Information Structure

Ling 567
February 22, 2010

Overview

- What is Information Structure?
- Topic and Focus
- Representations
- Information structure marking constructions
- Variable property mapping

Information structure

- Information structure is the study of how speakers structure sentences to convey new information linked to the preceding context by old information.

- Many (all?) languages have particular linguistic resources (morphological, syntactic or prosodic) which facilitate the encoding of information structure.

- Lambrecht (1996) describes information structure as that which distinguishes between “allosentences”: sentences with the same truth conditions that are nonetheless not felicitous in the same set of contexts.

    - Kim read a book.
    - It was Kim who read a book.
    - It was a book that Kim read.

Topic and Focus

- There is as yet little consensus in the literature about the range of information structural distinctions that must be drawn in an adequate model of language, nor on the definition of several key concepts.

- In the spirit of incremental development, we will begin by modeling each element in a sentence as topic-marked, focus-marked, or unmarked.

- Not worrying for now about:
    - Correlations between grammatical functions and info-str
    - Focus projection

Topic and Focus

- Topic and focus are properties of *linguistic expressions* that refer to referents, rather than of the referents themselves.

- Topic: A linguistic expression which refers to a known or inferable referent that the rest of the sentence provides more information about.

- Focus: The new information asserted by the speaker, against the background of that which is presupposed to be shared in the common ground (topic and tail).

    - Focus test: The part of the sentence answering an appropriate wh question is the focus: Who left? Ivan left.

- Tail: The rest of the sentence, that which is neither focus nor topic.

As-for topics

- In some languages/sentences, the topic element already plays a role in the clause (argument or adjunct).

- In other languages/sentences, the topic only plays the role of topic to the clause. It should be linked via a “_topic_p_rel”.

- In English, this is marked with “as for”.

- In other languages, as-for topics can take on ordinary topic marking.

Amerika ha supiido suketaa ga hayai
America TOP speed skater NOM fast

‘As for America, the speed skaters are fast.’

NB

- Every sentence is presumed to have a focus (possibly the entire sentence), but not every sentence has focus overtly marked.

- Not every sentence has a topic. Furthermore, in many languages at least, topics don’t need to be overtly marked as such.

- Our “unmarked” means “not overtly marked for topic or focus”, rather than “neither topic nor focus”.

Representations

- We will be representing information structure with a feature on individual variables, appropriate for both *ref-ind* and *event* variables.

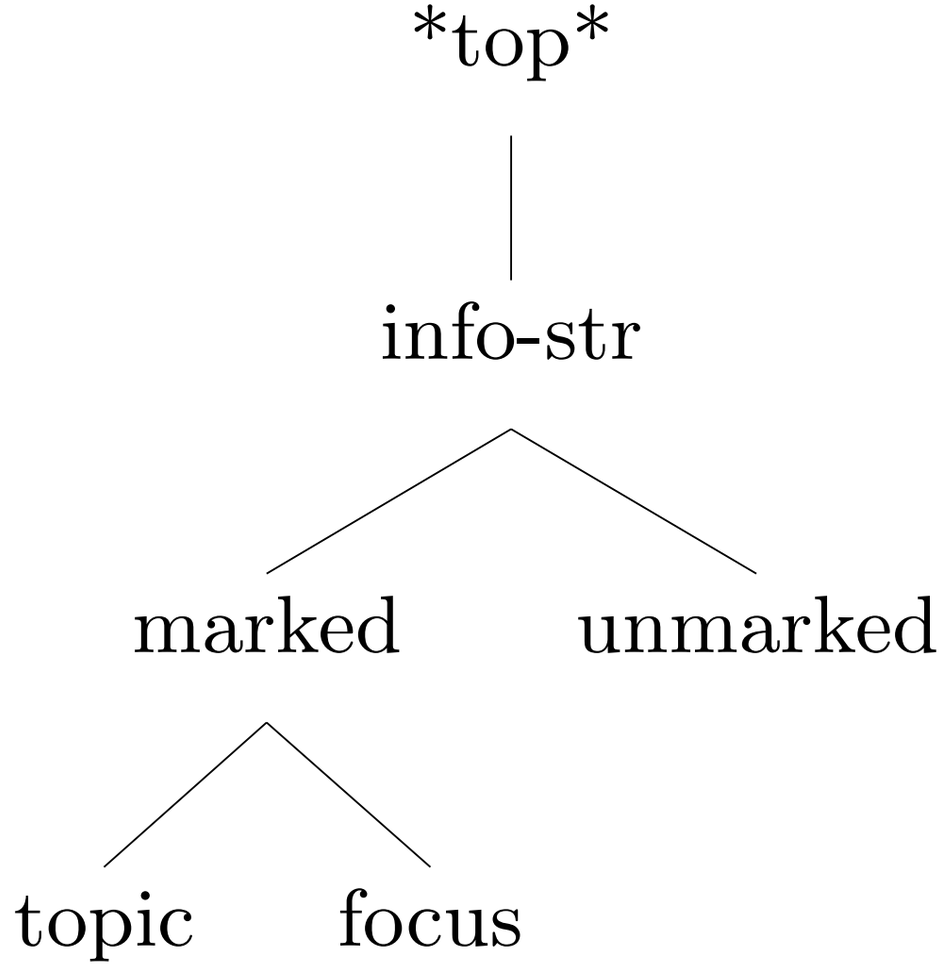
- Almost all elementary predications introduce a “characteristic variable” (the ARG0), so this will allow us to pick out individual parts of the sentence. (The exception is quantifiers.)

- This makes it available in the MRS, for translation purposes, and also available in the syntax, so it can be constrained by various means.

Representations

```
individual :+
   [ INFO-STR info-str ].

info-str := *top*.
marked := info-str.
unmarked := info.str.
topic := marked.
focus := marked.
```



Information structure marking constructions

- Distinguished position in the sentence: Sentence initial, pre-sentence, pre-verbal, post-verbal, post-sentence

- Clitics or adpositions

- Cleft constructions (outside the scope)

- Focus prosody (can pseudo-model)

- NB: In some languages, information structure is more grammaticalized than in others.

Distinguished position

- Pre- or post-verbal: add a rule type which insists on a lexical ([LIGHT +]) verb and constrains the INFO-STR of the non-head daughter. Cross-classify this type with all appropriate head-nexus phrases.

- Sentence initial position: Modify existing rules for this position to cosntrain INFO-STR of left daughter.

- Pre-sentence position: Add a new phrase structure rule (possibly a head-filler rule, possibly an as-for topic rule) which licenses one constituent before the rest of the clause and constrains the INFO-STR on that constituent.

Clitics or adpositions

- Clitics (modifiers?) constrain the INFO-STR of the head they modify.
- Russian “li” and Nishnaabemwin “na” are of this type.
- NB: Probably need two entries for “li” and “na” to capture ambiguity of verb-attached question clitics (between verb focus and sentence focus).
- Adpositions constrain the INFO-STR of their complement.
- In as-for topic constructions (see below), the adposition would serve as a modifier for the whole clause, introducing the “_topic_p_rel”.
- In other topic constructions, the adposition functions like a case-marking one, passing up the semantic information of its complement.

Focus prosody

- In many languages, prosodic emphasis on the focussed element might be the only systematic mark of information structure.

- Such prosody is not ordinarily marked in orthography, but for the purposes of this lab, we will mark it with a suffix -FP.

- The goal is to translate Russian “Ivan li kupil sobaku” to “Did Ivan-FP buy a dog?” and “Sobaku-li Ivan kupil” to “Did Ivan buy a dog-FP?”

- Model via a infl-ltol-add-only-no-ccont-lex-rule, constrained to apply last in the pipeline, which adds the -FP affix and contributes the information that the INDEX is [INFO-STR focus].

As-for topics and pro-drop

- If your language has both as-for topics and pro-drop, there is a decision to be made about how much ambiguity to accept.

- Strategy one: Allow both as-for topics and regular (clause-connected) topics, getting two parses whenever the position the topic is linked to is not overtly filled.

- Strategy two: Treat all topics as as-for topics, and write extra transfer rules for the MT demo.

Ohno ha hayai
Ohno TOP fast

‘As for Ohno, he is fast.’
‘Ohno is fast.’

Variable property mapping

- In MT, transfer rules rewrite MRS substructures to MRS substructures, changing PRED values, the number of elementary predications, and/or the linking of variables and handles between predications (and handle constraints).

- Transfer rules cannot rewrite feature values on variables (png, tam, sf, cog-st, info-str ...)

- This is handled by “variable property mapping”, or vpm.

- Each grammar is associated with a semi.vpm file which specifies the mapping between that grammar’s variable properties and an interlingual set.

- Once the semi.vpm file is invoked, only properties specifically handled are passed through. Everything else is suppressed.

Variable property mapping

- If a variable property is underspecified, the LKB will find all strings that are consistent with it.

- The information structure strategy outlined here has overt topics and foci marked as such, but everything else left as [INFO-STR info-str].

- In fact, we want an overt topic or focus only if one was given in the input.

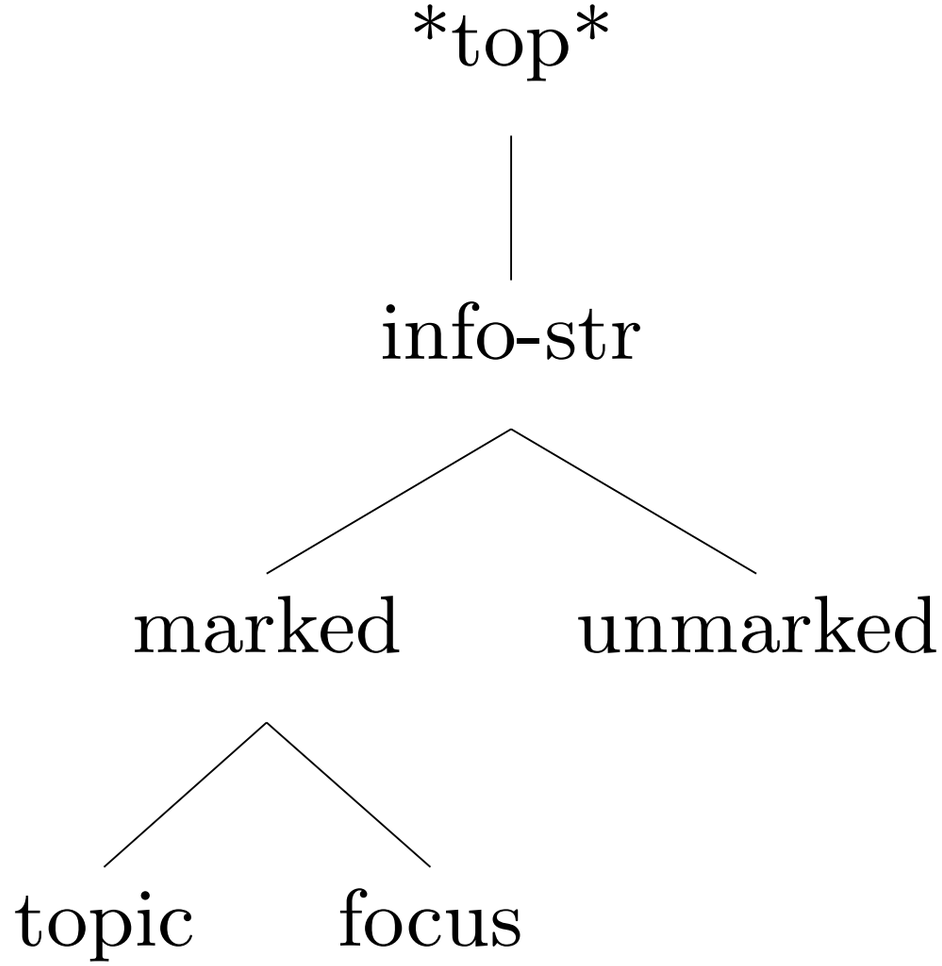
- Use vpm to map [INFO-STR info-str] to [INFO-STR unmarked].

- Worry about what happens when translating between a language that obligatorily marks topics/foci and one that doesn’t.

Representations

```
individual :+
   [ INFO-STR info-str ].

info-str := *top*.
marked := info-str.
unmarked := info.str.
topic := marked.
focus := marked.
```



Variable property mapping

- Most specific first
- <>: bidirectional
- >>: forward only
- <<: backward only
- *: whatever was matched
- [e]: event variable with no TENSE
- !: do not insert (forward)/match absence of property (backward)

```
E.TENSE : TENSE
    past        <> past
    present     <> pres
    future      <> fut
    real_tense <> tensed
    untensed    <> untensed
    *           >> untensed
    untensed    << *
    untensed    << [e]
```

INFO-STR and vpm

- Already encoded in provided semi.vpm (download from lab8 web page)

- Note that you might need to modify other aspects of this vpm, especially PNG.

```
INFO-STR : INFO-STR
  topic <> topic
  focus <> focus
  unmarked << info-str
  info-str >> unmarked
```

Overview

- What is Information Structure?
- Topic and Focus
- Representations
- Information structure marking constructions
- Variable property mapping