Morphology in Austronesian languages

Theodore Levin and Maria Polinsky

Keywords: Austronesian languages, Formosan languages, Malayo-Polynesian languages, Oceanic languages, infix, prefix, suffix, reduplication, root, stem, voice, category neutrality, applicative, causative

Summary
This chapter presents an overview of the major morphological properties of Austronesian languages. We present and analyze data that may bear on commonly discussed category neutrality of Austronesian and suggest that Austronesian languages do differentiate between core lexical categories. The chapter addresses the difference between roots and stems showing that Austronesian roots are more abstract than roots traditionally discussed in morphology. Austronesian derivation and inflexion rely on suffixation and prefixation; inflexion is also attested. Austronesian languages make extensive use of reduplication. In the verbal system, main morphological exponents mark voice distinctions as well as causatives and applicatives. In the nominal domain, the main morphological exponents include case markers, classifiers, and possession markers. Overall verb morphology is richer in Austronesian than nominal morphology. The chapter also presents a short overview of empirically and theoretically challenging issues in Austronesian morphology: the status of infixes and circumfixes, the difference between affixes and clitics, and the morphosyntactic characterization of voice morphology.

1 Setting the stage

This chapter presents an overview of major morphological properties of Austronesian languages (AN below). AN languages are usually divided into two branches: Formosan and extra-Formosan or Malayo-Polynesian. Malayo-Polynesian (MP below) further divides into two primary branches, Western Malayo-Polynesian (WMP) and Central-Eastern Malayo-Polynesian (CEMP); see Blust (2013) for a detailed discussion. WMP languages are spoken in Madagascar, Malaysia, Indonesia, the Philippines, parts of Taiwan, Thailand, Vietnam, and Cambodia.

AN languages are predominantly agglutinative. Within Western MP, the isolating type develops in Chamic under contact with the neighboring Mon-Khmer languages. In the Oceanic branch, languages spoken further east generally have more impoverished morphology. The AN family is very large (over 1,200 languages, according to the Ethnologue) and spans an immense geographical area, from Madagascar and Southeast Asia to the Pacific, so any generalizations concerning the whole family are bound to remain rather shallow. Nevertheless, a number of properties are characteristic of AN languages: the relevance of stems in word-formation and inflection; the predominance of prefixes over suffixes; productive inflexion, extensive reduplication; articulated voice systems; and articulated systems of possession marking.
2 Main morphological units

In the morphological makeup of AN, researchers recognize roots, stems, affixes, and clitics. Some researchers also distinguish a special class of particles but those are often co-extensive with clitics. Distinguishing clitics and affixes is not always straightforward, and different authors do not always agree on the differentiating criteria. In a number of cases, the distinction between derivational and inflectional morphology is subject to debate, especially in WMP (see De Guzman 1994 on Tagalog).

2.1 Roots and stems

Following Blust (2013:360), we use the term ‘stem’ to refer to free morphemes that are capable of being affixed, and the term ‘root’ for a smaller, more abstract, “submorphemic unit that is defined by recurrent association but not by contrast”. Stems are also referred to as ‘bases’ in the literature. This definition differs slightly from the more commonly adopted definition of root as a morpheme which is not an affix (Baker and Bobaljik 2008: 24).

A typical AN root has the shape –CVC which carries a generalized meaning and can combine with so-called “formatives” (thematic affixes) to produce stems. Because of their generalized meaning, most AN roots are category-neutral and can be used to derive nouns, verbs, adjectives, etc. This category neutrality may be perceived as a sign that AN languages lack a noun-verb distinction (Broschart 1997; Gil 2000, 2009; Kaufmann 2009). For instance, the Tongan manako can be used verbally (“to like, want’), nominally (“desire”), adverbially (“willingly”), or adjectivally (“wanted, desirable”).

However, some operations in AN syntax are sensitive to the noun-verb distinction. Using Tongan again as an example, the transitivizing suffix -i can combine only with those stems that are lexically specified as verbal, cf. the contrast between (1a) and (1b). For a non-verbal stem to combine with -i it first needs to be verbalized with the prefix faka- which forms stative verbs from nouns or adjectives; these verbs can then be transitivized by adding -i, (2b).

\[(1)\]
\[
\begin{align*}
\text{a. } & \text{manako ‘want’ -- manako’i ‘desire’} \\
\text{b. } & \text{motu ‘island’ -- *motu’i}
\end{align*}
\]

\[(2)\]
\[
\begin{align*}
\text{a. } & \text{motu ‘island’ – faka-motu ‘live an island way of life’} \\
\text{b. } & \text{faka-motu -- faka-motu’i ‘separate, cut someone off’}
\end{align*}
\]

Co-occurrence with the prefix faka- is also category-sensitive; the prefix derives statives from non-verbal stems and causative verbs from intransitive verbal stems (Churchward 1953: 253-255).

2.2 Affixes

AN languages are head-initial, and many of them, including languages spoken at the geographical extremes of the family, are verb-initial or predicate-initial, i.e., VSO or
VOS. SVO as the neutral word order is found across Malayic and Micronesian languages and some Melanesian languages. With the exception of some Western Melanesian languages that have long been in contact with Papuan languages and have developed verb-final orders, including postpositions (Lynch et al. 2011: 41, 49-50, 87; Crowley 2002: 37; Lichtenberk 1983b), no other word order types are considered basic within the family. Consistent with the head-initial properties of AN languages, these languages also have more prefixes than suffixes, especially in the MP branch. AN languages are also noted for their use of productive infixation.

Compounding, though attested, is not a particularly productive form of morphologically complex word-formation in AN languages. Therefore, we do not discuss it in any detail here. It is found more frequently in Oceanic and eastern Indonesian languages (see, e.g., Elbert and Pukui 1979, Crowley 1982 for discussion).

From a morpho-phonological standpoint, the relation between stems and derived words is formally transparent, with few instances of ablaut (though see Section 5.2 for an example), umlaut, suprasegmental or subtractive morphology (see Blust 2013: Ch. 6 for discussion of each of these phenomena in AN). However, the process of affixation can at time be obscured by other morpho-phonological alternations. The most common processes observed in morphophonology include metathesis, nasal substitution (Dempwolff 1934-1938; Blust 2013; Edwards 2018 on Amarasi; Pater 1999), and reduplication (see 2.3 below). Apocope, vowel reduction, and haplology, though less well-attested within AN also render certain affixations processes less than transparent.

### 2.2.1 Prefixation

Most AN prefixes are mono- or bi-syllabic and typically contain the *a* vowel. The most common prefixes that appear on verbs across most AN languages mark the following categories: inchoative, causative, stative, active voice, intransitive (or detransitivization), instrumental voice, and collective action. As the present paper is intended to serve as a wide overview of AN morphology a comprehensive list of these morphemes is not provided (see Blust 2013: 371-381 for individual examples and reconstructions). The prefix *maN*- where *N* represents a homorganic nasal, is particularly widespread in western MP as the marker of the active/agent voice, serving a similar function to the infix `<um>`, discussed in Section 3.1. Examples of Indonesian *meN*- are presented below:

\[
\begin{align*}
(3) & \quad \text{a. } \textbf{meng-ajar} \text{ ‘teach’} & \text{Bahasa Indonesia} \\
& \quad \text{b. } \textbf{men-curi} \text{ ‘steal’} \\
& \quad \text{c. } \textbf{mem-baca} \text{ ‘read’} \text{ (Sneddon 1996:13-4)}
\end{align*}
\]

In some western MP languages, the equivalent of *maN*- is solely the homorganic nasal *N*- (e.g., in Balinese, see Arka 2004).

Another widely observed affix is the prefix *pa*- and its cognates, widely used for causative formation. In general, voice affixes of almost all Formosan and western MP languages have both verbal and nominalizing functions, but the verbal functions are primary (Blust 2013: 456). Valency-changing morphology such as voice and causativization is discussed in more detail in Section 3 below.
Beyond the formation of verbs, AN languages use prefixation to derive ordinal numerals from cardinal numerals, e.g., Marshallese ka-ruo ‘second’, Seediq (Taiwan) tege-kingal ‘first’. Possession on nominals is often indicated by prefixes although for a number of languages it is debatable whether these are actually prefixes or clitics (see Sections 2.4 and 5.4).

### 2.2.2 Suffixation

The main suffixes observed throughout the family mark applicative (or circumstantial) voice on verbs, see Section 3.1, based on the active (agent voice) form of the verb in Malagasy:

1. **n-a-mono** / **n-a-mono-an’**
   - PST-AV-kill
   - PST-AV-kill-APPL
   - ‘killed’

2. **n-i-vidi** / **n-i-vidi-anan’**
   - PST-AV-buy
   - PST-AV-buy-APPL
   - ‘bought’

3. **n-a-fina** / **n-a-fen-an’**
   - PST-AV-hide
   - PST-AV-hide-APPL
   - ‘hid’

Because of the co-occurrence of the active prefix and the applicative suffix some researchers have proposed to treat the applicative morphology as circumfixal (see fn. ii). However, the applicative forms are clearly derived in a stepwise process with a fixed ordering (Blust 2013:399). Across AN, patient voice and future tense suffixes are also commonly attested. Again, a full list of frequently occurring suffixes is outside the scope of this brief overview, but see Section 3 for further discussion of valency-changing morphology.

Another common verbal suffix is the transitivizing -'i, widely attested in Oceanic languages, see (1) and (2) above and discussion in Section 3.4.

Suffixation can also be found in the nominal domain. In Oceanic, we find suffixes deriving nouns from verbs, e.g., -Vnga. Examples from the Kwaio language (Solomon Islands) are given below.

4. **eno-nga**
   - sleep-NMLZ
   - ‘sleeping’

5. **leka-nga**
   - go-NMLZ
   - ‘trip’

6. **to’oru-ngaa**
   - live-NMLZ

---

**Kwaio**
‘living’ (Keesing 1985:77)

The choice of allomorph above is conditioned by the length of the base. Disyllabic stems are nominalized with –nga; trisyllabic roots are nominalized with –ngaa (Lichtenberk 2011).

Finally, suffixes are commonly used to encode possessors on nouns; see Section 4.4 below for further discussion on possession marking.

2.2.3 Infixation

The two main infixes that recur across AN languages are the marker of the active (actor) voice, which often has the inchoative meaning as well, and the marker of the perfective aspect (which sometimes also functions as a nominalizer). The former is reconstructed as the Proto-AN *-um-, the latter as *-in- (Blust 2013). Consider the Tagalog examples:

(6) a. lakad ‘walk’ – 1<um>akad ‘to walk’
   b. sulat ‘write’ – s<um>ulat ‘to write’
   c. basa ‘read’ – b<um>asa ‘to read’

(7) a. lakad ‘walk’ – l<in>akad ‘to walk’
   b. sulat ‘write’ – s<in>ulat ‘to write’
   c. basa ‘read’ – b<in>asa ‘to read’

Productive infixation of the kind attested above is, for the most part, found in the languages of Taiwan, the Philippines, and northern Borneo and Sulawesi (Adelaar and Himmelmann 2005); but see Davis (2003) on the productive use of <in> as a nominalizer in Hoava (Solomon Islands).

Infixation is sensitive to phonological environment; infixes are found with C-initial stems. However, when the stem is V-initial the relevant morphemes may appear as prefixes, cf. (6), (7) and (8):

(8) a. ulán ‘rain’ – um-ulán; in-ulán ‘to rain’
    b. aral ‘teach’ – um-aral; in-aral ‘to teach’

This variability of placement, dictated by the phonological environment, has led many to the conclusion that infixes are not a unique category of affix – distinct from prefixes or suffixes – but, rather, that infixes should be modeled as prefixes or suffixes whose realization within the root/stem is conditioned by phonological operations that occur after affixation, e.g. metathesis (e.g. Halle 2001 and sources therein).iv

Tagalog speakers also show variability in infix placement in loanwords. Native words in Tagalog lack initial consonant clusters. However, Spanish loanwords retain these clusters. Some speakers place in the infix after the onset, while others intersperse the infix between onset consonants (see Zuraw 2007 and references therein).

As the Tagalog data above illustrate, some languages show one allomorph of the prefix/infix elements <um> and <in> regardless of placement (but see fn. iv for a possible explanation). Other languages frequently show two (or more) allomorphs
depending on the position of the morpheme; most commonly, the infixal form is realized as -VC-, but the prefixal form is realized as C-. Toba Batak (Sumatra) illustrates this for <um>, as in tagis ‘weeping’: t<um>o-tagis ‘to weep, cry’, but inum ‘drinking’: m-inum ‘to drink’. Still other languages display different patterns of allomorphy for <um> and <in> despite their identical -VC- form. Paiwan (Taiwan) retains the -VC- of <in> in prefixal position, even in environments where <um> is realized as m- as in m-alap ‘take, pick up’: in-alap ‘one’s catch in hunting’ (Blust 2013: 384).

2.3 Reduplication

Another environment where the interaction of morphology and phonology is particularly clear is that of reduplication. Reduplication thus involves the repetition of all or part of a root or stem sometimes accompanied by some phonological alteration(s) – As reduplication is an instance of affixation, it is quite common to find that it interacts with other instances of affixation occurring before some and after others. The ordering of affixation can have consequences for the form and placement of the reduplicant.

Reduplication is probably the most pervasive morpho-phonological process in AN languages, especially in the western part of the family. Full reduplication (repetition of the entire root/stem) and/or partial reduplication (repetition of a subpart of the root/stem) are productive processes in almost all western AN languages. The functions of reduplicative processes are, however, quite variable across languages.

Full reduplication is often used to indicate distributivity or plurality as applied to entities or to events; habitual or repeated activities; ongoing events; intensity and emphasis; and increase/decrease of size/amount. Related to these two latter functions is the use of reduplicated forms as augmentatives (correlating with the marking of an increasing size or amount) and diminutives or pejoratives (correlating with the marking of decrease). Some functions of full reduplication are illustrated below. In (9), full reduplication expresses the plurality of entities, in (9) full reduplication expresses a diminutive (or possibly a similarity relation), and in (9) it expresses approximation.

(9) a. rumah ‘house’; rumah-rumah ‘houses’ Botolan Sambal
    b. anak ‘child’; anak-anak ‘doll’ Bahasa Malay
    c. mbirij ‘black’; mbirij-mbirij ‘blackish’ Karo Batak

(Blust 2013: 419-20)

AN languages also show a number of partial reduplication patterns. The full inventory of attested patterns is too great to discuss in any detail here (see Blust 2013:406-31 for a comprehensive overview), so we illustrate only with the more pervasive patterns of partial reduplication below.

So-called Ca-reduplication occurs when the first consonant is copied and prefixed to the base, along with a default vowel. This pattern of reduplication frequently forms numerals and/or deverbal nouns. As indicated in the name of the phenomenon, the default vowel is most often a, but it can also be a schwa or other vowel. Examples include Puyuma (Taiwan) kədan ‘whet’: ka-kədan ‘whetstone’ (Blust 2013) or Balinese daar ‘eat’: daa- daar-an ‘food’ (Barber 1979, Clynnes 1995). In vowel-initial bases this fixed
vowel is the reduplicant, as in Thao m-iup ‘blow on’: a-iup ‘tube used to blow on the fire’ (Chang 1998). CaC-reduplication is a variant of Ca-reduplication reported for Taba (Makian, Kayoa, and southern Halmahera). This pattern also derives instrumental nouns, as in tek ‘scoop up water’: tak-tek ‘water scoop’ (Bowden 2001).

CV(C)-reduplication – also called monosyllabic or syllable reduplication – occurs when the reduplicant consists of a syllable prefixed to the base. This syllable can be either light (monomoraic; necessarily CV) or heavy (bimoraic). In Philippine languages, it is common to have both heavy and light syllable reduplication. Tagalog displays both light and heavy CV(C)-reduplication, as in bili ‘buy’: pag-bili-bili ‘selling’ and bili ‘buy’: bí-bili ‘will buy’. In light CV-reduplication, if the first syllable of the base is heavy, the reduplicant is lightened as in Tagalog pag-la-lakbáy ‘travelling’ where the coda of the stem syllable lak is not copied into the reduplicant. Conversely, in heavy CV(C)-reduplication, if the first syllable of the base is light, the vowel in the reduplicant will be stressed, as in the Tagalog example above, or lengthened, or the syllable will be closed; see Thurgood 1997 on Bontok (Luzon). As with Ca-reduplication, vowel-initial stems copy only the vowel in the surface realization, as in Bunun (Taiwan) ma-a-asik ‘keep sweeping’ or ma-u-uktic ‘keep cutting’ (Blust 2013: 425).vi The examples above show that the reduplicant is usually a copy of some base-initial segment. However, there are also examples where the reduplicant copies other parts of the base. In Madurese, the base-final syllable is copied as in les-toles ‘write (more than once)’ or ku-buku ‘books’ (Davies 1999). A particularly striking instance of CVC-reduplication is discontinuous reduplication where the reduplicant form is determined by segments of the base that are not contiguous. For example, in Palawan, diminutive partial reduplication is obtained by copying the first CV of the base and the final C of the base: bajúʔ ‘clothing’; báiʔ-baʔů ‘child’s clothing’, libun ‘woman’: lin-libun ‘girl’ (Revel-Macdonald 1979). Similar patterns are reported for Nakanai (New Britain) (Johnston 1980) and Ulumuar Malay (Nuger 2010).vii

Lastly, CV(C)CV-reduplication also called disyllabic or foot reduplication occurs when the reduplicant copies a combination of two syllables (i.e. a foot) prefixed to the base. The second syllable of the reduplicant is frequently monomoraic, regardless of the shape of the corresponding syllable in the base. Lauje (Sulawesi) illustrates this contrast in mađel-alenda ‘rather long’ (Adelaar and Himmelmann 2005: 123). Given the size of the reduplicant, foot reduplication can be indistinguishable from full reduplication unless a base longer than two syllables is used. However, a fuller range of data shows that the two patterns do diverge. In a number of AN languages, foot reduplication occurs in addition to full base reduplication. Consider Lauje again, as in mong-oŋtong-oŋtong ‘watch (for some time)’. In full reduplication the second syllable is closed, unlike in foot reduplication. Palauan exhibits a complex pattern of prefixal foot reduplication. The reduplicant is CieCV(Cz)- as in mə-saul ‘tired’: mə-sesu-saul ‘sort of tired’, mə-dakt ‘afraid’: mə-dédak-dakt ‘sort of afraid’ (Josephs 1975). Suffixal foot reduplication is attested in Hawaiian (Pukui and Elbert 1971), for example, aloha ‘love, affection’: àloha-loha ‘express affection’. Suffixal foot reduplication is also reported in Paiwan (Ferrell 1982) and Manam (Lichtenberk 1983b).

Given the prevalence of prefixation in the AN languages, it is unsurprising that most reduplicants are prefixes, as well. However as the Hawaiian foot reduplication data above show, reduplicant suffixification is also attested. Suffixal CV(C)-reduplication is reported in
Chamorro, as in ña.laŋ ‘hungry’: ña.laŋ- ‘very hungry’ (Topping 1973) and Yapese qa.thuk- ‘to mix’: maq.thuk- ‘mixed up’ (Jensen 1977). Thao, Rukai (Taiwan), Pazeh (Chang 1998), Balinese (Clynes 1995), and Siraya (Taiwan) (Adelaar 2000) are also reported to display suffixal reduplication. Reduplicative infixation is also reported. CV-reduplicants appear infixed to the stem in Thao (Blust 2013) and Xarâcûù (New Caledonia; Moyse-Faurie 1995). Examples from the latter include atî<fî>rî ‘have confidence in’, and jikiè ‘rich’: ji<kî>kiè ‘very rich’ (see also Elbert 1988 on Rennellese).

Finally, some western AN languages have been claimed to permit triplication: the affixation of a reduplicant two times. Examples from Thao include m-apa ‘carry’: apa-apa-apa-n ‘be carried’, shkash ‘fear’: makit-shka-shka-shkash ‘slowly overwhelmed by a sense of apprehension or foreboding’ (Blust 2001). Blust also provides the example zumzum ‘hold in one’s mouth’: za-za-zumzum ‘keep holding in one’s mouth’. Here, Ca-reduplication appears to have occurred twice. (Müller-Gotama 2001 provides possible examples from Sundanese.)

2.4 Clitics

Traditionally, clitics are characterized as morphemes that show contradictory behavior in their syntactic and phonological properties. From the point of view of syntax they are independent elements; from the point of view of phonology, their realization depends on another word or phrase.

A distinction can made between peripheral and second-position clitics. Peripheral clitics appear at the beginning or end of the constituent to which they belong. These clitics are found in almost all western AN languages. Second-position (Wackernagel) clitics usually occur after the first constituent of the phrase to which they belong, and are only attested in Formosan and western MP languages. Consider the following data from Tagalog. Its basic word-order is verb-initial; therefore, second position clitics normally occur immediately after the verb (10a,b). However, when a negative or other adverbial element appears in pre-verbal position, as in (10c), the clitics will also precede the verb.

(10)  a. Ibinigay=ko=na ang pera kay Charlie Tagalog IV.PFV.give=1.SG.GEN=PRF CM money DAT Charlie
      ‘I already gave the money to Charlie.’

   b. Nagtatrabaho=rin=ho=ba=kayo roon? AV.IPFV.work=also=HON=Q=2.PL.NOM there
      ‘Are you working there too, sir?’

   c. Hindi=pa=man=lamang=tuloy nakakapag-almusal si Juan.
      NEG=IPFV=even=only=as.result eat-breakfast NOM Juan
      ‘As a result, Juan hasn’t even had breakfast yet.’ (Kroeger 1998:1-2)

Both peripheral and second-position clitics are usually unstressed and form a prosodic unit with either the following word (proclitics) or the preceding word (enclitics). Because of their positional variability, it is generally easy to distinguish second position clitics from affixes. This distinction is less clear in the case of peripheral clitics. There are three
types of evidence employed to distinguish clitics from affixes: (i) Clitics generally do not trigger morphophonological alternations of the stem. (ii) Clitics tend to be less selective than affixes with regard to the category of their hosts. (iii) Clitics can show variability with regard to their position within morphologically complex elements. The order of affixes, however, is often rigid. Despite these diagnostics, there are some cases where the clitic-affix distinction remains uncertain. With respect to pronominal clitics Blust (2013: 403) notes, “In general, atonic monosyllabic pronouns that attach to an independent word are treated as clitics rather than affixes, although the basis for this decision is rarely explicit.”

In terms of their meaning and functions, AN clitics represent the following main types: (i) pronominal, (ii) aspectual/modal, (iii) negative, (iv) determiner, (v) emphatic, and (vi) clitics which encode question markers, politeness markers, as well as mirativity and evidentiality. In Tagalog, pronominal clitics are seen in =ko `1SG.GEN’ (10a) and =kayo ‘2PL.NOM’ (10b); see also Section 5.3 for a fuller discussion of pronominal forms. Aspectual clitics =na (10a) and =pa (10c) indicate perfective and imperfective aspect respectively. Finally, the clitic =ba encodes a question in (10b).

3 Valency-changing morphology

Valency alternations involve the addition of logical arguments to or subtraction of those arguments from a clause, as well as manipulations of the mapping of logical arguments into different clausal constituents.

3.1 Voice and transitivity

3.1.1 Voice

Perhaps no morphosyntactic category has been as much investigated in AN languages as voice. AN languages have constructions that resemble passives of more familiar Indo-European languages, as well as antipassive constructions. However, so-called symmetric voice systems are the primary focus of interest, and their analysis is still quite controversial, as discussed below.

English-like passives are found in numerous AN languages, though they are rare in Melanesian languages. The logical object is realized as subject, the logical subject is demoted to an oblique, adpositional phrase, and passive morphology appears on the verb:

(11) a. saya di-jemput (oleh dia).
    1SG PASS-meet by 3SG
    ‘I was met by him.’ (Sneddon 1996:248)

b. i koohete-tia a Pani e Huia
    T/A scold-PASS PERS.DET Pani by Huia
    ‘Pani was scolded by Huia.’ (Bauer 1993:396)
Some languages allow such passivization freely and extend it to intransitive predicates as well, cf. in Hawaiian:

(12) a. ua komo-hia ka mana’o i loko ona Hawaiian
  PRF enter-PASS DET thought to inside 3SG.POSS
  ‘A thought occurred to him.’
b. ua hae-hia ka ʻīlio
  PRF bark-PASS DET dog
  ‘The dog was angry.’ (Elbert and Pukui 1979: 86)

In some cases the passive morphology on the verb is zero. For example, alongside the marked passive as in (11a), Indonesian has a so-called bare passive (Chung 1976). Altogether, in Indonesian and Malay, researchers distinguish morphologically-marked active (with the prefix man-), morphologically-marked passive with di-, bare active, and bare passive. As a consequence of null morphology, it may be difficult to distinguish passives from null argument constructions. For example, subjects and objects can be dropped in Marshallese for some verb classes, making the analysis of the ‘passive’ in (13) unclear (similarly in Hoava, see Davis 2003, and Fijian, with Schütz and Nawadra 1972 arguing against the passive analysis and Kikusawa 1998 arguing in its favor).

(13) John e=naaj ɱwij~ɱwij rainin. Marshallese
  John 3SG.AGR=FUT operate-INTR today
  (a) ‘John will operate today.’
  (b) ‘John will be operated on today.’ (Willson 2010: 233, citing Bender 1969)

Even in languages with overt passive morphology, several strategies may be used (as in Marshallese, Willson 2010: 238), or the marking may be ambiguous with other verbal categories, such as transitive and perfective. (See Willson 2010 for Marshallese, Cook 1996, Mosel and Hovdaughen 1992: 198-204, 729-743 for Samoan.)

In the antipassive construction, it is the logical object that is demoted and may be left unrealized, and again special morphology often appears on the verb (Polinsky 2017). Antipassives are less commonly discussed amongst AN languages, in part because different researchers often use different operational criteria by which they define the antipassive and in some instances antipassives are conflated with intransitives or durative/imperfective forms. Chamorro displays a transitive-antipassive alternation, whereby the logical object is realized within an adpositional phrase and special morphology (man- or fan- depending on mood) appears on the verb (Cooreman 1988, Chung 1998). Even in languages where the presence of antipassives is not contested, their formation is lexically restricted (cf. Chung 1998: 37-39 for Chamorro).

(14) a. Un-patek i ga'lagu. Chamorro
  2S.SBJ-kick the dog
  ‘You kicked the dog.’
b. Mamatek hao gi ga'lagu.
  ANTIP.kick 2S.NOM LOC dog
  ‘You kicked at the dog.’ (Cooreman 1988: 578)
Pseudo noun-incorporation constructions and constructions with oblique object in Polynesian languages, like Niuean and Tongan, have also been treated as antipassive constructions (Polinsky 2017), as they have actor voice clauses in symmetrical voice-systems (Aldridge 2012; see immediately below for discussion).

The symmetrical voice systems of Western AN languages are not limited to more familiar active-passive or transitive-antipassive alternations, at least not straightforwardly. In these systems, there is no morphologically unmarked voice, and no argument is demoted to an oblique (Kroeger 1993, Foley 2008, Pearson 2005). Instead, symmetrical voice systems are characterized by the fact that a single argument of the clause – possibly a non-core argument – is privileged in certain ways. This argument may be in a certain linear position or receive a particular morphological marking. Dedicated morphology on the verb indicates which argument of the verb was chosen for this special status. Furthermore, this privileged argument is often the only one that can be wh-questioned, relativized, or focused—the restriction often referred to as the subject-only constraint (Schachter and Otanes 1972, Keenan 1976, Schachter 1976).

Consider the symmetrical voice alternation in Squiliq Atayal (Taiwan):

(15)  

|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
|   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |   |
| (15) | a. | 
| b. | Niq-un na’ Tali’ qu’ quilih qasa. | Patient Voice | eat-PV GEN Tali QU fish that | ‘Tali ate the fish.’ |

(Erlewine et al. 2017: 374; citing Liu 2004)

In each example, one argument of the verb (in italics) is in sentence-final position preceded by the marker qu’. Voice morphology on the verb (in bold) reflects this choice of argument.

The number of voices in a given language varies. Philippine and Formosan languages have been reported to have four or five distinct voices. Malagasy is usually analyzed as having three voices. Malay/Indonesian languages are frequently reported to have just two symmetrical voices, though they also have a more diverse inventory of active-passive alternations, as well.

3.1.2 Transitivity

Generally speaking, transitivity refers to a predicate’s ability to take a direct object, or “close object” as it is referred to in the literature on some Oceanic languages (Lynch et al. 2011). The issue of how transitivity is marked in AN languages, especially Formosan and
Western MP languages is particularly controversial and is connected to the status of symmetrical voice systems (see also Section 5.1 on case-alignment). Some treatments take voice affixes to be indicative a verb’s transitivity, while others maintain that voice morphology is orthogonal to the issue of transitivity. This latter position is most clearly supported in Split-S systems, like Balinese (Arka 2004) and Achenese (Legate 2014), whereby verbs that are clearly intransitive – they cannot take a direct object – can appear with either actor or patient voice forms depending on the thematic role of the sole argument, for example,

(18)  
   a. Ia ulung  
       3  ov.fall.down  
       ‘(S)he fell down.’
   b. Ia ng-eling  
       3  AV-cry  
       ‘(S)he cried.’ (Arka 2004: 40)


The distinction is more clear-cut in Oceanic languages where symmetrical voice systems do not arise. Oceanic languages frequently display formal marking for transitivity on verbs. Transitive predicates are marked with a suffix that can be derived from the Proto-Oceanic form *-i; see the Tongan examples in (1) and the following examples from Tigak (Papua New Guinea):

(19)  
   a. nonol ‘be thinking’ –  nol-i ‘think about’  
   b. visvis ‘fight’  –  vis-i ‘hit’
   c. kalkalum ‘look’  –  kalum-i ‘see’ (Lynch 1998:85)

In Proto Oceanic, *-i was generally added to an intransitive root with a final consonant. However, most Oceanic languages have lost these final consonants. Therefore, when the consonant is retained before a transitive affix, it is treated as part of the suffix, for example, consider Fijian:

(20)  
   taŋi / taŋi-ða  
   weep / weep-TR.3SG.OBJ  
   ‘weep’ ’cry for’ (Lynch et al. 2011:44)

In Nakamanaga (Papua New Guinea), the transitivizing suffix is –gi, and in Lenakel (Vanuatu) it is –in (Lynch 1998). Ulithian (Micronesia) shows considerable allomorphy in the transitive suffix, including /-ðI/, /-lI/, /-ŋI/, /-sI/ and /-xI/. These forms also show vowel harmony whereby /i/ becomes /u/ when affixed to a stem ending in schwa or a back vowel. The language also shows the forms /-fi/, /-mi/, and /-ri/ which are insensitive to the preceding vowel. All these variations are claimed to be lexically idiosyncratic (Lynch et al. 2011).

In Xărācūù it is reported that a single suffix, –ri, is used to form both transitive and applicative verbs (Moyse-Faurie 1995). A similar pattern is found in Saliba (Papua New Guinea; Margetts 1999).
Many Oceanic languages also display a transitivity alternation that employs reduplication (Lynch et al. 2011, Blust 2013). Here, intransitive forms are morphologically more complex than their transitive counterparts, and are marked with partial reduplication. This is illustrated in the Tolai language (Papua New Guinea); compare *iu ‘to wash’: *iuiu ‘to bathe’, *tumu ‘to write down’: *tutumu ‘to write’, or *kal ‘to dig up’: *kakal ‘to dig’ (Blust 2013: 684-685) and in Fijian, as in *cula ‘sewn’: *cula-cula ‘sew’ or *rabe ‘kick’: *rabe-rabe ‘do a lot of kicking’ (Dixon 1988).

### 3.2 Causativization

Causative constructions are valency-increasing constructions that serve to indicate that the subject causes someone/something to bring about an event, causes a change in state, or allows for the occurrence of an event or change of state. Cross-linguistically, causativization can target both transitive and intransitive predicates, thought there are language-specific constraints on the predicates that can be causativized.

Causativization is frequently signaled by reflexes of the proto-AN prefix *pa/paka- (Blust 2013), as in Amis (Taiwan) *pa-, and Roviana (Solomon Islands) *va-.

(16) a. **Pa-rakat kaku t-u paliding.** *Amis*
   CAUS-walk 1.S. Nom Dat-CN car
   ‘I drive the car.’

b. **Pa-ka’en kaku ci panay-an t-u pawli.**
   CAUS-eat 1.S. Nom PPN Panay-Dat Dat-CN banana
   ‘I feed Panay banana.’ (Wu 2006: 298)

(17) a. **Va-mae-a sa magu.** *Roviana*
   CAUS-come-3sg Def knife
   ‘Give me the knife!’

b. Keke tootoso **va-la-i-u ri pa Vira Haba.**
   One time CAUS-go-TR-1sg Loc Vira Harbor
   ‘One time, they made me go to Vira Harbor.’ (Corston-Oliver 2011:482)

Zeitoun and Huang (2000), suggest that *paka- (attested in, e.g., Paiwan paka-, Chamorro faha-, and Fijian vaka-) may be analyzed as multi-morphemic, consisting of the causative prefix *pa- ‘causative’ and the stative prefix *ka- (see Blust 2013: 372-380 for further discussion).

### 3.3 Applicativization

Like causativization, applicativization is a valency-increasing operation. However, in this case it is an oblique argument of the verb that is promoted to function as a grammatical object also called the applied argument. The applied argument, sometimes called a remote object in the Oceanic literature (Lynch et al. 2011: 44), can bear one of many thematic roles including benefactee, malefactee, goal, location, instrument, etc. Some languages utilize distinct morphemes across a number of applicative constructions.
Applicatives are divided into two groups cross-linguistically: high and low (Pylkkänen 2008). High applicatives can occur with intransitive predicates, implicate no relationship between the applied argument and the logical object and can treat either argument as the grammatical object for the purposes of syntactic operations. Low applicatives only occur with transitive predicates, they implicate a relationship between the applied argument and the logical object (frequently one of possession), and the applied argument is privileged for syntactic operations reserved for the grammatical object.

In Oceanic languages, applicativization is signaled by the use of a suffix related to the Proto Oceanic *-aki(ni) (Harrison 1982, Lynch et al. 2011), cf. in Niuean:

(18)  
a. Kua hele aki tuai e ia e titipi e falaoa.  
\[ \begin{align*} 
\text{PERF cut} & \text{APPL PERF ERG 3SG ABS knife ABS bread} \\
\text{Niuean} & \\
\text{‘He has cut the bread with the knife.’} 
\end{align*} \]

b. Kua hele tuai e ia e falaoa aki e titipi.  
\[ \begin{align*} 
\text{PERF cut} & \text{PERF ERG 3SG ABS bread with ABS knife} \\
\text{‘He has cut the bread with the knife.’} & \text{(Seiter 1980: 243-4)} 
\end{align*} \]

It is common for the form of the applicative affix to bear formal similarity to adpositions in the same language. (18b) shows a prepositional oblique marked with aki. whereas. (18a) shows that aki can appear within the predicate. When it does, the instrument, like the theme, is marked with absolutive case.

It has also been suggested that additional symmetrical voices beyond agent and patient can be analyzed as applicative constructions (Aldridge 2004). In these cases, an adjunct or indirect object is introduced as the highest internal argument. The co-occurrence of agent and/or patient voice morphology along with applicative morphology is particularly clear in Malay/Indonesian languages, e.g. Balinese (Arka 2004) and Bahasa Indonesia (Sneddon 1996). Javanese illustrates a pattern where the applicative suffix –ake can appear with both active and passive constructions; this suffix freely combines with active verbs, indicated by the homorganic nasal prefix N-, (19), or passive verbs, with the prefix di, (19).

(19)  
a. Ani n-ulis-ake Tono layang  
\[ \begin{align*} 
\text{Ani} & \text{ACT-write-APPL Tono letter} \\
\text{Javanese} & \\
\text{‘Ani wrote Tono a letter’} 
\end{align*} \]

b. Tono di-tulis-ake Ani layang  
\[ \begin{align*} 
\text{Tono} & \text{PASS-write-APPL Ani letter} \\
\text{‘Tono was written a letter by Ani’} & \text{(Nurhayani 2012:1-2)} 
\end{align*} \]
4 Verbal agreement and related phenomena

4.1 Verbal agreement

Descriptively, agreement is a pattern of morphological variation on verbal elements conditioned by the grammatical features (person, number, gender, case, etc.) of noun phrases in a clause.

AN languages with agreement generally mark agreement with the subject (regardless of their case-marking alignment; see Section 5.1 on alignment), and that marking is achieved using two basic strategies: affixal verbal agreement, common, for example, in Melanesian languages (see Yamada 2006 for representative languages and examples), and clitics (in Micronesian, some dialects of Fijian, Rotuman, Polynesian, as well as some Melanesian languages). Micronesian languages and some Melanesian languages also show agreement with the object (see Song 1994 for an overview of the Nuclear Micronesian data). Compare in Puluwat (Micronesia), which illustrates a common pattern of agreement:

\[\text{(20) Wurumwo ya yáékékél-ee-ɾ átekkit mákk Puluwat Wurumwo 3SG.SBJ teach-TRANS-3PL.OBJ children writing ‘Wurumwo taught the children writing.’ (Elbert 1974: 86)}\]

In Melanesian languages subject-agreement markers are often portmanteau forms which combine with the expression of the tense/aspect/mood categories of the verb (Lynch et al. 2011). In other languages, while subject agreement is independent of tense/aspect/mood, its form is nevertheless conditioned by such specifications, e.g. Palauan (Georgopoulous 1991, Nuger 2016).

Verbal agreement is less well-attested in western MP languages, though see Legate (2014) and sources therein on Acehnese; however a number of researchers have argued that symmetrical voice morphology (discussed in Section 3.1.1), in fact, reflects a form of agreement, namely case-agreement. On such accounts, voice reflects the outcome of independent syntactic processes of case-assignment and syntactic movement (e.g., Pearson 2005; Rackowski 2002; Rackowski and Richards 2005).

Two questions are of interest in relation to the morphosyntax of agreement in AN. First, a number of AN languages are pro-drop, but it is not yet clear if there is a correlation between the availability of pro-drop and the availability of agreement. For example, Micronesian languages, which have relatively rich agreement, also have pro-drop, but so do agreement-poor Melanesian languages such as Cheke Holo (Palmer 2009a). Presumably, the conditions on subject drop or topic drop may differ depending on whether or not a given language has agreement, but these conditions still need to be better studied. The potential relationship between pro-drop and rich agreement has been very widely investigated in non-AN languages, and AN languages have the potential to inform the debate.

The second issue has to do with the categorical nature of agreement markers: are they affixes or clitics? For subject markers, it is generally assumed that they are clitics, often on the basis of separability from the verb. For Rotuman, which on the surface seems to employ subject suffixes, den Dikken (2003: Ch. 6), following Vamarasi (2002), argues
that these are also clitics. His main argument is that the apparent “pronominal suffixes” attach to any element on their immediate left and do not select for a particular category of host. He also cites diachronic evidence that some pronominal suffixes develop from clitics. However, that does not necessarily mean that a clitic cannot change its category. Object markers are generally assumed to be suffixes (cf. Song 1994).

As mentioned above, the distinctions between affixation and cliticization have lately generated a lively debate in theoretical linguistics. We would like to emphasize the rich empirical potential offered by AN languages in this area.

4.2 Wh-Agreement

Wh-agreement is a phenomenon whereby canonical verbal agreement is obscured or altered when the argument cross-referenced by agreement is relativized, topicalized, or focused. The pattern is found in a variety of AN languages, mostly in Western MP languages, and has been well described for Chamorro (Chung 1982, 1994, 1998), Palauan (Georgopoulos 1991), Tukang Besi (Sulawesi; Donohue 1999), and Malagasy (Pearson 2005).

The wh-agreement patterns of Chamorro can be used as a representative example of the effect. In Chamorro transitive, realis clauses, subject agreement is realized as a prefix on the verb as in (21), where ha- cross-references the subject si Juan. However, when the subject is wh-questioned (21), ha- is not realized on the verb. Instead, the verb carries the infix <um>. Similarly, if the object is wh-questioned (21), ha- again is not realized; instead, the infix <in> appears.

(21) a. Ha-fa'gasi si Juan i kareta.  
   Chamorro  
   3SG-wash DET Juan DET car  
   ‘Juan washed the car’

b. Hayi f<um>a'gasi i kareta?  
   who WH.SUBJ.wash DET car  
   ‘Who washed the car?’

c. Hafa f<in>a'gase-nña si Henry pära hagu?  
   what WH.OBJ.wash.RED-3S DET Henry for you  
   ‘What is Henry washing for you?’ (Pearson 2005:410)

Wh-agreement shows a clear relationship to other patterns of ‘anti-agreement’ phenomena cross-linguistically (see, e.g., Ouhalla 1993 on Berber and Baier 2017 for a recent comprehensive overview). Furthermore, Pearson (2005) has argued that wh-agreement systems are related to symmetrical voice systems (see Section 3.1.1). The voice morphology of a language like Malagasy or Tagalog may, in fact, be wh-agreement morphology. This connection is particularly clear in Chamorro. The Chamorro subject wh-agreement marker <um>/mu- is cognate with the Tagalog actor voice marker <um>, and the object wh-agreement marker <in> is cognate with the patient voice marker <in> (Topping 1973, Donohue and Maclachlan 1999, Pearson 2005). The ubiquity of voice in many Western MP languages is then attributed to the ubiquity of topicalization (or focus) in every clause in these languages. Wh-agreement remains an active area of research in AN linguistics.
5. Nominal morphology

5.1 Case

Generally AN nouns do not inflect for case. Case-marking is expressed by particles, if at all, which can be analyzed as either case-marking clitics or prepositions. A number of researchers specifically argue that they are prepositions (see Broschart 1994, which also includes a review of earlier research). Two arguments support this conclusion. First, case-marking particles are often homophonous with the actual prepositions used in a given language. For example, the prepositions in Māori (New Zealand) are i ‘in, to’, ki ‘toward, at’, e ‘from, by’, and a ‘of’. Of these, i also marks direct objects, ki, indirect (possibly dative) objects, and e passive by-phrases. The second argument in favor of treating these markers as prepositions and not pure case-markers, comes from the fact that the oblique case forms do not combine with any prepositions, an unexpected distributional pattern if some forms with case-markers are to be governed by prepositions. There is no consensus in the literature on what these elements are, and their status is unlikely to be uniform across different languages.

In Philippine, Formosan, and some Oceanic languages, the exact form of nominal marking can be affected by definiteness and/or specificity and noun type. Frequently, unique forms are used to mark proper names, and in some cases indefinite or non-specific nouns are marked differently than definite/specific ones.

AN languages show a variety of case-alignment patterns. The majority of AN languages show a nominative-accusative alignment: the subject of an intransitive verb and the subject of a transitive verb appear in the same case (nominative), to the exclusion of the object (accusative). Tahitian in (22) is representative. Subjects are unmarked, and objects occur with the accusative particle ‘i.’

(22) a. te ma'ue nei te mau manu.  
   ASP fly ASP DET PL bird  
   ‘The birds are flying.’

b. ’ua ’ite te tamaiti ’i te mau manu.  
   PERF see DET child ACC DET PL bird  
   ‘The child saw (the) birds.’

Numerous AN languages show neutral alignment: noun phrases show no overt case marking. In such languages, grammatical functions are distinguished by agreement and/or word order alone. Lewo (Vanuatu) is an example:

(23) a. sira Ø-puyu Ø-pa ne Ø-tol metava.  
   girl 3SG.SBJ-climb 3SG.SBJ-go now 3SG.SBJ-reach above  
   ‘The girl climbed on up to the top.’

b. omami me-muni wii  
   1PL.EXCL 1PL.EXCL.SBJ-drink water  
   ‘We drank water.’ (Early 1993: 70, 73)
Some AN languages are clearly ergative; the subject of an intransitive verb and the object of transitive verb appear in the same case (absolutive), to the exclusion of the transitive subject, which is marked ergative, for example,

(24) a. na'e 'alu 'a Sione ki he ako.  
PST go ABS John to OBL school  
'John went to school.'

b. na'e tua'i 'e Sione 'a Mele.  
PST call ERG John ABS Mary  
'John called Mary.'

Ergative alignment is found in Western Polynesian languages, Roviana (Corston 1996, Corston-Oliver 2011), several languages of New Caledonia (Bril 1997, 2002; Moyse-Faurie and Ozanne-Rivierre 1983), and Melanesian languages such as Motu (Lister-Turner and Clark 1930; Dixon 1994: 58), Hula (Pat 1996, Ball 2007), and Sinaugoro (Tauberschmidt and Bala 1992).

Still other languages show a split-intransitive alignment whereby intransitive subjects are not uniformly marked. The distinction frequently correlates with the unergative/unaccusative distinction. Unergative subjects are marked like transitive subjects, while unaccusative subjects are marked like transitive objects. Split-intransitive analyses have been proposed for Acehnese (Durie 1985, 1987), Balinese (Arka 2004), and Tukang Besi (Donohue 1999) among others.

A number of AN languages show alignment patterns that are difficult to describe in these established terms. Philippine, Formosan languages, and Malagasy have been particularly subject to controversy as to what kind of alignment they represent. Consider the following examples from Tagalog:

(25) a. b<in>ili ng babae ang isda  
<ASP>buy NG woman ANG fish  
'The woman bought fish.'

b. b<um>ili ang babae ng isda  
<ASP>buy ANG woman NG fish  
'The woman bought fish.'

c. d<um>ating ang babae  
<ASP>arrive ANG woman  
'The woman arrived.'

The status of the markers *ang* (see also (10) above) and *ng*, which occur with common nouns, is subject to an ongoing debate concerning the relationship between verbal morphology and nominal-argument marking. When the verb takes the infix *–in*, the marker *ang* appears on the object; when the verb takes the marking *–um–*, *ang* appears on the subject. Some researchers suggest that Tagalog and related Philippine languages are ergative; the marker *ang* marks the absolutive case, while *ng* marks ergative and/or oblique case (Payne 1982; De Guzman 1988; Manning 1996; Wechsler and Arka 1998; Aldridge 2004, 2006, 2008; see Erlewine et al. 2017 for a recent review). On the other hand, Rackowski (2002) and Rackowski and Richards (2005) propose that Tagalog is a nominative-accusative language. Under this approach, the markers *ang* and *ng* are not
indicators of case at all. Instead, the main function of *ang* is then to mark the highest structural argument in a given configuration. As mentioned in Section 4.2, a similar approach is proposed for Malagasy by Pearson (2005).

### 5.2 Number and gender

Plurality of nominals can be indicated in a number of ways across AN languages. These include affixation; reduplication, as in (26); and the use of special pluralizer words, which appear to be similar to classifiers (Dryer 1989; Lynch at al. 2011: 174), cf. (26):

(26) a. pulau / *pulau*-pulau  
    island / RED-island  
    ‘island’ / ‘islands’ (Sneddon 1996: 21)

b. e faiaoga / e *tau* faiaoga  
    ABS teacher / ABS PL teacher  
    ‘(the) teacher’ / ‘(the) teachers’ (Seiter 1980: 38)

Polynesian languages distinguish singular and plural forms of nouns that refer to humans by root modification. For instance, in Māori, the vowel undergoes ablaut to indicate plurality; singular *tangata* ‘person’ and *tuahine* ‘(man’s) sister’ lengthen to form plural *tiingata* ‘people’ and *tuiihine* ‘(man’s) sisters’ (Bauer 1993: 106; Lynch et al. 2011: 38). However such pluralization strategies are optional. Given the appropriate context, plural interpretations can be achieved in the absence of this morphology.

Very few AN languages have gender/noun class. In those languages that have gender distinctions, the agreement in gender is registered on the determiner. For example, Teo (Papua New Guinea) has three noun classes: *e*-class (personal names, people with high social status, pets), *a*-class (all other humans, vertebrates, landmarks), and *o*-class (plants, amorphous masses) (Mosel 2007). Some AN languages also show gender distinctions in pronominal form. The Windesi dialect of Wandamen (Western New Guinea) is reported to make an animate/inanimate distinction (Anceaux 1961). In the Mataram-Selong dialect of Sasak (Lombok), the low register second person distinguishes male and female addressee as *ante* (male), and *kamu* (female) (Austin 2000). In addition, Sellato (1981) has reported a ‘three gender’ system of personal pronouns used by several groups of nomadic Punan who inhabit the Muller-Schwaner Mountains of southeast Borneo.

### 5.3 Pronominal forms

Most AN languages make a distinction between inclusive and exclusive forms in the 1st person plural. The former groups the speaker and the addressee, the latter groups the speaker and others to the exclusion of the addressee(s). This distinction is found in all paradigms of person pronouns (subject, object, possessor, etc.). Some exceptions to this generalization can be found in Western MP languages (see Stevens 1968, Donohue and Smith 1998, Robson 2002). In languages without an inclusive/exclusive distinction, pronominal number distinctions have for the most part collapsed entirely, as in Javanese and Madurese (Blust 2013).
Western MP pronominal systems usually do not include dual or trial forms – referring to groups of two and three, respectively. Instead, they frequently make only a singular/plural distinction if one is made at all. Many Oceanic languages, however, do make use of dual and sometimes trial/paucal pronouns. Paucal pronouns pick out small groups, usually between 3 and 10. In most Oceanic languages, the dual is derived by adding the numeral ‘two’, and the trial/paucal by adding the numeral ‘three’ to the pronominal stem, although this can be obscured by phonological processes. The Boumaa Fijian pronominal paradigm presented below illustrates both clusivity distinctions in the 1st person plural as well as a 4-way number distinction:

<table>
<thead>
<tr>
<th></th>
<th>SINGULAR</th>
<th>DUAL</th>
<th>PAUCAL</th>
<th>PLURAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>INCL.</td>
<td>EXCL.</td>
<td>INCL.</td>
<td>EXCL.</td>
</tr>
<tr>
<td>1ST</td>
<td>Yau</td>
<td>'etaru</td>
<td>'eirau</td>
<td>'etatou</td>
</tr>
</tbody>
</table>
| 2ND    | 'i'o     | 'emudrau| 'emudou| 'emunuu
| 3RD    | 'ea      | (i)rau | (i)ratou| (i)ra  |

(Dixon 1988; Harley and Ritter 2002:494)

Other AN languages known to recognize more than a singular/plural pronominal number distinction are found in central and western Borneo. Most languages in the area have a dual form, and many have a trial form as well. Most striking of all are Kenyah languages that show a five-way person distinction – singular, dual, trial, quadral and plural numbers – in addition to having an inclusive/exclusive distinction (Wong and Mantenuto 2017; Smith 2013, and further references therein).

5.4 Possession marking

AN languages commonly categorize entities as directly vs. indirectly possessed. Semantically, direct possession corresponds roughly to inalienable possession, also referred to as obligatory, inherent, subordinate, or realized possession. Beyond truly inalienable entities like body parts, the semantics of (in)alienability is not entirely predictable; it has been subject to much discussion in the literature (see Lynch 1973, 1997, Lichtenberk 1983a, 1985, 2005, Wilson 1982, Bickel and Nichols 2013, Nichols and Bickel 2013, and references therein). Indirect possession includes everything that can be alienably possessed, and is also referred to as dominant or unrealized possession. Morphologically, the distinction in possession types is marked in diverse ways across AN. In Drehu (New Caledonia), inalienable possession is marked by an affix on the head noun indicating that the possessor and alienably possessed nouns have only a freestanding possession marker:

<p>| | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>la</td>
<td>pengö-</td>
<td>ng</td>
<td>/ keme-</td>
</tr>
<tr>
<td>DET</td>
<td>love</td>
<td>manner-</td>
<td>INAL.1SG</td>
<td>father-</td>
</tr>
<tr>
<td></td>
<td>‘my manner’</td>
<td>‘their father’</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

a. la pengö- | ng / keme- | hun
| DET manner- | INAL.1SG | father- | INAL.3PL |

b. la ihnim | i angeic
| DET love | PRP 3SG
Inalienable possession marking may be obligatory, with 3rd singular typically being the default, citation form. Alienable possession marking is never obligatory.xii

In Polynesian languages, the distinction between inalienable and alienable possession is represented as the contrast between two series, the o series corresponds roughly to inalienable possession, and the a series corresponds roughly to alienable possession, for example,

(29)  
\begin{align*}
\text{a. } & \text{te Pukapuka a Heremaia} & \text{Māori} \\
& \text{DET book A Jeremiah} \\
& \text{‘The Book of Jeremiah’ (written by him)} \\
\text{b. } & \text{te Pukapuka o Hōhua} \\
& \text{DET book O Joshua} \\
& \text{‘The Book of Joshua’ (written about him) (Bauer 1997: Ch. 12)}
\end{align*}

Within alienable possession, many Oceanic languages further distinguish several categories based on salient properties of objects (see Lichtenberk 1983a for an overview and Bender and Beller 2006 for an overview and historical reconstruction). The most common, and rather simple, system is one that divides entities into food, drink, and everything else. Micronesian languages have a more articulated classification (cf. Dyen 1965, Benton 1968 for Chuukese; Rehg 1981 for Pohnpeian; Lee 1975 for Kosrae).xiii The classification into categories such as ‘food’, ‘drink’, ‘general’, etc., is encoded by freestanding expressions inside the noun phrase which are indexed for the person and number of the possessor; we will gloss them as CLF. Compare in Iaai (New Caledonia):xiv

(30)  
\begin{align*}
\text{a. } & \text{bele-n kaiɔ} & \text{Iaai} \\
& \text{CLF.DRINK-3SG.POSS water} \\
& \text{‘his/her water’} \\
\text{b. } & \text{hanii-n wɔɔ} \\
& \text{CLF.FOOD-3SG.POSS fish} \\
& \text{‘his/her fish’} \\
\text{c. } & \text{api-n meie} \\
& \text{CLF.GEN-3SG.POSS fire} \\
& \text{‘his/her fire’ (Ozanne-Rivierre 1976: 189)}
\end{align*}

The actual category of these classificatory expressions has been subject to debate. Most researchers agree that they are syntactically heads; their order in the noun phrase follows the general headedness principles of a language. In most AN languages they precede the noun denoting the possessum, although in VSO Micronesian languages they follow the noun. Most researchers agree that these expressions are different from sortal and measure classifiers familiar from such languages as Chinese or Thai (see, however, den Dikken 2003: Ch. 2 for an argument for them being more similar to the familiar classifiers than one would assume).
Unlike the better-known Southeast Asian classifiers, AN classifiers are not obligatory in counting, their inventory is more limited than that of familiar classifier languages, and most importantly, they do not serve to individuate and atomize nouns (cf. also Palmer and Brown 2007: 203).

Some researchers suggest that these words are a special closed class of nouns which take pronominal possessive marking and nominal dependents (Palmer and Brown 2007; Palmer 2009b). This approach relies on the parallelism between the classifiers and inalienably possessed nouns, which are also indexed for the person and number of their possessor using the same marking. The alternative, proposed by Lichtenberk (1983a, 2009), is that these expressions should be considered “relational classifiers”, thus functional elements, whose main purpose is to individuate the relation between the possessor and possessum under indirect possession. Lichtenberk’s main morpho-syntactic argument against treating these words as nouns comes from the fact that they are typically monosyllabic/monomoraic, while all other lexical nouns in Oceanic are disyllabic and/or bimoraic. Thus, “classifiers” do not meet the minimal nominal word criterion (Lichtenberk 2009: 385).

**Further readings**
An excellent overview of AN morphology, with extensive historical reconstructions, can be found in Blust (2013: Chapter 6). Lynch et al. (2011: Chapter 3) offer an overview of morphology in Oceanic languages, followed by several grammar sketches. Adelaar and Himmelmann (2005: Chapter 5) presents an overview of morphology in Formosan and Western MP languages, also followed by several grammar sketches.

**Abbreviations**

**Acknowledgements**
We would like to thank Marcel den Dikken, Michael Yoshitaka Erlewine, and Paulina Lyskawa for comments on previous versions of this chapter. Unless stated otherwise, illustrative examples are from the authors’ field notes. This work was supported in part by NSF grant BCS-1619857.

**References**


---

**Theodore Levin**
National University of Singapore
*tedlevin@nus.edu.sg*

**Maria Polinsky**
University of Maryland
*polinsky@umd.edu*

---

ii A final affixal category that we do not discuss in detail here is the *circumfix/refix*, a prefix-suffix unit. It is not immediately obvious whether these elements should be treated as single morphemes or as combinations of two, prefix and suffix, that occur together frequently or always. Some ‘true’ circumfixes are claimed to exist in the AN literature. Schachter and Otanes (1972) describe Tagalog (the Philippines) *pag- … -an* as an Object-Voice counterpart of Actor-Voice verbs prefixed by *mag-* , and Malagasy (Madagascar) demonstratives (as discussed in Rajemisa-Raolison 1971, *a.o.*) have been described as phrasal circumfixes.

iii See Blust (2013: 389-92) for discussion of the plural infix *-ar-* (which is significantly less well represented in modern AN languages).

iv This may, in fact, be an over-simplification of the Tagalog data because Tagalog has been argued to lack true V-initial stems. Such stems are said to begin with a glottal stop.
[?] that is not represented in the orthography (Schachter and Otanes 1972). Nevertheless, other languages with unambiguously V-initial roots, like Chamorro (Guam, Northern Mariana Islands) and Toba Batak, confirm this generalization.

v Typically AN CVC-reduplication is analyzed as involving length, not stress.

vi See Benton 1971 on a notable exception in Pangasinan (Luzon). In this language, the first CV is reduplicated even if the stem is V-initial, yielding an infixed reduplicant, as in amigo ‘friend’: a<m>mi>migo ‘friends’.

vii It is rare that both CV- and Ca-reduplication apply to the same set of lexical bases in a given language (but see Chang 1998 and Blust 2001 on Thao (Taiwan) and Li and Tsuchida 2001 on Pazeh (Taiwan)). Therefore, it is possible that Ca-reduplication and CV-reduplication might be variants of one another.

viii In recent literature, a distinction has been made between phonological and syntactic clitics (see e.g. Zwicky and Pullum 1983, Zwicky 1985, Anagnostopoulou 2003, Preminger 2009, Nevins 2011, Kramer 2014, Harizanov 2014). For our present purposes, we restrict our discussion to phonological clitics.

ix The ‘voice system’ of Western AN languages is also referred to as the ‘focus system’, among other terms. See Blust (2002, 2013) and Ross and Teng (2005) for an overview of terminological use in the literature.

x See Section 2.4 for additional discussion of cliticization in AN.

xi Regardless of the treatment of ang and ng, and related markers in other Western AN languages, these languages also display case-markers for oblique elements.

xii The complete absence of possessive markers is unusual, and such languages are rare; Toqabaqita (Solomon Islands) is a language with no such marking (Lichtenberk 2008).

xiii It seems that the more articulated Micronesian classification is subject to attrition. For example, Benton and Dyen recognize over two dozen classificatory expressions in traditional Chuukese. However, in her fieldwork with Chuukese conducted in 2000, the second author of this chapter found only the following: general, inanimate mobile, inanimate able to grow, small/intimate, drinkable, edible (raw), edible (cooked), animate female, and animate male.

xiv The form of the possessive morpheme is phonologically conditioned (Ozanne-Rivierre 1976: 149); the allomorph -n appears after the long e and after the long/short i, and unless the consonant preceding i is palatalized, the allomorph -n appears elsewhere.