JMORF/MORS — Morpho-Syntax

Non-referential NPs, Expletives, and Extraposition

Francis Bond
Palacký University
https://fcbond.github.io/
bond@ieee.org

Lecture 10
Location: KB-1.11

Overview

- Review: the copula ***be***
- Existentials (*There is an X*, ...)
- Dummy ***it***
- Complementizers and CP subjects
- Extraposition (*It worries me that X*, ...)
- Idioms (*X takes advantage of Y*, ...)
- Problems

Review

Introduction

- In Chapter 10, we met the passive ***be***.
- Passive ***be*** is just a special case — that ***be*** generally introduces [PRED +] constituents.
- Today, we'll introduce another ***be***, which occurs in existential sentences starting with ***there***, e.g. *There is a monster in Loch Ness.*
- Then we'll look at this use of ***there***.
- Which will lead us to a more general examination of NPs that don't refer, including some uses of ***it*** and certain idiomatic uses of NPs.

Chapter 10 Entry for *be*

$$\left\langle \text{be}, \left[\begin{array}{l} \textit{be-lxm} \\ \text{ARG-ST} \left\langle \boxed{1}, \left[\left[\left[\begin{array}{ll} \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \left[\begin{array}{ll} \textit{verb} & \\ \text{FORM} & \textit{pass} \end{array} \right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \left\langle \boxed{1} \right\rangle \\ \text{COMPS} & \left\langle \right\rangle \end{array} \right] \end{array} \right] \\ \text{SEM} & \left[\text{INDEX} \ s \right] \end{array} \right] \right] \right] \right\rangle \\ \text{SEM} \quad \left[\begin{array}{ll} \text{INDEX} & s \\ \text{RESTR} & \left\langle \right\rangle \end{array} \right] \end{array} \right] \right\rangle$$

This takes only passive lexemes: $\left[\text{FORM} \ \textit{pass} \right]$.

Copula (Generalized)

$$\left\langle \text{be}, \left[\begin{array}{l} \textit{be-lxm} \\ \text{ARG-ST} \left\langle \boxed{1}, \left[\left[\left[\begin{array}{ll} \text{SYN} & \left[\begin{array}{ll} \text{HEAD} & \left[\text{PRED } +\right] \\ \text{VAL} & \left[\begin{array}{ll} \text{SPR} & \left\langle \boxed{1} \right\rangle \\ \text{COMPS} & \langle\rangle \end{array} \right] \end{array} \right] \\ \text{SEM} & \left[\text{INDEX } s\right] \end{array} \right] \right] \right] \right\rangle \\ \text{SEM} \quad \left[\begin{array}{ll} \text{INDEX} & s \\ \text{RESTR} & \langle\rangle \end{array} \right] \end{array} \right] \right\rangle$$

- Allow also (**pred**icative) adjectives ***fond***, ***sleeping*** are $[\text{PRED } +]$
- ***mere*** is $[\text{PRED } -]$
- Most adjectives can be either

Existentials

Existentials

- The ***be*** in *There is a page missing* cannot be the same ***be*** that occurs in sentences like *Pat is tall* or *A cat was chased by a dog*. Why not?
- So we need a separate lexical entry for this ***be***, stipulating:
  - Its SPR must be ***there***.
  - It takes two complements, the first an NP and the second an AP, PP, or (certain kind of) VP.
  - The semantics should capture the relation between, e.g. *There is a page missing* and *A page is missing*.
- More examples:

  (1) *There was a person at the door.*

  (2) *There are many people fond of linguistics.*

  (3) *There are people looking at us.*

Lexical Entry for the Existential *be*

$$\left\langle \text{be}, \left[\begin{array}{l} \textit{exist-be-lxm} \\ \text{ARG-ST} \left\langle \left[\begin{array}{c}\text{NP}\\ \text{FORM } \textit{there}\end{array}\right], \boxed{1}, \left[\left[\left[\begin{array}{ll} & \text{HEAD } \left[\text{PRED } +\right] \\ \text{SYN} & \text{VAL } \left[\begin{array}{ll}\text{SPR} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle\rangle\end{array}\right] \\ \text{SEM} & \left[\text{INDEX } s\right]\end{array}\right]\right]\right] \right\rangle \\ \text{SEM} \quad \left[\begin{array}{ll}\text{INDEX} & s \\ \text{RESTR} & \langle\rangle\end{array}\right] \end{array}\right] \right\rangle$$

Questions About the Existential *be*

- What type of constituent is the third argument?
- Why is the third argument [PRED +]?
- Why is the second argument tagged as identical to the SPR of the third argument?

$$\left\langle \text{be}, \left[\begin{array}{ll} \textit{exist-be-lxm} & \\ \text{ARG-ST} & \left\langle \begin{bmatrix}\text{FORM} & \textit{there}\end{bmatrix}, \boxed{1}, \begin{bmatrix}\text{SYN} & \begin{bmatrix}\text{HEAD} & [\text{PRED} & +]\end{bmatrix} \\ & \begin{bmatrix}\text{VAL} & \begin{bmatrix}\text{SPR} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle\rangle\end{bmatrix}\end{bmatrix} \\ \text{SEM} & [\text{INDEX} \ s\,]\end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix}\text{INDEX} & s \\ \text{RESTR} & \langle\rangle\end{bmatrix} \end{array}\right] \right\rangle$$

There Are Questions Left

- What is the contribution of this ***be*** to the semantics of the sentences it occurs in?
- Can all [PRED +] predicates appear as the third argument in existentials?
- How do we rule out * ***There was a greyhound a good runner***?

$$\left\langle \text{be}, \begin{bmatrix} \textit{exist-be-lxm} \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{NP} \\ \text{FORM} & \textit{there} \end{bmatrix}, \boxed{1}, \begin{bmatrix} \begin{bmatrix} \text{SYN} & \begin{bmatrix} \text{HEAD} & [\text{PRED} & +] \end{bmatrix} \\ & \text{VAL} & \begin{bmatrix} \text{SPR} & \langle \boxed{1} \rangle \\ \text{COMPS} & \langle\rangle \end{bmatrix} \end{bmatrix} \\ \text{SEM} & [\text{INDEX} & s] \end{bmatrix} \end{bmatrix} \right\rangle \\ \text{SEM} & \begin{bmatrix} \text{INDEX} & s \\ \text{RESTR} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

The Entry for Existential *there*

$$\left\langle \text{there}, \begin{bmatrix} \textit{pron-lxm} \\ \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{FORM} & \textit{there} \\ \text{AGR} & \begin{bmatrix} \text{PER} & \textit{3rd} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \textit{none} \\ \text{INDEX} & \textit{none} \\ \text{RESTR} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

Questions About Existential *there*

- Why do we call it a pronoun?
- Why don't we give it a value for NUM?
- What does this entry claim is ***there***'s contribution to the semantics of the sentences it appears in?
- Is this a correct claim?

$$\left\langle \text{there}, \begin{bmatrix} \textit{pron-lxm} \\ \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{FORM} & \textit{there} \\ \text{AGR} & \begin{bmatrix} \text{PER} & \textit{3rd} \end{bmatrix} \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \textit{none} \\ \text{INDEX} & \textit{none} \\ \text{RESTR} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

Dummy *it*

Other NPs That Don't Seem to Refer

(4) *It sucks that the Rockies lost the series.*

(5) *It is raining.*

(6) *Andy took advantage of the opportunity.*

(7) *Lou kicked the bucket.*

What About *It follows that you are wrong*?

- This is an example of **extraposition**.

- To analyze it, we need:
  - An analysis of this use of ***that***[*]
  - Entries for verbs that take clausal subjects[*]

    (8) *That you are wrong follows.*
  - A lexical entry for dummy ***it***
  - A rule to account for the relationship between pairs like (8) and (9)

    (9) *It follows that you are wrong.*

[*]We need these anyway (independently motivated)

The Entry for Dummy *it*

$$\left\langle \text{it}, \left[\begin{array}{ll} \textit{pron-lxm} & \\ \textsc{syn} & \left[\textsc{head} \left[\begin{array}{ll}\textsc{form} & \textit{it} \\ \textsc{agr} & \textit{3sing}\end{array}\right]\right] \\ \textsc{sem} & \left[\begin{array}{ll}\textsc{mode} & \textit{none} \\ \textsc{index} & \textit{none} \\ \textsc{restr} & \langle\rangle\end{array}\right] \end{array}\right] \right\rangle$$

Questions About Dummy *it*

- How does it differ from the entry for dummy ***there***? Why do they differ in this way?

- Is this the only entry for ***it***?

$$\left\langle \text{it}, \begin{bmatrix} \textit{pron-lxm} \\ \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{FORM} & \textit{it} \\ \text{AGR} & \textit{3sing} \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \textit{none} \\ \text{INDEX} & \textit{none} \\ \text{RESTR} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

Complementizers and CP Subjects

A New Type of Lexeme: Complementizers

comp-lxm:

$$\begin{bmatrix} \textsc{syn} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{comp} \\ \textsc{agr} & \textit{3sing} \end{bmatrix} \\ \textsc{val} & \begin{bmatrix} \textsc{spr} & \langle\rangle \end{bmatrix} \end{bmatrix} \\ \textsc{arg-st} & \left\langle \begin{bmatrix} \textsc{s} \\ \textsc{index} & \textit{s} \end{bmatrix} \right\rangle \\ \textsc{sem} & \begin{bmatrix} \textsc{index} & \textit{s} \\ \textsc{restr} & \langle\rangle \end{bmatrix} \end{bmatrix}$$

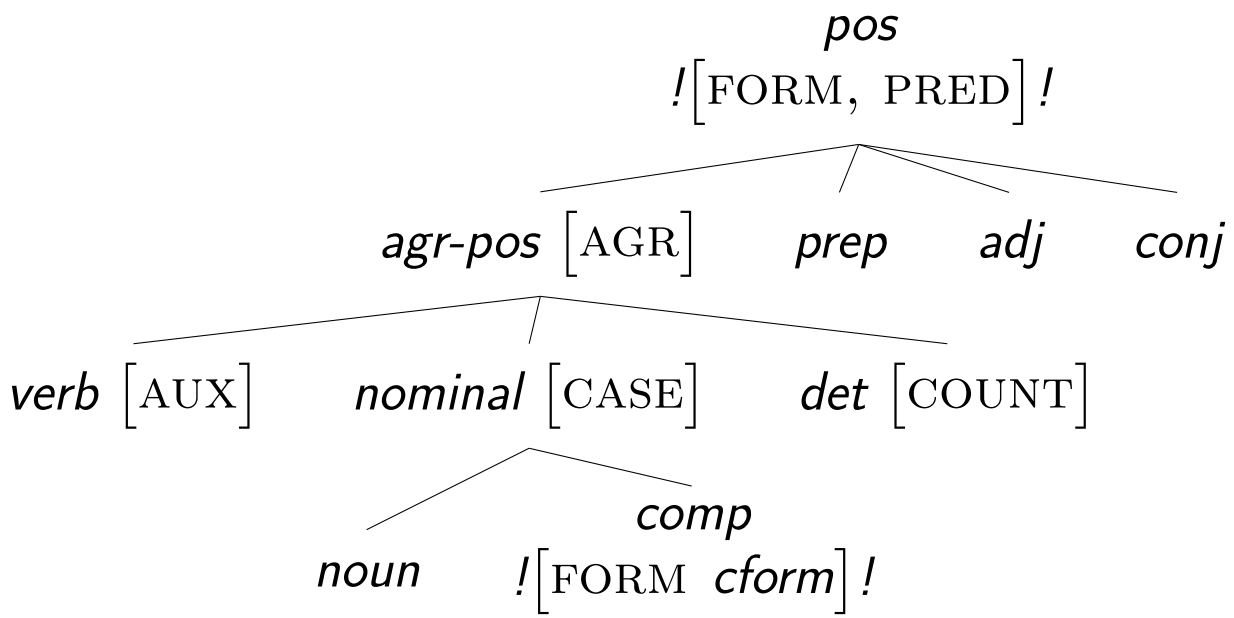
Questions About the Type *comp-lxm*

- Why does it stipulate values for both SPR and ARG-ST?
- Why is its INDEX value the same as its argument's?
- What is its semantic contribution?

comp-lxm:

$$\begin{bmatrix} \textsc{syn} & \begin{bmatrix} \textsc{head} & \begin{bmatrix} \textit{comp} \\ \textsc{agr} & \textit{3sing} \end{bmatrix} \\ \textsc{val} & \begin{bmatrix} \textsc{spr} & \langle\rangle \end{bmatrix} \end{bmatrix} \\ \textsc{arg-st} & \left\langle \begin{bmatrix} \textsc{s} \\ \textsc{index} & \textit{s} \end{bmatrix} \right\rangle \\ \textsc{sem} & \begin{bmatrix} \textsc{index} & \textit{s} \\ \textsc{restr} & \langle\rangle \end{bmatrix} \end{bmatrix}$$

The Type *comp*



The Lexical Entry for Complementizer *that*

$$\left\langle \text{that}, \begin{bmatrix} \textit{comp-lxm} \\ \textsc{arg-st} & \left\langle \left[\textsc{form} \ \textit{fin} \right] \right\rangle \\ \textsc{sem} & \left[\textsc{mode} \ \textit{prop} \right] \end{bmatrix} \right\rangle$$

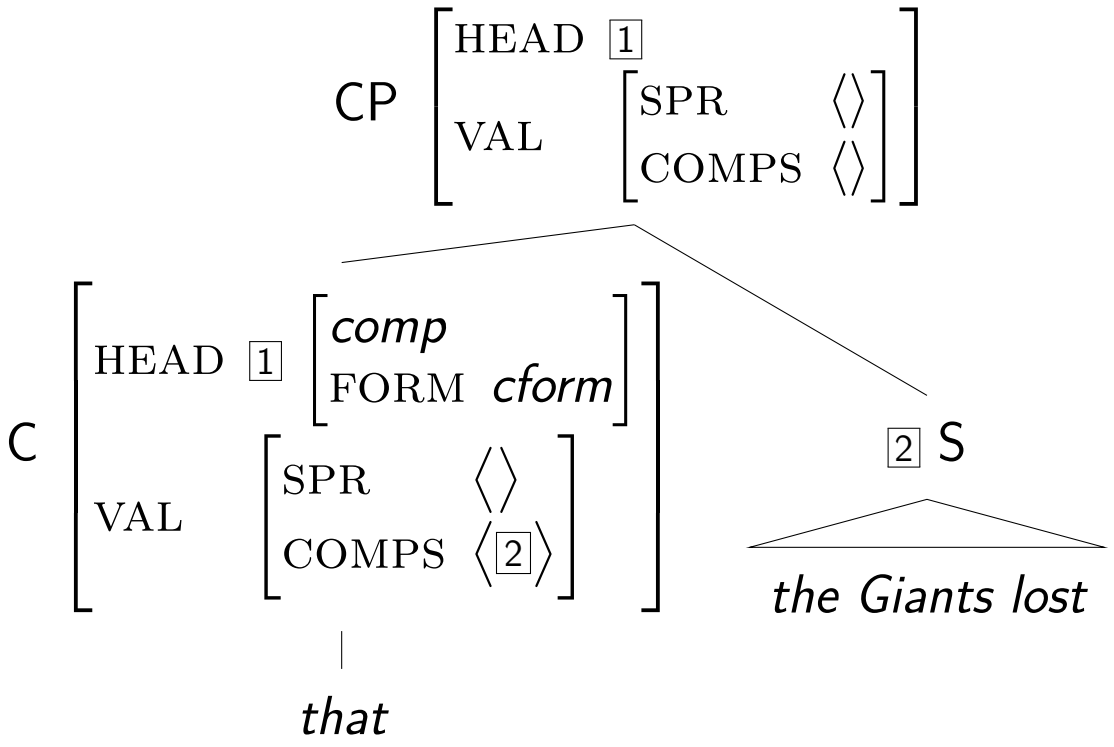
...with Inherited Information Filled In

$$\left\langle \text{that}, \begin{bmatrix} \textit{comp-lxm} & & \\ \text{SYN} & \text{HEAD} & \begin{bmatrix} \textit{comp} & \\ \text{FORM} & \textit{cform} \\ \text{AGR} & \textit{3sing} \end{bmatrix} \\ & \text{VAL} & \begin{bmatrix} \text{SPR} & \langle\rangle \end{bmatrix} \\ \text{ARG-ST} & \left\langle \begin{bmatrix} \text{S} & \\ \text{INDEX} & \textit{s} \\ \text{FORM} & \textit{fin} \end{bmatrix} \right\rangle & \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \textit{prop} \\ \text{INDEX} & \textit{s} \\ \text{RESTR} & \langle\rangle \end{bmatrix} & \end{bmatrix} \right\rangle$$

- Question: Where did [FORM *cform*] come from?

Structure of a Complementizer Phrase

(10) *that the Giants lost*



Sample Verb with a CP Subject

$$\left\langle \text{matter}, \begin{bmatrix} \textit{siv-lxm} \\ \text{ARG-ST} \left\langle \left[\text{SEM} \left[\text{INDEX } \boxed{1}\right]\right] \right\rangle \\ \text{SEM} \begin{bmatrix} \text{INDEX } s \\ \text{RESTR} \left\langle \begin{bmatrix} \text{RELN} & \textbf{\textit{matter}} \\ \text{SIT} & s \\ \text{MATTERING} & \boxed{1} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

Note: the only constraint on the first argument is semantic.

A Problem

- We constrained the subject of ***matter*** only semantically. However...
- CP and S are semantically identical, but we get:

  (11) *That Bush won matters* vs. **Bush won matters*

- Argument-marking PPs are semantically identical to their object NPs, but we get:

  (12) *The election mattered* vs. **Of the election mattered*

- So we need to add a syntactic constraint.

$$\left\langle \text{matter}, \left[\begin{array}{ll} \textit{siv-lxm} & \\ \text{ARG-ST} & \left\langle \left[\begin{array}{ll} \text{SYN} & \left[\text{HEAD} \quad \textit{nominal}\right] \\ \text{SEM} & \left[\text{INDEX} \ \boxed{1}\right] \end{array}\right] \right\rangle \\ \text{SEM} & \left[\begin{array}{ll} \text{INDEX} & s \\ \text{RESTR} & \left\langle \left[\begin{array}{ll} \text{RELN} & \textbf{\textit{matter}} \\ \text{SIT} & s \\ \text{MATTERING} & \boxed{1} \end{array}\right] \right\rangle \end{array}\right] \end{array}\right] \right\rangle$$

- S and PP subjects are generally impossible, so this constraint should probably be on *verb-lxm*.

Extraposition

Extraposition (at Last)

- Extraposition alters word order so that a relatively "heavy" constituent appears to the right of its canonical position.

(13) a. *That you were wrong follows.*
b. *It follows that you were wrong.*

(14) a. *That I mistyped it was frustrating.*
b. *It was frustrating that I mistyped it.*

(15) a. *Did that this happened surprise you?*
b. *Did it surprise you that this happened?*

Post-Inflectional Lexical Rule: *pi-rule*

$$\textit{pi-rule}: \left[\begin{array}{ll} \textit{l-rule} & \\ \textsc{input} & \left\langle / \boxed{0}, \left[\begin{array}{ll} \textit{word} & \\ \textsc{syn} & \left[\begin{array}{ll} \textsc{head} & / \boxed{1} \\ \textsc{val} & \left[\textsc{mod} \ \boxed{A}\right] \end{array}\right] \end{array}\right] \right\rangle \\ \textsc{output} & \left\langle / \boxed{0}, \left[\begin{array}{ll} \textit{word} & \\ \textsc{syn} & \left[\begin{array}{ll} \textsc{head} & / \boxed{1} \\ \textsc{val} & \left[\textsc{mod} \ \boxed{A}\right] \end{array}\right] \end{array}\right] \right\rangle \end{array}\right]$$

We use this when we expect the structure to be largely formed, and we want to change the order of things, ...

The Extraposition Lexical Rule

$$\left[\begin{array}{l} \textit{pi-rule} \\ \textsc{input} \quad \left\langle \mathrm{X}, \left[\textsc{syn} \left[\textsc{val} \left[\begin{array}{ll}\textsc{spr} & \langle \boxed{1}\ \mathrm{CP}\rangle \\ \textsc{comps} & \boxed{\mathrm{A}}\end{array}\right]\right]\right]\right\rangle \\ \textsc{output} \quad \left\langle \mathrm{Y}, \left[\textsc{syn} \left[\textsc{val} \left[\begin{array}{ll}\textsc{spr} & \langle \mathrm{NP}\ \left[\textsc{form}\ \textit{it}\right]\rangle \\ \textsc{comps} & \boxed{\mathrm{A}} \oplus \langle \boxed{1}\rangle\end{array}\right]\right]\right]\right\rangle \end{array}\right]$$

- Why is the type *pi-rule*?
- Why doesn't it say anything about the semantics?
- Why is the COMPS value $\boxed{\mathrm{A}}$ not $\langle\rangle$?

Extraposition with Verbs Whose comps Lists Are Nonempty

(16) *It worries me that war is imminent.*

(17) *It occurred to Pat that Chris knew the answer.*

(18) *It endeared you to Andy that you wore a funny hat.*

Idioms

Another Nonreferential Noun: *advantage*

$$\left\langle \text{advantage}, \begin{bmatrix} \textit{massn-lxm} & & \\ \text{SYN} & \begin{bmatrix} \text{HEAD} & \begin{bmatrix} \text{FORM} & \textit{advantage} \\ \text{AGR} & \textit{3sing} \end{bmatrix} \end{bmatrix} \\ \text{SEM} & \begin{bmatrix} \text{MODE} & \textit{none} \\ \text{INDEX} & \textit{none} \\ \text{RESTR} & \langle\rangle \end{bmatrix} \end{bmatrix} \right\rangle$$

The Verb That Selects *advantage*

$$\left\langle \text{take}, \left[\begin{array}{ll} \textit{ptv-lxm} & \\ \text{ARG-ST} & \left\langle \text{NP}_i, \left[\text{FORM } \textit{advantage}\right], \left[\begin{array}{ll}\text{FORM} & \textit{of}\\ \text{INDEX} & \textit{j}\end{array}\right] \right\rangle \\ \text{SEM} & \left[\begin{array}{ll} \text{INDEX} & \textit{s} \\ \text{RESTR} & \left\langle \left[\begin{array}{ll} \text{RELN} & \textbf{\textit{take_advantage}} \\ \text{SIT} & \textit{s} \\ \text{EXPLOITER} & \textit{i} \\ \text{EXPLOITED} & \textit{j} \end{array} \right] \right\rangle \end{array} \right] \end{array} \right] \right\rangle$$

Our Analyses of Idioms and Passives Interact...

- We generate

  (19) *Advantage was taken of the situation by many people.*
  (20) *Tabs are kept on foreign students.*

  with roughly the same meanings as

  (21) *Many people took advantage of the situation.*
  (22) *Someone keeps tabs on foreign students.*

- But not:

  (23) *Many people were taken advantage of.*

- Why not?

Overview

- Review: the copula ***be***
- Existentials (***there***, ***be***)
- Dummy ***it***
- Complementizers and CP subjects (***that***)
- Extraposition (***it***, LR)
- Idioms (***take_advantage***, ...)
- Problems

Problems

P1: *there* and Agreement

The analysis of existential ***there*** sentences presented so far says nothing about verb agreement.

A. Consult your intuitions (and/or those of your friends, if you wish) to determine what the facts are regarding number agreement of the verb in ***there*** sentences.

   Consider these sentences: which ones do you think are ok?

   (24) There is a cat on the mat.
   (25) There are a cat on the mat.
   (26) There is cats on the mat.
   (27) There are cats on the mat.

Give an informal statement of a generalization covering these facts, and illustrate it with both grammatical and ungrammatical examples. [*Note: Intuitions vary regarding this question, across both individuals and dialects. Hence there is more than one right answer to this question.*]

For me (Francis), only (24) and (27) are ok; for many speakers (26) is also ok.

B. How would you elaborate or modify our analysis of the ***there*** construction so as to capture the generalization you have discovered? Be as precise as you can.

P2: Passing Up the Index

A. Give the RESTR value that our grammar should assign to the sentence in (i). Be sure that the SIT value of the **smoke** predication is identified with the ANNOYANCE value of the **annoy** predication.

(i) *That Dana is smoking annoys Leslie.*

[*Hint: This sentence involves two of the phenomena analyzed in this chapter: predicative complements of* ***be*** *and CP subjects.*]

B. Draw a tree for (i). Use abbreviations for node labels, but show the index on each node.

C. Explain how the SIT value of the **smoke** predication gets identified with the ANNOYANCE value of the **annoy** predication. Be sure to make reference to lexical entries, phrase structure rules, and principles, as appropriate.

P3: An Annoying Problem

Assume that the lexical entry for the verb ***annoy*** is the following:

(28)
$$\left\langle \text{annoy} \;,\; \left[\begin{array}{l} \textit{stv-lxm} \\ \text{ARG-ST}\ \left\langle \left[\text{SEM}\ \left[\text{INDEX}\ \boxed{1}\right]\right], \text{NP}_i \right\rangle \\ \text{SEM}\ \left[\begin{array}{l} \text{INDEX}\ \ s \\ \text{RESTR}\ \left\langle \left[\begin{array}{ll} \text{RELN} & \textbf{\textit{annoy}} \\ \text{SIT} & s \\ \text{ANNOYED} & i \\ \text{ANNOYANCE} & \boxed{1} \end{array}\right] \right\rangle \end{array}\right] \end{array}\right] \right\rangle$$

A. What constraints are imposed on the lexical sequences that result from applying the **3rd-Singular Verb Lexical Rule** to this entry (including those that involve inheritance of constraints from the entry's supertypes)?

B. What constraints are imposed on lexical sequences that result from applying the **Extraposition Lexical Rule** to your answer to part (A)?

C. Draw a tree structure for the sentence in (29). You should show the value of all SEM features on all of the nodes, as well as the SPR and COMPS features for *annoys*.

(29) *It annoys Lee that Fido barks.*

D. The lexical entry for ***annoy*** allows NP subjects as well, as in (30). Why doesn't the grammar then also license (31)?

(30) *Sandy annoys me.*

(31) **It annoys me Sandy.*

Acknowledgments and References

- Course design and slides borrow heavily from Emily Bender's course: *Linguistics 566: Introduction to Syntax for Computational Linguistics*
  http://courses.washington.edu/ling566

- This lecture complements Sag et al. (2003, Chapter 11)

References

Ivan A. Sag, Tom Wasow, and Emily Bender. 2003. *Syntactic Theory: A Formal Introduction*. CSLI Publications, Stanford, second edition.