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


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

*Table 1 Referential Level. Annotation units: Referring expressions*

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

*Table 2 Lexical Level. Annotation units: Content words*
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-acc-sub1-new [l-new ]
[r-new ]
There was a widow called Men Bekung.

(2) Ia nongos di sisin alas-é gedé.
live at side forest-def big
l-new l-new l-new l-new
r-given [r-unused-unknown]
[l-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
Because her skin was white (light?) and her face was round like the moon, she was given the name Ni Bulan Kuning.

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

In the middle of the forest Ni Bulan Kuning got lost and was then separated from her mother.
(10) Ni Bulan Kuning
jejah pesan atin-né.
Ms. Bulang Kuning
scared very heart-3

[1-given-same ] l-new l-acc-part
[r-given ] (idiom)

Ni Bulan Kuning was very worried.

(11) Mara ia
inget dapetang-a,
when 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 jejah-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 ]

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

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

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

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

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

Ni Bulan Kuning became god.
(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 Bulan Kuning 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 Mémé 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).
### Table 3 Label inventory for information structure

<table>
<thead>
<tr>
<th>Tag</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Focus (F)</td>
<td>The part of a sentence that answers the (current) QUD</td>
</tr>
<tr>
<td>Focus Domain (~)</td>
<td>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)</td>
</tr>
<tr>
<td>Background</td>
<td>The non-focal part of a focus domain (that which is already mentioned in the current QUD)</td>
</tr>
<tr>
<td>(Aboutness) Topic (T)</td>
<td>Referential entity in the background; ideally part of a coreference chain</td>
</tr>
<tr>
<td>Contrastive topic (CT)</td>
<td>A focused topic; signals a discourse strategy (Büring 2003)</td>
</tr>
<tr>
<td>Not-at-issue content (NAI)</td>
<td>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)</td>
</tr>
</tbody>
</table>

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

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

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

One day, Ni Bulan Kuning joined in searching for firewood in the middle of the forest.
In the middle of the forest Ni Bulan Kuning got lost and was then separated from her mother.

> > 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 ]

[focus domain ]

Ni Bulan Kuning was very worried.

> > a5''': Mara ia inget when 3 realize

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).

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

> > > a6: Ni Bulan Kuning jejeh pesan atin-né. Ms.Bulan Kuning scared very heart-3

[topic ] [focus]

[focus domain ]

Ni Bulan Kuning was very worried.

> > > a6: Ni Bulan Kuning jejeh pesan atin-né.

Ms.Bulan Kuning scared very heart-3

[topic ] [focus]

[focus domain ]

Ni Bulan Kuning was very worried.

> > 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.
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

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

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

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

q15: [What did the voice say?]
   > a13': NiBulan Kuning tengkejut ningeh sabda
       Ms. Bulan Kuning surprised hear voice

When Ms. Bulan Kuning heard the voice, she was surprised.
> > a15: sawirêh Ni Bulan Kuning nyupat I Raksasa.
because Ms. Bulan Kuning av-purify Mr. Raksasa

[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: Raksana dadi Betara

Raksana become god

[topic ]

[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 domain]

*Ni Bulan Kuning was given the power to treat people.*

> a17’: Jani Ni BulanKuning mulih

now Ms. Bulan Kuning go home

[topic ]

[focus domain]

> a17’’: lantas ngalih Mémén Bekung.

then av-look.for mother Bekung

[focus ]

*Then she went home and saw Mémén Bekung.*
References


