



The Leverhulme Trust

Information Structure and Syntactic Choices in Kelabit

Charlotte Hemmings
University of Oxford

Workshop on Prominence in Austronesian, ANU, 9-10th December 2019

Introduction



- In this paper, I explore the role of **information structure** in determining **syntactic choices** in Kelabit, a Western Austronesian language of Northern Sarawak:
 - **Voice** construction
 - **Word Order**
 - **Case marking** of pronouns
- The aim of this paper:
 - To illustrate how **information structure** interacts with syntactic choices
 - To consider the implications for **prominence** in WAN

Outline



1. Features of Kelabit Grammar
2. Information Structure
3. Information Structure and Word Order in Kelabit
4. Information Structure and Case Marking in Kelabit
5. Information Structure and Voice in Kelabit
6. Conclusions



The Leverhulme Trust

Features of Kelabit Grammar

Background



- **Kelabit** is a Western Austronesian (WAn) language spoken mainly in the Fourth and Fifth divisions of Sarawak, Malaysia (Martin 1996).
- It is part of the **Apad Uat** subgroup of Northern Sarawak which also includes Lun Bawang/Lundayeh, and Sa'ban (Kroeger 1998).
- Data is based on **fieldwork** in **Bario** from 2013-2019 and consists of elicited **grammaticality judgements** and **naturalistic corpus examples**.



Symmetrical Voice



- Western Austronesian languages are known to have **symmetrical voice** alternations
 - Alternations in the mapping of arguments to functions without demotion/detransitivisation

Actor Voice

Undergoer Voice

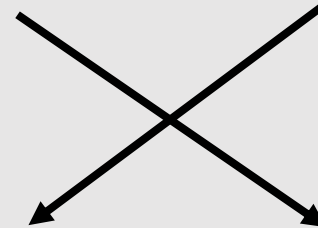
Semantic Roles

A

P

A

P



Syntactic Functions

SUBJ

CORE

SUBJ

CORE

Transitive

Transitive

Symmetrical Voice



(1a) Actor Voice

Nekuman *buah' kaber*

PFV.AV.eat pineapple

'The man ate pineapple'

la'ih sineh
man DEM

Subject

(1b) Undergoer Voice

Kinan *la'ih sineh* *buah' kaber*

PFV.UV.eat man DEM pineapple

'The man ate pineapple'

GFs are distinguished by **word order** and optional pre-subject particles (*teh* & *neh*)

Subject

Syntactic Choice 1



AV vs UV



Kelabit Word Order

(2a) **La'ih sineh** nekuman *bua' kaber*
 man DEM PFV.AV.eat pineapple
 'The man ate the pineapple'

(2b) **bua' kaber* nekuman **la'ih sineh**

(2c) **Bua' kaber** kinan *la'ih sineh*
 pineapple pfv.uv.eat man dem
 'The man ate pineapple'

(2d) **la'ih sineh* kinan **bua' kaber**

The **subject** can appear pre-verbally, as well as following the *non-subject core argument*

Kelabit Word Order



(3a) Nekuman **la'ih sineh** *bua' kaber*
PFV.AV.eat man DEM pineapple
'The man ate pineapple'

(3b) *Kinan *bua' kaber la'ih sineh*
pfv.uv.eat pineapple man
FOR: 'The man ate pineapple'

In AV, **VSO order** is also possible & attested

Kelabit Word Order



	SVO	VOS	VSO
AV	✓ = AVU	✓ = VUA	✓ = VAU
UV	✓ = UVA	✓ = VAU	

Default word order differs depending on the **voice construction**

All else being equal, there is a preference for **actor** before **undergoer**



Hanging Topic Construction

- It is also possible to have **hanging topics** in the left periphery (co-referenced by a pronoun).

(4a) Paul kedieh, nekuman bua' ebpuk t=ieh
 Paul EMPH.3SG AV.PFV.eat fruit passion PT=3SG.NOM
 'As for Paul, he ate passion fruit'

(4b) Bua' ebpuk suk na'ah ih, kinan Paul n=idih
 fruit passion REL afore UV.PFV.eat Paul PT=DEM
 'As for the passionfruit, Paul ate it.'

Focus Fronting



- It is possible for focus information to be fronted in an **inversion construction** using the particle *teh* (cf. ang-inversion in Tagalog)

(5a) **Peter** *teh* suk kuman *bua'* *kaber*
Peter PT REL AV.eat pineapple
'Peter was the one who ate the pineapple'

(5b) *Bua'* *kaber* *teh* kinan *Peter*
Pineapple PT UV.PFV.eat Peter
'Pineapple was the thing that Peter ate'

Syntactic Choice 2



SVO vs VOS vs VSO

Hanging topic

Focus fronting



Austronesian Case Marking

- In more conservative WAn languages, case-marking is used to indicate the **function** of an argument within the **voice system**.
- Typically, three case distinctions are assumed for **pronouns** (cf. Kroeger 1993)

- **NOM** – subjects
- **GEN** – non-subject actors
- **OBL** – obliques and non-subject undergoers

	actor	undergoer
AV	NOM	OBL
UV	GEN	NOM



Kelabit Case Marking

- Kelabit does not have **OBL** case and **NOM/GEN** alternate as a means of marking non-subject actors:

(6a) **Undergoer Voice**

Seni'er

kuh

t=ieh

UV.see

1SG.GEN

3SG.NOM

'I saw him'

(6b) **Undergoer Voice**

Seni'er

uih

t=ieh

UV.see

1SG.NOM

PT=3SG.NOM

'I saw him'

	actor	undergoer
AV	NOM	NOM
UV	GEN/NOM	NOM

differential actor marking

Syntactic Choice 3



NOM vs *GEN* for UV
actors

Summary



- Kelabit has **symmetrical voice** alternations:
 - The first choice is **AV** vs **UV**
- These allow different arguments to be mapped to **subject** with **flexible word order**
 - The second choice is SVO, VOS, VSO and **marked constructions**
- Pronouns are case-marked and the non-subject actor can be **differentially marked**
 - The third choice is **GEN** vs **NOM**
- Q: What is the role of **information structure** in these choices?



The Leverhulme Trust

Information Structure

Information Structure

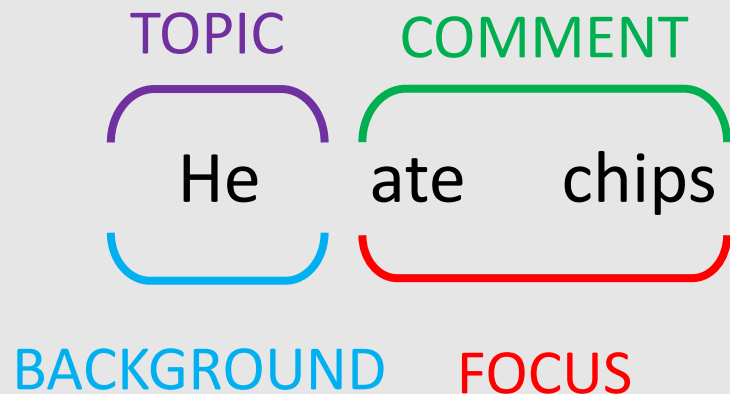


- **Information structure** can be understood as a formal mechanism for facilitating effective information exchange and update (Dalrymple and Nikolaeva 2011, Erteschik-Shir 2007).
- Among the most important information structure roles are **topic** and **focus**:
 - **Topic** is an entity that the speaker **identifies** and **about** which a proposition is made (Krifka 2008)
 - **Focus** is the **informative** part of the proposition and indicates the presence of **alternatives** (Krifka 2008)

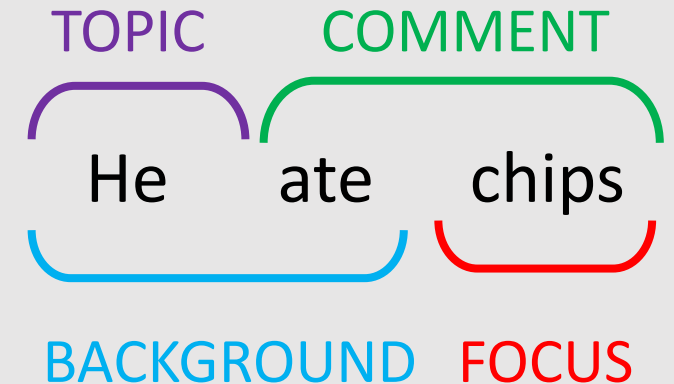
Information Structure

- These allow us to divide the information according to two major distinctions:
 - Topic vs Comment
 - Focus vs Background

CONTEXT: What did Peter do?



CONTEXT: What did Peter eat?





The Leverhulme Trust

Information Structure & Word Order in Kelabit



Word Order and Information Structure

- **Word order** can be used as a strategy to mark **information structure** in Kelabit by placing **focus** information before the **background**.
- We can see this by looking at **grammaticality judgements** in contexts where the focus is clearly identifiable
 - e.g. question-answer pairs, **negative contrast**
- Typically, focus is placed in **initial position**
 - For negative contrast, this involves **teh inversion** constructions

Narrow Focus

CONTEXT: Did Andy hit John yesterday?

(12a) **Not Andy...**

[Paul]_{focus}	teh	suk	nemupu'	<i>ieh</i>
Paul	PT	REL	PFV.AV.hit	3SG.NOM

'It was Paul who hit him (John)'

(12b) **Not John...**

[Paul]_{focus}	teh	suk	pinupu'	<i>neh</i>
Paul	PT	REL	UV.PFV.hit	3SG.GEN

'It wasn't John, it was Paul that he (Andy) hit'

A focused **subject**
appears pre-verbally with
SVO order

Predicate Focus



CONTEXT: did Andy hit John yesterday?

(13c) Andy didn't hit John

[nemepag	<i>Paul</i>] _{focus}
AV.PFV.slap	Paul

'He slapped Paul'

t=ieh

PT=3SG

(13d) Andy didn't hit John...

[pipag	<i>uih</i>] _{focus}
UV.PFV.slap	1SG.NOM

'I slapped John'

t=ieh

PT=3SG.NOM

When the **predicate** is focused, we get VOS order



Narrow Focus on Adjunct

Context: did Andy hit John yesterday?

A focused **adjunct** is
fronted

(14a) **Not yesterday...**

[edto ma'un]_{focus}

day before

t=ieh

PT=3SG.NOM

pinupu'

UV.PFV.hit

neh

3SG.GEN

'It was the day before that he hit him'

(14b) [edto ma'un]_{focus}

day before

t=ieh

PT=3SG.NOM

nemupu'

AV.PFV.hit

ieh

3SG.NOM

'It was the day before that he hit him'

Alternative Orders



(15a) **Andy didn't hit John...**

#**ieh**

3SG.NOM

FOR: 'he slapped Paul'

[nemepag	<i>Paul</i>] _{focus}
----------	--------------------------------

AV.PFV.slap	Paul
-------------	------

(15b) **Not yesterday...**

#pinupu'

UV.PFV.hit

FOR: 'he hit him the day before'

neh

3SG.GEN

t=ieh

PT=3SG.NOM

[edto ma'un]_{focus}

day.before

Corpus Examples



(16a) Buro neh **Guru'** Paul mey Ba Kelalan
away PT teacher Paul go Ba Kelalan
'Guru' Paul transferred to Ba Kelalan'

Jadi epat **neh kelas ih:** Primary 1, Primary 2, Primary 3, Primary 4
but four PT class PT Primary 1, Primary 2, Primary 3, Primary 4
'But there were four classes: Primary 1, Primary 2, Primary 3, Primary 4'

[**uih**]_{focus}

1SG.NOM

teh

PT

ngajar

AV.teach

deh

3PL

neto'

PT

'I was the one left teaching them'



Corpus Examples



Context: Pear story retelling – the man sees three boys wandering past eating pears and wonders if they have stolen his basket...

(16b) [kuman *buɑ'* *ih* **tupu**] focus
AV.eat fruit PT only

'They are just eating fruit'



t=ideh

PT=3PL.NOM

Summary



Context	Word Order	Voice
Narrow Contrast on Actor	[A] teh VU	Actor Voice
Narrow Contrast on Undergoer	[U] teh VA	Undergoer Voice
Predicate Contrast on Verb+Undergoer	[VU] teh A	Actor Voice
Predicate Contrast on Verb+Actor	[VA] teh U	Undergoer Voice
Narrow Contrast on a Time Adverbial	[X] teh AVU	Actor Voice

Focus > Background

This is common in languages which follow the **Principle of Newsworthiness** (Mithun 1992)



Word Order and Information Structure

- However, there is no one-to-one link between **position** and **information structure role**:
 - **Non-subject arguments** can be focused *in situ* so long as the subject is clause final
 - **SVO AV clauses** are judged to be acceptable in most contexts
 - The same **information structure context** can prompt different word order choices.



Narrow Focus on in-situ non-subject

Context: did Andy hit John yesterday? NO...

(17a) **Not John...**

nemupu'

AV.PFV.hit

'He hit Paul'

[Paul]_{focus}

Paul

t=ieh

PT=3SG.NOM

(17b) **Not Andy...**

pinupu'

UV.PFV.hit

'Paul hit him.'

[Paul]_{focus}

Paul

t=ieh

PT=3SG.NOM

Narrow Focus does
not necessarily trigger
SVO word order



SVO AV clauses acceptable in many contexts

(18) **Focused Undergoer**

Q. Kuman *enun* **t=ieh**?
 AV.eat what PT=3SG.NOM?
 ‘What is he eating?’

A. neh **ieh** kuman [*bua' kaber neh*]_{focus}
 DEM 3SG.NOM AV.eat fruit pineapple DEM
 ‘He is eating pineapple’



SVO AV clauses acceptable in many contexts

(19) Focused Predicate

Q. naru' *enun* **Peter**?

AV.do what Peter?

'what is Peter doing?'

A. neh **Peter**

DEM Peter

[kuman *bua'* *kaber*]_{focus}

AV.eat fruit pineapple

'Peter is eating pineapple'



SVO AV clauses acceptable in many contexts

(20) **Focused Sentence**

Q. Kapeh tebey’?
how actually
‘What happened?’

SVO in AV is not only
associated with narrow
focus on the actor

A.	[nih	Peter	kuman	<i>bua’</i>	<i>kaber</i>	<i>nedih</i>] _{focus}
	DEM	Peter	AV.eat	fruit	pineapple	3SG.POSS

‘Peter is eating pineapple’



Same IS context – Different Word Order

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

(23a) **Ideh** ngalo *labo puur*. **Ideh** ngalo *manuk* [...] **ideh** peh ngalo *labo l'ek* meto' (AV SVO)

(23b) Ngalo **tideh** *labo puur*. Ngalo **tideh** *manuk* [...] Ngalo **tideh** *labo l'ek* meto' (AV VSO)

(23c) Metanur *labo puur* **nideh**. Metanur *manuk* **nideh** [...] kineh **tideh** metanur *labo l'ek*
(AV VOS)

(23d) Tu'en *deh* metanur teh *labo puur*. Tu'en *deh* metanur teh *manuk* [...] Tu'en *deh*
metanur ayu' teh *labo l'ek*. (UV VOS)

Summary



- Fronting can be used as a strategy for marking **information structure** – specifically indicating the status of information as **focused**.
- The **voice alternations** can facilitate this in mapping different arguments to different functions.
- However, there is no **one-to-one link** between position, function and information structure role
- Both subjects and non-subjects can be topic and focus and both pre-verbal and post-verbal positions can be associated with these roles.



The Leverhulme Trust

Differential Marking

Case and Information Structure



- As in other languages (Witzlack-Makarevich & Seržant 2018), the choice of **case marking** in Kelabit is triggered by **information structure**:
 - GEN pronouns are used in naturalistic discourse when the actor is a **continuing topic**.
 - NOM pronouns represent **focus/contrastive** actors.



GEN as continuing topic

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

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

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

The GEN actor is a
 continuing topic

NB: the **undergoer**
 is not necessarily
 given/topical



NOM as focus/contrastive

(25) En *kuh* ni'er *ieh* naru' ih
 UV 1SG.GEN AV.see 3SG.NOM AV.make DEM
 'I'd watch her [my great aunt] doing it'

Naru' **n=uih** *petaa* *ba'o* *rawir*
 Av.make PT=1SG.NOM bead.cap beed rawir
 'Then I'd make my own orange bead cap'

Kayu' inih, senuuk *uih* *neh.*
 Like DEM UV.PFV.string 1SG.NOM DEM

'Like that one, I strung that [pointing to the bead cap on the table].'

The **NOM actor** is contrastively focused
 ...and the **undergoer** is the topic

Topic Test



- If you establish the actor as a **hanging topic**, then GEN is preferred:

(26a) **Paul kedieh,** kinan *neh* bua' ebpuk
Paul 3SG.EMPH UV.PFV.eat 3SG.GEN fruit passion
'As for Paul, he ate the passion fruit'

GEN = ✓

(26b) **#Paul kedieh,** kinan *ieh* bua' ebpuk
Paul EMPH.3SG UV.PFV.eat 3SG.NOM fruit passion
FOR: 'As for Paul, he ate passion fruit'

NOM = ✗

Focus Test



- If you make the actor the **answer to a wh-word**, NOM is preferred:

Context: who saw him?

(27a)	seni'er	<i>uih</i>	<i>t=ieh</i>	NOM = ✓
	UV.PFV.see	1SG.NOM	PT=3SG.NOM	
	'I saw him'			

(27b)	*seni'er	<i>kuh</i>	<i>t=ieh</i>	GEN = ✗
	UV.PRF.see	1SG.GEN	PT=3SG.NOM	
	'I saw him'			

Contrast Test



- If actor is **contrasted**, NOM is preferred:

(28a) Pinupu' *uih* **t=ieh** pu'un, am dih iko
UV.PFV.hit 1SG.NOM PT=3SG .NOM first NEG DEM 3SG.NOM

'I hit him first, not you' (i.e. you didn't hit him first))

NOM = ✓

(28b) #Pinupu' *kuh* **t=ieh** pu'un, am dih iko
UV.PFV.hit 1SG.GEN PT=3SG .NOM first NEG DEM 3SG.NOM

'I hit him first, not you' (i.e. you didn't hit him first))

GEN = ✗

Summary



	Expectedness	Information Structure
GEN ACTOR	expected	A = continuing topic
NOM ACTOR	unexpected	A = focus/contrastive

This is common in languages with **differential actor marking**

(Fauconnier 2011, Fauconnier & Verstraete 2014, McGregor 2010, Witzlack-Makarevich & Seržant 2018)



Case and Information Structure

- However, **NOM** marked **actors** in UV are not the only way of expressing **contrastive focus**.

(19) **Uih** teh ne-ngimet *inih* keneh
1SG.NOM PT PFV-AV.hold DEM he.said

‘I am the one holding this [the ceiling] up, he said (and not anyone else).’

- Similarly, not all **continuing topics** are **GEN** marked – since subjects can also be topics and are in **NOM** case

Summary



- There is **no one-to-one link** between form and information structure
- But a **marked construction** (NOM actor in UV) can be used as a strategy to convey a marked information structure reading.
- This construction may be relatively rare because it depends on a context where the **actor is contrastively focused** *and* the **undergoer is a topic** – both of these are unexpected mappings (Lambrecht 1994)



The Leverhulme Trust

Voice Construction



Voice and Information Structure

- A long-standing question is what determines **voice choice** and whether this is linked to **information structure** (see e.g. Chen & McDonnell 2019)
- We have already seen that voice can interact with information structure in facilitating different **word orders** or **case-marking** options.
- However, there is again no one-to-one link between **grammatical function** and information structure role
- That is, the choice of voice construction is not determined purely by the status of the subject as either **topic** or **focus**



Voice and Information Structure

- Nonetheless, the use of AV vs UV may be preferred in certain information structure contexts.
- To explore this, I used the **unhappy rats translation task** in which there are several short paragraphs for speakers to translate.
 - Each paragraph contains a test sentence (“cats chase rats” or “my sister kicked my dog”)
 - The context differs to establish different semantic arguments as topic, focus, background etc.

Voice and Information Structure



- 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 whereby the default for **actors** = **topic**, and the default for **undergoers** = **focus**



Voice and Information Structure

- In their study of Tagalog, they found the following:
 - If only the undergoer has a non-default mapping (topic), **uv is preferred**
 - If only the actor has a non-default mapping (focus), **AV is preferred**
 - If both actor & undergoer have default mappings, voice choice is determined by other parameters (e.g. **the definiteness of the undergoer**)
 - If both actor & undergoer have non-default mappings, the focality of the actor appears to be more prominent and **AV is preferred**.
 - Non-default mappings may also be expressed using **word order/ marked constructions** rather than through voice choice alone.

Unhappy Rats – Generic Undergoer



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





Unhappy Rats – Definite Undergoer



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



Voice and Information Structure

- In contexts 1 & 2, both actor & undergoer have non-default mappings → AV 
- In all focus contexts the actor has a non-default mapping → AV 
- In contexts where the undergoer has a non-default mapping of topic → UV 
- In contexts where both arguments have default mappings - voice choice is determined by different criteria → AV for generic undergoers, UV for definite undergoers 

Summary



- Voice choice in Kelabit may be triggered by **non-default mappings** between semantic roles and information structure just as in Tagalog
- However, the study reaffirms that the **information status** of the privileged argument itself does not determine voice choice alone
- In fact, much as in naturalistic corpora, **UV constructions** appear most frequently in the unhappy rats translation task in contexts where the **actor is a topic**, regardless of the status of the **undergoer**.
- Consequently, it is the information status of **the clause as a whole** that is important



The Leverhulme Trust

Conclusion

Conclusion



- In this paper, I explored the interaction between **word order, case marking, voice** and **information structure** in Kelabit.
- I showed that information structure can play a role in **unexpected syntactic choices**, e.g. verb-initial word order in AV, differential use of NOM in UV
- This shows that **symmetrical voice languages** can be affected by the same information structure considerations as **ergative** and **accusative** languages.

Conclusion



- However, there is no **one-to-one** correlation between word-order, case form, voice and information structure.
- Instead they combine and interact to express **information status in context**.
- This supports treating symmetrical voice as a **syntactic alternation** in the mapping of arguments to functions, rather than an **information-structurally** driven alternation in the encoding of topic/focus.
- But one that can interact with information structure as an important source of **prominence**



The Leverhulme Trust

Many Thanks!