From file cards to file folders: Nested aboutness and Tagalog nominal marking
This paper presents part of ongoing work on the morphosyntactic realization of nominal marking in Tagalog vis-a-vis information and discourse structure. Here, I argue that Tagalog topic marking shows how aboutness can nest hierarchically with progressive domain restriction.

The puzzle. Tagalog nominals are introduced by one of three markers: *ang*, *ng*, or *sa*. Traditional analyses have treated *ang* as marking nominative or absolutive case (Kroeger 1993, Aldridge 2004, a.o.), or as identifying the "pivot"—the argument indexed by verbal voice morphology (Schachter 1976). More recent work has argued that *ang* marks topichood rather than case (Richards 2000, V. Chen 2025, a.o.), but this leaves open the question of what kind of topic *ang* encodes, and how to account for its full distribution.

The distribution of *ang* presents an empirical puzzle for any unified account. First, *ang* appears in non-pivot contexts, such as in hanging topics (1) or exclamatives (2).

(1) *Ang babae$_i$, t⟨um⟩ingin ang lalaki sa kaniya$_i$.*
ang woman ⟨PRF.AV⟩look *ang* man *sa* 3S
'The woman, the man looked at her.'

(2) *Ang init!*
ang heat
'It's so hot!'

In (1), both the hanging topic *ang babae* and the pivot *ang lalaki* surface with *ang*, despite only the latter participating in pivot agreement—the verb surfaces in Agent Voice (AV). In (2), there is no verbal predicate at all—yet *ang* appears. Most strikingly, multiple *ang*-phrases can co-occur within a single clause, as in (1). If *ang* marks a unique grammatical relation such as case or pivothood, how can it appear multiple times? I propose that *ang* encodes aboutness—specifically, that it marks nominals that serve as aboutness topics in the sense of Reinhart (1981). Crucially, I argue that aboutness can nest hierarchically through progressive domain restriction, and that Tagalog morphologically marks each level of this nested structure. This provides rare overt evidence for the hierarchical organization of discourse that has been proposed in Questions Under Discussion (QUD) frameworks (Roberts 2012).

***Ang* marks aboutness.** I argue that *ang* encodes aboutness in Reinhart's (1981) sense: the entity that the sentence is about, which functions as the address where propositions are stored in the common ground. Both *ang*-marked pivots and hanging topics pass standard diagnostics for aboutness topics, such as continuation (3) and discourse frames (4).

(3) a. *T⟨in⟩ingn-an ng lalaki ang babae$_i$. Tapos, um-alis siya$_i$.*
⟨PRF⟩look-LV *ng* man *ang* woman after AV-leave *ang*.3S
'The man looked at the woman. Then, she left.'

b. *Ang babae$_i$, t⟨um⟩ingin ang lalaki sa kaniya$_i$. Tapos, um-alis siya$_i$.*
ang woman PRF.AV-look *ang* man *sa* 3S after AV-leave *ang*.3S
'The woman, the man looked at her. Then, she left.'

(4) Discourse frame: "Tell me about the woman."

A1.✓ *T⟨in⟩ingn-an ang babae ng lalaki.*
⟨PRF⟩look-LV *ang* woman *ng* man
'The man looked at the woman.'

A2.# *T⟨um⟩ingin ang lalaki sa babae.*
⟨PRF.AV⟩look *ang* man *ng* woman
'The man looked at the woman.'

A3.✓ *Ang babae$_i$, t⟨um⟩ingin ang lalaki sa kaniya$_i$.*
ang woman PRF.AV-look *ang* man *sa* 3S
'The woman, the man looked at her.'

In (3), *siya* refers to the *ang*-marked woman in both constructions. In (4), *ang*-marked responses (A1, A3) are felicitous answers while the *sa*-marked one (A2) is not. These diagnostics establish that *ang* encodes aboutness, regardless of pivothood.

Nested aboutness and domain restriction. Reinhart (1981) proposed that aboutness topics specify the address where propositions are stored in the common ground—like file cards in a library catalog. If both *ang babae* and *ang lalaki* are aboutness topics, how can a single clause have two addresses? I propose aboutness nests through progressive domain restriction: from file cards to folders containing sub-folders. Each aboutness topic restricts the domain of discourse for subsequent material:

(5) For n nested aboutness topics $t_1 ... t_n$, D_0 = unrestricted discourse domain:
$D_n = \{x \ : \ x \in D_{n-1} \ \wedge \ \textsc{relevant-to}(x, t_n)\}$ where $D_0 \supseteq D_1 \supseteq D_2 \supseteq ... \supseteq D_n$.

Each aboutness topic establishes a progressively more restricted domain: inner topics must denote entities in the domain established by outer topics. An *ang*-phrase identifies an entity x as the aboutness topic in some domain D, and updates the domain to restrict it to entities relevant to x. Formally:

(6) $[\![\text{ang NP}]\!] = \lambda D.\lambda P_{\langle e,t\rangle}.\iota x[\text{NP}(x) \wedge x \in D \wedge \textsc{about}(x) \wedge P(x)]$
$\wedge \ D := D \cap \{y : \textsc{relevant-to}(y, x)\}$.

For instance, the initial discourse domain D_0 for (1) would be first restricted to a $D_1 = D_0 \cap \{y : \textsc{relevant-to}(y, \text{the woman})\}$—the set of entities relevant to the woman. Then, *ang lalaki*, identifies the man as another aboutness topic and further restricts the domain to $D_2 = D_1 \cap \{y : \textsc{relevant-to}(y, \text{the man})\}$. The domain is thus restricted to entities relevant to both the woman and the man. The proposition is thus 'filed' under D_2.

This connects to Roberts' (2012) QUD framework: an aboutness topic in a sentence restricts the QUD to questions about that entity (establishing D_1). Later topics in that sentence establish sub-QUDs within that domain ($D_2 \subseteq D_1$). Tagalog morphologically realizes each level of this QUD hierarchy with *ang*, making nested discourse structure overtly visible.

Possession and pragmatics. Domain restrictions predict when nested *ang*-phrases are felicitous, such as with inalienable possession. In (7), the innermost topic, the color, must be that of the hair, which in turn must be that of the woman, all without explicit possessive *ng*.

(7) *Ang babae, ang buhok, matingkad ang kulay.*
ang woman *ang* hair bright *ang* color
'The woman, her hair, it's bright.' (*Lit.* 'The woman, the hair, the color is bright.')

Crucially, domain restriction is pragmatic. A sentence like (7) would normally be infelicitous if the hair under discussion were not intended to refer to the woman's. Judgments for this improve with pragmatic relevance, e.g., the woman is a hairdresser, showing that $D_{n+1} \subseteq D_n$ can be determined by semantic relations or contextual relationships.

Conclusion. Tagalog *ang* encodes aboutness through domain restriction, with nested domain restriction obtained through multiple *ang*-phrases in one sentence. This enriches our typology of discourse-sensitive nominal marking: languages often have no overt topic marker, or just one unique topic marker per sentence; Tagalog can have multiple. This raises questions about constraints on nesting depth and cross-linguistic patterns. Japanese *wa* and Korean *nun* mark topics but don't overtly distinguish nested levels, suggesting nested aboutness may be universal (as domain restriction) but morphological realization is parametric.

References

Aldridge, Edith. 2004. "Ergativity and Word Order in Austronesian Languages." Doctoral dissertation, Cornell University.

Chen, Victoria. 2025. "The Syntax of Philippine-Type Alignment: Insights from Case-Marking." *Natural Language & Linguistic Theory* 43 (February): 1839–98. https://doi.org/10.1007/s11049-025-09655-7.

Kroeger, Paul. 1993. "Phrase Structure and Grammatical Relations in Tagalog". *Dissertations in Linguistics*. CSLI Publications.

Reinhart, Tanya. 1981. "Pragmatics and Linguistics: An Analysis of Sentence Topics." *Philosophica* 27 (0). https://doi.org/10.21825/philosophica.82606.

Richards, Norvin. 2000. "Another Look at Tagalog Subjects." In Formal Issues in Austronesian Linguistics, edited by Ileana Paul, Vivianne Phillips, and Lisa Travis, vol. 49. Studies in Natural Language and Linguistic Theory. Springer Netherlands. https://doi.org/10.1007/978-94-017-1580-5.

Roberts, Craige. 2012. "Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics." *Semantics and Pragmatics* 5 (December). https://doi.org/10.3765/sp.5.6.

Schachter, Paul. 1976. "The Subject in Philippine Languages: Topic, Actor, Actor-Topic, or None of the Above?" In *Subject and Topic*, edited by Charles Li. Academic Press.