

VOICE AND PRONOMINAL FORMS IN KAYAN (UMA NYAVING)

Alexander D. Smith
*National University of
Singapore*
<smithad@hawaii.edu>

Michael Yoshitaka Erlewine
*University of Helsinki / National
University of Singapore*
<mitcho@nus.edu.sg>

Carly J. Sommerlot
*National University of
Singapore*
<cjsomms@nus.edu.sg>

Abstract

We provide a description of the basic clausal syntax of the Kayan language of Borneo (Austronesian) as spoken in Uma Nyaving (Sarawak, Malaysia), with particular emphasis on the inventory of voice and pronominal forms and their interactions. We show that this variety of Kayan includes two distinct types of analytic passive constructions, an undergoer voice construction specifically encoding first or second person agent features, and a maximally four-way distinction between pronominal forms. We highlight similarities to voice and pronominal forms in related language varieties of the region, and also discuss the potential historical relationship between different passive marker forms. Uma Nyaving Kayan's voice and pronominal system exemplifies a profile of grammatical features that have been described as "Central Bornean type."

Keywords: analytic passives, undergoer voice, pronominal forms, Kayan, Central Bornean type
ISO 639-3 codes: ree

1 Introduction

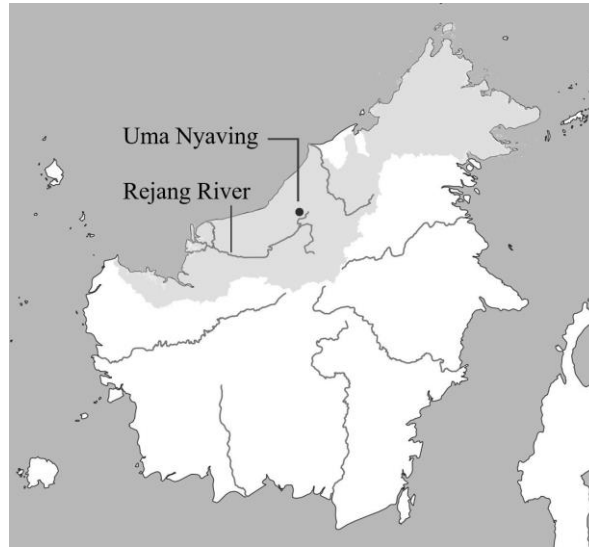
Kayan is a dialectically heterogeneous language of Borneo, with distinct varieties in both Sarawak (Malaysia) and West, East, and North Kalimantan (Indonesia). Dialect boundaries typically correspond to the river systems where the language is found, e.g., Baram river Kayan (ISO 639-3: kys), Rejang river Kayan (ISO 639-3: ree), Kayan river Kayan (ISO 639-3: xkn), and others. Kayan varieties belong to the larger Kayanic subgroup, which includes Murik-Merap (Murik or Ngorek along with Merap) and Segai-Modang (Smith 2017b). In this paper, we describe the basic clausal syntax of Uma Nyaving Kayan, a variety of the Rejang river Kayan dialect, with a focus on its voice system and its interaction with the pronominal system.¹ The Uma Nyaving community is located within the Sungai Asap resettlement area in the upper Rejang river. Map 1 shows the location of Uma Nyaving and its proximity to the Rejang river.

Like other varieties of Kayan, Uma Nyaving Kayan (hereafter, simply: "Kayan"), exhibits a basic S(Aux)VO word order (as in (1a)) with limited derivational verbal morphology. The language distinguishes multiple passive constructions, which all involve preverbal, non-affixal markers. We discuss and defend the use of the term "passive" for these constructions later. These passive constructions include the *an*-marked passive (1b), "agent-inflected passives" with the markers *ak* and *im* which specifically encode first-singular and second-singular agents, respectively (1c), and a periphrastic passive related to a verb meaning 'be hit' or 'get,' *ga?* (1d). There are no morphological passives in Kayan.²

¹ Uma Nyaving data are from Smith, Sommerlot & Erlewine 2023 (Kaipuleohone archive deposit) unless otherwise noted.

² Our glosses follow the Leipzig glossing rules (Haspelmath & Bickel 2015), with the following additions: N- and M- for verbal prefixes (§2.2); RED = reduced, SHORT = short for pronominal forms (§3).

Map 1: Location of Uma Nyaving within Malaysian Borneo



(1) Basic word order and multiple analytic passives:

- | | |
|--|--|
| <p>a. <u>Active Clause</u>
 <i>aso?</i> <i>anih</i> <i>ma?</i>ət <i>akuj</i>
 dog PROX bite 1SG
 ‘This dog bit me.’</p> | <p>c. <u>Agent-inflected passive</u>
 <i>iha?</i> { <i>ak</i> / <i>im</i> } <i>məte?</i>
 3SG PASS.1SG PASS.2SG hit
 ‘He was hit by me / you.’</p> |
| <p>b. <u>an passive</u>
 <i>akuj</i> <i>an</i> (<i>aso?</i> <i>anih</i>) <i>ma?</i>ət
 1SG PASS (dog PROX) bite
 ‘I was bitten (by this dog).’</p> | <p>d. <u>ga?</u> passive
 <i>akuj</i> <i>ga?</i> <i>ma?</i>ət ((<i>kəlan</i>) <i>aso?</i> <i>anih</i>)
 1SG get bite ((because) dog PROX)
 ‘I got bitten (by this dog).’</p> |

We schematize the basic word orders of these constructions for an eventive bivalent verb in (2), bolding the characteristic passive markers. There is no marked morphology associated with active clauses, as we discuss further below. Note that agents of *ga?* constructions (2d) require the preposition *kəlan* when not immediately postverbal.

(2) Word order schemas for active and passive clause types:

- | | |
|---|---|
| <p>a. <u>Active:</u></p> | S/Agent (Aux) V O/Undergoer |
| <p>b. <u>an passive:</u></p> | S/Undergoer (Aux) an (Agent) V |
| <p>c. <u>Agent-inflected passive:</u></p> | S/Undergoer (Aux) ak/im V |
| <p>d. <u>ga?</u> passive:</p> | S/Undergoer (Aux) ga? V (Agent) – or – ... (<i>kəlan</i> Agent) |

Kayan exhibits grammatical features that have been described as ‘Central Bornean-type’ (Clayre 1996, 2014, Kroeger & Smith to appear, Sommerlot to appear), in contrast to the well-studied categories of so-called Philippine-type and Indonesian-type grammar languages (see e.g. Himmelmann 2005). In particular, these languages have a two-way voice distinction (actor/undergoer, or active/passive), an analytic passive marked by a non-affixal, preverbal marker, distinct genitive pronouns (but no case marking outside of this), and limited verbal morphology (in particular, lacking applicative suffixes found in Indonesian-type languages). Both the *an* and *ga?* passives, which utilize preverbal, non-affixal markers, fit into this profile, and Kayan follows many other Central Bornean-type languages in having distinct genitive pronouns, which we detail in Section 3.³

³ There remains some uncertainty in the literature around the features that define ‘Central Bornean-type’ grammars. See Sommerlot to appear for a discussion of the variation found within this voice profile.

We will also highlight that the *gaʔ* passive frequently expresses subject affectedness, making its properties quite parallel to periphrastic passives related to a verb meaning ‘get’ in many other languages of the region, including the familiar *kena* construction in Malay (see e.g. Nomoto & Wahab 2012).

Our discussion is informed by and complements prior descriptions of other Kayan varieties, as in Clayre 1996 (focusing on Baram Kayan), Blust 1997 (focusing on Uma Juman, a variety of Rejang Kayan), Soriente 2013 (also focusing on Uma Pu, a variety of Baram Kayan), and Clayre 2014 (focusing on the Baram Kayan varieties Uma Pu and Uma Peliau). We are not aware of any prior work specifically on Uma Nyaving Kayan, making this paper the first to describe this variety, although our results by and large accord with these previous descriptions of related varieties, as we highlight throughout. In addition, the *gaʔ* passive to our knowledge has not been previously described in any Kayan variety.

Our paper is organized as follows. We first describe the general word order and other salient features of the basic morphosyntax of Kayan in section 2, and then introduce the various pronominal forms and their distribution in section 3. A description of the pronominal forms is necessary since pronominal behavior interacts in non-trivial ways with voice forms in Kayan. We then introduce the passive constructions in section 4. In particular, we will argue that the *an*-marked and agent-inflected passives are closely related, both in their synchronic syntactic properties and in their forms, whereas the *gaʔ* passive is more distinct. We will also discuss our motivation for describing these constructions as varieties of passives at the top of section 4. We conclude in section 5 by situating Kayan’s voice and pronominal inventories within the broader syntactic typology of western Austronesian languages.

2 Basic clausal morphosyntax

2.1 Word order of simple clauses

We begin with an overview of the word order of simple clauses in Kayan. Kayan clauses are typically subject-initial, as seen in (3). Other arguments of transitive verbs appear after the verb, as in (3a–c). In all examples in this section, the subject is the predicate’s actor (or else an experiencer of a mental state predicate as in (3e)), so we may describe these examples as actor-oriented or in the active voice.

(3) Simple clauses with unmarked word order:

- a. *akuj ninəŋ {ikaʔ / nipaʔ} dahələm deh*
 1SG see 2SG / snake yesterday
 ‘I saw {you / a snake} yesterday.’
- b. *{tamak (< taman-k) / dahaʔ} maʔe akuj sin*
 father-1SG.SHORT / 3PL give 1SG money
 ‘{My father / they} give me money.’
- c. *hari-k ɲitən akuj*
 sibling-1SG.SHORT pinch 1SG
 ‘My sibling pinches me.’
- d. *akuj aləm təpuruŋ*
 1SG PROG run
 ‘I am running.’
- e. *akuj takut*
 1SG afraid
 ‘I am afraid.’

- f. *aso?* *anən* *tudu*
 dog DIST sleep
 ‘The dog sleeps.’

Kayan has a number of passive constructions, described in section 4. Notice that, in contrast, there is no morphology specifically identifying the clauses as in (3) as being in the active voice. There are also no morphological case distinctions on core arguments.

Temporal auxiliaries, modals, and negators may precede the predicate, as we see in (4). The basic word order may thus be described as S(Aux)VO. Like many other languages of the region, examples such as in (3) without any temporal auxiliaries are temporally underspecified and can be interpreted as past, present, or future, depending on the context. When tense is underspecified in this way, we generally use the present tense for the English translations. Temporal adjuncts can also serve to specify the temporal specification for the clause, as in (3a). In passive clauses, auxiliaries must precede the passive marker, as in (4c).

(4) Auxiliaries precede the predicate:⁴

- a. {*naŋaŋ* / *iha?*} {*au* / *dʒan*} *te ha pasəl*
 Nanyang / 3SG AUX / NEG go LOC store
 ‘{Nanyang / he} {went / did not go} to the store.’
- b. *akuŋ dʒan dəŋ kuman anih*
 1SG NEG ABLE eat PROX
 ‘I cannot eat this.’
- c. *kajo? anən au an nəvəŋ*
 wood DIST AUX PASS fell
 ‘That tree is already cut down.’

Although subjects are generally preverbal, certain predicates allow for postverbal subjects. Stative predicates like (*mərən*) *lasu* ‘hot’ in (5a), *biaŋ* ‘broken’ in (5b), and *la?e* ‘tired’ in (5c) frequently occur precede their non-agentive subjects. This contrasts with dynamic intransitive verbs, which do not allow for a predicate-initial order. The active verb *tepuruŋ* ‘run’ in (6), for example, does not allow for a postverbal subject in either the full form *akuŋ* or the reduced clitic form *kuj*.

(5) Predicate-initial stative clauses:

- a. (*mərən*) *lasu ata? anən*
 (very) hot water DIST
 ‘This water is (very) hot.’
- b. *biaŋ kəliŋi anən*
 broken window DIST
 ‘The window is broken.’
- c. (*mərən*) *la?e=kuj*
 (very) tired=1SG.RED
 ‘I am (very) tired.’

⁴ Many sentences with the auxiliary *au* invite English translations with the adverb ‘already,’ but we choose the broad gloss AUX as we are not confident about its exact semantics. *dəŋ* as in (4b) is a possibility modal.

(6) Ungrammatical predicate-initial dynamic intransitive:

- * *təpuruŋ* (a)*kuj*
 run 1SG
 Intended: ‘I am running.’

This possibility for stative descriptions to be predicate-initial has also been described for two other languages of northern Borneo, Sa’ban, and Lundayeh (Clayre 2014).

Active subjects can also follow an auxiliary, resulting in Aux-SVO word order, when in reduced, enclitic forms, as seen in (7a, b). We discuss these reduced pronominal forms in section 3.2. In contrast, non-pronominal subjects cannot follow auxiliaries; see (8). Note that a variant of example (8) with the subject ‘that tree’ in clause-initial position is grammatical, as presented earlier in (4c).

(7) Reduced subject pronoun following the auxiliary:

- a. *au=kuj* *pə-pənu* *bakol* *anih* *dahin* *baha*
 AUX-1SG.RED CAUS-full basket PROX with uncooked.rice
 ‘I already filled the basket with rice.’
- b. *dʒan=haʔ* *te* *ha* *pasəl*
 NEG=2SG.RED go LOC store
 ‘He did not go to the store.’

(8) Ungrammatical full noun subject following the auxiliary:

- * *au* *kajoʔ* *anən* *an* *nəvəŋ*
 AUX wood DIST PASS fell
 Intended: ‘That tree is already cut down.’

As noted earlier, a clause can include multiple preverbal auxiliaries. In such cases, there is flexibility in the position of reduced form subjects. See the position of *kuj* in (9).

(9) Flexibility of reduced subject placement with multiple auxiliaries:

- a. *dʒan=kuj* *pajan* *pə-pənu* *bakol* *anih* *dahin* *baha*
 NEG=1SG.RED yet CAUS-full basket PROX with uncooked.rice
- b. *dʒan* *pajan=kuj* *pə-pənu* *bakol* *anih* *dahin* *baha*
 NEG yet=1SG.RED CAUS-full basket PROX with uncooked.rice
 Both: ‘I have not yet filled the basket with rice.’

2.2 Verbal morphology

As noted earlier, active clauses in Kayan are not identified by any voice-marking morphology and the language overall utilizes only a limited inventory of affixal morphology. In our data set, verbal morphology includes i) N-, a homorganic nasal that derives verbs from nouns and adjectives, ii) *pə-*, a causative prefix, and iii) *mə-*, a verbal prefix that derives verbs from nouns. This morphology reflects familiar morphology reconstructed to Proto-Malayo-Polynesian (PMP), including PMP *man-, *pa-, and *ma- (See Blust 2013, Wolff 1973, Ross 2009, and Adelaar 2005a for more on Proto-Austronesian and PMP verbal morphology).⁵

The homorganic nasal, N-, attaches to both nouns and adjectives. Its phonological realizations are similar to those described in other Austronesian languages with active homorganic nasal substitution, such as Ilocano (Rubino 2000), Indonesian (Sneddon 1996), Mapun (Collins, Collins & Hashim 2001), and many others. Where the stem begins with non-sonorant consonant, the nasal assimilates and deletes the stem-initial

⁵ In many languages (including Kayan), *manN- is reflected with only the nasal element N- (Blust 2004).

consonant. The nasal surfaces as *ŋ*- before vowel-initial stems and as *ŋə*- before sonorant-initial stems. Several examples showing N-'s morpho-phonological behavior are shown in (10).

(10) Some derived verbs with n-:

<i>takut</i> 'afraid'	→	<i>nakut</i> 'to frighten someone'
<i>biaŋ</i> 'broken (glass)'	→	<i>miaŋ</i> 'to break (glass)'
<i>pate</i> 'corpse'	→	<i>mate</i> 'to die'
<i>anak</i> 'child'	→	<i>ŋanak</i> 'to give birth'
<i>luvaŋ</i> 'hole'	→	<i>ŋəluvaŋ</i> 'to make a hole'

N- combines with some stems to produce causative or inchoative forms, increasing the valency of the base form. This is exemplified with the adjective *takut* and the noun *pate* in (11) and (12).

(11) N-causative from 'afraid':

a.	<i>akuj</i>	<i>takut</i>
	1SG	afraid
	'I am afraid.'	

b.	<i>ikaʔ</i>	<i>nakut</i>	<i>akuj</i>
	2SG	N-afraid	1SG
	'You scare me.'		

(12) N-inchoative from 'corpse':

a.	<i>akuj</i>	<i>ŋinəŋ</i>	<i>pate</i>	<i>anən</i>
	1SG	see	corpse	DIST
	'I saw that dead person.'			

b.	<i>asoʔ</i>	<i>anən</i>	<i>au</i>	<i>mate</i> (< N- <i>pate</i>)
	dog	DIST	AUX	N-corpse
	'That dog has died.'			

Verbs may also be derived with the prefix *mə*-, but in our data set only nouns take this prefix. Blust 1997 describes *mə*- as attaching to nominal roots, but also that it only derives stative verbs. In our data set, most uses are indeed stative, but at least one example, *mə-pa* 'to sweep', seems to express a dynamic meaning, so Blust's earlier description appears to not hold of modern day Uma Nyaving. A couple examples with *mə*-derived predicates are shown here in (13). Note that in (13a), the prefix surfaces without a schwa on vowel-initial roots. No examples of *mə*- on a two-syllable consonant-initial root are found in our data set.

(13) Use of the *mə*- verbalizer:

a.	<i>wit</i>	<i>anih</i>	<i>au</i>	<i>m-asəp</i>
	dishes	PROX	AUX	M-dirt
	'These dishes are already dirty.'			

b.	<i>akuj</i>	<i>aləm</i>	<i>mə-pa</i>	<i>alim</i>
	1SG	PROG	M-broom	room
	'I am sweeping the room.'			

Note that both *mə*- and N- here are both simply verbal derivational morphology. There are similar (and probably cognate) prefixes in nearby Malayic languages where they (*meN-* / *N-*) have been described as active voice prefixes (see e.g. Sneddon 1996, Cole, Hermon & Yanti 2008). However, as we discuss in section 4, *mə*- and *N-* in Kayan remain even in passive constructions, suggesting that their function is not to mark active voice. This is not uncommon in languages of Borneo; for example, Sommerlot (to appear) reports that several Land Dayak languages also retain their nasal prefixes in passive constructions.

The last morpheme we discuss is the causative prefix *pə*-, which typically attaches to either intransitive verbs or adjectives, such as *tudu* 'sleep' → *pətudu* 'put to sleep' and *lasu* 'hot' → *pəlasu* 'to heat something'. Some words that take the causative prefix in our data are shown here in (14), with some uses exemplified in (15–16).

(14) Some *pə*- causatives:

<i>tudu</i> ‘sleep’	→	<i>pə-tudu</i> ‘to put to sleep’
<i>lægak</i> ‘fall’	→	<i>pə-lægak</i> ‘to drop’
<i>pənu</i> ‘full’	→	<i>pə-pənu</i> ‘to fill’
<i>lasu</i> ‘hot’	→	<i>pə-lasu</i> ‘to heat’
<i>həŋəm</i> ‘cold’	→	<i>pə-həŋəm</i> ‘to cool’
<i>takut</i> ‘afraid’	→	<i>pə-takut</i> ‘to frighten’

(15) Use of the *pə*- causative of ‘sleep’

a.	<i>iha?</i>	<i>aləm</i>	<i>tudu</i>	b.	<i>akuj</i>	<i>pə-tudu</i>	<i>iha?</i>
	3SG	PROG	sleep		1SG	CAUS-sleep	3SG
	‘He is sleeping.’				‘I put him to sleep.’		

(16) Use of the *pə*- causative of ‘fall’

a.	<i>iha?</i>	<i>lægak</i>	b.	<i>ak</i>	<i>pə-lægak</i>	<i>biək</i>	<i>anih</i>	<i>həp</i>
	3SG	fall		PASS.1SG	CAUS-fall	bag	PROX	about.to
	‘He falls.’			‘The bag is dropped by me.’				

Morphological causatives contrast with analytic causatives formed from the verb *na* ‘to make,’ with the latter expressing more indirect causation.

(17) Morphological and analytic causatives:

a.	<i>akuj</i>	<i>pə-tədu</i>	<i>iha?</i>	b.	<i>akuj</i>	<i>na</i>	<i>iha?</i>	<i>tudu</i>
	1SG	CAUS-sleep	3SG		1SG	make	3SG	sleep
	‘I put him to sleep.’				‘I made him sleep.’			

Central Bornean type languages, as described by Clayre (1996) and others, are also notable for lacking applicative suffixes of the type observed in nearby, so-called Indonesian-type languages. Kayan seems to pattern like Central Bornean-type languages in this regard. For example, the suffix *-kan* in Indonesian-type languages would affix to the base forms in (14) to form causatives; in Kayan, only the prefix *pə*- is used, as exemplified above. We have not encountered any other applicative suffixes in our work with Kayan either, and previous descriptions of other varieties of Kayan additionally make no mention of such suffixes. For example, applicative morphology is not utilized in dative alternations like that in (18).

(18) No applicative morphology for dative alternation:

a.	<i>akuj</i>	<i>matə</i>	<i>kət</i>	<i>kuman</i>	<i>mən</i>	<i>hina?</i>	<i>dahələm</i>	<i>deh</i>
	1SG	send	thing	food	DAT	mother	yesterday	
	‘I sent food to mother yesterday.’							
b.	<i>akuj</i>	<i>matə</i>	<i>(mən)</i>	<i>hina?</i>	<i>kət</i>	<i>kuman</i>	<i>dahələm</i>	<i>deh</i>
	1SG	send	(DAT)	mother	thing	eat	yesterday	
	‘I sent mother food yesterday.’							

Lastly, Blust 1997 describes additional verbal morphology in the Uma Juman variety of Rejang Kayan not found in our data set, including a *pə*- ‘reciprocal’ prefix, a *pək*- variant of the causative prefix, and an *<əm>* infixal allomorph of *mə*-. The lack of these forms in our experience with Uma Nyaving is likely a gap in our dataset. For example, numerous variants of *pə*- are found in languages throughout the Austronesian world, including in other Kayan dialects (Clayre & Cubit 1974).

3 Pronouns

In this section we describe the Kayan pronominal system, including the pronominal paradigm, the function of full, reduced, genitive, and short form pronominals, as well as restrictions on their position and form. The inventory that we describe is similar to that described in Clayre 2014 (p. 140 table 7) for Uma Pu and Uma Peliau, two Baram river dialects of Kayan, but we will describe and illustrate their contexts of use in greater detail here.

Kayan’s pronominal system shares many characteristics with that of neighboring languages, and fits within an areal typology that transitions from more conservative and complex systems in the Philippines and North Borneo to reduced systems south of Sabah. The inventory includes full and reduced (enclitic) forms, as well as a set of short suffixal forms used for possessors, but lacking distinct oblique or topic forms, unlike many of the more conservative Sabahan languages to the north (Lobel 2016). The use of genitive forms for agents in passives is additionally found in other north-central Sarawak languages like Melanau, Berawan, and Kenyah (Clayre 1996, Soriente 2013), as well as a few languages spoken further south, like Banyaduq (Land Dayak) and Ahe (Malayic) (Sommerlot 2020, to appear). The use of short suffixal forms as possessors is additionally found in Murik, and fossilized remnants of such pronouns exist in both Penan and Sa’ban (Clayre 1996). In contrast, pronominal paradigms in languages further to the south, including most Malayic and Land Dayak languages, are generally further reduced, with many languages showing no distinction in forms at all, or only distinguishing between a full form and a shorter genitive form (Sommerlot 2020, to appear).

3.1 Pronominal Paradigms

Kayan’s pronominal forms, specifically the singular pronouns, express a maximally four-way distinction based on their contexts of use. We will call these forms full, reduced, genitive, and short forms.⁶ In examples throughout, we gloss the non-full forms RED, GEN, SHORT. We discuss the use of these different forms in this section.

The singular pronoun inventory is shown in table 1. As we will show, only third-singular pronouns exhibit the distinction between reduced and genitive forms.⁷ Note that these pronouns appear to be limited to animate referents; see example (20).

Table 1: Singular Pronominals

	full	reduced	genitive	short
1SG	<i>akuj</i>	<i>=kuj</i>	<i>=kuj</i>	<i>-k</i>
2SG	<i>ika?</i>	<i>=ka?</i>	<i>=ka?</i>	<i>-m</i>
3SG	<i>iha?</i>	<i>=ha?</i>	<i>=na?</i>	<i>-n na?</i>

For non-singular numbers, Kayan distinguishes dual, paucal, and plural series, as in table 2. Some non-singular pronouns—i.e. first person inclusive dual, first person inclusive plural, and first person exclusive plural—have reduced forms as well. The other non-singulars have only full forms. These reduced forms, shown immediately following the full form in table 2, also function as genitive forms. Categories with no reduced forms use the full form in genitive contexts. No short forms exist for the nonsingular pronouns.

⁶ Our full/reduced/genitive/short labels correspond to sets IA/IB/IIA/IIB respectively in Clayre 2014. Clayre 1996 makes a two-way distinction between what she calls focus or nominative forms (our full and reduced) and genitive forms (our genitive and short forms).

⁷ Clayre 2014 lists *ikam* as the second person plural, whereas we record Uma Nyaving Kayan *kəlo?*. The Uma Nyaving plural arose from a reinterpretation of the paucal, itself formed from a fusion of the second person singular and the numeral ‘three’ (*ika? təlo? → kəlo?*). Such reinterpretations that result in a loss of the original plural are common in Borneo (Smith 2017a). Additionally, our inventory in table 1 and that in Clayre 2014 (p. 140 table 7) differ in the description of a *na?* non-human form. See our discussion around example (20) below.

Table 2: Nonsingular Pronominals

	dual	paucal	plural
1INCL	<i>itu / =tu?</i>	<i>təlo?</i>	<i>itam / =tam</i>
1EXCL	<i>kawa?</i>	<i>kalo?</i>	<i>kame? / =ame?</i>
2	<i>kua?</i>	<i>kəlo?</i>	<i>kəlo?</i>
3	<i>dahawa?</i>	<i>dahalo?</i>	<i>daha?</i>

The form a pronoun takes is determined by a number of syntactic, semantic, and phonological factors, which we turn to next.

3.2 Full and Reduced Pronouns

We first discuss the distribution of full and reduced pronouns. Full forms have the widest distribution, as they are allowed in all nominal positions. This contrasts with the reduced pronominal forms, which are much more limited. For example, a subject cannot be a reduced form in clause-initial position (19a), but may appear following a functional element such as the relative complementizer *aləŋ* as in example (19b) or following an auxiliary as in (21a).

(19) Positional restrictions on reduced, enclitic forms:

- a. {*aku*_j / *=*ku*_j} *pə-tudu* *iha?*
 1SG.FULL / =1SG.RED CAUS-sleep 3SG
 ‘I put him to sleep.’
- b. *an*_{ih} *surat* *aləŋ* {*ika?* / =*ka?*} *ma?*_e *mən* *aku*_j
 PROX letter REL 2SG.FULL / =2SG.RED give DAT 1sg
 ‘This is the letter that you gave to me.’

Recall that we described the third-singular pronoun as exhibiting the greatest number of distinctions in form, which Clayre 2014 independently corroborates for the Uma Pu and Uma Peliau, two Baram river dialects; see her table 7. We give the relevant forms as full *iha?*, reduced =*ha?*, genitive =*na?*, and short -*n na?* (see section 3.4). Against this backdrop, we note the existence of utterances such as (20) where the form *na?* is in sentence-initial position, where we predict pronouns to appear in their full form. The difference is that the subject here is an inanimate demonstrative pronoun *na?*, rather than the reduced form of the (animate) third-singular pronoun, which is simply homophonous.

(20) Inanimate demonstrative *na?*:

- na?* *au* *ak* *na?*_{ɟəm} (< N-*ta?*_{ɟəm})
 that AUX PASS.1SG N-sharp
 ‘That (inanimate) is already sharpened by me.’

This example also serves to show that the full/reduced alternation is not simply a reflection of some constraint that bans light, monosyllabic words in sentence-initial position, in relation to observed universal tendencies for prosodic hosts to appear at utterance edges (cf e.g. McCarthy & Prince 1993, Branau to appear).

As objects of verbs or prepositions, reduced form pronouns are judged as marginal, with full forms clearly preferred. This is exemplified in (21). We note, however, that reduced pronouns are possible with stative clauses with postverbal subjects, as in (5c), repeated as (22).

(21) Preference for full forms in object position:

- a. *dʒan* {*ihaʔ* / *=haʔ*} *maʔe surat mən* {*akuʃ* / *ʔʔ=kuʃ*}
 NEG 3SG.FULL / =3SG.RED give letter DAT 1S.FULL / =1SG.RED
 ‘He did not give a letter to me.’
- b. *hari-k* *ɲitən* {*akuʃ* / *ʔʔ=kuʃ*}
 sibling-1SG.SHORT pinch 1SG.FULL / =1SG.RED
 ‘My sibling pinches me.’

(22) Stative clause with postverbal reduced form subject:

- (*mərəm*) *laʔe=kuʃ*
 (very) tired=1SG.RED
 ‘I am (very) tired.’

3.3 Genitive Pronouns

In this subsection and the next, we discuss the form of nominal possessors, which follow their possessed nouns. Although the full and reduced pronominal forms introduced above appear at first glance to also be available as nominal possessors, we observe a curious distinction in the third-singular.

Although the enclitic third-singular form following auxiliaries must be *=haʔ* (related to the full form *ihaʔ*), as introduced above, third-singular nominal possessors have a distinct enclitic form *=naʔ* instead. These forms are in complementary distribution, as shown in (23–25). We refer to this *=naʔ* form as a genitive, and we similarly gloss other enclitic forms in these environments as genitive as well.

(23) Ungrammaticality of *naʔ* in post auxiliary position:

- au / dʒan* {*=haʔ* / **=naʔ*} *te ha pasəl*
 AUX / NEG =3SG.RED / =3SG.GEN go LOC store
 ‘He {already went / did not go} to the store.’

(24) Ungrammaticality of *haʔ* as a possessor:

- asoʔ* {**=haʔ* / *=naʔ*}
 dog =3SG.RED / =3SG.GEN
 ‘His dog.’

The genitive form is also used to express non-subject agents in *an* passives. This will be discussed in more detail in section 4.1.

(25) Preference of the genitive as non-subject agent in a passive construction:

- kanən anən an* {**=haʔ* / *=naʔ*} *kuman*
 rice DIST PASS =3SG.RED / =3SG.GEN eat
 ‘That rice was eaten by him.’

3.4 Short Form Pronouns

The short form pronouns have the most limited distribution. While they, like the genitive form described above, are used to indicate possession, they can only do so while satisfying two specific constraints. First, these forms are subject to a phonological constraint; they are only licit when attaching to words that end in either *-ʔ* or *-n*. When attached, they replace the final consonant of the root, as demonstrated in (26).

(26) Reduced and short pronominal possessors:

		genitive	short	
a.	1SG	<i>buloʔ=kuj</i>	<i>bulo-k</i>	‘my hair’
	2SG	<i>buloʔ=kaʔ</i>	<i>bulo-m</i>	‘your hair’
	3SG	<i>buloʔ=naʔ</i>	<i>bulo-n naʔ</i>	‘his/her hair’
b.	1sg	<i>dʒaʔan=kuj</i>	<i>dʒaʔa-k</i>	‘my chin’
	2SG	<i>dʒaʔan=kaʔ</i>	<i>dʒaʔa-m</i>	‘your chin’
	3SG	<i>dʒaʔan=naʔ</i>	<i>dʒaʔa-n naʔ</i>	‘his/her chin’

Attempts to use short form possessors with roots with other codas result in ungrammaticality, as in (27). For these, only the genitive form is available to indicate a possessive relationship.

(27) Other body parts disallowing short form possessors:

		genitive	short	
a.	1SG	<i>anit=kuj</i>	* <i>ani-k</i>	‘my skin’
	2SG	<i>anit=kaʔ</i>	* <i>ani-m</i>	‘your skin’
	3SG	<i>anit=naʔ</i>	* <i>ani-n naʔ</i>	‘his/her skin’
b.	1SG	<i>tudək=kuj</i>	* <i>tudə-k</i>	‘my leg’
	2SG	<i>tudək=kaʔ</i>	* <i>tudə-m</i>	‘your leg’
	3SG	<i>tudək=naʔ</i>	* <i>tudə-n naʔ</i>	‘his/her leg’

In addition to this phonological constraint, the short form is only available for expressing inalienable possessors. Inalienably possessed nouns in Kayan include body parts, as in (26–27), and kinship terms as in (28):

(28) Kinship terms also allow short form possessors:

		genitive	short	
	1sg	<i>harin=kuj</i>	<i>hari-k</i>	‘my sibling’
	2sg	<i>harin=kaʔ</i>	<i>hari-m</i>	‘your sibling’
	3sg	<i>harin=naʔ</i>	<i>hari-n naʔ</i>	‘his/her sibling’

For nominal possessors to appear in the short form, both the phonological and semantic constraints must be met. For example, if a noun adheres to the phonological constraint but its possessor expresses an alienable possession relation, the short form cannot be used as shown here in (29).

(29) Short forms unavailable for alienable possession:

		genitive	short	
	1sg	<i>asoʔ=kuj</i>	* <i>aso-k</i>	‘my dog’
	2sg	<i>asoʔ=kaʔ</i>	* <i>aso-m</i>	‘your dog’
	3sg	<i>asoʔ=naʔ</i>	* <i>aso-n naʔ</i>	‘his/her dog’

Short forms may also be used following a limited set of prepositions and other functional elements. For instance, in example (30a), a subject appears as a short form following a relative pronoun built from the root *avan* ‘place; location,’ used to form locative relative clauses. In example (30b), the preposition *dahin* ‘with’ takes a short form suffix to yield *dahiŋk*, meaning ‘with me.’

(30) Short form pronominals following *avan* ‘place’ and *dahin* ‘with’:

- a. *kəde anoʔ havam (< ha-avan-m) məle davən anən*
store which LOC-place-2SG.SHORT N-buy clothes DIST
‘At which store did you buy those clothes?’
- b. *lakeʔ anən te ha pasər dahiək (< dahin-k)*
man DIST go LOC store with-1SG.SHORT
‘That man goes to the store with me.’

The examples in (31) show that the cause-encoding preposition *kəlan* can host a short form pronoun, resulting in *kəlam* ‘because of you,’ as well as the related, inherently undergoer-oriented lexical verb meaning ‘to be caught,’ yielding *kəlak* ‘caught by me’.⁸

(31) Short form pronominals following *kəlan* ‘because of’:

- a. *akuʔ ləgak kəlam (< kəlan-m)*
1SG fall because-2SG.SHORT
‘I fell because of you.’
- b. *masiək anən kəlak (< kəlan-k)*
fish DIST be.caught-1SG.SHORT
‘That fish was caught by me.’

Although the first and second person short forms are clearly distinct from the regular genitives (= *kuj* : -*k* and = *kaʔ* : -*m*), the third person short and genitive forms cannot always be distinguished. In particular, following words that end with *-n* such as *dʒaʔan* ‘chin,’ we predict both the genitive enclitic and short suffixal forms to yield the homophonous *dʒaʔa-n naʔ* and *dʒaʔan=naʔ* for ‘his/her chin,’ as attested in (26). In other environments, however, the two can be clearly distinguished. For example, both *buloʔ=naʔ* and *bulo-n naʔ* are both acceptable for ‘his/her hair,’ with the surface distinction reflecting the fact that the final consonant is replaced by the short form but not by the genitive form. These limited contexts where the two forms can be distinguished play a crucial role in ascertaining the constraints on the short form, for instance through the distinction between *asoʔ=naʔ* and **aso-n naʔ* in (29) which supports our inalienability constraint.

In our data, both the regular genitive and the short form are licit as possessors for words like those in (26) and (28). For example, there appears to be no difference in usage between *buloʔ=kuj* and *bulo-k* for ‘my hair.’ However, it appears that short forms tend to be volunteered more often than the regular genitives when both are possible.

The phonological constraint on short forms—that they only appear on roots that end in either *ʔ* or *n*—is notable, as these phonemes don’t tend to form an exclusive natural class with one another. Comparative analysis, however, suggests that this phonological constraint arose through historical processes and was originally more natural. Blust 1997, citing data from a different variety of Rejang Kayan spoken in Uma Juman, describes the short forms as attaching only to words that end in a glottal stop. In that variety, surface *n*-final words such as *dʒaʔan* ‘chin’ are analyzable as morphologically complex (*dʒaʔaʔ-n*), and could be elicited without final *-n* in certain contexts. This follows from the historical source of *-n*, which, according to Blust, was the Proto-Malayo-Polynesian (PMP) genitive case marker **ni*. PMP **ni* reduced to *n* and became a suffix on vowel-final roots. A similar process also affected the PMP first person genitive **ku* and second person genitive **mu*, which surfaced as consonantal suffixes on vowel-final roots. In Kayan, vowel-final roots were

⁸ Clayre 2014 describes this process applying to a few other undergoer-oriented verbs (e.g. *kelenghe-k* ‘heard by me’, *hu’e-k* ‘sent by me’) in Baram river dialects, although no sentence data is provided. Such constructions are both restricted to only a few lexical verbs and include verbs that differ from those found in Uma Nyaving (*kəriəŋ* ‘to hear’ and *matə* ‘to send’), so we cannot generalize to Uma Nyaving based on Baram river Kayan. Nevertheless, there may be a larger set of inherently undergoer-oriented verbs that are able to take short form non-subject agents without overt voice marking.

later closed with a glottal stop, altering the phonological environments where *-k*, *-m*, and *-n* surfaced (Blust 2002).

This historical background can help us make sense of the synchronic phonological condition. We hypothesize that, in a pre-Kayan stage, *-k*, *-m*, and *-n* affixed only to vowel-final roots and appeared in their longer forms *ku* (later appearing as Kayan *kuj*), *-mu* (later replaced by Kayan *kaʔ*), and *-ni* (later fused with **ia* as Kayan *naʔ*) elsewhere.⁹ After the insertion of final glottal stops, however, they came to appear only on glottal stop-final roots. In Uma Nyaving, some of these words have been further reanalyzed with the *-n* element fully incorporated into the root. The licit form *dʒaʔaʔ* in the Uma Juman variety of Kayan, for example, is not recognized in Uma Nyaving and may only appear with a final *n* as in *dʒaʔan*. The same is true for other *n*-final words such as Uma Naving *matan* ‘eye’ and *harin* ‘sibling’, which are ungrammatical if pronounced with a final glottal stop, **mataʔ* or **hariʔ*, even though the latter are the grammatical bare stem forms in other varieties of Kayan (Blust 1997, Clayre 1996). We suggest that these historical processes explain the initially somewhat surprising phonological constraints on short forms in Uma Nyaving.

4 Passives

We now turn to the description of the different undergoer voice constructions in Kayan, which we will describe as different types of passives. We will refer to each passive construction based on the form of their passive marker: the passive with the marker *an*; the passives with the markers *ak* and *im*, which specifically encode first-singular and second-singular agents, respectively; and finally the periphrastic passive with the marker *gaʔ*, related to a lexical verb meaning ‘be hit’ or ‘get.’ We repeat our illustration of the basic word order of each of these passive constructions, for a basic bivalent verb, in (32), with the four passive markers in bold.

- (32) Word order schemas for active and passive clause types: = (2)
- | | | |
|----|--------------------------|--|
| a. | Active: | S/Agent (Aux) V O/Undergoer |
| b. | <i>an</i> passive: | S/Undergoer (Aux) <i>an</i> (Agent) V |
| c. | Agent-inflected passive: | S/Undergoer (Aux) <i>ak/im</i> V |
| d. | <i>gaʔ</i> passive: | S/Undergoer (Aux) <i>gaʔ</i> V (Agent) – or – ... (<i>kəlan</i> Agent) |

As the schemas in (32) indicate and we show in detail below, each passive clause type has its own distinct characteristics. Most notably, an optional agent appears between the *an* marker and verb in the *an* marked passive, but instead post verbally in the *gaʔ* passive, with a preposition *kəlan* if not immediately postverbal. The *gaʔ* passive is also distinct in expressing subject affectedness semantics, as we will show below.

Our choice to refer to these undergoer voice constructions as “passives” is supported by crosslinguistic work on the typology of passive constructions. For example, consider the three properties in (33) that Legate (2021) gives as characteristics of “canonical passives,” then discussing constructions satisfying a subset of these criteria as “noncanonical passives.” By these and similar criteria, all three constructions we describe here can be described as different types of (noncanonical) analytic passives.¹⁰

⁹ Ultimately, these short forms reflect the PMP genitive pronominals, **ku* ‘first person singular genitive’, **mu* ‘second person singular genitive’, as well as the genitive case marker **ni*. The reduction from **-ku*, **-mu*, and **ni* to only the consonant is also not unique to Kayan. Similar process of vowel deletion resulting in *-k*, *-m*, *-n* and similar paradigms are attested throughout the family, but are generally thought to be parallel developments (Blust 2013: pp. 483–486).

¹⁰ Among other prominent descriptions, Siewierska (2013) gives five properties, adding to Legate’s that “it contrasts with another construction, the active” (clearly true of the passive constructions here) and “the construction is pragmatically restricted relative to the active.” We are not able to comment here on any pragmatic restrictions associated with the *an*-marked and agent-inflected passives here, although we do discuss semantic effects of the use of *gaʔ*, in section 4.4 below. Zúñiga & Kittilä (2019: p. 83) gives four properties for “prototypical passives,” adding to Legate’s that “Syntactic valency is one less than in the active diathesis (e.g., the verb is monovalent when its active counterpart is bivalent),” which is implicitly satisfied by agent promotion.

(33) Properties of canonical passives: (Legate 2021: p. 158)

- a. Agent demotion: The agent is semantically present but is not syntactically present as a noun phrase in its characteristic thematic position. Instead, the agent is either interpreted as existential (‘someone’) or associated with a ‘by’-phrase.
- b. Theme promotion: The theme [undergoer] raises from its low syntactic position associated with the interpretation as a theme to the grammatical subject position.
- c. Morphological marking: The verbal morphology is distinct from the active voice.

First, both the *an*-marked construction and *gaʔ*-marked construction make agents optional, interpreted as existential (‘someone’) if unspecified. The agent-inflected constructions with *ak* and *im* are always specified for first-singular or second-singular agents, but we will argue that these constructions are best understood, both in their synchronic syntax and historical development, as morphological variants of the *an*-marked construction with a specified agent. We note that agents, if present, appear without any preposition in the *an* construction and optionally in the *gaʔ* construction; this, then, is the only sense in which these constructions may be described as non-canonical passives. Turning to the second criterion of “theme promotion,” all three passives in (32) involve promoting the theme (the conventional object of an active construction) to be in subject position. Finally, all three passives involve dedicated morphological marking, albeit not a verbal affix.¹¹

4.1 The *an*-Marked Passive

The most productive passive construction in Kayan involves the preverbal marker *an*, with the general word order “S/Undergoer (Aux) *an* (Agent) V...”. Passive constructions with this general word order are well attested amongst languages of central Borneo. For example, such constructions with strikingly similar passive marker forms are found in Kenyah Lebo’ Vo’ (*ən*; Smith, Laing & Tang 2022), Murik (*ən*; Clayre 1996: pp. 79–80), Penan (*an*; Clayre 1996: p. 79), Sa’ban (*an* or *aro*’; Clayre 1996: pp. 77–78, 80). Passive constructions with similar word orders, but with more divergent passive marker forms, are described for various Land Dayak languages (Connell 2013, Sommerlot to appear), Salako (Malayic; Adelaar 2005b), various Punan varieties, Kajang, and different dialects of Kenyah (Smith 2017c, Césard, Guerreiro & Soriente 2015).

Several examples of the *an*-marked passive are shown in (34). Example (34a) shows an active clause for comparison and (34b–34d) show various passives with the same verb *kuman* ‘eat.’ Example (34b) demonstrates an *an* passive with no overt agent. If expressed, the *an* passive agent must immediately follow the *an* marker, as a full pronoun, proper name, or full noun phrase in (34c) and as genitive pronouns in (34d). Notice that the reduced form of the third-singular pronoun, *haʔ*, is ungrammatical in (34d), but the genitive *naʔ* may be used.¹² Finally, (34e) shows that the agent of an *an* passive cannot be postverbal; this will contrast from the behavior of *gaʔ* passives, in section 4.4.

(34) *an*-constructions and agent form and position:

- a. {*akuʃ* / *ikaʔ* / *ihaʔ* / *naŋaŋ*} *kuman* *kanən* *anən*
 1SG / 2SG / 3SG / Nanyang eat rice DIST
 ‘I / You / He / Nanyang ate the rice.’ (active baseline.)
- b. *kanən* *anən* *an* *kuman*
 rice DIST PASS eat
 ‘That rice was eaten.’
- c. *kanən* *anən* *an* {*akuʃ* / *ikaʔ* / *ihaʔ* / *naŋaŋ* / *hinak* (< *hinan-k*)} *kuman*
 rice DIST PASS 1SG / 2SG / 3SG / Nanyang / mother-1SG.SHORT eat
 ‘That rice was eaten by me / you / him / Nanyang / my mother.’

¹¹ It is clear from the discussion in these works that “morphological marking” includes the use of additional functional morphology that does not affix to the lexical verb itself, such as the passive markers here. Siewierska (2013) and Zúñiga & Kittilä (2019: p. 91) refer to such passives as “analytic” passives.

¹² Recall that there is also a homophonous inanimate demonstrative *naʔ*, which is not an enclitic; see (20).

- d. *kanən anən an* {=*kuj* / =*ka?* / *na?* / *=*ha?*} *kuman*
 rice DIST PASS 1SG.GEN / 2SG.GEN / 3SG.GEN / *3SG.RED eat
 ‘That rice was eaten by me / you / him.’
- e. *akuj an* {*aso?* *anih*} *ma?ət* {**aso?* *anih*}
 1SG PASS dog PROX bite dog PROX
 ‘I was bitten by this dog.’

The passive marker and the agent, when expressed, must appear immediately before the verb. They cannot be separated from the verb and thus must follow any auxiliaries.

(35) Strict verb-adjacency of *an* and the agent:

- bakol anih* {**an=kuj*} *dʒan* {**an=kuj*} *pajan* {*an=kuj*} *pə-pənu*
 basket PROX {PASS=1SG.GEN} NEG {PASS=1SG.GEN} yet {PASS=1SG.GEN} CAUS-full
dahin baha
 with rice
 ‘The basket has not yet been filled with rice by me.’

Recall that Kayan has two causative constructions, one involving the prefix *pə-* and another involving an analytic causative marker *na*; see (17). The *an* passive can target the causee of a morphological causative, but not that of an analytic causative:

(36) *an* can target *pə-* causatives, but not *na* causatives:

- a. *iha?* *an=kuj* *pə-tudu*
 3SG PASS=1SG.GEN CAUS-sleep
 ‘He was put to sleep by me.’
- b. * *iha?* *an=kuj* *na tudu*
 3SG PASS=1SG.GEN make sleep
 Intended: ‘He was made to sleep by me.’

The subject (undergoer) may stay in postverbal position where the voice construction is used to highlight a result state, as in (37) as well as in (16b). This reflects the general availability of postverbal subjects for stative descriptions, as we saw in (5). Such sentences therefore have a word order at first glance similar to that of active clauses. However, they allow for omitted agents (37a), unlike active clauses, and expressed agents must appear between the *an* marker and verb (37b), showing that they pattern with other *an* examples.

(37) *an*-constructions with postverbal subjects:

- a. *au an təvəŋ kajo? anən*
 AUX PASS fell tree DIST
 ‘The tree is already cut down.’
- b. *an=kuj pə-takut lake? anən*
 PASS=1SG.GEN CAUS-afraid man DIST
 ‘That man is scared by me.’ / ‘I scare that man.’

Finally, we note that the *an* passive with a second-person agent is also used to express imperative clauses. The use of passive voice forms in imperatives has been noted for some other Austronesian languages (Blust 2013: pp. 507–508). We will see further examples of imperatives involving passive morphosyntax in Kayan in example (42) and in Merap (also Kayanic) in example (45).

(38) Imperative clause with *an*:

an=ka? *jinəŋ* *akuŋ*
 PASS=2SG.GEN look 1SG
 ‘Look at me!’

4.2 Agent-inflected passives

Next we turn to the undergoer voice constructions with the markers *ak* and *im*, which specifically encode first-singular and second-singular agents, respectively. Because they necessarily encode specific (speaker or addressee) agents, they do not exhibit the agent demotion property of canonical passives as in (33a); however, they do satisfy the two other properties of theme [undergoer] promotion and morphological marking, making them candidates for the label of “non-canonical passives.” Furthermore, we will show below that these agent-inflected constructions behave identically to *an*-marked passives in their syntactic distribution and range of use, modulo the pre-determined choice of agent. Despite their status as a non-canonical passive, then, we argue that these agent-inflected constructions with *ak* and *im* are thus best understood in relation to—and in some sense, morphologically opaque, specific instances of—the *an*-marked passive construction. In this section, we describe the properties of the agent-inflected passive; we then discuss the relationship between this construction and the *an* passive in detail in section 4.3.

We first present a minimal pair between a first-singular-inflected passive with *ak* versus an *an* construction with a first-singular agent pronoun, in (39). As this minimal pair illustrates, the agent inflected passive generally has the same word order as that of *an*-marked passives, described above. The agent-inflected passive marker must follow auxiliaries and be immediately before the verb, as seen in (40). The undergoer subject of this construction may be postverbal, as in (41).

(39) Parallel *an* and agent-inflected passives:

a. <i>ika?</i> <i>ak</i> <i>jinəŋ</i>	b. <i>ika?</i> <i>an=kuŋ</i> <i>jinəŋ</i>
2SG PASS.1SG see	2sg PASS=1SG.GEN see
‘You are seen by me.’	‘You are seen by me.’

(40) Agent-inflected passive marker must immediately precede verb:

bakol *anih* (**ak*) *dʒan* (**ak*) *pajan* (*ak*) *pə-pənu* *dahin* *baha*
 basket PROX (PASS.1SG) NEG (PASS.1SG) yet (PASS.1SG) CAUS-full with rice
 ‘The basket has not yet been filled with rice by me.’

(41) Inflected construction with postverbal subject:

ak *pə-takut* *lake?* *anən*
 PASS.1SG CAUS-afraid man DIST
 ‘I scare that man / that man is scared by me.’

Finally, we observe that the second-singular inflected marker *im* is also used in imperative clauses, again parallel to this exceptional use of the *an*-marker in (38).

(42) Imperative with agent-inflected passive marker:

im *jinəŋ* *akuŋ*
 PASS.2SG look 1SG
 ‘Look at me!’

In summary, we have yet to identify any syntactic differences between the inflected constructions and their corresponding *an* forms, other than the greater flexibility of *an* to encode different types of agents or no agent at all when not inflected.

Finally, we note that there is no identifiable third-person singular inflected passive marker. We return to this question again in the next section, where we address the historical relationship between the *an* passive and the agent-inflected passives.

4.3 On the relationship between the agent-inflected inflected and an-marked passives

Earlier work on Kayan offers only fragmented discussions of these passive constructions. Blust (1997), for example, discusses the markers *ak* and *im*, but makes no mention of the *an* passive with first- and second-person agents. He considers the agent-inflected passive forms to be independent innovations, and analyzes Uma Juman *ən* as a genitive marker, not as a voice marker. Clayre (2014) briefly discusses these forms for the Uma Pu and Uma Peliau dialects of Kayan, where she describes agent-inflected passive markers as derived through a synchronic process involving *an* and the short form pronominals: “First and second singular pronouns, *-k* and *-m*, of set IIB [our short forms] fuse with *en* to form *ek* and *im* respectively” (p. 141). (In the orthography of Clayre 2014, *e* = [ə].) The behavior of these short form inalienable possession suffixes was discussed earlier in §3.4 and are shown again here in (43), along with examples of the short forms attaching to the preposition *dahin* ‘with’, as well as a table showing the derivation of agent-inflected passive markers for Kayan Uma Nyaving utilizing the same suffixes in (44):

(43) Examples of short form pronouns:

a.	<i>matan</i>	‘eye’	b.	<i>dahin</i>	‘with’
	<i>mata-k</i>	‘my eye’		<i>dahi-k</i>	‘with me’
	<i>mata-m</i>	‘your eye’		<i>dahi-m</i>	‘with you’
	<i>mata-n na?</i>	‘her/his eye’		<i>dahi-n na?</i>	‘with her/him’

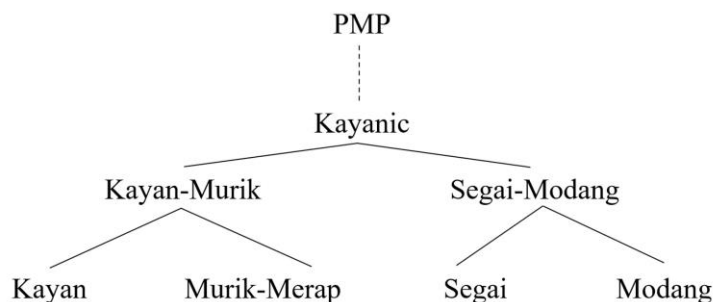
(44) Derivation for inflected markers:

an=kuj **an-k* → *ak*
an=ka? **an-m* → *im*
(*an=na?* **an-n na?* → *-n na?*)

We agree with Clayre’s suggestion that the agent-inflected passive markers are derived from short form pronominals suffixing to the general passive marker *an*. Although similar short forms are found in many languages of the region, Kayanic appears to be unique in that these short forms may also attach to a voice marker, making these Kayan examples particularly interesting. However, irregularities in the output of this morphological process pose some issues for a synchronic account. This complication is seen in the shape of the second person inflected marker, *im*, which under a synchronic analysis would be expected to surface as *am* (*an-m* → **am*). (The same issue arises with the dialects described in Clayre 2014; see the forms *en* vs second-singular *im* in the quote above.) We therefore analyze these agent-inflected passive markers as historically derived from short-form affixation, but synchronically unanalyzable forms.

We support this view by considering the forms of passive markers in related languages of the Kayanic group. For reference, we reproduce the family tree for the Kayanic subgroup in Figure 1.

Figure 1: Organization of the Kayanic subgroup (from Smith 2017b)



We begin the comparative analysis with Merap. In Merap, a Kayanic language spoken in North Kalimantan, we note a cognate inflected marker *am* which, like Uma Nyaving *im*, is used in imperative constructions as shown in (45).¹³

(45) Merap *am* inflected second person marker in an imperative sentence (Smith 2017c field notes):

hrejə am ηaʔaʔ bãə niə
 try PASS.2SG open door DIST
 ‘Try to open that door!’

Imperatives tend to utilize the passive voice, not only in Kayanic but in Austronesian languages more generally (Blust 2013: pp. 507–508). Merap *am* is derivable from *an-m* and could be a parallel innovation, motivated by the same morpho-phonological processes that motivated the inflected construction in Kayan, or it could be a shared retention from the common ancestor to Kayan and Merap. Either way, the Merap second person inflected marker is clearly derived from **an-m*, and suggests a similar source for Kayan *im*.

To explain the irregularity in vowel quality, we turn to Kelai, a Segai-Modang (Kayanic) language in East Kalimantan. In Kelai, the passive is formed with an immediately preverbal marker, just as in Kayan and Merap, but the shape of the Kelai passive marker is *en*, not *an* or *ən* (46).¹⁴

(46) Kelai *en* passive (Smith 2017c field notes):

*seʔ en (*koj) bəwp (koj)*
 3sg PASS 1SG hit 1SG
 ‘He is hit by me.’

A major difference between the Kelai *en* passive and Kayan *an* passive is that the passive agent in Kelai, if expressed, is postverbal; see (46). However, if Kayan (Uma Nyaving) *an*, Kayan (Uma Juman) *ən*, Merap *an*, and Kelai *en* all have the same historical source—a putative proto-Kayanic **in*—then the *im* irregularity in Kayan can be understood as an irregular retention of that earlier form. Kayan and Merap reflexes of the passive marker reflect vowel reduction, a common phonological process that targets functional vocabulary. We hypothesize that, in Proto-Kayan-Muric, the passive marker was therefore pronounced variably as **in* or **an*. During this period, the second person form stabilized as *im* in Kayan, while the basic passive marker and first person form stabilized as *an* and *ak*, respectively, then further reduced to *ən* and *ək* in the Baram river dialects described by Clayre (2014). The hypothesized historical derivation of the passive marker forms in Kayanic is shown in (47). Note that although we propose to view Kayan inflected passive markers as synchronically unanalyzable, we do not take a position on the status of similar constructions Merap or Kelai, which might be analyzed as reflecting active synchronic processes.

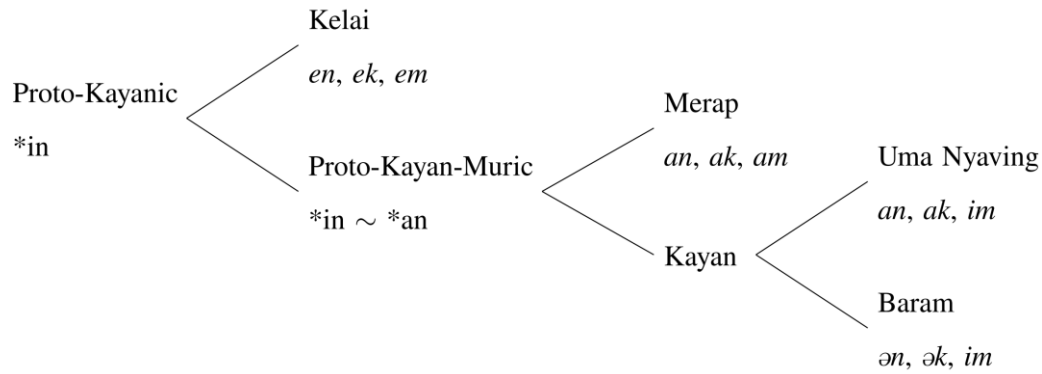
¹³ Note additionally that example (45) appears to be an instance of “crossed control,” as the passive agent (the addressee) is interpreted as the agent of ‘try.’ See for example Vander Klok & Paul 2021 and citations there for more on crossed control in other languages of the broader region.

¹⁴ Kelai *en* is also used as a genitive and dative, and may even combine with cognate short-forms as *ek* (en-k) ‘to me’ and *em* (en-m) ‘to you’ (Smith 2017c field notes).

(i) *seʔ hay pnejk ek (< en-k) / em (< en-m)*
 3sg give broom DAT-1SG.SHORT / DAT-1SG.SHORT
 ‘He gave the broom to me / to you.’

If such uses are a retention, it suggests the possibility that oblique prepositional phrases (a *by*-phrase), placed before the verb, may be a potential historical source for *an*-type passive constructions in Kayan as well.

(47) Suggested historical derivation for Kayanic passive markers:



4.4 Periphrastic passive with *gaʔ*

We have also encountered another, distinct passive construction with the marker *gaʔ*, which to our knowledge has not been described in previous work on Kayan. A first example is presented in (48). The *gaʔ* passive appears to be related to the lexical verb *gaʔ* meaning ‘to be hit, struck by, get’ as in (49). As a lexical verb, *gaʔ* typically takes an affected undergoer subject, with a causer or theme appearing post verbally. Due to this connection, we will gloss all instances of *gaʔ* as ‘get,’ including in the *gaʔ* passive construction.

(48) The *gaʔ* periphrastic passive:

akuʃ gaʔ maʔət ((kəlan) asoʔ anih)
 1SG get bite ((because) dog PROX)
 ‘I got bitten (by this dog).’

(49) The lexical verb *gaʔ*:

- a. *akuʃ gaʔ ikaʔ*
 1SG get 2SG
 Possible uses include: ‘I got bumped into by you.’ / ‘I caught (the flu) from you.’
- b. *akuʃ gaʔ batoʔ*
 1SG get rock
 ‘I got hit by a rock.’

Periphrastic passive constructions derived from lexical verbs meaning ‘get’ or ‘be hit’ are common in the region. Examples include constructions involving the form *sioʔ* in Kenyah (Smith, Laing & Tang 2022), *bok* in Punan (Soriente 2013, Césard, Guerreiro & Soriente 2015), *duʔk* in Bidayuh Biatah (Sommerlot, Smith & Erlewine 2023), and *kuniʔ* in Beaye (Sommerlot to appear), as well as of course the familiar *kena* construction in Malay (see e.g. Nomoto & Wahab 2012). We note that undergoer voice constructions with a similar lexical source are well attested across languages of Mainland Southeast Asia as well; see e.g. Gerner 2003 and Prasithratsint 2004.

Agents are optional in *gaʔ* constructions, as in the *an*-marked passive above, but they must appear following the verb, either immediately postverbally or else with the preposition *kəlan*, as in (49) and (50a). We show the unavailability of agents between *gaʔ* and the verb, with or without *kəlan*, in (50). We also note that pronominal agents must be introduced by *kəlan* rather than bare, as shown in (51).

(50) *gaʔ* agent must be postverbal:

- a. *akuʃ gaʔ mətəʔ (kəla-n naʔ)*
 1SG get hit because-3SG.SHORT
 ‘I got hit (by him).’

- b. * *akuʔ gaʔ* {=*haʔ* / =*naʔ* / *kəlan-naʔ*} *məteʔ*
 1SG get 3SG.RED / 3SG.GEN / because-3SG.SHORT hit
 Intended: ‘I got hit by him.’

(51) Pronominal agents require *kəlan*:

- a. * *ihaʔ gaʔ məteʔ akuʔ*
 3SG get hit 1SG
 Intended: ‘He got hit by me.’
- b. *ihaʔ gaʔ məteʔ kəlak* (< *kəlan-k*)
 1SG get hit because-1SG.SHORT
 ‘He got hit by me.’

As shown in example (52), the *gaʔ* marker cannot appear in a sentence already marked with *an*. In this case, our consultant explicitly noted that *gaʔ* and *an* cannot be used in the same sentence.

(52) *gaʔ* and *an* cannot cooccur:

- * *uma=naʔ gaʔ an dahaʔ nutoŋ* (< N-*tutoŋ*)
 house=3SG.GEN get PASS 3PL N-burn
 Intended: ‘His house got burned by them.’

Although most verbs may form passives with both *an* and *gaʔ*, some verbs seem to prefer *gaʔ* and are judged as ungrammatical if *an* is used. Such a case is shown here in (53) with the verb *maʔət* ‘to bite’ judged as ungrammatical with *an*.

(53) *gaʔ* preference over *an*:

- a. *akuʔ gaʔ maʔət*
 1SG get bite
 ‘I got bitten.’
- b. * *akuʔ an maʔət*
 1SG PASS bite

Our investigation to date suggests that the *gaʔ* passive may often express unexpected, sudden, unpredictable, or adversative semantics, whereas the *an* and related agent-inflected passives carry no such semantics. In this regard, the *gaʔ* passive appears to be similar to the Malay periphrastic passive formed with *kena* (Nomoto & Wahab 2012). Our consultant’s comments and translations suggest that, at least in some examples, the use of *gaʔ* adds an adversative flavor to a sentence indicating that someone was put down to sleep. A contrast between the *an* passive with a neutral context in (54a) and the *gaʔ* passive in (54b) was explicitly pointed out by our speaker in elicitation.

(54) Adversative readings of *gaʔ*:

- a. *ihaʔ an=kuj pə-tudu*
 3SG PASS=1SG.GEN CAUS-sleep
 ‘He was put to sleep by me.’ (neutral)
- b. *ihaʔ gaʔ pə-tudu kəlak* (< *kəlan-k*)
 3SG get CAUS-sleep because-1SG.SHORT
 ‘He was put to sleep by me.’ (adversative)

The marker *gaʔ* was rejected with certain verbs where our consultant could not reconcile the ‘surprising’, ‘unpredictable’, or ‘unexpected’ interpretation from *gaʔ* with the event described. For example, in (55) the verb *majo* ‘to clean something’ cannot be paired with *gaʔ* (55c) but is accepted with *an* (55b). The clashing semantics in (55c) was singled out by our consultant as causing the ungrammaticality, since one does not ‘unpredictably,’ ‘unexpectedly,’ or ‘adversely’ clean one’s house. These interpretations are tenuous, however, and more research is needed to fully flesh out these restrictions.

(55) Mirative interpretation of *gaʔ*:

- a. *akuʃ majo uma=kuʃ*
 1SG clean house=1SG.GEN
 ‘I cleaned my house.’
- b. *uma=kuʃ an dahaʔ majo*
 house=1SG.GEN PASS 3PL clean
 ‘My house was cleaned by him.’
- c. * *uma=kuʃ gaʔ majo*
 house=1SG.GEN get clean
 Intended: ‘My house got cleaned.’

With certain verbs, the unpredictable or surprising semantics of *gaʔ* is redundant, resulting in no difference in interpretation between the *an* and *gaʔ* passives in certain contexts like that in (56) with the verb *pə-takut* ‘to frighten’.

(56) Neutral interpretation of *gaʔ*:

- a. *akuʃ an=kaʔ pə-takut*
 1SG PASS=2SG.GEN CAUS-afraid
- b. *akuʃ gaʔ pə-takut kəlam (< kəlan-m)*
 1SG get CAUS-afraid because-2SG.SHORT
 Both: ‘I get scared by you.’

In summary, such evidence points to some sense of unexpectedness or adversity associated with *gaʔ*, in contrast to the other constructions which seem to be more neutral in their semantics and pragmatics. These effects warrant further investigation in the future.

5 Conclusion

We have provided an updated description of Kayan voice, which we describe as having three passive voice strategies. Two of these, the *an*-marked and the agent-inflected passives, are historically related through a morphophonological process involving pronominal suffixes whose affixal boundaries have been obscured through historical sound change, resulting in the contemporary agent-inflected forms *ak* and *im*. A third strategy involves the newly-described *gaʔ* construction, formed through the grammaticalization of a lexical verb meaning ‘to be hit, struck by, get,’ a common source for periphrastic passives in the region. The *gaʔ* passive construction sometimes encodes unexpected or adversative semantics, but more work is needed to fully understand its use.

Kayan voice is distinct from other, better-described Austronesian voice systems, such as in so-called Philippine-type or Indonesian-type languages (see e.g. Himmelmann 2005), nor does it reflect a “transitional” position between the two (see e.g. Hemmings 2015, Utsumi 2022). Instead, the Kayan voice system patterns with several nearby languages, like other north and central Sarawak languages, in having features described as ‘Central Bornean-type’: a two-way voice system, analytic passives, distinct genitive pronouns but otherwise no case distinctions, and an absence of applicative suffixes. Languages that have been included under the Central Bornean umbrella include Berawan, Penan, Melanau, Lun Bawang, varieties of Kenyah, Murik,

Sa'ban, and Land Dayak as well as previously described varieties of Kayan (Clayre 1996, 2014, Soriente 2013, Kroeger & Smith to appear). However, there is considerable variation within this category (Sommerlot to appear), and not all these languages share some of the features described here for Kayan. First, all three passive constructions in Uma Nyaving Kayan utilize pronominal agents that occur in their genitive form; not all Central Bornean-type languages exemplify this, but it has been described for Berawan, Penan, Melanau, Murik, and Kenyah (Clayre 1996, 2014, Kroeger & Smith to appear). Second, the *an*-marked passive in Kayan exhibits “S/Undergoer PASS (Agent) V...” order,¹⁵ which is additionally found in Murik, Kenyah, Sa'ban, and Lun Bawang (Clayre 1996, 2014) as well as several Land Dayak languages (Connell 2013, Sommerlot to appear). Other languages, however, have more variable word order than Kayan, and others instead exhibit “S/Undergoer PASS V (Agent) ...” order similar to the *ga?* passive; see for example Berawan and Melanau (Clayre 1997, Blust 1988).

We hope that our detailed description of Kayan Uma Nyaving voice offers a template for future investigations into the voice systems of Central Bornean-type languages. Further work into the grammars of languages of the region are necessary in order to uncover the uniformities and points of diversity amongst these languages, as well as their synchronic and diachronic relations to the better studied Philippine- and Indonesian-type grammars.

Acknowledgements

We gratefully acknowledge our language consultant Elrica Anak Blabun for sharing her language with us. This work was also informed by discussions with audiences at SEALS 32 and APLL 15. This work was supported by the Singapore Ministry of Education and the National University of Singapore under grants A-8000132-00-00 and A-8001136-00-00. Erlewine's work was supported by a fellowship at the Helsinki Collegium for Advanced Studies.

References

- Adelaar, Alexander & Nikolaus P. Himmelmann (eds.). 2005. *The Austronesian languages of Asia and Madagascar*. Routledge.
- Adelaar, K. Alexander. 2005a. The Austronesian languages of Asia and Madagascar: a historical perspective. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 1–43. Routledge.
- Adelaar, K. Alexander. 2005b. *Salako or Badamea: sketch grammar, text and lexicon of a Kanayatn dialect in West Borneo*. Vol. 2 (Frankfurter Forschungen zu Südostasien). Wiesbaden: Harrassowitz. vii+328.
- Blust, Robert. 1988. Sketches of the morphology and phonology of Bornean languages 2: Mukah Melanau. In Hein Steinhauer (ed.), *Papers in western Austronesian linguistics*, 151–216. Canberra: Research School of Pacific Studies, Department of Linguistics, Australian National University.
- Blust, Robert. 1997. Sketches of the morphology and phonology of Bornean languages 1: Uma Juman (Kayan). In Hein Steinhauer (ed.), *Papers in Bornean and western Austronesian languages*, 2, 7–122. Canberra: Pacific Linguistics.
- Blust, Robert. 2002. Formalism or phoneyism? The history of Kayan final glottal stop. In K. Alexander Adelaar & Robert Blust (eds.), *Between worlds: linguistic papers in memory of David John Prentice*, 29–37. Canberra: Pacific Linguistics.
- Blust, Robert. 2004. Austronesian nasal substitution: A survey. *Oceanic Linguistics* 43.1:73–148.
- Blust, Robert. 2013. *The Austronesian languages*. 2nd ed. Canberra: Pacific Linguistics.
- Branan, Kenyon. to appear. *The Left Edge Ban*. Oxford University Press.
- Césard, Nicolas, Antonio Guerreiro & Antonia Soriente. 2015. *Petualangan Unjung dan Mbui Kuvong Sastra lisan dan Kamus Punan Tuvu' dari Kalimantan*. Jakarta: École Française d'Extrême-Orient & Kepustakaan Populer Gramedia.
- Clayre, Beatrice. 1996. The changing face of focus in the languages of Borneo. In Hein Steinhauer (ed.), *Papers in Austronesian linguistics*, 3:51–88. Canberra: Pacific Linguistics.

¹⁵ On our account, the agent-inflected passives instantiate this order as well.

- Clayre, Beatrice. 1997. Verb affixes in Berawan. In Cecilia Odé & Wim Stokhof (eds.), *Proceedings of the Seventh International Conference on Austronesian Linguistics*, 233–255. Leiden, The Netherlands: Brill.
- Clayre, Beatrice. 2014. A preliminary typology of the languages of middle Borneo. In Peter Sercombe, Michael Boutin & Adrian Clynes (eds.), *Advances in research on linguistic and cultural practices in Borneo* (Monograph series 14). Phillips, Maine: Borneo Research Council.
- Clayre, Beatrice M. & L. E. Cubit. 1974. An outline of Kayan grammar. *Sarawak Museum Journal* 22:43–91.
- Cole, Peter, Gabriella Hermon & Yanti. 2008. Voice in Malay/Indonesian. *Lingua* 118.10:1500–1553.
- Collins, Millard A., Virginia R. Collins & Sulfilix A. Hashim. 2001. *Mapun-English dictionary*. Manila: Summer Institute of Linguistics - Philippines, Inc.
- Connell, Timothy M. 2013. A sketch grammar of Matéq: a Land Dayak language of West Kalimantan, Indonesia. University of Canterbury MA thesis.
- Gerner, Matthias. 2003. Passive of affect in Kam (Dong) and other Kadai languages: the missing link from synchrony. *Australian Journal of Linguistics* 23.1:35–70.
- Haspelmath, Martin & Balthazar Bickel. 2015. The Leipzig glossing rules.
- Hemmings, Charlotte. 2015. Kelabit voice: Philippine-type, Indonesian-type or something a bit different. *Transactions of the Philological Society* 113:383–405.
- Himmelman, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Alexander Adelaar & Nikolaus P. Himmelman (eds.), *The Austronesian languages of Asia and Madagascar*, 110–181. Routledge.
- Kroeger, Paul & Alexander D. Smith. to appear. Non-Malayic languages of Borneo. In K. Alexander Adelaar & Antoinette Schapper (eds.), *The Oxford guide to the Malayo-Polynesian languages of Southeast Asia*. Oxford University Press.
- Legate, Julie Anne. 2021. Noncanonical passives: a typology of voices in an impoverished universal grammar. *Annual Review of Linguistics* 7:157–176.
- Lobel, Jason. 2016. *North Borneo sourcebook*. University of Hawai'i Press.
- McCarthy, John & Alan Prince. 1993. Generalized alignment. In Geert Booij & Jaap van Marle (eds.), *Yearbook of morphology*, 79–153. Dordrecht: Kluwer.
- Nomoto, Hiroki & Kartini Abd. Wahab. 2012. *Kena* adversative passives in Malay, funny control, and covert voice alternation. *Oceanic Linguistics* 51:360–386.
- Prasithratsint, Amara. 2004. The adversative passive marker as a prominent areal feature of Southeast Asian languages. In *Papers from the Eleventh Annual Meeting of the Southeast Asian Linguistics Society*, ed. by Somsonge Burusphat, 583–598. Tempe, Arizona: Arizona State University, Program for Southeast Asian Studies.
- Ross, Malcolm. 2009. Proto-Austronesian verbal morphology: a reappraisal. In K. Alexander Adelaar & Andrew Pawley (eds.), *Austronesian historical linguistics and culture history: a festschrift for Robert Blust*, 295–326. Canberra: Pacific Linguistics.
- Rubino, Carl Ralph Galvez. 2000. *Ilocano dictionary and grammar: Ilocano-English, English-Ilocano* (PALI Language Texts). Honolulu: University of Hawai'i Press.
- Siewierska, Anna. 2013. Passive constructions. In Matthew S. Dryer & Martin Haspelmath (eds.), *The world atlas of language structures online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.
- Smith, Alexander D. 2017a. Reconstructing Proto-Kenyah pronouns and the development of a true five number system. In Hsiu-chuan Liao (ed.), *Issues in Austronesian historical linguistics* (JSEALS Special Publications 1), 48–66. Honolulu: University of Hawai'i Press.
- Smith, Alexander D. 2017b. *The languages of Borneo: a comprehensive classification*. Doctoral Dissertation, Department of Linguistics, University of Hawai'i.
- Smith, Alexander D. 2017c. *Fieldnotes from Borneo collected between 2014 and 2017*. ms.

- Smith, Alexander D., Louise Ping Laing & Xuepei Tang. 2022. *Kenyah Lebo' Vo', Long Tikan elicitation sessions*. Kaipuleohone Language Archive. <https://hdl.handle.net/10125/102361>
- Smith, Alexander D., Carly J. Sommerlot & Michael Yoshitaka Erlewine. 2023. *Uma Nyaving Kayan elicitation sessions*. Kaipuleohone Digital Language Archive. <https://hdl.handle.net/10125/104786>
- Sneddon, James Neil. 1996. *Indonesian: a comprehensive grammar*. Routledge.
- Sommerlot, Carly J. 2020. *On the syntax of West Kalimantan: asymmetries and A'-movement in Malayic and Land Dayak languages*. Doctoral Dissertation, Department of Linguistics, University of Texas, Arlington.
- Sommerlot, Carly J. to appear. Voice in Land Dayak languages. *Oceanic Linguistics*.
- Sommerlot, Carly J., Alexander D. Smith & Michael Yoshitaka Erlewine. 2023. *Biatah, Bidayuh elicitation sessions*. Kaipuleohone Language Archive. <https://hdl.handle.net/10125/104690>
- Soriente, Antonia. 2013. Undergoer voice in Borneo Penan, Punan, Kenyah, and Kayan languages. In K. Alexander Adelaar (ed.), *Voice variation in Austronesian languages: linguistic studies of Indonesian and other languages in Indonesia*, 175–203. NUSA 54.
- Utsumi, Atsuko. 2022. Semantic typology of voice systems in Western Malayo-Polynesian languages. *STUF: Language Typology and Universals* 75.1:129–160.
- Vander Klok, Jozina & Ileana Paul. 2021. What categorial ambiguity doesn't tell us about crossed control: commentary on Jeoung 2020. *Language* 97.3:e276–e292.
- Wolff, John. 1973. Verbal inflection in Proto-Austronesian. In Andrew B. Gonzales (ed.), *Parangal kay Cecilio Lopez: essays in honor of Cecilio Lopez on his seventy-fifth birthday*, 71–91. Quezon City: Linguistic Society of the Philippines.
- Zúñiga, Fernando & Seppo Kittilä. 2019. *Grammatical voice*. Cambridge University Press.