

removes their clothes)
gram·mar /'græm·ər, ˌɡræm·ər/
about how words change

Computational Construction Grammar for Open Source Collaborative Grammar Development (part 3)

Remi van Trijp

Sony Computer Science Laboratory Paris
remi@csl.sony.fr

Katrien Beuls

VUB AI-Lab
katrien@ai.vub.ac.be

Our Dream

- Coding the grammars of the world
 - Open source development of grammars
 - Create challenges and “problem sets”
 - Create a pool of solutions and design patterns
 - Set up an international network



Challenge: Tense-Aspect-Modality

<--> Verb Phrase

- First challenge has been launched & is almost finished: the verb phrase
 - English
 - Dutch
 - Italian
 - Spanish
 - Catalan
 - Russian
 - ...
- Results will be publicly announced in September at the International Conference on Construction Grammar



Semantics-Driven Approach

- Every time they utter a sentence, speakers from almost every known language...
 - Locate situations in time
 - He sings. vs. He sang.
 - Indicate the flow of the situation
 - He sings. vs. He was singing.
 - Express the speaker's commitment
 - I will sing. vs. I might sing.
- Yet, Tense-Aspect-Modality is practically ignored in many language technologies

Semantics-Driven Approach

- Every time they utter a sentence, speakers from almost every known language...
 - Locate situations in time
 - He sings. vs. He sang.
 - Indicate the flow of the situation
 - He sings. vs. He was singing.
 - Express the speaker's commitment
 - I will sing. vs. I might sing.
- Yet, Tense-Aspect-Modality is practically ignored in many language technologies

Example from Google Translate

- Google Translate vs. Human Translator:
 - Ik had het moeten weten.
I should have known.
I should have known.
 - Ik had het kunnen weten.
I should have known.
I could have known.
 - Ik had het geweten kunnen hebben.
I should have known.
I should have been able to know.

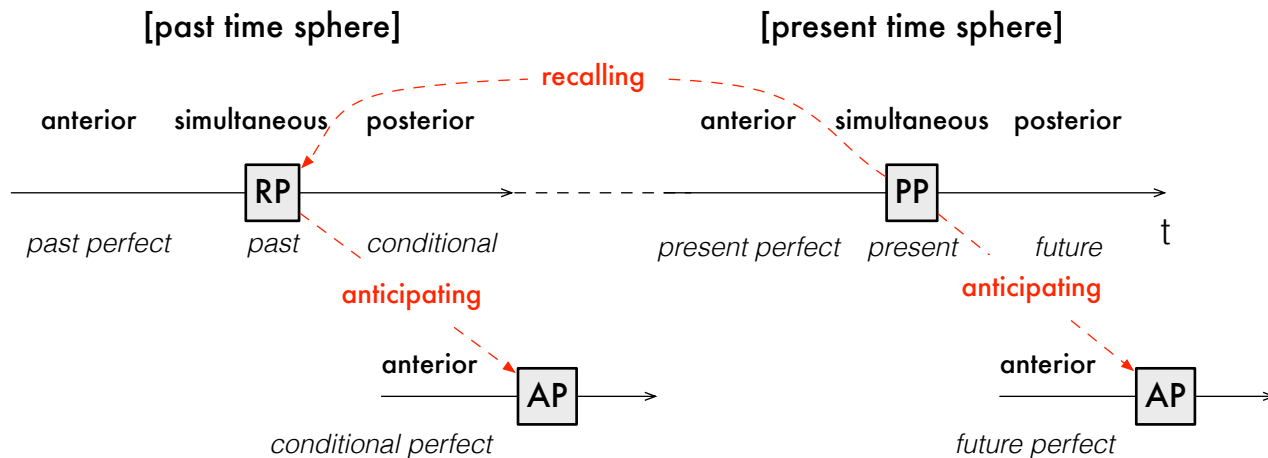
A DESIGN PATTERN FOR TENSE-ASPECT-MODALITY

Key Ideas

- Cognitive linguistics
 - Grammatical constructions
 - Distinction
- Formal grammar (e.g. LFG and HPSG)
 - TAM as feature-value pairs
 - Formal encoding in the verb phrase
- Computational linguistics
 - Efficient processing techniques
(head-corner processing)

Design Pattern

- Separation between “what” to say (conceptualization) and “how” to say it (grammar)



Head-driven grammar

- Main verbs “trigger” the grammar to build a verb phrase
- Grammatical items (auxiliaries, morphemes, ...) are then incrementally added to the verb phrase

**AN EXAMPLE FOR THE ENGLISH
VERB PHRASE**

English Verb Phrase

- Highly structured “pattern”

SEM: Modality/future > Perfect > Progressive > Lexical meaning

SYN: Modal-auxiliaries > Aux-HAVE > Aux-BE > Main verb

E.g. He must have been singing.

English Verb Phrase

- Most slots are “optional” ...
 - He sings.
 - He is singing.
 - He has sung.
 - He will sing.
 - He will be singing.
 - He will have sung.
 - He might have been singing.
 - ...

English Verb Phrase

- But not independent of each other...
 - Auxiliary BE puts the lexical verb in its “ing-form”
He sings. vs. He is singing.
 - Auxiliary HAVE puts the next word in its “past participle” form
He has sung. vs He has been singing.
 - Modal/Future auxiliaries put the next word in its “infinite” form
He will sing. vs He might be singing.

English Verb Phrase

- Inflection, which marks tense and agreement with the subject, is carried by the first form...
 - He sings.
 - He is singing.
 - He has been singing.

English Verb Phrase

- ... but in the case of negation, only auxiliaries can be inflected
 - He hasn't sung.
 - He wasn't singing.
 - (?) He sang not. (= extremely archaic)
- If there is no auxiliary, “do-support” is called
 - He didn't sing.

English Verb Phrase

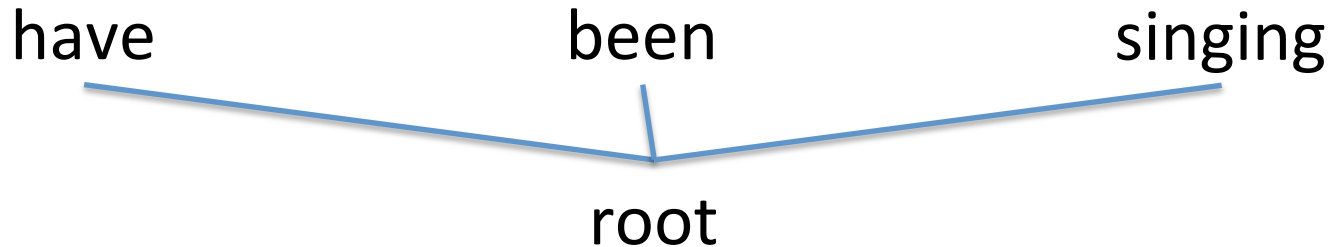
- The first auxiliary can be disconnected from the rest of the VP in questions
 - Who sang?
 - What were you singing?
- If there is no auxiliary, DO-support is called again
 - What did you sing?

Challenges

- Handle uncertainty about
 - Which slots are going to be filled
 - Which form must express tense and agreement
- Capture the systematicity of the VP
 - Avoid listing all possible combinations
 - Make the same system compatible with simple declarative utterances as with questions, negated utterances, and so on

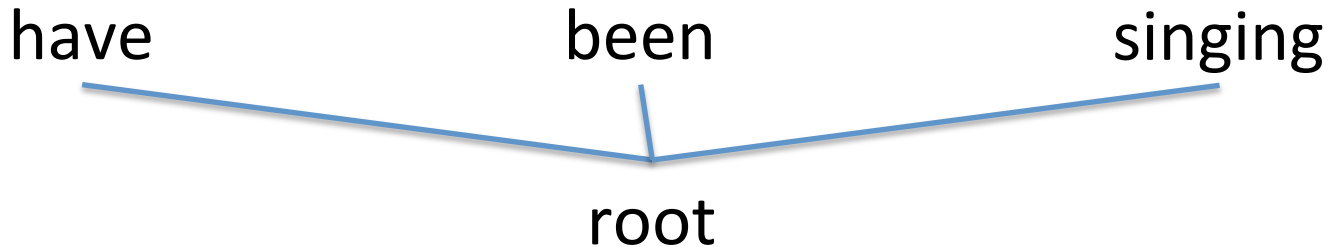
An FCG implementation: Parsing

- Morphological and lexical constructions analyze word forms
- Two classes of verbs: auxiliaries and lexical verbs

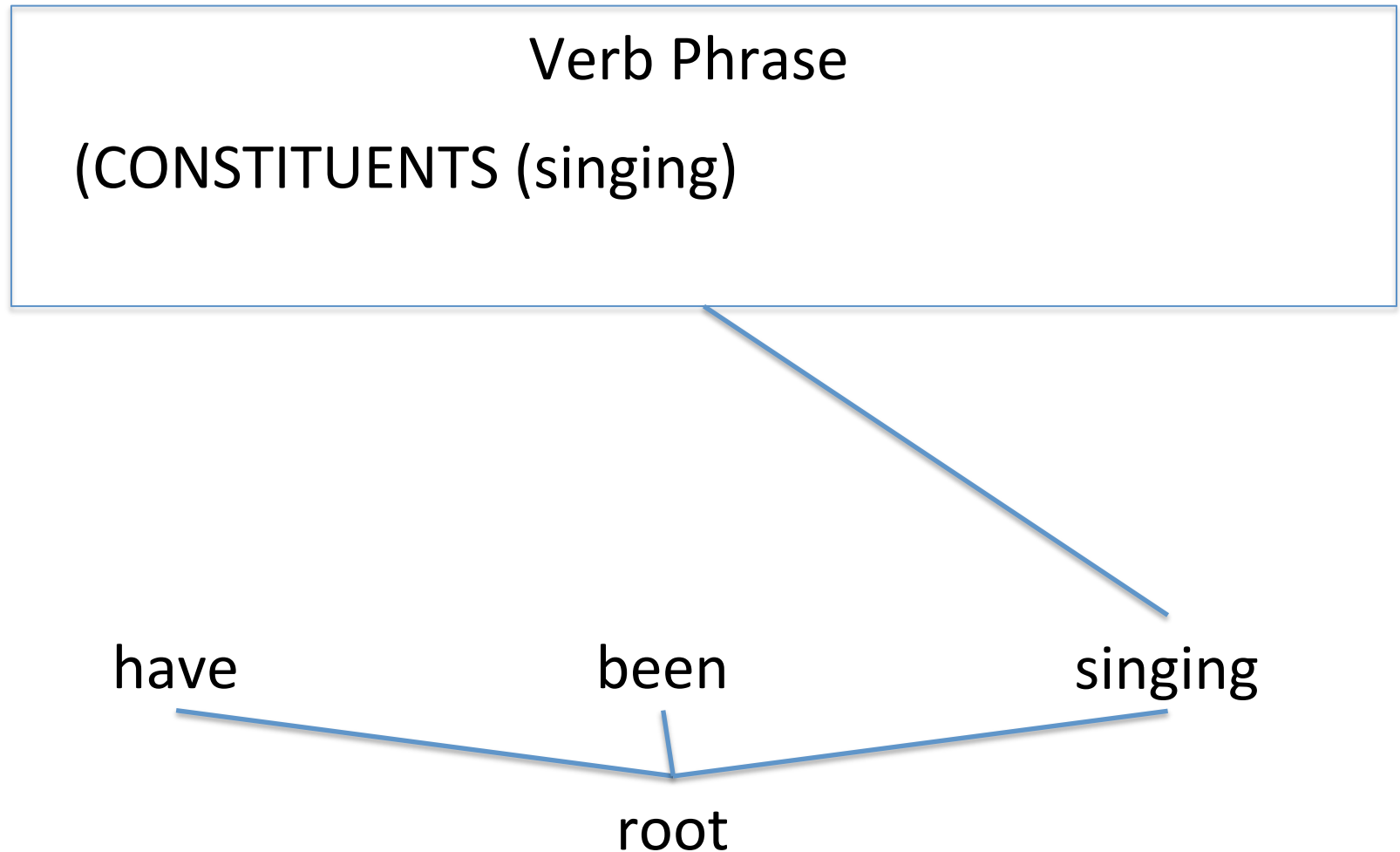


An FCG implementation: Parsing

- Verb phrase (VP) construction is triggered by the lexical verb form “singing”
- VP keeps track of its constituents



An FCG implementation: Parsing



An FCG implementation: Parsing

Verb Phrase
(CONSTITUENTS (**been** singing))

Other constituents
subscribe themselves
to the VP

have

been

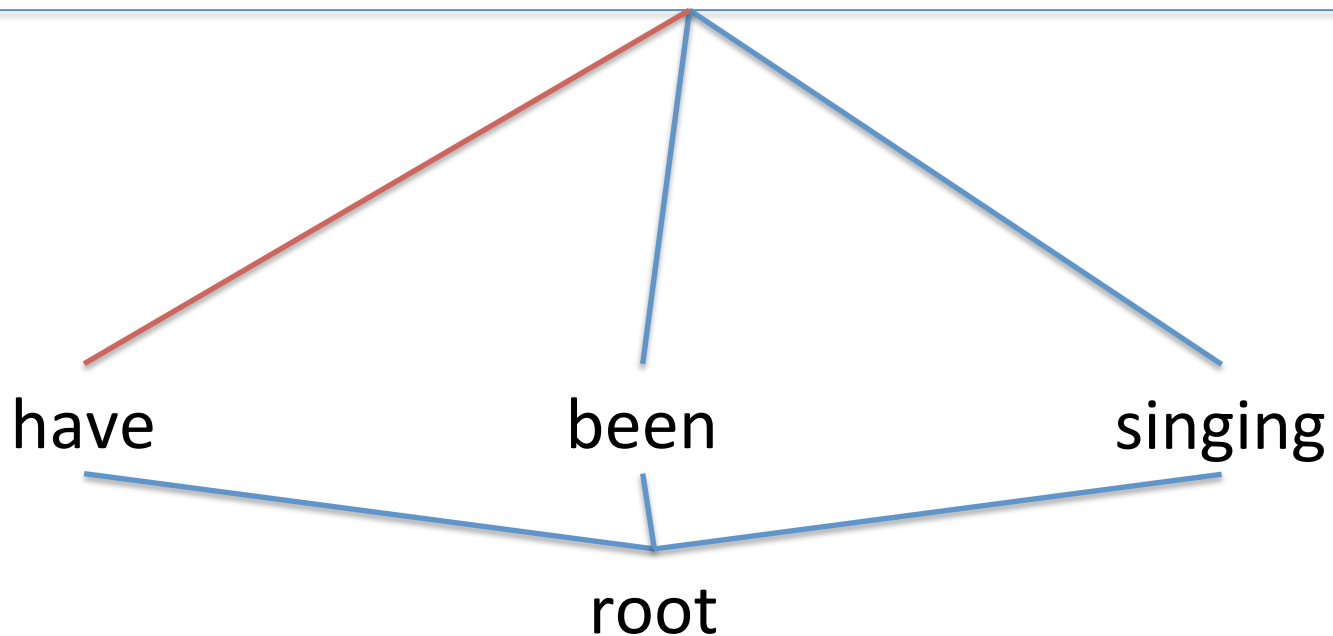
singing

root

```
graph TD; VP["Verb Phrase<br/>(CONSTITUENTS (been singing))"] ---|red line| been[been]; have[have] ---|blue line| root[root]; singing[singing] ---|blue line| root; root ---|blue line| been;
```

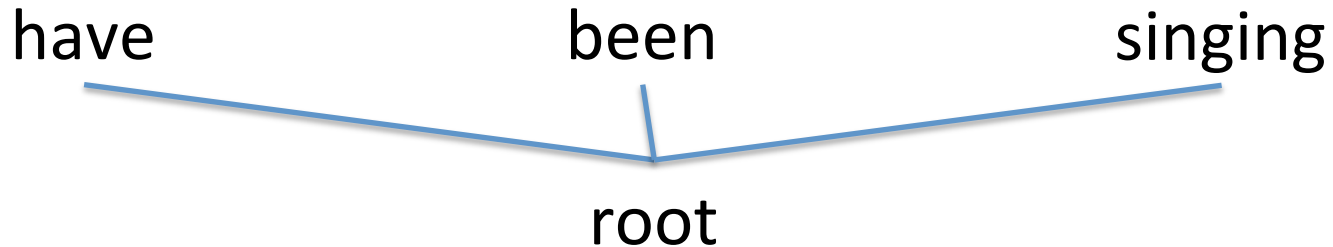
An FCG implementation: Parsing

Verb Phrase
(CONSTITUENTS (**have** been singing))

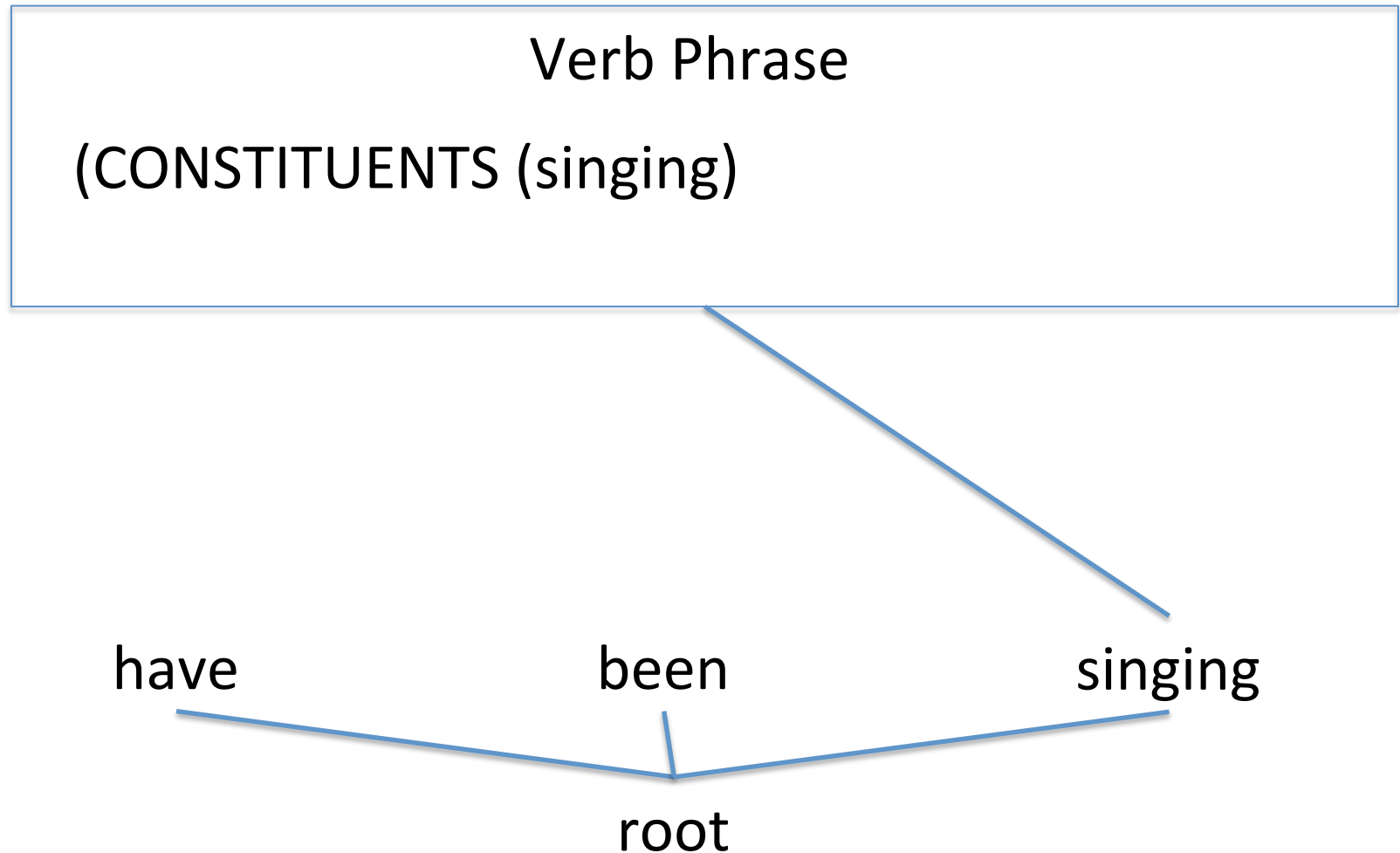


An FCG implementation: Production

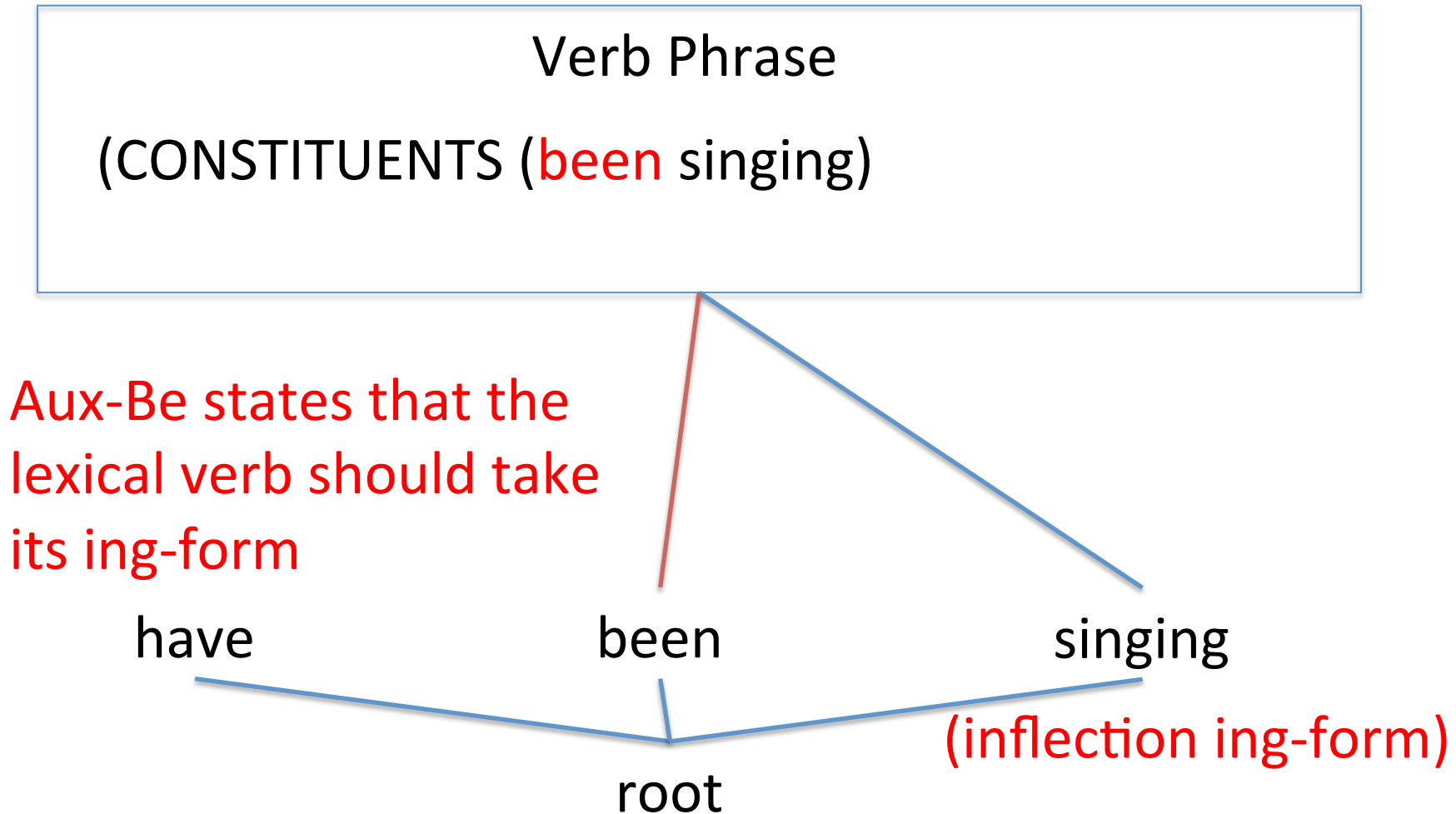
- Morphological and lexical constructions analyze word forms
- Two classes of verbs: auxiliaries and lexical verbs



An FCG implementation: Production



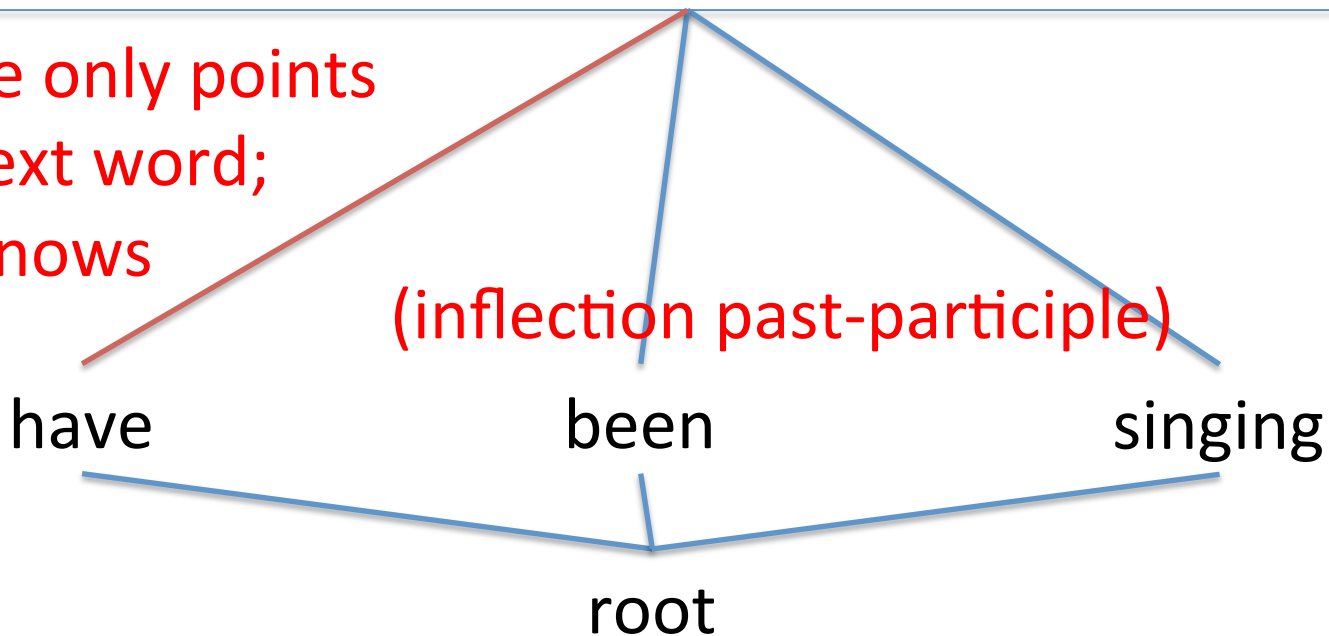
An FCG implementation: Production



An FCG implementation: Production

Verb Phrase
(CONSTITUENTS (**have** been singing))

Aux-have only points
to the next word;
the VP knows
which
one
this is.



Demonstration

TRANSLATION BETWEEN ENGLISH AND DUTCH

About the Grammars

- Basic grammar (handwritten) consists of only 7 grammatical constructions and a handful of lexical-morphological constructions
- Machine learning for automatically expanding the inventory
- Engineering solutions for optimizing processing
- Number of grammars grows linearly instead of exponentially because of separation between conceptualization and grammar
- Inference mechanisms to inform users about meanings that could not be translated, or meanings that are needed for proper choices

Future Challenges (Demo)

ROBUSTNESS AND FLUIDITY

Future Challenges

CONCLUSIONS AND DISCUSSION

Open Source Grammars

- Goal: an international network of grammar developers
 - Common set of challenges
 - Common pool of design patterns/solutions
 - Common list of conceptual primitives
 - Common repository of grammars
- Current phase: expansion of the FCG userbase

Open Source Grammars

- New training materials are in preparation
 - Online video lectures will be launched on 20 June 2014 at www.fcg-net.org
 - The FCG website will be transformed to a community site in the Fall of this year
- User feedback
 - Please contact us at info@fcg-net.org with feedback and comments about this tutorial and our (future) training materials

Open Source Grammars

- New training materials are in preparation
 - Online video lectures will be launched on 20 June 2014 at www.fcg-net.org
 - The FCG website will be transformed to a community site in the Fall of this year
- User feedback
 - Please contact us at info@fcg-net.org with feedback and comments about this tutorial and our (future) training materials

New Challenge

- New challenge: argument structure
 - Will be launched at the Workshop on Computational Construction Grammar
 - Satellite of the International Workshop on Construction Grammar 8
(3-6 September 2014, Osnabrueck, Germany)

Thank you!

- Questions?
 - info@fcg-net.org
 - remi@csl.sony.fr
 - katrien@ai.vub.ac.be
- Many thanks to Luc Steels, Emilia Casademont, Miquel Cornudella, Yana Knight, Michael Spranger, Paul Van Eecke and Pieter Wellens