Reconsidering the Semantic Role(s) of Predicate Arguments

PRELIM Task 2 / SEMANTICS

FINAL PRESENTATION

Benjamin Van Durme and Kyle Rawlins,
Craig Harman, ***Drew Reisinger***,
Rachel Rudinger, Frank Ferraro

Theta roles (linguistics)
Semantic roles (comp. ling.)

Verbal meanings consist of information like:

Who did *what* to *who(m)?*

What happened to *which* individual?

…

Theta roles fill in the *who* (and some of the *what*)

Agent, **Patient**, **Theme**…

PropBank

(22) Frameset **hit** "strike"

Arg0: hitter

Arg1: thing hit, target

Arg2: instrument of hitting

Ex1: Agentive subject: "[$_{\text{Arg0}}$ He$_i$] digs in the sand instead of [$_{\text{Arg0}}$ *trace*$_i$*] *hitting* [$_{\text{Arg1}}$ the ball], like a farmer," said Mr. Yoneyama. (wsj_1303)

Ex2: Instrumental subject: Dealers said [$_{\text{Arg1}}$ the shares] were *hit* [$_{\text{Arg2}}$ by fears of a slowdown in the U.S. economy]. (wsj_1015)

Ex3: All arguments: [$_{\text{Arg0}}$ John] *hit* [$_{\text{Arg1}}$ the tree] [$_{\text{Arg2}}$ with a stick].

Palmer, Martha, Daniel Gildea, and Paul Kingsbury. "The proposition bank: An annotated corpus of semantic roles." Computational linguistics 31.1 (2005): 71-106.

Dowty ('91)

Dowty argued for the notion of:

proto-Agent and *proto*-Patient

Verb arguments only *tend* to have certain basic properties, and these *correlate* in Agent/Patient like ways

Arguments with more Agent properties tend to be **SUBJECT**, those with more Patient properties, **OBJECT**

Dowty, David. "Thematic proto-roles and argument selection." Language (1991): 547-619.

Dowty's Properties

Proto-Agent and Proto-Patient properties from Dowty (1991)

Proto-Agent properties	Proto-Patient properties
Volitional involvement in the event or state	Undergoes change of state
Sentience (and/or perception)	Incremental theme
Causing an event or change of state in another participant	Causally affected by another participant
Movement (relative to another participant)	Stationary relative to movement of another participant
Exists independently of the event named by the verb	Does not exist independently of the event, or not at all

- Decomposing proto-roles to properties (c.f. VerbCorner)
- Traditional thematic roles emerge as prototypical configurations of properties.

Hartshorne, Joshua K., Claire Bonial, and Martha Palmer. "The VerbCorner Project: Toward an Empirically-Based Semantic Decomposition of Verbs." EMNLP. 2013.

Kako ('06)

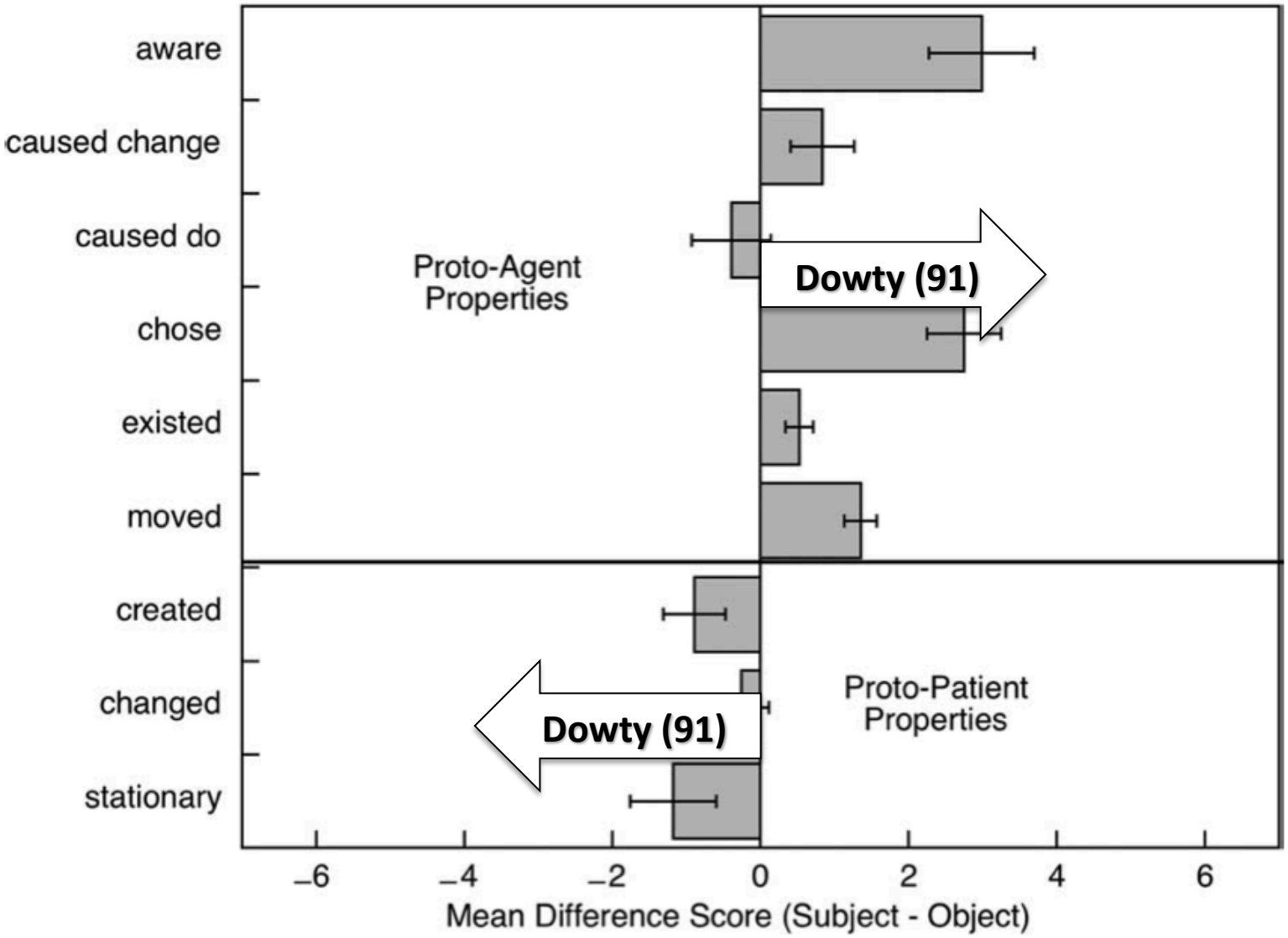
How stable are proto-role feature judgments when experimentally tested for a range of verbs?

The rom found the zarg.

- How likely is it that *the rom* chose to be involved in *finding*?
- How likely is it that *the rom* moved?
- …

Kako, Edward. "Thematic role properties of subjects and objects." Cognition 101.1 (2006): 1-42.

Kako's Findings



What do we want to do?

- Confirm prior work by Dowty, Kako
  - Test Dowty's introspective claims empirically
  - Evaluate models of syntactic realization beyond Dowty's
- Proof-of-concept: annotating binary properties
  - Are judgments good? Stable?
- Demonstrate common information in various semantic resources

Outline of Projects

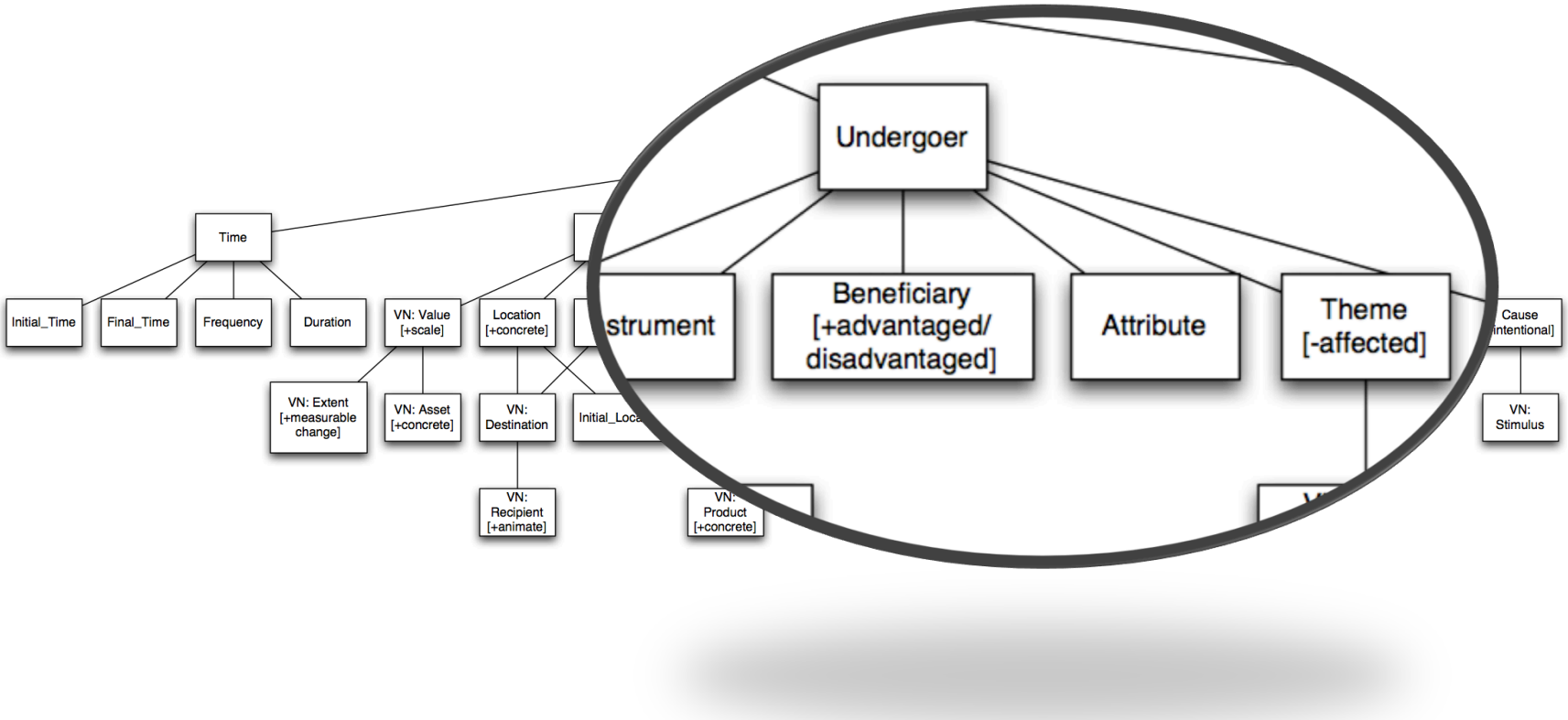
- Automatically build corpus of annotated properties with VerbNet and SemLink
- Annotate proto-role properties using Mechanical Turk
  - Real sentences from PropBank
  - Artificial sentences with nonsense (nonce) arguments

What is VerbNet?

- Large lexicon of English verbs, categorized into 274 different classes
- VerbNet's thematic roles can be organized into a hierarchy with property inheritance
- Each verb class is associated with some subset of these thematic roles and corresponding syntactic realization

Kipper-Schuler, K. (2005), "VerbNet: A broad-coverage, comprehensive verb lexicon." PhD. Thesis. Computer and Information Science Dept., University of Pennsylvania. Philadelphia.

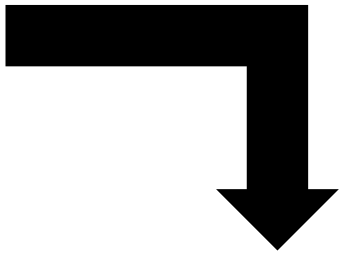
VerbNet Hierarchy



Bonial, Claire, William Corvey, Martha Palmer, Volha V. Petukhova, and Harry Bunt. "A hierarchical unification of LIRICS and VerbNet semantic roles." Semantic Computing (ICSC), 2011 Fifth IEEE International Conference on, pp. 483-489. IEEE, 2011.

Properties from VerbNet

- **Actor**: Participant that is the instigator of an event.
- **Cause**: Actor in an event (that may be animate or inanimate) that initiates the event, but that does not act with any intentionality or consciousness; it exists independently of the event. -LIRICS
- **Agent**: Actor in an event who initiates and carries out the event intentionally or consciously, and who exists independently of the event. -LIRICS
- ***Co-Agent***: *Agent who is acting in coordination or reciprocally with another agent while participating in the same event (specific to events with symmetrical participants).*[5]
- ***Stimulus***: *Cause in an event that elicits an emotional or psychologic*
- **Undergoer**: instigator of
- **Instrument** by an agent formed; it e



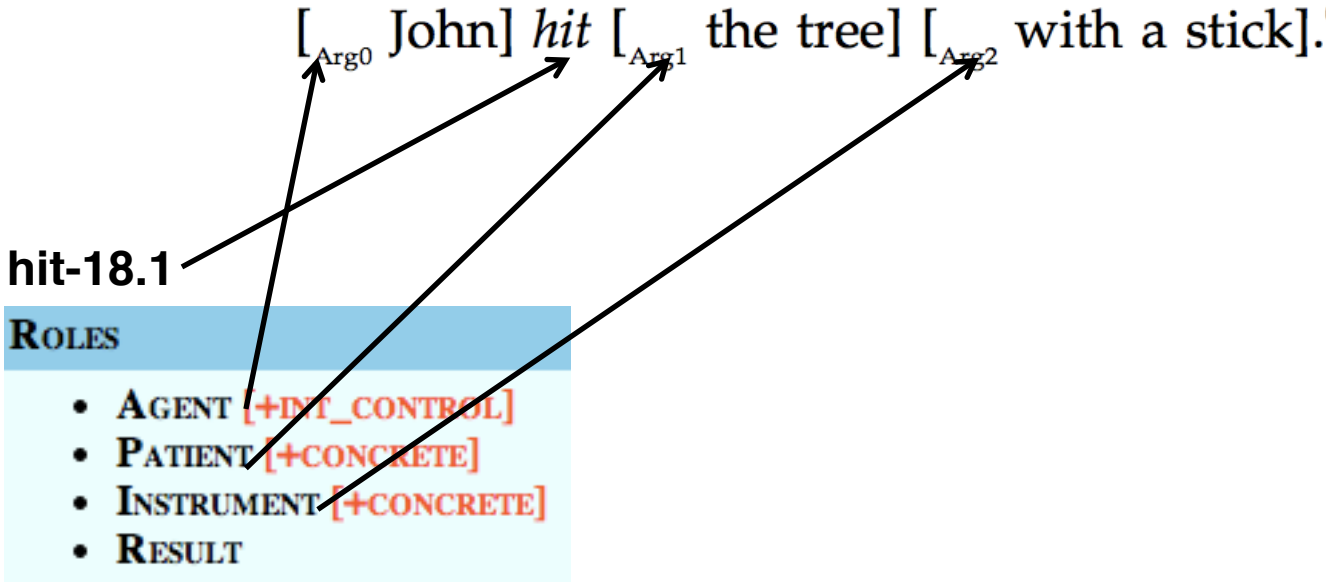
VN Role	Properties			
actor	instigates			
cause	instigates	causes	exists.independently	
agent	instigates	has.intent	exists.independently	
co-agent	instigates	has.intent	exists.independently	symmetric.participants
stimulus	instigates	causes	causes.perceptual-response	exists.independently
undergoer	not.instigates			
instrument	not.instigates	exists.independently	manipulated.by.an.agent	used.in.intentional.act

Extracting Properties from VerbNet

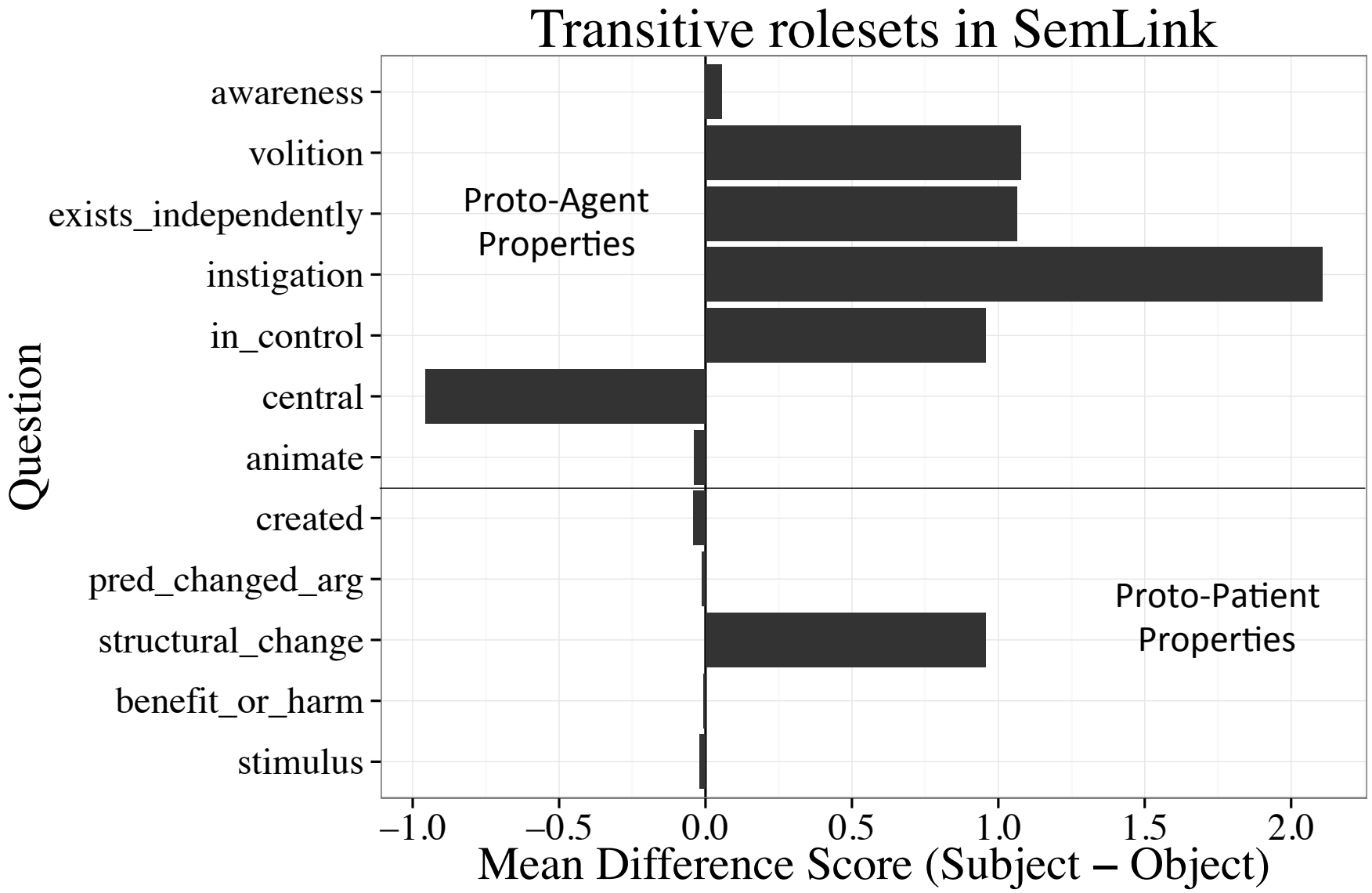
- Break down VerbNet role definitions into a simple discrete set of Dowty-like binary properties
- Two purposes:
  - To generate set of questions that elicit judgments about these properties for MTurk annotation
  - To induce a set of properties on any VerbNet-annotated data (e.g. SemLink)

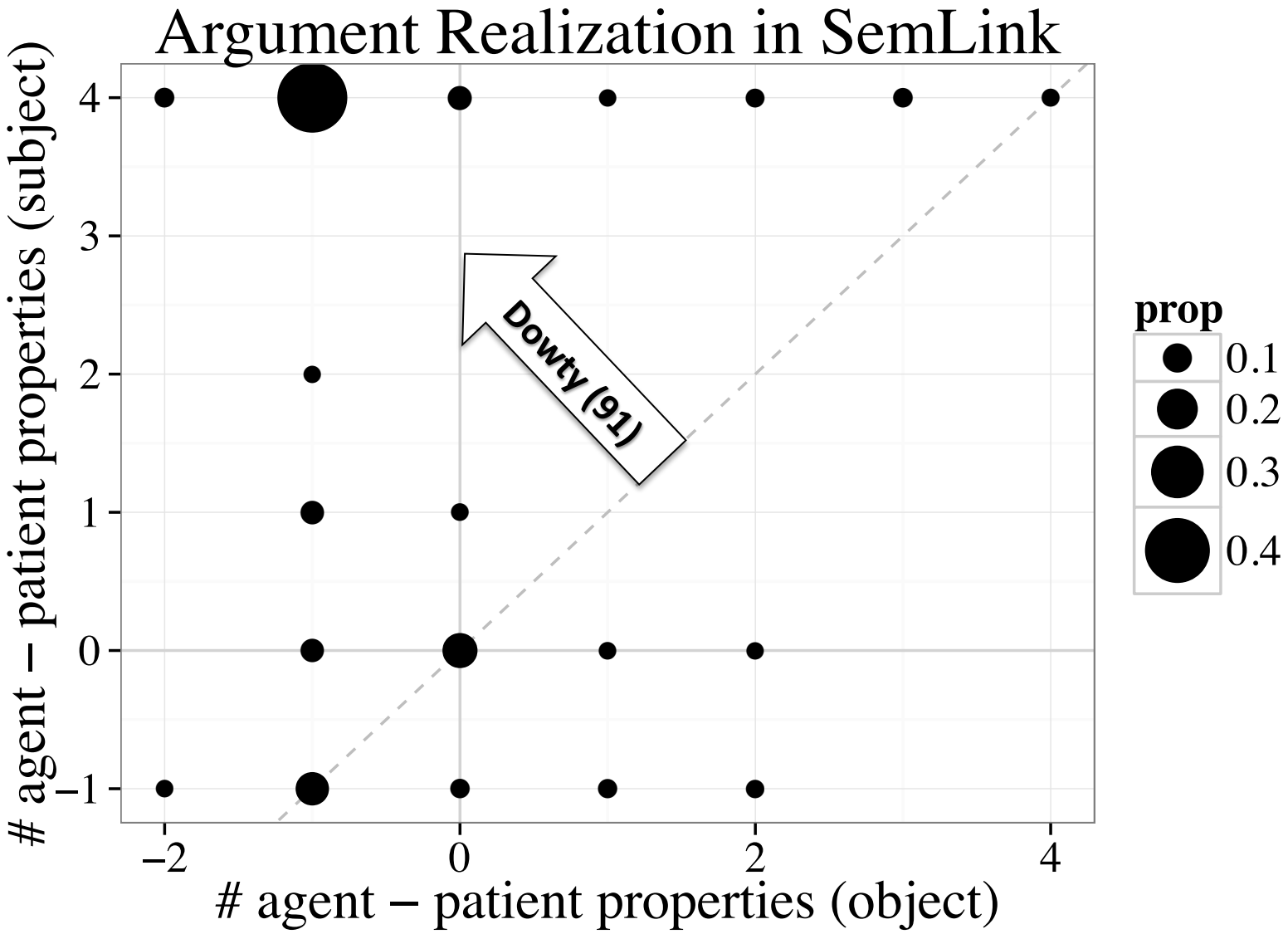
SemLink

- Partial VerbNet layer on top of PropBank



Palmer, Martha. "Semlink: Linking propbank, verbnet and framenet." Proceedings of the Generative Lexicon Conference. 2009.


Transitive rolesets in SemLink
awareness
volition
exists_independently
instigation
in_control
central
animate
created
pred_changed_arg
structural_change
benefit_or_harm
stimulus
Question
Proto-Agent Properties
Proto-Patient Properties
–1.0
–0.5
0.0
0.5
1.0
1.5
2.0
Mean Difference Score (Subject – Object)


Argument Realization in SemLink
agent – patient properties (subject)
agent – patient properties (object)
4
3
2
1
0
−1
−2
0
2
4
Dowty (91)
prop
0.1
0.2
0.3
0.4

Mechanical Turk Project

Sentence 1

An assassin in Colombia killed a federal judge on a Medellin street .

For a federal judge:

- How likely or unlikely is it that a federal judge chose to be involved in the killing?

  very unlikely | somewhat unlikely | not enough information | somewhat likely | very likely

---

- How likely or unlikely is it that the killing caused a change in a federal judge?

  very unlikely | somewhat unlikely | not enough information | somewhat likely | very likely

---

Mechanical Turk Project

- How likely or unlikely is it that The antibody was/were aware of being involved in the killing?

(•) very unlikely ( ) somewhat unlikely ( ) not enough information ( ) somewhat likely ( ) very likely

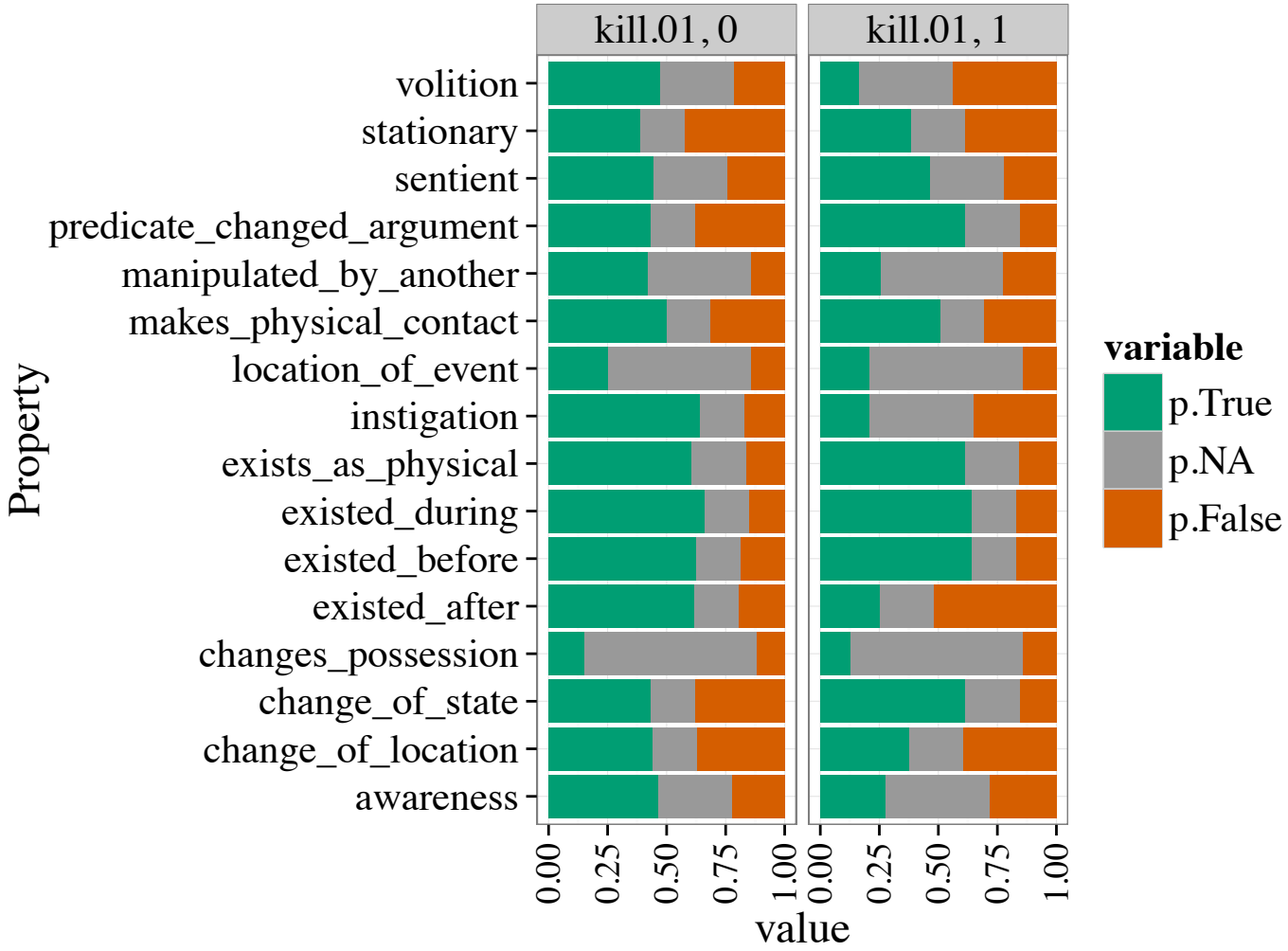
Does this question make sense? (•) No ( ) Yes

Argument Properties

- 16 properties annotated by Turkers on 16 rolesets
- Properties derived from various sources
  - VerbNet
  - Dowty (‘91), Kako (‘06), Grimm (‘10)
- ~25 properties identified but not used
- HIT can be extended easily to new properties

Grimm, Scott. "Semantics of case." Morphology 21.3-4 (2011): 515-544.

HIT Results: kill.01



Nonce-Argument Project

Task

Sentence 1

The neeglur killed the bogrub .

For the bogrub:

- How likely or unlikely is it that the the bogrub was/were altered or somehow changed during or by the end of the killing?
  ○ very unlikely ○ somewhat unlikely ○ not enough information ○ somewhat likely ○ very likely

- How likely or unlikely is it that the bogrub made physical contact with someone or something else involved in the killing?
  ○ very unlikely ○ somewhat unlikely ○ not enough information ○ somewhat likely ○ very likely

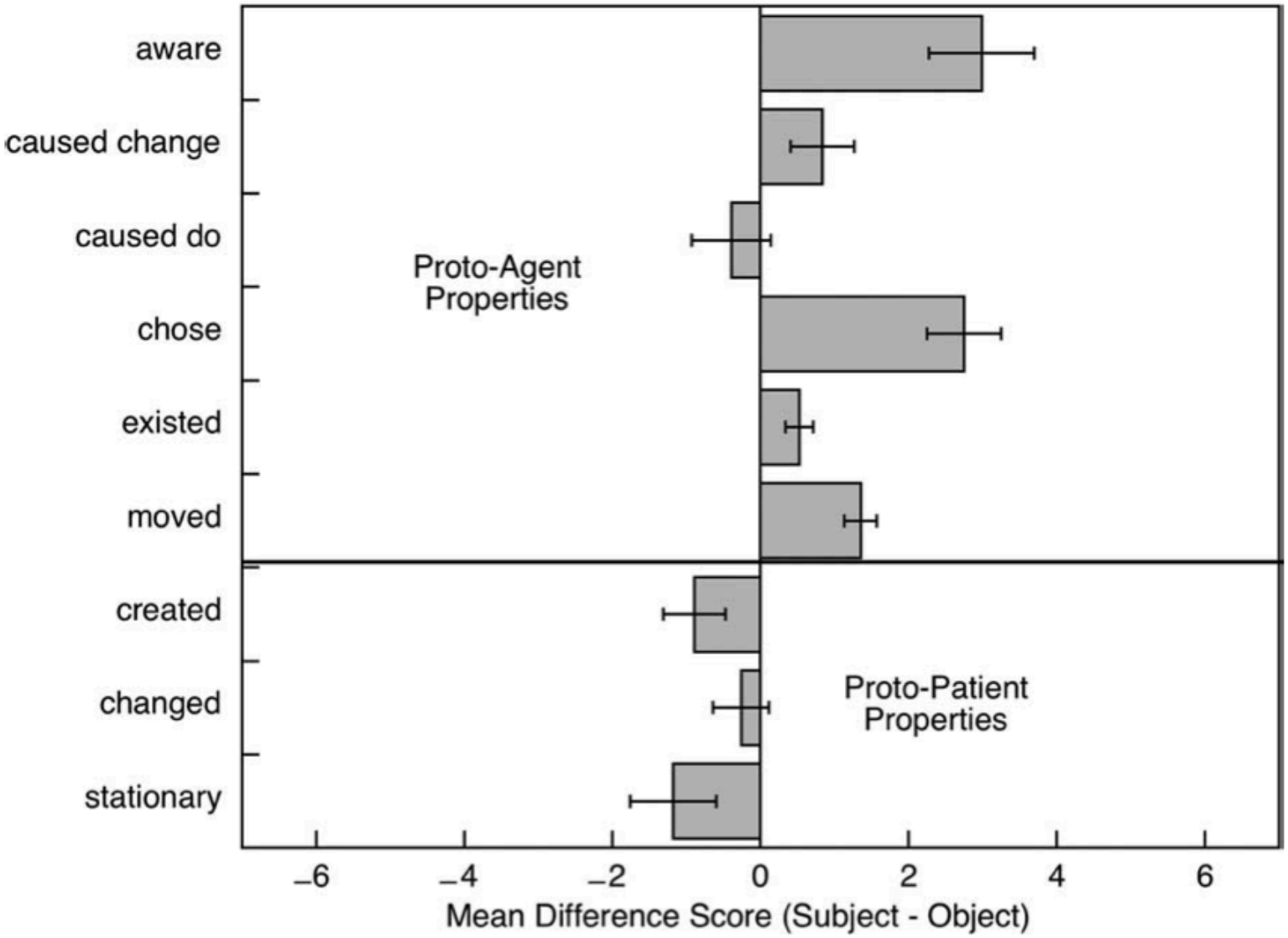
- How likely or unlikely is it that the bogrub was/were aware of being involved in the killing?
  ○ very unlikely ○ somewhat unlikely ○ not enough information ○ somewhat likely ○ very likely

- How likely or unlikely is it that the bogrub changes location during the killing?
  ○ very unlikely ○ somewhat unlikely ○ not enough information ○ somewhat likely ○ very likely

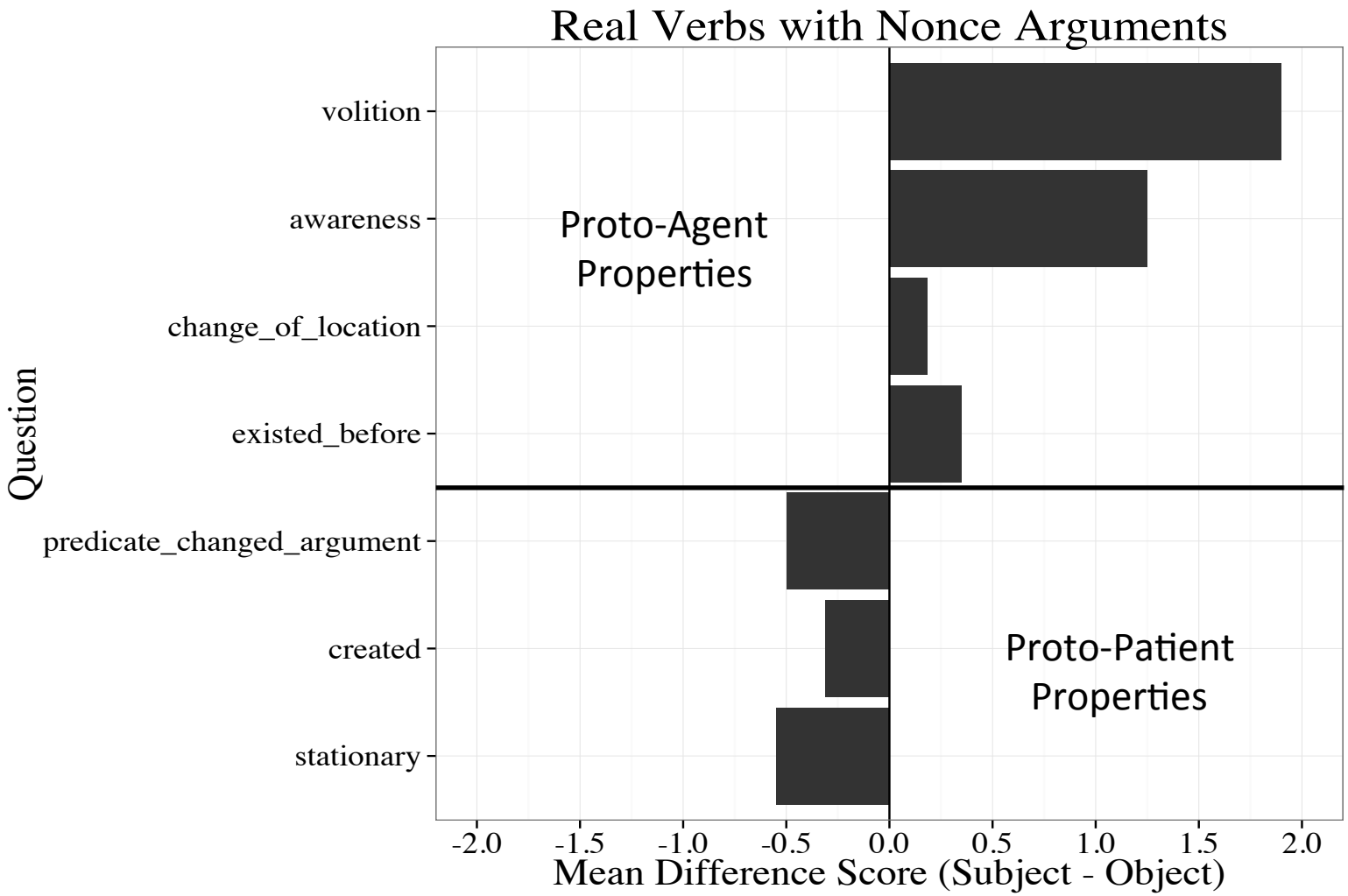
Motivation

- Easy to run, since we already have infrastructure
- Natural comparison with previous work
  - VerbCorner
  - Kako (‘06)
- Sanity check on MTurk annotation methodology
- Comparison between type- and token- level annotation.

Kako's Findings



Comparison with Kako



Conclusions

- We can use Mechanical Turk for direct annotation of predicate argument properties.
- The properties captured by VerbNet thematic roles support Dowty's proto-role hypothesis.
- Consistent Turker annotation of nonce argument properties further validates work by Dowty and Kako.

Future Directions

- Data-driven methods for selecting proto-role properties
- Extension to FrameNet frame elements
- Predictive modeling of PropBank labels/ VerbNet roles from properties
- Integration of symbolic and connectionist models of semantic compositionality (under NSF INSPIRE grant 'Gradient symbolic computation' BCS-1344269)