CHAPTER 12

Morphology of Formosan Languages

Elizabeth Zeitoun

12.1 Introduction

The present chapter provides a succinct and selective overview of the (synchronic) morphology of the Austronesian languages of Taiwan. The Formosan languages are agglutinative and exhibit rich morphology, with affixation and reduplication being two major morphological processes. While there are no more than 20 grammatical affixes, most Formosan languages feature a large number of lexical affixes that encode different meanings that may or may not change the lexical category of the base to which they attach. In particular, they may encode adverbial concepts, as in Mantauran Rukai *mata-sialalra-ae* 'listen well' (< *mata-...-ae* 'certainly, absolutely') (Zeitoun 2007, p. 184). Reduplication

is not always "a morphological process relating the base form of a morpheme or stem to a derived form that may be analysed as being constructed from the base via affixation (or infixation) of phonemic material *which is necessarily identical in whole or in part to the phonemic content of the base form*." (italics my emphasis) as assumed by Marantz (1982, p. 437) and may pose challenges to morphological theories. Readers are referred to the descriptions of particular languages elsewhere in this handbook.

Most of this chapter is dedicated to a discussion of morphological units (§12.2) and morphological processes (§12.3). Other topics, such as nominal (§12.4) and verbal morphology (§12.5), as well as lexical categories (§12.6) are mentioned in passing, as they are being covered in more detail in other chapters contained in this handbook.

The position taken here is that the segmentation of words into smaller units offers advantages for our understanding of the Formosan languages, which, as mentioned above, are predominantly agglutinative. Most notably, it allows us to understand the mechanisms that underlie the formation of words further reflected in their behavior. Thus, a morpheme-based morphological approach is adopted for practical reasons. It views the morpheme as the most basic meaningful unit of a language, following Bloomfield (1933) and Hockett (1954), both advocates of the Item-and-Arrangement (IA) analysis. This is also the view assumed in most studies, even those that do not mention their precise theoretical background. Starosta (2003), as a strong supporter of "seamless morphology", is the sole author in Formosan linguistics to have argued against such a structuralist approach, suggesting that words (including compounds) are related to each other through pairwise dependency relations and form-meaning analogies. Despite his major contribution to our understanding of the morphosyntactic typology of the Formosan languages, I will not follow his analysis here.

12.2 Morphological Units

Morphological units include morphemes and their allomorphs (§12.2.1), roots, stems (§12.2.2), words (§12.2.3), and affixes, and clitics (§12.2.4).

12.2.1 *Morphemes and Their Allomorphs*

Morphemes represent the minimal meaningful units of a language and are divided into two categories, bound and free, which further subsume lexical vs. grammatical morphemes. There is no one-to-one correspondence between these two classifications. As shown in Figure 12.1., lexical morphemes include

roots and affixes, while grammatical morphemes consist of function words, affixes, and clitics. Definitions and illustrations are given in sections 12.2.2–12.2.4.

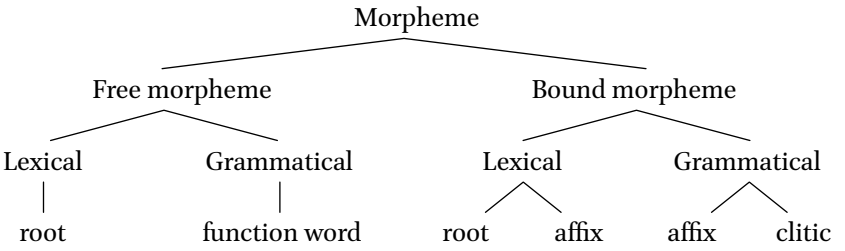


FIGURE 12.1 A classification of morphemes, roots, affixes, and clitics BASED ON ZEITOUN ET AL. 2015, P. 73

Two notes are in order. First, across the Formosan languages (and even across dialects), there is no exact correspondence between free and bound morphemes, and it is important to understand that different factors (phonological, morphological, etc.) must be taken into consideration to determine the (in)dependent status of each morpheme (see §12.2.1.2). Second, lexical and grammatical morphemes share some phonotactic and morphological properties, but again, differences are found across languages or dialects. For instance, lexical roots can undergo affixation, e.g., Squliq Atayal ***cin***-*bzyok* 'possess pig(s)' (< *bzyok* 'pig'), *s*⟨***in***⟩*ok*-***an*** 'odor' (< *sok* 'smell'), and reduplication, e.g., Squliq Atayal ***l***~*lukus* 'a lot of clothes' (< *lukus* 'clothes'); or they may also be part of a compound, e.g., Squliq Atayal *mqwas biru'* 'student' (< *mqwas* 'read, sing (AV)' + *biru'* 'write'). On the one hand, function words are not generally subject to any of these morphological processes, and if they are, it is only under very restricted conditions. On the other hand, they can undergo cliticization, e.g., Squliq Atayal ***nyux***=*saku' m-aras ruku'* [PROG=1SG.NOM AV-bring umbrella] 'I am taking my umbrella' (L. Huang 2022).

Morphemes exhibit different types of allomorphs: phonological, grammatical, and lexical. These allomorphs can be easily identified across the Formosan languages and are briefly discussed in turn below.

Most phonological allomorphs include voice affixes, which may exhibit different forms according to the kind of root to which they attach. In Kuljaljau Paiwan, for instance, AV-marked verbs are most commonly formed by the infix ⟨*em*⟩, e.g., *k*⟨***em***⟩*esa* 'cook (AV)', *dj*⟨***em***⟩*ekec* 'make adhere, alight (AV)', *q*⟨***em***⟩*aung* 'be distressed/weep (AV)'; its conditioned allomorphs include ⟨*en*⟩

after a labial initial (*p*, *b*, *v*, *m*), e.g., *b⟨**en**⟩usbus* 'drizzle, spray (AV)',[1] *v⟨**en**⟩eli* 'buy'; *me-* with a root whose initial segment is *ng*, e.g., ***me**-nganga* 'molest (AV)', ***me**-ngidju* 'give off heat (AV)' and *m-*, before a vowel-initial stem, e.g., ***m**-edjek* 'burn sth/be aflame (AV)', ***m**-alap* 'take (AV)' (Ferrell & Tjakisuvung forthcoming). In some languages, such as Pazeh-Kaxabu, Kanakanavu, and Saisiyat, voice affixes may assimilate to the first or last vowel of the base, or less commonly to the first consonant. In Pazeh-Kaxabu, the prefix *mu-* attaches to roots with a vowel *u* or *a* in the first syllable of the verb, e.g., ***mu**-k**u**dung* 'hit, strike (AV)', ***mu**-b**a**xa* 'give (AV)'; *me-* to those whose first vowel is *e*, e.g., ***me**-depex* 'study (AV)'; and *mi-* to those whose first vowel is *i*, e.g., ***mi**-k**i**ta* 'see (AV)'. In Saisiyat and Kanakanavu, vowel assimilation accounts for the allomorphy of the UVP suffix. In both languages, the UVP suffix (Saisiyat *-en* and Kanakanavu *-~~u~~n*) exhibits three allomorphs that assimilate to the previous vowel; cf. Saisiyat *-en*/*-on*/*-in*, as in *shebet-**en*** 'beat (UVP)', *ko:k**o**(:)-**on*** 'shave (UVP)' and *mari'-**in*** 'take (UVP)', and Kanakanavu *-~~u~~n*/*-un*/*-in*, as in *ka~~u~~n-~~u~~n* 'eat (UV)', *patup**u**n-**un*** 'throw (UV)', *pa'apic**i**-**in*** 'cut (UV)'. Saisiyat is also prone to sibilant assimilation (Blust 1995), and the UVC prefix *shi-* has the allomorph *si-* when the root/stem starts with an interdental fricative (Zeitoun et al. 2015); cf. ***shi**-ngoip* 'forget (UVC)' and ***si**-s**i**'ael* 'eat (UVC)'.

The stative prefix *ma-* and its alternate form *ka-*, which are found across many Formosan languages, can be treated as grammatical allomorphs. Stative verbs are marked by *ma-* or Ø, but in their root forms (when marked as causative, irrealis, and reciprocal, for instance), they are prefixed by *ka-* (Zeitoun & Huang 2000)

(1) Pazeh (Blust 1999, pp. 347–348)
***ma**-ngesen* 'afraid'
*pa-**ka**-ngesen* 'to frighten'

(2) Mantauran Rukai (Zeitoun 2007, p. 201)
***ma**-poli* 'be white'
*pa-**ka**-poli* 'to whiten (lit. to make white)'

An illustration of a lexically conditioned allomorph is found in Mantauran Rukai. In this dialect, the plural is marked on [+human] common nouns through the prefixation of *a-* to the base, e.g., ***a**-tamatama* 'middle-aged men'

1 When the vowel following the infix is *e*, that *e* is commonly deleted, e.g., *k⟨em⟩eljang ~ k⟨em⟩ljang* 'understand', *p⟨en⟩esis ~ p⟨en⟩sis* 'pull out hair'.

(< *tamatama* 'middle-aged man'), ***a-valrovalro*** 'young women' (< *valrovalro* 'young woman, maiden'). With the noun *lalake* '(one's) child', the plural is not the expected form ***a-lalake* but rather *la⟨**a**⟩lake* '(own) children', *a-* and ⟨*a*⟩ 'PL' being lexically conditioned allomorphs. It can be hypothesized that *lalake* 'child' is the reduplicated form of PAN *aNak and this explains the insertion of ⟨*a*⟩ rather than its prefixation.

Fewer than a dozen cases of suppletion have been reported in Formosan languages (see Zeitoun 2007, Zeitoun et al. 2015). One involves the complementary distribution in Saisiyat of the verbs *rima'* and *'osha'* 'to go'. The former occurs only in indicative/affirmative clauses and the latter only in non-indicative/negative clauses (Zeitoun et al. 2015, pp. 75–76). Compare (3a–c):

(3) Tungho Saisiyat (based on Zeitoun et al. 2015, pp. 75–76)

a. *yako* ***rima'/*'osha'*** *lamsong.*
1SG.NOM go[AV] Nanchuang
'I go/went to Nanchuang.'

b. *yako* *'okay* ***'osha'/*rima'*** *lamsong.*
1SG.NOM NEG:LNK go Nanchuang
'I did not go to Nanchuang.'

c. ***'osha'/*rima'*** *lamsong!*
go:IMP.AV Nanchuang
'Go to Nanchuang!'

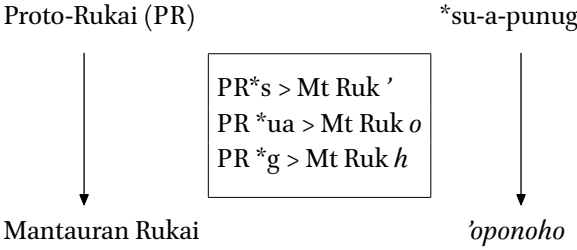
12.2.2 *Roots and Stems*

A root consists of a single and synchronically underived morpheme (i.e., containing no derivational morphemes), e.g., Paiwan *maca*, Tsou *mcoo* 'eye' (< PAN *maCa 'eye').[2] It is important to stress the notion of "synchronicity": In Mantauran Rukai, the verb *kane* 'eat' (< PAN *kaen 'eat') is a root just like *'oponoho* 'Mantauran (people/village)' or *valrovalro* 'young woman, maiden'. However, it is clear that while *kane* cannot be further divided, *valrovalro* represents the lexicalization of a previously reduplicated but no longer identifiable root, cf. **valro*. In the case of *'oponoho*, things are more complicated. Synchronically speaking, *'oponoho* is an indivisible morpheme that Mantauran speakers are unable to segment; as such, it can be treated as a root. However, it is possible to

2 Our usage of "root" here follows the conventional definition and differs from that of Blust (2013, p. 360), who defines a "root" as referring to a unit that cannot occur in isolation and recurrently exhibits "submorphemic form-meaning associations".

identify each of its components by comparing it with forms found in the other Rukai dialects,[3] e.g., Budai Rukai *su-a-punugu*, Tanan Rukai *so-a-'onogo* [from-REAL-place.name] lit. 'from Punugu/Ponogo', which can be reconstructed to Proto-Rukai, knowing that Mantauran Rukai has undergone a number of sound changes, as shown in (4):

(4)



Thus, while Mantauran Rukai *'oponoho* is analyzable on comparative grounds as meaning 'from Ponogo', this analysis is not possible based on the synchronic Mantauran Rukai data alone. Thus, in this dialect, it is a synchronically indivisible root.

As shown in Figure 12.1, roots are divided into free and bound roots. It is difficult to decide a priori how to identify bound vs. free roots across the Formosan languages as a certain lexical root might be a free morpheme in one language but bound in another. Three examples will be given here to illustrate this point. First, in some Formosan languages, bare (dynamic) verbs are free roots. For instance, in Tona Rukai, root form *kane* is found after the conjunction *la* 'and'. Similarly, Thao *kan* and Saisiyat *si'ael* 'eat' occur as bare/root forms in AV imperative or negative clauses, e.g., Thao *kan afu!* 'Eat your rice!' (said to a child who will not eat) (Blust 2003, p. 444), Saisiyat *si'ael!* 'Eat!'. In other languages, such as Isbukun Bunun, Kanakanavu, and Tsou, dynamic verbs are bound roots. Thus, Isbukun Bunun ***kaun*, Kanakanavu ***kaun(ʉ)*, Tsou ***anʉ* are never unaffixed; these verbs must always be marked for voice, e.g., Isbukun Bunun *maun* 'eat (AV)', *kaun-un* 'eat (UVP)' (Li 2018), Kanakanavu *k⟨um⟩a~kaun(ʉ)* 'eat (AV.IPFV)', *kaun-ʉn* 'eat (UVP)', Tsou *bonʉ* 'eat (AV)', *an-a* 'eat (UVP)'. Second, in many Formosan languages, stative verbs are bound roots, e.g., Tona Rukai |agi'i| vs. *ma-agi'i* '(to be) good, fine', Thao |qitan| vs. *ma-qitan* 'good, beautiful', Saaroa |vacangʉ| vs. *ma-vacangʉ* 'good'. In Kavalan and Isbukun Bunun, however, stative verbs are free roots, cf. Kavalan *nngi* 'good', *mnet* 'spicy, hot (flavour)', as in

3 With the exception of Maga, there is no distinction between *u* and *o* in the Rukai dialects, and orthographical differences reflect speakers' pronunciation.

mnet anem=na [spicy mind=3SG.GEN] 'He/She is wicked' (Li & Tsuchida 2006, p. 208), Isbukun Bunun *sial* 'goodness' vs. *ma-sial* 'good', *davus* 'wine' vs. *ma-davus* 'sweet'. Third, in most Formosan languages, kinship terms are free roots, e.g., Saaroa *ina'a*, Paiwan *kina*, Kanakanavu *cina*, Isbukun Bunun *tina* 'mother'; however, in most Rukai dialects (with the exception of Maga and Tona), kinship terms are bound morphemes, cf. Mantauran and Budai Rukai |*ina*|, Tanan Rukai |tina| 'mother', and must be suffixed by a genitive pronoun, as in Mantauran Rukai *ina=li* and Tanan Rukai *tina=li* 'my mother'.

A stem consists of a root having undergone a derivational process (affixation or reduplication), but it may be identical in form to the root. Taking again the examples mentioned above, Isbukun Bunun *maun* 'eat (AV)', *kaun-un* 'eat (UVP)' (L. Li 2018), Kanakanavu *k⟨um⟩a~kaʉn(ʉ)* 'eat (AV.IPFV)', *kaʉn-ʉn* 'eat (UVP)', and Tsou *bonʉ* 'eat (AV)', *an-a* 'eat (UVP)' all represent stems. Stem modification is scarcely mentioned, but an interesting case can be illustrated with the alternation of PAN *d with *p (and the reflexes of these two phonemes in the modern languages) in the numeral 'two' when it is suffixed with *-N. Compare PAN *duSa 'two' and PAN *ma-puSa-N 'twenty' (< PAN *ma- 'tens', *puSa- 'two (bound form)', *N 'recurrence') (Zeitoun, Teng & Ferrell 2010).

12.2.3 *Words*

Words consist of a stem, which may be simple (viz. a root), e.g., Mantauran Rukai *pato'o* 'tell', Thao *taun* 'house', or complex (viz. a reduplicated or affixed root), e.g., Puyuma *kualeng-an* 'illness', Saisiyat *ma-ka~k-si'ael* [AV-RED-eat~eat] 'marry' (lit. 'eat lunch together'). Phonological words represent the domain of stress assignment and are subject to different phonological processes. Morphological words are composed of a phonological word with one (or more) clitic(s). In some languages, clitics are unstressed (e.g., Rukai, Puyuma, Kanakanavu), while in others, they can be stressed (e.g., Saisiyat).

On the phonological level, the Formosan languages are subject to a minimal word constraint, but differ cross-linguistically. Some languages (e.g., Takibakha Bunun, Paiwan, or Atayal) exhibit monosyllabic roots, which generally consist of a (phonetically) long vowel (i.e., the syllable must be bimoraic), e.g., Takibakha Bunun *sak* [sa:k] 'smell', *kan* [ka:n] 'fish'. If this root undergoes affixation, the vowel becomes short, cf. Takibakha Bunun *sakun* [sakun] 'smell (UVP)'. In most languages, including Pazeh-Kaxabu, Saisiyat, Rukai, and Puyuma, the roots of morphological (or content) words are disyllabic, e.g., Saisiyat *talek* 'cook', Katripul Puyuma *dawa* 'millet', Tanan Rukai *'ongolo* 'drink'. In Kanakanavu, words are trimoraic, e.g., *vantuku* 'money', *sarone* 'man, male'. If a base is disyllabic, the vowel of the first syllable will be long, and that of the second short; when it undergoes affixation, or if it is cliticized, the first vowel

is short. Compare *manu* [má:nu] 'child' with *manu-in* [manúin] 'his/her/their child' and *manu=isi* [manúisi] 'this child'. In Saaroa, lexical words are trisyllabic,[4] e.g., *vatu'u* 'stone', *ama'a* 'father', but when derived, the base loses its final syllable, cf. *ama=isa* 'his/her father'.

In all the Formosan languages, grammatical words tend to be monosyllabic (e.g., Puljetji Paiwan *ni* 'GEN', Saisiyat *ki* 'COM', Tona Rukai *la* 'CONJ'), but some may be disyllabic (e.g., Mantauran Rukai *mani* 'then (CONJ)'), or even trisyllabic if they represent the combination of different morphemes (sometimes non-identifiable), e.g., Kanakanavu *nakai* 'but (CONJ)', Thao *matsahay* 'COM' (< *mat* 'CONJ', *sahay* 'here'), Mantauran Rukai *alakai* 'because' (< *=i* '3SG.GEN').

12.2.4 *Affixes and Clitics*

Affixes form a class of monosyllabic and disyllabic morphemes morphologically and phonologically bound to their hosts. Phonologically, they may either undergo processes such as assimilation, dissimilation, resyllabification, and vowel deletion or induce morphophonemic alternations of the base to which they attach. As shown above, in Kanakanavu, the UV suffix *-ʉn*, as in *kaʉn-ʉn* 'eat (UV)' (< |*kaʉn*| 'eat'), becomes *-in* when it attaches to a root with *i* in the penultimate syllable, as in *arivʉ-in* 'hold (UV)' (< |*arivi*| 'hold'). Dissimilation is found in Tsou, with the change of the glottal fricative *h* to a velar stop *k* when adjacent to *s* (Tung 1964, Tsuchida 1976, Li 1977). Compare, for instance, *s⟨m⟩oh'o* 'hatch (AV)' and *sko'-a* 'hatch (UVP)'. Vowel deletion is observed in different environments. In Tkdaya Seediq, it shows up in bases that undergo suffixation, such as *adis* 'bring (IMP.AV)' vs. *des-i* 'bring (IMP.UVP)' (Li 1977, p. 401); in Saisiyat, however, it is induced by prefixation in examples like the following: *'a-s⟨m⟩i'ael* 'be eating/will eat (AV.PROG/AV.IRR)' (< *s⟨om⟩i'ael* 'eat (AV)'). In Mantauran Rukai, the suffixation of the imperative suffix *-a* triggers the alternation between *o* and *lr*, e.g., *maavanao* 'bathe' vs. *maavanalr-a* 'bathe!' (see Li 1977). Morphologically, affixes become part of the word to which they attach. This was shown above with the Kanakanavu example *manu-in* [manúin] 'his/her/their child'.

The number of affixes range from only about two dozen in Kaxabu to over three hundred in Saisiyat. They include prefixes, infixes, suffixes, and circumfixes; in most languages, prefixes significantly outnumber the other types of affixes. There are only two productive infixes, AV ⟨*um*⟩, PFV ⟨*in*⟩ (and their respective variants), and many languages preserve the reflexes of the PAN fos-

4 Pan (2012, p. 32) mentions that disyllabic content words are rare.

silized infixes *⟨al⟩, *⟨aR⟩, and *⟨aN⟩ (see Li & Tsuchida 2009). Infixes usually occur after the first consonant of the stem, e.g., Mayrinax Atayal *c⟨um⟩aping* 'sweep (AV)', Puljetji Paiwan *c⟨em⟩avu* 'wrap (AV)', Isbukun Bunun *m⟨**in**⟩aun* 'ate (AV:PFV)', Saisiyat *s⟨**in**⟩i'ael* 'ate (PFV.UVP)'. Contrary to expectation, however, the infix ⟨*in*⟩ occurs after a cluster of prefixes made up of the realis *ka-* and the UVC prefix *shi-* (or the allomorph *si-*) in Saisiyat, as in *ka-sh-⟨in⟩*, e.g., *ka-sh-r⟨**in**⟩akep* 'grabbed (PFV.UVC)', *ka-s-s⟨**in**⟩apoeh* 'swept (PFV.UVC)'.

Affixes consist of grammatical affixes and lexical affixes. Grammatical affixes are rather restricted in terms of inventory (about twenty can be identified in each language) but carry a full range of functions. They may indicate verb class, voice, mood, aspect, modality, and/or plurality; they may induce valency change, verbalization, or nominalization. A few are found more specifically with numerals, for instance, the reflexes (*-l* and *-n*) of the recurrence suffix reconstructed as PAN *N (e.g., Budai Rukai *ma-pusa-le* 'twenty' < *drusa* 'two'). Some exclusively attach to verbs and others to nouns. Lexical affixes (mostly prefixes) provide an additional meaning and may combine with grammatical affixes.

One affix found in some Formosan languages is *na-* 'DIR', which never attaches on its own to a base but co-occurs with other locative, orientational, and directional affixes, e.g., Thao *maku-**na**-sahay* [AV:go-DIR-over.there] 'go over there' (Blust 2003, p. 110), *m-u-**na**-faw* [AV-go-DIR-up] 'go up, as a squirrel running up a branch to escape' (ibid., p. 38); Saisiyat *ka-sh-**na**-koraeh* [walk(.along.a.trail)-step.on-DIR-stride.over] 'cross a river', *ma-sh-**na**-'abe'* [AV-fall.down-DIR-a.lot] 'rain a lot'.

Determining the function of a specific grammatical affix or pinning down the meaning of a lexical prefix is sometimes complicated by polysemy and homonymy, which may occur within the same language or across different languages. In Nanwang Puyuma, for instance, the prefix *i-* is polysemous: it is used to express diverse meanings such as wearing (e.g., *m-i-kabung* [AV-wear-hat] 'wear a hat'), possession (e.g., *m-i-paisu* [AV-have-money] 'have money'), instrumentality (e.g., *m-i-pitaw* [AV-use-hoe] 'use a hoe'), and existence (e.g., *m-i-riwanes* [AV-exist-rainbow] 'there is a rainbow') (Teng 2014). Grammatical affixes might carry portmanteau functions. The infix ⟨*in*⟩ encodes perfectivity, as in Saisiyat *s⟨om⟩⟨**in**⟩i'ael* 'ate (AV.PFV)', but it can also serve simultaneously as perfective and either undergoer voice or nominalization (e.g., *s⟨**in**⟩i'ael* 'what was eaten (UVP.PFV)/ food (PFV:PAT.NMLZ)'). Across languages or dialects, it may carry many diverse functions or just one, and cross-linguistic comparison might be flawed if these functions are not properly identified. Teng & Zeitoun (2016), for instance, show that what had been analyzed as different UV voices in Saaroa are actually nominalizations, the infix ⟨*in*⟩ occur-

ring only in nominal derivations in this language. Within the same language, many grammatical and lexical affixes are homophonous, and it is sometimes difficult to identify the function and/or meaning of each of these prefixes. Consider the occurrence of the different *k-* prefixes in Saisiyat (5) and Atayal (6):

(5) Tungho Saisiyat (Zeitoun et al. 2015, p. 85)

a. *k-* 'stative'

*pa-**k**-bain*	'make lazy'	< *bain* 'lazy', *pa-* 'CAUS'

b. *k-* 'eat'

*pa-**k**-si'ael*	'make … eat lunch'	< *k⟨om⟩si'ael* 'eat lunch (AV)'
vs. *pa-si'ael*	'make … eat, feed'	< *s⟨om⟩si'ael* 'eat (AV)'

c. *k-* 'crush'

*pa-**k**-tel*	'make … pluck/cut'	< *k⟨om⟩tel* 'pluck/cut (AV)'

d. *k-* 'walk'

*pa-**k**-lobih*	'make … return'	< *lobih* 'return', *lo-* 'walk', |bih| 'turn back'

(6) Squliq Atayal (Huang & Hayung 2018, p. 28)

a. *k-* 'stative'

***k**-slaq*	'muddy'	< *slaq* 'mud'

b. *k-* 'grow'

***k**-phpah*	'to blossom'	< *phpah* 'flower'

c. *k-* 'wear'

***k**-yamil*	'wear shoes'	< *yamil* 'shoes'

d. *k-* 'toward'

***k**-suruw*	'go backward'	< *suruw* 'behind'

e. *k-* 'come from'

***k**-'ulay*	'come from Wulai'	< *'ulay* 'Wulai'

Co-occurrence restrictions with voice affixes (and in particular AV affixes) may allow us to distinguish different homophonous prefixes (Zeitoun et al. 2015, Wu 2008). Compare, for instance, Saisiyat *ki-pazay* 'harvest rice' vs. **k⟨om⟩i-pazay* (*ki-* 'harvest, gather'), and *ki-'oer* 'dig foundations' vs. *k⟨om⟩i-'oer* 'dig foundations (AV)' (< *ki-* 'to dig') (Zeitoun et al. 2015, p. 86).

Clitics are phonologically bound but grammatically free morphemes. In contrast with affixes, they are not selective of their hosts and do not induce the phonological changes mentioned above. In Saisiyat, for instance, while the prefix *no-* 'INST.NMLZ/UVC' yields vowel harmony, deletion of the vowel ⟨*o*⟩ in the ⟨*om*⟩ infix, and resyllabification, e.g., ***noe-h⟨m⟩**iwae'* [nœh.mi.wæʔ] 'use

to saw, saw (n.)' (< *h⟨oem⟩iwae'* 'to saw (AV)'), its clitic counterpart *nom=* 'INST.NMLZ/UVC' does not, cf. ***nom=h⟨oem⟩iwae'*** [nom.hœ.mi.wæʔ] 'use to saw, saw (n.)'.

The Formosan languages exhibit very few clitics, usually no more than twenty. Clitics include proclitics, enclitics, and circumclitics. Proclitics precede the base, as in Nanwang Puyuma ***ku=****salretrag-ay*... [1SG.GEN=pour-UVL] 'I poured ... (UVL)', while enclitics follow it, as in Nanwang Puyuma *me-lra~lriputr=****ku*** [AV-RED~wrap=1SG.NOM] 'I am wrapping ... (AV.PROG)'. Circumclitics are found scarcely and consist of two indivisible morphemes preceding and following the root/stem, e.g., the possessive *'inoka=...=****a*** in Tungho Saisiyat, as in ***'inoka****=ripon=****a*** *kinaat* 'Japanese book' (Zeitoun et al. 2015, p. 91).

Clitics carry various grammatical functions: they include pronouns; demonstratives; negators; and also markers of mood, aspect, and evidentiality. They differ from affixes in that they attach to phrases rather than roots/stems. Different types of clitics can be distinguished: phrasal clitics (7a), head-adjacent clitics (Lin 1996, Li 2010), as in (7b), and second-position clitics, as in (7c).

(7) Kavalan (H. Chang & Lee 2002, p. 354)
a. [*m-ringi tu repaw*]=***ay*** *wasu*
AV-look.after OBL house=REL dog
'the dog that looks after the house'

Isbukun Bunun (Lin 1996, p. 38)
b. ***na***=[*haiap saikin tu ku-sain-tin saia*]
IRR=know 1SG.NOM LNK come-here-DEM.PROX 3SG.NOM
'I know that he/she will come.' (rather than *'I will know that he/she will come.')

Nanwang Puyuma (Teng 2008, p. 98)
c. *m-uka=****mu*** *m-utrangi-a kan temuu i,...*
AV=go=2PL.NOM AV-visit-PROJ OBL.SG your.grandparent TOP
'When you go to visit your grandmother, ...'

Though the notion of "second-position clitic" is important to characterize nominative pronouns in Formosan languages, it must be stressed that this concept is actually relative and changing because each language has its own morphosyntactic system. In Puyuma, for instance, nominative pronouns attach to the negator in negative AV clauses, as shown in (8a). Hence, they can be treated as second-position clitics. In negative UV clauses, however, nominative pronouns

never attach to the negator. Rather they follow the verb that appears after the negator; the aspectual marker can encliticize to the negator, as in (8b).

(8) Nanwang Puyuma (Teng 2008, p. 33)
a. *adri=**ku**=driya t⟨em⟩alam m-u-isatr dra sasudang.*
NEG=1SG.NOM=IPFV ⟨AV⟩try AV-go-up OBL.IND boat
'I have never got on a boat.'

b. *adri=**la** tu=pa-drua-i=**ku** kantu ruma'.*
NEG=PFV 3.GEN=CAUS-come-UVL=1SG.NOM OBL.DEF/3.PSR house
'He did not cause me to come to his house.'

Affixes and clitics may occur together, the former usually closer to the base than the latter, as in (9a–b), but affixes may also precede or follow clitics, as in (9c–d). In (9c), the verb is first nominalized |*kaun*| 'eat' > *na=kaun-un* 'that will be eaten, food', and then further verbalized through the prefixation of *ka-* 'make'. In (9d), the proper noun consists of a compound *'okay=a=boa:*, which is verbalized through the addition of *shin-* 'to call'. The derived verb is obligatorily marked as UVP, through the addition of the voice marker *-en*.

(9) a. Tungho Saisiyat
*tawmo' nisho' si'ael-**en**=**ay**?*
banana 2SG.GEN eat-UVP=QST
'Did you eat the banana?'

b. Nanwang Puyuma (Teng 2008, p. 69)
*m-inatray kadru **tu**=k⟨**in**⟩iedreng-an.*
AV-die there 3.PSR=⟨PFV⟩lie-LOC.NMLZ
'It died there on its bed.'

c. Isbukun Bunun (L. Li 2018, p. 82)
***ka**-[**na**=kaun-un]*
make-IRR=eat-UVP
'to make food that will be eaten'

d. Tungho Saisiyat (Zeitoun 2015, p. 91)
*yako **shin**-['**okay**=**a**=**boa(:)**]-**en**.*
1SG.NOM call-Okay=LNK=Boā-UVP
'I am called Okay Boā.'

12.3 Major Morphological Processes

There are three major morphological processes in Formosan languages, affixation (§12.3.1), reduplication (§12.3.2), and compounding (§12.3.3), the first two of which are the most productive. Incorporation has also been reported but only sporadically (§12.3.4).

12.3.1 *Affixation*

Three distinct issues related to affixation are discussed below: affix inventory and ordering (§12.3.1.1), prefix harmony (§12.3.1.2), and prefix hopping (§12.3.1.3). Deaffixation is briefly mentioned in §12.3.1.4.

12.3.2 *Affix Inventory and Ordering*

Across the Formosan languages, roots can usually take up to three affixes, though in others, up to four or five affixes can co-occur. Languages in which the maximum is three include Paiwan, Bunun, Saaroa, Kanakanavu, Amis, Rukai, and Atayal, cf. Puljetji Paiwan ***si-ka-i-****maza* [UVC-DIR-LOC-here] 'be here (UVC); reason to be here' (W. Huang 2012, p. 143), Isbukun Bunun ***pa-tin-****tua-****un*** [CAUS-suddenly-open-UVP] 'open' (L. Li 2018, p. 56), Mantauran Rukai ***ni-pakini-pa-****'ongol-****a****=iae* [CNC-all-CAUS-drink-all=1SG.OBL] 'even if all (of them) invited me to drink …', Mayrinax Atayal ***pa-ka-****isiting-****un*** [CAUS-STAT-short-UVP] 'shorten (UVP)' (L. Huang 2000, p. 384). Four affixes can co-occur on a single base in Nanwang Puyuma, cf. Nanwang Puyuma ***in-u-k-****isatr-****an*** [PFV-go-DIR-up-LOC.NMLZ] 'place to which one has gone up before' (Teng 2008, p. 31), and five in Saisiyat ***ka-sh-k⟨in⟩on-in-****awaeh* [REAL-UVC-⟨PFV⟩push-toward-open] 'opened (UVC.PFV)'.

While three to five prefixes can be found in a row in certain languages, only one or two suffixes at most can co-occur, including the reconstructed PAN recurrence suffix *N, usually followed by *-an* 'UVL/LOC.NMLZ', e.g., Saisiyat *sh⟨in⟩oe'-ha-****l-an*** 'shoot once (UVL.PFV)'.

With respect to the two infixes, the usual ordering is ⟨*um*⟩ 'AV' before ⟨*in*⟩ 'PFV', cf. Mayrinax Atayal *l⟨****um****⟩⟨****in****⟩anguy* 'swam (AV.PFV)', Saisiyat *r⟨****om****⟩⟨****in****⟩ae'oe:* 'drank (AV.PFV)'. In Kanakanavu, the perfective infix ⟨*in*⟩ precedes ⟨*um*⟩, e.g., *c⟨****in****⟩⟨****m****⟩~~u~~'~~u~~ra* 'saw (AV.PFV)', *s⟨****in****⟩⟨****m****⟩a'~~u~~* 'played (AV.PFV)'.

12.3.3 *Prefix Harmony*

As first noted by Nojima (1996) for Bunun, a particular prefix occurring on the first verb in a series of two may anticipate the semantics of the main verb, or the same prefix must occur on two (or more) serial verbs. This is a phenome-

non[5] that has been reported in Bunun (10), Siraya (11), Thao, and Tsou (Tsuchida 2000, Blust 2003, Adelaar 1997, 2004).

(10) Isbukun Bunun (Nojima 1997, pp. 17–18)
a. ***mis****-utmag* ***mis****-busuk.*[6]
burn-carelessly burn-intoxicated
'(He) carelessly became drunk.'

b. ***kis****-asu-a=s* *mabananaz=tia* ***kis****-laupa.*
stab-immediately-UVL=OBL man=that stab-stab
'Immediately after that, the man stabbed (the woman).'

(11) Siraya (Adelaar 2004)
a. ***paka****-lpux=kaw* ***paka****-kuptix* *ĭau-an-da*
ANTP-can=2SG.NOM CAUS-purify 1SG-LOC-ADVS
'[if you wish] you are able to purify me' (viii:2) (p. 339)

b. *...ru* ***hu****-bäw-aw=mau* *m-ĭt* *ta* *ăta*
when drink-new-IRR.UVP=1SG.OBL AV-drink NOM this
h*⟨m⟩****u****-lam* *ĭmumi-än...*
⟨AV⟩ANTP-with 2PL-LOC
'... when I drink it new with you ...' (xxvi:29) (p. 334)

In other Formosan languages, this phenomenon is only found sporadically, but a few instances are worth mentioning. In Rukai, the sole prefix that induces prefix harmony is the verbalizer *to-* 'make, produce, build', which must occur on the denominal verb and the cardinal numeral that follows, as in *to*-Noun *to*-Numeral:

(12) Tanan Rukai
a. ***to-a-lalak****=ako* ***to-drosa.***
produce-REAL-child=1SG.NOM produce-two
'I have two children.' (Lit. 'I gave birth to two children.')

5 Different terms have been used to refer to this phenomenon, viz. "prefix harmony" (Tsuchida 2000), "affix echoing" (Blust 2003), and "anticipating sequences" (Adelaar 2004).

6 This is a cognate form found in many Formosan languages, cf. Saisiyat *boshok* 'drunk'. It is prefixed by the stative *ma-* in many Formosan languages, cf. Tona Rukai *ma-bosoko*. In Isbukun Bunun, it is prefixed by *mis-* 'burn', which also occurs on the first verb.

b.* ***to-a-lalak***=*ako* ***drosa.***
produce-REAL-child=1SG.NOM two

In Saisiyat, when a composite numeral is followed by another verb, there is a concordance between the choice of the lexical prefix and the verb, either through the repetition of the same prefix or through the selection of a semantically close verb type. The example in (13) shows that *kish-* 'to read aloud, sing' on a denumeral verb can be followed by the verb *kishkaat* 'study' (< *kish-* 'read aloud, sing' + *k⟨om⟩aat* 'write (AV)') or *maatol* 'sing (AV)' (because of the concordance in meaning) but not with another verb (e.g., *t⟨om⟩ortoroe'* 'teach (AV)') (i.e., no form and/or meaning concordance).

(13) Tungho Saisiyat (Zeitoun et al. 2015, p. 96)

a. *yako* ***kish****-posha-l* ***kishkaat.***
1SG.NOM read.aloud-two-RECUR study/read
'I read (aloud) twice.'

b. *yako* ***kish****-posha-l* ***maatol.***
1SG.NOM read.aloud-two-RECUR AV:study/read
'I read (aloud) twice.'

c.* *yako* ***kish****-posha-l* ***t⟨om⟩ortoroe'.***
1SG.NOM read.aloud-two-RECUR ⟨AV⟩teach

Other types of affixal concordance are found; one has to do with the occurrence of the causative *pa-* on verbs following either a manipulative verb, as in (14a–b), or a stative causative verb, as in (14c), in analytic causative constructions.

(14) a. Mantauran Rukai (Zeitoun 2007, p. 240)
pa-'adhi'adhil=*ine* *aamae* *iinae*
caus-DYN.NFIN:endure=3SG.OBL own.father own.mother
pa-kelrakelrange *a-olrolai.*
CAUS-DYN.NFIN:beat PL-child
'My father forced my mother to beat the children.'

b. Nanwang Puyuma (Teng 2008, p. 172)
tu=***pasisi****-ay*=*ku* ***pa-karun.***
3.GEN=force-UVL=1SG.NOM CAUS-work[AV]
'He/She forced me to work.'

c. Isbukun Bunun (L. Li 2018, p. 354)
pi-nungsiv-un=*ku* *'ubuh=a* ***ma-pa-sabah.***
STAT.CAUS-quiet-UVP=1SG.OBL infant=NOM.DIST AV-CAUS-sleep
'I make/made that child sleep quietly.'

The other has to do with the occurrence of ⟨*in*⟩ on serial verbs, the first of which is a verb of position (note that on such verbs, ⟨*in*⟩ encodes the progressive rather than perfectivity) in Saisiyat.

(15) Tungho Saisiyat (Zeitoun et al. 2015, p. 342)
a. *korkoring* *m⟨**in**⟩aywawaak* *m⟨**in**⟩ae'rem.*
child AV⟨PROG⟩lie.down AV⟨PROG⟩sleep
'The child is lying down sleeping.'

b.* *korkoring* *m⟨**in**⟩aywawaak* *mae'rem.*
child AV⟨PROG⟩lie.down AV:sleep

c.* *korkoring* *maywawaak* *m⟨**in**⟩ae'rem.*
child AV:lie.down AV⟨PROG⟩sleep

I will not include in this discussion here the notion of "voice harmony", which is discussed by Chang in this handbook (Chapter 24).

12.3.4 *Prefix Hopping in Saisiyat*

Prefix hopping refers to the occurrence of a lexical or grammatical prefix on the preceding morpheme if it is a negator, as in (16a) and (16b), or a case marker (see Zeitoun et al. 2015, p. 95ff. for details). In (16a), the morpheme *k-*, which marks stativity but never occurs when the verb is in the indicative (i.e., affirmative declarative), cf. *raam* 'know', attaches to the negative sequence consisting of the negator *'oka'* 'NEG' merged with the ligature *'i=*, thus yielding *'okik*. In (16b), the morpheme *k-* means 'to eat', as shown in the verb form *k-⟨om⟩si'ael* 'eat lunch' (< *k-* 'eat', *si'ael* 'eat', ⟨*om*⟩ 'AV') and is attracted to the negator that precedes the verb. Prefix hopping does not happen if the prefix is preceded by another prefix, for instance the causative *pa-*, as in (16a') or the irrealis prefix *'a-* as in (16b). Saisiyat is the sole language in which this phenomenon has been reported.

(16) a. Tungho Saisiyat (Zeitoun et al. 2015)
yako *'oki**k*** ***raam*** *maatol.*
1SG.NOM NEG:LNK:STAT know AV:sing
'I do not know how to sing.'

a'. *yako* *kayzaeh* *kay=**pa-k-raam*** *hisia.*
1SG.NOM can NEG:LNK=CAUS-STAT-know 3SG.ACC
'I cannot let him/her know.' (p. 381)

b. *yako* *k⟨om⟩osha: lasia,* ***"a-k-⟨om⟩si'ael*** *rini." kikraami*
1SG.NOM ⟨AV⟩say 3PL.NOM IRR-eat-⟨AV⟩eat here however
lasia *raeiw=ila* ***'oki**k* *si'ael=ila.* *rini.* [...]
3PL.NOM leave[AV]=COS NEG:LNK:eat eat[AV]=COS here
'I said to them, "Come and eat lunch here!" However, they ran away and did not eat lunch here ...' (p. 577)

12.3.5 *Deaffixation in Kaxabu*

Deaffixation has only been reported in Kaxabu (see Chen 2016, Lim 2022, Lim & Zeitoun, this handbook, Chapter 51). It refers to a process whereby affixes behave as clitics or independent (grammatical/lexical) words; note that affixes, clitics, and free morphemes coexist at the moment in this language. The UVP suffix *-en* usually appears a clitic, as shown in (17a), though it is still used as a suffix on verbs ending in *s*, as in (17b). The causative *pa-ka-* occurring on stative verbs (18a) is also found as a free morpheme taking nominal complements, as in (18b).

(17) Kaxabu (forthcoming)
a. *sinaw='en* *naki=lia.*
wash=UVP 1SG.GEN=COS
'It has been washed by me.'

b. *kudis-en* *naki=lia.*
cut-UVP 1SG.GEN= COS
'It has been cut by me.'

(18) Kaxabu (forthcoming)
a. ***pa-ka**-tahayak!*
CAUS-STAT-tired
'Thank you!' (Lit. '(I) made (you) tired.')

b. ***paka*** *nita* *umut* *ma-ziah!*
make 1PL.INCL.GEN rice.cake STAT-cooked
'Cause our rice cake to be cooked!'

12.3.6 *Reduplication*

Reduplication is a highly productive process in Formosan languages (see Lee 2007, Lee, this handbook, Chapter 13). As mentioned in the introduction, contrary to what is assumed in general linguistics (see for instance Marantz 1982, p. 437), it consists of the (not necessarily conform) copy of the base, with a maximum of two syllables being reduplicated.

There are at least five types of reduplication. (1) Lexicalized consists of a fossilized and usually no longer identifiable base that has undergone reduplication, e.g., Isbukun Bunun *bunbun* 'banana', Saisiyat *borbor* 'jew's harp', Thao *karkar* 'to chew', Nanwang Puyuma *kamangkamang* 'big spider' (Teng 2008). (2) Simple reduplication consists of the reduplication of the base in part or in whole. Partial reduplication involves the reduplication of a light or heavy syllable or just the first consonant of a syllable and can take the shapes C-, CV-, CCV-, CGV-, CVC-, CVG-, CVV-, and C*a*- reduplication. (3) Serial reduplication refers to the reduplication of the base that has already undergone reduplication (the two patterns of reduplication are distinct), e.g., Saisiyat ***ka*~*koti*~*kotih*** 'pinch one another' (< *kotih* 'to pinch') (Zeitoun et al. 2015, p. 134), Mantauran Rukai *ma-**ta**~**tobi**~tobi* 'cry for one another' (< *tobi* 'to cry') (Zeitoun et al. 2007, pp. 61–62), Nanwang Puyuma ***ra*~*ruma*~*ruma*** 'every house' (< *ruma* 'house') (Teng 2008, p. 46), Paiwan *ma-**pa**~**pana**~panaq* 'shoot one another' (< *panaq* 'shoot'). (4) Discontinuous reduplication consists of two reduplicated noncontinguous segments from the same base, e.g., Saisiyat ***paeh***~*pae*-***hae***~*hangih* 'keep on crying together' (< *h⟨oem⟩angih* 'cry (AV)'), ***pak***~*pa*-***ka***~*kaas* 'keep on biting each other' (< *k⟨om⟩aas* 'bite (AV)'). (5) Triplication refers to the reduplication of the base twice in a unitary process (Blust 2001, 2003), e.g., Thao ***apa*~*apa*~*apa*-(*a*)*n*** 'be carried (UVL)' < *apa* 'carry' (Blust, 2003, p. 196), Mantauran Rukai ***mena*~*mena*~*menanae*** 'every day' < *menanae* 'one day' (Zeitoun 2007, p. 62).

There are two noteworthy aspects of reduplication that deserve to be mentioned. First, in some languages, affixes can participate in reduplication. In Saisiyat, CVC reduplication copies affixes contained in the base form, including grammatical affixes such as *⟨om⟩*/*ma-*/*m-* 'AV', *ma*-C*a*- (~*pa*-C*a*-) 'REC', e.g., *mang~mangoip* 'keep forgetting (AV)' (< *mangoip* 'forget (AV)'), *so~smi'ael* 'be eating (AV.PROG)' (< *s⟨om⟩i'ael* 'eat (AV)'), *pak~pa-ka~koring* 'keep on beating each other' (< *k⟨om⟩oring* 'beat (AV)'), and verbalizers (e.g., *kash-* 'step' as in ***kash***~*kash-'abo'* 'keep on entering'). Second, verbs are more prone to undergoing reduplication than nouns, some of which can never be reduplicated, for instance proper names (cf. Saisiyat *kizaw*, vs. *****kiza***~*kizaw*), family names (e.g., Mantauran Rukai *(la)pangolai* '(La)pangolai' vs. **(la)**pango***~*pangolai*), and true toponyms, cf. Mantauran Rukai *'oponoho* 'Mantauran' (vs. **'opono*~***pono***-

ho) as opposed to Tsou *la~lauya* 'village name (place of maples)' (< *lauya* 'maple') (Wright 1996, p. 56). With a few exceptions, which include numerals (e.g., Isbukun Bunun ***ta****~tini* 'one (person)') and interrogative content words (e.g., (Central) Amis ***cima****~cima* 'anybody' (< *cima* 'who'), grammatical words do not undergo reduplication.

With both nouns and verbs, the meaning that is most commonly obtained through reduplication is plurality, which subsumes quantification, collectivity and distributivity, repetition, and/or continuity (see Zeitoun & Wu 2006, Lee 2007, Lee, this handbook, Chapter 13).

12.3.7 *Compounding*

Compounds are made up of two or more words. Usually the meaning of compounds is straightforward and predictable. In some cases, the meaning of a compound is opaque, as in Mantauran Rukai *lelepe mavoroko* [bean monkey] 'green beans'. In some languages, they form one morphological word, as in Saaroa (cf. *ti='acangʉralhʉ* (rather than **ti'i 'acangʉralha* [faeces star] 'meteor')) and Kanakanavu (cf. *manu=marisinatʉ* [manumarisinatʉ] (rather than ***[ma:nu] [marisinatʉ]) 'student' (< *manu* 'child', *marisinatʉ* 'study')); in other languages, the two parts of the compounds represent two phonological words on their own, as in Mantauran Rukai (cf. *vanidho kipingi* [student clothes] '(student) uniform').

Compounding has not been widely studied in Formosan linguistics, as it raises a number of challenges. One problem is whether to include two words with an intervening ligature, e.g., Saisiyat *'aeim=**a**=pizosan* [plum=LNK =hairy]. The other concerns the distinction that needs to be made between compounds and nominal complements. They differ in that the head of the compound cannot undergo ellipsis in coordination, while that of a nominal complement can; compare (19a–b), where the head is indicated in bold. Marking of possession is not found in either of the nominal components of the compound but is indicated on the head of a nominal complement (19c–d):

(19) Mantauran Rukai

a. Compound
ovale ***kipingi*** *la* *dha'olo* ***kipingi***
hair clothes CONJ rain clothes
'a pullover and a raincoat'

a'.* *ovale* *Ø* *la* *dha'olo* ***kipingi***
hair Ø CONJ rain clothes

b. Nominal complement
vila'a=ni *dha'ane* *la* *Ø ta-ve'ek-ae*
side=3SG.GEN house CONJ Ø LOC.NMLZ-pig-LOC.NMLZ
'the side of the house and that of the pigpen'

c. Compound
ovale(*=***ni**) *kipingi*
hair=3SG.GEN clothes
'pullover'

d. Nominal complement
ta-ve'ek-ae *'adhingi*(=***ni***)
LOC.NMLZ-pig-LOC.NMLZ inside(=3SG.GEN)
'inside the pigpen'

Compounding is more productive in some languages (e.g., Bunun, Kanakanavu, Saaroa) than others (e.g., Saisiyat, Rukai, Paiwan). In some languages like Mantauran Rukai, the ordering of the two nouns in a compound can be inverted (cf. *kavale koli'i* [shoe sun] ~ *koli'i kavale* [sun shoe] 'sandals'), but in many others, the order is fixed, cf. Saisiyat *lapwar boay* [guava fruit] 'guava' (and not **boay lapwar*), Isbukun Bunun *davus bunun* [wine person] 'native wine, millet wine' (and not **bunun davus*) (L. Li 2018, p. 123), Saaroa *cucu=mi'a'a* 'store' (< *cucu'u* 'person', *mi'a'a* 'sell') (and not **mi'a'a=cucu'u*), Kanakanavu *manu=nanakʉ* 'little girl' (< *manu* 'child', *nanakʉ* 'woman, female') (rather than **nanakʉ=manu*).

Compounds can be made up of a noun and a verb, as in Kanakanavu *cau=mu'uma* [person=AV.weed] 'farmer', Isbukun Bunun *kusbai lumah* [fly house] 'plane' (L. Li 2018, p. 123) or two (simple or complex) nouns as a type of compound with at least two complex nouns, as in Mantauran Rukai (20). They mostly refer to new concepts regarding cultural material.

(20) Mantauran Rukai
ta-pa-ka-tee~telek-ae *'adhama~dhamai*
LOC.NMLZ-CAUS-STAT-RED~cold-LOC.NMLZ side.dish~RED
a-'ongol-ae
OBJ.NMLZ-drink-OBJ.NMLZ
'refrigerator' (lit. 'place where side dishes and drinks are kept cold')

They can also consist of a numeral and a noun. Compounds including the numeral 'one' occur recurrently across the Formosan languages, e.g., Saisiyat

'aehae' halapaw [one bed] 'couple',[7] Tsou *amo=coni* [father=one] 'uncle', Kanakanavu *cani=pininga* [one=space.surrounding.the.house] 'household'.[8]

12.3.8 *Incorporation*

Incorporation is a process whereby a nominal complement (or object), which is syntactically independent, is morphologically part of the verb (Baker 1988). This is not a phenomenon that has been widely discussed in Formosan languages. It has been recognized in two quite different constructions.

In the first, illustrated by Tsou (21b), the case marker that otherwise normally precedes a nominal argument, as in (21a), is absent and the noun *oko* 'child' can be treated as incorporated onto the verb.

(21) Tsou

a. *mi=ta* *eobako* ***ta/*Ø*** *oko.*
AV.REAL=3SG.NOM hit.AV OBL/*Ø child
'He is beating a child.'

b. *eaa* *Ø/*ta/*to* *oko*
have Ø/*OBL/*OBL child
'have a child'

In the second, found in Northern Paiwan (22), the case marker is incorporated into the verb phrase, which is obligatorily a verb encoding motion, location, or path (Wu 2020). Knowing, however, that affixation and compounding can be confused with each other—since some prefixes originate from verbs—it is difficult to determine whether the following two examples represent instances of compounding or affixation, and I will leave this question open for now.

(22) Paiwan (Wu 2020, p. 103)
*k⟨em⟩asi-**tjay**-Palang=aken* *a* *mantjez.*
⟨AV⟩from-OBL-Palang=1SG.NOM LNK AV.come.back
'I come back from Palang's place.'

---

7 Interestingly, in Saisiyat, affixes attach to each side of the compound, rather than on the first (or the second) noun. Compare: ***pak**-'aehae' halapaw-**en*** [CAUS:STAT-one bed-UVP] 'make (two persons become) one couple (UVP)', and not ****pak**-'aehae'-**en** halapaw-**en*** [CAUS:STAT-one-UVP bed] or ****pak**-'aehae' halapaw* [one CAUS:STAT-bed-UVP].

8 Few compounds with a numeral other than 'one' have been reported in Formosan languages so far.

12.4 Nominal Morphology

On the one hand, as far as nominal morphology is concerned, no Formosan language makes any gender distinction. On the other hand, many exhibit a common/personal noun dichotomy, which may be manifested in a number of ways, most notably the case-marking system. With a few exceptions, the Formosan languages do not inflect for case; rather, case markers usually precede the noun and encode semantic distinctions that allow us to distinguish between common nouns (e.g., 'dog', 'house', 'food') and personal nouns (including kinship terms and proper nouns). Compare for instance Paiwan *ta djulis* 'Chenopodium formosanum' and *tjay Muninung* 'for Muninung' in (23a–b).

(23) Puljetji Paiwan (based on Huang 2012)
a. *uri=v⟨en⟩ava=(a)ken* ***ta djulis.***
IRR=⟨AV⟩make.liquor=1SG.NOM OBL Chenopodium.formosanum
'I am going to make liquor from *Chenopodium formosanum*.' (p. 99)

b. *na=v⟨en⟩eli=anga=(a)ken taicu a vangavangan* ***tjay Muninung.***
PFV=⟨AV⟩buy=COS=1SG.NOM OBL.this LNK toy OBL Muninung
'I already bought this toy for Muninung.' (p. 15)

In Nanwang Puyuma, the common/personal noun dichotomy is expressed through the different word formations of cardinal numerals when they modify nouns. Cardinal numerals modifying common nouns are suffixed by *-a* and undergo CV- or CVCV-reduplication, while those modifying personal nouns are prefixed by the personal classifier *mia-*.

(24) Nanwang Puyuma (Teng 2008, p. 73)
a. ***unian*** *dra* ***pa~pat-a*** *ami dra trabukan.*
not.exist OBL.INDF RED~four-NPERS year OBL.INDF boy's house
'There is no house belonging to a four-year-old boy.'

b. *mi-walak dra* ***mia-pat*** *dra walak.*
have-child OBL.INDF PERS-four OBL.INDF child
'They have four children.'

Within common nouns, a further distinction may also be made between human and nonhuman and animate and inanimate. This is characterized in many

languages by the occurrence of different numerals and quantifiers (e.g., 'few', 'many'), as illustrated in (25) and (26). In (25a–b), this distinction is encoded with the occurrence of a different prefix on the numeral. In (26a–b), the quantifier |*kalra*| 'many' takes a different prefix (cf. *ma-* 'NHUM' vs. *ta-* 'HUM') according to the type of nominal argument to which it corresponds.

(25) a. Kavalan (Li & Tsuchida 2006)
kin-turu ***a*** *sunis=ku.*
HUM-three NOM child=1SG.GEN
'I have three children.' (p. 488)

b. *pa-qann=iku* *tu* *wasu=ku* ***u-turu****(='ay).*
CAUS-eat=1SG.GEN OBL dog=1SG.GEN NHUM-three(=REL)
'I feed my three dogs.' (p. 78)

(26) a. Tanan Rukai
d⟨o⟩are-a ***ma-kalra*** *dokolro!*
⟨SUBJ⟩make-IMP NHUM-many glutinous.cake
'Prepare many glutinous cakes!'

b. *w-a-ongo~ongolo* *bava* ***ta-kalra****.*
ACT-REAL-RED~drink wine HUM-many
'Many people are drinking.'

Plurality is expressed by two different means in Formosan languages, reduplication and affixation, depending on the language and on the type of noun. In many Formosan languages, with the exception of common nouns undergoing nominalization (e.g., 'tree' > 'forest'), plurality is primarily marked on human nouns, as in (27a–b). The second process consists of the addition of an affix on the base, as shown in (28a–b) (Zeitoun 2009).

(27) a. Puljetji Paiwan (based on W. Huang 2012)
ra~malje~maljeng *tiamadju.*
RED~elder 3PL.NEUT
'They are elders.' (p. 241)

b. *tucu* *uri=b⟨en⟩aqebaq* *ti* *kina* *ta* ***kina~kina*** *a*
today IRR=⟨AV⟩pound NOM.PN mother OBL RED~breeder LNK
vasa.
taro

'Today, mother is going to pound breeders of taro (to make material for *qavay* 'round dumplings with filling').' (p. 245)

(28) Puljetji Paiwan (based on W. Huang 2012, p. 192)
a. ***lja****-taljialep=aken.*
PL-family.name=1SG.NOM
'I am from the Taljalep family.'

Mantauran Rukai
b. ***a****-tamatama=nomi.*
PL-middle-aged.man=2PL.NOM
'You (PL) are middle-aged men.'

12.5 Verbal Morphology

The Formosan languages are characterized by highly complex verbal morphology, of which voice is one of the most important features. A distinction is generally made between actor voice (AV) and undergoer voice (UV), which may further subsume patient (UVP), locative (UVL), and circumstantial (UVC) voices. Consider (29a–d), with the verb in bold and the subject in italics. In (29a), the verb is marked with the infix ⟨*em*⟩, and the actor =*aken* is mapped to subject. In (29b), the verb is infixed with ⟨*in*⟩ 'PFV.UVP', the undergoer *madrusa a c*⟨*em*⟩*akav* 'two thieves' is the subject and the actor *nazua a kisac* 'that police officer' is a non-subject core argument marked by the genitive. In (29c), it is the theme *azua a c*⟨*em*⟩*akav* 'that thief' that is mapped to subject, and the verb is suffixed with *-an* 'UVL'. In (29d), the theme *azazua a paday* 'the rice' is the subject, and the verb is prefixed with *si-* 'UVC'.

(29) a. Puljetji Paiwan (W. Huang 2012)
*na=****k*⟨*****em*****⟩*****an*****=***aken ta demangasan.*
PFV=⟨AV⟩eat=1SG.NOM OBL goat
'I ate goat (meat).' (p. 6)

b. ***dj*⟨*****in*****⟩*****adjas*****=***anga a madrusa a c*⟨*em*⟩*akav*
⟨PFV.UVP⟩catch=COS NOM two.persons LNK ⟨AGT.NMLZ⟩steal
nazua a kisac.
GEN:that LNK police
'That police officer caught two thieves.' (p. 10)

c. ***p⟨in⟩acun-an*** *niamadju azua a c⟨em⟩akav.*
⟨PFV⟩see-UVL 3PL.GEN NOM:that LNK ⟨AGT.NMLZ⟩steal
'They saw that thief.' (p. 10)

d. ***si-ulaviq***=*anga azazua a paday.*
UVC-overflow=COS NOM:that LNK rice
'The rice overflowed.' (p. 8)

Voice morphology is undergoing attrition in moribund languages. Only three voices are found in Thao (AV, UVP, and UVL) (30) and Saaroa (AV, UVP, and UVC), and two in Kanakanavu (AV and UV) (31), Kavalan (AV and UVP/UVL) (32), and Kaxabu (AV and UV).[9]

(30) a. Thao (Blust 2003)
yaku ***k⟨m⟩an*** *lapat.*
1SG.NOM ⟨AV⟩eat guava
'I ate guava.' (p. 971)

b. *a*=*kan-in suma izay a buna.*
IRR=eat-UVP other that LNK sweet.potato
'Someone will eat that sweet potato.' (p. 444)

(31) a. Kanakanavu
ni-k⟨um⟩aʉn=*cu*=*ku kamsia.*
PFV-⟨AV⟩eat=COS=1SG.NOM candy
'I have already eaten candies.'

b. ***ni-kaʉn***=*maku kamsia.*
PFV.UVP-eat=1SG.GEN.NSA candy
'I ate (the) candy.'

9 UVL was found in one instance, with the loss of grammatical relations (and in particular, the notion of non-subject actor):
(i) Kaxabu
yaku=*a* ***aba*** *kaidisiw* ***a-ituku***=***'an.***
1SG.NEUT=LNK father be.there IPFV-sit=UVL
'Father sits there.'

(32) a. Kavalan (based on Li & Tsuchida 2006, p. 260)
q⟨m⟩an *tu kukuy a sunis a yau haw?*
⟨AV⟩eat OBL candy NOM child LNK that QST
'That child is eating candy, isn't he?'

b. *t⟨em⟩meq tu qawpiR a tina=na,*
⟨AV⟩roast OBL sweet.potato NOM mother=3SG.GEN
qannan*=na a sunis=na.*
eat:UVP =1SG.GEN NOM child=3SG.GEN
'The mother roasted sweet potatoes, and her child ate them.'

Rukai is the sole language to exhibit an active/passive voice dichotomy (see Zeitoun 2007, forthcoming).

(33) Tona Rukai
a. ***wasilape*** *kake ma'aka-koa-nga.*
ACT:REAL:search 1SG.NOM all-thing-SUP
'I look(ed) for all the things.'

b. ***kyasilape*** *koso* (*nakoa*).
PASS:REAL:search 2SG.NOM (1SG.OBL)
'You are/were looked for (by someone/me).'

Voice interacts closely with lexical and grammatical aspect, mood, modality, and negation. In many languages, voice is identical in form to nominalization, and, at times, it can be very difficult to distinguish these two mechanisms from each other. Voice goes along with verb agreement, transitivity, and valency-increasing and -decreasing devices, all of which are dealt with in other chapters herein.

12.6 Lexical Categories

Lexical categories are at the margin of morphology and syntax and are briefly covered in this section.

There are two open classes in Formosan languages, nouns and verbs. In some languages (e.g., Rukai), the distinction between what constitutes a noun and what constitutes a verb is rather clear-cut; in others, though, lexical roots are treated as categorially neutral (Bril 2017). There are well-established closed classes, e.g., pronouns, demonstratives, case markers, clausal and interclausal

elements (including topic markers, linkers, coordinators, and (co-)subordinators), and interjections. Other word classes have been more difficult to define either because of their morphosyntactic homogeneity or because of their cross-linguistic scarcity. For instance, not all languages include a well-formed set of adverbs (see P. Li 2021 for an overview of this topic). In Mantauran Rukai, for instance, adverbial concepts are expressed mostly by affixes occurring on the verb, while it is rendered by a verb in other languages (34a–b).

(34) Mantauran Rukai (Zeitoun 2007)

a. ***ma'ati****-kane~kane=lrao, ka poelre-ka=li.*
a.little-RED~eat=1SG.NOM NEG all-NEG=1SG.GEN
'I ate very little, I did not eat it all.' (p. 184)

b. *m-o-a maavanao la=ko* ***'ara****-'apece la=ko*
SUBJ-go-IMP bathe if=2SG.GEN early-sleep if=2SG.GEN
'ara*-'omaca la=ko ka-somikace.*
early-wake.up if=2SG.GEN STAT-healthy
'Go take a shower, sleep early, rise early, and you will feel refreshed.' (p. 185)

What may be defined as a noun at the root level (e.g., yesterday, next year) might actually function as an adverb on the morphosyntactic level. In Paiwan (35), as in many other languages, temporal nouns, functioning as adverbs on the syntactic level, are not case-marked.

(35) Puljetji Paiwan (W. Huang 2012, p. 127)
uri=vaik ti Kuljelje a s⟨em⟩a-cadja ***nutiav,*** *sa*
IRR-leave NOM Kuljelje LNK ⟨AV⟩go-far.away tomorrow so.that
ku=kainuli-an timadju ta djalan.
1SG.GEN=pray-SUBJ.UVC 3SG.NEUT OBL road
'Kuljelje is going far away tomorrow, so I pray for him for (his) journey.'

Particles include numerous types of functionally ill-defined elements (e.g., final particles), as in Wulai Atayal *la* 'COS' (36a), which may or may not co-occur with other elements as clitics, as in (36b), whereby the particle *la* merges with the coordinating conjunction *ru'*, as shown by the deletion of the vowel *a*.

(36) Wulai Atayal (L. Huang 1995)
a. *n-wah-an=maku'* *m-cisan* *Ulay* ***la.***
PFV-come-UVL=1SG.GEN AV-play Wulai COS
'I have visited Wulai already.' (p. 276)

b. *nanu' s-panga'=nya'* ***l=ru',*** *ras-un=nya'*
so UVC-carry.on.back=3SG.GEN COS=CONJ carry-UVP=3SG.GEN
squ' ska' hlahuy.
LOC middle forest
'Therefore, he (the bear) carried her (the woman) into the forest.' (p. 273)

Very few prepositions are found in the Formosan languages. Languages having been reported to exhibit prepositions include Bunun, Kanakanavu, Paiwan, and Saisiyat. An illustration is given in (37).

(37) Isbukun Bunun (L. Li 2018, p. 161)
a. *na='asa* *Alang=a* *tu* *ma-suhais* *sui* ***sia***
IRR=want Alang=DIST.NOM LNK AV-come.back money LOC
Abus=tia.
Abus=DIST.OBL
'Alang must return money to Abus.'

Paiwan (W. Huang 2012, p. 44)
b. *tjuruvu* *a* *pusalasaladj* ***i*** *ta* *quma.*
many.HUM NOM helpers LOC OBL field
'There are many helpers (or workers) in the field.'

Finally, there are no morphosyntactic grounds for distinguishing adjectives as a distinct class, and they are thus included among stative verbs (38), and modals have not been treated as a separate class except in Isbukun Bunun (39) (L. Li 2018).

(38) Mantauran Rukai (Zeitoun 2007, p. 136)
a. ***ma-ecelrange*** *dhona'i* *ali'i=dha.*
STAT-black that tooth=3SG.GEN
'His/Her teeth are black.'

b. ***ma-olripi****-na-ka=i* *'ina'i* *vaha=nai*
STAT-disappear-still-NEG=3SG.GEN this language=1PL.EXCL.GEN

'oponoho.
Mantauran
'Our language has not disappeared yet.'

(39) Isbukun Bunun (L. Li 2018)
a. ***mahtu*** *kasu* *ka-davus* *'aupa?*
can 2SG.NOM make-wine QST
'Can you make wine?' (p. 411)

b. ***masu*** *kasu* (*tu*) *maun tamaku.*
no.need 2SG.NOM (LNK) AV:eat cigarette
'You do not need to smoke.' (p. 461)

12.7 Conclusion

The Formosan languages exhibit rich morphology and are all the more interesting because they pose challenges to morphological theories. This chapter has only provided a selective overview of different morphological topics, concentrating on a discussion of morphological units and morphological processes, with an excursus onto nominal/verbal morphology and lexical categories.

References

Adelaar, Alexander (2004). The coming and going of "lexical prefixes" in Siraya. *Language and Linguistics* 5(1), 333–361.

Anderson, Stephen R. (2005). *Aspects of the theory of clitics*. Oxford studies in theoretical linguistics 11. Oxford: Oxford University Press.

Baker, Mark (1988). *Incorporation: A theory of grammatical function changing*. Chicago, IL: University of Chicago Press.

Bloomfield, Leonard (1933). *Language*. New York, NY: Holt and Co.

Blust, Robert (1999). Notes on Pazeh phonology and morphology. *Oceanic Linguistics* 38(2), 321–365.

Blust, Robert (2003). *Thao dictionary*. Language and Linguistics Monograph Series A5. Taipei: Institute of Linguistics, Academia Sinica.

Blust, Robert (2013). *The Austronesian languages*. Asia-Pacific Linguistics Open Access Monographs A-PL 008. Canberra: The Australian National University. [https://openresearch-repository.anu.edu.au/bitstream/1885/10191/6/Blust-2013-AustronesianLanguages.pdf].

Bril, Isabelle (2017). Roots and stems: Lexical and functional flexibility in Amis and Nêlêmwa. *Studies in Language* 41(2), 358–407.

Chang, Henry Y. (this handbook). Serial Verb Constructions and other complex constructions in Formosan languages. In Paul Jen-kuei Li, Elizabeth Zeitoun and Rik De Busser (Eds.), *Handbook of Formosan languages: The Indigenous languages of Taiwan*. Leiden: Brill.

Chang, Henry Y., and Amy Pei-jung Lee. 2002. Nominalization in Kavalan. *Language and Linguistics* 3(2), 349–368.

Chang, Laura M. (1998). Thao reduplication. *Oceanic Linguistics* 37(1), 277–297.

Chen, Victoria (2016). When synthetic meets analytic: A note on structural borrowing in Kaxabu Pazeh (Squib). *Oceanic Linguistics* 55(2), 669–677.

Ferrell, Raleigh and Drungdrung Tjakisuvung (forthcoming). *Paiwan texts*. MS.

Hockett, Charles F. (1954). Two models of grammatical description. *Word* 10, 210–234.

Hsin, Tien-hsin (2000). Aspects of Maga Rukai phonology. PhD dissertation. Storrs, CT: University of Connecticut.

Huang, Hui-chuan Jennifer (this handbook). Phonology. In Paul Jen-kuei Li, Elizabeth Zeitoun and Rik De Busser (Eds.), *Handbook of Formosan languages: The Indigenous languages of Taiwan*. Leiden: Brill.

Huang, Lillian M. (1995). The syntactic structure of Wulai and Mayrinax Atayal: A comparison. *Bulletin of the National Taiwan Normal University* 40, 261–294.

Huang, Lillian M. 黃美金 (2022). Taiyayu cilei ji qi jiaoxue 泰雅語詞類及其教學 [*Word classes and applications to language teaching in Atayal*] Taiwan nandao yuyan congshu III-2 臺灣南島語言叢書 III-2 [Series on Formosan Languages III-2]. 臺北 [Taipei]: Caituan faren yuanzhuminzu yuyan yanjiu fazhan jijinhui 財團法人原住民族語言研究發展基金會 [Foundation for the Research and Development of Indigenous Languages].

Huang, Wei-chen (a.k.a. Ng, I-sin) (2012). A study of verbal morphology in Puljetji Paiwan. MA thesis. Hsinchu: National Tsing Hua University.

Lee, Amy Pei-jung (2007). A typological study of reduplication in Formosan Languages. PhD dissertation. Colchester: University of Essex.

Lee, Amy Pei-jung (this handbook). Reduplication. In Paul Jen-kuei Li, Elizabeth Zeitoun and Rik De Busser (Eds.), *Handbook of Formosan languages: The Indigenous languages of Taiwan*. Leiden: Brill.

Li, Lilian Li-ying (2010). Clitics in Nantou Isbukun Bunun (Austronesian). MA thesis. Puli, Nantou: National Chi Nan University.

Li, Li-ying Lilian (2018). A grammar of Isbukun Bunun. PhD dissertation. Hsinchu: National Tsing Hua University.

Li, Paul Jen-kuei (1977). Morphophonemic alternations in Formosan languages. *Bulletin of the Institute of History and Philology, Academia Sinica* 48(3), 375–413.

Li, Paul Jen-kuei (2021). Adverbs in the Austronesian languages of Taiwan. *Asian Languages and Linguistics* 2(1), 80–109.

Li, Paul Jen-kuei and Shigeru Tsuchida (2006). *Kavalan dictionary*. Language and Linguistics Monograph Series A19. Taipei: Institute of Linguistics, Academia Sinica.

Li, Paul Jen-kuei and Shigeru Tsuchida (2009). Yet more PAn infixes. In Bethwyn Evans (Ed.), *Discovering history through languages: Papers in honour of Malcolm Ross*, (pp. 345–362). Pacific Linguistics 605. Canberra: The Australian National University.

Lim, Hong-sui 林鴻瑞 (2022). Gehawuyu cankao yufa: yige jidu binwei de taiwan nandao yu 噶哈巫語參考語法: 一個極度瀕危的臺灣南島語 [A grammar of Kaxabu, a moribund Formosan language]. Boshi lunwen 博士論文 [PhD dissertation]. Nantouxian pulizhen 南投縣埔里鎮 [Puli, Nantou]: Guoli jinan guoji daxue 國立暨南國際大學 [National Chi Nan University].

Lim, Hong-sui and Elizabeth Zeitoun (this handbook). Pazeh-Kaxabu. In Paul Jen-kuei Li, Elizabeth Zeitoun and Rik De Busser (Eds.), *Handbook of Formosan languages: The Indigenous languages of Taiwan*. Leiden: Brill.

Lin, Hsiu-Hsu (1996). Isbukun phonology: A study of its segments, syllable structure and phonological processes. MA thesis. Hsinchu: National Tsing Hua University.

Marantz, Alec (1982). Re reduplication. *Linguistic Inquiry* 13(3), 435–482.

Nojima, Motoyasu (1996). Lexical prefixes of Bunun verbs. *Journal of the Linguistic Society of Japan* 110, 1–27.

Pan, Chia-jung (2012). A grammar of Lha'alua: An Austronesian language of Taiwan. PhD dissertation. Cairns: James Cook University.

Starosta, Starosta (2003). Do compounds have internal structure? A seamless analysis. In Rajendra Singh & Stanley Starosta (Eds.), *Explorations in seamless morphology*, (pp. 116–147). New Delhi, Thousand Oaks and London: Sage Publications. Reprinted in Elizabeth Zeitoun (Ed.) (2009). *Formosan Linguistics: Stanley Starosta's contributions*, (pp. 139–169). Language and Linguistics Monograph Series C6-1. Taipei: Institute of Linguistics, Academia Sinica.

Teng, Stacy Fang-ching (2008). *A reference grammar of Puyuma, an Austronesian language of Taiwan*. Pacific Linguistics 595. Canberra: The Australian National University.

Teng, Stacy Fang-ching (2014). Grammaticalization of predicative possession in Nanwang Puyuma and as a basis for reconstruction in PAn. *Oceanic Linguistics* 53(1), 141–159.

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

Tsuchida, Shigeru (1976). *Reconstruction of Proto-Tsouic phonology*. Study of Languages and Cultures of Asia and Africa, Monograph Series No. 5. Tokyo: Tokyo Gaikokugo Daigaku.

Tsuchida, Shigeru (2000). Lexical prefixes and prefixes harmony in Siraya. In Videa DeGuzman and Byron Bender (Eds.), *Grammatical analysis: Morphology, syntax and*

semantics, (pp. 90–109). Oceanic Linguistics Special Publication 29. Honolulu, HI: University of Hawai'i Press.

Tsukida, Naomi (2005). Seediq. In Nikolaus Himmelmann and Alexander Adelaar (Eds.), *The Austronesian languages of Asia and Madagascar*, (pp. 291–325). Routledge Language Family Series. London and New York, NY: Routledge.

Tung, T'ung-ho, with the assistance of S.H. Wang, T.K. Kuan, T.F. Cheng and Margaret M. Yan (1964). *A descriptive study of the Tsou language, Formosa*. Institute of History and Philology Special Publications, No. 48. Taipei: Institute of History and Philology, Academia Sinica.

Wu, Chunming. 2020. Two types of noun incorporation in Northern Paiwan. In Henry Y. Chang and Hui-chuan J. Huang (Eds.), *Papers from the Austronesian Formal Linguistics Association 25: AFLA 25*, (pp. 97–117). JSEALS Special Publication No. 5. Honolulu, HI: University of Hawai'i Press.

Zeitoun, Elizabeth (2007). *A grammar of Mantauran Rukai*. Language and Linguistics Monograph Series A4–2. Taipei: Institute of Linguistics, Academia Sinica.

Zeitoun, Elizabeth (2009). Reassessing the reconstruction of plural affixes in PAn: Evidence from the Formosan languages. In Alexander Adelaar and Andrew Pawley (Eds.), *Austronesian historical linguistics and culture history: A festschrift for Bob Blust*, (pp. 359–372). Pacific Linguistics 601. Canberra: The Australian National University.

Zeitoun, Elizabeth (forthcoming). *A comparative study of Rukai*. MS.

Zeitoun, Elizabeth, Tai-hwa Chu and Lalo a Tahesh Kaybaybaw (2015). *A study of Saisiyat morphology*. Oceanic Linguistics Special Publication, No. 40. Honolulu, HI: University of Hawai'i Press.

Zeitoun, Elizabeth and Lillian M. Huang (2000). Discussion on *ka-*, an overlooked marker of verbal derivation in the Formosan languages. *Oceanic Linguistics* 39(2), 391–414.

Zeitoun, Elizabeth, Stacy F. Teng and Raleigh Ferrell (2010). Reconstruction of '2' in PAn and related issues. *Language and Linguistics* 11(4), 853–884.

Zeitoun, Elizabeth and Chen-huei Wu (2006). Reduplication in Formosan languages. In Henry Y. Chang, Lillian M. Huang and Dah-an Ho (Eds.), *Streams converging into an ocean: Festschrift in honor of Professor Paul Jen-kuei Li on his 70th birthday*, (pp. 97–142). Language and Linguistics Monograph Series W-5. Taipei: Institute of Linguistics, Academia Sinica.

Zwicky, Arnold M. (1985). Cliticization and particles. *Language* 61(2), 283–305.

Zwicky, Arnold M. and Geoffrey K. Pullum (1983). Cliticization vs. inflection: English *n't*. *Language* 59(3), 502–513.