States, events, and generics: computational modeling of situation entity types

CIS Talks LMU München, December 2016

Annemarie Friedrich, Saarland University (now CIS)
joint work with Alexis Palmer (now University of North Texas)
and Manfred Pinkal (Saarland University)

Thanks!

Alexis Palmer

Manfred Pinkal

Melissa Peate Sorensen

Liesa Heuschkel

Kleio-Isidora Mavridou

Christine Bocionek

Fernando Ardente

Ruth Kühn

Ambika Kirkland

Damyana Gateva

Discourse modes [Smith, 2003]

Prof. Dr. Origin at Saarland University came into his office one morning and was very surprised by the results of an experiment he had started the day before. He called in his assistants to inspect the hen and the egg that were the subject of his experiments…

The chicken or the egg causality dilemma is commonly stated as "which came first, the chicken or the egg?" To ancient philosophers, the question about the first chicken or egg also evoked the questions of how life and the universe in general began. …

In my opinion, the results of Prof. Dr. Origin's group are highly interesting, but they do by no means solve the philosophical question of how life and the universe began. I believe that much more research is needed, and that the field of biology alone will not be able to answer this question.

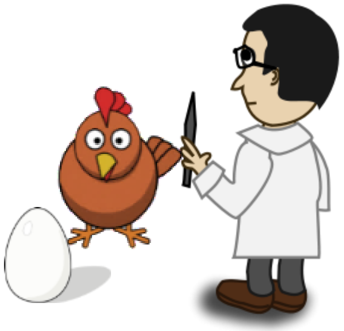
Discourse modes

Prof. Dr. Origin at Saarland University came into his office one morning and was very surprised by the results of an experiment he had started the day before. He called in his assistants to inspect the hen and the egg that were the subject of his experiments…

NARRATIVE

The chicken or the egg causality dilemma is commonly stated as "which came first, the chicken or the egg?" To ancient philosophers, the question about the first chicken or egg also evoked the questions of how life and the universe in general began. …

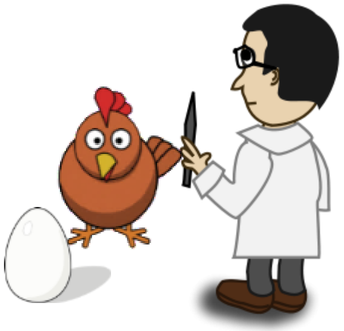
In my opinion, the results of Prof. Dr. Origin’s group are highly interesting, but they do by no means solve the philosophical question of how life and the universe began. I believe that much more research is needed, and that the field of biology alone will not be able to answer this question.



Discourse modes

Prof. Dr. Origin at Saarland University came into his office one morning and was very surprised by the results of an experiment he had started the day before. He called in his assistants to inspect the hen and the eggs that were the subject of his experiments…

NARRATIVE



The chicken or the egg causality dilemma is commonly stated as "which came first, the chicken or the egg?" To ancient philosophers, the question about the first chicken or egg also evoked the questions of how life and the universe in general began. …

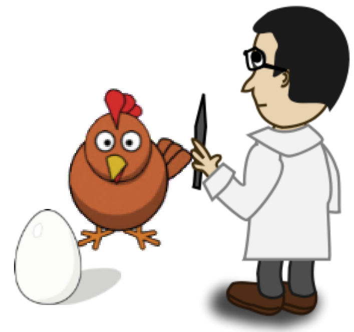
INFORMATION

In my opinion, the results of Prof. Dr. Origin's group are highly interesting, but they do by no means solve the philosophical question of how life and the universe began. I believe that much more research is needed, and that the field of biology alone will not be able to answer this question.

Discourse modes

Prof. Dr. Origin at Saarland University came into his office one morning and was very surprised by the results of an experiment he had started the day before. He called in his assistants to inspect the hen and the egg that were the subject of his experiments…

NARRATIVE



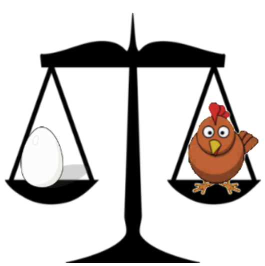
The chicken or the egg causality dilemma is commonly stated as "which came first, the chicken or the egg?" To ancient philosophers, the question about the first chicken or egg also evoked the questions of how life and the universe in general began. …

INFORMATION

In my opinion, the results of Prof. Dr. Origin's group are highly interesting, but they do by no means solve the philosophical question of how life and the universe began. I believe that much more research is needed, and that the field of biology alone will not be able to answer this question.

ARGUMENT

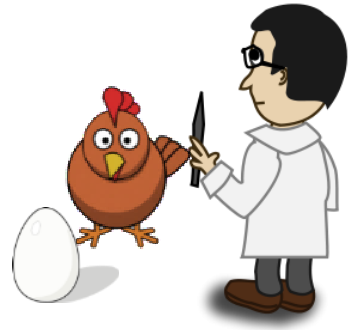
COMMENTARY



Discourse modes

Prof. Dr. Origin at Saarland University came into his office one morning and was very surprised by the results of an experiment he had started the day before. He called in his assistants to inspect the hen and the egg that were the subject of his experiments…

NARRATIVE



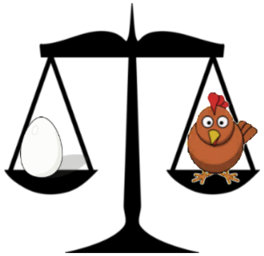
The chicken or the egg causality dilemma is commonly stated as "which came first, the chicken or the egg?" To ancient philosophers, the question about the first chicken or egg also evoked the questions of how life and the universe in general began. …

INFORMATION

In my opinion, the results of Prof. Dr. Origin's group are highly interesting, but they do by no means solve the philosophical question of how life and the universe began. I believe that much more research is needed, and that the field of biology alone will not be able to answer this question.

ARGUMENT

COMMENTARY



one text

≈ one genre

one passage

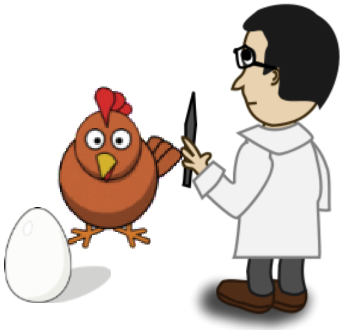
≈ one discourse mode

Discourse modes & situation entity types

Prof. Dr. Origin at Saarland University came into his office one morning and was very surprised by the results of an experiment he had started the day before. He called in his assistants to inspect the hen and the egg that were the subject of his experiments…

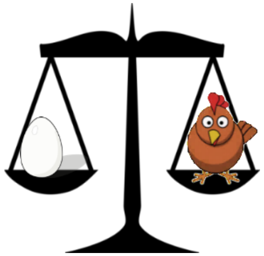
The chicken or the egg causality dilemma is commonly stated as "which came first, the chicken or the egg?" To ancient philosophers, the question about the first chicken or egg also evoked the questions of how life and the universe in general began. …

In my opinion, the results of Prof. Dr. Origin's group are highly interesting, but they do by no means solve the philosophical question of how life and the universe began. I believe that much more research is needed, and that the field of biology alone will not be able to answer this question.



NARRATIVE

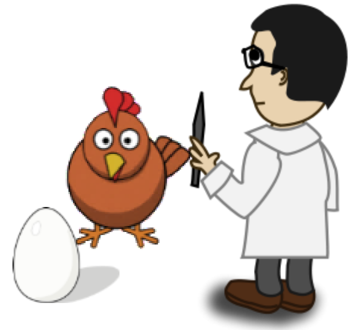
INFORMATION



ARGUMENT
COMMENTARY

Discourse modes & situation entity types

Prof. Dr. Origin at Saarland University came into his office one morning and was very surprised by the results of an experiment he had started the day before. He called in his assistants to inspect the hen and the egg that were the subject of his experiments…



NARRATIVE

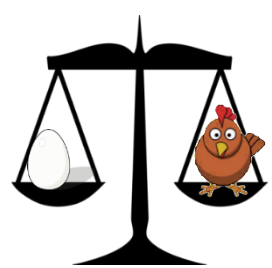
STATE

EVENT

The chicken or the egg causality dilemma is commonly stated as "which came first, the chicken or the egg?" To ancient philosophers, the question about the first chicken or egg also evoked the questions of how life and the universe in general began. …

INFORMATION

In my opinion, the results of Prof. Dr. Origin's group are highly interesting, but they do by no means solve the philosophical question of how life and the universe began. I believe that much more research is needed, and that the field of biology alone will not be able to answer this question.

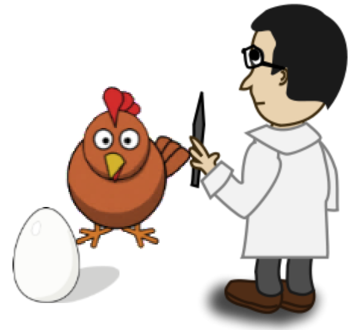


ARGUMENT

COMMENTARY

Discourse modes & situation entity types

Prof. Dr. Origin at Saarland University came into his office one morning and was very surprised by the results of an experiment he had started the day before. He called in his assistants to inspect the hen and the egg that were the subject of his experiments…



NARRATIVE

State

Event

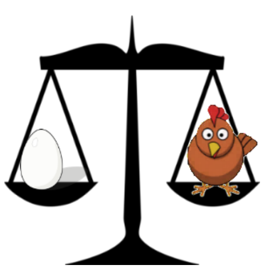
The chicken or the egg causality dilemma is commonly stated as "which came first, the chicken or the egg?" To ancient philosophers, the question about the first chicken or egg also evoked the questions of how life and the universe in general began. …

INFORMATION

Generic Sentence

Generalizing Sentence

In my opinion, the results of Prof. Dr. Origin's group are highly interesting, but they do by no means solve the philosophical question of how life and the universe began. I believe that much more research is needed, and that the field of biology alone will not be able to answer this question.

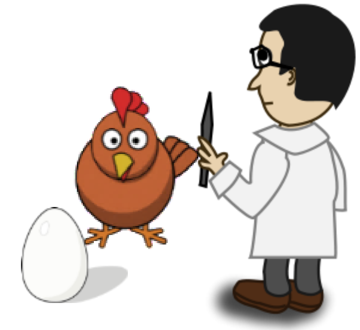


ARGUMENT

COMMENTARY

Discourse modes & situation entity types

Prof. Dr. Origin at Saarland University came into his office one morning and was very surprised by the results of an experiment he had started the day before. He called in his assistants to inspect the hen and the egg that were the subject of his experiments…



NARRATIVE

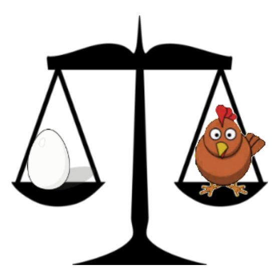
State
Event

The chicken or the egg causality dilemma is commonly stated as "which came first, the chicken or the egg?" To ancient philosophers, the question about the first chicken or egg also evoked the questions of how life and the universe in general began. …

INFORMATION

Generic Sentence
Generalizing Sentence

In my opinion, the results of Prof. Dr. Origin's group are highly interesting, but they do by no means solve the philosophical question of how life and the universe began. I believe that much more research is needed, and that the field of biology alone will not be able to answer this question.



ARGUMENT
COMMENTARY

State, Event, Abstract Entities, Generic / Generalizing Sentences

Situation entity types

- Inventory of **aspectual clause types** motivated by a theory of discourse [Smith 2003]

What is clause-level aspect?

- **aktionsart**

  [Vendler 1957, Bach 1986]

state	The ship is in motion.
event	The ship moved.
process	The ship is moving.

- **habituals / genericity** [Krifka et al. 1995]

John cycles to work.
Students like coffee.

Why model these phenomena?

- understand temporal relations in discourse
- distinguish between / extract different types of knowledge
- identify different modes of discourse

Situation entity types: summary

coercion to STATE:
negation, modality, future, perfect, conditionality, subjectivity

	Julie likes Cooper. Julie did **not** kill the mouse.
	Julie met Cooper two years ago.
	..., said the zookeeper.
GENERIC SENTENCE	**Owls** are nocturnal animals.
GENERALZING SENTEN	Julie often teases Cooper.
IM	
QU	s on your slides?

What are the major differences between these types?

Situation entity types: summary

Does the verb express an **event** or a **state**?

lexical aspectual class

	Julie likes Cooper. Julie did **not** kill the mouse.
EVENT	Julie met Cooper two years ago.
- REPORT	..., said the zookeeper.
GENERIC SENTENCE	**Owls** are nocturnal animals.
	Julie often teases Cooper.
	Catch the mo ?
	Why are the on your slides?

Does the sentence talk about a **particular referent** or a **kind/class**?

genericity

Does something happen repeatedly or once?

habituality

Data sets and annotation procedure

MASC
25,000 clauses
essays, letters, fiction, technical, travel, news …

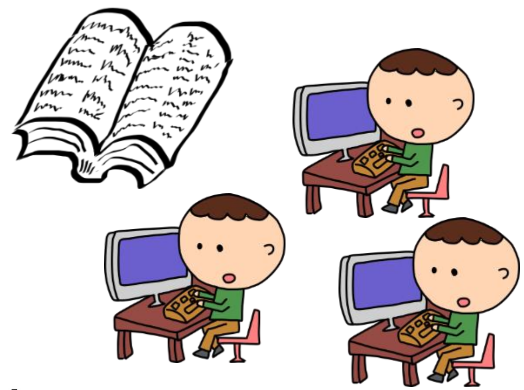
Wikipedia
10,000 clauses
botany, animals, sports, biographies, science, …

segmentation into clauses (SPADE)

Annotators label

- situation entity type
- genericity of main referent
- lexical aspectual class of main verb
- habituality of main verb

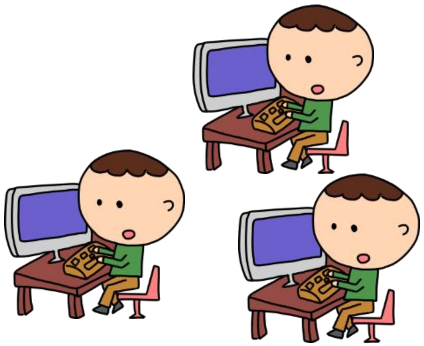
training phase
+ manual



gold standard = majority vote over labels of 3 annotators

(about 10% of automatically created segments marked as "NO SITUATION")

Inter-annotator agreement



Fleiss' κ

= how much agreement beyond chance was reached

Fleiss' κ: features		MASC / Wiki
aspectual class	stative, dynamic, both	0.69 / 0.64
main referent	generic, non-generic, cannot decide	0.69 / 0.65
habituality	episodic, static, habitual, cannot decide	0.55 / 0.67

Inter-annotator agreement

	% in gold standard		Fleiss' κ
Situation entity type	MASC	Wikipedia	Krippendorff's diagnostics
State	49.8	24.3	0.67
Event	24.3	18.9	0.74
Report	4.8	0.9	0.80
Generic Sentence	7.3	49.7	0.68
Generalizing Sentence	3.8	2.5	0.43
Question	3.3	0.1	0.91
Imperative	3.2	0.2	0.94
undecided	2.4	2.1	-

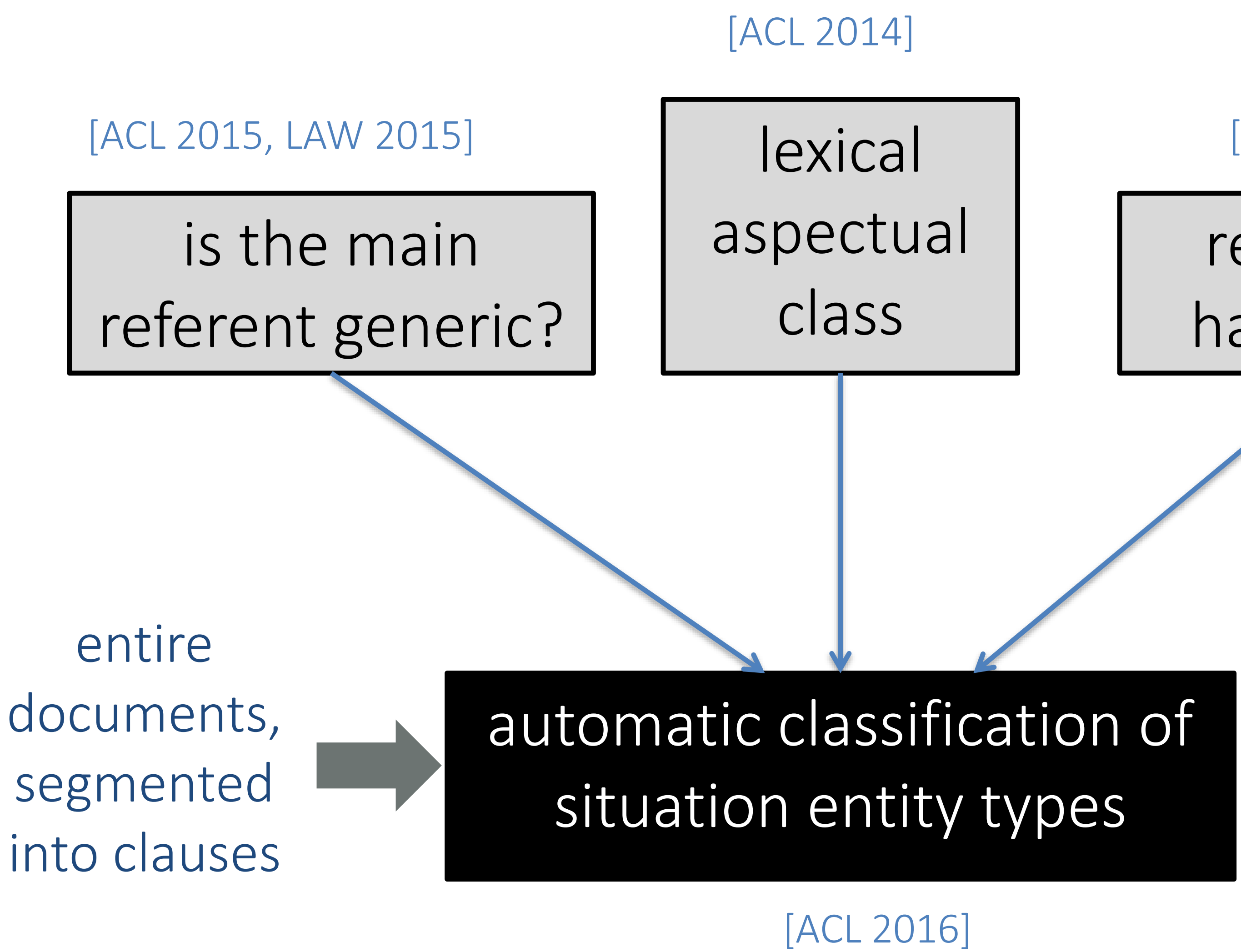
Related work

- **modeling of Vendler classes**
  - state, activity, accomplishment, achievement
  - Italian [Zarcone & Lenci 2008], German [Hermes et al. 2015]
  - stative vs. dynamic, completedness [Siegel & McKeown 2000]

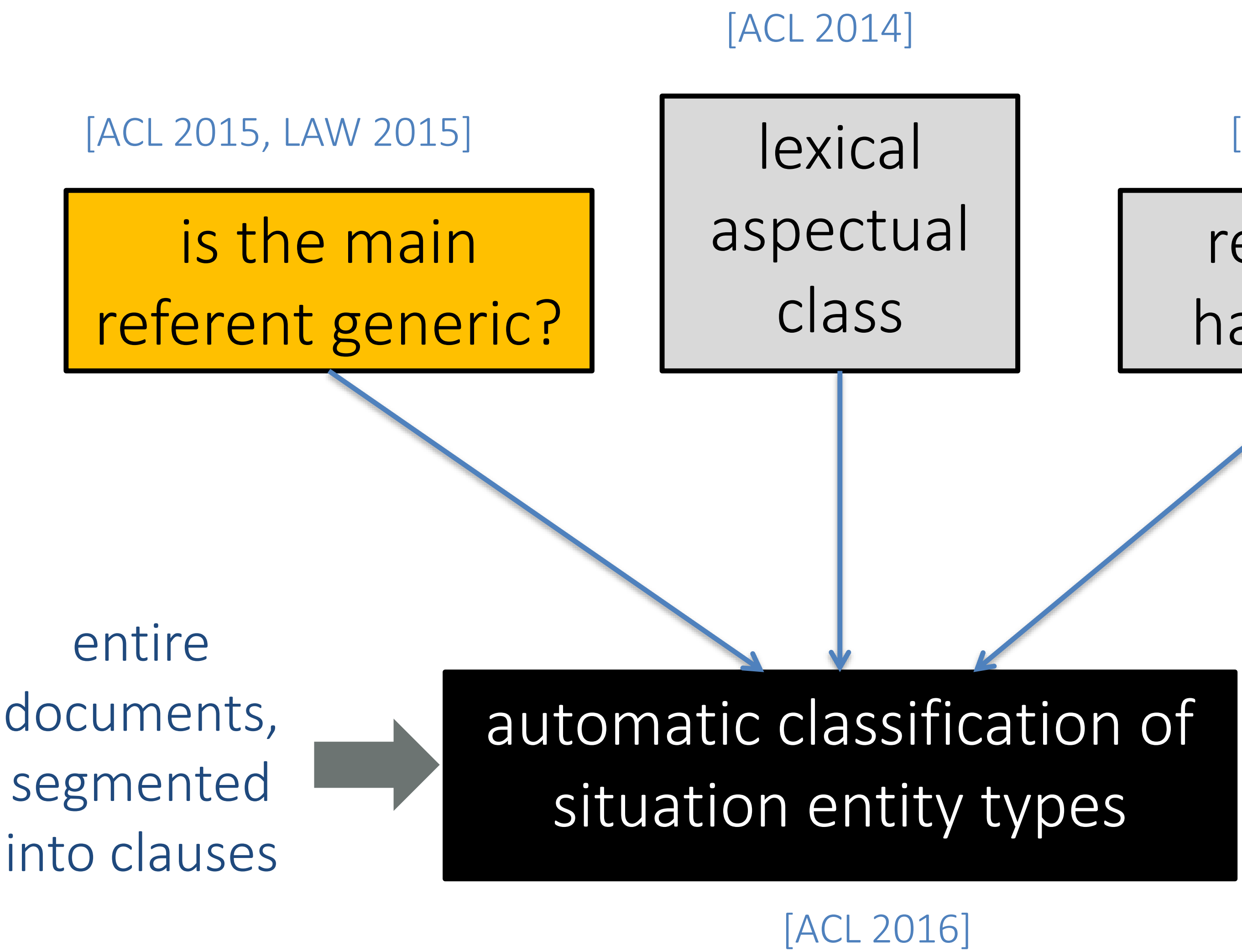
- **modeling genericity**
  - identifying genericity of NPs / reference to kinds [Reiter & Frank 2010]
  - recognizing habituals [Mathew & Katz 2009]

- **labeling situation entities** [Palmer et al. 2007]
  - maximum entropy model, features: pos tags, words, linguistic
  - data set: 20 texts / 4391 clauses, Brown corpus, κ=0.52

Computational modeling of situation entity types

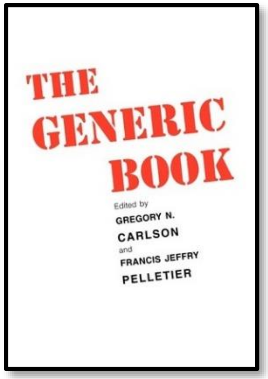


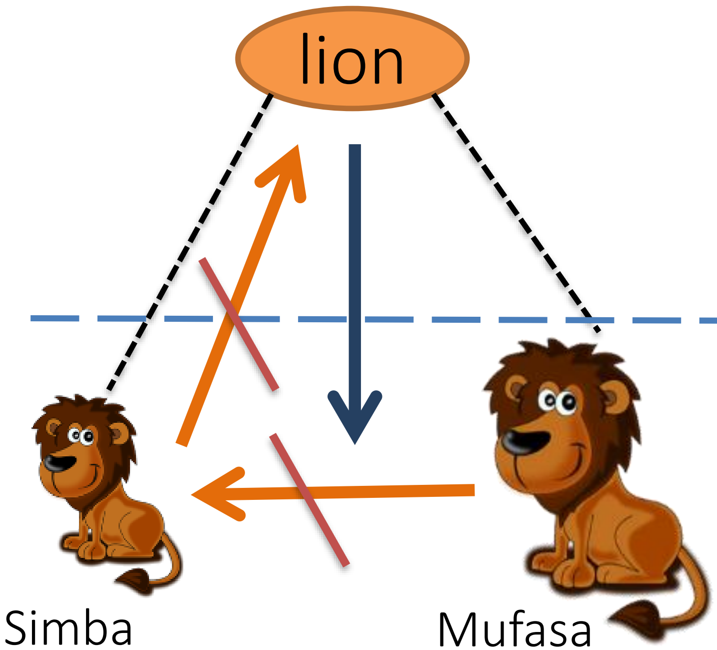
Computational modeling of situation entity types



Genericity

Krifka, Manfred, et al.
Introduction to genericity.
In *The Generic Book* (1995).





different
entailment properties

<u>Lions</u> are dangerous.

kind-referring
generic

<u>Mufasa</u> is dangerous.
<u>Simba</u> is dangerous.

non-generic

Reference to kinds

	kind-referring	non-kind-referring
definite NPs	The lion is a predatory cat.	The cat chased the mouse.
indefinite NPs	Lions eat meat.	Dogs were barking outside.
quantified NPs	Some (type of) dinosaur is extinct.	Some dogs were barking outside.
proper names	Panthera leo persica was first described by the Austrian zoologist Meyer.	John likes ice cream.

form of NP not sufficient

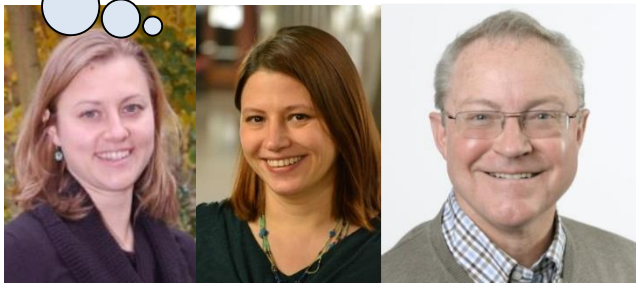
clause / context matters

Discourse-sensitive approach

WIKIPEDIA
The Free Encyclopedia

[**The recent year's growth twigs**]
are green and turn dark brown.

It‘s impossible to label this without discourse context!



Discourse-sensitive approach

WIKIPEDIA
The Free Encyclopedia

[**Sugar maples** generic] also have a tendency to color unevenly in fall.
[**The recent year's growth twigs** generic] are green and turn dark brown.

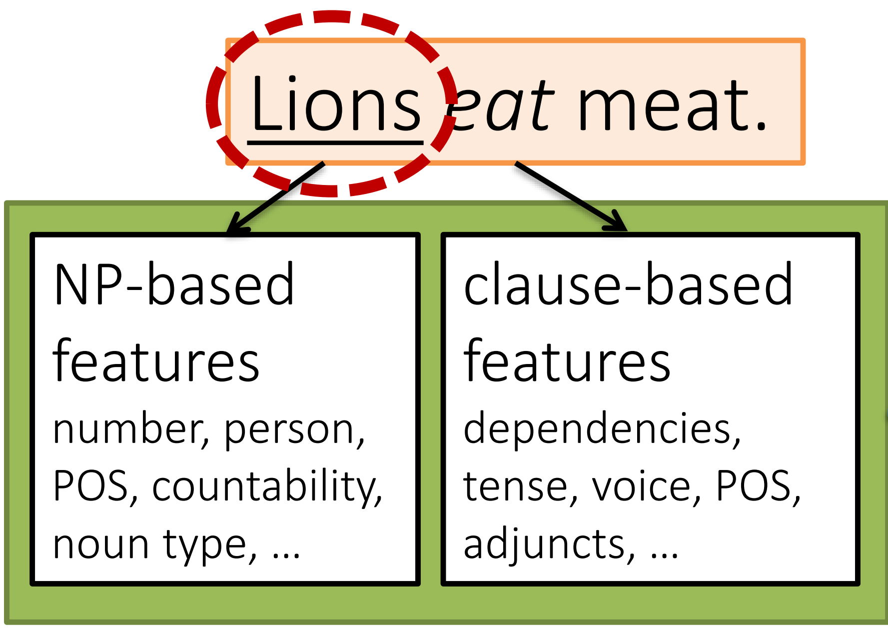
genericity labeling of noun phrases in entire texts
→ sequence labeling task

Baseline: identifying generic noun phrases

Data: ACE-2 & ACE-2005

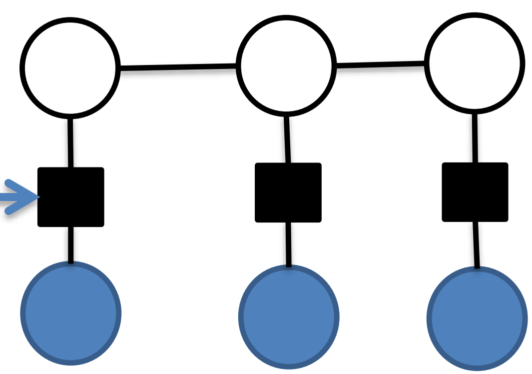
→ largest corpora annotated with NP-level genericity to date, ~40k NPs

- SPC = specific / non-generic
- GEN = generic
- USP = underspecified

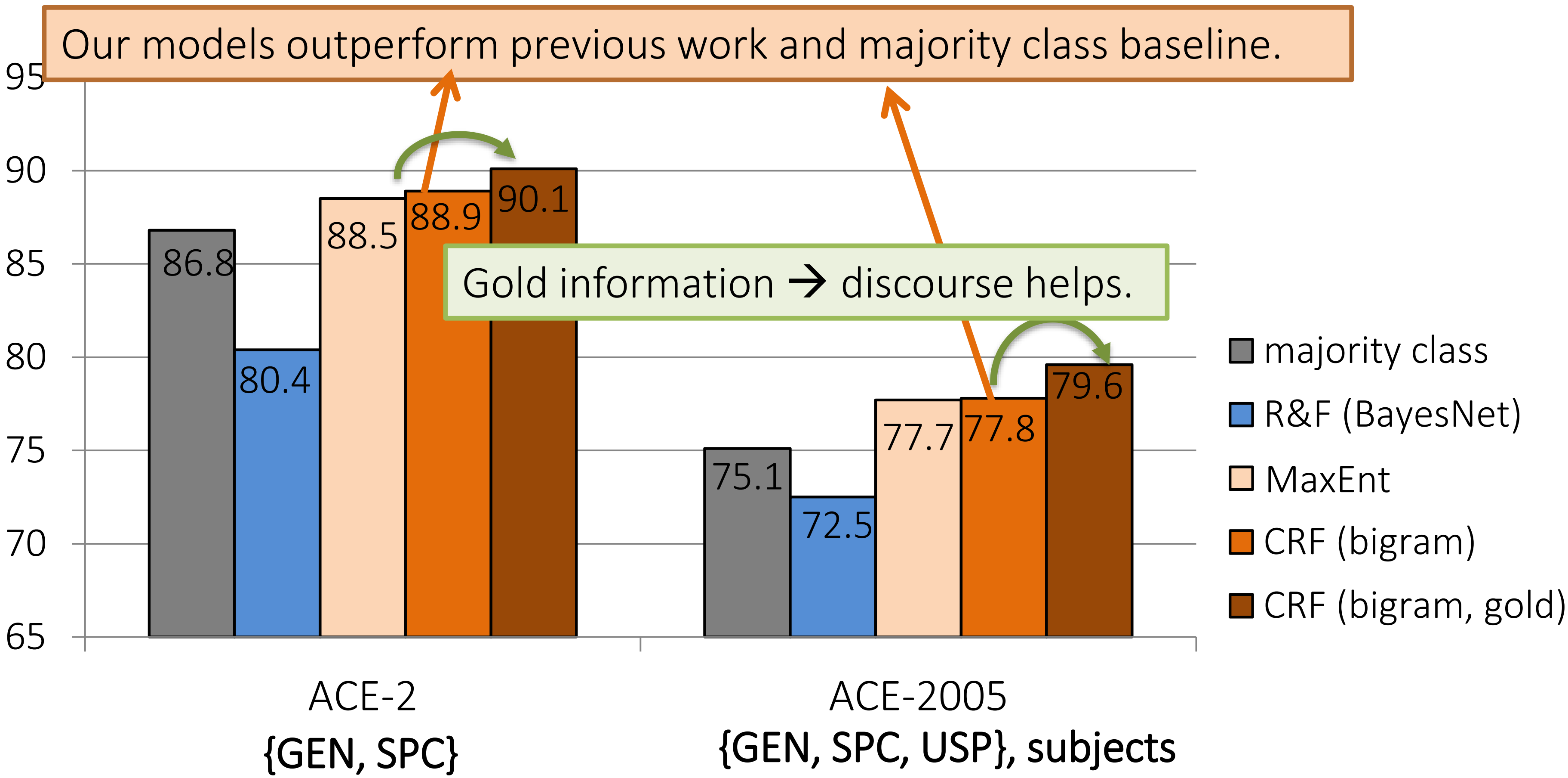


Bayesian network [Weka]
[Reiter & Frank ACL 2010]
local decision for each NP

Linear Chain CRF
[Friedrich & Pinkal ACL 2015]
labels assigned to other NPs/clauses influence the decisions



Accuracy: ACE-2 and ACE-2005

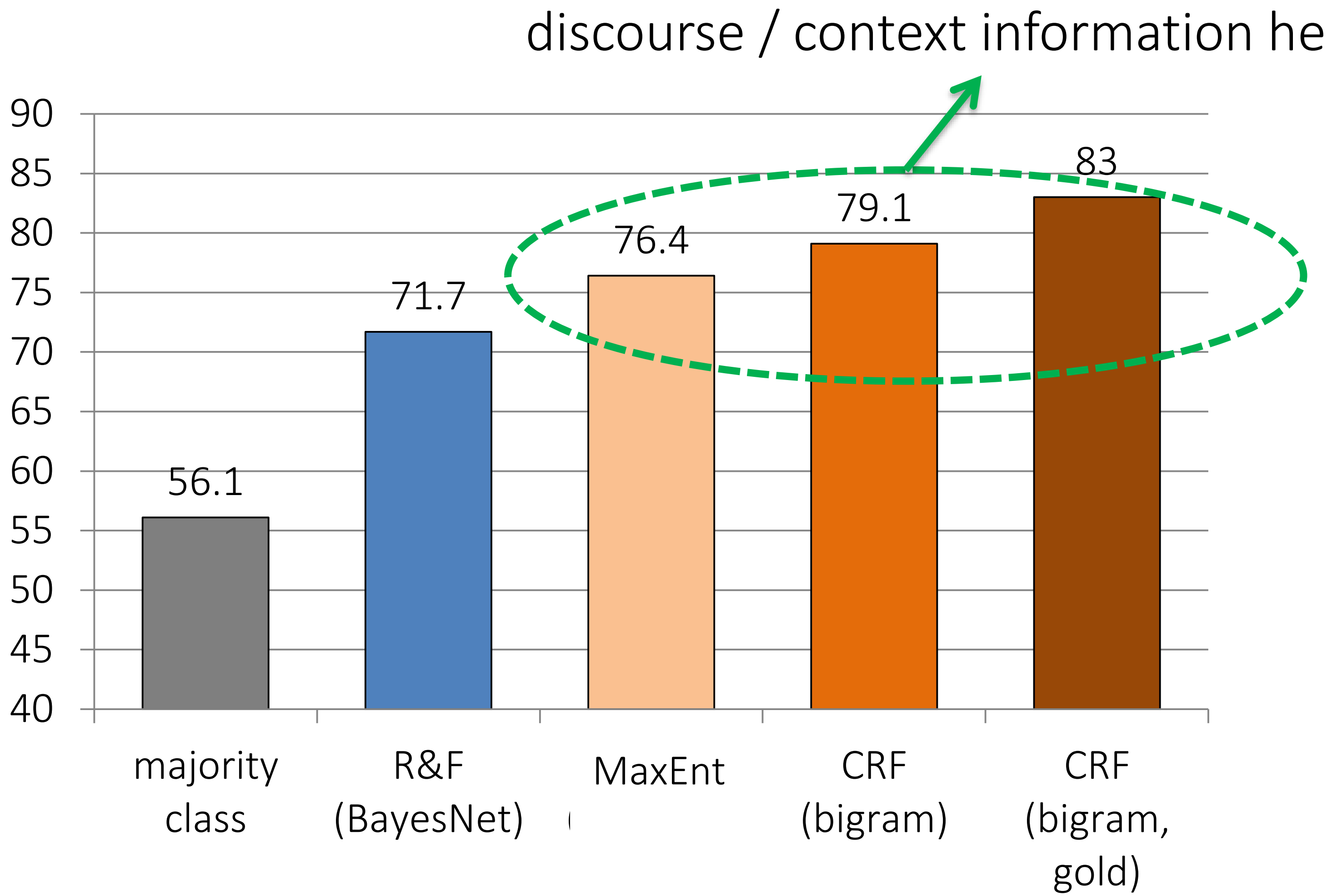


Few generic instances. [Friedrich et al. LAW 2015]

Problems in annotation guidelines, mix genericity and specificity.

→ *Officials reported…* (USP) → is non-generic, non-specific! → SPC

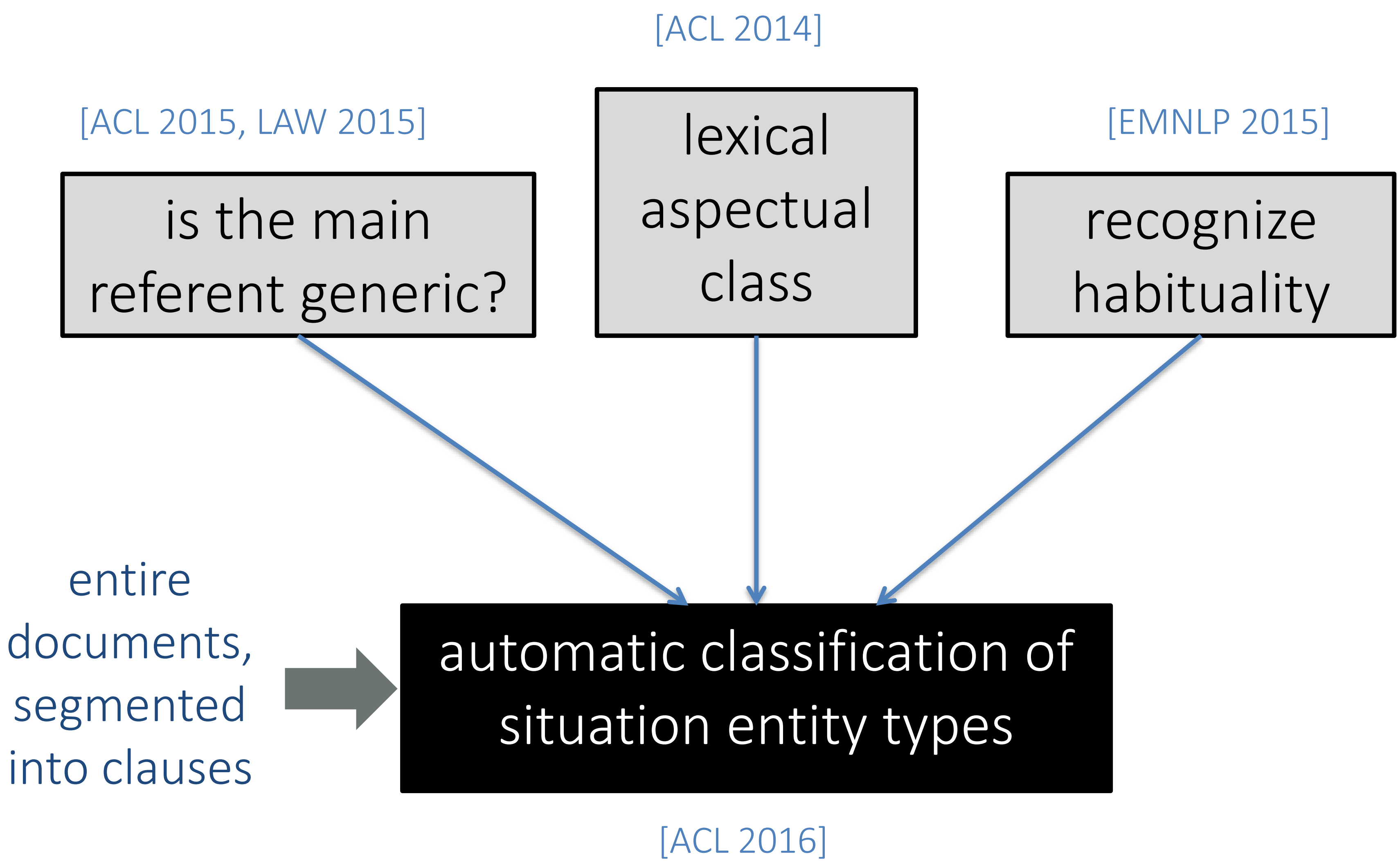
Accuracy: Wikipedia data (main referent)



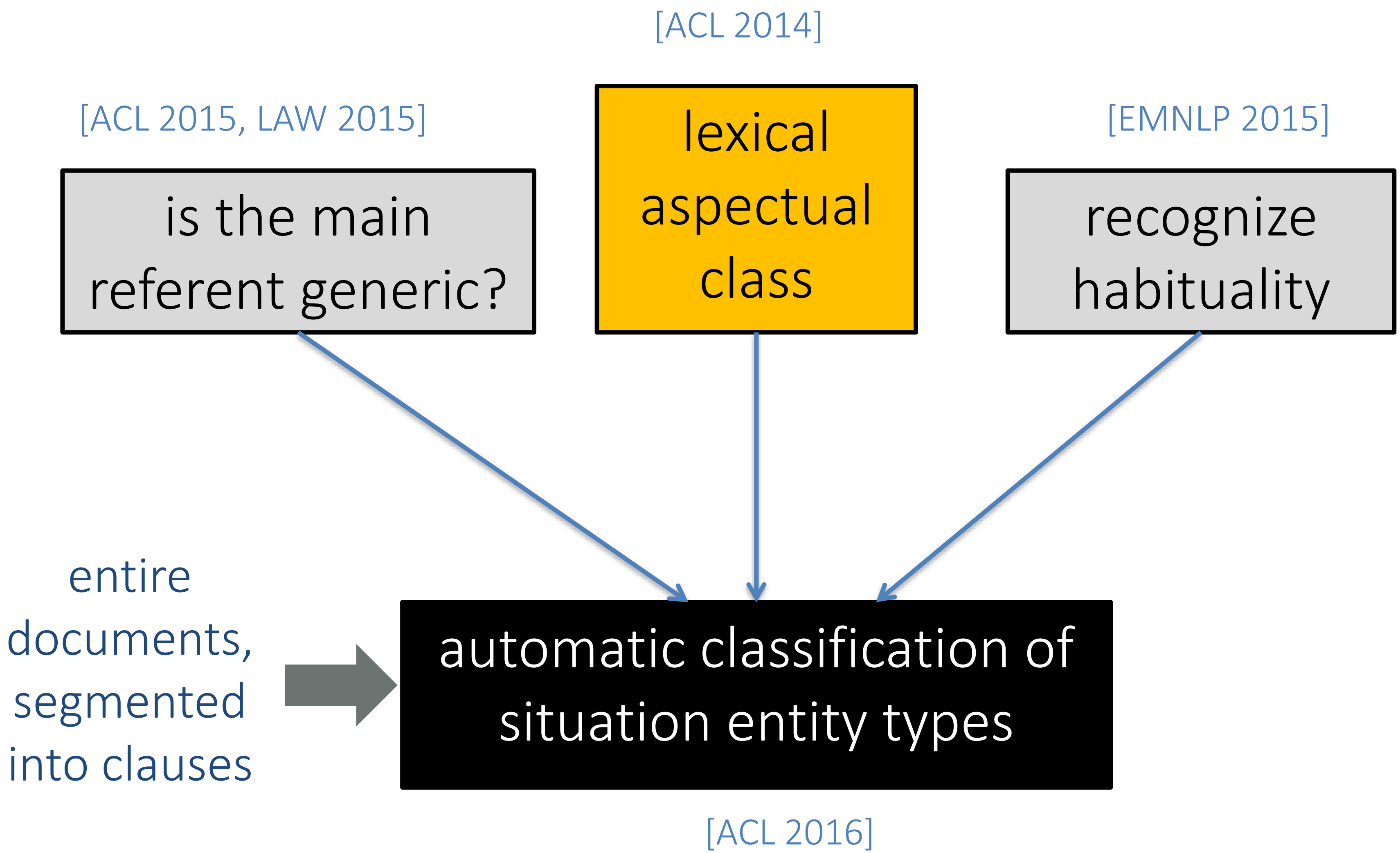
{generic, non-generic}

all differences statistically significant

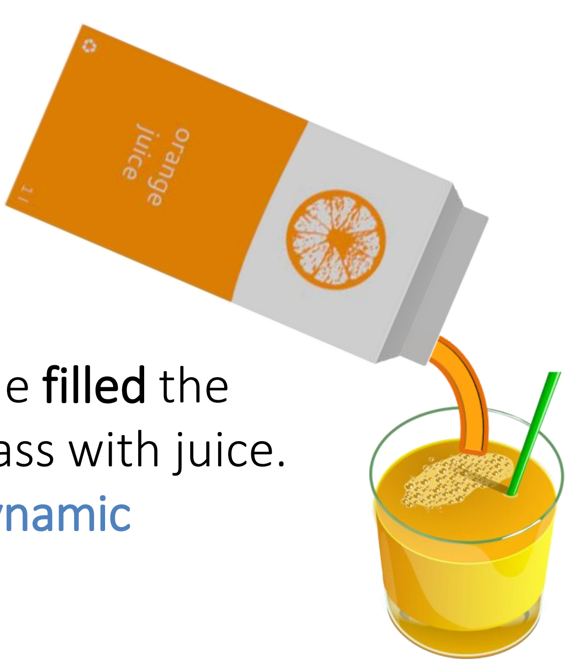
Computational modeling of situation entity types



Computational modeling of situation entity types



Lexical aspectual class



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

Juice **fills** the glass.
stative

The glass **was filled** with juice.
both interpretations possible

Bach [1986]: time schemata of <u>sentences</u>

- eventuality type
  - state
  - non-states
    - process
    - event

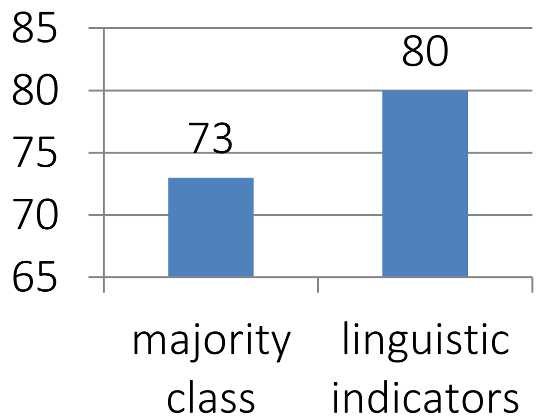
Vendler [1957]: time schemata of <u>verbs</u>
lexical aspect / aktionsart

states	*love, own*	stative
activities	*run*	
accomplishments	*write a letter*	dynamic
achievements	*realize*	

Predicting fundamental aspectual class

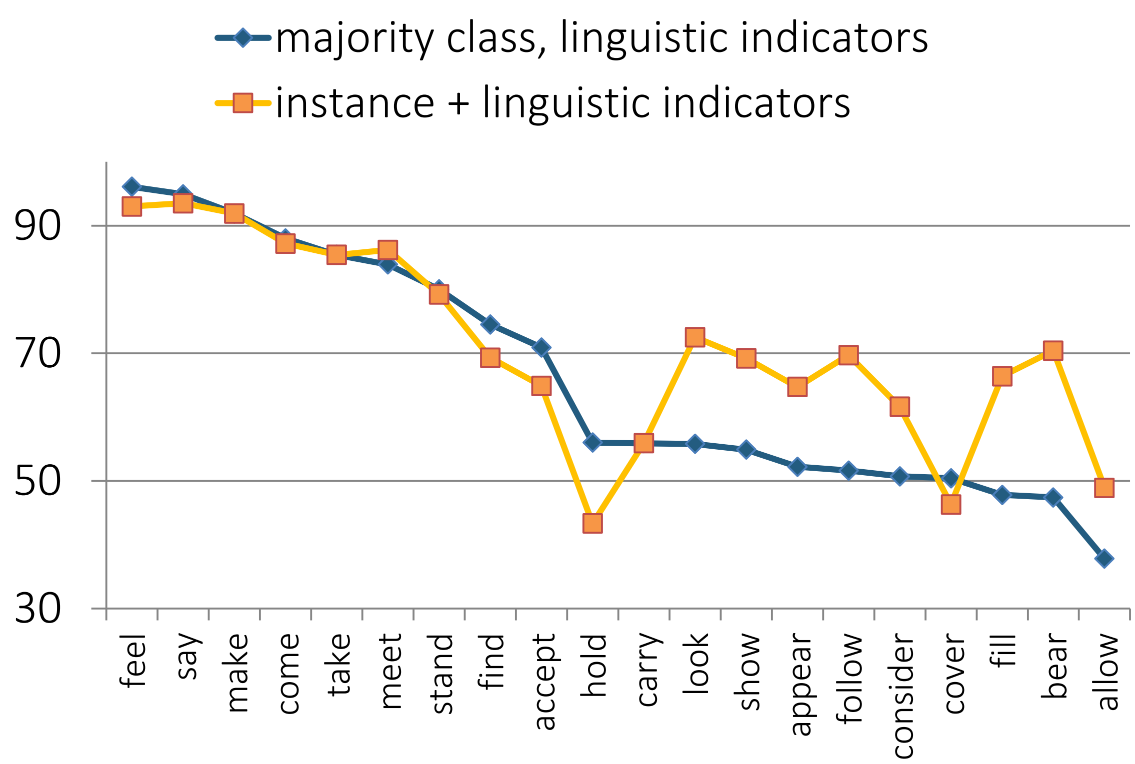
John will **love** this cake!	stative	*John love cake*
John has **kissed** Mary.	dynamic	*John kiss Mary*
John **drives** to work.	dynamic	*John drive to work*

- **Features**: linguistic indicators → patterns how verb types behave in a large parsed corpus, e.g., how often they occur with the progressive, with certain adverbials, ... [Siegel & McKeown 2000]
- **Data set:** MASC letter, essays, news annotated for aspectual class on clause level: *The glass is **filled** with juice. stative*
- **Finding**: linguistic indicators generalize across verb types



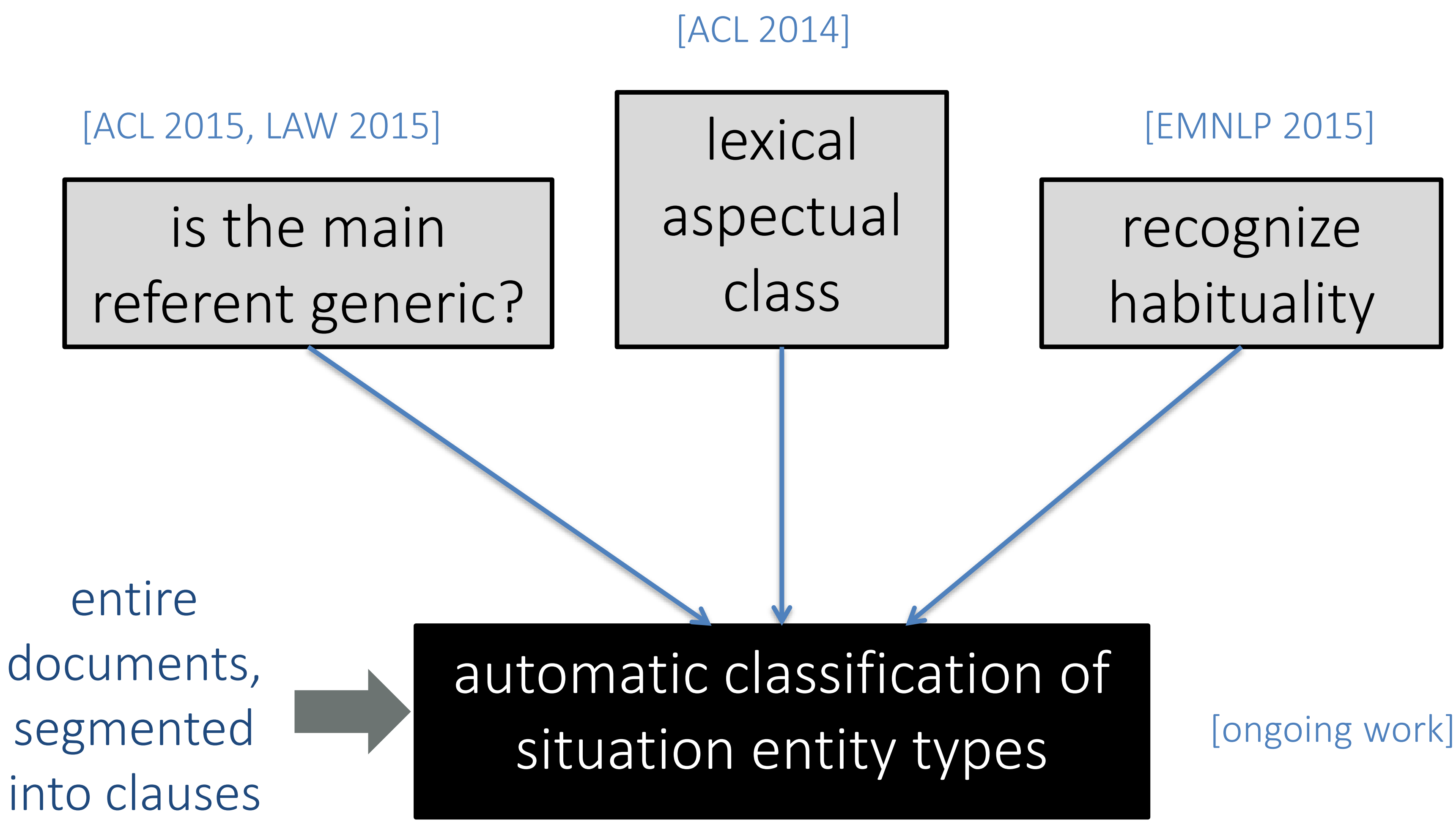
Predicting fundamental aspectual class

- **Problem**: features are type-based → performance never better than guessing the majority class per verb type
- **Solution**: add in instance-based features (syntactic-semantic features reflecting context of each verb occurrence)

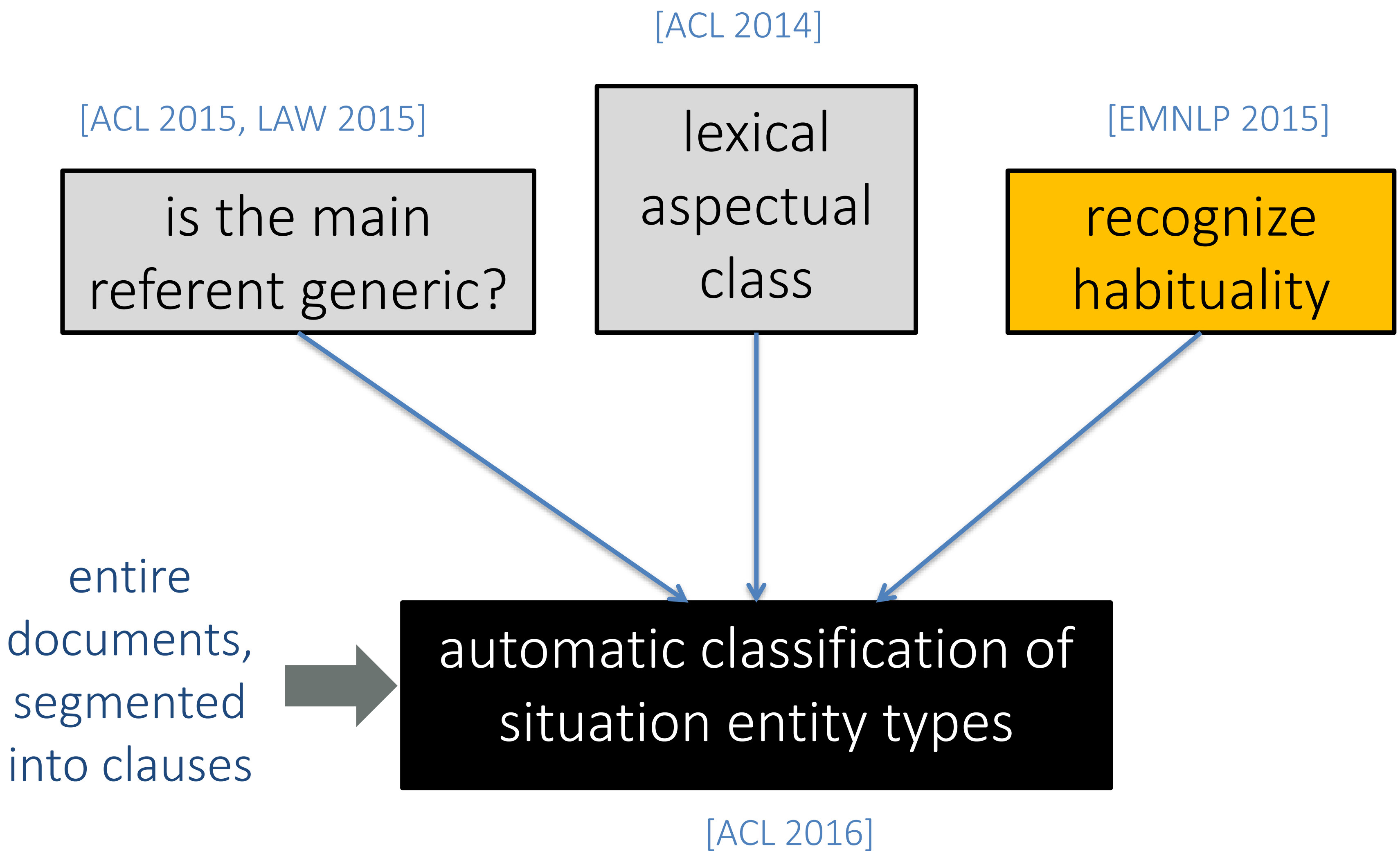


Brown corpus
130 sentences per verb type, leave-one-out CV
labels: stative, dynamic, both readings possible

Computational modeling of situation entity types



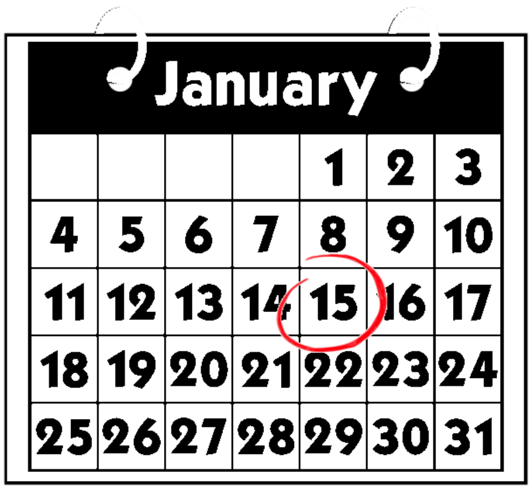
Computational modeling of situation entity types



Habituality

episodic

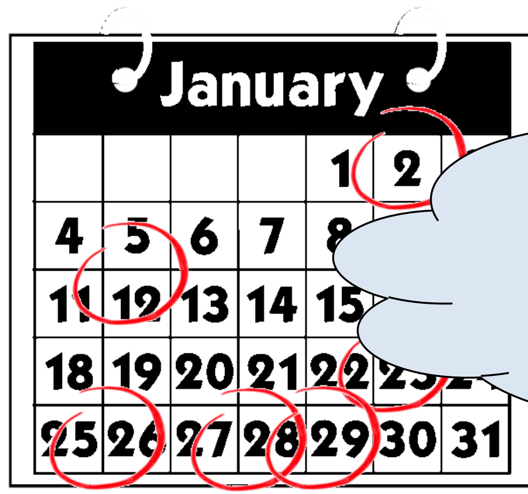
a particular event



John went swimming yesterday!

habitual

generalization over situations, exceptions are tolerated



Bill often goes swimming.

Bill **likes** coffee.
Bill **didn't** go swimming.
Bill **can** swim.

Mathew & Katz [2009]

A three-way classification of clausal aspect

clausal aspect		lexical aspect
episodic	Bill **drank** a coffee after lunch.	*dynamic*
habitual	Bill *usually* **drinks** coffee after lunch.	*dynamic*
	Italians **drink** coffee after lunch.	*dynamic*
	Sloths *sometimes* **sit** on top of branches.	*stative*
	John *never* **drinks** coffee.	*dynamic*
static	Bill **likes** coffee.	*stative*
	Bill *can* **swim**.	*dynamic*
	Bill *didn't* **drink** coffee yesterday.	*dynamic*
	Mary *has* **made** a cake.	*dynamic*

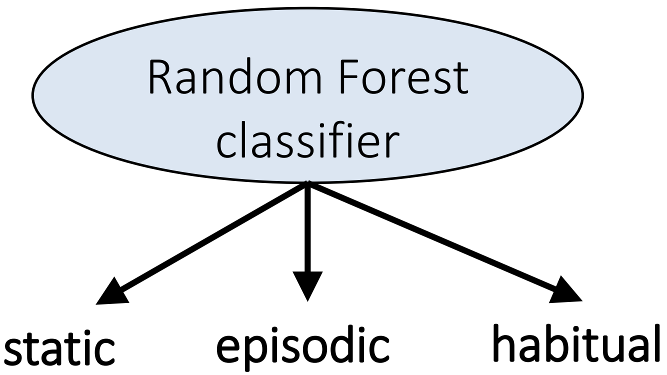
[Friedrich & Pinkal, EMNLP 2015]

Automatic classification of clausal aspect

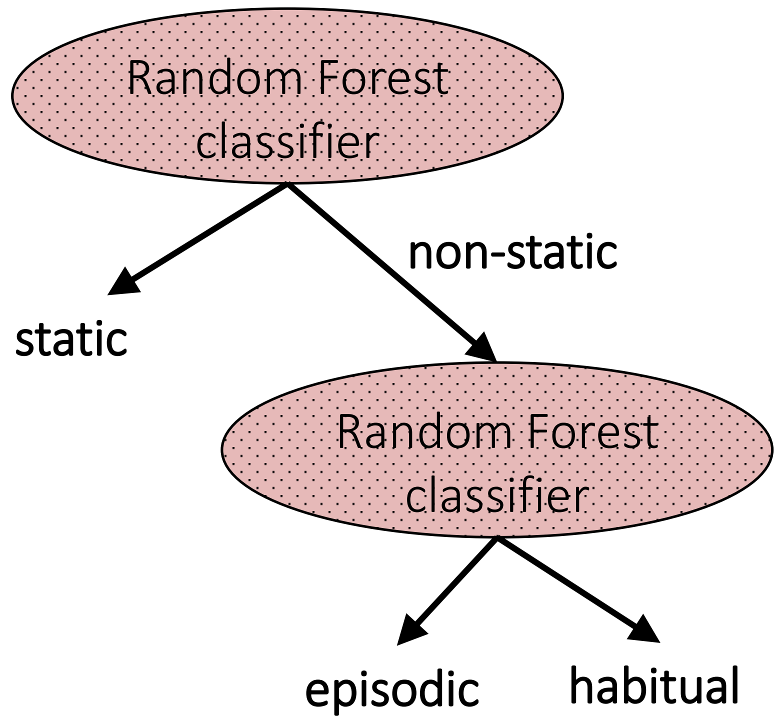
Features:

- instance-based features
- type-based features (linguistic indicators)

Joint model



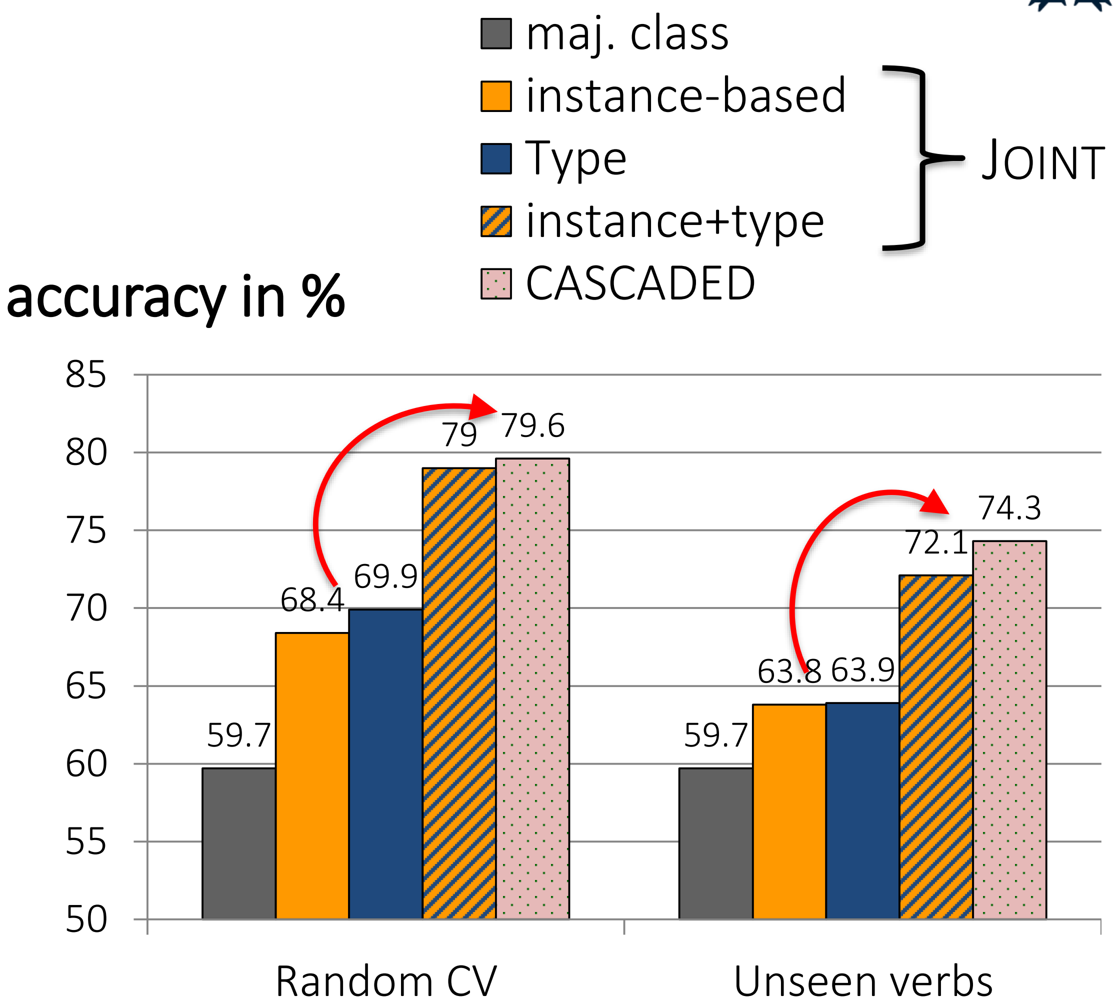
Cascaded model



Automatic classification of clausal aspect

102 texts, 10355 clauses
3 annotators, κ=0.61

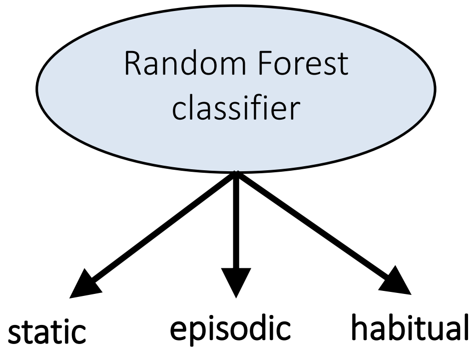
60% **static**
20% **episodic**
20% **habitual**



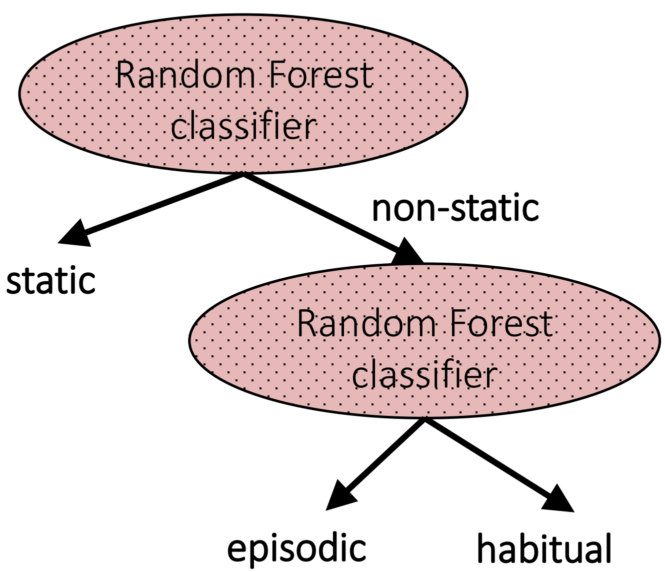
Both instance- and type-based features are needed!

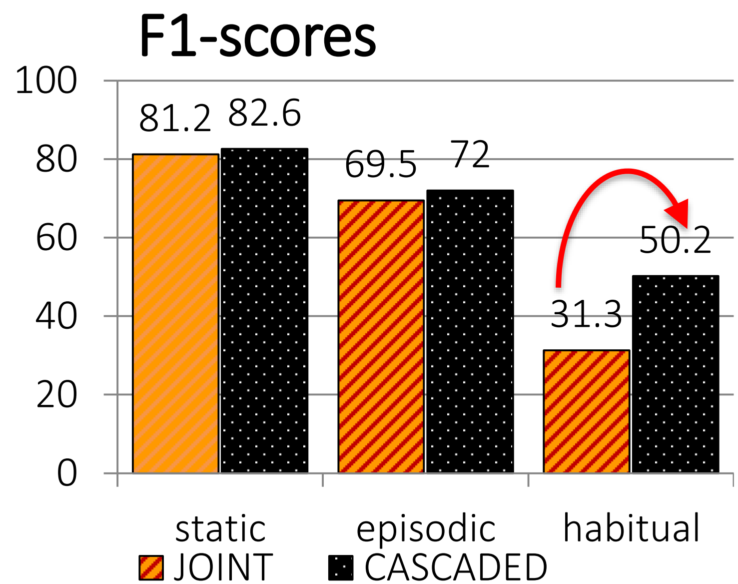
Automatic classification of clausal aspect

JOINT MODEL



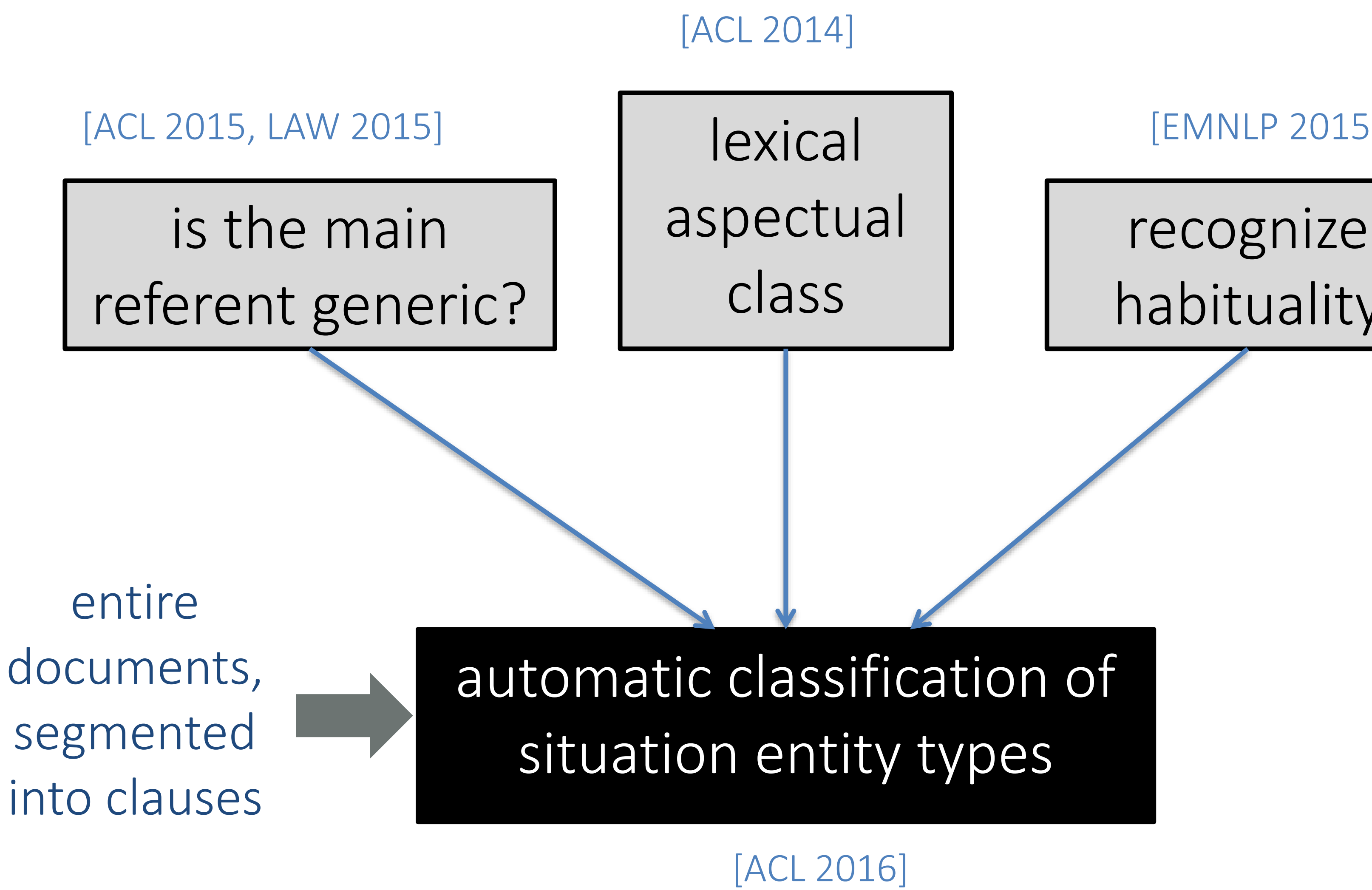
CASCADED MODEL



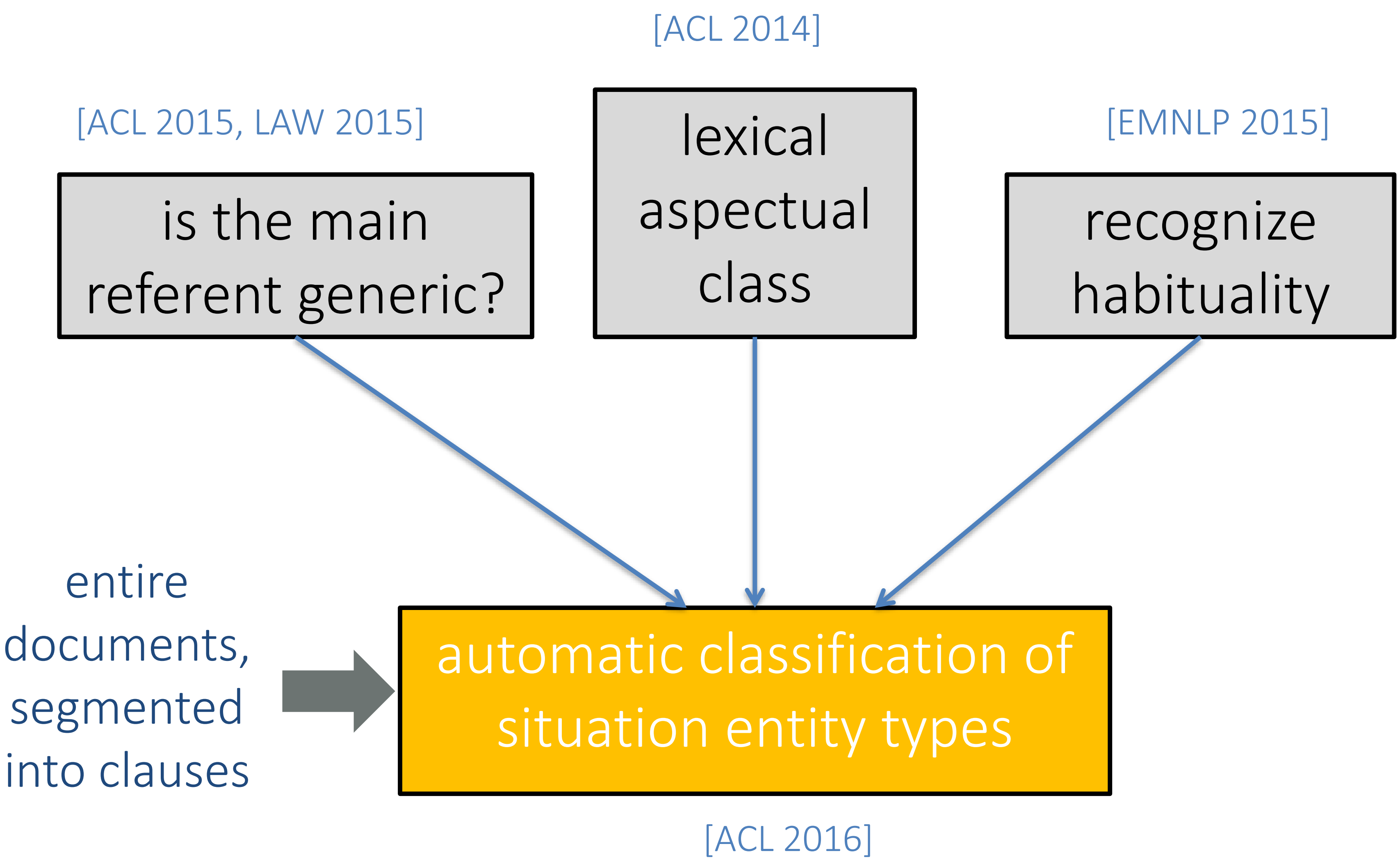


Cascaded model improves identification of habituals in free text.

Computational modeling of situation entity types



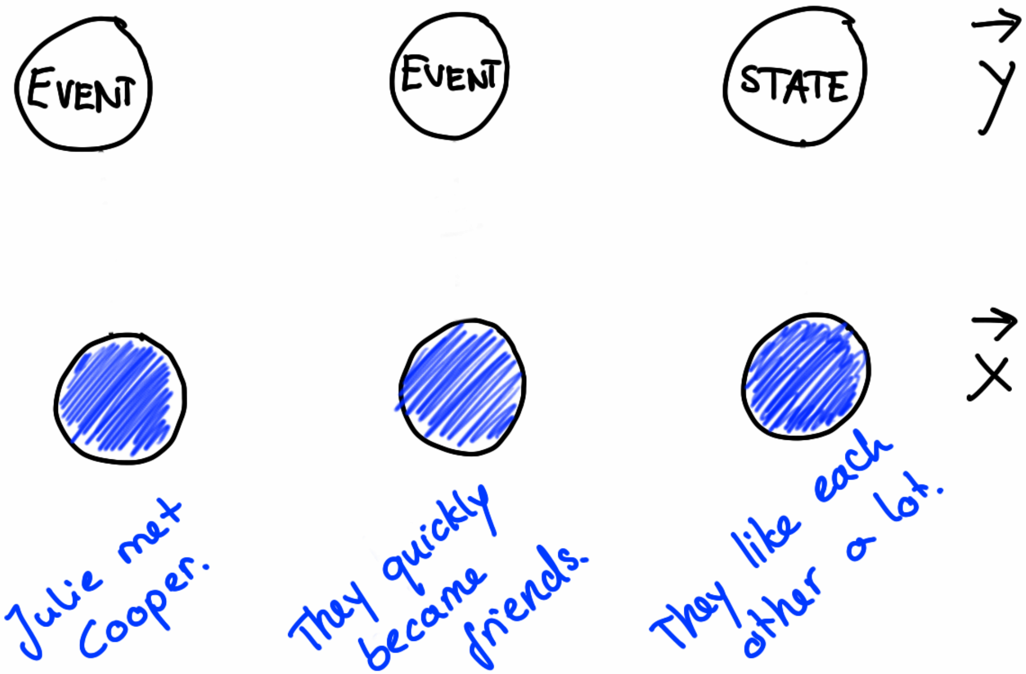
Computational modeling of situation entity types



Conditional random field (CRF)

- text document = sequence of clauses
- $\vec{y}$ = sequence of situation entity type labels
- $\vec{x}$ = features representing the clauses

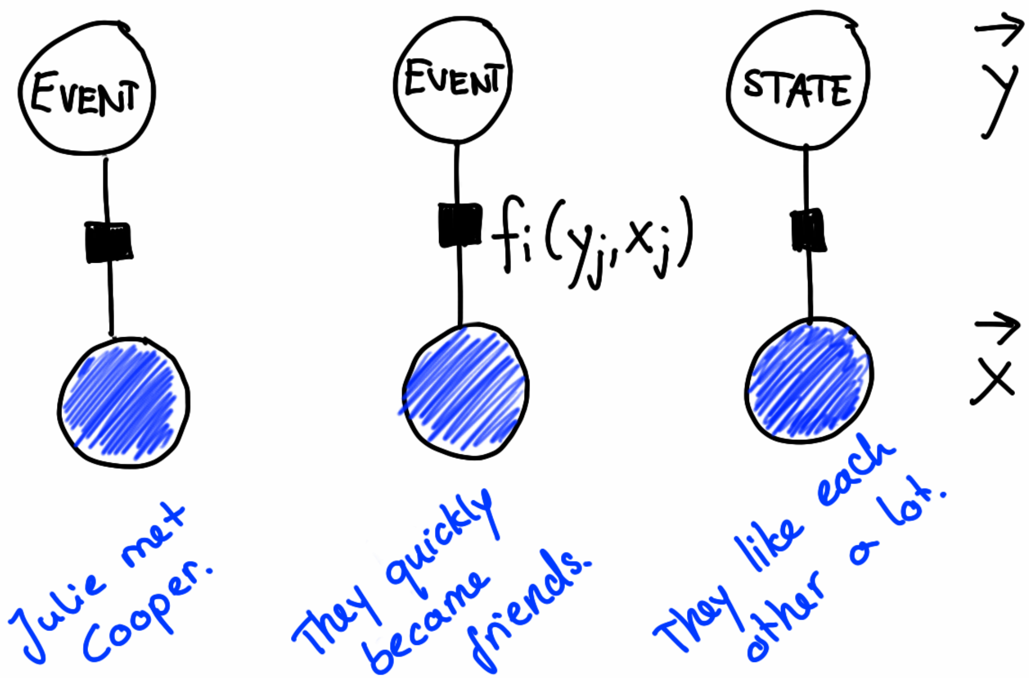
$$P(\vec{y}|\vec{x}) = \frac{1}{Z(\vec{x})} \exp\left(\sum_{j=1}^{n} \sum_{i=1}^{m} \lambda_i f_i(y_{j-1}, y_j, \vec{x}, j) \right)$$



Conditional random field (CRF)

- text document = sequence of clauses
- $\vec{y}$ = sequence of situation entity type labels
- $\vec{x}$ = features representing the clauses
- λ_i= weight for feature function f_i
- $f_i(y_j, x_j)$ = clause / type
  → MaxEnt

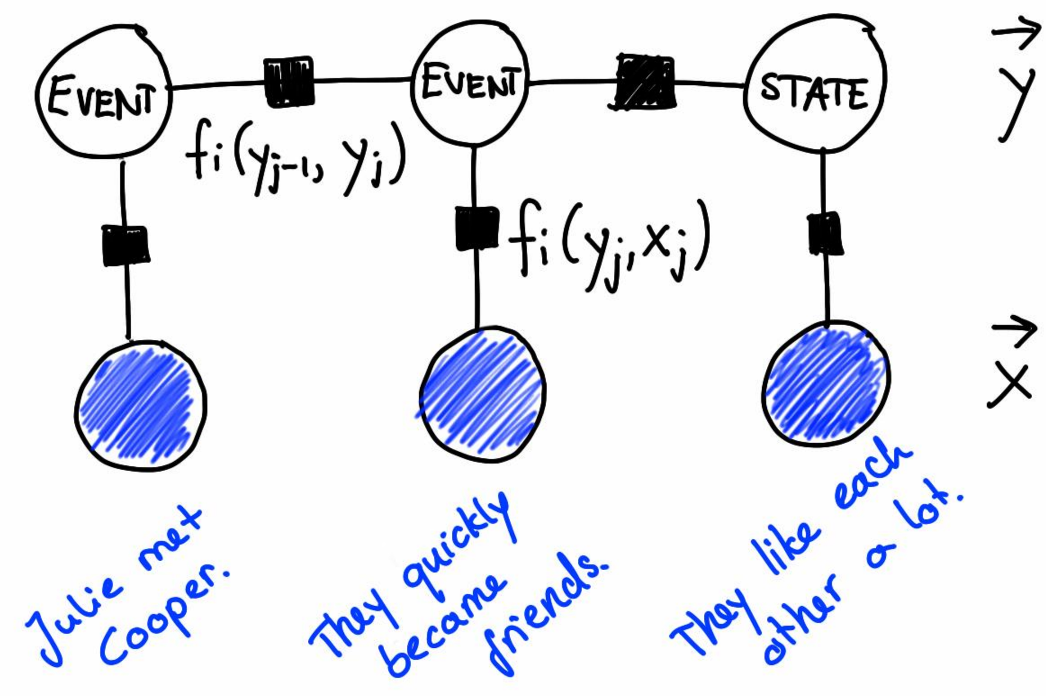
$$P(\vec{y}|\vec{x}) = \frac{1}{Z(\vec{x})} \exp\left(\sum_{j=1}^{n}\sum_{i=1}^{m} \lambda_i f_i(y_{j-1}, y_j, \vec{x}, j)\right)$$



Conditional random field (CRF)

- text document = sequence of clauses
- $\vec{y}$ = sequence of situation entity type labels
- $\vec{x}$ = features representing the clauses
- λ_i = weight for feature function f_i
- $f_i(y_j, x_j)$ = clause / type
  → MaxEnt
- $f_i(y_{j-1}, y_j)$ = type / type
  → CRF

$$P(\vec{y}|\vec{x}) = \frac{1}{Z(\vec{x})} \exp\left(\sum_{j=1}^{n}\sum_{i=1}^{m} \lambda_i f_i(y_{j-1}, y_j, \vec{x}, j)\right)$$



Situation entity types

- Which parts of the clause are most important to distinguish the types?
  [Friedrich & Palmer 2014b], [Friedrich et al. 2015], [Smith 2003]

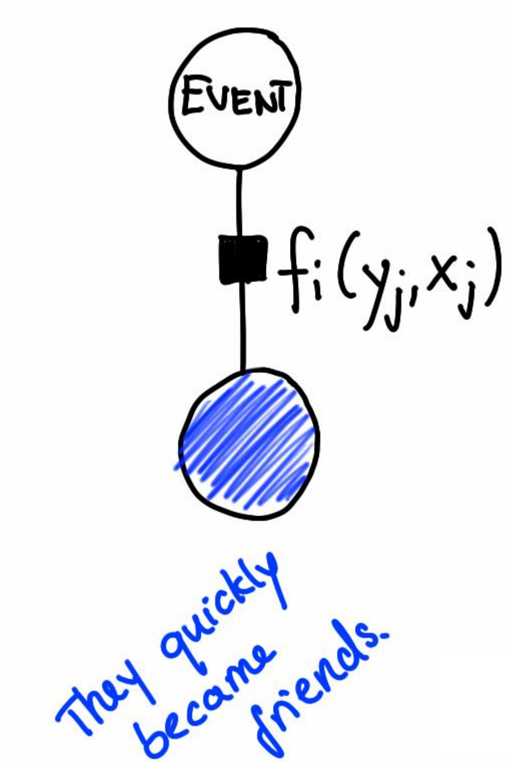
main verb → verb that heads the clause

Julie **likes** Cooper.	STATE
Julie **met** Cooper.	EVENT
Julie **teases** Cooper.	GENERALIZING SENTENCE

main referent → subject of main verb (what the clause is about)

Julie is an owl.	STATE
Owls are nocturnal animals.	GENERIC SENTENCE

Features for clauses



- **pos:** part of speech tags
- **bc:** Brown word clusters

  [Turian et al. 2010]

- **mv**: main verb
  - tense, voice, progressive, perfect, lemma, WordNet hypernyms, ...
- **mr**: main referent
  - lemma, determiner type, noun type, number, person, countability, WordNet, dependency relations, ...
- **cl**: *clause*
  - adverbs, conditional, modal, negated, ...

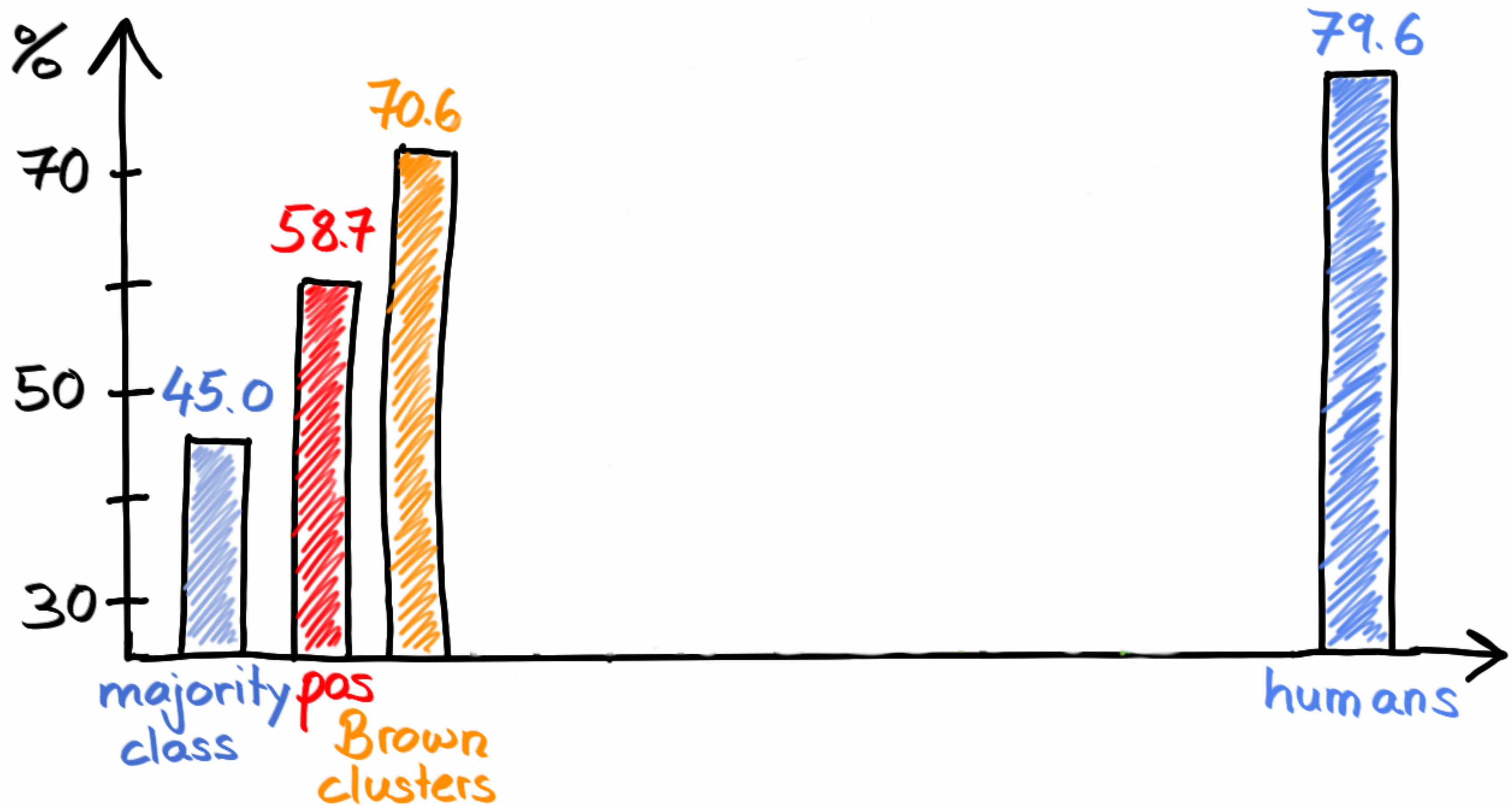
How well does it work?

Results: Impact of different feature sets

- Accuracy. Wiki+MASC dev set (80% of data), CRF, 10-fold CV.

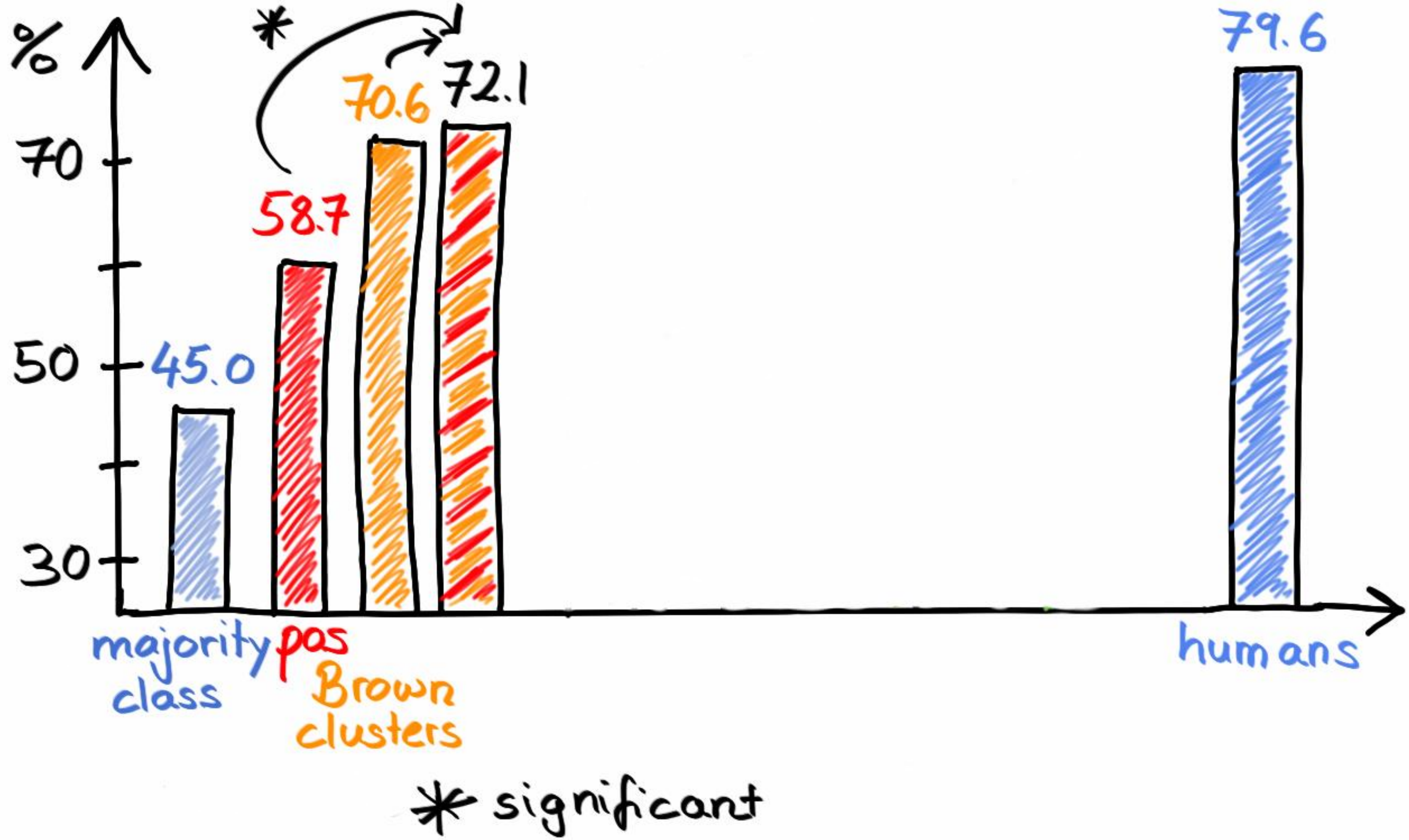
Results: Impact of different feature sets

- Accuracy. Wiki+MASC dev set (80% of data), CRF, 10-fold CV.



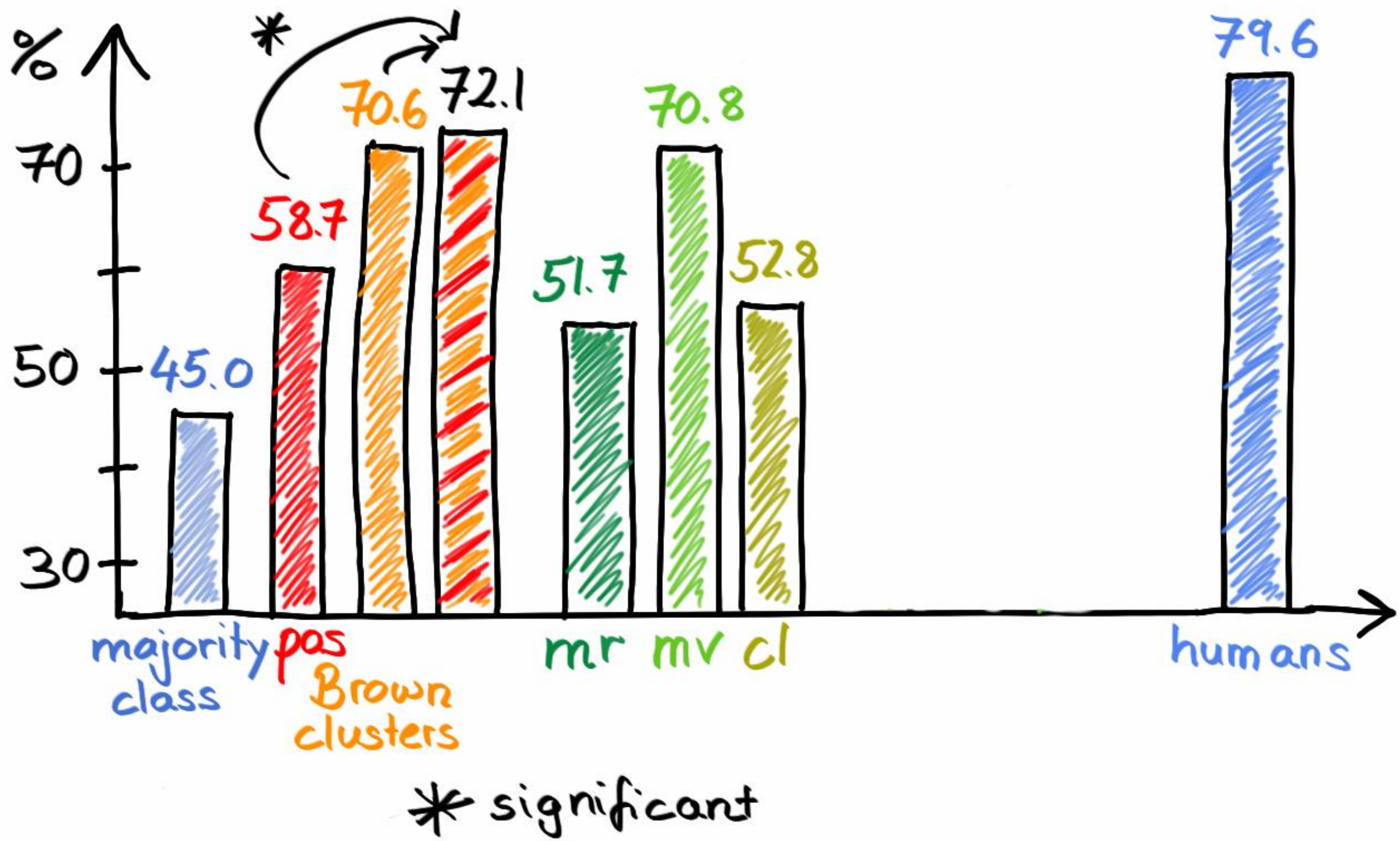
Results: Impact of different feature sets

- Accuracy. Wiki+MASC dev set (80% of data), CRF, 10-fold CV.



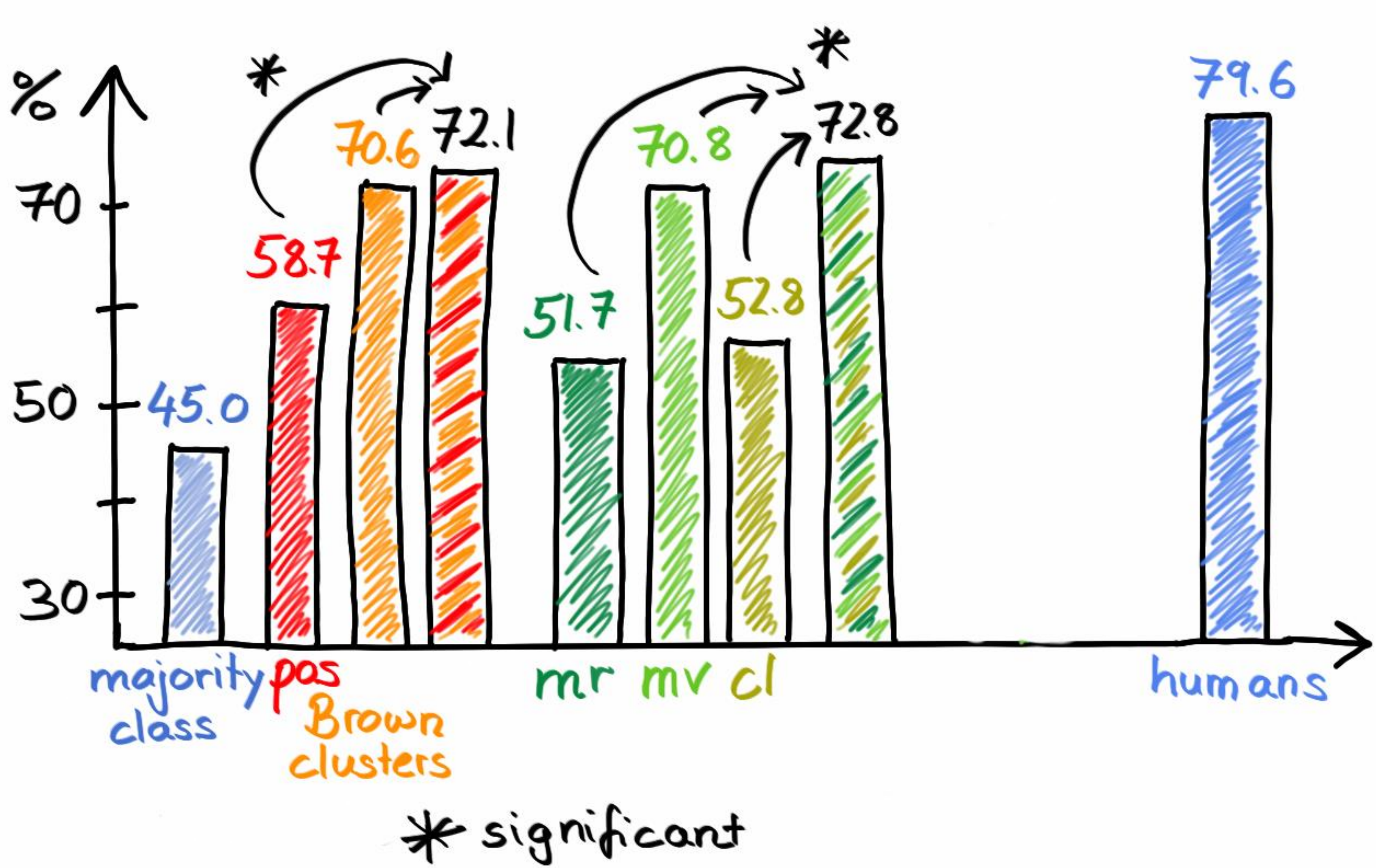
Results: Impact of different feature sets

- Accuracy. Wiki+MASC dev set (80% of data), CRF, 10-fold CV.



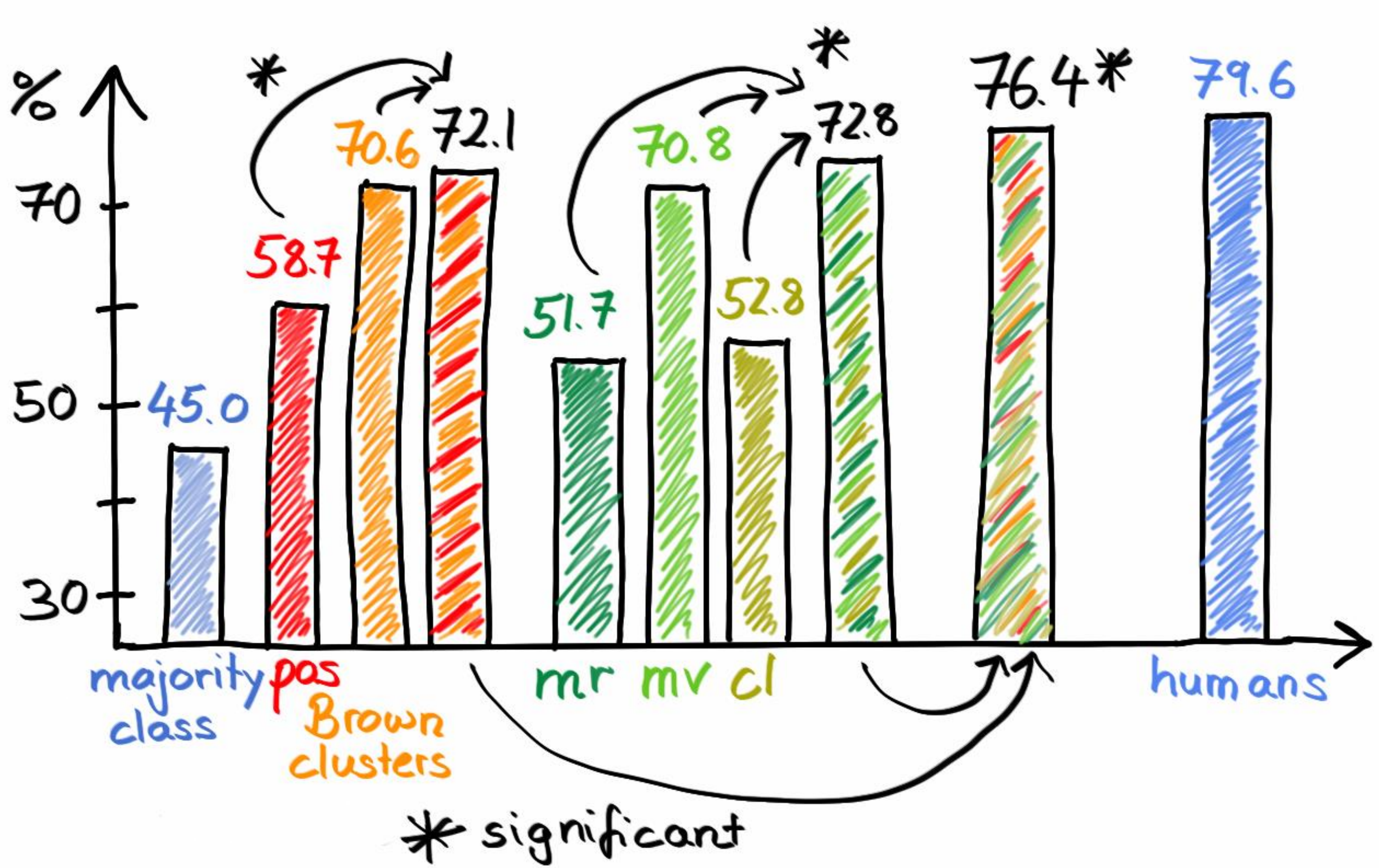
Results: Impact of different feature sets

- Accuracy. Wiki+MASC dev set (80% of data), CRF, 10-fold CV.



Results: Impact of different feature sets

- Accuracy. Wiki+MASC dev set (80% of data), CRF, 10-fold CV.



Results on heldout test set (20% of data)

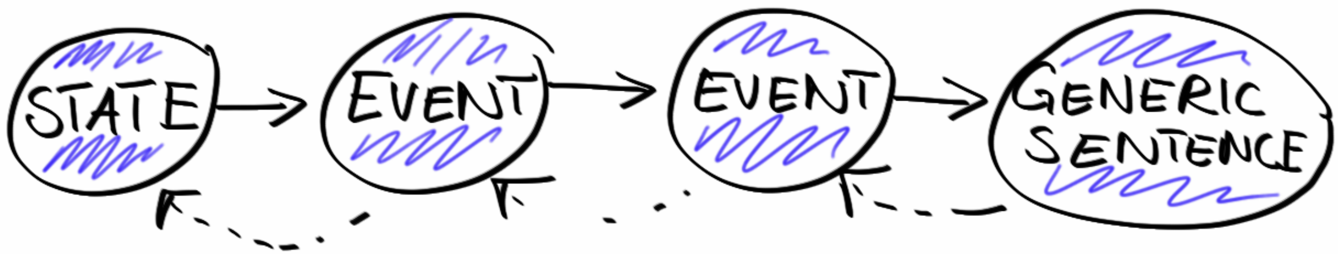
- Training on entire MASC+Wiki development set.

	macro-average			
feature set	**P**	**R**	**F**	**accuracy**
majority class (STATE)	6.4	14.3	8.8	44.7
pos+Brown	67.6	60.6	63.9	69.8
mr+mv+cl	69.9	61.7	65.5	71.4
all	**73.4**	**65.5**	**69.3**	**74.7**

Is sequential information important?

as claimed by Palmer et al. [ACL 2007]

... and if yes, when?



Maximum entropy model vs. conditional random field

situation entity type	MaxEnt	CRF
STATE	79.1	**80.6**
EVENT	77.5	**78.6**
REPORT	78.2	**78.9**
GENERIC SENTENCE	61.3	**68.3**
GENERALIZING SENTENCE	25.0	**29.4**
IMPERATIVE	72.3	**75.3**
QUESTION	84.4	**84.4**
macro-avg. F1	68.7	71.2
accuracy	74.1	76.4

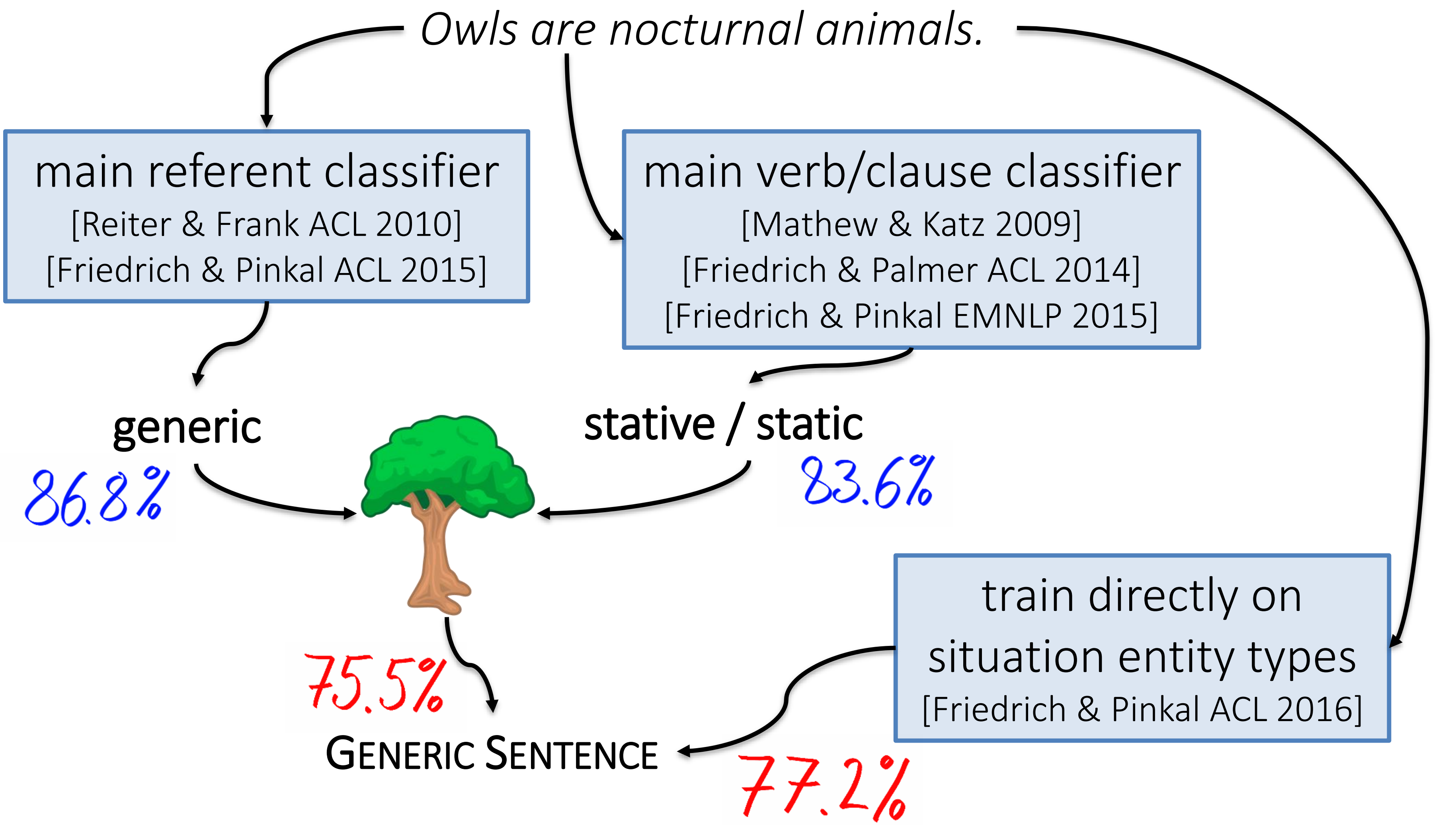
significant

How genre-dependent is this task?

- How important is it to have in-genre training data?
  - helpful, ≈ + 5% accuracy / F1
- Is it a good idea to add out-of-genre training data?
  - YES! 49.0 → 64.0 (macro-average F1)
  - system gets better at identifying infrequent types
- Statistics per type / genre: see [Friedrich, Palmer & Pinkal ACL 2016]

Pipelined model for situation entity types?

- State, Event, Generic Sentence, Generalizing Sentence



Lessons learned

- situation entity type classification task is (somewhat) difficult even for humans
- system performs well when comparing to human performance (76% vs. 80%)
- our system performs well across genres
- some types are infrequent in particular genres
  - adding out-of-domain training data helps to identify them
- a wide range of syntactic-semantic features are useful for this task
- sequential information useful for identifying „generic contexts“

Open questions

- integration of aspectual information into temporal relation identification systems?
- leveraging modeling of aspect for MT?
- semi-/unsupervised acquisition of aspectual information, e.g., from parallel corpora? **deep learning?**
- crowdsource relevant annotations?
- pre-processing step for argumentation mining, user-guided summarization etc.?