DURel Annotation Tool
Measuring Patterns of Contextual Word Meaning over Time

May 14, 2021

Dominik Schlechtweg and Maike Park
Institute for Natural Language Processing, University of Stuttgart
Leibniz-Institute for German Language, Mannheim
Motivation

- common problem:
  - given: set of word uses (corpus)
  - searched: their meanings and their relations

- relevant for:
  - historical linguistics
  - lexicography
  - digital humanities

- common approach: researcher scans corpus himself
  - tedious
  - subjective
  - no protocol
  - bias

- solution: DURel Annotation Tool¹
  - online interface
  - upload word uses for annotation
  - well-established protocol for contextual word meaning annotation

¹https://www.ims.uni-stuttgart.de/data/durel-tool
Data

Table 1: Sample of diachronic corpus.

<table>
<thead>
<tr>
<th>Year</th>
<th>Year</th>
<th>Text</th>
</tr>
</thead>
<tbody>
<tr>
<td>1824</td>
<td>1842</td>
<td>and taking a knife from her pocket, she opened a vein in her little <em>arm</em>, And those who remained at home had been heavily taxed to pay for the <em>arms</em>, ammunition; and though he saw her within reach of his <em>arm</em>, yet the light of her eyes seemed as far off</td>
</tr>
<tr>
<td>1860</td>
<td></td>
<td>and though he saw her within reach of his <em>arm</em>, yet the light of her eyes seemed as far off</td>
</tr>
<tr>
<td>1953</td>
<td></td>
<td>overlooking an <em>arm</em> of the sea which, at low tide, was a black and stinking mud-flat</td>
</tr>
<tr>
<td>1975</td>
<td></td>
<td>twelve miles of coastline lies in the southwest on the Gulf of Aqaba, an <em>arm</em> of the Red Sea.</td>
</tr>
<tr>
<td>1985</td>
<td></td>
<td>when the disembodied <em>arm</em> of the Statue of Liberty jets spectacularly out of the</td>
</tr>
</tbody>
</table>
(A) [...] and taking a knife from her pocket, she opened a vein in her little arm, and dipping a feather in the blood, wrote something on a piece of white cloth, which was spread before her.

(D) It stood behind a high brick wall, its back windows overlooking an arm of the sea which, at low tide, was a black and stinking mud-flat [...]

Annotation
Scale

4: Identical
3: Closely Related
2: Distantly Related
1: Unrelated

Table 2: DURel relatedness scale.
Figure 1: Word Usage Graph of English *arm*.
Clustering

Figure 2: Word Usage Graph of English arm.
Lexical Semantic Change
Case study: Polysemy across Language Varieties

- Baldissin et al. (submitted)
- Diatopic lexical semantic variation in Spanish
- extend DURel framework to onomasiological questions
Case study: Polysemy across Language Varieties

Figure 4: Word Usage Graph of Spanish *colectivo* and *guagua*.
Case study: Polysemy across Language Varieties

Spain

Argentina
Case Study: Lexical Semantic Change Discovery

- Kurtyigit et al. (2021)
- focus on change discovery:
  - discovering novel word senses over time in a diachronic corpus pair
  - evaluating visualizations of the annotated data from a lexicographer’s point of view (how intuitive is it? are clusters conclusive? annotations reliable?)
- results:
  - high-quality predictions, high inter-annotator agreement
  - useful visualizations of clusters and relations
  - detection of previously undescribed changes that weren’t included in dictionaries
Case Study: Lexical Semantic Change Discovery

Figure 6: Word Usage Graph of German Zehner.
Case Study: Lexical Semantic Change Discovery

$t_1$

$t_2$
Upcoming case study

- target: words that recently underwent semantic change (semantic neologisms)
- common problem: difficult to detect
  - quantitative criteria of (semi-)automatic approaches are geared towards the lexicalization process of new words (Falk et al., 2014; Fišer & Ljubešić, 2016; Klosa & Lüngen, 2018)
  - tend to rely on frequency measures
- common problem: difficult to define
  - what does a majority of speakers perceive as new meaning (and what not)?
  - which type of corpus data represents common language the most adequately?
Targeting something really tricky to find

► target: verbs from a dictionary of neologisms in Contemporary German language

► specific problems:
  ► semantic change ranges from widening/narrowing to metaphoric transfer
  ► (the meaning of) a verb changes and spreads slower
  ► verbs tend to be extremely infrequent in corpora used for the detection of neologisms (covering five to 10 years)

► data: newspaper texts (representing use of general German language)

► objective:
  ► detect infrequent novel senses of verbs in small corpora
  ► evaluate the semantic proximity between old and new senses (long-term goal: enhancing the objectivity of lexicographical decisions)
Conclusion

- **inter-subjectivity:**
  - avoids experimenter bias through standard protocol and annotation by multiple humans
  - inter-annotator agreement gives measure of reliability

- **simple:**
  - the judgment of use pair relatedness is an intuitive task for annotators generally yielding high agreement (Erk et al., 2013; Schlechtweg et al., 2018)
  - annotated data can be visualized as semantic relatedness graphs on 2D plots

- **preparation-lean:**
  - researchers only need to sample word uses

- **grounded in theory:**
  - relatedness judgments have theoretical basis in cognitive semantics (Blank, 1997; Schlechtweg et al., 2018)

- **flexible:**
  - clustering algorithm and parameters can be changed after annotation, avoiding re-annotation
References I


