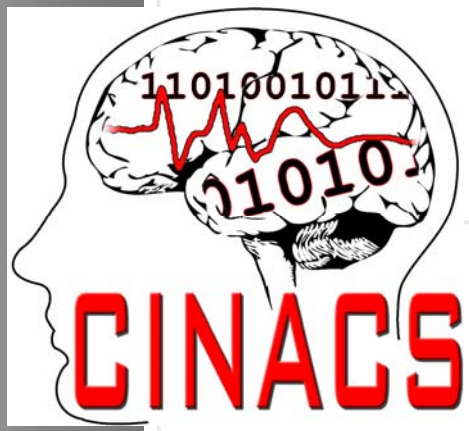
CINACS Graduate Research Group :: Hamburg University

Integrating Cross-Modal Context for Syntactic Disambiguation



03 September 2007
CINACS Summer School, Beijing
Patrick McCrae
Supervisors: Prof. Wolfgang Menzel, Prof. Maosong Sun

Before I start: Let's take a quick look back at last year's Summer School presentation

CINACS Graduate College :: Hamburg University

CINACS PhD Project
Cross-Modal Compensation in Syntactic Parsing

China Week: Hamburg, 26 Sept 2006
Patrick McCrae
Supervisor: Prof. Wolfgang Menzel

Presentation Structure

- Motivation
- Approach
- Research Status
- Future Directions

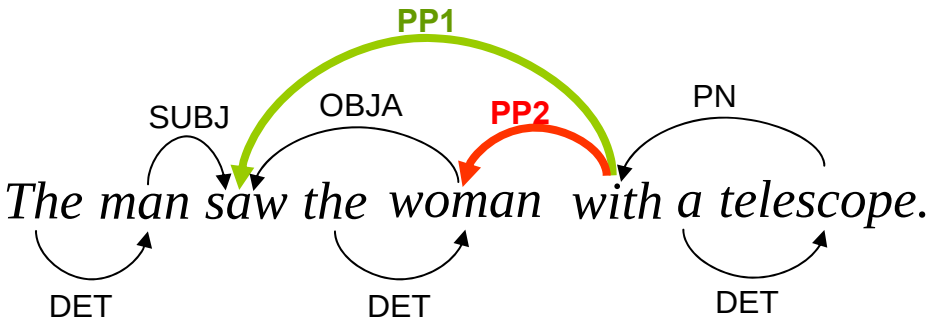
Presentation Structure

- Motivation
- Approach
- Research Status
- Future Directions

Cross-modal context provides extra-sentential knowledge in human sentence processing

Observation

- Ambiguity is omnipresent in natural language.
- Human communication still succeeds (in most cases).
- Human parsing is remarkably robust (quite in contrast to most NLP applications).



Conclusions

1. Humans integrate information from sources other than the linguistic utterance (**extra-sentential knowledge**).
2. Syntactic attachment preferences can be influenced by **cross-modal context**.

Crain & Steedman (1985): Context as constructed by the hearer influences grammaticality judgements

Context 1: Complement-inducing

A psychologist was counseling a married couple. One member of the pair was fighting with him but the other one was nice to him.

Context 2: Relative-inducing

A psychologist was counseling two married couples. One of the couples was fighting with him but the other one was nice to him.

Target Sentence	Context	Considered ungrammatical
Reduced Relative *The psychologist told the wife that he was having trouble with to leave her husband.*	Relative-inducing	22%
	Complement-inducing	50%
Complement *The psychologist told the wife that he was having trouble with her husband.*	Relative-inducing	54%
	Complement-inducing	12%

Conclusions

1. Grammaticality judgements – and hence syntactic processing – can be ***context-sensitive***.
2. An adequate model of human sentence processing needs to include ***context-dynamic preferences*** rather than static rules.

Tanenhaus et al. (1995): Visual context immediately affects the early stages of sentence processing in PP attachment

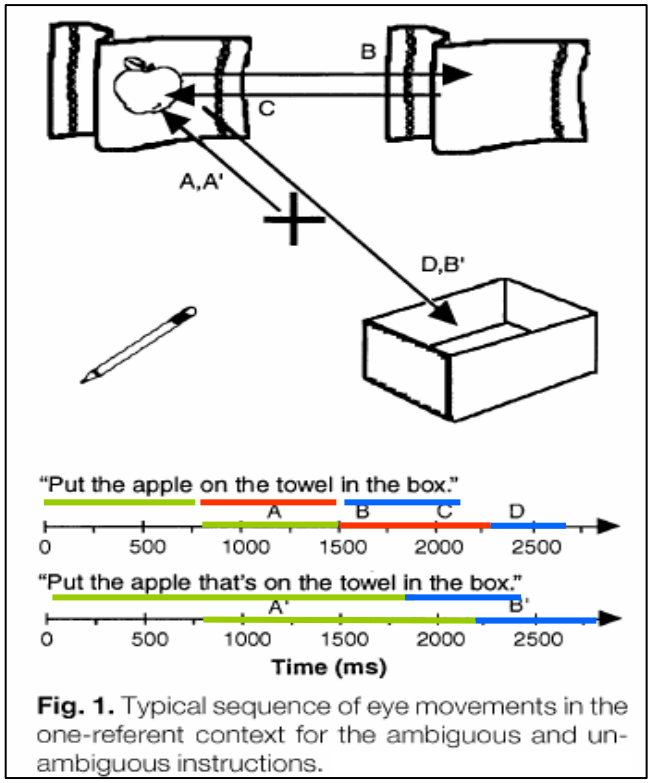


Fig. 1. Typical sequence of eye movements in the one-referent context for the ambiguous and unambiguous instructions.

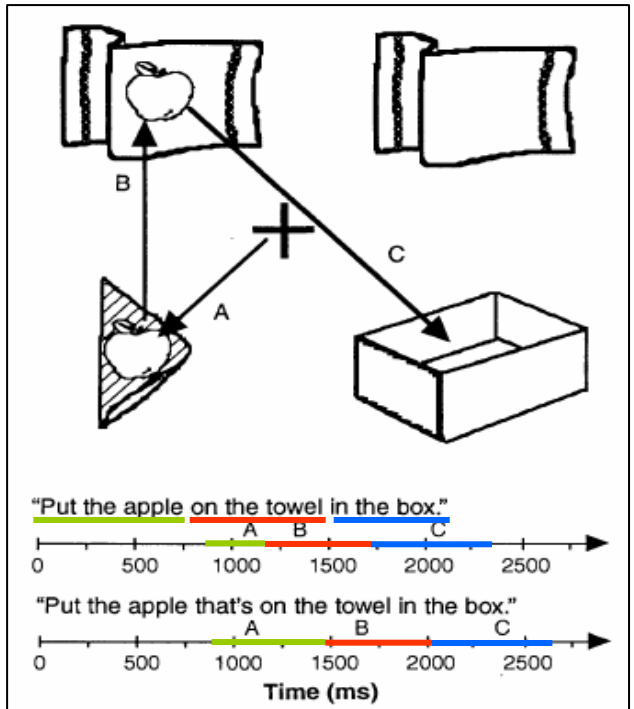


Fig. 2. Typical sequence of eye movements in the two-referent context.

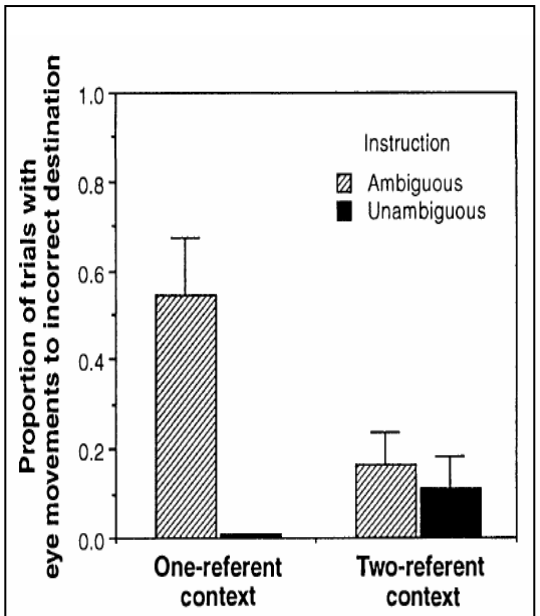


Fig. 3. Proportion of trials in which participants looked at the incorrect destination.

Conclusions

1. "(...) people seek to establish reference (…) during the ***earliest moments*** of linguistic processing."
2. "(…) referentially relevant ***nonlinguistic information immediately affects*** the manner in which the linguistic input is initially structured."

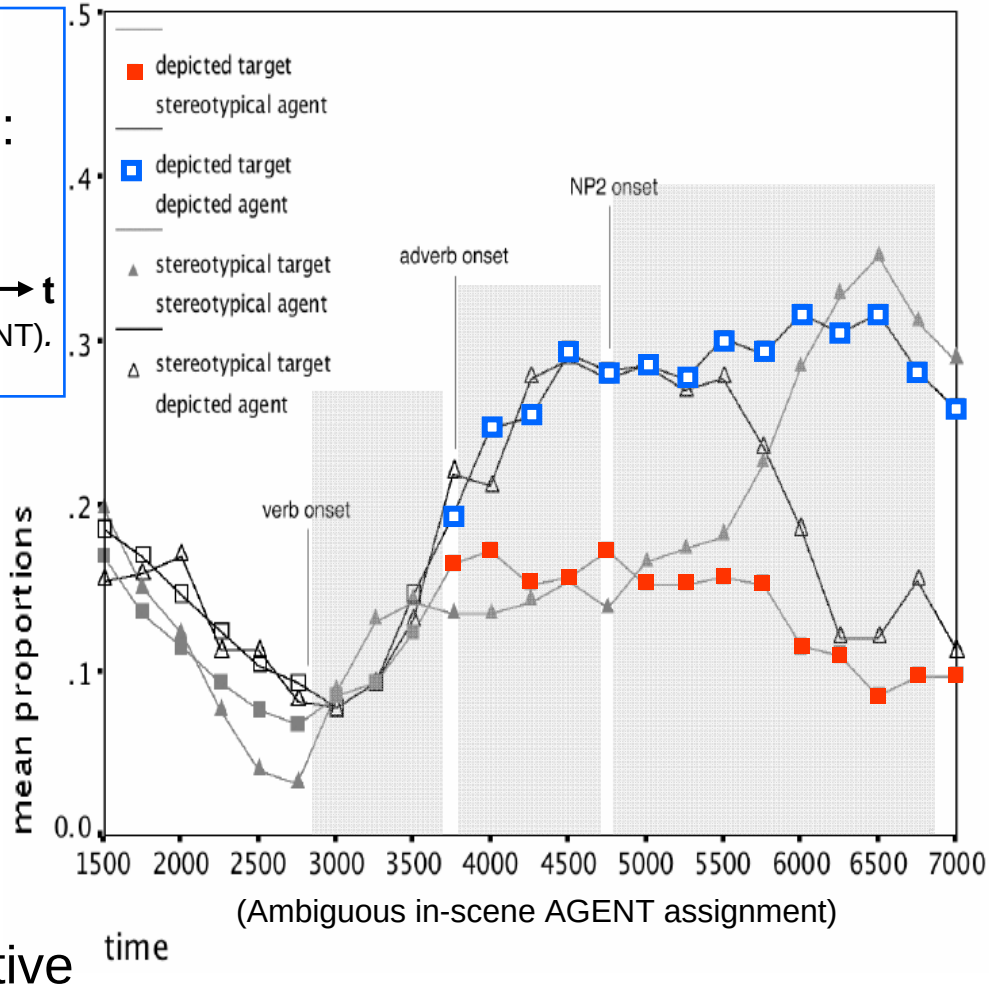
Knoeferle (2005): Visual context has a stronger influence on thematic role expectations than stored world-knowledge

Investigate the relative strength of visual context in AGENT anticipation:

Den Piloten bespitzelt gleich der Zauberer.

t

The pilot (PATIENT)*_spies on_soon_the wizard* (AGENT).



(Ambiguous in-scene AGENT assignment)

'spy-on' + scene → wizard

'spy-on' + world knowledge → detective

Presentation Structure

- Motivation
- **Approach**
- Research Status
- Future Directions

Overall Research Goals

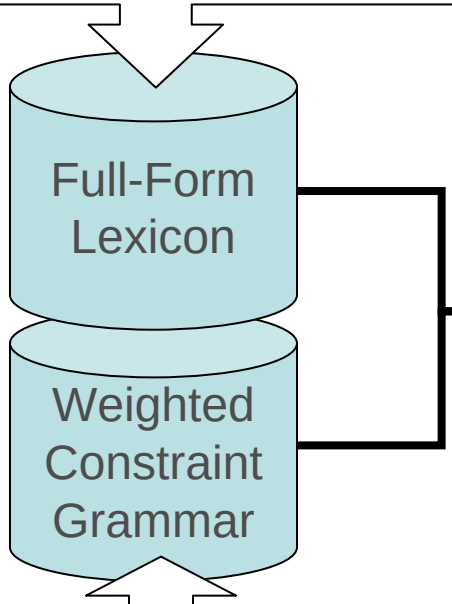
1. Achieve an understanding of the parameters involved in the cross-modal influences upon syntactic processing.

2. Design a computational model for the interaction between cross-modal context and language processing.

3. Implement the model in a parsing architecture capable of reproducing cross-modal context influences upon syntactic parsing.

The CDG parser permits to model syntactic preferences in a weighted constraint dependency grammar

```
Mann := [ cat: NN, base: Mann,
case: nom_dat_acc, person: third,
gender: masc, number: sg, sort:
human ];
```



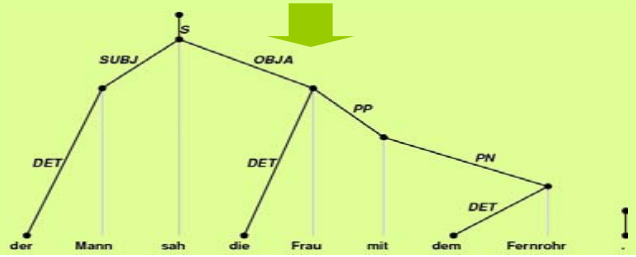

```
{X:SYN}: Subject_verb_agree: 0.1:
X.label = SUBJ
-> ( X^person = X@person
     & X^number = X@number
     & X@case = nom );
```

Der Mann sah die Frau mit dem Fernrohr.

The man saw the woman with the telescope.

CDG Key Features

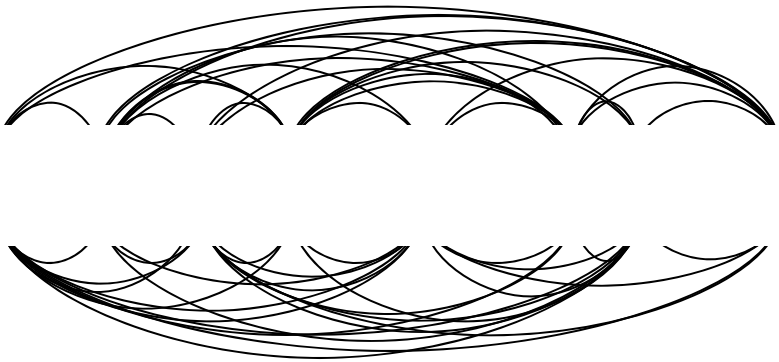
- Constraint-based rather than rule-based
- Models syntactic preferences via weighted constraints
- Extensible functional scope by integration of external plug-in components
- Full language coverage
- Highest reported accuracy on unrestricted written German input



Weighted constraint dependency parsing in CDG is a three-step process: Step 1

1. Creation of the Maximal Net

- Every word is connected to every other word in the sentence via labelled edges which represent the grammatical dependency.
- The number of connecting edges between a word pair corresponds to the number of possible edge labels in the syntax.



Weighted constraint dependency parsing in CDG is a three-step process: Step 2

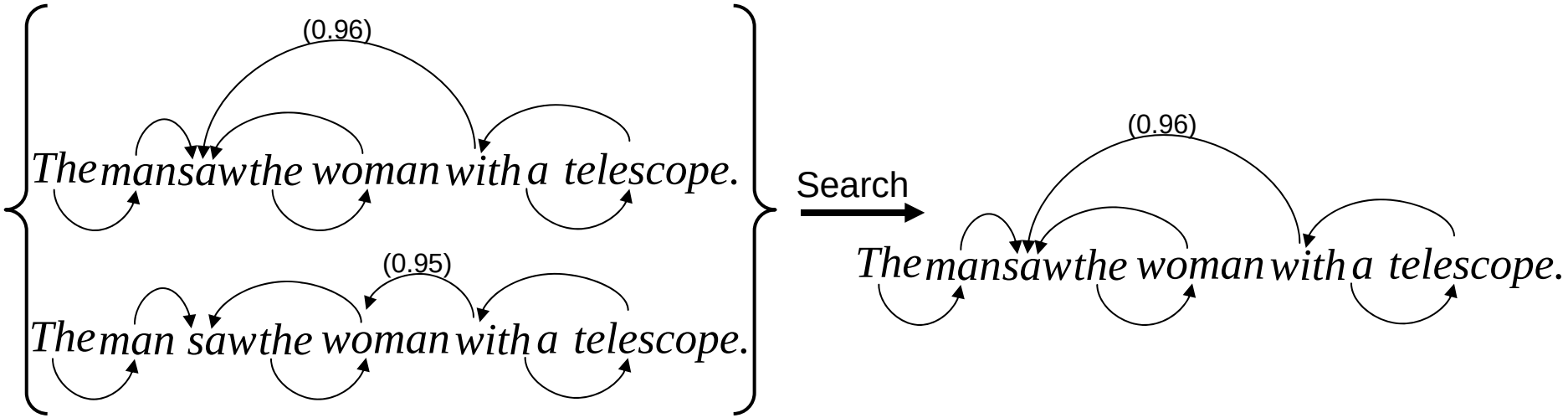
2. Test for Constraint Satisfaction

- Constraints are applied to single edges and edge pairs.
- Every edge receives a score according to the constraint weights given in the grammar.
- Edge connections sanctioned by the grammar receive low scores and will most likely not contribute to the optimal parse.

Weighted constraint dependency parsing in CDG is a three-step process: Step 3

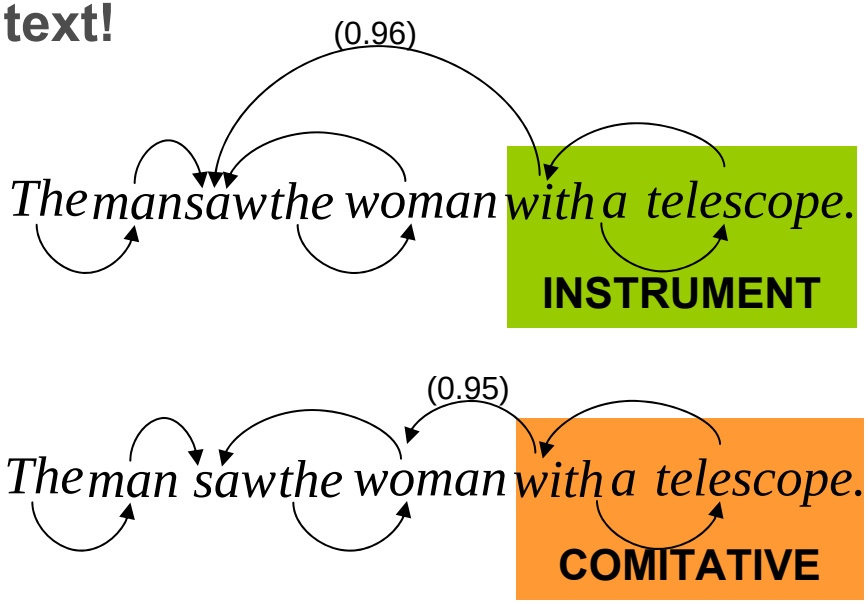
3. Search for the Optimal Solution

- A parse's overall score is the product of the weights of all constraints which have been applied to the sentence. The global optimum is the one with a score closest to 1.
- Due to the large size of the search space optimisation heuristics need to be applied. The global optimum may therefore not always be found.



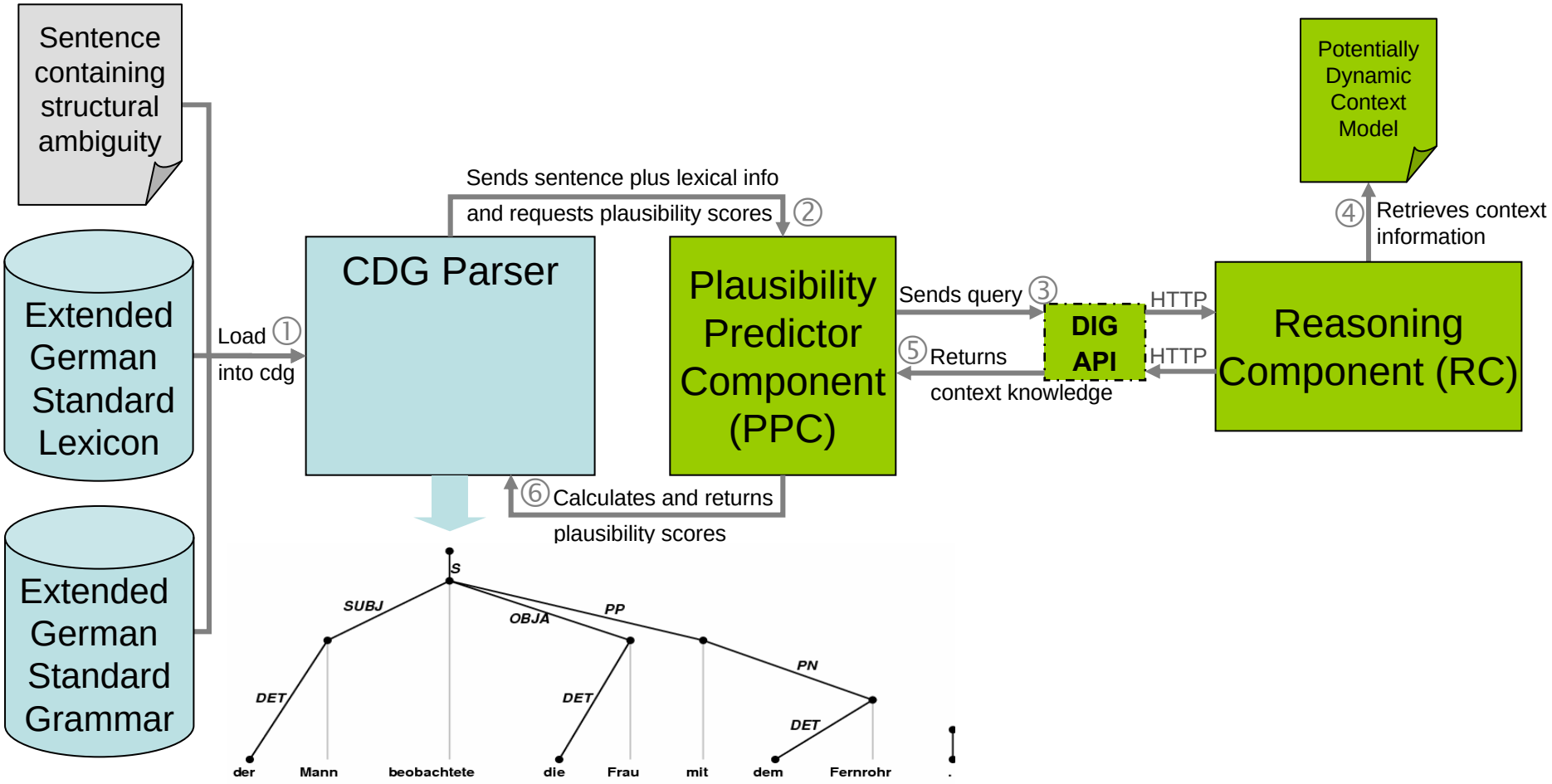
Key Challenge: context-sensitive disambiguation is *not* a matter of grammaticality alone but also of semantics

Both parses are grammatically correct. We need to detect whether one of the two may be favoured semantically by utterance context!

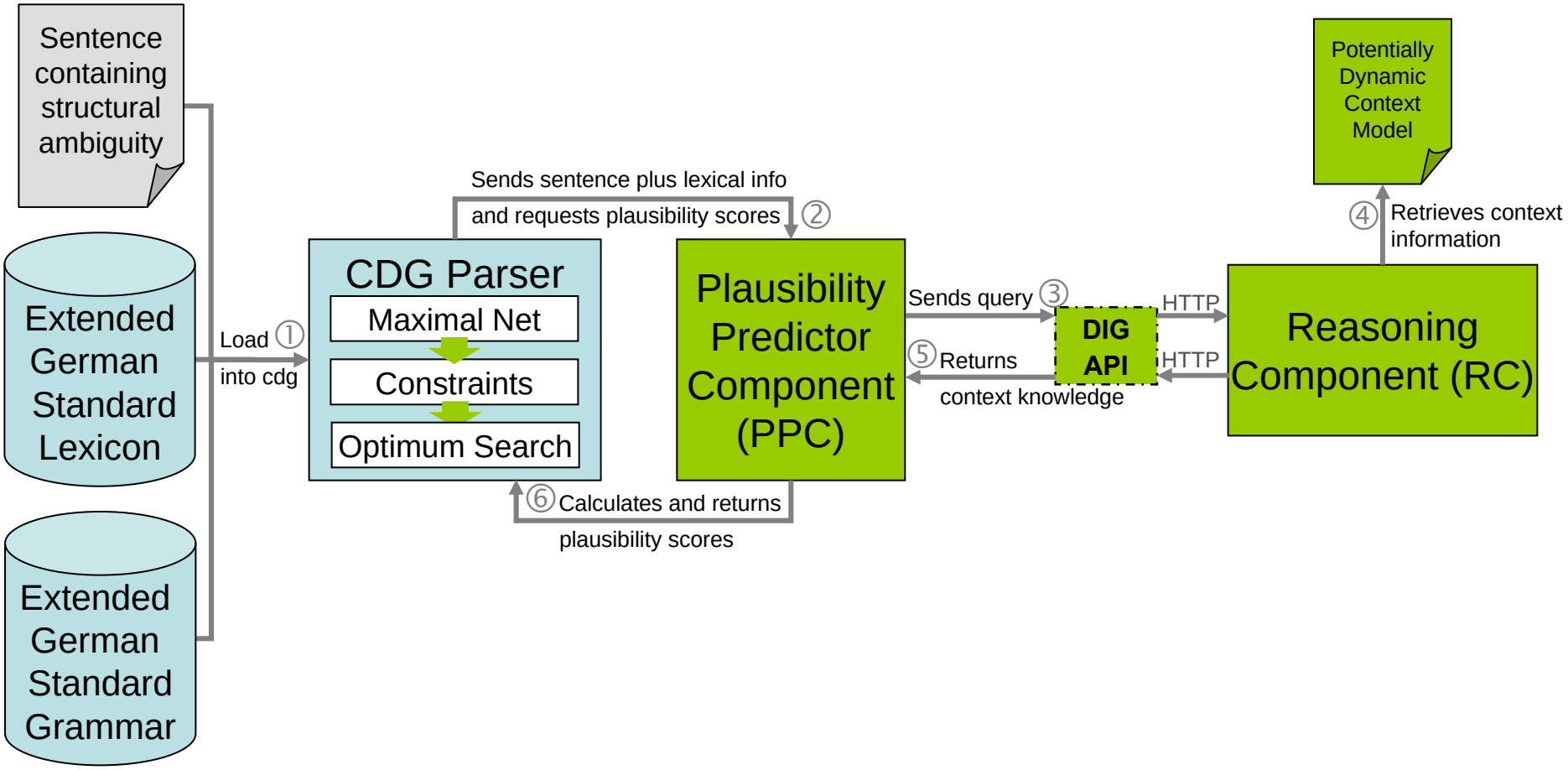


Working assumption: In sentence processing, humans use cross-modal context as one of several constraining factors to assess the plausibility of thematic role assignments.

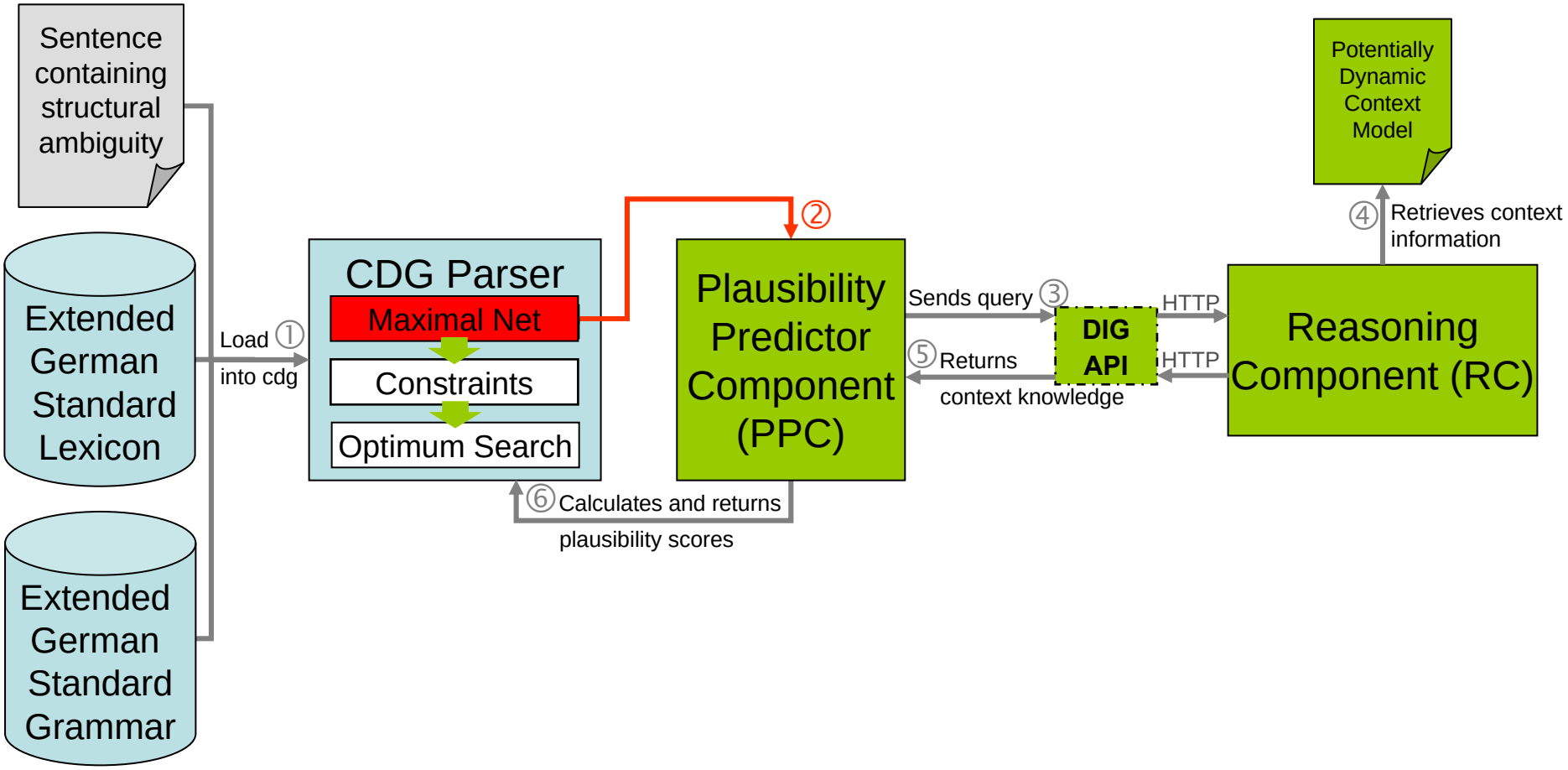
Putting it all together: We propose a cognitively motivated architecture to integrate cross-modal context into parsing



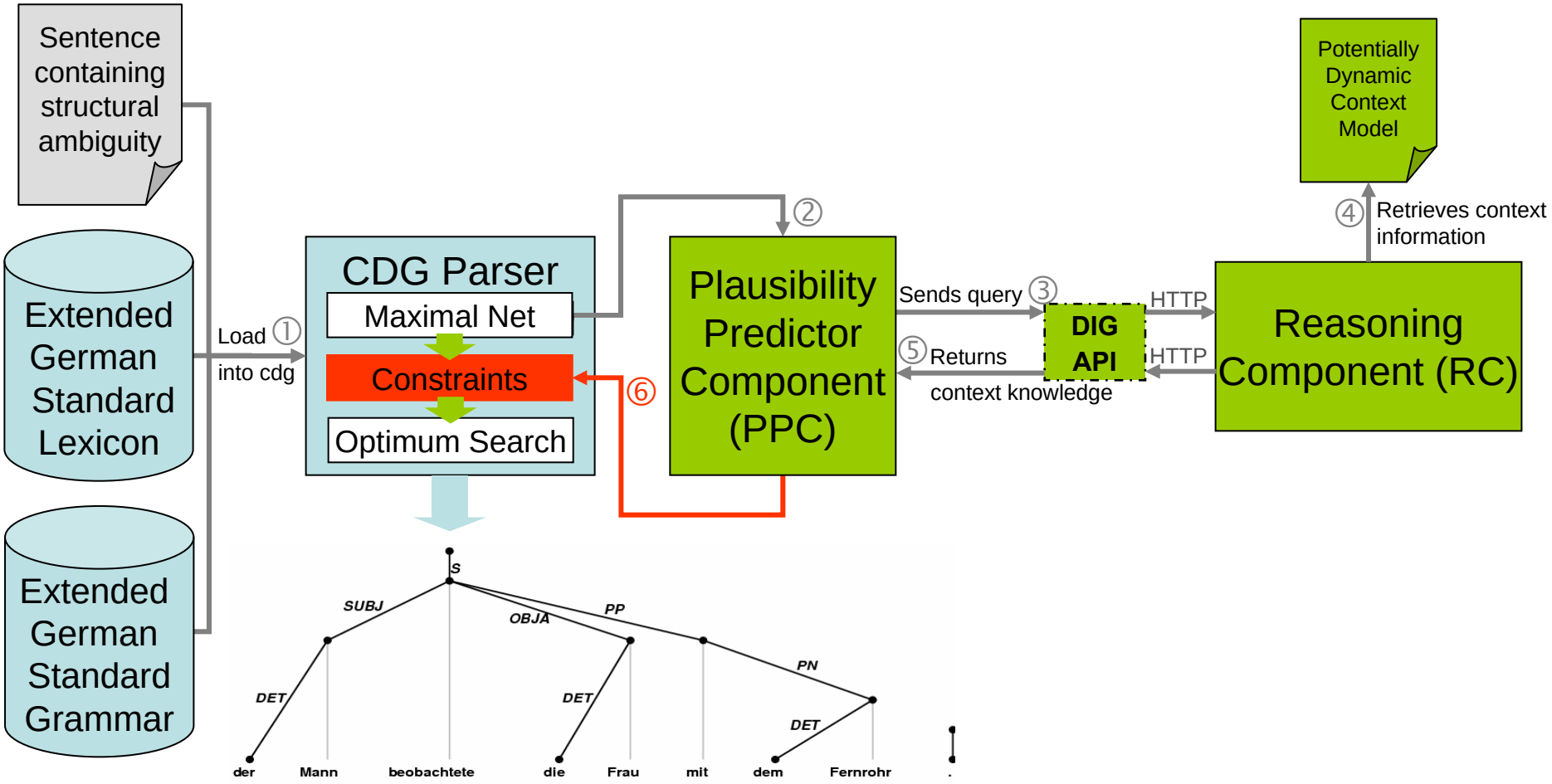
Parsing integration is achieved by obtaining plausibility predictions prior to checking for constraint satisfaction



Parsing integration is achieved by obtaining plausibility predictions prior to checking for constraint satisfaction



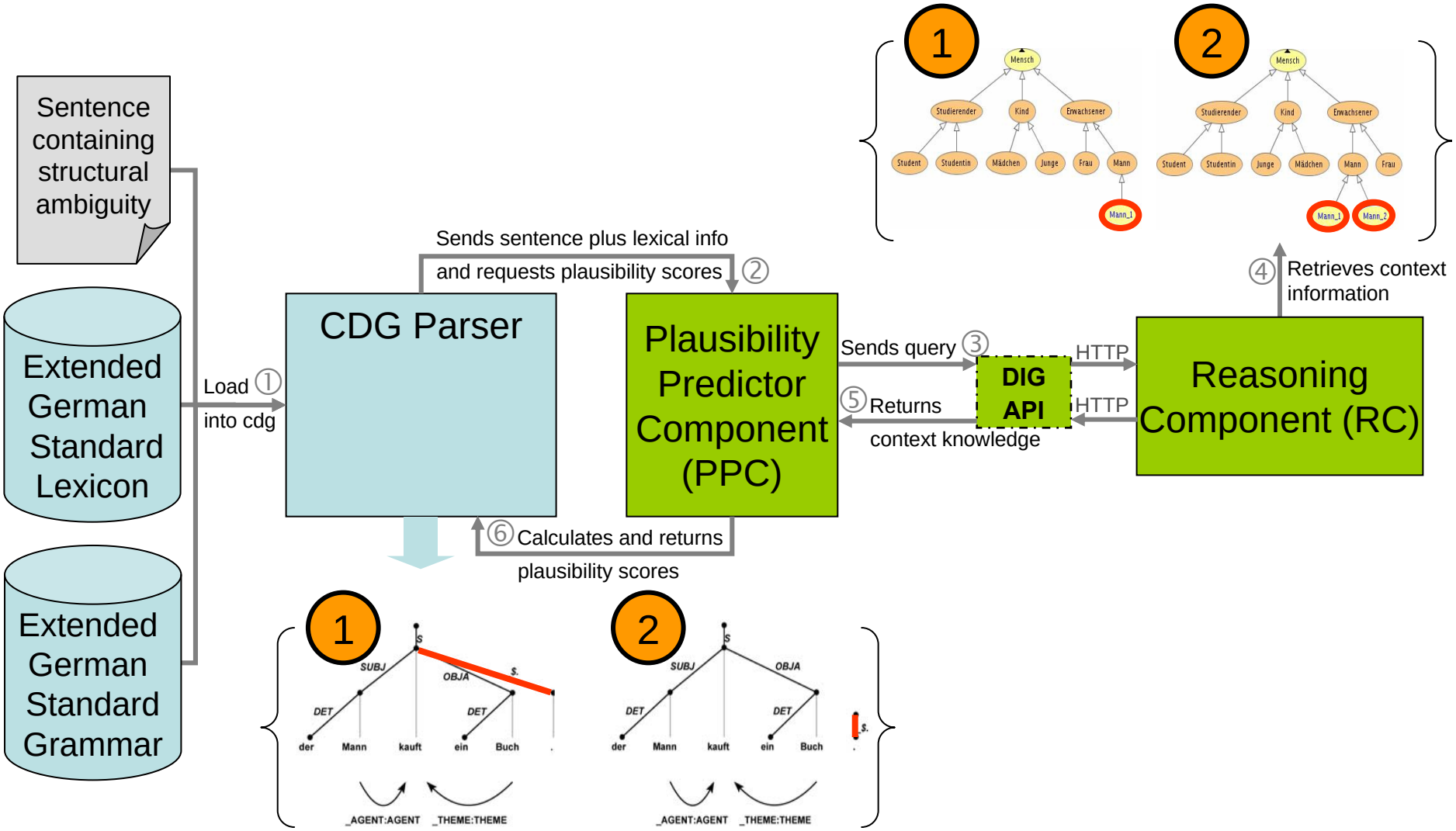
Plausibility prediction prior to checking for constraint satisfaction permits integration into the parsing process



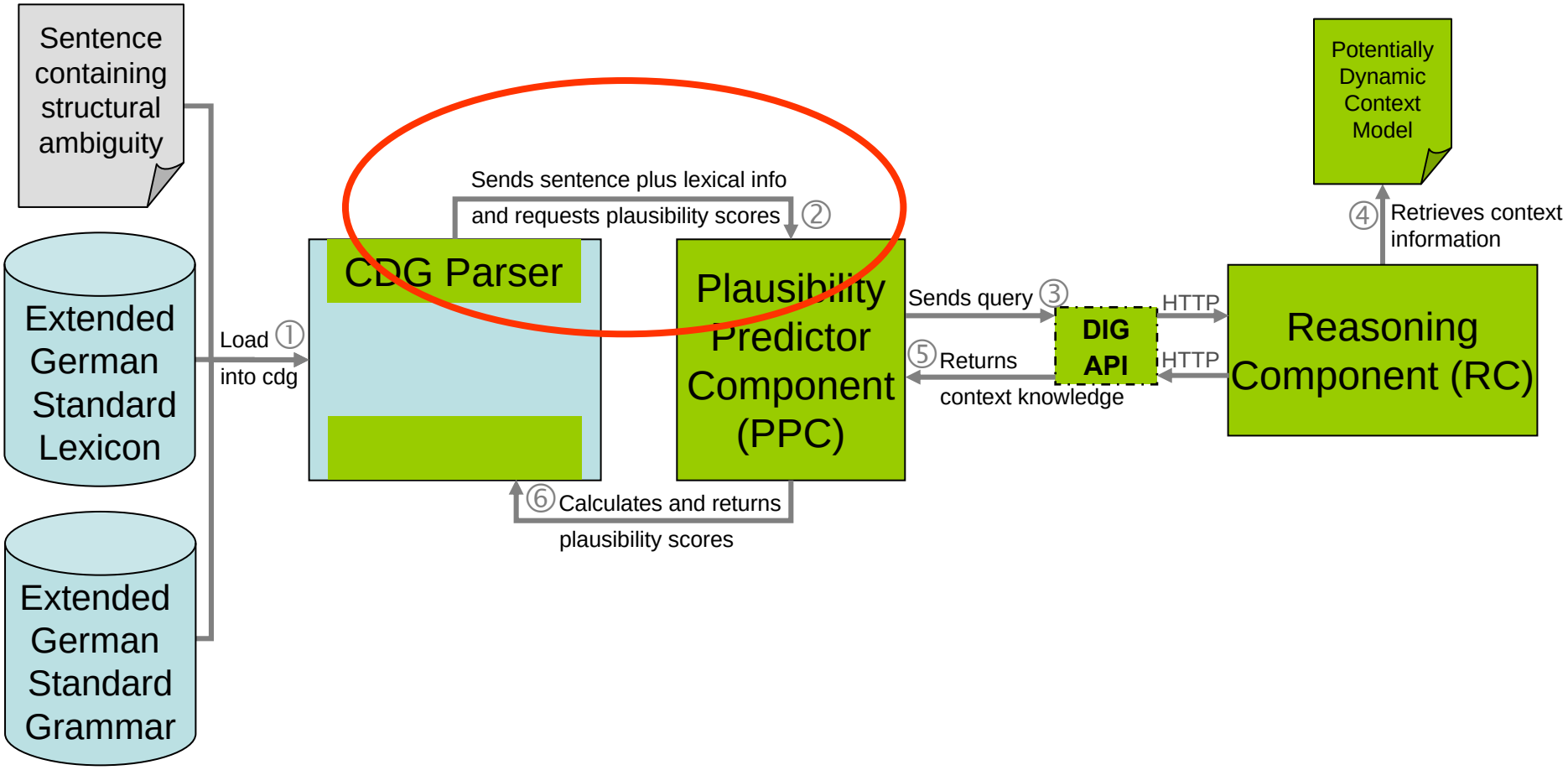
Presentation Structure

- Motivation
- Approach
- **Research Status**
- Future Directions

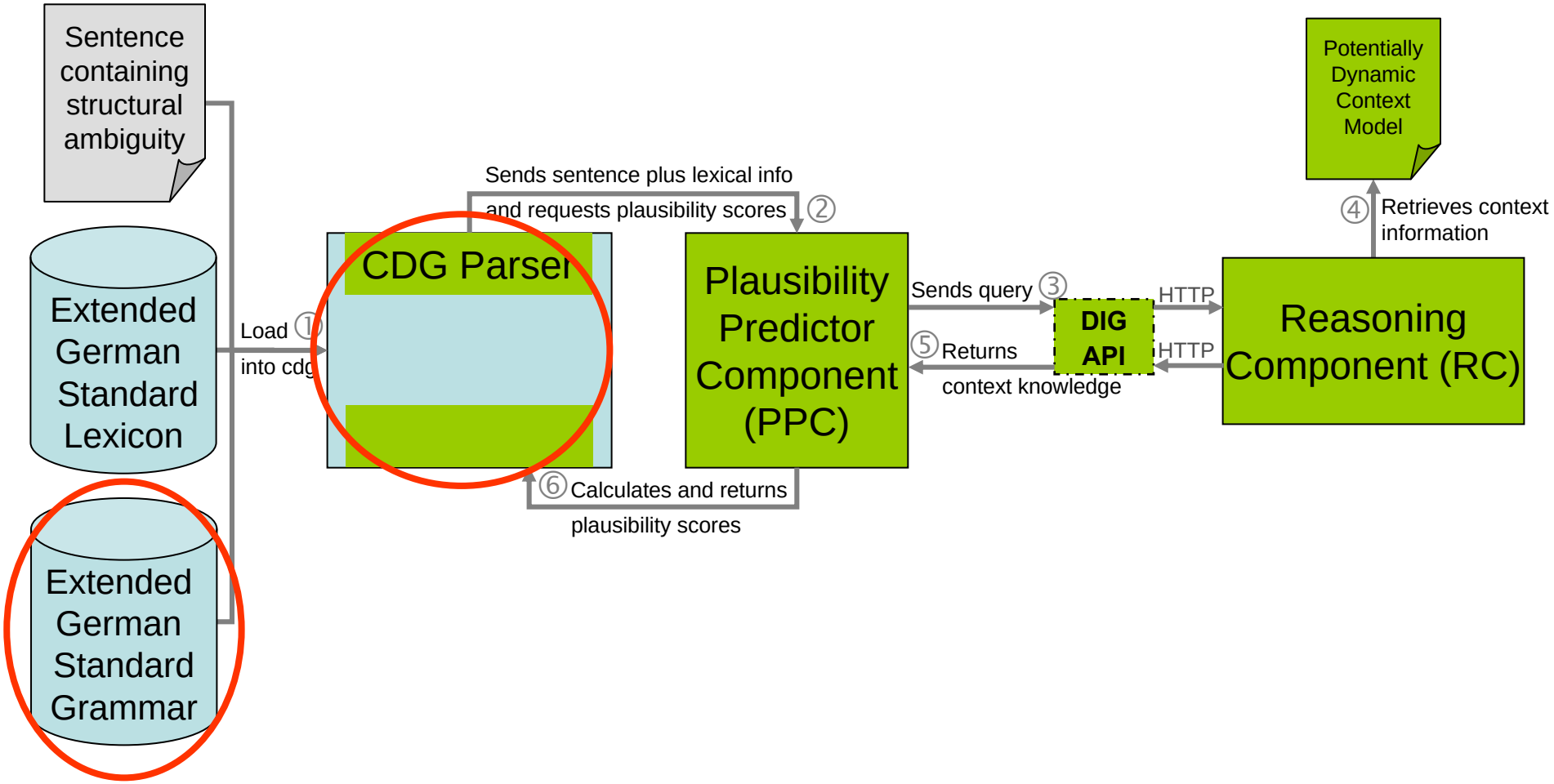
A technical proof of concept has demonstrated syntactic attachment decision making based on the Context Model



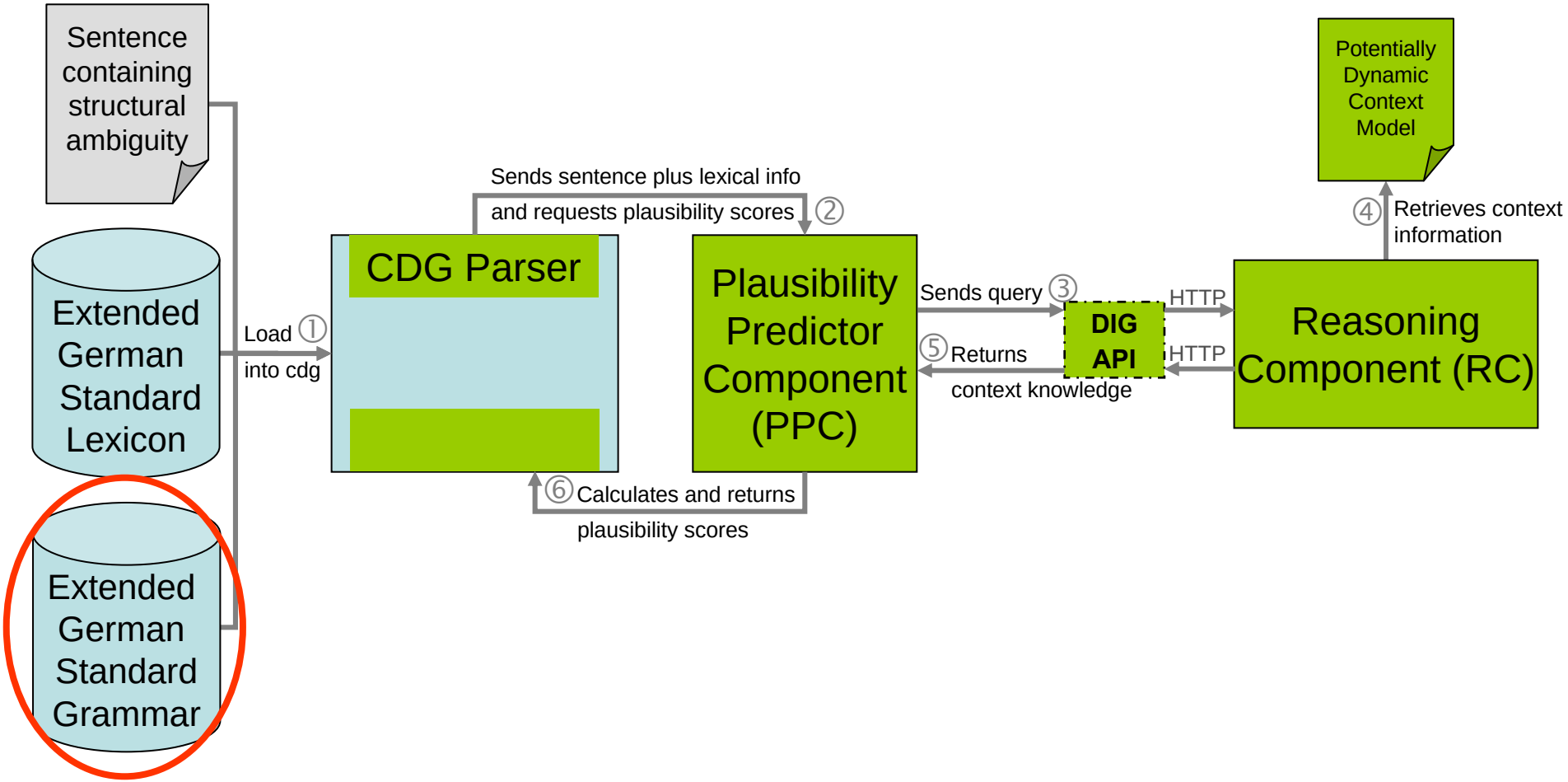
Challenges: CDG only submits the sentence's words but not *all* homonyms from the lexicon to the predictor



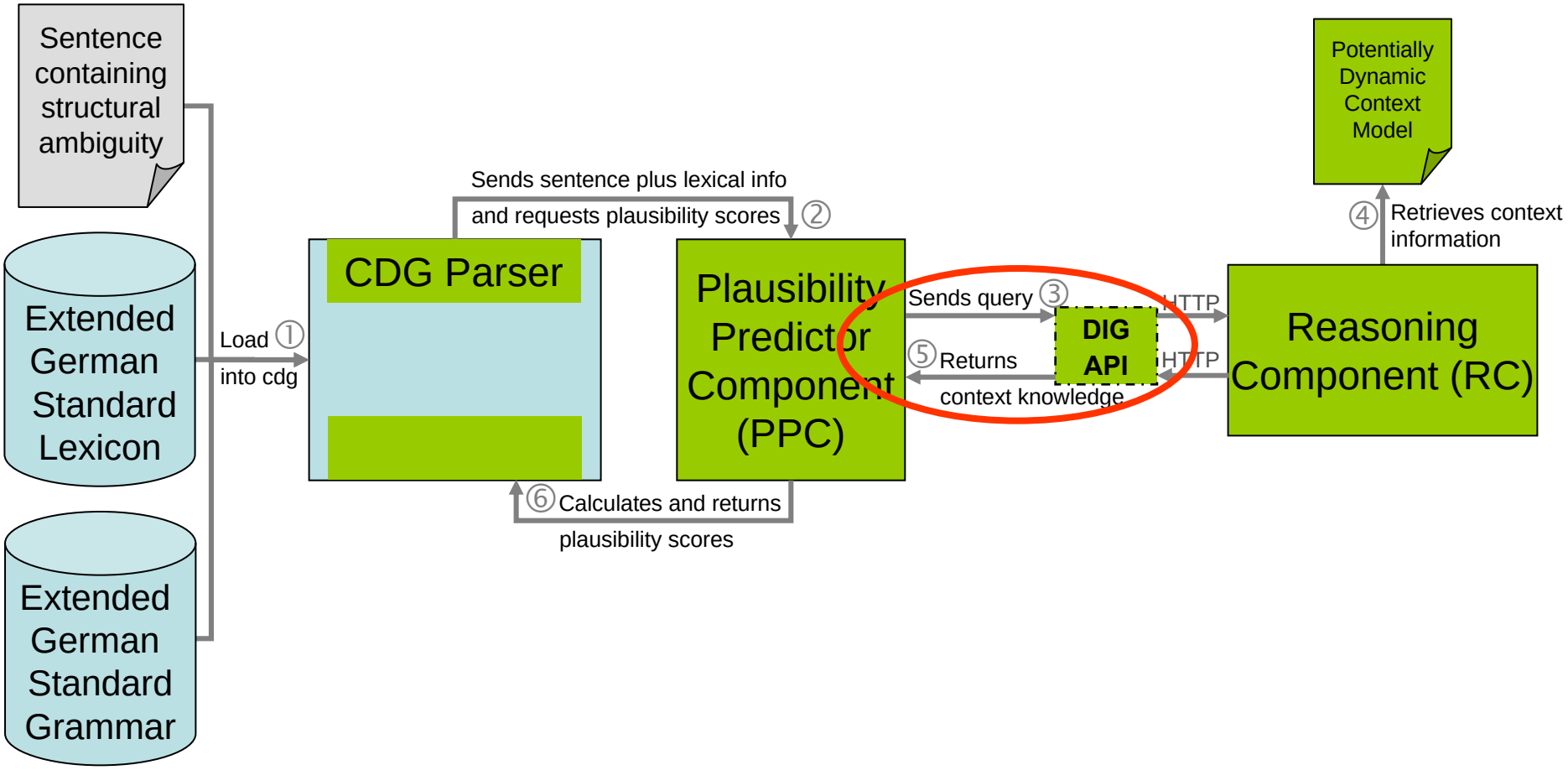
Challenges: CDG's grammar only permits access to unary node features rather than binary relations (edge features)



Challenges: CDG assigns syntactic labels only



Challenges: The DIG API has a fairly limited query syntax



The Context Integration Architecture should permit to approach notoriously hard German syntax ambiguities

- **PP Attachment**

  *Der Mann sah die Frau **mit dem Fernrohr**.*
  The man saw the woman with the telescope.

- **Coordination**

  *In diesem Haus wohnen schöne **Männer und Frauen**.*
  In this house live beautiful men and women.

- **Genitive-Dative Ambiguity in Feminine Singular Nouns**

  *Er sah, dass die Tochter **der Biologin** ein Buch kaufte.*

  He saw { *the daughter buying the biologist a book* / *the biologist's daughter buying a book* }.

- **Subject-Object Ambiguity in Plural Nouns**

  *Das sind die Kinder, **die die Lehrer** gesehen haben.*

  These are the children that { *saw the teachers.* / *the teachers saw.* }.

Central research questions to address with the target architecture in place

How does cross-modal context integration …

- … affect parsing accuracy inside and outside the modelled domain?
- … respond to different kinds of knowledge representation in the Context Model?
- … perform over a range of systematically varied constraint weightings?
- … affect parsing efficiency, i. e.: *system performance*?
- … perform over a range of different syntactic ambiguity phenomena?
- … affect parsing robustness against ill-formed input?

Publications so far can be downloaded from my personal webpage on the CINACS website

McCrae, Patrick (2007): "Integrating Cross-Modal Context for PP Attachment Disambiguation"

In: Proceedings of Third International Conference on Natural Computation, Vol. 3, pp. 292 - 296. Publisher: IEEE. ICNC'07. Haikou (Hainan), China; 24 - 27 August 2007.

McCrae, Patrick & Menzel, Wolfgang (2007): "Towards a System Architecture for Integrating Cross-Modal Context in Syntactic Disambiguation "

In: Proceedings of NLPCS 2007 – Fourth International Workshop on Natural Language Processing and Cognitive Science. Edited by: Bernadette Sharp and Michael Zock. Publisher: INSTICC Press.
NLPCS 2007. Funchal (Madeira), Portugal; 12 - 13 June 2007.

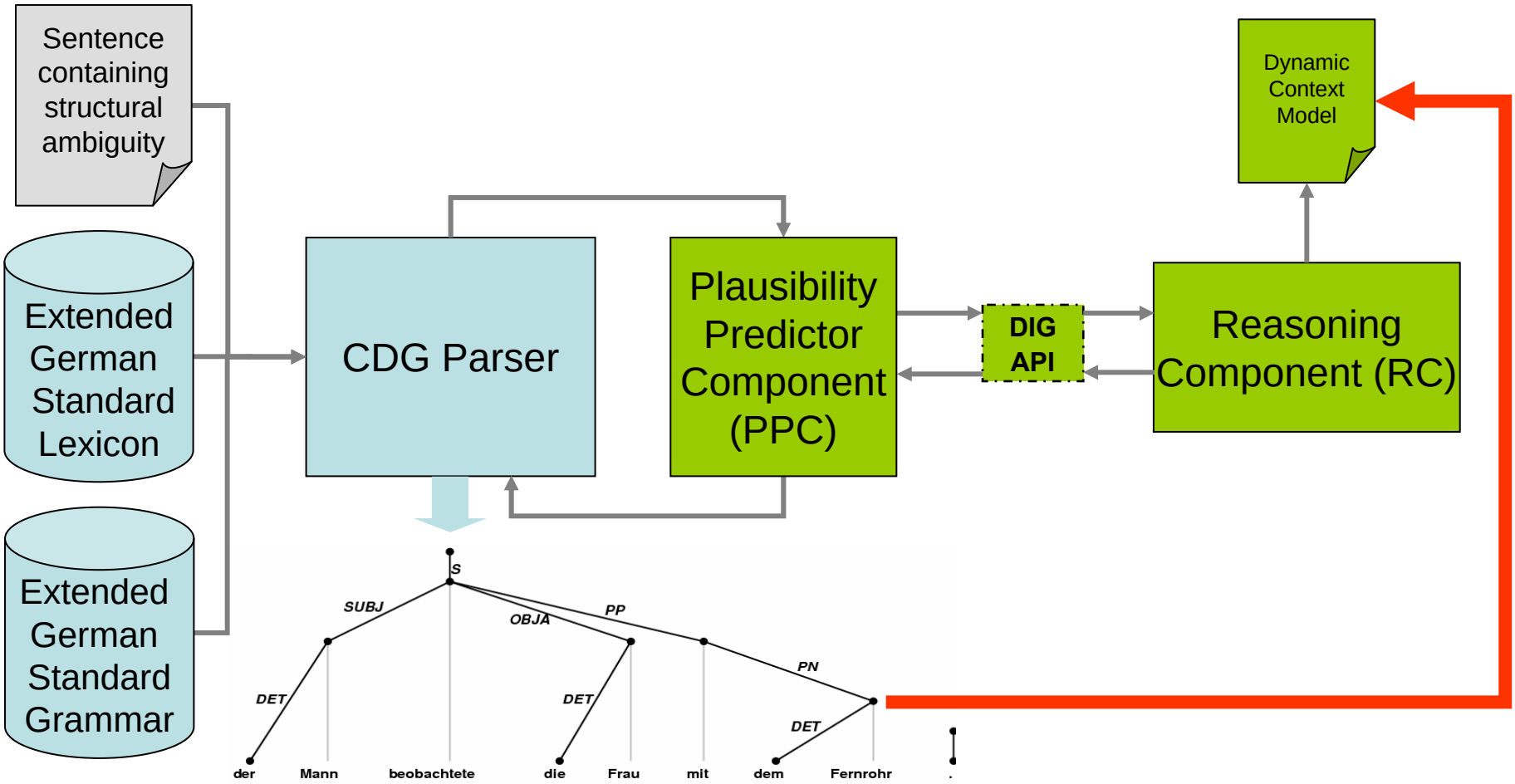
Conclusions

- A cross-modal model of human sentence processing must achieve integration of information from various modalities into parsing.
- A predictor calculates context-sensitive plausibility scores for thematic role assignment based on queries to a propositional context model. Weighted constraints integrate these plausibility scores to direct preferences in syntactic dependency parsing.
- We hypothesise that context integration will significantly and substantially improve the accuracy of structural ambiguity resolution.
- First experimental results are expected for the end of this year.

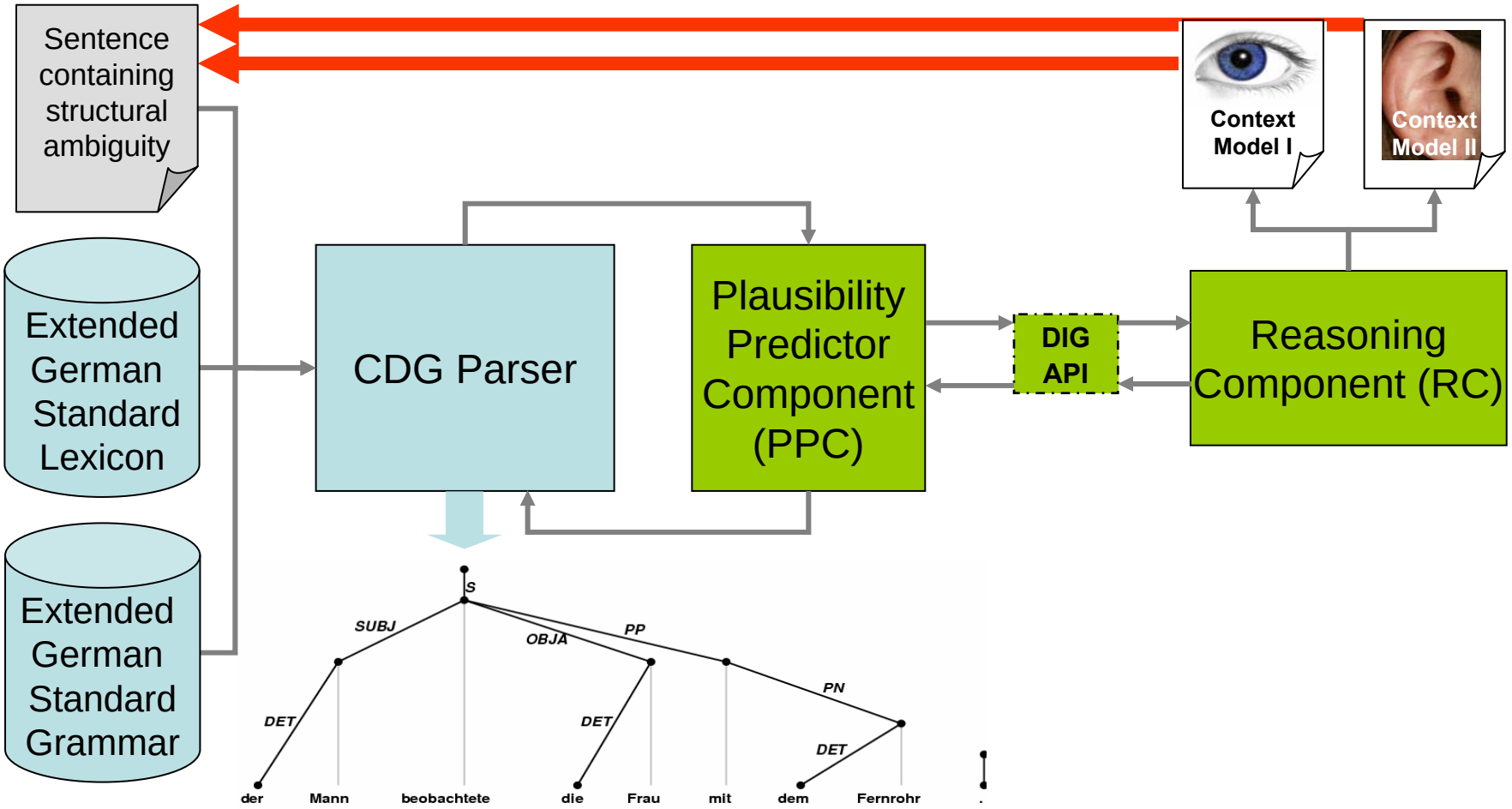
Presentation Structure

- Motivation
- Approach
- Research Status
- Future Directions

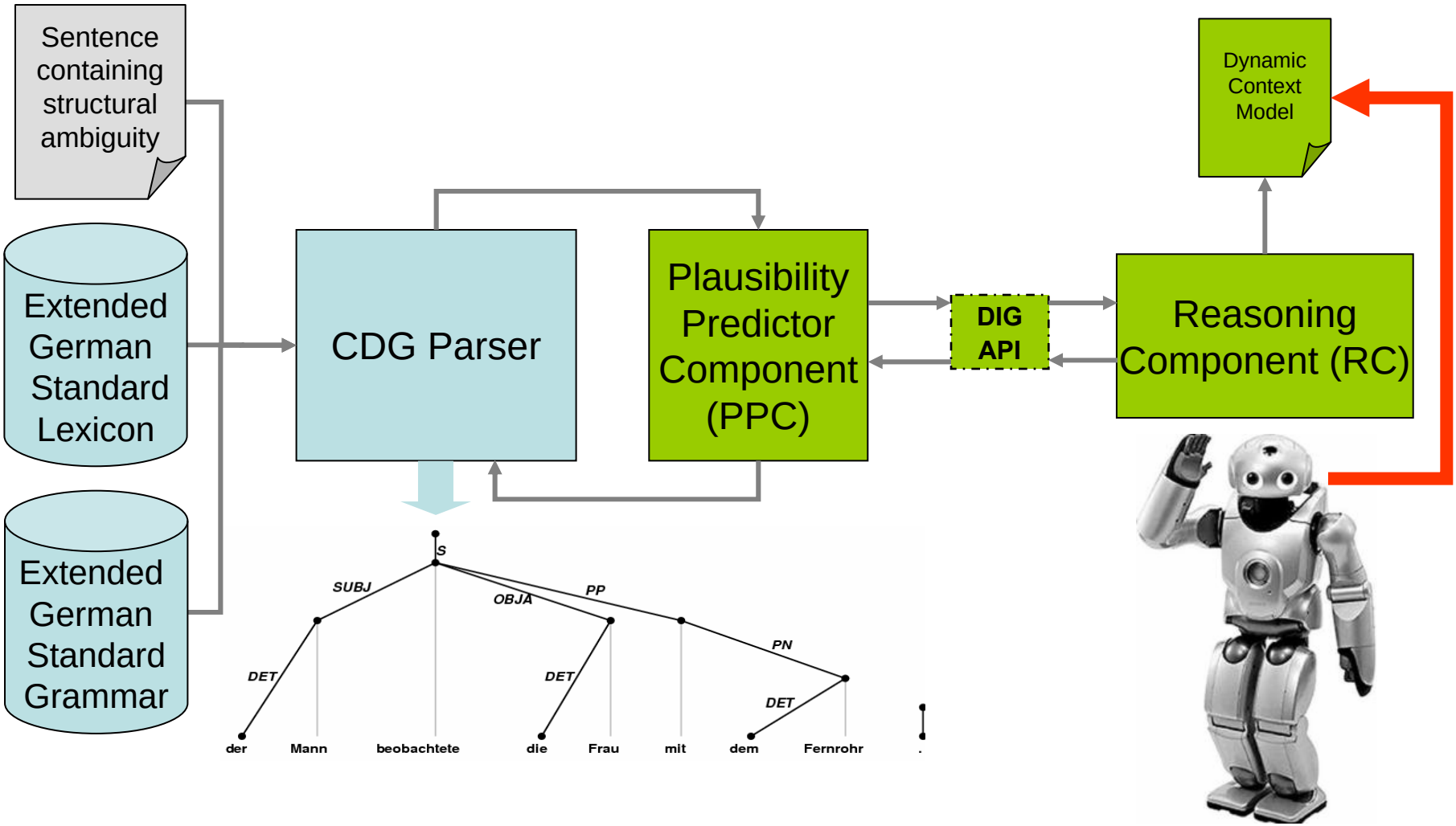
1. Discourse Analysis: Incrementally build the Context Model from parse results of preceding sentences



2. Cross-modal Communication: Integrate modality-specific context models to assess the impact of each modality



3. Sensor Fusion: Select the contextually most plausible parse based on data from a multi-sensor environment



Thank you for your time and attention

Patrick McCrae

CINACS Graduate Research Group
Department of Informatics, Hamburg University
Vogt-Kölln-Straße 30
22527 Hamburg
Germany

e-mail: patrick.mccrae@informatik.uni-hamburg.de
phone: +49 . 40 . 428 83 - 23 60

References

Crain, Stephen & Steedman, Mark (1985). On not Being Led up the Garden Path - the Use of Context by the Psychological Syntax Parser. In D. Dowty, L. Karttunen and A. Zwicky (Eds.), *Natural language parsing: Psychological, computational, and theoretical perspectives* (pp. 320 – 358). Cambridge University Press.

Tanenhaus, Michael, Spivey-Knowlton, Michael. J., Eberhard, Kathleen M. et al. (1995). Integration of visual and linguistic information in spoken language comprehension. SCIENCE (Volume 268), 16 June 1995, pp. 1632 – 1634.

Knoeferle, Pia S. (2005). The Role of Visual Scenes in Spoken Language Comprehension: Evidence from Eye-Tracking (PhD Thesis). Saarbrücken: Universität des Saarlandes.

Foth, Kilian & Menzel, Wolfgang (2006). The Benefit of Stochastic PP Attachment to a Rule-Based Parser. In *Proceedings of the 21st International Conference on Computational Linguistics*. Sydney: Coling-ACL-2006.

Bechhofer, Sean (2003). The DIG Description Logic Interface: DIG/1.1. In *Proceedings of the International Workshop on Description Logics (DL 2003).*