Non-referential NPs, Expletives, and Extrapolation
Overview

• Existentials
• Extraposition
• Idioms
Where We Are, and Where We’re Going

• Last time, we met the passive *be*.
• Passive *be* is just a special case -- that *be* generally introduces [PRED+] constituents (next slide).
• Today, we’ll start with 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*

\[
\begin{align*}
\langle be, \quad \text{ARG-ST} \quad \langle [1], \quad \text{SYN} \quad \langle \text{HEAD} \quad \langle \text{verb} \quad \text{FORM} \quad \langle \text{pass} \quad \text{VAL} \quad \langle \text{SPR} \quad \langle [1] \quad \text{COMPS} \quad \langle \rangle \rangle \quad \text{SEM} \quad \langle \text{INDEX} \quad s \rangle \quad \text{RESTR} \quad \langle \rangle \rangle \rangle \quad \text{SEM} \quad \langle \text{INDEX} \quad s \rangle \rangle
\end{align*}
\]
Copula (generalized)

\[ be-lxm \]

\[ \langle be, \langle ARG-ST \langle 1, SYN \langle HEAD \langle PRED + \rangle \rangle \rangle, VAL \langle SPR \langle COMPS \langle \rangle \rangle \rangle \rangle, SEM \langle INDEX \langle s \rangle \rangle \rangle, RESTR \langle \rangle \rangle \]
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*. 
Lexical Entry for the Existential *be*

\[
\langle \text{be}, \langle \begin{array}{c}
\text{ARG-ST} \\
\text{SEM}
\end{array} \rangle \rangle
\]

\[
\begin{array}{c}
\text{exist-be-lxm} \\
\text{NP} \\
\text{SEM}
\end{array}
\]

\[
\begin{array}{c}
\text{FORM there}, \langle 2 \rangle \\
\text{VAL} \\
\text{RESTR} \langle \rangle
\end{array}
\]

\[
\begin{array}{c}
\text{PRED} + \\
\text{COMPS} \langle \langle 2 \rangle \rangle \\
\text{INDEX s}
\end{array}
\]
Questions About the Existential \textit{be}

- What type of constituent is the third argument?
- Why is the third argument \([\text{PRED} +]\)?
- Why is the second argument tagged as identical to the SPR of the third argument?
- What is the contribution of this \textit{be} to the semantics of the sentences it occurs in?
- Can all \([\text{PRED} +]\) predicates appear as the third argument in existentials?
- How do we rule out \textit{*There was a greyhound a good runner}?
The Entry for Existential *there*

\[ \langle \text{there} , \begin{bmatrix} \text{pron-lxm} \\ \text{SYN} \\ \text{SEM} \end{bmatrix} \begin{bmatrix} \text{FORM} & \text{there} \\ \text{AGR} & \text{PER} \ 3rd \\ \text{MODE} & \text{none} \\ \text{INDEX} & \text{none} \\ \text{RESTR} & \langle \rangle \end{bmatrix} \rangle \]
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?

```plaintext
⟨pron-lxm

SYN [there , ]

HEAD [FORM there]

AGR [PER 3rd]

MODE none

INDEX none

RESTR ⟨⟩ ⟩
```
Other NPs that don’t seem to refer

• *It sucks that the Rockies lost the series.*

• *It is raining.*

• *Andy took advantage of the opportunity.*

• *Lou kicked the bucket.*
What we need to deal with examples like *It follows that you are wrong*

- A lexical entry for this dummy *it*
- An analysis of this use of *that*
- Entries for verbs that take clausal subjects (as in *That you are wrong follows*)
- A rule to account for the relationship between pairs like *That you are wrong follows* and *It follows that you are wrong*
The Entry for Dummy *it*

\[
\langle \text{it}, \rightline{\text{pron-lxm}} \rangle
\]

SYN

\[
\left[ \begin{array}{c}
\text{HEAD} \\
\text{AGR} \\
\text{FORM} \\
\end{array} \right]
\]

3\text{sing}

it

SEM

\[
\left[ \begin{array}{c}
\text{MODE} \\
\text{INDEX} \\
\text{RESTR} \\
\end{array} \right]
\]

none

none

\langle \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*?

\[
\begin{array}{c}
\text{pron-lxm} \\
\text{SYN} \\
\text{SEM} \\
\end{array}
\begin{array}{c}
\langle \text{it,} \rangle \\
\begin{array}{c}
\text{HEAD} \\
\text{MODE} \\
\text{INDEX} \\
\text{RESTR} \\
\end{array}
\begin{array}{c}
\text{FORM} \text{it} \\
\text{AGR} 3\text{sing} \\
\text{none} \\
\langle \rangle \\
\end{array}
\end{array}
\]
A New Type of Lexeme: Complementizers

\[ \text{comp-lxm :} \]

\[
\begin{align*}
\text{ARG-ST} & : \langle \left[ \text{INDEX} \ s \right] \rangle \\
\text{SEM} & : \left[ \text{INDEX} \ s \right] \\
\text{SYN} & : \text{HEAD} \left[ \text{comp} \right. \\
\text{VAL} & : \text{AGR} \left[ 3\text{sing} \right. \\
\end{align*}
\]
Questions About the Type $\textit{comp-lxm}$

- 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?
The Type $comp$

```
pos

[FORM, PRED]
```

```
agr-pos
```

```
AGR
```

```
verb
```

```
nominal
```

```
det
```

```
AUX
```

```
CASE
```

```
COUNT
```

```
noun
```

```
comp
```

```
FORM cform
```
The Lexical Entry for Complementizer *that*
...and with inherited information filled in

Question: Where did [FORM cform] come from?
Structure of a Complementizer Phrase

```
Structure of a Complementizer Phrase

[ \[ HEAD \ 2 \ 
   VAL \ 
   \] ]

C

[ \[ word \ 
   \] ]

[ \[ HEAD \ 
   \] ]

[ \[ VAL \ 
   \] ]

that

the Giants lost
```
Sample Verb with a CP Subject

\[
\langle \text{matter} , \rangle \\
\begin{array}{c}
\text{SEM} \\
\text{ARG-ST} \\
\text{siv-lxm}
\end{array}
\begin{array}{c}
\langle \text{SEM [INDEX [1]]} \rangle \\
\text{INDEX} \quad s \\
\text{RESTR} \\
\langle \text{RELN} \\
\text{SIT} \\
\text{MATTERING} \quad \text{matter} \quad s \quad \text{[1]} \rangle
\end{array}
\]

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:
    \textit{That Bush won matters} vs. *\textit{Bush won matters}*
  • Argument-marking PPs are semantically identical to their object NPs, but we get:
    \textit{The election mattered} vs. *\textit{Of the election mattered}*
  • So we need to add a syntactic constraint.

\[
\begin{align*}
\text{ARG-ST} & : \left[ \begin{array}{c}
\text{siv-lxm} \\
\text{SEM} \\
\text{INDEX} \\
\text{RESTR}
\end{array} \right] \\
\text{SEM} & : \left[ \begin{array}{c}
\text{SYN} \left[ \begin{array}{c}
\text{HEAD nominal} \\
\text{INDEX} \ []
\end{array} \right] \\
\text{RELN} \\
\text{MATTERING}
\end{array} \right] \\
\text{INDEX} & : \left[ \begin{array}{c}
\text{s} \\
\text{SIT} \\
\text{MATTERING} \\
\text{m} \\
\end{array} \right]
\end{align*}
\]

• S and PP subjects are generally impossible, so this constraint should probably be on *verb-lxm*. 
Why is the type *$\pi$-rule*?

Why doesn’t it say anything about the semantics?

Why is the COMPS value $[A]$, not $< >$?
Extraposition with Verbs whose COMPS Lists are Nonempty

• *It worries me* that war is imminent.

• *It occurred to Pat* that Chris knew the answer.

• *It endeared you to Andy* that you wore a funny hat.
Another Nonreferential Noun

\[ \langle \text{advantage} , \langle \text{massn-lxm} \rangle \rangle \]

\[
\text{SYN} \quad \text{HEAD} \quad \text{AGR} \\
\quad \text{FORM} \quad \text{advantage} \quad 3\text{sing} \\
\quad \text{MODE} \quad \text{none} \\
\text{SEM} \quad \text{INDEX} \quad \text{none} \\
\quad \text{RESTR} \quad \langle \rangle
\]
The Verb that Selects *advantage*

\[
\langle \text{take} , \langle \text{INDEX } s \rangle \rangle
\]

\[
\text{SEM} \quad \text{RESTR} \langle \langle \text{RELN} \text{ SIT } \text{ EXPLOITER } i \rangle, \langle \text{EXPLOITED } j \rangle \rangle
\]

\[
\langle \text{ARG-ST} \langle \text{NP}_i , \langle \text{FORM advantage} \rangle, \langle \text{FORM of } \text{INDEX } j \rangle \rangle \rangle
\]

\[
\langle \text{ptv-lxm} \rangle
\]
Our analyses of idioms and passives interact...

• We generate

  Advantage was taken of the situation by many people.
  Tabs are kept on foreign students.

• But not:

  Many people were taken advantage of.

• Why not?
Overview

• Existentials (*there, be*)
• Extraposition (*that, it, LR*)
• Idioms
Reading Questions

- We added the type *pi-rule* and it is allowed to "muck around with argument structure". That seems to break the nice clean cut distinction between i-rules and d-rules. Does it add the NP[form it] to the SPR or does it simply take it out of the [A] and put it on the SPR list and add the constraint [FORM it]? Also how do we tell the differences between *pi-rules* and *d-rules*?
Reading Questions

• Why is *advantage* shown as being 3rd person, while *tabs* is not?
### cn-lxm

<table>
<thead>
<tr>
<th>SYN</th>
<th>HEAD [noun]</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEM</td>
<td>AGR [PER 3rd]</td>
</tr>
<tr>
<td>ARG-ST</td>
<td>MODE / ref</td>
</tr>
<tr>
<td></td>
<td>INDEX i</td>
</tr>
<tr>
<td></td>
<td>FIRST DP_i</td>
</tr>
<tr>
<td></td>
<td>REST / ( )</td>
</tr>
</tbody>
</table>

### cntn-lxm

<table>
<thead>
<tr>
<th>ARG-ST</th>
<th>[COUNT +], ...</th>
</tr>
</thead>
</table>

### massn-lxm

<table>
<thead>
<tr>
<th>ARG-ST</th>
<th>[COUNT -], ...</th>
</tr>
</thead>
</table>
Reading Questions

• I'm not sure I quite understand which words are valued PRED + and which are valued PRED -. Specifically, why is be PRED +?

• How does the Extraposition Lexical Rule plural verbs from taking dummy it as a SPR?

• What are some of the other ways idioms are treated in grammars? As far as allowing the semantic meanings under the covers, etc.
Reading Questions

• What about *Well, there’s the cat*...?

• Does *exist* have the same lexical entry as existential *be*?

• I'm confused about the finite *S* that is the part after *that* as defined on page 340. What is the *S*? The definition says is a phrase headed by a finite verb so why isn't it just a VP?
Reading Questions

- It seems like we're getting a lot of new features that aren't always relevant. How does the person building the grammar trade off accuracy versus intelligibility and generalization? Is the goal always to find a way to ensure accuracy, and if so how about the trade-off between lexical entry complexity, more rules, etc.? Is this a case by case basis and somewhat intuitive, or is there an actual metric to quantify the entropy (or something like that) of the grammar?