Grammatical Profiling for Semantic Change Detection

Mario Giulianelli, Andrey Kutuzov, Lidia Pivovarova

Research Seminar in Language Technology, November 18th, 2021

Contents

1. Motivation
2. Task definition
3. Grammatical profiling for semantic change detection
4. Morphosyntax methods for semantic change detection
5. Yes, it works: results
6. Qualitative analysis
   - When it works?
   - When it does not work?
7. Category importance
8. Summing up

Motivation

- Semantics, morphology and syntax are strongly interdependent
- Semantic change <–> changes in the distribution of grammatical features (no matter the causal direction)

Motivation

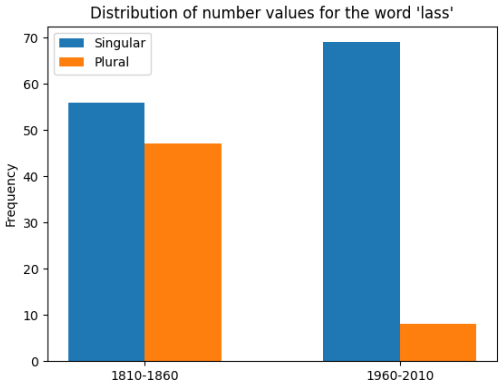
- Semantics, morphology and syntax are strongly interdependent
- Semantic change <–> changes in the distribution of grammatical features (no matter the causal direction)

English noun '*Lass*':

1. Older times: 'YOUNG WOMAN' sense more dominant ('*lasses are dancing*')
2. 20th century: 'SWEETHEART' sense more dominant ('*the young hero and his lass*')

A sharp decrease in plural usages!

Are there systematic correlations between diachronic semantic change and morphosyntactic changes?



(English corpora of SemEval 2020)

Contents

1. Motivation
2. Task definition
3. Grammatical profiling for semantic change detection
4. Morphosyntax methods for semantic change detection
5. Yes, it works: results
6. Qualitative analysis
   - When it works?
   - When it does not work?
7. Category importance
8. Summing up

Task definition

- Data:
  - A *corpus* compiled of documents from at least two time periods
  - A list of *target words*
- Task:
  - *Rank target words* according a degree of semantic shift
  - in an *unsupervised* way, using only corpus.

Example: SemEval dataset

Changed			Not-changed		
beatus	1	1.870	hostis	0	1.890
poena	1	1.800	dubius	0	1.851
senatus	1	1.490	imperator	0	1.638
civitas	1	1.437	sapientia	0	1.433
voluntas	1	1.407	simplex	0	1.428
sacramentum	1	1.131	scriptura	0	1.427
humanitas	1	1.119	nobilitas	0	1.270
sanctus	1	1.110	honor	0	1.144
dolus	1	0.975	fidelis	0	1.137
nepos	1	0.880	jus	0	1.115
virtus	1	0.863	acerbus	0	0.817
necessarius	1	0.816	adsumo	0	0.720
regnum	1	0.668	dux	0	0.561
itero	1	0.667	potestas	0	0.548
salus	1	0.553	credo	0	0.519
licet	1	0.509	cohors	0	0.250
oportet	1	0.496	ancilla	0	0.143
pontifex	1	0.492	consul	0	0.128
sensus	1	0.402	titulus	0	0.126
templum	1	0.295			
consilium	1	0.160			

Semantic change detection tasks

Standard subtasks

- Subtask 2: ranking (graded change detection)
- Subtask 1: binary classification (any senses gained or lost?)

1. *SemEval* dataset: both subtasks, English, German, Latin, Swedish [Schlechtweg et al., 2020]
2. *EvaLita* dataset: Subtask 1 only, Italian [Basile et al., 2020]
3. *RuShiftEval* dataset: Subtask 2 only, Russian [Kutuzov and Pivovarova, 2021]

Semantic change detection tasks

Standard subtasks

- Subtask 2: ranking (graded change detection)
- Subtask 1: binary classification (any senses gained or lost?)

1. *SemEval* dataset: both subtasks, English, German, Latin, Swedish [Schlechtweg et al., 2020]
2. *EvaLita* dataset: Subtask 1 only, Italian [Basile et al., 2020]
3. *RuShiftEval* dataset: Subtask 2 only, Russian [Kutuzov and Pivovarova, 2021]

In total, 274 manually annotated words from 3 Indo-European language groups: Italic, Germanic and Slavic.

Contents

1. Motivation
2. Task definition
3. Grammatical profiling for semantic change detection
4. Morphosyntax methods for semantic change detection
5. Yes, it works: results
6. Qualitative analysis
   - When it works?
   - When it does not work?
7. Category importance
8. Summing up

Profiling

- ***Grammatical profile***: a set of preferred morphological forms for a word

Profiling

- ***Grammatical profile***: a set of preferred morphological forms for a word
- Profiling, i.e. studying profile differences, is used in corpus linguistics to reveal subtle semantic differences: [Hanks, 1996], [Gries and Divjak, 2009], [Janda and Lyashevskaya, 2011], etc.

Profiling

- ***Grammatical profile***: a set of preferred morphological forms for a word
- Profiling, i.e. studying profile differences, is used in corpus linguistics to reveal subtle semantic differences: [Hanks, 1996], [Gries and Divjak, 2009], [Janda and Lyashevskaya, 2011], etc.
- Theoretical roots in construction grammar, cognitive linguistics and priming theory

Profiling

- ***Grammatical profile***: a set of preferred morphological forms for a word
- Profiling, i.e. studying profile differences, is used in corpus linguistics to reveal subtle semantic differences: [Hanks, 1996], [Gries and Divjak, 2009], [Janda and Lyashevskaya, 2011], etc.
- Theoretical roots in construction grammar, cognitive linguistics and priming theory
- **It is theoretically plausible that diachronic word meaning change is reflected in a change of profiles**

Profiling

- ***Grammatical profile***: a set of preferred morphological forms for a word
- Profiling, i.e. studying profile differences, is used in corpus linguistics to reveal subtle semantic differences: [Hanks, 1996], [Gries and Divjak, 2009], [Janda and Lyashevskaya, 2011], etc.
- Theoretical roots in construction grammar, cognitive linguistics and priming theory
- **It is theoretically plausible that diachronic word meaning change is reflected in a change of profiles**
- However, the majority of previous methods for semantic change detection rely on distributional word embeddings

Profiling

- ***Grammatical profile***: a set of preferred morphological forms for a word
- Profiling, i.e. studying profile differences, is used in corpus linguistics to reveal subtle semantic differences: [Hanks, 1996], [Gries and Divjak, 2009], [Janda and Lyashevskaya, 2011], etc.
- Theoretical roots in construction grammar, cognitive linguistics and priming theory
- **It is theoretically plausible that diachronic word meaning change is reflected in a change of profiles**
- However, the majority of previous methods for semantic change detection rely on distributional word embeddings
- The recent exception is [Ryzhova et al., 2021]: case and number frequency distributions to reveal diachronic change in Russian nouns

Our findings in short

- Tracing changes in the morphosyntactic categories does outperform count-based distributional models.
- Morphological and syntactic categories are complementary.
- Useful categories are language-dependent.
- Predictions are interpretable.

https://aclanthology.org/2021.conll-1.33/

Contents

1. Motivation
2. Task definition
3. Grammatical profiling for semantic change detection
4. Morphosyntax methods for semantic change detection
5. Yes, it works: results
6. Qualitative analysis
   - When it works?
   - When it does not work?
7. Category importance
8. Summing up

Morphosyntax methods for semantic change detection

Basic procedure 1/2

- Target historical corpora are tagged and parsed with *UDPipe*
- *A profile* is a dictionary of frequencies of morphosyntactic categories for each target word
- A dictionary is treated as a vector of values and semantic change estimated as *cosine distance* between feature vectors in two time periods
- **Basic procedure:**

`Tense=Pres|VerbForm=Part`	50
`Mood=Ind|Tense=Past|VerbForm=Fin`	24
`Tense=Past|VerbForm=Part|Voice=Pass`	17
`VerbForm=Inf`	9
`Mood=Ind|Tense=Pres|VerbForm=Fin`	1
`Tense=Past|VerbForm=Part`	1

$\longrightarrow$ cosine

Morphosyntax methods for semantic change detection

Basic procedure 1/2

- Target historical corpora are tagged and parsed with *UDPipe*
- *A profile* is a dictionary of frequencies of morphosyntactic categories for each target word
- A dictionary is treated as a vector of values and semantic change estimated as *cosine distance* between feature vectors in two time periods
- **Basic procedure:**

`Tense=Pres|VerbForm=Part`	50
`Mood=Ind|Tense=Past|VerbForm=Fin`	24
`Tense=Past|VerbForm=Part|Voice=Pass`	17
`VerbForm=Inf`	9
`Mood=Ind|Tense=Pres|VerbForm=Fin`	1
`Tense=Past|VerbForm=Part`	1

$\longrightarrow$ cosine

Feature vectors are actually time-dependent grammatical profiles.