S-structure Features for Information Structure Analysis

Louise Mycock & John Lowe

Faculty of Linguistics, Philology and Phonetics
University of Oxford
Outline

• Atomic discourse functions versus information features
• Information features
  • Previous approaches: [NEW] and [PROM]
  • Our salience features: [ABOUT] and [UPDATE]
  • Our newness features: [DISC_NEW] and [HEAR_NEW]
• Salience features and syntax: a crosslinguistic perspective
• Other possible features and the remaining challenge
Discourse functions in LFG

Dalrymple & Nikolaeva (2011): an s-structure contains a feature DF whose value (TOPIC, FOCUS, BACKGROUND or COMPLETIVE) specifies the discourse function of a referent, which in turn determines the role that the relevant meaning constructor will bear at the level of i-structure.

On this view, information structure (I.S.) categories are atomic.
Issues with atomic discourse functions

- An inventory of seemingly unrelated categories
- Empirically and theoretically well grounded?
- Are four types of discourse function sufficient?
- An approach which serves to obscure generalizations
In Nandi (Creider & Creider 1983), a difference in topic type determines c-structure position.

(1) Q: Who watched the lamb?
A: *ki:*rip *la:*kwɛ:t [*a:*rtɛ:t]TOPIC
   watched child.NOM lamb.OBL
   ‘The child watched the lamb’

(2) Q: What did the child do?
A: *ki:*rip *a:*rtɛ:t [*la:*kwɛ:t]TOPIC
   watched lamb.OBL child.NOM
   ‘The child watched the lamb’
Topic: atomic?

In Nandi (Creider & Creider 1983), a difference in topic type determines c-structure position.

\[(3) \text{[la:kwɛ:t } \text{kò]} \text{SWITCH}_\text{TOPIC} \; \text{ki:rìp} \; \text{a:rtɛ:t} \]

\[
\begin{align*}
\text{child.OBL} & \quad \text{PRT} & \quad \text{watched} \quad \text{lamb.OBL}
\end{align*}
\]

‘(As for) the child(, s/he) watched the lamb’
In Nandi (Creider & Creider 1983), a difference in topic type determines c-structure position.

(4) \textit{kikw}n\textit{ẹ} n\textit{ẹ} kityo ak\textit{ẹ} \textbf{[e:m\textbackslash:n\textbackslash:n]}\textsc{contin_topic}

is.amazing just another country.that
‘It’s just amazing, that country is something else’

\textbf{[ko an\textbackslash:ɛ: ak an\textbackslash:ɛ:]\textsc{switch_topic} \Lambda:me\textbackslash:n\textbackslash:ne}

\textsc{PRT} me and me I.live.at

\textit{komsit\textbackslash:ɔ p\textbackslash:v pe:lko\textbackslash:tin\textbackslash:ɛ}

side.this of Belgut

‘And as for me, I live right in this part of Belgut’
I.S. categories: discourse functions

**TopicE** (Topic Establisher): a ‘new’ topic, which did not have the status of topic in the immediately preceding discourse context; a category including Switch Topics.

**TopicC** (Continuing Topic): a current (ongoing), constant or repeated topic.

**FocusNI** (Focus New Information): provides information which is not assumed to be shared by the interlocutors.

**Backg.Info** (Background Information): not of primary importance to the discourse.
(5) “A long time ago there was a man whose daughter was in a dance.”

[Ja k’a rme’al]TopicE xukoj pa xajoj xin Tukun ‘(As for) his daughter, he entered (her) in the dance of Tecun’

y [ja rme’al]TopicC xoki Malincha. ‘and the daughter played the part of Malincha’

(Aissen 1992, citing Shaw 1971)
Switch Topic, Focus

In Garrwa, clause-initial position is a position of prominence, hosting a switch topic (6) or focus (7).

(6) “the reference to these two characters comes after a lengthy ‘aside’ concerning a type of dance some other characters were performing.”

[Mushin 2006]
In Garrwa, clause-initial position is a position of prominence, hosting a switch topic (6) or focus (7).

(7) Q: **What** did you get in the west?

A: \[wurumul\]_{FocusNI} =ngayu wurdumba=yi
bait =1SG get=PAST

bayangarri
west.side
‘I got bait over in the west’

(Mushin 2005)
(8) Q: What did Kathryn buy at the market?  
A: [Apples]_{FocusNI}.

[TopicC: Kathryn; Backg.Info: bought at the market]

(9) ‘Because [man]_{TOPIC} does not ever make [a sacrifice]_{FOCUS} for a woman …’
    aor nahīī kab hii de-gaa
    and not ever EMPH give-FUT.MSG
    ‘nor will (he)_{TopicC} ever make (a sacrifice for a woman)_{Backg.Info}.’

(Butt & King 1997)
Another approach: information features

In key work on I.S. in LFG, Choi (1999) and Butt & King (1996) propose to decompose discourse functions, defining them in terms of binary information features.

In both cases, they restrict their classificatory system to two information features: [NEW] and [PROMINENT].

Their resulting classifications are different.
**Information features: Choi (1999)**

<table>
<thead>
<tr>
<th>[NEW]</th>
<th>“new or novel information”</th>
</tr>
</thead>
<tbody>
<tr>
<td>[PROM]</td>
<td>“prominence or salience”, “singled out among potential alternatives”</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>[+NEW]</th>
<th>[−NEW]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+PROM] CONTRASTIVE FOCUS</td>
<td>TOPIC</td>
</tr>
<tr>
<td>[−PROM] COMPLETIVETE FOCUS</td>
<td>TAIL</td>
</tr>
</tbody>
</table>
Information features: Butt & King (1996)

- **NEW**: new or novel information
- **PROM**: prominence or salience

<table>
<thead>
<tr>
<th></th>
<th>[+NEW]</th>
<th>[−NEW]</th>
</tr>
</thead>
<tbody>
<tr>
<td>[+PROM]</td>
<td>Focus</td>
<td>Topic</td>
</tr>
<tr>
<td>[−PROM]</td>
<td>Compleative Information</td>
<td>Background Information</td>
</tr>
</tbody>
</table>
We retain the notion of ‘prominence’ with respect to the relative status of units of information, which can be speaker determined, but we refer to this as **saliency**.

We reserve **prominence** for reference to the means by which an item may be ‘highlighted’, e.g. by occupying a particular syntactic position (c-structure), or by possessing particular prosodic properties (p-structure).
Our information features: salience

**ABOUT**
The proposition expressed is about this meaning, in the sense that it represents a matter of current concern and is the pivot for truth-value assessment.

**UPDATE**
This meaning provides an information update that develops the communication due to its novel information structure status relative to the current discourse context.
Our information features: newness

**DISCOURSE NEWNESS (DISC_NEW±)**
A meaning that has not previously been evoked in the discourse.

**HEARER NEWNESS (HEAR_NEW±)**
A meaning that cannot reasonably be assumed to be already known to the addressee.

These two features are based on proposals in Prince (1992), which reframes her earlier taxonomy (Prince 1981).
Discourse functions defined in terms of our information features

<table>
<thead>
<tr>
<th>Discourse Function</th>
<th>salience</th>
<th></th>
<th>newness</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABOUT</td>
<td>UPDATE</td>
<td>DISC_NEW</td>
</tr>
<tr>
<td>TopicE</td>
<td>+</td>
<td>+</td>
<td>±</td>
</tr>
<tr>
<td>FocusNI</td>
<td>−</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>TopicC</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>Backg. Info</td>
<td>−</td>
<td>−</td>
<td>±</td>
</tr>
</tbody>
</table>
## Classic information structure distinctions defined

<table>
<thead>
<tr>
<th></th>
<th>ABOUT +</th>
<th>ABOUT −</th>
<th>ABOUT −, UPDATE +</th>
<th>ABOUT + and/or UPDATE −</th>
<th>UPDATE +</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Topic</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Comment</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Focus</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>(Back)Ground</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>‘Newsworthy’</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Dalrymple & Nikolaeva (2011) claim that “semantic structure [is] the proper level for the representation of features relevant for the definition of discourse referents”.

We do not employ the s-structure feature $DF$ with its set of possible atomic values.

Instead, we propose a set of binary-valued features based on the information features that have just been introduced.
S-structure features and the syntax–I.S. interface

These s-structure features are available to associate particular c-structure positions not only with feature specifications that define ‘classic’ discourse functions like Focus, but also with A-V pairs which capture other important facts about the syntax–I.S. relationship of the type presented in Herring (1990) or those discussed in Gundel (1987).
S-structures for Topics (Nandi)

(2) TopicC

child: [ABOUT+, UPDATE−]

(3) TopicE

child: [ABOUT+, UPDATE+]

Ordering principle:
[ABOUT+, UPDATE+] > [ABOUT−] > [ABOUT+, UPDATE−]
S-structures for Garrwa

(6) **TopicE**

those two:

\[
\begin{bmatrix}
\text{ABOUT} & + \\
\text{UPDATE} & +
\end{bmatrix}
\]

(7) **FocusNI**

bait:

\[
\begin{bmatrix}
\text{ABOUT} & - \\
\text{UPDATE} & +
\end{bmatrix}
\]

Ordering principle: \([\text{UPDATE}+] > [\text{UPDATE}−]\)
(8) Q: **What** did Kathryn buy at the market? 
A: **[Apples]**FocusNI.

TopicC: Kathryn  [ABOUT+, UPDATE–]
Backg.Info: bought at the market  [ABOUT–, UPDATE–]

(9) **aor nahīī kab hii de-gaa**
and not ever EMPH give-FUT.MSG
‘nor will (he)**TopicC** ever make (a sacrifice for a woman)**Backg.Info.’

TopicC: he  [ABOUT+, UPDATE–]
Backg.Info: a sacrifice for a woman [ABOUT–, UPDATE–]
Crosslinguistic generalizations
(Herring 1990)

<table>
<thead>
<tr>
<th>Word order (GFs)</th>
<th>Tendencies in relation to I.S.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVO</td>
<td>• topics in initial position</td>
</tr>
<tr>
<td>SOV</td>
<td>• topics in initial position</td>
</tr>
<tr>
<td></td>
<td>• focus is immediately preverbal</td>
</tr>
<tr>
<td></td>
<td>• backgrounded, predictable information including TopicC may appear postverbally</td>
</tr>
<tr>
<td>VSO, VOS, OVS, ‘free’</td>
<td>• focus–topic order, but any Switch Topic is initial</td>
</tr>
</tbody>
</table>
Capturing crosslinguistic generalizations using our features

<table>
<thead>
<tr>
<th>Word order (GFS)</th>
<th>Generalizations: s-structure features</th>
</tr>
</thead>
</table>
| SVO              | • initial: [ABOUT+]  
                  | [ABOUT+] > [ABOUT−] |
| SOV              | • initial: [ABOUT+]  
                  | [ABOUT+] > [ABOUT−]  
                  | • immediately preverbal: [ABOUT−, UPDATE+]  
                  | • final: [UPDATE−, DISC_NEW−, HEAR_NEW−] |
| VSO, VOS, OVS, ‘free’ | • initial: [ABOUT+, UPDATE+]  
                          | [UPDATE+] > [UPDATE−] |
Gundel’s (1987) universals of topic-comment structure

“in all languages, syntactic constructions which code topic-comment structure have one of three basic functions:

(i) to place new ... topics at the beginning of the sentence;
(ii) to place topic (either old or new) before comment;
(iii) to place focus at the beginning of the sentence and old, already established topics at the end.”
Gundel’s (1987) universals of topic-comment structure

(i) place new topics at the beginning of the sentence
   initial: [ABOUT+, UPDATE+]
(ii) place topic (either old or new) before comment
    [ABOUT+] > [ABOUT−]
(iii) place focus at the beginning of the sentence and old, already established topics at the end
    initial: [ABOUT−, UPDATE+] ... final: [ABOUT+, UPDATE−]

“There is apparently no language which has a construction whose primary function is to place new topics at the end of a sentence.”

*final: [ABOUT+, UPDATE+]
A possibility to explore

Other features to capture (the introduction/maintenance/shift of) the deictic centre, or possibly related concepts (perspective, point of view, speaker empathy, narrative perspective).

Such features may help in the analysis of, for example, the obviative/proximate distinction (see Dahlstrom 2003 on Meskwaki) or maybe mirativity (Delancey 2012).
Conclusion

Using the proposed binary-valued information features, we can define four key discourse functions and also capture important crosslinguistic generalizations about ordering according to I.S. status.

Result: a more fine-grained approach to I.S. categories based on well-grounded binary features.
The remaining challenge

How do we integrate these proposals with an approach to the formal representation of i-structure within the LFG architecture?

Lowe & Mycock presentation at *The Syntax and Information Structure of Unbounded Dependencies* Workshop on Sunday
Acknowledgements

The authors gratefully acknowledge financial support from Ministerio de Ciencia e Innovación Grant FFI2011-23046 (P.I. Alex Alsina).


