Annotating Corpus Data with Information Structure: An Example from Balinese

Arndt Riester (Institute for Natural Language Processing, University of Stuttgart)

The 2nd International Workshop on Information Structure in Austronesian Languages

Tokyo, Feb 11-13 2015

A. Information status annotation (*RefLex* scheme, cf. Baumann & Riester 2012, R & B in prep.)

Tag	Description	Uniqueness and Context
r-given-sit	Symbolic deixis	Entity unique in text-external context
r-environment	Gestural deixis	
r-given	Coreference anaphora	Entity unique in previous discourse context
r-given-displaced	... remote antecedent (> 5 clauses)	
r-cataphor	Cataphora	Entity unique in upcoming discourse cont.
r-bridging	Bridging / associative anaphora	Entity unique in previous frame context
r-bridging-contained	Bridging anaphor with embedded antecedent	Entity unique in global context
r-unused-unknown	Hearer-known, discourse-new	
r-unused-known	Hearer-unknown, discourse-new	
r-new	Discourse-new item	non-unique
+generic	Generic or non-specific item	
+predicate	Used in predicative construction	

Table 1 Referential Level. Annotation units: Referring expressions

Tag	Description	Cognitive status (cf. Chafe 1994)
l-given-same	Repetition	active, i.e. salient concepts
l-given-syn	Synonym of previous item	
l-given-super	Hypernym of previous item	
l-given-whole	Holonym of previous item	
l-accessible-sub	Hyponym of previous item	semi-active, i.e. derivable concepts
l-accessible-part	Meronym of previous item	
l-accessible-stem	Partial overlap with previous item	
l-new	Unrelated within last 5 clauses	inactive concepts

Table 2 Lexical Level. Annotation units: Content words

(The following Balinese text was kindly provided to me by Asako Shiohara. Annotations added, AR.)

Bulan Kuning

(1) Ada koné anak luh balu madan Mén Bekung.
exist hearsay person woman widow named mother Bekung
l-new **l-acc-sub** **l-acce-sub** **l-new** **[l-new** **]**
[r-new **]**
There was a widow called Men Bekung.

(2) Ia nongos di sisin alas-é gedé.
3 live at side forest-def big
l-new **l-new** **l-new** **l-new**
r-given **[r-unused-unknown]**
[r-bridging-contained **]**
She lived beside the big forest.

(3) Gegina-né sai-sai ngalih saang ka alas-é.
job-3gen often av-look.for firewood to forest-the
l-new **l-new** **l-acc-part** **l-given-same**
r-bridging-cont **r-new+generic** **[r-given** **]**
Her work was to go to the forest every day looking for firewood.

(4) ka-critaang jani ia maan nuduk anak cerik pusuh.
pass-tell now 3 get av-find child small small
l-new **l-new** **l-new** **l-acc-sub** **l-new**
r-given **[r-new** **]**
It is told now that she found a very small child.

(5) Sawireh kulit-ne putih tur mua-né bunter buka bulan-é
because skin-3 white and face-3 round like moon-def
l-acc-part **l-new** **l-acc-part** **l-new** **l-new**
r-bridging-cont **r-bridging-cont** **r-unused-known**

(6) lantas   ka-adanin   Ni   Bulan   Kuning.
    then   pass-name   Ms.   Bulan (moon)   Kuning (yellow)
             l-new   **l-new**   **l-given-same** **l-new (l-given-syn → putih?)**
             [r-unused-unknown+predicate   ]

Because her skin was white (light?) and her face was round like the moon, she was given the name Ni Bulan Kuning.

(7) Sedek dina   anu   Ni Bulan Kuning   ajaka   bareng
    when   day   unknown   Ms. Bulan Kuning with   together
           l-new **l-new**   **[l-given-same   ]**
    [r-new   ]   **[r-given   ]**

    ngalih   saang   ka   tengah   alas-é.
    av-look for   firewood   to   center   wood-def
    l-given-same   **l-given-same**   **l-new**   **l-given-same**
                                                                   r-given-same
    r-given+generic   **[r-bridging-contained   ]**

One day, Ni Bulan Kuning joined in searching for firewood in the middle of the forest.

(8) Di   tengah   alas-é   Ni Bulan Kuning   paling
    at   center   forest-def   Ms. Bulan Kuning   lost
              l-given-same   **l-given-same**   **[l-given-same]** **l-new**
                                 r-given
    [r-given   ] **[r-given   ]**

(9) tur   palas   ngajak   meme-n-ne.
    and   separate   with   mother-ins-3
              l-new   **l-new**
                                      r-given-displaced

In the middle of the forest Ni Bulan Kuning got lost and was then separated from her mother.

(10) Ni Bulan Kuning   jejeh   pesan   atin-né.

     Ms. Bulang Kuning   scared   very   heart-3

     [l-given-same   ]   **l-new**   **l-acc-part**

     [r-given   ]   **(idiom)**

     Ni Bulan Kuning was very worried.

(11) Mara ia   inget   dapetang-a,

     when 3   realize   find-3act

                   l-new   **l-new**

             r-given

(12) ibane   suba   saup-a   teken   I Raksasa

     oneself   already   take-3act by   Mr. Giant

                    l-new   **[l-new   ]**

     r-given   **[r-unused-unknown]**

     When she realized she found herself held by I Raksasa (the Giant).

(13) Ia   ngetor   baan   jejeh-ne

     3   tremble   because of   fear-3gen

           l-new   **l-acc-stem**

      r-given   **[r-bridging-contained+generic]**

     She trembled because she was scared.

(14) Ia   sayangang-a   pesan   baan   I Raksasa.

     3   love-3act   very   by   Mr. Giant

           l-new   **[l-given-same]**

      r-given   **[r-given   ]**

     She was very much loved by I Raksasa.

(15) I Raksasa   ngelah   manik   sakti   telung   besik

     Mr. Giant   av-have   stone   magic   three   item

     [l-given-same]   **l-new**   **l-new**   **l-given-super**

     [r-given   ]   **[r-new**   **]**

(16) luire: manik   api,   manik   yeh,   manik   angin.
     that is: stone   fire   stone   water   stone   wind
               l-giv-same l-new   **l-giv-same l-new**   **l-giv-same l-new**
               [r-unused-unknown][r-unused-unknown][r-unused-unknown]

I Raksasa had three manik (jewels, magic stones), that is, the fire magic stone, the water magic stone, and the wind magic stone.

(17) Sawireh   ia   sayangang-a,
     because   3   love-3act
                              l-given-same
     r-given

(18) orahin-a   ia   kagunan   manik-é   maketetelu   ento.
     tell-3act   3   use   stone-the   three   that
     l-new   **l-new**   **l-give-same**
                   r-given   **[r-given   ]**
                             [r-bridging-contained   ]

Because she was loved, she was told the use of those three magic stones.

(19) Katuju   I Raksasa   luas,
     when   Mr. Giant   go out
     [l-given-same]   **l-new**
     [r-given   ]

(20) lantas plaibang-a   manike   makejang teken   I Bulan Kuning.
     then   run with-3act   stone-def all   by   Ms. Bulan Kuning
            l-new   **l-giv-same**   **[l-given-same   ]**
                                    [r-given   ] [r-given   ]

When I Raksasa happened to be out, then all the magic stones were taken by Ni Bulan Kuning (and she ran away with them).

(21) Saget teka I Raksasa sahasa nguber Ni Bulan Kuning.

suddenly come Mr.Giant return av-chase Ms. Bulan Kuning

l-new **l-given-same** **l-new** **l-new** **[l-given-same** **]**

[r-given **]** **[r-given** **]**

Suddenly I Raksasa returned to chase Ni Bulan Kuning.

(22) Makir-é Ni Bulan Kuning bakatanga tekén I Raksasa

when Ms.Bulan Kuning catch-3act by Mr.Giant

[l-given-same **]** **l-new** **[l-given-same]**

[r-given **]** **[r-given** **]**

(23) lantas sabat-a baan manik-é ento kanti I Raksasa mati

then throw-3act with stone-def that until Mr. Giant die

l-new **l-giv-same** **[l-given-same]** **l-new**

[r-given **]** **[r-given** **]**

When she was nearly caught by I Raksasa, then he got these stones thrown at himself until I Raksasa died.

(24) Ni Bulan Kuning tengkejut ningeh sabda

Ms. Bulan Kuning surprised hear voice

[l-given-same **]** **l-new** **l-new** **l-new**

[r-given **]** **r-new**

(25) sawiréh Ni Bulan Kuning nyupat I Raksasa.

because Ms. Bulan Kuning av-purify Mr. Giant

[l-given-same **]** **l-new** **[l-given-same]**

[r-given **]** **[r-given** **]**

Ni Bulan Kuning was surprised to hear a voice saying that Ni Bulan Kuning had purified I Raksasa.

(26) Raksasa dadi Betara

Raksana become god

l-given-same **l-new** **l-new**

r-given **r-new+predicate**

(27) | tur | mawali | buin | ka | suarga. |
|---|---|---|---|---|
| and | return | again | to | heaven |
| **l-new** | | | | **l-new** |
| | | | | **[r-unused-known]** |

I Raksasa became a god and returned to heaven.

(28) | Ni Bulan Kuning | kaicén | kasaktian | maubad-ubadan. |
|---|---|---|---|
| Ms. Bulan Kuning | pass-give | power | treat |
| **[l-given-same   ]** | **l-new** | **l-new** | **l-new** |
| **[r-given   ]** | | **[r-unused-unknown+generic   ]** | |

Ni Bulan Kuning was given the power to treat people.

(29) | Jani | Ni BulanKuning | mulih | lantas | ngalih | Mémén Bekung. |
|---|---|---|---|---|---|
| now | [Ms. Bulan Kuning] | go home | then | av-look.for | mother Bekung |
| | **[l-given-same   ]** | **l-new** | | **l-new** | **[l-new   ]** |
| | **[r-given   ]** | | | | **[r-given-displaced]** |

Then she went home and saw Mémen Bekung.

av: actor voice, 3: 3rd person, gen: genitive, act: actor

B. Annotation of discourse structure, Question under Discussion, information structure

➢ Goal: Transform a text into one (or sometimes several smaller) discourse tree(s); in the following indicated by indentations

➢ Distinguish *at-issue* from *not-at-issue* content.

➢ Add *Questions under Discussion* (QUDs, cf. Roberts 1996, Büring 2003). These represent the non-terminal nodes (q).

➢ Assertions (a) represent terminal nodes (e.g. a5' and a5'' are two parallel answers to q5).

Tag	Definition
Focus (F)	The part of a sentence that answers the (current) QUD
Focus Domain (~)	A piece of discourse (containing a focus) that recurs elsewhere (possibly as a question), differing only with respect to the focus (Rooth 1992, Büring 2008)
Background	The non-focal part of a focus domain (that which is already mentioned in the current QUD)
(Aboutness) Topic (T)	Referential entity in the background; ideally part of a coreference chain
Contrastive topic (CT)	A focused topic; signals a discourse strategy (Büring 2003)
Not-at-issue content (NAI)	Material that does not answer the current QUD but represents optional information, e.g. non-restrictive modifiers, parentheticals, appositives, evidentials (e.g. *I believe that*), speaker-oriented adverbs (e.g. *luckily*) etc. (cf. Potts 2005, Simons et al. 2010)

Table 3 Label inventory for information structure

The same story again

QUDs and information structure annotations added, AR.

q0: {How does the story begin?}

> a0:	Ada	koné		anak	luh	balu	madan Mén Bekung.
	exist	hearsay		person	woman	widow	named mother Bekung
		[not-at-issue]					
	[focus						**]**

There was a widow called Men Bekung.

> q1: {What about her (in general)?}

>> a1':	Ia	nongos	di	sisin	alas-é	gedé.
	3	live	at	side	forest-def	big
	[topic]	**[focus**				**]**
	[focus domain					**]**

She lived beside the big forest.

> > a1'': Gegina-né sai-sai ngalih saang ka alas-é.

  job-3gen often av-look.for firewood to forest-the

  (topic?)

  [focus   ]

  [focus domain   ]

  Her work was to go to the forest every day looking for firewood.

> q2: {What happened to her on a particular day?}

> > a2: ka-critaang jani ia maan nuduk anak cerik pusuh.

  pass-tell now 3 get av-find child small small

  [not-at-issue   ] **[topic]** **[focus   ]**

                                     [focus domain   ]

  It is told now that she found a very small child.

> > q3: {What happened with the child?}

> > > a3: Sawireh kulit-ne putih tur mua-né bunter buka bulan-é

          because skin-3white and face-3 round like moon-def

          [not-at-issue   ]

          lantas ka-adanin Ni Bulan Kuning.

          then pass-name Ms. Bulan Kuning

          [focus   ]

  Because her skin was white and her face was round like the moon, she was given the name Ni Bulan Kuning.

q4: {What happened when Bulan Kuning was older?}

> a4: Sedek dina anu Ni Bulan Kuning ajaka bareng

  when day unknown Ms. Bulan Kuning with together

                             [topic   ] **[focus**

  [focus domain

  ngalih saang ka tengah alas-é.

  av-look for firewood to center wood-def

                                               ]

                                               ]

  One day, Ni Bulan Kuning joined in searching for firewood in the middle of the forest.

> q5: {What happened in the center of the forest?}

>>a5':

Di	tengah	alas-é	Ni Bulan Kuning	paling
at	center	forest-def	Ms.Bulan Kuning	lost
			[topic	**] [focus]**

[focus domain]

>>a5'':

tur	palas	ngajak	meme-n-ne.
and	separate	with	mother-ins-3

[focus]

]

In the middle of the forest Ni Bulan Kuning got lost and was then separated from her mother.

>>q6: {*How did she feel about this?*}

>>>a6:

Ni Bulan Kuning	jejeh	pesan	atin-né.
Ms.Bulang Kuning	scared	very	heart-3
[topic	**] [focus**		**]**

[focus domain **]**

Ni Bulan Kuning was very worried.

>>a5''':

Mara	ia	inget
when	3	realize

[not-at-issue]

dapetang-a, ibane	suba	saup-a	teken	I Raksasa
find-3act oneself	already	take-3act	by	Mr. Giant

[focus]

When she realized she found herself held by I Raksasa (the Giant).

>>q7: {*How did she feel about this?*}

>>>a7:

Ia	ngetor	baan	jejeh-ne
3	tremble	because of	fear-3gen
[topic]	**[focus**		**]**

[focus domain **]**

She trembled because she was scared.

q8: {*How did the relationship with Raksasa continue?*}

>a8':

Ia	sayangang-a	pesan	baan	I Raksasa.
3	love-3act	very	by	Mr. Giant
[topic]	**[focus**		**]**	

[focus domain **]**

She was very much loved by I Raksasa.

> *q9: {What about Raksasa?}*

> > a9':

I Raksasa	ngelah	manik	sakti	telung	besik
Mr. Giant	av-have	stone	magic	three	item
[topic]		**[focus**			**]**
[focus domain					**]**

> > *q10: {What kind of stones?}*

> > > a10':

luire:	manik	api,
that is:	stone	fire,
		[focus]
	[focus domain	**]**

> > > a10'':

manik	yeh,
stone	water,
	[focus]
[focus domain]	

> > > a10''':

manik	angin.
stone	wind
	[focus]
[focus domain	**]**

I Raksasa had three manik (jewels, magic stones), that is, the fire magic stone, the water magic stone, and the wind magic stone.

> > *q11: {What did she learn about the magic stones?}*

> > > a11:

Sawireh	ia	sayangang-a,			
because	3	love-3act			
[not-at-issue			**]**		
orahin-a	ia	kagunan	manik-é	maketetelu	ento.
tell-3act	3	use	stone-the	three	that
[focus]	**[topic]**	**[focus]**			**]**
[focus domain					**]**

Because she was loved, she was told the use of those three magic stones.

11

q12: {What happened with the stones one day when Raksasa was out?}

> a12': Katuju I Raksasa luas,

when Mr.giant go out

[focus domain

lantas plaibang-a manike makejang teken I Bulan Kuning.

then run with-3act stone-def all by Ms.Bulan Kuning

[focus] [topic]

]

When I Raksasa happened to be out, then all the magic stones were taken by Ni Bulan Kuning, (and she ran away with them).

> *q13: {What happened to her then?}*

> > a13': Saget teka I Raksasa sahasa nguber Ni Bulan Kuning.

suddenly come Mr.Giant return av-chase Ms.Bulan kuning

[focus] [topic]

[focus domain]

Suddenly I Raksasa returned to chase Ni Bulan Kuning.

> > *q14: {What happened to Raksasa when he (almost) caught her?}*

> > > a14: Makir-é Ni Bulan Kuning bakatanga tekén I Raksasa

when Ms. Bulan Kuning catch-3act by Mr.Giant

[focus domain

lantas sabat-a baan manik-é ento kanti I Raksasa mati

then throw-3act with stone-def that until Mr. Giant die

[focus]

]

When she was nearly caught by I Raksasa, then he got these stones thrown at himself until I Raksasa died.

> > a13'': NiBulan Kuning tengkejut ningeh sabda

Ms. Bulan Kuning surprised hear voice

[topic] [not-at-issue] [focus]

[focus domain]

> > *q15: {What did the voice say?}*

12

> > > a15:   sawiréh   Ni Bulan Kuning   nyupat   I Raksasa.

because   Ms.Bulan Kuning   av-purify   Mr.Raksana

[focus   **]**

Ni Bulan Kuning was surprised to hear a voice saying that Ni Bulan Kuning had purified I Raksasa.

q16: {What happened to Raksana?}

> a16':   Raksasa   dadi   Betara

Raksana   become   god

[topic]   **[focus**   **]**

[focus domain

> a16'':   tur   mawali   buin   ka   suarga.

and   return   again   to   heaven

[focus   **]**

                                     ]

I Raksasa became a god and returned to heaven.

q17: {What happened to Bulan Kuning in the end?}

> a17':   Ni Bulan Kuning   kaicén   kasaktian   maubad-ubadan.

Ms.Bulan Kuning   pass-give   power   treat

[topic   **]**   **[focus**   **]**

[focus domain   **]**

Ni Bulan Kuning was given the power to treat people.

> a17'':   Jani   Ni BulanKuning   mulih

now   Ms. Bulan Kuning   go home

[topic   **]**   **[focus**   **]**

[focus domain

> a17''':   lantas   ngalih   Mémén Bekung.

then   av-look.for   mother Bekung

[focus   **]**

                                     ]

Then she went home and saw Mémen Bekung.

References

Baumann, Stefan & Arndt Riester (2012). Referential and Lexical Givenness: Semantic, Prosodic and Cognitive Aspects. In: G. Elordieta & P. Prieto (eds.) *Prosody and Meaning.* Berlin: Mouton de Gruyter, pp. 119-161.

Büring, Daniel (2003). On D-Trees, Beans, and B-Accents. *Linguistics & Philosophy* 26(5): 511-545.

Büring, Daniel (2008). What's New (and What's Given) in the Theory of Focus? In: *Proceedings of the 34th Annual Meeting of the Berkeley Linguistics Society*, pp. 403-424.

Chafe, Wallace (1994). *Discourse, Consciousness, and Time.* University of Chicago Press

Potts, Christopher (2005). *The Logic of Conventional Implicatures.* Oxford University Press.

Riester, Arndt & Stefan Baumann (in preparation). RefLex Scheme – Annotation Guidelines. URL: http://www.ims.uni-stuttgart.de/institut/mitarbeiter/arndt/

Roberts, Craige (1996). Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. *OSU Working Papers in Linguistics* 49. The Ohio State University.

Rooth, Mats (1992). A Theory of Focus Interpretation. *Natural Language Semantics* 1(1): 75-116.

Simons, Mandy, Judith Tonhauser, David Beaver and Craige Roberts (2010). What Projects and Why. In: Proceedings of Semantics and Linguistic Theory (SALT-20), pp. 309-327. Vancouver.