CoreLex: Recap and Related Work

Paul Buitelaar
Insight Centre for Data Analytics
Data Science Institute, NUI Galway
Overview

some basics: word senses, systematic polysemy, metonymy, coercion

Generative Lexicon model of lexical semantics (Pustejovský, 1995)

CoreLex (Buitelaar, 1998)

related work & future directions
Some Basics
Dictionary definitions for the noun ‘ball’

ball¹

a spherical or approximately spherical body or shape. He rolled the piece of paper into a ball.

a round or roundish body, of various sizes and materials, either hollow or solid, for use in games, as baseball, football, tennis, or golf.

a game played with a ball, esp. baseball: The boys are out playing ball.

Military. a. a solid, usually spherical projectile for a cannon, rifle, pistol, etc., as distinguished from a shell.  b. projectiles, esp. bullets, collectively.

Horticulture. a compact mass of soil covering the roots of an uprooted tree or other plant.

Literary. a planetary or celestial body, esp. the earth.

Mathematics. (in a metric space) the set of points whose distance from the zero element is less than, or less than or equal to, a specified number.

ball²

a large, usually lavish, formal party featuring social dancing and sometimes given for a particular purpose, as to introduce debutantes or benefit a charitable organization.

Informal. a thoroughly good time: Have a ball on your vacation!

http://dictionary.reference.com/browse/ball
Word Senses for the noun ‘ball’

spherical body or shape

lavish, formal party featuring social dancing
Unrelated Word Senses

The **ball** went over the fence - ARTIFACT

The **ball** went on into the late hours - EVENT
Related Word Senses

The **ball** went over the fence - **ARTIFACT**

The boys are out playing **ball** - **EVENT**

![Soccer players on a field](image-url)
Regular Polysemy

unrelated senses (homonymy) vs. related senses (regular polysemy)

“Polysemy of the word A with the meanings \(a_i\) and \(a_j\) is called regular if, in the given language, there exists at least one other word B with the meanings \(b_i\) and \(b_j\), which are semantically distinguished from each other in exactly the same way as \(a_i\) and \(a_j\) and if \(a_i\) and \(b_i\), \(a_j\) and \(b_j\) are nonsynonymous.”

unrelated senses (homonymy) vs. related senses (regular polysemy)

“Polysemy of the word A with the meanings $a_i$ and $a_j$ is called regular if, in the given language, there exists at least one other word B with the meanings $b_i$ and $b_j$, which are semantically distinguished from each other in exactly the same way as $a_i$ and $a_j$ and if $a_i$ and $b_i$, $a_j$ and $b_j$ are nonsynonymous.”

Systematic Polysemy

“If more than two words share the same sense distribution (i.e. have the same set of senses), then this may indicate a pattern of systematic polysemy.”

(Buitelaar, 1998) - adapted from (Apresjan, 1974)

Referred to in the literature as:
‘regular polysemy’ (Apresjan, 1974)
‘logical polysemy’ (Pustejovsky, 1991; 1995)
‘systematic polysemy’ (Nunberg & Zaenen, 1992)
Systematic Polysemy – examples by (Bierwisch, 1983)


Die Schule, Die Universität, Das Theater, ...
...steht neben dem Sportplatz.
...macht einen Ausflug.
...langweilt ihn.
...ist aus der Geschichte Europas nicht wegzudenken.
Systematic Polysemy – other examples

The Boston office has been newly decorated
The Boston office was founded in 1985
The Boston office called

The cherry blossomed early this year
He likes a cherry every now and then
I like your cherry shirt

I always read the New York Times
I always buy the New York Times
The New York Times was sold last year
Systematic Polysemy and Metonymy

The Boston office called.

(A staff member of) the Boston office called.

‘staff part-of office’

more in general: ‘person part-of organization’

Metonymy: interpret a part as representing the whole
By "coercion" we mean the function or relation that maps from the explicit argument to the intended implicit argument in cases of metonymy, such as the mapping from "Shakespeare" to "the plays of Shakespeare" in the interpretation of the sentence "I like to read Shakespeare."


The Boston office called.

Semantic type of office (‘organization’) coerced into ‘person’ by verb to call
Coercion and Discourse Referents

The Boston office called. They signed a new contract.

Co-reference resolution between office (‘person’) and they (‘person’) metonymy, coercion, systematic polysemy of office
Peter bought a car. The engine runs well.

Bridging as coercive accommodation between a car and the engine
metonymy of engine (‘part-of car’)

Underspecified Discourse Referents

[A long book heavily weighted with military technicalities]_{NP: event-physical_object-content}, in this edition it is neither so long nor so technical as it was originally.

- a long book... takes long to read event
- heavily weighted... physical weight physical-object
- military technicalities... information is technical content
Generative Lexicon
“... there is a system of relations that characterizes the semantics of nominals very much like the argument structure of a verb ... Essentially the **qualia structure of a noun** determines its meaning as much as the list of arguments determines a verb’s meaning.”


**Infer an event from the lexical semantics of book** as represented by its Qualia Structure (Pustejovsky 1995)

*I began the book*

*I began (reading) the book*
Qualia Structure

The qualia structure, inspired by Moravcsik’s (1975) interpretation of the *aitia* (cause) of Aristotle. 

The modes of explanation associated with a word or phrase in the language:

- **Formal**: the basic category of which distinguishes the meaning of a word within a larger domain.
- **Constitutive**: the relation between an object and its constituent parts.
- **Telic**: the purpose or function of the object, if there is one.
- **Agentive**: the factors involved in the object’s origins or “coming into being”.

From slides by James Pustejovsky
Qualia Structure

<table>
<thead>
<tr>
<th>Type</th>
<th>Definition</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td>Formal</td>
<td>is-a, hyponymy</td>
<td>physical-object, content</td>
</tr>
<tr>
<td>Constitutive</td>
<td>part-of, meronymy</td>
<td>section, ...</td>
</tr>
<tr>
<td>Telic</td>
<td>what is the object used for</td>
<td>read, ...</td>
</tr>
<tr>
<td>Agentive</td>
<td>how did the object originate</td>
<td>write, ...</td>
</tr>
</tbody>
</table>
Qualia Structure – Lexical Conceptual Paradigm

**book**

ARGSTR =

ARG1 = \( x:information \)

ARG2 = \( y:phys\_obj \)

QUALIA =

information.phys_obj_lcp

“dotted type”

FORMAL: \( \text{hold}(y,x) \)

TELIC: \( \text{read}(e,w,x.y) \)

AGENTIVE: \( \text{write}(e',v,x.y) \)
Qualia Structure as Knowledge Graph – Formal

phys-obj

communication

book

isa
Qualia Structure as KG – Constitutive
Qualia Structure as KG – Agentive/Telic

- event
- reading
- writing
- printing
- phys-obj
- communication
- book

is_a:
- reading isa event
- writing isa reading
- book isa phys-obj

hasProperties:
- reading hasCreationProcess book
- printing hasProductionProcess book
- communication hasProductionProcess book

is_used_for:
- reading isUsedFor book
They printed some very interesting books.
They printed some very interesting books.
Qualia Structure as Underspecified Semantic Type

physical-object-communication-event

book

cover

section

Formal

Constitutive

Telic

Agentive

„book“

write

read
Qualia Structure – Type Coercion, Metonymy

_I began the book_

_I began (reading) the book_

_Type coercion:_ direct-object of _begin_ requires an ‘event’

_Metonymy:_ _book_ (‘physical-object’, ‘communication’) part-of _reading_ (‘event’)

Towards CoreLex

Aggregate all types that can be reached through the Knowledge Graph into Underspecified Semantic Type

CoreLex approach based on systematic analysis of sense distribution across WordNet nouns
CoreLex – Qualia Structure acquisition

CoreLex is an attempt to **automatically acquire underspecified lexical semantic representations** that reflect systematic polysemy. 

Underspecified lexical semantic representations (semantic types) can be viewed as **shallow Qualia Structures**.
CoreLex – Systematic Polysemy

Systematic polysemy can be empirically studied in WordNet by observing sense distributions.

*If more than two words share the same sense distribution (i.e. have the same set of senses), then this may indicate a pattern of systematic polysemy.*
“... semantic tagging should be a first step in the interpretation process by assigning each lexical item a representation of all of its systematically related 'senses'. Further semantic processing steps derive discourse dependent interpretations from this representation. Semantic tags are therefore more like pointers to complex knowledge representations, which can be seen as underspecified lexical meanings.”


[A long book heavily weighted with military technicalities]NP:event-physical_object-content , in this edition it is neither so long nor so technical as it was originally.
CoreLex overview

- Systematic Polysemy
- Motivation
- Theory
- Generative Lexicon
- CoreLex
- Data
- Corpora
- Underspecified
- Semantic Tagging
- Application
Semantic Lexicon organized by concept/meaning/sense
Lexical Database: machine readable
Freely available from http://wordnet.princeton.edu/
“... initial idea was to provide an aid to use in searching dictionaries conceptually, rather than merely alphabetically ... WordNet instantiates hypotheses based on results of psycholinguistic research ... In anomic aphasia, there is a specific inability to name objects. When confronted with an apple, say, patients may be unable to utter “apple,” even though they will reject such suggestions as shoe or banana, and will recognize that apple is correct when it is provided.”

WordNet - synsets

Word meaning expressed in WordNet by use of sets of synonyms (synset)
synset \{board, plank\} represents board in the meaning of plank
synset \{board, committee\} represents board in the meaning of committee
WordNet – synsets for table

WordNet Search - 3.1
- WordNet home page - Glossary - Help

Word to search for: table

Display Options: (Select option to change) ▼ Change

Key: "S." = Show Synset (semantic) relations, "W." = Show Word (lexical) relations

Display options for sense: (gloss) "an example sentence"

Noun

- S: (n) table, tabular array (a set of data arranged in rows and columns) "see table 1"
- S: (n) table (a piece of furniture having a smooth flat top that is usually supported by one or more vertical legs) "it was a sturdy table"
- S: (n) table (a piece of furniture with tableware for a meal laid out on it) "I reserved a table at my favorite restaurant"
- S: (n) mesa, table (flat tableland with steep edges) "the tribe was relatively safe on the mesa but they had to descend into the valley for water"
- S: (n) table (a company of people assembled at a table for a meal or game) "he entertained the whole table with his witty remarks"
- S: (n) board, table (food or meals in general) "she sets a fine table", "room and board"

Verb

- S: (v) postpone, prorogue, hold over, put over, table, shelf, set back, defer, remit, put off (hold back to a later time) "let's postpone the exam"
- S: (v) table, tabularize, tabularise, tabulate (arrange or enter in tabular form)
Synsets are organized hierarchically: generalization/hypernym, specialization/hyponym

{entity}
...
{whole, unit}
{building material}
{lumber, timber}
{board, plank}

hypernym

hyponym
WordNet – hierarchy example chemical

- **S. (n) chemical, chemical substance** (produced by or used in a reaction involving changes in atoms or molecules)
  - *direct hyponym / full hyponym*
  - *direct hypernym / inherited hypernym / sister term*
- **S. (n) material, stuff** (the tangible substance that goes into the makeup of a physical object) "coal is a hard black material"; "wheat is the stuff they use to make bread"
  - *direct hyponym / full hyponym*
  - *direct hypernym / inherited hypernym / sister term*
- **S. (n) substance, matter** (that which has mass and occupies space) "an atom is the smallest indivisible unit of matter"
  - *direct hyponym / full hyponym*
  - *direct hypernym / inherited hypernym / sister term*
- **S. (n) physical entity** (an entity that has physical existence)
  - *direct hyponym / full hyponym*
  - *direct hypernym / inherited hypernym / sister term*
- **S. (n) entity** (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
  - *direct hyponym / full hyponym*
WordNet – hierarchy example *food*

- **S**: (n) *food*, *nutrient* (any substance that can be metabolized by an organism to give energy and build tissue)
  - *direct hyponym / full hyponym*
  - *part meronym*
  - *direct hypernym / inherited hypernym / sister term*
    - **S**: (n) *substance*, *matter* (that which has mass and occupies space) "an atom is the smallest indivisible unit of matter"
      - *direct hyponym / full hyponym*
      - *direct hypernym / inherited hypernym / sister term*
    - **S**: (n) *physical entity* (an entity that has physical existence)
      - *direct hyponym / full hyponym*
      - *direct hypernym / inherited hypernym / sister term*
    - **S**: (n) *entity* (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
      - *direct hyponym / full hyponym*
## WordNet – statistics (version 3.1)

### Number of words, synsets, and senses

<table>
<thead>
<tr>
<th>POS</th>
<th>Unique Synsets</th>
<th>Total Strings</th>
<th>Total Word-Sense Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>117798</td>
<td>82115</td>
<td>146312</td>
</tr>
<tr>
<td>Verb</td>
<td>11529</td>
<td>13767</td>
<td>25047</td>
</tr>
<tr>
<td>Adjective</td>
<td>21479</td>
<td>18156</td>
<td>30002</td>
</tr>
<tr>
<td>Adverb</td>
<td>4481</td>
<td>3621</td>
<td>5580</td>
</tr>
<tr>
<td>Totals</td>
<td>155287</td>
<td>117659</td>
<td>206941</td>
</tr>
</tbody>
</table>

### Polysemy information

<table>
<thead>
<tr>
<th>POS</th>
<th>Monoseymous Words and Senses</th>
<th>Polyseymous Words</th>
<th>Polyseymous Senses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>101863</td>
<td>15935</td>
<td>44449</td>
</tr>
<tr>
<td>Verb</td>
<td>6277</td>
<td>5252</td>
<td>18770</td>
</tr>
<tr>
<td>Adjective</td>
<td>16503</td>
<td>4976</td>
<td>14399</td>
</tr>
<tr>
<td>Adverb</td>
<td>3748</td>
<td>733</td>
<td>1832</td>
</tr>
<tr>
<td>Totals</td>
<td>128391</td>
<td>26896</td>
<td>79450</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>POS</th>
<th>Average Polysemy Including Monoseymous Words</th>
<th>Average Polysemy Excluding Monoseymous Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>1.24</td>
<td>2.79</td>
</tr>
<tr>
<td>Verb</td>
<td>2.17</td>
<td>3.57</td>
</tr>
<tr>
<td>Adjective</td>
<td>1.40</td>
<td>2.71</td>
</tr>
<tr>
<td>Adverb</td>
<td>1.23</td>
<td>2.50</td>
</tr>
</tbody>
</table>
WordNet to CoreLex steps

Noun_1

Basic Type_1

Systematic Polysemous Class_1

Noun_n

Basic Type_n

Systematic Polysemous Class_n
Basic Types

obs 00012670 abstraction 0
act 00016649 act 0 human_action 0 human_activity 0
gst 00004473 causal_agent 0 cause 0 causal_agency 0
anm 00006850 animal 0 animate_being 0 beast 0 brute 0 creature 0 fauna 0
art 00012607 artifact 0 artefact 0
atr 00017556 attribute 0
cel 00003711 cell 0
chem 085017331 compound 0 chemical_compound 0
chem 08505286 chemical_element 0 element 0
com 000018599 communication 0
con 05405491 consequence 0 effect 0 outcome 0 result 0 upshot 0
ent 00002400 entity 0
ev 00016459 event 0
fod 00011263 food 0 nutrient 0
frm 00014556 shape 0 form 0
grp 05315837 biological_group 0
grp 00017008 group 0 grouping 0
gsp 05119847 social_group 0
grs 05116476 people 0
hum 00004856 person 0 individual 0 someone 0 mortal 0 human 0 soul 0
lfr 00000728 life_form 0 organism 0 being 0 living_thing 0
lme 08322690 linear_measure 0 long_measure 0
loc 00014314 location 0
log 0540515 location 1
meas 00019060 measure 0 quantity 0 amount 0 quantum 0
mic 097400781 microorganism 0
nnt 00009912 natural_object 0
nat 05715416 body_of_water 0 water 0
net 05720252 lend 0 dry_lend 0 earth 0 ground 0 solid_ground 0 terra_firma 0
pit 00008894 plant 0 flora 0 plant_life 0
phn 00019295 phenomenon 0
phys 00004650 object 0 inanimate_object 0 physical_object 0
pos 00017394 possession 0
pro 08239098 process 0
prt 05654477 part 0 piece 0
psy 00012517 psychological_feature 0
qud 08310215 definite_quantity 0
qui 08310433 indefinite_quantity 0
rel 00017852 relation 0
spc 00015245 space 0
sta 00015437 state 0
sub 00010368 substance 0 matter 0
tme 00055837 time_period 0 period 0 period_of_time 0 amount_of_time 0
tme 00022094 time_unit 0 unit_of_time 0
tme 00014882 time 0
## Systematic Polysemous Classes

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>acp</td>
<td>act attribute process psychological-feature state</td>
</tr>
<tr>
<td>acr</td>
<td>act attribute event relation state</td>
</tr>
<tr>
<td>acs</td>
<td>act state</td>
</tr>
<tr>
<td>aes</td>
<td>act event state</td>
</tr>
<tr>
<td>aev</td>
<td>act event</td>
</tr>
<tr>
<td>age</td>
<td>act causal agent</td>
</tr>
<tr>
<td>agh</td>
<td>causal agent human</td>
</tr>
<tr>
<td>agl</td>
<td>causal agent location</td>
</tr>
<tr>
<td>age</td>
<td>causal agent animal</td>
</tr>
<tr>
<td>agp</td>
<td>causal agent psychological-feature</td>
</tr>
<tr>
<td>agt</td>
<td>causal agent</td>
</tr>
<tr>
<td>anf</td>
<td>animal food</td>
</tr>
<tr>
<td>ann</td>
<td>animal artifact natural-object</td>
</tr>
<tr>
<td>anp</td>
<td>animal psychological-feature</td>
</tr>
<tr>
<td>aqu</td>
<td>artifact quantity-definite quantity-indefinite</td>
</tr>
<tr>
<td>ara</td>
<td>artifact attribute psychological-feature state</td>
</tr>
<tr>
<td>arg</td>
<td>artifact group</td>
</tr>
<tr>
<td>arh</td>
<td>artifact human</td>
</tr>
<tr>
<td>arp</td>
<td>artifact psychological-feature state</td>
</tr>
<tr>
<td>art</td>
<td>artifact state</td>
</tr>
<tr>
<td>atc</td>
<td>attribute communication phenomenon psychological-feature state</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>
From Synsets to Basic Types

book 1. {publication} => artifact
2. {product, production} => artifact
3. {fact} => communication
4. {dramatic_composition, dramatic_work} => communication
5. {record} => communication
6. {section, subdivision} => communication
7. {journal} => artifact

book : ‘artifact-communication’
Basic Types to Systematic Polysemous Classes

‘artifact-communication’

amulet annals armband arrow article ballad bauble beacon bible birdcall blank blinker boilerplate book … catalog catalogue chart chevron clout compact compendium convertible copperplate copy cordon corker … guillotine homophony horoscope indicator journal laurels lay ledger loophole marker memorial nonsense novel … pamphlet pastoral paternoster pedal pennant phrase platform portrait prescription print puzzle radiogram rasp recap riddle rondeau … statement stave stripe talisman taw text … transcription trophy trumpery … wrapper yardstick
Systematic Polysemous Classes – other examples

animal-natural_object
alligator broadtail chamois ermine leopard muskrat ...

natural_object-plant
algarroba almond anise baneberry butternut candlenut...

action-artifact-group_social
artillery assembly band dance gathering institution ...
Systematic Polysemous Classes – other examples

action-attribute-event-psychological
appearance decision deviation impulse outrage …

possession-quantity_definite
cent centime dividend gross penny real shilling
Non-Systematic Classes – example

action-animal-artifact

bat drill fly hobby ruff solitaire spat
Partly-Systematic Classes – example

**action-geographical_location**
bolivia caliphate charleston chicago clearing emirate michigan prefecture repair ...

**systematic polysemy**
clearing, repair, wheeling

**homonym**
bolivia, charleston, chicago, michigan

caliphate, emirate, prefecture
CoreLex vs. WordNet

CoreLex

WordNet

'animal-artifact-natural-object'

'act-state-substance'
CoreLex as Semantic Lexicon

CoreLex is available from
http://www.cs.brandeis.edu/~paulb/CoreLex/corelex.html

course-grained semantic lexicon (~40,000 nouns, 126 semantic types)
underspecified semantic tagging (126 underspecified semantic types vs.
~60,000 synset based senses)
related work & future directions
Related work – Peters and Peters, 2000

<table>
<thead>
<tr>
<th>Relation</th>
<th>Examples</th>
<th>No of words covered</th>
</tr>
</thead>
<tbody>
<tr>
<td>holdfast – control</td>
<td>lock, clasp, clench</td>
<td>5</td>
</tr>
<tr>
<td>picture – painting</td>
<td>etching, engraving, fresco</td>
<td>5</td>
</tr>
<tr>
<td>music – dance</td>
<td>waltz, rumba, bolero</td>
<td>20</td>
</tr>
<tr>
<td>communication system – broadcast</td>
<td>radio, television, wireless</td>
<td>3</td>
</tr>
<tr>
<td>path – travelling</td>
<td>crossing, walk, promenade</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1: UB pair Artefact - Action

Related work – Tomuro, 2000

Figure 3: Parts of WordNet trees ARTIFACT and MEASURE

Table 3: Examples of Automatically Extracted Systematic Polysemy

<table>
<thead>
<tr>
<th>Top relation class</th>
<th>Relation</th>
<th>Common Words</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACTION-LOCATION</td>
<td>[ACTION, POINT]</td>
<td>“drop”, “circle”, “intersection”, “dig”, “crossing”, “bull’s eye”</td>
</tr>
<tr>
<td>ARTIFACT-GROUP</td>
<td>STRUCTURE, PEOPLE</td>
<td>“house”, “convent”, “market”, “center”</td>
</tr>
<tr>
<td>ARTIFACT-SUBSTANCE</td>
<td>FABRIC, CHEMICAL_COMPOUND</td>
<td>“acetate”, “nylon”, “acrylic”, “polyester”</td>
</tr>
<tr>
<td>COMMUNICATION-PERSON</td>
<td>VOICE, SINGER</td>
<td>“soprano”, “alto”, “tenor”, “baritone”</td>
</tr>
</tbody>
</table>
Related work – Navigli, 2006; Snow et al., 2007

Figure 1: Sense clusters for the noun *bass*; the eight WordNet senses as clustered into four groups in the SESEVAL-2 coarse-grained evaluation data

Related work – Boleda et al., 2012

Table 5: Meta alternations and their average precision values for the task. The random baseline performs at 0.313 while the frequency baseline ranges from 0.255 to 0.369 with a mean of 0.291. Alternations for which the model outperforms the frequency baseline are in boldface (mean AP: 0.399, standard deviation: 0.119).

<table>
<thead>
<tr>
<th>Meta Alternation</th>
<th>AP</th>
</tr>
</thead>
<tbody>
<tr>
<td>grs-psy</td>
<td>0.709</td>
</tr>
<tr>
<td>pro-sta</td>
<td>0.678</td>
</tr>
<tr>
<td>fod-plt</td>
<td>0.645</td>
</tr>
<tr>
<td>psy-sta</td>
<td>0.630</td>
</tr>
<tr>
<td>hum-prt</td>
<td>0.602</td>
</tr>
<tr>
<td>grp-psg</td>
<td>0.574</td>
</tr>
<tr>
<td>grs-log</td>
<td>0.573</td>
</tr>
<tr>
<td>act-evt</td>
<td>0.539</td>
</tr>
<tr>
<td>evt-psy</td>
<td>0.526</td>
</tr>
<tr>
<td>act-tme</td>
<td>0.523</td>
</tr>
<tr>
<td>art-pho</td>
<td>0.520</td>
</tr>
<tr>
<td>act-pro</td>
<td>0.513</td>
</tr>
</tbody>
</table>

Table 6: Sample targets for meta alternations with high AP and mid-coherence values.

- **grs-psy**: democracy, faculty, humanism, regime
- **pro-sta**: bondage, dehydration, erosion, urbanization
- **psy-sta**: anaemia, delight, pathology, sensibility
- **hum-prt**: bun, contractor, peter, subordinate
- **grp-psg**: category, collectivism, socialism, underworld

Figure 2: Average Precision and Coherence ($\kappa$) for each meta alternation. Correlation: $r = 0.743$ ($p < 0.001$)
Related work – Alonso et al., 2013

What is the selected sense for the next example of the word Spain?

_The weather in Spain is generally very hot in the summer._

- Location (the place)
- Organization (institutions, people, etc.)
- Both Location and Organization

<table>
<thead>
<tr>
<th>Dot type</th>
<th>$\bar{A}_o \pm \sigma$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>eng:animeat</td>
<td>0.86 ± 0.24</td>
<td>0.69</td>
</tr>
<tr>
<td>eng:artinfo</td>
<td>0.48 ± 0.23</td>
<td>0.12</td>
</tr>
<tr>
<td>eng:contcont</td>
<td>0.65 ± 0.28</td>
<td>0.31</td>
</tr>
<tr>
<td>eng:locorg</td>
<td>0.72 ± 0.29</td>
<td>0.46</td>
</tr>
<tr>
<td>eng:prores</td>
<td>0.5 ± 0.24</td>
<td>0.10</td>
</tr>
<tr>
<td>da:contcont</td>
<td>0.32 ± 0.37</td>
<td>0.39</td>
</tr>
<tr>
<td>da:locorg</td>
<td>0.73 ± 0.37</td>
<td>0.47</td>
</tr>
<tr>
<td>spa:contcont</td>
<td>0.36 ± 0.3</td>
<td>0.42</td>
</tr>
<tr>
<td>spa:locorg</td>
<td>0.52 ± 0.28</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Table 2: Averaged observed agreement and its standard deviation and alpha
Future Directions

sense embeddings


Named Entities


systematic polysemy in domains (take a corner)


multilingual synsets / senses (BabelNet)

metaphor

discourse analysis (underspecified discourse referents)
Thank you!