Type disambiguation of English *-ment* derivatives

Gabriella Lapesa[1], Lea Kawaletz[2], Marios Andreou[2] Max Kisselew[1], Sebastian Pado[1], Ingo Plag[2]
University of Stuttgart[1], Heinrich-Heine University Düsseldorf[2]

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-lexicalized words, using **manual annotation** and **corpus-based modeling**

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***

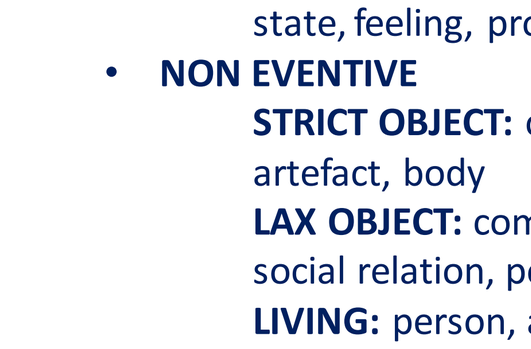
Corpus-based modeling

Distributional Semantics (DS):

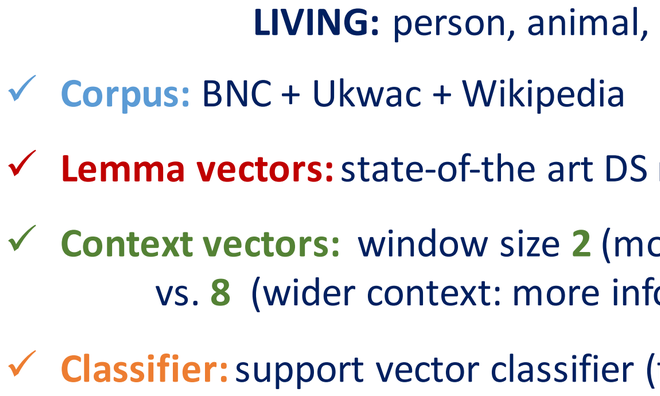
difference in meaning = difference in distribution

DS meaning → word vector

list of words which occur in the context of a target

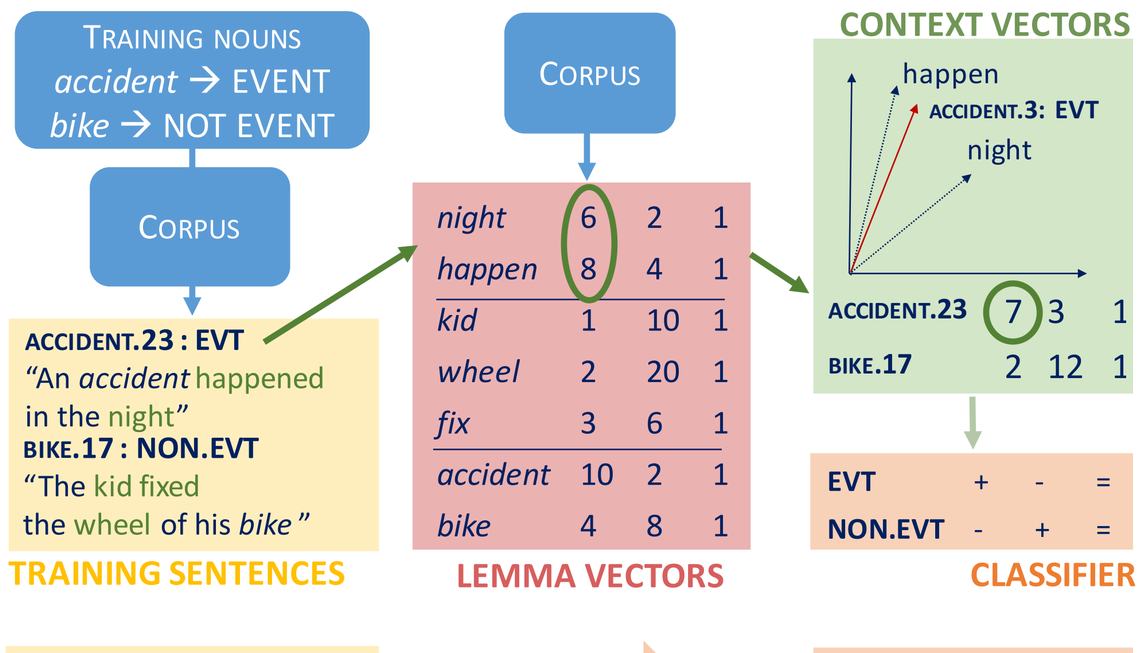


DS **successful** in modeling word similarity **but…**
known challenges: polysemy and unattested or low frequency words



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

Workflow: DS & Machine Learning

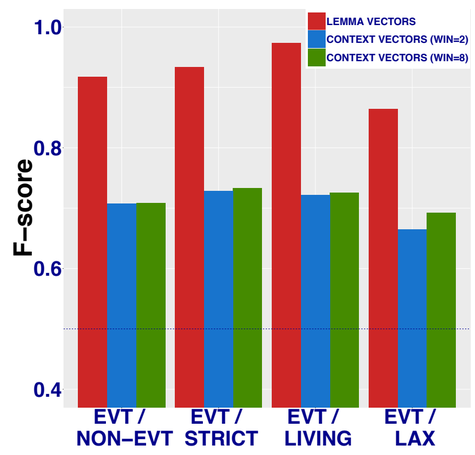


Experiments & Results

- ✓ **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

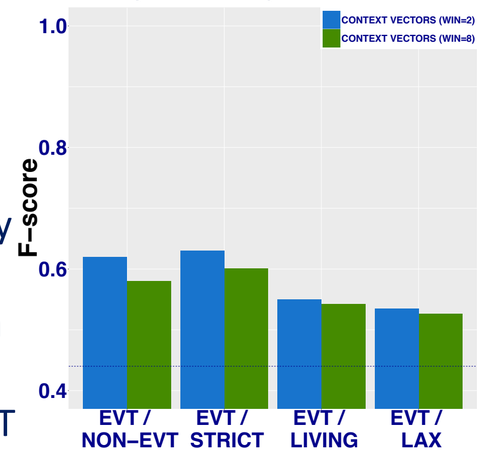
ANALYSIS 1: can we correctly classify training nouns and their contexts as EVT vs. NON EVT?
Result: yes, very well

- 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: 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

- Firth, J. (1957). A synopsis of linguistic theory 1930-1955. Studies in Linguistic Analysis (special volume ofthe Philological Society), 1952-59, 1-32.
- Harris, Z. (1954). Distributional structure. Word, 10(23), 146–162.
- Kipper-Schuler, K. 2005. VerbNet: A broad-coverage, comprehensive verb lexicon. PhD. Thesis. Computer and Information Science Dept., University of Pennsylvania. Philadelphia, PA.
- Levin, B. (1993): English verb classes and alternations. A preliminary investigation. Chicago: University of Chicago Press.
- Omer Levy and Yoav Goldberg (2014) Dependency-based word embeddings. Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Short Papers), pages 302–308.
- Lieber, R. (2016). English nouns: The ecology of nominalization. Cambridge: Cambridge University Press.
- Plag, I., Andreou, M., & Kawaletz, L. (forthcoming). A frame-semantic approach to polysemy in affixation. In The lexeme in descriptive and theoretical morphology.
- Rainer, Franz (2014): Polysemy in derivation. In Rochelle Lieber, Pavol Štekauer (Eds.): The Oxford Handbook of Derivational Morphology. 1. ed. Oxford: Oxford Univ. Press (Oxford handbooks in linguistics), pp. 338–353.
- Schütze, H. (1998). Automatic word sense discrimination. Computational Linguistics, 27(1), 97–123.

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