small and round one of my hand. To



refer to the side of a finger, one will say *nu-kapi-da-ma* (1sg-hand-ROUND-CL:SIDE.OF), reversing the order of morphemes *-ma-* and *-da* (see Aikhenvald, 2017b).

#### 1.2.4 Reduplication

In a number of Arawak languages, reduplication is used with verbs to mark intensive, repetitive, or continued action, or plurality of participants. In Ashaninka Tambo, a Kampa language, the disyllabic roots *sheva* 'move' undergoes full reduplication to describe continuous shaking movement of a sitting female (Mihas, 2017: 787).

- (4) o-shevasheva~t-ak-a                      o-saik-ak-e                      kipatsi-ki    *Ashaninka Tambo*  
 3NM.S-move.body~EP-PFV-REAL    3NM.S-be.at-PFV-REAL    ground-LOC  
 'she was shaking, sitting on the ground'

Reduplication of the stem initial CV- is a feature of Terêna and Palikur; Waujá, Bahuana and also Terêna display CVCV partial reduplication. In Baure, the stem final -CV(C) can be reduplicated, e.g. *hiš* 'pull', *hiš-hiš* 'pull quickly' (Danielsen, 2007: 96); similar patterns are found in Ignaciano and Warekena of Xié.

In general, reduplication is a feature of verbs. In Baure it also applies to adjectives. Some languages have several kinds of reduplication with different meanings, e.g. Terêna CV-: *há-harara kóye* 'it is very red'; CVCV-: *tuvo'tuvo kóye* '(it) is moving many times' (Ekdahl and Grimes, 1964, Ekdahl and Butler, 1979).

Reduplication of the stem-final CV syllable in Warekena of Xié has different meanings dependent on the type of verb and its transitivity (Aikhenvald, 1998). Reduplication applies to all verbs (except those with a monosyllabic root, e.g. *ma* 'say, do', *we* 'leave'). Reduplication with transitive verbs may indicate (a) intensity of action, e.g. *ni-wayata-ta* (3pl-speak-RED) 'they talked for a long time', (b) multiplicity of O; (c) the multiplicity of A. If the verb is A=S ambitransitive, reduplication indicates unusual multiplicity of subjects (accompanied with a plural marker, otherwise optional) (see Aikhenvald and Dixon, 2011 on the semantics of reduplication and its correlations with grammatical relations). The unreduplicated verb in its transitive use is at (5). The object, *yudubu*, can be omitted, and then the verb will be used intransitively.

- (5) ema    ni-fuwiya    (yudubu)                                      *Warekena of Xié*  
 tapir    3pl-whistle    (song)  
 'Tapirs whistle (a song)'

In (6), the reduplicated verb refers to many tapirs whistling.

- (6) ni-fuwiya-ya                      ema-nawi                                      *Warekena of Xié*  
 3pl-whistle-RED                      tapir-PL  
 'Many tapirs are whistling.'

In strictly transitive verbs, reduplication refers to a multiplicity of O. (7) shows a strictly transitive verb *-eda* 'perceive, see, hear, feel', unreduplicated.

- (7) nida                      yuřana                      utřipie                      ni-wayata                      *Warekena of Xié*  
 3pl+perceive    voice                      bird                      3pl-speak  
 'They can hear the voice of birds singing'

In (8) the same verb *-eda* occurs reduplicated, reflecting the multiplicity of objects. This example cannot be interpreted as 'very many people can hear the voice of birds'.

- (8)    *nida-da-hã*                    *yujana*                    *utjipie ni-wayata*                    *Warekena of Xié*  
       3pl+perceive-PAUS    voice                    bird    3pl-speak  
       'They can hear the voice of very many birds singing'

The causative marker *-ta* can also undergo reduplication, with an implication of intensive causation, as in *kune-ta* (fear-CAUS) 'scare', *kune-ta-ta* (fear-CAUS-RED) 'scare very much'. Reduplication of stative intransitive verbs indicates the intensity of the state or condition, e.g. *apate* 'be cold', *apate-te* 'be very cold'; *akune* 'be dangerous', *akune-ne* 'be very dangerous'. Reduplication of intransitive active verbs indicates the multiplicity of subject referents, and multiplicity of action. In (9), reduplication of the final syllable of the active intransitive verb *-mita* 'fly' indicates that the action was performed many times by many birds.

- (9)    *utjipie*                    *ni-mita-ta-wa*                    *Warekena of Xié*  
       bird                    3pl-fly-RED-NONACCOMPLISHED  
       'Very many birds are flying (to and fro)'

A reduplicated verb may have a somewhat unpredictable meaning compared with the corresponding non-reduplicated one, as in *mutjita* 'bite', *mutjita-ta* 'gnaw', *tapa* 'go, pass', *tapa-pa* 'walk'; *wiyua* 'die, faint', *wiyu-yua* 'faint of drunkenness'.

Inherently reduplicated nouns are typically animal names which are onomatopoeic in origin, e.g. Baure *sorisori*, Baniwa of Içana *bokokoli*, Tariana *pupuli* 'owl', Tariana *kerekere* 'sparrow' (Danielsen, 2007: 96, Ramirez, 2001, Aikhenvald, 2003). Reduplication and even triplication has been described for ideophones and expressives in Kampa languages (e.g. Mihas, 2012).

### 1.2.5 Compounding

Arawak languages tend to have productive noun compounding, predominantly of endocentric type with part-whole or genitive relationship between the nominal components, e.g. Alto Perené *inka-nija* (rain-water) 'rainwater', *bametanitsi-panko* (teaching-house) 'school' (Mihas, 2015: 402; forthcoming).

Nominal compounds can contain bound (or obligatorily possessed) and free nouns. An example of a combination of two free nouns in Paresi-Haliti is *kamae kokoiya* (sun hawk) 'harpy eagle'; *zohitya-kate* (cashew-tree.of) 'cashew tree' contains a free and a bound noun, and *e-tsiri n-ahi* (3sg-head.of-LINKER-bone.of) 'his skull, lit. bone of the head' contains two bound nouns connected with a linker (Brandão, 2014: 177-8). In Piapoco, from the Uapuí subgroup of North Arawak, compounds always contain bound nouns, e.g. *i-üwi-ápuná* (3masc.sg-ear-path) 'his ear canal' (Klumpp, 1990: 166-7).

In numerous languages with multiple classifier systems (§3.3), some of the classifiers may have the same segmental form as bound or obligatorily possessed nouns. For instance, the classifier *-ya* 'skin, extended items' in Baniwa of Içana, Kurripako and Tariana — used on nouns, with adjectives, number words, and in possessive constructions — has the same form as the obligatorily possessed noun *-(i)ya* 'skin' (see Aikhenvald, 2007b; forthcoming on North Arawak languages; Mihas, 2015: 414 and forthcoming on Kampa languages, Danielsen, 2007: 96-100 on Baure).<sup>5</sup>

In Northern Kampa languages, compounding process can occur recursively, and a compound can consist of up to four roots, as in Alto Perené *incha-shi-tonki-ro-nto-shi* (plant-leaf-bone-NMLZ-NMLZ.NM-leaf) 'the herb used for making a female fat (lit. 'the herb for the

bony females', which is fed to the girls after they are placed in a menarche hut at the onset of their first menses) (Mihas, 2017: 789).

Verbal compounding tends to be restricted in its productivity. In Baure, verb compounds are limited to just one type, with *ke* 'do, say, go', as the first verb, and another verb as the second verb in the compound. Some verbal affixes, including the directional/intentional *-pik*, may have grammaticalized from the verb *-pik* 'come' (Admiraal and Danielsen, 2014: 108). A number of verbal enclitics marking aspect and manner of action in Tariana have grammaticalised within verb-compounding structures under the influence of structurally similar patterns in the neighbouring unrelated East Tucanoan languages, e.g. Tariana *-sita* 'finish, manage to do', =*sita*, *sta* 'completive marker', *whyume* 'be last', =*whyume* 'lastly, last time' (see Aikhenvald, 2016b, for a summary).

Semi-productive verb-verb compounds are a feature of some Kampa languages. In Alto Perené, the second component in a verbal compound provides purpose or manner specification to the whole verb, e.g. *-menko-tashi-* (construct.platform-roast) 'construct a grilling platform for cooking meat'. Many compounds are non-compositional, e.g. *-ñ-a-pinhatsa-* (see-EPENTHETIC-obey) 'respect', *-ñ-a-shintsti-* (see-EPENTHETIC-be.strong) 'have contempt' (Mihas, 2015: 159-60). Compounds in Northern Kampa languages may involve adjectives, adverbs, and members of closed classes, such as interrogative pronouns.

Productive incorporation of nouns into a verb is a feature of many languages south of the Amazon and also of Palikur, spoken in the Brazilian state of Amapá and French Guiana — §4.4.

## 2 Word classes

### 2.1 Nouns, verbs, and adjectives

All Arawak languages have open classes of nouns and verbs. Inflectional categories of nouns are discussed in §3; categories of verbs are addressed in §4. A common and stable feature of Arawak languages is marking participants — including subjects and in many instances, also objects and other grammatical roles — on the verb with bound morphemes (see Aikhenvald, 1999: 88-9, 2018a; Mihas, 2017). Prefixes mark the subject of a transitive verb (A) and of an intransitive active verb (S<sub>a</sub>), and the possessor on nouns. In at least two thirds of the languages, personal suffixes or enclitics express the object of a transitive verb (O), and the subject of stative verbs (S<sub>o</sub>) and/or the subject of non-verbal predicates.<sup>6</sup> A few highly synthetic languages (including those from the Kampa subgroup in Peru) employ suffixes to cross-reference the object and also the recipient or an oblique. The majority of Arawak languages (with the exception of Tariana) do not employ cases for marking grammatical relations.

**Table 1 A composite statement of A/S<sub>a</sub>/possessor prefixes and O/S<sub>o</sub> suffixes/enclitics in Arawak languages (Aikhenvald, 1999: 83; 2002: 289; 2018a)**

	prefixes		suffixes	
	sg	pl	sg	pl
person				
1	* <i>nu-</i> or * <i>ta-</i>	* <i>wa-</i>	* <i>-n(u),-na</i> * <i>-te</i>	* <i>-wa</i>
2	* <i>pi-</i>	* <i>(h)i</i>	* <i>-pi</i>	* <i>-hi</i>
3nfem	* <i>ri-</i> , <i>i-</i>	* <i>na-</i>	* <i>-ri</i> , <i>-i</i>	* <i>-na</i>
3fem	* <i>thu-</i> , <i>ru-</i> , <i>u-</i>	* <i>na-</i>	* <i>-thu</i> , <i>-ru</i> , <i>-u</i>	* <i>-na</i>
'impersonal' /reflexive	* <i>pa-</i>			
indefinite or non- focused A/S <sub>a</sub>	* <i>i-</i> , <i>a-</i> (?)			
general S <sub>o</sub> /O	-	-	* <i>-ni</i>	

The system in Table 1 distinguishes two numbers (singular and plural) and three persons (first, second and third, with feminine and non-feminine gender distinctions in third person singular in the majority of languages). The form *ta-/te* for first person singular pronoun is one of the defining features of the *Ta-Arawak* languages (mentioned in note 1). The person marking system includes values additional to three persons. The polysemous prefix *pa-* used on nouns and on verbs has impersonal meaning 'one, generic person' in several languages of the Upper Rio Negro-Orinoco Basin (including Tariana, Baniwa of Içana-Kurripako dialect continuum, Warekena of Xié, Bare, Cabiari and Maipure). Yavitero and Baniva of Guiana, two closely related languages spoken within the same area, add to this the meaning of coreferential possessor (on nouns). In Mawayana and Wapishana, in the Basin of Rio Branco, the Arawak languages of Xingu, and Paresi-Haliti, the reflexes of the prefix *pa-* occur only on nouns with the meaning of coreferential possessor, 'one's own'. The same prefix marks third person in Bahuana and possibly in Manao. In Yucuna and in Palikur, *pa-* has a reciprocal meaning; and in Guarekena the same prefix is used as a third person feminine marker (see Aikhenvald, 2018a for the discussion of the sources, and the historical developments of *pa-*; Green and Green, 2016: 206 on Palikur).

A number of Arawak languages (see Aikhenvald, 1999: 88; 1995a) have a further vocalic prefix *\*i-* or *\*a-* which marks a non-specified or focussed subject (A/S<sub>a</sub>) and unspecified, or indefinite, possessor. There is also some evidence in favour of a general or dummy suffix (or enclitic) *\*-ni* marking S<sub>o</sub>/O, attested across the family, whose origin and development is a matter for further study. The first person plural prefix in Terêna and in a number of Kampa languages has developed inclusive reference (see Aikhenvald, 2018b: 20-21 for some details).

In just a few languages south of the Amazon, the forms of subject prefixes on verbs and on nouns have minor differences. For instance, the prefix *pa-* in Amuesha is used only as a third person possessor marking on nouns but not on verbs. In Nomatsiguenga it marks third person feminine subject just on verbs. In this language, non-feminine subject-marking prefixes on verbs and possessor-marking prefixes on nouns have a somewhat different set of allomorphs (see Shaver, 1996: 34-5, and also Wise, 1971, for a full statement).<sup>7</sup>

Composition of the adjective class varies. In many languages adjectives share properties both with nouns and with verbs, in addition to special features of their own. In Baure, adjectives are an open class, similar to nouns in that they can occur with plural, diminutive, augmentative, and locative suffixes. Similarly to verbs, adjectives can undergo reduplication to mark emphasis, and occur with the approximative marker *-so-* 'about, more or less'; they can be nominalised with the same marker as the verbs. Within a noun phrase, adjectives take a classifier agreeing with the head noun, e.g. *him čo-po-ča* (fish big-CL:TINY.OBJECT-AUG) 'very big fish(es)' (Danielsen, 2007: 165; the classifier for tiny iobjects refers to the fish). A similar situation has been described for the closely related Ignaciano (Olza Zubiri et al., 2001: 179-88).

Alto Perene has a small closed class of underived adjectives, covering the meanings of size, age, quantification, and similarity, and an open class of derived adjectives. All adjectives take locative case markers and the plural number marker *-paye*, just like nouns. They are similar to verbs in that they can function as heads of intransitive and transitive predicates, and take markers of person aspect and reality status. They can take special degree suffixes (not found with other word classes), and show gender agreement with the head noun in a noun phrase (Mihas, 2015: 120).

The small adjective class in Paresi-Haliti covers size, age, value, and physical properties. Adjectives are similar to stative verbs in that they can occur with verbal prefixes and take some verbal inflections. But unlike verbs, they can modify nouns directly, e.g. *hati*

*kalore* (house big) 'big house', while verbs have to take the adjectivizer *-re* to be used as modifiers to nouns (Brandão, 2014: 130-1).

Tariana has a closed class of twenty-nine adjectives, with two subclasses: adjectives referring to dimension, age, and colour are noun-like, while adjectives of value and physical properties share similarities with stative verbs and with nouns; all adjectives can be used as modifiers in noun phrases, in contrast to verbs and nouns which have to be adjectivised to be used this way. Adjectives in some Kampa languages, e.g. Apurucayali, appear to be noun-like (Payne, 1981: 18; see also Wise, 1986: 578). In Guajiro, adjectives are considered a subclass of stative verbs (Alvarez, 1994).

## 2.2 Word-class changing derivations

Word-class changing derivations include numerous nominalizing markers, the most widespread across the family being the reflex of proto-Arawak *\*-ri* 'action nominalization', also used as a relativiser and adjectiviser on verbs. Causative prefixes and suffixes occur on nouns as derivational markers forming transitive verbs, e.g. Paresi-Haliti *itxo* 'hoe', *a-itxo-tya* (CAUS-hoe-THEMATIC) 'to weed', Tariana *sie* 'firewood', *-sie-ta* (firewood-CAUS) 'put on fire', *payuma* traditional soap', *-payuma-ta* 'apply traditional soap' (Brandão, 2014: 268; Aikhenvald, 2003: 274). The ubiquitous prefix *ka-* 'attributive' discussed in §1.2.1 derives possessive stative verbs from obligatorily possessed nouns. The marker of non-specified possessor on obligatorily possessed nouns (reflex of Proto-Arawak *-tʃi* or *\*-hVi*) forms deverbal action nominalizations in many languages of the family, e.g. Warekena of Xié *nu-naware* 'my village', non-possessed form *naware-ʃi* 'village (in general)'; *-yanata* 'write', deverbal nominalization *nu-yanata-ʃi* 'my writing' ((3.2.1)).<sup>8</sup>

## 2.3 Affixes shared by nouns and by verbs

Affixes which can occur on nouns, verbs, and also adjectives, typically include markers of degree (intensifier, augmentative, and approximative) and diminutive. In Baniwa of Içana, the augmentative *-pi* can occur on a noun, e.g. *tsiíno-pi* (dog-AUG) 'a big dog' and on a transitive verb, indicating the size of the object, as in (10) (Ramirez, 2001: 179).

- |      |  |                    |                        |
|------|--|--------------------|------------------------|
| (10) | no-kapa-pi<br>1sg-ver-AUG<br>'I saw a big dog' | tsiíno<br>cachorro | <i>Baniwa of Içana</i> |
|------|--|--------------------|------------------------|

The intensifier *-pero/-piro* in Alto Perené can occur on nouns and on verbs (Mihás, 2015: 385). In Tariana, the approximative *-iha* can be used with nouns, verbs and adjectives. The diminutive *-tiki/-tuki* can occur with nouns and with verbs, with a slight difference in form and in meaning. When used with nouns, the diminutive agrees in number with the noun head. The form *-tiki/-tuki* has the meaning of singular, with *-tupe* used for plural, as in *nu-itu-tuki* (1sg-daughter-DIM.SG) 'my little (or beloved) daughter', *nu-yanape-tupe* (1sg-children-DIM.PL) 'my little (or beloved) children'. There is no number agreement for *-tiki/-tuki* on verbs; it describes an action performed a little bit, or almost performed but not quite, as in (11) and (12) (see also Aikhenvald, 2003: 192-4, 179; 518).

- |      |  |                |
|------|--|----------------|
| (11) | nu-hwa-tuki-mahka<br>1sg-fall-DIM-REC.P.NONVIS<br>'I almost fell down' | <i>Tariana</i> |
| (12) | pi-dia-tuki<br>2sg-return-DIM  | <i>Tariana</i> |



other words and morphemes (as in Baniwa Hohôdene *ri-kwa-da-ka* (3sgnf-be.enough-CL:ROUND-DECLARATIVE) 'four' (literally, 'it is enough')). 'Five' typically includes the word for 'hand', and 'ten' tends to be expressed as 'two hands'. The root meaning 'one' is also used in the meaning of 'other'.<sup>10</sup>

Palikur is unusual for the family, and for Amazonia in general, in having a large system of native number words, partly seven-based, and partly ten-based, going up to 199 (French Creole words are used for counting from 200 up) (Green and Green, 1972, 2016).

### 3 Grammatical categories of nouns and noun structure

Grammatical categories of nouns across Arawak languages are possession, number, gender, classifiers, case, and nominal tense.

#### 3.1 Possession

Nouns divide into obligatorily — or inalienably — possessed and optionally — or alienably — possessed. Many languages have a closed class of nouns that cannot occur in possessive constructions, including natural phenomena in Waujá, Nanti, and Baniwa of Içana, and proper names and place names in Paresi-Haliti (Richards, 1973; Michael, 2013; Aikhenvald, 2013: 16; Brandão, 2014: 176). Obligatorily possessed nouns are body parts, kinship terms and a few important possessions, e.g. 'name' and 'house'. They always occur with a possessive prefix (the forms are in Table 1). If the possessor is not specified, these nouns occur in an unpossessed form marked with a reflex of proto-Arawak *-tʃi* or *\*-hVi*, e.g. Alto Perené *no-ina* (1sg-wife) 'my wife', *ina-tsi* (wife-NPOSS) 'wife (in general)', Paresi-Haliti *no-tiho* (1sg-face) 'my face', *tiho-ti* (face-NPOSS) 'face (in general)', Piro *no-myo* (1sg-hand) 'my hand', *myo-ti* (hand-NPOSS) 'hand (in general)', Bare *nu-nene* (1sg-tongue) 'my tongue', *nene-hei* 'tongue in general' (Mihás, 2015: 337, Rowan and Burgess, 1979: 73, Hanson, 2010: 107, Aikhenvald, 1995a: 13).

In a number of North Arawak languages, the non-possessed suffix on nouns is accompanied with the indefinite prefix *i-*, e.g. Baniwa of Içana *nu-whida* (1sg-head), *i-hwida-tti* (INDEF-head-NPOSS) 'head in general' (Aikhenvald, 1995b, Taylor, 1990). The suffix is also used to derive action nominalizations from verbs. The two functions share a common functional core: the non-possessed suffix makes a bound noun into a free form. When used with a verb, it allows a bound verbal root — typically used with suffixes — to occur without them. In a number of languages, including Tariana and Nanti, the non-possessed suffix (called 'alienator' by Michael, 2013: 154 and 'absolute' by Payne, 1991: 379) is no longer productive and survives in just a few nouns (e.g. Nanti *panko-tsi*, Tariana *pani-si* (house-NPOSS) 'a house'). The non-possessed suffix *-ti* is restricted to just a few forms in Mojo Ignaciano (Olza Zubiri et al., 2001: 40-2), e.g. *nu-yava* (1sg-axe) 'my axe', *yava-ti* 'axe in general'; the innovated suffix *-re* is used more productively in a similar function, e.g. *nu-situ* 'my sand flea', *situ-re* 'sand flea in general'). Baure innovated the unpossessed marker, *-ko/-ok*, used with a restricted number of inalienably possessed nouns, e.g. *ni-toer* 'my field', *toero-ko* 'a field' (Danielsen, 2007: 121-2).

Optionally, or alienably, possessed nouns take a possessive prefix and an additional suffix — a reflex of *\*-ne*, *\*-te*, *\*-re*, *\*-i/e*, or *\*-na*. The choice of the suffix may be partly motivated by the semantics of a noun. In Warekena of Xié ans Bare, the majority of artefacts take the possessive suffix *-re*, e.g. Wakekena of Xié *neyupa* 'paddle', *nu-neyupa-re* 'my paddle'; *bitsa* 'hammock', *nu-bitsa-re* 'my hammock', Bare *kaku* 'fish net', *nu-kaku-re* 'my fish net'. The suffix *-nel/-ne* in Bare is used with the majority of nouns of various semantic groups, including loans, e.g. *matseta* 'knife', *nu-matsteta-ne* 'my knife'.

In Piro, nouns with human or animate referents, or having referents of cultural significance, take the suffix *-ne*, e.g. *fima* 'fish', *n-fim-ne* 'my fish', *tfali* 'fishing net', *n-tfal-ne* 'my fishing net'. Nouns which take *-re* are instrument nominalizations containing the nominaliser *-pi*, e.g. *fifyapi* 'clothes brush', *n-fifyap-re* 'my clothes brush'. The suffix *-te* applies to a large heterogeneous class of nouns, including loanwords, e.g. *rimiryo* 'cure' (from Spanish *remedio*), *n-rimiryo-te* 'my cure' (Hanson, 2010: 107-14). In Paresi-Haliti, nouns with inanimate reference take the possessive suffix *-ne*, as in *hito* 'bow', *hito-ne* 'his bow'. Most animate nouns take the suffix *-za* (the reflex of Proto-Arawak *\*-te*), e.g. *halawa* 'a type of bird', *halawa-za* '(his) bird'; the possessed suffix *-la/-ra* is used for all other nouns.

Paresi-Haliti developed an additional innovative feature. Possessed nouns take a special suffix marking first person singular possessor, e.g. *malamala-ti* (lung-NPOSS), *no-malamala-i* (1sg-lung-1sg.POSS) 'my lung', *no-halawa-xi* (1sg-type.of.bird-1sg.POSS) 'my bird' (Brandão, 2014: 168-70, on the distribution of suffixes; and §4.1).

The choice of possessive suffixes *-ne* and *-te* in Warekena of Xié correlates with the distance from the speaker, e.g. *tsinu* 'dog', *nu-tsinu-ne* (1sg-dog-CLOSE.POSS) 'my dog (close to me)', *nu-tsinu-te* (1sg-dog-DISTAL.POSS) 'my dog (far from me)'.

In some Kampa languages, possessive suffixes are chosen partly based on the meaning of the noun, and partly by its syllabic length. In Alto Perené, the possessive suffix *-ri* is often used with artefacts and plant names, e.g. *n-itsa-ri* 'my robe'. In other cases, the choice of possessive suffix is conditioned phonologically: the suffix *-ni* is used with mono- and disyllabic roots, e.g. *no-mapi-ni* 'my stone', and *-te* with trisyllabic roots, e.g. *no-chacopi-te* 'my arrow' (Mihas, 2015: 335-6; see also Michael, 2008: 300, on Nanti). In Tariana, possessive suffixes are no longer productive; they survive in a few nouns referring to traditional artefacts, flora, and fauna (Aikhenvald, 2003: 131-2).

In a few inalienably possessed nouns in Paresi-Haliti, the non-possessed suffix can occur together with a possessive suffix, expressing an additional possessor. The non-possessed suffix transforms the erstwhile inalienably possessed noun into an alienably possessed one. For instance, the form *ete-ti* 'flesh-NONPOSS' means 'meat, flesh', and *ete-ti-ra* (flesh-NONPOSS-POSS) means 'meat that belongs to me; meat that I am eating', *tsiri-ti-ra* (head-UNPOSS-POSS) means 'head (of an animal I killed) that belongs to me'. This 'double possession' marking is an innovation (Brandão, 2014: 176; see also Rowan and Burgess, 1979: 73; and §1.2.1 on prefix stacking in Amuesha to express double possession).

### 3.2 Number

Number marking on nouns and noun phrases is optional, and correlates with the nominal hierarchy: explicit number marking is more likely to be used with human and animate referents than with inanimates (as predicted by Smith-Stark, 1974). Plural number is fairly uniform across the family; typical markers are reflexes of *\*-nal/-ni/-nai* 'animate/human plural' and *-pe* 'inanimate/animate non-human plural'. In Bare, *-ñu* is used as an optional number marker on nouns with human or animate referents, and *-be* marks plural of inanimates (Aikhenvald, 1995). In Baniwa of Icana and Kurripako, the suffix *-nai* is optionally used to mark plural on nouns with human and animate reference, and *-pe* is used with other nouns; an additional suffix *-nai-pe* is used with nouns denoting blood kin (Bezerra, 2005, Ramirez, 2001). Some languages, e.g. Piro, have just the marker *-ne* (Hanson, 2010: 116). Northern Kampa languages have the reflexes of *\*-pe* in Ashaninka *-pe*, Pajonal *-paeni* and Pichis *paini* (Mihas, 2017: 793).

A number of Ashéninka varieties including Alto Perené have the plural marker *-ite* used with nouns referring to humans and supernatural beings. Tariana has a complex system of number marking partially calqued from East Tucanoan languages. Resígaro is the only



Arawak language with dual marking on nouns and pronouns, developed as a consequence of language contact with the unrelated Bora (Aikhenvald, 2001).

In a number of Arawak languages of the Uapuí subgroup with multiple classifier systems (including Tariana, Baniwa of Içana and Resígaro), nouns with an inanimate referent have a collective meaning. A noun classifier in a derivational function has to be added for singulative reference, e.g. Tariana *de:ri* 'banana (collective)', *de:ri-pi* (banana-CL:LONG) 'a banana'. Only a noun with a classifier can be pluralized, e.g. *de:ri-pi-pe* (banana-CL:LONG-pl) 'bananas'. This feature is also found in neighbouring languages, especially East Tucanoan, Bora and Witotoan (see Aikhenvald, 2014, for further details).

### 3.3 Genders and classifiers

Most Arawak languages distinguish masculine and feminine genders in third person cross-referencing affixes and clitics, singular personal pronouns, demonstratives, nominalizations, and gender-sensitive kinship terms, e.g. Palikur *amepi-yo* 'thief (woman)', *amepi-ye* 'thief (woman)', Achagua *ka-muru-ka-i* (ATTR-manage-NOM-masc.sg) 'male hunter', *ka-muru-ka-u* (ATTR-manage-NOM-fem.sg) 'female hunter' (Meléndez, 1998: 60-1, 77), Tariana *nu-phe-ri* 'my elder brother', *nu-phe-ru* 'my elder sister'. Gender is also expressed in agreement markers on adjectives. Noun phrases with two agreeing modifiers, a demonstrative and an adjective, from Piapoco are at (14)-(15) (Klumpp, 1990: 114-15; 86)

- |      |   |              |                                       |                |
|------|---|--------------|---------------------------------------|----------------|
| (14) | yái<br>DEM.PROX.nonfem.sg<br>'this big buck deer' | néri<br>deer | manuí-iri-ca<br>big-nonfem.sg-ARTICLE | <i>Piapoco</i> |
| (15) | úái<br>DEM.PROX.fem.sg<br>'this big doe deer'     | néri<br>deer | manuí-íchúa-ca<br>big-fem.sg-ARTICLE  | <i>Piapoco</i> |

Gender markers are a relatively stable category across the family. The markers go back to Proto-Arawak third person singular suffixes and prefixes feminine (*r*)*u*, masculine (*r*)*i*. Third 'neuter' gender was innovated in Palikur, and in Ignaciano (where gender choice correlates with the sex of the speaker: Rose, 2013). Palikur distinguishes three genders in demonstratives, cross-referencing prefixes and suffixes, and independent pronouns, and two genders in adjectives, derivational suffixes, and a few verbal forms (Aikhenvald and Green, 1998).

Gender assignment tends to be relatively transparent: female referents belong to feminine gender, and all the rest to the masculine (or nonfeminine) gender, which is also the unmarked choice if the gender is unknown. In Kampa languages inanimates tend to be classified as belonging to the feminine gender. Gender choice can be determined by the role of the entity in myths: in a number of Kampa languages, *cashiri* 'moon' is masculine (in contrast to many other inanimates) because it is believed to have been a mythical man. In Palikur, heavenly bodies, thunder and lightning are masculine for the same reason (Payne, 1989: 130; Aikhenvald and Green, 1998). Gender choice in Lokono correlates with the attitude of the speaker (Pet, 1987: 26-7): for instance, females and despicable males are treated as feminine, and so are males who are not of the speaker's tribe.

A number of Arawak languages have lost gender distinctions in agreement and cross-referencing, in all likelihood under the influence of neighbouring languages without genders. The loss of genders in the Arawak languages of Xingu (Waurá, Mehinaku, Yawalapiti, †Kustenaú) is attributed to the influence of other languages of the region, from the Tupí and Carib families (Seki, 1999). The loss of gender in Amuesha may be due to Quechua influence

(Adelaar, 2006). Lack of gender in Bahuana may be due to the influence of *Língua Geral* (Ramirez, 1992). *Terêna* and *Kinikinao* may have lost genders as a consequence of their contact with speakers of *Tupí-Guaraní* and of *Kadiweu* (a *Guaicuruan* language). Many of the 'gender-less' languages employ classifiers as noun categorization devices, and preserve gender distinctions in some nominalizations and derivations.

More than half of the languages have complex multiple classifier systems (see Aikhenvald, 2017c, 2012: 295-303, for a typological perspective). Classifiers occur on number words, verbs, on adjectives, and on nouns themselves. They categorise the noun in terms of its shape, consistency, and animacy. The classifier *-mashi* 'flat, broad' in *Alto Perené* occurs on an adjective (16), a noun (17), and on a verb categorising the intransitive subject 'manioc plants' (18) (Mihas, 2017: 792).

- (16) antaro-**mashi**                      kishi                      *Alto Perené: classifier on an adjective*  
 big-CL:FLAT.BROAD    grassland  
 'The grassland is a vast territory'
- (17) antami-**mashi-ki**                      *Alto Perené: classifier on a noun*  
 forest-CL:FLAT.BROAD-LOC  
 '(through) the forested area'
- (18) o-naria-**mashi-t-aty-a**                      kaniri *Alto Perené: classifier on a verb*  
 3NM.S-lie-CL:FLAT.BROAD-EP-PROG-REAL    manioc  
 'The manioc plants are lying flat in the field [because of the wind]'

In *Baniwa* of *Içana-Kurripako*, classifiers used with adjectives distinguish non-feminine and feminine forms, creating a typologically unusual instance of an interaction of gender and classifier system, as shown in (19)-(20) (Ramirez, 2001; Taylor, 1991; Aikhenvald 2007b, Bezerra, 2005).

- (19) apa-ita                      nawiki                      maka-dari                      *Baniwa of Içana*  
 one-CL:HUMAN                      person                      big-CL.ADJ:HUMAN  
 'one big person'
- (20) apa-ma                      inaru                      matsia-dazu                      *Baniwa of Içana*  
 one-CL:FEM    woman                      beautiful-CL.ADJ.FEM  
 'one beautiful woman'

In most multiple classifier languages of the family, classifiers on nouns are used as a productive derivational device, e.g. *Tariana di-tape* (3sgnf-medicine), *di-tape-dapana* (3sgnf-medicine-CL:HABITAT), 'hospital' (lit. medicine-house), *di-tape-dapanite* (3sgnf-medicine-CL:HABITAT+CL:ANIMATE) 'a person of medicine house, a medical worker'.

Multiple classifier systems are a feature of South Arawak languages (including *Terêna*, *Baure*, and *Ignaciano*), *Xinguan* languages and *Paresi-Haliti*, *Kampa*, and *Amuesha*. North of the Amazon, classifiers in multiple contexts are found in *Baniwa* of *Içana-Kurripako*, *Tariana*, *Resígaro*, and *Wapishana*. *Palikur* has more than a dozen classifiers which have different semantics and form depending on whether they are used on number words, verbs, or adpositions (Aikhenvald and Green, 1998).

A number of North Arawak languages — including *Yucuna*, *Bahuana*, *Achagua*, *Warekena* of *Xié*, *Baniva* of *Guainia*, and possibly the extinct *Maipuré* — have only numeral classifiers. Many classifiers transparently derive from obligatorily possessed nouns, e.g.

Baniwa of Içana *-ya* 'classifier for foldable extended items'. *-ya* 'skin'. The number of classifiers vary from several dozen in Baniwa of Içana, Tariana, Kampa and South Arawak languages to about ten in Achagua and Yucuna. Classifiers are absent from Piapoco, Bare, Guarequena, Chamicuro, Lokono, Guajiro (or Wayyu-naiki), Añun, and Garifuna (see Aikhenvald, 1999: 83, 2012: 296-8; forthcoming). Bound class nouns in Apurina and in Piro may be considered an incipient system of noun categorization (Facundes, 2000; Hanson, 2010: 101).

The distribution of multiple classifier systems in Arawak languages overlaps with the two regions with complicated systems of classifiers — north-west Amazonia with links to north-eastern Peru, and southern Amazonia (Brazil and adjacent areas of Bolivia) (Aikhenvald, 2012: 301-2). Classifier systems in Tariana and Resígaro bear the impact of neighbouring East Tucanoan and Bora languages, respectively. Language contact has undoubtedly played a role in the emergence of classifier systems in the family.

### 3.4 Case

The majority of Arawak languages do not employ cases for marking core grammatical relations. The only exception is Tariana, from the multilingual Vaupés River Basin Linguistic area. Here, core cases *-naku/-nuku* 'topical non-subject' and *-nel/-nhe* 'focal subject' were developed under the influence of the neighbouring East Tucanoan languages (Aikhenvald, 2002).

Tariana has an instrumental-comitative case (*-(i)ne*) and one locative marker (*-se*) covering direction, location, and source; one locative case is a feature of Baure (*-ye*: Danielsen, 2007: 150-1), Kampa languages (e.g. *-ki* in Alto Perené: Mihas, 2015: 352-353; *-ku* in Nanti: Michael, 2008: 302). Piro adds *-yma* 'comitative-instrumental' and *-yehi* 'in the vicinity of' to the multifunctional locative case *-ya* (Matteson, 1965; Hanson, 2010: 149-152). In addition to the instrumental-locative *-ã* and *-takote* 'next to, in the vicinity of', Apurina has a causal, a comitative, and a temporal case on nouns (Facundes, 2000: 259-62).

A larger set of locative cases (locative 'in, at', allative 'to', and ablative 'from') is a feature of Amuesha (in addition to a dative-benefactive case), Bare, Piapoco, and Achagua (Duff-Tripp, 1997: 35-36). Baniwa of Içana-Kurripako distinguish two locatives — a general *-riku*, and *-naku* 'on the surface of', in addition to a perlative *-wa*.

In many North Arawak languages, two locative cases can combine in one word to express complex locative meanings, e.g. Baniwa *awakada-riku-hre* (jungle-LOC-TOWARDS) 'towards the inside of the jungle', Bare *damakaru-kua* (jungle-LOC+PERLATIVE) 'along the jungle', *nisa-ní-waku* (1sg+canoe-POSS-PERLATIVE+LOC) 'along on my canoe' (Aikhenvald, 2002, 107). Locative case markers often lexicalize and develop into derivational affixes on inherently locational nouns, e.g. Bare *damakaru-ku* (jungle-LOC) 'jungle, in the jungle', Tariana *awakada-riku* (jungle-DER/LOC) 'the inside of the jungle'.<sup>11</sup>

### 3.5 Nominal tense

Nominal tense is a recurrent feature for many languages of Amazonia (Aikhenvald, 2012: 158-163). Many Arawak languages have independent grammatical marking of tense on a noun phrase, or 'nominal tense' (see Nordlinger and Sadler, 2004). Tariana, Baniwa of Içana-Kurripako, and Piapoco distinguish nominal past and nominal future, e.g. Tariana *panisi-mikiri* (house-NOM.PAST-masc.sg) 'former or destroyed house', *panisi-pena* (house-NOM.FUT) 'future house', Baniwa of Içana *nu-ita-ni-mi* (1sg-canoe-POSS-NOM.PAST) 'my former canoe (burnt or abandoned)', *nu-ita-ni-wa* (1sg-canoe-POSS-NOM.FUT) 'my future canoe (not yet bought or built)' (Taylor, 2010: 208; see also Matteson, 1965: 95-97, for past and future markers on nouns in Piro). Present-tense reference of a noun is typically formally unmarked.

Ashéninka Perené distinguishes a number of values of nominal past (and no nominal

future). The markers are *-ni* 'ceased existence' (also used on verbs, with the meaning of distant past: §2.3), *=ranki* 'temporal precedence, change of entity's property', and *-nta* 'ceased human possessive relationship'. Nouns with non-past reference are formally unmarked. (21) illustrates different marking of tense and reality status on the verb and on the noun (Mihas, 2014: 9; 2017: 793-4).

- (21) No-n-kinkitsa-t-ako-t-e-ro                      inaa-ni                      *Ashéninka Perené*  
 1A-IRR-tell-EP-APPL-EP-IRR-3.NON.MASC.O    mother-CEASED.EXIST  
 'I will tell about my deceased mother'

Nominal past in Mawayana, a moribund Arawak language spoken in the frontier corner of Brazil, Guyana, and Suriname, developed as a result of its contact with the dominant Waiwai and Trio, from the Carib family (Carlin, 2006: 324-5). Nominal tense is absent from Wapishana, an Arawak language closely related to Mawayana, and from Palikur, Guajiro, and Añun.

### 3.6 The structure of a noun and further nominal categories

In many Arawak languages, similarly to other highly synthetic languages of Amazonia, nouns can be morphologically complex (see Aikhenvald, 2017a). A nominal word in Alto Perené and many other Kampa languages may include up to nine suffix positions, including markers of intensification, diminutive, number, and tense. A noun in Piro may include seven suffixal slots, with meanings covering nominal tense, aspect, frustrative modality, sequence, repetition, 'interest', and 'calling attention to the head word' (Matteson, 1965: 95-98). A noun in Tariana can have up to thirteen post-root slots (suffixes and enclitics), covering genders, classifiers, number, evaluation (pejorative), approximative, diminutive and augmentative, tense, extralocality, and case (see Aikhenvald, 2017b: 730-1).

Further nominal categories include intensifiers, diminutives, and markers of attitude and affection, e.g. Tariana *=yana* 'pejorative', Mawayana *=kwe* 'affective', and similarity, e.g. Mawayana *-ni* 'similative' (Carlin, 2006: 325).

## 4 Verb categories and verb structure

The verb is the most complicated part of the grammar of every Arawak language, and the only obligatory constituent in a clause. The internal organization of a verbal word follows the principle of templatic organization, with a number of slots for suffixes and enclitics which are mutually exclusive; the number of structural slots ranges from five in Wapishana to twenty in Tariana (Tracy, 1974; Aikhenvald, 2016b; see also Payne, 2002: 486 on Asheninka; Mihas, 2017: 796-7 on Kampa languages). Two or three prefix positions on a verb are a feature of South Arawak languages, Kampa, Paresi-Haliti and Xinguan Arawak (§1.1.1).

In the history of individual languages, a thematic suffix (often segmentally the same as valency changing suffixes *-ta* and *-ka*) may have to be added to the verbal root to form the verbal base. The choice of a thematic syllable tends to correlate with the subclass of verbs; for instance, in Paresi-Haliti, spoken south of the Amazon in the state of Mato Grosso in Brazil, the formative *-tya/-tsa* occurs with transitive and ditransitive verbs, and *-ka* occurs with active intransitive and also with transitive verbs (Brandão, 2014: 227-9; Rowan and Burgess, 1979). In Guajiro, transitive verbs occur with the thematic markers *-jV*, *-tV*, *-lV*, *nV*, and *-kV*, while active intransitive verbs take only *-tV* (Alvarez, 1994). In Bare, transitive verbs can take thematic suffixes *-sa* and *-ka*, while active intransitive take *-ka*, *-da* and *-ña*. Thematic markers do not occur on stative intransitive verbs.

### 4.1 Verb classes and the marking of grammatical relations

Verbs divide into active and stative. Active verbs can be transitive, with two core arguments (A, the transitive subject, and O, the object), or ditransitive, or active intransitive, with one core argument,  $S_a$ . Stative verbs, with one core argument,  $S_o$ , are typically intransitive (some languages have transitive stative verbs, e.g. Tariana *harame* 'be afraid (of someone)'; see also Danielsen and Granadillo, 2008: 403). As shown in Table 1, A is typically marked with the same prefixes as  $S_a$ , and O is expressed with the same suffixes or enclitics as  $S_o$ . This is a type of split-S system, illustrated in (24)-(25), using the example of Warekena of Xié (further discussion is in Aikhenvald, 2012: 212-214). In (24), the prefix marks the A, and the enclitic marks the O.

- (22) pi-muta-mia=yu *Warekena of Xié*  
 2sgA-call-PERF=3sgfO  
 'You called her'

The subject of intransitive active verbs, including verbs of motion is also marked with prefixes. An example is at (23).

- (23) nu-fumia-wa<sub>Sa</sub>      yu-tapapa-mia *Warekena of Xié*  
 1sg-spouse-FEM<sub>Sa</sub>      3sgfSa-walk.around-PERF  
 'My wife walked around'

An example of an  $S_o$  verb with its only core argument marked with a pronominal clitic is in (24).

- (24) anuana-mia=yu<sub>So</sub>      nu-fumia-wa<sub>So</sub> *Warekena of Xié*  
 be.sick-PERF=3sgf<sub>So</sub>      1sg-spouse-FEM<sub>So</sub>  
 'My wife was sick'

The semantic content of active and stative verb classes may be different even for closely related languages. Baniwa of Içana-Kurripako display a fluid S-marking with just a few verbs: the same intransitive verb can take an  $S_a$  prefix if it refers to a controlled activity of 'crying', as in (25), and an  $S_o$  enclitic if it refers to an uncontrolled event, 'raining', as in (26) (Aikhenvald, 1995a: 165-6, see also Danielsen and Granadillo, 2008: 410, for some similar examples in Baure).

- (25) li-idza-ka *Baniwa of Içana-Kurripako*  
 3sgnf-cry-PROG  
 'He is crying'
- (26) idza-ka=ni *Baniwa of Içana-Kurripako*  
 rain-PROG=3sgnf  
 'It is raining'

The Split-S system has been lost from Xinguan languages and Paresi-Haliti, and from a number of languages north of the Amazon (including Guajiro, Añun, and Wapishana). A small number of stative verbs in Paresi-Haliti take the  $S_o$  first person singular subject marker -i, with the same form as the one employed on nouns marking the first person singular possessor (§3.1) — another innovation in this language. In Palikur, the split S patterns survives only on the interrogative 'who' in the predicate slot (Aikhenvald and Green, 1998: 469).

Piro displays a typologically unusual pattern of split-S marking: pronominal prefixes mark the A/S of verbs, and suffixes mark the O of transitive verbs and the subject of non-verbal predicates (Hanson, 2010: 208-16). Similarly, in Baure nouns and adjectives in the predicate slot take cross-referencing suffixes (also used for marking objects of transitive verbs), while most stative verbs take prefixes to mark the subject, just like other intransitive and transitive verbs (Danielsen, 2007; Danielsen and Granadillo, 2008).

Kampa languages have an unusual system of fluid S-marking. The subject of a transitive verb is always marked with a cross-referencing prefix, and the object with a suffix. The subject of any intransitive verb can be marked with either a prefix, or a suffix, depending on the topicality of the S argument. If the subject is a well-established topic, it is marked with a suffix, and if it is a new topic, it will tend to be marked with a prefix (Mihás, 2017: 803-804; Payne, 2002: 495). Kampa languages allow up to two object suffixes, one of which tends to be the direct object, and the other the beneficiary or an additional argument added through an applicative derivation (example (32) in §4.2).

Guajiro and Añun have two verbal conjugations: suffixal, or 'analytic', and prefixal, or 'synthetic'. Cross-referencing suffixes in an analytic verbal conjugation refer to the subject (A/S) of a verb which can be considered of low affect. Cross-referencing prefixes are used in the prefixal conjugation with transitive verbs with a referential definite object to refer to the A of a transitive verb of high affect; there, cross-referencing suffixes refer to the O. In subordinate clauses, prefixes mark A/S<sub>a</sub> and suffixes mark O/S<sub>o</sub>, reflecting the archaic Arawak pattern (Alvarez, 1994; Olza Zubiri and Jusayú, 2012; Patte, 1989)

#### 4.2 Verbal categories

Typical verbal categories include tense, aspect, evidentiality, modalities (desiderative, frustrative, meaning 'do in vain'), and valency changing derivations — passives, reflexives, reciprocals, causatives, and applicatives (§4.2). The opposition of realis-irrealis is obligatory in Kampa and in Terêna (and its close relative Kinikinao). Realis covers events which took place in the past; irrealis covers future events, commands, unrealised events, and negated statements; negative commands are cast in realis (Aikhenvald, 2012: 190-1; Mihás, 2017: 796; Butler and Grimes, 1964; see also Danielsen, 2007: 267-8 on the irrealis in Baure).

All Arawak languages have complex systems of aspects. The inventory of aspects in most Kampa languages include perfective, terminative, completive, customary, progressive, durative, inchoative, and iterative, in addition to special markers of perfective and imperfective (Mihás, 2017: 796). Baure has a 'departitive' suffix *-wana* which refers to the action done before or while departing. Modal meanings include intentional, desiderative, optative, purposive, dubitative, apprehensive, and frustrative. Negation is often marked on the verb with a suffix, and/or the common Arawak prefix *ma-* (see the papers in Michael and Granadillo, 2014).

Evidentiality (grammaticalised marking of information source) is a key feature of Tariana. There are five specifications — visual, nonvisual, inferred, assumed, and reported fused with three tenses (present, recent past, and remote past). This system was developed as the result of contact with East Tucanoan languages (Aikhenvald, 2002, 2016b, 2017b). Reported evidentials are attested in Baniwa of Içana-Kurripako, Piapoco, Achagua and Resígaro north of the Amazon, and in a variety of languages south of the Amazon, including Waujá, Terêna, Piro, Ignaciano, and Amuesha. Bahuana (Ramirez, 1992: 64-5) distinguishes a reported and an inferred evidential, and Nanti has quotative, reported and inferred evidential (each expressed with a clause-level clitic: Michael, 2008: 323-325).

Verbal suffixes may express emotional states and attitudes, e.g. Perené *-matsi* 'compassion' and *-amampi* 'premonition'. Kampa, Amuesha, and South Arawak languages have verbal affixes indicating the direction of the action, e.g. Terêna *-op/-po/p* 'motion back

to a point of origin', Baure *-pik* 'directional', Alto Perené *-ai* 'distal' (Ek Dahl and Grimes, 1964; Danielsen, 2007: 265; Mihas, 2015: 397). Wapishana has verbal suffixes marking intermittent action, aimless action, and 'reversative (undoing something)' (Tracy, 1974). Verbal suffixes and enclitics may have very specific meanings, e.g. Kampa *-aman* '(do) early in the morning', *-nink* '(do) late in the afternoon' (see Wise, 1986: 590-591), Tariana *=bisi* '(throw) into the fire', *=thepi* '(throw) into the water', *=ñu* 'step on something and feel sharp pain'. Verbal affixes with specific meanings are a typical feature of polysynthetic languages of Amazonia and other regions. A variety of valency-changing derivations and numerous patterns of noun incorporation are a feature of many languages south of the Amazon.

#### 4.2 Valency-changing derivations

Arawak languages have complex sets of valency-changing derivations. There is typically just one passive; antipassive is a feature of some Kampa languages, including Asheninka and Alto Perené (Payne, 2002: 492; Mihas, 2015: 309). Reflexives and reciprocals are expressed with valency-reducing derivations on the verb (a summary is in Aikhenvald, 1999: 90-93). A striking feature of Kampa languages is having several causatives, and up to six applicative derivations, including comitative, benefactive/malefactive, beneficiary-recipient, presential, separative ('away from'), and instrumental.

Two causative markers in Asheninka differ in the type of causation they express, and the participation of the causee (in agreement with the parameters in Dixon, 2012: 268-280). The causative prefix *-oi-* in (27) implies direct physical contact and direct action of the causer, with the causee having no control (Payne, 2002: 497).

- (27) n-oi-pithok-ak-e-ri no-tomi Asheninka  
 1sg-CAUS-turn-PERF-MODE-3masc  
 'I turned my son over (e.g. speaking of a baby or an immobile son)'

The causative suffix *-aka(g)* illustrated in examples (2), (3), and (28) expresses indirect coercion or manipulation, with the causee having partial control.

- (28) no-pithok-aka-ak-e-ri no-tomi Asheninka  
 1sg-turn-CAUS-PERF-MODE-3masc  
 'I made my son turn over'

Applicative derivations in Kampa languages can apply to intransitive verbs making them transitive. Or they can apply to transitive verbs, promoting a peripheral argument to direct object and rearranging the valency. The inventory of applicatives across Kampa languages include the following (Mihas, 2017: 798). The semantic role of oblique argument promoted to core is indicated in brackets:

- generalized *-ako~-ko* (location, source, stimulus, topic 'about'),
- sociative-causative *-aka(g)~-ka* (co-actor/companion; causee),
- beneficiary/maleficiary *-vent~-ben* (beneficiary, maleficiary, 'on behalf of, because of, instead of', topic 'about'),
- beneficiary-recipient *-(a)nont* (beneficiary-recipient, maleficiary-recipient, self-beneficiary-recipient),
- intent *-ashi~-asi* (goal),
- instrument/reason *-ant*,
- separative *-(a)pitha~pitsa~pi* (source '[away] from'),
- presential *-imo~-mo* (companion, allative 'to', apudessive 'in front of').

Nomatsiguenga also has an allative applicative *-te*. Ashaninka and Kakinte have a further applicative *-iment* 'with reference to'. An example of the instrumental applicative, from Nanti, is at (29) (Michael, 2008: 286-288; N is an unspecified nasal consonant).

- (29) *i=oseroNk-aNt-ak-a=ro* *Nanti*  
 3mascS=slice-APPL.INSTR-PERF-REAL.CLASS.A=3nonmascO  
 'He carved with it (a knife)

Presential applicative makes someone in whose presence the action is performed into an object. In Nanti, this applies only to intransitive verbs.

- (30) *o=pok-imo-ak-i=na* *Nanti*  
 3nonmasc=come-APPL.PRES-PERF-REAL.I=IO  
 'She arrived where I was' (lit. she arrived in my presence)

Separative applicative makes a point of 'separation', or an entity taken away, into an object. This applies to verbs of any transitivity. In (31), from Alto Perené (Mihas, 2015: 301), the separative applicative marked with *-apitsa-* is used with a transitive verb.

- (31) *eero* *p-a-apitsa* *Alto Perené*  
 NEG.IRR 2A-take-APPL.SEP-EP-REAL-1sgPOSSR-3nonmascO  
*iroñaaka* *no-vanki-ri*  
 now 1sg.POSS-plant-POSS  
 'He won't take away from me my crops'

Both the possessor and the entity taken away from the possessor ('crops') are cross-referenced on the verb. Complex applicative systems are a feature of Amuesha and Piro (Wise, 2002); they are largely absent from North Arawak languages which tend to have just a causative marker.

#### 4.4 Noun incorporation

Highly synthetic languages, such as Kampa, Amuesha, Piro, Apurina, Baure, Ignaciano, and Palikur, have productive noun incorporation. Similarly to many other Amazonian languages, in most Arawak languages only obligatorily possessed nouns tend to be incorporated. Palikur has a closed set of seven body parts which can be incorporated into the verb if they have the function of the object of a transitive verb or the subject of a stative verb. In (32), the object 'eye' is incorporated into the verb (Aikhenvald and Green, 1998: 452).

- (32) *kuri ig hakis-ota-ne han akiw* *Palikur*  
 now 3masc rub-eye-CONTnf thus again  
 'He continued rubbing his eyes again' (lit. eye-rubbing)

If the possessor of the body part is not coreferential with the subject, the possessor becomes the direct object. This is known as 'possessor raising', or 'possessor ascension'. In (33), also from Palikur, the possessor is cross-referenced on the verb with the direct object suffix *-gi* 'him'.

- (33) *ig-kis hapis patuk-ot-bet-h-e-gi* *Palikur*



3masc-PL shoot burst-eye-MULTIPLE-INT-COMPL-3masc  
 'They shot his eyes out' (lit. they eye-shot-him)

Intransitive verbs which undergo noun incorporation typically express qualities, as in (34).

- (34) eg barew-kug *Palikur*  
 3fem clean-foot  
 'She is clean-footed' (i.e., her feet are clean)

Similar examples from Nanti are in Michael (2008: 161-164). In some Kampa languages, such as Alto Perené, nouns of most classes, including those with animate reference, can be incorporated (Mihas, 2017: 800; 2015: 309-315). Incorporated body parts can become lexicalised; then the meaning of the whole cannot be deduced from the meaning of the parts, as in Palikur *nah barew-wok* (I clean-hand) 'I am poor, destitute' (lit. I am clean-handed) (Aikhenvald and Green, 1998: 455).

Classificatory noun incorporation (in the sense of Mithun, 1984) is a feature of Apurina and also of Piro: a noun with generic reference, in the function of object (O) or intransitive subject (S), can be incorporated into a verb delimiting the semantic field of the O or S noun (Facundes, 2000: 338; see also Aikhenvald, 2012: 197). In (35), from Apurinã, the classificatory noun *-pe* 'pulp of' is incorporated into the verb, limiting the reference of the noun *komuru* 'manioc' to 'manioc pulp'. The object itself, *komuru* 'manioc', can be overtly mentioned, but is optional (Facundes, 2000: 338).

- (35) ata (komuru-pe) usonãka-pe-ta-ka *Apurinã*  
 1pl (manioc-pulp.of) dry-PULP.OF-VERBALISER-CAUS  
 'We put the manioc pulp to dry'

A classificatory incorporand can be used without an accompanying noun. In (36), the classificatory noun *xiti* 'earth of' is incorporated into the verb, to anaphorically refer to the noun *kikio* 'field' previously mentioned in the text:

- (36) ata yotika-xiti-ta *Apurinã*  
 1pl burn-EARTH.OF-VERBALISER txa-ru AUX-3masc.sgOBJ  
 '(First, we prepare the field (*kiko*) by cutting down the trees, then) we set it (the field) on fire'

Over time, a classificatory noun may grammaticalize into a verbal classifier, losing its nominal features. Many verbal classifiers in Palikur and in Baure come from grammaticalized body parts; synchronically, they are distinct (a number of criteria for distinguishing them are set out in Aikhenvald and Green, 1998: 451-455). Nouns referring to location can be incorporated in Baure into transitive and into intransitive verbs, e.g. *-'ači-* 'other place' (followed by the non-possessed suffix *-ko*) in (37) (Danielsen, 2007: 209).

- (37) ver pi-tora-'ači-ko-pa *Baure*  
 PERF 2sg-find-other.place-NPOSS-INTENT pi-avinon 2sg-husband  
 'You will find your husband in another place'

Incorporating a postposition is a means of increasing valency of an intransitive verbs, or rearranging valency of a transitive verb. Five of seventeen postpositions in Paresi-Haliti can



number-marking on nouns, fitting in with the Tucanoan 'mould'. The system of multiple classifiers in Tariana bears the impact of Tucanoan languages. A number of structural features of classifier system which Tariana shares with closely related Baniwa of Içana-Kurripako (spoken outside the Vaupés River Basin linguistic area) are also found in East Tucanoan, Bora, Witotoan, and Guahiboan languages, suggesting wider patterns of areal diffusion (Aikhenvald, 2002; 2003a, b; 2012, 79-82; 2016b).

In many instances, the exact source of contact-induced features is hard to ascertain, due to historical merger of various, perhaps unrelated groups and the presence of substrata from extinct languages (typical for the Amazon Basin: Aikhenvald, 2012: 27). Amuesha bears traces of contact with Quechua and also other, now hard-to-identify languages. Palikur has numerous features unusual for the family; only some of them can be accounted for by Carib influence. The modern Palikur came about as a result of merger of several groups who originally spoke different languages (Diana Green, p.c.; Green and Green, 2013: 214-216). Yucuna, another unusual Arawak language, was also, originally, a conglomerate of several groups many of which are said to have spoken completely different languages in the past (Fontaine, 2008: 48-50, 83-84). Some of the features of Yucuna can be accounted for by East Tucanoan influence; others may well be due to the impact of other, unidentified, languages. The interaction of genetic inheritance, language contact patterns — only some of which may still be recovered —, and independent innovations is what makes Arawak languages dauntingly diverse.

#### Abbreviations:

A transitive subject	INT intensive
APPL applicative	INTENT intentional
APPL.INSTR instrumental applicative	IRR irrealis
APPL.PRES presential applicative	LOC locative
APPL.SEP separative applicative	M masculine
ATTR attributive	masc nonmasculine
AUG augmentative	NEG negative
AUX auxiliary	NEG.IRR negative irrealis
CAUS causative	NM nonmasculine
CEASED.EXIST ceased existence	NMLZ nominalizer
CL CLASSIFIER	NOM.FUT nominal future
CL:FEM classifier for feminines	NOM.PAST nominal past
CL.ADJ adjectival classifier	NONFEM nonfeminine
CL.ADJ.FEM adjectival classifier for feminines	NONMASC nonmasculine
COMPL completive	NPOSS non-possessed
CONTnf continuative non-feminine	O object
DEM.PROX proximal demonstrative	OBJ object
DER derivational	PAUS pausal
DIM diminutive	PERF perfect
DIM.SG diminutive singular	PFV perfective
EP EPENTHETIC	pl plural
fem feminine	POSS possessive
FEM feminine	POSSR possessor
FUT future	PROG progressive
HAB habitual	REAL realis
IMPFV imperfective	REAL.I realis class 1
INDEF indefinite person	REC.P.NONVIS recent past non-visual
	RED reduplication

REL.FUT relative future	sg singular
S intransitive subject	sgnf singular nonfeminine
S <sub>a</sub> subject of an active intransitive verb	THEM thematic
S <sub>o</sub> subject of a stative intransitive verb	

### Further readings

Aikhenvald, A. Y. 2016a. Oxford Bibliography Online: Arawak languages (refereed updateable resource with summaries and evaluation for each entry). General Editor: Mark Aronoff. New York: Oxford University Press (online resource).

### References

- Adelaar, W. F. H. 2006. The Quechua impact on Amuesha, an Arawak language of the Peruvian Amazon. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.), *Grammars in contact: a cross-linguistic typology* (pp. 290–312). Oxford: Oxford University Press.
- Admiraal, Femmy and Swintha Danielsen. 2014. Productive compounding in Baure. In S. Danielsen, K. Hanss, and F. Zúñiga (Eds.), *Word formation in South American languages* (pp. 79-112). Amsterdam: John Benjamins.
- Aikhenvald, A. Y. 1995a. *Bare*. Languages of the World/Materials 100. Munich: Lincom Europa.
- Aikhenvald, A. Y. 1995b. Person marking and discourse in North Arawak languages. *Studia Linguistica*, 49, 153-195.
- Aikhenvald, A. Y. 1998. Warekena. In D. C. Derbyshire and G. K. Pullum (Eds.), *Handbook of Amazonian Languages*, Volume 4 (pp. 215-439). Berlin: Mouton de Gruyter.
- Aikhenvald, A. Y. 1999. The Arawak language family. In R. M. W. Dixon and A. Y. Aikhenvald (Eds.), *The languages of the Amazon* (pp. 65-105). Cambridge: Cambridge University Press.
- Aikhenvald, A. Y. 2001. Areal diffusion, genetic inheritance and problems of subgrouping: a North Arawak case study. In A. Y. Aikhenvald and R. M. W. Dixon (Eds.), *Areal diffusion and genetic inheritance: case studies in comparative linguistics* (pp. 167-94). Oxford: Oxford University Press.
- Aikhenvald, A. Y. 2002. *Language contact in Amazonia*. Oxford: Oxford University Press.
- Aikhenvald, A. Y. 2003. *A grammar of Tariana, from north-west Amazonia*. Cambridge: Cambridge University Press.
- Aikhenvald, A. Y. 2003b. Mechanisms of change in areal diffusion: new morphology and language contact, *Journal of Linguistics*, 39, 1-29. - iff neded
- Aikhenvald, A. Y. 2004. The adjective class in Tariana. In R. M. W. Dixon and A. Y. Aikhenvald (Eds.), *Adjectives: a cross-linguistic typology* (pp. 97-124). Oxford: Oxford University Press.
- Aikhenvald, A. Y. 2007a. Classifiers in multiple environments: Baniwa of Içana/Kurripako: a North Arawak perspective. *International Journal of American Linguistics*, 27, 475-500.
- Aikhenvald, A. Y. 2007b. Typological dimensions in word formation. In T. Shopen (Ed.), *Language Typology and Syntactic Description*, Volume 3 (pp. 1-65). Cambridge: Cambridge University Press.
- Aikhenvald, A. Y. 2008b. Versatile cases. *Journal of Linguistics*, 44, 565-603.
- Aikhenvald, A. Y. 2012. *The languages of the Amazon*. Oxford: Oxford University Press.
- Aikhenvald, A. Y. 2013. Amazonia: linguistic history. In Peter Bellwood (Ed.), *Encyclopaedia of migrations*, volume 1 (pp. 384-91). Routledge-Wiley.
- Aikhenvald, A. Y. 2014. Number and noun categorization: a view from north-west Amazonia. In Gerrit J. Dimmendaal and Anne Storch (Eds.), *Number – constructions and*

- semantics. Case studies from Africa, Amazonia, India and Oceania* (pp. 33-55). Amsterdam: John Benjamins.
- Aikhenvald, A. Y. 2016b. Language contact and word structure: a case study from north-west Amazonia. In A. L. Berez-Kroeker, D. M. Hintz and C. Jany (Eds.). *Language Contact and Change in the Americas. Studies in honor of Marianne Mithun* (pp. 297-313). Amsterdam: John Benjamins.
- Aikhenvald, A. Y. 2017a. Polysynthetic structures of Lowland Amazonia. In M. Fortescue, M. Mithun and N. Evans (Eds.). *The Oxford Handbook of Polysynthesis* (pp. 284-311). Oxford: Oxford University Press.
- Aikhenvald, A. Y. 2017b. Tariana, an Arawak language from north-west Amazonia. In M. Fortescue, M. Mithun and N. Evans (Eds.). *The Oxford Handbook of Polysynthesis* (pp. 713-734). Oxford: Oxford University Press.
- Aikhenvald, A. Y. 2017c. A typology of noun categorization devices. In A. Y. Aikhenvald and R. M. W. Dixon (Eds.). *The Cambridge Handbook of Linguistic Typology* (pp. 361-404). Cambridge: Cambridge University Press.
- Aikhenvald, A. Y. 2018a. Disentangling a versatile prefix: the nature and development of a polysemous marker in Arawak languages. *International Journal of American Linguistics*, 84, 1-49.
- Aikhenvald, A. Y. 2018b. "Me, us and others : Expressing the self in Arawak languages of South America. In M. Huang and K. Jaszczolt (Eds.). *Expressing the Self: Cultural Diversity and Cognitive Universals* (pp. 13-39). Oxford: Oxford University Press.
- Aikhenvald, A. Y. and R. M. W. Dixon. 2011. Non-ergative associations between S and O. In Aikhenvald, Alexandra Y. and R. M. W. Dixon. *Language at large. Essays on syntax and semantics* (pp. 143-69). Leiden: Brill.
- Aikhenvald, A. Y. and D. Green. 1998 Palikur and the Typology of Classifiers. *Anthropological Linguistics*, 40, 429-480.
- Aikhenvald, A. Y. forthcoming. A view from the north: genders and classifiers in North Arawak languages of the Upper Rio Negro.
- Alvarez, J. 1994. *Estudios de Lingüística Guajira*. Maracaibo: Gobernación del Estado Zulia, Secretaría de Cultura.
- Bendor-Samuel, J. T. 1966. Some prosodic features in Terêna. In C. Bazell et al. (Eds.), *In Memory of J. R. Firth* (pp. 30-39). London: Longmans.
- Bezerra, Z. A. 2005. Gramática Kuripako. Mimeographed. Missão Novas Tribos do Brasil.
- Brandão, A. P. B. 2014. A reference grammar of Paresi-Haliti (Arawak). PhD dissertation. The University of Texas at Austin.
- Carlin, E. B. 2006. Feeling the need: the borrowing of Cariban functional categories into Mawayana (Arawak). In A. Y. Aikhenvald & R. M. W. Dixon (Eds.). *Grammars in contact: a cross-linguistic typology* (pp. 313-332). Oxford: Oxford University Press.
- Danielsen, S. 2007. *Baure: An Arawak language of Bolivia*. Leiden, The Netherlands: CNWS.
- Danielsen, S. & T. Granadillo. 2007. Agreement in two Arawak languages: Baure and Kurripako. In S. Wichman & M. Donohue (Eds.), *The typology of semantic alignment* (pp. 396-411). Oxford: Oxford University Press.
- Derbyshire, D. C. 1986. Comparative survey of morphology and syntax in Brazilian Arawakan. In D. C. Derbyshire and G. K. Pullum (Eds.), *Handbook of Amazonian Languages*, Vol. 1 (pp. 469-566). Berlin: Mouton de Gruyter.
- Dixon, R. M. W. 2010. *Basic Linguistic Theory*. Volume 2. *Grammatical topics*. Oxford: Oxford University Press.
- Dixon, R. M. W. 2012. *Basic Linguistic Theory*. Volume III. *Further grammatical topics*. Oxford: Oxford University Press.

- Ekdahl, E. M. and Butler, N. E. 1979. *Aprenda Terêna*. Vol. 1. Brasilia: SIL.
- Ekdahl, M. and Grimes, J. E. 1964. Terena verb inflection. *International Journal of American Linguistics*, 30, 261-8.
- Facundes, S. 2000. The language of the Apurinã people of Brazil. PhD diss., State University of New York at Buffalo.
- Fontaine, L. 2008. *Paroles d'échanges et règles sociales chez les indiens yucuna d'Amazonie colombienne*. Paris: L'Harmattan.
- Green, D. and H. Green. 2016. *Kagta Yuwitekne. Parikwaki-Parantunka. Dicionário Palikur-Português*. SIL Brasil: www.sil.org.
- Green, H. and Green, D. 1972. *Surface structure of Palikur grammar*. Brasilia: SIL.
- Green, L. and D. Green. 2013. *Knowing the day, knowing the world. Engaging Amerindian through in public archaeology*. Tucson: The University of Arizona Press.
- Hanson, R. 2010. A grammar of Yine (Piro). PhD thesis, La Trobe University, Melbourne.
- Klumpp, D. 1990. *Piapoco grammar*. Colombia: SIL.
- Lathrap, D. W. 1970. *The Upper Amazon*. New York: Praeger Publishers.
- Matteson, E. 1965. *The Piro (Arawakan) language*. Berkeley and Los Angeles: University of California Press.
- Medeiros, M. 1990. Posse nominal na língua Mehinaku, paper presented at the Annual meeting of Asociación de Linguística y Filología da America Latina, Campinas.
- Meléndez Lozano, M. A. 1998. *La lengua Achagua. Estudio gramatical*. Colciencias: Universidad de los Andes.
- Michael, L. D. 2008. Nanti evidential practice: Language, knowledge, and social action in an Amazonian society. PhD diss., Univ. of Texas at Austin.
- Michael, L. D. and T. Granadillo. 2014. (Eds.). *Negation in Arawak languages*. Leiden: Brill.
- Mihas, E. 2012. Ideophones in Alto Perené (Arawak) from Eastern Peru. *Studies in Language*, 36, 300–343.
- Mihas, E. 2014. Nominal and verbal temporal morphology in Ashéninka Perené (Arawak). *Acta Linguistica Hafniensia: International Journal of Linguistics*, 45, 1: 1–30.
- Mihas, E. 2015. *A grammar of Alto Perené (Arawak)*. Berlin: De Gruyter.
- Mihas, E. 2017. The Kampa subgroup of the Arawak language family. In A. Y. Aikhenvald & R. M. W. Dixon (Eds.), *The Cambridge Handbook of Linguistic Typology* (pp. 782–814). Cambridge: Cambridge University Press.
- Mihas, E. Forthcoming. Genders and classifiers in Kampa (Arawak) languages of Peru.
- Mithun, M. 1984. The evolution of noun incorporation. *Language*, 60, 847–94.
- Nordlinger, R. and L. Sadler. 2004. Nominal tense in cross-linguistic perspective. *Language* 80, 776–806.
- Olza Zubiri, J. and M. Â. Jusayú. 2012. *Gramática de la lengua guajira. Morfosintaxis*. Caracas: Universidad Católica Andrés Bello.
- Olza Zubiri, J., C. Nuni de Chapi, and J. Tube. 2001. *Gramática Moja Ignaciana (morfosintaxis)*. San Cristóbal, Venezuela: Univ. Católica del Táchira.
- Payne, D. L. 1981. *Phonology and morphology of Axininca Campa*. Dallas: SIL and University of Texas at Arlington.
- Payne, D. L. 1991. A classification of Maipuran (Arawakan) languages based on shared lexical retentions. In D. C. Derbyshire and G. K. Pullum (Eds.), *Handbook of Amazonian Languages*, Vol. 3 (pp. 355–499). Berlin: Mouton de Gruyter.
- Payne, D. L. 2002. Causatives in Asheninka: The case for a sociative source. In Masayoshi Shibatani (Ed.), *Causatives. The Grammar of Causation and Interpersonal Manipulation* (pp. 485–505). Amsterdam: John Benjamins.
- Payne, J. K. 1989. *Lecciones para el aprendizaje del idioma Ashéninka*. Série Lingüística Peruana No. 28. Yarinacocha: Instituto Lingüístico del Verano.

- Patte, M. 1989. *Estudio descriptivo de la lengua Añun (o 'Paraujano')*. San Cristóbal: Universidad Católica del Táchira.
- Pet, W. J. A. 1987. Lokono Dian, The Arawak Language of Suriname: A sketch of its grammatical structure and lexicon. PhD dissertation, Cornell University.
- Ramirez, H. 1992. *Le Bahuana, une nouvelle langue de la famille Arawak*. Chantiers Amerindia. Supplement 1 au No. 17, d *Amerindia*, Paris.
- Ramirez, H. 2001. *Uma gramática do Baniwa do Içana*. Manaus: Universidade Federal do Amazonas.
- Reinoso Galindo, Andrés E. 2002. *Elementos para una gramática de la lengua piapoco*. Bogotá: Ministério de cultura.
- Richards, J. 1973. Dificuldades na análise de possessão nominal na língua Waurá. *Série Lingüística*, 1, 11-29.
- Richards, J. 1988. A estrutura verbal Waurá. *Série Lingüística*, 9, 197-218.
- Rose, Françoise. 2013. Los generolectos del mojeño. *Liames*, 13, 115–134.
- Romling da Silva, Glauber. 2016. Predicados monovalentes paresi-haliti (Arawak): paralelismo entre nomes e verbos. In F. Queixalós and D. M. Gomes (Eds.), *O sintagma nominal em línguas amazônicas* (pp. 291-328). Campinas: Pontes editores.
- Rowan, O. and Burgess, E. 1979. Paresis grammar, *Arquivo Lingüístico* 149. Brasilia: SIL.
- Shaver, H. 1996. *Diccionario Nomatsiguenga-Castellano Castellano Nomatsiguenga*. Perú: Ministério de Educación, ILV.
- Smith-Stark, S. 1974. The plurality split. *Papers from the Annual Regional meeting of the Chicago Linguistic Society*, 10, 657-671.
- Socorro Sánchez, M. 2005. Morfología y syntaxis del Baniva. PhD dissertation, Universidad de los Andes.
- Taylor, G. 1991. *Introdução à língua Baniwa do Içana*. Campinas: Editora da Unicamp.
- Taylor, G. 2010. O “Caduco” e o “Frustrativo” nas Línguas Baniwa do Içana e Nheengatu (Alto Rio Negro, Brasil). In E. Carlin & S. van de Kerke (Eds.), *Linguistics and Archaeology in the Americas* (pp. 207–214). Leiden: Brill.
- Tracy, F. V. 1974. An introduction to Wapishana verb morphology. *International Journal of American Linguistics*, 40, 120-6.
- Wise, M. R. 1971. *Identification of participants in discourse: A study of aspects of form and meaning in Nomatsiguenga*. SIL Publications in Linguistics and Related Fields 28. Norman: University of Oklahoma and SIL.
- Wise, M. R. 1976. Apuntes sobre la influencia inca entre los amuesha, factor que oscurece la clasificación de su idioma. *Revista del Museo Nacional*, 42, 355-66.
- Wise, M. R. 1986. Grammatical Characteristics Of Preandine Arawakan Languages Of Peru. In D. C. Derbyshire and G. K. Pullum (Eds.), *Handbook of Amazonian Languages*, Vol. 1 (pp. 567-642). Berlin: Mouton de Gruyter.
- Wise, M. R. 1990a. Afijos causativos y comitativos en idiomas de la familia arawaka maipurán. In Rodolfo Cerrón-Palomino & Gustavo Solís Fonseca (Eds.), *Temas de lingüística Amerindia. Primer Congreso Nacional de Investigaciones Lingüístico-Filológicas* (pp. 291-307). Lima: ConcyTec.
- Wise, M. R. 1990b. Valence-changing affixes in Maipurán Arawakan Languages. In Doris L. Payne (Ed.), *Amazonian Linguistics: Studies in Lowland South American Languages* (pp. 89-116) Austin: University of Texas Press.
- Wise, M. R. 2002. Applicative affixes in Peruvian Amazonian languages. In Mily Crevels, Simon van de Kerke, Sérgio Meira, & Hein van der Voort (Eds.), *Current Studies on South American Languages. Selected papers from the 50th International Congress of Americanists in Warsaw and the Spinoza Workshop on Amerindian Languages in Leiden, 2000* (pp. 329-344). Leiden: CNWS.

## Endnotes

<sup>1</sup> The genetic unity of Arawak languages was first recognized by Father Filippo Salvatore Gilij as early as 1783. The recognition of the family was based on a comparison of pronominal cross-referencing prefixes in Maipure, a now extinct language from the Orinoco Valley, and in Mojo, a South Arawak language from Bolivia. The limits of the family were established by the early twentieth century. The internal classification of Arawak languages remains a matter of some debate. A number of grammatical and lexical traits distinguish languages spoken to the north of the River Amazon from those spoken to the south. Well-established subgroups include

(a) Kampa in Peru and adjacent regions of the state of Acre in Brazil;

(b) South Arawak languages in Brazil and Bolivia;

(c) Xinguan Arawak languages;

(d) a few small North Arawak groupings in Brazil, Colombia and Venezuela including

(i) the *Ta*-Arawak subgroup — with Guajiro (or Wayyu-naiki), Añun (or Parauhano) in the region of Peninsula Guajira in Venezuela and Colombia, Garifuna (or Black Carib) in Central America, and Lokono (alternatively known as Dian, or Arawak) in Guyana, French Guyana and Suriname (and also the long-extinct Taino, the language of the first indigenous group encountered by Columbus)

(ii) The Uapuí subgroup in the Upper Rio Negro Basin (consisting of Tariana, Baniwa of Içana-Kurripako continuum, Piapoco, Cabiari, Resígaro, and Guarequena).

An alternative name for Arawak family is Maipurán. The term *Arawakan* was earlier used for the combination of a well established genetic subgrouping (known as *Maipurán*) and a number of other groups not demonstrably related and is hence avoided by the majority of qualified experts in the field and linguists with sound comparative-historical background. A detailed discussion of the family is in Aikhenvald (1999, 2002, 2012: 32-6); a comprehensive bibliography and an up-to-date classification is in Aikhenvald (2016a). Throughout this article, languages are referred to with the names in the quoted sources.

<sup>2</sup> Wise (1990a: 306) offers examples of cognacy between a valency increasing prefix in one language and a suffix with a similar meaning in another, e.g. Apurina *mi-* 'comitative applicative' and Ashaninka *-imo* 'comitative', Asheninka *-imo* 'presential, comitative, benefactive applicative'. At present, these correlations cannot be considered conclusive, since most of these morphemes are monosyllabic and their meaning ranges are rather broad.

<sup>3</sup> See Wise (1986), Aikhenvald (2012a: 135-6) on variable affix placement in Amazonian languages.

<sup>4</sup> An example of variable morpheme order in Amuesha is in Wise (1986: 583). In a few other instances in Kampa languages, changing position of the suffixes has no semantic effect (see Wise, 1986: 583 and Mihas, 2015: 103). Variable morpheme order is a feature of a number of highly synthetic languages of Lowland Amazonia (Aikhenvald, 2017a).

<sup>5</sup> This similarity led some researchers (e.g. Admiraal and Danielsen, 2014) to surmise that classifiers in their derivational functions can be considered a type of compounding. Compounds may form one or several phonological words, and one grammatical word (see Aikhenvald, 2007a: 30-3 and references there for a classification of nominal compounds).

<sup>6</sup> Non-verbal predicates may include nouns, adjectives, manner adverbs and time words.

<sup>7</sup> Shared person marking on the noun and on the verb is a recurrent feature of many Amazonian languages (a summary is in Aikhenvald, 2012: 176). This polyfunctionality led some researchers to assume that these prefixes must be proclitics since they can be conceived as displaying freedom of movement (e.g. Brandão, 2014). This is not a convincing argument, since in no Arawak language do personal prefixes display any properties associated with



clitics of any sort. Based on this one shared feature between nouns and verbs in Paresi, Romling da Silva (2016) argues in favour of nouns being a type of intransitive verbs.

<sup>8</sup> A cognate of the same marker appears in Alto Perené *-rentsi*, Machiguenga *agaNtsi* (Mihas, 2015: 426; Michael, 2008: 303).

<sup>9</sup> For Arawak languages it would be inaccurate to say that bound pronouns arose as a consequence of grammaticalization of currently attested independent personal pronouns. Extant independent personal pronouns can be transparently traced to personal prefixes accompanied by a derivational formative. Their forms and exact use vary. In languages with no pronominal suffixes or enclitics (such as Bare), independent pronouns can be used in non-subject functions (Aikhenvald, 1995b: 32). The largest system in the family, with four types of independent pronouns (each of which can occur together with bound pronominal prefixes and suffixes) has been described for Alto Perené, a Kampa language, by Mihas (2015: 128-34, 608-11). It is of course quite possible — even likely — that at a much earlier stage the pronominal suffixes could have resulted from grammaticalization of erstwhile pronouns which have been lost.

<sup>10</sup> Aikhenvald (2012: 351, 358), Payne (1991: 414, 421-2).

<sup>11</sup> Double case marking in Tariana (Aikhenvald, 2003: 158-162) may mark two grammatical functions of the noun within the main and the embedded clause.