

# Word order phenomena in grammar design

Antske Fokkens

Department of Computational Linguistics  
Saarland University



# Outline

- 1 General Topic and Thesis overview
- 2 Word Order Library in the Matrix
  - Auxiliaries
- 3 Current and near future work
- 4 Related research questions
- 5 The future



## Topic

- Working title: Word Order phenomena in Grammar Design
- First part of thesis: a "complete basic" word order library for the grammar matrix
- Second part of the thesis: to be determined



# Timeline

- End of summer:
  - basic word order library finished
  - second part of thesis decided upon
- Summer 2010: bulk of PhD research done



## Previous state of word order library

- The six basic word orders: SOV, SVO, VSO, VOS, OVS, OSV
- Verb-final, Verb-initial, free word order
- Determiner-Noun order
- Auxiliary-verbal complement (working for harmonic order)



## Additions to Word order library

- A v2-analysis (Warlpiri based)
- Auxiliary-verb analyses
  - "Disharmonic" word order (where head-comp is different for verb-object and auxiliary-vcomp)
  - V-complements, notably verbal clusters



# Auxiliaries: Implementations

- Auxiliaries can precede or follow the verb for all basic word orders, but v2
- For V2 languages, the auxiliary that heads the sentence is in second position, else their order is free
- For many combinations, word order depends on the kind of complement the auxiliary has (v, vp, or s)
- Verb clusters (where all verbs must be adjacent) are formed for:
  - Verb-initial and verb-final languages
  - (Optionally) for VSO and OSV, and free word order languages



## Auxiliaries: Questions

- What are auxiliaries?
- Word order and auxiliary complement (s, vp, v) are highly related: is this right?
- Harmonic word order is straight-forward, non-harmonic order for auxiliaries having a v-compl is tricky: but (how much) does this occur?
- V-clusters seem a natural phenomena: does the analysis reflect this?
- The verbal cluster feature registers whether a specific element has been picked up by the head daughter: should it be more general?





## Auxiliaries: To Do

- For (at least) free word order: it should be possible to have auxiliaries that may precede or follow their complement
- For other word orders: the possibility to have complement-preceding and complement-following auxiliaries
- For v2 languages: verbal cluster in second or final position:
  - This occurs in Germanic languages: is the Austronesian based v2 analysis right for them?



## Current work

- Checking word order related phenomena of negation, yes-no questions (+ add Germanic v2)
- Typological reading on auxiliary structures
- Typological research on general word order phenomena (WALS)



# Phenomena for the near future

- Adpositions
- Adjectives
- Ditransitive verbs
- Free subject-verb order, Free object-verb order
- Word order interactions with optional complements



## Potential further extension of word order page

- Adverbs
- Subordinate clauses
- Compositional semantics and how to cope with discontinuity
- Interaction between word order and other phenomena (e.g. Turkish fixed word order when nouns do not bear case marking)



## Typological questions

- What function does word order have?
- What limits are (typically) found in free word order languages?
- The 2nd position:
  - Are there languages that are 'predicate 2', instead of just v2?
  - What about the 2nd position for clitics in Slavic languages?

How to use Typology?



# Language Description Questions

- Do language description methods provide insight to possibilities and limits, and frequencies of particular word order phenomena?
- To what extent is the questionnaire influenced by HPSG, LKB, or particular implementations?
- Can more general ('free') description techniques be incorporated in the questionnaire, and be used to derive an implementation?



## 'Flexible' restrictions on word order

- How free is free word order? How to incorporate notions such as 'gradient grammaticality'?
- The influence of context, information structure and prosodic factors in acceptability
- Can statistics be used when there are significant difference in acceptability?
  - To guide generation and only form 'natural sounding' strings?
  - To reduce spurious ambiguity and support disambiguation in parsing
- Do statistics correlate with the facility to process a string for human beings?



## Possible future directions

- Continue expanding the word order library: subordinates, relative clauses, adverbs, Finnish definiteness, Turkish interactions with morphology (more possibilities than time and space in one PhD)
- "Play with" other tools TRALE, XLE
- Gradient grammaticality/Interaction with information structure (theoretically)





## Other possible future directions

- Gradient grammaticality and parsing of ungrammatical strings: which constraints are often violated?
- Optimality for free word order and quasi-free word order languages (what is the ideal implementation of free word order?)
  - Generation from free word order
- Psychological plausible parsing

