Resolving Shell Nouns

Ph.D. Dissertation Defence
Advisor: Graeme Hirst

by

Varada Kolhatkar

Department of Computer Science
University of Toronto
varada@cs.toronto.edu

Shell nouns

(Vendler 1968; Halliday and Hasan 1976; Ivanic 1991; Asher 1994; Francis 1994; Schmid 2000, inter alia)

> Shell noun resolution: Identifying shell content of a shell noun phrase in the given context

shell content

The municipal council will have to decide whether to balance the budget by raising revenue or cutting spending. The council will have to come to a resolution by the end of the month. This issue is dividing communities across the country.

shell noun phrase

Goal of the research

Treatment of shell nouns from a computational linguistics perspective

- Research questions

  - To what extent are speakers of English able to interpret shell nouns?

  - How can we develop a computational system to resolve shell nouns?

  - To what extent can the knowledge derived from the linguistics literature help in this process?

Examples of shell nouns

Schmid provides a list of 670 shell nouns

fact	issue	problem	principle	decision
thing	concept	reason	notion	phenomenon
idea	rumour	legend	message	possibility
belief	plan	truth	theory	thought
order	trend	argument	proposal	certainty

Why do we care?

- Ubiquity of shell nouns

  - *fact*, *idea*, *problem*: among 100 most frequently occurring nouns in the BNC (Schmid 2000)

- Functions in discourse

  - **Characterize** and **label** information in the context

- Potential applications

  - Discourse understanding, text summarization, non-factoid question answering, ESL learning (Francis 1988; Flowerdew 2003; Hinkel 2004)

State of the art in CL

● Fair amount of attention in Linguistics
(Vendler 1968; Halliday and Hasan 1976; Ivanic 1991; Asher 1994; Francis 1994; Schmid 2000, inter alia)

... but largely ignored in CL

Current challenge in anaphora resolution:
Going beyond nominal anaphora
(Byron 2004; Poesio 2011)
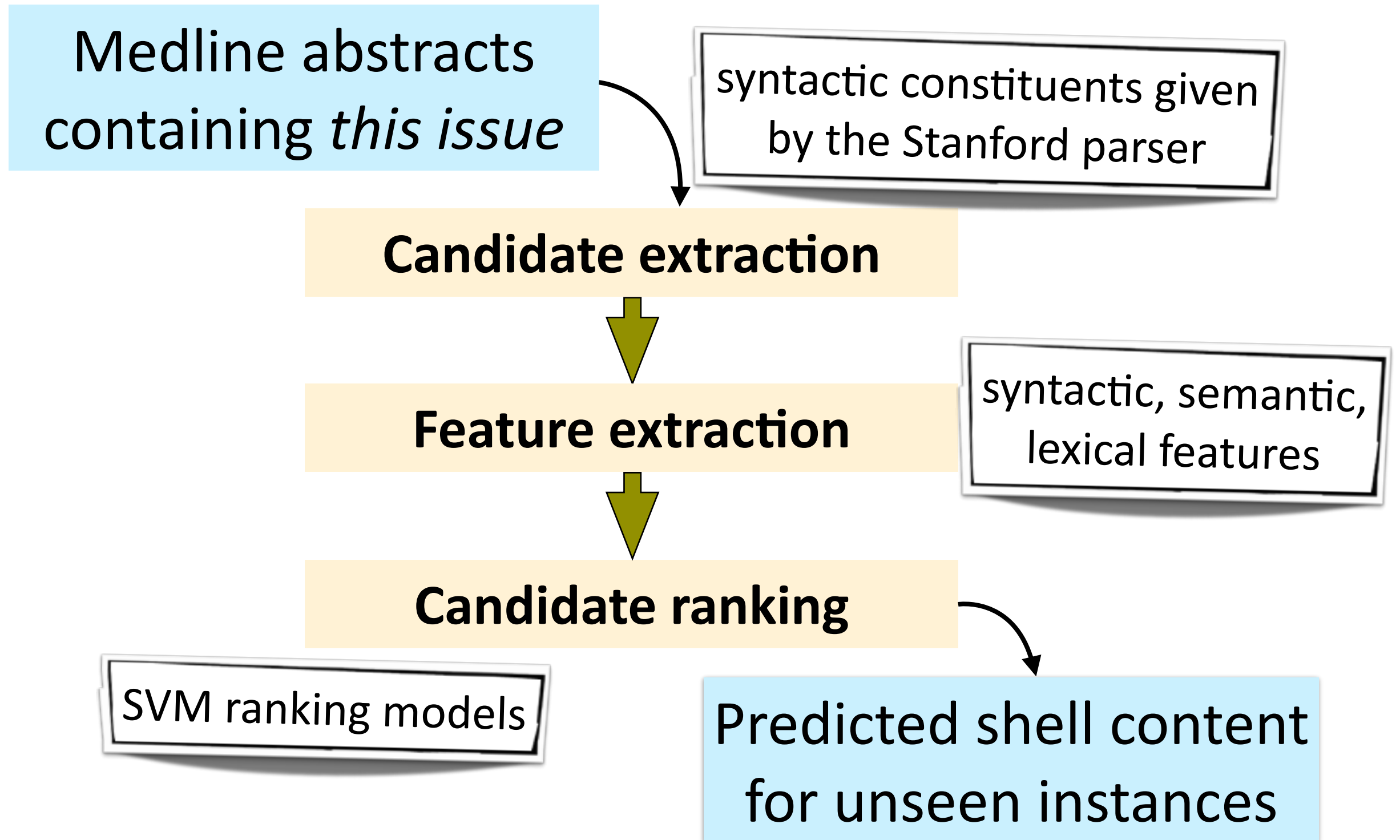
Pilot study

Chapter 3
(Kolhatkar and Hirst 2012)

Focus

Resolve *this issue* instances in the Medline domain

- Data

  - 183 *this issue* instances

  - Agreement: **0.86** (Krippendorff 2013)

Resolution algorithm

Medline abstracts containing *this issue*

syntactic constituents given by the Stanford parser

Candidate extraction

Feature extraction

syntactic, semantic, lexical features

Candidate ranking

SVM ranking models

Predicted shell content for unseen instances

Pilot study: summary

Feasible to reliably annotate and resolve the shell noun phrase *this issue* in the Medline domain.

Random
Adjacent sentence
SVM ranker

Exact shell content matching
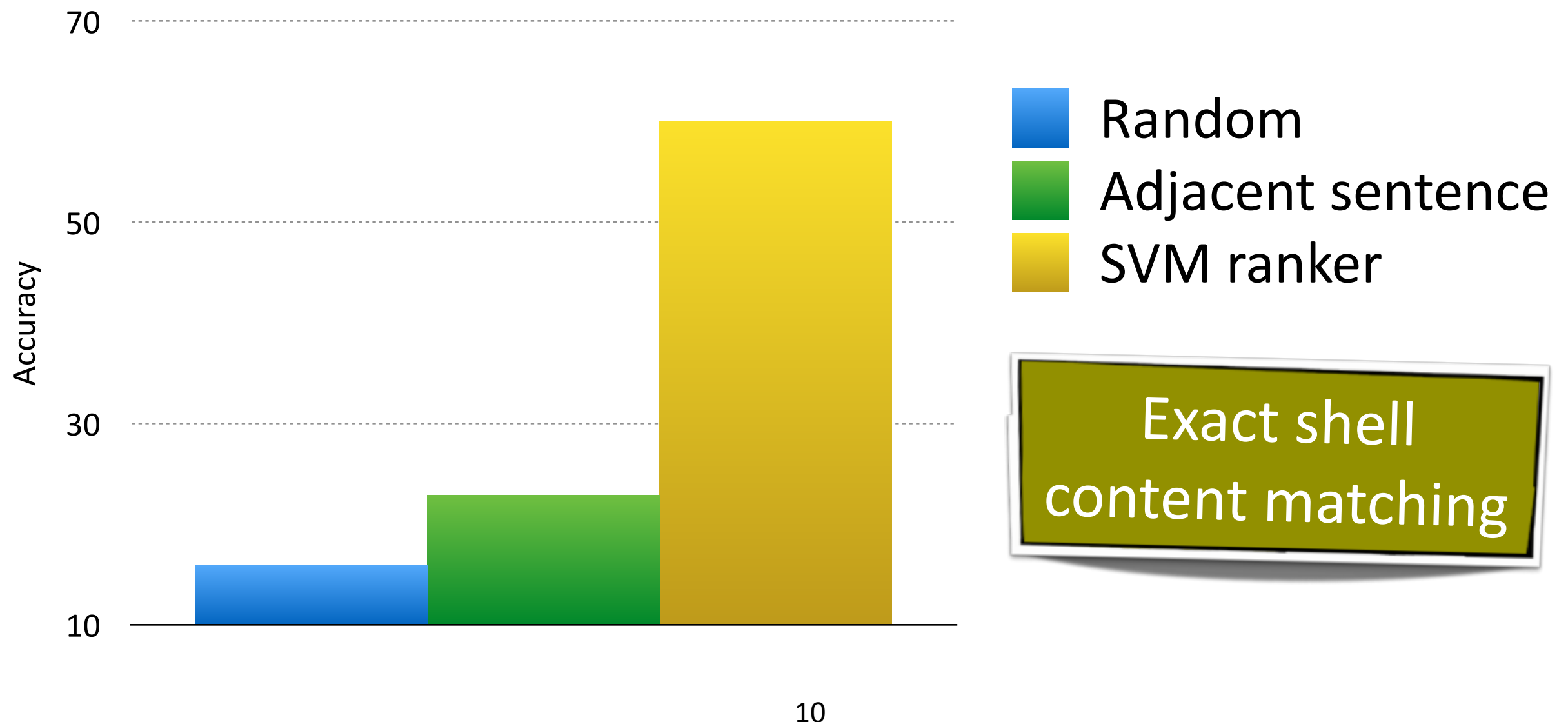
Generalizing to other shell nouns

- Goal
  - A variety of shell nouns
  - Broader domain
- Primary challenges
  - Idiosyncrasies
  - No annotated data
  - A variety of constructions

Different types of usages

The municipal council will have to decide whether to balance the budget by raising revenue or cutting spending. The council will have to come to a resolution by the end of the month. This issue is ...

The issue that this country and Congress must address is how to provide optimal care for all without limiting access for the many.

A bad idea does not harm until someone acts upon it.

Mathis is the cover subject of this week's issue of Sports Illustrated.

Different types of usages

✓ The municipal council will have to decide whether to balance the budget by raising revenue or cutting spending. The council will have to come to a resolution by the end of the month. This issue is ...

The issue that this country and Congress must address is how to provide optimal care for all without limiting access for the many.

A bad idea does not harm until someone acts upon it.

Mathis is the cover subject of this week's issue of Sports Illustrated.

Different types of usages

anaphoric (ASN)

✓ The municipal council will have to decide whether to balance the budget by raising revenue or cutting spending. The council will have to come to a resolution by the end of the month. This issue is …

cataphoric (CSN)

✓ The issue that this country and Congress must address is how to provide optimal care for all without limiting access for the many.

A bad idea does not harm until someone acts upon it.

Mathis is the cover subject of this week's issue of Sports Illustrated.

Different types of usages

anaphoric (ASN)

✓ The municipal council will have to decide whether to balance the budget by raising revenue or cutting spending. The council will have to come to a resolution by the end of the month. This issue is …

cataphoric (CSN)

✓ The issue that this country and Congress must address is how to provide optimal care for all without limiting access for the many.

indefinite shell content

✗ A bad idea does not harm until someone acts upon it.

Mathis is the cover subject of this week's issue of Sports Illustrated.

Different types of usages

anaphoric (ASN)

✓ The municipal council will have to decide whether to balance the budget by raising revenue or cutting spending. The council will have to come to a resolution by the end of the month. This issue is …

cataphoric (CSN)

✓ The issue that this country and Congress must address is how to provide optimal care for all without limiting access for the many.

indefinite shell content

✗ A bad idea does not harm until someone acts upon it.

✗ Mathis is the cover subject of this week's issue of Sports Illustrated.

non-shell noun usage

Resolving Cataphoric Shell Nouns (CSNs)

Chapter 4
(Kolhatkar and Hirst 2014)

CSN patterns

N-be-clause

N-clause

Pattern	Example
N-be-to	**Our plan is *to*** *hire and retain the best managers we can*.
N-be-that	**The major reason is *that*** *doctors are uncomfortable with uncertainty*.
N-be-wh	Of course, **the central issue is *whether*** *animal testing is cruel*.
N-to	**The decision *to*** *disconnect the ventilator* came after doctors found no brain activity.
N-that	These challenges do not undermine **the fact *that*** *museums are on a high*.
N-wh	If there ever is **any doubt whether** *a plant is a poppy or not*, break off a stem and squeeze it.
N-of	**The concept *of*** *having an outsider as Prime Minister* is outdated.

CSNs: a semantic phenomenon

CSN

effect

One reason that 60 percent of New York City public-school children read below grade level is that many elementary schools don't have libraries.

cause = shell content

- Identify that *reason* expects two arguments: cause and effect

- Identify that the shell content is given in the cause argument

- Identify the syntactic constituent representing cause

Shell noun families

- Where can we find this kind of semantic knowledge?

  Answer:
  Schmid's semantic families

- Schmid groups together different usages of 670 shell nouns into 79 semantic families

Shell noun families

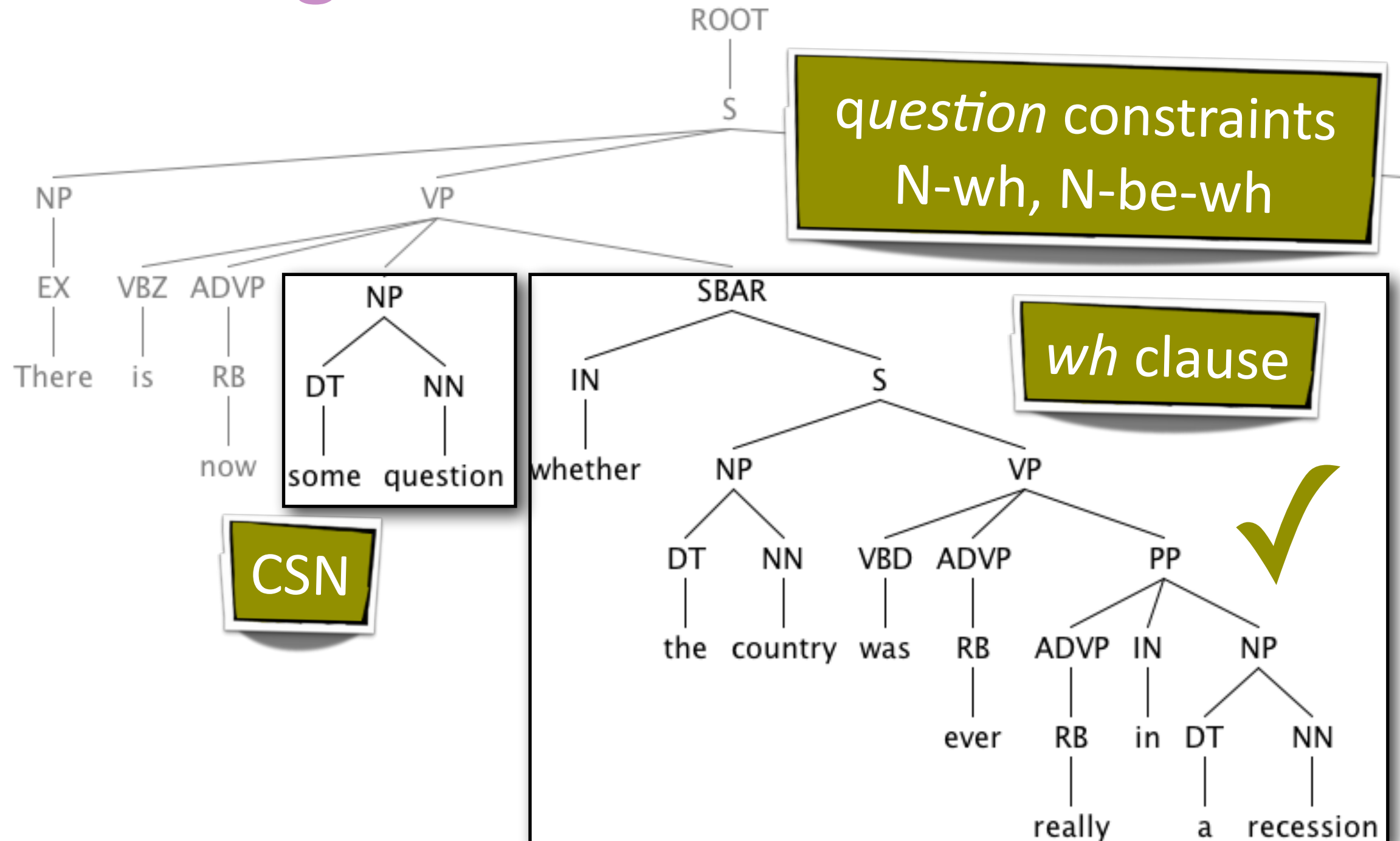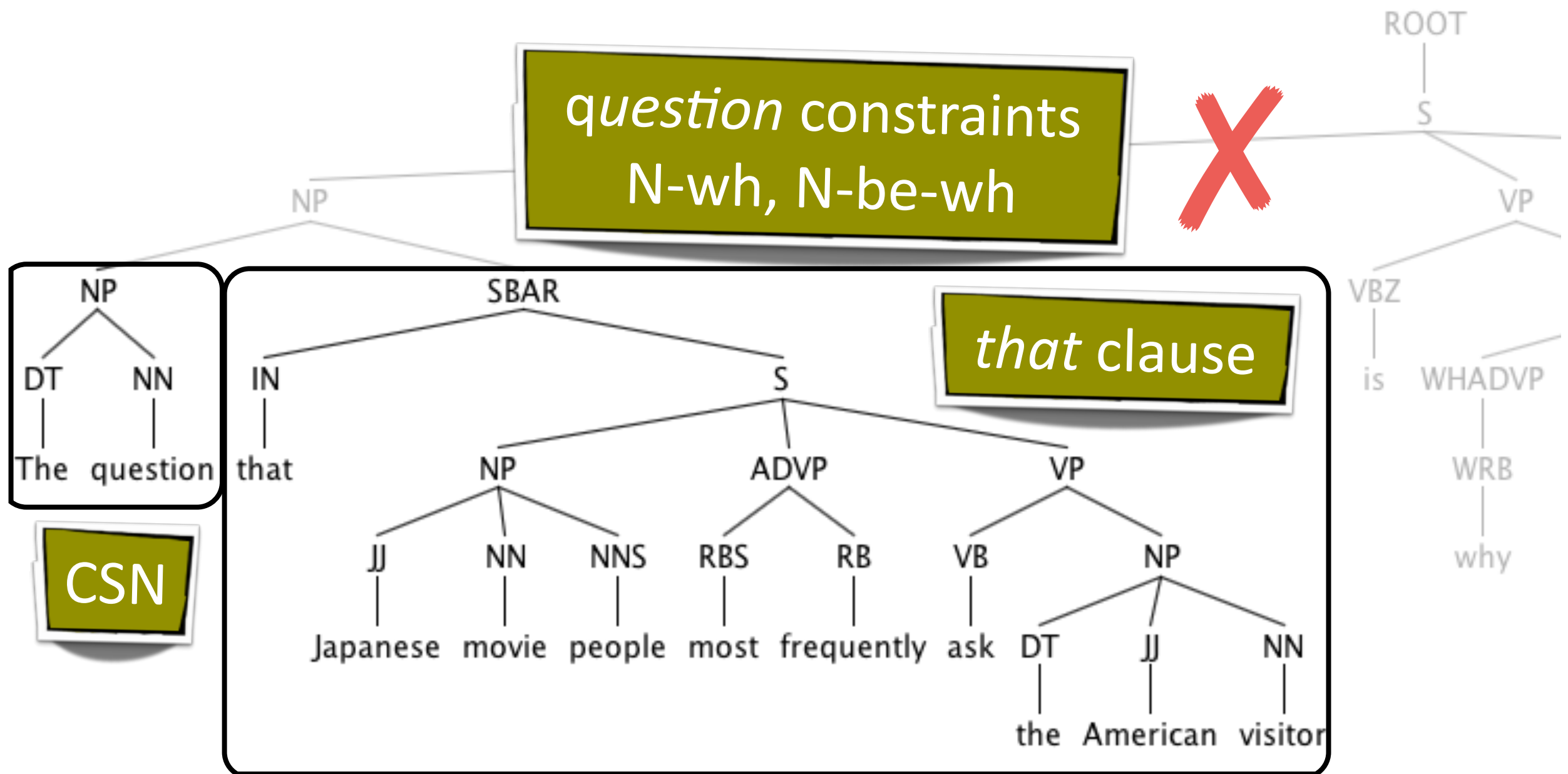| _Idea_	
Shared semantic features	[MENTAL], [CONCEPTUAL]
frame	Mental; focus on propositional content of IDEA
Nouns	_Point, idea, position, issue, theory, notion, thought, principle, rule, subject, image, myth, law, theme, concept, secret, scenario, wisdom, hypothesis, thesis,..._
Pattern	_N-be-that, N-that_

Using families for CSN resolution

There is now some **question** whether the country was ever really in a recession.

Using families for CSN resolution

The **question** that Japanese movie people most frequently ask the American visitor is why …

Evaluation data

Idea family

Semantic features: [mental], [conceptual]

Frame: men~~~~~~~~~ of IDEA

idea, issue, concept

Nouns: *idea*~~~~~~~~ *ory*, ...

Patterns: *N-be-that/of*, *N-that/of*

Trouble family

Semantic features: [eventive], [attitudinal], [manner], [deontic]

Frame: general eventive

problem, trouble, difficulty

Nouns: *problem*, *trouble*, *difficulty*,

Patterns: *N-be-to*

Thing family

Semantic features: [factual]

Frame: general factual

fact, phenomenon

Nouns: *fact*~~~~~~~~ *business*

Patterns: *N-that, N-be-that*

Plan family

Semantic features: [mental], [volitional], [manner]

Frame:

decision, plan, policy

Nouns:

Patterns: *N-be-to/that*, *N-to/that*

Problem family

Semantic features: [factual], [attitudinal], [impeding]

Frame: general factual

rouble, *difficulty*, *point*, *thing*, *snag*,

Patterns: *N-be-that/of*

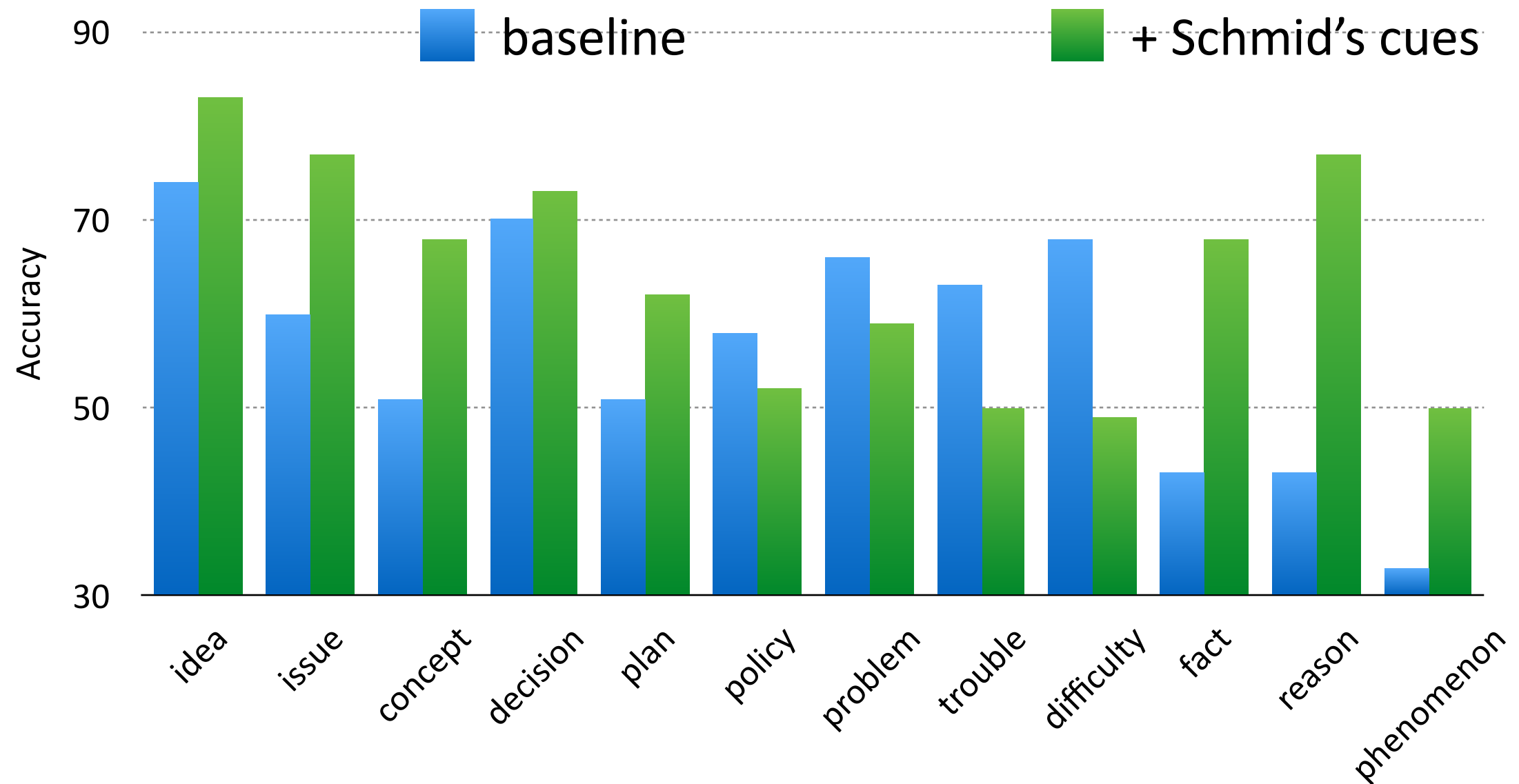## *Reason* family

Semantic features: [factual], [causal]

Frame: causal~~~~~~~focus on CAUSE

Nouns: *reason*~~~~~ reason ~~~~~ *thing*

Patterns: *N-be-that/why, N-that/why*

20

CSN resolution results

Semantic knowledge from Schmid's families is helpful in resolving CSNs

Legend: ■ baseline ■ + Schmid's cues

Y-axis: Accuracy (90, 70, 50, 30)

X-axis categories: idea, issue, concept, decision, plan, policy, problem, trouble, difficulty, fact, reason, phenomenon
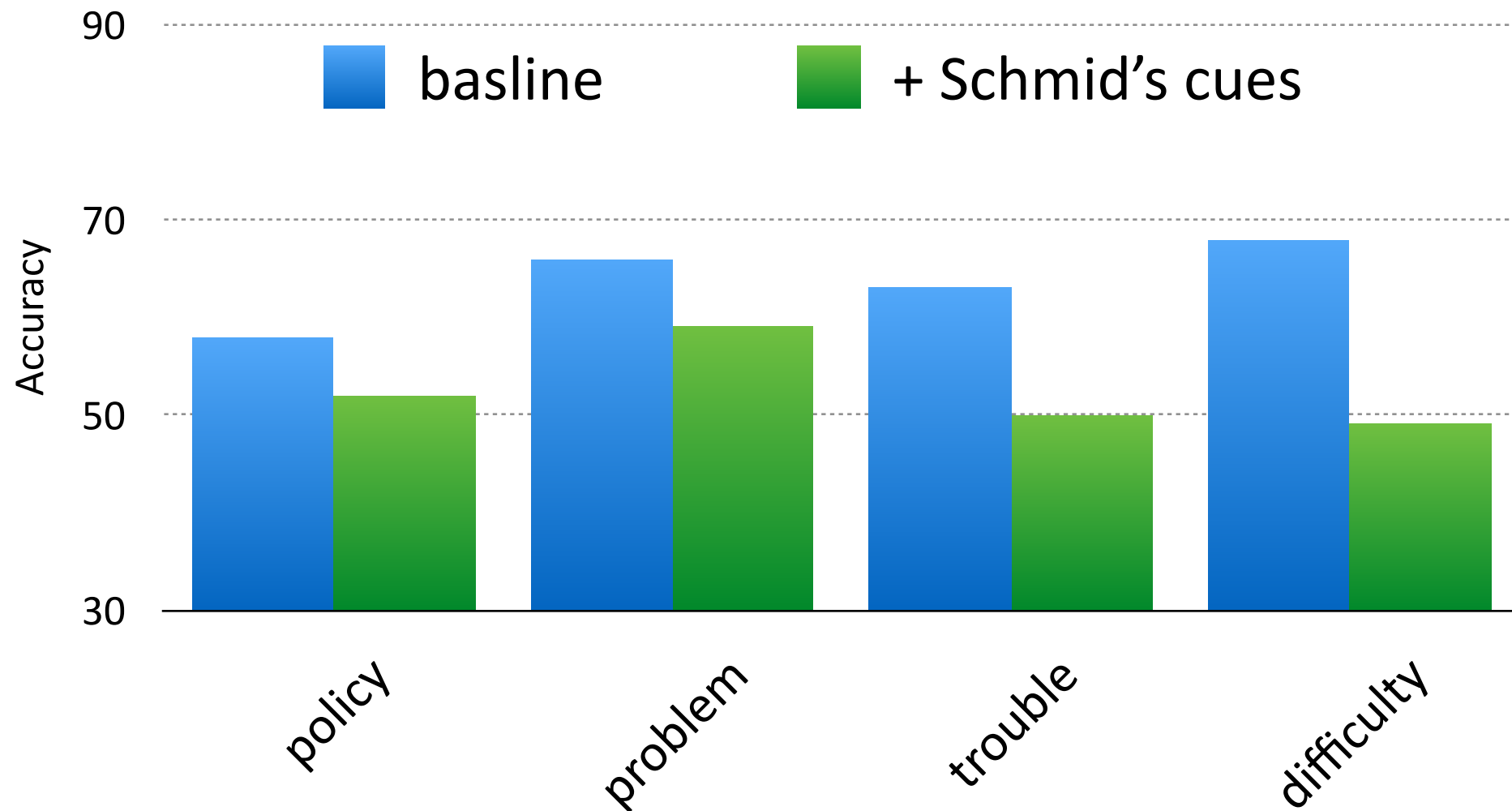
CSN resolution results

Schmid's framework particularly helps in resolving nouns with strict expectations
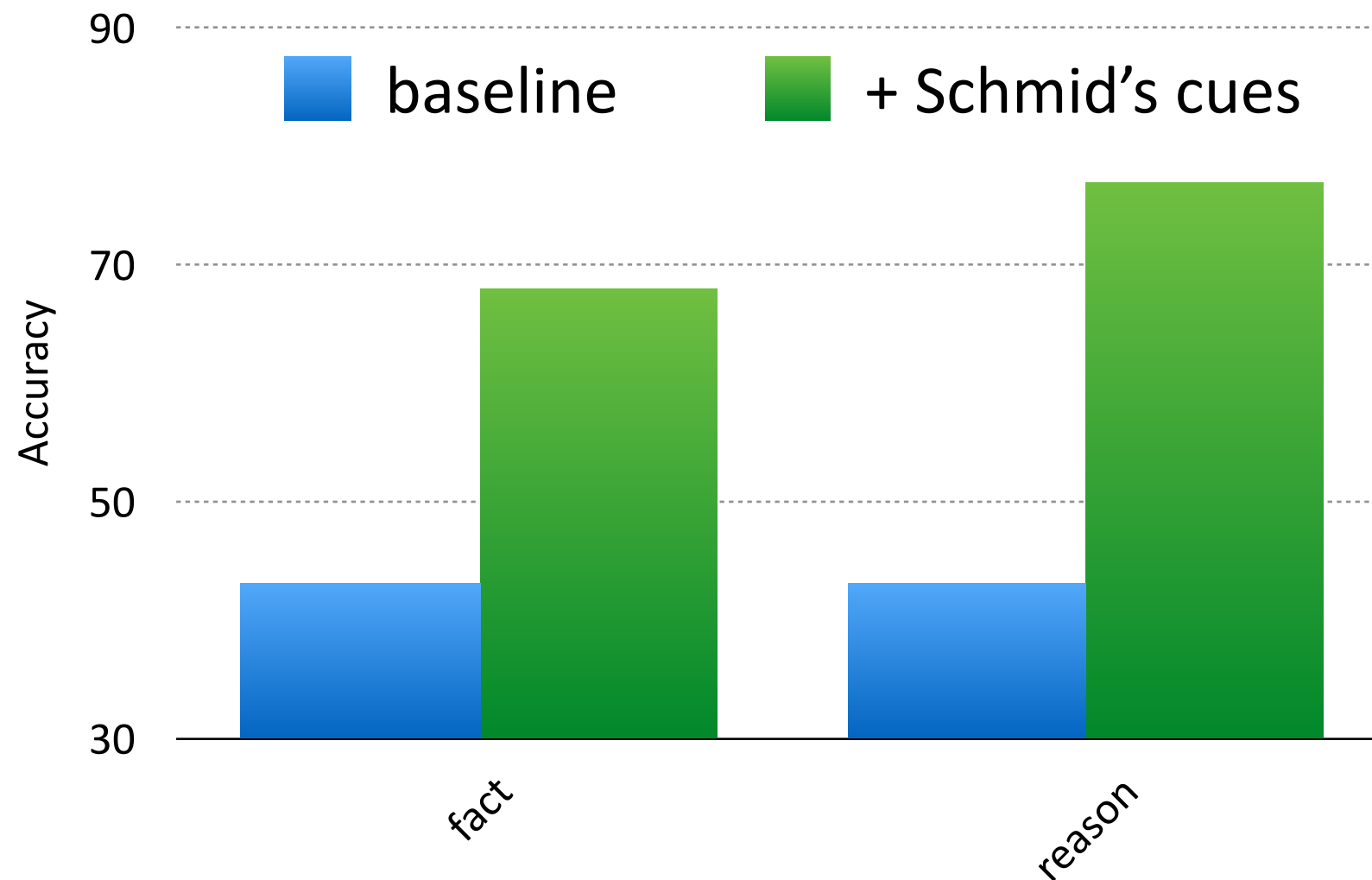
CSN resolution results

Schmid's cues were deleterious for more flexible nouns

Resolving Anaphoric Shell Nouns

Chapter 5
(Kolhatkar et al. 2013a, 2013b)

ASNs and CSNs similarities

ASN example

The municipal council had to decide whether to balance the budget by raising revenue or cutting spending. The council had to come to a resolution by the end of the month. This issue was dividing communities across the country.
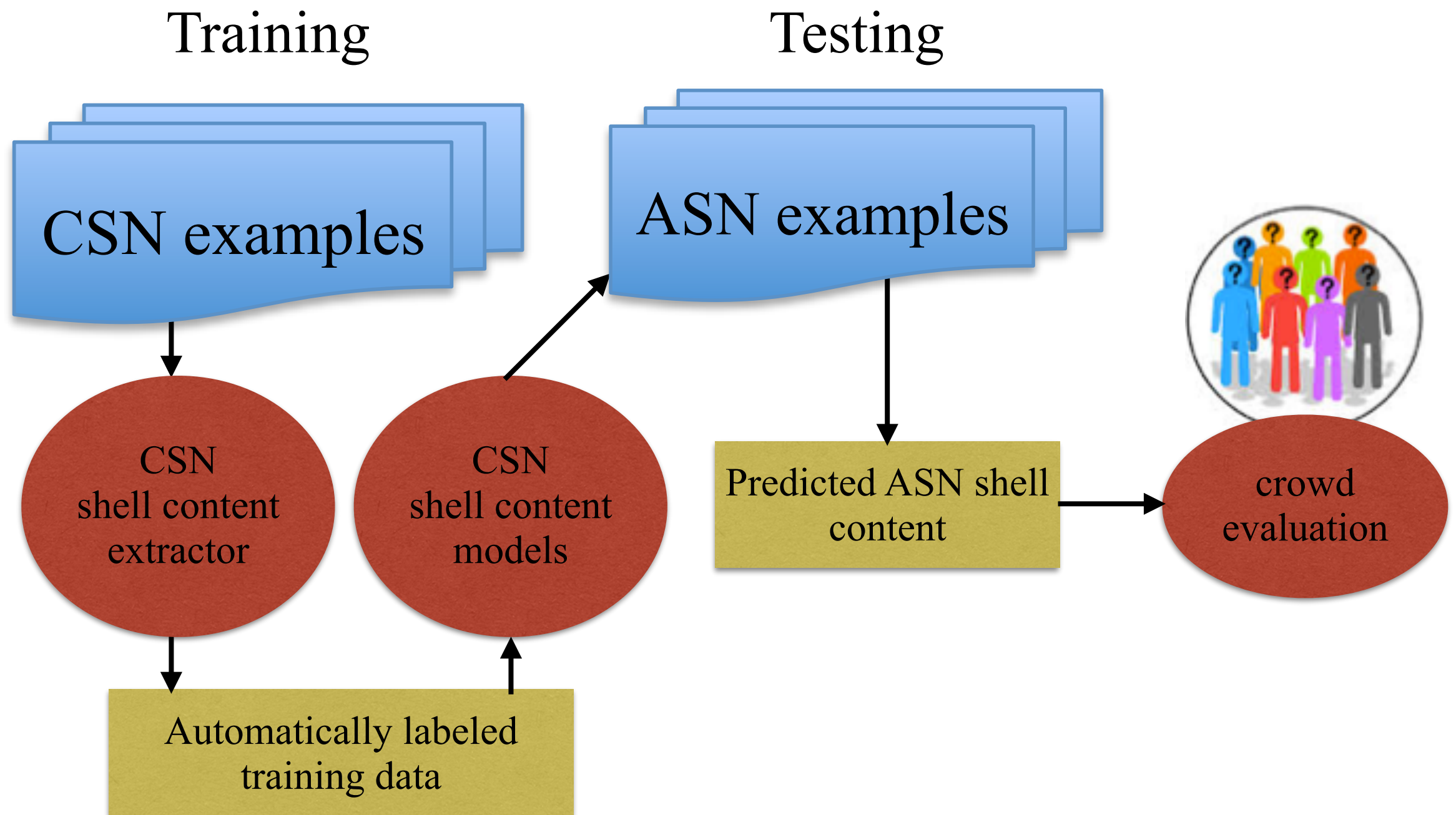
whether clause in both cases

CSN example

Of course, the central, and probably insoluble, issue is whether animal testing is cruel.

Hypothesis

CSN shell content and ASN shell content share some linguistic properties, and hence linguistic knowledge encoded in CSN shell content will help in interpreting ASNs.

Overview

Annotating ASNs to their shell content (Kolhatkar et al. 2013a)

The ASN corpus

The New York Times

- Base corpus: The NYT corpus
  (Sandhaus 2008)

- ~475 instances per 6 selected shell nouns
  fact, reason, issue, decision, question, possibility

- Total: 2,323 ASN instances

Annotation tasks

ASN instances from the NYT

CrowdFlower Expt. 1

Identify the sentence containing shell content

CrowdFlower Expt. 2

Identify the precise shell content

Annotated ASN Corpus

Crowdsourcing does best with simple tasks
(Madnani et al. 2010; Wang et al. 2012)

CrowdFlower expt. 1

Identifying the sentence containing the shell content

(a3) New York is one of only three states that do not allow some form of audio-visual coverage of court proceedings. **(a2)** Some lawmakers worry that cameras might compromise the rights of the litigants. **(a1)** But a 10-year experiment with courtroom cameras showed that televised access enhanced public understanding of the judicial system without harming the legal process. **(b)** New York's backwardness on this issue hurts public confidence in the judiciary...

CrowdFlower expt. 2

Identifying the precise shell content

New York is one of only three states that do not allow some form of audio-visual coverage of court proceedings. Some lawmakers worry that cameras might compromise the rights of the litigants. But a 10-year experiment with courtroom cameras showed that televised access enhanced public understanding of the judicial system without harming the legal process. New York's backwardness on this issue hurts public confidence in the judiciary...

CrowdFlower expt. 2

Identifying the precise shell content

New York is one of only three states that do not allow some form of audio-visual coverage of court proceedings.

Select one of the options (required)

- ○ None
- ○ one of only three states
- ○ some form of audio-visual coverage
- ○ some form of audio-visual coverage of court proceedings
- ○ that do not allow some form of audio-visual coverage of court proceedings
- ○ audio-visual coverage of court proceedings
- ○ some form
- ○ New York is one of only three states that do not allow some form of audio-visual coverage of court proceedings.
- ○ only three states that do not allow some form of audio-visual coverage of court proceedings
- ○ one of only three states that do not allow some form of audio-visual coverage of court proceedings
- ○ allow some form of audio-visual coverage of court proceedings

10 top-ranked candidates given by Kolhatkar et al. 2013b

How far can we get using CSN models?

Metric and baseline

- Success at n (S@n)

  - Proportion of instances where the crowd's answers occur within our ranker's first n choices

  - S@1 is standard precision

- Baseline

  - Consider the crowd-annotated sentence as the correct shell content

Ranker evaluation
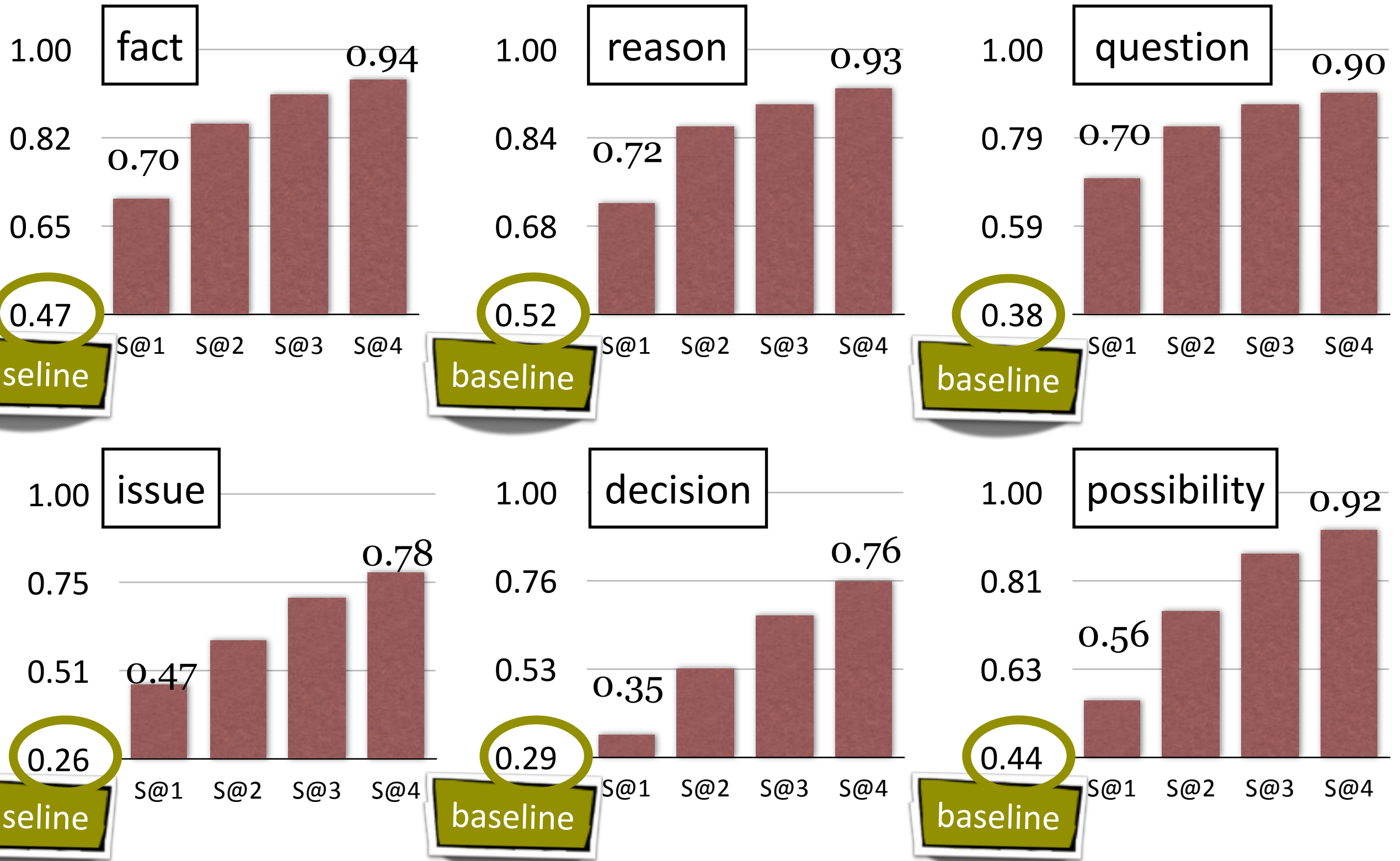
Summary of contributions

- First work that sheds light on shell nouns from a computational linguistic perspective

- First step towards resolving abstract anaphora

- Three resolution systems and four reliably-annotated corpora to further pursue this line of research

Future directions

- Short-term future directions

  - Shell noun resolution for other languages
    (with Heike Zinsmeister and Stefanie Dipper)

  - One SVM ranker for all shell nouns
    (with Alexander Schwing)

- Long-term future directions

  - Clustering shell nouns with similar semantic
    expectations similar to verb clustering

  - Identifying shell chains

End-to-end shell noun resolution

Resolving six CSNs with rules
(Kolhatkar et al. 2013a)

Generalizing CSN resolution
(Kolhatkar and Hirst 2014)

Resolving the same six ASNs using CSN shell content as training data
(Kolhatkar et al. 2013a)

Annotating ASNs
(Kolhatkar et al. 2013b)

Schmid's definition of shell nouns

- Characterization

- Concept-formation
  - Allow speakers to encapsulate the complex chunks of information in temporary nominal concepts with clear-cut conceptual boundaries

- Linking
  - Interpret two groups of linguistic elements together, as being related to and even dependent on each other

Schmid's definition of shell nouns

Shell nouns
e.g., *idea*, *fact*, *problem*

⟷

Full content nouns
e.g., *teacher*, *cat*,
journey

Pronouns with
anaphoric function
e.g., *she*, *it*, *this*, *that*

stable and rich denotation

limited potential for characterization

relatively constant relationship to the
experience they encapsulate as a concept

no concept-forming effects

suited for exophoric reference

create links of referential
identity or co-reference

Schmid's definition of shell nouns

- Combine the three functions of characterization, concept-formation and linking, which are otherwise performed separately, each by different types of linguistic elements

- They perform these functions in a fine-tuned balance between conceptual stability and informational flexibility

Schmid's definition of shell nouns
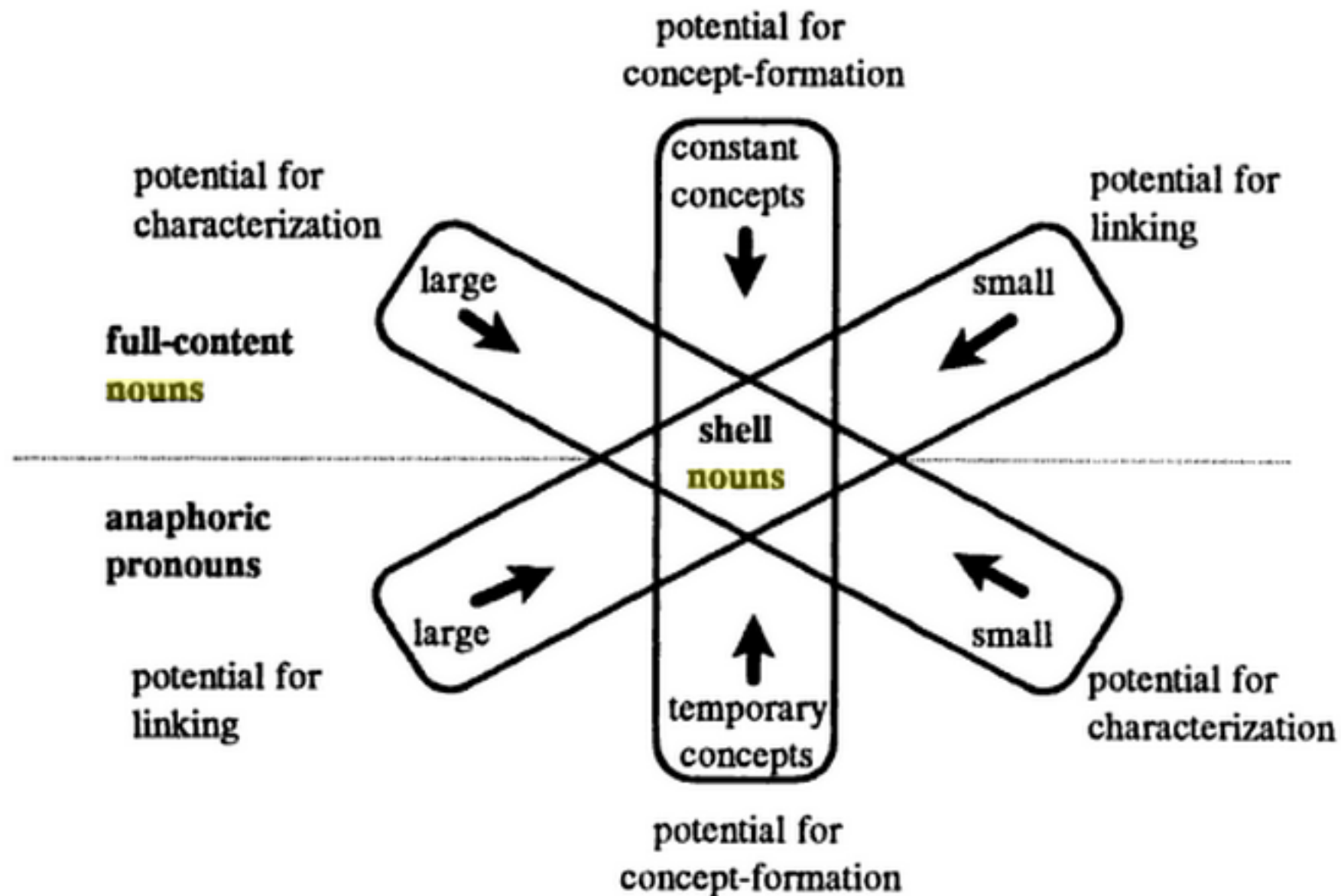
Ideas regarding proper evaluation of shell nouns

- Evaluation metric based on the success of the system at identifying the type of the objects that are really referred to instead of the words that evoke those objects

  - This problem is challenging for anaphoric demonstratives, as the type of the referent has to be identified from the predicative context

  - For shell nouns, the type of the referent is encoded in the shell nouns themselves

Identifying just the type of the referent is not enough

The teacher erased the solutions before John had time to copy them out, as he had momentarily been distracted by a band playing outside.

A. This fact infuriated him, as the teacher always erased the board quickly and John suspected it was just to punish anyone who was lost in thought, even for a moment.

B. This fact infuriated the teacher, who had already told John several times to focus on class work.

Identifying just the type of the referent is not enough

Shell content for A

The teacher erased the solutions before John had time to copy them out, as he had momentarily been distracted by a band playing outside.

A. This fact infuriated him, as the teacher always erased the board quickly and John suspected it was just to punish anyone who was lost in thought, even for a moment.

B. This fact infuriated the teacher, who had already told John several times to focus on class work.

Identifying just the type of the referent is not enough

The teacher erased the solutions before John had time to copy them out, as he had momentarily been distracted by a band playing outside.

A. This fact infuriated him, as the teacher always erased the board quickly and John suspected it was just to punish anyone who was lost in thought, even for a moment.

B. This fact infuriated the teacher, who had already told John several times to focus on class work.

Ideas regarding proper evaluation of shell nouns

- Identifying the words, i.e., syntactic constituent, representing the required object type itself is challenging, as there is no one-to-one correspondence between a semantic concept and its syntactic shape

  - A concept like *issue* can take many different syntactic forms such as verb phrases, noun phrases, sentences, clauses

Ideas regarding proper evaluation of shell nouns

- Consider first *n* crowd answers rather than the top answer

- Extrinsic evaluation

  - E.g., to what extent shell noun resolution helps ESL learners

Automatically determining whether a noun is flexible

Noun	Proportion							total
	N-be-to	*N-be-that*	*N-be-wh*	*N-to*	*N-that*	*N-wh*	*N-of*	
problem	2	29	2	6	19	3	38	31,189
reason	-	9	-	59	15	15	2	41,230
fact	-	4	-	-	90	-	5	81,825

Flexible nouns tend to occur with a variety of patterns

48

The history of language is the history of a process of abbreviation.

— Friedrich Nietzsche

CrowdFlower Confidence

Annotator	Trust	Answer
A	0.75	"a
B	0.75	"a
C	1.0	"a

CrowdFlower Confidence

Annotator	Trust	Answer
A	0.75	"a
B	0.75	"a
C	1.0	"a

score for "a_2"=1.75

CrowdFlower Confidence

Annotator	Trust	Answer
A	0.75	"a
B	0.75	"a
C	1.0	"a

score for "a_1" = 0.75

score for "a_2" = 1.75

Crowd's answer:
"a_2" with confidence $0.7 = 1.75/(1.75+0.75)$
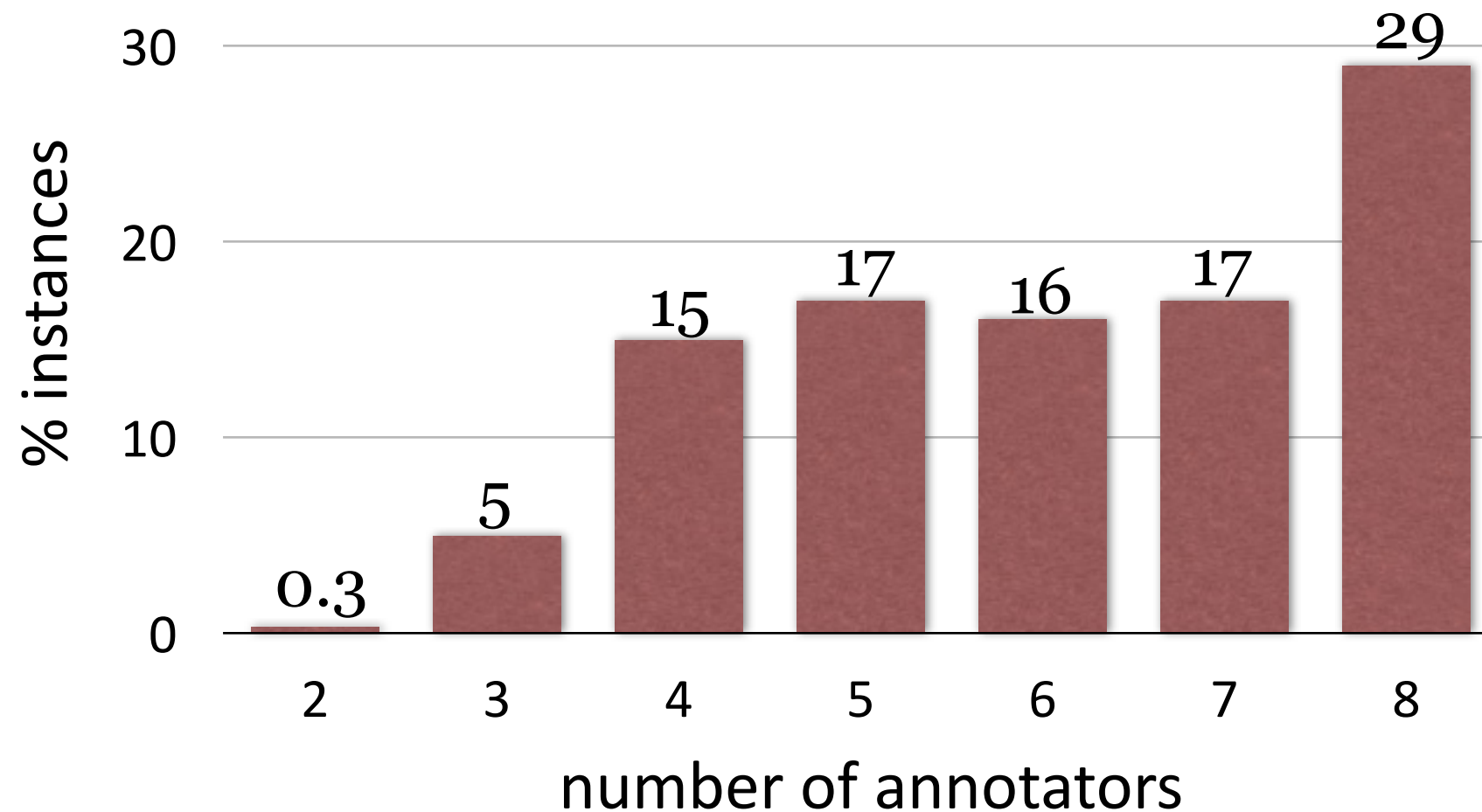
50

Satisfaction Level
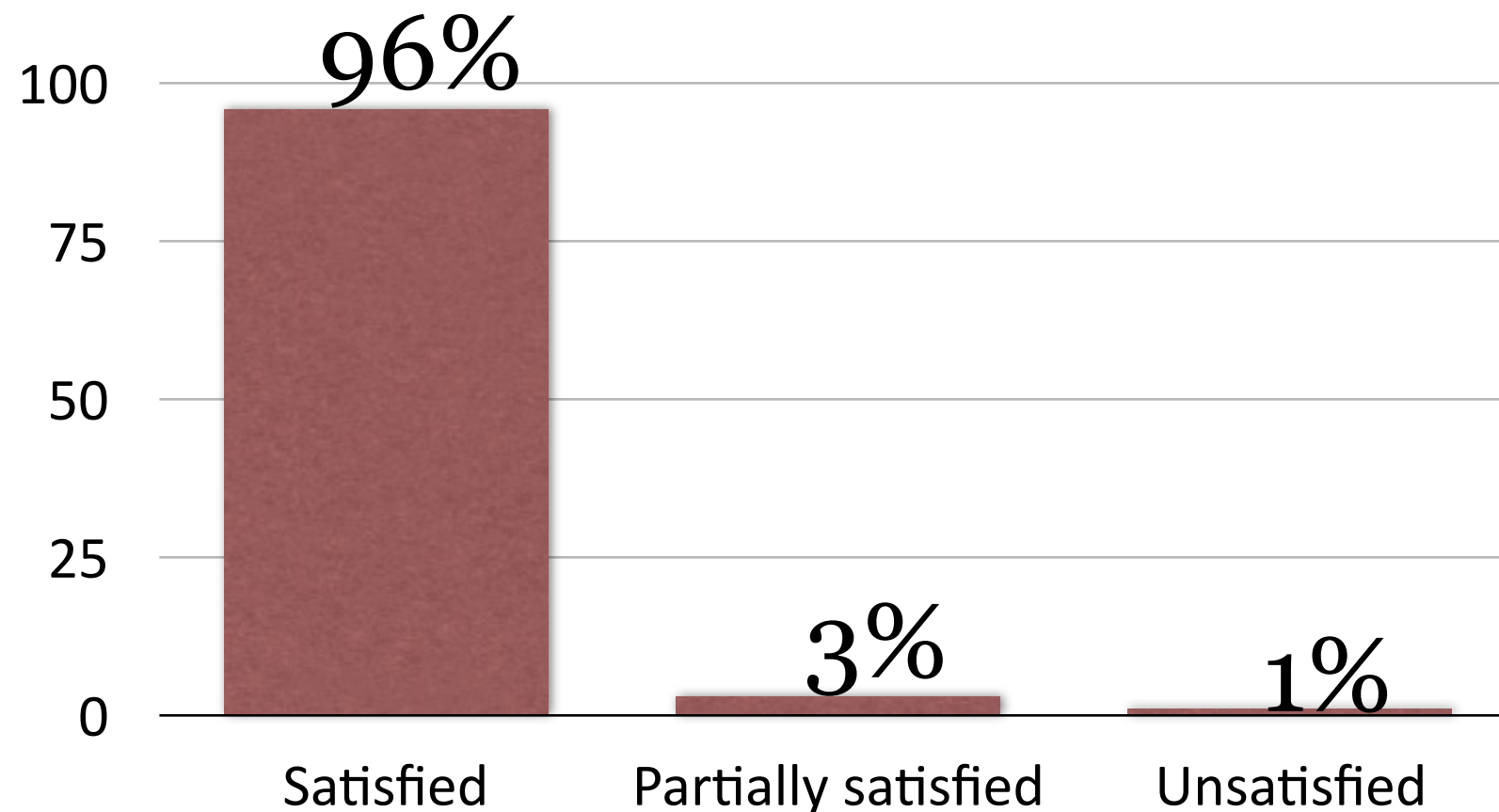
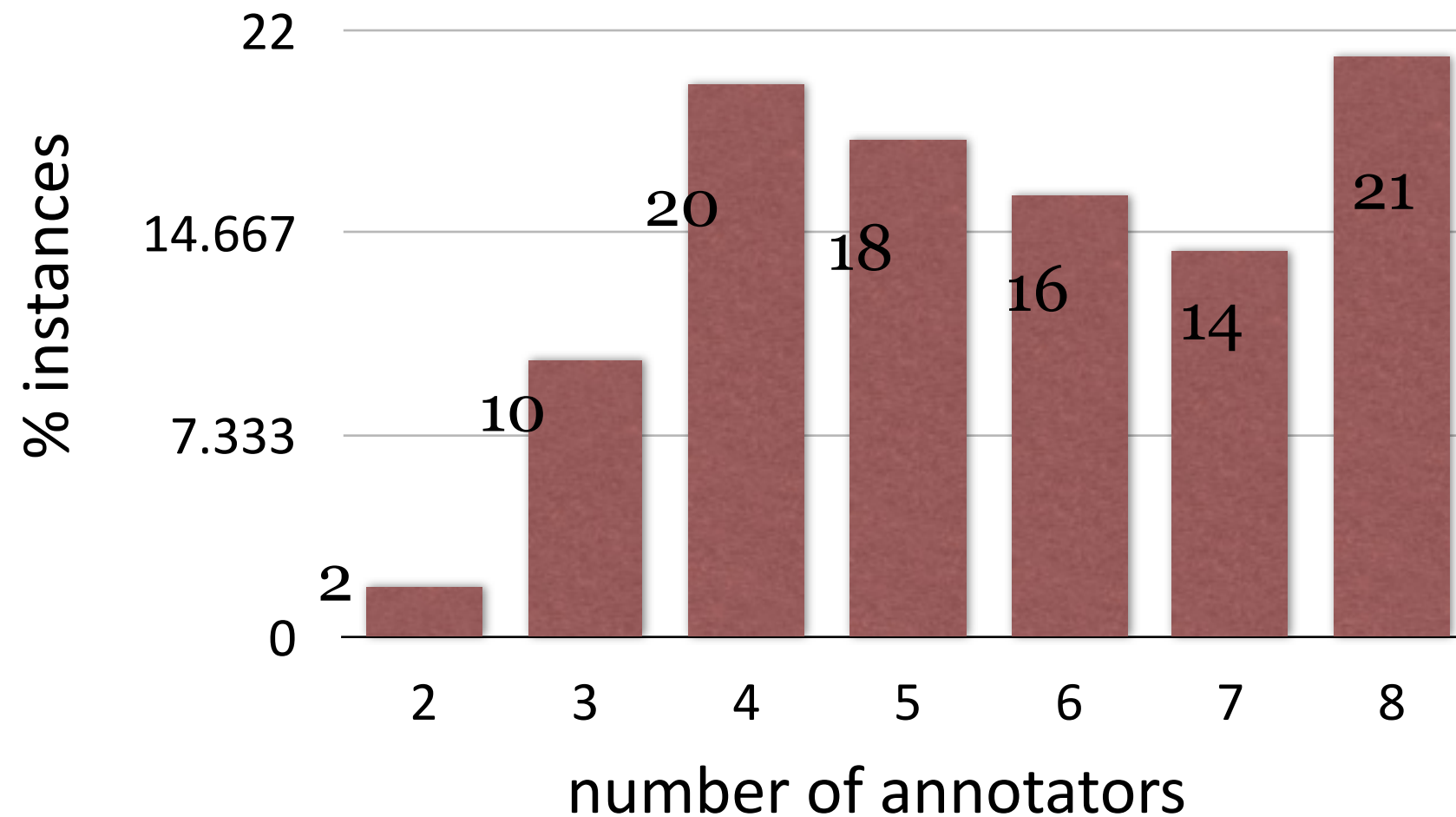Only 2% of the instances were labeled *None*

Evaluators were generally satisfied with the provided options.
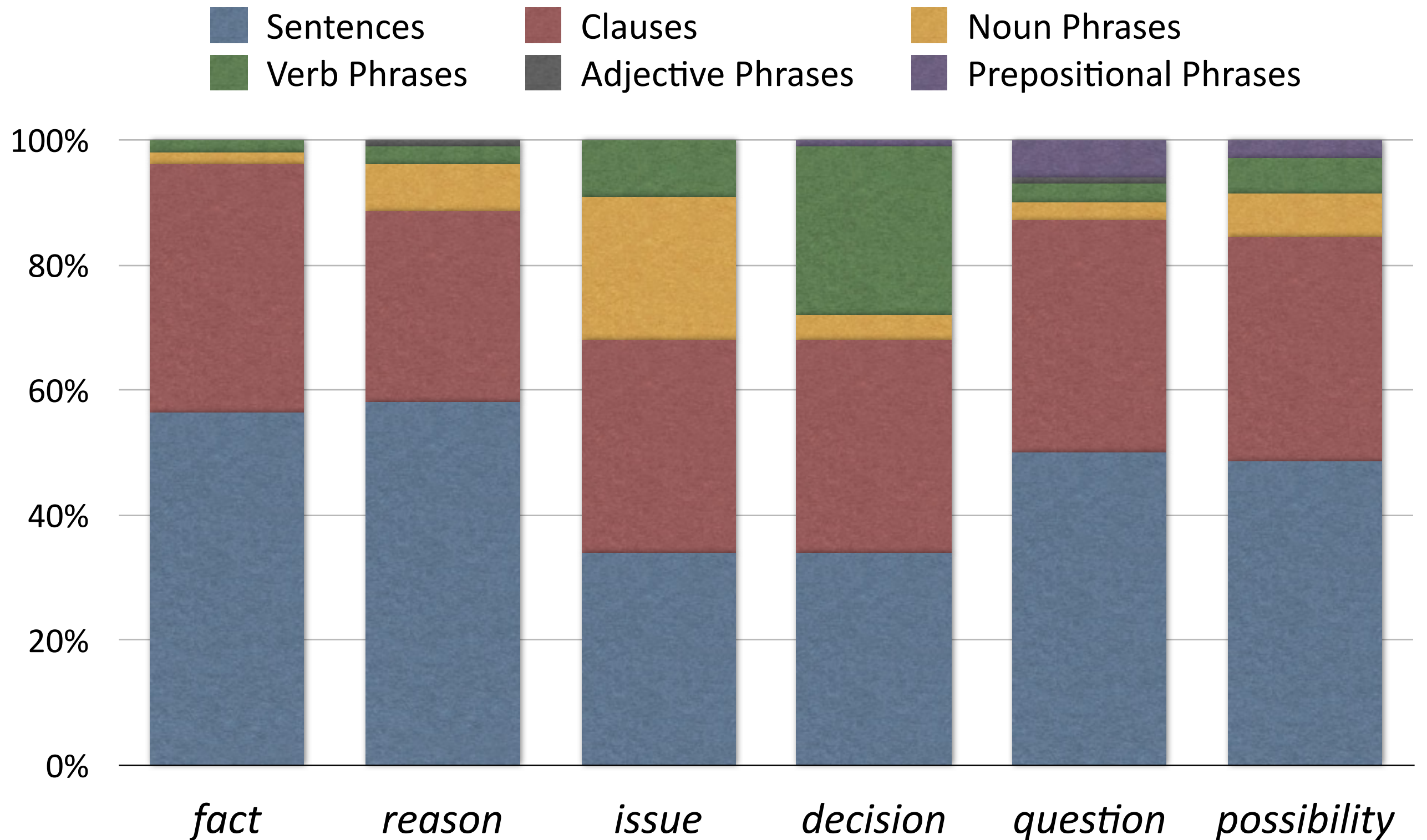
Head agreement

About 94% of the time, at least 4 annotators agreed on the head of the antecedent.

Exact agreement

About 89% of the time, at least 4 annotators agreed on the exact antecedent

Hard Examples

The teacher erased the solutions before John had time to copy them out, as he had momentarily been distracted by a band playing outside.

- This fact infuriated him, as the teacher always erased the board quickly and John suspected it was just to punish anyone who was lost in thought, even for a moment.

- This fact infuriated the teacher, who had already told John several times to focus on class work.

Hard Examples

Several Vatican officials said, however, that any such talk has little meaning because the church does not take sides in elections. But the statements by several American bishops that Catholics who vote for Mr. Kerry would have to go to confession have raised the question in many corners about whether this is an official church position.

The church has not addressed this question publicly and, in fact, seems reluctant to be dragged into the fight..."

Hard Examples

Any biography of Thomas More has to answer one fundamental question. Why? Why, out of all the many ambitious politicians of early Tudor England, did only one refuse to acquiesce to a simple piece of religious and politica opportunism? What was it about More that set him apart and doomed him to a spectacularly avoidable execution?

The innovation of Peter Ackroyd's new biography of More is that he places the answer to this question outside of More himself.