1 What is Language and Do Animals (Besides Humans) Have It?

A proper conception of language is central to Cognitive Science:

By these two methods we can also recognize the difference between man and animals. For it is a very remarkable thing that there are no men, not even the insane, so dull and stupid that they cannot put words together in a manner to convey their thoughts. On the contrary, there is no other animal, however perfect and fortunately situated it may be, that can do the same. And this is not because they lack the organs, for we see that magpies and parrots can pronounce words as well as we can, and nevertheless cannot speak as we do, that is, in showing that they think what they are saying.

... Note also that we should not confuse speech, and all those signs which in the practice of human beings convey thoughts, with the natural sounds and movements that indicate passions, and can be imitated by machines as well as animals.

Discourse on Method (1644), page 572

A **linguistic theory** is a definition of the concept ‘language’.

- on a per-grammar basis, this is called **particular grammar**
- the things that is shared by all languages is called **universal grammar**

Natural languages contrast with artificial languages

- Morse code
- C++
- Klingon, Esperanto, Hobbit-talk...

The C++ compiler sure knows what is and what isn’t in its language!

1.1 How Something Could *Not* Be In a Language

If you know English, you know

1. English words (e.g. as listed in the OED)
2. English sentences (i.e. sequences of selections from the OED that are English sentences)

Q. Which of the 120 permutations of the words “CNN will conduct three debates” are also English sentences?

- Negative results from this kind of ‘experiment’ are marked with asterisks
- leads to concept of language as infinite subset of sequences of dictionary words
1.2 Parts of speech

Word substitutability – part of speech a predictive claim about new sentences

<table>
<thead>
<tr>
<th>Name</th>
<th>Aux</th>
<th>V</th>
<th>D</th>
<th>N_{+}-plural</th>
</tr>
</thead>
<tbody>
<tr>
<td>CNN</td>
<td>will</td>
<td>conduct</td>
<td>three</td>
<td>debates</td>
</tr>
<tr>
<td>CNN</td>
<td>will</td>
<td>conduct</td>
<td>three</td>
<td>concerts</td>
</tr>
<tr>
<td>CNN</td>
<td>will</td>
<td>conduct</td>
<td>three</td>
<td>inquiries</td>
</tr>
<tr>
<td>CNN</td>
<td>will</td>
<td>conduct</td>
<td>three</td>
<td>hearings</td>
</tr>
<tr>
<td>CNN</td>
<td>will</td>
<td>conduct</td>
<td>three</td>
<td>operations</td>
</tr>
</tbody>
</table>

any plural noun can occupy that sentence-final position.

With nine substitution classes \(\{V, A, Adv, N, Name, D, P, Prn, Aux\}\) the number of possible category sequences goes up sharply as sentences get longer:

\[
\begin{align*}
9 \\
81 \\
729 \\
6561 \\
59049 \\
531441 \\
4782969 \\
43046721 \\
387420489 \\
3486784401 \\
31381059609 \\
282429536481 \\
2541856828329 \\
22876792454961 \\
205891132094649 \\
1853020188851841 \\
1667718169966569 \\
15009463529699121 \\
1350851717672992089 \\
1215766545905692880 \\
\ldots
\end{align*}
\]

Most of these don’t correspond to any possible sentences:

1. Name — Adv — A — Aux — Aux
2. P — P — D — Aux — Prn
3. Name — Prn — A — N — P
4. A — Aux — Adv — Prn — N
5. P — Name — P — A — V
6. Aux — Adv — Prn — A — Aux
7. P — Name — P — Adv — Name
8. D — Name — A — P — Adv
10. P — Aux — Aux — Prn — V

Which of these should have stars in front of them?
1.3 Constituency

Having part-of-speech categories makes our theory shorter (it doesn’t involve the dictionary) but is essentially the same theory as “All sequences of English words from the dictionary are sentences.”

<table>
<thead>
<tr>
<th>Phrase</th>
<th>Parse Tree</th>
</tr>
</thead>
<tbody>
<tr>
<td>Carol should cast a vote</td>
<td>Name</td>
</tr>
<tr>
<td>She should cast a vote</td>
<td>Prn</td>
</tr>
<tr>
<td>The woman should cast a vote</td>
<td>determiner phrase (DP)</td>
</tr>
<tr>
<td>The African-American woman should cast a vote</td>
<td>DP</td>
</tr>
</tbody>
</table>

**phrase substitutability**

The one person who has a prayer of changing things in this country should cast a vote

DP is just the tip of the iceberg

| VP | cast a vote, rise to the occasion |
| AP | despondent about her boyfriend, eager for victory |
| NP | candidate, candidate for president |
| PP | about her boyfriend, for victory |
| CP | that she’s going to St. Tropez, if W. will win |

Phrases XP headed by word of category X. e.g. NPs don’t have determiners at the beginning, but rather nouns.

The number of conceivable binary trees for sentences of length \( n \) grows at an alarming rate: 0, 1, 2, 5, 16, 48, 144, 432, 1296... which **constituency** claims are correct?

(For more info see [http://www.ling.upenn.edu/~beatrice/syntax-textbook](http://www.ling.upenn.edu/~beatrice/syntax-textbook))

**noun substitution** pronouns can substitute for noun phrases

(1)

| (a) The little boy fed the cat | → He fed her |
| (b) Black cats detest green peas. | → They detest them |

**location substitution** adverbs like *here* or *there* can substitute for prepositional phrases

(2)

| (a) Put it on the table. | → Put it there. |
| (b) Put it over on the table | → Put it there. |

**so-substitution** *so* can substitute for adjective phrase

(3)

| (a) I am very happy and Linda is so, too |
| (b) I am very fond of my nephew * and Linda is so of her niece |

**question formation** questioned fragments must be of matching syntactic type

| (a) Noun phrase | What do you like? |
| (b) Prepositional phrase | How did the cat stroll across the porch? |
| (c) | Where did Ali Baba go? |
| (d) Adjective phrase | How did Ali Baba return? |
| (e) Adverb phrase | How did they do? |

The cats. Cats with long fluffy tails.
With a confident air
On a long journey. To New York.
Wiser than before. Fairly jet-lagged.
Not badly. Surprisingly well. Much better than they had expected.
**VP deletion test** if the deleted VP is — in the ...will too construction, the antecedent must be a full VP

Henry has [vp passed the wine to Falstaff], and Bardolph will — too.  
* Henry has passed the wine to Falstaff, and Bardolph has — Pistol.

(4)

Petruccio will [vp find a wife] and Hortensio will — too.  
* Petruccio will find a wife, and Bianca will — a husband

**coordination test** usually, only full constituents can be coordinated with constituents of the same type using and, or.

The scribe wrote a book and the king wrote to his wife  
* The scribe wrote [dp a book] and [pp to his wife]

### 1.3.1 Which constituency is correct?

(5) A.

```
S
 /\   
|  \  
NP  V_{transitive}  NP
   /    
  D N  D N
 the secretary  the letter
```

B.

```
S
 /\   
|  \  
NP  VP
   /    
  D N  V_{transitive}  NP
       /    
      D N  D N
 the secretary  drafted  the letter
```

### 1.4 Relationship to What You Learned in School

(6) Some rules prescribe what you should and should not do:

a. Don’t start a sentence with a conjunction.

b. Don’t use contractions.

c. Don’t use sentence fragments.

d. Don’t end a sentence with a linking verb.

e. Don’t use dangling participles.

f. Don’t end a sentence with a preposition.

g. Don’t use an object pronoun for a subject pronoun in a conjoined noun phrase.

h. Don’t use a plural pronoun to refer back to a singular noun like everyone, no-one, someone, and

i. Don’t split infinitives.

j. Use whom, not who, as the object of a verb or preposition.
(7) These are bad prescriptive grammar, but still English:
   a. Over there is the guy who I went to the party with. violates (f), (j)
   b. Bill and me went to the store. violates (g)

   a. Over there is guy the who I went to party the with.
   b. Over there is the who I went to the party with guy.
   c. Bill and me the store to went.

(8) These are just not English:
   a. Over there is guy the who I went to party the with.
   b. Over there is the who I went to the party with guy.
   c. Articles precede the nouns they belong with.

(9) Example descriptive rules:
   a. Relative clauses follow the noun that they modify.
   b. Prepositions precede their objects.

The descriptions in (9) have the status of laws, which in linguistics are called ‘generalizations.’ It would be nice to have a theory from which these generalizations follow and explains why they are true while at the same time defining the infinite set of English utterances.

1.4.1 Productivity – longest English sentence?
   a. I am happy to teach you today.
   b. I am very happy to teach you today.
   c. I am very very happy to teach you today.
   d. I am very very very happy to teach you today.

(10)...

(11) Mary thinks that Mary thinks that Mary thinks that Mary thinks that ...
    class was great today.

1.5 Finite State Grammar

The idea is that after “I am” you are in one of a finite number of states differentiating alternative elaborations of the sentence being defined (as a member of English).

Similarly with after “Mary thinks”

performance limitations this ‘generative’ process cannot actually be way people produce or comprehend language

- our lungs prevent the utterance of sentences more than 1 million words long
- drunkenness messes up out ability to make nouns and verbs agree

1.5.1 Example: Swahili

(* the Swahili−1 automaton from Gazdar and Mellish’s Natural Language Processing *)

let swahili1 =
  { states = stateset_of_list (range 1 5); (* there are five states total *)
    transitions = moveset_of_list [  
    Move (1,"ni",2); (* ni, u, a and tu are the subject morphemes *)
    Move (1,"u",2);
    Move (1,"a",2);
    Move (1,"tu",2);
    Move (1,"wa",2);
    Move (2,"ta",3); (* ta, na me, li are the tense morphemes *)
  }
Move (2, "na", 3);
Move (2, "me", 3);
Move (2, "li", 3);

Move (3, "ni", 4); (* ni, ku, m, tu and wa are the object morphemes *)
Move (3, "ku", 4);
Move (3, "m", 4);
Move (3, "tu", 4);
Move (3, "wa", 4);

Move (4, "penda", 5); (* here are some swahili verb stems *)
Move (4, "piga", 5);
Move (4, "sumbua", 5);
Move (4, "lipa", 5));
start = (IntSet.singleton 1);
final = (IntSet.singleton 5)}

Figure 1: A very small fragment of Swahili morphology

Any grammar having only rules of the form \( A \rightarrow bC \) where \( A, B \) are nonterminals and \( b \) is a terminal has a corresponding finite state machine. Given a string, if a path can be found through the machine, the string is generated by the grammar and vice versa.

There are some languages that cannot be recognized by finite state machines.

<table>
<thead>
<tr>
<th>formal language</th>
<th>(non-)examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>sequence of ( a )'s followed by an equal number of ( b )'s</td>
<td>( \in ) ( ab, aabb, aaabbb, ... )</td>
</tr>
<tr>
<td></td>
<td>( \notin ) ( aabbb, aaaaaaaaaaaaaabbb )</td>
</tr>
</tbody>
</table>

Call the number of \( a \)'s in the sentence being analyzed \( n \). A finite state machine would need to ‘remember’ this number \( n \) while waiting for the end of the \( b \)'s. But, by definition, a finite state machine will only have enough states to remember some fixed number of \( a \)'s. Hence there exists neither a finite-state grammar nor a finite state machine for the language \( a^n b^n \).
1.6 English is just like $a^n b^n$

Consider

(12)  
   a. The cat died.
   b. The cat the dog chased died.
   c. The cat the dog the rat bit chased died.
   d. The cat the dog the rat the elephant admired bit chased died

If $A = \{\text{the cat, the dog, the rat, the elephant, the kangaroo, \ldots}\}$ and $B = \{\text{chased, bit, admired, ate, befriended, \ldots}\}$ it’s clear that (12) have the structure

(13) $a^n b^{n-1}$ died

Chomsky and Miller (1963) argued that the obligatory paired dependencies presented by either...or, if...then or the agreement between verbs and subjects can nest inside one another to an arbitrary depth.

(14) Anyone who feels that if so many more students whom we haven’t actually admitted are sitting in on the course than we have that the room had to be changed, then probably auditors will have to be excluded, is likely to agree that the curriculum needs revision.

This is analogous to the reversal language $\{xx^R | x \in \{a, b\}^*\}$, which also cannot be generated by a finite automaton.

Grammar (15) does generate the language $a^n b^n$. It is a context-free grammar, and as such is capable of deriving any number of center-embeddings.

(15)  
   $S \rightarrow a b$  
   $S \rightarrow a S b$

There is a corresponding machine that accepts the context-free languages: the push down automaton.

1.7 The Chomsky Hierarchy

Let $A, B$ be nonterminals, $b$ a terminal, $\alpha$ a nonempty sequence of either kind of symbol, and $\gamma, \delta$ possibly empty sequences of either kind of symbol. The Chomsky Hierarchy is a classification of languages in a subset relationship. Each language has corresponding class of machine that recognizes it.

<table>
<thead>
<tr>
<th>language</th>
<th>rule type</th>
<th>machine</th>
</tr>
</thead>
<tbody>
<tr>
<td>finite-state</td>
<td>$A \rightarrow bC$</td>
<td>finite state machine</td>
</tr>
<tr>
<td>context-free</td>
<td>$A \rightarrow \alpha$</td>
<td>push down automaton</td>
</tr>
<tr>
<td>context-sensitive</td>
<td>$\gamma A \delta \rightarrow \gamma \alpha \delta$</td>
<td>linear bounded automaton</td>
</tr>
<tr>
<td>unrestricted</td>
<td>no restriction</td>
<td>Turing machine</td>
</tr>
</tbody>
</table>

The cardinality of the set of languages definable by a grammar formalism is called its generative capacity. Natural languages are believed to reside somewhere between context-free and context-sensitive.
2 Readings

Could it be that the difference between finite-state and phrase-structure captures the distinction between communicative animals and language-using people?

The Faculty of Language: What Is It, Who Has It and How Did it Evolve
by Marc D. Hauser, Noam Chomsky and W. Tecumseh Fitch
Science 298 1569-1579 22 November 2002

The faculty of language: what’s special about it?
by Steven Pinker and Ray Jackendoff
Cognition 95(2) March 2005 201–236

The Evolution of the Language Faculty: clarifications and implications.
W. Tecumseh Fitch, Marc D. Hauser and Noam Chomsky.

The nature of the language faculty and its implications for the evolution of language
(Reply to Fitch, Hauser and Chomsky)
by Ray Jackendoff and Steven Pinker
Cognition 97(2) March 2005 211-225