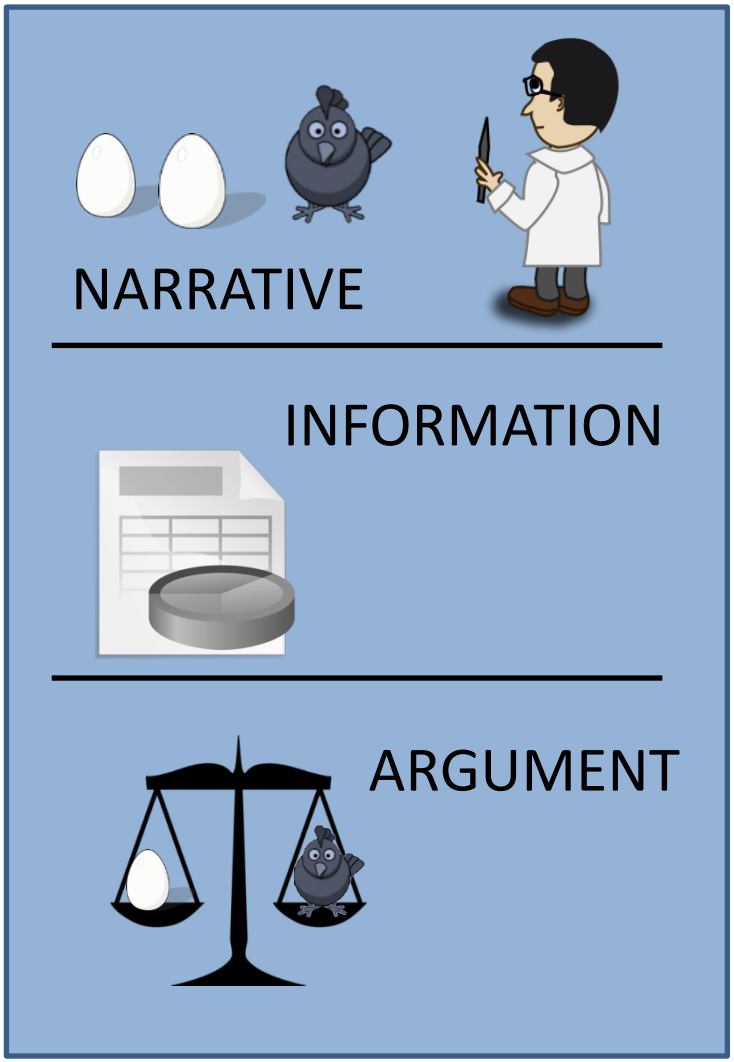
Linguistic Annotation Workshop (LAW VIII)
Dublin, August 2014

Situation entity annotation

Annemarie Friedrich & **Alexis Palmer**
Department of Computational Linguistics
Saarland University

Modes of discourse [Smith 2003]



Different passages of a text can have different discourse modes.

one text ≈ one genre

one text ≠ one discourse mode

Modes of discourse [Smith 2003]:

Situation entity types

NARRATIVE

EVENT, STATE

REPORT

EVENT, STATE, general statives

DESCRIPTION

EVENT, STATE, ongoing EVENT

general statives

INFORMATION

ARGUMENT

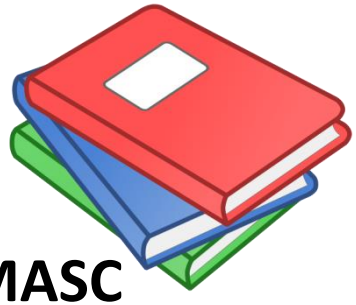
FACT, PROPOSITION, general statives

Related work

- Palmer et al. [2007]:
  - first labeled data set for SEs
  - ~6000 clauses
  - no annotation manual
  - Cohen's $\kappa = 0.54$
- Stede & Peldzsus [2012]:
  - illocutionary status of clauses in causal relations

    ~pragmatic role, e.g. REPORT, DIRECTIVE, COMMITMENT

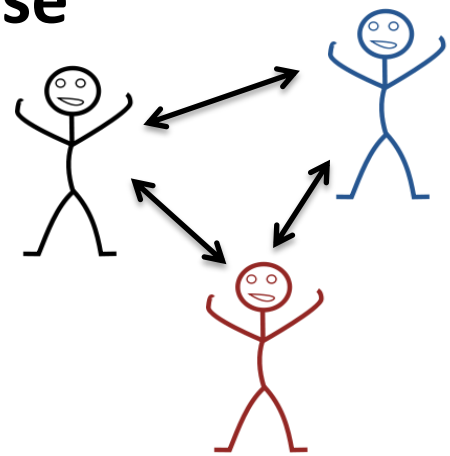
Overview of this work:



MASC

(automatic) segmentation

training phase
+ manual



inter-annotator agreement

Situation entity (SE) types

- genericity of main referent
- fundamental aspectual class
- habituality

Feature-based annotation

intra-annotator consistency

Motivation of annotation study

foundation for analysis of the theory of Discourse Modes [Smith 2003]

training, development, evaluation of automatic systems for classifying SEs and related tasks

assess the applicability of SE type classification as described by Smith [2003]

borderline cases? human agreement?

Situation entity types (SE types)

Yesterday, Mary bought a cat. **EVENT**

Now she owns four cats. **STATE**

} eventualities

Susie often feeds Mary's cats. **GENERALIZING SENTENCE**

Cats are very social animals. **GENERIC SENTENCE**

} general statives

SE types: abstract entities

here: clausal complements
of factive / implicative verbs

Susie **knows** **STATE**

that Mary loves her cats a lot. **FACT** objects of knowledge

Susie **believes** **STATE**

that the cats also love Mary. **PROPOSITION**

objects of belief

SE types: speech act types [Palmer et al. 2007]

Did you see my cats? **QUESTION**

Don’t forget to feed the cats! **IMPERATIVE**

Derived situation entity types

coerce **EVENTs** to **STATEs**:

negation, modality, future / perfect tense,
conditionality, subjectivity

Susie will feed the cats.
Susie has not fed the cats.
If Susie has forgotten the cats,
they might be hungry now.

Derived SE types

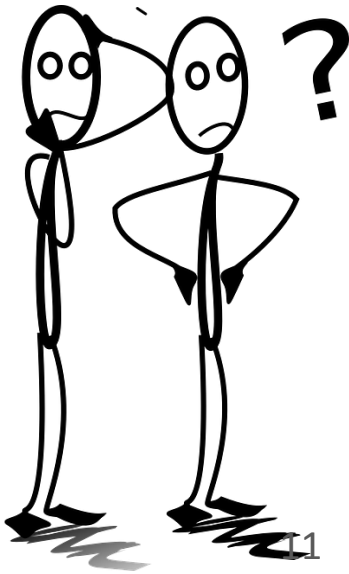
general statives are not subject to such coercion:

Susie never feeds Mary's cats. **GENERALIZING SENTENCE**

Cats might be the most popular pet. **GENERIC SENTENCE**

SE types: summary

Eventualities	STATE	*Mary likes cats.*
	EVENT	*Mary fed the cats.*
	- REPORT	*..., Mary said.*
General Statives	GENERALIZING SENTENCE	*Mary often feeds my cats.*
	GENERIC SENTENCE	*Cats are always hungry.*
Abstract Entities	FACT	*I know that Mary fed the cats.*
	PROPOSITION	*I believe that Mary fed the cats.*
Speech Acts	QUESTION	*Does Mary like cats?*
	IMPERATIVE	*Don't forget to feed the cats!*



Data: Manually Annotated SubCorpus (MASC) of Open American National Corpus

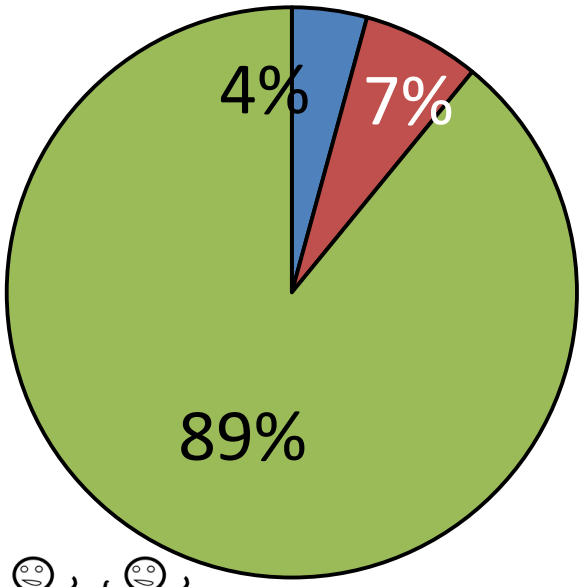
- ✓ additional types of annotation available
- ✓ open distribution of annotations
- ✓ wide range of genres

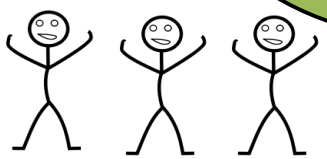
MASC section	# of situations (segments)	average # tokens per segment
news	3455	9.9
jokes	2563	6.9
letters	1851	11.1

Segmentation

SPADE [Soricut & Marcu 2003]
+ heuristic post-processing
+ manual correction

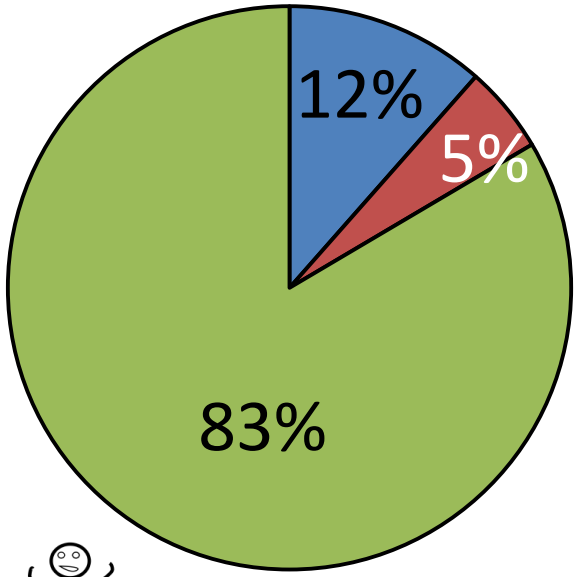
marked as **NO SITUATION**
by at least one annotator
(e.g. headlines, names, dates)

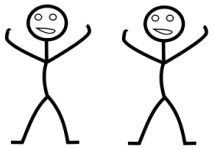




MASC news: 2823 segments
2515 situations for analysis

merged to other segment
by at least one annotator





MASC news, jokes, letters: 9428 segments
7869 situations for analysis

Feature-driven annotation

1. label “easy” cases: speech acts, lexically-triggered abstract entities, other clear-cut cases
2. determine **feature values**

genericity of main referent	fundamental aspectual class	habituality

3. use feature values to assign

Situation entity (SE) types

\+ Options for indicating uncertainty, multiple SE types / feature values.

Which features distinguish the SE types from each other?

Feature-driven annotation

1. label “easy” cases: speech acts, lexically-triggered abstract entities, other clear-cut cases
2. determine **feature values**

genericity of main referent	fundamental aspectual class	habituality

3. use feature values to assign

Situation entity (SE) types

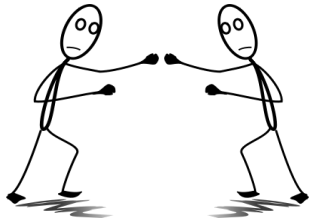
\+ Options for indicating uncertainty, multiple SE types / feature values.

Advantages

✓ easier to convey annotation scheme

✓ get partial information

✓ analyze disagreements



Situation Entity Types : Annotation

File: training_test_mixed.txt

		simply "the Saar",
9	ST	as is frequently referred to) did not exist as a unified entity.
10	ST	Until then, some parts of it had been Prussian
11	ST	while others belonged to Bavaria.
12	EV	The inhabitants voted to rejoin Germany in a plebiscite
13	EV	held in 1935.
14	ST	From 1947 to 1956 the Saarland was a French-occupied territory(the "Saar Protectorate") separate from the rest of Germany.
15	ST	Between 1950 and 1956, Saarland was a member of the Council of Europe.
16		In 1955, in another plebiscite, the inhabitants were offered independence,
17		but voted instead for the territory to become a state of West Germany.
18	seg_prob	
19	seg_prob	MARS
20	ST	Mars is the fourth planet from the Sun and the second smallest planet in the Solar System.
21	ST	Named after the Roman god of war,
22	GEN_STAT, GENERAL	it is often described as the "Red Planet"
23	ST, GEN_STAT, GENERAL	because the iron oxide prevalent on its surface gives it a reddish appearance.

Features

Main Referent

- [x] specific
- [ ] generic
- [ ] expletive
- [ ] can't decide
- [ ] not the grammatical subject

Aspectual Class of main verb

- (•) stative
- ( ) dynamic (eventive)
- ( ) both readings possible
- ( ) can't decide

Habituality of main verb

- [ ] episodic
- [ ] habitual
- [x] static
- [ ] can't decide

Segmentation Problems

- [ ] no situation
- [ ] includes text that does not belong to this situation
- [ ] multiple situations

Situation Entity Types

- [x] STATE
- [ ] EVENT
  - [ ] REPORT
- [ ] GENERAL STATIVE
  - [ ] GENERALIZING SENT.
  - [ ] GENERIC SENTENCE
- [ ] ABSTRACT ENTITY
  - [ ] FACT
  - [ ] PROPOSITION
  - [ ] RESEMBLANCE
- [ ] SPEECH ACT
  - [ ] IMPERATIVE
  - [ ] QUESTION

- [ ] Not done
- [ ] I'm not sure here

Comments:

Feature: genericity of main referent

What is this clause about? → usually the grammatical subject

SPECIFIC (≈ non-generic)

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

Mary likes cats.
The cats broke the TV.
WWF protects animals.
That she didn't answer upset me.
Knitting this scarf took me two days.

GENERIC

kind-referring / class-referring NPs
generic concepts

Cats eat mice.
Lions in captivity have trouble to produce offspring.
Dinosaurs are extinct.
Security is an important issue.
Knitting a scarf is generally fun.

distinguishes GENERIC SENTENCEs from other SE types
(in combination with other features)

[Friedrich & Palmer, ACL 2014]

Feature: fundamental aspectual class

distinguishes
EVENTs from STATEs

feature of the entire clause, marks main verb.

Juice **fills** the glass.
STATIVE

The glass **was filled** with juice.
BOTH readings possible

She **filled** the glass with juice. **DYNAMIC**

Feature: habituality

feature of the entire clause, marks main verb.

distinguishes EVENTs from general statives.

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

Mary feeds her cats every morning. (**habitual**: regularity)

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

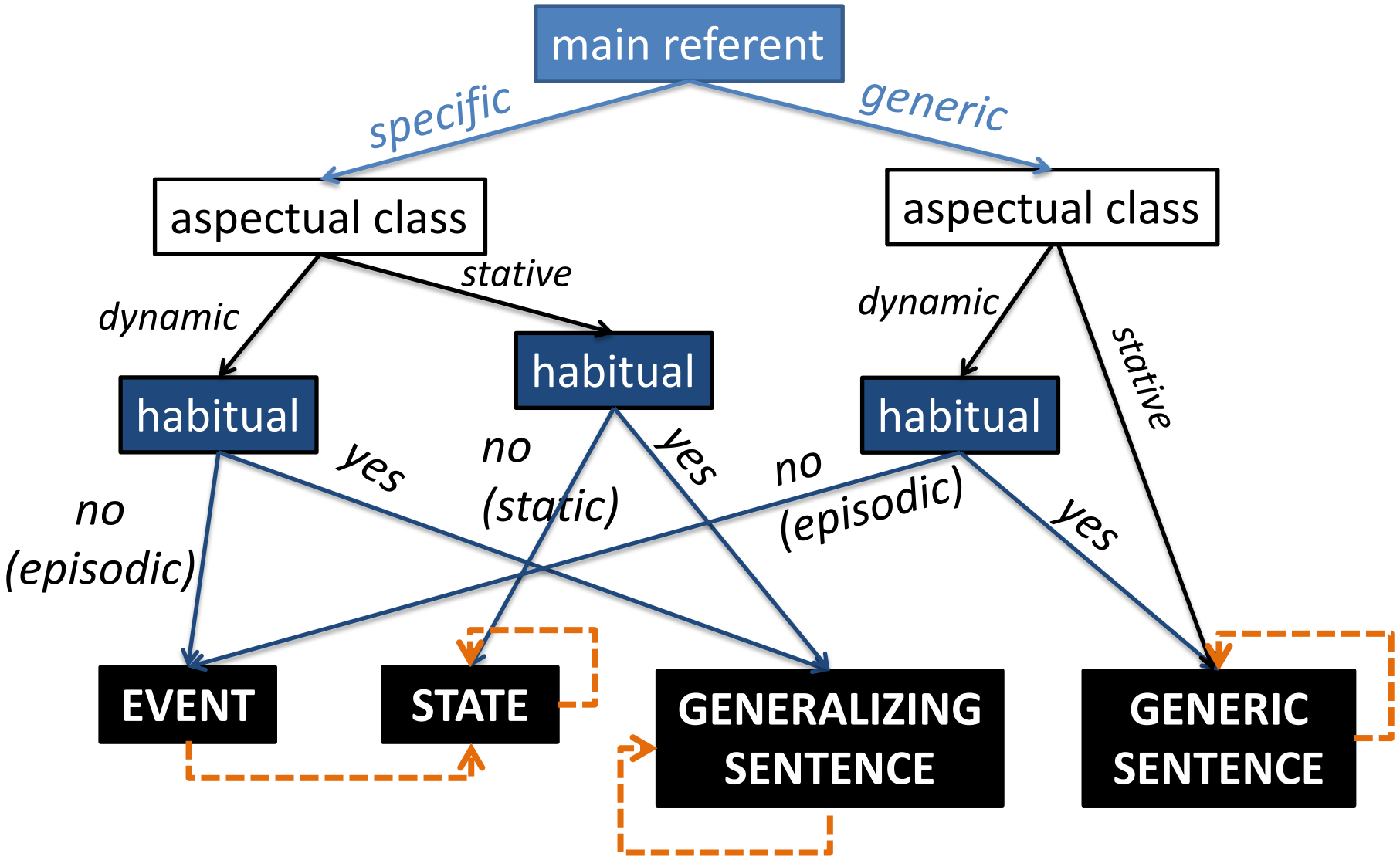
Mary owns four cats. (**static**: for STATEs)

Features – broader perspective

corpus data for sub-tasks studied in the NLP community for which no large data sets are available

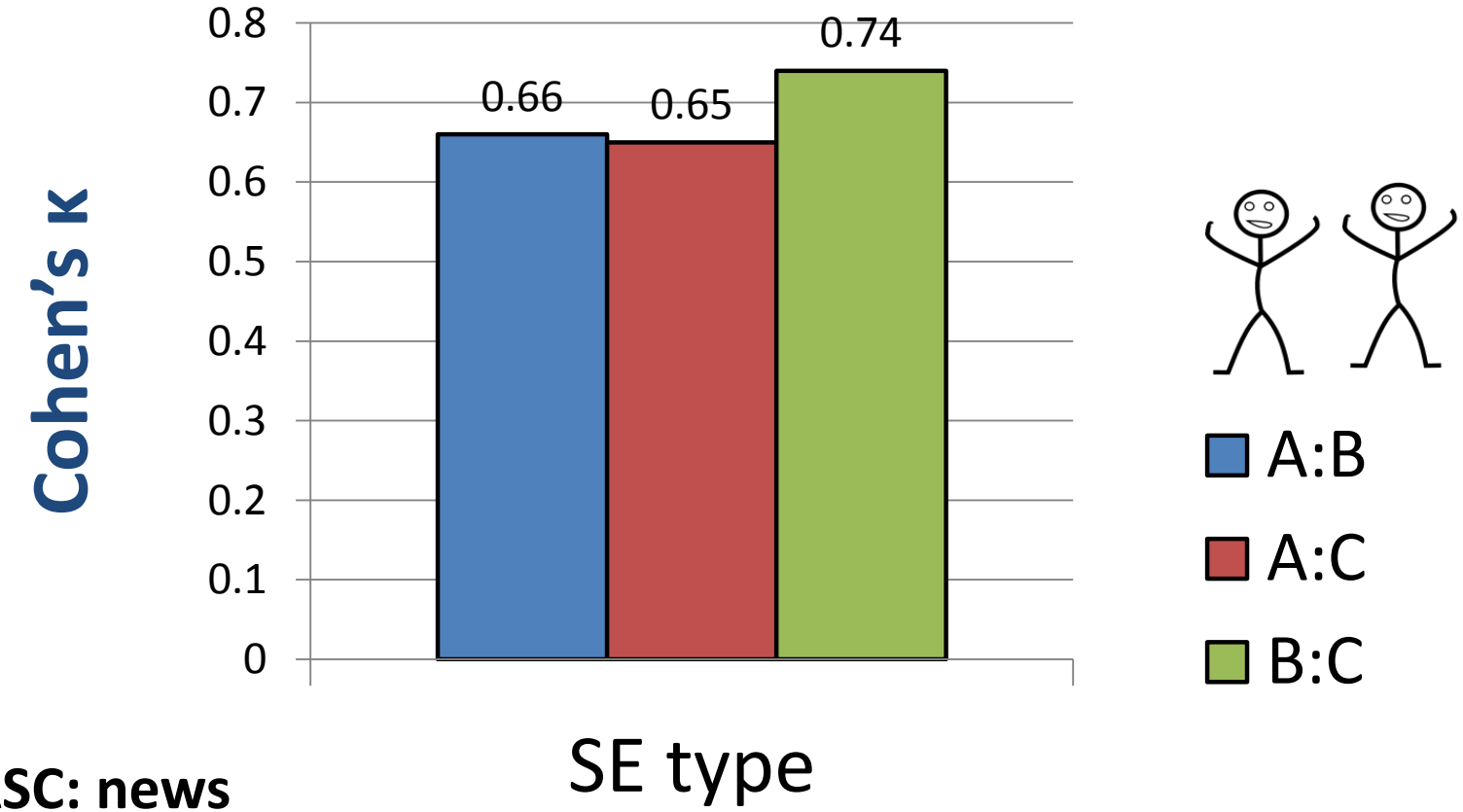
- automatic classification of fundamental aspectual class [Siegel & McKeown 2000, Friedrich & Palmer 2014] with the aim of improving temporal discourse processing [UzZaman et al. 2013, Bethard 2013, Costa & Branco 2012]
- identifying generic noun phrases [Reiter & Frank 2013]
- identifying habitual vs. episodic sentences [Mathew & Katz 2009]

Features & SE types



SE types: inter-annotator agreement

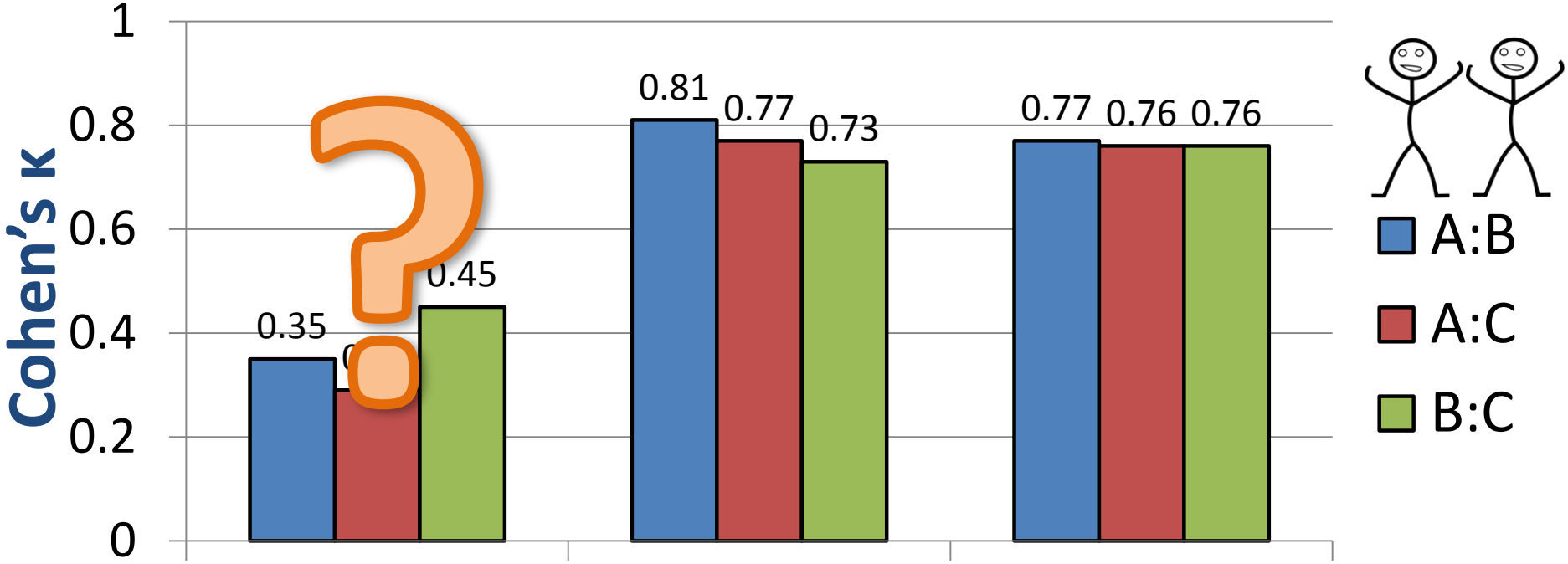
STATE, EVENT, GENERIC SENTENCE, GENERALIZING SENTENCE



MASC: news

Features: inter-annotator agreement

MASC: news



main referent	aspectual class	habituality
specific	*stative*	*episodic*
generic	*dynamic*	*habitual*
expletive	*both*	*static*

Feature: genericity of main referent (inter-annotator agreement)

183 clauses: B &C agree, A disagrees			
92%: B & C →specific, A → generic			
40%: misunder-standing by A	**30%**: multiple readings	**30%**: other	

As a governor, I'll make sure <u>that every kid in New York</u> has the same opportunity.

<u>you</u> in letters → generic or addressee?

Comparing B and C: ($\kappa = 0.45$)

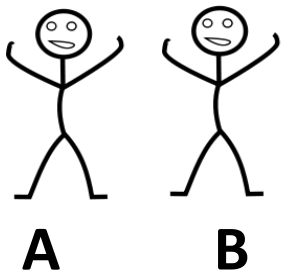
- 2358 segments : specific by both
- 122 segments: generic by at least one
- 43 segments: generic by both

➡ very few cases, cannot draw conclusions on reasons for low κ yet.

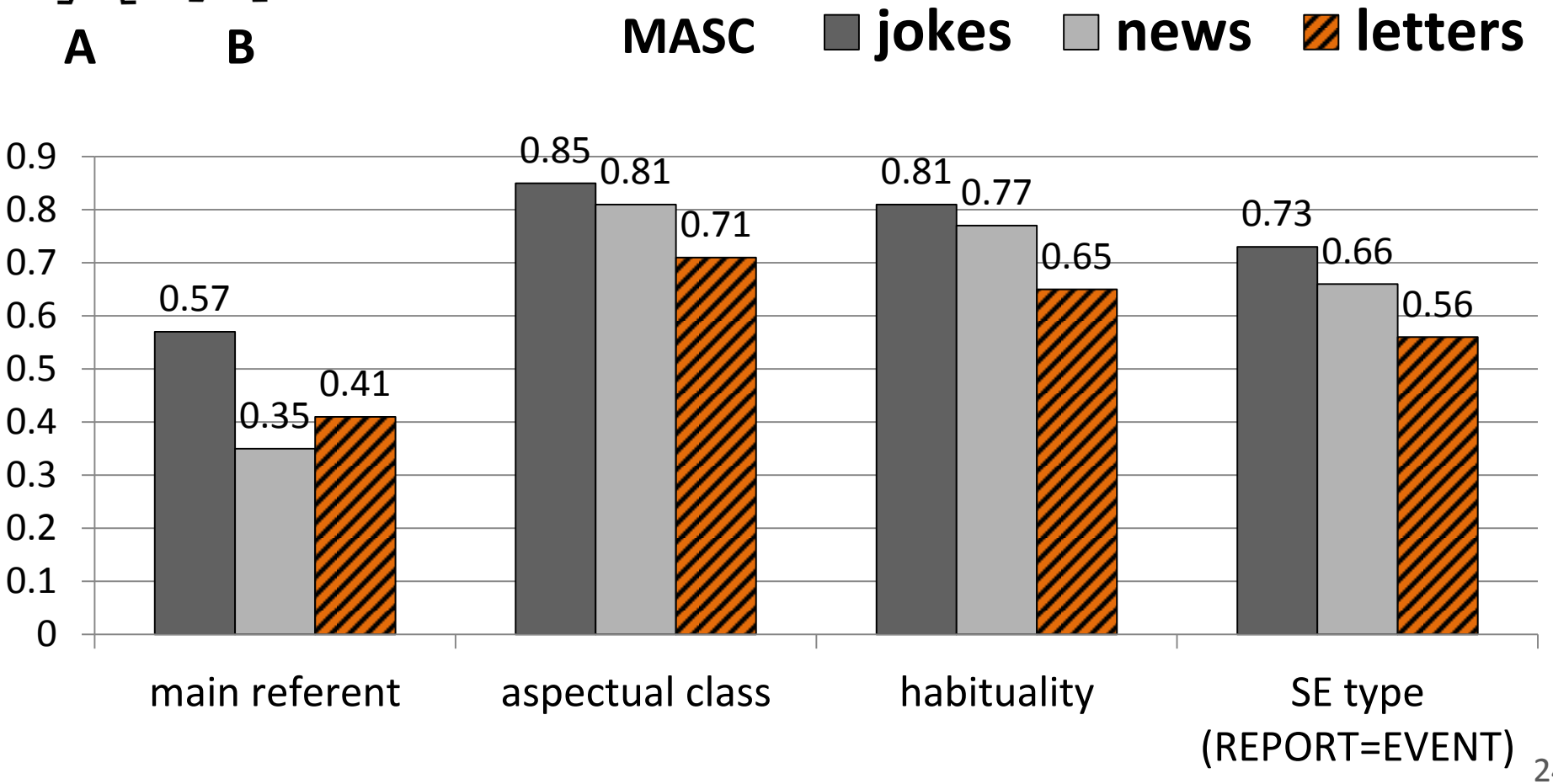
➡ follow-up study with data targeting generics in progress

➡ annotators with different preferences:
identify ambiguous cases

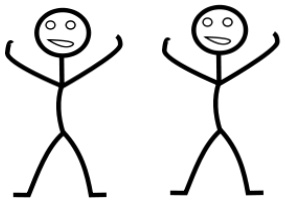
Inter-annotator agreement: genres



Why is agreement lower on letters subsection?

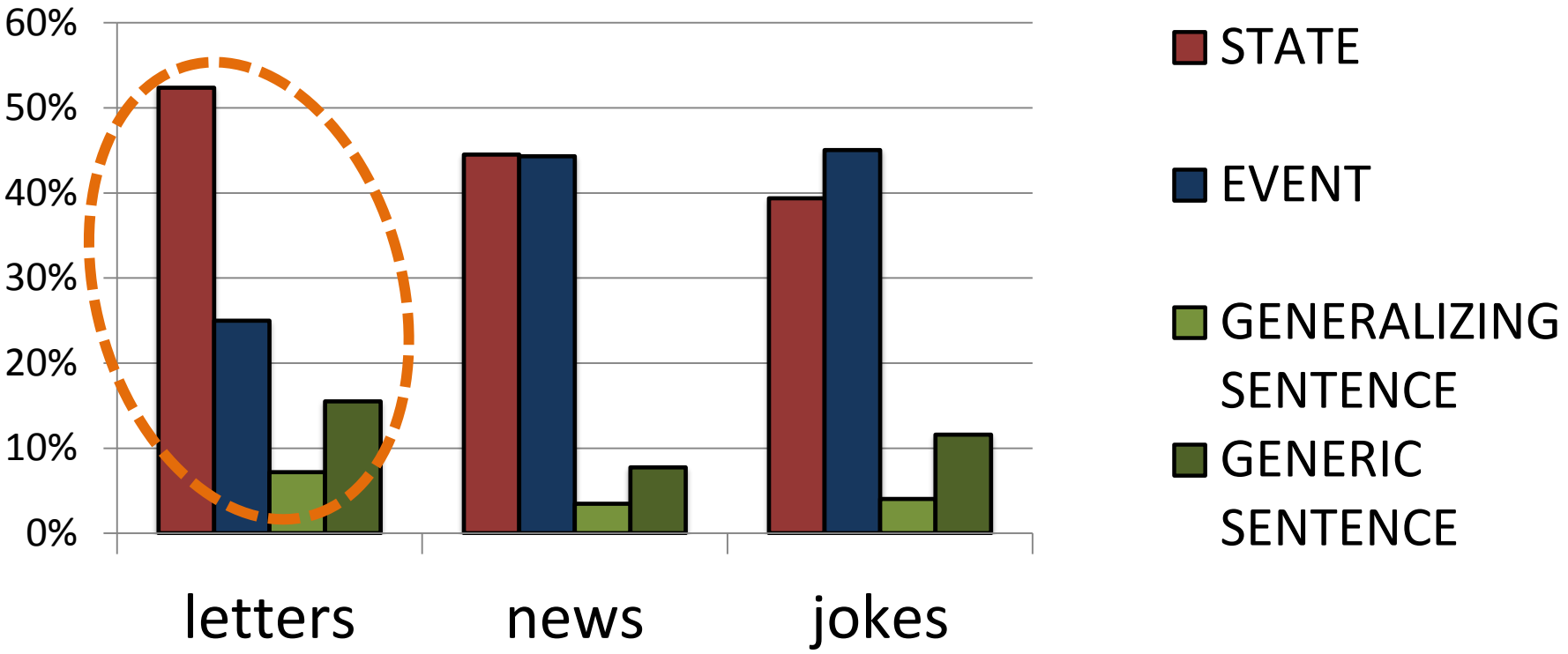


Distribution of SE types: genres



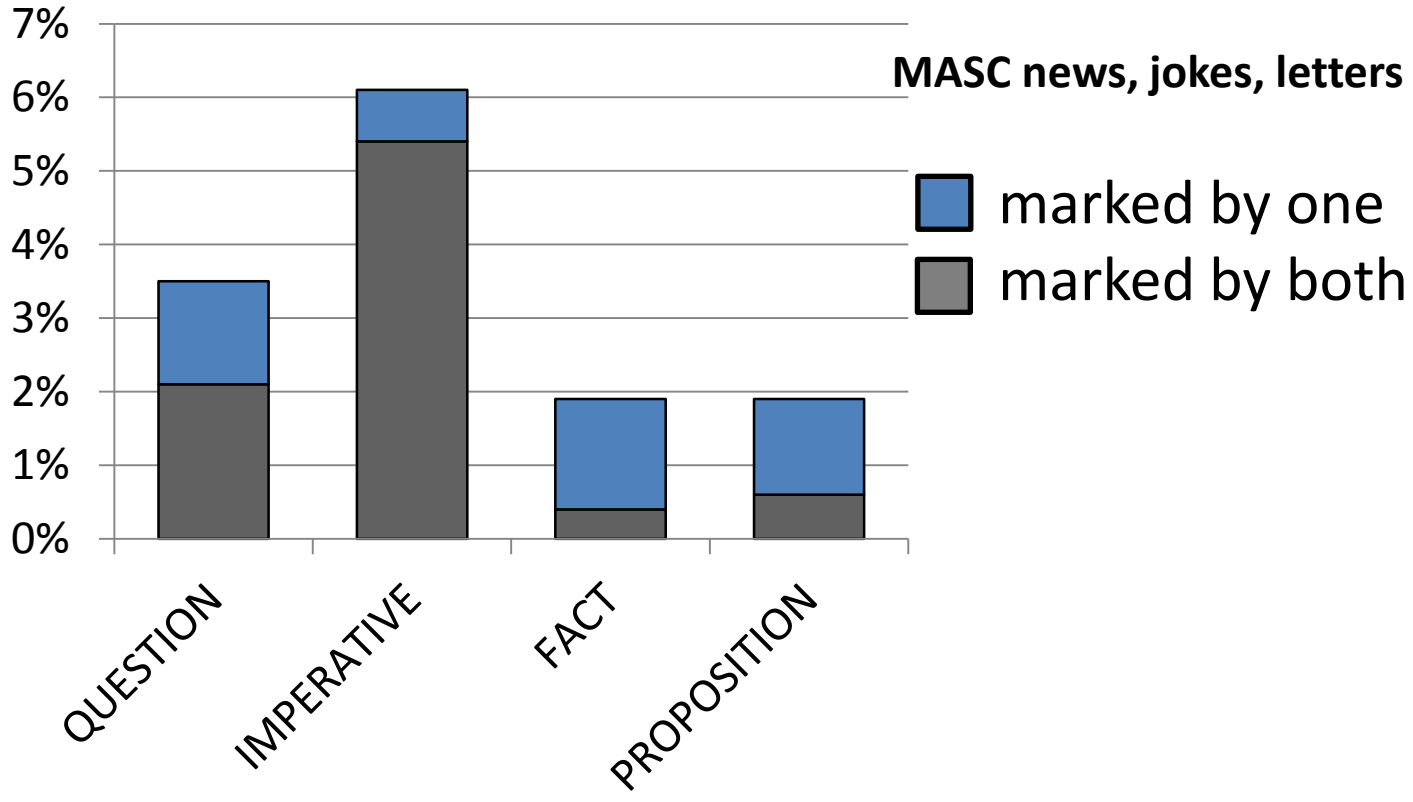
A B

average of SE labels assigned



letters has fewer events, more general statives

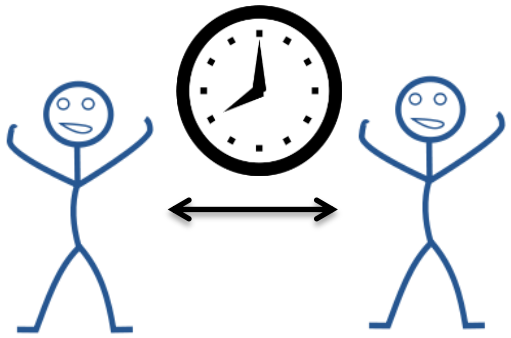
% of situations marked as speech acts / abstract entities:



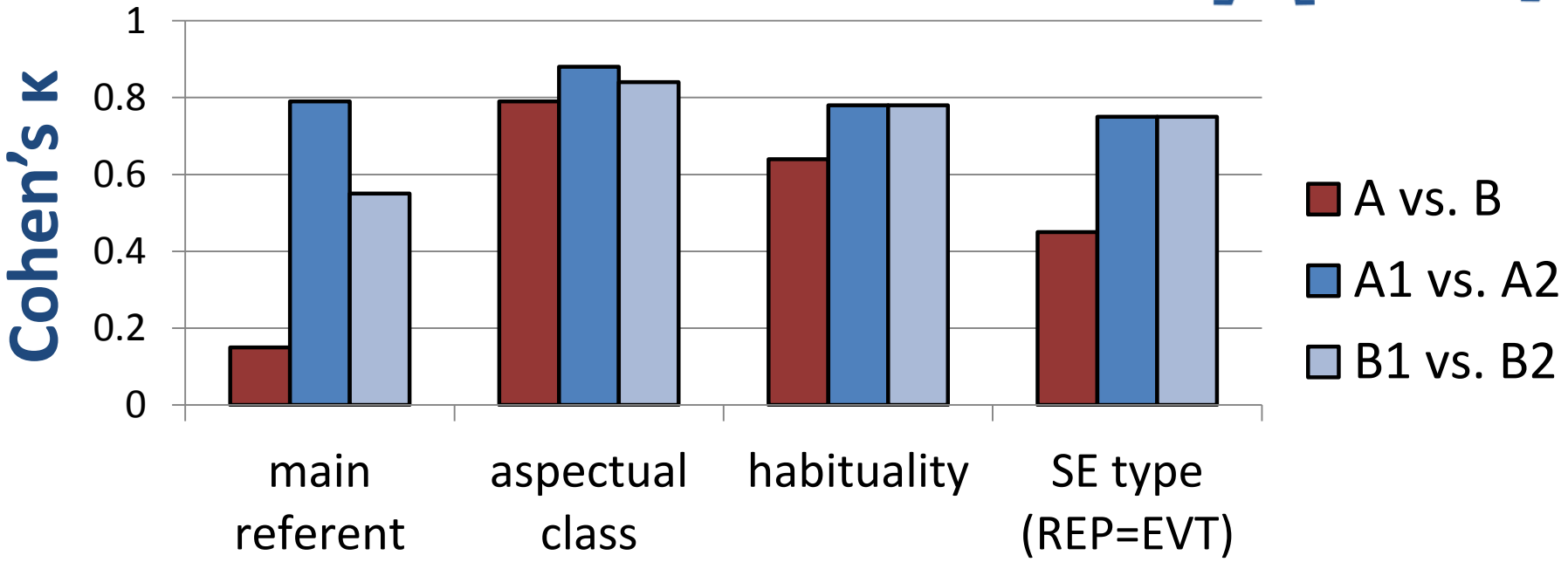
indirect questions?

no satisfying agreement yet – lacking recall?

Intra-annotator consistency



11 (5 news, 5 letters, 1 jokes) documents, 600 segments
(lowest agreements on SE type)



→agreement with oneself > agreement with other annotator

→different understanding of some cases

→some noisy cases: annotators *do* disagree with themselves

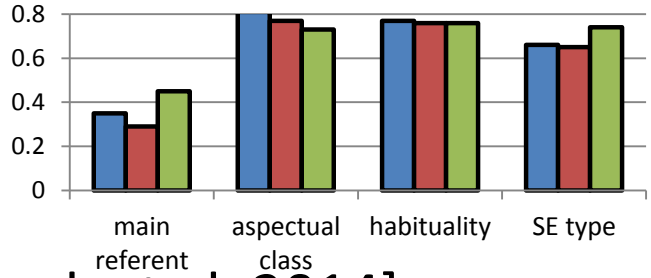
(but: hardest part of data set, total % of noise on SE type level << 20%)

Conclusions & future work

- Annotation guidelines for situation entity types:
  - substantial agreement achieved for SE type, aspectual class & habituality
  - part of disagreements: hard cases

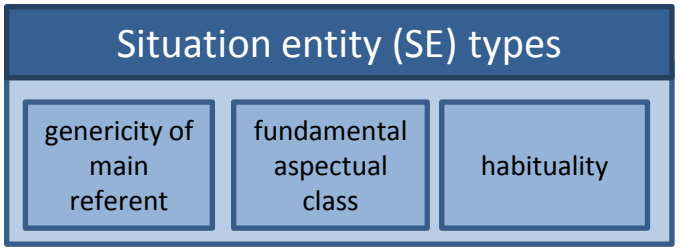
    → leverage for training

    [Beigman Klebanov & Beigman 2009, Plank et al. 2014]



- Feature-based approach
  - helps annotators during annotation
  - detailed analysis of annotator disagreements
  - identify problems in guidelines

    → follow-up study on genericity



Thanks to

Bonnie Webber

Andreas Peldzsus

Manfred Pinkal

Ambika Kirkland

Ruth Kühn

Fernando Ardente

References

Beata Beigman Klebanov and Eyal Beigman. 2009. **From annotator agreement to noise models.** Computational Linguistics, 35(4):495–503.

Steven Bethard. 2013. **ClearTK-TimeML: A minimalist approach to TempEval 2013**. In Second Joint Conference on Lexical and Computational Semantics (* SEM), volume 2, pages 10–14.

Francisco Costa and António Branco. 2012. **Aspectual type and temporal relation classification.** In Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics (EACL), pages 266–275.

Annemarie Friedrich and Alexis Palmer. 2014. **Automatic prediction of aspectual class of verbs in context.** In Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL). Baltimore, USA.

Nancy Ide, Christiane Fellbaum, Collin Baker, and Rebecca Passonneau. 2010. **The manually annotated subcorpus: A community resource for and by the people.** In Proceedings of the ACL 2010 conference short papers, pages 68–73.

Alexis Palmer, Elias Ponvert, Jason Baldridge, and Carlota Smith. 2007. **A sequencing model for situation entity classification.** Proceedings of ACL 2007.

References (ctd)

Nils Reiter and Anette Frank. 2010. **Identifying generic noun phrases.** In Proceedings of the 48th Annual Meeting of the Association for Computational Linguistics (ACL).

Eric V Siegel and Kathleen R McKeown. 2000. **Learning methods to combine linguistic indicators: Improving aspectual classification and revealing linguistic insights.** Computational Linguistics, 26(4):595–628.

Carlota S Smith. 2003. **Modes of discourse: The local structure of texts.** Cambridge University Press.

Radu Soricut and Daniel Marcu. 2003. **Sentence level discourse parsing using syntactic and lexical information.** In Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology-Volume 1, pages 149–156. Association for Computational Linguistics.

Naushad UzZaman, Hector Llorens, Leon Derczynski, Marc Verhagen, James Allen, and James Pustejovsky. 2013. Semeval-2013 task 1: **Tempeval-3: Evaluating time expressions, events, and temporaectual classification and revealing linguistic insights.** Computational Linguistics, 26(4):595–628.