

Symmetrical Voice Systems and Precategoriality in Philippine

Languages

William A. Foley

University of Sydney

1.0 The Philippine Voice System

The analysis of voice in Philippine languages, typically called the ‘focus system’ (e.g. Ramos 1974; Ramos and Bautiste 1986) has been a source of contention for nearly a hundred years (Blake 1906), and this shows no sign of letting up (Schachter 1976; McGinn 1988; Gerdts 1988; Shibatani 1988; Guilfoyle, Hung and Travis 1992; Kroeger 1993). The present paper is yet another contribution to this ongoing debate, but endeavors to enlarge its scope in two ways: first, by linking the peculiar voice properties of these languages to other unusual typological properties, specifically the problematic nature of a distinction between root classes of noun and verb, and secondly, by demonstrating that this cluster of typological properties may not be unique to Philippine languages, but are also found in some Amerindian languages of the Northwest Coast of North America, particularly Wakashan languages. The wider perspective made available through the comparison with the Wakashan languages allows us new opportunities to think about Philippine languages and I will develop these insights into a coherent theoretical account of their typology.

The nature of the basic facts which underlie all this debate can be seen in the following sentences from Tagalog (the sentences are slightly unnatural in

that Philippine languages, like most languages, do have a strong preference for no more than one full NP per clause, but they are grammatical and do illustrate the nature of the voice system well):

- (1) (a) *b-um-ili* *ng* *isda* *sa* *tindahan* *ang* *lalake*
 VC-buy CORE fish OBL store man

‘The man bought fish in the store.’

- (b) *bi-bilh-in* *ng* *lalake* *sa* *tindahan* *ang* *isda*
 IRR-buy-VC CORE man OBL store fish

‘The man will buy the fish in the store.’

- (c) *bi-bilh-an* *ng* *lalake* *ng* *isda* *ang* *tindahan*
 IRR-buy-VC CORE man CORE fish store

‘The man will buy fish in the store.’

- (d) *ipam-bi-bili* *ng* *lalake* *ng* *isda* *ang* *salapi*
 VC-IRR-buy CORE man CORE fish money

‘The man will buy fish with the money.’

- (e) *i-bi-bili* *ng* *lalake* *ng* *isda* *ang* *bata*
 VC-IRR-buy CORE man CORE fish child

‘The man will buy fish for the child.’

Note that the root of the verb *bili* ‘buy’ (reduplicated in (b)-(e) to *bibili* to indicate irrealis modality, i.e. ‘will buy’) takes a range of different affixes: *-um-*, *-in-*, *-an*, *ipaN-* and *i-*. The choice of each affix is determined by the semantic properties of whichever NP is preceded by *ang*. If the *ang* NP is causer, performer or initiator of the event (let’s abbreviate this as the actor), in this cause the buyer, the verb takes the infix *-um-* (1a). On the other hand, if the *ang* NP is the participant affected or undergoing a change in location or ownership as a result of the event (what I will call the undergoer (Foley and Van Valin (1984))), in these examples the thing bought, the verb takes the suffix

-*in* (1b). Other possibilities are also available: the *ang* NP can denote the place whether the event is accomplished (the suffix *-an*; (1c)), the instrument used to accomplish the event (the prefix *ipaN-*; (1d)) or the person who benefits from the event (the prefix *i-*; (1e)). In each case it is the affix on the verb which marks the semantic role of the *ang* NP: hence the description by Blake nearly a hundred years ago (1906) of this phenomenon as the ‘expression of case by the verb’, rather than its usual locus on the noun or in the noun phrase.

Blake’s description, while very insightful, is, however, not complete. In addition to the verbal affixes, Tagalog does possess true case markers, as is clear from the glosses given to the prepositions *ng* and *sa* above. But the Tagalog case system is really a very simple binary one, contrasting CORE NPs, those prototypically subcategorised by verbs like actors and undergoers, with OBLique NPs, not so subcategorized, such as locations, beneficiaries, goals, etc. NPs are divided into [\pm proper] and case marked as follows:

| | | |
|-----------|-----------|-----------|
| (2) | CORE | OBL |
| [+proper] | ni Juan | kay Juan |
| [-proper] | ng lalake | sa lalake |

PROnominal NPs are inflected for case:

| | | CORE | OBL |
|----|----------|-------|--------|
| | 1 | ko | akin |
| SG | 2 | mo | iyó |
| | 3 | niya | kaniya |
| | 1excl | natin | atin |
| PL | 1include | namin | amin |
| | 2 | ninyo | inyo |
| | 3 | nila | kanila |

| | | |
|-----------|------------|-------|
| PROX | nito | dito |
| NR DISTAL | niyan | diyan |
| FR DISTAL | niyon/noon | doon |

The CORE pronominals never occur with the preposition *ng* or *ni*, but the OBLique pronominals when functioning as clause level constituents typically do.

If the CORE/OBLique distinction, *ng/sa* etc, exhausts the NP case marking distinction, then what exactly is the function of *ang*? Note that it too, has its [+proper] NP variant and a full set of corresponding pronouns (i.e. those that can replace a *ang* NP and syntactically function identically):

- (3) NP [+proper] si Juan
 [-proper] ang lalake

PROnominals

| | | |
|-----------|--------|---------|
| | 1 | ako |
| SG | 2 | ka/ikaw |
| | 3 | siya |
| | 1 excl | tayo |
| DL | 1 incl | kami |
| | 2 | kayo |
| | 3 | silá |
| PROX | | ito |
| NR DISTAL | | iyan |
| FR DISTAL | | iyon |

The behavior here certainly looks like case marking parallel to the pattern in (2) and familiar from case marking systems from languages around the world. This had led some analysts to dub the *ang* forms as the NOMinative case,

noting the fact that it is these NOMinative forms which show agreement with the verb via the affixation of *-um-*, *-in-*, etc, and, crosslinguistically, it is the NOMinative case forms in languages which prototypically show agreement with the verb, as in English:

(4) NOMinative ACCusative

| | | | |
|---|----|------|------|
| 1 | SG | I | me |
| | PL | we | us |
| 3 | SG | she | her |
| | PL | they | them |

(a) I(SG) *am* (SG) going to visit them (PL)

(b) *I (SG) *are* (PL) going to visit them (PL)

(c) We (PL) *are* (PL) going to visit her (SG)

(d) *We (PL) *is* (SG) going to visit her (SG)

This analysis would then suggest a three way nominal case system for Tagalog: CORE-[+NOM], CORE-[-NOM], and OBLique.

2.0 What the Philippine voice system is not

2.1 Arguments against the active-passive analysis

While there are good reasons to question whether *ang* is a case marker at all, lets follow where the analysis leads us. If *ang* in the examples in (1) really represents a NOMinative case, then the contrast between (1a) and (1b) looks very much like an active-passive alternation in English: the actor is marked as NOMinative by *ang* in (1a), while the undergoer is marked as NOMinative in (1b). But there are two immediate problems with this analysis. First, the actor of the so-called passive sentence in (1b) remains a CORE NP marked by *ng* in marked contrast to the OBLiquely marked actors of canonical passives crosslinguistically (e.g. the preposition *by* in English). Second, again contrary

to canonical passives, there is no unmarked active form, but both the putative active and passive voices are indicated by verbal morphemes: *-um-* for active and *-in* for passive. This is highly unusual; normally the active voice is morphologically and syntactically unmarked, the basic inflectional form of the verb, while the passive is marked; compare English *buy* versus *is bought*. Finally, examples (1c) through (1e) present still further problems for the analysis. In each of these cases it is erstwhile OBLique NPs, locatives, instrumentals or benefactives, arguments not subcategorized by the verb, which have assumed NOMinative case and the verb appears with specific affixes indicating their semantic function. Again, these are non-actor NPs in NOMinative case, so following the above analysis these need to be analyzed as passives. This would require postulating that Tagalog has four passive voices in contrast to English's one (actually rather more than that because there are a number of other voice types which also allow non-subcategorized arguments to assume NOMinative case not illustrated here; see Schachter and Oanes (1972)). This is typologically a highly marked situation. It is true that many languages (e.g. Bantu) allow non-subcategorized arguments to assume NOMinative case in passives, but they do this first via applicativization which makes the non-subcategorized argument a CORE NP of the verb and then by passivization which allows it to assume NOMinative case, as in this Indonesian example:

- (5) sawah itu di-tanam-i padi oleh Hasan
 rice fields DET PASS-plant-APPL rice by PN

‘The rice field was planted with rice by Hasan.’

First, the APPLicative suffix *-i* allows the non-subcategorized erstwhile locative argument *sawah* ‘rice field’ to assume CORE NP status and then the prefix *di-*PASS allows it to assume NOMinative case, realized by preverbal position in Indonesian. Note the actor in the passive follows the verb and is marked with

the OBLique agentive preposition *oleh*. This is crosslinguistically the expected pattern, but the Tagalog situation is fundamentally different. For example in (1c) there is only the voice affix *-an*. This would need to be analyzed as doing double duty, both applicativization and passivization. But outside of Philippine languages, languages simply don't do this, and this typological fact must force us to question seriously whether the passive analysis is really the right one for these languages.

It is important to note that this unusual typological picture is not restricted just to the geographical area of the Philippines, but is widespread among the Western Austronesian languages from Taiwan to Western Indonesia and is diagnostic of what is known as 'Philippine-type' languages among Austronesian specialists. There are good reasons to suspect that proto-Austronesian was itself a language of this type (Wolff 1973, 1979; Starosta, Pawley and Reid 1982). Atayal (Egerod 1965, 1966; Rau 1992; Huang 1993) of Taiwan and Kimaragang (Kroeger 1988, 1996) both illustrate the same structural patterns of voice as Tagalog:

(6) Atayal (Huang 1993)

(a) *t-m-tu* *tali* *ghuniq*
 VC-crush PN tree
 'The tree crushed Tali.'

(b) *tʔ-un* *ghuniqu* *tali*
 crush-VC tree PN
 'The tree crushed Tali.'

(c) *byiq-an-mu* *hiraʔ* *gulih* *tali*
 give-VC-1SG CORE yesterday fish PN
 'I gave a fish to Tali yesterday.'

- (d) *s-ʔagan- mu gulih sgariʔ gani*
VC-take 1SG CORE fish net this
‘I caught the fish with this net.’
- (e) *s-ʔagan-mu gulih Tali*
VC-take-1SG CORE fish PIN
‘I’ll catch a fish for Tali.’

In this dialect of Atayal there are no obligatory NP case marking prepositions (although other dialects do have them), so that word order is much more rigidly fixed than in Tagalog. The actor always immediately follows the verb (if it is a pronoun it is encliticized as in (6 c-e)), unless it is NOMinative, in which case, it is clause final (6a). The verb is affixed with the usual voice affixes, all of which are in fact cognate with their Tagalog equivalents (*-um-*: *-m-*: *-in-*; *-un* < PAN*-«n; *-an*: *-an*; *i-*: *s-* < PAN **Si-*). The same basic patterns are found: (1) subcategorized arguments remain CORE NPs (i.e. ϕ case marked) in all voice forms of the verb; (2) there are no unmarked verbal forms, i.e. an active form; and (3) nonsubcategorized arguments can freely assume NOMinative case without going through an APPLicative derivation first via the affixes *-an* or *s-*. While it might be objected that the verb ‘give’ does subcategorize a recipient argument, as in (6c) such an analysis seems wrong for Atayal and Tagalog, as may be the case for many Austronesian languages. Returning to (6c), if *Tali* were not NOMinative, it could occur with an overt case marker and that case would be the OBLique *te*, not the CORE *naʔ*, indicating that the recipient argument is probably not subcategorized. Also, there are other examples of verbs affixed with *-an*, in which the locative argument could not be plausibly argued to be subcategorized:

- (7) *nanu*[?] [?]*by-an* *tali*
 what sleep-VC PN
 ‘Where does Tali want to sleep?’

(8) Kimaragang (Kroeger 1988)

- (a) *momoli* *okuh* *do* *tasin*
 m-poN-boli 1SG CORE [-DEF] salt
 VC-TR-buy
 ‘I am going to buy some salt.’

- (b) *amu* *kuh* *bili-on* *itih* *tasin* *ditih*
 NEG 1SG CORE buy-VC this salt this
 ‘I won’t buy this salt.’

- (c) *siongah* *pinomolian* *nuh* *dilo* *gampa* *nuh*
 where -in-poN-bili-an 2SG CORE that CORE machete 2SG POSS
 PAST-TR-buy-VC
 ‘Where did you buy your machete?’

- (d) *n-i-boli* *kuh* *it* *siin* *ku* *dot* *tasi*
 PAST-VC-buy 1SG CORE [+DEF] money 1SG POSS CORE [-DEF] salt
 ‘I bought salt with my money.’

- (e) *isai* *boli-an* *nuh* *ditih* *tubat* *ditih*
 who buy-VC 2SG CORE this medicine this
 ‘Who will you buy medicine for?’

Kimaragang, like Tagalog and unlike Atayal, has NP case markers and like Tagalog has a three way contrast: e.g. *it* CORE-NOM, *dit* CORE-NON-NOM and *sid* OBLique; or *okuh* 1SG CORE-NOM, *kuh* 1SG CORE-NON-NOM and *dogon* OBLique (the CORE case markers also distinguish [\pm DEF], *it* CORE-NOM [+DEF] versus *ot* CORE-NOM [-DEF], an important point I will have cause to return to later). Again, the general typological picture is like that of Tagalog and Atayal: (1) subcategorized arguments remain core in all voice

forms of the verb (the forms *kuh* and *nuh* in (7b-e); (2) all verbal forms are overtly marked for voice, i.e. the semantic role of the NOMinative NP, there being no unmarked verbal form corresponding to the English active voice and (3) non-subcategorized arguments again directly assume NOMinative through a single verbal affix, *-an* or *i-*, there being no intermediate APPLicative derivation. Again, the affixes involved are directly cognate with the Tagalog ones: *-um-*: *m-~um-*; *-in-*: *-on* < PAN **-<n*; *-an-*: *-an*; and *i-*: *i-*. The only seemingly significant difference between Kimaragang and the other two languages concerns the example in which the beneficiary is marked NOMinative (7e). In Tagalog and Atayal examples this participant type was marked with the prefix **Si-* (Tg *i-*, At *s-*), the same affix used for instruments, but in Kimaragang, it occurs with *-an*, the locative voice marker. But the beneficiary voice marker actually shows considerable variation among Philippine type languages (and within them!) between **-an* and **Si-*; even in Tagalog a number of stems will show beneficiary voice with *-an*.

These three languages, drawn from widely disparate areas, present a consistent typological picture of Philippine type languages, and one which seems quite incompatible with an active-passive analysis of these clausal alternations. Many other arguments against such an analysis have been presented by previous researchers, so I will not repeat them here, but for some of these see Schachter's justly famous papers on Tagalog grammatical relations (Schachter 1976, 1977).

2.2 Arguments against the ergative-antipassive analysis

In recent years, problems of this sort have led researchers to propose an alternative, an ergative-absolutive analysis of Philippine clausal structure (Cena 1977, De Guzman 1979, Gerdts 1988). In this analysis, the NPs which

were described as bearing NOMinative case in the active-passive analysis, are now claimed to carry ABSolute case (this is typologically plausible, because ABSolute NPs are the ones most likely to agree with the verb in ergative-absolutive languages) and the erstwhile non-NOMinative core NPs are now actually taken to be ERGatively marked. Example (1b) would now be analyzed as:

- (9) *bi-bilh-in* *ng* *lalake* *sa* *tindahan* *ang* *isda*
 IRR-buy-VC ERG man OBL store ABS fish
 ‘The man will buy the fish in the store.’

As this is now an ergative language, the form with the undergoer marked with *ang* would be the basic clause type, with it in ABSolute case another actor in ERGative case. The clause type in which the actor is in ABSolute case is now a derived form, the antipassive, a functional equivalent of passive in ergative languages: the ABSolute NP of (9) becomes ERGative and the ERGative NP, ABSolute:

- (10) *b-um-ili* *ng* *isda* *sa* *tindahan* *ang* *lalake*
 ANTI-buy ERG fish OBL store ABS man
 ‘The man bought fish in the store.’

The *-um-* is analyzed as an antipassive marker which derives an intransitive verb from the transitive *bilh-in* in (9). Note that *-um-* prototypically occurs with intransitive verbs:

- (11) *p-um-unta* *ang* *lalake*
 go ABS man
 ‘The man went.’

The ergative analysis has a great advantage over the active-passive one in that it neatly accounts for why non-subcategorized arguments can assume ABSolute status without a prior APPLicative derivation (examples 1c-1e). It is well known that true ergative languages have a global constraint against

double ABSolutive NPs in a clause. As a result whenever applicativization applies to give core NP status, specifically ABSolutive case, to a previously non-subcategorized oblique NP, a simultaneous antipassive applies to force the prior ABSolutive to assume OBLique case, as in these Chamorro examples (Cooreman 1988):

- (12) (a) *hu-tuge'* *i* *katta*
 1SG ERG-write ABS letter
 'I wrote the letter.'
- (b) *hu-tugi'-i* *i* *patgon* *ni* *katta*
 1SG ERG-write-APPL ABS child OBL letter
 'I wrote the child the letter.'

(12a) presents *tuge* 'write' as a simple transitive verb with two subcategorized arguments: *hu-* 1SG ERG and *i katta* ABS letter. (12b) presents a derived applicative verb form with *-i* APPL 'to/for'. The newly introduced CORE participant *i patgon* ABS child now assumes ABSolutive case, and due to the global blocking constraint against double ABSolutive NPs, this derivation now simultaneously forces the previous ABSolutive NP *i katta* to now appear as OBLique *ni katta*, an antipassive process. Because the whole process is one simultaneous derivation, there is no need for a separate antipassive morpheme: APPLicativization is antipassivization, pure and simple.

Adapting this to the Tagalog examples in (1c-e) and ceteribus paribus the Atayal and Kimaragang examples in (6c-e) and (8c-e), we would now claim that *-an* and *i-* are not voice markers at all, simply APPLicative affixes, and the case marking exhibited in the examples is simply the result of a process parallel to that described for Chamorro:

- (13) (a) bi-bilh-*in* ng lalake sa tindakan ang isda
 IRR-buy ERG man OBL store ABS fish
 ‘The man will buy the fish in the store.’
- (b) bi-bilh-*an* ng lalake ng isda ang tindakan
 IRR-buy-APPL ERG man ERG fish ABS store
 ‘The man will buy fish in the store.’

Here *-an* is simply an APPLicative marker which makes *tindakan* a non-actor CORE argument, hence case marked ABSolutive, and forces the prior ABSolutive NP *isda* into ERGative case, via the no double ABSolutive constraint.

As attractive as the ergative analysis might be, it too runs into serious descriptive problems. First, if examples like (10) are really antipassives they should exhibit the crosslinguistically well attested typological properties of antipassives. Unfortunately they do not. Comparing the verb forms of (10) and (11), we see that the antipassive and true intransitive verbs are morphologically identical. This is virtually never true of antipassives. While the verbs of antipassive *are* intransitive, they always carry morphological markers of derivation to indicate that they have been derived from transitive verbs, and these are always distinct from simply intransitive verb inflection, as in Dyirbal (Dixon 1972):

- (14) (a) Intransitive verb

 bayi yara bani- u
 DET.ABS man.ABS come-TNS
 ‘The man came.’

- (b) antipassive *derived* intransitive verb

 bayi yara bugun d3ugumbil-gu
 DET.ABS man.ABS DET.DAT woman-DAT

bural-*Na-* u

see-ANTI-TNS

‘The man saw the woman.’

Note the prototypical Dyirbal antipassive contains the obligatory derivational suffix *-Na(y)* which the intransitive verb necessarily lacks. (Some languages like Eskimo may have zero allomorphs of antipassive affixes, but these are *allomorphs* of morphemes overt in other environments, so they do not constitute a counterexample to the claim that derived antipassive verb forms are never identical with true underived intransitive verbs.) This is quite different from the situation in Tagalog and suggests that (10) is not a true antipassive construction.

Continuing along these lines, the prototypical effect of an antipassive is to delink the undergoer of the verb from the status of core argument and reassign it to an oblique function. Note above in (14b), the antipassive construction forces the seen participant to appear in the dative case, an oblique case in Dyirbal.

Now again consider Tagalog example (10), repeated here as (15):

| | | | | | | | |
|------|-------------------|-----|------|-----|----------|-----|--------|
| (15) | b- <i>um</i> -ili | ng | isda | sa | tindahan | ang | lalake |
| | ANTI-buy | ERG | fish | OBL | store | ABS | man |

‘The man bought fish in the store.’

If (15) is to be considered a true antipassive then *ng isda* must be construed as an OBLique constituent. On the face of it, the ERGative case marking might be taken as an argument against this, as ERGative case is otherwise the case of the CORE actor participant. But, in fact, in many languages ERGative case is homophonous with the OBLique INSTRumental case marker, and this could easily be the case here, especially given the fact that the proposition marking true instrumental NPs in Tagalog is indeed *ng*. It also true that *ng* marked undergoers are rather syntactically impotent in comparison to *ng* marked

actors, and this relative syntactic impotency could be taken as a further argument for their OBLique status and an antipassive analysis for (15). There is, however, one argument that suggests that *ng* marked undergoers are not OBLique constituents, but CORE NPs like *ng* marked actors. Tagalog allows true OBLique constituents to precede the verb in a topicalization construction:

- (16) sa tindahan bi-bilh-*in* ng lalake ang isda
 OBL store IRR-buy ERG man ABS. fish
 ‘In the store the man will buy the fish.’

This construction is completely prohibited for CORE NPs like ergatively marked actors:

- (17) *ng lalake bi-bilh-*in* sa tindahan ang isda
 ERG ma IRR-buy OBL store ABS fish

And it is equally prohibited for *ng* marked undergoers in putative antipassive constructions like (15):

- (18) *ng isda b-*um*-ili sa tindahan ang lalake
 ERG fish buy OBL store ABS man

This strongly argues that *ng* marked undergoers are still subcategorized arguments of the verb and not truly OBLique NPs as an antipassive analysis of (15) would require.

This constraint cannot be put down to a simple constraint against *ng* marked NPs in clause initial position. Some verbs in putative antipassive constructions idiosyncratically assign either ERGative or OBLique case to the non-actor argument, depending on its definiteness:

- (19) *um*-akyat ng/sa puno ang bata
 climb ERG/OBL tree ABS child
 ‘The child climbed a/the tree.’

However, regardless of whether the NP is marked with *ng* or *sa* it remains a CORE argument, for in neither case can it assume the preverbal topicalization position of true OBLique NPs:

(20) (a) *ng puno *um*-akyat ang bata
 ERG tree climb ABS child

(b) *sa puno *um*-akyat ang bata
 OBL tree climb ABS child

(Grammatical in the meaning ‘The child climbed *up* the tree’.)

This conclusively demonstrates that the *ng* marked NPs of constructions like (15) are not OBLique and therefore these constructions cannot be true antipassives.

While the evidence in Tagalog for the CORE status of *ng* marked NPs is somewhat indirect, in some other Philippine type languages it is much more transparent and hence persuasive. For example, in Kimaragang (Kroeger 1996), there is the normal pattern of multiple voices with different verbal affixes to mark voice (see examples in (8)). If the system is a true ERGative-ABSolute one, then again the form with *-m-* marking the actor as ABSolute would have to be taken as an antipassive. These putative antipassives, however, co-occur with an additional derivational suffix lacking in all other voice types:

(21) (a) mangalapak okuh do niyuw
 m-poN-lapak 1SG ABS ERG [-DEF] coconut
 VC-?-split

‘I will split a coconut.’

(b) lapak-*on* kuh it niyuw
 split-VC 1SG ERG ABS[+DEF] coconut

‘I will split the coconut.’

| | | | | | | |
|-----|------------------|---------|-----------|---------|-----------|-------|
| (c) | lapak- <i>an</i> | kuh | do | niyuw | it | wugok |
| | split-VC | 1SG ERG | ERG[-DEF] | coconut | ABS[+DEF] | pig |

‘I will split a coconut for the pigs.’

Crucially, these prefixes *poN-* ~*po-* occur in sentences like (21a), voice constructions in which the actor is marked ABS, putative antipassives. But remarkably the function of these prefixes is to mark alternations in the semantic role of the non-actor argument. Kimaragang like English exhibits alterations in the choice of the undergoer NP like *load hay in the truck* and *load the truck with hay*. Constructions like the former, with the theme as CORE undergoer and the locative as OBLique occur with the prefix *po-*, while constructions with both arguments as CORE undergoer NPs take *poN-*:

| | | | | | | |
|----------|-------------|---------|----------|------|-----|---------|
| (22) (a) | ϕ-po-suwang | okuh | ditih | sada | sid | pata'an |
| | -UND-enter | 1SG ABS | this ERG | fish | OBL | basket |

‘I will put this fish in the basket.’

| | | | | | | |
|-----|----------------------|---------|------------|---------|------------|------|
| (b) | monuwang | okuh | do | pata'an | do | sada |
| | <i>m</i> -PON-suwang | | | | | |
| | -UND-enter | 1SG ABS | ERG [-DEF] | basket | ERG [-DEF] | fish |

‘I will fill a basket with fish.’

| | | | | | | | |
|----------|------------|---------|------------|-------|-----|-------|----------|
| (23) (a) | ϕ-pa-ta'ak | okuh | do | siin | sid | tanak | kuh |
| | -UND-give | 1SG ABS | ERG [-DEF] | money | OBL | child | 1SG POSS |

‘I give money to my child.’

| | | | | | |
|-----|---------------------|---------|------------|-------|----------|
| (b) | mana'ak | okuh | di | tanak | kuh |
| | <i>m</i> -poN-ta'ak | | | | |
| | -UND-give | 1SG ABS | ERG [+DEF] | child | 1SG POSS |
| | do | siin | | | |
| | ERG [-DEF] | money | | | |

‘I will give my child money.’

PoN- is also used to mark patients as undergoer, as (21a) demonstrates. Sentences like (21a) and (22)-(23) present a strong challenge to the ergative analysis and specifically the analysis of *-(u)m-* marked verbs as antipassives. In antipassive constructions, the undergoer must be an OBLique constituent, yet in Kimaragang the verb is obligatorily marked in *-(u)m-* inflected verbs for the specific semantic role of the undergoer. Crosslinguistically, verbal marking is exclusively restricted to CORE arguments, so the Kimaragang undergoer in (21a) must be a CORE argument. But if it is, again on crosslinguistic evidence, (21a) cannot be properly analyzed as an antipassive construction.

Nor will an analysis of *poN-* as an APPLicative marker work either. First, why would an APPLicative marker be needed in (21a)? The verb *lapak* ‘split’ is already a transitive verb subcategorizing an actor and an undergoer, as (21b) clearly shows. Why would an APPLicative be needed in (21a) with only the same two arguments? Consider, then, for the sake of argument that there are two *poN-*’s, a verbal derivational prefix in (21a) and a true APPLicative in (22b) and (23b). *PoN-* in the latter two examples is taken as an APPLicative suffix deriving a ditransitive verb from a formerly monotransitive root by presenting a formerly OBLique locational argument as a CORE argument. The main problem with this analysis is that it sacrifices what we saw with Tagalog data is the strongest evidence for the ergative analysis in the first place: namely, that given the universal one ABSolutive NP per clause constraint operative in true ergative languages, an APPLicative derivation will necessarily trigger antipassive, forcing the previous ABSolutive NP into OBLique case. Kimaragang has examples parallel to Tagalog *-an* marked verbs that seem to fit this description:

(24) (a) itih pe'es i-ta'ak dih kamaman sid dogon
 this ABS knife give ERG [+DEF] uncle OBL 1SG OBL
 'My uncle will give this knife to me.'

(b) taak-*an* okuh dih kamamen do pe'es
 give 1SG ABS ERG [+DEF] uncle ERG [-DEF] knife
 'My uncle gave me a knife.'

(24a) is the basic construction if the language is analyzed as ergative, with the actor in the ERGative case, the theme in the ABSolute, and the recipient in the OBLique. The suffix *-an* is analyzed as an APPLicative promoting the recipient to core status and deriving a ditransitive verb. Due to the one ABSolute per clause constraint, a simultaneous antipassive applies, throwing the theme into an OBLique case, in this case, the ERGative, as with Tagalog *ng*. Note importantly, no *poN-* is used in the derivation: the antipassive is implicated and simultaneous with APPLicative formation, and the single suffix *-an* is sufficient. Now note what happens with *-m-* marked verbs:

(25) mana'ak okuh dih kamaman do pe'es
 m-poN-ta'ak
 UND-give 1SG ABS ERG [+DEF] uncle ERG [-DEF] knife
 'I gave my uncle a knife.'

Note that there are two ERGative NPs, the theme and the recipient, and the prefix *poN-*, but no suffix *-an*. Example (25) is extremely hard to account for in the ergative/antipassive analysis. Note that the recipient occurs in the ERGative case, not OBLique. In order for that to be possible it needs to have gone through a stage of being a CORE argument, i.e. licensed by *-an*, but *-an* is nowhere to be found, only *poN-*. It begs credibility to suppose *poN-* is an allomorph of *-an*; one a prefix, the other a suffix. Further they can actually co-occur in certain verbal forms *poN-lapak-an* 'split things someplace'.

Finally, there are verb forms with *poN-* which occur in basic clauses, i.e. clauses in which the undergoer is in ABSolutive case:

- (26) *nunuh* *ot* *pana'ak* *nuh*
 poN-ta'ak
 what *ABS* *UND-give* *2SG* *ERG*
 'What would you give?'

nunuh 'what' is in ABSolutive case, but if *poN-* is an APPLicative then by the one ABSolutive NP per clause constraint *nunuh* should be ergative with the understood but ellipsed recipient in ABSolutive. The fact that this is not the case suggests either that *poN-* is not an APPLicative or Kimaragang is not an ergative language. So (25) as an antipassive is now wholly unaccounted for; we would expect by normal derivational processes (27) instead:

- (27) **t-um-aak-an* *okuh* *dih* *kamaman* *do* *pe'es*
 give-APPL *1SG* *ABS* *ERG [+DEF]* *uncle* *ERG [-DEF]* *knife*
 'I gave my uncle a knife.'

in which the antipassive triggered by the applicative *-an* first forces the theme *pe'es* 'knife' into ERGative case, and the second antipassive with the voice affix *-(u)m-* delinks the formerly ABSolutive recipient argument *kamaman* 'uncle' from CORE status and places it in OBLique ERGative case. Of course (27) is not what we really get; instead we find (25), with a prefix *poN-* that seems to verbally crossreference NPs as undergoers, a function of CORE grammatical relations no less! This all leads to the single conclusion that Kimaragang should not be analyzed as an ergative language.

Other Philippine type languages too present arguments against the ergative analysis via interactions with APPLicative derivations. Consider these data from Sama (Walton 1986):

- (28) (a) b'lli ku taumpa ma onde'
 buy 1SG ERG shoe(s) OBL child
 'I bought the shoes for the child.'
- (b) N-b'lli aku taumpa ma onde'
 ANTI-buy 1SG ABS shoes OBL child
 'I bought shoe(s) for the child.'

On the face of it, Sama would seem to be the ideal Philippine candidate for an ergative language. For the first time, we find a ϕ marked underived verbal form, one used with the undergoer as ABSolutive, a putative basic ergative construction, and another form derived with *N-*, presenting the actor as ABSolutive, and hence, an antipassive. But (28b) presents one obvious immediate problem for the antipassive analysis: the undergoer remains ϕ case marked as it was when it was the ABSolutive NP in (28a); it completely fails to take the OBLique case marker *ma*. Again, it is crosslinguistically the case that antipassives force undergoer NPs into OBLique cases, and it is a overwhelmingly the case that OBLique NPs are marked by some overt predictor/adposition/case marker, not by ϕ . ϕ marking is a property of CORE NPs and the fact that *taumpa* 'shoe(s)' in (28b) is so marked suggests that it indeed is a CORE NP and hence (28b) is not an antipassive.

Further evidence comes from interaction with APPLicative derivations. consider these examples:

- (29) (a) b'lli ku taumpa ma si Andi
 buy 1SG ERG shoe(s) OBL PROP PN
 'I bought the shoes for Andy.'
- (b) N-b'lli aku taumpa ma si Andi
 ANTI-buy 1SG ABS shoe(s) OBL PROP PN
 'I bought shoes for Andy.'

- (30) (a) b'lli-an ku si Andi taumpa
 buy-APPL 1SG ERG PROP PN shoe(s)
 'I bought Andy some shoes.'
- (b) N-b'lli-an aku si andi taumpa
 ANTI-buy-APPL 1SG ABS PROP PN shoe(s)
 'I bought Andy some shoes.'

(29a, b) are the expected basic ergative and antipassive constructions respectively. (30a) is a common examples of applicative derivation: *si Andi* which was oblique in (29), marked with *ma* is promoted to core status by the suffix *-an*. (30b) is the crucial example; it is the Sama equivalent of the unattested Kimaragang example (27). First APPLicativization applies, marked by *-an*, and then antipassivization with *N-*. While it is true that the no double ABSolute constraint holds for (30b), only *aku* 1SG is ABSolute, it remains problematic why both the beneficiary and the themes remain ϕ case marked, i.e. behave like CORE arguments, when the double operation of two antipassives should put both into OBLique cases. It is much more difficult in the Sama case to claim homonymy between a ϕ marked core ERGative case and an OBLique case, first, because OBLique cases are almost never ϕ marked, but require an overt marker; and second, because the Sama OBLique marker for instrumental NPs is in fact an overt preposition *maka*. We must conclude therefore that the antipassive analysis fails for Sama as it has for Tagalog and Kimaragang.

Returning to Tagalog now, there are a couple of other arguments that can be advanced against the ergative analysis. Proponents of the ergative analysis claim that forms in which the actor of the clause is marked by *ang* ABS are antipassives, and antipassivization is accomplished through deriving

intransitive verbs from basic transitive ones. There are a number of serious problems with this view. First, it depends on a sharp, well defined notion of transitivity, dividing verbs into transitive and intransitive classes.

Unfortunately, this notion is anything but clear in Philippine languages. Not a few semantically intransitive verbs occur with voice suffixes diagnostic of transitive verbs, namely *i-*, *-in* and *-an*: *i-kasal* ‘get married’, *pawis-an* ‘sweat’, *kilabot-an* ‘be terrified’, *antuk-in* ‘fell sleepy’, *langgam-in* ‘be infested with ants’, *ma-lamig-an* ‘feel cold’. Nor is this restricted to Tagalog: witness the Kimaragang intransitive verbs *tuuwan* ‘be thirsty’, *losu-an* ‘feel hot’, *sogit-on* ‘feel cold’, *o-weeg-an* ‘be flooded’ (Kroeger 1988). Thus, we cannot simply gloss *-um-* as simply ‘intransitive’ and be done with it: there are a number of different voice affixes that semantically intransitive verbs take. This is because the voice affixes signal subtle semantic properties of the NP marked by *ang*, as well as semantic properties of the verb (see Ramos 1974; de Guzman 1978). They are not global markers of a syntactic distinction of transitivity, which is in fact very problematic in Tagalog.

Related to this fact are the particular semantics supplied to roots by *-um-*. It occurs in paradigmatic alternation with a number of other affixes for the derivation of verbal forms, all of which mark their sole core argument with *ang*:

- (31) (a) *pula* ‘red’ → *ma-pula* ‘be redish’ → *p-um-ula* ‘become red’ → *mam-*
 (p)*ula* ‘blush’
- (b) *um-abot* ‘reach for’ *mag-abot* ‘hand to’
 bangka ‘boat’ *mam-(b)angka* ‘go boating’
 bus ‘bus’ *mag-bus* ‘ride a bus’
 payat ‘thin’ *p-um-ayat* ‘become thin’

Again simply glossing an affix like *-um-* or any of the other affixes in (31) as ‘intransitive’ ignores the rich derivational and semantic functions they serve. These data cast grave doubt on the claim that *-um-* is an intransitivizer or antipassivizer for it is clearly in paradigmatic alternation with the other affixes in (31), so that if it is, they must be too. But while some languages may have more than one antipassive, the differences are largely syntactic, not the semantic subtleties expressed, for example, in the contrast between *-um-*, *mag-*, *ma-* and *maN-*. This strongly suggests that *-um-* is not a marker of antipassive, but simply a verbal derivational suffix, subject to semantic conditions and in paradigmatic contrast to other such affixes.

2.3 Evidence against the unergative-unaccusative analysis

I have now considered both an active-passive and ergative-antipassive analysis of the Philippine type voice alternations typified by Tagalog (1), Atayal (6) or Kimaragang (8) and marshalled a number of arguments to reject both. The question remains where does this now leave us in trying to understand these languages? There is, of course, a third type, the unergative-unaccusative split type language, and while Tagalog and Kimaragang (Kroeger 1990) unquestionably do have some properties of such languages in their patterns of verbal morphology (Drossard 1984), it is clear that this type is also inappropriate because both unergative and unaccusative intransitive verbs marked their single core argument with *ang*, although they commonly do take different verbal affixes, typically *-um-* for unergatives and *ma-* for unaccusatives:

(32) (a) unergative *punta* ‘go’

| | | |
|--------------------|------------|--------|
| p- <i>um-</i> unta | <i>ang</i> | lalake |
| go | | man |

‘The man went.’

(b) unaccusative *lanta* ‘wither’

| | | | |
|--------------------|------------|------------|-----------------|
| <i>ma-la-lanta</i> | <i>ang</i> | <i>mga</i> | <i>bulaklak</i> |
| IRR-wither | | PL | flower |

‘The flowers will wither.’

While the hypothesized unergative-unaccusative split does explain some of the variations we find in verbal morphology (as it should if, as we claimed earlier, verbal morphology reflects subtle semantic differences in verbal meaning and the semantics of participants), the fact that the sole CORE argument of both verb types is marked with *ang* indicates that Tagalog is not *syntactically* of this type. Again Philippine languages seem to slip through the crack of the available syntactic typologies for voice and case marking systems.

3.0 The symmetrical voice type

Let me offer now the possibility that Philippine languages actually belong to a distinct syntactic type, one not covered by any of the three investigated thus far. I will call these languages the symmetrical voice type. To understand more clearly what this type entails, let me summarize first the properties of asymmetrical voice languages typified by nominative-accusative English and ergative-absolutive Dyirbal. Both of these languages have a single NP per clause, the NOMinative in English and the ABSolute in Dyirbal which is selectively treated grammatically vis-a-vis all the other NP types; let’s call this type of NP, the PIVot (see Dixon 1979 and Foley and Van Valin 1984 for further discussion and justification of this notion). There is a marked preference in each language as to which NP should function as the PIVot; this is the actor for a nominative-accusative language like English and the undergoer for an ergative-absolutive language like Dyirbal. This is indicated by the fact that the clause type is the basic form if these NPs are selected as PIVot and the verb is morphologically unmarked for voice. In any other NP type, i.e. undergoer for

English or actor for Dyirbal, is selected as PIVot, a marked clause type occurs and a special voice affix on the verb must be used, passive or antipassive.

Hence, voice choice in English and Dyirbal is asymmetrical, because one NP is selectively preferred for PIVot choice over all others, a preference clearly signalled in (the lack of) verbal morphology and congruent NP case marking.

Philippine type languages contrast in being symmetrical voice language. No one NP type is preferred for PIVot choice (i.e. *ang* marked NPs in Tagalog); regardless of which choice is made, all are signalled by some overt verbal voice morpheme (e.g. *-um-*, *-in*, *-an*, *i-*, etc). Further, other than the superposition of *ang* marking on the NP choice for PIVot, no alterations accrue to the case marking of the NPs in the clause, in marked contrast to the radical rearrangements of case marking required by marked voice options like passive or antipassive in asymmetrical languages.

Let me explore the lexical implications of this proposed split. In an asymmetrical language like English or Dyirbal, the unmarked voice choice (active or ergative) corresponds to the basic lexemic form of the verb:

(33) (a) Fred smashed the pumpkin.

smash <actor, undergoer>
|
PIVot

(b) Egbert gave a python to Mildred.

give <actor, undergoer, locative>
|
PIVot

For both these verbs and innumerable others, the basic lexeme form, *smash*, *give*, etc, corresponds to the construction in which the actor is PIVot, active voice. In order to make other choices, a special lexical derivational process (see Bresnan 1982; Foley and Van Valin 1984) needs to apply to the basic lexical entities above. The PASSive suffix *-en* functions to delink the actor from the

argument structure of the lexeme, so that it is no longer a possible PIVot choice. The result is a marker derived verb form in which the PIVot choice must be one of the other arguments:

- (34) (a) The pumpkin was smashed (by Fred).

smash-ed <undergoer>
|
PIVot

- (b) The python was given to Mildred

give-en < undergoer, locative>
|
PIVot

- (c) Mildred was given a python

give-en < undergoer, locative>
|
PIVot

Now consider the lexical implications of the symmetrical analysis. The Tagalog root for ‘give’ is *bigay*. However, unlike English *give*, it is never used in this unmarked form. It always requires an overt voice affix to signal which NP is PIVot:

| | |
|---|---|
| <i>mag-bigay</i> <actor, undergoer, locative> PIVot | <i>i-bigay</i> <actor, undergoer, locative> PIVot |
| <i>bigy-an</i> <actor, undergoer, locative> PIVot | |

But there is still another crucial difference between English and Tagalog related to this. As examples (1d, e), (6d, e) and (8d, e) above and (35) below (Schachter and Otane 1972) demonstrate, the choice of PIVot in Tagalog, unlike English is not restricted to subcategorized arguments:

- (35) (a) p-in-ag-halu-*an* ni Rosa ng asukal
 PERF-stire-VC CORE PN CORE sugar

sa kape ang kapitara
 OBL coffee PIV coffee pot

‘Rosa stirred sugar into the coffee in the coffee pot.’

(b) p-in-ag-sulat-*an* ko ang desk na ito
 PERF-write-VC 1SG.CORE PIV desk LIG this

‘I wrote on this desk.’

These facts really pose a dilemma. While we can plausibly argue that, for example, the English verb *give* has the argument structure <actor, undergoer, locative>, i.e. subcategorizes the NP fulfilling these roles, because among other things, each of these can assume PIVot status, one cannot make a similar argument for Tagalog or other Philippine languages. In these languages, the NPs that can assume PIVot function are by means restricted to NPs that could be claimed to be subcategorized or semantically selected by the verb, as the above examples illustrate. No plausible theory of verbal lexical semantics would want to argue that instruments or outer locatives and temporal adjunct NPs are semantically selected by verbs, yet they can be PIVot choices in these languages (Schachter and Otnes 1972: 310-321). If this is the case, how can we argue a Tagalog root like *bili* ‘buy’ has an argument structure like <actor, undergoer> or even <actor, undergoer, locative>? Do we really want to claim on the basis of the examples in (1) that its argument structure is <actor, undergoer, locative, instrument, beneficiary> or, more strikingly, that the argument structure of Tagalog *halu* ‘stir (35a) is <actor, theme, locative₁, locative₂>? I think not; such a claim would make a shambles of any claim of the theoretical usefulness of a crosslinguistically viable notion of argument structure.

What might be the alternative? Perhaps, roots like *bigay* ‘give, *halu* ‘stir’, *bili* ‘buy’, etc do not entail argument structure at all, merely some generalized

conceptual structure paraphrasable as ‘giving by X of Y to Z’ or ‘stirring by A of B into C at D’, with potential spacio-temporal coordinates understood. True argument structure as we understand it crosslinguistically would only be introduced when the roots are derived with the voice markers: postulation of an argument structure array and choice of a PIVot are simultaneous. This is why verbs in clauses are never unmarked for voice: without a voice affix they lack argument structure and hence cannot govern NPs. The coemergence of argument structure and PIVot choice is also why semantically very marginal NPs like outer locatives can be PIVots: when they are to be PIVots, verb forms are derived which actually subcategorize them, i.e. include them in the argument structure so they are properly governed.

This proposal raises yet another possibility. If roots like *bigay* ‘give’ and *halu* ‘stir’ lack argument structure in their basic lexical forms, what then are the criteria for considering them lexical verbs? Not much as we shall see, and this typological fact, no lexical distinction between noun and verb roots will turn out, I will argue, to be the ultimate basis of the symmetrical voice typology as a whole. The claim here is that Tagalog roots are basically *precategorical*, neither noun nor verb. Many roots which correspond to entities and would therefore be nouns in languages with a sharp lexical noun/verb contrast can really be used as verbs, simply with the addition of a voice affix:

| | | | | |
|------|---------|----------|-------------|-------------------------|
| (36) | abogado | ‘lawyer’ | mag-abogado | ‘become a lawyer’ |
| | bus | ‘bus’ | mag-bus | ‘ride a bus’ |
| | kamay | ‘hand | mag-may-an | ‘shake hands’ |
| | langgam | ‘ant’ | langgam-in | ‘be infested with ants’ |
| | anak | ‘child’ | mang-anak | ‘give birth to’ |
| | lubid | ‘rope’ | lubir-in | ‘be made into rope’ |

(Note that while the English noun *rope* can be used as a verb too, it is done without derivation and is semantically transparent as part of a general syntactic process, N(instrument)→V(use instrument). The Tagalog processes, however, are neither semantically transparent nor formally predictable, as befits a lexically based derivational process.) Even OBLiquely case marked NPs can be verbalized via a voice affix:

- (37) sa Maynila ‘at Manila’ p-um-a-sa Maynila ‘go to Manila’
 sa Kastila ‘Spanish’ mag-sa Kastilya ‘be like a Spaniard’
 sa inyo ‘your’ s-um-a inyo ‘be yours’

When we turn to the other side of the equation, deriving nouns from roots which semantically denote events, prototypical verbs in languages with a clear noun/verb contrast, we find a remarkable fact: clear cases of lexical derivational processes of nominalization are actually impossible to find in the language, in spite of the fact that nominalization is an extremely common derivational process crosslinguistically. All morphemes which could be claimed to derive nominals are also used with derived verb forms: there are *no* unique nominalization affixes, a highly salient typological fact. Consider these examples (Schachter and Otones 1972: 97-106):

- (38) (a) *maN-*

| | | | |
|------|--------|-------------|--------------|
| ibig | ‘love’ | mangingibig | ‘lover’ |
| tahi | ‘sew’ | mananahi | ‘dressmaker’ |
| basa | ‘read’ | mambabasa | ‘reader’ |

- (b) *-in*

| | | | |
|------|---------|----------|--------------|
| aral | ‘study’ | aral-in | ‘a lesson’ |
| awit | ‘sing’ | awit-in | ‘a song’ |
| bili | ‘buy’ | bilih-in | ‘a purchase’ |

(c) *-an*

hiram ‘borrow’ hiram-an ‘place for borrowing’

tagu ‘hide’ tagu-an ‘hiding place’

giik ‘thresh’ giik-an ‘threshing place’

Other verbal derivational affixes, for example the CAUSative prefix *pa-*, can also be used as seeming nominalizers:

| | | | |
|------|-------------|------------------------------------|---|
| (39) | verbal | | nominal |
| | mag-pa-luto | ‘make someone cook something’ | pa-luto ‘something caused to be cooked’ |
| | mag-pa-abot | ‘make someone hand something over’ | pa-abot ‘something caused to be handed over’ |
| | mag-pa-gawa | ‘make someone make something’ | pa-gawa ‘something caused to be made’ |

Even affixes which derive activity nominals are homophonous with affixes widely used in verbal derivations:

| | | | | | |
|----------|------------------------------|----------|---------|------|---------|
| (40) (a) | bi-bigy-an | ko | silá | ng | regalo |
| | IRR-give-VC | 1SG CORE | 3PL PIV | CORE | present |
| | ‘I will give them presents.’ | | | | |

| | | | |
|-----|--------------------------------------|-----|---------|
| (b) | bigay-an | ng | regalo |
| | give | GEN | present |
| | ‘The giving of one another presents’ | | |

| | | | | | | |
|----------|----------------------------------|----------|------|---------|-----|-------|
| (41) (a) | i-pag-lu-luto | niya | ng | pagkain | ang | bata |
| | VC- -IRR-cook | 3SG CORE | CORE | food | PIV | child |
| | ‘She cooked food for the child.’ | | | | | |

| | | | |
|-----|---------------------------|-----|---------|
| (b) | pag-lu-luto | ng | pagkain |
| | cook | GEN | food |
| | ‘The cooking of the food’ | | |

This pattern of homophony is an extremely puzzling one. While, as Himmelmann (1991) points out, there may be prosodic differences between the verbal and nominal uses of the affixes, not unlike English *súbject* (N) versus *subjéct* (V), still the pervasive homophony is extremely striking and needs to be addressed. Nor is it restricted to Tagalog as Starosta, Pawley and Reid (1980, 1982) demonstrate, but is widespread among Philippine type and Western Austronesian languages generally. There are three possible ways to account for this homophony. One is to claim that they are basically nominalizing affixes; the use of them to derive verb voices is deceptive in that these derived voice forms are in fact nothing but nominalizations. This is an old and very venerable analysis of Philippine type clause structure, one that goes back to the nineteenth century, but still has considerable current support (see, for examples, Starosta, Pawley and Reid 1982; Naylor 1995). In such an analysis a so called basic transitive clause like (40a) actually is an equational clause with a predicative nominization headed by a verb with its associated dependents followed by a subject, so that (40a) would actually be more accurately translated as ‘the will be given to by me of the present was they’. There are indeed sound arguments for this analysis, hence its longevity, specifically the ease with which any so called verb can function nominally simply with the addition of a NP marking preposition:

- (42) (a) ang bi-bigy-*an* ko ng regalo ang bata
 PIV IRR-give-VC 1SG CORE CORE present PIV child
 ‘It’s the child that I will give the presents.’
 (literally) ‘The child is the will be given to by me of the presents.’

(b) *i-bi-bigay* ko ang regalo sa mga d-*um*-ating
VC-IRR-give 1SG CORE PIV present OBL PL arrive
‘I will give presents to the ones who come.’

(literally) ‘The will be given to by me of presents are the arrivers.’

But the nominalizing analysis immediately runs afoul of the precategoriality of Tagalog roots. Claiming these affixes as nominalizers only makes sense if we can plausibly claim they are verbal roots in the first place, which are then derived into nominals via affixation. However, many of the roots of these so called nominalizations denote prototypical entities, which would need to be assigned to a noun class. Why would one derive a noun (i.e. nominalize) from something which already is a noun? Examples: *abogado* ‘lawyer’, *mag-abogado* ‘become a lawyer’, *Tagalog* ‘Tagalog’, *mag-Tagalog* ‘speak Tagalog’, *anay* ‘termite’, *anay-in* ‘be infested with termites’, *adobo* ‘traditional Philippine meat dish’, *adoboh-in* ‘braise’, *bangka* ‘boat’, *mam-(b)angka* ‘go boating’; *tiket* ‘ticket’, *tikit-an* ‘give a ticket to’, *kamay* ‘hand’, *kamay-an* ‘shake hands with’. In order to uphold the nominalization analysis for *anay-in* ‘be infested with termites’, for example, we would have to claim that a verb stem *anay* ‘be infested with termites’ is first derived by ϕ affixation from the noun root *anay* ‘termite’, and then the derived verb stem is further nominalized by *-in*. This seems a curiously roundabout and unmotivated analysis, especially given the fact there is no evidence at all for a verbal form *anay*. The nominalization analysis needs to postulate this completely unattested stem, and this ad hocness should make us seriously question its viability.

Another possible approach to the homophony of verbal and nominal derivational morphology has been suggested by Himmelmann (1991). In essence he extends the claim of the precategoriality of Tagalog roots to fully derived words as well; a similar claim has been made by Gil (1993). Again

there is much in Tagalog grammar that supports this analysis. As we have seen (42a, b), Tagalog verbs can freely function as NPs without any further derivation, simply being preceded by an NP case marking preposition. Taking these verbs as the heads of these NPs they can, by normal theories of phrasal projection, be claimed to be functioning like nouns. Thus, the same fully derived word can function as apparently different word classes, depending on the syntactic structure it occurs in, either as a noun (42a) or a verb (40a), suggesting that the word itself lacks specification as either a noun or a verb. As attractive as this account is, it too runs into serious empirical difficulties. First, as Himmelmann (1991) himself points out sometimes a derived noun form and the verbal form to which it would be closest semantically and formally are not formally identical (Schachter and Otnes, 1972: 99-100): *bilihin* (N) ‘something to buy’ *bilhín* (V) ‘NP buys NP (PIVOT)’; *kumpunihin* ‘something to repair’, *kumpunihín* ‘NP repairs NP(PIVOT)’. Clearly these differences are related to the prosodic differences between nominal and verbal derivation, but if there is no lexical difference between these two forms, however we wish to label it, how can one state it? One cannot, as Himmelmann does, treat it as a problem of the phonology only, for the phonology on its own has no access to the kind of semantico-syntactic information necessary for the proper assignment of the stress. Hence some kind of distinct syntactic labelling for the two words is necessary for the phonology to assign stress properly, indicating a lexical syntactic distinction does exist. Again, we could choose to label this distinction in anyway we like, but given the meanings, noun versus verb does seem an obvious choice. Nor do the derived noun and the closest related verb form necessarily require the same affix: *alagain* (N) ‘something to take care of’, *alagaan* (V) ‘NP takes care of NP (PIVOT)’; *aralin* (N) ‘something to study’, *pagaralan* (V) ‘NP studies NP (PIVOT)’; *burdahin* (N)

‘something to embroider’ *burdahan* (V) ‘NP embroiders NP (PIVOT)’. There is also good syntactic distributional evidence for a distinction between nominal and verbal derived words. Consider the form *awitin*, completely homophonous between a noun ‘song’ and a verb ‘NP sings NP (PIVOT)’. Interestingly, they do not both have the same syntactic combinatorial possibilities. As a nominal form, it has two ways of forming pronominal possessors, either with the OBLique form of the pronoun preceding it, linked with the ligature *-ng*, *iyong awitin* 2SG OBL-LIG song, ‘your song’, or with a pronoun of the CORE series following it *awitin mo* song 2SG CORE ‘your song’. As a verb, however, it may only have a pronoun of the CORE series preceding or following it: *huwas mo-ng awit-in* NEG IMP 2SG CORE-LIG sing ‘don’t sing it!’ or *awit-in mo* sing 2SG CORE ‘sing it!’. It may never be directly preceded by a pronoun of the OBLique series: **huwag iyo-ng awit-in* NEG IMP 2SG OBL-LIG sing. In order to account for this distributional difference we have to claim a syntactic distinction between a homophonous derived nominal and verbal form *awitin*, ultimately rejecting the assertion that the precategoriality of roots could be expanded to derived forms as well.

Having argued against both the nominalization and the categorially undifferentiated analysis of Tagalog affixes, there only remains the third possibility advanced above: verbalization. In essence, the Tagalog affixes *-um-*, *-in*, *-an* and *i-* derive verbs primarily and any nominal uses are strictly secondary. This is the analysis I wish to argue for here, but before proceeding with it, it is important to see the Tagalog system in a wider typological perspective, and so let me turn to this now.

3. A comparative typological excursion: Northwest Coast languages

3.1 The lack of an noun/verb distinction in Northwest Coast languages

As strikingly unusual Philippine type languages might be, they are not completely unique. The problematic status of a noun/verb distinction was noted at least as long ago as 1911 by Boas (Boas 1966[1911]) for the Amerindian languages of the Northwest Coast of North America and has continued to be argued about by scholars working with these languages (Swadesh 1939; Sapir and Swadesh 1946; Boas 1947; Kuipers 1968; Jacobsen 1979; Kinkade 1983; Broschart 1991; Sasse 1993; Jelinek and Demers 1994). All three genetic groups in the area, Wakashan, Salish and Chimakuan, are said to exhibit a lack of a strong lexical contrast between noun and verb, as was claimed above for Tagalog roots. Consider these famous Nootka examples from Swadesh (1939: 78-9):

- (43) (a) *mamu·k-ma* *qu·ʔas-ʔi*
 work-3SG INDIC man-DET
 ‘The man is working.’
- (b) *qu·ʔas-ma* *mamu·k-ʔi*
 man-3SG INDIC work-DET
 ‘The working one is a man.’

Note that the root *mamu·k-* ‘work’ denotes an activity and would correspond to a prototypical verb, while *qu·ʔas* ‘man’ denotes an entity, a person, and as such is a prototypical member of a noun class. In (43a) everything is as we might expect: the mood marker, a prototypical verbal inflection, attaches to *mamu·k-* ‘work’ and the DETerminer, a prototypical modifier of nouns, attaches to *qu·ʔas-* ‘man’. But our expectations are turned on their head in (43b): here

qu-ʔas- ‘man’ is inflected for mood and *mamu-k-* ‘work’ takes the DETerminer!

These data argue that there is no sharp word class distinction between these two words; prototypical verbal and nominal inflections and modifications can apply to both. Parallel structures are found in Tagalog:

- (44) (a) *mag-tra-trabaho* *ang* *lalake*
 VC-IRR-work PIV man
 ‘The man will be working.’
- (b) *ang* *lalake* *ang* *mag-tra-trabaho*
 PIV man PIV VC-IRR-work
 ‘The one who will be working is the man.’

The pivot marking preposition *ang* is the Tagalog equivalent of the Nootka DETerminer -ʔi. There is no mood suffix like Nootka *-ma* to mark predicates in Tagalog; clause initial position is sufficient. Words of a wide semantic range can function as clause initial predicates in Tagalog, with undifferentiated syntactic behavior:

- (45) (a) *titser* *ang* *lalake*
 teacher PIV man
 ‘The man is a teacher.’
- (b) *um-alis* *ang* *lalake*
 VC-leave PIV man
 ‘The man left.’
- (c) *ma-taas* *ang* *lalake*
 tall PIV man
 ‘The man is tall.’
- (d) *na-sa* *Maynila* *ang* *lalake*
 OBL Manila PIV man
 ‘The man is in Manila.’

Just as the case with Nootka (examples (43a, b) above and below) (Swadesh 1939):

- (46) (a) $\text{ʔi} \cdot \text{h} \cdot \text{ma} \cdot$ $\text{qu} \cdot \text{ʔ} \text{as} \cdot \text{ʔi}$
 large-3SG INDIC man-DET
 ‘The man is large.’
- (b) $\text{ʔu} \cdot \text{k}^{\text{w}} \text{i} \cdot \text{ma}$ $\text{qu} \cdot \text{ʔ} \text{as} \cdot \text{ʔi}$
 to him-3SG INDIC man-DET
 ‘He does to, stands in relative to the man.’

And Straits Salish (Jelinek and Demers 1994), is another Northwest Coast language claimed to lack a lexical noun/verb distinction:

- (47) (a) $\text{t}^{\text{h}} \text{il} \ll \text{m} \cdot \text{l} \ll \text{-sx}^{\text{w}}$
 sing-PAST-2SG NOM
 ‘You sang.’
- (b) $\text{si}^{\text{h}} \text{em} \cdot \text{l} \ll \text{-sx}^{\text{w}}$
 chief-PAST-2SG NOM
 ‘You were a chief.’
- (c) $\text{sey}^{\text{h}} \text{si} \cdot \text{l} \ll \text{-sx}^{\text{w}}$
 afraid-PAST-2SG NOM
 ‘You were afraid.’
- (d) $\text{ʔ}^{\text{h}} \ll \text{-x}^{\text{w}} \text{otq} \ll \text{m} \cdot \phi \cdot \text{s} \ll \text{n}$
 to-waterfall-PRES-1SG NOM
 ‘I’m going to Bellingham.’

3.2 Symmetrical voice in the Northwest Coast: Kwakwaka

While these data support the contention that Tagalog patterns syntactically like the Northwest Coast languages, other languages claimed to be lacking a lexical noun/verb distinction, it is important to remember the reason why this typological fact became important in the first place. This salient typological

fact was claimed to be linked to that other very peculiar typological property of Philippine-type languages—their symmetrical voice structure. If roots are precategorial, i.e. neither noun nor verb, then the lack of any unmarked voice form of the verb finds a ready explanation: deriving a voice orientation of a verb requires deriving a verb as well, a process efficiently done simultaneously. It also entails a marked difference in lexical structure between asymmetrical languages like English and a symmetrical language like Tagalog. The lexeme *give* in English is a verb, with a corresponding argument structure <actor, undergoer, locative>. The Tagalog root *bigay* ‘give’ however is precategorial; it lacks a true argument structure, but does have a precategorial semantic structure like ‘the giving of something to someone by someone’.

If Northwest Coast languages, like Tagalog, have precategorial roots, then we would expect to find among them symmetrical voice languages. And indeed we do, in the first of these languages to be described as lacking the noun/verb distinction, Kwakwala (Boas 1911) a language of the same family as Nootka. Kwakwala is remarkably like Tagalog in many respects, and by exploring and developing the implications of these similarities, we will attempt to delimit more precisely just what is meant by a symmetrical voice language.

Kwakwala like Tagalog is essentially predicate initial with all NPs case marked by prepositions, which along with DETerminers are phonologically encliticized to the previous full phonological word (a process, interestingly, also found in the Miaoli dialect of Atayal (Huang 1993)). The basic case markers are *i*~ ϕ PIVot, *x*5 ACCusative and *s* INSTRumental (all homophonous with the corresponding bound verbal pronominal enclitics for third person). The PIVot is also crossreferenced by a pronominal agreement enclitic to the first element of the clause, the verb, if it occurs first; otherwise, an auxiliary (Boas 1947;

Levine 1980a, b); Anderson (1984). Consider this example from Anderson (1984: 24):

- (48) (a) k^wix^ʔid-i-da b«g^wan«ma-x5-a q'asa-s-is
 clubbed-PIV-DET man-ACC-DET otter-INSTR-3SG POSS
 t'alwag^wayu
 club

‘The man clubbed the sea otter with his club.’

- (b) la-i ax5^ʔid-i-da c'adaga-x5-a 'u^ʔ«lq^wʔi
 AUX-3 take-PIV-DET woman-ACC-DET dishes

‘Then the woman takes the dishes.’

(Note that the PIVot case marker and the third person enclitic pronoun are identical.) Other than the encliticization of the NP markers and the cross referencing of the PIVot, the system is much like Tagalog. Tagalog too requires clitic pronouns to precede the main verb, if there is a preceding auxiliary or adjunct phrase:

- (49) hindi/bukas ko sila bi-bigy-*an* ng regalo
 NEG/tomorrow 1SG CORE 3PL PIV IRR-give-VC CORE present

‘I will not/tomorrow I will give them presents.’

Further, Kwakwala has a system of multiple voices: an unmarked verbal form, in which the actor is PIVot; a verb form suffixed with *-su^ʔ*, in which the undergoer is PIVot; and a verb form suffixed with *-ayu*, in which OBLique NPs assume PIVot status (Levine 1980a):

- (50) (a) n«p'id-i-da g«nan«m-x5-a guk^w-s-a t'is«m
 throw-PIV-DET child-ACC-DET house-INSTR-DET rock

‘The child hit the house with a rock.’

- (b) n«p'id-*su^ʔ*-i-da guk^w-s-a g«nan«m-s-a t'is«m
 throw-VC-PIV-DET house-INSTR-DET child-INSTR-DET rock

‘The child hit the house with a rock.’

- (c) $n\ll p'id-ayu-i-da$ $t'is\ll m-x5-a$ guk^w-s-a $g\ll nan\ll m$
 throw-VC-PIV-DET rock-ACC-DET house-INSTR-DET child
 ‘The child hit the house with a rock.’

Each verb form, ϕ , $-su^?$ or $-ayu$, corresponds to a different PIVot choice, whose semantic role is indicated by the proper suffix. The PIVot is realized in Kwakwala by position, immediately following the main verb (if an enclitic pronoun to an AUXiliary, it will precede the verb) and the case marker $-i$. Like Tagalog voice affixes, these same suffixes are used derivationally to form what look like prototypical nouns; $m\ll x$ - ‘strike’ $m\ll n-ayu$ ‘drumsticks’; $G\ll ls$ - ‘paint’ (V), $G\ll ls-ayu$ ‘paint (N)’; $yaq'\ll nd$ - ‘talk’, $yaq'\ll nd-ayu$ ‘topic of conversation’; λax^w - ‘fish’ $\lambda aw-ayu$ ‘fish trap’; $n\ll p-x^?id$ throw-PUNC ‘throw’, $n\ll p-x^?id-ayu$ ‘a projectile’; $^?n\ll w$ - ‘aim’, $^?no^?-ayu$ ‘medicine for rubbing’; $n\ll p-x^?id$ - throw-PUNC ‘throw’, $n\ll p-x^?id-su^?$, ‘a target’; $ha^?mx^?id$ - ‘eat’, $ha^?mi^?yid-su^?$ ‘food’; there is even an equivalent suffix to Tagalog $-an$ for deriving putative nouns which describe the place where the activity or event occurs: $n\ll k$ - ‘steam’, $n\ll g-as$ ‘place for steaming’; $t\ll c'$ - ‘warm oneself’, $t\ll c'-as$ ‘place where one warms oneself’; $han\lambda$ - ‘shoot’, $han\lambda-as$ ‘place of shooting’; $t'\ll m$ ‘sew’ $t'\ll m-as$ ‘place of sewing’ (Boas 1947). This suffix can also be used rather like another voice marker, to indicate locational participants as PIVot, exactly parallel to Tagalog $-an$ (Levine 1980a: 243):

- (51) $lala^?-as-ux5^w-da$ $^?w\ll i^?nag^wis-s-a$ $b\ll g^wan\ll m$
 go-VC-DEIC-DET village-INSTR-DET man
 ‘The village is gone to by the man.’

And perhaps most strikingly parallel to Tagalog, any fully derived verb with its governed arguments can function as a NP simply with a preceding

Levine 1980a):

- in (52d) -«ʹ is an allomorph of -suʹ used with experiential and perception verbs (Levine 1980b). In each of the examples in (52), the verb form following the DETerminer is a fully finite one that could serve as the main predication of clause if it were in clause initial position. Like Tagalog there are no nominalizing affixes used.

It is important to note the points of similarity and difference between Tagalog and Kwakwala to get a fuller typological overview of symmetrical voice systems. Note firstly that non-subcategorized OBLique arguments like locatives and instruments are promoted directly to PIVot status without going through any intermediate lexical derivation of APPLcativization. This seems to be an important defining property of symmetrical voice systems, provided, of course, the languages do not otherwise exhibit clear defining properties of asymmetrical ergative-absolutive languages which also show this property,

due to their no double ABSolutive constraint. We have already argued that Tagalog is clearly not an ergative-absolute language, and Kwakwala too exhibits no clear characteristics of this language type: the presence of an overt ACCusative case preposition, for example, being a strong argument against classifying the language as an ergative one.

One difference between Kwakwala and Tagalog concerns the morphological paradigm of the voice affixes. In Tagalog all verb forms bear overt affixes; there is no unmarked verb form, a point I argued was crucial to the overall typology of symmetrical voice languages. But note in Kwakwala, the verb form which indicates the actor is PIVot bears no overt affix; does this indicate that their form is unmarked and so constitutes an argument against the symmetrical analysis which asserts there are no unmarked voice types, in contrast to the situation in asymmetrical languages? The answer is no, and it revolves crucially around making a distinction between unmarkedness and simply the lack of an overt morpheme, in essence, claiming that ϕ is just one option in a paradigmatic cell of options and no more basic than the others. In other words we are distinguishing between the basic lexemic form and a derived form marked by no overt morpheme. There is good evidence to believe the latter is the case in Kwakwala. Note that in English we have the basic lexemic form *hit*, but that any derivative use of this is overtly signalled: *be hit*, *hitting*, *hitter*, *hittee*, etc. Crucially, nominalized uses like the last two, in which the argument structure of the verb had been altered, are derivationally marked. Now consider Kwakwala. Note that the same form which indicates the actor is PIVot is also the agentive nominalization form ‘hitter’ (52a) or ‘thrower’ (52c). And it is the ϕ marking which distinguishes this from the ? nominalization ‘hittee’ (52b) or ‘seen’ (52d). This strongly suggests that ϕ is in

paradigmatic contrast with *-su*[?], *-‘*, (or *-ayu* in still other cases). The forms are not really unmarked in the sense of being basic, underived; they are marked, but by ϕ .

Finally, note that actors not functioning as PIVot in Tagalog are still case marked as CORE NPs with *ng*, but in Kwakwala they appear to be marked as OBLique with *s*. This may be a difference between the two languages, so that the property of non-PIVot actors remaining CORE may not be a defining property of symmetrical voice languages. However, we should inquire as to whether Kwakwala *s* is unambiguously an OBLique case marker. First, it does mark NP arguments of some verbs which are clearly subcategorized by the verbs in question and hence CORE participants:

- (53) (a) *n*«*p-x*[?]*id-i-da* *b*«*g*^w*an*«*m-s-a* *t*'*is*«*m*
 throw-PUNC-PIV-DET man-INSTR-DET rock
 'The man threw the rock.'
- (b) *c*'«-*s-a* *x*«*g*«*me* *la-q*
 give-INSTR-DET comb go-3
 'She gave him a comb.'
- (c) «*xc*'*odi*[?]-*s* *la-x*5-*a* 'oq'*we*
 put-INSTR go-ACC-DET dish
 '(He) put it into a dish.'

Perhaps, even more problematic for a claim for the oblique status of INSTRumental NPs is the fact that the language, like Tagalog, possesses three distinct series of enclitic pronominal forms, one corresponding to the PIVot, one to the *x*5 ACC marked NPs, and one for *s* INSTR NPs (note that the third person enclitics are identical to the corresponding NP case markers):

| | | PIV | x5 ACC | s INSTR |
|---|---------|-----------------------------------|--------|-----------------------------------|
| | SG | -«n(‘) | ϕ | -«n(‘) |
| 1 | PL INCL | -«nts | ϕ | -«nts |
| | PL EXCL | -«nu [?] x5 ^w | ϕ | -«nu [?] x5 ^w |
| 2 | | -«s | -o‘ | -os |
| 3 | | ϕ~i | -q~-x5 | -s |

Crucially, when first or second person actor does not function as *PIVot*, it is still realized as a bound pronominal affix on the verb (Boas 1947: 286):

- (54) (a) hi[?]m-«n ‘oma «x[?]ix5sd«-s«[?]-wi
DEIC-1SG INSTR much like-VC-DEIC
‘That is what is liked very much by me.’
- (b) hi[?]m-«n duG^w-«‘i-da ixp«ma‘a
DEIC-1SG INSTR see-VC-PIV-DET good play
‘That good play was seen by me.’

Data like (54) strongly argue that non-PIVot actors in Kwakwala are indeed CORE arguments. OBLique NPs are not realized by bound or enclitic pronominals crosslinguistically. And note further the behavior of the Kwakwala INSTRumental enclitic pronominals is exactly parallel to Tagalog CORE *ng* pronominals; both encliticize to the first major constituent of the clause. Indeed as remarked earlier, Tagalog OBLique *sa* pronominals *never* encliticize. The again very striking parallelism between Tagalog and Kwakwala on this feature argues that non-PIVot actors in both languages should be assigned the same syntactic status: CORE NPs.

3.3 Summary of symmetrical voice system properties

On the basis of the pervasive similarities between Philippine languages and Kwakwala we may propose the following properties as diagnostic of symmetrical voice languages:

1. Voice forms which make nonsubcategorized arguments like locatives and instrumentals directly accessible to PIVot function without going through an intermediate APPLicative derivation. This is criterial for the symmetrical voice type provided the language is not otherwise of the ergative-absolutive type, in which case the constraint against double ABSolutive NPs may be operative.
2. No neutral voice forms of the verb, all being equally marked as derived forms, although one of the voice affixes may be a zero morpheme. A logical consequence of property (2) would be symmetrical voice languages in which semantic classes of verbs vary as to the PIVot choice signalled by zero, in contrast to Kwakwala in which zero always indicates actor as PIVot. If the zero morpheme is just a lack of overt phonological material in paradigmatic opposition with other parallel morphemes having overt phonological realizations, then it is fully to be expected that various verbal semantic classes may differ as to which voice variant bears the zero form. Such languages do in fact exist; for example Bilaan of the Southern Philippines (Abrams 1961; Rhea 1972):

(55) *klang* 'cut'

- | | | | | |
|-----|-----------------------------|------|------|------------------------------------|
| (a) | klang-gu | kayo | balò | zero indicates instrument as PIVot |
| | cut-1SG CORE | wood | bolo | |
| | 'I cut wood with the bolo.' | | | |

(b) k-*an*-lang-gu kayo (-)(a)n- indicates undergoer as PIVot
 -cut-1SG CORE wood
 'I cut the wood.'

(c) k-*am*-lang agu kayo (-)(a)*m*- indicates actor as PIVot
 -cut 1SG PIV wood
 'I cut wood.'

(56) *bat* 'throw'

(a) bat-gu batu di gumnè zero indicates undergoer as PIVot
 throw-1SG CORE rock LOC house
 'I throw the rock at the house.'

(b) *n*-bat-gu batu i gumnè (-)(a)n- indicates locative as PIVot
 throw-1SG CORE rock PIV house
 'I throw a rock at the house.'

(c) *m*-bat agu batu di gumnè (-)(a)*m*- indicates actor as PIVot
 throw 1SG PIV rock LOC house
 'I throw a rock at a house.'

(57) *subè* 'go upstream'

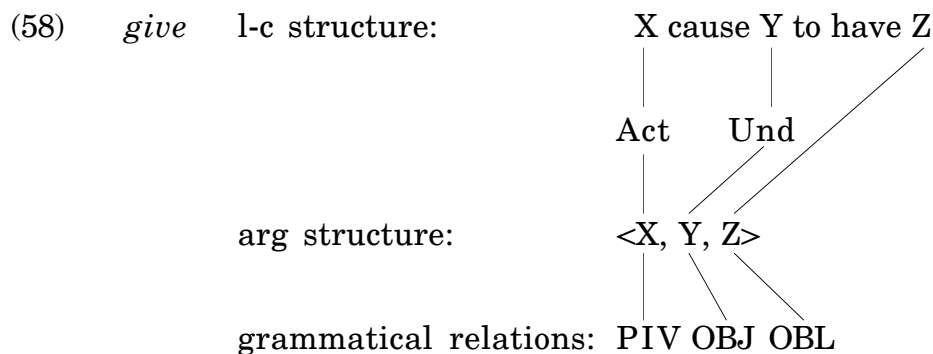
(a) subè agu zero indicates actor as PIVot
 go upstream 1SG PIV

(b) s-*n*-ube-gu yéél én (-)(a)*n*- indicates locative as PIVot
 -go upstream-1SG CORE river DEIC
 'I go upstream on that river.'

3. Actor participants when they fail to appear as PIVot continue to appear as CORE NPs. The same applies to undergoer participants in voice constructions in which non-subcategorized erstwhile OBLique participants like locative or instrumentals function as PIVot (compare examples (1d) and (50c)).

4. There is a lack of a clear noun/verb contrast in the lexical roots of the language.

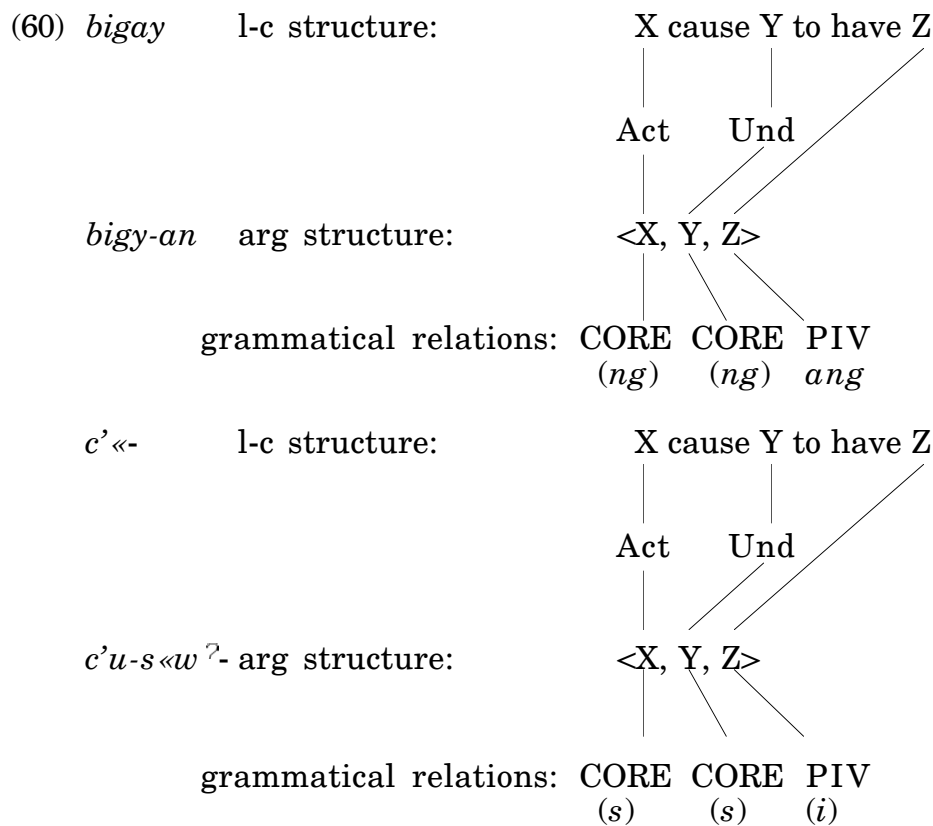
It is this last feature which I have identified as the lynch pin for the whole symmetrical voice type: if roots are precategorial, neither verb nor noun, they necessarily lack lexical properties of verbs, such as subcategorized participants in argument structures. Choosing a participant as PIVot requires an argument structure array or participants to choose among, and this in turn requires a verb. In symmetrical languages with precategorial roots the whole process is done simultaneously with one derivation. All words in all languages have a semantic representation of lexical conceptual structure, and let us assume these are roughly similar across languages, so that English *give*, Tagalog *bigay* and Kwakwala *c'«-* all would have a lexical conceptual structure roughly along the lines of 'X causes Y to have Z'. Because English *give* is a verb this maps directly onto an argument structure, along the lines of (58):



Tagalog and Kwakwala, on the other hand, with precategorial roots do not have an argument structure tied to the root form; there is only the l-c structure:

- (59) *bigay* l-c structure: X cause Y to have Z
 'give'
 c'«- l-c structure: X cause Y to have Z

Argument structures accrue to the roots when they are derived as verbs and at the same time pivot selection is made:



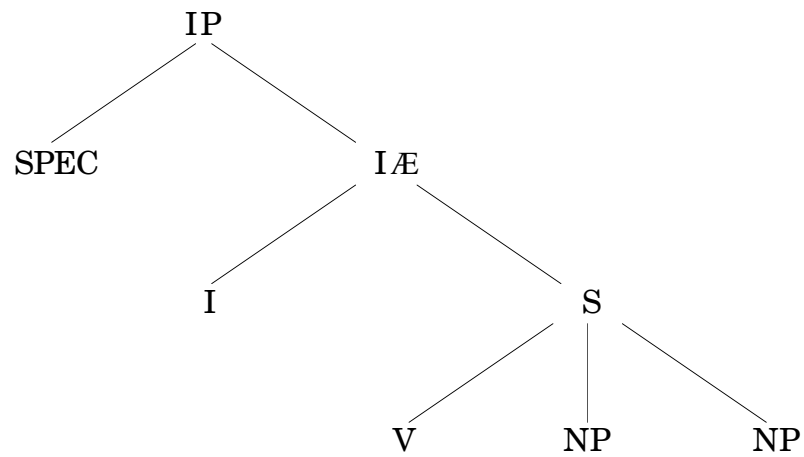
A major typological break, right down to the structure of the lexicon, is thus claimed between symmetrical and asymmetrical languages. This obviously has important further typological and theoretical implications, what will now be addressed.

4.0 Theoretical implications

Tagalog and Kwakwala's similarities go much deeper than just morphological and lexical; they have largely parallel clausal constituent structures as well.

Kroeger (1992) convincingly demonstrates that the following is the basic constituency of a Tagalog clause

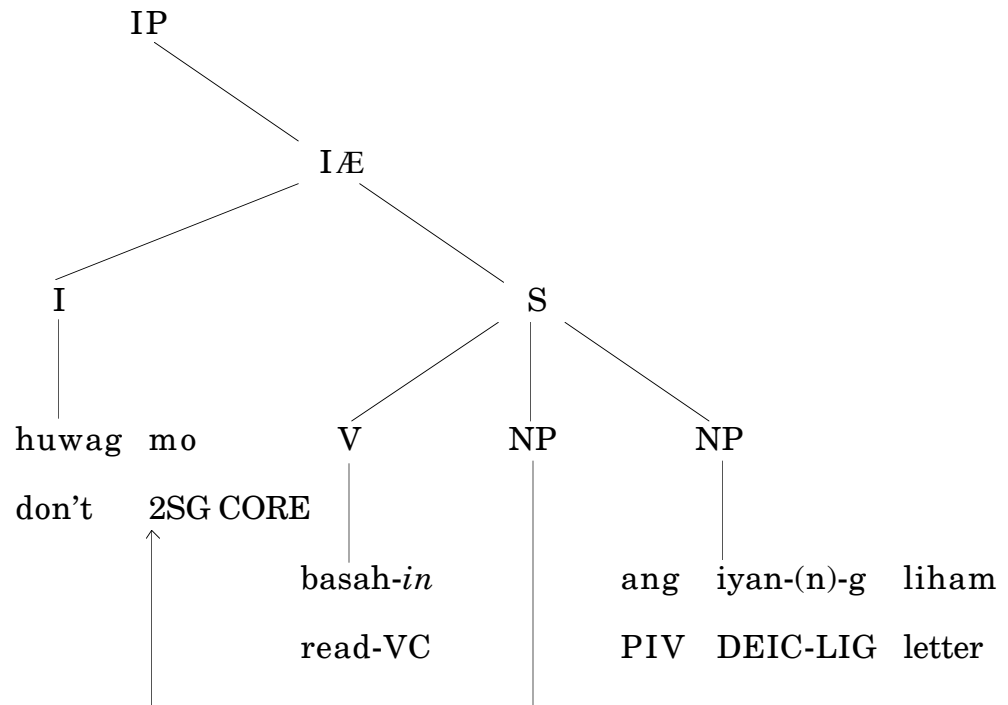
(61)



The crucial distinction here is between the I constituent which projects the IP constituent and then V which projects the S. The I corresponds to what corresponds to the set of AUXiliaries in the language, and the V, the normal predicating verb (Schachter and Otnes 1972: 523):

- (62) (a) huwag mo-ng basah-*in* iyan-(n)g liham
 don't 2SG CORE-LIG read-VC DEIC-LIC letter
 'Don't read that letter.'

(b)

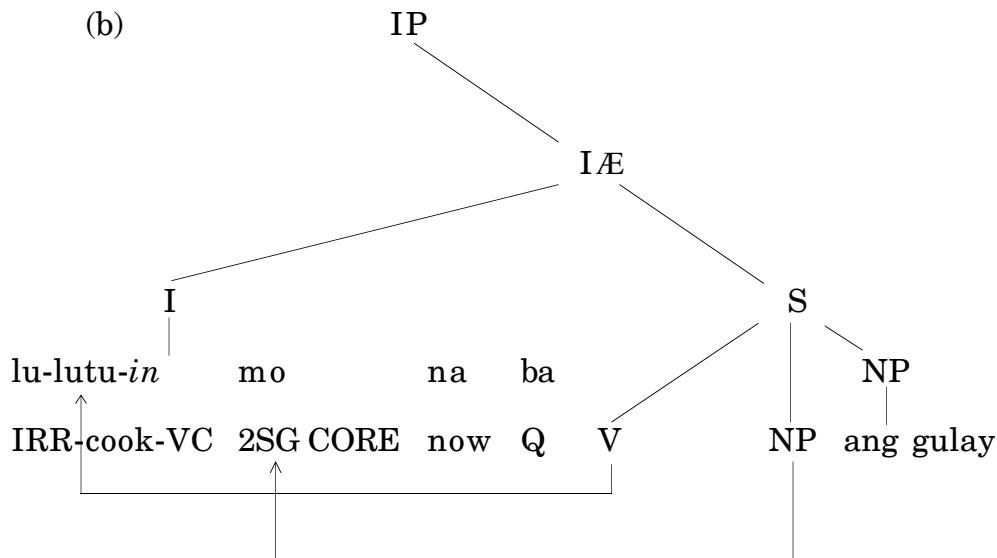


Note that the clitic placement rule is that they immediately follow the elements in I. This is true of all enclitics in Tagalog, not just pronominals (Schachter and Otnes 1972: 523):

- (63) huwag mo kaya-ng tawag-in ang duktur
 don't 2SG CORE perhaps-LIG call-VC PIV doctor
 'Perhpas I should call the doctor.'

If there is no overt auxiliary, the verb occupies the I position and clitics follow it:

- (64) (a) lu-lutu-*in* mo na ba ang gulay?
 IRR-cook-VC 2SG CORE now Q PIV vegetables
 'Will you be cooking the vegetables now?'



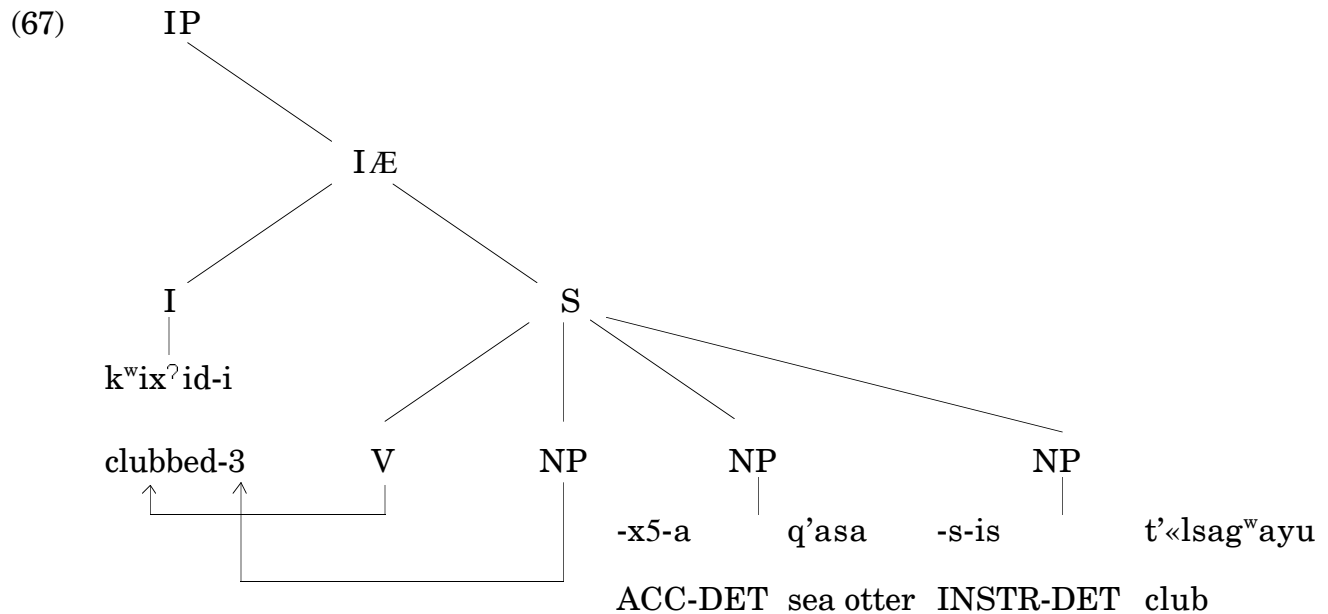
Based on the description and examples presented in Boas (1947), Levine (1980a, b) and Anderson (1984), it seems highly likely that Kwakwala clauses have much the same constituency. Pronominals encliticize to AUXiliaries preceding the predicator verb (Anderson 1984):

- (65) la-i ax^{5?}id-i-da c'«daqa-x5-a 'u[?]«lq^{w?}i
 AUX-3 take-PIV-DET woman-ACC-DET dishes
 'Then the woman takes the dishes.'

But if there is no AUXiliary, again the pronominal encliticizes to the clause initial verb:

- (66) k^wix[?]id-i-x5-a q'asa-s-is t'«lwag^wayu
 club-3-ACC-DET otter-INSTR-3SG POSS club
 'He clubbed the sea otter with his club.'

This suggests an analysis much like Tagalog:



The strong parallelism in basic constituency between Tagalog and Kwakwala could be seen to just be a coincidence, but I will argue that it is not, but is linked to the other typological properties of the two languages, precategoryality and symmetrical voice. While the set of AUXiliaries may be small in Tagalog (*huwas* 'don't' may be the only clear case), this is not true of other Philippine languages which may have a large set (see Tung's (1964) description of the Formosan language Tsou). Starosta, Pawley and Reid (1982) insightfully make a major point of the fundamental importance of the AUXiliaries in reconstructing the clausal syntax of Proto-Austronesian. To see the potential importance of this let us quickly look at an Austronesian language with precategoryal roots and obligatory AUXiliaries, Tongan. Unlike Tagalog and Kwakwala, Tongan is largely an isolating language, lacking their

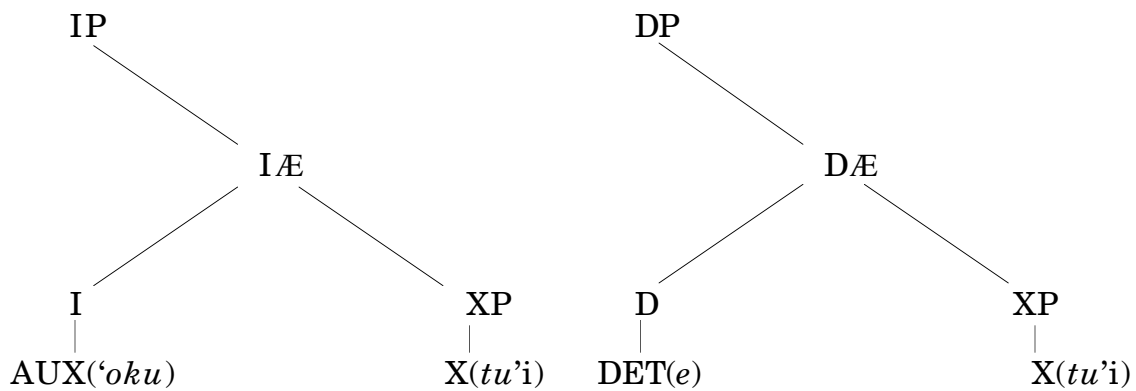
by now familiar morphological patterns of derivations from precategorical roots. Rather what distinguishes between a nominal and a verbal or predicative reading of a precategorical root like *tu'i* 'leading' is the presence of preceding overt tense/aspect AUXiliaries for the verbal/predicative usage and a preceding DETerminer for the nominal (Broschart 1997):

- (68) (a) ko e tu'i o Tonga
 PROP DET lead POSS Tonga
 'The King of Tonga'
- (b) 'oku kei tu'i 'a Sioni
 PRES still lead ABS John
 'John still rules.'

This pattern is found for all roots, even those which denote prototypical entities or events:

- (69) (a) 'oku fefine kotoa e kau lelé
 'PRES woman all DET HUM PL run
 'The ones running are all female.'
- (b) 'oku lele e kau fefiné
 PRES run DET HUM PL woman
 'The women are running.'

What is crucial in Tongan in determining verbal or nominal usage are the phrasal heads TNS/ASP AUXiliary (or I) and DETerminer. Let us argue, following much recent work (e.g. Abney 1987), that they are the heads of their respective phrases, so that the phrases are their projections:



The heads of the XP complements of I and D are precategorial, not specified for any part of speech category like N, V or A. By principles of projection the XP too is unspecified for any word class category; all categorial information about its semantic interpretation as, for example, entity or event, comes from the phrasal head I or D.

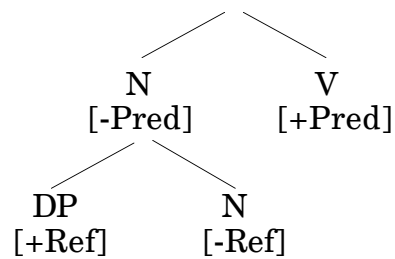
Consider what this claim entails in semantic terms. Since Aristotle, sentences have been analyzed into a predicate, saying something newsworthy about something, and a subject, the thing about which something is said. What makes something newsworthy or at least notable is change or being unexpected, so that predication will be prototypically be associated with roots which denote non-time-stable situations (Givon 1984), e.g. events, actions, happenings, the usual semantics of verbs in languages with a clear verb class. On the other hand, the subject is what is held constant across various predications, newsworthy descriptions we might make of it. This constancy requires descriptive stability and is prototypically linked to roots which describe entities, the defining properties of which are taken to be constant through time: a *kookaburra* is a kind of bird, regardless of whatever time we might talk about it in or whatever predications we might make of it. Thus, through a discourse, the predication about a subject comprises variable information, but the subject itself remains a constant. But the Tongan facts present an immediate

challenge to this description. What are *tu'i* 'lead', *fefine* 'woman' or *lele* 'run'—stable or variable?

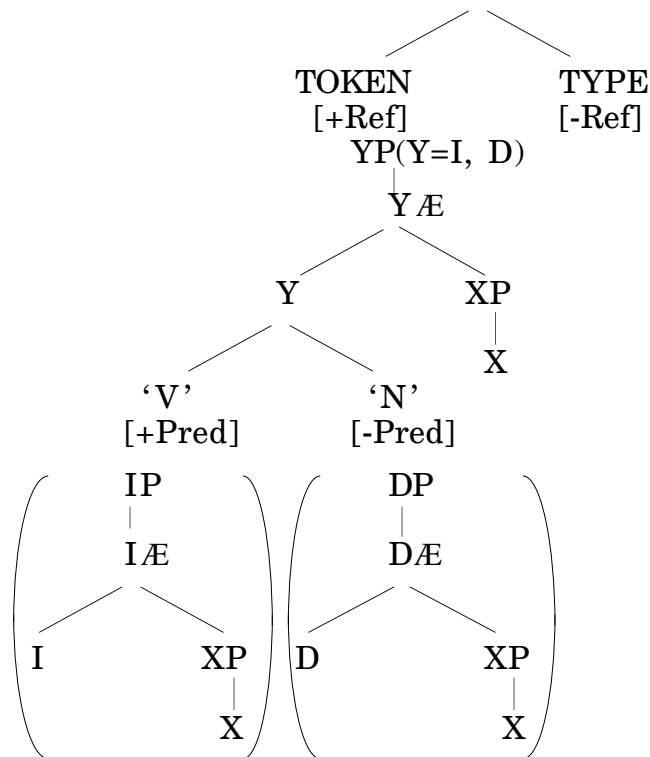
The Tongan facts indicate they can be both, depending on what they combine with. This suggests the notions of changeability or constancy are not in the roots themselves, but in the I or D phrasal heads. Note all roots can be predicates regardless of their semantic descriptions of entity or event, provided they are complements of an I AUXiliary; the notions of changeability [+Pred] versus constancy [-Pred] does not distinguish among them. In view of these facts, Broschart (1997) suggests an alternative to the variable (predicate) versus constant (subject) way of describing clause structure that may be more insightful for Tongan. He points out in predicate logic terms, as opposed to discourse textual ones, the predicate is a constant, the unity of the event described, while the associated arguments are variables, e.g. *give* <x, y, z>; so every act of giving implies someone gives something to some else, but it is absolutely unimportant for the TYPE of event described ('giving') who it is that gives (me, the Buddha, Mikhail Gorbachev) or what is given (a dollar, a mustard seed, a peace prize, etc). Here, of course, the variables have nothing to do with lack of time stability, rather we have a contrast between what is essential and what are possible realizations, or TYPE (act of giving) TOKEN (various participants involved in act of giving). Note that it is the I of IP or D of the DP which fixes the TOKEN of a particular TYPE: *fefine* 'woman' is a TYPE, so that any female human will fall within its denotation, but *e fefine* 'the woman' is a TOKEN, a particular woman presumed uniquely identifiable in the context. *lele* 'run' is a TYPE, any act of running falling within its description, while *'oku lele PRES* run 'is running' is a TOKEN, a specific instantiation of running taking place right now. Broschart (1997) uses the feature [+Ref] (*e fefine*, *'oku lele*) versus [-Ref] (*fefine*, *lele*) to capture this

TYPE/TOKEN distinction (perhaps [\pm Ref] may not be the most perspicacious here, especially in view of its application to roots denoting events like *lele* ‘ran’, but in the absence of any clear alternative I will stick with it here). Note importantly, that roots which denote events and are semantically variable in the textual-discoursal viewpoint are the constants of the predicate logic viewpoint. The verb *eat* is constant in denoting innumerable acts of eating, involving an infinite variety of eaters and things eaten and occurring across the full span of time and space, but individual acts of eating can be instantiated in a particular discourse context via the specification by I within the temporal-reality continuum: *is eating, ate, will eat, might eat*, etc. Roots denoting entities, which are the variables in a predicate logic formalism, are the constants in a textual-discourse framework; the DP *the snake* is constant in the sense that the entity in question always has the properties of a snake and can be taken as such throughout the discourse.

Broschart (1997) makes a fundamental contrast between languages which divide their basic clausal constituents according to the feature [\pm Pred] (noun/verb languages) or [\pm Ref] (TYPE/TOKEN languages). [\pm Pred] languages are the familiar European languages like Latin or Spanish with a clear N/V contrast. [+Pred] corresponds to verbs, those non-time stable roots which are prototypical predicates, and so function in their most morphologically unmarked usage. [-Pred] covers nouns, the denotationally constant roots which are prototypical participants and occur in these functions in their unmarked usage. There may be a secondary [\pm Ref] distinction to contrast *man* from *the man*, but [\pm Pred] is the dominant feature in such languages:



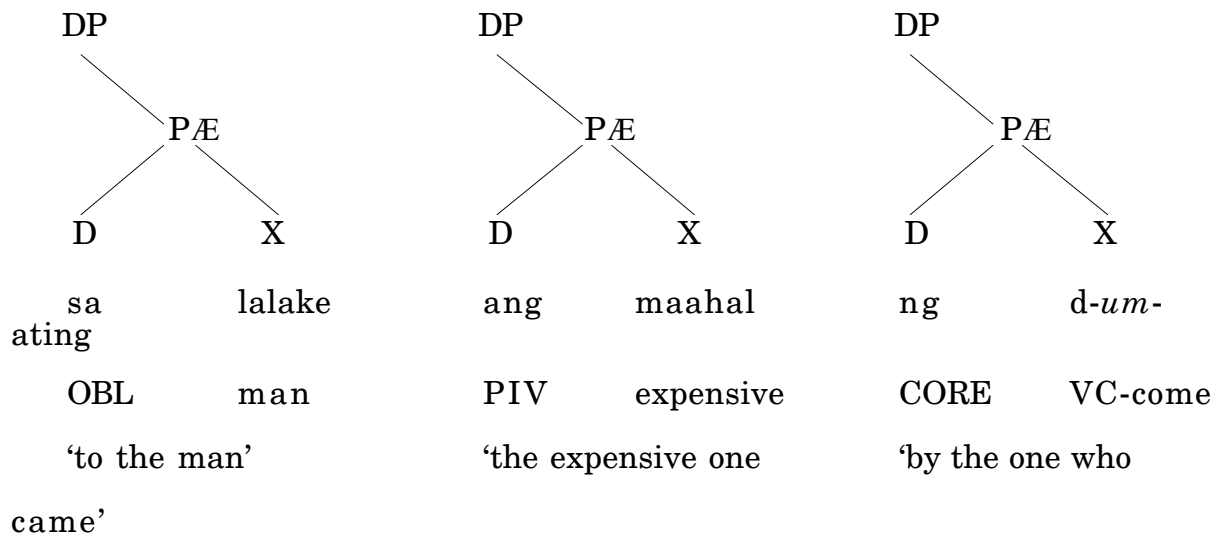
Tongan and such languages lacking the noun/verb contrast have $[\pm\text{Ref}]$ (TYPE: $[-\text{Ref}]$, TOKEN: $[\text{+Ref}]$) as the primary distinction, subsuming both the tense/aspect specification of roots via I in IP and their denotational narrowing by D in DP. This contrasts roots instantiated in a IP or DP projection from those not so instantiated. A secondary contrast then distinguishes those roots governed by I with a projected IP ($[\text{+Pred}]$ because of the lack of time stability expressed by tense/aspect AUXiliaries) from those governed by D with a DP projection ($[-\text{Pred}]$):



Clear cases of noun/verb contrast languages then draw a sharp distinction between inherently predicable, tense-aspect inflectable $[\text{+Pred}]$ roots and those which are not ($[-\text{Pred}]$), while TYPE/TOKEN languages have their primary contrast between syntactic units which are general in denotation and not

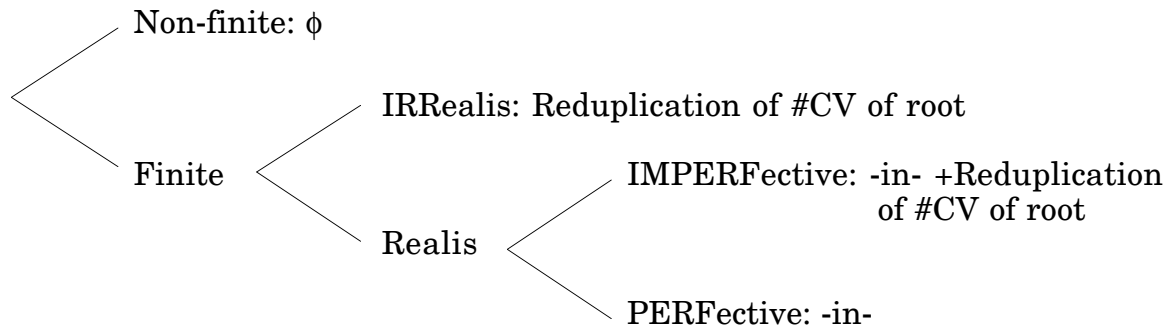
specifically instantiated, [-Ref] roots, and fully instantiated XP phrases ([+Ref]), via I or D heads, which provide these specifics.

How does Tagalog (and by extension, Kwakwala) fit into this typology? Essentially, it is like Tongan, a TYPE/TOKEN language in which the dominant division for syntactic units is [\pm Ref], and this pattern might be diagnostic of Austronesian languages generally. As we have seen, Tagalog NPs are always preceded by a preposition which marks their function, *ang*, *ng*, *sa*, or the distinctions signalled by these prepositions are realized in distinct pronominal forms, e.g. *ako*, *ko*, *akin* 1SG. These prepositions are the D head which projects the DP and transforms the [-Ref] root to a [+Ref] syntactic unit:



The semantics of these Tagalog prepositions strongly supports their analysis as D; it has long been known (Bell 1978; Schachter 1976, 1977; Adams and Manaster-Ramer 1988) that in addition to their case marking properties, they signal notions of referentiality like definiteness and indefiniteness, properties well known crosslinguistically to accrue to the category D. Also, as we have seen, in Kimaragang (Kroeger 1988), there are distinct forms of the case marking prepositions exactly along the lines of referentiality distinctions: PIV: *ih* [+Def], *oh* [-DEF]: CORE: *dih* [+DEF] *do* [-DEF], again strongly indicating they should be analyzed as D.

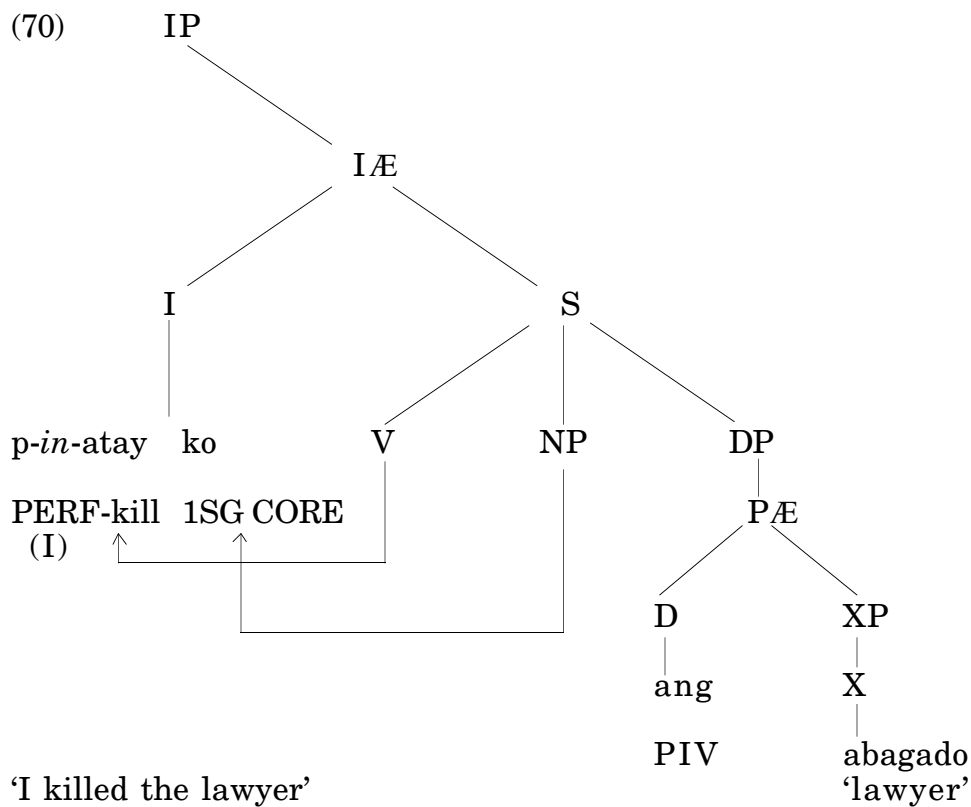
The category I is not as transparent in Tagalog, because unlike isolating Tongan, Tagalog is morphologically rich and words can be overtly inflected for tense/aspect without the need for an I AUXiliary. The basic Tagalog tense aspect system shows some allomorphic variation according to voice derivation, but the basic system is as follows:



So: *bili* ‘buy’

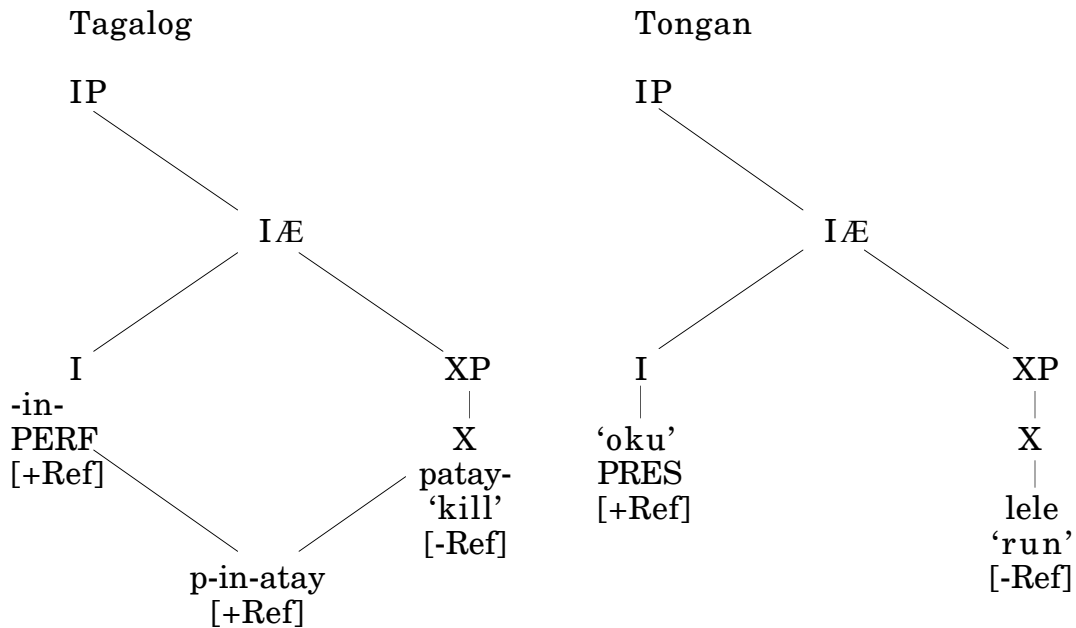
| | | | | |
|--------|-------------|-------------|----------------|---------------|
| NFN | b-um-ili | bilh-in | bilh-an | i-bili |
| IRR: | bi-bili | bi-bilh-in | bi-bilh-an | i-bi-bili |
| IMPER: | b-um-i-bili | b-in-i-bili | b-in-i-bilh-an | i-b-in-i-bili |
| PERF: | b-um-ili | b-in-ili | b-in-ilh-an | i-b-in-ili |

As (61) shows, there is also a structural position for the I in a Tagalog clause which takes an exocentric S complement. If there is an overt I, the AUXiliary *huwag* ‘don’t’, in the I position, the following verb must occur in the non-finite form, which being ϕ , just the voice affix, could arguably be analyzed as not carrying I inflection at all. Verbs inflected for I via tense-aspect marking do not follow AUXilaries and occur in the I position themselves. Thus, unlike Tongan which has no word level units which are [+Ref] because it lacks the requisite morphology, Tagalog does: any word which carries tense-aspect specification is [+Ref] because it is specified for I. In turn, because IP is a projection of I, any word carrying I specification must occur in the clausal I slot:



'I killed the lawyer'

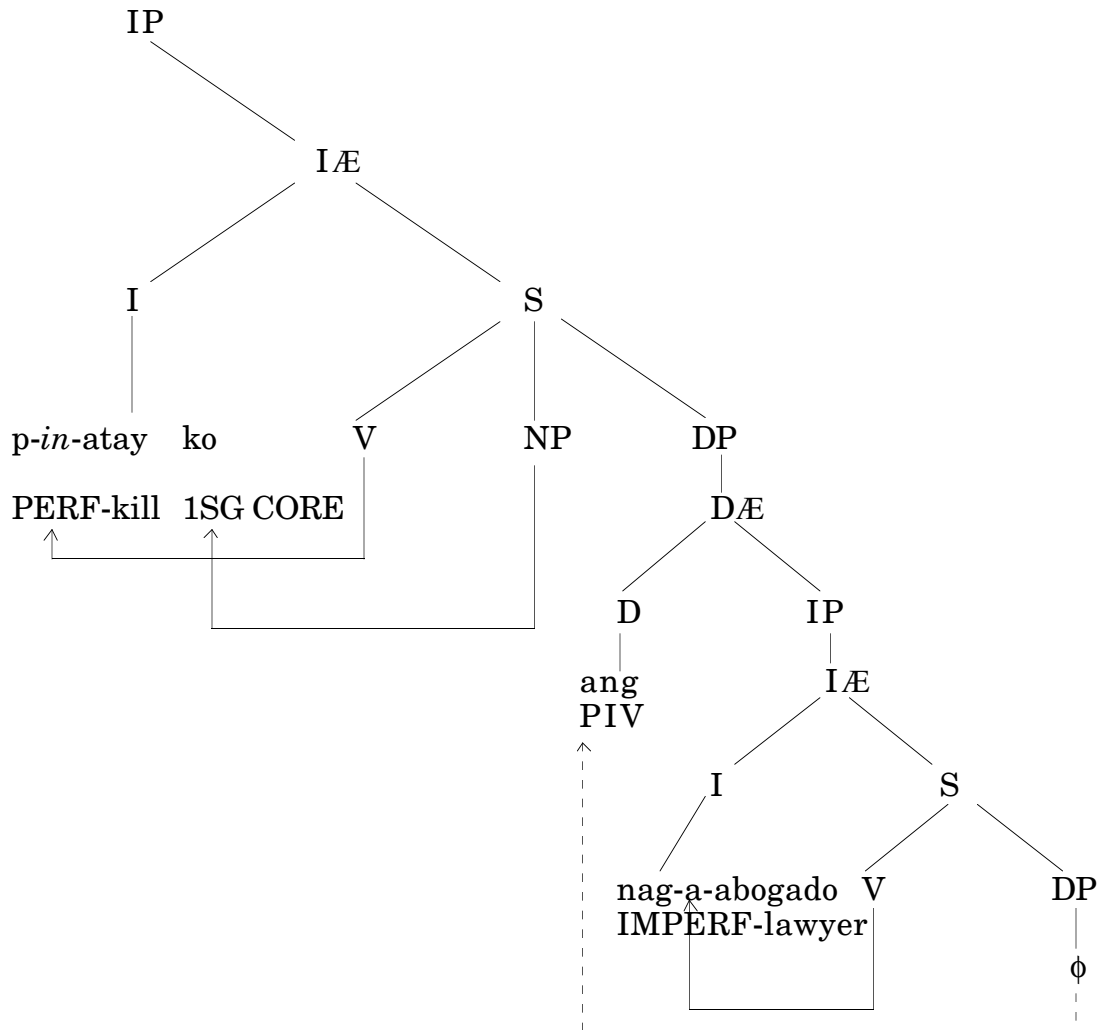
Thus, while the root *patay* 'killing' is precategorial [-Ref] X, and so too is the non-finite form *patay-in* 'kill' <x, y> with the voice affix *-in*, the form inflected for tense-aspect is fully instantiated, via the category I, as [+Ref] albeit realized inflectionally, at the word level, rather than phrasally, as in Tongan.



Now contrast example (70) with (71):

- (71) p-*in*-tay ko ang nag-a-abogado
 PERF-kill 1SG CORE PIV IMPERF-lawyer

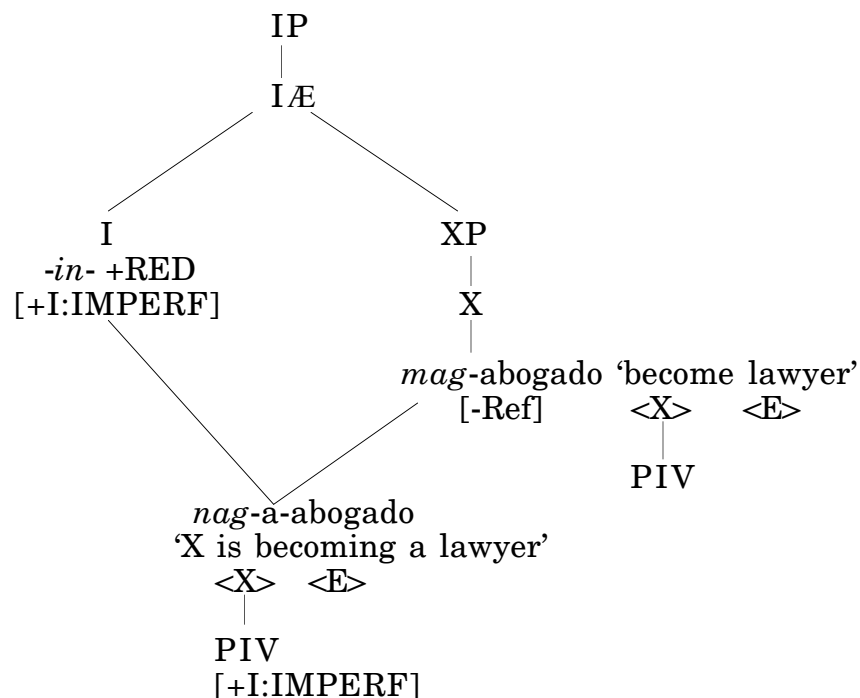
‘I killed the one becoming a lawyer.’



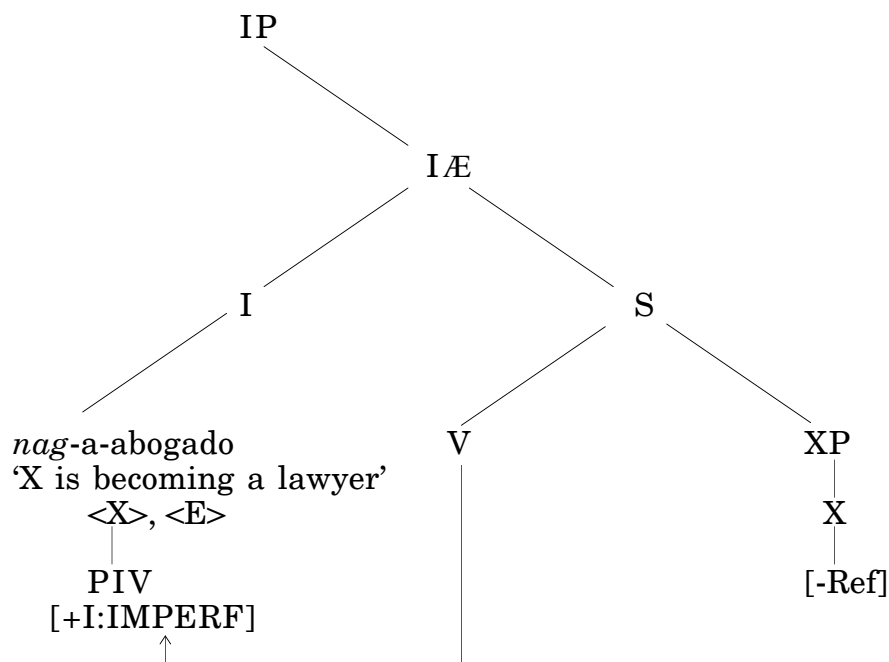
How do we account for the semantic differences between the sentences in (70) and (71)? In both cases we have the [-Ref] root *abogado* ‘lawyering’. In (70) it is referential simply by being the complement of the D *ang*. In Higginbotham’s (1985) terms, we could analyze the D as the instantiating <R> variable for *abogado* and providing its specification:

abogado ‘lawyering’ > *mag*-abogado ‘become lawyer’ <X> <E>
 [-Ref] [-Ref] |
 PIV

This verbal word is still completely [-Ref], uninstantiated, and will describe any act of becoming a lawyer, with any participant, anywhere in the space-time continuum or even imaginary worlds. It is the presence of the <E> variable which unambiguously makes *mag-abogado* a verb. While I claim roots like *abogado* are precatatorial, derived forms like *magabogado* are not. At the word level Tagalog clearly has verbs. The <E> variable indicates that the word describes an event, and events are by definition non-time stable, so that such words are the prototypical forms instantiated by I and predicates [+Pred], all features ascribed to the class of verbs. Note that the semantic difference between *abogado* ‘lawyer’ and *mag-abogado* ‘become a lawyer’ is entirely due to the <E> variable added to the latter via *mag-*. The fact that *abogado* lacks this variable and so cannot be interpreted as an event does *not* make it a noun. It simply lacks this variable and remains precatatorial. If the operation which established *mag-abogado* ‘become a lawyer’ as a verb is the addition of the variable <E>, it would have to be the establishment of the variable <R> which would make *abogado* ‘lawyering’ a noun. But as we have seen, *abogado* on its own lacks the variable <R>; it must be the complement in a DP to a D head (*ang*, etc) for this to be established and specified. Next the form *mag-abogado* ‘become lawyer’ is instantiated via tense-aspect inflection and becomes [+Ref] in the sense of [+I:IMPERF]:



Because the verb form is specified for I, it must occupy the I clausal position; what will remain in the S complement is the single DP filling the X variable in its argument structure and specified by *mag*- as PIVot:



Note that X, the PIVot argument of the verb in I, the one becoming the lawyer, is still [-Ref]. Furthermore, looking back at example (71), we see in the actual

sentence the <X> argument of the verb is missing. In fact, expressing it overtly

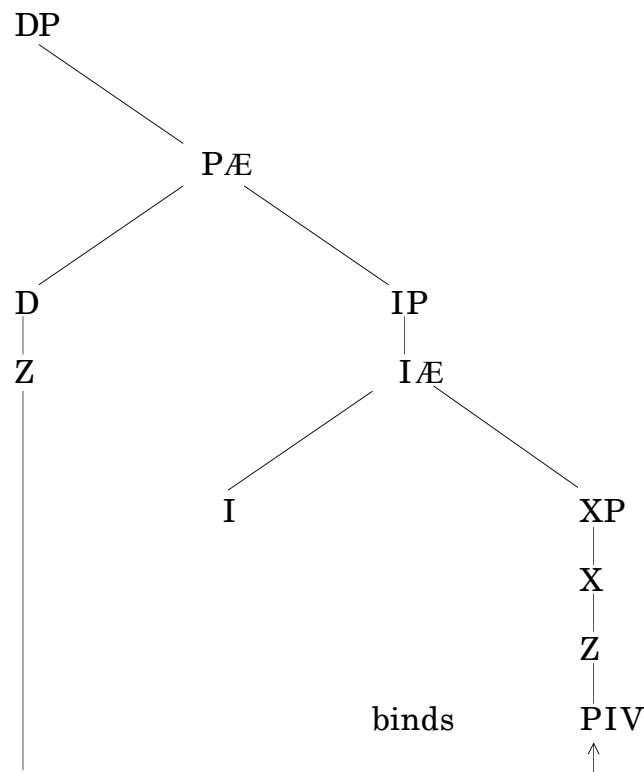
PIV

results in an ungrammatical sentence:

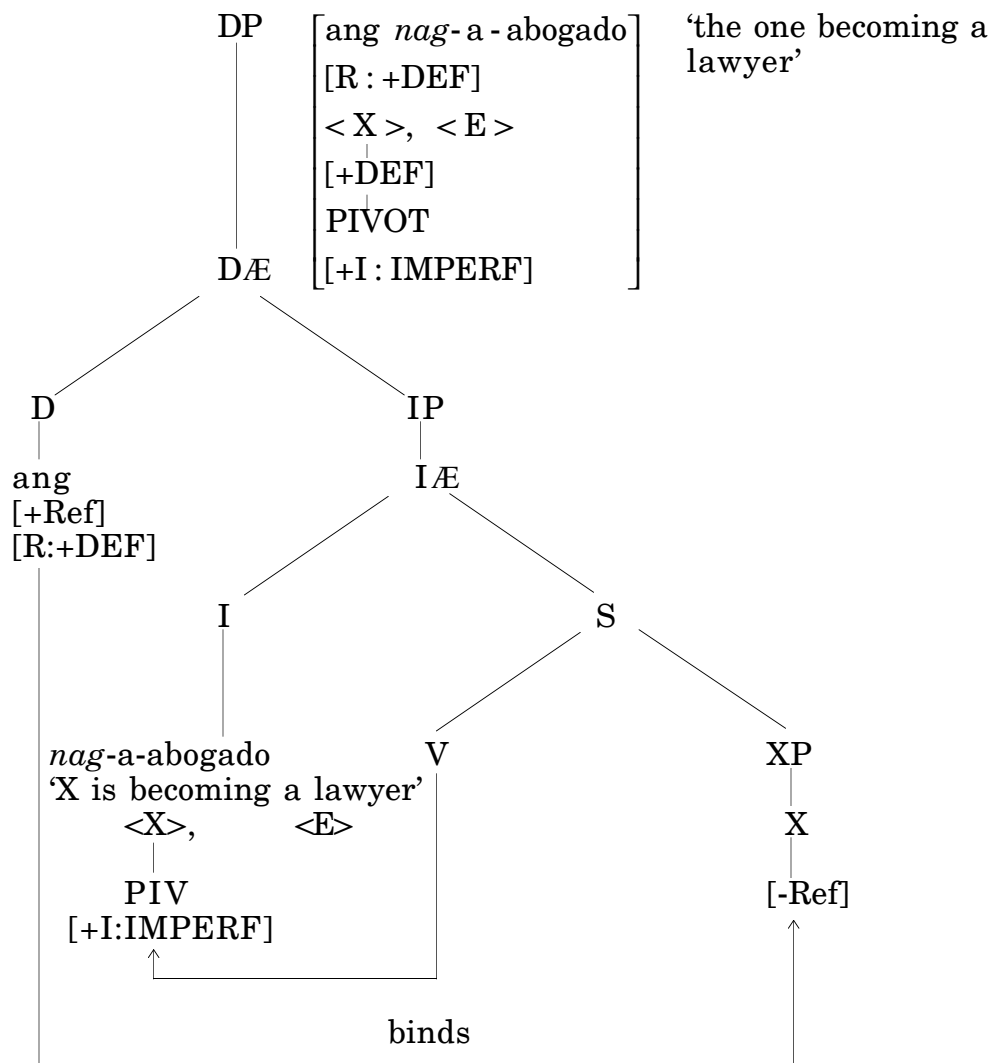
(72) *p-in-atay ko ang nag-a-abogado siya
 PERF-kill 1SG CORE PIV IMPERF-lawyer 3SG PIV

‘I killed him, becoming a lawyer.’

But, of course, the argument structure must be satisfied, so there must be a connection between its suppression and the highest DP projection. Essentially, there is an overall constraint in Tagalog that the D in a DP projection must bind the variable PIVot argument of its IP complement:



Exactly the same constraint is found in Straits Salish, another TYPE/TOKEN language (Jelinck and Demers 1994). Essentially the D head provides the [+Ref] features of the overall DP and because of the binding constraint ensures that the referent of the [+Ref] DP is functioning in the role of the PIV of the verb, which, being bound, is suppressed:



Note that this analysis predicts that DPs can only be formed with IPs as complements if the referent of the entire DP corresponds to the PIVot argument of the I specified verb. This is indeed the case: (the root *kita-* 'see' takes the voice affixes *maka-/ma-* for experience as PIVot versus stimulus as PIVot):

- (73) (a) p-*in*-atay ko ang mga *naka-kita* ko
- PERF-kill 1SG CORE PIV PL PERF-see 1SG CORE
- 'I killed the ones who saw me.'

(b) *p-in-atay ko ang mga na-kita ako
 PERF-kill 1SG CORE PIV PL PERF-see 1SG PIV
 ‘I killed the ones who saw me.’

(73a) is fine because the referent of the DP is the PIVot of the IP complement, ‘the ones who saw me’, with voice affix *maka-* for experiencer as PIVot. (73b) is bad because of referent of the DP is the same, but the PIVot of the IP is *ako* 1SG, ‘the one seen’, indicated by the voice affix *ma-*, stimulus as PIVot.

This analysis of Tagalog has a great deal to recommend it, but let me present just three more arguments in its favor. One of its central claims is that the presence of the <E> variable in a verb form is the result of its derivation by a voice affix and that the root itself lacks such specification and remains [-Ref]. If this is true, a logical consequence is that any form which contains a <E> variable, ie an event semantics, must have first undergone a voice derivation and, specifically, that event nominalizations or gerunds must show such traces. This is indeed the case in Tagalog: the derivation of gerunds is subject to allomorphic variation, but these variations are *directly determined by the allomorphic variation in the voice affixes* (Schachter and Otnes 1972: 160-161):

-um- verbs take *pag-*

| | | | |
|------------|---------|------------|-----------|
| d-um-ating | ‘come’ | pag-dating | ‘coming’ |
| um-alis | ‘leave’ | pag-alis | ‘leaving’ |
| s-um-ulat | ‘write’ | pag-sulat | ‘writing’ |

mag- verbs take *pag-* + Reduplication of #CV

| | | | |
|----------|---------|------------|------------|
| mag-aral | ‘study’ | pag-a-aral | ‘studying’ |
| mag-aral | ‘study’ | pag-a-aral | ‘studying’ |

| | | | |
|-----------|--------|--------------|-----------|
| mag-bigay | ‘give’ | pag-bi-bigay | ‘giving’ |
| mag-luto | ‘cook’ | pag-lu-luto | ‘cooking’ |

maN- verbs take *paN-* + Reduplication of #CV

| | | | |
|--------------|--------------|-------------|-----------------|
| mam-(b)angka | ‘go boating’ | pamambangk | ‘going boating’ |
| man-(t)akot | ‘scare’ | pananakot | ‘scaring’ |
| mang-anak | ‘give birth’ | panganganak | ‘giving birth’ |

ma- verbs take *pag-* + *ka-*

| | | | |
|-----------|---------------------|---------------|-----------------------|
| ma-tunaw | ‘melt’ | pag-ka-tunaw | ‘melting’ |
| ma-buyo | ‘get involved’ | pag-ka-buyo | ‘getting involved’ |
| ma-hirati | ‘become accustomed’ | pag-ka-hirati | ‘becoming accustomed’ |

These data strongly argue that the roots must first be derived as verbs via the voice affixes *-um-*, *mag-* etc, before they are turned into gerunds. There is no other way to account for the clear allomorphic patterns. This only makes sense if the derivation via voice is crucial to the semantics of the gerunds, and, as I have claimed, it is, for the derivation via voice is what adds the <E> variable to the precategorial <-Ref> semantics of the roots, a semantic variable which is crucial to the event interpretation of the gerund. Note there is no evidence in English for the interaction of voice with event nominalization, but this is what would be expected, given the claim that English roots like *give*, *cook*, *melt* etc are already verbs and thereby carry the <E> variable.

Second, because of the constraint that an aspectually inflected verb (i.e. specified for I) must occupy the I head position of the IP, it follows that Tagalog must necessarily lack a constituent VP in such cases. Kroeger (1992)

demonstrates that true phrasal constituents when functioning as predicates present PIVot pronominals either following the whole phrase predicate or enclitic to the head:

- (74) (a)

| | | |
|-------|-----|---------|
| buhat | sa | Maynila |
| from | OBL | Manila |

 siya
_{XP} 3SG.PIV

‘He’s from Manila.’

- (b) buhat siya sa Maynila
 from 3SG PIV OBL Manila

‘He’s from Manila.’

- (75) (a) _{XP} [opisyal sa hukbo] siya
 officer OBL army 3SG PIV

‘He’s an officer in the army.’

- (b) opisyal siya sa hukbo
 officer 3SG PIV OBL army

‘He’s an officer in the army.’

- (76) (a) _{XP} [takot sa kulog] siya
 fear OBL thunder 3SG PIV

‘He’s afraid of the thunder.’

- (b) takot siya sa kulog
 fear 3SG PIV OBL thunder

‘He’s afraid of the thunder.’

But if the seeming head is an inflected verb, this alternation is no longer possible: the enclitic PIVot pronoun must immediately follow the verb:

- (77) (a) *um-akyat* siya sa puno
 VC-climb 3SG PIV OBL tree

‘He climbed the tree.’

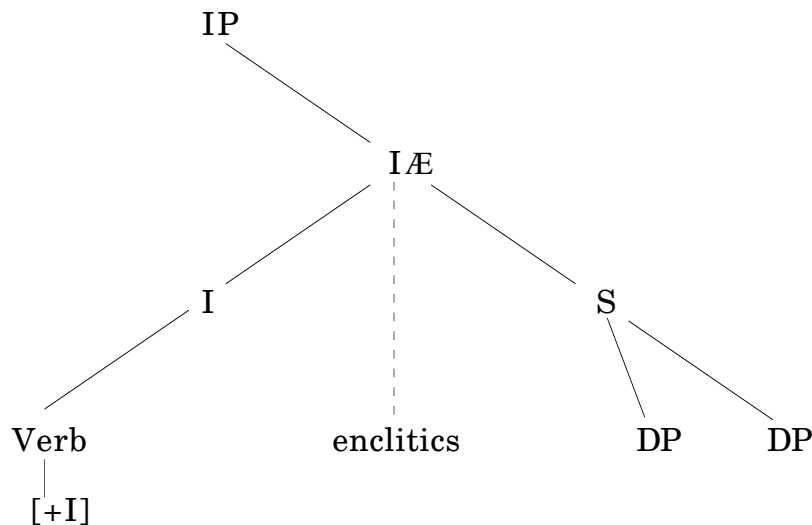
| | | | | |
|---------------------|-----------|-----|-------|---------|
| (b) * _{XP} | [um-akyat | sa | puno] | siya |
| | VC-climb | OBL | tree | 3SG PIV |

| | | | | |
|----------|----------|---------|------|-------|
| (78) (a) | b-um-ili | siya | ng | gatas |
| | VC-buy | 3SG.PIV | CORE | milk |

‘He bought some milk.’

| | | | | |
|---------------------|-----------|------|--------|---------|
| (b) * _{XP} | [b-um-ili | ng | gatas] | siya |
| | VC-buy | CORE | milk | 3SG PIV |

The marked contrast between the examples in (74)-(76) with a non-verbal head and those in (77) and (78) with a verbal head strongly argues that Tagalog does indeed lack a phrasal VP projection containing a verb head and associated DP complements. This is in fact predicted by the analysis proposed here which stipulates that an aspectually inflected verb is specified for I and must therefore occupy the I head position of the projected IP and cannot therefore form an immediate constituent with its subcategorized DP arguments:



Finally, the analysis claims that a [-Ref] root denoting entities lacks an inherent <R> variable and that this is provided by the D head in the DP phrase. This entails a claim that an entity denoting word can only be referential when it is the complement of a D head. In fact Tagalog entity denoting words always

occur with a D head except in one construction, the possessive/existential structure with *may*, *maynoon* ‘there is/have’ and *wala* ‘there isn’t/don’t have’:

- (79) (a) *may* *pera* *si* *Juan*
 money PROP PIV *John*
 ‘John has money.’

- (b) *may* *pera* *sa* *mesa*
 money OBL *table*
 ‘There is money on the table.’

Note that *pera* ‘money’ is necessarily indefinite in these structures, pragmatically non-referential. If it is to taken as referential, a different structure involving a DP must be employed:

- (80) (a) *kay* *Juan* *ang* *pera*
 PROP OBL *John* PIV *money*
 ‘John has the money.’

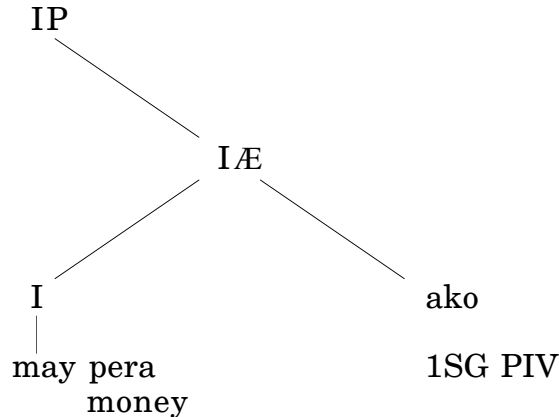
- (b) *na-sa* *mesa* *ang* *pera*
 VC-OBL *table* PIV *money*
 ‘The money is on the table.’

The crucial question is what is the syntactic status of *pera* in sentences like those of (79). The placement of enclitics is a valuable diagnostic clue here; note that they can occur following *pera*:

- (81) (a) *may* *pera* *ba* *si* *Juan*
 money Q PROP PIV *John*
 ‘Does John have money?’

- (b) *may* *pera* *ako*
 money 1SG PIV
 ‘I have money.’

This demonstrates that *may pera* can be taken as a constituent, and given the placement rule for enclitics that they follow the first constituent dominated by IP, we must analyze it as the I head:



Thus, *pera* remains [-Ref] precategoryal in these existential/possessive constructions, part of a complex predicate with *may* as the head. Conclusive evidence that this analysis is correct is provided by a nearly synonymous alternative construction, in which *pera* ‘money’ overtly functions as the [-Ref] root of a derived fully inflected verb involving the voice affix *mag-* and the derivational prefix *ka-*

- (82) nag-ka-ka-pera ako
 IMPERF.VC-PER-money 1SG PIV
 ‘I’m coming into some money.’

The syntactic status of *pera* ‘money’ is much the same in both (81b) and (82), a precategoryal [-Ref] root. In both cases it occupies the I head position by virtue of a larger *verbal* syntactic structure, headed by *may* in (81b) and aspect inflection (I) in (82). Thus, as our analysis predicts, there are no examples of [-Ref] precategoryal roots ever functioning as head of phrases; roots can only be referential, i.e. linked to the variable <R>, when they function as complements to D heads. The fact that indefinite [-Ref] possessed entities are actually assigned to an I constituent shows the importance of the constraint in the overall typology of the language and is further strong support for the

precategorial analysis of Tagalog roots proposed here. It is hard to imagine why a language would make such a sharp structural break between truly indefinite non-referential expressions and specific or definite referential ones unless the wider structural patterns of the language mandated it. It is of course, our contention that the basic patterns on which Tagalog clause structures are built do in fact require such a split.

5.0 Conclusion

In this paper I have argued that the Philippine voice system belongs to neither of the familiar active-passive or ergative-absolutive type, but belongs to a third type, the symmetrical voice system. Diagnostic properties of the symmetrical voice type include: lack of an unmarked verbal form, all forms being equally morphologically derived; the ability of non-subcategorized participants like locatives or instrumentals to freely assume pivot or subject status via their unique voice marking affixes; and the CORE status of actor and undergoer arguments regardless of the voice type chosen, in marked contrast to passive or antipassive constructions which present the actor or undergoer respectively as oblique constituents. Symmetrical voice in turn was linked to the precategorial property of roots, the lack of a noun/verb distinction. A parallel constellation of properties was found in the Wakashan language Kwakwala, strongly arguing that this was not a random assemblage of traits, but a coherent structural type with interrelated properties. It was then demonstrated that the precategorial [-Ref] roots actually get instantiated [+Ref] through being complements to heads D or I which assign the crucial categorial semantic features of <R> or <E> respectively. Unlike languages with a noun/verb distinction, the roots lack these semantic features which are ultimately definitional of the noun/verb distinction itself (<R> for nouns and <E> for verbs). Predication is essentially a function of the <E> variable: lack of time stability and hence newsworthiness is

clearly entailed by eventhood. It is this semantic fact which is the key to the link between precategoriality and symmetrical voice. Predication is always predication of something, so verbs always have a relational component to their meaning, ‘predication about X’, and the unmarked choice for the role of X is the unmarked voice choice: e.g. predication of actor (active; unmarked) versus predication of undergoer (passive: marked) in active-passive languages, with the opposite choice in ergative-antipassive languages. But if the language lacks basically predicated roots, i.e. is precategorial, then, of course, it must lack a unmarked predicated orientation. If all predicated forms are derived from neutral roots by morphological means, then the predicated orientation, i.e. what participant role is being predicated of, will also be so derived—in short, a symmetrical voice system. Furthermore, because there was no unmarked predication orientation in the first place, there is no need to alter the syntactic status of other participants; they can remain CORE or OBLique depending on their inherent semantics and relation to the meaning of the root. Thus, it would seem to be in the nature of universal grammar that symmetrical voice systems require precategoriality. The reverse correlation, however, does not hold: languages with precategoriality of roots exist without symmetrical voice systems: Tongan is a good example, which lacks voice as a grammatical category altogether. The addition of symmetrical voice systems as a third voice system type along with active-passive and ergative-antipassive opens up new possibilities for understanding voice as a grammatical category. Only further research can tell us just how useful an addition it will turn out to be.

References

- Abney, S. 1987. The English noun phrase in its sentential aspect. Unpublished PhD dissertation, Massachusetts Institute of Technology.
- Abrams, N. 1961. Word base classes in Bilaan. *Lingua* 10. 391-402.
- Adams, K. and A. Manaster-Ramer. 1988. Some questions of topic/focus choice in Tagalog. *Oceanic Linguistics* 27. 79-102.
- Anderson, S. 1984. Kwakwala syntax and the government-binding theory. *Syntax and Semantics* 16. 21-75. New York: Academic Press.
- Bell, S. 1978. Two differences in definiteness in Cebuano and Tagalog. *Oceanic Linguistics* 17. 1-11.
- Blake, F. 1906. The expression of case of the verb in Tagalog. *Journal of the American Oriental Society* 27. 183-189.
- Boas, F. 1911. Kwakiutl. In F. Boas, ed., Handbook of American Indian languages. *Bureau of American Ethnology Bulletin* 40, 423-557. Washington: Smithsonian Institution.
- Boas, F. 1947. Kwakwala grammar with a glossary of the suffixes. *Transactions of the American Philosophical Society* 37. 203-377.
- Boas, F. 1966[1911]. Introduction to The Handbook of American Indian Languages. Lincoln: University of Nebraska Press.
- Bresnan, J., ed. 1982. The mental representation of grammatical relations. Cambridge, MA: MIT Press.
- Broschart, J. 1991. Noun, verb and PARTICIPATION. A typology of the noun/verb-distinction. In H. Seiler and N. Premper, eds., Partizipation. Das Sprachliche Erfassen von Sachverhalten. Tübingen: Narr.
- Broschart, J. 1997. Why Tongan does it differently: categorial distinctions in a language without nouns or verbs. *Linguistic Typology* 1. 123-165.

- Cena, R. 1977. Patient primacy in Tagalog. Paper presented at the Annual Meeting of the Linguistic Society of America. Chicago.
- Cooreman, A. 1988. The antipassive in Chamorro: variations on the theme of transitivity. In M. Shibatani, ed., *Passive and voice*, 561-593. Amsterdam: Benjamins.
- De Guzman, V. 1978. Syntactic derivation of Tagalog verbs. *Oceanic Linguistics Special Publications* 16. Honolulu: University of Hawai'i Press.
- De Guzman, V. 1979. Morphological evidence for primacy of patient as subject in Tagalog. Paper presented at the Annual Meeting of the Linguistic Society of America. Los Angeles.
- Dixon, R. 1972. The Dyirbal language of North Queensland. Cambridge: Cambridge University Press.
- Drossard, W. 1984. Das Tagalog als Repräsentant des aktiven Sprachbaus. Tübingen: Narr.
- Egerod, S. 1965. Verb inflection in Atayal. *Lingua* 15. 251-282.
- Egerod, S. 1966. Word order and word classes in Atayal. *Language* 42. 346-369.
- Foley, W. and R. Van Valin. 1984. Functional syntax and universal grammar. Cambridge: Cambridge University Press.
- Gerdts, D. 1988. Antipassives and causatives in Ilokano: evidence for an ergative analysis. In P. McGinn, ed., *Studies in Austronesian linguistics*, 295-321. Athens, OH: Ohio University Press.
- Gil, D. 1993. Tagalog semantics. *Berkeley Linguistics Society* 19. 390-403.
- Givon, T. 1984. Syntax. A functional typological introduction. Amsterdam: Benjamins.
- Guilfoyle, E., H. Hung and L. Travis. 1992. SPEC of IP and SPEC of VP: two subjects in Austronesian languages. *Natural Language and Linguistic Theory* 10. 375-414.

- Higginbotham, J. 1985. On semantics. *Linguistic Inquiry* 16. 547-593.
- Himmelman, N. 1991. The Philippine challenge to universal grammar.
Institute for Linguistics, University of Köln Working Papers (n.s.) 15.
- Huang, L. 1993. A study of Atayal syntax. Taipei: Crane Publications.
- Jacobson, W. 1979. Noun and verb in Nootkan. *British Columbia Provincial Museum Heritage Records* 4. 83-155.
- Jelinek, E. and R. Demers. 1994. Predicates and pronominal arguments in Straits Salish. *Language* 70. 697-736.
- Kinkade, D. 1983. Salish evidence against the universality of 'noun' and 'verb'.
Lingua 66. 25-46.
- Kroeger, P. 1988. Verbal focus in Kimaragang. *Pacific Linguistics* A78. 217-240.
- Kroeger, P. 1990. Stative aspect and unaccusativity in Kimaragang Dusun.
Oceanic Linguistics 29. 110-131.
- Kroeger, P. 1992. Phrase structure and grammatical relations in Tagalog.
Stanford: Center for the Study of Language and Information.
- Kroeger, P. 1996. The morphology of affectedness in Kimaragang Dusun.
Pacific Linguistics A84. 33-50.
- Kuipers, A. 1968. The categories verb-noun and transitive-intransitive in English and Squamish. *Lingua* 21. 610-626.
- Levine, R. 1980. On the lexical origin of the Kwakwala passive. *International Journal of American Linguistics* 46. 240-258.
- Levine, R. 1980. Passives and controllability in Kwakwala. *Glossa* 14. 139-167.
- McGinn, R. 1988. Government and case in Tagalog. In R. McGinn, ed., *Studies in Austronesian linguistics*, 275-293. Athens, OH: Ohio University Press.
- Naylor, P. 1995. Subject, topic and Tagalog syntax. In D. Bennett, T. Bynon and G. Hewitt, eds., *Subject, voice and ergativity: selected essays*. London: School of Oriental and African Studies, University of London.

- Ramos, T. 1974. The case system of Tagalog verbs. *Pacific Linguistics* B27.
- Ramos, T. and M. Bautista. 1986. Handbook of Tagalog verbs: inflections, modes and aspects. Honolulu: University of Hawai'i Press.
- Rau, D. 1992. A grammar of Atayal. Taipei: Crane Publications.
- Rhea, M. 1972. Prefocus and verbal orientation in Sarangani Bilaan. *Philippine Journal of Linguistics* 2. 35-42.
- Sapir, E. and M. Swadesh. 1946. American Indian grammatical categories. *Word* 2. 103-112.
- Sasse, H. 1993. Das Nomen - eine universale Kategorie? *Sprachtypologie und Universalienforschung* 46. 187-221.
- Schachter, P. 1976. The subject in Philippine languages: topic, actor, actor-topic or none of the above. In C. Li, ed., Subject and topic, 491-518. New York: Academic Press.
- Schachter, P. 1977. Reference-related and role-related properties of subjects. *Syntax and Semantics* 8. 279-306. New York: Academic Press.
- Schachter, P. and F. Otanes. 1972. Tagalog reference grammar. Berkeley: University of California Press.
- Shibatani, M. 1988. Voice in Philippine languages. In M. Shibatani, ed., Passive and voice, 85-142. Amsterdam: John Benjamins.
- Starosta, S., A. Pawley and L. Reid. 1982. The evolution of focus in Austronesian. *Pacific Linguistics* C75. 145-170.
- Swadesh, M. 1939. Nootka internal syntax. *International Journal of American Linguistics* 9. 77-102.
- Tung, T. 1964. A descriptive study of the Tsou language, Formosa. Institute of History and Philology Special Publications 48. Taipei: Academica Sinica.
- Walton, C. 1986. Sama verbal semantics: classification, derivation and inflection. Manila: Linguistic Society of the Philippines.

- Wolff, J. 1974. Verbal inflection in Proto-Austronesian. In A. Gonzales, ed.,
Parangal kay Cecilio Lopez, 71-91. Philippine Journal of Linguistics
Special Monographs 4. Quezon City: Linguistic Society of the Philippines.
- Wolff, J. 1979. Verbal morphology and verbal sentences in Proto-Austronesian.
In P. Naylor, ed., Austronesian studies: papers from the Second Eastern
Conference on Austronesian languages, 153-168. Ann Arbor: Center for
South and Southeast Asian Studies, University of Michigan.