Two types of alignment change in nominalizations: Austronesian and Japanese

Edith Aldridge   Yuko Yanagida
Academia Sinica   University of Tsukuba

Abstract
This paper investigates two instances of alignment change which both resulted from reanalysis of a nominalized embedded clause type, in which the external argument was marked with genitive case and the internal argument was focused. We show that a subject marked with genitive case in the early development of Austronesian languages became ergative-marked when object relative clauses in cleft constructions were reanalyzed as transitive root clauses. In contrast to this, the genitive case in Old Japanese nominalized clauses, marking an external argument, was extended to mark all subjects. This occurred after adnominal clauses were reanalyzed as root clauses. Japanese underwent one more step in order for genitive to be reanalyzed as nominative: the reanalysis of impersonal psych transitive constructions as intransitives.

    With these two case studies of Austronesian and Japanese, we show that reanalysis of nominalization goes in either direction, ergative or accusative, depending on the syntactic conditions involved in the reanalysis.

Keywords: active-inactive, cleft, nominative case, psych predicate, split-ergativity, syntactic reconstruction, unaccusative

1. Introduction
The term "alignment" refers to the distribution of case markers on subjects and objects. When referring to alignment, grammatical functions are often identified as "A" for transitive subject, "S" for intransitive subject, and "O" or "P" for transitive object. In this paper, we use the term "subject" to refer collectively to "A" and "S" roles, while we use the term "object" for the "O/P" relation. Modern Standard Japanese is a language with accusative alignment, i.e., "A" and "S" arguments take the same case marking, so both transitive and intransitive subjects have nominative case, while transitive objects have a different case, accusative, as shown in (1).

(1)   Modern Standard Japanese
    a. *Hanako ga   kita*
       Hanako NOM   came
       'Hanako came.'
    b. *Hanako ga   ringo o   tabe-ta*
       Hanako NOM   ringo ACC   eat-PST
       'Hanako ate an apple.'

In contrast to this, in the ergative language Dyirbal, intransitive subjects have the same marking as transitive objects, which is referred to as "absolutive", while transitive subjects take a different case, ergative.

(2)   Dyirbal (Dixon 1994: 161)
  a. *yabu*   *banaga-nʸu*
     mother.ABS   return-NONFUT
     'Mother returned.'
  b. *nguma*   *yabu-**nggu**   *bura-n*
     father.ABS   mother-ERG   see-NONFUT
     'Mother saw father.'

This paper is also concerned with split-ergativity. Split-ergativity typically involves a combination of the preceding two types and is manifested by the existence of two transitive clause types: one ergatively aligned and one exhibiting the accusative pattern. In Indo-Aryan languages, imperfective clauses are accusatively aligned, while perfective clauses are ergative. The difference in alignment can be seen in the verbal agreement. In the accusative clause type in (3a), the verb agrees with the external argument, while in the ergative clause type in (3b), agreement is with the internal argument, showing that these two NPs are the ones marked with nominative (absolutive)[1] case.

(3)   Hindi (Mahajan 1990: 72–3)
  a. **raam**   *roTii*   *khaataa*   *thaa*.
     Ram(M).NOM bread(F)   eat.IPFV.M was.M
     'Ram (habitually) ate bread.'
  b. *raam-ne*   **roTii**   *khaayii*   *thii*.
     Ram(M).ERG   bread(F).NOM eat.PFV.F   was.F
     'Ram ate bread.'

A fourth alignment type which is relevant to this paper is active alignment, in which all external arguments are marked with ergative case, while internal arguments take absolutive/nominative case. Crucially, this means that the language has two intransitive clause types, an active one with an ergative subject and an inactive one with an absolutive subject.

(4)   Basque (Rosen 1996)
  a. **Edu-k**   *liburua*   *erosi*   *du*.
     Edu-ERG   book.ABS   buy   do.3SG.3SG
     'Edu bought the book.'
  b. **Toni-k**   *kurritu*   *du*.
     Toni-ERG   run   do.3SG
     'Toni ran.'
  c. **Toni**   *goiz*   *iritsi*   *da*.
     Toni.ABS   early   arrive   be.3SG
     'Toni arrived early.'

From a diachronic perspective, an ergative clause type has been argued to emerge in accusative languages when either a detransitivized or embedded clause type in which an external argument

---

[1] In this paper, we treat absolutive case as equivalent to nominative, though there are some ergative languages for which this is not true (Legate 2003, Aldridge 2004, and others).

is marked with non-nominative case or an adposition is reanalyzed as a finite transitive clause type. For example, the Indo-Aryan ergative alignment in the perfective aspect is generally traced to a construction in Sanskrit built on the participle *-ta* (Proto-Indo-European *-to),[2] exemplified in (5b). Note the case on the external argument, glossed as 'instrumental'.

(5) Classical Sanskrit (Klaiman 1978: 205)

    a. *naro*   *vedān*   *paṭhati*
        man.NOM.SG   Veda.ACC.PL.M   recites.3SG
        'The man recites the vedas.'

    b. *narena*   *vedāḥ*   *paṭhi-tāḥ*
        man.INS.SG   Veda.NOM.PL.M   recite-TA.PL.M
        'The man recited the vedas.'

Another source of ergative alignment is the reanalysis of embedded nominalizations as transitive verbal clauses. Johns (1992) proposes that transitive clauses in the Inuit language Inuktitut are derived synchronically from nominalizations. First note that possessors are marked with the same case as transitive subjects, glossed as 'relative'.

(6) Inuktitut

    a. *anguti-up*   *nanuq*   *kapi-ja-a-0*
        man-REL   bear.ABS   stab-PASS.PTCP-3SG/3SG
        'The man stabbed the bear.'   (Johns 1992: 61)

    b. *anguti-up*   *qimmi-a*
        man-REL   dog-3SG
        'the man's dog'   (Johns 1992: 69)

Johns proposes that transitive verbs combine with a passive participle, *-ja* in (6a), which nominalizes the verb root. The external argument is merged within the nominal projection and assigned genitive case, while the internal argument is base generated outside the nominalization in subject position. The verbal projection functions as a nominal predicate in a copula construction, which can be literally translated as 'The bear is the man's stabbed one.' Gildea (1998) has proposed a similar derivation of ergative clauses from nominal predicates in copula constructions in Carib languages.

    In this paper, we examine two cases of alignment change from an embedded nominalization. One type is the case of accusative to (split-)ergative alignment in Austronesian languages, developing earlier proposals by Starosta et al. (1982), Ross (2009), and others. The other type is the case of active to accusative alignment in Japanese. In both Austronesian and Japanese, a transitive clause exhibiting non-accusative alignment was found in a type of cleft construction. In Austronesian languages, the embedded nominalization was reanalyzed as a finite verbal clause with an ergative subject and nominative object. Given the condition of ambiguity required for reanalysis to take place, we argue that this cleft was not fully biclausal. Rather, the presupposition was a reduced (and nominalized) relative clause in which the focused constituent

---

2 Various proposals have been made regarding the structure of the participle construction and the course of the reanalysis, but all are in agreement that this construction supplied the input to that reanalysis (Benveniste 1952; Cardona 1970; Pray 1976, Anderson 1977; Payne 1980; Bubenik 1989; Hook 1991; Butt 2001; Bynon 2005; Haig 2008, 2010; Dahl 2016; Butt & Deo 2017; and others).

could value nominative case in the matrix clause but still be spelled out in its argument position inside the reduced relative. Such a construction can be found in the Formosan language3 Budai Rukai. Note that the verb is nominalized, and the external argument has genitive case. Nominative case appears on the direct object, but this constituent surfaces inside the nominalized clause rather than the canonical clause-initial focus position. The surface word order is consequently equivalent to a monoclausal declarative clause, which facilitated the reanalysis.

(7)   Budai Rukai (Chen 2008: 82)
        *Ta-badh-ane   ki   tina-ini   **ka**   **laimai***
        NONFUT-give-NMLZ   GEN   mother-3SG.GEN   NOM   clothes
            ki   lalake-ini.
            OBL   child-3SG.GEN
        'The clothes are what the mother gave her child.'

The other type of alignment change that we consider is from active to accusative alignment. Modern Standard Japanese is a language with accusative alignment, as we noted at the beginning of this section. Old Japanese of the 8th century, however, exhibited active alignment in nominalizations. The genitive *ga* marks the external arguments of active verbs which were also high in animacy.4

(8)   Old Japanese
        a. *Saywopimye no kwo ga   **pire** puri-si   yama*   (MYS 868)
            Sayohime GEN child GEN scarf   wave-PST.ADN mountain
            'the mountain where the child Sayohime waved her cloth'
        b. ***yama miti wo***   *kimi ga   kopem-aku*   (MYS 4225)
            mountain road OBJ you GEN   cross-NMLZ
            'The mountain road is what you crossed over.'

Yanagida (2006, 2007) and Yanagida & Whitman (2009) argue that zero-marked objects, as in (8a) are non-specific and (pseudo-)incorporated into the verb. Following Baker (1988), they analyze (8a) as a derived intransitive. The canonical transitive construction appears in OSV as in (8b). Yanagida & Whitman (2009) and Yanagida (2012, 2018b) argue that OSV originates in a copula construction in which *wo* functions as a copula and is later reanalyzed as an accusative case marker.
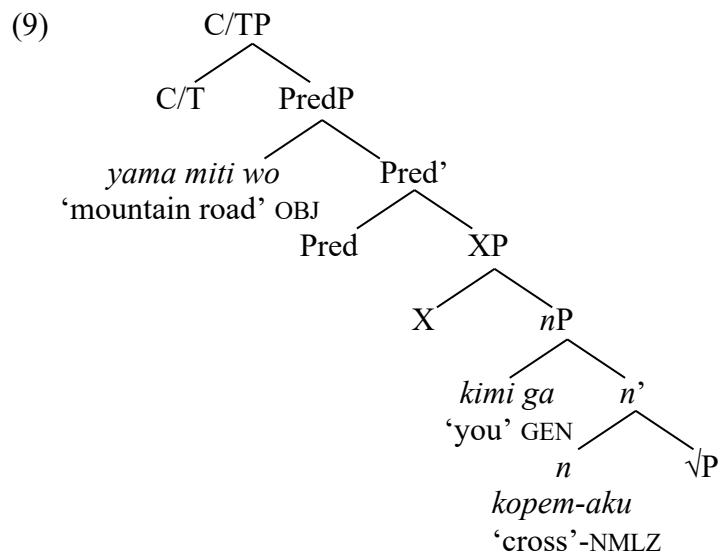
    In this paper, we adopt the Minimalist model of syntax put forth in Chomsky (2000 and subsequent works) and propose that Austronesian and Japanese have in common the characteristic that the transitive subject is marked with genitive case in nominalizations, and this case is licensed by the functional category *n* that selects a category-neutral √P. Focused objects obligatorily move to the specifier of a predication phrase (PredP) either overtly (Old Japanese) or

---

3 The term "Formosan languages" refers collectively to the Austronesian languages spoken in Taiwan. Formosan languages have a special status in Austronesian historical linguistics, given that Taiwan is the homeland of the Austronesian language family, and multiple early branches of the Austronesian family tree are represented there. Consequently, Formosan languages are the focus of the discussion in this paper on Austronesian languages. However, the term "Formosan languages" does not refer to a subgroup of this language family, since several different subgroups are represented among the Formosan languages, as we will make clear in §2.
4 Old Japanese examples are taken from *Man'yoshu* (abbreviated as MYS), the oldest collection of Japanese verse compiled in the mid-8th century A.D.

covertly (Austronesian).5 For example, in (8b), the subject marked with *ga* appears in Spec(*n*P), and the object marked with *wo* moves to [Spec, PredP]. The projection labeled "XP" houses functional material inside the nominalization like aspect.

(9)
```
                C/TP
              /      \
          C/T        PredP
                   /       \
          yama miti wo      Pred'
        'mountain road' OBJ   /    \
                          Pred      XP
                                  /    \
                                 X      nP
                                      /    \
                                 kimi ga    n'
                                'you' GEN   /  \
                                           n    √P
                                       kopem-aku
                                       'cross'-NMLZ
```

We propose that reanalysis of a nominalized embedded clause as a finite verbal clause is explained by a categorial change of the nominal head *n* to the verbalizing head *v*. In Austronesian, genitive became ergative case in transitive clauses due to this reanalysis. Intransitive clauses were unaffected, because the change took place in cleft constructions with focused objects, which are necessarily transitive. The result was accusative alignment in Japanese because the erstwhile genitive case *ga* started to appear on subjects of intransitive verbs before transitive clauses were fully reanalyzed as having nominative-accusative alignment. This was due to the reanalysis of causer arguments in impersonal psych constructions as theme subjects of unaccusative predicates. Nominative-accusative alignment resulted after the categorical change because *ga* was found on both transitive and intransitive subjects and was thus reanalyzed as nominative case.

The following two sections examine these two changes in detail, beginning with the Austronesian change from accusative to ergative in §2. Section 3 examines the change from active to accusative alignment in Japanese.

2. Genitive to ergative in Austronesian
This section discusses a change from accusative to ergative alignment in Austronesian languages. We propose that this change took place in an early branch of the Austronesian language family at a time when these languages were still spoken exclusively in Taiwan, the homeland of this language family. For this reason, evidence for our proposal comes from languages spoken currently in Taiwan.

---

5 Frellesvig, Horn & Yanagida (2015) argue that OJ has differential object marking (DOM) associated with specificity. While zero-marked objects inside VP, as in (8a) are non-specific and pseudo-incorporated into the verb, *wo*-marked objects, as in (8b) undergo Object Shift (OS) (cf. Yanagida 2006) to the outer edge of *v*P, where they are assigned accusative and at the same time receive a specific interpretation. Here we assume that *wo*-marked objects move to Spec(PredP).

Most of the Austronesian languages of Taiwan, as well as in the Philippines, exhibit a non-accusative type of alignment in their morphological case marking. In these languages, different verbal affixes seem to signal which DP in the clause has nominative case. In the following Paiwan examples, the infix *<em>* appears when the subject is the nominative argument, regardless of whether the clause is intransitive (10a) or transitive (10b). The aspect marker *<in>* appears in perfective clauses when an internal argument in a transitive clause has nominative case, the theme in (10c) and the goal in (10d). The applicative suffix *-an* is additionally required for a goal or locative argument to be given nominative case (10d).6

(10)　　　Southern Paiwan7
　　　a.　*G<**em**>uregoh　　a　　　　vatu.*
　　　　　<AV>bark　　　　NOM.CN　　dog
　　　　　'The dog is barking.'
　　　b.　*T<**em**>alem　ti　　　　ina　　　　ta　　　　qarizang.*
　　　　　<AV>plant　　NOM.PN　　mother　　OBL.CN　　bean
　　　　　'Mother plants beans.'
　　　c.　*S<**in**>aqis　　　a　　　　u-itong　　　ni　　　　ina.*
　　　　　<TR.PFV>make　　NOM.CN　　1SG.GEN-clothing　GEN.PN　　mother
　　　　　'Mother made my clothes.'
　　　d.　*P<**in**>avay-**an**　　ni　　　　ama　ta　　　　tjakit　a　　　　kakedrian.*
　　　　　<TR.PFV>give-APPL　　GEN.PN　　father　OBL.CN　　knife　NOM.CN　　child
　　　　　'Father gave the child a knife.'

Given that these languages seem to have two transitive clause types – (10b) versus (10c, d) – they are sometimes described in the literature as having a "symmetrical voice" system, i.e., the affixes *<em>* and *<in>* on the verb do not induce a change in argument structure, as would be the case in an asymmetrical voice alternation such as passive (cf. Himmelmann 2005 and references therein). However, characterizing this system in terms of voice presupposes that the nominative argument is a subject, which is a dubious claim since the nominative DP does not always perform the expected subject functions in the areas of binding, raising, or control (Schachter 1976; Payne 1982; Gerdts 1988; de Guzman 1988; Shibatani 1988, and many others).8 Instead, we adopt for a more typologically broad view and treat the Philippine and Formosan "voice system" as a type of split-ergative alignment, given that two transitive clause types can be characterized in terms of an alignment split, (10c) and (10d) being ergatively aligned, and (10b) aligning accusatively. See also Payne (1982), Gerdts (1988), de Guzman (1988), Mithun (1994), Liao (2002), Aldridge (2004, 2008, 2012), Chang (2011), and others for various ergative analyses. However, we follow tradition in Austronesian linguistics and refer to nominative-accusative transitive constructions like (10b) as "actor voice" (AV) to reflect the fact that the subject has nominative case.

---

6 There is a second applicative that selects instruments, beneficiaries, and transported themes. For simplicity, we use only the locative applicative to illustrate the applicative clause type.

7 Unless given a specific citation, Austronesian examples in this paper are taken from Edith Aldridge's field notes.

8 The only purported subject property clearly exhibited by nominative DPs is the ability to undergo relativization (Keenan & Comrie 1977). However, the restriction that only nominative DPs can undergo relativization and other types of A'-movement is a characteristic of syntactic ergativity and not of subjects in accusative languages. See Aldridge (2004, 2008) for discussion.

One principal concern in this section is with the syncretism between ergative and genitive case. In (11) the possessor in a DP takes the same *ni* case as the subject in the ergative clauses (10c, d).

(11)   Southern Paiwan
        nanemanemanga   ni   ina
        thing   GEN.PN   mother
        'Mother's things'

Following Starosta et al. (1982) and Ross (2009), we argue that the syncretism between ergative and genitive case is the result of the reanalysis of embedded nominalizations, specifically cleft constructions, as finite verbal clauses. In the cleft, the external argument had genitive case, supplied by the nominalizing morphology in the embedded clause, and the focused constituent was marked with nominative case, valued with the matrix T. As for the actor voice construction, we propose that this is a retention from Proto-Austronesian (PAn), which we take to be a language with accusative alignment, following Aldridge (2015, 2016).

2.1 Previous connections between Austronesian clause structure and nominalization
Since Starosta et al. (1982), it has been noted that affixes marking finite ergative verbs in Philippine and most Formosan languages have a diachronic connection to nominalizers. Kaufman (2009) even goes so far as to propose that Tagalog ergative clauses are synchronically built on nominalizations. For example, in Tagalog, a verb can project an ergative clause, as in (12a), or it can refer to an individual or set of individuals, as in (12b).

(12)   Tagalog
        a. *B<**in**>ili   ng   babae   ang   isda.*
            <TR.PFV>buy   GEN.CN   woman   NOM.CN   fish
            'The woman bought the fish.'
        b. *ang   b<**in**>ili*
            NOM.CN   <TR.PFV>buy
            'what was bought'

In contrast, some Austronesian languages have a distinction between nominalizing and verbalizing affixation. As an example, the Puyuma finite transitive verb in (13a) takes the suffix *-aw*, an affix which never appears on a verb in a nominalization. Note that (13a) is an ergative clause in which the object has nominative case. Ross (2009, 2012) and Aldridge (2015, 2016) propose that the affixes marking ergative verbs in Puyuma have a diachronic source unrelated to nominalization. This can be seen in (13b), showing a nominalized relative clause. The verb takes the perfective aspect marker *<in>* and the nominalizer *-an*. The transitive suffix *-aw* is never found in a nominalized clause.

(13)   Puyuma
        a. *tu=trakaw-**aw**   na   paisu kan   isaw*
            3.GEN=steal-TR   DEF.NOM   money SG.OBL   Isaw
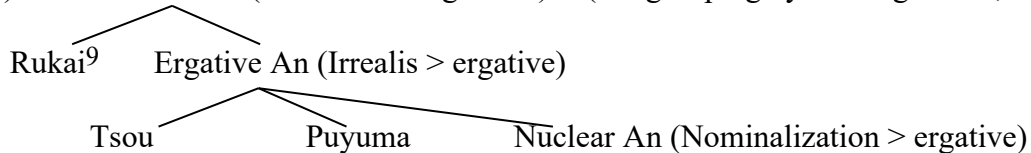   'Isaw stole the money.'   (Teng 2008: 147)

7

b. *ala   amuna   sadru   [[tu=tr<**in**>ekelr-**an**]   na   asi]*
   maybe because   many   3=<PFV>drink-NMLZ   DEF.NOM milk
   'Maybe because the milk he drank is a lot.'   (Teng 2008: 105)

Cognates of the morphemes surfacing only on nominalized verbs in Puyuma appear on finite verbs in Philippine languages and most other Formosan languages like Paiwan. Note that *-an* has been reanalyzed as a locative applicative in Paiwan. When this affix appears on a transitive verb, a goal or locative argument surfaces with nominative case, as we pointed out for (10d), which is repeated below as (14).

(14)   Southern Paiwan
   *P<**in**>avay-**an**   ni   ama   ta   tjakit   a   kakedrian.*
   <TR.PFV>give-APPL   GEN.PN   father OBL.CN   knife   NOM.CN   child
   'Father gave the child a knife.'

The dichotomy between verbal and nominal affixation is found in Puyuma and also in Rukai. Another Formosan language, Tsou, marks ergative clauses with a set of affixes similar to those found in Puyuma but has lost the nominalizing set. A common characteristic of all three of these languages is the fact that morphemes like <*in*> and *-an* are not found on finite verbs in matrix clauses. On this basis, Ross (2009, 2012) argues that these three Formosan languages – Puyuma, Tsou, and Rukai – reflect a diachronic stage predating the reanalysis of nominalizing morphemes like <*in*> and *-an* as verbal affixes. He proposes that the reanalysis took place in a language that he calls Proto-Nuclear Austronesian, which is a daughter of PAn and sister to the other three languages. The Nuclear Austronesian (NAn) subgroup today encompasses all Austronesian languages except for Rukai, Puyuma, and Tsou. Aldridge (2015, 2016, 2018) refines Ross' subgrouping hypothesis by identifying an origin for the ergative alignment found in Puyuma and Tsou. This change took place in a language she calls Proto-Ergative Austronesian (PEAn), which is sister only to Rukai. On this proposal, PAn is reconstructed with accusative alignment, which is retained in the Rukai dialects. We adopt this subgrouping proposal in this paper.

(15)   Austronesian (Accusative alignment)   (Subgrouping by Aldridge 2015, 2016)

   Rukai[9]   Ergative An (Irrealis > ergative)

         Tsou   Puyuma   Nuclear An (Nominalization > ergative)

We briefly summarize the alignment of PEAn, since it is also inherited by the NAn languages which are the main focus of this paper. Aldridge (2015, 2016, 2018) proposes that ergative alignment first emerged in Austronesian languages in irrealis mood, while realis clauses remained accusatively aligned. Thus, PEAn was a language with split-ergative alignment. Ergative alignment was later extended to realis clauses in Puyuma and Tsou, but PNAn inherited the split.

   Evidence for the historical connection between ergativity and irrealis mood can be seen in the affixes marking ergative clauses in Puyuma. Puyuma displays the same type of split-ergativity as

---

9 This proposal is in agreement with Starosta's (1995, 2001) claims that Rukai is a primary subgroup of PAn, though there are significant differences between the bases for the two claims.

is found in Philippine and other Formosan languages, with accusative alignment found in AV clauses. But the transitive and applicative affixes attaching to ergative verbs have a different diachronic source. Aldridge (2018) develops an earlier insight by Starosta (1995) in proposing that the affixes *-u* and *-i* found on irrealis ergative verbs in Puyuma are reflexes of an incorporated definite determiner and preposition, respectively. Incorporation of the determiner and preposition left the bare NPs without case-marking, so these NPs had to exceptionally value nominative case with T, resulting in the emergence of nominative objects in these constructions. This in turn was due to the lack of accusative case in irrealis clauses, this clause type being a detransitivized syntactic environment, as noted by Hopper & Thompson (1980). Consequently, the object could only value nominative case. The ergative affixes were later extended to realis verbs in Puyuma and Tsou, as can be seen by the resemblance between the two sets of affixes in (16), the realis transitive and applicative suffixes containing the *-u* and *-i* of the irrealis (exemplified by the imperative) paradigm.

(16) | Puyuma | AV | TR | APPL (adapted from Teng 2008) |
|---|---|---|---|
| Realis | *<em>V* | V-*aw* (< -*a* + -*u*) | V-*ay* (< -*a* + -*i*) |
| Imperative | V | V-*u* | V-*i* |

Note that AV verbs are affixed with the same *<em>* infix found on AV verbs in Paiwan and other Formosan languages. (17a) shows an example from Southern Paiwan. A reflex of this morpheme is also found in the accusative language Tona Rukai as the prefix *w-* marking active dynamic verbs, as in (17b).
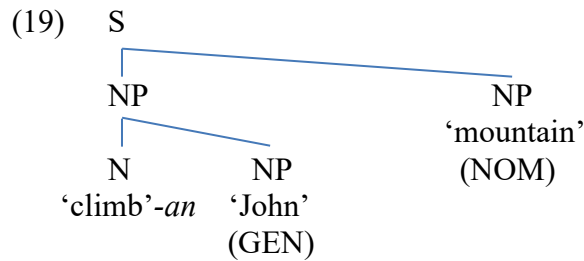
(17) a. Southern Paiwan
    *G<**em**>uregoh   a   vatu*.
    <AV>bark   NOM.CN   dog
    'The Dog is barking.'
   b. Tona Rukai
    ***w**-a-thenay*   *ki*   *tatay   namia*
    ACT.DYN-REAL-sing   NOM.PN   father   1PL.EXCL.OBL
    'Our father sings.'

It is thus uncontroversial that this affix was found in PAn, reconstructed by Wolff (1973) as **<um>* and by Ross (2009) as **M-*. Consequently, it is only the ergative verbs that were innovated, first in irrealis mood in PEAn and subsequently in realis mood, after the reanalysis of nominalizations in PNAn.

    Returning to the discussion of PEAn, Aldridge proposes that ergative alignment was limited to irrealis mood in PEAn, while realis clauses were still accusatively aligned. So it is not surprising that NAn languages retain verbal affixes like *-u* and *-i* in irrealis mood only. For example, Paiwan uses the same *-u* suffix on ergative imperative verbs.
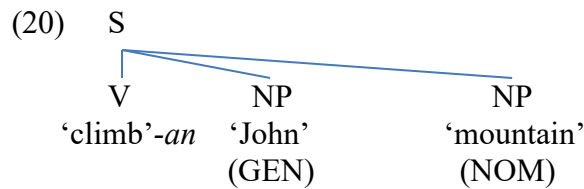
(18) Southern Paiwan
    *Santapav-**u**   i   qinaljan!*
    build-IMP   P   village
    'Build (it) in the village!'

Turning now to the innovation which produced realis ergative clauses in PNAn, Starosta et al. (1982) propose that the input structure to the reanalysis was a cleft construction comprised of a focused NP as subject and a nominalized relative clause as predicate.

(19)   S

```
         ┌──────────────────────┐
        NP                      NP
     ┌───────┐              'mountain'
     N       NP               (NOM)
  'climb'-an 'John'
            (GEN)
```

        'The place where John climbed is the mountain.'   (Starosta et al. 1982: 157)

The authors claim that biclausal copula constructions like (19) were ultimately reanalyzed as transitive verbal clauses like (20). Nominalizers like *-an consequently became verbal affixes.

(20)   S

```
     ┌─────────┬──────────────┐
     V         NP             NP
  'climb'-an  'John'       'mountain'
             (GEN)          (NOM)
```

        'John climbed the mountain.'   (Starosta et al. 1982: 157)

However, the detailed steps in the reanalysis are not spelled out; nor is a clear motivation for the change identified. According to the traditional approach to reanalysis put forth by Langacker (1977) and assumed in more recent work by Hopper & Traugott (1993), Harris & Campbell (1995), and others, syntactic reanalysis takes place in the presence of structural ambiguity. Harris and Campbell (1995: 72) specifically refer to "the patterns which have the potential for multiple structural analyses, and which thus provide the input to reanalysis". However, it is not clear how the requisite multiple structural representations could have been generated in the case in question, given the clear divergences in interpretation between a declarative dynamic transitive clause with a presupposed nominative argument and a stative copular predication in which the nominative argument receives an identificational focus interpretation. This potential problem could be avoided if the reanalysis never took place, and verbal clauses were still built on copula structures like (19), as Kaufman (2009) proposes for modern Tagalog. However, the unlikelihood of such a synchronic analysis is easily demonstrated by the fact that the nominative argument is by no means restricted to a clause peripheral position. Nominative DPs tend to surface in the base generated argument positions, as in (21a) and (21b), but they can also undergo scrambling, as in (21c).

(21)   Southern Paiwan
        a. *T<em>alem*   ***ti***   ***ina***   *ta*   *qarizang.*
            <AV>plant   NOM.PN   mother   OBL.CN   bean
            'Mother plants beans.'

b. *P<**in**>avay-an*   *ni*   *ama*   *ta*   *tjakit*   ***a***   ***kakedrian***.
      <TR.PFV>give-APPL   GEN.PN   father   OBL.CN   knife   NOM.CN   child
      'Father gave the child a knife.'

c. *S<in>aqis*   ***a***   ***u-itong***   *ni*   *ina*.
      <TR.PFV>make   NOM.CN   1SG.GEN-clothing   GEN.PN   mother
      'Mother made my clothes.'

Another difference involves information structure. In Philippine and Formosan languages, the DP with nominative case is typically definite, expressing given information, and not focused. This can be seen by contrasting the oblique object in (21a) and the nominative object in (21c) above. The nominative object in the ergative clause in (22a) is definite, while the genitive object in the actor voice clause in (22b) is indefinite and generally nonspecific. Nominative objects cannot be nonspecific.

(22)   Tagalog
     a.   *B<in>ili*   *ng*   *babae*   ***ang***   ***isda***.
        <TR.PRV>buy   GEN.CN   woman   NOM.CN   fish
        'The woman bought the fish.'
     b.   *B<um>ili*   *ang*   *babae*   ***ng***   ***isda***.
        <AV.PRV>buy   NOM.COM   woman   GEN.CN   fish
        'The woman bought a/*the fish.'

In summary, it is clear that a change from nominalization to verbal clause has taken place. In the next subsection, we propose an analysis of this change from cleft constructions containing nominalized relative clauses to monoclausal finite verbal clauses, accounting for the following changes: 1) the change from nominal to verbal status in the verbs; 2) the loss of biclausality; and 3) the loss of focus on the nominative DP.

2.2 Reanalysis
In this subsection, we propose our analysis of how biclausal cleft constructions were reanalyzed as matrix transitive clauses in PNAn. Specifically, we propose that the ergative clause type in the majority of Formosan and Philippine languages, i.e., NAn languages, ultimately traces its origin to a fully biclausal cleft construction of the type proposed by Starosta et al. (1982). But we add evidence for an intermediate stage in which the biclausal cleft was reanalyzed as an in-situ cleft construction (in the sense of Whitman 1997) in which the presupposition was a reduced nominalized embedded clause with a genitive subject. The focused constituent was the theme object in the embedded clause, but the reduced nature of the nominalization allowed this argument to value nominative case in the matrix clause while surfacing in its thematic position in the embedded clause. The surface position of the focused theme inside the relative clause in turn facilitated the reanalysis of the cleft as a monoclausal construction, since this weakened the evidence for a biclausal structure with the focused constituent in the canonical clause-initial focus position. The loss of the cleft structure in turn led to the final stage, specifically the reanalysis of the nominalization as a declarative verbal clause, since the lack of evidence for the cleft structure removed the evidence for the focus interpretation for the object.

     Given the lack of written records for Austronesian languages reflecting significant time depth, we rely solely on comparative evidence to support our proposal. Consequently, evidence for the type of construction in which the reanalysis might have occurred should be sought in the

extra-NAn languages, identified by Ross (2009, 2012) as Rukai, Puyuma, and Tsou, which do not reflect this change. Secondly, it needs to be made clear that only cleft, or other focus constructions, in which one argument values nominative case external to the nominalization can be considered as input to the reanalysis. This is because nominative case is not available in other types of nominalized constructions, so there is no evidence of ergative alignment inside the nominalization. The direct object in the Puyuma nominalized clause in (23a) has the same oblique case that a direct object has in an AV clause like (23b).

(23)   Puyuma
     a.   *k<em>adru*   *[**ku**=k<in>a-sagar-an*   ***dra***   ***suan]***
          <AV>there   1.SG.GEN=<PFV>KA-like-NMLZ   OBL   dog
          'My loving of dogs is like that.'   (Teng 2008: 142)
     b.   *tr<em>akaw* ***dra***   ***paisu***   *i*   *isaw*
          <AV>steal   INDF.OBL   money   SG.NOM   Isaw
          'Isaw stole money.'   (Teng 2008: 147)

Nominative case appears only on DPs valuing their case with finite T. Example (24a) shows a monoclausal AV clause, where the subject has nominative case. (24b) is a biclausal cleft construction in which an interrogative pronoun is the focused constituent and values nominative case in the matrix clause. The rest of the construction comprises a nominalized relative clause. Note further the nominalizing morphology on the verb and genitive case on the subject in the embedded clause in (24b).

(24)   Puyuma
     a.   *bəray=ku*   *ḍa*   *kuraw* *ḍa*   *ŋiaw*   (Tan 1997: 11)
          give=1.SG.NOM   OBL.INDF   fish   OBL.INDF   cat
          'I gave a fish to a cat.'
     b.   *a*   *manay* *[nu=b<**in**>əray*   *kan*   *aʈuŋ]*   (Tan 1997: 116)
          NOM   what   2.SG.GEN=<NMLZ>give   OBL   aʈuŋ
          'What is the thing that you gave to Aʈuŋ?'

The order posited by Starosta et al. (1982), where the nominative DP surfaces in clause-final position, obtains if the relative clause is topicalized. This construction is also clearly biclausal, since the matrix predicate *i namali* 'SG.NOM my.father' appears following the copula. The topicalized clause is also followed by a topic marker.

(25)   Puyuma (Teng 2008: 192)
          [na   *pu-ka-lrikudran*   *m-inatray] i,*   *amau*   *i*   *namali*
          DEF.NOM   CAUS-KA-behind   AV-die   TOP   COP   SG.NOM   my.father
          'The one who died is my father.'

In what follows, we propose that the reanalysis in question began in a focus construction that was already monoclausal, the type of construction that Whitman (1997) calls an "in-situ" cleft,[10] in which the focused constituent surfaces in its argument position inside a nominalized clause. Our proposal is in large part based on Aldridge's (2017) account of this reanalysis. First,
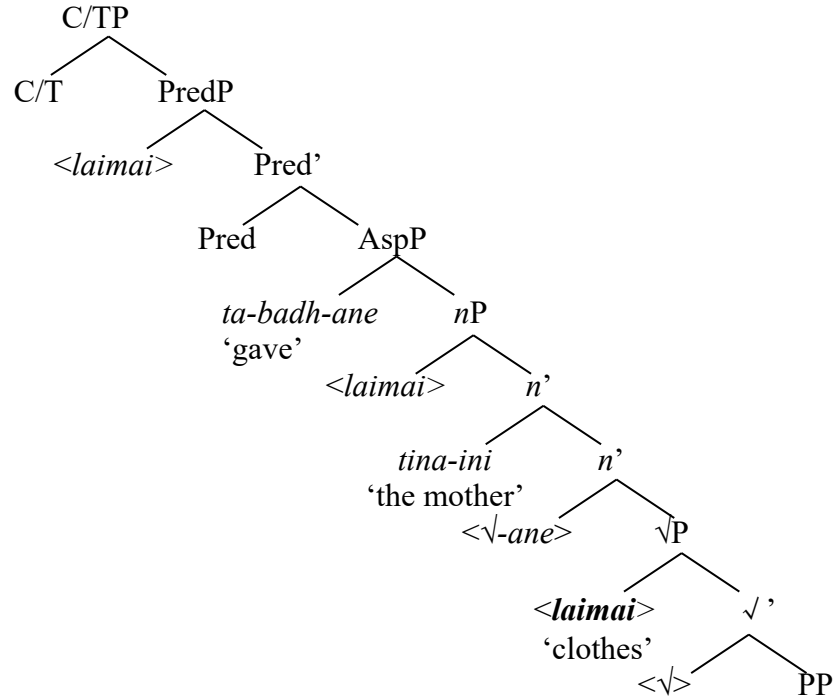
---

[10] Whitman proposes such an analysis for focus constructions in Early Middle Japanese and Modern Sinhala.

Aldridge observes with Chen (2008) that the Budai dialect of Rukai has an in-situ focus construction in which the verb is nominalized and the focused constituent appears with nominative case, as shown in (26). Chen refers to this construction as "objective voice", since the theme argument always has nominative case. But he also acknowledges that the affix *-ane* is a nominalizer. A key point to note is that the focused constituent is both preceded and followed by material internal to the nominalized clause.

(26)   Budai Rukai (Chen 2008: 82)
        Ta-badh-ane   *ki*   *tina-ini*   **ka**   **laimai**
        NONFUT-give-NMLZ   GEN   mother-3SG.GEN   NOM   clothes
            ki   *lalake-ini*.
            OBL   child-3SG.GEN
        'The clothes are what the mother gave her child.'

Aldridge proposes that the focused constituent moves to [Spec, PredP], where it functions as the matrix subject. The embedded reduced relative clause serves as its predicate. She derives the surface word order by positing additional movements of relative clause-internal material into the matrix clause so that it precedes the focused constituent in surface order. However, she does not present any evidence for these subsequent movements, so we propose a simpler alternative in which the focused constituent surfaces in-situ in its argument position in the reduced relative clause. Specifically, we maintain the proposal that the focused constituent undergoes movement to the specifier of PredP in order to value nominative case and check a focus feature, but post-syntactically it is spelled out in its base position inside the nominalized clause. With Aldridge (2017), we assume that the surface position of the focused theme following the embedded genitive subject facilitates the reanalysis of this construction as monoclausal. Specifically, the fact that the nominative DP surfaces inside the relative clause reduces the evidence for a biclausal cleft construction.

(27)

```
                        C/TP
                   ┌──────┴──────┐
                 C/T           PredP
                         ┌───────┴───────┐
                     <laimai>          Pred'
                                  ┌───────┴───────┐
                                Pred            AspP
                           ta-badh-ane      ┌────┴────┐
                             'gave'     <laimai>     n'
                                              ┌───────┴───────┐
                                          tina-ini           n'
                                        'the mother'    ┌──────┴──────┐
                                                    <√-ane>          √P
                                                            ┌─────────┴─────────┐
                                                        <laimai>               √'
                                                        'clothes'         ┌─────┴─────┐
                                                                        <√>          PP
```

Indirect support for this proposal comes from a similar construction in Sinhala. Slade (2011, 2018) shows that Sinhala has a biclausal cleft construction in which the focused constituent occupies a clause-peripheral position, while the presupposition is expressed as a participial relative clause. The focused constituent is marked with the emphatic particle *ya*. The participle inflection is indicated by the gloss 'E'. Note the accusative case on the embedded subject, suggesting that this clause is reduced and nonfinite, so nominative case is not available for the subject. This is illustrated by the Literary Sinhala construction in (28a). This construction has been reanalyzed as monoclausal, as shown by the Modern Sinhala example in (28b), where the focused constituent appears in-situ inside the clause. Note that participle inflection still appears on the verb, but this clause is finite, as can be seen by the presence of nominative case on the subject. The function of the erstwhile participle now is to mark the scope of the focus. (28c) shows that in the absence of a clause-internal focus constituent, the verb takes the unmarked finite inflection, glossed as '-A'.[11]

(28) a. Literary Sinhala (Slade 2011: 46)
          [mā   yanne]   gamaṭa^F ya
          I.ACC   go.PRS.E   village.DAT   YA
          'It is to the village that I go.'
      b. Modern Colloquial Sinhala (Slade 2011: 44)
          mamə gaməṭə^F (y/tamay)   yanne
          I.NOM village.DAT   EMPH   go.PRS.E
          'It is to the village that I go.'

---

11 See also Gair (1983), Kishimoto (1992, 2005), and Hagstrom (1998) for other analyses of Sinhala in-situ clefts.

c. Modern Colloquial Sinhala (Slade 2011: 45)
mamə gamətə yanna
I.NOM village.DAT go.PRS.A
'I go to the village.'

There is evidence that the Budai in-situ cleft was also reanalyzed from a biclausal cleft like (28a), in which the focused constituent must surface in a position external to the nominalized clause. Such constructions were shown for Puyuma in (24b) and (25). Tanan Rukai has the same type of construction. A focused object appears in clause-initial position and cannot surface inside the nominalized clause. Note further the suffix *-ani* on the verb in (29), which is cognate with the nominalizer *-ane* in Budai, as well as the *-an* nominalizer in Puyuma. The perfective infix *<in>* also appears on the nominalized verb. Recall that *<in>* is found only in nominalizations in extra-NAn languages like Puyuma and Rukai.

(29)   Tanan Rukai (Li 1973: 109)
*kay   'aysu   b<**in**>aay-**ane**   naku-a   ina   marudrang.*
this   money   <PFV>give-NMLZ   1SG-ACC   that   old.man
'This money was given to me by that old man.'

Our position that the Budai in-situ cleft represents an innovation is demonstrated by a merger that resulted in a restriction such that only the theme argument can be focused in the in-situ cleft. In Tanan and other dialects of Rukai, relative clauses in which the gap is a theme are formed on just the nominalizer *-ani*. The prefix *a-* is also added to express imperfective aspect, as shown in (30a). The combination of *ta-* and the nominalizer *-ani* appears when a goal or locative argument is extracted, as in (30b).

(30)   Tanan Rukai
a.   *w-aga=su   sa   aga   sa   [a-kane-ane=ta   ki   maum]*
PST-cook=2.SG   INDF   food   INDF   IPFV-eat-NMLZ=1.PL.INCL   P   night
'Did you cook dinner (lit: 'Did you cook what we will eat tonight')?'
b.   *[ludhaa   kuadra   ta-turavai-ane-li]   ka   daili*
tomorrow   that   TA-work-NMLZ-1SG.GEN   NOM   far
'Where I am working tomorrow is far away.'

In contrast to this, the prefixes *ta-* and *a-* in Budai in-situ clefts have a temporal or aspectual sense, glossed by Chen (2008) as 'non-future' and 'future', as shown in (31a) and (31b), respectively. Note that *ta-* in the Tanan locative relative in (31b) can be used even with prospective aspect in the embedded clause, indicating that it does not have the 'non-future' sense that it does in Budai.

(31)   Budai Rukai (Chen 2008: 91)
a.   ***Ta**-alup-ane   ki   tara-alupu   ka   lrava.*
NONFUT-hunt-NMLZ   GEN   AGT.NMZL-hunt   NOM   flying.squirrel
'The flying squirrel is what a hunter captured.'

b. *A-alup-ane   ki   tara-alupu   ka   lrava.*
   FUT-hunt-NMLZ   GEN   AGT.NMZL-hunt   NOM   flying.squirrel
   'The flying squirrel is what a hunter will capture.'

Furthermore, Chen (2008) shows that the Budai in-situ cleft is only acceptable when a theme argument is focused, as is the case in both (31a) and (31b). Hence, it can be seen that a merger has taken place in the Budai in-situ cleft construction resulting in the loss of the goal/locative relativization strategy after the reanalysis of *ta-* from a locative relativizer to an aspect marker. Given that the change in question involves a merger, the Budai in-situ cleft is clearly an innovation, since mergers are irreversible. Note, however, that this merger was also accompanied by a split. Specifically, the locative relativization strategy is retained but is only found in true relative clauses containing a gap, as in (32). Given that all Rukai dialects employ the *ta- -ane* strategy for forming relative clauses on locative gaps, this is clearly a retention. The fact that the in-situ cleft is only found in Budai, and that its formation involved a merger makes it clear that this construction must be an innovation.

(32)   Budai Rukai (Chen 1999: 18)
     Ma-kaeLa   ku   [ta-tualath-ane-ta].
     STAT-different NOM   NMLZ-originate-NMLZ-1PL.GEN
     'You and I come from different places.' (Lit: 'Our origins are different.')

In this way, Budai Rukai is similar to Sinhala in possessing an innovative cleft construction in which the focused constituent surfaces in its argument position inside the clause which expresses the presupposition. The verb in this clause type is also marked with an inflection distinct from a typical finite verb. However, the Budai in-situ cleft is not completely parallel to the Modern Sinhala focus construction in (28b). The Modern Sinhala in-situ cleft is fully monoclausal and finite, as evidenced by the presence of nominative case on the subject rather than the focused constituent, so the verbal inflection no longer signals embedding but functions merely to mark the scope of the focus. In contrast to this, the verb is still nominalized in the Budai in-situ cleft. This is demonstrated in part by the fact that the embedded subject has genitive case. Like other Rukai dialects, Budai is an accusative language, so subjects surface with nominative case in finite verbal clauses. Furthermore, the in-situ cleft differs from finite clauses in certain interpretive properties. According to Chen (2008), the active[12] voice sentence in (33a) is clearly past tense, and the action has been completed. In contrast, the in-situ cleft construction in (33b) can be either present or past tense; the event is not completed; and it can also express a habitual situation.

(33)   Budai Rukai
    a. *Wa-kane   ku   babui   ka   cumai.*   (Chen 2008: 77)
       NONFUT-eat   ACC   boar   NOM   bear
       'The bear ate a boar.'
    b. *Ta-kane-ane   ki   cumai ka   babui.*   (Chen 2008: 88)
       NONFUT-eat-NMLZ GEN   bear   NOM   boar
       'The boar is/was what the bear (usually) ate (at).'

---

12 Recall that Rukai dialects are accusatively aligned. So this construction is a typical active voice construction and not an "actor voice" construction of the type found in the split-ergative Formosan and Philippine languages.

Chen analyzes this construction aspectually as a derived state, containing both a culmination point and a consequent state. The embedded reduced relative clause expresses the culmination, and the nominalizing morpheme *-ane* maps this culmination to a consequent state. Additional evidence that this construction is aspectually complex comes from the location and function of the perfective aspect marker. (34) shows that this morpheme, *-nga*, follows the nominalizer *-ane* on the verbal complex, suggesting that it is located structurally outside of the nominalization. Chen (2008) proposes that its semantic function is to focus on the culmination point expressed by the nominalization.

(34) Budai Rukai (Chen 2008: 96)
      *Ta-kane-ane-**nga**   ki   cumai ka   babui.*
      NONFUT-eat-NMLZ-PFV   GEN   bear   NOM   boar
      'The boar is what a bear has already eaten (at).'

The Budai in-situ cleft represents an important intermediate step in the reanalysis of nominalizations to verbal clauses, because the nominative theme surfaces in a position internal to the nominalized clause. Thus, the case marking and word order in this construction are identical to what is found in ergative constructions in NAn languages. However, the clause is still nominalized, and the nominative DP continues to receive a focus interpretation. Consequently, it is clear that this construction has not yet been reanalyzed as a neutral declarative clause type.

   It is this step in the reanalysis that we turn to in the following discussion. We propose that the final crucial step in the reanalysis was the loss of the nominalizing inflection on the verb and the concomitant loss of focus on the object in the in-situ cleft. Recall that the Sinhala in-situ cleft construction retains the erstwhile participle marking *-E* on the verb, even though this verb is fully finite and the subject surfaces with nominative case, as shown in (35a). This contrasts with the neutral finite inflection *–A* shown in (35b). The function of *-E*, then, is clearly to mark this construction as a type of cleft containing a focused constituent.

(35)   Modern Colloquial Sinhala (Slade 2011: 44–45)
      a. *mamə gaməʈə^F (y/tamay) yann**e***
         I.NOM village.DAT   EMPH   go.PRS.E
         'It is to the village that I go.'
      b. *mamə gaməʈə   yann**a***
         I.NOM village.DAT   go.PRS.A
         'I go to the village.'

In the following discussion, we present evidence that the earliest ergative clauses which were fully finite verbal constructions did not have the nominalizer *-ane* inflection on the verb. We further propose that the change in the verbal morphology was a consequence of the loss of the focus interpretation for the nominative DP. These two changes together resulted in the reanalysis of the erstwhile nominalization as a neutral transitive clause type. The evidence for this proposal comes from other Formosan languages which have ergative constructions in which only a theme argument can surface with nominative case, just as we showed for Budai in-situ clefts. Unlike the Budai in-situ cleft, however, this new construction has no nominalizing morphology on the verb, and the nominative DP does not receive a focus interpretation.

According to Zeitoun & Teng (2016), the two Formosan languages Kanakanavu and Saaroa have an ergative clause type which was reanalyzed from a nominalization, but a reflex of the PAn nominalizing suffix *-an does not appear on the verb. Interestingly, in this ergative clause type, which they label "undergoer voice", only the theme can surface with nominative case. Such an example is shown in (36a). The fact that this construction is diachronically connected to a nominalization is evidenced by the presence of the *<in>* perfective aspect marker, found only in nominalized clauses in extra-NAn languages like Puyuma and Tanan Rukai. Additionally, the external argument in the clause is marked with genitive case. In contrast to this, if a different internal argument, e.g., a goal or location, has nominative case, then a nominalization marked with *-an* must be used, as in (36b). Interestingly, a theme relative clause is also nominal when it is marked with *-an*, as in (36c).

(36)　　　Kanakanavu (Teng & Zeitoun 2016)
　　　a.　*c<**in**>apa=maku　　　'alam.*　　　　　　　　　　(p. 138)
　　　　　<PFV.UV>roast=1SG.GEN　meat
　　　　　'I roasted meat.'
　　　b.　*cikiringa cakuran=ia, ni-pe-pacal-**an**-in　　　vavulu.*　(p. 145)
　　　　　side.river=TOP　　　　PFV-CAUS-die-LOC.NMLZ-3GEN　wild.pig
　　　　　'As for the riverside, it is the place where he killed wild pigs.'
　　　　　(lit. 'As for the riverside, (it) his pig-killing place.')
　　　c.　*sua　　[ni-kalʉ'-**a(n)**=maku=ia] 'a:cu　　ni-ara-[a]ka.*　(p. 146)
　　　　　NOM　PFV-like/love-NMLZ=1SG.GEN=TOP　PFV-INCH-bad
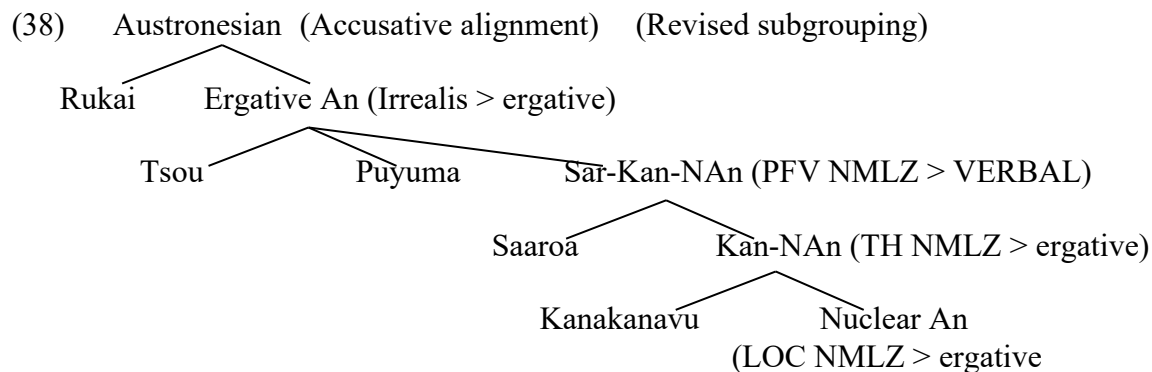　　　　　'As for my lover, s/he is dead.' (lit. 'As for the one I loved, s/he is dead.')

Clearly, what has happened is that the in-situ cleft constructions in which the theme is focused has been reanalyzed as a fully verbal root clause. When that happened, however, the verb could no longer take the nominalizing suffix *-an(e)*. We propose that the loss of the nominalizer *-an(e)* was motivated by the loss of the focus interpretation for the nominative DP. Specifically, once the in-situ cleft construction was reanalyzed as monoclausal, there was no surface evidence for a cleft (or focus) structure, so the association with focus was lost altogether. This in turn resulted in the loss of the nominalizer, since one role of this affix was to mark the scope of the focus interpretation, like the participle in Sinhala.

In a subsequent stage in the development of the Austronesian "voice system", the locative nominalizations were in turn reanalyzed as verbal, and the erstwhile nominalizer *-an* was reanalyzed as a locative/goal applicative, as can be seen in the Paiwan example in (37), as well as other NAn languages today. We assume that the specialization of *-an* as a locative applicative in verbal clauses was due to the prior existence of an ergative clause type in which a theme has nominative case, like the Kanakanavu example in (36a).

(37)　　Southern Paiwan
　　　　*P<**in**>avay-**an**　　ni　　ama　ta　　tjakit　a　　　kakedrian.*
　　　　<TR.PFV>give-APPL　GEN.PN　father　OBL.CN　knife　NOM.CN　child
　　　　'Father gave the child a knife.'

Zeitoun & Teng (2016) revise Ross' (2009, 2012) NAn subgroup on the basis of these facts in Kanakanavu and Saaroa. They propose two new subgroups which reflect the ongoing process of

reanalysis of erstwhile nominalized clauses into finite root clauses. Incorporating their subgrouping hypothesis into the view that we assume in this paper, we obtain the following divisions. Put simply, new subgroups are posited in which the innovations are taking place incrementally. The transition from nominalization to verbal clause has only just begun in Saaroa. The perfective aspect marker, which is restricted to nominalizations in Rukai and Puyuma, has been extended to verbal AV clauses. Theme nominalizations in all aspects were reanalyzed as ergative clauses in Kanakanavu. The transition from nominalization to verbal clause is fully complete in NAn, where the PAn nominalizer *-an has been reanalyzed as a locative applicative.

(38)   Austronesian (Accusative alignment)   (Revised subgrouping)

Rukai   Ergative An (Irrealis > ergative)

Tsou   Puyuma   Sar-Kan-NAn (PFV NMLZ > VERBAL)

Saaroa   Kan-NAn (TH NMLZ > ergative)

Kanakanavu   Nuclear An
(LOC NMLZ > ergative

2.3 Summary
In this section, we proposed an analysis of the development of split-ergative alignment from embedded nominalizations in Austronesian languages. Following Ross (2009), Teng & Zeitoun (2016), and Zeitoun & Teng (2016), we proposed that the reanalysis began in a subgroup of Austronesian which excludes Rukai, Puyuma, and Tsou. The input to the reanalysis was an in-situ cleft construction in which a focused theme argument occupied its thematic position inside the nominalized clause expressing the presupposition. The embedded nominalization was a reduced relative clause, allowing the focused constituent to value nominative case with matrix T. This yielded an ergative case-marking pattern inside the relative clause in which the external argument had genitive case while the direct object was marked with nominative case.

The surface position of the nominative object inside the nominalization reduced the evidence for a biclausal structure and led to a multiple-step reanalysis from nominal to verbal clause: 1) the reanalysis of the in-situ cleft as a monoclausal construction; 2) the loss of focus for the nominative object; and 3) the loss of the nominalizing suffix on the verb in ergative clauses with nominative themes.

In the final stage of the reanalysis, nominalizations involving extraction of internal arguments other than themes, e.g., locative constituents, were also reanalyzed as finite root clauses, yielding the type of split-ergativity manifested by the majority of Formosan and Philippine languages today, i.e., the "voice system", in which different nominal arguments appear with nominative case, depending on the inflection attaching to the verb.

An anonymous reviewer points out that positing multiple innovations provides a more solid basis for subgrouping than positing just a single innovation. In this section, we developed Ross' (2009) proposal that NAn is defined on the basis of one innovation, i.e., the reanalysis of nominalizations as finite verbal clauses. We explicitly spelled out multiple incremental stages in this process, and also elucidated plausible motivations for each. This makes our proposal not only a more articulated and empirically supported account of the changes reflected in NAn but

also provides additional justification for the existence of this subgroup in showing that it is defined by a series of independent innovations rather than just one,

In the next section, we show how active alignment in Old Japanese was reanalyzed as accusative. The active alignment was found exclusively in nominalized clauses, so this change was also related to reanalysis of nominalized clauses as finite matrix clauses. But the outcome was different, due to the specific syntactic conditions related to that reanalysis.

3. Genitive to nominative in Japanese

In Old Japanese (OJ; 8th century), *ga*, the ancestor of Modern Japanese nominative case, was a genitive case, marking the subject of what Yanagida & Whitman (Y&W; 2009) identify as a "nominalized clause," represented by the adnominal form of predicates. The subject of a main declarative clause, traditionally labelled *shushi* 'conclusive', was *zero*-marked and never took *ga*. Y&W (2009) argue that most subordinate clauses in OJ were nominalizations in which the subject was marked with genitive *ga/no* or *zero*, depending both on the semantics of the subject NP and on the semantics of verbs. It is widely recognized that the genitive *ga* came to be a nominative case after the adnominal form was reanalyzed as the main clause predicate form, which occurred sometime after the 16 century (cf. Yanagida 1985 Tsuboi 2001). Importantly, OJ employed two types of genitive, *ga* and *no*. The traditional analysis fails to account for why *ga*, but not *no,* became a nominative case in Modern standard Japanese.

In this section, we argue that the genitive *ga* did not simply become a nominative case due to the merger of adnominal and conclusive clauses, but that some peculiar psych predicates, which Yanagida (2018a) labels "impersonal psych transitives" played a crucial role in the development of the nominative *ga* in the history of Japanese. For periodization, we follow Frellesvig (2010):

Old Japanese (OJ)	700–800
Early Middle Japanese (EMJ)	800–1200
Late Middle Japanese (LMJ)	1200–1600
Early Modern Japanese (EModJ)	1600–1800

3.1 Two types of genitive markers in Old Japanese
Modern Japanese (ModJ) displays a textbook example of a nominative-accusative case-marking system. *Ga* marks nominative case and *o* marks accusative case. Transitivity does not affect the case marking on the subject, as shown in (39).

(39)   Modern Japanese
    a. *watasi* **ga** *niwa de kusa* **o** *katta*
       I   NOM   garden LOC   grass ACC   cut
       'I removed the grass in the garden.'
    b. *ume no hana* **ga** *saita*
       plum GEN blossom NOM   at.peak
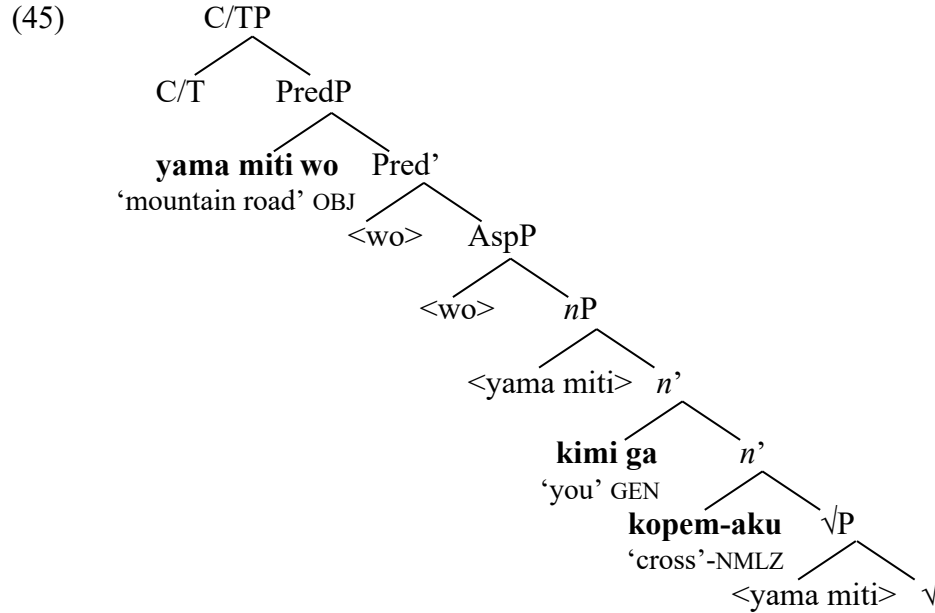   'The plum blossoms are at their peak.'

In OJ, the subject of a main declarative verb in the conclusive form is morphologically *zero*-marked; that is, nominative case in OJ is *zero*.

(40)   Old Japanese (MYS 1943; MYS 834)
       a. *ware   kusa   tore-ri*
           I   grass   remove-PFV.CONCL
           'I removed grasses.'
       b. [*ume no   pana*]   *ima   sakari nar-i*
           plum GEN blossom   now   at.peak be-CONCL
           'The plum blossoms are now at their peak.'

Ga, the ancestor of Modern Japanese nominative case, is a genitive case in OJ. There are, in fact, two distinct genitive markers: *ga* and *no*. The possessors of noun phrases are marked either with *ga* or *no* as in (41).

(41)   Old Japanese (MYS 4303; MYS 4191)
       a. [*wa=**ga** sekwo   **ga** yadwo*]
           I=GEN   lover   GEN house
           'my lover's house'
       b. [***ayu   no**   si=ga   pata*]
           sweetfish GEN   it=GEN   fin
           'sweetfish's fins'

The alternation between *ga* and *no* depends on the place of the NP in the animacy hierarchy (42) (Silverstein 1976; Comrie 1981; Dixon 1994: 85).

(42)   first/second person > third person >proper nouns> human > animate > inanimate
         ga/ga   ga   ga   no/zero no/zero   no/zero

In OJ, personal pronouns have two distinct forms: the weak or clitic forms of personal pronouns (primarily monosyllabic forms such as *(w)a* (1SG), *na* (2SG), *si* (3)) occur only with *ga* but never with *no*. The corresponding strong pronouns (*(w)are* (1SG), *nare* (2SG), and *sore* (3)) are only used in conclusive clauses, and never appear with genitive case, that is, **(w)a/na-re=ga/no. Ga* also occurs with proper nouns and kinship terms such as 'mother' and 'child', higher on the animacy hierarchy. Common NPs, lower on the hierarchy, are marked with the other genitive particle *no* or *zero*. These two types of genitive case also mark the subject in adnominal/nominalized clauses.

    Adnominal/nominalized clauses display active alignment. The use of *ga* depends not only on the semantics of the NPs, but also on the semantics of the predicates. Transitive subjects as in (43a–b) and active intransitive subjects as in (43c) are marked with *ga*, whereas inactive intransitive subjects as in (44) are marked either with *no* or *zero*. (There are, however, quite a number of instances of *no* which mark the subject of a transitive verb.)

(43)   Old Japanese (MYS 868; MYS 4225; MYS 4357)
       a. *Saywopimyeno   kwo   **ga**   pire   puri-si   yama*
           Sayohime   GEN child   AGT   scarf wave-PST.ADN mountain
           'the mountain where the child Sayohime waved her cloth'

   b. *yama   miti wo   kimi **ga** kopem-aku*
      mountain road OBJ you   AGT cross-NMLZ
      'The mountain road is what you crossed over.'
   c. *wagimokwo   **ga**   naki-si   so   [o]mopayu*
      my.wife   AGT   cry-PST.ADN   FOC   long.for
      'I long for my wife, who cries ...'

(44)   Old Japanese (MYS 3837; MYS 822)
   a. [*midu **no** tama ni   nitu-ru*]   *mimu*
      water GEN pearl DAT   resemble-ADN   see
      '(I) see water which resembles a pearl.'
   b. *wa=ga sono ni   [ume no   pana] ti-ru   pisakata no ame ywori*
      I=GEN garden LOC plum GEN blossom fall-ADN Epithet GEN sky from
      'in my garden plum blossoms fall from the sky'

As discussed in detail in Yanagida (2006, 2018b) and Y&W (2009), OJ has two types of objects. *Zero*-marked objects, as in (43a), are non-specific, and *wo*-marked objects, as in (43b), are specific. Syntactically, these two types of objects occur in different structural positions. *Zero*-marked objects appear immediately adjacent to the verb, whereas objects marked with *wo* necessarily move over the subject, resulting in [O wo S ga V] word order. Y&W (2009) argue that *zero*-marked objects are (pseudo-)incorporated into the verb, while the [O wo S ga V] transitive clause (43b) originates in a copula construction in which *wo* functions as a copula and was later reanalyzed as an accusative case marker. The projection containing the original copula was reanalyzed as AspP which selects a nominalized *v*P. Y&W (2009) further propose the hypothesis that earlier Japanese had a right-branching copula. This is because at the OJ stage, the aspectual auxiliary *ari* 'be', the modal auxiliary *e* 'be able', and the negative marker *na* all appear to the left of the lexical verb. Following the present syntactic framework, we propose that the OSV transitive clause (43b) involves overt movement of the object to [Spec, PredP] in structure parallel to (27) in Austronesian languages. This is represented in (45) (only bolded words are pronounced).

(45)

```
                C/TP
              /      \
          C/T        PredP
                    /      \
          yama miti wo     Pred'
          'mountain road' OBJ  /    \
                          <wo>    AspP
                                 /    \
                             <wo>     nP
                                     /    \
                           <yama miti>    n'
                                        /     \
                                   kimi ga     n'
                                   'you' GEN  /    \
                                      kopem-aku    √P
                                      'cross'-NMLZ /   \
                                           <yama miti>   √
```

3.2 After OJ: A change from active to accusative alignment

This section provides empirical evidence to show that Japanese underwent a significant change from OJ to Early Modern Japanese (EModJ). *Ga* used to mark canonical transitive subjects in OJ was almost lost in Early Middle Japanese (EMJ). After the agentive *ga* was lost, learners of OJ were presented with scant evidence that the object moves to the left of the subject. As a result, object movement was lost. In EMJ, a transitive clause appears with the order [S no O wo V], where the subject is marked with the other genitive particle *no*, as in (46).

(46) Early Middle Japanese (*Papakigi; Genji*)
[*ki   no   miti no   takumi*] ***no*** *yorodu no   mono* ***wo*** *tukuri idasu mo*
wood GEN tool GEN craftsman GEN various GEN   thing OBJ make out EXCLM
'The craftsman invents various things.'

The canonical nominative-accusative pattern (39a) in Modern Japanese, however, did not emerge until intransitive subjects had been fully marked with *ga* in Early Modern Japanese (EModJ).

Harris & Campbell (H&C) (1995) discuss a possible scenario for the shift from active to accusative alignment through extension whose constraints are formulated as the Complementarity Principle (H&C 1995: 259). As illustrated in Table 1, an active case B marking active intransitive and transitive subjects is extended to mark inactive intransitive subjects (They use the term "donor-recipient" relations).

Table 1: Hypothetical example of alignment change (H&C 1995: 258)

	Direct Object	Intransitive Subject		Transitive Subject
		Inactive	Active	
Before change: Active	A	A	B	B
After change: Accusative	A	B	B	B

23

A closer examination of the data in Japanese, however, shows that *ga* did not simply change from active to nominative by extending *ga* to inactive intransitive subjects. Use of *ga* became highly infrequent as an NP subject marker in EMJ before it was reanalyzed as nominative case in EModJ. Table 2 indicates the frequency of *ga* and *no* marking the subject of a verb through OJ to EModJ.[13]

Table 2:Distribution of *ga/no* marking the subject of a verb (CHJ)

	OJ *Man'yoshu*	EMJ *Genji* (1010)	EModJ *Toraakirabon* (1642)
Subject *ga* +verb	615(40%)	57 (4%)	1622 (76%)
Subject *no*+verb	957(60%)	1361 (96%)	504 (24%)
Total	1572	1418	2126

Table 2 reveals that the agentive *ga* was almost lost in EMJ (only 4% with respect to *no*). There are two major triggering events that led to the drastic increase of *ga* in EModJ. First, by the time of EModJ, the conclusive form of predicates was lost and the adnominal form was reanalyzed as the main clause predicate form. Second, the cause argument marked with *ga* in some particular object experiencer constructions was reanalyzed as the theme S argument of an unaccusative predicate (see §3.3).

Yamada (2000) examines the occurrences of nominative *ga* used in main clauses by counting the number of *ga* in the relatively colloquial LMJ text *Amakusa Heike*.[14] His findings are given in Table 3. The distribution of *ga* in LMJ differs from OJ and EMJ; *ga* appears on intransitives, in particular, unaccusative verbs and rarely marks the subject of a transitive verb.

Table 3: *Ga* in main clauses in LMJ (*Amakusa Heike* 1592, Yamada 2000)

	Transitive	Unergative	adjective	unaccusative	total
Ga	2(2%)	13(16%)	15(18%)	54(64%)	84(100%)
No	1(25%)	1(25%)	2(50%)	0(0%)	4(100%)
zero	41(33%)	18(15%)	33(27%)	18(15%)	123(100%)

Given the data in Table 3, Yamada (2000) suggests that the use of nominative *ga* in main clauses started out in LMJ by marking the subject of an unaccusative verb.

In order to investigate what types of predicates occur with *ga*, Yanagida (2017) collected the data from *Toraakirabon Kygen*, a half century after *Amakusa Heike,* which is also a relatively colloquial collection of texts, made up of kyogen (comic) plays. 50 high frequency verbs were selected out of a total of 169 verbs which appear with a *ga*-marked subject. All 2263 instances of each verb are then classified into transitive, unergative and unaccusative.[15] The frequency of occurrence of these 50 verbs selected in the text is given in Table 4.

---

[13] The quantitative study given in §3.2 and §3.3 is based on the data taken by Yanagida (2017) from the Corpus of Historical Japanese (CHJ) produced by the National Institute of Japanese Language and Linguistics, through OJ to Early Modern Japanese (EModJ). The data are taken from *Man'yoshu* (95,743 words), *Genji Monogatari* (431,130 words) (1010), and *Toraakirabon Kyogen* (207,253 words) (1642). The CHJ has no grammatical markup; thus only string searches are possible. The data in Table 2 are limited to noun+*ga/no* immediately preceding the verb.

[14] *Amakusa Heike* is a romanized translated version of the tale of *Heike* published in1592.

[15] Due to the design of the corpus, 2263 instances of verbs represent those that occur within 10 words after noun+*ga/no*. It is therefore not precisely the total occurrence of verbs with subjects marked with *ga*.

Table 4: *Ga* in EModJ (*Toraakirabon Kyogen* 1642, CHJ)

	Transitive: 20 Agent	Unergative: 5 Agent	Unaccusative: 25 Theme	Total 50
Subject=*ga*	237 (10%)	214 (10%)	1812 (80%)	2263 (100%)
Occurrences of Verbs	4479(23%)	2942(15%)	11784 (61%)	19205(100%)

The data show that the frequency of *ga* marking theme arguments (80%) is significantly higher than *ga* marking agent arguments (20%). Some examples are given in (47).

(47) Early Modern Japanese (*Toraakirabon Kyogen* 1642)
　　a. ***ame ga　furu***
　　　 rain NOM　fall
　　　 'The rain falls.'
　　b. ***mizu ga　de-ta***
　　　 water NOM　come.out-PST
　　　 'Water came out.'
　　c. ***Utubo no ke ga　nuke-ta***
　　　 Utubo GEN hair NOM　fall.out-PST
　　　 'The hair of Utubo fell out.'

Despite this text containing more unaccusative verbs (61%) than transitive/unergative verbs, the overall data are consistent with Yamada's claim that nominative *ga* started marking the non-human theme argument of an unaccusative verb rather than the agent argument of a transitive verb.

　　Recall that in OJ *are* and *sore* are the strong forms of their corresponding weak/clitic forms: *a=*(1SG) and *so=*(3). These strong pronouns are never marked with *ga* in OJ and EMJ, but there is no such restriction in EModJ. Given that the strong pronouns only appear in main declarative clauses in OJ and EMJ, examples like (48) show that *ga* was reanalyzed as a nominative case marker by the time of EModJ.

(48)　Early Modern Japanese (*Toraakibon Kyogen* 1642)
　　a. ***are ga　kane no ne　　o　　kiki-tara ba…***
　　　 that NOM　bell GEN sound　ACC　　hear-AUX if
　　　 'If that person hear the sound of the bell…'"
　　b. ***sore ga　ta　he　mizu o　　ireteoku***
　　　 that NOM　field LOC　water ACC　　put
　　　 'That person put water into the field.'

To summarize this section, the present study shows that nominative *ga* emerged through the historical processes indicated below:

　　a.　*Ga* marks agent arguments of transitive or active intransitive verbs in OJ (700-800).
　　b.　Agentive *ga* decreased drastically in its frequency in EMJ (800-1200).
　　c.　*Ga* started marking the theme arguments of unaccusative predicates

in LMJ (1200-1600).
d. *Ga* was extended to mark subjects in transitive clauses in EModJ (1600-1800).

Given that subjects in EMJ are predominantly marked with *no* rather than *ga* (see Table 2), an important question posed by the data is why *ga*, but not *no*, came to be a nominative case marker in Modern standard Japanese and why nominative *ga* started out by marking unaccusative subjects rather than transitive subjects. This question has never been raised by the traditional Japanese grammarians. In the following sections, we propose that some peculiar psych predicate constructions played a crucial role in the emergence of nominative *ga* in EModJ.

3.3 Reanalysis
This section discusses what Yanagida (2018a) labels the "impersonal psych transitive" with an unexpressed first person object experiencer. It is shown that this particular object experiencer construction found in OJ and EMJ has a causative structure in which the subject marked with *ga* denotes a CAUSE and the predicate takes a vestigial causative suffix. We propose that reanalysis of *ga*-marked cause arguments as theme arguments of unaccusatives triggered a shift from active *ga* to nominative *ga*.

3.3.1 Psych predicate constructions
Psych predicate constructions show peculiar properties across languages. To illustrate the point, consider (49).

(49) a. Little kids fear dogs.   (SE verb)
     b. Dogs frighten little kids.   (OE verb)

The psych verbs in (49) involve the same theta-roles, experiencer and theme (or sometimes called stimulus), but the structural positions are reversed in the two constructions. In (49a), the experiencer is in the subject position (SE verb), while in (49b) the experiencer is in the object position (OE verb). As is widely discussed in the literature, this reversal of theta roles associated with psych predicates challenges Baker's (1988) Uniformity of Theta Assignment Hypothesis (UTAH), which states that identical thematic arguments should have identical structural positions. To solve this problem, Pesetsky (1995) formulated a thematic hierarchy, as in (50).

(50) Thematic Hierarchy (Pesetsky 1995)
     Agent > Cause > Experiencer > Theme/Subject Matter

According to Pesetsky (1995), the subject of an OE verb is interpreted as a causer, which induces a change of state reading. Given (50), OE verbs do not lead to a violation of the UTAH, because a causer is ranked higher than an experiencer. The fact that object experiencer verbs have a causative structure is widely attested across languages. For example, in Modern Japanese object experiencer verbs appear with the causative suffix -*se*.

(51)   Modern Japanese
     a. *Taroo ga   inu o   kowaga-tta*   (SE verb)
        Taroo NOM dog ACC   fear-PST
        'Taroo feared the dog.'

26

b. *Inu ga   Taroo o   kowagara-**se**-ta*   (OE verb)
   dog NOM Taroo ACC   fear-CAUS-PST
   'The dog frightened Taroo.'

In Assamese, cited by Woolford (2008:24), the object experiencer construction contains the light verb *make/do.*

(52)   Assamese (Eastern Indo-Aryan language)
   a. *gan-tu-e   xap-tu-k   khogal korile*
      song-CLASS-ERG snake-CLASS-DAT anger made/did
      'The song angered the snake.'
   b. *boroxun-e Ram-ok   xant   korile*
      rain-ERG   Ram-DAT calm   made/did
      'The rain calmed Ram.'

(52a–b) show that the subject is the external causer argument of the light verb *korile* 'make/do' and thus is assigned ergative in Assamese.

3.3.2 Impersonal psych transitive constructions in Old Japanese
Turning again to OJ, Kikuta (2012) addresses the problem of Y&W's (2009) hypothesis that *ga* is an active case, pointing out that OJ *ga* marks the theme subject of an experiencer verb, such as *wasur-* 'forget' *omop-* 'think', *mi* 'see' etc. A closer examination of Kikuta's counterexamples, however, shows that they systematically appear with the auxiliary verb *yu* (stem *ye-*), as illustrated in (53).

(53)   Old Japanese (MYS 4407; MYS 3191)
   a. ***imo   ga   *kopisiku   wasura-**ye**-nu-kamo*
      lover   CAUS miss   forget-GET-NEG-Q
      'Did I miss my dear and cannot forget her?'
      (My dear made me forget her, didn't she?)
   b. *yama   kopyeni-si   **kimi   ga**   omopo-**yu**-raku-ni*
      mountain cross-PST   you   CAUS   think-GET-NMLZ-LOC
      'You came to my mind (You made me think about you)
       as I was crossing over the mountains.'

The auxiliary *yu* is traditionally analyzed as deriving middles, passives and potentials, and the non-passive use of *yu* predominantly appears with psych verbs, as in (54) (cf. Koji 1980).

(54) a. *wasur-* 'forget' > *wasura-**yu***
     b. *omop-* 'miss' > *omopa-**yu***

Whitman (2008) proposes that *-yu* is related to the acquisitive verb *u* (stem *e-*) 'get', which is the source of transitivity alterations in *-e-* in OJ and later stages of the language. If *yu* relates to the acquisitive verb *u* 'get', as suggested by Whitman, experiencer middles such as (53) may have an original transitive source. That is, (53a–b) can be analyzed as object experiencer predicates whose subjects serve as causers, not as theme arguments, as assumed by Kikuta (2012).

27

Psychological adjectives, such as 'sad' or 'painful' as in (55–56) have similar properties. These psych adjectives occur with the suffix -*si*.

(55) Old Japanese (MYS 4391)
  [*wa=ga kopisu namu*] ***imo*** ***ga*** *kana-**si** sa*
   I=AGT love AUX.ADN maiden CAUS sad-do NMLZ
   'The maiden I love made me sad.'

(56) Old Japanese (MYS 4338; MYS 1007)
  a. [*papa wo panarete yuku*] ***ga*** *kana-**si** sa*
    mother OBJ part go.ADN CAUS sad-do NMLZ
    'I am sad about parting from my mother.'
    (Parting from my mother made me sad.)
  b. [*tada pitori-kwo ni a-ru*] ***ga*** *kuru-**si** sa*
    only one-child DAT be-ADN CAUS pain-do NMLZ
    'It is painful to me to be the only child…'"

The suffix-*si* attaching to the adjective is homophonous with the infinitive form of the causative light verb *su* 'do'. We hypothesize that the suffix *si* has a verbal origin corresponding to English verb 'do', and that it introduces a cause argument marked with *ga*. The subject marked with *ga*, most frequently an adnominal clause as in (56), is interpreted as a causer which triggers a change of the psychological state of a first-person experiencer. These particular object experiencer predicates in OJ have the following characteristics (cf. Yanagida 2018a):

  a. They are *impersonal* in that a first person experiencer is necessarily unexpressed.
  b. The predicates are inflected with the vestigial causative light verb.
  c. An argument marked with *ga* is necessarily interpreted as a cause, but not as an experiencer.
  d. A clausal argument is marked with *ga* but never by *no*.

Impersonal psych verbs have been the subject of much discussion in historical linguistics (e.g., Fischer & van der Leek 1983; Allen 1995; Malchukov 2008; Malchukov & Siewierska 2011). Malchukov (2008) proposes that theme S intransitives in Native American languages evolved as a result of a reanalysis of "transimpersonal experiencer constructions" (the term first used by Haas 1941) with an object experiencer. We propose that a shift from genitive *ga* to nominative *ga* in Japanese involved a somewhat similar process: cause arguments in impersonal experiencer constructions were reanalyzed as the theme arguments of unaccusatives. The reanalysis of this type of psych transitive predicates provides a straightforward explanation for why *ga,* but not *no,* came to be a nominative case marker. It also accounts for why *ga* started out marking the subject of an unaccusative verb, as observed by Yamada (2000) (see Table 3–4).

3.3.3 After OJ: Psych predicate constructions
After agentive *ga* was lost in EMJ, the cause arguments of OE psych predicates continued to appear with *ga*. In particular, adnominal clauses marked with *ga* as in (57) became widespread in EMJ, as pointed out by Ohno (1977, 1978) and Yamada (2010). The data taken from the CHJ

reveal that in EMJ, there is a total of 261 tokens of *ga*-marked clauses, 135 of which are followed by a psych predicate as in (57).[16]

(57) Early Middle Japanese (*Genji, Kocho*)
[*yono supe ni   kaku suki   tamape-ru kokorobape wo miru*] ***ga***   *wokasiu mo*
life   last LOC such infatuated HON-ADN heart   ACC see   CAUS funny   also
*apareni mo **oboyu-ru** kana*
pitiful also seem-AUX EXCLM
'Seeing him infatuated with a woman in his last years made me think him
funny and pitiful.'

Table 5 now indicates the distribution of *ga*-marked subjects with non-psych and psych adjectives from OJ to EModJ.

Table 5: *Ga* (nominal/clausal arguments) with adjectives (CHJ)[17]

periodization	OJ (700-800)	EMJ (1010)	EModJ (1642)
ga+Non-psych adjective	0/0	1/1	318/98
ga+OE psych adjective	4/32	4/77	18/17

The subject of a non-psych adjective is not marked with *ga* in OJ and EMJ. This is expected since *ga* marks only the subject of an active verb in this period. Importantly, the data suggest that a cause argument of an OE psych adjective continued to be marked with *ga* after agentive *ga* was lost in EMJ. Some examples are given in (58–59).

(58) Early Middle Japanese (*Genji, Potaru*)
*Gen **ga**   ukari-**si**   sama ni pa nazurapu-beki   kepai   nara-ne-do...*
Gen CAUS annoying-DO thing LOC TOP compare-AUX.ADN appearance be-not-though
'Although I should not compare (it) with how much Gen was annoying (to me)...'

(59)   Early Middle Japanese (*Genji, Kocho; Genji, Wakana*)
a. [*kokorobape womi-ru*]   ***ga***   *woka-**si**-u mo*
kindness   ACC see-ADN CAUS thankful-do EXCLM
'I am thankful for your kindness. (Your kindness made me feel thankful).'
b. [*notamapu to   kiku*]   ***ga***   *itopo-**siku***
say   that hear.ADN CAUS   sad-do.INF
'I am sad to hear her say that. (It made me feel sad to hear her say that.)'

(58–59) have a causative structure parallel to (55–56) in OJ: both NPs and clausal complements marked with *ga* are interpreted as a cause, and a first person experiencer is necessarily unspecified in these constructions.

   From EMJ to EModJ, there was a significant change. The subject of a non-psych adjective, which was predominantly *zero* marked in OJ and EMJ, appears with *ga* in EModJ. Psych-adjectives also changed in important respects. Many of the impersonal OE adjectives in OJ and

---

16 The auxiliary *yu* attached to the psych verbs in (53) was replaced by *-ru* in EMJ.
17 The quantitative data taken from the CHJ are limited to nominals or clauses marked with *ga* immediately preceding an adjective; they do not represent the total number of the relevant instances.

EMJ, such as *wosi* 'sad' in (60), came to take an overt experiencer marked with the dative *ni* after the vestigial causative suffix *si* was lost.

(60)   Early Modern Japanese (*Toraakirabon Kyogen* 1642)
     onna   ni   nani ga   osikarafu
     woman DAT   what NOM   sad
   'What is the woman sad about?' (What is regrettable to the woman?)

In (60), the *ga*-marked argument is now interpreted as a theme, but not as a cause. The experiencer subject marked with *ni* was an innovation. The data in (60) suggest that the cause argument of an OE adjective was reanalyzed as the theme argument of a SE predicate, as syntactically represented in (61).[18]
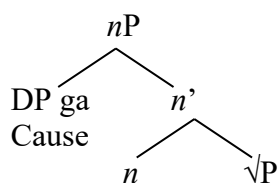
(61) Active   >   Nominative
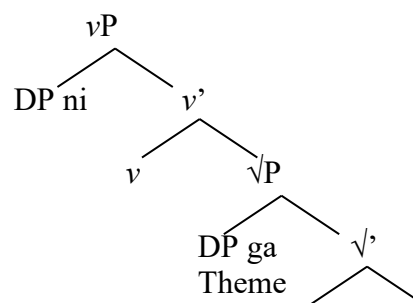
The cause argument of an OE psych predicate appears in [Spec, *n*P] in the same way as the agent argument of a transitive/unergative verb. This was reanalyzed as the internal theme argument after the nominal head *n* became verbalizing *v*. Once *ga* came to mark a theme argument, it was extended to mark all types of unaccusative subjects.

3.4 Summary
In §3, we showed that genitive *ga*, the ancestor of Modern Japanese nominative case, was used as an active case particle marking transitive and active intransitive subjects in OJ. A canonical transitive clause had a type of cleft structure in which the object moves to [Spec, PredP]. We then provided evidence to show that an active-to-accusative shift did not occur through extension based on donor-recipient relations, as hypothesized by Harris & Campbell (1995:258). Agentive *Ga* was once almost lost before it came to mark unaccusative subjects. The fluctuation in the frequency of *ga*, as shown in Table 2, implies that *ga* did not become a nominative case in one fell swoop after the merger of adnominal and conclusive predicates. Instead, many steps took place before transitive clauses were fully reanalyzed as having a nominative-accusative alignment. This is summarized as follows:

  Step I: Agenitive *ga* was lost in EMJ, which triggered the loss of object movement.
  Step II: The cause argument marked with *ga* was reanalyzed as a theme argument.
  Step III: *Ga* was extended to mark all types of unaccusative subjects.

---

18 From a cross-linguistic perspective, a historical development of OE predicates into SE predicates is well attested (cf. van Gelderen 2014). Yanagida (forthcoming) argues in detail that dative subject constructions in Modern Japanese arose as a byproduct of a change occurring from active-inactive to accusative alignment.

Step IV: *Ga* was extended to mark transitive subjects.

The impersonal psych predicates whose argument was marked with *ga* are semantically transitive but syntactically intransitive in that an object experiencer never surfaced in object position. Because of this peculiar property they were reanalyzed as unaccusatives with the sole argument marked with *ga*. Subsequently, they were further reanalyzed as transitive verbs whose subject experiencer is marked with the dative particle *ni*, which is an innovation in EModJ.

4. Conclusion

This paper has detailed two instances of alignment change, one from accusative to split-ergative in Austronesian languages and the other from active to accusative in Japanese. Both of the changes were related to the reanalysis of embedded nominalizations as finite matrix clauses. The different outcomes were induced by different syntactic conditions obtaining in the constructions in the two cases.

The Austronesian case involved an in-situ cleft construction in which a focused theme argument valued nominative case with matrix T while the external argument received genitive case internal to the nominalization. This ergative case-marking pattern was inherited directly in the descendant languages after reanalysis of the nominalized clauses as verbal.

In contrast, Old Japanese exhibited active alignment in nominalized clauses, the genitive case marker *ga* appearing on all external argument subjects which were high in animacy. The overall frequency of *ga*-marked subjects decreased in Early Middle Japanese, but *ga*-marking was consistently maintained on the cause argument in psych constructions. In Early Modern Japanese, impersonal psych predicates were reanalyzed as unaccusative, with the concomitant result that the subject in this construction was understood as a theme internal argument rather than an external argument. Once *ga* came to mark internal arguments, it could only be interpreted as a nominative marker rather than an active case.

In this way, we have shown that there is no predetermined route for alignment change. The outcome of the change is dependent on the particular syntactic environment in which the change takes place.

Abbreviations
Abbreviations not found in the Leipzig Glossing list are as follows
ACT=active
ADN=adnominal
AGT=agent
AV=actor voice
CN=common noun
CONCL= conclusive
DYN=dynamic
EMPH=emphatic
EXCLM=exclamative
HON=honorific
INCH=inchoative
NONFUT=non-future
OE=object experiencer
P=preposition
PN=personal name
SE=subject experiencer
UV=undergoer voice.

Digitalized Texts
The Corpus of Historical Japanese (CHJ), the National Institute of Japanese Language and Linguistics, https://maro.ninjal.ac.jp/

References
Aldridge, Edith. 2004. *Ergativity and word order in Austronesian languages*. Ithaca, New York: Cornell University dissertation.
Aldridge, Edith. 2008. Generative approaches to ergativity. *Language and Linguistics Compass: Syntax and Morphology* 2(5). 966-995.
Aldridge, Edith. 2012. Antipassive and ergativity in Tagalog. *Lingua* 122. 192–203.
Aldridge, Edith. 2015. A Minimalist approach to the emergence of ergativity in Austronesian languages. *Linguistics Vanguard* 1(1). 313–326.
Aldridge, Edith. 2016. Ergativity from subjunctive in Austronesian languages. *Language and Linguistics* 17(1). 27-62.
Aldridge, Edith. 2017. Intransitivity and the development of ergative alignment. In Jessica Coon, Diane Massam & Lisa Travis (eds.), *The Oxford handbook of ergativity*, 501–529. Oxford: Oxford University Press.
Aldridge, Edith. 2018. Reconstructing Proto-Austronesian alignment. Paper presented at the 20th Diachronic Generative Syntax conference (DIGS 20), York University, UK.
Allen, Cynthia. 1995. *Case marking and reanalysis*. Oxford: Oxford University Press.
Anderson, Stephen. 1977. On the mechanisms by which languages become ergative. In Charles Li (ed.), *Mechanisms of syntactic change*, 317–363. Austin: University of Texas Press.
Baker, Mark. 1988. *Incorporation: A theory of grammatical function changing*. Chicago: University of Chicago Press.
Benveniste, Emil. 1952. La construction passive du parfait transitif. *Bulletin de la Société de Linguistique de Paris* 48(1)[=1974. 192-202].

Bubenik, Vit. 1989. On the origins and elimination of ergativity in Indo-Aryan Languages. *Canadian Journal of Linguistics* 34(4). 377–398.

Butt, Miriam. 2001. A reexamination of the accusative to ergative shift in Indo Aryan. In Miriam Butt & Tracy Holloway King (eds.), *Time over matter: Diachronic perspectives on morphosyntax*, 105–141. Stanford: CSLI Publications.

Butt, Miriam & Ashwini Deo. 2017. Developments into and out of ergativity: Indo-Aryan Diachrony. In Jessica Coon, Diane Massam & Lisa Travis (eds.), *The Oxford handbook of ergativity*, 531–552. Oxford: Oxford University Press.

Bynon, Theodora. 2005. Evidential, raised possessor and the historical source of the ergative construction in Indo-Iranian. *Transactions of the Philological Society* 103(1). 1–72.

Cardona, George. 1970. The Indo-Iranian construction Mana (Mama) Kriam. *Language* 46. 1–12.

Chang, Henry Y. 2011. Transitivity, ergativity, and the status of O in Tsou. In Jung-hsing Chang (ed.), *Language and Cognition: Festschrift in Honor of James H-Y. Tai on His 70th Birthday*, 277–308. Taipei: Crane Publishing.

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

Chen, Cheng-Fu. 2008. *Aspect and tense in Rukai: Interpretation and interaction*. University of Texas, Austin, dissertation.

Chomsky, Noam. 2000. *Minimalist inquiries*. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays in Minimalist syntax in Honor of Howard Lasnik*, 89-155. Cambridge, MA: MIT Press.

Comrie, Bernard. 1981. *Language universals and linguistic typology: Syntax and morphology*. Oxford: Blackwell.

Dahl, Eystein. 2016. The origin and development of the Old Indo-Aryan predicated *-tá* construction. In Eystein Dahl & Krzysztof Stroński (eds.) *Indo-Aryan ergativity in typological and diachronic perspective*, 63–110. Amsterdam: John Benjamins.

De Guzman, Videa P. 1988. Ergative analysis for Philippine languages: An analysis. In Richard McGinn (ed.), *Studies in Austronesian linguistics*, 323–345. Athens, Ohio: Ohio University Center for International Studies.

Dixon, R.M.W. 1994. *Ergativity*. Cambridge: Cambridge University Press.

Fischer, Olga & van der Leek, Frederike. 1983. The demise of the Old English impersonal construction. *Journal of Linguistics* 19, 337–368.

Frellesvig, Bjarke. 2010. *A history of the Japanese language*. Cambridge: Cambridge University Press.

Frellesvig, Bjarke, Stephen, Horn & Yuko Yanagida. 2015. Differential object marking: A corpus based study. In D. Haug, et al. (eds.), *Historical linguistics: Current issues in linguistic theory*, 195–211. Amsterdam: John Benjamins.

Gair, James W. 1983[1998]. Non-configurationality, movement, and Sinhala focus. Paper presented at the Linguistic Association of Great Britain, Newcastle, September 1983. [Published in Gair 1998:50–64].

Gair, James W. 1998. *Studies in South Asian linguistics: Sinhala and other South Asian languages*. New York: Oxford University Press.

van Gelderen Elly. 2014. Changes in psych-verbs: A reanalysis of little *v*. *Catalan Journal of Linguistics* 13. 99–122.

Gerdts, Donna B. 1988. Antipassives and causatives in Ilokano: Evidence for an ergative analysis. In Richard McGinn (ed.), *Studies in Austronesian Linguistics*, 295–321. Athens, Ohio: Ohio University Center for International Studies.

Gildea, Spike. 1998. *On Reconstructing grammar*. New York: Oxford University Press.

Hagstrom, Paul. 1998. *Decomposing questions*. MIT dissertation.

Haig, Geoffrey. 2008. *Alignment change in Iranian Languages*. Berlin: Mouton de Gruyter.

Haig, Geoffrey. 2010. Alignment. In Silvia Luraghi & Vit Bubenik (eds.), *Continuum companion to historical linguistics*, 250–268. New York: Continuum International Publishing Group.

Haas, Mary R. 1941. Tunica. In Franz Boas (ed.), *Handbook of American Indian Lianguages*, 9–143. New York:Augustin.

Harris, Alice & Lyle Campbell. 1995. *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.

Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 110–181. New York: Routledge.

Hook, Peter. 1991. On identifying the conceptual restructuring of passive to ergative in Indo-Aryan. In Madhav M. Deshpande & Saroja Bhate (eds.), *Pāninian Studies: Professor S. D. Joshi Felicitation Volume*, 177–199. University of Michigan: Center for South and Southeast Asian Studies.

Hopper, Paul & Sandra Thompson. 1980. Transitivity in grammar and discourse. *Language* 56. 251–299.

Hopper, Paul & Elizabeth Traugott. 1993. *Grammaticalization*. Cambridge University Press.

Johns, Alana. 1992. Deriving ergativity. *Linguistic Inquiry* 23. 57–88.

Kaufman, Daniel. 2009. Austronesian nominalism and its consequences: A Tagalog case study', *Theoretical Linguistics* 35(1). 1–49.

Keenan, Edward & Bernard Comrie. 1977. Noun phrase accessibility and Universal Grammar. *Linguistic Inquiry* 8(1). 63–99.

Kikuta, Ciharu. 2012. Jodai nihongo no ga-kaku nituite [On the case marker *ga* in Old Japanese] *Dosisha Daigaku Jinbun Gakkai* [The Literary Association], Doshisha University 89, 89–123.

Kishimoto, Hideki. 1992. LF pied piping: Evidence from Sinhala. *Gengo Kenkyu* 102. 46–87.

Kishimoto, Hideki. 2005. *Wh*-in-situ and movement in Sinhala questions. *Natural Language and Linguistic Theory* 23. 1–51.

Klaiman, Miriam H. 1978. Arguments against a passive origin of the IA ergative. In *Chicago Linguistic Society: Papers from the 14th Regional Meeting*, 204–216. Chicago: Chicago Linguistic Society.

Langacker, Ronald W. 1977. Syntactic reanalysis. In Charles Li (ed.), *Mechanisms of syntactic change*. Austin: University of Texas Press, 57–139.

Legate, Julie. 2003. *Warlpiri: Theoretical Implications*. Cambridge, MA.: MIT doctoral dissertation.

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

Liao, Hsiu-chuan. 2002. The Interpretation of *tu* and Kavalan ergativity. *Oceanic Linguistics* 41(1). 140–158.

Mahajan, Anoop. 1990. *The A/A' distinction and movement theory*. MIT dissertation.

Malchukov, Andrej. 2008. Split intransitives, experiencer objects and transimpersonal constructions: (re-)establishing the connection. In Mark Donohue & Søren Wichmann (eds.), *The typology of semantic alignment*, 76–100. Oxford: Oxford University Press.

Malchukov, Andrej & Anna Siewierska. 2011. *Impersonal constructions*. Amsterdam: John Benjamins.

Mithun, Marianne. 1994. The implications of ergativity for a Philippine voice system. In Barbara Fox & Paul Hopper (eds.), *Voice: Form and function*, 247–277. Amsterdam: John Benjamins.

Ohno, Susumu. 1977. Shukaku joshi *ga* no seiritsu [The development of the nominative case particle ga], *Bungaku* 45:102–117.

Ohno, Susumu. 1978. *Bunpoo to goi*. [Grammar and Lexicon]. Tokyo: Iwanami Shoten.

Payne, John R. 1980. The decay of ergativity in Pamir Languages. *Lingua* 51. 147–186.

Payne, Thomas. 1982. Role and reference related subject properties and ergativity in Yup'ip Eskimo and Tagalog. *Studies in Language* 6(1). 75–106.

Pesetsky, David. 1995. *Zero syntax*. Cambridge, MA: The MIT Press.

Pray, Bruce R. 1976. From passive to ergative in Indo-Aryan. In Manindra K. Verma (ed.), *The notion of subject in Indo-Aryan languages*, 195–211. Madison: University of Wisconsin (South Asian Studies, Publication series 2).

Rackowski, Andrea. 2002. *The Structure of Tagalog: Specificity, voice, and the distribution of arguments*. MIT dissertation.

Rackowski, Andrea & Norvin Richards. 2005. Phase edge and extraction: A Tagalog case study. *Linguistic Inquiry* 36(4). 565–599.

Rosen, Carol. 1996. LING 401: Typology. Course notes, Cornell University.

Ross, Malcolm. 2009. Proto Austronesian verbal morphology: A reappraisal. In K. Alexander Adelaar & Andrew Pauley (eds.), *Austronesian historical linguistics and culture history: A festschrift for Robert Blust* (Pacific Linguistics 601) , 295–326. Canberra: Research School of Pacific and Asian Studies, Australian National University.

Ross, Malcolm. 2012. In defense of Nuclear Austronesian (and against Tsouic). *Language and Linguistics* 13(6). 1253–1300.

Schachter, Paul. 1976. The subject in Philippine languages: Topic, actor, actor-topic, or none of the above. In Charles Li (ed.), *Subject and topic*, 491–518. New York: Academic Press.

Shibatani, Masayoshi. 1988. Voice in Philippine languages. In Masayoshi Shibatani (ed.), *Passive and voice*, 85–142. Amsterdam: John Benjamins.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Dixon R.M.W. (ed.), *Grammatical categories in Australian languages*, 112–171. Canberra: Australian Institute of Aborignal Studies.

Slade, Benjamin. 2011. *Formal and philological inquiries into the nature of interrogatives, indefinites, disjunction, and focus in Sinhala and other Languages*. University of Illinois Ph.D. dissertation.

Slade, Benjamin. 2018. History of focus-concord constructions and focus-associated particles in Sinhala, with comparison to Dravidian and Japanese. *Glossa* 3. 1–28.

Starosta, Stanley. 1995. A grammatical subgrouping of Formosan languages. In Paul J.-K. Li, Cheng-hwa Tsang, Ying-kuei Huang, Dah-an Ho, Chiu-yu Tseng (eds.), *Austronesian studies relating to Taiwan*, 683–726. Taipei: Academia Sinica.

Starosta, Stanley. 2001. Reduplication and the subgrouping of Formosan languages. Paper presented at the International Symposium on Austronesian Cultures: Issues relating to Taiwan,

Academia Sinica. Published in Elizabeth Zeitoun (ed.), *Formosan Linguistics: Stanley Starosta's Contributions*, vol. 2, 801-834. Taipei: Language and Linguistics, 2009.

Starosta, Stanley, Andrew K. Pawley & Lawrence A. Reid. 1982/2009. The evolution of focus in Austronesian. In Amran Halim, Lois Carrington & S.A. Wurm (eds.), *Papers from the third International Conference on Austronesian Linguistics. Vol. 2: Tracking the travellers* (Pacific Linguistics C-65). Canberra: Research School of Pacific and Asian Studies, Australian National University, 145–170 (republished in Elizabeth Zeitoun (ed.), *Formosan linguistics: Stanley Starosta's contributions. Vol. 2: Publications on Formosan languages* (Language and Linguistics Monograph Series C6–65). Taipei: Institute of Linguistics, Academia Sinica, 297-328 [with an expanded version of the paper, with the same title, published for the first time in the same volume, 329–481]).

Tan, Cindy Ro-lan. 1997. *A study of Puyuma simple sentences*. Taipei: National Taiwan Normal University MA Thesis.

Teng, Stacy Fang-ching. 2008. *A reference grammar of Puyuma*. Canberra: Pacific Linguistics.

Teng, Stacy F & Elizabeth Zeitoun. 2016. The noun-verb distinction in Kanakanavu and Saaroa: Evidence. *Oceanic Linguistics* 55(1). 134–161.

Tsuboi, Yoshiki. 2001. *Nihongo katuyo taikei no hensen* (Historical change in the Japanese conjugation system). Tokyo: Kasama Shoin

Whitman, John. 1997. *Kakarimusubi* from a comparative perspective. In Ho-min Sohn & John Haig (eds.), Japanese/Korean linguistics, vol. 6, 161–178. Stanford: Center for the Study of Language and Information.

Whitman, John. 2008. The source of the bigrade conjugation and stem shape in pre-Old Japanese. In Bjarke Frellesvig & John Whitman (eds.), *Proto-Japanese*, 159-174. Amsterdam: John Benjamins.

Wolff, John. 1973. Verbal inflection in Proto-Austronesian. In Andrew Gonzales (ed.), *Essays in honor of Cecilio Lopez on his seventy-fifth birthday*, 71–91. Quezon City: Linguistic Society of the Philippines.

Woolford, Ellen. 2008. Differential subject marking at argument structure, syntax and PF. In Helen de Hoop & Peter de Swart (eds.), *Differential subject marking*, 17–40. Dordrecht: Springer.

Yamada, Masahiro. 2000. Shugo hyôji ga no seiryoku kakudai no yôso [The expansion of the use of the subject denotor *ga*: A comparison between the original text of the Tale of Heike and Amakusaban Heike]. *Kokugogaku* 51(1). 1–14.

Yamada, Masahiro. 2010. *Kakujoshi ga noTsujiteki Kenkyu* [A diachronic study of the case particle *ga*]. Hituzi:Tokyo.

Yanagida, Seiji. 1985. *Muromachi jidai o kokugo* [The language of Muromachi period] Tokyo: Tokyo Do.

Yanagida, Yuko. 2006. Word order and clause structure in Early Old Japanese. *Journal of East Asian Linguistics* 15. 37–68.

Yanagida, Yuko. 2007. Miyagawa's (1989) exceptions: An ergative analysis. *MIT Working Papers in Linguistics* 55: 265–276.

Yanagida, Yuko. 2012. The syntactic reconstruction of alignment and word order: The case of Old Japanese. In Ans van Kemenade & Nynke de Haas (eds.), *Historical Linguistics 2009*, 107–128. Amsterdam: John Benjamins.

Yanagida, Yuko. 2017. Genitive/active to nominative case in Japanese: the role of complex experiencer constructions. Paper presented at *the 23rd International Conference on Historical Linguistics*, The University of Texas at San Antonio, San Antonio.

Yanagida, Yuko. 2018a. Differential subject marking and its demise in the history of Japanese. In I. Seržant & A. Witzlack-Makarevich (eds), *Diachrony of differential argument marking*. 403–425. Berlin: Language Science Press.

Yanagida, Yuko. 2018b. Differential argument marking and object movement: A typological perspective: In Kunio Nishiyama, Hideki Kishimoto & Edith Aldridge (eds.), *Topics in Theoretical Asian Linguistics*, 181–205. John Benjamins.

Yanagida, Yuko. forthcoming. The origin of dative subjects and psych predicate constructions in Japanese. *Journal of Historical Linguistics*.

Yanagida, Yuko & John Whitman. 2009. Alignment and word order in Old Japanese. *Journal of East Asian Linguistics* 18. 101–144.

Zeitoun, Elizabeth & Stacy F. Teng. 2016. Reassessing the position of Kanakanavu and Saaroa among the Formosan languages. *Oceanic Linguistics* 55(1). 162–198.