A sensorimotor characterisation of syntax, and its implications for models of language evolution

Alistair Knott

Department of Computer Science, University of Otago

What's the best piece of news you can imagine receiving as a language evolution researcher?

What's the best piece of news you can imagine receiving as a language evolution researcher?

Here's one candidate:

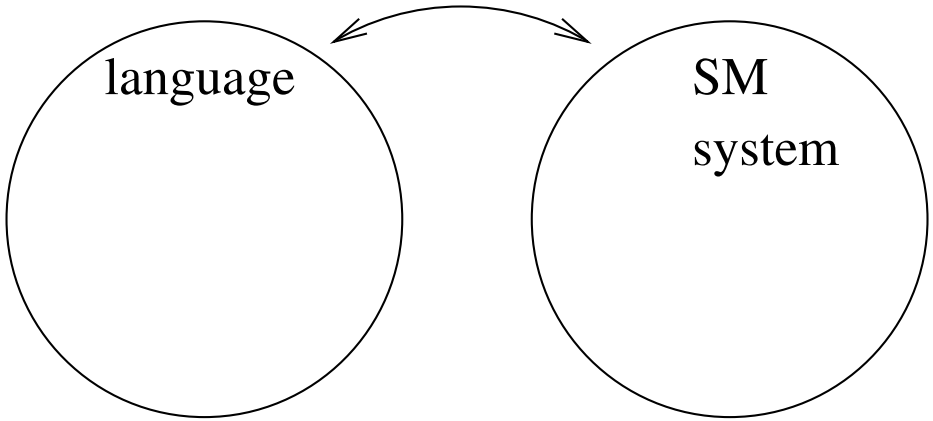
> *'There is a very direct interface between language and the sensorimotor system.'*

What's the best piece of news you can imagine receiving as a language evolution researcher?

Here's one candidate:

> *'There is a very direct interface between language and the sensorimotor system.'*

Fodor's idea: language and the SM system are modules.

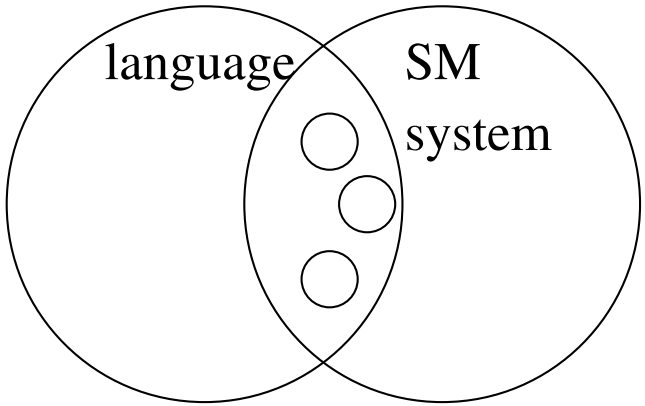


What's the best piece of news you can imagine receiving as a language evolution researcher?

Here's one candidate:

> *'There is a very direct interface between language and the sensorimotor system.'*

An alternative idea: language and the SM system share machinery.

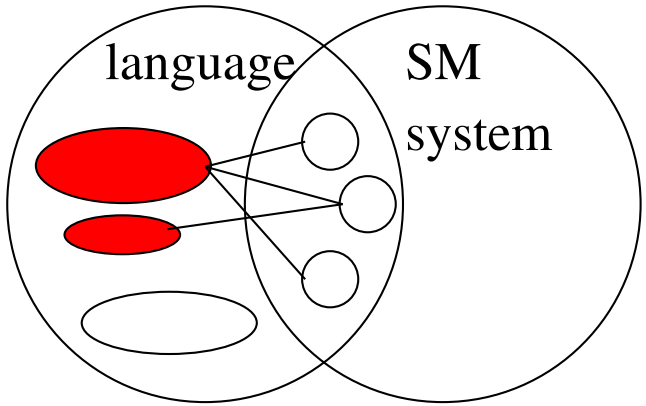


What's the best piece of news you can imagine receiving as a language evolution researcher?

Here's one candidate:

> *'There is a very direct interface between language and the sensorimotor system.'*

An alternative idea: language and the SM system share machinery. Some language circuits just 'read out' SM representations.

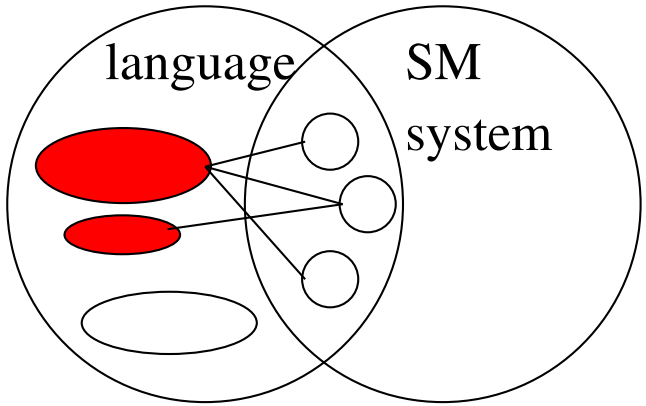


What's the best piece of news you can imagine receiving as a language evolution researcher?

Here's one candidate:

> *'There is a very direct interface between language and the sensorimotor system.'*

An alternative idea: language and the SM system share machinery. Part of the syntactic structure of language is sensorimotor in origin.

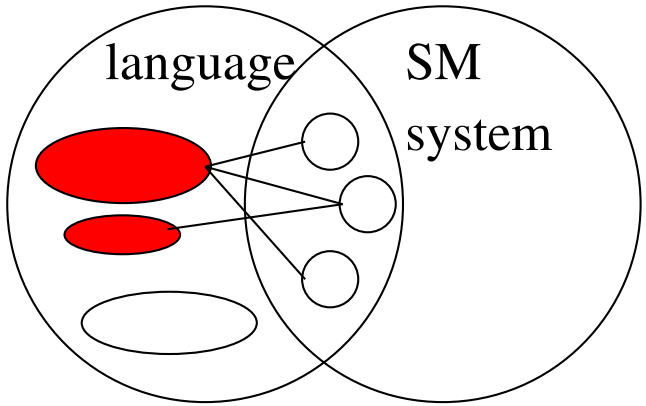


What's the best piece of news you can imagine receiving as a language evolution researcher?

Here's one candidate:

> *'There is a very direct interface between language and the sensorimotor system.'*

Why would this be good news for language evolution researchers?

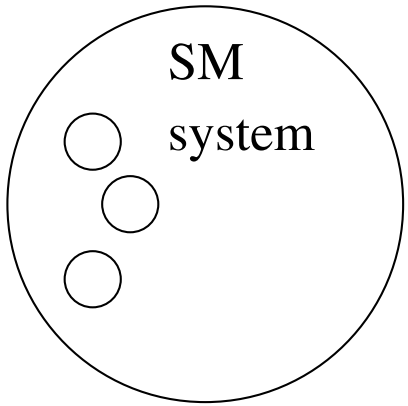


What's the best piece of news you can imagine receiving as a language evolution researcher?

Here's one candidate:

> *'There is a very direct interface between language and the sensorimotor system.'*

(1) It makes it very likely that some linguistic mechanisms evolved as biological adaptations of the SM system.

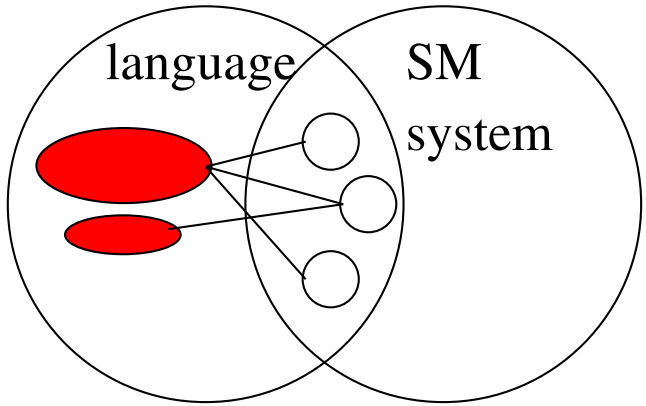


What's the best piece of news you can imagine receiving as a language evolution researcher?

Here's one candidate:

> *'There is a very direct interface between language and the sensorimotor system.'*

(1) It makes it very likely that some linguistic mechanisms evolved as biological adaptations of the SM system.

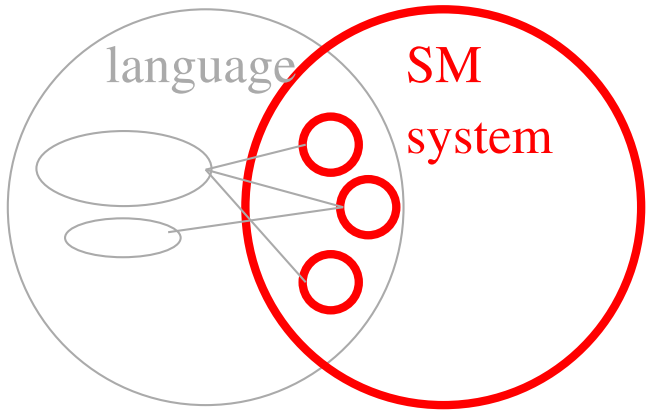


What's the best piece of news you can imagine receiving as a language evolution researcher?

Here's one candidate:

> *'There is a very direct interface between language and the sensorimotor system.'*

(2) We can use the *modern SM system* to approximate the platform from which language evolved.



Aim of the talk

I'll propose that for *simple*, *concrete* sentences, syntactic structure really is largely sensorimotor in origin.

Aim of the talk

I'll propose that for *simple*, *concrete* sentences, syntactic structure really is largely sensorimotor in origin.

I'll argue for a particularly strong hypothesis:

> *The syntactic structure of a simple sentence reporting a concrete episode in the world can be understood as a description of the sensorimotor processes involved in apprehending that episode.*

Aim of the talk

I'll propose that for *simple*, *concrete* sentences, syntactic structure really is largely sensorimotor in origin.

I'll argue for a particularly strong hypothesis:

> *The syntactic structure of a simple sentence reporting a concrete episode in the world can be understood as a description of the sensorimotor processes involved in apprehending that episode.*

This would be particularly good news for language evolution researchers.

An objection

‘Languages of the world are very different from one another, but we all have the same sensorimotor processes.’

To maintain the strong hypothesis, we must adopt a syntactic theory in which differences between sentences reporting the same episode in different languages are relatively superficial.

An objection

‘Languages of the world are very different from one another, but we all have the same sensorimotor processes.’

To maintain the strong hypothesis, we must adopt a syntactic theory in which differences between sentences reporting the same episode in different languages are relatively superficial.

Perhaps surprisingly, the syntactic framework we are drawn towards is some form of Chomskyan generative grammar.

- Sentences have a ‘surface’ structure and an ‘underlying’ structure.
- It’s the *underlying* structure which can be interpreted as a description of SM processing.

Method

Take a simple example episode: **a man grabs a cup**.

Develop two models:

- A **syntactic model** of the (underlying) structure of sentences reporting this episode in different languages, according to some variety of Chomskyan syntax.
- A **sensorimotor model** of the processes which allow the episode to be perceived or executed, according to current psychology/ neuroscience.

Question: do these models overlap?

I argue that there are formal similarities between the syntactic and sensorimotor models, which are strong enough to require explanation.

A syntactic model of *The man grabbed a cup*

I assume a version of Minimalism (Chomsky, 1995).

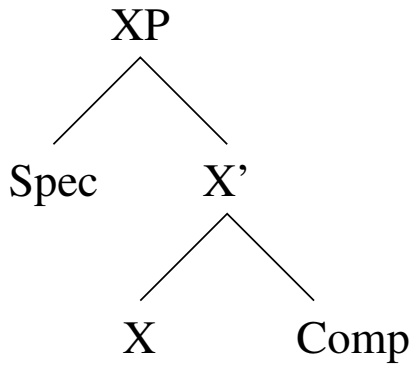
In Minimalism, a sentence needs to be described at two different syntactic levels:

- Phonetic form (PF) represents the surface form of the sentence.
- Logical form (LF) is 'the level of syntactic representation which interfaces with the semantic system'.

I'll assume that all translations of *The man grabbed a cup* have the same LF.

The building blocks of syntactic structures

The basic unit of phrase structure is the X-bar schema.

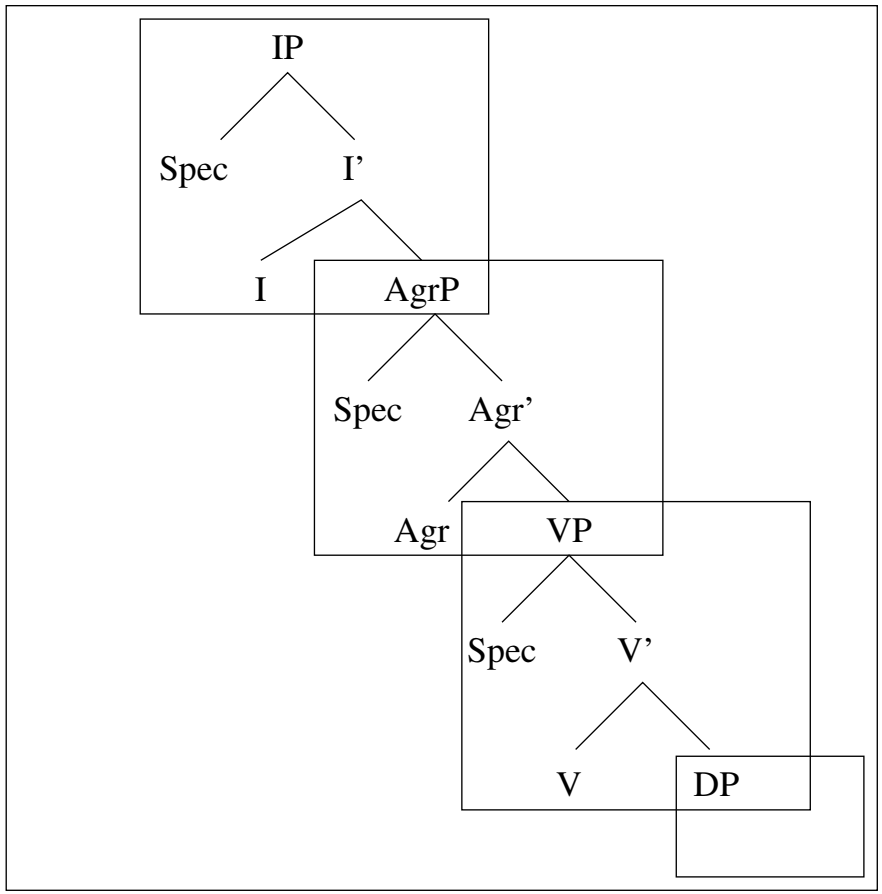


Each word in a sentence contributes an XP structure.

- The head of the structure (X) is the word itself.
- The structure also has slots for a specifier (Spec) and a complement (Comp).
- These slots can (recursively) be occupied by other XPs.

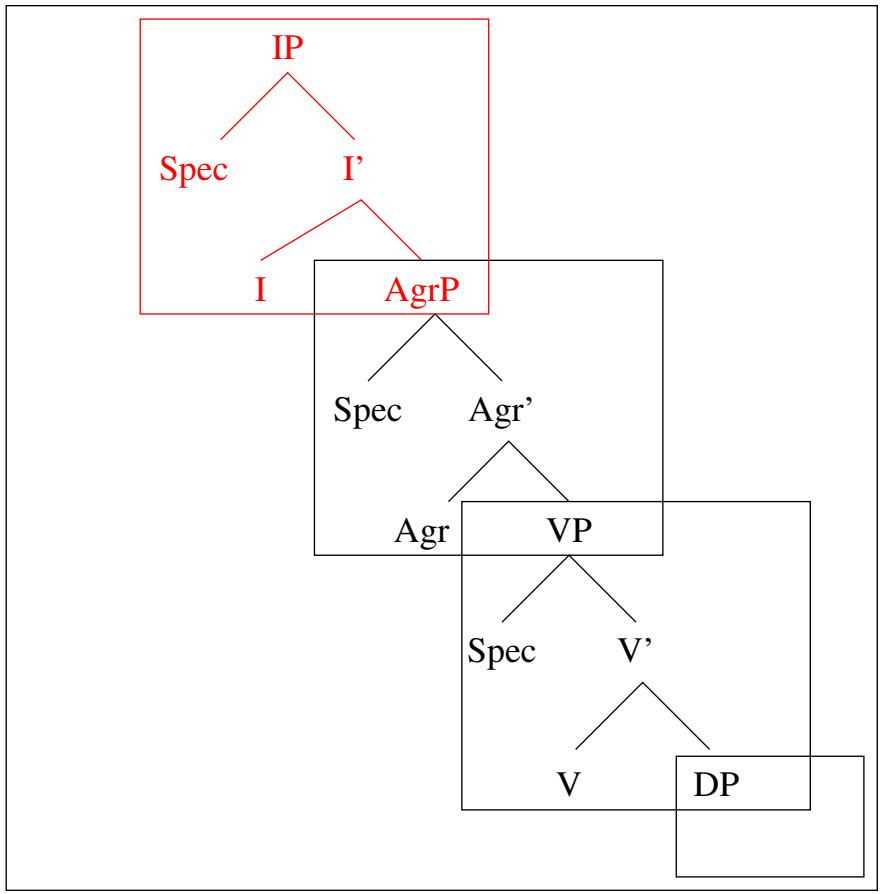
LF structure of *The man grabbed a cup*

The LF of our sentence is a right-branching pattern of XPs.



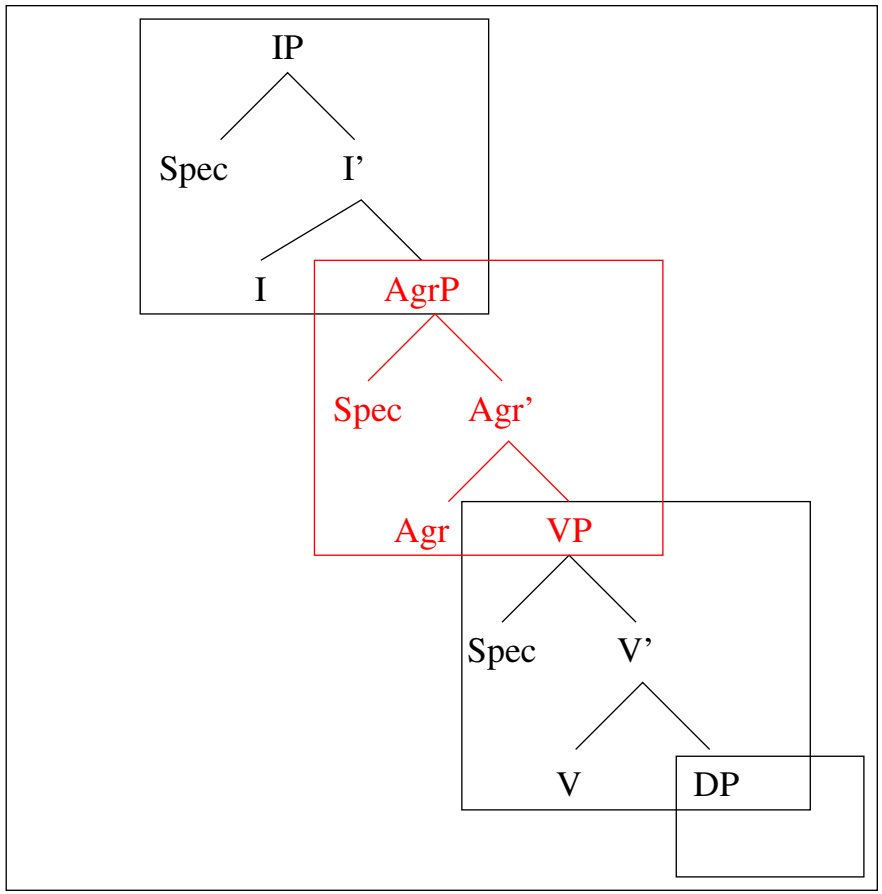
LF structure of *The man grabbed a cup*

IP is associated with the subject.



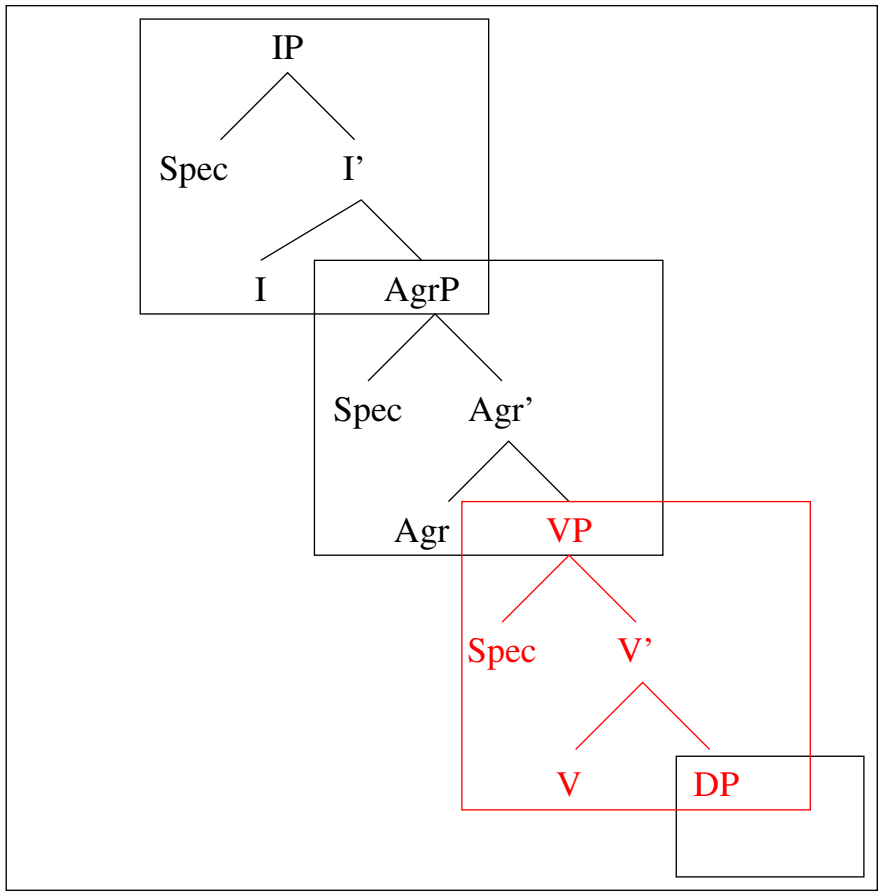
LF structure of *The man grabbed a cup*

AgrP is associated with the object.



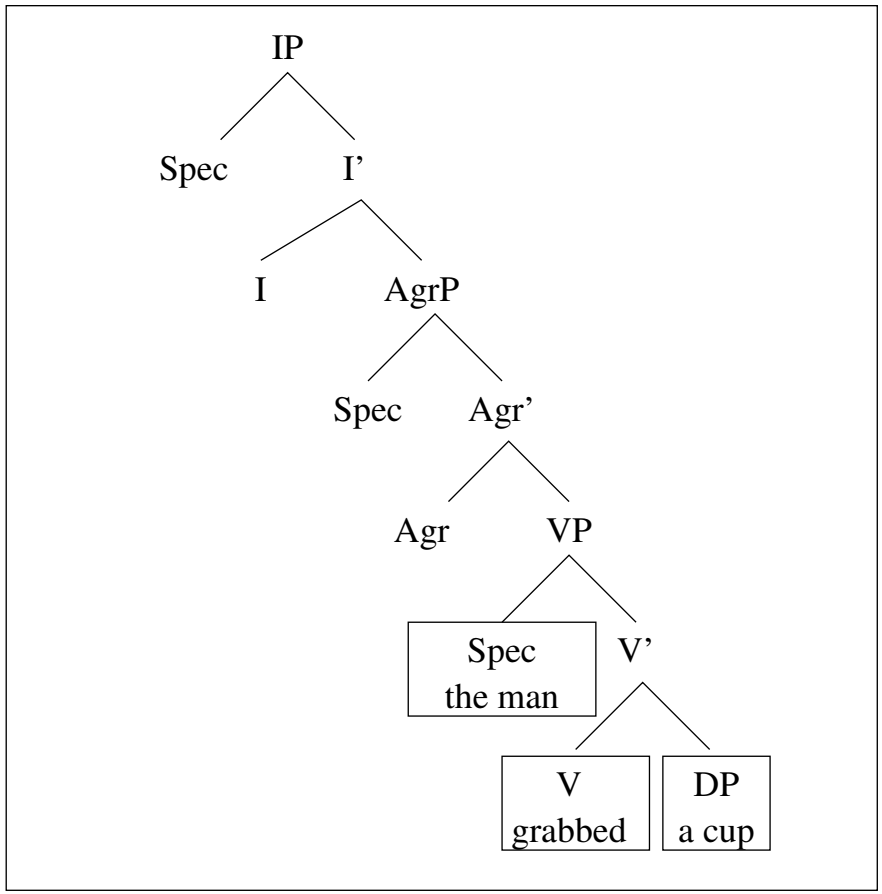
LF structure of *The man grabbed a cup*

AgrP is associated with the verb.



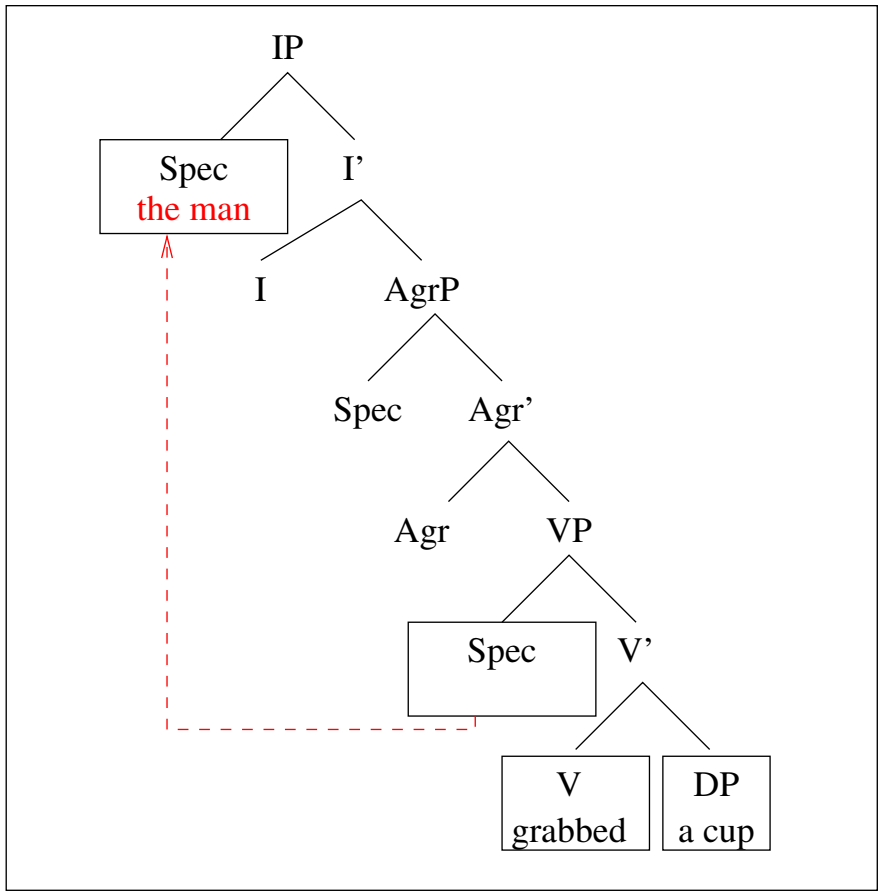
LF structure of *The man grabbed a cup*

VP is headed by the verb, and introduces the subject and object.



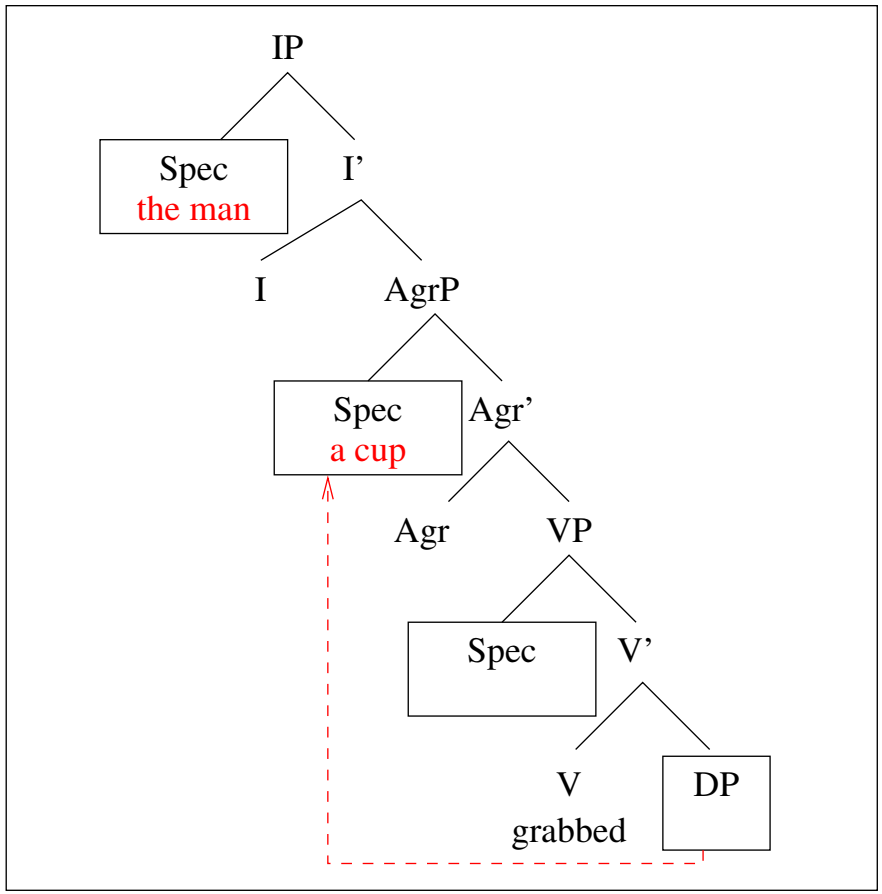
LF structure of *The man grabbed a cup*

The subject raises to [Spec,IP] to get Case (assigned by I)



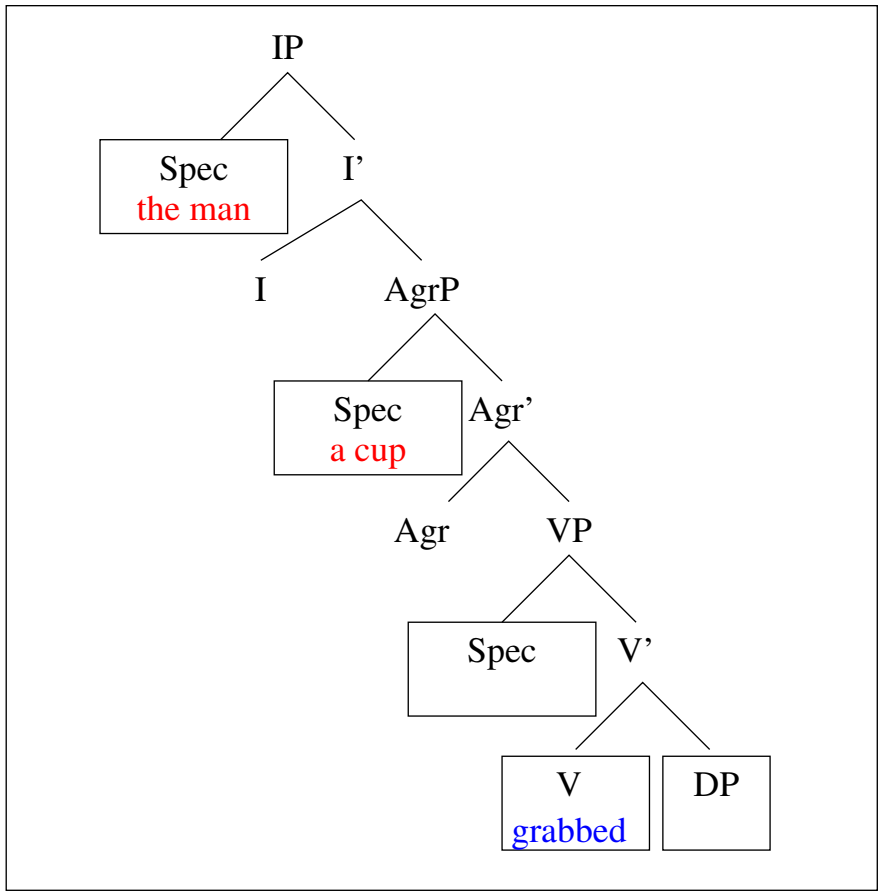
LF structure of *The man grabbed a cup*

The object raises to [Spec,AgrP] to get Case (assigned by Agr)



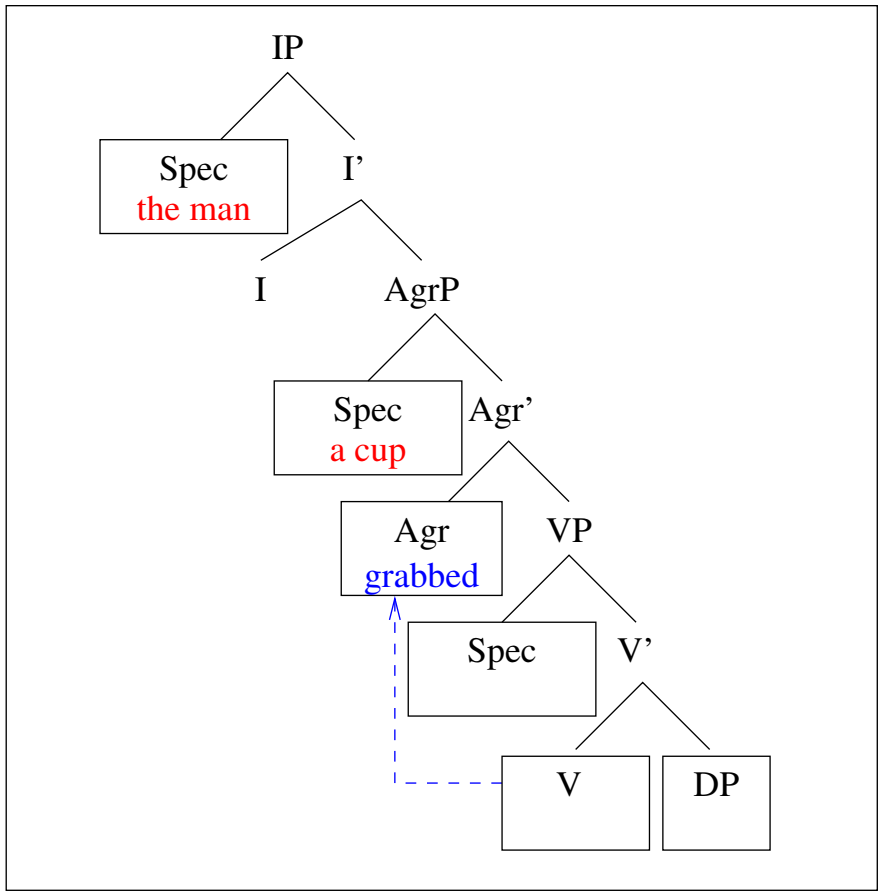
LF structure of *The man grabbed a cup*

The verb raises successively to Agr and I.



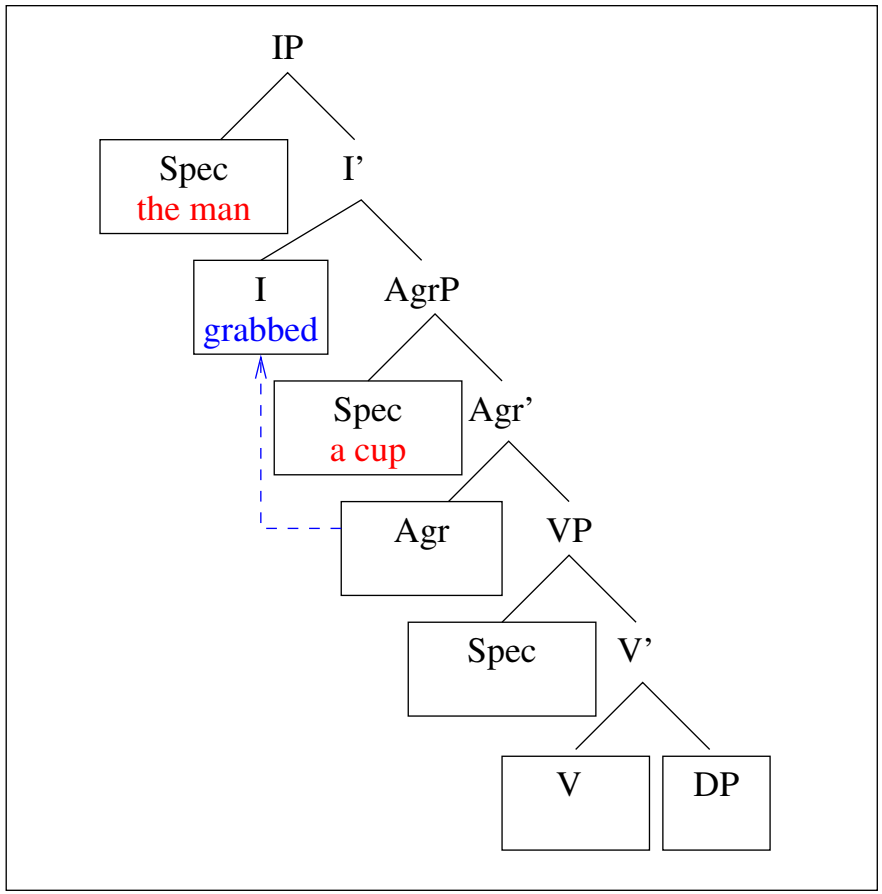
LF structure of *The man grabbed a cup*

The verb raises successively to Agr and I.



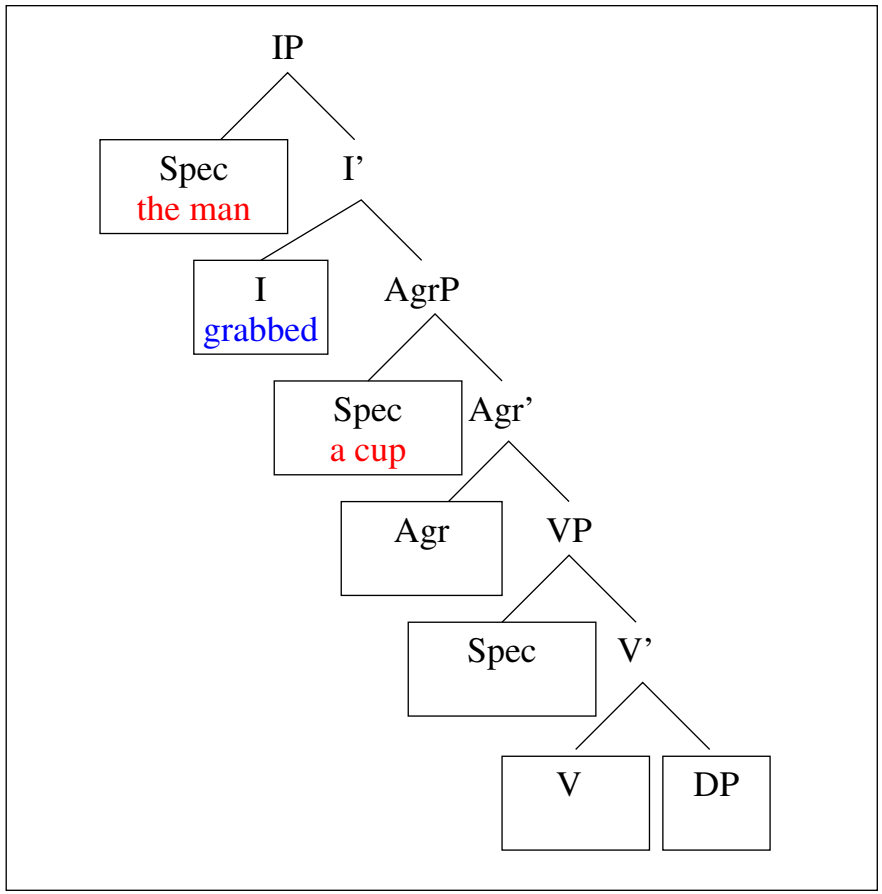
LF structure of *The man grabbed a cup*

The verb raises successively to Agr and I.



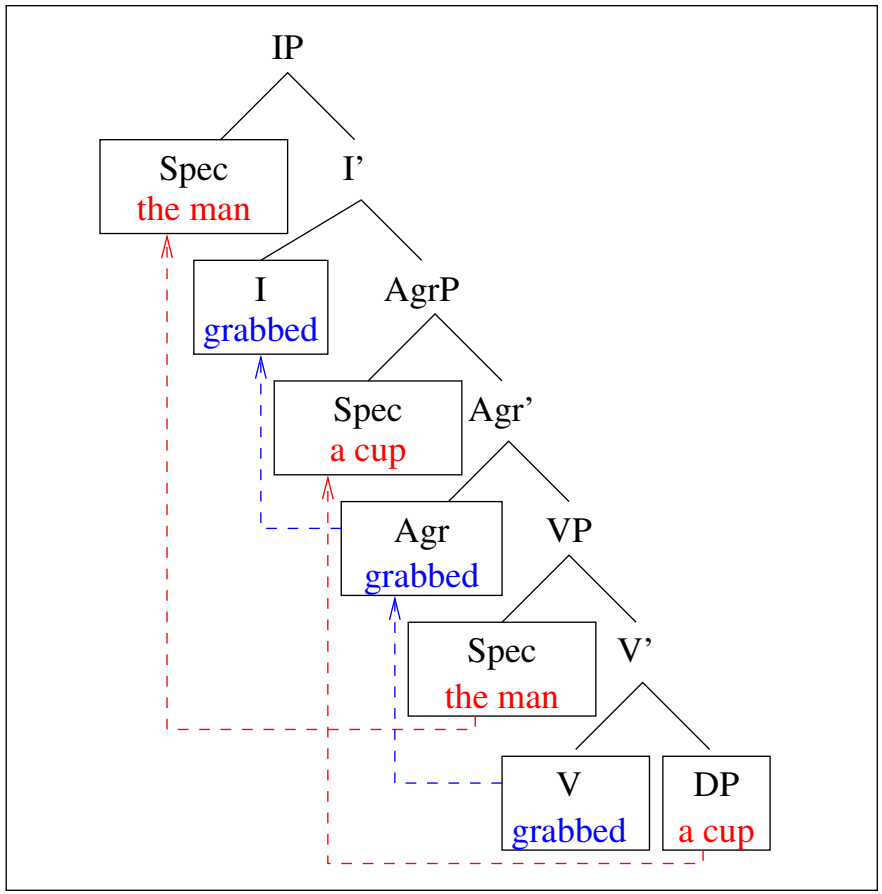
LF structure of *The man grabbed a cup*

The verb raises successively to Agr and I.



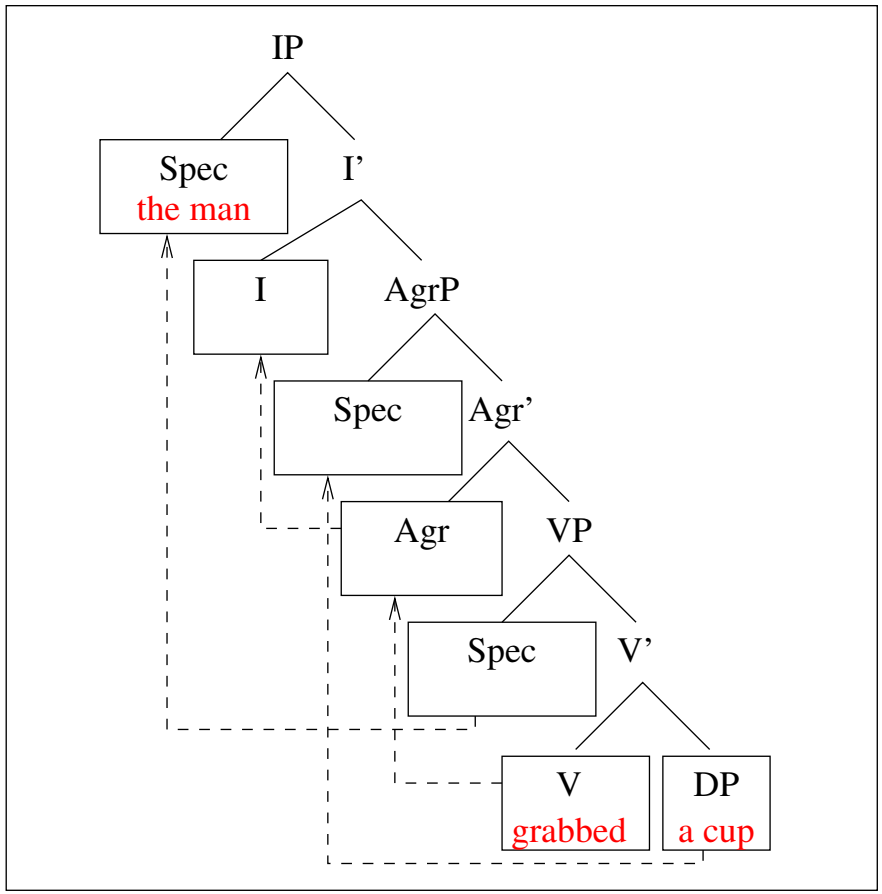
LF structure of *The man grabbed a cup*

Languages have different conventions about when to pronounce the different elements of LF.



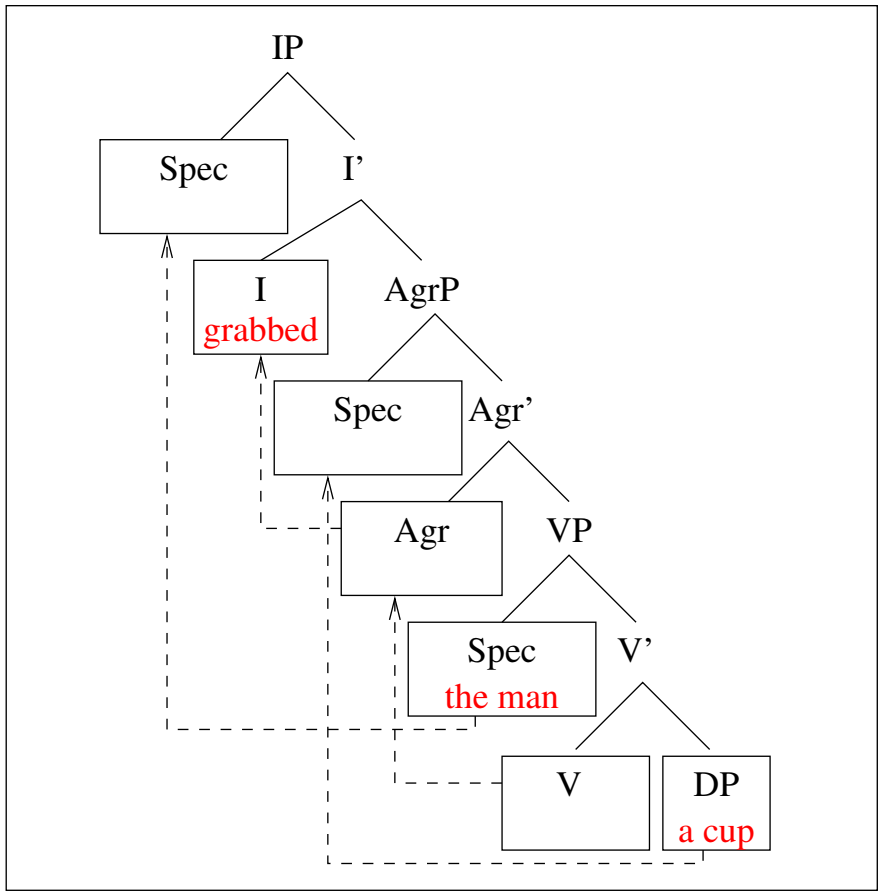
LF structure of *The man grabbed a cup*

English PF looks like this:



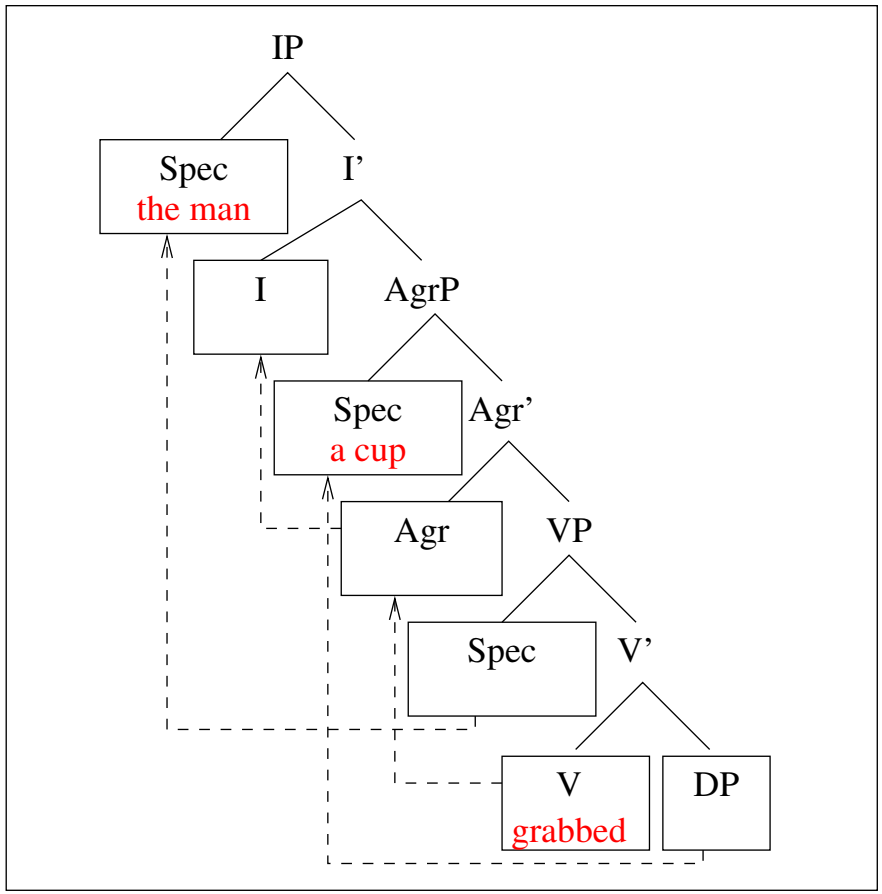
LF structure of *The man grabbed a cup*

Māori PF looks like this:



LF structure of *The man grabbed a cup*

Japanese PF looks like this:

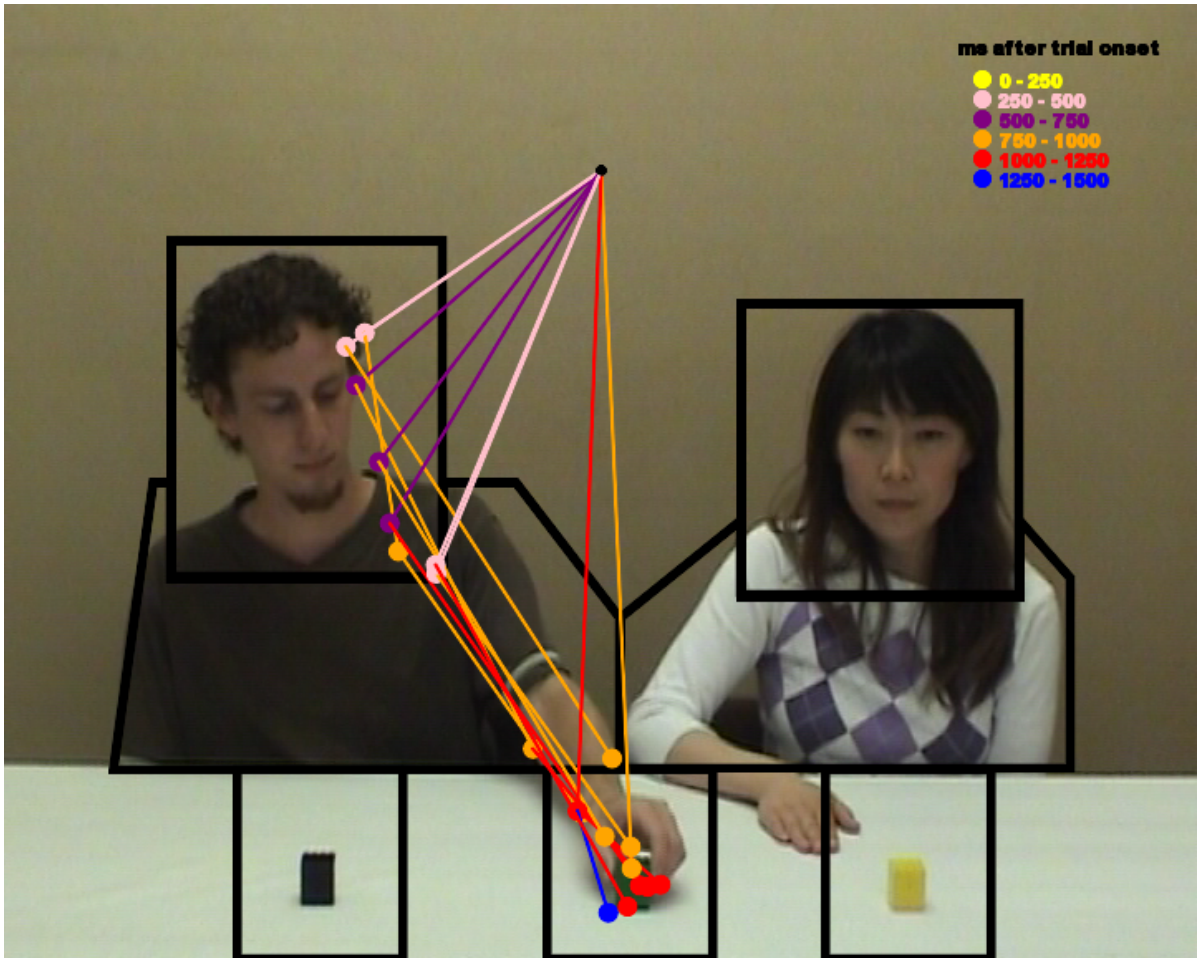


A sensorimotor model of reach-to-grasp actions

1. I argue that the processes involved in perceiving or executing a reach-to-grasp action have a strong *sequential structure*:

- The observer must first attend to the agent
- and then attend to the target
- and only then dynamically monintor the action to completion.

Some evidence from an eye-tracking study



(Webb *et al.* 'Eye movements during transitive action observation have sequential structure', Acta Psychologica 2010)

A sensorimotor model of reach-to-grasp actions

2. I argue that both agent and target are *reattended to* when the action is experienced.

- During action monitoring, the agent is reattended to *as a dynamic entity*.
- At the end of the action, the target is reattended to *as a motor state*.

I suggest that these actions of reattention are crucial for the development of cross-modal object representations.

SM signals during experience of a reach-to-grasp

Context	Action signals	Reafferent signals
C1	*attend_agent*	*attending_to_agent*
C2	*attend_cup*	*attending_to_cup*
C3	*grasp*	*attending_to_agent*
C4		*attending_to_cup*

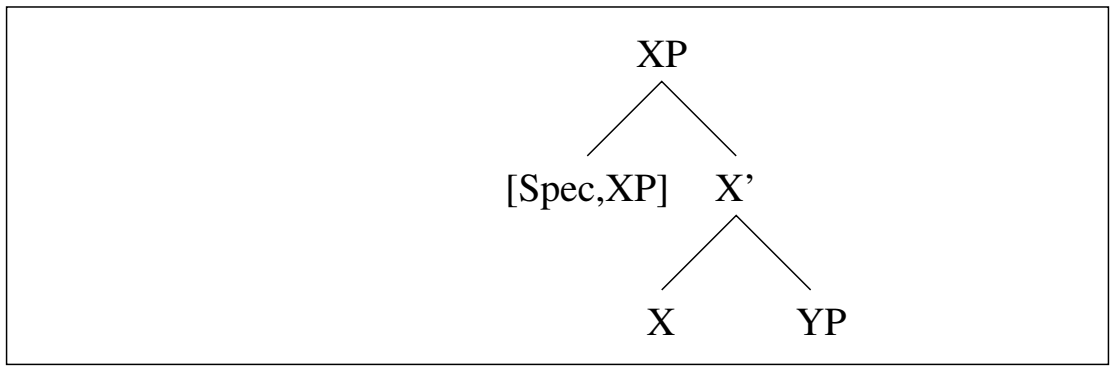
A SM characterisation of LF

My proposal (roughly):

> *The LF of 'The man grabbed a cup' describes the sequence of SM operations/states evoked during experience of a cup-grabbing event.*

SM interpretation of an XP schema

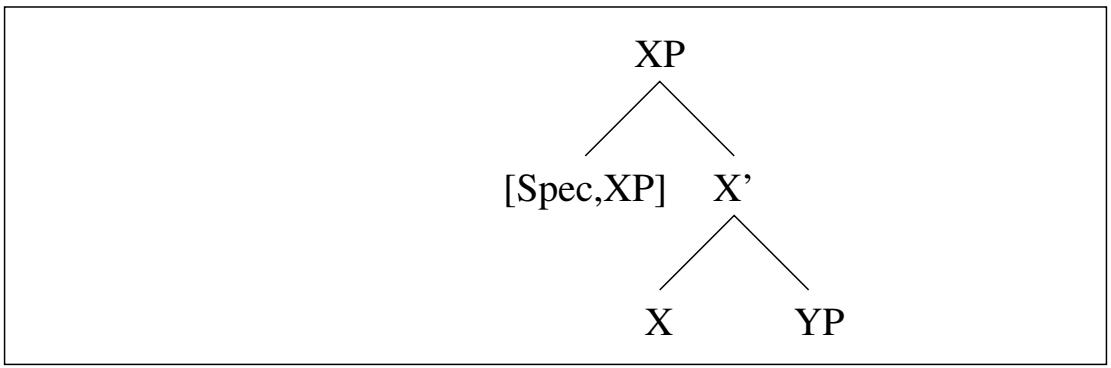
I express this proposal by giving a general SM interpretation of an XP schema.



SM interpretation of an XP schema

I express this proposal by giving a general SM interpretation of an XP schema.

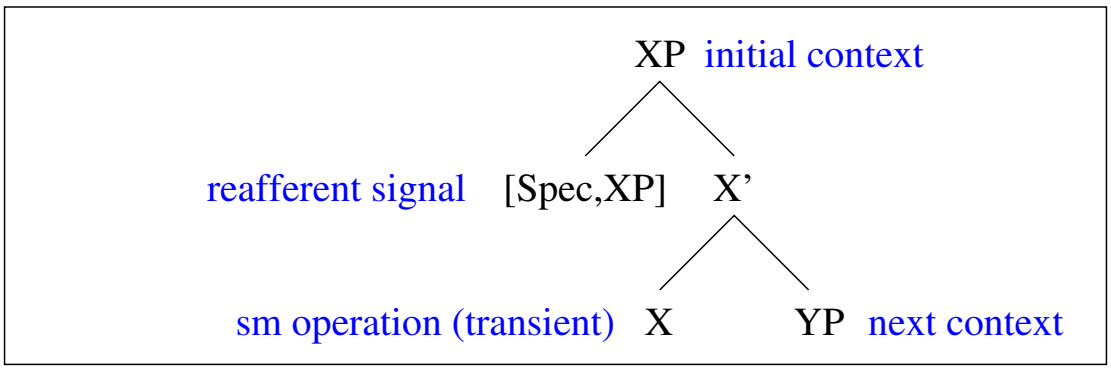
- An XP in an LF structure denotes a *SM operation*, executed in a certain *initial context*, which generates a *reafferent consequence*, and establishes a *new context*.



SM interpretation of an XP schema

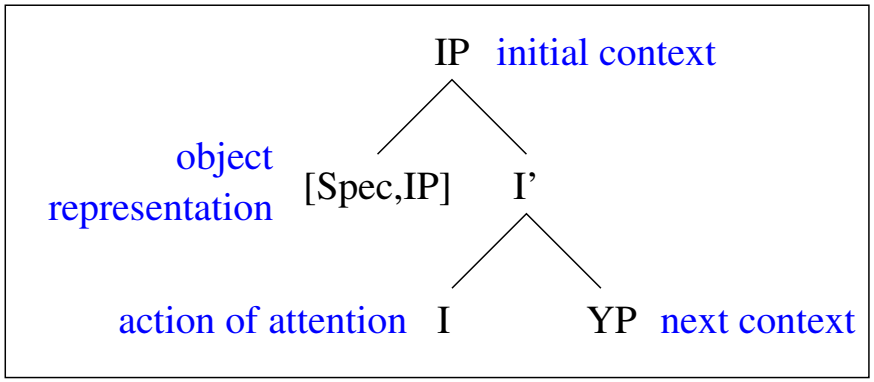
I express this proposal by giving a general SM interpretation of an XP schema.

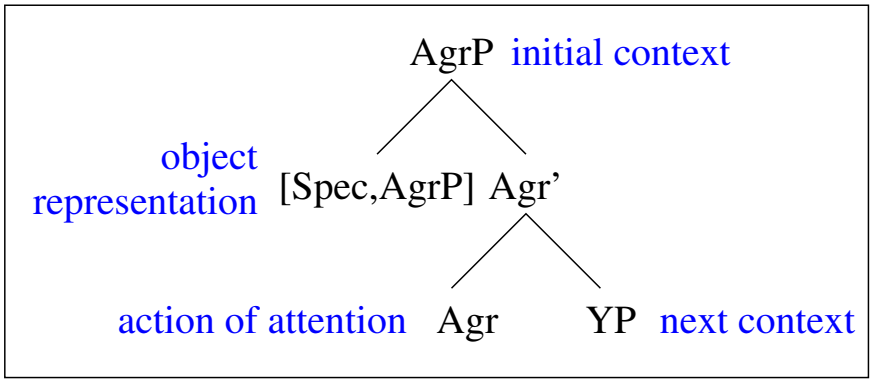
- An XP in an LF structure denotes a *SM operation*, executed in a certain *initial context*, which generates a *reafferent consequence*, and establishes a *new context*.
- Each item in the XP schema has a SM interpretation.



XPs denoting attentional operations

IP and AgrP denote attentional operations, and their reafferent consequences.



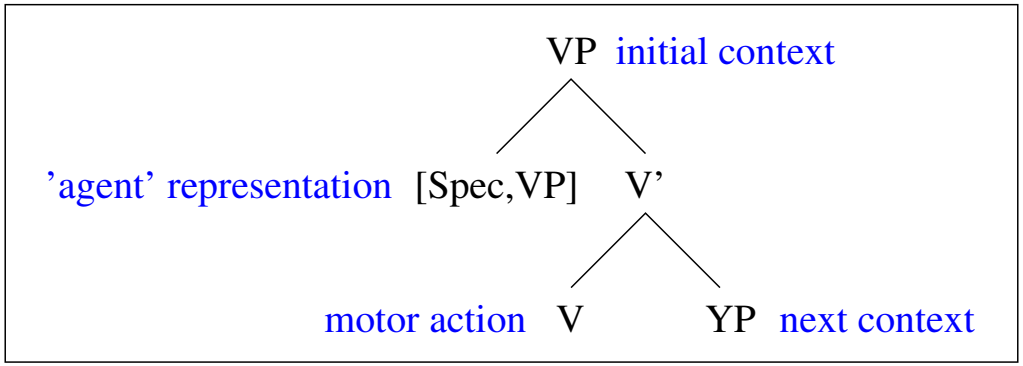


In these XPs:

- The head (which 'checks' a piece of verb morphology) denotes an action of attention.
- The specifier (the position occupied by a DP) denotes the sensory representation of an object.

XPs denoting motor operations

VP denotes a motor operation, and its reafferent consequences.

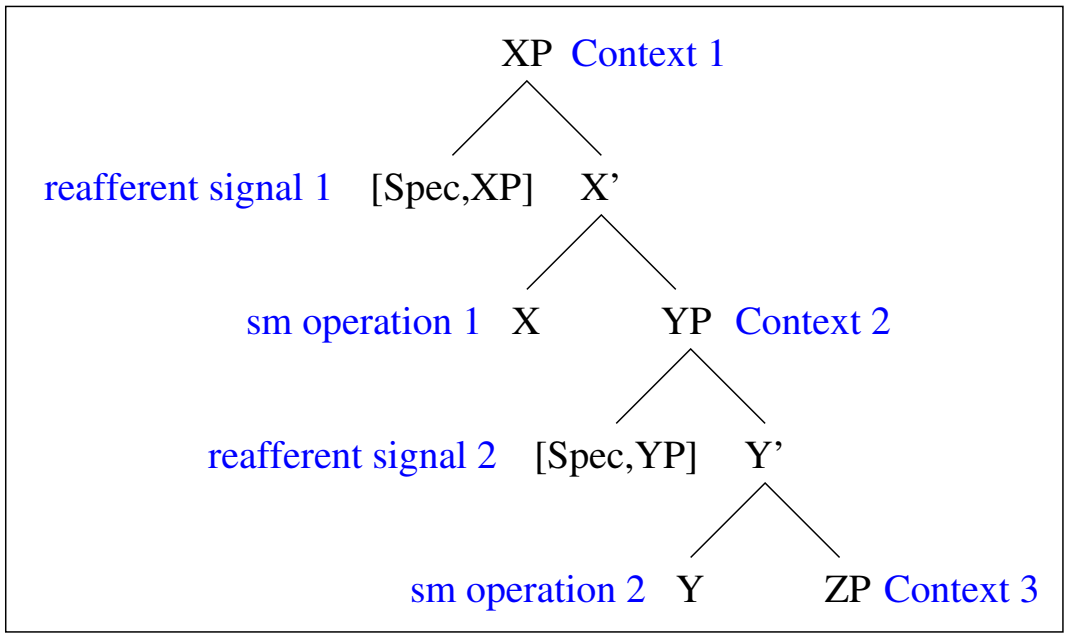


- The V head denotes a motor programme.
- The specifier of VP denotes the sensory representation of an agent.

A right-branching structure of X-bar schemas

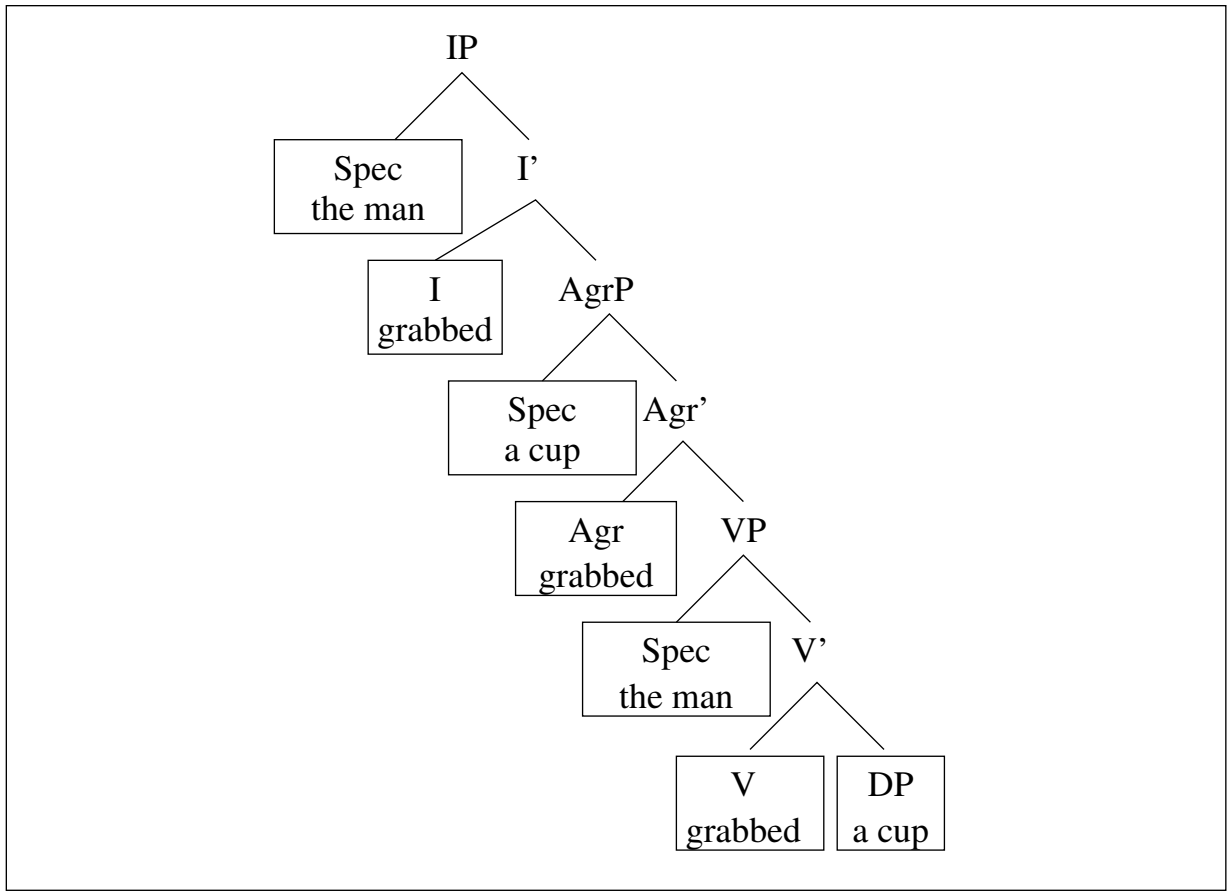
In a right-branching X-bar structure, the next context of one XP is the initial context of its complement XP.

So a right-branching structure of XPs describes a *sequence* of sensorimotor operations.



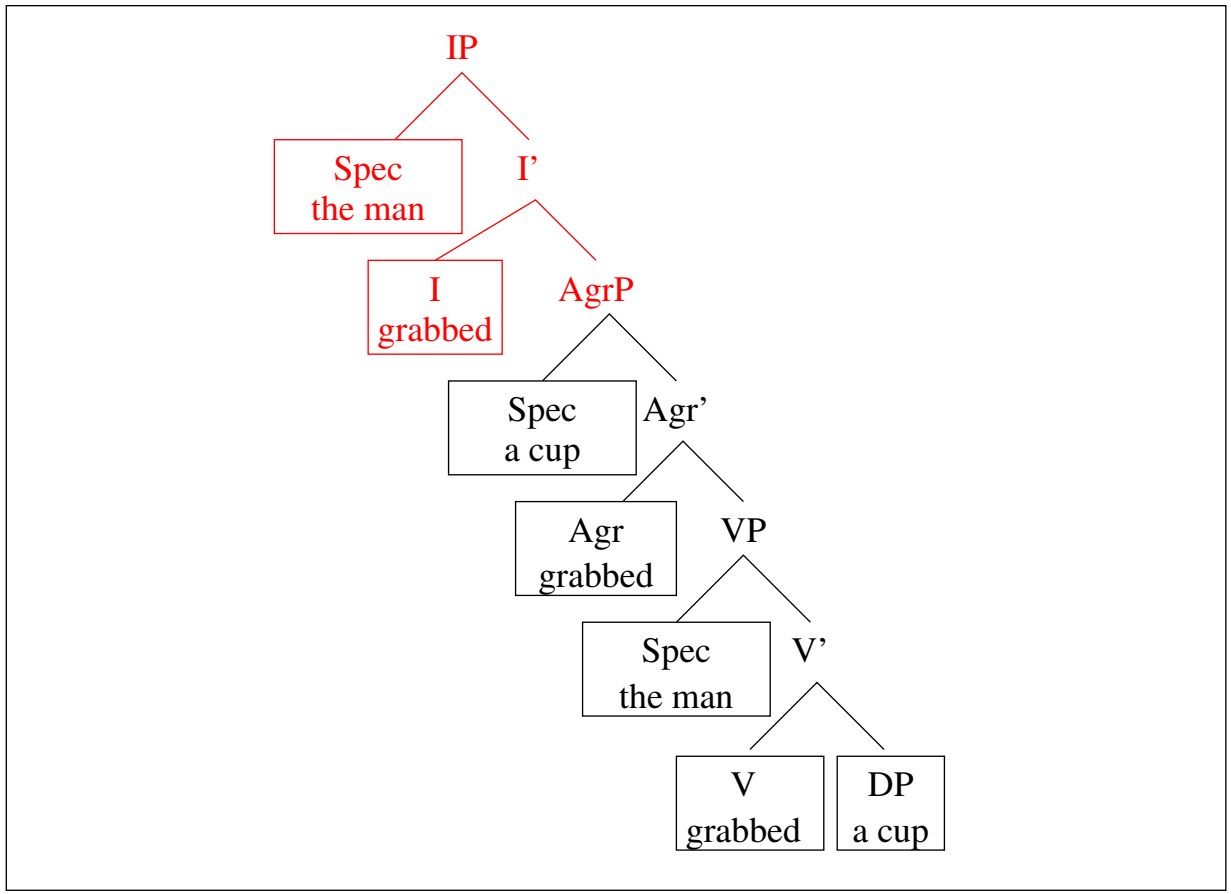
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



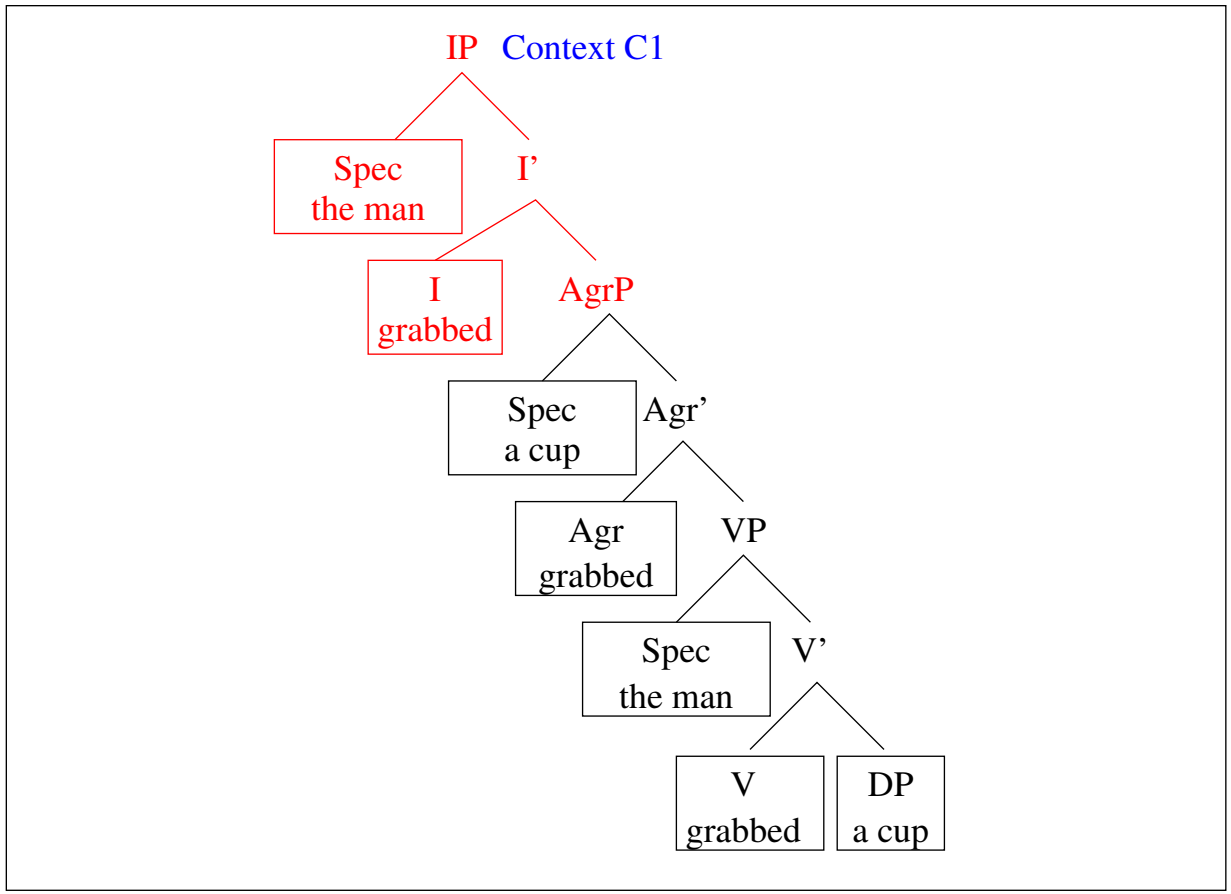
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



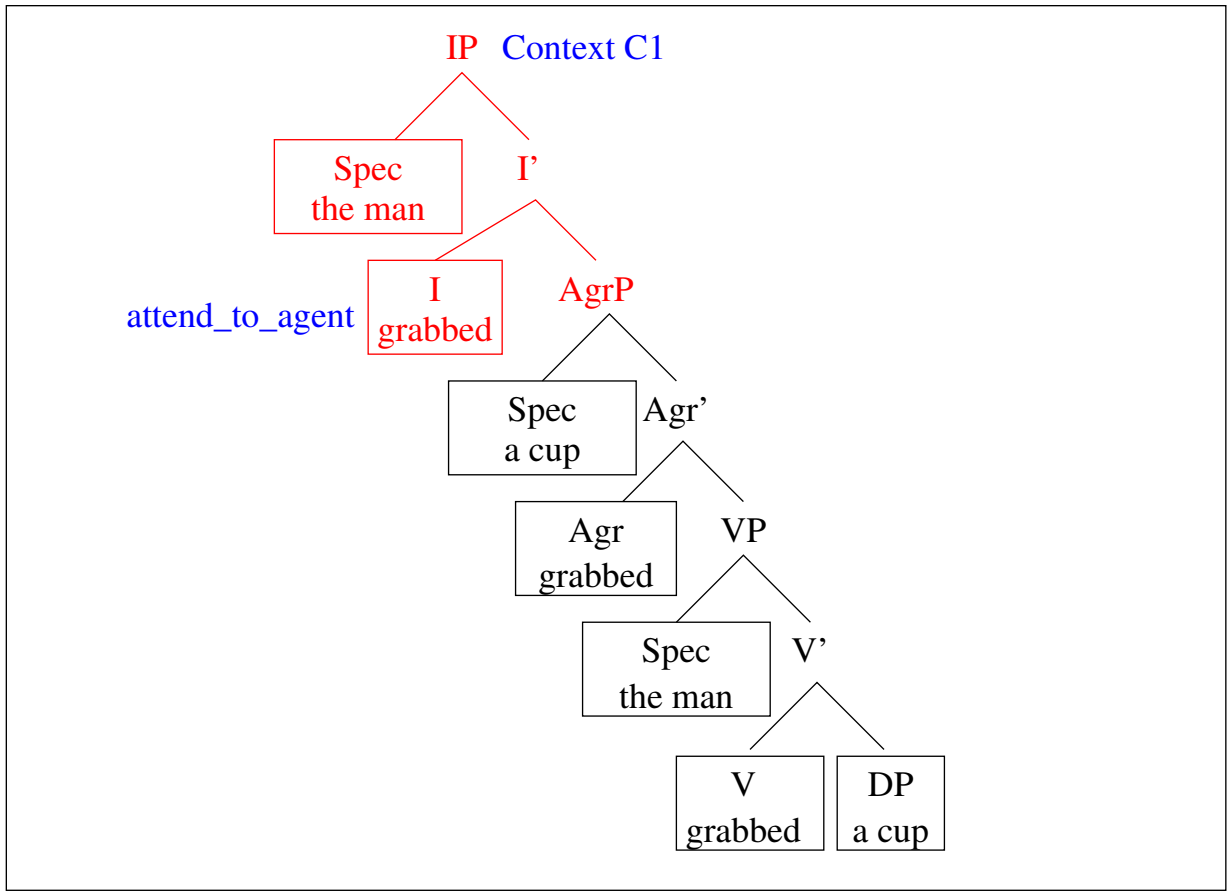
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



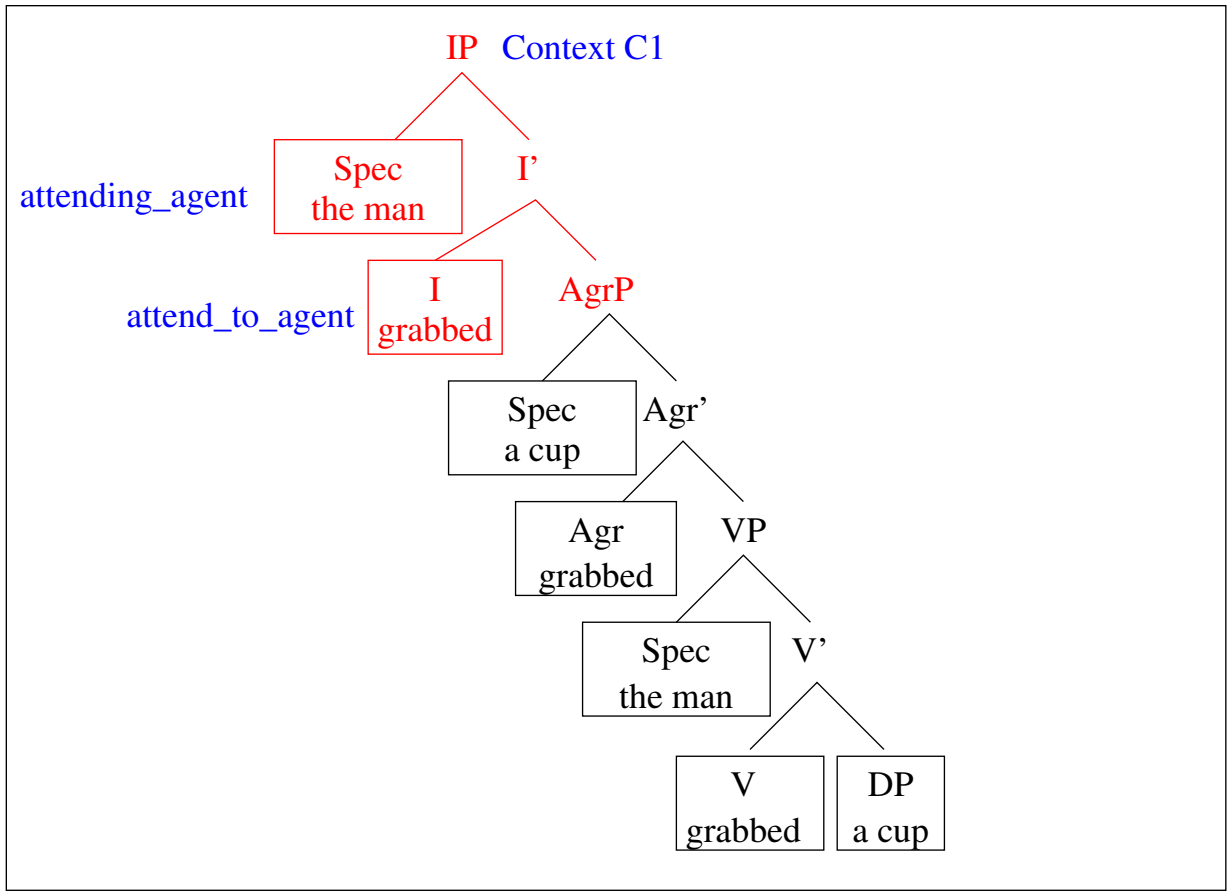
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



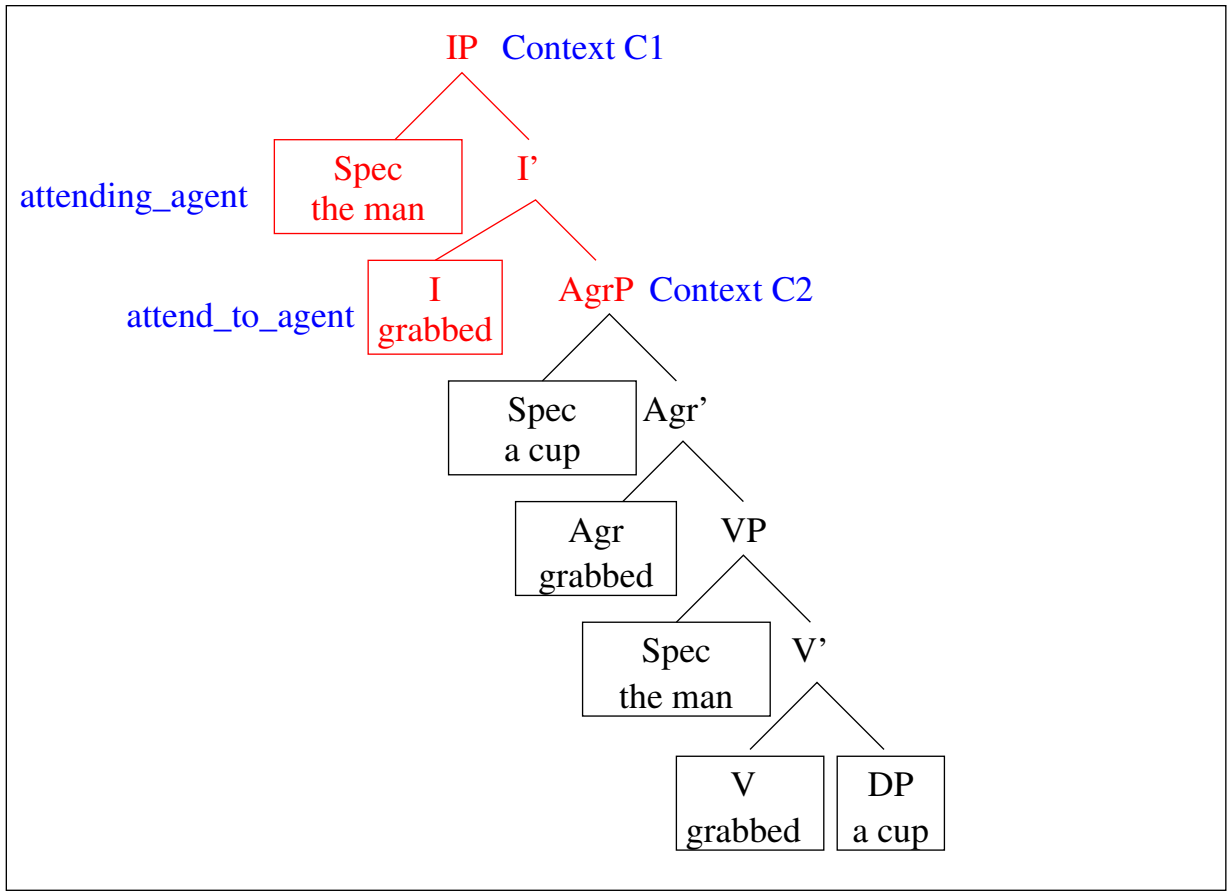
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



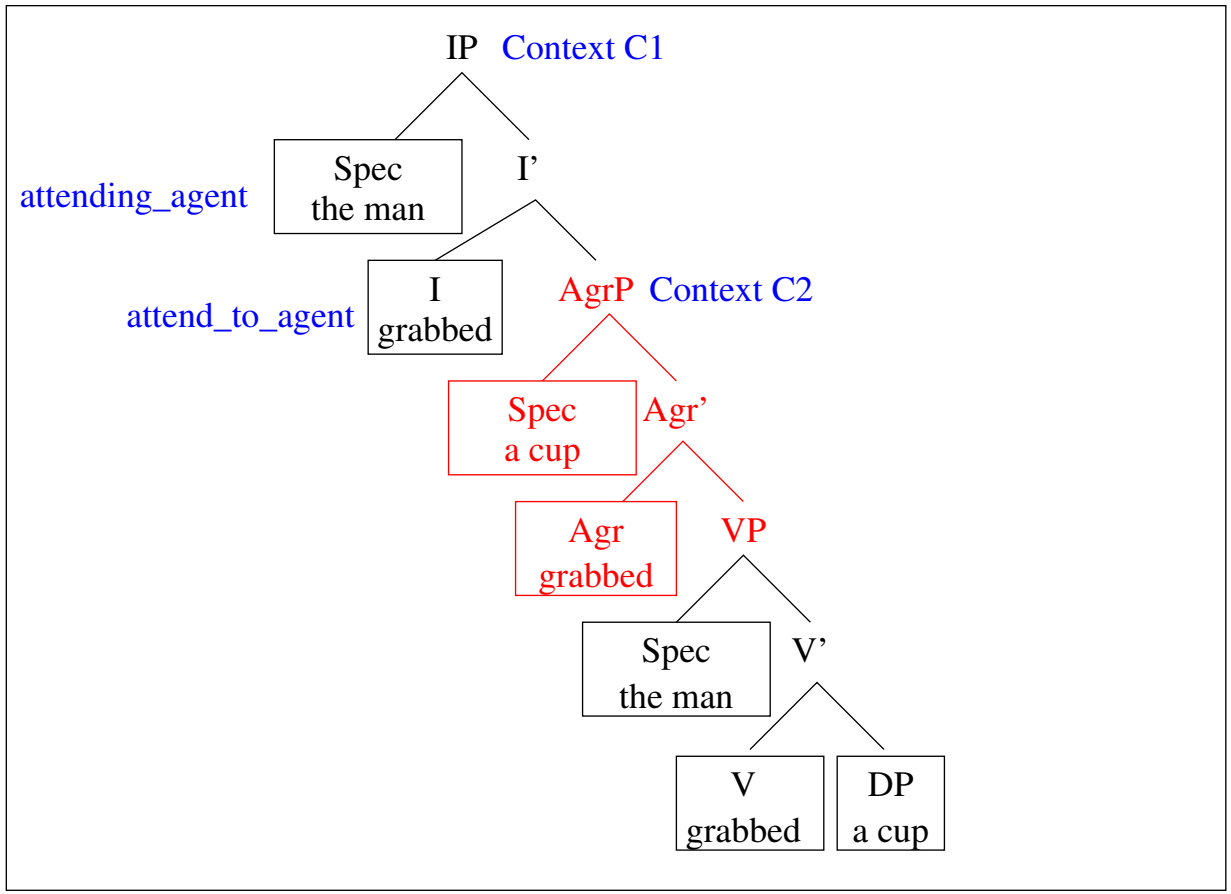
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



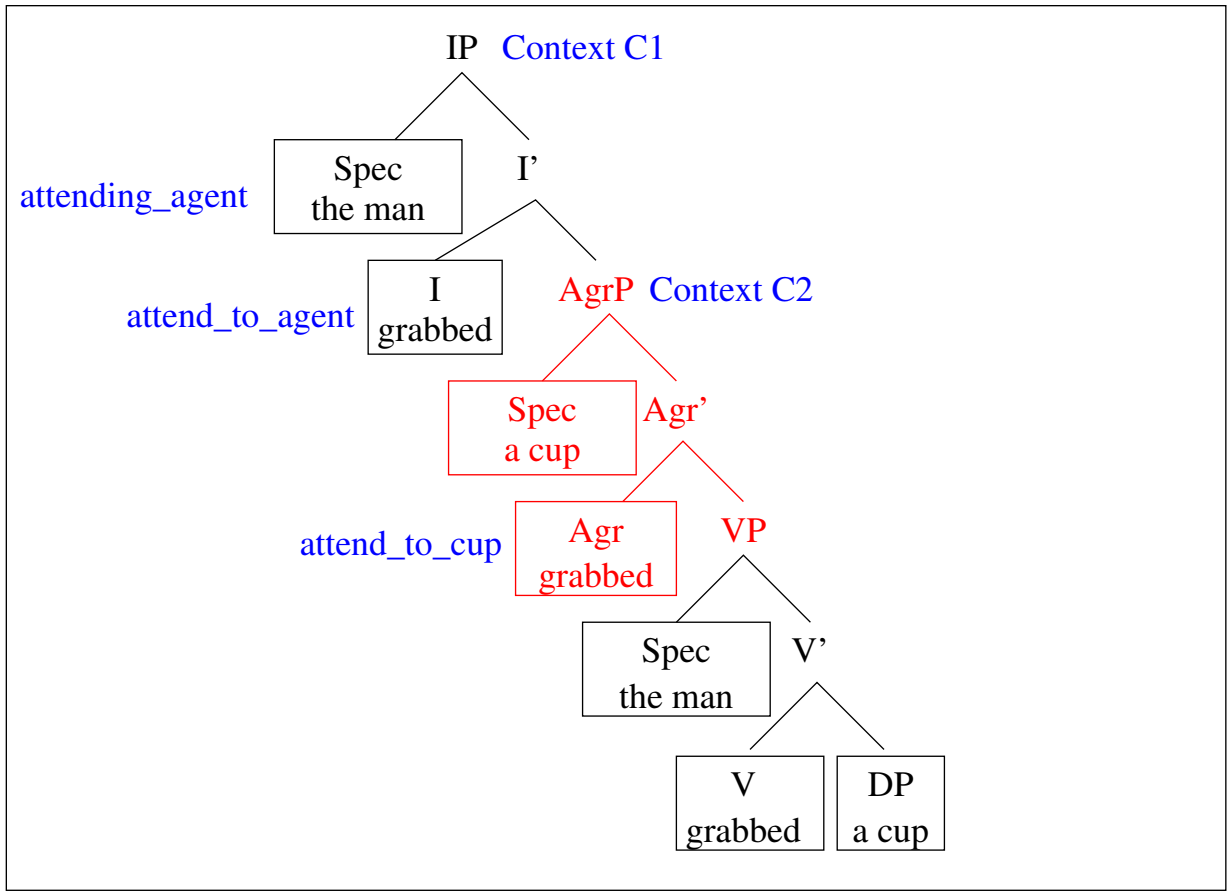
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



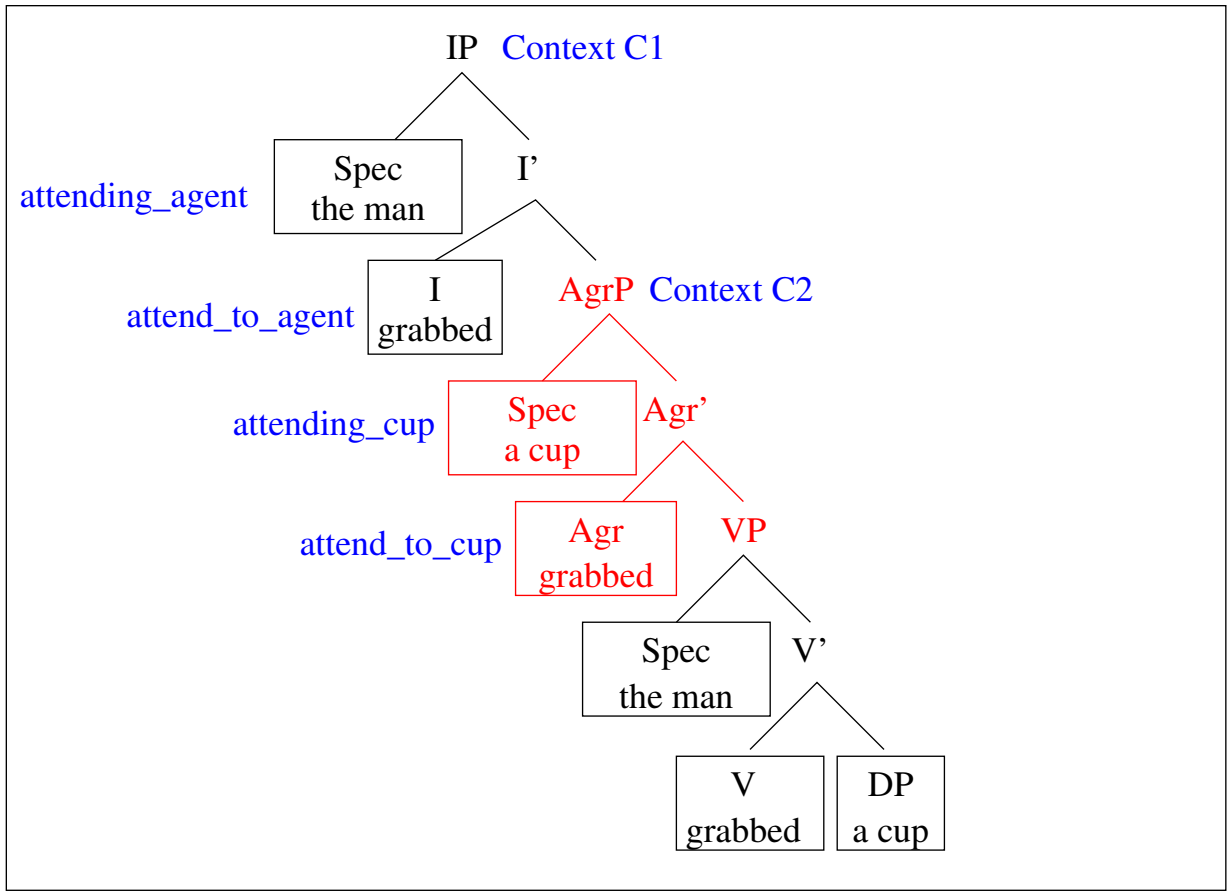
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



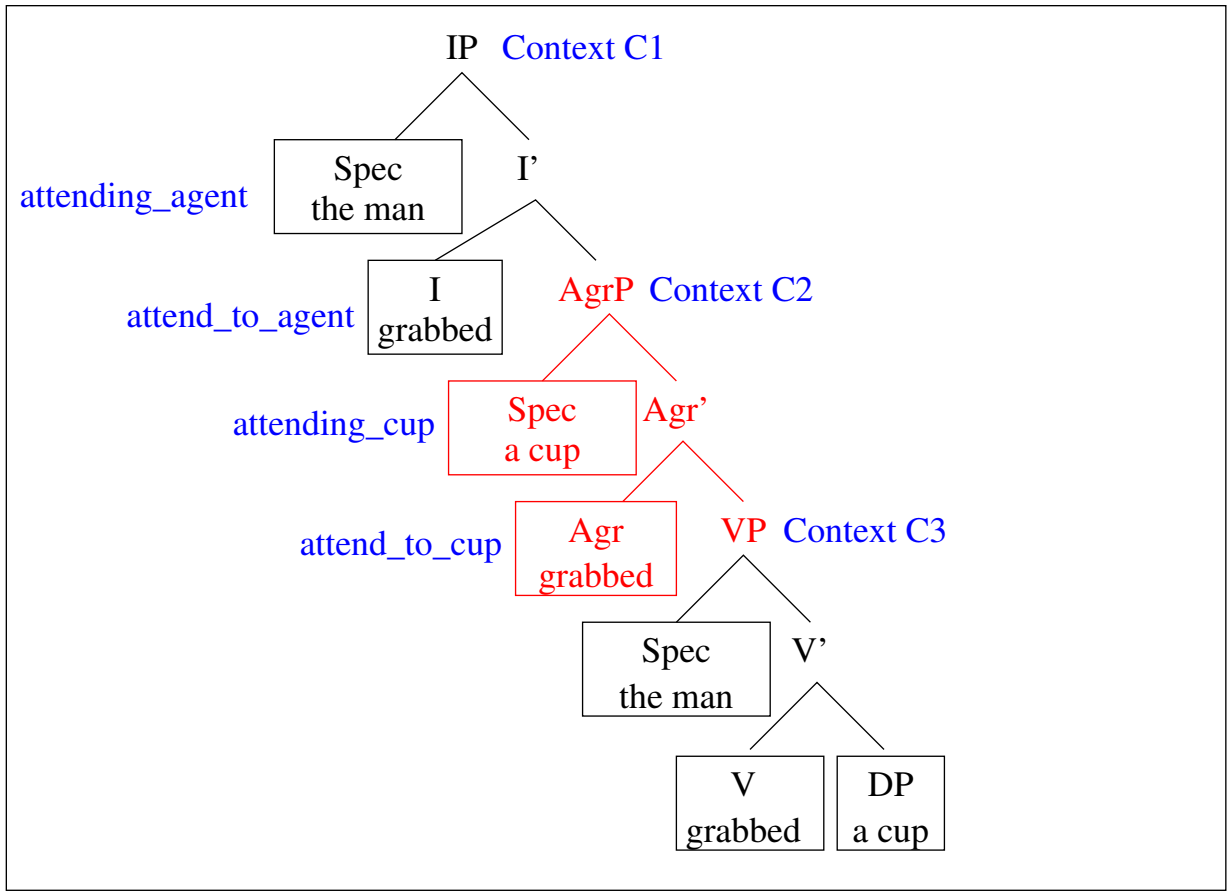
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



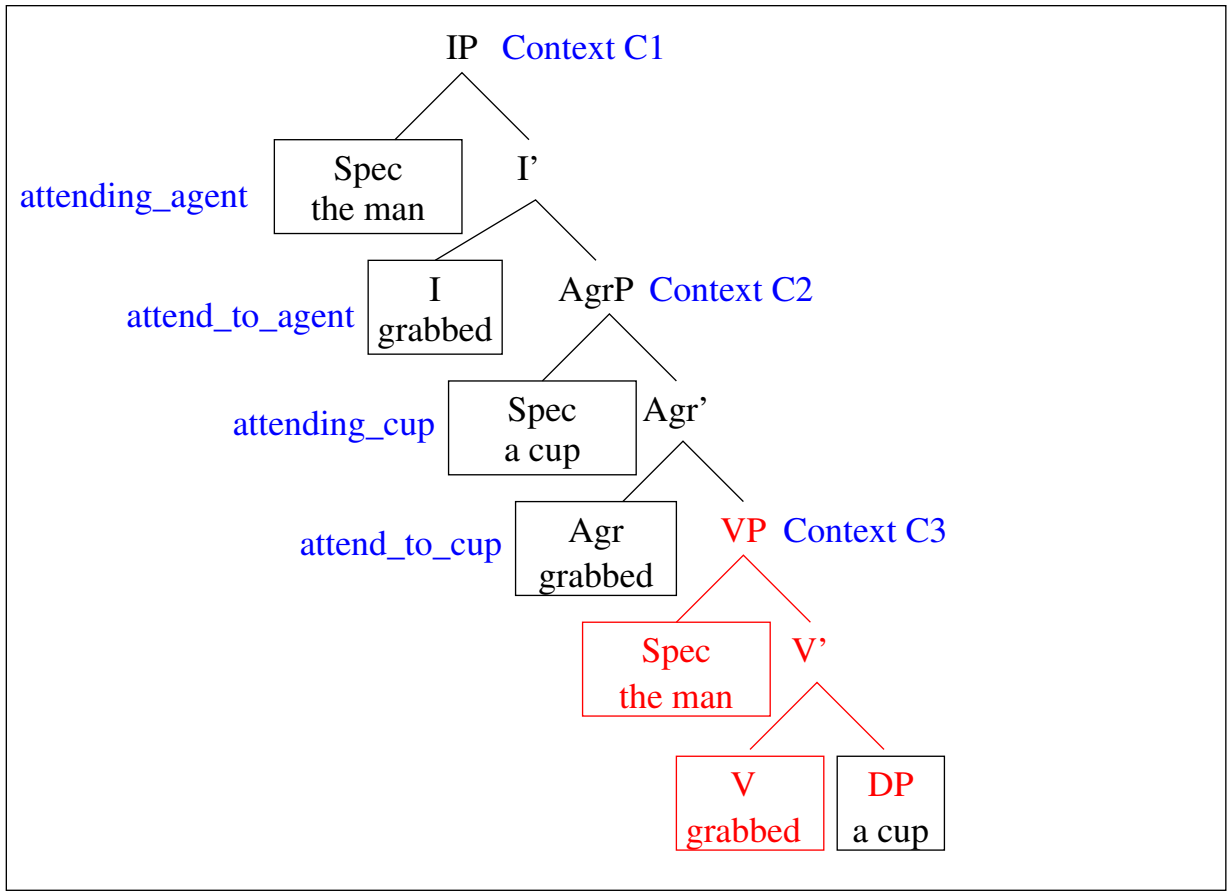
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



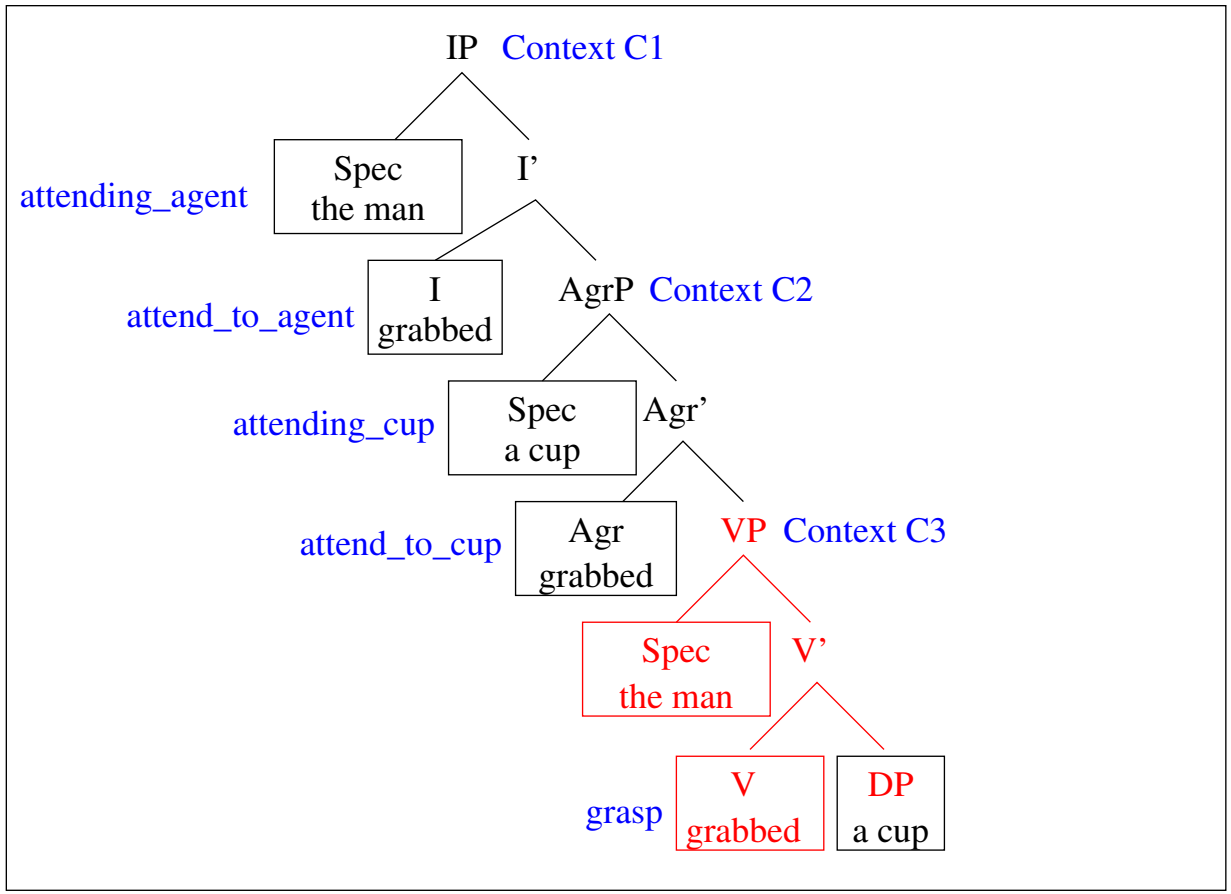
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



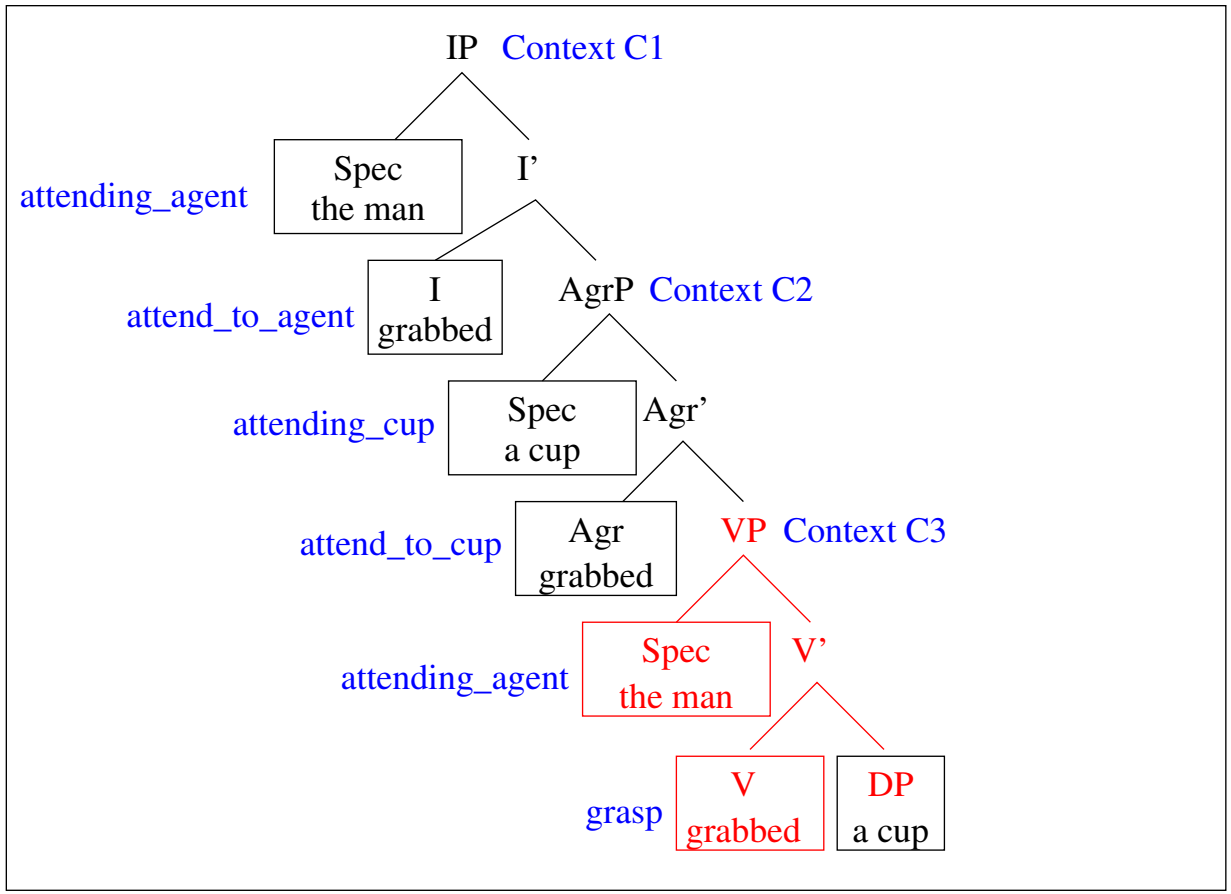
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



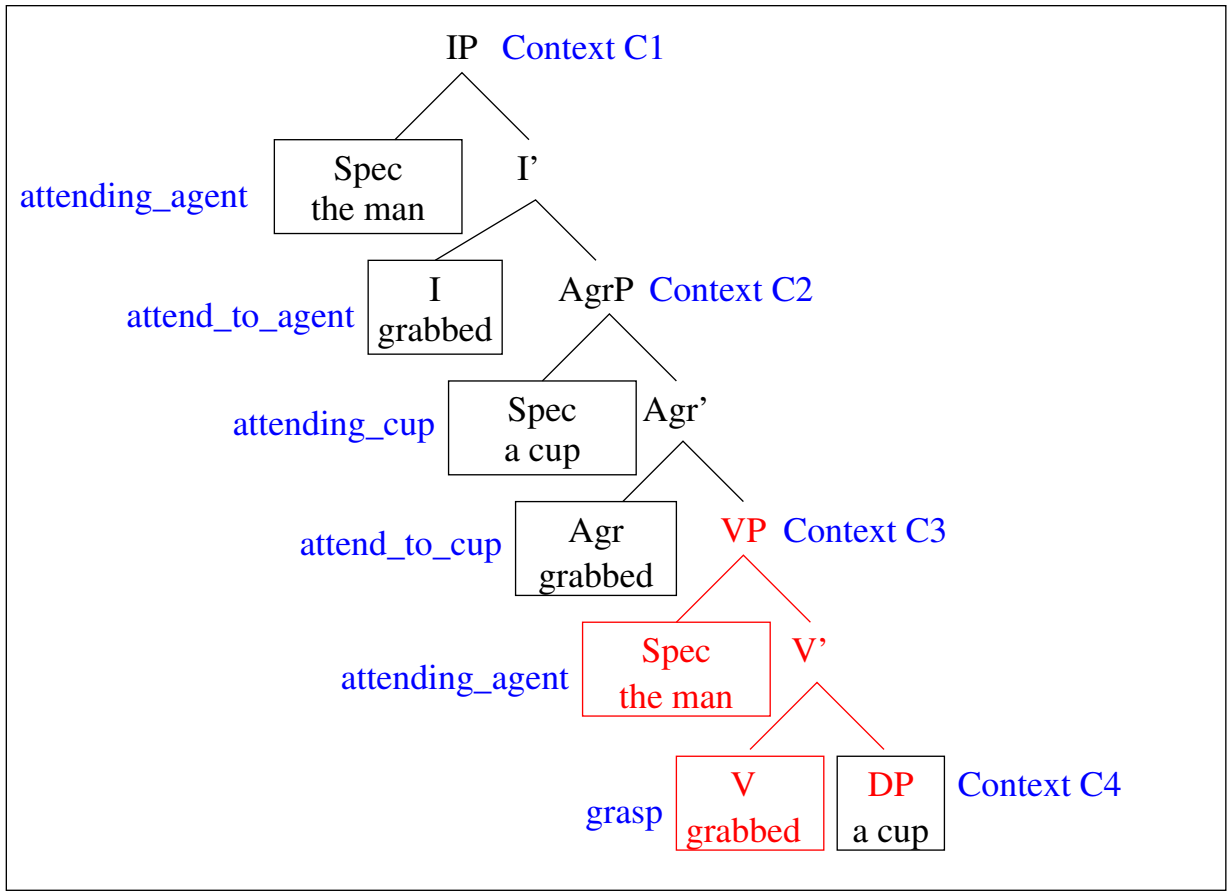
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



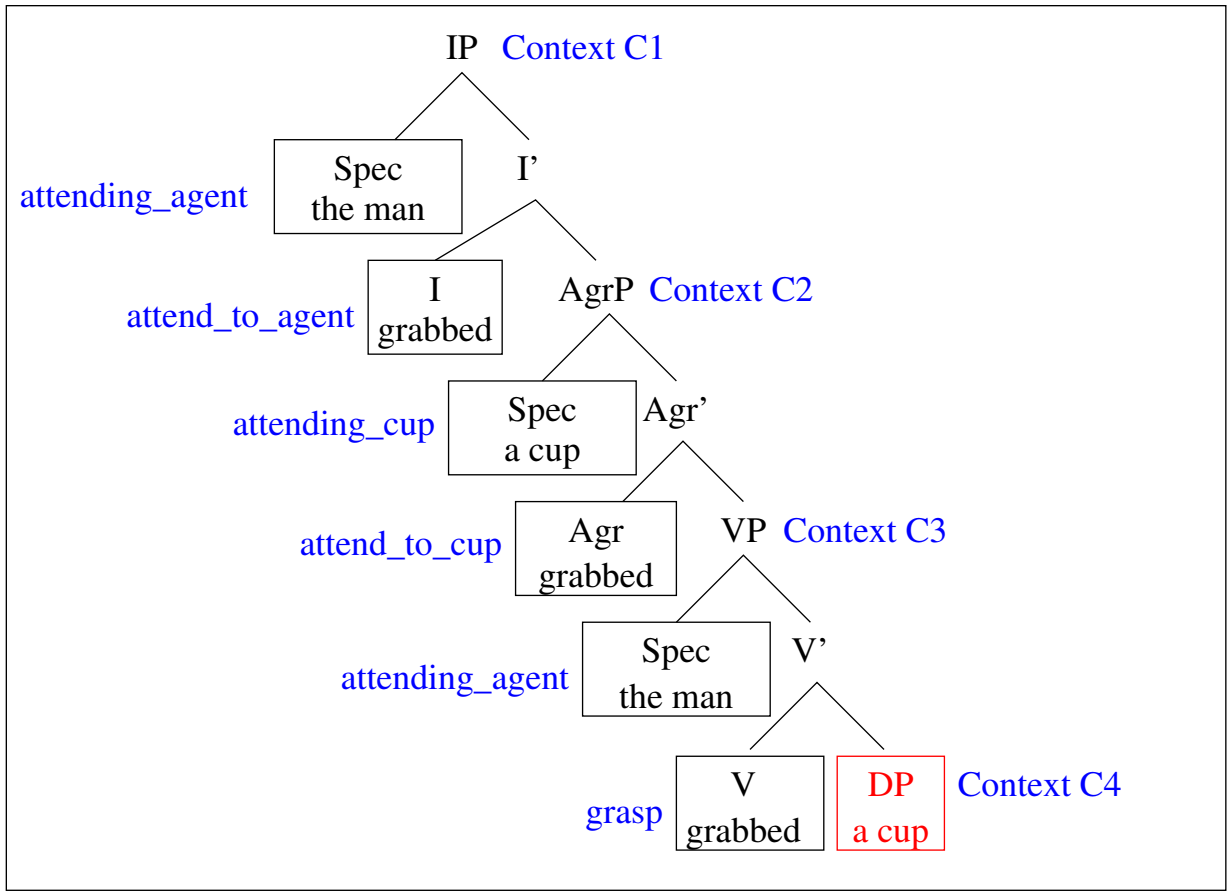
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



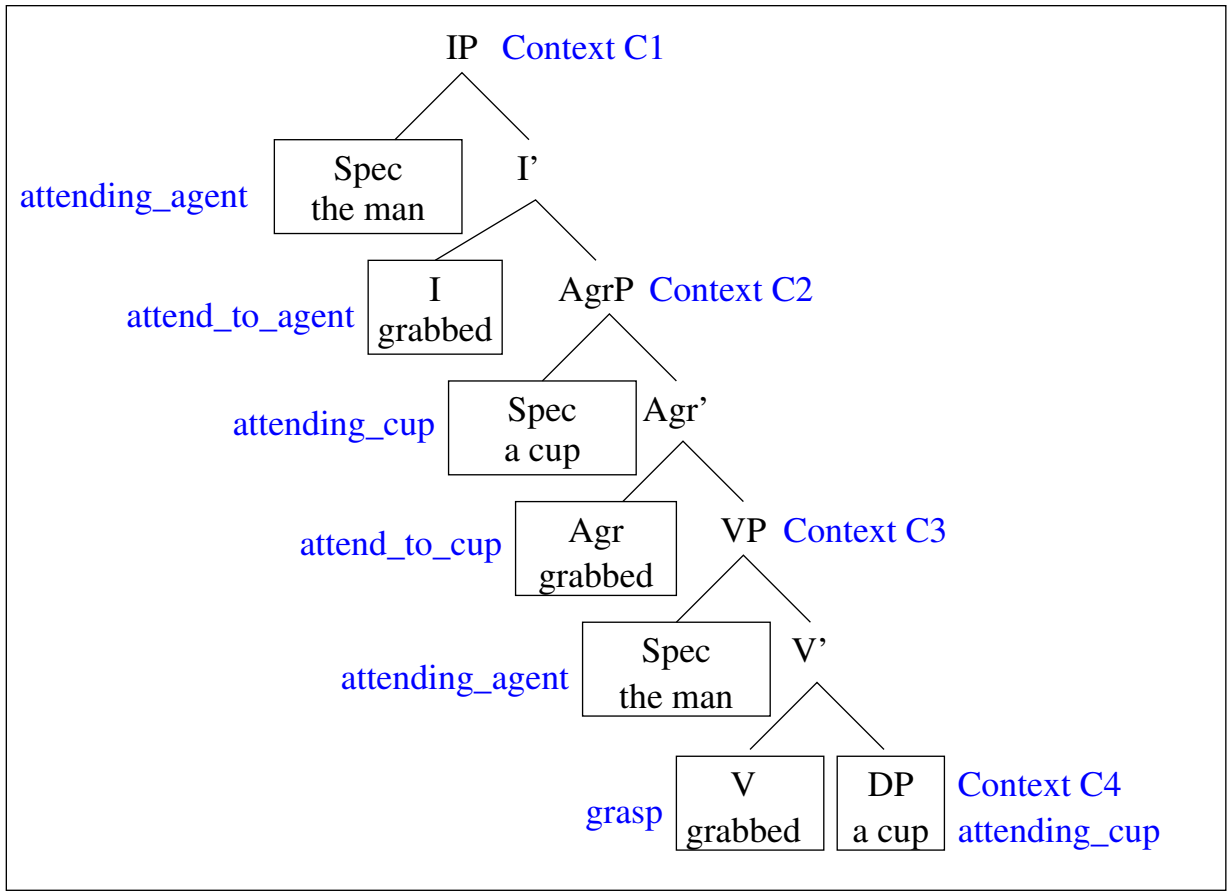
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



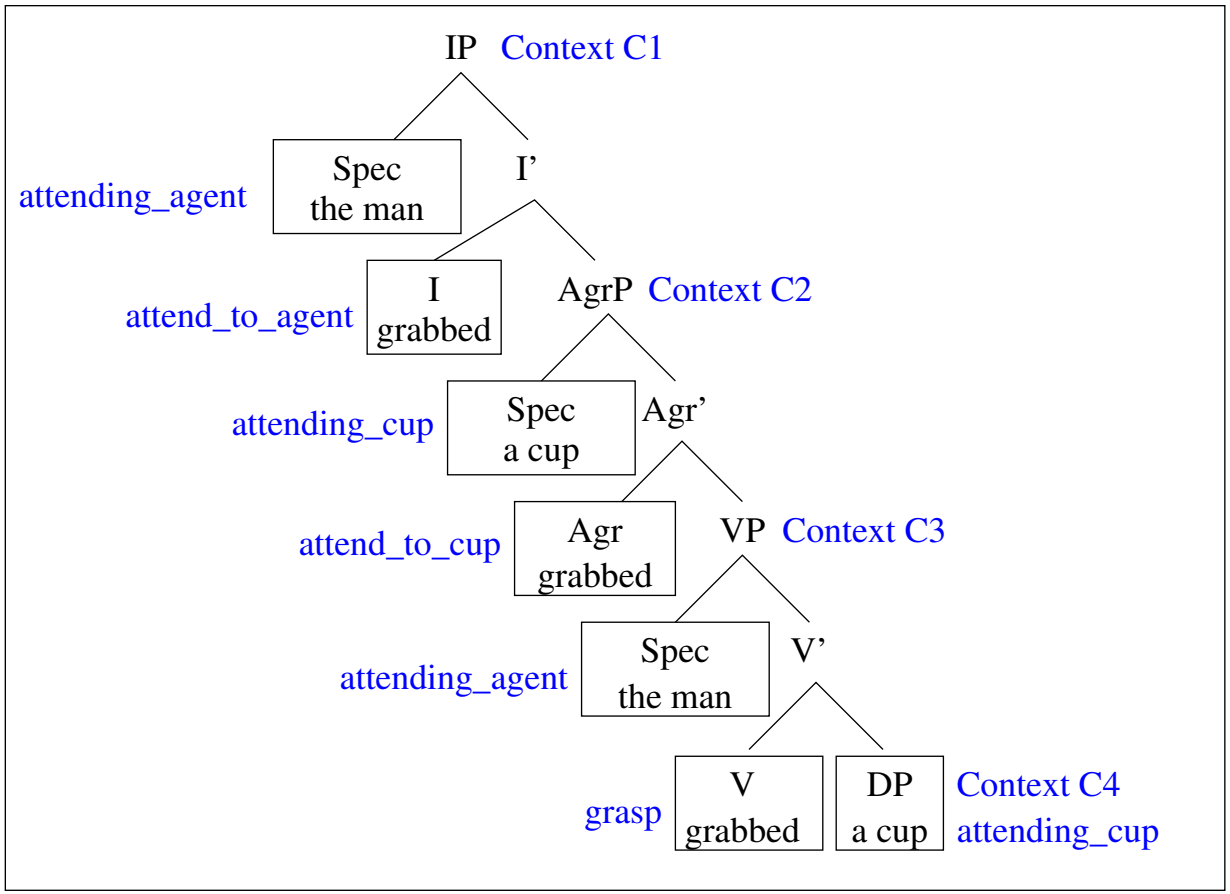
SM interpretation of the LF of *The man grabbed a cup*

The four LF projections map onto the four stages of the SM sequence.



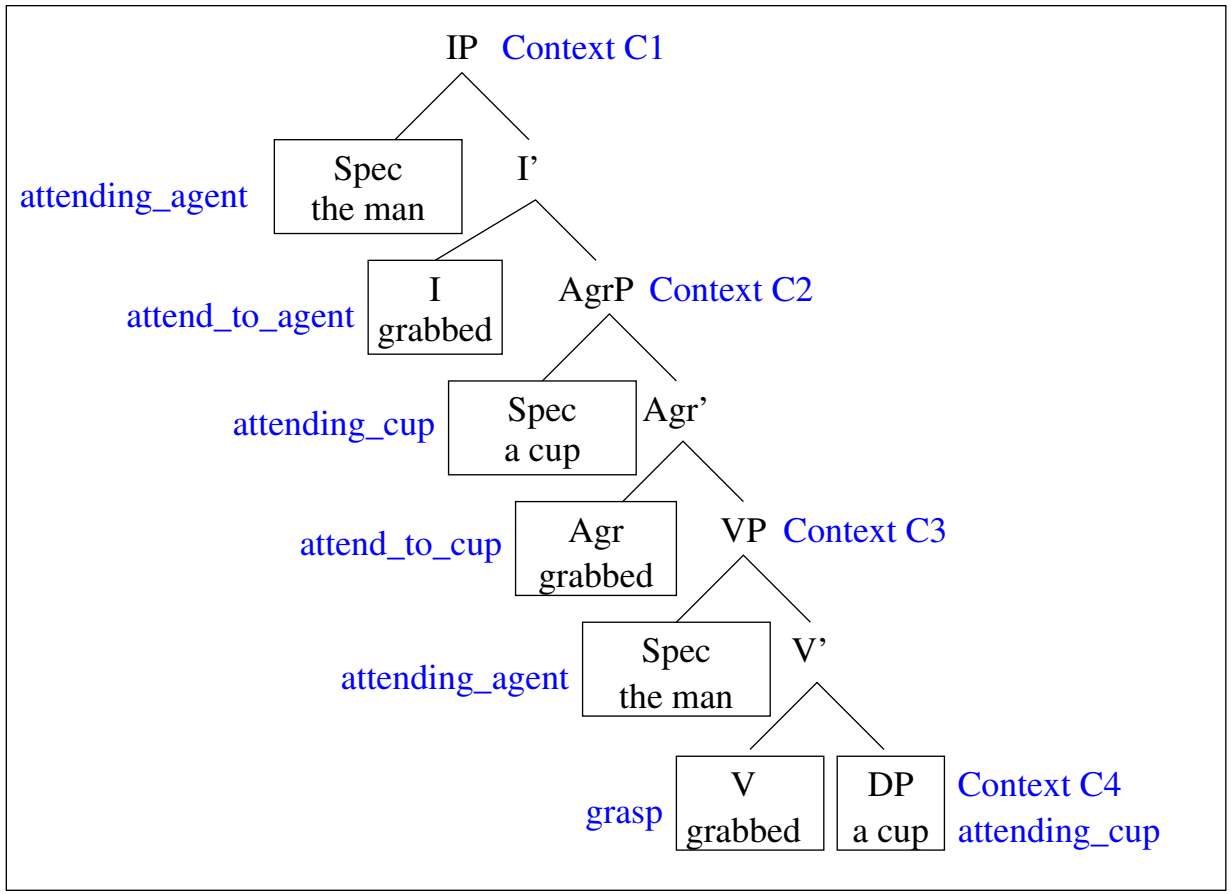
SM interpretation of the LF of *The man grabbed a cup*

So: what's the SM characterisation of movement?



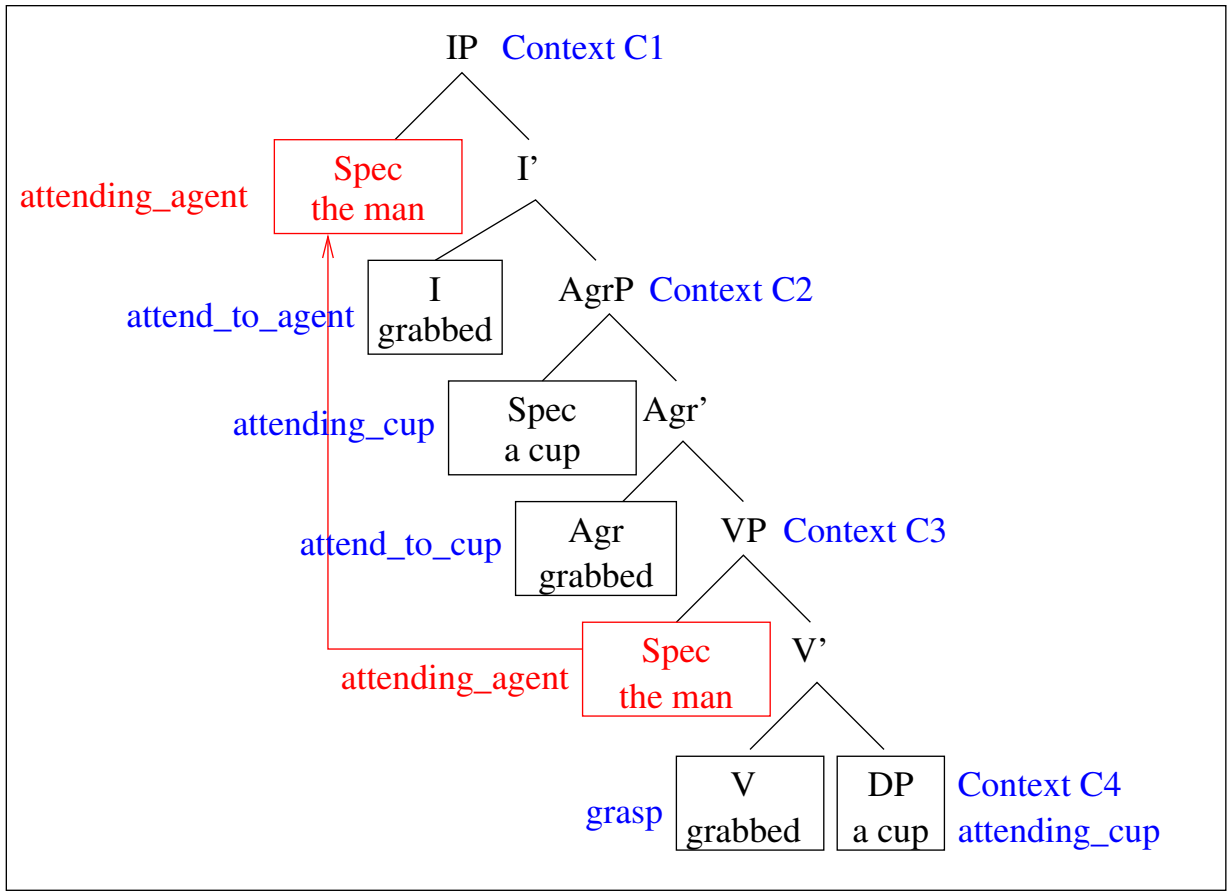
SM interpretation of the LF of *The man grabbed a cup*

DP raising reflects operations of *re-attention to agent and patient*.



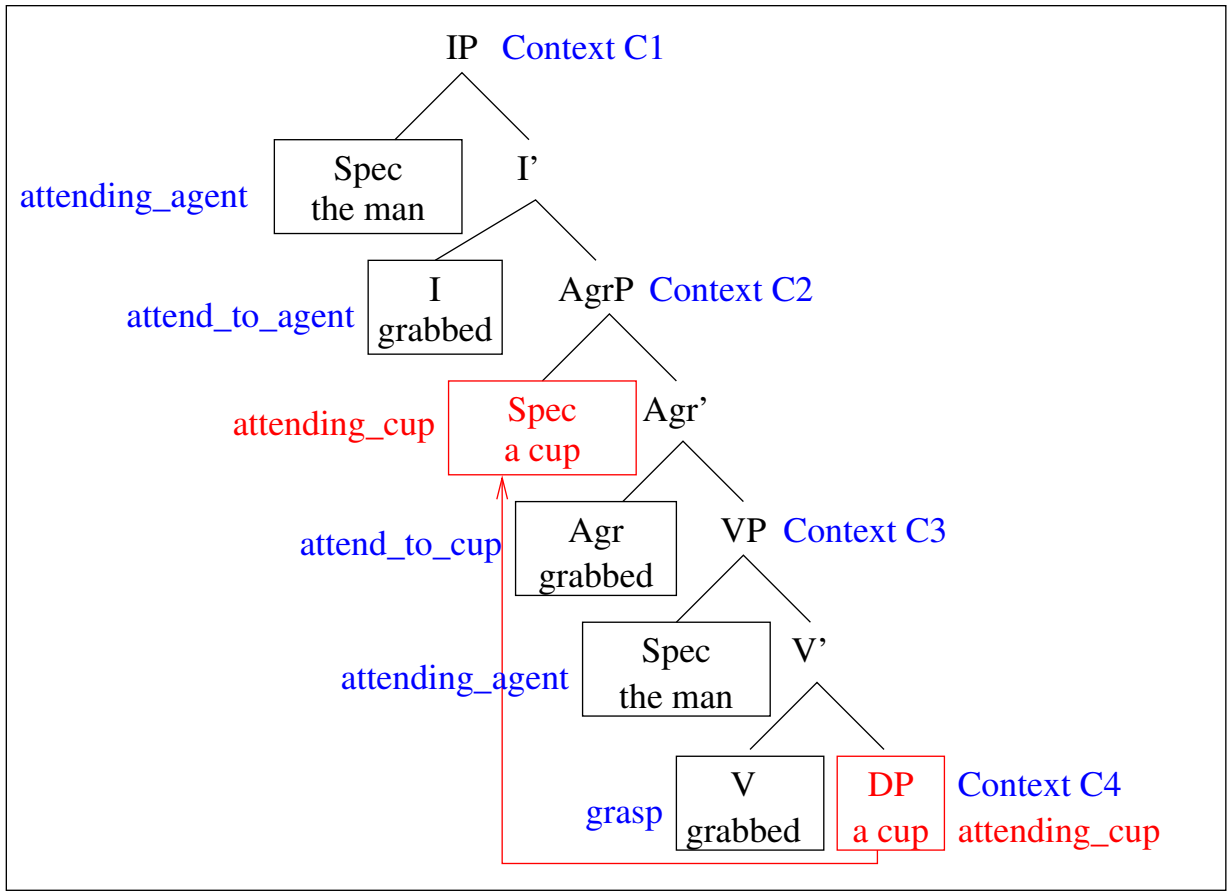
SM interpretation of the LF of *The man grabbed a cup*

DP raising reflects operations of *re-attention to agent and patient*.



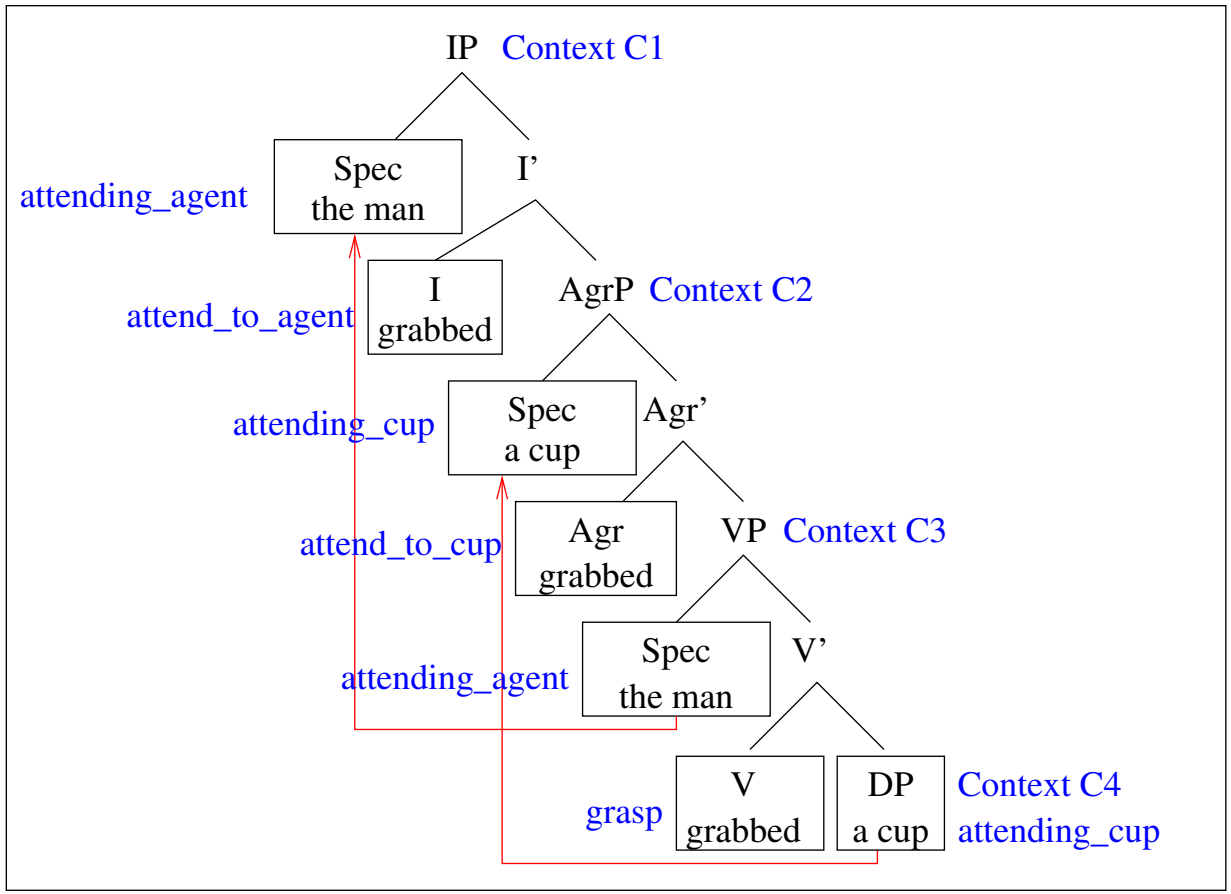
SM interpretation of the LF of *The man grabbed a cup*

DP raising reflects operations of *re-attention to agent and patient*.



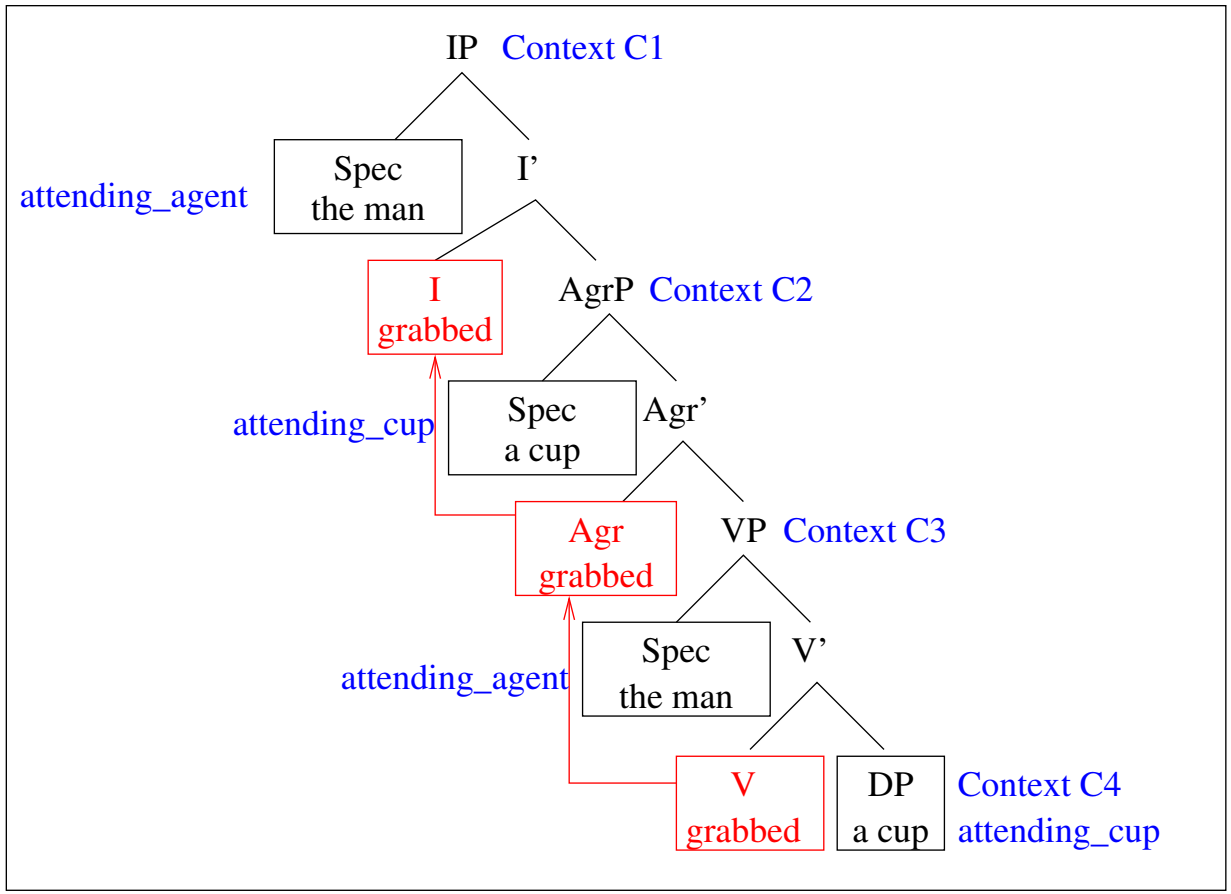
SM interpretation of the LF of *The man grabbed a cup*

SM interpretation of 'DPs must raise to check Case':
Objects must be attended to before being involved in action monitoring.



SM interpretation of the LF of *The man grabbed a cup*

There's a neat SM interpretation of V-Agr-I raising too.



Language processing and syntactic development

Interpreting LF as describing a SM sequence allows us to think of Minimalism as a processing model, as well as just a model of ‘syntactic competence’.

Language processing and syntactic development

Interpreting LF as describing a SM sequence allows us to think of Minimalism as a processing model, as well as just a model of 'syntactic competence'.

E.g. there's a natural model of sentence generation:

- An agent generating a (concrete) sentence internally rehearses a SM sequence stored in working memory, producing words as behavioural side-effects.

Language processing and syntactic development

Interpreting LF as describing a SM sequence allows us to think of Minimalism as a processing model, as well as just a model of 'syntactic competence'.

E.g. there's a natural model of sentence generation:

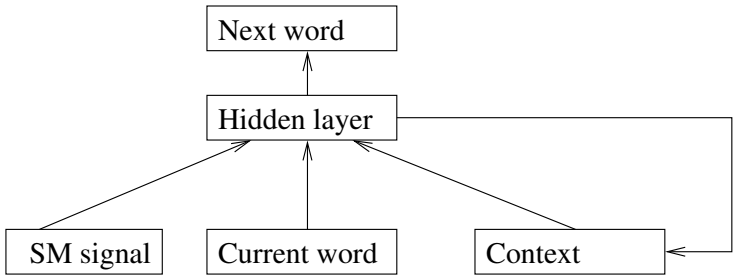
- An agent generating a (concrete) sentence internally rehearses a SM sequence stored in working memory, producing words as behavioural side-effects.

And of syntactic development:

- The agent learns which SM signals to 'read out' as words, and which to suppress.

Language processing and syntactic development

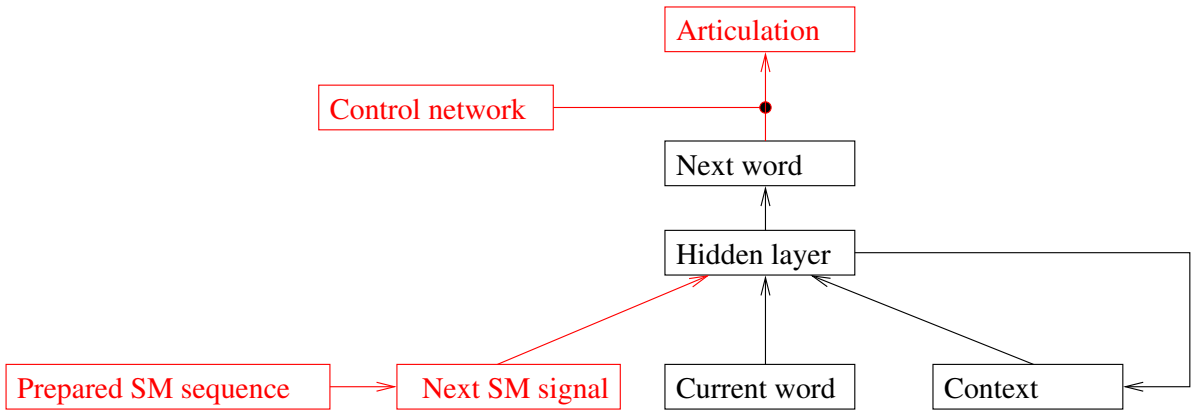
I have developed a neural network model of this generation/learning mechanism, which can handle idioms as well as productive syntactic constructions. (Joint work with Martin Takac and Lubica Benuskova.)



The model should look familiar to connectionist linguists.

Language processing and syntactic development

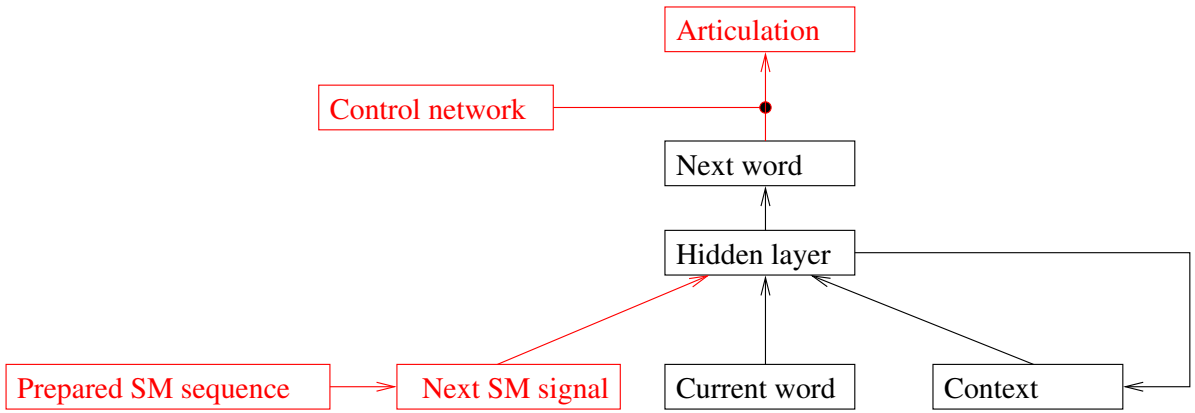
I have developed a neural network model of this generation/learning mechanism, which can handle idioms as well as productive syntactic constructions. (Joint work with Martin Takac and Lubica Benuskova.)



The model should look familiar to connectionist linguists.

Language processing and syntactic development

I have developed a neural network model of this generation/learning mechanism, which can handle idioms as well as productive syntactic constructions. (Joint work with Martin Takac and Lubica Benuskova.)



The model should look familiar to connectionist linguists. But it can also be interpreted by Minimalist linguists as a model of how speakers learn to set 'UG parameters' governing word order.

Beyond grabbing cups

I've just finished a book, called *Sensorimotor cognition and natural language syntax*, which goes into detail about all the models just described. (See my publications webpage.)

Beyond grabbing cups

I've just finished a book, called *Sensorimotor cognition and natural language syntax*, which goes into detail about all the models just described. (See my publications webpage.)

The book also extends the SM interpretation of LF to several other constructions:

- The internal structure of noun phrases
- Predicative clauses
- Quantification
- Relative clauses.

Implications for an account of language evolution

Implications for an account of language evolution

Basic proposal: a lot of the syntax of concrete sentences is actually sensorimotor in origin.

Implications for an account of language evolution

Basic proposal: a lot of the syntax of concrete sentences is actually sensorimotor in origin.

A possible timeline for language evolution:

Implications for an account of language evolution

Basic proposal: a lot of the syntax of concrete sentences is actually sensorimotor in origin.

A possible timeline for language evolution:

- Early hominins have a fairly modern SM system, including a mechanism for storing and replaying SM sequences.

Implications for an account of language evolution

Basic proposal: a lot of the syntax of concrete sentences is actually sensorimotor in origin.

A possible timeline for language evolution:

- Early hominins have a fairly modern SM system, including a mechanism for storing and replaying SM sequences.
- A circuit evolves for mapping static SM signals to overt behaviours, allowing the production of simple word sequences.

Implications for an account of language evolution

Basic proposal: a lot of the syntax of concrete sentences is actually sensorimotor in origin.

A possible timeline for language evolution:

- Early hominins have a fairly modern SM system, including a mechanism for storing and replaying SM sequences.
- A circuit evolves for mapping static SM signals to overt behaviours, allowing the production of simple word sequences.
- A new circuit evolves, co-opting the simple language circuit and the SM sequence rehearsal mechanism, which generates word sequences *from SM sequences*.