AUSTRONESIAN ERGATIVITY
TRACED THROUGH TWO CYCLES

Joseph C. Finney
Monterey, CA
<jcfinney@redshift.com>

0 Introduction
This is a study in the diachronic syntax of the Austronesian (AN) languages. It presents a
paradigm of the cycle of ergativity, as seen in that family. We shall be concerned with the
syntactic casemarking and the semantic Case relations of the core terms of the verb, in all its
voices. Like most in the field, we’ll use the accepted definitions from Dixon (1994) that a
language is “accusative” (NomAc) if S, the single essential term of intransitive sentences, is
marked the same as A, the agent (actor, doer) of the transitive sentences, while O, the object-
patient-undergoer bears a different marking, defined as Accusative; and a language is
“ergative” (AbErg) if S is marked the same as O, while A bears a different marking, defined
as Ergative. Dixon (1994, 183) is to be commended for using the concept “moving a
language around the cycle of change”, and that concept is basic to our presentation here.
Dixon’s book defined concepts and basic orientation assumed in the present work.

We’ll study how languages move around in cycles. Examine our table (diagram, a
short unnumbered page) first and use it as orientation in reading the paper.

As we are specially interested in Polynesian and the complex development within
it, let’s summarize the successive phases of syntax that come down to it from our
theoretically reconstructed PAN.

There are two syntactic states (NomAc and AbErg) in each cycle. Each state lasts
for a stage.

There are two phases (A and B) in each syntactic stage (NomAc or AbErg).

That makes four phases in each full cycle (1A, simple NomAc; 1B, NomAc with
passive; 2A, simple AbErg; 2B, AbErg with antipassive).

1. The earliest AN we can reconstruct began as NomAc in PAN (easily reconstructable
from Basay and other languages considered archaic). Starosta (who died recently)
thought Rukai preserves PAN’s NomAc (though it might possibly have gone from
PAN NomAc through AbErg and back to NomAc).

2. AbErg in many Formosan and Philippine tongues, and through one branch to
Paoe’s new NomAc. In parallel at greater speed AbErg lasts to about Proto
Eastern Oceanic.

---

1 The following abbreviations will be used for language groups: AN Austronesian, MP Malayo-
Polynesian, CEMP Central-Eastern Malayo-Polynesian, CP Central Pacific, PN Polynesian,
EPN East Polynesian. All may be preceded by P for Proto-. In hypothetical words, V and C
mean any vowel and any consonant. PCM preposition-or-casemarker. TAM tense-Aspect
marker, RED reduplication, DP determiner (noun) phrase, Lin linker. In Tagalog I’ll spell the
ergative casemarker nang as such and not “ng” as is customary.
3. NomAc from roughly Proto Eastern Oceanic through Central Pacific, Fijian, and the earliest reconstructable phase of Proto PN.

4. AbErg by the final phase of Proto PN (through a complex series of changes as proposed by Finney). Staying AbErg in Tongan, Samoan, and largely in the Outliers.

5. Rapanui, Easter Island (parallel to the mainline East PN) staying AbErg but becoming NomAc now, ending cycle 2 and beginning cycle 3 now.

6. NomAc in Proto East PN and persisting as in Hawaiian, Tahitian, Maori. These are 1B of cycle 3.

7. Maori is still within NomAc 1B cycle 3, but its predominant use of passive voice (more often than active) is a sign that it is moving toward AbErg syntax 2A. The passive voice, growing in the present phase, will become the Ergative in the future phase 2A, if the Maori language survives so long.

AN is a good family in which to analyze the diachronic changes of ergativity because for at least one living group, East Polynesian, reconstructions show two complete ancestral cycles (four changes of ergativity state, eight phase-changes), from PAE down: NomAc to AbErg to NomAc to AbErg to NomAc. And because one EPN language is irrevocably headed for AbErg again in a third cycle.

In applying the Dixon test to Stage 2 Ergative AN languages, which construction (voice) shall we use as the transitive? The Agent-Focus, or one of the Non-Agent-Focuses (NAF)? As we shall show in examining Mayrinax, the answer is clearly the Patient Focus. That’s the one that is the canonical Non-Agent Focus (and is in English, too). In any AbErg language, the Patient Absolutive term is the only one that will survive and have a descendant as a core term in the NomAcc 1A of the new cycle (where it will be the Accusative Direct Object). Analogously, in English the canonical passive is the Patient Passive “An award was given (to the teacher) (by the Board)” and not, for example, the Dative Passive “The teacher was given an award (by the Board).” A complete cycle, shown in the Table, has four stages. They are: 1A, simple NomAc; 1B, NomAc with a passive voice; 2A, simple AbErg; 2B, AbErg with an antipassive voice [whence we keep repeating]. Could we begin just as well with simple AbErg and end four stages down? I think not. The path 1B to 2A is simple, but 2B to 1A is more complex and variable and is better put at the end.

A complete cycle, shown in the Table, has four stages: 1A, simple NomAc; 1B, NomAc with a passive voice; 2A, simple AbErg; 2B, AbErg with an antipassive voice. Could we begin just as well with simple AbErg and end four stages down: I think not. The path 1B to 2A is simple, but 2B to 1A is more complex and variable, and is better put at the end.

A good recent discussion on the general concepts is Christopher Manning’s Ph.D. dissertation, published by Stanford (1996). His research was on Inuit, a West Greenland Eskimo language, but he discusses broader issues well. He considers argument structure (something closer to semantics than to pure syntax) the major determinant of syntax in languages throughout the world. He observes that two paths have been proposed for sliding into ergativity: one, from passive voices; the other, from nominalization (which is not discussed in this present paper).
Some events we’ll see are (1) evidence about PAN from Basay, Atayalic and other archaic tongues; (2) dramatic return to NomAc through mixed ergativity ending Cycle 1, in van den Berg’s series of villages (and confirmed in Central Pacific); (3) complex changes out of Fiji-like NomAc in the PPN period; (4) many types of syntactic change in Polynesian languages; (5) return to NomAc through mixed ergativity in Rapanui (right now!); and (6) one NomAc East PN language (Maori) taking an irrevocable step toward AbErg syntax early in a third cycle.

The term “Focus” refers to the DetP in Absolutive case, in phase 2A or 2B, which, like all Absolutive DetP, descends from ancestral Nominative case subject of a passive in stage 1B. There are three or more because stage 1B normally has three or more passive constructions. It also refers to the “Focus affixes” on the verb which identify those roles of each Abs term as an Agent, a Patient, or a Locative or some other role. Each corresponds to a type of passive voice. In stage 1B, Agent Focus (AF) Abs DetP had been Nominative subjects of active voice sentences, while Non-Agent Focus (NAF) Abs DetP had been Nominative subjects of passive sentences in the recent stage 1B.

The same basic principles should underlie events in all cycles, but very early in the second detectable cycle of An languages, the Focus affixes in MP had severely eroded, and the term Focus is not used. One could label Tongan’s main syntactic (AbErg) type as NAF, or label Fijian’s or East Polynesian’s NomAc syntax as AF, but nobody does.

Translations of Ergative sentences (stage 2A) into English. If the Ergative agent is translated as the English subject, it shows the new NomAcc syntax that will appear two stages later, 1A, when the old Ergative case is reanalyzed as the new Nominative case. But if the Absolutive patient is translated as the subject, it shows the old passive voice of the NomAc sentence that appeared earlier in stage 1B. English is a language in stage 1B.

1 General characteristics of AN languages

Major words from most ancient forward were CVCCVC > CVCVC > CVCV (in most MP languages and many others now).

The greatly prevailing word order is SVO, and therefore adpositions are prepositions, not postpositions.

All AN languages have Verb Object order, and so they have prepositions, left of the DetP, not postpositions.

In many AN languages, adjectives are not sharply distinguished from verbs. As an example, in Nanumanga, where I did field work, “te tangata fano” translates “the running man”, while “te tangata e fano” translates “the man runs or is running”, in which “e” is the most neutral tense-aspect marker. In many AN languages “ne” marks past or present perfect tense, and in some “kV” marks future. Many languages in cycle 1 have subject markers and object markers on the verb. They are obviously reduced forms of personal pronouns, reflecting older word order and syntax. If third person we commonly see both the noun and the affix.

The word “focus” means “voice”. The focus markers on the verb show its voice. The suffix -un shows that the subject is a patient. The suffix -an showed originally that the subject was a location, but in some languages has come to be used also for a patient subject. Both those markers may be old phrases meaning “at it”. Agent subjects in the archaic languages may be shown by prefixes and infixes at the verb.
PAn demonstratives distinguish specific/nonspecific (which Huang calls Rf-Nrf), not definite/indefinite. When English “a” means “a certain” it is specific in AN. In the disagreement between Starosta and Blust, I may occasionally refer to (Austronesian minus Rukai) as Mainline AN.

Some events we’ll cover are (1) evidence about PAn from Basay and other archaic tongues; (2) dramatic return to NomAc through mixed ergativity ending Cycle 1, in van den Berg’s series of villages; (3) complex changes out of Fiji-like NomAc in the PPN period; (4) many types of syntactic change in Polynesian languages; (5) return to NomAc through mixed ergativity in Rapanui (right now!); and (6) one East PN language’s taking an irrevocable step toward AbErg in a third cycle.

We shall see evidence to support a prehistory (1) that the earliest reconstructable PAn had NomAcc syntax; (2) that before moving toward ergativity PAn had to develop passive voice(s) (1B) and developed at least three of them, much like the ones that English has; (3) that through small, identifiable steps, the passive voices became ergative voices in 2A; and (4) by the time the Malayo Polynesian speakers left Formosa, nearly all the AN tongues had changed at least to 1B and some to 2A; and (5) the AbErg, stage 2A was ancestral to all the AN languages but the most highly archaic ones.

2 Basic principles of linguistic change

Lenition (progressive weakening of sounds toward zero) is an inconspicuous process, but it’s the great engine of linguistic change. As phonological clues disappear, redundancy falls, and scarcely enough stimuli remain for interpretation, words are inserted. Unstressed words (especially pronouns, and prepositions or casemarkers) become reduced to clitics, which then are reduced to affixes and then to zero.

Space concepts (and space prepositions) are the source of abstract concepts. The prepositions “from [a physical source]” come to mark agents (sources of action): at first in passives (1B) and then in the ergatives to which the 1B passives are reanalyzed in the formation of stage 2A.

As agent subjects are in some sense the source of the action, and accusative object patients are in some sense the destination or goal or theme of the action, it’s not surprising that ancestral “from, by” prepositions come to mark agents, and ancestral “at” or “to” prepositions come to mark patients, receivers of the action.

Though we tend to think categorically, in major changes of status many syntactic and phonological changes take place in small steps, which are inconspicuous. A change from AbErg to NomAc syntax, or from NomAc to AbErg, takes place as a series of smaller steps.

This standard way to develop an Ergative syntax is to reanalyze a passive voice (stage 1B) to Ergative (thereby entering stage 2A).

When a syntactic innovation occurs, it may begin as a synchronic transformation, secondarily generated. When the innovation is well established, it may become the one that is generated directly, and then the older syntax, so far as it remains, comes to be generated from the newer one. We shall see in Mayrinax, 2A, that the newer NAF patient-focus construction comes to be the standard, and the older (disappearing) AF sentences are generated secondarily and are not the ones to be used in the Dixon test for ergativity.

Nominalization had been proposed by Starosta, Pawley and Reid (1981) as a source of Ergative syntax in AN. That’s not implausible. In some AN languages (notably