Lecture 2: Grammar

Cognition, Language & Communication 2013

Jelle Zuidema
ILLC
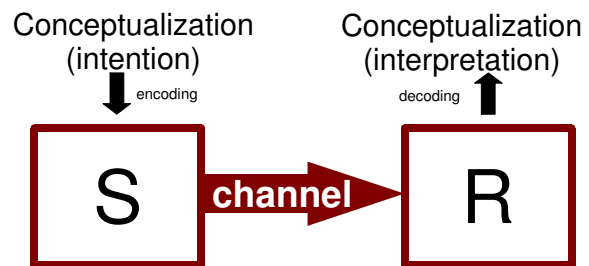zuidema@uva.nl

Claude Shannon: an engineer's perspective

- 1916-2001
- MSc 1937: Boolean algebra in computers
- PhD 1940: Population genetics
- 1948 Information Theory
- Mechanical mouse, Rocket powered flying discs, "Ultimate Machine"

Recap

- Duality of patterning and recursive hierarchical phrase-structure seem to be unique to human language
- Many of the other claimed 'unique design features' are shared, to some degree, with other animals, including arbitrariness, displacement, discreteness, stimulus freedom, vocal learning, cultural transmission and (a rudimentary form of) compositionality
- Combination of design features is certainly unique
- Difficult to get quantitative, precise statements about how different language is

Models of Communication

Conceptualization
(intention)

↓ encoding

Conceptualization
(interpretation)

↑ decoding

S →channel→ R

Plan for today

- How can we make more precise what's special about human language and its ability convey complex messages?
    – Models of communication
    – Roots of generative grammar

## A biologist's perspective

Millikan (ref Noble, 1998)

10
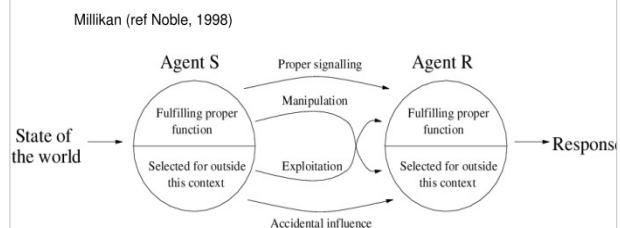
A biologist's perspective

- Accidental influence: e.g., pig scares mouse
- Exploitation: e.g., cheetah catches injured gazelle
- Manipulation: e.g., broken wing display
- Proper signalling: e.g., bee dance

- Index: a signal that cannot be faked because its intensity is physically connected to the quality being signalled.

- Handicap
- Cue
- Signal
- Sign
- Icon

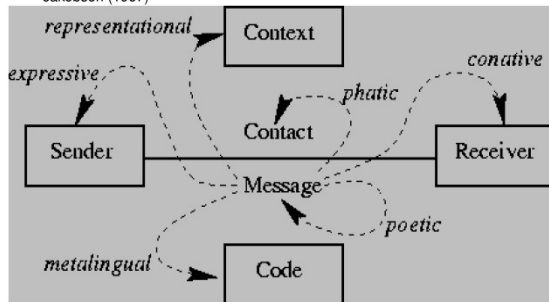### Shannon / information theory

(Weaver, 1949)

- Three levels of analysis:
  - Technical level
  - Semantic level
  - Effectiveness level
- At the technical level, the content of communicative act is irrelevant; the source is viewed as a stochastic process;
- Shannon's concept of information: reduction in uncertainty about the source;
- (Note: a subjectivist interpretation of probabilities)
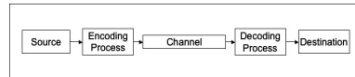
Functions of communication

Jakobson (1967)

What's special about the 'code' used in human language?

- Shannon's information theory allows us to quantify the amount of information transferred in various species, but has little to say about qualitative differences in the type of messages conveyed;
- Jakobson's and Millikan's models organize questions about the proximate and ultimate functions of communication, but have again little to say about the actual content of human communication in a comparative perspective.
- Symbolism, compositionality

Functions of communication

- Expressive, e.g., *ouch!*
- Representational, e.g., room 2.02 is over here
- Phatic, e.g., *how are you?*
- Conative, e.g., imperatives
- Poetic, e.g., *absence of evidence*
- Metalingual, e.g. definitions
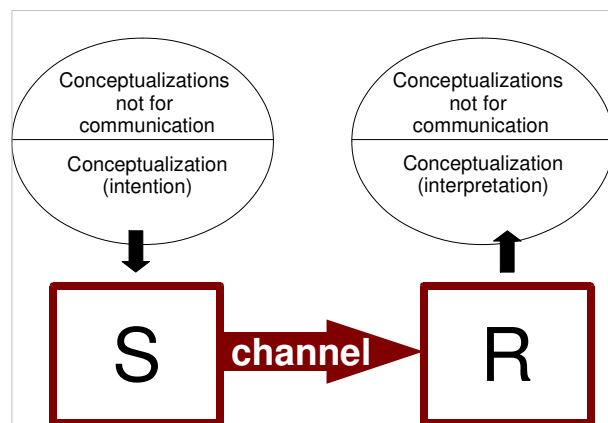
Symbolicism

[Even] single-symbol utterances in young children go beyond primate calls in important respects [...]. Perhaps the most important difference is the non-situation-specificity of human words.

The word *kitty* may be uttered by a baby to draw attention to a cat, to inquire about the whereabouts of a cat, to summon the cat, to remark that something resembles a cat, and so forth.

Other primates' calls do not have this property. A food call is used when food is discovered (or imminently anticipated) but not to suggest that food be sought. A leopard alarm call can report the sighting of a leopard, but cannot be used to ask if anyone has seen a leopard lately [...] *(Jackendoff, 2002)*

Generative grammar

beware: Symbols

- A signal whose form is unrelated to its meaning (Maynard Smith & Harper, 2003)
- [...] the reference of a symbol is a detached representation, while a signal refers to a cued representation. [...] a signal refers to something in the outer environment or to the emotional state of the signaler, while a symbol refers to the inner world. (Gardenfors 2003)
- A sign is a symbol when it refers to its object by virtue of a law. [...] The symbolic rule may have been formulated a priori by convention, or a posteriori by cultural habit. (Peirce/Desemedt).

Noam Chomsky

- Born 1928 in Philadelphia
- 1955 Logical Structure of Linguistic Theory
- 1957 Syntactic Structures
- 1965 Aspects of the Theory of Syntax
- 1968 The Sound Pattern of English (with Halle)
- 1975 The Logical Structure of Linguistic Theory
- 1981 Lectures on Government and Binding
- 1995 The Minimalist Program
- Now: Professor emeritus at MIT

5

Compositionality

It is astonishing what language can do. With a few syllables it can express an incalculable number of thought, so that even a thought grasped by a terrestrial being for the very first time can be put into a form of words which will be understood by someone to whom the thought is entirely new.

This would be impossible, were we not able to distinguish parts in the thoughts corresponding to the parts of a sentence, so that the structure of the sentence serves as the image of the structure of the thoughts. (Frege, 1923)

**Autonomy of syntax, generative methodology**

(Chomsky, 1957)

(1)   Colorless green ideas sleep furiously.

(2)   *Furiously sleep ideas green colorless.

(3)   have you a book on modern music?

(4)   the book seems interesting.

(5)   *read you a book on modern music?

(6)   *the child seems sleeping.

$$grammaticality : \Sigma^* \mapsto \{\,yes, no\,\}$$

6

Which generative model?

- Chomsky's proposed methodology of building "generative models"
- needed: a mathematically precise formalism for generating sequences of words
- Most popular such formalism in the 1950s? Shannon's Markov models!
- Hidden Markov Model (HMM) = Finite-state machine + probabilities.
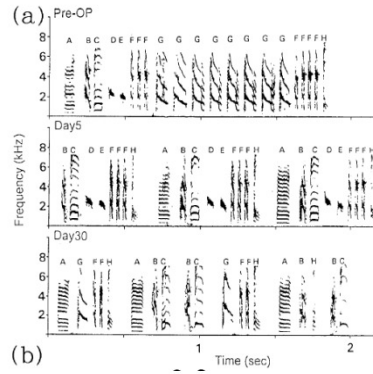
Shannon 1948

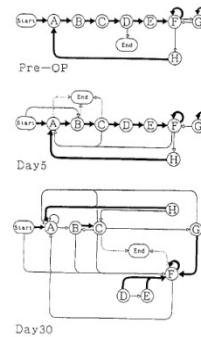Approximations of English based on character transition probabilities:

0-order: XFOML RXKHRJFFJUJ ZLPWCFWKCYJ FFJEYVKCQSGHYD QPAAMKBZAACIBZLHJQD

1st-order: OCRO HLI RGWR NMIELWIS EU LL NBNESEBYA TH EEI ALHENHTTPA OOBTTVA NAH BRL

2nd-order: ON IE ANTSOUTINYS ARE T INCTORE ST BE S DEAMY ACHIN D ILONASIVE TUCOOWE AT TEASONARE FUSO TIZIN ANDY TOBE SEACE CTISBE

3d-order: IN NO IST LAT WHEY CRATICT FROURE BIRS GROCID PON-DENOME OF DEMONSTURES OF THE REPTAGIN IS REGOACTIONA OF CRE


(Okanoya & Yamaguchi, 1997)


Shannon 1948

Approximations of English based on word transition probabilities:

1st-order: REPRESENTING AND SPEEDILY IS AN GOOD APT OR COME CAN DIFFERENT NATURAL HERE HE THE A IN CAME THE TO OF TO EXPERT GRAY COME TO FURNISHES THE LINE MESSAGE HAD BE THESE

2nd-order: THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED
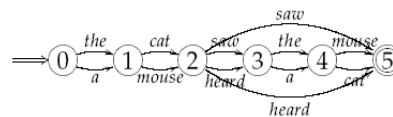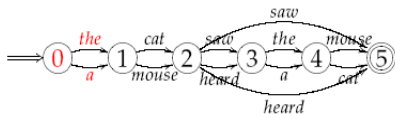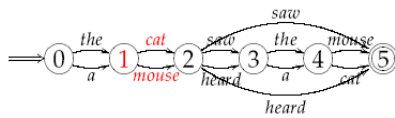

0→the 1
0→a 1

Finite-state Automata are inadequate

(Chomsky, 1957)
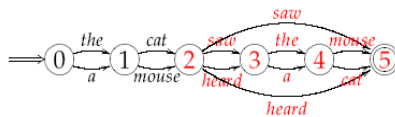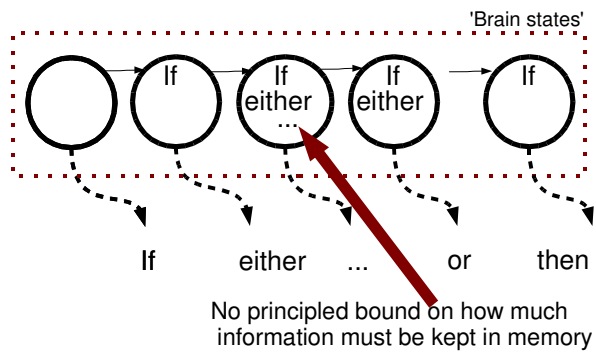
Let S1, S2, S3, S4, S5 be simple declarative sentences in English. Then also

(2) If S1, then S2.
(3) Either S3 or S4.
(4) The man who said that S5, is arriving today

are sentences of English.

0→the 1
0→a 1
1→cat 2
1→mouse 2

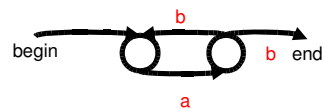'Brain states'

If   either   ...   or   then

No principled bound on how much information must be kept in memory

0→the 1
0→a 1
1→cat 2
1→mouse 2
2→saw
2→heard
2→saw 3
2→heard 3
3→the 4
3→a 4
4→cat
4→mouse

Simplest example of a "finite-state language":

$(ab)^n$

E.g. ab, abab, ababab, abababab

Simplest example of a "context-free language":

$a^n b^n$

E.g. ab, aabb, aaabbb, aaaabbbb, ...

**(Context-Free) Phrase Structure Grammars**

- Sentence → NP + VP
- NP → T + N
- VP → Verb + NP
- T → the
- N → man, ball, etc.
- Verb → hit, took, etc.

---

(1) a. Gilligan claims that Blair deceived the public.
  b. Gilligan claims that Campbell helped Blair deceive the public.
  c. Gilligan claims that Kelly saw Campbell help Blair deceive the public.
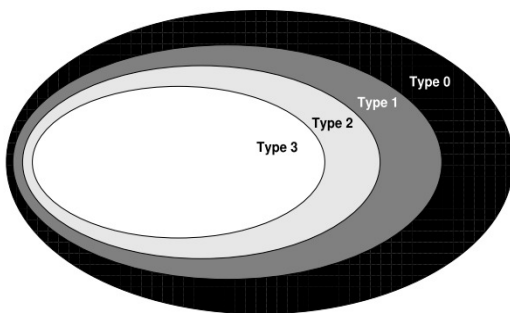   (tail recursion)

Chomsky Hierarchy

3. Finite state grammars	$A \to a, A \to aB$	$(ab)^n, a^n b^m$
2. Context-free grammars	$A \to \gamma$	$a^n b^n$
1. Context-sensitive grammars	$\alpha A \beta \to \alpha \gamma \beta$	$a^n b^n c^n$
0. Unrestricted grammars	$\alpha \to \gamma$	$\{a^n b^m c^l | l = n * m\}$

The Chomsky Hierarchy
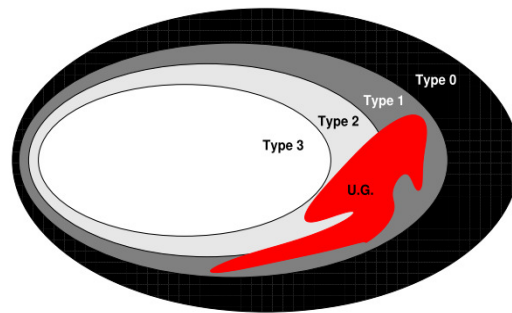
The Chomsky Hierarchy

Universal Grammar

- The set of possible natural languages;
- The innate contribution to everyone's "knowledge of language".
- The initial state of the "language acquisition device";
- The universal blueprint underlying all language + the "language acquisition device".