

This paper will be published in Pepper et al. (eds.) (2022)

Please cite as:

Rose, Françoise & An Van linden. 2022. The derivational use of classifiers in Western Amazonia. In Steve Pepper, Francesca Masini & Simone Mattiola (eds.), *Binominal lexemes in cross-linguistic perspective: Towards a typology of complex lexemes* [Empirical Approaches to Language Typology 62], 237-276. Berlin: De Gruyter Mouton. [<https://doi.org/10.1515/9783110673494-008>]

Table of contents

1	Introduction	2
2	Nouns and binominals	5
2.1	Simple nouns: bound nouns and independent nouns	6
2.2	Adnominal possession construction	7
2.3	Nouns with derivation suffixes	8
2.4	Binominals with a preposition.....	9
2.5	N-N compounds	9
2.6	N N juxtaposition	11
2.7	N-CLF derived nouns.....	11
2.8	Why N-CLF formations deserve closer attention	12
3	Classifiers vs. bound nouns	13
3.1	Inventories, form and meaning.....	13
3.2	Phonology and prosody	16
3.3	Syntax.....	18
3.4	Functions	22
3.4.1	Functions on nouns.....	22
3.4.2	Functions in verb forms.....	26
4	Conclusions and diachronic implications.....	29
5	References	32

The derivational use of classifiers in Western Amazonia

Françoise Rose¹ and An Van linden²

Abstract: Western Amazonian languages stand out in having classifiers that – in addition to the well-established classifier environments – also appear as derivational devices on nouns. Since classifiers are commonly assumed to originate in nouns, classifier languages confront us with an analytical problem in the domain of binominals, i.e. how to distinguish between the derivational use of classifiers on nouns, and noun-noun compounds. The present paper addresses this problem on the basis of primary data from Harakmbut (isolate, Peru) and Mojeño Trinitario (Arawak, Bolivia), two unrelated Western Amazonian languages. As a factor bearing on this problem, we show that in both languages the noun/classifier distinction is blurred by the fact that there is a class of nouns that share many features with the canonical classifiers, i.e. that of bound nouns. In this paper, we discuss how noun-classifier derivation differs from noun-noun compounding, or classifiers from bound nouns for that matter, at the phonological, semantic and syntactic levels in both languages.

1 Introduction³

Amazonian nominal categorization systems are known for challenging the traditional view of the distinction between noun classes and classifiers (Payne 1987; Grinevald and Seifart 2004). Among these, Western Amazonian classifiers are particularly interesting for two reasons. First, they typically form multiple-classifier systems (Aikhenvald 2000), with the same set of classifiers occurring in different syntactic environments. Second, their classifiers are multifunctional (Krasnoukhova 2012), in having the three functions of categorization, derivation and, less importantly, agreement. This paper focuses on the derivational use of classifiers in Western Amazonian languages, previously noted in the literature (Payne 1987; Aikhenvald 2000: 220; Seifart and Payne 2007; Petersen de Piñeros 2007; Krasnoukhova 2012: 209; Brandão 2016; Wojtylak 2016), but not yet extensively described, neither in individual languages nor cross-linguistically.⁴ It will deal more specifically with the derivational use of classifiers on nominal roots,⁵ which is functionally equivalent to noun-noun compounding, and thus participates in the formation of binominal lexemes, i.e. “lexical items that consist primarily of two nominal

¹ Françoise Rose, Dynamique du Langage, UMR5596, CNRS & Université de Lyon (France). 16 avenue de Berthelot, 69007 Lyon, France, francoise.rose@cnrs.fr.

² An Van linden, University of Liège (Belgium), KU Leuven (Belgium) & Collegium de Lyon (France). Department of Modern Languages, University of Liège, Place Cockerill 3 - 5, 4000 Liège, Belgium, an.vanlinden@uliege.be.

³ Authorship of this paper is shared jointly. Work on Mojeño Trinitario has been supported by the ASLAN Laboratoire d'excellence, Lyon, France. Work on Harakmbut has been made possible by mobility grants and postdoctoral grants from the Research Foundation, Flanders (FWO) (2009–2014) and the Fund for Scientific Research (FNRS) (2015–2016), as well as research grants from the research council of the KU Leuven (GOA/12/007 & C14/18/034). The paper was revised during Van linden's research stay in Lyon, supported by the Collegium de Lyon and the ASLAN Laboratoire d'excellence. We thank Olga Krasnoukhova, as well as the editors and referee for helpful comments on earlier drafts.

⁴ But see Seifart (2005: 106-122) for a discussion of the derivational uses of noun classes on nouns in Miraña.

⁵ Classifiers can also be used as nominalizing devices on verbs, see Krasnoukhova (2012: 210).

constituents and whose function is to name a (complex) concept that involves an unstated (or underspecified) relation between two entities” following the definition of Pepper (2020: 1, see also the introduction to this volume). The functional resemblance of classifiers on nouns and binominal compounds is illustrated in (1)-(2) from Harakmbut (ISO code AMR, see below), with classifier-derived nouns (hereafter N-CLF) in (1), and noun-noun compounds (N-N) in (2).⁶

Harakmbut

- | | | | |
|-----|----|---|-------|
| (1) | a. | <i>siro-pi</i> [metal-CLF:stick] ‘knife’ (cf. Hart 1963: 1) | N-CLF |
| | b. | <i>siro-pu?</i> [metal-CLF:cylindrical;hollow] ‘metal tube’ (cf. Hart 1963: 1) | N-CLF |
| (2) | a. | <i>ndumba-kuwa</i> [forest-dog] ‘bush dog’ (Helberg 1984: 252; Tripp 1995: 194) | N-N |
| | b. | <i>ãwīt-ku</i> [giant.otter-head] ‘giant otter’s head’ (Hart 1963: 3) | N-N |

The constructional approach to binominal lexemes that sees roots and affixes as end-points on a continuum (Pepper 2020: 15) is particularly useful to investigate the use of classifiers in binominal lexeme formation. Indeed, since classifiers are commonly assumed to originate in nouns (Mithun 1986: 395; Aikhenvald 2000), more specifically in compounds (cf. Seifart 2010), and are often difficult to distinguish in a straightforward manner (Dixon 1986: 106), classifier languages confront us with an analytical problem in the domain of binominals: how can we distinguish between the derivational use of classifiers on nouns, and noun-noun compounds? This boils down to discussing the empirical realization of the types **cls** “classifier” and **cmp** “compounding” of Pepper’s (2020: 145-169) typology of binominals (see Section 2). The nominal roots found in compounds are also themselves considered to be on the continuum between affixes and canonical nominal roots. Addressing this empirical question therefore results in studying the language-internal competition between classifier derived nouns (N-CLF) and noun-noun compounds (N-N), and involves tackling more general and theoretical questions like the definition of classifiers, and the distinction between derivational affixes and bound roots (Lieber and Štekauer 2009).

The present paper addresses the competition between two binominal constructions (N-CLF and N-N) through the comparison of two Western Amazonian languages: Mojeño Trinitario and Harakmbut. These languages are genetically unrelated and not in contact, and are being studied by the authors on the basis of primary data collected in the field. Mojeño Trinitario (ISO code TRN) is an Arawak language spoken in the lowlands of Bolivia, in and around the old missionary town of Trinidad and in the Isiboro-Sécure territory. It is spoken by around 3000 speakers (Crevels and Muysken 2009) and is endangered by the gradual loss of inter-generational

⁶ **Abbreviations:** 1: first person; 2: second person; 3: third person; ACC: accusative; AN: animate; ACT: active; ART: article; CLF: classifier; DEM: demonstrative; DERIV: derivative; DEP: dependent verb form; DIM: diminutive; DIST: distal; DIST.PST : distant past; F: feminine; FAN: addictive; GEN: genitive; GRN: general relational noun; HAB.A: habitual actor; IMP: imperative; IND : indicative; INS: instrumental; IRR: irrealis; LOC: locative; M: masculine; NH: non-human; NOM: nominative; NPF: noun prefix; NPOSD: non-possession; NVOL: non-volitional; PFV: perfective; PL: plural; PLURACT: pluractional; PREP: preposition; PSD: possessed; REC.PST: recent past; SG: singular; SOC: sociative causative; SPAT: spatial; SP.P.NZ: specific patient nominalizer; SUBST: substitutive; TRNS: transitiviser; VPL: verbal plural

transmission. Investigation on Mojeño benefits from previous work on the language (Gill 1957), on its sister language Mojeño Ignaciano (especially Olza et al. 2002), on a variety spoken in the 17th century in Jesuit missions (Marbán 1702), and on historical work on the Arawak family (Payne 1991a *inter alia*). A basic introduction to Mojeño Trinitario is Rose (2015). Harakmbut is a Peruvian Amazonian language spoken in the departamentos of Cusco and Madre de Dios. It is considered an isolate (Tovar 1961; Loukotka 1968; Lyon 1975; Helberg 1984; Wise 1999: 307), as Adelaar’s (2000, 2007) proposal of a genetic link with the Brazilian Katukina family still awaits further corroboration. Earlier work on Harakmbut has mainly focused on the most vital dialect, i.e. Amarakaeri⁷ (ISO code AMR) (Hart 1963; Helberg 1984, 1990; Tripp 1976, 1995), as does the present paper; Van linden (2022) presents a basic description. With about 1000 speakers left (Moore 2007: 46), the language is highly endangered, also aggravated by parents’ reluctance to pass on the language to their children.

Because the boundary between classifiers and nouns is not straightforward in these two languages, we will compare classifiers and nouns within and beyond binominal lexemes in order to distinguish compounding and derivation through classifiers in binominals. In general terms, classifiers are morphemes providing an overt categorization of nominals (Grinevald Craig 2004: 1016). They encode “some salient perceived or imputed characteristics of the entity to which an associated noun refers” (Allan 1977: 285). The two classifier systems differ in that Mojeño Trinitario is a multiple and multifunctional classifier language with an extensive set of classifiers formally rather distinct from nouns, but showing almost the exact same syntactic distribution, while Harakmbut has a small set of classifiers that are formally identical to nominal roots but show a distribution distinct from that of nouns. In this language, classifiers are mainly used as verbal classifiers;⁸ they are hardly used as categorizing devices on nouns. Yet, both languages show patterns of word formation in which classifiers function as derivational affixes, which are at first sight not easy to distinguish from noun-noun compounds.

In this paper, we will first present the different types of nouns and binominals found in the two languages (Section 2). Importantly, we will introduce the class of bound nouns (Section 2.1), which contribute significantly to the analytical problem focused on here. In Section 3, we will compare classifiers and bound nouns in the two languages at different levels of analysis, thus proposing a methodology to distinguish between N-CLF and N-N in any classifier language. We will start with the result of our analyses, providing an inventory of the sets of classifiers and bound nouns in the languages studied, and describe their form and meaning (Section 3.1). We will then discuss how classifier-derived nominals compare to noun-noun compounds, or classifiers to bound nouns for that matter, at several levels, viz. phonological/prosodic, syntactic and semantic (Sections 3.2 to 3.4). In the process, we will also discuss the theoretical question of

⁷ Speakers of the Amarakaeri variety regard the label ‘Amarakaeri’ as a derogatory term, as it means ‘(fierce) murderer’, going back to an ancient story about the origin of the different ethnolinguistic groups of the Harakmbut people. They prefer to call their variety ‘Arakmbut’.

⁸ Or, following Passer (2016: 17), “verb” classifiers rather than “verbal” ones (e.g. Aikhenvald 2000: chap. 6).

whether classifiers are thing-morphs, i.e. morphs that denote a thing (Pepper 2020: 12), a concept extended from that of thing-roots, i.e. roots that denote an (animate or inanimate) physical object (Haspelmath 2012). Thing-morphs themselves cover both thing-roots and thing-affixes. Finally, we will summarize our findings, and elaborate on their diachronic implications (Section 4).

2 Nouns and binominals

This section presents the different types of binominal lexemes (or constructions) found in Mojeño Trinitario and Harakmbut. After presenting the two classes of simple nouns (Section 2.1), we list the different types of binominals (Sections 2.2 to 2.7), with special attention to N-N compounds (Section 2.5) and classifier-derived nouns, the phenomenon central to this paper (Section 2.7). Section 2.8 elaborates why classifier-derived nouns deserve closer attention. Section 3 will then focus on word formation, i.e. the creation of one-word lexemes, restricting the discussion to the **cmp** and **cls** constructions of Pepper's (2020) typology.

First we give a preliminary quantitative account of the different types of binominals attested in the two languages, on the basis of the list of 100 complex concepts designed by Pepper (2020: 391-392). We classified the data collected with this list into Pepper's nine types of binominals (Pepper 2020: 145-169), listed below with simplified definitions:

- **jxt**: juxtaposition of two separate thing-roots without any additional element
- **cmp**: compounding of two thing-roots in a single word
- **der**: derivation from a thing-root with a thing-affix that contributes some semantic content
- **cls**: thing-root with a classifier, where the denotatum of the binominal is different from that of the base (the classifier is used to derive a new meaning rather than for classification)
- **prp**: head and modifier are independent lexemes, and an additional lexeme forms a constituent with the modifier
- **gen**: head and modifier are independent lexemes, with an additional word-class preserving morpheme attached to the modifier
- **adj**: head and modifier are independent lexemes, with an additional word-class changing morpheme attached to the modifier
- **con**: head and modifier are independent lexemes, with an additional word-class preserving morpheme attached to the head
- **dbl**: head and modifier are independent lexemes, with additional morphemes attached to both.

For Mojeño Trinitario, 90 items have been collected for 88 concepts. Out of these, 27 items are binominals, instantiating five types (**con**, **der**, **cmp**, **cls**, **jxt**) of Pepper's nine types (Pepper

2020: 145-169), as detailed in Table 1.⁹ Most notable is the fact that 10 of the 27 binominals involve the derivational use of a classifier: it is the most common binominal structure, and the topic of the present paper. The items that are not binominals are either simple forms (often borrowings) or deverbal nominalizations, not counting as binominals.¹⁰

Out of the 78 Harakmbut data items collected for 72 entries of Pepper’s (2020: 391-392) list, 29 are binominals. These instantiate four of the nine types, as presented in Table 1; there are no examples of **jxt**, **prp**, **adj**, **con** and **dbl**. The predominant type is **cmp** (15 out of 29 items), while **cls** and **gen** are instantiated 6 times each.¹¹ It should be noted, however, that a few **cmp** and **gen** examples also involve classifiers, which – if ranged with **cls** – would yield a 28% share of **cls** (8 out of 29 binominals). Non-binominal items include 19 simple forms, two descriptive phrases, and 28 verb-based items. Taken together, therefore, noun-noun compounds and classifier-derived nouns constitute two major devices to form binominals in these two Western Amazonian languages.

	jxt	cmp	der	cls	gen	con
TRN (27)	11%	19%	22%	37%	0%	11%
AMR (29)	0%	52%	7%	21%	21%	0%

Table 1: Mojeño Trinitario and Harakmbut’s binominals in Pepper’s list of 100 complex concepts

Before presenting the various types of binominals found in Mojeño Trinitario and Harakmbut, it is important to distinguish two classes of simple nouns, bound and independent nouns.

2.1 Simple nouns: bound nouns and independent nouns

In both languages, noun roots are easily distinguished from verb roots or other words classes, mainly on the basis of their morphological potential. As in many Amazonian languages (Krasnoukhova 2012), there is a crucial distinction between two classes of nominal roots, i.e. bound nouns and independent nouns.¹² This is a morphological distinction, in that bound nouns

⁹ Section 2.4 discusses a marginal construction of the **prp** type, which was not illustrated in the Mojeño Trinitario translations of the list of 100 complex concepts.

¹⁰ Interestingly, among the 20 simple native forms, five suspiciously show final syllables homophonous with classifiers (without being synchronically segmentable), and out of the 28 deverbal nominalizations, 12 involve classifiers as a derivational device. These additional uses of classifiers go beyond the topic of this paper.

¹¹ Note that these numbers are different from the counts in Pepper (2020: 441, 479). For instance, Pepper overlooked some spatial markers in his analysis, which inflated his **cmp** type (19 items) at the expense of his **gen** type (3 items).

¹² These classes are sometimes called obligatorily vs. non-obligatorily possessed nouns, or inalienable vs. alienable nouns.

require some morphology to constitute a word, unlike independent nouns.¹³ The former category is illustrated in (3) and (5); the latter in (4) and (6).

Trinitario

- | | | | |
|-----|----|--|------------------|
| (3) | a. | <i>wiye</i> [ox] ‘ox’ | independent noun |
| | b. | <i>n-wiye-ra</i> [1SG-ox-PSD] ‘my ox’ | |
| (4) | a. | <i>n-juma</i> [1SG-sickness] ‘my sickness’ | bound noun |
| | b. | <i>jma-re</i> [sickness-NPOSD] ‘sickness’ | |

Harakmbut

- | | | | |
|-----|----|---|------------------|
| (5) | a. | <i>pagŋ</i> [father] ‘father’ | independent noun |
| | b. | <i>ndoʔ-edn pagŋ</i> [1SG-GEN father] ‘my father’ | |
| (6) | a. | <i>ndoʔ-edn-ndik</i> [1SG-GEN-name] ‘my name’ | bound noun |
| | b. | <i>wa-ndik</i> [NPF-name] ‘name’ | |

Independent nouns may occur as nominal heads without affixes, cf. *wiye* in (3) and *pagŋ* in (5). Bound nouns, by contrast, never occur as nominal heads without affixes, cf. *-juma* in (4) and *-ndik* in (6). In Mojeño Trinitario, bound nouns take possession-related affixes to reach wordhood status (Rose 2020). Typically they take a person prefix for their possessor, cf. (4a), but a sub-class of bound nouns, including *-juma* in (4), instead take some derivational affix, e.g. the non-possession suffix *-re* in (4b), when no possessor is specified. Independent nouns, in turn, do not need extra affixes when unpossessed (3a). But when they are possessed, a sub-class of them, including *wiye* in (3b), require a derivational suffix (such as *-ra*) in order to take person prefixes. The situation is somewhat different in Harakmbut, where bound nouns require a noun prefix, *wa-* or *e-*, to obtain independent nominal status, cf. (6b).¹⁴ Unlike independent nouns (5b), they can also attach to adnominal modifiers to form one prosodic word, as in (6a) with a genitive-marked pronoun (cf. Van linden 2021, 2022).

This paper will show that bound nouns are more often found in compounds and incorporated in verb forms than independent nouns, thus showing a distribution roughly similar to that of classifiers. Classifiers constitute yet a different category; they are morphologically bound elements, but not nouns. They will be systematically compared to bound nouns in Section 3.

2.2 Adnominal possession construction

Mojeño Trinitario and Harakmbut have rather different adnominal possession constructions, exemplified in (7) and (8) respectively. In terms of Pepper’s typology of binominals, Mojeño Trinitario adnominal possession instantiates the **con** type, because it has an additional marker on the head, whereas the Harakmbut construction exemplifies the **gen** type, because it has an additional marker on the modifier.

¹³ This is why, in this paper, bound nouns cited within the text are preceded by a hyphen.

¹⁴ These noun prefixes have been analysed as nominalizing prefixes in Van linden (2022).

- (7) TRN *to* *t(a)-og'e* *to* *kwoyu*
 ART.NH 3NH-body ART.NH horse
 ‘the body of the horse’
- (8) AMR *apetpet-en* *hak*
 jaguar-GEN house
 ‘the jaguar’s den’

In Mojeño Trinitario (7), the adnominal possession construction is made up of two noun phrases (NPs), each consisting minimally of a noun preceded by an article, indicating that the noun is specific. N1 expresses the possessee and takes a person prefix referring to the possessor expressed by N2.¹⁵ In Harakmbut (8), in which NPs minimally consist of a noun only, the order in adnominal possession constructions is that of possessor – possessee, and genitive case is marked on the possessor noun (cf. Tripp 1995: 195).

In addition to structures with a genitive marker in Harakmbut, Pepper (2020: 441, 479) also analyses one-word structures including a spatial element linking two noun roots as instances of the **gen** type. An example is given in (9).

- (9) AMR *wa-mbaʔ-taʔ-meh* [NPF-hand-SPAT:base-hump] ‘wrist’

In (9), the spatial affix *-taʔ-* links two thing-roots: a wrist is a hump at the base of the hand. As the first noun root semantically modifies the second noun root and as the spatial affix does not change the word class of the modifier noun, structures like (9) are indeed similar to genitive constructions like (8). Other spatial linking elements in Harakmbut, like *-(o)k-* and *-ti-*, point to different spatial configurations of the component elements (cf. Hart 1963).

2.3 Nouns with derivation suffixes

Both languages have the **der** type of binominals, made of a nominal root and a derivational morpheme. Mojeño Trinitario uses a number of derivational suffixes, five of which are exemplified in Pepper’s (2020) list. They are the diminutive *-gira* (10), the substitutive *-ra’o* (11) (also used to derive ‘step-father’ from ‘father’), the additive *-more ~-mre* ‘fan of’ (12), the ‘habitual actor’ *-eru* (13), and the non-possessed *-re* (normally used to allow a bound noun to occur without a person prefix, cf. (4b); in Section 2.1) resulting in a non-compositional meaning in (14).

¹⁵ A sub-class of independent nouns, viz. those that cannot take a possessive prefix, use a juxtaposed generic bound noun *-ye’e* that carries the possessive prefix (Rose 2015). This sub-class is labelled “non-directly possessable nouns” in Rose (2020).

Trinitario

- (10) *kwoy-gira* [horse-DIM] ‘foal’
- (11) *viya-ra'o* [Lord-SUBST] ‘chieftain’¹⁶
- (12) *'san-ti-mre* [field-NPOSD-fan] ‘farmer (lit. fan of field)’
- (13) *tyuraj-eru* [mud-HAB.A] ‘potter’
- (14) *chut-re* [head-NPOSD] ‘skull’

Harakmbut also has a number of derivational suffixes, like *-eri* in (15); nominal bases suffixed by *-eri* refer to animate entities living in or coming from the place denoted by the nominal base. The two examples of the **der** type in Pepper’s (2020) list involve the diminutive suffix *-si?po* as in (16).

Harakmbut

- (15) *Porto-lus-eri* [Puerto-Luz-AN] ‘people living in/coming from Puerto Luz’
- (16) *wa-mbo-si?po* [NPF-youngster-DIM] ‘boy’

2.4 Binominals with a preposition

Binominals of the **prp** type, i.e. with an additional lexeme forming a constituent with the modifier, are found in Mojeño Trinitario only. This infrequent construction uses the preposition *te* (with very broad semantics), linking the head element in initial position to the modifying noun that follows. It should be noted that the first noun is always an independent noun borrowed from Spanish, such as *manteka* and *eskina* in (17).¹⁷

Trinitario

- (17) a. *manteka te jimo* [butter PREP fish] ‘fish grease’
- b. *j-mu-ena eskina-no te plasa* [DEM-NH.PL-DIST corner-PL PREP square]
‘those corners of the square’

2.5 N-N compounds

Both languages allow compounding of two nominal roots in a complex nominal word (**cmp** type). It is striking that the two noun classes (bound and independent) are not distributed evenly across N1 and N2 in N(1)-N(2) compounds, but rather show the same skewed distribution. In Table 2, we have split the **cmp** type of Table 1 according to different noun classes in N1 and N2 positions. Independent nouns (I in Table 2) are only rarely found as N2, whereas bound nouns (B in Table 2) frequently occur as N2.

¹⁶ *Viya* is a noun lexicalized from *vi-* 1PL and *iya* ‘father’ meaning both ‘Sir’ and ‘Lord’ (to refer to Christian god).

¹⁷ This construction, unattested in Old Mojeño, most likely emerged due to contact, facilitated by the formal resemblance between Spanish and Mojeño Trinitario prepositions *de* and *te*.

	I-I	I-B	B-B	B-I
Mojeño Trinitario (5)	0	1	4	0
Harakmbut (15)	0	5	9	1

Table 2: Types of N-N compounds according to morphological class of N1 and N2 in Pepper’s list of 100 complex concepts

In both languages, most typical N-N compounds, such as (18) and (19), are endocentric compounds in which N2 is the semantic head and N1 is semantically subordinate. The semantic head of a compound is identified on the basis that the concept expressed by the compound is a sub-class of the concept denoted by the head: “the whole compound must be a hyponym of its head” (Scalise and Fábregas 2010: 111).

(18) TRN *kasiki-yeno* [cacique-wife] ‘cacique’s wife’

(19) AMR *kaymãri-mbogŋ* [zungaro-lip] ‘lip of a zungaro fish’

In both languages, N2 is more rarely an independent noun; examples are given in (20) and (21). In Trinitario, N1 is the semantic head in such cases.

(20) TRN *’nuuku-mari* [hole-stone] ‘cave’

(21) AMR *wa-taʔpi-widn* [NPF-spine-stone] ‘kidney’ (Tripp 1995: 130b)

In Mojeño Trinitario, the selection of the gender value of the article (non-human, human plural, masculine singular or feminine singular) signals the morphological head of the compound, i.e. the element of the binominal which defines the formal properties of the compound as a lexical item (Scalise & Fábregas 2010: 124).¹⁸ Either N1 or N2 can be the morphological head. In (18), the compound *kasikiyeno* triggers a feminine singular form of the article, like *yeno* does, even though *kasiki* is masculine. However, in other endocentric compounds, it is N1 that determines the form of the article. In Harakmbut, there is no such way of determining the morphological head of N-N compounds.

In addition, we have also documented some exocentric compounds, i.e. compounds which are not hyponyms of any of their components (cf. Bauer 2001: 700), as in (22).

(22) TRN *ñi paku-miro* [ART.M dog-face] ‘a (male) being with a human body and a dog head’

In (22), neither *paku* ‘dog’ nor *miro* ‘face’ can be considered the semantic head of the compound, as the compound refers neither to a type of dog nor to a type of face, but – through metonymy – to a type of human being. That is, the semantic head ‘lies outside’ the component elements of the compound. This compound has no morphological head either: while both *paku* and *miro* trigger non-human agreement, *pakumiro* instead takes human agreement (see the singular masculine article *ñi* in (22)). The Harakmbut example (19) can also be used as an

¹⁸ Note that the morphological head of a binominal is by definition a sub-part of a lexeme and should not be confused with the syntactic head of a noun phrase, which is a full word.

exocentric compound. In that case, it refers to a person whose lips resemble these of a zungaro fish. In fact, metonymy based on animal body part nouns like (19) produces great nicknames among the Harakmbut community. Several aspects of N-N compounds will be described more in-depth in Section 3.

2.6 N N juxtaposition

Mojeño Trinitario also has a binominal construction of the **jxt** type, with two nominal words and without additional material, e.g. (23)-(24).

(23) TRN *ñi* *'chane* *'jiro*
 ART.M person man
 'the man'

(24) TRN *no* *n-jañono-no* *trinrano-no*
 ART.PL 1SG-relative-PL Trinitario-PL
 'my Trinitario relatives'

This construction differs from N-N compounds in that the two nominal roots do not form a single word, but are simply juxtaposed. Both nouns are separate phonological and morphological words, as shown by the plural marking on both N1 and N2 in (24). Moreover, unlike in most N-N compounds, N1 is the head, and N2 the modifier (N2 is not the head of a separate noun phrase as it lacks an article).¹⁹ This construction also differs from the adnominal possession construction both semantically and formally. Semantically, it does not involve a relation of possession between two nouns, but codes a variety of other relations between N1 and N2. For instance, N2 can be a hyponym (23), a synonym, a noun expressing substance or material, or a noun qualifying N1 (24); N1 can be a measure term with the countized element in N2, or N1 can be a title used with the proper name in N2. Formally, N1 is not possessed, unlike in the adnominal possession construction (see Section 2.2).

2.7 N-CLF derived nouns

Both languages have binominals based on a nominal root plus a classifier that does not 'classify' this root, but rather derives a new stem with a different meaning (see Section 3.4.1 for contrasting examples). We have analysed these examples as denominals based on a classifier (**cls**). The derivational use of classifiers is not characteristic of prototypical classifier systems (Krasnoukhova 2012: 209), but has been mentioned before for some classifier languages (Payne 1987: 28-29; Aikhenvald 2003: 84, 225). The classifier immediately follows the noun root to form a complex noun stem in both languages, cf. (25)-(26), and thus occupies the same slot as

¹⁹ There are examples of N2 preceded by an article in the semantic subtypes where either N2 expresses substance/material, or N1 is a measure term.

N2 in N-N compounds. In Mojeño Trinitario, some nouns derived with a classifier additionally take the derivational suffix *-rV* (with vowel harmony) between the root and the classifier, as in (27).²⁰

- (25) TRN a. *to yuk-pi* [ART.NH fire-CLF:fili] ‘a candle’
b. *to giorre-pi* [ART.NH snake-CLF:fili] ‘a worm’
c. *to 'o'e-pi* [ART.NH rainbow-CLF:fili] ‘an eel’
- (26) AMR a. *peraʔ-po* [rubber-CLF:sphere] ‘plastic ball’ (Hart 1963: 5)
b. *siro-po* [metal-CLF:sphere] ‘tin can’ (Hart 1963: 1)
c. *aymōrō-po* [honey-CLF:sphere] ‘bee’
- (27) TRN a. *n-iyépé-re-ku* [1SG-foot-deriv-CLF:path] ‘my footprint’
b. *v-emtone-re-pi* [1pl-labor-deriv-CLF:fili] ‘our way/process of working’

Semantically, the addition of the classifiers on these noun roots causes substantial changes of meaning, e.g. from ‘fire’ to ‘candle’ in (25a). Nevertheless, the semantic import of classifiers in N-CLF derived nouns is typically less specific than that of nouns in N-N compounds, whether independent or morphologically bound. In both languages, and presumably across languages, classifiers tend to have more general semantics than nouns, as they denote shapes, qualities or substances (see Aikhenvald (2000: 271-305) for an overview of the semantics of classifiers).²¹ This is in line with Lieber and Štekauer’s (2009: 5) semantic criterion to distinguish bound roots from derivational affixes, i.e. that “roots in some sense have more semantic substance than affixes.” At the same time, as is apparent from examples (25)-(26), in both languages classifiers have a semantic effect on nouns that is far less abstract than that of the – clearly derivational – diminutive suffixes in (10) and (16) or the ‘habitual actor’ suffix in (13) above; a candle in (25a) is an instance of a long flexible object related to fire. We will elaborate on the semantic import of classifiers in Section 3.4 and relate this to the question of whether classifiers are thing-affixes.

2.8 Why N-CLF formations deserve closer attention

In the previous sections, we have presented two classes of simple nouns and seven binominal constructions in Mojeño Trinitario and Harakmbut, following Pepper’s (2020) typology. Although it is the least frequent binominal type in Pepper’s (2020) cross-linguistic study, the **cls** type is not marginal in South America, unlike in other macro areas (2020: 170). More specifically, it is found in three of the ten South American languages of the sample, viz. in Mojeño Trinitario, Harakmbut and Murui Huitoto (Wojtylak 2016), all spoken in Western Amazonia. Krasnoukhova (2012: 209) also states that the derivational function of classifiers “is

²⁰ The derivational suffix *-rV*, although it can surface as *-re*, differs from *-re* NPOSD in that it applies to any noun class, while *-re* NPOSD only applies to a lexically determined subset of bound nouns when they are used without a possessive person prefix.

²¹ This is not true of Mojeño Trinitario unique classifiers, which have the same meaning as that of a lexical noun, from which they differ in form, see classifier *-pewo* ‘CLF:foot’ vs *-iype* ‘foot’.

very prominent among languages with a multifunctional classifier system” in South America. This type of binominal construction thus deserves closer examination, as it is cross-linguistically rare but at the same time areally pervasive..

The constructional approach to binominal lexemes adopted here is a fruitful way of pursuing this goal because it facilitates the comparison of classifier-derived nouns with N-N compounds. The previous discussion has indicated two reasons why comparing these two phenomena is key. One is that the semantic contribution of classifiers to binominals is not very different from that of nouns in N-N compounds, and the other is that classifiers take up the same position in N-CLF formations as N2 in N-N compounds.

Section 3 will investigate the similarities and differences between classifiers and bound nouns in greater detail. It will show that the distinction between classifier-derived nouns and N-N compounds is difficult to make in the two languages studied for different reasons: in Mojeño Trinitario because classifiers and bound nouns show a very similar distribution also beyond binominal constructions, and in Harakmbut because all classifiers are formally identical to bound nouns, some of which have a very general, abstract meaning.

3 Classifiers vs. bound nouns

In this section we distinguish classifiers from bound nouns. Section 3.1 presents the inventories of the two categories in Mojeño Trinitario and Harakmbut, and the criteria underlying them regarding form/meaning mapping.. In addition, the inventories result from meticulously comparing the behaviour of classifiers and bound nouns in larger morpho-syntactic units. Specifically, we will compare classifiers and bound nouns in terms of their phonological and prosodic integration in morphologically complex nouns (Section 3.2), their syntactic distribution within and beyond the noun phrase (Section 3.3), and finally their functions when occurring on nouns and in verb forms (Section 3.4). We posit that, in any classifier language, classifiers and bound nouns will differ along at least one of these dimensions of analysis.

3.1 Inventories, form and meaning

Distinguishing classifiers from bound noun roots is not straightforward. This section will spell out some considerations used in each language to identify elements as belonging to one or the other category.

Mojeño Trinitario has 31 classifier suffixes, listed in the appendix of Rose (2019a). Most of these are of a CV form and lack any obvious relationship to a noun (28). Some have a long history; *-pi* ‘CLF:filiform’ (28a), for instance, has been reconstructed for proto-Arawak, along with its lexical source **pi* ‘snake’ (Payne 1991b: 248) (note that (28b) is not the reflex of **pi*). Others show a formal and semantic resemblance to a noun (29). Finally, some have the same form as and a meaning related to that of a noun (30). When a classifier is formally related to or similar to a noun, its meaning is more abstract and general, often denoting shape or localization.

A subset of classifiers have allomorphs depending on whether they are stem-final (on nouns, numerals and stative verbs), or stem-internal in active verbs, such as *-mo* ~ *-me* ‘CLF:fabric’ in (31a) and (31b) respectively.

Trinitario

- (28) a. *-pi* ‘CLF:fili’ (for thin, long, flexible items)
b. *giore* ‘snake’
- (29) a. *-ju'e* ~ *-je* ‘CLF:interior’
b. *-ju'e* ‘stomach’
- (30) a. *-miro* ‘CLF:face’ (referring to either faces or places in front of a ground)
b. *-miro* ‘face’
- (31) a. *t-jitu-mo* [3-be_thick-CLF:fabric] ‘it is thick (of a fabric, for ex.)’
b. *s-oktáya-me-ko* [3f-step_on-CLF:fabric-ACT] ‘she is stepping on the blanket (for ex.)’

Bound nouns form a large sub-class of the Mojeño Trinitario nouns (32)–(33), with two morphological sub-classes (Rose 2020). Bound nouns of the first class obligatorily take a possessive prefix when used as the head of a noun phrase (32). Bound nouns of the second class generally take a possessive prefix when used as the head of a noun phrase (33a), except when they carry a derivational suffix for non-possession, like *-ti* in (33b) or *-re* in (4b).²² Semantically, these nouns denote parts of wholes and kinship relationships, as well as items of clothing and personal accessories, some bodily excretions, personal attributes and a few artefacts. Nominal roots are at least (underlyingly) disyllabic in Mojeño Trinitario. They do not show allomorphy, although their form with or without a prefix can have a different surface realization, due to rhythmic syncope (33) (see Section 3.2).

Trinitario

- (32) a. *n-amri* [1SG-grandchild] ‘my grandchild’
b. **amri* ‘grandchild’
- (33) a. *n-yowo* [1SG-axe] ‘my axe’
b. *ywo-ti* [axe-NPOSD] ‘axe’

To distinguish classifiers from bound roots in a consistent way, we have used the following methodology for Mojeño Trinitario. If a ‘suspect’ item used in word formation is formally and/or semantically distinct from a similar element that is used as the head of a noun phrase, it is considered a classifier. This is the case for the classifier *-mu'i* (~ *-m'i*) used to refer to various aspects of the environment (e.g. time, looks) (34a), the underlying form of which differs from that of its lexical nominal source *-imu'i* (~ *-im'i*) ‘physical property’ (34b). This means that, conversely, if the suspect element shows neither a formal nor a semantic distinction between its use in word formation and as the head of a noun phrase, it is considered a noun. This is the case with *-chupu* (~ *-chpu*) ‘trunk’ in (35a) and (35b)).

²² The distribution of these suffixes is lexically determined.

Trinitario

- (34) a. *to n-ijare-m'i*
 ART.NH 1SG-name-CLF:setting
 'my birthday'
- b. *to ta-em'i ma 'chane (taem'i < ta+ imu'i)*
 ART.NH 3NH-physical_property ART.M²³ person
 'the shape of a man'
- (35) a. *to manka-chpu*
 ART.NH mango-trunk
 'the mango tree trunk'
- b. *to ta-chupu (to) manka*
 ART.NH 3NH-trunk ART.NH mango
 'the trunk of a mango tree'

Harakmbut has a much smaller inventory of classifiers, about 13, all of which are formally identical to a bound noun. They are monosyllabic, mostly of a CV form, which is not distinctive of bound nouns nor classifiers, as there are a few independent nouns with the same syllabic structure, e.g. *ho* 'peach palm'. Examples are given in (36) to (43), with the (a)-examples representing the classifiers, and the (b)-examples their formally identical bound nouns.

Harakmbut

- | | |
|--|---|
| (36) a. <i>-mba?</i> CLF:hand;leaf | b. <i>-mba?</i> 'hand', 'leaf'; 'hand/leaf-shape' |
| (37) a. <i>-pe</i> CLF:disk | b. <i>-pe</i> 'jaw, chin, cheek'; 'sth disk-like' |
| (38) a. <i>-pa</i> CLF:rod | b. <i>-pa</i> 'penis'; 'rod' |
| (39) a. <i>-pu?</i> CLF:cylindrical;hollow | b. <i>-pu?</i> 'bamboo'; 'tube' |
| (40) a. <i>-nda</i> CLF:fruit | b. <i>-nda</i> 'fruit'; 'fruit shape (e.g. grapefruit)' |
| (41) a. <i>-po</i> CLF:sphere | b. <i>-po</i> 'something round' |
| (42) a. <i>-pi</i> CLF:stick | b. <i>-pi</i> 'something stick-like' |
| (43) a. <i>-wẽ</i> CLF:liquid | b. <i>-wẽ</i> 'river; liquid' |

This formal identity of classifiers and bound nouns bears heavily on the analytical problem of distinguishing between N-N compounds and N-CLF formations in Harakmbut. As items like (36) to (43) show two different types of semantic extensions in different syntactic environments, we analysed the same form as instantiating two different morphological categories, classifier and noun. For a number of items (about half), e.g. in (36) to (39), the bound noun has a more specific

²³ Note that the masculine singular human article shows two forms, depending on the gender of the speaker (i.e. genderlects, see Rose 2013). The form *ni* in examples (22), (23), (53) and (58b) is uttered by a woman, whereas the form *ma* in example (34b) is uttered by a man (Rose 2013).

meaning referring to a body or plant part in addition to a more abstract meaning referring to a shape identical to their classifier counterparts. Whenever these items in binominals involve their more specific meaning, they are analysed as bound nouns. For the other half, including (40) to (43), however, there is no difference in meaning between the two categories. The meaning of the bound nouns is somewhat atypical for nouns, as they refer in a general way to entities with the shape or substance denoted by the classifier. Items like (40) to (43) and items like (36) to (39) when carrying an abstract meaning are attributed categorial status depending on their syntactic distribution and function in comparison to other – less ambiguous – items of the sets of classifiers and bound nouns. However, as these two categories share the function of word formation when occurring on nouns (see Section 3.4), even this paradigmatic approach cannot guarantee the correct analysis. Consider the difference between (44) and (45), for example.

Harakmbut

(44) *wã-õh-wẽ* [NPF-nose-liquid] ‘nostril’

(45) *kumo-k-wẽ* [barbasco-SPAT:separation-(CLF:)liquid] ‘barbasco juice’

In (44), the binominal can be analysed as an exocentric compound, as ‘nostril’ refers to neither a type of nose, nor a type of liquid, but to the place where liquid (mucus) leaves the nose. As the N-CLF formations in Harakmbut are never semantically exocentric, unlike N-N compounds, *-wẽ* is analysed as a bound noun in (44). In (45), by contrast, either analysis for *-wẽ* is acceptable, as spatial linkers are found between two nominal roots in **gen** binominals (like (9)) as well as between a nominal root and a classifier in **cls** binominals (see Pepper 2020: 479).

Bound nouns form a large sub-class of the Harakmbut nouns, referring to inalienably possessed entities, such as body parts, plant parts, and landscape parts, as well as basic shapes or qualities of entities (Van linden 2021). They also include kinship terms. The 73 items identified as “shape morphemes” by Hart (1963), 45 to 50 of which Payne (1987: 36) analyses as classifiers, are all bound nouns but they do not exhaust the class (kinship terms, for instance, are not included by either author). However, in our analysis, only a small subset of these also function as classifiers, specifically verb classifiers (see Section 3.4).

We now move on to comparing N-N compounds and N-CLF derived nouns, or bound nouns and classifiers, at different levels of analysis. Phonology and prosody do not provide clear criteria for distinguishing bound nouns and classifiers for the two languages examined here (Section 3.2), but the study of the syntactic distribution of classifiers and bound nouns (Section 3.3) as well as their functions in some of these syntactic environments (Section 3.4) does .

3.2 Phonology and prosody

Phonology and prosody are not instrumental in distinguishing classifier-derived nouns from nominal compounds in the two languages under study.

In Mojeño Trinitario, classifiers and bound nouns in compounds behave identically as part of the word for phonotactics and rhythmic syncope (Rose 2019b). Phonotactics in Mojeño Trinitario resolves hiatus at morpheme boundaries within the word by several processes (diphthongization,

or deletion of one of the vowels, accompanied by palatalization or labio-velarization of the preceding consonant in some environments). These processes apply in all words to solve hiatus, including between a noun and its classifier (46) as well as between the two nominal roots of a compound noun (47).

Trinitario

- (46) a. *vtseramo* vi-tsera+omo [1pl-tear-CLF:liquid] ‘our tears’ N-CLF
 b. *sawariomo* saware+omo [tobacco-CLF:liquid] ‘tobacco juice’ N-CLF
 (47) a. *kwoyichko* kVwoyu+ichVko [horse-excrement] ‘horse excrement’ N-N
 b. *wakaechkopa* waka+ichVko-pa [cow-excrement-CLF:mass] ‘cow dung’ N-N

Metrical parsing in Mojeño Trinitario is iambic for most word classes, and applies iteratively from left to right, with stress falling on the final foot of the word (Rose 2019b). A rather pervasive process of rhythmic syncope makes each vowel in a weak metrical position subject to deletion (except the final vowel, which is always maintained). This process applies in all words, and thus forms evidence that classifier-derived nouns (48) and binominal compounds (49) are single words. Stressed syllables have been underlined, and the syncopated vowels are in bold in the underlying representation.

Trinitario

- (48) a. *sponji* sVponi+ji [corn-CLF:amorph] ‘corn field’ N-CLF
 b. *Trinram'i* Trinra+mu'i [Trinidad-CLF:setting] ‘the festival of Trinidad’ N-CLF
 (49) a. *swotonepgi* sV-wotone+pi [3f-button-ankle] ‘malleolus (outer ankle)’ N-N
 b. *kwoysumu* kVwoyu+sumu [horse-snout] ‘mounting ox’ N-N

In Harakmbut, classifiers and bound nouns also behave identically as part of the prosodic word for stress placement. (There are no specific phonotactic processes to be mentioned here, and Harakmbut does not show rhythmic syncope.) Examples are in (50) and (51), in which the stressed syllables have been underlined.

Harakmbut

- (50) a. *siro* [metal] ‘metal; machete’
 b. *siro-pi* [metal-CLF:stick] ‘knife’ (cf. Hart 1963: 1) N-CLF
 (51) a. *tare* [manioc] ‘manioc’
 b. *tare-mba?* [manioc-hand;leaf] ‘manioc leaf’ N-N

While in the morphologically simple nouns in (50a) and (51a), the stress falls on the first syllable, in the complex nouns (50b) and (51b) the stress falls on the last syllable of these disyllabic noun roots. Examples (50) and (51) provide evidence that both compounds and nouns derived with a classifier form single prosodic words, with the main stress falling on the penultimate syllable (cf. Van linden 2022).

3.3 Syntax

The syntactic distribution of classifiers and bound nouns in Mojeño Trinitario and Harakmbut is compared in Table 3. In both languages, bound nouns are found in at least the same syntactic environments as classifiers (however, sometimes with distinct functions, as laid out in Section 3.4). The crucial difference in their distribution is that classifiers are not accepted as the head of a noun phrase. This is in line with the idea that “bound roots can be distinguished from affixes only by virtue of also occurring as free forms” (Lieber and Štekauer 2009: 5). The distribution of classifiers in Harakmbut is much more restricted than has been claimed in earlier work, which was typically not based on primary data but on Hart’s (1963) paper about “shape morphemes”.²⁴ It is certainly more restricted than in Mojeño Trinitario, whose system clearly corresponds to what Aikhenvald (2000) labels a ‘multiple classifier system’.

Syntactic environment	Mojeño Trinitario		Harakmbut	
	CLF	bound N	CLF	bound N
as NP head	✗	✓	✗	✓
on nouns	✓	✓	✓	✓
on numerals	✓	✓	✗	✓
on adjectives	✓	✓	✓	✓
in verbs	✓	✓	✓	✓

Table 3: The syntactic distribution of classifiers and bound nouns in Mojeño Trinitario and Harakmbut

In Mojeño Trinitario, classifiers cannot be used as the head of a noun phrase (52). For instance, the classifier *-pi*, which prototypically classifies ‘filiform’ objects like ropes, belts, or snakes, can be defined as classifying “long, thin, flexible items” (25a). It cannot refer by itself to a long, thin and flexible item (52). By contrast, bound nouns can be used as the head of a noun phrase, like *yeno* in (53), also illustrated in a compound in (18).

Trinitario

(52) **to* *ta-pi* (*to*) *yuku* *ART CLF
 ART.NH 3NH-CLF:fili ART.NH fire
 ‘a long, thin and flexible piece of fire’

²⁴ For example, Derbyshire and Payne (1990: 246, 260) mistakenly state that the classifier system in Amarakaeri/Harambut is a primarily verb-incorporated system that has developed non-gender concordial functions, and that classifiers have a nominalizing function (1990: 267). Aikhenvald (2000: 123), in turn, wrongly mentions languages of the “Harakmbet” family as examples of languages with large sets of numeral classifiers (without giving a reference).

(53) *su ñi-yeno ñi kasiki* ART N
 ART.F 3M-wife ART.M cacique
 ‘the wife of the cacique’

The situation in Harakmbut is different due to the formal identity of classifiers and bound nouns. The form *-pa* in (54a) can be analysed as a classifier or as a bound noun occurring on a noun; in (54b) it is a noun and can hence function as the head of a noun phrase in a clause. It thus behaves in the same way as the unequivocal bound noun *-ayʔ* in (55). That is, on nouns the difference between classifiers and bound nouns is almost impossible to make in Harakmbut (see also Section 3.1).

Harakmbut

- | | | | |
|------|----|---|--------------|
| (54) | a. | <i>hak-pa</i> [house-(CLF:)rod] ‘rafter’ (cf. Hart 1963: 3) | N-CLF /= N-N |
| | b. | <i>wa-pa</i> [NPF-penis;rod] ‘penis; rod’ | N |
| (55) | a. | <i>wa-kuʔ-ayʔ</i> [NPF-head-bone] ‘skull’ | N-N |
| | b. | <i>wa-ayʔ</i> [NPF-bone] ‘bone’ | N |

In both languages, bound nouns can attach to the right of numerals (56a) and (57a), but only Mojeño Trinitario numerals can take classifiers (56b). In the examples, the stressed syllables have been underlined to support the identification of prosodic words.

Trinitario

- (56) a. *api-pgienu* api+**pgienu** [two-neck] ‘two necks’ NUM-N
 b. *no api-na-no* (*‘chañ-ono*) NUM-CLF
 ART.PL two-CLF:human-PL person-PL
 ‘two persons’

Harakmbut

- | | | | | |
|------|----|--|--|-------|
| (57) | a. | <i>ĩh-tõ-ẽ-ỹ</i>
1SG.IND-SOC-be-1.IND
'I have two hands' | <i>mbotta?</i> -mba?
two- hand | NUM-N |
| | b. | <i>ĩh-tõ-ẽ-ỹ</i>
1SG.IND-SOC-be-1.IND
'I have two hands' | <i>mbotta?</i> <i>wa-mba?</i>
two NPF- hand | NUM N |

Numerals in Mojeño Trinitario are bound roots: they obligatorily combine with a classifier, an independent or bound noun or a multiplicative. Classifiers are found much more frequently on numerals than bound nouns; and while in elicitation, all classifiers can be affixed to numerals, in the texts collected in the field the classifier *-na* for humans is almost exclusively found,

Adjectives do not require the presence of classifiers or bound nouns in either language, as exemplified in (58a) and (59). They nevertheless can combine with a nominal root or a classifier to form a single prosodic word in both languages. The rare Mojeño Trinitario examples all involve bound nouns, which lose their possessive prefix in this combination (58b). Classifiers also are (rather rarely) found on adjectives, either for ‘agreement’ (58c) or for word-class changing derivation resulting in a lexical nominalization (58d).

(58) a. *to* *'chope* *smeno* ADJ N
ART.NH big forest
'the big forest'

b. *ñi* *'moperu* *'chope-**chuti*** *t-kowo* *te* *tajunorokku* ADJ-N=ADJ
ART.M youngster big-**head** 3-bathe PREP creek
'the big-headed boy bathes in the creek' (Ibáñez Noza et al. 2007: 181)

c. *to* *'chope-**gie*** *wkugi* ADJ-CLF=ADJ
ART.NH big-CLF:**cyl** tree
'the big trunk'

d. *to* *'chope-**'e*** ADJ-CLF=N
ART.NH big-CLF:**convex**
'a drum'

²⁵ Numeral constructions with independent nouns show only one structural type, i.e. that with numeral and head noun forming two distinct words (cf. Van linden 2022).

Harakmbut

- (59) a. *kuwa* *uru-nda* N ADJ
 dog beautiful-NDA
 ‘a/the beautiful dog’
- b. *mbi?igŋ* *wa-mboro* N ADJ
 fish WA-big
 ‘a/the/ some big fish’

(60) Harakmbut

- a. *a-yok-i* *sal* *uru-wettone-ta-nda* ADJ-N=N
 1SG.IMP-give-1.IMP salt beautiful-woman-ACC-NDA
 ‘I (should) give salt to the beautiful woman.’
- b. *ndo?* *ih-ẽ-ỹ* *mboro-?i-nda* - ADJ-N=ADJ
 1SG 1SG.IND-be-1.IND big-foot-NDA²⁶
 ‘I am big-footed’
- c. *ih-tõ-ẽ-ỹ* *wa-?i* *mboro-nda* N ADJ=NP
 1SG.IND-SOC-be-1.IND NPF-foot big-NDA
 ‘I have big feet.’
- d. *mboro-po-nda* [big-CLF:sphere-NDA] ‘fat, big’ ADJ-CLF=ADJ

In Harakmbut, adjectives combine with both independent (60a) and bound nouns (60b), but the result is quite different. Combinations with independent nouns, like *wettone* ‘woman’ in (60a), result in a nominal word form.²⁷ Combinations with bound nouns, by contrast, result in an adjectival word form, like ‘big-footed’ in (60b). Crucially, bound nouns that function as a nominal head, as in (60c), cannot fuse with the adjective root. Very rarely, classifiers are also found on adjectives (60d), which yields complex adjectives, similarly to the pattern with bound nouns illustrated in (60b). That is, unlike in Mojeño Trinitario, classifiers on adjectives are not used for agreement or for deadjectival nominalization.

Finally, in both languages bound nouns and classifiers are incorporated in verb forms. In Mojeño Trinitario, they both attach either to stative verbs or to active verbs immediately after the root, as in (61a) and (61b). Classifiers are associated with nominal expressions of different grammatical roles: the unique argument of intransitive verbs, the patient of transitive verbs, or obliques as in (61a) (Rose 2019a). Incorporated bound nouns also correspond – in the counterpart clause without incorporation – to the unique argument of an intransitive root, to an oblique, or to the patient argument of a transitive root (61b). In Harakmbut, classifiers (62a) and bound nouns

²⁶ The analysis of the suffix *-nda* remains unclear; this is why the gloss just repeats the form itself.

²⁷ Note also that adjectives do not obligatorily fuse with independent nouns. The distribution of constructions might relate to the referential properties of the NP (see Van linden 2022). More research is needed here.

(62b) also share the same slot in the morphological template of the verb, but they precede the verb root in that language. Both types of incorporated elements are associated with the unique argument of an intransitive root, as in (62a), or the patient argument of a transitive root, as in (62b).

- (61) TRN a. *n-semo-pi-ko* [1SG-be_angry-CLF:**fili**-ACT] ‘I am angry at these words.’
 b. *t-vi-o'i-ri-ko* [3-take_out-**fruit**-PLURACT-ACT] ‘(s)he collects fruits.’
 (62) AMR a. *o-poʔ-sak-on* [3SG.IND-CLF:**sphere**-break-PFV.NVOL] ‘it (the pot) has broken.’
 b. *ih-mbaʔ-teḡ-me-y* [1SG-**hand**-cut-REC.PST-1.IND] ‘I cut my hand.’

In conclusion, classifiers and bound nouns have a very similar syntactic distribution in Mojeño Trinitario, with just one environment restricted to nouns, i.e. that of head of a noun phrase. In Harakmbut, classifiers share fewer syntactic environments with bound nouns. Both categories are frequently found on nouns and incorporated in verbs in the two languages under study, which is why we turn to the functions they have in these environments in the next section.

3.4 Functions

In both languages studied here, classifiers and bound nouns share the syntactic environments of occurring on nouns (Section 3.4.1) and being incorporated in verb forms (Section 3.4.2). Focusing on how their functions in these environments differ we will contribute to the debate on whether classifiers constitute thing-affixes, a question we link up with headedness.

3.4.1 Functions on nouns

Table 4 compares the functions of classifiers and bound nouns when appended to nouns in Mojeño Trinitario and Harakmbut. While both can be used to create new lexical items, classifiers have an additional function of qualification (or property-assignment), especially in Mojeño Trinitario. This is why Mojeño Trinitario classifiers can be described as multifunctional classifiers (Krasnoukhova 2012): the same set is used for qualification, derivation (and some agreement).

Functions on N	Mojeño Trinitario		Harakmbut	
	CLF	bound N	CLF	bound N
word formation	✓	✓	✓	✓
qualification	✓	✗	(✓)	✗

Table 4: The functions of classifiers and bound nouns on nouns
in Mojeño Trinitario and Harakmbut

The literature on classifiers on nouns does not usually tease apart different functions of classifiers on nouns. Instead, the functions of classifiers on nouns are either vaguely referred to

collectively as “classification/ categorization” (Seifart 2010: 725) or subsumed under the cover term “derivation” (Payne 2007, Wojtylak 2016, Aikhenvald 2000). Only a few authors briefly mention distinct functions of classifiers on nouns. For instance, Contini-Morava and Kilarski (2013: 268-269) distinguish between two major semantic functions of nominal classification: “The first, ‘expansion of the lexicon’, involves the use of nominal classification markers to create nouns. [...] The second type [is] ‘differentiating referents’”. In the same vein, Brandão (2016: 279) mentions that in Paresi, classifiers on nouns either derive new nouns or not. Interestingly, Pepper (2020: 148-154) also explicitly distinguishes two functions of classifiers on nouns and only includes the derivational use of classifiers on nouns in his typology of binominals, i.e. structures where the denotatum of the binominal is different from that of the base (see Section 2). In the present paper, the following semantic criteria are used to distinguish the two major functions of classifiers: if N-CLF designates an instance of the type denoted by the nominal root, we are dealing with qualification; if N-CLF denotes an entity (or type) different from that denoted by the nominal root, we are dealing with derivation, subsumed under word formation in Table 4, so as to include the process of compounding observed for bound nouns.

Mojeño Trinitario examples of classifiers on nouns are presented in (63)-(64). When classifiers are used on nominal roots for derivation, they serve to derive a complex concept in a generally semantically transparent way (63a). However, they can also be used on nominal roots to categorize the referent of the root, in terms of shape or material (64).

Trinitario

- | | | | |
|------|----|---|---------------|
| (63) | a. | <i>yuk-pi</i> [fire-CLF:fili] ‘candle’ | derivation |
| | b. | <i>gióre-pi</i> [snake-CLF:fili] ‘worm’ | derivation |
| (64) | a. | <i>mári-si</i> [stone-CLF:sphere] ‘round stone’ | qualification |
| | b. | <i>tsera-(o)mo</i> [tear-CLF:liquid] ‘tear’ | qualification |

Example (63) shows the use of the classifier *-pi* ‘CLF:filiiform’ for derivation. In (63a), ‘candle’ is not a type of fire: the classifier-derived word refers to some entity distinct from that referred to by the nominal root. In (63b), the derivation is not compositional, as the referent of the root (‘snake’) already belongs to the shape class expressed by the classifier: the resultant meaning is lexicalized. Example (64) shows the use of classifiers for qualification. In (64a), the use of *-si* on *mári* ‘stone’ does not change the meaning of the word but highlights physical characteristics of the referent (shape/material). The use for qualification is less crucial to the overall meaning of N-CLF; it often seems redundant, as in (64b). Senft discusses this use as follows: “The classifier that refers to a nominal referent may [...] highlight a special (shade of) meaning which then extracts one special referent out of the sum of possible extralinguistic referents the noun can refer to if it is not specified by this classifier” (Senft 2000: 36).

It should be noted that some Mojeño Trinitario roots are obligatorily classified: they cannot stand by themselves without a classifier, such as **mopo*, **wayo*, and **tére* shown in (65d) and (66b, e).

Trinitario

- (65) a. *mópo-si* [bee_related-CLF:sphere] ‘bee’ derivation
b. *mop-ji* [bee_related-CLF:amorph] ‘beeswax’ derivation
c. *mop-omo* [bee_related-CLF:liquid] ‘honey’ derivation
d. **mopo*
- (66) a. *wáyo-si* [deer_fly-CLF:sphere] ‘deer fly’ qualification
b. **wayo*
c. *tére-pi* [belt-CLF:fili] ‘belt’ qualification
d. *tére-mo* [belt-CLF:fabric] ‘woven belt’ qualification
e. **tére*

Yet, the same criteria apply for distinguishing between derivation and qualification; N-CLF in (65) yields lexemes with meanings distinct from that of the root, while N-CLF in (66) refers to an element of the type expressed by the root itself. Note that in the case of **mopo*, which can be combined with distinct classifiers (65) (unlike **wayo*), the gloss used is general enough to accommodate the different derivatives.²⁸ In contrast with these two uses of Mojeño Trinitario classifiers on nominal roots, bound nouns in N-N compounds are only used for word formation (see Section 2.5).

In Harakmbut, classifiers are generally used on nouns for derivation (67a), but there is a marginal use of classifiers on proper names that arguably rates as qualification. Specifically, the second author noted that members of the family hosting her referred to her as *Anpi* as in (67b).

Harakmbut

- (67) a. *wã-ðh-pi* [NPF-nose-CLF:stick] ‘beak (of a bird)’ (cf. Hart 1963: 2)
b. *An-pi* [An-CLF:stick] ‘An, who is slender (or stick-shaped) (and whom we hold dear)’

It turned out that uses like (67b) are only acceptable in contexts of knowing the person well and having a good relationship with them. They highlight the overall physical appearance of that person, and function as terms of endearment at the same time. Insofar as the criterion proposed above to distinguish between the two functions of classifiers on nouns concerns denotation only (and not connotation as well), this classifier use on proper names instantiates qualification. While this use is highly socially constrained, the derivational use of classifiers, illustrated in (67a) with the same classifier *-pi*, is not; it constitutes a productive process of word formation. Bound nouns in compounds, in turn, are only used for word formation (see Section 2.5).

The Mojeño Trinitario data allow us to look into the notion of headedness in N-CLF formations. The examples in (68) suggest that the specific function of the classifier is crucial in that domain.

²⁸ In cases of obligatorily classified roots like **mopo* ‘bee-related’, it could be questioned to what extent the nominal root still is a thing-root, as it does not really denote a physical object (Haspelmath 2012: 115), but rather a semantic domain.

Trinitario

- (68) a. *to wayo-si* [ART.NH deer_fly_{NH}-CLF:sphere] ‘the deer fly’ qualification
 b. *su choka-si* [ART.F blond_haired_F-CLF:sphere] ‘the blond woman’ qualification
 c. *to Peru-pa’i* [ART.NH Pedro_M-CLF:ground] ‘Pedro’s land’ derivation

In (68a) and (68b), in which the classifier *-si* is used for qualification, it is the nominal root that functions as the morphological head, as it triggers the appropriate gender value on the article, i.e. non-human for (68a) vs. feminine for (68b). The classifier *-si* for spherical items only highlights some aspect of the referent, i.e. by classifying it as an insect in (68a) or by focusing on its most relevant body part, the head, in (68b). In (68c), in which the classifier *-pa’i* is used for derivation, by contrast, the classifier is the morphological head as it determines the gender value of the article (if the nominal root were the head, the article would show masculine gender, agreeing with *Pedro*). Note that this morphological headedness does not imply that classifiers can function as the syntactic head of NPs; in (68c), the classifier *-pa’i* is the morphological head of the word *Perupa’i*, and it is the complex word form *Peru-pa’i* that is the head of the NP, not just the classifier *-pa’i*.

This formal evidence of headedness in Mojeño Trinitario also informs the discussion of whether classifiers rate as thing-affixes. When used for derivation, we tend to follow Pepper’s (2020: 10-12) onomasiological approach (after Štekauer 2000) and analyse classifiers as heads – and hence thing-affixes – in N-CLF formations, with the nominal root as semantically subordinate, which squares with the paraphrases we provided for classifier-derived nouns above and in Section 2.7. This analysis meshes well with the nominalizing function of classifiers on adjectives, as in (58d), and verbs, as in (69) (compare with (65c)).

- (69) TRN *to t-ijr-omo* [ART.NH 3-be_hot-CLF:liquid] ‘breakfast/dinner, lit. hot liquid’

When used for qualification, by contrast, we argue that classifiers do not function as heads and are not thing-affixes. They clearly have a modifying function (cf. Mithun 1986), and we propose to term them “property-affixes”, as they assign a (temporary or inherent) property to the referent of the nominal root, or to the type denoted by the root, characterizing it in terms of shape, quality or substance, in a way that is formally different from the prototypical property-roots in these languages, i.e. adjectives. This analysis fits with the agreement function of classifiers on modifiers (58c) and the classifying function of classifiers in verbs (see next section). We can thus conclude that the categorial status of classifiers as thing-affixes depends on the function they have on nominal roots.

The status of the classifier as semantic and morphological head in N-CLF words where the classifier plays a derivational role does not prevent the nominal root (semantically a modifier) to remain referential. This is illustrated in (70), where the nominalization *to nnujre* ‘(the thing) that I chewed’ modifies the nominal root *saware* ‘tobacco’, which is in turn the modifier part of the classifier-derived word *sawariomo* ‘tobacco juice’ (it is obviously not the juice that was chewed, as the juice precisely results from the process of chewing leaves).

Trinitario

- (70) *n-es-cho* *to* *sawari-omo,* *éto-na*
 1SG-give_drink-ACT ART.NH tobacco-CLF:liquid one-CLF:GEN
kchara to sawari-omo to n-nu-j-re
 spoon ART.NH tobacco-CLF:liquid ART.NH 1SG-chew-CLF:amorph-SP.P.NZ
 ‘I gave her the tobacco (juice), one spoon of tobacco juice that I had chewed.’

3.4.2 Functions in verb forms

Moving on to the syntactic environment of incorporation in verbs, Table 5 shows that the uses of classifiers and bound nouns are not functionally equivalent. That is, classifiers and bound nouns do not participate in the same functional types of noun incorporation as defined by Mithun (1984). In both Mojeño Trinitario and Harakmbut, classificatory noun incorporation (Type IV) is exclusively found with classifiers, while lexical compounding (Type I) is restricted to incorporation of nouns. Type II and Type III incorporation show variation between the languages studied. This section discusses these four types of noun incorporation, and their availability in the two languages examined.

Types of noun incorporation (Mithun 1984)	Mojeño Trinitario		Harakmbut	
	CLF	bound N	CLF	bound N
Type I: lexical compounding	✗	✓	✗	✓
Type II: manipulation of case	✓	✓	✗	✓
Type III: backgrounding in discourse	✓	✗	✓	✓
Type IV: classifying with ‘coreferential’ NP	✓	✗	✓	✗

Table 5: The functions of classifiers and bound nouns in verbs

In both languages, Type I noun incorporation, which serves to create new lexemes for “name-worthy” activities (Mithun 1984: 848) and derives intransitive predicates from transitive ones, is found exclusively with nouns (71)-(72).

- (71) TRN *t-vi-o'i-ri-ko* [3-take_out-**fruit**-PLURACT-ACT] ‘(S)he collects fruits.’ Type I with N
 (72) AMR *oʔ-ndagɨ-ka* [3SG.IND-**path**-make] ‘(S)he is making a path.’ Type I with N

In (71), the transitive verb stem *-vi* ‘take out’ is combined with the bound noun *-o'i* ‘fruit’ to yield an intransitive verb that denotes an “institutionalized” activity (Mithun 1984: 849), i.e. fruit-picking. The incorporated noun bears the semantic relationship of patient to its host verb. In (72), the bound noun *-ndagɨ* ‘path’ is incorporated into the transitive verb stem *-ka* ‘make;do’ to form the intransitive verb ‘path-make’.

Type II noun incorporation, which affects the valency structure of the whole clause (Mithun 1984: 856), is found with both classifiers (73) and bound nouns (74) in Mojeño Trinitario, while it is restricted to the latter category in Harakmbut (75).

Trinitario

- (73) *n-eja-j-ko* *to* *tyuraji*
 1SG-sit-CLF:**amorph**-ACT ART.NH mud
 ‘I am (heap)-sitting in the mud.’ Type II with CLF

- (74) *na-ech-kute-cho-po* *eto* *povre* *sorare*
 3PL-cut-**hindleg**-ACT-PFV 3NH poor animal
 ‘They cut off the hind leg of the poor animal.’ Type II with N

Harakmbut

- (75) *mbe-ku-ti-kot-uy-ne* *apoare?-a* *ta?mba-ya*
 3SG>1/2SG-**head**-SPAT:up-fall-DIST.PST-IND papaya-NOM swidden-LOC
 ‘A papaya fell on my head in the swidden long ago.’ Type II with N

The examples with nouns, (74) and (75), involve possessors being advanced to object status, which position is vacated by the incorporated body part (cf. Mithun 1984: 857-858). The construction with the classifier in (73) is somewhat different, and not discussed in Mithun (1984). Specifically, the classifier functions as an applicative marker (Rose 2019a), which promotes in (73) the locative argument of an intransitive verb to object position; the resulting construction is formally transitive.

Type III noun incorporation is used to background known or incidental participants in discourse (Mithun 1984: 859). It is restricted to classifiers in Mojeño Trinitario (78). In Harakmbut it is found with both classifiers (77) and bound nouns (78).

Trinitario

- (76) *p-eja-pue-gi-a* [2SG-sit-CLF:**ground**-ACT-IRR] ‘(Please) sit down (on the floor).’
 Type III with CLF

Harakmbut

- (77) *pera* *o-n-ka* *ãñĩ* *o-mbewik-po* *eskalera-te*,
 pear 3SG.IND-SPAT-do FILLER 3SG.IND-go.up-DEP ladder-LOC
ãñĩ *o-ma-nda-e-a* *ãñĩ* *kanasta-yo* [...]
 FILLER 3SG.IND-VPL-CLF:**fruit**-get-TRNS FILLER basket-LOC
 ‘He is picking pears, eh, going up on a ladder, eh, he is taking/collecting them (the fruits), eh, in a basket.’ (spontaneous speech) Type III with CLF

- (78) *apetpet-ʔidn ih-waway-me-y ndumba-yo. ken ndoʔ-edn*
 jaguar-tooth 1SG.IND-find-REC.PST-1.IND forest-LOC then 1SG-GEN
wa-mambuy-ta ih-ʔidn-yok-me-y
 NPF-same_sex_sibling-ACC 1SG.IND-**tooth**-give-REC.PST-1.IND
 ‘I found a jaguar’s tooth in the forest. Then I gave it (the tooth) to my sister.’

Type III with N

In (76), the incorporated classifier *-pue* introduces a non-topical participant (cf. the “absolute” function of classifiers in Grinevald and Seifart 2004) which is immediately retrievable from the context of the speech event (exophoric retrieval). In (77), the first clause contains a full NP introducing the topical argument *pera* ‘pear’; in the second one, anaphoric reference to the pears is realized by the classifier *-nda* for fruit (prototypically grapefruit). In (78), the bound noun *-ʔidn* ‘tooth’ likewise anaphorically refers to the full NP *apetpetʔidn* ‘jaguar’s tooth’ in the previous clause. It should be noted that the use of the incorporated items in (77)–(78) is not literally anaphoric, since they are non-referential, but they “retain the entity in question within the arena of discourse”, because “incorporated nouns, not salient constituents in themselves, do not obstruct the flow of information, yet their presence is sufficient to narrow the scope of the verb” (Mithun 1986: 381–382).

Finally, Type IV noun incorporation, also termed classificatory noun incorporation because the incorporated element classifies a more specific external NP present in the clause (Mithun 1984: 863), is restricted to classifiers in both languages. Examples are given in (79) and (80).

Trinitario

- (79) *t-eja-me-re-ko te pjo ñi-ye'e estera*
 3-sit-CLF:**fabric**-PLURACT-ACT PREP DEM 3M-GRN mat
 ‘He is (fabric)-sitting on a mat.’

Type IV with CLF

Harakmbut

- (80) *idn-pa-a i-ma-nda-kot-a-y palta*
 tooth-CLF:rod-INS 1SG.IND-VPL-**CLF:fruit**-fall-TRNS-1.IND avocado
 ‘I am making the avocados fall with a hook-shaped rod.’

Type IV with CLF

Example (79) contains the verb *-eja* ‘sit’ and the classifier *-me*, which specifies the shape of the locative argument (*estera*). (Note that (79) is syntactically different from (73) in that the locative argument is coded as a prepositional phrase, while in (73) it is coded as an object noun phrase. The classifier in (79) has no effect on the valency of the verb.) Similarly, in (80) *-nda* characterizes the O-argument of the verb in terms of shape, expressed by the external NP *palta*.

In both languages, classificatory noun incorporation constructions are typically used to introduce new topics, or re-activate aforementioned ones.

In conclusion, looking at the functional types of incorporated classifiers and bound nouns has enabled us to describe these two categories with greater precision. The two languages allow only bound nouns in Type I noun incorporation²⁹ (lexical compounding) and only classifiers in Type IV (classificatory noun incorporation). The data also corroborate two diachronic hypotheses proposed by Mithun: (i) Types I to IV form an implicational hierarchy for the development of noun incorporation (1984: 874), and (ii) classifiers originate in nouns (1986: 395). Type II for Mojeño Trinitario and Type III for Harakmbut form truly transitional stages, in which both bound nouns and classifiers are allowed.

The differential availability of classifiers and bound nouns for the different types of noun incorporation also bears on the question of whether classifiers are thing-affixes. Type I most clearly involves thing-morphs, analogously to N-N compounding, but is restricted to nouns in both languages.³⁰ Type IV, in turn, most clearly involves what we have called “property-affixes”, and is restricted to classifiers. Indeed, their use in Type IV is semantically analogous to their qualifying use on nouns. Taking into account these most clear-cut types, the data suggest that incorporated in verb forms, classifiers are more likely property-affixes than thing-affixes.

4 Conclusions and diachronic implications

This paper has focused on binominal lexemes in two Western Amazonian languages, Mojeño Trinitario and Harakmbut, and has tackled the analytical problem of distinguishing between classifier-derived nouns and noun-noun compounds, that is, the types **cls** and **cmp** in Pepper’s (2020) typology of binominals. This entailed scrutinizing the distinction between classifiers and bound nouns in the two languages, as these elements have a similar semantic effect on the noun root they combine with in the binominal types studied and occupy the same (rightmost) slot in them. This in turn required us to look far beyond complex nouns.

Methodologically, our analysis has benefitted from our comparative approach, systematically searching for paradigmatic oppositions both across and within the languages under investigation. While our phonological and prosodic analyses highlighted the similarity between the two binominal types studied (both forming single word forms, Section 3.2), looking at the form and meaning of ‘suspect’ items (Section 3.1) as well as comparing the syntactic distribution of classifiers and bound nouns (Section 3.3), and their functions on nouns and when incorporated in verb forms (Section 3.4) proved crucial in drawing the boundary between classifiers and bound

²⁹ There are a few exceptions here; the independent noun *hak* ‘house’ also occurs in Type I noun incorporation in Harakmbut.

³⁰ Note that nouns in Type I noun incorporation are analysed here as thing-morphs in spite of being non-referential (cf. Mithun 1984: 849). This is by analogy with noun-noun compounds, like *apple juice*, in which *apple* is non-referential as well, but still a thing-morph. In our discussion on the categorial status of classifiers, we thus abstract away from referentiality.

nouns. The criteria that emerged from our study as most important to this distinction are summarized in Table 6. We are confident that our methodology can be applied to other Western Amazonian languages as well, where derivational use of classifiers abounds (Section 2.8). In particular, it could be useful in clarifying the status of so-called “repeaters” (Grinevald Craig 2004: 1026), i.e. bound forms formally similar to some nouns and used with the same morphosyntactic distribution as classifiers. Repeaters are not prototypical classifiers in that, sharing their semantics with a single noun, they are too specific to be truly used for categorization. Instead of analysing repeaters as a special type of classifiers, we would consider them simply as (bound) nouns.

	Mojeño Trinitario	Harakmbut
CLF formally or semantically distinct from N	✓	✗
CLF have a qualifying function on N; bound nouns do not	✓	(✓)
CLF cannot function as NP head; bound nouns can	✓	✓
CLF have a categorizing function when incorporated in verbs; bound nouns do not	✓	✓

Table 6: Criteria to distinguish between classifiers and bound nouns in Mojeño Trinitario and Harakmbut

For Mojeño Trinitario, a multiple and multifunctional classifier language with an extensive set of classifiers that has almost the exact same syntactic distribution as bound nouns, the impossibility for classifiers to function as head of a noun phrase together with the formal distinctness of classifiers and bound nouns allowed us to distinguish between classifier-derived nouns and noun-noun compounds. Our findings thus challenge Gill’s (1957) claim that in Mojeño Trinitario some bound nouns are used as classifiers. They are also more precise than what Admiraal and Danielsen (2014) and Facundes and Freitas (2015) concluded for two other Arawak languages, Baure and Apuriña, i.e. that classifiers or classificatory nouns are a sub-type of bound nominal roots. For Harakmbut, a verbal classifier language with a small set of classifiers, the binominals studied turned out to be one of the few syntactic environments/functions for which classifiers and bound nouns are not in complementary distribution. As Harakmbut classifiers are formally identical to bound nouns, it proved at times impossible to distinguish between classifier-derived nouns and noun-noun compounds. Typically, items with very general semantics, merely referring to a shape, quality or substance, like *-pi* in (1a), were analysed as classifiers in binominals.

Our study has also reflected on the status of classifiers as thing-affixes. Not surprisingly, our functional approach suggests a nuanced answer. The Mojeño Trinitario data provided formal

evidence that in classifier-derived nouns classifiers are the morphological head, determining gender agreement on the article. In classifier-qualified nouns, by contrast, it is the noun that is the morphological head. Taking into account the semantic and formal similarity between classifier-derived nouns and noun-noun compounds (in which N2 is the head) in the two languages studied, this extra piece of formal evidence of headedness convinced us that, in their derivational use, classifiers are thing-affixes, which is in line with the onomasiological approach taken in Pepper (2020: 10-12), inspired in turn by Štekauer (2000). Additional evidence comes from nominalization in Mojeño Trinitario, where (unlike in Harakmbut) classifiers are used to derive nouns from verbs (69). However, in all other syntactic environments and functions studied, classifiers are more aptly analysed as what we labelled “property-affixes”. This holds especially for their qualifying use on nouns (Section 3.4.1) and their use in Type IV ‘noun’ incorporation (cf. Mithun 1984) (Section 3.4.2), functions that are excluded for bound nouns. Arguably, it also holds for their agreement uses on numerals and adjectives in Mojeño Trinitario, where they categorize the head noun rather than nominalize their host, while bound nouns on the same host types always yield nominals (Section 3.3). We can thus conclude that classifiers are prototypically property-morphs and only rate as thing-morphs when they share the same position and function as other prototypical thing-morphs, i.e. nouns, as in binominals.

Finally, we turn to the diachronic implications of our study. Our data corroborate the hypothesis that classifiers originate in nouns (Mithun 1986: 395; Aikhenvald 2000: 353-361), more specifically bound nouns in both languages. This development is in line with the general direction of grammaticalization from roots into affixes. In addition, our data suggest that bound nouns developed the types of noun incorporation I to IV proposed in Mithun (1984) in that same order (Section 3.4.2). The origin of classifiers is still visible in synchrony in Harakmbut, with pervasive polyfunctionality of the classifiers/bound nouns. We hypothesize that the classifiers merely referring to shapes or substances (like *-pi* ‘CLF:stick’ in (1a)) have gone further down the grammaticalization pathway than those whose corresponding bound noun still has a more specific meaning (like *-puʔ* ‘CLF:cylindrical;hollow’ in (1b)). The idea is that items like *-pi* lost their more specific meaning (semantic bleaching), with only the more schematic classifier meaning remaining. Harakmbut classifiers do not include any item (anymore) for which the more specific meaning is the only one available when used as a bound noun; that is, they have already undergone semantic generalization. Unfortunately, there is little material available for further diachronic or comparative work on this. In Mojeño Trinitario, the noun-to-classifier pathway is confirmed by cases of classifier/noun homonymy (e.g. *-miro* in (30)) and cases showing phonetic erosion and semantic bleaching of the classifier vis-à-vis its corresponding bound noun (e.g. *-ju'e ~-je* in (29)). For other cases, recourse must be made to reconstructions of the nominal sources of classifiers (e.g. *-pi* ‘CLF:fili’ (Section 3.1)). It remains to be investigated whether more nominal sources of classifiers can be reconstructed on the basis of comparison with sister languages. Comparing the two languages, we seem to be dealing with a different historical depth of the

emergence of classifiers, with Mojeño Trinitario showing an ‘older’ system including also some more recently integrated classifiers.

Our data also support the diachronic hypothesis that the development of nouns into classifiers took place in morphologically complex forms. Admiraal and Danielsen (2014: 90) and Croft (2017: 427) have pointed to noun-noun compounding as source constructions for classifiers, Mithun (1984, 1986) to noun incorporation, and Payne (2007: 472) and Seifart (2010: 729) to both. This paper suggests that the locus of change for the languages examined is the syntactic environments shared by classifiers and nouns, i.e. mainly noun-noun compounding and noun incorporation, but also combinations with numerals and adjectives. These latter two types of structures are generally underdescribed (Admiraal and Danielsen 2014 is an exception to this), and have therefore been overlooked as possible source environments for classifiers. We believe that these four morphologically complex structures instantiate a general template consisting of a host root and an element to its right. In the languages studied here, the elements attached to the different types of host are predominantly bound nouns and classifiers. This is why we conclude that classifiers developed from bound nouns in Mojeño Trinitario and Harakmbut in a general morphologically complex source template. Our data thus nicely fit Seifart’s (2010: 729) generalization that “classifiers of different types seem to diachronically ‘piggy back’ on existing constructions although the details of how this might generalize both within and across different classifier types remain still unclear.”

5 References

- Adelaar, Willem. 2000. Propuesta de un nuevo vínculo genético entre dos grupos lingüísticos indígenas de la Amazonía occidental: Harakmbut y Katukina. In Luis Miranda Esquerre (ed.), *Actas del I Congreso de Lenguas Indígenas de Sudamérica*, vol. 2, 219–236. Lima: U. Ricardo Palma.
- Adelaar, Willem. 2007. Ensayo de clasificación del katawixí dentro del conjunto harakmbut-katukina. In Andres Romero-Figueroa, Ana Fernández Garay, Angel Corbera Mori (eds.), *Lenguas indígenas de América del Sur: Estudios descriptivo-tipológicos y sus contribuciones para la lingüística teórica*, 159–169. Caracas: Universidad Católica Andrés Bello.
- Admiraal, Femmy & Danielsen, Swintha. 2014. Productive Compounding in Baure (Arawakan). In Swintha Danielsen, Katja Hannß & Fernando Zúñiga (eds.), *Word Formation in South American Languages*, 79–122. Amsterdam: Benjamins.
- Aikhenvald, Alexandra. 2000. *Classifiers. A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.
- Allan, Keith. 1977. “Classifiers.” *Language* 53 (2): 285–311.

- Bauer, Laurie. 2001. Compounding. In Martin Haspelmath, Ekkehard König, Wulf Oesterreicher & Wolfgang Raible (eds.), *Language Typology and Language Universals*, Vol. I, 695–707. Berlin: Walter de Gruyter.
- Brandão, Ana Paula. 2016. “A Incorporação de Nomes e Classificadores Em Paresi-Haliti (Aruák).” *Liames* 16 (2): 271–83.
- Contini-Morava, Ellen, and Marcin Kilarski. 2013. “Functions of Nominal Classification.” *Language Sciences* 40: 263–299.
- Crevels, Mily & Pieter Muysken. 2009. *Lenguas de Bolivia: presentación y antecedentes*. In Mily Crevels & Pieter Muysken (eds.), *Lenguas de Bolivia*, 13–26. La Paz: Plural Editores.
- Croft, William. 2017. “Classifier Constructions and Their Evolution: A Commentary on Kemmerer.” *Language, Cognition and Neuroscience* 32 (4): 425–27.
- Derbyshire, Desmond C. & Doris L. Payne. 1990. Noun classification systems of Amazonian languages. In Doris Payne (ed.), *Amazonian linguistics: Studies in lowland South American languages*, 243–271. Austin: University of Texas Press.
- Dixon, R.M.W. 1986. Noun Classes and Noun Classification in Typological Perspective. In Colette Craig (ed.), *Noun Classes and Categorization*, 105–112. Amsterdam & Philadelphia: Benjamins.
- Facundes, Sidney & Freitas, Marília Fernandez Pereira de. 2015. De compostos nominais produtivos a um sistema incipiente de classificação nominal em Apurinã (Aruák) [From productive noun compounding to AN incipient noun classification system in Apurinã (Arawak)]. *Moara* 43 (2). 23–50.
- Gill, Wayne. 1957. *Trinitario Grammar*. San Lorenzo de Mojos: Misión Nuevas Tribus.
- Grinevald, Colette & Frank Seifart. 2004. Noun classes in African and Amazonian languages: Towards a comparison. *Linguistic Typology* 8 (2). 243–285.
- Hart, Raymond. 1963. Semantic components of shape in Amarakaeri Grammar. *Anthropological Linguistics* 5 (9). 1–7.
- Haspelmath, Martin. 2012. How to compare major word-classes across the world’s languages. *UCLA Working Papers in Linguistics, Theories of Everything* 17, Article 16. 109–130.
- Helberg Chávez, Heinrich Albert. 1984. *Skizze einer Grammatik des Amarakaeri*. Tübingen: Tübingen University PhD dissertation.
- Helberg Chávez, Heinrich Albert. 1990. Análisis funcional del verbo amarakaeri. In Rodolfo Cerrón Palomino & Gustavo Solís Fonseca (eds.), *Temas de lingüística amerindia*, 227–249. Lima: Concytec.
- Ibáñez Noza, Eulogio, Pedro Fabricano Noe, Marcelo Guaji Noe, Claudio Guaji Jare, Bartola Guaji Jare, and Nemecio Yuco Parada. 2007. *Gramática Mojeña Trinitaria*. Tomo I. Trinidad, Beni: Centro Social y Comunitario “Ipeno Imutu”, Cabildo Indigenal de Trinidad.
- Krasnoukhova, Olga. 2012. *The Noun Phrase in the languages of South America*. Utrecht: LOT.

- Lieber, Rochelle & Štekauer, Pavol. 2009. Introduction: Status and Definition of Compounding. In Rochelle Lieber & Pavol Štekauer (eds.), *Oxford Handbook of Compounding*, 3–18. Oxford: Oxford University Press.
- Loukotka, Cestmír. 1968. *Classification of South American Indian Languages*. Los Angeles: Latin American Center, UCLA.
- Lyon, Patricia Jean. 1975. Dislocación tribal y clasificaciones lingüísticas en la zona del río Madre de Dios. In: *XXXIX Congreso Internacional de Americanistas, Lima 1970. Actas y Memorias, Vol. 5.*, 185–207. Lima: Instituto de Estudios Peruanos.
- Marbán, Pedro. 1702. *Arte de la lengua Moxa, con su Vocabulario, y Cathecismo*. Lima: Imprenta Real de Joseph de Contreras.
- Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60 (4). 847–879.
- Mithun, Marianne. 1986. The convergence of noun classification systems. In Colette Craig (ed.), *Noun Classes and Categorization*, 379–397. Amsterdam & Philadelphia: Benjamins.
- Moore, Denny. 2007. Endangered languages of lowland tropical South America. In Matthias Brenzinger (ed.), *Language diversity endangered*, 29–58. The Hague: Mouton.
- Olza Zubiri, Jesús, Conchita Nuni de Chapi & Juan Tube. 2002. *Gramática Moja Ignaciana*. Caracas: Universidad Católica Andres Bello.
- Passer, Matthias Benjamin. 2016. (What) Do verb classifiers classify? *Lingua* 174. 16–44.
- Payne, Doris L. 1987. Noun Classification in the Western Amazon. *Language Sciences* 9 (1). 21–44.
- Payne, David. 1991a. A classification of Maipuran (Arawakan) languages based on shared lexical retentions. In Desmond Derbyshire and Geoffrey Pullum (eds.), *Handbook of Amazonian Languages*, Vol. 3, 355–499. Berlin & New York: Mouton de Gruyter.
- Payne, David. 1991b. Clasificadores nominales: La interacción de la fonología, la gramática y el léxico en la investigación comparativa del Maipuran. *Revista Latinoamericana de Estudios Etnolingüísticos* 6. 241–257.
- Pepper, Steve. 2020. *The typology and semantics of binominal lexemes. Noun-noun compounds and their functional equivalents*. Oslo: University of Oslo PhD dissertation.
- Petersen de Piñeros, Gabriele. 2007. “Nominal Classification in Uitoto.” *International Journal of American Linguistics* 73 (4): 389–409.
- Rose, Françoise. 2013. “Los Generolectos Del Mojeño.” *Liames* 13: 115–34.
- Rose, Françoise. 2015. Mojeño Trinitario. In Mily Crevels & Pieter Muysken (eds.), *Lenguas de Bolivia*, vol 3, *Oriente*, 59–97. La Paz: Plural Editores.
- Rose, Françoise. 2019a. “From Classifiers to Applicatives in Mojeño Trinitario: A New Source for Applicative Markers.” *Linguistic Typology* 23 (3): 435–66. <https://doi.org/10.1515/lingty-2019-0024>
- Rose, Françoise. 2019b. Rhythmic syncope and opacity in Mojeño Trinitario. *Phonological Data and Analysis* 1 (2): 1–25.

- Rose, Françoise. 2020. "The possessive noun classes of Mojeño Trinitario." Presented at the *Atelier Morphosyntaxe*, laboratoire Dynamique du Langage, Lyon, France, October 20.
- Scalise, Sergio & Antonio Fábregas. 2010. The head in compounding. In Sergio Scalise & Irene Vogel (eds.), *Cross-disciplinary issues in compounding*, 109–126. Amsterdam: Benjamins.
- Seifart, Frank. 2005. *The structure and use of shape-based noun classes in Miraña (North West Amazon)*. Nijmegen: Max Planck Institute for Psycholinguistics PhD dissertation.
- Seifart, Frank & Payne, Doris. 2007. Nominal classification in the North West Amazon: Issues in areal diffusion and typological characteristics. *International Journal of American Linguistics* 73. 281–287.
- Seifart, Frank. 2010. Nominal Classification. *Language & Linguistics Compass* 4 (8). 719–736.
- Senft, Gunter. 2000. "What do we really know about nominal classification systems." In *Systems of Nominal Classification*, Cambridge University Press. Language Culture and Cognition Series 4.
- Štekauer, Pavol. 2000. Beheading the word? Please, stop the execution. *Folia Linguistica* 34 (3–4). 333–356.
- Tovar, Antonio. 1961. *Catálogo de las lenguas de América del Sur*. Buenos Aires: Editorial Sudamericana.
- Tripp, Robert. 1976. Los verbos Amarakaeri. Datos Etno-Lingüísticos: Colección de los archivos del ILV 33. Lima: Instituto Lingüístico de Verano.
- Tripp, Robert. 1995. *Diccionario amarakaeri-castellano*. Yarinacocha: Ministerio de Educación & Instituto Lingüístico de Verano.
- Van linden, An. 2021. "Where alienability accounts fall short: bound nouns in Harakmbut." Presented at the 54th Annual Meeting of the Societas Linguistica Europaea (SLE54), Online conference. <https://doi.org/10.17605/OSF.IO/UDVEW>.
- Van linden, An. 2022. Harakmbut. In Patience Epps and Lev Michael (eds.), *Amazonian Languages, An International Handbook. Language Isolates*, Volume 1: Aikanã to Kandozi-Shapra [Handbücher zur Sprach- und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science (HSK) 44/1], 437–477. Berlin: de Gruyter Mouton. <https://doi.org/10.1515/9783110419405-010>
- Wise, Mary R. 1999. Small language families and isolates in Peru. In R. M. W. Dixon & Alexandra Y. Aikhenvald (eds.), *The Amazonian languages*, 307–340. Cambridge: Cambridge University Press.
- Wojtylak, Katarzyna I. 2016. "Classifiers as Derivational Markers in Murui (Northwest Amazonia)." In *Word-Formation Across Languages*, edited by Lívia Körtvélyessy, Pavol Štekauer, and Salvador Valera, Cambridge Scholars Publishing. Newcastle-upon-Tyne.