

6-2017

Saisiyat Morphology

Daniel Kaufman
CUNY Queens College

How does access to this work benefit you? Let us know!

Follow this and additional works at: https://academicworks.cuny.edu/qc_pubs

 Part of the [East Asian Languages and Societies Commons](#), and the [Morphology Commons](#)

Recommended Citation

Kaufman, Daniel. "Saisiyat Morphology: A Review Article." *Oceanic Linguistics*, vol. 51, no. 1, 2017, pp. 278-293. Reprinted in CUNY Academic Works.

This Article is brought to you for free and open access by the Queens College at CUNY Academic Works. It has been accepted for inclusion in Publications and Research by an authorized administrator of CUNY Academic Works. For more information, please contact AcademicWorks@cuny.edu.

Review Article

Saisiyat Morphology

Daniel Kaufman

QUEENS COLLEGE, CUNY, AND ENDANGERED LANGUAGE ALLIANCE

1. OVERVIEW. Zeitoun, Chu, and kaybaybaw (2015) present an in-depth description of the morphology of the Tungho dialect of Saisiyat, a highly endangered Formosan language spoken in northwest Taiwan.¹ According to the publisher's description, the book was motivated by the fact that there were few studies on Saisiyat morphology, despite it being one of the better described Formosan languages, and that Saisiyat's morphological complexity had been overlooked in the past. The book's first author, Elizabeth Zeitoun, is a leading figure in Formosan linguistics who has made many groundbreaking contributions to the description of several Formosan languages. The second author, Tai-hwa Chu, has served as Zeitoun's Saisiyat research assistant over the last two decades; and the third author, Lalo a tahesh kaybaybaw, is also a Saisiyat speaker and teacher in the Saisiyat communities of Tungho and Nanchuang. From the outset, it is very encouraging to see the native speaker collaborators who contributed so much of their expertise to the work treated as coauthors, despite the first author having done most or all of the actual writing.

The book constitutes a tremendous contribution to our understanding of Formosan languages and, tragically, the type of research on which the work depends may no longer be possible in another ten to twenty years. While attempts at revitalization are being undertaken, the prognosis offered by the authors is grim; Saisiyat, they write, "will most certainly disappear in the long term" (7).

With regard to the data, most of the examples cited in the book appear to have been elicited by the authors targeting the various points under discussion. While it would have been optimal to include a link or reference to the primary data, this is still a common lacuna in much current descriptive work. The book presents a methodical inventory of morphological paradigms and patterns, and as can be expected in a book of this scope, there are many points that warrant further scrutiny. I highlight these points in the hope that some of the open questions can be revisited while there are still fluent speakers to aid in their resolution. For reasons of space, I only focus on the more difficult analytical aspects of the work, with a cursory review of other areas.

The book is organized into the following chapters:

- Ch. 1, the introduction, offers a succinct demographic picture of the modern communities, a review of the literature, and a brief background of thirteen of the language

1. Elizabeth Zeitoun, Tai-hwa Chu, and Lalo a tahesh kaybaybaw. 2015. *A study of Saisiyat morphology*. Oceanic Linguistics Special Publications No. 40. Honolulu: University of Hawai'i Press. xxvi + 630 pp. ISBN 978-0-8248-5042-5. \$40.00, paper.

consultants, in addition to setting out the (standard structuralist) descriptive framework to be applied.

- Ch. 2 presents a sketch of the entire grammar, a welcome addition considering the length of the entire work. The chapter also covers Saisiyat phonology, which presents many challenges of its own.
- Ch. 3 is an overview of morphological units and morphological processes, covering affixation, cliticization, compounding, and one of the more difficult to describe systems of reduplication in any Formosan language.
- Ch. 4 examines word classes in Saisiyat, a generally contested area in Austronesian languages. Among lexical categories (excluding purely functional elements), the authors argue for the existence of nouns, verbs, prepositions, adverbs, and numerals as distinct categories in Saisiyat. A practical approach is taken to identifying lexical categories in which morphological and syntactic criteria are combined when necessary.
- Ch. 5 investigates nouns, nominalizations, pronouns, and demonstratives.
- Ch. 6 examines verbal morphology and is naturally one of the most complex chapters in the book. Fortunately, the chapter opens with a good bird's eye view of the voice, mood, and aspect paradigm, which stands out for some of its unusual features.
- Ch. 7 examines negation, which is a rather rich area in Saisiyat as mood, aspect, lexical category, and the dynamic/stative distinction are all reflected by the various negative markers. The chapter does an admirable job of sorting out the many different markers and their distribution. Surprisingly, assimilated Chinese loans are discussed here for the first time in the description of negative function words. While there are not many obvious Chinese loans in the cited examples in the rest of the book, negation, for whatever reason, is a functional area that has attracted a considerable amount of Chinese input.
- Ch. 8 looks at valence-adjusting operations, specifically, the formation of causatives, reflexives, and reciprocals, including nominal reciprocal relations.
- Ch. 9 describes the various nominalization patterns of Saisiyat, including gerunds, action/state nominals, and the various argument nominals.
- Ch. 10, the final chapter, gives a comprehensive account of composite verbs, that is, verbs that are formed with a lexical prefix of the type *kish-* 'to read', *pil-* 'to cook', and so forth. The chapter represents one of the most complete descriptions of a lexical prefix system in any Formosan system—cf. Nojima (1996) for Bunun, and Kaybaybaw (2009) for Saisiyat—the further study of which may very well lead to etymologies for what have become purely functional prefixes in other branches of the family.

In the following, I take a closer look at some elements of the description with an eye toward improving on the analysis. This should not be construed as a critical review but rather a modest attempt at the kind of rigorous interrogation that all thoughtful grammatical descriptions deserve.

2. MORPHOPHONOLOGY: SYNCOPE AND REDUPLICATION.

Saisiyat shows deletion of a medial weak vowel (*e* and *o*) in the configuration VCe/oCV : for example, /koməpə/ → [kompə] 'squeeze'.² It is said that schwa is different and that

deletion in *CeCVC* stems patterns differently, but no evidence is given for any differences here with *CoCVC* stems. The general description of the deletion facts is somewhat roundabout. The authors state: “In verbs with *CeCVC* stems which cannot undergo CV-reduplication the vowel of the AV infix <*om*> cannot be deleted, cf. *k<om>pe*: ‘squeeze (AV)’ ~ *ko-k<om>pe*: ‘be squeezing (AV)’. In verbs with a CV(C) onset, as in (2.82), CV-reduplication yields the deletion of the vowel of the AV infix, cf. *k<om>ita* ‘see (AV)’ ~ *ko-k<m>ita* ‘be seeing (AV)’” (53).

This could be described more insightfully as deletion taking place when licensed by general phonotactics. Impossible forms like **'a-t<m>kaw* [ʔatmkaw] and **'a-sh<m>bet* [ʔaʃmbət] are trivially ruled out by the allowable syllable types of the language rather than the lack of reduplication, as suggested above. The more interesting problem, which is not discussed, is how vowels are chosen for deletion when deletion of either an affix vowel or root vowel yields a phonotactically permissible output. We can examine this with two words with CV-reduplication and the <*om*> infix. The difference between *kepe* and *kita*’ is that *e* (/ə/) is a weak vowel that can be targeted by syncope while *i* is not. When <*om*> is inserted, the *e* in the first syllable of *kepe* is deleted. When the infix is added to *kita*’, the environment is similar but there is no weak vowel to delete, as shown below:

(1)		NO SYNCOPE	ROOT-DELETION
	/k<om>əpə:/ →	*[koməpə:]	[kompə:]
	/k<om>itaʔ/ →	[komitaʔ]	*[komtaʔ]

After CV-reduplication of these forms, the word-medial *o* in /kokomitaʔ/ is deleted.

(2)		NO SYNCOPE	ROOT-DELETION	AFFIX-DELETION
	/ko-k<om>əpə:/ →	*[kokoməpə:]	[kokompə:]	*[kokməpə:]
	/ko-k<om>itaʔ/ →	*[kokomitaʔ]	*[kokomtaʔ]	[kokmitaʔ]

Why do we not get a parallel pattern of deletion in the two forms when *[kokməpə:] is also a phonotactically permissible output? One analysis that suggests itself is that syncope is a cyclic rule that applies after the addition of each relevant morpheme. In other words, syncope does not apply over an underlying representation such as /kokoməpə:/ but rather, it applies to /kokompə:/, after the word medial schwa has already been deleted, as shown in the following derivation with ordered rules:

(3)		/ʃəbət/ ‘hit’	/kitaʔ/ ‘see’
i.	< <i>om</i> > infixation	ʃoməbət	komitaʔ
ii.	syncope	ʃombət	—
iii.	CV-reduplication	ʃoʃombət	kokomitaʔ
iv.	syncope	—	kokmitaʔ

A far thornier problem is suggested by data where epenthesis appears to feed syncope, with each process depending on the other (55):

(4)	/mon-ʃəpat/	→ [monoʃpat],	*[monʃəpat]
	/k<om>on-kəpə:/	→ [komonokpə:],	*[komonkəpə:]

2. I cite forms directly in the book’s orthography, but I use IPA for my own phonological representations. The relevant orthography–IPA correspondences are: *e* = ə, *sh* = ʃ, *oe* = œ, ‘ = ʔ, *ng* = ŋ. I have changed bold font to italic in direct quotes, since italics are much more commonly used in citing language data.

The authors write, “When verbs with CeCVC stems are attached with prefixes of the form CVC-, a vowel is inserted after the prefix to avoid three consonants occurring in a row” (55). This is meant to explain how an underlying form like /mon-ʃəpat/ ‘forty’ surfaces as [monɔʃpat]. There is no reason, though, why a /CVC-CeCVC/ structure should yield a CCC cluster in the first place; the schwa-deleting syncope process should simply be blocked here, as it is elsewhere. Nonetheless, the authors suggest that *o* is inserted to repair an abstract form like /monʃpat/, thus deriving [monɔʃpat]. However, for the schwa to be lost from the root /ʃəpat/ ‘four’ by syncope, it must be preceded by the epenthetic vowel after /mon-/ ‘decade’:

(5) EPENTHESIS SYSCOPE
 mon-ʃəpat → monoʃpat → monɔʃpat

But the epenthesis would only be motivated by a /nʃp/ cluster that came about by syncope. This derivation, thus, resembles the mythical ouroboros serpent that eats its own tail. If these facts are correct, they pose an interesting problem of circularity for a derivational account and an economy problem for a parallel Optimality Theoretic account, as we should not find deletion of one vowel and insertion of another without improving the markedness of the output.

To add to the mystery, another form without deletion is cited as *monʃepat*, but the underlying difference is not made clear. Finally, it is unclear how syncope applies to vowels other than *e* and *o*. Some examples are given of /a/ and /i/ deletion (55), but this appears sporadic, without any rules given for its application.

Certain function words form a phonological word with following material, and, in these cases, syncope applies across words. For example, the weak vowel *e* is deleted in *tənon* ‘weave’ when preceded by negation: /ʔoki tənon/ NEG weave → [ʔokitnon]. Confusingly, this sandhi phenomenon is represented orthographically in the book as *ʔokit non*, with the initial consonant of the second element as part of the preceding word. This is motivated by the fact that the representation *tənon* violates a phonotactic constraint against such tautosyllabic clusters. But the spaces should represent morphological word boundaries rather than phonological word boundaries, on par with the other symbols in this representation (“-” for affixes and “=” for clitics).

As a result of this convention, we end up with many unseemly interlinearized glosses such as the following, where *terai* ‘blind date’ is a single root:

(6) ʔokit raiʔ-i...
 Neg:Lig:blind date blind date-Neg.UVP

Again, the facts are by no means simple, as the phonological rebracketing appears to have had real morphological consequences, as well. Stative and dynamic verbs, which may be unmarked in their bare forms, are negated differently (Yeh 2000, Zeitoun 2001). Stative predicates are negated with *ʔokik* and dynamic predicates are negated with *ʔokip*, most likely a result of a historical reanalysis that took the first consonant of a following *ka- STAT and *pa- CAUS/DYN as belonging to the preceding negation marker. The fact that the stative/dynamic distinction shows up on negation in this way, even in cases where it would not appear on a following predicate, shows that the phonological rebracketing has led to real morphological rebracketing. There is no evidence, however, that this has happened with the initial consonants of lexical roots such as the *t* of *terai* above.

Reduplication has been a rich area of research in several Formosan languages, as many of these languages display diverse and complex patterns. The book largely builds on Zeitoun and Wu's (2005) presentation of the Saisiyat facts and makes certain improvements to it. Most importantly, the authors distinguish between reduplication "structures" (underlying morphemes) and "patterns" (surface realizations). In the earlier work, Zeitoun and Wu (2005:53) state that "in Saisiyat we cannot specify the prosodic unit of a reduplicant since semantic meanings carried out by CVC-/CV-/CVV-/CV-reduplication overlap partially/totally." Each surface pattern was, thus, treated as a distinct process. In the current work, the authors state (117) that an underlying CVC-reduplication process "subsumes different patterns of reduplication, CCV-reduplication, CVV-reduplication and CV reduplication, variants defined by the forms of the roots to which they are attached." In practice, however, the authors seem to have not made the full leap to underlying representations and continue to refer to the surface patterns as if they were distinct morphemes. For example:

CVC-reduplicated dynamic verbs can either be given a continuous or repetitive interpretation [...] Verbs undergoing CCV- (3.114) or CVV-reduplication (3.115) carry a continuous interpretation. We suspect that both patterns of reduplication might actually carry out the same meaning as CVC- (and CV-) reduplication (i.e., repetition), but we have not found such examples to support this hypothesis (122–23).

It is completely clear, though, from the text that CCV reduplication comes about as a result of syncope (deletion of the weak vowels *o* and *e* in a VC__CV environment). (In this sense, the current work is a step backward from Zeitoun and Wu [2005], where CCV reduplication was not recognized as an independent pattern.) The authors also make clear that CVV comes from reduplicating a base with a geminate vowel. How, then, could the allomorphs of a reduplication morpheme, which are plainly conditioned by phonology, have different meanings? Of course, it would be ideal to have examples of the full range of meanings with each allomorph, but this must be taken as an accidental gap. The other choice is tantamount to claiming that the apparent phonological conditioning is illusory. This issue rears its head in several places in this chapter, for example, "verbs undergoing CVC-, CCV- and CVV-reduplication carry out a continuous or repetitive interpretation; nouns, on the other hand, are given a diminutive meaning (they only undergo CVC-reduplication)" (122). Again, this should not be possible if the three patterns are phonologically conditioned allomorphs. As it turns out, the identification of CCV-reduplication depends entirely on the verbal AV infix <om>, and this precludes it from occurring with nouns, by their definition. With regard to the lack of CVV on nouns, no examples are given of this type of reduplication for nominal stems with CVV-initial shape. The statement that nouns only undergo CVC-reduplication, thus, needlessly undercuts the analysis of CVC-, CCV-, and CVV- as allomorphs. There are also some misclassifications of patterns here. Under "Specific CV-reduplication," for instance, we find CVC reduplication, *tom-tomay-nonak* (130).

While there is no shortage of analytical challenges, the data seem to suggest the following underlying reduplication morphemes attaching to verbs: σ_{μ} PROGRESSIVE; $\sigma_{\mu\mu}\sigma\sigma$ REPETITIVE; yet the authors resist positing underlying morphemes as prosodic templates to

be filled by segmental material from the base despite there being good precedent for such analyses in Austronesian languages (for example, Hayes and Abad 1989 for Ilokano).

A separate category is set up for CVCV reduplication (section 3.5.2.3), which is claimed to require a particular morphological “trigger” in the form of the suffixes *-an* or *-in*. But as the only stems this is exemplified with have the shape CVCV, it is unclear how this pattern would apply to stems with an initial closed syllable.

Finally, the authors recognize a distinct type of reduplication that they term “discontinuous” (136), claiming that the reduplicated segments in forms like *pakpakakoring* below are not contiguous (glossing mine).

- (7) pak-pa-ka-koring
 σ_{H} -pa-Ca-koring
 ITER-RECP-RECP-beat
 ‘keep on beating each other’

The segment in question is *k*, which originates in the initial consonant of the root and is copied in the outermost reduplication process that yields *pak-*. But to claim that this is a distinct reduplication phenomenon is to cling too tightly to the surface. In serial derivational terms, the discontinuity is an epiphenomenon: *Ca*-reduplication applies to the root yielding *kakoring*, the *pa-* prefix is added to this stem and finally the iterative morpheme, an empty heavy syllable σ_{H} , is filled from the following material to give *pak-*. In contrast to the term “discontinuous,” each process is entirely local when viewed serially.

There is a somewhat unfortunate tendency to dismiss published examples that deviate from the authors’ data as erroneous. For instance, “the CV-reduplication sub-pattern has been wrongly described before. Yeh (2003:125) provided the following forms ... which were unfortunately not accepted and/or recognized by our informants when we were working on reduplication (cf. Zeitoun and Wu 2005:7)” (125). Returning, however, to Zeitoun and Wu’s original paper, we find a plausible reason for the discrepancy: Yeh worked with speakers of the Taai dialect, in addition to the Tungho dialect they investigated. While it is very possible that other linguists have been in error in describing the facts of Saisiyat, there must also be an affordance given to dialectal and idiolectal variation, especially in the case of endangered languages whose full range of variation we will likely never understand.

3. LEXICAL CATEGORIES. There are some basic questions regarding lexical categories specific to Austronesian languages: Can all roots take voice morphology descended from earlier Austronesian * $\langle\text{um}\rangle$ AV, * $\langle\text{en}\rangle$ PV, * $\langle\text{an}\rangle$ LV, *Si- CV? Can such words function as both predicates and arguments? Do such words differ from underived nominal stems in their distribution? By defining verbs from the outset as words that are marked for voice, modality, and aspect (139), the question of what type of roots may take such marking in the first place is largely avoided. But more substantive criteria are also presented. Crucially, it is shown that nominal roots like *timo* ‘salt’ must be verbalized by a prefix like *hin-* ‘use’ before they can take voice, modality, and aspect morphology. This is an interesting contrast with Philippine languages, which make little to no use of lexical verbalizing prefixes of this type. Possessive marking and negation are other areas in

which bare nominal stems are distinguished from voice inflected forms, the latter of which cannot take the possessive circumfix, but can take the irrealis negator *'amkay*.

In regard to verb classes, it is claimed (148) that “any type of verb, with the exception of stative verbs, can be overtly marked as AV.” But a few pages earlier (143–44), a class of roots is posited that do not take voice marking in the actor voice. This is important, as the absence of voice marking in the actor voice paradigm has consequences for the reconstruction of Austronesian verb classes as well as the aspectual paradigm (for example, Ross 2015).

The case markers are also introduced in the same chapter, and Saisiyat again presents unique challenges in this department. The marker *ka* is analyzed with three very different functions—NOMINATIVE, ACCUSATIVE, and LIGATURE—but the distinction in several examples between the ligature and accusative functions is not always clear, as in the following, where *kaehoey* is an object, but is preceded by what is glossed as the ligature (168, ex.4.60).

- (8) noe-h<m>iwa' ka kaehoey pa-'ila'ino'-on=ila?
 UVC?-<AV>saw LIG wood CAUS-go:where-UVP=CS
 Lit. ‘Where has gone what will be used to cut wood?’ / ‘Where has the saw gone?’

It seems that this usage is unrelated to the more typical ligature examples cited (for example, *hiza ka korkoring* that LIG child ‘that child’). Other examples look like they could reflect something like the case harmony found in complex noun phrases in Puyuma (Teng 2008). As this is the only section devoted to the function and distribution of the ligature, it is unfortunate that its use in the associated examples is not explained at all. To be fair, this is more of a syntactic problem than a morphological one, but the inclusion of relevant data begs for further analysis.

4. THE PREFIX *ka-*. The treatment of *ka-*, from Proto-Austronesian **ka-* (Ross 1995; Zeitoun and Huang 2000; Blust 2003), which I have argued originates from a predicate meaning ‘have’ (Kaufman 2011, 2012), poses some of the usual problems encountered with related languages. It is interpreted, for instance, as a realis marker in *ka-* *-an* nominalizations, despite lacking any observable REALIS/IRREALIS distinction in words such as ‘pipe’, ‘hunting place’, ‘place to invite for dinner’. Later we find that *ka-* is analyzed as REALIS/NEUTRAL in locative nominalizations, but the same *ka-* appears in almost every other kind of nominalization (state, instrument, agentive) as well. By the end of the book we are left with *ka-* being glossed as REALIS, IRREALIS, INSTRUMENTAL NOMINALIZATION, EVENT/ACTION NOMINALIZATION, and STATIVE. The multifunctional nature of *ka-* poses challenges in many Austronesian languages, but there seems to be an opportunity to make some generalizations over the functions cited here, especially by viewing *ka-* as an incipient marker of nominalization more generally.

5. WORD ORDER. One of the most notable characteristics of Saisiyat from an Austronesian perspective is its unusual word order, and although the book is narrowly focused on morphology, some discussion of syntax is unavoidable. Saisiyat is the only

Formosan language, as far as I can tell, where the regular position of a possessor precedes the possessum and the regular position of a genitive marked NAV agent precedes its predicate. The most plausible cause for this is the influence of Chinese, as suggested by Yeh (1991, 2000). As this is likely a more recent change, we would expect some variability in positioning, but the authors state (207) that possessors must precede the possessum (as in Chinese), contradicting an earlier statement (140) that “both nouns and verbs can be preceded or followed by genitive phrases.” The data support the latter position, with examples such as *kasnaw noka 'anhi* ‘soup GEN bamboo.shoot ‘soup of bamboo shoot’ (103). These are viewed as compounds, but it is not entirely clear what differentiates a compound from a phrasal construction with a possessor, as they contain a noun phrase with genitive marking. Note also that Ogawa and Asai’s (1935) materials from earlier Saisiyat present it as a language with more canonical Formosan word order properties: the predicate precedes the subject and the possessum precedes the possessor:

- (9) Soebae:oh ila ka tial noka kabinao.
 big INCPT NOM abdomen GEN girl
 ‘The abdomen of the girl was getting bigger.’ (Ogawa and Asai 1935:123)

Later in the book, we also find postposed genitive possessors (366, ex. 7.14b), throwing the earlier claim into question.

- (10) Sho’o ’okik kala ’oya’ nisho’ sh<om>ir’ael=ay?
 2S.NOM NEG PLUR.LOC mother 2S.GEN <AV>live=QST
 ‘Aren’t you living at your mother’s place?’

Interestingly, genitive phrases are often fronted in the modern language, but they cannot serve as predicates (208), for which the dedicated possessor circumclitic construction is used. So, while the position of these phrases appears historically innovative, other aspects of their syntax conform to general Austronesian patterns (cf. Kaufman 2009a on the ban against genitive case predicates in Philippine languages).

6. DEMONSTRATIVES. Saisiyat has a relatively large inventory of demonstratives, and these are classified by the authors according to two deictic features, [\pm visible] and [\pm far], over three distal categories taken as primitives: proximate, medial, and distal. This seems to imply that there are no less than six categories of distance (proximate [$-$ far], proximate [$+$ far], medial [$-$ far], medial [$+$ far], etc.). I have never encountered a precedent for such a highly articulated six-way deictic system, and one is naturally suspicious if these categories do not perhaps encode other evidential functions, or proximity to the hearer (as is common in Philippine languages), or perhaps even temporal functions, as has been argued for Tsou case markers (Chang 2011). The system is also uneven in a way that arouses suspicion: there is an invisible proximate, and no visible medial or distal.

This section also highlights what can be considered a more general weakness of the book, an overuse of constructed paradigmatic examples. There are four full pages devoted to lengthy constructed examples showing that each demonstrative can play the same syntactic roles, a generalization that could have been stated in a single sentence. More importantly, because the examples are constructed, we are still at a loss as to how this complex system of demonstratives actually functions in discourse.

7. **THE VERBAL PARADIGM.** Within the indicative paradigm (see table 1), we find conservative forms for the actor and patient voices but several innovations in the locative and circumstantial voices (248):

TABLE 1. INDICATIVE PARADIGM

Indicative	AV	UVP	UVL	UVC
PERFECTIVE	<i>M<in>STEM</i>	<i><in>STEM</i>	–	<i>ka-sh<in>STEM</i>
NEUTRAL	<i>M-STEM</i>	<i>STEM-en</i>	<i><in>STEM-an</i>	<i>shi-STEM</i>
IMPERFECTIVE	<i>Red-M-<in>STEM</i>	<i>Red-STEM-en</i>	<i>ka-Red-STEM-an</i>	<i>shi-Red-STEM</i>
IRREALIS	<i>'am=M-stem</i>	<i>ka-STEM-en</i>	<i>ka-STEM-an</i>	<i>Ca-STEM</i> <i>ka-STEM</i> <i>nom=<om>STEM</i> <i>no-<m>STEM</i>

For one, there is claimed to be no perfective form in the locative voice, most likely a result of this voice having been nearly eliminated from the language altogether (253). More unexpected, however, is the locative voice's use of *<in>* in the neutral form and *ka-* in the imperfective, making it seem very much like nominalizations examined in ch. 9. This is, of course, recognized by the authors, and they justify the distinction based on a diagnostic given in footnote 88 (280), namely, that only nominals can take a possessor in the true possessor case. They argue that the same form *<in>STEM-an* can function both as a verb and a noun, depending on its syntactic position. When a form like *s<in>i'ael-an* is in subject position, as in (12a) below, it must be a noun and, thus, its agent can be expressed either in the genitive or possessor case. When the same form is in predicate position, however, as in (12b), it is claimed that the agent can only be expressed in the genitive case, as shown.

- (11) a. *ma'an'inmana'a s<in>i'ael-an ray katiltatimae'an*
 1S.GEN/1S.POSS <PERF>eat-UVL LOC market place
 'The market place is where I eat.'
- b. *hini katiltatimae'an ma'an/*?'inmana'a s<in>i'ael-an*
 this market place 1S.GEN / 1S.POSS <PERF>eat-UVL
 'The market place is where I eat.'

This is an interesting diagnostic briefly introduced earlier in chapter 4 (140), and one that could have been applied more widely to the other ambiguous forms in the verb and nominalization paradigms. What is apparently missing is a diagnostic frame that could be used to identify the various inflections, most crucially, the neutral and irrealis forms. Presumably, the *<in>STEM-an* form is treated as neutral because verbal uses of a plain *STEM-an* were not attested, but, certainly, there are examples of *<in>STEM-an* being used in perfective contexts, with *<in>* being glossed as perfective, as in the following (281, ex.6.34a):

- (12) *ni 'oya' s<in>apoeh-an, 'okik kayzaeh, ma'an tashi-'izaeh-en*
 GEN mother <PERF>sweep-UVL NEG.LIG:STAT good 1S.GEN tidy-again-UVP
 'Mother swept the floor but not properly, so I tidied (the house) again.'

Perhaps what is meant by neutral is not infinitival but rather a general form that can be used in any aspect. This seems to be corroborated by many examples but is never dis-

cussed explicitly, outside of the suggestive statement (331) that, “for some reasons, most aspectual markers only surface in AV clauses.”

Another apparent Saisiyat innovation in the paradigm above is the adoption of *ka-* in the perfective of the circumstantial voice. It is not clear if the presence of *ka-* here is subject to dialectal or other sorts of variation, as Yeh (2002) has the conservative *sh<in>i-* in the same function, but this may be characteristic of the Taai dialect. The presence of *ka-* in this function leads to the somewhat egregious analysis of *ka-* as both a realis marker (in CV) and an irrealis marker (in all voices, including CV). What seems more likely is that *ka-* has been reanalyzed as a general nominalizer (a view that is well supported by its wider distribution) and that both the locative and circumstantial voices are in the process of being reanalyzed as nominalizations. The anomalous presence of *ka-* in the imperfective of the locative voice may be explained by the same logic; otherwise, we are left with a situation where *ka-* must be a marker of imperfective, realis, and irrealis in a single paradigm.

Finally, we can note that the circumstantial voice irrealis resembles a historical hodgepodge of forms where one might expect ***ka-shi-* or ***shi-ka-* on analogy with the rest of the paradigm. The implications of Ca-reduplication in this instrumental function have been discussed by Blust (1998) and commented on briefly in Kaufman (2009b:207–9). Blust points out that a reduplicative morpheme **Ca-* can be reconstructed as an instrumental nominalizer and that this undermines Starosta, Pawley, and Reid’s (1982) analysis of the circumstantial voice marker **Si-* as an instrumental nominalization reanalyzed as a verb. If Saisiyat **Ca-* forms truly belong in the verbal paradigm, then this at least shows that bona fide participant nominalizations can make their way into a verbal paradigm. On the other hand, it is not at all demonstrated that the *Ca-* form really belongs together with *shi-* as merely two aspectual inflections of the same voice. While all the *Ca-* forms have a strict instrumental interpretation, the forms with *shi-* appear to have a far wider range of uses that are very familiar from other cognates of early Austronesian **Si-*. It is a considerable deficiency of this chapter that the meaning and use of the various voice forms are barely explained. We find, for instance, the following two examples, whose use of circumstantial voice *shi-* defies explanation on the basis of anything given in the text (291, ex. 6.45c and ex. 6.45b):

- (13) a. *yako* ’am=paewhiil *ka* ’ima=’akoy ’ima=sobaeoh,
 IS.NOM IRR=choose ACC AGTNMZ=many AGTNMZ=big
shi-k-’ol’ola’an, *yako* *kayni*’
 UVC-STAT-small IS.NOM refuse
 ‘I am choosing many big ones. Small ones, I do not want them.’

- b. *hiza* *kabinao*’ *raam* *tal’izaeh*, *yao* *k<om>ita*’
 that young.woman know well-dressed IS.NOM <AV>see

shi-k’al’alak *sharara*’=ila
 UV-STAT-young like=CS

‘That young woman knows how to dress herself. I saw her and found her young. I started to like her.’

We also find constructions with the locative and circumstantial voice whose translations make no reference to a locative or circumstantial argument. For example, we would

expect the following constructions (316, ex.6.73b and 6.74b, respectively), to mean ‘I’m going to swim in it’ and ‘I’m going to swim with/for it’.

- (14) a. ma’an lalangoy-ana=ila b. ma’an lalangoy-anay=ila
 1S.GEN swim-OPT.UVL=CS 1S.GEN swim-OPT.UVC=CS
 ‘I am going to swim.’ ‘I am going to swim first.’

As noted above, the consistency of interpretations across aspectual inflections is key to determining the structure of the paradigm. If there is only some functional overlap rather than total identity between several of the irrealis and realis forms in the circumstantial voice, the analysis of the verbal paradigm would require revision. The authors note that they struggled with the inclusion of the *nom*=<*om*>STEM in the circumstantial irrealis, and this indeed seems to be a completely different beast, bearing actor voice morphology and an otherwise unidentified proclitic that is glossed tentatively as “UVC?” (undergoer voice circumstantial). Again, it is necessary to show that these innovative forms really share all the diverse functions associated with *shi-* to make the case that they are bona fide circumstantial voice markers and not narrow instrumental constructions.

Relatedly, the Saisiyat circumstantial voice has been reported to have a rather unique function among Formosan languages in serving as a nonfinite form specifically for intransitive predicates, as exemplified below (Chou 2016:195, glossing adapted).

- (15) sia s<in>wa’-en ma’an **si’-osha**
 3S.NOM <PRF>permit-PV 1S.GEN UVC-go
 ‘He was permitted to go by me.’

Unfortunately, we find no mention of this function, but certain otherwise unexplained usages in the text may have a similar source, for instance, the examples below (292, ex.6.46b⁷; 475 ex. 8.83c), where *shi-* marks intransitive verbs.

- (16) a. lasia rosha’ m<in>ae’rem ma’an **shi-p-ranaw=ila**
 3P.NOM two <PROG>AV:sleep 1S.GEN UVC-DYN-bathe=CS
 ‘They were both sleeping so I went to bathe.’
 b. ... yao bazae hilsia **shi-p-hae-hangih** ...
 1SG hear 3P.ACC UVC-DYN-RED-cry
 ‘...I heard them cry...’

Another mystery exists in the use of the patient voice hortative *-aw*. The authors state that this form, unlike the other hortative, does not actually have a typical hortative function (exhorting a first person subject, as in “Let’s...”). Rather, it is claimed without much comment to oddly require a third person subject. The evidence given for this is the following ungrammatical template (313, ex.6.66) whose intended meaning is not clear.

- (17) *yako /*yami /*’ita /*sho’o/ *moyo ta-V-aw!
 1S.NOM /1PE.NOM /1PL.NOM T/2S.NOM/ 2P.NOM HORT-V-OPT.UVP

This example must have been constructed in error, as we would not expect the notional subject of a patient voice verb to be in the nominative case. (It is in fact unclear from the following examples what the case frames for hortative verbs are and if they differ from indicative patient voice.) If the nominative case pronouns are intended to be patients, it would scarcely improve the situation as the resulting meaning would be reflexive ‘Let’s

V me/us/you?'. Furthermore, we find examples later in the book (449, ex.8.33b) that do, in fact, seem to contain a first person subject of a patient voice hortative, like:

- (18) ta-pa-baeiew-aw ka korkoring ma'an
 HORT-CAUS-buy-UVP NOM child 1S.GEN
 'I will let the child buy (something).'

Saisiyat possesses several verb classes that, to some extent, appear to be inherited from Proto-Austronesian (Ross 2015). They are identified by their morphological behavior and, in certain cases, their semantics. The authors are unsure of how to treat one distinction with relation to the actor voice verbs: "We hesitate to treat *m-* and *mo-* as allomorphs of <*om*>, but such a generalization is easily obtained in view of the (many identical) morphological patterns shared by this class of verbs" (261). Considering that *m(o)-* is only found with bilabial-initial stems, it is unclear what gives rise to this hesitation (although fn.79 suggests Yeh 2003 also resists an allomorphy account). There is ample precedent in Austronesian languages for reduction, externalization, and deletion of actor voice *-<*um*> with bilabial-initial stems (Blust 2009:371–72; Zuraw and Lu 2009). An unfortunate result of this hesitation is that the door is left open for functional and morphological differences to accrue between <*om*> and *m-*. One is given the impression, for instance, that <*in*> has a perfective function when cooccurring with <*om*> but a progressive function when cooccurring with *m-*. This may very well be the case across verb classes, but there should be no such semantic difference if *m-/mo-* is conditioned by a bilabial-initial stem. This problem is identical to the one encountered with some of the reduplication patterns earlier. Either two surface forms are in an allomorphic relation or they arise from two different morphemes; there is little room for middle ground.

The verb classes cannot be completely predicted on the basis of root meanings but there are few to no surprises once roots are properly categorized into their classes. That is to say, once we understand that *raam* 'know' belongs to the \emptyset -*k*-class (that is, shows a zero/*k*-alternation), we can be sure that it will appear zero in bare form but that the *k-* will arise in further morphological expansions, as in *pa-k-raam* CAUS-STA-know. The same is true for all of the prefixes that identify classes. Yet every new morpheme is introduced by a chart that contains the full array of classes and internal prefixes. This creates the impression of a language with massively complex paradigms where, in actuality, the vast majority of these forms are predictable. Much paper could have been saved by only including unpredictable morphological combinations. Having said that, the charts leave no room for doubt that such forms do occur and were checked by the authors, which is surely an advantage.

8. NOMINALIZATIONS. The authors show that derived nouns (for example, *s<in>i'ael* <PERF.PATNMLZ>eat) pattern syntactically with underived nouns (for example, *'aelaw* fish). Unfortunately, they do not show that derived nouns pattern very differently from their indicative verbal counterparts. For instance, it is shown that derived and underived nouns can (i) occur in argument position, (ii) be preceded by case markers, (iii) cooccur with demonstrative pronouns, (iv) be put in possessive case with the circumclitic *'inko*= *=a*, and finally, (v) be modified by another noun. But of these five diagnostics, only (iv) was introduced in the main discussion of verb/noun differences in the language.

We are, thus, unsure whether (i), (ii), (iii), and (v) really serve to identify the nominalization process without ungrammatical examples of verbs in the diagnostic contexts. This is important because, as the authors note (481), “most nominalizers are identical in form to voice/aspect affixes (see Yeh 2000c, 2003, 2011). Thus, it is sometimes difficult to determine whether a word is verbal or nominal.” As seen in the examples below, it is only morphosyntactic context (and glossing) that differentiates forms like patient voice/nominalization *ka-en* (483, ex.9.14a,b).

- | | | |
|------|-----------------------------------|---|
| (19) | VERBAL PATIENT VOICE FORM | PATIENT NOMINALIZATION |
| | hini 'aelaw 'inisho' ka-si'ael-en | 'izi'='i 'ash-'ashay ka ka-si'ael-en! |
| | this fish 2S.DAT IRR-eat-UVP | NEGIMP=LIG RED-waste ACC IRR-eat-PATNMZ |
| | 'This fish is for you to eat.' | 'Don't keep on wasting food!' |

To add to the above confusion between verbs and argument nominalizations, there is a third homophonous category that is stealthily introduced as “clausal nominalizations” or “syntactic nominalizations.” It took the reviewer several readings to understand that the authors treat forms like the patient voice/nominalization above alternatively as “syntactic nominalizations” when they occur in a relative clause, stating that “the marking of the verb within the relative clause corresponds to the noun with which it agrees; the same nominalizers are used for argument and clausal nominalization” (481). This creates a massive circularity in the diagnostics used to differentiate putative “argument nominalizations” from “nominalized clauses.” To take but one example, when an *-en* form is associated with a genitive agent, it is treated as a nominalized clause, but when a genitive agent is absent, it is treated as an argument nominalization. The diagnostic “may contain a genitive agent” is then taken to distinguish argument nominalizations from clausal nominalizations. As far as I can tell, this problem plagues every purported distinction between the two types of nominalization. The only way in which the diagnostic properties of clausal and argument nominalizations avoid being tautological is if containing a genitive agent, for instance, sets all other properties of the form along the lines of a clausal nominalization rather than an argument nominalization. We would predict, then, that an *-en* form could appear in argument position but not if it is associated with a genitive agent. Unfortunately, nothing of the sort is demonstrated.

A point seemingly not mentioned in the text but implicit in the paradigm tables is that argument nominalizations differ from verbal forms in lacking an imperfective aspect. This would suggest that forms inflected for imperfective aspect cannot take on all the nominal properties of perfective or irrealis inflected forms. While this may have simply been an oversight, it should be discussed explicitly.

The gerund category, exemplified below (489), also suffers from some of the same problems described for the posited clausal nominalizations.

- | | |
|------|---------------------------------------|
| (20) | Kizaw t<om>ortoroe' ka ('am=)t<om>non |
| | Kizaw <AV>teach ACC GER.IRR=<AV>weave |
| | 'Kizaw is teaching weaving.' |

Regarding gerund forms, we are told (489) that, “in the realis, they may be identical in form to AV verbs. [...] In the irrealis, they are marked by 'a-/am= 'Ir?.” Unfortunately, this latter feature is also shared by indicative actor voice verbs, as 'a-/am are general aspect markers. In a footnote (491, fn.119) it is suggested that 'ima= (glossed as an agent

nominalizer) stands in contrast with *'am=*, as the former appears to require an agent nominalization reading rather than a gerund interpretation:

- (21) Kizaw t<om>ortoroē' ka 'ima=t<om>non
 Kizaw <AV>teach ACC AGTNMZ=<AV>weave
 'Kizaw is teaching those who are weaving.'

But this marker, too, occurs in plain indicative actor voice predicates (where it is glossed as a progressive marker), as seen below (344, ex.6.108), and, thus, it cannot be treated as a nominalizer.

- (22) kamatomortoroē' 'ima=t<om>ortoroē' ka shayshiat=a kapa'yakai'.
 teacher PROG=<AV>teach ACC Saisiyat=LIG speech
 'The teacher is teaching Saisiyat.'

It appears, then, that a morphological category that may not exist in Saisiyat has been foisted on to a clearly distinct syntactic construction that corresponds roughly with gerunds from more familiar languages. The morphology recruited for this context is simply actor voice, and this may correspond with the use of AV forms in embedded clauses in other Formosan languages (see Aldridge 2004, *inter alia*).

As mentioned earlier, Saisiyat has almost completely conflated the patient and locative voices in the verbal paradigm but is said to maintain this distinction in nominalization. When we inspect the meanings of locative nominalization, however, we find that it is difficult to discern a locative meaning in many of them. For instance (501), the locative form of 'write' would in most Austronesian languages with a locative voice refer to what was written on, or someone who was written to, but *k<in>aat-an* <PERF>write-LOCNMZ is translated as '(that) which was written'. The same holds for constructions built on *talek* 'cook' and others. The authors have correctly classified *-an* as both a patient and locative nominalizer, but there are perhaps forms listed as patient nominalizations that are best understood as locative nominalizations, as this category in Austronesian refers to more than a place, strictly speaking. For instance, 'leftovers' is the food that has been eaten from, 'footprint' is where someone has stepped, 'that which was salted' is where salt was put, and so forth. It is perhaps through these pseudo-object meanings that the locative has come to expand into the territory of the patient voice.³

9. SUMMARY. Clearly, a remarkable amount of energy was expended in the taxonomic side of the grammar, positing an impressive array of features and classes for every imaginable morphological contrast. Less energy, however, was spent on getting to the bottom of the larger patterns of use for individual forms. As a concrete example, we are shown a vast array of voice forms with their definitions and interrelations but it is largely left open as to what factors go into selecting one voice over another in actual discourse or in particular syntactic contexts. In this sense, the work interprets morphology as wholly distinct from syntax. If the book has any underlying structural weakness, it is due to putting the taxonomic cart before the functional horse. It is my feeling that further investiga-

3. See Blust and Chen (to appear) for more examples of locative voice subsuming patient voice historically in Formosan and Malayo-Polynesian languages. It is interesting to note, though, that in Saisiyat, the patient voice is far more robust than the locative voice, in distinction, apparently, to all other cases, which attest a patient > locative replacement.

tion could yield a considerable rearrangement of the paradigms and definitions of the functional morphology presented here but, more importantly, a crucial foundation for further morphosyntactic exploration has been laid.

In closing, I would mention one particular rich area for further research that has been surprisingly neglected in the Saisiyat literature except for a few stray comments, and that is the role of language contact in Saisiyat, and the changes that have taken place from Ogawa and Asai's (1935) documentation to the modern language. While the word order of modern Saisiyat looks almost nothing like other Formosan languages, deviations do not always appear to be induced by Chinese. At the same time, the morphological simplification and lexical borrowing that we might expect with heavy contact is seemingly nowhere to be found. This area of study is still in its infancy and I believe many interesting questions may be solved by delving into the sparse but important historical record collected at the beginning of the last century.

I have only been able to scratch the surface of the rich body of complex morphological phenomena documented by Zeitoun, Chu, and kaybawbaw in their over 600-page work. I hope it is clear that, despite whatever critiques have been made above, the work under review is an outstanding contribution to our understanding of Saisiyat in particular and Formosan languages more generally. Hopefully, the authors will continue their laudable work and further fill in the descriptive and analytical gaps while it is still possible.

REFERENCES

- Aldridge, Edith. 2004. Ergativity and word order in Austronesian languages. PhD diss., Cornell University.
- Blust, Robert. 1998. Ca-reduplication and Proto-Austronesian grammar. *Oceanic Linguistics* 37:29–64.
- . 2003. Three notes on early Austronesian morphology. *Oceanic Linguistics* 42:438–78.
- . 2009. *The Austronesian languages*. Canberra: Pacific Linguistics.
- Blust, Robert, and Victoria Chen. To appear. The pitfalls of negative evidence: “Nuclear Austronesian,” “Ergative Austronesian,” and their progeny. *Language and Linguistics*.
- Chang, Henry Y. 2011. Tsou case markers revisited: Visibility, proximity, and beyond. In *Diversity of languages: Papers in honor of Professor Feng-fu Tsao on the occasion of his retirement*, ed. by Raung-fu Chung, Wei-tien Dylan Tsai, and Hsiu-hsue Liu, 91–123. Taipei: The Crane Publishing Co.
- Chou, Yi-ming Marc. 2016. Two types of object control in Saisiyat: A movement-based approach. *Linguistics and Literature Studies* 4(3):188–202.
- Hayes, Bruce, and May Abad. 1989. Reduplication and syllabification in Ilokano. *Lingua* 77(3–4):331–74.
- Kaufman, Daniel. 2009a. Austronesian nominalism and its consequences: A Tagalog case study. *Theoretical Linguistics* 35(1):1–49.
- . 2009b. Austronesian typology and the nominalist hypothesis. In *Austronesian historical linguistics and culture history: A festschrift for Robert Blust*, ed. by Alexander Adelaar and Andrew Pawley, 187–215. Canberra: Pacific Linguistics.

- . 2011. Exclamatives and temporal nominalizations in Austronesian. In *Nominalization in Asian languages: Diachronic and typological perspectives*, ed. by Foong Ha Yap and Janick Wrona, 721–54. Amsterdam/Philadelphia: John Benjamins.
- . 2012. PAN *ka- and predicate classes. Paper presented at AFLA 19, Taipei: Academia Sinica.
- Kaybaybaw, Ya'aw Kalahae. 2009. A morphological and semantic study on word formation in SaySiyat. MA thesis, National Hsin-Chu University of Education.
- Nojima, Motoyasu. 1996. Lexical prefixes of Bunun verbs. *Gengo Kenkyu: Journal of the Linguistic Society of Japan* 110:1–27.
- Ogawa, Naoyoshi, and Erin Asai. 1935. *The myths and traditions of the Formosan native tribes*. Taipei: Taihoku Imperial University.
- Ross, Malcolm. 1995. Proto-Austronesian verbal morphology: Evidence from Taiwan. In *Austronesian studies relating to Taiwan, number 3*, ed. by Paul Jen-kuei Li, Cheng-we Tsang, Dah-an Ho, Ying-kuei Huang, and Chiu-yu Tseng, 727–91. Taipei: Academia Sinica.
- . 2015. Reconstructing Proto Austronesian verb classes. *Language and Linguistics* 16(3):279–345.
- Starosta, Stanley, Andrew Pawley, and Lawrence A. Reid. 1982. The evolution of focus in Austronesian. In *Papers from TICAL*, vol. 2, ed. by Amran Halim, Lois Carrington and S.A. Wurm, 201–216. Canberra: Pacific Linguistics.
- Teng, Stacy Fang-ching. 2008. *A reference grammar of Puyuma, an Austronesian language of Taiwan*. Canberra: Pacific Linguistics.
- Yeh, Marie M. 1991. Saisiyat structure. MA thesis, Tsing Hua University.
- . 2000. *A reference grammar of Saisiyat*. Series on Formosan Languages 2. Taipei: Yuanliou. [In Chinese.]
- Zeitoun, Elizabeth. 2001. Negation in Saisiyat: Another perspective. *Oceanic Linguistics* 40:125–34.
- Zeitoun, Elizabeth, and Lillian M. Huang. 2000. Concerning ka-, an overlooked marker of verbal derivation in Formosan languages. *Oceanic Linguistics* 39:391–414.
- Zeitoun, Elizabeth, and Chen-huei Wu. 2005. Saisiyat reduplication revisited. *Concentric: Studies in Linguistics* 31(2):31–56.
- Zuraw, Kie, and Yu-an Lu. 2009. Diverse repairs for multiple labial consonants. *Natural Language and Linguistic Theory*, 27:197–224.