

# AN APPLICATIVE PROCESS PERSISTING THROUGH AUSTRONESIAN: SOURCE OF VOICE DETERMINATION?

Joseph C. Finney

## Overview

This is a study in diachronic syntax of the Austronesian (AN) languages. We shall be concerned with the casemarking (syntactic case) and the semantic Case relations of the core terms of the verb, in all its voices.

We begin by reviewing my comparative study of Atayal (in the North Formosan group) and Malagasy (in the Western Malayo-Polynesian group of AN). The conclusion is that forms -an/-un/-in, verbal focus affixes, are underlyingly ANA/UNA/INA, i.e., locative preposition A/U/I followed by the demonstrative NA (“that”). The focus affix is a grammaticalization of a phrase, “at that”. This conclusion, if true, has profound consequences on our understanding of the diachronic development of the AN system.

In a few languages, the allative preposition KA/KU/KI is grammaticalized in another form of the phrase. In a few others, demonstratives of other distances (NI “this” and DA “that, remote, out of sight”), grammaticalize in the same way.

In AN as in IndoEuropean (IE), both third-person pronouns and Determiners, including articles, typically arise from demonstrative pronouns.

Here’s an example of a simple use of the phrase (not grammaticalized) from Ross (1997); it’s in Tigak, a Meso-Melanesian Western Oceanic language of New Ireland.

pikoi ina    iai  
bark PREP tree  
“tree bark, bark of a tree”

Even here, the fusion of preposition *i* and article (determiner) *na* into a single word is a slight grammaticalization. The optional use of the *i* “at” preposition for “of” is found in what I call Greater Central-Eastern MP, i.e., Central and Eastern MP languages plus adjacent Wolio.

The thrust of this paper is to examine some of the Rukai languages of Formosa for cognate forms and their place in AN’s diachronic syntax. Rukai

has come to be of special interest because of Starosta's (1994) proposal that Rukai was the first to break off from the main line of AN.

The present study finds that both Tanan Rukai and Mantaoran Rukai indeed have casemarked pronouns (or verbal suffixes descended from casemarked pronouns) which obviously descend from phrases consisting of a locative preposition (a/u/i) plus pronouns. These have rather limited applicative functions in Rukai and they do not serve as makers of voice change. Nevertheless, Tanan Rukai has such a formation with some passive uses, even though an unrelated construction is the passive voice. Rukai languages show some elements of the focus affixes found in the other major branches of AN. These facts raise the question: did Rukai have the focus voice-changing system and lose the essential part of it? Or did the other AN languages develop the system from a common ancestor not shared by Rukai? I make a judgment on the issue.

Now let's look at the facts: and begin with examining the prepositions, casemarkers, and demonstratives of AN languages in general, and PAN, that can form the relevant PPs and then grammaticalize. We note that casemarkers commonly arise from prepositions, and that in diachronic studies it is convenient to speak of PCMs (prepositions-casemarkers).

### **Prepositional phrase: a-na = "at that", "there"**

The table below (based on evidence that may be given in another paper) shows the composition of prepositions and casemarkers (henceforth PCM) and Demonstratives (Dem) / Determiners (Det) in Proto-Austronesian (henceforth PAN) and Austronesian (AN) languages generally. Choose any consonant from column 1 plus any vowel from column 2 to make a PCM, and add a Dem/Det from column 3, which may be fused with the PCM to form a casemarked Dem/Det. Each Dem/Det refers to a distance (from this-here to that-there). Glosses show the likely PAN semantics.

1. Consonant	2. Vowel	3. Dem/Det
zero (at,in)	i (person,	ni (this, near me)
glottal (topic)	proper name)	a
n (from,of,by)	u (unmarked,	tu
k (to,toward)	neutral)	na (that)
s (not Gen/Erg)	a (nonspec,	da (that, out of
d (not Gen/Erg)	common noun)	sight)
t (not Gen/Erg)		

With the disappearance of the initial glottals in most AN languages, the glottal case (Topic > Nominative/Absolutive) becomes indistinguishable from the zero case, which is locative but comes to take accusative uses in

some languages. We'll call the zero-form locative and the k-form allative. In AN as in other language families, the allative form can get dative use, and the locative and allative forms can get accusative use. AN allatives also (perhaps through "turn into") acquire use as predicate nouns or presentatives; and from that use k- can come to acquire nominative use.

In some Formosan languages, such as Li's Tanan Rukai (1973) an element -a is added, said to be an accusative marker, forming object pronouns inia and idaa (p 83).

The forms from the table that we'll deal mostly with in this study are ana (< zero + a + na), una (< zero + u + na), and ina (< zero + i + na); all of which have the basic locative meaning "at that" or "there" in the hypothetical PAN and in some languages today. Those are the forms most often grammaticalized, though in a few languages forms beginning k- (allative "to") or ending -Da (remote) come to be grammaticalized in the same way.

In Li's table of feature classification of determiners in Tanan Rukai (1973, 87) one portion shows:

	-NM	+NM
-pers	ka	ka
+pers	ki	ku

The table is interesting in grouping vowels i and u together against a, and it does so without invoking a schwa vowel. Locatives are more likely to be non-persons, and patients are more likely to be persons. I suggest that this is the reason why, of the grammaticalized constructions that arose in AN from "at that", the ina / una passive voices (with non-low vowel) promoted accusative patients to subjects, and the ana passive/circumstantial voices (with the low vowel a) promoted mostly locatives. Another way to say this is that in the voice-changing affixes of verbs in such languages as Atayal, Tagalog, and Malagasy, the -pers PCMs with vowel a were used mainly for semantic locatives and circumstantials, and the +pers PCMs with vowels i and u came to be used for most person-like semantic patients. Tanan's distinction between ku marking Nominatives and ki marking non-Nominatives raises questions I don't wish to discuss in this paper. Let's only note that in Polynesian tongues ko < \*ku has nominative-like functions (marking presentative and predicate noun), while ki keeps its normal use as preposition "to".

In Li's -NM, -pers slot, sa substitutes for ka if it is +spec. That doesn't affect what we have said about the vowels.

Here are a few examples of the basic phrase, without grammaticalization, in attested languages.

The word ina occurs in some Polynesian languages without grammaticalization, with the meaning, “in or at that place” or (by extension) “in or at that time”. Carroll & Soulik’s Nukuoro dictionary, 1973, defines it as “to it”. It is interesting that “in or at that (place)” is exactly the most ancient PAN meaning that is reconstructable from Rukai and from Amis, as we shall see. For Hawaiian, Puku’i and Elbert (1957/1965,94) define ina as “Here, now, then” (i.e., in or at that place or time):

ina| iho ke ala  
Here is the road.

ina| aku ka hale.  
Yonder’s the house.

ina| aku a ho|’eha mai  
[He will] soon be back.

From *Tanan Rukai*, Li (*op. cit.*, 88)  
iakai biliŋ inia aŋatu  
exist high that tree  
[-NM] = Acc/Loc  
“(It) is high up **on that** tree.”

Malagasy shows a-ni contracted to an- with its simple meaning, in this sentence quoted (Pearson & Paul, 1996, 131) by Pearson (1996) from Ileana Paul (1996).

Misy lehilahy any an- trano  
(A-NI)  
exist man Deic Obl- house  
“There is a man **in the** house.”

When the function of the demonstrative is to serve as a Determiner, one of its effects may be to make the NP definite, specific, and referential.

In (perhaps) all languages, time locatives arise from place locatives. So it’s no surprise to find ana and ina forms referring to time. Rukai languages have (Zeitoun, pers. comm.) adverb ana “still”. Nuclear Polynesian languages have e VERB ana for continuative aspect. Both of those a-na could easily arise from “**at that** time”.



## Grammaticalization of the Phrase: Introduction

We'll begin by showing the syntactic similarity of the ana voices of Malagasy (a Malayo-Polynesian language) and Atayal (a North Formosan language). The Austronesian (henceforth AN) family has its westmost member, Malagasy (Madagascar), off the coast of Africa, and extends East 205 degrees, more than half-way around the world, to Rapanui (Easter Island), which is at the longitude of Colorado's west border and is ruled by Chile. The most popular classification makes Malayo-Polynesian (MP) one major division of AN, with the other three major branches crowded together on what may be the ancestral AN homeland, Taiwan. So, Malagasy and Atayal are not at all closely related to each other.

Comparing Malagasy with Atayal, I'll show that the main body of Austronesian languages had ancestrally a form ana, with variants una and ina, which was applicative (in the sense of taking an oblique adjunct and making it a core term), and which, as a suffix on verbs, marks a passive or ergative voice. Then, examining the Rukai group, we'll find the element functioning applicatively but not marking passive or ergative voice. The Rukai languages show the forms a-na and i-na in various uses, all easily interpretable as variations on a phrase meaning "at that". Some of the usages can be classed as applicative functions, but without any of the voice-changing effect that is seen in languages such as Mayrinax and Malagasy. So I'll argue that Rukai must have been the first group to separate from Mainline Austronesian, separating before the construction acquired a voice-changing function. Note that Starosta (without using the word Mainline) has argued that Rukai was the first to break off, though the reasons he advances differ from those presented here. Li groups Rukai among the Paiwanic (South Formosan) languages, while Tsuchida and Dyen put it in the Tsouic (Central Formosan) group.

The world is full of Nominative-Accusative languages (henceforth NomAc or "accusative"). But pure Absolutive-Ergative languages (AbErg or "ergative") are rare. So diachronically the oscillation is between pure NomAc and mixed languages, and it's convenient to classify the mixed ones as AbErg. Many of the conflicting studies have dealt with Tagalog, a tongue described by Schachter and Otnes (1972) in an extensive grammar that never uses the concept "subject" (Schachter's added a 1996 monograph titled *The Subject in Tagalog: Still None of the Above*). Not to get involved in the controversy, I'll refer in this paper to "passive/ergative" constructions.

## Syntax Common to Malagasy and Atayal

In a paper given at the annual meeting of the Austronesian Formal Linguistics Association in April, 1997, I showed evidence that the verb's suffixes for non-active voices [passive/ergative] in Malagasy are ANA and

INA, and that they are cognate with suffixes *-an* and *-un*, respectively, in Atayal. Malagasy belongs to the Malayo-Polynesian (MP) primary branch of the AN family. Atayal is a set of dialects or languages in the North Formosan group. In the basic syntax, Malagasy and Atayal are more like each other than either is like other languages of MP (e.g., Tagalog) or languages of other MP branches in Formosa. The fact that Malagasy and Atayal, not closely related, have this basic similarity suggests that the common fundamental points of their syntax are very ancient, and originated at a time close to Proto-AN. Li (1997, 1994) says correctly: “The Mayrinax dialect of Atayal is conservative not only in phonology and morphology, but also in syntax.”

**A-NA voice:** In this pattern, the noun in focus (Topic or Nominative or Absolutive Case) can be semantically either a patient or a “circumstantial” such as a locative.

*Malagasy, Keenan (1996)*

n	i	vidy	ana	[n-]Rabe	ilay	satroka	Raso
past	active	buy	ANA	Rabe	that	hat	Raso
		circum	Gen.agent	patient	Nom/Absol		beneficiary

“Raso was bought-for that hat by Rabe”

*Mayrinax Atayal, Huang (1994)*

qilap	-an	ni'	yaya'	ku'	paga'=su'
sleep	ANA	Gen.	Mother	Nom.Spec.	bed =2SBGen.
	agent		syntactic	Nom/Absol	
			semantic	locative	

“Mother slept on your bed.”

“Your bed was slept on by Mother.”

hihip	-an	ni'	yaya'	ku'	'ulaqi'
kiss	ANA	Gen.	mother	Nom.Spec.	child
		agent		syntactic	Nom/Absolutive
			semantic	patient	

“Mother (kissed, is kissing) the child.”

“The child (is, was) kissed by Mother.”

The identity of word-order (preferred or required) and the similarity of the syntactic cases are clear in these examples, and are not shared by many languages more closely related to Malagasy, or more closely related to

Atayal. In both Malagasy and Atayal, the ana form can be semantically patient or locative, while the ina or una form can only be patient. In both languages, the ina/una form typically is applied to a simple form of the verb, while the ana is more often applied to a verb that already has affixes (often the infix -in- in Atayal). There are differences, of course. In Malagasy, Keenan calls the ina voice (with patient subject) “passive” and calls the ana voice “circumstantial”, but where the ana verb’s syntactic subject is semantically patient, and other elements are consistent therewith, he calls it a variant of the “passive” and not “circumstantial”.

English has a circumstantial passive, too: “This house has been lived in. The matter has been looked into.” English forbids a direct object in that voice. (\*This house has been given a party in.) But Malagasy and Mayrinax regularly have (and Malagasy perhaps requires) a direct object in the circumstantial, as Keenan notes by saying that that voice increases the valence of the verb from two to three terms.

Is the basic form Vna or only Vn? The basic diachrony of all languages is phonological lenition to zero, balanced by insertion of lexical items. So that’s a presumption for the full form as ancestral. Keenan reconstructs the full forms in his analyses, and shows phonological deletion of the a when the agent occurs just right of it. The full form occurs in sentences in which the agent is deleted. The full form is used by native speakers as a citation form. A synchronic linguistic analysis treats the -a as epenthetic; and that may be the convenient synchronic analysis, but if so, it’s diachronically a reanalysis.

Here are some examples from Lisa Travis (pers. comm.) giving full lexical forms ending in the full ana:

ny vy no anaovana ny antsy  
the iron PRT circ.make-ANA the knife  
“It’s with iron that one makes knives.”

no sasana ny lamba.  
wash-ANA the clothes  
“The clothes were washed.” [clothes as patient: passive  
use of the ANA form]  
[There is no INA form of this verb.]

nasasana lamba ny savony  
wash-ANA clothes the soap  
“The soap was used to wash the clothes.”  
[The soap was washed-clothes with.]

Establishing that the ancestral form of the Philippine and Formosan “focus affixes” is **Vna** and not simply **Vn** is a key point in the diachronic reconstruction set forth here. It enables us to relate the forms in Malagasy and the other languages with those in Rukai.

### Applicative Function in Malagasy

Keenan (1996) states that the circumstantial (ANA) voice increases the valence of the verb from two to three by adding a core term. He doesn't use the word “applicative”, but valence increase seems to fit the definition of that word, as taking what would otherwise be an oblique adjunct and making it a core term. The sentences given here in which the (passive/ergative) agent is omitted show that even core terms can have zero realization in some languages. The passive transformation could add a term if it were derived from an intransitive verb, which would have to be an active (unergative) one.

### General Comments on the Rukai Languages

It is agreed that Rukai consists of three groups (perhaps languages). One consists of Manturan alone. Another is Tona and Maga, which are geographically close to Mantaauran; the third is Tanan and Budai/Labuan, which are geographically far from the other two groups and far from each other. It's on the basis of the reported mutual intelligibility that I define three languages. Here is how Zeitoun (1996) describes them:

“At first glance, Mantaauran should not be regarded as a peculiar language... In having oblique proper nouns suffixed with -ina, it behaves (to some extent) like two other Rukai dialects (Maga and Tona). In Maga, for instance, oblique proper nouns and kinship terms must be suffixed with -a(na).”

pa-ŋulu      kiki              bvaa ipul-a.  
Caus-drink 1S.FreeNom wine (name)-Obl  
“I let Ipulo drink wine.”

“It [Mantaauran] differs quite drastically from all these languages, however, in that it has developed a ‘verb-object’ agreement marking: third person (oblique) participants are cross-referenced on the verb by means of a pronominal clitic which is identical in form and use to the suffix found on the noun in [a Mantaauran sentence which has ina where the Maga example has a].” Her footnote adds:

“Though ina in Mantaauran and -a(na) /-ana in Maga/Tona do not show cognicity, I have shown in Zeitoun (1995,231) that -ina is a locative form, used to case-mark an oblique personal NP and contrasts with two

other (locative) forms -aə and -anə, which are suffixed to different roots: locative/place nouns (as in Havai-ane ‘the girl’s family’ < Havai ‘girl’). In Maga or Tona, the same form -a(na) /-anə is found in all locative (argument or non-argument) NPs.”

In pers. comm., Zeitoun says, “Syntactically, [Mantauran] differs from the other Rukai dialects and the Formosan languages as a whole: (1) it lacks the nominal case marking commonly found in most of these languages. (2) it has bound pronouns which tend to coalesce with the verb stem. (3) the third person pronoun has been reanalyzed as a ‘non-agent’ agreement marker....There are some major distinctions in the use of -a(na) in Maga, -anə in Tona, against that of -inə in Mantauran, summarized as follows. -a(na)/anə never attach to verbs. Hence the construction \*V-a(na) is not found. They always occur with nouns marked/understood as [+Oblique, +human], as in: ‘I gave Ipulo-Obl a book.’ In a conjoined NP (e.g., I gave a book to Mary and John), in Maga and Tona the suffix will occur on the two nouns. But Mantauran prefers to have the suffix moved onto the verb, and then establish an agreement in plurality with the conjoined NP it corefers to.”

All this is rich in reported data and in opinion. I gratefully accept the data. I respect the opinion and I agree with some of it. I draw my own conclusions about the direction of diachronic change.

Why is the demonstrative syllable (**na**) deletable after **a** and not after **i**? The optional deletion must have come at a fairly late time when the **ana** element was losing its analyzability as two morphemes. Possibly because the **i**, used with persons, needs to be specific, while the **a**, used more for locatives and other non-persons, seemed more of an oblique adjunct and less likely to be specific. There may have been some inherent incompatibility between the non-specificity and the demonstrative pronoun which inherently bears some specificity. A similarity is found in Polynesian, where the verbal suffix -ina can take the form -a but not -i. These matters are murky and these suggestions, very tentative.

### Basic Construction in Common to Malagasy and Rukai

We look for, and find, and shall show evidence for the shared presence in the two language groups, of a basic syntactic construction with these characteristics.

The basic element is a phrase **i na** and alternatively **a na**, preposition “at” + demonstrative pronoun “that” (with a very limited number of alternatives to each). They tend to become case-marked demonstratives and sometimes determiners. The reference can be to persons or non-persons; if a distinction is made between persons and non-persons, **i** is for persons and **a** is for non-

persons. The semantic roles of the entities can expand to a broad range. If a distinction is made between patients and things that could be oblique adjuncts (locatives and other circumstantials), **i** is for patients and **a** is for locatives and other circumstantials. Syntactically, the referenced NP can be anything but an agent-subject. The construction early came to be used applicatively, i.e., to take what would otherwise be an oblique adjunct and make it a core term or argument. In many languages, the source form continues to be found, perhaps in the same sentences, as a normal locative, often as **a-ni** or **a-na**.

One place the phrase occurs is just right of the verb (or the verb plus any suffixed nominative pronouns). Within the broad range of ancestrally allowable uses, individual languages may delete some uses and specialize in others.

### Overview of the Construction in Rukai languages

The one thing that the Rukai languages do not do, that is done by other Formosan and Phillipine languages such as Mayrinax, Malagasy, and Tagalog, is to coindex the INA-UNA-ANA demonstrative with the NP in the Topic-Nominative-Absolutive case. That's why the Rukai languages don't develop the passive or ergative voice cognate with what those other languages have. Tanan Rukai has a passive voice but it's marked very differently and is unrelated to such voices in the other languages. Zeitoun, Huang, and others have shown that particles in Formosan languages typically mark not voice, tense, or aspect alone, but some combination of voice and tense/aspect. One can see the general paths that re-analysis and change must have followed.

Li's **Tanan** (and presumably Budai) kept many of the ancient prepositions, but as casemarkers or casemarked articles (Determiners). Tanan (at least for **k-**) groups **u** and **i** together as +person, in opposition to **a** for -person. Therein it foreshadows the grouping of **u** as alternative to **i** in contrast with **a** in what were to become the voice-changing verbal suffixes.

Tanan's **\*ana** survives mainly as a suffix forming nouns from verbs. Some such nouns are clearly locative nouns denoting the place where the action takes place. That must have been how the noun-formation began. But this process has generalized into forming other verbal nouns, such as the names of the actions. It is also true that some of these verbal nouns, in NP NP clauses, tend to show some passive meaning, even though Tanan also has a true passive voice. So the verbal nouns may (or may not) have been ancestral to the **-an** suffix for circumstantial passives-ergatives as found in other Formosan and Philippine languages, as Li's mentor, Starosta, believes. A weakness in that theory is that it seems unrelated to the truer passive-

ergative suffix **-in**, **-un** that cries out for an explanation related to that of **-an**.

Tanan's main and prominent survivor of the key phrase, of course, is the demonstrative **ina**. Zeitoun has correctly said that Mantauran's **ina** comes from an ancestral locative construction. It seems clear that the same is true of Tanan's **ina**. From its ancient locative case relation it has expanded to a general non-Nominative case, sometimes called Accusative. Examples show it for motion to or from and even for agent "by" (but never the subject-topic-Nominative). It's used both as a complete Determiner Phrase and as a Determiner introducing a NP.

An interesting point in Tanan is that INA occurs in several alternate forms, used interchangeably: **inia**, **ina**, and **na** (synonyms for middle distance), and **iDaa** (for farthest distance). Reference to the table of distance-markers in AN languages shows that what has happened in Tanan is the merger of **ni** and **na** distances. (The same merger occurred in Amis, a member of the South Formosan group.) Tanan's proximal distance is **ki-kay** ("this"), which we need not discuss here. The simple form **na** with zero casemarker is intriguing; one could speculate about the possibility of a non-locative origin in an ancestral zero-marked accusative, but I know of no relevant evidence.

Tanan's INA forms are classified as independent object third person pronouns (Li, 1973, 83, given below), while Zeitoun calls Mantauran's Bound Object when they follow the verb. This is merely a well-recognized difference in diachronic stage. Tanan's independent pronouns are an earlier stage which, in the post-verbal position, easily are weakened to Bound Objects through the usual processes of lenition and grammaticalization.

The most general rule for occurrence of the INA forms is the same in Tanan as in Mantauran: they occur in non-Nominative case form (syntactic case).

**Mantauran** (as visible in Zeitoun's publications) has, like Tanan, discarded most uses of **ana** in favor of **ina**. But Mantauran has sharply reduced the use of the INA element from its maximum use in Tanan (and presumably proto-Rukai) by imposing a constraint against its use for non-persons. The use is required when the referenced NP is a person. It's optional when the logically apparent object is a nonperson but its expressed possessor is a person. Otherwise its use has been deleted and not replaced with **ana**. I see no need to declare that either the suffix to verbs or the suffix to nouns is the original position from which the other has been derived. Examples from other languages, AN and non-AN, show that location of the former PP demonstrative right of the verb is common and can be the point of origin of such PP both diachronically and synchronically. (That's also true of direct objects, which the former PP become.) Position of the demonstrative after a

noun, though less common, is also found in a variety of languages, with a somewhat different denotation. There's no reason why this demonstrative PP can't have been present in both positions from time immemorial. What is specific to Mantauran is the fuzzing of the two uses to the point where the speaker often has a choice between them. As for Zeitoun's term "agreement-marker" for the INA at the verb's right, this is a term that becomes applicable when the former independent pronoun (as seen in Tanan) becomes an enclitic and then a suffix on the verb. I don't view this as a drastic change: only as a common occurrence of the normal processes of lenition and grammaticalization found in languages throughout the world.

Zeitoun's use of the term "agreement marker" differs in two ways from the conventional use. (1) In the usual case, ("The man goes", "Homo it", versus "The men go", "Homines eunt") the number marker for nouns has nothing phonologically in common with the number marker for verbs, while in Zeitoun's examples the two are identical. (2) In the normal use of the term, marking for number is required on both words, while in Zeitoun's Mantauran examples, it is only in exceptional cases that both words receive the marking.

For **Maga** and **Tona** I have seen no materials, and must rely on Zeitoun's pers. comm. Both dialects of this language seem to have done the opposite of Mantauran and Tanan-Buai. Maga-Tona has abolished **ina** and kept **ana**. Even for **ana** its use as an argument of the verb has disappeared. The use remaining is that after nouns. In that position the former demonstrative phrase has become a clitic, and perhaps now a suffix, attached to the noun to which it refers.

### Specifics of "ina, ana" in Tanan Rukai

Tanan Tukai's i is cognate with the i identified as a locative PCM in Mayrinax Atayal by Li (1994). Some other AN languages have it as a or u. But in Tanan the i as a casemarker has fused with the demonstratives into a single word ina, inia, iDaa, constituting the accusative (non-Nominative) demonstratives varying in distance; and the prepositional syntax is no more.

Li's (1973) work on the Tanan Rukai language was based in his Ph.D. dissertation completed that year under Starosta at Hawaii. Its table of determiners (p 87) sheds light on the subject of the present paper. Simplified by omission of the non-demonstrative half (casemarked articles), it's as follows, and the table of personal pronouns follows it. It's easy to see that the demonstratives and the third-person pronouns have a common ancestry.



## DEMONSTRATIVES

	+nigh -far	-nigh -far	-nigh +far
-NM	ki-kay	i-ni-a, i-na, na	i-Da-a
+NM	kay(vay)	ku-a-ni	ku-a-Da

## PERSONAL PRONOUNS

	Independent		Dependent	
	Subject	Object	Subject Mkr	Possessive
1sg	ku-n-aku	n-aku-a	-aku, -n-aku	-li
2sg	ku-su	mu-su-s	-su	-su
3sg def	ku-a-ni	i-ni-a	zero	-i-ni
indef	ku-a-Da	i-Da-a	zero	-i-Da
1pl exc	ku-nai	nai-a	-nai	-nai
inc	ku-ta	mi-ta-a	-ta	-ta
2 pl	ku-numi	numi-a	-numi	-numi
3 pl def	ku-l-i ni	l-i-ni-a	zero	-l-i-ni
indef	ku-l-i-Da	l-i-Da-a	zero	-l-i-Da

In this language, an interesting mixture of the ancient and the altered, are many points that we can't discuss. The contrasting use of "this/that" demonstratives to show definite-indefinite is not found in any other AN language that I know. Because the PAN casemarkers for Topic (quasi-Nominative) and Locative became the same in Proto-Rukai (by loss of glottal), and because the old Locative marker became the casemarker for the whole non-Nominative case form, a new Nominative casemarker was needed, and Rukai is not the only branch to make do with k.

We could call the non-Nominative set Loc/Acc, as each of these Determiners, written as a single word, consists obviously of a preposition/casemarker (PCM) (ancestrally locative i or allative ki) followed by a pronoun which was ancestrally a demonstrative of one distance or another and is also capable of being used as a third person singular pronoun. The remote form is i-Da-a, the intermediate i-ni-a, and the proximate ki-kay. Both ina and na, used by Starosta (1974), are given as

alternate forms of i-ni-a, though they look like \*i-na-a (cf. ina which we'll see in Mantauran). Both \*ni and \*na, as well as \*Da (Nuclear Polynesian laa), are members of the set of demonstratives for distances (degrees of "this" and "that") which are found throughout the AN family. Tanan non-Nominative demonstratives begin with i- no matter whether patient or locative, and no matter whether personal or impersonal. But two other Rukai languages (Maga and Tona) have a- forms. As we've noted, it's common to see accusative casemarkers derived from locative prepositions. In Hawaiian-Tahitian-Maori, where this has happened, the accusative and the oblique prepositional constructions can still be distinguished syntactically, though they are marked the same. But in Tanan Rukai, the PAN prepositions have deteriorated, surviving only as casemarkers, determiners, and prefixes on demonstratives. The four "prepositions" listed by Li (111), "with", "at", "from", "to", are from verbs "use", "stay", "come from", and "arrive". What has happened in Tanan is that all cases but Nominative have combined into a single oblique case form which Li calls Accusative. Ancestrally it was (at least for the pronouns) locative (AN preposition i) but it's come to be used for the accusative and all other cases but the subject. Other PAN prepositions survive as casemarked Determiners (ka, ki, ku, sa).

So we see here in Rukai all the raw materials for what developed into the voice-marking suffixes for verbs in Malagasy and in Atayal. But they take no such function in any Rukai language.

Here are some of Li's sentences using the INA equivalents:

wacɪɪlaku      iDaa  
looked-I   [at]   that  
              "I looked at that." (p 89)

iakai      ina      daanli      tubi-tubi kay Lulay  
stayed [at] that   house-my crying   this child  
              "This child stayed in my house crying." (p 159)

wacɪɪl-su      iDaa      aʔuy  
saw-you      that      fire  
              "You saw that fire." (p 161)

wa-laub      ini-a      daan      ku-ani      umas  
burned      that      house that      man  
              "That man burned that house." (p 190)

Starosta (1974), based on interviews in 1972, described the causative constructions in Rukai (Tanan) and five other Formosan languages. He did the interviews in the Mandarin Chinese language, in which he has some proficiency. Li assisted him but was not listed as a co-author.

Here are three sentences from Starosta (1974, 321-322) with my analysis added in square brackets. In Starosta's line-pairs of grammatical gloss, the first line is his syntactic "Case Form" and the other his semantic "Case Relation". In these sentences, all causative, I'll use "pivot" for the one who is the patient of the causation and the doer of the resultant act.

wa-ranaw	iDa	adadam	ko	damay	(ani?alay	toalay	
make	INA	bird	CM	Damay	fly	from	
(nonpassive)	(AC)			(NM)			(L)(src)
(PAT)	(AGT)			(LOC)			
[		pivot	Nom/Absol	]			

obola	akila	tarumak
mountain	to	Tanan
	(L)	(gol) (LOC)

"Damay made the bird fly from the mountain to Tanan."

mua	si	la?	ŋ	si	kauLiva	iDa	Lulay	mua	vilvil	nia	aŋato	miya
go	find	tell		INA	child	go	shake	tree				
+imp	+imp	+imp			+AC	+imp	-fin	+AC				
					+DAT			+PAT				
					[pivot]							

"Go and find the child and tell him to go shake the tree."

ko	laciŋ	boada	amiya	na	maroDaŋ	sa	Lima	ka	?ayso	ina	Lulay	CM
Lacing	give	cause	INA	old	man	five	dollars	INA	child			
+NM	+imper	+fin		+AC			+AC		+AC			
+AGT				+BEN			+OBJ		+DAT			
				[pivot]								

"Lacing had the old man give the child five dollars."

Starosta's article, from which the examples are taken, dealt exclusively with causative constructions in six Formosan languages. In these sentences, the pivot is an INA-marked noun which may be human or not. It's the patient of the causation and the agent of the resultant doing.

The dative beneficiary fits the pattern of Keenan's "circumstantial" in which the valence of the verb is increased from two (subject and direct object) to three. In the sentences, in which the INA noun is the pivot, which clause is it in? Probably the upper (left) clause, if the construction is something on the order of an equi deletion; the small clause to the right is marked non-finite in one and seems to be non-finite in the other, too.

Notice that in Tanan the INA has only at most an applicative function, adding a core noun slot, and does not signal a change in voice. A reasonable

hypothesis is that Tanan Rukai broke off early from the common main line of AN which later came to use the INA and ANA construction for voice changing. Tanan has a passive, which does not resemble the non-active voices of other AN languages.

Notice that in all the Tanan sentences, the noun that is governed by INA, and is the patient of the upper clause, comes immediately to the right of INA. That is where we expect it to be, if, as we surmise, the earliest meaning of I-NA (or A-NA) is “at that”. If *i* or *a* is a preposition or casemarker, and *na* is a demonstrative which can serve as a Determiner, we expect the relevant noun to follow immediately to the right. In the three Tanan sentences quoted here, the order is: VERB INA NOUN. One has a preposed topic.

The ina in Tanan Rukai (and presumably the ana in Maga and Tona Rukai) thus stay fairly close to their PAN origin as locative prepositional phrases, while broadening their semantic scope. It seems that they have an applicative role, adding valence to the verb, but the extent to which they do so is uncertain, as the tests for argumenthood have not been done. The quotation from Zeitoun (above) implies that some of the INA nouns are arguments, but others remain oblique adjuncts. The sentences in which Starosta shows two syntactic accusatives, one of them semantically dative, can be compared with double-object verbs in English, which may or may not be regarded as the result of an applicative process.

### Specifics of “ina, ana” in Mantauran Rukai

From among various Mantauran sentences given in Zeitoun (1966), the following ones follow the Tanan word order and represent what I believe was the word order most used in the ancestral language. I follow Zeitoun in writing the INA elements as suffixes on the verbs, though they must have been phrases of separate words, “at the”, in the ancestral period in which they began to grammaticalize. BO and BG are bound object and bound genitive. In Mantauran the “object” with which the INA phrase agrees is not necessarily the accusative patient. So it should not surprise us that Zeitoun labels it oblique, though she also calls it verb-object agreement (as Li does in Tanan). The being to whom the “ina” refers must be human and may be only the possessor of a NP.

*Zeitoun 1996, 15, 27*

maðalam	-il-inə	[taotao la anao] ðipolo
	INA	
love	-3P.BO	Taotao and Anao Thipolo

I propose the ancestral syntax:  
 VERB PREP DEM/DETNP (OBJECT) NOM/ABS/TOPIC  
 “Thipolo loves Taotao and Anao.”

1996, 18, 35b

olalamə	-ŋa	-inə		lalakə	-ni		
run		-Perf	-3S.BO	child		-3S.BG	
		INA		NOM/ABS/TOP		his	VERB
PREP	DEM/DET		NP(OBJECT)				

“His children ran away on him.”

These sentences preserve the normal order, in which the NP is preceded by a preposition and a demonstrative or determiner. In the label PREP, of course, I’m calling it what it was ancestrally (and still is in some languages, as in Malcolm Ross’s example cited). In this regard Mantauran is close to PAN, in which the i/a/u syllable was (as in Atayal) a preposition with mostly locative function; but when we see the I-NA element attached to the verb as a suffix and labeled Bound Object, we also see a transition away from locative and toward accusative. Zeitoun’s many examples show the I-NA-marked NP with a wide variety of syntactic and semantic roles, excluding only agent and the subject/topic in the unmarked case. The Mantauran system may be called applicative, i.e., taking elements that would normally be oblique adjuncts and promoting them to terms by giving them the special marking.

Mayrinax and Malagasy have changed the word order slightly. With voice change, the noun referenced with INA (patient or locative) is the syntactic subject and is no longer co-indexed with the former pronoun NA nor is it the object of the former preposition I / A; so it no longer forms a phrase with INA and no longer needs to come just right of INA. That makes it possible to insert the passive-ergative agent immediately right of the verb in those languages. A slight change in word-order reflects a profound change in the syntax.

One position the INA suffix can take is a less common one: suffixed to its NP. That may be a late development, motivated and reinforced by its ability in Mantauran to avoid a certain ambiguity that Zeitoun discusses. [Mantauran doesn’t casemark its nouns; and so, if two human nouns occur in the sentence, it’s not clear which one is the non-Nominative, i.e., coindexed with the INA demonstrative suffixed to the verb.] Li (1994, 20) notes in a non-Rukai language, Isbukun Bunun, nominative and oblique casemarkers usually suffixed to their nouns, and suggests that they may have developed from demonstratives. In English we occasionally put a locative demonstrative (“here”, “there”) after a noun. A common expression

in Hawaiian is “Hawaii nei” (“this Hawaii” or “Hawaii here”), which is an affectionate expression.

Rukai languages show some elements of the affix system found in other Formosan and Western MP languages. Verbs in active voice in Mantauran and other Rukai languages have a prefix wa-. Zeitoun (pers. comm.) suggests, and I agree, that this may be homologous with the m- prefixes of the active voice in those other languages. But the most striking thing in the syntax of languages such as Malagasy and Atayal is the use of ANA and INA to show passive or ergative voices. The essential diachronic development was that the ANA and INA phrases “at that” came to coindex the oblique “that” element with the unmarked Topic. And that’s the one thing that cannot be done with those elements in Rukai. There is nothing in Rukai that looks as though Rukai had had that key element and lost it. What seems the simpler explanation is that Rukai’s ancestor separated from Mainline Austronesian before the latter developed the applicative element into a voice-changing one.

Even in sentences in Malagasy and in Mantauran that have only one non-Nominative term, an applicative process may have increased the valence of the verb; in such cases, from one (intransitive with an oblique adjunct) to two.

### **Proposed Sequence: from only Applicative in Rukai to Applicative plus Voice-Changing in Mainline**

From evidence that I’ll discuss elsewhere, I suggest that the unmarked Topic-Nominative-Absolutive case began in earliest PAN as a left, sentence-initial topic that was not syntactically a part of the clause or sentence. Overtly it was marked by ?a, ?u, or ?i, which differed from the locative marker in having initial glottal. This syntax was ancestral to all AN languages. By late PAN, still ancestral to all AN, it was optionally moved to the right edge of the sentence, where it became part of the clause and began to be syntactically related to the other components.

For reasons that have not been identified, in sentences having the applicative INA, ANA, UNA, the newly incorporated Topic became coindexed with the element that had been promoted from oblique adjunct to core term by the applicative process. That made the unmarked Topic, Nominative, Absolutive from the semantic patient, locative, or other role that would normally be an adjunct. In such constructions the Genitive-Ablative preposition nV, which already meant “from” and “of”, became the preposition for the passive agent or the casemarker for the ergative agent. About this time the INA-ANA-UNA forms began optionally to elide the final -a in some phonetic contexts. The fact that Malagasy keeps the final

vowel shows that its disappearance was not complete by the time that Proto-MP separated from the other main branches of Mainline AN.

### **An Undecided Question: Which Paradigm for the a / u / i Alternations?**

There are two possible paradigms that might explain the variation among the three vowels in the i-na, u-na, i-na elements. PAN is believed to have had four vowels: a, i, u, and e. Ross's (1992) table shows how all reconstructed PAN phonemes are reflected in the reconstructed Proto-languages of nine Formosan groups and Proto MP. Lawrence Reid (pers. comm.) states that the i-u variation in the voice-changing suffixes (in, un) corresponds with the accepted reflexes of the schwa in those languages. It has not been proposed that there was a preposition of locative or other use consisting phonologically of the schwa.

The other paradigm of vowel variation in particles is based on findings in Chamorro and Paiwan. Topping's Chamorro-English Dictionary (1975) includes: "nu Particle, used to mark non-focus elements of a construction. Usually occurs as ni which is a contracted form of nu plus i." Paiwan is a language in which \*t has become tj and \*s has become t. Ferrell's (1985) Paiwan Dictionary shows "tu construction marker preceding noun phrase which is neither in Focus, nor Agent"; and also shows it as tua, page 13. He also shows ni, nu, and nu-a, all glossed "belonging to, of". The ni entry says it's used "with personal name; cf. ti [which is the honorific < \*si]". All this seems to show that for several different monomoraic prepositions or casemarkers, PAN had vowel u as the unmarked state, with ua > a for nonspecific common nouns and ui > i for persons. Probably a and i were articles. So an attractive hypothesis is that the three-vowel alternation in the INA/UNA/ANA elements belongs to this set.

It remains to be seen whether a reasonable diachronic hypothesis could relate both these paradigms to the voice-changing elements discussed here. We have seen in Li's Tanan Rukai (with another initial consonant, k) that u and i are grouped together as +person, against a, -person. It is not inconceivable that languages choosing between the two high vowels for schwa's outcome might be constrained to make the same choice for the INA/UNA element. It is possible that a hypothesis of a reanalysis at some point in time could bring the paradigms together.

### **Summary**

1. I have reviewed a paradigm showing that the most ancient monosyllabic prepositions and casemarkers in Austronesian, still found as such in many languages today, consist of a consonant with the prepositional or casemarking content (such as locative or allative) plus a vowel that shows +person, -specific, and the like.

2. I have reviewed a paradigm identifying AN demonstratives denoting various distances from “this here” to “that there far away”.
3. I have shown that phrases of PCM + demonstrative still occur throughout the AN world with meanings such as “at that” = “there”.
4. I have shown evidence that the voice-changing affixes on verbs in Malagasy have the basic forms ana and una; and that their function can be thought of as basically applicative. As the AN/UN/IN verbal focus affixes in Philippine and Formosan languages are clearly cognate with those of Malagasy, they, too, must have ancestrally been ANA-UNA-INA.
5. I have shown that in Mayrinax Atayal, a very conservative Formosan language, the syntax of non-active voices shows basic likeness to that in Malagasy. As the languages are not in the same basic branch of AN, the likeness must preserve very ancient syntax and morphology, close to Proto-AN. In particular, Mayrinax -an has syntax strikingly like Malagasy -ana, and Mayrinax -un like Malagasy -ina. One similarity is that the ana forms make the NP of unmarked case either a patient or a circumstantial (locative, etc.), while the forms with the other vowel make the NP of unmarked case a patient. The suffixes on the verb may be considered applicative in the sense of adding to the valence of the verb, taking an additional argument or core term that might else be an oblique adjunct.
6. In this paper we have examined the Rukai languages in general. We have found forms, both ina and ana, which resemble the corresponding Malagasy elements both phonologically and (partially) in usage and function. The differences in function are such that it is reasonable to believe that the Rukai languages represent an earlier usage that could develop into the usage and function found for these forms in Malagasy. The hypothesis is suggested that these forms are cognate. Tanan and Mantauran keep only the ina forms as casemarked pronouns, and use the ana forms mostly for formation of nouns (at first location nouns) from verbs.
7. In this paper we have examined the Tanan Rukai language as reported by Li. It keeps the forms from ina (also forms ini and iDa) as casemarked non-nominative pronouns. Their usage may be considered applicative.
8. In this paper we have examined the Mantauran Rukai language as reported by Zeitoun. As compared with Tanan, Mantauran deletes the ini/ina (and iDa) element except in relation to persons. Further, Mantauran has lenited and grammaticalized the element from an independent pronoun to a clitic or suffix on the verb. The directions of all these differences are consistent only with the hypothesis that Tanan represents the older situation and Mantauran the innovative.



9. The moderate applicative uses of INA in the Rukai languages do not change the voice. From this and other evidence, the conclusion is suggested that Rukai did not undergo some basic syntactic and morphological changes that took place in the ancestors of all other AN languages.

## Needs

What is most needed now for Rukai is a series of systematic studies on the languages in which the **ina** element is a possible argument of the verb: namely, Tanan, Budai-Labauan, and Mantauran. Such studies need to use appropriate synchronic tests to determine which **ina** phrases, if any, are core terms (arguments) of the verb, and which ones, if any, are mere oblique adjuncts. The results can clarify to what extent the Rukai languages used the element applicatively, as it is used, for example, in Malagasy. The result could help to clarify the diachronic evolution of syntax in the early phases of the AN language family.

## References

- Carroll, Vern & Topias Soulik. 1973. *Nukuoro Lexicon*. Honolulu, University Press of Hawaii.
- Ferrell, Raleigh. 1981. *Paiwan Dictionary*. Pacific Linguistics C-73.
- Huang, Lillian M. 1994. *A Study of Mayrinax Syntax*. Taipei: Crane.
- Keenan, Edward L. 1996. Morphology is structure: a Malagasy test case. In Pearson & Paul, 1996, 92-112.
- Li, Paul Jen-kuei. 1973. *Rukai Structure*. Special Publication 64. Taipei: Institute of History and Philology, Academia Sinica.
- \_\_\_\_\_. 1994. Paper read at the 7th International Conference on Oceanic Linguistics, Leiden.
- Pearson, Matthew and Ileana Paul (eds.). 1996. *The Structure of Malagasy*. UCLA Occasional Papers in Linguistics, Number 17.
- Pearson, Matthew. 1996. Domain phrases and topic arguments in Malagasy existentials. In: Pearson & Paul, 1996.
- Paul, Ileana. 1996. Specificity, existentials, and word order in Malagasy. Ms., Linguistics Dept., McGill University.
- Puku'i, Mary Kawena and Samuel H. Elbert. 1957. *Hawaiian-English Dictionary*. University of Hawaii Press.

Ross, Malcolm. 1992. The sound of Proto-Austronesian: an outsider's view of the Formosan evidence. *Oceanic Linguistics* 31:23-64.

\_\_\_\_\_. 1997. Early Oceanic adjectives. Paper read at the Third International Conference on Oceanic Languages, University of Waikato, New Zealand.

Schachter, Paul 1996. The Subject in Tagalog: Still None of the Above. *UCLA Occasional Papers in Linguistics* No. 15. Los Angeles: Department of Linguistics UCLA.

Schachter, Paul & Fe Otanes. 1972. *Tagalog Reference Grammar*. Berkeley: Univ. of California Press.

Starosta, Stanley. 1974. Causative verbs in Formosan languages. *Oceanic Linguistics* 13:279-369.

\_\_\_\_\_. 1994. Paper read at the 7th International Conference on Oceanic Linguistics, Leiden.

Topping, Donald, M. Pedro, M. Ogo, & B. C. Dungca. 1975. *Chamorro-English Dictionary*. Honolulu, University of Hawaii Press.