

## Information structure and syntactic choices in Northern Sarawak

Charlotte Hemmings, University of Oxford  
charlotte.hemmings@ling-phil.ox.ac.uk

### 1. Introduction

- ❖ This paper investigates the role of information structure in determining syntactic choices when expressing notionally transitive clauses in three languages of Northern Sarawak: Kelabit (Bario), Sa'ban (Long Banga) and Lun Bawang (Ba Kelalan)
- ❖ These are Western Austronesian languages of the Apad Uat subgroup (Kroeger 1998)
- ❖ Like other Western Austronesian languages, they have SYMMETRICAL VOICE alternations:

(1) *Kelabit*

a. **Actor Voice**

Nengelaak    *nuba'*                    **tesineh nedih.**  
PFV.AV.cook    rice                                    mother 3SG.POSS  
'Her mother cooked rice.'

b. **Undergoer Voice**

Linaak            *tesineh nedih*                    **nuba'.**  
PFV.UV.cook    mother 3SG.POSS                    rice  
'Her mother cooked rice.' (elicitation, fieldnotes)

(2) *Sa'ban*

a. **Actor Voice**

**Súel nah**    maan    *bii'*  
Girl    DEM    AV.eat rice  
'That girl eats/is eating rice'

b. **Undergoer Voice**

inaan    *súel nah*    **bii'**    **ceh**    **ai**  
UV.eat    girl    DEM    rice    2SG    DEF  
'That girl ate rice' (elicitation, fieldnotes)

(3) *Lun Bawang*

a. **Actor Voice**

ne'    nukat                    *kelatih*                    **uih**                    nalem  
PFV.go AV.dig                    worms                    1SG.NOM                    yesterday  
'I went to dig up worms yesterday'

b. **Undergoer Voice**

Tinukat            *uih*                    **kelatih**                    **dih**                    feh  
UV.PFV.dig    1SG.NOM                    worms                    DEM                    PT  
'I already dug up the worms' (elicitation, fieldnotes)

Table 1. Grammatical Functions in AV and UV

	<b>Actor</b>	<b>Undergoer</b>
ACTOR VOICE	subject	object
UNDERGOER VOICE	object	subject

- ❖ The position of the object is fixed after the verb, but the subject has flexible WORD ORDER and can appear either pre- or post-verbally:

(4) *Sa'ban Word Order*

- a. **Súel nah** maan *bii'*  
 girl DEM AV.eat rice  
 'That girl eats rice'
  
- b. Maan *bii'* **súel nah**  
 AV.eat rice girl DEM  
 'The girl eats rice'

- ❖ Across the Apad Uat languages, UV clauses are generally verb-initial and AV clauses are generally SVO by default (Clayre 2014).

- ❖ Finally, in Kelabit and Lun Bawang we find examples of DIFFERENTIAL CASE MARKING. The expected case-marking pattern in Western Austronesian is summarised in Table 2:

Table 2. Case-Marking in Western Austronesian

	<b>Actor</b>	<b>Undergoer</b>
ACTOR VOICE	NOM	OBL
UNDERGOER VOICE	GEN	NOM

- ❖ However, in Kelabit UV constructions both NOM and GEN alternate as a means of expressing UV actors:

(5) *Kelabit Differential Actor Marking*

- a. **Undergoer Voice (GEN actor)**  
 Seni'er      *kuh*      **t=ieh**  
 UV.see      1SG.GEN      PT=3SG.NOM  
 'I saw him'
  
- b. **Undergoer Voice (NOM actor)**  
 Seni'er      *uih*      **t=ieh**  
 UV.see      1SG.NOM      PT=3SG.NOM  
 'I saw him'

- ❖ Similarly, in the dialect of Lun Bawang spoken in Ba Kelalan, OBL and NOM alternate as a means of expressing undergoers in both AV and UV:

(6) *Lun Bawang Differential Undergoer Marking*

a. **Actor Voice**

<b>Yudan</b>	nemefet	<i>keneh</i>
Yudan	AV.PFV.hit	3SG.OBL
'Yudan hit him'		

b. **Yudan**            nemefet            *ieh*  
 Yudan            AV.PFV.hit        3SG.NOM  
 'Yudan hit him'

c. **Undergoer Voice**

Bifet	<i>iYudan</i>	<b>keneh</b>
UV.PFV.hit	Yudan	3SG.OBL
'Yudan hit him'		

d. Bifet            *iYudan*            **ieh**  
 UV.PFV.hit        Yudan            3SG.NOM  
 'Yudan hit him'

- ❖ Consequently, speakers make syntactic choices in terms of VOICE, WORD ORDER, and CASE FORM when expressing transitive events.
- ❖ Q1: since information structure is known to affect these choices in other languages, does information structure play a role in determining syntactic choices in Northern Sarawak?
- ❖ Q2: do the three languages have any differences in terms of information structural considerations?
- ❖ Q3: what does this mean for our analysis of symmetrical voice languages?

## 2. Information Structure

- ❖ Information structure can be understood as a formal mechanism for facilitating effective information exchange or update (Dalrymple & Nikolaeva 2011, Erteschik-Shir 2007)
- ❖ Information is essentially 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 well-defined roles cross-linguistically are TOPIC and FOCUS (Krifka 2008, Lambrecht 1994: 127, 213):

*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.

- ❖ These are identified in Kelabit, Sa'ban and Lun Bawang using the following methods:
  - Information structure DIAGNOSTIC TESTS and grammaticality judgements
  - By analysing examples in context in a NATURALISTIC TEXT CORPUS (Hemmings 2017)
  - Using the 'unhappy rats' TRANSLATION TASK in which the same sentence is presented for translation in differing information structure contexts (Latrouite & Riester 2018).

### 3. Information Structure and Syntactic Choices in Northern Sarawak

#### 3.1 Word Order

- ❖ Fronting can be used as a means of expressing both narrow focus on the subject or predicate focus on the verb + object:

(7) *Kelabit*

**Context: 'Did Andy hit John yesterday?' No...**

- a. [Paul]<sub>focus</sub>    teh    suk    nemupu'    *ieh*  
 Paul            PT    REL    PFV.AV.hit    3SG.NOM  
 'It was Paul who hit him (John)'
- b. [Paul]<sub>focus</sub>    teh    suk    pinupu'    *neh*  
 Paul            PT    REL    PFV.UV.hit    3SG.GEN  
 'It was Paul that he (Andy) hit'
- c. [nemepag    Paul]<sub>focus</sub>    **teh=ieh**  
 AV.PFV.slap    Paul            PT=3SG.NOM  
 'He (Andy) slapped Paul'
- d. [pipag        *uih*]<sub>focus</sub>    **teh=ieh**  
 UV.PFV.slap    1SG.NOM        PT=3SG.NOM  
 'I slapped him (John)'
- e. [edto    ma'un]<sub>focus</sub>    **teh=ieh**    nemupu'    *ieh*  
 day    before            PT=3SG.NOM    AV.PFV.hit    3SG.NOM  
 'It was the the day before yesterday that he hit him' (elicitation, fieldnotes)

(8) *Sa'ban*

a. **Context: who is drinking water?**

[Súel    rah    nah]<sub>focus</sub>    si'    méroop    *pei'*    lang    nah  
 woman old    DEM            PT    AV.drink    water    plain    DEM  
 'The old woman is drinking the water'

b. **Context: what is the woman doing?**

[Méroop    *pei'*]<sub>focus</sub>    **yeh**    **nah**  
 AV.drink    water            3SG    DEM  
 'She's drinking water'

c. **Context: did Semion hit you? No...**  
 [Peter]<sub>focus</sub> si' mamal éek ai  
 Peter one AV.hit 1SG PT  
 'It was Peter who hit me'

d. **Context: was it a man that you hit? No...**  
 [si' súel] si' jamal éek ai  
 one woman one UV.hit 1SG PT  
 'It was a woman that I hit'  
 cf. #jamal éek si' súel

(9) *Lun Bawang*  
**Context: Did Yudan hit Bulan yesterday? No...**  
 [iGituen]<sub>focus</sub> luk nemefet keneh  
 Gituen REL AV.PFV.hit 3SG.OBL  
 'It was Gituen who hit her' (elicitation, fieldnotes)

- ❖ These orders place focus before background information as is common in languages that follow the PRINCIPLE OF NEWSWORTHINESS (Mithun 1992)
- ❖ However, topic-comment structure can also affect word order (see appendix), the *in-situ* object can also have a narrow focus reading and basic order AV clauses are ok in most contexts.
- ❖ So no one-to-one link between word order and information structure but it can be used as a strategy

### 3.2 Differential Marking

- ❖ The differential use of NOM in Kelabit UV constructions marks a contrastively focused actor. This appears to be used when the undergoer is also the topic:

(10) *Kelabit*  
**Context:** discussing necklaces made by the speaker and the speaker's aunt. Pointing to a particular necklace on the table...  
 [Kayu' inih]<sub>topic</sub>, senuuk [uih]<sub>focus</sub> neh.  
 Like DEM UV.PFV.string 1SG.NOM DEM  
 'Like that one, I strung that' (text, BAR08092014CH\_02)

(11) *Kelabit*  
 a. **Context: who saw him?**  
 Seni'er [uih]<sub>focus</sub> t=ieh  
 UV.PFV.saw 1SG.NOM PT=3SG.NOM  
 'I saw him' (\*kuh)

b. **Context: fight over who hit some third person first...**  
 Pinupu' [uih]<sub>focus</sub> t=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) (\*kuh)

- ❖ In cases of UV with GEN marked actors, the actor is typically the continuing topic.

(12) *Kelabit*

**Context:** story about Dayang Beladan, a turtle who had her prized instrument stolen by a monkey. She now has her instrument back but has spotted the monkey again...

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

Nukab            [neh]<sub>topic</sub>            **bubpu'**            **daan**  
 UV.PFV.open    3SG.GEN            door                hut  
 'Opened the door to the hut'

Nalap            [neh]<sub>topic</sub>            **dteh**    **kayuh**  
 UV.PFV.fetch    3SG.GEN            one                stick  
 'Picked up a piece of wood'

Nulin            [neh]<sub>topic</sub>            **kuyad**            **sineh**  
 UV.PFV.throw    3SG.GEN            monkey            DEM  
 'And threw it at the monkey'

(13) *Kelabit*

**Context:** Once upon a time there were two people. One was called Peter. One was called Paul. Peter ate a pineapple...

[Paul            kedieh]<sub>topic</sub>,            kinan            neh            **bua'**    **ebpuk**  
 Paul            3SG.EMPH            UV.PFV.eat    3SG.GEN            fruit    passion  
 'As for Paul, he ate the passion fruit' (\*ieh)

- ❖ This conforms to patterns of DIFFERENTIAL ACTOR MARKING cross-linguistically (Fauconnier 2011, Fauconnier & Verstraete 2014, McGregor 2010, Witzlack-Makarevich & Seržant 2018)

- ❖ Differential use of OBL in the Lun Bawang dialect of Ba Kelalan is unusual because it is used both when the undergoer is an object (in AV) and when the undergoer is a subject (in UV). It remains to be further explored but may mark the undergoer as topical:

(14) *Lun Bawang*

- a. **Context:** story about the crow and the argus pheasant

dih    **Bungkaak**    nenaat            ki=Tuwau            feh  
 and    crow            AV.PFV.paint    OBL=argus.pheasant    PT

naru'            keneh            roo'-roo'            taga [...]  
 AV.make            3SG.OBL            good-REDUP            pretty  
 'and so Crow painted Argus Pheasant to make him beautiful'  
 (folk story, BAK20171101CH\_03)

- b. Context: My dog and my sister don't get on well...
- |             |            |  |            |        |         |              |
|-------------|------------|--|------------|--------|---------|--------------|
| <b>Uko'</b> | <b>uih</b> |  | <b>dih</b> | megai' | mangang | <i>keneh</i> |
| dog         | 1SG.NOM    |  | DEM        | always | AV.bark | 3SG.OBL      |
- 'My dog always barks at her'

(15) *Lun Bawang*

- a. **Context: Who hit Bulan?**

<b>Yudan</b>	nemefet	<i>keneh</i>
Yudan	AV.PFV.hit	3SG.OBL

'Yudan hit her' (\*ieh)

- b. **Context: why did he hit her?**

Bifet	<i>ieh</i>	<b>keneh</b>	ngaceku	<b>ieh</b>	melalid
UV.PFV.hit	3SG.NOM	3SG.OBL	because	3SG.NOM	ADJ.naughty

'He hit her because she was naughty' (\*ieh)

- c. **Context: Did Yudan hit Bulan? No...**

<b>i=Gituen</b>	luk	nemefet	<i>keneh</i>
NOM=Gituen	REL	AV.PFV.hit	3SG.OBL

'It was Gituen who hit her' (\*ieh)

❖ The NOM undergoer may be less topical or (at least in UV) focus:

(16) *Lun Bawang*

**Context:** A boy has arrived in the video and is about to steal some pears. He sees a man picking fruit up in the tree...

Dih	<b>ieh</b>	nier	<i>ieh</i>	nge'	luun
DEM	3SG.NOM	AV.see	3SG.NOM	there	upon

'He looks at him [the man] up there'

A	<b>delai</b>	<b>dih</b>	nier	<i>keneh</i>
NEG	man	DEM	AV.see	3SG.OBL

'But the man isn't looking at him' (pear story, BAK20190227CH\_01)

(17) *Lun Bawang*

- a. **Context: who did Yudan hit?**

<b>i=uih</b>		bifet	<i>ieh</i>
NOM=1SG.NOM		UV.PFV.hit	3SG.NOM

'He hit me' (\*keneh)

- b. **Context: Did Yudan hit Bulan? No...**

<b>uih</b>	luk	bifet	<i>ieh</i>
1SG.NOM	REL	UV.PFV.hit	3SG.NOM

'He hit me' (\*keneh)

❖ It seems that OBL forms are preferred when the undergoer is topical.

- ❖ This is in keeping with other patterns of differential object marking cross-linguistically (Dalrymple & Nikolaeva 2011, Iemmolo 2010).

### 3.3 Voice Construction

- ❖ Voice interacts with word order and differential marking since only subjects can appear pre-verbally and, at least for Kelabit, DAM is restricted to UV. Thus, information structure may determine voice choice if speakers wish to use marked constructions.
- ❖ However, it is not the case that voice (i.e. the mapping of arguments to functions) is determined by information structure or the role of the “subject” as topic or focus.
- ❖ This is evident from the previous discussion, which showed that a) different word orders within AV or UV can be associated with different focus-background structures (see 3.1) and b) different case-forms of pronouns within a voice construction can be associated with different information structure roles (see 3.2).
- ❖ To explore whether the voices are associated with information structure considerations, and whether the languages differ from one another, I asked 6 speakers of Kelabit, 6 speakers of Sa’ban and 4 speakers of Lun Bawang to translated 12 sets of paragraphs: the Unhappy Rats Translation Task (Latrouite & Riester 2018).
- ❖ The first 6 paragraphs repeated the sentence ‘cats chase rats’ in different information structure contexts. In this case, actor and undergoer are both indefinite and generic. The second set of 6 paragraphs repeated the sentence ‘my sister kicked my dog’ in different information structure contexts. In this case, the actor and undergoer were definite.
- ❖ The Lun Bawang translations exclusively used default AV constructions, regardless of whether the actor or undergoer was topical:

(18) *Lun Bawang*

- a. **Context Actor = Topic:** It’s not only birds and cats that my sister hates and kicks...

**Ieh**            pana’    nupak            uko’    uih            dih  
 3SG.NOM    also    AV.kick            dog    1SG.NOM    DEM  
 ‘She also kicks my dog’

- b. **Context: Undergoer = Topic:** 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...

Idih    **kinanak decur uih**            nengipak    neneh    bura’-bura’  
 and    sibling    girl    1SG.NOM    AV.PFV.kick    3SG.OBL    hard-REDUP  
 ‘And my sister kicked him so hard’



- ❖ In Kelabit, UV constructions are found in both sets but slightly more frequently when the undergoer is definite. In general, it is used when the actor is a continuing topic.<sup>1</sup>

(19) *Kelabit*

- a. **Context: Actor = Topic:** It's not only birds and cats that my sister hates and kicks...

Tu'en *neh* metey' **teh** **uku'** **kudih** meto'  
 UV 3SG.GEN AV.kick PT dog 1SG.POSS as.well  
 'She also kicks my dog'

- b. **Context: Undergoer = Topic:** My dog is the poorest dog...

Kineh meto' **suk** **kenanak kudih** nemetey' *ieh* kail-kail  
 like as.well REL sibling 1SG.POSS AV.PFV.kick 3SG.NOM hard-REDUP  
 'My sister also kicked him so hard'

- ❖ In Sa'ban, UV constructions are mainly found when the undergoer is definite. Like Kelabit, they are also used when the actor is a continuing topic and the undergoer is backgrounded/contrasted. However, they are also found in contexts where the undergoer is the topic and the actor is contrasted.<sup>2</sup>

(20) *Sa'ban*

- a. **Context: Actor = Topic:** It's not only birds and cats that my sister hates and kicks...

**Kuu'** **éek** **ai** pun an *yeh* moté'  
 dog 1SG DEF also UV 3SG AV.kick  
 'She also kicks my dog'

- b. **Context: Undergoer = Topic:** My dog is the poorest dog...

an *aréen súel éek* **ai** moté' *pasel-pasel* **yeh**  
 UV sibling girl 1SG DEF AV.kick hard-REDUP 3SG  
 'My sister also kicked him so hard'

- c. **Context: 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.

Nonoh tah **aréen súel éek** **ai** moté' *kuu' éek* *ai* ma'ét-ma'ét  
 like PT sibling girl 1SG DEF AV.kick dog 1SG DEF hurt-REDUP  
 'My sister also kicked my dog very badly.'

<sup>1</sup> NB. VSO AV constructions were also used as an alternative in these contexts where the undergoer was generic.

<sup>2</sup> NB. SVO UV constructions were used to mark the status of the undergoer as contrastively focused.

#### 4. Conclusion

- ❖ Information structure can play a role in unexpected syntactic choices, e.g. verb-initial word order, differential use of NOM or OBL case.
- ❖ Despite morphosyntactic differences, symmetrical voice languages are affected by similar information structure considerations to ergative/accusative languages.
- ❖ However, there is no one-to-one correlation between word-order, case form, voice and information structure, rather they combine and interact to express information in context.
- ❖ This supports treating symmetrical voice as a syntactic alternation in the mapping of arguments to functions.
- ❖ The most morphosyntactically conservative languages (i.e. Lun Bawang) may not be the most functionally conservative (i.e. Sa'ban) and the voice systems may interact with information structure in different ways.

#### Acknowledgements

I gratefully acknowledge the support of the Leverhulme Trust (ECF-2016-425) and the Kelabit, Lun Bawang and Sa'ban communities.

#### References

- 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.
- 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.
- Hemmings, Charlotte. 2017. "Documentation of the Kelabit Language, Sarawak, Malaysia." SOAS, Endangered Languages Archive. <https://elar.soas.ac.uk/Collection/MPI1029735>.
- Iemmolo, Giorgio. 2010. Topicality and differential object marking: Evidence from Romance and beyond. *Studies in Language* 34(2), 239-272.
- Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55(3-4), 243-276.
- 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 & 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*. Berlin: Language Science Press.

- McGregor, William B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120(7), 1610-1636.
- Mithun, Marianne. 1992. Is Basic Word Order Universal. In Doris L. Payne (ed.), *Pragmatics of Word Order Flexibility*, 15-61. Amsterdam: Benjamins.
- 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*. Berlin: Language Science Press.

## Appendix

- ❖ Establishing/Contrastive Topics can also be realised via a hanging topic construction:

(21) *Kelabit*

**Context:** It's not only wolves and foxes that catch rats...

[useyng	peh] <sub>topic</sub>	tu'en	deh	ngenep	teh	labo i'eyk
cat	also	UV	3PL.GEN	AV.catch	PT	rat

'Cats also catch rats' (translation task)

(22) *Sa'ban*

**Context:** beginning of the story of Jabori – both Jabori and his mum have been introduced...

[déeh	nai] <sub>topic</sub>	mluen	miskin	déeh	nai
3DU	DEM	live	poor	3DU	DEM

'Those two, they were poor' (text, LBA20190312CH\_02)

(23) *Lun Bawang*

**Context:** Asai makes the rivers, Tin Berene makes the land. They first started abroad and now they have returned to Borneo...

[iTin	Berene	nih] <sub>topic</sub>	na	ieh	metueh
Mother	Berene	DEM	NEG	3SG.NOM	ADJ.strong

'Tin Berene, she wasn't strong' (text, BAK20190221CH\_02)