

## Language

### 1. Why is language important for Information Science?

communication  
window to thought

systems:

information retrieval  
interfaces  
    explanation facilities

text:

    abstraction  
    report generation  
intelligent tutoring systems  
communication management systems

## 2. What is language?

vs other means of communication  
bee dance  
primate gestures  
music & dance  
programming languages

Hockett 1960

semanticity & arbitrariness of units  
displacement in time and space  
discreteness & productivity

many to many mapping  
meaning - utterances

infinite sets of meanings & utterances  
power of  
formal systems  
modularity

Structure

phonology	sound
syntax	structure
semantics	meaning
pragmatics	use

### 3. Chomsky: Formalization of Linguistics

productivity & regularity  
competence vs. performance

grammar as formalism:

distinguish

grammatical vs. ungrammatical  
(as a compiler does)

linguistic intuitions -> grammar

acceptability

paraphrase / synonymy

ambiguity

lexical

structural

adequacy criteria:

observational

descriptive

explanatory

distinguish?

describe differences?

explain reasons for?

## 4. Syntactic Structure

### Phrase Structure

(psychological evidence)

pause structure

speech errors - follow phrase structure

reading times

Graf & Torrey (1966)

line by line paragraph reading

"transformations" (Chomsky)

active/passive

questions (yes/no, wh-questions)

relative clauses

### Rewrite Rules

terminal and nonterminal symbols

$A \rightarrow B + c$

noun phrase  $\rightarrow$  determiner adjective noun

$\langle \text{integer} \rangle ::= \langle \text{digit} \rangle \mid \langle \text{integer} \rangle \langle \text{digit} \rangle$

Chomskian Hierarchy (of languages)  
(from simplest to most complicated)

3. regular or finite state grammar
2. context free
1. context sensitive
0. transformational

(capital letters are non-terminal symbols, lower case are terminal symbols)

3.  $A \rightarrow B + c$  (at most one of each)
2.  $A \rightarrow$  any combination of terminal and non-terminal symbols
1.  $c_1 A c_2 \rightarrow c_1$  anything  $c_2$   
where  $c_1$  and  $c_2$  are optional contexts that are not changed but determine applicability of the rule, which is otherwise like a context free rule
0. anything  $\rightarrow$  anything  
no restrictions at all,  
elements of left hand side can be reordered or deleted, on right hand side new elements can be reintroduced without restrictions

deep structure vs. surface structure

phrase structure rules (context free)  
transformational rule

## 5. Relation between Thought & Language?

Thought = Language:

Behaviorist Proposal

subvocal speech

Smith et al., 1947 curare

Thought depends on Language

(linguistic determinism)

Whorf: linguistic relativity

but deeper interpretation of this....

Language depends on Thought

phrase & propositional structure are related

perception -> terminology

Rosch, 1973,

Dani color "learning"

preferred word order (SVO etc.)

Greenberg (1963)

SOV 44%

SVO 35%

VSO 19%

VOS 2%

OVS, OSV -

Thinking is independent of Language

(modularity)

Fodor: "mentalese"

-> two issues

acquisition, uniqueness question

comprehension

What does Language Understanding Require?

knowledge of world

physical

social

speaker's model of world

conversational context

physical context

goals

## 6. Language Acquisition & Innateness/Uniqueness Issue

brain differentiation

ease of learning

earlier

less conscious

less variability

no evidence of deliberate teaching or reinforcement

developmental regularities

timing

acquisition patterns

critical periods

cross language regularities

cross language developmental regularities

language universals & parameter setting

(A-over-A)

(pro-drop)

rules and the past tense

correct irregulars, overgeneralize, correct  
Rumelhart & McClelland (1986) PDP model  
Pinker & Prince (1988)

data didn't fit

learning didn't fit

meaning (rang vs. ringed)

\* problem of scaling up

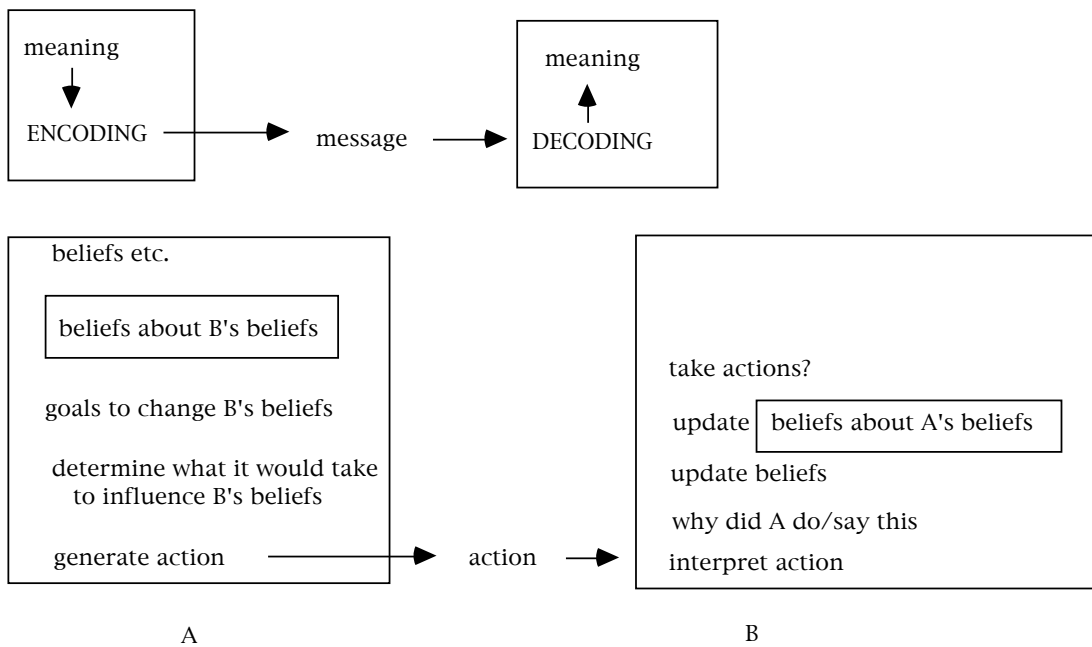
## Language Comprehension

tool for communication:  
(vs. Shannon & Weaver Model)

gist  
implication  
context

plan-based understanding

Shannon & Weaver Model vs. Communication Model



Example: Wilensky's UC system

literal understanding  
indirect speech acts  
domain understanding

Plan Based Understanding

"Do you have a watch?"

"Do you know where the bus for downtown stops?"

"Is there a gas station around here?"

Answer requires: "Why do you want to know?"

Outline of Plan Based Understanding

Plan schemas

steps in the plan

items involved in, required for plan

pointers from items to plans they are involved in  
watch -> time telling plan

pointers from plans to larger plans of which they are subplans  
telling time ->  
various activities that require knowing what time it is

"Do you have a watch?"

"Don't worry, your next meeting has been moved back a half  
an hour."

identify what plans speaker may be executing that require  
the information they are seeking

## Language Understanding:

### Structure

phonology	sound
syntax	structure
semantics	meaning
pragmatics	use

### 1. Perception:

lexical access and retrieval  
resolution of lexical ambiguity

### 2. Parsing:

constituent & sentence construction  
evidence for.... e.g.,  
Jarvella, 1971,  
Caplan 1972

*Now that artistists are working fewer hours  
oil prints are rare.*

*Now that artistists are working in  
oil prints are rare.*

### 3. Semantic processing

deriving propositional structure  
immediacy of interpretation  
integration of syntactic and semantic processing

Frame-based understanding:

Tarzan: "Jane fruit eat"

eat

isa            action

agent        animal or agent or activity..

capable of consuming things

object       food or something

capable of being consumed

constraint   agent capable of consuming object

preferences:

deep explicit methods

statistical methods

#### 4. Utilization:

integration of sentence meaning with existing knowledge:

inferences

of reference, pronominal reference

of unstated, implied information

given/new

text structure

coherence

bridging inferences

pragmatics

sentence utilization and interpretation in context:

modeling the world

physical

social

models of agents

(plan-based understanding)

conversational context

Integration into existing knowledge

Clark & Haviland 1970's "given / new" distinction,  
supposition vs assertions:

subject predicate

stress

definite/indefinite articles

structures such as fronting

"It was a chair that she bought"

evidence:

Haviland & Clark 1974

"given an alligator.....The alligator

"wanted an alligator.....The alligator

Loftus & Zanni 1975

"a/the broken headlight"

inference of the unstated:

McKoon & Ratcliff 1981

pounded boards -> hammer

Just & Carpenter 1987

swept -> broom

pronoun reference:

Just & Carpenter 1987

gender

grammatical role

recency

general knowledge

need for hierarchical representation of discourse

Grosz 1975

assembly and reassembly of pumps

text structure

Kintsch & Keenan 1973

reading time a function of number of propositions

Kintsch & van Dijk 1978

hierarchical structure

influences memory

bridging inferences (Haviland & Clark, 1974)

capacity limit ~4 propositions

-> problems with distant inferences

coherence

influences recall

## Pragmatics - early insights:

Searle: Speech Acts

representatives

directives

commissives

expressives

declarations

Indirect Speech Acts

H's ability to perform A

S's wish that H will do A

referring to H's doing A

H's desire or willingness to do A

reasons for doing A

\*not imperatives

\*not ambiguous

\*conventional forms

\*not idioms

\*can be used literally

\*not additional sentence meaning

but additional speaker meaning

applications:

database interfaces

CSCW (computer supported cooperative work)

Grice

conversational implicature

Talking is rational, purposive behavior based on common goals, interactions of participants, continuation of transaction

Cooperative Principle

"Make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged."

Conversational Maxims:

1. Quantity:  
enough, not too much
2. Quality:  
don't say what you believe is false or for which you lack information
3. Relation:  
be relevant
4. Manner:  
be brief, orderly, avoid obscurity, ambiguity