Constituent structure in a Tagalog text

Randy J. LaPolla
La Trobe University

This paper is an inductive look at the constituents found in a randomly selected Tagalog text, Bob Ong’s *Alamat ng Gubat* (Makati City, MM: Visual Print Enterprises, 2004). The analysis is based on the full text, but we will only be able to go through the first few lines of the text here, which we will do one by one, and discuss the structures found in each line of the text in bullet format after the relevant line. At the end of the paper we will bring up some important questions about the structures found in Tagalog based on this text.

1  

*Alamat ng Gubat*

[alamá̱t ng gúbat]_{Y/TITLE}

legend REL jungle

(The) Legend of the Jungle

• 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* ([n̥ŋ]), if the word following this marker is a common noun, or *ni*, if the word following this marker is a proper name. This sort of structure manifests the following set of modifying functions: part-whole (inalienable possession), possession (alienable), “possession” with locative and abstract nouns (such as in line 1), adverbial modification, and predicate-argument structure.

• In this structure the head of the phrase is the first constituent, and the modifying element follows *ng* or *ni*.

2  

*Noóng únā̱ng 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/na*. This structure manifests a much more varied set of modifying functions than 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

---

1 See table at end for abbreviations used. 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, e.g. “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. The book should be consulted for that level of structure.

2 I use quotes around “adjectival” here and elsewhere in the paper, as there is much controversy about the definition of form classes in Tagalog. See for example Himmelmann 2008 for discussion.
existant in an existential predicate), indirect quotes, certain types of possessive modification, nonpossessive 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.

3. \( sa \text{ isáng } \text{liblíb } \text{na kahárián} \)

\[
[sa \text{ [isá-ng } \text{liblíb } \text{na } \text{kahárián}^3]_{\text{X/LC}}_{\text{LP}}
\]

in a remote kingdom

4. \( sa \text{ ilálim ng dágat} \)

\[
[sa \text{ [ilálim } \text{ng } \text{dágat]}_{\text{X/LC}}_{\text{LP}}
\]

at the bottom of the sea,

- The third type of phrase, found 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 non-direct arguments.\(^4\)
- The “preposition” \( sa \) can take a single word or an X phrase or a Y phrase as complement. In line 3 it takes an X phrase as complement and in line 4 it takes a Y phrase as complement. It is somewhat problematic to call \( sa \) a preposition, 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 clearly nominal in a grammatical sense.
- There are two LOC phrases in these two lines, 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. \( ay \text{ may nakatíráng magandá } \text{at mabait } \text{na siséna} \)

\[
[ay \text{ [may } \text{naka-tírá=ng]}
\]

at the bottom of the sea,

---

\(^3\) The word \( \text{kahárián} \) is formed from the root \( \text{hári 'king'} \), plus the two affixes \( \text{ka- ASSOC and -an LFS} \). The two are independent affixes, but are commonly used together to form abstract nouns, e.g. \( \text{kaalám} \) 'wisdom' (< \( \text{alám} \) 'know'), \( \text{kagandáhan} \) 'beauty' (< \( \text{gandá} \) 'beauty'), \( \text{kabuháyan} \) 'livelihood' (< \( \text{búhay} \) '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, e.g. for \( \text{kahárián} \), the sense is ‘a place where (people) have the same king’, so it seems \( \text{ka-} \) would be affixed first, then \( \text{-an} \).

\(^4\) For locational predications, \( \text{nasa} \), rather than \( \text{sa} \), is generally used, e.g. \( \text{Nasa gubat siya 'He is in the forest.'} \)
there was a beautiful and nice mermaid who lived there.

- *Ay* at the beginning of line 5 marks the fact that the element before this marker is the topic, and the element following the marker is the predicate of (comment about) that topic. Lines 2-4 all relate to this predicate. By “topic” here, were are talking of the grammatical pivot (grammatically privileged argument) of the construction, the argument singled out for special morphosyntactic treatment. As this pivot is chosen by the speaker on the basis of what the clause is about, it is called the topic.

- The predicate (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 or an X phrase or Y phrase as complement. In this case it takes an X phrase as complement. If *doon* ‘there’ compounds with *may* in the predicate (e.g. in line 27: *mayroon* *s’yang karamdámán* ‘he has an illness’), then the combined form *mayroon* is linked to the existant (the thing that exists) in an X phrase. This structure is also used for possession, with the possessor as topic, as in in 27.

- A fifth 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

\[ \text{Péро wala siyang kinaláman} \]
\[ \text{péro} \quad \text{[wala?} \quad \text{<siya>_{TOP}}=\text{ng} \quad \text{[kinaláman^5}_{EX}]_{X/PRED} \]

But she doesn't have anything to do

7

\[ \text{sa kwéntong ito.} \]
\[ \text{[sa} \quad \text{[kwénto}=\text{ng} \quad \text{ito]_{X/LC}}_{LP} \]

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 existant can appear as part of the predicate. This structure is also used for (negative) possession, as in this example, with the possessor as topic. It is also possible for the existant to appear as an *ang*-marked topic (see below on *ang*), such as later in the story, where the protagonist (Tong) says “*Wala po ako =ng pílak . . .*” [N.EXIST POL 1sgT=LNK silver] ‘I don’t have any silver . . .’.

- 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

---

5 The word *kinaláman* ‘involvement’ is formed from the root *álám* 'know', plus *ka-* ASSOC and -an LFS, together forming *kaálaman* 'knowledge, wisdom', and then the infix -ín- RPUT appears within the *ka-* prefix. The resulting form *kinaalaman* then obligatorily reduces to *kinaláman*. 


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-tción-an} na lang natin}]_{Y}  
therefore SPEC ACTNOM-REDUP-emphasis-LFS CSM only/just 1plinclNT

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) marks the constituent following it as topic of the clause. The particle ang 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  
ay si Tong,  
ay [si [Tong]}_{X}  
FT SPEC PN  
is Tong,

- 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 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 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 (the oblique marker for proper names) would imply it is an oblique argument, which it isn’t, and so si is used here.

10  
ang pinakabátang anák ni Háring Talangká?  
[ang [{pinaká-báta=ng anák}]_{X} ni [{hári=ng talangká?}]_{Y}X/TOPi  
SPEC SUPER-young=LNK child REL king=LNK crab  
the youngest child of King Crab,

---

6 Cf. tutuúnan 'concentration', but it seems the order of affixing is first to add pag-, forming pagtutuon, and then -an is added to that.
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,

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

kundi? ang magkasakit.
kung [[hindi?]PRED [ang magka-sakit]TOPYL>CC>PRED]CLP
COMP NEG SPEC POSPREF-illness
other than to be the one who become sick.

• The passage from line 7 up 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 pansín, with ang pinakabátang anák ni Háring Talangká? modifying si Tong as an appositional (non-restrictive) 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 (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 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.

At íto na ngá ang dahilán
at [[[ito na nga]PREDj] {ang [dáhil-án}
CONJ this CSM EMPH SPEC because-LFS

kung bákit isáng áraw
[kung [bákit [isá=ng áraw]XTOPz]
COMP why one=LNK day
and it is the reason why one day

ay biglá na lang ipinatáwag
ay [{biglá? na lang i-in+pa-táwag}PREDi
FT sudden CSM only CON-RPUT-CAUS-call

si Tong ng kanyáng inang reyna.
{si Tong}TOPi [ng [kanyá=ng ina=ng reyna]X]YA>YCL>PREDz CLNCL>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 ones preceding it (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: *ító na ngá ang dahlán kung bákit isáng áraw ay biglá na lang ipinatáwag si Tong ng kanyáng inang reyna* has *ító na ngá* as predicate and *ang dahlán kung bákit isáng áraw ay biglá na lang ipinatáwag si Tong ng kanyáng inang reyna* as topic, with this topic including the subordinated modifier *kung bákit isáng áraw ay biglá na lang ipinatáwag si Tong ng kanyáng inang reyna*.

• 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, *bigla 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


PN child SPEC 2sgPOSS=LNK father

Tong, Child, your father

19

ay may karamdáman," wíka ng reyna.


FT EXIST ASSOC-feeling-LFS language REL queen

has an illness", said the queen.

• Lines 18-19 form a single clause where 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.

• The topic of the higher structure (the clause that has the quote embedded as the predicate) is a Y phrase, *wíka ng reyna*, literally ‘the language of the queen’, which is not marked as a topic, but seems to form an equative clause with the quote. (In line 27 we have the same structure, and it is marked overtly as the topic. See LaPolla and Poa 2005 on speech act constructions in Tagalog.)

20

"Hindi na s'ya nakákalangóy.

[hindiʔ 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 (*hindiʔ 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.
Kailángan mong umáhon ngayón din papúnta sa lúpa

You need to get up right now (and)

Pa-púnta sa lúpa

go to the land

úpang kumúha ng púso ng ságing—
to get the heart of the banana—

ang tánging prutas na makakapágpagaling sa kanya."

the only fruit that can cure him.

Sumagót si Tong, "Ngúnit ináng réyna,

Tong answered, "But, Mother Queen,

7 The combination of ngayon ‘now’ and din ‘also’ means ‘right now’.
"Dáhil nga mayroon s'yang karamdáman!", ang sagót ng rénya.

The reply (answer) of the queen was "Because he has an illness!".

In 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 Actor Topic, and the quote is an unmarked non-topic direct 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 rénya 'the answer of the queen', is the topic, and the quote itself is the predicate for this topic, the two forming an equative clause (see LaPolla & Poa 2005).

Summary and Conclusion
We have identified the following phrase types:

• X phrase: links elements of many types in a modificational relationship. The head cannot be identified morphosyntactically.

• Y phrase: links elements in a basically possessive relationship. The semantic head always appears in initial position. The relationship between the predicate and the non-topical direct arguments is also marked as a Y phrase.

• LOC phrase: marked by what appears to be like a preposition, but can take X or Y phrase (or single word) as complement.

• Topic phrase: a single word or an X phrase or a Y phrase can appear as topic, marked by ang or, in the case of pronouns, have the topic form, or in the case of proper human names, marked by si.

• CONJ phrase: conjoins two elements at any level.

• 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.

The question then is do these phrases correspond with the types of phrases we find in many other languages, such as noun phrase, verb phrase, and preposition phrase, or do we have 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 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 would be negative. In that case, how then should we describe Tagalog?

My answer is that we should describe it on its own terms, as I did above, and not try to fit it into Indo-European grammatical categories when that 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>1plincNT</td>
<td>first person inclusive non-topic pronoun</td>
<td>second position clitic pronoun</td>
<td>natin</td>
</tr>
<tr>
<td>3sgNT</td>
<td>3rd person singular non-topic pronoun</td>
<td>second position clitic pronoun</td>
<td>niya ~ n'ya</td>
</tr>
<tr>
<td>3sgT</td>
<td>3rd person singular topic pronoun</td>
<td>second position clitic pronoun</td>
<td>siya ~ s'ya</td>
</tr>
<tr>
<td>AC</td>
<td>Ang complement (element that follows ang)</td>
<td>derivational prefix</td>
<td>pag-</td>
</tr>
<tr>
<td>AM</td>
<td>appositive modifier</td>
<td>derivational prefix</td>
<td>ka-</td>
</tr>
<tr>
<td>ACTNOM</td>
<td>action nominalizer ('act of Ving')</td>
<td>derivational prefix</td>
<td>pag-</td>
</tr>
<tr>
<td>ASSOC</td>
<td>associative marker, marks reciprocal or joint activity, e.g. kasama ‘go together’, kapatid ‘brother’</td>
<td>derivational prefix</td>
<td>pag-</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</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 (one formed with the conjunction at)</td>
<td></td>
<td></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 ('precicely', 'truely')</td>
<td>second position clitic particle</td>
<td>nga</td>
</tr>
<tr>
<td>EX</td>
<td>existant (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>marker for showing that the clause has a fronted topic (it appears before the predicate—a marked construction)</td>
<td>particle (occurs between topic and predicate)</td>
<td>ay</td>
</tr>
<tr>
<td>INHERENT.ABLE</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>suffix</td>
<td>-an ~ -han</td>
</tr>
<tr>
<td>LOC</td>
<td>locative</td>
<td>particle</td>
<td>sa</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Description</td>
<td></td>
<td></td>
</tr>
<tr>
<td>--------------</td>
<td>-------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>LNK</td>
<td>linker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MOD</td>
<td>modifier phrase</td>
<td></td>
<td></td>
</tr>
<tr>
<td>NC</td>
<td>noun complement</td>
<td></td>
<td></td>
</tr>
<tr>
<td>N.EXIST</td>
<td>negative existential and possessive marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>OS</td>
<td>ongoing state marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POL</td>
<td>politeness marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PRED</td>
<td>predicate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PROSP</td>
<td>prospective aspect marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>POSPREF</td>
<td>shows possession of referent of root</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REL</td>
<td>relational marker</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPAT</td>
<td>Realis perfective actor topic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>RPUT</td>
<td>Realis perfective undergoer topic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>REDUP</td>
<td>reduplication (if of first syllable of predicate, marks imperfective—on-going or future)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SIT.ABLE</td>
<td>‘for’, ‘in order to’, to make a situation come about</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SPEC</td>
<td>specific—marks form as identifiable (often marks topic); si is used a proper name that is the topic within a clause and also when referring to a person using their name</td>
<td></td>
<td></td>
</tr>
<tr>
<td>STAT</td>
<td>stative</td>
<td></td>
<td></td>
</tr>
<tr>
<td>SUPER</td>
<td>superlative marker</td>
<td></td>
<td></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></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>marks a phrase formed of items linked by REL ng [naŋ]</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

References
