

# Is Malayo-Polynesian a primary branch of Austronesian?

## A view from morphosyntax

Victoria Chen,<sup>1</sup> Jonathan Kuo,<sup>2</sup> Maria Kristina S. Gallego<sup>3,4</sup> and Isaac Stead<sup>5</sup>

<sup>1</sup>Victoria University of Wellington | <sup>2</sup>National Taipei University of Technology | <sup>3</sup>Australian National University | <sup>4</sup>University of the Philippines Diliman | <sup>5</sup>Max Planck Institute for Evolutionary Anthropology

An understudied morphosyntactic innovation, reanalysis of the Proto-Austronesian (PAN) stative intransitive prefix \**ma-* as a transitive affix, offers new insights into Austronesian higher-order subgrouping. Malayo-Polynesian is currently considered a primary branch of Austronesian, with no identifiably closer relationship with any linguistic subgroup in the homeland (Blust 1999, 2009/2013; Ross 2005). However, the fact that it displays the same innovative use of *ma-* with Amis, Siraya, Kavalan and Basay-Trobiawan and shares the merger of PAN \*C/t with this group suggests that Malayo-Polynesian and East Formosan may share a common origin – the subgroup that comprises the four languages noted above. This observation points to a revised subgrouping more consistent with a socio-historical picture where the out-of-Taiwan population descended from a seafaring community expanding to the Batanes and Luzon after having developed a seafaring tradition. It also aligns with recent findings in archaeology and genetics that (i) eastern Taiwan is the most likely starting point of Austronesian dispersal (Hung 2005, 2008, 2019; Bellwood 2017; Bellwood & Dizon 2008; Carson & Hung 2018) and (ii) that the Amis bear a significantly closer relationship with Austronesian communities outside Taiwan (Capelli et al. 2001; Trejaut et al. 2005; McColl et al. 2018; Pugach et al. 2021; Tätte et al. 2021). Future investigation of additional shared innovations between Malayo-Polynesian and East Formosan could shed further light on their interrelationships.

**Keywords:** Austronesian higher-order subgrouping, Malayo-Polynesian, East Formosan, syntactic reanalysis, Austronesian diaspora, Out-of-Taiwan Hypothesis, triangulation

## 1. Introduction

There has been a consensus in the literature that all Austronesian languages spoken outside Taiwan belong to a single subgroup, Malayo-Polynesian (MP). This subgroup is defined by a series of innovations shared across extra-Formosan languages (e.g., Mills 1975; Dahl 1973; Blust 1977, 1993a, 1999, 2001, 2009/2013; Liao 2011; Ross 2002, 2005, 2012; a.o.).<sup>1</sup>

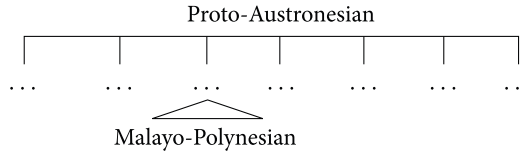
- (1) Major innovations defining Malayo-Polynesian<sup>2</sup>
  - a. PAn \*C/t > PMP \*t (merger)
  - b. PAn \*N/n > PMP \*n (merger)
  - c. PAn \*S/h > PMP \*h (merger; with irregular loss of PAn \*s > zero)
  - d. PAn \*l > R/\_Vj
  - e. PAn \*mu '2PL.GEN' > PMP \*mu '2SG.GEN'
  - f. Morphological innovations: PMP \*maN- 'Actor Voice morpheme', \*paN- 'distributive', \*paR- 'durative, reciprocal', \*maR- 'Actor Voice morpheme'
  - g. Metatheses: PAn \*-CVS > variation between \*-hVC and \*-CVh (with \*S > h)
  - h. Irregular changes: PAn \*biRbiR > PMP \*bibiR 'lips'; PAn \*Siwa > PMP \*siwa 'nine'; PAn \*paNudaN > PMP \*paNdan 'pandanus' (among others)

While these innovations offer solid evidence that all Austronesian languages spoken outside the homeland descend from a single origin, they do not provide direct evidence for MP's external affiliation – namely, its relationship with the languages spoken in the homeland. Does MP share a closer relationship with a particular Formosan subgroup? From a socio-historical point of view, a MP-speaking population would in principle belong to one of the Taiwanese indigenous groups, giving rise to the subgrouping in Figure 1.

---

1. See Bellwood (1984–5, 2005a, 2005b, 2017) and Blust (1984–1985, 1999) for specific arguments for Taiwan as the center of the Austronesian diaspora.

2. PAn and PMP in (1) stand for “Proto-Austronesian” and “Proto-Malayo-Polynesian”, respectively. Sources of each innovation: (1a)–(1c): Mills (1975); Blust (1999, 2001, 2009/2013); Ross (2005); (1b)–(1c): Blust (1999, 2001, 2009/2013); Ross (2005); (1d): Blust (2001); (1e): (Blust 1977, 2009/2013); Ross (2005); (1f): Ross (2002); Liao (2011); Blust (2009/2013); (1g): Blust (2001, 2009/2013); (1h): (Blust 2001, 2009/2013).



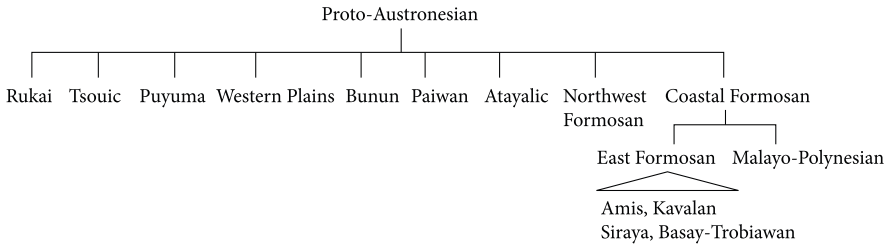
**Figure 1.** Malayo-Polynesian as a secondary subgroup

Such a connection may be adduced by means of exclusively shared innovations between MP and some Formosan subgroup. If no linguistic evidence indicates such a connection, we would have to assume either (a) the migration ex-Taiwan took place before any distinctive innovations had developed, or (b) the ancestor of MP originated as a distinct speech community when reaching Taiwan, then moved off without leaving any modern descendants. Given the relatively long pause in Taiwan before the settlement of Luzon (500–1000 years, see Bellwood 2007, 2017; Bellwood et al. 2011; Bellwood and Dizon 2013; Hung 2005; Hung and Bellwood 2010; Gray et al. 2009), both scenarios are less likely, unless no positive evidence indicates other alternatives.<sup>3</sup>

However, linguistic evidence for the origin of MP has remained vague. From the perspective of sound change, two Formosan subgroups are most likely to be its closest relatives: (i) Bunun, which shares the mergers of PAN \*C/t and \*N/n with MP, and (ii) East Formosan (EF), a subgroup comprising four coastal languages that share the merger of PAN \*C/t with MP. Both mergers are commonly treated as drifts, as neither constitutes an exclusively shared innovation. Accordingly, MP is considered a higher-order subgroup that bears no particularly close relation with any Formosan language (e.g., Blust 1999; Ross 2009, 2012).

In this paper, we present evidence for an alternative view: MP is not an Austronesian primary branch, but a subgroup under a larger branch that also subsumes East Formosan. Support for this claim comes from an understudied functional variation of the PAN morpheme \*ma- among higher-order Austronesian subgroups. While its stative intransitive function is attested across western Austronesian, an innovative use is exclusively shared between MP and EF. We argue accordingly that EF may be the closest relative of MP in the homeland and together with it constitutes a single daughter of Austronesian, as in Figure 2. Not only does this proposal better reflect the out-of-Taiwan dispersal, but it also points to a scenario more consistent with that expected from a socio-historical perspective (Figure 1) as well as recent findings in genetics and archaeology.

3. One other potential explanation for the lack of an MP-Formosan connection is that the closest relatives of MP that remained in Taiwan later became extinct as a result of language expansions. We will show in this paper that this scenario is also unlikely.



**Figure 2.** Proposed Austronesian higher-order subgrouping

The remainder of this paper is structured as follows. In the following section, we review recent proposals for Austronesian higher-order subgrouping, focusing on the treatment of Malayo-Polynesian. Section 3 surveys two distinct uses of \*ma- among higher-order Austronesian languages and their distribution. Sections 4 and 5 present arguments for one of the two uses of \*ma- as a true case of shared innovation. Section 6 presents additional evidence for the subgrouping in Figure 2 and inferences from sister fields. Section 7 concludes. Except where otherwise indicated, the data presented in this paper comes from primary fieldwork over a period of five years from 2015 to 2020.

## 2. A note on Austronesian higher-order subgrouping

We begin with an overview of MP’s interrelationships with other higher-order subgroups. Current controversies in Austronesian primary-level subgrouping boil down to three questions, (2a)–(2c).

- (2) a. Whether or not all Austronesian languages except Rukai, Tsou, and Puyuma constitute a single primary branch (Ross 2009, 2012).
- b. Whether or not all Austronesian languages except Rukai constitute a single primary branch (Starosta 1995; Aldridge 2016, to appear).
- c. Whether or not MP is an independent primary branch (Blust 1999; Ho & Yang 2000; Ross 2020).

These proposals yield distinct interpretations of the position of MP: (a) MP as an Austronesian primary branch (Proposal 1), (b) MP as a first-order subgroup of a primary branch (Proposal 2), and (c) MP as a secondary or lower order subgroup of a primary branch (Proposal 3).<sup>4</sup>

4. This summary excludes earlier proposals that offer no specific evidence for their subgroupings (e.g., Dyen 1965; Ferrell 1969; Reid 1982; Harvey 1982) and those that have been critically

Proposal 1 treats MP as a first-order offshoot of PAN along with nine other branches located in the homeland (Blust 1999), as seen in Figure 3. This treatment draws on the absence of exclusively shared phonological innovations between MP and any Formosan subgroups. Accordingly, MP is considered an independent primary branch.

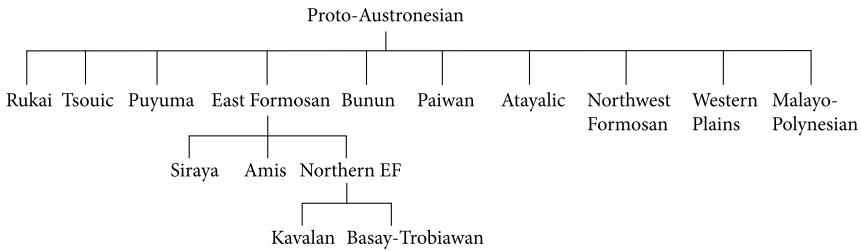


Figure 3. Austronesian higher-order subgrouping (Blust 1999)

Proposal 2 treats MP as a first-order subgroup of a primary branch comprising all Austronesian languages except Rukai, Tsou and Puyuma (Figure 4). This proposal draws on the fact that Proto-Malayo-Polynesian displays a morphological paradigm that manifests nominalizer-voice affix homophony. Based on the assumption that this feature reflects a single shared innovation, MP is placed under the primary branch defined by this alleged post-PAN innovation.

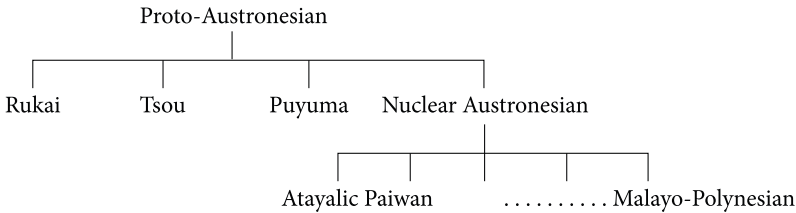


Figure 4. Ross (2009, 2012)

Proposal 3 argues instead that MP shares a closer relationship with Siraya, Kavalan and Bunun (Ho & Yang 2000) (Figure 5), building on the assumption that the merger of PAN \*C and \*t observed in all these languages reflects a single shared innovation.

---

reviewed in recent work (e.g., Sagart 2004; Aldridge 2016). See Ross (2012); Blust (2009/2013); Blust (2014); Blust and Chen (2017) and Chen (2017) for a more detailed discussion.

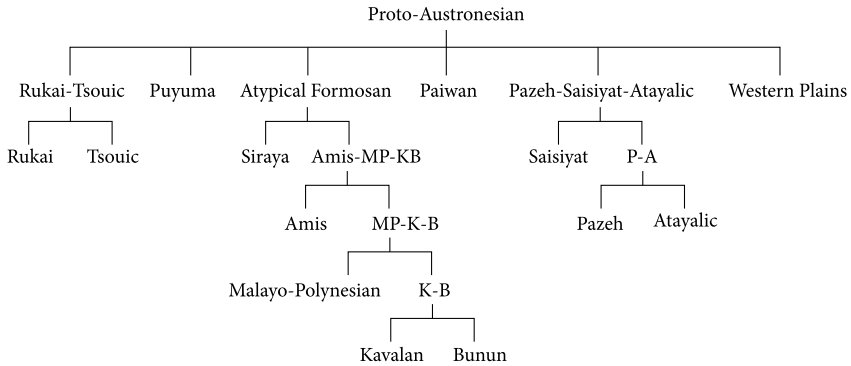


Figure 5. Ho & Yang (2000)

Since nominalizer-voice affix homophony is found across the majority of higher-order subgroups and the merger of PAN \*C/t remains controversial as a criterion for subgrouping, neither proposal offers a definite answer to the origin of Malayo-Polynesian.

Turning now to the internal subgrouping of Malayo-Polynesian, two competing proposals are worth noting. According to the traditional view (Blust 1983–1984, 1991, 1999), PMP underwent a binary split into two primary branches, Western Malayo-Polynesian (WMP) and Central-Eastern Malayo-Polynesian (CEMP), as in Figure 6.

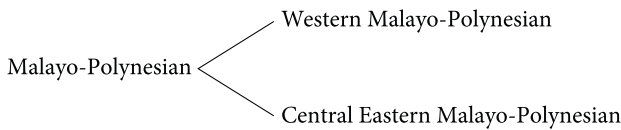


Figure 6. MP primary-level subgrouping (Blust 1983–1984, 1993b, 1999)

However, as Ross (1995: 67) points out, given the standard view of MP migration patterns, WMP as a group only reflects the speech of all of the communities that remained behind after the migration of the Central-Eastern speech community. Blust (1984–1985) also acknowledges the absence of definite evidence defining WMP as a coherent subgroup. Building on this agreement is a more recent proposal that PMP is better viewed as having undergone a rapid split into no fewer than nine primary branches (Smith 2017), as illustrated in Figure 7. We

adopt this proposal as it aligns with recent archaeological findings (Bellwood 1988, 1989, 2007; Kirch 2002; Ward et al. 1998).<sup>5</sup>

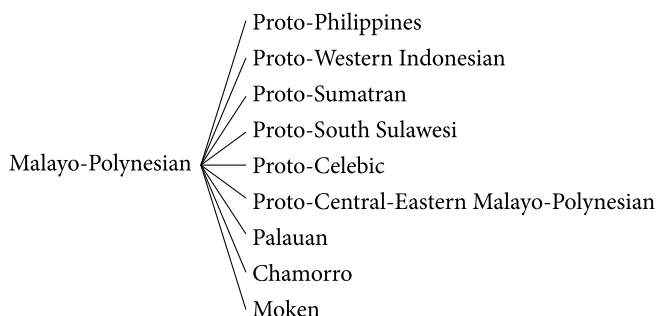


Figure 7. MP primary-level subgrouping (Smith 2017)

With this background in mind, we are now ready to turn to the functional variation of the morpheme *\*ma-* in western Austronesian and its implications for the origin of MP.

### 3. Two functions of *ma-* in western Austronesian

A comparative look at *ma-* in higher-order western Austronesian languages reveals two distinct functions: (i) *ma-* as a stative prefix used in intransitives, and (ii) *ma-* as a Patient Voice-like affix used in transitives. While the former is attested across western Austronesian, the latter is rare in Taiwan and attested only in a subset of Malayo-Polynesian languages.<sup>6</sup> We begin with a look at the distribution of each function, and then discuss what this implies for subgrouping.

5. As Smith (2017) notes, archaeological evidence suggests that the settlement of Neolithic agriculturalists in Borneo, Java, Sumatra, Sulawesi, Timor, Halmahera, the Marianas and Palau (between 4,000 and 3,500 BP) occurred not much later than the initial settlement of the Philippines (4,500 BP). This contradicts the subgrouping proposal in (8) and suggests a rapid expansion after the settlement of Luzon.

6. We do not discuss other attested functions of *ma-* in individual languages here, as this is not directly relevant to our subgrouping argument. See Ross (2015) and Himmelmann (2004) for details.

### 3.1 *ma-* as a stative intransitive prefix

The stative intransitive use of *ma-* is well-documented in the literature (see Himmelmann 2004; Huang and Sung 2008; Blust 2009/2013; Ross 2015; Blust and Trussel 2020 and the literature cited in these works). Across western Austronesian, this affix commonly combines with adjectival verbs and forms stative intransitive clauses. In such constructions, the sole argument – usually a theme or an experiencer – bears subject-marking, analogous to other Actor Voice-marked intransitives. This case pattern is illustrated by the Tagalog examples (3a)–(3c). To remain analysis-neutral, we gloss subject-marking as pivot throughout this paper.

(3) Tagalog

- a. *Ma-rikít ang babae.*  
 STAT.AV-be.beautiful PIVOT woman  
 ‘The woman is beautiful.’ (ma-clauses)
- b. *K<um>anta ang babae.*  
 <AV>sing PIVOT woman  
 ‘The woman sang.’ (Actor Voice clauses)
- c. *D<um>ating ang babae.*  
 <AV>arrive PIVOT woman  
 ‘The woman arrived.’ (Actor Voice clauses)

This function of *ma-* is uncontroversially reconstructable to Proto-Austronesian, as it is attested across all primary branches regardless of the exact subgrouping adopted (Ho 1998; Blust 1999; Starosta 1995; Sagart 2004; Ross 2009). The data below illustrates its use in 12 selected languages representing different higher-order subgroups. Parentheses following the language name indicate the language’s subgrouping affiliation under Blust’s (1999) (Figure 3) and Smith’s (2017) proposals (Figure 7) for Malayo-Polynesian subgrouping.<sup>7</sup>

- (4) a. Puyuma (Puyuma)  
*Ma-liyay na bangsaran.*  
 AV.STAT-be.drunk DEF.PIVOT young.man  
 ‘The young man is drunk.’
- b. Paiwan (Paiwan)  
*Na=ma-Leva-Leva=mun?* (Chang 2006: 269)  
 PRF=AV.STAT-RED-be.joyful=2PL.PIVOT  
 ‘Have you been happy?’

7. We indicate Ivatan’s subgrouping affiliation as “Batanic”, as the relationship of Batanic languages with other Philippine languages remains controversial (see Ross 2005, 2020; Blust 2019, 2020; and Liao 2020 for discussion).



- c. Bunun (Bunun)  
*Ma-diqla bahi.* (de Busser 2009: 328)  
 AV.STAT-be.bad dream.PIVOT  
 ‘The dream is bad.’
- d. Saisiyat (Northwestern Formosan)  
*Hae:wan ma-skes.* (Zeitoun et al. 2015: 514)  
 night AV.STAT-be.cold  
 ‘At night it is cold.’
- e. Ivatan (Batanic, MP)  
*Ma-pteng qo tao qandelak.* (Reid 1966: 127)  
 AV.STAT-be.hungry PIVOT man tomorrow  
 ‘The man will be hungry tomorrow.’
- f. Seediq (Atayalic)  
*M-sibus beyuq ka walu ga.* (ODFL)  
 AV.STAT-sweet very PIVOT honey that  
 ‘That honey is very sweet.’
- g. Rukai (Rukai)  
*Ma-adraw kay awlru-su.* (ODFL)  
 AV.STAT-be.big PIVOT head-2SG.POSS  
 ‘Your head is big.’
- h. Tagalog (Philippines, MP)  
*Ma-ta-talino ang=mga=bata=ng Intsik.* (Kroeger 1991: 24)  
 AV.STAT-PL-be.smart PIVOT=PL=child=LK Chinese  
 ‘The Chinese children are bright.’
- i. Chamorro (Chamorro, MP)  
*Ma-guf i geftao na biha.* (Chung 2020: 159)  
 AV.STAT-be.happy the generous LK old.woman  
 ‘The generous old woman was happy.’
- j. Palauan (Palauan, MP)  
*Ng kmal mle me-rau.* (Chedaol Biblia, Matthew 19:22)  
 3SG.S very AUX.PST AV.STAT-be.rich  
 ‘He was very rich.’
- k. Tamambo (CEMP, MP)  
*Glas mo ma-bila.* (Jauncey 1997: 135)  
 glass 3SG STAT.INTR-shatter  
 ‘The glass is shattered.’
- l. Tukang Besi (Western Indonesian, MP)  
*No-mo-nini.* (Donohue 1999: 157)  
 3R-STAT.INTR-cold  
 ‘They are getting cold.’

Given this distribution,  $ma_{\text{INTR}}$ - is uncontroversially a retention from Proto-Austronesian. See Blust (2009/2013), Ross (2015) and Blust and Trussel (2020) for the same assumption.

### 3.2 *ma-* as a Patient Voice-like affix used in transitives

A second function of *ma-* is found in a subset of western Austronesian languages, where the affix appears in transitive clauses featuring a genitive initiator and a theme in subject-marking.<sup>8</sup> Consider the data below from Amis (Formosan) and Tagalog (Malayo-Polynesian). Examples (5a) and (6a) illustrate the affix's intransitive use; (5b) and (6b) demonstrate the transitive use with a genitive-marked initiator.

- (5) Amis (East Formosan)
- a. *Ma-curah ku lumaq.*  
 MA-burn PIVOT house  
 'The house burned.' (ma- as a stative intransitive affix)
- b. *Ma-curah ni Kulas ku lumaq.*  
 MA-burn GEN Kulas PIVOT house  
 'Kulas burned the house.' (ma- as a transitive affix)
- (6) Tagalog (Malayo-Polynesian)
- a. *Na-sunog ang bahay.*  
 MA.REAL-burn PIVOT house  
 'The house burned.' (ma- as a stative intransitive affix)
- b. *Na-sunog ni Ivan ang bahay.*  
 MA.REAL-burn PN.GEN Ivan PIVOT house  
 'Ivan accidentally burned the house.' (ma- as a transitive affix)

In both languages, the presence of the genitive initiator alters the sentence from stative to eventive/dynamic. The GEN-pivot case frame is reminiscent of the canonical Patient Voice (PV) construction, which displays the same case pattern. Consider the examples in (7)–(8).<sup>9</sup>

8. We use the term “initiator” here to refer to noun phrases with either the agent or causer role, as the genitive-marked argument in the *ma*-construction is not restricted to (animate) agents.

9. Voice markers such as *-en* and *ma-* in Amis have been reported to bear default TAM readings (future versus perfective). See Tsukida (1993, 2008) for details. In Tagalog and many other languages of the Philippines and northern Borneo, *ma-* (and its equivalents) inflects for mood and surfaces as *na-* in realis clauses. See Himmelfmann (2004) for details.

## (7) Amis

- a. *Ma-curah ni Kulas ku lumaq.*  
 MA-burn GEN Kulas PIVOT house  
 'Kulas burned the house.' (ma-construction)
- b. *Curah-en ni Kulas ku lumaq.*  
 burn-PV GEN Kulas PIVOT house  
 'Kulas will burn the house.' (Patient Voice clause)

## (8) Tagalog

- a. *Na-sunog ni Ivan ang babay.*  
 MA.REAL-burn PN.GEN Ivan PIVOT house  
 'Ivan (accidentally) burned the house.' (ma-construction)
- b. *S<in>unog ni Ivan ang babay.*  
 burn <PV.PRF> PN.GEN Ivan PIVOT house  
 'Ivan burned the house.' (Patient Voice clause)

Sharing the same case frame, these two constructions differ only in subtle semantic interpretations. In Amis, *ma*-marked clauses denote less volitional events with an emphasis placed on the affectedness of the undergoer (Tsukida 1993; Wu 2006; see also Huang and Sung 2008 for a similar description for Kavalan). In contrast, the *en*-marked Patient Voice construction usually features a volitional initiator (Wu 2006: 39–41, 269). Similarly, *ma*-marked constructions in Tagalog often bear an accidental or abilitative reading (Himmelmann 2004, 2006), as opposed to its *PV*-marked construction, which denotes volitionality.<sup>10</sup> Given their shared case frame, much literature has described *ma*-marked transitive clauses as a type of Patient Voice construction (e.g., Amis: Tsukida 2008; Wu 2006; Kavalan/Amis: Huang & Sung: 2008; Ivatan: Reid 1966; Proto-Paitanic: Lobel 2013; Itbayaten: Yamada 2014; Yami: Rau & Dong 2006; Cebuano: Tanangkingsing 2009; inter alia).

Crucially, genitive-marked initiators are impossible with *ma*- constructions in most Formosan languages, where initiators can only be incorporated as an optional oblique phrase. Consider the data below from Paiwan and Puyuma,

10. In some Philippine languages, *ma<sub>TR</sub>*-marked clauses can bear an abilitative reading and the actual interpretation (abilitative vs. accidental) is determined by context. However, the accidental/abilitative reading is not attested in available descriptions in Chamorro and Palauan, two isolated MP primary branches. This suggests that these two readings could be secondary innovations. Here, we focus on only the diachronic implications of the morpheme's reanalysis from an intransitive intransitive morpheme to a transitive affix.

which represent two different primary branches.<sup>11</sup> Each represents a distinct Austronesian primary branch.<sup>12</sup>

- (9) a. Paiwan  
*Ma-takeDus a za kupu ni 'aLuay (tay Kalalu)*  
 AV.STAT-touch PIVOT that cup PN.POSS 'aLuay (PN.OBL Kalalu)  
 'aLuay's cup was unintentionally touched (by Kalalu).' (Chang 2006: 214)
- b. Puyuma  
*Ma-binga=ku dra tu-nirengayan kan Siber.*  
 AV.STAT-annoy=1SG.PIVOT ID.OBL 3.POSS-word LK Siber  
 'I was annoyed by Siber's words.'
- c. Puyuma  
 \**Tu<sub>i</sub>=ma-binga=ku (kan Siber/kana tu-nirengayan kan*  
 3.GEN<sub>i</sub>=MA-annoy=1SG.PIVOT (SG.GEN Siber/CN.GEN 3.POSS-word LK  
*Siber)i.*  
 Siber)i  
 (intended: 'I was annoyed {by him/her/Siber/Siber's words}.')

Within Formosan, the transitive use of *ma-* is attested only in four coastal languages: Amis (10), Kavalan (11), Basay-Trobiawan (12), and Siraya (13).<sup>13</sup> These languages constitute all members of Blust's (1999) East Formosan branch defined by the merger of PAN \*j and \*n.

- (10) Amis (East Formosan)  
*Ma-curah ni Kulas ku lumaq.*  
 MA-burn GEN Kulas PIVOT house  
 'Kulas burned the house.'
- (11) Kavalan (East Formosan)  
*Ma-ziut=na ya taquq nay ta paRinan.*  
 MA-hang= [3SG.GEN] PIVOT chicken that LOC tree  
 'He hung the chicken on the tree.' (Huang & Sung 2008: 161)

11. According to primary fieldwork on Nanwang Puyuma and available descriptions (Teng 2008), neither structure is possible for *ma*-marked clauses, as in (9c).

12. Note that the [pivot – oblique] case frame is available also in languages with *ma*<sub>TR</sub>-. Consider: Tagalog *Na-inis ng/sa (GEN/OBL) bata angale* 'The woman was annoyed by the child.' Amis: *ma-èsam kura tamdaw tura (OBL) lalangaw* 'The person was annoyed with the fly.'

13. All Basay data cited in this paper comes from the Trobiawan dialect. Following previous work (Blust 1999; Li 2004; Sagart 2004), we refer to this language as "Basay-Trobiawan".

- (12) Basay-Trobiawan (East Formosan)  
*Ma-unu=isu ma-tavan=na tama-isu?* (Li 2014: 26)  
 MA-what=2S.GEN MA-head.hunt=ASP father-2S.POSS  
 ‘Why have you beheaded your own father?’
- (13) Siraya (East Formosan)  
*Ma-i-riney=eta hia tu-turo ki rata.* (Adelaar 2011: 88)  
 MA-LOC-make=1PL.GEN here RED-three PIVOT tabernacle  
 ‘Let’s make here three tabernacles.’

This observation allows for two generalizations: (i) given the limited distribution of *ma<sub>TR</sub>*- in Formosan, this function is likely to be innovative, and (ii) given its presence in all members of EF it is most economical to assume *ma<sub>TR</sub>*- to be a retention from Proto-East-Formosan. This strengthens East Formosan as an independent subgroup (Blust 1999; Li 2010) and undermines previous disagreements (Sagart 2004, 2014).

A similar function of *ma*- is attested across multiple Malayo-Polynesian primary branches: Western Indonesian, Palauan, Chamorro and Philippines. Whether or not Philippine languages constitute a single coherent primary branch remains controversial (Reid 1982; Blust 2017, 2020; Smith 2017; Liao 2020; Ross 2020; Zorc 2020). The fact that *ma<sub>TR</sub>*- is attested in various Philippine subgroups thus indicates that this function may in fact be distributed across more than four MP primary branches. Examples (14a)–(14f) present a list of data that illustrate this distribution.<sup>14</sup> See also Table 1 for a sample of our survey results.

- (14) a. Chamorro (Chamorro, MP)  
*Ma-yamak i batalan ni napu.* (Chung 2020: 213)  
 MA-break the plank GEN wave  
 ‘The plank was broken by the waves.’
- b. Palauan (Palauan, MP)  
*A ngikel a me-ka er a ngalek.* (Gibson 1993: 143)  
 DL fish DL MA-eaten GEN DL child  
 ‘The child is eating the fish.’
- c. Arta (Philippines, MP)  
*Saya [ma-alap didi ama=mi=ti ] a laman.*  
 that [MA-get PL.GEN father=1PL.GEN=SPC ] LK boar  
 ‘That is the wild pig which our fathers can get.’ (Kimoto 2017: 215)

14. While this function is common in western Austronesian, we have not found a similar use in CEMP languages, where reflexes of *ma*- have lost the ability to license an optional genitive initiator. See Evans and Ross (2001) and work cited there.

- d. Papar (WI, MP)  
*Pai' too' dinó, niyá ma-akan kadaingan.*  
 NEG put.OV-IMP there later MA.NPST-eat children.(GEN)  
 'Don't put it there, the children might accidentally eat it.' (Lobel, pers. comm.)
- e. Murut Nabaay (WI, MP)  
*Kai' bulii' tiyo' ginó, indák am ma-akan nu dalaing.*  
 NEG put.LV-IMP that there or.else MA.NPST-eat GEN children  
 'Don't put it there, it might (accidentally) get eaten by the child.' (Lobel, pers. comm.)
- f. Bisaya (WI, MP)  
*Racun ino na-akan nu anak titinó.* (Lobel, pers. comm.)  
 poison that MA.NPAST-eat GEN child that  
 'That poison, it got accidentally eaten by that child.'
- g. Tatana (WI, MP)  
*Ino na-akán nu bogók dodinái'.* (Lobel, pers. comm.)  
 that MA.PST-eat GEN child earlier  
 'That got (accidentally) eaten by the child earlier.'
- h. Sungai Karamuak (WI, MP)  
*Racun diri na-akan di tanák.* (Lobel, pers. comm.)  
 poison that MA.PST-eat GEN child  
 'That poison, it accidentally got eaten by the child.'
- i. Sinabû (WI, MP)  
*Racun sití, na-akan ni anak-ku.* (Lobel, pers. comm.)  
 poison that MA.PST-eat GEN child-1SG.GEN  
 'That poison, it got (accidentally) eaten by my son/daughter.'

**Table 1.** Two functions of *ma-* in selected western Austronesian languages

	Subgrouping affiliation	Intransitive <i>ma-</i>	Transitive <i>ma-</i>	Source(s)
Puyuma	Puyuma	YES	NO	Teng (2008); Cauquelin (2015); primary data
Paiwan	Paiwan	YES	NO	Chang (2000, 2006)
Saisiyat	Northwestern Formosan	YES	NO	Zeitoun et al. (2015); Yeh (2018)
Atayalic	Atayalic	YES	NO	Huang (2000); Huang & Wu (2018)
Pazeh	Western Plains	YES	NO	Lin (2000)

Table 1. (continued)

	Subgrouping affiliation	Intransitive <i>ma-</i>	Transitive <i>ma-</i>	Source(s)
Rukai	Rukai	YES	NO	Zeitoun (2000, 2007, 2018)
Bunun	Bunun	YES	NO	De Busser (2009); Li (2018)
Tsou	Tsouic	YES	NO	Chang & Pan (2018); Wang (2014)
Saaroa	Tsouic	YES	NO	Pan (2012)
Kanakanavu	Tsouic	YES	NO	Sung (2018); Wild (2018)
Amis	East Formosan	YES	YES	Wu (2006); Brill (2017); primary data
Kavalan	East Formosan	YES	YES	Li & Tsuchida (2006)
Siraya	East Formosan	YES	YES	Adelaar (2011)
Basay-Trobiawan	East Formosan	YES	YES	Liu (2007); Li (2014)
Yami	Batanic (/Philippines), Malayo-Polynesian (MP)	YES	YES	Rau & Dong (2006)
Itbayaten	Batanic(/Philippines), MP	YES	YES	Yamada (2014)
Ivatan	Batanic(/Philippines), MP	YES	YES	Reid (1966)
Ilocano	Philippines, MP	YES	YES	Rubino (1997)
Arta	Philippines, MP	YES	YES	Kimoto (2017)
Tagalog	Philippines, MP	YES	YES	Himmelman (2004, 2006); primary data
Cebuano	Philippines, MP	YES	YES	Tanangkingsing (2009)
Proto-Paitanic	Greater Dusunic, Western Indonesian, MP	YES	YES	Lobel (2013), pers. comm.
Sinabu	Paitanic, Greater Dusunic, Western Indonesian, MP	YES	YES	Lobel pers. comm.

Table 1. (continued)

	Subgrouping affiliation	Intransitive <i>ma-</i>	Transitive <i>ma-</i>	Source(s)
Bisaya Sabah	Bisaya-Lotud, Greater Dusunic, Western Indonesian, MP	YES	YES	Lobel pers. comm.
Sungai Karamuak	Dusunic, Greater Dusunic, Western Indonesian, MP	YES	YES	Lobel pers. comm.
Papar	Papar, Greater Murutic, Western Indonesian, MP	YES	YES	Lobel pers. comm.
Murut	Murutic, Greater Murutic, Western Indonesian, MP	YES	YES	Lobel pers. comm.
Tatana	Tatana, Greater Murutic, Western Indonesian, MP	YES	YES	Lobel pers. comm.
Tidung Sumbol- Dungusan	Tidung, Greater Murutic, Western Indonesian, MP	YES	YES	Lobel pers. comm.
Old Balinese	Western Indonesian, MP	YES	YES	Beratha (1992)
Chamorro	Chamorro, MP	YES	YES	Gibson (1980); Chung (2020)
Palauan	Palauan, MP	YES	YES	Gibson (1993)

Note in particular that  $ma_{\text{TR}}$ - is attested across the Batanic languages, which are spoken on a group of islands scattered between Taiwan and Luzon (Figure 8). The fact that  $ma_{\text{TR}}$ - is attested in all four languages (15a)–(15d) is important for interpreting its chronology, as Proto-Batanic could be argued to be one of the early descendants of Proto-Malayo-Polynesian given its geographic location and conservatism. However, this question remains unsettled. See recent debates in Ross (2005, 2020); Blust (2019, 2020); Liao (2020); Reid (1982, 2020) and Zorc (2020).<sup>15</sup>

15. According to primary fieldwork, *ma-* can, but does not always, denote non-volitional or abilitative reading in Batanic languages. The actual interpretation of a sentence is still context-based. This is why examples (15a)–(15d) do not bear a non-volitional reading.



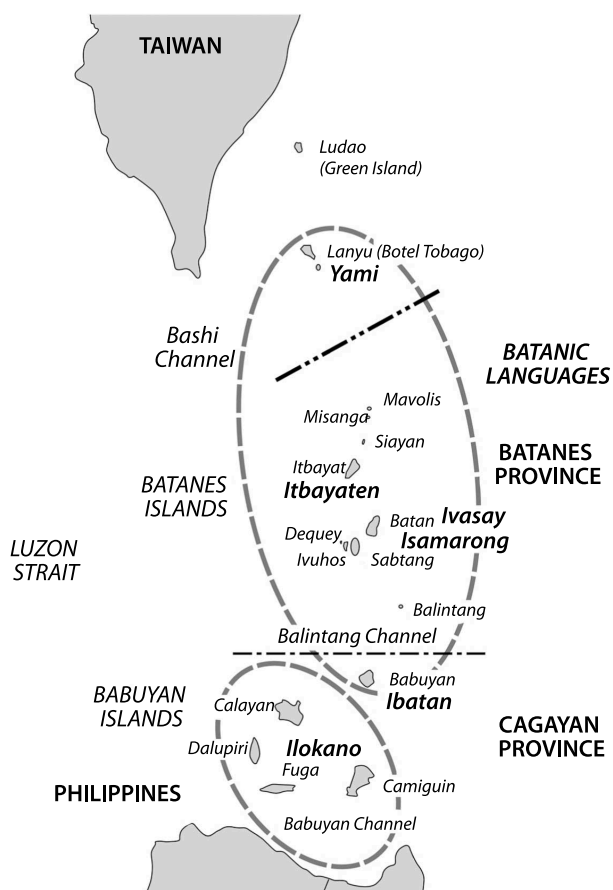


Figure 8. Distribution of the Batanic languages

- (15) a. Ivatan (Batanic)  
*Na-boyaw ni Adod o bago saya.* (Maree 2007: 223)  
 MA-chase GEN Adod PIVOT boar those  
 ‘Those pigs were chased away by Adod.’
- b. Ivatan (Batanic)  
*Ma-voyaw=mo qo manok.* (Reid 1966: 125)  
 MA-chase=2SG.GEN PIVOT chicken  
 ‘The chicken is being chased by you.’
- c. Itbayaten (Batanic)  
*Na-tta=ko si Orsing dawi.* (Yamada 2014: 72)  
 MA-see=1SG.GEN PIVOT Orsing there  
 ‘I saw Orsing there.’

d. Yami (Batanic)

*Ma-kala=ta*                      *o*                      *mogis nio?*                      (Rau & Dong 2006: 115)

MA-find=1PL.INCL.GEN PIVOT rice    2PL.GEN

‘Can we find your rice?’

#### 4. Proposal: $ma_{TR}$ - as a single, shared innovation of EF and MP

Given the observation above, an important question arises: is  $ma_{TR}$ - reconstructable to PAn, as is the stative intransitive function? If not, does it reflect a single innovation? As seen earlier,  $ma_{TR}$ - is attested in East Formosan and four MP primary branches. Possible interpretations of this distribution are laid out in (16). Here, we use the term “drift” in its accepted sense, i.e., parallel and independent innovations (Sapir 1921; Jespersen 1922; Andersen 1990; McMahon 1994; Croft 2000, 2006; Baxter et al. 2006; Steels & Szathmáry 2018; inter alia).<sup>16</sup>

- (16) a. **Scenario I (retention):** Both functions were retentions from PAn; the stative function was stable, while the transitive function has been lost in most primary branches.
- b. **Scenario II (single innovation):** The transitive use of  $ma$ - reflects a single innovation that took place prior to the split of East Formosan and Malayo-Polynesian.
- c. **Scenario III (two independent innovations, i.e., drift):** The transitive use of  $ma$ -reflects two independent innovations in Proto-East-Formosan and Proto-Malayo-Polynesian respectively.
- d. **Scenario IV (multiple independent innovations, i.e., drift):** The transitive use of  $ma$ - reflects multiple innovations in various East Formosan and Malayo-Polynesian subgroups.
- e. **Scenario V (borrowing):** The distribution of  $ma_{TR}$ - is the result of contact between East Formosan and Malayo-Polynesian.

In what follows, we present specific arguments for Scenario II.

##### 4.1 Against Scenario I ( $ma_{TR}$ - as a PAn retention)

Interpreting  $ma_{TR}$ - as a PAn retention is disfavored given its absence in any other higher-order Austronesian subgroup. Interpreting this function as a retention

---

16. The reanalysis proposed here is thus not a unique case. See Lavidas (2007) and Kulikov (2003) for two similar cases of change in Greek and Vedic where an originally intransitive/anti-causative construction was reanalyzed as a transitive causative.

thus forces an undesirable assumption:  $ma_{TR}$ - has been lost independently in the majority of Austronesian primary branches, while the stative function of the same morpheme remains stable across these branches. This scenario also entails a further undesirable assumption:  $ma_{TR}$ - is unstable and prone to loss in eight of the 10 primary branches, but remains highly resistant to change in all East Formosan and Malayo-Polynesian subgroups. Finally, the fact that  $m$ -initial morphemes canonically denote an Actor Voice case frame in western Austronesian (Blust 2009/2013; Ross 2015) suggests that  $ma_{TR}$ -, which denotes a Patient Voice case frame, is more likely to be innovative.

Accordingly, Scenario I is disfavored.

#### 4.2 Against Scenarios III and IV ( $ma_{TR}$ - as the outcome of drift)

Two other potential alternatives are to analyze  $ma_{TR}$ - either as (a) the outcome of two independent changes occurring in Proto-East-Formosan (PEF) and Proto-Malayo-Polynesian (PMP) or as (b) the result of multiple innovations occurring at lower-level subgroups of EF and MP.

Neither proposal is ideal. The fact that  $ma_{TR}$ - is attested in all members of East Formosan (§3) indicates that this innovation is best reconstructed to Proto-East Formosan. Similarly, the presence of  $ma_{TR}$ - in four of the nine Malayo-Polynesian primary branches – in particular in two isolated branches, Chamorro and Palauan, as well as all Batanic languages, strongly suggests that this function can be traced back to PMP. This argues against interpreting  $ma_{TR}$ - as the outcome of multiple independent innovations in EF and MP subgroups (Scenario IV).

Analyzing  $ma_{TR}$ - as two independent developments in PEF and PMP (Scenario III) is disfavored for similar reasons, as it forces the undesirable assumption that PAN \* $ma$ - underwent two highly similar innovations in two immediate descendants (and before further split of both branches), but not within any other eight branches. Moreover, given the presence of PAN \* $C/t$  merger in both EF and MP, we must assume PEF and PMP underwent two identical changes: innovation of  $ma_{INTR}$ - and the merger of PAN \* $C$  and \* $t$  into  $t$ .

Not only do both scenarios contradict the Principle of Economy, but they also leave a few independent pieces of evidence for an EF-MP connection unexplained. See §6 for details.

#### 4.3 Against Scenario V ( $ma_{TR}$ - as a case of structural borrowing)

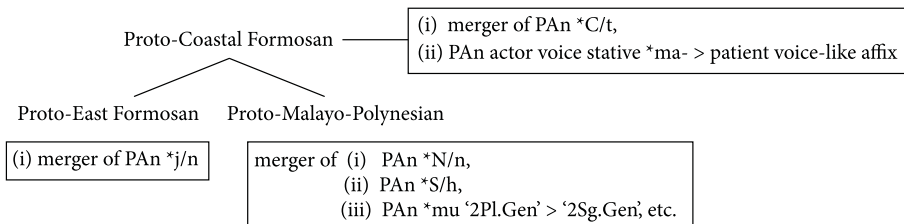
Could the distribution of  $ma_{TR}$ - reflect borrowing between EF and MP as a product of contact? We argue that this scenario is highly unlikely. Structural borrowing

has been reported to be far less common than lexical borrowing, and it typically occurs under conditions of intensive contact and bilingualism (see, e.g., Haugen 1950; Weinreich 1953; Thomason and Kaufman 1988). If  $ma_{TR}$ - is indeed a product of contact, we would expect to see other contact-induced changes in the lexicons of EF and MP languages. However, none of the four East Formosan languages show significant borrowing from MP languages, and vice versa.

The fact that  $ma_{TR}$ - adheres to the inflectional paradigms of EF and MP is further evidence against a borrowing hypothesis. In MP languages,  $ma$ - inflects for mood and surfaces as  $na$ - in the realis, as seen in (14f)–(14i) and (15). This follows from the general morphological rule observed with other  $m$ -initial morphemes, such as reflexes of PMP \* $maR$ - ‘actor voice durative’ and PMP \* $maN$ - ‘actor voice distributive’. In Formosan, this MP-specific inflectional rule is not attested.

## 5. East Formosan as the closest relative of Malayo-Polynesian

We argue accordingly that the transitive use of  $ma$ - is best viewed as an innovation which occurred prior to the split of EF and MP, and has been inherited by both branches. This proposal points to the subgrouping scenario in Figure 9, with the proposed ancestor of EF and MP labeled *Coastal Formosan*. Given the absence of PAN \* $j$  and \* $n$  in MP, we assume that MP moved out of Taiwan before Proto-East-Formosan underwent this change.<sup>17</sup>



**Figure 9.** Working hypothesis: EF and MP as sisters under a single primary branch

The current proposal yields two broader implications. First, the East Formosan languages are the closest relatives of Malayo-Polynesian in Taiwan. Second, all Austronesian primary branches are represented on Taiwan, with

17. This proposal partially coincides with two earlier insights: (i) Amis falls under an Austronesian primary branch that also subsumes extra-Formosan languages (Harvey 1982), and (ii) PMP was a language spoken by multiple small groups of migrant seafarers from southeastern Taiwan (Reid 2016: 132).

Malayo-Polynesian being a subbranch of one of the nine Formosan branches.<sup>18</sup> This proposal is consistent with the fact that East Formosan constitutes the only Austronesian language group in Taiwan distributed around the coastline of the island, as seen in Figure 10.

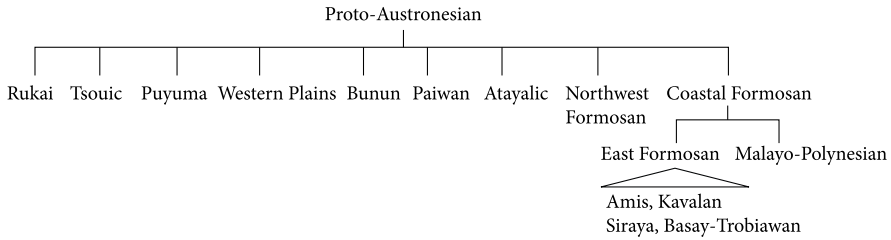


**Figure 10.** Distribution of East Formosan languages

The fact that the EF languages are spread along the east coast of Taiwan from the northern tip to the southwestern coast suggests that EF speakers may have expanded along the coastline of the island by sea.<sup>19</sup> We term this proto-language *Coastal Formosan*. This distribution, along with the known seafaring traditions of the East Formosan groups (Mabuchi 1960; Ferrell 1969; Li 2004, 2012), offers a reasonable origin story for Malayo-Polynesian as a speech community that was enabled to split off and spread to new environments by seafaring technology already possessed by the parent community. This proposal implies a nine-branch tree for Austronesian higher-order subgrouping (Figure 11).

18. The current evidence also argues against placing Bunun within an EF-MP language group (Ho & Yang 2000), as the transitive use of *ma-* is not attested in Bunun.

19. Several authors have noted that East Formosan speaking communities share certain cultural similarities, such as the matrilineal kinship system shared by the Amis, Kavalan and Siraya (Li 2004: 372, note 17; Shepherd 1993: 458, note 88; Ferrell 1969: 56). This constitutes tentative evidence for a common ethnic origin for speakers of these language groups.



**Figure 11.** Revised Austronesian higher-order subgrouping

## 6. Further evidence for EF-MP connection

Before concluding, we present additional support for the proposal we have outlined. Section 6.1 presents lexical evidence for an EF-MP connection; §6.2 discusses recent findings in sister fields consistent with the current proposal.

### 6.1 Potential lexical innovations defining Proto-Coastal-Formosan

If EF and MP are indeed descended from a single ancestor, we would expect more evidence for their common origin. One piece of evidence comes from the cognacy of Amis' existential negator *awa*. As seen in Table 2, neither EF nor MP languages possess a reflex of the PAn existential negator \**uka*, which is widely attested in Formosan (Austronesian Comparative Dictionary (ACD); Lin 2011). Moreover, an innovative form is attested in the EF language Amis and in five MP languages of the Philippines (Table 3), demonstrating a potential case of *replacement innovation* indicating their shared origin. This cognacy is difficult to explain if Amis and Philippines languages belong to two distinct primary branches.

**Table 2.** Reflexes of Proto-Austronesian existential negator \*uka\*

Saisiyat	Northwestern Formosan	<i>'oka</i>
Seediq	Atayalic	<i>uka</i>
Thao	Western Plains	<i>uka</i>
Bunun	Bunun	<i>uka</i>
Tsou	Tsouic	<i>uk'a</i>
Saaroa	Tsouic	<i>'uka'a</i>
Rukai	Rukai	<i>okaodho</i> (near comparison)
Atayal	Atayalic	<i>ungat</i> (near comparison)

\* *Near comparison* refers to comparisons in the Austronesian Comparative Dictionary for which “the observed similarity appears too great to attribute to chance, but because of imprecise agreement the reconstruction of a well-defined form is not yet possible.” (Blust & Trussel 2020)

**Table 3.** Replacement innovation of PAN \*uka in EF and MP languages\*

Amis	East Formosan	<i>awa</i>
Ivatan	Batanic/Philippines, Malayo-Polynesian	<i>ava</i>
Ibatan	Batanic/Philippines, Malayo-Polynesian	<i>aba</i>
Ilokano	Philippines, Malayo-Polynesian	<i>awan</i>
Isneg	Philippines, Malayo-Polynesian	<i>awan</i>
Casiguran Dumagat	Philippines, Malayo-Polynesian	<i>ewan</i>

\* Although this set of correspondences only constitutes near comparisons, reflexes of PAN \*uka are not attested in any EF or MP languages. Kavalan (Li & Tsuchida 2006), Siraya (Adelaar 2011) and Basay-Trobiawan (Li 2014) all possess an innovative existential negator etymologically unrelated to PAN \*uka. This suggests that \*uka may have been replaced by a distinct word before EF and MP split off.

Two other replacement innovations we have observed are PAN \*kuti ‘vulva, vagina’ (replaced by EF-MP \*puki) and \*qelud ‘housepost, pillar’ (replaced by EF-MP \*SadiRi). Consider Tables 4–7. To the best of our knowledge, neither PAN word has a reflex in EF and MP languages.

**Table 4.** Reflexes of PAn \*kuti ‘vulva, vagina’

Papora	Western Plains	<i>huci/kuci</i>
Thao	Western Plains	<i>kuti</i>
Bunun	Bunun	<i>kuti</i>
Paiwan	Paiwan	<i>kutji</i>
Puyuma	Puyuma	<i>kuti</i>

**Table 5.** PAn \*kuti replaced by EF-MP \*puki

Amis	East Formosan	<i>poki</i>
Ibang	Philippines, Malayo-Polynesian	<i>fuki</i>
Atta	Philippines, Malayo-Polynesian	<i>puki</i>
Kapampangan	Philippines, Malayo-Polynesian	<i>púki</i>
Tagalog	Philippines, Malayo-Polynesian	<i>púki</i>
Bikol	Philippines, Malayo-Polynesian	<i>púki</i>
Yakan	Philippines, Malayo-Polynesian	<i>puki</i>
Berawan	Western Indonesian, Malayo-Polynesian	<i>pukkeh</i>
Javanese	Western Indonesian, Malayo-Polynesian	<i>puki</i>
Acehnese	Chamic, Western Indonesian, Malayo-Polynesian	<i>pukɔə</i>
Moken	Moken, Malayo-Polynesian	<i>pukui</i>
Banggai	Celebic, Malayo-Polynesian	<i>uki</i>
Palauan	Palauan, Malayo-Polynesian	<i>wuk</i>
Liò	CEMP, Malayo-Polynesian	<i>puki</i>
Erai	CEMP, Malayo-Polynesian	<i>ui</i>
Sula	CEMP, Malayo-Polynesian	<i>poki</i>
Loniu	OC, CEMP, Malayo-Polynesian	<i>pwi-n</i>
Tarpia	OC, CEMP, Malayo-Polynesian	<i>piki</i>
Gitua	OC, CEMP, Malayo-Polynesian	<i>pugi</i>

**Table 6.** Reflexes of PAn \*qelud ‘housepost, pillar’

Saisiyat	Northwestern Formosan	<i>Kæ-’Lor</i>
Pazeh	Northwestern Formosan	<i>urut</i>
Paiwan	Paiwan	<i>qeluz</i>
Thao	Western Plains	<i>qrus</i>



Table 7. PAn \*qelud replaced by EF-MP \*SadiRi

Amis	East Formosan	<i>salili</i>
Ilokano	Philippines, Malayo-Polynesian	<i>adígi</i>
Tagalog	Philippines, Malayo-Polynesian	<i>halígi</i>
Aklanon	Philippines, Malayo-Polynesian	<i>halígi(h)</i>
Berawan	Western Indonesian, Malayo-Polynesian	<i>dəkkih</i>
Iban	Western Indonesian, Malayo-Polynesian	<i>diri</i>
Totoli	Celebic, Malayo-Polynesian	<i>olii</i>
Balantak	Celebic, Malayo-Polynesian	<i>orii'</i>
Rotinese	CEMP, Malayo-Polynesian	<i>di</i>
Manggarai	CEMP, Malayo-Polynesian	<i>siri</i>
Tetun	CEMP, Malayo-Polynesian	<i>rii</i>
Buli	SHWNG, CEMP, Malayo-Polynesian	<i>li</i>
Numfor	SHWNG, CEMP, Malayo-Polynesian	<i>rir</i>
Manam	OC, CEMP, Malayo-Polynesian	<i>ariri</i>
Lau	OC, CEMP, Malayo-Polynesian	<i>lili</i>

EF and MP also share a number of lexical items unattested in any other primary branch.<sup>20</sup> Consider the sample list in Tables 8–12. All items are attested in Amis and Kavalan (two better described EF languages) and various MP primary branches, but not in any other Formosan subgroup. Note in particular that the regular sound correspondences shown between Amis and Kavalan suggest these lexical items existed prior to the split of Proto-East-Formosan.

20. A preliminary search of the ACD has revealed more than 30 lexical items that have a wide distribution in Malayo-Polynesian but are observed only in EF languages in Taiwan. These include \*kawit 'hook', \*baba 'to ride pick-a-back', \*betik 'to snap the fingers', \*qasawa 'spouse', \*ma-baSaW 'cooled off', \*pikpik 'sound of patting or tapping', \*tabaN 'head trophy', \*tekep 'a cover; to cover with a flat surface', \*qaluR 'current, deep channel in the middle of a river', and \*pekpek 'to beat, hit'. We remain agnostic about the status of these items, as they may be PAn retentions with limited reflexes.

**Table 8.** Cognates of \*bubu ‘conical bamboo basket trap for fish’. This is a possible replacement innovation of PAN \*bubu ‘grandparent/grandchild (reciprocal term of address).’ Reflexes of this word are only attested in Formosan

Amis	East Formosan	<i>fofo</i>	A fish trap for catching fish and crabs in rivers
Kavalan	East Formosan	<i>bubu</i>	conical bamboo basket trap for fish
Ilokano	Philippines, Malayo-Polynesian	<i>bobo</i>	kind of large bow net used to catch fresh-water shrimps
Tagalog	Philippines, Malayo-Polynesian	<i>bubo</i>	fish trap
Aklanon	Philippines, Malayo-Polynesian	<i>bobo</i>	fish trap
Kelabit	Western Indonesian, Malayo-Polynesian	<i>bubuh</i>	bamboo basket trap for fish
Bintulu	Western Indonesian, Malayo-Polynesian	<i>buwew</i>	conical bamboo fish trap
Malagasy	Western Indonesian, Malayo-Polynesian	<i>vovo</i>	kind of basket used for fishing
Moken	Moken, Malayo-Polynesian	<i>bubey</i>	fish trap
Palauan	Palauan, Malayo-Polynesian	<i>bub</i>	trap (usually for fish)
Hawu	CMP, CEMP, Malayo-Polynesian	<i>wuwu</i>	fish trap
Rotinese	EMP, CEMP, Malayo-Polynesian	<i>bufu</i>	fish trap
Buli	SHWNG, CEMP, Malayo-Polynesian	<i>pup</i>	fish trap
Kowiai	CMP, CEMP, Malayo-Polynesian	<i>fuf</i>	kind of fish trap
Proto-Oceanic	CEMP, Malayo-Polynesian	*pupu	conical bamboo basket trap for fish

**Table 9.** Cognates of \*Rabiqi ‘late afternoon, evening, evening meal’

Amis	East Formosan	<i>lafī</i>	‘evening’
Kavalan	East Formosan	<i>Rabi</i>	‘evening, dinner; the evening meal’
Isneg	Philippines, Malayo-Polynesian	<i>xabi</i>	‘evening, night’
Ifugaw	Philippines, Malayo-Polynesian	<i>labi</i>	‘night’
Cebuano	Philippines, Malayo-Polynesian	<i>gabi’i</i>	‘night’
Supan	Western Indonesian, Malayo-Polynesian	<i>gabpi-n</i>	‘night’
Mongondow	Celebic, Malayo-Polynesian	<i>gobii</i>	‘night’
Bimane	CMP, CEMP, Malayo-Polynesian	<i>awi(na)</i>	‘yesterday’
Wandamen	SHWNG, CEMP, Malayo-Polynesian	<i>ravi-nena</i>	‘afternoon’
Ron	SHWNG, CEMP, Malayo-Polynesian	<i>rob</i>	‘night’
Numfor	SHWNG, CEMP, Malayo-Polynesian	<i>rob</i>	‘night’
Proto-Oceanic	CEMP, Malayo-Polynesian	*Rapi	‘afternoon, evening, yesterday’

**Table 10.** Cognates of \*buSaw ‘cold, of leftover food; leftover from a meal’

Amis	East Formosan	<i>fasaw</i>	‘cooled off’
Kavalan	East Formosan	<i>basaw</i>	‘fever went down, abated; to become cold’
Itbayaten	Batanic/Philippines, Malayo-Polynesian	<i>vahaw</i>	‘idea of being cold (of food esp.)’
Ilokano	Philippines, Malayo-Polynesian	<i>baaw</i>	‘left over, cold rice; tepidity, coolness’
Pangasinan	Philippines, Malayo-Polynesian	<i>baaw</i>	‘cooked rice’
Tagalog	Philippines, Malayo-Polynesian	<i>bahaw</i>	‘left-over food, especially boiled or steamed rice’
Bikol	Philippines, Malayo-Polynesian	<i>bahaw</i>	‘cold, referring only to food once served hot’
Aklanon	Philippines, Malayo-Polynesian	<i>bahaw</i>	‘cold rice; cool off, get cool (said of food)’
Mansaka	Philippines, Malayo-Polynesian	<i>baaw</i>	‘food prepared for a trip’

Table 10. (continued)

Bintulu	Western Indonesian, Malayo-Polynesian	<i>pa-vaw</i>	'cold, of hot food that has gotten cold'
Agutaynen	CEMP, Malayo-Polynesian	<i>baw</i>	'breakfast, morning snack'

Table 11. Cognates of \*laRiw 'run, run away, flee, escape'

Amis	East Formosan	<i>laliw</i>	'escape'
Kavalan	East Formosan	<i>m-RaRiw</i>	'run, run away'
Itbayaten	Batanic/Philippines, Malayo-Polynesian	<i>yayo-h</i>	'race'
Hanunóo	Philippines, Malayo-Polynesian	<i>lagiw</i>	'running'
Abaknon	Philippines, Malayo-Polynesian	<i>lahi</i>	'to run, run away'
Cebuano	Philippines, Malayo-Polynesian	<i>lagiw</i>	'run away'
Maranao	Philippines, Malayo-Polynesian	<i>lagoy</i>	'rush, hurry'
Samal	Philippines, Malayo-Polynesian	<i>lahi-lahi</i>	'to run'
Iban	Western Indonesian, Malayo-Polynesian	<i>lari</i>	'run away'
Malay	Western Indonesian, Malayo-Polynesian	<i>lari</i>	'escape'
Kambera	CMP, CEMP, Malayo-Polynesian	<i>lao</i>	'to run'
Hawu	CMP, CEMP, Malayo-Polynesian	<i>rai</i>	'run, run away'
Soboyo	CMP, CEMP, Malayo-Polynesian	<i>labi</i>	'to run, run away'
Ron	SHWNG, CEMP, Malayo-Polynesian	<i>farar</i>	'to run'

Table 12. Cognates of \*bangeSiS 'fragrant/fragrance'

Amis	East Formosan	<i>fangsis</i>	'sweet'
Kavalan	East Formosan	<i>bangsis</i>	'fragrant'
Ilokano	Philippines, Malayo-Polynesian	<i>bang'i</i>	'to smell of toast'
Cham	Western Indonesian, Malayo-Polynesian	<i>bangi</i>	'used of all agreeable sensations: good, tasty, redolant, etc'
Simalur	Western Indonesian, Malayo-Polynesian	<i>fangi</i>	'odor, fragrance'
Old Balinese	Western Indonesian, Malayo-Polynesian	<i>wangi</i>	'fragrance'
Ngadha	CMP, CEMP, Malayo-Polynesian	<i>vangi</i>	'to smell'

## 6.2 Support from sister fields

In what follows, we summarize recent findings in genetics and archaeology which yield consistent inferences with the current proposal.

### 6.2.1 Genetics

Recent genetic research has suggested a close connection between East Formosan and Malayo-Polynesian populations. Capelli et al. (2001) report that Y chromosomes from the Amis but not other non-East Formosan aboriginal communities of Taiwan are distributed throughout selected Austronesian communities outside Taiwan. McColl et al. (2018), Tätte et al. (2021) and Pugach et al. (2021) reach a similar conclusion that the Amis speakers cluster most closely with Philippine speakers in Y-DNA and mtDNA. Consider the following data from Chen et al. (2011: 44), which shows that the genetic distance between the Amis and Filipinos was shorter than that between the Amis and other Formosan tribes.<sup>21</sup>

**Table 13.** Genetic distance between Formosan and Philippines populations

Population compared	Genetic distance
Amis – Philippines	0.025
other Formosan aboriginals – Amis	0.068
other Formosan aboriginals – Philippines	0.073

Trejaut et al. (2005) reports two consistent findings: (i) among all Formosan populations tested, the Amis in particular are more closely related to island South-east Asian populations than to populations from mainland East Asia, and (ii) Y-chromosome haplogroup B4a1a occurs frequently among the Amis and the Yami (Batanic, Malayo-Polynesian) but not in any other Formosan communities (Paiwan, Puyuma, Rukai) examined in the study. Contra the conventional view in linguistics that Malayo-Polynesian does not have a closer relationship with any Formosan group, this suggested link between East Formosan and Malayo-Polynesian finds a home for the out-of-Taiwan population in Taiwan, offering a scenario that is potentially more compatible with the perspective of human expansion.

21. Capelli et al. (2001) did not specify the exact Filipino community examined in their study, and the Amis was the only East Formosan population examined in this study.

### 6.2.2 *Archaeology*

Recent archaeological findings further support a closer link between East Formosan and Malayo-Polynesian. In particular, evidence from material culture shows strong links between early Malayo-Polynesian sites on the Batanes islands and Luzon, and the areas of Taiwan where East Formosan languages are spoken today, and where Proto-Coastal-Formosan was presumably spoken.

Early sites on the Batanes islands, dated at the earliest to around 2200 BCE, have yielded numerous pottery sherds characterized by red-slipped or plain surfaces, rare cord-marking and circle stamped decoration, and occasional carination (joining of a rounded base to a main vessel with inward sloping sides) (Bellwood & Dizon 2005, 2008, 2013). This pottery has strong similarities with that found in coastal southeastern Taiwan. Sites of similar age or slightly earlier than 2200 BCE, such as Donghebei (Chu 1990), Fushan and Chaolaiqiao (Hung 2008), share the characteristics already mentioned, as well as very similar everted and concave rim profiles, indicating that similar vessel forms and techniques were used in their manufacture. These sites are regarded as the type sites of the Fushan archaeological culture of the middle Neolithic phase on the eastern coast of Taiwan (Kuo 2019).

The “package” of material culture which arrived in the Batanes with Neolithic settlers shows other strong links to southeast Taiwan and the Fushan culture (Hung 2005, 2008; Kuo 2019). Geochemical composition of nephrite jade artifacts from early settlement sites in the Batanes and northern Philippines shows the raw material was sourced from the Fengtian quarry in Hualien, eastern Taiwan (Hung & Iizuka 2013; Hung et al. 2007). Many of these artifacts are also clearly linked to the Fushan culture in terms of their design and workmanship.

Several authors have already suggested in the archaeological literature that these similarities in material culture indicate that it was people bearing the Fushan culture sites who moved southwards into the Batanes and Luzon (e.g., Hung 2005; Kuo 2019; Carson & Hung 2018). The East Formosan languages are also concentrated in southeast Taiwan, overlapping with the zone in which Fushan culture sites are found. Because of this, and because it is also clear from the archaeological evidence that the Neolithic colonization process moved north to south from eastern Taiwan to the Philippines (Bellwood et al. 2011), it seems reasonable to identify the Fushan culture with a possible Proto-Coastal-Formosan speech community.

## 7. Conclusion

Although the Austronesian peoples' long pause in Taiwan (Bellwood 2007, 2017; Bellwood et al. 2011; Gray et al. 2009) suggests that Malayo-Polynesian (MP) languages might be expected to belong to a primary branch located in the homeland, the lack of precise phonological evidence for their origin has led to the view that MP constitutes a primary branch. In this paper, we have presented novel evidence that MP is the sister of East Formosan, a subgroup comprising four languages (Amis, Siraya, Kavalan, Basay-Trobiawan) distributed around the coastline of Taiwan. Support for this claim comes from an understudied syntactic reanalysis shared exclusively between EF and MP: reanalysis of the PAn stative intransitive \*ma- into a transitive affix. The fact that all members of East Formosan reflect the same merger of PAn \*C/t with MP lends further support to this proposal. This revised subgrouping is not only supported by a number of lexical innovations between EF and MP, but it also aligns with two recent findings in sister fields: (i) eastern Taiwan was the starting point of Austronesian expansion (Hung 2005, 2008, 2019; Bellwood 2017; Bellwood & Dizon 2008; Carson & Hung 2018), and (ii) East Formosan speakers are genetically closer to Austronesian communities located outside Taiwan (Capelli et al. 2001; Trejaut et al. 2005; McColl et al. 2018; Pugach et al. 2021; Tätte et al. 2021).

## Funding

This project was supported by a Marsden Fast Start Grant (#MFP-VUW2012) awarded to the first author and a Chiang Ching-Kuo Foundation Research Grant (#RG021-A-16) in which the first author was a project collaborator.

## Acknowledgements

We thank Peter Bellwood, Robert Blust, Geoff Chambers, Shu-chao Chen, Beth Evans, Shuanfan Huang, Hsiao-chun Hung, Ritsuko Kikusawa, Benedict King, Miriam Meyerhoff, Laurie Reid, Malcolm Ross, Alex Smith, Benjamin van der Voorn, Sandra Auderset, Marie-France Duhamel and Mary Walworth, as well as two anonymous *Diachronica* reviewers for comments and feedback. Thank you also to Jason Lobel for data on Borneo languages. Thank you to Atrung Kagi, Lisin Kalitang, Sairin bin Kamis, Malinus bin Jaliji, Aisah binti Abdurasik, Jaimin bin Tangkisan and Pinggung alias Jacob bin Pong, for sharing their language with us.

## Abbreviations

ASP	aspect	AV	actor voice
AUX	auxiliary		

CEMP	Central-Eastern Malayo-Polynesian	PL	plural
		PMP	Proto-Malayo-Polynesian
CMP	Central Malayo-Polynesian	PN	proper name
CN	common noun	PRF	perfective
EMP	Eastern Malayo-Polynesian	PST	past tense
GEN	genitive	PV	patient voice
ID	indefinite	RED	reduplication
INTR	intransitive	REAL	realis
IRR	irrealis	SG	singular
LK	linker	SHWNG	South Halmahera-West New Guinea
LOC	locative		
OBL	oblique	STAT	stative
NPST	non-past	TR	transitive
PAn	Proto-Austronesian		

## References

- Adelaar, Alexander. 2011. *Siraya: Retrieving the phonology, grammar, and lexicon of a dormant Formosan language*. Berlin & Boston: De Gruyter Mouton.  
<https://doi.org/10.1515/9783110252965>
- Aldridge, Edith. 2016. Ergativity from subjunctive in Austronesian languages. *Language and Linguistics* 17(1). 27–62.
- Aldridge, Edith. To appear. Syntactic conditions on accusative to ergative alignment change in Austronesian languages. *Journal of Historical Linguistics*.
- Andersen, Henning. 1990. The structure of drift. In Henning Andersen & Konrad Koerner (eds.), *Historical Linguistics 1987. Papers from the 8th International Conference on Historical Linguistics*, 1–20. Amsterdam: Benjamins. <https://doi.org/10.1075/cilt.66.02and>
- Baxter, Gareth, Richard Blythe, William Croft & Alan McKane. 2006. Utterance selection model of language change. *Physical Review E* 73(4). 46–118.  
<https://doi.org/10.1103/PhysRevE.73.046118>
- Bellwood, Peter. 1984–1985. A hypothesis for Austronesian origins. *Asian Perspectives* 26. 107–117.
- Bellwood, Peter. 1988. *Archaeological research in south-eastern Sabah*. Kota Kinabalu: Sabah Museum and State Archives.
- Bellwood, Peter. 1989. Archaeological investigations at Bukit Tengkorak and Segarong, southeastern Sabah. *Bulletin of the Indo-Pacific Prehistory Association* 9. 122–162.  
<https://doi.org/10.7152/bippa.v9i0.11286>
- Bellwood, Peter. 2005a. *First farmers: The origins of agricultural societies*. Oxford: Blackwell.
- Bellwood, Peter. 2005b. Coastal south China, Taiwan, and the prehistory of the Austronesians. In Chung-Yu Chen & Jian-Guo (eds.), *The archaeology of south-east coastal islands of China conference*, 1–22. Taipei: Executive Yuan, Council for Cultural Affairs.
- Bellwood, Peter. 2007. *Prehistory of the Indo-Malaysian archipelago*, revised ed. Canberra: Australian National University e Press. <https://doi.org/10.22459/PIMA.03.2007>
- Bellwood, Peter. 2017. *First islanders: Prehistory and human migration in Island Southeast Asia*. New Jersey: John Wiley & Sons. <https://doi.org/10.1002/9781119251583>



- Bellwood, Peter & Eusebio Dizon. 2005. The Batanes archaeological project and the Out of Taiwan hypothesis for Austronesian dispersal. *Journal of Austronesian Studies* 1. 1–32.
- Bellwood, Peter & Eusebio Dizon. 2008. Austronesian cultural origins: Out of Taiwan, via the Batanes Islands, and onwards to Western Polynesia. In Alicia Sanchez-Mazas, Roger Blench, Malcolm D. Ross, Ilia Peiros & Marie Lin (eds.), *Past human migrations in East Asia: Matching archaeology, linguistics and genetics*, 23–39. UK: Routledge.
- Bellwood, Peter, Geoffrey Chambers, Malcolm Ross & Hsiao-chun Hung. 2011. Are “cultures” inherited? Multidisciplinary perspectives on the origins and migrations of Austronesian-speaking peoples prior to 1000 BC. In B. Roberts & M. Vander Linden (eds.), *Investigating archaeological cultures: Material culture, variability, and transmission*, 321–354. New York: Springer. [https://doi.org/10.1007/978-1-4419-6970-5\\_16](https://doi.org/10.1007/978-1-4419-6970-5_16)
- Bellwood, Peter & Eusebio Dizon. 2013. The Batanes Islands and the prehistory of Island Southeast Asia. In Peter Bellwood & Eusebio Dizon (eds.), *4000 years of migration and cultural exchange: The archaeology of the Batanes Islands, Northern Philippines (Terra Australis 40)*, 235–239. Canberra: Australian National University ePress. <https://doi.org/10.22459/TA40.12.2013.13>
- Beratha, Ni Luh Sutjiati. 1992. Evolution of verbal morphology in Balinese. PhD dissertation, The Australian National University.
- Blust, Robert. 1977. The Proto-Austronesian pronouns and Austronesian subgrouping: A preliminary report. *Working Papers in Linguistics* 9(2). 1–15. Honolulu: Department of Linguistics, University of Hawai‘i at Mānoa.
- Blust, Robert. 1983–1984. More on the position of the languages of eastern Indonesia. *Oceanic Linguistics* 22–23. 1–28.
- Blust, Robert. 1984–1985. The Austronesian homeland: A linguistic perspective. *Asian Perspectives* 26. 45–67.
- Blust, Robert. 1991. The Greater Central Philippines hypothesis. *Oceanic Linguistics* 30. 73–129. <https://doi.org/10.2307/3623084>
- Blust, Robert. 1993a. \*S metathesis and the Formosan/Malayo-Polynesian language boundary. In Øyvind Dahl (ed.), *Language – a doorway between human cultures: Tributes to Dr. Otto Chr. Dahl on his ninetieth birthday*, 178–183. Oslo: Novus.
- Blust, Robert. 1993b. Central and Central Eastern Malayo-Polynesian. *Oceanic Linguistics* 32. 241–293. <https://doi.org/10.2307/3623195>
- Blust, Robert. 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*, 31–94. Taipei: Institute of Linguistics, Academia Sinica.
- Blust, Robert. 2001. Malayo-Polynesian: New stones in the wall. *Oceanic Linguistics* 40. 151–55. <https://doi.org/10.2307/3623269>
- Blust, Robert. 2009/2013. *The Austronesian languages*. Canberra: Pacific Linguistics.
- Blust, Robert. 2014. Some recent proposals concerning the classification of the Austronesian languages. *Oceanic Linguistics* 53(2). 300–391. <https://doi.org/10.1353/ol.2014.0025>
- Blust, Robert. 2017. The linguistic history of Austronesian-speaking communities in Island Southeast Asia. In Peter Bellwood, *First islanders: Prehistory and human migration in island Southeast Asia*, 190–197. Oxford: Wiley Blackwell.
- Blust, Robert. 2019. The Austronesian homeland and dispersal. *Annual Review of Linguistics* 5. 417–434. <https://doi.org/10.1146/annurev-linguistics-011718-012440>

- Blust, Robert. 2020. Response to comments on “The resurrection of Proto-Philippines”. *Oceanic Linguistics* 59. 450–79. <https://doi.org/10.1353/ol.2020.0020>
- Blust, Robert & Victoria Chen. 2017. The pitfalls of negative evidence: ‘Ergative Austronesian’, ‘Nuclear Austronesian’ and their progeny. *Language and Linguistics* 18(4). 577–621.
- Blust, Robert & Stephen Trussel. 2020. Austronesian comparative dictionary. Online: <http://www.trussel2.com/> Austronesian Comparative Dictionary.
- Bril, Isabelle. 2017. Roots and stems in Amis and Nêlêmwa (Austronesian). In Eva van Lier (ed.), *Studies in Language* 41(2). 358–407. <https://doi.org/10.1075/sl.41.2.04bri>
- Capelli, Cristian, James F. Wilson, Martin Richards, Michael P. Stumpf, Fiona Gratrix, Stephen Oppenheimer, Peter Underhill, Vincenzo L. Pascali, Tsang-Ming Ko & David B. Goldstein. 2001. A predominantly indigenous paternal heritage for the Austronesian-speaking peoples of insular Southeast Asia and Oceania. *American Journal of Human Genetics* 68. 432–43. <https://doi.org/10.1086/318205>
- Carson, Mike & Hsiao-chuen Hung. 2018. Learning from paleo-landscapes: Defining the land-use systems of the ancient Malayo-Polynesian homeland. *Current Anthropology* 59(6). 790–813. <https://doi.org/10.1086/700757>
- Cauquelin, Josiane. 2015. *Nanwang Puyuma-English dictionary*. Taipei: Academia Sinica.
- Chang, Anna. 2000. *A reference grammar of Paiwan* (in Chinese). Taipei: Yuanliu.
- Chang, Anna. 2006. *A reference grammar of Paiwan*. PhD dissertation, The Australian National University.
- Chang, Yung-li & Chia-jung Pan. 2018. *A sketch grammar of Tsou* (in Chinese). New Taipei: Council of Indigenous Peoples.
- Chen, Yaofeng, Shuzhuo Chen & Muzhu Xu. 2011. Homeland or transfer station? Genetic studies’ contribution to our understanding of Austronesian diaspora (in Chinese). *Humanities and Social Sciences Newsletter Quarterly* 12(3). 41–49.
- Chen, Victoria. 2017. *A re-examination of the Philippine-type voice system and its implications for Austronesian primary-level subgrouping*. PhD dissertation, University of Hawai‘i.
- Chu, C. 1990. *Investigation on the Neolithic sites in the estuary of Mawuku River, Taidong* (in Chinese). Unpublished Masters thesis, National Taiwan University, Taipei.
- Chung, Sandra. 2020. *Chamorro grammar*. California Digital Library University of California.
- Croft, William. 2000. *Explaining language change: An evolutionary approach*. London: Longman.
- Croft, William. 2006. The relevance of an evolutionary model to historical linguistics. In O. Nedergaard Thomsen (ed.), *Competing models of linguistic change: Evolution and beyond*, 91–132. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.279.o8cro>
- Dahl, Otto Christian. 1973. Proto-Austronesian. *Scandinavian Institute of Asian Studies Monograph Series*, No. 15. London: Curzon Press.
- De Busser, Rik. 2009. *Towards a grammar of Takivatan Bunun: Selected topics*. PhD dissertation, La Trobe University.
- Donohue, Mark. 1999. *A grammar of Tukang Besi*. Amsterdam: Mouton de Gruyter. <https://doi.org/10.1515/9783110805543>
- Dyen, Isidore. 1965. *A lexicostatistical classification of the Austronesian languages*. *Indiana University Publications in Anthropology and Linguistics. Memoir of the International Journal of American Linguistics* 19. Baltimore: The Waverly Press.
- Evans, Bethwyn & Malcolm Ross. 2001. The history of Proto-Oceanic \*ma-. *Oceanic Linguistics* 40. 269–290. <https://doi.org/10.1353/ol.2001.0018>

- Ferrell, Raleigh. 1969. *Taiwan aboriginal groups: Problems in cultural and linguistic classification*. Taipei: Academia Sinica.
- Gibson, Jeanne. 1980. Clause union in Chamorro and in Universal Grammar. PhD dissertation, University of California, San Diego, CA.
- Gibson, Robert. 1993. Palauan causatives and passives: An incorporation analysis. PhD dissertation, University of Hawai'i at Mānoa.
- Gray, Russell, Alexei Drummond & Simon Greenhill. 2009. Language phylogenies reveal expansion pulses and pauses in Pacific settlement. *Science* 323(5913). 479–483. <https://doi.org/10.1126/science.1166858>
- Harvey, Mark. 1982. Subgroups in Austronesian. In Amran Halim, Lois Carrington & Stephen A. Wurm (eds.), *Papers from the Third International Conference on Austronesian Linguistics* 2. 47–99. Canberra: The Australian National University.
- Haugen, Einar. 1950. The analysis of linguistic borrowing. *Language* 26. 210–231. <https://doi.org/10.2307/410058>
- Himmelman, Nikolaus. 2004. On statives and potitives in Western Austronesian (mostly Tagalog). In Paul Law (ed.), *Proceedings of AFLA 11. ZAS Papers in Linguistics* 34. 103–19. Berlin: Zentrum für Allgemeine Sprachwissenschaft, Typologie und Universalienforschung (ZAS). <https://doi.org/10.21248/zaspil.34.2004.206>
- Himmelman, Nikolaus. 2006. How to miss a paradigm or two: Multifunctional *ma-* in Tagalog. In Felix Ameka, Alan Dench & Nicholas Evans (eds.), *Catching language: The standing challenge of grammar writing*, 487–526. Berlin: Mouton de Gruyter.
- Ho, Dah-an. 1998. Genetic relationships among the Formosan languages (in Chinese). *Chinese Studies* 16(2). 141–71.
- Ho, Dah-an & Hsiu-fang Yang. 2000. *Prologue: Austronesian and Formosan languages*. In Lilian Huang (ed.), *Formosan languages grammar series* (in Chinese), 1–25. Taipei: Yuanliu.
- Huang, Lilian. 2000. *A reference grammar of Atayal* (in Chinese). Taipei: Yuanliu.
- Huang, Lilian & Hsin-sheng Wu. 2018. *Atayal. A sketch grammar of Tsou* (in Chinese). New Taipei: Council of Indigenous Peoples.
- Huang, Shu-Ping & Li-May Sung. 2008. The undergoer focus *ma-* in Kavalan. *Oceanic Linguistics* 47. 159–84. <https://doi.org/10.1353/ol.0.0010>
- Hung, Hsiao-chuen. 2005. Neolithic interaction between Taiwan and Northern Luzon: The pottery and jade evidence from the Cagayan Valley. *Journal of Austronesian Studies* 1. 109–34.
- Hung, Hsiao-chuen. 2008. Migration and cultural interaction in southern coastal China, Taiwan and the northern Philippines, 3000 BC to AD 100: The early history of the Austronesian-speaking populations. PhD dissertation, The Australian National University.
- Hung, Hsiao-chuen. 2019. Ancient interactions between eastern Taiwan and the Philippines: From the Palaeolithic Age to Iron Age. *Field Archaeology of Taiwan* 19(2). 87–116.
- Hung, Hsiao-Chun, Yoshiyuki Iizuka, Peter Bellwood, Kim Dung Nguyen, Bérénice Bellina, Praon Silapanth, Eusebio Dizon, Rey Santiago, Ipoi Datan, & Jonathan H. Manton. 2007. Ancient jades map 3,000 years of prehistoric exchange in Southeast Asia. *Proceedings of the National Academy of Sciences* 104(50). 19745–19750. <https://doi.org/10.1073/pnas.0707304104>

- Hung, Hsiao-chuen & Peter Bellwood. 2010. Movement of raw materials and manufactured goods across the South China Sea after 500 BCE: From Taiwan to Thailand, and back. In Berenice Bellina, E.A. Bacus, T.O. Pryce & J.W. Christie (eds.), *50 years of archaeology in Southeast Asia: Essays in honour of Ian Glover*, 235–245. Bangkok: River Books.
- Hung, Hsiao-chuen & Yoshiyuki Iizuka. 2013. The Batanes Nephrite Artefacts. In Peter Bellwood & Eusebio Dizon (eds.), *4000 Years of Migration and Cultural Exchange: The Archaeology of the Batanes Islands, Northern Philippines*, 149–168. ANU Press. <https://doi.org/10.22459/TA40.12.2013.09>
- Jauncey, Dorothy. 1997. A grammar of Tamambo, the language of western Malo, Vanuatu. PhD dissertation, The Australian National University.
- Jespersen, Otto. 1922. *Language, its nature, development, and origin*. New York: H. Holt.
- Kimoto, Yukinori. 2017. A grammar of Arta: A Philippine Negrito language. PhD dissertation, Kyoto University.
- Kirch, Patrick. 2002. *On the road of the winds: An archaeological history of the Pacific islands*. Berkeley: University of California Press.
- Kroeger, Paul. 1991. Phrase structure and grammar relations in Tagalog. PhD dissertation, Stanford University.
- Kulikov, Leonid. 2003. The labile syntactic type in a diachronic perspective: The case of Vedic. *SKY Journal of Linguistics* 16. 93–112.
- Kuo, Su-chiu. 2019. *New frontiers in the neolithic archaeology of Taiwan (5600–1800 BP)*. Singapore: Springer Singapore. <https://doi.org/10.1007/978-981-32-9263-5>
- Lavidas, Nikolaos. 2007. *Transitivity alternations in diachrony: Changes in argument structure and voice morphology*. Newcastle: Cambridge Scholar Publishing.
- Li, Lilian Li-ying. 2018. A grammar of Isbukun Bunun. PhD dissertation, National Tsing Hua University.
- Li, Paul Jen-kuei. 2004. Origins of the East Formosans: Basay, Kavalan, Amis and Siraya. *Language and Linguistics* 5. 363–376.
- Li, Paul Jen-kuei. 2010. Origins and dispersal of the Formosan natives in the east coast. *Taiwan Journal of Indigenous Studies* 3(4). 1–9.
- Li, Paul Jen-kuei. 2012. The ethnic groups in Lanyang plain at an early stage, as seen in the language data (in Chinese). *Kiwulan International Symposium*, 13–26. Yilan: Archives of Yilan History.
- Li, Paul Jen-kuei. 2014. *Texts of the Trobiawan dialect of Basay*. Asian and African Lexicon Series 56. Tokyo: Research Institute for Languages and Cultures of Asia and Pacific, Tokyo University of Foreign Studies.
- Li, Paul Jen-kuei & Shigeru Tsuchida. 2006. *Kavalan dictionary*. Taipei: Institute of Linguistics, Academia Sinica.
- Liao, Hsiu-chuan. 2011. Some morphosyntactic differences between Formosan and Philippine languages. *Language and Linguistics* 12(4). 845–876.
- Liao, Hsiu-chuan. 2020. A reply to Blust (2019) “The Resurrection of Proto-Philippines.” *Oceanic Linguistics* 59. 426–449. <https://doi.org/10.1353/ol.2020.0019>
- Lin, Shu-yi. 2011. Reconstructing negative morphemes in Proto-Austronesian: Evidence from Formosan languages. MA thesis, National Taiwan Normal University.
- Lin, Ying-chin. 2000. *A reference grammar of Pazeh* (in Chinese). Taipei: Yuanliu.
- Liu, Dorinda Tsai-Hsiu. 2007. Basay nominal constructions. *Working Papers in Linguistics* 38(5). 1–29. Department of Linguistics, University of Hawai‘i.

- Lobel, Jason. 2013. Philippine and north Bornean languages: Issues in description, subgrouping, and reconstruction. PhD dissertation, University of Hawai'i.
- Mabuchi, Tōichi. 1960. The aboriginal peoples of Formosa. In George Peter Murdock (ed.), *Social structure in Southeast Asia*, 127–40. Chicago: Quadrangle.
- McCull, Hugh, Fernando Racimo, Lasse Vinner, Fabrice Demeter, Takashi Gakuhari, J. Victor Moreno-Mayar, George van Driem, ... Eske Willersley. 2018. The prehistoric peopling of Southeast Asia. *Science* 361. 88–92. <https://doi.org/10.1126/science.aat3628>
- McMahon, April. 1994. *Understanding language change*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139166591>
- Mills, Roger. 1975. Proto South Sulawesi and Proto Austronesian phonology. PhD dissertation, University of Michigan at Ann Arbor.
- Online Dictionary of Formosan languages (ODFL). <http://e-dictionary.apc.gov.tw>
- Pan, Jia-jung. 2012. A grammar of Lha'lua, an Austronesian language of Taiwan. PhD dissertation, James Cook University.
- Pugach, Irina, Alexander Hübner, Hsiao-Chun Hung, Matthias Meyer, Mike Carson, & Mark Stoneking. 2021. Ancient DNA from Guam and the peopling of the Pacific. *Proceedings of the National Academy of Sciences* 118(1). e2022112118. <https://doi.org/10.1073/pnas.2022112118>
- Rau, Victoria & Maa-neu Dong. 2006. *Yami Texts with reference grammar and dictionary*. Institute of Linguistics, Academia Sinica.
- Reid, Lawrence. 1966. An Ivatan Syntax. *Oceanic Linguistics Special Publication*. Honolulu: Pacific and Asian Linguistics Institute, University of Hawai'i.
- Reid, Lawrence. 1982. The demise of Proto-Philippines. In Amran Halim, Lois Carrington & Stephen A. Wurm (eds.), *Papers from the Third International Conference on Austronesian Linguistics* 2. 201–216. Canberra: The Australian National University.
- Reid, Lawrence. 2016. Accounting for variability in Malayo-Polynesian pronouns. *Journal of Historical Linguistics* 6(2). 130–64. <https://doi.org/10.1075/jhl.6.2.01re>
- Reid, Lawrence. 2020. Response to Blust, “The Resurrection of Proto-Philippines”. *Oceanic Linguistics* 59 (1/2). 374–393. <https://doi.org/10.1353/ol.2020.0017>
- Ross, Malcolm. 1995. Some current issues in Austronesian linguistics. In Darrell Tryon (ed.), *Comparative Austronesian dictionary* 1. 45–120. Berlin: Mouton de Gruyter.
- Ross, Malcolm. 2002. The history and transitivity of Western Austronesian voice and voice marking. In Faye Wouk & Malcolm Ross (eds.), *The history and typology of Western Austronesian voice systems*, 63–78. Canberra: Pacific Linguistics.
- Ross, Malcolm. 2005. The Batanic languages in relation to the early history of the Malayo-Polynesian subgroup of Austronesian. *Journal of Austronesian Studies* 1(2). 1–23.
- 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*, 295–326. Canberra: Pacific Linguistics.
- Ross, Malcolm. 2012. In defense of Nuclear Austronesian (and against Tsouic). *Language and Linguistics* 13(6). 1253–1330.
- Ross, Malcolm. 2015. Reconstructing Proto Austronesian verb classes. *Language and Linguistics* 16(3). 279–345.
- Ross, Malcolm. 2020. Comment on Blust “The Resurrection of Proto-Philippines.” *Oceanic Linguistics* 59. 366–73. <https://doi.org/10.1353/ol.2020.0016>
- Rubino, Carl Ralph Galvez. 1997. A reference grammar of Ilocano. PhD dissertation, University of California, Santa Barbara.

- Maree, Rundell. 2007. Ibatan: A grammatical sketch of the language of Babuyan Claro Island. *Linguistic Society of the Philippines Special Monograph Issue* 53. Manila: Linguistic Society of the Philippines.
- Sagart, Laurent. 2004. The higher phylogeny of Austronesian and the position of Tai-Kadai. *Oceanic Linguistics* 43(2). 411–444. <https://doi.org/10.1353/ol.2005.0012>
- Sagart, Laurent. 2014. In defense of the numeral-based model of Austronesian phylogeny, and of Tsouic. *Language and Linguistics* 15(6). 859–82.
- Sapir, Edward. 1921. *Language: An introduction to the study of speech*. New York: Harcourt, Brace and Company.
- Shepherd, John Robert. 1993. *Statecraft and political economy on the Taiwan frontier, 1600–1800*. Stanford: Stanford University Press.
- Smith, Alexander. 2017. The Western Malayo-Polynesian Problem. *Oceanic Linguistics* 56(2). 435–490. <https://doi.org/10.1353/ol.2017.0021>
- Starosta, Stanley. 1995. A grammatical subgrouping of Formosan languages. In Paul Li, Cheng-hwa Tsang & Chiu-yu Tseng (eds.), *Austronesian studies relating to Taiwan*, 683–726. Symposium Series of the Institute of History and Philosophy, Academia Sinica, no. 3. Taipei: Academia Sinica.
- Steels, Luc & Eörs Szathmáry. 2018. The evolutionary dynamics of language. *Biosystems* 164. 128–37. <https://doi.org/10.1016/j.biosystems.2017.11.003>
- Sung, Li-May. 2018. *A sketch grammar of Kanakanavu* (in Chinese). New Taipei: Council of Indigenous Peoples.
- Tanankingsing, Michael. 2009. A functional reference grammar of Cebuano. PhD dissertation, National Taiwan University.
- Tätte, Kai, Ene Metspalu, Helen Post, Leire Palencia-Madrid, Javier R. Luis, Maere Reidla, Anneliis Rea, Erika Tamm, Everett J. Moding, Marian M. de Pancorbo & Ralph Garcia-Bertrand. 2021. The Ami and Yami aborigines of Taiwan and their genetic relationship to East Asian and Pacific populations. *European Journal of Human Genetics* 29, 1092–1102. <https://doi.org/10.1038/s41431-021-00837-6>
- Teng, Stacy Fang-ching. 2008. *A reference grammar of Puyuma, an Austronesian language of Taiwan*. Canberra: Pacific Linguistics.
- Thomason, Sarah & Terrence Kaufman. 1988. *Language contact, creolization and genetic linguistics*. Berkeley: University of California Press. <https://doi.org/10.1525/9780520912793>
- Trejaut, Jean A., Toomas Kivisild, Jun Hun Loo, Chien Liang Lee, Chun Lin He, Chia Jung Hsu, Zheng Yuan Li & Marie Lin. 2005. Traces of archaic Mitochondrial lineages persist in Austronesian speaking Formosan populations. *PLoS Biology* 3. e247. <https://doi.org/10.1371/journal.pbio.0030247>
- Tsukida, Naomi. 1993. The use of *-en* form in Fataan-Amis. *Asian and African Linguistics* 22. 123–140.
- Tsukida, Naomi. 2008. Verb classification in Amis. In Mark Donohue & Søren Wichman (eds.), *The typology of semantic alignment*, 277–294. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199238385.003.0011>
- Ward, Jerome V., J. Stephen Athens and Carol Hotton. 1998. Holocene pollen records from Babeldaob island, Palau, Western Caroline islands. Paper presented at the Annual Meeting of the Society for American Archaeology, Seattle, March 29.
- Wang, Chao-li. 2014. *The vocabulary of the Tsou language* (in Chinese). Chiayi: Lantan.
- Weinreich, Uriel. 1953. *Languages in contact: Findings and problems*. The Hague: Mouton.
- Wild, Ilka. 2018. Voice and transitivity in Kanakanavu. PhD dissertation, Universität Erfurt.

- Wu, Joy Jing-lan. 2006. Verb classification, case marking, and grammatical relations in Amis. PhD dissertation, State University of New York at Buffalo.
- Yamada, Yukihiro. 2014. *A grammar of the Itbayat language of the Philippines*. Himeji, Japan.
- Yeh, Marie Mei-li. 2018. *A sketch grammar of Saisiyat* (in Chinese). New Taipei: Council of Indigenous Peoples.
- Zeitoun, Elizabeth. 2000. *A reference grammar of Rukai* (in Chinese). Taipei: Yuanliu.
- Zeitoun, Elizabeth. 2007. *A grammar of Mantauran (Rukai)*. Taipei: Institute of Linguistics, Academia Sinica.
- Zeitoun, Elizabeth. 2018. *A sketch grammar of Rukai* (in Chinese). New Taipei: Council of Indigenous Peoples.
- Zeitoun, Elizabeth, Tai-hwa Chu, & Lalo a Tahesh Kaybaybaw. 2015. *A study of Saisiyat morphology*. Oceanic Linguistics Special Publication no. 40. Honolulu: University of Hawaii Press.
- Zorc, David. 2020. Reactions to Blust's "The Resurrection of Proto-Philippines." *Oceanic Linguistics* 59. 394–425. <https://doi.org/10.1353/ol.2020.0018>

## Résumé

Innovation morphosyntaxique peu étudiée, la réanalyse du préfixe intransitif statif proto-austronésien (PAN) \*ma- en affixe transitif, offre de nouvelles perspectives sur le groupement linguistique dont émerge l'austronésien. Le malayo-polynésien, le sous-groupe linguistique qui comprend toutes les langues austronésiennes parlées en dehors de Taïwan, est actuellement considéré comme une branche primaire sans étroite relation discernable avec quelque sous-groupe linguistique du foyer d'origine (Blust 1999, 2009/2013; Ross 2005). Cependant, ce sous-groupe affiche la même utilisation innovante de ma- que quatre langues de l'est de Formose, dispersées sur le littoral (l'amis, le siraya, le kavalan, le basay-trobiawan) et partage la fusion des phonèmes \*C/t proto-austronésiens avec ce groupe, ce qui suggère la possibilité d'une origine commune pour ces groupes linguistiques de l'est de Formose et le malayo-polynésien. Cette observation conduit à postuler un sous-groupe plus conforme à la représentation socio-historique selon laquelle la population issue de Taïwan descendrait d'une communauté de navigateurs dont la tradition maritime s'étendait des Batanes à l'île de Luzon. Cela correspond également aux découvertes récentes en archéologie et en génétique selon lesquelles (i) l'est de Taïwan est probablement le point de départ de la dispersion austronésienne (Hung 2005, 2008, 2019; Bellwood 2017; Bellwood & Dizon 2008; Carson & Hung 2018) et (ii) les Amis manifestent une relation spécifiquement plus étroite avec les communautés austronésiennes en dehors de Taïwan (Capelli et al. 2001; Trejaut et al. 2005; Tätte et al. 2021; Pugach et al. 2021). Un examen ultérieur des innovations communes aux langues malayo-polynésiennes et de l'est de Formose pourrait aider à clarifier la nature de leurs liens.

## Zusammenfassung

Die Reanalyse des proto-austronesischen (PAN) Stativpräfix \*ma- als Transitivaffix ist eine wenig beachtete morphosyntaktische Innovation und bietet neue Erkenntnisse in Bezug auf die übergeordnete Gruppierung des Austronesischen. Das Malayo-Polynesische, der Zweig

der alle austronesischen Sprachen außerhalb Taiwans umfasst, wird momentan als primärer Zweig gesehen, ohne identifizierbares näheres Verhältnis zu den Zweigen in Taiwan (Blust 1999, 2009/2013; Ross 2005). Es zeigt aber den gleichen innovativen Gebrauch von ma- und die Verschmelzung von PAN \*C/t wie vier ostformosische Sprachen (Amis, Siraya, Kavalan und Basay-Trobiawa), die verstreut an der Küste Taiwans liegen. Dies deutet darauf hin, dass der ostformosische und der malayo-polynesische Zweig einen gemeinsamen Ursprung haben könnten. Diese Beobachtung deutet auf eine neue Gruppierung hin, die besser mit einem sozio-historischen Szenario übereinstimmt, in dem die Bevölkerung außerhalb Taiwans von einer seefahrender Gruppe abstammt, die in die Batanen und Luzon expandierte. Dies passt auch zu jüngsten Erkenntnissen aus Archäologie und Genetik, die zeigen, (i) dass der Osten Taiwans wahrscheinlich der Ausgangspunkt für die Ausbreitung des Austronesischen ist (Hung 2005, 2008, 2019; Bellwood 2017; Bellwood & Dizon 2008; Carson & Hung 2018) und (ii) dass die Amis ein viel engeres Verhältnis mit austronesischen Gruppen außerhalb Taiwans haben denn innerhalb (Capelli et al. 2001; Trejaut et al. 2005; McColl et al. 2018; Pugach et al. 2021; Tätte et al. 2021). Zukünftige Studien zu mehr gemeinsamen Innovationen zwischen dem Malayo-Polynesischen und dem Ostformosischen könnten weiteres Licht auf ihr Verhältnis werfen.

## Address for correspondence

Victoria Chen  
School of Linguistics and Applied Language Studies  
Victoria University of Wellington  
VZ 302, 26/28 Kelburn Parade  
WELLINGTON 6012  
New Zealand  
victoria.chen@vuw.ac.nz

## Co-author information

Jonathan Kuo  
Department of English  
National Taipei University of Technology  
kuojona@ntut.edu.tw

Isaac Stead  
Max Planck Institute for Evolutionary  
Anthropology  
isaac\_stead@eva.mpg.de

Maria Kristina S. Gallego  
University of the Philippines Diliman  
The Australian National University  
kristina.gallego@anu.edu.au



## **Publication history**

Date received: 17 March 2021

Date accepted: 16 January 2022

Published online: 16 May 2022