Computational Psycholinguistics

Lecture 5: Ambiguity Resolution in Parsing: Determinism, Parallelism

Afra Alishahi

November 24, 2008

Deterministic Parsing

- The issue of time: human parsing is fast but backtracking is slow, thus we should avoid it.
  - Use as much information as needed to ensure that the right decision is made
  - Build syntactic analyses only when there is sufficient grounds to guarantee that it is the correct one
  - If an incoming input does not fit into the structure constructed so far, then the parser fails (no re-analysis)

➡ Allows for predicting garden-path effects

Deterministic Parser of Marcus

- Marcus (1980) suggested a deterministic approach
  - Left-to-right, bottom-up, up to three items lookahead before making a decision
- Advantage: fast, clear prediction of garden-paths
- Disadvantage:
  - Not fully incremental: potentially large lookahead items are left on the stack
  - Problematic with head-final languages
  - Garden-path effects are not a matter of degree

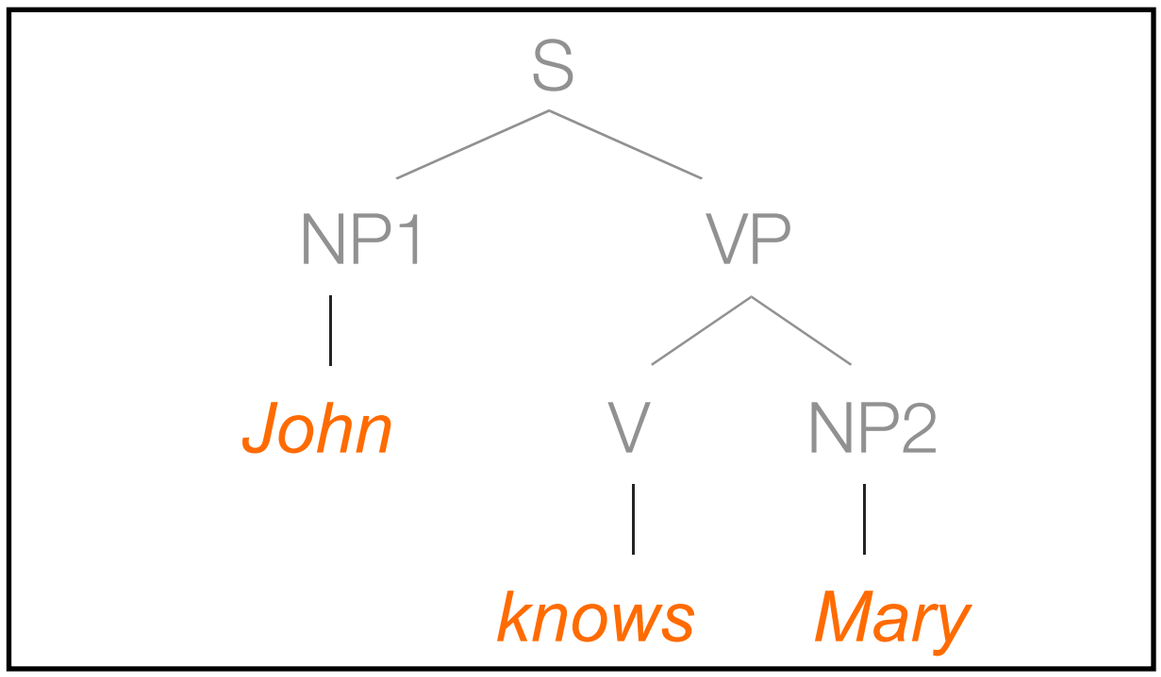
Underspecification

- Can local ambiguity be handled using underspecified representations?
  - Representations allow some ambiguity to remain, and be later removed without destructive re-parsing
- Description Theory (Marcus, Hindle, Fleck, 1983)
  - Using tree *descriptions* instead of trees, as a set of dominance and precedence relations

D-theory: An Example

- Trees are described as a set of nodes, and a set of precedence and dominance relations:

John knows Mary



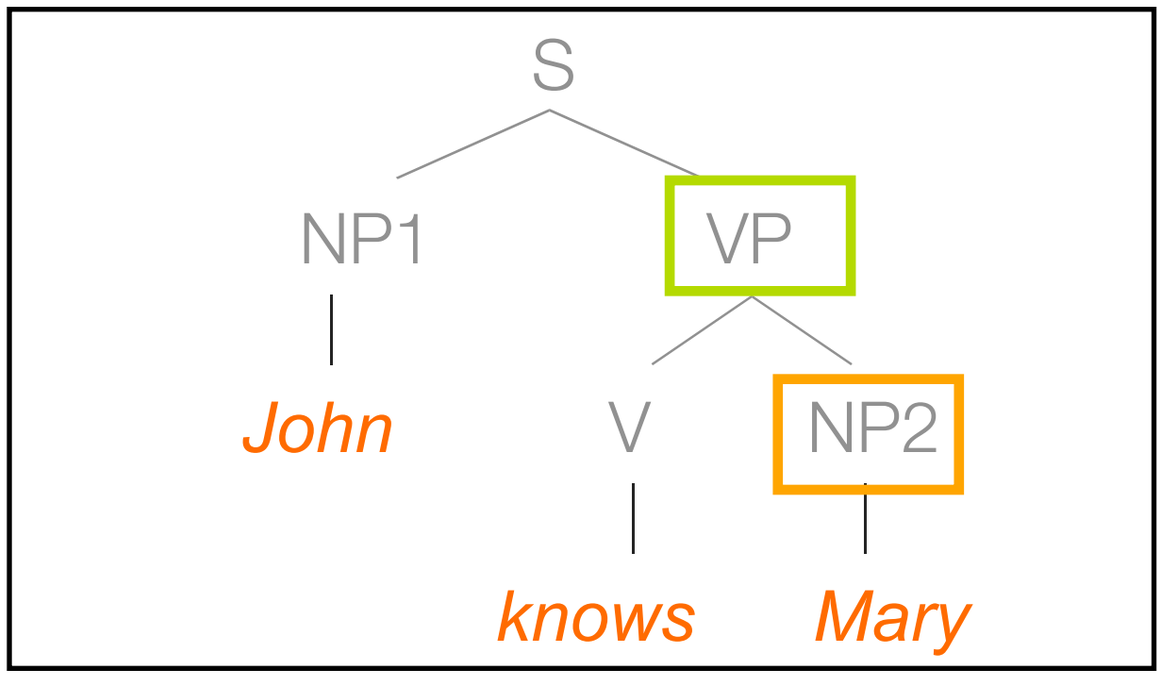
{dom(S,NP1), dom(S,VP), dom(S,V), dom(S,NP2), prec(NP1,VP), dom(VP,V), …}

Monotonic Parsing

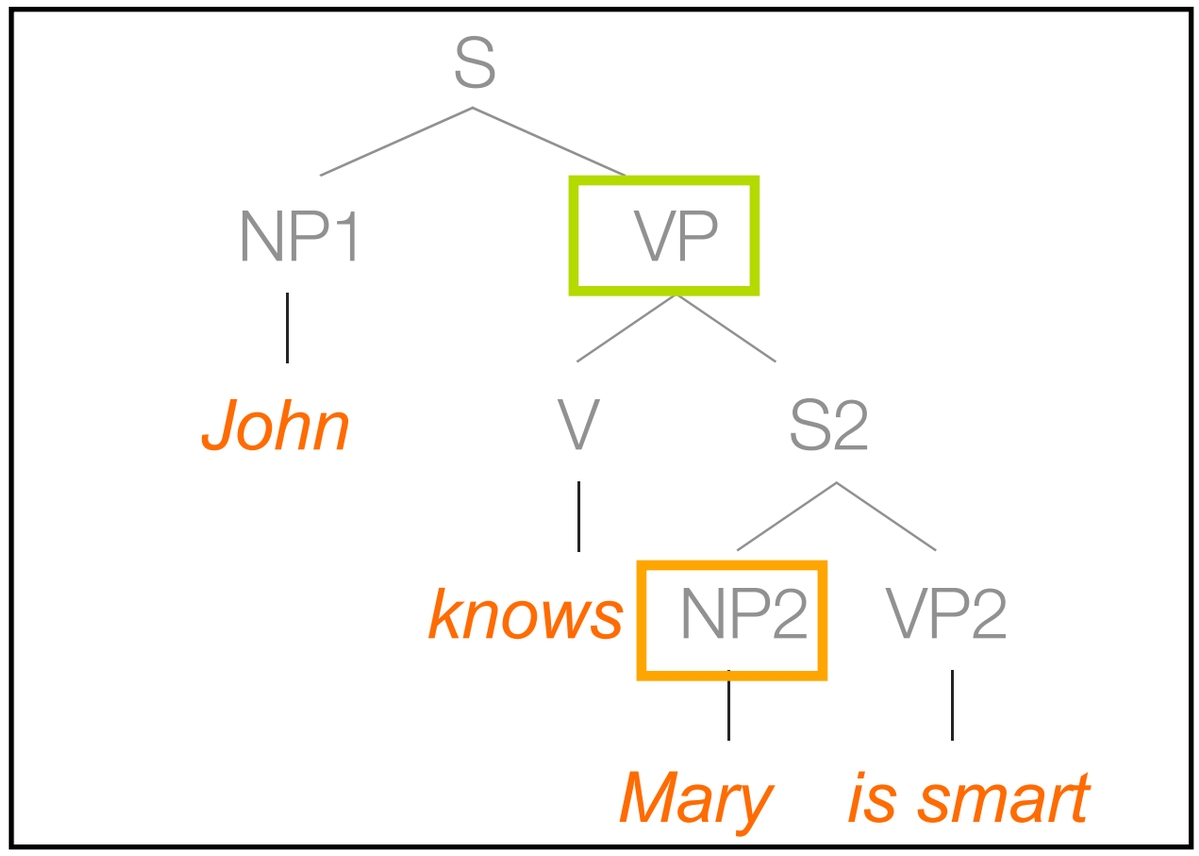
- Perform reanalysis without destructive backtracking (hence monotonic)
  - Structural revisions only require adding new precedence and dominance relations
  - removing a relation is not needed
- Predict processing difficulty when non-monotonic reanalysis is needed
- Weinberg (1994), Gorrel (1995), Sturt & Crocker (1996)

Monotonic Parsing: An Example

"John knows Mary …"



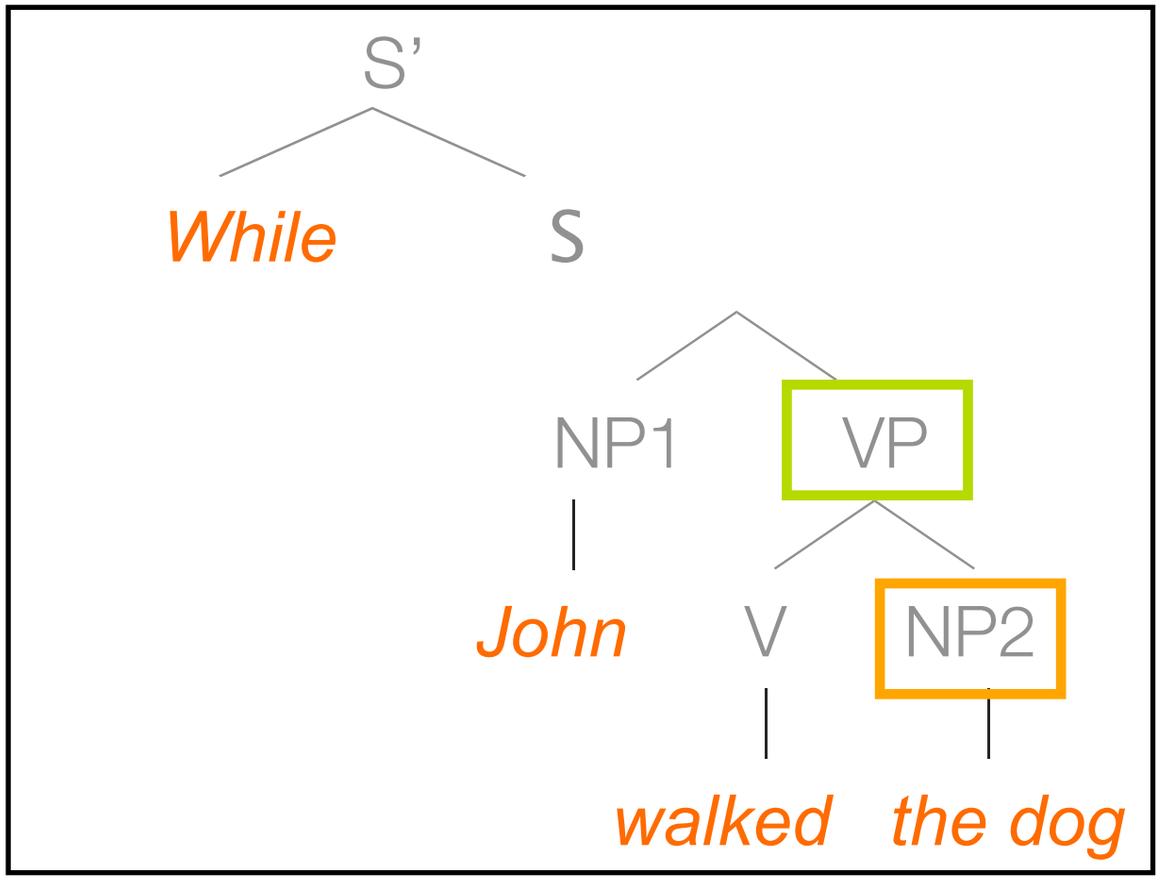
"John knows Mary is smart."



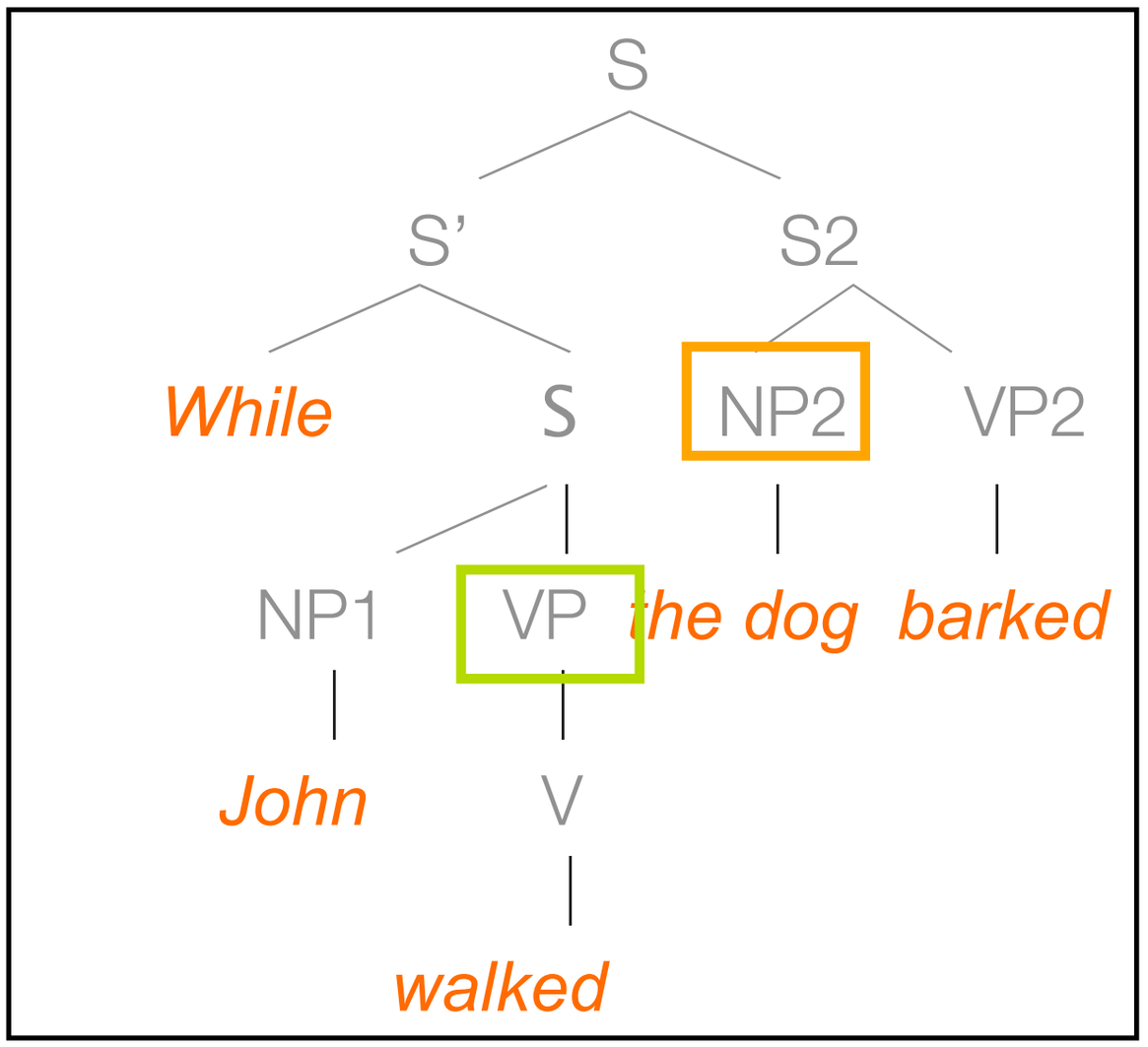
{dom(S,NP1), dom(S,VP), dom(S,V), dom(S,NP2), prec(NP1,VP), dom(VP,V),..}
dom(VP,S2), dom(S2,NP2), prec(NP2,VP2)…}

Parsing Difficulty

"While John walked the dog ..."



"While John walked the dog barked."



➡ Dominance / precedence relations are not preserved.

Parallel Parsing

- Assumption: people have the ability to construct alternative syntactic analyses in parallel
  - When ambiguity is encountered, pursue all possible options instead of choosing among them
- No reanalysis is needed
  - When one parse fails, it is eliminated from consideration
  - The correct parse is taking place in parallel

Infinite Parses

- Full parallelism (where every analysis is pursued) is not psychologically possible.

“I believe ...”

“I believe the daughter ...”

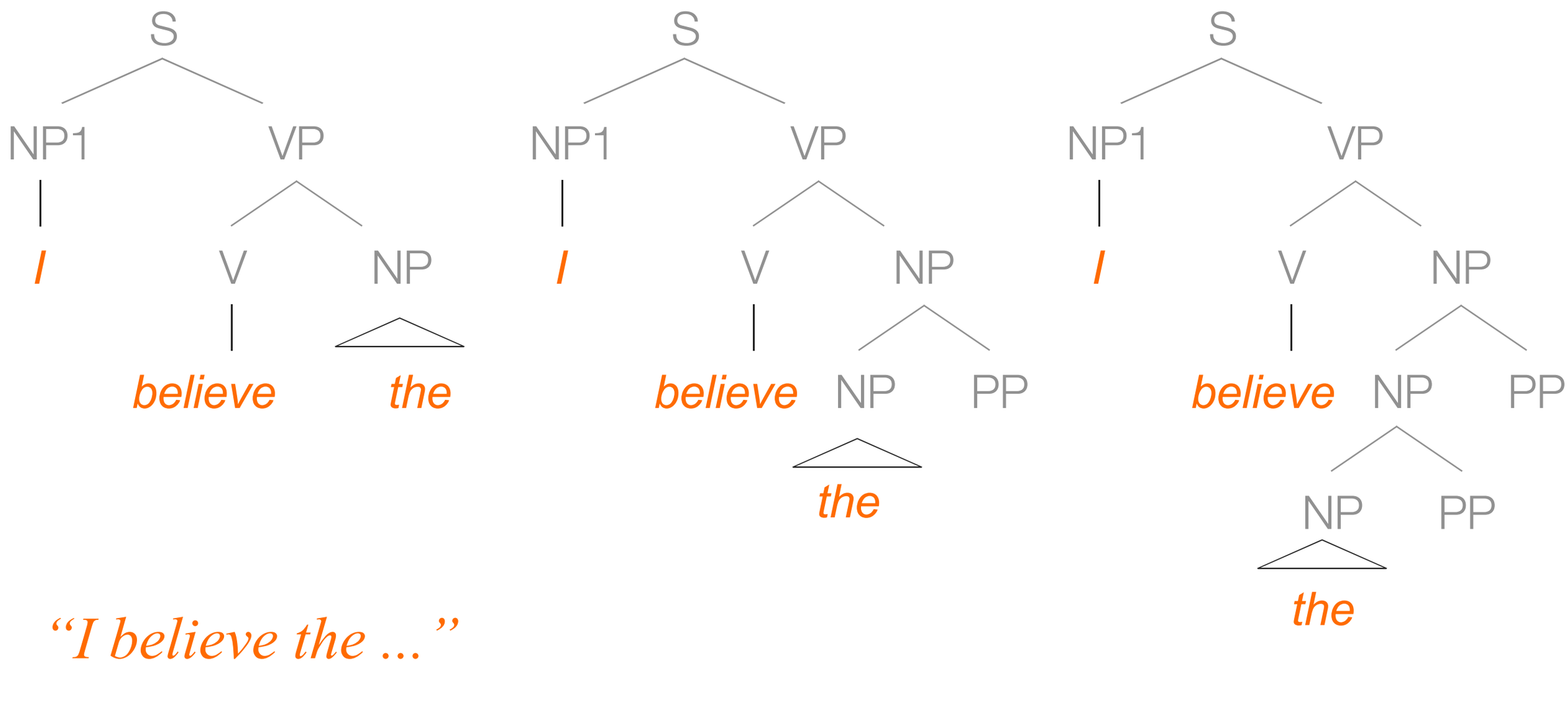
“I believe the daughter of the sister ...”

“I believe the daughter of the sister of the colonel.”

“I believe the daughter of the sister of the colonel is my aunt.”

Infinite Parses

- Full parallelism (where every analysis is pursued) is not psychologically possible.



"I believe the ..."

Bounded, Ranked Parallelism

- Full parallelism is not cognitively plausible:
  - Memory requirements for a full parallel parser can easily exceed human memory resources.
  - It does not explain the garden-path effects.
- Alternative suggestions:
  - Bounded parsing: there number of analyses that can be considered in parallel are limited.
  - Ranked parsing: analyses are ordered according to some measure (where rank shows preference).

Ranking the Parses

- Ranking determines which analyses to pursue in parallel and which ones to discard
  - Bounded parser will pursue highly ranked analyses
- Predictions:
  - Correct discarded analyses are difficult garden paths.
  - Correct low-ranked analyses are easy garden paths.
- Gibson (1991): rank according to a set of principles based on memory load.

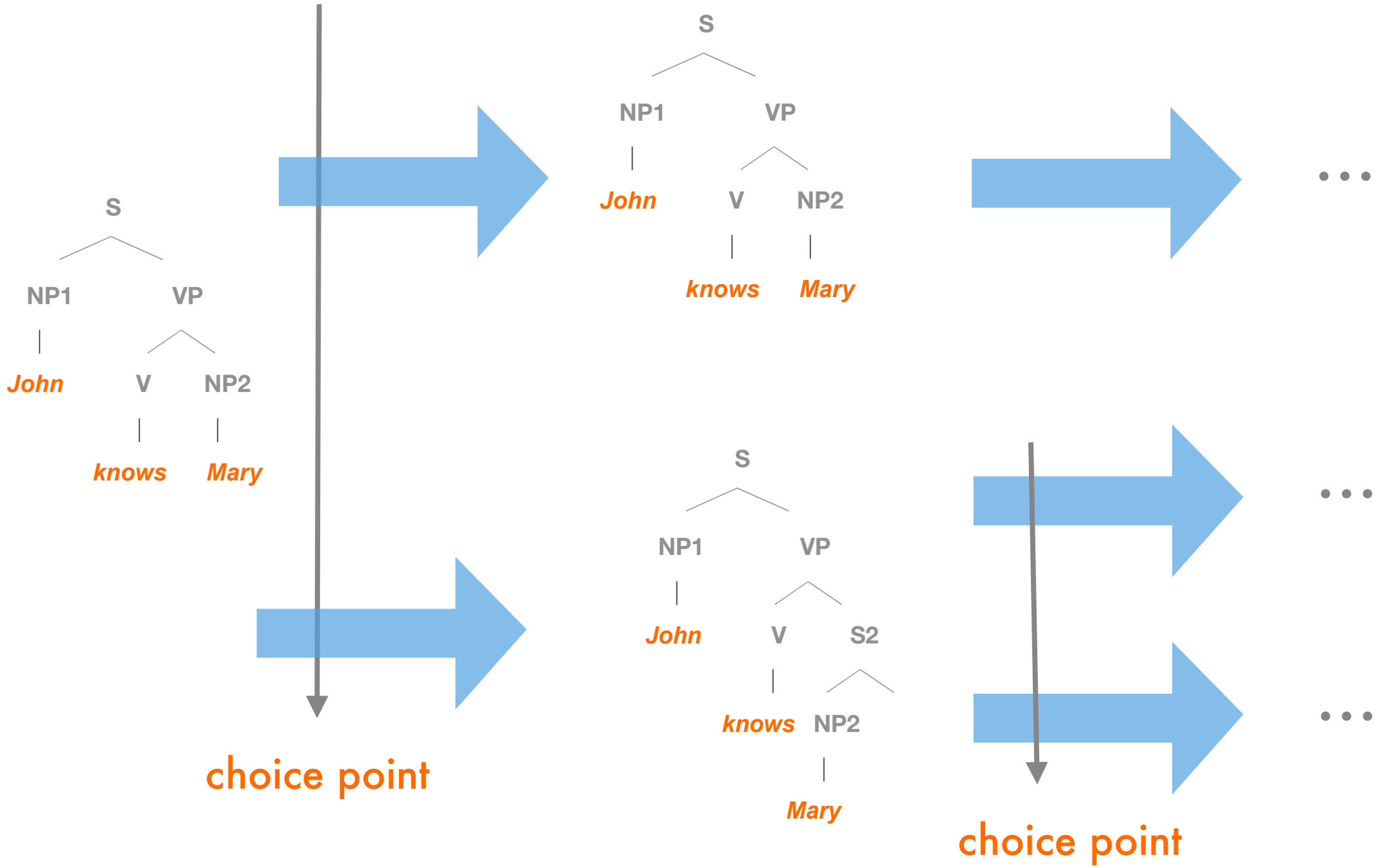
Momentary Parallelism

- Altman (1988):
  - All possibilities are considered at each choice point
  - Only one survives and is pursued
- Advantages:
  - Permits the use of semantic and pragmatic knowledge to assist in resolving local ambiguity
  - Limits the explosion of multiple analyses

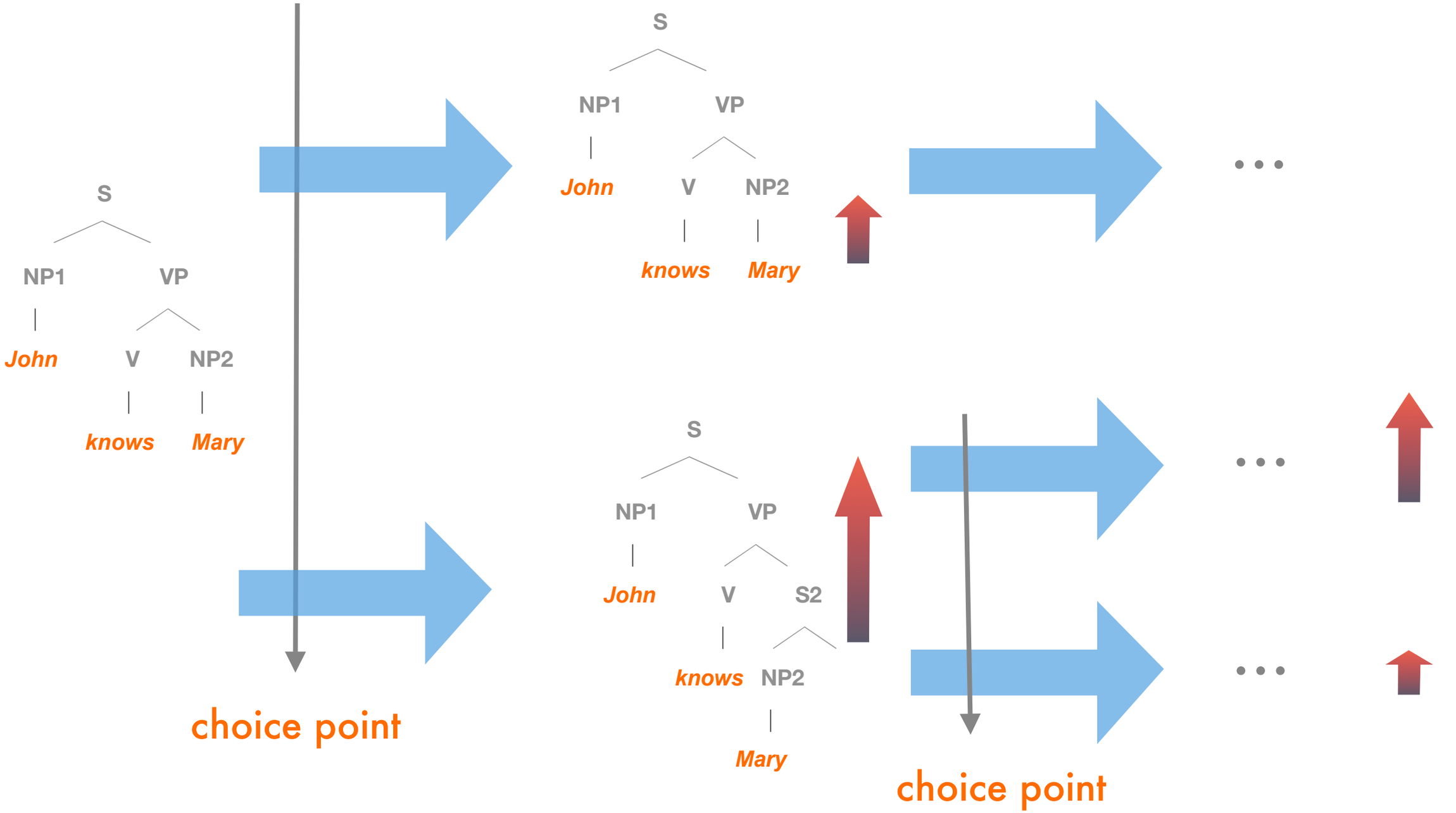
Competitive Activation

- A different approach:
  - Pursue multiple analyses in parallel
  - Allow these structures to compete with each other in the ranking process
    - E.g., MacDonald, Pearlmutter, Seidenberg (1994), Trueswell & Tanenhaus (1994), Stevenson (1994)
- MacDonald et al (1994): each analysis has an activation level
  - Total activation is fixed for all analyses
  - Increase in activation of one => decrease in the other

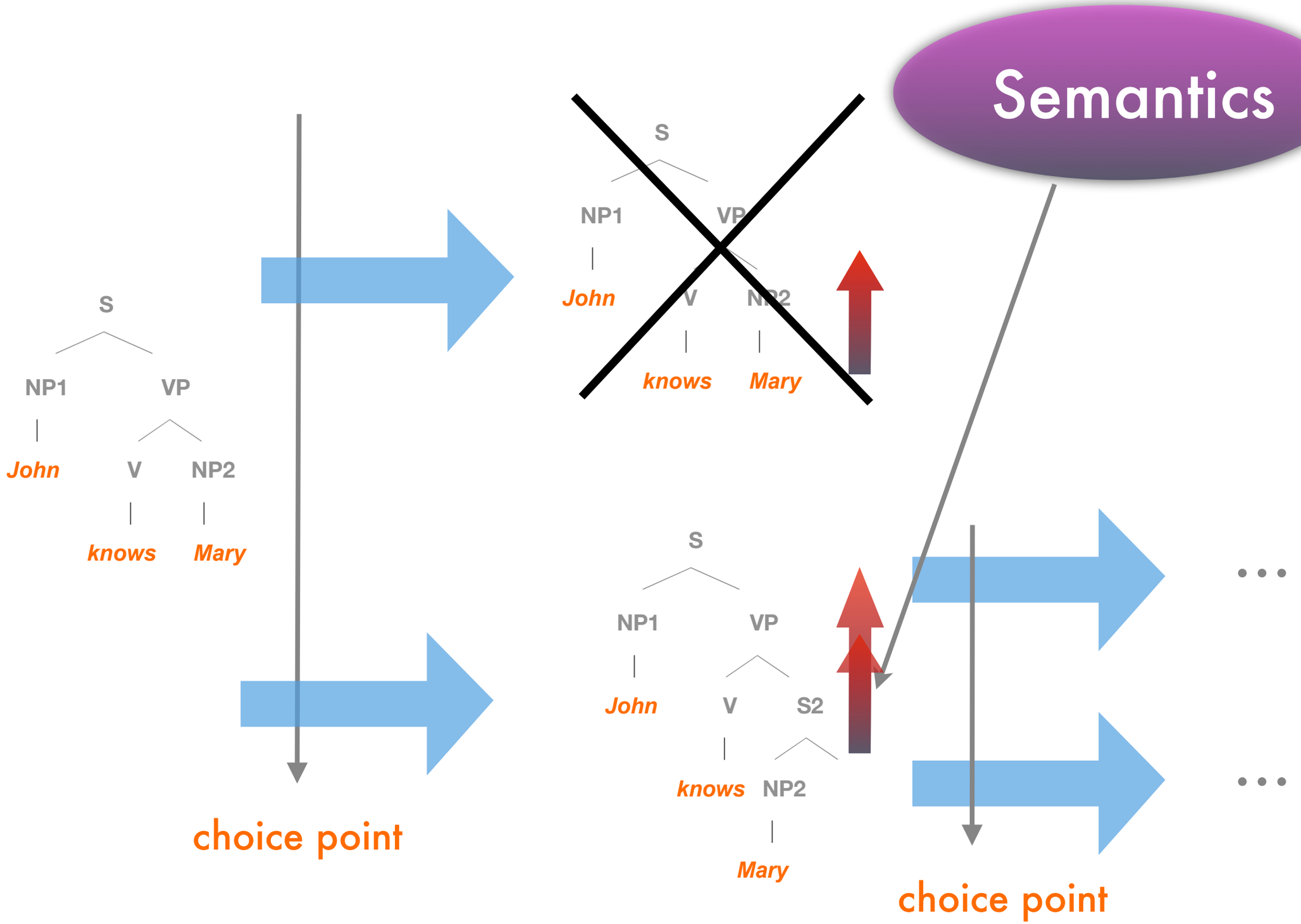
Full Parallelism



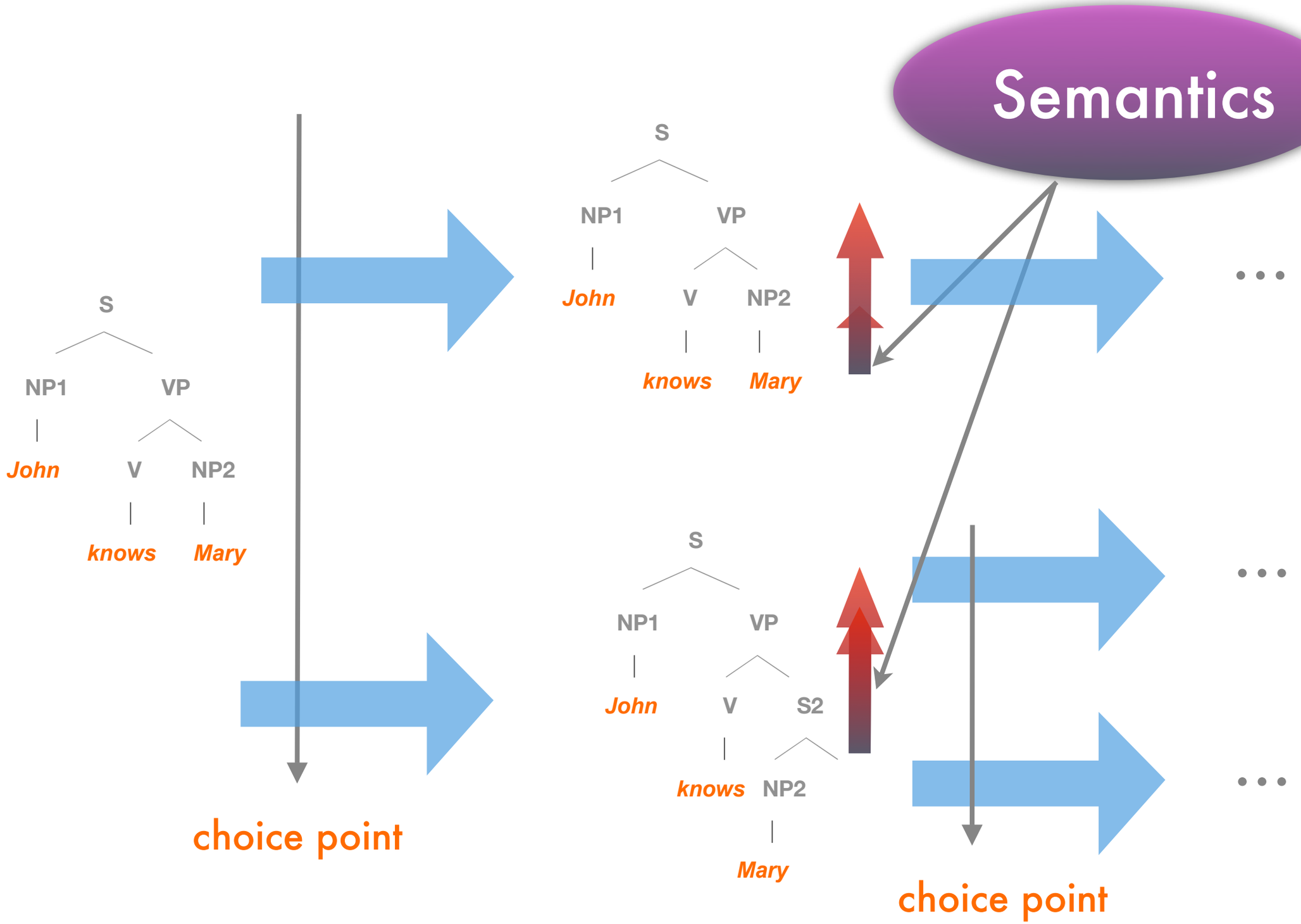
Ranked Parallelism



Momentary Parallelism



Competitive Activation



Modularity vs. Interaction

- Which knowledge source is used when?
- Modular architecture
  - Lexical access precedes parsing, which in turn precedes semantic processing, and so on.
  - E.g., Frazier (1984)
- Interactive architectures
  - A single parsing process combines various sources of knowledge (e.g., lexical, syntactic, semantic...)
  - E.g., Altman (1988), MacDonald et al. (1994)

Mapping to Processing Difficulty

- Consider:

  "The fossil examined ..."

  "The archaeologist examined ..."

- Linking hypotheses?
  - Modular models: the main clause reading is systematically preferred to the reduced-relative.
  - Interactive models: there is no such systematic preference; semantic fit resolves the ambiguity.
  - ➡ Multiple-constraint approach