Association Norms of German Noun Compounds

Overview: Noun Compounds, Compositionality, and Associations

German Noun Compounds
- Compounds: combinations of two or more simplex words; examples for German: Ahornblatt ‘maple leaf’, Nähmaschine ‘sewing machine’, Rotkohl ‘red cabbage’
- Basis: 450 concrete, depictable noun compounds from von der Heide and Borgwaldt (2009)

Compositionality
- Interest: degrees of compositionality of German noun compounds
  - cognitive and computational research on
  - lexical-semantic properties of compounds and their constituents
  - semantic relatedness between the compounds and their constituents
  - psycholinguistic research on compounds in the mental lexicon:
    - are compounds stored as single units, or are they stored decomposed?
    - are the meanings of the constituents activated during processing?

Association Norms
- Motivation: association norms have a long tradition in psycholinguistic research to investigate semantic memory
- Implicit notion: associates reflect meaning components of words
- Assumption: associations are useful for research on degrees of compound compositionality, to identify salient properties of the compound components
- Result: new lexical-semantic resource – associations of compounds and their constituents
- General procedure to collect associations and obtain association norms:
  - present target stimuli to participants in an experiment
  - participants provide associate responses, i.e., words spontaneously called to mind
  - quantification over target-association pairs: association norms

Obtaining Association Norms from a Web Experiment
- Web-experiment: 442 compound nouns and their constituents → 996 target stimuli:
  - example stimuli: Ahornblatt, Ahorn, Blatt; Nähmaschine, nähen, Maschine; Rotkohl, rot, Kohl;
  - random division into 12 separate experimental lists of 83 nouns each
- Data: 268 participants, between 14 and 28 for each data set:
  - maximum number of responses: 3 associations; average: 2.6 associations;
  - collection: 46,989/29,221 association tokens/types in 17,906 trials
  - 861 trials without response (mostly for cranberry morphemes, e.g., Him in Himbeere)
- Create association norms: for each stimulus (i.e., the compounds and their constituents), quantify over all responses in the experiment
- Two modes: perform quantification
  - (i) considering only the first response in each trial, and
  - (ii) considering all responses, disregarding the order of the associates (presented below);
- Background: association chain effects (McEvoy and Nelson, 1982), e.g., tree – leaf – to float

Part-of-Speech Tag Distribution of Responses

<table>
<thead>
<tr>
<th>N</th>
<th>ADJ</th>
<th>V</th>
<th>OTHER</th>
</tr>
</thead>
<tbody>
<tr>
<td>33,322</td>
<td>6,835</td>
<td>5,264</td>
<td>1,827</td>
</tr>
<tr>
<td>TOTAL FREQ</td>
<td>71%</td>
<td>14%</td>
<td>11%</td>
</tr>
</tbody>
</table>

| Dose ‘can’ | 92% | 3% | 5% | 0% |
| Rotenschlüssel ‘clef’ | 96% | 2% | 0% | 3% |
| Faden ‘thread’ | 43% | 10% | 44% | 3% |
| Türklinke ‘door handle’ | 54% | 3% | 38% | 5% |
| Zitrone ‘lemon’ | 20% | 74% | 3% | 3% |
| Wolfsichl ‘woolen scarf’ | 37% | 49% | 13% | 1% |

Association Frequencies for Example Noun Compounds and their Constituents (mode (ii): all associations)

 caso Study: Compositionality of German Noun Compounds
- Goal: explore whether associations to compounds provide insight into salient properties that in turn could be useful for computational models of compound compositionality
- Method: simple association overlap measure to predict the degree of compositionality of the experiment compound nouns, i.e., use proportion of shared associations of a compound and a constituent in comparison to the total number of associations of the compound
- Example: Ahornblatt receives a total of 39 associations, out of which it shares 31 with the first constituent Ahorn and 14 with the second constituent Blatt.
- Thus, the predicted degrees of compositionality are 31/39 = 0.79 for Ahornblatt-Ahorn, and 14/39 = 0.36 for Ahornblatt-Blatt.
- The predicted degrees of compositionality (system scale: 0 to 1) are compared against the mean compositionality judgements (human scale: 1=opaque to 7=strong compositionality) by von der Heide and Borgwaldt, using the Spearman rank-order correlation coefficient.
- Surprisingly successfully correlation: \( r_s = 0.5228, p < .00001 \)