Language: Psycholinguistics

Language

→ The batter was trying to hit a home run.

Different levels of Analysis:

$[\text{The batter}]_{NP}$ $[\text{was trying to hit a home run}]_{VP}$ SYNTAX

The batter was try -ing to hit a home -run MORPHOLOGY

[t] [h] [e] [b] [a] [t] [t] [e] [r] [w][a] [s] . . . . PHONOLOGY

How is language produced

Air moves from the lungs.

Sounds are produced by impeding and or interrupting the flow of air.

-b lips

-s tongue behind teeth restricts air

-sh tongue at palate restricts air

The vocal cords produce sounds by vibration.

Phonology

phones - acoustically different speech sounds

phonemes- phones that make a difference in meaning

bat vs. pat

allophones - phones that are different, but do not make a difference in meaning.

The [p]'s are allophones in English.

p^hat vs spot

Sound Categorization

Sounds can be categorized by how they are produced (features):

place of articulation - Where the airflow is restricted.

Lips- [p], [b] Lips/Teeth-[f], [v] Alveolar Ridge- [t], [d]

manner of production- How the airflow is restricted

full closure [p], [b], [t] partial closure [f], [z]

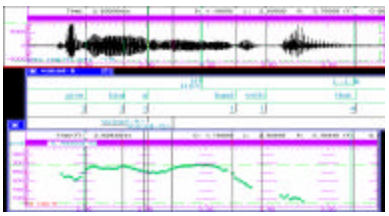
voicing- Whether the vocal chord vibrate

[g] vs [k] [z] vs [s]

…segmentation isn't easy

Difficulty of Speech Perception

-Words aren't easily segmentable. Phonemes can take on features of neighboring phonemes in what is known as **coarticulation**. There's a large amount of variance in the signal.



Categorial Perception

We hear speech sounds as members of a category. We are sensitive to differences between categories but insensitive to differences within categories.

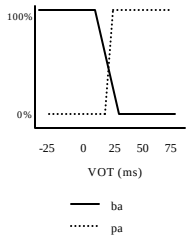
[b] vs. [p]

[b] is voiced, [p] is unvoiced

In the word "pit" voicing starts after 60 msec (**VOT**), in "bit" it's about 0 msec.

Will listeners hear "pit" or "bit" if we systematically vary the VOT?

Categorical Perception



- There is no graded membership. People hear either one phoneme or the other.

- Members of the same category with different VOT's are difficult to discriminate.

Syntax

"Language makes infinite use of finite means."

A finite set of elements (words) can be combined to form an almost infinite number of larger strings (sentences).

1) Syntax provides structure that tells us who did what to whom…

An elephant chased a lion.

A lion chased an elephant.

2) A set of rules that determines which sentences are good and which are bad.

John likes the Yankees

*John like the Yankees

Constituents

Words appear to be grouped together into constituents that are assembled into a hierarchy. Distributional analysis allows us to figure out what the constituents are:

x likes x.

John likes **the teacher**.
The boy likes **the teacher**.
The small boy likes **the teacher**.

The teacher likes **John**.
The teacher likes **the boy**.
The teacher likes **the small boy**.

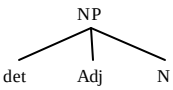
* John the boy the teacher "like" is not in the same category as "the boy"

NP → the teacher, the boy, John, the small boy. . . .

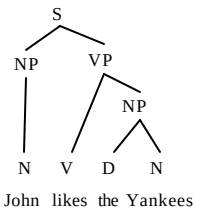
Phrase Structure Rules

NP → (det) A* N

"A noun phrase consists of an optional determiner, followed by any number of adjectives, followed by a noun"



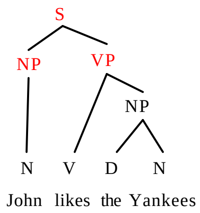
Phrase Structure



S → NP VP
VP → V NP
NP → (D) N

Rewrites rules restrict the composition of syntactic constituents

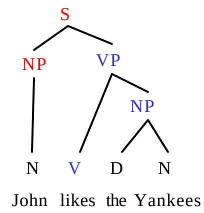
Phrase Structure



S → NP VP
VP → V NP
NP → (D) N

Rewrites rules restrict the composition of syntactic constituents

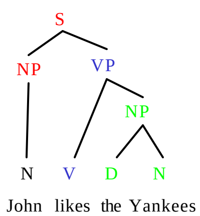
Phrase Structure



S → NP VP
VP → V NP
NP → (D) N

Rewrites rules restrict the composition of syntactic constituents

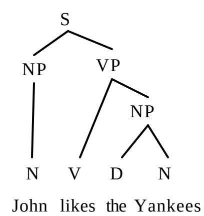
Phrase Structure



S → NP VP
VP → V NP
NP → (D) N

Rewrites rules restrict the composition of syntactic constituents

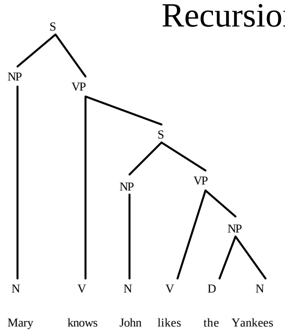
Recursion



VP → V S
S → NP VP
VP → V NP
NP → (D) N

Rewrites rules restrict the composition of syntactic constituents

Recursion



VP → V S
S → NP VP
VP → V NP
NP → (D) N

Language Acquisition

How do kids learn this stuff ?

All kids can do it:

“ It is a very remarkable fact that there are none . . .without even excepting idiots, that they cannot arrange different words together, forming of them a statement by which they make known their thoughts; while on the other hand, there is no other animal, however perfect and fortunately circumstanced it may be, which can do the same.”

-Descartes, 1611

Language is Hard

Figuring out the right rules is difficult.

John ate food John ate.

John is too stubborn to talk to Bill John is too stubborn to talk to.

Question formation:

The man is a fool Is the man a fool?

The man who is amusing is a fool *Is the man amusing is a fool?

The man who is amusing is a fool Is the man who is amusing a fool?

The man is a fool who is amusing *Is the man is a fool who amusing?

The correct rule has to refer to the hierarchy of the sentence.

Language Learning

Some aspects of language seem to innate although exactly what kids bring to the learning process is highly controversial.

Arguments for the innateness of language:

- Critical Period
- Children can invent languages
  - Creoles
  - Sign Language

Psychological Rules

One claim is that kids are learning an algebraic system of rules

past tense: X + "ed" = past tense

"turn" + "ed" = "turned"

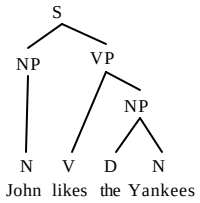
Kids erroneously apply "ed" to irregular verbs (runned, goed, etc…). These are called **over-regularization errors**, and they suggest that kids aren't just imitating what they hear.

Might be evidence for rules although connectionist networks can learn to do the same thing.

Are rules real or the product of a system that produces outputs that only appear to be governed by rules?

How do we figure out Sentence Structure?

Listen to the entire sentence first and then assign a structural analysis to it.



Advantages:

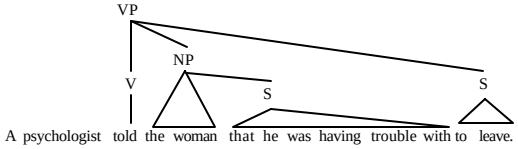
You won't make mistakes or get led down the garden path

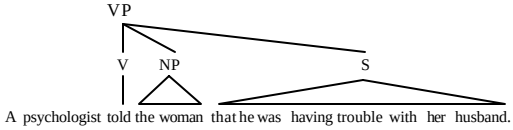
Disadvantages:

You'd need a huge amount of memory. The world's longest sentence was in James Joyce's *Ulysses*. 40 pages long

Local Ambiguity

The problems people have with interpretation suggests that they build structure "on the fly".





Processing Language

How do we process language?

-Modularity

-Interactionist

Two areas of Language

- Processing Sentences
- Processing Words

Modularity

Modules are (Fodor, 1983):

1) Domain Specific - they only operate over a certain set of inputs (e.g. syntax, color vision, etc.)

2) Informationally encapsulated - Other parts of the system cannot access processes within module, only ouput

3) Innate

4) Mandatory - Once given an input, modules run automatically

Interactionist view - Syntactic and Lexical Access involves considering multiple sources of information simultaneously, and making choices that satisfy the constraints based on this information.

Parsing

How do people make parsing decisions?

Strategy 1: Modularity approach. Syntactic principles are used to parse a structure incrementally.

Strategy 2: Constraint Based approach. A listener uses multiple sources of information to parse a sentence.

Syntax first

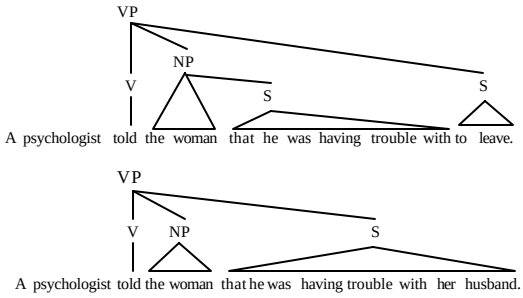
Frazier (1987) claimed that the parser is modular. Parsing is guided by a few basic syntactic principles. Non-syntactic information is only used when a mistake is made.

1) **Late Closure** - New parts of a sentence are attached to the phrase or clause that is currently being processed.

John had a date with the sister of the teacher who was arrested.

2) **Minimal Attachment** - New parts of a sentence are attached so that the sentence results in the fewest nodes possible

Local Ambiguity



Constraint Based Models

Listeners use lots of different sources of information in their first pass parse of a sentence, not just syntax (Tanenhaus et al, 1995; MacDonald, Pearlmutter & Seidenberg,1998)

Semantics can play a role:

1) The defendant examined by the lawyer turned out to be unreliable.

2) The evidence examined by the lawyer turned out to be unreliable.

Constraint Based Models

Linguistic Context:

A jockey raced a horse past the barn and a horse past the lake. The horse raced past the barn fell.

Non-linguistic Context

Constraint Based vs. Modular

What do listeners know and when do they know it?

The evidence suggests that non-syntactic information is used very early on in processing, but the modularity hypothesis isn't necessarily ruled out.

Words

referent- The thing in the world that is referred to.

…but this can't be all there is to word meaning

Venus is the evening star.

Venus is Venus.

Frege (1892) proposed that words have **senses**..meaning how we frame or understand a referent. How do we access meaning when we hear a word?

Words with Multiple Meanings

bank- financial institution, side of a river

bug - insect, a spying device

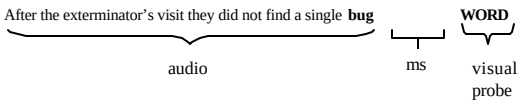
How do we know which meaning to access?

For several weeks after the exterminator's visit they did not find a single **bug** in the apartment.

Modularity hypothesis: all meanings are activated…then context is used to select the right one.

Interactionist hypothesis: Only the "insect" meaning will be activated, because the context biases the interpretation.

Lexical Access (Swinney 1979)



INSECT faster than CUSTOM after 0ms

SPY faster than CUSTOM after 0ms

INSECT faster than SPY after 300 ms

Evidence for the modularity hypothesis.

Lexical Access

However other studies have shown that other factors can affect lexical access.

Meaning dominance - The relative frequency of the individual meanings of a word play a role in whether they're accessed.

bank - financial institution > side of the river

pitcher- ball player = container

Context interacts with frequency to determine lexical access.

I opened a checking account at a bank.

The fisherman waited by the bank.

Lexical Access

Multiple sources of information (lexical frequency and context) seem to affect lexical access.

Dominant meaning is easiest to access, but context can provide evidence for one meaning over another.

Language

What is the structure of language?

How do kids figure out this structure?

How do we figure out the structure of a sentence when we hear it?

How do we access the meaning of words?