Notes on Tagalog nominalism

NIKOLAUS P. HIMMELMANN

Dan Kaufman's brilliant paper articulates the nominalist hypothesis in very clear and up-to-date terms. For the first time, it provides the outlines of a formalized account of this hypothesis which shows the underlying coherence of a range of widely known and much studied morphosyntactic features of Tagalog.

My two comments come from a very sympathetic reader interested in further clarifying and strengthening the analysis rather than debunking it. The first comment will deal with the nature of the syntactic flexibility of content words in Tagalog (relating in particular to Kaufman's section 3.1). The second one pertains to explanations for the ungrammaticality of genitive predicates in Tagalog (relating mostly to Kaufman's sections 3.2 and 4.1).

1. Syntactic flexibility and the nature of lexical categories

In the introduction to his section 3 ("Categoriality on the phrasal and clausal levels"), Kaufman (henceforth K) notes: "The most celebrated symmetricality in Tagalog syntax is on the phrasal and clausal levels . . . , namely, the ability of all word types to appear both in argument position . . . or in the clause initial predicate position." A minor, but not completely irrelevant problem in this wording is the phrase "all word types" which K uses in place of the phrase "all content words" or a similar one found in the previous literature on this topic. The issues at stake here are the proper delimitation of the class of words exhibiting the astonishing syntactic flexibility and their syntactic category. For the following remarks, I take *content words* to be the cover term for all open class items with the exception of personal names.

Everyone, including K I would think, would agree that, of course, not literally *all* word types of Tagalog evince the syntactic flexibility referred to in the quote above. Clearly, various types of closed class function words, including both the phrase-marking proclitics *ang*, *nang* and *sa* as well as the second position clitics do not belong here. Possibly somewhat less clear is the issue with regard to pronouns and deictics because these share some of the flexibility of content words. Still, these items form closed classes with quite a few special properties (both syntactic and semantic) which distinguishes them from the large class of content words. Hence, I fully agree with K's comments regarding the special status of deictics at the end of his section 3.1 (K's examples 31 and 32), but I fail to see the relevance for the overall argument because I would not have counted them among the content words to begin with.

In this regard, it may be useful to note that personal names in Tagalog also quite clearly do not belong to the class of content words. Among other things, personal names are obligatorily marked by a special set of phrase-markers (*si*, *ni*, *kay*), and while the distribution of personal name phrases overlaps in many respects with that of content word phrases, there are also enough differences to require a category of their own. Thus, for example, oblique *kay* can co-occur with oblique *sa* to form phrases indicating a source (e.g. *sa kay Maria* "from Maria"). See also Himmelmann (2005: 146).

But the preceding points are basically irrelevant for the main issue discussed in K's section 3.1, namely the question of the proper syntactic category of Tagalog content words. This question can be tackled from two points of view, language-internally and cross-linguistically. Cross-linguistically, the argument would roughly go like this: The class of Tagalog content words shows many morphosyntactic properties generally associated with nouns. Hence, it makes sense to assign them to this class, especially if one works within the framework of an explicit (formalized) grammatical theory which uses a universalist set of basic syntactic categories. While not made explicit, this line of argument appears to be of major import to K's account.

Language-internal evidence for assigning content words to a particular syntactic category rests essentially on the possibility of further syntactic subcategorization. Otherwise, as convincingly argued in Walter's (1981) classic study on the lack of a noun/verb distinction, it does not make

sense (from a purely language-internal perspective) to classify a large open class of content words as either nouns or verbs. Instead, the words in such a class would have to be assigned to a single macro-category (as in K's option (20b) at the beginning of section 3.1).

In line with this consideration, K uses the occurrence of a class of words with the suffix −*i* in southern Tagalog dialects as an argument to assign the category *noun* to the large majority of content words in Tagalog. The *i*-suffixed words in these dialects are restricted to predicative (and more specifically, imperative) function and hence prima facie qualify as verbs. It is not quite clear to me, whether and to what extent it is really legitimate and useful to introduce dialectal evidence here, as one would still want to see an analysis which fully accounts for those varieties of Tagalog – including the standard(ized) variety – where no *i*-suffixed forms occur and where all content words (properly delimited along the lines indicated above) appear to be able to occur in all argument and predicate functions and positions (but see below for an important qualification).

Otherwise, I fully agree with K's argument. In fact, the argument is much stronger for many other Philippine-type languages where word forms restricted to predicative (verbal) positions and functions are more varied and more robustly attested. In Bisayan languages, for example, there are generally two morphological classes of content words, ones suffixed with –*i* and ones suffixed with –*a*, which are restricted in this way. In these languages, the verbal forms do not only occur in imperatives but also as dependent or controlled elements in various kinds of subordinating constructions (hence the forms are often labelled *subjunctive*). The following example is from Cebuano where predicates negated with *waláʔ* have to occur in subjunctive form:

(1) *waláʔ*   *niya*   *lutu-**a***
     NEG.PST   3.SG.GEN   cook-PV.SUBJ
     He did not cook it. (Wolff 1972: 1121)

In fact, and this is what I consider to be one of the two most difficult empirical challenges to the nominalist hypothesis (cp. Himmelmann 2005: 141),[1] it may be the case that even in standard Tagalog the position

---

[1]   The other one relates to the fact that undergoer voice subjunctives also quite generally occur with genitive marked actors, cp. Kaufman 2009, section 5.2.

of the controlled predicate in at least some types of control constructions is restricted to voice-affixed forms. That is, it is not clear to me whether in the constructions illustrated by the following two examples,[2] it is at all possible to use unaffixed lexical bases in the controlled slots (i.e. the slots occupied by *humukay* and *gamitin*, respectively):

(2) **Iniyutus** nya**ng** **humukay** nang apat na
   i -in -utos niyá -ng um-hukay nang apat na
   CV-RLS(UG)-order 3.SG.GEN-LK AV-excavation GEN four LK
   malalim na balón
   ma-lalim na balón
   ST-depth LK well
   He ordered four deep wells dug (he ordered someone to dig four wells) (Bloomfield 1917: 24)

(3) ang tuláy ay **ipinagbawal** **gamitin**
   ang tuláy ay i -in -pag-bawal gamit-in
   SPEC bridge PM CV-RLS(UG)-GER-forbidden use -PV
   The bridge was condemned (forbidden to be used) (English 1986)

If controlled predicate positions can indeed only be filled by voice-marked forms, the ability to occur in these positions could be used to divide the large class of content words into more noun-like and more verb-like (= voice-affixed) members (in the same way as K uses the *i*-suffixed forms in southern Tagalog dialects).

   Note, however, that there are two problems with this line of argument. First, and to my mind somewhat surprisingly given the non-negligible amount of formal literature on Tagalog, it is still the case that not much is really known about control and complex predicate constructions in Tagalog apart from the very perceptive, but far from comprehensive discussion in Kroeger (1993: 71–107). From the little work I myself have tried to do on this topic it has become clear to me that this is one of those topics where extensive corpus research has to be combined with careful syntactic experimentation[3] in order to achieve empirically viable results

---

2   Example (2) illustrates a 'normal' Tagalog control construction, example (3) something which could be analyzed as a complex predicate construction (note the missing ligature/complementizer!).

3   Of the type advocated by Schütze (1996) and illustrated by some of the contributions in Kepser & Reis (2005).

(i.e. there is hardly any use in doing quick checks of a few invented examples with one or two native speakers!).

Second, even if the careful empirical research just mentioned were to show that there is indeed a requirement to use voice-affixed forms in this position, it may still be the case that this does not reflect a syntactic constraint but rather can be accounted for in purely semantic terms. Given the recent attempts to account for control phenomena mostly in semantic terms (e.g. Jackendoff & Culicover 2003), this does not seem to be a completely unreasonable suggestion.

2. Constraints on genitive predicates

As convincingly argued by K in his section 3.2, the famous extraction asymmetries of Tagalog are related to the fact that Tagalog generally disallows genitive-marked phrases to function as predicates.[4] That is, the widely used examples of the type

(4) ***Si**no ang binili-∅ ang libro?
    who SPEC RLS(UG):buy-PV SPEC book
    (For, 'Who bought the book?')

are ungrammatical for reasons completely orthogonal to the issue of extraction. The examples which would actually test for extraction, i.e. (K's 48)

(5) ***Ni**no ang binili-∅?
    GEN:who SPEC RLS(UG):buy-PV
    (For, 'Who bought (it)?' / 'Whose was the bought thing')

---

[4] In fact, as discussed further below, all constructions where a genitive-marked phrase would precede its head are disallowed. An exception are certain types of (mostly temporal) adjuncts which appear to be introduced by the genitive proclitic *nang* (cp. Kaufman 2009, section 3.4):

(i) Nang ala-una ay umalis sila
    ?? o'clock-one PM AV:leave 3.PL
    'At one o'clock, they left'

The analysis of this construction is still a matter of debate, including the question of whether the initial *nang* is indeed an instance of the genitive proclitic. For the remainder of this section, this construction is ignored.

are ungrammatical because genitive predicates are generally ungrammatical in Tagalog as shown by examples such as (K's 38b and 49):

(6)   ***Ni**   Juan   ang   koponan
          GEN   Juan   SPEC   team
      (For, 'The team is Juan's')

However, as noted by K, in the latter case the genitive proclitic can be replaced by a dative proclitic (K uses *oblique* instead of *dative/locative*) and then the sentence becomes grammatical (K's 39):

(7)   Kay=Juan   ang   koponan
      DAT=Juan   SPEC   team
      'The team is Juan's'

This is just an instance of a very general pattern according to which genitive marking is replaced by dative or locative marking when modifier precedes modified. Thus, instead of *ang libro niya* for 'her book' one can also say *ang kaniya-ng libro* (SPEC 3.SG.DAT-LK book). Consequently, the question arises whether the real examples to be tested for delimiting extraction constraints in Tagalog should not rather involve dative/locative phrases, i.e.

(8)   ?*(sa)   kanino   ang   binili-∅?
         LOC   DAT:who   SPEC   RLS(UG):buy-PV
      (For, 'Who bought (it)?')

Example (8) is actually grammatical as a possible expression for 'who has the bought thing/the bought thing belongs to whom'. But I have no natural attestations of examples of this type where the extracted phrase refers to the actor. And I have also not been able to test it properly with native speakers for this reading. Hence, I have to leave this question undecided.

   Regardless of how this empirical issue is resolved, the question remains of how to explain the robust constraints on the extractability of genitive phrases in Tagalog. K suggests that an explanation may be found in relation to the restrictions on possessor extraction widely attested in the languages of the world (his section 4.1), i.e. that Tagalog is just particularly rigorous with regard to a constraint that is found generally in the languages of the world.

While I agree that this may be part of the story, it seems to me that the restrictions on extracting genitives is part of a more general pattern in Tagalog which also includes the linking particle *na/-ng* and the complementizer *na*. The linker, by its very nature, has to come *in between* the two constituents it connects. Hence, the inversion of *balón na malalim* 'deep well' is not **na malalim balón*, but *malalim na balón*. As for complement clauses, it is generally impossible in Tagalog to have them precede the matrix predicate, as illustrated here with K's (25):

(9) a. Sinabi-∅   nila   na   hindi   sila
         RLS(UG):say-PV   3.PL.GEN   LK   NEG   3.PL
         mag-bá-basketbol
         AV-RDP-basketball
         'They said that they wouldn't play basketball.'
     b. *na hindi sila magbábasketbol sinabi nila

It seems obvious to me that complementizer *na*, linker *na/-ng* and genitive *nang* form a family of semantically and structurally related elements and constructions where these elements are positionally restricted to occur *in between* the two constituents linked by them.[5] Hence, I would suggest that a comprehensive explanation of the constraints on the extraction of genitive-marked phrases in Tagalog should also be able to account for the fact that the same restriction also pertains to the complementizer *na* and the linker *na/-ng*.

3.   Conclusion

In concluding these brief remarks, I would like to stress a point implicit in the title of this squib. I believe that Tagalog represents the most radical manifestation of nominalist syntax in Austronesian languages. While nominalist traits can be found in many Austronesian languages, it appears to be the case that only in (standard) Tagalog almost all differences

---

5   It is most likely that the complementizer *na* and the linker *na/-ng* both have the same historical source, but they differ in their variants, as briefly discussed by K in his section 3.1. It has also been suggested that genitive *nang* actually includes the linker (i.e. consists of *na + ang*), but this is highly speculative.

between nominal and verbal predicates have disappeared and given rise to the exclusive use of the equational clause type. That is, for (standard) Tagalog it may be possible to argue that all predication involves nominal predicates and hence nominalist syntax. For other languages, including possibly southern Tagalog dialects, the analysis for nominal predicates in equational structures needs to be complemented with an analysis of a more verbal predicate type (as it stands, it is unclear how K would analyse *i*-suffixed predicates in southern Tagalog and the more clearly verbal dependent forms attested in most other Philippine-type languages illustrated by the Cebuano example in (1) above). Such an analysis would have to be able to account for the fact that, at least in some languages and constructions, the marking of the arguments of the more verb-like predicates is identical with the marking of nominal predicates (i.e. in particular the genitive-marking of actors of undergoer voice verbs). But, at the same time, it has to allow verbal heads also to occur with crosslinguistically more canonically marked arguments (e.g. as direct ('accusative') complements or, in undergoer voice constructions, as oblique agents), as such marking is widespread in the symmetrical voice languages of Indonesia.

Abbreviations

AV	ACTOR VOICE	PN	PERSONAL NAME
CV	CONVEYANCE VOICE	PST	PAST
DAT	DATIVE	PV	PATIENT VOICE
GEN	GENITIVE	RLS	REALIS
GER	GERUND	RDP	REDUPLICATION
IN	INCLUSIVE	SG	SINGULAR
LK	LINKER	SPEC	SPECIFIC ARTICLE
LOC	LOCATIVE	ST	STATIVE
NEG	NEGATION	SUBJ	SUBJUNCTIVE
PL	PLURAL	UG	UNDERGOER
PM	PREDICATE MARKER		

Westfälische Wilhelms-Universität Münster
nikolaus.himmelmann@uni-muenster.de

References

Bloomfield, Leonard. (1917) *Tagalog texts with grammatical analysis*, Urbana: The University of Illinois.

English, Leo J. (1986) *Tagalog-English dictionary*, Manila: National Book Store.

Himmelmann, Nikolaus P. (2005) 'The Austronesian languages of Asia and Madagascar: Typological overview', in Alexander Adelaar & Nikolaus P. Himmelmann (eds), *The Austronesian languages of Asia and Madagascar*, 110–181, London: Routledge.

Jackendoff, Ray, and Peter W Culicover. (2003) 'The semantic basis of control', *Language* **79**.3.517–556.

Kaufman, Daniel. (2009) 'Austronesian typology and the nominalist hypothesis', In Adelaar, A. and Pawley, A. (eds.), *Festschrift for Robert Blust*, Canberra: Pacific Linguistics.

Kepser, Stephan & Marga Reis (eds). (2005) *Evidence in Linguistics*, Berlin: Mouton de Gruyter.

Kroeger, Paul R. (1993) *Phrase Structure and Grammatical Relations in Tagalog*, Stanford: Stanford University Press.

Schachter, Paul and Otanes, Fay T. (1972) *Tagalog reference grammar*, Berkeley: University of California Press.

Schütze, Carson T. (1996) *The Empirical Base of Linguistics*, Chicago: The University of Chicago Press.

Walter, Heribert. (1981) *Studien zur Nomen-Verb-Distinktion aus typologischer Sicht*, München: Fink.

Wolff, John U. (1972) *A Dictionary of Cebuano Visayan*, 2 vols., Ithaca: Cornell University Southeast Asia Program.