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?
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 \textit{a}

- Is this a fully specified description?
- What features are unspecified?
- How many word structures can this entry license?
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

\[
\langle \text{cat} ,
\begin{array}{c}
\text{SYN} \\
\text{VAL} \\
\text{SEM}
\end{array}
\rangle
\]

\[
\begin{array}{c}
\langle \text{cat} ,
\begin{array}{c}
\text{HEAD} \\
\text{AGR} [', \text{GEN} \text{d neut}] \\
\text{D}
\end{array}
\rangle
\end{array}
\]

\[
\begin{array}{c}
\langle \text{AGR} [\text{COUNT} + \text{INDEX} \text{ref} \text{INDEX} \text{k}]
\rangle
\end{array}
\]

\[
\begin{array}{c}
\langle \text{RELN} \text{cat} 
\rangle
\end{array}
\]

© 2003 CSLI Publications
Description of Word Structures for cat
Description of Word Structures for \textit{a}

\[
\begin{array}{c}
\text{word} \\
\text{SYN} \\
\text{VAL} \\
\text{SEM} \\
\end{array}
\begin{array}{c}
\text{det} \\
\text{AGR} \ 3\text{sing} \\
\text{COUNT} + \\
\text{MODE} \text{ none} \\
\text{INDEX} \ j \\
\text{RESTR} \left( \left[ \text{RELN} \ a \right] \right) \\
\end{array}
\begin{array}{c}
\text{HEAD} \\
\text{COMPS} \left( \right) \\
\text{SPR} \left( \right) \\
\text{MOD} \left( \right) \\
\end{array}
\begin{array}{c}
\text{a} \\
\end{array}
\]
Building a Phrase
Constraints Contributed by Daughter Subtrees
Constraints Contributed by the Grammar Rule
A Constraint Involving the SHAC

```
[phrase
  SYN [ VAL [ SPR ⟨⟩]]]
```

```
[7
  word
    [det 3sing
det [ 3sing
        AGR
        GEND neut]
        COUNT +
        COMPS ⟨⟩
        SPR ⟨⟩
        MOD ⟨⟩]
    [SEM
      MODE none
      INDEX k
      MODE
      RESTR [RELN a BV k]]]
]
```

```
[7D
  word
    [noun
      AGR [ 3sing
          GEND neut]
          COUNT +
          INDEX k]
      COMPS ⟨⟩
      MOD ⟨⟩]
    [SEM
      MODE ref
      INDEX k
      MODE
      RESTR [RELN cat INSTANCE k]]]
]
```
Effects of the Valence Principle
Effects of the Head Feature Principle
Effects of the Semantic Inheritance Principle
Effects of the Semantic Compositionality Principle
Is the Mother Node Now Completely Specified?

**Phrase**

- **SYN**
  - **HEAD**: 6
  - **VAL**
    - **COMPS**: [3]
    - **MOD**: [4]
- **SEM**
  - **MODE**: 8
  - **INDEX**: k
  - **RESTR**: A ⊕ B

**Word**

1. **SYN**
   - **HEAD**: det
     - **AGR**: 2
     - **COUNT**: +
   - **VAL**
     - **COMPS**: [⟩⟩]
     - **SPR**: ⟨⟩
     - **MOD**: ⟨⟩
   - **SEM**
     - **MODE**: none
     - **INDEX**: k
     - **RESTR**: A

2. **SYN**
   - **HEAD**: noun
     - **AGR**: 2
     - **3sing**: GEND neut
   - **VAL**
     - **COMPS**: [⟩⟩]
     - **SPR**: ⟨⟩
     - **MOD**: ⟨⟩
   - **SEM**
     - **MODE**: ref
     - **INDEX**: k
     - **RESTR**: B

---

© 2003 CSLI Publications
Lexical Entry for *slept*

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

**SYN**
- **HEAD**: *verb*
- **NP**: \( m \) \( \text{nom} \)
- **SPR**: \( \text{AGR} \ 9 \) \( \text{CASE} \)
- **COMPS**: \( \langle \rangle \)
- **MOD**: \( \langle \rangle \)
- **INDEX**: \( s_1 \)
- **MODE**: prop

**SEM**
- **RESTR**: \( \langle \text{SLEEPER} \ m, \ldots \rangle \)
- **RELN**: sleep
Another Head-Specifier Phrase

Key

- HSR
- SHAC
- Val Prin
- HFP
- SIP
- SCP

© 2003 CSLI Publications
Is this description fully specified?
Does the top node satisfy the initial symbol?
RESTR of the S node

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

S

NP

D

the

N

photos

PP

of

NP

D

the

N

suspect

VP

V

disappeared

ADV

yesterday
Head Features from Lexical Entries

\[ S \]
\[ NP \]
\[ HEADdet \] the
\[ HEADnoun \] photos
\[ HEADprep \] of
\[ HEADdet \] the
\[ HEADnoun \] suspect
\[ VP \]
\[ HEADverb \] disappeared
\[ HEADadverb \] yesterday
\[ PP \]
the photos of the suspect disappeared yesterday
Valence Features: Lexicon, Rules, and the Valence Principle

The diagram illustrates the valence features of a sentence. Each node in the tree represents a word, with its corresponding valence features. The keys for the nodes are:

- **Lexicon**
- **Val.** (Valence)
- **Rules**

The sentence analyzed is "the photos of the suspect disappeared yesterday." The valence features for each word are shown in the diagram, indicating the syntactic role and the corresponding valence features.

The sentence structure is as follows:
- **the** (Lexicon)
- **photos** (Lexicon)
- **of** (Val.)
- **the** (Lexicon)
- **suspect** (Lexicon)
- **disappeared** (Lexicon)
- **yesterday** (Lexicon)
Required Identities: Grammar Rules

$S$

$NP$

$D$

the

$N$

photos

$PP$

of

$NP$

$D$

the

$N$

suspect

$VP$

$V$

disappeared

$ADV$

yesterday
Two Semantic Features: the Lexicon & SIP

```
the photos of the suspect disappeared yesterday
```

```
MODE prop INDEX s3

MODE ref INDEX j

MODE none INDEX j

MODE ref INDEX j

MODE ref INDEX k

MODE prop INDEX s3

MODE prop INDEX s3

MODE prop INDEX s4

MODE none INDEX s4

the

MODE ref INDEX j

photos

MODE ref INDEX k

of

MODE none INDEX k

MODE ref INDEX k

the

MODE ref INDEX k

suspect
```
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

RELN the
BV j

the

B

RELN photo
INST j
CONTENT k

photos

C()

of

D

RELN the
BV k

the

E

RELN suspect
INST k

suspect

disappeared

yesterday
An Ungrammatical Example

What’s wrong with this sentence?
An Ungrammatical Example

What’s wrong with this sentence?

So what?
An Ungrammatical Example

The Valence Principle

*S

NP

[CASE acc]

them

VP

[SPR ⟨⟩]

V

sent

NP

[SPR ⟨⟩]

NP

us

D

a

N

letter
An Ungrammatical Example

Head Specifier Rule

*S

[CASE acc]

them

[SPR ⟨1⟩ ]

V

sent

[SPR ⟨1⟩ NP[nom]]

NP

us

NP

D

N

a

letter

contradiction
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