

Article

# On the Overlooked Diversity of Clause Structures and Argument Structures in Non-Indo-European Languages

Randy J. LaPolla 

Centre for Liberal Arts and Social Sciences, Nanyang Technological University Singapore, Singapore 639818, Singapore; randylapolla@protonmail.com

**Abstract:** This article responds to a conference call for papers that makes universalist assumptions about clause structures, assuming all languages in the world basically follow the same organizing principles in terms of clause structure, argument structure, and alignment. The article presents data from Tagalog to show how different a language can be from the assumed universal organizing principles to make the point that by imposing an Indo-European framework on non-Indo-European languages, we are overlooking the true diversity of language forms found in the world's languages.

**Keywords:** argument structure; clause structure; grammatical alignment; head-marking; Tagalog; Philippine-type languages; Mandarin Chinese

## 1. Introduction

The call for papers for a recent conference on argument coding patterns<sup>1</sup> states that “Argument coding patterns consist of bound markers indicating the semantic and syntactic dependency of the arguments from their verb and are either argument-bound (flagging or dependent-marking) or verb-bound (indexing or head-marking)...”. This is a universalist statement assuming all languages are like the major Indo-European languages in terms of having only the two options given here, head-marking or dependent marking. It also states, “Transitive and intransitive verbs are also relatively stable cross-linguistically in terms of their alignment options (ergative, accusative or a mixture of the two).” This assumes there are only two alignment types cross-linguistically or within a language, another non-empirical universalist assumption based on Indo-European languages. Because of such common universalist assumptions, linguists regularly try to force non-Indo-European languages into an Indo-European Procrustean bed. This is unfortunate, as we then overlook the true diversity of patterns and organizational principles manifested in the world's languages. A corollary of this type of view is a view of grammar as autonomous from speakers; all morphology is seen as systematic, even in recent work on variable marking, yet in many languages the role-marking morphology is not paradigmatic and systematic, but semantic marking used when the speaker feels the need to reduce possible ambiguity, such as using an agentive or anti-agentive marker when there are two animate referents mentioned in a clause (see LaPolla 1992, 1995 for early discussions of non-systemic anti-agentive and agentive marking, respectively). This means there is a need to recognize that some languages do not have an alignment.

## 2. Mandarin Chinese

Although Mandarin Chinese is one of the non-Indo-European languages that is often forced into an Indo-European-type analysis, there is a long history of scholars working on Chinese showing that the clause structure of Mandarin Chinese does not have one of the standard assumed alignments, or even a fixed argument structure, as it has not grammaticalized a system for constraining the identification of the roles of the main participants in the clause, and the unmarked clause is simply topic-comment (e.g., Chao [1948] 1967, [1955] 1976, [1959] 1976, 1968; Lü 1979; LaPolla 1990, 1993, 2009; LaPolla and Poa 2006).<sup>2</sup>



**Citation:** LaPolla, Randy J. 2024. On the Overlooked Diversity of Clause Structures and Argument Structures in Non-Indo-European Languages. *Languages* 9: 135. <https://doi.org/10.3390/languages9040135>

Academic Editors: Jeanine Treffers-Daller, Chenglong Huang, Youjing Lin, Ziwo Lama and Elvis Yang Huang

Received: 11 May 2023  
Revised: 4 December 2023  
Accepted: 19 March 2024  
Published: 8 April 2024



**Copyright:** © 2024 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

In this paper I want to focus on Tagalog, a Malayo-Polynesian language spoken in the Philippines, which has a very different clause structure from Indo-European languages, and does not mark arguments in the Indo-European way, and is also quite different from Chinese in structure.

The question is how many more systems might we discover if we work inductively on natural language data and stop imposing the Indo-European system on non-Indo-European languages?

### 3. Tagalog

#### 3.1. The Tagalog Clause

As shown in LaPolla (2014) (see also LaPolla 2019, 2023), Tagalog clause structure does not have any constituent phrase categories we can identify as “noun phrase” and “verb phrase”; the constituents found are constructions that link elements together, but the same constructions can be used both for reference and predication. All lexical items can be used either predicatively or referentially, so there is no grammatical distinction between noun and verb (Himmelman 2004).<sup>3</sup> Non-topical core arguments<sup>4</sup> in Tagalog are not distinguished from each other, regardless of whether they represent actors or different types of undergoers; they are represented the same way, linked to the predicating element in a possessive phrase (a “*ng* [nən] phrase” or as a possessive pronoun, which is a second-position clitic following the head of the phrase, whether the phrase is predicative or referential), not by the traditional IE argument marking strategies mentioned in the EDAP2023 call for papers, and this possessive phrase (which is the totality of the predicate) is in apposition to the Topic<sup>5</sup> of the clause, as both have the same reference.<sup>6</sup> Using Indo-European terminology, the clause is an equational copula clause, but there is no copula. For example,<sup>7</sup>

- (1) *Pwede ko bang kunin yun leaves...?*  
 [pwede ko ba=ng kuhin-in]PRED [yung leaves]TOP  
 can 1sgposs Q=LNK take-UT that+LNK leaves  
 ‘Can I take them(,) the leaves...?’  
 (<https://delishably.com/beverages/How-to-Make-Malunggay-Tea-Home-Made-Moringa-Tea>, accessed on 25 April 2023)

Tagalog clauses are consistently focus-initial (unmarked clauses are predicate-initial), so the unmarked information structure is Comment–Topic (not Topic–Comment). As argued in LaPolla (2019), although the Topic usually occurs at the end of the clause (if it is not a pronoun and not omitted and not focal), the Theme, or starting point of the clause is still important because of the information packed into the Theme to assist the addressee in projecting (anticipating) the speaker’s communicative intention. The predicate in many cases marks aspect, realis/irrealis, and often the semantic role of the Topic of the clause, though in natural conversation much of the marking can be left out. In (1), the agent argument appears as a possessive (non-Topic) pronoun (*ko*) modifying the predicate (*pwede bang kunin*)—the pronoun is a second position clitic, and so appears after the first word of the predicate, i.e., in the middle of the predicate phrase, which takes the form of a “*na/=ng* linker phrase”, that is, the two parts of the predicator are linked together by the linker *na/=ng*<sup>8</sup>. Here, although the linker is marking the linkage of *pwede* and *kunin*, the linker actually appears on the question marker (*ba*) rather than on *pwede* in this context because the question marker is also a second position clitic. This marking on an element that is not actually being linked is neither dependent-marking nor head-marking, an option not given in the universalist statement quoted above. If this example was not a question, the linker would appear on the non-Topic pronoun: *pwede ko=ng kunin*. Only if nothing appeared between the first and second word of the predicate would the linker appear on an element of the predicate: *pwede=ng kunin*. Note that this structure is the same for elements of the predicate and also for elements of reference phrases; see, for example, the Topic in (2), below, where again the linker linking *anak* and *lalaki* (*anak na lalaki* ‘son’) is not on either of these words but appears on the possessive pronoun (*ko=ng*; the possessive pronoun

appears after *anak* because it is a second-position enclitic, but it is actually modifying the entire phrase *anak na lalaki* ‘son’):

- (2) *Ano ang tawag sa asawa ng anak kong lalaki?*  
 [ano]PRED [ang tawag sa asawa ng anak ko=ng lalaki]TOP  
 what SPEC call LOC spouse POSS child 1sgPOSS=LNK male  
 ‘What is my son’s spouse called?’  
 (<https://dayuhanglalaki.blogspot.com/2021/09/tawag-sa-asawa-ng-kapatid-mong-lalaki.html>, accessed on 5 May 2023)

Compare, for example, the following two examples:

- (3) *Kaibigan lang siya ng tatay ko.*  
 kaibigan lang siya ng tatay ko  
 [friend only <3sgT>TOP POSS father 1sgPOSS]PRED  
 ‘He is only my father’s friend’.
- (4) *Tinanong lang siya ng tatay ko.*  
 t<in>anong lang siya ng tatay ko  
 [<RNAT>asking only <3sgT>TOP POSS father 1sgPOSS]PRED  
 ‘My father only asked him.’

In these examples, *kaibigan ng tatay ko* ‘my father’s friend’ and *tinanong ng tatay ko* ‘asked by my father’ are phrases of the same type in Tagalog, though in English we translate them very differently. (See LaPolla (2014) for more on the different types of phrases in Tagalog, and Naylor (2005) and references therein on the isomorphy of referential and predicative phrases.)

As some Topic pronouns are also second-position clitics, the Topic actually can appear within the predicate, as it does in (3) and (4), and if it appears between the two parts of a “*na/=ng* linker phrase”, the linker linking the two discontinuous parts of the predicate appears on the Topic pronoun rather than on any element of the predicate, as with *siya=ng* in (5):

- (5) *Pero wala siyang kinalaman sa kwentong ito.*  
 pero [wala? <siya>TOP=ng kinalaman]PRED  
 but N.EXIST 3sgT=LNK involvement  
 But she doesn’t have anything to do
- [sa [kwento=ng ito]]LP  
 loc story=LNK PROXT  
 with this story. (Bob Ong 2003, *Alamat ng gubat*, Visprint, Inc., Manila, p. 1)

From an Indo-European point of view this is a difficult construction to understand, as an element (the Topic) that is not part of the predicate not only appears in the middle of the predicate, but also takes the relational marker (linker) that links the two parts of the predicate together. This example also shows the peripheral locative structure, with the general peripheral marker *sa* followed by another “*na/=ng* linker phrase”. So in this example the same structure, a “*na/=ng* linker phrase”, is used for both the predicator and for a peripheral argument.

Another point about this “*na/=ng* linker phrase” is that it does not fit the usual clear grammatical head-dependent structure found in most languages. In many cases the two elements of the “*na/=ng* linker phrase” can be reversed, e.g., in (5), the expression *kwento=ng ito* could also be said as *ito=ng kwento*, both meaning ‘this story’.

Returning to example (1), the agent appears as a non-topical argument, as the leaves (what is being taken out) is the Topic of the clause. The predicate is marked for Undergoer Topic (UT). A literal translation would be ‘is that which I can take out the leaves?’. The Topic also takes the form of a “*na/=ng* linker phrase”, with the demonstrative pronoun linked to the Topic with the linker *na/=ng*, so we can see that the predicate phrase and the Topic phrase have the same structure. This equational copula clause-like structure is

clearer in cases where we can find the same role marking in the predicator and the Topic reference phrase:

- (6) *Binilhan ng lalaki ng bigas ang tindahan ng lola.*<sup>9</sup>
- |                  |                 |        |           |        |      |
|------------------|-----------------|--------|-----------|--------|------|
| [b<in>ili-han    | ng              | lalaki | ng        | bigas] | PRED |
| <RNAT>buying-LFS | POSS            | male   | POSS      | rice   |      |
| [ang             | tinda-han       | ng     | lola]     | TOP    |      |
| SPEC             | merchandise-LFS | POSS   | old.woman |        |      |
- ‘The man bought rice in the old woman’s store’ (Lit: ‘The old woman’s store is the place of the man’s rice buying.’)

Here both the predicate and the Topic are marked with the Location-Forming Suffix, and the man and the rice, as non-topical arguments, are linked to the predicate in a possessive construction using the “ng possessive linker phrase”. The predicate is also marked with an infix as realis-perfective and as having a non-actor Topic.

### 3.2. The Topic in Tagalog

The Topic in Tagalog also differs from the general concept of topic in Indo-European languages, in that in Tagalog there is a much freer range of possibilities of what can appear as Topic. Generally, almost any referent associated with the situation in some way, whether a core or peripheral argument semantically or even a very indirectly affected referent, can be the Topic of the clause. It is usually one that is identifiable to the hearer, but not always (Adams and Manaster-Ramer 1988). This is much freer than in, for example English, where the possibilities are limited to just two direct arguments and adjuncts, and so very often what is the Topic in Tagalog, what the clause clearly is about, will not be even a notional topic in the English translation, as in the following example (using the Location Topic Construction—compare the use of the locative expression for the one affected in the English translation):

- (7) *Huwag mong ubusan ng gasoline si Ricky.*
- |        |             |            |               |      |      |        |     |
|--------|-------------|------------|---------------|------|------|--------|-----|
| [huwag | mo=ng       | ubusan     | ng gasoline]  | PRED | [si  | Ricky] | TOP |
| NEGIMP | 2sgPOSS=LNK | finish-LFS | POSS gasoline |      | SPEC | PN     |     |
- ‘Don’t use up all the gasoline on Ricky.’

That the Topic is what the clause is about can be seen from the fact mentioned in the quote given in footnote 6. Here again, the *na/ng* linker linking the two parts of the predicate (*huwag* and *ubusan*) appears on the possessive pronoun (*mo*) referring to the actor of the clause, which appears in second position because it is a second-position clitic (it is actually modifying *huwag na ubusan* as a whole), and the gasoline is linked to the predicate in a “ng possessive phrase”. So there is a complex structure of linkage<sup>10</sup>, with the overall predicate being a “ng possessive phrase”,<sup>11</sup> and within that there is one other possessive phrase (with *mo*) and one “*na/ng* linker phrase”. The Topic here takes the form *si* instead of *ang* or *‘yung* because it is a proper name.

### 3.3. Derivational Marking Related to Participants

The different choices of Topic are not active–passive but simply different ways of profiling the event (Foley and Van Valin 1984, §4.3); there is no change in transitivity or markedness, and there is no “basic” form; all are derived, though frequency counts have shown a strong preference for the use of Undergoer Topic constructions (Adricula 2023), though it depends on genre (Longacre 1996). This system is similar to the choice of A construction vs. O construction in Jarawara (Dixon 2000, 2004) depending on what is considered the topic of the clause, though Tagalog allows for many more choices for the Topic than Jarawara: actor Topic (8), patient Topic (9), conveyance Topic (a theme, a benefactive, or an instrument) (10), or locative Topic (a goal, source, or stative (essive) locative argument) (11).

- (8) *Bumili kanalang ng bago. Kamahal pa ng 2ndhand bikes ngayon.*  
 [b<um>ili <ka>TOP na lang ng bago]PRED  
 <AT>buying 2sgT CSM only POSS new  
 ‘Buy a new one. Secondhand bikes are still expensive now.’  
 ([https://www.reddit.com/r/RedditPHCyclingClub/comments/12x1rbh/gusto\\_ko\\_sanang\\_bumili\\_ng\\_mountainpeak\\_or\\_devel/](https://www.reddit.com/r/RedditPHCyclingClub/comments/12x1rbh/gusto_ko_sanang_bumili_ng_mountainpeak_or_devel/), accessed on 25 April 2023)
- (9) *Saan mo binili ang mga punlang itatanim natin?*  
 [Saan mo b<in>ili]PRED [ang mga punla=ng i-ta-tanim natin]TOP  
 where 2sgPOSS <RNAT>buying SPEC pl seedling=LNK CT-REDUP-plant 1plinclPOSS  
 ‘Where did you buy the seedlings we will plant?’  
 (<https://quizizz.com/admin/quiz/5dd3b41a5d5cc0001c874d9c/panghalip>, accessed on 30 April 2023)
- (10) *ang mapagbibilhan anyang nakatawa ay ibibili ko ng hikaw at singsing.*  
 [i-bi-bili ko ng hikaw at singsing]  
 CT-REDUP-buying 1sgPOSS POSS earrings and ring  
 ‘(with the money from) the ones that can be bought I will buy earrings and a ring.’  
 (<https://www.tagaloglang.com/ang-buto-ng-atis/>, accessed on 30 April 2023)
- (11) *Binilhan ako ni Mama ng cake!*  
 [b<in>ili-han <ako>TOP ni Mama ng cake]PRED  
 <RNAT>buying-LFS 1sgT POSS mom POSS cake  
 ‘Mom bought me a cake!’ (<https://www.youtube.com/watch?v=6J7LSxbFCA4>, accessed on 25 April 2023)

These are sometimes talked about as different symmetrical voices, but they are actually derivations, not inflections (Himmelmann 2004, 2005, 2008), and some of the same affixes often appear on referential phrases involving object words, not just on predicates or action words, e.g., *basura* ‘garbage’, *basurahan* ‘garbage bin, garbage truck’, using the same Locative-Forming Suffix (LFS) that we saw in the predicates of examples (6), (7), and (11). In most cases they change the meaning and use of the form, e.g., *lakas* ‘strong’ > *lakasan* (with LFS added and with syncopated vowel) ‘strengthen’, *langgam* ‘ant’ > *langamin* ‘be infested with ants’ (with patient-undergoer suffix; Himmelmann 2004, p. 1480). As Himmelmann points out (Himmelmann 2004, p. 1480), there is no formal evidence to support the common view that adding the affixes to action words is inflection while adding them to object words is derivation. They are all derivation (see also Rubino 1998). Himmelmann states (Himmelmann 2004, p. 1481), “...there are no productive inflectional paradigms for voice, as suggested by the commonly used ‘paradigmatic’ examples in the literature. Instead, derivations from all kinds of bases are only partially predictable on the basis of their semantics and exhibit a large number of idiosyncrasies, which again suggests derivation rather than inflection.”

Similar to the generalization given in footnote 6, Himmelmann (2004, p. 1481) argues that the affixes “change the orientation of a given base in such a way that it may be used to refer to one of the participants involved in the state of affairs denoted by the base ... In this view, *-um-* is an actor orienting infix which derives from a base such as *tango* ‘nod, nodding in assent’ a word *tumango* which could be glossed as ‘one who nods, nodder’. This expression no longer directly denotes the action of nodding, but rather the participant who nods. That is, in the Tagalog clause ... *tumango ang unngo* ‘The monkey nodded in assent’, both *tumango* and *unngo* refer to the same entity. Imitating the equational structure of this clause it could be rendered as ‘nodd-er in assent (was) the monkey’ ... Note, however, that Tagalog voice affixes are not nominalising in a morphosyntactic sense, since they do not change the syntactic category of the base...”

The so-called actor voice infix *-um-* can also be used when there is no actor, as with certain natural processes, such as *umulan* ‘rain (falling)’ (< *ulan* ‘rain’) and with certain processes, such as *pumuti* ‘become white, bleached’ (< *puti* ‘white’). Due to these problems and also to the fact that they are derivational affixes, they cannot be seen as agreement affixes, as is common in the literature on Tagalog.

In Tagalog there is no neutralization we can call S, nor even neutralization of semantic roles that would form a single grammatical category of Actor or Undergoer, in terms

of the derivations, as different types of actor and undergoer (e.g., with different degrees of intention, agentivity, transitivity, and/or affectedness) involve different derivations depending on the nature of the action, the Topic of the clause, and sometimes the nature of the affected participant. For example, in an intransitive clause, different events involving the same root can take different affixes, e.g., given the root *dulas*, *madulas* can be used for ‘slip (unintentionally)’ and *dumulas* can be used for ‘slide (i.e., slip intentionally)’,<sup>12</sup> and the marking of an intransitive actor can be different from that of a transitive actor, even with the same root, e.g., *labas* ‘outside’: *lumabas* ‘(the one who) comes/goes out’ vs. *maglabas* ‘(the one who) brings/takes out’. In one case the difference in the affix represents a difference in the direction of action: *bumili* ‘(one who) buys’, *magbili* ‘(the one who) sells’. The prefix *mag-* can also be used to express greater frequency or greater intensity, e.g., *bumasa* ‘read’, *magbasa* ‘to read a lot/study’, and they can even be used together for even greater intensity: *mag-um-aral* ‘study diligently’ (examples adapted from [Himmelmänn 2005](#), p. 365).

And in some cases of the use of the *maN-* actor marking orientation prefix, the difference is in whether the action is directed at a single undergoer or distributed over several undergoers, e.g., *pili* ‘choice’, *mamili* (*maN* “actor” prefix + *pili*) ‘(one who) chooses among several items’; *takot* ‘fear’, *manakot* (*maN* “actor” prefix + *takot*) ‘(one who) frightens several people’ ([Himmelmänn 2004](#), p. 1476). Possibly related to this use of *maN-* is its use in creating words that refer to people who frequently carry out particular actions, such as professionals (with reduplication of the first syllable of the root), e.g., *mangbabasa* ‘reader’ < *basa* ‘reading’, *manggagamot* ‘physician’ < *gamot* ‘medicine, cure’ (examples from [Himmelmänn 2005](#), p. 373). From all this it is clear these are not agreement inflections.

There are also three different types of undergoer marking, which can be used to represent a range of semantic types.

The facts given above have not stopped many people from trying to force Tagalog into a nominative–accusative or ergative–absolutive alignment pattern by manipulating made-up examples to make the pattern look consistent, and by calling possessive pronouns ergative, but as mentioned above, the different constructions for each Topic type are not either of these assumed alignments, but are a different sort of structure altogether, which we can call the Philippine type. With each different construction, as we saw in the examples above, one participant is singled out as the Topic, and the non-topical participants appear either in a possessive phrase as part of the predicate, if they are direct arguments, or in an adjunct phrase, if they are oblique arguments. This is not at all the sort of standard alignment system assumed in the quotes given above.

#### 4. Conclusions

This short paper is just to make a simple point: we should not assume all languages work the same way, whether it is in terms of alignment, marking, or any other aspect of the language. We should be open to the diversity of possibilities, and document them faithfully when we find them, working inductively through natural language data, and not force non-Indo-European languages into an Indo-European mold. This holds for writing reference grammars and also for teaching the language. There is no need to force concepts like “noun”, “noun phrase”, “subject”, etc. on the language if they are not appropriate. Readers and students will be much better off if the language is described in a way that reflects the way people actually speak the language.

**Funding:** This research received no external funding.

**Data Availability Statement:** The data used are either from publicly available websites or from the book *Alamat ng gubat*, by Bob Ong, Visprint, Inc., Manila, 2003.

**Conflicts of Interest:** The author declares no conflict of interest.

## Abbreviations

Abbreviation	Meaning	Position	Form
1plinclPOSS	first-person inclusive possessive (and non-Topic) pronoun	following head (predicator or reference phrase)	<i>natin</i>
1sgPOSS	first-person singular possessive pronoun	second-position clitic	<i>ko</i>
1sgT	first-person Topic pronoun	second-position clitic	<i>ako</i>
2sgPOSS	second person singular possessive pronoun	second-position clitic	<i>mo</i>
2sgT	second-person Topic pronoun	second-position clitic	<i>ka</i>
3sgT	third-person Topic pronoun	second-position clitic	<i>siya</i>
AT	actor Topic	derivational infix after first consonant of root	<i>-um-</i>
CSM	Change-of-state aspect marker	second-position clitic	<i>na</i>
CT	Conveyance Topic	derivational prefix on root	<i>i-</i>
LFS	Location-Forming Suffix (forms elements that represent locations); when the word with this suffix is the predicate, the Topic of the clause is a location (“locative focus”)	derivational suffix on root	<i>-an ~ -han</i>
LNK	Linker	clitic (occurs on first item) ~ particle (occurs between two items linked)	<i>-ng ~ na</i>
LOC	General peripheral argument marker	when not in focus, generally occurs at the end of the clause followed by a reference phrase	<i>sa</i>
LP	Locative phrase	adjunct in which oblique arguments appear, <i>sa</i> + reference phrase	
N.EXIST	Negative existential	predicator	<i>wala?</i>
NEGIMP	Negative imperative	appears in utterance-initial position	<i>huwag</i>
pl	plural	appears optionally before reference phrase with plural referent	<i>mga</i> [manga]
POSS	Possessive marker	particle (occurs between two items linked; <i>ni</i> is used to link a proper name)	<i>ng</i> [nan], <i>ni</i>
PN	Personal name		
PRED	Predicate		
PROXT	Proximate Topic deictic pronoun		<i>ito</i>
Q	Question-marking particle	second-position clitic	<i>ba</i>
REDUP	Reduplication	usually first syllable of root to mark imperfective or planned event	
RNAT	Realis non-actor Topic	derivational infix appearing after initial consonant of predicate or before vowel initial	<i>-in-</i>
SPEC	Marker of specific referent	occurs before reference phrase	<i>ang ~ si ~ 'yung</i>
TOP	Topic		
UT	undergoer-Topic marker	derivational suffix	<i>-in</i>

## Notes

- 1 “Explaining the cross-linguistic distribution of argument-coding patterns” (EDAP2023), Universität Potsdam Campus Am Neuen Palais, 21–23 March 2023; <https://sites.google.com/view/edap2023/home> (accessed on 13 October 2022).
- 2 Some scholars have talked about this as a neutral alignment, but it is not a type of alignment, it is a complete lack not only of alignment, but also of any system for identifying the roles of the main participants in discourse, which is what the function of relational marking and alignment is. This also includes also not having grammaticalized constructions that manifest word-order constraints or any other restricted neutralizations of semantic roles and pragmatic functions for syntactic purposes (i.e., what is normally talked about as grammatical relations, e.g., “subject” and “direct object”; LaPolla 1993, 2023; Van Valin and LaPolla 1997, chap. 6).
- 3 But see Himmelmann (2008) for arguments against a precategorical analysis and for arguments for morphological word classes that do not correspond with English noun and verb. See also LaPolla (2010) for a discussion of Himmelmann’s approach in Chinese.
- 4 See footnote 5 for the definition of Topic in Tagalog. The other core arguments in the clause are non-topical arguments, which appear as possessive pronouns or in possessive phrases linked by *ng* [nan]. There is no structural justification for distinguishing between the use of the possessor marking for arguments modifying the predicator (and forming a possessive phrase with it that constitutes the predicate of the clause) and for modifiers of the head of referential phrases and forming possessive phrases with them. For this reason both are glossed POSS.

- 5 Following best practice in typology, for language-specific (descriptive) categories and constructions I will capitalize the initial letters of the name of the category or construction, but for comparative concepts I will not capitalize the first letter. So, for example “Actor” refers to the language-specific grammatical category manifesting a particular neutralization of semantic roles in the language under discussion, while “actor” refers to the comparative concept of the one who performs an action. As there are no universal or cross-linguistic grammatical categories, descriptive and comparative concepts need to be kept distinct. In the case of Topic, it is a Tagalog-specific grammatical status, as it is an argument given special morphosyntactic treatment, as well as a pragmatic status, as it is what the clause is about (cf. Lambrecht (1994) on topic, what the clause is about—and usually part of the presupposed information—vs. focus, the information evoked by an assertion that cannot be supplied by the addressee). The Topic can appear as a second-position clitic topical pronoun (there are different sets of pronouns for topical vs. non-topical referents, the latter being the possessive pronouns), or as a reference phrase at the end of the clause if it is not focal, marked by a demonstrative plus linker, usually *ang* or *‘yung*, or at the beginning of the clause if it is focal. That is, the Topic is not identified by its position in the clause, but by its marking, unlike, for example, in Chinese, where being preverbal is enough to identify an element as a Topic. The Tagalog Topic is often referred to as “subject” in much of the literature, but the question of grammatical relations in Tagalog is quite controversial (Schachter 1976, 1977; Schachter and Otones 1972; Naylor 1980; Foley and Van Valin 1984, §4.3), and I see no structural justification in Tagalog for using that term, or “direct object”, or “noun phrase” or “verb phrase”.
- 6 “[...]ny predication minus its topic can function as a nominalization understood to denote what would be the topic of that predication” (Adams and Manaster-Ramer 1988, p. 81).
- 7 Four lines are used in the examples, with the first line as spoken and the morpheme analysis below that, because of the infixes and as there are sometimes morphophonemic sound changes that appear in the first line, but do not appear in the morpheme analysis, as in this example, where *kuhin-in* becomes *kunin*, and *‘yung* becomes *yun* before *leaves* due to assimilation, and as in example (6), with the dropping of *-i* from the root *bili* ‘buy’ when the locative forming suffix *-han* is suffixed to the root (which is also modified by the infix *-in-*), resulting in *binilhan*.
- 8 In the “*na/=ng* linker phrase”, *na* is used when the word before it ends in a consonant other than *-n*, *=ng* is used for open syllables (as in *ba=ng* in the example above) and syllables ending in *-n*. Below we will see the possessive linker is also written *ng*, but it is pronounced [nɑŋ] in this case and is a stand-alone word. We will call that the “*ng* possessive linker phrase”.
- 9 As mentioned below, the Location-Forming Suffix can mark a range of argument types. The structure is the same, but the interpretation is different. For example, compare (6) with *Binilhan ng lalaki ng bigas sa palengke ang lola*. ‘The man’s buying of the rice in the market was (for) the old woman.’ The only real difference is the understanding of the Topic referent as a location vs. as a human, leading to a benefactive understanding of the clause in the case of the latter.
- 10 The linking structures can be quite complex, with multiple embeddings and overlaps, at least in written Tagalog. See LaPolla 2014 for examples from one published text.
- 11 This is the general case for predicates. As it is a linker phrase, and *ng* is not case marking or a preposition, the two elements of the phrase cannot be separated, e.g., the element following *ng* cannot appear in clause-initial position the way Topics and *sa* peripheral phrases can when they are in focus.
- 12 This is not the sort of “split S” found in some languages, such as Acehnese (Durie 1985, 1987), as the affixes are not pronominal agreement; they are derivational affixes that change the nature of the event rather than select an undergoer vs. actor argument.

## References

- Adams, Karen L., and Alexis Manaster-Ramer. 1988. Some Questions of topic/focus choice in Tagalog. *Oceanic Linguistics* 27: 79–101. [CrossRef]
- Adracula, Norielle. 2023. Examining voice choice in Tagalog: A corpus of web-based Tagalog. *Proceedings of the Linguistic Society of America* 8: 5533. [CrossRef]
- Chao, Yuen Ren. 1967. *Mandarin Primer: An Intensive Course in Spoken Chinese*. Cambridge: Harvard University Press. First published 1948.
- Chao, Yuen Ren. 1968. *A Grammar of Spoken Chinese*. Berkeley and Los Angeles: University of California Press.
- Chao, Yuen Ren. 1976. Notes on Chinese grammar and logic. In *Aspects of Chinese Sociolinguistics: Essays by Yuen Ren Chao*. Edited by Anwar S. Dil. Stanford: Stanford University Press, pp. 237–49. First published 1955.
- Chao, Yuen Ren. 1976. How Chinese logic operates. In *Aspects of Chinese Sociolinguistics: Essays by Yuen Ren Chao*. Edited by Anwar S. Dil. Stanford: Stanford University Press, pp. 250–59. First published 1959.
- Dixon, R. M. W. 2000. A-constructions and O-constructions in Jawawara. *International Journal of American Linguistics* 66: 22–56. [Cross-Ref]
- Dixon, R. M. W. 2004. *The Jarawara Language of Southern Amazonia*. Oxford: Oxford University Press.
- Durie, Mark. 1985. *A Grammar of Acehnese*. Dordrecht: Foris.
- Durie, Mark. 1987. Grammatical relations in Acehnese. *Studies in Language* 11: 365–99. [CrossRef]
- Foley, William A., and Robert D. Van Valin, Jr. 1984. *Functional Syntax and Universal Grammar*. Cambridge: Cambridge University Press.

- Himmelman, Nikolaus P. 2004. Tagalog (Austronesian). In *Morphology. An International Handbook on Inflection and Word Formation*. Edited by Geert Booij, Christian Lehmann, Joachim Mugdan and Stavros Skopeteas. Berlin: de Gruyter, pp. 1473–90.
- Himmelman, Nikolaus P. 2005. Tagalog. In *The Austronesian Languages of Asia and Madagascar*. Edited by Karl Alexander Adelaar and Nikolaus P. Himmelman. London: Routledge, pp. 350–76.
- Himmelman, Nikolaus P. 2008. Lexical categories and voice in Tagalog. In *Voice and Grammatical Relations in Austronesian Languages*. Edited by Peter K. Austin and Simon Musgrave. Stanford: CSLI Publications, pp. 247–93.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.
- LaPolla, Randy J. 1990. Grammatical Relations in Chinese: Synchronic and Diachronic Considerations. Ph.D. dissertation, University of California, Berkeley, CA, USA.
- LaPolla, Randy J. 1992. Anti-ergative marking in Tibeto-Burman. *Linguistics of the Tibeto-Burman Area* 15: 1–9.
- LaPolla, Randy J. 1993. Arguments against ‘subject’ and ‘direct object’ as viable concepts in Chinese. *Bulletin of the Institute of History and Philology* 63: 759–813.
- LaPolla, Randy J. 1995. Ergative marking in Tibeto-Burman. In *New Horizons in Tibeto-Burman Morpho-Syntax*. Edited by Yoshio Nishi, James A. Matisoff and Yasuhiko Nagano. Senri Ethnological Studies 41. Osaka: National Museum of Ethnology, pp. 189–228.
- LaPolla, Randy J. 2009. Chinese as a Topic-Comment (not Topic-Prominent and not SVO) language. In *Studies of Chinese Linguistics: Functional Approaches*. Edited by Janet Xing. Hong Kong: Hong Kong University Press, pp. 9–22.
- LaPolla, Randy J. 罗仁地. 2010. Feilübin Tagaluo yu (Tagalog) de cilei fanchou (The lexical categories of the Tagalog language of the Philippines). In *Yuyanxue Luncong*. Beijing: Peking University, vol. 41, pp. 1–14.
- LaPolla, Randy J. 2014. Constituent structure in a Tagalog text. *Language and Linguistics* 15: 761–74. [\[CrossRef\]](#)
- LaPolla, Randy J. 2019. Arguments for seeing Theme-Rheme and Topic-Comment as separate functional structures. In *Systemic Functional Language Description: Making Meaning Matter*. Edited by James R. Martin, Yaegan J. Doran and Giacomo Figueredo. London: Routledge, pp. 162–86.
- LaPolla, Randy J. 2023. Grammatical Relations. In *Cambridge Handbook of Role and Reference Grammar*. Edited by Delia Bentley, Ricardo Mairal Usón, Wataru Nakamura and Robert D. Van Valin, Jr. Cambridge: Cambridge University Press, pp. 269–29.
- LaPolla, Randy J., and Dory Poa. 2006. On describing word order. In *Catching Language: The Standing Challenge of Grammar Writing*. Edited by Felix Ameka, Alan Dench and Nicholas Evans. Berlin: Mouton de Gruyter, pp. 269–95.
- Longacre, Robert E. 1996. Focus in Philippine languages and some similar structures in English. In *Pan Asiatic linguistics: The 4th International Symposium on Languages and Linguistics, Mahidol University, Salaya, Thailand, 8–10 January 1996*. Bangkok: Institute of Language and Culture for Rural Development, Mahidol University at Salaya, pp. 30–34.
- Lü, Shuxiang. 1979. *Hanyu Yufa Fenxi Wenti (Questions in the Analysis of Chinese Grammar)*. Beijing: Commercial Press.
- Naylor, Paz Buenaventura. 1980. Linking, relation-marking, and Tagalog syntax. In *Austronesian syntax. Papers from the Second Eastern Conference on Austronesian Languages*. Edited by Paz Buenaventura Naylor. Ann Arbor: Center for South and Southeast Asian Studies, pp. 33–49.
- Naylor, Paz Buenaventura. 2005. On the stative predicate: Tagalog “existentials” revisited. In *Current Issues in Philippine Linguistics and Anthropology Parangal Kay Lawrence A. Reid*. Edited by Hsiu-chuan Liao and Carl R. Galvez Rubino. Manila: The Linguistic Society of the Philippines and SIL Philippines.
- Rubino, Carlos R. G. 1998. The morphological realization and production of a nonprototypical morpheme: The Tagalog derivational clitic. *Linguistics* 36: 1147–66.
- Schachter, Paul. 1976. The subject in Philippine languages: Actor, topic, actor-topic, or none of the above. In *Subject and Topic*. Edited by Charles N. Li. New York: Academic Press, pp. 491–518.
- Schachter, Paul. 1977. Reference-related and role-related properties of subjects. In *Syntax and Semantics, Vol. VIII: Grammatical Relations*. Edited by Peter Cole and Jerrold M. Sadock. New York: Academic Press, pp. 279–306.
- Schachter, Paul, and Fei T. Otnes. 1972. *A Tagalog Reference Grammar*. Berkeley and Los Angeles: University of California Press.
- Van Valin, Robert D., Jr., and Randy J. LaPolla. 1997. *Syntax: Structure, Meaning, and Function*. Cambridge: Cambridge University Press.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.