Semantic Change Detection

Anna Marakasova

TU Wien

Diachronic Dynamics of Lexical Networks - DYLEN (with Uni Wien and ÖAW)

Goals:

(1) Investigate network-based methods for semantic shift detection,

(2) Define factors that bring about semantic change.

Diachronic Dynamics of Lexical Networks - DYLEN (with Uni Wien and ÖAW)

Goals:
(1) Investigate network-based methods for semantic shift detection,
(2) Define factors that bring about semantic change.

Data:
(1) Austrian Media Corpus (AMC), > 30 years, 11 billion tokens.
(2) The corpus of Austrian parliamentary records (ParlAT), 20 years, 75 million tokens.

Outline

1. Semantic change detection methods
2. Current challenges
3. Our experiments
   a. Evaluation of different methods on various datasets
   b. Frequency change effect
   c. Graph-based approach
4. Conclusion

1. Semantic change detection

Computational detection of semantic change

- A huge increase of interest within the last few years [Tahmasebi et. al. 2018, Kutuzov et. al. 2018].

Computational detection of semantic change

- A huge increase of interest within the last few years [Tahmasebi et. al. 2018, Kutuzov et. al. 2018].
- Mainly between diachronically sliced text corpora, but also synchronic change between literary genres / topic domains.

Computational detection of semantic change

- A huge increase of interest within the last few years [Tahmasebi et. al. 2018, Kutuzov et. al. 2018].

- Mainly between diachronically sliced text corpora, but also synchronic change between literary genres / topic domains.

  # 'nice' - to describe an ignorant or foolish person,
  'nervous' - of or belonging to the nerves.

Computational detection of semantic change

- A huge increase of interest within the last few years [Tahmasebi et. al. 2018, Kutuzov et. al. 2018].
- Mainly between diachronically sliced text corpora, but also synchronic change between literary genres / topic domains.

  # 'nice' - to describe an ignorant or foolish person,
  'nervous' - of or belonging to the nerves

  # 'mouse' in biological context vs. tech context,
  'apple' in food/diet context vs. tech context.

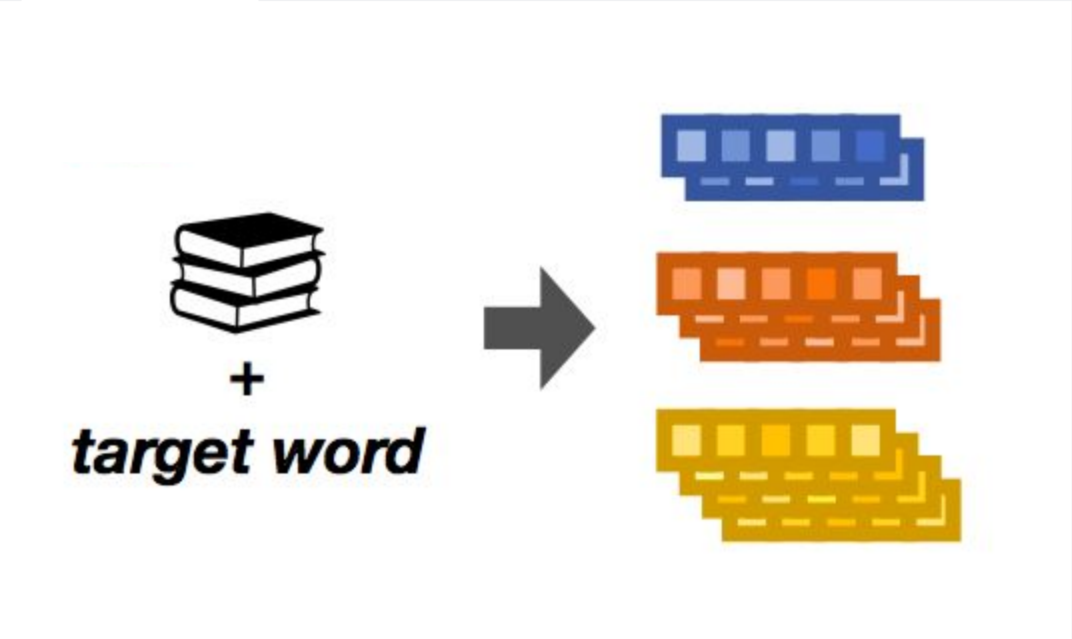
Computational detection of semantic change

Semantic representation + alignment technique + change measure

- Word embeddings models: average (type) embeddings and contextualized (token) embeddings + dynamic word embeddings.
- Alignment: orthogonal procrustes (OP), vector initialization (VI), temporal referencing (TR), canonical correlation analysis (CCA).
- Change measure: cosine distance (CD), euclidean distance (ED), local neighborhood distance (LND)

Contextualized word embeddings for LSC

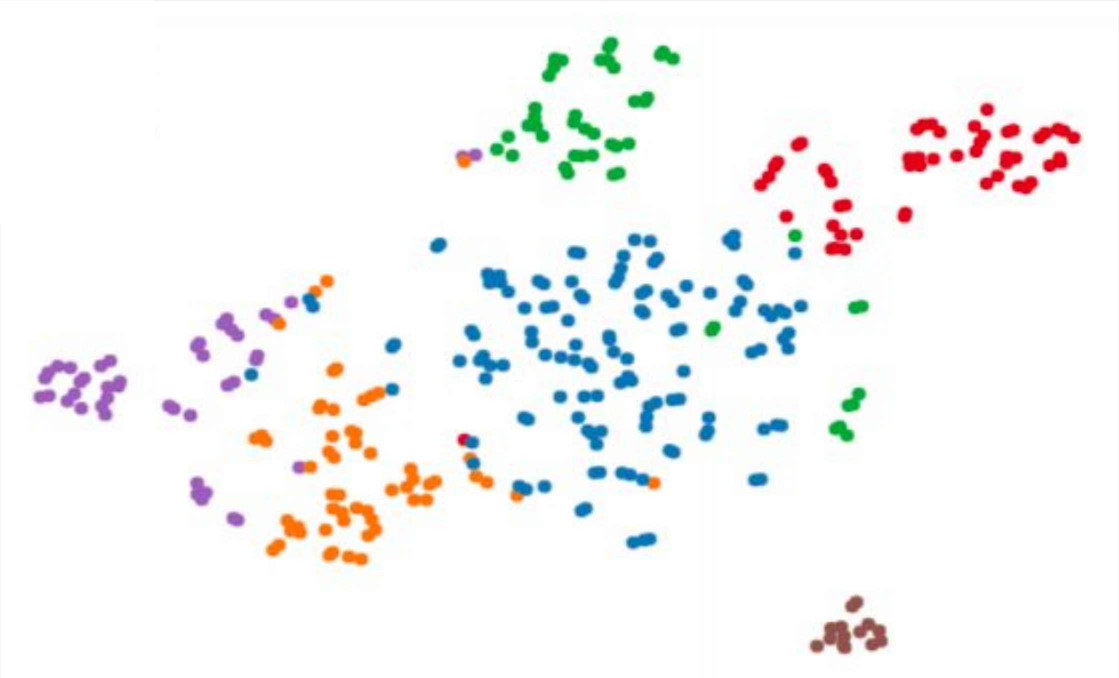
First step:



* a picture from Mario Giulianelli

Contextualized word embeddings for LSC

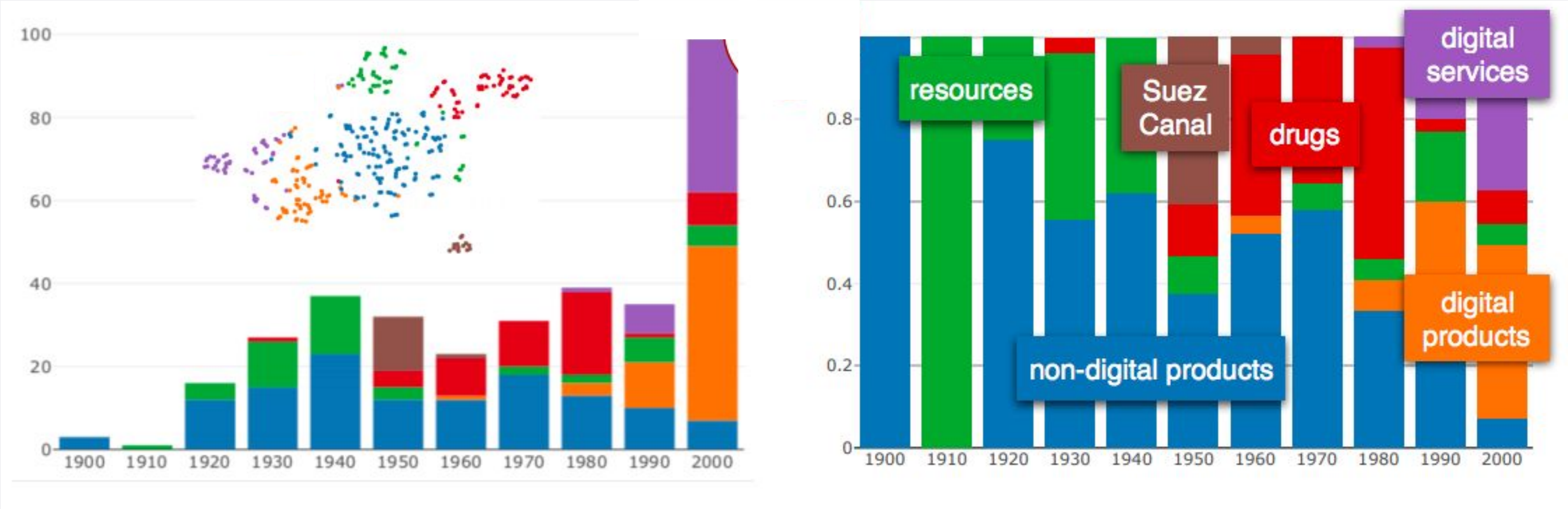
Second step:



* a pictures from Mario Giulianelli

Contextualized word embeddings for LSC

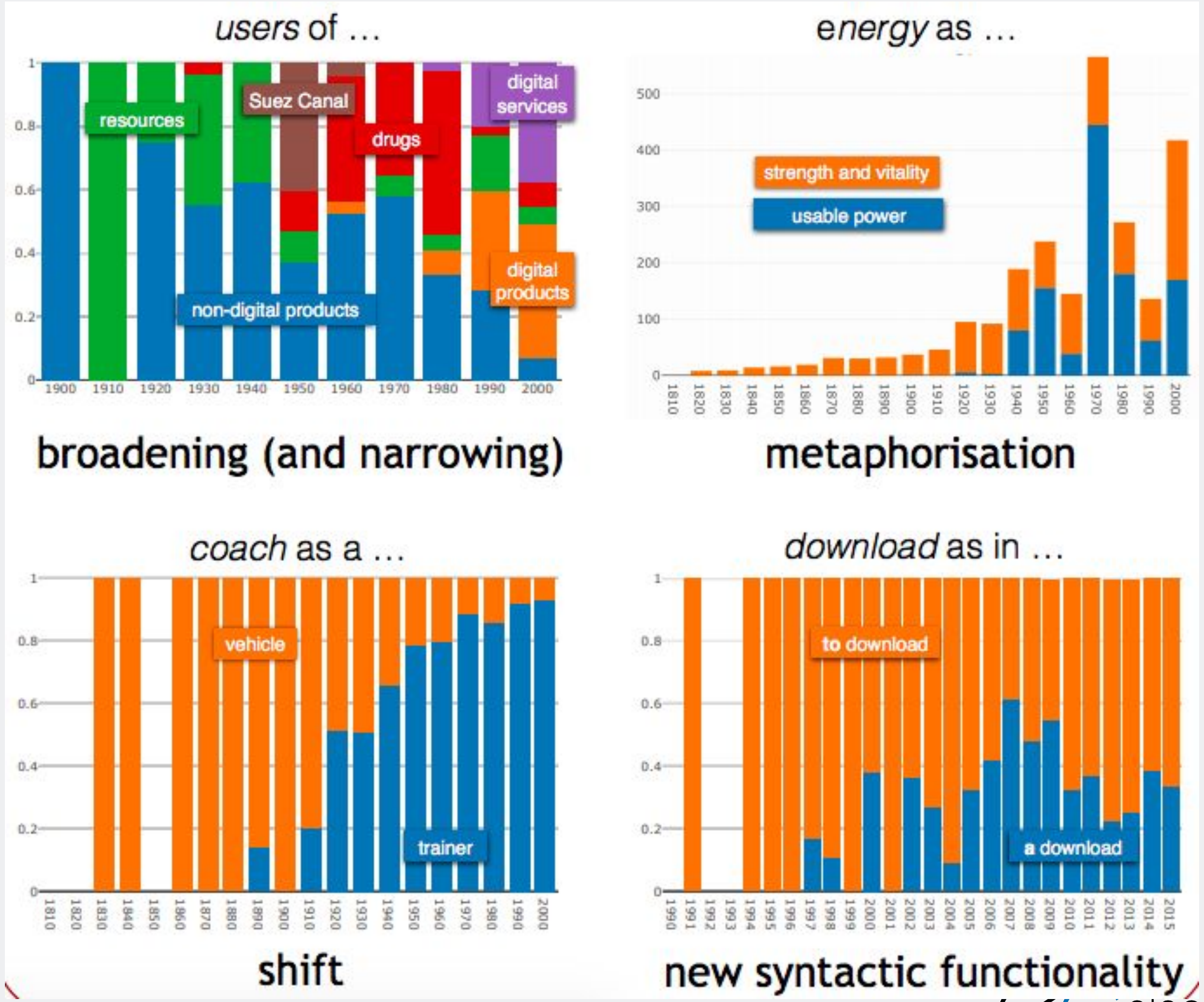
Third step:



* Pictures from Mario Giulianelli

Contextualized word embeddings for LSC

- Usage and cluster distributions are compared with:
  - Jensen-Shannon distance (JSD)
  - Average pairwise distance between token embeddings (APD)
  - Average cosine distance (CD)



SemEval-2020, task 1

- Unsupervised Lexical Semantic Change (LCS) Detection for German, English, Latin and Swedish [Schlechtweg et. al. 2020]:

SemEval-2020, task 1

- Unsupervised Lexical Semantic Change (LCS) Detection for German, English, Latin and Swedish [Schlechtweg et. al. 2020]:

  - subtask 1: binary classification,
  scoring measure - accuracy;
  - subtask 2: ranking,
  scoring measure - Spearman's ρ with the gold rank.

SemEval-2020, task 1. Data

Gold standard data:

A manually annotated semantic change of target words between two corpora based on word sense frequency distributions.

SemEval-2020, task 1. Data

Gold standard data:

A manually annotated semantic change of target words between two corpora based on word sense frequency distributions.

	t_1			t_2		
Senses	**Chamber**	**Biology**	**Phone**	**Chamber**	**Biology**	**Phone**
# uses	12	18	0	1	11	18

SemEval-2020, task 1. Data

Gold standard data:

	size	nouns/verbs/ adjectives	inter-annotator agreement	LSC subtask 1	LSC subtask 2	frequency difference
English	37	33/40/0	0.69	0.43	0.24	-0.29
German	48	32/14/2	0.59	0.35	0.31	0.00
Swedish	31	23/5/3	0.58	0.26	0.16	0.00

SemEval-2020, task 1. Methods

- Static type embeddings:
  - one word --> one vector
  - do not model word senses

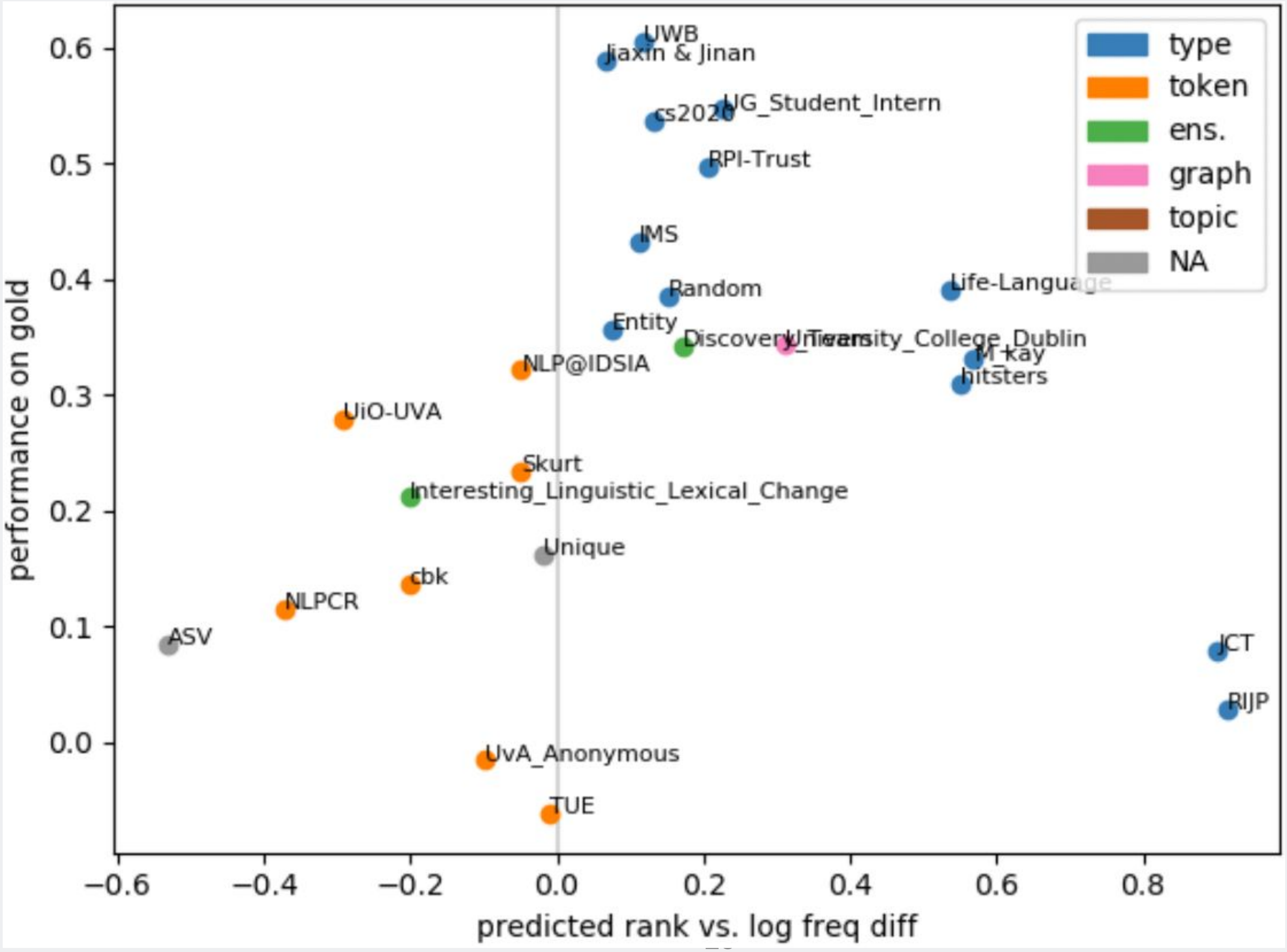
SemEval-2020, task 1. Methods

- Static type embeddings:
  - one word --> one vector
  - do not model word senses
- Token, i.e. contextualised, embeddings:
  - one usage --> one vector
  - require fine-tuning on annotated data.

SemEval-2020, task 1. Methods

- Static type embeddings:
  - one word --> one vector
  - do not model word senses
- Token, i.e. contextualised, embeddings:
  - one usage --> one vector
  - require fine-tuning on annotated data.
- Current graph-based semantic network and topic model approaches are far from achieving state-of-the-art results.

SemEval-2020, task 1



SemEval-2020, task 1

Team	Subtask 1					System
	Avg.	EN	DE	LA	SV	
UWB	.687	.622	.750	.700	.677	type
Life-Language	.686	.703	.750	.550	.742	type
Jiaxin & Jinan	.665	.649	.729	.700	.581	type
RPI-Trust	.660	.649	.750	.500	.742	type
UG_Student_Intern	.639	.568	.729	.550	.710	type
DCC	.637	.649	.667	.525	.710	type
NLP@IDSIA	.637	.622	.625	.625	.677	token
JCT	.636	.649	.688	.500	.710	type
Skurt	.629	.568	.562	.675	.710	token
Discovery_Team	.621	.568	.688	.550	.677	ens.
Count Bas.	.613	.595	.688	.525	.645	-

Team	Subtask 2					System
	Avg.	EN	DE	LA	SV	
UG_Student_Intern	.527	.422	.725	.412	.547	type
Jiaxin & Jinan	.518	.325	.717	.440	.588	type
cs2020	.503	.375	.702	.399	.536	type
UWB	.481	.367	.697	.254	.604	type
Discovery_Team	.442	.361	.603	.460	.343	ens.
RPI-Trust	.427	.228	.520	.462	.498	type
Skurt	.374	.209	.656	.399	.234	token
IMS	.372	.301	.659	.098	.432	type
UiO-UvA	.370	.136	.695	.370	.278	token
Entity	.352	.250	.499	.303	.357	type
Random	.296	.211	.337	.253	.385	type

SemEval-2020, task 1

Team	Subtask 1					System
	Avg.	EN	DE	LA	SV	
UWB	.687	.622	.750	.700	.677	type
Life-Language	.686	.703	.750	.550	.742	type
Jiaxin & Jinan	.665	.649	.729	.700	.581	type
RPI-Trust	.660	.649	.750	.500	.742	type
UG_Student_Intern	.639	.568	.729	.550	.710	type
DCC	.637	.649	.667	.525	.710	type
NLP@IDSIA	.637	.622	.625	.625	.677	token
JCT	.636	.649	.688	.500	.710	type
Skurt	.629	.568	.562	.675	.710	token
Discovery_Team	.621	.568	.688	.550	.677	ens.
Count Bas.	.613	.595	.688	.525	.645	-

Team		Subtask 2					System
		Avg.	EN	DE	LA	SV	
UG_Student_Intern		.527	.422	.725	.412	.547	type
Jiaxin & Jinan		.518	.325	.717	.440	.588	type
cs2020		.503	.375	.702	.399	.536	type
UWB		.481	.367	.697	.254	.604	type
Discovery_Team		.442	.361	.603	.460	.343	ens.
RPI-Trust		.427	.228	.520	.462	.498	type
Skurt		.374	.209	.656	.399	.234	token
IMS	.575	.372	.301	.659	.098	.432	type
UiO-UvA	.618	.370	.136	.695	.370	.278	token
Entity		.352	.250	.499	.303	.357	type
Random		.296	.211	.337	.253	.385	type

post evaluation fine-tuning

SemEval-2020, task 1

User	Entries	Date of Last Entry	Team Name	All	English	German	Latin	Swedish
				SPR ▲	SPR ▲	SPR ▲	SPR ▲	SPR ▲
akutuzov	9	04/07/20	UiO-UVA	0.618 (1)	0.605 (1)	0.695 (5)	0.561 (1)	0.610 (2)
jenskaiser	4	05/06/20	in vain	0.575 (2)	0.461 (2)	0.782 (1)	0.390 (9)	0.668 (1)
mpoemsl	6	03/23/20	UG Student Intern	0.545 (3)	0.422 (4)	0.725 (3)	0.487 (4)	0.547 (7)
jinan	34	03/24/20	Jiaxin & Jinan	0.499 (4)	0.295 (14)	0.735 (2)	0.388 (10)	0.579 (5)
p.cassotti	13	03/15/20	Random	0.496 (5)	0.304 (11)	0.722 (4)	0.395 (8)	0.562 (6)
lucas.rettenmeier	5	06/09/20	#hitsters	0.483 (6)	0.356 (7)	0.679 (7)	0.300 (18)	0.597 (3)
pribanp	4	03/11/20	UWB	0.455 (7)	0.365 (5)	0.687 (6)	0.181 (20)	0.587 (4)
mgruppi	13	03/28/20	RPI-Trust	0.432 (8)	0.292 (16)	0.567 (11)	0.455 (7)	0.415 (9)
belerico	9	10/18/20	INSID&S LAB	0.419 (9)	0.352 (8)	0.640 (9)	0.340 (14)	0.343 (13)

RuShiftEval-2021

- A shared task on semantic change detection for Russian.

RuShiftEval-2021

- A shared task on semantic change detection for Russian.
- Time periods: pre-Soviet (1700-1916); Soviet (1918-1990); post-Soviet (1991-2016).

RuShiftEval-2021

- A shared task on semantic change detection for Russian.
- Time periods: pre-Soviet (1700-1916); Soviet (1918-1990); post-Soviet (1991-2016).
- A ranking task, i.e. target words should be ranked according to the strength of their meaning change:
  - between pre-Soviet and Soviet periods;
  - between Soviet and post-Soviet periods;
  - between pre-Soviet and post-Soviet periods.

RuShiftEval-2021

User	Entries	Date of Last Entry	Team Name	RuSemShift1 (pre-Soviet:Soviet) ▲	RuSemShift2 (Soviet:post-Soviet) ▲	RuSemShift3 (pre-Soviet:post-Soviet) ▲	Average score ▲
myrachins	7	02/28/21	GlossReader	0.781	0.803	0.822	0.802 (1)
UsrD7	5	03/01/21	DeepMistake	0.798	0.773	0.803	0.791 (2)
netvad	5	03/01/21	DeepMistake	0.794	0.773	0.799	0.789 (3)
davletov-aa	1	02/28/21	DeepMistake	0.749	0.801	0.788	0.779 (4)
vanyatko	2	02/25/21		0.678	0.746	0.737	0.720 (5)
aryzhova	8	02/26/21		0.469	0.450	0.453	0.457 (6)
SyrielleM	10	02/28/21	Discovery Team	0.455	0.410	0.494	0.453 (7)
pribanp	10	02/26/21	UWB	0.362	0.354	0.533	0.417 (8)
dschlechtweg	10	03/01/21		0.419	0.373	0.383	0.392 (9)
jenskaiser	10	02/28/21		0.430	0.310	0.406	0.382 (10)

2. Current challenges

Challenges

- What is meaning and what is semantic change?

Challenges

- What is meaning and what is semantic change?
- Models' unreliability: random initialization, sensitivity to the corpus size and frequency.

Challenges

- What is meaning and what is semantic change?
- Models' unreliability: random initialization, sensitivity to the corpus size and frequency.
- Datasets' reliability (size and annotation quality).

Challenges

- What is meaning and what is semantic change?
- Models' unreliability: random initialization, sensitivity to the corpus size and frequency.
- Datasets' reliability (size and annotation quality).
- Dataset (language and genre) dependent parameter tuning.

3. Our experiments

Aims

- Find a more reliable solution for unsupervised modeling of LSC.

Aims

- Find a more reliable solution for unsupervised modeling of LSC.
- Control for various factors that might be misinterpreted as LSC, i.e. model artefacts, frequency change, particular genre and language.

Aims

- Find a more reliable solution for unsupervised modeling of LSC.
- Control for various factors that might be misinterpreted as LSC, i.e. model artefacts, frequency change, particular genre and language.
- Investigate methods for modeling and evaluating short-term semantic shifts (not covered here).

Experiments

Datasets: DYLEN (Austrian media), SemEval (German, English, Swedish), RuShiftEval (Russian), RuSemShift (Russian).

Semantic change over time

DYLEN data

- Centuries: language evolution.
- Decades: evolution of socio-cultural concepts, genres, styles.
- Years and months: impact of real-world events onto society.
- Weeks and days: spread of lexical innovations, variation among communities.

Experiments

Datasets: DYLEN (Austrian media), SemEval (German, English, Swedish), RuShiftEval (Russian), RuSemShift (Russian).

1. Evaluation of different methods on various datasets.
2. Frequency change effect in LSC.
3. Graph-based approach for LSC.

3.a Evaluation of different methods on various datasets

LSC (ranking)

	Mean	Std
SGNS+OP+CD	**0.51**	0.052
SGNS+LND	0.47	0.045
GloVe+OP+CD	**0.51**	0.055
GloVe+VI+CD	0.44	0.062
GloVe+LND	0.48	0.042
BERT+APD	0.43	0.056
BERT+CD	0.37	0.048
BERT+JSD	0.35	0.059

Some insights learnt

- Different degree of LSC in gold standard data affects model's performance.

Some insights learnt

- Different degree of LSC in gold standard data affects model's performance.

- The prevalence of one sense appears to be challenging for the models that use contextualised embeddings.

3.b Frequency change effect

Semantic change and frequency

The laws of semantic change [Hamilton et. al. 2016]:

- rates of semantic change scale with a negative power of word frequency;

Semantic change and frequency

The laws of semantic change [Hamilton et. al. 2016]:

- rates of semantic change scale with a negative power of word frequency;

Counter findings:

- negative correlation between meaning change and frequency is largely an artefact of the models [Dubossarsky et. al. 2017];
- semantic change is higher for low-frequency words and for words with strong changes in frequency [Schlechtweg et. al. 2020].

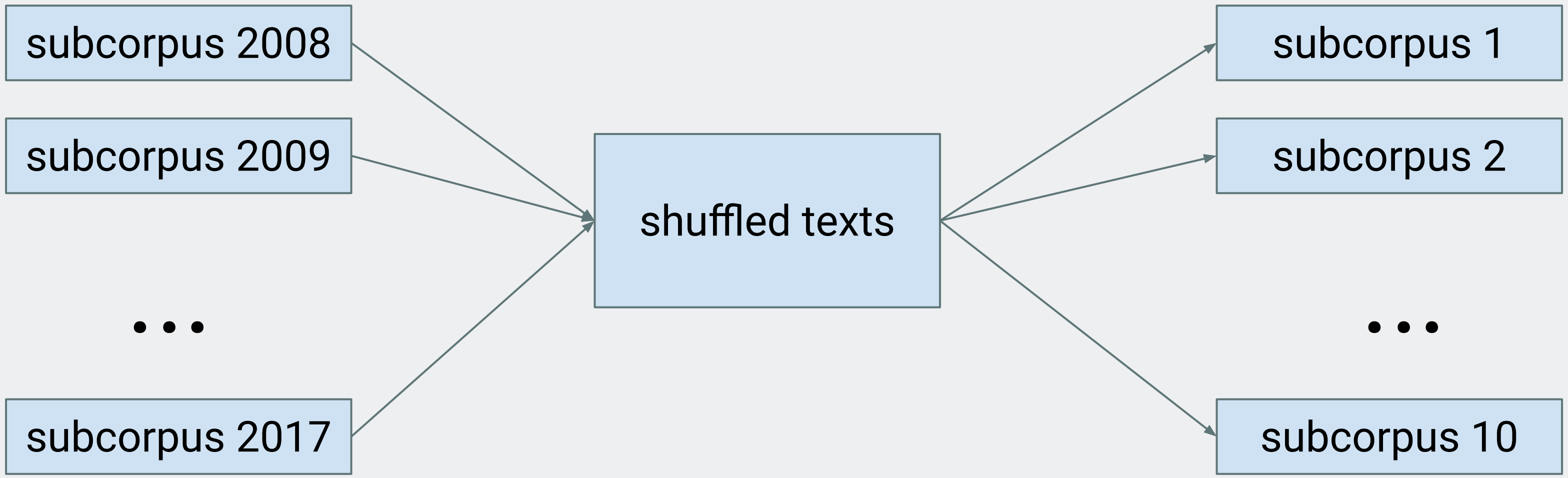
Semantic change analysis (1)

Frequency		VI_o	VI_m	LND	RDCG
Low	**0.18 (0.09)**	0.54 (0.11)	0.56 (0.12)	0.62 (0.20)	0.67 (0.24)
Medium	**0.89 (0.32)**	0.61 (0.12)	0.57 (0.11)	0.62 (0.21)	0.66 (0.24)
High	**3.12 (0.89)**	0.64 (0.14)	0.56 (0.08)	0.64 (0.20)	0.66 (0.23)

Semantic change analysis (2)

Frequency		VI_o	VI_m	LND	RDCG
Low variance	**0.12 (0.04)**	0.55 (0.12)	0.65 (0.07)	0.61 (0.20)	0.65 (0.23)
Medium variance	**0.88 (0.21)**	0.54 (0.12)	0.59 (0.07)	0.63 (0.21)	0.67 (0.25)
High variance	**1.70 (0.98)**	0.51 (0.12)	0.59 (0.12)	0.56 (0.21)	0.60 (0.23)

Control condition



Following [Dubossarsky et. al. 2017, Rettenmeier 2020], we measure the variance explained by frequency using linear mixed effects model.

Control condition

		VI_o		VI_m		LND		RDCG	
		SGNS	GloVe	SGNS	GloVe	SGNS	GloVe	SGNS	GloVe
Explained variance	**Original**	23%	38%	25%	38%	59%	60%	62%	61%
	Random	18%	33%	7%	9%	48%	49%	53%	49%
Semantic change	**Original**	0.59							
	Random	0.30							

Frequency factor

- Absolute difference between log-transformed relative frequencies in two subsequent years.

  => frequency time series

- Autocorrelation based dissimilarity measure [Montero et al., 2014].

Frequency correlation

Frequency	VI_o		VI_m		LND		RDCG	
	SGNS	GloVe	SGNS	GloVe	SGNS	GloVe	SGNS	GloVe
All	0.68	0.54	0.44	0.36	0.33	0.34	0.31	0.31

Our findings

- No clear pattern of semantic change and frequence development.

  -- significant frequency increase is not necessarily followed by significant change in usage,
  -- relatively constant frequencies over time do not imply stability of contextual meaning.

Our findings

- No clear pattern of semantic change and frequence development.

  -- significant frequency increase is not necessarily followed by significant change in usage,
  -- relatively constant frequencies over time do not imply stability of contextual meaning.

- Deliberate model design leads to the models of semantic change that encode less frequency information.

3.c Graph-based approach

Motivation

- Follow an intuitive semantic representation on the level of words' senses.

Motivation

- Follow an intuitive semantic representation on the level of words' senses.
- Provide a more robust solution for LSC detection (i.e., by avoiding noisy alignment of word embeddings, etc.).

Motivation

- Follow an intuitive semantic representation on the level of words' senses.
- Provide a more robust solution for LSC detection (i.e., by avoiding noisy alignment of word embeddings, etc.).
- Possibility to examine what exactly has changed and to detect a change point.

Motivation

- Follow an intuitive semantic representation on the level of words' senses.
- Provide a more robust solution for LSC detection (i.e., by avoiding noisy alignment of word embeddings, etc.).
- Possibility to examine what exactly has changed and to detect a change point.
- A straightforward approach to visualise LSC analysis for various applications (in humanities).

Approach 1

Two words are similar if they are related to the similar words [Glen and Widom 2002]

Approach 1

Two words are similar if they are related to the similar words [Glen and Widom 2002]
=> compare semantic neighbourhoods of subcorpora:

Approach 1

Two words are similar if they are related to the similar words [Glen and Widom 2002]
=> compare semantic neighbourhoods of subcorpora:

1) Word vector representation for each subcorpus.

Approach 1

Two words are similar if they are related to the similar words [Glen and Widom 2002]
=> compare semantic neighbourhoods of subcorpora:

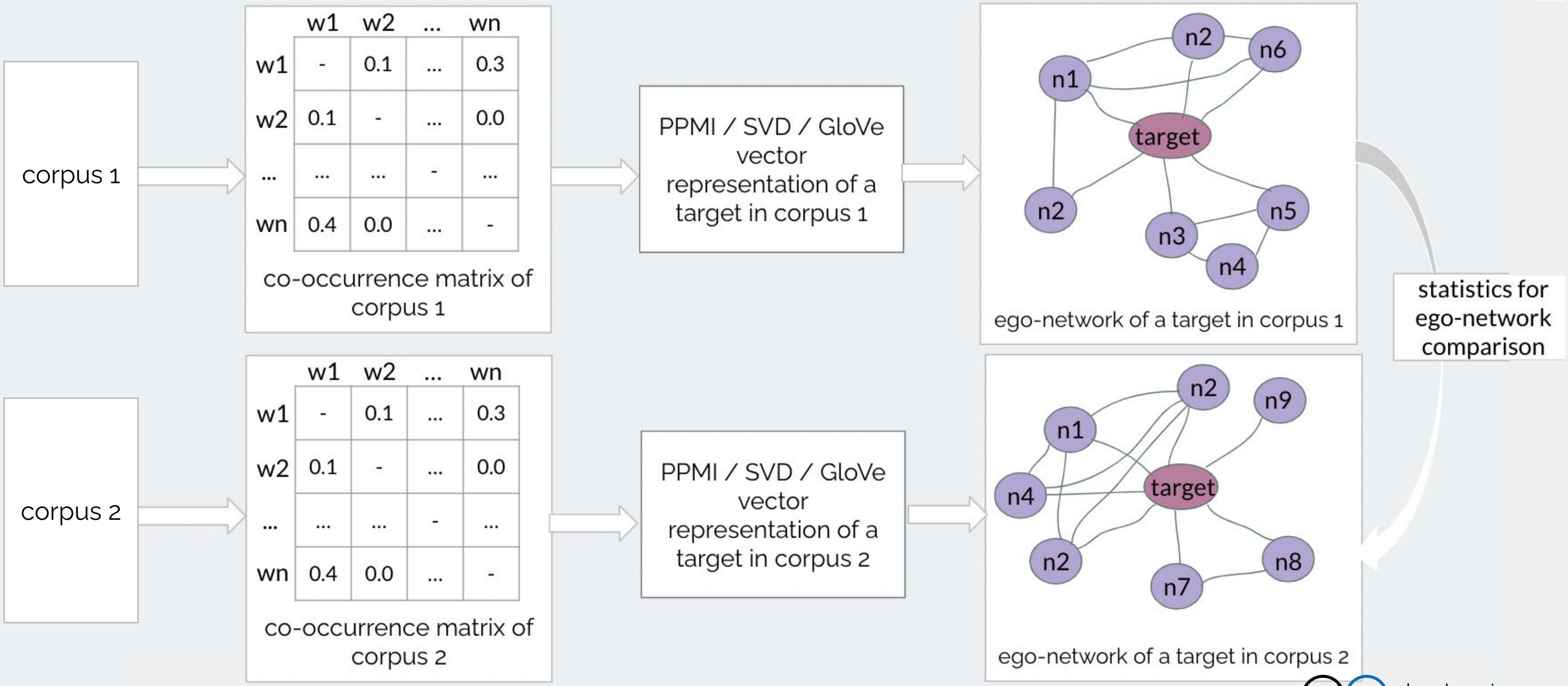
1) Word vector representation for each subcorpus.
2) Ego-network of semantically most related words.

Approach 1

Two words are similar if they are related to the similar words [Glen and Widom 2002]
=> compare semantic neighbourhoods of subcorpora:

1) Word vector representation for each subcorpus.
2) Ego-network of semantically most related words.
3) Trace dynamics of ego networks.

Approach 1


corpus 1
w1 w2 ... wn
w1 - 0.1 ... 0.3
w2 0.1 - ... 0.0
... ... - ...
wn 0.4 0.0 ... -
co-occurrence matrix of corpus 1
PPMI / SVD / GloVe vector representation of a target in corpus 1
n1 n2 n6 target n2 n3 n4 n5
ego-network of a target in corpus 1
statistics for ego-network comparison
corpus 2
w1 w2 ... wn
w1 - 0.1 ... 0.3
w2 0.1 - ... 0.0
... ... - ...
wn 0.4 0.0 ... -
co-occurrence matrix of corpus 2
PPMI / SVD / GloVe vector representation of a target in corpus 2
n1 n2 n9 n4 target n2 n7 n8
ego-network of a target in corpus 2

Approach 1. Compare ego-networks

- Neighbourhood similarity.
  Statistics to compare node sets in two time periods.

Approach 1. Compare ego-networks

- <u>Neighbourhood similarity</u>.
  Statistics to compare node sets in two time periods.

- <u>Dynamics of network features.</u>
  Compare the difference of graph metrics values.

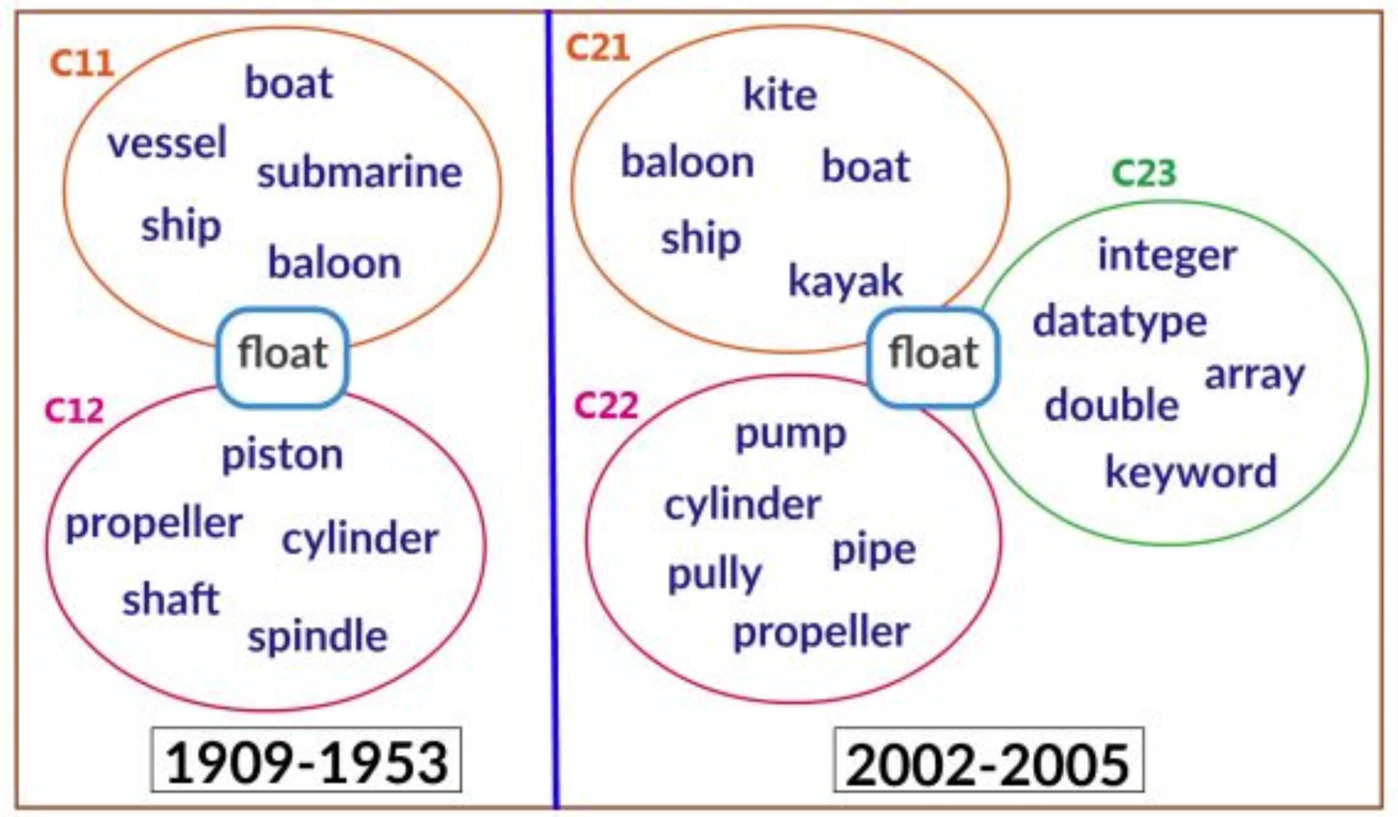
Approach 1. Compare ego-networks

- Neighbourhood similarity.
  Statistics to compare node sets in two time periods.

- Dynamics of network features.
  Compare the difference of graph metrics values.

- Clustering (word sense induction).
  Compare distribution of clusters among time periods.

Approach 1. Compare ego-networks. Clustering



[Jana et. al. 2019]

Approach 1. Compare ego-networks. Clustering

1. Chinese Whispers clustering --> sense clusters for each time period

<table>
<tr><th>Cluster</th><th>1997</th><th>2001</th><th>2009</th><th>2016</th><th>2017</th></tr>
<tr><td>Integration1</td><td>Entwicklung, Bereich, EU, Europa, Politik, Unternehmen, Maßnahme, Schaffung, Strategie, Verbesserung, Zusammenarbeit, Reform, Schwerpunkt, Lösung, Kooperation, Gesellschaft

“General”</td><td>Bereich, Entwicklung, Arbeitsmarkt, EU, Maßnahme, Förderung, Bildung, Unternehmen, Verbesserung, Schaffung, Lösung, Schwerpunkt, Zusammenarbeit, Projekt, Kooperation

“General”</td><td>Bereich, Thema, Bildung, Entwicklung, Lösung, Arbeitsmarkt, Maßnahme, Förderung, Gesellschaft, Umsetzung, Initiative, Schwerpunkt, Politik, Kompetenz, Bedeutung, Verbesserung

“General”</td><td>Arbeitsmarkt, Bildung, Maßnahme, Herausforderung, Thema, Gesellschaft, Lösung, Entwicklung, Schule, Beschäftigung, Förderung, Politik, Initiative, Bereich, Unterstützung

“General”</td><td rowspan="2">Bildung, Arbeitsmarkt, Migration, Flüchtling, Maßnahme,, Gesellschaft, Mindestsicherung, Herausforderung, Lösung, Zuwanderung, Entwicklung, Förderung, Politik, Asylberechtigte, Deutschkurs, Zuwanderer, Initiative, Sicherheit, Zusammenarbeit, Zusammenleben

“General + migration”</td></tr>
<tr><td>Integration2</td><td></td><td>Zuwanderer, Ausländer, Zuwanderung, Integrationsvertrag, Migrantin

“Migration”</td><td>Migrant, Zuwanderer, Migration, Zuwanderung, Migrationshintergrund, Migrantin, Integrationspolitik, Deutschkenntnis

“Migration”</td><td>Flüchtling, Asylberechtigte, Mindestsicherung, Deutschkurs, Asylwerber, Migration, Integrationsmaßnahm e, Asyl, Asylverfahren, Integrationsminister, Zuwanderung, Schutzberechtigte

“Migration”</td></tr>
</table>

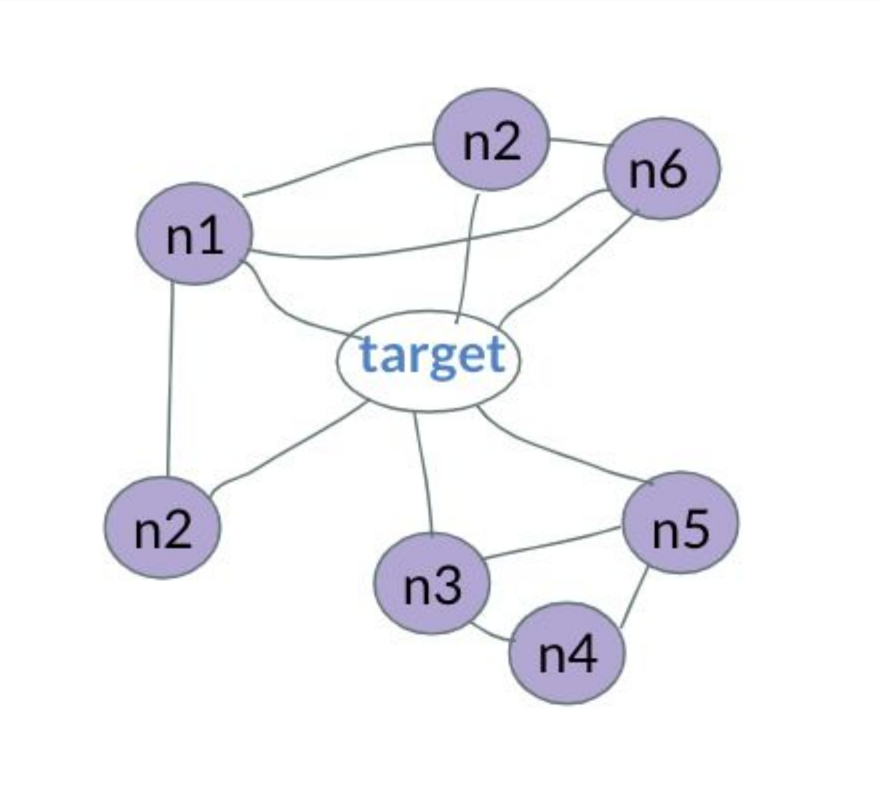
Cluster	1997	2001	2003	2008	2018
Klima1	Bundeskanzler, Viktor_Klima, Kanzler, SPÖ, Schüssel, ÖVP, Franz_Vranitzky, Bundesregierung, Vizekanzler, Regierung, Finanzminister, Wolfgang_Schüssel, Haider, FPÖ, Pensionsreform, Reform, Politik, Minister, Jörg_Haider, Rudas, Koalition, Klestil, Diskussion, Partei “Chancellor Klima”	SPÖ, FPÖ, Politik, ÖVP, Partei, Regierung, Haider, Bundeskanzler, Schüssel, Kritik, Koalition, Westenthaler, Grüne, Freiheitliche, Gusenbauer, Politiker, Opposition, Debatte, Zusammenhang, Bundesregierung, Sozialdemokrat, Klubobmann, Abgeordnete “Political atmosphere”	Politik, SPÖ, Regierung, ÖVP, Auswirkung, Bevölkerung, Natur, Politiker, Problem, Wasser, Maßnahme, Gefahr, Partei, Kanzler, Schüssel, Situation, Koalition, Region, Entwicklung, Schutz, Einfluss, Zusammenhang, Land, Atmosphäre, FPÖ, Grund, Kritik, Grüne “Political atmosphere”	Klimawandel, Klimaschutz, Umwelt, Energie, CO2, Treibhausgas, Pflanze, Landwirtschaft, Maßnahme, Politik, Kohlendioxid, Energieträger, Natur, CO, Erwärmung, Region, Bevölkerung, Umweltminister, Umweltschutz, Wasser, Auswirkung, CO2-Emission “Climate change”	Klimawandel, Umwelt, Klimaschutz, Landwirtschaft, Erwärmung, Erderwärmung, Temperatur, Pflanze, Energie, Klimapolitik, Klimaerwärmung, Natur, CO2, Forscher, Auswirkung, Dürre, Treibhausgas, Grad, Lebensraum, Wasser, Region, Biodiversität, Maßnahme, Wärme “Climate change”
Klima2			Klimaanlage, Klimaautomatik, Fensterheber, Nebelscheinwerfer “Air conditioning”		

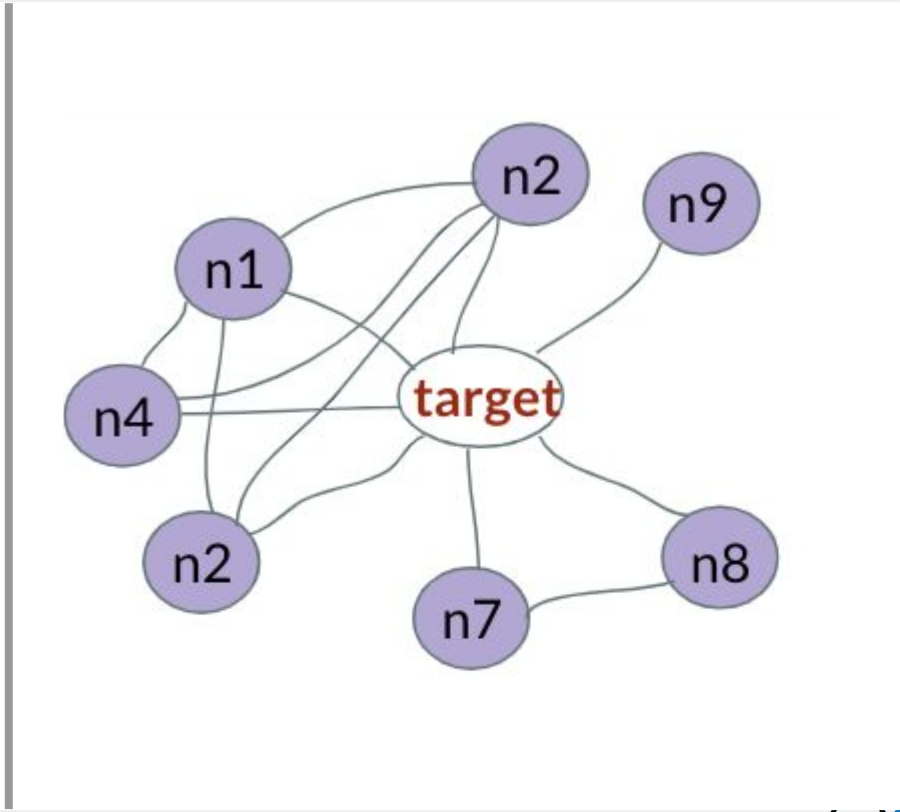
Approach 1. Compare ego-networks. Clustering

1. Chinese Whispers clustering --> sense clusters for each time period

Approach 1. Compare ego-networks. Clustering

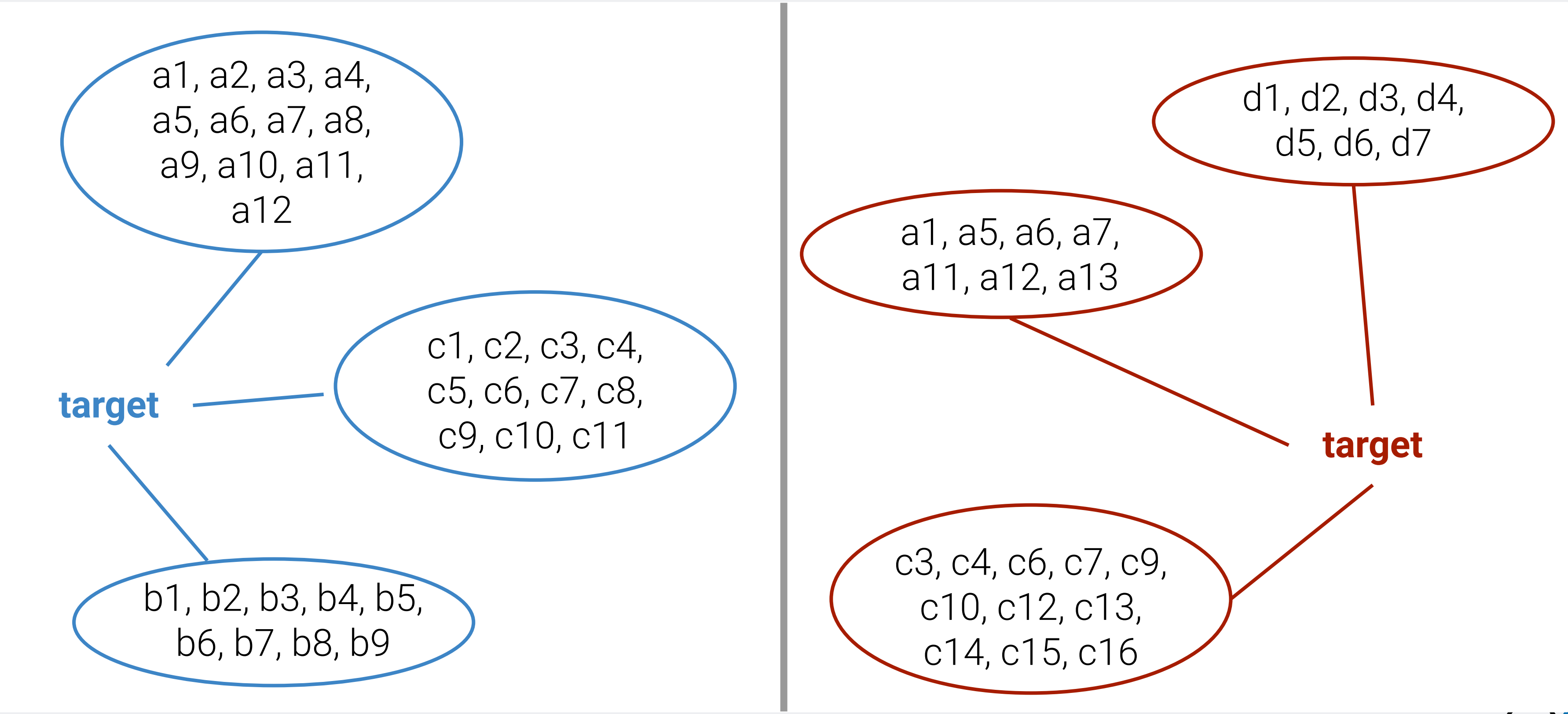
1. Chinese Whispers clustering --> sense clusters for each time period





Approach 1. Compare ego-networks. Clustering

1. Chinese Whispers clustering --> sense clusters for each time period



corpus 1 **corpus 2**

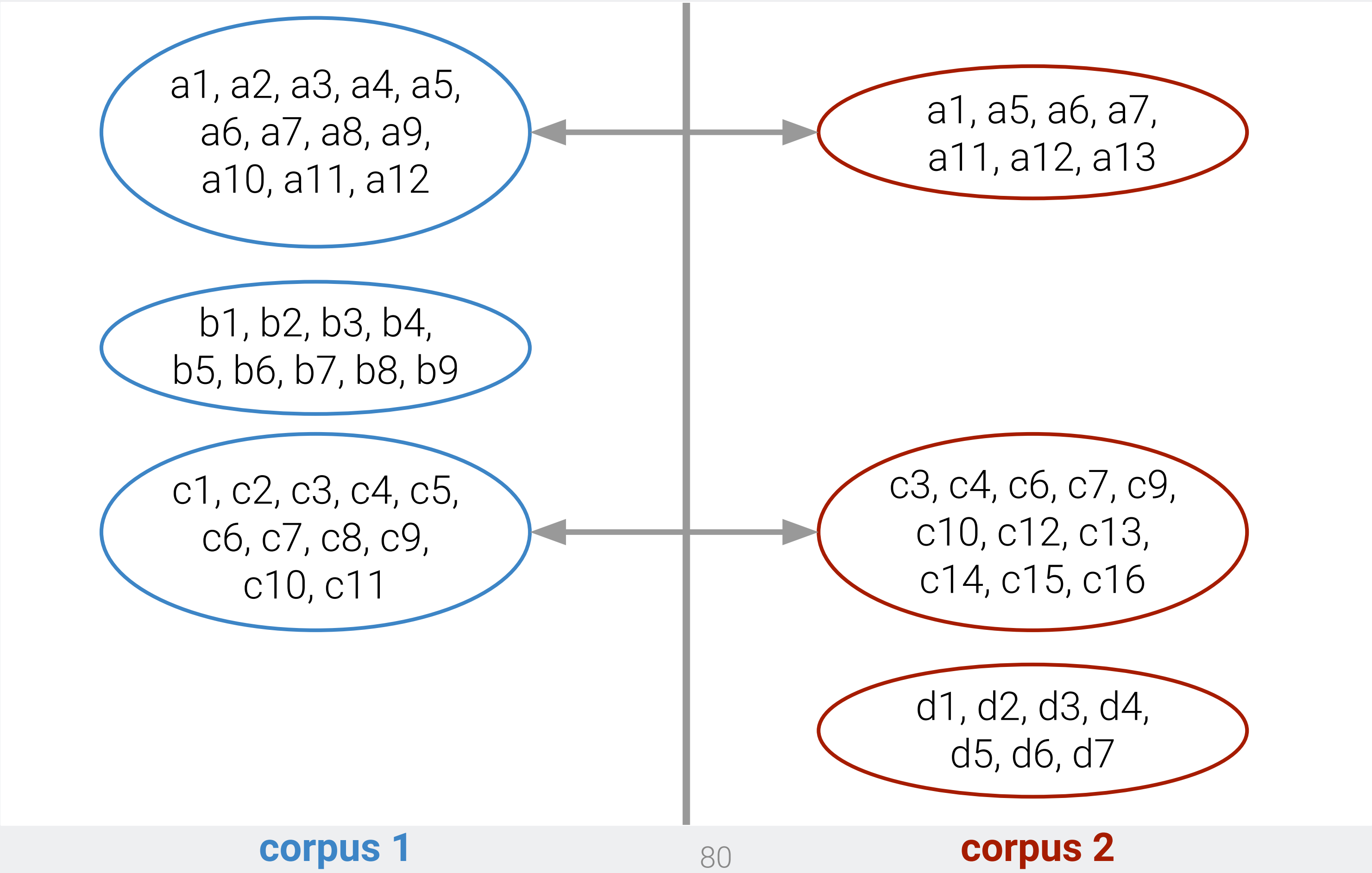
Approach 1. Compare ego-networks. Clustering

1. Chinese Whispers clustering --> sense clusters for each time period

2. Affinity propagation clustering --> correspondence of sense clusters.

Approach 1. Compare ego-networks. Clustering

2. Affinity propagation clustering --> correspondence of sense clusters.



TU WIEN Informatics

electronic commerce

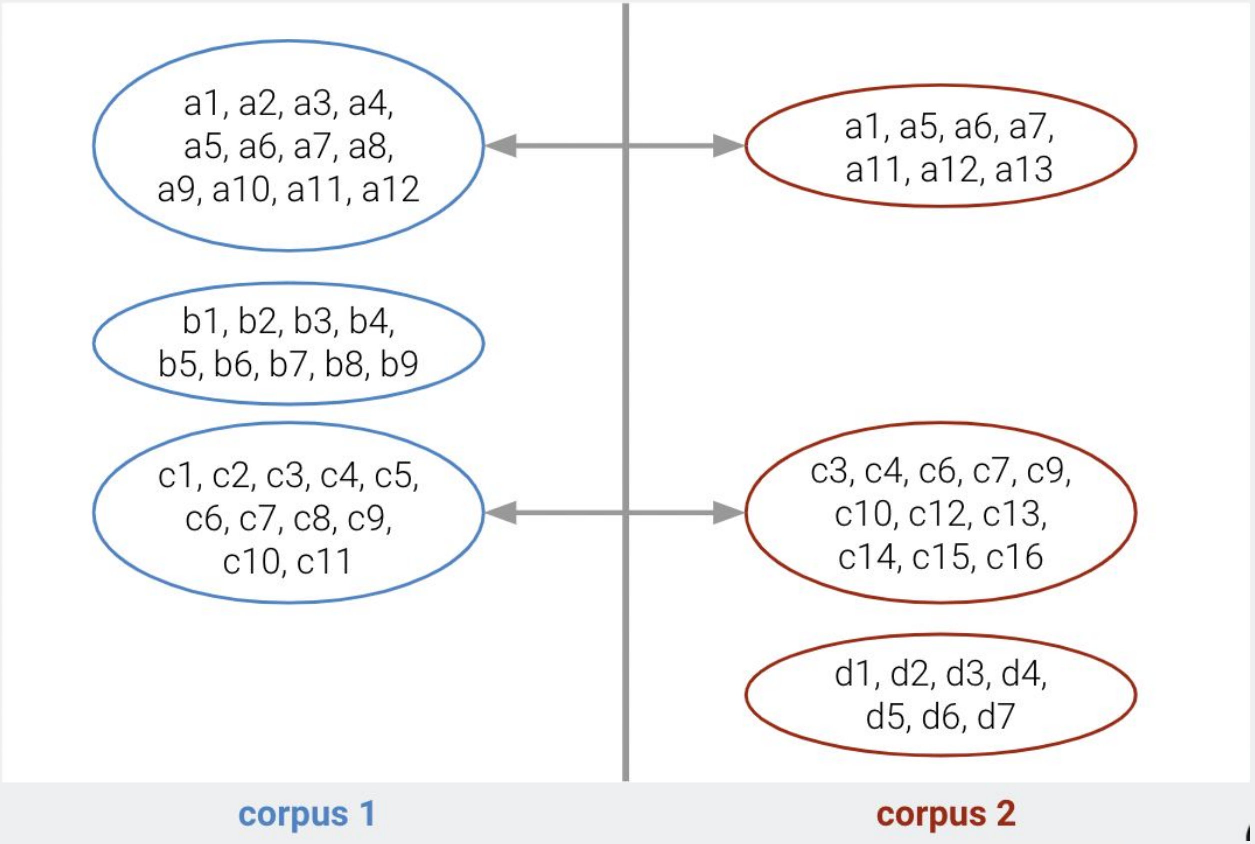
Approach 1. Compare ego-networks. Clustering

1. Chinese Whispers clustering --> sense clusters for each time period
2. Affinity propagation clustering --> correspondence of sense clusters.
3. Compute clusters' similarity -- ?

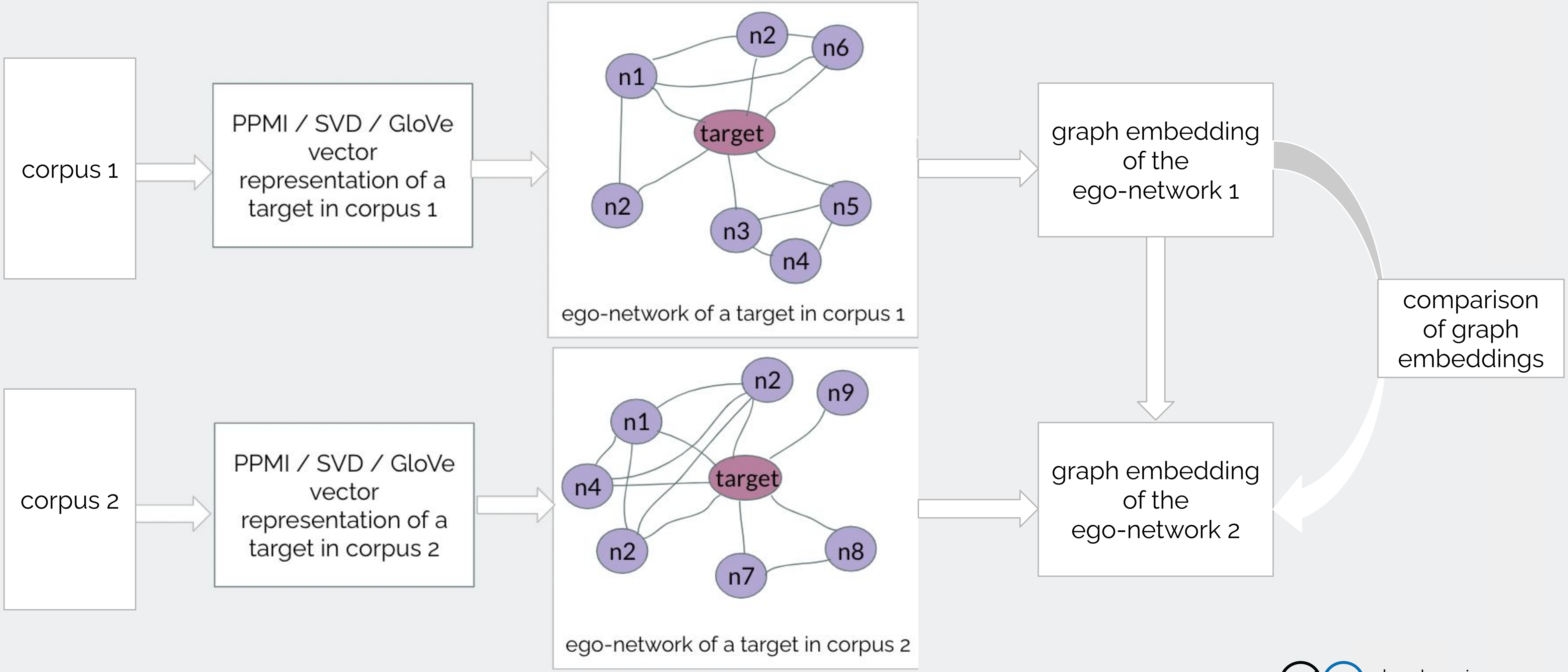
Approach 1. Compare ego-networks. Clustering

3. Compute clusters' similarity -- ?

- Jaccard index or rankDCG --> similarity of corresponding clusters
- Proportion of words in emerged / lost clusters
- Proportion of emerged / lost clusters

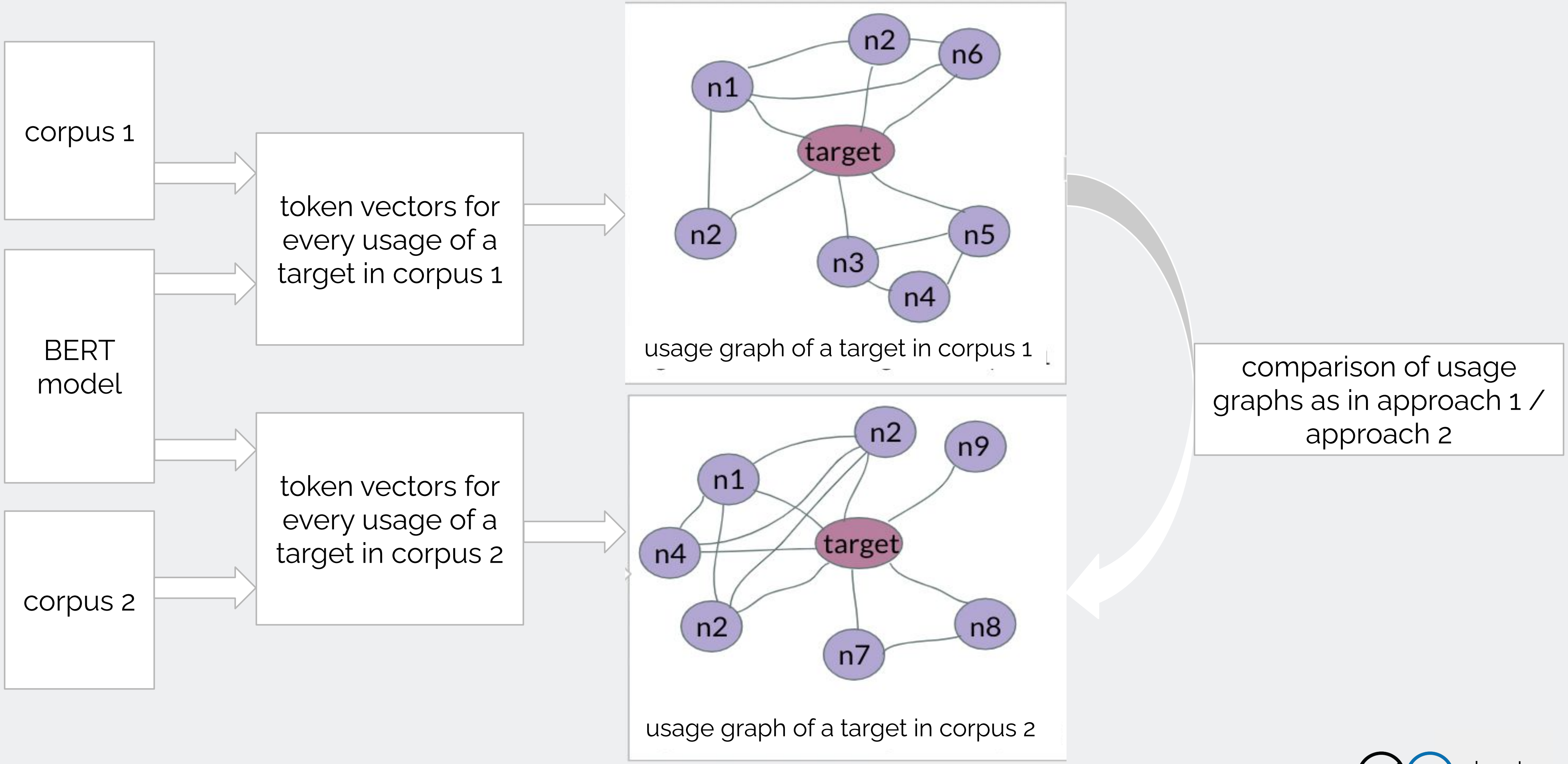


Approach 2



corpus 1
PPMI / SVD / GloVe vector representation of a target in corpus 1
n1
n2
n6
target
n2
n3
n4
n5
ego-network of a target in corpus 1
graph embedding of the ego-network 1
comparison of graph embeddings
corpus 2
PPMI / SVD / GloVe vector representation of a target in corpus 2
n1
n2
n9
n4
target
n2
n7
n8
ego-network of a target in corpus 2
graph embedding of the ego-network 2

Approach 3


corpus 1
BERT model
corpus 2
token vectors for every usage of a target in corpus 1
token vectors for every usage of a target in corpus 2
n1
n2
n6
target
n2
n3
n5
n4
usage graph of a target in corpus 1
n2
n9
n1
n4
target
n2
n7
n8
usage graph of a target in corpus 2
comparison of usage graphs as in approach 1 / approach 2

Results (ranking)

	DYLEN	SemEval EN	SemEval DE	SemEval SV	RuShift Eval 1	RuShift Eval 2	RuShift Eval 3	RuSem Shift 1	RuSem Shift 2	Mean	Std
GloVe+cl+ rankDCG	0.31	0.40	0.42	0.46	0.43	0.52	0.42	0.61	0.44	0.44	0.078
GloVe+dn2v+CS	0.49	0.53	0.52	0.58	0.54	0.52	0.59	0.55	0.53	**0.54**	**0.029**
GloVe+dn2v+UD	0.53	0.52	0.56	0.54	0.57	0.58	0.62	0.57	0.58	**0.56**	**0.028**
BERT+dn2v+ rankDCG	0.32	0.37	0.27	0.34	0.31	0.29	0.33	0.30	0.26	0.31	0.032
BERT+dn2v+UD	0.42	0.32	0.49	0.44	0.37	0.32	0.41	0.44	0.46	0.41	0.056
BERT+dn2v+CS	0.38	0.48	0.52	0.50	0.49	0.51	0.55	0.46	0.51	0.49	0.045

Results (ranking)

	DYLEN	SemEval EN	SemEval DE	SemEval SV	RuShift Eval 1	RuShift Eval 2	RuShift Eval 3	RuSem Shift 1	RuSem Shift 2	Mean	Std
GloVe+cl+ rankDCG	0.31	0.40	0.42	0.46	0.43	0.52	**0.42**	0.61	0.44	0.44	0.078
GloVe+dn2v+CS	0.49	0.53	0.52	0.58	0.54	0.52	**0.59**	0.55	0.53	0.54	0.029
GloVe+dn2v+UD	0.53	0.52	0.56	0.54	0.57	0.58	**0.62**	0.57	0.58	0.56	0.028
BERT+dn2v+ rankDCG	0.32	0.37	0.27	0.34	0.31	0.29	**0.33**	0.30	0.26	0.31	0.032
BERT+dn2v+UD	0.42	0.32	0.49	0.44	0.37	0.32	**0.41**	0.44	0.46	0.41	0.056
BERT+dn2v+CS	0.38	0.48	0.52	0.50	0.49	0.51	**0.55**	0.46	0.51	0.49	0.045

4. Conclusion

Graph-based approach

- More stable performance across different datasets.

Graph-based approach

- More stable performance across different datasets.
- Cluster-based embeddings.

Graph-based approach

- More stable performance across different datasets.
- Cluster-based embeddings.
- Various graph embeddings architectures and network related parameters.

Graph-based approach

- More stable performance across different datasets.
- Cluster-based embeddings.
- Various graph embeddings architectures and network related parameters.
- Dynamic graph embeddings for multiple time / domain point analysis.

Thank you!

Anna Marakasova
anna.marakasova@tuwien.ac.at