Automatically Determining the Semantic Gradation of German Adjectives

Peter F. Schulam
Department of Computer Science
Princeton University
United States
pschulam@princeton.edu

Christiane Fellbaum
Department of Computer Science
Princeton University
United States
fellbaum@princeton.edu

Abstract
The semantics of gradable adjectives, in particular their order with respect to one another on a given scale, are not sufficiently represented in lexical resources such as wordnets. Thus, it is not clear whether superb expresses the quality of “goodness” equally, more, or less strongly than excellent or great. Sheinman and Tokunaga (2009) determine the relative gradation of English adjectives by applying lexical-semantic patterns that hold between members of pairs of similar descriptive adjectives to corpus searches. The patterns identify one member of such pairs as the one that expresses a stronger, or more intense, degree of the property denoted by the scale than the other member. By iteratively applying these patterns to a set of adjectives, Sheinman and Tokunaga (2009) arrive at a uniform score for each adjective that allows them to place it at an appropriate point on a scale. We extend the AdjScales method (Sheinman and Tokunaga 2009) to some frequent and salient German adjectives to test its crosslingual robustness. Our work has consequences for automatic text understanding and generation, lexicography and language pedagogy.

1 Introduction
Adjectives remain a relatively ill-understood category. This is reflected in their representation in language resources, which typically lack explicit indications of the degree or intensity with which adjectives express a common property. Thus Roget’s 21st Century Thesaurus identifies both acceptable and superb as synonyms of good 1, but native speakers easily agree that the sentences Her work was acceptable/good/superb express different meanings.

1http://www.thesaurus.com/browse/good

Automatic text understanding and generation systems must be able to differentiate among such adjectives as well.

1.1 Adjectives in WordNet
The most widely used lexical resource, WordNet, organizes English adjectives into “dumbbell” structures, consisting of two polar adjectives and a number of adjectives that are semantically similar to one of each of the poles (Gross et al. 1989; Gross and Miller 1990). Polar adjective pairs such as long and short, called “direct” antonyms by Gross and Miller (1990), label two salient opposed points on a scale such as “length” (Bierwisch 1987). Extended and abbreviated are ”semantically similar” to long and short, respectively. These adjectives are called “indirect” antonyms of the polar adjectives. Figure 1 shows an example of a WordNet “dumbbell”.

Figure 1: An example of a WordNet “dumbbell structure” (Gross and Miller 1990)

A major shortcoming of WordNet’s treatment of adjectives is that it does not distinguish among the
strength of the similar adjectives: what is the degree to which they express the property of the scale? I.e., what is their relative position with respect to the polar adjectives and one another?

GermaNet abandons the use of WordNet’s dumbbell structure and adopts a hierarchical structure akin to that which is used for organizing nouns and verbs (Hamp and Feldweg 1997). The root nodes, like temperaturspezifisch (“temperature-specific”) refer to the shared property and, in some cases, to the domain of the adjectives in the tree. These terms are the direct superordinates of the polar adjectives. Each polar adjective dominates a subtree of additional adjectives, which are claimed to be increasingly more specific as one descends the tree.

While the Princeton WordNet does not distinguish between the relative intensities of similar adjectives, it encodes a coarse scale by focusing on antonymous adjectives. GermaNet, however, completely abandons a scalar approach in adopting a hierarchical structure. Schulam (2009) questions this move away from bipolarity by showing the strong co-occurrence patterns of German bipolar adjectives — the phenomenon that underpins the Princeton WordNet organization. Furthermore, rejecting a bipolar representation fails to account for the scalar properties of many adjectives.

2 Gradability

Our goal is to assign gradable adjectives to their relative positions on a scale. Kennedy (1999) points out that “a defining characteristic of gradable adjectives is that there is some gradient property associated with their meaning with respect to which objects in their domains can be ordered.” A necessary characteristic of gradable adjectives is that they “ascribe to their head nouns values of (typically) bipolar attributes and consequently are organized in terms of binary oppositions” (Gross and Miller 1990). The bipolar adjectives, as well as their similar adjectives, naturally allow themselves to be placed on unidimensional scales, with each pole of the dimension corresponding to one of the two attributes.

2.1 Property Scales

Languages tend to provide words for referring to intermediate values between opposing semantic poles. Consider the phrase “The bread is warm”. If we drew a straight line through the lexicalized “points” associated with the words cold, lukewarm, and hot, the word warm might lie on the line somewhere between lukewarm and hot.

Lexical resources indicate that speaker introspection provides only fuzzy judgments about where gradable adjectives fall on a scale. Nevertheless, language provides for ways to detect these subtle gradations via lexical-semantic patterns. Our goal is to empirically identify such patterns for German adjectives and to propose a method for correctly placing them on their respective scales.

3 AdjScales

Sheinman and Tokunaga (2009) propose AdjScales, a method that accepts a group of similar adjectives as input and returns as output a unified scale on which the initial adjectives, as well as additional similar ones, are situated in order of increasing strength.

The fundamental preparatory step is the discovery of scaling patterns (Sheinman and Tokunaga 2009). AdjScales uses pattern extraction queries of the form \(a \ast b\), \(a\) and \(b\) are seed words and \(\ast\) is a wildcard. Sheinman and Tokunaga perform such queries using a yahoo search engine API that allows for searches and the collection of “snippets”, small text excerpts for each result returned from a search engine. These snippets can then be compiled into a database of excerpts containing patterns of interest. AdjScales selects seed words in a supervised manner using two seeds \(seed_1\) and \(seed_2\) such that \(seed_2\) is stronger-than \(seed_1\). Sheinman and Tokunaga selected 10 seed word pairs from Gross and Miller (1990) that intuitively display a clear gradation along the same scale.

After successfully creating a database of snippets, AdjScales extracts patterns of the form \([prefix_x, x inflex_y, y postfix_y]\) “where \(x\) and \(y\) are slots for words or multiword expressions” (Sheinman and Tokunaga 2009). Pattern candidates must be consistent with respect to the order in which all instances display the seed words. “Intense” patterns display the weaker word first and the more intense word second, while “mild” patterns do the opposite. Valid patterns must also be supported by at least three seed pairs, and they must repeat twice in extracted sentences. Finally, the patterns must be supported by seed pairs describing different properties.

Pattern extraction is followed by several steps: input, scale relaxation, extension, scaling, and scales unification.

The input step selects at least two similar adjec-
tives. Scale relaxation divides the input adjective set into two antonymous subsets using WordNet’s dumbbell structure (Sheinman and Tokunaga 2009).

The extension step proposes adjectives belonging to the same scale as the input word, based on members of a WordNet dumbbell structure.

Scaling involves iteratively placing pairs of adjectives from the extended input set into the extracted patterns. Sheinman and Tokunaga employ a weighting algorithm that uses the page hit counts returned from a search engine when searching for the complete phrase formed by the adjective pair and the extracted pattern. They use this algorithm to apply a uniform score representing intensity to each adjective (Sheinman and Tokunaga 2009). This allows for the adjective in each of the subdivided groups to be ordered according to increasing intensity as indicated by their score.

After the two subdivided groups have been independently ordered, the scales are unified. The two independently unified scales are merged at the “mild” ends of the spectrum.

4 AdjScales in German

Adapting the pattern extraction process to German adjectives involved selecting suitable seed words, choosing an accessible and extensive corpus in which we could search for patterns, and selecting patterns from the data returned from the pattern extraction queries.

4.1 Seed Word Selection

From a list of 35 antonymous adjective pairs identified by Deese (1964) we selected five antonym pairs as candidate seeds and translated them into German: kalt-heiß, dunkel-hell, schnell-langsam, traurig-glücklich, and stark-schwach. The pairs represent a variety of different properties to ensure that our extracted patterns would apply to a broad range of semantically diverse adjectives.

Next we paired each of the members of the five antonymous pairs with another adjective from the same scale that we intuitively judged to be more mild or more intense. For example, for kalt (cold) we chose the milder adjectives kühl (cool) to complete the seed pair.

We compiled a list of similar adjectives for each of the members of the five antonymous pairs by using the graphical user interface GermaNet Explorer, which allowed us to search the adjective trees in GermaNet. After compiling a list of similar adjectives for each of the members of the pairs, we performed queries using COSMAS II.

We searched for sentences containing both a translated Deese adjective and one of the corresponding similar adjectives. After iterating through all similar adjectives for each of the Deese adjectives, we chose the most suitable pairing based on the size and diversity of the results returned. The final seed pairs can be found in Table 1. Table cells filled with “XXX” indicate that no appropriate adjective was discovered.

4.2 Pattern Extraction

To extract patterns, we performed queries in COSMAS II that searched for co-occurrences of the seed pairs within the same sentence regardless of their relative order. We exported the results to text files and processed them using simple python scripts.

We first separated the results for each pair of adjectives into files containing “mild” patterns and files containing “intense” patterns. We then removed all results in which the seed words were connected only by und, as this pattern does not indicate the relative strength of the adjectives that it joins.

The final list of aggregated patterns is shown in Tables 2 and 3.

---

Table 1: Complete list of seed words chosen for this study.

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Kalt</td>
<td>Kühl</td>
<td>Heiß</td>
<td>Warm</td>
</tr>
<tr>
<td>Dunkel</td>
<td>Düster</td>
<td>Hell</td>
<td>Grell</td>
</tr>
<tr>
<td>Schnell</td>
<td>Hastig</td>
<td>Langsam</td>
<td>Schleppend</td>
</tr>
<tr>
<td>Traurig</td>
<td>Bitter</td>
<td>Glücklich</td>
<td>Zufrieden</td>
</tr>
<tr>
<td>Stark</td>
<td>Stabil</td>
<td>Schwach</td>
<td>XXX</td>
</tr>
</tbody>
</table>

Table 2: List of discovered intense patterns.

<table>
<thead>
<tr>
<th>#</th>
<th>Intense Patterns</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>X, fast Y</td>
</tr>
<tr>
<td>2</td>
<td>X, nicht jedoch Y</td>
</tr>
<tr>
<td>3</td>
<td>X, zwar nicht Y</td>
</tr>
<tr>
<td>4</td>
<td>X und oft Y</td>
</tr>
<tr>
<td>5</td>
<td>X sogar Y</td>
</tr>
<tr>
<td>6</td>
<td>X, aber nicht Y</td>
</tr>
</tbody>
</table>

---

2cold-hot, dark-bright, fast-slow, sad-happy, strong-weak

3http://www.hytex.tu-dortmund.de/ressourcen.html
4http://www.ids-mannheim.de/cosmas2
4.3 Pattern Testing

To test our extracted patterns we used a python script to submit queries containing a pattern in which the adjective slots were filled with test adjective pairs that were different from those used to initially extract the patterns. A test was considered successful when the search returned more results for intense patterns when the adjective pairs were ordered mild-intense and when the search returned more results for mild patterns with the adjective pairs were ordered intense-mild.

Many pairs of test adjectives did not return any hits for both the appropriate and inappropriate ordering for both mild and intense patterns. On the other hand, a number of pairs produced successful results. For example, the pair laut (loud) and ohrenbetäubend (deafening) returned successful results. The results of this test can be seen in tables 4 and 5. Given that some pairs of adjectives produced successful results, we believe that the failure of other test cases should not be ascribed to any shortcomings of the method employed but rather to the limited scope of our study. Further work with a larger number of distinct adjectives is needed.

5 Discussion

We demonstrated the robustness of the AdjScales process proposed for English by Sheinman and Tokunaga (2009) by successfully adapting it to German. While the sample set of adjectives used to extract patterns from our test corpus was relatively small — resulting in the aggregation of patterns that could only be applied within a limited scope — our application of the patterns to the selected adjective pairs yielded results mirroring those of Sheinman and Tokunaga (2009). Some of the patterns that we induced for German were translational equivalents of the English patterns, while others seem specific to German. Further extensions to additional languages are planned.

The broader implications of our study is that the recreation of AdjScales strongly supports the underlying semantic analysis. The lexical organization of scalar adjectives into bipolar adjectives (Bierwisch 1987; Gross and Miller 1990) extends across languages that recognize this lexical category. However, the labeling of values along the scale relative to the two poles may well differ crosslinguistically. While linguists, lexicologists and psychologists have long taken this for granted, we believe our contribution to be both novel and important in that it provides empirical evidence for understanding the meanings of specific adjectives as members of a scale. AdjScales can furthermore provide powerful tools for computational linguists and Natural Language Processing applications as well as supply the foundation for the development of more effective language reference tools.

Acknowledgments

We thank the Institut für Deutsche Sprache for making the COSMA corpus available. Fellbaum’s work was supported by U.S. National Science Foundation grant IIS 0705155.
References


