**Type disambiguation of English -ment derivatives**

Gabriella Lapesa¹, Lea Kawaletz², Marios Andreou² | Max Kisselew¹, Sebastian Pado¹, Ingo Plag²

University of Stuttgart¹, Heinrich-Heine University Düsseldorf²

---

### Derivation and polysemy

Derived words tend to be polysemous

- We focus on *ment* derivatives:
  - Eventive (assessment) vs. non eventive (pavement) reading
  - Context often determines the reading of a derived word
  - Some nominalizations, however, remain ambiguous even in context

**Question:** to which extent does context determine the readings of derived words?

- We try to answer this question for newly derived, non-lexicized words, using manual annotation and corpus-based modeling

---

### Corpus-based modeling

**Distributional Semantics (DS):**

difference in meaning = difference in distribution

DS meaning → word vector

- list of words which occur in the context of a target

**LEMMA VECTORS**

wear
t-shirt
tie
t-shirt
tie

**CONTEXT VECTORS**

weart-shirt
judge
lawyer
law

**Strategy:** use context words to approximate the meaning of the target when the target vector is unreliable (low frequency) or unattested

---

### Experiments & Results

- **Training nouns:** WordNet
  - EVENTIVE:
    - state, feeling, process, phenomenon, event, act
  - NON EVENTIVE:
    - STRICT OBJECT: object, substance, food, location, artefact, body
    - LAX OBJECT: communication, quantity, relation, social relation, possession
  - LIVING: person, animal, plant

- **Corpus:** BNC + Ukwac + Wikipedia
- **Lemma vectors:** state-of-the-art DS model
- **Context vectors:** window size 2 (most immediate context) vs. 8 (wider context: more info, more noise?)
- **Classifier:** support vector classifier (tested in many configurations)
  - F-score: high if the classifier captures all instances of class A and not too many non-instances of A

---

### The -ment dataset

**Selection of sentences** containing low-frequency -ment nouns:

- **Sources:** COCA, GloWBE, WebCorp, BNC, Wikipedia, Google, ...
- **Data:** 401 tokens, 56 types
- **Base verbs:** change-of-state, force, psych, putting

**Annotation:**

- EVENTIVE: In many places, emplacement of granite plutons is synchronous to volcanic eruptions
- NON EVENTIVE: I set down the scrap of doll's dress, a bedragglement of loose lace hem
- AMBIGUOUS: [...] when it is evoked by a man who has suffered its most horrible debauchments

---

### Workflow: DS & Machine Learning

**TRAINING SENTENCES**

<table>
<thead>
<tr>
<th>Accident</th>
<th>Bike</th>
<th>Event</th>
</tr>
</thead>
</table>
| Accident.23: EVT | “An accident happened in the night” | BIKE.17: NON.EVT
| The kid fixed the wheel of his bike |

**TEST SENTENCES**

<table>
<thead>
<tr>
<th>Accident.23: EVT</th>
<th>Bike</th>
<th>Night</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accident.3: EVT</td>
<td>night</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>happen</td>
<td>8</td>
<td>4</td>
<td>1</td>
</tr>
<tr>
<td>wheel</td>
<td>2</td>
<td>20</td>
<td>1</td>
</tr>
<tr>
<td>fix</td>
<td>3</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>accident</td>
<td>10</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>bike</td>
<td>4</td>
<td>8</td>
<td>1</td>
</tr>
</tbody>
</table>

**Analysis 1:**

- **Question:** can we correctly classify training nouns and their contexts as EVT vs. NON EVT?
  - **Result:** yes, very well
    - All classifiers are above baseline
    - Performance matches linguistic intuitions
    - Context vectors: lower performance than lemma vectors (task more difficult: target not used)
    - Slight preference for larger windows

**Analysis 2:**

- **Question:** based on context, can we correctly classify -ment instances?
  - **Result:** yes, but the task is difficult
    - (Lemma vectors are not available)
      - All classifiers above baseline
      - Small window the best choice: newly derived words have more informative immediate context than training nouns
    - AMBIGUOUS mostly classified as EVT

**DS vectors can successfully disambiguate newly derived -ment nominalizations.** Future work:

- qualitative evaluation of the predictions
- quantitative investigation of the factors which determine the predictions of the classifier (e.g., frequency)

---

The authors gratefully acknowledge financial support by the Deutsche Forschungsgemeinschaft Collaborative Research Centers 732 (project B9: Lapesa, Kisselew, Pado) and 991 (project C08: Kawaletz, Andreou, Plag).

Contacts: gabriella.lapesa@ims.uni-stuttgart.de, lea.kawaletz@hhu.de
References


• Harris, Z. (1954). Distributional structure. Word, 10(23), 146–162.


Contacts: gabriella.lapesa@ims.uni-stuttgart.de, lea.kawaletz@hhu.de, Marios.Andreou@uni-duesseldorf.de, max.kisselew@ims.uni-stuttgart.de, pado@ims.uni-stuttgart.de, Ingo.Plag@uni-duesseldorf.de