8. Language

Language Comprehension and integration

Memorization of structures (Behavioral psychology)
From experience, we memorize structures: which can be words or types of phrases that go together. In this case a fundamental role is played by memory, to store statistics

Elements in S-R approaches to language
This approaches claim we learn which words come after each other. they also allow for some abstraction. This is more compact than storing single word's associations.

Abstract elements in sentence level
Longer phrases seemed to resolve into combinations of shorter ones, which in turn could resolve in single words.

Meaning in S-R approaches to language

- **Substitution View:** words directly replaced objects in the world. The meaning of a word was simply the physical response it triggered, similar to the object itself.
- **Disposition view:** a word's meaning is our readiness to respond in a particular way.

Abstract rules (Chomsky grammar)
We learn abstract rules that specify constraints on grammaticality

Transformational grammar
Grammar constitutes a theory of the speaker's underlying linguistic knowledge.

Sentence meaning is captured by its structure
Sentences are represented as trees, with the Chomsky hierarchy. As sentence meaning is related to its structure, it is possible to disambiguate ambiguous strings of words.

Derivational theory of psychological complexity
I identified an objective metric on which complexity can vary, and examined correlations between value of this metric and human behavior. The metric was the number of transformation needed to transform a kernel sentence to a given sentence. The entire sentence was taken as a unit, and a set of pre-set rules (a grammar) was used to see if the given sentence can be derived from the kernel sentence. No incremental parsing was used.

Parsing-based accounts
Parsing is an incremental process by which we try to assign a structure to what we have read/heard. In some cases we experience local ambiguity that slows down our comprehension. Such ambiguity is, in many cases, local and is disambiguated by later content. Parsing-based accounts try to explain why some tentative interpretations are preferred and why some sentences are more difficult.

Insensitivity of the Human Sentence-Processing System to Hierarchical Structure Frank and Bod (2011)

Problem
Traditional AI language models relied on strict grammatical rules. They wanted to see if models that simply learned from data, without explicit grammar rules, could better predict how humans understand sentences.

Methodologies
The researchers compared three model types trained on the "Wall Street Journal" corpus (using Part-of-Speech tags).

Training

- **Probabilistic Phrase-Structure Grammars (PSGs)**: These used complex, hierarchical grammar rules.
- **Markov Models (N-grams)**: These predicted the next word based only on the few words before it (sequential).
- **Echo State Networks (ESNs)**: A type of neural network that learns sequential patterns, potentially over longer distances.

Evaluation

- **Linguistic Accuracy:** How well they predicted the next word's Part-of-Speech.
- **Psychological Accuracy:** Crucially, how well their "surprise" at unexpected words matched human reading times (measured with eye-tracking).

Results

Sequential Models are linguistically accurate: Simpler Markov models performed as well as complex grammatical PSGs in predicting word sequences.

Neural Networks Best Predict Human Behavior: ESNs (neural networks) were the most accurate at predicting how humans actually read, meaning their "surprise" aligned best with human reading difficulties.