Modelling semantic property acquisition from single linguistic exposures

Marco Baroni*, Alessandro Lenci†, Brian Murphy*, Massimo Poesio*

CIMeC (University of Trento)*, Department of Linguistics (University of Pisa)†

23rd July 2008

- People can learn concept properties from a **single linguistic exposure**
- They must recognise properties as **interesting** or not; and also guess the **concept-property relation**
- We hypothesise that much of the necessary information is accessible in the **immediate textual context**:
  - the superordinate category of the concept
  - the (potential) property lexeme
  - the connecting text pattern
- Two computational challenges
  - can **relevant** properties be distinguished from irrelevant ones?
  - can the property **type** (relation to the concept) be determined?

- *Assuming* speaker has **substantive prior knowledge** of previously learned concepts, we investigate the following questions:
  - What **kinds of prior knowledge** are most informative for property learning?
  - Is property learning easier/harder for **certain categories** of concepts?

Evidence

- Properties: *cat has tail*
  - Has the potential property already been seen in previous concepts? With which type(s)?
  - We use a variation on the McRae/Wu & Barsalou typology: category, component, function, action, behaviour, contextual setting, associated entity, perceptually simple quality, perceptually complex quality
- Patterns: *cat has tail*
  - Has the pattern already been seen with previous concepts-property pairs? With which type(s)?
  - *StruDEL* parser (Baroni & Lenci 2008) to filter/generalize patterns

- Type distributions
  - Does the relative frequency of type relations help (e.g., *component* relations are much more frequent than *category* relations)?
- Category-conditioned distributions
  - Does the category of the concept help (e.g., *function* is more common type for tools than animals)?

The probabilistic model

- Given a single text instance in which a concept occurs with property x and connecting pattern y, we assign type t according to the following formula:

$$t = \begin{cases} \underset{t}{\operatorname{argmax}}\, P(T = t|X = x, Y = y) & \text{if } > r \\ \mathit{irrelevant} & \text{otherwise} \end{cases}$$

- Assuming conditional independence of x and y given t, posterior probability of t is given by:

$$P(T = t|X = x, Y = y) = \frac{P(X = x|T = t)P(Y = y|T = t)P(T = t)}{\sum_{u \in T} P(X = x|T = u)P(Y = y|T = u)P(T = u)}$$

- Estimation by simple maximum likelihood, with add-1 smoothing (fully incremental implementation should be straightforward)

- Different assumptions about available prior evidence implemented in following variations of model:
  - **property**: uniform $P(t)$, uniform $P(y|t)$
  - **pattern**: uniform $P(t)$, uniform $P(x|t)$
  - **property, pattern**: uniform $P(t)$
  - **property, pattern, type**: $P(t)$ estimated from distribution in prior set
  - **property, pattern, category**: property and pattern terms are estimated separately for concepts belonging to different categories
  - **property, pattern, type, category**: as above with category-conditioned estimation of types
  - **baseline**: properties are assigned most likely type for category prior set (*function* for tools, *component* for mammals) – this is not a trivial baseline!

Prior knowledge

- Priors estimated from properties and patterns associated with 127 common concrete concepts (animals, tools. . .), all attested in corpus of child-directed speech
- Type-conditioned property distribution and type distribution estimated from informant responses in concept description norms of McRae et al. 2005
- Type-conditioned pattern distribution estimated from large Web-derived corpus (Ferraresi et al., 2008), extracting (with StruDEL parser) patterns connecting selected concepts and their norms-attested properties

Test set

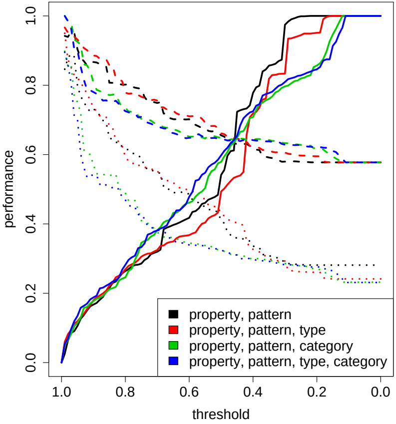
- 339 unseen test concepts from same categories of prior set, sampled from 3 frequency ranges (low: *bulbul*, medium: *cormorant*, high: *dove*)
- For each test concept, about 10 tuples concept + pattern + (potential) property randomly extracted from corpus and tested as single exposures
- Gold standard created by manual annotation of test tuples for relation type

Evaluation procedure

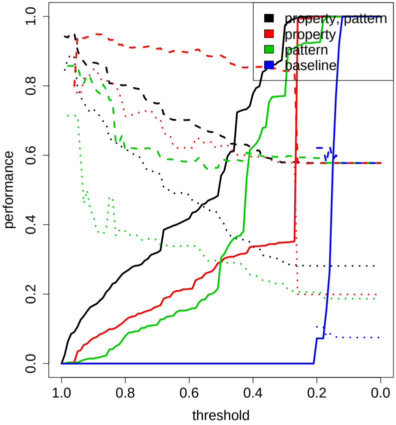
- Type labels assigned by rule: pick most likely type, or *irrelevant* if probability below threshold r
- We evaluate by letting threshold r range across possible p-values (from 1 to 0), and computing performance measures at each
- Performance measures:

recall	proportion of relevant properties recognised at $p > r$ to true total relevant properties: **continuous lines**
r-precision	proportion of *relevant* properties among those with probability $p > r$: **dashed lines**
c-precision	proportion of relevant properties with *correct* type among those with $p > r$: **dotted lines**

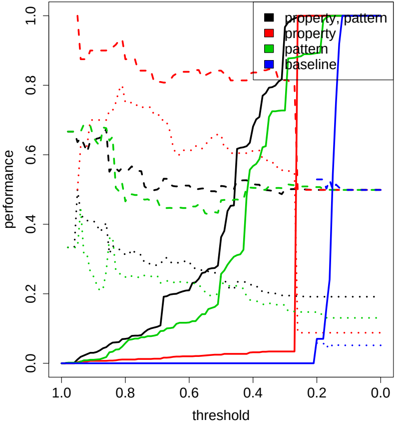
Does knowledge about concept categories, and type distribution help?



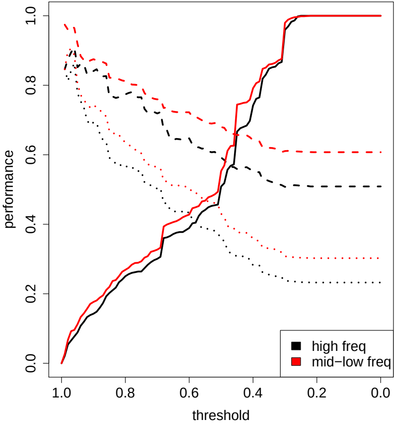
What is the relative contribution of pattern and property knowledge?



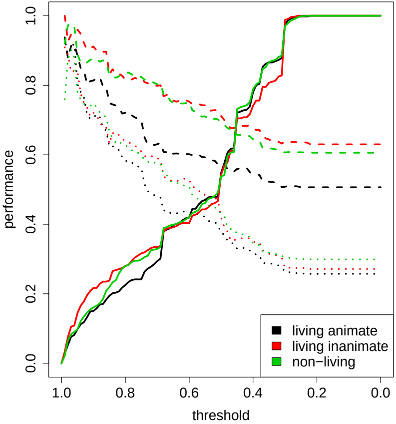
How does the model generalise to novel property lexemes?



Performance on lower frequency concepts?



Performance across categories?



Conclusions, Further Work

- Pattern and property contribute independent information, giving considerably higher performance than a non-trivial baseline
- Category specific knowledge and relative frequency of types do not improve performance
- Text describing low frequency words seems more informative than that for high frequency
- Further analysis to do on effect of property type and category

References

P. Bloom. 2000. *How children learn the meanings of words*. Cambridge: The MIT Press.

A. Ferraresi, E. Zanchetta, M. Baroni & S. Bernardini. 2008. Introducing and evaluating ukWaC, a very large Web-derived corpus of English. In *Proceedings of the WAC4 Workshop*.

M. Baroni & A. Lenci. 2008. Concepts and properties in word space. *Italian Journal of Linguistics*.

K. McRae, G. Cree, M. Seidenberg & C. McNorgan. 2005. Semantic feature production norms for a large set of living and nonliving things. *Behavior Research Methods* **37**: 547-559.

G. Murphy. 2002. *The big book of concepts.* The MIT Press.

M. Poesio and A. Almuhareb. 2008. Extracting concept descriptions from the Web: The importance of attributes and values. To appear in Buitelaar and Cimiano (eds.) *Bridging the Gap between Text and Knowledge*.

G. Vigliocco, D. Vinson, W. Lewis, and M. Garrett. 2004. Representing the meanings of object and action words: The featural and unitary semantic space hypothesis. *Cognitive Psychology* **48**: 422-488.

Relation types

category: *horse* is a *mammal*

component: *cat* has *a tail*

function: *hammer* is for *pounding*

action: *brush* is used by *rotating*

behaviour: *mosquito* engages in *biting*

contextual setting: *dog* is in *kennel*

associated entity: *beret* is associated with *officer*

perceptually simple quality: *banana* is *yellow*

perceptually complex quality: *dress* is considered *elegant*

out: (all adjectives, verbs and nouns occurring in the near of target concept but not denoting a property)