

Austronesian Nominalism and its consequences: A Tagalog case study*

DANIEL KAUFMAN

‘From thought we say think, but from justice
we can say nothing’

—Democritus

1.0. Introduction: the Austronesian voice system

The Austronesian languages comprise a vast family with some 1,268 members spanning half of the globe, from Rapanui (Easter Island) on its Eastern boundary, to Malagasy (Madagascar) on its Western boundary, and from the Aboriginal languages of Taiwan (known as ‘Formosan languages’) on its Northern boundary to Maori (New Zealand) in the

* I thank audiences at AFLA XV, the CUNY Graduate Center and NYU where I presented parts of the present paper. A special debt of gratitude goes to the organizers of Beyond Focus Austronesian workshop at the Zentrum für Allgemeine Sprachwissenschaft whose invitation prompted me to begin thinking about these issues more seriously. I also thank, Edith Aldridge, Marcel den Dikken, Manfred Krifka, John Whitman, Malcolm Ross, Nikolaus Himmelmann and Paul Kroeger for helpful discussion and critiques, none of whom, needless to say, should share any blame for the ideas presented here. Finally, I express a great debt of gratitude to Myrna Galvan for patiently sharing all her intuitions and judgments.

A note on transcription: I employ here a variant of the transcription system introduced in Wolff, Centeno & Rau (2005) where long vowels are marked by an acute accent and the final glottal stop is marked by a grave accent. The genitive case marker, written as *ng* in Filipino orthography, is transcribed here as pronounced, i.e. *nang*.

Abbreviations: AV – actor voice, BEG – begun aspect, CV – conveyance voice, DAT – dative, DEP – dependent, DST – distributive, GEN – genitive case, IMPRF – imperfective, LIM – limitative, LNK – linker, LOC – locative, LV – locative voice, NOM – nominative (default) case, OBL – oblique case, PL – plural, POSS – possessive, PST – past, PV – patient voice, RSLT – resultative, STA – stative.

South. These languages are as diverse as they are widespread. In some areas we find predicate initial languages with a complex voice system that has been described variously as patterning along ergative, symmetric and accusative lines (Philippines, Taiwan); in other areas we can find highly isolating SVO languages with no voice morphology to speak of (parts of East Timor and East Indonesia); in yet other areas we find agglutinating nominative-accusative SOV languages with person agreement (Papua New Guinea). Despite this great variety, we can now say quite definitively which areas are morphosyntactically conservative and which areas are innovative. Beginning with Dempwolff (1934–38), much has been learnt about the origin and spread of these languages through rigorous use of the comparative method. It is now certain that the Austronesians originated in present day Taiwan and expanded southwards, populating the Philippines and then the Indo-Malaysian archipelago from which they later expanded eastwards into the Pacific. The Formosan languages comprise anywhere from three (Sagart 2004) to nine (Blust 1999) primary subgroups of the Austronesian family with all other Austronesian languages outside of Taiwan belonging to a single primary subgroup, termed Malayo-Polynesian (Blust 1977). It is now well accepted that most of the Formosan languages, together with the Philippine languages, preserve aspects of Austronesian morphosyntax which have disappeared in most languages outside of these areas (Ross 2002, forthcoming, Blust 2002 and references therein). Most notable of these features is the complex voice system, which has given rise to much debate and discussion in both typological and syntactic circles. The relevant portion of the reconstructed voice paradigm is shown in (1).

(1) Proto-Austronesian voice morphology (Ross 2002, Wolff 1973)

Agent Voice	*⟨ <i>um</i> ⟩
Patient Voice	*- <i>en</i>
Locative Voice	*- <i>an</i>
Conveyance Voice	* <i>Si</i> -

Each voice selects its corresponding argument (or adjunct) as the subject of the clause as illustrated in (2) with Tagalog. The subject is marked with the case marker *ang*, which I gloss here as NOMINATIVE (without implying the system is nominative-accusative). In (2)a, an actor voice clause, we see the patient marked with genitive case, the locative phrase marked

with oblique case and a benefactive phrase argument marked with a preposition plus oblique case; in (2)b, patient voice corresponds to nominative case on the patient; in (2)c, locative voice corresponds to nominative case on the locative, and in (2)d, the so called conveyance voice corresponds to nominative case on the benefactive.¹ The meanings of all four sentences are essentially equivalent but differ in the interpretation of definiteness and effectedness of the arguments, most notably in that the object of an actor voice form is interpreted as indefinite while the nominative phrase is interpreted as definite.²

- (2) a. k<um>áin nang=dagà sa=pinggan pára sa=áso **ang=púsa**
 <AV:BEG>eat GEN=rat OBL=plate for OBL=dog NOM=cat
 ‘The cat ate a rat on the plate for the dog.’
- b. k<in>áin-Ø nang=púsa **ang=dagà** sa=pinggan pára sa=áso
 <BEG>eat-PV GEN=cat NOM=rat OBL=plate for OBL=dog
- c. k<in>áin-an nang=púsa nang=dagà **ang=pinggan** pára
 <BEG>eat-LV GEN=cat GEN=rat NOM=plate for
 sa=áso
 OBL=dog
- d. i-k<in>áin nang=púsa nang=dagà sa=pinggan **ang=áso**
 CV-<BEG>eat GEN=cat GEN=rat OBL=plate NOM=dog

The most celebrated aspect of the complex voice system is its interaction with extraction, specifically, question formation, topicalization and relativization. The constraint, which is quite general to languages which preserve the original voice system and even many which have simplified it, is often referred to as the “subjects-only” restriction on extraction (Keenan 1976). It is exemplified with Tagalog in (3)–(6) (based on Richards 2000). In (3), the argument selected by the voice morphology is the agent and hence it can be questioned in the apparent cleft-like structure required for argument interrogatives (i.e. *who*, *what*, *which*). Extracting the agent from a patient voice clause, as in (3)b is ungrammatical.

¹ The conveyance voice (in the terminology of Wolff 1973) corresponds to a wide range of meanings including benefactive, instrumental and objects moving away from speaker (as in the themes of actions such as ‘give’, ‘push’, ‘throw’, etc.).

² It is rarely mentioned that in Tagalog oblique phrases must also be interpreted as definites. It is not clear how far this pattern generalizes to related languages beyond Tagalog and should thus probably not be taken as a property of Philippine-type voice systems.

Likewise in (4)a, extraction of the patient from a patient voice clause is seen to be legitimate but extraction of the same argument from an actor voice clause is ungrammatical, as shown in (4)b.

- (3) a. Sino ang=b⟨um⟩ili nang=téla?
 who NOM=⟨AV:BEG⟩buy GEN=cloth
 ‘Who bought the cloth?’
 b. *Sino ang=b⟨in⟩ili-∅ ang=téla?
 who NOM=⟨BEG⟩buy-PV GEN=cloth
- (4) a. Ano ang=b⟨in⟩ili-∅ nang=babáe?
 what NOM=⟨BEG⟩buy-PV GEN=woman
 ‘What did the woman buy?’
 b. *Ano ang=b⟨um⟩ili ang=babáe?
 what NOM=⟨AV:BEG⟩buy NOM=woman

Topicalization, which does not require the apparent cleft-like structure used for argument interrogatives, is subject to the same constraint. In (5)a, the actor of an actor voice clause is legitimately topicalized to the left periphery followed by the topic marker *ay*. In (5)b, we see that the topicalization of the patient from the same type of clause is ungrammatical.

- (5) a. Ang babáe ay b⟨um⟩ili nang=téla
 NOM=woman TOP ⟨AV:BEG⟩buy GEN=cloth
 ‘The woman, bought cloth.’
 b. *Nang=téla ay b⟨um⟩ili ang=babáe
 GEN=cloth TOP ⟨AV:BEG⟩buy NOM=woman

Finally, the examples in (6) show that the same constraint holds for relativization. Relative clauses are typically head initial or predicate initial in Philippine languages, with the notional head being connected to the predicate material by the nasal linker. In (6)a, we see the actor of an actor voice clause successfully relativized while in (6)b we see the ungrammaticality of relativizing the patient of an actor voice clause.

- (6) a. Ang=babáe=ng b⟨um⟩ili nang=téla
 NOM=woman=LNK ⟨AV:BEG⟩buy GEN=cloth
 ‘the woman who bought the cloth’
 b. *Ang=téla=ng b⟨um⟩ili ang=babáe
 NOM=cloth=LNK ⟨AV:BEG⟩buy NOM=woman

1.1. *Austronesian voice as nominalization*

The above restrictions have been analyzed in a multitude of different ways, but in broad terms, recent generative analyses appear to have settled on the idea that the ungrammatical extractions result from a locality violation (Richards 2000, Rackowski 2002, Aldridge 2002 et seq, Maclachlan & Nakamura 1997, Pearson 2005, Chung 1998, among others). Abstracting away from differences, these accounts all rely on a derivation in which the subject must always be in a higher position at the point where extraction is triggered, essentially rendering lower DPs invisible. Here, I will sketch out an analysis which does not crucially rely on locality in the sense employed above. Rather, I follow an older intuition which views all predication in conservative Austronesian languages as inherently copular (Lopez 1928, Seiter 1975, De Wolf 1988, Naylor 1980) and all predicates as inherently nominal (Capell 1964, Starosta, Pawley & Reid 1982)³. Cases of bad extraction represent extraction from NP/DP, a highly marked operation cross-linguistically (Horn 1974). *Prima facie* evidence for a nominal interpretation of voice forms can be seen in the interpretations of the independent DPs in (7), corresponding to the four voices seen above.

- (7) a. ang=b<um>ili
 NOM=<AV:BEG>buy
 ‘the buyer/one who bought’
- b. ang=b<in>ili-∅
 NOM=<BEG>buy-PV
 ‘the (thing) bought’
- c. ang=b<in>il-han
 NOM=<BEG>buy-LV
 ‘the (place) bought at’
- d. ang=i-b<in>ili
 NOM=CV-<BEG>buy
 ‘the one bought for’

³ A connection between nouns and verbs had also been made in the earlier Austronesianist literature by van der Tuuk (1864–1867), Adriani (1893) and Scheerer (1924).

If voice marked forms are nominalizations then the sentences in (2) could be built from copular clauses that equate the nominal predicate with the nominative marked subject. On this interpretation, the translations in (8) are a more literal rendering of the Tagalog examples in (2) (simplified).

- (8) a. k<um>áin nang=dagà **ang=púsa**
 <AV:BEG>eat GEN=rat NOM=cat
 ‘The cat was the eater of a rat.’
- b. k<in>áin-∅ nang=púsa **ang=dagà**
 <BEG>eat-PV GEN=cat NOM=rat
 ‘The rat was the eaten one of the cat.’
- c. k<in>áin-an nang=púsa nang=dagà **ang=pinggan**
 <BEG>eat-LV GEN=cat GEN=rat NOM=plate
 ‘The plate was the cat’s eating place of the rat.’
- d. i-k<in>áin nang=púsa nang=dagà **ang=áso**
 CV-<BEG>eat GEN=cat GEN=rat NOM=dog
 ‘The dog was the cat’s “eating benefactor” of the rat’

This interpretation corresponds with analyses of other, possibly more conservative Austronesian languages. Ferrell (1982), for instance, describes the voice forms of Paiwan, a Formosan language, as shown in Table 1.

Starosta, Pawley & Reid (1982) argue that the nominal appearance of Austronesian voice forms stems from a historical reanalysis of nominalizations as canonical predicates (verbs in their analysis). This hypothesis has found additional support from recent historical work (Ross forthcoming) and further typological evidence (Kaufman forthcoming). Here, I follow to its logical conclusion the intuition that the structures in (7) are

Table 1. *The function of Paiwan voice forms (Ferrell 1982: 17, 106, Ross 2002: 38)*

	verb form	nominalization
<i>k<əm>an</i>	actor voice neutral	‘eater’, ‘someone who eats’
<i>kan-ən</i>	patient voice neutral	‘food’, ‘something to be eaten’
<i>k<in>an</i>	patient voice perfective	‘consumed food’, ‘something eaten’
<i>kan-an</i>	location voice neutral	‘place where one eats’
<i>si-kan</i>	circumstantial voice neutral	circumstantial voice neutral

basic, i.e. not derived headless relatives. I take the effect of Austronesian voice morphology to be akin to that of thematic nominalization in English with actor oriented *-er* and patient oriented *-ee* (Marchand 1969, Bauer 1983, Barker 1998). Barker (1998) notes that a verb like *amputate* in English specifies two arguments, a subject, which corresponds to the agent in the amputation event, and a direct object, which corresponds to a body part of an individual which is to be removed. There is no argument of *amputate*, however, which refers to the individual whose limb is amputated (and thus no sentence **The doctor amputated John*). Barker goes on to note that this restricted argument structure does not interfere at all in the formation of *amputee*, which refers precisely to the individual in question:

“... the meaning of the verb amputate guarantees the existence of a person undergoing amputation, even though there is no syntactic argument that corresponds to this participant ... the fact that the person undergoing amputation is a participant of every amputation event is sufficient to enable a set of amputation events to characterize the *-ee* noun amputee: for each amputation event *e*, there exists an individual *x* which is a participant in *e* such that *x* is (becomes) an amputee. Thus amputee is episodically linked to the meaning of amputate despite the fact that there is no corresponding syntactic argument position.” (Barker 1998: 714)

This is precisely the nature of the Austronesian voice markers, which can now be easily related to their nominal characteristics. Any Tagalog lexical root can take any voice so long as the conceptual representation of the root provides for the relevant participant.⁴ This is illustrated with five

⁴ Although I take this to be intrinsically correct, it is also a simplification. Several works have attempted to classify Tagalog roots according to the voice affixes they cooccur with and there has been little agreement on how many classes exist or if such a classification is even justified. Kroeger (1998) makes the interesting claim that “nominal” and “verbal” roots in Tagalog evince different patterns with the voice markers as a result of their derivational verbalizing function with nominal roots. Unfortunately, several of the empirical claims regarding what morphology particular roots can and cannot appear with do not appear correct. Crucially, it must be realized that the voice system is fully productive and can create novel combinations of roots and voices for specialized meanings. Claims of *unpredictable* gaps in the voice paradigm turn out to be either attested or simply cases of blocking by more specific forms (Kaufman 2007). Principled gaps, such as **bigay-in* give-PV, are predictable on the basis of the semantics, parallel to unacceptable *-ee* formations in English.

Table 2. *Some Tagalog voice markers and their interpretation*

	<i>i-</i> CONVEYANCE	<i>-in</i> PATIENT	<i>-an</i> LOCATIVE	$\langle um \rangle$ ACTOR
$\sqrt{t\acute{a}wa}$ [[laugh]]	<i>i-t\acute{a}wa</i> 'laugh about X'	–	<i>taw\acute{a}n-an</i> 'to laugh at X'	<i>t\langle um \rangle\acute{a}wa</i> 'X to laugh'
$\sqrt{b\acute{ig}ay}$ [[give]]	<i>i-b\acute{ig}ay</i> 'give X (theme)'	–	<i>b\acute{ig}y-an</i> 'give to X'	<i>b\langle um \rangle\acute{ig}ay</i> 'X to give out' <i>mag-b\acute{ig}ay</i> 'X to give'
$\sqrt{l\acute{a}kad}$ [[walk]]	<i>i-l\acute{a}kad</i> 'walk X somewhere'	<i>lak\acute{a}r-in</i> 'walk X (a distance)'	<i>lak\acute{a}r-an</i> 'walk to X'	<i>l\langle um \rangle\acute{a}kad</i> 'X to walk'
$\sqrt{p\acute{u}tol}$ [[cut]]	<i>i-p\acute{u}tol</i> 'cut with X'	<i>put\acute{u}l-in</i> 'cut X'	<i>put\acute{u}l-an</i> 'cut from X'	<i>p\langle um \rangle\acute{u}tol</i> 'X to cut'
$\sqrt{s\acute{u}lat}$ [[write]]	<i>i-s\acute{u}lat</i> 'write (with) X'	<i>sul\acute{a}t-in</i> 'to write X'	<i>sul\acute{a}t-an</i> 'write on X'	<i>s\langle um \rangle\acute{u}lat</i> 'X to write'

roots in Table 2. A root with a canonically monovalent meaning such as $\sqrt{t\acute{a}wa}$ [[laugh]] is able to take not only the actor voice but also the conveyance voice, corresponding to the theme or cause of laughing, and the locative voice, corresponding to a participant who is laughed at.⁵ On the other hand, $\sqrt{b\acute{ig}ay}$ [[give]] has three conceivable participants, Agent, Theme and Receptient, corresponding to the actor voice, conveyance voice and locative voice, respectively.^{6,7} The patient voice is ungrammat-

⁵ The problems inherent in treating these morphemes as applicatives are outlined in Kaufman (to appear). Three features which are expected of nominalizations but odd for applicatives are the following:

- (i) If the conveyance voice *i-* (and, on some accounts the locative voice, *-an*) are applicatives, we expect them to cooccur with transitive *-in* (cf. Foley 2008).
- (ii) We do not necessarily expect applicative objects to be obligatorily promoted to subject. Applicatives are understood to license new objects but the 'objects' of *i-* and *-an* predicates never surface as anything but subjects (cf. Pearson 2005: 408 fn. 20 regarding Malagasy).
- (iii) There is nothing which should prevent multiple applicatives as commonly found in Bantu and more innovative Austronesian languages which show unambiguous applicatives (e.g. Tukang Besi, Donohue 1999).

⁶ Richards (2000) and Rackowski (2002) argue against a direct dependency between voice morphology and theta-roles in favor of treating unexpected voice/theta-role correspon-

ical with \sqrt{bigay} because the Theme of [[give]] undergoes movement away from the Agent, an object interpretation which is only compatible with conveyance voice. The root $\sqrt{lákad}$ [[walk]] can take all four voices. The conveyance voice corresponds to an object being conveyed by walking, the patient voice to the distance walked, the locative voice to the destination and the actor voice to the agent of the walking.

In the following I will attempt to unite elements from three quite separate strands of research into Tagalog morphosyntax, the historical, the typological, and the generative, in the service of furthering our understanding of several key linguistic properties of Philippine languages which have attracted the attention of Austronesianists and non-Austronesianists alike.

New here is the idea that certain nominal properties on the level of the word and clause find their source on the level of the root. A peculiar feature of Tagalog which has only recently come to light is the fact that all roots obtain an essentially nominal interpretation when used independently (Himmelmann 2008). I take this to be indicative of the fact that Tagalog lacks a verbal category altogether, accounting for both the interpretations of bare root and the projection of nominal syntax throughout the clause.

2.0. On lexical categories

One of the often overlooked reasons lexical categories are of such interest is that they are absolutely unnecessary from the perspective of

dences as quirky case. It is certainly true that there are roots whose voice paradigms appear exceptional but this is again precisely what we find in nominalizations. With English agentive *-er* we find exceptional cases such as *roaster* (as in chicken) and *walker* (instrument) and with patient oriented *-ee* we find *retiree*, *attendee*, *refugee*. Barker attempts to unify the semantics of *-ee* formations as all involving sentience with lack of volitional control. Nonetheless, there are always outliers to semantically based generalizations.

⁷ To form a transitive predicate, \sqrt{bigay} requires prior affixation of *pag-*, which is characterized as an inner causative by Travis (2000). The combination of *pag-* and $\langle um \rangle$ is spelled out as *mag-*.

logical form. Indeed, there are no lexical categories, as such, in predicate calculus and the language of logic and yet, it appears that no natural language can make do without them (see Evans & Osada 2005 and Baker 2003 for critical reviews of allegedly a-categorial languages). Lexical categorization can thus only be seen as a unique characteristic of natural language.

On the other hand, it is a widely held although too rarely examined belief that all languages possess at least the basic categories of noun and verb. In practice, these categories tend to be identified by shallow morphological criteria in the tradition of Dionysius Thrax, i.e. a word which is marked for tense versus a word which is marked for plurality. As a result, all languages which instantiate the semantic categories tense and number on the word level are typically described as possessing nouns and verbs. Of course, it is more precise to say that such a language only instantiates tense and number as word-level morphology. The labels “verb” and “noun” only become useful when they indicate correlations across multiple domains, and only then if their categorization is unpredictable on independent grounds. For instance, if a set of roots show identical restrictions on what type of morphology they may take and are further restricted in their syntactic potential, it makes perfect sense to bestow a category label on them. If, however, we simply define verbs as those words which host tense/aspect morphology, and this morphology can occur on any lexical stem, then we are simply giving a superfluous label to words which happen to bear tense/aspect marking. Furthermore, it is not the case that semantic, morphological and syntactic categories must line up in the way we predict on the basis of more familiar languages (Hengeveld 1992, Broschart 1997). Nouns and verbs can be understood to represent alignments of categories on at least three different levels: nouns canonically denote objects, take number, case, gender/class morphology and typically play an argumental role in the clausal syntax. Verbs, on the other hand, canonically denote events, take aspect, tense, mood morphology and typically play a predicational role in the clausal syntax. Himmelmann (2008: 249) avoids the use of “noun” and “verb” in his discussion of Tagalog lexical categories precisely because the alignment of semantic, morphological and syntactic categories appear to be quite different from that of English and more familiar languages (see also Gil 1993, 1995, 2000, Foley 1998).

2.1. *Root and word*

The definition of the term root is somewhat contentious in contemporary morphological theory. In its traditional sense, “root” simply refers to the smallest unanalyzable lexical portion of a word (Bloomfield 1917, Hockett 1958). In theories which endorse semantic decomposition (Marantz 1991 et seq, Borer 2004), the root, as such, is a purely abstract element which contains conceptual meaning but lacks category, event structure and argument structure. For the purposes of this paper, we use “root” in its more traditional sense to refer to simplest, unanalyzable surface forms. Only after we discuss the syntax and interpretation of these simplest surface forms will we be able to speculate as to what types of more abstract elements lie behind these surface forms.

All lexical roots in Tagalog may appear on their own, unadorned by voice or aspect morphology. I make a crass division here between two kinds of roots based on their conceptual structure.⁸ On the one hand there exist roots whose entire denotation can be captured reasonably well in a snapshot. I count such basic roots as *bato* ‘rock’, *púsà* ‘cat’, *táo* ‘person’ as belonging to this class. Among those roots whose denotation cannot be captured so easily there exist several subtypes. There are those which could be elucidated by adding a timeline, among which I count roots like *takbo* run, *patay* kill, *inom* drink, and those whose elucidation would not be aided by the addition of a timeline, among which I count such roots as *ama* father, *íbig* love, *lungkot* sadness.

We will be concerned in this section with the *bato* ‘rock’ type, which we can term “simple entity-denoting” and the *takbo* ‘run’ type, which we term “simple event-denoting”.⁹ A striking generalization about Tagalog

⁸ Because of the focus of this paper, we also restrict our examination to nouns and verbs, leaving out discussion of property denoting words.

⁹ Roots can thus be considered as typed here as I take concepts like $\llbracket \text{jump} \rrbracket$, which contain transitions, to be incoherent without the dimension of time. They are roots which, in Barker’s (1998: 717) terms are, “associated with a set of eventualities that can serve as qualifying events”. This dimension is unnecessary, and indeed infelicitous for simple entity denoting roots like $\llbracket \text{rock} \rrbracket$. Note that the presence of a timeline is completely independent of how this timeline ultimately contributes to the semantics of a surface form. The interface between the timeline in conceptual structure and the ultimate denotation I take to be the proper domain of event decomposition.

Table 3. *Tagalog root meanings*

English	Tagalog	Root	Root Translation
'walk'	<i>l<um>ákad</i>	<i>lákad</i>	'a walk, an errand'
'run'	<i>t<um>akbo</i>	<i>takbo</i>	'a run, pace'
'eat'	<i>k<um>áin</i>	<i>káin</i>	'eating, meal'
'think'	<i>mag-isip</i>	<i>ísip</i>	'thought, thinking'
'kill'	<i>p<um>atay</i>	<i>patay</i>	'corpse'
'see'	<i>ma-kità</i>	<i>kità</i>	'visible thing'
'destroy'	<i>ma-sirà</i>	<i>sirà</i>	'destroyed part'
'break'	<i>ma-básag</i>	<i>básag</i>	'a break'
'teach'	<i>mag-túrò</i>	<i>túrò</i>	'lesson, teaching'
'burn'	<i>ma-súnog</i>	<i>súnog</i>	'fire'
'say'	<i>mag-sábi</i>	<i>sábi</i>	'what is said'
'buy'	<i>b<um>ili</i>	<i>bili</i>	'price bought for'
'fall'	<i>ma-húlog</i>	<i>húlog</i>	'a fall, thing dropped'
'take'	<i>k<um>úha</i>	<i>kúha</i>	'taken object'

is that all roots, on the surface, including those which we think of as prototypically event-denoting, appear to denote entities (Himmelman 2008, Dahl 1973: 120, Cena 1977).¹⁰ This can be seen quite clearly in Table 3. The second column lists the Tagalog translational equivalents of the English verbs in the first column. These forms are all combinations of a root with voice morphology.¹¹ The third column lists the root of these translational equivalents and the fourth column shows the translation of the root when used independently.

Focusing on the third and fourth columns, we see that in each case the denotation of the independent root refers to an argument within the event, e.g. *patay* 'a corpse', *sirà* 'a destroyed part', *básag* 'a break', or to the action itself, *lákad* 'walk, errand' *takbo* 'run, pace'. To briefly exemplify how these roots are typically used in discourse, three of the roots in Table 3 are shown in sentences culled from the internet in (9)–(11).

¹⁰ Himmelman (2008: 275) enumerates the following possible relationships between event denoting roots and their surface interpretation: "(a) the state which ensues from the successful performance of the action (similar to the past participle in English); (b) the result or the typical cognate object of the action (similar to object(ive) nominalisations in English); or (c) the name of the action (similar to an action nominalisation in English)."

¹¹ The particular choice of voice is arbitrary for many of the items but more restricted in the case of others. In particular, predicates which are inherently non-volitional/unaccusative require *ma-* in Tagalog (i.e. *makità*, *masirà*, *mabásag*, *masúnog*, *mahúlog*).

In each case, the predicate refers to a particular instantiation or a generic interpretation of walking, sleeping or cooking.¹² As discussed by Himmelmann (2008: 278), these roots cannot be predicated of the notional subject, as shown in (13).¹³ This indicates that the entity denoting character of Tagalog roots is not derived from their syntactic position but is rather present whether they surface independently in any position.

- (13) a. %lákad=siya b. %túlog=siya
 walk=3s.NOM sleep=3s.NOM
 c. %lútò=siya
 cook=3s.NOM

There exists an interesting pattern in how the denotations of surface roots are derived from the underlying concepts which they presumably signify on a more abstract level. For monovalent roots such as *lákad* ‘walk’, *takbo* ‘run’, *túlog* ‘sleep’, the root denotes an instantiation, final result, or purpose of the action.¹⁴ With semantically bivalent roots such as *patay* [[kill]] ‘corpse’, *lútò* [[cook]] ‘cooked dish’, or *básag* [[break]] ‘a break’, the root denotes the patient or theme of the action.¹⁵ What could be responsible for this? If, as claimed by Marantz (1991), Kratzer (1996), Chomsky (1995) and others, the so-called external argument (Williams 1994) is introduced not by lexical heads themselves but rather by a higher functional projection corresponding roughly to Dowty’s (1979) ‘DO’ operator, then it makes sense that a root could not access an external argument within its smallest domain. It follows naturally then that if a bare root is entity denoting, it must denote one of its internal arguments, e.g. Patient, Theme, etc. (cf. Marantz 1997).¹⁶

¹² Note that the same reading is found with English *sleep*. In English, however, there is very little regularity in this pattern (cf. Clark & Clark 1979).

¹³ Although *túlog siya* is infelicitous, *tulog siya* with vowel length removed from the root *túlog* is completely natural. A prosodic morpheme which removes vowel length from roots creates resultative states.

¹⁴ Note that bare roots do *not* denote event nominalizations, which are expressed with additional morphology (*pag-* with reduplication if required by the stem), e.g. *pag-lákad* ‘walking’, *pag-takbo* ‘running’, *pag-túlog* ‘sleeping’.

¹⁵ That is, roots whose denotations are incoherent without the presence of a second argument.

¹⁶ *Prima facie* counter evidence to this claim exists in Tagalog with roots such as *gúro* ‘teacher’, *tánod* ‘guard’. These roots, however, are demonstrably not derived from

We can at this point also ask why all lexical roots which ultimately form event denoting words must denote entities in their bare form. In the theory of Distributed Morphology, roots, as such, are strictly abstract entities which never surface without the addition of further functional morphology. This theory is very apt for capturing the nature of Semitic-type consonantal roots, which form a wide range of semantically related words of various lexical categories but which never surface independently. For instance, the root $\sqrt{\text{LMD}}$ in Hebrew is involved in all the words in (14), but $\sqrt{\text{LMD}}$ cannot surface on its own.

- (14) $\sqrt{\text{LMD}}$ [[learn]]
- | | | | |
|----------------|---------------|---------------|----------------------|
| lemed | limud | lamdan | melumad |
| ‘learning’ | ‘study’ | ‘learner’ | ‘learned’ (A) |
| melamed | talmid | lamad | talmud |
| ‘teacher’ | ‘student’ | ‘learned’ (V) | ‘s.t. to be studied’ |

If Semitic is a transparent representation of a more general morphological reality, then we can imagine all roots to be like Semitic consonantal roots, unable to appear without the addition of categorizing morphology. On such a theory, what I have been referring to here as bare roots in Tagalog would not be bare at all but would carry a null affix which adds the category *n* and derives the nominal-like interpretation. This affix must

event-denoting roots to begin with. This can be seen from two independent diagnostics. First, many bivalent event denoting roots can take a prosodic morpheme which removes vowel length from a root and produces a resultative, e.g. *tápos* ‘finish, end’ *tapos* ‘finished’. Agent denoting roots like *tánod* can never take this morpheme, i.e. there is no **tanod* for ‘guarded’ or **gurò* for ‘taught, tutored’. The second piece of evidence is that these roots are often not the same roots used for forming corresponding event denoting predicates. For instance, the corresponding event denoting predicate which best describes the action characteristic of a teacher is not formed with the root *gúrò* but rather with the unrelated (although coincidentally similar) *túrò* ‘teach, point’. As expected, when used independently, this event denoting root refers to an internal argument, not the external argument, as seen in (ii)b.

- | | |
|---|--|
| <p>(i) a. mag-gúrò
AV-teacher
‘to be a teacher/study to be a teacher’</p> | <p>b. gúrò
teacher
‘teacher’</p> |
| <p>(ii) a. mag-túrò
AV-teach/point
‘to teach’</p> | <p>b. túrò
teach/point
‘a teaching, lesson taught, something pointed to’</p> |

attach to roots when they surface independently, otherwise, we would expect to find examples such as those in (13) to be grammatical. When these same abstract roots are employed to form event-denoting predicates, they presumably attach to a null v , which both projects the verbal category and adds an eventive interpretation. On this theory, event denoting and bare entity denoting derivations of a root could be represented as in (15).

- | | | | |
|---------|--|----|--|
| (15) a. | nag-lútò- \emptyset_V =siya
AV.BEG-cook=3S.NOM
'S/he cooked' | b. | lútò- \emptyset_N
cook
'cooked dish' |
|---------|--|----|--|

This begs the question of why the categorial head responsible for the verbal category and eventive semantics, v , is unavailable for independent roots. It is difficult to imagine a principled way of allowing its presence in forms such as (15)a but disallowing it in (15)b. Following Fabb (1984) and Pesetsky's (1995) analysis of certain combinatorial facts of English morphology, we could say that the null v simply selects for affixed stems and not roots, or alternatively, that affixing of v demands the further addition of aspect and voice morphology.¹⁷ But as will be shown in the

¹⁷ Among these two tacks, the latter approach would be far more promising than the former. It could be argued that v adds morphosemantic features which cannot be saturated without projecting further morphosyntactic structure, namely, Voice and Aspect phrases. This would correlate nicely with the agentive and eventive functions of v , respectively. Because there is no overt word-level case, number or gender morphology in Tagalog, it could then be argued that the addition of n simply has no visible consequences. It could further be predicted that a simplex v word could also exist were it to satisfy its morphosemantic requirements higher in the derivation. This could potentially account for the iterative construction, one of the only contexts where an unaffixed form is able to obtain an eventive interpretation and take an external argument subject (*siya* 'S/he').

- (i) Súlat=siya nang=súlat
write=3S.NOM GEN=write
'S/he's writing and writing'

Nonetheless, I reject this approach as it is unable to capture the overwhelming evidence that even apparent verbs affixed with Voice and Aspect morphology still have nominal properties, as will be discussed extensively below.

following sections, even voice and aspect inflected forms show nominal behavior, a fact which I argue is best captured by treating even these apparent verbs as nominals.¹⁸ There also exists interesting evidence from code mixing phenomena which argues that all Tagalog roots are nominal whether or not they are further embedded under voice/aspect morphology. As shown in (16), voice and aspect morphology regularly create event denoting predicates from unambiguous nominal English stems. The same phenomenon can be seen in the numerous Spanish borrowings into Tagalog. In cases where we might expect an infinitive or inflected verb to be borrowed for an event denoting predicate we instead find that it is the nominal form which is consistently borrowed, as seen in (17)a form *trabaja* ‘work’ and (17)b from *parada* ‘stopping’.

- (16) a. *mag-ice-cream* b. *mag-basketbol*
 AV-ice-cream AV-basketball
 ‘eat ice cream’ ‘play basketball’
- (17) a. *mag-trabaho* b. *p<um>arada*
 AV-work <AV>stop
 ‘to work’ ‘to park’

¹⁸ It is even more problematic for an approach like Baker’s (2003) which views roots as inherently verbal, nominal or adjectival, generally corresponding to root meaning. Baker (2003: 53) makes the claim that operations such as incorporation and causativization disambiguate zero derivations due to the “Proper Head Movement Generalization” (Li 1990): “A lexical head A cannot move to a functional head B and then to a lexical head C”. This prevents a verbal root from nominalizing (movement to a functional head) and then incorporating (movement to a lexical head). Likewise, it prevents a nominal root from verbalizing and then causativizing. Tagalog appears to confound this claim with the two productive formations in (i) and (ii). In (i) the voice/aspect form *magka-* ‘to have’ is able to take notionally ‘verbal’ roots while retaining their entity denoting meanings. In (ii), the causative *pa-* attaches to notionally ‘verbal’ roots but maintains the entity denoting meaning found with the bare root, e.g. *gawà* ‘thing made’.

- (i) *mag-ka-lákad* (ii) *pa-gawà*
 AV-HAVE-walk CAUS-make
 ‘to have a walk/errand’ ‘thing caused to be made’

A complementary piece of evidence comes from a variety of Filipino English called *konyo* or *kolehiyala* English.¹⁹ Here we find precisely the converse situation: Tagalog roots being borrowed into English syntax. As we may expect if Tagalog independent roots are nominal, these formations require overt verbalization when used as verbs in English syntax. This verbalization is carried out with the verb *make*, a prototypical light verb used for similar purposes across languages. The phenomenon is exemplified in (18).²⁰

- (18) a. Let's make *pasok* ('enter') *na* to our class!
 b. Wait *lang!* i'm making *kain* ('eat') *pa!*
 c. Come on *na*, we can't make *hintay* ('wait') anymore!

Thus we see that while the voice/aspect morphology has the power to create event denoting predicates from unambiguous English nouns in a Tagalog context, Tagalog roots require English light verbs for the same function when used in an English context. This makes sense if Tagalog surface roots are always nominal and voice morphology naturally takes nominal complements. An alternative analysis of these facts on which the lexical stems of Tagalog voice/aspect marked words contain a null verbalizer must bear the burden of explaining why this verbalizer can attach to English stems in a Tagalog syntactic context (e.g. *mag-ice-cream*) but why neither the English nor the Tagalog verbalizer can attach to Tagalog stems in an English syntactic context (e.g. *We can't *hintay* anymore).

In the following sections we examine several properties of Tagalog phrase structure showing that the nominal properties of surface roots projects far beyond the word level.

¹⁹ It is characteristic of these varieties simply because they are the only forms of Filipino code mixing which regularly create English verbal predicates from Tagalog roots. See Bautista (1996) for an overview of varieties of Philippine English and code mixing.

²⁰ This is an excerpt from a humorous piece entitled "the 10 *konyo* commandments" which pokes fun at the speech of wealthy Anglophone Filipinos and was widely distributed over the internet. More examples can be found in Bautista (1996). The clitics *na* and *pa* are aspectual ('already', 'still', respectively) and the clitic *lang* is delimitive 'only/just'.

The analysis in (20)a is by far the most common one in the generative literature but is rarely argued for explicitly against (20)b or c.²¹ Here, I will argue that (20)b or c are, on the whole, correct and that (20)a is unfounded. For some Austronesian languages, there appears to be good evidence that (20)a is correct, or at least that it represents one option. For instance, in Malagasy, there exist functional elements which appear specific to introducing relativizations and the cleft portion of wh-/focus constructions (*izay* and *no*) and which thus break the symmetry generally found in Tagalog and Philippine languages.²² Himmelmann (2008), however, argues against a headless relative analysis for Tagalog as there is no independent evidence for complicating the syntax by positing an otherwise undetectable asymmetry between “verbally” headed and “nominally” headed phrases. To take a simple example, the voice/aspect inflected words can, and very commonly do, appear in the position of unambiguous nominals and allow modification by canonical DP internal material, as seen in (21)–(24).

- (21) Iyong dalawa=ng ma-ganda=ng pinsan=mo
 that:LNK two:LNK STA-beauty=LNK cousin=2S.GEN
 ‘Those two beautiful cousins of yours.’ (unambiguous nominal)

²¹ There is a relatively large literature on the analysis of wh- questions and focus constructions in Austronesian languages as involving pseudo-clefts in subject position (Richards 1998, Aldridge 2002, Georgopoulos 1991, Chang 2000). Nonetheless, these works do not consider the possibility that the constituent in subject position is smaller than a relative clause, e.g. on par with a regular DP. As discussed above, there exists direct evidence against a simple DP analysis in Malagasy (cf. Potsdam 2006: 2176), but not in Philippine languages. Aldridge (2004: 318–321) presents clitic placement facts to argue for the biclausality of sentences such as (19)b. However, the generalization that clitics can never escape a case phrase is able to account for all the same data without positing an asymmetry.

- (i) Kotse=mo(*=ko) ang=t<in>ign-an(=ko)
 car=2S.GEN=1S.GEN NOM=<BEG>look-LV=1S.GEN
 ‘Your car is what I looked at.’
- (ii) T<in>ign-an=ko(*=mo) ang=kotse(=mo)
 <BEG>look-LV=1S.GEN=2S.GEN NOM=car=2S.GEN
 ‘I looked at your car.’

²² See Ntelitheos (2006) for arguments against (20)c in Malagasy. He brings certain binding facts to bear on the issue which cannot be fully addressed here.

- (22) Iyong dalawa=ng ma-ganda=ng s<um>ayaw
 that:LNK two=LNK STA-beauty=LNK <AV:BEG>dance
 ‘Those two beautiful (ones who) danced.’ (actor voice)
- (23) Iyong dalawa=ng ma-ganda=ng <in>áwit-Ø=mo
 that:LNK two=LNK STA-beauty=LNK <BEG>sing-PV=2S.GEN
 ‘Those two beautiful (ones which) you sang.’ (patient voice)
- (24) Iyong dalawa=ng ma-ganda=ng t<in>ign-an=mo
 that:LNK two=LNK STA-beauty=LNK <BEG>look-LV=2S.GEN
 ‘Those two beautiful (ones who) you looked at.’ (locative voice)

Note the presence of the nasal “linker” or “ligature” between each word level member of the phrase headed by the demonstrative. This linker is a functional element common to almost all Philippine languages and is found between all elements in a modificational relationship including relative constructions. The view taken here is that the linker is best viewed as a functional element signaling Predicate Modification (Heim & Kratzer 1998, see also Chierchia & Turner’s 1988 JOIN operator). As seen in (25), the linker is also found in a different function as a complementizer. It could then be the case that while the linker marks a lower modificational relationship between the demonstrative, numeral and adjective, it indicates a higher, clausal relationship between the adjective and a relative containing the following voice marked words in (22)–(24). There are two pieces of evidence, however, that there is no such asymmetry between (21), on the one hand, and (22)–(24) on the other.

The nasal linker has two allomorphs conditioned by the preceding segment: /ŋ/ post-vocally and /na/ post-consonantly. The post-consonantal linker /na/ can also appear before (non-reduced) complement clauses and restricted relatives even when the preceding segment is vocalic, as shown in (25), and is indeed slightly preferred.

- (25) S<in>ábi-Ø=nila (%ŋ/na) hindi=sila mag-bá~basketbol
 <BEG>say-PV=3P.GEN LNK NEG=3P.NOM AV-IMPRF~basketball
 ‘They said that they wouldn’t play basketball.’

In smaller modification contexts such as that between a numeral and an unambiguous noun, we find that the post-consonantal /na/ allomorph is only possible in postvocalic contexts when there is contrastive focus on what follows. Thus, (26) is only possible with /na/ when ‘the two

TEACHERS' are being contrasted with two of something else (similar to the interpretation of 'the two who are teachers' in English, suggesting that a relative clause may underlie this structure, i.e. 'these two who are teachers').

- (26) Ito ang=dalawa (ŋ/#na) gúrò
 this NOM=two LNK teacher
 'these are the two teachers'

We can now compare the choice of post-vocalic linker before unambiguous nouns as in (26), and voice/aspect marked words, as seen in (27). As indicated below, the same focus reading is required for /na/ to appear before a voice/aspect inflected word as required before a simple noun.

- (27) Ito ang=dalawa (ŋ/#na) nag-tú~túrò
 this NOM=two LNK AV.BEG-IMPRF~teach
 'these are the two (who are) teaching'

Although a full account of linker allomorphy remains elusive (see Richards 1999 for many interesting observations), inasmuch as it is linked to the syntax (directly or indirectly via prosodic structure) the symmetry between (26) and (27) argues against any major structural difference between voice/aspect inflected words and unambiguous nouns in a DP context along the lines of (20)a.²³

The second piece of evidence mentioned above relates to possessors. If all complements of case marking determiners contain either an overt or null nominal head then we would not expect to see syntactic variation between overtly headed NPs and null headed ones. Interestingly, such variation exists. Although there is a good deal of syntactic symmetry across word classes in Tagalog, not all words classes are able to license possessors phrases without contextual coercion.²⁴ Property denoting words are

²³ Note that the judgments reported here are not exactly those reported in Richards (1999). Richards claims that both allomorphs are possible before what he analyzes as reduced relatives, only /na/ is possible before (non-control) complement clauses, and only /ŋ/ is felicitous NP internally.

²⁴ See van Eijk & Hess (1986), Davis & Matthewson (1999), Demirdache & Matthewson (1996) for the same diagnostic applied to Salish languages, which appear to share many typological properties with Austronesian. The ability to license possessors in Tagalog is linked directly to the aspectual stability of the base in addition to contextual factors. The

one of the types of elements which do not license possessors out of the blue, as is clear from the two examples in (28). If there existed a null nominal head in (22)–(24) which was responsible for licensing the DP internal material, then it should be able to license the genitive in (28)a, just as the overt head licenses the genitive in (28)b.

- (28) a. ang=basag (%nang=babáe)
 NOM=break/RSLT GEN=woman
 ‘the (woman’s) broken one’
- b. ang=basag na bentána nang=babáe
 NOM=break/RSLT LNK window GEN=woman
 ‘the woman’s broken window’

We can conclude that whatever is responsible for the flexibility, it is probably not a null nominal head. This leaves us with the two other hypotheses in (20): either the lexical heads belong to a single macro-category, (20)b, or functional heads are simply less selective in Tagalog than in more familiar languages, (20)c.

The claim in (20)b, that all words are of the same macro-category has been made most recently by Himmelmann (2008), who posits the “syntactic uniformity hypothesis for content words” stating that “content words are categorially indistinct with regard to syntactic category” (Himmelmann 2008: 264).²⁵ The evidence reviewed below suggests that although bare roots and voice/aspect inflected forms appear to share much in common, it is not the case that all content words are categorially indistinct.

more stable and given the base, the more felicitous a possessor. Although this deserves far more attention, we can exemplify with the contextual coercion of a possessor on an actor voice verb. The actor voice form in (i) can be possessed, as in (ii), only if it is known that the speaker has a child studying in the University of the Philippines.

- (i) nag-yu~yúpi
 AV.BEG-IMPRF~U.P.
 ‘studying at the University of the Philippines.’
- (ii) Ito ang=nag-yú~yúpi=ko
 this NOM=AV.BEG-IMPRF~U.P.=1s.GEN
 ‘This is my University of the Philippines studying one’

²⁵ He does show however that property denoting roots must be classed into at least two categories for the purposes of morphology (see also Wolff 1993).

Table 4. *Proto-Austronesian Voice/Aspect Paradigm* (Ross 2002: 33)

	Actor	Patient	Location	Circumstantial
Indic. Neutral	$\langle um \rangle \checkmark$	$\checkmark -\partial n$	$\checkmark -an$	<i>Si-</i> \checkmark
Indic. Perfective	$\langle umin \rangle \checkmark$	$\langle in \rangle \checkmark$	$\langle in \rangle \checkmark -an$	<i>Si-</i> $\langle in \rangle \checkmark$
Indic. Durative	$\langle um \rangle R-\checkmark$	<i>R-</i> $\checkmark -\partial n$	<i>R-</i> $\checkmark -an$	<i>Si-R-</i> \checkmark
Non-indic. Atemporal	\checkmark	$\checkmark -u, \checkmark -a$	$\checkmark -i$	$\acute{a}n-i + \checkmark, \checkmark -\acute{á}ni$

Recall that the nominal view of the voice system appears corroborated by historical reconstruction. Comparative evidence suggests that while the progenitors of the voice forms discussed throughout here were nominalizations, there also existed another paradigm which probably constituted the original verbs of Proto-Austronesian. These forms are what Wolff (1973) first reconstructed as the “dependent” paradigm of the Proto-Austronesian voice system. The change which has come to typify modern Philippine languages was the reinterpretation of the nominalizations as canonical event denoting predicates (Starosta, Pawley & Reid 1982).²⁶ The relevant portion of the reconstructed Proto-Austronesian voice/aspect paradigm is shown in Table 4, an expanded version of (1). (Ross’s “circumstantial” corresponds to what we have been referring to as the “conveyance” voice.) While the top three rows, labeled “indicative” following Ross (2002), are taken here to represent aspectually inflected thematic nominalizations, the bottom row appears to have a closer connections to the proto-verbs.

What is critical for us here is that indicative forms appear throughout Philippine languages as both canonical nominalizations in argument position as well as aspect inflected predicates but non-indicative forms do not have such a free distribution. Ross (2002: 37) states, “. . . in Philippine-type languages which retain non-indicative verb forms (and Tagalog doesn’t), a non-indicative form derived from a root may only occur in the predicate slot, and the same must have been true of PAn.” Interest-

²⁶ Evidence that this reinterpretation took place after the break up of Austronesian is found in three Formosan languages, Tsou, Rukai and Puyuma, which appear to have not made the nominalization to event predicate reanalysis (see Ross to appear).

ingly, there are dialects of Tagalog, notably that of Batangas, which do preserve the original verbal forms. These forms have also been historically marginalized in Batangas Tagalog but unlike in Standard Tagalog, they are still regularly found in imperatives where they alternate with the nominal forms found in Standard Tagalog. Two morphosyntactic distinctions between the indicative (nominal) and non-indicative (verbal) forms are telling. First, the non-indicative forms behave like imperative verbs cross-linguistically in requiring deletion of the imperative addressee (in the case of a singular addressee). In contrast, the addressee is nearly obligatorily present with nominal imperatives, as shown in (29).

- (29) a. Buks-an*(=mo) ang=pintuan!
 open=LV=2S.GEN NOM=door
 ‘Open the door!’ (“indicative” – nominal)
- b. Buks-i*(=mo) ang=pintuan!
 open=LV.DEP=2S.GEN NOM=door
 ‘Open the door!’ (“non-indicative” – verbal)

This is an expected difference between the two forms as verbs are the more natural vehicles of illocutionary force and thus expected to be able to license addressee obviation. More important here however is the different behavior of the two forms when embedded within a DP. It is an oft-repeated fact of Tagalog that seemingly any type of word can be made into a DP by the addition of one of the case markers. Even an indicative imperative can find itself embedded under an *ang* phrase in all dialects of Tagalog, as shown in (30)a. What has not been noticed is the inability of the non-indicative forms to do the same, as seen in (30)b. These forms are restricted to appear in predicate position in line with Ross’s (2002) generalization above.

- (30) a. Pintuan ang=buks-an=mo!
 door NOM=open-LV=2S.GEN
 ‘Open the DOOR!’
- b. *Pintuan ang=buks-i!
 door NOM=open-DEP.LV

The ungrammaticality of (30)b is the first gap mentioned at the outset of this section and it shows that the ability for a voice/aspect forms to appear in a canonically nominal domain in Tagalog is not due to a

complete blindness of the functional material in DP towards its lexical complements. Rather, it is quite possible that the freedom we find above is due in part to the lexical heads within DP sharing a (nominal) macro-category.

The second piece of evidence that unambiguous nouns and voice/aspect marked words share a category comes from the behavior of deictics, and oblique phrases more generally. As seen in (31), “bare” deictics are found either in post-predicate position or left-peripheral focus position when functioning as adjuncts. However, they are not entirely felicitous as the lexical heads of DP, as shown in (32)a. To appear in this position they take what appears to be voice/aspect morphology, as shown in (32)b (Schachter & Otnes 1972, Aldridge 2004: 324–325, Mercado 2005, among others).²⁷

- (31) a. Nag-áral=sila díto b. Díto=sila nag-áral
 AV:BEG-study=3P.NOM here here=3P.NOM AV:BEG-study
 ‘They studied here.’ ‘They studied HERE.’

- (32) a. ??ang=díto b. ang=nan-díto
 NOM=here NOM=AV.DST.BEG-here
 ‘the one here’

This is further evidence that “the syntactic uniformity hypothesis for content words” cannot be an across the board phenomenon but is rather focused on two word classes, bare roots and voice/aspect forms (descendent from the Proto-Austronesian indicative (nominal) paradigm).

²⁷ Specifically, deictics appear to take the actor voice distributive in the perfective (*nan-*) or progressive (*nan-REDP*), or the patient voice non-volitional in either of the same two aspects (*na-* and *na-REDP*, respectively). This distinction also appears in oblique phrases which take *sa* as arguments but prefer *nása* as predicates. The *na-* formant is interpreted by Aldridge (2004: 324) as instantiating a verbal category which is required for prepositional phrases to function as predicates. It must be noted however that when the subject is headed by a motion denoting word, a *sa* marked oblique can appear in the predicate position, as shown in (i).

(i) Sa=Manilà ang=punta=nila
 OBL=Manila NOM=go=3P.GEN
 ‘Their going was to Manila.’

It is an interesting and very general fact of Philippine languages that predicates such as those in (36)b are impossible to express in the genitive case. This is exemplified below where the genitive in post-predicate position in (38)a is contrasted with the ungrammatical genitive predicate in (38)b. To express a possessor predicate, the oblique case must instead be employed, as shown in (39), which, in addition to possession, allows for a very general locative reading, e.g. with, for, etc.

- (38) a. Ang=koponan ni=Juan b. *Ni=Juan ang=koponan
 NOM=team GEN=Juan GEN=Juan NOM=team
 ‘Juan’s team’ (For, ‘The team is Juan’s’)

- (39) Kay=Juan ang=koponan
 OBL=Juan NOM=team
 ‘The team is Juan’s’

The second type of predicational genitive, shown in (37) above, has no genitive equivalent in Tagalog at all, either as modifier or predicate, as seen in (40).²⁸ Rather, relations of this type are expressed with the existential, as in (41).

- (40) *manga=libro nang=ma-laki=ng halaga
 PL=book GEN=STA-size=LNK worth
 (For, ‘books of great worth’)
- (41) manga=libro=ng may=ma-laki=ng halaga
 PL=book=LNK EXT=STA-size=LNK worth
 ‘books of great worth’

Restrictions on genitive predicates are quite common cross-linguistically. DPs which can take genitive case when serving as modifiers or arguments must often take a dative or oblique case when serving as

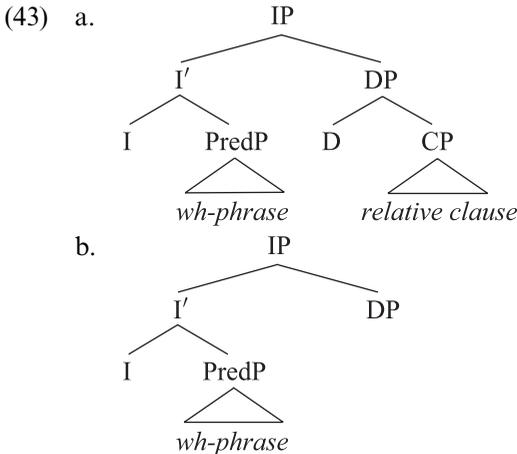
²⁸ Note also that while genitive *nang* may denote several relations in Tagalog, causation is not one of them. Compare the difference between *nang* and English *of* in (i), where an oblique is required instead of a genitive to indicate cause.

(i) Na-matay=sila (sa=/*nang=)gútom
 PV.NVOL-die=3P.NOM OBL/GEN=hunger
 ‘They died of hunger.’

predicates, as in Tagalog. In other languages, genitive predicates can only appear as part of larger DPs which include the possessum. Such is the case described for Una, a Papuan language spoken in Irian Jaya (Louwerse 1988 via Dryer 2007), as shown in (42), where a possessor without a possessum is ungrammatical.

- (42) A yina Karba *(yina)
 that food Karba food
 ‘That food is Karba’s (food)’

I claim here that this general ban on genitive predicates is the basis for the Austronesian extraction restrictions discussed above. This follows from two simple points. First, it has been argued convincingly that focused and interrogative DPs (‘what’ and ‘who’) surface in the predicate position of a pseudo-cleft structure in Philippine type languages (see Paul 2000, 2001, Georgopolous 1991, Aldridge 2002, Potsdam 2006, Oda 2002, Gerassimova & Sells 2008, among others). The basic structure convincingly argued for by Potsdam (2006) for Malagasy is shown in (43)a, although he notes that (43)b – more in line with what has been argued for here regarding Tagalog – is also a possibility for Malagasy in many cases (see also Paul 2000, 2001).



The second, oft-ignored point is that Tagalog argument interrogatives are case marked, as shown clearly by the forms in Table 5.

Table 5. *Case marked interrogatives in Tagalog*

	NOM	GEN	OBL
'What'	ano	nang=ano	saan
'Who'	sino	nino	kanino

On the basis of case preservation alone, we can rule out many of the examples in the literature which are claimed to exemplify a peculiarly Austronesian restriction on extraction. Taking a sentence such as that in (44) as a basis of comparison, the extraction restriction is generally exemplified with ungrammatical sentences such as that in (45). The immediate problem with (45), however, is orthogonal to extraction restrictions; a nominative interrogative corresponds to an argument which should receive genitive case. The result is independently ungrammatical because the clause has three nominative marked constituents and is, in a sense, an incoherent concatenation of the two grammatical copular sentences shown in (46) and (47). Note that the interrogative in (46) must be equated with the patient of buying and not the agent because, as is argued here, this is the denotation of the voice/aspect marked form itself.

- (44) B<in>ili-Ø **ni**=Boboy ang=libro
 <BEG>buy-PV GEN=Boboy NOM=book
 'Boboy bought the book.'
- (45) ***Sino** ang=b<in>ili-Ø ang=libro?
 NOM:who NOM=<BEG>buy-PV NOM=book
 (For, 'Who bought the book?')
- (46) **Sino** ang=b<in>ili-Ø?
 NOM:who NOM=<BEG>buy-PV
 'Who was bought?'
- (47) Ang=b<in>ili-Ø ang=libro
 NOM=<BEG>buy-PV NOM=book
 'What was bought was the book'

More to the point of the extraction restriction than is the example shown in (48), with a genitive case interrogative in predicate position correctly corresponding to what would be a genitive marked agent of the patient voice form *binili*. But this, too, is ungrammatical, because Taga-

log strictly bans genitive predicates, as we have already seen above.²⁹ Thus, in accounting for the ungrammaticality of (49) – a very basic fact of Tagalog and Philippine languages more generally *which is completely orthogonal to the voice system* – we have already accounted for (48).

(48) ***Nino** ang=b<in>ili-Ø?
 GEN:who NOM=<BEG>buy-PV
 (For, ‘Who bought (it)?’ / ‘Whose was the bought thing’)

(49) ***Ni**=Juan ang=koponan
 GEN=Juan NOM=team
 (For, ‘The team is Juan’s’)

In the next section, I flesh out what I take to be the minimum requirements for an explanatorily adequate derivation of the Tagalog clause which unites the root-level, word-level and clause-level facts examined above.

4.0. Towards a nominal analysis

The main contribution of the present work is to bring to light several non-trivial connections between Tagalog morphosyntax and ostensibly universal features of nominal morphosyntax and thereby open a new vista for explaining extraction restrictions. The formalization of this may proceed in any number of ways and I am committed more to the structural analogy itself rather than any particular formalization. Nonetheless, in this section I sketch out the bare essentials of one proposal which can begin to account for both the root interpretation facts as well as several facts of clausal syntax, to be reviewed below.

At the heart of the current proposal is a lack of *v*, the categorizing head which creates verbs.³⁰ This forces lexical roots to merge with *n*

²⁹ The genitive interrogatives *nino* ‘who’ and *nang=ano* ‘what’ are most commonly found in echo questions and may only be used in-situ.

³⁰ Recall that Batangas Tagalog still possesses robustly verbal forms in the imperative. I take these forms to be instantiations of *v* which, for whatever reason, are incompatible with declaratives.

(or a^{31}) and hence adopt entity denoting meanings when appearing independently. It also accounts for the projection of a predicate phrase with several nominal characteristics. I take the key differences between n and v to be the following:

(50) n properties:

- (i) A Possessor is projected [Spec, n]
- (ii) Association with genitive case
- (iii) No inherent capacity for an event variable

v properties:

- (i) An Agent is projected in [Spec, v]
- (ii) Assignment of accusative case to object
- (iii) Inherent capacity for an event variable

The idea that a language can possess a defective v category or lack it altogether is not new. In this respect, the current proposal most resembles that of Johns' (1992) for Eskimo with similar ideas having been proposed by Bok-Bennema (1991), Alexiadou (2001), Nash (1996) to account for ergativity more generally.

As a result of lacking v , the L(exical)-syntax of a canonical eventive predication will look like (51). A root merges with its complement and then with a category determining head n . The categorizing nP phrase projects a possessor in its specifier. The complement is assigned genitive case by being within the domain of nP . For convenience, it is indicated as theme but post-predicate genitive phrases may obtain a wide range of interpretations (theme, instrumental, cause, temporal adjunct).

(51) nP [POSS _{n} '[n \sqrt{P} [$\sqrt{[THEME]]}$]]

Because n has no capacity for an event variable, the structure in (51) will only be able to express an *internal* "slice" of the event if spelled out without the addition of further operators as the external argument of the

³¹ For reasons of space, I do not discuss adjectival categories in Tagalog. I take there to be only one true surface adjective in Tagalog, corresponding to Himmelmann's type B and Sabbagh's (2005) unaccusative adjectives, which are formed by deletion of vowel length in the root. The other type of property denoting word, corresponding to Himmelmann's type A and Sabbagh's unergative adjective, I take to be a stative formed by the combination of *ka-* HAVE with actor voice $\langle um \rangle$, and thus derived similarly to the forms discussed above.

event has not been merged. If an element is merged in [Spec,*n*P] it is assigned genitive case and construed as a possessor without any implications of agency, as shown by the interpretation of the potentially independent phrases in (52).³²

- (52) a. súlat ni=Juan b. patay ni=Juan
 write GEN=Juan kill GEN=Juan
 ‘Juan’s letter’ ‘Juan’s killed person’

At this point, Voice, the locus of the voice morphology, is merged, taking *n*P as its complement as shown in (53).

- (53) VoiceP[(AGENT-θ) VoiceP’[Voice⁰ [_{nP}[POSS_n’[_n √P [√ [THEME]]]]]]

The root raises to Voice to pick up one of the four voices and thereby fixes its reference to one of the participants in its denotation. Recall that I am treating Voice here not on par with English active/passive alternations but rather more like *-er/-ee* nominalizations. This can be either due to the different nature of Philippine Voice or a natural outcome of the combination of Voice+*n*. Following Kratzer (1994), I take the specifier of VoiceP to be associated with agentivity. The [Spec,VoiceP] position represents a second chance for an external argument to pick up an agentive theta-role, given the lack of *v*P in Tagalog. By raising from *n*P to VoiceP, the possessor is interpreted as an agent.

Although not directly relevant here, I posit AspectP and NumberP above VoiceP to which the predicate head raises and picks up aspect morphology and number agreement, respectively.³³ Although Tagalog may

³² Note that agency is difficult to interpret even for roots which whose agents can be “easily reconstructed” (√kill, √destroy) in the terms of Marantz (1997).

³³ I abstract away from the possibility of multiple aspect phrases, as in Travis (2000), who argues for a lower and higher aspect heads in Tagalog based on the different positions of the morphs indicating begun and imperfective aspect. Regarding NumP, there are two number agreement morphemes in Tagalog spelled out *si-* and ⟨*ang*⟩, the latter of which is obsolete in spoken Tagalog. I assume that both markers indicate plurality of the predicate head itself, which is then associated with the subject via predication at TP. While ⟨*ang*⟩ is a general number marker appearing with all voices and indicating plurality of the predicate head, *si-* is restricted to appearing with the actor voice (specifically with *mag-*). Taking these morphemes to actually be plural markers rather than number agreement, we predict that a morpheme restricted to indicating agent plurality will also be restricted to appearing on actor voice predicates. For a different account of number marking with *si-*, see Aldridge (2004).

be peculiar in allowing nominal aspect (but see Malchukov 2004 for other non-Austronesian examples), the canonical nominal projections are present as well. Above NumberP we find DemP, which hosts demonstratives, and CaseP, which can host the nominative case marker, in the case of equational sentences (with two *ang* phrases). This structure predicts that nominal elements such as numerals, demonstratives, etc. can appear in the upper range of the predicate phrase and this is indeed what we find. Note the similarity between the sentences in (54) and (55).³⁴

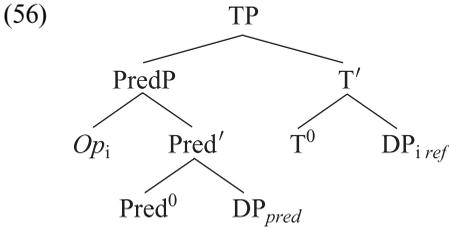
(54) S<um>ayaw diyan ang=manga=pinsan=ko
 <AV:BEG>dance there NOM=PL=cousin=1S.GEN
 ‘My cousins danced there.’

(55) Iyong dalawa=ng s<um>ayaw diyan
 that=LNK two=LNK <AV:BEG>dance there
 ang=manga=pinsan=ko
 NOM=PL=cousin=1S.GEN
 ‘Those two who danced over there are my cousins.’

These functional elements, culminating in the DP, are understood here to all be included within the extended projection of *n*. The DP is daughter to the Predicate Phrase (PredP), which I take here (unorthodoxly) to be base generated in [Spec,TP]. The predicate proper is thus a DP complement of Pred⁰. The specifier of PredP is a null operator which is coindexed with the *ang* phrase. The *ang* phrase subject I take to be base generated as the complement of a null copular T and coindexed with a null operator in [Spec,PredP], as illustrated in (56) in its most minimal form.³⁵ Above TP is a CP layer which is involved in adjunct questions (*where, when, why, how*) and topicalization.

³⁴ There is, as can be expected, a difference in the predicational vs. specificational reading between these two sentences but I am not aware of any evidence for substantial differences in their underlying structure.

³⁵ Pearson (2005) also makes use of a null operator which is coindexed with the Malagasy equivalent of the Tagalog *ang* phrase. Otherwise, however, the two proposals differ considerably in that Pearson argues for a *wh*-agreement type analysis (Chung 1998) while I propose a copular structure as the basis for all primary predication with a very different basis for ruling out the classic cases of bad extraction in Austronesian.



This structure is most saliently divergent from other proposals in the following ways: (i) the basic order is base generated and not derived via predicate fronting, (ii) the subject, rather than the predicate, is the complement of T^0 , (iii) there is no c-command relation between DP_{ref} (the subject) and the genitive phrases contained within DP_{pred} . I take (i) to be supported by the complete lack of linear evidence for predicate fronting in Tagalog, or any other Philippine languages I am aware of. Strikingly unlike the case of verb movement in Romance languages, the predicate always precedes the subject even in contexts where we might expect raising to be impossible (i.e. all forms of subordinate clauses).³⁶ Regarding (ii), it has already been argued by Massam (2000) and others for certain Austronesian languages that the predicate ultimately resides in [Spec,TP]. If we furthermore take predication to be an inherently reversible relation akin to set intersecton (see Den Dikken 2008 and references therein), then the structural relationship between predicate and subject in (56) may not be a concern in and of itself. Regarding (iii), the command relationship between the predicate internal arguments and the subject are not clear from (56) and I take this to in fact be an advantage of the structure. All previous proposals for the phrase structure of Tagalog have been largely motivated by a variety of binding facts as diagnosed by the distribution of

³⁶ In fact, there is evidence that this order is due to a configuration below TP. Small clauses, for instance, which are demonstrably smaller than TP, unanimously show predicate initial order, as shown in (i).

- (i) a. i-t<in>uring=ko=ng kaibígan si=Susan
 CV-<BEG>consider=1S.GEN=LNK friend NOM=Susan
 'I considered Susan a friend'
- b. *i-t<in>uring=ko=ng si=Susan kaibígan
 CV-<BEG>consider=1S.GEN=LNK NOM=Susan friend

reflexive anaphors, bound variable readings, condition C and crossover effects (Byma 1987, Guilfoyle, Hung & Travis 1992, Kroeger 1993, Richards 2000, Rackowski 2002, Aldridge 2004, Wegmuller 1998, among others). Due to space constraints, I am not able to offer a full discussion of this data but I would argue that the lack of clarity in the literature regarding Tagalog binding facts can be connected to the murkiness of binding relations in copular clauses more generally (see Jacobsen 1994 and references therein). However, the current proposal makes a strong and easily testable claim which is that the binding relations between the arguments of an event denoting clause should be replicable with an unambiguous and underived nominal predicate.

Examples of the type in (57) are often used to illustrate the asymmetric nature of reflexive binding in Tagalog, beginning with Schachter (1976, 1977), who captured the reflexive binding facts by direct reference to thematic roles/argument structure rather than surface case relations.

- (57) a. S<in>ampal-Ø ni=Juan ang=sarili=niya
 <BEG>slap-PV GEN=Juan NOM=self=3S.GEN
 ‘Juan slapped himself.’
- b. *S<in>ampal-Ø nang=sarili=niya si=Juan
 <BEG>slap-PV GEN=self=3S.GEN NOM=Juan

What has gone unnoticed is the near identical relation between possessors of unambiguous nominal predicates and the subject, shown in (58).³⁷ Thus, regardless of the best analysis of these facts, a unified explanation

³⁷ Note, however, that some speakers do not judge (58)b to be as bad as (59)b, suggesting a secondary role for agency. In fact, agency and intentionality can even be seen to fully subvert the general pattern. In (1) (from the internet), the context in (a) licenses a genitive agent reflexive in (b) which all speakers queried find acceptable.

- (1) a. Na-*hypnotize*=siya dáti at ang=*task*=niya ay ligáw-an
 PV.NVL.BEG-hypnotize=3S.NOM before and NOM=*task*=3S.GEN TOP court-LV
 ang=sarili=niya.
 NOM=self=3S.GEN
 ‘He was hypnotized and his task was to court himself...’
- b. Na-basted=siya nang=sarili=niya.
 PV.NVL.BEG-busted=3S.NOM GEN=self=3S.GEN
 ‘He was rejected by himself.’

for (58) and (59), the minimal “verbal” pair, appears necessary and would come for free under the present account.³⁸

- (58) Kaaway ni=Tyson ang=sarili=niya
 enemy GEN=Tyson NOM=self=3S.GEN
 ‘Tyson’s enemy is himself.’
 ?*Kaaway nang=sarili=niya si=Tyson
 enemy GEN=self=3S.GEN NOM=Tyson
 ‘Tyson’s enemy is himself.’
- (59) K<in>a~kaaway-∅ ni=Tyson ang=sarili=niya
 <BEG>IMPRF~enemy-PV GEN=Tyson NOM=self=3S.GEN
 ‘Tyson makes himself an enemy.’
 *K<in>a~kaaway-∅ nang=sarili=niya si=Tyson
 <BEG>IMPRF~enemy-PV GEN=self=3S.GEN NOM=Tyson

Another interesting parallel between genitive agents and possessors relates to weak crossover. Aldridge (2004: 161) claims that a structure such as that in (60) induces weak crossover, as expected on her account. In fact, coreference is perfectly possible in such sentences and this is fully expected given the present copular analysis. The possibility of coreference in (60) is parallel to that in the English (61)b, as opposed to (61)a; in neither case does the interrogative bind its trace across a bound pronoun.³⁹

- (60) Sino_i ang=s<in>ampal-∅ nang=asáwa=niya_i
 NOM:who NOM=<BEG>slap-PV GEN=spouse=3S.GEN
 ‘Who_i was slapped by his/her_i spouse?’
- (61) a. Who_i does his_{*i/j} mom love *t*_i?
 b. Who_i *t*_i is his_{i/j} mom’s favorite?

³⁸ Aldridge (2004: 257) proposes TP fronting over the subject to [Spec,CP] to derive the canonical predicate initial order with non-verbal predicates. As this movement can only be conceived of as A'-movement, we expect reconstruction to a position where reflexive binding of the subject by a possessor within the predicate would be impossible.

³⁹ The lack of Weak Crossover effects in (60) can be compared to the impossibility of coreference in (i), which uncontroversially does instantiate movement of the oblique phrase to CP. Paradoxically, however, oblique phrases do not appear to be generated beneath the subject and thus, although (i) instantiates movement, it should not involve the WCO configuration either.

(i) Kaníno_i s<um>ampal ang=asáwa=niya_{*i/j}?
 OBL:who <AV:BEG>slap NOM=spouse=3S.GEN
 ‘Who did his/her spouse slap?’

The binding data is obviously much richer than can be discussed here and must be await further work for full integration into the present analysis. In the following we turn to other consequences of the proposed structure: islands, coordination, floating quantifiers and secondary predication.

4.1. *Islands*

Earlier, we traced the ungrammaticality of genitive phrase “extraction” from the fact that genitive phrases are generally bad predicates in Philippine languages and argument interrogatives are required to be in predicate position. But not only do genitives make for bad predicates, they are bad topics, as well, as seen in (62).

- (62) a. Ang=libro ay b<in>ili-∅ ni=Boboy
 NOM=book TOP <BEG>buy-PV GEN=Boboy
 ‘Boboy bought the book.’
 b. *Ni=Boboy ay b<in>ili-∅ ang=libro
 GEN=Boboy TOP <BEG>buy-PV NOM=book

It is known, however, that possessor extraction is a widely restricted phenomenon (Keenan & Comrie 1977, 1979, Comrie & Keenan 1979) and has special requirements when permitted (see Gavruseva 2000 for some ideas on what these restrictions consist of). Although the comparative data regarding possessor extraction and extraction from NP more generally are quite complex, we can note widespread restrictions on similar structures across a wide range of languages, not least of which is English, as shown in (63) and (64), the functional analogues to the structure proposed here for Tagalog.⁴⁰

- (63) a. These workers are employees of Ronaldo
 b. *[Of Ronaldo]_i, these workers are employees t_i

⁴⁰ It is a curious fact that even *of*-phrase extractions deemed to be grammatical in English, as in (i), are considerably degraded (in my own judgment) without preposition stranding, as shown in (ii). I am not aware of any discussion of these facts in the literature.

(i) Who(m)_i did you see a picture [of t_i]?
 (ii) ?[Of who(m)]_i did you see a picture t_i ?

- (64) a. Ronaldo is an employer of these workers
 b. *[Of these workers]_i, Ronaldo is an employer t_i

Similar extractions of possessors are categorically ungrammatical in Hebrew, requiring pied-piping of the entire DP (regardless of whether it is definite or indefinite), shown in (65) (see also Landau 1999). The same holds true in (related) Levantine Arabic, shown in (66).

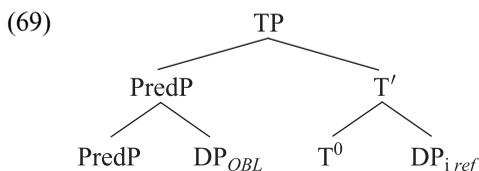
- (65) a. [et=ha=bayt šel mi]_i raita t_i ?
 ACC=DEF=house of who see.PST.2s
 ‘Whose house did you see?’
 b. *[šel mi]_i raita [(et=ha=)bayt t_i]?
 of who saw.PST.2s OBJ=DEF=house
- (66) a. [be:t mi:n]_i šuft t_i ?
 house who see.PST.2s
 ‘Whose house did you see?’
 b. *mi:n_i šuft [be:t t_i]?
 who see.PST.2s house

Extraction from DP famously requires dative case in Hungarian (Szabolcsi 1983), as seen in (67) and (68), similar to the requirement on possessor predicates in Tagalog discussed in (38) and (39) above.

- (67) (a) Mari- \emptyset vendég-e- \emptyset
 the Mari-NOM/GEN guest-POSS.3s
 ‘Mary’s guest’ (Szabolcsi 1983)
- (68) Mari-nak a vendég-e- \emptyset
 Mary-DAT the guest-POSS.3s
 ‘Mary’s guest’ (Szabolcsi 1983)

The “case shift” from genitive to dative in Hungarian extraction appears to agree with the latter’s looser connection to NP associate. This is further corroborated by Den Dikken’s (1999) finding that dative possessors are never seen to trigger agreement within DP. It should come as no surprise then that oblique case phrases can be extracted without any problem in Tagalog. For proponents of a strict locality based approach to extractability, this has been accounted for by treating these phrases as prepositional, and therefore not subject to the same constraints which are faced by DPs (Aldridge 2002, Richards 2000, Rackowski 2002). As noted

by Gerassimova and Sells (2008: 196–197), such an interpretation of the Tagalog *sa* phrase does not jibe well with the facts. First of all, the oblique is sensitive to the $[\pm\text{person}]$ distinction in its complement (surfacing as *kay* with $[\text{+person}, -\text{pl}]$ and *kina* with $[\text{+person}, \text{+pl}]$ complements), just like uncontroversial nominative and genitive case markers. Second, the oblique and genitive are interchangeable in contexts such as comparatives. Third, there exist unambiguous prepositions such as *pára* ‘for’, *búhat* ‘from’, *mulà* ‘from’ and *hanggang* ‘until’, all of which require oblique phrase complements. Fourth, oblique marks definite patients in certain actor voice constructions, an unlikely function for a preposition. Finally, we can add that prepositions would be expected to take case phrase complements but *sa=ang= OBL=NOM=* and *sa=nang= OBL=GEN=* are impossible in Tagalog. The extractability of obliques is seen here as a function of their attachment as adjuncts to PredP, that is, outside the DP island, as illustrated in (69).⁴¹



The full consequences of (69) for the syntax of oblique phrases have yet to be worked out and must remain for later work but the connection between high attachment within PredP and extractability appears firm. As shown in Kaufman (2007), those adjuncts which have no effect on aspectuality are extractable while those which do (e.g. durative and manner adverbs) are not.⁴²

⁴¹ High attachment of locatives as event modifiers has a precedent in Barbiers (1995). If this can be argued to be a function of their interpretation rather than linked to their status as PPs then the analysis could carry over to Tagalog without difficulty. Again, the binding facts are tricky but it appears that linear order plays a role in all but reflexive binding, which is far more tied to argument structure (Kroeger 1993).

⁴² Case was furthermore shown in Kaufman (2007) to be a bad predictor for extractability. Certain actor voice objects may take oblique case with a specific or partitive interpretation but these oblique objects cannot be extracted. Conversely, some high adjuncts, notably temporal and clausal adjuncts are introduced with genitive case but do allow extraction.

4.2. *Coordination*

The structure in (56), similar to right branching specifier approaches, e.g. Guilfoyle, Hung & Travis (1992), makes a clear prediction regarding the constituency of the clause: the nominative phrase should not form a constituent with the predicate while excluding any genitive phrase. We expect the coordination facts in (70) to follow and this is what precisely what is reported by Kroeger (1993), as seen in (71) and (72) (the clisis in (71) is not responsible for its grammaticality).

- (70) a. [[Pred Gen] & [Pred Gen] Nom]
 b. *[[Pred Nom] & [Pred Nom] Gen]

(71) Hú~hugás-an=ko at pu~punas-an=mo
 IMPRF~wash-LV=1S.GEN and IMPRF~wipe-LV=2S.GEN
 ang=manga=pinggan
 NOM=PL=plate
 ‘I’ll wash and you dry the dishes.’ (Kroeger 1993: 34)

(72) ?*Ni-luto-∅ ang=pagkain at h<in>ugas-an
 BEG-cook-PV NOM=food and <BEG>wash-LV
 ang=manga=pinggan ni=Josie
 NOM=PL=plate GEN=Josie
 (For, ‘Josie will cook the food and wash the dishes.’)
 (Kroeger 1993: 34)

It is also an advantage of the structure in (56) that it predicts the unmarked subject final word order without having to resort to scrambling. Although scrambling in the post-predicate domain is very much a feature of Tagalog and other Philippine languages, the tight constituency of the genitive constituent with the predicate has been noted repeatedly.⁴³

⁴³ It has also been noted that genitive agents (of non-actor voice predicates) appear to be more tightly bound to the predicate than genitive phrase objects (of actor phrase predicates). I have no good explanation for this fact but it could correlate with the raising of genitive agents to VoiceP on this account.

Aldridge (2002) notes that one problem for the Guilfoyle, Hung & Travis (1992) account of Tagalog with a rightward specifier hosting the *ang* phrase is that it incorrectly predicts complement clauses will be generated to the left of the subject. Here, I privilege the coordination facts over this important consideration as I find obligatory extraposition of CP to be a plausible solution to the relative positioning of subjects and complement clauses.

4.3. *Floating quantifiers and secondary predication*

It has been observed that floating quantifiers and secondary predicates in Tagalog always associate with the *ang* phrase, as seen in (73) and (74) (Kroeger 1993, Schachter 1994).

(73) [S⟨um⟩ú~súlat na lahat] nang=manga=líham
 <AV.BEG>IMPRF~write LNK all GEN=PL=letter
 sa=manga=kaibígan ang=manga=bátà
 OBL=PL=friend NOM=PL=child
 ‘All the children write letters to friends.’

(74) a. Nag-háin na lasing si=Maria nang=isdà
 AV.BEG-serve LNK drunk NOM Maria GEN=fish
 ‘Maria served the fish drunk.’
 b. #I⟨ni⟩háin na lasing ni=Maria ang=isdà
 CV<BEG>serve LNK drunk GEN=Maria NOM=fish
 (‘The fish was served drunk’)

Unlike more familiar cases of quantifier float, Tagalog possesses what may be better termed “sinking quantifiers”. A quantifier may either appear with the DP it modifies or linked to the predicate head, as seen above. Intermediate positions are ungrammatical. This is predicted if Tagalog lacks quantifier floating altogether. The predicate head and the subject are both nominals in a copular relation and thus quantifying over one will entail quantifying over the other.⁴⁴ The same relation holds with secondary predicates. On this analysis, the more literal rendering of (74)a is ‘Maria was the drunk server of fish’ and (74)b, ‘The fish were the drunk servees of Maria’. The puzzling distribution of quantifiers and secondary predicates is thus seen to be predicted trivially on the nominalist analysis.

⁴⁴ Subtle differences may exist depending on whether the quantifier attaches to the predicate head or the subject but this has never been investigated.

5.0. Conclusion

A rather radical analysis of Tagalog has been presented here which makes no use of the verbal category. While similar analyses of other ergative patterning languages have been proposed, they have also faced certain difficulties (see, for instance, Sadock's 1999 critique of a verbless analysis of Eskimo). It could very well be the case that similar problems will force a compromise in the Tagalog case as well.⁴⁵ The strength of the evidence presented, however, argues for at least some nominal component within the predicate phrase whether or not voice/aspect forms must be treated exactly on par with unambiguous nouns.

Besides helping to account for the interpretation of roots, another advantage of maintaining a nominal analysis is that, for the first time, it is possible to find a *clear* parallel between extraction restrictions in Austronesian and more familiar languages such as English, as shown in §4.1. There is another notable advantage which relates to the typology of ergative languages. As documented by Dixon (1994) and Palancar (2002), ergative case almost always shows a syncretism with another "peripheral case", typically instrumental, ablative or genitive. It has been suggested that these syncretisms correspond to various historical sources for ergativity (Plank 1979, Garrett 1990, among others). Manning (1996) makes a rough cut between ergativity which as has arisen from reanalysis of passive, corresponding to instrumental/ablative case syncretism versus that which has arisen from reanalysis of nominalization, corresponding to genitive case syncretism:

"I believe that historical origin could be a good guide in subdividing the types of ergative languages, although the matter would require much further investigation. Making an initial cut between ergativity arising from a perfective or passive origin (reinterpreting an oblique instrumental or agent as the ergative NP) seems promising. . . . In contrast [to Trask's (1979) typology and predictions], I am suggesting that many languages where ergativity arises from nominalization are syntactically ergative (whereas the ergativity in the Indic Indo-European languages,

⁴⁵ There are asymmetries between voice/aspect forms and bare roots which I have not been able to treat here for lack of space. Some of these are discussed by De Guzman (1996).

for example, seems superficial from the point of view of syntactic behavior).” (Manning 1996: 21)

Surprisingly, the different status of nominalization versus passivization suggested by Manning for syntactic ergativity can even be shown to have reflexes in English. Witness the highly marked genitive extraction in (75) compared to the perfectly natural extraction of a passive agent in (76).⁴⁶

- (75) a. Juan was an employee of Rizal.
 b. *?Of whom was Juan an employee?
- (76) a. Juan was employed by Rizal.
 b. By whom was Juan employed?

The extent to which ergative-instrumental/ablative syncretic languages lack syntactic ergativity remains to be seen. It is a promising start however that the classically syntactically ergative languages, Mayan, Eskimo and Austronesian, all share the genitive-ergative syncretism while Basque, an ergative languages with no unexpected extraction asymmetries, shows an ergative-ablative syncretism. Further work should reveal the value of this typological generalization and consequently the extent to which an a-verbal analysis of genitive-ergative syncretic languages should be maintained.

Cornell University
dak37@cornell.edu

References

- Adriani, Nicolaus. 1893. *Sangireesche spraakkunst*. Leiden.
- Aldridge, Edith. 2002. Nominalization and *Wh*-movement in Seediq and Tagalog. *Language and Linguistics* 3, 2:393–426.
- Aldridge, Edith. 2004. Ergativity and Word Order in Austronesian Languages. Ph.D. dissertation, Cornell University.

⁴⁶ In Kaufman (forthcoming) I show that the breakdown of the nominal predication system in a number of Indonesian language corresponds with the development of an extractable prepositional agent in place of the unextractable genitive agent of Philippine languages.

- Alexiadou, Artemis. 2001. *Functional structure in nominals: Nominalization and ergativity*. (Linguistics Today 42). Amsterdam: John Benjamins.
- Baker, Mark. 2003. *Lexical Categories: Verbs, Nouns and Adjectives*. Cambridge University Press.
- Barkers, Sjeff. 1995. *The Syntax of Interpretation*. Dissertation Leiden University. Holland Academic Graphics, The Hague.
- Barker, Chris. 1998. Episodic -ee in English: A Thematic Role Constraint on New Word Formation. *Language* 74: 695–727.
- Bauer, Laurie. 1983. *English word-formation*. Cambridge: Cambridge University Press.
- Bautista, Ma. Lourdes. 1996. Varieties of Philippine English, In *Readings in Philippine Sociolinguistics*, Bautista, Ma. Lourdes (ed.), Manila: De La Salle University.
- Bloomfield, Leonard. 1917. *Tagalog Texts with Grammatical Analysis*. 3 vols. Urbana, Ill.: University of Illinois.
- Blust, Robert A. 1977. The Proto-Austronesian pronouns and Austronesian subgrouping: a preliminary report. *University of Hawaii Working Papers in Linguistics* 9/2: 1–15.
- Blust, Robert. 1999. Subgrouping, circularity and extinction: some issues in Austronesian comparative linguistics. In E. Zeitoun and P. J.-K. Li, eds. 31–94.
- Blust, Robert. 2002. Notes on the history of ‘focus’ in Austronesian languages. In Faye Wouk & Malcolm Ross, eds. *The History and Typology of Western Austronesian Voice Systems*, 63–78, Canberra: Pacific Linguistics.
- Bok-Bennema, Reineke. 1991. *Case and agreement in Inuit*. Dordrecht: Foris.
- Borer, Hagit. 2004. *Structuring Sense*. Oxford: Oxford University Press.
- Broschart, Jürgen. 1997. Why Tongan does it differently: Categorical distinctions in a language without nouns and verbs. *Linguistic Typology* 1:123–165.
- Byma, Gary. 1987. *Government and Binding in Tagalog: an Ergative Analysis*. Ph.D. dissertation, University of Calgary.
- Capell, Arthur. 1964. Verbal systems in Philippine languages. *Philippine Journal of Science* 93:231–249.
- Cena, R. M. 1977. Patient primacy in Tagalog. Paper presented at the LSA Annual Meeting, Chicago.
- Chang, M. Y. 2000. On Tsou wh-questions: movement or in situ? *Language and Linguistics* 1, 1–18.
- Chierchia, G. & R. Turner. 1988. Semantics and Property Theory. *Linguistics and Philosophy* 11, 261–302.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Chung, Sandra. 1998. *The Syntax of Agreement*. Chicago: University of Chicago.
- Comrie, Bernard and Keenan, Edward. 1979. Noun phrase accessibility revisited. *Language* 65: 649–664.
- Dahl, Otto C. 1973. *Proto-Austronesian*. The Scandinavian Institute of Asian Studies, Monograph Series 25. Lund and London: Curzon Press.
- Davis, Henry and Lisa Matthewson. 1999. On the Functional Determination of Lexical Categories. *Révue Québécoise de Linguistique* 27:2.
- De Guzman, Videia P. 1996. Lexical categories in Tagalog. In Mark Alves, ed., *Pan-Asiatic Linguistics: Proceedings of the Fourth International Symposium on Languages and Linguistics, vol. 1*, pp. 307–96. Mahidol University.
- De Wolf, Charles. 1988. Voice in Austronesian languages of Philippine type: passive, ergative, or neither? In Maayoshi Shibatani, ed. *Passive and Voice*, 143–193. Amsterdam: Benjamin.

- Dempwolff, Otto. 1934–38. Vergleichende Lautlehre des Austroneischen Wortschatzes (3 vols), Beihefte Zeitschrift für Eingeborenensprache 15, 17, 19, Berlin: Deitrich Reimer Verlag.
- Den Dikken, Marcel. 1999. On the structural representation of possession and agreement: The case of (anti-)agreement in Hungarian possessed nominal phrases. In I. Kenesei, ed. *Crossing Boundaries: Theoretical advances in Central and Eastern European Languages*, 137–178. Amsterdam: John Benjamins.
- Den Dikken, Marcel. 2008. *Relators and Linkers: The Syntax of Predication, Predicate Inversion and Copulas*. Cambridge: MIT Press.
- Dixon, Robert. 1994. *Ergativity*. Cambridge University Press.
- Donohue, Mark. 1999. *A Grammar of Tukang Besi*. Amsterdam: Mouton de Gruyter.
- Dowty, David. 1979. *Word Meaning in Montague Grammar*. Reidel, Dordrecht.
- Dryer, Matthew S. 2007. “Clause types”. In *Clause Structure, Language Typology and Syntactic Description*, Vol. 1, edited by Timothy Shopen, pp. 224–275. Second Edition. Cambridge University Press.
- Evans, Nicholas & Osada, Toshiaki. 2005. Mundari: The myth of a language without word classes. *Linguistic Typology* 9/3:351–90.
- Fabb, Nigel. 1988. English suffixation is constrained only by selectional restrictions. *Natural Language and Linguistic Theory*, 6:527–539.
- Ferrell, Raleigh. 1982. *Paiwan Dictionary*. Canberra: Pacific Linguistics.
- Foley William A. 1998. Symmetrical voice systems and precategoriality in Philippine languages. Paper presented at the 3rd LFG conference, Brisbane, Australia.
- Foley, William. 2008. The place of Philippine languages in a typology Of voice systems. In *Voice and Grammatical Relations in Austronesian Languages*, Peter Austin and Simon Musgrave, eds. Stanford: CSLI.
- Garrett, Andrew. 1990. The origin of split NP ergativity. *Language* 66: 261–296.
- Gavruseva, Elena. 2000. “On the Syntax of Possessor Extraction”. *Lingua* 110, 743–772.
- Georgopoulos, Carol. 1991. *Syntactic Variables*. Dordrecht: Kluwer Academic Publishers.
- Gil, David. 1993. Tagalog semantics. *Proceedings of BLS 19*. 390–403.
- Gil, David. 1995. Parts of Speech in Tagalog. In *Papers from the Third Annual Meeting of the Southeast Asian Linguistics Society*, Mark Alves (ed.). 67–90. Arizona State University, Program for Southeast Asian Studies.
- Gil, David. 2000. Syntactic categories, cross-linguistic variation and universal grammar. In *Approaches to the Typology of Word Classes*, Petra Vogel and Bernard Comrie, eds. Berlin: Walter de Gruyter: 173–216.
- Gerassimova, Veronica and Peter Sells. 2008. Long-Distance Dependencies in Tagalog: The Case for Raising. In Charles B. Chang and Hannah J. Haynie, eds. *Proceedings of the 26th West Coast Conference on Formal Linguistics*, 190–198. Somerville, MA: Cascadilla Proceedings Project.
- 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.
- Heim, Irene and Angelika Kratzer. 1998. *Semantics in Generative Grammar*. Oxford, Basil, Blackwell.
- Hengeveld, Kees. 1992. *Non-verbal Predication: Theory, Typology, Diachrony*. Berlin: Mouton de Gruyter.
- Himmelmann, Nikolaus. 2008. Lexical categories and voice in Tagalog. In Peter K. Austin and Simon Musgrave, eds. *Voice and Grammatical Functions in Austronesian Languages*. Stanford: CSLI.

- Hockett, Charles F. 1958. *A course in modern linguistics*. New York: Macmillan.
- Horn, G. M. 1974. The Noun Phrase Constraint. Ph.D. dissertation, MIT.
- Johns, Alana. 1992. Deriving Ergativity. *Linguistic Inquiry*, 23:1, pp. 58–87.
- Kaufman, Daniel. 2007. Genitive alignment and Malayo-Polynesian languages. Talk presented at 'Beyond 'Focus' and Ergativity: Towards a More Comprehensive View of Austronesian Morphosyntax' Zentrum für Allgemeine Sprachwissenschaft, Berlin.
- Kaufman, Daniel. Forthcoming. Austronesian typology and the nominalist hypothesis. In Adelaar, A. and Pawley, A. (eds.), *Festschrift for _____*.
- Keenan, Edward and Comrie, Bernard. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8. pp. 63–69.
- Keenan, Edward and Comrie, Bernard. 1979. Data on the Noun Phrase Accessibility Hierarchy. *Language* 55, 333–351.
- Keenan, Edward. 1976. Remarkable Subjects in Malagasy. In Charles Li, ed., *Subject and Topic*, 249–299. New York: Academic Press.
- Kratzer, Angelika. 1996. Severing the external argument from its verb. In *Phrase Structure and the Lexicon*. J. Rooryck and L. Zaring, eds. Dordrecht, Kluwer: 109–137.
- Kroeger, Paul R. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford: Stanford University Press.
- Kroeger, Paul. 1998. Nouns and verbs in Tagalog: a reply to Foley. Paper presented at the third LFG Conference, Brisbane, Australia.
- Landau, Idan. 1999. *Elements of control*. Ph.D. dissertation, MIT, Cambridge, MA.
- Li, Yafei. 1990. X⁰-binding and verb incorporation. *Linguistic Inquiry* 21:399–426.
- Lopez, Cecilio. 1928. *Comparison of Tagalog and Iloko*. Ph.D. dissertation, University of Hamburg.
- Louwse, John. 1988. *The morphosyntax of Una in Relation to Discourse Structure*. RSPAS, Canberra: Australian National University.
- Maclachlan, Anna and Masanori Nakamura. 1997. Case-checking and Specificity in Tagalog. *The Linguistic Review* 14:307–333.
- Malchukov, Andrej. 2004. *Nominalization/Verbalization*. LINCOM Studies in Language Typology. Munich: Lincom Europa.
- Manning, Christopher. 1996. *Ergativity: Argument Structure and Grammatical Relations*. Stanford: Center for the Study of Language and Information.
- Marantz, Alec. 2001. Words and Things. Handout, MIT.
- Marantz, A. 1997. No Escape from Syntax: Don't Try Morphological Analysis in the Privacy of Your Own Lexicon. In A. Dimitriadis, L. Siegel, et al., eds., *University of Pennsylvania Working Papers in Linguistics*, Vol. 4.2, Proceedings of the 21st Annual Penn Linguistics Colloquium, pp. 201–225.
- Marchand, Hans. 1969. *The categories and types of present-day English word-formation*. Munich: C. H. Beck.
- Massam, Diane. 2000. VSO and VOS: aspects of Niuean word order. In Andrew Carnie and Eithne Guilfoyle, eds., *The syntax of verb-initial languages*: 97–116.
- Matthewson, Lisa and Demirdache, Hamida. 1996. On the Universality of Syntactic Categories. *Proceedings of the North East Linguistic Society* 25, 79–93.
- Mercado, Raphael. 2005. XP Predicate-Fronting in Tagalog. Talk given at AFLA XII.
- Nash, Léa. 1996. The internal ergative subject hypothesis. *NELS* 26, 195–209.
- Naylor, Paz Buenaventura. 1980. Linking, relation marking, and Tagalog syntax. In Naylor ed. 1980:35–50.

- Ntelitheos, Dimitris. 2006. The Morphosyntax of Nominalizations: A Case Study. Ph.D. dissertation, University of California, Los Angeles.
- Oda, Kenji. 2002. Wh- questions in V-initial languages. MA Thesis, University of Toronto.
- Palancar, Enrique. 2002. *The Origin of Agent Markers*. Berlin: Akademie Verlag.
- Partee, B. H., & Borschev, V. B. 2000. Genitives, Relational Nouns, and the Argument Modifier Distinction. In *ZAS Papers in Linguistics* 17, eds. Fabricius-Hansen, C., Lang, E., & Maienborn, C., 177–201. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Sprachtypologie und Universalienforschung.
- Paul, Ileana. 2000. *Malagasy Clause Structure*. Ph.D. dissertation, McGill University.
- Paul, Ileana. 2001. Concealed pseudo-clefts. *Lingua* 111: 707–727.
- Pearson, Matthew. 2005. The Malagasy Subject/Topic as an A' element. *Natural Language and Linguistic Theory* 23, 381–457.
- Pesetsky, David. 1995. *Zero Syntax: Experiencers and Cascades*. Cambridge: MIT Press.
- Plank, Frans. 1979. Ergativity, Syntactic typology and universal grammar: some past and present viewpoints, in *Ergativity: Towards a theory of grammatical relations*, Frans Plank (ed.). London: Academic Press.
- Potsdam, Eric. 2006. More Concealed Pseudoclefts and the Clausal Typing Hypothesis. *Lingua* 116, 2154–2182.
- Rackowski, Andrea. 2002. *The Structure of Tagalog: Specificity, Voice, and the Distribution of Arguments*. Ph.D. dissertation, MIT.
- Richards, Norvin. 1998. “Syntax versus Semantics in Tagalog wh-extraction,” in *UCLA Occasional Papers in Linguistics* 21: Recent Papers in Austronesian Linguistics, Matthew Pearson, ed. Los Angeles: UCLA Department of Linguistics, 259–276.
- Richards, Norvin. 1999. “Complementizer cliticization in Tagalog and English,” in *Toronto Working Papers in Linguistics* 16.2: Proceedings of AFLA 6, Carolyn Smallwood and Catherine Kitto, eds. Toronto: Toronto Working Papers in Linguistics, 297–312.
- Richards, Norvin. 2000. Another Look at Tagalog Subjects. In: Ileana Paul, Vivianne Phillips and Lisa Travis, eds. *Formal Issues in Austronesian Linguistics*, pp. 105–116. Dordrecht: Kluwer.
- Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice marking. In: Faye Wouk & Malcolm Ross, eds. *The History and Typology of Western Austronesian Voice Systems*, 63–78, Canberra: Pacific Linguistics.
- Sabbagh, Joseph. 2005. *Non-verbal argument structure: evidence from Tagalog*. Ph.D. dissertation, MIT.
- Sadock, Jerrold M. 1999. The Nominalist Theory of Eskimo: A Case Study in Scientific Self-Deception. *International Journal of American Linguistics* 65.4: 383–406.
- Sagart, L. 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43, pp. 411–440.
- Schachter, Paul. 1977. Reference-related and Role-related properties of subject. In Peter Cole & Jerrold Sadock (eds.), *Syntax and Semantics 8: Grammatical Relations*. New York: Academic Press.
- Schachter, Paul. 1976. The Subject in Philippine Languages: Topic, actor, actor-topic, or none of the above. In Charles Li, ed., *Subject and Topic*, 491–518. New York: Academic Press.
- Schachter, Paul. 1994. The Subject in Tagalog: Still None of the Above. *UCLA Occasional Papers in Linguistics* 15.

- Schachter, P. and Fe T. Otnes. 1972. *Tagalog Reference Grammar*. Berkeley, CA: University of California Press.
- Scheerer, Otto. 1924. On the Essential Difference Between the Verbs of the European and the Philippine Languages. *Philippine Journal of Education* 7 pp. 1–10.
- Seiter, William. 1975. Information questions in Philippine languages. Ms., University of California, San Diego.
- Starosta, Stanley, Andrew Pawley & Lawrence Reid. 1982. The evolution of focus in Austronesian. In *Papers from the Third International Conference on Austronesian Linguistics 1: Tracking the Travellers*. Halim, Amran, Carrington, Lois, and Wurm, S. A. (eds.). Canberra: Pacific Linguistics, 145–170.
- Szabolcsi, Anna. 1983. The possessor that ran away from home. *The Linguistic Review* 3:89–102.
- Tuuk, H. N. van der. 1864–1867. *Tobasche Sprachkunst*, 2 vols, Amsterdam: Muller.
- Trask, R. L. 1979. On the origins of ergativity. In *Ergativity: Towards a theory of grammatical relations*, Frans Plank (ed.), 385–404. London: Academic Press.
- Travis, Lisa. 2000. Event Structure in Syntax. In C. Tenny and J. Pustejovsky, eds., *Events as Grammatical Objects*, 145–185. Stanford: Center for the Study of Language and Information.
- van Eijk, Jan, and Thom Hess. 1986. Noun and verb in Salishan. *Lingua* 69:319–31.
- Wegmüller, Ursula 1998. Sentence Structure and Ergativity in Tagalog. (Arbeitspapiere, 36.) Bern: Universität Bern, Institut für Sprachwissenschaft.
- Williams, Edwin. 1994. *Thematic Structure in Syntax*. Cambridge: The MIT Press.
- Wolff, John. 1973. Verbal Inflection in Proto-Austronesian. In: A. B. Gonzalez, ed. *Parangal Kay Cecilio Lopez*, pp. 71–91. Quezon City: Philippine Journal of Linguistics (Special monograph 4).
- Wolff, John U. 1993. Why roots add the affixes with which they occur. In *Topics in Descriptive Austronesian Linguistics* ed. G. P. Reesink. 217–244. Leiden: Vakgroep Talen en Culturen van Zuidoost-Azië en Oceanië. Semaian 11.
- Wolff, John U., Ma. Theresa, C. Centeno and Der-Hwa V. Rau. 2005. *Pilipino through Self-Instruction*. Second Edition. Cornell University Southeast Asia Program, Ithaca, New York: Cornell University.
- Wouk, Faye & Malcolm Ross. 2002. *The History and Typology of Western Austronesian Voice Systems*. Canberra: Pacific Linguistics.