Syntax: combining words to build sentences
Ling 106, Maribel Romero
Oct 17, 2001

How do speakers of a language put together a finite number of discrete elements (e.g., words) to generate infinite number of sentences?

1 Memorization

1.1 Lookup table
Each sentence would be associated with a message via a lookup table of some sort. Speakers would simply memorize some (mysterious) relationship between sentences and meanings.

<table>
<thead>
<tr>
<th>Sentence_1</th>
<th>Message_1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sentence_2</td>
<td>Message_2</td>
</tr>
<tr>
<td>Sentence_3</td>
<td>Message_3</td>
</tr>
<tr>
<td>Sentence_n</td>
<td>Message_n</td>
</tr>
</tbody>
</table>

1.2 Problems

- Novel sentences, infinite number of sentences.
- Grammaticality judgments.

   A genuinely novel sentence would have the same status as an ungrammatical string.

   (1)  
   a. *Tiger the ate basketball the. (ungrammatical)
   b. The tiger ate the basketball. (novel)

2 Words and Modes of Combination

- The smallest meaningful unit is the word. Speakers learn word meanings plus some method of combining those words to form sentences.
- For example, recall the syntactic rules for Propositional Logic:

  - If $\phi$ is a formula in PL, then $\neg \phi$ is a formula in PL too.
    “It is not the case that $\phi$.”
  - If $\phi$ and $\psi$ are formulae in PL, then $(\phi \land \psi)$ is a formulae in PL.
    “$\phi$ and $\psi$.”
  - If $\phi$ and $\psi$ are formulae in PL, then $(\phi \lor \psi)$ is a formulae in PL.
    “$\phi$ or $\psi$.”
– If $\phi$ and $\psi$ are formulae in PL, then $(\phi \rightarrow \psi)$ is a formulae in PL.
  “if $\phi$ then $\psi$.”
– If $\phi$ and $\psi$ are formulae in PL, then $(\phi \leftrightarrow \psi)$ is a formulae in PL.
  “$\phi$ if and only if $\psi$.”
– Nothing else is a formula in PL.

• What is the simplest device that could compute the method of combining words necessary for natural language?

• We can think of a sentence as a chain of words and a speaker as a device which consists of a finite number of mental predispositions, each associated with a rule that allows the speaker to produce a word and move to a new mental predisposition. That is, a speaker’s grammar may be equivalent to a finite state automaton (FSA).

3 Example of finite state automaton: automatic doors

Finite state automaton (FSA) is the simplest model of computation. Many useful devices can be modeled using FSA.

• Behavior of automatic doors:
  The door can be in one of two states: open or closed.
  If it is closed and a person is standing on the pad in front of the doorway, the door opens.
  If it is closed and a person is standing on the pad to the rear of the doorway, it remains closed.
  If it is closed and a person is standing on neither pad, it remains closed.
  If it is closed and people are standing on both pads, it remains closed.
  If it is open and a person is standing on the pad in front of the doorway, it remains open.
  If it is open and a person is standing on the pad to the rear of the doorway, it remains open.
  If it is open and if a person is standing on neither pad, it becomes closed.
  If it is open and if people are standing on both pads, it remains open.

• Behavior of automatic doors in tabular form.

<table>
<thead>
<tr>
<th></th>
<th>NEITHER</th>
<th>FRONT</th>
<th>REAR</th>
<th>BOTH</th>
</tr>
</thead>
<tbody>
<tr>
<td>CLOSED</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OPEN</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
• Behavior of automatic doors in graph form.

States: OPEN, CLOSED.
Input signals: FRONT, REAR, NEITHER, BOTH.

4 A grammar as a finite state automaton

• We could view a grammar for natural language as a finite state automaton. E.g., the following FSA can generate several English sentences.

• Which of the following sentences are generated by this FSA?

(2) a. A happy boy eats sad ice cream.
b. The boy eats ice cream.
c. A boy ate dogs.
d. A happy boy ate hot dogs.
e. One ate candy.
f. One happy girl eats candy.
g. One happy happy girl eats hot dogs.
h. One happy girl eats hot hot dogs.
i. A sad happy dog ate ice cream.
5 A potential problem

- FSA theory fails to capture the fact that some words pattern alike.

Words fall into classes defined by intersubstitutivity. Two words belong to the same grammatical category if one can be substituted for the other in a sentence in a way that preserves grammaticality.

(3) a. The happy boy XXXX the girl.
b. The happy XXXX kisses the girl.
c. The happy boy kisses XXXX girl.
d. The XXXX boy kisses the girl.

<table>
<thead>
<tr>
<th>Grammatical Category</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proper Nouns</td>
<td>John, Bill, Mary</td>
</tr>
<tr>
<td>Pronouns</td>
<td>he, she, it, they</td>
</tr>
<tr>
<td>Nouns</td>
<td>dog, cat, tiger, ball</td>
</tr>
<tr>
<td>Intransitive Verbs</td>
<td>walk, sleep, snore</td>
</tr>
<tr>
<td>Transitive Verbs</td>
<td>see, kiss, hug</td>
</tr>
<tr>
<td>Ditransitive Verbs</td>
<td>give, send, donate</td>
</tr>
<tr>
<td>Propositional Verbs</td>
<td>know, claim, believe</td>
</tr>
<tr>
<td>Determiners</td>
<td>the, some, every, a</td>
</tr>
<tr>
<td>Prepositions</td>
<td>with, to, in, on</td>
</tr>
<tr>
<td>Adjectives</td>
<td>tall, short, bogus</td>
</tr>
<tr>
<td>Adverbs</td>
<td>quickly, calmly, always</td>
</tr>
<tr>
<td>Complementizers</td>
<td>that, if, whether</td>
</tr>
<tr>
<td>Conjunctions</td>
<td>and, or, but</td>
</tr>
</tbody>
</table>

When people learn language, they do not learn which word follows which other word. Rather, they learn which grammatical category follows which other category.

- Finite state automata have a corresponding rule formalism: \( A \rightarrow aB \)

1. \( A \) is a single symbol (corresponding to a state) called a 'non-terminal symbol'.
2. \( a \) corresponds to a lexical item.
3. $B$ is a single non-terminal symbol.

- $S \rightarrow detB$
- $S \rightarrow detC$
- $B \rightarrow adjB$
- $B \rightarrow adjC$
- $C \rightarrow nounD$
- $D \rightarrow transitiveVerbE$
- $E \rightarrow noun$