



Information Structure and Grammaticalization in Tagalog

Author(s): Latrouite, Anja

Source: Proceedings of the International Workshop on Information Structure of Austronesian Languages, 10 April 2014, pp.159-170.

Published by: ILCAA, TUFS

Permanent URL: <http://hdl.handle.net/10108/75997>

The Prometheus-Academic Collections are a repository of academic research. In them are found the research and educational achievements of the Tokyo University of Foreign Studies and also historical materials which have been preserved. The University shares such academic achievements on the Internet to ensure an open university. Copyright is retained by each author, academic society, association, publisher, and/or other rights owners. The whole or parts of the materials can be used (referred to, reproduced, printed), with copyright acknowledged, for academic and personal use only.

Information Structure and Grammaticalization in Tagalog

Heinrich-Heine Universität Düsseldorf

Anja Latrouite

latrouite@phil.uni-duesseldorf.de

0. Prelude

Information-structural notions play an important role in the description of Austronesian languages like Tagalog. Their word order is not free, but in large parts dictated by information-structure notions like topic and focus (Kaufman 2006), and the choice of referential expressions is guided by topichood (Nagaya 2007). Moreover, the grammatical voice-system identifies a subject which is considered a grammaticalized topic (Shibatani 1991) as it still exhibits many topic-like features, e.g. (i) it is not limited to direct semantic arguments of the verb, (ii) it is per default understood as definite or at least specific, and (iii) it occurs and is banned in many of the same constructions as Japanese topics (Katagiri 2006).

The goal of this paper is to distangle the different kinds of *topic* vs. *foci* related to different levels, give an overview of marking strategies, explore the interaction between discourse-/sentence-/clause-level IS notions and discuss Undergoer-orientedness with respect to the choice of the grammatical topic.

1. Information Structure: Terminology & Phenomena

Studies in Information Structure deal with different ways of establishing and enriching the *Common Ground* shared by the speaker and the addressee. We often find a distinction between works concerned with **pragmatic topics/foci** and **semantic topics/foci**. The former is mainly about **discourse coherence**, the second is about the **sentential content**, i.e. about constructing the predication structure and the scopal relations within the sentence, as summed up in table 1. According to Krifka and Musan (2012) often ‘one and the same device can be used for packaging (information, AL) as well as for constructing the content’ (of a sentence, AL).

	<u>Pragmatics</u>	<u>Semantics</u>
TOPIC	OLD /GIVEN FILE-CARD UNDER WHICH NEW INFORMATION IS STORED	PREDICATED OVER WIDE SCOPE FRAME-SETTER
FOCUS	(relatively) NEW PRAGMATICALLY NON-RECOVERABLE FOCUS OF ATTENTION (INTENDED BY SPEAKER)	SENSITIVE TO CERTAIN PARTICLES NARROW SCOPE

Table 1: Topic and Focus as semantic and pragmatic notions

There are two views on the role of contrast. The first is that focus always evokes a set of alternatives (Krifka & Musan 2012), i.e. all uses of focus go back to contrast, the second is that contrast is independent of both focus and topic (Erteschik-Shir 2009). For the purpose of this paper I take the second view.

As pointed out by Gundel et al. (1993) and Hedberg (2013), topicality as *givenness* may be either relational or referential. An element may be *relationally given* with respect to another one on the semantic, syntactic or pragmatic level, or *referentially given* in the sense that there is ‘a relation between an **intended referent** of a linguistic expression (...) and its informational status in the

memory/attention states in the hearer’s mind’ (Hedberg 2013:2). Topicality (or in-focus status, as Gundel et al. put it) can be established by (i) a preceding reference in the same sentence, (ii) a referent focused in a preceding clause (existential or cleft sentence) or (iii) the event expressed in a preceding clause. Reference as such is complex. McFarland (1978) characterises reference as a linguistic process that is constrained by the nature of the referent (specific, generic, non-specific), by the speaker’s knowledge about the referent (a file of prior knowledge about a referent) and the hearer’s knowledge (as assumed by the speaker). Reference and referentiality are central to many approaches to voice and subject choice in Tagalog. The idea is that subject choice and morphosyntactic coding in Tagalog is primarily determined by the referentiality of the theme argument (of a two-place predicate). There are various variations and elaborations of San José’s (1610) rule of thumb that an indefinite or non-specific theme argument of a two-place predicate cannot be subject (*ang*-marked), while a definite or specific theme argument has to be subject. A more recent and more sophisticated proposal along these lines is that the noun phrase type and the definiteness hierarchy play a major role (cf. Sabbagh 2012, who takes differential object marking and case marking exceptions are taken into account).

In Latrouite (2011, 2012) I suggest that subject choice and morphosyntactic coding in Tagalog is determined by referentiality (including animacy) considerations of the theme argument, event semantics (ES) and information structure (IS) and that the three levels are hierarchically ordered: IS > ES > RS. The central question to be asked in this context is then: *which aspects of IS are meant here?*

3. Coding of Information Structure in Tagalog

The languages of the world display various linguistic means for structuring information with respect to (i) **the discourse-level**, (ii) **the sentence-level** and (iii) **the clause-level**. The linguistic means may be **lexical, morphological, syntactic or prosodic**. In the first part I sum up findings that have been brought forward with respect to the discourse level.

3.1 Structuring Information on the Discourse-Level (via referential expressions)

On the discourse-level it is important to keep track of referents. Table 2 gives an overview of the noun markers used in Tagalog. Common nouns and personal names take a different set of noun markers that display the same three-fold distinction with respect to case as personal pronouns and demonstrative pronouns. Animacy plays an important role in Tagalog. Personal pronouns are (usually) restricted to animate referents, while inanimate referents are usually taken up by demonstrative pronouns.

	Nominative	Genitive	Dative
Common Nouns	<i>ang</i>	<i>ng</i>	<i>sa (object only)</i>
Personal names singular	<i>si</i>	<i>ni (actor only)</i>	<i>kay (object only)</i>
Personal Pronoun 3s	<i>siya</i>	<i>niya (actor only)</i>	<i>sa kaniya(object only)</i>
Dem. Pronoun Prox.	<i>ito</i>	<i>nito</i>	<i>dito</i>

Table 2: Tagalog noun markers

In his own corpora Nagaya (2006: 95) finds that, disregarding animacy considerations, Tagalog speakers may use personal pronouns for global discourse topics or continued topics and demonstrative pronouns for non-topical arguments. Note that the the grammaticalized topic (the nominative-marked argument or subject) in the sentences in (1) is not necessarily the global discourse topic, i.e. a non-discourse topic coded via a demonstrative pronoun, as in (1b), may be the sentence topic/subject. The notion *topic* in Nagaya (2006) seems to be used in the filecard sense of the notion, i.e. in (1) it is the filecard of *bata* ‘the child’ under which the speaker intends

the hearer to store all new information. The choice of referential expression thus signals which referent the speaker wants the hearer to focus on and follow in the subsequent chain of sentences. According to Nagaya, both the zero realization in (1c) and the demonstrative pronoun in (1b) imply that *bata* ‘child’ is the global discourse topic. Sentences like (1d) with two third person pronouns are found to be less acceptable. The reason seems to be that reference tracking becomes harder.

- (1) a. *H<in>a~hanap ng bata yun-g nanay niya.*¹

IPFV<RLS>_{UV}look.for GEN child Dem.NOM-LK mother 3SGEN

‘The child was looking for his/her mother.’

- b. *Tapos na-kita niya ito sa kusina.*

Then UV.RLS-see 3SGEN DP_{NOM} DAT kitchen

‘Then (s)he found her/ this one in the kitchen.’

- c. *Tapos na-kita niya Ø sa kusina.*

Then UV.RLS-see 3SGEN DAT kitchen

‘Then (s)he found (her/this one) in the kitchen.’

- d. **Tapos na-kita niya siya sa kusina.*

Then UV.RLS-see 3SGEN 3SNOM DAT kitchen

Intended: ‘Then (s)he found (her/this one) in the kitchen.’

My own consultants reject (1c) and accept (1d). It is possible that they construe ‘mother’ as the global topic and have a preference for construing the grammaticalized topic as the global discourse topic without further context. Note that reference to the speaker and the hearer is always marked by personal pronouns regardless of global topichood, showing that pronouns simply mark the in-focus (attention) state of the respective argument, not global discourse topichood per se. However, in-focus, i.e. very topical, arguments tend to outrank arguments lower on the Givenness Hierarchy for global topichood. Given their prominence, it does not come as a surprise that in contrast to third person pronouns, first and second person pronouns may not be left out, i.e. the speaker and the hearer may not be coded as zero. This is shown in (2) for the first person singular. Native speakers clearly prefer for *ako* to be overt, even if it can be inferred. Note that the same holds for the addressee (*mo* ‘you’), even though the verb is imperative, so that the addressee has to be the Actor argument of the sentence.

- (2) *Tawag-an mo *(ako).*

Call-TV 2SGEN 1SNOM

‘You call me.’

As mentioned above, inanimate arguments are usually referred to by demonstrative pronouns, but may be referred to by a personal pronoun, *if* – and according to Nagaya (2006) only if – they are the global discourse topic, as in (3b) and (3c) taken from an excerpt about the best gift the speaker has ever received.

¹ Glosses: AV: Actor voice; ACC: accusative; GEN: genitive; DAT: dative; DEM: demonstrative pronoun; IPFV: imperfective; msc: masculine; LK: linker; LV: locative voice; NOM: nominative; P: plural; OBL: oblique; PAST: past tense; PRT: particle; RLS: realis; s: singular; ST: stative; TOP: topic; TV: theme voice; UV: Undergoer voice. Infixes are marked by < > and separate the initial consonant of the stem (C_{stem}) from the rest of the verb stem. Glosses in subscript and square brackets are not morphologically marked but implicit to the form.

- (3) a. *B<in<igy-an ako ng gift ng mother ko, father ko saka kuya ko sila-ng tatlo.*
 <RLS>give-LV 1SNOM GEN gift GEN mother 1SGEN father 1SGEN and older brother 1SGEN 3PNOM-LK three
 'My mother, my father and my older brother, the three of them, gave me a gift.'
- b. *Libro siya ng fairy tales.*
 book 3SNOM GEN fairy tales
 'It was a book of fairy tales.'
- c. *Tapos simula 10 year old ako hanggang ngayon b<in>a~basa ko pa rin siya.*
 Then around 1SNOM until now <RLS>_{UV} IPFV~read 1SGEN still also 3SNOM
 'Since I was ten until now I have still been reading it.' (Nagaya 2006: 97)

Nagaya argues that zero coding is used for animate third person pronouns, if they express non-topics on the global discourse level, regardless of sentence-level topichood. The idea is that zero coding backgrounds arguments and that in (4) the information is intended by the speaker to be stored under the file card of *ako* 'I', not under *the Japanese*. As the referent of the zero coded argument has to be inferred from the mention of *Japan* in (4a), it is less given than the speaker, i. e. it is lower on the Givenness Hierarchy.

- (4) a. (...) *nung nasa japan lagi ako-ng k<in>a~kausap Ø in japanese.*
 ... when in Japan always 1SNOM-LK IPFV~<RLS>_{TV}.talk.to
 'When (I was) in Japan, (they) were always talking to me in Japanese.'
- b. *tapos siguro mga dalawa tatlong beses na ako-ng Ø t<in>anong for directions*
 then probably PL two three-LK times LK 1SNOM-LK <RLS>_{UV}ask
 'Later I was probably asked two or three times for directions (by them).' (Nagaya 2007: 98)

In (5b) a new referent is introduced into the pear story by the phrase *tatlong bata* '(the) three children', turned into a part of the **common ground**, but not into the global topic. The global discourse topic is still 'the boy' (*siya* 'he'), which is shown by the fact that the new referent is coded as zero in the subsequent sentence (5d).

- (5) a. *na-hi~hirap-an na siya.*
 ST-IPFV~poor-LV PRT 3SNOM
 'He was having problems.'
- b. *may may naka-kita sa kanya-ng tatlo-ng bata-ng,*
 exist exist AV-see DAT 3SDAT-LK three-LK child-LK
 'There were three children who saw him.'
- c. *tatlong bata na nag-namamasyal.*
 three-LK child LK AV.RLS-roam
 'Three kids who were roaming around.'
- d. *t<in>ulung-an Ø siya-ng i-lagay yun-g prutas sa kaing.*
 <RLS>help-LV 3SNOM-LK TV-put DEM.NOM-LK fruit DAT basket
 '(They) helped him (=the boy) put the fruit into the basket. (Nagaya 2007:94)

As to be expected, zero coding does not seem to be admissible for an animate referent that stays the global discourse topic, as exemplified in (6) for the argument *Wendy*, which is taken up over and over again by *siya* 'she' rather than zero coded. ?? indicates that speakers highly disprefer zero coding.

The newly introduced referent *Flor* in (6d) as non-discourse topic, however, may be coded by zero, as shown in (6e). Due to the information-structural reference tracking function of the pronoun, (6e) and (6f) are not understood to be ambiguous. Rather speakers understand right away that Wendy is the one invited by Flor and without money.

- (6) a. *Kahapon p<um>unta si Wendy sa UP.*
 yesterday <AV>_{RLS}go.to NOM W DAT UP
 'Yesterday Wendy went to UP.'
- b. *Nag-lakad ??(siya) papunta doon.*
 AV.RLS-walk 3SNOM to there
 'She walked there.'
- c. *dahil wala ??(siya)ng pera.*
 due.to NEG 3SNOM-LK money
 'because she did not have money.'
- d. *Tapos, naka-salubong ??(siya) ni Flor.*
 then AV.RLS-scome across 3SNOM GEN F
 'Afterwards she came across Flor.'
- e. *Ni-yaya Ø ??(siya)-ng mag-meryenda.*
 UV.RLS-invite 3SNOM-LK AV-snack
 '(Flor) invited her to (have a) snack.'
- f. *eh wala nga ??(siya)ng pera*
 NEG PRT 3SNOM-LK money
 'She really did not have money.'
- g. *Kaya, ni-libre ??(siya) ni Flor.*
 SO UV.RLS-free 3SNOM GEN F
 'So Flor paid for her.' (Nagaya 2007:87)

Zero coding is admissible and preferred for inanimate referents that are not the global discourse topic, but the local topic in a topic chain of a paragraph, as in the following example taken from the pear story, which is primarily not about a box and its adventures, but a boy.

- (7) a. *B<in>uhat niya ang isa-ng kahon.*
 <RLS>_{TU}lift 3SGEN NOM one-LK box
 'He lifted up one box.'
- b. *D<in>ala niya Ø ngayon sa bayk*
 <RLS>_{UV}carry 3SGEN now DAT bike
 'He carried (it) then to the bike.'
- c. *I-s<in>akay niya Ø*
 TV<RLS>put 3SGEN
 'He put (it) down.'
- d. *Um-alis na siya ngayon.*
 AV.RLS-leave PRT 3SNOM now
 'He left then.' (Nagaya 2007: 95)

Summing up, we can conclude that:

➔ Global topichood on the discourse level is not tied to grammaticalized topichood.

- ➔ Given and activated arguments may be coded by zero, pronouns and demonstratives, but only pronouns and demonstratives are used to mark global topics on the highest level; pronouns seem to outrank demonstratives for the coding of global discourse topic status.
- ➔ Zero coding is chosen for given arguments not intended to be the global discourse topic.
- ➔ Apart from global discourse topics and sentence topics, there may be local topics in a discourse.

3.2 Syntactic and Prosodic Means of Structuring Information on the Sentence-Level

Tagalog has two syntactic ways of marking a frame-setting topic (8), arguments being marked by the topic marker *ay* and adjuncts optionally so, and two ways to mark contrastive and exhaustive focus (9) (cf. Katagiri 2006). The different strategies depend on the nature of the phrase fronted, i.e. whether it is a nominative argument of the verb, as in the cleft construction in (9a), or not, as in (9b) and (9c).²

- (8) a. *Siya* **(ay)** *t<um>a~tawa.*
 3sNOM TOP <AV>_{RLS}IPFV~laugh
 'He was laughing.'
- b. *Kahapon* **(ay)** *t<um>awa siya.*
 yesterday TOP <AV>_{RLS}laugh3sNOM
 'Yesterday, he laughed.'
- (9) a. *SIYA* **ang** *t<um>a~tawa.*
 3sgNOM NOM <AV>_{RLS}IPFV~laugh
 'He was the one who was laughing.'
- b. *KAHAPON* *siya* *t<um>awa sa kaniya.*
 yesterday 3sNOM <AV>_{RLS}laughDAT 3sOBL
 'YESTERDAY he laughed at him.'
- c. *SA* *KANIYA siya* *t<um>awa kahapon.*
 DAT 3sOBL 3sNOM <AV>_{RLS}laugh yesterday
 'AT HIM he laughed yesterday.'

A sentence-initial LDP (=left detached position) is for setting a frame for a proposition and can be iterated (10b). The sentence-initial Pre-Core Slot is for contrastive, exhaustive foci, and cannot be iterated. The order is always TOPIC>FOCUS, as shown in (10a, c).

² Both structures are not good in answer to a simple information question like

- | | |
|---|--|
| Q: <i>Ano ang g<in>awa mo kahapon?</i>
what NOM <RLS> _{TV} do 2sGEN yesterday | 'What did you do yesterday?' |
| ?A1: <i>Kahapon ay t<um>ulong ako sa mga bata.</i>
yesterday TOP <AV> _{RLS} help 1sNOM DAT PL child | 'Yesterday, I helped the children.' |
| #A2: <i>Kahapon ako t<um>ulong sa mga bata.</i>
yesterday 1sNOM <AV> _{RLS} help DAT PL child | 'It was yesterday that I helped the children.' |
| #A3: <i>Sila ang t<in>ulong-an ako kahapon.</i>
3pNOM NOM <RLS>help-LV 1sNOM yesterday | 'They were the ones I helped yesterday.' |
| #A4: <i>Sila ay t<in>ulong-an ako.</i> | 'Them, I helped.' |
| A5: <i>T<um>ulong ako sa mga bata kahapon.</i> | 'I helped the children yesterday.' |

- (10) a. *Kahapon ay ANG BABAE ang t<um>awa.*
 yesterday TOP NOM woman NOM <AV>laugh
 'Yesterday, it was THE WOMAN who laughed.'
- b. *Kahapon ay siya ay t<um>awa.*
 Yesterday TOP 3sNOM TOP <AV>laugh
 'Yesterday, he laughed.'
- c. **ANG BABAE ang sa kaniya ay t<um>awa.*
 NOM woman NOM DAT 3sgOBL TOP <AV>laugh
 'THE WOMAN was the one who laughed at him.'

Interestingly, if there are two contrasted arguments of the verb, fronting is not possible, and both contrastive constituents stay in situ and are marked prosodically, as shown in (11). Kaufmann's (2006) explanation for this is that there is only one syntactic focus position and if it is occupied, then the rest has to be interpreted as comment/ presupposed, so neither arguments ends up in the focus position but they both stay in-situ.

- (11) *B<in>a~basa ang LIBRO ni JOHN.*
 <RLS>_{TV}IPFV.read NOM book GEN John
 'JOHN read a BOOK.'

Even if the two contrasted elements do not both have direct argument status, a double focus fronting strategy is out, as shown in (12).

- (12) **SA KANIYA SIYA ang t<um>awa.*
 DAT 3sOBL 3sNOM NOM <AV>_{RLS}laugh
 'AT HIM HE was the one who was laughing.'

If the non-Actor is a semantic argument of the verb, even topic fronting of this phrase is rejected in a sentence with a focused Actor, as shown by Latrouite (2011).

- (13) #*Kay Lina (ay), SIYA ang t<um>awa.*
 DAT Lina TOP 3sNOM NOM <AV>_{RLS}laugh
 ?'(As for) at Lina HE was the one who laughed.'

Actors however, may be frame-setting topics in sentences with focused Undergoers.

- (14) *Si Lina (ay), SILA ang t<in>awa.*
 NOM Lina TOP 3pNOM NOM <UV>_{RLS}laugh
 '(As for) Lina, THEY were laughed at (by her).'

A theory of IS should be able to explain these restrictions on syntactic structure and the interaction with prosody. Latrouite (2011, 2013) suggests that the facts above are tightly linked to the voice system and the nature of subject selection in Tagalog. Those arguments that are the most prominent on the referential structural [RS], the event-structural [ES] and/or information-structural level turn into the subject. The levels are assumed to be ordered (IS > ES > RS), so that IS is the most important-level. If the Actor and the Undergoer arguments are both prominent on the information-structural level, the Undergoer argument outranks the Actor for subjecthood, this is why the Undergoer may not be topical in the Actor-focus sentence, but the Actor may be topical in the Undergoer-focus sentence.

3.3 Structuring Information prosodically

On the core-level, focus is discussed with respect to focus-sensitive particles, while topic is – as already mentioned – discussed with respect to subjecthood. Prosodically the sentence-final position is marked (bears nuclear stress) and easily receives a (contrastive) focus reading. As the following sentences from Kaufmann (2006: 193) show, the focus-sensitive clitic *rin* ‘also’ refers to the sentence-final referential phrase in the sentence.

- (15) a. *Bukod kay Ricky, i-p<in>a-kilala ko rin kay Paolo si John.*
 besides DAT R TV<RLS>CAUS.know 1sGEN also DAT P NOM J
 ‘Besides Ricky, I also introduced **John** to Paolo.’ (Focus Implication: I also introduced R to P)
- b. *Bukod kay Ricky, i-p<in>a-kilala ko rin si John kay Paolo.*
 besides DAT R TV<RLS>CAUS.know 1sGEN also DAT P NOM J
 ‘Besides Ricky, I also introduced John to **Paolo**.’ (Focus Implication: I also introduced J to R)

The examples show that non-canonical word order may be licensed by focus, e.g. a sentence-final genitive-marked Actor, is licensed if the Actor is focused as in the following correction sequence in (16) (Kaufmann 2006: 194).

- (16) A: *Kinausap ng bawan propesor ang mga estudyante?*
 ‘Did all the professors talk to the students?’
- B: *Hindi. Hindi k<in>ausap ang mga estudyante ni Propesor Martinez.*
 NEG NEG <RLS>_{UV}talk.to NOM PL students GEN Prof. M
 ‘No, **Professor Martinez** did not talk to the students.’

Kaufmann’s conclusion is that two out of three focus marking strategies, i.e. oblique fronting (15) and prosodic movement (16), are not related to voice and subject choice. Only clefting enforces that the focus be the subject/grammaticalized topic. The subsequent sections deals with those aspects of information structure that are relevant to voice and subject selection.

4. The Grammaticalized Topic and Information Structure

Katagiri (2006) finds that **patient-orientedness** outranks (**discourse**) **topichood** for subject selection. But what is patient-orientedness? Taken in the narrow sense, a patient is an argument undergoing a change of state, but there is only a limited set of verbs that take true patient objects. On this interpretation, patient-orientedness cannot be considered as characterising a language-system as a whole. Usually ‘patient’ in ‘patient-orientedness’ is understood in a broader sense of ‘all Undergoer arguments affected by the action depicted by the verb.’ The patient-orientedness of Tagalog lies in the fact that, according to *the referentiality hypothesis*, a definite or specific Undergoer in Tagalog has to appear as the grammaticalized topic/subject, overruling all other considerations. However, there are exceptions to this, showing that strict patient-orientedness is indeed restricted to a limited class of verbs:

- (17) a. *Ba~basa ang bata ng/sa libro.*
 AV.IPFV~read NOM child GEN/DAT book
 ‘The child will read a/the book.’ (De Guzman 1999, cited from Katagiri 2005: 164)
- b. *D<um>a~dalo ako ng/sa meeting.*
 <AV>_{RLS}IPFV~attend 1sNOM GEN/DAT meeting
 ‘I attend a meeting/the meeting.’ (Bowen 1965: 222)

- c. *Na-alala ko nag-ba~basa ako nito.*³
 AV.RLS-remember 1SGEN AV.RLS-IPFV~read 1SNOM this.GEN
 'I remember reading this.'
- d. *Nang-ha~harana ang binata ng/sa dalaga.*
 MANG.RLS-IPFV~serenade NOM young man GEN/DAT lady
 'The young man serenades ladies/ the lady.'
- e. *Naka-kita ang bata ng/*sa aksidente.*
 AV.RLS-visible NOM child GEN/DAT accident
 'The child saw a/*the accident.' (cf. Schachter & Otones 1972: 382)

The data demonstrate that *sa*-marking is only possible if a theme argument may be interpreted as specific with a certain AV-form, but it is not merely an overt specificity marker: *sa*-marking is subject to processing considerations as is differential object marking in general (cf. Latrouite 2011, 2012). Given the general tendency that specific themes have to be realized as subjects and induce UV, the interesting question is: what are the conditions under which an Actor may be subject despite a specific theme? Latrouite (2011, 2012) makes claims about these conditions. The basic idea is that the subject argument is selected because it is the most prominent and that prominence can be evaluated on three ordered levels, the level of referentiality, the level of event structure and the level of information structure: **IS prominence > ES prominence > Referential Prominence**. The specificity restriction with respect to the non-subject Undergoer(, which is related to the level of referentiality,) may only be overridden if the Actor is more prominent than the Undergoer in event-structural or information-structural terms. **EVENT-STRUCTURAL PROMINENCE** is about the inherent orientation of a verb stem (result- vs. process/activity-oriented vs. neutral verbs); one argument may be more topical with respect to the e.

An argument of a certain verb is event-structurally prominent due to the fact that

- (a) when decomposing the predicate into meaning components, the specific meaning component only provides specific information on this one argument; (e.g. the object experiencer verb *tumakot* 'frighten' from *takot* 'fear'; the fear is attributed to the theme and Actor voice in basic sentences is strongly dispreferred (18)).
- (b) it delimits the runtime or measures out the event (e.g. for incremental theme verbs 'to run', 'to write', 'to build' etc.).

- (18) * *T<um>akot siya kay Juan.*
 AV.RLS-fear 3SNOM DAT J
 Intended: 'He frightened Juan.'

Certain verbs thus have an inherent orientation. However, the inherent orientation may be overridden, e.g. by explicit emphasis on the activity part of the run-time of the event (cf. Latrouite 2011). This is mostly achieved by aspect choices, i.e. imperfective marking.

³ (context: I remember that I read (this= Abante bed time stories), it was funny because these journals (they) belonged to the kumpare of my Dad, I just borrowed (them) pretending to use (them) in school.)

<http://www.pinoyexchange.com/forums/showthread.php?t=441236&page=5>

- (19) a. *T<um>a~takot ng mga negosyante ang rallies.*
 <AV>[REAL]IPFV~fear GEN PL entrepreneur NOM rallies
 'The rallies are frightening/frighten (the) entrepreneurs.'
 (simplified from Pilipino Star Ngayon, December 12, 2000, Mag-rally or tumahimik)
- b. *T<um>a~takot⁴ sa kanya ang pagpapalaglag sa kanya-ng magiging anak.*
 <AV>[REAL]IPFV ~fear DAT 3SDAT NOM abortion DAT 3SOBL-LK develop-LK child
 'The abortion of her developing child frightens/is frightening her.'
 (www.panitikanngpilipinas17.blogspot.com/)

As we have seen above, **INFORMATION-STRUCTURAL PROMINENCE** is achieved by placing an argument into the clause-initial focus or topic position. For certain verbs like *pumatay* 'to kill' which are inherently strongly Undergoer oriented, this is the only possibility to get Actor voice. (20a) shows that the Actor voice form of *pumatay* in a basic sentence is strongly dispreferred; (20b) shows a clefted Actor voice sentence with a focused Actor argument; (20c) and (20d) show clauses in which the Actor is the topic. While most native speakers readily accept (20c), my consultants differ in whether they accept (20d) or not. However, when looking at texts the structure in (20d) is indeed frequently found in Tagalog texts, as the examples mentioned in Sabbagh (2012) and cited below in (21) show.

- (20) a. **P<um>atay si Boyet kay Juan.*
 <AV>[RLS]dead NOM B DAT J
 Intended: 'Boyet killed Juan.'
- b. *Si Boyet ang p<um>atay kay Juan.*
 NOM B NMZ <AV>[RLS]-dead DAT J
 'Boyet was the one who killed Juan.'
- c. *S<in>ampal ko ang lalaki-ng p<um>atay kay Juan.*
 <UV.RLS>spank 1s.GEN NOM MAN-LK <AV>[RLS]dead DAT J
 'I spanked the man who killed Juan.' (Nagaya 2007:348/49)
- d. *?Siya ay p<um>atay kay Juan.*
 3SNOM TOP C <AV>[RLS]-dead DAT J
 ?'He killed Juan.'

The prediction based on the ideas presented here is that examples containing specific *ng*-marked Undergoers should be restricted to sentences with Actors that are information-structurally prominent or event-structurally prominent, i.e. they should contain verbs that are either inherently Actor-oriented or neutral with respect to orientation. The examples in (21) from Sabbagh (2012) are a point in case. They show that the definiteness of the Undergoer does not outrank information-structural considerations in the choice of the grammaticalized topic. In (21a) the Actor is a preposed topic and Actor voice of the clearly inherently Undergoer-oriented verb *magbukas* 'to open' is acceptable despite the presence of a definite Undergoer. Interestingly the subsequent sentences in (21) show that the mere fact that the Actor appears in the preverbal position, and not in the basic postverbal position, is often enough to license Actor voice despite a definite Undergoer. This suggests that the preverbal position in itself is information-structurally significant, even if the Actor only appears in this position because it is a clitic (21c, 21d).

- (21) a. *ako ay na-tulog, k<um>ain, nag-bukas nito-ng kumpyuter,*
 1SNOM TOP AV.RLS-sleep <AV>[RLS].eat, AV.RLS-open this.GEN-LK computer
na-nood ng telebisyon, na-tulog uli.
 AV.RLS-watch GEN television AV.RLS-sleep again
 'I slept, ate, **opened this computer**, watched television, and then went back to

⁴ better with *naka-takot* (p.c. R. Kvist)

sleep again.’ (<http://pasapinoy.com/2011/03/21/historyador-iv-bagabag/>)

- b. *At kaya gusto ko-ng ma-nalo ng award na ito*
and so want 1s.GEN-LK AV-win GEN award LK this ‘
‘And so I want to win this award [=Comedy actress award]’ (<http://m.pep.ph/moblie/news>)
- c. *Hindi ba kayo nag-kita ng asawa ni Col. Adante?*
not Q 2pNOM AV.RLS-see GEN spouse GEN Col. Adante
‘Have you not met Col. Adante’s wife?’
(<http://www.pinoyexchange.com/formus/printthreadphp?t=345875&pp=40&page=43>)
- d. *Hindi matanda-an ni Noel kung saan at kailan sila nag-kita ng*
not remember-LV GEN Noel COMP where and when 3pNOM AV.RLS-meet GEN
pintor na kaibigan ni Allyssa na siya-ng g<um>uhit sa larawan-g ito.
painter LK friend GEN Allyssa LK 3sNOM-LK <AV>_{RLS}draw DAT picture-LK DP.NOM
‘Noel can’t remember where or when they met Allyssa’s painter friend who drew this picture.’
(<http://www.iluko.com/article.aspx?articleid>)

If information-structural prominence on the sentence level makes an argument more likely to turn into the grammaticalized topic and may overrule patient-orientedness, then we should also find that placing an argument into the prosodically prominent sentence-final focus position should license Actor subjects with specific or definite Undergoers or inherently Undergoer-oriented verbs. Discussing Undergoer-oriented pseudo verbs like *gusto*, which usually require the Undergoer to be marked by nominative, Schachter and Otnes (1972) already mentioned that certain speakers accept nominative-marked Actor arguments with this verb, if the Actor is realized sentence-finally and prosodically marked.

- (22) a. *??Gusto si Charlie ng mansanas.*
like NOM C GEN apple
Intended: ‘Charlie likes apples.’
- b. *Gusto ng mansanas si Charlie.*
like GEN apple NOM C
‘CHARLIE likes apples.’ (cf. S & O 1972)

Similarly, we find sentences in Sabbagh’s (2012) corpora that exhibit sentence-final Actor arguments in Actor voice sentences with definite Undergoers.

- (23) a. *Tanging p<um>ansin kay Elias si Maria Clara.*
Only <AV>_{RLS} notice DAT E NOM MC
‘Only Maria Clara noticed Elias.’ (A34 <http://monmen.multiply.com/journal/item/24>)
- b. *Pagkaraa’y isa-isang nag-basa ng kanila-ng tula ang akin-g mga estudyante.*
then TOP onebyone-LK <AV>_{RLS}read GEN their-LK poem NOM my-LK PL student
‘Then one by one my students read their poems.’
(<http://sandalilamang.blogspot.com/2007/02/pagtula.html>)

The data here suggest that *ay*-topicalization, clitic fronting, prosodic movement to the end of the sentence and subordination can serve similar purposes with respect to the selection of the Actor argument as the subject and in licensing Actor voice in ‘unexpected’ cases. A larger corpus study of the phenomena presented here is certainly called for, as well as a suitable framework in order to represent the similarities and differences of the different information-structural constructions exemplified above. Both is a desideratum of future research.

References:

- Bloomfield, Leonard. 1917. (with Alfredo Viola Santiago) *Tagalog Texts with Grammatical Analysis*. University of Illinois studies in language and literature, 3.2-4. Urbana, IL: University of Illinois.
- Bowen, Donald J. (ed.) 1965. *Beginning Tagalog*. Berkeley/Los Angeles: University of California Press.
- Gundel, Jeanett, N. Hedberg, and R. zacharski. 1993. Cognitive Statuses and the form of referring expressions. In *Language* 69 (2). 274-307.
- Hedberg, Nancy. *Applying the Givenness*. Hierarchy. Paper submitted for the 1st Workshop on 'Cross-linguistic perspectives on the information structure in Austronesian languages'. Tokyo, 13-16. Dec. 2013.
- Katagiri, Masumi. 2006. http://www-01.sil.org/asia/philippines/ical/papers/katagiri-topic_tagalog.pdf
- Kaufman, Daniel. 2005. Aspects of pragmatic focus in Tagalog. I Wayan Arka and Malcom. Ross (eds.) *The many faces of Austronesian voice systems: some new empirical studies*, 175-196.
- McFarland, C. 1978. 'Definite objects and subject selection in Philippine languages.' In Edrial-Luzares, Casilda & Austin Hale (eds.) *Studies in Philippine Linguistics*, Vol. 2. Manila: Linguistic Society of the Philippines, 139-182.
- Nagaya, Naonori. 2006. Preferred referential expressions in Tagalog. Tokyo University Linguistic Papers (TULIP) 25 (2006) 83-106.
- Nagaya, Naonori. 2007a. 'Information structure and constituent order in Tagalog.' *Language and Linguistics* 8(1): 343-372.
- Latrouite, Anja. 2011. *Voice and case in Tagalog: the coding of prominence and orientation*, Heinrich-Heine-Universität: PhD Dissertation.
- Latrouite, Anja. 2013. Referential Properties, Event Structure and the Syntax-Semantics Interface. Talk given at the 6th *Austronesian Papuan Languages and Linguistics Conference* 2013. London, SOAS
- Latrouite, Anja. 2012. 'Differential Object Marking in Tagalog.' In Lauren Eby Clemens and Greg Scontras (eds), *Proceedings of AFLA 18*, Harvard University
- Sabbagh, Joseph. 2012. Specificity and Objecthood in Tagalog. Manuscript- UT Arlington. ling.auf.net/lingbuzz/001647
- San José, de Francisco Blancas. 1610. *La Regla y Arte de la Lengua Tagala*.
- Schachter, P. & F. Otones. 1972. Tagalog reference grammar. University of California Press, Berkeley, CA.