Information Structure and Constituent Order in Tagalog*

Naonori Nagaya
The University of Tokyo

The goal of this paper is to spell out the way syntax and pragmatics interact with each other inside and outside the clause in Tagalog. Inside the clause, different constructions are employed to express different types of focus structure: a canonical construction for predicate focus and sentence focus, a cleft construction for argument narrow focus, and a fronting construction for adjunct narrow focus. In every case the clause-initial position is included in the actual focus domain. Outside the clause, on the other hand, a topical/presupposed element is optionally preposed to the sentence-initial position called the left-detached position. It will be shown that this sentence-initial position is clearly distinguished from the clause-initial position. The traditionally recognized contrast between sentence and clause is indisputable and unequivocal in this language.

Key words: Tagalog, information structure, focus structure, constituent order, cleft, fronting

1. Introduction

This paper aims to elucidate aspects of the interplay between morphosyntax and information structure of Tagalog within the framework of Role and Reference Grammar with special reference to its theories of clause structure and information structure (Foley and Van Valin 1984, Van Valin 1993, Van Valin & LaPolla 1997, Van Valin 2005).

* An earlier version of this paper was presented at the 2005 International Course and Conference on Role and Reference Grammar, held at Academia Sinica, Taipei, in June 2005 (Nagaya 2005). Many thanks to my consultants for Tagalog, who shared their precious native-speaker intuitions with me. Thanks, also, to Tooru Hayasi, Daphne Lindaya, Joe V. Lopresti, Yoshiki Nishimura, Ricardo Ma. Nolasco, Toshio Ohori, Masayoshi Shibatani, Tasaku Tsunoda, and Robert D. Van Valin Jr. for their indispensable comments and suggestions. I am also grateful to two anonymous reviewers for their eye-opening and detailed comments on this paper. Of course, responsibility for any errors is purely my own. Special thanks to Hsiu-chuan Liao for sending an electronic copy of her dissertation to me, to Elizabeth Zeitoun for her dedicated effort to edit this volume, and to Danilo Bv. Tuazon and Susie Tuazon for providing the hospitality and spiritual help that made this research possible. Lastly, I wish to acknowledge my gratitude to the Japan Society for the Promotion of Science for supporting this research financially.
This paper has a twofold goal and revolves around clausal boundaries. In the first part, I will confine the discussion to the syntax-pragmatics interface inside the clause. After some basic facts and concepts are reviewed in §2, it will be argued in §3 that different focus structures are expressed by different constructions in Tagalog. In broad focus a focal predicate occupies the clause-initial nucleus, whereas in narrow focus, a focused non-predicating element appears in the clause-initial position by means of a cleft construction for an argument, and a fronting construction for an adjunct. This means that in any case the clause-initial position is included in the actual focus domain. This will be examined in §4.

In the second part, I will expand the scope of the discussion to include the entire sentence, taking into account a sentence-initial optional position called the left-detached position, which houses a topical or presupposed element. In §5, I will present a description of characteristics of the left-detached position in Tagalog and demonstrate how this position can be distinguished from the clause-initial position, by employing syntactic tests including enclitic placement and a pause.

In this way, syntax and pragmatics interact differently inside and outside the clause. Clause-externally, the focus domain obligatorily includes the clause-initial position; clause-externally a topical/presupposed element optionally occupies the sentence-initial position. I will deal with this issue and conclude this paper in §6.

Some terminological notes are in order before turning to the main body of the analyses. In this paper, the term ‘topic’ is employed only in the sense of Gundel (1988) and Lambrecht (1994), although in the literature of Philippine linguistics it has been used for what I call absolutive in this paper (for example, Schachter & Otanes 1972, Schachter 1976, 1977). Likewise, the terms ‘focus’, ‘focus structure’, ‘broad focus’, ‘predicate focus’, ‘sentence focus’, and ‘narrow focus’ are used strictly in Lambrecht’s sense. Note that although it includes ‘focus’ in its name, a ‘focus system’ is the term for voice/valence-marking morphology found in Tagalog and other Philippine-type languages, and is irrelevant to the focus at issue in the paper.

2. Background


In this section, some basic facts and concepts in Tagalog will be reviewed: clause structure in §2.1, and the ‘focus system’ and relativization in §2.2. Some special
constructions relevant to this paper will be introduced in §2.3.

2.1 Layered structure of the clause in Tagalog

The theory of clause structure in Role and Reference Grammar is called the layered structure of the clause (Van Valin & LaPolla 1997:Ch. 2, and Van Valin 2005: Ch. 1). In this theory, the clause is composed of the core and the periphery. The core includes the nucleus and its core argument(s), which correspond to the predicate and the argument(s), respectively. The periphery refers to a syntactic unit housing non-arguments or adjuncts of the clause such as a PP adjunct and an adverb. The sentence is an even larger syntactic unit which may include multiple clauses in complex sentences or an additional sentence-initial position.

There are two types of pragmatically motivated optional position: a precore slot and a left-detached position. The precore slot is the position in which question words appear in languages in which they do not occur in situ, and is inside of the clause but outside of the core. The left-detached position is the location of sentence-initial elements, which are set off from the clause by a pause, and is outside of the clause but within the sentence.

An English sentence with all these elements is given in (1) and its constituent projection in Figure 1 (the example is cited from Van Valin & LaPolla 1997:36).

(1) Yesterday, what did Robin show to Pat in the library?

Now let us consider the Tagalog example in (2) and the constituent projection in Figure 2. In the nucleus is the verb \textit{binili}, the core argument positions are filled individually with the pronominal enclitic \textit{ko} and the noun phrase \textit{ang bulaklak}, and the periphery includes the PP adjunct \textit{para kay Weng} and the bare NP adverb \textit{kahapon}.
Naonori Nagaya

(2) \(B \text{ili=} \text{ko=} \text{na} \ ang=\text{bulaklak} \ para \ kay=\text{Weng} \ \text{kahapon.}\)

\[\text{PF:bought=}1\text{SG.ERG=already} \ \text{Abs=flower} \ \text{for} \ \text{DAT=Weng} \ \text{yesterday}\]

‘I have already bought the flower for Weng yesterday.’

\[\text{SENTENCE} \rightarrow \text{CLAUSE} \rightarrow \text{CORE} \rightarrow \text{PERIPHERY}\]

\[\text{NUC} \rightarrow \text{ARG} \rightarrow \text{ARG} \rightarrow \text{PP} \rightarrow \text{ADV}\]

\[B \text{ili=} \text{ko=} \text{na} \ ang=\text{bulaklak} \ para \ kay=\text{Weng} \ \text{kahapon.}\]

**Figure 2:** Tagalog layered structure of the clause

In Tagalog, the nucleus is prototypically a verb but can be nonverbal. The verb has a special feature called the ‘focus system’, which will be discussed in connection with relativization in the following subsection. The nucleus is located in the core-initial position and is clause-initial, too, as long as the precore slot is null.

The core argument position is occupied by a pronominal enclitic or a noun phrase. Noun phrases are introduced by a determiner (Reid & Liao 2004a, b). Both noun phrases and pronominal enclitics inflect for case: absolutive,\(^1\) ergative/genitive,\(^2\) and dative.

It is of significance to point out that in Tagalog enclitics including the pronominal enclitic are so-called second position clitics, and have to appear in the second position of the clause (Kroeger 1993, 1998). They include na ‘already’, pa ‘still’, lang/lamang ‘only, just’, nanan ‘indeed’, and ba ‘question’ in addition to the pronominal enclitic. Clitics and hosts thereof are linked by ‘=’ in this paper.

The periphery includes a PP adjunct headed by a preposition such as para ‘for’ and dahil ‘because’, and a bare NP adverb such as kahapon ‘yesterday’ and bukas ‘tomorrow’.

Examples with the precore slot and left-detached position will be given later in the paper. The syntactic and pragmatic contrast between them plays an important role in

---

1. The absolutive case marks the single core argument in an intransitive clause (S), and the undergoer core argument in a transitive clause (UT). As will be shown later in §5, a left/right-detached noun phrase is also marked in this case.

2. The ergative case and the genitive case are formally identical but functionally different. The ergative case marks the actor core argument in a transitive clause (AT). In contrast, the genitive case is used to introduce a possessive phrase and an adverbial phrase.
understanding the distinction between sentence and clause, which is the issue of the second part of this paper.

2.2 ‘Focus system’ and relativization

Tagalog and other so-called Philippine-type languages are known for their voice/valence-marking verbal morphology called the ‘focus system’. This can be analyzed as an equipollent verb derivational system, where both transitive and intransitive verbs are derived from the same stem by adding different affixes (Nagaya 2006a, c, Haspelmath 1993 for equipollent derivation, also Nichols et al. 2004). Verbs are marked by one of the four affixes: m-/-um- (Actor Focus (AF)), -in (Patient Focus (PF)), -an (Locative Focus (LF)) and i- (Circumstantial Focus (CF)). Again, note that ‘focus’ of the focus system is distinct from ‘focus’ of focus structure.

Basically, AF verbs are intransitive, except for meteorological verbs, and the single core argument is marked in the absolutive case. In contrast, PF, LF, and CF verbs are either intransitive or transitive: an undergoer is encoded in the absolutive case, while an actor, if any, is marked in the ergative case.

One of the important functions of these verbal affixes is relativization. In Tagalog, only the core argument in the absolutive case (S, or UT) can be relativized in a clause. For example, the single core argument ang bata ‘child’ in the intransitive clause (3a) can be relativized as in (3b). (Henceforth, the relative clause is put in parentheses.)

\[(3) \quad a. \quad \text{P<um>unta} \quad \text{ang=bata} \quad \text{sa=simbahan}. \quad \text{AF:went} \quad \text{ABS=child} \quad \text{DAT=church} \]
\[\quad \text{‘The child went to the church.’} \]
\[\quad b. \quad \text{bata=ng} \quad \text{[p<um>unta} \quad \text{[S] sa=simbahan]} \]
\[\quad \text{child=LIN} \quad \text{AF:went} \quad \text{DAT=church} \]
\[\quad \text{‘the child who went to the church’} \]

In the transitive clause (4a), the undergoer argument ang isda ‘fish’ can be a pivot for relativization as in (4b), whereas the actor core argument ng bata ‘child’ cannot be extracted as in (4c)

\[\text{\underline{3} Note that the productive AF marker mag- is analyzed to be composed of m- + pag-, and mang-, m- + pang-.}\]
(4) a. B<in>ili ng=bata ang=isda.
   PF:bought ERG=child ABS=fish
   ‘The child bought the fish.’

b. isda=ng [b<in>ili ng=bata [Uₜ]]
   fish=LIN PF:bought ERG=child
   ‘the fish which the child bought’

c. * bata=ng [b<in>ili [Aₜ] ang=isda]
   child=LIN PF:bought ABS=fish
   Intended for ‘the child who bought the fish’

However, different verbal affixes can be used to relativize different participants. The
actor argument above becomes a pivot when the AF form is employed instead of the PF
form, as in (5).

(5) bata=ng [b<um>ili [S] ng=isda]
   child=LIN AF:bought GEN=fish
   ‘the child who bought (the) fish’

As shown above, only the S or Uₜ can be a pivot for relativization in Tagalog. It follows
that relativization in Tagalog shows syntactic ergativity (see alsoBlake 1988).

2.3 Some important constructions

I will demonstrate in the following sections that different focus structures are
expressed by different constructions, which are overviewed in advance here.

[1] Cleft construction. A cleft construction in Tagalog is an intransitive clause
where its single core argument is a headless relative clause, and its nucleus is a noun
phrase coreferential with the gap in the headless relative clause. See example (6) and
Figure 3.

(6) Si=Boyet, ang=[p<um>atay [S] kay=Juan].
   ABS=Boyet ABS=AF:killed DAT=Juan
   ‘The one who killed Juan is Boyet.’

---

4 As one of the reviewers mentioned, there are some exceptions to this generalization (Ceña
1979). See Reid & Liao (2004a, b) for discussion of relativization in Philippine languages and
also see Liao (1999-2000) for more discussion of relativization.
This cleft analysis is borne out by the fact that a clause which can appear only as a relative clause can also appear as the core argument of a cleft construction. Let us consider (7), which is ungrammatical: the clause *pumatay kay Juan* cannot be used in non-relative environments (Schachter & Otanes 1972:299ff., Schachter 1976:517, Schachter 1977:303-304).

(7) *P<um>atay si=Boyet kay=Juan.
AF:killed ABS=Boyet DAT=Juan
Intended for ‘Boyet killed Juan.’

Compare this example with (8), where the same clause is employed as a relative clause modifying a noun phrase, yielding a grammatical sentence.

(8) S<in>ampal=ko ang=lalaki=ng [p<um>atay kay=Juan].
PF:spanked=1SG.ERG ABS=man=LIN AF:killed DAT=Juan
‘I spanked the man who killed Juan.’

Now let us look at (6) again, which is grammatical, even though *p<um>atay kay Juan* is employed. This means that relativization is involved in example (6). That is, example (6) is a cleft construction.

Other pieces of evidence for this analysis include the interpretation of a depictive secondary predicate (Nagaya 2004) and the position of an enclitic (Kroeger 1993, 1998).
[2] Fronting construction. In a fronting construction, an adjunct is fronted to the precore slot like WH-questions in English (cf. ‘emphatic inversion’ of Schachter & Otanes 1972:496ff.).

(9) \[ PrCS Ngayon]=ko=lang na-basa ang=e-mail=mo.  
\[ now=1SG.ERG=just PF:read ABS=e-mail=your  
‘I read your e-mail just now.’

It is significant that in this construction the enclitics are attached to the precore slot rather than the nucleus. See Figure 4:

```
<table>
<thead>
<tr>
<th>SENTENCE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLAUSE</td>
</tr>
<tr>
<td>PrCS</td>
</tr>
<tr>
<td>CORE</td>
</tr>
<tr>
<td>ARG</td>
</tr>
<tr>
<td>NUC</td>
</tr>
<tr>
<td>ARG</td>
</tr>
<tr>
<td>PRED</td>
</tr>
<tr>
<td>ADV PRO VNP</td>
</tr>
<tr>
<td>Ngayon=ko=lang na-basa ang=e-mail=mo.</td>
</tr>
</tbody>
</table>
```

**Figure 4:** Fronting construction

[3] Canonical construction. All the other constructions discussed in this paper are called canonical constructions. In a canonical construction, the ‘subject-predicate relationship’ is not clefted and the precore slot is null. An example of this type of construction has been given in (2).

3. **Focus structures and their realizations**

Now let us look at how information structure and morphosyntax interact with each other in Tagalog, relying on Lambrecht’s theory of focus structure adopted in RRG. The most important distinction in this theory is between narrow focus and broad focus. In narrow focus the focus falls only on a single constituent, while in broad focus the focus extends over more than one constituent. Broad focus is composed of two subtypes: predicate focus and sentence focus. The former is equivalent to the traditionally recognized ‘topic-comment’ organization of information in a clause. In the latter, the focus domain encompasses the entire clause and there is no topic in it.

In addition to those mentioned above, two types of narrow focus are recognized in this paper for the sake of convenience: argument narrow focus and adjunct narrow focus.
The whole taxonomy of these distinctions is summarized in Table 1.

**Table 1: Taxonomy of focus structure**

<table>
<thead>
<tr>
<th>Focus structure</th>
<th>Focus domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>Broad focus</td>
<td></td>
</tr>
<tr>
<td>Predicate focus</td>
<td>Predicate (+ non-predicating element(s))</td>
</tr>
<tr>
<td>Sentence focus</td>
<td>Entire clause</td>
</tr>
<tr>
<td>Narrow focus</td>
<td></td>
</tr>
<tr>
<td>Argument narrow focus</td>
<td>Single argument</td>
</tr>
<tr>
<td>Adjunct narrow focus</td>
<td>Single adjunct</td>
</tr>
</tbody>
</table>

In this section, it is argued that in Tagalog different types of focus structure are expressed by different constructions. I will discuss each of these types of focus structure individually: predicate focus in §3.1, sentence focus in §3.2, argument narrow focus in §3.3, and adjunct narrow focus in §3.4.

### 3.1 Predicate focus

The construction used to encode predicate focus is a canonical construction. See the answer in example (10).

(10) Q: Kumusta ang=kotse=mo? how ABS=car=your
     ‘How’s your car?’
A: Na-sira (ang=kotse=ko). [canonical]
     PF:broke down ABS=car=my
     ‘(My car) broke down.’

Note that in Tagalog ‘topic-comment’ organization is not reflected in constituent order. It is realized as ‘comment-topic’ order in this head-initial language as in (10A), although ‘topic-comment’ order is the most common order cross-linguistically (see Van Valin & LaPolla 1997:421).

### 3.2 Sentence focus

The canonical construction is also employed to express sentence focus, in which the actual focus domain is the entire clause. See the answer in example (11).

(11) Q: Ano ang=[nang-yari]?
     what ABS=AF:happened
     ‘What happened?’
     (LIT. ‘The thing that happened was what?’)
A: Na-sira ang=kotse=ko. [canonical]  
PF:broke down Abs=car=my  
‘My car broke down.’

Thus, both predicate focus and sentence focus are represented by the same construction. Namely, broad focus structure is marked by the canonical construction. This is illustrated in Figures 5 and 6. (10A) and (11A) are identical in the constituent projection but different in the focus structure projection.5

---

5 Note that Tagalog employs a presentational construction as well as a canonical construction to mark sentence focus structure (Nagaya 2005). The presentational construction is a subtype of existential sentence, where a newly-introduced participant is realized as a noun phrase or a headless relative clause.

May [naka-tayo sa=labas]. [presentational]  
exist AF:standing DAT=outside  
‘There is/was someone standing outside.’

Canonical and presentational constructions can describe much the same situation. But they are different in that the speaker assumes that the hearer already knows about a newly-introduced participant at the time of utterance in a canonical construction, but he/she does not in a presentational construction. That is why a proper noun is not allowed in a presentational construction. See Nagaya (2005) for details.

It is worth pointing out that, since a presentational construction is formed by relativization and relativization shows syntactic ergativity in Tagalog, what is introduced by this construction is restricted to S and UT, which are universally favored for new information (Du Bois 1987).
3.3 Argument narrow focus

The cleft construction serves to encode argument narrow focus. See the answer in example (12). This construction is obligatory in this focus structure.

(12) Q: Na-rinig=ko=na na-sira ang=motor=mo. 
   PF:heard=1SG.ERG=LIN PF:broke down ABS=motorcycle=your
   ‘I heard that your motorcycle broke down.’
A: Hindi. Ang=kotse=ko ang=[na-sira]. [cleft]
   No ABS=car=my ABS=PF:broke down
   ‘No. My car broke down.’
   (lit. ‘No. What broke down is my car.’)

3.4 Adjunct narrow focus

To express adjunct narrow focus, the fronting construction is employed. See the answer in example (13). This construction is obligatory in this focus structure.

(13) Q: Kailan=ba na-sira ang=kotse=mo?
   when=Q PF:broke down ABS=car=your
   ‘When did your car break down?’
A: [PrCS Kahapon] na-sira (iyon). [fronting]
   yesterday PF:broke down (it.ABS)
   ‘(It) broke down yesterday.’

The analyses of (12A) and (13A) are given in Figures 7 and 8. (12A) and (13A) are different in the constituent projection, but the same in the focus structure projection.
3.5 Summary

The discussion in this section is summarized in Table 2, where the clause-initial constituent is in boldface, and “CL” indicates an enclitic.

<table>
<thead>
<tr>
<th>Focus structure</th>
<th>Constructions</th>
<th>Constituent Order</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predicate focus</td>
<td>Canonical</td>
<td>V (CL) NP …</td>
</tr>
<tr>
<td>Sentence focus</td>
<td>Canonical</td>
<td>V (CL) NP …</td>
</tr>
<tr>
<td>Argument narrow focus</td>
<td>Cleft</td>
<td>NP (CL) ang=[V …]</td>
</tr>
<tr>
<td>Adjunct narrow focus</td>
<td>Fronting</td>
<td>ADV (CL) V …</td>
</tr>
</tbody>
</table>

An important consequence of this observation is that the clause-initial element in boldface is included in the focus in all the focus structures. In predicate focus and sentence focus, the nucleus houses a focal predicate in a canonical construction. In narrow focus, the nucleus accommodates a focal argument in a cleft construction, and the precore slot contains a focal adjunct in a fronting construction. I will give further evidence for this claim in the next section.
4. Clause-initial position as the obligatory focus domain

In this section, I will try to demonstrate that the clause-initial position in Tagalog is always in focus, by examining various types of constructions as a litmus test for the focus: exhaustive identification (§4.1), universal quantifier (§4.2), WH-question (§4.3), yes-no question (§4.4), correction (§4.5), and negation (§4.6). Each construction has a specific portion which should be in focus. For example, a WH-word has to be in focus in WH-questions. I will show that in any case the clause-initial position is (a part of) the focus domain each construction specifies.

4.1 Exhaustive identification

It is well known that an exhaustively identified element is in focus (cf. Kiss 1998). Exhaustive identification in Tagalog is marked by the enclitic \textit{lang}/lamang. The scope of this enclitic includes at least the clause-initial position. See the examples below.

(14) Na-tulog\=lang\=siya sa=bahay. [canonical]  
\text{PF:slept=only=3SG.ABS D AT=house}  
‘He/she only slept in the house.’

(15) Siya=lang ang=[na-tulog sa=bahay]. [cleft]  
3SG.ABS=only ABS=PF:slept D AT=house  
‘Only \textit{he/she} slept in the house.’  
(LIT. ‘The one who slept in the house was only him/her.’)

(16) [PrCS Sa=bahay]=lang=siya na-tulog. [fronting]  
\text{DAT=house=only=3SG.ABS PF:slept}  
‘He/she slept only \textit{in the house}.’

Note that the scope of exhaustive identification can be over the entire clause as well as the predicate in (14). In either case, the clause-initial position is included in the focus domain of exhaustive identification.

Because of this property, the enclitic \textit{lang} cannot exclusively identify an argument or adjunct outside the clause-initial position. See examples (17) and (18), whose possible interpretation is that the clause-initial position is included in the focus domain.

(17) I-text=mo=lang si=Daphne. [canonical]  
\text{CF:text=2SG.ERG=only ABS=Daphne}  
*‘Text only \textit{Daphne}.’  
ok ‘Just text Daphne.’
(18) **Nag-meryenda**=lang=siya sa=bahay. [canonical]
AF:had a snack=only=3SG.ABS DAT=house
*‘He/she had a snack only in the house.’
ok ‘He/she only had a snack in the house.’

For a single non-predicating element to be the focus of exhaustive identification, it is necessary to employ a cleft or a fronting construction like (19) and (20).

(19) **Si=Daphne**=lang ang=[i-text=mo]. [cleft]
ABS=Daphne=only ABS=CF:text=2SG.ERG
‘Text only **Daphne.**’
(LIT. ‘The one you (should) text is only Daphne.’)

(20) **[PrCS Sa=bahay]**=lang=siya nag-meryenda. [fronting]
DAT=house=only=3SG.ABS AF:had a snack
‘He/she had a snack only in the house.’

As shown above, the clause-initial position is always in the focus domain of exhaustive identification.

4.2 Universal quantifiers

The universal quantifier is also sensitive to information structure. There is a cross-linguistic tendency for the universal quantifier not to appear in the focus domain of narrow focus structure (Kiss 1998).

This also obtains in Tagalog. The universal quantifiers **lahat** ‘all’ and **kahit ano** ‘anything’ cannot occur in the nucleus of a cleft construction, as in (21) and (22).

(21) *Lahat ng=pagkain ang=[k<in>ain=ko]. [cleft]
all GEN=food ABS=PF:ate=1SG.ERG
(22) *Kahit ano ang=[k<in>ain=ko]. [cleft]
anything ABS=PF:ate=1SG.ERG

In contrast, these universal quantifiers can appear in a core argument position of a canonical construction.

(23) **K<in>ain=ko ang=lahat ng=pagkain.** [canonical]
P:ate=1SG.ERG ABS=all GEN=food
‘I ate all the food.’
(24) **K<in>ain=ko kahit ano.** [canonical]
P:ate=1SG.ERG anything
‘I ate anything.’
Like this, the universal quantifiers are excluded from the clause-initial position in Tagalog. This fact supports the claim that the clause-initial position is always in focus.

### 4.3 WH-questions

WH-questions are a typical example of narrow focus structure: a WH-word is narrowly focused in this type of sentence. In Tagalog, when an argument is questioned, a cleft construction is employed. See (26), which is a WH-question corresponding to (25).

(25) K<um>ain si=Jonalyn ng=hamburger kahapon. [canonical]
    AF:ate A BS=Jonalyn GEN=hamburger yesterday
    ‘Jonalyn ate hamburger yesterday.’

(26) Sino=ba ang=[k<um>ain ng=hamburger kahapon]? [cleft]
    who=Q A BS=AF:ate GEN=hamburger yesterday
    ‘Who ate hamburger yesterday?’
    (LIT. ‘Who was the one who ate hamburger yesterday?’)

A fronting construction is used when an adjunct is questioned, as in (27).

(27) [PrCS Kailan]=ba k<um>ain si=Jonalyn ng=hamburger? [fronting]
    when=Q AF:ate A BS=Jonalyn GEN=hamburger
    ‘When did Jonalyn eat hamburger?’

Thus, WH-questions are just a type of narrow focus construction in Tagalog. This is also the case in other Austronesian languages like Malagasy (Potsdam 2004).

It is noteworthy that Tagalog has what I call WH-verbs, whose root is a WH-word (Schachter & Otanes 1972:509).

(28) Nag-ano=sila dito?
    AF:did what=3PL.ABS here
    ‘What did they do here?’

(29) <In>ano=mo ang=damit=ko?
    PF:did what=2SG.ERG A BS=clothes=my
    ‘What did you do to my clothes?’

Since they obviously contain a focal verbal predicate in the nucleus, these WH-questions with WH-verbs are categorized as predicate focus in spite of being WH-questions. Nonetheless, these are the same as ordinary WH-questions I discussed here in the sense that the clause-initial element is exclusively the focus.
4.4 Yes-no questions

Yes-no questions are formed by adding the interrogative enclitic *ba* or just by intonation. Example (30) is a yes-no question sentence expressing predicate focus or sentence focus.

(30) Pu-punta=ba si=Melanie dito? [canonical]
    AF:will come=Q ABS=Melanie here
    ‘Will Melanie come here?’

Now let us turn to narrow focus. When a single argument is the focus of a yes-no question, a cleft is employed as in (31), which is asking whether or not it is Melanie that will come to the speaker’s side.

(31) Si=Melanie=ba ang=[pu-punta dito]? [cleft]
    ABS=Melanie=Q ABS=AF:will come here
    ‘Will Melanie come here?’
    (LIT. ‘Is the one who will come here Melanie?’)

In the case of a narrowly focused adjunct, it is fronted as in (32), which is asking whether or not it is to the speaker’s side that Melanie will come.

(32) [PCS Dito]=ba pu-punta si=Melanie? [fronting]
    here=Q AF:will come ABS=Melanie
    ‘Will Melanie come here?’

In this way, in Tagalog yes-no questions, the clause-initial position is always included in the focus of question.

4.5 Correction

Correction of someone’s belief is also one of the typical contexts of narrow focus structure. Let us look at the conversation below:

(33) (Person A joked about Ronald, who was going to celebrate his birthday soon. Then, Person B, Ronald’s wife, responded to Person A.)
    A: i. Kawawa=naman si=Ronald. [canonical]
       pitiful=indeed ABS=Ronald
       ‘Ronald is really pitiful.’
ii. Kasi, wala=ako sa=kaarawan=niya.  
because not exist=1SG.ABS DAT=birthday=his  
‘This is because I will be absent on his birthday.’

B: i. Ikaw ang=[kawawa]. [cleft]  
2SG.ABS ABS=pitiful  
‘You (> not Ronald) are pitiful.’  
(LIT. ‘The one who is pitiful is you.’)

ii. Kasi, hindi=ka maka-ka-kain ng=handa=ko.  
because not=2SG.ABS AF:can eat GEN=prepared food=my  
‘This is because you will not be able to eat the food I will prepare.’

Compare the utterance of Person A (33Ai) (predicate focus) with that of Person B (33Bi) (argument narrow focus). Person B is trying to correct Person A’s belief by employing the cleft construction.

The fronting construction as well as the cleft is used for correction.

(34) (The hearer came back to the meeting after using his cell phone outside the room. He sat down in a seat, but that seat was not the seat where he was seated before. So, the speaker told the hearer where he was seated before.)  
[PrCS Doon]=ka naka-upo. [fronting]  
there=2SG.ABS AF:were seated  
‘You were seated there (> not here).’

On the other hand, when the adjunct is in situ, the sentence is broad focus: the implication observed in (34) is not found in (35).

(35) Naka-upo=ka doon. [canonical]  
AF:were seated=2SG.ABS there  
‘You were seated there.’

Here is a pair of examples with a temporal adjunct. (36) is simply proposing that the speaker and the hearer meet tomorrow. But (37) is insisting that the speaker and the hearer meet tomorrow, not the day for which the meeting was previously scheduled.

---

6 The particle *kasi* is analyzed as a sentence-level clause-linkage marker (Van Valin & LaPolla 1997:476), and is not counted as a clause-initial element. See also note 9.
(36) Mag-kita=tayo bukas. [canonical]
AF:meet=1PL.INCL.ABS tomorrow
‘Let’s meet tomorrow.’

(37) [PRCS Bukas]=na=lang=tayo mag-kita. [fronting]
tomorrow=already=just=1PL.INCL.ABS AF:meet
‘Let’s meet tomorrow (> not another day).’

In every example, the clause-initial constituent is in the focus of correction.

4.6 Negation

In Tagalog, different scopes of negation are expressed by different constructions (see also Schachter & Otanes 1972:499-500, Ramos & Ceña 1990:172-173). In any case, the clause-initial position is included in the focus of negation.

(38) Hindi s<in>ipa ni=Mark si=Arius. [canonical]
not PF:kicked E RG=Mark ABS=Arius
‘Mark didn’t kick Arius (> Mark did not do anything to Arius, or Mark did something but kicking).’

(39) Hindi si=Mark ang=[s<um>ipa kay=Arius]. [cleft]
not ABS=Mark ABS=AF:kicked DAT=Arius
‘Mark didn’t kick Arius (> but someone else kicked Arius).’
(LIT. ‘The one who kicked Arius was not Mark.’)

(40) Hindi kahapon s<in>ipa ni=Mark si=Arius. [fronting]
not yesterday PF:kicked E RG=Mark ABS=Arius
‘Mark didn’t kick Arius yesterday (> but at some other time).’

Internal negation or narrow scope negation, where the scope of negation is only part of the clause (Van Valin & LaPolla 1997:45-46), requires a cleft or a fronting construction, as in (39) and (40). The position of the negative operator hindi in the layered structure of the clause needs more study. Note that it is possible that the negation has clausal scope in (38).

4.7 Summary

As shown above, the focus of exhaustive identification, WH-question, yes-no question, correction and negation includes the clause-initial position, i.e. the nucleus or precore focus slot, in any focus structure. In this sense, the clause-initial position is the obligatory focus domain in Tagalog. In broad focus, the predicate in a clause is in focus.
and occupies the clause-initial nucleus, while in narrow focus one of the non-predicating elements is focused and appears in the clause-initial position by means of a cleft or fronting construction.7

5. Left-detached position and sentence-clause distinction

The discussion so far is confined to the clause-internal correlation between clause structure and information structure. In this section, let us expand the range of the discussion to include the entire sentence and consider aspects of syntax-pragmatics interface at the level of sentence.

The center of contention in this regard is the left-detached position, which is an optional location for sentence-initial elements. In this section, first, I will present a syntactic and pragmatic description of this position briefly in §5.1. Then, I will demonstrate how this position behaves differently from the clause-initial position, with regard to syntactic behaviors in §5.2. Through these considerations, the validity of the traditional distinction between sentence and clause will be reevaluated in Tagalog.

5.1 Left-detached position

The left-detached position in Tagalog, which has already been described in Schachter & Otanes (1972:485ff.) and Foley & Van Valin (1984:126ff.), houses a topical/presupposed element. Let us examine examples (41) and (42).

(41) [sentence [LDP Si=May], [clause [PCS kailan]=ba=siya ba-balik dito]? 
ABS=May when=Q=3S G.ABS AF:will return here
‘As for May, when will she come here?’

(42) [sentence [LDP Bukas], [clause ikaw=na=lang ang=[bi-bili ng=ulam]].
 tomorrow 2SG.ABS=already=just ABS=AF:will buy GEN=food
‘Tomorrow, it is you who will buy food.’

7 In this connection, the flexibility of focus structure is worth mentioning. Van Valin (1999) has proposed the typology of rigid and flexible focus structure languages. The former, like Italian and French, has a restriction on the potential focus domain, whereas the latter does not, like English and Russian. According to Van Valin’s typology, Tagalog is categorized as a language of flexible focus structure, since the potential focus domain is not restricted to a subpart of the main clause. However, Tagalog is a language of rigid focus structure with regard to the obligatory focus domain: it has a restriction that the clause-initial position (= obligatory focus domain) must be included in the actual focus domain. This suggests that finer typology of rigid and flexible focus structure is possible.
In (41), the absolutive noun phrase *si May* is preposed to the left-detached position, and functions as a topic. The remaining clause is a Wh-question, in which the Wh-word *kailan* ‘when’ is fronted to the precore slot. In (42), the adverb *bukas* ‘tomorrow’ is dislocated in the left-detached position and sets the temporal framework for the remaining cleft clause. These analyses are represented in Figures 9 and 10.

**Figure 9:** Left-detached position (41)  
**Figure 10:** Left-detached position (42)

The left-detached position is usually set off from the clause by a pause\(^8\) (Van Valin & LaPolla 1997:36, Van Valin 2005:6). The pause is expressed by ‘,’ in this paper.

(43) (The speaker is reporting about the breakfast she has just eaten.)

\begin{itemize}
\item a. *Katatapos=lang=nami=ng mag-breakfast.*
\item *finished=just=1PL.EXCL.GEN=LIN AF:have breakfast*
\item ‘We have just finished our breakfast.’
\end{itemize}

\(^8\) The left-detached position is sometimes followed by the particle *ay*, coupled with a pause. According to Schachter & Otanes (1972:485), *ay*-construction is “characteristic of formal style, and is more common in writing, lectures, sermons, etc., than it is in ordinary conversation”.

362
b. \[LDP\] Ang=ulam=namin, [\textit{clause} longanisa, tinapa, corn beef, at atsara].
\hspace{1cm} ABS=dish=our longanisa tinapa corn beef and atsara
\hspace{1cm} ‘Our dishes are longanisa, tinapa, corn beef, and atsara.’

This position is a preferred site for a contrastive topic as shown in (44).

(44) (On the phone)
\begin{itemize}
  \item Nasaan=ka=na?
  \hspace{1cm} where=2SG.ABS=now ‘Where are you now?’
  \item LDP Kami, [\textit{clause} nasa labas=pa].
  \hspace{1cm} 1PL.EXCL.ABS be at outside=still ‘As for us, (we) are still outside.’
\end{itemize}

The examples presented so far all involve detached noun phrases. But a setting element as well can be located in the left-detached position to introduce the temporal or spatial framework to a discourse.

(45) [\textit{LDP} Kahapon], [\textit{clause} nag-lunch si=Daphne sa=KFC].
\hspace{1cm} yesterday AF:had lunch ABS=Daphne DAT=KFC
\hspace{1cm} ‘Yesterday, Daphne had lunch at KFC.’

(46) [\textit{LDP} Sa=palengke], [\textit{clause} b<um>ili ng=isda ang=lalake].
\hspace{1cm} DAT=market AF:bought GEN=fish ABS=man
\hspace{1cm} ‘In the market, the man bought some fish.’ (Foley and Van Valin 1984:128)

The left-detached position can also serve as a landing site for possessor raising.

(47) [\textit{LDP} Si=Juan], [\textit{clause} na-matay ang=tatay].
\hspace{1cm} ABS=Juan PF:died ABS=father
\hspace{1cm} ‘Juan, (his) father died.’

5.2 Distinction between sentence and clause

It should be clear by now that the left-detached position is pragmatically different from the clause-initial position. Then the logical question to ask is how this position can be distinguished from the clause-initial position. This issue is addressed in this section by employing seven tests.
5.2.1 Enclitic placement

As mentioned in §2.1 (see also Kroeger 1993, 1998), the enclitics in a clause follow the clause-initial element, as in (48).

(48) [LDP Bukas], [clause [PrCS saan]=ka=ba pu-punta]? tomorrow where=2SG.ABS=Q AF:will go

‘Tomorrow, where will you go?’

In contrast, the enclitics cannot accompany the left-detached element, as in (49).

(49) *[LDP Bukas]=ka=ba, [clause [PrCS saan] pu-punta]?

There is one exception to this generalization: bakit ‘why’. This WH-word seems to occupy the precore slot in WH-questions, but is not necessarily accompanied by an enclitic.9 The reason is unknown.

5.2.2 Uniqueness condition

In Tagalog, only one precore slot is available in each clause. I refer to this as the uniqueness condition (cf. Rizzi 1997). Example (50) is not grammatical; it includes more than one precore slot in the clause.

(50) *[clause [PrCS Ano=ng oras] [PrCS dito]=ka pu-punta]? what=L IN hour here=2SG.ABS AF:will come

On the other hand, the uniqueness condition does not hold in the case of the left-detached position. More than one left-detached position can occur in a sentence.

(51) [LDP Sa=palagay=ko], [LDP si=Paul], [clause patay=na]. DAT=opinion=my ABS=Paul dead=already

‘In my opinion, as for Paul, (he) was already dead.’

---

9 One of the reviewers suggested that kasi in (33) should be added to this statement. But I analyze this particle as a sentence-level clause-linkage marker, and do not regard it as an exception. See also note 6.
5.2.3 Semantic relatedness

Since it is inside a clause, the clause-initial element must be relevant to the semantics of the clause. A manner adverb and a depictive secondary predicate (Nagaya 2004) can appear in this position.

(52) \[ \text{[clause [PrCS Mabilis]=na=niya k\text{\textless}in\text{\textgreater}ain ang=pansit].} \]
\[ \text{quick=already=3SG.ERG PF:ate ABS=noodle} \]
‘He ate noodles quickly.’

(53) \[ \text{[clause [PrCS Nakahubad]=mo=ba k\text{\textless}in\text{\textgreater}ain ang=isda]?} \]
\[ \text{naked=2SG.ERG=Q PF:ate ABS=fish} \]
‘Did you eat the fish naked?’ (The hearer was naked.)

On the other hand, the left-detached element need not be relevant to the semantics of the clause.

(54) \[ \text{[LDP Ang=nais=ko], [clause malusog ang=mama=ko].} \]
\[ \text{ABS=wish=my healthy ABS=mother=my} \]
‘My wish is that my mother would be healthy.’

(55) \[ \text{[LDP Ang=[g\text{\textless}in\text{\textgreater}awa=ko]], [clause t\text{\textless}in\text{\textgreater}ulung-an=ko si=Mutya].} \]
\[ \text{ABS=PF:did=1SG.ERG LF:helped=1SG.ERG ABS=Mutya} \]
‘What I did was that I helped Mutya.’

A modal adverb and a sentential adverb also appear here.

(56) \[ \text{[LDP Siguro], [clause b\text{\textless}um\text{\textgreater}agsak si=Gaga sa=pagsusulit].} \]
\[ \text{probably AF:failed ABS=Gaga DAT=examination} \]
‘Probably, Gaga failed the exam.’

(57) \[ \text{[LDP Sa=kasamaang-palad], [clause b\text{\textless}um\text{\textgreater}agsak si=Gaga sa=pagsusulit].} \]
\[ \text{unfortunately AF:failed ABS=Gaga DAT=examination} \]
‘Unfortunately, Gaga failed the exam.’

5.2.4 Pause

The pause or intonational break cannot follow the clause-initial element, but it can accompany the left-detached element(s). Let us look at examples (41) and (42), repeated here as (58) and (59).
(58) \[ \text{LDP Si=May}, \ [\text{clause [PrCS kailan](*)=ba(*,)=siya(*,)} \ \text{ba-balik dito}? \]
\[
\text{ABS=May when=Q=3SG.ABS AF:will return here} \\
\text{‘As for May, when will she come here?’} \\
\]

(59) \[ \text{LDP Bukas}, \ [\text{clause ikaw(*,)=na(*,)=lang(*,)} \ \text{ang=[bi-bili ng=ulam]}. \]
\[
\text{tomorrow 2SG.ABS=already=just ABS=AF:will buy GEN=food} \\
\text{‘Tomorrow, it is you who will buy food.’} \\
\]

There is no pause following the nucleus and the precore slot.

5.2.5 Resumptive pronoun

The left-detached noun phrase can leave a resumptive pronoun coreferential with it in the following clause.

(60) \[ \text{LDP Si=Ian}, \ [\text{clause i-b\<in\>ili=mo(=siya)}]? \]
\[
\text{ABS=Ian CF:bought=2SG.ERG(=3SG.ABS)} \\
\text{‘As for Ian, did you buy (something) (for him)?’} \\
\]

(61) \[ \text{LDP Si=Flor}, \ [\text{clause na-kita(=niya)} \ \text{si=Weng}]. \]
\[
\text{ABS=Flor PF:met(=3SG.ERG) ABS=Weng} \\
\text{‘As for Flor, (she) met Weng.’} \\
\]

Van Valin & LaPolla (1997:36) say that “if the NP in it [= the left-detached position] functions as a semantic argument in the following clause, there must be a pronoun in the clause which refers to it”. But since Tagalog allows zero anaphora (cf. Nagaya 2006a, b), the resumptive pronoun is optional.

Such a resumptive pronoun is not available in a cleft construction.

(62) \*Si=Ian \ ang=[i-b\<in\>ili=mo=siya] \\
\text{ABS=Ian ABS=CF:bought=2SG.ERG=3SG.ABS}

5.2.6 Universal quantifier

In §4.2, I have pointed out that the universal quantifiers (lahat ‘all’ and kahit ano ‘anything’) cannot appear in the clause-initial position. In contrast, they are available in the left-detached position.

(63) \[ \text{LDP Lahat ng=pagkain}, \ [\text{clause k\<in\>ain=ko}]. \]
\[
\text{all GEN=food PF:ate=1SG.ERG} \\
\text{‘All the food, I ate (it).’} \\
\]
(64) [LDP Kahit ano], b<in>ili=niya.
    anything PF:bought=3SG.ERG
    ‘Anything, she bought (it).’

5.2.7 Symmetric property

The left-detached position and the clause-initial position differ in terms of symmetric property, too. In Tagalog, the detached position has a symmetric property regarding the position in which it appears: whether it appears to the left of the clause or to the right does not affect the pragmatic status of the detached phrase. Tagalog has a right-detached position, as well, which is also for a presupposed element.

(65) [clause <Um>upo=ka=muna dito], [RDP ang=sabi=ko].
    AF:sit=2SG.ABS=first here ABS=what is said=my
    ‘What I said is that you sit down here.’
(66) [clause Mahal-in=mo ang=kapwa=mo], [RDP ang=turo ni=Hesus].
    PF:love=2SG.ERG ABS=neighbor=your ABS=teachings GEN=Jesus
    ‘Love your neighbor, according to the teachings of Jesus.’

On the other hand, the clause-initial element does not have a symmetric property. When it appears in non-clause-initial positions, it may or may not be in focus.\(^{10}\)

5.2.8 Where do these differences come from?

As has been demonstrated above, the left-detached position and the clause-initial position (i.e. the nucleus or precore slot) behave differently at least with regard to these seven phenomena. See Table 3.

<table>
<thead>
<tr>
<th></th>
<th>Left-detached position</th>
<th>Clause-initial position</th>
</tr>
</thead>
<tbody>
<tr>
<td>Enclitic placement</td>
<td>*</td>
<td>√</td>
</tr>
<tr>
<td>Uniqueness condition</td>
<td>*</td>
<td>√</td>
</tr>
<tr>
<td>Semantic relatedness</td>
<td>*</td>
<td>√</td>
</tr>
<tr>
<td>Pause</td>
<td>√</td>
<td>*</td>
</tr>
<tr>
<td>Resumptive pronoun</td>
<td>√</td>
<td>*</td>
</tr>
<tr>
<td>Universal quantifier</td>
<td>√</td>
<td>*</td>
</tr>
<tr>
<td>Symmetric property</td>
<td>√</td>
<td>*</td>
</tr>
</tbody>
</table>

\(^{10}\) I have not been able to find a postcore slot in Tagalog so far. See Van Valin & LaPolla (1997: 37) and Van Valin (2005:5-6) for the postcore slot.
These results are not accidental but rather natural consequences of the characterization of the left-detached position and the clause-initial position in RRG: the left-detached position is clause-external, while the clause-initial position is clause-internal.

Let us take enclitic placement as an example. The pronominal enclitic, which is an argument of a predicate, can be located in the position immediately following the clause-initial position, because the position is still inside the clause. But when it is put after the left-detached position, it is outside the clause. This situation is not acceptable, because the semantic argument of the predicate must be included in the clause.

This explanation is applicable to the other phenomena: the pause reflects the clause boundary; the clause-initial element does not leave a resumptive pronoun because it is still inside the clause; since the left-detached element is out of the clause, it is free from the uniqueness condition and the semantic relatedness condition of the clause. Thus, the traditional sentence-clause distinction adopted in RRG can provide a satisfactory explanation for different behaviors of the two positions.

6. Conclusion

In this paper, I have discussed the interaction between syntax and pragmatics in Tagalog, and have claimed the following. At the clause-internal level, different constructions are employed to represent different types of focus structure: a canonical construction for broad focus, a cleft construction for argument narrow focus, and a fronting construction for adjunct narrow focus. In any case the clause-initial position is in focus and thus can be considered as the obligatory focus domain. At the clause-external level, a topical/presupposed element comes optionally in the left-detached position. But this position can be clearly distinguished from the clause-initial position with regard to syntax and pragmatics. To put it differently, the traditional contrast between sentence and clause is verified in this language. Thus, in Tagalog, syntax and pragmatics interact differently inside the clause than they do outside the clause: the focus obligatorily falls on the domain including the clause-initial position, and a topical/presupposed element is optionally preposed to the sentence-initial left-detached position.

Concerning such an interconnection between constituent order and information structure, Gundel (1988:229) proposes ‘two independent, and sometimes conflicting principles’:

(67) First Things First Principle
    Provide the most important information first.

(68) Given Before New Principle
    State what is given before what is new in relation to it.
The discussion in this paper reveals that these two principles are both realized in Tagalog but in different layers. Clause-internally, First Things First Principle (67) is the determining factor: the actual focus domain always includes the clause-initial position. Clause-externally, in contrast, Given Before New Principle (68) should be taken into account: a topical/presupposed element optionally occupies the sentence-initial position called the left-detached position. In light of my analysis, these two principles do not conflict but function in distinct layers of clause structure.

Thus, this paper presents all these findings and claims in an integrated way within the framework of RRG. In particular, it provides strong support for the validity of its theories of clause structure and information structure.

References


Law, 244-258. ZAS Papers in Linguistics 34. Berlin: ZAS.

[Received 2 January 2006; revised 1 June 2006; accepted 1 August 2006]
塔加洛語資訊結構及句法成分的排列方式

長屋尚典
東京大學

本文主要目的為探討句法及語用在塔加洛語子句內外的互動。在子句內，不同結構可以表達不同的焦點：典型結構可反應述語或整句焦點；分裂結構可反應論元的狹窄焦點；前移結構則可以反應修飾語的狹窄焦點。在這三種結構裡，實際焦點範疇涵蓋了子句最前面的位置。然而在子句外，主題或預想成分可以選擇性的前移到句子最前方所謂的「向左分離」位置，此與子句的前置位置有所不同。因此，傳統中所認爲的句子和子句之間的不同在這個語言裡有非常明確的區分。

關鍵詞：塔加洛語，資訊結構，焦點結構，句法成分的排列方式，分裂句，前移