

Annotation and automatic classification of situation entity types

Annemarie Friedrich

Alexis Palmer

Department of Computational Linguistics, Saarland University



One way to analyze text passages is in terms of the situation entities that they introduce to the discourse. In her work on modes of discourse, Smith (2003) distinguishes several situation entity types, which are expressed at the clause level. We create a corpus annotated with these types as well as relevant features, and aim to create automatic methods for labeling clauses with these features and types.

Situation entity types

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.*



Features

What is this clause about?

Genericity of main referent

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

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

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

Cats eat mice. Security is an important issue.

Aspectual class of main verb

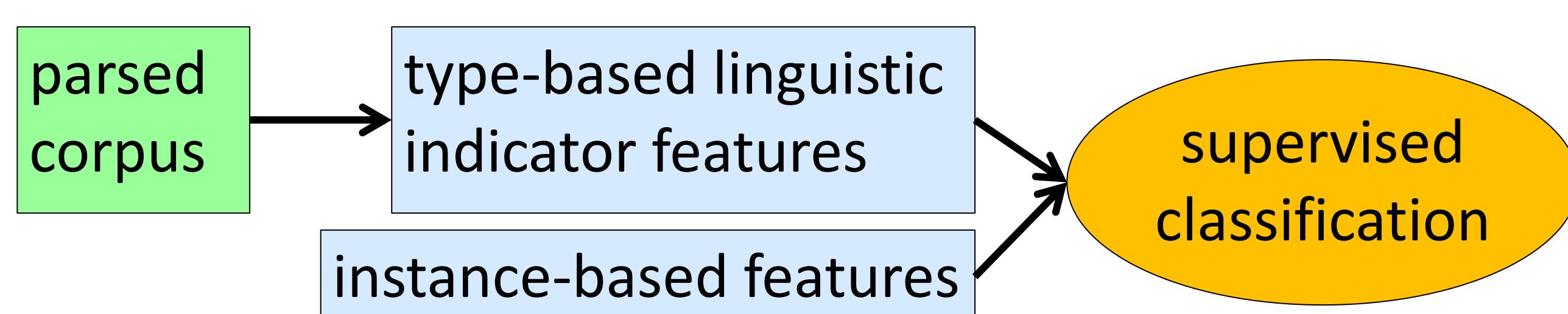
Juice fills the glass. (stative)

She filled the glass with juice. (dynamic)

The glass was filled with juice. (both)



Automatic prediction in context: (Friedrich & Palmer ACL 2014)



Findings:

- 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

Habituality of clause

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

Mary drives to work by car. (**habitual**: regularity)

Glass breaks easily. (**habitual**: regularity)

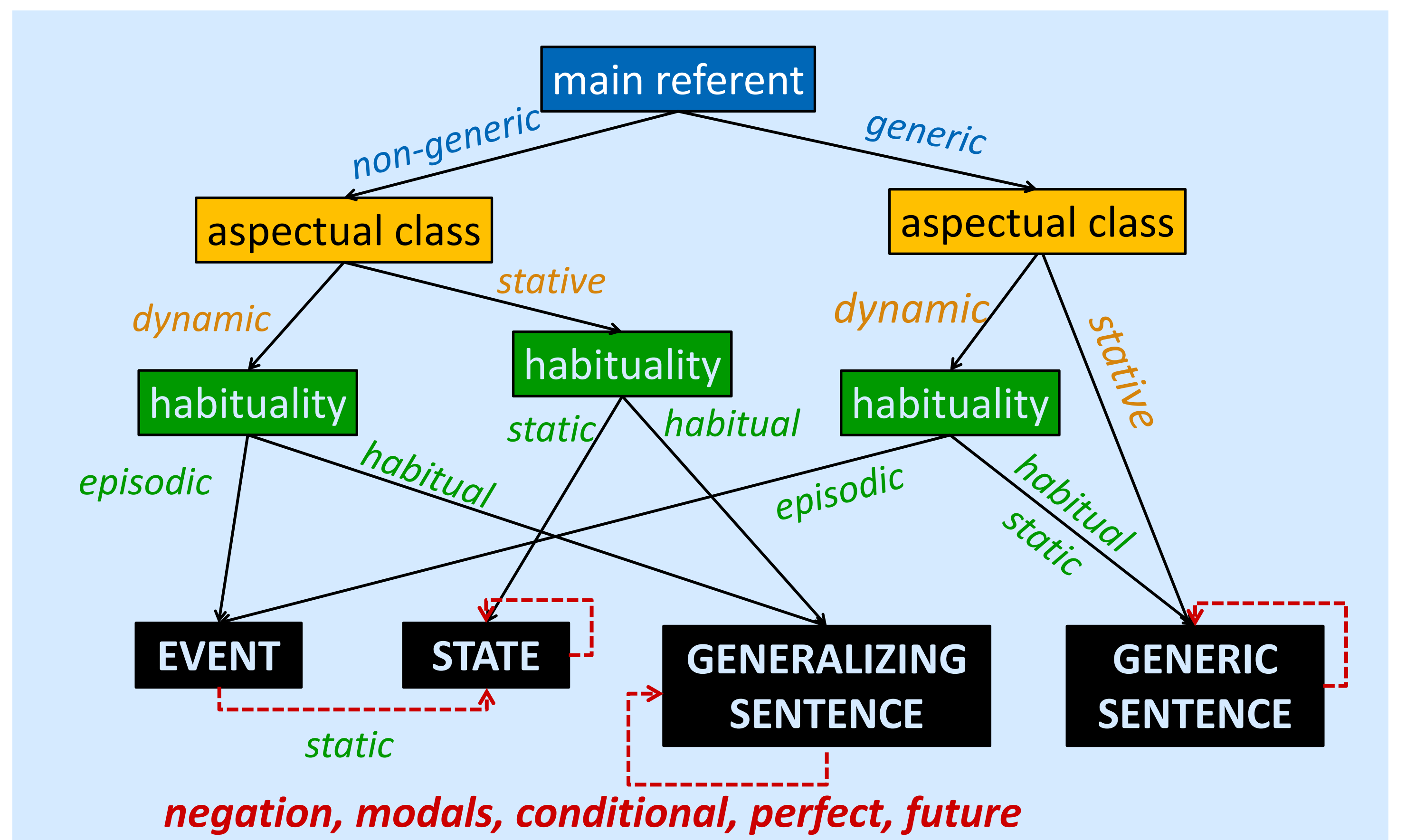
Mary owns four cats. (**static**)



Corpus annotation

Feature-driven annotation of situation entity types

- + easier to convey annotation scheme
- + get partial information
- + analyze disagreements



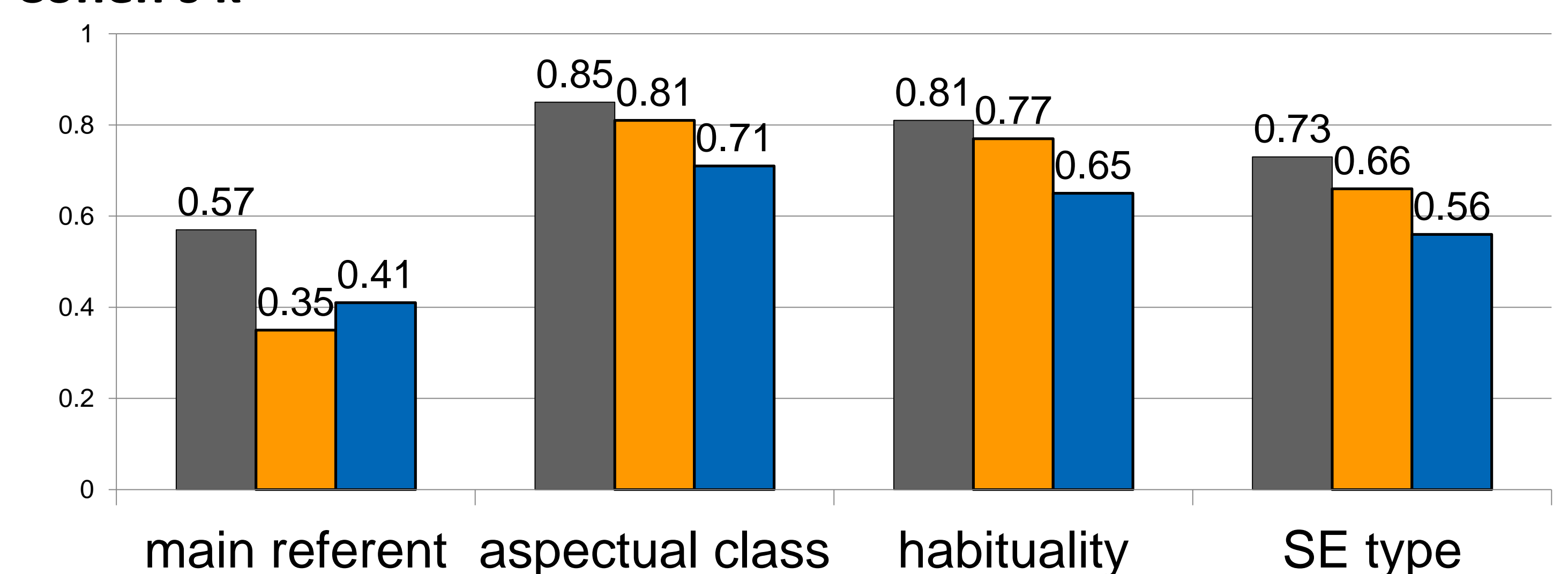
Data: MASC: various genres, other linguistic (syntactic and semantic) gold-standard annotations available.

Inter-annotator agreement

(Friedrich & Palmer LAW 2014)

Cohen's κ

■ jokes ■ news ■ letters



Goals of the project

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

References:

- 1) Carlota Smith. **Modes of Discourse: the local structure of texts**. 2003. Cambridge University Press.
- 2) Annemarie Friedrich & Alexis Palmer. **Situation entity annotation**. 2014. In *Proceedings of LAW VIII*, Dublin, Ireland.
- 3) Annemarie Friedrich & Alexis Palmer. **Automatic prediction of aspectual class of verbs in context**. 2014. In *Proceedings of ACL*, Baltimore, USA.