Cognitively Salient Relations for Multilingual Lexicography

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

Andrea Abel
EURAC
Bolzano
aabel@eurac.edu

Marco Baroni
CIMeC
University of Trento
marco.baroni@unitn.it

Abstract

Providing sets of semantically related words in the lexical entries of an electronic dictionary should help language learners quickly understand the meaning of the target words. Relational information might also improve memorisation, by allowing the generation of structured vocabulary study lists. However, an open issue is *which* semantic relations are cognitively most salient, and should therefore be used for dictionary construction. In this paper, we present a concept description elicitation experiment conducted with German and Italian speakers. The analysis of the experimental data suggests that there is a small set of concept-class–dependent relation types that are stable across languages and robust enough to allow discrimination across broad concept domains. Our further research will focus on harvesting instantiations of these classes from corpora.

1 Introduction

In electronic dictionaries, lexical entries can be enriched with hyperlinks to semantically related words. In particular, we focus here on those related words that can be seen as systematic *properties* of the target entry, i. e., the basic concepts that would be used to define the entry in relation to its superordinate category and coordinate concepts. So, for example, for animals the most salient relations would be notions such as "parts" and "typical behaviour". For a horse, salient properties will include the mane and hooves as parts, and neighing as behaviour.

Sets of relevant and salient properties allow the user to collocate a word within its so-called "word field" and to distinguish it more clearly from neighbour concepts, since the meaning of a word is not defined in isolation, but in contrast to related words in its word field (Geckeler, 2002). Moreover, knowing the typical relations of concepts in different domains might help pedagogical lexicography to produce structured networks where, from each word, the learner can naturally access entries for other words that represent properties which are salient and distinctive for the target concept class (parts of animals, functions of tools, etc.). We envisage a natural application of this in the automated creation of structured vocabulary study lists. Finally, this knowledge might be used as a basis to populate lexical networks by building models of concepts in terms of "relation sketches" based on salient typed properties (when an animal is added to our lexicon, we know that we will have to search a corpus to extract its parts, behaviour, etc., whereas for a tool the function would be the most important property to mine).

This paper provides a first step in the direction of dictionaries enriched with cognitively salient property descriptions by eliciting concept descriptions from subjects speaking different languages, and analysing the general patterns emerging from these data.

It is worth distinguishing our approach to enriching connections in a lexical resource from the one based on free association, such as has been recently pursued, e. g., within the WordNet project (Boyd-Graber et al., 2006). While we do not dispute the usefulness of free associates, they are irrelevant to

Coling 2008: Proceedings of the workshop on Cognitive Aspects of the Lexicon (COGALEX 2008), pages 94–101
Manchester, August 2008

our purposes, since we want to generate systematic, structured descriptions of concepts, in terms of the relation types that are most salient for their semantic fields. Knowing that the word *Holland* is "evoked" by the word *tulip* might be useful for other reasons, but it does not allow us to harvest systematic properties of flowers in order to populate their relation sketch: we rather want to find out that tulips, being flowers, will have *colour* as a salient property type. As a *location* property of tulips, we would prefer something like *garden* instead of the name of a country or individual associations. To minimise free association, we asked participants in our experiments to produce concept *descriptions* in terms of characteristic properties of the target concepts (although we are not aware of systematic studies comparing free associates to concept description tasks, the latter methodology is fairly standard in cognitive science: see section 2.2 below).

To our knowledge, this sort of approach has not been proposed in lexicography, yet. Cognitive scientists focus on "concepts", glossing over the fact that what subjects will produce are (strings of) words, and as such they will be, at least to a certain extent, language-dependent. For lexicographic applications, this aspect cannot, of course, be ignored, in particular if the goal is to produce lexical entries for language learners (so that both their first and their second languages should be taken into account).

We face this issue directly in the elicitation experiment we present here, in which salient relations for a set of 50 concepts from 10 different categories are collected from comparable groups of German and Italian speakers. In particular, we collected data from high school students in South Tyrol, a region situated in Northern Italy, inhabited by both German and Italian speakers. Both German and Italian schools exist, where the respective non-native language is taught. It is important to stress that the two communities are relatively separated, and most speakers are *not* from bilingual families or bilingual social environments: They study the other language as an intensively taught L2 in school. Thus, we move in an ideal scenario to test possible language-driven differences in property descriptions, among speakers that have a very similar cultural background.

South Tyrol also provides the concrete applicative goal of our project. In public administration and service, employees need to master both languages up to a certain standardised level (they have to pass a "bilingual" proficiency exam). Therefore, there is a big need for language learning materials. The practical outcome of our research will be an extension of ELDIT[1], an electronic learner's dictionary for German and Italian (Abel and Weber, 2000).

2 Related Work

Lexicographic projects providing semantic relations and experimental research on property generation are the basis for our research.

2.1 Dictionaries

In most paper-based general and learners' dictionaries only some information about synonyms and sometimes antonyms is presented. Newer dictionaries, such as the "Longman Language Activator" (Summers, 1999), are providing lists of related words. While these will be useful to learners, information about the *kind* of semantic relation is usually missing.

Semantic relations are often available in electronic resources, most famously in WordNet (Fellbaum, 1998) and related projects like Kirrkirr (Jansz et al., 1999), ALEXIA (Chanier and Selva, 1998), or as described in Fontenelle (1997). However, these resources tend to include few relation types (hypernymy, meronymy, antonymy, etc.). The salience of the relations chosen is not verified experimentally, and the same set of relation types is used for all words that share the same part-of-speech. Our results below, as well as work by Vinson et al. (2008), indicate that different concept classes should, instead, be characterised by different relation types (e. g., function is very salient for tools, but not at all for animals).

2.2 Work in Cognitive Sciences

Several projects addressed the collection of property generation data to provide the community with feature norms to be used in different psycholinguistic experiments and other analyses: Garrard et al. (2001) instructed subjects to complete phrases ("concept is/has/can. . . "), thus restricting the set of producible feature types. McRae et al. (2005) instructed their subjects to list concept properties without such restrictions, but providing them with some examples. Vinson et al. (2008)

[1] URL `http://www.eurac.edu/eldit`

gave similar instructions, but explicitly asked subjects not to freely associate.

However, these norms have been collected for the English language. It remains to be explored if concept representations in general and semantic relations for our specific investigations have the same properties across languages.

3 Data Collection

After choosing the concept classes and appropriate concepts for the production experiment, concept descriptions were collected from participants. These were transcribed, normalised, and annotated with semantic relation types.

3.1 Stimuli

The stimuli for the experiment consisted of 50 concrete concepts from 10 different classes (i. e., 5 concepts for each of the classes): *mammal* (dog, horse, rabbit, bear, monkey), *bird* (seagull, sparrow, woodpecker, owl, goose), *fruit* (apple, orange, pear, pineapple, cherry), *vegetable* (corn, onion, spinach, peas, potato), *body part* (eye, finger, head, leg, hand), *clothing* (chemise, jacket, sweater, shoes, socks), *manipulable tool* (comb, broom, sword, paintbrush, tongs), *vehicle* (bus, ship, airplane, train, truck), *furniture* (table, bed, chair, closet, armchair), and *building* (garage, bridge, skyscraper, church, tower). They were mainly taken from Garrard et al. (2001) and McRae et al. (2005). The concepts were chosen so that they had unambiguous, reasonably monosemic lexical realizations in both target languages.

The words representing these concepts were translated into the two target languages, German and Italian. A statistical analysis (using Tukey's honestly significant difference test as implemented in the R toolkit[2]) of word length distributions (within and across categories) showed no significant differences in either language. There were instead significant differences in the frequency of target words, as collected from the German, Italian and English WaCky corpora[3]. In particular, words of the class *body part* had significantly larger frequencies across languages than the words of the other classes (not surprisingly, the words *eye*, *head* and *hand* appear much more often in corpora than the other words in the stimuli list).

[2] URL `http://www.r-project.org/`
[3] URL `http://wacky.sslmit.unibo.it/`

3.2 Experimental Procedure

The participants in the concept description experiment were students attending the last 3 years of a German or Italian high school and reported to be native speakers of the respective languages. 73 German and 69 Italian students participated in the experiment, with ages ranging between 15 and 19. The average age was 16.7 (standard deviation 0.92) for Germans and 16.8 (s.d. 0.70) for Italians.

The experiment was conducted group-wise in schools. Each participant was provided with a random set of 25 concepts, each presented on a separate sheet of paper. To have an equal number of participants describing each concept, for each randomly matched subject pair the whole set of concepts was randomised and divided into 2 subsets. Each subject saw the target stimuli in his/her subset in a different random order (due to technical problems, the split was not always different across subject pairs).

Short instructions were provided orally before the experiment, and repeated in written format on the front cover of the questionnaire booklet distributed to each subject. To make the concept description task more natural, we suggested that participants should imagine a group of alien visitors, to each of which a particular word for a concrete object was unknown and thus had to be described. Participants should assume that each alien visitor knew all other words of the language apart from the unknown (target) word.

Participants were asked to enter a descriptive phrase per line (not necessarily a whole sentence) and to try and write at least 4 phrases per word. They were given a maximum of one minute per concept, and they were not allowed to go back to the previous pages.

Before the real experiment, subjects were presented an example concept (not in the target list) and were encouraged to describe it while asking clarifications about the task.

All subjects returned the questionnaire so that for a concept we obtained, on average, descriptions by 36.48 German subjects (s.d. 1.24) and 34.34 Italian subjects (s.d. 1.72).

3.3 Transcription and Normalisation

The collected data were digitally transcribed and responses were manually checked to make sure that phrases denoting different properties had been properly split. We tried to systematically apply the

criterion that, if at least one participant produced 2 properties on separate lines, then the properties would always be split in the rest of the data set.

However, this approach was not always equally applicable in both languages. For example, *Transportmittel* (German) and *mezzo di trasporto* (Italian) both are compounds used as hypernyms for what English speakers would probably rather classify as *vehicles*. In contrast to *Transportmittel*, *mezzo di trasporto* is splittable as *mezzo*, that can also be used on its own to refer to a kind of vehicle (and is defined more specifically by adding the fact that it is used for transportation). The German compound word also refers to the function of transportation, but *-mittel* has a rather general meaning, and would not be used alone to refer to a vehicle. Hence, *Transportmittel* was kept as a whole and the Italian quasi-equivalent was split, possibly creating a bias between the two data sets (if the Italian string is split into *mezzo* and *trasporto*, these will be later classified as hypernym and functional features, respectively; if the German word is not split, it will only receive one of these type labels). More in general, note that in German compounds are written as single orthographic words, whereas in Italian the equivalent concepts are often expressed by several words. This could also create further bias in the data annotation and hence in the analysis.

Data were then normalised and transcribed into English, before annotating the type of semantic relation. Normalisation was done in accordance with McRae et al. (2005), using their feature norms as guidelines, and it included leaving habitual words like "normally,", "often", "most" etc. out, as they just express the typicality of the concept description, which is the implicit task.

3.4 Mapping to Relation Types

Normalised and translated phrases were subsequently labelled for relation types following McRae et al.'s criteria and using a subset of the semantic relation types described in Wu and Barsalou (2004): see section 4.1 below for the list of relations used in the current analysis.

Trying to adapt the annotation style to that of McRae et al., we encountered some dubious cases. For example, in the McRae et al.'s norms, *carnivore* is classified as a hypernym, but *eats_meat* as a behaviour, whereas they seem to us to convey essentially the same information. In this case, we decided to map both to *eats_meat* (behaviour).

Among other surprising choices, the normalised phrase *used_for_cargo* is seen by McRae et al. as a function, but *used_by_passengers* is classified as denoting the participants in a situation. In this case, we followed their policy.

While we tried to be consistent in relation labelling within and across languages, it is likely that our own normalisation and type mapping also include a number of inconsistencies, and our results must be interpreted by keeping this important *caveat* in mind.

The average number of normalised phrases obtained for a concept presented is 5.24 (s.d. 1.82) for the German participants and 4.96 (s.d. 1.86) for the Italian participants; in total, for a concept in our set, the following number of phrases was obtained on average: 191.28 (German, s.d. 25.96) and 170.42 (Italian, s.d. 25.49).

4 Results

The distribution of property types is analysed both class-independently and within each class (separately for German and Italian), and an unsupervised clustering analysis based on property types is conducted.

4.1 Distributional Analysis

We first look at the issue of how comparable the German and Italian data are, starting with a check of the overlap at the level of specific properties. There are 226 concept–property pairs that were produced by at least 10 German subjects; 260 pairs were produced by at least 10 Italians. Among these common pairs, 156 (i. e., 69% of the total German pairs, and 60% of the Italian pairs) are shared across the 2 languages. This suggests that the two sets are quite similar, since the overlap of specific pairs is strongly affected by small differences in normalisation (e. g., *has a fur*, *has fur* and *is hairy* count as completely different properties).

Of greater interest to us is to check to what extent property *types* vary across languages and across concept classes. In order to focus on the main patterns emerging from the data, we limit our analysis to the 6 most common property types in the whole data set (that are also the top 6 types in the two languages separately), accounting for 69% of the overall responses. These types are:

- category (Wu/Barsalou code: *ch*; "*pear is a fruit*")

- (external) part (WB code: *ece*; "*dog has 4 legs*")
- (external) quality (WB code: *ese*; "*apple is green*")
- behaviour (WB code: *eb*; "*dog barks*")
- function (WB code: *sf*; "*broom is for sweeping*")
- location (WB code: *sl*; "*skyscraper is found in cities*")

Figure 1 compares the distribution of property types in the two languages via a *mosaic plot* (Meyer et al., 2006), where rectangles have areas proportional to observed frequencies in the corresponding cells. The overall distribution is very similar. The only significant differences pertain to category and location types: Both differences are significant at the level $p < 0.0001$, according to a Pearson residual test (Zeileis et al., 2005).

For the difference in location, no clear pattern emerges from a qualitative analysis of German and Italian location properties. Regarding the difference in (superordinate) categories, we find, interestingly, a small set of more or less abstract hypernyms that are frequently produced by Italians, but never by Germans: *construction* (72), *object* (36), *structure* (16). In the these cases, the Italian translations have subtle shades of meaning that make them more likely to be used than their German counterparts. For example, the Italian word *oggetto* ("object") is used somewhat more concretely than the extremely abstract German word *Objekt* (or English "object", for that matter) – in Italian, the word might carry more of an "artifact, man-made item" meaning. At the same time, *oggetto* is less colloquial than German *Sache*, and thus more amenable to be entered in a written definition. In addition, among others, the category *vehicle* was more frequent in the Italian than in the German data set (for which one reason could be the difference between the German and Italian equivalents, which was discussed in section 3.3). Differences of this sort remind us that property elicitation is first and foremost a verbal task, and as such it is constrained by language-specific usages. It is left to future research to test to what extent linguistic constraints also affect deeper conceptual representations (would Italians be faster than Germans at recognising superordinate properties of concepts when they are expressed non-verbally?).

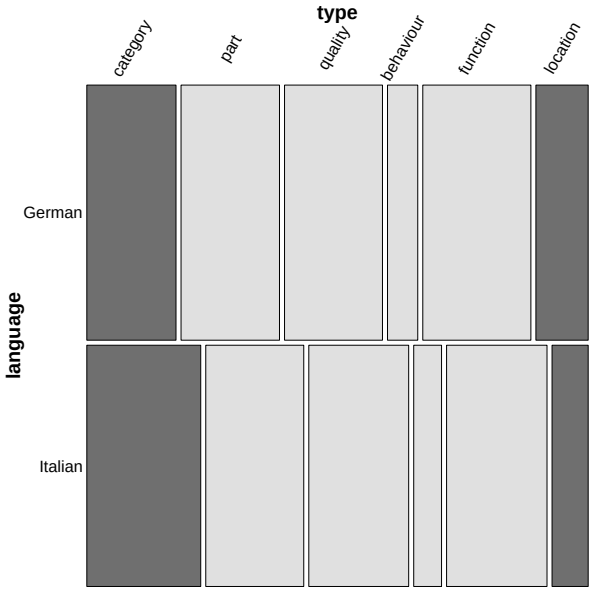

Figure 1: Cross-language distribution of property types

Despite the differences we just discussed, the main trend emerging from figure 1 is one of essential agreement between the two languages, and indicates that, with some caveats, salient property types may be cross-linguistically robust. We, thus, turn to the issue of how such types are distributed across concepts of different classes. This question is visually answered by the association plots in figure 2 on the following page.

Each plot illustrates, through rectangle heights, how much each cell deviates from the value expected given the overall contingency tables (in our case, the reference contingency tables are the language-specific distributions of figure 1). The sign of the deviation is coded by direction with respect to the baseline. For example, the first row of the left plot tells us, among other things, that in German behaviour properties are strongly over-represented in mammals, whereas function properties are under-represented within this class. Like in figure 1, shades of grey cue degrees of significance of the deviation (Meyer et al., 2003).

The first observation we can make about figure 2 is how, for both languages, a large proportion of cells show a significant departure from the overall distribution. This confirms what has already been observed and reported in the literature on English norms – see, in particular, Vinson et. al. (2008):

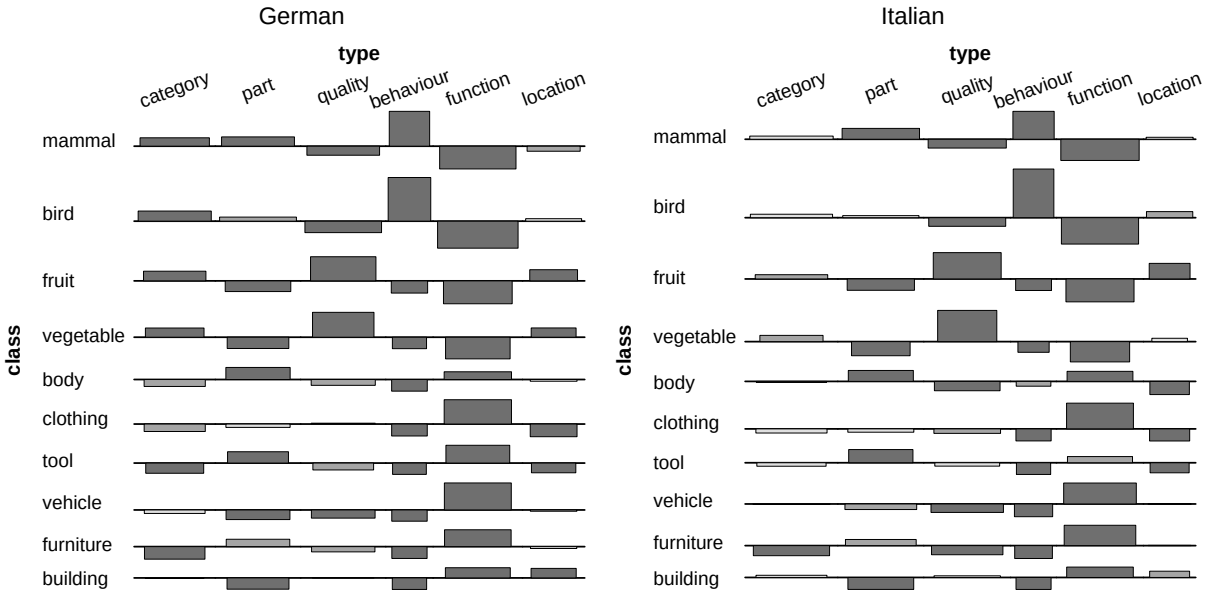

Figure 2: Distribution of property types across classes

property types are highly distinctive characteristics of concept classes.

The class-specific distributions are extremely similar in German and Italian. There is no single case in which the same cell is deviating significantly but in opposite directions in the two languages; and the most common pattern by far is the one in which the two languages show the same deviation profile across cells, often with very similar effect sizes (compare, e. g., the *behaviour* and *function* columns). These results suggest that property types are not much affected by linguistic factors, an intrinsically interesting finding that also supports our idea of structuring relation-based navigation in a multi-lingual dictionary using concept-class–specific property types.

The type patterns associated with specific concept classes are not particularly surprising, and they have been already observed in previous studies (Vinson and Vigliocco, 2008; Baroni and Lenci, 2008). In particular, living things (animals and plants) are characterised by paucity of functional features, that instead characterise all man-made concepts. Within the living things, animals are characterised by typical behaviours (they bark, fly, etc.) and, to a lesser extent, parts (they have legs, wings, etc.), whereas plants are characterised by a wealth of qualities (they are sweet, yellow, etc.) Differences are less pronounced within man-made objects, but we can observe parts as typical of tool and furniture descriptions. Finally, location is a more typical definitional characteristic of buildings (for clothing, nothing stands out, if not, perhaps, the pronounced *lack* of association with typical locations). Body parts, interestingly, have a type profile that is very similar to the one of (manipulable) tools – manipulable objects are, after all, extensions of our bodies.

4.2 Clustering by Property Types

The distributional analysis presented in the previous section confirmed our main hypotheses – that property types are salient properties of concepts that differ from a concept class to the other, but are robust across languages. However, we did not take skewing effects associated to specific concepts into account (e. g., it could be that, say, the property profile we observe for body parts in figure 2 is really a deceiving average of completely opposite patterns associated to, say, heads and hands). Moreover, our analysis already assumed a division into classes – but the type patterns, e. g., of mammals and birds are very similar, suggesting that a higher-level "animal" class would be more appropriate when structuring concepts in terms of type profiles. We tackled both issues in an unsupervised clustering analysis of our 50 target concepts based on their property types. If the postulated classes are not internally coherent, they will not form coherent clusters. If some classes should be merged, they will cluster together.

Concepts were represented as 6-dimensional vectors, with each dimension corresponding to one

of the 6 common types discussed above, and the value on a dimension given by the number of times that concept triggered a response of the relevant type. We used the CLUTO toolkit[4], selecting the `rbr` method and setting all other clustering parameters to their default values. We explored partitions into 2 to 10 clusters, manually evaluating the output of each solution.

Both in Italian and in German, the best results were obtained with a 3-way partition, neatly corresponding to the division into animals (mammals and birds), plants (vegetables and fruits) and objects plus body parts (that, as we observed above, have a distribution of types very similar to the one of tools). The 2-way solution resulted in merging two of the classes animals and plants both in German and in Italian. The 4-way solution led to an arbitrary partition among objects and body parts (and *not*, as one could have expected, in separating objects from body parts). Similarly, the 5- to 10-way solutions involve increasingly granular but still arbitrary partitions within the objects/body parts class. However, one notable aspect is that in most cases almost all concepts of mammals and birds, and vegetables and fruits are clustered together (both in German and Italian), expressing their strong similarity in terms of property types as compared to the other classes as defined here.

Looking at the 3-way solution in more detail, in Italian, the concept *horse* is in the same cluster with objects and body parts (as opposed to German, where the solution is perfect). The misclassification results mainly from the fact that for *horse* a lot of functional properties were obtained (which is a feature of objects), but none of them for the other animals in the Italian data. In German, some functional properties were assigned to both *horse* and *dog*, which might explain why it was not misclassified there.

To conclude, the type profiles associated with animals, vegetables and objects/body parts have enough internal coherence that they robustly identify these macro-classes in both languages. Interestingly, a 3-way distinction of this sort – excluding body parts – is seen as fundamental on the basis of neuro-cognitive data by Caramazza and Shelton (1998). On the other hand, we did not find evidence that more granular distinctions could be made based on the few (6) and very general types we used. We plan to explore the distribution across the remaining types in the future (preliminary clustering experiments show that much more nuanced discriminations, even among all 10 categories, can be made if we use all types). However, for our applied purposes, it is sensible to focus on relatively coarse but well-defined classes, and on just a few common relation types (alternatively, we plan to combine types into superordinate ones, e. g. external and internal quality). This should simplify both the automatic harvesting of corpus-based properties of the target types and the structuring of the dictionary relational interface.

Finally, the peculiar object-like behaviour of body parts on the one hand, and the special nature of horse, on the other, should remind us of how concept classification is not a trivial task, once we try to go beyond the most obvious categories typically studied by cognitive scientists – animals, plants, manipulable tools. In a lexicographic perspective, this problem cannot be avoided, and, indeed, the proposed approach should scale in difficultiese to even trickier domains, such as those of actions or emotions.

5 Conclusion

This research is part of a project that aims to investigate the cognitive salience of semantic relations for (pedagogical) lexicographic purposes. The resulting most salient relations are to be used for revising and adding to the word field entries of a multilingual electronic dictionary in a language learning environment.

We presented a multi-lingual concept description experiment. Participants produced different semantic relation type patterns across concept classes. Moreover, these patterns were robust across the two native languages studied in the experiment – even though a closer look at the data suggested that linguistic constraints might affect (verbalisations of) conceptual representations (and thus, to a certain extent, which properties are produced). This is a promising result to be used for automatically harvesting semantically related words for a given lexical entry of a concept class.

However, the granularity of concept classes has to be defined. In addition, to yield a larger number of usable data for the analysis, a re-mapping of the rare semantic relation types occurring in the actual data set should be conducted. Moreover, the stimuli set will have to be expanded to include, e. g., ab-

[4] URL `http://glaros.dtc.umn.edu/gkhome/cluto/cluto/overview`

stract concepts – although we hope to mine some abstract concept classes on the basis of the properties of our concept set (colours, for example, could be characterised by the concrete objects of which they are typical).

To complement the production experiment results, we aim to conduct an experiment which investigates the perceptual salience of the produced semantic relations (and possibly additional ones), in order to detect inconsistencies between generation and retrieval of salient properties. If, as we hope, we will find that essentially the same properties are salient for each class across languages and both in production and perception, we will then have a pretty strong argument to suggest that these are the relations one should focus on when populating multi-lingual dictionaries.

Of course, the ultimate test of our approach will come from empirical evidence of the usefulness of our relation links to the language learner. This is, however, beyond the scope of the current project.

References

[Abel and Weber2000] Abel, Andrea and Vanessa Weber. 2000. ELDIT—A Prototype of an Innovative Dictionary. In Heid, Ulrich, Stefan Evert, Egbert Lehmann, and Christian Rohrer, editors, EURALEX *Proceedings*, volume 2, pages 807–818, Stuttgart.

[Baroni and Lenci2008] Baroni, Marco and Alessandro Lenci. 2008. Concepts and Properties in Word Spaces. *Italian Journal of Linguistics*. To appear.

[Boyd-Graber et al.2006] Boyd-Graber, Jordan, Christaine Fellbaum, Daniel Osherson, and Robert Schapire. 2006. Adding Dense, Weighted Connections to WordNet. In *Proceedings of the Thirds International WordNet Conference*. Masaryk University Brno.

[Caramazza and Shelton1998] Caramazza, Alfonso and Jennifer R. Shelton. 1998. Domain–Specific Knowledge Systems in the Brain: The Animate–Inanimate Distinction. *Journal of Cognitive Neuroscience*, 10:1–34.

[Chanier and Selva1998] Chanier, Thierry and Thierry Selva. 1998. The ALEXIA system: The Use of Visual Representations to Enhance Vocabulary Learning. In *Computer Assisted Language Learning*, volume 11, pages 489–522.

[Fellbaum1998] Fellbaum, Christiane, editor. 1998. *WordNet: An Electronic Lexical Database*. Language, Speech, and Communication. MIT Press, Cambridge, MA.

[Fontenelle1997] Fontenelle, Thierry. 1997. Using a Bilingual Dictionary to Create Semantic Networks. *International Journal of Lexicography*, 10(4):275–303.

[Garrard et al.2001] Garrard, Peter, Matthew A. Lambon Ralph, John R. Hodges, and Karalyn Patterson. 2001. Prototypicality, Distinctiveness, and Intercorrelation: Analyses of the Semantic Attributes of Living and Nonliving Concepts. *Cognitive Neuropsychology*, 18(2):125–174.

[Geckeler2002] Geckeler, Horst. 2002. Anfänge und Ausbau des Wortfeldgedankens. In Cruse, D. Alan, Franz Hundsnurscher, Michael Job, and Peter Rolf Lutzeier, editors, *Lexikologie. Ein internationales Handbuch zur Natur und Struktur von Wörtern und Wortschätzen*, volume 21 of *Handbücher zur Sprach- und Kommunikationswissenschaft*, pages 713–728. de Gruyter, Berlin – New York.

[Jansz et al.1999] Jansz, Kevin, Christopher Manning, and Nitin Indurkha. 1999. Kirrkirr: Interactive Visualisation and Multimedia From a Structured Warlpiri Dictionary. In *Proceedings of the 5th Australian World Wide Web Conference (AusWeb'99)*, pages 302–316.

[McRae et al.2005] McRae, Ken, George S. Cree, Mark S. Seidenberg, and Chris McNorgan. 2005. Semantic Feature Production Norms for a Large Set of Living and Nonliving Things. *Behaviour Research Methods, Instruments & Computers*, 37(4):547–559.

[Meyer et al.2003] Meyer, David, Achim Zeileis, and Kurt Hornik. 2003. Visualizing Independence Using Extended Association Plots. In *Proceedings of DSC 2003*. Online at URL `http://www.ci.tuwien.ac.at/Conferences/DSC-2003/`.

[Meyer et al.2006] Meyer, David, Achim Zeileis, and Kurt Hornik. 2006. The Strucplot Framework: Visualizing Multi–Way Contingency Tables With vcd. *Journal of Statistical Software*, 17(3):1–48.

[Summers1999] Summers, Della, editor. 1999. *Longman Language Activator. The World's First Production Dictionary*. Longman, Harlow.

[Vinson and Vigliocco2008] Vinson, David P. and Gabriella Vigliocco. 2008. Semantic Feature Production Norms for a Large Set of Objects and Events. *Behaviour Research Methods*, 40(1):183–190.

[Wu and Barsalou2004] Wu, Ling–ling and Lawrence W. Barsalou. 2004. Grounding Concepts in Perceptual Simulation: I. Evidence From Property Generation. Unpublished manuscript.

[Zeileis et al.2005] Zeileis, Achim, David Meyer, and Kurt Hornik. 2005. Residual–Based Shadings for Visualizing (Conditional) Independence. Technical Report 20, Department of Statistics and Mathematics, Wirtschaftsuniversität, Vienna.