

This was published in Zariquiey et al. 2019

Please cite as:

Van linden, An. 2019. Nominalization in Harakmbut. In Roberto Zariquiey, Masayoshi Shibatani & David W. Fleck (eds.), *Nominalization in the Languages of the Americas* [Typological Studies in Language 124], 455-490. Amsterdam: John Benjamins.

<https://doi.org/10.1075/tsl.124.12lin>

WITH CORRIGENDA to the printed version indicated with track changes

Nominalization in Harakmbut

An Van linden (an.vanlinden@uliege.be)
University of Liège & KU Leuven

Abstract

This paper focusses on verb-based nominalization in Harakmbut (isolate, Peru), which falls into two formal types on the basis of the prefix used. The first type, using the nominalizing prefix *wa(?)*-, is restricted to participant nominalization and is predominantly used to produce nouns for NP-use. The second type, using the nominalizing prefix *e(?)*-, is mainly used for event nominalization and typically produces multi-word nominalizations. Depending on the constructions they occur in and additional suffixation they take, nominalizations with *e(?)*- can serve complementation as well as adverbial functions. Across the two formal types, multi-word nominalizations combine NP-like external syntax with verb-like internal syntax. The two nominalizing prefixes also serve a basic function in noun-based nominalization, lending independent status to obligatorily bound nouns.

1 The Harakmbut language and collection of data¹

Harakmbut is an underdescribed language from the Peruvian Amazon, spoken in a number of ‘native communities’ in the departamentos of Madre de Dios and Cusco. The communities are located on the Madre de Dios River and its upper tributaries, such as the Colorado River. Before I go into more detail about my own fieldwork, I will first summarize what has already been written about the language and its speakers.

The genetic affiliation of Harakmbut has been a topic of debate. The language has formerly been classified as an Arawak or Maipuran language by McQuown (1955) (see Hart 1963: 6) and Matteson (1972), but this has found little acceptance (Adelaar 2007: 39). Wise (1999: 307) states that Harakmbut is commonly accepted to be an isolate (cf. Dryer & Haspelmath 2013 in WALS). More recently, drawing on mainly lexical evidence, Adelaar (2000, 2007) has proposed that it is genetically related to the Brazilian Katukina family, which

¹ The research reported on in this paper has been made possible by mobility grants and a postdoctoral grant from the Research Foundation – Flanders (FWO), as well as a postdoctoral grant from the Fund for Scientific Research – FNRS and a research project grant from the research council of the University of Leuven (KU Leuven) (GOA/12/007). I thank an anonymous reviewer as well as the editors for their insightful comments on an earlier version of this paper. 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.

may be further linked to Macro-Jê. In addition, language contact should also be reckoned with, as Harakmbut exhibits a number of Western Amazonian grammatical features, as well as features characteristic of the Guaporé-Mamoré linguistic area in southwest Brazil and eastern Bolivia (Crevels & van der Voort 2008), close to the border with the Peruvian departamento of Madre de Dios, with one member language, Ese Ejja (Tacanan), also being spoken in Madre de Dios. Harakmbut has already been noted to share some grammatical features with Ese Ejja (Pozzi-Escot 1998: 93), and I also believe it shares features with other languages in the area like Cavineña and Kwaza, e.g. in the domains of aspect, associated motion, and, more pertinently to this paper, noun-based nominalization (see Section 5.3).

Previous discussions have also focussed on relations within the Harakmbut group. The main question is whether Harakmbut should be regarded as a single language with a number of dialectal variants or rather as a small language group or family consisting of distinct, related languages. The distinction of seven ethno-linguistic groups (Amarakaeri, Watipaeri, Arasaeri, Sapiteri, Kisambaeri, Pukirieri and Toyoeri) by the anthropologist Andrew Gray (1996: 7-9) might hint at the latter option, while linguists such as Helberg (1984, 1990), Wise (1999) and Adelaar (2007) agree on the former. The proposal that Harakmbut is in fact a single language is consistent with my language consultants' assessments. By now, the most vital varieties are the first two listed above. Of many of the other dialects only a handful of speakers – if any – are left, and very little information is available.

Earlier linguistic work on Harakmbut has mainly concentrated on the most vital dialect, i.e. Amarakaeri (Hart 1963; Helberg 1984, 1990; Tripp 1976, 1995). It should be noted that 'vital' is a relative term, as the number of speakers has been estimated at 1,000 by Moore (2007: 46), and I noted during my field stays that young parents are reluctant to pass on the language to their children, as it is felt to socially stigmatize them. Children are thus mainly brought up in Spanish, and acquire only a passive competence in Harakmbut. Young adults and speakers up to the age of fifty generally are bilingual in Harakmbut and Spanish. Speakers older than fifty are mainly monolingual in Harakmbut. My own fieldwork also focusses on the Amarakaeri dialect.² The data presented in this paper are drawn from audio recordings made in the native communities of Puerto Luz, San José del Karene and Shintuya, all with Amarakaeri informants, in the summers of 2010, 2011 and 2016. So far, I have mainly transcribed and analysed elicitation sessions with bilingual speakers, which implies that the bulk of the data used in Sections 3 to 5 does not represent spontaneous speech. If it does, this has been indicated in the example. The practical orthography used is IPA-based, and different from the community spelling.

2 Nominalization in Harakmbut

South American languages generally show a rich diversity of nominalization structures and functions. This also holds for Harakmbut, as this paper aims to show. Its main focus will be on verb-based nominalizations that do not feature person or mood marking. It is beyond the scope of this paper to discuss the nominalization of finite verb forms by the relativizing suffix *-niŋ* (or 'nominalizing' suffix, cf. Shibatani, this volume), which is detailed to some extent in Van linden (Forthc.).

² I would like to point out that the speakers of this variety regard the label *Amarakaeri* as a depreciating term; it is adapted from *wa-mba-arak-a-eri* (NMZR-VPL-kill-TRNS-AN), a verb-based nominalization meaning '(fierce) killer/murderer', 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 'Arak(m)but', as distinct from the Watipaeri variety, towards whose speakers they generally entertain feelings of enmity rather than brotherhood.

The data available in the literature (Tripp 1976, 1995; Helberg 1984) and my own fieldnotes indicate that (non-finite) verb-based nominalization in Harakmbut falls into two formal types, which can be distinguished on the basis of the prefix used. A first type uses the nominalizing prefix *wa(?)*-, illustrated in (1), while the second type features the nominalizing prefix *e(?)*-, exemplified in (2), which is also used in the citation form of verbs and other non-finite verb forms.³ I will show that these prefixes serve a basic function in noun-based nominalization as well. In the examples given, grammatical nominalizations (cf. Shibatani, this volume) are rendered between square brackets.

- (1) Jonas-tewapa o-niŋ-ka wa-wedn grĩngo-a
 Jonas-BEN 3SG.IND-BEN-make NMZR-lie foreigner-NOM
 ‘The foreigner makes a bed for Jonas.’
- (2) ndak õʔ-ẽ [e-mbaʔ-tiak, ãñĩ, keme]_{NMLZ}
 good 3SG.IND-be NMZR-VPL-come FILLER tapir
 ‘It is good that (, eh,) the tapirs have come.’

In (1), the prefix *wa-* attaches to the verb root *-wedn* ‘lie’ to form the noun ‘bed’. This resultant nominalized form functions as an argument participant, viz. it constitutes the direct object of the finite verb form *oniŋka*; it is left unmarked as is generally the case for inanimate O-participants (see Section 3). In (2), the nominalized form *embaʔtiak* consists of the nominalizing prefix *e-* and the verb base *-mbaʔtiak*. It serves as the verb phrase of the complement clause functioning as the subject of the commentative predicate *ndak õʔẽ* ‘is good’. Thus, while nominalization with *wa(?)*- derives a noun from a lexical verb and realizes participant nominalization in (1), nominalization with *e(?)*- yields an ‘action nominal’ (cf. Comrie & Thompson 2007: 343) from a predicate, containing also a noun phrase that corresponds to the subject of the verb stem (i.e. *keme*), and realizes event nominalization in (2). Both participant and event nominalization are common in South-American languages (Van Gijn et al. 2011: 10-13).

While (1) and (2) illustrate the predominant functions and uses of the two formal types of verb-based nominalization available in Harakmbut, they do not exhaust them. Specifically, nominalizations with *wa(?)*- are found to sometimes modify other nouns, in which function they are equivalent to relative clauses, having their own notional argument participants. Similarly, nominalization with *e(?)*- is not limited to deriving action nominals from predicates, since it is sometimes also used to derive participant nominalizations that function in the same manner as underived nouns (see (53) in Section 5.3 for an example). Together the two formal types of verb-based nominalization realize all three subtypes of subordinate clauses traditionally distinguished, with nominalizations with *wa(?)*- coding relative relations, while nominalizations with *e(?)*- are used to code both complement and adverbial relations.

The discussion is organized as follows. Section 3 discusses the basic features of Harakmbut grammar that are needed to analyse the nominalization data. Section 4 focusses on nominalization with *wa(?)*-, while Section 5 homes in on nominalization with *e(?)*-. Each section will discuss further subtypes of these formal types, with a focus on the internal and external syntax of the nominalized forms; it will be investigated to what extent these are verb-like or NP-like. It will also be examined to what degree the nominalized forms retain verbal categories and adopt nominal ones. With regard to the latter, it will be shown that both formal

³ The glottal stop has no phonemic value in Harakmbut (*pace* Helberg 1984: 22), but rather a suprasegmental one: it is optionally used to demarcate syllable boundaries when these lack consonantal onsets or codas (see Van Linden Forthc.).

types of nominalized forms use suffixes that are also used on underived nouns. Section 6, finally, recapitulates the major findings, and proposes some questions for further research.

3 Basic features of Harakmbut grammar

Before we delve into the analysis of nominalized forms, this section discusses some basic features of Harakmbut grammar that are crucial to determining their internal and external syntax. Specifically, it will concentrate on nominal categories, verbal categories and the coding of grammatical relations (based on Van linden Forthc.).

Harakmbut nouns can be marked for a number of categories; Table 1 visualizes the morphological template of the head of a noun phrase (NP). Table 2 presents the (rather extensive) set of case markers, two of which are polysemous with the instrumental function. Many of (the functions of) these markers have already been noted before by Helberg (1984: 436-444) and/or Tripp (1995: 194-200); this is also indicated in Table 2.

(complex) (pro)noun stem	Collective	Case	Focus ₁	Focus ₂
	<i>-(o)mey</i> COLL	(see Table 2)	<i>-nãÿõ</i> COND	<i>-nda</i> FOC
			<i>-yo</i> REST	

Table 1: Morphological template of the nominal head

Suffix	Case
<i>-ʔa~-a</i>	nominative ^{H, T}
	instrumental ^{H, T}
<i>-ere</i>	comitative ^{H, T}
	instrumental ^T
<i>-ta(h)</i>	accusative ^{H, T}
<i>-en~-edn~-wedn~-ʔedn</i>	genitive ^{H, T}
<i>-tewapa</i>	beneficiary ^T
<i>-(o)niŋ</i>	similative
<i>-apo</i>	reason
<i>-mbayo</i>	privative
<i>-yo, -ya, -taʔ, -te, -yon, -pen</i>	locative ^{H, T}

Table 2: Inventory of case suffixes (^H: Helberg (1984: 436-444); ^T: Tripp (1995: 194-200))

Furthermore, nouns lack the category of number and Harakmbut lacks articles, which would express definiteness or specificity. Instead, nouns pattern with a number of adnominal modifiers, like indefinite and demonstrative modifiers (e.g. *ken* in (4b)), as well as indefinite and cardinal quantifiers. The syntactic relation of adnominal possession is expressed by genitive case marking on the possessor (pro)noun, cf. (3).

- (3) *ndoʔ-edn siro*
1SG-GEN machete
'my machete'

Another type of adnominal construction that is pertinent to this paper (see Section 5.2.1) is modification by adjectives. In my data, adjectives appear in both continuous and discontinuous

NPs.⁴ In the first subtype – the only relevant one here – they occur in prenominal (4a) as well as postnominal position (4b).

- (4) (a) aʔ-yok-i sal uru-wettone-ta-nda
 1SG.IMP-give-1.IMP salt beautiful-woman-ACC-NDA
 ‘I (should) give salt to the beautiful woman.’
 (b) ih-yok-i sal ken wettone-tewapa uru-nda
 1SG-give-1.IND salt DIST woman-BEN beautiful-NDA
 ‘I give salt to that beautiful woman.’

(4a) and (4b) are translations of the same stimulus, but they show some interesting differences. For one, they show that R-participants of ditransitive events can receive either accusative marking (4a) or beneficiary marking (4b) (see below). A more important difference lies in the adjectival construction type. While in (4b) the adjective follows the head noun and the adjectival root is suffixed by *-nda*, in (4a) the adjectival root precedes the noun, and the *-nda* suffix is appended only after the case-marked noun. The NP in (4a) also shows phonological fusion; the stressed syllable nucleus is underlined. Both constructions feature the suffix *-nda*, whose function in (4) I am uncertain about, but it seems to be different from that in *Lupeanda* in (5) below, in which it is used as a focus marker appended to a nominative-marked noun (not modified by any adjective) (see Table 1). In (4), *-nda* does not mark focus; it seems to be required by the modification construction. In *mbiʔign-tonenda* in (5), it does not mark focus either; nouns suffixed by derivational affix *-tone* ‘adult, old, big’ are also found without suffix *-nda* in contexts similar to (5). I hypothesize that *-nda* basically serves to produce (independent) modifiers or to signal a modification relation (this hypothesis especially targets examples like (4) and infinitival nominalizations, see Sections 5.2.1 and 5.2.2).

- (5) Lupe-a-nda oʔ-te₋ŋ-me mbiʔign-tone-nda
 Lupe-NOM-FOC 3SG.IND-cut-REC.PST fish-big-NDA
 ‘Lupe herself cut the big fish.’

Harakmbut verbal morphology comprises inflectional as well as derivational categories. The former involve tense, (types of) aspect, mood, modality, evidentiality, and verbal argument marking. The latter include valency-changing categories like transitivizers and applicatives, as well as spatial elements and (types of) aspect. A number of these are illustrated in (6).

- (6) o-ma-niŋ-toʔ-tiak-me-te aypo
 3SG.IND-VPL-BEN-SOC-come-REC.PST-NVISINDIR.EVD food
 wa-mbet-ta Puerto-lus-yo
 NMZR-family-ACC Puerto-Luz-LOC
 ‘She took food to Puerto Luz for her family (members).’

The verb stem in (6) is intransitive (*-tiak* ‘come’), but its valency is increased by the sociative causative prefix *toʔ-*, which promotes *aypo* ‘food’ to direct object status, and by the benefactive applicative *niŋ-*, which promotes *wambet* ‘family’ to beneficiary status, receiving accusative marking. These derivational prefixes also appear on nominalized verb forms (e.g. (26), (35)). In (47), even the (inflectional) tense marker *-me* coding recent past, like in (5)-(6), is retained in the nominalized form.

⁴ I am not sure whether NPs whose elements are not adjacent are ‘merely’ discontinuous or rather appositional.

Verbal plural marking by *mba~ma~mã*- (phonologically conditioned allomorphs) is also retained in nominalized forms. This category serves to signal plurality of the action denoted by the verb or plurality of participants engaged in the event. In the latter function it works ergatively, indicating plurality of the S-participant in intransitive events (cf. nominalized form in (40)) and of the (applied) O-participant in (extended) transitive events, like in (6), which need not be expressed by external NPs. In (6), the A-participant brought food for more than one family member. In (12), the verbal plural marker indicates that the action of the nominalized verb is performed several times.

Furthermore, Harakmbut verbs show all four types of noun incorporation identified in Mithun (1984). It will become clear that incorporated nouns (types I to III) and verbal classifiers (type IV) are retained in nominalized forms, e.g. (19), (31), and (51).

Finally, grammatical relations are reflected by both head and dependent marking. The head marking system involves hierarchical indexation resulting in a configuration-based split (without direction marking), based on the position of the patient participant on the person hierarchy 1/2>3: while third-person patients are never indexed, speech act participant patients trigger distinct relational prefixes, viz. portmanteau prefixes indexing both agent and patient. This split amounts to accusative alignment in non-local configurations (involving a third person acting on another third person) and direct configurations (involving a speech act participant acting on a third person), as agents acting on third person patients on the one hand and the sole participants of intransitive clauses on the other are cross-referenced on the verb by the same set of prefixes (A>3-markers = S-markers).

The dependent marking system is different, but no less complex, as the three argument roles (S, A and O) show differential or optional marking in independent clauses (case vs. zero exponence). The marking of O-participants is animacy-based. Human and higher order animate Patient-like arguments carry accusative case marking (e.g. *Lisbet-ta* in (8)), while inanimate and lower order animate Os go unmarked (e.g. *mbi?igŋtonenda* in (5) and *aypo* in (6)). As illustrated in (4a) and (6), accusative case is also marked on human Recipient-like arguments in (applied) ditransitive clauses. However, (applied) R-participants are also found to carry beneficiary case marking (see (4b)).

The marking of A-participants is governed by both animacy and focus. Non-focal animate As are typically left unmarked, e.g. *mboerek* in (7), while inanimate As are marked, specifically by a case suffix analysed as nominative by Helberg (1984) and Tripp (1995), e.g. *kurudn-a* in (8). This type of differential A-marking is cross-linguistically recurrent (cf. Fauconnier 2011).

- (7) *sik-yo-edn-nda* *ãñĩ*, *mboerek* *o-n-ka*,
 black-LOC-?-NDA FILLER man 3SG.IND-SPAT-do
ãñĩ, [...] *pera*
 FILLER pear
 ‘Early in the morning, eh, a man is picking, eh, pears.’ (spontaneous speech)
- (8) *kurudn-a* *o-seŋ-pak-a* *Lisbet-ta*
 thunder- NOM 3SG.IND-crazy-VBZ-TRNS Lisbeth-ACC
 ‘The thunder drives Lisbeth crazy.’

Animate A-participants that are in argument focus also tend to go marked, e.g. *Lupe-a-nda* in (5), just like As that are in focus within the broader discourse context, cf. (9). Like (7) and (10)-(11) below, (9) is taken from my Pear story data, and is assumed to represent spontaneous

speech. In the first (dependent) clause, the A-argument is the boy who stole the pears (see (10)), while in the next one, there is a switch in A-participant to *muneyosi?po-a*.⁵

- (9) o-k-to-wa-po bisikleta-te; ken ãnĩ
 3SG.IND-SEPARATION-SOC-go-DEP bicycle-LOC then FILLER
 o-ndeh ãnĩ muneyo-si?po-a, ãnĩ,
 3SG.IND-meet FILLER girl-DIM-NOM FILLER
 ndaŋ-no-po-te
 path-(vital centre-CLF:round)_{middle}-LOC
 ‘He goes away with them [i.e. pears] on his bike; then, eh, a little girl crosses him, eh, in the middle of the road.’ (spontaneous speech)

S-participants are typically left unmarked, whether they have human referents, e.g. *wasi?po* in (10), or inanimate ones, e.g. *widn* in (11). Only very rarely are S-participants marked by ‘nominative’ case.

- (10) ken wa-si?po o-tiak-po ãnĩ
 then NMZR-(peel-CLF:round)_{child} 3SG.IND-come-DEP FILLER
 bisikleta-te ãnĩ o-ta-mbere ãnĩ kanasta
 bicycle-LOC FILLER 3SG.IND-APPL-steal FILLER basket
 ‘Then a child comes, eh, on his bike, eh, and he steals his [i.e. the pear picker’s], eh, basket.’ (spontaneous speech)
- (11) ken adnte ãnĩ o?-wedn ãnĩ widn, widn-tone-nda
 then far.away FILLER 3SG.IND-lie FILLER stone stone-big-NDA
 ‘Then, further down, eh, there lies, eh, a stone, a big stone.’ (spontaneous speech)

While the Harakmbut case marking system has been analysed as showing nominative-accusative alignment in earlier work (Helberg 1984; Tripp 1995), the observed patterns of optional A- and S-marking point to **a tripartite an optional ergative accusative** system of alignment, in which formal marking of S is highly constrained (cf. McGregor 2007, 2010) (but in this paper I have not yet adapted glossing of -a accordingly).

4 Nominalization with prefix *wa(?)*-

The first formal type of nominalization in Harakmbut discussed here is characterized by affixation of the nominalizing prefix *wa(?)*- to the verb stem. It is functionally restricted to participant nominalization, and it mainly produces heads of NPs that can occur in any participant slot in the higher clause. In addition, it also yields forms that show modification uses. Within this first formal type, a further distinction can be made on the basis of the presence of an additional suffix, viz. *-eri* ‘animate’. While *wa(?)*-nominalizations suffixed by *-eri* refer to animate entities, *wa(?)*-nominalizations without *-eri* have inanimate referents. The first subtype invariably involves agentive nominalization, while the second one realizes instrumental or objective nominalization (see Comrie & Thompson 2007).

⁵ In addition, the girl is newly introduced in (9). However, first-mention use does not suffice to trigger case marking on animate As; (7) instantiates the first mention of the pear-picker participant in the story, but here A is left unmarked.

4.1 Animate referents

Harakmbut morphology caters for a derivational process whereby verbs can be made into nouns denoting an animate entity that can be described as ‘one who “verbs”’. This type of participant nominalization is traditionally termed ‘agentive nominalization’ (cf. Comrie & Thompson 2007: 336). In Harakmbut, it involves affixation of the nominalizing prefix *wa(?)*- in conjunction with the suffix *-eri* ‘animate’ to the verb stem (see also Tripp (1976: 1), who labels *wa-* (<hua->) as a classifier). Examples are in (12) and (13).

- (12) *wa-mba-yorok-eri*
 NMZR-VPL-dream-AN
 ‘dreamer’
- (13) *wa-mationka-eri*
 NMZR-hunt-AN
 ‘hunter/one who hunted/one who hunts’

As can be gathered from these examples, the resultant noun need not be in an ‘Agent’ relationship with the verb it is derived from (cf. Comrie & Thompson 2007: 336). In (12), the noun is in an Experiencer relationship with the verb ‘dream’. Also, (12) retains verbal plural marking, which indicates here plurality of action; in Harakmbut culture a dreamer is a shaman-like figure who regularly receives dreams from the spirits.

Agentive nominalizations are found to serve two functions. Firstly, they can serve as head of an NP realizing any type of argument function (cf. Tripp 1976: 2; 1995: 194). In (14), for example, the form *wamationkaeri-ta* functions as direct object of the verb *oketea*, as signalled by the accusative case marker *-ta*.

- (14) *apetpet-a o-ket-e-a wa-mationka-eri-ta*
 jaguar-NOM 3SG.IND-run-ITER-TRNS NMZR-hunt-AN-ACC
 ‘The jaguar makes the hunter run.’

Secondly, agentive nominalizations can also be used to modify other nouns, and are thus functionally equivalent to relative clauses (this function is not described by Tripp 1976, nor Helberg 1984).⁶ In (15) the phrase *henpu wambakaerita* is functionally equivalent to a right-adjoined relative clause, restricting the reference of the head noun it modifies, viz. *arakmbutta* (cf. Andrews 2007: 214-217); note that basic word order in Harakmbut is (not strictly) OVS (Van linden Forthc.). In (16), the form *wamanokoteri* functions as a headless relative clause (lacking a nominal head like *arakmbutta* in (15)). It thus shows NP-use of a grammatical nominalization; it realizes a non-referential NP functioning as complement to the subject in a predicational copular clause.

- (15) *arakmbut-ta iʔ-uk-i [henpu*
 person-ACC 1SG-search-1.IND string.bag
*wa-mba-ka-eri-ta]*_{NMLZ}
 NMZR-VPL-make-AN-ACC
 ‘I am looking for the person who makes string bags.’

⁶ The main relativization strategy in Harakmbut involves suffixation of the finite verb form of the relative clause by *-niŋ* (Van linden Forthc.).

- (16) wa-mba-yorok-eri òʔ-ẽ
 NMZR-VPL-dream-AN 3SG.IND-be
 [wa-ma-no-kot-eri]_{NMLZ}
 NMZR-VPL-(vital.centre-fall)_{realize}-AN
 ‘The dreamer is one who knows many things.’

Example (15) is the most interesting one with regard to the external and internal syntax of agentive nominalization. In (15), the nominalized form is marked for accusative case. This use of the nominal category of case suggests that the action nominal has an NP-like external syntax. The notional direct object of the nominalized form (*henpu*) goes unmarked, just like inanimate direct objects in independent clauses (see Section 3). In addition, the nominalizations in (15) and (16) both retain verbal plural marking. This suggests that the internal syntax of agentive nominalizations is verb-like rather than NP-like; however, I have no examples with animate notional direct objects to bolster the argument.

Finally, it should be noted that the suffix *-eri* is also used in a productive noun-to-noun derivational process:⁷ 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 (17a) or a proper noun (17b). The derived nouns often serve as demonyms or gentilics (see also Tripp 1995: 193). In (17b), for example, the suffix is added to the name of the native community of Puerto Luz to denote its members. Note that the names of the Harakmbut ethno-linguistic groups mentioned in Section 1 also end in *-eri*.

- (17)(a) pato numba-eri
 duck forest-AN
 ‘duck, one that is in the forest/bush’ [Sp. *pato de monte*; Lat. *Sarkidiornis melonotos*]
 (b) Porto-lus-eri
 Puerto-Luz-AN
 ‘the people from Puerto Luz’

4.2 Inanimate referents

The second type of verb-based nominalization with *wa(?)*- uses no further marking, and produces nouns that refer to inanimate entities. These nouns can bear two different relationships to the verb they are derived from. In one type, the resultant nouns denote the instrument for “verbing”, and in the other, they denote the result or the typical or ‘cognate’ object of the action designated by the verb. The associated processes are traditionally called ‘instrumental’ and ‘objective’ nominalization respectively (cf. Comrie & Thompson 2007: 338-342). This semantic distinction does not correlate with a formal one.

The derivational process of forming “non-personal” nouns by adding the nominalizing prefix *wa(?)*- to a verb stem has also been noted by Tripp (1976: 1). His examples all constitute instrumental nominalizations, although his paraphrase for (18) reads “thing that discharges, i.e. a rifle” (1976: 2). Another example is (19). Note that (19) shows noun incorporation of type I; ‘head-dressing’ qualifies as a “name-worthy” activity (Mithun 1984: 849). However, this type of lexical compounding does not make the nominalized verb more ‘finite’.

⁷ Helberg (1984: 445) discusses the *-eri* suffix in a section on noun-based nominalization only, describing its meaning as ‘group of persons’; yet his examples include both verb-based and noun-based nominalizations. Example (17a) shows that *-eri* is not restricted to humans, but applies more generally to animate entities.

- (18) wa-potoŋ
 NMZR-discharge
 ‘a rifle/something for the purpose of discharging or firing’ (cf. Tripp 1976: 2)
- (19) wa-ku-ot
 NMZR-head-get.dressed
 ‘a hat/something to dress your head with’

Like agentive nominalizations, instrumental nominalizations can serve two different functions. In (20), for example, *wawedn* functions as the head of an NP realizing the direct object of the verb form *oniŋka* (see (1) in Section 2). In (21), by contrast, the nominalization serves to modify the indefinite pronoun *kate(pi)* ‘something’, and thus is functionally equivalent to a relative clause. This modification use is also described by Helberg (1984: 455) (unlike the NP-use). He analyses examples like (21) as purposive clauses, although the nominalized forms clearly modify nouns and thus function at NP-level. In my data, purposive subclauses operating at clause level invariably use finite verb forms marked for imperative mood and suffixed by *-po*, which signals the dependent status of the clause (see Van linden Forthc.).

- (20) Jonas-tewapa o-niŋ-ka wa-wedn grinŋo-a
 Jonas-BEN 3SG.IND-BEN-make NMZR-lie foreigner-NOM
 ‘The foreigner makes a bed for Jonas.’
- (21) kate=pi [wa-ka hak]_{NMLZ}
 something=INDET NMZR-make house
 ‘something to make a house with, like a beam’

In (21), the notional subject of the nominalized form is left implicit (generic reference), but the notional direct object is expressed (*hak*). Like in (15) above, its lack of extra markers points to the verb-like nature of the internal syntax of this nominalization.

In addition to instrumental nominalizations, prefixation of *wa(?)*- to a verb stem also produces objective nominalizations, i.e. nouns denoting the result or the typical or ‘cognate’ object of an action (cf. Comrie & Thompson 2007: 340-341). In (22a), the word for ‘language’ is construed as the ‘result of saying’ (there is no other lexical item available), and in (22b), the term ‘alphabet’ is construed as the ‘result of writing’ or the ‘cognate object of writing’. Arguably, these examples could also be analysed as instrumental nominalizations.

- (22) (a) arakmbut-en waʔ-aʔ
 person-GEN NMZR-say
 ‘the language of the people’, ‘the Harakmbut language’
- (b) or(oʔ)-edn wa-ma-ndoya
 1PL-GEN NMZR-VPL-write
 ‘our letters’, ‘our alphabet’

Examples (22a) and (22b) form true noun phrases with genitive-marked (pro)nouns functioning as possessive determiners of lexicalized nominalizations. Although semantically these genitive forms correspond to the subjects of the verbs from which the head nouns are derived, they arguably do not form part of the nominalization. This can even be upheld for examples like (23), which features a productive indirect causation construction in which the causing event is coded as the fixed phrase “Causer-GEN *waʔ-aʔ-te*” and the caused event is coded as a main clause.

- (23) Luis-en waʔ-aʔ-te Yesika o-mba-wedn-a
 Luis-GEN NMZR-say-LOC Yesika 3SG.IND-VPL-lie-TRNS
 mbiʔigŋ mbarak-te
 fish ground-LOC
 ‘Luis makes Yesika put the fish on the ground.’ (Literally: ‘At Luis’s words/speech,
 Yesika puts the fish on the ground.’)

While in cases like (23) the nominalization does refer to what the referent of the genitive-marked noun said, the fixedness of the expression and its high productivity support an analysis in terms of a one-word nominalization, as suggested by the literal translation. Note that the locative case suffix *-te* is also found on (multi-word) nominalizations with *e(?)*-, with such forms functioning as an adverbial time clause (see Section 5.2.1).

The affix used to derive nouns referring to inanimate entities from verbs also serves a function in noun-based nominalization, in fact a very basic one. Harakmbut common nouns divide into two groups on the basis of their morphological status, viz. potentially free nouns and obligatorily bound nouns (Van linden Forthc.). Whereas potentially free nouns can stand on their own as a word form, obligatorily bound ones require a nominalizing prefix to obtain independent nominal status, e.g. *wa(?)*- in (24) (but also *e(?)*-, see Section 5.3).

- (24) wa-mbaʔ
 NMZR-hand
 ‘hand’ (Helberg 1984: 437)

The set of bound nouns is semantically restricted to nouns that refer to parts of entities, such as (human or animal) 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 basic shapes or qualities of entities. This set has been identified as “shape morphemes” by Hart (1963) (and adopted as such by Helberg (1984: 243)), and analysed as classifiers by Payne (1987: 35-37). I will come back to nominalization of bound nouns in Section 5.3, where it will be placed in an areal perspective.

5 Nominalization with prefix *e(?)*-

The second formal type of nominalization in Harakmbut features the prefix *e(?)*- appended to the verb stem. This type is predominantly used for event nominalization and only marginally for participant nominalization. In event nominalization, *e(?)*-nominalizations are found to code complement as well as adverbial relations, which will be discussed in Sections 5.1 and 5.2 respectively. It will become clear that this semantic distinction has a formal correlate, in that – unlike in complement relations – the forms coding adverbial relations all feature an extra marker signalling the type of adverbial relation. However, what is shared in both types of subordination relations is that the nominalized forms retain very few – if any – inflectional verbal categories, and that they combine NP-like external syntax with verb-like internal syntax. In the latter respect, they pattern like nominalizations in Kakataibo (Valle & Zariquiey, this volume) and Cahita (Álvarez, this volume). As the prefix *e(?)*- is also used in the citation form of verbs, nominalization with *e(?)*- used in event nominalization functions will be termed ‘infinitival’ nominalization. Section 5.3, in turn, will concentrate on participant nominalization functions of *e(?)*-nominalizations.

5.1 Complement relations

This section takes a closer look at infinitival nominalization used to code complement relations in Harakmbut. This type of infinitival nominalization is found in syntactic environments in which a core argument noun phrase is called for. Not unexpectedly, therefore, these forms either show no further morphology, or they are marked for case, specifically accusative case, flagging the direct object of the main clause verb phrase. The discussion below is organized according to the semantic types of the complement-taking predicates that pattern with infinitival nominalization (based on Noonan 2007: 120-145).

5.1.1 Commentative predicates

Commentative predicates provide a comment on their complement proposition in that they express the speaker's attitudinal evaluation of the propositional content coded in the complement (Noonan 2007: 127-128). Examples include predicates expressing judgements (*be odd, be significant, be important*) or emotional reactions (*regret, be sorry, be sad*) (Noonan *ibid.*). These two types of commentative predicates are also found in Harakmbut, as illustrated in (25), which repeats (2), and (26) respectively.

- (25) ndak ðʔ-ẽ [e-mbaʔ-tiak, ãñĩ, keme]_{NMLZ}
 good 3SG.IND-be NMZR-VPL-come FILLER tapir
 'It is good that (, eh,) the tapirs have come.'
- (26) nduru-ŋ-nda ɪʔ-ẽ-ỹ, [tareʔ
 happy-NDA 1SG-be-1.IND manioc
 e-niŋ-to-tiak opudn-a]_{NMLZ}
 NMZR-BEN-SOC-come 2PL-NOM
 'I am very happy that you (pl) have brought manioc (for us).'

In both examples, the nominalized forms constitute the head of an action nominal, containing also a noun phrase that corresponds to the subject of the verb stem. In (25) and (26), these notional subjects are not coreferential with the matrix subjects; I have no examples in which coreference does obtain. In (25), the notional subject of the nominalization, *keme*, is left unmarked, just like S-participants of independent clauses; verbal plural marking is used here to indicate plurality of the S-participant. In (26), the notional A-participant *opudna* (which itself is optional in this case according to my informant) is marked for nominative case, which contrasts with the genitive case markers found on the notional subjects of the one-word participant nominalizations in (22)-(23). We can conclude that as far as argument marking is concerned, Harakmbut infinitival action nominals do not accommodate themselves to noun phrase syntax, but have a verb-like internal syntax instead.

If we focus on the retention of inflectional verbal categories in the infinitival action nominals in (25)-(26), we are led to assume that neither tense nor aspect is retained. Even though the propositional contents denoted by the nominalizations refer to events that reached completion before the moment of the attitudinal assessment, and thus are presupposed true (Van linden 2012: 54-62, cf. Noonan 2007: 128), neither infinitival form is marked for past tense or any aspectual category. Nevertheless, (25)-(26) illustrate that infinitival nominalizations do allow for (derivational) valency-increasing morphology, as (26) has two more arguments than (25), viz. an applied direct object *tareʔ* through sociative causative prefix *to-*, and an applied indirect object (implied, not overtly expressed) through the benefactive applicative prefix *niŋ-*, cp. (6) for an independent clause (Section 3).

5.1.2 Ability predicates

Ability predicates indicate the ability of the subject participant to carry out a particular State of Affairs. They are traditionally included in the class of modal predicates, and in many languages they take part in clause or lexical union, functioning as auxiliary verbs rather than complement-taking predicates (cf. Noonan 2007: 138-139). The type of ability that is expressed by means of infinitival nominalization in Harakmbut is that of ‘learned’ or ‘acquired’ ability; the expression of ‘intrinsic’ ability does not use nominalization (cf. distinction between constructions with *saber* versus *poder* in Spanish). Examples are given in (27)-(28).

- (27) ndo ïh-nõ-põ-ẽ-ỹ
 1SG 1SG-(vital.centre-CLF:round-be)_{know}-1.IND
 [e-ndopih]_{NMLZ}
 NMZR-swim
 ‘I am able to swim.’
- (28) ïh-nõ-põ-ẽ-nde-y
 1SG-(vital.centre-CLF:round-be)_{know}-ALREADY-1.IND
 [e-mba-tuk-e? tare?]_{NMLZ}
 NMZR-VPL-plant-ITER manioc
 ‘I already learned how to sow (a field of) manioc.’

As can be expected from the semantics of the complement relation, the notional subject of the action nominal is coreferential with the syntactic subject of the matrix, and has been equi-deleted (cf. Noonan 2007: 75-76). It can even be stated that in constructions like (27)-(28) finite forms of *ẽnõpõẽ* can transfer their subject selection to the infinitival nominalization. This possibility signals that *ẽnõpõẽ* in its acquisition of ability sense has moved already some way on the auxiliatation path proposed by Heine (1993: 58-66). A more detailed discussion of whether *ẽnõpõẽ* should be analysed as a complement-taking predicate or auxiliary here is beyond the scope of this paper.

The observation that *ẽnõpõẽ* constructions like (27)-(28) are semantically restricted to acquired ability contexts can be explained by the verb’s polysemy, which is illustrated in (29). The first finite form of *ẽnõpõẽ* in (29) functions as a knowledge or acquisition of knowledge predicate that patterns with a sentence-like complement featuring a different subject and finite verb form (see Van linden Forthc.). This second finite form, by contrast, functions as an ability predicate and patterns with a nominalized form whose notional subject has been equi-deleted. It is probable that the (acquisition of) knowledge sense of *ẽnõpõẽ* blocks further semantic extension to the sense of intrinsic ability. In addition to ‘(get to) know’, *ẽnõpõẽ* is also observed to express meanings like ‘think’ (propositional attitude predicate), ‘hope’ (desiderative predicate) and ‘feel’ (immediate (sensory) perception predicate).

- (29) wa-si?-po
 NMZR-(peel-CLF:round)_{child}
 õ-nõ-põ-ẽ-mẽ
 3SG.IND-(vital.centre-CLF:round-be)_{know}-REC.PST
 kuwa [e-ndopih]_{NMLZ}
 dog NMZR-swim
 õ-nõ-põ-ẽ
 3SG.IND-(vital.centre-CLF:round-be)_{know}
 ‘The child learned that dogs are able to swim.’

Taking into account the differences in semantic properties of modal versus commentative complement relations, I assume that the nominalized forms patterning with *ēnōpōē* will tend to show retention of fewer verbal categories than those patterning with commentative predicates. As described in more detail in Van linden (2012: 203-207),⁸ drawing on Noonan (2007) and Cristofaro (2003), modal relations involve a higher degree of semantic integration and semantic dependency than commentative relations, which explains why they are frequently observed to combine with reduced complement types across languages. In Harakmbut, the main formal distinction lies in the treatment of the notional subject of the infinitival nominalization (equi-deleted vs. overtly expressed). The retention of tense is not expected in the case of modal relations, and this is borne out by the Harakmbut data. The same is expected for inflectional aspectual categories. Note that iterative aspect, coded by suffix *-e*, is a derivational type of aspect in Harakmbut, as evidenced by its influence on word stress (see Van linden Forthc.). In (28), the presence of the iterative marker in *embatuke?* can be explained by the nature of the activity learned, which was not the planting of just one manioc seed, but the sowing of a whole field of manioc (i.e. planting repeatedly), which is even enhanced by the presence of verbal plural marking.

5.1.3 Immediate perception predicates

Immediate perception predicates name the sensory mode by which the main clause subject participant directly perceives the event referred to in the complement (cf. Noonan 2007: 142). Harakmbut has three different formal strategies to code complements of this type of predicate, one of which involves infinitival nominalization. This strategy is illustrated in (30) and (31).

(30) mboerek o-tiaway-tuy apetpet-ta
 man 3SG.IND-see-DIST.PST. **NVIS.INDIR.EVD** jaguar-ACC
 [e-arak ken-en wã-tō-ē-ta]_{NMLZ}
 NMZR-kill;hit 3-GEN NMZR-SOC-be-ACC
 ‘The man saw the jaguar attacking his wife.’

(31) ken-en nãŋ-a o-pē-ē-tuy
 3-GEN mother-NOM 3SG.IND-ear.canal-be-DIST.PST. **INDIR.EVDNVIS**
 apetpet-ta [e-kwiri?-min ken-en wa-yombu-ta]_{NMLZ}
 jaguar-ACC NMZR-brain-suck 3-GEN NMZR-daughter-ACC
 ‘The mother heard the jaguar sucking the brains of her daughter.’

While in (30), the subject participant becomes aware of the (horrible) event coded in the complement clause by seeing it happen, in (31) the subject participant relies on auditory perception. Nevertheless, the examples share the same syntactic structure; in both instances the subject of the complement proposition is marked for accusative case (*apetpet-ta*), which suggests that it is syntactically treated as the (animate) direct object of the complement-taking predicate, although semantically it is the entire event that is perceived (cf. Kirsner & Thompson 1976). While the notional subjects of the nominalized forms receive their marking from the matrix clause, the direct objects of the dependent clauses seem to receive it from the nominalized forms; in both examples they have animate referents and are marked for accusative case as well (*kenen wãtōē-ta* in (30); *kenen wayombu-ta* in (31)). It can be noted that the latter participant is treated as direct object by virtue of noun incorporation of type II in Mithun’s (1984) typology. Incorporation of the bound noun *wa-kwiri?* ‘brain(s)’ vacates the position of object of *e-min* ‘suck’, to which the possessor of the incorporated noun is advanced, i.e. *kenen wayombu-ta* (cf. Mithun 1984: 857-858). This type of NI is comparable to the applicative

⁸ See also Van linden & Davidse (2009: 178-180).

marking observed in (26), as it basically serves as a valency-increasing mechanism. Finally, as can be expected from a complement relation that shows a high degree of semantic integration and dependency (cf. Noonan 2007: 142-144), the nominalized forms show no retention of inflectional verbal categories.

Comparing the infinitival nominalization strategy with the other two strategies available in Harakmbut, I hypothesize that the former is dedicated to contexts of non-deliberate perception that do not involve first person singular matrix subjects that are coreferential with the direct object of the complement proposition. For these special first person contexts, Harakmbut speakers produced constructions with sentence-like complements. The third strategy, in which the complement takes the form of a finite relative clause with the direct object of the matrix as its antecedent, seems to be restricted to contexts of deliberate perception, focussing on how exactly the perceived action proceeds. As these alternative strategies do not feature infinitival nominalization, no further examples are provided.

5.1.4 *Desiderative predicates*

Desiderative predicates express a desire that the State of Affairs coded in the complement be actualized (Noonan 2007: 132). Of the three subtypes distinguished by Noonan (2007: 132-135), it is only the *want*-class that patterns with infinitival nominalization in Harakmbut. What may strike the reader immediately when considering examples (32)-(34) is that the nominalized form is marked for accusative case (by *-ta*, cf. Table 2). Among the complement relations coded by infinitival nominalization, the desiderative relation is the only one in which the nominalized complement occurs with the nominal category of case, and thus most clearly features NP-like external syntax. However, this coding pattern is unexpected in view of the animacy constraint on O-marking, as events are inanimate entities. I have no explanation for this (but see Section 5.2.1). Incidentally, desiderative relations form the only complement relation described by Tripp (1976:3; 1995: 216) and Helberg (1984: 360, 451-452). Both provide examples with *e?pak* ‘want, love’ as complement-taking predicate, cf. (32)-(34), and note that in such cases the infinitive functions as object.⁹

- (32) [e-kore-ta]_{NMLZ} on-pak-me
 NMZR-turn.back-ACC 3PL.IND-want-REC.PST
 o-wi-nok
 3SG.IND-rain-BECAUSE
 ‘They wanted to go back because it was raining.’
- (33) pa [e?wa-ta]_{NMLZ_1} i?pak-Ø [e-mationka-ta
 Q NMZR-go-ACC 2SG-want-DUB NMZR-hunt-ACC
 ndo?-ere]_{NMLZ_2?}
 1SG-COM
 ‘Do you (sg) want to go and hunt with me?’
- (34) [mbuttinda e-ma-n-a-ta]_{NMLZ_part1} ih-pak-i
 truth NMZR-VPL-SPAT-say-ACC 1SG-want-1.IND
 [opudn-ta]_{NMLZ_part2}
 2.PL-ACC
 ‘I want to tell you (pl) the truth.’

In examples (32)-(34), the notional subject of the nominalization is coreferential with the matrix subject, and has been equi-deleted. Equi-deletion also obtains in my other examples with

⁹ Like in the case of complements of perception predicates, however, it should be noted that semantically it is the whole event coded by the complement clause that functions as direct object of the desiderative predicate.

identical subjects in main and complement clause. In addition to these contexts, my data include two examples in which the matrix and complement proposition have different subjects, one of which shows infinitival nominalization, cf. (35).

- (35) mboerek oʔ-pak-me [e-niŋ-to-tiak-ta
 man 3SG.IND-want-REC.PST NMZR-BEN-SOC-come-ACC
 keme]_{NMLZ}
 tapir
 ‘The man wanted them to bring along tapir for him.’

In (35), the matrix subject is coreferential with the (applied) indirect object of the complement proposition. The same situation holds in the other example, which uses a sentence-like complement instead, whose verb is marked for imperative mood. However, similarly to the distribution of complementation strategies used to code immediate perception relations, this case involves a first person singular matrix subject that is coreferential with a non-subject argument in the complement proposition. More systematic research is needed to check whether this first person singular context is the odd one out in other areas of grammar as well.

While the external syntax of infinitival nominalizations functioning as desiderative complements is NP-like, their internal syntax looks verb-like. Example (34) is most informative in this respect: the notional primary object of the nominalized verb is accusative-marked (*opudn-ta*), like human primary objects are treated in independent clauses. Note that this constituent of the multi-word nominalization is separated from the others by the main verb (hence *NMLZ_part1* and *NMLZ_part2* in subscript). The notional object of the nominalized verb in (35) is left unmarked, just like objects referring to dead (and lower order) animals are in independent clauses.

Finally, with desiderative relations showing a high degree of semantic integration and dependency, much like the modal and immediate perception relations (cf. Noonan 2007: 142-144), no retention of inflectional verbal categories is expected. This expectation is borne out in my data. Again, derivational categories prove different than inflectional ones, with valency-increasing morphology appearing in example (35) (cp. (26)).

5.2 *Adverbial relations*

While the previous section focussed on infinitival nominalization serving to code complement relations (NP-use of event nominalizations), this section concentrates on infinitival nominalization used to signal adverbial relations (modification use of event nominalizations), thus carrying information on the circumstances of the main event in which the core argument participants engage. This function is formally reflected in the presence of extra markers on the nominalized form. Importantly, all of the markers observed are also found to occur on underived nouns, though they are not all case markers. This feature points to the external syntax of the nominalized forms being NP-like. In what follows, I will discuss the use of infinitival nominalization to code temporal, concessive, conditional and locative relations respectively.

5.2.1 *Temporal relations*

In Harakmbut all types of temporal relations are expressed through non-finite clause types, and a number of them use infinitival nominalization followed by the suffixes *-te* or *-anda*. I hypothesize that these suffixes are in complementary distribution, with *-te* used in contexts with different subjects in matrix and subordinate clause, cf. (36)-(37), and *-anda* being restricted to same-subject contexts, cf. (38)-(39). As the function of *-te* in the nominal domain, i.e. that of a locative case marker (see *bisikleta-te* in (9), Section 3), is semantically related to that in

infinitival nominalization, both being dedicated to spatio-temporal location, *-te* has kept its locative case gloss in the examples below. Since the function of *-anda* in the nominal domain (see Section 3) seems at first sight to be unrelated to the function it has in (38)-(39), I first tentatively provide a separate gloss (but see below).

- (36) siʔnoŋ oʔ-tay-on-me [apane
 baby 3SG.IND-sleep-PFV.NVOL-REC.PST grandfather
 e-n-mba-tiapak¹⁰-te]_{NMLZ}
 NMZR-SPAT-VPL-narrate-LOC
 ‘The baby fell asleep while the grandfather told him a story.’
- (37) noŋ-ok-a 2010 [e-tiak-te An]_{NMLZ}
 other-period-ADV 2010 NMZR-come-LOC An
 o-to-k-ka-me-y mbiʔigŋ
 1PL.EXCL-SOC-SEPARATION-do-REC.PST-1.IND fish
 ‘Last year, in 2010, when An came, we fished with her.’ (spontaneous speech)
- (38) [on-a oroʔ-ta eʔ-uk-anda]_{NMLZ} tiaway-we
 2SG-NOM 1PL-ACC NMZR-search-SIMUL.SS see-NEG
 õʔ-ẽ-ně
 1<>2SG-be-IND
 ‘While you (sg) are looking for us, you (sg) don’t find us.’
- (39) [pomelo e-tipit-anda]_{NMLZ} ih-mbaʔ-te.ŋ-me-y
 grapefruit NMZR-peel-SIMUL.SS 1SG-hand-cut-REC.PST-1.IND
 ‘I cut my hand while I was peeling a grapefruit.’

Both types have been noted by Tripp (1976), who does not define the distribution of the suffixes as explicitly, but does hint at it. Specifically, Tripp’s (1976: 9-10) description of the *-anda* type mentions its relation to the subject of the main clause; the verbal noun¹¹ in *-ada* “gives the circumstance or condition related to the subject”. Tripp (1976: 8-9) treats infinitival nominalization with *-da* separately from that with *-ada*,¹² and is less precise about its function, saying it indicates “the circumstance of the lexical predicate”. However, what is striking is that all the examples he provides feature verbs whose root ends in *a* (e.g. *-ka* ‘do’, *-wa* ‘go’). This suggests that suffixation of *-anda* to stems ending in *a* involves loss of one vowel segment. On infinitival nominalization with *-te*, Tripp (1976: 4) merely states that it yields a temporal noun denoting a simultaneous event. In the grammatical notes accompanying his dictionary, Tripp (1995: 216) does state explicitly that *-te* is used in different-subject contexts. As illustrated in (40) below, we will see that the *-te* type is not restricted to simultaneous events, but can also be used to code anterior events. Incidentally, Helberg (1984: 451) considers this to be the main function of infinitival nominalization with *-te* (he does not discuss forms in *-anda*).

While all my examples with *-anda* and the majority of those with *-te* denote simultaneous events, some of those with *-te* refer to events that are anterior to the main clause event. As exemplified in (40), the nominalized form carries an extra marker, viz. *-nde*, which has the specific temporal adverbial meaning of ‘already’.

¹⁰ Helberg (1984: 461) proposes a further morphological analysis of *-tiapak*.

¹¹ While Tripp (1976) uses the Spanish term *sustantivo verbal* (‘verbal noun’), it would be more accurate to describe the cases in Sections 5.1 and 5.2 as *action nominals* (cf. Comrie & Thompson 2007: 343).

¹² Tripp (1976) writes <(a)da>, while I represent the suffix as <(a)nda>. Tripp (1995: 12) does recognize the presence of a nasal sound in the suffix, but prefers not to represent this in writing. He states that <d> is pronounced as [ʔd], with the plosive being prenasalized, in syllable-initial position. The same holds for <tada> in Section 5.2.2 below. In my analysis, [nd] is a post-stopped allophone of /n/ preceding oral vowels (see Van linden Forthc.).

- (40) arakmbut on-mba-uk-me wandey-ta,
 person 3PL.IND-VPL-search-REC.PST wounded-ACC
 heridos, wandey-ta [taka e-mba-wa-nde-te]_{NMLZ}
 wounded wounded-ACC Taca NMZR-VPL-go-ALREADY-LOC
 ‘The Harakmbut looked for the wounded after the Taca (people) had gone.’

Example (40) thus indicates that infinitival nominalization with *-te* retains markers that express temporal adverbial meanings. On forms with *-anda*, by contrast, no such markers are attested in my data. The verbal plural category is also retained in (40), just like in other types of verb-based nominalization.

If we take a closer look at how the notional arguments of the nominalized forms are marked, we can observe that these use the same markers as canonical main clause arguments. An interesting example in this respect is (38), in which the matrix verb form is highly ambiguous (on referential obscurity in the verbal argument marking system in Harakmbut, see Van linden 2014). It appears that this referential ambiguity is resolved in the nominalized clause, with two case-marked personal pronouns preceding the nominalized form, which is in turn followed by the main clause verb phrase. The notional subject of the nominalized form is marked for nominative case (*on-a*), while the direct object is marked for accusative case (*oroʔ-ta*). In all other examples, no case marking is found on the notional arguments of the nominalized forms, but this absence of marking is no different from what would be the case in independent clauses: S-participants go unmarked (cf. (37), (40)), just like non-focal animate A-participants (cf. (36)) and inanimate O-participants (cf. (39)) (see Section 3). All of this leads to the conclusion that the internal syntax of infinitival nominalization with *-te* and *-anda* is more verb-like than NP-like.

Yet, the distribution of the ‘temporal’ suffixes in other domains of Harakmbut grammar points to the nominal character of the external syntax of the nominalized types looked at here. Both suffixes are also found to occur on underived nouns. Within the nominal domain the suffix *-te* functions as a locative case marker, cf. (9) above (see also Helberg 1984: 439; Tripp 1995: 196); in addition to spatial location, it is also used to express temporal location, e.g. *agosto-te* ‘in August’. The suffix *-anda* is also found on underived nouns, specifically as a morphologically complex suffix combining nominative case suffix *-a* with the focus marker *-nda*, cf. *Lupe-a-nda* in (5) above. This morphological complexity may be key to an alternative analysis of *-anda* suffixed to infinitival nominalizations. Taking into account the same-subject restriction of infinitival nominalizations with *-anda* on the one hand, and the function of *-nda* in adnominal modification contexts (see Section 3) on the other, the suffix might as well be analysed as a complex suffix comprising the nominative suffix *-a* and the general modifier suffix *-nda*, as presented in (41).

- (41) [e-waʔ-e-a-nda]_{NMLZ} ih-kot-me-y
 NMZR-go-ITER-NOM-NDA/MOD 1SG-fall-REC.PST-1.IND
 ‘I fell while I was walking.’ (Literally: ‘I, who was walking, fell’, or ‘I fell walking’)

In this alternative analysis, the infinitival nominalization is nominative-marked so as to indicate that the event denoted by the nominalized form specifies a circumstance related to the subject of the main event. In Shibatani’s (this volume) terms, the parsing of *-anda* in (41) points to modification use of a participant nominalization, while the one in (38)-(39) points to modification use of an event nominalization. The alternative analysis is corroborated by what is observed for the suffix *-tanda* in (44), Section 5.2.2. More generally, this proposal assumes that the external syntax of nominalizations differs from that of underived nouns in terms of the

differential/optional nature of case marking. While S-participants in independent clauses tend to go unmarked, infinitival nominalizations with *-anda* are used to modify any main clause subject, whether it be an S-participant (cf. (41), or a (focal or non-focal) A-participant (cf. (38)). Similarly, nominalizations functioning as O-participants of desiderative predicates also flout the animacy constraint on case-marking; as discussed in Section 5.1.4, these do carry accusative marking, in spite of their inanimate nature.

5.2.2 Concessive relations

A second type of adverbial relation that is coded by infinitival nominalization in Harakmbut is that of concession. In this type, the nominalized forms are suffixed by *-tanda*, as is illustrated in (42) and (43). Example (43) shows that this suffix can be used in different-subject contexts (*ẽ?ẽtanda*) as well as same-subject contexts (*embaukpaktanda*).

- (42) [sik-yo ẽ?-ẽ-tanda¹³]_{NMLZ} o-mbewik-me-ne
 black-LOC NMZR-be-CONC 1PL.INCL-go.up-REC.PST-IND
 Porto-lus-yo
 Puerto-Luz-LOC
 ‘Although it was already dark/night, we went up(river) to Puerto Luz (by canoe).’
- (43) [lus ẽ?-ẽ-tanda i e-mba-uk-pak-tanda]_{NMLZ}
 light NMZR-be-CONC and NMZR-VPL-hot-VBZ-CONC
 wa-si?-po on-mba-tay-mbedn
 NMZR-(peel-CLF:round)_{child} 3PL.IND-VPL-sleep-ALL.NIGHT
 ndak-a
 good-ADV
 ‘In spite of the light and the heat (although they felt hot), the children slept well all night.’

The examples above do not show retention of any inflectional verbal category, nor does any other example in my data. With regard to argument marking, example (43) suggests that the internal syntax of infinitival nominalization with *-tanda* is verb-like, as the notional subject of the first form (*lus*) is unmarked, like the subject of independent existential constructions (S-participants, see Section 3).

The availability of infinitival nominalization with *-tanda* in both different-subject and same-subject contexts is also observed by Tripp (1976: 10-12), but he does not attribute concessive semantics to it. Instead, he argues that the construction indicates “the circumstance or condition of the related predication”; his examples do not lend themselves well to a concessive interpretation either. Similarly, Helberg (1984: 471-472) analyses infinitival nominalization with *-tanda* as expressing the temporal relation of simultaneity. His examples include both different-subject and same-subject contexts, but do not seem to involve concessive linking. However, in later work, Tripp (1995: 216) notes that *-tada* (and *-ada*, cf. Section 5.2.1) frequently has an “adversative” meaning. Some of the examples adduced allow for a concessive interpretation as well.

In addition, Tripp (1976: 10-11) notes that “verbal nouns” in *-tanda* can also “describe the circumstance of the object of the related predication.” In his example (44), I believe that *-tanda* should be analysed further into *-ta-nda* (-ACC-NDA/MOD) along the same lines as my alternative analysis for *-a-nda* in (41), Section 5.2.1. This analysis is not proposed by Tripp, although he calls *-tada* a “complex suffix” further below (1976: 12).

¹³ The *e?*- prefix has nasal quality here through nasal spreading from the verb root *-ẽ* ‘be’.

- (44) ken y-ok-wek-po ndo
 then 1SG-SEPARATION-wound.with.arrow-DEP 1SG
 [eʔ-ti-mon-an-ta-nda]_{NMLZ}
 NMZR-UP-flee-PFV.VOL-ACC-NDA/MOD
 ‘Thus I pierced the animal that he had lost.’ [‘Then I pierced the escapee, viz. an animal/person that fled (from him) from high up.’ AVL] (Tripp 1976: 11, ex. 57; adapted spelling; my morpheme breaks and glosses)

Tripp (1976: 10-11) gives two more examples that could be analysed similarly to (44); I have not encountered any comparable example in my own data so far. Semantically, in cases like (44) – like in the other examples given by Tripp – no concessive relation holds between the nominalized event and the main clause event, but rather a general modification relation targeting a main clause participant. The nominalized form in (44) thus differs from those in (42)-(43) in involving participant nominalization rather than event nominalization. Finally, it should be noted that the *-tanda* suffix is only to be found on underived nouns when it is indeed further analysable into *-ta-nda* (-ACC-NDA), but not in its monomorphemic form (see (4a) in Section 3). This implies that for concessive infinitival nominalization with *-tanda* we cannot be as confident about its external syntax being NP-like as for, e.g., infinitival nominalization with *-te*.

5.2.3 Conditional relations

Infinitival nominalization is also used in Harakmbut to signal conditional relations between events. In these cases the nominalized forms are suffixed by *-nãỹõ*, irrespective of their semantic subtype. The construction in (45) exemplifies a reality condition, while that in (46) instantiates an unreality condition of the predictive subtype (the semantic classification adopted here is the one proposed by Thompson et al. (2007: 254-262)).

- (45) [eʔ-wi-nãỹõ]_{NMLZ} mbaʔa-we ïh-ẽ-ãpo-y
 NMZR-rain-COND work-NEG 1SG-be-FUT-1.IND
 ‘If it rains, I won’t work.’
- (46) [aya-nda, aya-nda e-mba-pe-nãỹõ]_{NMLZ}
 all-NDA all-NDA NMZR-VPL-eat-COND
 o-yok-i gayeta
 1>2SG.IMP-give-1.IMP biscuit
 ‘If you (sg) eat (up) everything, I’ll give you (sg) a biscuit.’

The nominalized forms in (45) and (46) share the same formal make-up featuring the nominalizing prefix *e(ʔ)-* and the conditional suffix *-nãỹõ*, but the verb forms in the main clauses are different. The reality condition construction in (45), which refers to a habitual situation, has a main clause verb form marked for indicative mood and future tense (*ïhẽãpoy*), while the predictive (unreality condition) construction in (46) contains an imperative verb form (*oyoki*, literally ‘I should give you’). However, other predictive examples in my dataset also show indicative future forms. In addition to reality conditions and predictive ones, my data include one counterfactual situation, which is given in (47).

- (47) [eʔ-wi-me-nãỹõ]_{NMLZ} mbaʔa-we ïh-ẽ-ỹ taʔmba
 NMZR-rain-REC.PST-COND work-NEG 1SG-be-1.IND swidden
 ‘If it had rained (yesterday), I would not have worked (on) the swidden.’

In (47), the nominalized form has an extra marker compared to the forms in (45) and (46), viz. the recent past tense marker *-me* (cf. (5)-(6) in Section 3). Of all the nominalized forms available in Harakmbut, formed with *wa(?)*- (Section 4) or *e(?)*- (Section 5), this subtype is the only one that shows retention of the verbal category of tense, or at least the value of recent past. The use of a past tense marker in a counterfactual construction can readily be explained in terms of the origins of counterfactuality (cf. Van linden & Verstraete 2008), and it also forms a cross-linguistically recurrent formal feature of this type of conditional construction (cf. Van linden 2004).

A third type of unreality condition, viz. hypothetical conditions, is illustrated in Helberg (1984: 464). As can be seen in (48), the nominalized form has the same formal make-up as the forms used in reality and predictive conditions, but the main clause verb form contains the modal suffix *-ipot*, which denotes future-oriented possibility and invariably combines with verbal argument markers of the dubitative mood paradigm.¹⁴

- (48) [eʔ-ti-kot-nãỹõ]_{NMLZ} oʔ-mon-ipot
 NMZR-UP-fall-COND 1PL-flee-FUT.POSS
 ‘If they discovered us, we would flee.’ (Helberg 1984: 464; adapted spelling; my glosses)

Conditional infinitival nominalization has also been described by Tripp (1976: 6) as taking the form of *e(?)*- + verb stem + *-nãỹõ*. One of his examples shows retention of the temporal adverbial marker *-nde*, viz. (49) below, which we have also observed for infinitival nominalization with *-te* (cf. (40) above).

- (49) [on e-k-waʔ-nde-nãỹõ]_{NMLZ}
 2SG NMZR-SEPARATION-go-ALREADY-COND
 mo-mba-arak-aʔ-∅
 3>1/2PL-VPL-kill-TRNS-DUB
 ‘If you (sg) go (and leave us) now, they will kill us.’ (Tripp 1976: 6, ex. (31); adapted spelling; my morpheme breaks and glosses)

Example (49) is not only informative with regard to the formal properties of the nominalized form, it also shows what form the notional subject of the nominalized form takes. Like in the case of other types of infinitival nominalization, the notional subject is unmarked here. As S-arguments are typically left unmarked in independent clauses as well (see Section 3), we can conclude that conditional infinitival nominalization has verb-like internal syntax.

In addition, example (49) is also interesting as it challenges my earlier generalizations on the main clause verb phrase in predictive conditional constructions. Specifically, its main clause verb form is not marked for imperative or indicative mood, but for dubitative mood (see also Tripp 1976: 7, ex. (33)); an imperative form in (49) would end in *-e*, while an indicative one would end in *-ne*; cf. Van linden 2014). On the basis of Tripp’s and my own examples, I hypothesize that mood marking in the main clause verb phrase of predictive conditionals is determined by the person category of the grammatical subject, with third person subjects being restricted to dubitative mood forms, while first person subjects take indicative or imperative verb forms. This hypothesis ties in with the circumstance that we are intrinsically unable to predict how others will react if a certain condition obtains, while we can be rather confident of what we would or should do. Of course, more data are needed to verify this hypothesis.

¹⁴ Tripp (1995: 222) characterizes *-ipot* as a future subjunctive form restricted to 1SG subjects. However, my data include examples with all person and number combinations. Helberg (1990: 239) attributes desiderative meaning to *-ipot* in addition to the meaning of possibility.

Finally, like the suffixes used in temporal infinitival nominalization, the suffix *-nãỹõ* is also found on underived nouns (and pronouns). An example is given in (50) below, in which the conditional suffix attaches to the noun *wambo* ‘youngster’. The elicitation stimulus was intended to yield a modal expression (of permission), which it did, but my language consultant additionally used a conditional construction to spell out the modal agent of the (permitted) action.

- (50) *wambo-nãỹõ* *mã-õʔõ-ndik* *õnʔ-ẽ*
 youngster-COND VPL-bathe-POT 3PL.IND-be
noŋ-ti-a *wẽʔẽỹ-ỹõ*
 (other-on)_{one}-ADV water-LOC
 ‘Youngsters can bathe in the river on their own.’ (Literally: ‘If they are youngsters, they can bathe in the river on their own.’)

In conclusion, the data analysed so far suggest that conditional infinitival nominalization fits the generalization of combining verb-like internal syntax with NP-like external syntax.

5.2.4 Locative relations

The last type of infinitival nominalization to be discussed here codes the adverbial relation of location. For this type of adverbial relation, nominalization constitutes only one of the three strategies available in Harakmbut (in addition to relative and main clause constructions). Unfortunately, I can only reproduce a case in point from Tripp (1976), as my own data do not include any instances. In (51) below, the nominalized form is marked for locative case by the suffix *-yo*.

- (51) [*Kereto e-n-pa-wedn-yo*]_{NMLZ} *oʔ-ey-wa-po*
 Kereto NMZR-SPAT-CLF:stick-lie-LOC 1PL-?-go-DEP
 ‘We went to the place where Kereto was lying ill.’ (Tripp 1976: 6, ex. (27); adapted spelling; my morpheme breaks and glosses)

The locative marker found on the nominalized form in (51) is also commonly appended to underived nouns, as illustrated in example (50) above (*wẽʔẽỹ-ỹõ*) (see also Table 2). This observation testifies to the external syntax of this locative infinitival nominalization being NP-like. The nominalized form shows no retention of inflectional verbal categories, but it does feature noun incorporation of type IV (cf. Mithun 1984), with verbal classifier *-pa* ‘(shape of a) stick’ categorizing the S-participant of the nominalized form; the sick person is stick-like in that he cannot move anymore. Its notional subject goes unmarked (*Ketero*) (like S-participants in independent clauses, see Section 3), which, like in the other types of infinitival nominalization, points to verb-like internal syntax.

5.3 Participant nominalization

While the majority of cases of nominalization with *e(ʔ)*- form action nominals from predicates or propositions, which in turn function as complements or circumstantial adjuncts to the main clause event, some cases just form nouns from lexical verbs. The most straightforward case is the formation of the citation form of verbs. In eliciting verbs using Spanish infinitives like

correr ‘run’ to gather vocabulary items, cf. (52), I invariably obtained forms in *e(?)*- from all my consultants.¹⁵ Another example is in (53), which is similar to Tripp (1976: 3, ex. (11)).

- (52) *e?*-ket
NMZR-run
‘(to) run’
- (53) *o?*-sot-me *e?*-wi-a
3SG.IND-make.wet-REC.PST NMZR-rain-NOM
‘The rain made him wet.’

In (53), the nominalized form *e?wi-a* functions as the (head of the) subject NP of the clause; it is even marked for nominative case, which is to be expected on the basis of its semantic properties and grammatical role, i.e. it is an inanimate A-participant (see Section 3). My data include a similar example with *e-digηpak* ‘(have) fever’. In any case, nominalized forms like *e?wi-a* in (53) refer to inanimate entities, which can be conceived of as the result of the action denoted by the verb stem. We can therefore conclude that nominalization with *e(?)*- can also result in objective nominalization, just like nominalization with *wa(?)*- (see Section 4.2 above).

Also, like in the case of nominalization with *wa(?)*-, the affix used to derive nouns referring to inanimate entities from lexical verbs serves a basic function in noun-based nominalization. In the same way as *wa(?)*-, *e(?)*- also attaches to bound nouns to produce independent nouns. Example (54a) forms a minimal pair with (24) above. While (24) yields the independent noun ‘hand’, the nominalization in (54a) yields the independent noun ‘leaf of a plant or tree’ (see also Helberg 1984: 254, 437), which has of course a shape very similar to that of a hand and also forms an upper extremity of a living body. In its noun-based nominalization function, then, *e(?)*- serves to produce the “citation form” of (a small set of) bound nouns; another example is *e?-pu* ‘bamboo’. Note that in specific construction types, i.e. those featuring adnominal modifiers that obligatorily precede the nominal head in continuous noun phrases, bound nouns can phonologically fuse with their modifier – with the nominalizing prefix being dropped (see Van Linden Forthc.). Absence of the nominalizing prefix is also observed in word formation processes like compounding; in (54b) the bound noun root *-mba?* attaches to the adjectival root *pay* ‘bitter’ to form a(n) (independent) compound noun denoting ‘tobacco’.

- (54) (a) *e*-mba?
NMZR-hand;leaf
‘leaf of a plant or tree’ (Helberg 1984: 437)
- (b) *pay*-mba?
bitter-hand;leaf
‘tobacco’

The two nominalizing prefixes studied here thus share the function of lending independent status to a set of inalienably possessed nouns (however, note that the set of morphologically bound nouns does not exhaust the set of inalienably possessed nouns, as for instance a number of kinship terms do not constitute bound nouns, e.g. *nāη* ‘mother’, *pagη* ‘father’). The very basic nature of this function possibly suggests that it may have formed the diachronic source for its function in verb-based nominalization.¹⁶ The morphological boundedness of verb roots and bases may have facilitated this development.

¹⁵ In view of this observation and further analysis of the verbal paradigms, I believe that all verb roots are obligatorily bound morphemes.

¹⁶ The functions of the nominalizing prefixes can be compared to two functions exhibited by classifiers in multiple classifier systems in North West Amazonian languages. In Bora-Miraña, for instance, classifiers can transform mass nouns into countable nouns as well as derive nouns from verbs (Seifart 2007). Aikhenvald (2000: 220-221)

The prefixes *wa(?)*- and *e(?)*- thus are competitors in noun-based nominalization, but at this stage I cannot say what factors exactly determine their distribution within the set of bound nouns. Cases like *-mba?* which combine with the two prefixes seem to be very infrequent. In general, the prefix *wa(?)*- is used for many more bound nouns than *e(?)*-, but their respective host classes do not seem to differ in terms of semantic properties like having animate versus inanimate possessors. Rather than *wa(?)*-, it is the less frequent prefix *e(?)*- that is more interesting from an areal perspective. Specifically, it is formally and functionally similar to the dummy noun prefix *e-* in Cavineña and other Tacanan languages (Guillaume 2008: 409-416). In addition, it is comparable to the semantically empty noun formative *e-* in Kwaza, which serves as “a noun formative to lend independent status to classifiers” (Van der Voort 2005: 397). In fact, Crevels & van der Voort (2008) identify the availability of a semantically empty noun formative root taking the form of *e-/i-* as an areal feature characteristic of the Guaporé-Mamoré region.

6 Conclusion

This paper has investigated (non-finite) verb-based nominalization in the Amarakaeri/Arakmbut variety of Harakmbut, with some excursions to noun-based nominalization. Within these two types, two formal subtypes have been distinguished. While in noun-based nominalization the two formal types share the same function but show a skewed distribution of frequency (or size of host class) (Section 5.3), in verb-based nominalization they show skewed distributions of functions. That is, whereas nominalization with *wa(?)*- is restricted to participant nominalization and is predominantly used to produce nouns for NP-use (typically one-word nominalizations), nominalization with *e(?)*- is mainly used for event nominalization (typically multi-word nominalizations). Table 3 summarizes the main findings on verb-based nominalization; the numbers refer to the examples given in Sections 4 and 5 above.

Table 3 bears out that the formal and semantic categories distinguished cross-cut each other. However, there is no ‘complete’ mismatch between formal and semantic categories, as the formal categories nicely carve up the domain of subordination, with *wa(?)*-nominalizations coding relative relations while *e(?)*-nominalizations are used to code complement and adverbial relations (but see the discussion of *-a-nda* in Section 5.2.1).

further lists Guahibo, Tucano and Tariana as examples in point. It is important to note, however, that the classifiers in these languages carry a specific semantic load (and are used in up to five classifier environments) whereas the Harakmbut prefixes are semantically empty (they only have the functional value of turning a bound noun into an independent nominal, or a verb base into a nominalization).

Formal type (prefix)	Extra suffix	One-word nominalization [type of participant NMLZ]	Multi-word nominalization		
			[subordinate type]	Use	Event/Part.
<i>wa(?)</i> -	<i>-eri-X</i>	Noun with animate referent: (12)-(13)-(14) [agentive]	(15) [relative] (16) [relative]	Mod NP	Part. Part.
	<i>-X</i>	Noun with inanimate referent: (18)-(19)-(20) [instrumental] (22)-(23) [objective]	(21) [relative]	Mod	Part.
<i>e(?)</i> -	–	(52) [citation form of verb]	(25)-(26) [commentative complement] (27)-(28)-(29) [ability complement] (30)-(31) [perception complement]	NP NP NP	Event Event Event
	<i>-X</i>	Noun with inanimate referent: (53) [objective]	–		
	<i>-ta</i>	–	(32)-(33)-(34)-(35) [desiderative complement]	NP	Event
	<i>-te/-yo</i> (LOC)	–	(36)-(37)-(40) [temporal adverbial] (51) [locative adverbial]	Mod Mod	Event Event
	<i>-a-nda</i>	–	(38)-(39)-(41) [temporal adverbial]/[relative]	Mod	Event/ Part.
	<i>-tanda</i>	–	(42)-(43) [concessive adverbial]	Mod	Event
	<i>-nãyõ</i>	–	(45)-(46)-(47)-(48)-(49) [conditional adverbial]	Mod	Event

Table 3: Forms and functions of verb-based nominalization in Harakmbut (-X refers to any ending that an underived noun can take)

Further generalizations that could be made pertain to the internal and external syntax of verb-based nominalization. All types of one-word participant nominalization showed NP-like external syntax (perhaps apart from the citation form); the use of genitive-marked nouns to signal the notional subjects of the nominalized forms in (22) and (23) in fact testified to these forms constituting one-word nominalizations, and hence lacking any internal syntax. Multi-word nominalizations, in turn, were all found to combine NP-like external syntax with verb-like internal syntax, just like, for example, nominalizations in Kakataibo (Valle & Zariquiey, this volume) and Cahita (Álvarez, this volume). Specifically, if notional subjects were expressed, they appeared either unmarked or marked for nominative case, just like S- or A-participants of independent clauses do. Similarly, notional direct objects appeared unmarked when referring to inanimate entities, and marked for accusative case when referring to animate entities, thus adhering to the principles governing differential O-marking in independent clauses

(see Section 3). With respect to the external syntax of event nominalizations, some reservation was already made for the concessive subtype (see Section 5.2.2), and it was hypothesized more generally that infinitival nominalizations – unlike underived nouns – disrespect the differential/optional character of case marking (see Section 5.2.1). Here, I would also like to question the status of the nominalized forms that take no extra suffix and are used to code the complement clauses of commentative, ability and immediate perception predicates. While commentative clauses still feature notional subjects taking the same form as S- or A-participants in independent clauses, the forms used in modal and immediate perception complements might be better analysed as infinitives rather than nominalizations. In the case of ability predicates, the subject of the complement proposition proved equi-deleted (see Section 5.1.2), which is typical of infinitival complements (see Noonan 2007: 67), and in the case of immediate perception predicates, the subject of the complement proposition has its case assigned by the main clause verb phrase (see Section 5.1.3). In fact, this construction comes close to an *Accusativus-Cum-Infinitivo* construction found in the complementation system of, for example, a fair number of Indo-European languages. Further investigation is needed here, also with regard to the status of the predicate *ẽnõpõẽ*, which – in its acquired ability sense – seems to be moving along the auxiliatio pathway proposed by Heine (1993).

A final topic that this paper touched upon only briefly is comparison with Harakmbut's neighbouring languages. Areality was brought into the discussion in Section 5.3 on noun-based nominalization, but the processes of verb-based nominalization were not placed in any comparative perspective. Comparison with other Peruvian Amazonian languages and with the languages included in the Guaporé-Mamoré region (Crevels & van der Voort 2008) will reveal to what extent the Harakmbut system stands out. Comparison with the Katukina-Kanamari system (Anjos Gonçalves da Silva 2011) will contribute to the debate on the hypothesized genetic link between the Katukina family and Harakmbut. Needless to say, these form interesting avenues for further research.

References

- Adelaar, Willem. 2000. Propuesta de un nuevo vínculo genético entre dos grupos lingüísticos indígenas de la Amazonía occidental: Harakmbut y Katukina. In *Actas del I Congreso de Lenguas Indígenas de Sudamérica*, Luis Miranda Esquerre (ed.), Vol. 2, 219-236. Lima: Universidad Ricardo Palma.
- Adelaar, Willem. 2007. Ensayo de clasificación del katawixí dentro del conjunto harakmbut-katukina. In *Lenguas indígenas de América del Sur: Estudios descriptivo-tipológicos y sus contribuciones para la lingüística teórica*, Andrés Romero-Figueroa, Ana Fernández Garay & Angel Corbera Mori (eds), 159-169. Caracas: Universidad Católica Andrés Bello.
- Aikhenvald, Alexandra Y. 2000. *Classifiers. A Typology of Noun Categorization Devices*. Oxford: Oxford University Press.
- Andrews, Avery. 2007. Relative clauses. In *Language typology and syntactic description: Vol. 2: Complex constructions* (2nd edition), Timothy Shopen (ed.), 206-236. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511619434.004>
- Anjos Gonçalves da Silva, Zoraide dos. 2011. *Fonología e Gramática Katukina-Kanamari* [LOT Series 281]. PhD dissertation, Vrije Universiteit Amsterdam.
- Comrie, Bernard & Thompson, Sandra A. 2007. Lexical nominalization. In *Language typology and syntactic description: Vol. 3: Grammatical Categories and the Lexicon* (2nd edition), Timothy Shopen (ed.), 334-381. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511618437.006>

- Crevels, Mily & Van der Voort, Hein. 2008. The Guaporé-Mamoré region as a linguistic area. In *From linguistic areas to areal linguistics*, Pieter Muysken (ed.), 151-79. Amsterdam: John Benjamins.
<https://doi.org/10.1075/slcs.90.04cre>
- Cristofaro, Sonia. 2003. *Subordination*. Oxford: Oxford University Press.
- Dryer, Matthew S. & Haspelmath, Martin (eds). 2013. *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. (Available online at <http://wals.info>)
- Fauconnier, Stefanie. 2011. Differential Agent marking and animacy. *Lingua* 121: 533-547.
<https://doi.org/10.1016/j.lingua.2010.10.014>
- Guillaume, Antoine. 2008. *A grammar of Cavineña* [Mouton Grammar Library no. 44]. Berlin & New York: Mouton de Gruyter.
<https://doi.org/10.1515/9783110211771>
- Gray, Andrew. 1996. *The Arakmbut: Mythology, Spirituality and History*. Oxford: Berghahn.
- Hart, Raymond. 1963. Semantic components of shape in Amarakaeri Grammar. *Anthropological Linguistics* 5(9): 1-7.
- Heine, Bernd. 1993. *Auxiliaries: Cognitive forces and grammaticalization*. New York: Oxford university press.
- Helberg Chávez, Heinrich A. 1984. Skizze einer Grammatik des Amarakaeri. PhD dissertation, Tübingen University.
- Helberg Chávez, Heinrich A. 1990. Análisis funcional del verbo amarakaeri. In *Temas de lingüística amerindia*, Rodolfo Cerrón Palomino & Gustavo Solís Fonseca (eds), 227-249. Lima: Concytec.
- Kirsner, Robert & Thompson, Sandra A. 1976. The role of pragmatic inference in semantics: A study of sensory verb complements in English. *Glossa* 10: 200-240.
- Matteson, Esther. 1972. Proto Arawakan. In *Comparative studies in Amerindian languages*, Esther Matteson, Alva Wheeler, Frances L. Jackson, Nathan E. Waltz & Diana R. Christian (eds), 160–242. The Hague: Mouton.
<https://doi.org/10.1515/9783110815009.160>
- McGregor, William B. 2007. Ergative marking of intransitive subjects in Warrwa. *Australian Journal of Linguistics* 27(2): 201-229.
<https://doi.org/10.1080/07268600701531351>
- McGregor, William B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120: 1610-1636.
<https://doi.org/10.1016/j.lingua.2009.05.010>
- 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.
<https://doi.org/10.1353/lan.1984.0038>
- Moore, Denny. 2007. Endangered languages of lowland tropical South America. In *Language diversity endangered*, Matthias Brenzinger (ed.), 29-58. The Hague: Mouton.
- Noonan, Michael. 2007. Complementation. In *Language typology and syntactic description: Vol. 2: Complex constructions* (2nd edition), Timothy Shopen (ed.), 52–150. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511619434.002>
- Payne, Doris L. 1987. Noun Classification in the Western Amazon. *Language Sciences* 9(1): 21-44.
[https://doi.org/10.1016/S0388-0001\(87\)80007-4](https://doi.org/10.1016/S0388-0001(87)80007-4)
- Pozzi-Escot, Inés (ed.). 1998. *El Multilingüismo en el Perú*. Cuzco: CBC-SID.

- Seifart, Frank. 2007. The prehistory of nominal classification in Witotan languages. *International Journal of American Linguistics* 73(4): 411-445.
<https://doi.org/10.1086/523772>
- Thompson, Sandra A., Longacre, Robert E. & Hwang, Shin Ja J. 2007. Adverbial clauses. In *Language typology and syntactic description: Vol. 2: Complex constructions* (2nd edition), Timothy Shopen (ed.), 237-300. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511619434.005>
- Tripp, Robert. 1976. Sustantivos verbales y frases de sustantivos verbales en Amarakaeri. In 14. Datos Etno-Lingüísticos: Colección de los archivos del ILV 50. Lima: Instituto Lingüístico de Verano.
- Tripp, Robert. 1995. *Diccionario amarakaeri-castellano*. Yarinacocha: Min. de Educación & Instituto Lingüístico de Verano.
- Van der Voort, Hein. 2005. Kwaza in a Comparative Perspective. *International Journal of American Linguistics* 71 (4): 365-412.
<https://doi.org/10.1086/501245>
- Van Gijn, Rik, Haude, Katharina & Muysken, Pieter. 2011. Subordination in South America: an overview. In *Subordination in native South American languages* [Typological Studies in Language 97], Rik van Gijn, Katharina Haude & Pieter Muysken (eds), 1-24. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.97.01van>
- Van linden, An. 2004. A typological study of counterfactuality. MA dissertation, University of Leuven (KU Leuven).
<https://doi.org/10.13140/RG.2.2.22541.59369>
- Van linden, An. 2012. *Modal adjectives: English deontic and evaluative constructions in diachrony and synchrony* [Topics in English Linguistics 75]. Berlin: De Gruyter Mouton.
<https://doi.org/10.1515/9783110252941>
- Van linden, An. 2014. Argument marking in Harakmbut: Looking for referential transparency. Paper presented at Syntax of the World's Languages VI, 8-10 September 2014, University of Pavia.
<https://doi.org/10.13140/2.1.5077.4400>
- Van linden, An. Forthcoming. Harakmbut. In *Amazonian Languages, An International Handbook*, Patience Epps & Lev Michael (eds). Berlin: de Gruyter Mouton.
- Van linden, An & Verstraete, Jean-Christophe. 2008. The nature and origins of counterfactuality in simple clauses: Cross-linguistic evidence. *Journal of Pragmatics* 40: 1865-1895.
<https://doi.org/10.1016/j.pragma.2008.03.008>
- Van linden, An & Davidse, Kristin. 2009. The clausal complementation of deontic-evaluative adjectives in extraposition constructions: A synchronic-diachronic approach. *Folia Linguistica* 43(1): 175-217.
<https://doi.org/10.1515/FLIN.2009.005>
- Wise, Mary R. 1999. Small language families and isolates in Peru. In *The Amazonian languages*, R. M. W. Dixon & Alexandra Y. Aikhenvald (eds), 307-340. Cambridge: Cambridge University Press.

Abbreviations

1	1 st person	INDET	indeterminate
2	2 nd person	INDIR.EVD	indirect evidential
3	3 rd person	INS	instrumental
>	'acts on'	ITER	iterative
ACC	accusative	LOC	locative
ADV	adverb(ializer)	MOD	modifier
AN	animate	NEG	negation
APPL	applicative	NMZR	nominalizer
BEN	beneficiary/ benefactive	NOM	nominative
CLF	classifier	NVIS	non-visual evidential
COLL	collective	NVOL	non-volitional
COM	comitative	PFV	perfective
CONC	concessive	PL	plural
COND	conditional	POT	potential
DEP	dependent verb form	PRIV	privative
DIM	diminutive	Q	question particle
DIST	distal	REAS	reason
DIST.PST	distant past	REC.PST	recent past
DUB	dubitative	REST	restrictive
EXCL	exclusive	SG	singular
FOC	focus	SIM	simulative
FUT	future	SIMUL	simultaneity
FUT.POSS	future-oriented possibility	SOC	sociative causative
GEN	genitive	SPAT	spatial prefix
IMP	imperative	SS	same subject
INCL	inclusive	TRNS	transitivizer
IND	indicative	VBZ	verbalizer
		VOL	volitional
		VPL	verbal plural