HOW BIG A SPOON SHOULD SYNTAX USE TO FEED SEMANTICS?

Aravind K. Joshi
University of Pennsylvania
Philadelphia PA
USA

ESSLLI 2008 Workshop: What Syntax Feeds Semantics?
Hamburg, August 13 2008

Outline

- Introduction
- Bigger spoon for CFG– LTAG
  - Derivation Tree and semantics computed from the derivation tree
- Flexible composition, Multicomponent LTAG, making the spoon bigger
    - Some applications
- Bigger spoon for a categorial grammar
- Interaction with discourse
- Summary

Introduction

- Formal systems to specify a grammar formalism
  - Start with primitives (basic primitive structures or building blocks) as simple as possible and then introduce various operations for constructing more complex structures
    - Conventional (mathematical) wisdom
  - Alternatively,

Introduction: CLSG

- Start with complex primitives which directly capture some crucial linguistic properties and then introduce some general operations for operations for composing them
  -- Complicate Locally, Simplify Globally (CLSG)

- CLSG approach is characterized by localizing almost all complexity in the set of primitives, a key property

Introduction: CLSG – localization of complexity

- **Specification of the finite set of complex primitives becomes the main task of a linguistic theory**

- **CLSG pushes all dependencies to become local, i. e. , they arise initially in the primitive structures to start with**

CLSG approach

- CLSG approach has led to several new insights into
  - Syntactic description
  - Semantic composition
  - Language generation
  - Statistical processing, Psycholinguistic properties
  - Discourse structure

Outline

- Introduction
- Bigger spoon for CFG– LTAG
    - Derivation Tree and semantics computed from the derivation tree
- Flexible composition, Multicomponent LTAG, making the spoon bigger
        - Some applications
- Bigger spoon for a categorial grammar
- Interaction with discourse
- Summary

Localization of Dependencies

- agreement: person, number, gender
- subcategorization: sleeps: null; eats: NP; gives: NP NP; thinks: S
- filler-gap: who did John ask Bill to invite e
- word order: within and across clauses as in scrambling and clitic movement
- function – argument: all arguments of the lexical anchor are localized

Localization of Dependencies

- word-clusters (flexible idioms): non-compositional aspect
  - take a walk, give a cold shoulder to
- word co-occurrences
- lexical semantic aspects
- statistical dependencies among heads
-
-

Strong lexicalization of CFG's

Given a CFG, G, we want to construct a grammar G' such that the elementary structures in G' (each associated with a lexical item)

(1) localize the dependencies

(2) structures generated by G' are the same as those generated by G

then it can be shown that the composition operation of substitution alone is not sufficient.

However, adding adjunction as another operation does the trick.

Thus adjunction arises in the process of lexicalizing a CFG!

Surprise: The resulting system is stronger than CFG's both syntactically and semantically

Lexicalized TAG: LTAG

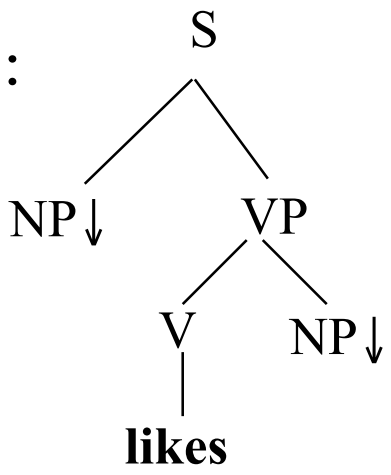
- **Finite set of elementary trees anchored on lexical items**
- **Elementary trees: Initial and Auxiliary**
- **Operations: Substitution and Adjoining**
- **Derivation:**
  - **Derivation Tree**
    - **How elementary trees are put together.**
  - **Derived tree**

LTAG: Some Formal Properties

- TAGs (more precisely, languages of TAGs) belong to the class of languages called mildly context-sensitive languages (MCSL) characterized by
  - polynomial parsing complexity
  - grammars for the languages in this class can characterize a limited set of patterns of nested and crossed dependencies and their combinations
  - languages in this class have the constant growth property, i.e., sentences, if arranged in increasing order of length, grow only by a bounded amount
  - this class properly includes CFLs

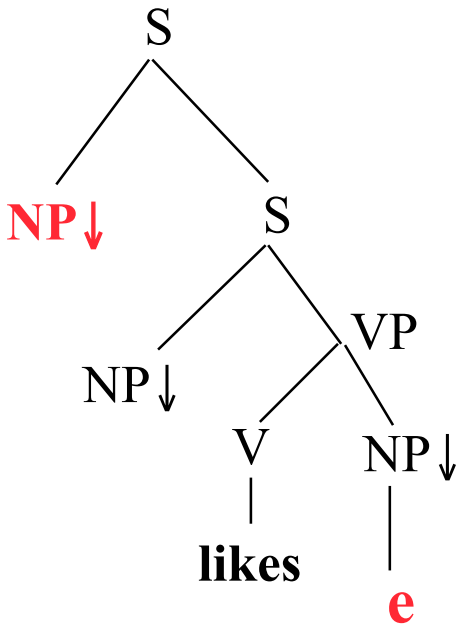
LTAG: Examples

α1:



transitive

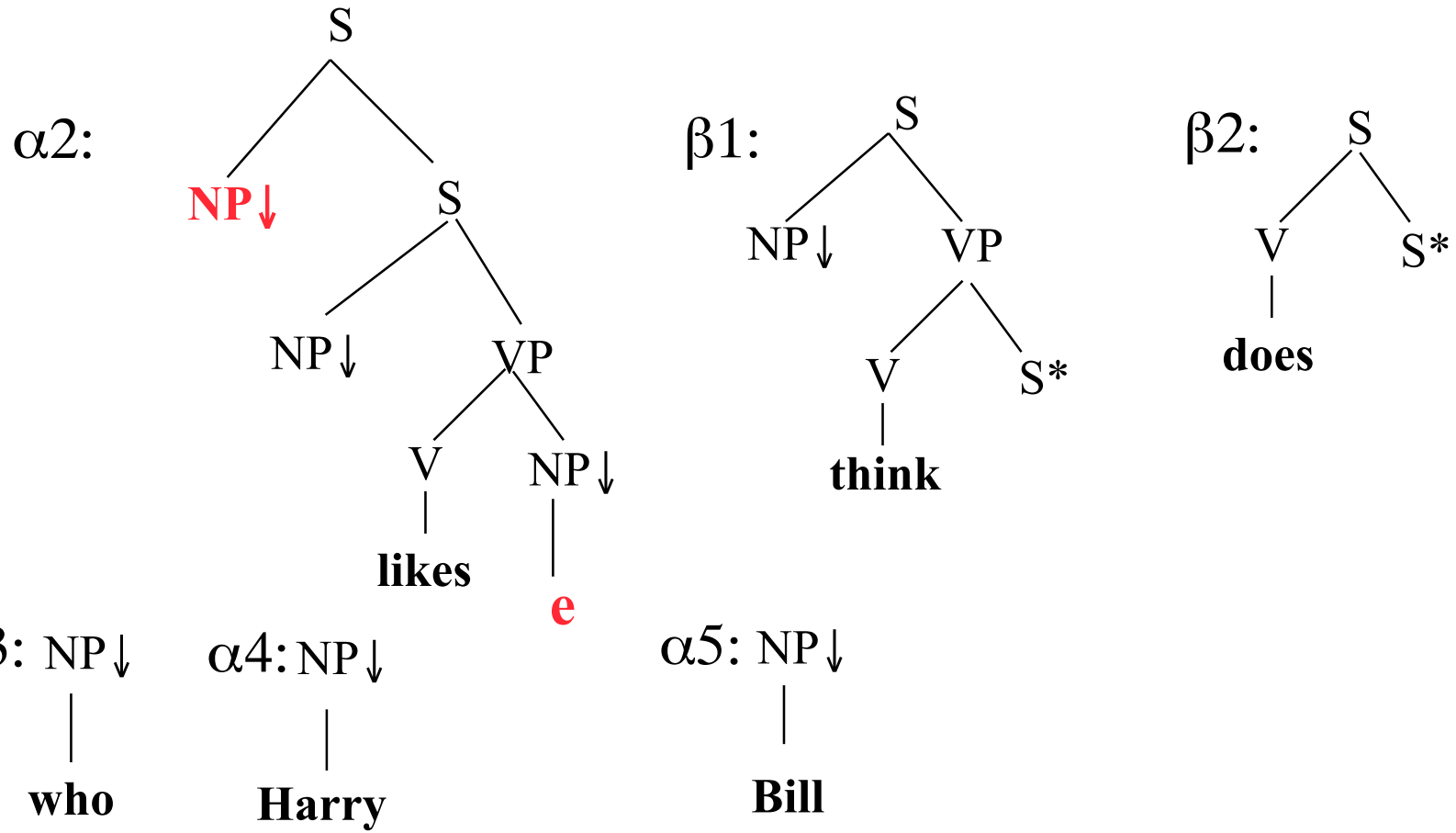
α2:



object extraction

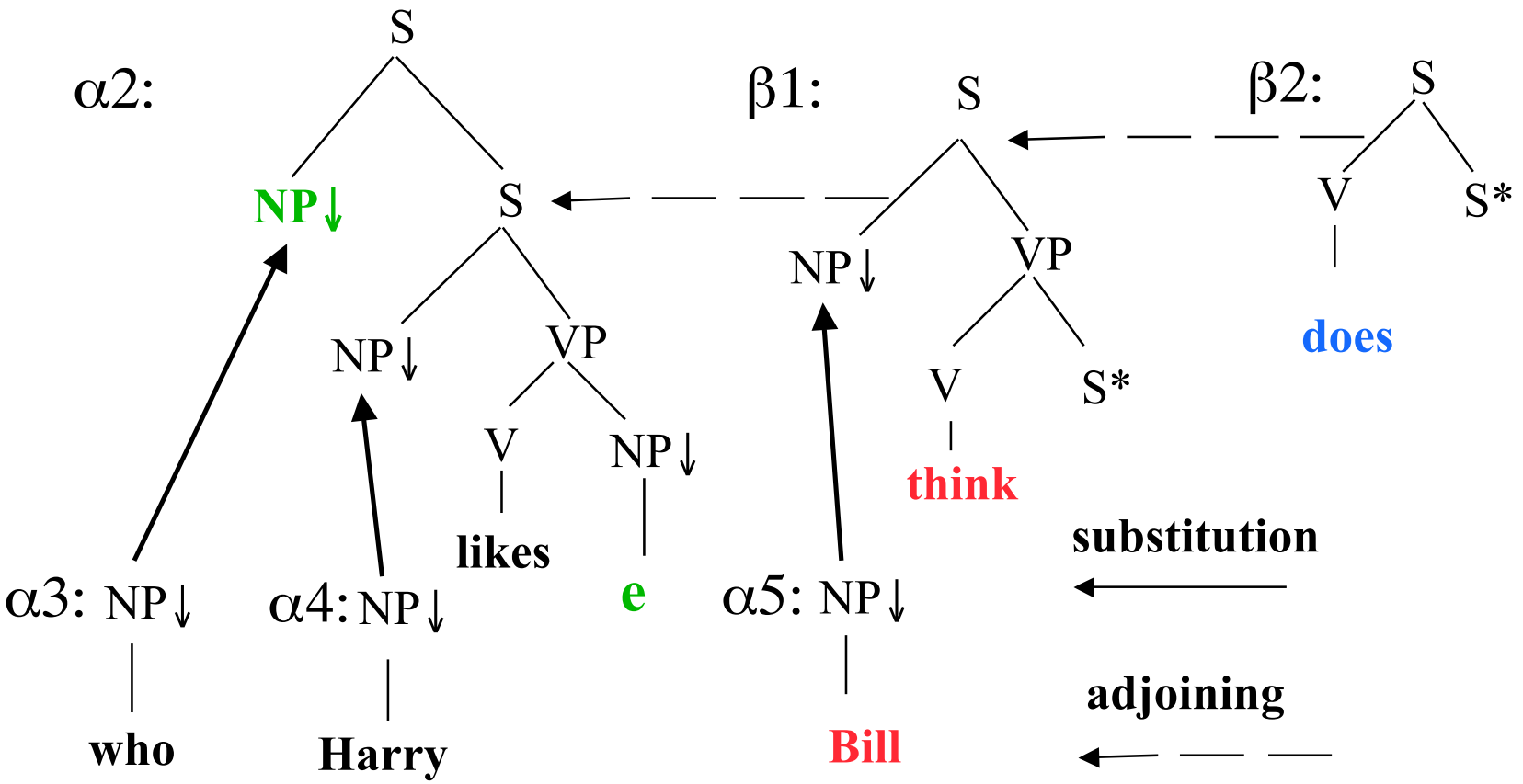
some other trees for 'likes' subject extraction, topicalization, subject relative, object relative, passive, etc.

LTAG: A derivation



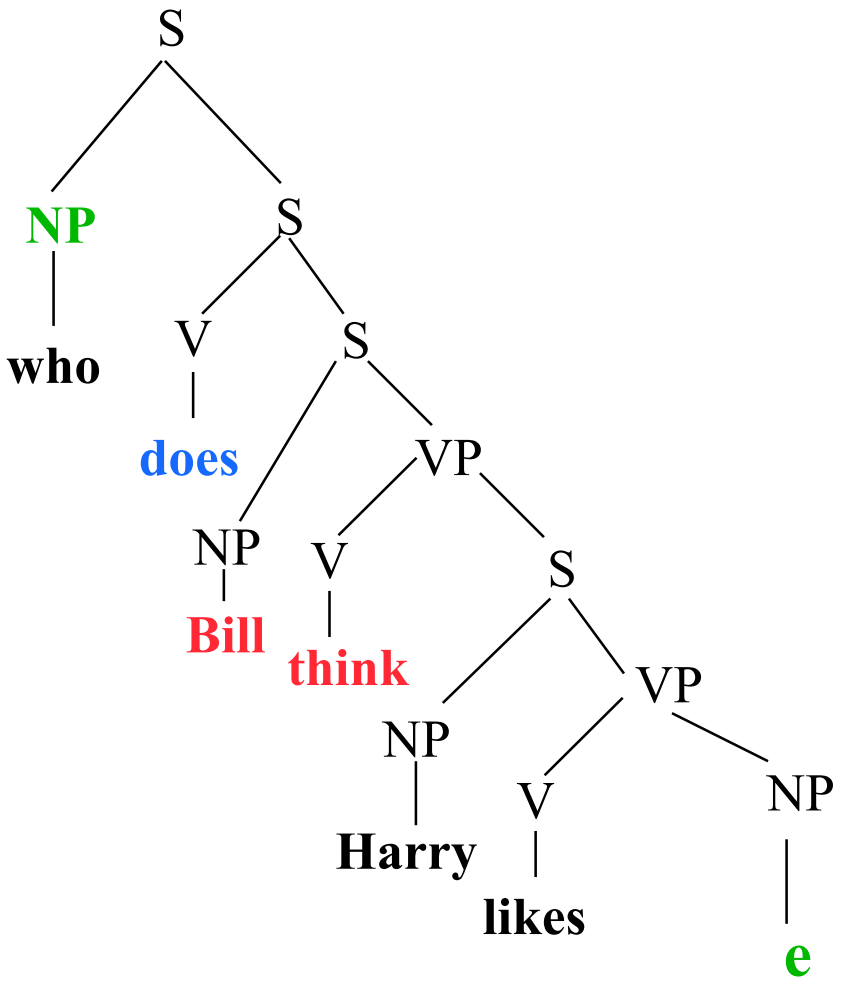
LTAG: A Derivation

who does **Bill think** Harry likes



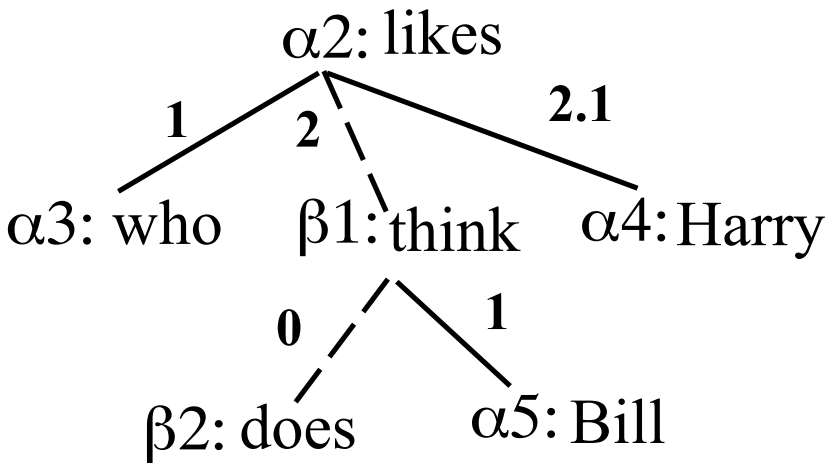
LTAG: Derived Tree

who does **Bill think** Harry likes



LTAG: Derivation Tree

who does Bill think Harry likes

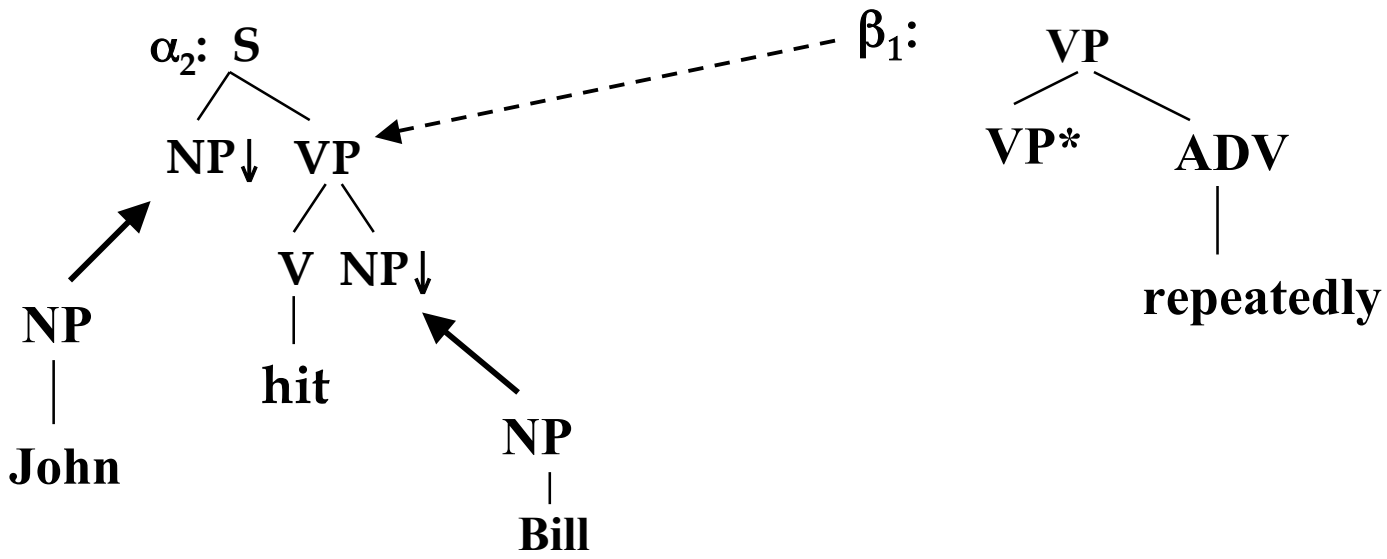


substitution

adjoining

- **Composition by lexical attachments (substitution and adjoining)**
- **The derivation tree shows what attaches to what and where**
- **Semantics to be defined on the derivation tree**
  - **-- need for additional information?**
- **Order of traversal of the nodes**

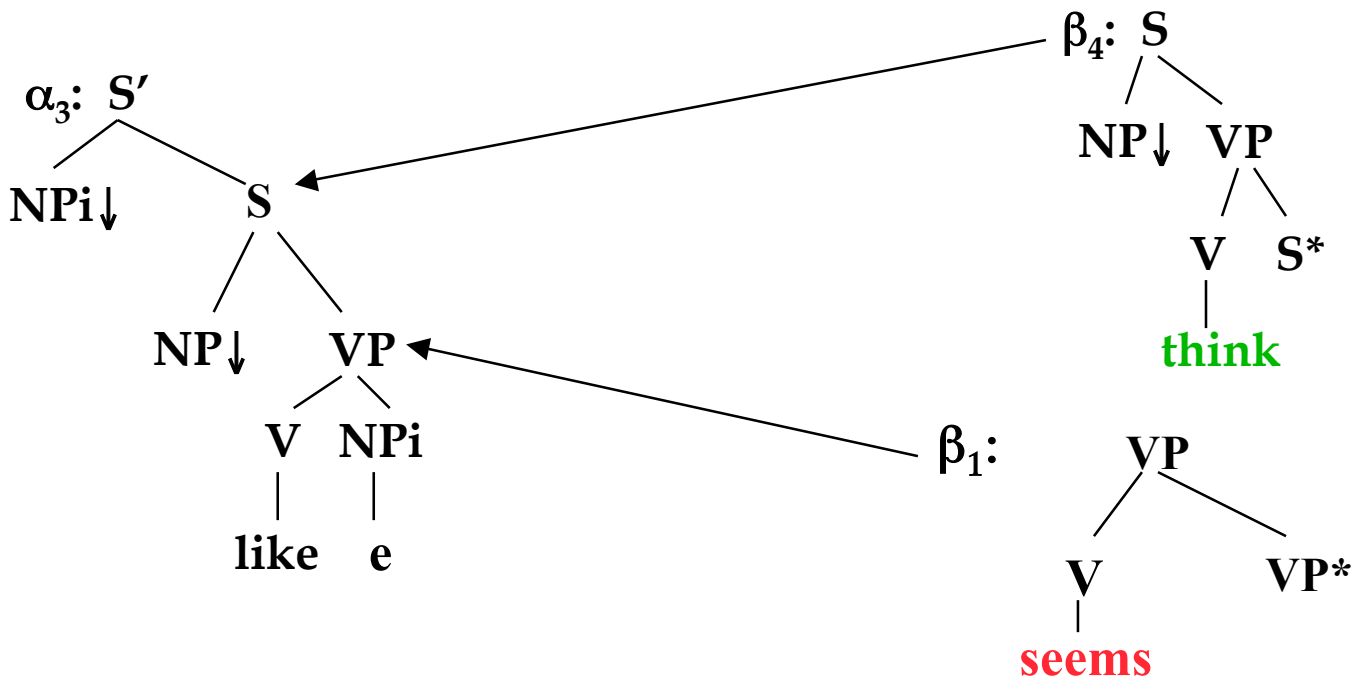
Composition defined by the derivation tree

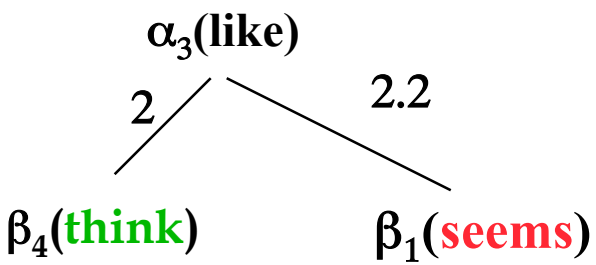


about: s_2
John: x_1
Bill: x_2
hit(s_1, x_1 , x_2)
repeatedly(s_2, s_1)

Attachments along the trunk (path from root to lexical anchor)

(who do you think John seems to like)





In the derivation tree seems and think are adjoined along the trunk

-- uniform convention for scoping--lower nodes before higher nodes along the trunk

Additional information on the derivation tree: Some alternatives

- Additional links
- Adding features
- Extend the use of the addresses in the derivation tree by adopting a uniform order of traversal of the tree
  -- post order traversal

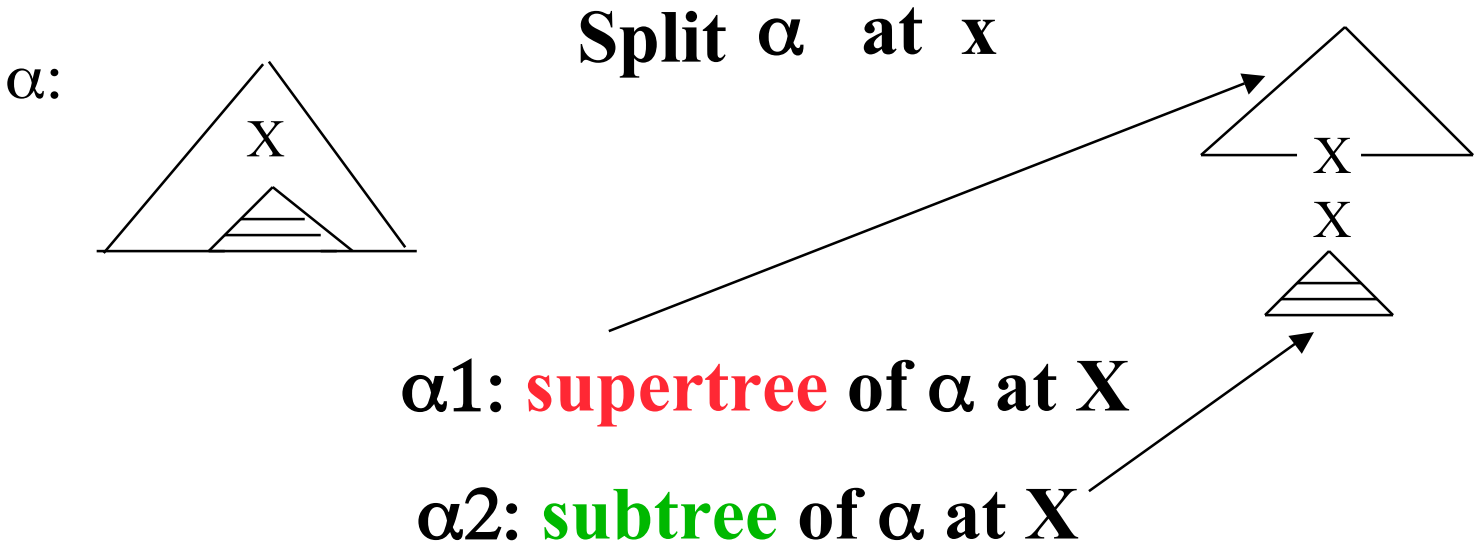
Joshi and Vijayshanker, 1999, Frank and van Genbirth, 2001, Kallmeyer and Joshi, 2003, Joshi, Kallmeyer and Romero, 2003, Gardent and Kallmeyer 2004, Kallmeyer and Romero, 2004, Kallmeyer and Romero, 2008, …

Outline

- Introduction
- Bigger spoon for CFG– LTAG
  - Derivation Tree and semantics computed from the derivation tree
- Flexible composition, Multicomponent LTAG, making the spoon bigger
  - Some applications
- Bigger spoon for a categorial grammar
- Interaction with discourse
- Summary

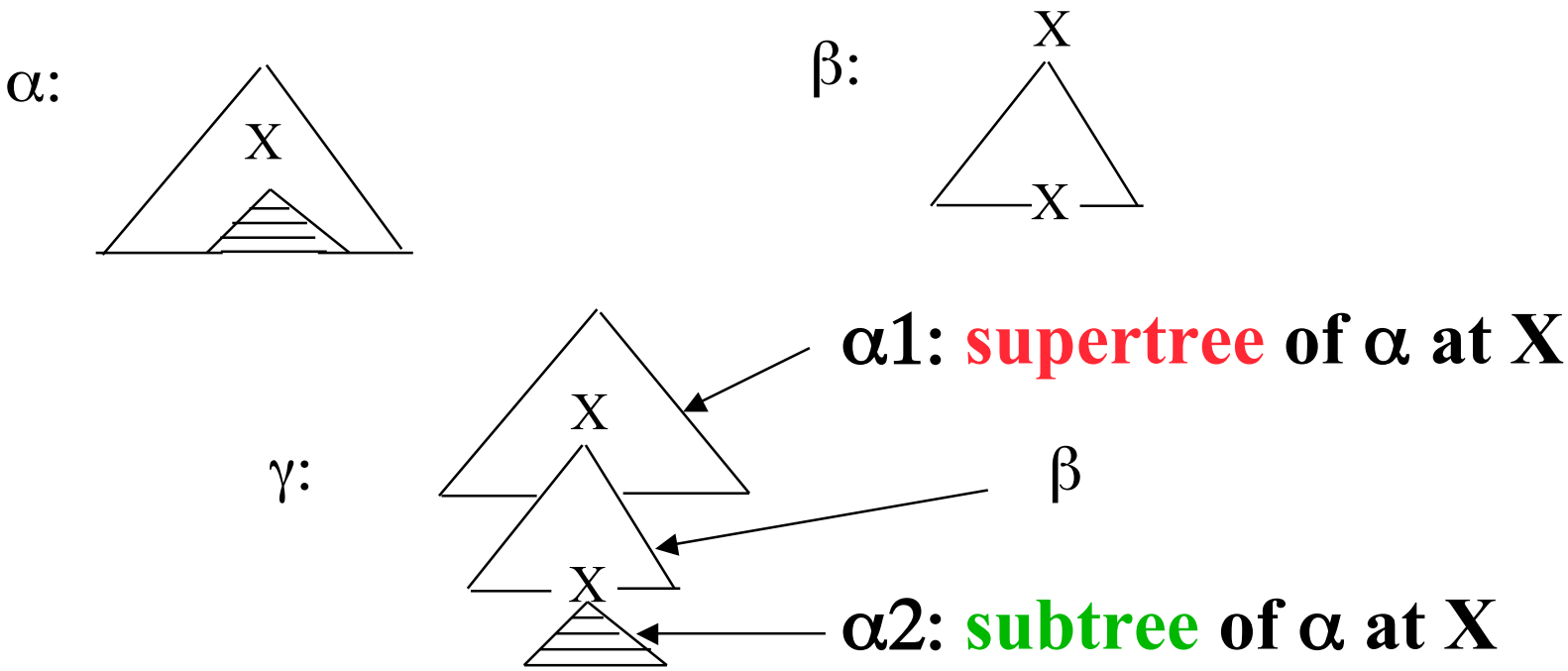
Flexible Composition

Adjoining as Wrapping



Flexible Composition

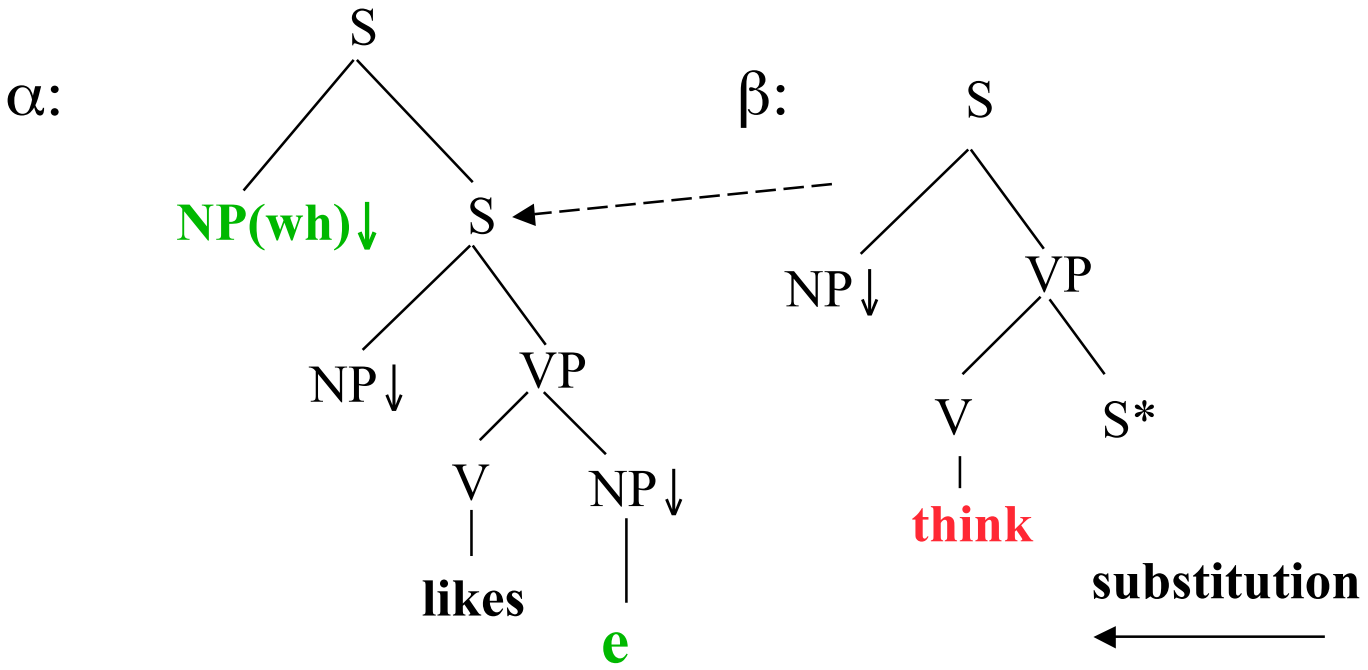
Adjoining as Wrapping



α wrapped around β i.e., the two components α1 and α2 are wrapped around β

Flexible Composition

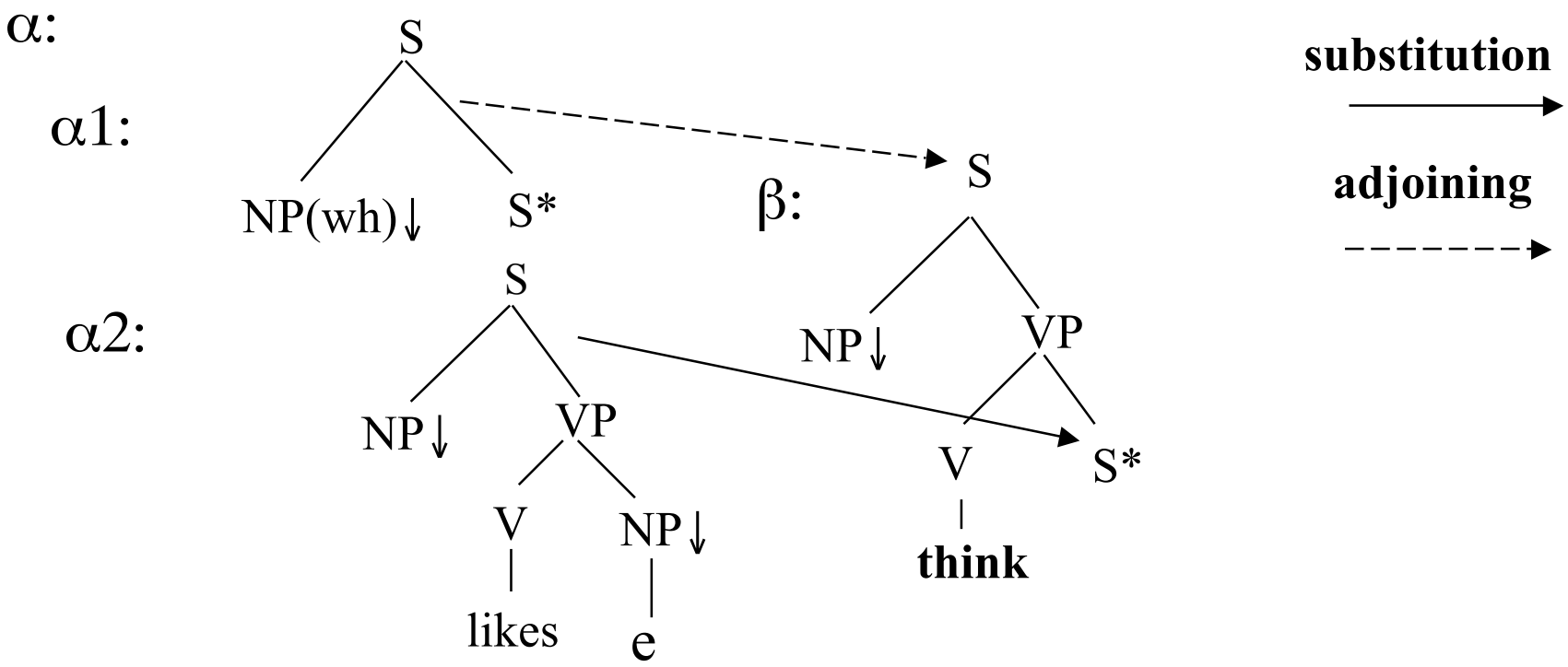
Wrapping as substitutions and adjunctions



- We can also view this composition as
  α wrapped around β
- Flexible composition

Flexible Composition

Wrapping as adjunction and reverse adjunction

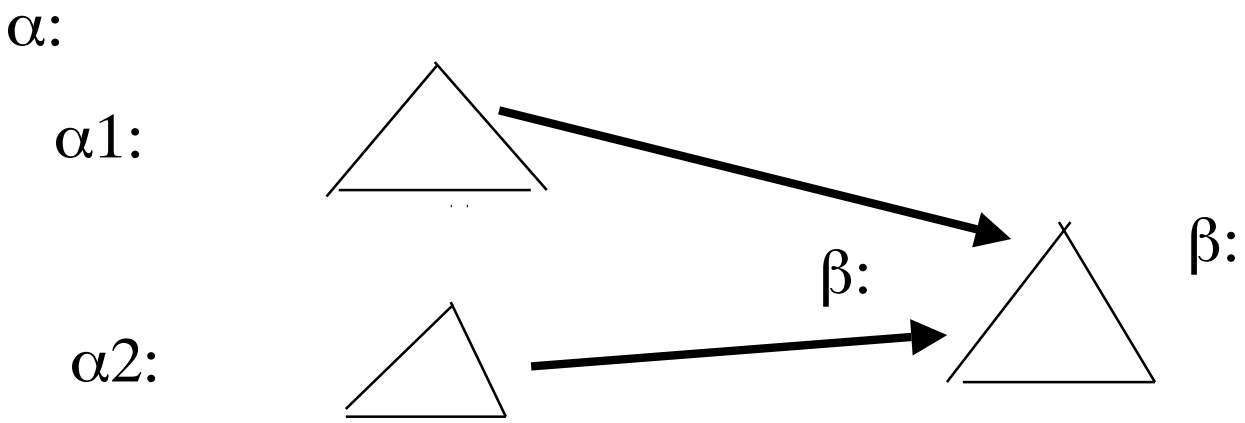


α1 and α2 are the two components of α
α1 attached (adjoined) to the root node S of β
α2 attached (reverse adjoined) at the foot node S of β

Leads to multi-component TAG (MC-TAG)

Multi-component LTAG (MC-LTAG)

(Making the spoon bigger)



The two components are used together in one composition step. Both components attach to nodes in β, an elementary tree. This preserves locality.

The representation can be used for both

-- predicate-argument relationships

-- non-p/a information such as scope, focus, etc.

Multicomponent LTAG (MC-LTAG)

Generalizing on the adjoining as wrapping perspective leads to MC-LTAG.

- **A lexical item may be associated with a finite set of trees, each tree in the set is a component**
- **Set of components together provides an extended domain of locality**
  - **The set of components together define one elementary object**
  - **The components are used together in one composition step with the individual components being composed by attachments**

Multicomponent LTAG (MC-LTAG)

- The representation can be used for both
  - -- predicate argument relationships
  - -- scoping information
- The two pieces of information are together before the single composition step
- However, after the composition there may be intervening material between the components

Tree-Local Multi-component LTAG (MC-LTAG)

- How can the components of MC-LTAG compose preserving locality of LTAG
- Tree-Local MC-LTAG
  - -- Components of a set compose only with an elementary tree or an elementary component
- Tree-Local MC-LTAGs are weakly equivalent to LTAGs
- However, Tree-Local MC-LTAGs provide structural descriptions not obtainable by LTAGs
- Increased strong generative power; hence supporting more semantics

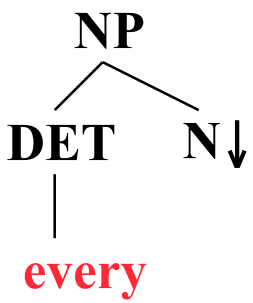
Scope ambiguities: Example

(every student hates some course)

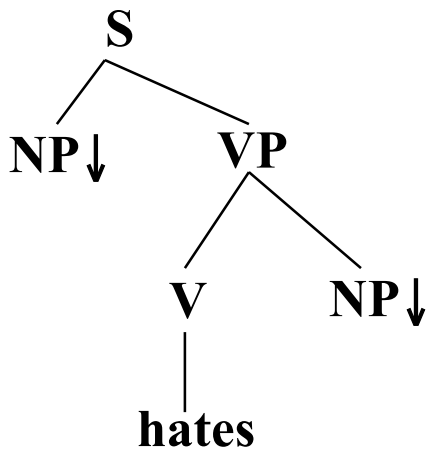
α1:

α11: S*

α12:



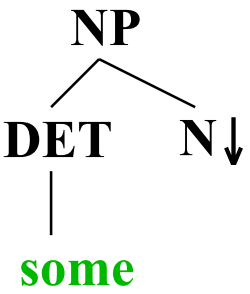
α3:



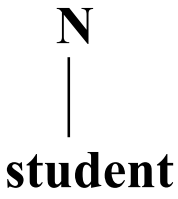
α2:

α21: S*

α22:



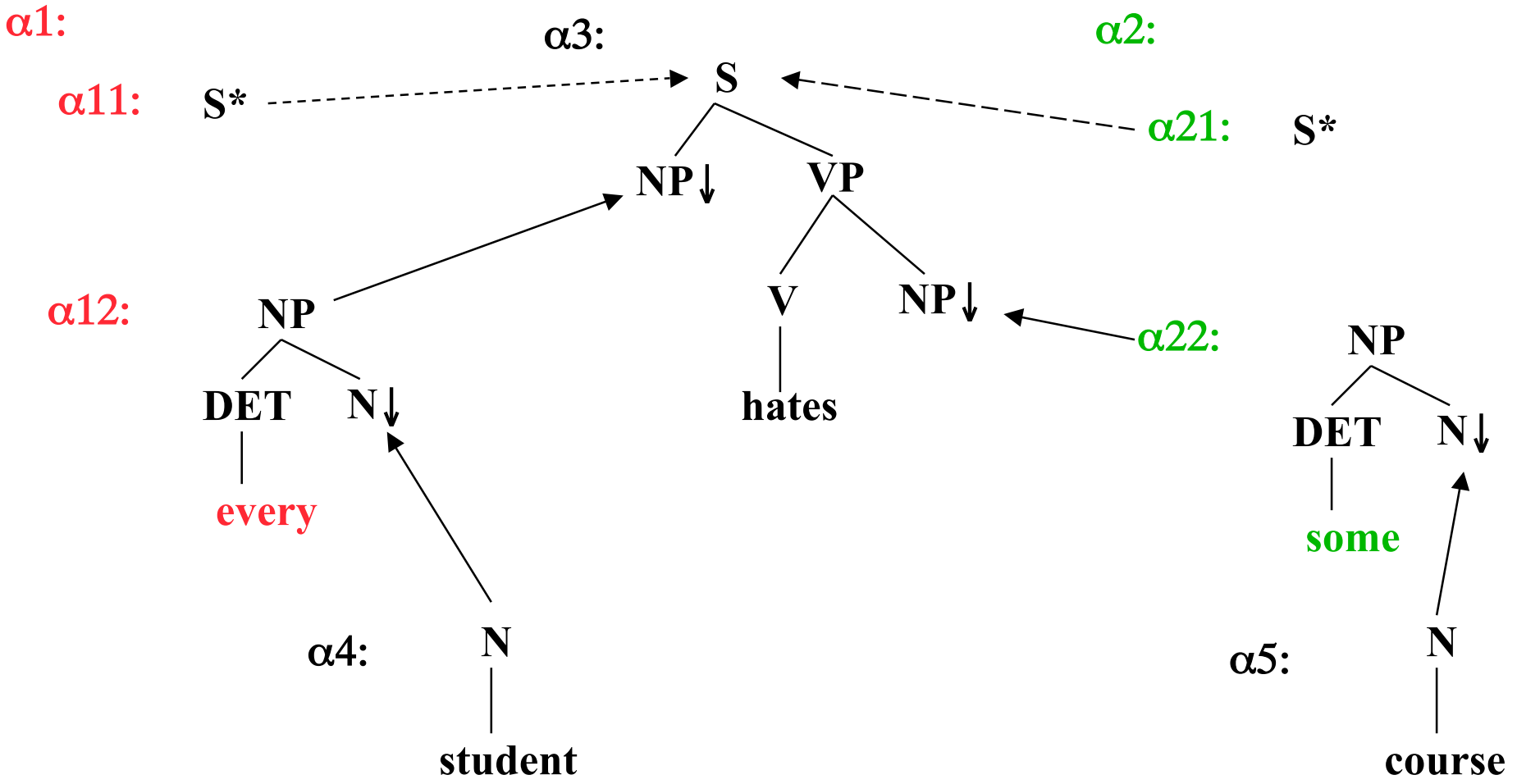
α4:



α5:

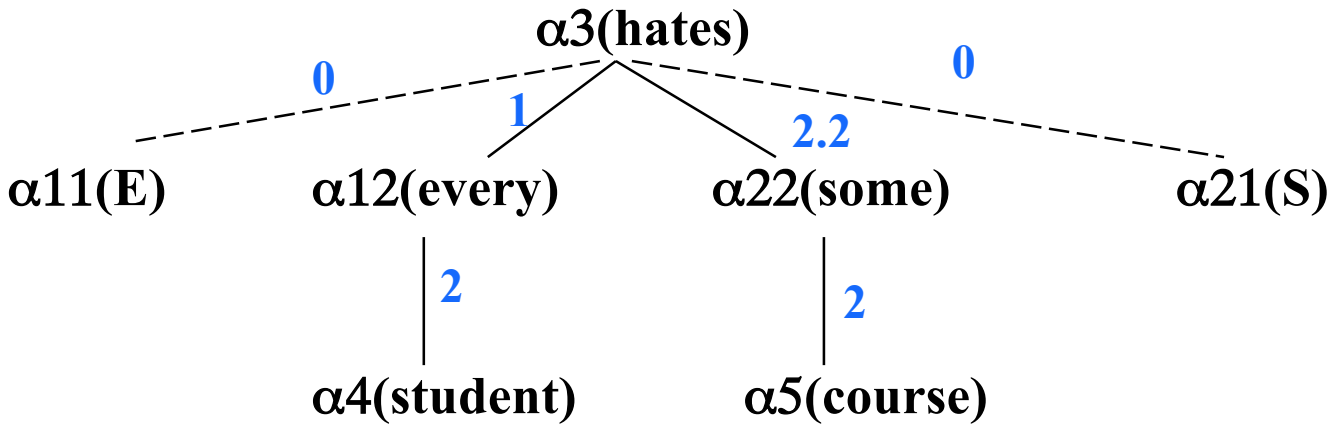
Derivation with scope information: Example

(every student hates some course)



Derivation tree with scope information: Example

(every student hates some course)



- α11 and α21 are both adjoined at the root of α3(hates)
  - multiple adjunctions at the same node

- They can be adjoined in any order, thus representing the two scope readings (underspecified representation)
- The scope readings represented in the LTAG derivation itself

Patterns of scope orderings

- **Adding features**

Kallmeyer and Romero, 2004,
Kallmeyer and Romero, 2008, …

Extend the use of the addresses in the derivation tree by adopting a uniform order of traversal of the tree
-- post order traversal

Other uses of tree-local MC-TAG

- **Misplaced adjectives**
- **Parentheticals**
- **Scrambling patterns**
- **Clitic movement**
-
-
-

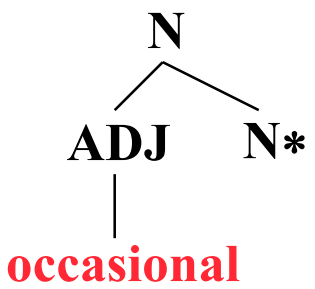
Misplaced Adjectives

(1) **An occasional sailor passed by**

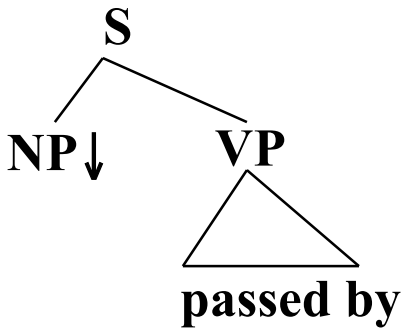
α1:

α11: S*

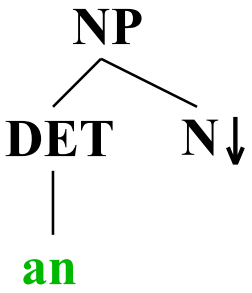
α12:



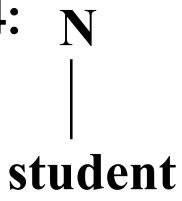
α2:



α3:

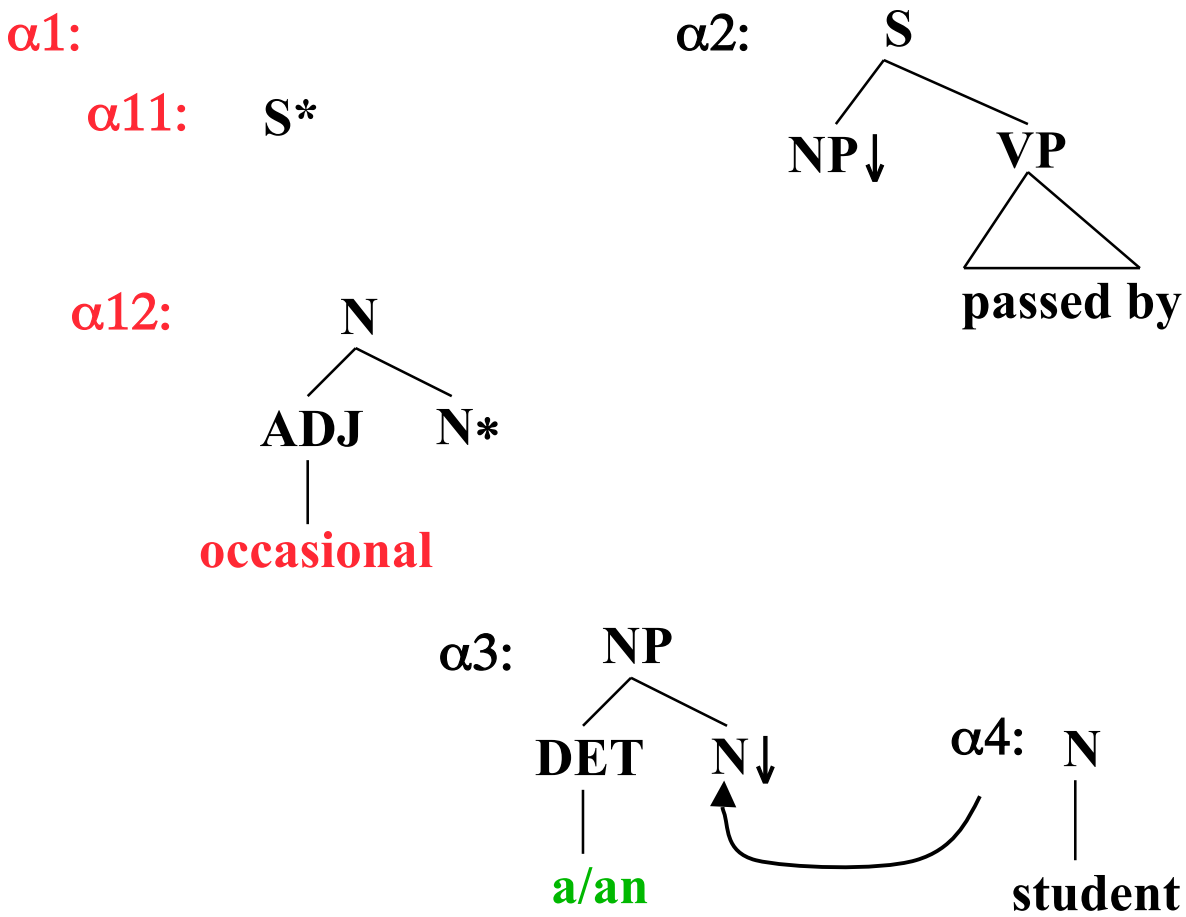


α4:



Misplaced Adjectives

(1) **An occasional sailor passed by**



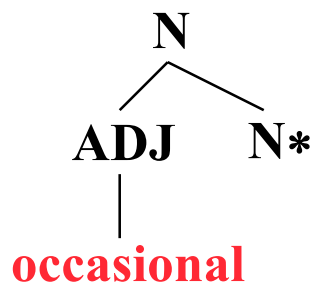
Misplaced Adjectives

(1) **An occasional sailor passed by**

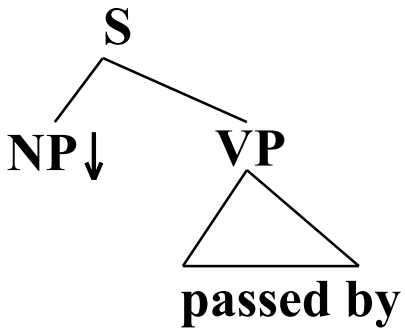
α1:

α11: S*

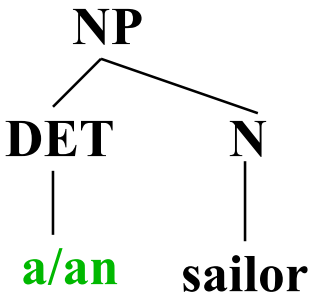
α12:



α2:

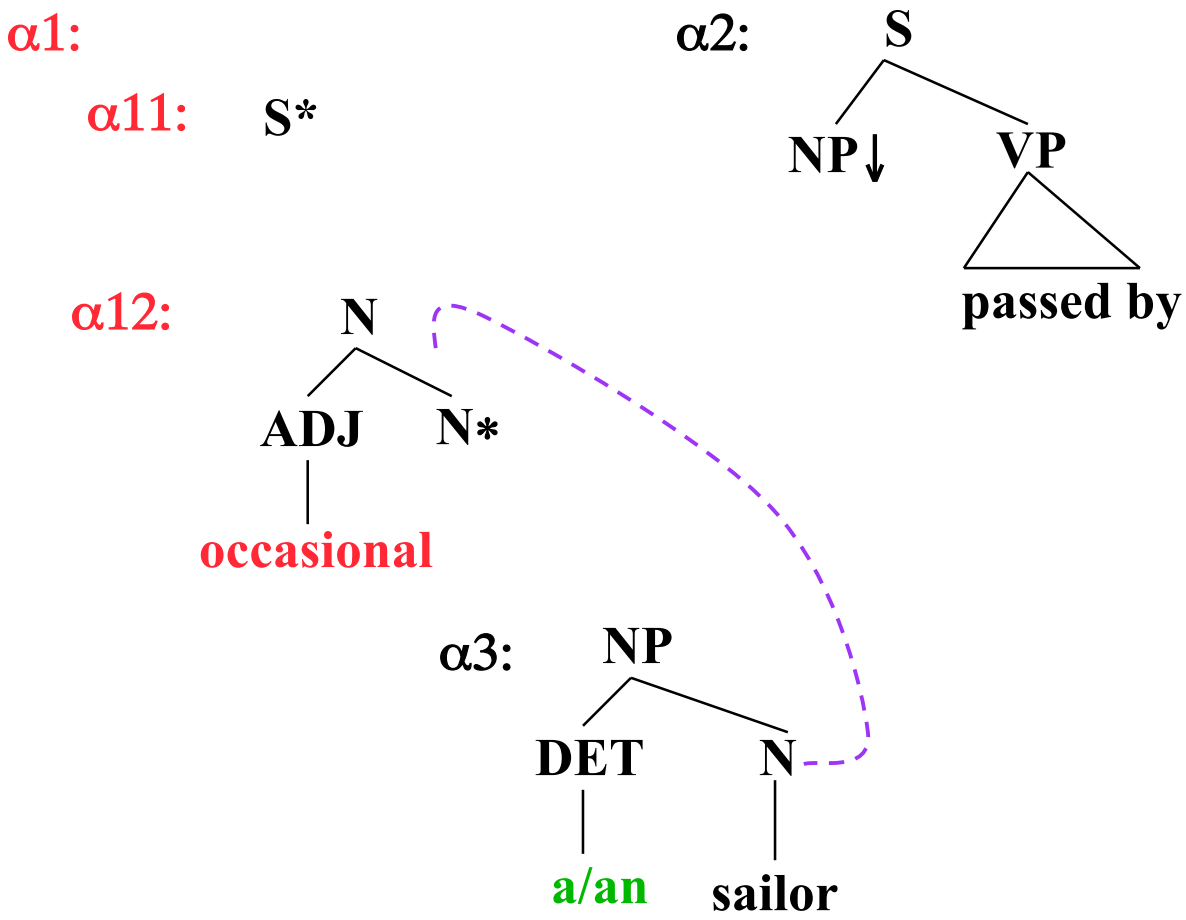


α3:



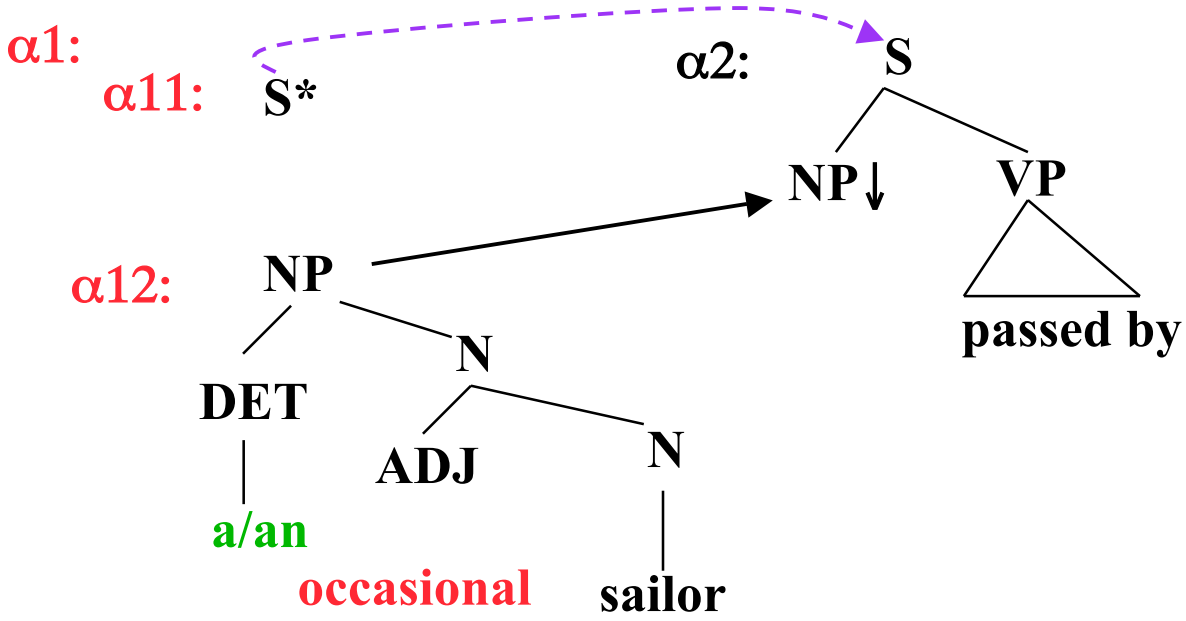
Misplaced Adjectives

(1) **An occasional sailor passed by**



Misplaced Adjectives

(1) An occasional sailor passed by



Parentheticals

(2) Hillary, Obama thinks, will win the primary

An extension of tree local MC-TAG is required-- sister adjoining, which was developed by David Chiang (2000) for another purpose. With this extension we still have the weak equivalence with the standard TAG

Scrambling Patterns

- **Embedding of complement clauses in German**

(1) Hans$_1$ Peter$_2$ Marie$_3$ schwimmen$_3$ lassen$_2$ sah$_1$

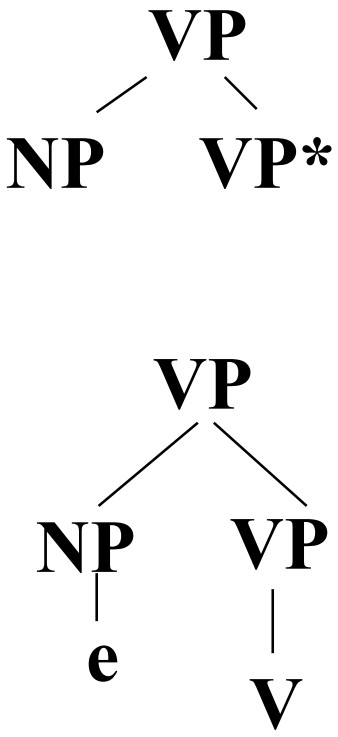
NP1 NP2 NP3 V3 V2 V1

(Hans saw Peter let/make Marie swim)

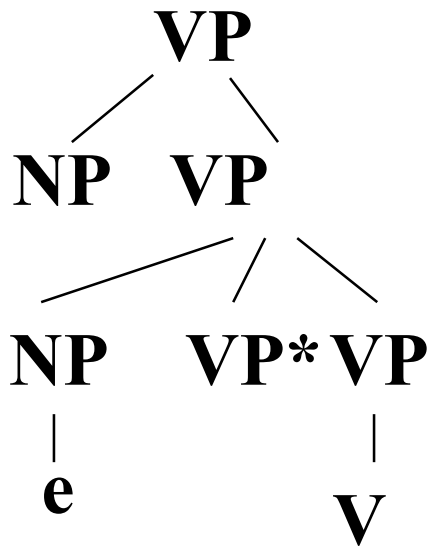
Scrambled versions of (1) permuting the NP's and keeping the order of V's fixed as in (1)

(*Proper names, instead of full NPs are used for convenience*)

Elementary Trees for a Scrambled Argument

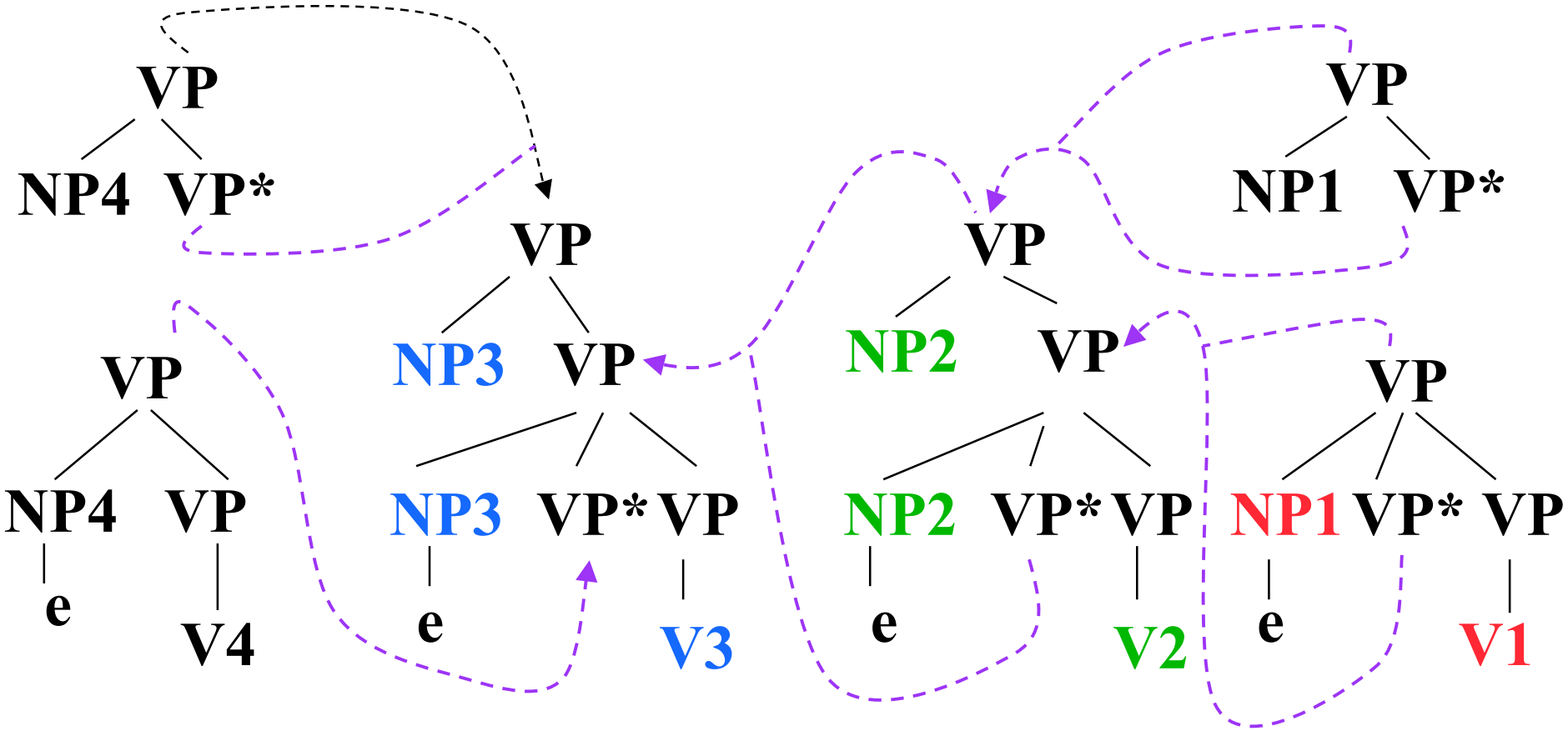


Multi-component Tree (domination constraint)



Standard single component tree

Scrambling: NP4 NP3 NP1 NP2 V4 V3 V2 V1



Clitic Climbing

- **Clitic placement can also be viewed as a word-order variation and described by using MC-TAG as in scrambling**

Bleam 1998, 2002, Chen-Main. 2007

Outline

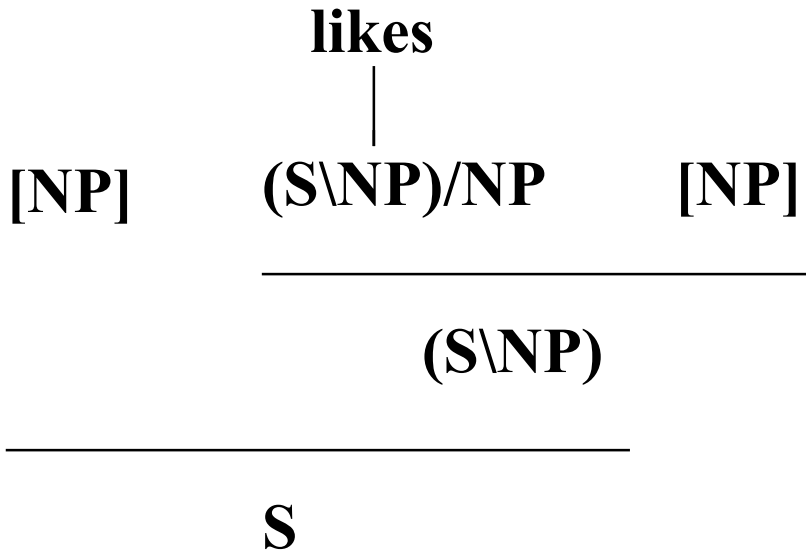
- Introduction
- Bigger spoon for CFG– LTAG
    - Derivation Tree and semantics computed from the derivation tree
- Flexible composition, Multicomponent LTAG, making the spoon bigger
        - Some applications
- Bigger spoon for a categorial grammar
- Interaction with discourse
- Summary

Bigger Spoon for Categorial Grammar

Partial proof trees as building blocks for a categorial grammar, Joshi and Kulick, Linguistics and Philosophy, 20, 1997

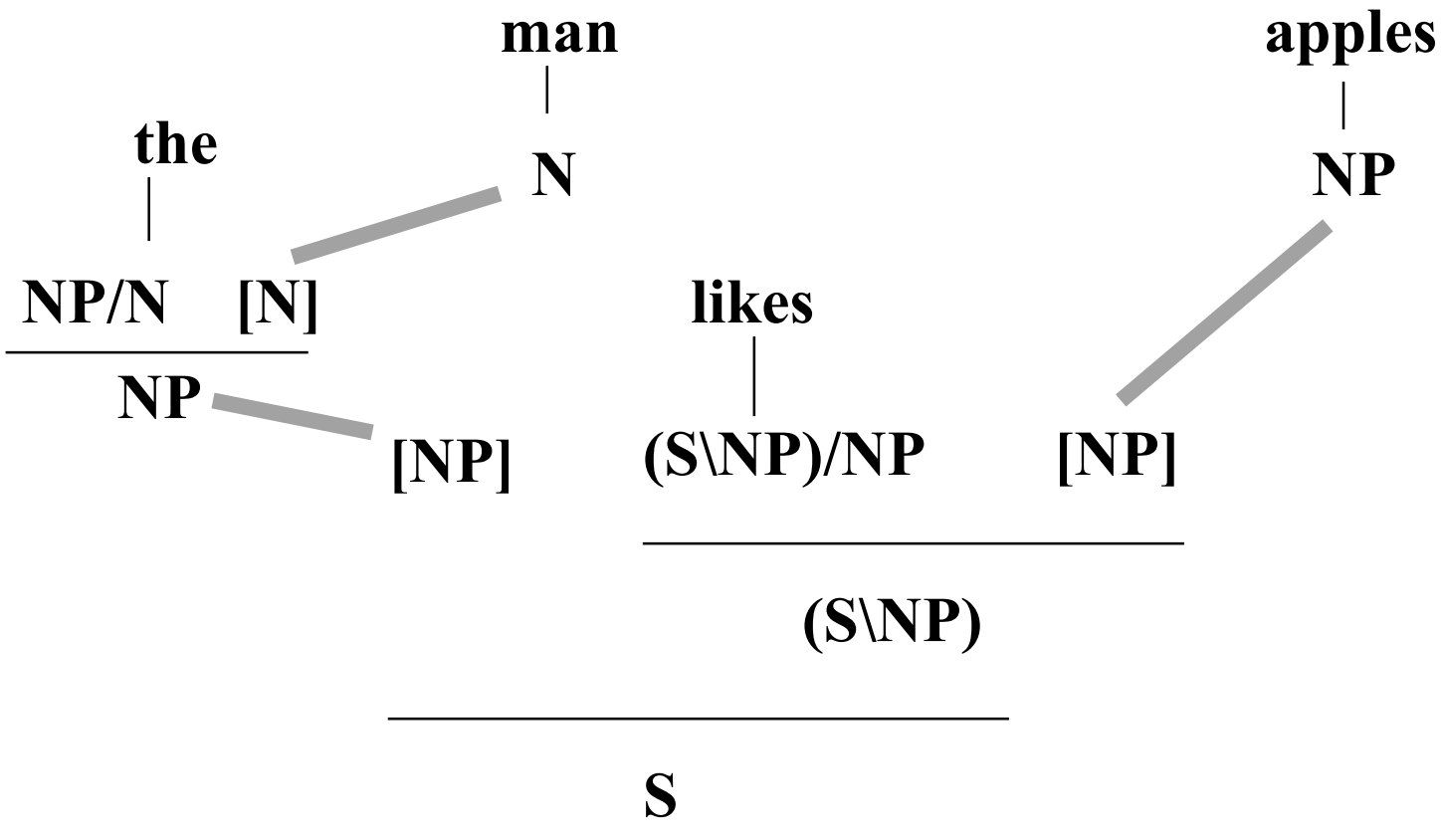
Partial proof trees, hybrid logic, and quantifier scope, Joshi, Kulick, and Kurtonina, ESSLLI 1999, Utrecht

Bigger Spoon for Categorial Grammar

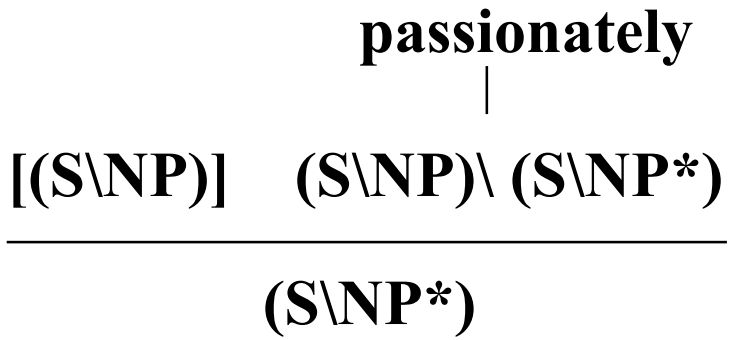


- Each lexical item is associated with one or more (basic) partial proof trees (PPT) obtained by unfolding arguments.
- (PPT) is a finite set -- the set of basic types.
- Informal description of the inference rule -- linking

Linking conclusion nodes to assumption nodes: an inference rule, stated informally



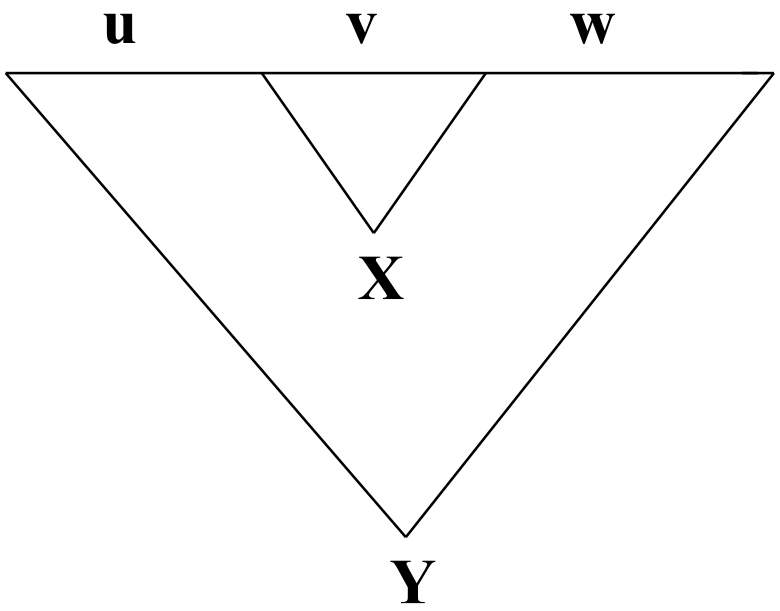
Bigger spoon



- No unfolding past an argument marked by *
- Thus unfolded arguments are only those which are the arguments of the lexical item.

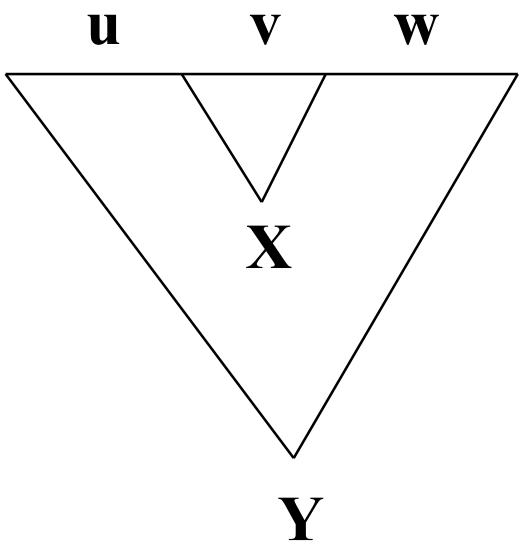
Stretching and linking – an inference rule

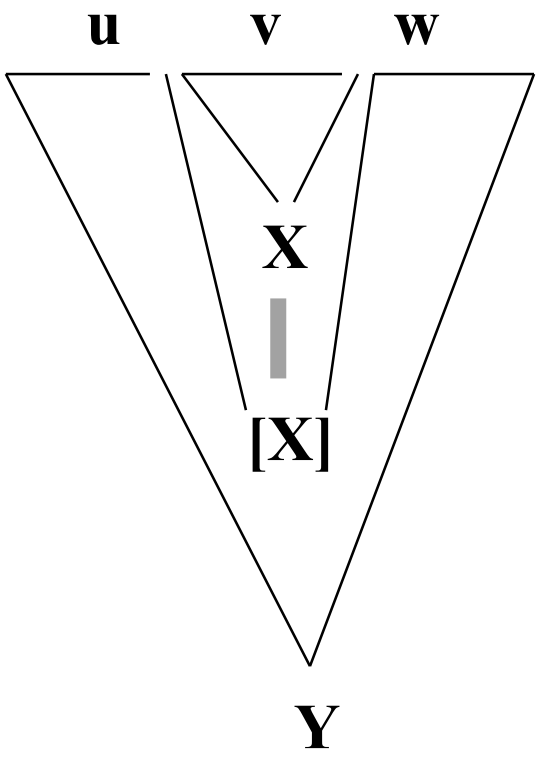
A proof tree can be *stretched* at any node.



A proof tree to be stretched at the node X.

Stretching a proof tree at node X





X is the conclusion from v

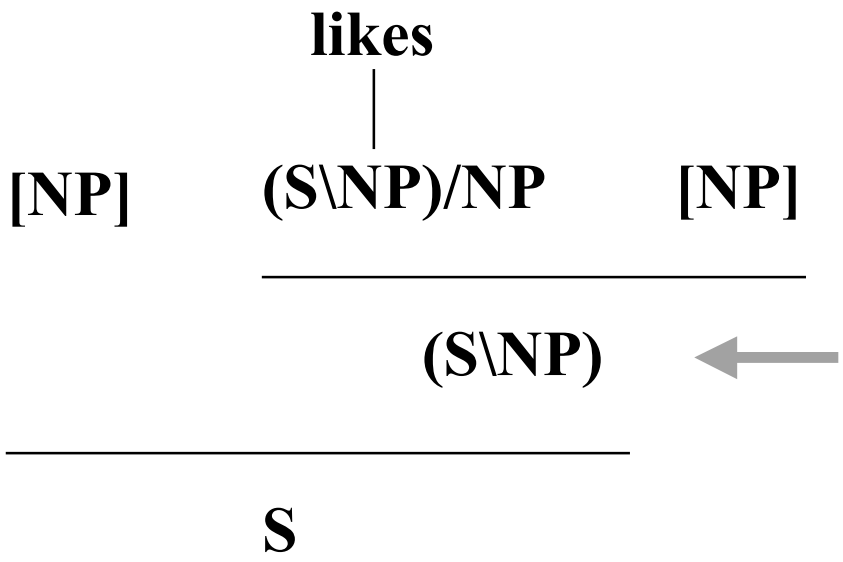
Y is the conclusion from

u [X] w

i.e., from u, assumption X and w

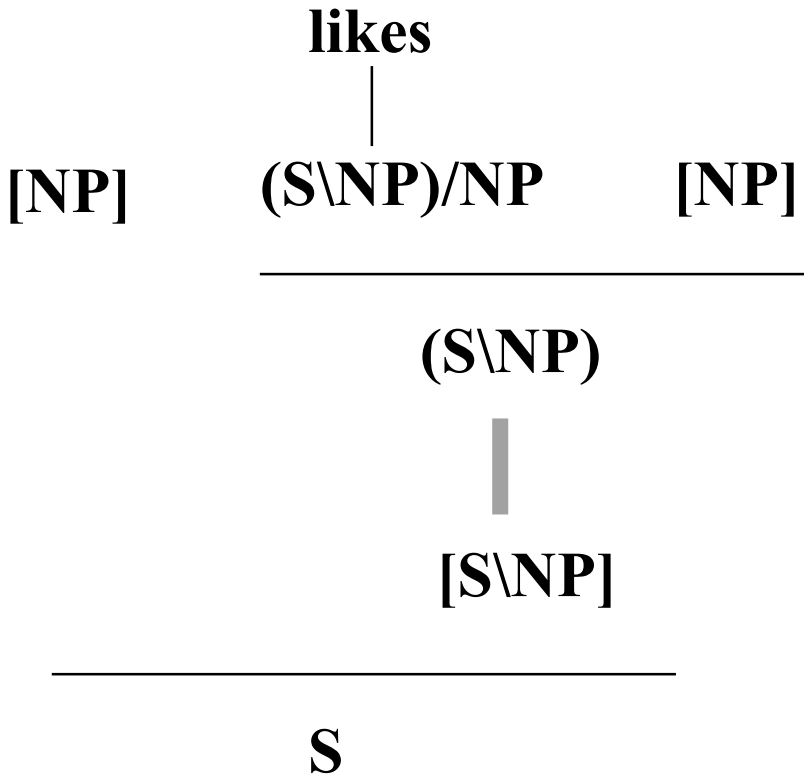
Linking X to [X] we have the original proof tree.

Stretching and linking -- an example

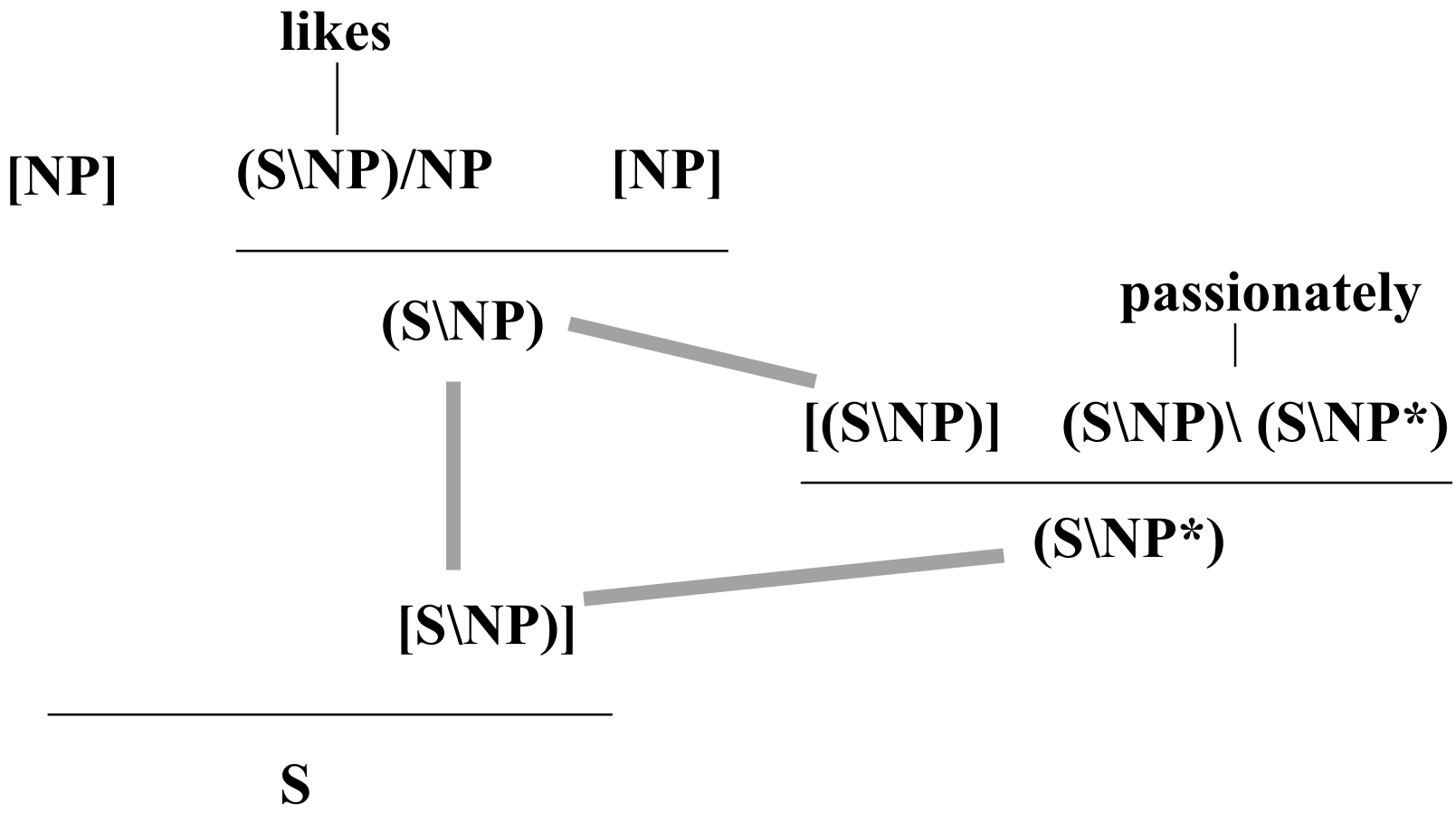


Stretching at the indicated node

Stretching and linking -- an example

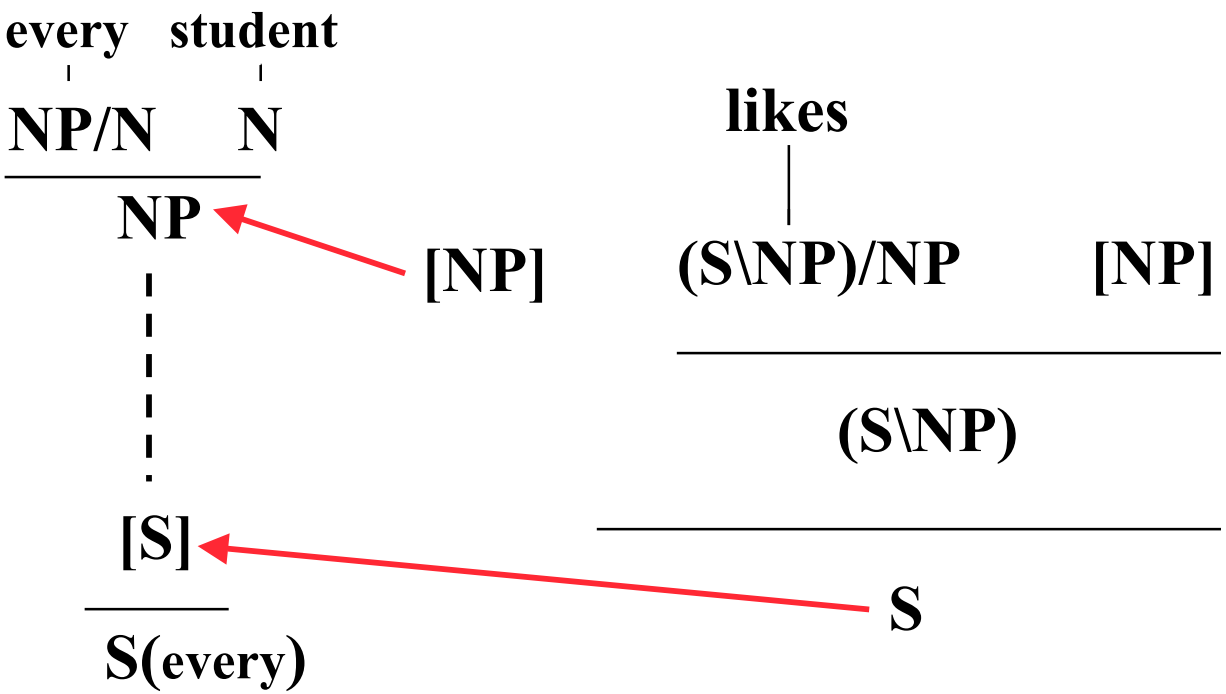


Stretching and linking -- an example

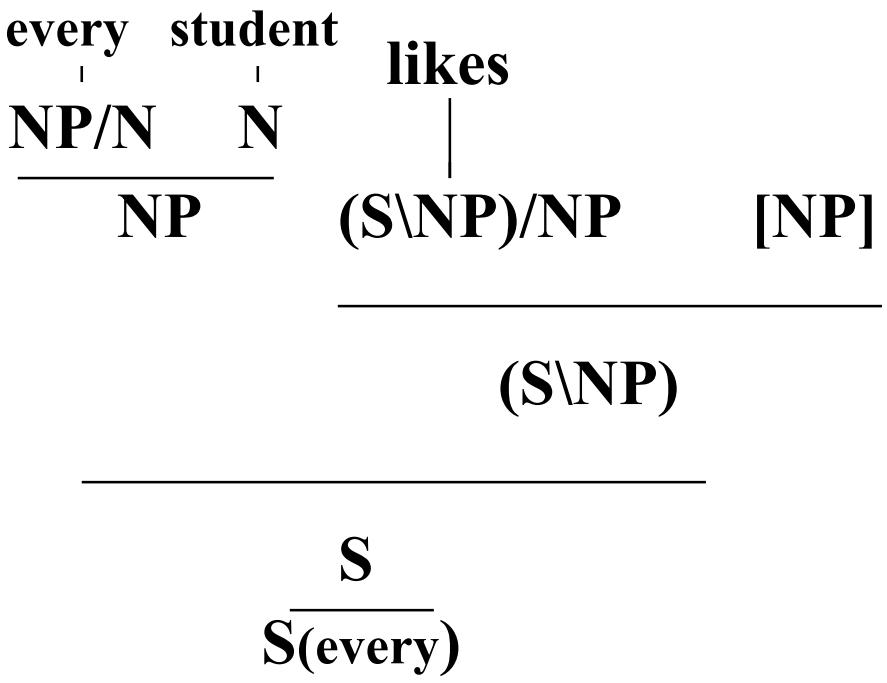


John likes apples passionately

Bigger Spoon for Categorial Grammar

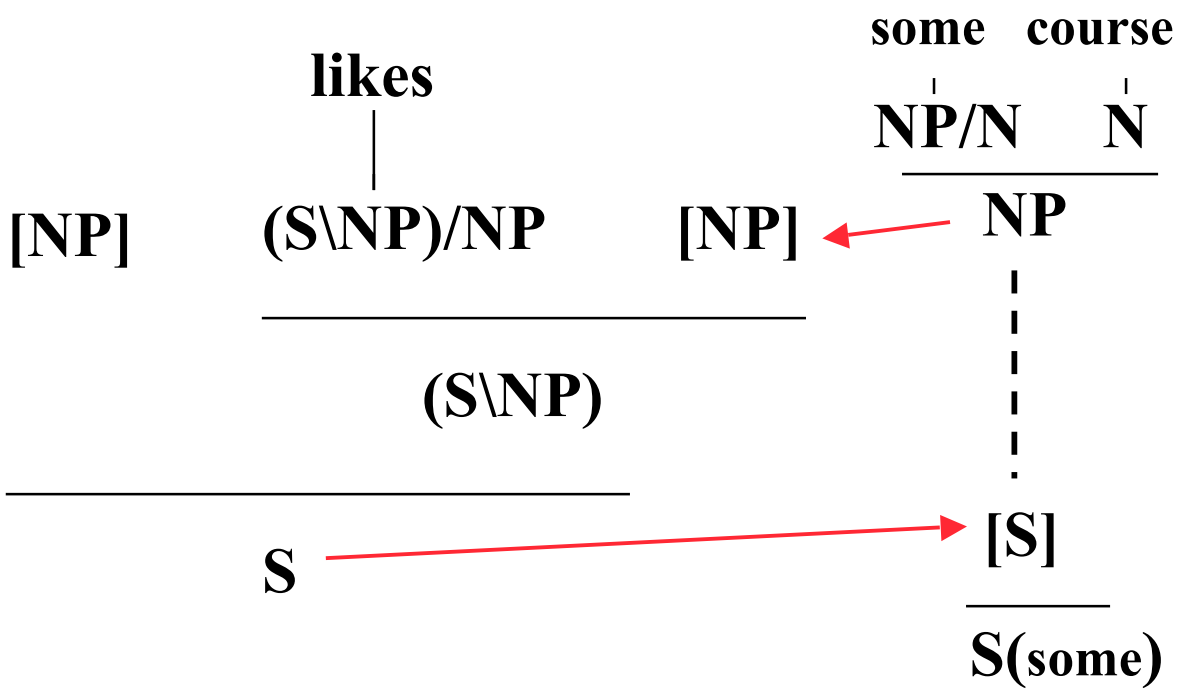


Bigger Spoon for Categorial Grammar



some > every

Bigger Spoon for Categorial Grammar



every > some

Outline

- Introduction
- Bigger spoon for CFG– LTAG
  - Derivation Tree and semantics computed from the derivation tree
- Flexible composition, Multicomponent LTAG, making the spoon bigger
  - Some applications
- Bigger spoon for a categorial grammar
- Interaction with discourse
- Summary

Interaction with Discourse

- Sometimes syntax should hold the spoon back form semantics for a while
  - --Avoid delivering a complete structure even when there is no ambiguity

    (1) John said Bill left

- Role of attribution in discourse
  - -- Illustrated with some examples from the Penn Discourse Treebank (PDTB)

PDTB Annotations: Attributions

- **Attribution features are annotated for**
  - **Explicit connectives**
  - **Implicit connectives**
  - **AltLex (Lexical phrases behaving as connectives)**

☞ **34% of discourse relations are attributed to an agent other than the writer.**

Attribution

- Attribution captures the relation of "ownership" between agents and Abstract Objects.
  But it is not a discourse relation!

- Attribution is annotated in the PDTB to capture:

(1) How discourse relations and their arguments can be *attributed to different individuals*:

  - When Mr. Green won a $240,000 verdict in a land condemnation case against the state in June 1983, [he says] *Judge O'Kicki unexpectedly awarded him an additional $100,000*.

  ⇒ Relation and Arg2 are attributed to the Writer.
  ⇒ *Arg1* is attributed to another agent.

Mismatch between sentence level semantics and discourse level semantics

There have been no orders for the Cray-3 so far, **though** the company says **it is talking with several prospects**.

Discourse semantics: contrary-to-expectation relation between "there being no orders for the Cray-3" and "there being a possibility of some prospects".

Sentence semantics: contrary-to-expectation relation between "there being no orders for the Cray-3" and "the company saying something".

Mismatch between sentence level semantics and discourse level semantics

Although takeover experts said they doubted Mr. Steinberg will make a bid by himself, the application by his Reliance Group Holdings Inc. could signal his interest in helping revive a failed labor-management bid.

Discourse semantics: contrary-to-expectation relation between "Mr. Steinberg not making a bid by himself" and "the RGH application signaling his bidding interest".

Sentence semantics: contrary-to-expectation relation between "experts saying something" and "the RGH application signaling Mr. Steinberg's bidding interest".

Attribution cannot always be excluded by default

- Advocates said the 90-cent-an-hour rise, to $4.25 an hour by April 1991, is too small for the working poor, **while** opponents argued that the increase will still hurt small business and cost many thousands of jobs.

Working with derivation trees can help as elementary trees corresponding to attributions may be easily included or left out as needed

Summary

- **Larger elementary structures as building blocks for localizing dependencies**
  - **-- adjunction (besides substitution) as a composition operation arises naturally**
- **Compositional semantics computed on the derivation tree on an LTAG and not on the derived tree**
  - **-- MRS type representation arises naturally**

Summary

- Multicomponent LTAGs arise naturally out of flexible composition
  - -- attachments for predicate argument composition distinguished from “scope” type composition
- Same game played out for categorial grammar
- Interaction with discourse
  - -- sometimes syntax has to hold back some spoonfuls from semantics
    - -- possible role for the derivation trees