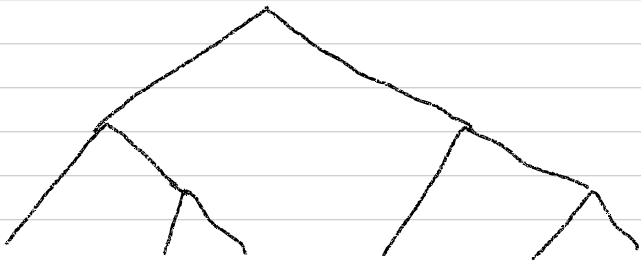
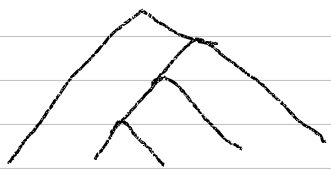
Immediate constituent analysis

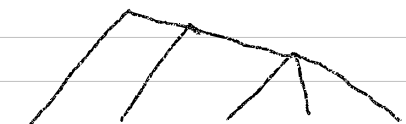
* Divide a sentence into its constituent elements
  - cut into two parts
  - then cut those parts into two
  - continue the segmentation
* Sometimes binary segmentation may not be possible, so more than two divisons may have to be made

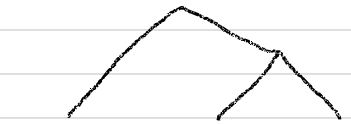


* Expansion
  - A set (or sequence) of elements is considered to be an expansion of another if it can be substituted for it
    - "The overconfident dirty-looking naughty kid" can be substituted with "a kid"
    - A kid,
      A naughty kid
      A dirty-looking naughty kid
      An overconfident dirty-looking naughty kid

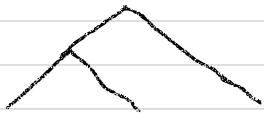
* Non-binary IC analysis example ≈ Ambiguity resolution







shirt is Egyptian



cotton is Egyptian

- IC analysis is just that — an analysis of a sentence. It does not say what other possible sentences can be formed in a language.
- Only generative models can do so, which is associated with transformational generative grammar. More on this later.

Richer information than IC analysis can be provided by Phrase Structure and PS Rules.

Phrase structure & Phrase structure rules

PS: The way parts of sentences/phrases are organised and related to one another
- a concept closely related to Chomsky's Transformational Generative Grammar

PS rules: describe the syntax of a language, how the constituent parts of the phrases are organized

↓↓

syntactic categories, including lexical and phrasal categories

e.g. $S \rightarrow NP\ VP$

i.e. A sentence consists of a noun phrase and a verb phrase.

$NP \rightarrow (Det)\ N_1$

$N_1 \rightarrow (AP)\ N_1\ (PP)$

where the constituents in round brackets are optional.

Generative grammar — Chomsky

- "a system of rules that in some explicit and well-defined way assigns structural descriptions to sentences"
- "an algorithm for specifying, or generating, all and only the grammatical sentences in a language"
- Not all models of grammar are generative. Some produce language by repeating memorized fragments or by probabilistic modeling.
- aims to explain the cognitive basis of language by formulating and testing explicit models of humans' subconscious grammatical knowledge
  - studies language as a part of cognitive science
  - studies the mental processes that allows humans to use language
  - descriptive more than prescriptive
  - distinguishes competence and performance
  - (language being innate and universal (vs. blank slate) (see. Universal Grammar)
- Different versions — Transformational Generative Grammar, Government and binding theory, Minimalist program, etc.

Universal Grammar

- language as an organ; pre-programmed for language
- humans' innate capacity to learn language
  - not a language but any language under specific contex or conditions
- "poverty of stimulus"
  - a human child is presented with limited, and insufficient data but can deduce from it a linguistic system
- critical period for the development of the language organ

However all these claims have been questions, not entirely refuted.

- FOXP2 gene, claimed to be responsible for encoding the ability to use language isn't actually always.