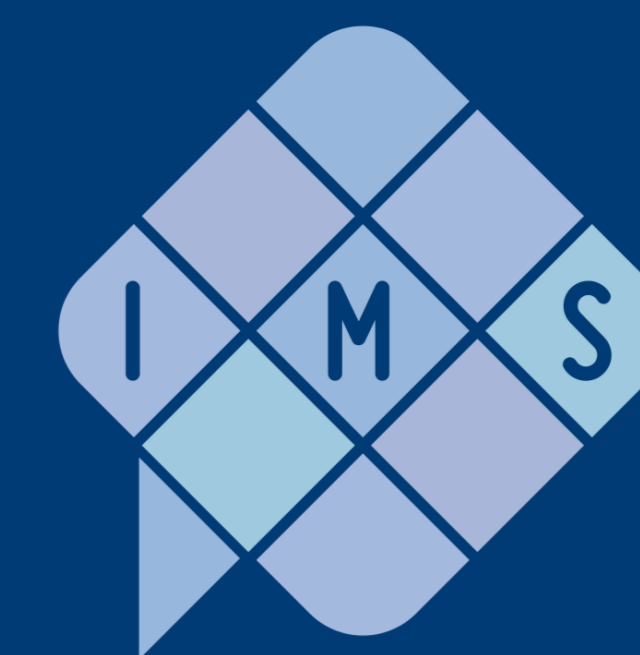




# Annotation and automatic classification of situation entity types



Annemarie Friedrich

Computerlinguistik, Universität des Saarlandes

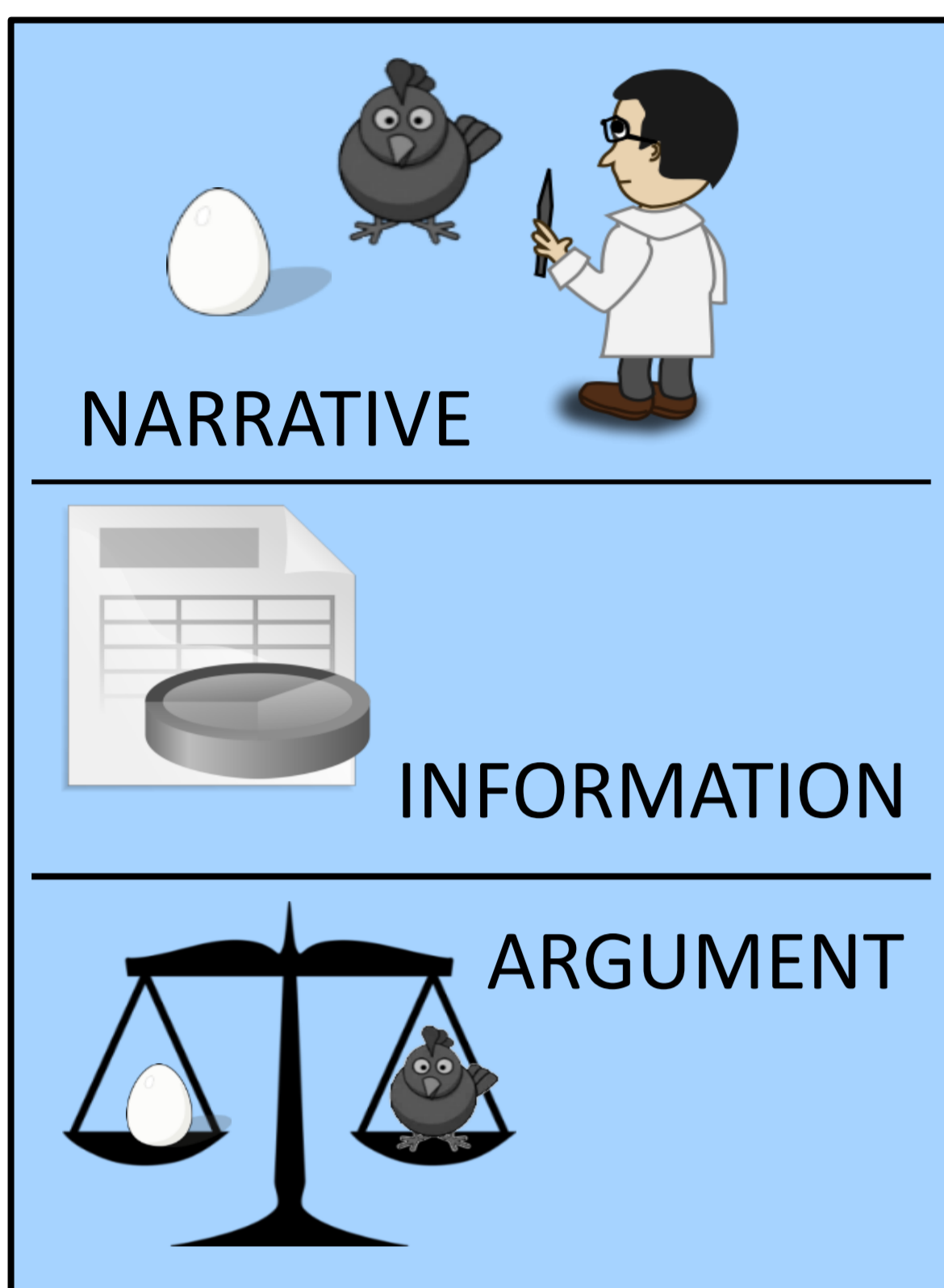
Alexis Palmer

IMS, Universität Stuttgart

<http://sitent.coli.uni-saarland.de>

DGFS Tagung 2015, Leipzig

## Discourse modes & situation entities [Smith 2003]



+ REPORT, DESCRIPTION

**STATE:** *Carl is a cat.*

**EVENT:** *Carl entered the room.*

**GENERALIZING SENTENCE:**

*Carl sometimes catches mice.*

**GENERIC SENTENCE:**

*Cats are popular pets.*

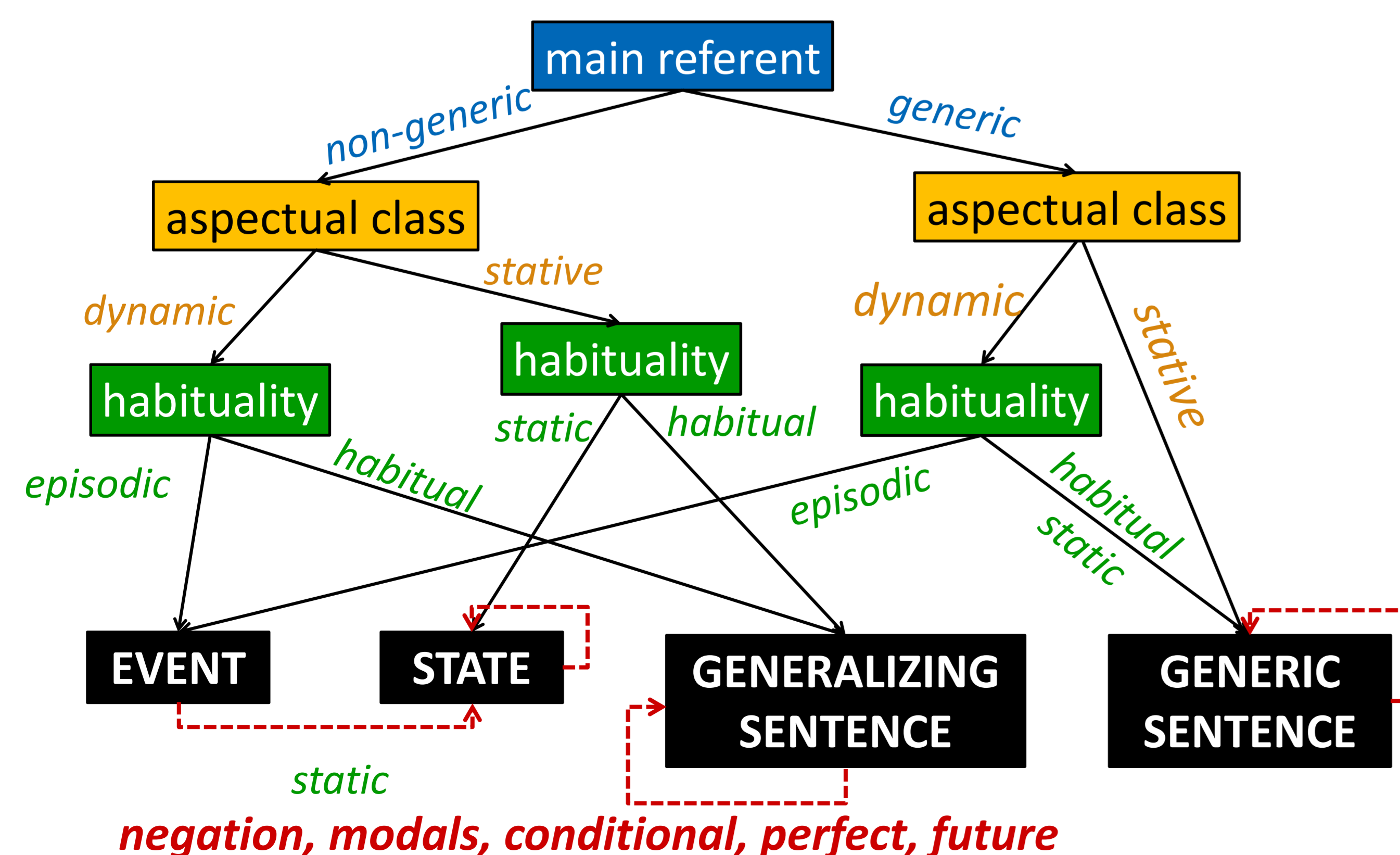
**ABSTRACT ENTITIES:**

*I know/believe  
that Mary likes cats.*

## Goals of the project

- assess the applicability of situation entity type classification: borderline cases? human agreement?
- training, development, evaluation of **automatic systems for classifying situation entities** and related tasks
- long-term: improving automatic (temporal) **discourse processing**, providing a foundation for analysis of the theory of **discourse modes** [Smith 2003]

## Situation types and their features



## Features: how to distinguish situation entity types

### Genericity of main referent

What is this clause about?

**particular** entity/group/company/organization/situation/process

*Mary likes cats. That she didn't answer upset me.*

**kind-referring/generic** NPs, **generic** concepts

*Cats eat mice. Security is an important issue.*

Krifka et al. (1995):

genericity;

Carlson (2005):

habitual sentences

### Habituality of clause

*Mary fed her cats this morning.* **episodic**: one-time event

*Mary drives to work by car.*

*Glass breaks easily.*

*Mary owns four cats.*

**habitual**: regularity

**static**

### Aspectual class of main verb

*Juice fills the glass.*

*She filled the glass with juice.*

*The glass was filled with juice.*

**stative**

**dynamic**

**both**

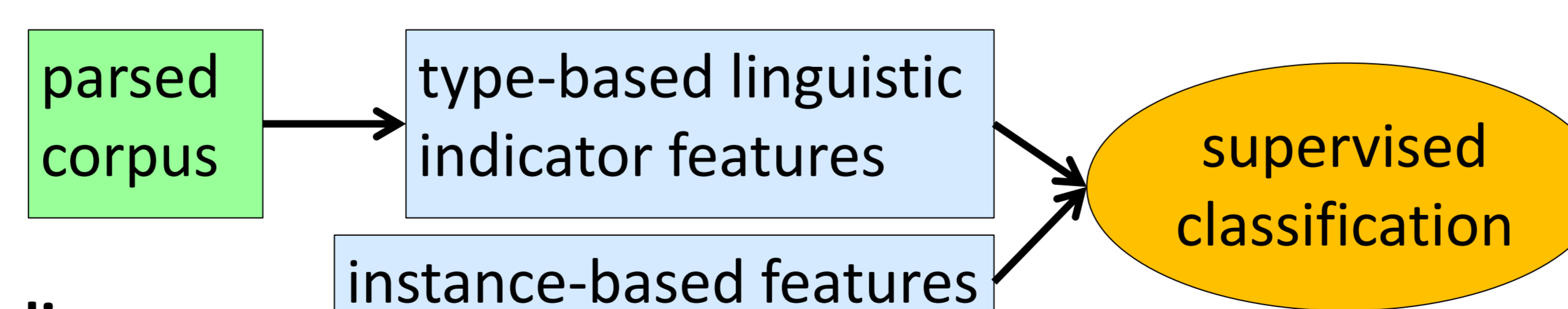


→ Vendler (1957): lexical aspect; Bach (1986): eventuality types

**Advantages of feature-driven annotation** [Friedrich & Palmer 2014a]

- ✓ easier to convey annotation scheme
- ✓ harness useful partial information
- ✓ analysis of disagreements

## Automatic prediction of aspectual class of verbs in context [Friedrich & Palmer 2014b]



### Findings:

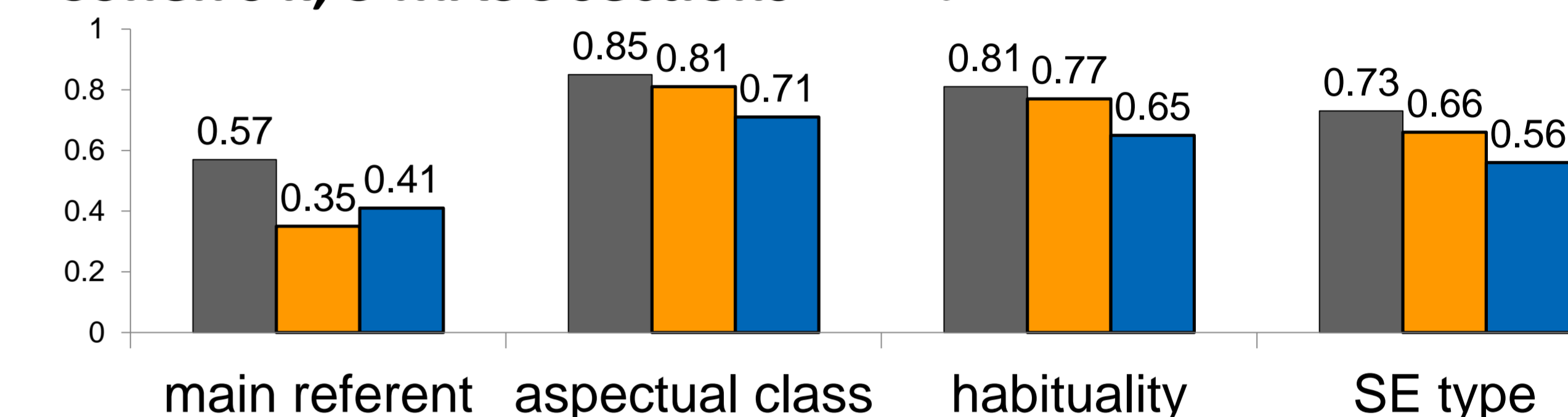
- accuracies between 80% and 90%
- verb-type based features generalize across verb types  
→ classifying instances of verbs unseen in training data
- especially for ambiguous verbs, instance-based features are essential

## Corpus annotation

**Data:** Manually Annotated SubCorpus (MASC) of the Open American National Corpus: various genres, other linguistic (syntactic/semantic) annotations available + **Wikipedia**.

**Status:** ≈ 40,000 double-/triple-annotated segments

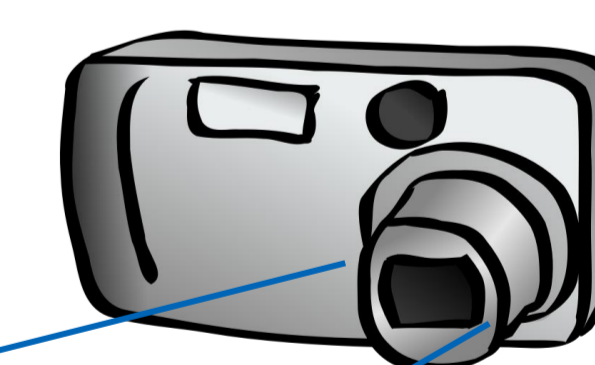
### Cohen's κ, 3 MASC sections



→ **substantial agreement** (except for main referent: reason is sparseness of generic main referents, agreement higher if balanced)

## Outlook

- **automatic prediction of**
  - genericity + habituality, situation entity types
  - aspectual class of light verbs:  
*have a heart attack* vs. *have a daughter*  
*make sense* vs. *make a cake*
- **situation entity types** = aspectual information  
= how speaker / writer presents a situation (≈ lens)
  - extend annotation scheme to other languages (planned: German, Chinese)
  - leverage information e.g. for
    - evaluation of translation quality
    - temporal processing



## References

- E. Bach. *The algebra of events*. 1986. Linguistics and philosophy 9 (1).
- G. N. Carlson. *Generics, habituals and iteratives*. 2005. Encyclopedia of Language and Linguistics, Elsevier.
- A. Friedrich & A. Palmer. *Situation entity annotation*. 2014a. LAW VIII, Dublin, Ireland.
- A. Friedrich & A. Palmer. *Automatic prediction of aspectual class of verbs in context*. 2014b. ACL, Baltimore, USA.
- M. Krifka et al. *Genericity: An Introduction*. 1995. The Generic Book.
- C. Smith. *Modes of Discourse: the local structure of texts*. 2003. Cambridge University Press.
- Z. Vendler. *Verbs and Times*. 1957. In: Linguistics in Philosophy, Cornell University Press.