

Diachronic development of the Nuclear Austronesian locative and patient applicatives

Edith ALDRIDGE

Academia Sinica

This paper proposes reconstructions and diachronic developments of the primary patient and locative voice (PV and LV, respectively) suffixes found in the Formosan and Philippine languages. The LV suffix grammaticalized from a noun meaning ‘place’ and formed nominalized relative clauses in Proto-Austronesian (PAN) on both theme/patient and locative positions. The PV suffix *-en was innovated in a daughter of PAN called Proto-Ergative Austronesian in order to express telic events in nominalized relative clauses. The centrality of the direct object in determining the boundedness of an event is what ensured that *-en would develop as a PV marker and not be extended to other voices. The reanalysis of the nominalizations as verbal matrix clauses in Proto-Nuclear Austronesian (Ross 2009) gave birth to the specialization of *-en and *-an as PV and LV, respectively. Given that *-en could only be used in bounded events, it replaced *-an in this environment, relegating the latter primarily to LV clause types.

1. Introduction¹

Like most other Formosan and Philippine languages, Tagalog exhibits a type of ergative alignment commonly referred to as a “voice system”. The term “voice” refers to the fact that different verbal affixes reflect which argument has absolutive/nominative case. In the perfective aspect, the infix <um> indicates that the subject has nominative case, as in (1a). This is referred to as “actor voice” (AV). The perfective aspect marker changes to <in> when nominative case appears on an internal argument in an ergative clause, i.e., in “non-actor voice” (NAV). When <in> appears with no additional voice marking on the verb in a monotransitive clause, the direct object has nominative case, as in (1b). This is referred to as “patient voice” (PV). When a PV clause is nonfinite or expresses a prospective event, the suffix -in appears instead of the infix, as in (1c). The applicative -an signals that the nominative argument is a goal or locative constituent, as in (1d). This is the “locative voice” (LV). The applicative i- is associated with a nominative instrument, beneficiary, or moved theme. (1e) shows an example with a moved theme. This is referred to as “circumstantial voice” (CV).

Tagalog

- (1) a. *B<um>ili ang babae ng isda.*
 <AV.PFV>buy NOM woman GEN fish

‘The woman bought (a) fish.’

¹ Unless otherwise cited, the Formosan language data are taken from my fieldnotes. Fieldwork on Rukai was supported by grants from the Chiang Ching-kuo Foundation for International Scholarly Exchange (JS015A-12), the University of Washington Nostrand Endowment, and the University of Washington Department of Linguistics. I am also indebted to the native speakers who supplied the data itself. I would also like to thank the organizers and participants at the 14th International Austronesian and Papuan Languages & Linguistics Conference organized by the Leibniz Centre General Linguistics and Humboldt University for feedback on the presentation version of this paper. Finally, I am grateful to two anonymous reviewers for comments on the first draft of this paper.

- b. *B<in>ili* *ng* *babae* *ang* *isda*.
 <NAV.PFV>buy GEN woman NOM fish
 ‘The woman bought the fish.’
- c. *Bi-bilh-in* *ng* *babae* *ang* *isda*.
 RED-buy-PV GEN woman NOM fish
 ‘The woman will buy the fish.’
- d. *B<in>igy-an* *ng* *babae* *ng* *isda* *ang* *lalaki*.
 <NAV.PFV>give-LV GEN woman GEN fish NOM man
 ‘The woman gave the man a fish.’
- e. *I-b<in>igay* *ng* *babae* *ang* *isda* *sa* *lalaki*.
 CV-<NAV.PFV>give GEN woman NOM fish to man
 ‘The woman gave the fish to the man.’

Because of the existence of more than one transitive clause type (AV and NAV), Tagalog and other Philippine languages are often characterized as having a “symmetrical” voice system (Himmelman 2005 and others). Crucially, this type of voice does not involve alternations in argument structure, external and internal arguments continuing to function as core arguments, regardless of case marking. Given that both AV and NAV are active and potentially transitive, the symmetrical voice approach substantively characterizes these languages as split-ergative, AV clauses being aligned as nominative/accusative and NAV as ergative/absolutive. For simplicity of exposition, I refer to this type of alignment as a “voice system” but point out here that symmetrical voice is effectively a type of split-ergativity.

This paper proposes diachronic origins for the PV and LV suffixes seen in (1c) and (1d), respectively. Contra other reconstructions of Proto-Austronesian (PAn) morphosyntax, I do not attribute a voice system to PAn. Consequently, the proto-forms of these affixes were not voice markers in PAn but rather developed diachronically from morphemes performing different functions. I propose that the LV marker *-an* be reconstructed as a nominalizer, grammaticalizing from a noun meaning ‘place’, which could also form relative clauses on VP-internal positions in PAn. As for the PV suffix *-in*, this is generally reconstructed with the form **-en*, /e/ a schwa, and attributed to PAn as a PV marker. However, I propose that this affix was a later development, emerging in a daughter of PAn called “Proto-Ergative Austronesian” (PEAn). Like **-an*, **-en* was also not a voice marker but rather contributed temporal information, surfacing in relative clauses in order to indicate the telicity of a prospective event. Its use solely in patient voice was due to the semantic and syntactic contribution of theme direct objects to telicity, given the prominence of the theme argument in this clause type.

Since **-an* formed relative clauses on a variety of non-subject gap positions in PAn, including theme/patient direct objects, its function overlapped with **-en* in terms of voice. I propose that the principle difference between them was aspectual. I propose that **-en* developed specifically in bounded events, while **-an* was neutral with respect to telicity. But after the emergence of **-en*, **-an* came to be restricted to unbounded events or relative clauses with locative gaps. This contrast between **-en* and **-an* in terms of aspect is retained in a number of Formosan languages, as I discuss in section 4. The specialization of **-en* and **-an* as PV and LV markers, respectively, continued to develop with the reanalysis of nominalized relative clauses as matrix verbal clauses in Proto-Nuclear Austronesian.

The next section provides undergirding for the roles of these two morphemes in nominalized relative clauses by arguing that all NAV relative clauses in PAn and PEAn were nominalized. I also argue for the lack of a voice system in PAn and the existence of the Ergative Austronesian subgroup, where the voice system was first innovated. Section 3 presents the analysis of *-an as a VP nominalizer in PAn, while section 4 argues for the later development of *-en as a marker of telicity in prospective events in PEAn.

2. High-order subgrouping and the nominal-to-ergative reanalysis

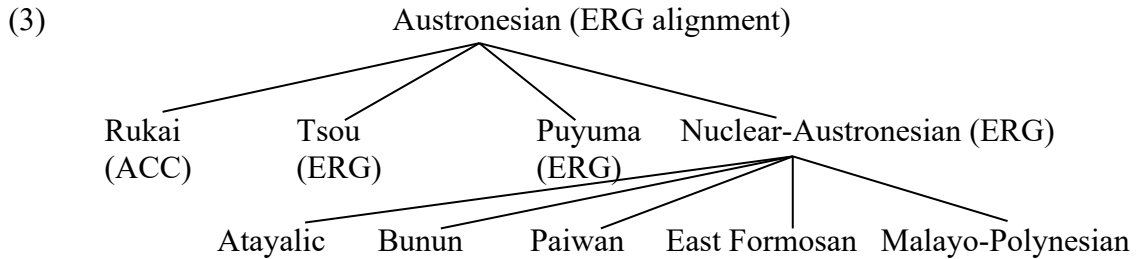
In sections 3 and 4, respectively, I argue that the LV and PV affixes developed diachronically from functional morphemes employed in nominalized relative clauses. This proposal is grounded within the Starosta et al. (1982) position that a subset of the voice affixes found in Philippine and Formosan languages have a nominal origin. This position was subsequently adopted by Ross (2009, 2012), who identifies the reanalysis of nominalized relative clauses as verbal matrix clauses as the defining innovation of the “Nuclear Austronesian” (NucAn) subgroup, which I discuss below. I adopt the NucAn hypothesis in this paper, but I argue against the widely-held view that PAn itself should be reconstructed with a voice system and follow instead my (2016, 2021a) position that PAn had accusative alignment, while the voice system was first innovated in Proto-Ergative Austronesian (PEAn). In section 4, I propose that development of the PV suffix *-en also be attributed to PEAn. In the current section, I concentrate on the nominal origins for the LV and PV affixes, as well as the lack of a voice system in PAn.

The classic reconstruction of Austronesian voice is generally credited to Wolff (1973), who posited a PAn voice system along the following lines.² The Tagalog examples in (1) exhibit reflexes of these morphemes. A form of this paradigm has been widely adopted by Austronesian historical linguistics, including Ross (1995), Blust (1999), Blust & Trussel (ongoing), and others. In this paper, I concentrate on the PV and LV markers. Note in particular the restriction of *-en to non-perfective aspect. This contrast between *-en and *<in> is also retained in Tagalog, as discussed in the previous section.

(2)		<u>AV</u>	<u>PV</u>	<u>LV</u>	<u>CV</u>
	Non-perfective	*M-	*-en	*V-an	*Si-
	Perfective	*<in>M-	*<in>	*<in>...-an	* Si-...<in>

Starosta et al. (1982) proposed that these affixes were employed in nominalizations in either PAn or Pre-PAn, only later being reanalyzed as verbal affixes. Ross (2009, 2012) shows that Puyuma, Rukai, and Tsou exhibit no evidence of the use of these affixes in verbal contexts; in Puyuma and Rukai they are clearly only found in nominal clauses, while in Tsou they have been lost altogether. Consequently, Ross separates these languages from a large subgroup called “Nuclear Austronesian” (NucAn) and proposes the reanalysis of embedded nominalizations as verbal matrix clauses as the defining innovation of this subgroup.

²This is only a partial paradigm, which does not include the non-indicative mood voice markers. Some of the forms also differ slightly from Wolff’s reconstruction, most closely resembling the version put forth by Ross (1995). *M- had three allomorphs, one of which was *.



However, Ross continues to maintain that PAN had a voice system, a type of ergative alignment, in both verbal (V) and nominal (N) environments, given that Rukai is the only extra-NucAn language with accusative alignment. The verbal realis paradigm is based primarily on evidence from Puyuma, but similar affixes are also found in Tsou. NucAn languages also retain this paradigm, but in irrealis contexts only. The NAV voice markers shown for Tagalog in (1) were used only in nominal environments in PAN.

(4)

	<u>AV</u>	<u>PV</u>	<u>LV</u>	<u>CV</u>
Non-perfective (N)	*M-	*-en	*-an	*Si-
Perfective (N)	*<in>M-	*<in>	*<in>...-an	*Si...<in>
Realis (V)	*M-	*-aw	*-ay	*-an-ay

I begin the discussion in this section with the evidence for positing the NucAn subgroup. In order to demonstrate that reanalysis of embedded nominalizations is the defining innovation of this subgroup, it is necessary to show that extra-NucAn languages retain this clause type only as nominalizations. Puyuma is such a language, matrix PV verbs in taking the suffix *-aw*. Note that (5a) is a matrix PV clause in which the theme argument has nominative case. The affix *-aw* never appears on a verb in a nominalization. In the nominalized relative clause in (5b), the verb takes the perfective aspect marker *<in>*, which only appears in nominalizations, and the nominalizer *-an*.

Nanwang Puyuma

- (5) a. *tu=trakaw-aw na paisu kan isaw*
 3.GEN=steal-PV DEF.NOM money SG.OBL Isaw
 ‘Isaw stole the money.’ (Teng 2008:147)
- b. *ala amuna sadru [[tu=tr<in>ekelr-an] na asi]*
 maybe because many 3=<PFV>drink-NMLZ DEF.NOM milk
 ‘Maybe because the milk he drank is a lot.’ (Teng 2008:105)

(5b) actually bears strong surface resemblance to the LV construction in Tagalog and other NucAn languages, though the argument extracted in (5b) is the theme and not a location.

Tagalog

- (6) *B<in>igy-an ng babae ng isda ang lalaki.*
 <NAV.PFV>give-LV GEN woman GEN fish NOM man
 ‘The woman gave the man a fish.’

In the following section, I argue that the LV marker developed historically from the nominalizer **-an* reflected in Puyuma (5b). In PAN, **-an* could be associated with either theme or locative gaps, while it developed into an LV marker in Proto-NucAn (PNucAn). NucAn languages also retain the perfective infix, but it is used in both embedded and matrix clauses, while it surfaces only in nominalizations in Puyuma.

Rukai exhibits a similar dichotomy between verbal and nominal clause types. The onglide in the realis prefix *wa-* shown in the verbal clause in (7a) is cognate with PAN *M-. The nominalized clause in (7b) shows the same nominalizing suffix as Puyuma *-an* and perfective aspect infix *<in>*.³ As in Puyuma, nominalization is required when the embedded object is the head of a relative clause. Li (1973) refers to fronting constructions like those in (7b) as “emphatic”. My consultants translate examples of this type as clefts, with the clause-initial NP in focus.

Tanan Rukai (Li 1973:108–109)

- (7) a. *ku lacing wa-baay naku-a sa lrima ka 'aysu.*
 NOM Lacing REAL-give 1SG-ACC OBL five LK money
 ‘Lacing gave me five dollars.’
- b. *kay 'aysu b<in>aay-an naku-a ina marudrang.*
 this money <PFV>give-NMLZ 1SG-ACC that old.man
 ‘This money was given to me by that old man.’

Ross (2009, 2012) proposes that the nominal voice markers were reanalyzed as verbal and replaced the erstwhile realis affixes in PNucAn. This accounts for the fact that these affixes now appear on verbal matrix verbs in NucAn languages like Tagalog. However, Ross’ assumption that this reanalysis replaced earlier verbal voice markers is problematic. First, Ross offers no motivation for the loss of the earlier affixes. Secondly, there is evidence against positing a replacement. Teng & Zeitoun (2016) report that the reanalysis of nominalizations as matrix clauses is taking place incrementally in Kanakanavu. Kanakanavu shows a two-way voice distinction between actor and patient voice in matrix clauses. The PV clause type employs erstwhile nominalizing morphology, *<in>* in the perfective aspect and *-un* (< *-en) in future contexts. The theme or patient argument has nominative case, while the external argument has genitive case. Teng & Zeitoun (2016) propose that PV clauses have completed the transition from nominal to verbal.

Kanakanavu (Teng & Zeitoun 2016:138)

- (8) a. *c<in>apa=maku 'alam.*
 <NAV.PFV>roast=1SG.GEN meat.NOM
 ‘I roasted meat.’
- b. *te:=maku cakup-un ca:u i:sa.*
 IPFV=1SG.GEN.AG stab-IPFV.PV person that
 ‘I will stab that person.’

However, the two applicative constructions (LV and CV) are still employed only in nominalized relative clauses. The following pair of examples show the nominalizer employed in a locative and theme relative clause, respectively, which are both formed on *-an*.

³ According to my own fieldnotes, *<in>* appears to have been lost in clausal nominalizations in Tanan, though it can be found in some lexical nominalizations.

Kanakanavu (Teng & Zeitoun 2016)

- (9) a. *cikiringa cakuran=ia, [ni-pe-pacal-an-in vavulu].* (p. 145)
 side.river=TOP PFV-CAUS-die-NMLZ-3.GEN wild.pig

‘As for the riverside, it is the place where he killed wild pigs.’
 (lit. “As for the riverside, (it) his pig-killing place.”)

- b. *sua [ni-kalu'-a(n)=maku=ia] 'a:cu ni-ara-[a]ka.* (p. 146)
 NOM PFV-like/love-NMLZ=1SG.GEN=TOP PFV-INCH-bad

‘As for my lover, s/he is dead.’ (lit. “As for the one I loved, s/he is dead.”)

The following shows that the CV applicative likewise only surfaces in nominalizations.

Kanakanavu (Teng & Zeitoun 2016:145)

- (10) *ka:lu i:si=ia si-po 'ocipi-in 'u:ru*
 wood this=TOP INS.NMLZ-cook-3SG.GEN cooked.rice

‘As for the wood, (it) was her rice-cooking instrument.’

The preceding examples from Kanakanavu show that the reanalysis from nominal to verbal clause in NucAn languages is currently ongoing, at least in this language. More importantly, Kanakanavu does not show evidence of ever having had the realis verbal voice affixes proposed by Ross (2009) in (4).⁴ The relevant portion of the Kanakanavu paradigm is given in (11). Kanakanavu does not have LV and CV affixes in verbal contexts, though Ross (2009, 2012) predicts that it should have reflexes of *-ay and *-an-ay, respectively.

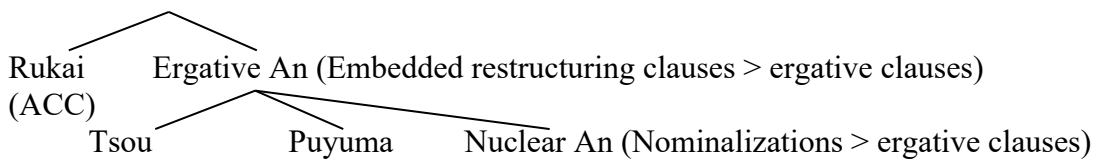
(11)	<u>AV</u>	<u>PV</u>	<u>LV</u>	<u>CV</u>
Non-perfective (N)	M-	-a(n)	ta...-a(n)	si-
Perfective (N)	<in>M-	<in>...-a(n)	<in>...-a(n)	si-
Non-perfective (V)	M-	-un	---	---
Perfective (V)	<in>M-	<in>	---	---

(Teng & Zeitoun 2016:138–139)

Kanakanavu, then, provides evidence against reconstructing a voice system for verbal clauses in PAN. Given both the lack of evidence and the lack of motivation for the replacement of earlier voice affixes in PNucAn, I assume the simpler approach in Aldridge (2016, 2021a) that these affixes simply did not exist in PAN. I therefore distinguish Rukai from the other languages on the basis of its inherited accusative alignment. In contrast, the Ergative Austronesian languages have innovated the ergative-type voice system. According to Aldridge (2016, 2021a), the realis voice affixes reconstructed by Ross (2009, 2012) first appeared in embedded clause types and were later extended to matrix clauses in Puyuma. Tsou retains them in embedded environments. I briefly sketch the innovations of these voice markers in section 4.

⁴ Saaroa, another language analyzed by Teng & Zeitoun (2016), does partially reflect these affixes. However, this does not constitute evidence for Ross’s (2009) reconstruction, since the preverbal placement of clitic pronouns strongly suggests an embedded origin for these clause types, as predicted by Aldridge (2016, 2021a). Starosta et al. (1982) propose that proclitic, as opposed to enclitic, pronouns are found in Austronesian languages in which embedded clauses have been reanalyzed as matrix clauses after deletion of the verb in the higher clause.

(12) Austronesian (ACC alignment)



On this view, PAn did not have a voice system in matrix clauses, though two (LV and CV) of the modern voice markers were employed in relative clauses in order to extract different arguments, as shown in (13). In this paper, I focus on theme and locative relative clauses and propose that both types were projected by the relativizing morpheme *-an. Evidence for this proposal is presented in section 3. In the remainder of this section, I present another argument for reconstructing PAn without a matrix voice system. I do not discuss the origin of the CV prefixal applicative *Si- in this paper but refer the reader to Peterson (1997, 2007) and Aldridge (to appear) for proposals for how this applicative developed from a verb meaning “to carry, to wear, to have” in a serial verb construction.

(13)		<u>AV</u>	<u>PV</u>	<u>LV</u>	<u>CV</u>
	Nominalization	---	*-an	*-an	*Si-
	Verbal	*M-	---	---	---

According to Aldridge (2021a, 2021b), relative clause formation in the Extra-NucAn languages provides additional evidence for this higher subgrouping hypothesis, specifically the proposal that PAn had accusative alignment rather than a voice system. Philippine and Formosan languages have a constraint on movement such that the nominative NP is free to move, but other NPs cannot undergo movement in clauses that have a nominative NP. For example, the theme object in a PV clause in Tagalog can become the head of a relative clause, as in (14a). But this is not possible for the genitive subject, as in (14b).

Tagalog

- (14) a. *isda-ng b<in>ili ng babae*
 fish-LK <NAV.PFV>buy GEN woman
 ‘fish that the woman bought’
- b. **babae-ng b<in>ili ang isda*
 woman-ng <NAV.PFV>buy NOM fish
 ‘woman who bought the fish’

The inverse is true in AV clauses where the subject has nominative case. Here, the nominative subject undergoes dislocation, as in (15a), but the genitive theme in (15b) is not eligible.

Tagalog

- (15) a. *babae-ng b<um>ili ng isda*
 woman-LK <AV.PFV>buy GEN fish
 ‘woman who bought a/the fish’
- b. **isda-ng b<um>ili ang babae*
 fish-LK <AV.PFV>buy NOM woman
 ‘fish that the woman bought’

Rukai clearly has the same constraint. As shown in (16), the subject can become the head of a relative clause. In contrast to object relative clauses, subject relative clauses in most Rukai dialects are verbal, containing a reflex of *M- in non-future tense or aspect. The verbal status of the embedded verb is indicated by the fact that the reflex of *M- surfaces in matrix clauses, as well as relative clauses.

Tona Rukai

- (16) a. *kusi'a ka wa-thenay ki tatay namia*
 yesterday TOP REAL-sing NOM father 1PL.INC
 ‘Our father sang yesterday.’
- b. *nani-ini [kudrai wa-the-thenay]?*
 who-3SG.GEN that REAL-RED-sing
 ‘Who is that one who is singing?’

In the next pair of examples, (17a) shows future marking in a declarative clause, while (17b) shows a corresponding relative clause. This future marker never surfaces in non-subject relative clauses, which are nominalized and cannot contain tense marking.

Tanan Rukai

- (17) a. *ludha ay-kela ku tina=li*
 tomorrow FUT-come NOM mother=1SG.GEN
 ‘My mom will come tomorrow.’
- b. *[kuadra ay-suwawaw] ka mukabarubarua*
 DEM FUT-clean TOP girl
 ‘The one who will clean is the girl.’

Specifically, an object cannot undergo movement in a verbal clause, which contains a nominative subject. Consequently, object relative clauses are nominalized, as shown above in (7b). An additional example is given below in (18). In addition to the nominalizer *-ane*, the verb also takes the imperfective prefix *a-* rather than future tense *ay-*. This asymmetry between verbal and nominal clauses is easily understood in terms of the Austronesian nominative extraction constraint. In short, the object cannot move when there is a subject with nominative case, i.e., in a matrix clause. This is why object relative clauses in Rukai must be nominalized. The subject in a nominalization has genitive case, which allows the object to move over it.

Tanan Rukai

- (18) *manema [kuadra a-senag-ane=ini]?*
 what DEM IPFV-sing-NMLZ=3SG.GEN
 ‘What is it that he/she will sing?’

The same is true in Puyuma, as noted by Teng (2008). Puyuma exhibits an ergative-type voice system like Tagalog. The reflex of PAN *M- marks AV verbs, as in the simple intransitive declarative clause in (19a). AV also surfaces in subject relative clauses, as shown in (19b).

Nanwang Puyuma (Teng 2008:135)

- (19) a. *ta-ka-kesi=ku*
 <AV>-RED-study=1SG.NOM
 ‘I am studying.’
- b. *a [ta-ka-kesi] =ku*
 INDEF.NOM <AV>-RED-study =1SG.NOM
 ‘I am a student.’ (lit: ‘I am one who studies.’)

However, object relative clauses must be nominalized, as in (20b). (20a) shows a verbal PV clause marked with *-aw* for contrast. As discussed above, *-aw* never surfaces in relative clauses. Returning to the point at hand, the requirement that Puyuma object relative clauses be nominalized is not predicted by the nominative extraction constraint. This is because the object in a PV matrix clause has nominative case, so it should be able to undergo movement, as shown for Tagalog in (14a).

Nanwang Puyuma

- (20) a. *tu=trakaw-aw na paisu kan isaw*
 3.GEN=steal-PV DEF.NOM money SG.OBL Isaw
 ‘Isaw stole the money.’ (Teng 2008:147)
- b. *ala amuna sadru [[tu=tr<in>ekelr-an] na asi]*
 maybe because many 3=<PFV>drink-NMLZ DEF.NOM milk
 ‘Maybe because the milk he drank is a lot.’ (Teng 2008:105)

Tsou also demonstrates that this is the case. Tsou has lost the clausal nominalizing morphology found in Puyuma and Rukai. Objects can undergo relativization directly in verbal clauses, as shown in (21b) for Tsou. It is clear that the embedded clause in (21b) is verbal, due to the presence of the same NAV auxiliary verb and transitive PV morphology as in the main clause in (21a).

Tsou

- (21) a. *i-ta teaph-a to kexpx ta pasuya 'e cxyx*
 NAV-3SG put.into-PV OBL backpack ERG PN ABS lunch.box
 ‘Pasuya put the lunch box into his backpack.’ (Chang 2011:282)
- b. *cuma na [i-he papas-a]*
 what ABS NAV-3PL cut-PV
 Lit. ‘What are the things they are cutting?’
 ‘What are they cutting?’ (Chang 2011:301)

Ross’ (2009, 2012) reconstruction of PAN with a voice system is unable to explain why Puyuma non-subject relative clauses must be nominalized. This is because the object in a transitive clause like (20a) has nominative case and consequently should be able to undergo movement. On the other hand, the nominalization requirement can be accounted for if PAN is reconstructed as an accusative language. PAN can then be analyzed like Rukai. As an accusative language, subjects in verbal PAN clauses regularly surfaced with nominative case, which prevented an object from moving in this context. Object relative clauses were then required to be nominalized in order to mark

the subject with genitive case so the object could move over it. Nominalized object relative clauses in Puyuma are then simply a retention from accusative PAn.⁵

To summarize the preceding discussion, PAn employed nominalizations in order to form relative clauses on non-subject positions, and this was a consequence of the subject movement constraint. In fact, it is not uncommon for languages with accusative alignment to employ different types of relative clause – either nominal or participial – in order to extract non-subjects as opposed to subjects. This fact provides indirect support for the reconstruction of PAn with accusative alignment, since nominalized object relative clauses are typically found in accusative languages. Languages of this sort include the Uto-Aztecan language Yaqui (González 2012), Cuzco Quechua (Muysken 2011, Cole & Hermon 2011, Hastings 2004), and most Altaic languages, like Turkish (Hankamer & Knecht 1976; Dede 1978; Kornfilt 1997, 2008; Aygen 2002; Cagri 2005). For example, finite declarative clauses in Altaic languages like Turkish exhibit accusative alignment with nominative subjects and accusative objects.

Turkish (Kornfilt 2007:309)

- (22) a. *Ali smav-ı geç-ecek.*
 Ali.NOM test-ACC pass-FUT
 ‘Ali will pass the test.’
- b. *Sen smav-ı geç-ecek-sin.*
 2SG.NOM test-ACC pass-FUT-2SG
 ‘You will pass the test.’

But relative clauses in these languages exhibit an asymmetry similar to Austronesian languages in which verbal inflection alternates depending on whether the head nominal is the embedded subject or not. In subject relatives the inflection is *-an*, but this changes to *-duğ* when a non-subject is extracted. Note further that the embedded subject has genitive case when another nominal has moved over it.

Turkish (Cagri 2005:6)

- (23) a. [*divan-da otur-an*] *bayan*
 sofa-LOC sit-SR lady
 ‘the lady who is sitting on the sofa’
- b. [*bayan-ın otur-duğ-u*] *divan*
 lady-GEN sit-NSR-3SG sofa
 ‘the sofa that the lady is sitting on’

⁵ An anonymous reviewer questions the use of morphosyntactic evidence for the purposes of subgrouping, an assumption commonly made by historical linguists. However, it is important to keep in mind the fact that arguments for subgrouping are based on establishing direction of change, and this in turn is based on identifying natural processes of change. I proposed natural changes accounting for the emergence of the voice system in PEAn in Aldridge (2016, 2021a). Additional proposals that direction of change can be established for diachronic syntax can be found in Harris & Campbell (1995), Gildea (1998), Roberts & Roussou (2003), Willis (2011), Barðdal & Eythórsson (2012), and Walkden (2013, 2014). Given that syntax (like phonology) is a rule-based linguistic system, both synchronic and diachronic variation are expected to be systematic and predictable. Consequently, as the field of diachronic syntax develops, it is anticipated that more and more evidence will be brought to bear on the naturalness and predictability of syntactic change undergirding its applicability in establishing direction of change.

A similar pattern is found in Cuzco Quechua. Subjects and objects surface with nominative and accusative case, respectively, in declarative clauses.

Cuzco Quechua (Hastings 2004:10)

- (24) *Juan waka-ta ranti-rqa-n.*
 Juan.NOM cow-ACC buy-PST-3SG

‘Juan bought a cow.’

But relative clauses display the same type of subject/non-subject asymmetry as in Turkish. Verbs in subject relatives are marked with with *-q*, as in (25a), while non-subject relatives show *-sqa* on the embedded verb, as in (25b). The subject in (25b) also surfaces with genitive case.

Cuzco Quechua (Hastings 2004:58)

- (25) a. [*wayna waka ranti-q*]
 boy cow buy-NMLZ(SUBJ)
 ‘the boy who bought the cow’
- b. [*wayna-q waka ranti-sqa-n*]
 boy-GEN cow buy-NMLZ(PST/NONSUBJ)-3SG

‘the cow which the boy bought’

The Uto-Aztecan language Yaqui likewise has accusative alignment in declarative clauses.

Yaqui (González 2012:71)

- (26) *U yoeme uka kari-ta jinu-k*
 DET man.NOM DET.ACC house-ACC buy-PFV

‘The man bought the house.’

But relative clauses display a three-way opposition reminiscent of an Austronesian-type voice system. Subject extraction is marked with *V-me*, as in (27a); object extraction is marked with *V-‘u*, as in (27b); and locative extraction shown in (27c) correlates with a different marker *V-‘epo/‘apo*.

Yaqui

- (27) a. *U yoeme [kari-ta jinu-ka-me] ousi tom-ek*
 DET man.NOM house-ACC buy-PFV-REL a.lot.of money-POSS
 ‘The man who bought the house has a lot of money.’ (González 2012:72)
- b. *U bisikleeta [in jinu-ka-‘u] sikili*
 DET bicycle 1SG.GEN buy-PFV-REL red
 ‘The bicycle that I bought is red.’ (González 2012:73)
- c. *Wa kari [nim bo‘e-pea-‘apo] ujyooli*
 DEM house 1SG.GEN sleep-DES-REL pretty
 ‘The house that I want to sleep in is pretty.’ (González 2012:78)

González (2012:84) argues that the locative relativizer derives from the locative adposition *-po*. Consequently, *-po* can be analyzed as an applicative, which affords argument status to a locative constituent.

To summarize this section, I have argued that PAn be reconstructed as an accusative language. But I agree with Ross (2009, 2012) that it had a dichotomy between verbal and nominal clausal projections, and nominalized relative clauses were reanalyzed as verbal matrix clauses in Proto-Nuclear Austronesian. In the next two sections, I propose diachronic origins in nominalized relative clauses for the two morphemes that developed into the PV and LV voice markers after the nominal-to-verbal reanalysis in PNucAn. In section 3, I discuss the morpheme which derived theme/patient and locative relative clauses in PAn, which I identify as *-an. *-an became the LV marker in PNucAn. I propose in section 4 that PV *-en was a later development, emerging first in Proto-Ergative Austronesian before becoming the PV marker in PNucAn.

3. Origin of the Nuclear Austronesian LV marker *-an

One conclusion of the previous section is that PAn did not have a voice system, per se, but it did employ different verbal morphology for forming subject and non-subject relative clauses. Specifically, subjects could be extracted in verbal clauses, because they have nominative case. But non-subjects could only undergo movement in nominalizations, and this is because the nominalization provided genitive case for the subject, which is what allowed the object to move over it. This section discusses the nominalizing affix *-an, which projected theme and locative relative clauses in PAn. Demonstrating that *-an was not only associated with locative gaps additionally constitutes evidence against reconstructing PAn with a voice system, since this *-an could not have been a voice marker. This proposal stands in stark contrast to Ross (2009, 2012), who assumes that *-an functioned only as a locative nominalizer in PAn, while PV *-en was used for theme extractions. The function of *-an as a VP nominalizer on both theme and locative positions is evidenced by the distributions of its reflexes in extra-Nuclear Austronesian (NucAn) languages.

Table 1. Reflexes of *-an in Extra-Nuclear Austronesian languages

	NMLZ	PV relative clause	LV relative clause	Source
Rukai	YES	YES	YES	Author's fieldnotes
Puyuma	YES	(YES)	YES	Teng (2008)
Tsou	YES	NO	NO	Tung (1964)
Sa'aroa	YES	NO	YES	Pan (2012)
				Teng & Zeitoun (2016)
Kanakanavu	?	YES	YES	Teng & Zeitoun (2016)

As can be seen in Table 1, *-an is found consistently in lexical nominalizations in all extra-NucAn languages. Kanakanavu is marked with “?” because I have no data at hand. Reflexes of *-an are also used widely to form relative clauses on both theme and locative positions. The distributions of *-an and *-en are subject to dialect variation in Puyuma, as I show in section 4. Tsou has lost the nominalized relative clauses inherited from PAn, though it reflects *-an in lexical nominalizations. As in Tsou, Sa'aroa has also extended the verbal PV marker to relative clauses with theme gaps. Given that the PV affix *-a* in both Tsou and Sa'aroa was innovated in Proto-Ergative Austronesian (PEAn), as I discuss in section 4, its use in relative clauses is likewise the result of an

innovation, with the result that only LV relatives continue to be nominalized with a reflex of *-an, which is *-ana* in this language.

In contrast to *-an, *-en is extremely rarely attested outside of NucAn languages, as shown in Table 2. The sources are the same as in Table 1, so I do not repeat these.

Table 2. Reflexes of *-en in Extra-Nuclear Austronesian languages

	Nominalization	PV relative clause	LV relative clause
Rukai	NO	NO	NO
Puyuma	(YES) ⁶	(YES)	NO
Tsou	NO	NO	NO
Sa'aroa	NO	NO	NO
Kanakanavu	NO	NO	NO

Only certain dialects of Puyuma reflect *-en in nominalizations. Otherwise, *-en is not reflected in extra-NucAn languages except in Kanakanavu, where it has only a verbal use, as I discuss below. The wide distribution of *-an and the relative paucity of *-en suggests strongly that only the former should be attributed to PAn. For this reason, I reconstruct *-en to PEAn and not to PAn, as I discuss in section 4. This temporal disconnect between the two affixes, as well as the wide employment of *-an in both PV and LV relative clauses, further argues against reconstructing them as voice markers. This is because voice markers would be expected to have originated as members of a voice paradigm in their respective PV and LV functions in the same synchronic stage of the languages' development. Consequently, I argue in this section for the reconstruction of *-an as a VP nominalizer which could be associated with both theme and locative gaps in PAn. In section 4, I propose that the original function of *-en was aspectual and not to mark voice.

In this section, I present evidence for reconstructing *-an as a VP nominalizer, which plausibly grammaticalized from a noun meaning 'place'. This is the function of its reflexes in lexical nominalizations in Formosan languages. In the following examples, a reflex of *-an attaches to a noun or verb and creates a noun referring to a place. The Tsou example in (28) names the place associated with a particular clan.

Tsou (Tung 1964:175)

(28) *luhtu-ána* (from *luhtu* (one branch of Tsou) + *-an > place name)

In Kavalan, the reflex of *-an attaches to a verb to create a locative noun.

Kavalan (Chang & Lee 2002:363)

(29) *yau uzis-an muzis*
is bathe-NMLZ bathe

'(He) is bathing in the bathroom.'

The Rukai examples in (30) are also locative nouns, created by attachment of the reflex of *-an to another noun.

⁶ Some dialects use *-en* to form theme nominalizations, but these are always accompanied by reduplication expressing aspectual information, as I discuss in section 4.

Mantauran Rukai (Zeitoun 2007:209)

- (30) a. *acilalr-ae* ‘pond’ < *acilai* ‘water’ + *-an
 b. *avai-nae* ‘bride’s family’ < *avai* ‘female’ + *-an

The preceding examples of lexical nominalizations are mostly taken from extra-NucAn languages, but this use of *-an is also found consistently in NucAn languages, as exemplified by Kavalan in (29).

In PAn clausal nominalizations, *-an formed relative clauses on a variety of gap positions. (31) and (32) show relative clauses formed on the embedded theme argument in Nanwang Puyuma.

Nanwang Puyuma (Teng 2008:136)

- (31) a. *senay* ‘sing’
 b. *s<in>enay-an*
 <PFV>sing-NMLZ
 ‘songs sung’
 c. *sa-senay-an*
 RED-sing-NMLZ
 ‘songs to be sung’

Nanwang Puyuma (Teng 2008:105)

- (32) *ala amuna sadru* [[*tu=tr<in>ekelr-an*] *na asi*]
 maybe because many 3.PRS=<PFV>drink-NMLZ DEF.NOM milk
 ‘Maybe because the milk he drank is a lot.’

Nanwang *-an* can also create nominals or relative clauses on locative gaps.

Nanwang Puyuma (Teng 2008:138)

- (33) a. *alup* ‘hunt’ *alup-an* ‘hunting ground’
 b. *takesi* ‘study’ *takesi-an* ‘school’

(34) shows that *-an* can create a gapless relative clause, presumably relativizing on the event variable.

Nanwang Puyuma (Teng 2008:142)

- (34) *kadru* [*ku=k<in>a-sagar-an*] *dra suan*
 <INTR>there 1SG.GEN=<PRV>KA-like-NMLZ OBL dog
 ‘My loving of dogs is like that.’

Blust & Chen (2017) propose that employment of *-an* to produce relative clauses on theme gaps in Nanwang nominalizations is the result of the loss of the reflex of *-en. This may in fact be true for Nanwang PV relative clauses, given that the Katripul dialect has a reflex of *-en used as a PV relativizer, as I discuss in the following section. However, loss of *-en does not account for the employment of *-an in forming gapless relative clauses. Nor does it account for the use of reflexes of *-an to extract themes in languages that also have a reflex of *-en, which is the case in many NucAn languages. For example, Central Amis does not employ a reflex of *-en in forming relative clauses but rather uses the reflex of *-an, as shown in (35a). Amis uses a reflex of *-en as a PV marker in matrix clauses, as shown in (35b). Consequently, even though this language

employs *-en for PV in verbal contexts, *-an is still used to form relative clauses on theme positions.

Central Amis

- (35) a. *ma-olah kako to-ya*
 IPFV.STAT-like NOM.1SG ACC-that
[mi-asip-an ni Panay inacila] a codad.
 IPFV.AV-read-OREL GEN Panay yesterday LK book
 ‘I like that book that Panay read yesterday.’ (T. Chen 2018:272)
- b. *faedet-en ni Panay ko-ya dateng i honi.*
 hot-PV GEN Panay NOM-that dish P moment
 ‘Panay heated that dish just now.’ (T. Chen 2018:48)

Another such language is the Atayalic language Seediq. Like Amis, a reflex of *-en marks PV in verbal contexts, as shown in (36a). But relative clauses produced on theme gaps are frequently marked with -an, as in (36b). Marking with -un (< *-en), as in (36c), is far less common.

Seediq

- (36) a. *wada bube-un na Pihu ka dangi=na*
 PFV hit-PV GEN Pihu NOM friend=3SG.GEN
 ‘Pihu hit his friend.’
- b. *laqi b<n>be-an na Pihu-ni laqi alang itsin*
 child <PFV>hit-LV GEN Pihu-DEF child village other
 ‘The child that Pihu hit is a child from another village.’
- c. *laqi wada bube-un na Pihu-ni sio sapah=na*
 child PFV hit-PV GEN Pihu-DEF next.to house=3SG.GEN
 ‘The child that Pihu hit is his next door neighbor.’

Kanakanavu, classified by Teng & Zeitoun (2016) as an extra-NucAn language, is like Amis, only -an being employed in relative clauses, which are nominalized in this language. *-en is also employed in this language but is found exclusively in verbal contexts. As discussed in section 2, Zeitoun & Teng (2016) propose that Kanakanavu reflects an intermediate stage in the reanalysis of nominalizations as verbal clauses. The Kanakanavu reflex of *-en marks PV in matrix clauses expressing imperfective aspect like (37a). But object relative clauses are formed on the reflex of *-an, as in (37b).

Kanakanavu (Zeitoun & Teng (2016))

- (37) a. *te:=maku cakup-un ca:u i:sa.* (p. 138)
 IPFV=1SG.GEN.AG stab-IPFV.PV person that
 ‘I will stab that person.’
- b. *sua [ni-kalu'-a(n)=maku=ia] 'a:cu ni-ara-[a]ka.* (p. 146)
 NOM PFV-like/love-NMLZ=1SG.GEN=TOP PFV-INCH-bad
 ‘As for my lover, s/he is dead.’ (lit. “As for the one I loved, s/he is dead.”)

These findings are completely consistent with my proposal that *-an projected nominalized relative clauses with theme or locative gaps in PAn and was not a voice marker. Amis, Seediq, and Kanakanavu can then be understood as retaining this property of *-an from PAn. In particular, the fact that a reflex of *-an is found in theme, as well as locative, nominalizations supports my proposal that the distinction between *-en and *-an did not involve voice.

Relatedly, De Busser (2009) shows that *-an* functions productively as a lexical nominalizer in Takivatan Bunun. Most commonly, *-an* forms nominalizations on locations or times, which is expected given its origin as a noun meaning ‘place’. The following examples show locations.

Takivatan Bunun (De Busser 2009: 254)

- (38) a. *a-sabaq-an*
STAT-sleep-NMLZ
‘a spot in the forest where a human is sleeping’
- b. *a-lukmuʔ-an*
STAT-squat.down-NMLZ
‘a spot in the forest where an animal is resting’

But there are also nominalizations formed on abstract results. This usage suggests a more general use for *-an as a nominalizer.

Takivatan Bunun (De Busser 2009: 254)

- (39) a. *bahiʔ-an*
to.dream-NMLZ
‘dream’
- b. *in-liskin-an*
VIA-think-NMLZ
‘thoughts’

De Busser points out that the reflex of *-en is far less productive in nominalizations. It can be used to form theme nominalizations, but he also notes that these are not lexical nominalizations, suggesting that their formation might be more recent than the examples with *-an* seen in (38) and (39).

Takivatan Bunun (De Busser 2009: 263)

- (40) a. *kuð-kuða-un*
RED-work-PV
‘things that need to be done’
- b. *siða-un*
take-PV
‘things that have to be taken’

Jeng (1977) mentions that some verbs in the Takbanuadh dialect of Bunun can mark PV clauses with both *-un* (< *-en) and *-an*. Although he does not attribute any difference in functionality to the two suffixes, it is clear that a reflex of *-an can surface in PV clauses in this language, supporting my reconstruction of *-an as being associated with both theme and location gaps.

Takbanuadh Bunun (Jeng 1977: 75)

- (41) a. *simaq* 'ampuk-**an** 'ista?
 who gather-LV him
 'Whom is he gathering?'
- b. *simaq* 'ampuk-**un** 'ista?
 who gather-PV him
 'Whom is he gathering?'

Reflexes of *-an are also used widely as nominalizers in Malayo-Polynesian languages. (42) shows examples of theme nominalizations in Standard Indonesian.

Indonesian (Sneddon 1996:31)

- (42) a. *tulis-an*
 to.write-NMLZ
 'writing'
- b. *tanam-an*
 to.plant-NMLZ
 'plant'
- c. *makan-an*
 to.eat-NMLZ
 'food'

Theme nominalizations can also be produced from adjective bases.

Indonesian (Sneddon 1996:34)

- (43) a. *asam-an*
 sour-NMLZ
 'pickles'
- b. *kotor-an*
 dirty-NMLZ
 'garbage, excrement'

In Indonesian, -an can also create nominalizations on actions.

Indonesian (Sneddon 1996:32)

- (44) a. *larang-an*
 prohibit-NMLZ
 'prohibition'
- b. *tembak-an*
 shoot-NMLZ
 'shooting'

Locative nominalizations also occur.

Indonesian (Sneddon 1996:32)

- (45) a. *giling-an*
to.mill-NMLZ
'mill'
- b. *bendung-an*
to.dam-NMLZ
'dam'

Lynch et al. (2002:70) additionally reconstruct *-an as a general nominalizer in Proto-Oceanic, as reflected in the following examples. (46a) shows an action nominalization, while (46b) exemplifies a locative nominalization.

- (46) a. *mate-a* (Vitu; Lynch et al. 2002:70)
die-NMLZ
'death'
- b. *habo-habotu-ana* (Roviana; Corston-Oliver 2002:472)
RED-sit-NMLZ
'chair'

In short, reflexes of *-an are widely attested as a nominalizer/relativizer for theme, locative, and event variable gaps, though reflexes of *-en are only rarely found in nominalizations. This wide association with different argument positions also argues against reconstructing *-an specifically as a locative voice marker. It further bears noting that a noun meaning 'place' is known to have grammaticalized into a relative clause forming morpheme in other languages as well. The non-subject relativizer in Late Archaic Chinese (LAC; 5th–3rd centuries BCE) was employed in the same range of functions as in Nanwang Puyuma. This morpheme *suo* was also a noun meaning 'place' and formed relative clauses on theme (47a), locative (47b), and event variable (47c) gaps.

- (47) a. 其所言者特未定 (Zhuangzi, Qiwulun)
[qi suo yan zhe] te wei ding.
they PLACE say DET but not uniform
'[What they have to say] is not uniform.'
- b. 其北陵, 文王之所避風雨也。 (Zuozhuan, Xi 32)
Qi bei ling [Wen Wang zhi suo [bi feng yu]] ye.
3.GEN north hill Wen king GEN PLACE escape wind rain STAT
'The north hill is [where the (Zhou) king Wen took shelter from the storm].'
- c. 所不與舅氏同心者, 有如白水。 (Zuozhuan, Wen 13)
[Suo bu yu jiu shi tong xin zhe] you ru bai shui.
PLACE NEG with uncle same mind COND be as white water
'If (I) am not loyal to my uncle, then let it be as the white water.'

Another language exemplifying the use of a 'place' noun as a relativizer is the Uralic language East Khanty (Potanina 2008). (48a) shows this noun functioning as a locative

relativizer, while (48b, c) demonstrate that it has also grammaticalized into a general nominalizer.

- (48) a. *kul wel-tā-l tayi-j-a wān tayı kit'-əs*
 fish kill-NPP-3SG place-EP-LAT near place exist-PST.3SG
 ‘That place where he was fishing was not far.’ (Potanina 2008:79)
- b. *lopəl-tə jə-min tayi*
 angle-NPP become-CONV place
 ‘fishing’ (= “place that has become fishing”) (Potanina 2008:80)
- c. *mā il-əlintā-l-əm iy-nə mən-t*
 1SG down-lie-PRES-1SG bear-LOC 1SG-ACC
nu-li-tə tayi əntə wu-t-am-a
 up-eat-NPP place NEG see-IMP.P-1SG-ILLAT
 ‘I lay down (so as) not to see the bear ripping me apart.’
 (= “the bear’s place ripping me apart”) (Potanina 2008:80)

Canaanite languages also employ a reflex of *ʔaθr ‘place’ as a relative particle (Wilson-Wright 2019, Holmstedt 2002, and references therein). (49) shows a relative clause formed on a theme gap in Biblical Hebrew.

- (49) *ʿal miškābī ballēlōt biqqaštī ʿēt še ʿāh^abā napšī*
 upon bed.my in.the.nights seek.1CS.PFV ACC REL-Ø-love.3FS.PFV soul.my
 ‘upon my bed at night I seek **who** (=the man that) my soul loves’ (Song 3.1; Holmstedt 2002:10)

This cross linguistic comparison enhances the plausibility of reconstructing *-an as a general relativizer grammaticalizing from a noun meaning ‘place’, given that a parallel process is found in multiple unrelated languages.

In this section, I argued that *-an be reconstructed to PAN as a nominalizer forming both lexical nominalizations and relative clauses that could be associated with a variety of non-subject gap positions. Nominalization of relative clauses formed on non-subject gaps was required in order to allow these arguments to undergo movement over the subject, which was not possible in verbal clauses, where the subject had nominative case. This introduces the question of why the CV applicative *Si- was needed in order to extract other pseudo-arguments like instruments, moved themes, or beneficiaries. This is because this functional category was necessary in order to introduce these arguments into the derivation, as per the proposal put forth by Pylkkanen (2002) for “high” applicatives. These applicatives are generated outside of the verb phrase and introduce an argument into the event as a whole. In other words, these arguments cannot be selected directly by the lexical verb but require the applicative in order to enter the derivation. This situation is illustrated by light verb/applicatives selecting transported themes or instruments in West African languages (Lord 1993, Sebba 1987, and others). This is the light verb *de* in Twi, which Lord (1993) argues grammaticalized from a verb meaning ‘hold/have/possess/own’. The point here is that the applied object is selected by the light verb and not the lexical verb. This is the type of construction which Peterson (1997, 2007) proposes to be the origin of the Austronesian CV applicative *Si-.

Twi (Lord 1993)

- (50) a. *o-de afoa ce boha-m*
 he-DE sword put scabbard-inside
 ‘He put the sword into the scabbard.’ (p.66)
- b. *o-de enkrante tya duabasa*
 he-DE sword cut branch
 ‘He cut off a branch with a sword.’ (p.67)

In contrast, *-an, which nominalized only the verb phrase, could produce relative clauses only on VP-internal gap positions, particularly themes and locations. These two functions of *-an were later differentiated into LV marked by *-an and PV marked by *-en in Proto-Nuclear Austronesian. The next section explores the development of *-en in Proto-Ergative Austronesian and proposes an explanation for the subsequent emergence of the division of labor between *-an and *-en.

4. Origin of patient and locative voice in NucAn languages

This section addresses the origin of the PV marker *-en and also the question of how the nominalizer *-an came to function primarily as the LV marker in verbal clauses in Proto-Nuclear Austronesian (PNucAn). This division of labor between PV and LV is illustrated once again by the Tagalog examples below. The reflex of *-en occurs when a theme or patient nominal has nominative case, while *-an indicates that a goal or locative argument is the nominative constituent.

- (51) a. *Bi-bilh-in ng babae ang isda.*
 RED-buy-PV GEN woman NOM fish
 ‘The woman will buy the fish.’
- b. *B<in>ilh-an ng babae ng isda ang tindahan=ko.*
 <NAV.PFV>buy-LVGEN woman GEN fish NOM store=1SG.GEN
 ‘The woman bought a/the fish at my store.’

I propose that the development of PV and LV markers in NucAn languages was a consequence of event structure. Specifically, *-en originated as a marker of telic aspect in prospective events in nominal relative clauses. The semantic and syntactic contribution of theme direct objects in expressing bounded events is what ensured that *-en surfaces only in PV clauses. In contrast, *-an did not contribute aspectual information in PAN but came to be associated with unbounded events and LV clauses after the development of *-en in telic PV clause types. Preliminary evidence for the aspectual difference between PV and LV can be seen in the following Tagalog examples. Specifically, the voice distinction between *-in* and *-an* can be nullified exceptionally, and either one can appear when the theme NP has nominative case. However, the former (52a) indicates that the object is fully affected by the event, while the latter (52b) induces a partitive interpretation.

- (52) a. *Ka-kain-in=ko ang isda.*
 RED-eat-PV=1SG.GEN NOM fish
 ‘I will eat (up) the fish.’

- b. *K<in>ain-an=ko ang isda.*
 <NAV.PFV>eat-LV=1SG.GEN NOM fish
 ‘I ate some of the fish.’

Such a contrast between *-an and *-en is reflected even more clearly in multiple Formosan languages. In Northern Amis, both PV *-en* and LV *-an* can be used with nominative themes, but the event is bounded and the object is more affected with *-en*, as in (53a). A partitive interpretation is induced by the presence of *-an* in (53b).

Northern Amis (Bril 2022:46)

- (53) a. *Kaen-en=tu k-iya buting.*
 eat-PV=PFV NOM-DEM fish
 ‘The fish has been eaten up.’ (“completely eaten”)
- b. *Kaen-an=tu k-iya buting.*
 eat-LV=PFV NOM-DEM fish
 ‘The fish has been eaten.’ (“may or may not be fully eaten”)

A similar alternation is observed in Mayrinax Atayal. The reflex of *-en in this language is *-un*, which is used in a bounded event. In contrast, a partitive interpretation obtains when the verb is suffixed with LV *-an*.

Marinax Atayal (Peng 2016:116)

- (54) a. *pag-hulaqiy-un ni Tiwas ku’ qusia’ ka’ hani.*
 become-ice-PV GEN Tiwas NOM.SPEC water LK this
 ‘Tiwas froze all of the water.’
- b. *pag-hulaqiy-an ni Tiwas ku’ qusia’ ka’ hani.*
 become-ice-LV GEN Tiwas NOM.SPEC water LK this
 ‘Tiwas froze some of the water.’

According to Tsukida (2009), another Atayalic language Truku uses PV *-un* and LV *-an* primarily to express temporal information. PV typically expresses future, as in (55a), while LV is used for other tenses or aspects, as in (55b). Note that the nominative NP in both examples is a location, showing that both *-un* and *-an* can be used to express LV. As for the association of *-un* with future, this is unsurprising given the fact that reflexes of *-en generally surface only in nonfinite or prospective events. This is also the distribution assumed for *-en in PAN (Wolff 1973, Ross 2009, Blust & Chen 2017, and others), a position that I, too, adopt below.

Truku (Tsukida 2009:219)

- (55) a. *taqi-un laqi ka hini*
 sleep-PV child NOM here
 ‘A/the child will sleep here.’
- b. *taqi-an laqi ka hini*
 sleep-PV child NOM here
 ‘A/the child sleeps here.’

As in Mayrinax and Amis, Truku also distinguishes *-un* and *-an* in terms of affectedness of a nominative theme, using PV *-un* when the object is fully affected and *-an* otherwise. In (56a), the trash is completely burned away, while in (56b), the milk is not completely

drunk. Note further that the tense/aspect interpretation is more flexible when the object is fully affected, with the interpretation of (56a) being ambiguous between future and habitual.

Truku (Tsukida 2009:229)

- (56) a. *peseqema-un laqi ka sudu*
burn-PV child NOM trash

‘A/the child burns the trash.’ / ‘A/the child will burn the trash.’

- b. *n-emah-an laqi ka unuh*
PFV-drink-PV child NOM milk

‘The milk has been drunk by a/the child. (The amount has decreased.)’

Thus, Amis, Mayrinax, and Truku use reflexes of both *-en and *-an in clauses with nominative theme arguments, with the distinction between them being aspectual: *-en marks bounded events, while *-an occurs in unbounded events. Consequently, it is unlikely that these two morphemes were originally PV and LV markers, respectively, contra Wolff (1973), Ross (1995, 2009, 2012), Blust (1999), Blust & Trussel (ongoing), and others.

Another reason to suspect a connection with aspect comes from the fact that the reflex of *-en surfaces predominantly in prospective events and does not co-occur with the perfective aspect marker <in>, as shown for Tagalog in (57a). <in> by itself marks perfective PV clauses, as in (57b).

Tagalog

- (57) a. *Bi-bilh-in ng babae ang isda.*
RED-buy-PV GEN woman NOM fish

‘The woman will buy the fish.’

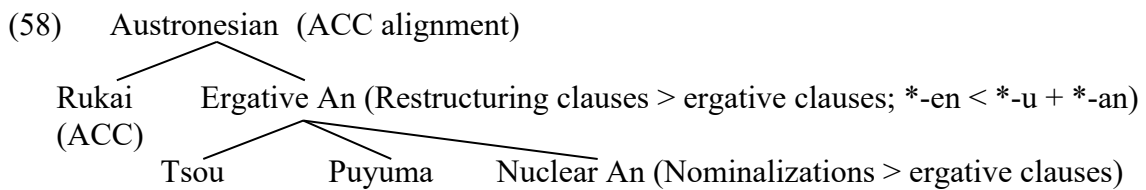
- b. *B<in>ili ng babae ang isda.*
<NAV.PFV>buy GEN woman NOM fish

‘The woman bought the fish.’

I propose that *-en was innovated precisely in order to mark a prospective event as telic. It was not used to mark perfective clauses, since these are telic by default.

It bears mentioning that Blust & Chen (2017) put forth an alternative view of the alternation between *<in> and *-en, proposing that there is a zero allomorph of PV *-en in perfective events. However, there is no independent evidence for the presence of a null PV marker. This proposal also clearly rests on the assumption that *-en functioned primarily as a PV marker, which encounters difficulty when confronted with languages like Amis, Mayrinax, and Truku, where reflexes of *-an also occur in clauses with nominative themes rather than locations. This analysis also does not capture the aspectual differences between *-en and *-an exhibited by these languages. My analysis, on the other hand, accounts naturally for the absence of *-en in the perfective aspect. The role of *-en was originally aspectual, signaling the telicity of an event which otherwise would be assumed to be unbounded, i.e., an event which is not yet realized. *-en is therefore obviated in perfective contexts, where the default interpretation is telic. Later in this section, I discuss why *-en surfaces only in PV clauses, while *-an developed into the NucAn LV marker.

In this section, I propose that *-en was a post-PAN development, emerging in Proto-Ergative Austronesian (PEAn) in order to express telicity in prospective events. It is reflected in some dialects of Puyuma and in most Philippine and Formosan NucAn languages as PV, but it is not found in Rukai or Tsou. Consequently, I treat it as an Ergative Austronesian (EAn) innovation.⁷ Although *-en is not reflected in Tsou, this is not problematic, since Tsou has lost the clausal nominalizations of PAN. It is also not a NucAn language, so it never underwent the reanalysis of nominalizations as matrix clauses. Consequently, *-en is not expected to be reflected in this language. Ergative Austronesian is thus defined by two innovations: 1) the reanalysis of embedded restructuring clauses as the ergative-type of voice system found in modern Tsou and Puyuma; and 2) the development of *-en.



My reconstructions for PEAn are shown in (59). As mentioned above, *-en only occurs in non-perfective, mostly prospective, events, and I assume that this is because perfective events were interpreted as bounded by default and thus did not require special marking. I further propose that *-en is composed of one of the PEAn PV markers *-u and the nominalizer *-an, accounting for its distribution in nominalizations, as I show below for Katripul Puyuma. (59) additionally shows the newly innovated voice system, which emerged in PEAn nonfinite clauses embedded under restructuring verbs. *-a marked embedded nonfinite verbs, while the LV marker *-i was a locative applicative grammaticalizing from an incorporated preposition. Tsou uses the bare nonfinite stem formed by *-a to mark PV, while Puyuma reflects the addition of a second PV marker *-u, which grammaticalized from a definite determiner that was incorporated to the nonfinite verb stem following *-a. I assume that these distinct reflexes of PV in Tsou and Puyuma reflect variation present in the parent language PEAn. *-u seems to have been the dominant variant, as *-au is widely attested among NucAn languages in irrealis mood. I discuss these affixes briefly below, but the reader is referred to Aldridge (2016, 2021a) for more detailed discussion.

(59) PEAn	<u>AV</u>	<u>PV</u>	<u>LV</u>	<u>CV</u>
Non-perfective atelic (N)	---	*-an	*-an	*Si-
Non-perfective telic (N)	---	*-en	*-an	*Si-
Perfective (N)	---	*<in>	*<in>...-an	*Si...<in>
Finite (V)	*M-	---	---	---
Nonfinite (V)	*M-...-a	*-a(u)	*a-i	*an-i

In the remainder of this section, I argue for the nominal origin of *-en, its use in bounded events, and how this use resulted in its developing into the NucAn PV marker,

⁷ Blust and Chen (2017) reconstruct PV *-en and LV *-an to PAN and propose that the two merged in Proto-Rukai. However, even they admit that their arguments are merely suggestive. Furthermore, reconstructing *-en and *-an as voice markers faces difficulties accounting for the paucity of nominalizations formed with *-en, the temporal functions of *-en and *-an, and the overlap between the two in terms of voice. Given the lack of a reflex of *-en in Rukai dialects, I assume that it was simply not present in Proto-Rukai.

while *-an came to be used predominantly as LV. I begin with evidence for the nominal origin of *-en.

As discussed in section 3, *-an was a general VP nominalizer in PAn. In NucAn languages, *-en is reflected primarily in verbal contexts, as demonstrated for Amis, Seediq, and Bunun in section 3. However, I also showed that *-en is not prohibited from appearing in nominalizations. It is found in one extra-NucAn language exclusively in nominalizations, as evidenced by the behavior of *-en in Katripul Puyuma, where it forms nominalized relative clauses on theme position. In Katripul, -en appears in prospective events, indicated by reduplication of the verb stem in (60). Completed events are marked only with the perfective infix <in>.

<u>Root</u>	<u><in>V PFV_N</u>	<u>RED-V-en PV_N</u>
(60) a. <i>kerutr</i> ‘dig’	<i>k<in>erutr</i> ‘thing dug’	<i>ka-kerutr-en</i> ‘thing to be dug’
b. <i>kezeng</i> ‘pull’	<i>k<in>ezeng</i> ‘thing pulled’	<i>ka-kezeng-en</i> ‘thing to be pulled’

(Blust & Chen 2017:606)

(61) provides full sentence examples of this type of relative clause. (61b) additionally shows that this type of relative clause can contain an overt agent, suggesting that this clause type is not passive.

Katripul Puyuma

(61) a. <i>ma-tra-trangis</i>	<i>na</i>	<i>verek</i>	[<i>na</i>	<i>tra-truwak-en</i>]	
AV-RED-cry	NOM	pig	NOM	RED-slaughter-PV.NMLZ	
‘The pig which is going to be slaughtered is crying.’ (Teng 2018:133)					
b. <i>p-u-a-lusu=ku</i>		<i>za</i>	<i>paisu</i>		
CAUS-MOT-IPFV-down=1S.NOM	IND.OBL	money			
[<i>zaku</i>	<i>tra-trima-en</i>	<i>za</i>	<i>'uma.</i>]		
1S.PRS.OBL	RED-buy-PV	OBL	farm	(Stacy Teng, p.c.)	
‘I am withdrawing some money with which I will buy a farm.’					

I assume that this relativizing function of *-en is a PEA retention and that *-en first appeared in nominalized relative clauses. Next, I turn to the formation of *-en from PV *-u and the nominalizer *-an by first establishing the development of *-u as a PV marker and its association with bounded events.

*-u is reflected as a PV marker in Puyuma in both realis and irrealis mood. The following examples show imperative constructions in Nanwang Puyuma. The PV marker is -u, while LV is marked by -i.

Nanwang Puyuma

(62) a. <i>pilang-u</i>	<i>i</i>	<i>temuu</i>			
take-PV.IMP	SG.NOM	your.grandmother			
<i>m-uka</i>	<i>i</i>	<i>drena-drenan</i>			
INTR-go	LOC	RED-mountain			
‘Take your grandmother to the mountains.’					
b. <i>puka-i</i>	<i>dra</i>	<i>tidrul</i>	<i>dra</i>	<i>samaya</i>	
put-LV.IMP	INDEF.OBL	wasp	INDEF.OBL	some	
‘Put some wasps (in).’ (Puyuma; Teng 2008:216)					

These suffixes are also found in realis clauses, but they are preceded by an additional vowel /a/. In Aldridge (2021a), I propose that *-a was a marker of embedded nonfinite clauses in PAn. The voice system reflected in Tsou and Puyuma was innovated in Proto-Ergative Austronesian in nonfinite restructuring clauses where the object could not value accusative case. This made the object dependent on nominative case from the higher clause, thus giving birth to the ergative-type alignment which characterizes the voice system of Tsou and Puyuma. I further propose that realis clauses in Puyuma were derived from these embedded nonfinite clauses after the loss of the auxiliary verb introducing them (following a similar proposal made by Starosta et al. 1982). Because of the co-occurrence with the reflex of *-a, /-u/ and /-i/ are pronounced as offglides in the examples in (63). -a does not appear in imperatives like (62a, b) since these developed from matrix clauses.

- (63) a. *tu=trakaw-aw na paisu kan isaw*
 3.GEN=steal-PV DEF.NOM money SG.OBL Isaw
 ‘Isaw stole the money.’ (Puyuma; Teng 2008:147)
- b. *tu=trakaw-ay=ku dra paisu kan isaw*
 3.GEN=steal-LV=1SG.NOM INDEF.OBL money SG.OBL Isaw
 ‘Isaw stole money from me.’ (Puyuma; Teng 2008:147)

As for the origins of the *-u and *-i voice markers, this can be accounted for on Starosta’s (1995) proposal that they were reanalyzed from the definite determiner *u and the locative preposition *i, which were incorporated to the verb under adjacency.

The preceding discussion presented evidence for reconstructing *-u as a PV marker in PEAn. Proposing that PV *-u grammaticalized from a definite determiner also helps to understand why its distribution should be limited to PV clause types. It is widely known that PV clauses in Formosan and Philippine languages are typically telic, and their objects are generally definite (Zeitoun 1992, 1996 for Tsou; Wu 2006, 2007; Huang & Sung 2008; Kuo 2016 for Amis; Peng 2016, S. Chen 2016, 2018 for Mayrinax Atayal; and Nolasco 2005 for Tagalog). Puyuma clauses marked with a reflex of *-u also provide evidence for the diachronic connection between *-u and definiteness. According to Teng (2008), objects in AV clauses tend to be indefinite, while objects in PV clauses, which are marked with a reflex of *-u, tend to be definite. Stacy Teng (p.c.) also reports that Puyuma PV clauses are generally telic.

Nanwang Puyuma (Teng 2008:147)

- (64) a. *trakaw dra paisu i isaw*
 <AV>steal INDEF.OBL money SG.NOM Isaw
 ‘Isaw stole money.’
- b. *tu=trakaw-aw na paisu kan isaw*
 3.GEN=steal-PV DEF.NOM money SG.OBL Isaw
 ‘Isaw stole the money.’

Such a connection between definiteness of the theme argument and telicity is not surprising, given the fact that syntactic and semantic properties of direct objects play a crucial role in delimiting events (Tenny 1987, 1994; Van Voorst 1988; Borer 1994; Ritter & Rosen 2000; Svenonius 2002; Ramchand 2008; Travis 2010; and others). The following examples show that a specific object in the English example in (65a) can provide an endpoint to a bounded event, while this is not possible in the case of a

nonspecific object in (65b). Specificity allows the object in (65a) to be measured out and fully affected over the course of the event.

- (65) a. Mary drank **a beer** in an hour.
b. Mary drank **beer** for an hour.

Consequently, it is not surprising that a definite determiner developed into a PV marker, given the role of definite or specific definite objects in building bounded events.

Another factor which led to the development of a designated PV marker in telic events was the case marking on the object, since theme direct objects have structural (nominative) case in PV clauses in Austronesian languages. This is because of the role played by structural case marking on themes in creating bounded events (De Hoop 1996; Borer 1994, 2005; Kiparsky 1998; Ritter & Rosen 2000; Van Hout 2000; Kratzer 2004; Csirmaz 2005, 2012; Basilico 2008; Travis 2010; among many others). For example, when an object in Finnish has accusative case, the object is fully affected and the event is interpreted as completed (bounded). Consequently, the bear referred to in (66a) is actually shot. But in (66b), where the object has partitive case, the bear is not necessarily hit by the shot; what is communicated is merely the fact that a shooting event with a bear as its target took place.

Finnish (Kiparsky 1998:267)

- (66) a. *Ammu-i-n* *karhu-n*.
shoot-PST-1SG bear-ACC
‘I shot a/the bear.’
- b. *Ammu-i-n* *karhu-a*.
shoot-PST-1SG bear-PART
‘I shot at a/the bear.’

In short, the definite determiner *u developed into a designated PV marker, because it is in this clause type where a definite/specific object with structural case combines with the predicate to create a bounded event. This characteristic of PV clauses is also related to the development of *-en in nominalized relative clauses in PEAn. I propose that the PEAn verbal PV marker *-u was also employed in nominal environments in order to mark telic events in prospective aspect. In this environment, *-u was followed by the nominalizer *-an, producing the new PV marker *-en. The correlation with telicity likewise ensured that *-en would only surface in PV clauses.

The changes and their motivations that I have proposed for the development of PV marking in Austronesian languages is summarized in the following table. *-u developed into a PV marker in PEAn via incorporation of the determiner marking a definite direct object in this clause type. The restriction of *-u to PV clauses was due to the combined contribution of the definiteness of the object with its structural nominative case marking. These two factors in turn conspired to produce a bounded interpretation for events expressed by this clause type, since it is precisely when definite/specific direct objects have structural case that they introduce an endpoint to an event. Consequently, it is only in PV clauses where marking with an erstwhile definite determiner (*-u) surfaces in conjunction with a nominative object in a telic event. *-u was extended to nominalized relative clauses in order to indicate the boundedness of a prospective event, since prospective events are not telic by default. *-u was followed by the nominalizer *-an in this environment, creating a nominal PV affix *-en. Relative clauses projected by *-en

were reanalyzed as verbal in PNucAn. The association with telicity ensured that *-en would continue to function as a PV marker and not surface in other voices.

Table 3. Diachronic development of PV marking

Stage	Change	Motivation
PEAn	*u DET > *-u PV _V	1. Incorporation of DET to V 2. Correlation of DEF, NOM object in telic event
	*-u + *-an > *-en PV _N	Mark telicity in prospective aspect
PNucAn	*-en PV _N > *-en PV _V	Reanalysis of nominal RC as verbal

In section 3, I proposed that *-an was a lexical nominalizer, as well as relative clause former on VP-internal gaps (both themes and locations) in PAn, having grammaticalized from a noun meaning ‘place’ in Pre-PAn. As a relativizer, PAn *-an did not contribute any aspectual information, but it ceased to be used in telic events after telic *-en developed in PEAn. These functions of *-en and *-an in nominal environments were inherited by PNucAn in verbal clauses, thus accounting for the aspectual distinction between *-en and *-an in PV clauses in Tagalog, Amis, and Atayalic languages shown in (52)–(56). The predominant use of *-an as an LV marker in NucAn languages developed as a consequence of the fact that nominative marking on objects in PV clauses led increasingly to a bounded interpretation for these events. These changes are summarized in Table 4.

Table 4. Diachronic development of LV marking

Stage	Change	Motivation
Pre-Pan	*an ‘place’ > *-an NMLZ	Common morphosyntactic changes
	*-an NMLZ > *-an REL _N	
PEAn	Loss of *-an REL _N TEL	Development of *-en REL _N TEL
PNucAn	*-an REL _N > *-an LV/PV _V	Reanalysis of nominal RC as verbal
	*-an LV/PV _V > *-an LV _V	Increase of telic PV with *-en

As shown in (60)–(61), Puyuma retains *-en only in nominal environments, because it is an extra-NucAn language and consequently has not undergone the reanalysis of nominalizations as verbal matrix clauses. Kananavu, Amis, and Seediq reflect intermediate stages. Clauses with *-en have been fully reanalyzed as verbal and consequently no longer appear in nominal environments. On the other hand, clauses with *-an have not yet completed this transition, and relative clauses projected by *-an continue to possess nominal characteristics. The behavior of *-an in Kananavu, Amis, and Seediq also argues against an alternative approach in which *-en and *-an are undergoing merger.⁸ Such an approach would assume that *-en and *-an were

⁸ Merger of *-en and *-an is proposed by Blust and Chen (2017) in order to explain the paucity of a reflex of *-en in extra-Nuclear Austronesian languages.

historically found as voice markers in both nominal and verbal clauses, so it would have to assume independent innovations in the different languages. Furthermore, merger provides no obvious explanation for why *-an is found primarily in nominal environments, as opposed to verbal.

It is only in languages which have completed the nominal-to-verbal transition where reflexes of *-en and *-an occur freely in both matrix and relative clauses. This is because the loss of the categorial distinction between matrix and relative clauses resulted in the employment of verbal clauses in both environments. Tagalog is such a language. (67) shows PV examples and (68) shows LV examples.

Tagalog

- (67) a. *Bi-bilh-in ng babae ang isda.*
 RED-buy-PV GEN woman NOM fish

‘The woman will buy the fish.’

- b. *isda-ng [bi-bilh-in ng babae]*
 fish-LK RED-buy-PV GEN woman

‘the fish that the woman will buy’

Tagalog

- (68) a. *Bi-biby-an ng babae ng isda ang lalaki.*
 RED-give-LV GEN woman GEN fish NOM man

‘The woman will give a fish to the man.’

- b. *lalaki-ng [bi-biby-an ng babae ng isda]*
 man-LK RED-give-LV GEN woman GEN fish

‘the man that the woman will give a fish to’

In languages like Tagalog, then, the behaviors of the voice affixes are completely parallel in matrix and relative clauses. The functions of PV and LV are also clearly differentiated, LV being employed in order to project relative clauses only on goal or location gaps and not on themes. This shows that in Tagalog the erstwhile nominalizations have been completely reanalyzed as verbal, since *-an is no longer able to project a relative clause on both locative and theme gaps.⁹

5. Conclusion

This paper has proposed diachronic origins for the Austronesian PV and LV suffixes *-en and *-an. I have shown that the primary distinction between the two is in telicity rather than voice. *-an was historically neutral with respect to aspect, having grammaticalized from a noun meaning ‘place’ to become a nominalizer and relative clause former in Proto-Austronesian (PAn). In contrast, PV markers developed in telic clause types in Proto-Ergative Austronesian (PEAn) due to the fact that these clauses featured definite direct objects marked with structural nominative case. The restriction

⁹ Others, e.g., Kaufman (2009), have also proposed that Tagalog lacks a categorial distinction between matrix and relative clauses, but Kaufman takes both to be nominal rather than verbal. Full discussion of this proposal is beyond the scope of the current paper, but my findings do suggest that clauses in Tagalog are not nominal, since they do not display the nominal behavior of uncontroversial nominalizations found in Puyuma, Rukai, and Kananavu discussed in this paper, in particular the ability of *-an to occur with theme gaps, as well as the lack of tense.

of affixes like *-en to telic clauses because of their definite, nominative objects led to the loss of the ability of *-an to appear in PV clauses expressing bounded events. This resulted eventually in the specialization of *-an as an LV marker in Nuclear Austronesian (NucAn) languages after the reanalysis of nominalized relative clauses as verbal matrix clauses and the proliferation of telic PV clauses with nominative objects.

This proposal accounts for a number of anomalies left unexplained if *-en and *-an are both attributed to PAn as voice markers. The first is the fact that *-en is only rarely attested outside of NucAn languages, in contrast to the ubiquitous employment of *-an in nominalizations. On my proposal, *-an was present in PAn, while *-en developed later in PEAn.

Another key fact accounted for only on my approach is the overlap between *-an and *-en in clauses with nominative themes. This suggests strongly that the two were not originally distinguished in terms of voice. Furthermore, the aspectual difference between them – *-en marking bounded events and *-an surfacing in unbounded events – is better accounted for on my proposal that PV *-en developed specifically in telic environments.

Finally, my proposal makes a unique contribution to Austronesian historical linguistics in simply offering an explanation for the existence of the cross linguistically unusual “voice” system, since this analysis argues for the historical origin and development of this paradigm from morphemes playing more conventional roles in erstwhile accusative languages.

Abbreviations

1	First person	2	Second person
3	Third person	ACC	Accusative
AG	Agent	AV	Actor voice
CAUS	Causative	COND	Conditional
CONV	Converb	CV	Circumstantial voice
DEF	Definite	DEM	Demonstrative
DES	Desiderative	DET	Determiner
EAn	Ergative Austronesian	ERG	Ergative
FS	Feminine singular	FUT	Future
GEN	Genitive	ILLAT	illative
IMP	Imperative	INC	Inclusive
INCH	Inchoative	INDEF	Indefinite
INS	Instrument	IPFV	Imperfective
LAT	Lative	LK	Linker
LOC	Locative	LV	Locative voice
MOT	Motion	N	Nominal
NAV	Non-actor voice	NEG	Negation
NMLZ	Nominalization	NOM	Nominative
NPP	Non-past participle	NSR	Non-subject relative
NSUBJ	Non-subject	NucAn	Nuclear Austronesian
OBL	Oblique	OREL	Object relative
P	Preposition	PAn	Proto-Austronesian
PART	Partitive	PEAn	Proto-Ergative Austronesian
PFV	Perfective	PL	Plural
PN	Personal name	PNucAn	Proto-Nuclear Austronesian
PRES	Present	PRS	Personal
PST	Past	PV	Patient voice
RC	Relative clause	REAL	Realis

RED	Reduplication	REL	Relative
SG	Singular	SPEC	Specific
SR	Subject relative	STAT	Stative
SUBJ	Subject	TEL	Telic
TOP	Topic	V	Verbal

References

- Aldridge, Edith. 2016. Ergativity from subjunctive in Austronesian languages. *Language and Linguistics* 17. 27–62.
- Aldridge, Edith. 2021a. Syntactic conditions on accusative to ergative alignment change in Austronesian languages. *Journal of Historical Linguistics* 11. 214–247.
- Aldridge, Edith. 2021b. The nature and origin of syntactic ergativity in Austronesian languages. In Jóhannes Gíslí Jónsson & Thórhallur Eythórsson (eds.), *Syntactic Features and the Limits of Syntactic Change*. Oxford: Oxford University Press. 265–300.
- Aldridge, Edith. To appear. The typology and diachrony of Philippine-type alignment. In James Crippen, Rose-Marie Dechaine & Hermann Keupdjio (eds.), *Move and Agree*. Amsterdam & New York: John Benjamins.
- Aygen, N. Gulsat. 2002. *Finiteness, Case and Clausal Architecture*. Ph.D. thesis. Harvard University, Cambridge, MA.
- Barðdal, Jóhanna & Thórhallur Eythórsson. 2012. Reconstructing syntax: Construction Grammar and the Comparative Method. In Hans C. Boas & Ivan Sag (eds.), *Sign-Based Construction Grammar*. Stanford: CSLI Publications. 257–308. <http://org.uib.no/iecastp/ReconstructingSyntax-CSLI-Revised.pdf>
- Basilico, David. 2008. The syntactic representation of perfectivity. *Lingua* 118. 1716–1739.
- Blust, Robert A. 1999. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In Elizabeth Zeitoun & Paul Jen-kuei Li (eds.), *Selected Papers from the eighth International Conference on Austronesian Linguistics* (Symposium Series of the Institute of Linguistics). Taipei: Institute of Linguistics (Preparatory Office), Academia Sinica. 31–94.
- Blust, Robert & Victoria Chen. 2017. The pitfalls of negative evidence: ‘Nuclear Austronesian’, ‘Ergative Austronesian’, and their progeny. *Language and Linguistics* 18. 577–621.
- Blust, Robert & Stephen Trussel. ongoing. *The Austronesian comparative dictionary*. Online version: <http://www.trussel2.com/acd/>
- Borer, Hagit. 1994. The Projection of Arguments. In E. Benedicto & J. Runner (eds.), *Functional Projections*. University of Massachusetts Occasional Papers 17. 19–47.
- Borer, Hagit. 2005. *Structuring Sense, vol. II: The normal course of events*. New York: Oxford University Press.
- Bril, Isabelle. 2022. Lexical restrictions on grammatical relations in voice constructions (Northern Amis). *STUF – Language Typology and Universals* 75. 21–71.

- Cagri, İlhan Merih. 2005. *Minimality and Turkish relative clauses*. Ph.D. thesis. University of Maryland, College Park.
- Chang, Henry Y. 2011. Transitivity, ergativity, and the status of O in Tsou. In Jung-hsing Chang (ed.), *Language and Cognition: Festschrift in Honor of James H-Y. Tai on His 70th Birthday*. Taipei: Crane Publishing. 277–308.
- Chang, Henry Yung-li & Amy Pei-jung Lee. 2002. Nominalization in Kavalan. *Language and Linguistics* 3. 349–368.
- Chen, Sihwei. 2016. Lexical aspect and aspectually unmarked predicates in Atayal. In Hiroki Nomoto, Takuya Miyauchi & Asako Shiohara (eds.), *AFLA 23: The Proceedings of the 23rd Meeting of the Austronesian Formal Linguistics Association*. Canberra: Asia-Pacific Linguistics. 65–80.
- Chen, Sihwei. 2018. *Finding Semantic Building Blocks: Temporal and modal interpretation in Atayal*. Ph.D. thesis. University of British Columbia, Vancouver.
- Chen, Tingchun. 2018. *Multiple Case Assignment: An Amis case study*. Ph.D. thesis. MIT, Cambridge, MA.
- Cole, Peter & Gabriella Hermon. 2011. Nominalization and case assignment in Quechua. *Lingua* 121. 1225–1251.
- Csirmaz, Aniko. 2005. *Semantics and phonology in Syntax*. Ph.D. thesis. MIT, Cambridge, MA.
- Csirmaz, Aniko. 2012. The case of the divisible phase: Licensing partitive case in Finnish. *Syntax* 15. 215–252.
- Corston-Oliver, Simon. 2002. Roviana. In Lynch et al. (eds.), *The Oceanic Languages*, Richmond, Surrey: Curzon. 467–497.
- De Busser, Rik L.J. 2009. *Towards a Grammar of Takivatan Bunun: Select topics*. Ph.D. thesis. La Trobe University, Bundoora, Australia.
- Dede, Müşerref. 1978. Why should Turkish relativization distinguish between subject and non-subject head nouns? In Jeri J. Jaeger, Anthony C. Woodbury, Farrell Ackerman, Christine Chiarello, Orin D. Gensler, John Kingston, Eve E. Sweetser, Henry Thompson & Kenneth W. Whistler (eds.), *Proceedings of the 4th annual meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society. 67–77.
- De Hoop, Helen. 1996. *Case Configuration and Noun Phrase Interpretation*. New York: Garland Press.
- Gildea, Spike. 1998. *On Reconstructing Grammar*. New York: Oxford University Press.
- González, Albert Álvares. 2012. Relative clauses and nominalizations in Yaqui. In Bernard Comrie & Zarina Estrada-Fernández (eds.), *Relative Clauses in Languages of the Americas: A typological overview*. 67–95. Amsterdam & Philadelphia: John Benjamins.
- Hankamer, Jorge & Laura Knecht. 1976. The role of the subject/non-subject distinction in determining the choice of relative clause participle in Turkish. *NELS* 4. 123–135.
- Harris, Alice & Lyle Campbell. 1995. *Historical Syntax in Cross-linguistic Perspective*. New York: Cambridge University Press.

- Hastings, Rachel. 2004. *The Syntax and Semantics of Relativization and Quantification: The case of Quechua*. Ph.D. thesis. Cornell University, Ithaca, NY.
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Alexander Adelaar & Nikolaus P. Himmelman (eds.), *The Austronesian Languages of Asia and Madagascar*. New York: Routledge. 110–181.
- Holmstedt, R. D. 2002. *The relative clause in Biblical Hebrew: A linguistic analysis*. Ph.D. thesis. The University of Wisconsin-Madison, Madison, WI.
- Huang, Shuping & Li-May Sung. 2008. The undergoer focus ma- in Kavalan and Amis. *Oceanic Linguistics* 47. 159–184.
- Jeng, Heng-hsiung. 1977. *Topic and Focus in Bunun*. Taipei: Institute of History and Philology, Academia Sinica.
- Kaufman, Daniel. 2009. Austronesian Nominalism and its Consequences: A Tagalog case study. *Theoretical Linguistics* 35. 1–49.
- Kiparsky, Paul. 1998. Partitive Case and Aspect. In Miriam Butt & Wilhelm Geuder (eds.), *The Projection of Arguments*. Stanford: CSLI.
- Kornfilt, Jaklin. 1997. On the Syntax and Morphology of Relative Clauses in Turkish. In Kamile İmer, Ahmet Kocaman & Sumru Özsoy (eds.), *Dilbilim Araştırmaları [Linguistic investigations]*. Ankara: Keçiçek Yayınları. 24–51.
- Kornfilt, Jaklin. 2007. Verbal and nominalized finite clauses in Turkish. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*, New York: Oxford University Press. 305–332.
- Kornfilt, Jaklin. 2008. Subject case and *Agr* in two types of Turkic RCs. In Suleyman Ulutaş & Cedric Boeckx (eds.), *Proceedings of WAFL 4*. Cambridge, MA: MITWPL 56. 145–168.
- Kratzer, Angelika. 2004. Telicity and the meaning of objective case. In Jacqueline Guéron & Jacqueline Lecarme (eds.), *The Syntax of Time*. Cambridge, MA: MIT Press. 389–423.
- Kuo, Jonathan Cheng-Chuen. 2016. The event structure of Amis intransitive manner/result verbs: A lexicalist analysis. In Hiroki Nomoto, Takuya Miyauchi & Asako Shiohara (eds.), *AFLA 23: The Proceedings of the 23rd Meeting of the Austronesian Formal Linguistics Association*. Canberra. 138–151.
- Li, Paul Jen-kuei. 1973. *Rukai Structure*. Taipei: Academia Sinica Institute of History and Philology.
- Lord, Carol. 1993. *Historical Change in Serial Verb Constructions*. Amsterdam & Philadelphia: John Benjamins.
- Lynch, John, Malcolm Ross, Terry Cowley. 2002. *The Oceanic Languages*. Curzon Press. Reprinted in 2011. London: Routledge.
- Muysken, Pieter. 2011. Relative clauses in Ecuadorian Quechua. In Rik van Gijn, Katharina Haude & Pieter Muysken (eds.), *Subordination in Native South-American Languages*. Amsterdam/Philadelphia: John Benjamins. 251–266.

- Nolasco, Ricardo. 2005. What ergativity in Philippine languages really mean. *Proceedings of the Taiwan-Japan Joint Workshop on Austronesian Languages*. 215–238.
- Pan, Chia-Jung. 2012. *A Grammar of Lha'alua: An Austronesian language of Taiwan*. Ph.D. thesis. James Cook University, Douglas, Australia.
- Peng, Shu-fang. 2016. *A Study of Verbal Morphology in Mayrinax Atayal*. MA Thesis. National Tsing Hua University, Hsinchu, Taiwan.
- Peterson, David A. 1997. The evolution of applicative constructions and Proto-Austronesian morphosyntax. *Berkeley Linguistics Society* 23. 278–289.
- Peterson, David A. 2007. *Applicative Constructions*. Oxford & New York: Oxford University Press.
- Potanina, Olga. 2008. Grammaticalization and relative clauses in Eastern Khanty. In Edward J. Vajda (ed.), *Subordination and Coordination Strategies in North Asian Languages*. Amsterdam & Philadelphia: John Benjamins. 77–84.
- Pylkkanen, Liina. 2002. *Introducing Arguments*. Ph.D. thesis. MIT, Cambridge, MA.
- Ritter, Elizabeth & Sara Rosen. 2000. Event Structure and Ergativity. In C. Tenny & J. Pustejovsky (eds.), *Events as Grammatical Objects*. Stanford: Center for the Study of Language and Information. 187–238.
- Ramchand, Gillian Catriona. 2008. *Verb meaning and the lexicon: A first-phase syntax*. New York: Cambridge University Press.
- Roberts, Ian & Anna Roussou. 2003. *Syntactic change: A Minimalist approach to grammaticalization*. Cambridge: Cambridge University Press.
- Ross, Malcolm. 1995. Reconstructing Proto-Austronesian Verbal Morphology: Evidence from Taiwan. In Paul J.-K. Li, Cheng-hwa Tsang, Ying-kuei Huang, Dah-an Ho & Chiu-yu Tseng (eds.), *Austronesian Studies Relating to Taiwan*. Taipei: Academia Sinica. 727–791.
- Ross, Malcolm. 2009. Proto Austronesian Verbal Morphology: A reappraisal. In Alexander Adelaar & Andrew Pawley (eds.), *Austronesian Historical Linguistics and Culture History: A festschrift for Robert Blust*. Canberra: Pacific Linguistics. 295–326.
- Ross, Malcolm. 2012. In defense of Nuclear Austronesian (and against Tsouic). *Language and Linguistics* 13. 1253–1300.
- Sebba, Mark. 1987. *Syntax of Serial Verbs: An Investigation into Serialisation in Sranan and other Languages*. Amsterdam & Philadelphia: John Benjamins.
- Sneddon, James Neil. 1996. *Indonesian: A comprehensive grammar*. London: Routledge.
- Starosta, Stanley. 1995. A grammatical subgrouping of Formosan languages. In Paul J.-K. Li, Cheng-hwa Tsang, Ying-kuei Huang, Dah-an Ho & Chiu-yu Tseng (eds.), *Austronesian Studies Relating to Taiwan*. Taipei: Academia Sinica. 683–726.
- Starosta, Stanley & Andrew K. Pawley & Lawrence A. Reid. 1982. The Evolution of Focus in Austronesian. In Amram Halim, Lois Carrington & S. A. Wurm (eds.), *Third International Conference on Austronesian Linguistics*. Pacific Linguistics, C-75. 145–170.

- Svenonius, Peter. 2002. Icelandic case and the structure of events. *Journal of Comparative Germanic Linguistics* 5. 197–225.
- Teng, Stacy Fang-ching. 2008. *A Reference Grammar of Puyuma*. Canberra: Pacific Linguistics.
- Teng, Stacy Fang-Ching. 2018. A sketch grammar of Puyuma [in Chinese]. Series on Formosan Languages, 13. Taipei: Council of Indigenous Peoples.
- Teng, Stacy F & Elizabeth Zeitoun. 2016. The noun-verb distinction in Kanakanavu and Saaroa: Evidence. *Oceanic Linguistics* 55. 134–161.
- Tenny, Carol Lee. 1987. *Grammaticalizing Aspect and Affectedness*. Ph.D. thesis. MIT, Cambridge, MA.
- Tenny, Carol. 1994. *Aspectual Roles and the Syntax-Semantics Interface*. Dordrecht: Kluwer Academic Publishers.
- Travis, Lisa deMena. 2010. *Inner Aspect*. Dordrecht: Springer.
- Tsukida, Naomi. 2009. *Sedekkugo (Taiwan) Bunpo* [Sedek Grammar]. Ph.D. thesis. University of Tokyo, Tokyo.
- Tung, T'ung-ho. 1964. *A Descriptive Study of the Tsou Language, Formosa*. Taipei: Institute of History and Philology, Academia Sinica.
- Van Hout, Angeliek. 2000. Event semantics in the Lexicon-Syntax Interface: Verb frame alternations in Dutch and their acquisition. In Carol Tenny & James Pustejovsky (eds.), *Events as Grammatical Objects: The converging perspectives of lexical semantics and syntax*. Stanford: CSLI Publications. 238–282.
- Voorst, Jan. van. 1988. *Event Structure*. Amsterdam: John Benjamins.
- Walkden, George. 2013. The correspondence problem in syntactic reconstruction. *Diachronica* 30. 95–122.
- Walkden, George. 2014. *Syntactic Reconstruction and Proto-Germanic*. Oxford University Press.
- Willis, David. 2011. Reconstructing last week's weather: Syntactic reconstruction and Brythonic free relatives. *Journal of Linguistics* 47. 407–446.
- Wilson-Wright, Aren M. 2019. The Canaanite languages. In Huehnergard, J. & Pat-El, N. A. (eds.), *The Semitic Languages*. London: Routledge. 509–532.
- Wolff, John. 1973. Verbal Inflection in Proto-Austronesian. In Andrew Gonzales (ed.), *Essays in Honor of Cecilio Lopez on His Seventy-Fifth Birthday*. Quezon City: Linguistic Society of the Philippines. 71–94.
- Wu, Jing-lan Joy. 2006. *Verb Classification, Case Marking, and Grammatical Relations in Amis*. Ph.D. thesis. State University of New York, Buffalo.
- Wu, Joy. 2007. “Voice” markers in Amis: A Role and Reference Grammar analysis. *Language and Linguistics* 8. 95–142.
- Zeitoun, Elizabeth. 1992. *A Syntactic and Semantic Study of Tsou Focus System*. MA thesis. National Tsing Hua University, Hsinchu, Taiwan.
- Zeitoun, Elizabeth. 1996. The Tsou temporal, aspectual, and modal system revisited. *Bulletin of the Institute of History and Philology* 67. 503–531.

- Zeitoun, Elizabeth. 2007. *A Grammar of Mantauran (Rukai)*. Taipei: Institute of Linguistics, Academia Sinica.
- Zeitoun, Elizabeth and Stacy F. Teng. 2016. Reassessing the position of Kanakanavu and Saaroa among the Formosan languages. *Oceanic Linguistics* 55.1: 162-198.