

The category of predicates and predicate phrases in Tagalog: Commentary on Kaufman's article

JOSEPH SABBAGH

Kaufman's article contains three major claims:

- (1) a. Roots in Tagalog are category neutral, though all roots have an essentially nominal interpretation underlyingly.
- b. Tagalog lacks the category *v*.
- c. The "Subject-only" restriction on extraction in Tagalog is the result of a constraint that prohibits genitive case-marked DP's from appearing in predicate position.

Claim (1a) in conjunction with claim (1b) yields the interesting consequence that Tagalog has no verb phrase category. In line with some of the work with Distributed Morphology (Marantz 1997), Kaufman assumes that category neutral roots must ultimately be categorized by being embedded under a functional element such as *v*, *n*, or *a*. A root embedded under *v* will function syntactically as a verb and will be part of a verb phrase, while a root embedded under *n* will function as a noun and will be part of a noun phrases. If there is no *v* in the lexicon of Tagalog, as Kaufman argues, then it follows that there can be no verb phrases (*vP*). Instead, Kaufman claims that all putative VP predicates are, in fact, (simple) NP predicates.

In the discussion below, I will point out a few empirical issues relating to these claims, which seem to me to be worth further investigation with respect to Kaufman's main hypothesis.

1. The category and interpretation of roots

Claim (1a) is based on the observation that when morphologically complex words (namely, those which are typically analyzed as verbs) are

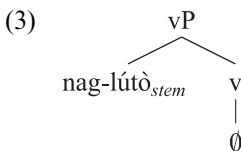
stripped away of their aspect and voice inflection, what is left behind is a root that can function on its own as a noun. Kaufman provides ample examples of this in his Table 3 (p. 9) (see also Himmelman 2008), and this list can easily be expanded upon. There are, however, at least two reasons to doubt that claim (1a) can accurately characterize the entire lexicon of Tagalog.

First, it is not clear whether all roots can function independently in the syntax. English's dictionary (English 1986) lists many roots alongside an affixed form with an identical meaning (e.g., *lakad/pag-lakad* ('walk')). However, English's dictionary also lists a number of roots that evidently occur only in an affixed form. Some examples are given in (2).

(2)

English	Tagalog	Root	Stem	Root Translation
'touch lightly'	<i>sumaláng</i>	(saláng)	pag-saláng	'light touch'
'hide'	<i>tumáago</i>	(tágo)	tagu-án	'hiding place'
'travel around'	<i>lumíbot/ maglibót</i>	(líbot)	pa-líbot	surroundings
'be in danger'	<i>mangánib</i>	(ngánib)	pang-ánib	'danger, crisis'
'adore, worship'	<i>sumambá</i>	(sambá)	pag-sambá	'worship, homage'

These examples raise a potential problem of the same nature that Kaufman points to with respect to the postulation of *v*. Specifically, Kaufman argues that if the lexicon of Tagalog contained *v*, which is responsible for the verbal category and eventive semantics, then one would need to explain why it evidently cannot attach to independent (i.e., bare) roots, but must instead attach to affixed stems. Considering forms like *nag-lútò* ('cook'), for instance, Kaufman assumes that *v* would have to attach outside of the stem as depicted in (3) (adapted from Kaufman's (15a)).



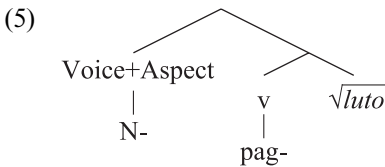
The same problem arises with the examples above where it appears that *n* must attach to affixed stems rather than directly to a root.

One possible solution to this problem would be to claim that the affixes which attach to the roots in the examples in (2) are themselves special spell outs of *n*. The roots in (2) could then be said to idiosyncratically select for such affixes, while other roots (such as *lútò* ‘cooked dish’) could be said to select for *n* which is phonetically null. A partial sketch of an analysis of this sort is given in (4).

(4) **Realizations of *n*:**

- a. $n \leftrightarrow /pag- / / \sqrt{salang}, \sqrt{samba}, \dots$ etc.
- b. $n \leftrightarrow / \emptyset / / \sqrt{lutò}, \sqrt{lakad}, \dots$ etc.

If one takes this approach, however, then a similar solution opens up for the cases of forms like *nag-lútò* ‘cook’. In particular, one could suppose that the morphology that makes up the stem is, itself, the spell out of the category *v*. Indeed, Travis (2000, 2002), Rackowski (2002), Aldridge (2004), and others take exactly this type of approach and propose that the affix *nag* is actually a morphologically complex form that consists of a underspecified nasal (N) which encodes (fused) voice and aspect and combines with the morpheme *pag-*, which is taken to be a morphological realization of *v*. For these authors, then, the representation of *nag-lútò* would be roughly as shown as in (5) (cf. (3)).



Now, just as one might claim that there are different realizations for *n*, depending on the nature of the root, one might likewise claim that there are different realizations for *v* depending on the nature of the root. Roots such as *lútò*, for instance, can be said to select for a *v* which is realized as *pag-*. Roots such as *saláng* and *tahímik* (see above), on the other hand, can be said to select for a *v* which is realized as \emptyset .¹

¹ In which case, *-um-* is the spell-out of Voice/Aspect.

In short, Kaufman's argument against *v* based on the claim that it would be difficult to imagine a principled reason why it cannot attach directly to roots is weakened by the fact that not all instance of his *n* are clearly capable of attaching to independent roots either.

A second issue concerning (1a) is that there are a number of roots that appear to be adjectival rather than nominal. Some examples are given below.²

(6)

English	Tagalog	Root	Root Translation
'uproot'	<i>bumaklás</i>	<i>baklás</i>	'uprooted'
'penetrate'	<i>bumaón</i>	<i>baón</i>	'embedded'
'dislodge'	<i>pumatós</i>	<i>patós</i>	'dislodged'
'get in front of'	<i>tumapát</i>	<i>tapát</i>	'in front of'
'keep quiet'	<i>tumahímik</i>	<i>tahímik</i>	'quiet'

Kaufman does suggest (p. 15) that roots may combine either with a *n* or with a *a*. It might therefore be stipulated that the roots in (6) are idiosyncratic in that they select to combine with *a* rather than *n*. While straightforward, this approach does have potentially significant implications for Kaufman's larger claims that all predication in Tagalog is inherently copular, and that all predicates are categorically DP.

Concretely, Kaufman's reasoning is that since all roots are interpreted nominally, it is straightforward to explain how this noun-ness is projected throughout the syntax eventually leading to the result that all predication ends up nominal. To put it in slightly different, more technical terms, since all roots combine with *n* (*v* being absent from the lexicon), the (functional) categories that the projection of *n* will be naturally allowed to combine with in its "extended projection" will be those such as D(et). Hence, all roots will ultimately project a DP structure. However, if roots

² There are evidently some roots that have either an adjectival or nominal interpretation/function (*takdá*' ('limitation' (n.)/'limited' (a.)). In some cases, the noun/adjective distinction can be marked by placement of stress.

such as those in (6) are restricted such that they must combine with *a* rather than *n*, it is not as straightforward that such projections will ultimately combine with functional categories such as *D*, given that *D* does not typically combine with adjectives.

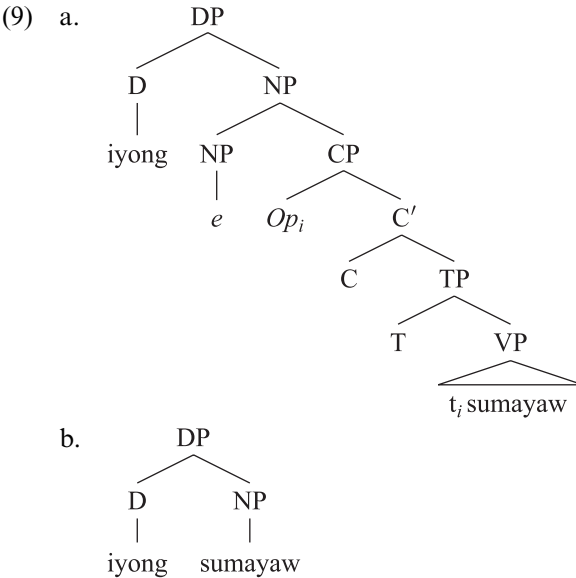
Two possibilities open up. One direction would be to say that Tagalog predicates come in one of two types, either adjectival (perhaps projections of DegP (see, e.g., Abney 1987)) or nominal (projections of DP). According to this view, the type of predicate will depend on the properties of the root. Alternatively, one might suppose that functional categories such as *D* in Tagalog are not selective about the category of the phrase that they combine with – i.e., whether it is *nP* or *aP*. If one takes this latter route, then much of the explanatory power of Kaufman’s original proposal would be lost since the projection of nominal syntax would no longer follow as a consequence of the properties of roots. The former approach therefore seems to be the most parsimonious with Kaufman’s proposal, but this conclusion seems to lead us away from the strong claim that he makes that all predicates in Tagalog are nominal in their underlying and surface form.

2. The category of the predicate phrase and extraction restrictions

Kaufman’s claim that Tagalog lacks *v* has implications for the controversial analysis of noun phrases like (7) (Kaufman’s example (22)). What is interesting about this and similar examples is that a putative verb (*s<um>ayaw*) is preceded by a case marker and various noun modifiers, though there is no overt noun in this structure with which these elements are associated. In what follows, I will refer to NP’s like (7) as “complex NP’s” to distinguish them from “simple NP’s” like (8) which are headed by unambiguous nouns.

- (7) *iyong (dalwa-ng) (maganda-ng) sumayaw*
that.LK two-LK beautiful-LK AV.BEG.dance
‘those (two) (beautiful) (ones who) danced’
- (8) *iyong (dalawa-ng) (maganda-ng) guro.*
that.LK two-LK beautiful-LK teacher
‘Those (two) (beautiful) teachers.’

One analysis of complex NP's, implicit in the work of those who assume a noun/verb distinction in Tagalog, involves positing a headless relative clause structure as shown in (9a). According to this analysis, the nominal adornments that appear optionally in (7) can be said to combine with the null NP head of the relative clause. Kaufman, however, argues that complex NP's have the simpler structure shown in (9b), in which *s<um>ayaw* is simply a noun which projects an NP structure. On this analysis, complex NP's are no different syntactically from simple NP's. Given this, the nominal adornments that can optionally appear in a complex NP like (7) can be said to combine directly to NP projected from *s<um>ayaw*.



Kaufman offers three arguments supporting an analysis complex NP's as simple NP's rather than as headless relative clauses. Rather than comment on these arguments directly, I would like to focus on a set of facts relating to the distribution of possessive pronouns within the NP which might be taken as an additional argument in support for analyzing constructions like (7) as simple NP's.

Consider first the examples in (10) of a simple NP containing a possessor argument. In (10a), a genitive marked possessive pronoun appears

post-nominally. In (10b), the possessor appears pre-nominally as an oblique marked pronoun.

- (10) a. ang guro ko
NOM teacher 1SG.GEN
'my teacher'
b. ang *aki*-ng guro
NOM 1SG.OBL-LK teacher
'my teacher'

The alternation between genitive and oblique possessor is quite general in Tagalog, and occurs in a number of other Austronesian languages as well.³ Important for our purposes, the alternation is also attested in complex NP's. Concretely, a "verb" in the patient voice (PV)-form which is contained in a complex NP can either be followed by a genitive pronoun which expresses the agent or may be preceded by an oblique form of the pronoun that expresses the agent.

- (11) a. ang *binili*- \emptyset ko
NOM BEG.buy-PV 1SG.GEN
'the thing I bought'
b. ang *aki*-ng *binili*- \emptyset
NOM 1SG.OBL-LK BEG.buy-PV
'the thing I bought'

This parallelism between simple NP's and complex NP's arguably suggests a similar syntactic structure.

Consider now a simple predicational sentences such as (12). Note that the "verb" predicate occurs in the patient voice (PV), and the agent is expressed as a genitive pronoun.

- (12) *Binili*- \emptyset ko ang sapatos.
BEG.buy-PV 1SG.GEN NOM shoes
'I bought the shoes.'

³ In Tagalog, this alternation only occurs with possessive pronouns.

According to Kaufman, (12) is a copular clause in which the putative VP predicate is actually a DP. That is, there is no categorical difference between *b<in>ili ko* in (11a) above and (12) (modulo, the appearance of the nominative case marker). If this is correct, then, based on analogy to the examples in (10) and (11), one might expect that the genitive pronoun in (12) should be able to surface as an oblique pronoun to the left of *b<in>ili*. This expectation is evidently born out, as (13) shows (see also, Naylor 1980, and Culwell-Kanarek 2005).

- (13) *Aki-ng binili-∅ ang sapatos.*
 1SG.OBL-LK BEG.buy-PV NOM shoes
 ‘I bought the shoes/The shoes are what I bought’

The pair in (12) and (13) is similar to Kaufman’s pair of sentences (54) and (55), given below as (14) and (15), respectively.

- (14) *Sumayaw diyan ang mga pinsan ko.*
 AV.BEG.dance there NOM PL cousin 1SG.GEN
 ‘My cousin dances there.’
- (15) *Iyong dalawa-ng sumayaw diyan ang mga pinsan ko.*
 that.LK two-LK AV.BEG.dance there NOM PL cousin 1SG.GEN
 ‘Those two who danced over there are my cousins.’

Kaufman points out that (13) and (14) differ interpretively in that (13) has a predicational reading while (14) has a specificational reading.⁴ Likewise, (13) could be characterized as a specification reading which contrasts with the predications reading of (12).

Overall, these distributional facts concerning genitive and oblique pronouns demonstrate that noun possessors and agents of “verbs” pattern alike syntactically. This type of parallelism is consistent with and offers potentially positive support for the claim predicates in Tagalog are nominal and represented syntactically as simple NP’s.

⁴ Here, I take “specificational” to refer to a sentence type with a fixed topic-focus structure (see, e.g., Mikkelsen 2003 and references therein). In (14), the phrase that precedes the nominative represents given or old information (i.e., the topic), while the nominative phrase represents the new information (i.e., the focus). Simple “predicational” sentences like (13), by contrast, have no fixed topic-focus structure.

2.1. *The extraction restriction*

While the facts discussed above may support the claim that all predication involves nominal predication, they also lead to a potential problem for his explanation of the “subject only” restriction on extraction. Kaufman attributes the inability of non-subjects to be extracted in Tagalog to a cross-linguistically common ban on genitive predicates, which – given the pseudo-cleft nature of *wh*-questions in Tagalog – explains why genitives may not be questioned (see Kaufman’s (44)–(49)). The inability of possessors to be extracted via topicalization (see Kaufman’s (62b)) is accounted for in terms of a cross-linguistically common restriction banning possessor extraction (see, e.g., Keenan & Comrie 1977, 1979).

As Kaufman points out, to express a possessor predicate, an oblique case form rather than a genitive form of the possessor must be used. The contrast between the following examples illustrate this point.

- (16) a. *Ni Juan ang koponan.
 GEN Juan NOM team
 (‘The team is Juan’s.’) (= Kaufman’s (38b))
 b. Kay Juan ang koponan.
 OBL Juan NOM team
 (‘The team is Juan’s.’) (= Kaufman’s (39))

As observed above, possessors of simple NP’s and agents in complex NP’s (in particular, those containing a “verb” in the PV-form) may be expressed as an oblique pronominal. Given this, one might expect that the subject only restriction could be evaded if the agent of a PV-form “verb” is realized in this manner, just as the ban on genitive predicates, (16a), can be circumvented by using an oblique form, (16b). This expectation is evidently not born out, as the ungrammaticality of the examples in (17) demonstrate.

- (17) a. *Kanino (ang) binili- \emptyset ang sapatos?
 who.OBL NOM BEG.buy-PV NOM shoes
 (‘Who bought the shoes?’) (Ok as, ‘Who were the shoes bought from?’)
 b. *(Sa) akin ay binili- \emptyset ang sapatos.
 OBL 1SG.OBL AY BEG.buy-PV NOM shoes
 (‘I bought the shoes.’) (Ok as, ‘From me, the shoes were bought.’)

Since these examples do not appear to violate either of the constraints that Kaufman proposes to account for the subject only restriction, it is unclear what their ungrammaticality stems from.⁵ Overall, despite the parallels between possessors of simple NP's and agents of PV-form predicates observed in the previous section, the contrast between (16) on the one hand and (17a, b) on the other suggests a significant difference that may be problematic for a parallel syntactic analysis of the sort that Kaufman proposes.

3. Other copular clauses

While Kaufman notes an interpretive difference between the sentences in (14) and (15) (see above), he claims there is no known evidence for positing a difference in their syntactic structure. There are, however, similar pairs of sentences that do contrast systematically. Consider, for instance, the predicational sentence in (18) and the so-called “focus construction” in (19). Note that in (19) the content of the predicate in the predicational sentence in (18) is expressed in a complex noun phrase.

(18) *Sumayaw diyan si Juan.*
 AV.BEG.dance there NOM Juan
 ‘Juan danced there.’

(19) *Si Juan ang sumayaw diyan.*
 NOM Juan NOM AV.BEG.dance there
 ‘It was Juan who danced there.’

For Kaufman, *s<um>ayaw diyan* would be presumably be a simple NP in both (18) and (19). (18) and (19) also make it seem as if the predicate (*ang*) *s<um>ayaw diyan* can be freely ordered with respect to the notional subject *si Juan*, which might be attributed to Kaufman’s claim that “predication is an inherently reversible relation akin to set intersection”.

However, the word order given in these sentences appears to be fixed. Specifically, unless it has been topicalized, the nominative noun phrase in

⁵ These examples may also raise a problem for the locality accounts of the extraction restriction which Kaufman cites.

(18) must follow the predicate phrase (as shown by (20a)) and in (19) the order of the two nominative marked noun phrases cannot be reversed (as shown by (20b)). (Note: Topicalized phrases always precede sentential negation. Hence, negation is used in these example to preclude the possibility that topicalization has occurred.)

- (20) a. *Hindi si Juan *sumayaw* diyan.
 not NOM Juan AV.BEG.dance there
 ('Juan didn't dance there.')
- b. *Hindi ang *sumayaw* diyan si Juan.
 not NOM AV.BEG.dance there NOM Juan
 ('It was Juan who danced there.')

Another difference between the sentences in (18) and (19) concerns the distribution of a special pronominal form, *siya* (meaning 'the one')⁶. This pro-form may occur within the second nominative noun phrases in sentences like (19), preceding *s<um>ayaw*. By contrast, the ungrammaticality of (21b) shows that this pro-form may not occur within the predicative phrase of simple predicational sentences like (18).

- (21) a. Si Juan ang *siya-ng sumayaw* diyan.
 NOM Juan NOM the.one-LK AV.BEG.dance there
 'It was Juan who danced there.'
- b. (**Siya-ng*) *sumayaw* diyan si Juan.
 the.one-LK AV.BEG.dance there NOM Juan
 'Juan danced there.'

A possible route to explaining the word order restrictions would be to connect them to information (i.e., topic-focus) structure differences between predicational and specificational sentence. If such an approach proves tenable, then the word order restrictions may ultimately be orthogonal to Kaufman's main hypothesis concerning the nominal status of all predication in the language. The syntactic status of the pronoun *siya* in (20a) is unclear at present. At the very least, however, the contrast between (21a) and (21b) suggests some significant difference relating to the

⁶ This pro-form is at least homophonous with the 3rd person singular nominative pronoun meaning 'he/she'.

internal structure of predicate phrases that occur in simple predicational sentence and those which occur as part of complex NP's. Ideally, these differences should be accounted for by a theory such as Kaufman's that analyzes the two as syntactically identical (for Kaufman, analyzed as simple NP's).

4. Concluding remarks

The question of whether or not Tagalog has a verbal category is clearly an important one. Typological research on lexical categories has suggested something of a hierarchy of lexical categories (Schachter and Shopen 2007). Concretely, if a language has only a single lexical category, it will be a noun. If a language makes a two category distinction, it will be between nouns and verbs. Only if it has a three way category distinction will a language distinguish between nouns, verbs, and adjectives. If Kaufman is correct that there is no verbal category in Tagalog, but if – as suggested in section 1 – there must at least be distinction between nouns and adjectives, then Tagalog would seem to be rare from a typological perspective.

Despite many ways in which nominal and putative verbal categories in Tagalog pattern alike, there remain a number of contexts, as noted in Sections 2 and 3 above, where they show divergent properties. Without a detailed analysis of these properties and how they arise, the data discussed here should not, of course, be taken as refutation of Kaufman's main hypothesis regarding the absence of a verbal category in Tagalog. Given the significance of the issues from both a typological as well as Tagalog specific perspective, the asymmetries noted here should receive careful investigation before it can be confidently concluded that Tagalog is distinguished by its complete absence of a verbal category.

Reed College
sabbagh@reed.edu

References

- [1] Abney, Steven. *The Noun Phrase in its Sentential Aspect*. PhD dissertation, MIT, Cambridge, MA.

- [2] Aldridge, Edith. 2004. *Ergativity and Word Order in Austronesian Languages*. PhD dissertation, Cornell, Ithica, NY.
- [3] Culwell-Kanarek, Nathan. 2005. Pre-verbal pronouns in Tagalog syntax. In Heinz and Nthlitheos, eds. *UCLA Working Papers in Linguistics no. 12: Proceedings of AFLA XII*.
- [4] English, Leo James. 1986. *Tagalog-English Dictionary*. Mandaluyong City, Philippines: National Bookstore.
- [5] Himmelmann, Nikolaus. 2008. Lexical categories and voice in Tagalog. In Peter K. Austin and Simon Musgrave, eds. *Voice and Grammatical Functions in Austronesian Languages*. CSLI, Stanford.
- [6] Keenan, Edward and Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8, pp. 63–69.
- [7] Kennan, Edward and Bernard Comrie. 1979. Data on the noun phrase accessibility hierarchy. *Language* 55, pp. 333–351.
- [8] Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In A. Dimitriadis, L. Siegal, et. al., eds. *University of Pennsylvania Working Papers in Linguistics*, Vol. 4.2, Proceedings of the 21st Annual Penn Linguistics Colloquium, pp. 201–225.
- [9] Mikkelsen, Line. 2003. *Copular Clauses: Specification, predication, and equation*. John Benjamins, Amsterdam.
- [10] Naylor, Paz B. 1980. Linking, Relation-Marking, and Tagalog Syntax. In Naylor, Paz B. (ed.), *Austronesian Syntax. Papers from the Second Eastern Conference on Austronesian Languages* [Ann Arbor 1976]. Ann Arbor: Center for South and Southeast Asian Studies (Michigan Papers on South and Southeast Asia No. 15), pp. 33–49.
- [11] Rackowski, Andrea. *The Structure of Tagalog: Specificity, Voice, and the Distribution of Arguments*. PhD dissertation, MIT, Cambridge, MA.
- [12] Schachter, Paul and Timothy Shopen. 2007. Parts of speech systems. In Shopen, ed. *Language Typology and Syntactic Description: Volume 1*. Cambridge: Cambridge University Press, pp. 1–60.
- [13] Travis, Lisa. 2000. The I-syntax s-syntax boundary: Evidence from Austronesian. In Ileana Paul, V. Phillips, and L. Travis, eds. *Formal Issues in Austronesian Linguistics*. Kluwer: Dordrecht.
- [14] Travis, Lisa. 2002. Agents and causes in Malagasy and Tagalog. In Andrea Rackowski and Norvin Richards, eds. *Proceedings of AFLA 8*. Distributed by MIT Working Papers in Linguistics, Cambridge, MA.