

Semantic Role Annotations for RRGparbank

Tatiana Bladier, Kilian Evang, Laura Kallmeyer

Heinrich Heine University Düsseldorf

TreeGraSP meeting #7
Düsseldorf, 16 February 2022

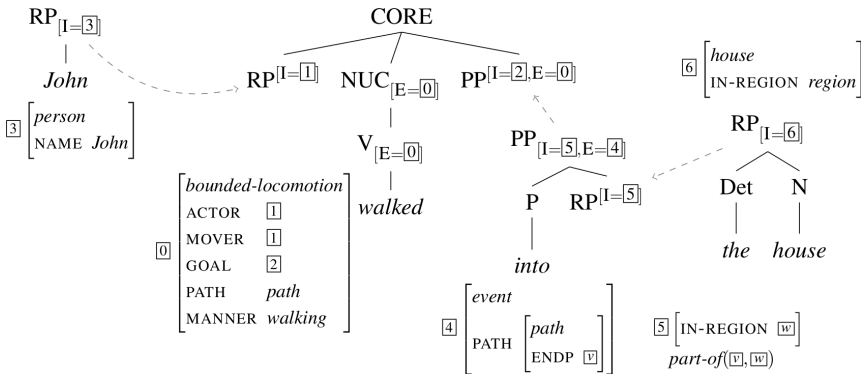


Introduction

In this talk

- ongoing effort to annotate RRGparbank with semantics roles
- why, how, current state

Motivation: towards Frame-semantic Parsing with RRG



Annotation Interface

Whatever	he	said	,	the	swift	answer	crushed	him	like	a	bludgeon	.
Topic	Agent	V										
						Agent	V	Patient				

<https://rrgparbank.phil.hhu.de>

Annotators

- since December: 2 student assistants à 8h/week
- Tanja and Kilian also annotate and adjudicate

What We Are Annotating

For now:

- *1984* in English and German
- only verbal predicates (V; not AUX, not OP, not N, not A, not ADV)
- all arguments
- adjuncts (-PERI), but *only* when a role from our inventory clearly fits (Time, Location, Extent, Purpose...); unclear cases are left unannotated

Choosing a Role Inventory

Our initial choice: VerbNet (Kipper-Schuler, 2005)

Affector	Context	Instrument	Product	Topic
Agent	Destination	Location	Purpose	Trajectory
Attribute	Duration	Manner	Recipient	Value
Axis	Experiencer	Material	Reflexive	
Beneficiary	Extent	Path	Result	
Cause	Final_Time	Patient	Source	
Co-Agent	Goal	Pivot	Stimulus	
Co-Patient	Initial_Location	Precondition	Theme	
Co-Theme	Initial_State	Predicate	Time	

Choosing a Role Inventory, cont.

Expectation

- roles are well defined
- can be applied across languages without explicit reference to frame

Reality

- pretty arbitrary choices are often necessary
- e.g., Agent vs. Experiencer, Patient vs. Theme...

Choosing a Role Inventory, cont.

Expectation

- decide by looking up frames in VerbNet

Reality

- low coverage (only 6 791 English verbs)
- “frames” are not defined in a semantically coherent way
- few, incomplete examples
- annotation is frustrating; agreement is low

VerbAtlas (Di Fabio et al., 2019)

- high coverage (13 767 verbal WordNet synsets)
- organized into 466 coarse-grained, semantically coherent frames
- each frame has a number of roles
- each (frame, role) pair is mapped to one or more PropBank (frame, role) pairs; these have specific descriptions!
- high coverage makes it easy to find the right frame for a verb
- granularity of frames makes it easy to find the right frame even for verbs that are not in WordNet (e.g., non-English verbs)
- mapping to PropBank makes it easy to unambiguously map arguments to roles

VerbAtlas Role Inventory

Agent	Co-Patient	Goal	Product	Stimulus
Attribute	Co-Theme	Instrument	Purpose	Theme
Beneficiary	Destination	Location	Recipient	Time
Cause	Experiencer	Material	Result	Topic
Co-Agent	Extent	Patient	Source	Value

Example

Example

(1) Die Uhren[?] schlugen[V] dreizehn[?]

Steps to determine roles

- ① *schlagen* $\hat{=}$ WordNet synset “strike (indicate (a certain time) by striking)”
- ② mapped to VerbAtlas SIGNAL_INDICATE frame
- ③ contains roles Agent, Recipient, Topic, Instrument, Location
- ④ mapped to (among others) PropBank roleset indicate.01
- ⑤ Agent \rightarrow ARG0 (“indicator”); Topic \rightarrow ARG1 (“thing indicated”)
- ⑥ thus, the roles are Agent and Topic

Frame Overview Document for Annotators

SIGNAL_INDICATE

Synsets

tip (indicate a certain time) by striking, threaten (to be a menacing indication of something:), semaphore (send signals by or as if by semaphore), stigmatize; stigmatise (mark with a stigma or stigmata), point (indicate the presence of (game) by standing and pointing with the muzzle), wigwag (send a signal by waving a flag or a light according to a certain code), tip (mark with a tip), bespot (mark with, or as if with, spots), block-out (indicate roughly), pluralize; pluralise (mark with a grammatical morpheme that indicates plural), bespeak; betoken; indicate; point; signal (be a signal for or a symptom of), beckon; wave (signal with the hands or nod), indicate; point; designate; show (indicate a place, direction, person, or thing; either spatially or figuratively), bode; portend; auspicate; prognosticate; omen; presage; betoken; foreshadow; augur; foretell; prefigure; forecast; predict (indicate by signs), crisscross (mark with a pattern of crossing lines), signalize; signalise; point-out; call-attention (point out carefully and clearly), ring; band (attach a ring to the foot of, in order to identify), line (mark with lines), whistle (give a signal by whistling), flag-down (signal to stop), flag (communicate or signal with a flag), groan; moan (indicate pain, discomfort, or displeasure), say (indicate), brand; trademark; brandmark (mark with a brand or trademark), signpost (mark with a signpost, as of a path), wigwag (signal by or as if by a flag or light waved according to a code), finger (indicate the fingering of musical scores for keyboard instruments), bless; sign (make the sign of the cross over someone in order to call on God for protection; consecrate), buoy (mark with a buoy), receipt (mark or stamp as paid), denote (be a sign or indication of), heliograph (signal by means of a mirror and the using the sun's rays), code (attach a code to), blaze (indicate by marking trees with blazes), spot (mark with a spot or spots so as to allow easy recognition), brand (burn with a branding iron to indicate ownership; of animals), wink (signal by winking), mark (designate as if by a mark), star; asterisk (mark with an asterisk), meter (stamp with a meter indicating the postage), mark (make or leave a mark on), point, the way (indicate the right path or direction), euphemize; euphemise (refer to something with a euphemism), stake; post (mark with a stake), check; check-off; mark; mark-off; tick-off; tick (put a check mark on or near or next to), shake (shake (a body part) to communicate a greeting, feeling, or cognitive state), spell; import (indicate or signify), dot (mark with a dot)

Roles and PropBank Mappings

SIGNAL_INDICATE	Agent	Recipient	Topic	Instrument	Location
beckon.01					
bode.01	omen	benefactive	good or bad		
foreshadow.01	predictor of future		prediction	means of prediction	
groan.01	groaner	groan at			
indicate.01	indicator	indicated to	thing indicated		
line.03					
mark.02	agent, marker		mark	instrument, marker	surface, marked-on
point.02	pointer		pointed at		
portend.01	thing portending		portend for	sign	
spell.02	theme		explanation		
wave.01	agent, causer	person waved at	entity moving		
wave.02	waver, giver of directions	receiver of directions	directions		
wink.01	winker	winked at		eye	

Examples from PropBank

- Still, AMR's stumble doesn't [bode]_v [well]_{Topic} for the rest of the industry.
- The new reports of sluggishness, which [*T*-2] were [foreshadowed]_v [*-1] by an earlier Labor Department report that manufacturing payrolls dropped by 105,000 in September, give the Fed another reason to further ease its grip on credit and lower interest rates.
- While Friday's plunging stock market prompted new fears about the economy's prospects, a little-known indicator that [*T*-1] has faithfully [foreshadowed]_v [the economy's ups and downs]_{Topic} by [exceptionally] long lead times]_{Instrument} points to a sustained rise in overall business activity.
- [He]_{Agent} often [groans]_v, he says [0] [*T*-3], at the "clutter" of sports stories about drugs, alcohol, gambling and some player's lament "about the misery millions [0] he is offered [*-1] [*T*-2] [*-1] to play the game]_{Recipient}."

<https://rrgparbank.phil.hhu.de/editor/static/doc/verbatlas-1.0.3.html>

Annotation Guidelines

https://www.dropbox.com/s/e4hgr8jj4fm1zr4/Annotation_Guidelines_Semantics.pdf?dl=0

- ① what to annotate (see above)
- ② how to annotate (choosing roles, see above)
- ③ specific guidelines wrt. syntax
 - determining the head of a constituent
 - annotating antecedents in control, raising, secondary predicates, purpose infinitives, *tough*-construction, relative clauses, etc.
 - “small clauses”
 - coordination and ellipsis
 - figurative use of predicates
 - non-role arguments: expletive pronouns, weather pronouns, multiword expressions

Heads

- (2)
- a. Winston Smith_[Theme] slipped_[V] through the door_[Destination]
 - b. all that was said of Goldstein seemed_[V] to him_[Experiencer] to be true_[Theme]

Annotating Antecedents

(3) *Object control*

- a. He_[Agent] asked_[V] me_[Recipient] to dance_[Topic]
- b. He asked me_[Agent] to dance_[V]

(4) *Subject control*

- a. He_[Agent] promised_[V] me_[Recipient] to dance_[Topic]
- b. He_[Agent] asked me to dance_[V]

(5) *Raising*

- a. He seems_[V] to be dancing_[Theme]
- b. He_[Agent] seems to be dancing_[V]

Annotating Antecedents, Cont.

(6) *Secondary predicates*

- a. He_[Agent] entered_[V] the room_[Destination], singing and dancing
- b. He_[Agent] entered the room, singing_[V] and dancing
- c. He_[Agent] entered the room, singing and dancing_[V]
- d. als Winston Smith_[Agent], das Kinn_[Theme] an die Brust_[Destination] gepresst_[V], um den rauhen Wind wo entgehen, rasch durch die Türen eines der Häuser des Victory-Blocks schlüpfte

Annotating Antecedents, Cont.

(7) *Purpose infinitives*

- a. als Winston Smith_[Theme], das Kinn an die Brust gepresst, um dem rauhen Wind_[Source] zu entgehen_[V], rasch durch die Türen eines der Häuser des Victory-Blocks schlüpfte
- b. dass es in der Küche nichts_[Patient] zu essen_[V] gab außer einem Stück Schwarzbrot
- c. wenn man_[Agent] es fertig brachte, dem Televisor ein Schnippchen zu schlagen, und ihn_[Theme] allein zu fassen_[V] bekam
- d. She_[Recipient] is too important for you_[Agent] to talk_[V] to

(8) *Temporal infinitives*

- a. He_[Theme] left_[V] after trashing_[Time] the room
- b. He_[Agent] left after trashing_[V] the room_[Patient]

Annotating Antecedents, Cont.

- (9) *tough-construction*
- a. He_[Stimulus] is hard to love_[V]

“Small Clauses”

- (10)
- a. zuerst sah_[V] man_[Experiencer] ihn_[Stimulus] sich durchs Wasser wälzen_[Stimulus]
 - b. you_[Experiencer] saw_[V] him_[Stimulus] wallowing_[Stimulus] along in the water

Coordinated Predicates

- (11)
- a. She_[Agent] was singing_[V] and dancing
 - b. She_[Agent] was singing and dancing_[V]
 - c. Mary_[Experiencer] loves_[V], and Ben hates, jazz_[Stimulus]
 - d. Mary loves, and Ben_[Experiencer] hates_[V], jazz_[Stimulus]

Coordinated Arguments

- (12)
- a. Die Frauen_[Theme], und vor allem die jungen_[Theme], gaben_[V] immer_[Time] die blind ergebenen Parteil Anhänger_[Attribute], die gedankenlosen Nachplapperer_[Attribute], die freiwilligen Spitzel_[Attribute] ab, mit deren Hilfe man weniger Linientreue aushorchen konnte
 - b. Sam_[Agent] and Alex_[Agent] married_[V]
 - c. Sam_[Agent] married Alex_[Co-Agent]

Elided Predicates

- (13) a. Sam_[Agent] married_[V] Alex_[Co-Agent], and Bill Kathy

Expletives and Weather Pronouns

- (14)
- a. I_[Experiencer] hate_[V] it when people yell_[Stimulus]
 - b. dass es ihm_[Experiencer] fast so_[Attribute] vorkam, als habe er sie niedergeschrieben
 - c. there seemed_[V] to be_[Theme] no colour in anything
 - d. It is raining_[V]

Figurative Use of Predicates

- (15) a. Die Augen_[Agent] bohrten_[V] sich_[Reflexive] tief_[Extent] in
Winstons Blick_[Destination]

Multiword Expressions

(16) *Verbal idioms*

- a. Don't spill_[V] the beans_[VID]
- b. He_[Patient] kicked_[V] the bucket_[VID]
- c. Are you_[Agent] pulling_[V] my leg_[VID]?
- d. Goldstein_[Agent] ließ_[V] seinen üblichen giftigen Angriff_[Theme] gegen die Lehren der Partei vom Stapel_[VID]
- e. You_[Theme] were supposed to stand_[V] to attention_[VID]

Multiword Expressions, Cont.

(17) *Light verbs*

- a. I_[Agent] 'm taking_[V] a bath_[LVC]
- b. He_[Theme] took_[V] a seat_[LVC]

Multiword Expressions, Cont.

(18) *Verb-particle construction*

- a. The wedding_[Theme] will take_[V] place on Saturday_[Time]
- b. She_[Agent] took_[V] part in the race_[Theme]
- c. Der Große Bruder_[Experiencer] sieht_[V] dich_[Stimulus] an!

Multiword Expressions, Cont.

(19) *Inherently reflexive verbs*

- a. Ich_[Experiencer] sah_[V] mich_[Reflexive] um
- b. He_[Agent] fanned_[V] himself_[Experiencer]

Annotation Status

- methodology
- progress
- next steps

Annotation Methodology

- pre-annotation using InVeRo-SRL
- 2 annotators + 1 adjudicator per sentence
- different combinations of annotators and adjudicators
- weekly meetings

Annotation Progress

- first 425 sentences, German and English, 2 annotators, not yet adjudicated
- but: still based on VerbNet, no pre-annotation
- with VerbAtlas, pre-annotation and more practice, we expect annotators to be able to annotate at least 100 sentences (50 English, 50 German) in 8h, probably more
- at this pace, German and English parts of RRGparbank will be fully annotated in about 2y

Next Steps

- fewer meetings
- annotate VerbAtlas frames as well?
- speed up annotation: hire more annotators?

References I

- Di Fabio, A., Conia, S., and Navigli, R. (2019). VerbAtlas: a novel large-scale verbal semantic resource and its application to semantic role labeling. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 627–637, Hong Kong, China. Association for Computational Linguistics.
- Kipper-Schuler, K. (2005). *VerbNet: A broad-coverage, comprehensive verb lexicon*. PhD thesis, University of Pennsylvania.