



Grammar Engineering: Applications & Complex Predicates vs. Modals vs. Auxiliaries

**Miriam Butt
(University of Konstanz)
and
Martin Forst (NetBase Solutions)**

Colombo 2014

Possible Applications

- LFG grammars encode deep linguistic information
- They are not easy to build since:
 - one needs to be a good grammar engineer
 - one needs to understand the structure of the language
- But very useful for applications which need deep linguistic knowledge
 - Computer Assisted Language Learning (CALL)
 - Question-Answering
 - Language Documentation
 - Treebanking

Possible Applications

Demos:

- CALL grammar for English
- Murrinh-Patha Electronic Resources
- Question-Answering System: PARC-Bridge
- ParGram Bank

The Verbal Complex

- The Verbal Complex in South Asian Languages (SALs) tends to be quite intricate.
- Multifunctionality of morphology
 - Complex stacking of morphology
- Multifunctionality of lexical items
 - many items can function as
 - » full verbs
 - » auxiliaries
 - » modals
 - » light verbs
- Understanding the difference and encoding it is not trivial but essential.

Auxiliaries

- Auxiliaries come in a variety of forms
- They also have a variety of functions
- Most common:
 - tense auxiliaries
 - aspect auxiliaries
 - passive auxiliaries
- Tense Auxiliaries situate an event in time
 - present (event coincides with speech time)
 - past (event precedes speech time)
 - future (event is projected to be after speech time)

Tense and Auxiliaries

- Not all languages have all three types of tense auxiliaries.
- In SALs, the tense auxiliaries typically are a form of the verb 'be'.
- Tense can also be marked on the verb via inflectional morphology.
- The future (auxiliary or morphology) often shares characteristics with modals:
 - the future indicates a possible future world/situation
 - a subtype of modals indicate possibility ('can', 'might', etc..)
 - so there is semantic overlap

Aspectual Auxiliaries

- Aspect situates an event with respect to some reference time.
- Aspect generally encompasses:
 - perfect
 - imperfect
 - progressive
- In SALs, aspectual auxiliaries typically denote the progressive and are derived from a verb like 'stay', 'live'.

Tense and Aspect via E, R, S

- It may be helpful to use a (simplified) version of the Reichenbachian/Kamp system to understand the differences.
- Three entities are postulated:
 - event time (E)
 - speech time (S)
 - reference time (R) [there can be multiple Rs)

Tense and Aspect via E, R, S

- Tense: Relationship between R and S
- Aspect: Relationship between E and R
- The systems interact (which is what makes tense/aspect complicated).

Present _____E,R,S_____>

Past _____E_____R,S_____>

Future _____R,S_____E_____>

Present Perfect _____E_____R,S_____>

Past Perfect _____E_____R_____S_____>

Tense and Aspect via E, R, S

- Past and Present Perfect look very similar, they also often merge into one another as part of language change (cf. German).
- But subtle difference:

Present	E,R & R,S (He is small.)
Past	E,R & R < S (He was small.)
Future	E,R & R > S (He will be small.)
Present Perfect	E < R & R,S (He has been small.)
Past Perfect	E < R & R < S (He had been small.)

Tense and Aspect via E, R, S

- The present and the progressive are also very similar.
- Some languages do not make a difference (e.g., German).
- In some languages the progressive is used when an event rather than a state is involved.
- The event is happening at speech time, but extends before and after it as well.

Present

He is small.

Progressive

He is running.

Tense and Aspect Generally

- Tense/Aspect (and Mood) is a very difficult topic.
- However, it is central for language processing.
- The previous slides present a simplified picture.
- But in figuring out a language, one should start with a simple picture and try to sort out things from there.

Tense and Aspect in LFG

- Tense and Aspect Auxiliaries are **functional elements**.
- They serve to say something more about the event/state described by the main verb.
- A simple way to model this is for auxiliaries to:
 - NOT contribute a PRED value
 - contribute information about Tense/Aspect in terms of feature values at f-str

Modals

- Modals serve to take an event and say something about:
 - whether there is an obligation to do the event (deontic modals; *must, should*)
 - possibility/ability for the event to happen (*can, might*)
- They are not functional elements, but are generally treated as full verbs.
- Modals generally take an XCOMP.
 - Kim has [to go to school].
 - Kim must [go to school].

Modals vs. Tense/Aspect

- Modals do not situate an event in time.
- Modals express information about the likelihood of an event happening.
- In LFG
 - modals contribute a PRED
 - have an XCOMP in their subcategorization frame
 - tense/aspect auxiliaries do neither
- See Bhatt et al. (2011) for a discussion of Urdu modals in LFG.

References

Bhatt, Rajesh, Tina Bögel, Miriam Butt, Annette Hautli, and Sebastian Sulger. 2011. Urdu/Hindi Modals On-line LFG 2011 Proceedings, Hong Kong. Pages 47-67.

<http://csli-publications.stanford.edu/LFG/16/lfg11-toc.html>.

Sebastian Sulger, Miriam Butt, Tracy Holloway King, Paul Meurer, Tibor Laczko, György Rakosi, Cheikh Bamba Dione, Helge Dyvik, Victoria Rosén, Koenraad De Smedt, Agnieszka Patejuk, Özlem Cetinoglu, I Wayan Arka and Meladel Mistica: ParGramBank: The ParGram Parallel Treebank. *Proceedings of ACL 2013 (Long Papers)*, Sofia, Bulgaria.

<http://aclweb.org/anthology/P/P13/P13-1054.pdf>.