Can Human Verb Associations help identify Salient Features for Semantic Verb Classification?

Sabine Schulte im Walde
Computational Linguistics
Saarland University
Saarbrücken, Germany

University of Sussex, May 2006
Overview

Assumption: Associations model aspects of verb meaning. This knowledge is needed for semantic verb classes.

1. Human verb associations - collection and analysis
2. Association-based verb classes and validation
3. Exploring semantic class features
4. Inducing verb classes with corpus-based features

Question: Can human verb associations help identify salient features for semantic verb classification?
Variety of Semantic Verb Classes

- Semantic verb classes with various instantiations of semantic similarity, e.g.
  - syntactic similarity / alternations (Levin, 1993):
    - buy, catch, earn, find, steal, ... (obtaining:get verbs)
  - synonymy (WordNet):
    - buy, purchase
  - situation-based agreement (FrameNet):
    - buy, purchase (commerce_buy) inherits from acquire, gain, get, obtain, procure, secure (getting)
Creation of Semantic Verb Classes (1)

- Resource-intensive vs. automatic methods
- Examples of automatic methods:
  - Merlo & Stevenson (CL Journal, 2001):
    classify 60 English verbs which alternate between an intransitive and a transitive usage into three classes; features model syntactic frame alternation proportions and heuristics for semantic role assignment
  - Stevenson & Joanis (CoNNL, 2003):
    classify into 13 Levin classes with 20 verbs each in a two-/three-class distinction; general feature space and manual/seed subsets for syntactic slots, tense, voice, aspect, animacy
Creation of Semantic Verb Classes (2)

• Examples of automatic methods (cont’d):
  » Korhonen, Krymolowski & Marx (ACL, 2003):
    48+26+57 polysemous Levin/Dorr/Korhonen classes;
    subcategorisation info from Briscoe/Carroll parser
  » Schulte im Walde (PhD, 2003; CL Journal, 2006):
    classifies 168 verbs into 43 FrameNet-style classes;
    subcategorisation frames, PPs, selectional preferences

• Classification and clustering parameters:
  verbs, classes, algorithm, features, etc.
Semantic Verb Classes: Features

- Model semantic similarity of interest
- Similarity at the syntax-semantics interface: *behaviour*
- Potentially salient features:
  - syntactic frames
  - prepositional phrases
  - argument role fillers
  - adverbial adjuncts, etc.
- Granularity of features
• **Goal**: human associations to identify salient features

• **Assumptions**:
  » associations model verb meaning aspects
  » theory-independent
  » variety of semantic verb relations
  » guidance to feature selection

• Empirical model for verb features (among others): window co-occurrence vs. syntactic frame fillers
Procedure

1. Collect human verb associations
2. Hierarchical clustering on experiment verbs; basis: verb associations (assoc-classes)
3. Validate clustering against GermaNet and FrameNet
4. Compare cluster features with corpus-based features
5. Hierarchical clustering on experiment verbs; basis: corpus-based features (corpus-classes)
6. Compare corpus-classes against assoc-classes
7. Evaluation of hypothesis
Hypothesis and Research Questions

• **Hypothesis:** The more associations are found as instantiations in a feature set, the better is a clustering as based on that feature type.

• **Do the human associations help identify salient features to induce semantic verb classes?**
  I.e., do the corpus-based results outperform previous results?

• **Are the same types of features salient for different types of semantic verb classes?**
Human Verb Associations: Collection and Analysis
Web Experiment: Material

- 330 German verbs
- Variety of semantic verb classes, possible ambiguity:
  - self-motion: gehen ‘walk’, schwimmen ‘swim’
  - cause: verbrennen ‘burn’, reduzieren ‘reduce’
  - experiencing: lachen ‘laugh’, überraschen ‘surprise’
  - communication: erzählen ‘tell’, klagen ‘complain’
  - body: schlafen ‘sleep’, abnehmen ‘lose weight’
- Variety of frequency ranges (1 < freq < 71,604)
- Random distribution: 6 data sets à 55 verbs, balanced for class affiliation and frequency ranges
Web Experiment: Procedure

schneien

kalt
rodeln
Schneemann
weiß
dämmern
Web Experiment: Data

- 299 accepted data files
- Participants per data set: between 44 and 54
- Number of trials: 16,445
- Number of associations per target verb: range 0-16, average: 5.16
- Responses: 81,373 tokens for 18,884 types
### Quantification over Association Types

**klagen** ‘complain, moan, sue’

<table>
<thead>
<tr>
<th>Term</th>
<th>Translation</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gericht</td>
<td>‘court’</td>
<td>19</td>
</tr>
<tr>
<td>jammern</td>
<td>‘moan’</td>
<td>18</td>
</tr>
<tr>
<td>weinen</td>
<td>‘cry’</td>
<td>13</td>
</tr>
<tr>
<td>Anwalt</td>
<td>‘lawyer’</td>
<td>11</td>
</tr>
<tr>
<td>Richter</td>
<td>‘judge’</td>
<td>9</td>
</tr>
<tr>
<td>Klage</td>
<td>‘complaint, lawsuit’</td>
<td>7</td>
</tr>
<tr>
<td>Leid</td>
<td>‘suffering’</td>
<td>6</td>
</tr>
<tr>
<td>Trauer</td>
<td>‘mourning’</td>
<td>6</td>
</tr>
<tr>
<td>Klagemauer</td>
<td>‘Wailing Wall’</td>
<td>5</td>
</tr>
<tr>
<td>laut</td>
<td>‘noisy’</td>
<td>5</td>
</tr>
</tbody>
</table>
Linguistic Analysis of Experiment Data

• Preference for morpho-syntactic category of responses?
  → distinguish major parts-of-speech:
    nouns, verbs, adjectives, adverbs

• Typical argument holders of verb valency?
  → investigate linguistic functions realised by nouns:
    empirical grammar model (Schulte im Walde, 2003)
### Morpho-Syntactic Distribution

<table>
<thead>
<tr>
<th></th>
<th>V</th>
<th>N</th>
<th>ADJ</th>
<th>ADV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Freq</td>
<td>19.863</td>
<td>48.905</td>
<td>8.510</td>
<td>1.268</td>
</tr>
<tr>
<td>Prob</td>
<td>25</td>
<td>62</td>
<td>11</td>
<td>2</td>
</tr>
<tr>
<td>Freq</td>
<td>9.317</td>
<td>23.524</td>
<td>4.983</td>
<td>802</td>
</tr>
<tr>
<td>Prob</td>
<td>24</td>
<td>61</td>
<td>13</td>
<td>2</td>
</tr>
</tbody>
</table>
Morpho-Syntactic Correlations/Tests

- Correlation +
  
  target verb frequency ⇔ verb/adverb responses

- Correlation -
  
  target verb frequency ⇔ noun responses

- Variation across verb classes
Syntax-Semantic Functions of Nouns

• Source: statistical grammar model

• Verb valency:
  » 38 syntactic subcategorisation frames
  » plus PP information (case+preposition) → 178 frames
  » subcategorised nouns

• Example: backen ‘bake’
  » frames: $\text{NP}_{\text{nom}}$
    $\text{NP}_{\text{nom}} \text{ NP}_{\text{acc}} \ldots$
  » filler examples for $\text{NP}_{\text{nom}} \ [\text{NP}_{\text{acc}}]$: Brot ‘bread’
    Kuchen ‘cake’ …
Excursus: Statistical Grammar Model

• Head-lexicalised probabilistic context-free grammar (Charniak, 1997; Carroll and Rooth, 1998)

• 35 million words of German newspaper corpora

• Unsupervised training by EM-Algorithm (Baum, 1972)

• Robust statistical parser LoPar (Schmid, 2000)

• Corpus-based quantitative lexical information: word frequencies, linguistic functions, head-head relations
Syntax-Semantic Functions: Analysis

• Typical conceptual roles which speakers have in mind
• Look up syntactic relationships between verb and nouns
• Example:

\[
\sum = 40.5 \\
\sum = 9 \\
\sum = 43.5
\]

- Kuchen (45)
- Brot (18)
- Plätzchen (10)
- Bäcker (8)
- Brötchen (8)
- Pizza (3)
- Mutter (1)

backen
## Functions: Distributions

<table>
<thead>
<tr>
<th>Function</th>
<th>TOKEN</th>
<th>TYPES</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>S</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S V</td>
<td>1,892</td>
<td>4</td>
</tr>
<tr>
<td>S V AO</td>
<td>1,054</td>
<td>2</td>
</tr>
<tr>
<td>S V DO</td>
<td>291</td>
<td>1</td>
</tr>
<tr>
<td>S V PP</td>
<td>608</td>
<td>1</td>
</tr>
<tr>
<td><strong>AO</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S V AO</td>
<td>3,239</td>
<td>7</td>
</tr>
<tr>
<td>S V AO DO</td>
<td>840</td>
<td>2</td>
</tr>
<tr>
<td>S V AO PP</td>
<td>692</td>
<td>1</td>
</tr>
<tr>
<td><strong>DO</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S V DO</td>
<td>270</td>
<td>1</td>
</tr>
<tr>
<td>S V AO DO</td>
<td>476</td>
<td>1</td>
</tr>
<tr>
<td><strong>PP</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S V PP:in_{Dat}</td>
<td>487</td>
<td>1</td>
</tr>
<tr>
<td><strong>Unknown noun</strong></td>
<td>10,663</td>
<td>22</td>
</tr>
<tr>
<td><strong>Unknown function</strong></td>
<td>24,536</td>
<td>50</td>
</tr>
</tbody>
</table>
Syntax-Semantic Frame Inspection

Sabine Schulte im Walde
No Linguistic Function in Grammar

- *backen* ‘bake’
  Ofen ‘oven’ (19), *Mehl* ‘flour’ (17), *Weihnachten* ‘Xmas’ (15)

- *fliegen* ‘fly’
  *Urlaub* ‘vacation’ (11), *Flügel* ‘wings’ (9)

- *anfangen* ‘begin’
  *Start* ‘start’ (14), *Motivation* ‘motivation’ (3)

- *enden* ‘end’
  *Feierabend* ‘leisure-time’ (4), *Rente* ‘retirement’ (2)
Window Co-Occurrence (1)

- Corpus data: 200 million word newspaper text
- Window (left+right): 5/20/50 words, excluding symbols
- Basis: association tokens

<table>
<thead>
<tr>
<th>window</th>
<th>pos (28%)</th>
<th>neg (72%)</th>
<th>all</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>99</td>
<td>55</td>
<td>68</td>
</tr>
<tr>
<td>20</td>
<td>100</td>
<td>69</td>
<td>78</td>
</tr>
<tr>
<td>50</td>
<td>100</td>
<td>75</td>
<td>82</td>
</tr>
</tbody>
</table>
Window Co-Occurrence (2)

- Distinction with respect to window frequency

<table>
<thead>
<tr>
<th>window</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>5</th>
<th>10</th>
<th>20</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>68</td>
<td>59</td>
<td>53</td>
<td>45</td>
<td>36</td>
<td>25</td>
<td>15</td>
</tr>
<tr>
<td>20</td>
<td>78</td>
<td>72</td>
<td>68</td>
<td>61</td>
<td>52</td>
<td>42</td>
<td>28</td>
</tr>
<tr>
<td>50</td>
<td>82</td>
<td>78</td>
<td>74</td>
<td>69</td>
<td>61</td>
<td>52</td>
<td>38</td>
</tr>
</tbody>
</table>
Association Analysis: Summary

- Morpho-syntactic distribution: nouns dominate
- Properties of nouns represent argument roles of verbs
- Scene information in addition to subcategorisation
- Co-occurrence counts to supplement argument counts
- Results depend on verb frequency and semantic class
- Usage of roles and window-based nouns for distributional verb descriptions
Association-based Verb Classes: Creation and Validation
## Association Overlap

### klagen / jammern ‘moan’

<table>
<thead>
<tr>
<th>German Word</th>
<th>English Translation</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frauen</td>
<td>‘women’</td>
<td>2 / 3</td>
</tr>
<tr>
<td>Leid</td>
<td>‘suffering’</td>
<td>6 / 3</td>
</tr>
<tr>
<td>Schmerz</td>
<td>‘pain’</td>
<td>3 / 7</td>
</tr>
<tr>
<td>Trauer</td>
<td>‘mourning’</td>
<td>6 / 2</td>
</tr>
<tr>
<td>bedauern</td>
<td>‘regret’</td>
<td>2 / 2</td>
</tr>
<tr>
<td>beklagen</td>
<td>‘bemoan’</td>
<td>4 / 3</td>
</tr>
<tr>
<td>heulen</td>
<td>‘cry’</td>
<td>2 / 3</td>
</tr>
<tr>
<td>nervig</td>
<td>‘annoying’</td>
<td>2 / 2</td>
</tr>
<tr>
<td>nölen</td>
<td>‘moan’</td>
<td>2 / 3</td>
</tr>
<tr>
<td>traurig</td>
<td>‘sad’</td>
<td>2 / 5</td>
</tr>
<tr>
<td>weinen</td>
<td>‘cry’</td>
<td>13 / 9</td>
</tr>
</tbody>
</table>
Association-based Clustering

- Agglomerative (bottom-up) hierarchical clustering
- Similarity measure: skew divergence
- Merging criterion: Ward’s method (sum-of-squares)
- Hierarchy cut: 100 classes
- Cluster analysis informs about
  » classes
  » verbs
  » class features, i.e. associations
Association-based Example Classes

<table>
<thead>
<tr>
<th>Class</th>
<th>Features</th>
</tr>
</thead>
<tbody>
<tr>
<td>bedauern <code>regret´, heulen </code>cry´, jammern <code>moan´, klagen </code>complain, moan, sue´, verzweifeln <code>become desperate´, weinen </code>cry´</td>
<td>Trauer <code>mourning´, weinen </code>cry´, traurig <code>sad´, Tränen </code>tears´, jammern <code>moan´, Angst </code>fear´, Mitleid <code>pity´, Schmerz </code>pain´, etc.</td>
</tr>
<tr>
<td>abnehmen <code>lose weight´, abspecken </code>lose weight´, zunehmen `gain weight´</td>
<td>Diät <code>diet´, Gewicht </code>weight´, dick <code>fat´, abnehmen </code>lose weight´, Waage <code>scale´, Essen </code>food´, essen <code>eat´, Sport </code>sports´, dünn <code>thin´, Fett </code>fat´, etc.</td>
</tr>
</tbody>
</table>
Validation: Procedure

**Claim:** A clustering as based on verb associations and a standard clustering setting compares well with existing semantic classes.

1. Lexical resources:
   - GermaNet (Kunze, 2000)
   - Salsa / FrameNet (Erk et al., 2003)

2. Extraction of sub-classifications of resources

3. Hierarchical clustering of verb subsets

4. Pair-wise evaluation (Hatzivassiloglou/McKeown, 1993)
Validation: Classes and Verbs

<table>
<thead>
<tr>
<th></th>
<th>classes</th>
<th>verb senses</th>
<th>verbs</th>
<th>amb</th>
</tr>
</thead>
<tbody>
<tr>
<td>GermaNet</td>
<td>33</td>
<td>71</td>
<td>56</td>
<td>1.3</td>
</tr>
<tr>
<td>FrameNet</td>
<td>38</td>
<td>145</td>
<td>91</td>
<td>1.6</td>
</tr>
</tbody>
</table>
Validation Results: GermaNet

32 classes;
F = 62.69%
(UB: 82.35%)
Validation Results: FrameNet

10 classes; F = 34.68% (UB: 60.31%)
Association-based Classes: Summary

- Considerable overlap between association-based classes and the lexical resources GermaNet and FrameNet
- Differences in validation for GermaNet vs. FrameNet:
  - types of semantic similarity
  - degrees of ambiguity
  - clustering parameters: number of verbs, etc.

→ Potential use of association-based classes as gold standard for clustering experiments

→ Associations provide guidance to feature selection
Exploring
Semantic Class Features
Exploring Semantic Class Features

- **Grammar-based relations** from statistical grammar:
  verb-noun pairs with nominal heads of NPs and PPs,
  verb-adverb pairs from adverbial modifiers

- **Co-occurrence window:**
  200-million word newspaper corpus,
  20-word window (left and right)
Exploring Semantic Class Features

<table>
<thead>
<tr>
<th>features</th>
<th>grammar relations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
</tr>
<tr>
<td>cov. (%)</td>
<td>3.82</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>features</th>
<th>co-occurrence: window-20</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>all</td>
</tr>
<tr>
<td>cov. (%)</td>
<td>66.15</td>
</tr>
</tbody>
</table>
Corpus-based Clustering

- **Experiment verbs:**
  agglomerative hierarchical clustering, evaluation against assoc-classes: *accuracy*

- **GermaNet:**
  random selection of 100 synsets, random hard version with 233 verbs, clustering and evaluation as above

- **FrameNet:**
  pre-release version from May 2005, random hard version with 406 verbs in 77 classes, clustering and evaluation as above
Corpus-based Clustering: Questions

• Do the results of the clusterings with respect to the underlying feature types correspond to the overlap of associations and feature types?

• Do the corpus-based feature types which were identified on the basis of the associations outperform previous clustering results?

• Do the results generalise over the semantic class type?
**Corpus-based Clustering: Results**

<table>
<thead>
<tr>
<th>frames</th>
<th>grammar relations</th>
</tr>
</thead>
<tbody>
<tr>
<td>f-pp</td>
<td>f-pref</td>
</tr>
<tr>
<td>Assoc</td>
<td>37.50</td>
</tr>
<tr>
<td>GN</td>
<td>46.98</td>
</tr>
<tr>
<td>FN</td>
<td>33.50</td>
</tr>
</tbody>
</table>

**co-occurrence: window-20**

<table>
<thead>
<tr>
<th></th>
<th>all</th>
<th>cut</th>
<th>ADJ</th>
<th>ADV</th>
<th>N</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Assoc</td>
<td>39.33</td>
<td><strong>39.45</strong></td>
<td>37.31</td>
<td>36.89</td>
<td>39.33</td>
<td>38.84</td>
</tr>
<tr>
<td>GN</td>
<td>51.53</td>
<td>52.42</td>
<td>50.88</td>
<td>47.79</td>
<td><strong>52.86</strong></td>
<td>49.12</td>
</tr>
<tr>
<td>FN</td>
<td>missing</td>
<td>32.84</td>
<td>31.08</td>
<td>31.00</td>
<td><strong>34.24</strong></td>
<td>31.75</td>
</tr>
</tbody>
</table>
Summary of Results

- No correlation between overlap of associations / feature types and respective clustering results (Pearson’s correlation, \( p > .1 \))
- Window-based features are not significantly worse than selected grammar-based functions; applying cut-offs has almost no impact
- Several cases of grammar-based and window-based features outperform frame-based features (i.e., previous work)
- Adverbs outperform frame-based features, even some nominals
- Most successful feature types vary for gold standards
- Significantly better results for GermaNet clusterings than for experiment-based and FrameNet clusterings
Outlook

• Which feature types are appropriate to model human associations?

• Which types of (semantic) verb classifications rely on which types of features?

• Which classification parameters (e.g., size of classes, ambiguity of verbs, empirical properties of verbs) influence the clustering result?

• How do the features and parameters differ with respect to a specific semantic verb class?