Charlotte Hemmings

Information structure and syntactic choices in Kelabit

Abstract: This paper investigates the role of information structure in determining three syntactic choices in the Kelabit language of Northern Sarawak: (1) voice choice within a symmetrical voice system, (2) word order within each voice construction and (3) differential case-marking of the undergoer voice (UV) actor. In each case, information structure is shown to be relevant: word order can be used to place focus before background information, nominative case is used for the UV actor when the actor is focused/contrasted and the undergoer is topic, and an unusual mapping between arguments and information structure roles (e.g. a focused actor or an undergoer topic) can trigger the choice of the respective voice construction. However, it is the relative prominence or unexpectedness of information that is important, rather than the status of the subject as topic or focus. Hence, there is no one-to-one link between particular grammatical functions and information structure roles. Instead, the different syntactic choices combine and interact to express information in context. This supports the view that prominence is relational and suggests that information structure is an important component of prominence in Western Austronesian.

1 Introduction

This paper investigates the role of information structure in determining syntactic choices in Kelabit, a Western Austronesian language spoken in Northern Sarawak (Kroeger 1998: 255). Like other Western Austronesian languages, Kelabit has a system of symmetrical voice alternations, in which the mapping of arguments to functions changes without affecting the resulting transitivity of the clause. Consequently, speakers make a choice of which voice construction to use when expressing notionally transitive events. Within each voice construction, speakers also have a choice of word order. Finally, there is a choice of case-marking for the actor in an undergoer voice construction. The aim of the paper is to illustrate how information structure interacts with these choices and to explore the implications of this for the notion of prominence in Western Austronesian.

Charlotte Hemmings, University of Oxford

https://doi.org/10.1515/9783110730753-004

The paper is structured as follows. §2 provides some background on Kelabit and introduces the voice system, flexible word order and differential case marking. §3 defines information structure in relation to other levels of prominence. §4 sets out a methodology for exploring the interaction between information structure and syntactic choices. §5 discusses the interaction between information structure and word order. §6 discusses the interaction between information structure and differential marking. §7 discusses the interaction between information structure and voice and §8 concludes.

2 Syntactic choices in Kelabit

Kelabit is a Western Austronesian language spoken by roughly 2,000–6,000 speakers, predominantly in Northern Sarawak, Malaysian Borneo. It is most closely related to the Lun Bawang and Sa'ban languages, which collectively form a subgroup, variously labelled as Apad Uat, Kelabitic and/or Dayic (Smith 2017, Blust 2013, Kroeger 1998, Hemmings 2020). The data used in this paper was collected over several years of fieldwork, and comprises elicitation as well as a naturalistic text corpus. The elicited data was collected in Bario. The naturalistic text corpus was collected from a variety of speakers in Bario, Pa Umor and Pa Dalih (see Hemmings 2017).

Kelabit has a number of interesting syntactic features which are typical of the languages of central Borneo (Clayre 2014, Soriente 2013). In this paper, I focus on the system of symmetrical voice, flexible word order across the voice constructions and the pattern of differential case marking found in undergoer voice. First and foremost, Kelabit like most other languages of the region, has a symmetrical voice system. This involves alternations in verbal morphology that indicate different mappings of arguments to functions but do not affect the syntactic transitivity of the resulting clause. In other words, symmetrical voice languages have more than one type of transitive clause with two or more core arguments (Himmelmann 2005, Riesberg 2014). This can be illustrated from Kelabit in (1) using the root *laak* 'cook'. Both (1a) and (1b) express the same event in which a mother is cooking rice. However, they differ in the verbal morphology and this corresponds to a different mapping of arguments to functions. To keep track of the data, I will refer to arguments using the macro-roles of ACTOR and UNDERGOER. These remain constant and refer to 'her mother' and 'rice' respectively in (1). I will refer to the function of arguments, which differs depending on the voice construction, using the terms SUBJECT

and NON-SUBJECT CORE ARGUMENT.[1] Throughout the paper, subjects are indicated in **bold**, and non-subject core arguments underlined:[2]

(1) Symmetrical Voice
   a. Actor Voice (AV)
      Nengelaak *nuba'* ***tesineh*** ***nedih.***
      PFV.AV.cook rice mother 3SG.POSS
      'Her mother cooked rice.'
   b. Undergoer Voice (UV)
      Linaak *tesineh* *nedih* ***nuba'.***
      PFV.UV.cook mother 3SG.POSS rice
      'Her mother cooked (the) rice.' (elicitation, fieldnotes)

The actor voice construction in (1a) is marked with a nasal prefix *neN-* that attaches to the root word *laak* 'cook' to form *nengelaak* 'AV.PFV.cook'.[3] The actor (*tesineh nedih* 'her mother') is mapped to subject and the undergoer (*nuba'* 'rice') is a non-subject core argument. In the undergoer voice construction in (1b), however, the undergoer (*nuba'*) is mapped to subject and the actor (*tesineh nedih*) is a non-subject core argument. This is formally marked by the *-in-* infix which results in the verb form *linaak* 'UV.PFV.cook'.[4] The function of arguments is reflected in their position: non-subject core arguments typically appear directly following the verb, whilst

1 I use the term subject for ease of presentation and for reasons further elaborated in Hemmings (2021a). I use the term non-subject core argument, rather than object, given that it includes non-subject actors in UV and actor objects are cross-linguistically rare. However, it should be noted that grammatical functions have long been controversial in Western Austronesian (Bickel 2010, Dryer 1997, Schachter 1976, 1996). Both 'subject' and 'non-subject core argument' can be understood as descriptive labels in the sense of Haspelmath (2010).

2 Glosses: 1 = first person, 2 = second person, 3 = third person, AV = actor voice, DEM = demonstrative, EMPH = emphatic, GEN = genitive, HT = hanging topic, IRR = irrealis, NEG = negation, NOM = nominative, PFV = perfective, PL = plural, POSS = possessor, PT = particle, REDUP = reduplication, REL = relativiser, SG = singular, UV = undergoer voice.

3 The *ne-* prefix marks perfective aspect. The nasal prefix *N-* is realised as *nge-* before roots beginning with approximants, *ng-* before vowel-initial roots (e.g. *imet* to *ngimet* 'AV.hold') and homorganic nasal substitution before roots beginning with obstruents (e.g. *perek* to *merek* 'AV.squeeze').

4 The *-in-* infix marks undergoer voice in the perfective aspect and has several allomorphs. If the root is vowel-initial, the prefix *n-* is used (e.g. *imet* to *nimet* 'UV.PFV.hold'). If the first vowel is schwa, i-ablaut is used (e.g. *perek* to *pirek* 'UV.PFV.squeeze') For further discussion of AV and UV allomorphs as well as the irrealis paradigms, see Hemmings (2016).

subjects have greater word order flexibility.[5] They also share a series of morphological and syntactic properties. For example, the AV actor and UV undergoer share the ability to be relativised on, raised, and controlled (Hemmings 2021a). The AV undergoer and UV actor do not have these subject properties but share a number of core argument properties distinguishing them from adjuncts and obliques. For example, the AV undergoer and UV actor are realised as nominal arguments and form a constituent with the verb which time adverbials cannot intervene between. In contrast, obliques are typically realised as PPs and can be separated from the verb + non-subject core argument complex by time adverbials (Hemmings 2021a). Consequently, I will assume the mapping from arguments to functions in Tab. 1.[6]

Tab. 1: Grammatical functions in AV and UV.

	subject	**non-subject core**
Actor voice	actor	undergoer
Undergoer voice	undergoer	actor

Thus, one syntactic choice available to speakers of Kelabit is the choice of voice construction, which allows for different mappings between arguments (actor and undergoer) and functions (subject and non-subject core).

Another choice is that subjects have relatively flexible word order. Whilst the non-subject core argument always follows the verb, and typically appears in the immediately post-verbal position, the subject argument can appear either pre-verbally, as in (2a) and (2d), or post-verbally, as in (2b), (2c) and (2e). When the subject is post-verbal it typically follows the non-subject core argument, as in (2b) and (2e). The only exception to this generalisation is that the actor subject can intervene between the verb and the non-subject undergoer in AV constructions, as in (2c). It is not possible, however, for the undergoer subject to intervene between the verb and non-subject actor in UV constructions, as shown in (2f). To capture these patterns, I will discuss word order in relation to the grammatical functions assumed in Tab. 1, referring to subjects as S, and non-subject core arguments as O. Note that this presentation differs from many typological approaches to word order, where S and O are semantically defined (see e.g. Greenberg 1963, Dryer 2013, Riesberg,

5 Though note that the actor can intervene between the verb and the undergoer in AV constructions, as illustrated in (2c) below. Regardless of the word order configuration, the identification of grammatical functions is unambiguous given the verbal morphology.

6 Kelabit also has an instrumental voice construction in which the instrument is mapped to subject and the verb takes the *peneN-* affix in realis/perfective. This construction is marginal and is not further discussed in this paper.

Malcher & Himmelmann 2019). Assuming these definitions, the possible orders in Kelabit are as follows:

(2) Actor Voice

a. SVO
La'ih ***sineh*** *ne-kuman* *bua'* *kaber.*
man DEM PFV-AV.eat fruit pineapple
'The man ate pineapple.'

b. VOS
Ne-kuman *bua'* *kaber* ***la'ih*** ***sineh.***
PFV-AV.eat fruit pineapple man DEM
'The man ate pineapple.'

c. VSO
Ne-kuman ***la'ih*** ***sineh*** *bua'* *kaber.*
PFV-AV.eat man DEM fruit pineapple
'The man ate pineapple.'

Undergoer Voice

d. SVO
Bua' ***kaber*** *kinan* *la'ih* *sineh.*
fruit pineapple UV.PFV.eat man DEM
'The man ate pineapple.'

e. VOS
kinan *la'ih* *sineh* ***bua'*** ***kaber*** ***dih.***
UV.PFV.eat man DEM fruit pineapple DEM
'The man ate the pineapple.'

f. VSO
kinan* ***bua'* ***kaber*** *la'ih* *sineh.*
UV.PFV.eat fruit pineapple man DEM
For: 'The man ate pineapple.' (elicitation, fieldnotes)

The possible word orders in (2) are not equally frequent. Firstly, AV constructions outnumber UV constructions across a range of genres, including folk stories and news reports (see Hemmings 2021b). Secondly, the frequency of word orders differs depending on the voice construction. For AV, the most frequent word order is SVO. Verb-initial orders, and particularly VSO constructions, are marked choices. For example, in a corpus of narratives collected using the pear story stimulus, of 74 AV constructions, only 2 (or 3%) had VSO order (Hemmings 2016: 457). For UV, however, the frequency of word orders differs according to text genre. In folk stories, UV clauses tend to have VOS order, which is typical of more conservative Western

Austronesian languages in the Philippines and Taiwan (Donohue 2007, Riesberg, Malcher & Himmelmann 2019). In other genres, such as news reports, SVO orders are most frequent (Hemmings 2016).[7] Post-verbal subjects tend to be definite. Consequently, VOS order in (2e) is acceptable when the undergoer is modified by a determiner, *bua' kaber dih* 'fruit pineapple DEM', but less natural with generic and indefinite undergoers, as in (1b).[8] Definite arguments tend to be more topic-worthy (Lambrecht 1994), which already suggests that information structure plays a role in determining the choice of word order (see §5).

Apart from alternations in basic order, there are two other constructions that allow information to appear clause-initially. The first is a hanging topic construction, in which the hanging topic is co-referenced by a pronoun in the main clause. The hanging topic can co-reference the subject, as in (3a-b). However, it may also correspond to a non-subject actor, (3c), and need not be linked to an argument of the main clause at all, as in (3d):

(3) Hanging Topic

a. *[Paul kedieh]$_{HT}$ nekuman bua' ebpuk* ***teh***[9]***=ieh***
Paul EMPH.3SG AV.PFV.eat fruit passion PT=3SG.NOM
'As for Paul, he ate passionfruit.'

7 Note that there may be translation influence as the news report data was translated from newspapers in English and Malay – both predominantly SVO languages – before being broadcast on the local radio station.

8 The example in (1b) was elicited but UV constructions do occur with indefinite undergoers in final position in the corpus (see Hemmings 2021b). These can occur either with the indefinite article *edteh* 'one/a' or as bare nouns, as in (1b). With bare nouns it appears that UV undergoers can have either a generic or definite interpretation. Overall, there is a preference for undergoers in UV constructions, particularly those in clause-final position, to be definite (see Hemmings 2021b: 602). This is common in Western Austronesian languages (see e.g. Latrouite & Riester 2018: 248).

9 As discussed in Hemmings (2016: 193–197), Kelabit has a series of sentence particles that typically occur in second-position, clause-finally or in both positions and are glossed PT. The exact semantics of these particles remains to be better understood, but they appear to serve a discourse function. Among the particles, *teh*, *neh* and *peh*, are particularly common and could be thought to express information structure, e.g. by separating focus and background in the case of *teh* (see §5) and expressing additive focus ("also") in the case of *peh* (see §7). Like pre-nominal case-marking particles in more conservative Western Austronesian languages, these particles typically precede the subject, and cannot occur before the non-subject core argument. However, unlike case-marking particles they do not occur in every clause. Hence, they are more like discourse particles than subject markers.

b. *[Bua' ebpuk suk na'ah ih]$_{HT}$ kinan Paul* ***neh***[10]***=idih***
fruit passion REL afore PT UV.PFV.eat Paul PT=DEM
'As for the passionfruit, Paul ate it.'

c. *[Paul kedieh]$_{HT}$ kinan neh* ***bua' ebpuk***
Paul 3SG.EMPH UV.PFV.eat 3SG.GEN fruit passion
'As for Paul, he ate the passionfruit.' (elicitation, fieldnotes)

d. *[Tapi bulu' sineh]$_{HT}$ kiteb neh* ***pa'up bukuh ih***[11]
but bamboo DEM UV.PFV.cut 3SG.GEN end edge DEM
'But that bamboo, he had cut both ends off.' (text, BAR17082014CH_08)

The topic phrase is typically separated from the rest of the clause by a short pause and information contained within the topic phrase is often contrasted.

The second construction is sometimes known as focus-fronting. It involves an initial constituent, followed by the particle *teh* and the rest of the clause:

(4) Focus Fronting

a. ***Peter*** *teh (suk) kuman bua' kaber*
Peter PT REL AV.eat fruit pineapple
'It was Peter who ate the pineapple.'

b. ***Bua' kaber*** *teh kinan Peter*
fruit pineapple PT UV.PFV.eat Peter
'It was pineapple that Peter ate.' (elicitation, fieldnotes)

The initial (or "fronted") phrase may be the subject, as in (4). However, it can also be an adjunct or larger constituent, as discussed in §5. The only constituent that cannot occur in initial position is the non-subject core argument when it is not part of the verb phrase.[12] Whenever *teh* occurs in verb-initial clauses, or clauses with an initial adjunct, it always precedes the subject (see Hemmings 2021a). For

10 The form *neh* is highly multifunctional since it is a sentence particle but also the GEN form of the 3SG pronoun (see below) as well as the distal demonstrative, shortened from *ineh* (which forms a paradigm with *inih/nih* 'proximal' and *idih/dih* 'medial'). These can be distinguished as homonyms on the basis that they differ in function as well as distribution: particles appear in second-position preceding nouns, pronouns appear in place of nouns and demonstratives follow the nouns they modify or occur clause-initially to express progressive action or serve a clause-linking function (see Hemmings 2016: 175–176).

11 The particle *ih* appears to have developed from the medial demonstrative *dih*. It follows the noun it modifies and indicates that the noun in question is definite/specific (see Hemmings 2016).

12 I assume that obliques could also appear in the initial position but cannot find examples in the corpus to confirm this. Note that obliques can be adjunct-fronted before the subject without *teh* (see Hemmings 2021a).

this reason, it is not clear whether focus-fronting constructions involve variation in basic word order, as in (2), but with the particle *teh* additionally marking the end of the focus constituent, or whether the structures in (4) should be analysed as (pseudo)-clefts with nominal predicates and a subject formed from a headless relative clause (cf. Potsdam 2006).[13] Whatever the correct analysis of the structure, it is clear that different constituents can appear in the initial position of the construction. Consequently, another choice available to speakers is to alter the order that information is presented in – either by mapping a different argument to subject, or by using hanging topic and focus fronting constructions.

A further choice available to speakers is case-marking. In more conservative Western Austronesian languages, case-marking appears on all nominal arguments and directly reflects the function of an argument within the symmetrical voice system. The typical case distinctions can be summarised in (5):[14]

(5) Western Austronesian Case-Marking
   a. NOM – subjects (e.g. AV actor, UV undergoer etc.)
   b. GEN – non-subject actors (e.g. UV actor)
   c. OBL – obliques and definite non-subject undergoers (e.g. AV undergoer)

Kelabit does not have case-marking on nominal arguments, whose function is instead reflected in word order, as outlined above. However, it maintains a reduced system of case-marking in a subset of the pronouns, 1SG, 2SG, 3SG and 3PL, with a formal opposition between NOM and GEN forms, as seen in Tab 2.[15]

13 Perhaps an argument in favour of the pseudo-cleft analysis for the examples in (4) is the fact that the relativiser *suk* optionally appears. As discussed in Hemmings (2016: 188–189), the main function of *suk* is to signal a relative clause in which the referent of a relativised noun is singular and specific. Plural, indefinite and non-specific referents are relativised using the relativiser *nuk*. Both particles can also be used as nominalisers, forming headless relative clauses. Consequently, under a pseudo-cleft analysis (4a) could be translated as 'the one who ate pineapple was Peter'. This is similar to the function of *yang* in Standard Indonesian (see e.g. Musgrave 2001: 13–14). Note that *suk* would probably be ungrammatical in (4b) since the pineapple is generic and non-specific, but can occur in focus-fronting constructions where the verb is UV, such as (14b). Since *suk/nuk* are not obligatory in focus-fronting constructions, however, it is possible that we are dealing with different types of construction when they occur.

14 Note that the case labels used in the literature depend on the analysis of verbal marking, which has been subject to long-standing debate. See Chen & McDonnell (2019) for an overview.

15 The pronominal system in Kelabit otherwise distinguishes singular, dual, paucal and plural number, makes an inclusive/exclusive distinction and has a generic pronoun *narih* (Hemmings 2016).

Tab. 2: Kelabit case-marked pronouns.

	nom	gen
1sg	uih	kuh
2sg	iko	muh
3sg	ieh	neh
3pl	ideh	deh

They are labelled NOM and GEN on the basis that they appear cognate with these forms in more conservative Austronesian languages (see Hemmings 2016: 324–325) and are typically used for the function of expressing subjects and non-subject actors respectively, as shown in (6):

(6) Case-Marking
   a. **Actor Voice**
      Uih *ni'er* *ieh*
      1SG.NOM AV.see 3SG.NOM
      'I saw him.'
   b. **Undergoer Voice**
      Seni'er *kuh* ***teh=ieh***
      UV.PFV.see 1SG.GEN PT=3SG.NOM
      'I saw him.' (elicitation, fieldnotes)

However, Kelabit differs from other Western Austronesian languages in two important ways. Firstly, definite non-subject undergoers are marked with NOM rather than OBL. This is contrary to what might be expected given that OBL has the function of marking non-subject undergoers in more conservative Western Austronesian languages, as shown in (5), and preserves this function in the closely related Lun Bawang language (Clayre 2005). The result is that two NOM marked arguments can occur in AV constructions like (6a). Secondly, and perhaps related to the possibility of double NOM clauses in AV, NOM and GEN are both possible as a means of expressing the UV actor:

(7) Case-Marking
   a. **Undergoer Voice (GEN actor)**
      Seni'er *kuh* ***teh=ieh***
      UV.PFV.see 1SG.GEN PT=3SG.NOM
      'I saw him.'

b. **Undergoer Voice (NOM actor)**
Seni'er *uih* ***teh=ieh***
UV.PFV.see 1SG.NOM PT=3SG.NOM
'I saw him.' (elicitation, fieldnotes)

Consequently, speakers also have a choice of case-form for the actor in UV constructions. This resembles the pattern of differential actor marking found in other languages, where an argument with the same semantic role and grammatical function can be encoded in more than one way (Fauconnier 2011, McGregor 2010, Witzlack-Makarevich & Seržant 2018). Since voice, word order and differential marking are known to be affected by information structure cross-linguistically, we will now explore what role information structure plays in determining these choices in Kelabit.

3 Defining information structure

In order to explore the role of information structure in determining syntactic choices in Kelabit, this section defines information structure, and the notions of topic and focus, discussing how these relate to the wider concept of prominence discussed in this volume. As discussed in Himmelmann & Primus (2015), the notion of prominence is used as an organising principle across different levels of linguistic structure. This includes prominence at the level of argument structure, where actors are generally assumed to be more prominent than other participant roles (Riesberg, Malcher & Himmelmann 2019, Riesberg & Primus 2015), prominence at the level of functional structure, where subjects are assumed to be more prominent than other functions (Keenan & Comrie 1979) and referential prominence, where accessible and given referents are assumed to be more prominent than inaccessible or new referents (Ariel 1990, Gundel, Hedberg & Zacharski 1993). Another level of structure for which prominence relations are relevant is information structure, which will be the focus for this paper.

Himmelmann & Primus (2015) argue that prominence is related to the focus of attention across these different levels of structure. Hence, what agents, subjects, topics, foci and accessible/given referents share is that they are most likely to be at the centre of our attention. Latrouite (2011) also discusses different levels of structure at which an argument can be more prominent than another. In addition to referential prominence and information-structural prominence, she lists event-structural prominence. Specifically, she notes that verbs can be subcategorised into those with undergoer-oriented event structures (e.g. change-of-state verbs), those

with actor-oriented event structures (e.g. manner of action verbs) and those that are neutral (e.g. punctual contact verbs like *to hit*) (Latrouite & Riester 2018: 250). She demonstrates that for undergoer-oriented verbs, the undergoer is more prominent than the actor and that this can also affect syntactic choices such as voice choice. However, she notes that information-structural prominence can outrank prominence at referential and event-structural levels (Latrouite 2011, Latrouite & Riester 2018).

Information structure is generally understood as a formal mechanism for facilitating effective information exchange or update (Dalrymple & Nikolaeva 2011, Erteschik-Shir 2007). Information is built up of propositions, structured according to what the hearer is presupposed to know, and what they are to learn as a result of communication. The most important, and hence prominent, information structure roles are TOPIC and FOCUS (Krifka 2008, Lambrecht 1994: 127, 213):

(8) Information Structure Roles
Topic
An entity that the speaker identifies and about which a proposition is made.
Focus
The informative part of the proposition that makes an utterance into an assertion and indicates the presence of alternatives.

Topic and focus are typically assumed to be mutually exclusive. Three types of focus domains are distinguished: *argument-focus* (or narrow focus on a single argument); *predicate-focus* (or focus on the verb and its complements) and *sentence-focus* (or broad focus on the entire clause/sentence) (Lambrecht 1994: 222–223).[16]

Using these notions, we can partition the clause into topic-comment and focus-background structures, where topic and focus are more prominent than comment and background respectively. Focus sometimes aligns with comment but need not necessarily. To illustrate consider (9a) and (9b):

16 Note that finer distinctions are sometimes made among different types of topic and focus depending on the semantic/pragmatic interpretations (see e.g. Van der Wal 2016 for discussion). For some contrast is also an important notion (e.g. Vallduví & Vilkuna 1998), though it remains controversial whether contrast should be treated as an independent information structure category or a conversational implicature (Lambrecht 1994: 291). See also Matić & Wedgwood (2013) for similar discussion and arguments against treating focus as a cross-linguistically stable phenomenon but rather as a heuristic for particular contexts and interpretations.

(9) Information Structure
   a. **Context: What did Peter do?**
      $[\text{He}]_{\text{topic}}$ [ate chips]$_{\text{comment}}$
      $[\text{He}]_{\text{background}}$ [ate chips]$_{\text{focus}}$
   b. **Context: What did Peter eat?**
      $[\text{He}]_{\text{topic}}$ [ate chips]$_{\text{comment}}$
      [He ate]$_{\text{background}}$ [chips]$_{\text{focus}}$

In both contexts, the entity identified by the speaker is Peter. Hence, [he] corresponds to the topic. The comment made about him is that he [ate chips]. In (9a), the verb and the object contribute new, informative information. Hence, the verb phrase [ate chips] is also the focus. In (9b), however, the actor and the verb are given in the context of the question, which leaves narrow focus on [chips]. Hence, focus and comment do not align. Consequently, both topic/comment and focus/background structures may be (independently) relevant in determining syntactic choices (cf. Arka & Sedeng 2018). The following section outlines the methodology used for analysing information structure in Kelabit.

4 Analysing information structure in Kelabit

In order to identify topic and focus in Kelabit, I combined elicitation, analysis of a naturalistic text corpus and stimuli-based tasks. Following Lambrecht (1994) and Van der Wal (2016), information structure DIAGNOSTIC TESTS were used to elicit grammaticality judgements on different structures in context. These were largely explored with a single consultant, who is a fluent speaker of Kelabit, English and Malay. Judgements were confirmed on separate occasions and discussed informally with other speakers or compared against the naturalistic text corpus. The first test involves paraphrasing with a hanging topic construction. This is considered a good diagnostic for topics (Lambrecht 1994: 152) and hence was used to assess the acceptability of structures when different arguments have topic status:

(10) Hanging Topic Test
   [As for Andy]$_{\text{HT}}$, he hit John. [Actor Topic]
   [As for John]$_{\text{HT}}$, he was hit by Andy. [Undergoer Topic]

The second test is the canonical test for focus and involves using question/answer pairs (Lambrecht 1994, Van der Wal 2016, Dik 1997, Krifka 2008, Matić & Wedgwood 2013). It is generally assumed that, in the answer to a question, the constituent that

replaces the wh-phrase is in focus (Van der Wal 2016). This test was used to assess the acceptability of structures in different focus domains:

(11) Question/Answer Test

Who hit John?	
[Andy]$_{\text{focus}}$ hit John	[Narrow Focus on Actor]
Who did Andy hit?	
Andy hit [John]$_{\text{focus}}$	[Narrow Focus on Undergoer]
What did Andy do?	
Andy [hit John]$_{\text{focus}}$	[Predicate Focus on Verb+Undergoer]
What happened to John?	
[Andy hit]$_{\text{focus}}$ John	[Predicate Focus on Verb+Actor]
What happened?	
[Andy hit John]$_{\text{focus}}$	[Sentence Focus]

The final test involves correction. A yes/no question is given as context and the constituent that is negated or corrected corresponds to the focus (see e.g. Dery 2007). This was used to establish the acceptability of different structures in different focus domains with the additional entailment of contrast:

(12) Negative Contrast Test

Did Andy hit John yesterday?	
No, [Paul]$_{\text{focus}}$ hit him	[Narrow Focus on Actor]
No, Andy hit [Paul]$_{\text{focus}}$	[Narrow Focus on Undergoer]
No, Andy [slapped Paul]$_{\text{focus}}$	[Predicate Focus on Verb+Undergoer]
No, [I slapped]$_{\text{focus}}$ John	[Predicate Focus on Verb+Actor]
No, [I slapped Paul]$_{\text{focus}}$	[Sentence Focus]

In addition to elicitation, naturally occurring examples of the structures were analysed in context using a NATURALISTIC TEXT CORPUS collected by the author over several periods of fieldwork between 2013–2019 (Hemmings 2017). It includes roughly 14 hours of recorded text material, comprising conversations, descriptions of cultural practices, folk stories, personal narratives, procedural text, radio broadcasts and song. Examples of relevant structures were identified in the corpus and analysed qualitatively, according to their informational structural context.

Finally, I used a translation task with multiple participants to identify which choices speakers make when translating the same sentence in different contexts. The task was developed by Latrouite & Riester (2018) to explore the role of information structure in voice choice in Tagalog. It involves 12 paragraphs that speakers were given to translate from English into Kelabit. The first six paragraphs repeated

the sentence 'cats chase rats' in different contexts. In this case, the undergoer is generic and indefinite. The second six paragraphs repeated the sentence 'my sister kicked my dog' in different contexts. In this case, the undergoer is definite. Both sets were included since definiteness – or referential prominence – is also known to affect voice choice in Western Austronesian (Kroeger 1993, Kaufman 2017).[17] The paragraphs were given to six native speakers of Kelabit, who were residing in Bario at the time of recording and aged between 55–70.[18] Speakers were given as much time as they wanted to translate from English and subsequently their written translations were discussed with the author in order to clarify which structures were used. In four cases, the translations were subsequently audio recorded. All speakers were fluent speakers of Kelabit and English. In the following sections, I explore what the data collected using these methods can tell us about the role of information structure in determining syntactic choices in Kelabit.

5 Information structure and word order

The first choice to be discussed is word order. I will demonstrate that information structure plays a role in word order choice since word order – and the focus-fronting construction in particular – can be used as a strategy to place focus before background. However, I will ultimately argue that word order is not fully determined by information structure but is also sensitive to other levels of prominence, such as argument role and grammatical function. To illustrate, I will first present data elicited using the negative contrast test. These data clearly show a preference for ordering focus before background, which results in different word order choices,

17 As discussed in Latrouite & Riester (2018), this task is useful as a means of collecting controlled and comparable data about syntactic choices in different information structure contexts from multiple speakers. However, one of the challenges is ensuring that the translators interpreted the context in the same way that you intended it, and in the same way as each other. Where speakers varied in terms of voice or construction choice, it is difficult to know whether this might be because they interpreted the context differently. It is also difficult to avoid translation bias entirely. For more discussion of the challenges associated with the different methods of data collection see Hemmings (2020).

18 There is a major dialect boundary between northern and southern varieties of Kelabit though finer dialect differences may exist (see Hemmings 2020 for discussion). All of the speakers in this study spoke northern varieties, including dialects of Bario, Pa Main, Pa Umor and Pa Ukat.

depending on the focus domain. When there is narrow focus on an actor or undergoer, the preference is to realise that argument as the subject in an initial position:[19]

(13) Negative Contrast
Context: 'Did Andy hit John yesterday?' No. . .
a. ***[Paul]***$_{focus}$ *teh suk nemupu' ieh.*
Paul PT REL PFV.AV.hit 3SG.NOM
'It was Paul who hit him (not Andy).' [Narrow Focus on Actor]
b. ***[Paul]***$_{focus}$ *teh suk pinupu' neh.*
Paul PT REL PFV.UV.hit 3SG.GEN
'It was Paul that he hit (not John).' [Narrow Focus on Undergoer]

In (13a), the actor is contrasted (Paul vs Andy). The information is expressed using an AV construction with SVO order. In (13b), the undergoer is contrasted (Paul vs John) and the information is expressed using a UV construction with SVO order. As discussed in §2, these could be considered pseudo-cleft constructions, where the predicate is nominalised, since the relativiser *suk* appears. Either way, narrow focus triggers a construction which places the focused actor or undergoer first.

In contrast to the SVO order triggered by narrow focus, predicate focus is expressed by VOS order, as illustrated in (14), where the verb (*mupu'* 'hit') and undergoer (*John*) or actor (*Andy*) are both contrasted against alternatives:

(14) Negative Contrast
Context: 'Did Andy hit John yesterday?' No. . .
a. *[nemepag Paul]*$_{focus}$ ***teh=ieh.***
AV.PFV.slap Paul PT=3SG.NOM
'He slapped Paul (as opposed to hitting John).'
[Predicate Focus on Verb+Undergoer]
b. *[pipag uih]*$_{focus}$ ***teh=ieh.***
UV.PFV.slap 1SG.NOM PT=3SG.NOM
'I slapped him (as opposed to Andy hitting him).'
[Predicate Focus on Verb+Actor]

19 As discussed in §2, non-subject core arguments always follow the verb and cannot appear initially. If an argument has narrow focus, the preference is for that argument to appear initially, which necessitates choosing the voice construction in which that argument can be mapped to subject (i.e. AV for actors and UV for undergoers). However, it is also possible for a non-subject core argument to have a narrow focus reading *in-situ* in VOS constructions, as shown in (21) below.

In (14a), the verb and the undergoer are contrasted (hitting John vs slapping Paul). This information is expressed using an AV construction with VOS order. In (14b), the verb and the actor are contrasted (Andy hitting vs the speaker slapping). This information is expressed using a UV construction and VOS order. Alternative orders are less acceptable, as indicated by #, or unacceptable in a given context, as indicated by */#:

(15) Negative Contrast
Context: 'Did Andy hit John yesterday?' No. . .
a. ***#ieh*** *[nemepag* *Paul]*$_{focus}$
3SG.NOM AV.PFV.slap Paul
For: 'He slapped Paul.'
b. **/#[pipag]* *neh* ***[Paul]***$_{focus}$
UV.PFV.slap 3SG.GEN Paul
For: 'He slapped Paul.' [Predicate Focus on Verb+Undergoer]

Sentence (15a) is an AV construction with SVO order – the most frequent transitive clause type. However, it is judged to be less acceptable, since the focus is ordered after the background. Similarly, (15b) is a UV construction with VOS order – the unmarked order for UV in out-of-the-blue elicitation contexts. Nonetheless, it is judged to be strange in this context since the focus elements do not form a single constituent. Hence, predicate focus results in a preference for VOS order.

Finally, if an adjunct is contrasted, the preference is to realise the adjunct initially:

(16) Negative Contrast
Context: 'Did Andy hit John yesterday?' No. . .
a. *[edto ma'un]*$_{focus}$ ***teh=ieh*** *nemupu'* *ieh*
day before PT=3SG.NOM AV.PFV.hit 3SG.NOM
'It was the day before yesterday that he hit him (not yesterday).'
b. *[edto ma'un]*$_{focus}$ ***teh=ieh*** *pinupu'* *neh*
day before PT=3SG.NOM UV.PFV.hit 3SG.GEN
'It was the day before yesterday that he hit him (not yesterday).'
c. *#pinupu'* *neh* ***teh=ieh*** *[edto* *ma'un]*$_{focus}$
UV.PFV.hit 3SG.GEN PT=3SG.NOM day before
For: 'It was the day before yesterday that he hit him (not yesterday).'
[Narrow Focus on Adjunct]

In (16), the time adverbial is contrasted (yesterday vs the day before). In this context, the preference is to realise the adjunct initially, either in an AV construction, as in

(16a), or in a UV construction, as in (16b). However, even though the unmarked position of adjuncts is clause-final in Kelabit, this order is judged to be strange when the adjunct is contrasted, as shown in (16c). Consequently, fronting is used as a strategy to place contrastive focus before background across different focus domains.

Similar patterns are found in wh-question and answer pairs. Narrow focus on an argument generally triggers SVO order:

(17) Question-Answer Pairs
   a. **Context: who hit John?**
      [Andy]$_{focus}$ *nemupu'* *John*.
      Andy AV.PFV.hit John
      'Andy hit John.' [Narrow Focus on Actor]
   b. **Context: who did Andy hit?**
      [John]$_{focus}$ *pinupu'* *Andy*.
      John UV.PFV.hit Andy
      'Andy hit John.' (elicitation, fieldnotes) [Narrow Focus on Undergoer]

In (17a), the actor corresponds to the wh-word. The answer is expressed using an AV construction with SVO order. In (17b), the undergoer corresponds to the wh-word and the answer is given in UV with SVO order.[20]

In fact, it is ungrammatical for question words, which are inherently focused, to appear clause-finally:[21]

(18) Question-Answer Pairs
   a. **Actor voice**
      Nemupu'* *John* *[iih]***$_{focus}$*?*
      AV.PFV.hit John who
      For: 'who hit John?'

20 Note that another common strategy for questions and answers is to use a cleft construction with the relativiser *suk/nuk* (see Hemmings 2016: 220–221). Only subjects can be clefted.

21 An interesting question is whether the corresponding answers would also be ungrammatical with these word orders. The answer is likely that they would be infelicitous but this remains to be further confirmed with native speaker judgements. As shown in (22), you can get narrow focus on an *in-situ* non-subject core argument, which could be clause-final if SVO word order is used. However, to have a clause-final subject (or VOS order) expressing the focus in answer to a wh-question would most likely be infelicitous. In elicitation sessions, narrow focus questions were typically answered with either SVO order when S = focus, or VOS order when O = focus.

b. **Undergoer Voice**
Pinupu'* *Andy* *[iih]$_{focus}$?***
UV.PFV.hit Andy who
For: 'who did Andy hit?' (elicitation, fieldnotes)

In contrast, predicate focus questions can be answered with VOS order:

(19) Question-Answer Pairs
a. **Context: What is he doing?**
[Kuman *bua'* *kaber* *nedih]$_{focus}$* ***teh=ieh***
AV.eat fruit pineapple 3SG.POSS PT=3SG.NOM
'He's eating his pineapple.' [Predicate Focus on Verb + Undergoer]
b. **Context: What happened to the fruit?**
[kinan *uih]$_{focus}$* ***neh=idih***
UV.PFV.eat 1SG.NOM PT=DEM
'I ate it.' (elicitation, fieldnotes) [Predicate Focus on Verb + Actor]

In (19a), the actor is given, whilst the verb and undergoer correspond to the focus. A possible answer is an AV construction with VOS order where the verb and undergoer are ordered first. Similarly, in (19b) the undergoer is given, whilst the verb and actor correspond to the focus. A possible answer is a UV construction with VOS order, placing the verb and actor first.[22]

Finally, a question about an adjunct is likely to be answered with the adjunct initially:

(20) Question-Answer Pairs
Context: When will Peter eat his pineapple?
[Na'an]$_{focus}$ *teh* ***Peter*** *kuman* *bua'* *kaber* *nedih*
later PT Peter AV.eat fruit pineapple 3SG.POSS
'Peter will eat his pineapple later.' [Narrow Focus on Adjunct]

In (20), the time adverbial corresponds to the focus and appears initially followed by an AV clause with SVO order. These preferences are summarised in Tab. 3:

22 Note though, as discussed in (22), an AV construction with SVO order is just as likely in answer to a predicate-focus question.

Tab. 3: Word Order Preferences by Focus Domain.

Context	Word Order	Voice
Narrow Focus on Actor	S (teh) VO	AV
Narrow Focus on Undergoer	S (teh) VO	UV
Predicate Focus on Verb+Undergoer	VO (teh) S	AV
Predicate Focus on Verb+Actor	VO (teh) S	UV
Narrow Focus on Adjunct	X (teh) SVO	AV or UV

Consequently, information structure plays a role in word order choice since different word orders can be used to maintain focus before background order in different focus domains. Moreover, the focus-fronting construction appears to be particularly associated with contexts where the initial element is overtly contrasted, since the particle *teh* is consistently used in these contexts.

However, though speakers may have preferences for focus before background orders in specific information structure contexts, as summarised in Tab. 3, word order is not fully determined by information structure, since the same information structure context can result in different word order choices. For example, non-subject arguments can also have a narrow focus reading *in-situ*:[23]

(21) Negative Contrast
Context: Did Andy hit John yesterday? No. . .

a. *Pinupu'* *[Paul]$_{focus}$* ***teh=ieh***
UV.PFV.hit Paul PT=3SG.NOM
'Paul hit him (not Andy).' [Narrow Focus on Actor]

b. *Nemupu'* *[Paul]$_{focus}$* ***teh=ieh***
AV.PFV.hit Paul PT=3SG.NOM
'He hit Paul (not John).' [Narrow Focus on Undergoer]

c. *Pinupu'* *neh* *[edto* *ma'un]$_{focus}$* ***teh=ieh***
UV.PFV.hit 3SG.GEN day before PT=3SG.NOM
'It was the day before that he hit him (not yesterday).' (elicitation, fieldnotes) [Narrow Focus on Adjunct]

A contrasted actor is not necessarily expressed as the subject of an AV construction. It can also be realised as the non-subject argument of a UV construction with VOS order, as in (21a). Similarly, a contrasted undergoer can be realised as the non-subject argument of an AV construction, as in (21b). Finally, a contrasted adjunct can

23 Similarly, non-subject core arguments can be questioned *in-situ* (see Hemmings 2016).

also appear inside the verb phrase, as in (21c), so long as a given topic appears as subject clause-finally. In these instances, the contrast may be marked prosodically rather than through word order, since prosody is another strategy employed by Western Austronesian languages to mark information structure (Latrouite & Riester 2018).[24] Consequently, narrow focus does not necessarily trigger SVO word order, or the use of a focus-fronting/pseudo-cleft construction.

Moreover, AV constructions with SVO order appear to be acceptable regardless of the focus domain:

(22) Question/Answer Pairs

a. **Context: What is he eating?**
neh ***ieh*** *kuman* *[bua'* *kaber* *neh]*$_{focus}$
DEM 3SG.NOM AV.eat fruit pineapple DEM
'He is eating the pineapple.' [Narrow Focus on Undergoer]

b. **Context: What is Peter doing?**
neh ***Peter*** *[kuman* *bua'* *kaber]*$_{focus}$
DEM Peter AV.eat fruit pineapple
'Peter is eating pineapple.' [Predicate Focus on Verb + Undergoer]

c. **Context: What's happening?**
[nih ***Peter*** *kuman* *bua'* *kaber* *nedih]*$_{focus}$
DEM Peter AV.eat fruit pineapple 3SG.POSS
'Peter is eating pineapple' (elicitation, fieldnotes) [Sentence Focus]

AV constructions with SVO order – the basic construction in terms of frequency – are not only used to express narrow focus on the actor, but can also express narrow focus on the undergoer, as in (22a); predicate focus on the verb and undergoer, as in (22b); and sentence focus, as in (22c). Hence, particular focus contexts do not necessarily trigger the use of focus-background orders but can also be expressed using basic word order. Again, it may be focus domains are distinguished prosodically when word order is not used as a strategy to mark the focus/background partition, though this is less obvious in contexts not associated with contrast. Alternatively, the choice of AV construction with SVO order may be determined by the fact that a demonstrative appears in initial position to express progressive aspect (see Hemmings 2016). This remains to be further explored.

24 These examples were elicited and not recorded. A full study on Kelabit prosody would be an important avenue for future research. However, see Himmelmann (2018) for discussion of prosody in Western Austronesian and the claim that prosodic stress may be less important in marking contrastive focus than in other languages.

Finally, the same information structure context can lead speakers to make alternative word order choices. This can be seen from the following passage in the translation task:[25]

(23) Translation Task
Cats are so aggressive. They chase squirrels. They chase birds. Some even chase dogs. I heard that cats were seen chasing a small kangaroo in Australia. They also chase rats, of course. But that is good.

The passage in (23) is about cats. Hence, in the underlined sentences the actor is the topic, whilst the verb and undergoer represent the comment. There is a direct parallelism between the undergoers, which could be thought to introduce an element of contrast (cf. Latrouite & Riester 2018). Finally, the particle 'also' introduces additive focus. The same context resulted in four different word order choices, as shown in (24):

(24) Translation Task
a. **AV SVO**
Ideh *ngalo* *labo* *puur.*
3PL.NOM AV.chase rodent squirrel
'They chase squirrels.'
Ideh *ngalo* *manuk* *[. . .]*
3PL.NOM AV.chase bird
'They chase birds.'
ideh *peh* *ngalo* *labo* *i'eyk* *meto'*
3PL.NOM PT AV.chase rodent rat too
'They also chase rats.'

25 The full breakdown of voice choice in the translation task is given in §7. In general, AV constructions had SVO order and UV constructions had VOS order in line with the basic word order for each construction outlined in §2. This particular context triggered a greater variety of word orders. For the test sentence 'they also chase rats', 2 speakers chose AV constructions with SVO word order, 3 speakers chose AV constructions with VSO order and 1 speaker chose a periphrastic UV construction with VOS order. These are all orders in which the actor topic precedes the focus undergoer (e.g. AVP or VAP). For the sentences 'they chase squirrels' and 'they chase birds' where there is a direct parallelism between the undergoers, all speakers used the same word order for both consecutive clauses (though this sometimes differed from the word order used in the test sentence): 3 speakers chose AV constructions with SVO word order, 1 speaker chose an AV construction with VOS order, shown in (24b), one speaker chose an AV construction with VSO order, shown in (24c), and 1 speaker chose a UV construction with VOS order, shown in (24d).

b. **AV VOS**
Metanur *labo* *puur* ***neh=ideh***
AV.chase rodent squirrel PT=3PL.NOM
'They chase squirrels.'
Metanur *manuk* ***neh=ideh*** *[. . .]*
AV.chase bird PT=3PL.NOM
'They chase birds.'
kineh ***teh=ideh*** *metanur* *labo* *i'eyk*
like.that PT=3PL.NOM AV.chase rodent rat
'They also chase rats.'

c. **AV VSO**
Ngalo ***teh=ideh*** *labo* *puur*
AV.chase PT=3PL.NOM rodent squirrel
'They chase squirrels.'
Ngalo ***teh=ideh*** *labo* *i'eyk* *meto'*
AV.chase PT=3PL.NOM rodent rat too
'They also chase rats.'

d. **UV VOS**
Tu'en *deh* *metanur* *teh* ***labo*** ***puur***
UV.IRR.do 3PL.GEN AV.chase PT rodent squirrel
'They chase squirrels.'
Tu'en *deh* *metanur* *teh* ***manuk*** *[. . .]*
UV.IRR.do 3PL.GEN AV.chase PT bird
'They chase birds.'
Tu'en *deh* *metanur* *ayu'* *teh* ***labo*** ***i'eyk***[26]
UV.IRR.do 3PL.GEN AV.chase EMPH PT rodent rat
'They also chase rats.'

In (24a), the underlined sentences are translated using the basic AV construction with SVO word order. In (24b), the first two clauses are translated using an AV construction with VOS order, placing focus before background. Finally, in (24c) and (24d) the sentences are translated using an AV construction with VSO order and a UV construction with VOS order respectively. Both result in verb-actor-undergoer order, which typically appears in naturalistic discourse when the actor is a continuing topic (see Hemmings 2016). It may be that speakers chose to use word order and

26 This is a periphrastic strategy for expressing undergoer voice using the UV irrealis form of the verb 'to do' *tu'en* + actor + AV.verb + undergoer. It is functionally equivalent to a morphological UV irrealis form (Hemmings 2016).

voice choice to reflect different aspects of the context: some highlighting the parallelism between the undergoers and others highlighting the fact that the actor is a continuing topic. Nonetheless, this reinforces the idea that contrast on the undergoer does not universally trigger UV with SVO word order, since no speakers elected to use that construction.[27]

Thus, word order can be used to indicate the focus of a sentence, and the focus-fronting construction is the preferred means of expressing contrastive focus. However, there is no one-to-one link between word order and the information structure role of particular grammatical functions, since other strategies such as prosody can also be used. Consequently, information structure is not the only determiner of word order in Kelabit, but rather the choice is also related to other levels of prominence, as is common in Western Austronesian (see e.g. Riesberg, Malcher & Himmelmann 2019). In particular, there is an overwhelming preference for orders in which the actor precedes the undergoer, which would explain why AV VOS (=VUA) and UV SVO (=UVA) constructions are less discourse frequent and largely restricted to contexts where the undergoer has a prominent information structure role (see §2).

6 Information structure and differential marking

The second choice is the case-marking of UV actors. As in other languages with differential actor marking (Dalrymple & Nikolaeva 2011, Witzlack-Makarevich & Seržant 2018), the choice of NOM vs GEN appears to be triggered by information structure. Specifically, GEN actors are topics, whilst NOM actors are contrastively focused. Moreover, the use of NOM-marked actors in UV appears to be restricted to contexts in which the undergoer is the topic. This analysis is supported by naturally occurring examples from the corpus as well as information structure diagnostic tests.

As discussed above, UV constructions are less frequent than AV constructions. However, they typically occur with pronominal actors that are GEN marked. For example, in a subset of the corpus that contains six pear story retellings (Chafe 1980); five traditional children's folk stories and three news reports (approx. 12,000

27 A reviewer points out that the contrast created by the parallelism is different from the corrective focus contexts discussed above. It may be that the use of focus-fronting constructions is restricted to contexts with overt contrast or correction, since cross-linguistically contrast is often associated with the use of marked constructions or special prosody (see e.g. Latrouite & Riester 2018: 261). This could explain why constructions in which the undergoer is focus-fronted are not used by any of the speakers in this context.

words), 548 of the voice-marked transitive clauses were AV clauses (or approximately 74%) and 183 were UV clauses (or approximately 25%).[28] Of the UV clauses, 122 clauses had pronominal actors where speakers had a choice between NOM or GEN: 119 had a GEN marked actor, and only 3 clauses had a NOM marked actor.[29] This suggests, at least in monologues and narratives, that GEN-marked actors are expected, whilst NOM pronouns represent a marked choice.

In naturalistic text, GEN actors typically encode (continuing) topics. This could be considered the proto-typical function of actors (Dalrymple & Nikolaeva 2011, Witzlack-Makarevich & Seržant 2018: 11, Du Bois 1987, 2003, McGregor 2010), as well as the proto-typical function of pronouns (Witzlack-Makarevich & Seržant 2018: 26). To illustrate, consider the four UV clauses in (25), which occur towards the end of a traditional folk story about the tortoise, Dayang Beladan, and her quest to retrieve a stolen mouth harp from a monkey:

(25) GEN as topic

a. *Nalap* *neh* ***pupu'***
UV.PFV.fetch 3SG.GEN hitting.implement
'She [Dayang Beladan] fetched something to hit with.'

b. *Nukab* *neh* ***bubpu'*** ***daan***
UV.PFV.open 3SG.GEN door hut
'Opened the door to the hut.'

c. *Nalap* *neh* ***dteh*** ***kayuh***
UV.PFV.fetch 3SG.GEN one stick
'Picked up a piece of wood.'

d. *Nulin* *neh* ***kuyad*** ***sineh***
UV.PFV.throw 3SG.GEN monkey DEM
'And threw it at the monkey.' (folk story, PDA10112013CH_01)

In each clause in (25), the actor is expressed using the 3SG.GEN pronoun *neh*. This refers to Dayang Beladan, the central character of the narrative as a whole and

28 The remaining 13 voice-marked clauses were in instrumental voice.

29 Moreover, two of the examples of NOM actors may not be convincing. One was uttered by a speaker from across the border, who may be speaking a different dialect of Kelabit in which GEN pronouns are simply not used. In another example, the actor is actually *edteh burur ideh* 'one of them' [lit. one body 3PL.NOM]. Hence, this is perhaps not the same construction. In addition to the materials discussed above, there is currently roughly 10 hours of naturalistic text in the corpus. Other genres have yet to be systematically analysed though a few naturally occurring examples from conversations are discussed below. It may be that the proportion of NOM actors is higher in multi-speaker recordings than the single-speaker recordings analysed here.

the topic of the four consecutive clauses. The status of the undergoers varies. In (25d), *kuyad sineh* 'that monkey', is modified by a definite demonstrative. The referent is active, having been mentioned in the preceding discourse, and continues to be discussed in the subsequent clauses. Hence, it could be considered topical. However, the undergoers in the other clauses are neither topical nor given, and don't necessarily convey prominent information: in (25a) the *pupu'* 'hitting implement' is generic; in (25b) the *bubpu' daan* 'door to the hut' is inactive and in (25c) the *kayuh* 'stick' is newly introduced with the indefinite marker *edteh* 'one'. None of these undergoers remain important in the subsequent discourse. Consequently, unlike in other Western Austronesian languages, the choice of UV does not appear to be motivated mainly by the given status/definiteness of the undergoer (Kroeger 1993, Kaufman 2017). Instead, a GEN marked actor serves to mark the actor as topic, regardless of the status of the undergoer.

In contrast to GEN actors, most naturally occurring examples of NOM actors in UV occur in multi-participant conversations. They signal that the actor is in focus or contrasted. A good example is the conversation in (26) which took place between two friends and the author. In the conversation, the speaker tells us about how she learnt to string beads from a great aunt. There are a selection of necklaces and bead caps on the table and towards the end of the conversation the speaker points to one as an example of a bead cap that she had made herself:

(26) NOM as focus

a. ***Uih*** *keli'* *naru'* *baney* *let* *uih* *i'it* *ngilad*
1SG.NOM know AV.make necklace from 1SG.NOM small past
nuuk *maya'* *edteh* *tetepuh* *menaken* *kuh* *keyh*
AV.string follow one grandparent sibling 1SG.GEN small
'I've known how to make necklaces since I was young, I used to string beads following a great aunt of mine.' [. . .]

b. *Nuuk* *teh* ***kedieh*** *petaa* *ngilad,* *petaa* *ba'o* *rawir*
AV.string PT 3SG.EMPH bead.cap past bead.cap bead type
'She would make bead caps in the past, of orange beads.'

c. *En* *kuh* *ni'er* ***ieh*** *naru'* *ih*[30]
UV 1SG.GEN AV.see 3SG.NOM AV.make DEM
'I'd watch her doing it.'

30 This is an example of a periphrastic UV construction, which involves the UV.IRR form of the verb 'to do', *tu'en*, or its shortened form *en*, and a lexical verb in AV form. The construction as a whole can be considered UV since the undergoer in (26), *ieh* '3SG.NOM', behaves as subject. For more details see Hemmings (2016).

d. *Naru'* ***neh=uih*** *petaa* *ba'o* *rawir*
AV.make PT=1SG.NOM bead.cap bead type
'Then I'd make my own orange bead cap.'

e. *Kayu' inih, senuuk* *uih* ***neh***
like DEM UV.PFV.string 1SG.NOM DEM
'Like this one, I strung that [pointing to the bead cap on the table].'

In the UV clause in (26e), the undergoer is the topic rather than the actor. It is highly activated since the bead cap is present on the table and the speaker is pointing to it. Moreover, (26e) is a hanging topic construction in which the undergoer is overtly expressed in the left-periphery, *kayu' inih* 'like this one'. The actor is part of the comment made about the bead cap. Moreover, it is contrasted against other possible actors, including the speaker's great aunt. In other words, the speaker wishes to emphasise that this particular bead cap was strung by her and not anybody else.

The following are additional examples from conversations:

(27) NOM as focus

a. *Suk apeh?* ***suk*** *belaan* *iko* *ih.*
REL which REL UV.IRR.say 2SG.NOM DEM
'Which one?' 'The one that you said!' (conversation, PUM18102013CH_01)

b. *Setu'uh neh, nalap* *ieh* ***teh=ieh.***
actually DEM UV.PFV.take 3SG.NOM PT=3SG.NOM
'But actually he took it.' (conversation, PDA07112013CH_01)

c. *Riak* ***uih*** *naru'* *seh* *diko* *keyh,*
future 1SG.NOM AV.make one 2SG.POSS PT
merey iten *ieh* ***sinih.***
AV.give UV.IRR.take 3SG.NOM DEM
'Later, I'll make one for you, I'll give this one for her to take.'
(conversation, BAR20190211CH_07)

In each case there is a sense that the actor is contrasted against other possible actors, i.e. the addressee in (27a), the other person accused of taking the item in (27b), and the recipient of the item in (27c). In addition, the undergoers are given and expressed using pronouns, definite demonstratives and relativising particles. They are what the speaker is talking about in each instance. Hence, NOM marked actors in naturalistic examples appear to mark an actor that is contrastively focused in contexts where the undergoer is the topic, whilst GEN marked actors are used in other contexts, most typically with continuing topics.

This analysis is supported by grammaticality judgements on information structure diagnostic tests. In contexts where the actor is a topic, GEN is preferred. For

example, GEN marked actors can function as the resumptive pronoun in a hanging topic construction, whilst NOM marked actors cannot:

(28) Hanging Topic
Context: Once upon a time there were two people. One was called Peter. One was called Paul. Peter ate a pineapple. . .
a. *[Paul kedieh]$_{topic}$, kinan neh **bua' ebpuk.***
Paul 3SG.EMPH UV.PFV.eat 3SG.GEN fruit passion
'As for Paul, he ate the passion fruit.'
b. **/#[Paul kedieh]$_{topic}$, kinan ieh **bua' ebpuk.***
Paul 3SG.EMPH UV.PFV.eat 3SG.NOM fruit passion
For: 'As for Paul, he ate the passion fruit.' [Actor Topic]

(28a) is a possible hanging topic construction, where the topic corresponds to the GEN marked actor. However, a NOM pronoun in this context is infelicitous, as shown in (28b).

Conversely, when the actor is in focus, NOM is preferred. For example, in narrow focus question/answer pairs, NOM will be used in the answer and GEN is infelicitous:

(29) Question/Answer Pairs
Context: who saw him?
a. *Seni'er [uih]$_{focus}$* ***teh=ieh***
UV.PFV.saw 1SG.NOM PT=3SG.NOM
'I saw him.'
b. **/#Seni'er [kuh]$_{focus}$* ***teh=ieh***
UV.PFV.saw 1SG.GEN PT=3SG.NOM
'I saw him.' [Narrow Focus on Actor]

Finally, NOM is also preferred when the actor is overtly contrasted. This can be seen from the contrasting judgements in (30), which were elicited in the context of two people arguing over who was the first to tag a given third person:

(30) Negative Contrast
Context: fight over who hit some third person first. . .
a. *Pinupu' uih* ***teh=ieh*** *pu'un am dih iko*
UV.PFV.hit 1SG.NOM PT=3SG.NOM first NEG DEM 2SG.NOM
'I hit him first, not you.' (i.e. you didn't hit him first)
b. **/#Pinupu' kuh* ***teh=ieh*** *pu'un am dih iko.*
UV.PFV.hit 1SG.GEN PT=3SG.NOM first NEG DEM 2SG.NOM
For: 'I hit him first, not you.' [Narrow Focus on Actor]

The NOM actor in (30a) is acceptable in this context. However, the GEN actor in (30b) is unacceptable. This supports the idea that NOM is felicitous when the actor is in contrastive focus.

In a different context, where the argument is not over who did the hitting, but rather who got hit, we find the opposite pattern: it is grammatical for the UV actor to be expressed using a GEN pronoun, but very odd to use the NOM form:

(31) Negative Contrast

a. ***Ieh*** *teh suk pinupu' kuh, am dih iko.*
3SG.NOM PT REL UV.PFV.hit 1SG.GEN NEG DEM 2SG.NOM
'He's the one I hit, not you.' (i.e. I didn't hit you)

b. */#***Ieh*** *teh suk pinupu' uih, am dih iko.*
3SG.NOM PT REL UV.PFV.hit 1SG.NOM NEG DEM 2SG.NOM
For: 'He's the one I hit, not you.' [Narrow Focus on Undergoer]

In (31), the undergoer of the hitting event is contrasted. This is realised via a focus-fronted construction which places the undergoer initially. The actor is given, however, and in this context the GEN pronoun is favoured and the NOM pronoun is unacceptable. The results of the tests are summarised in Tab. 4:

Tab. 4: Case and Information Structure in UV Actors.

Test	NOM ACTOR	GEN ACTOR
Hanging Topic Actor	X	✓
Focus Actor	✓	X
Contrasted Actor	✓	X
Contrasted Undergoer	X	✓

Consequently, information structure plays a clear role in the case-marking of UV actors: the unmarked choice, GEN, is associated with continuing topics, whilst the marked choice, NOM, is associated with focused/contrasted actors. This conforms to a common pattern of differential actor marking cross-linguistically (Fauconnier 2011, Fauconnier & Verstraete 2014, McGregor 2010, Witzlack-Makarevich & Seržant 2018).[31]

31 Note that a possible alternative analysis is that GEN pronouns are clitics that cannot be prosodically stressed. Hence, if contrast is expressed via prosody when a contrasted non-subject argument appears *in-situ*, as discussed above, then the NOM form has to be used. Regardless of whether this pattern is treated as differential case marking or a prosodic phenomenon, it is clear that NOM UV actors are restricted to contexts in which the actor is focused and the undergoer is topic. As will

However, like word order, there is no one-to-one link between case and information structure role outside of the specific contexts discussed above. Firstly, not all topics are GEN marked. Indeed, as in other languages, it is common for subjects to represent topics and these are always NOM marked, as seen in the hanging topic construction in (3a). Similarly, a UV construction with a NOM marked actor is not the only means of expressing narrow focus on actors: they can also be focus-fronted in an AV construction. In addition to the elicited examples in §5, there are also plenty of examples of this structure in the corpus. For example, (32) is from a narrative about animals meeting to decide on who is the strongest. A gecko stakes his claim by crawling up onto the ceiling and proclaiming that he is the one stopping the sky from falling in:

(32) Focus-Fronting

a. *Tulu* ***uih*** *na'am ngimet* *ceiling,* *lit* *tebpa* ***teh***
if 1SG.NOM NEG AV.hold ceiling suddenly fall.in PT
langit ***ih*** *keneh*
sky DEM he.said
'If I don't hold up the ceiling, it will fall in, he said.'

b. ***Uih*** *teh ne-ngimet* *inih* *keneh*
1SG.NOM PT PFV-AV.hold DEM he.said
'I am the one holding it up, he said.' (folk story, BAR17082014CH_06)

In (32b), the actor is contrasted against other possible claimants to the title of strongest animal. The undergoer is lexically given in the immediately preceding clause and expressed through a demonstrative pronoun, *inih*. Hence, the activation status of the arguments is not dissimilar to the conversation in (26). Yet the actor is contrasted in an AV construction rather than UV. A possible explanation for this choice is that the focus-fronting construction is really associated with contrastive topics, since (32b) is about the gecko rather than the ceiling. In contrast, NOM actors in UV appear to be restricted to cases where the actor is focused *and* the undergoer is the topic. This is a very unusual mapping between arguments and information structure roles (see Lambrecht 1994), which might explain why it occurs so infrequently. It would also suggest that the choice of case-marking, and indeed voice construction, not only depends on the information structure characteristics of the

be discussed in more detail in §7, this represents a very marked mapping between arguments and information structure roles. Hence, speakers may choose UV to reflect the marked status of the undergoer as topic, and NOM case to reflect the marked status of the actor. See Latrouite & Riester (2018: 255) for discussion of speakers expressing information-structure prominence of multiple arguments in a clause using different strategies.

argument encoded, but also on other relevant referents in the clause. Consequently, information structure plays a role in the choice of case-marking for UV actors but there is no one-to-one link between form and information structure role more generally.

7 Information structure and voice

The final choice is voice construction. A long-standing question in the literature is what determines voice choice in Western Austronesian (see Chen & McDonnell 2019 for an overview). We have seen from the discussion in §5 and §6 that voice interacts with word order and differential marking since only subjects can appear pre-verbally and differential actor marking is only found in UV constructions. Hence, information structure may determine voice choice if a speaker wishes to use a marked construction, such as focus-fronting or the NOM marked actor in UV. However, we have also seen ample evidence that voice – i.e. the mapping of arguments to functions – is not strictly determined by the information structure role of the subject as either topic or focus. This is evident from the fact that different word orders can be associated with different focus domains (see §5) and different case forms can be associated with different information structure roles for the UV actor despite having the same grammatical function (see §6). Nonetheless, it may be that the use of AV/UV is preferred in certain information structure contexts, as has been argued for other Western Austronesian languages (Latrouite & Riester 2018).

To explore this, I asked six speakers of Kelabit to complete the translation task outlined in §4. The idea was to identify whether the voice construction used in the translation changed depending on the information structure context. Latrouite & Riester (2018) argue that information structural prominence is a key factor in voice choice in Tagalog. They define prominence as having a non-default mapping and assume the following default mappings following Lambrecht (1994): actor → topic, undergoer → focus. This leads them to the following hypotheses, which they show to be broadly accurate for their sample of data in Tagalog:

(33) Hypotheses on Voice Choice in Tagalog (Latrouite & Riester 2018)
   a. If the undergoer has a non-default mapping (i.e. topic), UV is preferred
   b. If actor has a non-default mapping (i.e. focus), AV is preferred
   c. If both actor & undergoer have default mappings, voice choice is determined by other parameters (e.g. the definiteness of the undergoer)

d. If both actor & undergoer have non-default mappings, the focality of the actor appears to be more prominent and AV is preferred.
e. Non-default mappings may also be expressed via word order/marked constructions.

Assuming that information structural prominence is also relevant for voice choice in Kelabit, we can make the same general predictions.[32] The results of the translation task are summarised in Tabs. 5 and 6 respectively. They suggest that the predictions are borne out for Kelabit in the same way as for Tagalog, as discussed below.

Tab. 5: Unhappy Rats – Generic Undergoer.

Context	AV	UV
1. U = topic, V+A = comment/focus	5	1
2. U = topic, A = contrasted/focus, V = background	2	4
3. A = topic, V+U = comment/focus	4	2
4. A+U = contrasted/focus, V = background	5	1
5. All focus	5*	0
6. A = topic, U = contrasted/focus, V = background	5	1
	26/35	9/35

*one speaker did not translate this sentence

Tab. 6: Unhappy Dogs – Definite Undergoer.

Context	AV	UV
1. U = topic, A+V = comment/focus	6	0
2. U = topic, A = contrasted/focus, V = background	6	0
3. A = topic, V+U = comment/focus	1	5
4. All focus	6	0
5. A = topic, U = contrasted/focus, V = background	1*	4
6. A+U = topic, V = comment/focus	0	6
	20/35	15/35

*one speaker used a non-voice marked structure in the translation

32 We might assume differences between Kelabit and Tagalog in the following respects: firstly, as discussed in §6, the definiteness of the undergoer does not necessarily trigger UV choice in Kelabit. Hence, other parameters may determine voice choice when neither actor nor undergoer has information structure prominence (cf. 33c). Secondly, although I assume that both Kelabit and Tagalog can indicate marked choices via marked constructions in addition to voice choice, the nature of these constructions may vary (cf. 33e).

In context 1 and 2, the actor and the undergoer both have non-default mappings. The prediction is that this would prompt AV, since the focality of the actor outranks the topichood of the undergoer (cf. 33d). This is generally true for generic and definite undergoers. Consider (34):

(34) Unhappy Dogs, Definite Undergoer
Context 1: U = topic, A+V = comment/focus
My dog is the poorest dog in the world. He got abandoned as a puppy and almost starved. He got into an accident and lost a leg. Once an eagle attacked him. That is why there is a scar between his eyes. Also my sister kicked him so hard, when she was young, that he lost all trust in human beings

Kineh	*teh*	***kinanak***	***dedtur***	***kudih***	*nemetey'*	*ieh*
like.that	PT	sibling	girl	1SG.POSS	PFV.AV.kick	3SG.NOM

'My sister also kicked him.'

The paragraph in (34) is about the dog. The act of the sister kicking him is new information in the underlined sentence. Hence, the undergoer is the topic and the actor is part of the comment and focus. All speakers expressed this using an AV construction with basic SVO word order.

The only context in which the unexpected role of the actor did not trigger AV was context 2 for generic undergoers. Four out of six consultants chose UV. However, in all cases the UV structures had VOS order and, in three of the four cases, the contrasted actor was expressed as a hanging topic:

(35) Unhappy Rats, Generic Undergoer
Context 2: U = topic, A = contrasted/focus, V = background
It is not only wolves and foxes that threaten rats and catch them. Cats also catch rats and eat them afterwards

[Useyng	*peh]$_{topic}$,*	*debpen*	*deh*	*teh*	***labo***	***i'eyk***
cat	also	UV.IRR.catch	3PL.GEN	PT	rodent	rat

'Cats, they also catch rats.'

In (35), cats are overtly contrasted against wolves and foxes. Hence, in the underlined sentence the undergoer (rats) is part of the background, whilst the actor (cats) is in focus. As discussed above, both focus-fronted structures and hanging topic constructions are associated with contrast. The use of UV allows for the verb and actor to be focus-fronted, resulting in VOS order, and also for an actor in a hanging topic construction followed by the additive focus particle *peh*. This may motivate the unexpected choice of UV. It is also in keeping with the predictions, which state

that marked constructions may express marked information structures in place of voice choice alone (cf 33e).

In all-focus contexts the actor has a non-default mapping and the choice of AV is also predicted (cf 33b). This is invariably the case for both generic and definite undergoers:

(36) Unhappy Dogs, Definite Undergoer
Context 4: All Focus
People are so aggressive these days. Take yesterday, someone drilled a hole into a car to steal gasoline. Then someone pushed away the neighbour's kid so that he almost fell in front of a car. Also my sister kicked my dog very badly. The poor little guy cannot walk properly anymore.
Mey teh ***kinanak dedtur kudih*** *nemetey' uku' kudih*
and PT sibling girl 1SG.POSS PFV.AV.kick dog 1SG.POSS
kail-kail.
hard-REDUP
'And my sister kicked my dog very badly.'

The paragraph in (36) is about people and the bad things they do. In the underlined sentence, the actor (my sister), the verb (kick) and the undergoer (my dog) all represent new information. All consultants expressed this using AV and basic SVO word order.

Similarly, in context 4 for generic undergoers, where both the actor and the undergoer are contrasted, AV is predicted since the status of actor as contrastively focused is non-default. For most speakers, this was the case:

(37) Unhappy Rats, Generic Undergoer
Context 4: A+U = contrasted/focus, V = background
Life in the wilderness is pretty cruel. Lions chase and catch antelopes, sharks catch tunafish and happen to get caught and killed by humans themselves. Even here in the city these cruel laws of nature can be observed. (Our domestic) cats also chase and catch rats, and some also bring them home to continue playing with the bleeding creature.
Useyng *peh ngalo mey ngenep labo i'eyk.*
cat also AV.chase and AV.catch rodent rat
'Cats also chase and catch rats.'

The paragraph in (37) is about wild animals that chase and catch each other. In the underlined sentence, the actor (cats) and the undergoer (rats) are contrasted

against other pairs of predators and prey (e.g. lions and antelopes or sharks and tunafish). This is expressed using an AV structure with basic SVO order.[33]

In context 6 for definite undergoers, where the undergoer has the non-default information structure role of topic, UV is predicted (cf. 33a). This is borne out by the translation task where all speakers chose UV:

(38) Unhappy Dogs, Definite Undergoer
Context 6: A+U = topic, V = comment/focus
My sister and my dog do not really get along. My dog barks at my sister. He also bites her sometimes.

Kerepen	*neh*	*teh*	***ieh***	*temidteh*	*ih*
UV.IRR.bite	3SG.GEN	PT	3SG.NOM	sometimes	DEM

'He sometimes bites her.'

The paragraph in (38) is about the relationship between the sister and the dog. Consequently, in the underlined sentence, the actor (the dog) and the undergoer (the sister) are both topics. This is expressed using a UV construction with VOS order.

Finally, in contexts where both actor and undergoer have default mappings, i.e. contexts 3 and 6 for generic undergoers, and 3 and 5 for definite undergoers, voice choice is predicted to depend on other factors, such as the definiteness of the undergoer (or referential prominence) (cf. 33c). Interestingly, referential prominence seems to play a role for Kelabit too, since UV was the most frequent choice in the unhappy dogs translation, whilst AV was the most frequent choice in the unhappy rats translation:

(39) Unhappy Dogs, Definite Undergoer
Context 3: A = topic, V+U = comment/focus
My sister is so mean. She must be the meanest person in the world. She tricks people whenever she can. She steals money. She even stole my sister's money once. She also kicks my dog, whenever she feels like it.

Tu'en	*neh*	*metey'*	*teh*	***uku'***	***kudih***
UV.do	3SG.GEN	AV.kick	PT	dog	1SG.POSS

'She kicks my dog.'

33 One speaker translated this sentence using UV. This may have been triggered by the desire to express habitual aspect (see Hemmings 2016).

(40) Unhappy Rats, Generic Undergoer
Context 3: A = topic, V+U = comment/focus
Cats are silly creatures with nothing but nonsense on their minds. They climb up on curtains, they bring home mice. Cats also chase and catch big rats, when they are in the mood. Who wants to have a big rat in their house?

Useyng	*meto'*	*ngalo*	*mey*	*ngenep*	*labo*	*nuk*	*merar*
cat	too	AV.chase	and	AV.catch	rat	REL	big

'Cats also chase and catch big rats.'

The paragraph in (39) is a series of comments about the sister. Similarly, (40) is a series of comments about cats. Consequently, in both of the underlined sentence, the actor is the topic and the undergoer is part of the comment. This is the default mapping between arguments and information structure roles. In (39), where the undergoer is definite, the sentence is translated with UV and VOS word order. In (40), however, the undergoer is generic and this triggers AV. Thus, although a major distinction between Tagalog and Kelabit is that the definiteness of the undergoer is not the main factor motivating the choice of UV (see §6), referential prominence also plays a role in voice choice in Kelabit.

Consequently, Kelabit appears to follow similar patterns to Tagalog in that a non-default mapping between semantic role and information structure role can trigger the choice of AV vs UV. However, it is the relative prominence (or unexpectedness) of information that affects voice choice rather the information status of the subject as either topic or focus. This is in keeping with the Von Heusinger & Schumacher's (2019) definition of prominence as relational that is adopted in this volume. In fact, much as in the naturalistic corpus, UV constructions appear most frequently in the translation task when the actor is a topic, regardless of the status of the undergoer. Hence, there is no one-to-one mapping between grammatical function and information structure role either and instead voice choice appears to depend on the information structure context as a whole.

8 Conclusion

This paper examined the role of information structure in determining three syntactic choices in Kelabit: word order, differential case marking and voice. Using information structure diagnostic tests, analysis of a naturalistic text corpus and a translation task, it showed that information structure can motivate unexpected syntactic choices. Word order can be used to place focus before background, with focus-fronting constructions used to mark contrast. Similarly, the use of NOM actors

in UV appears to be licensed in contexts where the actor is contrastively focused *and* the undergoer is a topic. Finally, voice choice is affected by non-default mappings between arguments and information structure roles since the argument with the most unexpected mapping is likely to be realised as subject. This suggests that information structure is an important component of prominence in Austronesian. However, the paper also showed that there is no one-to-one link between word-order, case form, voice and information structure. Instead they combine and interact to express information in context, and this context as a whole affects syntactic choices.

References

Ariel, Mira. 1990. *Accessing noun-phrase antecedents.* London: Routledge.

Arka, I Wayan & I Nyoman Sedeng. 2018. Information structure in Sembiran Balinese. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages*, 139–175. Berlin: Language Science Press.

Bickel, Balthasar. 2010. Grammatical relations typology. In Jae Jung Song (ed.), *The Oxford Handbook of Linguistic Typology*, 399–444. Oxford: Oxford University Press.

Blust, Robert. 2013. *The Austronesian languages.* Canberra: Pacific Linguistics.

Chafe, Wallace L. 1980. *The pear stories: Cognitive, cultural, and linguistic aspects of narrative production.* Norwood, NJ: Ablex.

Chen, Victoria & Bradley McDonnell. 2019. Western Austronesian voice. *Annual Review of Linguistics* 5. 173–195.

Clayre, Beatrice. 2005. Kelabitic languages and the fate of 'focus': evidence from the Kerayan. In I. Wayan Arka & Malcolm Ross (eds.), *The many faces of Austronesian voice systems: some new empirical studies*, 17–57. Canberra: Pacific Linguistics.

Clayre, Beatrice. 2014. A preliminary typology of the languages of Middle Borneo. In Peter Sercombe, Michael Boutin & Adrian Clynes (eds.), *Advances in research on cultural and linguistic practices in Borneo*, 123–151. Phillips, Maine USA: Borneo Research Council.

Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press.

Dery, Jeruen E. 2007. Pragmatic focus and word order variation in Tagalog. *Language and Linguistics* 8(1). 373–402.

Dik, Simon C. 1997. *The theory of functional grammar*, Part 1: The structure of the clause. Berlin: Mouton De Gruyter.

Donohue, Mark. 2007. Word order in Austronesian from north to south and west to east. *Linguistic Typology* 11(2), 349–391.

Dryer, Matthew S. 1997. Are grammatical relations universal. In Joan Bybee, John Haiman & Sandra A. Thompson (eds.), *Essays on language function and language type dedicated to T. Givón*, 115–143. Amsterdam: John Benjamins.

Dryer, Matthew S. 2013. Order of Subject, Object and Verb. In Matthew S. Dryer & Martin Haspelmath (eds.), *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology.

Du Bois, John W. 1987. The discourse basis of ergativity. *Language: Journal of the Linguistic Society of America* 63(4). 805–855.

Du Bois, John W. 2003. Argument structure: grammar in Uue. In John W. Du Bois, Lorraine E. Kumpf & William J. Ashby (eds.), *Preferred argument structure: Grammar as architecture for function*, 11–60. Benjamins.

Erteschik-Shir, Nomi. 2007. *Information structure: The syntax-discourse interface*. Oxford: Oxford University Press.

Fauconnier, Stefanie. 2011. Differential Agent Marking and animacy. *Lingua* 121(3). 533–547.

Fauconnier, Stefanie & Jean-Christophe Verstraete. 2014. A and O as each other's mirror image? Problems with markedness reversal. *Linguistic Typology* 18(1). 3–49.

Greenberg, Joseph H. 1963. Some universals of grammar with particular reference to the order of meaningful elements. In Joseph H. Greenberg (ed.), *Universals of Language*, 73–113. Cambridge, MA: MIT Press.

Gundel, Jeanette K, Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69*(2)*. 274–307.

Haspelmath, Martin. 2010. Comparative concepts and descriptive categories in crosslinguistic studies. *Language* 86(3). 663–687.

Hemmings, Charlotte. 2016. T*he Kelabit Language, Austronesian voice and syntactic typology*. PhD Dissertation, Department of Linguistics, SOAS, University of London.

Hemmings, Charlotte. 2017. "*Documentation of the Kelabit Language, Sarawak, Malaysia*." SOAS, Endangered Languages Archive. http://hdl.handle.net/2196/00-0000-0000-000F-B667-4.

Hemmings, Charlotte. 2020. Methods in Language Documentation and Description: A Guide to the Kelabit Documentation Project. *Journal of Modern Languages* 30(1). 6–76.

Hemmings, Charlotte. 2021a. Subjects in Austronesian: Evidence from Kelabit. In I Wayan Arka, Ash Asudeh & Tracy Holloway King (eds.), *Modular Design of Grammar: Linguistics on the Edge*, 160–180. Oxford: OUP.

Hemmings, Charlotte. 2021b. When an antipassive isn't an antipassive anymore: The Actor Voice construction in Kelabit. In Katarzyna Janic & Alena Witzlack-Makarevich (eds.), *Antipassive: Typology, Diachrony, and Related Constructions*, 579–620. Amsterdam: John Benjamins.

Himmelmann, Nikolaus P. 2018. Some preliminary observations on prosody and information structure in Austronesian languages of Indonesia and East Timor. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Perspectives on information structure in Austronesian languages*, 347–374. Berlin: Language Science Press.

Himmelmann, Nikolaus P & Beatrice Primus. 2015. Prominence beyond prosody – A first approximation. In Amedeo de Dominicis (ed.), *pS-prominenceS: Prominences in Linguistics. Proceedings of the International Conference*, 38–58. Viterbo: DISUCOM.

Himmelmann, Nikolaus P. 2005. The Austronesian languages of Asia and Madagascar: Typological characteristics. In Alexander Adelaar & Nikolaus P. Himmelmann (eds.), *The Austronesian languages of Asia and Madagascar*, 110–181. London: Routledge.

Kaufman, Daniel. 2017. Lexical category and alignment in Austronesian. In Jessica Coon, Diane Massam & Lisa deMena Travis (eds.), T*he Oxford handbook of ergativity,* 589–628. Oxford: Oxford University Press.

Keenan, Edward L. & Bernard Comrie. 1979. Data on the Noun Phrase Accessibility Hierarchy. *Language* 55(2). 333–351.

Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3–4). 243–276.

Kroeger, Paul R. 1993. *Phrase structure and grammatical relations in Tagalog*. Stanford, CA: CSLI Publications.

Kroeger, Paul R. 1998. Language classification in Sarawak: A status report. *The Sarawak Museum Journal* 53(74). 137–173.

Lambrecht, Knud. 1994. *Information structure and sentence form: Topic, focus, and the mental representations of discourse referents*. Cambridge: Cambridge University Press.

Latrouite, Anja. 2011. *Voice and case in Tagalog: the coding of prominence and orientation*. Phd dissertation, University of Düsseldorf.

Latrouite, Anja & Arndt Riester. 2018. The role of information structure for morphosyntactic choices in Tagalog. In Sonja Riesberg, Asako Shiohara & Atsuko Utsumi (eds.), *Information structure in Austronesian languages*, 247–284. Berlin: Language Science Press.

Matić, Dejan & Daniel Wedgwood. 2013. The meanings of focus: The significance of an interpretation-based category in cross-linguistic analysis. *Journal of Linguistics* 49(1). 127–163.

McGregor, William B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120(7). 1610–1636.

Musgrave, Simon. 2001. *Non-subject arguments in Indonesian*. PhD Dissertation, Melbourne University.

Potsdam, Eric. 2006. The cleft structure of Malagasy wh-questions. In Hans-Martin Gärtner, Paul Law & Joachim Sabel (eds.), *Clause structure and adjuncts in Austronesian languages*, 195–232. Berlin: Mouton de Gruyter.

Riesberg, Sonja. 2014. *Symmetrical voice and linking in western Austronesian languages*. Berlin: De Gruyter Mouton.

Riesberg, Sonja, Kurt Malcher & Nikolaus P Himmelmann. 2019. How universal is agent-first? Evidence from symmetrical voice languages. *Language* 95(3). 523–561.

Riesberg, Sonja & Beatrice Primus. 2015. Agent prominence in symmetrical voice languages. *STUF – Language Typology and Universals* 68(4). 551–564.

Schachter, Paul. 1976. The subject in Philippine languages: Topic, Actor, Actor-Topic, or none of the above? In Charles N. Li (ed.), *Subject and topic*, 491–518. New York: Academic Press.

Schachter, Paul. 1996. The subject in Tagalog: Still none of the above. *UCLA Occasional Papers in Linguistics* 15, 1–61.

Smith, Alexander D. 2017. *The languages of Borneo: A comprehensive classification*. PhD Dissertation, University of Hawai'i at Manoa.

Soriente, Antonia. 2013. Undergoer voice in Borneo: Penan, Punan, Kenyah and Kayan languages. In Alexander Adelaar (ed.), *Voice variation in Austronesian languages of Indonesia*, 175–203. Jakarta: NUSA.

Vallduví, Enric & Maria Vilkuna. 1998. On rheme and kontrast. In Peter W. Cullicover & Louise McNally (eds.), *The limits of syntax*, 79–108. New York: Academic Press.

Van der Wal, Jenneke. 2016. Diagnosing focus. *Studies in Language* 40(2). 259–301.

Witzlack-Makarevich, Alena & Ilja A. Seržant. 2018. Differential argument marking: Patterns of variation. In Alena Witzlack-Makarevich & Ilja A. Seržant (eds.), *The Diachronic typology of Differential Argument Marking*, 1–40. Berlin: Language Science Press.