

Grammar Development with LFG and XLE

Miriam Butt University of Konstanz

Last Time

- Lexical Rules (Passive, Dative Shift)
- Different Types of Functional Equations
 - defining
 - constraining
 - negative
 - existential
- Testsuites

This Time: Lesson 4

- I. LFG Treatment of Adjuncts
 - Adjectives and Adverbs
 - PPs
- 2. PP OBL vs. PP Adjunct
- 3. Semantically heavy vs. light Ps
- 4. Tokenization: Sentence Punctuation

Tokenization

- So far we have parsed sentences that do not contain any punctuation.
- But this needs to be integrated into the grammar.
- Dealing with punctuation is part of tokenization.
 - capitalization
 - multiword items like San Francisco
 - abbreviations like *Mr., Dr.*
 - etc.

Tokenization

- Tokenizers in the ParGram grammars are generally implemented via finite-state transducers.
- They interact with the finite-state morphological analyzers (cascades of transducers).
- Creating good tokenizers is an art of its own.
 - Not part of this course.
 - The default tokenizers provided with XLE are very good.
 - But if working with different languages, will eventually have to build your own.

Adjuncts vs. Arguments

- So far we have parsed sentences that consisted of:
 - a verbal head/predicate
 - NP arguments of that verbal head
- Arguments have the property that each type may only appear once.
 - A clause does not have two subjects
 - or two objects, etc.

Adjuncts vs. Arguments

- But consider sentences like:
 - » The small, grey, intelligent dog devoured bones.
 - » The dog devoured large, delicious, juicy bones.
- Adjectives of the same type can appear multiple times.
- In principle, they can appear infinitively many times.
- In LFG adjectives are treated as Adjuncts.

Types of Adjuncts

- Other common types of adjuncts are adverbs and PPs.
 - » The very, very, very large, absolutely totally black dog appeared.
 - » The dog barked [in the garden] [under the tree] [on a rock].
- Recall that so far we have treated PPs as OBL.
- An OBL is a governable grammatical function.
- PPs can function both as OBL or as an Adjunct.
- Sometimes it is not easy to tell the difference.

Various kinds of PPs

- There are two types of OBL PPs
- Both types are characterized by the fact that the PP is required by the verb as an argument.
 - In one type the P expresses a (generally spatial) meaning:
 - » the gorilla put the banana in the tree
 - » *the gorilla put the banana
 - In the other type the P has no separate or only a weak/light meaning:
 - Ex: the meaning of serving somebody in a restaurant
 - » the maitre'd waited on the customer
 - » *the maitre'd waited on the customer on the boss
 - » the maitre'd waited (does not mean serve)

Various kinds of PPs

- PPs that function as Adjuncts:
 - the P expresses a meaning of its own
 - any number of PPs can be used
 - the verb does not require the PPs
- Adjunct Example
 - spatial sense of 'on'
 - wait with on not in the sense of serving a customer
 - » the maitre'd waited on the bench on the bank of the river
 - » the maitre'd waited

Adjunct PPs

- The Ps in Adjuncts are assumed to have a PRED.
- This is because they contribute their own semantics.
- In many languages, they also impose case restrictions on their arguments (like verbs do).

Case Requirements

- In German, Ps require their objects to have a specific case.
- Some Ps allow for more than one case and this can give rise to meaning differences.

Der Hund rannte in dem Garten. the.Nom dog ran in the.**Dat** garden 'The dog ran around in the garden.'

Der Hund rannte in den Garten. the.Nom dog ran in the.**Acc** garden 'The dog ran into the garden.'

Adjunct PPs

- The Ps in Adjuncts are assumed to have a PRED.
- The PRED subcategorizes for an object.
- For example:
 - (^ PRED) = `on<(^OBJ)>'

```
000
                             X 1 valid F-structure for S
kill prev next Commands Views 🗆 a 💷 c 💷 n 💷 s 💷 🗴
  lock F-structure #1
  "the dog slept on the bed"
                'sleep<[1:dog]>'
         PRED
                PRED 'dog'
1CASE nom, DEF +, NTYPE count, NUM sg, PERS 3
         SUBJ.
                   [PRED 'on<[9:bed]>'
                          PRED 'bed'
         ADJUNCT
                 { OBJ
                            ASE acc, DEF +, NTYPE count, NUM sg, PERS
                   PTYPE sem
       5MOOD indicative, TENSE past
```

OBL PPs

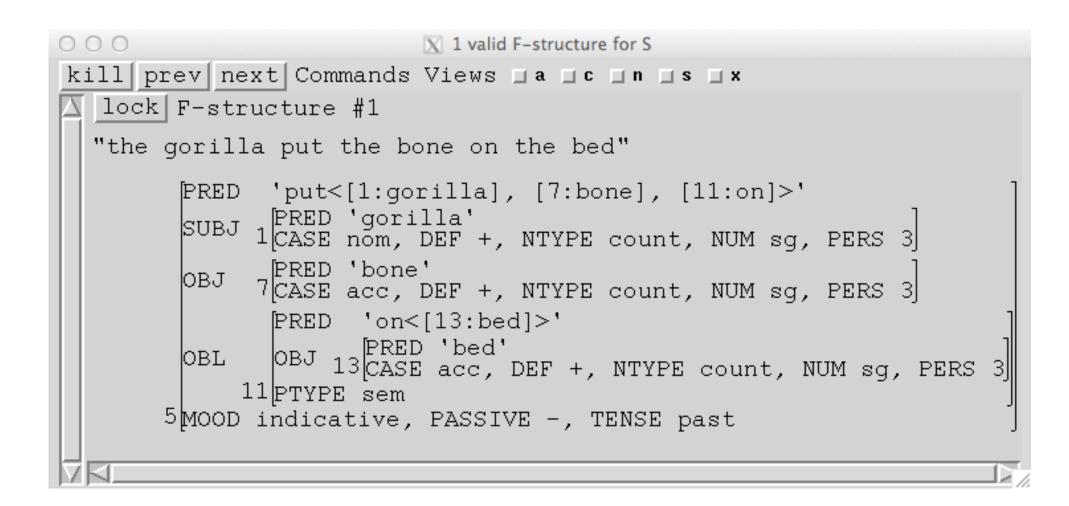
- The non-semantic (wait on) Ps in OBLs are assumed not to have a PRED.
- They are simply registered via a PFORM attribute.
- For example:

 $-(^{PFORM}) = on$

- The semantic Ps in OBLs are assumed to have a PRED.
- For example for put the banana in the tree – (^ PRED) = `in<(^OBJ)>'

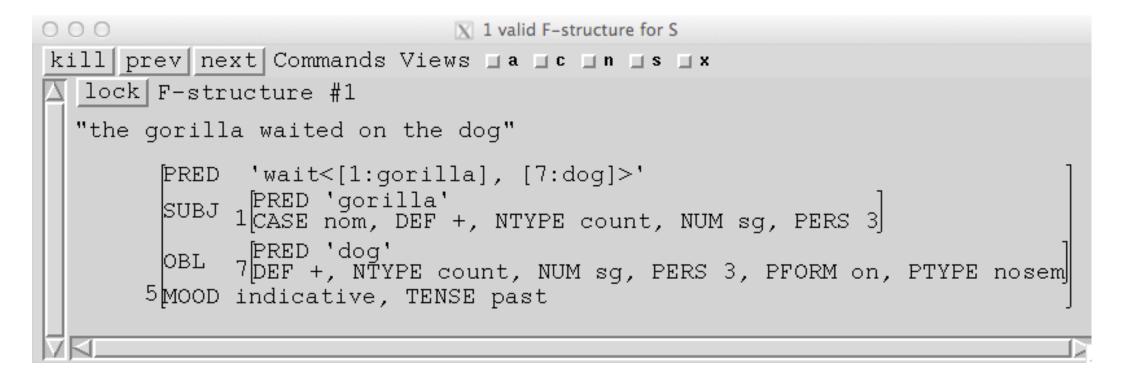


Example for semantic P with OBL:





Example for non-semantic P with OBL:



Dealing with Adjuncts

- Now let us return to the status of adjuncts in LFG.
- Recall that LFG allows only one value for any given attribute.
- **But:** Adjuncts can appear multiple times.
- Solution: These elements are projected into elements of set-valued attributes such as ADJUNCT, MOD, etc.
- LFG Notation: $\downarrow \in (\uparrow ADJUNCT)$
- XLE Notation: ! \$ (^ ADJUNCT)
 - down is an element of the adjunct set of the mother (up)



grammar3.lfg testsuite3.lfg

punctuation adjectives various types of PPs

Practical Work

- This concludes Lesson 4.
- The practical work you should do now is detailed in Exercise 4.
- You will practice with
 - adjectives and adverbs (adjuncts)
 - various types of PPs
 - punctuation