



Project
MUSE[®]

Today's Research. Tomorrow's Inspiration.

<http://muse.jhu.edu>

Squib

Is Puyuma a Primary Branch of Austronesian? A Reply to Sagart

Stacy F. Teng and Malcolm Ross

ACADEMIA SINICA AND AUSTRALIAN NATIONAL UNIVERSITY

Ross (2009) proposes the Nuclear Austronesian hypothesis, according to which the Formosan languages Puyuma, Rukai, and Tsou are each probably a primary branch of Austronesian and all Austronesian languages other than these three belong to a single, Nuclear Austronesian, branch defined by the nominalization-to-verb innovation originally proposed by Starosta, Pawley, and Reid (1981, 1982) for Proto-Austronesian itself. Sagart (2010) argues that there is evidence that Puyuma has also undergone the nominalization-to-verb innovation and is accordingly not a primary branch of Austronesian. In this short paper we show that Sagart's evidence is based on misanalyses of Puyuma data and that these data do not reflect the nominalization-to-verb innovation. Sagart's argument against the Nuclear Austronesian hypothesis does not stand up to closer scrutiny.

1. INTRODUCTION.¹ In his paper “Is Puyuma a primary branch of Austronesian?” which appeared in the June 2010 issue of *Oceanic Linguistics*, Laurent Sagart “point[s] out some problems” with the outline of early Austronesian phylogeny presented in Ross (2009). Starosta, Pawley, and Reid (1981, 1982) showed that nominalizing morphology was reanalyzed as verbal morphology in very early Austronesian. Thus the undergoer-voice affixes *-en ‘patient subject’, *-an ‘location subject’, and *Si- ‘circumstance subject’, along with the perfective infix *⟨in⟩, reflected in the verb forms of many Formosan and Philippine languages, were originally nominalizers. We will call this innovation the “nominalization-to-verb” (Nom-to-V) innovation.² Ross's phylogeny is based on the claim that this innovation had not occurred prior to Proto-Austronesian, as its discoverers had thought, since it is not reflected in Puyuma, Rukai, or Tsou. Instead, Ross suggests that it occurred in what he calls “Proto-Nuclear Austro-

1. We thank Robert Blust and Elizabeth Zeitoun for their helpful comments on an earlier draft of this squib. The two authors, of course, retain responsibility for its contents. It is difficult to describe our roles precisely, but the impetus for this paper came from Teng, who provided the diachronic phonological analysis of Puyuma in section 2 and the synchronic morphosyntactic analyses and the arguments in sections 3 and 5. Ross provided the comparisons in section 2 and the argument of section 4. The argument of section 2 was a joint effort.
2. Abbreviations used here other than those in the Leipzig rules are: Nom-to-V, nominalization-to-verb; PAN, Proto-Austronesian; PERS, personal; PMP, Proto-Malayo-Polynesian; RED, reduplication; UVC, undergoer voice with circumstantial or conveyance-like subject; UVL, undergoer voice with location-like subject; UVP, undergoer voice with patient-like subject.

nesian,” a language ancestral to all Austronesian languages other than the three just named. He assumes these three languages to be primary branches of Austronesian, coordinate with Proto-Nuclear Austronesian.

Ross (2009) focuses particularly on Puyuma. In his paper, Sagart alleges that Ross is wrong to believe that Puyuma does not reflect the Nom-to-V innovation. He proposes instead that Puyuma did undergo this innovation, but that its reflexes have been partly overlaid by subsequent innovations. Crucially, he claims that Puyuma retains reflexes of the Nom-to-V innovation and is therefore, *contra* Ross, not a primary branch of Austronesian.

Sagart also deals briefly with Tsou and Rukai verbal morphology. His discussion of Tsou adds nothing to published hypotheses, while his argument that the prefix *ki-* ‘passive’ has replaced reflexes of the undergoer-voice affixes **-en*, **-an*, and **Si-* in Rukai (Sagart 2009) requires separate treatment.³ He also comments that the Nuclear Austronesian hypothesis cannot account for shared lexical innovations attributed to Tsuchida’s (1976) Tsouic subgroup and to the various subgroups posited in Sagart (2004). A response to the latter is given by Winter (2010). The Tsouic question is complex and needs treatment elsewhere.

Curiously, Sagart implies that because the Nom-to-V innovation entailed only a single reanalysis, it is not of particular significance for subgrouping. The subgrouping value of an innovation, however, lies in the low probability of its independent, parallel occurrence in the languages attributed to the subgroup. Sagart remarks that the Nom-to-V innovation was “admittedly a complex reanalysis event that affected several nominalizers simultaneously.” This surely means that the probability of independent, parallel occurrence is low. The fact that the same complex set of outcomes is reflected in Nuclear Austronesian languages of Taiwan and the Philippines indicates that they are descended from a single interstage.⁴

Sagart’s main concern, however, is with the facts of Puyuma. He claims that certain facts present in the Puyuma data available to him are evidence that a language ancestral to Puyuma did undergo the Nom-to-V innovation. We argue here that these data are misanalyzed and that there is no evidence that a pre-Puyuma interstage underwent this innovation. Sagart’s evidence is of two kinds:

1. He argues that the infix *<in>*, occurring in verb stems and claimed by Teng (2008a) always to mark a nominalization in Nanwang Puyuma, in fact marks perfective finite verbs in one Nanwang Puyuma sociolect. If this is true, Puyuma reflects the Nom-to-V innovation, and Ross’s evidence that it is a primary branch of Austronesian evaporates.
2. He presents evidence that the undergoer-voice patient subject suffix **-en* is reflected in Puyuma verbs of the shape **paR-Numeral-en*. If this is true, then Puyuma would appear to have fossilized reflexes of **-en* in finite verbs, again reflecting the Nom-to-V innovation.

We discuss these two sets of data in reverse order, because the refutation of the second claim is a matter of historical phonology and is relatively simple, and also because the arguments against the first claim are morphosyntactic and enable us to more easily make the transition to our discussion of Sagart’s third section, which presents a hypothesis as to

3. Sagart’s wording could be taken to indicate that Zeitoun and Teng (2009) agree with his history. In fact, they remain neutral with regard to subgrouping (Zeitoun and Teng 2009:481).

4. This set of outcomes is also reflected much further afield among Austronesian languages outside Taiwan, but the evidence is often subtle, as it is overlaid by further typological changes.

how Puyuma, having allegedly undergone the Nom-to-V innovation, has since lost most of its manifestations.

2. DOES PUYUMA REFLECT PAN *paR-X-en OR *paR-X-[e]N? It is fairly easy to show on the basis of comparative evidence that what Sagart takes to be a Puyuma reflex of *paR-X-en, where X is a numeral below ten, is in fact a reflex of *paR-X-[e]N. He writes, “there can be no doubt that Puyuma forms like *parpuan* [‘do twice’] contain a fossilized UVP suffix *-en.” And later: “Puyuma -n reflects PAN *-n unambiguously, so that *parpuan* cannot be from *paR-puSa-N,” and is evidently taken to reflect *paR-puSa-en. This level of confidence is simply not justified by the data. The assumption that Puyuma -n *always* reflects PAN *-n and never PAN *-N is mistaken,⁵ as we show below.

First, however, we give evidence for the reconstruction of the PAN suffix *-[e]N, which was attached to numerals and apparently had a multiplicative or “recurrence” function (see also Zeitoun, Teng, and Ferrell 2010).⁶ As Li (2006:140–41) shows, decade numerals from twenty upward had the structure PAN *ma-X-[e]N. Examples are in table 1, and are chosen for their regularity in PAN (the relationship between PAN *duSa ‘2’ and *ma-puSa-N ‘20’ is irregular), and to illustrate the fact that the suffix was *-N after a vowel, *-eN after a consonant.⁷ PAN *ma- here is evidently not the indicative stative marker that occurs in many Formosan languages, alternating with plain stative *ka- (Zeitoun and Huang 2000), since the Mantauran Rukai decade terms begin in *ma-ma-*, alternating with *ka-ma-* (Zeitoun 2007:256): that is, the first *ma-* is stative, the second reflects PAN decade-forming *ma-. The Puyuma terms are those used for counting animals and objects (Cauquelin 1991), and here *ka-* intervenes between decade-forming *ma-* and the numeral.

PAN multiplicative *-[e]N is also reflected in other numeral-based forms in some of the Formosan languages in table 1:⁸

- | | | |
|---------------------------------|--|---------------------------------|
| (1) NANWANG PUYUMA ⁹ | | |
| pari-asa-l | | ‘do once’ |
| par-pua-n | | ‘do twice’ |
| par-telru-n | | ‘do three times’ |
| | | (Cauquelin 1991) |
| (2) MANTAUAN RUKAI | | |
| a. maka-tolro-lo | | ‘for three (days/months/years)’ |
| maka-ʔepate-le | | ‘for four (days/months/years)’ |
| maka-lrima-le | | ‘for five (days/months/years)’ |
| maka-eneme-le | | ‘for six (days/months/years)’ |

5. We retain *N for consistency with Sagart. As *N was almost certainly a lateral, we would have preferred, like Tsuchida (1976) and others, to use *L.

6. The term “recurrence” was suggested by Raleigh Ferrell (Elizabeth Zeitoun, pers. comm.).

7. A dash in the table indicates that the term does not reflect the reconstructed PAN form.

8. Orthography: in all Formosan languages except Tsou, *e* is [ɛ]; in Tsou *e* is [e]; *lr*, *dr*, *tr* are retroflex [l̥], [d̥], [t̥]; *lh* is a voiceless lateral fricative [l̥h]; *sh* is [ʃ]; and Thao *z* is [ð].

9. In Nanwang, members of this set above ‘do three times’ are unsuffixed. Sagart cites further members of this set in *-n* in other Puyuma dialects.

TABLE 1. THE FORMATION OF TERMS FOR ‘THIRTY’, ‘FORTY’, ‘FIFTY’ AND ‘SIXTY’ IN FORMOSAN LANGUAGES*

PAN	‘three’ *telu	‘thirty’ *ma-telu-N	‘four’ *Sepat	‘forty’ *ma-Sepate-eN	‘five’ *lima	‘fifty’ *ma-lima-N	‘six’ *enem	‘sixty’ *ma-nem-eN
Puyuma (Nanwang)	telru	ma-ka-telru-n	pat	ma-ka-pet-el	lrima	—	nem	ma-ka-nem-en
Rukai (Ianan)	tolro	ma-tolro-l	so?ate	ma-so?ate-l	lrima	ma-lrima-l	eneme	ma-neme-l
Rukai (Mantauran)	tolro	ma-ma-tolro-lo	pate	ma-ma-?epate-le	lrima	ma-ma-lrima-le	eneme	ma-ma-neme-le
Tsou	tueu	m-tueu-hu	suptu	m-suptu-hu	eimo	m-eimo-hu	nomu	mo-nmu-hu
Kanakanavu	tuulu	ma-tu-n	supate	ma-sepate-n	liima	ma-lima-en	neema	ma-enem-en
Saaroa	u-tulu	ma-tulu-lhu	upaate	ma-upate-lh	ku-lima	ma-lima-lh	eneme	ma-eneme-lh
Bunun (Takivatan)	tau	ma-tiu-n	pat	ma-sipate-un	hima	ma-hima?-un	num	ma-num-un
Thao	turu	ma-turu-z	shpat	ma-shpat-iz	rima	ma-rima-z	—	—
Saisiyat (Iaai)	tolro?	ma-tolro-l	shepat	—	—	—	—	—
Atayal (Mayrimax)	tuu?	ma-tuu-l	sepate	ma-spat-ul	?rima?	ma-rima-l	—	—
Seediq (Paran)	teru	m-teru-l	sepac	m-spat-ul	rima	m-rima-l	—	—

* Sources: Nanwang Puyuma: Cauquelin (1991), Teng (2008a); Mantauran Rukai: Zeitoun (2006); Tsou: Szakos (1994); Kanakanavu: Ho (1997); Takiva-tan Bunun: De Busser (2009); Thao: Blust (2003); Taai Saisiyat: Li (1978). Otherwise Li (2006) and Paul Jen-kuei Li, pers. comm.

- b. ?o-tolro-lo 'three times'
 ?o-?epate-le 'four times'
 ?o-lrima-le 'five times'
 ?o-eneme-le 'six times'
- c. ?ao-tolro-lo 'the third time'
 ?ao-?epate-le 'the fourth time'
 ?ao-lrima-le 'the fifth time'
 ?ao-eneme-le 'the sixth time'
- d. ta?a-tolro-lo 'three times as much'
 ta?a-?epate-le 'four times as much'
 ta?a-lrima-le 'five times as much'
 ta?a-eneme-le 'six times as much' (Zeitoun 2007:263–68)
- (3) TSOU
- a. ei-to-teu-hu 'three times'
 ei-so-spot-hu 'four times'
 ei-eo-emo-hu 'five times'
- b. ma?hi-to-teu-hu 'three hours'
 ma?hi-spot-hu 'four hours'
 ma?hi-emo-hu 'five hours'
- c. mi-teu-hi 'stay three days'
 mi-spot-hu 'stay four days'
 mi-emo-hu 'stay five days' (Szakos 1994)
- (4) THAO
- a. mu-turu-z 'do three times'
 mu-shpat-iz 'do four times'
 mu-rima-z 'do five times'
- b. kun-turu-z, lhun-turu-z 'do three times'
- c. lhin-turu-z-in 'be divided into three parts' (Blust 2003)

Some Formosan languages that do not reflect PAN *-[e]N in their decade terms nonetheless reflect it elsewhere:

- (5) KAVALAN
 turu 'three' qi-taru-n 'three times'
 (Li and Tsuchida 2006)
- (6) SIRAYA
 turu 'three' turu-l 'three times'
 pitu 'seven' pitu-l 'seven times' (Adelaar 2008)
- (7) PAIWAN
 a. celru 'three' maka-celru-l 'three times, three days'
 sepac 'four' maka-simac-el 'four times, four days'
 lrima 'five' maka-lrima-l 'five times, five days'
 unem 'six' maka-nem-el 'six times, six days'

- b. s(ɛm)an-celru-l 'do three times'
 s(ɛm)an-simac-el 'do four times'
 s(ɛm)an-lrima-l 'do five times'
 s(ɛm)an-nem-el 'do six times'

(Ferrell 1982; Egli 1990:152; Egli 2002)

Formosan reflexes of PAN *n and *N are shown in table 2. It is clear that the suffixes in table 1 and in (1) through (7) reflect PAN *-[e]N. Notably, however, Puyuma is the one language with two reflexes of *N, namely *n* and *l*. Of these, *l* is more frequent than *n* in Puyuma reflexes and, as noted by Ting (1978), we find, for example, Nanwang Puyuma *ʔasil* 'salty' (< PAN *qasiN), *ʔudal* 'rain' (< PAN *quzaN) but *bulran* 'moon' (< PAN *bulaN), *talrun* 'grass' (< PAN *CaluN), *a-bilrin* 'late' (< PAN *biliN 'follow'), *ayan* 'termite' (< PAN *SayaN; Blust 2001). We also find alternations between *l* and *n* among Puyuma dialects in certain lexical items: Pinaski *harimalaw* 'golden bug' vs. Rikavung *harimanamanaw*; Nanwang *nanibuan* 'offspring' vs. *lalivuan* in Pinaski, Ulivelivek, Katipul, and Kasavakan; and *ʔapelil* 'bitter' in Nanwang, Pinaski, and Ulivelivek vs. *ʔapelin* in Kasavakan, Rikavung, and Katipul.

At least one environment in which PAN final *-N has become Puyuma -n can be identified, namely, when the onset of the final syllable is the retroflex lateral -lɾ- (< PAN *l): cf. *bulran*, *talrun*, and *bilrin* in the previous paragraph, as well as *ma-ka-telru-n* 'thirty' (< PAN *ma-telu-N, table 1), *par-telru-n* 'do three times' (< PAN *paR-telu-N), and Tamalakaw Puyuma *ma-ka-walru-n* 'eighty' (Tsuchida 1980:267) (< PAN *ma-walu-N).¹⁰ The synchronic alternation in Nanwang between *tilril* 'book' and *tilrin* 'book' seems also to reflect a constraint on the sequence -lɾVl.

This constraint does not explain all instances where PAN *-N is reflected as Puyuma -n, but it does help to explain why we find mixed reflexes of *-[e]N in the decade terms: *ma-ka-telru-n* 'thirty' and *ma-ka-walru-n* 'eighty' reflect the constraint, whereas *ma-ka-pet-el* 'forty' (< *ma-Sepat-eN, table 1) doesn't. Unexpectedly, *ma-ka-nem-en* 'sixty' (< PAN *ma-nem-eN) has -n for -l, while the remaining Nanwang decade terms do not reflect *-N.¹¹ This may reflect an unidentified environment where *-N becomes -n, or it may reflect leveling of the suffix form across the paradigm. Tamalakaw *ma-ka-pitu-l* 'seventy' also reflects *-N as expected (Tsuchida 1980).

Similar considerations apply to the Puyuma reflexes of PAN *paR-X-[e]N 'do X times'. Nanwang *pari-asa-l* 'do once' and *par-telru-n* 'do three times' behave as expected, but *par-pua-n* 'do twice' doesn't. The remaining members of the set lack a reflex of *-[e]N.¹² In

TABLE 2. FORMOSAN REFLEXES OF PAN *n AND *N*

PAN	Puyuma	Rukai	Tsou	Kan.	*Saa.	Pai.	Bunun	Thao	Sai.	Atayal	Seediq	Siraya	Kav.
*n	n	n	n	n	n	n	n	n	n	n	n	n	n
*N	n, l	l	h	n	lh	l	n	z [δ]	l	l	l	l	n

* Kan. = Kanakanavu, Saa. = Saaroa, Pai. = Paiwan, Sai. = Saisiyat, Kav. = Kavalan.

10. Nanwang Puyuma has *ma-ka-walru* 'eighty' where Tamalakaw has *ma-ka-walru-n*.

11. They are *ma-ka-butaʔan* 'twenty' and *ma-ka-luatr* 'fifty' (not reflecting their PAN forebears *ma-puSa-N 'twenty' and *ma-lima-N 'fifty'); and *ma-ka-pitu* 'seventy' and *ma-ka-iwa* 'ninety' (each missing the PAN suffix of *ma-pitu-N and *ma-Siwa-N).

12. They are *par-pat*, *par-luatr*, *par-nem*, *par-pitu*, *par-walu*, and *par-iwa* (Cauquelin 1991).

the Katipul and Tamalakaw paradigms cited by Sagart, leveling has proceeded further. In Tamalakaw, we find *paR-asa-n* ‘do once’, *paR-puwa-n* ‘do twice’, *paR-telru-n* ‘do three times’, *paR-epat-en* ‘do four times’, *paR-nem-en* ‘do six times’, *paR-walru-n* ‘do eight times’, but *paR-pitu* ‘do seven times’ and *paR-iwa* ‘do nine times’, where expected *-l* has been lost. Katipul is similar, except that *par-pitu-n* has also acquired *-n*.

In the light of all this comparative evidence, it is almost inconceivable that Puyuma *-n* in numeral terms should reflect PAN **-en*, as Sagart proposes for ‘do X times’ terms, rather than PAN **-N*. There is, moreover, a further piece of evidence pointing away from **-en*, namely that in Puyuma, like several other Formosan languages (and presumably in PAN), each vowel in a sequence is a syllabic nucleus, and vowels are not deleted at syllable boundaries.¹³ Thus if *par-pua-n* ‘do twice’, and *par-telru-n* ‘do three times’ reflected PAN **paR-puSa-en* and **paR-telu-en*, respectively, their expected outcomes in Nanwang Puyuma would be *par-pua-én* and *par-telru-én*, with default word stress on the final syllable.

We think that the phonological evidence that Puyuma ‘do X times’ terms reflect PAN **paR-X-[e]N*, not PAN **paR-X-en*, is convincing in itself. However, Sagart argued for **paR-X-en* as the source of Puyuma *par-pua-n* ‘do twice’ and *par-telru-n* ‘do three times’ (but not *pari-asa-l* ‘do once’) on the grounds that the Puyuma forms are cognate with forms in Philippine languages. If this were true, then **paR-X-en* ‘do X times’ would be reconstructible to a language ancestral to both Puyuma and the Philippine languages, and perhaps to Proto-Austronesian (PAN). In Sagart’s words, “based on the agreement between Tamalakaw Puyuma and Philippine languages, Blust (n.d.) [1998] reconstructed PAN forms with an **-en* suffix for each [numeral].”¹⁴ But Blust (1998) is more circumspect than this. He reconstructs PAN **paR-* ‘do X times’,¹⁵ but places **paR-X-en*, reconstructed to a not yet determined interstage, in parentheses to indicate doubt, with the meaning ‘divide into X’. The evidence for the latter is found in entries for the numeral forms and consists of entries for Tamalakaw Puyuma and Bikol of southeastern Luzon, shown in table 3.

The meaning of the Bikol terms is not ‘do X times’ but ‘divide into X, send X at a time’, and this was evidently the meaning of Proto-Malayo-Polynesian (PMP) **paR-X-en*: cf. Ilokano *pag-dua-en* ‘do two at the same time, divide into two’ (Rubino 2000). Given this

TABLE 3. BLUST’S PUYUMA AND BIKOL FORMS

PAN		PUYUMA	BIKOL
		‘do X times’	‘divide into X, send X at a time’
<i>*paR-Septat-en</i>	(no gloss given; <i>*Septat</i> ‘four’)	<i>paR-pat-en</i>	<i>pag-apat-ón</i>
<i>*paR-enem-en</i>	‘do six times’	<i>paR-nem-en</i>	<i>pag-anom-ón</i>
<i>*paR-walu-en</i>	‘do eight times’	<i>paR-walru-n</i>	<i>pag-walo-ón</i>

13. This is true at least of Mantaoran Rukai, Kakanavu, Saaroa, Paiwan, Takavatan Bunun, Saisiyat, Kavalan, and Truku Seediq. Note that phonetic vowel sequences reflecting PAN **aw*, **ay*, and **ty* are not counted as sequences of phonological vowels and are tautosyllabic. Only in Thao and Atalay reflexes of PAN **-V-en* is **-e-* deleted; while in Tsou, **-V-* sometimes becomes a consonant.

14. Sagart cites forms for ‘one’, ‘four’, ‘six’, and ‘eight’; Blust (1998) is a work in progress and does not yet include forms for the other numerals.

15. Blust refers the reader to entries under individual numerals, but there does not seem to be evidence to support either the PAN form or its gloss. However, this is irrelevant to present concerns and will doubtless be rectified in the published version.

difference between the meanings of the Puyuma and the Bikol terms, we do not know why Blust gives ‘do six times’ and ‘do eight times’ as PAN glosses, yet glosses *paR-X-en as ‘divide into X’; probably the discrepancy simply reflects the fact that this is a massive work in progress. Certainly there is no clear evidence about the meaning of putative *paR-X-en in a language ancestral to both Puyuma and Bikol, and therefore no clear evidence that the Puyuma and Bikol terms are cognates rather than chance resemblances between Puyuma reflexes of PAN *paR-X-[e]N and Bikol reflexes of PMP *paR-X-en.

The derivational (especially verbal) morphology of both Formosan and Philippine languages is very complex, but the derivational morphology of Philippine languages reflects PMP innovations that render their morphologies rather different from those of Formosan languages, the latter in any case displaying considerable differences among themselves. Much of the complexity lies in prefix–suffix combinations like *paR-X-[e]N and *paR-X-en, in which the same affixes are used over and over again in different permutations. This makes the probability of chance resemblances quite high, and much work on the history of early Austronesian affixation is needed before we can be certain what is cognate with what.

The likelihood that the Puyuma and Bikol affix combinations do not reflect shared inheritance is increased by the fact that there is strong evidence in Blust (1998) for the reconstruction of PMP *paka-X ‘do X times’ (Actor Voice form: *maka-X), implying that it is rather unlikely that PMP *paR-X-en ever meant ‘do X times’.

3. DOES PUYUMA HAVE PERFECTIVE FINITE VERBS CONTAINING *in*?

We turn now to the first of Sagart’s two arguments for the proposition that Puyuma reflects the Nom-to-V innovation.¹⁶ Puyuma, like most Formosan and Philippine languages, has two voices, an actor voice (AV), in Puyuma main clauses always intransitive, and a transitive undergoer voice (UV). The undergoer voice occurs in three versions (UVP, UVL, and UVC), according to whether a patient, a location, or a circumstance (prototypically the instrument) is its subject. The subject is encoded in the nominative case, the actor of a UV verb in the genitive.¹⁷

We noted in section 1 that before the Nom-to-V innovation, **in* was a perfective nominalizer. Sagart claims that the infix *in* sometimes marks a perfective finite UV verb in one Nanwang Puyuma sociolect. If he is right, then this sociolect reflects the Nom-to-V innovation. Sagart’s examples of allegedly finite verbs containing *in* include both transitives, like (8), and intransitives, like (9). The upper row of interlinear glosses and the upper free English gloss are Sagart’s, the lower ones ours.¹⁸

16. Data in this section are from Nanwang Puyuma, but to the best of our knowledge other Puyuma dialects have cognate morphosyntactic constructions.

17. Formosan and Philippine voice systems have been described in various ways. This account follows Teng (2008a), which draws on Himmelmann (2005).

18. Sagart glosses *kana* as genitive, Teng (2008a:50) as oblique, since genitive and oblique noun phrases (but not pronouns) are identically marked in Nanwang Puyuma (Teng 2009). This difference in glossing is irrelevant to both Sagart’s and our arguments.

- (8) Tu=*p<in>auka=ku* kana ragan m-uka i Balangaw.
 3SG.GEN<PRF.UVP>send=1SG.NOM DEF.GEN priest AV-go LOC Taitung
 3SG.GEN<PRF.NMLZ>send=1SG.NOM DEF.OBL priest AV-go LOC Taitung
 ‘The priest sent me to Taitung.’ (Sagart 2009, [1])
 ‘the priest’s sending me to Taitung’
- (9) P<in>iwalak=*ku* i Puyuma.
 <PRF.UVP>have.child=1SG.NOM LOC Puyuma
 <PRF.NMLZ>have.child=1SG.NOM LOC Puyuma
 ‘I was born in Puyuma.’ (Sagart 2009, [3])
 ‘I am one who was born in Puyuma.’

We maintain that these examples represent two different constructions, both containing a nominalization formed with *<in>*.

Sagart’s ground for interpreting the forms with *<in>* in (8) and (9) as finite verbs is the 1SG nominative enclitic =*ku*. Teng (2008a:130) says that, while verbs with non-third person subjects obligatorily take a nominative enclitic, nominalizations in *<in>* never do. Sagart takes this to mean that in examples like (8) and (9), the forms containing *<in>* must be perfective finite verbs.

It is now clear that the statement that nominalizations in *<in>* never take a nominative enclitic is a slight overgeneralization: “slight” because nominalizations in *<in>* with a nominative enclitic like those in (9) and (10) do indeed occur, but are very rare in natural data. None are found in Teng’s five hours of natural text, and very few in Cauquelin’s 188 pages of interlinearized text (2008:137–325). Sagart cites two instances from Cauquelin (2008), at her sentences 15-13 and 16-34, both intransitive. Two more instances occur in 15-15.¹⁹ Transitive instances do not occur at all in Cauquelin’s texts.

The form *p<in>auka* in (8) is the nominalized UVP form of the verb *pauka* ‘send’, =*ku* is the pronominal enclitic referencing the undergoer (‘me’), and *tu=* encodes the possessor of a nominal (‘his sending ...’). Teng’s consultants insist that *tu=* can be replaced by the free possessor pronoun *nantu*, as in (10). This replacement is a crucial test of whether the following head is nominal or verbal, as *nantu* only co-occurs with a nominal head (Teng 2008a:49).

- (10) Nantu p<in>auka=*ku* kana ragan m-uka i Balangaw.
 3.POSS <PRF.NMLZ>send=1SG.NOM DEF.OBL priest AV-go LOC Taitung
 ‘the priest’s sending me to Taitung’

The construction in (8) is thus a noun phrase meaning ‘the priest’s sending me to Taitung’; that is, it is a nominalization of the clause shown in (11), which contains a perfective finite verb. In (11) the genitive proclitic references the actor of the verb (still ‘s/he’ = ‘the priest’) and cannot be replaced by *nantu* (**nantu pauka-aw=ku=la*).

- (11) Tu=*pauka-aw=ku=la* kana ragan m-uka i Balangaw.
 3SG.GEN=send-UVP=1SG.NOM=PRF DEF.OBL priest AV-go LOC Taitung
 ‘The priest sent me to Taitung.’

19. We count only forms in which *<in>* is infixated into the first syllable of the verb stem (which may itself be a prefix). We do not count forms like *mi-s<in>abung-an=ku* ‘I bring an offering’ (6-22), a verb composed of the derivational prefix *mi-* ‘have, carry’ and the noun *s<in>abung-an* ‘offering’, itself a nominalization of the verb *sabung* ‘make an offering’.

We turn now to the intransitive construction exemplified in (9). Unlike (8), the construction in (9) is a full clause. It is, however, an intransitive clause with a nominal predicate, and means (semi-literally) ‘I am one who was born in Puyuma’. The construction is the same as the clause in (12), where the predicate is the noun phrase *a lalak* ‘a child’. The enclitic =*ku* is the subject of both (9) and (12).

- (12) A lalak=**ku**.
INDF.NOM child=1SG.NOM
 ‘I am a child.’

The predicate in (9), *p<in>iwalak*, appears superficially not to be a noun phrase. However, Teng’s consultants feel that it is “incomplete” and that the example should preferably read *a p<in>iwalak=ku i Puyuma*, the indefinite nominative determiner *a* confirming that the predicate is indeed a noun phrase. Thus in (9) the syntactic host of the subject enclitic =*ku* is not *p<in>iwalak*, but the predicate noun phrase *a p<in>iwalak*, as described in Teng (2008a:31–32).²⁰

There is another reason why it is very unlikely that the predicate of (9) is a verb. There is no dispute over the fact that (9) is intransitive, but intransitive verbs in Puyuma have AV forms in <*em*> or one of its variants (Teng 2008a: 47–48, ch.8).²¹ If *p<in>iwalak* were a verb, then, under Sagart’s assumption that the Nom-to-V innovation had occurred at a stage ancestral to all Formosan languages, we would expect it to resemble in function the verb forms reflecting *<*in*> in other Formosan languages, that is, the Nuclear Austronesian languages of Taiwan. These forms are UVP, and therefore transitive, verbs (and indeed Sagart glosses *p<in>iwalak* as UVP). But (9) not only clearly has intransitive meaning; it also lacks the second, genitive actor, argument like *tu=* in (11) that a transitive verb requires.

In his table 1, labeled “UV markers in the Nanwang Puyuma sociolect studied by Josiane Cauquelin (simplified),” Sagart (2010:198) provides the affix paradigm shown here as in table 4. However, this is not an adequate analysis of either Cauquelin’s or Teng’s data. Neither Cauquelin (1991:46, 50–51, 2008:10–17, 24) nor Teng (2008a:130–31) analyzes forms in <*ininan* (henceforth “<*in22 and that the perfective aspect of a finite verb is marked by attaching the enclitic =*la*, illustrated in (11), to a neutral verb (Cauquelin 1991:47, Teng 2008a:117–18). Sagart does not tell us how the allegedly finite <*inla*.*

TABLE 4. SAGART’S TABLE 1

	UVP	UVL	UVC
Neutral	V-aw	V-ay	V-anay
Perfective	< <i>in <td>?</td> <td><<i>in </i></td></i>	?	< <i>in </i>

20. Sagart gives another similar example, his (4):

- (i) P<*in
 <PRF>have.child-UVC=1SG.NOM=already*

His gloss is ‘I have already given birth’, but its meaning is ‘I am one that has already been born’.

21. The variants and <*em*>, *me-*, *m-*, *ma-*, and zero: the choice of variant depends largely on the class of the verb (Teng 2008a:108, 120–23; Ross 2009:297, table 2).
 22. Neither Cauquelin nor Teng assumes that <*inan* encodes a UVC nominalization. There has been a great deal of lexicalization, and the division of labor between forms with and without -*an* defies straightforward analysis.

The very infrequency of *<in>*-forms with a nominative enclitic in Cauquelin's texts (four instances, as noted above) indicates that these are not, as Sagart claims, perfective finite verbs. However, since there is no third person enclitic (i.e., the third person nominative enclitic is zero), we also searched the texts for *<in>*-forms without a nominative enclitic that under his hypothesis might be finite verbs because they occur as predicates. Even in languages that have undergone the Nom-to-V innovation, numerous instances of **<in>*V and **<in>*V-*an* reflexes are nominalizations, and under Sagart's hypothesis we would expect the same to be true of Puyuma. We therefore ignored all *<in>*-forms that are preceded by a determiner or a free possessor pronoun, as these are unambiguously nominals, and also forms that the sentence structure and Cauquelin's translation require to be nominals (many of these are lexicalized and crop up quite often) or relative clauses (which are nominalizations by another name: Teng 2008a:101; Ross 2009:208–309). This left us with just one instance of an *<in>*-form as a predicate, shown in (13).²³

- (13) K<in>a-sagar dra birua dra kar-pali-pali.
 <PRF.NMLZ>ka-love INDF.OBL spirit INDF.OBL together-RED-have.evil.eye
 'He was one loved by the spirits who together have the evil eye.'
 (Cauquelin 2008:4-09; glosses ours)

There is, in fact, a constructionally similar example in Sagart's paper, his (6), reglossed and shown here as (14). However, Sagart analyzed =*ku* as an enclitic to the verb, whereas grammaticality requires *ku*= 'my' as a possessor proclitic to *rumaq* 'house'.

- (14) K<in>ibulras-an **ku**=rumaq.
 <PRF>borrow-NMLZ 1SG.GEN=house
 'My house was a borrowed one.'

In total, then, there are just five examples of predicate nominalizations in Cauquelin's texts, four with a nominative enclitic and one without. Such a low frequency speaks against these being finite perfective verbs. On the other hand, there are 113 instances of perfective =*la* in Cauquelin's published texts, a figure that is in line with Cauquelin's and Teng's analysis of this as the default perfective marker.²⁴ Predicate nominalizations like those in (9), (13), and (14) were doubtless the early Austronesian precursors of the Nom-to-V innovation, and it is easy to see how reanalysis could take place, but all the evidence says that reanalysis has not taken place in Puyuma and that these are still simply predicate nominalizations, and as such have a rather low token frequency.²⁵

Sagart claims that there are grammatical differences between Catholic and non-Catholic (Protestant and unconverted) sociolects of the Nanwang dialect of Puyuma, and that this

23. In *k<in>a-sagar*, *ka-* is a formative occurring in certain Puyuma verb classes (Ross 2009, table 2). Historically it marked statives, but it has no consistent semantic function in modern Puyuma. In the verb *kar-pali-pali*, the form *pali* means 'supernatural power'.

24. Based on a visual count of the printed texts (an electronic version is not available). Cauquelin writes *la* as a separate word and glosses it 'Asp' ('aspect'), but its postpredicate position guarantees that it is perfective =*la*. Its distribution across Cauquelin's texts is patchy: this is probably the result of varying discourse types.

25. We note that Kaufman (2009a, b) has recently proposed that Tagalog predicates (other than imperative forms in one dialect) are still nominal. If he is right, then the Proto-Nuclear Austronesian innovation was that the construction in (9), (13), and (14) was generalized to become the default predicative construction. The reanalysis of nominal predicates as verbs occurred later and independently in various daughter-languages.

may account for a difference between *in*-forms with nominative enclitics (in the Catholic sociolect) and their absence in the non-Catholic sociolect. The claimed difference is said by Cauquelin to be supported by the use of the portmanteau proclitic pronoun *kumu*- ‘1SG.GEN + 2SG.NOM’ among Catholics and its absence among non-Catholics, reported in an unpublished paper by Cauquelin. However, Teng has done fieldwork at Nanwang since 1995 and worked on Puyuma for about eleven of the intervening years, collecting data from Catholic, Protestant, and unconverted speakers, and does not see evidence of these sociolectal differences. The portmanteau form had not been previously reported by Cauquelin, and was first reported in Nanwang in Teng’s (2008b) survey of portmanteau proclitic pronouns across Puyuma dialects.²⁶ Sagart also claims that differences between Cauquelin’s and Teng’s data are due to different collection methods. It is not clear how this difference is related to the sociolectal difference, but we have shown above that the apparent difference with regard to *in*-forms is the result of incorrect analyses in Sagart’s paper.

4. WHAT HAPPENED TO PUYUMA? Sagart ends with a hypothesis about how Puyuma could have undergone the Nom-to-V innovation, yet retained none of its nominalizers (table 5) in finite verb forms. He suggests that the Nom-to-V innovation was followed by a two-stage process. First, the new finite verbs—formed with nominalizers—were replaced by periphrastic forms consisting of a preverb (an auxiliary) and a dependent verb formed with one of the UV affixes *-aw*, *-ay*, or *-anay*. Second, preverbs were lost, leaving the forms in *-aw*, *-ay*, and *-anay* as new finite forms. His evidence for the first stage is analogical: he claims that the same thing occurred in Tsou (a proposal rejected by Ross 2009). His evidence for the second stage is that actor (genitive) pronouns are procliticized to the verb in Puyuma. Clitic pronouns in Formosan languages are typically enclitics, but Starosta, Pawley, and Reid (1981, 1982) suggested that, with preverb loss, a pronoun encliticized to the preverb is left stranded and attaches itself to the following verb.

There is a significant snag in the evidence for the first stage. If finite verbs formed with nominalizers were replaced by a preverb + dependent verb construction, then we would expect the dependent verbs in that construction to have the forms shown in the bottom row of table 5. These are the forms that occur in Tsou, which has allegedly undergone the same innovation, and their reflexes, tabulated in appendix B of Ross (2009), indicate strongly that they formed the PAN dependent paradigm: they follow the negative preverb in Puyuma and Mayrinax Atayal, and are dependent (that is, follow a preverb) in Tsou, Saisiyat, Seediq, and Paiwan (they are also imperatives in some languages). Sagart’s hypothesis, however, asks us instead to accept that the pre-Puyuma dependent forms were ancestral to the Puyuma UV suffixes *-aw*, *-ay*, and *-anay*. These reflect the forms in the Finite Realis and Optative/hortative rows of table 5, and their reflexes in appendix B of Ross (2009) are *never* dependent verbs. In languages other than Puyuma they are optative/hortative (Mayrinax Atayal, Seediq, Thao, Paiwan, Haiyan Amis) or imperative (Kanakanavu, Ishbukun Bunun), but not dependent (they are also not future/

26. The presence of a larger set of portmanteau double pronouns in Tamalakaw Puyuma was noted by Tsuchida (1980). There is, moreover, a difference between Nanwang and Tamalakaw forms. Nanwang speakers use *kunu*- and *kanu*- interchangeably as ‘1SG.GEN + 2SG.NOM’, with a preference for *kanu*-, whereas Tamalakaw, as reported by Tsuchida, distinguishes between *kanu*- ‘1SG.GEN + 2SG.NOM’ and *kunu*- ‘2SG.GEN + 1SG.NOM’ (Teng 2008b).

irrealis, even though Sagart labels them thus). The presence in Puyuma of *-aw*, *-ay*, and *-anay* as both optative/hortative and realis affixes is more plausibly accounted for under the Nuclear Austronesian hypothesis of Ross (2009), according to which these forms were displaced in their realis function in languages other than Puyuma, Tsou, and Rukai by the nominalizing forms **-en*, **-an*, and **Si-*.

The evidence for Sagart's second stage, the procliticization of pronouns to the verb, is circumstantial. The history of pronoun cliticization remains puzzling, for several reasons. It seems likely that PAN had only one set of pronominal clitics (Ross 2006), which later diverged into genitive and nominative sets in various languages. In a majority of Formosan languages, pronouns are enclitics to the verb (or to the preverb if there is one), but in Paiwan and Puyuma, genitive (actor-of-transitive) pronouns are proclitics, and in the Maga and Tona dialects of Rukai, subject pronouns are proclitics. One is tempted to ask if this is an areal feature.²⁷ In other Puyuma dialects, however, there is a set of portmanteau proclitics to transitive verbs: they consist diachronically of a genitive–nominative clitic sequence, a phenomenon that is unique to Puyuma (Teng 2008b). In these circumstances, we think that the jury is still out on the history of Puyuma clitic pronouns and that no verdict can be safely based on their position.

5. REFLEXES OF PAN *Sa-? Sagart himself notes a potential consequence of his hypothesis: if indicative UVP forms reflecting PAN **-en* have been eliminated from Puyuma morphology but have left fossil traces, then we might also expect to find traces of PAN **-an* UVL and **Sa-* or **Si-* UVC. Sagart discusses in his footnote 13 the possibility that verbal **Sa-* is reflected in nominals like *a-iray-an* 'a mill' (*iray* 'to grind'), *a-imeŋ-an* 'a cell, detention room' (*imeŋ-u* 'to keep in custody'), and *a-kan-an* 'food' (*kan* 'eat') (Cauquelin 1991). His tentative argument is that, for example, a UVC verb form *a-iray*, presumably 'grind with s.t.', has been nominalized with *-an* to form *a-iray-an*. However, these nominalizations reflect the *Ca-...-an* pattern described by Teng (2008a:136–37) and require no further explanation. The prefix *a-* is the allomorph of *Ca-* that occurs with vowel-initial stems (Teng 2008a:41–42); in Puyuma the stem 'eat' is *ekan*, and hence also vowel-initial.

TABLE 5. PAN VERBAL AFFIXES*
(Excerpted from Ross 2009, table 1)

		AV	UVP	UVL	UVC
Nominalizers	Realis	*M-STEM	*STEM-en	*STEM-an	*Si-STEM
	Perfective realis	*M-⟨in⟩STEM	*⟨in⟩STEM	*⟨in⟩STEM-an	*S⟨in⟩i-STEM
Finite verbs	Realis	*M-STEM	*STEM-aw	*STEM-ay	*an-aySTEM
	Optative/hortative	*M-STEM-a	*STEM-aw	*STEM-ay	*an-aySTEM
	Dependent	*M-STEM	*STEM-a/i	*STEM-i	*an-iSTEM

* Notes: PAN **M-* took three forms: **um*, **ma-*, or zero (Ross 2009:297). On Wulai Atayal evidence, **an-ay* and **an-i* are reconstructed as preverbs, but their Puyuma, Mayrinax Atayal, and Seediq reflexes are suffixes.

27. It is almost the same as the area in which *ki-* passives occur (Zeitoun and Teng 2009). Other evidence for contact-induced change in Puyuma is provided by Paiwan-to-Puyuma and Puyuma-to-Paiwan lexical loans, and by the presence of the Rukai possessor suffix *-li* 'my' in certain Puyuma kin terms.

In Puyuma dialects other than Nanwang, there is a UVC nominalizer, but it is *i-* (< PAN *Si-), not *a-* (Ross 2009:304).

6. CONCLUSION. Sagart concludes, “what is certain is that a language ancestral to Puyuma had both UVP markers *-en (neutral) and *⟨in⟩ (perfective): this is sufficient to falsify Ross’s theory.” We hope we have shown that this is not certain. Indeed, it is wrong. The alleged fossilized reflex of *-en is a reflex of PAN *-[e]N and the supposed finite verbs in *⟨in⟩ are still nominalizations. Sagart’s article thus provides no evidence that Puyuma has undergone the Nom-to-V innovation, and therefore no evidence against its status as a primary branch of Austronesian.

REFERENCES

- Adelaar, K. Alexander. 2008. Siraya structure. Unpublished MS, University of Melbourne.
- Blust, Robert A. 1998. Austronesian comparative dictionary. Unpublished electronic files, University of Hawai‘i.
- . 2001. Malayo-Polynesian: New stones in the wall. *Oceanic Linguistics* 40:151–55.
- . 2003. *Thao dictionary*. Language and Linguistics Monograph Series A5. Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica.
- Cauquelin, Josiane. 1991. The Puyuma language. *Bijdragen tot de Taal-, Land- en Volkenkunde* 147:17–60.
- . 2008. *Ritual texts of the last traditional practitioners of Nanwang Puyuma*. Language and Linguistics Monograph Series A23. Taipei: Institute of Linguistics, Academia Sinica.
- De Busser, Rik L. J. 2009. Towards a grammar of Takivatan Bunun: Selected topics. PhD thesis, La Trobe University.
- Egli, Hans. 1990. *Paiwangrammatik*. Wiesbaden: Otto Harrassowitz.
- . 2002. *Paiwan Wörterbuch: Paiwan–Deutsch, Deutsch–Paiwan*. Wiesbaden: Otto Harrassowitz.
- Ferrell, Raleigh. 1982. *Paiwan dictionary*. Canberra: Pacific Linguistics.
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In *The Austronesian languages of Asia and Madagascar*, ed. by K. Alexander Adelaar and Nikolaus P. Himmelman, 110–81. London: Routledge.
- Ho, Dah-an 何大安. 1997. Kānākānāfū yū 卡那卡那富語 [The Kanakanavu language]. In *高雄縣南島語言 [The Formosan languages of Kaohsiung county]*, ed. by Paul Jen-kuei Li 李壬癸, 228–71. Kaohsiung: Kaohsiung County Government.
- Kaufman, Daniel. 2009a. Austronesian typology and the nominalist hypothesis. In *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, ed. by Alexander Adelaar and Andrew Pawley, 197–226. Canberra: Pacific Linguistics.
- . 2009b. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35:1–49.
- Li, Paul Jen-kuei. 1978. A comparative vocabulary of Saisiyat dialects. *Bulletin of the Institute of History and Philology* 49:133–99.
- . 2006. Numerals in Formosan languages. *Oceanic Linguistics* 45:133–52.
- Li, Paul Jen-kuei, and Shigeru Tsuchida. 2006. *Kavalan dictionary*. Taipei: Institute of Linguistics, Academia Sinica.

- Ross, Malcolm. 2006. Reconstructing the case-marking and personal pronoun systems of Proto Austronesian. In *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuai Li on his 70th birthday*, ed. by Henry Yungli Chang, Lillian M. Huang, and Dah-an Ho, 521–64. Language and Linguistics Monograph Series W5. Taipei: Institute of Linguistics, Academia Sinica.
- . 2009. Proto Austronesian verbal morphology: A reappraisal. In *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, ed. by Alexander Adelaar and Andrew Pawley, 295–326. Canberra: Pacific Linguistics.
- Rubino, Carl R. G. 2000. *Ilocano grammar and dictionary*. Honolulu: University of Hawai'i Press.
- Sagart, Laurent. 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43:411–44.
- . 2009. PAN morphology in phylogenetic perspective. Paper presented to the panel “Reconstructing the morphology of proto-Austronesian” at the Eleventh International Conference on Austronesian Linguistics, Aussois, France, June.
- . 2010. Is Puyuma a primary branch of Austronesian? *Oceanic Linguistics* 49:194–204.
- Starosta, Stanley, Andrew Pawley, and Lawrence A. Reid. 1981. The evolution of focus in Austronesian. Paper presented at the Third International Conference on Austronesian Linguistics, Bali.
- . 1982. The evolution of focus in Austronesian. In *Papers from the Third International Conference on Austronesian Linguistics*, vol. 2: *Tracking the travellers*, ed. by Amran Halim, Lois Carrington, and S. A. Wurm, 145–70. Canberra: Pacific Linguistics.
- Szakos, József. 1994. Die Sprache der Cou: Untersuchungen zur Synchronie einer austronesischen Sprache auf Taiwan. PhD thesis, Rheinische Friedrich-Wilhelms-Universität zu Bonn.
- Teng, Stacy Fang-ching. 2008a. *A reference grammar of Puyuma, an Austronesian language of Taiwan*. Canberra: Pacific Linguistics.
- . 2008b. The internal relationships among the Puyuma dialects: Evidence from morphosyntax. Paper presented at the Workshop on Classification and Dispersal of the Austronesians: Anthropological, Archaeological, Genetic and Linguistic Studies Relating to Taiwan, Academia Sinica, Taipei, Jan 3.
- . 2009. Case syncretism in Puyuma. *Language and Linguistics* 10:819–43.
- Ting, Pang-Hsin 丁邦新, 1978. Gū-Bēinán yǔ de nǐcè 古卑南語的擬測 [Reconstruction of Proto-Puyuma phonology]. *Bulletin of the Institute of History and Philology, Academia Sinica* 49:321–92.
- Tsuchida, Shigeru. 1976. *Reconstruction of Proto-Tsouic phonology*. Studies of Languages and Cultures of Asia and Africa, Monograph Series 5. Tokyo: Institute for the Study of Languages and Cultures of Asia and Africa.
- . 1980. Puyuma-go (Tamarakao hōgen) goi: Fu gohōgaisetsu oyobi tekisuto [Puyuma (Tamalakaw dialect) vocabulary: With grammatical notes and texts]. In *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.
- Winter, Bodo. 2010. A note on the higher phylogeny of Austronesian. *Oceanic Linguistics* 49:283–88.
- Zeitoun, Elizabeth. 2007. *A grammar of Mantaoran (Rukai)*. Language and Linguistics Monograph Series A4–2. Taipei: Institute of Linguistics, Academia Sinica.
- Zeitoun, Elizabeth, and Lillian M. Huang. 2000. Concerning *ka-*, an overlooked marker of verbal derivation in Formosan languages. *Oceanic Linguistics* 39:391–414.
- Zeitoun, Elizabeth, and Stacy F. Teng. 2009. From *ki-N* ‘get N’ in Formosan languages to *ki-V* ‘get V-ed’ (passive) in Rukai, Paiwan and Puyuma. In *Discovering history*

- through language: Papers in honour of Malcolm Ross*, ed. by Bethwyn Evans, 479–500. Canberra: Pacific Linguistics.
- Zeitoun, Elizabeth, Stacy F. Teng, and Raleigh Ferrell. 2010. Reconstruction of ‘2’ in PAN and related issues. Manuscript submitted to *Language and Linguistics*.