University of Stuttgart
Germany

CHANGE
IS KEY

Lexical Semantic Change Detection

January 19, 2024

Dominik Schlechtweg
Institute for Natural Language Processing, University of Stuttgart

Introduction

- Lexical Semantic Change Detection (Schlechtweg, 2023)
  - goal: automate the analysis of changes in word meanings over time

    (1) *Der zweyte Theil vom Bauernrechte ist schon lange aus der **Presse**;*
    'The second part of Farmers' Rights already left the **press**;'

    (2) *Alle Freiheiten suspendirt! die persönliche Freiheit wie die der **Presse**!*
    'All freedoms suspended! the personal freedom as well as the one of the **press**!'

Human Measurement of Lexical Semantic Change

A	1824	and taking a knife from her pocket, she opened a vein in her little **arm**,	☺
B	1842	And those who remained at home had been heavily taxed to pay for the **arms**, ammunition;	⚔
C	1860	and though he saw her within reach of his **arm**, yet the light of her eyes seemed as far off	☺
		. . .	
D	1953	overlooking an **arm** of the sea which, at low tide, was a black and stinking mud-flat	
E	1975	twelve miles of coastline lies in the southwest on the Gulf of Aqaba, an **arm** of the Red Sea.	
F	1985	when the disembodied **arm** of the Statue of Liberty jets spectacularly out of the	☺

Table 1: Sample of diachronic corpus.

Word Use Pairs

(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 [. . .]

Semantic Proximity Scale

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

Table 2: DURel relatedness scale.

Graph representation

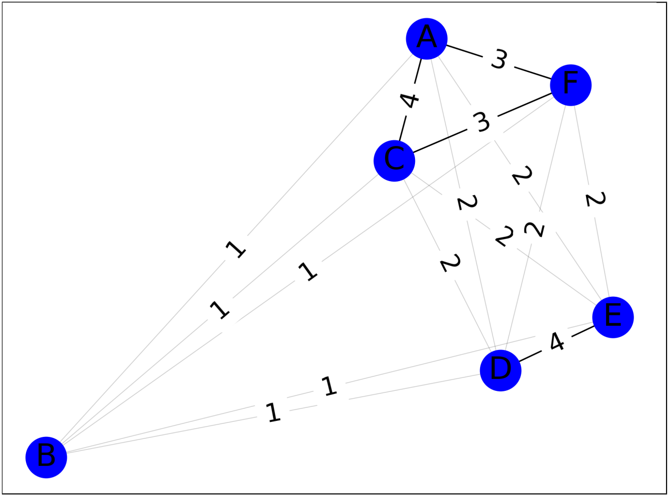


Figure 1: Word Usage Graph of English *arm*.

Clustering

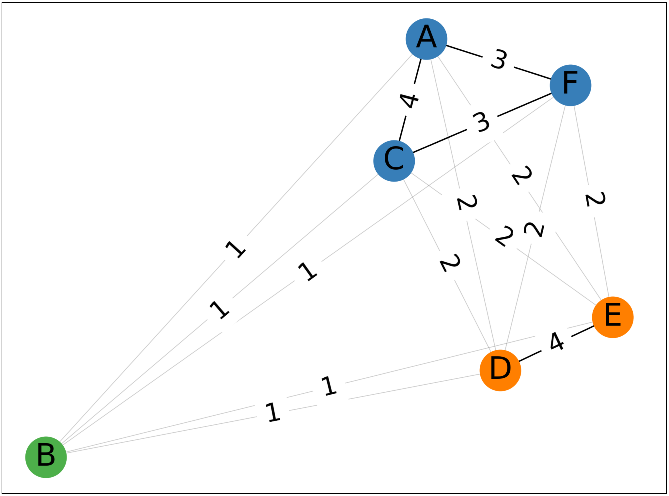
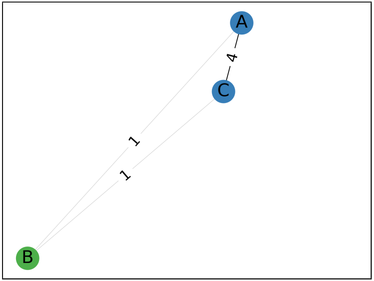
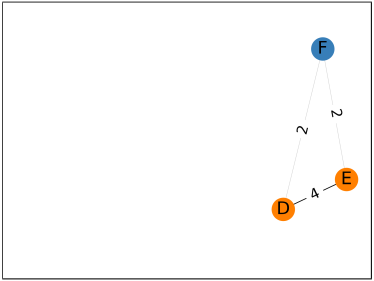


Figure 2: Word Usage Graph of English *arm*. $D = (3, 2, 1)$.

Lexical Semantic Change



$t_1,\ D_1 = (2, 0, 1)$



$t_2,\ D_2 = (1, 2, 0)$

Change Scores

- **binary change** (loss and gain of senses)
- **graded change** (changes in sense probabilities)

Evaluation Tasks

Task 1 Binary classification: for a set of target words, predict the binary change score

Task 2 Ranking: rank a set of target words according to their graded change score

(Schlechtweg et al., 2020)

Example: Swedish *ledning*[1]

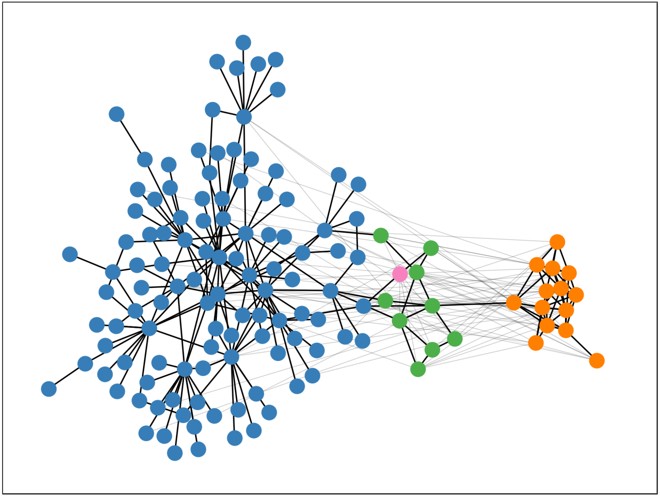


Figure 4: WUG of Swedish *ledning*.

[1] Datasets available at `https://www.ims.uni-stuttgart.de/data/wugs`

Example: Swedish *ledning*

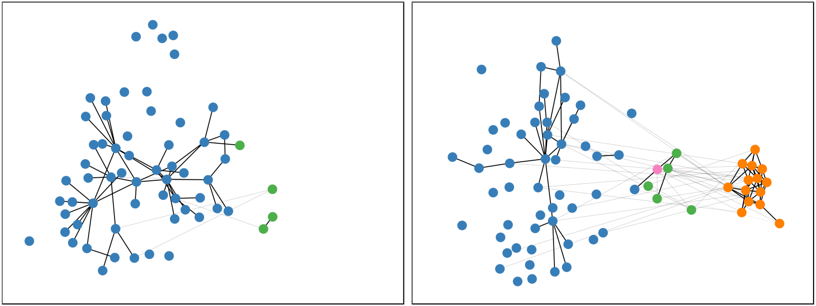


Figure 5: WUGs of Swedish *ledning*: subgraphs for 1st time period G_1 (left) and 2nd time period G_2 (right). $D_1 = (58, 0, 4, 0)$, $D_2 = (52, 14, 5, 1)$, $B(w) = 1$ and $G(w) = 0.34$.

Example: German *Eintagsfliege*

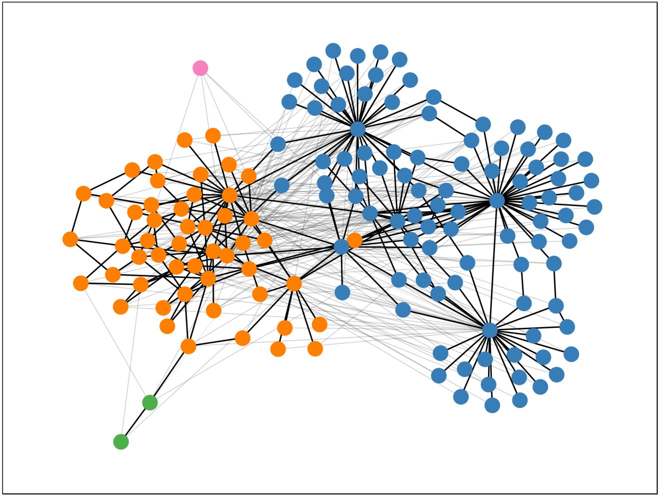


Figure 6: WUG of German *Eintagsfliege*.

Example: German *Eintagsfliege*

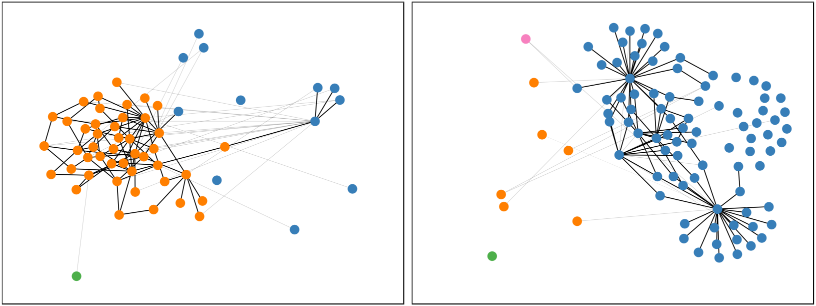


Figure 7: WUG of German *Eintagsfliege*: subgraphs for 1st time period G_1 (left) and 2nd time period G_2 (right). $D_1 = (12, 45, 0, 1)$, $D_2 = (85, 6, 1, 1)$, $B(w) = 0$ and $G(w) = 0.66$.

Summary of Annotation Steps

1. semantic proximity labeling
2. clustering
3. change measurement

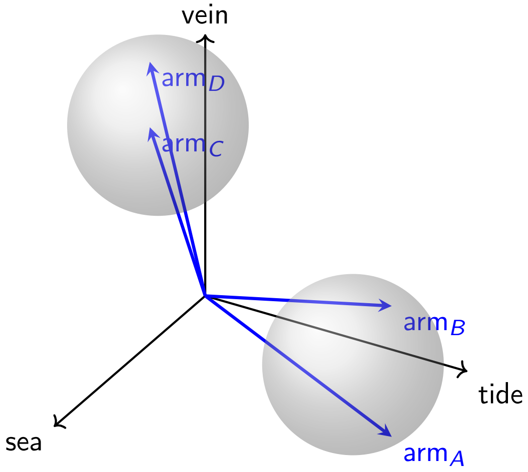
Summary of Annotation Steps with Tasks

1. semantic proximity labeling ↔ **Word-in-Context Task**
2. clustering ↔ **Word Sense Induction**
3. change measurement ↔ **Lexical Semantic Change Detection** (including previous tasks)

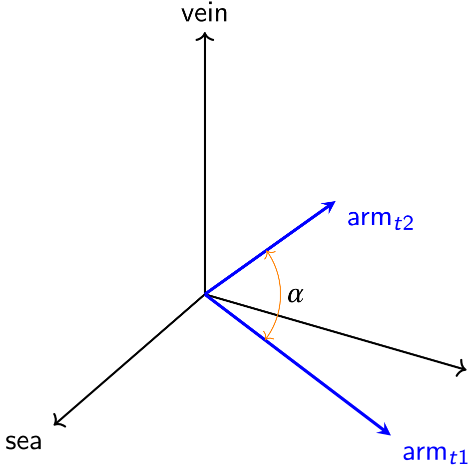
Computational Measurement of Lexical Semantic Change

- Typical token-based Model is composed by
  1. semantic proximity model (e.g. similarity between contextualized embeddings)
  2. clustering method (optional)
  3. change measure
  - model the human measurement process
  - one vector per word use (BERT, ELMo)
- Typical type-based Model is composed by
  1. semantic representation per word (type vector)
  2. alignment
  3. measure
  - do not model the human measurement process
  - one average vector per word (Word2Vec, GloVe)

Simple token-based Model



Simple type-based Model

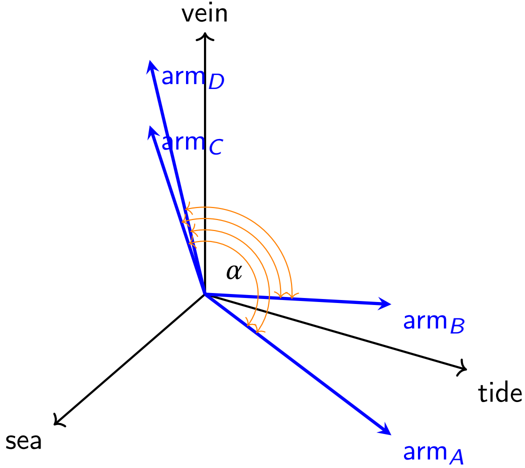


Results

Lang.	Binary	Graded
English	.70	.76
German	.70	.88
Latin	.79	.57
Swedish	.64	.75
Russian	-	.80
Spanish	.72	.74

Table 3: Current SOTA performances on LSCD tasks (Cassotti et al., 2023; Rachinskiy & Arefyev, 2022; Schlechtweg et al., 2020). Values give F1 for binary change and Spearman for graded change.

SOTA Model for graded change: APD



SOTA Model for binary change: APD + thresholding

Word	APD
word1	1.9
word2	1.7
word3	1.4
word4	1.1
word5	0.7
word6	0.6
word7	0.4

Thresholding for binary change prediction.

SOTA Model for binary change: APD + thresholding

Word	APD
word1	1.9
word2	1.7
word3	1.4
word4	1.1
word5	0.7
word6	0.6
word7	0.4

Thresholding for binary change prediction.

SOTA Model for binary change: APD + thresholding

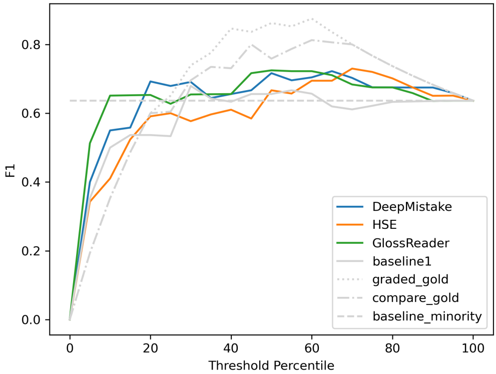


Figure 8: F1 scores over binarization thresholds based on percentiles on submitted Graded Change predictions for top four teams in evaluation phase 1 in Zamora-Reina et al. (2022).

Application

Discovery Task Classification: Decide for a **large set of unseen** words which ones lost or gained senses

(Kurtyigit et al., 2021)

Results

System	Performance
type	.71
token	.62
random	.35

Table 4: Performance type- and token-based compared to random baseline.

Discovered Change

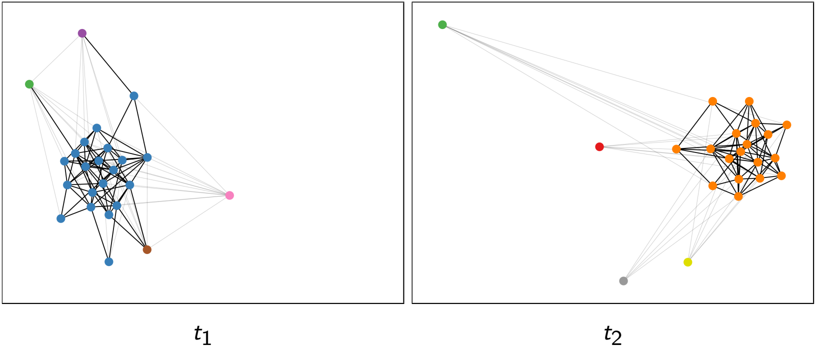


Figure 9: Word Usage Graph of German *Zehner*.

Two Uses of German *Zehner*

(1) *Man sieht also, daß die Striche nach den Tausenden, nach den Hunderten und nach den* ***Zehnern*** *gesetzt werden.*
'So you can see that the strokes are placed after the thousands, after the hundreds, and after the **tens**.'

(2) *Fußball-Toto : Kein Elfer ; 6* ***Zehner*** *mit je 3778 Mark ; 152 Neuner mit je 298 Mark.*
'Soccer lottery : No eleven ; 6 **tens** with 3778 Mark each ; 152 nines with 298 Mark each.'

Conclusion

- LSCD is a **valid and meaningful NLP task** which can be solved reasonably well with computers
- current models show medium to high performance
- WiC models for semantic proximity have lead to quantum leap in performance
- both prevalent model types discover new semantic changes with above-random probability
- **problem**: current SOTA models do not model the annotation process

Open Problems

- model the annotation process → improve WSI step
- improve data quality:
  1. add more annotations
  2. clean existing data sets
  3. use alternative annotation strategies
- multiple time points
- onomasiological perspective
- multi-word expressions
- types of change
- qualify detected changes
- measure causes

References I

Cassotti, P., Siciliani, L., de Gemmis, M., Semeraro, G., & Basile, P. (2023, July). Xl-lexeme: Wic pretrained model for cross-lingual lexical semantic change. In *Proceedings of the 61th annual meeting of the association for computational linguistics.* Online: Association for Computational Linguistics.

Kurtyigit, S., Park, M., Schlechtweg, D., Kuhn, J., & Schulte im Walde, S. (2021, aug). Lexical Semantic Change Discovery. In *Proceedings of the 59th annual meeting of the association for computational linguistics and the 11th international joint conference on natural language processing (volume 1: Long papers).* Online: Association for Computational Linguistics. Retrieved from https://aclanthology.org/2021.acl-long.543/

Rachinskiy, M., & Arefyev, N. (2022). GlossReader at LSCDiscovery: Train to select a proper gloss in english – discover lexical semantic change in spanish. In *Proceedings of the 3rd international workshop on computational approaches to historical language change.* Dublin, Ireland: Association for Computational Linguistics.

Schlechtweg, D. (2023). *Human and computational measurement of lexical semantic change* (Doctoral dissertation, University of Stuttgart, Stuttgart, Germany). Retrieved from http://dx.doi.org/10.18419/opus-12833

Schlechtweg, D., McGillivray, B., Hengchen, S., Dubossarsky, H., & Tahmasebi, N. (2020). SemEval-2020 Task 1: Unsupervised Lexical Semantic Change Detection. In *Proceedings of the 14th International Workshop on Semantic Evaluation.* Barcelona, Spain: Association for Computational Linguistics. Retrieved from https://www.aclweb.org/anthology/2020.semeval-1.1/

Zamora-Reina, F. D., Bravo-Marquez, F., & Schlechtweg, D. (2022). LSCDiscovery: A shared task on semantic change discovery and detection in Spanish. In *Proceedings of the 3rd international workshop on computational approaches to historical language change.* Dublin, Ireland: Association for Computational Linguistics. Retrieved from https://aclanthology.org/2022.lchange-1.16/