

Van linden, An. 2023. 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], 441-481. Berlin: de Gruyter Mouton. (version: 12/11/2022)

<https://doi.org/10.1515/9783110419405-010>

Harakmbut

An Van linden
University of Liège & KU Leuven

Contact details:

University of Liège

Department of Modern Languages: Linguistics, Literature & Translation

Place Cockerill 3 - 5, 4000 Liège, Belgium

an.vanlinden@uliege.be

Table of Contents

1	Classification, demographics, and sociolinguistic background	2
2	Phonology and phonetics	4
3	Morphological profile and basic word classes	7
4	The noun phrase	7
4.1	Morphological template of the head	7
4.2	Pronouns and demonstratives	9
4.3	Common nouns	11
4.4	Modification by demonstratives and indefinites	12
4.5	Attributive possession	12
4.6	Nominal number and quantification	13
4.7	Descriptive modification	14
4.8	Word formation	15
5	The verb phrase	17
5.1	Mood	17
5.2	Argument marking	18
5.3	Tense, evidentiality and modality	20
5.4	Aspect	20
5.5	Associated motion	21
5.6	Verbal plural	21
5.7	Valency-changing mechanisms	22
5.8	Spatial prefixes	23
5.9	Noun incorporation	23

6	Simple clauses.....	24
6.1	Basic constituent order	24
6.2	Alignment system.....	25
6.3	Negation.....	25
7	Clause-linking.....	26
7.1	Relative relations	26
7.2	Complement relations.....	27
7.3	Adverbial relations	28
8	Conclusion	29
	Acknowledgments.....	30
	References	30
	Abbreviations	33

Abstract

This chapter presents a description of Harakmbut, an Amazonian language spoken in southeast Peru, based on existing work as well as original fieldwork. It focuses on its most vital dialect, Arakmbut (Amarakaeri). The discussion of its phonology and phonetics highlights nasality as an important—yet not fully understood—phenomenon. The chapter also presents morphological templates for both (pro)nominal heads and finite verb forms. The description of the noun phrase revolves around the distinction between obligatorily bound nouns and potentially free ones, which leads to distinct morphosyntactic behaviour in noun modification, noun incorporation, and word formation. Contra earlier work, I argue that just a limited number of bound nouns (rather than the whole class) should be analyzed as classifiers. The discussion of the verb phrase homes in on the lack of referential transparency in person marking, as well as the abundance of inflectional and derivational morphology, including markers of associated motion and temporal adverbial markers. In the system of argument marking on dependents, the three argument roles (S, A, and O) show differential and/or optional marking. At the level of clause-linking, nominalization plays an important role in the expression of relative, complement, and adverbial relations.

1 Classification, demographics, and sociolinguistic background

Harakmbut is a Peruvian Amazonian language spoken in the regions of Cusco and Madre de Dios. The name of the language is an autonym and means ‘person, people’. The Harakmbut people still live in their traditional homeland, which covers the area drained by the Madre de Dios River and all its tributaries from its headwaters down until the mouth of the Inambari River. As can be seen in Figure 1 [map to be included], the southern border is formed by the Andes. The speakers live in a number of sometimes ethnically mixed *comunidades nativas* ‘native communities’, which are protected by national law. These communities border on the Amarakaeri Communal Reserve, also a protected area, which lies in the center of the Harakmbut homeland.

The Harakmbut are divided into several groups that settled into different areas of the homeland, have some different cultural practices, and speak different varieties of the language (see Gray 1996: 4–16). Table 1 presents data on their current locations and numbers.

Table 1: Dialects of the Harakmbut language

ETHNIC GROUPS/ LANGUAGES ¹	GLOTTO-CODE	NATIVE COMMUNITIES	ETHNIC POPULATION ²	VITALITY STATUS
Amarakaeri (preferred autonym: 'Arakmbut')	amar1274	Puerto Luz, Shintuya, San José Del Karene, Barranco Chico, Boca Inambari, Boca Ishiriwe, Puerto Azul, Masenawa, Kotsimba	1043	Highly endangered
Watipaeri	huac1244/ huac1245	Queros (Cu), Santa Rosa de Huacaria (Cu)	392	Highly endangered
Arasaeri	aras1241	Arazaeri	317	Highly endangered
Pukirieri	-	?	168	Highly endangered
Sapiteri Kisambaeri	sapi1239 kisa1267	Barranco Chico	47	Highly endangered
Toyoeri	toye1240	-	-	extinct

Detailed studies on the dialectal differences are lacking, but the dialects are reported to be mutually intelligible. Solís Fonseca (2003: 158) and Helberg (in prep.) divide them into two main groups: Watipaeri and Toyoeri are phonetically and lexically somewhat different from Amarakaeri/Arakmbut, Arasaeri, and Sapiteri. The data from Aza (1936) and Peck (1979 [1958]), however, suggest that the Arakmbut variety is different from the other four mentioned, which are in turn similar to each other (See Section 2). Of the last five dialects in Table 1 only a handful of fluent speakers, if any, are left.

The Harakmbut are traditionally farmers (slash-and-burn agriculture) and hunter-gatherers, but they started working gold in the 1970s (e.g., Aikman 2009). For a detailed ethnography, the reader is referred to Gray (1996, 1997a, 1997b). The sociolinguistic situation points to a rather low degree of vitality, as I have found many young parents reluctant to pass on the language to their children, so as to protect them from stigmatization as an indigenous person. Children are mainly brought up in Spanish and only acquire a passive competence in Harakmbut. However, efforts are being made to develop didactic materials in Harakmbut, and to implement a programme of Bilingual Intercultural Education funded by the national government. In addition, a number of speakers are trying to boost the communities' self-esteem by publishing on Harakmbut oral tradition and documenting cultural practices. In

¹ It should be noted that the speakers of the Amarakaeri variety regard the label *Amarakaeri* as a derogatory term; it is adapted from *wa-mba-arak-a-eri* (NMLZ-V.PL-kill-TRNS-AN), a deverbal nominalization meaning '(fierce) killer/murderer' (cf. Helberg 1996: 18, in prep.), which goes back to an ancient story about the origin of the different ethnolinguistic groups of the Harakmbut people. They prefer to call their variety 'Arakmbut', as distinct from the Watipaeri variety, towards whose speakers they generally entertain feelings of hostility rather than brotherhood. The etymology of the other ethnonyms is as follows: *wa-tipa-eri* (NMLZ-dig.out.step-AN) 'people from the steps dug into the hillside'; *arāsā-eri* (Arasa-AN) 'people living on the Arasa River' (i.e., the Marcatapa, a tributary of the Inambari); *pukiri-eri* (Pukiri-AN) 'people living on the Pukiri River'; *toyo(dn)-eri* (downriver-AN) 'people living downriver' (cf. Helberg 1996: 18). The etymology of *kisambaeri* and *sapiteri* is unknown; the latter is certainly different from *kapiteri* (*kapite-eri* (Kapite-AN) 'people living on Mount Kapite'), who are part of the Arakmbut group (pace Helberg, in prep.).

² Numbers cited by INEI (Instituto Nacional de Estadística e Informática) in 2007 (courtesy of Yesica Patiachi Tayori). The locations of the ethnic groups without any NC in Table 1 is the pre-contact location assumed in Gray (1996: 6).

general, young adults and speakers up to the age of fifty are bilingual in Harakmbut and Spanish, while speakers older than fifty are mainly monolingual in Harakmbut. In some communities there is language contact across different Harakmbut varieties, and with Matsigenka, Yine (both Arawakan), and Ese Ejja (Takanan).

The genetic classification of Harakmbut, sometimes also termed ‘Harakmbet’, ‘Hate’³ or (mistakenly) ‘Mashco’, has been a matter of dispute. McQuown (1955: 530) and Matteson (1972) classified it as an Arawakan or Maipuran language. Similarly, Greenberg (1960) and Voegelin and Voegelin (1977) placed it in the Andean-Equatorial phylum, specifically the Arawak subbranch of the Equatorial branch (see Payne 1987: 23–24). However, Payne (1991: 365–369) convincingly refuted this classification; since the 1960s, consensus has emerged that Harakmbut is an isolate (Tovar 1961; Loukotka 1968; Lyon 1975; Helberg 1984; Wise 1999: 307). Recently, Adelaar (2000, 2007) proposed that it is genetically related to the Brazilian Katukina family, which may be further linked to Macro-Jê. This proposal is mainly based on lexical evidence and awaits further corroboration. From an areal perspective, Harakmbut exhibits a number of Western Amazonian grammatical features, as well as features characteristic of the nearby Guaporé-Mamoré linguistic area in southwest Brazil and eastern Bolivia (Crevels & van der Voort 2008).

The present chapter is based on previous work as well as my own fieldwork on the Harakmbut language, both of which focus on the Amarakaeri/Arakmbut dialect, which has the highest number of speakers. Earlier work includes studies by Hart (1963), Helberg (1984, 1990) and Tripp (1976a, 1976b, 1976c, 1995). My own data were collected through audio recording during three field trips in 2010, 2011 and 2016 (about five months in total), in the native communities of Puerto Luz, San José del Karene, and Shintuya. The data used in this chapter come mainly from elicitation sessions with bilingual speakers; if taken from spontaneous speech, this has been indicated in the example.

2 Phonology and phonetics

The phonological analysis presented in this section largely follows Helberg’s (1984: 13–178, in prep.) description. Tables 2 and 3 present the consonant and vowel phonemes of the Amarakaeri or ‘Arakmbut’ variety; the corresponding orthographic symbols are given in angled brackets. The practical orthography, designed by the author and indicated in brackets where relevant, is IPA-based, and different from the community spelling. Based on written sources, the segmental inventories of the other varieties seem to differ only with respect to the phonemic status of the glottal fricative [h] <h> and the bilabial approximant [w] <w>.

Table 2: Consonant phonemes of Amarakaeri/Arakmbut

MANNER, PLACE	BILABIAL	ALVEOLAR	VELAR
STOP	p	t	k
NASAL	m <m, mb>	n <n, dn, nd>	ŋ <ŋ, gŋ>
TAP		r <r>	
FRICATIVE		s	
APPROXIMANT	w		

³ The term *hate* (*ate* in Arakmbut) originates in the sequence *Harakmbut-ha-te* (people-say-LOC) ‘in the language of the people, in the Harakmbut language’, in which the deverbal nominalization is attached to the preceding modifier noun (*hate* itself is not an independent form in the language). An alternative construal to this sequence, in the Arakmbut variety, is given in (37b).

Table 3: Vowel phonemes of Amarakaeri/Arakmbut

	FRONT	CENTRAL	BACK
HIGH	i, ĩ		u, ũ
MID	ε <e>, ě <ẽ>		ɔ <o>, õ <õ>
LOW		a, ã	

Both consonant and vowel phonemes underlie a fairly wide range of speech sounds. The plosives /p/, /t/, /k/ generally show a voiceless realization, unless they occur in intervocalic position, in which they tend to become voiced ([b], [d], [g]), as in (3)—cf. Helberg (1984, in prep.).⁴ They are unreleased in syllable-final position, as with [kʰ] in (1). Alveolar /t/, /n/, and /s/ are often palatalized to [tʃ], [nʃ] and [ʃ] when followed by high front vowels (older speakers do not always palatalize them); for the [nd] allophone of /n/ (see below) this results in [ndʒ]. In case the high front vowels are in turn followed by a different vowel, the former tend to be unpronounced (but need to be posited for morphological reasons), as with the second syllable in (1).⁵ Palatalization of /s/ is also triggered by ensuing high back vowels.

- (1) *e-ti-aʔ-pak* /ε.tiaʔ.pak/ [ε.ʔaʔ.pakʰ] V.CVVʔ.CVC ‘to narrate’
 NMLZ-SPAT:up-say-VBZ (syllables are separated by a dot symbol)

All three nasal consonants show allophonic variation with pre- and/or post-stopped variants. In the case of preceding oral vowels, /n/ and /ŋ/ are realized as pre-stopped [dn] and [gŋ] respectively; in the case of preceding nasal vowels, they are realized as [n] and [ŋ]. In the case of following oral vowels, /n/ and /m/ are realized as post-stopped [nd] and [mb] respectively; in the case of following nasal vowels, they are realized as [n] and [m] (cf. Helberg 1984, in prep.). The same allophonic distribution has been noted for Toyoeri and Sapiteri by Peck (1979 [1958]: 18–21), and for Watipaeri and Arasaeri in a schoolbook drawn up by a multi-dialect author team (Manqueriapa Vitente et al. 2012). However, there is also some free variation within and across speakers of Arakmbut, e.g., *wa-mbaʔ-neŋ* (NMLZ-hand-amount) ‘five’ is attested as both [waʰ.maʔnẽŋ] and [waʰ.mbaʔnẽŋ]. In addition, there is a nasal harmony system (2), in which the nasal root elements are in bold, and there are degrees of nasality in vowels. All these aspects of nasality await further analysis. In the practical orthography used here, only strong (phonemic) nasality in vowels is indicated with a tilde; weak nasality (through co-articulation) is not indicated, for example in the proper name *Morimõ*. The consonant nasal allophones are spelled according to their realization.

- (2) *wã-tõ-ě* *mě-tã-ě-ně*
 NMLZ-CAUS.SOC-be 3SG>1/2SG.IND-APPL-be-IND
 ‘This is my husband.’

The phonetic realization of the vowel phonemes is very close to that of the corresponding cardinal vowels, apart from the mid vowels, which have a more raised pronunciation than the open-mid cardinal vowels. Allophonic variation is most noticeable when two vowel segments are adjacent. When adjacent to low central vowels, for instance, mid front vowels tend to be raised considerably, often even to palatal semi-vowel realizations; for example, the verb form in (2) is often pronounced as [mẽʰtãŋnẽ]. Raising also takes place when mid front vowels

⁴ In Toyoeri and Sapiteri, lenition in intervocalic position has only been described for /k/ (Peck 1979 [1958]: 17).

⁵ In Toyoeri and Sapiteri, palatalization is only reported for /t/ when followed by high front vowels, which also go unpronounced when they are themselves followed by a different vowel (Peck 1979 [1958]: 23–25).

precede low central vowels, which may in turn lead to palatalization of the preceding consonant (3).

- (3) *kate-apo* [1] Raising of /ε/ adjacent to /a/: ['ka.tia.bə]
 what-REAS [2] Palatalization of /t/: ['ka.tʃa.bə]
 ‘why?’

Similarly, mid back vowels tend to be raised and realized as [w] when they precede low central or mid front vowels in a single consonant-initial syllable (4). Helberg (1984; in prep.) only notes the [w] allophone of the mid back vowels in syllable-initial position and does not attribute phonemic status to the bilabial approximant. More detailed study is in order here. In general, vowels are spelled phonemically, for example /ε/ <e> in (3), except for /i, ĩ/, which are spelled <y, ÿ> when realized as [j] adjacent to vowels in a single syllable, as with the final syllable in (4). Vowels in stressed syllables are slightly longer than those in unstressed syllables. Syllable-final vowels of unstressed syllables are often elided (5b), especially in rapid speech.

The canonical syllable structure is (C)(V)V(V)(C), with the optional vowel segments restricted to the back and high front vowels phonetically realized as a semi-vowels (4), or “swallowed” after inducing palatalization of the preceding consonant (1). Table 4 presents the segmental restrictions on syllable onsets and codas for the consonant system (however, /r/ is excluded in word-initial position); syllables that contain nuclear vowels only do not have any segmental restrictions.

- (4) *ĩ-nõ-põ-ẽ-ỹ* /ĩ.nõ.põẽĩ/ [ĩ.'nõ.pwẽj]
 1SG-vital.center-CLF:round-be-1.IND V.CV.CVVV
 ‘I know’

Table 4: Segmental restrictions on syllable onsets and codas (consonants)

Position	allowed in onset and coda	allowed in onset, not allowed in coda	allowed in coda, not allowed in onset
Consonant phonemes	/t/, /k/, /n/, /s/	/p/, /m/, /r/, /w/	/ŋ/

Two speech sounds have not been discussed so far; they are the glottal stop [ʔ] <ʔ> and fricative [h] <h>, which show dialectal variation. Helberg (1984: 22, in prep.) argues that the glottal stop has phonemic status in Arakmbut, but my data do not contain conclusive evidence in support of this analysis. Rather, I believe it has the same suprasegmental function Helberg (1984: 143–148, in prep.) describes for [h], which is to optionally demarcate syllable boundaries when these lack consonantal onsets or codas. Its non-phonemic status is corroborated by the observation that its production varies within and across speakers of Arakmbut. In addition, Peck (1979 [1958]) does not report it for Toyoeri and Sapiteri. While Helberg (1984: 143–148, in prep.) analyzes [h] as a suprasegmental element in Arakmbut, Peck (1979 [1958]: 2) posits phonemic status for it in Toyoeri and Sapiteri. A preliminary comparison of 35 lexical items based on Peck (1979 [1958]), Manqueriapa Vitente et al. (2012), and my own data shows that (in writing) in Watipaeri/Toyoeri/Sapiteri, <h> before <i> corresponds to [w] in Arakmbut, while in these same varieties <h> before <e> corresponds to [w], [ʔ], or Ø in identical phonetic environments in Arakmbut. Before <a>, <o> and <u> (nasal or oral), <h> is either retained in Arakmbut, or pronounced [ʔ], or left unpronounced (Ø). In syllable-final position, Watipaeri/Toyoeri/Sapiteri <h> tends to be retained in Arakmbut. More research is needed on this dialectal variation.

The domain of stress assignment in a phonological word differs according to word class. For nouns, the stress domain is the root plus derivational affixes; the main stress falls on the penultimate syllable. In examples (1)–(5), word stress has been indicated through the stress symbol ['] or underlined syllable nuclei. As shown in (3) and (5c), inflectional suffixes like case endings do not bear on stress assignment. By contrast, derivational suffixes attaching to nouns typically affect stress placement (5).

- (5) a. *wã-wẽ* [wã.wẽ] b. *wã-wẽ-ẽrĩ* [wã.'wẽ.rĩ]
 NMLZ-liquid NMLZ-liquid-AN
 ‘liquid, river’ ‘river spirit’
- c. *wã-wẽ-ẽrĩ-tã* [wã.'wẽ.rĩ.tã]
 NMLZ-liquid-AN-ACC
 ‘river spirit’ (direct object function)

Verb forms also show stress on the penultimate syllable, but they have a different stress domain, as not all derivational suffixes are included (e.g., the transitivizer *-a* in (6) does not influence stress placement). This topic also needs further research.

3 Morphological profile and basic word classes

In terms of phonological fusion, grammatical markers in Harakmbut are phonologically bound, and are thus concatenative in nature (see Bickel & Nichols 2007: 181). For example, all case and TAME markers attach to their respective host words and form a single phonological word together with their host, as with the complex words in (6). Harakmbut has no isolating or nonlinear formatives.

- (6) *Nãŋ oʔ-tay-a wa-siʔ-po-ta*
 mother 3SG.IND-sleep-TRNS NMLZ-peel-CLF:round-ACC
 ‘The mother puts the child to bed.’

With respect to exponence, that is, the degree to which categories cumulate into single formatives (Bickel & Nichols 2007: 188), case markers generally are monoexponential (see Sections 4.1 and 6.2), whereas TAME markers are polyexponential (see Sections 5.1 and 5.2). In terms of flexivity, Harakmbut has nonflexive formatives; allomorphy is never lexically determined (item-based), but phonologically conditioned. Harakmbut formatives are thus generally agglutinative in the sense of concatenative-nonflexive.

Harakmbut grammatically distinguishes the open parts-of-speech classes of nouns, verbs, adjectives, and adverbs, as well as the closed classes of pronouns, quantifiers, auxiliaries, interjections, polar-question particles, ideophones (Helberg 1984: 241–242), and clitics. In addition, a small number of words are flexible: *sik*, for instance, can function as a verb root (‘become dark’) as well as a noun, for example, when marked for locative case in the temporal expression *sik-yo* (black-LOC) ‘at night’.

4 The noun phrase

4.1 Morphological template of the head

Within the Harakmbut noun phrase, a number of categories are marked on the head element. Table 5 presents the morphological template of the head of a noun phrase, with a linear order of base-collective-case-focus₁-focus₂ (note the two distinct focus slots). The first slot is filled by the head element of the NP, which can be of four types, namely common and proper

nouns, as with *apetpet* and *Luis* in (7), respectively; pronouns, as with *ndo?* and *opudn* (9); or filler words such as *ānĩ* in (8).⁶

- (7) *Luis-ʔa-nda oʔ-arak-me apetpet-ta*
 Luis-NOM-FOC 3SG.IND-kill-REC.PST jaguar-ACC
 ‘Luis himself killed the jaguar.’

Table 5: Morphological template of the head of an NP⁷

BASE	COLLECTIVE	CASE	FOCUS ₁	FOCUS ₂	
common noun	-(o)mey COL ^{H, T}	-ʔa~-a	NOM ^{H, T}	-yo REST ^{H, T}	-nda FOC ^H
proper noun	T		INS ^{H, T}	-nāyō COND	
pronoun		-ere	COM ^{H, T}		
filler word			INS ^T		
		-ta(h)	ACC ^{H, T}		
		-en~-edn~-	GEN ^{H, T}		
		wedn~-ʔedn			
		-tewapa	BEN ^T		
		-(o)niŋ	SIM		
		-apo	REAS		
		-mbayo	PRIV		
		-yo; -ya; -	LOC ^{H, T}		
		taʔ; -te; -			
		yon; -pen			

- (8) *oʔ-wa-me-ne sabado-taʔ ānĩ-ŷō wē-ũk-yo*
 1PL.INCL-go-REC.PST-IND Saturday(Sp)-LOC FILLER-LOC river-hot-LOC
 ‘We went to, um, Aguas Calientes on Saturday.’ (spontaneous speech: anecdote)

The second slot is devoted to collective marking (see Section 4.6), which is only available to nouns and plural personal pronouns, like *opudn* in (9).

- (9) *ndo-a opudn-omey-tah on-to-mba-pe-apet*
 1SG-NOM 2PL-COLL-ACC 1<>2PL.DUB-CAUS.SOC-V.PL-eat-FUT.EPIST
 ‘I am going to invite you (PL) to eat.’ (Tripp 1995: 191; my segmentations & glosses)

The third slot hosts case suffixes. Case is monoexponential in Harakmbut, and it is inert: it is marked only once for the NP it has in its scope. In addition, its marking is symmetrical across nominal and pronominal systems; compare (7) with (9). To date, there is no evidence of case spreading or stacking.⁸ The syntax of case will be discussed in Section 6.2.

Harakmbut has a rather extensive set of case markers, two of which show syncretism, or rather polysemy, that is, the first two case suffixes listed in Table 5. The suffix *-a* marks nominative (7) or instrumental case (28). The marker *-ere* can be used to signal an instrument (like *-a*) (10), but additionally expresses accompaniment, or a comitative relationship.

⁶ Note that *ānĩ* does not have any lexical content; it only marks a pause or hesitation in speech.

⁷ H = Helberg 1984: 436–444; T = Tripp 1995: 194–200.

⁸ There are cases that superficially show case stacking of the pattern N₁-GEN N₂-GEN-COM N₃, which means ‘the N₃ of (both) N₁ and N₂’. In such structures, however, the comitative marker functions as a coordinative linker, yielding a conjoined possessor phrase.

- (10) *kumeh-ere-yo-nda on-mbaʔ-wek-me*
 bow-INS-REST-FOC 3PL.IND-V.PL-wound.with.arrow-REC.PST
 ‘They pierced it with an arrow only.’

The accusative suffix *-ta* is used to code primary objects (see Section 6.2), and genitive marking is used to signal the syntactic relation of attributive possession (see Section 4.5). The case ending used to signal the semantic role of beneficiary is rather long, *-tewapa*, and is illustrated in (25). The case ending coding similative adjuncts is *-niŋ*; adjuncts of reason feature the case ending *-apo*, as in (3) and (41). Harakmbut also has a privative case marker, *-mbayo*, which expresses the lack or absence of the referent of the head it is marked on (11).

- (11) *paŋ-ɲ-mbayo ʔ-ẽ-ɲ*
 father-PRIV 1SG-be-1.IND
 ‘I have no parents; I am without parents’

For adjuncts indicating locations or directions, Harakmbut has a number of locative case markers (Tripp 1995: 196), for example, *-yo* (8) and *-ya* (42), occurring on the same nouns. Two markers are used for temporal location also, for example, *-taʔ* (8) and *-te*. In spatial contexts, *-te* involves contact between figure and ground (12).

- (12) *ken ǎni on-ma-ndeh-po muneyo-siʔpo bisikleta-te=kon*
 3 FILLER 3PL.IND-V.PL-meet-DEP girl-DIM bicycle(Sp)-LOC=ADD
 ‘Then, um, they cross one another; the little girl is also on a bike.’ (spontaneous speech: Pear story)

Finally, the fourth and fifth slots are occupied by focus markers. The fifth includes just the focus marker *-nda*, as in (7) and (10). It is also found in adjectival constructions (see Section 4.7). Other focus markers can precede it in the fourth slot, for example, the restrictive focus marker *-yo* ‘only’ (10). Clitics, like the additive focus marker *=kon* in (12) above, have not been included in the template of the head of an NP, as they can attach to other word classes.

4.2 Pronouns and demonstratives

The Harakmbut paradigm of independent personal pronouns formally distinguishes between first, second, and third persons, and for the first two, it also distinguishes between singular and plural number (Table 6).

Table 6: Personal pronouns in Harakmbut (cf. Helberg 1984: 254; Tripp 1995: 198)

	SINGULAR	PLURAL
1	<i>ndoʔ</i>	<i>oroʔ</i>
2	<i>on</i>	<i>opudn</i>
3		<i>ken</i>

The third person pronoun *ken* (Table 6) also functions as a distal demonstrative modifier. As detailed in Table 7, the demonstratives distinguish between distal/remote (47) and proximal values (13), and also between pronoun and adjective/modifier function.

Table 7: Demonstratives in Harakmbut

	MODIFIER (DEPENDENT)	INDEPENDENT PRONOUN
PROXIMAL	<i>in</i>	<i>ine</i>
DISTAL	<i>ken</i>	<i>kene</i>

- (13) *ine* *ōʔ-ě-tā* *ine* *ine*
 PROX 3SG.IND-be-INFR PROX PROX
 ‘It must be this one, (this one, this one).’ (spontaneous speech: Family problems, San Roque et al. 2012)

Table 8 presents the interrogative pronouns and modifiers, which form the basis of a number of indefinite pronouns. Here the main distinction is a semantic one, that is, human versus non-human referents. The pronouns can be inflected for any case, as with *kate-apo* in (3).

Table 8: Interrogative pronouns/modifiers in Harakmbut

	MODIFIER (DEPENDENT)	INDEPENDENT PRONOUN
HUMAN	<i>mbeʔ</i> ‘which’	<i>mbeʔ</i> ‘who’
NON-HUMAN	<i>men/kate</i>	<i>kate</i> ‘what’

Two forms have been found for adnominal question words targeting non-human referents. Judging from (14) and (15), it may be hypothesized that *men* is used to ask for identification of one or more members of a set (‘which?’), while *kate* is used to ask for type specification (‘what type/sort of?’) (but see [18] in Section 4.3).

- (14) *men* *kōsō* *ya-poʔ-sak-on?*
 which pot 3SG.DUB-CLF:round-break-PFV.NVOL
 ‘Which pot broke?’
- (15) *kate* *aypo* *iʔ-pak-ika-Ø?*
 what food 2SG-want-HAB-DUB
 ‘What sort of food do you (SG) like?’

The interrogative forms in Table 8 also serve as indefinite ones, as in (16) and (20). Table 9 details that free-choice indefinites host the clitic =*piʔ* INDETERMINATE, which attaches only after possible case endings (16). Realis indefinites, by contrast, do not (*pace* Tripp 1995: 200).

Table 9: Indefinites in Harakmbut (N = nominal head)

	MODIFIER (DEPENDENT)	INDEPENDENT PRONOUN
REALIS	HUMAN <i>men, mbeʔ</i> ‘some’	<i>mbeʔ</i> ‘somebody’
	NON-HUMAN <i>men</i> ‘some’	<i>kate</i> ‘something’
FREE CHOICE	HUMAN <i>mbeʔ</i> N= <i>piʔ</i> ‘whichever N’	<i>mbeʔ=piʔ</i> ‘who(m)ever’
	NON-HUMAN <i>moning</i> N= <i>piʔ</i> ‘whichever N’	<i>kate=piʔ</i> ‘whatever’

- (16) *mboerek* *o-mba-yok-me* *tareʔ* *mbeʔ-ta=piʔ*
 man 3SG.IND-V.PL-give-REC.PST manioc somebody-ACC=INDET
 ‘The man gave manioc to whom(so)ever.’

Furthermore, the interrogative pronouns also serve as negative indefinite pronouns (human *mbe?*; non-human *kate*) when they occur with the negative particle or predicate negation. Finally, Harakmbut also has a third person reflexive pronoun (50b), and it lacks relative pronouns. (Relative clauses are discussed in Section 7.1.)

4.3 Common nouns

Harakmbut common nouns can be divided into two classes on the basis of their morphological status, that is, potentially free nouns versus obligatorily bound nouns. This distinction is relevant to noun modification, noun incorporation, and word formation. While potentially free nouns can stand on their own as a word, obligatorily bound ones require a prefix to obtain independent nominal status. The two prefixes used, *wa(?)*- and *e(?)*-, are analyzed as semantically empty nominalizers that derive independent nouns from bound ones;⁹ they also serve in deverbal nominalization (see Sections 4.8 and 7). The bound root *-mba?*, for example, gives rise to two distinct independent nouns whose referents show a similarity in shape and form an upper extremity of a living body, cf. (17).

- | | | | | | |
|------|----|----------------------------|--|----|----------------------------|
| (17) | a. | <i>wa-mba?</i> | | b. | <i>e-mba?</i> |
| | | NMLZ-hand | | | NMLZ-hand |
| | | ‘hand’ (Helberg 1984: 437) | | | ‘leaf’ (Helberg 1984: 437) |

Differences between free and bound nouns also show up in the noun phrase. When combined with adnominal modifiers that obligatorily precede the nominal head in continuous noun phrases,¹⁰ free nouns show a single construction type, while bound nouns show two: (i) one in which they attach to a nominalizing prefix and follow the modifier like free nouns (18a), and (ii) one in which they fuse with their modifier (18b). In (18a) the (interrogative) modifier-head structure is like that in (15) with the free noun *aypo* ‘food’.

- | | | | | |
|------|----|----------------------|----------------|----------------|
| (18) | a. | <i>kate</i> | <i>wa-ndik</i> | <i>ĩʔ-ě-Ø?</i> |
| | | what | NMLZ-name | 2SG-be-DUB |
| | | ‘What is your name?’ | | |
| | b. | <i>kate-ndik</i> | <i>ĩʔ-ě-Ø?</i> | |
| | | what-name | 2SG-be-DUB | |
| | | ‘What is your name?’ | | |

Similarly to (18b), bound nouns also attach to genitive-marked (pro)nouns (Section 4.5) and to quantifiers (Section 4.6). Information included in lexical entries in Tripp (1995: 266a, s.v. *día* ‘day’) suggests that they attach to demonstrative modifiers and deictic adjectives like *noŋ* ‘other’ as well. Another difference between the two noun types is that only bound nouns are incorporable into the verb (Section 5.9).

Semantically, the set of free nouns is fairly heterogeneous, whereas the set of obligatorily bound nouns shows more homogeneity. The latter refer to inalienably possessed entities, such

⁹ The finding that the two nominalizers lead to two different lexical items when attached to the same noun root (17) might challenge the claim that they are semantically empty. However, noun roots that combine with both nominalizers are few; I am aware of one other pair: *eʔ-pu?* ‘bamboo’ vs. *wa-pu?* ‘tube’. Like in (17), the two lexical items are similar in shape. Note that the cases of *-mba?* and *-pu?* defy the generalization that *wa(?)*- is linked exclusively with the expression of inherent possession by an animate entity.

¹⁰ Harakmbut also features discontinuous noun phrases, characterized by the presence of quantifiers or descriptive modifiers (see Section 4.7).

as body parts, plant parts, and landscape parts (cf. the class of *e*-nouns in Cavineña as described by Guillaume 2008: 409–416), as well as kinship terms and basic shapes or qualities of entities. This set (excluding kin terms) has been identified as “shape morphemes” by Hart (1963), and analyzed as classifiers by Payne (1987). However, I will argue that only a subset of the bound nouns also function as classifiers, that is, the nouns that show classificatory noun incorporation (see Section 5.9). Not a single bound noun functions as a classifier in possessive, numeral, attributive, locative, or demonstrative nominal constructions. Example (19) with head noun *kuwa* ‘dog’, a free noun root, illustrates that Harakmbut does not use classifiers in possessive, demonstrative, and adjectival modifier environments, nor does it feature a noun classifier, in the sense of classifier from Aikhenvald (2000: 204–207).

- (19) *mbeʔ-edn* *ỹã-tã-ẽ* *in* *kuwa* *uru-ndaʔ*
 who-GEN 3SG.DUB-APPL-be PROX dog beautiful-NDA
 ‘Whose is this beautiful dog?’

4.4 Modification by demonstratives and indefinites

Harakmbut lacks articles which would express definiteness or specificity but does have demonstrative and indefinite modifiers (cf. Section 4.2). These invariably precede their nominal head and do not show any type of agreement with it. Demonstrative modifiers typically realize definite reference, for example, the proximal modifier in (19). In (20), indefinite modifier *mbeʔ* (see Table 9) realizes indefinite, non-specific reference.

- (20) *mbeʔ* *wettone* *oʔ-tiak-me* *taʔmba-yo*
 some woman 3SG.IND-come-REC.PST swidden-LOC
 ‘Some woman came to the swidden.’

4.5 Attributive possession

In Harakmbut, the syntactic relation of attributive possession is reflected by dependent marking: (pro)nouns denoting the possessor are marked for genitive case; the possessum is unmarked (21). The order is that of possessor-possessum (Tripp 1995: 195).

- (21) *ndoʔ-edn* *nãŋ*
 1SG-GEN mother
 ‘my mother’

Free and bound nouns show distinct morphosyntactic behaviour. Free possessed nouns, like *nãŋ* in (21), use the pattern in which the possessor and possessum form two distinct phonological words; stressed syllable nuclei are underlined. Bound possessed nouns have two patterns at their disposal: (i) the default pattern shared with free nouns (22), and (ii) a pattern exclusively available to bound nouns (23a). Like in (18b), this pattern is characterized by the absence of a nominalizing prefix and by the (possessive) modifier and head noun forming a single phonological word; compare (23a) with (23b) and (23c).

- (22) *ndoʔ-edn* *wa-nda-po* *õ-mẽʔ-aʔ*
 1SG-GEN NMLZ-CLF:fruit-CLF:round 3SG.IND-liver-say
 ‘My belly is making noise.’ (lit. ‘liver-says’)

- (23) a. *arakmbut-edn-ndik* b. *arakmbut* c. *wa-ndik*
 people-GEN-name people;person NMLZ-name
 ‘native word’ ‘people’ ‘name’
 (‘name of the people’)

4.6 Nominal number and quantification

Harakmbut does not require number marking on nouns in any context (cf. Tripp 1995: 194). Strategies to express overtly that more than one referent is involved include collective marking, modification by numerals, and indefinite quantifiers, as well as verbal plural marking (see Section 5.6). This section focuses on the first two strategies, confined to the noun phrase.

Harakmbut is noted to have two collective suffixes, *-(o)mey* and *-kupo*, which attach to nouns and plural personal pronouns (Tripp 1995: 194, 198). No uses of *-kupo* as a collective suffix have been found so far (as a bound noun it means ‘buttocks’), and only few uses of *-(o)mey*, on nouns with human referents engaged in a joint activity (cf. Corbett 2000: 199), cf. (24).

- (24) *wettone?-mey* *mba-tiaway-we* *õn-mã-ẽ-mẽ-tẽ*
 woman-COLL V.PL-see-NEG 3PL.IND-V.PL-be-REC.PST-NFIRSTH
- wa-mationka-eri-ta*
 NMLZ-hunt-AN-ACC
 ‘The (group of) women didn’t find the hunters.’

The second strategy involves modification by numerals and indefinite quantifiers. Harakmbut has a restricted numeral system. Most speakers can only count up to five (Table 10). However, I also recorded numeral expressions up to twenty from a very skilled speaker, who used elaborate periphrastic constructions referring to fingers, hands, toes, and feet. These expressions are generally accepted to form the original system.

Table 10: Cardinal numbers in Harakmbut

1	<i>noŋ-ti-nda</i>	other-SPAT:up-NDA
2	<i>mbotta?</i>	two
3	<i>mbapa?</i>	three
4	<i>mbotta?-mbotta?</i>	two-two
5	<i>wa-mba?-nej</i>	NMLZ-hand-amount

Harakmbut also has a set of indefinite quantifiers, namely absolute quantifiers *suwing* ‘(a) few/little’ and *wakka* ‘many/much’, and universal quantifier *aya* ‘all’ (cf. Helberg 1984: 257); there are no counterparts of *no*, *most*, and *each/every*.

Quantifying modifiers obligatorily precede the nominal head in continuous noun phrases. Example (25) illustrates the construction type shared by free and bound nouns for indefinite quantifiers, in which *wakka* is suffixed with *-nda*; (26) shows the shared construction type for cardinal quantifiers, in which the numeral lacks *-nda*.

- (25) *õn-mã-wẽyã-mẽ* *wakka-nda* *aypo* *aya-tewapa-nda*
 3PL.IND-V.PL-cook-REC.PST much-NDA food all-BEN-NDA
 ‘They cooked enough food for everyone.’

- (26) *Ih-yok-i mbotta? kuwa Luis-ta*
 1SG-give-1.IND two dog Luis-ACC
 ‘I give two dogs to Luis.’

Bound nouns stand out in allowing for an additional construction type involving phonological fusion (27). Example (27a) instantiates the same type as (25), while (27b) shows fusion of the quantifying modifier and the noun root into one phonological word, having a single stress. Numeral modifiers fuse with bound nouns analogously (but do not show *-nda*).

- (27) a. *ĩh-tõ-ẽ-ỹ aya-nda wa-?idn*
 1SG-CAUS. SOC-be-1.IND all-NDA NMLZ-tooth
 ‘I have all my teeth.’
- b. *ĩh-tõ-ẽ-ỹ wakkq-?idn-a-nda*
 1SG-CAUS.SOC-be-1.IND many-tooth-EP.V-NDA
 ‘I have many teeth.’

In addition to continuous noun phrases, indefinite quantifiers are also attested in discontinuous noun phrases, which have not been found with numeral modifiers. In such cases, the indefinite quantifier does not carry the suffix *-nda*. Numerals and indefinite quantifiers can also be used independently, in which case both types take the suffix *-nda*, as in (25) with the indefinite quantifier ‘all’.

4.7 Descriptive modification

In Harakmbut, “semantic adjectives” (Dryer 2007: 168) form a distinct word class (see also Helberg 1984: 241; Tripp 1995: 197–198), showing specific morpho-syntactic characteristics that are not yet well understood. Adjectives modifying nouns do not show any agreement with their head (see also Tripp 1995: 197). They appear in both continuous and discontinuous noun phrases. In the first type, they occur in prenominal and postnominal position. The noun phrase type and position of adjectives seem to be determined by the referential properties of the noun phrase they are part of. Free and bound head nouns do not behave differently, which is in line with the generalization in Section 4.3.

Continuous noun phrases in which the adjective immediately follows the head realize different types of reference, and their more specific formal features seem to differ accordingly. Generic noun phrases, like the one in boldface in (28), feature adjectives that are prefixed by *wa(?)*- and carry no suffix. Specific NPs, either definite or indefinite (29), by contrast, require the *-nda* suffix on the adjective, which typically does not carry the prefix *wa(?)*-. This pattern is attested for both free head nouns (19) and *wa(?)*-prefixed bound nouns (29). In discontinuous NPs, the adjective also follows its head noun, but comes only after the finite verb. Such NPs realize indefinite reference, and the adjectives do not carry *wa(?)*-, but do take the suffix *-nda*, just like in the continuous noun phrase in (29).

- (28) *mba-e-a-ndik õ?-ẽ mbi?igŋ wa-mboro? kumo-a*
 V.PL-get-TRNS-POT 3SG.IND-be fish NMLZ-big barbasco-INS
 ‘One can catch big fish with barbasco.’ or ‘Big fish can be caught with barbasco.’
- (29) *ĩh-tõ-ẽ-ỹ wa-?i mboro?-nda*
 1SG-CAUS.SOC-be-1.IND NMLZ-foot big-NDA
 ‘I have big feet.’

Continuous noun phrases with prenominal adjectives are also restricted in terms of the type of reference they realize, and again their formal features differ accordingly. In non-referential noun phrases, for example, those serving as predicative nominals (30), adjectives carry the prefix *wa(?)*- and often also the suffix *-nda*. Specific definite (referential) noun phrases show a different pattern (31), without prefix *wa(?)*- and with *-nda* not suffixed to the adjective, but to the head noun. The noun phrase in (31) also shows phonological fusion (the stressed syllable nucleus is underlined), but in similar polysyllabic structures stress patterns are less clearly indicative of a single phonological word; more investigation is needed.

(30) *ken ðnʔ-ẽ wa-ndak-nda wa-mationka-eri*
 3 3PL.IND-be NMLZ-good-NDA NMLZ-hunt-AN
 ‘They are good hunters.’

(31) *aʔ-yok-i sal uru-wettone-ta-nda*
 1SG.IMP-give-1.IMP salt(Sp) beautiful-woman-ACC-NDA
 ‘I (should) give salt to the beautiful woman.’

So far, the analysis of adjective constructions has been imprecise about the functions of prefix *wa(?)*- and suffix *-nda*. While the function of *wa(?)*- is clear in endocentric and deverbal exocentric nominalizations (see Sections 4.3 and 4.8), its function in adjective constructions is less well understood. It is also glossed as nominalizing prefix for lack of a better alternative.

The analysis of the suffix *-nda* remains equally unclear. When attached to a case/focus₁-marked noun or pronoun, it functions as a focus marker (7). In adjective constructions, by contrast, it does not signal information focus or increased degree (*pace* Tripp 1995: 197). In view of its occurrence on several types of adnominal modifiers and also adverbs, as well as on nominalized verb forms coding the adverbial relation of simultaneity (66), its function might be that of producing a general modifier.

4.8 Word formation

A number of bound noun roots, sometimes analyzed as classifiers, have been noted to attach to other nouns or noun roots (cf. Hart 1963: 1–2; Helberg 1984: 247–249; Payne 1987: 36–37; Tripp 1995: 193), yielding morphologically complex nouns that can function as nominal heads. This section will discuss noun-based as well as verb-based word formation processes.

Compounding is illustrated in (32), which consists of two morphologically free nouns.

(32) *ndumba-kuwa*
 forest-dog
 ‘bush dog’ [*Speothos venaticus*, AVL] (Helberg 1984: 252; Tripp 1995: 194)

In the examples in (33), the final elements are all bound nouns directly attached to the preceding element.¹¹ Examples (33a) and (33b) are analyzed as modifier-head structures resulting from compounding (e.g., a manioc leaf is a type of leaf). In (33c), by contrast, the first element denotes a type of material, and the second element a type of shape, which

¹¹ Note that in (33b) the first bound root *-mbaʔ* does attach to the prefix *wa[ʔ]*- to attain independent nominal status.

together denote a type of material having a particular shape. (33c) is therefore analyzed as a classifier-derived noun (cf. Payne 1987; see Rose & Van linden 2022).

- (33) a. *tareʔ-mbaʔ* b. *wa-mbaʔ-ku* c. *peraʔ-po*
 manioc-hand NMLZ-hand-head rubber-CLF:round
 ‘manioc leaf’ ‘fingernail’ ‘[e.g., plastic] ball’ (Hart 1963: 5)

Bound morphemes denoting shape are also used in the formation of complex body part nouns. These terms often contain what has been called ‘linkers’ by Hart (1963: 6), that is, bound morphemes that link shape-denoting morphemes to bound noun roots, specifying their spatial configuration, for example, *-taʔ-* ‘base, against, towards’ in (34a), and *-ti-* ‘up, on top of’ in (34b), and which may also occur in pre-root slots in verb forms (see Sections 5.7 and 5.8).

- (34) a. *wa-mba-taʔ-meh-po* b. *wa-kpo-ku-ti-mbaʔ*
 NMLZ-hand-SPAT:base-hump-CLF:round NMLZ-eye-head-SPAT:up-CLF:hand
 ‘wrist’ ‘eyelid’

In addition, Harakmbut has a number of clearly derivational suffixes that attach equally to both free nouns and bound nouns prefixed by a nominalizer. Nominal bases suffixed by *-eri* come to refer to animate entities living in or coming from the place denoted by the nominal base, which can be a common noun (6b) or a proper noun (35). The derived nouns are often true demonyms or gentilics (35). The same suffix is used in deverbal nouns with animate referents.

- (35) *Porto-lus-eri*
 Puerto-Luz-AN
 ‘people living in/coming from Puerto Luz’

Finally, common and proper nouns can also be suffixed by derivational morphemes that characterize the referent of the nominal base in terms of age and/or size, that is, *-toneʔ* (adult, old, big) (36) and *-siʔpo* (young, age of a child, small) (cf. Tripp 1995: 193), analyzed as a diminutive suffix in (12).

- (36) *i-wa-y* *widn-toneʔ* *ō-mã-ẽ-nĩŋ* *kěyõn*
 1SG-go-1.IND stone-big 3SG.IND-V.PL-be-REL thither
 ‘I am going to where there are big stones’

Turning to verb-based derivation, Harakmbut is found to use the same affixes as in noun-based derivation. The two nominalizing prefixes also derive inanimate nouns from lexical verbs (see Van linden 2019). Prefixation of *wa(ʔ)-* is used for instrument (37a) and object nominalizations (37b) (cf. Comrie & Thompson 2007: 338–342), but prefixation of *e(ʔ)-* only for the latter type (37c).

- (37) a. *wa-wedn* b. *arakmbut-edn* *waʔ-aʔ-te*
 NMLZ-lie people-GEN NMLZ-say-LOC
 ‘bed’ (= something ‘in the language of the people’, ‘in the
 for the purpose of lying) Harakmbut language’

5.3 Tense, evidentiality and modality

Harakmbut distinguishes between present (zero-marked), future (-*apo*), recent past (-*me*) and distant past tense (-*uy*) on finite verb forms in Sf6 (cf. Tripp 1976a, 1995: 221–222; *pace* Helberg 1984: 277). Past forms are also obligatorily marked for evidentiality, that is, direct, which is zero-marked (8), versus indirect, or non-firsthand, evidential, marked by -(*a*)*te* suffixed to past tense markers -*me* (39) and -*uy* (42), or by portmanteau -*tuy* (44) (Tripp 1976a, 1995: 222; Helberg 1984: 277–279). Indirect evidential markers compete with the mood+agreement suffixes for Sf7 (39).

- (42) *hak-ʔudn-ya* *o-ti-kot-uy-ate* *wẽỹ-paʔ-a*
house-upper.back-LOC 3SG.IND-SPAT:up-fall-REM.PST-NFIRSTH tree-CLF:rod-NOM
‘A branch fell on the roof long ago.’ (speaker did not see it happen)

In addition to direct versus indirect evidentiality, signalling whether the speaker witnessed the described event or not (see Van linden 2020), Harakmbut also marks inferential evidentiality, with -*ta* (13), and epistemic modality, with -*et*, as in (9) and (38), in Sf7 (Section 5.1). Predictions carry indicative mood and future tense markers; future-oriented possibility is expressed by dubitative-marked verb forms suffixed by -*ipot*.

Expressions of root modality (except participant-inherent subtypes, see Section 7.2) feature periphrastic constructions with a non-finite form suffixed by potential -*ndik* immediately followed by auxiliary *ẽʔẽ* ‘be’. These can be used to express the speaker’s assessment of a state of affairs as (un)desirable or (un)acceptable (deontic modality), or to indicate (im)possibilities or necessities inherent in situations (dynamic modality), as in (28) and (49b). The non-finite verb forms can carry verbal plural, valency-changing, and negation morphology, but no nominalizing prefix. TAME and argument marking occur on the auxiliary. The language also has an apprehensive suffix, -*apey* (Helberg Chávez 1990: 240; Tripp 1995: 222), found with all person subjects, but more research is needed on the set of person markers it combines with. Lastly, although not strictly modal, the suffix -*ay* (Sf4) also deserves mention—(40b) above—which signals that an action was narrowly averted (cf. Tripp 1995: 220).

5.4 Aspect

Harakmbut verbs are found with an extended set of markers that code pluractionality, aspect, and specific temporal adverbial meanings. All of these have been noted in earlier work, but sometimes analyzed differently (Helberg 1984: 284–286; Tripp 1995: 220–221).

Pluractionality marking includes the habitual suffix -*ika* (Sf5), as in (15) and (47), and iterative -*e* (Sf1), as in (43) and (48).

- (43) *siʔnoŋ pa ya-waʔ-e-nde?*
baby Q 3SG.DUB-go-ITER-ALREADY
‘Does the baby walk already?’

Grammatical aspect seems to be limited to perfective aspect (Sf3). The perfective suffixes are special in that they also code (non-)volitionality, that is, whether the A/S-participant is intentionally involved in the action referred to or not. They naturally occur on telic events (44). Volitional events are marked with -*an~-adn*, for example, the breaking event in (44). In (14), the same verb combines with -*on~-odn~-un~-udn* in a non-volitional event, with a patientive S-argument. Perfective aspect is also marked on stative predicates, which get a change-of-state interpretation (37c).

- b. *mba-tiaway-ndik* *õʔ-ẽ-ne* *men-ok=piʔ*
 V.PL-see-POT 1PL.INCL-be-IND any-period=INDET
 ‘We can see each other whenever/at any time.’

- (50) a. *ken on-mba-arak-me* b. *ken on-mba-arak-me waʔ-ta*
 3 3PL.IND-V.PL-kill-REC.PST 3 3PL.IND-V.PL-kill-REC.PST 3.REFL-ACC
 ‘They killed each other.’ ‘They killed themselves.’

Passive constructions, finally, consist of a finite form of *ẽʔẽ* ‘be’ and a nominalized verb form using prefix *e(ʔ)*- (51). They are only available to third person subjects; SAP patients require transitive (inverse) constructions.

- (51) *wenpu* *õʔ-ẽ-me* *e-mba-sayuŋ-ka*
 string.bag 3SG.IND-be-REC.PST NMLZ-V.PL-wet-make
 ‘The string bags were made wet.’

5.8 Spatial prefixes

Spatial prefixes specify locative or directional circumstances of (participants in) the event denoted by the verb. Examples include *ti-* in (42), which indicates location high up; *on-~n-* in (44), which signals the relation of ‘in’, ‘inside’ or ‘to’ (Tripp 1976a: 8); and *ok-~k-* in (52), which expresses separation (Tripp 1995: 219). They also serve to increase the valency of a verb (see Van linden 2022), as in (52): *ok-* promotes the person from whom the boys stole the manioc to a core argument registered on the verb.

- (52) *wambo-ta* *i-mba-uk-i* *tareʔ*
 boy-ACC 1SG-V.PL-search-1.IND manioc
- men-ok-mbere-me-niŋ-ta*
 3PL>1SG-SPAT:separation-steal-REC.PST-REL-ACC
 ‘I am looking for the boys that stole manioc from me.’

5.9 Noun incorporation

Harakmbut shows all four types of noun incorporation identified in Mithun (1984). Incorporated noun stems occupy Pf4. They are typically morphologically bound nouns; one exception is the free noun (*h*)*ak* ‘house’, as in (53) and (67), which is restricted to type I noun incorporation. I will argue that only elements that occur in type IV noun incorporation are true classifiers.

Type I noun incorporation or lexical compounding serves to create new lexemes for “name-worthy” activities (Mithun 1984: 848), and derives intransitive predicates from transitive ones. An example is given in (53): the transitive verb stem *-yoŋ* is combined with the free noun (*h*)*ak* ‘house’ to yield an intransitive verb that denotes a name-worthy activity of hunters, that is, the destruction of their waiting huts. The incorporated noun bears the semantic relationship of patient to its host verb. An example with a body part incorporated noun is given in (22).

- (53) *wa-mationka-eri* *o-ak-yoŋ-me*
 NMLZ-hunt-AN 3SG.IND-house-destroy-REC.PST
 ‘The hunter hut-destroyed.’

6.2 Alignment system

In the Harakmbut system of argument marking on dependents, the three argument roles (S, A, and O) show differential or optional marking. O-participants show animacy-based differential marking. Human and higher-order animate O-participants are accusative-marked with *-ta(h)*, whereas inanimate and lower-order animate O-participants are left unmarked (see [52] for both types). Accusative case is marked on patient-like arguments in transitive clauses, as in (7) and (9), as well as recipient-like arguments in ditransitive clauses, as in (16), (26), and (31). This parallels the primary object system in head marking (Section 5.2). However, (applied) R-participants also take beneficiary case marking (39).

Differential A-marking is governed by both animacy and focus. Non-focal animate A-participants tend to go unmarked, as in (6), (24), and (56). Inanimate A-participants, by contrast, typically carry the case suffix *-a* analyzed as nominative in earlier work (54), and so do animate A-participants that are in argument focus (e.g., question-answer pairs), or in focus within the broader discourse context (cf. Fauconnier 2011), as in (57): the first clause features a boy as A-participant, just like the clause preceding (57), while the second clause shows a switch in A-participant, which gets marked with *-a*.

- (57) *āñĩ* *pera* *o-ta-ma-nda-mbere?* *āñĩ* *pero*
 FILLER pear(Sp) 3SG.IND-APPL-V.PL-CLF:fruit-steal FILLER but(Sp)
- tiaway-we* *ōʔ-ē* *mboerek-a* *no-kot-we*
 see-NEG 3SG.IND-be man-NOM vital.center-fall-NEG
 ‘Um, he [i.e., the boy] is stealing his [i.e., the man’s] pears, um, but the man does not see it; he [i.e., the man] does not realize it.’ (spontaneous speech: Pear story)

S-participants typically go unmarked, whether they refer to human, as in (20) and (46), or inanimate participants, as in (36) and (51), and irrespective of their thematic role. Very rarely, overt marking with *-a* is used, for instance on an inanimate S (42), highlighting both the agentivity and the unexpectedness of the subject (McGregor 2007).

While earlier work maintains that the dependent marking system shows nominative-accusative alignment (Helberg 1984; Tripp 1995), the discussion above suggests that it is better characterized as a tripartite system, featuring both differential and optional marking.¹² Turning to the head marking system, Section 5.2 showed that Harakmbut basically shows nominative-accusative alignment, with hierarchical effects interfering in local and inverse scenarios.

6.3 Negation

Standard negation is expressed by a periphrastic construction in which the lexical verb base is suffixed by *-we* and immediately followed by a finite form of auxiliary *ēʔē* ‘be’ (cf. Tripp 1995: 218). TAME and argument marking occur on the auxiliary, as in (24) and (57). Negative existentials contain the negative particle *ewe*, which also serves as the negative response item ‘no’, and a finite form of *ēʔē* ‘be’. Negation with the privative suffix *-mbayo* was illustrated in (11) above.

¹² Accordingly, it might be better to gloss *-a* as ergative rather than nominative, also in view of (i) optionality being more common for ergative than nominative marking (McGregor 2010), (ii) the constraints on marking of S (McGregor 2007), and (iii) the polysemy of instrumental/ergative markers observed elsewhere (cf. Blake 1977: 51).

Negation in imperative/hortative sentences is quite different. Second person prohibitives are not formed with *-we*, and show distinct marking depending on the rank of O on the referential hierarchy (58). Third person prohibitives, by contrast, use the standard negation pattern with the auxiliary showing the imperative person affixes in Table 14, or use apprehensive forms.

- (58) a. *o-arak-pete* b. *i-arak-kate*
 2>1-hit;kill-2>1.PROH 2SG-hit;kill-2.PROH
 ‘Don’t hit me!’ ‘Don’t hit him!’

7 Clause-linking

7.1 Relative relations

Given that many of the properties of Harakmbut relative clauses vary in terms of the function of the relativized noun phrase (NP_{rel}), that is, the referent of the noun phrase whose reference is delimited (NP_{mat}) in the relative clause (S_{rel}), this section is organized in terms of these functions (terminology from Andrews 2007). When the relativized noun phrase functions as a subject, two formal strategies are available for the relative clause: suffixation of finite verb forms by the relativizing suffix *-niŋ* (52), and agentive deverbal nominalization with *wa(?)*- and *-eri*, as in (59) (see Section 4.8, and Van linden 2019).

- (59) *arakmbut-ta iʔ-uk-i wenpu wa-mba-ka-eri-ta*
 person-ACC 1SG-search-1.IND string.bag NMLZ-V.PL-make-AN-ACC
 ‘I am looking for the person who makes string bags.’

In both (52) and (59), the verb form carries case marking signalling the function of NP_{mat}, and S_{rel} appears outside of NP_{mat}, i.e., as a right-adjoined relative clause. There is also an alternative construal of (52) showing an external embedded relative clause. All strategies show omission of NP_{rel}.

Omission of NP_{rel} is also required when NP_{rel} functions as O (60), but S_{rel} only features suffixation of its verb form by *-niŋ*. Example (60) indicates that this strategy has verb-like internal syntax, with A marked for nominative case. It also suggests why A is expressed overtly at all: the *-niŋ* suffix competes with the mood+agreement suffixes for Sf7, so that in the absence of the overt pronoun, A can be understood to have a third person plural referent (Table 13).

- (60) *iʔ-uk-i siro opudn-a on-yok-me[*-ne]-niŋ*
 1SG-search-1.IND machete 2PL-NOM 1<>2PL-give-REC.PST-REL
 ‘I am looking for the machete that you (PL) gave me’

Thirdly, when NP_{rel} functions as an oblique participant (locative or instrumental), NP_{mat} is immediately followed by the pronoun *ken*, which introduces S_{rel} (61) and can be analyzed as a resumptive pronoun. In addition, locative NPs_{rel} also use structures in which the deictic locative adverb *kēyōn* ‘thither’ functions as a nominal domain following rather than preceding the relative clause (36); these involve omission of NP_{rel}. Instrumental NPs_{rel} have another strategy at their disposal as well, which involves instrumental deverbal nominalization through prefix *wa(?)*- (see Section 4.8) and omission of NP_{rel}, similarly to (59).

(61) *mboerek ã-ĩrĩŋ-ã-mẽ widn ken toto-ta*
 man 3SG.IND-hide-TRNS-REC.PST stone 3/DIST evil.spirit-ACC

o-arak-me-niŋ
 3SG.IND-kill-REC.PST-REL
 ‘The man hid the stone with which he killed the evil spirit.’

7.2 Complement relations

Harakmbut uses two main formal types of clauses to express a fairly wide range of complement relations. Table 15 indicates which semantic type of complement-taking predicates (CTPs) use which formal type (classification based on Noonan 2007), and includes reference to examples given. None of the CTPs uses overt complementizers.

Table 15: Formal types of complement clauses with their semantic types of CTP

Sentence-like complement	Nominalization in <i>e(?)</i> -	Other
– Utterance CTP (63), (64)	– Commentative CTP (unmarked INF)	– Predicates of fearing: adverbial relation of reason
– Propositional attitude CTP (62)	– Acquired ability CTP (unmarked INF)	– Immediate perception CTP: relative clause when emphasis on deliberate perception of <i>how</i> the perceived event proceeds
– Knowledge/acquisition of knowledge (KAK) CTP	– Immediate perception CTP (unmarked INF)	
– Desiderative CTP: <i>hope</i> -class	– Desiderative CTP: <i>want</i> -class (65) (INF marked for ACC case)	
– Manipulative CTP (44)		
– Immediate perception CTP: only when A (CTP) = O (complement clause) = 1SG		

Firstly, all sentence-like (S-like) complements involve a shift from the current speaker’s deictic centre to that of the represented speaker, cognizer, or experiencer (62), unless the current speaker is involved in the complement clause, as S or O (63). This deictic shift is especially noticeable in contexts where the participants in main and complement clause are third persons relative to the current speaker and coreference obtains between A of the CTP and O of the complement clause, as in (62): the dependent verb form shows verbal argument marking specific to the 3SG>1/2SG inverse scenarios, rather than non-local forms (see Section 5.2).

(62) *mbokerek ð-nð-pð-ẽ-mẽ-tẽ*
 man 3SG.IND-vital.center-CLF:round-be-REC.PST-NFIRSTH

mbe-arak-apo-ne-a *apetpet-a*
 3SG>1/2SG-kill-FUT-IND-QUOT jaguar-NOM
 ‘The man thought the jaguar was going to kill him.’ (lit. ‘The man thought: “The jaguar is going to kill me.”’)

Verb forms in complements of propositional attitude CTPs carry the quotative suffix *-a* (62), just like those of utterance predicates in contexts of indirect speech representation (63).¹³ Example (64) presents the direct speech counterpart of (63); such structures are never marked by quotative *-a*. Note that *-a* also is a verbal root, meaning ‘say’, as in (63) and (64), which is most probably its diachronic source.

¹³ Other CTP types with S-like complements use *-a* less consistently.

(63) *mboerek me-n-a-me-ne* *mbe-arak-apo-ne-a*
 man 3SG>1/2SG-SPAT:on-say-REC.PST-IND 3SG>1/2SG-kill-FUT-IND-QUOT

ndumba-yo
 forest-LOC

‘The man told me he was going to kill me in the forest’

(64) *mboerek oʔ-a-me* *o-arak-apo-ne* *ndumba-yo*
 man 3SG.IND-say-REC.PST 1<>2SG-kill-FUT-IND forest-LOC
 ‘The man said: “I am going to kill you (SG) in the forest.”’

In addition to S-like complements, complement relations are also expressed by nominalized verb forms prefixed by *e(?)*-. Complements of the *want*-class stand out in that the nominalizations invariably carry accusative marking (65) (cf. Tripp 1976b:3; 1995: 216; Helberg 1984: 451–452) in spite of being inanimate (Section 6.2). This indicates that their external syntax is noun-like (but special in terms of O-marking); their internal syntax, however, is verb-like; compare *opudnta* in (65). The notional subject of the nominalization in (65) is coreferential with the matrix subject, and has been equi-deleted (see Van linden 2019).

(65) *mbuttinda e-ma-n-a-ta* *ih-pak-i* *opudn-ta*
 truth NMLZ-V.PL-SPAT:on-say-ACC 1SG-want-1.IND 2.PL-ACC
 ‘I want to tell you (PL) the truth.’

7.3 Adverbial relations

As for complement relations, Harakmbut also uses nominalization in *e(?)*- for the expression of adverbial relations, with suffixes specifying the semantic subtype (Van linden 2019). In addition, some adverbial relations use relative clauses with deictic adverbials, or clauses whose verb form signals dependency; see Table 16, which includes reference to examples given.

Table 16: Formal types of adverbial clauses, with their semantic types of adverbial relations

Nominalization in <i>e(?)</i> - suffix	relation	Relativization	Dependent verb in <i>-po</i>	Dependent verb in <i>-nok</i>
<i>-te</i> (LOC)	Time: different subject	– Reason: with deictic <i>kente</i> ‘there’ (same subject)	– General adverbial (IND mood) (12), (47), (55), (67)	– Reason
<i>-(a)nda</i>	Time: same subject (66)	– Manner: with deictic <i>kenpa</i> ‘like this’	– Purpose (IMP/PROH mood) (45), (68)	
<i>-tanda</i>	Concession (46)	– Location: with deictic <i>kēyōn</i> ‘thither’ (36)		
<i>-nāyō</i>	Condition			

Like the nominalizations coding complement relations, those coding adverbial relations also combine noun-like external syntax with verb-like internal syntax. In (66), the notional participants of the nominalization are case-marked like main clause participants would be (see

Section 6.2). The suffix *-(a)nda* in (66) signals a temporal relation of simultaneity between events with shared subjects.

- (66) *on-a oroʔ-ta eʔ-uk-anda tiaway-we ɔʔ-ẽ-ne*
 2SG-NOM 1PL-ACC NMLZ-search-SIMUL.SS see-NEG 1<>2SG-be-IND
 ‘While you (SG) are looking for us, you (SG) don’t find us.’

A second formal strategy involves relativization with deictic adverbials following the relative clause, as described for *kẽyõn* (36) in Section 7.1. Thirdly, person-marked verb forms can receive suffixes that indicate the dependent status of the clause. Suffixation by *-po* is very frequent, especially in narratives. The relationship it establishes can be interpreted in various ways (e.g., time, reason; Tripp 1976c). As shown in (67), *-po* occurs in Sf7 just like *-niŋ* REL: the *o*-prefix does not index a third person singular S-participant like in (55), but a first person plural S-participant (the distinctive agreement suffix is absent because of *-po*).

- (67) *ken-taʔ mbaysik o-tiak-po hak-yo o-tiak-me-ne*
 DIST-LOC dusk 1PL-come-DEP house-LOC 1PL.INCL-come-REC.PST-IND

a_las_sinko-taʔ ɔ-kỹẽ-mẽ-nẽ in hak-yo
 at.five(Sp)-LOC 1PL.INCL-arrive.from.trip-REC.PST-IND PROX house-LOC
 ‘Then we came home at dusk, we came at five o’clock, we arrived in this house.’
 (spontaneous speech: anecdote)

Purposive relations stand out in that they use the imperative rather than indicative paradigm in *-po*-clauses, and same-subject purposives invariably index first person singular subjects (45) (cf. Tripp 1976c: 4). What is puzzling is that these verb forms retain the mood+agreement suffix (Sf7), with the quotative suffix *-a* still preceding *-po* (45), unless the imperative suffix is *-eʔ* in different-subject purposives (68). Similarly, negative purposives use verb forms with the apprehensive suffix *-apey*, followed in turn by quotative *-a* and *-po*.

- (68) *arakmbut-a o-n-ka kumo kã-mã-õrõk-e-po*
 person-NOM 3SG.IND-SPAT:on- barbasco 3SG.IMP-V.PL-go.out-3.IMP-DEP
 do

mbiʔigŋ kuru-te
 fish surface-LOC
 ‘The people use barbasco so that the fish come (out) at the (water) surface.’

8 Conclusion

This chapter has presented a grammatical description of Harakmbut, more specifically the variety known as Amaraeri or Arakmbut (the preferred autonym). Most of the data are drawn from elicitation, so the description will benefit from future work on spontaneously produced narratives, for example on basic constituent order, alignment, information structure, and discourse. Lesser understood topics within phonetics and phonology include nasality and stress assignment. In addition, more research is needed on the other varieties of Harakmbut.

Typologically interesting features include the presence of two classes of common nouns, potentially free nouns and obligatorily bound ones, which show differences in terms of noun modification, noun incorporation and, word formation. When combined with adnominal modifiers that obligatorily precede the nominal head in continuous noun phrases, for example, free nouns show a single construction type, while bound nouns show two: (i) one in which

they attach to a nominalizing prefix and follow the modifier, just like free nouns do, and (ii) one in which they attach to their modifier and form one phonological word with it. The distinct morpho-syntactic behaviour of bound nouns is unlike what has been described for bound nouns in, for instance, the Arawakan language Mojeño Trinitario (Rose 2015). Taking into account data from noun incorporation as well, I have argued that only the set of bound nouns showing classificatory noun incorporation (type IV in Mithun 1984) should be analyzed as classifiers, which does not exhaust the whole class of bound nouns (*pace* Payne 1987).

Another interesting feature relates to argument marking. Head marking involves hierarchical indexation resulting in a scenario-based split (without direction marking), based on the position of the O-participant on the person hierarchy 1/2 >3. In local scenarios, there is substantial referential opacity, with a number of pragmatic skewing strategies (Heath 1998) at work. In non-local and direct scenarios, the scenario-based split amounts to accusative alignment (A>3-markers = S-markers). Dependent marking had been analyzed as a nominative-accusative system in earlier work. However, the constraints on the marking of intransitive subjects suggest that it is better analyzed as a tripartite system, showing differential O-marking and both differential and optional A-marking.

Acknowledgments

The research reported on in this paper 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 a research project grant from the research council of the University of Leuven (KU Leuven) (GOA/12/007). I thank both editors and an anonymous referee for helpful comments on earlier drafts, as well as Olga Krasnoukhova, Françoise Rose, and Leo Wetzels. Special thanks go to Lev Michael for his help on the phonology and phonetics section. Of course, any errors of fact or interpretation remain my own responsibility. Finally, my sincere thanks go to the Harakmbut people, who warmly welcomed me in their communities and patiently taught me their beautiful language.

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.
- Aikhenvald, Alexandra Y. 2000. *Classifiers. A typology of noun categorization devices*. Oxford: Oxford University Press.
- Aikman, Sheila. 2009. The contradictory languages of fishing and gold panning in the Peruvian Amazon. *Maritime Studies (MAST)* 8. 50–71.
- Andrews, Avery D. 2007. Relative clauses. In Timothy Shopen (ed.), *Language typology and syntactic description. Volume II: Complex constructions*, 206–236. Cambridge: Cambridge University Press.
- Aza, José Pío. 1936. *Vocabulario español-arasairi*. Lima: Sanmarti y Cia.

- Bickel, Balthasar & Johanna Nichols. 2007. Inflectional morphology. In Timothy Shopen (ed.), *Language typology and syntactic description. Volume III: Grammatical categories and the lexicon*, 169–240. Cambridge: Cambridge University Press.
- Blake, Barry. 1977. *Case-marking in Australian languages*. Canberra: AIAS Press.
- Comrie, Bernard & Sandra A. Thompson. 2007. Lexical nominalization. In Timothy Shopen (ed.), *Language typology and syntactic description. Volume III: Grammatical categories and the lexicon*, 334–381. Cambridge: Cambridge University Press.
- Corbett, Greville. 2000. *Number*. Cambridge: Cambridge University Press.
- Crevels, Mily & Hein van der Voort. 2008. The Guaporé-Mamoré region as a linguistic area. In Pieter Muysken (ed.), *From linguistic areas to areal linguistics*, 151–179. Amsterdam: Benjamins.
- Dryer, Matthew S. 2007. Noun phrase structure. In Timothy Shopen (ed.), *Language typology and syntactic description. Volume II: Complex constructions*, 151–205. Cambridge: Cambridge University Press.
- Fauconnier, Stefanie. 2011. Differential agent marking and animacy. *Lingua* 121. 533–547.
- Gray, Andrew. 1996. *The Arakmbut: Mythology, spirituality and history*. Oxford: Berghahn.
- Gray, Andrew. 1997a. *The last shaman*. Oxford: Berghahn.
- Gray, Andrew. 1997b. *Indigenous rights and development*. Oxford: Berghahn.
- Guillaume, Antoine. 2006. La catégorie du ‘mouvement associé’ en cavineña: apport à une typologie de l’encodage du mouvement et de la trajectoire, *Bulletin de la Société de Linguistique de Paris* 101(1). 415–436.
- Guillaume, Antoine. 2008. *A grammar of Cavineña*. Berlin & New York: Mouton de Gruyter.
- Guillaume, Antoine. 2016. Associated motion in South America: Typological and areal perspectives. *Linguistic Typology* 20(1). 81–177.
- Hart, Raymond. 1963. Semantic components of shape in Amarakaeri grammar. *Anthropological Linguistics* 5(9). 1–7.
- Heath, Jeffrey. 1998. Pragmatic skewing in 1 ↔ 2 pronominal combinations in Native American languages. *International Journal of American Linguistics* 64(2). 83–104.
- 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.
- Helberg Chávez, Heinrich Albert. 1996. *Mbaisik: En la penumbra del atardecer. Literatura oral del pueblo harakmbut*. Lima: Centro Amazónico de la Antropología y Aplicación Práctica.
- Helberg Chávez, Heinrich Albert. in prep. *Gramática pedagógica harakmbut*.
- 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 Rosalía Avals de Matos & Rogger Ravines eds.), *XXXIX Congreso Internacional de Americanistas, Lima 1970. Actas y Memorias*, vol. 5, 185–207. Lima: IEP.
- Manqueriapa Vitente, Alberto, José Antonio Dumas Ramos, Marisabel Dumas Ramos, Julián Dariquebe Jerehua, Walter Quertehuari Dariquebe, Nicolás Barrio Rivas, Natalia Barrio Visse, Miriam Shimbo Vera, Marleny Rodríguez Agüero, Rosa Sihui Erembuyo, Manuel Manuaje José Tije & Marcia Tije Capi. 2012. *Daka wamandoya harakmbut ate* [Writing correctly in the Harakmbut language]. Lima: Ministerio de Educación.

- Matteson, Esther. 1972. Proto Arawakan. In Esther Matteson, Alva Wheeler, Frances L. Jackson, Nathan E. Waltz & Diana R. Christian (eds.), *Comparative studies in Amerindian languages*, 160–242. The Hague: Mouton.
- McGregor, William B. 2007. Ergative marking of intransitive subjects in Warrwa. *Australian Journal of Linguistics* 27(2). 201–229.
- McGregor, William B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120. 1610–1636.
- McQuown, Norman A. 1955. Indigenous languages of Latin America. *American Anthropologist* 57. 501–570.
- Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60. 847–879.
- Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language typology and syntactic description. Volume II: Complex constructions*, 52–150. Cambridge: Cambridge University Press.
- Payne, Doris L. 1987. Noun classification in the western Amazon. *Language Sciences* 9(1). 21–44.
- Payne, David L. 1991. A classification of Maipuran (Arawakan) languages based on shared lexical retentions. In Desmond C. Derbyshire & Geoffrey K. Pullum (eds.), *Handbook of Amazonian languages*, vol. 3, 355–500. Berlin: Mouton de Gruyter.
- Peck, Charles. 1979 [1958]. Toyeri y sapiteri: Un informe preliminar de la fonología y el vocabulario. Yarinacocha: Ministerio de Educación & Instituto Lingüístico de Verano.
- Rose, Françoise. 2015. Mojeño Trinitario. In Mily Crevels & Pieter Muysken (eds.), *Lenguas de Bolivia. Volume III: Oriente*, 59–97. La Paz: Plural Editores.
- Rose, Françoise & An Van linden. 2022. The derivational use of classifiers in Western Amazonia. In Steve Pepper, Francesca Masini & Simone Mattioli (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>].
- San Roque, Lila, Lauren Gawne, Darja Hoenigman, Julia Colleen Miller, Alan Rumsey, Stef Spronck, Alice Carroll & Nicholas Evans. 2012. Getting the story straight: Language fieldwork using a narrative problem-solving task. *Language Documentation & Conservation* 6. 135–174.
- Shibatani, Masayoshi & Prashant Pardeshi. 2002. The causative continuum. In Masayoshi Shibatani (ed.), *The grammar of causation and interpersonal manipulation*, 85–126. Amsterdam: John Benjamins.
- Solís Fonseca, Gustavo. 2003. *Lenguas en la Amazonía peruana*. Lima: Visual Service S.R.L.
- Tovar, Antonio. 1961. *Catálogo de las lenguas de América del Sur*. Buenos Aires: Editorial Sudamericana.
- Tripp, Robert. 1976a. *Los verbos amarakaeri*. Lima: Instituto Lingüístico de Verano.
- Tripp, Robert. 1976b. *Sustantivos verbales y frases de sustantivos verbales en amarakaeri*. Lima: Instituto Lingüístico de Verano.
- Tripp, Robert. 1976c. *Las relaciones señaladas por -po y -nōk en amarakaeri*. Lima: Instituto Lingüístico de Verano.
- Tripp, Robert. 1995. *Diccionario amarakaeri-castellano*. Yarinacocha: Ministerio de Educación & Instituto Lingüístico de Verano.
- Van linden, An. 2019. Nominalization in Harakmbut. In Roberto Zariquiey, Masayoshi Shibatani & David W. Fleck (eds.), *Nominalization in the languages of the Americas*, 455–490. Amsterdam: John Benjamins. [<https://doi.org/10.1075/tsl.124.12lin>]
- Van linden, An. 2020. Constructional effects of indirect evidential marking in Harakmbut. *Functions of Language* 27(1): 7–28. [<https://doi.org/10.1075/fof.20004.lin>]

- Van linden, An. 2022. Spatial prefixes as applicatives in Harakmbut. In Sara Pacchiarotti & Fernando Zúñiga (eds.), *Applicative morphology: Neglected syntactic and non-syntactic functions* [Trends in Linguistics. Studies and Monographs (TiLSM) 373], 129-159. Berlin: De Gruyter. [<https://doi.org/10.1515/9783110777949-006>]
- 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.

Abbreviations

1	first person	IND	indicative
2	second person	INDET	indeterminate
3	third person	INFR	inferential
ACC	accusative	INS	instrumental
ADD	additive	ITER	iterative
ADV	adverb(ial)	LOC	locative
AM	associated motion	NEG	negation
AN	animate	NFIRSTH	non-firsthand evidential
APPL	applicative	NMLZ	nominalizer, nominalization
ASP	aspect	NOM	nominative
AVRT	avertive	NVOL	non-volitional
BEN	benefactive	PFV	perfective
CAUS.SOC	sociative causative	PL	plural
CLF	classifier	POT	potential
COLL	collective	PRIV	privative
COM	comitative	PROH	prohibitive
CONC	concessive	PROX	proximal/proximate
COND	conditional	Q	question particle/marker
DEP	dependent	QUOT	quotative
DIM	diminutive	REAS	reason
DIST	distal	REC.PST	recent past
DU	dual	REFL	reflexive
DUB	dubitative	REL	relativizer
EPIST	epistemic modal	REM.PST	remote past
EP.V	epenthetic vowel	REST	restrictive
EXCL	exclusive	SG	singular
FILLER	filler, word search	SIM	similative
FOC	focus	SIMUL	simultaneous
FUT	future	SPAT	spatial
GEN	genitive	SS	same subject
HAB	habitual	TRNS	transitivizer
IMP	imperative	VBZ	verbalizer
INCL	inclusive	VOL	volitional
INCORP.N	incorporated noun	V.PL	verbal plural