How the Grammar Works
Overview

• What we’re trying to do
• The pieces of our grammar
• Two extended examples
• Reflection on what we’ve done, what we still have to do
What We’re Trying To Do

• Objectives
  • Develop a theory of knowledge of language
  • Represent linguistic information explicitly enough to distinguish well-formed from ill-formed expressions
  • Be parsimonious, capturing linguistically significant generalizations.

• Why Formalize?
  • To formulate testable predictions
  • To check for consistency
  • To make it possible to get a computer to do it for us
How We Construct Sentences

- The Components of Our Grammar
  - Grammar rules
  - Lexical entries
  - Principles
  - Type hierarchy (very preliminary, so far)
  - Initial symbol (S, for now)

- We combine constraints from these components.

**Q:** What says we have to combine them?

**A:** The definition of well-formed structure
An Example

A cat slept.

• Can we build this with our tools?
• Given the constraints our grammar puts on well-formed sentences, is this one?
Lexical Entry for $a$

\[
\begin{bmatrix}
\text{word} \\
\text{syn} \\
\langle a, \\ \text{val} \\
\text{sem} \\
\text{mode} \\
\text{index} \\
\text{rest} \\
\end{bmatrix}
\begin{bmatrix}
det \\
\text{agr} \\
\text{count} \\
\text{comps} \\
\text{spr} \\
\text{mod} \\
\text{none} \\
\text{j} \\
\end{bmatrix}
\begin{bmatrix}
3\text{sing} \\
+ \\
\langle \rangle \\
\langle \rangle \\
\langle \rangle \\
\langle \rangle \\
\langle \rangle \\
\end{bmatrix}
\]
Lexical Entry for *cat*

- Which feature paths are abbreviated?
- Is this a fully specified description?
- What features are unspecified?
- How many word structures can this entry license?
Effect of Principles: the SHAC
Description of Word Structures for *cat*

```
word
  | HEAD |
  |   noun   |
  | AGR [2] 3sing GEND neut |
SYN       |   D   |
            | SPR |
            |   COMPS () |
            | MOD ()   |
SEM       |   MODE ref |
            | INDEX k |
            | RESTR ⟨⟨ RELN cat ⟩ ⟩ |
            | ⟨⟨ INSTANCE k ⟩ ⟩ |
```

*cg t*
Description of Word Structures for $a$

\[
\text{word} \begin{bmatrix} \text{word} \\ \text{HEAD} & \text{det} \\ \text{AGR} & \text{3sing} \\ \text{COUNT} & + \\ \text{COMPS} & \langle \rangle \end{bmatrix} \begin{bmatrix} \text{VAL} \\ \text{SPR} & \langle \rangle \\ \text{MOD} & \langle \rangle \end{bmatrix} \begin{bmatrix} \text{SEM} \\ \text{MODE} & \text{none} \\ \text{INDEX} & j \\ \text{RESTR} & \langle \left[ \text{RELN} a \right] \rangle \\ \text{BV} & j \end{bmatrix} a
\]
Building a Phrase

[ ]

[ ] [ ]
Constraints Contributed by Daughter Subtrees

word

SYN

head
det
aggr 3sing

COUNT +

comps ⟨ ⟩

mod ⟨ ⟩

val

mode none

INDEX j

restr ⟨ reln a ⟩

SEM

SEM

SEM

noun

aggr 3sing

count +

gend neut

D

spr ⟨ ⟨

aggr 2

count +

INDEX k

⟩ ⟩

comps ⟨ ⟩

mod ⟨ ⟩

mode ref

INDEX k

restr ⟨ reln cat ⟩

INDEX cat

instance k

reln

BV j
Constraints Contributed by the Grammar Rule

\[
\text{phrase} \\
\text{SYN} \ [ \text{VAL} \ [ \text{SPR} \langle \rangle ]] \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]

\[
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\]
A Constraint Involving the SHAC

\[
\begin{align*}
\text{phrase} & : \text{SYN} [\text{VAL} [\text{SPR} \langle \rangle]] \\
\end{align*}
\]
Effects of the Valence Principle

[Diagram of linguistic structure with labels and symbols]

- Phrase structure with labels SYN and VAL
- Word structure with det, AGR, COUNT, COMPS, MOD, RELN, and INSTANCE
- Semantic structure with MODE, INDEX, and RESTRICTION

[Detailed labels and annotations for each node in the diagram]
Effects of the Head Feature Principle

```
phrase
SYN [HEAD 6]
VAL [SPR ⟨⟩ COMPS 3 MOD 4]

word
SYN [det [HEAD AGR 2 COUNT + COMPS ⟨⟩ SPR ⟨⟩ MOD ⟨⟩]]
SEM [MODE none INDEX k RESTR ⟨RELN a BV k⟩]

word
SYN [noun [HEAD AGR 2 3sing GEND neut]]
SEM [MODE ref INDEX k RESTR ⟨RELN cat INSTANCE k⟩]
```
Effects of the Semantic Inheritance Principle
Effects of the Semantic Compositionality Principle

phrase

[HEAD 6]

[SPR ⟨ ⟩]

[COMPS 3]

[MOD 4]

[MODE 8]

[INDEX k]

[RESTR A ⊕ B]

word

[det]

[AGR 2]

[COUNT +]

[COMPS ⟨ ⟩]

[SPR ⟨ ⟩]

[MOD ⟨ ⟩]

[MODE none]

[INDEX k]

[RESTR A] [RELN a BV k]

word

[noun 3sing]

[AGR 2]

[GEND neut]

[COMPS 3⟨ ⟩]

[SPR ⟨ ⟩]

[MOD 4⟨ ⟩]

[MODE 8ref]

[INDEX k]

[RESTR B] [RELN cat INSTANCE k]
Is the Mother Node Now Completely Specified?

```
<table>
<thead>
<tr>
<th>phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEAD 6</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SYN</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPR ⟨ ⟩</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMPS 3</td>
</tr>
<tr>
<td>MOD 4</td>
</tr>
</tbody>
</table>

| MODE 8 |
| INDEX k |
| RESTR A ⊕ B |

<table>
<thead>
<tr>
<th>word</th>
</tr>
</thead>
<tbody>
<tr>
<td>det</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SYN</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEAD 6</td>
</tr>
<tr>
<td>AGR 2</td>
</tr>
<tr>
<td>COUNT +</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>COMPS ⟨ ⟩</td>
</tr>
<tr>
<td>SPR ⟨ ⟩</td>
</tr>
<tr>
<td>MOD ⟨ ⟩</td>
</tr>
</tbody>
</table>

| MODE none |
| INDEX k |
| RESTR A ⟨ RELN a ⟩ |

<table>
<thead>
<tr>
<th>word</th>
</tr>
</thead>
<tbody>
<tr>
<td>noun</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>SYN</th>
</tr>
</thead>
<tbody>
<tr>
<td>HEAD 6</td>
</tr>
<tr>
<td>AGR 2</td>
</tr>
<tr>
<td>3sing</td>
</tr>
<tr>
<td>GEND neut</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>SPR ⟨ 7 ⟩</td>
</tr>
</tbody>
</table>

| COMPS 3 ⟨ ⟩ |
| MOD 4 ⟨ ⟩ |

| MODE 8 ref |
| INDEX k |
| RESTR B ⟨ RELN cat INSTANCE k ⟩ |
```
Lexical Entry for *slept*

\[
\langle \text{slept}, \text{word} \rangle
\]

\[
\begin{align*}
\text{SYN} & : [\text{HEAD} \quad \text{verb}] \\
\text{VAL} & : [\text{SPR} \quad \langle \text{AGR} \quad \mathbb{9} \quad \text{CASE} \quad \text{nom} \rangle] \\
\text{COMP} & : [\langle \rangle] \\
\text{MOD} & : [\langle \rangle] \\
\text{INDEX} & : s_1 \\
\text{MODE} & : \text{prop}
\end{align*}
\]

\[
\begin{align*}
\text{SEM} & : [\text{RESTR} \quad \langle \text{RELN} \quad \text{sleep} \rangle] \\
\text{SIT} & : s_1, \ldots \\
\text{SLEEPER} & : m
\end{align*}
\]
Is this description fully specified?
Does the top node satisfy the initial symbol?
RESTR of the S node

\[ \langle \begin{bmatrix} \text{RELN} & a \\ \text{BV} & k \end{bmatrix}, \begin{bmatrix} \text{RELN} & \text{cat} \\ \text{INST} & k \end{bmatrix}, \begin{bmatrix} \text{RELN} & \text{sleep} \\ \text{SIT} & s_1 \\ \text{SLEEPER} & k \end{bmatrix}, \cdots \rangle \]
Another Example

The NOM photos of the suspect disappeared yesterday.
Head Features from Lexical Entries

$S$

$NP$

$[HEADdet]$ $the$

$NOM$

$[HEADnoun]$ $photos$

$PP$

$[HEADprep]$ $of$

$NP$

$[HEADdet]$ $the$

$[HEADnoun]$ $suspect$

$VP$

$[HEADverb]$ $disappeared$

$[HEADadverb]$ $yesterday$
Head Features from Lexical Entries, plus HFP

[HEADdet] the photos

[HEADdet] of

[HEADnoun] photos

[HEADadverb] disappeared

[HEADadverb] yesterday

[HEADdet] the suspect

[HEADdet] the

[HEADnoun] suspect

[HEADadverb] disappeared

[HEADadverb] yesterday
Valence Features:
Lexicon, Rules, and the Valence Principle

Key
- Lexicon
- Val. Prin.
- Rules

the photos of disappeared yesterday of the suspect
The NOM
[SPR ⟨ 2 ⟩]
N
[COMPS ⟨ 3 ⟩]
photos
3 P P
P
[COMPS ⟨ 4 ⟩]
of
4 N P
5 D
the
N
[SPR ⟨ 5 ⟩]
suspect
VP
[SPR ⟨ 1 ⟩]
6 V
disappeared
ADV
[MOD ⟨ 6 ⟩]
yesterday
Two Semantic Features: the Lexicon & SIP

```
the photos of the the suspect
```

disappeared yesterday
RESTR Values and the SCP

A ⊕ B ⊕ C ⊕ D ⊕ E ⊕ F ⊕ G

A ⊕ B ⊕ C ⊕ D ⊕ E

B ⊕ C ⊕ D ⊕ E

C ⊕ D ⊕ E

D ⊕ E

E ⊕ G

A

the

B

photos

of

C

D ⊕ E

E

suspect

the

disappeared

yesterday
An Ungrammatical Example

What's wrong with this sentence?
An Ungrammatical Example

What’s wrong with this sentence?
An Ungrammatical Example

The Valence Principle

* S

NP [CASE acc] them

V

[SPR ⟨[1]⟩]

sent

NP

[SPR ⟨[1]⟩ NP[nom]]

us

NP

D

a

N

letter
An Ungrammatical Example

HeadSpecifier Rule

*\(S\)

\[\begin{align*}
\text{NP} & \quad \text{[CASE acc]} \\
\text{them} & \\
\text{VP} & \quad \text{[SPR} \langle \{\text{NP}\} \rangle \text{]} \\
\text{V} & \quad \text{sent} \\
\text{NP} & \quad \text{us} \\
\text{NP} & \quad \text{D} \\
\text{a} & \quad \text{letter}
\end{align*}\]
Exercise in Critical Thinking

• Our grammar has come a long way since Ch 2, as we've added ways of representing different kinds of information:
  • generalizations across categories
  • semantics
  • particular linguistic phenomena: valence, agreement, modification

• What else might we add? What facts about language are as yet unrepresented in our model?
Overview

• What we’re trying to do
• The pieces of our grammar
• Two extended examples
• Reflection on what we’ve done, what we still have to do
• Next time: Review (or Ch 7?)