



**HAL**  
open science

# An Austronesian-type voice system in an Amazonian isolate? Comparing Movima and Tagalog

Katharina Haude

► **To cite this version:**

Katharina Haude. An Austronesian-type voice system in an Amazonian isolate? Comparing Movima and Tagalog. 2023. hal-04211901

**HAL Id: hal-04211901**

**<https://hal.science/hal-04211901>**

Preprint submitted on 22 Sep 2023

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# **An Austronesian-type voice system in an Amazonian isolate?**

## **Comparing Movima and Tagalog**

Katharina Haude

*Centre National de la Recherche Scientifique, SEDYL UMR 8202*

Revision 1, August 2023

### **Abstract**

The paper discusses some typologically rare structural similarities between Movima, a South American isolate, and the Austronesian language Tagalog. Both languages have a symmetrical voice system, and in both languages verbs and nouns are syntactically nearly equivalent. For Tagalog, it has been argued that the system is due to a basically equational sentence pattern with a nominal predicate (the “nominalist hypothesis”), and this explanation is also plausible for Movima. However, in contrast to some accounts of Austronesian languages, there is no evidence of a nominalizing origin of the Movima voice markers, which would help to explain the use of verbs in nominal domains. The article shows that the description of small, understudied, and even isolate languages is necessary to realize that some grammatical patterns may not be as uncommon as one may think, and can help to refine the ideas conceived about the better-known linguistic systems.

## **1 A Bolivian isolate and its parallels with an Austronesian language**

Movima is one of the approximately 12 linguistic isolates of the Guaporé-Mamoré linguistic area (see Crevels & van der Voort 2008), the south-western Amazonian region between Bolivia and Brazil. Movima is an endangered language, still spoken by a few hundred adults, but not transmitted as a first language anymore. The speakers of the language are today largely concentrated at one spot, which is the town Santa Ana del Yacuma, a former Jesuit mission in the Bolivian department Beni. When speakers move away, they tend to settle in larger cities, where they cease using the language. However, there are still some speakers in small settlements in the region, e.g. in Coquinal near the Lago Rogaguado. The most distant

settlement that hosts a small Movima speaker community, near the town San Borja about 300km away from Santa Ana, was founded only a few decades ago. No obvious dialectal differences can be discerned between speakers living in these different places. Neither has it ever been possible to discern a genealogical relationship between Movima and any other language, be it close or distant, so that the isolate status of Movima is undisputed (Seifart & Hammarström 2019).

As pointed out by Campbell (2019), a language isolate must be considered an independent language family consisting of one single member. On the one hand, this facilitates the task of description, as there is no question of which language of the family should be described first or which language of a family is more conservative or innovative. On the other hand, describing an isolate is difficult because it is not clear what can be expected, apart from areal phenomena. In principle, a researcher describing an isolate has to start from scratch, especially when there are only few previous descriptions. In the case of Movima, these were largely limited to the work by the SIL linguists Judith and Robert Judy in the 1960s (e.g. Judy & Judy 1967) and to a small elicitation-based study on classifiers by Grinevald (2002).

Therefore, when I investigated the morphosyntax of Movima in the early 2000s, it was all the more surprising to find that the closest correspondences to the Movima grammatical patterns seemed to occur in the western Austronesian language Tagalog. The way in which Tagalog is analyzed by some linguists (in particular, Himmelmann 1991; Himmelmann 2008; Kaufman 2009a; Kaufman 2009b) seems to make the unusual properties of Movima fall in place: as will be argued below, it is the idea that an apparently transitive clause has as its predicate a participant-oriented nominal – similar to an active or passive participle – whose morphology signals the semantic role of the single argument of the clause.

In providing a description of these parallels, the paper shows once again why the study of isolates, being one-member language families, is so important. So far, symmetrical voice is often presented as being unique to western Austronesian languages. The discovery of a highly similar system in a different, areally remote language family shows that symmetrical voice does not only exist in this family and that it can arise independently of a genealogical or areal connection. Moreover, the existence of a well-documented language family with similar traits makes it possible to look for the possible diachronic origin of this pattern in the isolate. This study, therefore, first points out the similarities and differences between the Movima direct-inverse and the Philippine-type symmetrical voice system, and then tries to find evidence showing that the Movima system may have arisen through a similar diachronic process.

This paper is structured as follows. Section 2 gives an outline of Movima and Tagalog syntactic patterns. In 2.1, the structural properties of basic clauses in both languages are presented. In 2.2, the two transitive clause patterns of Movima are introduced and compared to the symmetrical voice alternations of Tagalog. Section 2.3 describes an important syntactic parallel between the two languages: the restriction on the extraction of the non-subject (here illustrated with headed relative clauses), which the voice alternation helps to circumvent. The findings of Section 2 are summarized in 2.4. Section 3, then, turns to another striking parallel between Movima and Tagalog: the syntactic flexibility of nouns and verbs. That nouns and verbs are distinct lexical categories in the two languages is argued in 3.1. Section 3.2 describes the use of nouns in predicate function, and 3.3 illustrates the occurrence of verbs in determiner phrases (DPs). Section 4 contains an attempt to explain the diachronic origin of the Movima system along similar lines as proposals for Tagalog which argued that all predicates are nominal and that the voice markers may originate from nominalizers (4.1). In 4.2, it is shown that voice marking is not necessary for Movima verbs in order to occur in a nominal function, and 4.3 argues that an origin as a nominalizer can only be proposed for one single voice marker, if at all. It is concluded that (overt morphological) nominalization may not be a prerequisite for such a system to arise. Section 5 summarizes the findings.

Note that not much of what is shown here is essentially new. The Movima facts have been discussed in several publications (e.g. Haude 2006; Haude 2019). The Tagalog data were assembled from different available sources, indicated in the relevant sections. What is new, however, is the more systematic comparison between the two systems (although, given the topic of this volume, the focus is on Movima, the linguistic isolate).<sup>1</sup> The many differences

---

<sup>1</sup> I am deeply indebted to the Movima speakers who allowed me to study their language and provided the data on which my research is based. The study was supported by the research lab *Structure et Dynamique des Langues* (SEDYL, CNRS UMR8202, INALCO, IRD 135). It is a late outcome of the RHIM project (“Referential Hierarchies in Morphosyntax”) of the ESF/EuroCores/EuroBABEL programme (Deutsche Forschungsgemeinschaft, HA-5910/1-1, 2009-2013) and has also benefited from the research operation *Morphosyntax in Discourse* (Axe3-GL5 of the LabEx EFL, ANR-10-LABX-0083 and IdEx Université Paris Cité – ANR-18-IDEX-000, 2020-2024). The analyses presented here owes a great deal to discussions I had over the years with Isabelle Brill, Werner Drossard, Spike Gildea, Nikolaus Himmelmann, Daniel Kaufman, Yury Lander, Francesc Queixalós, and Hans-Jürgen Sasse, among others. I wish to thank two anonymous reviewers and the editors of the present volume for their encouraging remarks on a previous version of this paper. All mistakes and shortcomings are my own responsibility.

between the two languages – to be expected given that they are unrelated and areally distant – are not scrutinized. Some of these differences, such as the discourse properties of the syntactically privileged argument, merit particular attention, but must be left for later research.

## 2 Alignment patterns and extraction restrictions

### 2.1 Intransitive and transitive clauses

Movima basic clauses are predicate initial. An intransitive clause consists minimally of the content word representing the predicate, usually accompanied by one or more particles; in (1), the unexpressed argument is represented as  $\emptyset$ .<sup>2</sup>

- (1) *chi:~chi* [Ø] *jayna rey*  
 MD~go\_out DSC EPIST  
 ‘(They) came out then.’ [JGD\_160808-Fundacion\_2 005]

The argument (S, represented in square brackets) can be expressed by a full referential phrase (henceforth, DP) consisting of a determiner and a content word, as in (2), or by a bound pronoun encliticized to the predicate through what I call “external cliticization” (marked by a double hyphen, --, and explained below in the context of transitive clauses), as in (3).<sup>3</sup>

- (2) *chi:~chi* [os *rulrul*]  
 MD~go\_out ART.N.PST jaguar  
 ‘A jaguar came out.’ [Balvina 144]

- (3) *ji<wa:~>wa[--us]*  
 come<MD~>--3M.AB  
 ‘He came.’ [Balvina 234]

---

<sup>2</sup> Unless otherwise indicated, all Movima examples are taken from a corpus of spontaneous oral discourse of approximately 26,000 clauses (130,000 words, 30 hours) collected between 2001 and 2012 during fieldwork in Santa Ana del Yacuma. The segments in parentheses after the translation indicate the source of the examples (usually speaker acronym, date, and short title of the text).

<sup>3</sup> The argument may also be represented by a pronominal copula preceding the predicate. This sentence type, however, is not considered “basic”, and therefore not included here (Haude 2018a; Haude 2018b).

The referential element of the argument expressions – the article in (2) and the pronoun in (3) – encode number (singular vs. plural/mass, ‘PL’), humanness and sex (human male, ‘M’, human female, ‘F’, non-human ‘N’), and spatial location (present at the speech situation or generic vs. absent from the speech situation, ‘AB’); determiners (here, articles) furthermore encode ongoing vs. ceased (‘PST’) existence. There is no tense marking on the verb, and unless there is an article indicating past tense (Haude 2011a), the tense forms chosen in the English translations of the examples follow the context in which the example occurred. Likewise, the Movima article does not indicate (in)definiteness, and so, the English translation of a DP as either definite or indefinite is based on the context.

Canonical basic intransitive clauses have a verb as their predicate, as in (1)–(3); as will be shown, however (3.2), also nouns can function as predicates of intransitive clauses. There is no agreement morphology on the predicate: cliticized pronouns represent the argument and occur *instead* of a DP, not in addition to it.

Transitive clauses contain maximally two core argument expressions. There is no case marking differentiating them. Rather, the arguments are distinguished by constituent order, which is reflected by linear ordering, morphophonological processes and by different syntactic properties. Due to the direct-inverse alternation leading to two transitive clause patterns in which the semantic roles are swapped (see 2.2), it is not helpful to characterise the syntactic arguments in terms of semantic roles. Therefore, I use the terms “internal” and “external argument” to distinguish their structural position as either internal or external to the predicate phrase.

The internal argument is linked to the predicate through “internal cliticization” (Haude 2006: 97–101), which has the following characteristics: the referential element (pronoun or determiner) is attached to the host in such a way that it forms a phonological unit with it, and it requires a preceding vowel (on a consonant-final host, the epenthetic vowel *-a* is inserted). The external argument, by contrast, is phonologically independent when expressed by a DP, like S of an intransitive clauses (see (2)); when expressed by a pronoun, the pronoun is attached to the host through “external cliticization” (Haude 2006: 101–103), which is also the way in which a pronoun is attached to the predicate of an intransitive clause (see (3)). This type of cliticization has no prosodic effect on the host and no further requirements, but the external cliticization of a vowel-initial pronoun to a consonant-final host leads to resyllabification, turning the consonant into the onset of the first syllable of the pronoun.

Below are examples of transitive clauses in which the arguments are expressed either by a DP or by a pronoun. For clarity, in these examples the internal argument is rendered in boldface and the external argument is inserted in brackets. Note that in all these clauses the predicate is marked as ‘direct’, indicating that the internal argument represents the actor and the external one the undergoer (see 2.2).

In (4), which contains two transitive predicates, the internal argument is represented by a pronoun, while the external argument remains unexpressed (here represented as  $\emptyset$ ), its referent being inferable from the context. This is a rather common situation (Haude 2014).

- (4) *bat-na=is*      [ $\emptyset$ ] *n-os*                      *ba<kwa~>kwa=isnos*      *Ana*  
 put-DR=3PL.AB                      OBL-ART.N.PST    head<INAL~>=ART.F.PST    Ana  
*che pek-na=is*      [ $\emptyset$ ]  
 and lift-DR=3PL.AB  
 ‘They put (it) on Ana’s (i.e. the statue’s) head and lifted (her).’  
 [JGD\_160808-Fundacion\_2 361]

In (5), the internal argument is expressed by a pronoun and the external argument by a DP; this is the most frequent constellation of argument expressions in transitive clauses (Haude 2014).

- (5) *tikoy-na=us*      [*os*              *rulrul*]  
 kill-DR=3M.AB    ART.N.PST    jaguar  
 ‘He killed the jaguar.’    [PMP\_HRR\_etal\_210908 085]

In (6), both arguments are expressed as a DP, which, in line with crosslinguistic tendencies (DuBois 1987), is not very common in actual discourse.

- (6) *jul<a>ra=is*                      *mowi:maj*      [*is*      *kanicha:na*]  
 outwin<DR>=ART.PL    Movima                      ART.PL    Canichana  
 ‘The Movimas outwon the Canichanas.’      [JGD\_160808-Fundacion\_2 233]

In (7), both arguments are expressed by a pronoun, an equally rare situation (Haude 2014). Note that the ‘obviative’ element *k-* on the externally cliticized pronoun in (7) occurs because the internal argument is or includes a third person referent (e.g. it refers to a third person or to the first person plural

exclusive; see Haude 2006: 279-280); it is, therefore, not a morphological particularity of pronouns expressing the external argument of a transitive clause.

- (7) *jiwa-łe-na=us[--k-i 'ne]*  
 come-CO-DR=3M.AB--OBV-3F  
 'He brought her.' [EAO\_Cbba 116]

In (8), finally, the highly uncommon situation is displayed in which the internal argument is expressed as a DP and the external argument as a bound pronoun. Since a DP usually refers to a third person, the pronoun encoding the external argument in these cases always bears the obviative marker *k-*. Therefore, there is never a resyllabification with a preceding consonant-final noun, so that it is impossible to tell whether the pronoun is encliticized to the noun or not. Here, the pronoun is represented as a clitic for the sake of consistency.

- (8) *yok-na=is*                      *pa:ko[--k-as]*  
 catch-DR=ART.PL      dog--OBV-3N.AB  
 'The dogs attacked it.' [EAO Jaguar 152]

The examples show that the external argument of the transitive clause (which has the P role in the above examples) is encoded in the same way as the S argument of an intransitive clause (as mentioned, the obviative marker *k-* is not a morphological property of pronouns in transitive clauses). The internal argument (which has the A role in the above examples), in turn, is encoded like a possessor. Also a possessor DP or pronoun is linked to the possessed noun through “internal cliticization” (Haude 2006: 97–101): the referential element (pronoun or determiner) is attached to the host in such a way that it forms a phonological unit with the host; on a consonant-final host, internal cliticization triggers the insertion of the epenthetic vowel *-a*. Example (9) illustrates this with two possessed DPs: the first DP (*us alwaj-a*) is possessed by another DP (*'nes majniwa*), which in turn is marked as possessed by a bound pronoun (= *'ne*). The brackets indicate the layered structure of this example.

- (9) *[[[us alwaj-a]='nes majniwa]='ne]*  
 ART.M      spouse-LV=3F      offspring\_of=3F  
 'her daughter's husband' [EAO Neighbours 011]



Thus, it can be stated that a basic intransitive and a basic transitive clause both contain a predicate and an optional argument expression that is located externally to the predicate phrase (as reflected by its looser phonological attachment, for instance); a transitive clause additionally contains an obligatory argument expression, which is encoded like a possessor and located within the predicate phrase. To summarize this, the structure of a basic Movima transitive clause is depicted in (10a) and that of an intransitive clause in (10b).

- (10) a. [[transitive verb] [argument]]<sub>PREDICATE</sub> [argument]  
 b. [intransitive verb]<sub>PREDICATE</sub> [argument]

Let us now turn to Tagalog. Also in Tagalog basic clauses, the predicate is in initial position. One argument in a transitive clause is encoded in the same way as the S of an intransitive clause, while the other argument is encoded like a possessor. In contrast to Movima, where the syntactic arguments are distinguished by constituent order, in Tagalog they are differentiated by segmental morphology: the DP that aligns with S of an intransitive clause contains a different determiner (the ‘nominative’: *ang* for common nouns, *si* for proper nouns) than the other DP, which is encoded like a nominal possessor and marked by ‘genitive’ *ng* (for common nouns) or *ni* (for proper nouns). There are also different sets of pronouns marking either one or the other relation. Accordingly, the linear order of the arguments is freer in Tagalog than in Movima (Schachter & Otnes 1972: 83). The following examples of Tagalog illustrate an intransitive clause (11), a transitive clause (12), and for comparison with transitive argument encoding, a possessed DP (13).

- (11) Tagalog intransitive clause (Schachter & Otnes 1972: 60; glosses added)

*G<um>ising [ang bata].*  
 <AV.PFV>awake NOM child  
 ‘The child awoke.’

- (12) Tagalog transitive clause (Schachter & Otnes 1972: 60; glosses added)

*G<in>ising ng ingay [ang bata].*  
 <UV.PFV>awake GEN noise NOM child  
 ‘A noise awakened the child.’

(13) Tagalog possessed DP (Kroeger 1993: 32)

*ang sungay ng kalabaw*

NOM horn GEN buffalo

‘the buffalo’s horn’

Even though in Tagalog the ordering of the nominal constituents after the predicate is not fixed, Schachter & Otnes (1972: 60) analyze the genitive argument, independently of its surface position in a transitive clause, as being “part of the Tagalog predicate”. An excerpt from their schematic representation of a Tagalog basic transitive and intransitive sentence (Schachter & Otnes 1972: 66) is presented in (14a) and (14b), respectively (with their tree structure replaced by brackets; note that Schachter & Otnes label the predicate-internal argument “complement” and the external argument “topic”).

(14) a. [[transitive verb] [complement]]<sub>PREDICATE</sub> [topic]

b. [intransitive verb]<sub>PREDICATE</sub> [topic]

Since the determiners that distinguish the arguments in a Tagalog transitive clause are different morphemes, they are often analyzed as case markers (see Kroeger 1993). They do not, unlike their Movima counterparts, encode any ontological properties of the referents, apart from marking the difference between common nouns and proper names. A further difference to Movima is that the Tagalog nominative phrase is generally interpreted as definite, while this is not the case for the Movima external argument.

So, the essence of what was shown above is that both in Movima and in Tagalog, the transitive clause contains one argument that shares its formal properties with the single argument of an intransitive clause (in Movima, through constituency; in Tagalog, through case marking). The other argument in a transitive clause shares its formal properties with those of a possessor.

So far, this is not very unusual cross-linguistically. These properties correlate, however, with the presence of multiple transitive constructions, a phenomenon that is generally described as a specificity of western Austronesian languages (e.g. Zúñiga & Kittilä 2019) and which I will now turn to.

## 2.2 Two transitive clause patterns

Movima has two transitive constructions (henceforth called “voices” to facilitate the comparison), which are labelled “direct” and “inverse” for reasons that will be made clear below in this section.<sup>4</sup> The above examples of transitive clauses were all of the direct type. In the direct voice, the internal argument encodes the actor and the external argument encodes the undergoer. In the inverse voice, illustrated in (15) (with a pronoun and a DP) and (16) (with two bound pronouns) below, the semantic relations are arranged in the opposite way. The difference is only signalled by morphological marking on the verb: in basic main clauses the inverse is marked by *-kay*, while the direct voice is marked by the suffix *-na* or its base-internal allomorph *-a-* (Haude 2006: 323–325).

- (15) *bu'ni yok-kay-a=us [os rulrul]*  
perhaps catch-INV-LV=3M.AB ART.N.PST jaguar  
'Perhaps he got caught by a jaguar.' [EAO Jaguar 213]

- (16) *jayna rey alwani-kay-a=us[--k-us]*  
DSC EPIST speak-INV-LV=3M.AB--OBV-M.AB  
'He had talked to him first.' [JAO Naye 056]

In Tagalog, too, verbal marking, generally referred to as “voice”, indicates the semantic roles of the clausal arguments. The system contains one actor voice (AV), which identifies the subject as the actor, and several undergoer voices (UV), in which the verbal marking identifies the subject as the patient, beneficiary, or other non-actor participant in the event. The most productive undergoer voice (Himmelman 2005: 367) is the patient voice, shown in (17a). The contrasting actor voice is given in (17b). (The Tagalog voice markers, unlike the Movima direct and inverse markers, also encode aspect.)

---

<sup>4</sup> The label “voice” is adequate for the Movima direct-inverse alternation to the extent that under particular circumstances (namely relatively equal ranking on the person/animacy hierarchy), speakers can choose between the two constructions.

(17) Tagalog transitive clause (adapted from Nagaya 2012: 50)

- a. UV: *K<in>ain ng bata [ang tinapay].*  
<PV.PFV>eat GEN child NOM bread  
'The child ate the bread.'
- b. AV: *K<um>ain [ang bata] ng tinapay.*  
<AV.PFV>eat NOM child GEN bread  
'The child ate some bread.'

The languages differ with regard to the factors on which the choice between the two transitive voices is based. In Movima, the alternation between direct and inverse is determined by a referential hierarchy: the referent that is higher in hierarchies of person (1 > 2 > 3), animacy (human > non-human animate > inanimate), and discourse topicality (topical > less topical), is encoded as the internal argument, while the lower-ranking referent is encoded as the external argument. Sentences corresponding to the Tagalog examples (12) and (17b), where the human event participant ('child') is encoded in the nominative and the inanimate event participant ('noise', 'bread') in the genitive, would not occur in Movima spontaneous speech: in Movima, the inanimate event participant in such a scenario would obligatorily be encoded as the external argument (which corresponds to the Tagalog nominative) and the human participant would be encoded as the internal argument (corresponding to the Tagalog genitive). When both participants rank equal in the person/animacy hierarchy, in Movima the internal argument encodes the more topical (i.e. discourse-given) referent, which is usually represented by a pronoun and interpreted as definite. Finally, in Movima there is a grammatical restriction on person encoding, according to which a higher-ranking (i.e. first or second) person must be encoded as the internal argument.

In Tagalog, by contrast, these hierarchies do not play a significant role, and the criteria for the choice of which event participant to encode as nominative are less easy to identify. One thing that seems to be clear, however, is that the Tagalog nominative phrase is usually interpreted as definite, as reflected by the translations of the above examples. In Movima, the interpretation of a DP as definite or indefinite depends on the context. In accordance with the referential hierarchies, however, the internal argument is more likely to be interpreted as definite than the external argument.

According to Zúñiga & Kittilä (2019: 149), the above-mentioned differences between inverse and symmetrical-voice systems “justify the separate treatment most studies have given western Austronesian and non-Austronesian voice systems”. Nevertheless, the syntactic properties of the two argument phrases in a Movima and a Tagalog transitive clause are astonishingly similar. We already saw the similarities in the basic sentence patterns. A further remarkable property of both systems is that relativization (and related operations that can be termed “extraction”) may only involve the external (in Tagalog, the nominative) argument. This will be illustrated in 2.3 for headed relative clauses.

### 2.3 Extraction restrictions

In Movima, only the external argument can be relativized (see Haude 2019; Haude In press). Headed relative clauses are introduced by the particle *di*’; the relativized argument is gapped, i.e., it does not appear inside the relative clause. This is shown in (18) with a transitive direct, and in (19) with a transitive inverse predicate.<sup>5</sup>

(18) *[os            tolkosya] di’ yey-na=us*  
 ART.N.PST girl            REL want-DR=3M.AB  
 ‘(He didn’t look for) some girl that he liked.’            [LTC\_020906\_5 350]

(19) *[isnos        kwe:ya] di’ yey-kay-a=us*  
 ART.F.PST woman        REL want-INV-LV=3M.AB  
 ‘(There was) a woman who liked him.’            [ERM\_150806 595-596]

This restriction on relativization is also a well-known trait of the symmetrical-voice system of Tagalog. In Tagalog, only the nominative argument may be relativized (and otherwise extracted; Kroeger 1993). When the nominative argument encodes the actor, the verb is marked for actor voice, and when the nominative argument encodes an undergoer, an undergoer (patient, locative, or beneficiary) voice is chosen. In (20a), the undergoer is relativized, and in (20b), the actor is relativized. The relativized argument is in the nominative, and verbal voice marking indicates the semantic role of the relativized argument.

---

<sup>5</sup> The use of a neuter article for human referents, as in (19), is possible with non-specific referents.

(20) Tagalog relativization (Kroeger 1993: 211)

a. UV: *ang bata ng [t<in>uksu ni Josie]*  
NOM child LK <PV.PFV>tease GEN Josie  
'the child who Josie teased'

b. AV: *ang bata ng [na-nuksu kay Josie]*  
NOM child LK AV.PFV-tease DAT Josie  
'the child who teased Josie'<sup>6</sup>

The Movima direct and inverse affixes, therefore, resemble the Tagalog voice affixes in that they allow the relativization of the external argument independently of the semantic role of that argument. Since both the external argument of Movima and the nominative argument of Tagalog align with S of an intransitive clause and have exclusive access to relativization (as well as other to extraction processes not discussed here such as topicalization; see Haude 2019), in the remainder of this paper the Movima external and the Tagalog nominative argument will for both languages be labelled “subject”, and the internal argument of Movima and the genitive-marked argument of Tagalog transitive clauses will henceforth be labelled “non-subject argument”.

As was mentioned above, the occurrence of the inverse voice in order to permit the extraction of the actor is much less productive in Movima than the use of the actor voice in Tagalog, however. This is because the Movima direct-inverse alternation is heavily governed by hierarchies of person and animacy (Haude 2014; Haude In press): when the referents of the two arguments of a transitive clause are on different positions in the hierarchy (e.g. when the clause describes an event with a human and an inanimate participant), the higher-ranking event participant must be encoded as the internal argument. Hence, the direct-inverse opposition cannot be fully exploited to allow relativization and other extraction operations: when the actor outranks the undergoer, the actor cannot be encoded as the external argument, and, therefore, cannot be extracted. Movima has an additional, detransitivizing operation for this purpose, which functions as an antipassive.<sup>7</sup> In (21), there is a human actor (‘people’) and an inanimate

---

<sup>6</sup> Proper-noun undergoers in the active voice are automatically marked as dative (see e.g. Latrouite 2011: 66). The change of the verb root *tuksu* to *nuksu*, with a homorganic nasal, results from morphophonological assimilation (S. Riesberg, p.c.).

<sup>7</sup> See Haude (In press) for a detailed analysis of the conditions of choice between antipassive and inverse.

undergoer ('the church'). The actor DP *is juyeni* is relativized. The relativizing particle *di'* is followed by the particle *kwey*. The presence of this particle has the effect that the predicate of the relative clause becomes intransitive, with the extracted argument as its S argument. The undergoer is encoded by an oblique phrase, marked by the prefix *n-*.

- (21) *[is juyeni] ... di' kwey pul-na n-as ele:siya*  
 ART.PL person REL DETR sweep-DR OBL-ART.N church  
 '(the) people (who work and) who sweep the church' [EAO\_Summary 015]

Thus, both in Tagalog and in Movima only the subject can be extracted, and the choice between different transitive constructions allows the extraction of the subject independently of its semantic role. However, the restrictions imposed by the person/animacy hierarchy in Movima distinguish the Movima system from the Tagalog one.

## 2.4 Summary

The formal similarities between Tagalog and Movima transitive clauses that were described above can be schematized as in (22). In both languages, there is one argument that is internal to the predicate phrase and one, the 'subject', that is external to it (aligning with the single argument of an intransitive clause). In (22a), it is shown that the Tagalog undergoer voice and the Movima direct voice assign the undergoer (U) role to the subject and the actor (A) role to the non-subject argument. In (22b), it is shown that the Tagalog actor voice and the Movima inverse voice assign the A role to the subject and the U role to the non-subject argument (although in Tagalog the assignment of the U role to the genitive argument is less systematic than in Movima: in some cases, the undergoer of the actor voice is encoded in the dative case; see (20b)).

- (22) a. [[Verb<sub>UV/DR</sub>] [A]]<sub>PRED</sub> [U]<sub>SUBJECT</sub>  
 b. [[Verb<sub>AV/INV</sub>] [U]]<sub>PRED</sub> [A]<sub>SUBJECT</sub>

Since the external argument shares its syntactic properties with the single argument of the intransitive clause, the voice alternation in both languages can be described as leading to a split-alignment pattern: in the Movima direct and the Tagalog undergoer voice, the undergoer argument aligns with the single argument of the intransitive clause, therefore showing an ergative pattern (A vs. S/U). In the actor/inverse voice, by contrast, the actor is encoded like

the single argument of the intransitive clause, so this voice shows an accusative alignment pattern (U vs. S/A).

At the same time, and this is a further parallel, despite the structural equivalence between the two transitive constructions, both languages show a bias towards the ergative construction. In both languages the voice that assigns the undergoer role to the subject is more frequent in natural discourse (Hopper & Thompson 1980; Himmelmann 2005: 368; Haude 2014). In Movima, the direct voice can be considered the default transitive construction, as it can override the person/animacy hierarchy and is preferred when both event participants are equally ranked (Haude 2014); furthermore, as mentioned above, to extract an actor that outranks the undergoer, a detransitivizing (antipassive) operation is chosen rather than the inverse (Haude In press). In Tagalog, in turn, the actor voice shows signs of being less transitive (Nagaya 2012) (see also the appearance of the dative in (20b)).

The following section deals with a further property shared by both languages: the syntactic flexibility of nouns and verbs.

### **3 Syntactic flexibility of nouns and verbs**

#### **3.1 Nouns and verbs as lexical categories**

Both in Movima and in Tagalog, verbs and nouns are different lexical categories. In Tagalog, verbs are distinguished from nouns in that they are marked for aspect, voice and mood (Schachter & Otnes 1972: 65; Himmelmann 2005: 362). In Movima, verbs are not systematically marked for the typical verbal categories (tense, modality, aspect, evidentiality), but there are some morphemes that are exclusive to verbs and others that are exclusive to nouns (Haude 2006: 108–109). For instance, only nouns can be combined with a verbalizer *-tik*, which derives an intransitive verb that denotes an event typically associated with the entity denoted by the noun. Nouns can be incorporated into verbs, but verbs cannot be incorporated (at least not in the same way). A good rule of thumb is that all Movima nouns, independently of their morphological shape, can in principle be marked as possessed by an internal enclitic, while verbs can be combined with an internal enclitic only if they are overtly marked as transitive by a direct or inverse marker. Furthermore, in adverbial and complement clauses, verbal predicates are marked differently than nominal ones: nominal predicates undergo reduplication, while verbal predicates receive a suffix *-wa* (Haude 2011b). Adjectives form an intermediate category with some tending to be more verb- and others more noun-like (Haude 2006: 112–113). Most



adjectival predicates in subordinate clauses are, like nouns, reduplicated, and can therefore be analyzed as nominals. Therefore, while the differences especially between intransitive verbs, unpossessed nouns and adjectives are often not immediately apparent, it is usually possible to assign a given content word to a lexical category.

### 3.2 Nouns as main-clause predicates

While verbs and nouns thus have different morphological properties both in Movima and in Tagalog, syntactically the distinction between these lexical classes is limited. In both languages a noun can function as a predicate and a verb can occur as the lexical element of a DP. The following examples illustrate intransitive clauses with nominal predicates in Movima. In parallel to the examples (1)–(3) above, example (23) is an intransitive clause whose subject is not overtly expressed (represented here as  $\emptyset$ ); in (24), the subject is expressed by a full DP; and in (25), the subject is a bound pronoun. There is no copula: the noun alone functions as predicate.

(23) *jayna n-os potmo-wa=us, jayna rulrul [∅]*  
 DSC OBL-ART.N.PST get\_up-NMZ.EVT=3M.AB DSC jaguar  
 ‘Then when he got up, (he was a) jaguar already.’ [LYO\_250808\_2 170]

(24) *bo ja’ jutpa [is maniwanra=is]*  
 REAS just arrow ART.PL weapon=3PL.AB  
 ‘... because their weapons (were) just arrows.’ [HRR\_120808-602]

(25) *bo rey rulrul[--as] rey*  
 REAS EPIST jaguar--3N.AB EPIST  
 ‘... because it (was) a jaguar.’ [PMP\_HRR\_etal\_210908 169]

A notable exception to the possibility of Movima nouns to function as predicates involves possessed and proper nouns. While it is possible for a possessed or proper noun to function as the predicate of a basic clause, as in (26), this is generally avoided. Example (26) is one of the rare corpus examples of a possessed noun as predicate.

- (26) *lat rey lavabat-a=as [os be~bet-kwa] jayna*  
 EV EPIST shade-LV=3N.AB ART.N.PST RED~skin-ABST DSC  
 ‘The hide (was) its soul (lit. shade), you see.’<sup>8</sup>  
 [HRR\_120808-tigregente 232]

The structure of the sentence in (26) is parallel to that of a transitive clause, in which one argument is expressed by an internal enclitic and the other one by a phonologically independent DP. However, it is not possible to express the external argument of the possessed nominal predicate by an enclitic pronoun, which is possible with verbal predicates (see (3) and (7)) and with unpossessed common nominal predicates (see (25)). Instead, the argument of a clause with a possessed nominal predicate can be expressed by a fronted pronoun, as in (27), which is the standard way to form equational clauses.<sup>9</sup>

- (27) *asko lavabat-a=os Buscha*  
 PRO.3N.AB:COP shade-LV=ART.N.PST Buscha  
 ‘It was Buscha’s soul.’ [HRR\_120808-tigregente 597]

In Tagalog as well, basic clauses can have a noun as their predicate, as in (28). As Schachter & Otnes (1972: 95) state, “[a]n unmarked noun is a nominal of invariant form that may occur [...] in predicate position.” Furthermore, in Tagalog, just as in Movima, “an unmarked noun in predicate position usually expresses either an INDEFINITE or a GENERIC meaning: respectively, ‘a member of the class \_\_\_’ or ‘the class \_\_\_’” (Schachter & Otnes 1972: 95; capitalization in original).

- (28) Tagalog nominal predicate (Schachter & Otnes 1972: 61; glosses added)  
*artista [ang babae]*  
 actress NOM woman  
 ‘The woman (is) an actress.’

<sup>8</sup> The example stems from a mythological story about a person who transforms into a jaguar; the person is referred to by the “non-human” pronoun =*as* here.

<sup>9</sup> The free pronoun is the syntactic predicate in this construction and can be analyzed as a copula (Haude 2018a; Haude 2019).

In Tagalog, also proper nouns can occur as predicates (Schachter & Otones 1972: 93–94); however, this seems to be uncommon, and in that case the noun tends to be preceded by the nominative marker (in (29) *si*, the nominative marker for proper names).<sup>10</sup> This is something that never happens in Movima, where a main-clause predicate is defined by the absence of a determiner.

(29) Tagalog proper noun as predicate (Schachter & Otones 1972: 94)

*Si Rosa [ang paborito kong kklase].*

NOM proper\_name NOM favourite my class

‘My favorite classmate is Rosa.’

Thus, in both languages, nonverbal predicates tend to be non-specific. This is actually not surprising: proper names and definite descriptions – which possessed nouns often are – do generally not make for good predicates (Frege 1892; see also Launey 2004: 242; Kaufman 2018: 209).

### 3.3 Verbs in DPs

This section illustrates the parallel noun-verb status in DPs in Movima and Tagalog. While in both languages, a typical DP contains a noun, also a verb can occur inside a DP without any morphological modification.

In Movima, the interpretation of a “verbal DP” is absolutely predictable: the DP refers to the event participant that would be encoded as the subject of the corresponding main clause. The examples below show Movima DPs containing an intransitive (30), a transitive direct (31), and a transitive inverse verb (32). In the case of transitive verbs, the role of the participant – actor or undergoer – is specified by the direct or inverse marking. Consequently, a DP with a direct verb (31) refers to the undergoer, a DP with an inverse verb (32) to the actor. A DP with an intransitive verb, like *jiwa:wa* in (30), refers to the participant that would be encoded as the single argument of the verb in predicate function (see (3) above).

---

<sup>10</sup> Without the noun-phrase marker, the word *Rosa* would simply denote the name, but would not refer to a person (Schachter & Otones 1972: 94).

(30) *treynta [is ji<wa:~>wa ney]*  
 thirty ART.PL come<MD~> here  
 ‘Thirty (people were who) came here.’ [CCT\_120907\_1 013]

(31) *bo to:mi [os yey-na=us jayna]*  
 REAS water ART.N.PST want-DR=3M.AB DSC  
 ‘Water was what he wanted then.’ [EAO\_240807\_vbr 133]

(32) *ji<wa:~>wa [os rey joy-te-kay-a=us ney]*  
 come<MD~> ART.N.PST MOD go-CO-INV-LV=3M.AB here  
 ‘The one who took him with him came here.’ [CCT\_120907\_1 100]

Note that verbs in DPs only characterize the event participant within the scope of the actual instantiation of the event. When an action is carried out habitually or professionally, the agent nominalizer *-pa* is used (Haude 2006: 475-477), as illustrated in (33) with an intransitive and in (34) with a direct-marked, transitive verbal base.

(33) *ittijkarim-pa*  
 1.INTR work-NMZ.AG  
 ‘I am a worker; I habitually work.’ [NAO\_FSG\_300706\_1 337]

(34) *istolkara-na:-pa*  
 ART.PL rob-DR-NMZ.AG  
 ‘(the) thieves; (the) (ones who) habitually steal’ [JMH\_160806\_1 146]

Also in Tagalog, voice-marked verbs are uniquely interpretable as referring to the event participant indicated by the voice marker. As Himmelmann puts it, “it is clear that the subject of an actor voice form is an agent of some kind and that the action denoted by the voice-marked form is in some way related to the THING or ACTION denoted by the root” (Himmelmann 2008: 288–289; small caps in original). As in Movima, “in this use it is also a specific instance of the action denoted by the root that is being referred to by the voice-marked form” (Himmelmann 2008: 287), i.e. the referent of the DP is characterized just by its participation in the particular event. The nominal use of Tagalog verbs is illustrated in (35a) and (35b) for the actor voice and the patient voice, respectively (Kaufman 2009a: 5).

(35) Tagalog verbs in DPs (AV and PV)

- a. *ang b<um>ili*  
NOM <AV.PFV>buy  
'the buyer/one who bought'
- b. *ang b<in>ili*  
NOM <PV.PFV>buy  
'the bought (thing)'

The use of a verb inside a DP in Movima usually goes together with the use of a nonverbal predicate, as in (30)–(31) above, the effect being that of argument focus (see Haude 2021). Propositionally, a clause with a nonverbal predicate and a DP containing a verb are identical to a canonical basic clause. Consider the basic clause in (36), whose propositional content is similar to that of (31) above.

- (36) *yey-na=us [is mo'incho]*  
want-DR=3M.AB ART.PL manioc\_beer  
'He wanted manioc beer.' [LTC\_020906\_5 381]

The pragmatic effect described above for Movima seems to be similar in Tagalog. As Schachter & Otones (1972: 151) put it, "[s]entences with nominalizations in topic position often involve a contrastive meaning. The element being contrasted is expressed in such sentences by the predicate."<sup>11</sup> See (37) for illustration.

- (37) Tagalog (adapted from Schachter & Otones 1972: 151; gloss of *na-kita* from Latrouite 2011: 67)
- Hindi, si Rosa [ang na-kita niya].*  
no NOM Rosa NOM UV.REAL-visible PRO.3.GEN  
'(Did he see Maria?) No, it's Rosa he saw.'

---

<sup>11</sup> Schachter & Otones (1972) analyze the occurrence of a verb inside a DP as a nominalization.

A difference between Tagalog and Movima is that when a Tagalog verb root, i.e. a form without voice marking, occurs inside a DP, it may denote a state or an event, and not necessarily (as in Movima) a participant in the state or event. A Tagalog DP containing a bare verb (i.e. a root) is interpreted according to the kind of predicate it occurs with. A root denoting an event, for instance, “may be used as the subject of predicates which denote the manner in which an event/action took place” (Himmelmann 2008: 278), as in (38). This is never the case in Movima.

(38) Tagalog (English 1986, cited by Himmelmann 2008: 278)

*bigla-an* [ang kanyá-ng alis]  
 sudden-? NOM 3.SG.DAT-LK departure  
 ‘His departure (act of leaving) was sudden.’

In Movima, by contrast, a bare verb in a DP never denotes an event. In order to denote an event or state, a predicate must be overtly morphologically derived; verbs are derived with the ‘event nominalizer’ *-wa*, as in (39).

(39) *chonsi* [os lat joy-**wa**=is]  
 noon ART.N.PST EV go-NMZ.EVT=3PL.AB  
 ‘They left at noon’, lit. ‘Their leaving was (at) noon.’  
 [HRR\_2009\_tape1\_B 323]

Thus, the interpretation of a DP containing a verb is straightforward in Movima. The DP always refers to the event participant that is the subject of the same verb in predicate function: the undergoer of a direct-marked verb, the actor of an inverse-marked verb, and the single participant of an intransitive verb, i.e. either an actor or an undergoer depending on the verb’s morphological and/or semantic properties (see Haude 2012).

While the parallels between Movima and Tagalog syntax are striking also in this regard, there are also some differences. First of all, the lexical flexibility in Tagalog is weaker than in Movima: unlike Movima, Tagalog differentiates the referential vs. predicative use of some words through a stress shift (Himmelmann 1991); and in Tagalog, the meaning of a verb unmarked for voice in referential function is less predictable than in Movima, as its interpretation may depend on the context.

## 4 A nominalist account for Movima?

### 4.1 The “nominalist hypothesis”

Kaufman’s analysis of Tagalog follows the tradition of analyzing Tagalog verbs as nominal elements, which started early in the 20<sup>th</sup> century (see Sasse 2009: 169). Under this view, “all predication in conservative Austronesian languages [is] inherently copular ... and all predicates [are] inherently nominal” (Kaufman 2009a: 5). Also Schachter & Otnes (1972: 62) describe the structure of a basic clause in Tagalog as comparable to that of an equational (or copular) clause in a language like English. Even though there is no complete consensus on this perspective, analyzing Tagalog clause structure in this way makes sense:

- if both verbal and nominal predicates are analyzed as forming equational clauses, no bias towards one formally more “basic” construction has to be assumed;
- there is no need to analyze a verb inside a DP as either a (zero marked) relative clause or a (zero marked) nominalized form;
- the extraction restrictions can be explained by the equivalence of the non-subject argument of a transitive clause and a nominal possessor;
- taking verbal voice markers as indicators of the semantic role of the subject – similar to active or passive participles – accounts for the existence of multiple “transitive” constructions (see below).

Under this analysis, then, a basic Tagalog undergoer- (here, patient-) voice clause like the one in (40a), and a basic actor-voice clause as in (40b), can be paraphrased with nominal constructions (Kaufman 2009a: 6).

(40) Tagalog nominal interpretation of voice-marked verbs

a. UV: *k<in>áin-u nang=púsa [ang=dagà]*

<PV.PFV>eat-PV GEN=cat NOM=rat

‘The rat was eaten by the cat’, lit. ‘The rat was the eaten one of the cat.’

b. AV: *k<um>áin nang=dagà [ang=púsa]*

<AV.PFV>eat GEN=rat NOM=cat

‘The cat ate a rat’, lit. ‘The cat was the eater of a rat.’

The analysis of Austronesian clause structure as being basically equational receives support from evidence that at least some Austronesian voice markers originate from nominalizations or are cognate with nominalizing morphology (Starosta, Pawley & Reid 1982; Ross 2002: 38). Hence, historical and comparative evidence suggests that the nominal appearance of Austronesian voice forms may stem from the reanalysis of nominalizations as canonical predicates.

Sasse (2009: 168) points out that this tendency is more widespread than one might think. Based on historical data from Eastern Aramaic, he argues that the use of participles (which can be considered participant-oriented nominalizations) is a convenient means to “freely distribute content words around in both argument and predicate positions”, since participles “are not associated with a prototypical ontological category”. When this is taken to an extreme, he argues, a language may do away with its finite predicates altogether: “[i]f a language ... abandons finite verb forms altogether and remains with participles and verbal nouns constructed according to the ‘nominal sentence’ type ... such a language would be able to freely distribute content words around in both argument and predicate positions, resulting in a very weak noun/verb distinction. This is precisely the situation we find in Tagalog”. Sasse states that usually, however, the development does not go that far, since languages tend to make a grammatical distinction between lexical classes that denote different ontological categories. Therefore, Sasse regards the Austronesian pattern as “unique” (Sasse 2009: 179) in that respect. With what we now know from Movima, the western Austronesian voice system may still be exceptional, but it can no longer be said to be unique.

Since Movima is an isolate, any attempt to explain the origin of its direct-inverse system must remain hypothetical. However, a “nominalist account” is clearly possible for Movima as well.<sup>12</sup> If Movima verbs are analyzed as nominals, no difference needs to be made between verbal and nominal predicates. Voice marking on bivalent verbs indicates the orientation of the (nominal) predicate towards the actor or the undergoer. Therefore, Movima clauses can be paraphrased in the same way as the Tagalog examples in (40). As an illustration in parallel to the Tagalog examples in (40), example (41) shows the nominal paraphrase of the direct transitive clause cited in (5) above, and (42) shows the nominal paraphrase of the inverse transitive clause in (15) above.

---

<sup>12</sup> This suggestion for Movima was first made by W. Drossard (p.c.) in 2005, who referred us to crucial publications on this point (Himmelman 1991; Sasse 1991).



(41) *tikoy-na=us [os rulrul]*  
 kill-DR=3M.AB ART.N.PST jaguar  
 ‘He killed the/a jaguar’, lit. ‘The/a jaguar was what he killed/his killed one.’  
 [PMP\_HRR\_etal\_210908 085]

(42) *bu’ni yok-kay-a=us [os rulrul]*  
 perhaps catch-INV-LV=3M.AB ART.N.PST jaguar  
 ‘Perhaps he got caught by a jaguar,’ lit. ‘... a jaguar was what caught him/his catcher.’  
 [EAO Jaguar 213]

In the following two subsections I turn to the question of whether there is any evidence that the Movima direct and inverse markers originate from nominalizers, as has been claimed for the Tagalog voice markers. I will show that intransitive verbs and verbs without any voice marking can be used in DPs (4.2), which means that the nominal interpretation of Movima verbs does not depend on their morphology. Still, a case might be made for the direct marker *-na*, but that even here the evidence is very limited (4.3).

#### 4.2 Intransitive verbs

Before arguing for a possible nominalizing origin of the direct marker 4.3, it is necessary to point out that most Movima verb roots cannot be combined with a direct or inverse marker at all. Direct and inverse marking is only possible for the class of “inherently bivalent roots” (Haude 2006: 339–340) or for bases derived by a valency-increasing marker such as causative, benefactive or applicative (Haude 2006: 348). All other verb bases are of the “inherently monovalent” type. The formal diagnostic for monovalent bases is that the addition of a suffix *-na* does not derive a direct transitive verb, but a place noun (see also 4.3). Consider, for instance, the monovalent verb base *tijka:rim* ‘work’ combined with *-na* in (43). Place nouns derived with *-na* are possessed and therefore, cannot function as the predicate of a basic clause with a pronominal enclitic, but must be combined with the free pronoun that acts as copula (see 3.2).

(43) *no-kos tijkarim-na=’ne*  
 OBL-ART.N.AB work-NMZ.LOC=3F  
 ‘at/of her work place’ [EAO Dialogue 056]

All verbs that are not marked as direct or inverse, be they semantically mono- or bivalent, are syntactically intransitive. Their orientation, i.e. the role of their argument as either actor or undergoer, is usually inferable from their morphology. Bivalent roots can take not only direct or inverse, but also a number of other voice markers (Haude 2012), as shown in the paradigm in Table 1.

Table 1. Movima voice markers (A=actor, U=undergoer)

| transitivity | marker         | meaning              | role of subject |
|--------------|----------------|----------------------|-----------------|
| transitive   | <i>-na/-a-</i> | direct               | U               |
|              | <i>-kay</i>    | inverse              | A               |
| intransitive | <i>-chet</i>   | reflexive/reciprocal | A+U             |
|              | <i>-'i</i>     | resultative          | U               |
|              | <i>-ete</i>    | agentive             | A               |
|              | <RED~>         | middle               | A+U             |

Any verb marked by one of the suffixes in Table 1 can occur in a nominal environment. In (44), this is shown for the bivalent root *bat-* ‘put’ combined with the ‘agentive’ suffix *-ete*, and in (45), it is shown for the same root with the ‘reflexive/reciprocal’ voice *-chet*.

- (44) *kas rey so<kak><te~>t-e=[us rey bat-e:te]*<sup>13</sup>  
 NEG.COP EPIST other<IRR><NMZ.ST~>-CLF.person=ART.M EPIST put-AGT  
 ‘It wasn’t only one man who contributed (i.e. there were several).’  
 [GBM\_Ganado 107]

- (45) *jaysot sinko suwe:ro [is bat-chet]*  
 seem five drip ART.PL put-R/R  
 ‘It seemed like five drips were sitting there (on my body, after a surgery).’  
 [EAO\_Cbba 219]

Interestingly, the suffix *-'i*, which marks the resultative voice, can be analyzed as a phonologically conditioned dummy suffix, and verbs with this suffix can be interpreted

<sup>13</sup> The argument of a subordinate, nominalized predicate is always internally cliticized, also when the predicate is intransitive.

alongside with unmarked bivalent verb bases, which denote a resultative state (Haude 2012), such as *ji:sa* ‘made’ in (46). These verbs can occur inside a DP just as well as their direct-marked counterparts, as shown in this example.

- (46) *beyra [kis ji:sa-na=i], beyra [kis ji:sa]*  
 little ART.PL.AB make-DR=3PL little ART.PL.AB make  
 ‘Little (pottery) is what they make, little is made.’ [CCT\_120907\_1 011]

Also monovalent bases, which do not participate in the voice alternation, can occur in nominal environments, like any lexical element in Movima. The referent of the nominal expression is the single event participant, or more precisely, the participant that would be encoded as the single argument of the same verb in predicate function. For instance, the argument of the verb *tijka:rim* ‘work’, which is an active verb, is the actor (i.e. the working person), and this is also the role of the referent of a DP containing this verb, as shown in (47). No other interpretation (e.g. as an event nominalization) is possible.<sup>14</sup>

- (47) *bo so<kak>t-e [kinos ney tijka:rim]*  
 REAS other<IRR>-CLF.person ART.F.AB DEF work  
 ‘Because this one who works is the only one.’ [PMC\_Pere 010]

Thus, in Movima, voice markers are not necessary to permit the use of a verb in a nominal domain. And in Tagalog, too, verb roots can occur without voice marking, and unmarked verb roots denoting two-participant events can denote the result of an action (Himmelman 2008: 275). Unlike Movima, where verbs with and without voice marking denote an event participant, unmarked Tagalog roots denoting one-participant events can denote the event itself (see (38)), and many Tagalog verb roots can convey more than one meaning. The crucial commonality between Tagalog and Movima, however, is that also in Tagalog, roots unmarked for voice behave grammatically “like all other content words. That is, in principle they fit into every position open to content words” (Himmelman 2008: 275).

---

<sup>14</sup> See Haude (2019: 237) on a few but partly systematic exceptions, which include incorporating verbs: they undergo a regular shift in orientation when occurring in a nominal environment. For no verb, however, is there any restriction to occur in a nominal environment.

It can be concluded from this that the possibility to use a verb in a DP is independent of its morphological marking. Still, in the quest of a potential nominalizing origin of Movima direct-inverse marking, the following section will zoom in on one case in which the direct marker can be traced back to a straightforward nominalizer.

### 4.3 The direct marker *-na* as a nominalizer

Given that the Tagalog voice markers are said to originate, at least in part, from nominalizers, it is tempting to also assume a nominalizing source for the Movima direct and inverse markers. Also synchronically, these markers can be analyzed as deriving a participant-oriented nominal, similar to the suffixes *-ee* and *-er* in English. However, there is no clear evidence for such an origin. Only for the direct suffix *-na* there is a (weak) sign that it might be related to a nominalizer: there is one word that contains the suffix *-na* and denotes the undergoer of a two-participant event, but it must synchronically be classed as a possessed noun: like a possessed noun, but unlike a transitive verb, it cannot occur as a predicate whose subject is expressed by a pronominal enclitic (see 3.2).

This is the word *ka:na* ‘food; what is eaten’. It can be assumed to be historically composed of a root *\*ka:-* ‘eat’ (cf. the synchronic verb root *kay-* ‘eat’) and a suffix *-na*, which derives a noun denoting the undergoer of the event. That this word is a noun and not a direct-marked transitive verb is evident from the fact that it only occurs in nominal domains, such as the DP in (48). As with all possessed nouns (see (27) above), the only possibility for this word to function as a predicate with a pronominal subject is in combination with the pronominal copula, as in (49).

- (48) *ji:sa-ti*            [*kos kana=a*]  
 make-IMP.DR ART.N.AB food=3N  
 ‘Prepare its (= the pet’s) food!’ [JGD\_130907-13 069]

- (49) *asko*            *kana=’ne*  
 PRO.3N.AB:COP food=3F  
 ‘That was her food (or: what she ate).’ [EAO Basket 006]

Furthermore, like nouns, but unlike (voice-marked) verbs, the word *ka:na* can be incorporated into a verb, as in (50).

- (50) *ona loy sal-a-ka:na*  
 let's<sub>see</sub> ITN look\_for-DR-food  
 'Let's see, (I)'ll look for food.'<sup>15</sup> [EGA\_Dialogue 109]

The synchronic verb with the meaning 'eat' in Movima is *kay~kay*, whose root is monovalent and which behaves like any other verb containing the 'middle' reduplication (see Table 1): it is intransitive and actor-oriented, i.e. when used as a predicate, as in (51), the single argument represents the actor. Likewise, when this verb occurs in a nominal environment, as in (52), it refers to the actor. The undergoer can only be expressed as an oblique phrase, as in (53).

- (51) *kay~kay [is wa:ka]*  
 MD~eat ART.PL cow  
 'The cows eat.' [GCM\_290806\_4 089]

- (52) *wo'ray [is kay~kay]*  
 bat ART.PL eat~MD  
 'Bats (are the ones who) eat.' [EAO\_Wo'ray 004]

- (53) *jayna kay~kay[--is] n-is is~is-ra*  
 DSC MD~eat--3PL.AB OBL-ART.PL RED~roast-CLF.meat  
 'Then they ate roasted meat.' [HRR\_120808-tigregente 548]

As with other monovalent roots, such as *joy-* 'go', the suffixation of *-na* to the root *kay-* derives a locational noun, as in (54).

- (54) *a'ko lat joy-na=is, a'ko kay-na=is*  
 PRO.3N:COP EV go-NMZ.LOC=3PL.AB PRO.3N:COP eat-NMZ.LOC=3PL.AB  
 'That (is) where they went, that (is) where they ate.' [EAO\_Llamada hija 023]

Since the roots *\*ka:-* and *kay-* 'eat' are most probably related, it may be the case that the direct suffix was first of all a nominalizer that derived words that could only occur in nominal

---

<sup>15</sup> Verbs with an incorporated undergoer argument are intransitive, and their subject can remain unexpressed. Here, the first person is understood from the context.

environments.<sup>16</sup> The form *ka:na* ‘food/something to eat’ can be regarded as a fossilized nominal that never made it to main-clause predicate status, while an intransitive verb took over the predicative function.

Thus, it remains an open question whether voice marking in Movima necessarily had its source in nominalization in order to have arrived at the present-day system. Maybe there are simply languages in which there is no categorial correlation between lexical categories and syntactic functions (see Himmelmann 2008: 263).<sup>17</sup> The orientational morphology of Movima that resembles Tagalog voice marking may be an effect of the importance of referential hierarchies for Movima argument encoding.

## 5 Summary and conclusion

Movima, a linguistic isolate from the south-western Amazon, has morphosyntactic properties that are rather uncommon cross-linguistically, but that show noteworthy similarities with phenomena seen in some languages from other, non-related and geographically remote families, especially Tagalog (western Austronesian). Both Tagalog and Movima have more than one transitive construction, in which verbal “voice” morphology assigns the semantic role(s) to the subject (or, in Movima, both) argument(s). Only one of the arguments of a transitive clause (the “subject”) can be extracted, e.g. relativized, while the other argument is encoded inside the predicate phrase, in the same way as a nominal possessor. According to Kaufman (2009a), this phrase-structural patterning is precisely the reason for the extraction restriction in Tagalog, an analysis that can without any problem be extended to Movima.

Also beyond the parallel encoding of nominal possessor and the non-subject argument in a transitive clause, in both languages lexical categories are syntactically flexible: nouns can function as predicates, and verbs can occur in DPs without any morphological modification.

---

<sup>16</sup> The locational nominalizer *-na* in (47) is probably not directly related to the direct-voice suffix: When a locational noun is derived from a bivalent verb – which is done with the complex form *-kwina* –, the verb base retains the direct marker (e.g. *dewaj-na-kwina=Ø* [see-DR-NMZ.LOC=1SG] ‘where I saw (you/him/her etc.)’ (see Haude 2006: 399–401).

<sup>17</sup> A similar situation has been described for other native American languages such as Salishan languages (Jelinek & Demers 1994), Classical Nahuatl (Launey 1994; Launey 2004), some Tupi-Guaranian languages (Queixalós 2006; da Cruz & Praça 2019) and Enlhet (Kalisch 2010). Potentially there are more languages in South America of this type that have not been studied in sufficient depth yet.

The voice marking on (in Movima, transitive) verbs indicates which event participant is denoted by the verb.

The lexical flexibility of verbs and the function of the voice markers invites the hypothesis that in both languages, the voice markers originate from oriented participant nominalizations, i.e. nominalizations that denote a participant in a particular event, similar to participles. Historical and comparative research on Austronesian languages has suggested that some voice markers in this language family are indeed historically related to nominalizers, but given the isolated status of Movima and the nearly inexistent historical documentation, it is not to be expected that any such evidence will ever be found for this language here.<sup>18</sup>

There are some important differences between the Movima and the Tagalog syntactic systems. The choice between the transitive voices is strongly determined by hierarchies of person and animacy in Movima, which is not the case when it comes to the choice between actor- and undergoer voices in Tagalog. Interestingly, due to this hierarchy, Movima has a stronger ergative and less symmetrical syntax than Tagalog: the inverse cannot be as freely used for extraction of the actor as the Tagalog actor voice, because the inverse may only be used if the undergoer outranks the actor on the person/animacy scale. In order to extract a high-ranking actor argument, Movima uses an antipassive operation, which does not exist in Tagalog.

Another difference lies in the discourse use of the voice alternation. The Tagalog nominative argument usually encodes a definite referent, while in Movima, a definite referent (being of higher discourse topicality) is encoded as the internal rather than as the external argument. More research is needed to explain this fundamental difference in two systems that are otherwise so similar.

We hope to have shown how useful it is to study lesser-known, endangered, and especially isolated languages: the more languages we have detailed knowledge about, the better we can see that what is deemed “exotic” may not be so rare after all, and the parallels that can be found in genealogically and areally unrelated systems can help us identify the degrees of variation that occur among such systems.

---

<sup>18</sup> It might be worthwhile, though, to investigate the Movima lexicon in more detail. Interestingly, many if not most content words show signs of being historically composite (see Haude 2006: 351 for examples). This suggests that Movima lexical items have undergone a significant development in the past, which may have blurred an earlier, more pronounced noun-verb distinction. .

## Symbols and abbreviations (based on the Leipzig Glossing Rules)

= internal cliticization; -- external cliticization; ~ reduplication; < > infixation; 1=first person; 3=third person; AB=absential; ABST=absolute state; AG= agent; AGT=agentive; ART=article; AV=actor voice; CLF=classifier; CO=co-participant; COP=copula; DAT=dative; DEF=definite; DETR=detransitivizer; DR=direct; DSC=discontinuous; EPIST=epistemic; EV=evidential; EVT=event; F=feminine; GEN=genitive; IMP=imperative; INAL=inalienable; INTR=intransitive; INV=inverse; IRR=irrealis; ITN=intention; LK=linker; LOC=locational; LV=linking vowel; M=masculine; MD=middle; N=neuter; NEG=negation; NMZ=nominalization; NOM=nominative; OBL=oblique; OBV=obviative; PFV=perfective; PL=plural; PRO=free pronoun; PST=past; PV=patient voice; REAL=realis; REAS=reason; RED=reduplication; REL=relativizer; R/R=reflexive/reciprocal; S=subject; SG=singular; ST=state; UV=undergoer voice.

## References

- Campbell, Lyle (ed.). 2019. *Language Isolates*. Oxon, New York: Routledge.
- 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* (Studies in Language Companion Series 90), 151–179. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- Cruz, Aline da & Walkíria Neiva Praça. 2019. Innovation in nominalization in Tupí-Guaraní languages: A comparative analysis of Tupinambá, Apyãwa and Nheengatú. In *Nominalization in Languages of the Americas* (Typological Studies in Language 124), 625–655. Amsterdam/Philadelphia: John Benjamins Publishing Company.
- DuBois, John W. 1987. The discourse basis of ergativity. *Language* 63(4). 805–855.
- Frege, Gottlob. 1892. Ueber Begriff und Gegenstand. *Vierteljahrsschrift für wissenschaftliche Philosophie* 16(2). 192–205.
- Grinevald, Colette. 2002. Nominal classification in Movima. In Mily Crevels, Simon van der Kerke, Sérgio Meira & Hein van der Voort (eds.), *Current Studies on South American Languages*, 216–239. Leiden: CNWS.
- Haude, Katharina. 2006. *A Grammar of Movima*. Nijmegen: Radboud University Doctoral dissertation. <https://repository.uibn.ru.nl/bitstream/handle/2066/41395/41395.pdf>.
- Haude, Katharina. 2011a. Tense marking on dependent nominals in Movima. In Renate Musan & Monika Rathert (eds.), *Tense Across Languages*. Berlin/Boston: De Gruyter. [http://halshs.archives-ouvertes.fr/docs/00/67/22/48/PDF/haude-TenseAcrossLanguages-revision\\_1.pdf](http://halshs.archives-ouvertes.fr/docs/00/67/22/48/PDF/haude-TenseAcrossLanguages-revision_1.pdf).
- Haude, Katharina. 2011b. Referring to states and events: subordination in Movima. In Rik van Gijn, Katharina Haude & Pieter Muysken (eds.), *Subordination in South American Languages* (Typological Studies in Language 97), 141–168. Amsterdam: John Benjamins Publishing Company. <https://doi.org/10.1075/tsl.97.06hau>.
- Haude, Katharina. 2012. Undergoer orientation in Movima. In Gilles Authier & Katharina Haude (eds.), *Ergativity, Valency and Voice* (Empirical Approaches to Language Typology 48), 159–287. Berlin/New York: De Gruyter. [http://halshs.archives-ouvertes.fr/docs/00/69/20/77/PDF/08\\_Haude.pdf](http://halshs.archives-ouvertes.fr/docs/00/69/20/77/PDF/08_Haude.pdf).



- Haude, Katharina. 2014. Animacy and inverse in Movima: a corpus study. *Anthropological Linguistics* 56(3–4). 294–314.
- Haude, Katharina. 2018a. Nonverbal predication in Movima. In Simon E. Overall, Rosa Vallejos & Spike Gildea (eds.), *Nonverbal Predication in Amazonian Languages* (Typological Studies in Language 122), 217–244. Amsterdam: John Benjamins Publishing Company. <https://halshs.archives-ouvertes.fr/halshs-01848221v1>.
- Haude, Katharina. 2018b. A topic-marking cleft? Analyzing clause-initial pronouns in Movima. In Evangelia Adamou, Katharina Haude & Martine Vanhove (eds.), *Information Structure in Lesser-Described Languages: Studies in Prosody and Syntax* (Studies in Language Companion Series 199), 217–244. Amsterdam: John Benjamins Publishing Company. <https://halshs.archives-ouvertes.fr/hal-01875573v1>.
- Haude, Katharina. 2019. Grammatical relations in Movima: alignment beyond semantic roles. In Alena Witzlack-Makarevich & Balthasar Bickel (eds.), *Argument Selectors: New Perspectives on Grammatical Relations* (Typological Studies in Language 123), 213–256. Amsterdam: John Benjamins Publishing Company. <https://halshs.archives-ouvertes.fr/halshs-02049736v1>.
- Haude, Katharina. 2021. Clefting and nominal predication: Two focus-marking constructions in Movima. *Faits de Langues* 52(1). 117–138. <https://doi.org/10.1163/19589514-05201006>.
- Haude, Katharina. In press. Between symmetrical voice and ergativity: Inverse and antipassive in Movima. *International Journal of American Linguistics*.
- Himmelman, Nikolaus P. 1991. *The Philippine Challenge to Universal Grammar* (Arbeitspapier 15 (Neue Folge)). Universität zu Köln: Institut für Sprachwissenschaft.
- Himmelman, Nikolaus P. 2005. Tagalog. In Alexander Adelaar & Nikolaus P. Himmelman (eds.), *The Austronesian Languages of Asia and Madagascar*, 350–376. London/New York: Routledge.
- Himmelman, Nikolaus P. 2008. Lexical categories and voice in Tagalog. In Peter K. Austin & Simon Musgrave (eds.), *Voice and Grammatical Functions in Austronesian Languages*, 247–293. Stanford, CA: CSLI Publications.
- Hopper, Paul J. & Sandra A. Thompson. 1980. Transitivity in grammar and discourse. *Language* 56(2). 251–299.
- Jelinek, Eloise & Richard A. Demers. 1994. Predicates and pronominal arguments in Straits Salish. *Language* 70(4). 697–736. <https://doi.org/10.2307/416325>.
- Judy, Judith E. & Robert A. Judy. 1967. Movima. In *Bolivian Indian Grammars* (Summer Institute of Linguistics Publications in Linguistics and Related Fields 16), 353–408. Oklahoma: SIL publications.
- Kalisch, Hannes. 2010. Los constituyentes de la cláusula enlhet (enlhet-enenlhet) Esbozo de una cláusula omnipredicativa. *Amerindia* 33/34. 109–150.
- Kaufman, Daniel. 2009a. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35(1). 1–49.
- Kaufman, Daniel. 2009b. Austronesian typology and the nominalist hypothesis. In Alexander Adelaar & Andrew Pawley (eds.), *Austronesian Historical Linguistics and Culture History: A Festschrift for Bob Blust*, 187–215. Canberra: Pacific Linguistics.
- Kaufman, Daniel. 2018. Austronesian predication and the emergence of biclausal clefts In Indonesian languages. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on Information Structure in Austronesian Languages*, 207–245. Berlin: Language Science Press.
- Kroeger, Paul. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford, CA: CSLI Publications.

- Latrouite, Anja. 2011. *Voice and Case in Tagalog: The Coding of Prominence and Orientation*. Düsseldorf: Heinrich-Heine-Universität Düsseldorf.
- Launey, Michel. 1994. *Une grammaire omniprédicative: Essai sur la morphosyntaxe du nahuatl classique*. Paris: CNRS Editions.
- Launey, Michel. 2004. The features of omnipredicativity in Classical Nahuatl. *STUF - Language Typology and Universals* 57(1). 49–69. <https://www.degruyter.com/document/doi/10.1524/stuf.2004.57.1.49/html>.
- Nagaya, Naonori. 2012. On the syntactic transitivity of Tagalog actor-focus constructions. *NINJAL Research Papers* 4. 49–76.
- Queixalós, Francesc. 2006. The primacy and fate of predicativity in Tupi-Guaraní. In Ximena Lois & Valentina Vapnarsky (eds.), *Root Classes and Lexical Categories in Amerindian Languages*, 249–287. Vienna: Peter Lang.
- Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. In Fay Wouk & Malcolm Ross (eds.), *The History and Typology of Western Austronesian Voice Systems* (Pacific Linguistics 518), 17–62. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National Univ.
- Sasse, Hans-Jürgen. 1991. Predication and sentence constitution in universal perspective. In Dietmar Zaefferer (ed.), *Semantic Universals and Universal Semantics*, 75–95. Berlin: Floris.
- Sasse, Hans-Jürgen. 2009. Nominalism in Austronesian: a historical typological perspective. Comments on Daniel Kaufman's "Austronesian Nominalism and its consequences." *Theoretical Linguistics* 35(1). 167–181.
- Schachter, Paul & Fé Otanes. 1972. *Tagalog Reference Grammar*. Berkeley, CA: University of California Press.
- Seifart, Frank & Harald Hammarström. 2019. Language isolates in South America. In *Language Isolates*, 260–286. Oxon, New York: Routledge.
- Starosta, Stanley, Andrew Pawley & Lawrence A Reid. 1982. The evolution of focus in Austronesian. In Amran Halim, Lois Carrington & Stephen A. Wurm (eds.), *Papers from the Third International Conference on Austronesian Linguistics, vol. 2: Tracking the travellers*, 145–170. Canberra: Pacific Linguistics.
- Zúñiga, Fernando & Seppo Kittilä. 2019. *Grammatical Voice*. Cambridge University Press. <https://doi.org/10.1017/9781316671399>.