<table>
<thead>
<tr>
<th>Title</th>
<th>Constituent structure in a tagalog text</th>
</tr>
</thead>
<tbody>
<tr>
<td>Author(s)</td>
<td>LaPolla, Randy J.</td>
</tr>
<tr>
<td>Citation</td>
<td>LaPolla, R. J. (2014). Constituent structure in a tagalog text. Language and linguistics, 15(6), 761-774.</td>
</tr>
<tr>
<td>Date</td>
<td>2014</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/10220/24087">http://hdl.handle.net/10220/24087</a></td>
</tr>
<tr>
<td>Rights</td>
<td>© 2014 The Author (Language and Linguistics). This paper was published in Language and Linguistics and is made available as an electronic reprint (preprint) with permission of The Author (Language and Linguistics). The paper can be found at the following official DOI: [<a href="http://dx.doi.org/10.1177/1606822X14544619">http://dx.doi.org/10.1177/1606822X14544619</a>]. One print or electronic copy may be made for personal use only. Systematic or multiple reproduction, distribution to multiple locations via electronic or other means, duplication of any material in this paper for a fee or for commercial purposes, or modification of the content of the paper is prohibited and is subject to penalties under law.</td>
</tr>
</tbody>
</table>
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. 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) Alamá t ng Gú bat
[alamá t ng gú bat] 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.


2 When written alone, ng is pronounced [ŋaŋ]. 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 ([n̥aŋ]), 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’ 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 únâng panahón,
    [noón=ng ún=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 hāri (‘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, kaālaman ‘wisdom’ (< alām; ‘know’), kagandāhan ‘beauty’ (< gandā; ‘beauty’), and kabuhāyan ‘livelihood’ (< būhay; ‘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 kaharián, the sense is ‘a place where (people) have the same king’, so it seems ka- would be affixed first, then -an.
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.

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) Pé ro wala siyang kinaláman
pé ro   [wala’<siya>LG]X 3sgT=LNK involvement
but   N.EXIST   N.EXIST 3sgT=LNK involvement

But she doesn’t have anything to do

(7) sa kwéntong ito.
[sa [kwénto=ng ito]X1C]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 káalaman (‘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 pansin
káya [ang [{pag-tu-tuón-an} na lang natin}Y
therefore SPEC GER-REDUP-emphasis-LFS CSM only/just 1pinclNT
ng pansin],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/iyung* can be followed by a single word, an X phrase, or a Y phrase. In line 8, two Y phrases (*pagtutuúnan ng pansin* 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) ay si Tong,
  ay [si [Tong]
  FT SPEC 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 Talangká?

\[
\text{ang} \quad \text{pinakabátang anák} \quad \text{ni} \quad \text{Háring Talangká?}
\]

\[
\text{SPEC} \quad \text{SUPER-young=LNK} \quad \text{child} \quad \text{REL} \quad \text{king=LNK} \quad \text{crab}
\]

the youngest child of King Crab,

(11) na túlad ng maráming hári

\[
\text{na} \quad \text{túlad ng} \quad \text{maráming hári}
\]

\[
\text{LNK} \quad \text{just.like} \quad \text{REL} \quad \text{STAT-many=LNK} \quad \text{king}
\]

who, like many kings,

(12) ay walá ng ibáng papél sa kwénto

\[
\text{ay} \quad \text{walá} \quad \text{ng} \quad \text{ibáng papél} \quad \text{sa} \quad \text{kwénto}
\]

\[
\text{FT} \quad \text{N.EXIST=LNK} \quad \text{other=LNK} \quad \text{role} \quad \text{LOC} \quad \text{story}
\]

has no role in (the) story

(13) kundí? ang magkasakit.

\[
\text{kundí} \quad \text{ang} \quad \text{magkasakit}
\]

\[
\text{COMP} \quad \text{NEG} \quad \text{SPEC} \quad \text{POSPREF-illness}
\]

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 Talangká?* is the predicate for the fronted topic *ang pagtutuúnan na lang natin ng pansin*, with *ang pinakabátang anák ni Háring Talangká?* modifying *si Tong* as an appositional modifier; *pinakabátang anák ni Háring Talangká?* is a Y phrase, which includes the X phrases *pinakabátang anák* and *Háring Talangká?*.

*Háring Talangká?* 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 *walá ng ibáng papél sa kwénto kundí? ang magkasakit*, the latter being itself a clause complex with two clauses, the first of which has *Háring Talangká?* as the (sub)topic and the second of which (a subordinate clause marked by *kung*) has *ang magkasakit* as the topic.

(14) At íto na nga ang dahilán

\[
\text{At} \quad \text{íto} \quad \text{na} \quad \text{nga} \quad \text{ang} \quad \text{dahil-an}
\]

\[
\text{CONJ} \quad \text{PROXT} \quad \text{CSM} \quad \text{EMPH} \quad \text{SPEC} \quad \text{because-LFS}
\]
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 kanyá ng inang réyna* has *itó na ngá* as predicate and *ang dahilán kung bákit isáng áraw ay biglá na lang ipinatáwag si Tong ng kanyá ng inang réyna* as topic, with this topic including the subordinated modifier *kung bákit isáng áraw ay biglá na lang ipinatáwag si Tong ng kanyá ng inang réyna* as topic, with the reference to the Queen.

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.

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 \textit{ang}, and the predicate takes the form of an existential phrase. This sort of structure can be used for interrogative word questions, such as [\textit{Ano} [\textit{ang gusto mo}]]? [\textit{what SPEC want 2sg\textsuperscript{NT}}] ‘What do you want?’; for attributive predication, such as [\textit{Titsér} [\textit{ako}]] ‘I am a teacher’; and for identificational predication, such as [\textit{Ito} [\textit{ang gusto ko}]] [\textit{PROXT SPEC want 1sg\textsuperscript{NT}}] ‘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, \textit{wí ka ng ré yna}—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.
\[
\text{[hindíʔ na <siya\textsubscript{TOP} nakáka-langóy\textsubscript{PRED}]}
\text{NEG CSM 3sg\textsuperscript{T} INHERENT.ABLE-SWIM}
\]
‘(He) is not able to swim anymore.

In line 20, again, two elements (\textit{hindíʔ} and \textit{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
\[
\text{[kailángan <mo>=ng [um-áhon ngayón din\textsuperscript{11}]\textsubscript{PRED}]}
\text{need 2sg\textsuperscript{NT}=LNK RPAT-get.up now also}
\]
You need to get up right now (and)

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

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

(24) ang táning prutas na makakapágpagaling sa kanya.’
\[
\text{[ang [tangi=ng prutas na [makáka-pag-pa-galing [sa kanya\textsubscript{LP=MOD=AM}]\textsubscript{Y}]
\text{SPEC only=LNK fruit LNK SIT.ABLE-GER-CAUS-recover LOC 3sg\textsuperscript{POS}}}
\]
the only fruit that can cure him.’

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

\begin{footnotesize}
\begin{enumerate}
\item The combination of \textit{ngayon} (‘now’) and \textit{din} (‘also’) means ‘right now’.
\end{enumerate}
\end{footnotesize}
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, makakapápgagaling 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) hindi ba’t talagá namang hindi nakákalangóy12
[hindíiez ba at talagá namán=ng hindíí namang=ng hindi? 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>TOP=ng [ka-ramdám-an]EX/PRED]PREDz]PREDz]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 hindi nakákalangóy ‘never could swim’ differs from hindi 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

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
<th>Part of speech</th>
<th>Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>1pinclNT</td>
<td>first-person inclusive non-topic pronoun</td>
<td>second-position clitic pronoun</td>
<td>natin</td>
</tr>
<tr>
<td>3sgNT</td>
<td>third-person singular non-topic pronoun</td>
<td>second-position clitic pronoun</td>
<td>niya ~ n’ya</td>
</tr>
<tr>
<td>3sgT</td>
<td>third-person singular topic pronoun</td>
<td>second-position clitic pronoun</td>
<td>siya ~ s’ya</td>
</tr>
<tr>
<td>AC</td>
<td><em>ang</em> complement (element that follows <em>ang</em>)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AM</td>
<td>appositive modifier</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GER</td>
<td>gerund (‘act of Ving’)</td>
<td>derivational prefix</td>
<td>pag-</td>
</tr>
<tr>
<td>ASSOC</td>
<td>associative marker, marks reciprocal or joint activity</td>
<td>derivational prefix</td>
<td>ka-</td>
</tr>
<tr>
<td>CAUS</td>
<td>causative</td>
<td>derivational prefix</td>
<td>pa-</td>
</tr>
<tr>
<td>CC</td>
<td>clause complex</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CL</td>
<td>marks a clause that appears embedded within the predicate or topic of another clause</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COMP</td>
<td>complementizer</td>
<td>clause-initial particle</td>
<td>kung</td>
</tr>
<tr>
<td>CON</td>
<td>conveyance; marks a ‘conveyed’ topic</td>
<td>derivational prefix</td>
<td>i-</td>
</tr>
<tr>
<td>CONJ</td>
<td>conjunctive marker (can conjoin clauses or any other constituents)</td>
<td>particle (appears between conjuncts)</td>
<td>at</td>
</tr>
<tr>
<td>CONJP</td>
<td>conjunction phrase (formed with the conjunction <em>at</em>)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CP</td>
<td>complement phrase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISLOC</td>
<td>distal locative pronoun</td>
<td>pronoun</td>
<td>doon</td>
</tr>
<tr>
<td>CSM</td>
<td>change of state marker</td>
<td>second-position clitic particle</td>
<td>na</td>
</tr>
<tr>
<td>EMPH</td>
<td>emphatic marker (‘precisely’, ‘truly’)</td>
<td>second-position clitic particle</td>
<td>nga</td>
</tr>
<tr>
<td>EX</td>
<td>existent (thing in an existential or negative existential clause that exists or does not exist)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EXIST</td>
<td>existential and possessive</td>
<td>predicator</td>
<td>may</td>
</tr>
<tr>
<td>FT</td>
<td>links a predicate with a fronted topic</td>
<td>particle (occurs between topic and predicate)</td>
<td>ay</td>
</tr>
<tr>
<td>INHERENTABLE</td>
<td>marks an inherent ability</td>
<td>derivational prefix</td>
<td>nakáka-</td>
</tr>
<tr>
<td>LC</td>
<td>locative complement phrase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LFS</td>
<td>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’)</td>
<td>derivational suffix</td>
<td>-an ~ -han</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Meaning</td>
<td>Part of speech</td>
<td>Form</td>
</tr>
<tr>
<td>--------------</td>
<td>----------------------------------------------</td>
<td>---------------------------------</td>
<td>--------------------</td>
</tr>
<tr>
<td>LOC</td>
<td>locative particle</td>
<td>particle</td>
<td>sa</td>
</tr>
<tr>
<td>LNK</td>
<td>linker clitic (occurs on first item) ~</td>
<td>second-position clitic</td>
<td>po</td>
</tr>
<tr>
<td>LP</td>
<td>locative phrase (phrase marked by sa)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>mod</td>
<td>modifier phrase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.EXIST</td>
<td>negative existential and possessive predicator</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OS</td>
<td>ongoing state marker</td>
<td>inflectional prefix</td>
<td>naka-</td>
</tr>
<tr>
<td>POL</td>
<td>politeness marker</td>
<td>second-position clitic</td>
<td>po</td>
</tr>
<tr>
<td>PRED</td>
<td>predicate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROSP</td>
<td>prospective aspect marker</td>
<td>prefix</td>
<td>pa-</td>
</tr>
<tr>
<td>POSPREF</td>
<td>shows possession of referent of root (magkasakit 'has illness')</td>
<td>prefix</td>
<td>magka-</td>
</tr>
<tr>
<td>PROXT</td>
<td>proximate topic pronoun</td>
<td>demonstrative pronoun</td>
<td>ito</td>
</tr>
<tr>
<td>Q</td>
<td>interrogative topic marker</td>
<td>second-position clitic</td>
<td>ba</td>
</tr>
<tr>
<td>REL</td>
<td>relational marker</td>
<td>particle (occurs between two linked items; ni is used before proper names; nina for more than one name)</td>
<td>ng [nan], ni</td>
</tr>
<tr>
<td>RPAT</td>
<td>realis perfective actor topic</td>
<td>derivational infix appearing after initial consonant of predicate or before vowel initial</td>
<td>-um-</td>
</tr>
<tr>
<td>RPUT</td>
<td>realis perfective undergoer topic</td>
<td>derivational infix appearing after initial consonant of predicate or before vowel initial</td>
<td>-in-</td>
</tr>
<tr>
<td>REDUP</td>
<td>reduplication (if of first syllable of predicate, marks imperfective)</td>
<td>derivational prefix</td>
<td>maká-</td>
</tr>
<tr>
<td>SIT.ABLE</td>
<td>'for', 'in order to', to make a situation come about</td>
<td>derivational prefix</td>
<td>maká-</td>
</tr>
<tr>
<td>SPEC</td>
<td>specific—marks form as identifiable (often marks topic) or simply instantiated</td>
<td>particle (si before a proper name; sina if more than one name)</td>
<td>ang, si</td>
</tr>
<tr>
<td>STAT</td>
<td>stative</td>
<td>derivational prefix</td>
<td>ma-</td>
</tr>
<tr>
<td>SUPER</td>
<td>superlative marker</td>
<td>derivational prefix</td>
<td>pinaka-</td>
</tr>
<tr>
<td>TOP</td>
<td>topic phrase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>VOC</td>
<td>vocative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>X</td>
<td>marks a phrase formed of items linked by LNK ng ~ na</td>
<td>derivational prefix</td>
<td></td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Meaning</td>
<td>Part of speech</td>
<td>Form</td>
</tr>
<tr>
<td>--------------</td>
<td>---------</td>
<td>---------------</td>
<td>------</td>
</tr>
<tr>
<td>Y</td>
<td><em>marks a phrase formed of items linked by <code>rel</code> ng [nan]</em></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**References**


[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
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，這些常用的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