

Formosan Languages and Linguistic Typology^{*}

Malcolm Ross and Stacy Fang-ching Teng

The Australian National University

There is quite a longstanding convention whereby Formosan languages (the Austronesian languages of Taiwan) are described using a framework and terminology developed by linguists working on the Austronesian languages of the Philippines. Linguists using this terminology talk, for example, about the ‘focus system’ of verbal constructions and about the ‘topic’ of the clause. Because this terminology is (i) unfamiliar to linguists working outside Taiwan and the Philippines and (ii) deficient in certain respects (it is often unclear, for example, what the function of the ‘topic’ is), descriptions of Formosan and Philippine languages using this terminology are sometimes somewhat opaque and rather difficult for other linguists to read.

In this paper we attempt to re-frame the description of certain aspects of the morphosyntax of Formosan languages in terms more familiar to typologists and to linguists working in other parts of the world. Among other things, the notion of ‘topic’ is re-examined, and the ‘focus system’ is reformulated in a framework derived from the work of the typologist William Croft. A re-framing of this kind should lead to two main benefits: increased comprehensibility and increased differentiation among the descriptions of Formosan languages (which are sometimes treated as being more similar to each other than they appear to be).

Key words: Puyuma, Formosan languages, linguistic typology, grammar

1. Introduction

Much description of Formosan languages has been and continues to be carried out using a framework and terminology first developed by linguists working on the Austronesian languages of the Philippines. Paying tribute to this origin, we shall refer to

^{*} The text of this paper was largely written by the first author. The Puyuma data used in the paper were provided by the second author, who is completing a doctoral dissertation on Puyuma at the Australian National University. The data are drawn from the large corpus of Puyuma oral texts that she has collected, transcribed, and analyzed. Many of the ideas in the paper are hers or are the outcome of our discussions about Puyuma. We would like to thank Shuanfan Huang and an anonymous reviewer, whose comments we found most helpful in our revision of the paper.

these as the **Philippinist** framework and terminology, despite the fact that they are also used in describing Formosan languages.

We would like to argue that the time has come to abandon the Philippinist approach in Formosan studies (just as it has been abandoned by many working in the Philippines) and to adopt ways of talking about Formosan languages that are more familiar to people working on languages outside Taiwan and the Philippines and particularly to linguists working on typological comparison.

The Philippinist approach has been applied to the description of every Formosan language except Rukai. It cannot be applied to Rukai because the language does not have the features that the Philippinist framework is designed to describe, and this betrays the problem with the Philippinist approach. It is not designed to describe grammars in general, but only grammars of a structural type dubbed “Philippine-type languages” by Himmelmann (2002) and found in a limited area within the Austronesian family.¹

We have three reasons for proposing the abandonment of the Philippinist approach, and all three arise from its limited applicability.

- (1) a. Philippinist terminology is opaque. For example, Philippinists use the terms ‘topic’ and ‘focus’ in ways quite different from modern usage among linguists working in other parts of the world. At the same time Philippinists avoid a raft of terms that other linguists commonly use. An obvious example is ‘subject’.
- b. The difference between the Philippinist and other approaches is not simply one of terminology. It is also one of framework, and this has the consequence that certain insights about the grammars of Formosan languages are either missed or ignored.²
- c. Application of the Philippinist approach to Formosan languages (except Rukai) implies that they are more similar to each other and to the languages of the Philippines than is really the case. Giving up the Philippinist approach should make some of the interesting differences among them much clearer.

The bottom line is that Philippine-type languages are typologically interesting in a number of ways, but they are often neglected in the typological literature and will only become better known if they are described using a framework and terminology that the

¹ This area embraces Taiwan, the Philippines, parts of northern Borneo, a portion of northern Sulawesi, and Madagascar.

² The first author has used Philippinist terminology in various publications, but has found it increasingly difficult to use when wanting to say things that transcend its framework.

rest of the linguistic world understands.

The Philippinist approach originated in a description of the southern Philippine language Maranao by McKaughan (1958). It dates from late in the Structuralist era, before the arrival of either typological or Chomskyan universal categories, and at a time when the goal was to describe a language on the basis of its own categories.³ McKaughan's concern was to describe Maranao without imposing on it categories derived from the analysis of English (McKaughan 1973), and this he did successfully. It was a logical next step to apply the same framework to other Philippine languages and then to Formosan languages which shared the features that McKaughan's approach was designed to capture.

2. The Philippinist framework

2.1 Major features

Before we go on to outline how we would prefer to talk about Formosan languages, we shall briefly describe the Philippinist approach. The paradigm of examples in (2) is typical of those used to illustrate Philippinist analyses. It is drawn from the introduction to Ferrell's (1982:31) dictionary of Paiwan (southern Taiwan). Such paradigms are obviously elicited and bear only a limited resemblance to sentences in real discourse, but they are central to the Philippinist approach and allow an illustration of its major features.

(2) Paiwan (southern Taiwan)

- a. *q<m>ahup a tsautsau tua vavuy i (tua) gadu tua vuluq*
 <AF>hunt TPC man OBL pig LOC (OBL) mountain OBL spear
 'The man hunts the pigs in the mountains with a spear.'
- b. *qahup-en nua tsautsau a vavuy i (tua) gadu tua vuluq*
 hunt-OF GEN man TPC pig LOC (OBL) mountain OBL spear
 'The man hunts the pigs in the mountains with a spear.'
- c. *qahup-an nua tsautsau tua vavuy a gadu tua vuluq*
 hunt-LF GEN man OBL pig TPC mountain OBL spear
 'The man hunts the pigs in the mountains with a spear.'
- d. *si-qahup nua tsautsau tua vavuy i (tua) gadu a vuluq*
 IF-hunt GEN man OBL pig LOC (OBL) mountain TPC spear
 'The man hunts the pigs in the mountains with a spear.'

³ Describing a language on the basis of its own categories was a laudable aim in our view (see §3.1), but doing so in a way which hindered comparison across languages was unfortunate.

The free glosses of all four examples are the same according to Ferrell. We return to this below. The interlinear glosses are ours, but based as far as possible on Ferrell. Ferrell (1982:30) writes, “All possible semantic roles for NP’s in the Paiwan sentence are subsumed under only four overtly-marked categories”, listed as

- (A) Agent/Actor
- (O) Object/Goal/Patient
- (R) Referent: spatial/temporal locus, indirect object, beneficiary
- (I) Instrument/Cause/Motivation/Origin

‘The verb ... is obligatorily inflected with an affix which specifies which particular one of these four possible semantic role categories is “in focus” in the sentence.’ These affixes are labeled AF (for Agent Focus), OF (for Object Focus) and so on in (2). The term **focus** is characteristic of Philippinist descriptions, and does not have the meaning that ‘focus’ has acquired in the study of information structure, e.g. in the work of Lambrecht (1994). Instead, the item which Ferrell terms ‘in focus’ is, we suggest, better labeled ‘subject’ (or ‘nominative’ or ‘absolutive’). In Paiwan, a common NP that is ‘in focus’ is introduced by the NP marker *a*, as (2) shows. Thus in (2a) *q<m>alup* is glossed ‘<AF>hunt’, since the infix <*m*> indicates that the ‘in focus’ item *a tsautsau* ‘the man’ is the Agent. Confusingly, Philippinists often call the ‘in focus’ item the ‘topic’ (!), here glossed ‘TPC’ in (2). Ferrell himself does not use ‘topic’, and there is good evidence (Ferrell 1982:34) that he had an intuitive understanding of the function of the ‘in focus’ item and knew that it had nothing to do with the category labeled ‘focus’ in general linguistics today.

In (2b) *qalup-en* is glossed ‘hunt-OF’, since the suffix *-en* indicates that the ‘topic’, *a vavuy* ‘the pig’, is the Patient.⁴ It will be clear that if the ‘topic’ resembles a subject, then the ‘focus’ affixes on the verb like <*m*> and *-en* have a function similar to that voice (active, passive) morphemes in a language like English. And somewhat as the agent in an English passive is introduced by the preposition *by*, so the agent of an OF verb in Paiwan is marked by the genitive NP marker *nua* (the term ‘genitive’ is Ferrell’s: the form also marks a possessor).

In (2c) and (2d) the pattern is similar to the one in (2b). The ‘semantic role’ of the topic is indicated by the ‘focus’ affix on the verb, and the agent is marked by the genitive NP marker. Thus *a gadu* ‘the mountain’ is (locative) topic of (2c), and *a vuluq* ‘the spear’ is (instrument) topic of (2d).

In all four sentences in (2), common noun phrases that are neither topic nor agent

⁴ Ferrell’s use of ‘Object’ here is typical of earlier Philippinist work; a more recent tendency is to use ‘Goal’, but this often has an allative sense in modern linguistics.

are marked with the NP marker *tua*, which we gloss ‘oblique’. A locative oblique may be marked more explicitly with the preposition *i*, and the NP marker may be omitted.⁵

Formosanists will question the identical translation which Ferrell provides for all four sentences. First, they will expect the topic usually to be rendered with an English definite, i.e. they would expect ‘the spear’ in (2d). This seems to be an oversight on Ferrell’s part. Second, they may expect the oblique-marked patient of the AF verb in (2a), *tua vavuy* to be rendered as indefinite (‘pigs’), since the patient of an independent AF clause is interpreted as indefinite in many Philippine-type languages. Paiwan is an exception to this generalization: the oblique-marked patient of an AF verb may refer to a definite entity (Anna Hsiou-chuan Chang, pers. comm.).

It is obvious from this account that the relationship between NP markers (topic, genitive, oblique) and ‘focus marking’ on the verb is crucial to Paiwan morphosyntax. The cases encoded by NP markers are also encoded in personal pronoun sets, as the examples in (3), from Ferrell (1972), show. In Paiwan a free pronoun or its enclitic variant (= *aken* in (3a)) is used as ‘topic’, and a set of genitive proclitics mark genitive agents (*ku*= in (3b-d)).⁶

- (3) a. *q<m>alup=aken tua vavuy i (tua) gadu*
 <AF>hunt=TPC:1S OBL pig LOC (OBL) mountain
 ‘I hunt boar on the mountain.’
- b. *ku=qalup-en a vavuy i (tua) gadu*
 GEN:1S=hunt-OF TPC pig LOC (OBL) mountain
 ‘I hunt boar on the mountain.’
- c. *ku=qalup-an a gadu tua vavuy*
 GEN:1S=hunt-LF TPC mountain OBL pig
 ‘I hunt boar on the mountain.’
- d. *ku=si-qalup a vuluq tua vavuy*
 GEN:1S=IF-hunt TPC spear OBL pig
 ‘I hunt boar with a spear.’

In general, other Philippine-type languages have morphosyntactic systems similar enough to Paiwan to be described using similar terminology, although there are considerable differences of detail. For example, Paiwan pronouns are atypical, as genitive clitic pronouns are enclitic (not proclitic) in many languages, and other languages often have

⁵ This means, according to Ferrell, that *i tua gadu*, *tua gadu* and *i gadu* are all acceptable for ‘in the mountains’.

⁶ There are also free genitive pronouns, as well as free oblique pronouns, but these do not concern us here (see Egli 1990:154-157).

a set of enclitic topic pronouns as well as a free set.

It is clear from (2) and (3) that clauses with OF, LF, and IF verbs are more similar in structure to each other than any of them is to AF verbs, as all three have a genitive agent, whereas a clause with an AF verb has a topic agent. Since there are generalizations to be made which cover clauses with OF, LF and IF verbs, Formosanists quite often use the cover-term ‘non-agent focus’ (NAF). The first Formosanist to do so seems to have been Tsuchida (1976).

The major features of the Philippinist approach are:

- (4) a. Presenting the ‘focus’ forms as a (usually four-member) paradigm.
- b. Using the term ‘focus’ for the system whereby certain verbal affixes allegedly mark the semantic role of the ‘in focus’ or ‘topic’ NP.
- c. Using the term ‘topic’ for the NP singled out in this way.
- d. Naming the verb forms with labels like ‘agent focus’ (or ‘actor focus’), ‘object focus’ (or ‘goal focus’ or ‘patient focus’), ‘referential focus’ (or, more usually, ‘locative focus’) and ‘instrument/beneficiary focus’.
- e. Ignoring the transitivity of voice forms.

There is some variation in terminology among linguists who employ a Philippinist approach. This is especially true among Formosanists, for whom, for example, ‘topic’ rarely has the sense of (4c) and ‘referential’ is not a synonym of ‘locative’ as it was in the early days of the Philippinist approach (see the quotation from McKaughan below). Furthermore, (4e) does not apply to all practitioners.

2.2 A little historiography

The central feature of the Philippinist framework is (4a), since it defines a particular way of understanding the system. The new understanding and its accompanying terminology emerged gradually among members of the Summer Institute of Linguistics (SIL) working in the Philippines. In the early days of its emergence, there were two terminological schools: one talked about the ‘topic’ but retained the term ‘voice’ (which in the English-language literature goes back at least to Blake 1906a, b), and the other talked about ‘focus’ and the ‘in focus’ NP. Eventually the two terminologies merged, and ‘topic’ and ‘focus’ came to be used together.

French (1988:4) cites McKaughan (1958:18) as the first account of a focus system in what we are calling the Philippinist framework. McKaughan originated the term ‘topic’ but retained ‘voice’.⁷

⁷ French (1988:4) quotes this passage rather inaccurately.

The syntactic relations between the substantive phrases and the verb center are indicated by the case marking particles...and also by certain morphemes included in the verb. We call the latter *voice* marking affixes. The voice marking affixes indicate the specific syntactic relations between the *topic* ... and the verb. The case marking particles indicate the syntactic relations between any substantive phrase and the verb. The following syntactic relations between the topic and the verb are marked by verb inflection: (1) *subjective* [actor] indicating that the topic is the originator of the action; (2) *objective* [goal or patient] indicating that the topic is the goal of the action; (3) *referential* [locative] indicating that the topic is the beneficiary or location of the action; and (4) *instrumental* indicating that the topic is the means used to bring about the action. All verbs are inflected to indicate one of these syntactic relations even when the topic is not actualized by a substantive phrase (or a substitute). (McKaughan 1958:18)

McKaughan's terms for the four focus forms are unfamiliar to today's Philippinists, and we have added the more familiar terms in brackets. McKaughan was writing at a time when morphosyntactic case and semantic role were often confused; hence his use of 'subjective' and 'objective' for 'actor' and 'goal' or 'patient'. He felt that the term 'subject' should be reserved for the agent, and this caused him to choose the term 'topic' for the subject-like noun phrase (McKaughan 1973:206).

At about the same time, Thomas (1958) was writing about Mansaka (a Central Philippine language of Mindanao).⁸ He belongs to the 'focus' school, and provides a paradigm using the more familiar terms 'actor focus', 'goal focus' and 'referent focus'.⁹ Almost simultaneously Dean (1958) uses almost the same terms in describing Bilaan (South Mindanao), speaking of "'voice" or "focus"' as well as 'actor focus' and 'goal focus' (Bilaan only has a two-way voice contrast).

Phyllis Healey (1960:19) was perhaps the first to use 'topic' and 'focus' together, in her Agta grammar when she writes about 'the topic or item in focus' and 'focusing', but ironically the presentation is not otherwise in a Philippinist framework (it lacks (4a)).

In *Papers in Philippine Linguistics*, a collection of papers by SIL members published in 1964 as volume 3 of the journal *Oceanic Linguistics*, Phyllis Healey's

⁸ Blust (2002:64) points to a footnote of Phyllis Healey's (1960:103) in which she attributes the first use of 'focus' to Alan Healey (1958), but this piece is not presented in a Philippinist framework and we cannot find the term 'focus' in it. Reid (1966:9-10), however, refers to unpublished notes with the same title where he says the term is used.

⁹ Mansaka referent focus combines the functions exercised by locative and instrument focus in other Philippine-type languages.

‘topic’ and ‘focus’ usage won the day, adopted by Forster (1964), Shand (1964), Morey (1964), Miller (1964), and Newell (1964). These publications are decidedly Philippinist and present the ‘focus’ forms of each language as a paradigm. Some writers use McKaughan’s terms (‘subjective’ etc.), whilst others follow Thomas, Dean & Pike (1963) and use ‘actor’, ‘goal’, and so on. The latter have won out in the literature. French (1988) offers a number of quotations showing that some Philippinists took the term ‘focus’ to mean that the ‘topic’ has an information-structure function similar to the ‘focus’ in its modern sense. This is unfortunate, as it is clearly wrong.

As far as we can tell, the first scholar to apply a Philippinist analysis to Formosan languages was Ferrell (1972), who provides very brief sketches of the verb systems of Paiwan, Amis, Bunun, Squaliq Atayal, and Tsou.¹⁰ He writes:

The Paiwanic languages more closely resemble Philippine ones in indicating the relationship between the verb and the noun phrases of a sentence by both verbal affixes and noun phrase-initial construction-marking particles or corresponding pronoun forms, thus “focusing” the sentence upon the actor (AF), object or goal (OF), locative or benefactive referent (RF), or instrument (IF) of the action. (Ferrell 1972:121)

3. Describing Puyuma

3.1 Preliminaries

Our approach here is to describe aspects of one Formosan language, Puyuma, and to provide some commentary on our choice of terms as or after we first use them, so that the reader has already seen a sample of their usage in practice. Little of our approach is new. What is new, perhaps, is that we would like to displace Philippinist terminology in its entirety.¹¹

It is beyond the scope of this paper to provide a detailed history of previous and present approaches to Philippine-type verbal systems, but such a survey is provided by Liao (2004).

The choice of Puyuma was determined by circumstances (see the Acknowledgements).

¹⁰ The first two major descriptions of Formosan languages in English were Tung (1964) on Tsou, which does not employ a Philippinist framework, and Li (1973) on Rukai, the one Formosan language without a ‘focus’ system. An early work on Bunun, Wu (1969), is also not a Philippinist work. Despite the title *Topic and Focus in Bunun*, Jeng (1977) is not written in a Philippinist framework but in one based on the Case Grammar of Fillmore (1968).

¹¹ Reid & Liao (2004) offer a similar attempt at a non-Philippinist account of Philippine languages, but their terminology is largely drawn from Stanley Starosta’s Lexicase framework.

A well analyzed data set from any other extant Formosan language except Rukai would probably have served our purpose just as well.¹²

Much of the east coast of Taiwan is occupied by speakers of Amis, the Formosan language with the largest number of speakers. Almost contiguous with Amis at its southernmost extent is Puyuma, with just a few thousand speakers. The border between Amis and Puyuma today is effectively the city of Táidōng (台東市), with the Puyuma settlement of Nánwáng (南王里) on its southern outskirts. The data used here are drawn mostly from the second author's corpus of Nanwang texts.

Blust (1999) divides Austronesian languages into ten phylogenetic groups on phonological grounds. One group, Malayo-Polynesian, includes all languages spoken outside Taiwan.¹³ The members of the other nine groups are (or were) all found in Taiwan, and Puyuma is the lone member of one of these groups. There are dialects within Puyuma, but scholars disagree about how to classify them (Huteson 2003). There is consensus, however, that there is a major division between Nanwang, together with nearby settlements, and the Katipol dialect of villages and townships a few kilometres further south. The major descriptive works on Puyuma to date are Tsuchida (1980, 1983) on Tamalakaw, a variety of Katipol, and Cauquelin (1991b, a) and Huang (2000) on Nanwang.¹⁴

There are considerable commonalities among the Formosan languages other than Rukai, but each language has its peculiarities. A Puyuma peculiarity is that its 'focus' or voice morphology is rather different from that of other Formosan languages. Possible reasons for this are discussed in Appendix A.

The theoretical underpinnings of this analysis are drawn from William Croft's Radical Construction Grammar, outlined by Croft (2001). There are a number of flavors of Construction Grammar (Kay & Fillmore 1999, Fillmore 1988, Goldberg 1995, 2002, Michaelis & Lambrecht 1996, Michaelis 2000, Croft 2001), all sharing the fundamental claim that the grammar of a language consists of a taxonomy of constructions, and that a construction is a form-meaning pairing. To this Croft (2001:32-33) adds further claims:

- All constructions and categories are language-specific, and that there is no such thing as a universal (crosslinguistic) construction. There is, for example, no universal Subject-Predicate construction (a claim that is important for the description of Philippine-type languages).¹⁵

¹² One of the advantages of the framework proposed here is that it can also be applied to Rukai.

¹³ Yami, spoken on Orchid Island (Lányǔ 蘭嶼鄉, just west of the southern tip of Taiwan), is Malayo-Polynesian.

¹⁴ Ting (1978) includes comparative wordlists for another four subdialects.

¹⁵ This claim has also been made by Dryer (1997) and in somewhat different terms by Kibrik

- The primitive units of language are constructions. This means that constructions are not defined by the units that occur in them. Instead, it is constructions that define their component slots and their fillers. Since parts of speech—noun, adjective, verb etc.—are defined by the constructions of each language, there are also no universal parts of speech.

Under these assumptions typology might appear impossible. What criteria are there for deciding which construction in Language A can legitimately be compared with which construction in Language B? For Croft, the answer is connected with the view of a construction as a form-meaning or form-function pairing. Constructions to be compared are those having a given meaning or performing a given function in the two languages.

These underpinnings explain why we do not pay much attention to arguments for universal syntactic categories like ‘subject’ or ‘voice’. We think that these are useful terms denoting units with similar functions in many languages, but we do not think that they are universal syntactic categories.

3.2 Two verbal constructions

The paradigm in (5) resembles those in (2) and (3). The examples are shown with two sets of glosses. The first set matches the categories of the Paiwan examples in (2) and (3). The second set represents the analysis we are putting forward here, and is explained and justified below.¹⁶

(1997). Dryer maintains that there are no universal grammatical relations, but only similarities among language-particular grammatical relations for which functional, cognitive, and semantic explanations may be sought. Croft (2001:32-33) extends Dryer’s position to constructions in general.

¹⁶ In the orthography used here, *T*, *D*, *L* are retroflex, and the apostrophe (‘) represents a glottal stop. The vowel *e* represents schwa ([ə]). Abbreviations used in (non-Philippinist) glosses of Puyuma examples are: :1S/:1EP/:1IP 1st person singular/exclusive plural/inclusive plural, :2S/:2P 2nd person singular/plural, :3S/:3P 3rd person singular/plural, ASP aspect, CAUS causative, COP copula, D: disjunctive (free), DEF definite, GEN: genitive, GER gerund, HORT hortative, IMPF imperfect, INDEF indefinite, IRR irrealis, ITR intransitive, LOC: locative, NEG negator, NOM: nominative, NSR nominalizer, OBL: oblique, P: possessor, PERF perfect, PERS personal, PROG progressive, REAL realis, REDUP reduplication, S singular, TPC topic, TR1 first transitive, TR2 second transitive, TR3 third transitive.

- (5) a. *tusuk=ku Da Lutung*
 <AF>pierce=TPC:1S OBL:INDEF monkey
 <ITR>pierce=NOM:1S OBL:INDEF monkey
 ‘I speared a monkey.’
- b. *ku=tusuk-aw na Lutung*
 GEN:1S=pierce-OF TPC monkey
 GEN:1S=pierce-TR1 NOM monkey
 ‘I speared the monkey.’
- c. *ku=tusuk-ay Da da’um nantu Tanguru’ kana Lutung*
 GEN:1S=pierce-LF OBL:INDEF needle TPC:P:3 head OBL:DEF monkey
 GEN:1S=pierce-TR2 OBL:INDEF needle NOM:P:3 head OBL:DEF monkey
 ‘I pierced the monkey’s head with a needle.’
- d. *ku=tusuk-anay na derederan Da Lutung*
 GEN:1S=pierce-IF TPC spear.type OBL:INDEF monkey
 GEN:1S=pierce-TR3 NOM spear.type OBL:INDEF monkey
 ‘I speared monkeys with the *derederan* (= kind of spear).’

The second set of glosses in (5) reflects an analysis in which the four focus constructions of the Philippinist analysis become two verbal clause types: intransitive and transitive. The actor focus construction is treated as INTRANSITIVE, whilst the ‘object’, locative and instrument/beneficiary focus constructions are treated as constructions with subtypes of TRANSITIVE verb. We return to these subtypes in §3.5.

The ‘topic’ is re-labeled NOMINATIVE or ‘subject’. In Puyuma it is always definite. Using Dixon’s (1979) O and A to refer respectively to the undergoer and actor arguments of a transitive clause and S to refer to the sole argument of an intransitive clause, we can say that O appears in the nominative case and A in the genitive, as illustrated in (5b-d). When a lexically transitive verb like *tusuk* ‘pierce, spear’ occurs intransitively, as in (5a), its S, corresponding to the A, appears in the nominative case. Its patient, corresponding to the O of a TR1 verb, appears in the oblique (‘patient’ here is sometimes theme, but not location, beneficiary, or instrument). If the clause is independent, then the patient must be indefinite. The syntactic status of this oblique patient has been a source of controversy in Philippine and Formosan linguistics, and we return to this in §3.6.

The two constructions are presented schematically in (6) and (7).¹⁷ Parentheses represent optional constituents of a construction. Additional oblique adjunct NPs like *Da da’um* (OBL:INDEF needle) in (5c), are not shown in these schemas.

¹⁷ Abbreviations not used elsewhere are TRVERB transitive verb, ITRVERB intransitive verb.

(6) Transitive construction:

- a. GEN:A= TrVERB =NOM:O (OBL NP:A)
- b. GEN:A= TrVERB (OBL NP:A) (NOM NP:O)

(7) Intransitive construction:

- a. ITRVERB =NOM:S (OBL NP:patient)
- b. ITRVERB (OBL NP: patient) (NOM NP: S)

The order of NPs is not fixed in either construction, and the commonest configurations are in any case that there is no full NP or only one in a clause. In any case, an oblique patient only occurs in (7) if the verb is lexically transitive.

The schemas above allow certain configurations of pronominals and full noun phrases that have not been represented in the examples so far. We return to these in §3.4.

Some of our readers are doubtless asking why we have presented these constructions as differing in transitivity rather than in voice. Why not refer to O voice and A voice? Or to ‘basic’ and ‘antipassive’ voices? The reason is one of simplicity. The intransitive construction in (5a) can certainly be regarded as differing in voice from the transitive construction in (5b-d). Across languages, there are two kinds of voice opposition whose members differ in transitivity. The first kind of opposition is usually labeled ‘accusative’, the second ‘ergative’. An accusative system has active (transitive) and passive (intransitive) voices. An ergative system has transitive¹⁸ and antipassive (intransitive) voices. The difference between them is that in the accusative opposition, the transitive member (the active) has an A subject, the intransitive member (the passive) a subject (S) corresponding to the O of the active, whilst in the ergative opposition, the transitive member has an O subject, the intransitive member (the antipassive) a subject (S) corresponding to the A of the transitive. Puyuma has an ergative opposition, since the transitive construction has an O subject, the intransitive a subject (S) corresponding to the A of the transitive.¹⁹

There is, however, a problem with calling the Puyuma intransitive an antipassive.

¹⁸ There is no widely accepted label for this voice. We take the label ‘transitive’ from Payne (1997:219).

¹⁹ That is, O is marked in the same way as S. This is the textbook criterion for ergativity. A number of Philippine-type languages have been claimed as ergative in this sense: see Blake (1988), Gerds (1988), De Guzman (1988), Kroeger (1993), Mithun (1994), Manning (1996), Starosta (1999), Liao (2004).

If we operate on the premise that constructions are language-specific, then the would-be antipassive in Puyuma is the intransitive construction. That is, an allegedly antipassive clause like (5a) is not different in structure from the clause in (8), which is intransitive but not antipassive.

- (8) *sa-senay i walegan*
 <ITR>PROG-sing NOM:S Walegan
 ‘Walegan is/was singing.’

In this respect, the Puyuma system differs from a prototypical ergative system like that of the Australian language Dyirbal, illustrated in (9). Here, the verbs in the intransitive construction in (9a) and in the transitive construction in (9b) are both unmarked for voice, whilst the verb in the antipassive voice construction in (9c) takes a special antipassive marker *-ŋa-*.²⁰

- (9) Dyirbal:
- a. *ŋuma banaga-ju*
 father:ABS return-NONFUT
 ‘Father returned.’
 - b. *yabu ŋuma-ŋgu bura-n*
 mother:ABS father-ERG see-NONFUT
 ‘Father saw mother.’
 - c. *ŋuma bura-ŋa-ju yabu-gu*
 father:ABS see-ANTIPASS-NONFUT mother-DAT
 ‘Father saw mother.’ (Dixon 1994:10, 13)

In Puyuma, the verb is always marked for transitivity/voice, but the antipassive and the intransitive are marked in the same way, in the examples above with ** (there being no dedicated antipassive marker). It is the verbs in transitive constructions that are marked differently.²¹ Note, incidentally, that changing the labels does not help this problem to go away. If we refer to ‘O voice’ and ‘A voice’, then the ‘A voice’ in Puyuma still is the intransitive construction, and vice versa.

The tabulation in (10) summarizes the alignment differences between Dyirbal and

²⁰ Abbreviations used only in Dyirbal examples are: ABS absolutive, ANTIPASS antipassive, DAT dative, ERG ergative, NONFUT non-future.

²¹ Furthermore, in a prototypical ergative system the A is marked by a dedicated ergative case (Dyirbal *-ŋgu* in (9b), whilst in Puyuma it is marked by the genitive. But this is not central to the issue of how the transitive construction is described.

Puyuma (English being added as an example of an accusative language).

(10)

	Dyirbal	Puyuma	English
Case-marking	S + O	S + O	S + A
Verbal morphology	ITR + OVOICE	ITR + AVOICE	ITR + AVOICE

The ‘Case-marking’ row records the familiar alignments. The ‘Verbal morphology’ row records the facts we have just observed. In Dyirbal the intransitive and O (transitive) voice pattern together (they are unmarked). In Puyuma the intransitive and A (intransitive) voice pattern together (they are marked with **). In English the intransitive and A (active/transitive) voice pattern together (they are unmarked).

Clearly, in an instance like (5), there is a voice opposition between the intransitive in (5a) and the transitive in (5b) (and less directly the transitives in (5c), (5d)). But there is no transitive voice corresponding to an intransitive like the one in (8). We could, of course, impose a distinction between ‘antipassive’ (or ‘A voice’) clauses and other intransitives on Puyuma, but it would be a distinction based on evidence from other languages, not from Puyuma, and would run contrary to the premise that constructions are language-specific and should be described in terms specific to the language.

3.2.1 A note on the term ‘subject’

Some readers may find our use of the terms ‘subject’ and perhaps ‘nominative’ odd in the context of a voice system that is ergatively aligned.

There are essentially two modern usages of the term ‘subject’. In one, it is reserved for the description of accusative languages, and is applied to the single core argument of an intransitive clause and the actor argument of a transitive clause, which are identically marked. This is often the usage among typologists (e.g., Payne 1997:129-133), who are faced with the plethora of different arrangements of grammatical relations both within and across languages (Andrews 1985:104-119).

The other usage prevails among syntacticians who, like Manning (1996:17), treat the subject ‘as the privileged term in a system of grammatical relations’. For Manning the subject in an accusative language is as defined in the previous paragraph, and ‘subject’ in an ergative language is applied to the single core argument of an intransitive clause and the O argument of a transitive clause, which are identically marked. In this sense, the term ‘subject’ fits the Puyuma ‘topic’ splendidly, and we adopt ‘subject’ here in order to avoid the clumsiness of ‘nominative NP’. We also adopt ‘nominative’ rather than ‘absolutive’ because of the former’s association with ‘subject’ and because it has

been commonly used in Formosan linguistics (see the survey in Starosta 1999).²²

The term ‘subject’ was in regular use in Philippine linguistics before McKaughan coined ‘topic’, and McKaughan (1973:208) himself attempted to re-introduce it in a paper in which he sought to retract the use of ‘topic’ and apologized ‘for confusing the issue by calling these subjects the “topic” of the sentence.’ Kroeger (1993) follows McKaughan (1973) and adopts ‘subject’ in his account of Tagalog, but many others writing about Philippine languages have stuck with ‘topic’. Formosanists have tended to talk about the ‘subject’ of the ‘in focus noun phrase’.

We would like to emphasize that our choices of ‘subject’ and ‘nominative’ are not intended to assert that the Puyuma grammatical relation they denote has any universal status. They are chosen because they are familiar among linguists and we think it is fairly easy to remember how we are applying them to Puyuma.

3.3 Transitivity and discourse

Not surprisingly, on the basis of text frequency, independent clauses with an O subject, i.e. transitive clauses, are basic in discourse. The subject of such a clause must be definite and, in general, if there is a definite patient, it must be the subject.²³ Since definiteness is often an indicator that the definite participant has already been introduced into the discourse, this means that transitive subjects in Puyuma tend to track O topicality, whereas in English they tend to track A topicality and continuity. Hence *DikeTan* ‘sticky rice’ is introduced as an oblique in (11a). In (11b) it is the topic (in the sense of Lambrecht 1994, i.e. it is topicalized) and is the subject of the verb *ta=iLang-aw* ‘we grind it’. It is then also subject of *ta=Lubuk-aw* ‘we pack it’ in (11c) and of *ta=Te’el-aw* in (11d). It is the subject because it is the definite O, where English subjects would track the A ‘we’, as the gloss shows.

- (11) a. *na binariyaw i a sinanga Da DikeTan*
 NOM:DEF sticky.rice.cake TPC NOM:INDEF product OBL:INDEF sticky.rice
 ‘*Binariyaw*, it is a kind of sticky rice product.
 b. *aw na DikeTan i, ta=iLang-aw Da enay,*
 and NOM:DEF sticky.rice TPC GEN:1IP=grind-TR1 OBL:INDEF water
 And the sticky rice, we grind it with water,

²² The term ‘nominative’ is derived from Latin *nomen* ‘name’, and denotes the form of a noun that is used to ‘name’ its referent, i.e. the citation form of the noun. Our usage of ‘nominative’ is consistent with this.

²³ There are exceptions, but they occur only in elicited material.

- c. *aw ta=Lubuk-aw Da Lubuk,*
 and GEN:1IP=sack-TR1 OBL:INDEF sack
 and we pack it in a sack,
- d. *aw ta=Ta-Te'el-aw Da barasa i,*
 and GEN:1IP=Ca-press-TR1 OBL:INDEF stone TPC
 and when we press it with a stone,
- e. *mu-teres tu=enay*
 ANTICAUS-filter NOM:P:3P=water
 its water is filtered out.'

3.4 Case-marking

As we have seen, Puyuma makes a three-way case distinction between NOMINATIVE, marking the subject, GENITIVE, marking a non-subject A, and OBLIQUE, marking among other things a non-subject patient.

The label 'oblique' is appropriate because the oblique case is often used for NPs that are clearly not a core argument of the verb, as shown in (12a). Here the core arguments are the (unmentioned) nominative O monkeys and the genitive A *ta* = 'we', whilst the instrument *kana derederan* '(with) the spears' is oblique. In (12b) the core arguments are the nominative O *na barasa* 'the stone' and genitive A *tu* = 'he', leaving the location/goal *kana kaLi* '(into) the river' as an oblique. In the relative clause in (12c) there are two obliques, neither of which is the A, and at least one of which cannot be a core argument.

- (12) a. *ta=tusuk-aw kana derederan*
 GEN:1IP=pierce-TR1 OBL:DEF spear
 'We pierced them [the monkeys] with the spears.'
- b. *tu=ba-bulu-an na barasa kana kaLi*
 GEN:3=REDUP-throw-TR3 NOM:DEF stone OBL:DEF river
 'He'll throw the stone into the river.'
- c. *iDi na aDi kiberay kan Tayban Da bini*
 this.NOM NOM:DEF NEG GET OBL:S Tayban OBL:INDEF seed
 'This person, who didn't get seeds from Tayban.'

In intransitive clauses like (12c) and (13a), the indefinite oblique noun phrases *Da bini* 'seeds' and *Da paisu* 'money' can be identified on semantic grounds as the patients of their respective clauses. However, nothing distinguishes them morphosyntactically from the instrument adjunct *Da puaTemeL* '(with) medicine' in (13b) or the comitative

adjunct *Da=tu sa-ra'ip-an* '(with) colleagues of hers' in (13c).²⁴

- (13) a. *Takaw=ku Da paisu*
 <ITR>steal=NOM:1S OBL:INDEF money
 'I stole money.'
- b. *aDi=ku karuwa m-inaTay Da puaTemeL*
 NEG=NOM:1S can ITR-die OBL:INDEF medicine
 'I can't die with medicine.' (i.e., 'I can't kill myself with medicine.')
- c. *an m-u-ra'i-ra'ip la Da=tu sara'ipan i*
 when ITR-go-REDUP-workplace PERF OBL:INDEF=GEN:3 colleague TPC
 'When she went to the workplace with colleagues of hers ...'

Another fact which distinguishes obliques from nominative and genitive arguments is that only a nominative (in (14a), (14c)) or a genitive (in (14b)) may control the omitted argument of a following verb in a serial verb construction (a purposive construction in (14a) and (14b), a manipulative construction in (14c)).

- (14) a. *m-u-ruma'=ku i ruma' [kirunguT kan nanali]*
 ITR-go-home=NOM:1S LOC home <ITR>take.care OBL my.mother
 'I went home to take care of my mother.'
- b. *tu=Lugas-aw me-na'u*
 GEN:3=lift-TR1 ITR-see
 'He lifted it to see.'
- c. *tu=bau-bau-aw=ku m-uka i takesian*
 GEN:3=REDUP-push-TR1=NOM:1S ITR-go LOC school
 'She kept pushing me to go to school.'

In Puyuma an oblique argument cannot be the controller of an omitted argument. However, as Chang (2003) points out, certain Formosan languages do allow an A-voice manipulative verb in a sentence analogous to (14c), and this verb takes a seemingly oblique-marked O ('me') which controls the omitted argument of the following verb. There is then a case for labeling the oblique-marked O as accusative rather than oblique and for generalizing this label to all oblique-marked Os (if Puyuma were such a language, *Da paisu* 'money' in (13a) would be an accusative). This raises the question, how does one distinguish in such a language between an accusative and an oblique? But this question does not apply to Puyuma.

²⁴ Further such adjuncts are the instruments *Da enay* '(with) water' in (11b) and *Da barasa* '(with) a stone' in (11d) and the locative *Da Lubuk* '(in) a sack' in (11c).

The nominative/genitive distinction is manifest in the pronoun paradigms in Table 1.²⁵

Table 1: Puyuma bound and oblique pronouns

	1S	2S	3S	1IP	1EP	2P	3P
Nominative	= <i>ku</i>	= <i>yu</i>	Ø	= <i>ta</i>	= <i>mi</i>	= <i>mu</i>	Ø
Genitive	<i>ku</i> =	<i>nu</i> =	<i>tu</i> =	<i>ta</i> =	<i>mi</i> =	<i>mu</i> =	<i>tu</i> =
Oblique	<i>kanku</i>	<i>kanu</i>	<i>kantaw</i>	<i>kanta</i>	<i>kaniam</i>	<i>kanemu</i>	<i>kantaw</i>

Nominative =*ku* is illustrated in the intransitive clauses in (13a) and (14a) above and in a transitive clause in (14c).

The genitive proclitics have been liberally illustrated in earlier examples. Neither the nominative enclitics nor the genitive proclitics are omissible when they mark an argument of the verb. However, there is no third person nominative form.²⁶ This has the consequence that a nominative A NP is not cross-referenced on the verb. Genitive A cross-referencing is mentioned below.

The NP markers, shown in Table 2, collapse the distinction between genitive and oblique.²⁷ Puyuma NP markers are similar to those of other Formosan and Philippine languages in making a distinction between common and personal nouns, but somewhat unusual in distinguishing between definite and indefinite common nouns. Puyuma shares with Tsou the feature of distinguishing only between nominative and oblique, and with Amis, Thao, and Saisiyat the feature of having an invariable marker for location nouns. There are also several sets of possessor pronouns which additionally mark the case of their NP. These are described in Appendix B.

Table 2: Puyuma NP markers

Noun type:	Common		Personal		Location
	Definite	Indefinite	Singular	Plural	
Nominative	<i>na, ina</i>	<i>a</i>	<i>i</i>	<i>na</i>	<i>i</i>
Oblique	<i>kana</i>	<i>Da</i>	<i>kan</i>	<i>kana</i>	

²⁵ A nominative pronoun is encliticized to the first element of the clause; a genitive pronoun is procliticized to the verb.

²⁶ If (8) lacked its subject NP *i Walegan*, and read simply *sa-senay* ‘s/he is/was singing’, it would still be a full clause.

²⁷ One could call the oblique NP markers ‘genitive-oblique’, but we follow Teng in calling the ‘oblique’ because (i) the combined term is clumsy and (ii) three out of four incorporate *ka-* (see Table 2), which also occurs in oblique pronouns (see Table 1).

The absence of dedicated genitive NP markers leads to the situation illustrated in (15). The A NP *kan nanali* (OBL:S my.mother) is necessarily marked as oblique, but it is cross-referenced on the verb by the genitive proclitic *tu=* GEN:3.²⁸

- (15) *aDi la tu-bu 'ut-i=ku kan nanali*
 NEG PERF GEN:3=stop-TR2=NOM:1S OBL:S my.mother
 'My mother didn't stop me.'

Verbal cross-referencing is rare in Philippine-type languages, but its occurrence in Puyuma has a straightforward diachronic explanation. In some ancestor of Puyuma, as in other Formosan languages, the three-way case distinction between nominative, genitive and oblique pronominals also applied to NP markers. At some stage, the functions of genitive NP markers were taken over by the oblique markers and the genitives were lost. This led to a potential ambiguity, whereby the (core) A NP and oblique NPs with various functions shared the same case marking. The genitive proclitic, which in any case occurred when there was no full A NP, came to be retained with a full A NP in order to disambiguate the A NP from other obliques.

This means that oblique NPs have two morphosyntactically distinct functions in Puyuma: (i) as the patient of an actor subject intransitive verb or as an adjunct (there is no distinction between these, as noted above); and (ii) as the A of a transitive verb, when it is cross-referenced by the third person genitive proclitic *tu=* on the verb.

3.5 Verbal morphology

Table 3 shows the basic verbal morphology of Puyuma. The labels in the left-hand column are fairly obvious, except for the Projective, which has a form separate from the Realis only in the intransitive, and is used to express a hortative 'let me'/'let us' or desiderative 'I/we want to', or a purposive after a motion verb.²⁹

²⁸ The same thing happens with possessive NPs. The oblique possessor NP is cross-referenced by a genitive pronoun: *nantu ruma' kan Pasiyar* /NOM:her house OBL:S Pasiyar/ 'Pasiyar's house'.

²⁹ The Realis is used in hortative clauses in transitive voices; with a stative the Progressive also expresses a temporary state; and the Irrealis is also used in prohibitions.

Table 3: Summary of Puyuma verbal morphology

	INTRANSITIVE	TRANSITIVE		
		1	2	3
Imperative	ROOT	ROOT- <i>u</i>	ROOT- <i>i</i>	ROOT- <i>an</i>
Realis	<i>M</i> -ROOT	ROOT- <i>aw</i>	ROOT- <i>ay</i>	ROOT- <i>anay</i>
Realis progressive	<i>M-Ca</i> -ROOT	<i>Ca</i> -ROOT- <i>aw</i>	<i>Ca</i> -ROOT- <i>ay</i>	<i>Ca</i> -ROOT- <i>anay</i>
Realis gerund	(as Realis)		< <i>in</i> >ROOT(- <i>an</i>)	
Negative	(as non-neg.)		ROOT- <i>i</i>	ROOT- <i>an</i>
Irrealis	<i>Ca</i> -ROOT		<i>Ca</i> -ROOT- <i>i</i>	<i>Ca</i> -ROOT- <i>an</i>
Irrealis gerund	(as Irrealis)		<i>Ca</i> -ROOT- <i>an</i>	
Projective	<i>M</i> -ROOT- <i>a</i>	(TRANSITIVE same as Realis)		
Negative hortative	<i>M-Ca</i> -ROOT	(no TRANSITIVE forms)		

There are, as noted in §3.2, three transitive forms labeled TR1, TR2, and TR3, corresponding respectively to the Philippinists' 'object' (patient/goal), location, and instrument/beneficiary focus. The effect of these derived forms is to allow referents in different semantic roles to become the O. This is the function of applicative morphemes in accusative languages, with the difference that the O in an accusative language is the direct object, but in an ergative language it is the subject. Few scholars seem to have considered the possibility that an ergative voice system might also have applicatives. A notable exception was Starosta (1986), who used the term 'recentralization' for the situation in Philippine-type languages whereby derived alternant forms of the verb like those with TR1, TR2, and TR3 in Puyuma allow different referents with various semantic roles to become O subject. Payne (1994:323-324, 1997:54) mentions in passing that such derivations might be regarded as applicatives. Croft (2003:227) reserves the term 'applicative' for accusative languages, but recognizes that the derivational process in Philippine-type languages is functionally similar to applicativization.

We are happy to call the suffixes which form transitive verbs in Puyuma 'applicatives'.³⁰ They do differ in other respects from typical applicatives, however. First, it is common in an applicative system for there to be an unmarked transitive, usually with a patient as direct object; the addition of applicative suffixes allows other roles, e.g. location, instrument, to be 'promoted' to object. This seems to be the norm in Bantu and Indonesian-type Austronesian languages. In Puyuma and in Philippine-type

³⁰ Ross (2002:21) considered labeling O voice (= transitive) alternations as applicatives but decided in favor of the conventional terminology of 'patient voice', 'location voice' and 'circumstantial voice' precisely because they are conventional in Formosan and Philippine studies. We make the opposite decision here, as our goal is to use terms that emphasize functional similarities across languages.

languages in general, on the other hand, there is an applicative affix on every transitive form, and this is also true of many Oceanic Austronesian languages, e.g. Standard Fijian.³¹

We could have called the Philippinists' patient, location, and instrument/beneficiary focus quite simply 'patient applicative', 'location applicative' and 'instrument/beneficiary applicative'. Instead, however, we have here labeled them TR1, TR2, and TR3. The reason for this is simple (and applies right across Philippine-type languages): the terms 'patient', 'location', and 'instrument/beneficiary' refer only to the semantic roles which are considered prototypical of the noun phrases that occupy the subject slots. In reality, there is a great deal of variation. With human referents, a 'locative' subject is often 'dative'. In many cases, an 'instrumental' referent is the 'theme', i.e. an object which moves as a result of the event. This is true of (16a) (see also (12b)). Examples (16b-c) illustrate other uses of TR3, all with some concept of movement.

- (16) a. *ku=tuLud-anay na sarekuDan kana temumuwan*
 GEN:1S-pass-TR3 NOM:DEF stick OBL:DEF offspring
 'I passed the stick to the offspring.'
- b. *ku=remere-anay na kinsas*
 GEN:1S=press.down-TR3 NOM:DEF police
 'I fought the policeman down (to the ground).'
- c. *tu=Tukul-anay tu=Dakur*
 GEN:3P=carry.on.back-TR3 NOM:P:3P=back
 'He hunched his back.'

Our assignment of TR1, TR2, and TR3 is not arbitrary, however. There is a general tendency for the subjects of verbs in TR1 to be permanently or radically affected by the action, e.g. to be bent or broken. The subjects of verbs in TR2 are less affected and less likely to undergo permanent change: They may be swept or cleaned, or they may receive something. The subjects of verbs in TR3 are the least affected: they may be used or moved or may be beneficiaries.

We have chosen to use the term 'gerund' in Table 3 rather than 'nominalization' because these morphological constructions are not simply lexical nominalizations but forms regularly used in relative clauses (if this is the appropriate term: see §3.7.1) where the participant in the main clause event is not the A of the relative clause. Thus in (17) the gerund *in-abak-an* can be translated into English as something like 'what [the child] was packed into'.

³¹ There is a substantial literature on applicatives in Oceanic languages, although they are often not labeled applicatives but 'direct' and 'indirect' transitives: cf. Pawley 1973 and Evans 2003.

- (17) *tu=laseD-aw dar i Talu-Talun [na in-abak-an*
 GEN:3=hide-TR1 often LOC REDUP-grass NOM:DEF GER:REAL-pack-GER
kana walak] na paDekan
 OBL:DEF child NOM:DEF backpack
 ‘She often hid it in the field, the backpack in which she had packed the child.’

The suffix *-an* may be dropped from the realis gerund (but not from its irrealis counterpart).

Gerunds are also used in certain complement clauses, but these lie beyond the scope of this paper.

Table 3 is only a summary of Puyuma verbal morphology. Transitives are formed regularly, but which of the three forms will occur with a given root is unpredictable. Intransitives, however, vary considerably. There are morphological classes of verbs, each with its own intransitive paradigm. These classes are summarized in Appendix C.

3.6 Voice and transitivity: unravelling a controversy

There has been some debate in the literature as to whether Philippine-type ‘focuses’ are ‘really’ voices or not (see De Guzman 1997). The argument centres around whether the ‘focuses’ are derivationally or inflectionally related to each other, since, it is claimed, voice must be inflectional. It has been argued that Philippine-type ‘focuses’ are derivational and therefore not voices. They are said to be derivational, because (as we noted earlier) one cannot predict for a given verb-root which ‘focuses’ it will have. Other arguments have also been used. We do not intend to engage with these here, because we take the view that if ‘voice’ is to have any cross-linguistic meaning, it must be defined in semantic and functional terms. Crudely, we can say that voice consists of alternate constructions which place either the A or the O in the subject position.³² From this viewpoint, Philippine-type ‘focuses’ are clearly voices: the precise mechanisms which voice constructions employ are irrelevant.

This means that the intransitive and transitive sentences in (18a) and (18b), repeated from (5), are in a voice relationship with each other.

³² For an account of voice across languages and some theoretical underpinnings, see Croft (2001: Ch. 8)

- (18) a. *tusuk=ku Da Lutung*
 <ITR>pierce-NOM:1S OBL:INDEF monkey
 ‘I speared a monkey.’
 b. *ku=tusuk-aw na Lutung*
 GEN:1S=pierce-TR1 NOM:DEF monkey
 ‘I speared the monkey.’

We have used the term ‘lexically transitive verb’ in this paper for verbs like *tusuk*, a verb with a paradigm of minimally two members: a transitive and an intransitive voice in a predictable semantic relationship (the term was apparently coined in this sense by Payne 1994:324, writing about the Philippine language Cebuano, and is also used by Huang 2002).

We have argued in §3.4 that there is no morphosyntactic difference between the oblique patient *Da Lutung* in (18a) and any other adjunct in an intransitive clause. The question we want to ask here is whether an intransitive-voice clause (i.e. an intransitive clause whose verb is the intransitive voice of a lexically transitive verb) is interpreted any differently from an intransitive clause whose verb is not lexically transitive. In other words, does the paradigmatic relationship of voice make any difference to the interpretation of the intransitive clause? There is some evidence that it does.

We have noted that the oblique patient of an intransitive-voice independent clause is indefinite. There are a few instances in the corpus, however, where what looks suspiciously like the oblique patient of an intransitive verb is definite (remember from Table 2 that Puyuma NP markers distinguish between definite and indefinite). What is interesting is that none of these verbs has a corresponding TR1 form with an O/patient subject. That is, there is no way to express the putative definite patient except as an oblique in the intransitive clause. In (19a) *kanDini* (OBL.this) is definite, and appears to be the patient of *ma-rengay-ta* ‘we talk about’. But it is not. There is indeed a corresponding TR1 verb *rengay-aw*, but its O/patient subject is the person talked to, not the thing talked about. In (19b) the situation is even simpler. The putative patient *kan temutaw* (OBL:S his.grandma) is not the O/patient of *laman* ‘have pity on’, because this verb, like several that express emotional attitudes, has no transitive counterpart.

- (19) a. *an ma-rengay=ta kanDini i, a barasa aDi=ta*
 when ITR-tell=NOM:1IP OBL.this TPC NOM:INDEF stone NEG=NOM:1IP
ka-ka
 HORT-<ITR>say
 ‘When we are talking about this, we won’t call it a stone.’

- b. *iDu na walak la i, laman la kan temutaw ...*
 NOM.that NOM:DEF child PERF TPC pity PERF OBL:S his.grandma
 ‘That child, he had pity on his grandmother...’

We tentatively conclude from these and other examples that the indefiniteness of the oblique patient of an intransitive clause is associated with the fact that the verb has a paradigmatically related transitive counterpart. Although the patient of an intransitive clause like (18a) is encoded in the same way as an adjunct like *Da enay* in (11b), the fact that a verb like *tusuk* ‘pierce’ in (18a) has a transitive counterpart *tusuk-aw* in (18b) ...

- forces the use of a transitive clause like (18b) if the patient is definite, and thereby blocks the occurrence of an intransitive clause with an oblique definite patient resembling the clauses in (19); and ...
- causes a clause like (18a) to be interpreted as an intransitive construction with a patient, not as an intransitive construction with an adjunct.

If these statements are true, then we need to recognize an ‘extended intransitive’ construction (EIC) with an oblique patient (cf. 1994:120-124). This is a subtype of the intransitive construction, and was anticipated in the schematic presentation in (7).

The idea that a single surface form, the intransitive construction plus an oblique-marked noun phrase, may encode different constructional meanings may seem odd, but it is probably true of all languages that there are constructions which express a particular meaning when they occur with a certain class of verbs. Sometimes these classes are large, sometimes they are smaller, and occasionally they have only a single member, as in the English double-object constructions in *Joe forgave Sally her sins* or *Joe envied Bob his wealth*, where the construction has a verb-specific meaning (Croft 2003) different from the prototypical double-object construction in *Joe gave Bob his wealth*.

There has been vigorous debate in the literature, some of it summarized by Ross (2002:24-31), as to whether EICs in Formosan and Philippine languages are intransitive or transitive. Kroeger (1993) argues that in Tagalog they are transitive; Starosta (1999) argues that in Atayal they are intransitive, and Reid & Liao (2004:441-442), discussing the typology of Philippine languages in general (including Tagalog), treat them as intransitive. We limit our discussion here to Puyuma, as it is likely that different languages require different analyses.

To understand the problem, we need to remind ourselves that a construction is a pairing of form and meaning. The Puyuma EIC looks identical to an ordinary intransitive construction with an oblique-marked adjunct. This is a statement about its form.

However, because the (intransitive) verb of an EIC has a transitive counterpart, the meaning of the EIC differs from that of an intransitive construction with an oblique-marked adjunct, and the oblique-marked argument of the EIC is interpreted as a patient. Whereas an oblique-marked adjunct like *Da enay* in (11b) can be omitted without a fundamental change in argument structure, the omission of the indefinite patient from an EIC clause will cause the hearer to search the discourse context for the appropriate referent. Thus the omission of *Da Lutung* from (18a) would have the hearer searching for the referent that the speaker speared, i.e. for the patient of the verb. Because the oblique-marked argument of the EIC is its patient, the EIC has been labeled ‘semantically transitive’ or ‘pseudo-transitive’ by some commentators. Hence when we look at the form (syntax) of the EIC, it is intransitive, but when we look at its meaning, it seems transitive-like.

This finding suggests that the status of TR1 differs from that of TR2 and TR3, since it is only the O subject, i.e. the patient, of the TR1 verb that appears as the oblique argument of an intransitive verb. Thus TR1 is semantically the ‘unmarked’ or default transitive—a fact that is hardly surprising in the light of the observation in §3.5 that in many languages it is the patient transitive that is least marked morphologically, whilst a locative transitive or an instrument transitive takes an applicative affix.

The interpretation of an oblique patient as indefinite occurs in independent clauses. Where a dependent verb is forced by its construction into the intransitive form, the patient may be definite. This happens in (14a), where the second verb of a serial verb construction is obligatorily intransitive and *kan nanali* ‘my mother’ is (obviously) definite, and in (20), where the A of the relative clause must be subject because it is coreferential with the head noun, forcing *kanaDi kana tuLu-a-ami* ‘these third-years’ into the oblique case.

- (20) *amau taita [na pa-la-laDam kanaDi kana tuLu-a-ami]*
 COP D:1IP NOM:DEF CAUS-REDUP-know OBL.these OBL:DEF three-LINKER-year
 ‘It is we who teach these third-years.’

Having found that there is an extended intransitive construction in Puyuma, it is not surprising to find that there is a corresponding extended transitive construction (ETC), i.e. a transitive construction like the one shown schematically in (6) but with the addition of an oblique-marked indefinite patient. Similar constructions are recognized by Reid & Liao (2004:445-446) in Philippine languages. Indeed, we can now see that (5d), repeated here as (21), exemplifies such a construction and that *Da Lutung* is the oblique patient.

- (21) *ku=tusuk-anay na derederan Da Lutung*
 GEN:1S=pierce-TR3 NOM:DEF spear.type OBL:INDEF monkey
 ‘I speared monkeys with the *derederan* (= kind of spear).’

Since the ETC has an oblique patient, and therefore never a patient subject, its verb will always be TR2 or TR3.

We also see the ETC in (22b), where the referent which is the O subject with TR1 in (22a) survives as an indefinite oblique when the beneficiary becomes O subject with TR3 in (22b).

- (22) a. *ku=Takaw-aw na paisu*
 GEN:1S=steal-TR1 NOM:DEF money
 ‘I stole the money.’
 b. *ku=Takaw-anay i nanali Da paisu*
 GEN:1S=steal-TR3 NOM:S my.mother OBL:INDEF money
 ‘I stole money for my mother.’

Again the indefinite patient *Da paisu* ‘money’ in (22b) cannot be replaced by the corresponding definite noun phrase: if a patient is definite, it must become the subject of the clause.

Across languages, verbs of giving are the strongest candidates for occurrence in constructions with three arguments, and it is not surprising that the transitive form of the Puyuma verb ‘give’ occurs in the ETC, treating the theme *Da ruma* ‘a house’ as a patient:

- (23) *tu-beray-ay Da ruma’ kan walegan i pilay*
 GEN:3S=give-TR2 OBL:INDEF house OBL:S Walegan NOM:S Pilay
 ‘Walegan gave Pilay a house.’

As with the EIC, the ETC is identical in form to a simple transitive clause with an oblique-marked adjunct like the instrument adjunct in (24a) and the duration adjunct in (24b).

- (24) a. *na binariyaw i, ta-LibuT-anay Da Labilu*
 NOM:DEF binariyaw TPC GEN:1IP-wrap-TR3 OBL:INDEF *Labilu*
Da mangede’
 OBL:INDEF tender
 ‘The *binariyaw* (= sticky rice), we wrap it with tender *Labilu* (= kind of leaf).’

- b. *na paTaka i, ta-iyam-ay Diya Da saya*
 NOM:DEF meat TPC GEN:1IP-salt-TR2 IMPF OBL:INDEF one
Da wari-an
 OBL:INDEF day-NSR
 ‘The meat, we salt it for one day...’

And again, it is the verb and its argument structure that determines the interpretation of the construction. With the ETC, it is the relationship between one transitive and another as in (22) that determines that the oblique-marked noun phrase in (22b) is interpreted as a patient.

3.7 The Subject Construction Hierarchy

We have seen that Puyuma case-marking and voice are ergatively organized. However, as many writers have pointed out, in a given language some constructions may be ergatively organized, others accusatively. Croft (2001:155) proposes a ‘Subject Construction Hierarchy’, based on Kazenin’s (1994) work and his own:

- (25) *The Subject Construction Hierarchy:*
 coordination < purpose < relativization < verb agreement < case marking

The hierarchy ‘defines an implicational scale such that for any construction on the scale, if the construction patterns ergatively, then all the constructions to the right of it on the scale also pattern ergatively; if the construction patterns accusatively, then all the constructions to the left of it on the scale also pattern accusatively’ (Croft 2001:155), or, as Kazenin observes, pattern neither way. English patterns accusatively for all features on the hierarchy; Dyirbal, on the other hand, patterns ergatively for all of them. As Croft (2001:156) shows, many languages lie somewhere between these two extremes.³³

We have demonstrated only that Puyuma patterns ergatively for the rightmost feature in the Subject Construction Hierarchy, namely case marking. If we interpret Puyuma clitic pronouns as ‘verb agreement’, then Puyuma also patterns ergatively for this category.

³³ The hierarchy is, as Croft says, incomplete. There are other functionally definable constructions that occur in a variety of languages and which are not (yet) included in the hierarchy. They include quantifier float, secondary predicate agreement, and raising, all of which Kroeger (1993:22-30) shows to be ergatively aligned in Tagalog. (In fact, from Kroeger’s description, Tagalog seems to pattern ergatively for all features on the hierarchy, although this is not his analysis, as he regards both O and A voice clauses in Tagalog as transitive.) Where these extra features fit into the hierarchy is not entirely clear.

3.7.1 Stative verbs, modification and relative clauses

The next category on the Subject Construction Hierarchy is relativization, which raises interesting issues.

There is no morphosyntactic division between dynamic intransitive and stative intransitive verbs in Puyuma, and there is no separate adjective class. All the sentences in (26) show the same intransitive construction. However, they are ranged semantically along a cline from (26a), which is clearly dynamic, to (26c) and (26d), which are clearly stative. There is no constructional difference between dynamic and stative verbs.

- (26) a. *sa-senay i walegan*
 <ITR>PROG-sing NOM:S Walegan
 ‘Walegan is/was singing.’
 b. *ma-tengaDaw i nanali*
 ITR-sit NOM:S my.mother
 ‘My mother sat down/is/was seated.’
 c. *ma-kiteng i pilay*
 ITR-small NOM:S Pilay
 ‘Pilay is/was small.’
 d. *bulay ina ruma’*
 ITR.small NOM:DEF house
 ‘The house is/was clean.’

There is also no morphological difference between them, as there are also dynamic verbs in *ma-* and in zero (see Appendix C). Many stative verbs belong to Class 3b or 3c; they have progressive and irrealis forms just as dynamic verbs do. For example, the dynamic verb *ma-Tangis* ‘weep’ has the realis progressive *ma-Ta-Tangis* and the irrealis *ka-Ta-Tangis*. Similarly, the stative verb *ma-kiteng* ‘be small’ has the realis progressive *ma-ka-kiteng*³⁴ and the irrealis *ka-ka-kiteng*.

There is thus no constructional difference between the predication of an action (by a dynamic verb) and the predication of a property (by a stative verb). The latter covers the functional domain of a predicate adjective in many other languages.

Similarly, there is also no constructional difference between **modification** by an intransitive action and modification by a property. Examples (27a) and (27b) both contain dynamic verbs, whilst (27c) and (27d) are stative. All four examples manifest the relative construction, whereby the relative clause precedes or follows (more commonly

³⁴ Meaning ‘becoming small’.

the latter) the modified noun and agrees with the noun in case and definiteness. This often results in the relative clause being marked by the same case marker as the modified noun. (The glosses of agreeing case markers are in large capitals.)

- (27) a. *iDi [na aDi kiberay kan tayban Da bini]*
 this.NOM NOM:DEF NEG get OBL.S Tayban OBL:INDEF seed
 ‘This (person), [who didn’t get seeds from Tayban]...’
- b. *kiumal=ta [Da mi-alup] [Da mi-a-kelep kaDini]*
 ITR.ask=NOM:1IP OBL:INDEF ITR-hunt OBL:INDEF ITR-PROG-nest here
 ‘We asked (some spirits) [who hunt] and [who live here].’
- c. *amau a Denan [a ma-kiteng] ka*
 COP NOM:INDEF mountain NOM:INDEF ITR-small <ITR>say
 ‘It was said that it was a [small] mountain.’
- d. *na suan [na ma-Tina] [na uTeuTem]*
 NOM:DEF dog NOM:DEF ITR-big NOM:DEF ITR:black
 ‘the [big] [black] dog’

The construction in (27c) and (27d) is the default construction for modification by a property. That is, the same construction is used for modification by an action—the typical domain of the relative clause—and modification by a property. There is, in other words, no morphosyntactic distinction between a relative clause and any other modification within a noun phrase, and one might well refer to both of them as the **modification construction**.³⁵

The relative clauses in (27) are all instances in which the modified noun is the S of the relative clause. The relative clauses are essentially identical to their main-clause counterparts.

When the modified noun is the O of the relative clause, a different construction is used. This is the gerund construction, as in (17), and the A may take the form of a possessor pronoun (possessor, not genitive, as the possessor-of-nominative pronoun in (28a) shows: cf. Appendix B). Again, the relative clause agrees in case and definiteness with the modified noun (cf. (17)), but the case marker of the relative clause may be omitted if a possessor pronoun precedes the verb, as a possessor pronoun also indicates

³⁵ A small objection to this conclusion might be raised, but we think it is invalid. Puyuma has a comparative prefix *mara-* which may be attached to a stative verb: *mara-arii* ‘faster’, *mara-inaba* ‘better’, *mara-ma-Tina* ‘bigger’, *mara-ma-kiteng*. It is prefixed to the A voice marker if there is one. The presence of *mara-* does not suggest a separate word class, however, as the presence of a comparative strategy in any language is conditioned by the semantic feature of gradability. Properties are generally gradable, and so may be subject to a comparative strategy.

the case of the noun phrase (see (28b) and (28c)):

- (28) a. *ala amuna saDu [nantu T<in>ekeL] na asi*
 maybe because ITR.many NOM:DEF:3S <GER:REAL>drink NOM:DEF milk
 ‘Maybe because the milk [he drank] was a lot.’
 b. *tu=uisaT-ay [ta=in-uisaT-an] na taleb*
 GEN:3=get.onto-TR2 NOM:P:3=GER:REAL-get.onto-NSR NOM:DEF raft
 ‘They got onto the raft [that they had got onto (before)].’
 c. *ulaya [ku=ra-rengay-an] a saya*
 exist NOM:-:1S=IRR-tell-NSR NOM:INDEF one
 ‘There’s one thing [I’m going to say].’

The question we asked in relation to the Subject Construction Hierarchy was whether relative clauses pattern ergatively or accusatively. The functionally significant feature of subjects in coordinate, purposive and relative constructions is that arguments are tracked across clauses: ‘The three behavioral constructions in the Subject Construction Hierarchy all involve a null expression of an argument in one clause that is coreferential with an argument in the other clause...’ (Croft 2001:159). If, in a clause-type with null argument expression, a clause with a null O has the same construction as a clause with a null S, then that clause-type patterns ergatively. If, on the other hand, it is a clause with a null A that has the same construction as a clause with a null S, then that clause-type patterns accusatively (Kazenin 1994). By this criterion, Puyuma relative clauses pattern accusatively, as one construction (the intransitive) is used when S or A is relativized, and a different construction (the gerund) is used when an O is relativized.

There is one more question to be addressed in relation to headless relative clauses like those found in (27b) and (29).

- (29) a. *arii [na mi-Tepa sanga’]*
 ITR.fast NOM:DEF ITR-aim.at <ITR>make
 ‘[Those who have professional skill in doing this] are fast.’
 b. *[na kinuwayan m-ekan Da TaLun] i, amau tu=walak*
 NOM:DEF before ITR-eat OBL:INDEF grass TPC COP NOM:3S=child
 ‘[The one who ate grass first] is its child.’

Because Puyuma relative clauses, at least in the intransitive, are preceded by the same case markers as other noun phrases, one could argue that a headless relative clause is not a relative clause but simply a noun phrase. One could further argue that relative clauses in general are simply noun phrases in apposition to ‘their head noun’. This position

receives some support from the fact that the position of the relative clause is not fixed. It is especially mobile when it is short, such that the three orders in (30) are all acceptable.

- (30) a. *na suan na ma-Tina na uTeuTem*
 NOM:DEF dog NOM:DEF ITR-big NOM:DEF ITR.black
 ‘the big black dog’
 b. *na ma-Tina na uTeuTem na suan*
 c. *na ma-Tina na suan na uTeuTem*

We do not need to argue the case one way or the other, as the categories ‘noun phrase’ and ‘relative clause’ are language-specific. The important thing is that the same construction is used for property modification and action modification (i.e. relativization), and that this construction itself has the same structure as a noun phrase.

3.7.2 Purposive serial verb constructions

There is no dedicated purpose construction in Puyuma. As (14a), (14b) illustrate, purpose is expressed by the serial verb construction, which also has a variety of other functions we shall not discuss here. The verbs in a serial verb construction share an argument. That is, an argument of the serialized (second) verb is coreferential with (‘controlled’ by) an argument of the first and has null expression, as in (14a-c). Any serialized verb must be intransitive, and so its null-expressed argument is always its subject (S). The O/patient of a serialized verb, like *kan nanali* ‘my mother’ in (14a), is expressed as an oblique.

Insofar as the null-expressed argument of a serialized verb is always S, and this often corresponds semantically with the A of a transitive verb, serialized verbs (including the purposive) are accusatively aligned.

3.7.3 Coordinate null expression

‘Coordinate null expression’ refers to coordinate clause constructions in which an argument of a clause has null expression when it is coreferential with an argument of a preceding clause. In an English sentence like {*John hit David and then ran away.*}, the null-expressed subject of the second clause is coreferential with the subject/A of the first: It is John who ran away. Coordination patterns accusatively. If the subject of the second clause is expressed by a pronoun, however, as in {*John hit David and then he ran away.*}, then the coreferentiality of the pronoun is potentially ambiguous: It may be

either John or David who ran away.

In a language like Dyirbal, where coordination patterns ergatively, the O ‘David’ would be the (absolutive) subject of the first clause, and the null-expressed subject of the second clause would be coreferential with it; i.e. it would be David who ran away.

Coordination in Puyuma does not pattern like either English or Dyirbal, as there is, strictly speaking, no coordinate null expression. Instead, it patterns like English coordination with the pronoun in the second clause, i.e. coreferentiality is potentially ambiguous. Thus in (31) it may be either Walegan or Pilay who ran away.

- (31) *tu=pukpuk-aw i pilay kan walegan aw p<en>uwar*
GEN:3P=beat-TR1 NOM Pilay OBL Walegan and <ITR>run.away
‘Walegan beat Pilay and then Walegan/Pilay ran away.’

Recall that Puyuma bound pronouns, whether nominative or genitive, cannot be omitted (§3.4). Recall also that there is no third person nominative pronoun (Table 1). The product of this situation is that there are no choices in Puyuma with regard to pronominal or null expression, and the construction in (31) carries both meanings.

3.7.4 Summary: the Subject Construction Hierarchy

The Subject Construction Hierarchy is an implicational scale whereby, if any construction in the hierarchy patterns accusatively, all the constructions to the left of it on the scale also pattern accusatively or pattern neither way (§3.7).

Puyuma observes the hierarchy. The two clause-internal constructions, case marking and verbal agreement, pattern ergatively. Two of the inter-clausal constructions, relativization and purpose (the serial verb construction in Puyuma) pattern accusatively. The leftmost construction on the hierarchy, coordination, patterns neither way.

4. Concluding thoughts: other Formosan (and Philippine) languages

Our first goal in this paper was to describe a Formosan language in a framework less opaque than the Philippinist approach. The reader will be the judge of whether we have succeeded. Our second goal was to provide insights that the Philippinist approach misses. We think we have achieved this in several respects. An important one was to show that, constructionally, there is no distinction between an ‘actor focus’ clause and an intransitive clause in Puyuma. The Philippinist approach misses this insight because it places the ‘actor focus’ in a paradigm with the transitive ‘focuses’ and separates it from the intransitive construction.

Our third goal was to propose an approach that would clarify differences among Formosan languages (and among Philippine-type languages in general). We believe that our approach will allow for this. Is there, for example, a distinction between an ‘actor focus’ clause and an intransitive clause in other Formosan languages? This is a question to which there is not yet a clear answer. We have not only shown how the morphosyntax of the Puyuma voice system functions. We have also touched on the syntax-discourse interface and indicated that (i) the oblique patient argument of a lexically transitive verb in an intransitive clause is always indefinite, and (ii) if a lexically transitive verb has a definite undergoer (O), the undergoer must be the subject. This has obvious discursal implications, mentioned in §3.3: Transitive subjects in Puyuma tend to track undergoer topicality.

We are confident that a similar approach to other Formosan languages will throw up important differences in this regard, which the Philippinist approach has overlooked. As we mentioned earlier, Anna Hsiou-chuan Chang’s work has shown that in Paiwan the oblique patient argument of a lexically transitive verb in an actor-subject clause is not always indefinite. It may follow from this that the subjects of undergoer-subject clauses are not the sole carriers of undergoer topicality. Huang (2002) suggests that in this regard Tsou functions much like Puyuma, but Seediq decidedly does not. In Tsou narratives patients of actor-subject clauses are overwhelmingly non-referential, whilst undergoers of undergoer-subject clauses are largely referential and definite, i.e. as in Puyuma subjects of undergoer-subject clauses tend to track undergoer topicality. In Seediq narratives, on the other hand, undergoers of actor-subject clauses are not significantly less referential or significantly less definite than those of undergoer-subject clauses and it is difficult to see what determines which voice speakers choose (Huang 2002:675-676, 685-687).

Shuanfan Huang (pers. comm.) points out on the basis of discourse studies that Squliq Atayal and Kavalan resemble Seediq in this regard. This means that our account of the Puyuma extended intransitive construction in §3.6 cannot be applied to these languages without considerable modification, since the corresponding structure in these languages encodes a different meaning; i.e., there is a different form-meaning pairing and therefore a different construction. It may also prove that an analysis of actor-subject clauses as intransitive is unsustainable for these languages. Explicit accounts of the differences between the functions of the ‘focuses’ in various Formosan and Philippine languages would be a welcome step away from Philippinist practice, which has tended to concentrate on the morphosyntax of the ‘focuses’ and to overlook their usage in discourse. They would also be invaluable in uncovering the history of the ‘focuses’ and their functions.

Appendices

Table 4: Summary of Puyuma verbal morphology

	INTRANSITIVE	TRANSITIVE		
		1	2	3
Realis	M-ROOT	ROOT- <i>aw</i>	ROOT- <i>ay</i>	ROOT- <i>anay</i>
Realis progressive	M- <i>Ca</i> -ROOT	<i>Ca</i> -ROOT- <i>aw</i>	<i>Ca</i> -ROOT- <i>ay</i>	<i>Ca</i> -ROOT- <i>anay</i>
Realis gerund	(as Realis)		< <i>in</i> >ROOT(- <i>an</i>)	
Imperative	ROOT	ROOT- <i>u</i>	ROOT- <i>i</i>	ROOT- <i>an</i>
Negative	(as non-negative)		ROOT- <i>i</i>	ROOT- <i>an</i>
Irrealis	<i>Ca</i> -ROOT		<i>Ca</i> -ROOT- <i>i</i>	<i>Ca</i> -ROOT- <i>an</i>
Irrealis gerund	(as Irrealis)		<i>Ca</i> -ROOT- <i>an</i>	
Projective	M-ROOT- <i>a</i>	(TRANSITIVE same as Realis)		
Negative hortative	M- <i>Ca</i> -ROOT	(no TRANSITIVE forms)		

Table 5: Summary of Proto Austronesian verbal morphology

	ACTOR VOICE	UNDERGOER VOICE		
		Patient subject	Location subject	Instrument subject
Realis	< <i>um</i> >ROOT	ROOT- <i>en</i>	ROOT- <i>an</i>	<i>Si</i> -ROOT
Realis perfective	< <i>umin</i> >ROOT	< <i>in</i> >ROOT	< <i>in</i> >ROOT- <i>an</i>	<i>Si</i> -< <i>in</i> >ROOT
Realis progressive	< <i>um</i> > <i>Ca</i> -ROOT	<i>Ca</i> -ROOT- <i>en</i>	<i>Ca</i> -ROOT- <i>an</i>	<i>Si</i> - <i>Ca</i> -ROOT
Irrealis	ROOT	ROOT- <i>u</i> , ROOT- <i>a</i>	ROOT- <i>i</i>	<i>án</i> - <i>i</i> + ROOT (ROOT- <i>áni</i>)
Projective	< <i>um</i> >ROOT- <i>a</i>	ROOT- <i>aw</i>	ROOT- <i>ay</i>	<i>án</i> - <i>ay</i> + ROOT (ROOT- <i>ánay</i>)

A. Voice-marking morphology and its history

A comparison of Puyuma verbal morphology in Table 4 with reconstructed Proto Austronesian (PAN) verbal morphology in Table 5 (based on Ross 2002) reveals that the two systems are fairly similar in organization but that certain morphemes reconstructed for PAN are missing in Puyuma or display a radically different distribution in Puyuma.

Assuming that the PAN reconstruction is correct, the following statements can be made about Puyuma forms:

1. Intransitive forms reflect PAn quite closely (Puyuma *M* reflects PAn **<um>*).
2. Transitive realis forms in *-aw*, *-ay* and *-anay* (i.e. except the gerund) reflect PAn projective forms.
3. Imperative forms in *Ø*, *-u*, *-i* and *-an* reflect PAn irrealis forms.
4. The realis and irrealis gerunds reflect the PAn realis perfective and realis progressive forms respectively. Otherwise, PAn perfective forms in **<in>* are lost.
5. The PAn undergoer voice realis affixes **-en*, **-an* and **Si-* are lost, except where **-an* is retained in gerunds.

Similar changes appear to have taken place in Tsou and Saaroa (but not in their assumed close relative Kanakanavu), but these have not been investigated.

There are two possible explanations for the differences between PAn and Puyuma, and we make no attempt to choose between them:

Incorrect reconstruction The PAn reconstruction in Table 5 is incorrect and Puyuma (and Tsou and Saaroa) reflect PAn more faithfully than other Formosan languages. Since the PAn forms listed as lost in 5 are all considered to have been nominalizing morphemes in pre-PAn, it is possible that they retained this status in PAn and only acquired realis verbal status after the ancestor of Puyuma had separated from other early Austronesian languages. Furthermore, the forms listed in 2 and 3 are considered to have been the major verbal morphemes in pre-PAn, and it is possible that they have retained this status in Puyuma. If this hypothesis is correct, then Tsou and Saaroa would probably represent another early branch of the tree. For a summary of hypotheses about pre-PAn and PAn verbal morphology, see Ross (2002).

Loss of morphemes The PAn reconstruction in Table 5 is approximately correct, and Puyuma has undergone the changes and losses listed above. Essentially, PAn irrealis forms have become imperatives (this is plausible), and PAn undergoer-voice projectives have extended their function to the realis, displacing the PAn forms. It is not obvious how might this have happened.

B. Possessor pronouns

Table 6: Puyuma possessor pronouns (P/ = Possessor of)

	1S	2S	3S	1IP	1EP	2P	3P
P/NOM 1	<i>ku=</i>	<i>nu=</i>	<i>tu=</i>	<i>ta=</i>	<i>niam=</i>	<i>mu=</i>	<i>tu=</i>
P/NOM 2	<i>nanku</i>	<i>nanu</i>	<i>nantu,</i> <i>nantaw</i>	<i>nanta</i>	<i>naniam</i>	<i>nanemu</i>	<i>nantu,</i> <i>nantaw</i>
P/OBL/DEF	<i>kanku</i>	<i>kanu</i>	<i>kantu</i>	<i>kanta</i>	<i>kaniem</i>	<i>kanemu</i>	<i>kantu</i>
P/OBL/INDEF	<i>Daku</i>	<i>Danu</i>	<i>Datu</i>	<i>Data</i>	<i>Daniam</i>	<i>Danemu</i>	<i>Datu</i>

At first sight, Puyuma seems to have a remarkable array of possessor pronouns which also mark the case of their noun phrase as nominative or oblique and, for the obliques, as definite or indefinite. These are the distinctions made by the common NP markers in Table 2, and closer inspection of the possessor pronouns in Table 6 shows that all sets except the possessor-of-nominative proclitics consist of quite transparent combinations of the common NP markers in Table 2 with the possessor-of-nominative proclitics. The possessor-of-nominative proclitics are similar, but not identical, to the genitive proclitics in Table 1.

Thus the set labeled ‘P/NOM 2’ consists of the nominative definite common noun marker *na* and the clitics in the row ‘P/NOM 1’. If the clitic-initial consonant is non-nasal, then *-n-* is inserted, e.g. *na-n-ku* 1S. If the clitic-initial consonant is *m-*, then *-ne-* is inserted: this applies only to *na-ne-mu* 2P. If the clitic-initial consonant is *n-*, then there is no insertion, as in *na-nu*.³⁶

The set labeled ‘P/OBL/DEF’ consists of the *ka-* element of the oblique definite common noun marker *kana* and the clitics in the row ‘P/NOM 1’. The same insertion rules apply.

The set labeled ‘P/OBL/INDEF’ consists of the oblique indefinite common noun marker *Da* and the clitics in the row ‘P/NOM 1’. The same insertion rules apply.

The two sets of possessor-of-oblique forms shown in Table 6 have longer variants not shown in the table. They are formed by combining *ka-* and *Da-* respectively with the set labeled ‘P/NOM 2’. Thus the longer 1S forms are *ka-nanku* and *Da-nanku*.³⁷

³⁶ These insertion rules do not necessarily reflect the actual historical processes that led to today’s forms.

³⁷ Only third person forms with *-tu* occur in these sets, not with *taw*.

C. Intransitive verbal morphology

Table 7: Puyuma intransitive morphology

		Imperative	Realis	Irrealis	Realis progressive
1a	‘bathe’	<i>dirus</i>	<i>dirus</i>	<i>da-dirus</i>	<i>da-dirus</i>
1b	‘swim’	<i>Languy</i>	<i>me-Languy</i>	<i>La-Languy</i>	<i>me-La-Languy</i>
1c	‘go’	<i>uka</i>	<i>m-uka</i>	<i>a-uka</i>	<i>m-a-uka</i>
2a	‘tell’	<i>rengay</i>	<i>ma-rengay</i>	<i>ra-rengay</i>	<i>ma-ra-rengay</i>
2b	‘be drunk’	<i>ka-Dangan</i>	<i>ma-Dangan</i>	<i>ka-Da-Dangan</i>	<i>ma-Da-Dangan</i>
3a	‘give’	<i>beray</i>	<i>beray</i>	<i>ba-beray</i>	<i>ba-beray</i>
3b	‘be hot’	<i>ka-bias</i>	<i>bias</i>	<i>ka-ba-bias</i>	<i>ba-bias</i>
3c	‘be good’	<i>ka-inaba</i>	<i>inaba</i>	<i>ka-i-na-naba</i>	<i>i-na-naba</i>
4a	‘fly’	<i>u-bii</i>	<i>m-u-bii</i>	<i>u-a-bii</i>	<i>m-u-a-bii</i>
4b	‘have money’	<i>pi-paisu</i>	<i>mi-paisu</i>	<i>pi-a-paisu</i>	<i>mi-a-paisu</i>
4c	‘hear’	<i>ki-lengaw</i>	<i>ki-lengaw</i>	<i>ki-a-lengaw</i>	<i>ki-a-lengaw</i>

Table 3 provides a summary of Puyuma verbal morphology. However, Puyuma verbs fall into a number of morphological classes in which the algebraic morphemes *M-* and *Ca-* are manifested in different ways. A single example from each class is given in Table 7 in order to illustrate the intransitive forms in each class (the transitive forms do not differ by class). The realis progressive form is predictable from the realis and irrealis forms.

Classes 1a, 1b, and 1c are variants of a single class, conditioned by the root-initial segment: The roots of Class 1b begin with nasal or liquid, those of Class 1a with another consonant, and those of 1c with a vowel. Class 1 verbs are dynamic, and usually have transitive counterparts, although the semantic relationship between the intransitive and transitive forms is not always predictable. Thus *dirus-aw* ‘wash-TR1’ means ‘wash (someone else)’ but *dirus* ‘wash-ITR’ means ‘wash (oneself), bathe, swim’.

Class 2 verbs share *ma-* as their variant of *M-*, but in Class 2b a prefix *ka-* surfaces when there is no *M-*. It is not clear that *ka-* has a synchronic function. Some Class 2 verbs are dynamic, others stative. In particular, many Class 2b verbs are stative.

Class 3 verbs share \emptyset as their variant of *M-*, but in Class 3b, like 2b, *ka-* again surfaces when there is no *M-*. In Class 3c, reduplication affects the syllable after initial *i-*. Some Class 3 verbs are dynamic, others stative.

Classes 4a, 4b, and 4c each have a derivational prefix, respectively *u-*, *pi-* and *ki-*, which occurs across the whole paradigm. The irrealis and realis progressive are formed

not by *Ca-* reduplication but by the insertion of *-a-* between the prefix and the root.

Class 4a manifests *M-* in the same way as Class 1c, by prefixing *m-*. Its members are verbs of movement, e.g. *u-arak* ‘dance’, *u-burek* ‘return’, *u-Tereb* ‘fall’.

In Class 4b *M-* and *pi-* coalesce as *mi-*. Its members express the existence or ownership of whatever is expressed by the root, e.g. *mi-kataguin* ‘take/have a spouse’, *mi-riwanes* ‘have a rainbow’, *mi-ruma* ‘have a family’, *mi-kabung* ‘wear a hat’.

Class 4c resembles Class 3 in having \emptyset for *M-*. When *ki-* is prefixed to a noun, the resulting verb means ‘obtain NOUN’, e.g. *ki-apuT* ‘pick flowers’ (*apuT* ‘flower’, *ki-kuraw* ‘1sh’ (*kuraw* ‘1sh’. When *ki-* is prefixed to a verb root, the subject is patient or recipient, e.g. *ki-tuLuD* ‘receive, catch’ (*tuLuD* ‘pass to’), *ki-aDas* ‘get lifted’ (*aDas* ‘lift’).

References

- Andrews, Avery D. 1985. The major functions of the noun phrase. *Language Typology and Syntactic Description*, ed. by Timothy Shopen, 62-154. Cambridge: Cambridge University Press.
- Blake, Barry J. 1988. Tagalog and the Manila-Mt Isa axis. *La Trobe Working Papers in Linguistics* 1:77-90.
- Blake, Frank R. 1906a. Contributions to comparative Philippine grammar. *Journal of the American Oriental Society* 27:317-396.
- Blake, Frank R. 1906b. Expression of case by the verb in Tagalog. *Journal of the American Oriental Society* 27:183-189.
- Blust, Robert A. 1999. Notes on Pazeh phonology and morphology. *Oceanic Linguistics* 38.2:321-365.
- Blust, Robert A. 2002. Notes on the history of ‘focus’ in Austronesian languages. *The History and Typology of Western Austronesian Voice Systems*, ed. by Fay Wouk and Malcolm Ross, 63-78. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Cauquelin, Josiane. 1991a. *Dictionnaire puyuma-français*. Paris: École française de l’extrême-orient.
- Cauquelin, Josiane. 1991b. The Puyuma language. *Bijdragen tot de Taal-, Land- en Volkenkunde* 147:17-60.
- Chang, Henry Yungli. 2003. AF verbs: transitive, intransitive or both? Paper prepared at the Symposium on the Notion ‘Verb’ in Formosan Languages. Taipei: Academia Sinica.
- Croft, William. 2001. *Radical Construction Grammar: Syntactic Theory in Typological Perspective*. New York: Oxford University Press.

- Croft, William. 2003. Lexical rules vs constructions: a false dichotomy. *Motivation in Language: Studies in Honour of Gunter Radden*, ed. by Hubert Cuyckens, Thomas Berg, Rene Dirven, and Klaus-Uwe Panther. Amsterdam: John Benjamins.
- De Guzman, Videa P. 1988. Ergative analysis for Philippine languages: an analysis. *Studies in Austronesian Linguistics*, ed. by Richard McGinn, 323-345. Athens: Center for Southeast Asian Studies, Ohio University Center for International Studies.
- De Guzman, Videa P. 1997. Verbal affixes in Tagalog: inflection or derivation? *Proceedings of the Seventh International Conference on Austronesian Linguistics*, ed. by Cecilia Odé and Wim Stokhof, 303-325. Amsterdam: Rodopi.
- Dean, James C. 1958. Some principal grammatical relations of Bilaan. *Oceania Linguistics Monographs* 5:83-90.
- Dixon, R. M. W. 1979. Ergativity. *Language* 55.1:59-138.
- Dixon, R. M. W. 1994. *Ergativity*. Cambridge: Cambridge University Press.
- Dryer, Matthew. 1997. Are grammatical relations universal? *Essays on Language Function and Language Type*, ed. by Joan Bybee, John Haiman, and Sandra A. Thompson, 113-143. Amsterdam: John Benjamins.
- Egli, Hans. 1990. *Paiwangrammatik*. Wiesbaden: Otto Harrassowitz.
- Evans, Bethwyn. 2003. *A Study of Valency-changing Devices in Proto Oceanic*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Ferrell, Raleigh J. 1972. Verb systems in Formosan languages. *Langues et techniques*, ed. by Jacqueline M. C. Thomas and Lucien Bernot, 121-128. Paris: Klincksieck.
- Ferrell, Raleigh J. 1982. *Paiwan Dictionary*. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Fillmore, Charles J. 1968. The case for case. *Universals in Linguistic Theory*, ed. by Emmon Bach and Robert T. Harms, 1-88. New York: Holt, Rinehart and Winston.
- Fillmore, Charles J. 1988. The mechanisms of 'Construction Grammar'. *BLS* 14:35-55. Berkeley: Berkeley Linguistic Society.
- Forster, Janette. 1964. Dual structure of Dibabawon verbal clauses. *Oceanic Linguistics* 3.1:26-48.
- French, Kileen Matsuda. 1988. The focus system in Philippine languages: an historical overview. *Philippine Journal of Linguistics* 19.1:1-27.
- Gerdts, Donna B. 1988. Antipassives and causatives in Ilokano: evidence for an ergative analysis. *Studies in Austronesian Linguistics*, ed. by Richard McGinn, 295-321. Athens: Ohio University Press.
- Goldberg, Adele E. 1995. *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

- Goldberg, Adele E. 2002. Surface generalizations: an alternative to alternations. *Cognitive Linguistics* 13.4:327-356.
- Healey, Alan. 1958. Notes on Yogad. *Oceania Linguistics Monographs* 5:111-120.
- Healey, Phyllis M. 1960. *An Agta Grammar*. Manila: Bureau of Printing.
- Himmelman, Nikolaus P. 2002. Voice in western Austronesian: an update. *The History and Typology of Western Austronesian Voice Systems*, ed. by Fay Wouk and Malcolm Ross, 7-16. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Huang, Lillian M. 2000. *Bēinányǔ Cānkǎo Yǔfǎ* [Puyuma Reference Grammar]. Taipei: Yuanliu.
- Huang, Shuanfan. 2002. The pragmatics of focus in Tsou and Seediq. *Language and Linguistics* 3.4:665-694.
- Huteson, Greg. 2003. Report on language proficiency, language use, and language attitudes among the Puyuma. Manuscript.
- Jeng, Heng-hsiung. 1977. *Topic and Focus in Bunun*. Institute of History and Philology Special Publication No. 72. Taipei: Institute of History and Philology, Academia Sinica.
- Kay, Paul, and Charles J. Fillmore. 1999. Grammatical constructions and linguistic generalizations: the *What's X doing Y?* construction. *Language* 75.1:1-33.
- Kazenin, Konstantin I. 1994. Split syntactic ergativity: toward an implicational hierarchy. *Sprachtypologie und Universalienforschung* 47:78-98.
- Kibrik, Aleksandr. 1997. Beyond subject and object: towards a comprehensive relational typology. *Linguistic Typology* 1.3:279-346.
- Kroeger, Paul. 1993. *Phrase Structure and Grammatical Relations in Tagalog*. Stanford: CSLI.
- Lambrecht, Knud. 1994. *Information Structure and Sentence Form: Topic, Focus and the Mental Representations of Discourse Participants*. Cambridge: Cambridge University Press.
- Li, Paul Jen-kuei. 1973. *Rukai Structure*. Institute of History and Philology Special Publication No. 64. Taipei: Institute of History and Philology, Academia Sinica.
- Liao, Hsiu-chuan. 2004. *Transitivity and Ergativity in Formosan and Philippine Languages*. Honolulu: University of Hawaii dissertation.
- Manning, Christopher D. 1996. *Ergativity: Argument Structure and Grammatical Relations*. Stanford: CSLI.
- McKaughan, Howard. 1958. *The Inflection and Syntax of Maranao Verbs*. Manila: Bureau of Printing.
- McKaughan, Howard. 1973. Subject versus topic. *Parangal Kay Cecilio Lopez*, ed. by Andrew B. Gonzalez, 206-213. Philippine Journal of Linguistics Special Monograph

- No. 4. Quezon City: Linguistic Society of the Philippines.
- Michaelis, Laura. 2000. Aspectual meaning as constructional meaning. *Proceedings of the BFG00 Conference Workshops*, ed. by Miriam Butt and Tracy Holloway King. Stanford: CSLI.
- Michaelis, Laura A., and Knud Lambrecht. 1996. Toward a construction-based theory of language function: the case of nominal extra position. *Language* 72.2:215-247.
- Miller, Jeanne. 1964. The role of verb stems in the Mamanwa kernel verbal clauses. *Oceanic Linguistics* 3.1:87-100.
- Mithun, Marianne. 1994. The implications of ergativity for a Philippine voice system. *Voice: Form and Function*, ed. by Barbara Fox and Paul J. Hopper, 247-277. Amsterdam: John Benjamins.
- Morey, Virginia. 1964. Distribution restrictions on co-occurrence of aspect and focus morphemes in Ata verbs. *Oceanic Linguistics* 3.1:69-86.
- Newell, Leonard E. 1964. Independent clause types of Batad Ifugao. *Oceanic Linguistics* 3.1:171-199.
- Pawley, Andrew K. 1973. Some problems in Proto-Oceanic grammar. *Oceanic Linguistics* 12.1-2:103-188.
- Payne, Thomas E. 1994. The pragmatics of voice in a Philippine language: actor-focus and goal-focus constructions in Cebuano narrative. *Voice and Inversion*, ed. by Talmy Givón, 317-364. Amsterdam: John Benjamins.
- Payne, Thomas E. 1997. *Describing Morphosyntax: A Guide for Field Linguists*. Cambridge: Cambridge University Press.
- Pike, Kenneth L. 1963. A syntactic paradigm. *Language* 39.2:216-230.
- Reid, Lawrence A. 1966. *An Ivatan Syntax*. Oceanic Linguistics Special Publication No. 2. Honolulu: Pacific and Asian Linguistics Institute, University of Hawaii.
- Reid, Lawrence A., and Hsiu-chuan Liao. 2004. A brief syntactic typology of Philippine languages. *Language and Linguistics* 5.2:433-490.
- Ross, Malcolm. 2002. The history and transitivity of western Austronesian voice and voice-marking. *The History and Typology of Western Austronesian Voice Systems*, ed. by Fay Wouk and Malcolm Ross, 17-62. Canberra: Pacific Linguistics, Research School of Pacific and Asian Studies, Australian National University.
- Shand, Jean. 1964. Categories and markers of tense, focus and mode in Ilianen Manobo. *Oceanic Linguistics* 3.1:58-68.
- Starosta, Stanley. 1986. Focus as recentralisation. *FOCAL I: Papers from the Fourth International Conference on Austronesian Linguistics*, ed. by Paul Geraghty, Lois Carrington, and S. A. Wurm, 73-95. Canberra: Australian National University.
- Starosta, Stanley. 1999. Transitivity, ergativity, and the best analysis of Atayal case marking. *Selected Papers from the Eighth International Conference on Austronesian*

- Linguistics*, ed. by Elizabeth Zeitoun and Paul Jen-kuei Li, 371-392. Taipei: Institute of Linguistics, Academia Sinica.
- Thomas, David D. 1958. Mansaka sentence and sub-sentence structures. *Philippine Social Sciences and Humanities Review* 23:339-358.
- Ting, Pang-Hsin. 1978. Gǔ-Běinányǔ de nǐchè [Reconstruction of Proto-Puyuma phonology]. *Bulletin of the Institute of History and Philology, Academia Sinica* 49.3:321-392.
- Tsuchida, Shigeru. 1976. *Reconstruction of Proto-Tsouic Phonology*. Studies of Languages and Cultures of Asia and Africa, Monograph Series No. 5. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa, Tokyo University of Foreign Studies.
- Tsuchida, Shigeru. 1980. Puyuma-go (Tamarakao hōgen) goi: fu gohōgaisetsu oyobi tekisuto [Puyuma (Tamalakaw dialect) vocabulary: with grammatical notes and texts]. *Kuroshio no Minzoku, Bunka, Gengo* [Ethnology, Cultures and Languages along the Black Current], ed. by Kuroshio Bunka no Kai (Black Current Cultures Committee), 183-307. Tokyo: Kadokawa Shoten.
- Tsuchida, Shigeru. 1983. Puyuma-English index. *Working Papers in Linguistics, University of Tokyo* 83:10-63.
- Tung, T'ung-ho. 1964. *A Descriptive Study of the Tsou Language*. Taipei: Institute of History and Philology, Academia Sinica.
- Wu, Peter. 1969. A descriptive analysis of Bunun language. Term paper. Washington, DC: Georgetown University.

[Received 28 March 2005; revised 17 August 2005; accepted 31 August 2005]

Malcolm Ross
Research School of Pacific and Asian Studies
and Centre for Research on Language Change
The Australian National University
Canberra, ACT 0200
Australia
malcolm.ross@anu.edu.au

Stacy Fang-ching Teng
stacyteng@yahoo.com.tw

台灣南島語及語言類型

Malcolm Ross 鄧芳青

澳洲國立大學

長久以來，許多學者在描寫台灣南島語時，慣於使用菲律賓南島語學者發展出來的理論框架及慣用語。例如，這些學者們談論動詞性結構的「焦點系統」(focus system) 或是子句的「主題」(topic)。用這些慣用語來描述菲律賓及台灣南島語常造成其他語言學者理解上的困難，因為這些慣用語 (一) 對非研究台灣或菲律賓南島語的學者而言是不熟悉的；(二) 在某些方面並不完備，例如，無法清楚界定所謂「主題」(topic) 的功能為何。

本文試著用語言類型學者或者其他非研究台灣南島語學者較為熟悉的用語及理論框架，來重新描述台灣南島語的某些構詞及句法特徵。其中，我們會重新檢視所謂「主題」(topic) 的概念，也會以語言類型學者 William Croft 的理論架構重新描述所謂的「焦點系統」(focus system)。這種重新的描述可帶來兩個優點：增加各個台灣南島語研究作品的可理解度，並且凸顯各語言的相異度。

關鍵詞：卑南語，台灣南島語，語言類型學，語法

