Featured Article

Language & Linguistics 語言暨語言學

Language and Linguistics
15(6) 761–774

DOI: 10.1177/1606822X14544619
lin.sagepub.com

SAGE

Constituent Structure in a Tagalog Text*

Randy J. LaPolla

Nanyang Technological University

If there are no classes of noun or verb in Tagalog, how can there be noun phrases and verb phrases? This paper contributes to the discussion on form classes in Tagalog by taking a detailed, inductive, line-by-line look at the structures and constituents found in a randomly selected Tagalog text, to create a typology of the structures found therein. It is shown that, while there are very obvious constructions with generally clearly differentiated functions, they do not correspond with noun phrases and verb phrases in Indo-European languages, as it cannot be said that one form is used for predication and another for reference.

Key words: Austronesian languages, form classes, linguistic typology, phrase structure, Tagalog

1. Introduction

This paper presents an inductive examination of the constituents found in a randomly selected Tagalog text, Bob Ong's *Alamat ng Gubat*.[1] The analysis is based on the full text, but we are able only to go through the first few lines of it here, which we do individually, sequentially, and discussing the structures found in each line. At the end of the paper, we pose some important questions about the structures found in Tagalog based on this text.

2. The text

(1) *Alamát ng Gúbat*
[alamát ng gúbat]$_{\text{Y/TITLE}}$
legend REL[2] jungle
(The) Legend of the Jungle

---

* This paper was presented as a keynote address to the 10th Philippine Linguistics Congress, University of the Philippines, Diliman, Quezon City, December 10–12, 2008. I would like to thank all those who commented on the paper at that time, as well as Carl Rubino and Ricardo Nolasco for their very helpful feedback on drafts of this paper following its presentation at the Congress.

1 Ong, Bob. 2004. *Alamat ng Gubat*. Makati City, Philippines: Visual Print Enterprises. 'Bob Ong' is the pen name (real name unknown) of an author known for using vernacular language to poke fun at life in the Philippines. His works are considered classics of Pinoy culture (see http://en.wikipedia.org/wiki/Bob_Ong).

2 When written alone, *ng* is pronounced [nɑŋ]. Please refer to the table at the end of this paper for a guide to glossing abbreviations used herein. An acute accent marks a stressed syllable, and a macron means the pitch stays high for that syllable (the stress marks and the glottal stops do not appear in the Tagalog orthography). Spanish and English loan words are in italics in the morpheme analysis line. Phrases are marked for their syntactic type and their functions in the text; for example, 'Y/TITLE' marks the phrase as a Y phrase that is acting as the title of the story. I have not been able to maintain the paragraph breaks that appear in the published version of this text, which should be consulted directly for a review of that level of structure.

The first type of phrase we find here in line 1 (the title of the book) is what I will be calling a 'Y' PHRASE (or relator phrase): one formed by *ng* ([nɑŋ]), if the word following the marker is not a common name, or *ni*, if the word following the marker is a human, proper name (*nina* when more than one proper name follows). This sort of structure manifests the following set of modifying functions: part–whole (inalienable possession), possession (alienable), 'possession' with locative and abstract concepts (such as in line 1), 'adverbial'[3] modification, and predicate-argument structure when the argument is not the topic of the clause or a locative argument marked by *sa*. In this structure, the head of the phrase is the first constituent, and the modifying element follows *ng* or *ni*/*nina*.

(2) Noóng únāng panahón,
[noón=ng *úna*=ng panahón]$_{X/TOP}$
that.time=LNK first=LNK time/epoch/weather
long ago (at the time of the first epoch),

The second type of phrase found is what I will call the 'X' PHRASE (or linker phrase): one formed by *-ng* (following an open syllable) or *na* (following a closed syllable). This structure manifests a much more varied set of modifying functions than does the Y phrase: 'adjectival' modification, numeral and measure modification, relative clause modification (restrictive and non-restrictive), demonstrative modification, (intensifier) 'adverbial' modification, intra-predicate structure (e.g. the relationship between a positive or negative existential and an existent—the thing that exists—in an existential predicate), indirect quotes, certain types of possessive modification, non-possessive modification, and the relationship between a predicate and its arguments in certain types of referential use (i.e. when they together form part of a (higher) clause argument). Unlike the Y phrase, this sort of structure does not link predicates and arguments in clauses acting as main clauses, but only marks relationships within clausal constituents.

The semantics of this sort of phrase are often difficult to determine: in many cases, it seems to be simply marking the fact that the elements form a phrase. The grammatical head of the phrase cannot consistently be identified by position, as in many cases the two (or more) elements can be reversed (*magangdang babae* / *babaeng maganda*; 'beautiful woman').

(3) sa isáng liblíb na kaharián
[sa [isá=ng liblíb na kaharián[4]]$_{X/LC}$]$_{LP}$
LOC one=LNK remote LNK kingdom
in a remote kingdom

[3] I use quotation marks around 'adverbial' here, and around 'adjectival' later, as there is much controversy about the definition of form classes in Tagalog. See for example Himmelmann (2008) and LaPolla (2010) for further discussion of this. Here, I am attempting to contribute to this discussion by looking at phrases rather than words, as, if there are no form classes of noun and verb, how can there be noun phrases and verb phrases?

[4] The word *kaharián* is formed from the root *hári* ('king'), plus the two affixes *ka-* (ASSOC) and *-an* (LFS). The two are independent affixes, but are commonly used together to express abstract concepts; for example, *kaálaman* 'wisdom' (< *alám*; 'know'), *kagandáhan* 'beauty' (< *gandá*; 'beauty'), and *kabuháyan* 'livelihood' (< *búhay*; 'life'). The two affixes are used together for this sense (there is no **kaalam* or **alaman*), though it seems there would have been an order of affixing; for example, with *kahárian*, the sense is 'a place where (people) have the same king', so it seems *ka-* would be affixed first, then *-an*.

(4) sa ilálim ng dágat,
[sa [ilálim ng dágat]$_{Y/LC}$]$_{LP}$
LOC bottom REL ocean
at the bottom of the sea,

The third type of phrase, found here in lines 3 and 4, is the locative phrase (LOC), which is marked by *sa*. The LOC phrase can be used for many sorts of locational and directional senses (e.g. allative, ablative) and for most other arguments not appearing in X or Y phrases.[5]

The 'preposition' *sa* can take a single word, an X phrase, or a Y phrase as its complement. In line 3 it takes an X phrase as its complement, and in line 4 it takes a Y phrase as its complement. It is somewhat problematic to call *sa* a preposition (as e.g. Himmelmann (in press) appears to do), because normally, a preposition is preposed to a noun phrase. This is not clearly the case here, as neither the X nor Y phrase is obviously nominal in a grammatical sense, but, as the form is preposed to something, I will continue to use the term 'preposition'.

There are two LOC phrases in lines 3 and 4, and they are both functioning at the clause level. That is, the second one is not embedded as a modifier within the first one, and could appear after the predicate rather than before it, as it is here.

(5) ay may nakatiráng magandá at mabaít na siréna.
ay [may [naka-tirá=ng
FT EXIST OS-live=LNK

[[[ma-gandá at ma-baít]$_{CONJP}$ na]$_{MOD}$ *siréna*]$_{X}$]$_{X/EX}$]$_{PRED}$
STAT-beauty CONJ STAT-goodness LNK siren(mermaid)
there was a beautiful and nice mermaid who lived there.

A fourth type of construction is formed by the linker *ay*, seen here at the beginning of line 5. It marks the fact that the element before it is part of the same construction (the clause) as the element following the marker, which is always the predicate of the clause. Lines 2–4 all relate to the predicate in line 5.[6] This construction contrasts with clauses in which all elements follow the predicate. The item fronted can be a locative/temporal expression (often a scene-setting element) or the topic of the clause (often contrastive). 'Topic' here refers to the grammatical pivot (grammatically privileged argument) of the construction, the argument singled out for special morphosyntactic treatment (when realis, irrealis, conveyance, or locational affixes are used on the predicate, they co-reference the semantic role of the topic—this is the so-called 'focus system' of Philippine languages, and the topic itself, if a pronoun, takes a special form, or, if a lexical form, takes a marker of specificity). In Tagalog, this argument is the topic in the pragmatic sense of being what the clause is about (see Lambrecht 1994 on the definition of 'topic'), and so it is also appropriate to call it the topic.

---

[5] For locational predications, *nasa*, rather than *sa*, is generally used; for example, *Nasa gubat siya* ('He is in the forest').

[6] I did not use brackets to mark off this phrase because it is so long, and the initial bracket might have confused the reader when discussing the first line.

Much is made of the fact that a clausal argument preceded by *ng* cannot appear in the pre-predicate position of an *ay* phrase (see e.g. Kaufman 2009), but there is confusion concerning the difference between a clausal argument preceded by *ng*, on the one hand, and arguments preceded by *sa* or *ang*, which can appear in the pre-predicate position of an *ay* phrase, on the other hand, owing to the assumption that *ang*, *sa*, and *ng* are all the same type of marker, often called case markers. If, instead, we see *ng* as a linker (which links two elements in a Y phrase and requires two elements to be used), rather than as a case marker, much like the *=ng*/*na* linker and *ay*, then there is no confusion about why we do not find arguments standing alone with *ng* in any position, just as we don't wonder why elements preceded by *=ng*/*na* don't stand alone, as *ng* creates a Y phrase in which the two elements are the head and the modifier—in this case, the predicate and the relevant argument. That is, I am arguing that, just as, for example, we would not expect to achieve the particular modificational relationship between *dágat* and *ilálim* in *ilálim ng dágat* ('bottom of the ocean') without them being in the order they are in and linked by *ng*, we shouldn't expect to achieve the particular relationship between *kumúha* and *ságing* (i.e. predicate and argument) in *kumúha ng ságing* ('get a/the banana') without them being in the order they are in and linked by *ng*. This is why there can be no 'extraction' of this sort of phrase.

The predicate in line 5 (everything in this line after *ay*) takes the form of an existential phrase (EX), based on the existential *may*; this word can take a single word, an X phrase, or a Y phrase as its complement. In this case, it takes an X phrase as its complement. If the remote demonstrative *doon*/*roon* compounds with *may* in the predicate (e.g. in line 27, later: *mayroon s'yang karamdáman* ('he has an illness')), then the combined form *mayroon* is linked to the existent in an X phrase. This structure is also used for possession, with the possessor as topic, as in line 27.

A sixth type of structure, seen in the latter part of line 5, is the conjunction phrase (CONJP), marked by *at*; it conjoins two elements of any level. In this example, the conjunction phrase is embedded within an X phrase, and links two 'adjectival' modifiers.

(6) Péro wala siyang kinaláman
péro [walaʔ <siya>$_{TOP}$=ng [kinaláman[7]]$_{EX}$]$_{X/PRED}$
but N.EXIST 3sgT=LNK involvement
But she doesn't have anything to do

(7) sa kwéntong ito.
[sa [*kwénto*=ng ito]$_{X/LC}$]$_{LP}$
LOC story=LNK PROXT
with this story.

In line 6, we find a negative existential phrase, where the predicate is based on the negative existential *wala*. As with the positive existential phrase, in this type of phrase the existent can appear as part of the predicate. This structure is also used for (negative) possession, as in this

[7] The word *kinaláman* ('involvement') is formed from the root *alám* ('know'), plus *ka-* (ASSOC) and *-an* (LFS), together forming *kaálaman* ('knowledge, wisdom'), and then the infix *-in-* (RPUT) appears within the *ka-* prefix. The resulting form **kinaalaman* reduces to *kinaláman* accordingly.

example, with the possessor as topic. See line 12, later, for a further example. It is also possible for the existent to appear as an *ang*-marked topic (see below on *ang*), as in *Walá na ang sakít* [N.EXIST CSM SPEC pain] 'The pain is gone'.

Because the personal pronouns are second-position clitics when they are functioning as arguments, and not functioning as predicates, they often occur linearly (but not grammatically) within the predicate, even if they represent the topic, as in this case, where the predicate is *walang kinaláman*, but the pronoun representing the topic appears after the first element of the predicate. This shows that the elements of the X phrase making up the predicate do not need to be contiguous, and that the linker *ng* ~ *na* is not necessarily marking a relationship between the element carrying the linker and the element following it.

In line 7, we have another LOC phrase with an X phrase complement.

(8) Káya ang pagtutuúnan na lang natin ng pansín

káya	[ang	[{pag-tu-tuón-an[8]	na	lang	natin}$_{Y}$	
therefore	SPEC	GER-REDUP-emphasis-LFS	CSM	only/just	1pinclNT	
ng	pansín]$_{Y}$]$_{TOPz}$					
REL	attention					

Therefore the one we will be focusing our attention on [lit: our focusing of attention]

In line 8, we find an *ang* phrase: the particle *ang* (or *si* for personal names / *sina* for more than one personal name) marks the constituent that is the topic of the clause as specific. It is historically a demonstrative followed by the linker (Reid 2000, 2002), and, in conversation, is often replaced by a form of the remote demonstrative *iyon* plus the linker: *'yung* (*iyung*). The particle *ang*/*'yung* can be followed by a single word, an X phrase, or a Y phrase. In line 8, two Y phrases (*pagtutuúnan ng pansín* and *pagtutuúnan natin*) overlap, and are both broken up by two second-position clitics, *na* and *lang*. The form of the pronoun, *natin*, shows it participates in the Y phrase, and so does not require the use of *ng* to mark the relationship with *pagtutuúnan*.[9]

(9) ay si Tong,

ay	[si	[Tong]
FT	SPEC[10]	PN

is Tong,

[8] See *tutuúnan* ('concentration'), but it seems the order of affixing is, first, to add *pag-*, forming *pagtutuon*, and then *-an* is added to that.

[9] Although personal pronouns (e.g. *natin* here) and demonstrative pronouns, such as *noon* in line 2, form possessive phrases when they follow reference to a referent, similar to expressions preceded by *ng* in Y phrases, as in *pagtutuúnan natin* or *aso natin* ('our dog'), they are morphosyntactically more free than expressions preceded by *ng* in Y phrases, appearing often as second-position clitics and phrase-initially and taking the *-ng*/*na* linker when linked to preceding phrases (e.g. *Nagulat akong noong nakita kita*; 'I was surprised when I saw you'), and they can be used without a possessive sense as simply a non-topical argument.

[10] A reviewer questioned why I use SPEC for the gloss of this form and *ang*, suggesting I might use 'Topic' instead. I use SPEC following Himmelmann (2008), but also because it does not always mark a topic; it simply marks the referent as identifiable, as in line 9.

Here, again, *ay* marks the constituent before it as the topic, and the constituent following it as the predicate. The clause formed by lines 8 and 9 is similar to a cleft construction, where what would otherwise be the topic is the focus (in this case, the predicate), and what would otherwise be the predicate acts as topic. Although the proper name *Tong* does not function as a topic here, it takes the particle *si*, which generally marks personal names that function as topics. This is because proper names (in the singular) must be marked by *si* or *ni* or *kay*, except when used as vocatives, but the use of *ni* (which is functionally equivalent to *ng*, but is used before personal names) would imply its participation in a Y phrase, which is not the case here, and *kay* (roughly, the equivalent of *sa* for proper names) would imply it is a locative argument, which it isn't, and so *si* is used here.

(10) ang pinakabátang anák ni Háring Talangkáʔ
[ang [[pinaká-báta=ng anák]$_{X}$ ni {{Hári=ng Talangkáʔ}$_{X/TOPi}$]$_{Y}$
SPEC SUPER-young=LNK child REL king=LNK crab
the youngest child of King Crab,

(11) na túlad ng maráming hári
[na [[túlad ng [ma-dámi=ng hári]$_{X}$]$_{Y}$]
LNK just.like REL STAT-many=LNK king
who, like many kings,

(12) ay waláng ibáng papél sa kwénto
ay {<<waláʔ=ng [ibá=ng *papél*]$_{X/EX}$>$_{X/PRED}$ [sa [*kwénto*]$_{LC}$]$_{LP}$
FT N.EXIST=LNK other=LNK role LOC story
has no role in (the) story

(13) kundíʔ ang magkasakít.
kung [[hindíʔ]$_{PRED}$ [ang magka-sakít]$_{TOP}$]$_{CL}$>$_{CC}$}$_{PREDi}$]$_{AM}$]$_{PREDz}$
COMP NEG SPEC POSPREF-illness
other than to be the one who becomes sick.

The passage from line 7 to line 13 involves several intertwined phrases: *si Tong*, *ang pinakabátang anák ni Háring Talangkáʔ* is the predicate for the fronted topic *ang pagtutuúnan na lang natin ng pansín*, with *ang pinakabátang anák ni Háring Talangkáʔ* modifying *si Tong* as an appositional modifier; *pinakabátang anák ni Háring Talangkáʔ* is a Y phrase, which includes the X phrases *pinakabátang anák* and *Háring Talangkáʔ*.

Háring Talangkáʔ also forms an X phrase with (is modified by) the non-restrictive modifier *na túlad ng maráming hári* and is also the topic of the predicate *waláng ibáng papél sa kwento kundíʔ ang magkasakit*, the latter being itself a clause complex with two clauses, the first of which has *Háring Talangkáʔ* as the (sub)topic and the second of which (a subordinate clause marked by *kung*) has *ang magkasakít* as the topic.

(14) At íto na ngá ang dahilán
at [[{íto na nga}$_{PREDj}$ {ang [dáhil-an
CONJ PROXT CSM EMPH SPEC because-LFS

(15) kung bákit isáng áraw
[kung [bákit [isá=ng áraw]$_{X/TOPz}$
COMP why one=LNK day
and it is the reason why one day

(16) ay biglá na lang ipinatáwag
ay [{biglaʔ na lang i-in+pa-táwag}$_{PREDi}$
FT sudden CSM only CON-RPUT-CAUS-call

(17) si Tong ng kanyáng ináng réyna.
{si Tong}$_{TOPi}$ [ng [kanyá=ng iná=ng *réyna*]$_{X}$]$_{A}$]$_{Y/CL/PREDz}$]$_{CL}$]$_{CP}$]$_{AC}$}$_{TOPj}$
SPEC PN REL 3sgPOSS=LNK mother=LNK queen
Tong's mother, the Queen, suddenly summoned him (had him called).

Line 14 begins with the conjunction *at*, which, here, links the following clauses (lines 14–17) with the preceding ones (lines 7–13).

In lines 14–17, again, we have a very complex structure, where the (fronted) topic of the highest level structure is a complex structure: *itó na ngá ang dahilán kung bákit isáng áraw ay biglá na lang ipinatáwag si Tong ng kanyáng inang réyna* has *itó na ngá* as predicate and *ang dahilán kung bákit isáng áraw ay biglá na lang ipinatáwag si Tong ng kanyáng inang réyna* as topic, with this topic including the subordinated modifier *kung bákit isáng áraw ay biglá na lang ipinatáwag si Tong ng kanyáng inang réyna*.

In line 16, we have the fronted topic marker, followed by the predicate of the embedded clause, which is itself a full clause, with a predicate and topic.

In the predicate of the embedded clause, *biglá* and *ipinatáwag* seem to form a phrase (even though they are separated by the second-position clitics), but there is no morphological marking of their relationship.

The representation of the actor of the embedded clause is an X phrase, but the whole of it forms a Y phrase with the predicate, and the Y phrase is interrupted by the topic.

(18) 'Tong, anák, ang iyóng amá
[[Tong]$_{VOC}$ [anák]$_{VOC}$ [ang [iyó=ng amá]$_{X}$]$_{TOPi}$
PN child SPEC 2sgPOSS=LNK father
'Tong, Child, your father

(19) ay may karamdáman', wíka ng réyna.
ay [may [ka-ramdám-an]$_{EX}$]$_{PREDi}$]$_{PREDz/QUOTE}$ [wíkaʔ ng *réyna*]$_{Y/TOPICz}$
FT EXIST ASSOC-feeling-LFS language REL queen
has an illness', said the Queen.

Lines 18–19 form another clause type, an equative clause (with no copula). In this instance, the predicate is an embedded quote, and the topic of that predicate is the quoting phrase. The quote starts with a vocative, and, within the quote, there is a fronted topic that takes the form of an X

phrase marked by *ang*, and the predicate takes the form of an existential phrase. This sort of structure can be used for interrogative word questions, such as [*Ano*] [*ang gusto mo*]? [what SPEC want 2sgNT] 'What do you want?'; for attributive predication, such as [*Titser*] [*ako*] 'I am a teacher'; and for identificational predication, such as [*Ito*] [*ang gusto ko*] [PROXT SPEC want 1sgNT] 'What I want is this'.

The topic of the higher structure (the clause that has the quote embedded as the predicate) is a Y phrase, *wíka ng réyna*—literally, '(the) language of the queen'. This phrase is not marked as a topic, but, in line 27 we have the same structure, and it is marked overtly as the topic. (See LaPolla & Poa 2005, on speech act constructions in Tagalog.)

(20) 'Hindí na s'ya nakákalangóy.
[hindíʔ na <siya>$_{TOP}$ nakáka-langóy]$_{PRED}$
NEG CSM 3sgT INHERENT.ABLE-swim
'(He) is not able to swim anymore.

In line 20, again, two elements (*hindíʔ* and *nakákalangóy*) seem to form a phrase (separated by the second-position aspect clitic and the topic), but there is no morphological marking of the relationship between the two elements.

(21) Kailángan mong umáhon ngayón din
[kailángan <mo>=ng [um-áhon ngayón din[11]]]$_{X/PRED}$
need 2sgNT=LNK RPAT-get.up now also
You need to get up right now (and)

(22) papúnta sa lúpa
[pa-púnta]$_{PRED}$ [sa lúpa]$_{LP}$
PROSP-go LOC land
go to the land

(23) úpang kumúha ng púso ng ságing—
úpang [um+kúha ng [[púsoʔ ng ságing]$_{Y}$
in.order.to RPAT+get REL heart REL banana
to get the heart of the banana—

(24) ang tánging prutas na makakapágpagaling sa kanya.'
[ang [tangi=ng *prutas* na [makáka-pag-pa-galing [sa kanya]$_{LP}$]$_{MOD}$]$_{X}$]$_{AM}$]]$_{Y}$
SPEC only=LNK fruit LNK SIT.ABLE-GER-CAUS-recover LOC 3sgPOSS
the only fruit that can cure him.'

In line 21, *mo* forms a Y phrase with *kailangan*, but this phrase is intertwined with the X phrase formed by *kailangan* and *=ng umáhon ngayón din papúnta*, due to the nature of *mo* as a

[11] The combination of *ngayon* ('now') and *din* ('also') means 'right now'.

second-position clitic. This X phrase functions as the predicate of the clause. The second-person pronoun takes the form *mo* because it is a non-locative argument, but not the topic, of *kailangan*. The second person does seem to be the topic of the predicate *umáhon*, though, as it is marked for 'actor topic' and it is assumed that it is the addressee that will get up.

In line 22, we have another clause, which might be seen as part of a serial construction with the predication in line 21.

In line 23, we have a subordinate clause giving the reason why Tong has to get up and go to the land. Again, there is no overt topic, though, as the predicate is marked as an actor topic, we assume the actor (Tong) is the topic intended.

Line 24 is an appositional modifier, modifying *ságing* ('banana'). It takes the form of a complex X phrase marked by *ang*. One element in the X phrase is a predicate plus a *sa* phrase, *makakapágpagaling sa kanya* ('able to cure him'), which, in the context, is seen as acting as a modifier of *prutas* ('fruit'). That is, it functions like a relative clause, but the structure it forms with *prutas* is just like any other X phrase. Again, there is nothing in the structure that identifies *prutas* as a grammatical head, so we identify *prutas* as the element being modified simply by inference; that is, it makes more sense in the context than the other way around.

(25) Sumagót si Tong, 'Ngúnit ináng réyna,
[um+sagót]$_{PREDi}$ [si Tong]$_{TOPi}$ [[ngúnit [iná=ng *réyna*]$_{X/VOCATIVE}$
RPAT+answer SPEC PN however mother=LNK queen
Tong answered, 'But, Mother Queen,

(26) hindí ba't talagá namang hindí nakákalangóy[12]
[hindíʔ ba at talagá namán=ng hindíʔ nakáka-langóy]$_{X/PREDz}$
NEG Q CONJ really also=LNK NEG INHERENT.ABLE-swim

ang amáng hári?'
[ang [amá=ng hári]$_{X}$]$_{TOPz}$]$_{QUOTE}$
SPEC father=LNK king
isn't it the case that my father, the king, actually can't swim (anyway)?'

(27) 'Dáhil nga mayroon s'yang karamdáman!', ang sagót ng rénya.
[[dáhil nga [may-doon <siya>$_{TOPi}$=ng [ka-ramdám-an]$_{EX}$]$_{PREDi}$]$_{PREDz/QUOTE}$
because EMPH EXIST-DISLOC 3sgT=LNK ASSOC-feeling-LFS

[ang sagót ng *réyna*]$_{TOPz}$
SPEC answer REL queen
The reply (answer) of the queen was, 'Because he has an illness!'

[12] Notice that *hindí nakákalangóy* 'never could swim' differs from *hindí na nakákalangóy* 'can't swim anymore' in line 20 only in the use of the change of state marker *na* in the latter.

In lines 25–27, we have two different speech act constructions. In the clause that makes up lines 25–26, the predicate is the quoting expression, here inflected for the actor topic, and the quote is an unmarked, non-topic, non-*sa* argument. Within the quote, there is a clear predicate-topic construction, though the predicate is rather complex.

In line 27, the structure is quite different, as the quoting expression, *ang sagót ng réyna* ('the answer of the queen'), is the topic, and the quote itself is the predicate for this topic, the two forming an equative clause (compare line 19, above).

3. Summary and conclusion

In the text analysis above, we have identified the following phrase types:

- **X phrase**: links elements of many types in a modificational relationship. The head cannot be identified consistently using word order.
- **Y phrase**: links elements in an essentially possessive relationship. The semantic head (modified element) always appears in the initial position. The predicate and a non-topical, non-*sa*-marked argument in a non-equative clause also form a Y phrase.
- **LOC phrase**: marked by what appears to be similar to a preposition, but which takes an X or Y phrase (or single word) as its complement.
- **Topic phrase**: a single word, X phrase, or a Y phrase (including a whole clause) can appear as a topic, marked by *ang*/*'yung*, or, in the case of pronouns, have the topic form, or, in the case of proper human names, be marked by *si*/*sina*.
- **CONJ phrase**: conjoins two elements at any level.
- ***ay* phrase**: links a topic or locative or temporal element with the predicate when it appears before the predicate, contrasting with clauses where the topical elements appear after the predicate.
- **Unmarked phrase**: existential *may* can take an element within the predicate without overt marking of the relationship (*may* might include what was, historically, a linker). Certain other elements seem to be able to be combined into a predicate without overt marking of linkage as well (see lines 16 and 20, earlier).

The question, then, is whether these phrases correspond with the types of phrases we find in many other languages, such as noun phrases, verb phrases, and preposition phrases, or is this a system that works differently?

When writing reference grammars of languages, we often will have chapters on the noun phrase and the verb phrase, with sections within each chapter on the structure of that particular phrase type. If we were to write a reference grammar of Tagalog, could we legitimately have a chapter, for example, on the noun phrase, with a section on the structure of the noun phrase, where that structure is significantly different from the structure we would describe in the chapter on the verb phrase? It seems, from the discussion above, that the answer to this question would be 'no'. In which case, how then should we describe Tagalog?

My conclusion is that we should describe it on its own terms, as I have endeavored to do here, and not try to fit it into any a priori grammatical categories when such an approach is not appropriate.

Tagalog glossing abbreviations

Abbreviation	Meaning	Part of speech	Form
1pinclNT	first-person inclusive non-topic pronoun	second-position clitic pronoun	*natin*
3sgNT	third-person singular non-topic pronoun	second-position clitic pronoun	*niya ~ n'ya*
3sgT	third-person singular topic pronoun	second-position clitic pronoun	*siya ~ s'ya*
AC	*ang* complement (element that follows *ang*)		
AM	appositive modifier		
GER	gerund ('act of Ving')	derivational prefix	*pag-*
ASSOC	associative marker, marks reciprocal or joint activity	derivational prefix	*ka-*
CAUS	causative	derivational prefix	*pa-*
CC	clause complex		
CL	marks a clause that appears embedded within the predicate or topic of another clause		
COMP	complementizer	clause-initial particle	*kung*
CON	conveyance; marks a 'conveyed' topic	derivational prefix	*i-*
CONJ	conjunctive marker (can conjoin clauses or any other constituents)	particle (appears between conjuncts)	*at*
CONJP	conjunction phrase (formed with the conjunction *at*)		
CP	complement phrase		
DISLOC	distal locative pronoun	pronoun	*doon*
CSM	change of state marker	second-position clitic particle	*na*
EMPH	emphatic marker ('precisely', 'truly')	second-position clitic particle	*nga*
EX	existent (thing in an existential or negative existential clause that exists or does not exist)		
EXIST	existential and possessive	predicator	*may*
FT	links a predicate with a fronted topic	particle (occurs between topic and predicate)	*ay*
INHERENT.ABLE	marks an inherent ability	derivational prefix	*nakáka-*
LC	locative complement phrase		
LFS	location-forming suffix (forms elements that represent locations); when the word with this suffix is the predicate, the topic of the clause is a location ('locative focus')	derivational suffix	*-an ~ -han*

Abbreviation	Meaning	Part of speech	Form
LOC	locative	particle	*sa*
LNK	linker	clitic (occurs on first item) ~ particle (occurs between two items linked)	*-ng* ~ *na*
LP	locative phrase (phrase marked by *sa*)		
mod	modifier phrase		
N.EXIST	negative existential and possessive	predicator	*walaʔ*
OS	ongoing state marker	inflectional prefix	*naka-*
POL	politeness marker	second-position clitic	*po*
PRED	predicate		
PROSP	prospective aspect marker	prefix	*pa-*
POSPREF	shows possession of referent of root (*magkasakit* 'has illness')	prefix	*magka-*
PROXT	proximate topic pronoun	demonstrative pronoun	*ito*
Q	interrogative marker	second-position clitic	*ba*
REL	relational marker	particle (occurs between two linked items; *ni* is used before proper names; *nina* for more than one name)	*ng* [nɑŋ], *ni*
RPAT	realis perfective actor topic	derivational infix appearing after initial consonant of predicate or before vowel initial	*-um-*
RPUT	realis perfective undergoer topic	derivational infix appearing after initial consonant of predicate or before vowel initial	*-in-*
REDUP	reduplication (if of first syllable of predicate, marks imperfective)		
SIT.ABLE	'for', 'in order to', to make a situation come about	derivational prefix	*maká-*
SPEC	specific—marks form as identifiable (often marks topic) or simply instantiated	particle (*si* before a proper name; *sina* if more than one name)	*ang*, *si*
STAT	stative	derivational prefix	*ma-*
SUPER	superlative marker	derivational prefix	*pinaka-*
TOP	topic phrase		
VOC	vocative		
X	marks a phrase formed of items linked by LNK *ng* ~ *na*		

Abbreviation	Meaning	Part of speech	Form
Y	marks a phrase formed of items linked by REL *ng* [nɑŋ]		

References

Himmelmann, Nikolaus. 2008. Lexical categories and voice in Tagalog. *Voice and Grammatical Relations in Austronesian Languages*, ed. by Peter Austin & Simon Musgrave, 247–293. Stanford: CSLI.

Himmelmann, Nikolaus. (in press.) Notes on noun phrase structure in Tagalog. *Special issue of Australian Journal of Linguistics*, ed. by Michael Ewing & Simon Musgrave. http://www.uni-muenster.de/imperia/md/content/allgemeine_sprachwissenschaft/dozenten-unterlagen/himmelmann/himmelmann_np_dp_pp_structure_20_12.pdf.

Kaufman, Daniel. 2009. Austronesian nominalism and its consequences: a Tagalog case study. *Theoretical Linguistics* 35.1:1–49.

Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic*, *Focus*, *and the Mental Representations of Discourse Referents*. Cambridge & New York: Cambridge University Press.

LaPolla, Randy J. 2010. Feilübin Tagaluoyu (Tagalog) de cilei fanchou [The lexical categories of the Tagalog language of the Philippines]. *Yuyanxue Luncong* [*Essays on Linguistics*] 41:1–14. Beijing: The Commercial Press.

LaPolla, Randy J., & Dory Poa. 2005. Direct and indirect speech in Tagalog. Workshop on Direct and Indirect Speech, Research Centre for Linguistic Typology, June 8, 2005. Melbourne: La Trobe University. http://tibeto-burman.net/rjlapolla/papers/Direct_and_Indirect_Speech_in_ Tagalog.pdf

Reid, Lawrence A. 2000. Sources of proto-oceanic initial prenasalization: the view from outside Oceanic. *Grammatical Analysis: Morphology*, *Syntax*, *and Semantics: Studies in Honor of Stanley Starosta*, ed. by Videa P. De Guzman & Byron W. Bender, 30–45. Honolulu: University of Hawai‘i Press.

Reid, Lawrence A. 2002. Determiners, nouns, or what? Problems in the analysis of some commonly occurring forms in Philippine languages. *Oceanic Linguistics* 41.2:295–309.

[Received 2 September 2013; revised 3 April 2014; accepted 8 April 2014]

Division of Linguistics and Multilingual Studies
Nanyang Technological University
HSS-03-45
14 Nanyang Avenue
Singapore 637332
randylapolla@ntu.edu.sg

菲律賓塔伽洛語 (Tagalog) 句子成分的結構

羅仁地
南洋理工大學

語言學家對菲律賓塔伽洛語 (Tagalog) 詞類的分析至今仍有爭論。本文卻從另一個角度探討塔伽洛語詞類的問題：如果塔伽洛語沒有詞類（如名詞、動詞），那就不可能有名詞組和動詞組。為了探討這個問題，本文通過歸納法，分析一篇塔伽洛語長篇語料，逐行詳細地列清該語料中的組織成分。分析的結果顯示，語料中雖然明顯呈現幾種常用的結構，這些常用結構也明顯地呈現不同的功能，但這些常用的結構及其功能卻不同於印歐語系語言的名詞組和動詞組的結構及其相對的功能，即：不同結構雖然有不同的功能，但沒有專門用來作謂語的結構，也沒有專門用來指事的結構。

關鍵詞：菲律賓塔伽洛語，詞類，南島語系語言，詞組結構，語言類型學