Information Structure, Word Order and Differential Marking in Kelabit

Charlotte Hemmings
University of Oxford

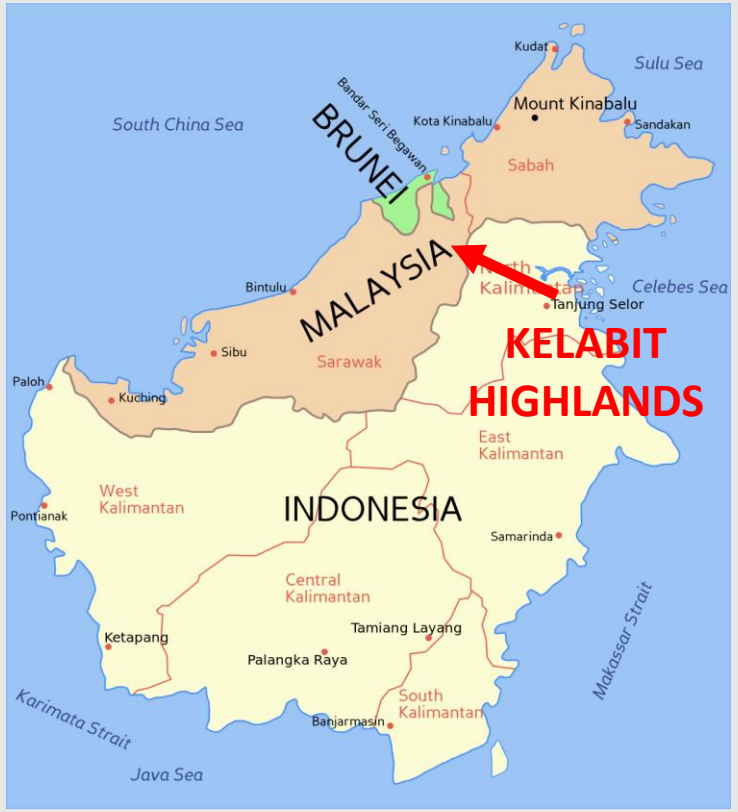
SLE 29th Aug-1st Sept 2018

Introduction

- **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 2013-2017.



Introduction

UNIVERSITY OF OXFORD
The Leverhulme Trust

- Like other WAn languages, Kelabit has a system of **symmetrical voice** alternations which allow different mappings from arguments to functions without changes in the resulting transitivity

- The voice alternations correlate with **word order** – there is a fixed post-verbal position for the non-subject core argument, whilst the subject is more flexible

- Finally, there is a reduced system of **case-marking** in the pronominal system – in undergoer voice (UV), NOM and GEN forms are used differentially to mark non-subject actors.

Introduction

UNIVERSITY OF OXFORD

The Leverhulme Trust

- Hence, in expressing two participant events, speakers make a choice of **voice**, word order and case.

- The aim of this paper:

  - To consider how **information structure** interacts with these syntactic choices
  - To consider the implications for WAn voice systems and information structure more generally

Outline

1. Voice, Word-order and Case-marking in Kelabit

2. Information Structure and Word Order

3. Information Structure and Differential Marking

4. Conclusions

The Leverhulme Trust

Voice, Word Order and Case-Marking

Kelabit Voice

UNIVERSITY OF OXFORD
The Leverhulme Trust

- In order to discuss word-order and case-marking it is necessary to introduce the Kelabit system of **symmetrical voice** alternations.

- Symmetrical voice is typical of WAn and gives speakers different ways of expressing events with (at least) two semantic arguments:
  - actor and undergoer.

- The constructions differ in their verbal morphology which corresponds to a different mapping of arguments to functions:
  - **subject** and *non-subject core argument*.

Kelabit Voice

(1a) **Actor Voice**

Nengelaak *nuba'* **tesineh nedih**

PFV.AV.cook rice mother 3SG.POSS

'Her mother cooked rice'

Subject

(1b) **Undergoer Voice**

Linaak *tesineh nedih* **nuba'**

PFV.UV.cook mother 3SG.POSS rice

'Her mother cooked rice'

Subject

Root = *laak*

AV = ne*N*-

UV = *-in-*

Question

what determines voice selection?

Kelabit Word Order - AV

(2a) **Pre-verbal**

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

'The man ate pineapple.'

(2b) **Post-object**

Ne-kuman	*bua'*	*kaber*	**la'ih**	**sineh**	ngimalem.
PFV-AV.eat	fruit	pineapple	man	DEM	yesterday

'The man ate pineapple yesterday.'

(2c) **Clause-final**

Ne-merey	*nuba'*	ngen	edteh	anak	**la'ih**	**sineh**
PFV-AV.give	rice	to	one	child	man	dem

'That man gave rice to a child'

The **subject** is flexible:

- Pre-verbal
- Post-object
- Clause-final

The *non-subject core* argument is post-verbal

Kelabit Word Order - UV

UNIVERSITY OF OXFORD
The Leverhulme Trust

(3a) **Pre-verbal**

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

'The man ate pineapple.'

b. **Post-object**

kinan	*la'ih*	*sineh*	**bua'**	**kaber**	ngimalem
UV.PFV.eat	man	DEM	fruit	pineapple	yesterday

'The man ate pineapple yesterday.'

c. **Clause-final**

Kinan	*John*	ngimalem	**neh**	**bua'**	**kaber**	**nedih**
UV.PFV.eat	John	yesterday	PT	fruit	pineapple	3SG.POSS

'John ate his pineapple yesterday'

The **subject** is flexible:
- Pre-verbal
- Post-object
- Clause-final

The *non-subject core* argument is post-verbal

Kelabit Case-Marking

UNIVERSITY OF OXFORD
The Leverhulme Trust

- In Kelabit, **case-marking** is found for a subset of the pronominal paradigm:

	NOM	GEN
1SG	uih	kuh
2SG	iko	muh
3SG	ieh	neh
3PL	ideh	deh

The forms are labelled NOM and GEN as they appear cognate with case-marked pronouns in other WAn languages

NOM = subjects
GEN = UV actors

Kelabit

- **NOM** is used for both subjects and non-subjects in AV:

(4a) **AV Actor**

Uih	ni'er	*ieh*
1SG.NOM	AV.see	3SG.NOM

'I see him.'

(4b) **AV Undergoer**

Ieh	ni'er	*uih*
3SG.NOM	AV.see	1SG.NOM

'He sees me.'

Kelabit

- **NOM** and **GEN** alternate as a means of expressing UV actors

(5a) **Undergoer Voice**

Seni'er	*kuh*	**t=ieh**
UV.see	1SG.GEN	PT=3SG.NOM
'I saw him'		

Differential Actor Marking (DAM)

(5b)

Seni'er	*uih*	**t=ieh**
UV.see	1SG.NOM	PT=3SG.NOM
'I saw him'		

Summary

UNIVERSITY OF OXFORD
The Leverhulme Trust

- Speakers have various **syntactic choices** at their disposal when expressing transitive events (in addition to prosody):

  - **Voice**: actor subject or undergoer subject
  - **Word order**: initial vs post-verbal vs final
  - **Case**: NOM VS GEN (for UV actor)

- Q: to what extent does **information structure** affects these choices?

The Leverhulme Trust

Word Order
& Information Structure

Information Structure

UNIVERSITY OF OXFORD

The Leverhulme Trust

- **Information structure** can be understood as a formal mechanism for facilitating effective information exchange or 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): topic-comment

  - Focus is the **informative** part of the proposition and indicates the presence of **alternatives** (Krifka 2008): focus-background

Word Order & Information Structure

UNIVERSITY OF OXFORD
The Leverhulme Trust

- It is well-known that information structure can affect **word-order**:

  - In many languages, there is a tendency to place topic before comment (Lambrecht 1994)

  - In other languages, focus comes before background – “the principle of newsworthiness (PoN)” (Mithun 1992)

Kelabit Word Order

UNIVERSITY OF OXFORD

The Leverhulme Trust

- Kelabit follows the PoN to a certain extent, since there is a tendency to place focus/contrasted information in initial position in both **narrow focus** and **predicate focus** contexts.

- However, there is no one-to-one link between position & information structure role:
  - non-subject arguments can be focused in the immediately post-verbal position
  - initial subjects may also be topics

- This can be seen in **spontaneous examples** as well as **information structure tests** (Q&A, corrective focus negation).

UNIVERSITY OF OXFORD

The Leverhulme Trust

Subject Initial with Narrow Focus

Question & Answer Test

UNIVERSITY OF OXFORD
The Leverhulme Trust

(6) **Focused Actor**

Q. **lih** nemupu' *John*?
who AV.PFV.hit John
'Who hit John?'

A. [**Andy**]$_{focus}$ nemupu' *John*
Andy AV.PFV.hit John
'Andy hit John'

Focused Undergoer

Q. **lih** pinupu' *Andy*?
who UV.PFV.hit Andy
'Who did Andy hit?'

A. [**John**]$_{focus}$ pinupu' *Andy*
John UV.PFV.hit Andy
'Andy hit John'

It is ungrammatical for *wh*-words to appear clause-finally

Negation Test (corrective focus)

Context: did Andy hit John yesterday?

Na'am	**Andy**	nemupu'	*John*	ngimalem…
NEG	Andy	PFV.AV.hit	John	yesterday

'Andy didn't hit John yesterday…'

(8a) **Contrasted Actor**

[**Paul**]$_{focus}$	**teh**	**suk**	nemupu'	*ieh*
Paul	PT	REL	PFV.AV.hit	3SG.NOM

'It was Paul who hit him (John)'

(8b) **Contrasted Undergoer**

[**Paul**]$_{focus}$	**teh**	**suk**	pinupu'	*neh*
Paul	PT	REL	UV.PFV.hit	3SG.GEN

'It was Paul that he (Andy) hit'

An inversion construction places the focused argument in initial position

Negation Test (corrective focus)

Context: did Andy hit John yesterday?

Na'am	**Andy**	nemupu'	*John*	ngimalem…
NEG	Andy	PFV.AV.hit	John	yesterday

'Andy didn't hit John yesterday…'

The same strategy is used to correct a time adjunct

Focus < Background

(8c) **Contrasted Adjunct**

[**edto ma'un**]$_{focus}$	**t=ieh**	pinupu'	*neh*
day before	PT=3SG.NOM	UV.PFV.hit	3SG.GEN

'It was the day before that he hit him'

(8d)

#pinupu'	*neh*	**t=ieh**	[**edto ma'un**]$_{focus}$
UV.PFV.hit	3SG.GEN	PT=3SG.NOM	day before

For: 'he hit him the day before' (not yesterday)

The Leverhulme Trust

Predicate Initial with Predicate Focus

Question & Answer Test

(9) **Focused predicate**

Q. **Enun** tu'en *neh*?

what UV.IRR.do 3SG.GEN

'What is he doing?'

Verb-initial order is a possible response to predicate focus questions

A. [Kuman *bua'* *kaber* *nedih*]$_{focus}$ **t=ieh**

AV.eat fruit pineapple 3SG.POSS PT=3SG.NOM

'He's eating his pineapple'

Negation Test (corrective focus)

Context: did Andy hit John yesterday?

Na'am	**Andy**	nemupu'	*John*	ngimalem…
NEG	Andy	PFV.AV.hit	John	yesterday

'Andy didn't hit John yesterday…'

Verb-initial order is used when the predicate has corrective focus

(11a) **Predicate Focus**

[nemepag	*Paul*]$_{focus}$	**t=ieh**
AV.PFV.slap	Paul	PT=3SG

'He slapped Paul'

(11b) *[pipag]$_{focus}$ *neh* [**Paul**]$_{focus}$

UV.PFV.slap 3SG.GEN Paul

For: 'he slapped Paul'

(11c) **#ieh** [nemepag Paul]$_{focus}$

3SG.NOM AV.PFV.slap Paul

For: 'he slapped Paul'

Spontaneous Examples

(13a) [kuman *bua'* *ih* **tupu**]$_{focus}$ **t=ideh**

AV.eat fruit PT only PT=3PL.NOM

'They are just eating fruit'

Focus particles

(13b) pengeh ineh, am dadan, mirat edteh anak i'it bah

after DEM NEG long INTR.appear one child small EXCL

'Not long afterwards, a small boy appeared'

[ngimet *edteh* *tupi*]$_{focus}$ **t=ieh**

AV.wear one hat PT=3SG.NOM

'He was wearing a hat'

Focus < Background once a discourse topic is established/continuing

Mid-summary

UNIVERSITY OF OXFORD

The Leverhulme Trust

- There is a tendency for **initial position** to be associated with focus information and final position to be associated with given information

- However…

- **Initial-position** is not the only option for expressing focus!
- Initial arguments can also be **topics**

- We can see this using the same diagnostic tests

The Leverhulme Trust

Post-verbal focus and initial topics

Question & Answer Test

(14) Focused Undergoer

Q.	Nekuman	*enun*	**teh**	**Peter**	ngimalem?
	AV.PFV.eat	what	PT	Peter	yesterday
	‘What did Peter eat yesterday?’				

Non-subject arguments can be questioned *in-situ*

A.	Nekuman	[*bua’* *kaber*]$_{focus}$		[**t=ieh**]$_{topic}$	ngimalem
	AV.PFV.eat	fruit	pineapple	PT=1SG.NOM	yesterday
	‘What did Peter eat yesterday?’				

Post-verbal Focus

,

Negation Test (corrective focus)

Context: did Andy hit John yesterday?

Na'am	**Andy**	nemupu'	*John*	ngimalem…
NEG	Andy	PFV.AV.hit	John	yesterday

'Andy didn't hit John yesterday…'

(15a) **Corrected Undergoer**

nemupu'	[*Paul*]$_{focus}$	[**t=ieh**]$_{topic}$
AV.PFV.hit	Paul	PT=3SG.NOM

'He hit Paul'

(15b) **Corrected Actor**

pinupu'	[*Paul*]$_{focus}$	[**t=ieh**]$_{topic}$
UV.PFV.hit	Paul	PT=3SG.NOM

'Paul hit him.'

A post-verbal element can also be corrected/ contrasted

Focus information can occur in initial position, but also post-verbal position

Topic-Initial Order

(16) **Actor Topic**

Q. naru' *enun* **Peter**?
AV.do what Peter?
'what is Peter doing?'

A. neh [**Peter**]$_{topic}$ [kuman *bua'* *kaber*]$_{focus}$
DEM Peter AV.eat fruit pineapple
'Peter is eating pineapple'

A predicate focus question can also be answered with topic > comment order

Initial subjects can have narrow focus but also be topics

Summary

- Hence, **focused subjects** always appear in initial position, but **non-subjects** can also be focused, and initial subjects can also be **topics**.

- Thus, there is no one-to-one link between position and information structure role

- An important question for future research is what the difference is between a focused subject in initial position and a focused object in post-verbal position?

- Let's now look at **differential marking**

The Leverhulme Trust

Differential Marking & Information Structure

Differential Marking & Information Structure

UNIVERSITY OF OXFORD
The Leverhulme Trust

- Differential marking is also known to correlate with **information structure** (among other factors):
  - Differential object marking (DOM) often overtly marks topical objects
  - Differential actor marking (DAM) often overtly marks focused/contrasted actors

(Dalrymple and Nikolaeva 2011, Fauconnier and Verstraete 2014, Iemmolo 2010, McGregor 2010, Witzlack-Makarevich and Seržant 2018)

Differential Marking in Kelabit

UNIVERSITY OF OXFORD
The Leverhulme Trust

- In Kelabit, the choice of NOM and GEN appears to follow a similar pattern to DAM:
  - GEN pronouns mark continuing topics (the default function of actors and pronouns)
  - NOM pronouns indicate focus/contrast
- This can be seen in **spontaneous examples** from the documentary corpus and is further supported by information structure **diagnostic tests**

GEN as continuing topic

(18) 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 given topic and has high topic continuity

NOM as focus/contrastive

(19)

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]

The NOM actor is contrasted against other possible actors

The undergoer is the topic

Hanging Topic Test

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

(20a)	**Paul**	**kedieh**,	kinan	*neh*	bua’	ebpuk	GEN = ✓
	Paul	3SG.EMPH	UV.PFV.eat	3SG.GEN	fruit	passion	
	‘As for Paul, he ate the passion fruit’						
(20b)	**#Paul**	**kedieh**,	kinan	*ieh*	bua’	ebpuk	NOM = ✗
	Paul	EMPH.3SG	UV.PFV.eat	3SG.NOM	fruit	passion	
	FOR: ‘As for Paul, he ate passion fruit’						

Question-Answer Test

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

Context: who saw her?

(21a)	seni'er	***uih***	**t=ieh**	NOM = ✓
	UV.PFV.see	1SG.NOM	PT=3SG.NOM	
	'I saw her'			
(21b)	*seni'er	*kuh*	**t=ieh**	GEN = ✗
	UV.PRF.see	1SG.GEN	PT=3SG.NOM	
	'I saw her'			

Contrast Test

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

(22a)	Pinupu’	*uih*	**t=ieh**	pu’un,	am	dih	iko	NOM = ✓
	UV.PFV.hit	1SG.NOM	PT=3SG .NOM	first	NEG	DEM	3SG.NOM	
	‘I hit him first, not you’							
(22b)	*Pinupu’	*kuh*	**t=ieh**	pu’un,	am	dih	iko	GEN = ✗
	UV.PFV.hit	1SG.GEN	PT=3SG .NOM	first	NEG	DEM	3SG.NOM	
	‘I hit him first, not you’							

Summary

- In the context of UV actors, GEN marks **topics** and NOM marks **focus**.
- However, there is also no one-to-one link between form and **information structure** since NOM can also mark subjects. These can be focus (in initial position) but also topics.
- Moreover, the use of NOM to mark focused actors in UV is seemingly linked to contexts where the actor is focused *and* the undergoer is the primary topic.
- Hence, differential marking may not only depend on information structure characteristics of the argument encoded, but also on **other relevant referents**.

The Leverhulme Trust

Conclusion

Conclusion

UNIVERSITY OF OXFORD

The Leverhulme Trust

- In this paper, I have explored how case-marking and word order are affected by **information structure** in Kelabit.

- I have shown that **initial position** is often associated with focus, and **clause-final** position with given information (or continuing topics).

- Similarly, I have argued that the choice of GEN vs NOM for UV actors is determined by **information structure**, since GEN pronouns reflect continuing topics, and NOM pronouns focus/contrast.

Conclusion

UNIVERSITY OF OXFORD
The Leverhulme Trust

- However, there is no one-to-one link between **function**, position or form and **information structure** role:

  - Both subjects and non-subject actors can be topics
  - Both subjects and non-subjects can be focused (in different positions)
  - Initial and post-verbal positions are associated with both topic and focus (only clause-final position is strictly associated with givenness)
  - NOM case can be associated with both topic and focus

Conclusion

- Consequently, **information status** is neither uniquely determined by voice, nor by word order nor by case-marking, but via a combination of the three.

- The particular encoding typically depends on **global information structure** properties, i.e. the status of both actor and undergoer (Latrouite and Riester 2018).

- Thus, expression of information status in Kelabit involves a complex interaction between syntax, semantics and morphology

Conclusion

UNIVERSITY OF OXFORD

The Leverhulme Trust

- And the **voice system** is independent of the level of information structure…

- But allows different configurations of word order and morphological encoding that reflect different pragmatic readings.

The Leverhulme Trust

Many Thanks!