Proto-Austronesian Interrogative Pronouns and Their Development
Edith Aldridge

Abstract: This paper reconstructs forms for Proto-Austronesian (PAn) 'who' and 'what' from which the corresponding interrogative pronouns found across Formosan languages can be straightforwardly derived. One key aspect of these reconstructions is that PAn interrogative pronouns consisted of an operator attached to an indefinite base. This shows PAn interrogative pronouns to belong to a consistent morphological paradigm, as well as to be interpreted in terms of an operator-variable configuration, as is true of *wh*-words cross linguistically. Another fundamental characteristic of my reconstructions is the division of interrogative pronouns into two semantic classes determined by the operators. The operator in the case of 'who' was a determiner and created a category construed with sets of individuals, while the operator creating 'what' resulted in a form construed with sets of properties, a commonly found typological distinction between 'who' and 'what'. The diverse array of forms found in Formosan languages can then be accounted for in terms of renewal of these word-internal operators from case markers in response to the opacity in the earlier composition induced by phonological and morphosyntactic changes taking place in the respective languages.
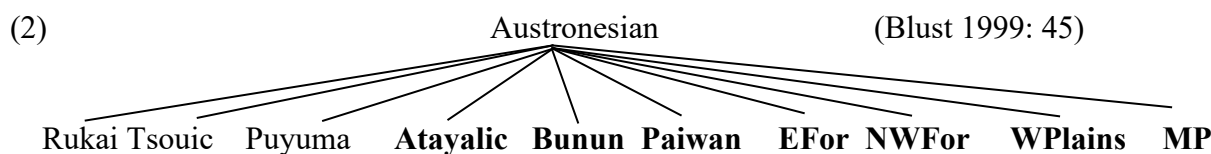
Keywords: reconstruction, syntactic change, wh-word, operator, variable, Formosan languages

1. Introduction

This paper proposes Proto-Austronesian (PAn) reconstructions of the interrogative pronouns for 'who' and 'what'. I take as my starting point the following pattern identified by Blust (2009) and Blust & Trussel (ongoing), referred to as "B&T" in the remainder of this paper. They point out that the form for 'who' in numerous Formosan languages[1] is composed of a nominative case marker for personal names *si and a base which they reconstruct as *ima 'who'.

(1)   WHO   (Blust 2009: 510)
    Thao   t(i)-ima
    Bunun   s(i)-ima
    Amis   c(i)-ima
    Paiwan   t(i)-ima
    Truku   ima

The reconstruction of *ima makes sense on Blust's (1999) subgrouping hypothesis, given that reflexes of *ima are found in seven of the ten first-order subgroups he assumes, but not in Rukai, Tsouic, or Puyuma.
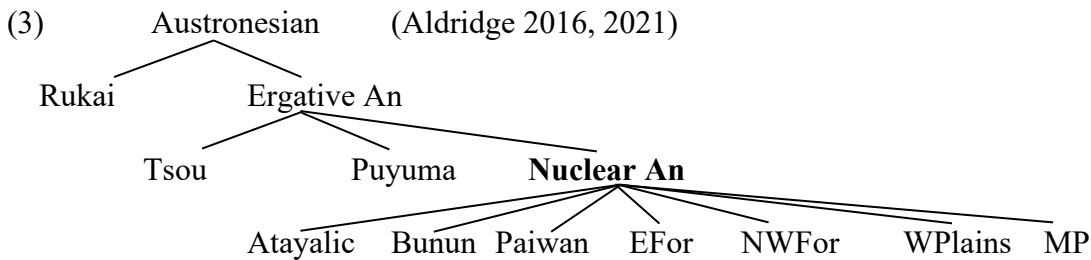
(2)   Austronesian   (Blust 1999: 45)

    Rukai Tsouic Puyuma **Atayalic Bunun Paiwan EFor NWFor WPlains MP**

However, the limitation of *ima to these subgroups is actually more compatible with Ross' (2009) Nuclear Austronesian (NucAn) hypothesis, according to which all high-order Austronesian subgroups except Rukai, Tsou, and Puyuma reflect the reanalysis of nominalized relative clauses as transitive matrix clauses.[2] Reflexes

---

[1] The term "Formosan languages" refers to Austronesian languages spoken in Taiwan. Given the consensus that Taiwan was the homeland of Proto-Austronesian, data from Formosan languages is essential in reconstructing this proto-language.

[2] I do not discuss this innovation in this paper, but the reader is referred to Aldridge and Yanagida (2021) for further argumentation for and discussion of this innovation.

of *ima are found in all and exactly the daughters of Proto-Nuclear Austronesian (PNucAn), which strongly suggests that the creation of this interrogative pronoun was another innovation defining this subgroup. The nominative case marker for personal names *si is also reflected only in NucAn languages. (3) shows the subgrouping hypothesis I adopt in this paper, which is based on Aldridge (2016, 2021). This family tree also includes NucAn. In this paper, I propose that the creation of *ima 'who' was a second innovation defining this subgroup.

(3) Austronesian (Aldridge 2016, 2021)

Rukai Ergative An

   Tsou   Puyuma **Nuclear An**

     Atayalic Bunun Paiwan EFor NWFor WPlains MP

The list in (4) shows that neither *ima nor *sima is reflected in extra-Nuclear languages. I also include the forms for 'what', as I will propose reconstructions for both forms in (5).

(4)

	WHO	WHAT	
Budai Rukai	anə-anə	manə-manə	(Chen 1999)
Tanan Rukai	anə-a	manə-ma	(author's fieldnotes)
Tsou	si-a	cu-ma	(Chang & Pan 2018)
Nanwang Puyuma	manay	manay	(Teng 2008)
Katripul Puyuma	əman	əman	(Teng 2018)
Ulribulribuk Puyuma	manu		(Ting 1978)

My PAn reconstructions are shown in (5). I argue for how these reconstructions can derive each of the forms shown in (4) in section 3.

(5) PAn
 a. *a-nu   WHO (D+INDEF = 'some/which one')
 b. *ma-nu WHAT (V+ INDEF = 'some/what kind')
 c. *i-nu   WHERE (P+INDEF = 'at some/which place')

In addition to the requisite sound changes, I also argue for the morphological compositions of these forms in section 2. First, all three forms in (5) are built on a base *nu, which I assume was the indefinite element in interrogative pronouns in PAn. As I show in section 2, interpreting interrogative pronouns involves an operator-variable construction. Based on Di Sciullo's (2005) analysis of English *wh*-words, I propose that the operator and indefinite were both contained within the pronoun itself in PAn. The operator in the form for 'who' is an indefinite determiner *a.[3] The PAn form for 'what' also contains an operator prefix *ma-, which I propose in section 2 was the stative verbal affix in PAn, with the result that the word for 'what' expresses the 'state of being something'. Formation of the form for 'where' *i-nu is similar, including an incorporated locative preposition. B&T also provide the same derivation. Since this reconstruction is uncontroversial, I do not discuss it further in this paper. I include it here merely to demonstrate that all of the interrogative pronouns which are reconstructable to PAn can be placed into a paradigm built on the indefinite element *nu.

  Another key property of PAn interrogative pronouns for 'who' and 'what' was their division into different semantic types. *anu 'who' quantified over sets of individuals, i.e. the semantic type <e>. In contrast, *ma-

---

[3] B&T reconstruct a similar form *anu but analyze it as an indefinite 'what-cha-ma-callit'. Though reflexes of *anu as 'what' can be found in Malayo-Polynesian languages, it is not reflected unambiguously as 'what' or as an indefinite like 'something' in Formosan languages. Since it is reflected only as 'who/someone' in extra-NucAn languages, this is the gloss I choose. In section 2, I also show that there is no reason to reconstruct separate forms for interrogative and indefinite functions, since Formosan interrogative pronouns reflect both of these functions.

nu was associated with a set of properties or kinds, i.e. semantic type <e,t>. I show in section 2 that this dichotomy is widely found in languages both in and outside of the Austronesian family.

This morphological analysis also has positive consequences for the diachronic development of the forms for 'who' and 'what' in Austronesian languages. I show in sections 3 and 4, respectively, how new forms were created in Formosan languages through processes of renewal that added new operator morphemes to an indefinite base.

This paper is structured as follows. In section 2, I present morphological evidence for my reconstructions of these two pronouns, both the origins of the three morphemes in the reconstructions of 'who' and 'what' in (5) and my proposal that these two interrogative pronouns belong to different semantic types. Section 3 shows how the forms currently in use in extra-NucAn languages could have been derived from the reconstructions given in (5). Section 4 proposes reconstructions of 'who' and 'what' in PNucAn and briefly sketches how the PNucAn forms developed from the PAn reconstructions.

2. Morphosyntactic Properties of PAn Interrogative Pronouns

This section argues for the morphological aspects of my PAn reconstructions of 'who' and 'what' through a combination of evidence from Austronesian and non-Austronesian languages. My proposed reconstructions for 'who' and 'what' are repeated below. In addition to the indefinite base *nu, there are two key aspects of these reconstructions: 1) the form for 'who' contains an incorporated determiner *a; and 2) the form for 'what' has a verbal prefix *ma-. This section offers evidence for reconstructing these three morphemes with these uses in the forms for 'who' and 'what'.

(6) PAn
    a.   *a-nu   WHO (D+INDEF = 'some/which one')
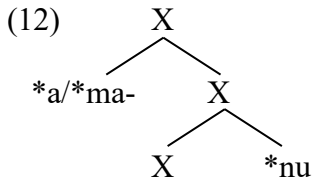   b.   *ma-nu   WHAT (V+ INDEF = 'some/what kind')

First, I propose that *a was an indefinite determiner in PAn. The most direct evidence for this comes from the fact that a reflex of *a functions as a nominative case marker in Puyuma for indefinite common nouns, as in (7a). *a is also reflected in Paiwan as a nominative case marker for common nouns, as in (7b). Note that Ross (2006) and B&T additionally reconstruct a determiner *a. The combination of the determiner *a and the indefinite *nu created a form with the category D that could mean either 'someone' or 'which one' depending on the context, as I discuss below.

(7) a.   Drua-drua   me-na'u-a   **a**   **trau**.
          RED-come   AV-see-SBJV   INDEF.NOM   person
          'Many people came to see.'   (Nanwang Puyuma; Teng 2008: 113)
    b.   S<in>aqis   **a**   **'u-itong**   ni   ina.
          <NAV.PFV>make   NOM.CN   1SG.GEN-clothing   GEN.PN   mother
          'Mother made my clothes.'   (Southern Paiwan)

I propose that the prefix *ma- in the form for 'what' was the stative prefix *ma-, which is reconstructed by Kaufman (2011) as an existential verb. The examples in (8) show reflexes of *ma- forming stative verbs from roots expressing properties. The function of *ma- is to assert that the subject has the property expressed by the root to which *ma attaches.

(8) a.   **Ma**-bini   salaman.
          STAT-full   bowl
          'The bowl is full.'   (Pazeh; Zeitoun & Huang 2000: 406)
    b.   **Ma**-beleng   ko-ani   angato.
          STAT-tall   NOM-that   tree
          'That tree is tall.'   (Tanan Rukai; Li 1973: 214)

c. **Ma**-ganda   ang   babae.
STAT-beautiful NOM   woman
'The woman is beautiful.'   (Tagalog)

When attaching to the indefinite *nu, I propose that *ma- created a predicate expressing some indeterminate property, in other words producing an indefinite or interrogative pronoun meaning 'some kind' or 'what kind'.

It may be wondered how a verb can serve as an argument in a clause. This is unproblematic for my analysis, since it is well established that verbs in Philippine and Formosan languages can be used in the same way as common nouns, surfacing in argument position following case markers. Kaufman (2009) even goes so far as to claim that all Tagalog predicates are inherently nominal.

(9)   Tagalog (Kaufman 2009: 5)
ang   b<um>ili
NOM.CN   <AV.PFV>buy
'the buyer/one who bought (something)'

Consequently, the category difference between V and N in these languages is negligible, to the extent that it exists at all. What is important for my purposes is that the word for 'what' has the semantic type of a property or predicate.

My proposal that PAn 'who' and 'what' are bimorphemic is also in line with the widely held view that *wh*-words are interpreted in an operator-variable configuration. At least since Chomsky (1977), these operator-variable constructions have been viewed in Generative Grammar as being derived through movement of the *wh*-phrase to the clausal periphery. In the resulting configuration, the clause-initial *wh*-phrase functions as the operator, and the base position of the *wh*-phrase (indicated by the copy in angle brackets) within the clause is the variable bound by the *wh*-operator. The necessity for treating *wh*-questions as operator-variable constructions arises from the fact that interrogative pronouns are not referential but are rather quantifiers ranging over sets of possible answers to these questions, 'who' in English quantifying over sets of humans and 'what' quantifying over sets of things.

(10)   [CP What [C' did [TP you [VP do <what>]]]]?

Di Sciullo (2005) proposes that operator binding in English takes place at the word level. The interrogative [WH] feature in English interrogative pronouns like *who* and *what* is itself the operator which quantifies over the indefinite part of the word. See also Tsai (1994) for a word-internal analysis of operator-variable binding in English *wh*-words.

(11)   X
          ╱   ╲
      wh-   X
            ╱   ╲
          X   o/at

My proposal for PAn 'who' and 'what' is of this type, where the determiner and verbal prefix function as word-internal quantifiers. Unlike English, Austronesian interrogative pronouns are not marked uniformly as interrogative by a feature like [WH]. Consequently, I analyze the quantifiers in PAn interrogative pronouns as existential rather than interrogative. Quantification by the determiner yields a set of indefinite/unspecified individuals, while quantification by the existential verbal prefix yields a set of unspecified properties. See also Karttunen (1977), who proposes that *wh*-phrases are existentially quantified NPs.

(12)
```
              X
            /   \
        *a/*ma-   X
                /   \
               X    *nu
```

My analysis also entails that PAn *wh*-words were used not only interrogatively but also as indefinites. This is evidenced by the fact that interrogative pronouns can be used as indefinites in Formosan languages. (13a) shows Tanan Rukai *manema* 'what'[4] used as a negative polarity item in the scope of negation. In the scope of a modal, Tsou *cuma* 'what' expresses the indefinite 'something', as in (13b). Forms for 'who' also are indefinites, as shown for Seediq *ima* in (13c). When the triggers for non-interrogative interpretations like negation and modals are absent, the default interpretation is interrogative, as in (13d). I assume in this case that the pronoun is bound by a null interrogative operator base generated in the left periphery of the clause. Thus, the existential quantification within the pronoun merely serves to introduce a variable which then is bound by different operators in the clause in order to determine the interpretation of the pronoun, as proposed by Cheng (1991), among many others.

(13) a. Ko-ani   tama-tama   ka   kadrua thingal   sa   **manema**.
          NOM-this   man   TOP   not   know   INDEF   what
          'This man does not know anything.'   (Tanan Rukai)
      b. Mo   asonx   bonx   no   **cuma** ko'ko….
          AV   maybe   eat.AV OBL   what   so
          'Maybe he ate something, so….'   (Tsou; Chang 1998: 173)
      c. Ini=ku   sukuxun   **ani**   **ima**.
          NEG=1SG.NOM   like   even   who
          'I don't like anyone.'   (Seediq; Chang 1996: 75)
      d. Kusu   u-duri-duri-su   sa   **manema**?
          2SG.TOP   ACT-RED-plant-2S.NOM   OBL   what
          'What are you planting?' '   (Tanan Rukai)

There is also cross linguistic support for the proposal that different interrogative pronouns can be distinguished morphologically on the basis of different types of affixes and are not necessarily marked uniformly by an interrogative feature. In the Niger-Congo language Wolof, a noun class marker incorporates into the interrogative or indefinite morpheme. The word for 'who' includes the class marker for singular human nouns, while the word for 'what' incorporates one of the class markers for inanimate objects.

(14)   <u>Wolof</u> (Martinovic 2015)
      a.   k-an
              CM-Q
              'who'
      b.   l-an
              CM-Q
              'what'

This situation is mirrored in many Bantu languages, where the word for 'who' is marked for noun class 1, the class for singular human nouns, while 'what' belongs to a class for inanimate things.

There is also crosslinguistic evidence for a semantic and/or categorial difference between 'who' and 'what'. Wang (1958/2004) reconstructs Old Chinese interrogative pronouns with different initials, suggesting a difference in category.

            Old Chinese interrogative pronoun reconstructions (Wang 1958/2004: 333)

---

[4] In Rukai orthography, 'e' is used to transcribe schwa.

(15) a. [ʑ-] series: 誰 shéi [*ʑǐwəi] 'who', 孰 shú [*ʑǐwəuk] 'which'
      b. [ɣ-] series: 何 hé [*ɣa] 'what', 奚 xī [*ɣǐe] 'what'

The [ʑ-] initial in the words for 'who' and 'which' can be understood as an incorporated determiner. Note that the demonstrative 是 shì 'this' is reconstructed by Guo (1986: 57) with the same initial, specifically [*ʑǐe]. Furthermore, the word for 'who' has the distribution of a DP in that it can serve as a possessor followed by the genitive case marker.

(16) a. 臣殺其君，誰之過也？   (*Guoyu*, Lu 15.0)
          Chén   shā qí   jūn, shéi zhī guò yě?
          advisor kill 3.GEN lord who   GEN fault STAT
          'If an advisor kills his lord, whose fault is it?'
      b. 誰之劍也？   (*Zuozhuan*, Ding 10)
          Shéi zhī   jiàn yě?
          who   GEN   sword STAT
          'Whose sword is (this)?'

In contrast to this, the word for 'what' does not function as a possessor but attaches directly to a noun as a modifier.

(17) a. 何害也？   (*Zuozhuan*, Xiang 27)
          Hé   hài   yě?
          what harm STAT
          'What harm is there?'
      b. 是何物也？   (*Zuozhuan*, Zhao 21)
          Shì   hé   wù   yě?
          this what thing STAT
          'What thing is this?'

The different syntactic distributions of 'who' and 'what' in Old Chinese can be understood if 'who' is a D and 'what' is a common noun, which is a predicative category.

    Although I have not found direct evidence for the verbal status of 'what' cross linguistically, there is clear indication that words for 'what' have the semantic type of predicates or properties, while forms for 'who' are associated with individuals. For example, Heim (1987) shows that there is an asymmetry between *who* and *what* in English existential constructions such that *what* can be the pivot in an existential construction, while *who* cannot.[5]

(18) a.   What is there in Austin?
      b.   *Who was there in the room when you got home?   (Heim 1987: 27)

Building on Heim (1987), Nicolae & Scontras (2011, 2014) propose that movement of *who* leaves a trace of type <e>, while *what* leaves a trace of type <e,t>. In other words, after movement, *who* quantifies over sets of individuals, while *what* quantifies over sets of properties. Only a property can serve as the pivot in an existential construction. They further show that Tagalog interrogative pronouns mirror the behavior of English *wh*-words in existential constructions. The word for 'what' can serve as the pivot, as in (19a), while this is not possible for 'who' in (19b).

---

[5] Heim marks this sentence as marginally acceptable, but for me it is ungrammatical as an existential construction.

(19) Tagalog (Nicolae & Scontras 2011: 128)

a. **Ano** ang mayroon sa bahay ni Juan?
    what   NOM exist.there in house GEN.PN Juan
    'What is there in Juan's house?'

b. *__Sino__ ang mayroon sa bahay?
    who   NOM exist.there in house
    '*Who is there in the house?'

Note that 'what' can also serve as the pivot in an existential construction in Old Chinese, while this is unattested for 'who'.

(20) a. 由也果，於從政乎何有？ (*Analects*, Yongye)
    Yóu yě guǒ,   yú cóng zhèng   hū   **hé**   yǒu?
    Zi Lu TOP decisive   to work government EXCL what exist
    'Zi Lu is decisive, but what is there (about decisiveness) (to use) in government?'

b. 王如好貨，與百姓同之，於王何有？ (*Mencius*, Liang Hui Wang 2)
    Wáng rú hào huò, yǔ bǎi xìng   tóng zhī,
    king if like goods with commoner share 3.GEN
        yú wáng **hé** yǒu?
        to king what exist
    'If you (your majesty) like (to collect) goods, then if you share them with the people, what (problem) is there for you?'

Further evidence for the category and/or semantic type difference between 'who' and 'what' in Austronesian languages comes from the phenomenon of *wh*-incorporation. Lin (2013) clearly demonstrates that there is an asymmetry between 'what' and 'who' in terms of incorporation in Formosan languages. Specifically, 'what' can undergo incorporation, as in (21a), but 'who' cannot, as in (21b). The word for 'who' can only be used as a free form, as in (21c), showing dislocation of 'who' to clause-initial position. See additionally Chen (1999) for discussion of *wh*-incorporation in Rukai.

(21) a. Mi-**maan** ci-panay?
    AV.do-what NOM.PN-Panay
    'What is Panay doing?' (Amis; Lin 2013: 91)

b. *Ma-**cima**=tu ku   tayni-ay?
    AV-who=PFV NOM.CN come-FAC
    'Who has come?' (Amis; Lin 2013: 89)

c. **Cima**=tu ku   tayni-ay?
    who=PFV NOM.CN come-FAC
    'Who has come?' (Amis; Lin 2013: 89)

Lin adopts Nicolae & Scontras' (2011) analysis that 'who' is type <e>, while 'what' is type <e,t>. This accounts for the incorporation asymmetry, since it is well known that only bare common nouns or NPs, which are type <e,t>, can undergo incorporation (Mithun 1984, Baker 1988, Gerdts 2001, Massam 2001, and others).

   A final argument for the different semantic types of 'who' and 'what' comes from an observation by Idiotov (2007). In many languages, the distinction between these pronouns lies less in animacy and more in specificity. When questioning individual entities – regardless of their animacy – 'who' is used. A commonly observed phenomenon is the use of 'who' to ask for a name, which can be the name of a person, as in the Uralic language Avam Nganasan in (22a), or the name of a non-person, e.g. a village, as in the Bantu language Ngome in (18b).

(22) a. siļi/*maa   n'im-ti? (Avam Nganasan; Valentin Goussev, p.c.; Idiotov 2007: 73)
          who/*what   name-3SG
          'What is his name? (lit. 'Who is his name?')'
      b. ngando   íyĕ   nda?   (Ngome; Rood 1958:xxi; Idiotov 2007: 77)
          CM9.village CM9.this   who
          'What is the name of this village? (lit.: 'Who is this village?')'

Blust (2009) has identified the same phenomenon in many Austronesian languages like Standard Indonesian. The word for 'who' rather than 'what' is used to question a person's name.

       Standard Indonesian (Blust 2009: 510)
(23) Siapa   nama   anda?
      who   name   2SG
      'What is your name? (lit. 'Who is your name?')'

In contrast, when asking for a kind or classification, then 'what' is used, as in the Northern Tungusic language Poligus Evenki. What is questioned here is the clan that the listener belongs to and not the specific identity of the individual.

(24) e:kun bi-si-nni?   bi   Kurkogir   bi-si-m
      what be-PRS-2SG   1SG   PROP   be- PRS-1SG
      'What are you? I am a Kurkogir (i.e., I belong to the Kurkogir clan, the name of my family is Kurkogir)'
      (Poligus Evenki; Konstantinova 1968:73; Idiotov 2007: 53)

Another example of the more specific nature of 'who' comes from its interchangeability with the selective interrogative pronoun 'which' in some languages. Citing Holst (2001: 131) and Nau (1999: 145-147), Idiotov (2007: 9) points out that the non-selective interrogative pronoun *kas* can be used for both 'who' and 'what' in Latvian. However, the selective interrogative pronoun *kurš* can also be used non-selectively, in which case it can only mean 'who' and not 'what'.

       On Idiotov's (2007) analysis, this difference between 'who' and 'what' is captured in terms of the features [IDENTIFICATION] for 'who' and [CLASSIFICATION] for 'what'. The types of sets each is associated with is given in parentheses. The set for 'who' is the set of proper names, i.e. individual entities, which in my analysis is type <e>, while 'what' is associated with sets of common nouns, i.e. classes or properties, which are formally type <e,t> on the analysis I adopt.

(25) a. WHO → [PERSON + IDENTIFICATION (+proper name)]
      b. WHAT → [INANIMATE THING + CLASSIFICATION (+common noun)] (Idiotov 2007: 19)

This section has presented morphosyntactic evidence for my reconstructions of 'who' and 'what' in PAn. I have given arguments from both Austronesian languages and languages outside this family for reconstructing 'who' as a D associated with sets of specific individuals and 'what' as a predicate ranging over sets of properties. In the next section, I show how the synchronic variation among extra-NucAn languages can be accounted for by natural phonological and morphological changes.

3. Accounting for the Variation in Extra-Nuclear Austronesian Languages

In this section, I account for the synchronic variation among the forms for 'who' and 'what' in the extra-NucAn languages Rukai, Tsou, and Puyuma.

(26)   Austronesian

  Rukai   Ergative An

        Tsou   **Puyuma**   Nuclear An
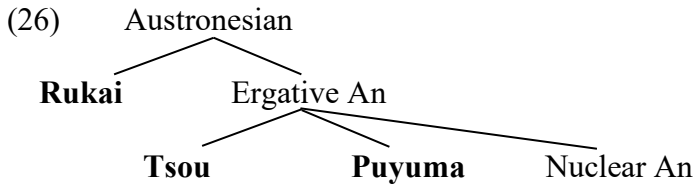
My reconstructions for PAn are repeated below.

		PAn	
(27)	a.	*a-nu	WHO (D+INDEF = 'some/which one')
	b.	*ma-nu	WHAT (V+INDEF = 'some/what kind')

I argue that the forms observed in Rukai, Tsou, and Puyuma can be derived straightforwardly from these PAn reconstructions through a limited set of natural changes. On the phonological level, these are all processes which have been argued to exist in each of the languages in question. In addition, I propose two morphological changes. One is the reanalysis of the word for 'what' as a new indefinite base, after which it came to be used to express both 'what' and 'who'. The other morphosyntactic change is the innovation of new operator prefixes, which I assume was motivated by the loss of transparency in the morphological composition of the pronoun in question after it underwent sound changes or reanalysis as the new indefinite base. After discussing extra-NucAn languages in this section, I briefly show how the same processes can account for subsequent changes in NucAn languages in section 4.

3.1.Rukai

I propose that Proto-Rukai inherited the PAn words in reduplicated form with a reduced vowel in the second syllable.

		Proto-Rukai	
(28)	a.	*anə-anə	WHO
	b.	*manə-manə	WHAT

I have not identified a motivation for the reduplication, but reduplication is a common process in this and other Austronesian languages. Furthermore, all of the dialects shown in (29) reflect it, so it must be attributed to the proto-language. My reconstructions are also consistent with those proposed by Li (1973).

		WHO	WHAT
(29)	Budai	*anə-anə*	*manə-manə*
	Tanan	*anə-a*	*manə-ma*
	Labuan	*anə-a*	*manə-ma*

Regarding the reduction of the stem-final vowel, the change from PAn *-u to Proto-Rukai *-ə is reasonable, given that stress in PAn would have been on the first (penultimate) syllable, as argued recently by Smith (2023), in accordance with Blust (1997, 2013) and others. Li (1973: 21) also points out that, although stress in Tanan Rukai is usually on the final syllable, it surfaces on the penultimate syllable in the case of interrogative pronouns.

    In the forms in Tanan and Labuan, the final syllable /nə/ has been deleted. A motivation for the loss of /nə/ in Tanan and Labuan is suggested by Li (1973, 1977), who proposes that word-final *n(ə) following /a/ was lost in Tanan. This process also accounts for the change in the nominalizing suffix from Proto-Rukai *-anə to -*a* in Tanan.

3.2. Tsou

My reconstructions for Proto-Tsou are shown below.

<pre>
                    Proto-Tsou
(30)   a.   *si-ã (< *si + *anə)   WHO
        b.   *cu-mã (< *Cu + *manə)   WHAT
</pre>

I propose that these proto-forms reflect the attachment of new operator prefixes – both case markers – concomitant with the loss of the final two segments /nə/. Regarding how this syllable was lost, I assume first with Wright (1996: 54) that vowels in unstressed syllables in Tsou tend to be reduced, and the final syllable /nə/ would have been unstressed, as discussed in the preceding subsection. I next follow Tsuchida (1975: 216), who argues that PAn *a coalesced with a following /n/ or /nə/ in word-final position, resulting in a word-final nasalized /ã/. Given that stress in Tsou nearly always falls on the penultimate syllable (Wright 1996: 50), I assume that the stress shifted to the preceding syllable after loss of the final /nə/. The stress shift is an important step in the historical development, since Tsuchida (1975) also shows that PAn stressed *a is reflected as /o/ in Tsou, while unstressed nasalized *ã[6] is reflected as /a/. The forms in modern Tsou are *sia* 'who' and *cuma* 'what'.

The next question to be addressed is the source of the prefixes shown in (30). I suggest that these derive historically from case markers, nominative for 'who' and non-nominative for 'what'. One of the several nominative case markers in Tsou is *si*, as shown in (31a). Tsou case markers encode deictic information such as distance from the speaker and also visibility. The deictic information conveyed by *si* is that the referent of the argument is visible and only somewhat distant from the interlocutors.

<pre>
                    Tsou (Chang & Pan 2018: 34, 45, 46)
(31)   a.   Mi=cu   bxvnx   si   'ume.
                    AV=PFV   bloom   NOM.VIS.MED   plum
                    'The plums have bloomed.'
        b.   Zou   skayavangx   na   sico.
                    be   Atayal   NOM   DEM.MED
                    'That (person) is Atayal.'
        c.   I-si   nocmu'-i   to   oko   'o   peisu-si.
                    NAV-3SG.IVIS   loose-APPL   OBL   child NOM   money-3SG.IVIS
                    'The child lost his/her money.'
</pre>

Before turning to the form for 'what', one more matter regarding *si* should be clarified. This marker is homophonous with the nominative case marker for personal names *si in Proto-NucAn (PNucAn, discussed in section 1 and again in section 4). However, the person marking function had not yet developed in Proto-Tsou, since Tsou case marking makes no distinction on the basis of person. Consequently, *si* in Tsou cannot be a reflex of *si NOM.PN. I assume that the Tsou case marker grammaticalized from a demonstrative. Note that Tsou also has a demonstrative *sico/sica* 'that one', which is used to indicate referents medium distance from the speaker, as shown in (31b). I assume that the first syllable is historically a demonstrative *si and that the case marker *si* grammaticalized from this demonstrative. Additional evidence that Tsou historically had a demonstrative *si comes from the 3rd person pronominal form *si*, as shown in (31c), which can also reasonably be assumed to have grammaticalized from a demonstrative. If PAn also had this demonstrative, then it may have been the diachronic source for the NOM.PN *si in NucAn languages, which would mean that Tsou *si* NOM.VIS.MED is cognate with *si in PNucAn, but Tsou does not reflect the person marking function of *si.

Turning now to *cu-* in *cuma* 'what', there is no object case marker *cu* in modern Tsou, but I suggest that this might be a reflex of Ross' (2006) *Cu, *C being a dental affricate, object case marker in PAn.

Regarding the motivation for prefixing case markers to the interrogative pronouns, I assume that this was the result of incorporation of *ã and *mã into the case markers after the loss of the original root /nə/. There is

---

[6] The parent language that Tsuchida attributes the nasalization to is Proto-Tsouic, a subgroup consisting of Tsou, Kanakanavu, and Saaroa. The status of this subgroup is controversial, having been brought into question by Chang (2006), Ross (2009, 2012), Zeitoun and Teng (2016), and others. In light of this controversy, I only consider historical developments which can clearly be attributed to Tsou.

also a ready syntactic environment for this incorporation to take place, since Proto-Tsou was surely a *wh*-in-situ language, as evidenced by the modern Tsou examples in (32). The frequent occurrence of 'who' in subject position and 'what' in object position would have resulted in the attachment of *ã and *mã, respectively, to the nominative *si and non-nominative *Cu case markers. The incorporation may also have been facilitated by phonological pressure to create bisyllabic forms.

(32) Tsou (Chang 1998: 102)
  a. I-si eobaka-a **no** **sia** 'e mo'o?
      NAV-3SG hit-PV OBL who NOM Mo'o
      'Who hit Mo'o?'
      Lit: 'Mo'o was hit by whom?'
  b. Te-ta mo-si ta pangka **no** **cuma** 'e pasuya?
      MOD-3SG AV-put OBL table OBL what NOM Pasuya
      'What will Pasuya put on the table?'

Incidentally, Rukai is also a *wh*-in-situ language.

(33)
  a. W-a-kela ku **ane-ane**.
      ACT-NONFUT-come NOM who
      'Who came?'   (Budai Rukai; Chen 1999: 44)
  b. Li-senay-su ku **mane-mane**.
      FUT-sing-2SG.NOM OBL what
      'What will you sing?'   (Budai Rukai; Chen 1999: 40)

In the next section, I additionally show examples of Puyuma *wh*-in-situ. *Wh*-in-situ is also widely found in NucAn languages. Given that *wh*-in-situ is found in every high-order Austronesian subgroup, this syntactic characteristic must also be reconstructed to PAn.

3.3. Puyuma

I propose that Proto-Puyuma inherited *manu 'what' directly from PAn, but it was extended to be used for 'who' as well as 'what'. Furthermore, I suggest that there was variation in Proto-Puyuma, allowing for the reduction of the final vowel in some dialects. I assume that this variation dates to Proto-Ergative Austronesian, Proto-Tsou inheriting only *manə.

       Proto-Puyuma
(34)   *manu/*manə WHO/WHAT

The evidence for final /u/ comes from the Ulribulribuk dialect, where *manu* is found for 'who' (Ting 1978: 385). Ting gives no form for 'what', but the forms for 'who' and 'what' are identical in all the other dialects, so I assume that this form is used for both 'who' and 'what' in Ulribulribuk as well. The lack of evidence for final /u/ in Nanwang and Katripul suggests that they inherited the reduced variant *manə.

    This reconstruction allows for a straightforward account of the Nanwang and Katripul forms. Puyuma dialects differ from PAn in having final stress (Tsuchida 1980; Teng 2008, 2018). I suggest that the loss of the final vowel in Katripul was a strategy for accommodating the stress. Although schwa can accept stress in modern Puyuma, Smith (2023) argues that this was not possible for PAn. I suggest that Proto-Puyuma was closer to PAn in this regard.

(35)
  a. Katripul   *əman*
  b. Nanwang   *manay*
  c. Ulribulribuk *manu*

Turning next to Nanwang, I suggest that the source for the nuclear vowel /a/ in the final syllable was vowel harmony. According to Teng (2008), schwa in Nanwang often takes on the quality of a vowel in a neighboring syllable. As for the final glide, I suggest that it may have been the third person pronoun *i or *ia (see B&T for reconstruction). This is suggested by a parallel in some Rukai dialects. Mantauran, Maga, and Tona do not have forms specifically for 'who' and 'what'. Rather, a cleft construction is used in which the presupposition is packaged as a headless relative clause, while the focused constituent is expressed as a noun possessed by a third person pronoun. In Mantauran, for example, the word for 'name' is suffixed with the pronoun -i in order to ask a 'who' type *wh*-question. Literally, this sentence means, "What is their name, the one who will cook tonight?"

Mantauran Rukai

(36) inay   maongo   anga-i   [ta-mo-aha'a   adhamay]?
      this   night   name-3SG.INDEF   NMLZ-ACT-cook   meal
      'Who (is it who) is going to cook supper tonight?'

I suggest that the Nanwang -*y* may have a similar origin. See also Idiatov (2007: 396-398) for another proposal that reconstructs interrogative pronouns in some Formosan languages as having an incorporated pronoun, though the forms he proposes are quite different from mine. Space limitations prevent a detailed discussion of his reconstructions here.

The motivation for the merger of the two interrogative pronouns into a single form is difficult to pinpoint. However, the homophony between 'who' and 'what' necessitated the creation of new operators in order to distinguish them. This role was taken on by case markers, the distinction between type <e> 'who' and type <e,t> 'what' being accomplished by the person distinction in the case marking system in Proto-Puyuma. The following examples show the opposition between personal and non-personal markers for nominative case in Nanwang, which is similar in the other dialects of Puyuma. Clearly, the personal case marker indicates arguments which refer to specific individuals, i.e. referents of type <e>, while the non-personal marker selects common nouns or NPs, which have the semantic type of properties or predicates, i.e. <e,t>.

Nanwang Puyuma

(37) a.   tr<em>akaw   dra   paisu **i**   **isaw**
            <AV>steal   INDF.OBL   money SG.NOM.PN   Isaw
            'Isaw stole money.'   (Puyuma; Teng 2008: 147)
      b.   drua-drua   me-na'u-a   **a**   **trau**
            RED-come   INTR-see-SBJV INDEF.NOM   person
            'Many people came to see.'   (Nanwang Puyuma; Teng 2008: 113)

With these semantic types in mind, when the case markers attach to interrogative pronouns, they function as operators, picking out members of sets of individuals in the case of personal markers and picking out members of sets of properties in the case of non-personal markers. These functions then completely mirror those I proposed for PAn *a and *ma- in section 2. Given that the case markers can distinguish the forms for 'who' and 'what', there is no need for distinct base forms, facilitating merger in order to simplify the pronominal system. However, case markers have not been incorporated into the interrogative pronouns, which I assume is due, at least in part, to the fact that the interrogative pronoun is bisyllabic.

Nanwang Puyuma (Teng 2008: 224-225)

(38) a.   mi-walak   **i**   **manay**?
            have-child   SG.NOM.PN   WH
            'Who gave birth to a child?'
      b.   **a**   **manay**   tu=edad   kana   kutrem?
            INDEF.NOM.CN   WH   3.SG=color   DEF.OBL   cloud
            'What color is the cloud?'

As to why it was *manu/*manə 'what' that became the new indefinite instead of *anu/*anə 'who', I presently have no clear answer to this question but suggest that both possibilities existed in Proto-Ergative Austronesian. Note that the NucAn language Kavalan extended *anə: *ti-ana* 'who' and *ni-ana* 'what' (Hsieh 2018: 111-115), *ti* and *ni* both being case markers.
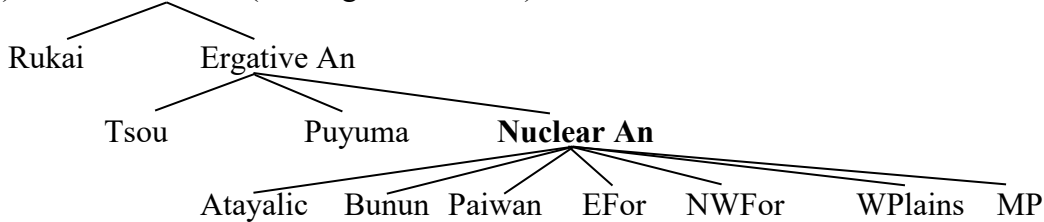
In this section, I have accounted for the developments of PAn interrogative pronouns in extra-NucAn languages through straightforward diachronic processes. The sound changes I identified are all compatible with known synchronic or diachronic phonological processes in the languages in question. The morphological changes I proposed are equally supported by the available evidence and natural processes of morphosyntactic change. One is renewal of the operators quantifying over the indefinite pronouns in response to opacity in the morphological composition of the interrogative pronouns. The presence of such operators is justified by the need to distinguish type <e> 'who' from type <e,t> 'what', as I argued in section 2. Another response to the opacity in the forms was the merger of the functions of 'who' and 'what' into a single form, the new indefinite base, from the erstwhile 'what'. This change also necessitated the innovation of new operators, which were available in the case marking system.

In the next section, I propose reconstructions for PNucAn 'who' and 'what' and show how similar diachronic processes account for the developments of these forms in NucAn languages.

4. Proto-Nuclear Austronesian Reconstructions

This section briefly considers the forms for 'who' and 'what' in NucAn languages. Lack of space prevents a detailed exposition of how the forms developed in each language, so I concentrate only on the general patterns which guide my reconstructions.

(39)   Austronesian (Aldridge 2016, 2021)

Rukai   Ergative An

      Tsou   Puyuma   **Nuclear An**

           Atayalic   Bunun Paiwan   EFor   NWFor   WPlains   MP

The analysis proposed in 3.3 for Puyuma suggests a straightforward explanation for the forms found in NucAn languages. First, note that the forms for both 'who' and 'what' in NucAn languages (shown in bold in 40) contain the syllable /ma/. This can be accounted for if the new indefinite form *manə reflected in Puyuma was reduced to a monosyllabic form after the attachment of new case markers, nominative personal case markers for 'who' and non-nominative common noun case markers for 'what'. The case marker attaching to form 'who' is *i NOM.PN, which was inherited from Proto-Ergative Austronesian. This is clear from the fact that *i is also reflected in Puyuma, as shown in (37a) and (38a). I assume that the truncation of *manə to /ma/ took place after the case markers were reanalyzed as part of each of the words. The motivation could have been the creation of a disyllabic word or simply the loss of word-final syllables with weak vowels.

(40)		WHO	WHAT	
	Budai Rukai	anə-anə	manə-manə	(Chen 1999)
	Tsou	si-a	cu-ma	(Chang & Pan 2018)
	Nanwang Puyuma	manay	manay	(Teng 2008)
	Truku	ima	manu	(Li & Hsu 2018)
	Paiwan	ima	ne-ma	(H.-C. Chang 2018)
	Thao	ti-(i)ma	nu-ma	(Jean 2018)
	Isbukun Bunun	si-(i)ma	ma-az	(Huang & Shih 2018)
	Takivatan Bunun	si-maq	maq	(De Busser 2009)
	Amis	ci-(i)ma	ma-an	(Wu 2018)
	Siraya	ti-(i)mang	mang	(Lee 2017)

After the innovation of *ima 'who', which is reflected directly in Paiwan and Atayalic languages like Truku, new forms for 'who' were created by attachment of the new personal nominative case marker *si, reflected as *si*, *ti*, and *ci* in different NucAn languages. Recall from section 1 that Blust (2009) and B&T propose that *si attached to *ima in numerous languages through drift. I adopt this proposal, with the caveat that this innovation took place only in NucAn languages, since it is not reflected in extra-NucAn languages (the unbolded languages in 35). As I pointed out in section 3.2, *si-* in Tsou is not a marker for personal nouns. Though I assume it to be related to the personal nominative case marker *si, Tsou reflects an earlier stage in this grammaticalization process and does not reflect the personal nominative marker directly.

As for *ima 'who', I propose that this is a second subgroup defining innovation for NucAn. But this proposal depends on demonstrating that *ima was the base to which *si attached in the NucAn languages other than Paiwan and Atayalic languages. These other NucAn languages seem to reflect the attachment of the personal nominative case marker *si directly to *ma. However, it is also possible, as Blust (2009) and B&T propose, that *si attached to *ima, followed by reduction of the identical vowel sequence. There is in fact some evidence in Paiwan that *ima was the base. Ferrell's (1982: 100) dictionary lists the reduplicated form *t(i)-ima-ima* meaning 'several persons'. The second occurrence of the pronoun shows *ima* since the pronoun is now not in the environment for deletion of the initial vowel. If the base form of the pronoun were /ma/, then the reduplicated form would be expected to surface as *ti-ma-ma*. This provides evidence that *si attached to *ima and not *ma, which in turn means that a reflex of *ima is found in all daughters of PNucAn. Consequently, the creation of *ima can be considered to be an innovation identifying the NucAn subgroup.

To summarize, my reconstructions for PNucAn are shown below. The new forms for 'who' and 'what' were both created by adding case markers to the indefinite base *manə. Adding the personal case marker *i resulted in the form for 'who', while language specific object marking was attached to form 'what'. A wide variety of affixes in the forms for 'what' can be seen in the table in (35). Space limitations prevent a detailed discussion of the developments of these forms, so this topic must be left to future research.

Proto-Nuclear Austronesian
(36)   *manə   WHO/WHAT
        *i-ma (< *i PN.NOM + manə)   WHO
        *ACC-ma   WHAT

5. Conclusion

This paper has proposed reconstructions for Proto-Austronesian 'who' and 'what' from which the forms found in modern Formosan languages can be straightforwardly derived. First, I argued that all of these forms consist of an operator attached to an indefinite base, a proposal which receives broad cross linguistic support. I also showed that 'who' and 'what' are associated with different semantic types. As is the case cross linguistically, 'who' is associated with sets of individuals, while 'what' is associated with sets of properties. The diverse array of forms found in Formosan languages can then be accounted for via renewal of these word-internal operators from case markers in response to the opacity in the earlier composition induced by phonological and morphosyntactic changes taking place in the respective languages. This proposal not only accounts for a greater number of forms than earlier approaches by Blust (2009) and Blust & Trussel (ongoing), it additionally is able to identify a new innovation defining the Nuclear Austronesian subgroup first proposed by Ross (2009).

References

Aldridge, Edith. 2016. Ergativity from subjunctive in Austronesian languages. *Language and Linguistics* 17.1: 27-62

Aldridge, Edith. 2021. Syntactic conditions on accusative to ergative alignment change in Austronesian languages. *Journal of Historical Linguistics* 11.2:214-247. https://doi.org/10.1075/jhl.20016.ald

Aldridge, Edith, and Yuko Yanagida. 2021. Two types of alignment change in nominalizations. *Diachronica* 38.3: 314-357

Baker, Mark C. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press

Blust, Robert. 1997. Rukai stress revisited. *Oceanic Linguistics* 36.2: 398-403

Blust, Robert. 1999. Subgrouping, circularity and extinction: Some issues in Austronesian comparative linguistics. In *Selected papers from the eighth International Conference on Austronesian Linguistics* (Symposium Series of the Institute of Linguistics), ed. Elizabeth Zeitoun, and Paul Jen-kuei Li, 31-94. Taipei: Institute of Linguistics, Academia Sinica

Blust, Robert. 2009/2013. *The Austronesian languages*. Canberra: Research School of Pacific and Asian Studies (2nd edn. College of Asia and the Pacific), Australian National University

Blust, Robert, and Stephen Trussel. ongoing. The Austronesian comparative dictionary. https://www.trussel2.com/ACD/. Accessed 23 August 2023

Chang, Chung-liang. 1996. *A study of Seediq interrogatives* [in Chinese]. MA thesis, National Tsinghua University, Hsinchu, Taiwan

Chang, Hsiu-chuan. 2018. *A sketch grammar of Paiwan* [in Chinese]. Series on Formosan Languages, 13. Taipei: Council of Indigenous Peoples

Chang, Melody Ya-Yin. 1998. WH-*constructions and the problem of* WH-*movement in Tsou*. MA Thesis, National Tsinghua University, Hsinchu, Taiwan

Chang, Henry Y. 2006. Rethinking the Tsouic subgroup hypothesis: A morphosyntactic perspective. In *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*, ed. Henry Y. Chang, Lillian M. Huang, and Dah-an Ho, 565-593. Taipei: Institute of Linguistics, Academia Sinica

Chang, Yung-li, and Chia-jung Pan. 2018. *A sketch grammar of Tsou* [in Chinese]. Series on Formosan Languages, 13. Taipei: Council of Indigenous Peoples

Chen, Cheng-Fu. 1999. Wh-*words as interrogatives and indefinites in Rukai*. MA Thesis, National Taiwan University

Cheng, Lisa Lai-Shen. 1991 *On the typology of wh-questions*. Dissertation, Cambridge, MA: Massachusetts Institute of Technology

Chomsky, Noam. 1977. On *wh*-movement. In *Formal syntax*, ed. Peter W. Culicover, Thomas Wasow, and Adrian Akmajian, 71-132. New York: Academic Press

De Busser, Rik L. J. 2009. *Towards a grammar of Takivatan Bunun: Select topics*. Dissertation, La Trobe University

Di Sciullo, Anna Maria. 2005. *Asymmetry in morphology*. Cambridge, MA: MIT Press

Ferrell, Raleigh. 1969. *Taiwan aboriginal groups: Problems in cultural and linguistic classification*. Taipei: Institute of Ethnology, Academia Sinica

Gerdts, Donna B. 2001. Incorporation. In *Handbook of morphology*, ed. Andrew Spencer, and Arnold M. Zwicky, 84-100. Oxford: Blackwell Publishing

Guo, Xiliang. 1986. *Hanyu gu yin shouce* [Handbook of Old Chinese phonology]. Beijing: Beijing University Press.

Heim, Irene. 1987. Where does the definiteness restriction apply? Evidence from the definiteness of variables. In *The representation of (in)definiteness*, ed. Eric Reuland, and A.G.B. ter Meulen, 21-42. Cambridge, MA: MIT Press

Holst, Jan Henrik. 2001. *Lettische grammatik*. Hamburg: Helmut Buske.

Hsieh, Fu-Hui. 2018. *A sketch grammar of Kavalan* [in Chinese]. Series on Formosan Languages, 13. Taipei: Council of Indigenous Peoples

Huang, Hui-chuan, and Chao-Kai Shih. 2018. *A sketch grammar of Bunun* [in Chinese]. Series on Formosan Languages, 13. Taipei: Council of Indigenous Peoples

Idiatov, Dmitry. 2007. A typology of non-selective interrogative pronominals. Dissertation, University of Antwerp. http://idiatov.mardi.myds.me/papers/PhD_Idiatov_2007_Final_Version+ERRATA.pdf. Accessed 29 May 2023

Jean, Shih-lang. 2018. *A sketch grammar of Shao* [in Chinese]. Series on Formosan Languages, 13. Taipei: Council of Indigenous Peoples

Kaufman, Daniel. 2009. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35.1: 1-49

Kaufman, Daniel. 2011. Exclamatives and temporal nominalizations in Austronesian. In *Nominalization in Asian languages: Diachronic and* typological *perspectives*, ed. Foong Ha Yap, Karen Grunow-Hårsta, and Janick Wrona, 721-754. Amsterdam & New York: John Benjamins

Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1: 3-44

Konstantinova, O.A. 1968. Эвенкийский язык [Evenki]. In Viktor V. Vinogradov et al. (ed.), Языки народов СССР [Languages of the peoples of the USSR], Vol. 5: Монгольские, тунгусо-маньчжурские и палеоазиатские языки [Mongolic, Tungus-Manchu and Paleoasiatic languages], 68-87. Leningrad: Nauka.

Lee, Hsu-Fen. 2017. *A study on nominalization in Siraya*. MA Thesis, National Kaohsiung University

Li, Paul Jen-kuei. 1973. *Rukai structure.* Taipei: Academia Sinica Institute of History and Philology

Li, Paul Jen-kuei. 1977. The internal relationships of Rukai. *Bulletin of the Institute of History and Philology* 48: 1-92

Li, Pei-jung, and Lowking Wei-Cheng Hsu. 2018. *A sketch grammar of Truku* [in Chinese]. Series on Formosan Languages, 13. Taipei: Council of Indigenous Peoples

Lin, Dong-yi. 2013. *Interrogative constructions in Kavalan and Amis*. Dissertation, University of Florida

Martinović, Martina. 2015. *Feature geometry and head-splitting: Evidence from the morphosyntax of the Wolof clausal periphery*. Dissertation, University of Chicago

Massam, Diane. 2001. Pseudo noun incorporation in Niuean. *Natural Language and Linguistic Theory* 19: 153-197

Mithun, Marianne. 1984. The evolution of noun incorporation. *Language* 60.4: 847-894

Nau, Nicole. 1999. Was schlägt der Kasus? Zu Paradigmen und Formengebrauch von Interrogativpronomina. *Sprachtypologie und Universalienforschung* 52(2): 130-150.

Nicolae, Andreea C., and Scontras, Gregory. 2011. How does who compose? In *Proceedings from AFLA 18*, ed. Lauren Eby Clemens, Gregory Scontras, and Maria Polinsky, 125-139. Cambridge, MA: Harvard University

Rood, N. 1958. *Ngombe-Nederlands-Frans woordenboek* (Dictionnaire ngombe-néerlandais-français). Tervuren: Musée royal du Congo belge.

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-Kuei Li on his 70th birthday*, ed. Henry Y. Chang, Lillian M. Huang, Dah-an Ho, 521-563. Taipei: Language and Linguistics

Ross, Malcolm. 2009. Proto Austronesian verbal morphology: A reappraisal. In *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, ed. Alexander Adelaar, and Andrew Pawley, 295-326. Canberra: Pacific Linguistics

Ross, Malcolm. 2012. In defense of Nuclear Austronesian (and against Tsouic). *Language and Linguistics* 13.6: 1253-1330

Shih, Cindy Peiru. 2008. *Interrogative constructions in Plngawan Atayal*. M.A. thesis, National Taiwan Normal University

Smith, Alexander. 2023. Reconstructing non-contrastive stress in Austronesian and the role of the mora in stress shift, gemination, and vowel shift. *Diachronica* 40.1: 111-152

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

Ting, Pang-hsin. 1978. Reconstruction of Proto-Puyuma phonology. *Bulletin of the Institute of History and Philology Academia Sinica* 49.3: 321-392

Tsai, W.-T. Dylan. 1994. *On economizing the theory of A-bar dependencies*. Dissertation, MIT

Tsuchida, Shigeru. 1975. *Reconstruction of proto-Tsouic phonology*. Dissertation, Yale University

Tsuchida, Shigeru. 1980. Puyuma (Tamalakaw dialect) vocabulary. In *Kuroshio no minzoku, bunka, gengo*, ed. Eikichi Ishikawa, 183-307. Tokyo: Kadokawa

Wang, Li. 1958. *Hanyu shi gao* [Lectures on the history of the Chinese language]. Beijing: Zhonghua Publishing. Reprinted in 2004. Beijing: Zhonghua Shuju.

Wright, Richard Albert. 1996. *Consonant clusters and cue preservation in Tsou*. Dissertation, UCLA

Wu, Jing-lan. 2018. *A sketch grammar of Amis* [in Chinese]. Series on Formosan Languages, 13. Taipei: Council of Indigenous Peoples

Zeitoun, Elizabeth, and Lillian M. Huang. 2000. Concerning ka-, an overlooked marker of verbal derivation in Formosan languages. *Oceanic Linguistics* 39.2: 391-414

Zeitoun, Elizabeth, and Stacy F. Teng. 2016. Reassessing the position of Kanakanavu and Saaroa among the Formosan languages. *Oceanic Linguistics* 55.1: 162-198

Edith Aldridge is a Research Fellow at Academia Sinica. Her research focuses on comparative and diachronic syntax, with language concentrations in Austronesian languages, Chinese, and Japanese. Her recent publications in historical linguistics include 'The typology and diachrony of Philippine-type alignment' (to appear, *Move and Agree*, JB), 'Alignment change' (To appear, *The Wiley-Blackwell Companion to Diachronic Linguistics*), 'Syntactic conditions on accusative to ergative alignment change in Austronesian languages' (2021, *Journal of Historical Linguistics*) 'The nature and origin of syntactic ergativity in Austronesian languages' (2021, *The Nature of Syntactic Change: Formal Approaches*, OUP), 'The syntax of pronoun fronting in Late Archaic Chinese negated clauses' (2021, *Journal of East Asian Linguistics*), 'Two types of alignment change in nominalizations' (with Yuko Yanagida, 2021, *Diachronica*), 'Subject/non-subject movement asymmetries in Late Archaic Chinese' (2019, *Glossa*), 'C-T Inheritance and the Left Periphery in Old Japanese' (2018, *Glossa*), and 'Intransitivity and the development of ergative alignment' (2017, *The Oxford Handbook of Ergativity*, OUP).