

Double Trouble: The Problem of Construal in Semantic Annotation of Adpositions

Jena Hwang



Na-Rae Han



Vivek Srikumar



Archana Bhatia



Tim O'Gorman



Nathan Schneider

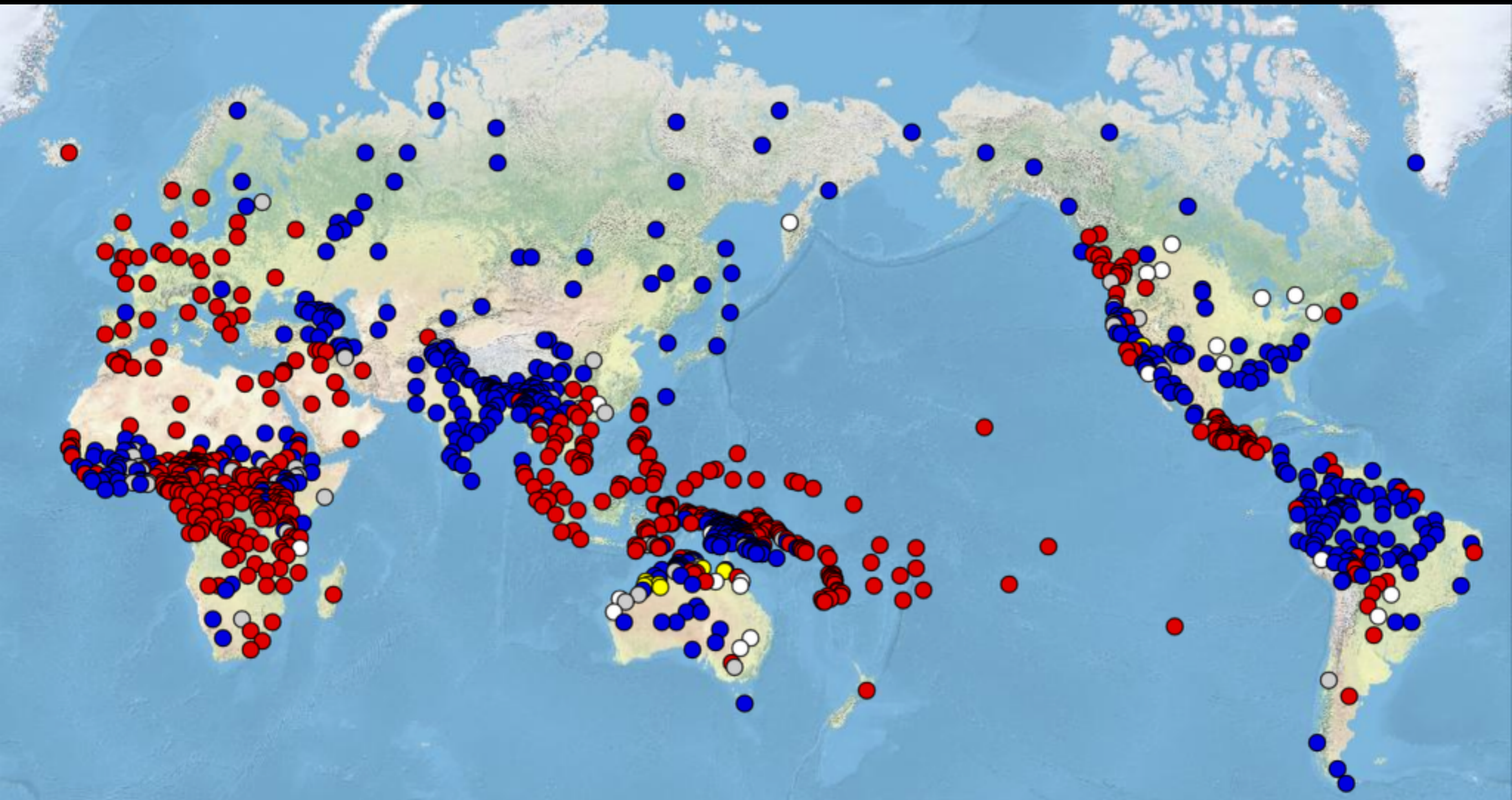


Most languages have **adpositions.**



adposition = preposition
| postposition





Feature 85A: Order of Adposition and Noun Phrase
Dryer in *WALS*, <http://wals.info/chapter/85>

We know PPs are challenging
for syntactic parsing.

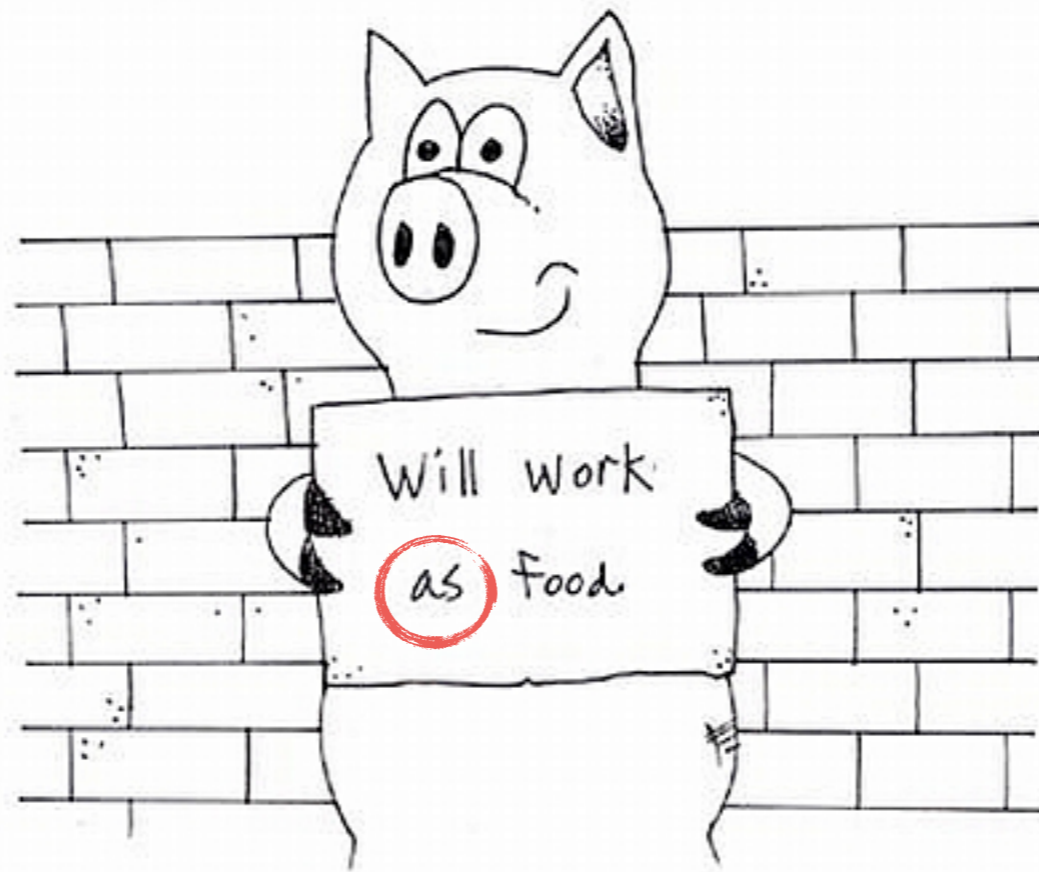
a talk **at** the workshop **on** prepositions

But what about the meaning *beyond*
linking governor & modifier?

“I study preposition semantics.”



Adpositions have semantics?!



Due to his grammar mistake, Wilbur found a position. It just wasn't the one he wanted.

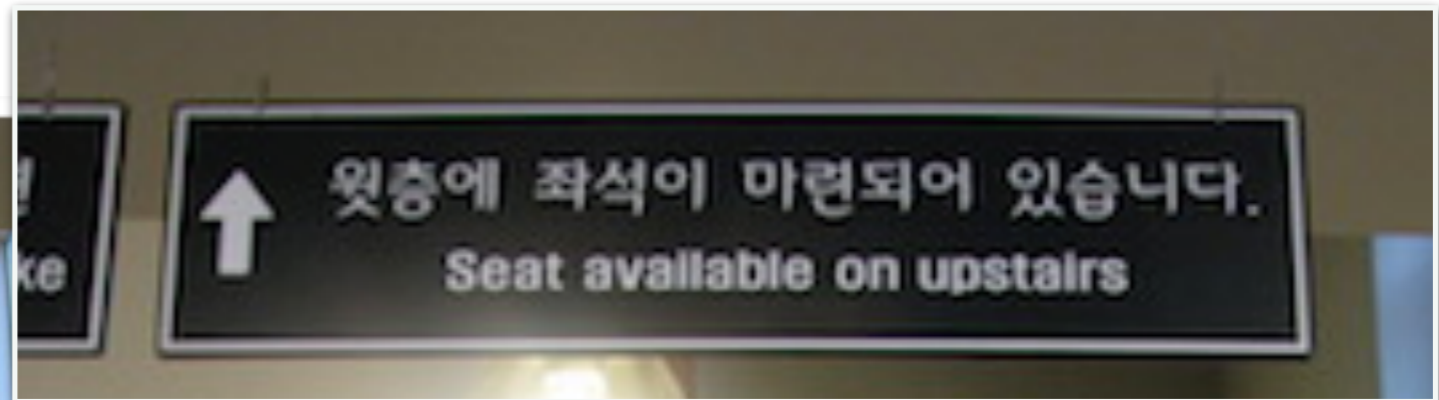
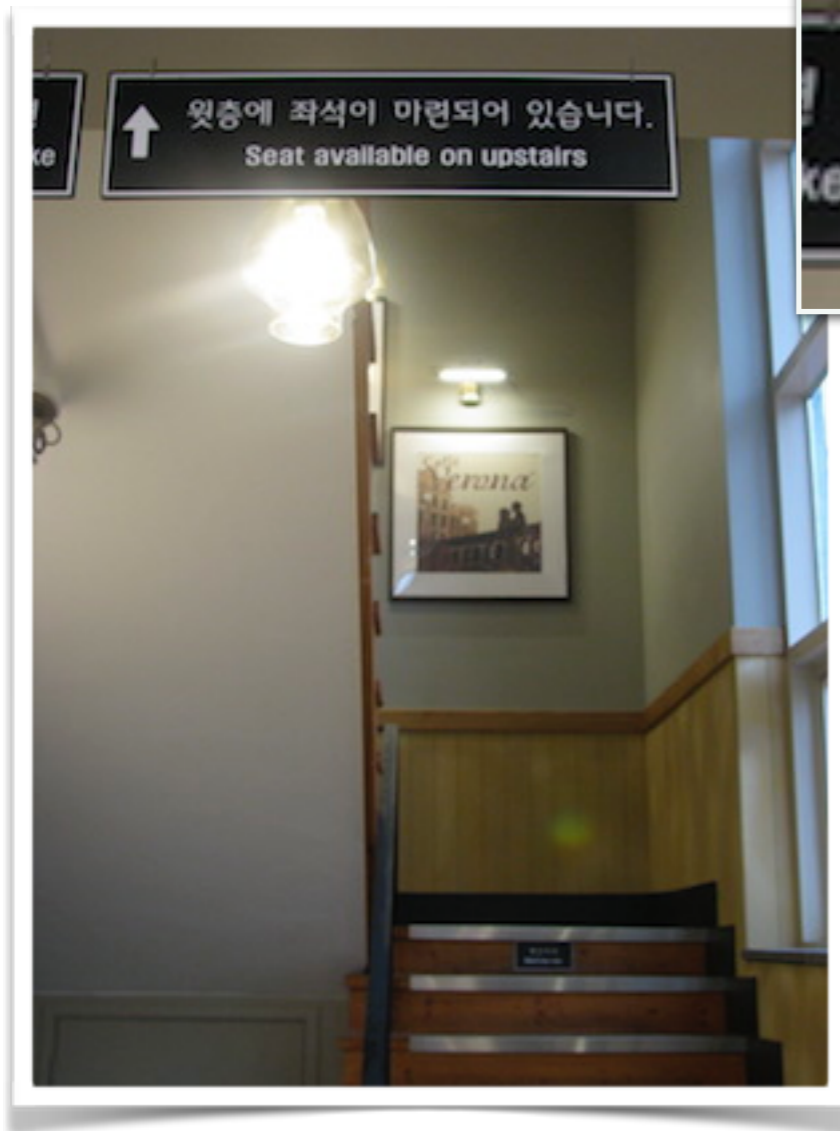
Polysemy

- With great frequency comes great **polysemy**.
- *in*
 - ▶ **in** the box
 - ▶ **in** the afternoon
 - ▶ **in** love, **in** trouble
 - ▶ **in** fact
 - ▶ ...

Cross-linguistically interesting

- Small number of grammatical categories
- Language-specific partitioning of functions
- Translations are many-to-many

Bewildering to learn in an L2



Shared functions

They ran **to** the roof **for** a quick escape.



DESTINATION PURPOSE



They made **for** the roof **to** escape the cops.

Design Principles

1. Coverage

2. Cross-linguistic adequacy

Design Principles

- 1. Coverage:** Annotate all adposition types and tokens in a corpus.
- 2. Cross-linguistic adequacy**

Design Principles

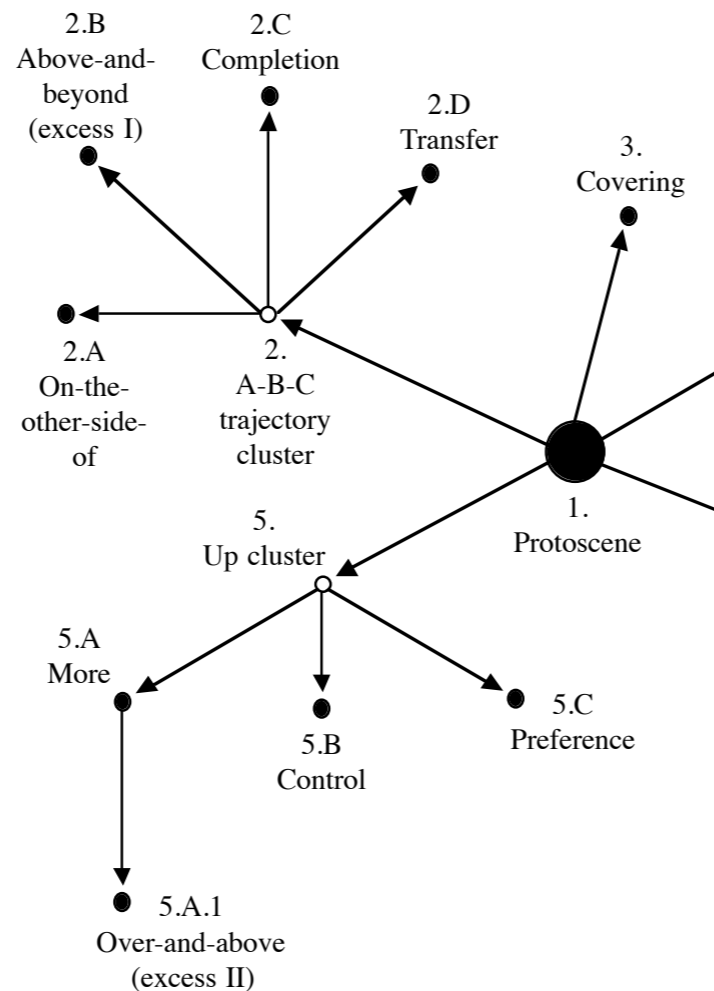
- 1. Coverage:** Annotate all adposition types and tokens in a corpus.
- 2. Cross-linguistic adequacy:** Our semantic functions should be as language-independent as possible.

Senses vs. Supersenses

over

• preposition

- » 1. extending directly upwards from: *I saw flames of heat.*
 - above so as to cover or protect: *an oxygen tent over the pasta.*
 - extending above (an area) from a vantage point
- » 2. at a higher level or layer than: *his flat was over the ground.*
 - higher in rank than: *over him is the financial director.*
 - expressing authority or control: *editorial control over the press.*
 - expressing preference: *I'd choose the well-known over the unknown.*
 - expressing majority: *the predominance of Asia over Europe.*
 - higher in volume or pitch than: *he shouted over the noise.*
- » 3. higher or more than (a specified number or quality): *she has been married for over a year.*
- » 4. expressing passage or trajectory across: *she trooped over the field.*
 - beyond and falling or hanging from: *he toppled over the wall.*
 - at the other side of; beyond: *over the hill is a small village.*
- » 5. expressing duration: *you've given us a lot of help over the years.*
- » 6. expressing the medium by which something is done: *he spoke over the loudspeaker.*
 - by means of the resistant surface constituted by: *the car rolled over the sand.*
- » 7. on the subject of: *a heated debate over unemployment.*



fine-grained details

lexeme-specific

(extensive linguistic & AI research
on space & time)

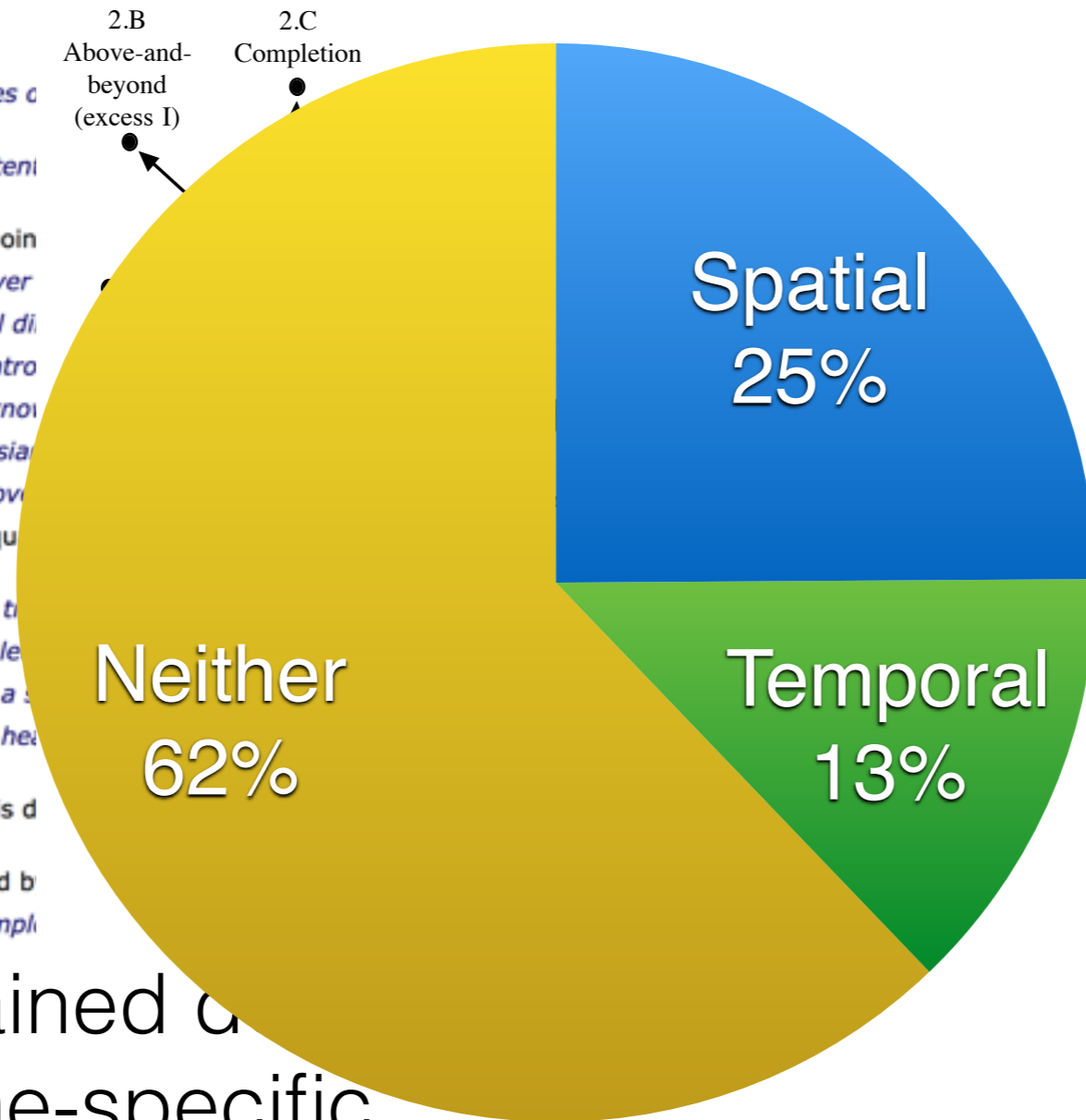
Senses vs. Supersenses

over

• **preposition**

- » 1. extending directly upwards from: *I saw flames of heat.*
 - above so as to cover or protect: *an oxygen tent over the pasta.*
 - extending above (an area) from a vantage point
- » 2. at a higher level or layer than: *his flat was over the ground floor.*
 - higher in rank than: *over him is the financial director.*
 - expressing authority or control: *editorial control over the press.*
 - expressing preference: *I'd choose the well-known over the unknown.*
 - expressing majority: *the predominance of Asia over Europe.*
 - higher in volume or pitch than: *he shouted over the noise.*
- » 3. higher or more than (a specified number or quantity): *she has been married for over a year.*
- » 4. expressing passage or trajectory across: *she topped the hill over the valley.*
 - beyond and falling or hanging from: *he toppled over the cliff.*
 - at the other side of; beyond: *over the hill is a small village.*
- » 5. expressing duration: *you've given us a lot of heat over the coffee.*
- » 6. expressing the medium by which something is done: *the speaker spoke over the loudspeaker.*
 - by means of the resistant surface constituted by: *the car rolled over the ground.*
- » 7. on the subject of: *a heated debate over unemployment.*

N = 4073



fine-grained &
lexeme-specific

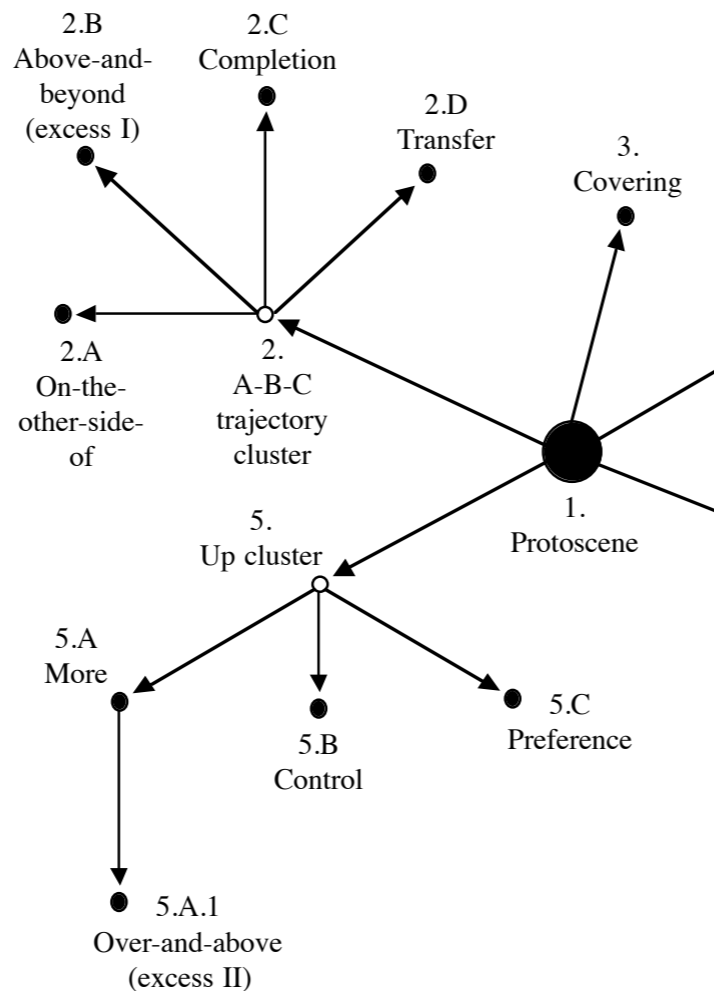
(extensive linguistic & AI research
on space & time)

Senses vs. Supersenses

over

• preposition

- » 1. extending directly upwards from: *I saw flames of heat.*
 - above so as to cover or protect: *an oxygen tent over the pasta.*
 - extending above (an area) from a vantage point
- » 2. at a higher level or layer than: *his flat was over the ground.*
 - higher in rank than: *over him is the financial director.*
 - expressing authority or control: *editorial control over the press.*
 - expressing preference: *I'd choose the well-known over the unknown.*
 - expressing majority: *the predominance of Asia over Europe.*
 - higher in volume or pitch than: *he shouted over the noise.*
- » 3. higher or more than (a specified number or quality): *she has been married for over a year.*
- » 4. expressing passage or trajectory across: *she trooped over the field.*
 - beyond and falling or hanging from: *he toppled over the wall.*
 - at the other side of; beyond: *over the hill is a small village.*
- » 5. expressing duration: *you've given us a lot of help over the years.*
- » 6. expressing the medium by which something is done: *he spoke over the loudspeaker.*
 - by means of the resistant surface constituted by: *she looked over her shoulder.*
- » 7. on the subject of: *a heated debate over unemployment.*



fine-grained details
lexeme-specific

cross-lexical classes; coarse;
interpretable names like TOPIC

(extensive linguistic & AI research
on space & time)

Preposition Supersenses

LOCATION

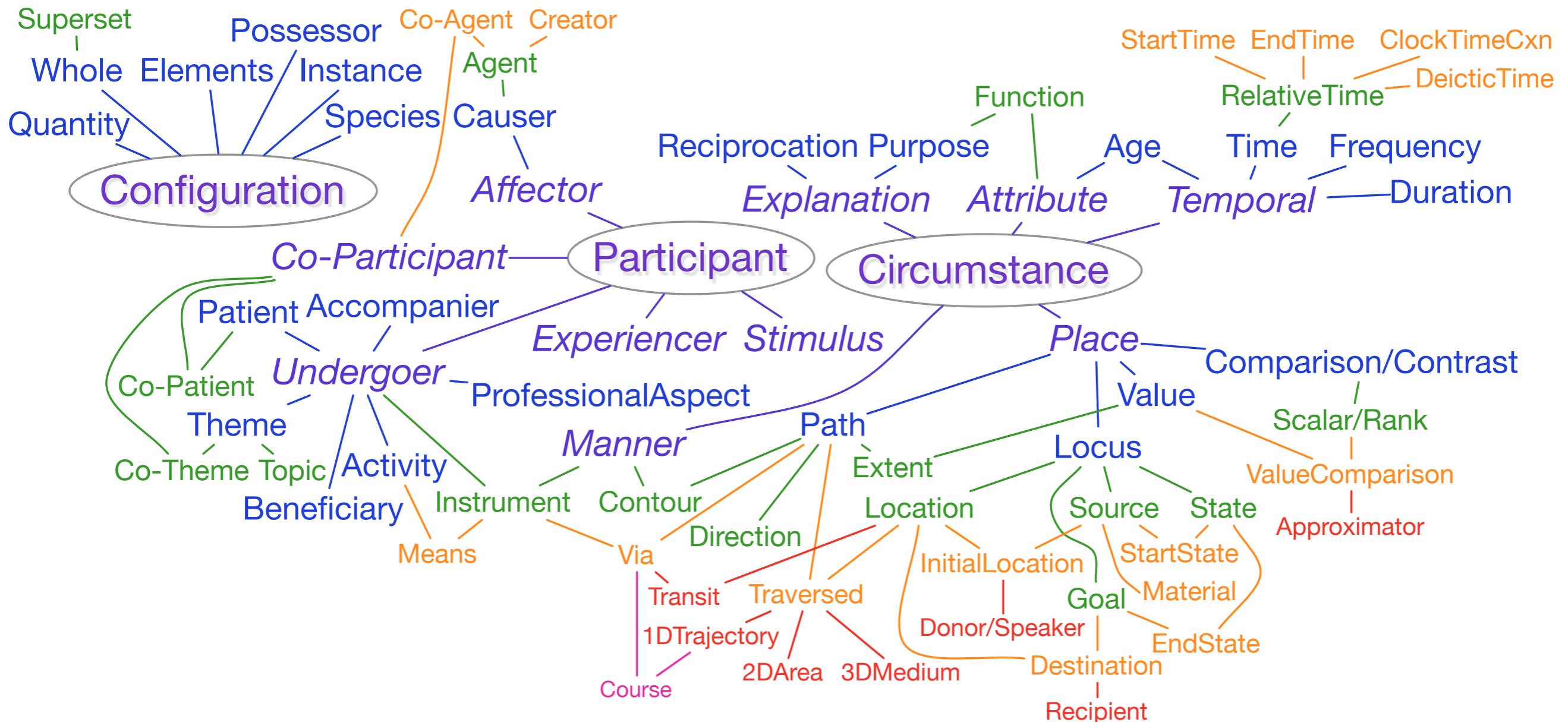
We met **in** Paris **at** a shop **on** a street **by** the Seine

TIME

at 6:00 **in** the evening **on** Saturday.

Supersense Hierarchy 1.0

[LAW 2015]



75 preposition supersense categories <http://tiny.cc/prepwiki>

English Annotation in STREUSLE

[LAW 2016]

- **Online reviews corpus** previously annotated for multiword expressions and noun & verb supersenses. 55,000 words, including 4,250 preps.
- **Comprehensive annotation:** first dataset with **all** prepositions (types+tokens) semantically annotated
 - ▶ Sentences not hand-selected
 - ▶ Sentences fully annotated
 - ▶ Preposition types not constrained by a lexicon (labels generalize)
 - ▶ All sentences seen by multiple annotators

Comparing resources

[LAW 2016]



| | | | | | | |
|----------------|----|-----|---|---|---|---|
| TPP | ✓ | (✓) | | ✓ | | The Preposition Project (Litkowski & Hargraves 2005, SemEval 2007 shared task) |
| D+ | 7 | | | ✓ | ✓ | TPP senses for 7 preposition types in PropBank WSJ data (Dahlmeier et al. 2009) |
| Tratz | 34 | (✓) | ✓ | ✓ | | Annotator-optimized revised senses for 34 TPP SemEval prepositions (Tratz 2011) |
| S&R | 34 | ✓ | | | | 32 hard clusters of TPP senses for 34 SemEval prepositions (Srikumar & Roth 2013) |
| Ours | ✓ | ✓ | ✓ | ✓ | ✓ | Preposition supersenses (Schneider et al. LAW 2015, 2016) |

A Vexing Problem

- Drawing clean boundaries between semantic categories is always difficult.
- But we were surprised by the frequency of apparent **overlaps** between semantic role labels.
- These overlaps proved pervasive in the other languages we looked at.



Destination/Location

- The prepositions **to**, **into**, **onto**, and **for** explicitly encode **DESTINATION**.
- **DESTINATION** masquerading as static **LOCATION**:
 - ▶ Put the pen **in** the box. (= *into*)
 - ▶ He threw his cards **on** the table. (= *onto*)
 - ▶ The ball rolled **behind** the trash can.
- Extremely productive for motion/caused motion!
- We could stipulate one or the other, but annotators would still get confused.

Fictive Motion

- In the other direction, we know that static locative relations can be described using dynamic language (Talmy 1996):
 - ▶ The road runs **through** the trees.
 - ▶ I heard him from the room next door.
 - ▶ The school is around the corner.
- In assigning a semantic label, is it sufficient to “choose sides” between the static nature of the spatial scene, and the dynamic way that relation is portrayed by the preposition?

Stimulus/Topic

- Another conundrum:
 - I thought **about** getting my ears pierced.: **TOPIC** (cf. *know, talk, read*)
 - I feared getting my ears pierced: **STIMULUS** (cf. *see, hurt*)
 - I was scared **about** getting my ears pierced: **???**
- Again, two labels are competing for semantic territory.
- Should we add more categories with double inheritance? (Problem: Proliferation of categories.)
- Should we just allow annotators to specify multiple labels if they're unsure? (Problem: Would create inconsistency.)

Construal

- Assumption thus far:
preposition token's semantics = role in a scene

...Topic

- ▶ I thought **about** getting my ears pierced.

Topic

- But it's not always so simple:

...Stimulus

- ▶ I was scared **about** getting my ears pierced.

Topic

Construal

- Observation: The preposition can frame or **construe** the situation in a way that differs from the predicate or scene.
- Solution: Allow tokens to receive **two labels** from the hierarchy, one for the **scene role** and one for the preposition's semantic **function**, when warranted.

Construal

- In fact, Stimulus can be interpreted differently by different prepositions:

...**Stimulus**

- ▶ I was scared **by** the bear.

Causer

...**Stimulus**

- ▶ I was scared **about** getting my ears pierced.

Topic

Experiencer Dative

- Experiencers can be realized as recipients/datives:

...**Experiencer**

- ▶ The bear felt scary **to** me.

Recipient

- In some languages, this is the main way **EXPERIENCERS** are realized:

- ▶ koev **li** ha-roš. [Hebrew]
Hurts **to.me** the-head ‘My head hurts.’

- ▶ mujh-**ko** garmii lag rahii hai. [Hindi]
I-**DAT** head feel PROG PRESS ‘I’m feeling hot.’



Employment

- The **PROFESSIONALASPECT** label is used for employer–employee and other professional relationships.
- It participates in several different preposition construals:

...**ProfAsp**

Beneficiary

▶ He works **for** XYZ Inc.
at

Location

...**ProfAsp**

Source

▶ He's **from** XYZ Inc.
with

Accompanier

Null Functions?

- Sometimes it's hard to tell whether the adposition has any semantic contribution:

...**Stimulus**

?

- ▶ I'm angry **with** my mom.
*mad

...**Topic**

?

- ▶ She's interested **in** politics.
*fascinated



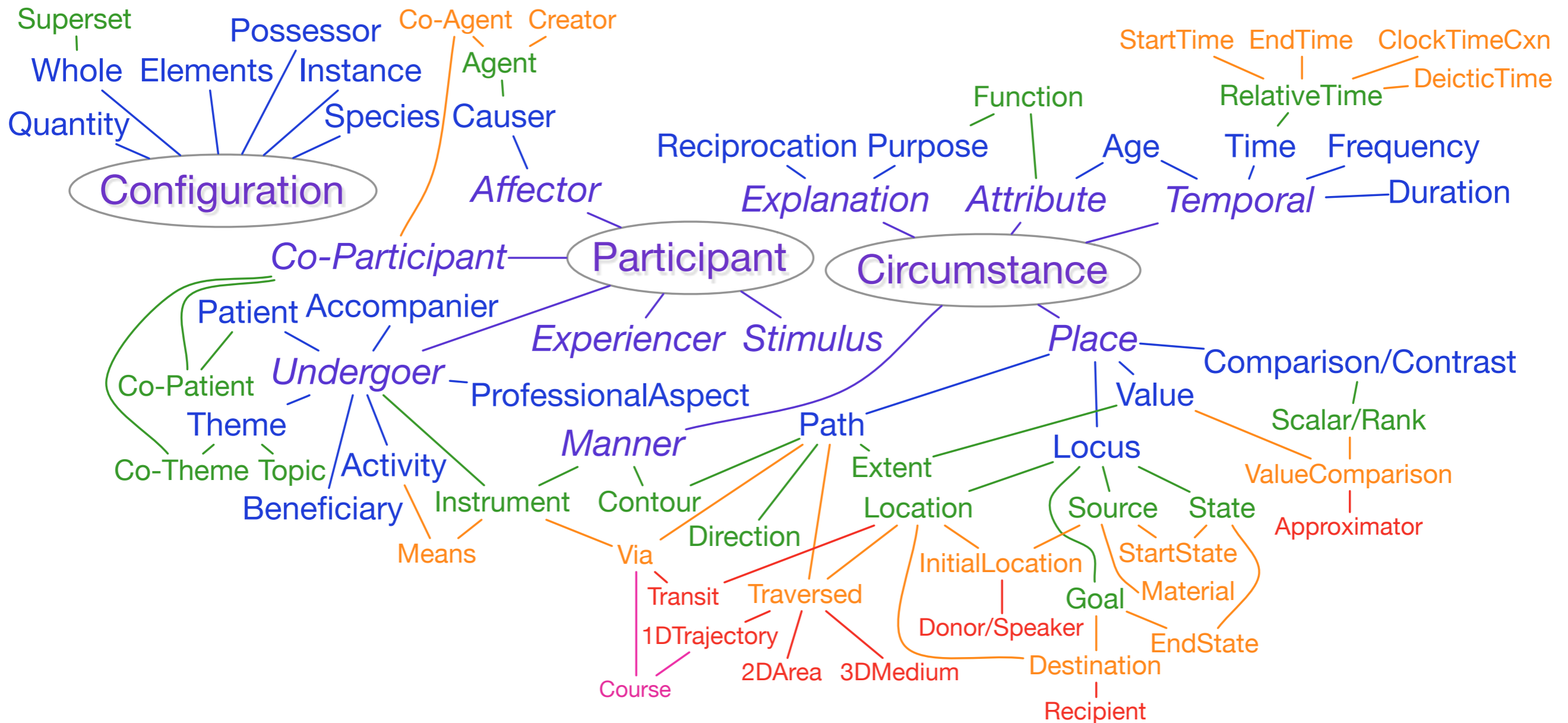
Postposition or Conjunction?

- The Korean marker **-wa** can have a comitative (**ACCOMPANIER**) meaning:
 - ▶ cheolsunun youngmi**wa** gilul geoleotta
‘Cheolsu walked the streets with Youngmi’
 - ▶ Cheolsunun youngmi**wa** chalul masyeotta
‘Cheolsu drank tea with Youngmi’
- But it can also mean ‘**and**’:
 - ▶ keopi**wa** chalul masija
‘Let’s drink coffee **and** tea’
- Our semantic inventory is limited to figure–ground relations. Would require labels for coordination semantics to cover **-wa** where it means ‘and’.

Ongoing & Future Work

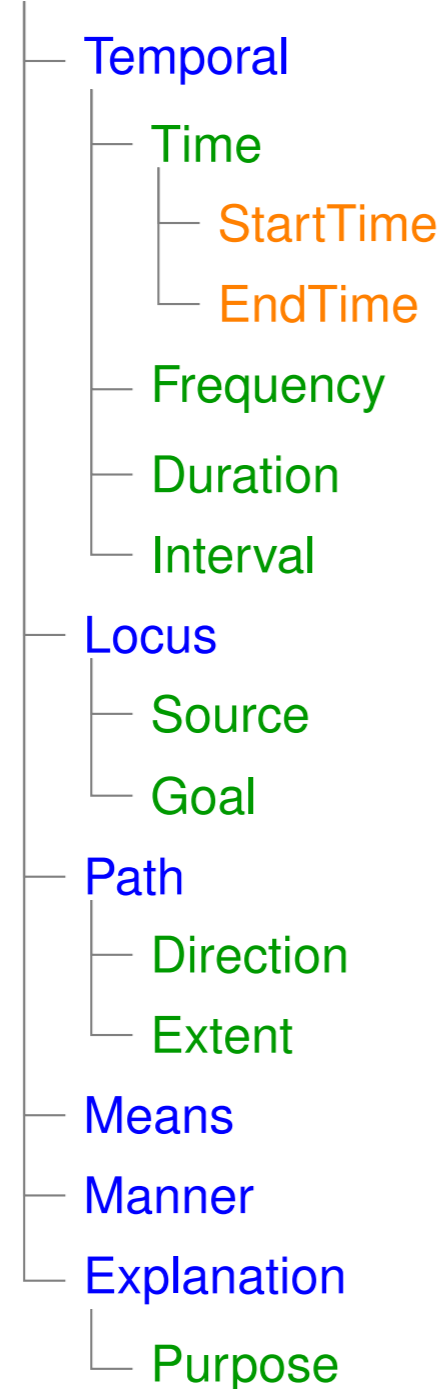
Hierarchy 1.0

[LAW 2015]



Hierarchy 2.0

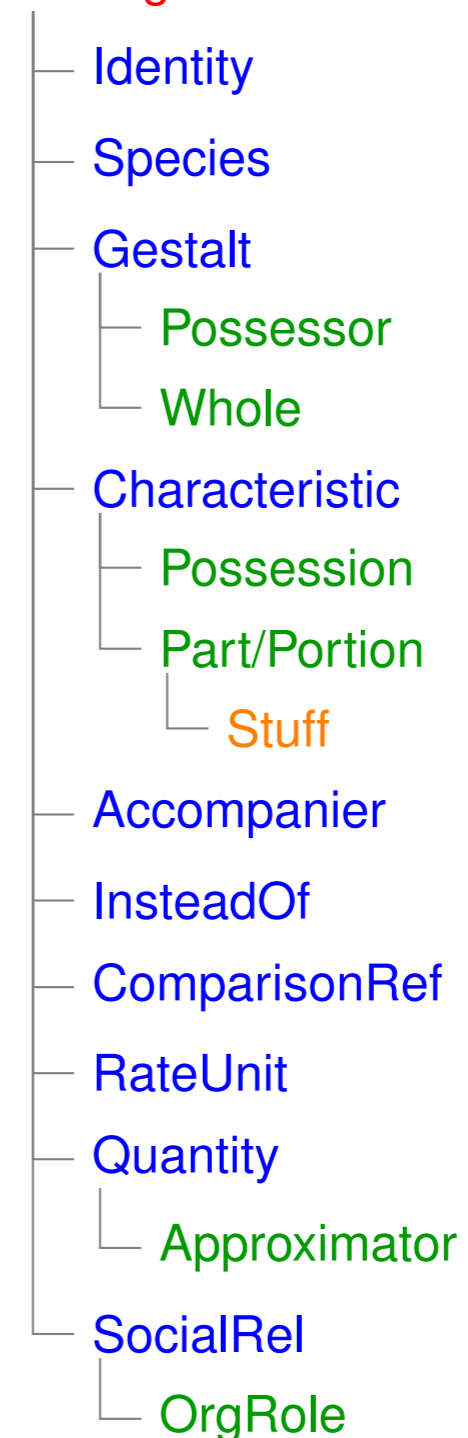
Circumstance



Participant



Configuration



Next Steps

- Annotation:
 - Updating the English reviews corpus
 - Monolingual Hebrew, Hindi, Korean data
 - Parallel data (*Little Prince*)
- Questions:
 - What construals are possible in what languages?
 - Can separating scene role from function better account for translation?
 - How well can the role and function be predicted automatically?



Thanks to

Martha Palmer, Ken Litkowski, Omri Abend, Katie Conger, Meredith Green, Michael Ellsworth, Paul Portner, Bill Croft

