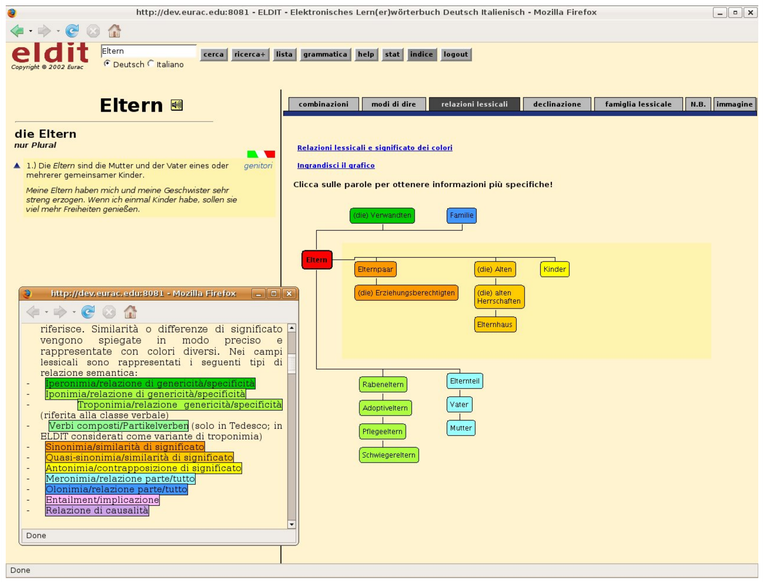
Cognitively Salient Semantic Relations in and for Concept Descriptions

Gerhard Kremer
gerhard.kremer@unitn.it
CIMeC, University of Trento, Italy

Introduction

Background:
An electronic learner's dictionary for German and Italian, which visualises (and links) semantically related words



- Covers small number of entries
- Relation types and entries are defined manually, with no empirical cognitive basis

Which are the most salient semantic relations in human minds?

Production Experiment

Elicitation of concept descriptions
(in line with property generation, semantic norm collection)

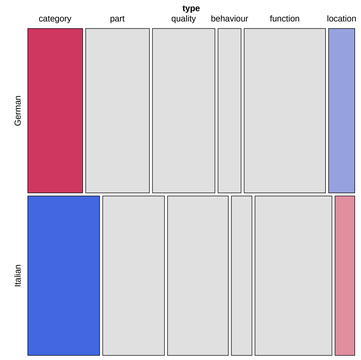
Stimuli:
5 concepts per class (mammal, bird, fruit, vegetable, body part, clothing, tool, vehicle, furniture, and building)

Participants:
73 German and 69 Italian native speakers
(students at high–schools at the average age of 17)

Data:
Phrases are split, transcribed in English, and normalised; then mapped to relation types

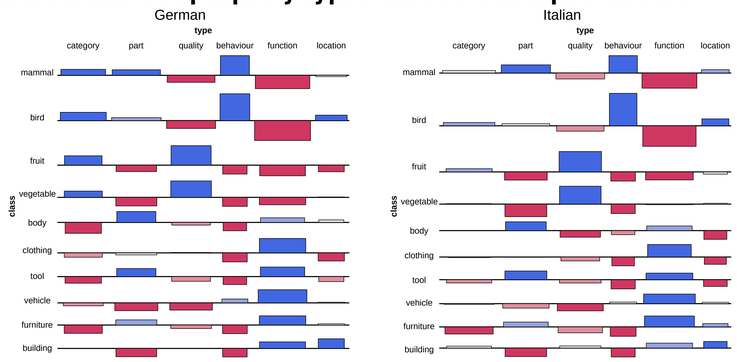
Results

Distributions of the six most common property types across languages:



- Possible reasons for sign. differences concerning category: usage of *oggetto* vs. *Sache*
- Salient property types were cross–linguistically robust

Distribution of property types across concept classes:



- Property types are highly distinctive characteristics of classes
- Cross–language similarities

Unsupervised clustering by these property types:

Best results for 3–way partition

In German:
- animals (+ head, pineapple)
- plants (+ skyscraper, tower)
- objects and body parts

In Italian:
- animals (+ head, pineapple, chemise)
- plants (+ truck)
- objects and body parts

Future Work

- Investigate distributions across less frequent property types
- Define granularity of classes
- Abstract objects
- Perception Experiment:
  Do humans also perceive the found relations as being more salient?
- Automatically harvesting instances of salient semantic relation types (from corpora, WordNet, etc., using seeds–pattern approach)

Finally: Resource for dictionary extension (also useful for concept descriptions/definitions, etc.)

Collaborators

Marco Baroni — CIMeC—Center for Mind/Brain Sciences
Andrea Abel — EURAC—European Academy Bolzano

References

McRae, K., Cree, G. S., Seidenberg, M. S., and McNorgan, C. (2005).
Semantic Feature Production Norms for a Large Set of Living and Nonliving Things.
Behaviour Research Methods, Instruments & Computers, 37(4):547–559.

Pantel, P. and Pennacchiotti, M. (2006).
Espresso: Leveraging Generic Patterns for Automatically Harvesting Semantic Relations.
In *Proceedings of the International Conference on Computational Linguistics and the Annual Meeting of the ACL*, pages 113–120, Sydney, Australia. ACL.

Vinson, D. P. and Vigliocco, G. (2008).
Semantic Feature Production Norms for a Large Set of Objects and Events.
Behaviour Research Methods, 40(1):183–190.

Wu, L. and Barsalou, L. W. (2004).
Grounding Concepts in Perceptual Simulation: I. Evidence From Property Generation.
Unpublished manuscript.