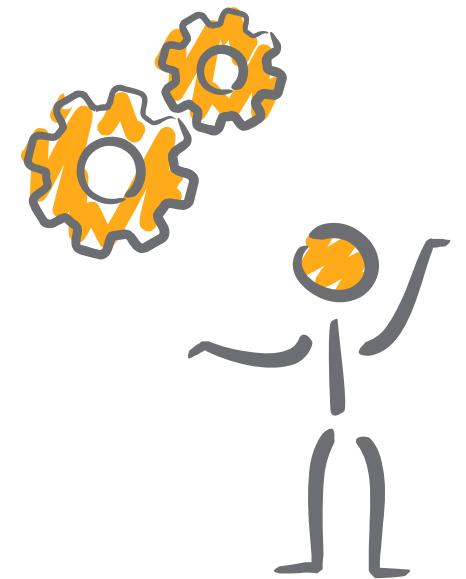


1st International Workshop on Computational Approaches to Historical Language Change

LChange'19



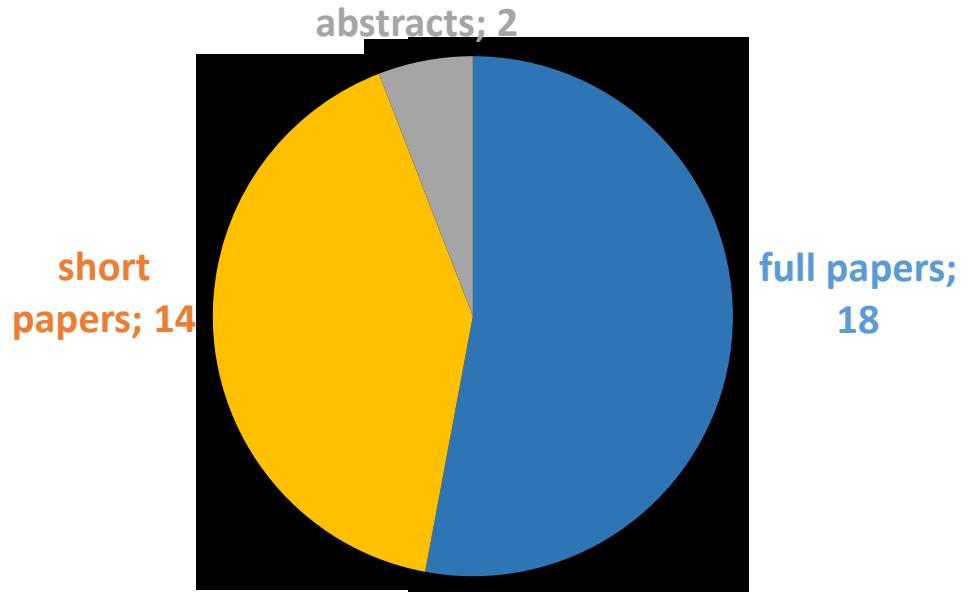
Nina Tahmasebi, Lars Borin, Adam Jatowt, Yang Xu
In conjunction with ACL2019



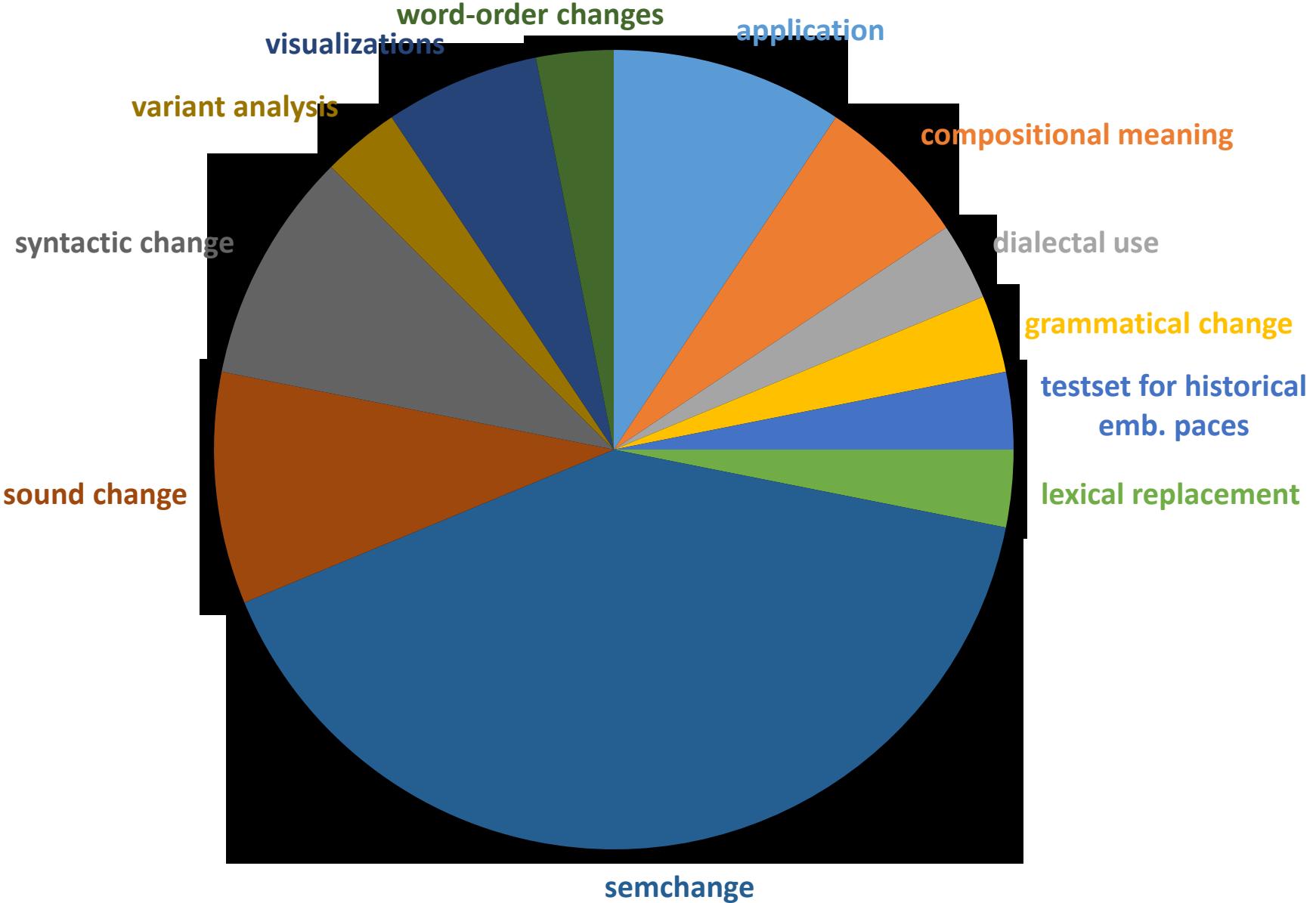
Florence, Italy, August 2, 2019
LChange'19 • ACL2019 • N. Tahmasebi

Submission info

- 55 submissions
- 34 accepted
- 57 reviewers!



Distribution of topics



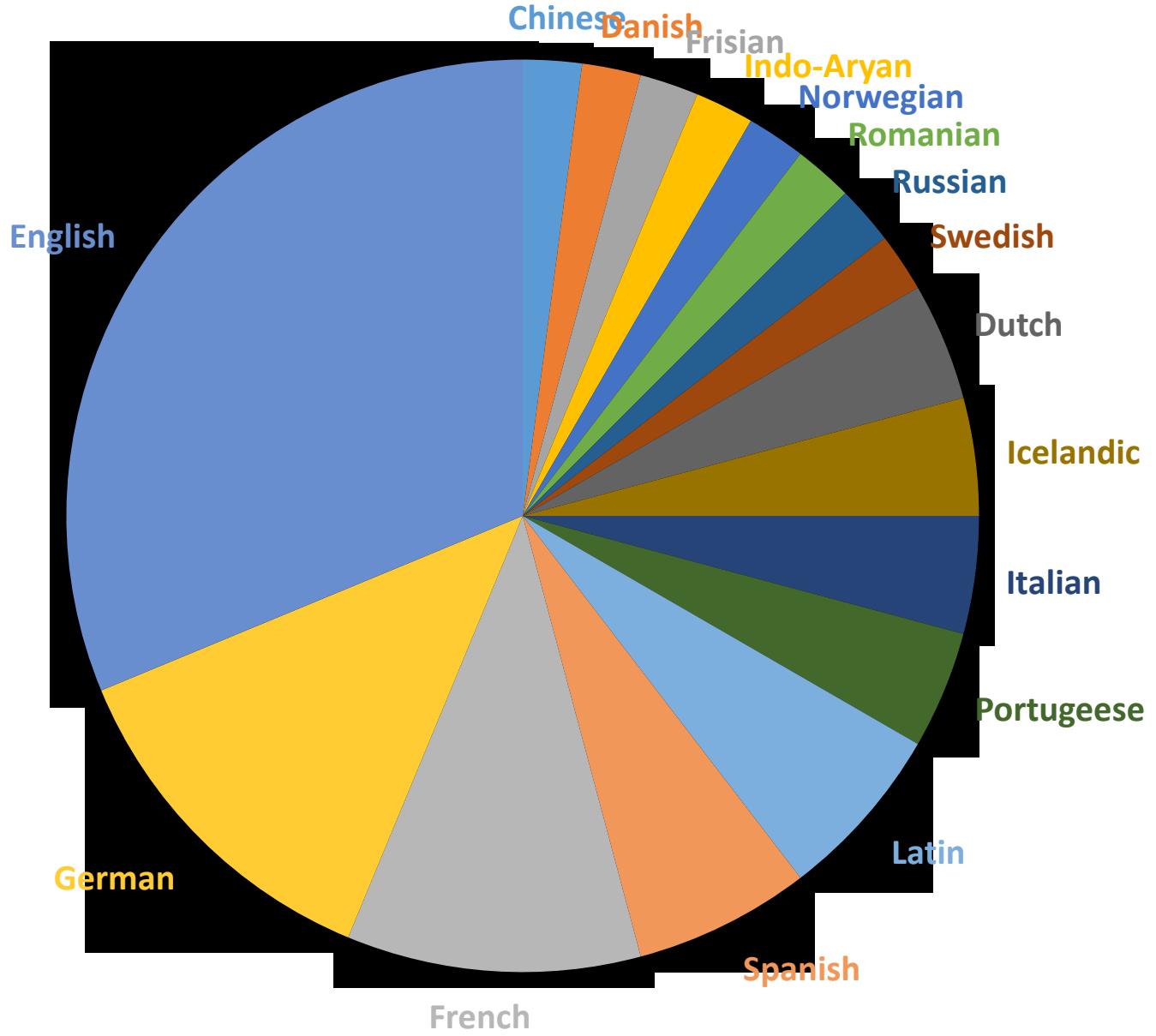
Data sets

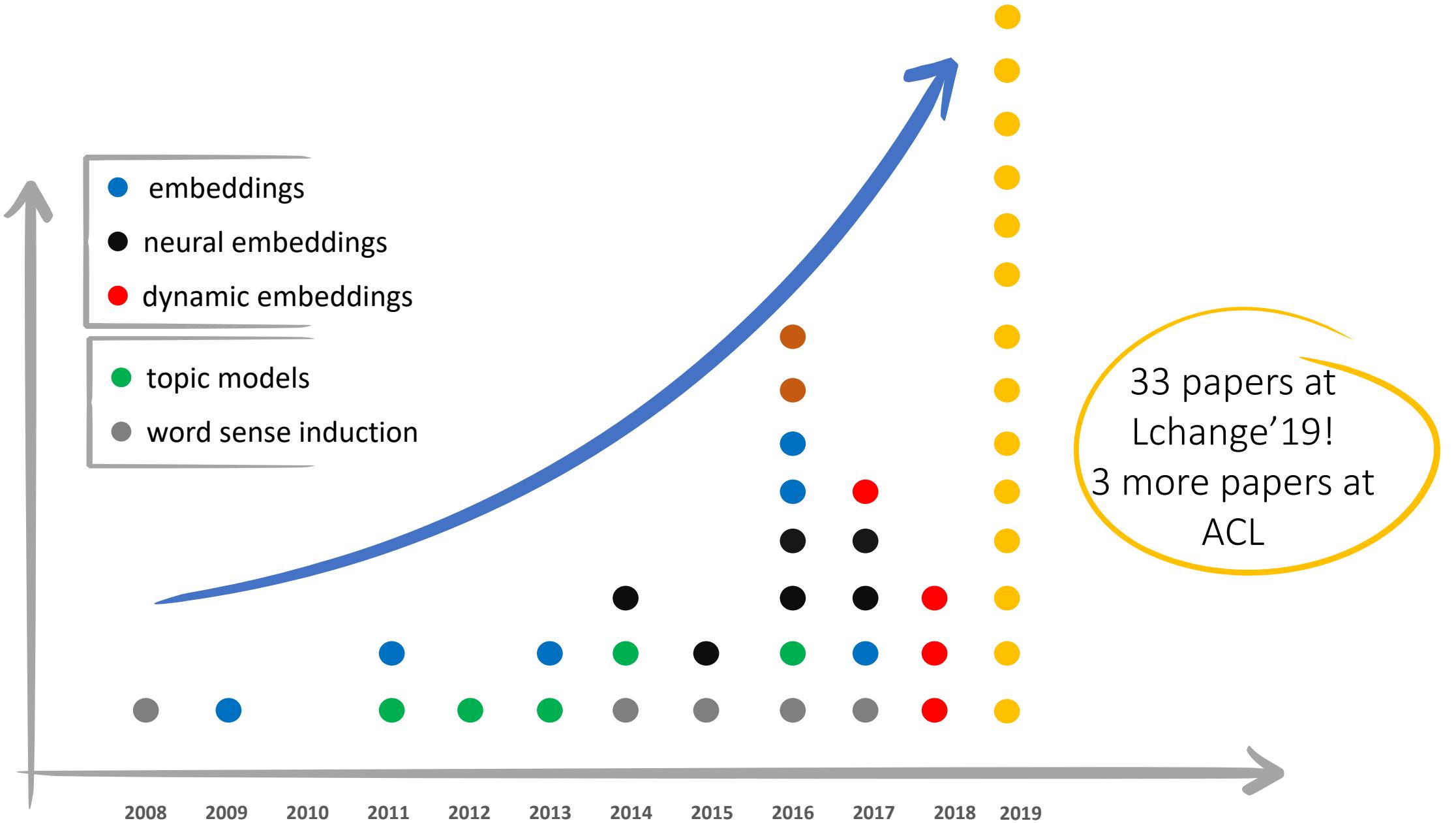
Table 3

Datasets used for diachronic conceptual change detection. Non-English .

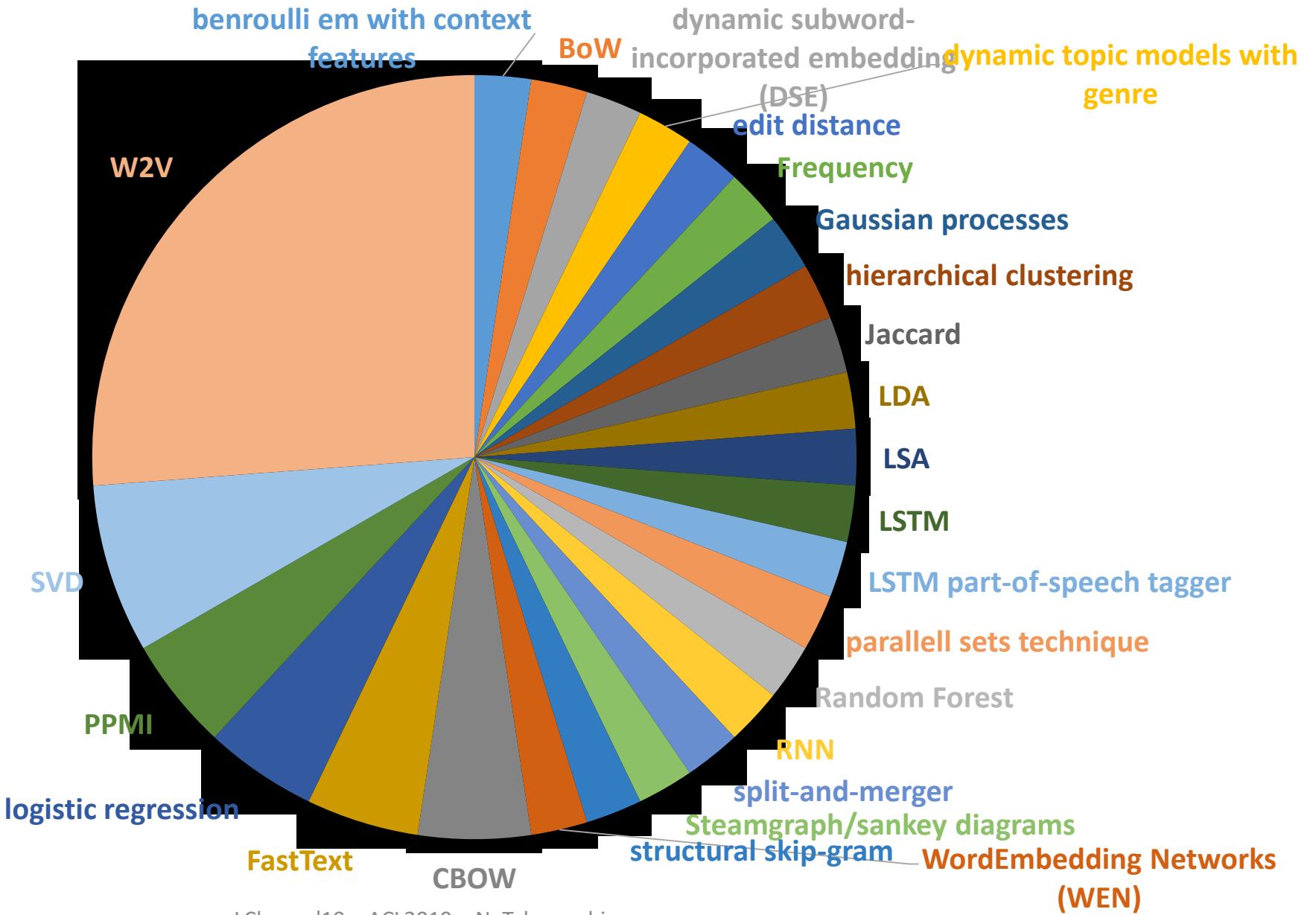
Sagi, Kaufmann, and Clark (2009a)	Helsinki corpus
Gulordava and Baroni (2011)	Google Ngram
Wijaya and Yeniterzi (2011)	Google Ngram
Lau et al. (2012)	British National Corpus (BNC), ukWaC
Cook et al. (2013)	Gigawords corpus
Cook et al. (2014)	BNC, ukWaC, Sibol/Port
Mihalcea and Nastase (2012)	Google books
· Basile et al. (2016)	Google Ngram (Italian)
· Tang, Qu, and Chen (2013, 2016)	Chinese People's Daily
Kim et al. (2014)	Google Ngram
Kulkarni et al. (2015)	Google Ngram, Twitter, Amazon movie reviews
Mitra et al. (2015)	Google Ngram, Twitter
Hamilton, Leskovec, and Jurafsky (2016b)	COHA, Google Ngram
· Eger and Mehler (2016)	COHA, Süddeutsche Zeitung, PL ⁷⁶
Azarbonyad et al. (2017)	New York Times Annotated Corpus, Hansard
· Rodda, Senaldi, and Lenci (2016)	Thesaurus Linguae Graecae
Frermann and Lapata (2016)	DATE corpus
Takamura, Nagata, and Kawasaki (2017)	Wikipedia (English and Japanese)
Kahmann, Niekler, and Heyer (2017)	Guardian (non-public)
Tahmasebi and Risse (2017a)	Times Archive, New York Times Annotated Corpus
Bamler and Mandt (2017)	Google Ngram, State of the Union addresses, Twitter
Yao et al. (2018)	New York Times (non-public)
Rudolph and Blei (2018)	ACM abstracts, ML papers ArXiv, U.S. Senate speech

Languages worked on

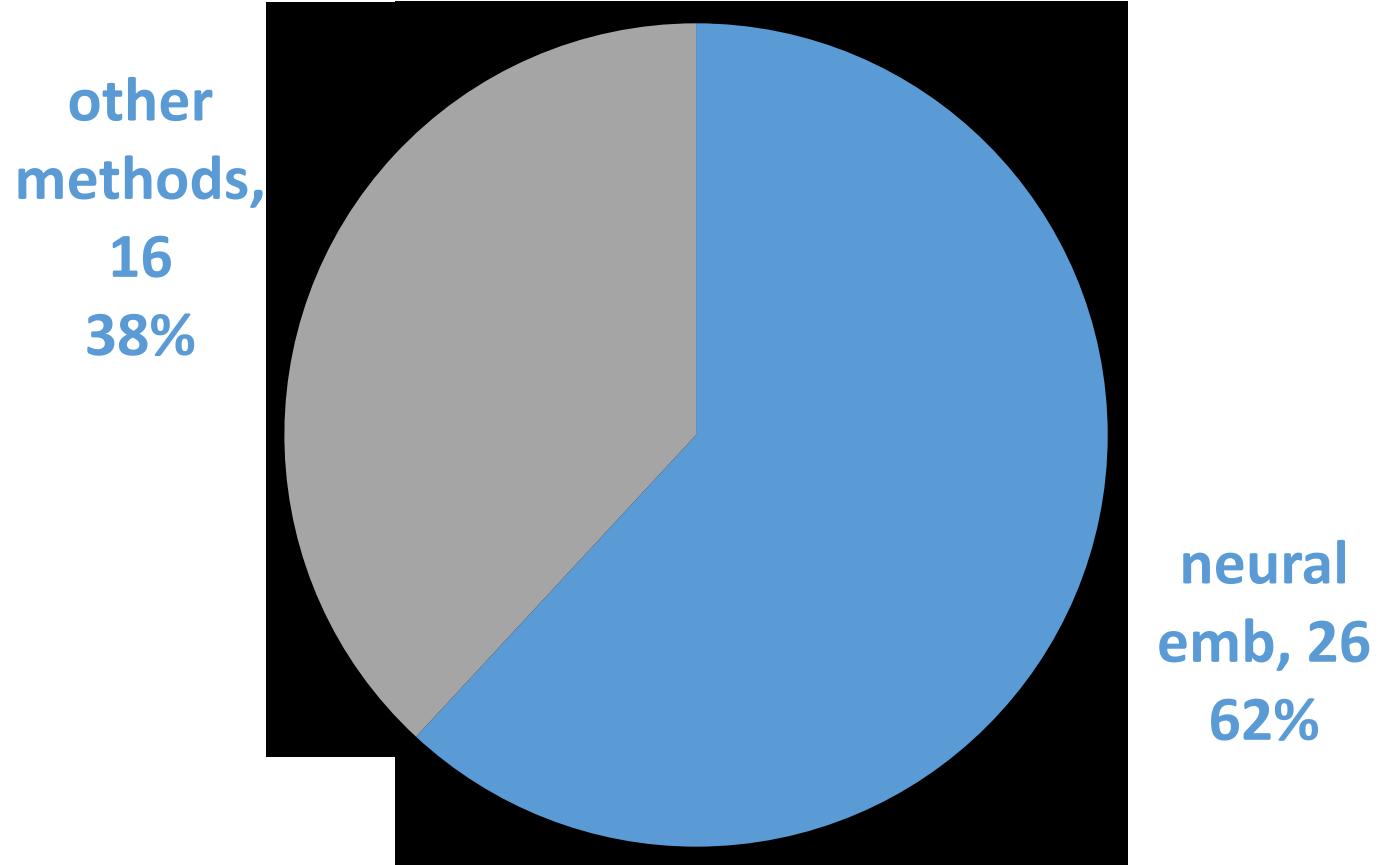




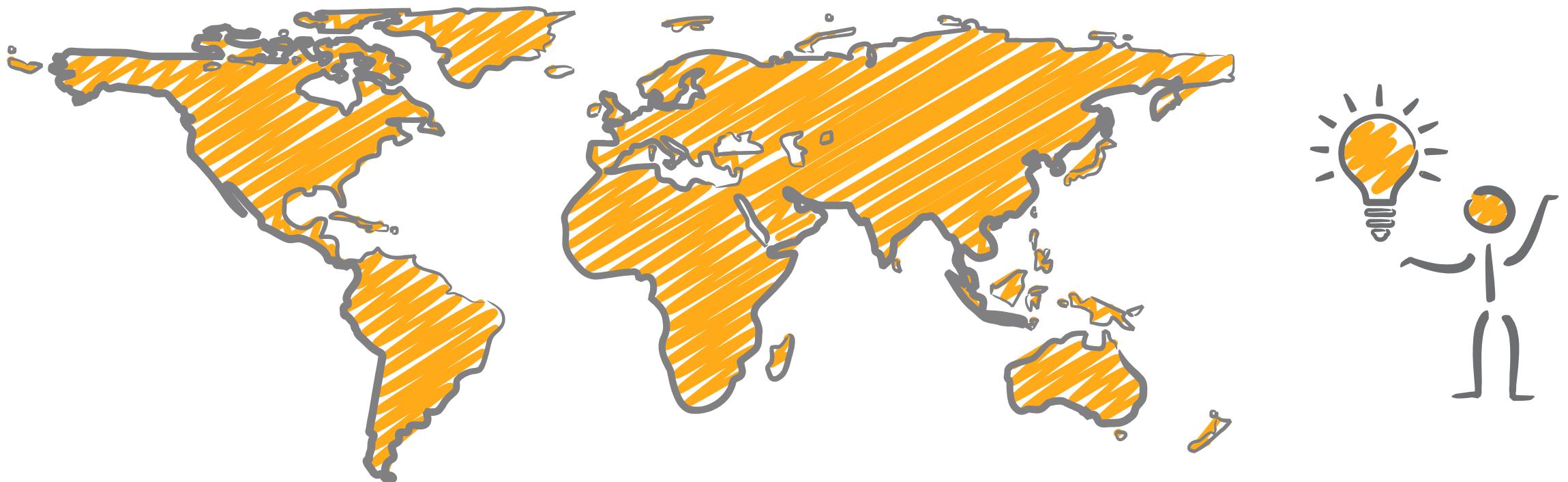
Methods



Methods



Unique meeting



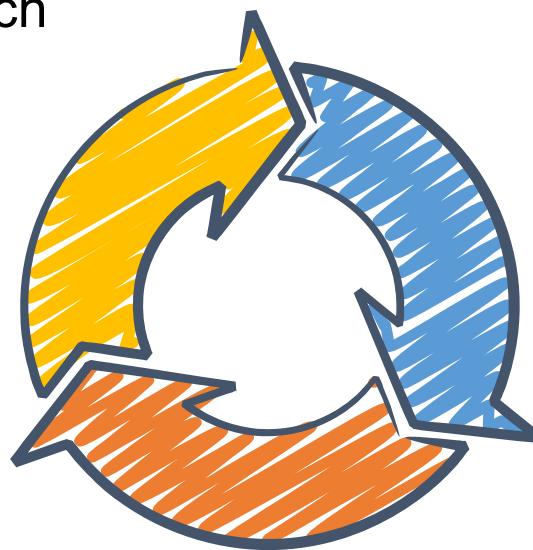


Let's talk

Bring together our communities

Quantitative approach

- Empirical
- Corpus-based
- Larger-scale
- Lower accuracy



Users

- Who need the end-result for their work

Qualitative approach

- Empirical
- Corpus-based
- Theory-driven
- Smaller-scale
- High-quality
- Case-studies

Future work:

- Evaluation (First keynote)
- Semantic change – sense-differentiated
- Methods for smaller data
- Moving forward: not reconstructing change but planning for future
- Connect computational methods with linguistic theory (Second keynote)
- Solve problems that exist

	prechosen		top	entity (S)ingle/ (P)airs	eval. method (M)anual/ (A)utomatic	time	# classes	classes	modes	
	# pos	# neg							time / sense	aware / diff
Sagi, Kaufmann, and Clark (2009a)	4	0		S	M	569y	4	2	broad./narrow.	no no
Gulordava and Baroni (2011)	0	0	100 ⁵⁴	S	M	40y	2	1	change	no no
Tang, Qu, and Chen (2013)	33	12		S	M	59	59	3	B/N/novel/change ⁵⁵	no no
Kim et al. (2014)	0	0	10/10 ⁵⁶	S/P ⁵⁷	M	110	110	1	change	yes ⁵⁸ no
Kulkarni et al. (2015)	20	0	20 ⁵⁹	S	M/A	105y/12y/2y	21/13/24	1	change	yes no
Hamilton, Leskovec, and Jurafsky (2016b)	28	0	10 ⁶⁰	S/P	M	200/190	20	1	change	no no
Rodda, Senaldi, and Lenci (2016)	0	0	50	S	M	1200y	2	1	change	no no
Eger and Mehler (2016)	0	0	21 ⁶¹	S/P	M	200/190	20/19	1	change	no no
Basile et al. (2016)	40	0		S	M	170	17	1	change	yes no
Azarbonyad et al. (2017)	24	0	5/5 ⁶²	S	M	20/11	2/2	1	change	no no
Takamura, Nagata, and Kawasaki (2017)	10	0	100/20 ⁶³	S/P	M	- ⁶⁴	2	1	change	no no
Kahmann, Niekler, and Heyer (2017)	4	0		S	M	≤ 1 ⁶⁵	48	1 ⁶⁶	change	no no
Bamler and Mandt (2017)	6	0	10	S/P	M ⁶⁷	209/230/7	209/230/21	1	change	no no
Yao et al. (2018)	4/1888 ⁶⁸	0		S	M/A	27	27	1	change	no no
Wijaya and Yeniterzi (2011)	4	2		S	M	500 ⁶⁹	500	2 ⁷⁰	change novel	yes yes ⁷¹
Lau et al. (2012)	5	5		S	M	43 y	2	1	novel	no yes
Cook et al. (2013)	0	0	30	S	M	14	2	1	novel	no yes
Cook et al. (2014)	7/13	50/164		S	M	43y/17y	2/2	1	novel	no yes
Mitra et al. (2015) ⁷²	0	0	69/50	S	M/A	488/2	8/2	3	split/join/novel ⁷³	no yes
Frermann and Lapata (2016)	4	0	200	S	M/A	311	16	2	change/novel	no yes
Tang, Qu, and Chen (2016) ⁷⁴	197	0		S	M	59	59	6	B/N/novel/change ⁷⁵	no yes
Tahmasebi and Risse (2017a)	35	25		S	M	222y	221	4	novel,B/N,stable	yes yes

Survey of Computational Approaches to Lexical Semantic Change,
Tahmasebi, Borin & Jatowt, 2018, <https://arxiv.org/abs/1811.06278>



Towards Computational Lexical Semantic Change Detection

[University of Gothenburg](#)

[Chalmers University of
Technology](#)

[Stockholm University](#)



2019 – 2022
The Swedish Research Council

Wp1: Swedish word sense induction

Wp2: Semantic change

Wp3: Lexical replacements

Wp4: Applications

WP*: Evaluation

- Integrated in all work packages

Sign up for our
news-list to keep posted

Today's schedule

Test your presentations
15 mins before your
session!

Don't forget to send
them to us!

9:00 – 9:15 Introduction

9:15 – 10:30 Haim Dubossarsky: Semantic Change in the Time of Machine Learning, Doing it Right! (Keynote)
(Chair: Lea Frermann)

10:30 – 10:45 Coffee Break

10:45 – 12:30 Session 2 (Chair: Richard Johansson)

12:30 – 13:30 Lunch Break

13:30 – 14:30 Claire Bowern: Semantic Change and Semantic Stability: Variation is Key (Keynote)
(Chair: Dominik Schlechtweg)

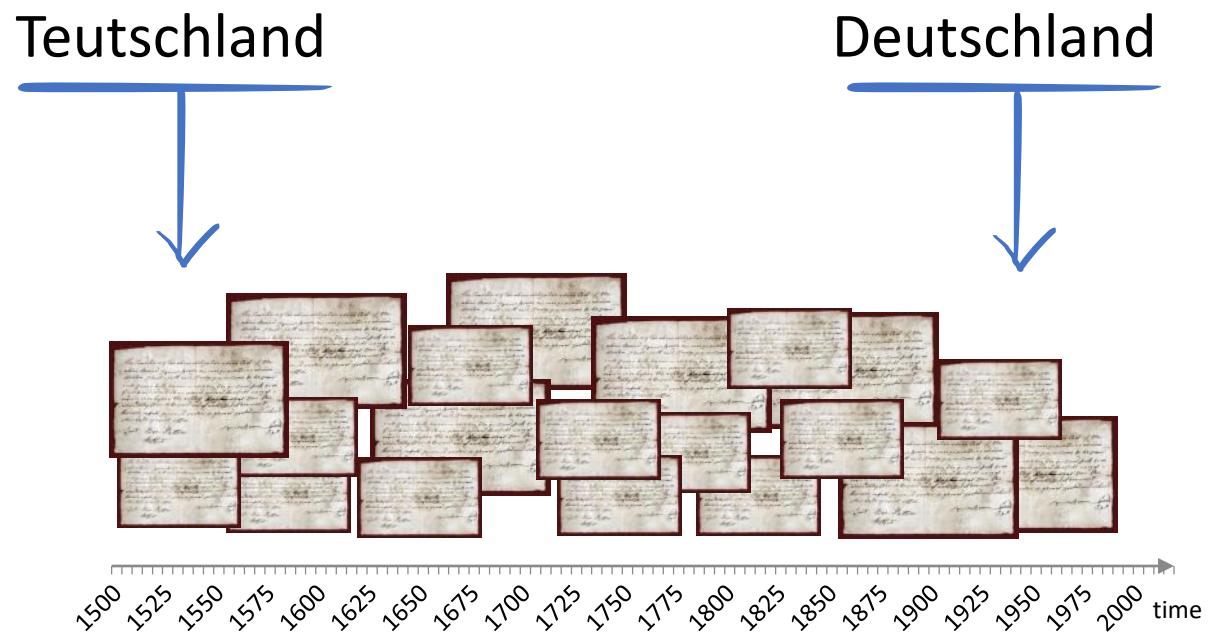
14:30 – 16:00 Poster Session with coffee

16:00 – 16:40 Session 5 (Chair: Simon Hengchen)

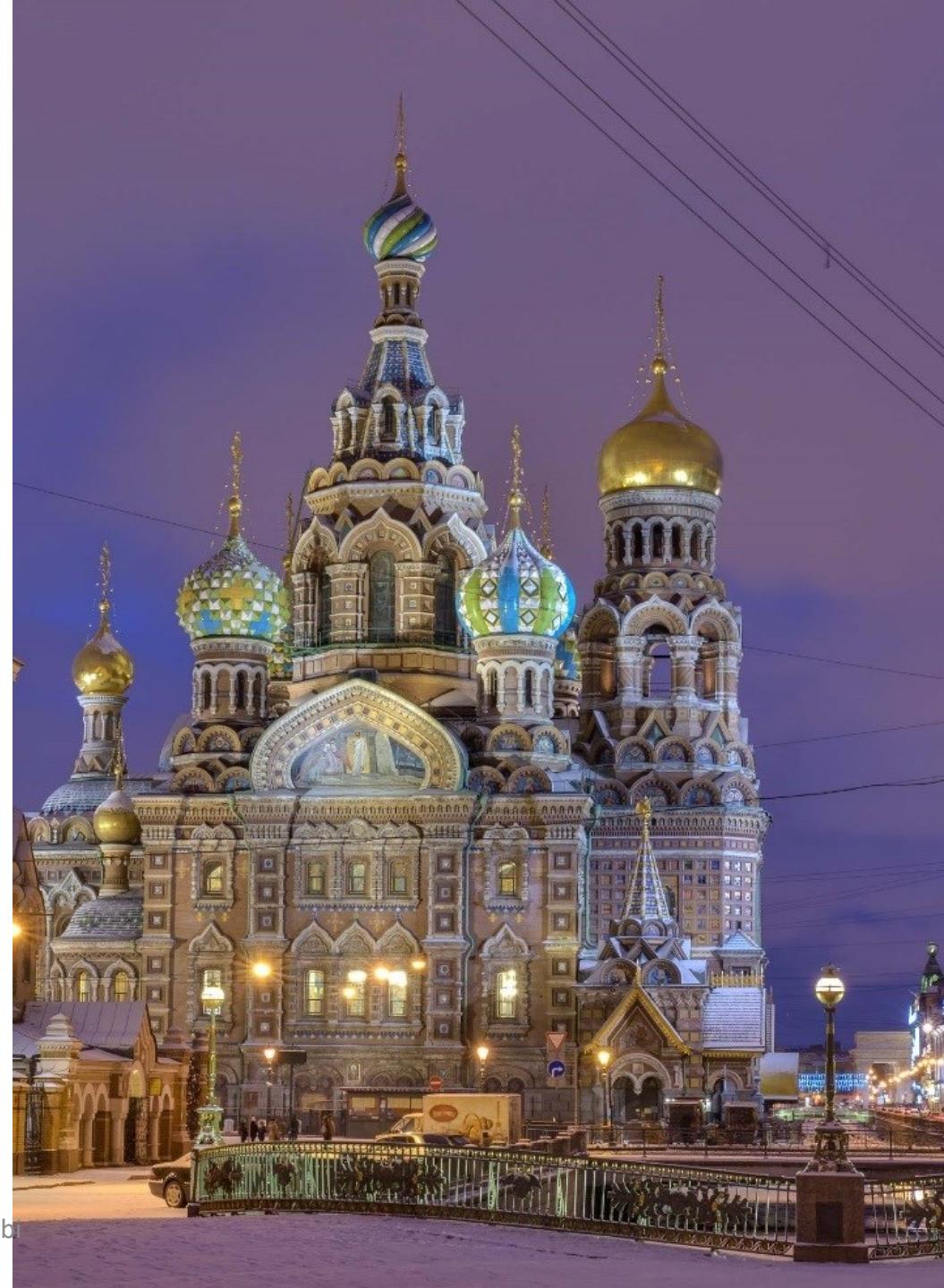
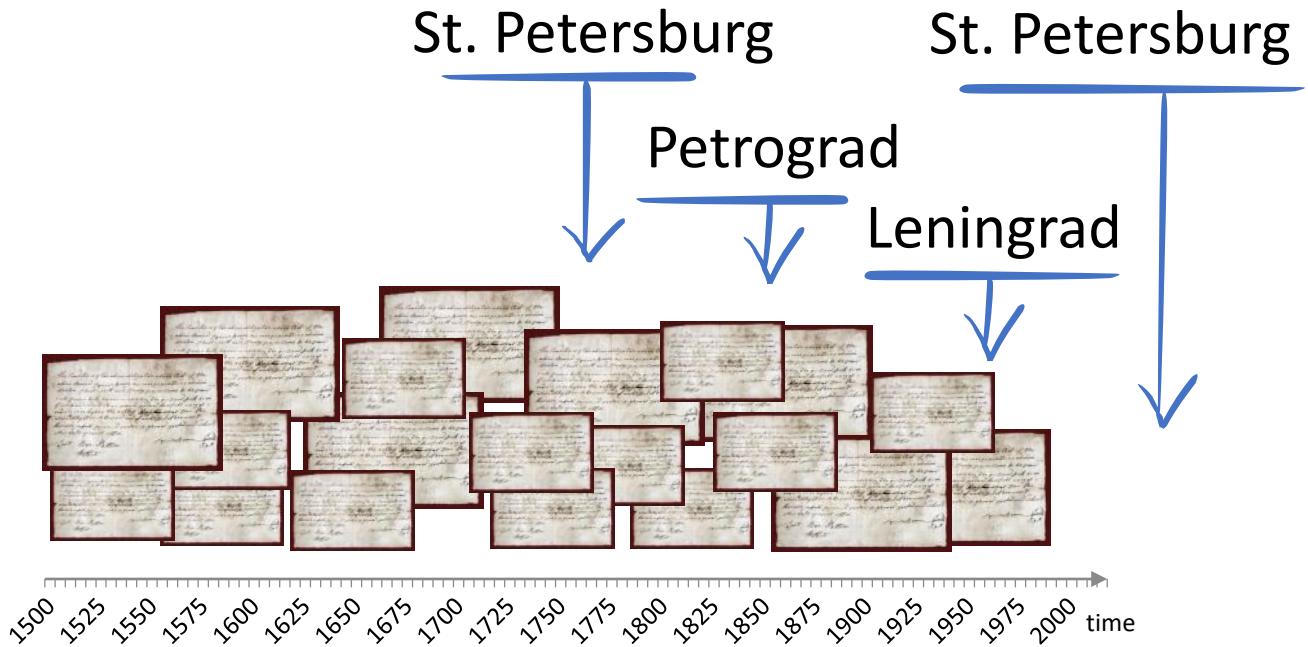
16:40 – 17:15 Discussion and Closing

Lexical Semantic Change from a computational perspective

Spelling change

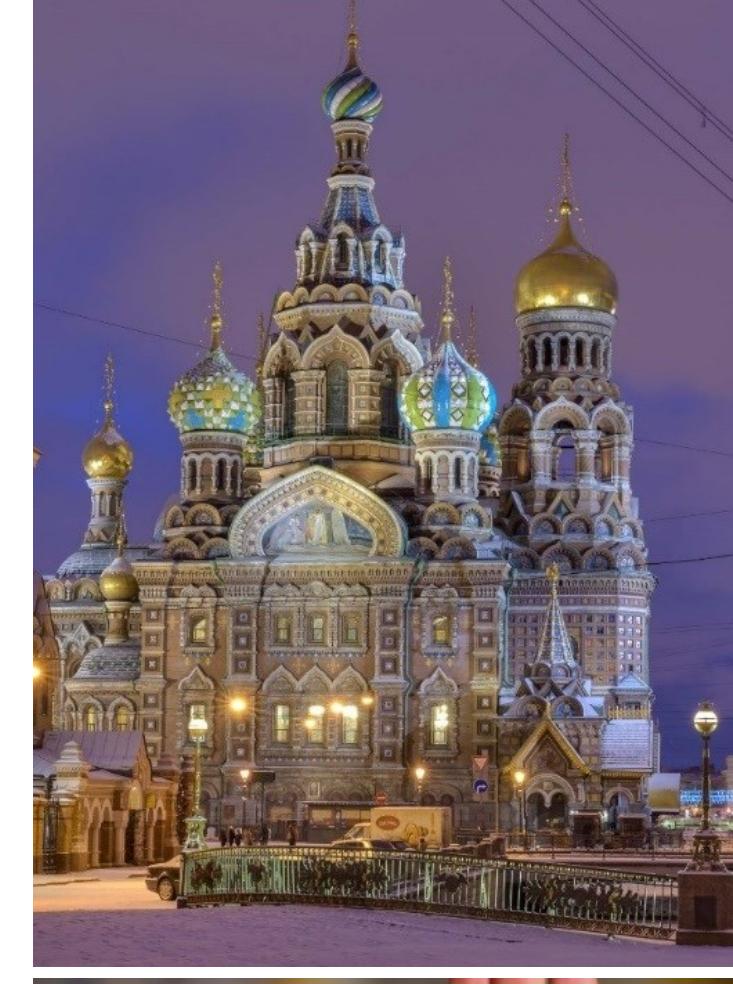
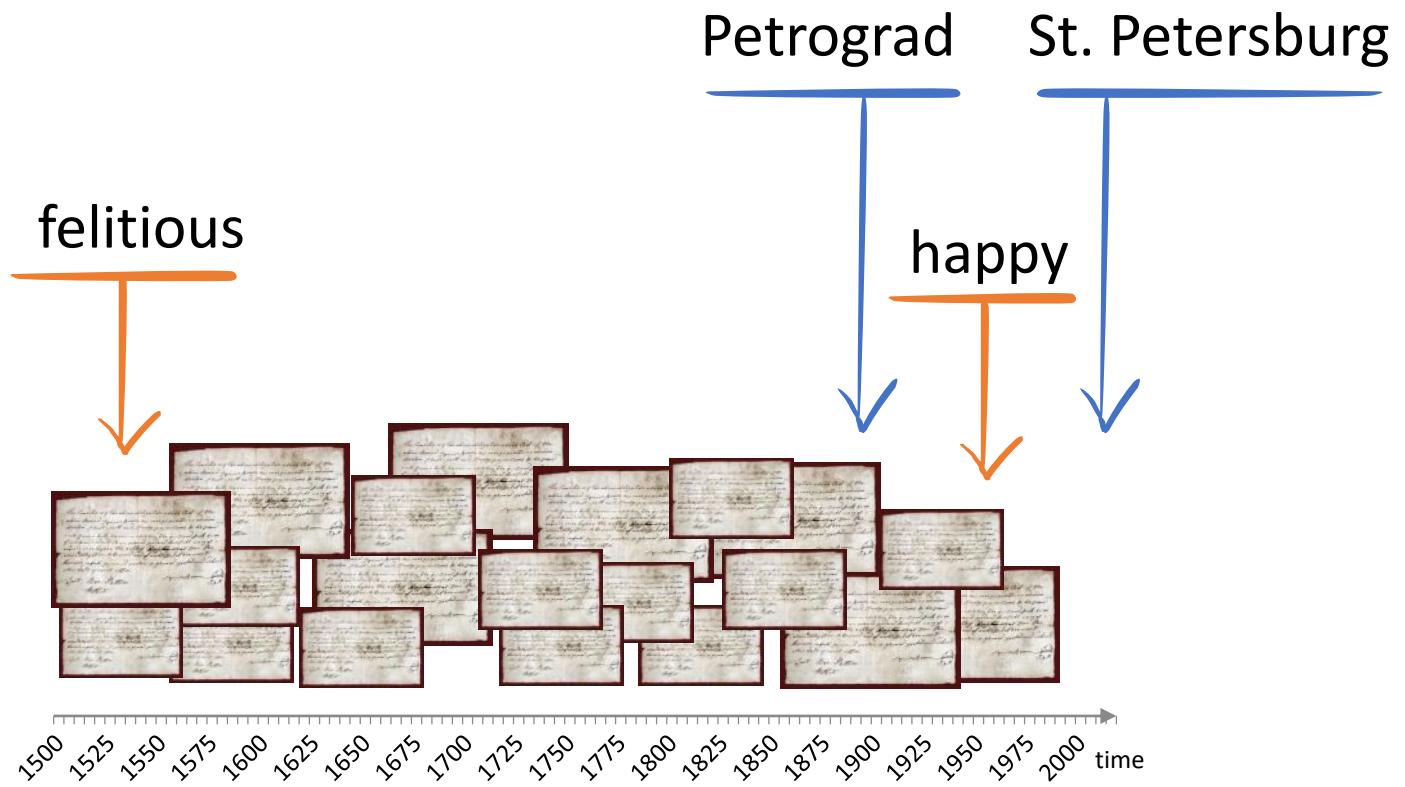


Lexical replacement: Named entity change



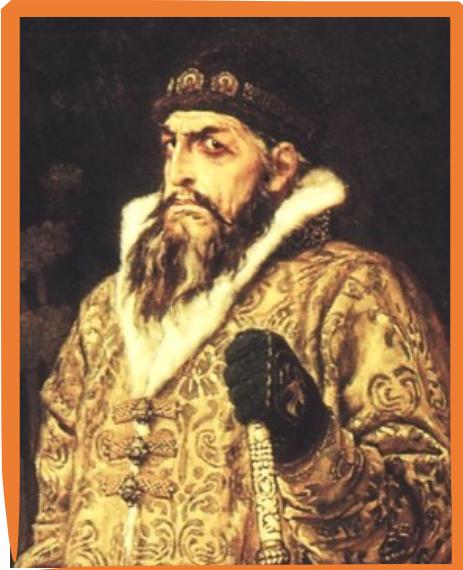


Lexical replacement:



awesome

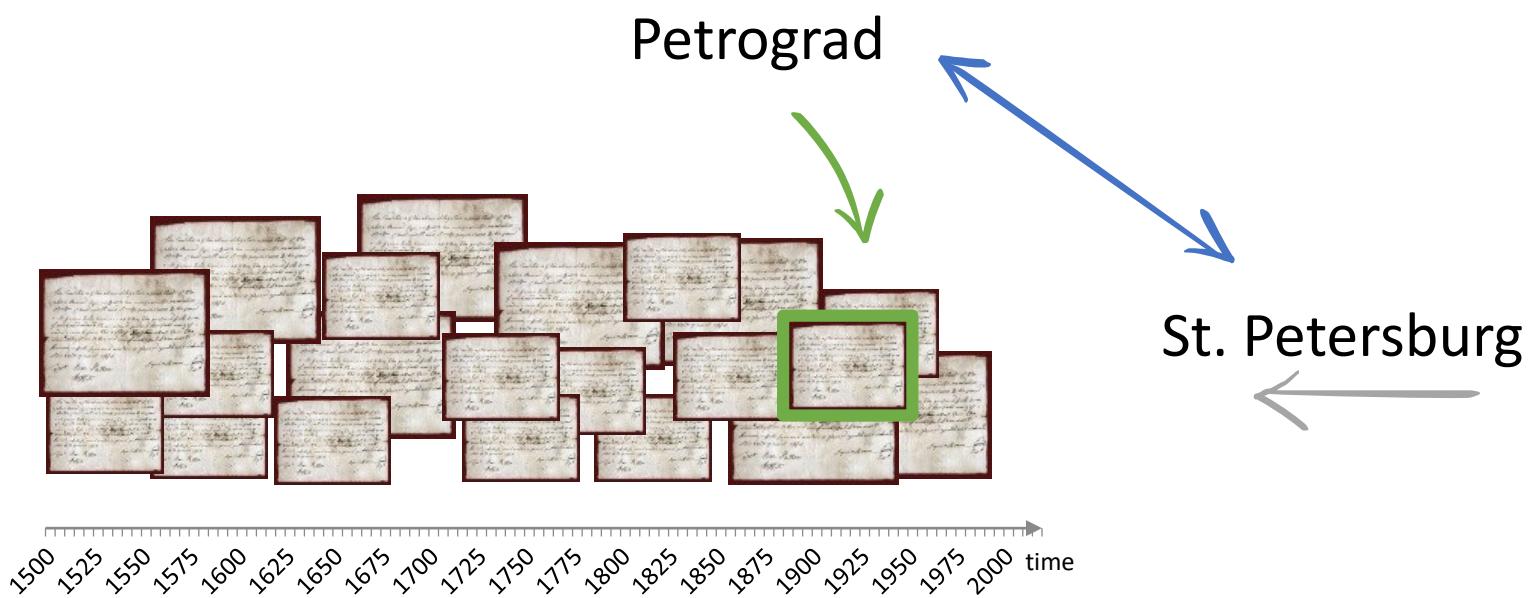
He was an
awesome leader!



He was an
awesome leader!



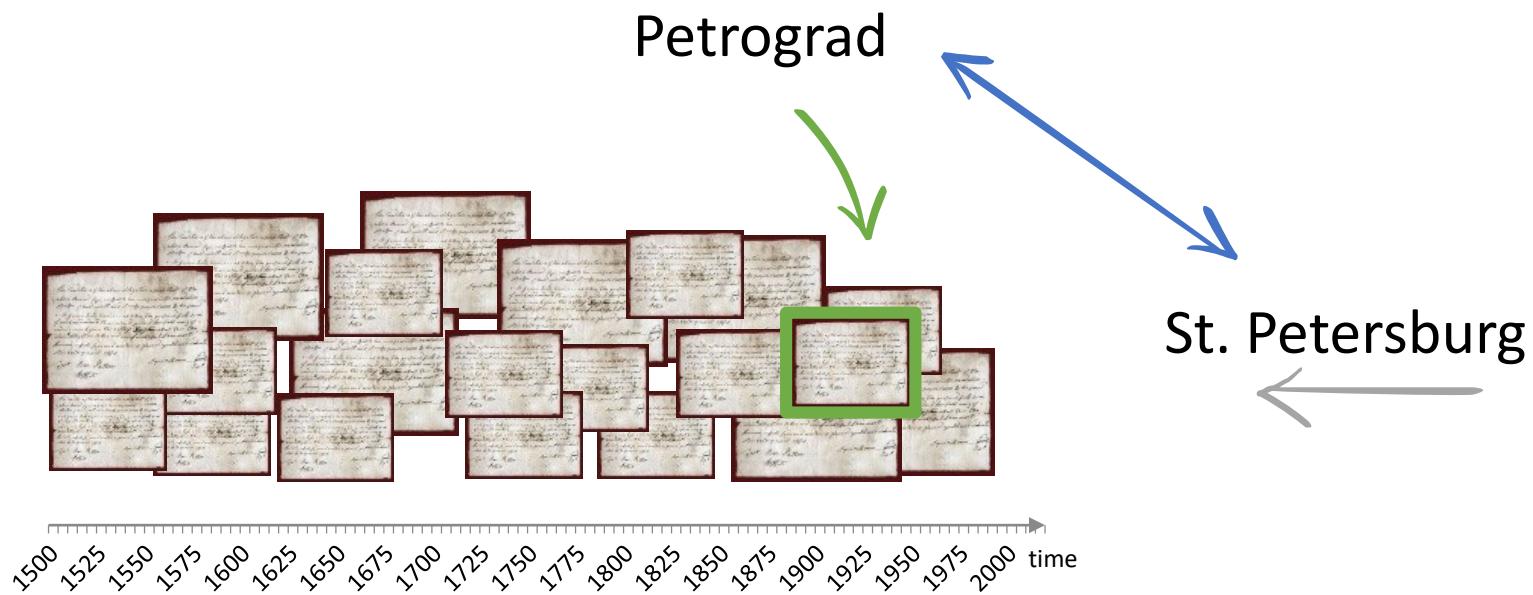
What is the problem?

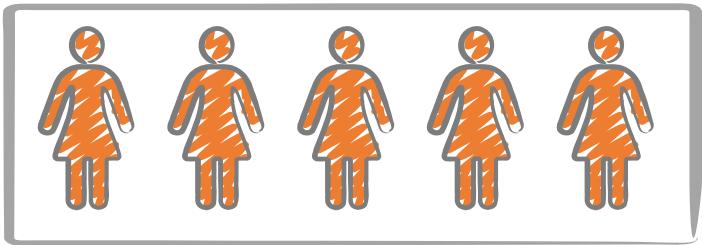
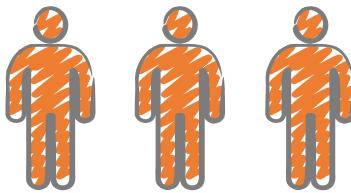


What is the problem?

Finding

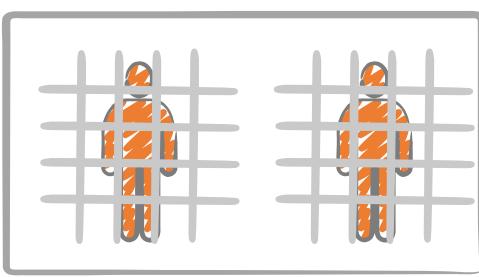
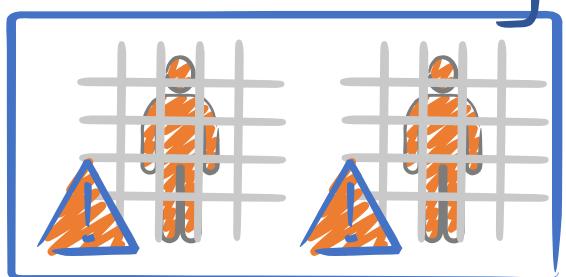
Interpreting



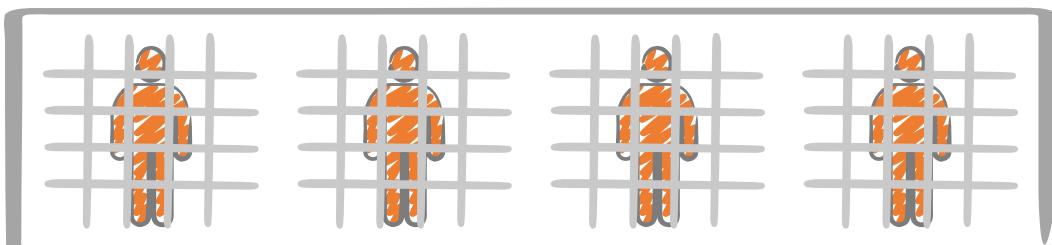


← girl

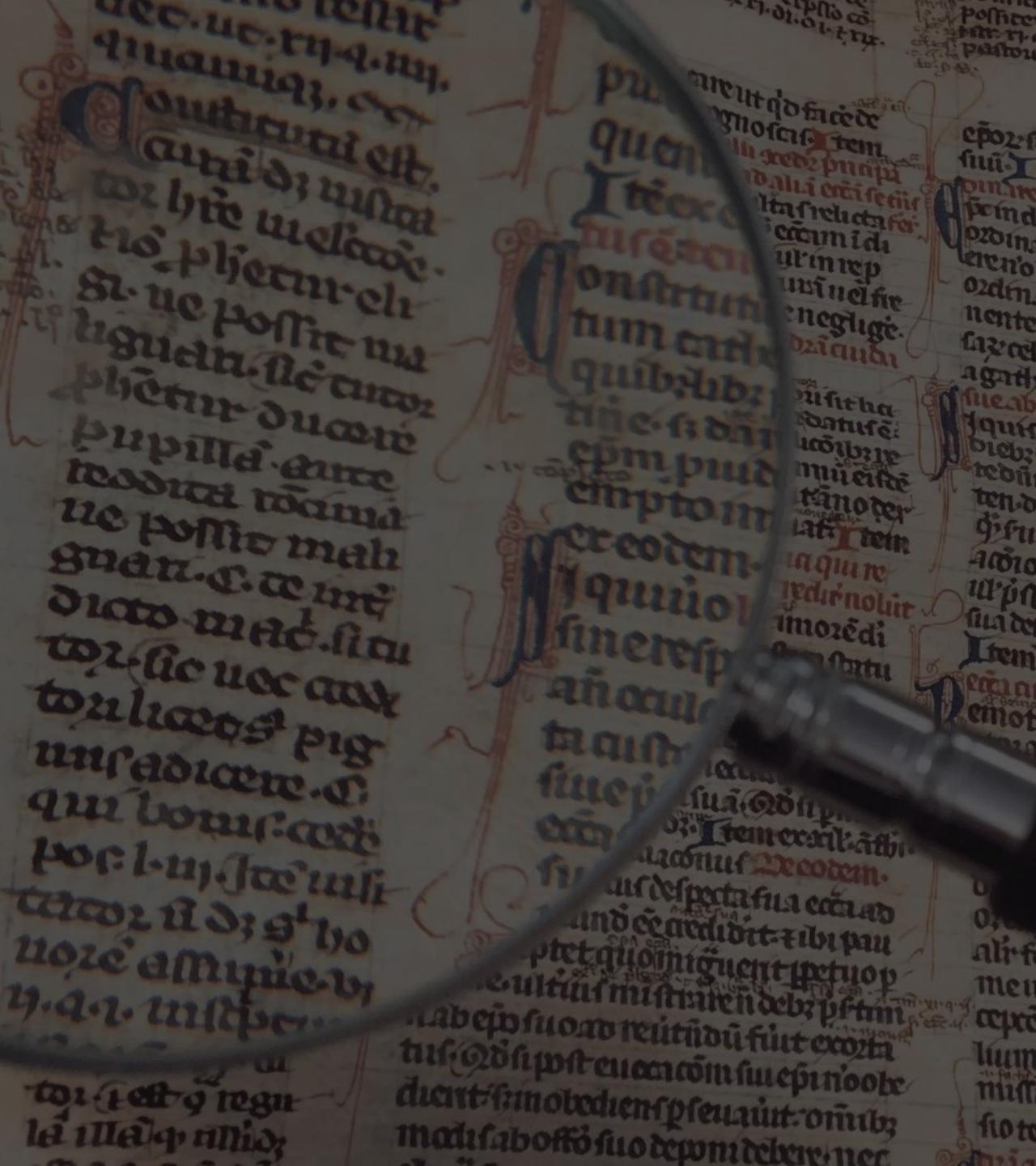
Wolf 'varg'



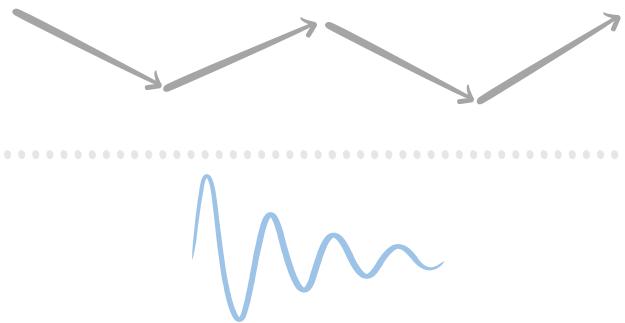
← criminal



Evaluation



Evaluation



signal change



collective text



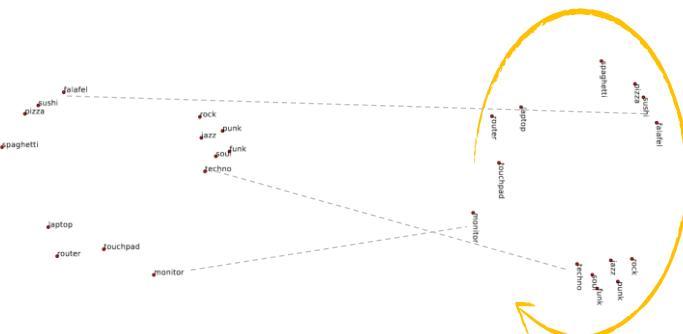
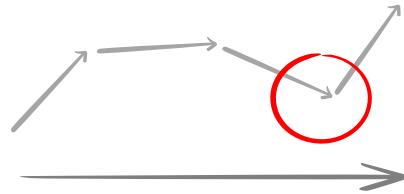
individual text



individual



LSC – individually trained embedding spaces



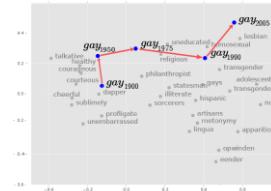
Vector space image:
Nieto Pina and
Johansson, RANLP'15

Track an individual
word w over time

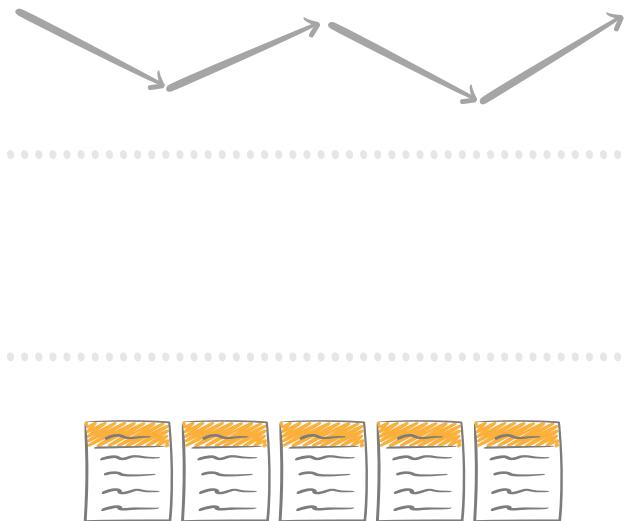
Change
point/degree
detection

multiple
time points
align

Single-point
embedding space
 t_i



Evaluation



signal change

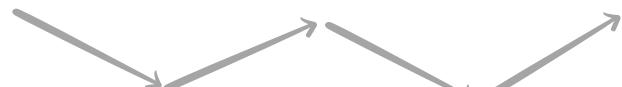
collective text

minimum

optimum

medium

Evaluation



signal change



individual text



individual

minimum

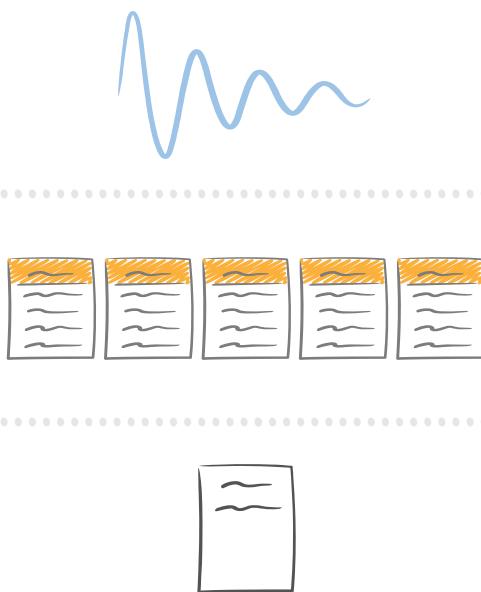
optimum

medium



?

Evaluation



signal
topic, cluster, vector...

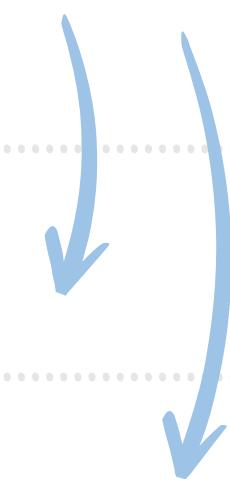
collective text

individual text

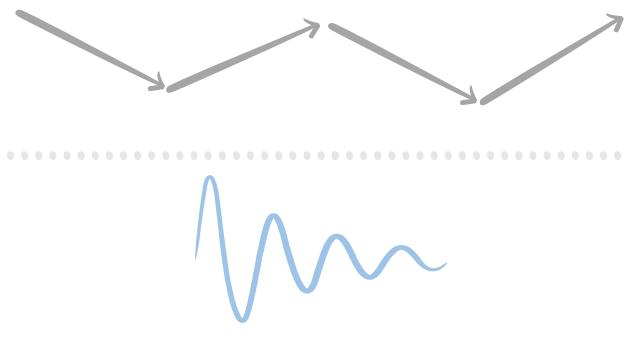
minimum

optimum

medium



Evaluation



signal change



collective text

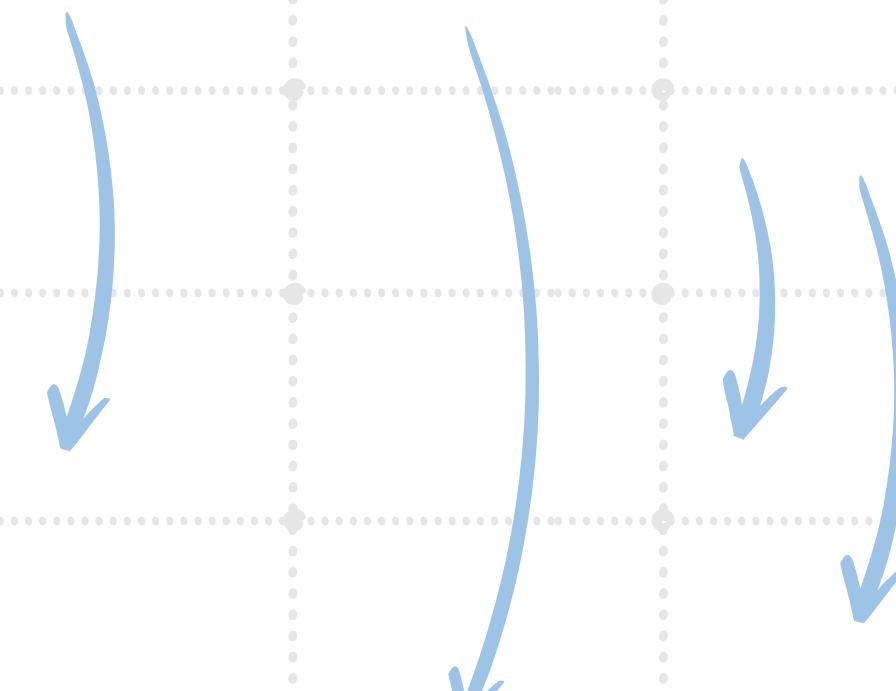


individual text



individual

minimum optimum medium



Evaluation

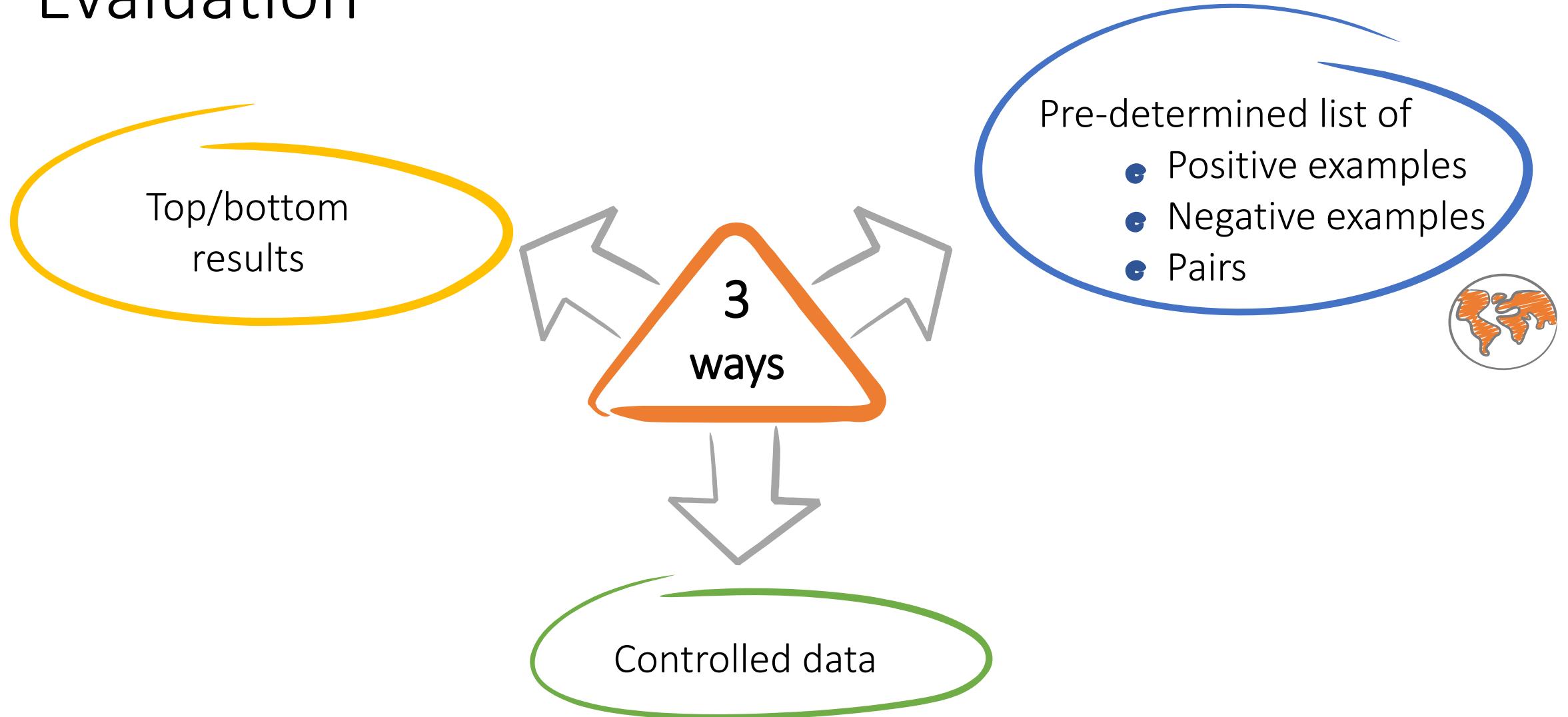


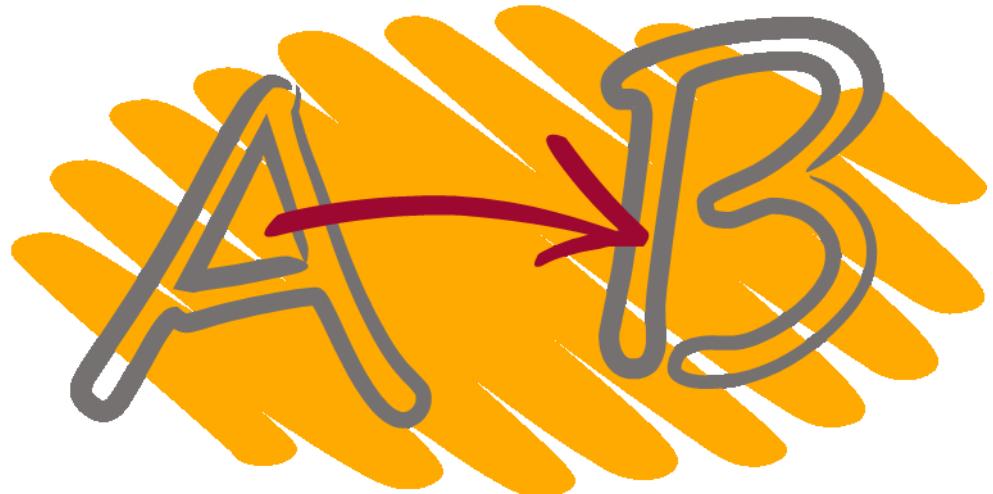
signal change



Currently, we do none
of the previous

Evaluation





<https://languagechange.org/semeval>
semeval2020lexicalsemanticchange@turing.ac.uk

SemEval 2020 Task 1:

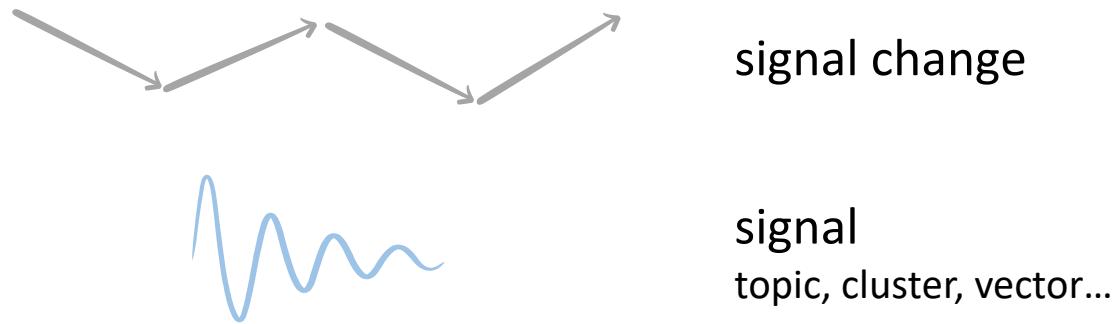
Unsupervised Lexical Semantic Change Detection

Dates:
Trial data July 31, 2019
Training data Sept. 4, 2019
Test data Dec. 3, 2019
Evaluation Jan. 10, 2020
Evaluation Jan. 31, 2020
Paper subm. Febr. 23, 2020

English
German

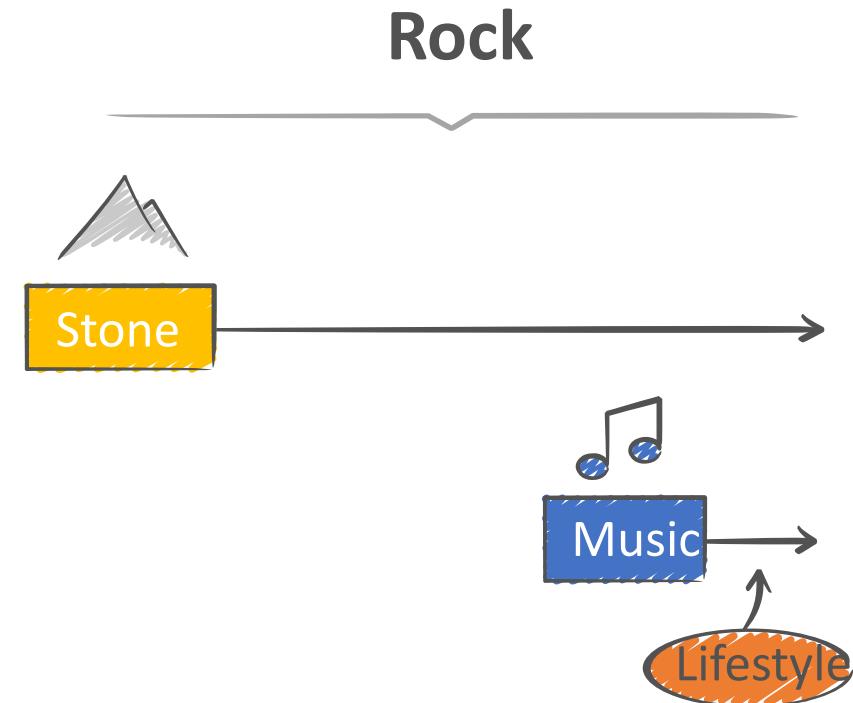
Latin
Swedish

Singal and Signal change

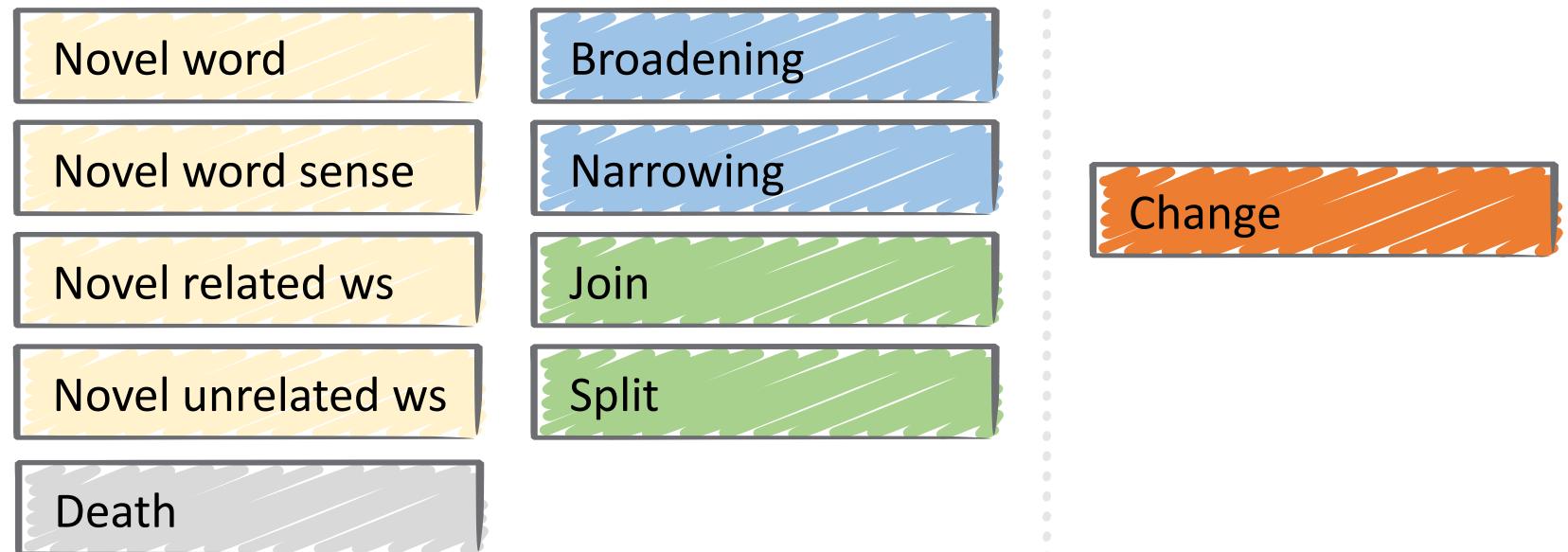
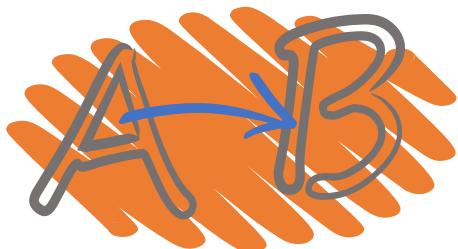


Interpretable change signal

Find changes **automatically**:
find **what** changes, **how** it
changed and **when** it changed



Change type

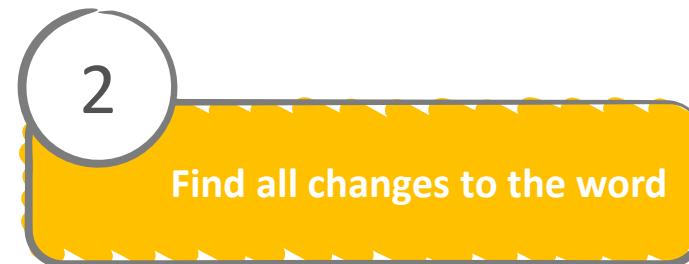
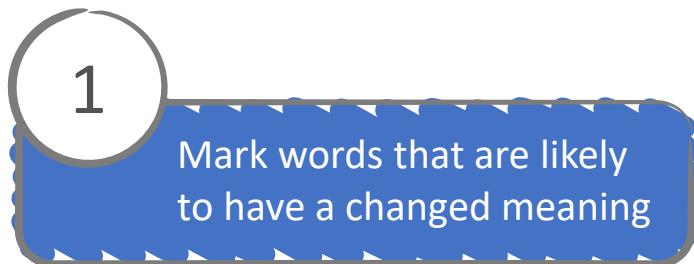


Sense-differentiated

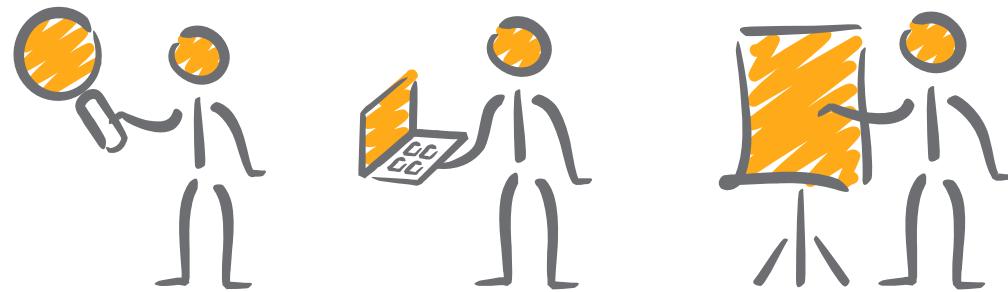
Single-sense

Return to individual instances:

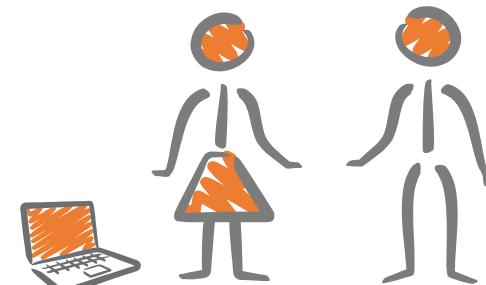
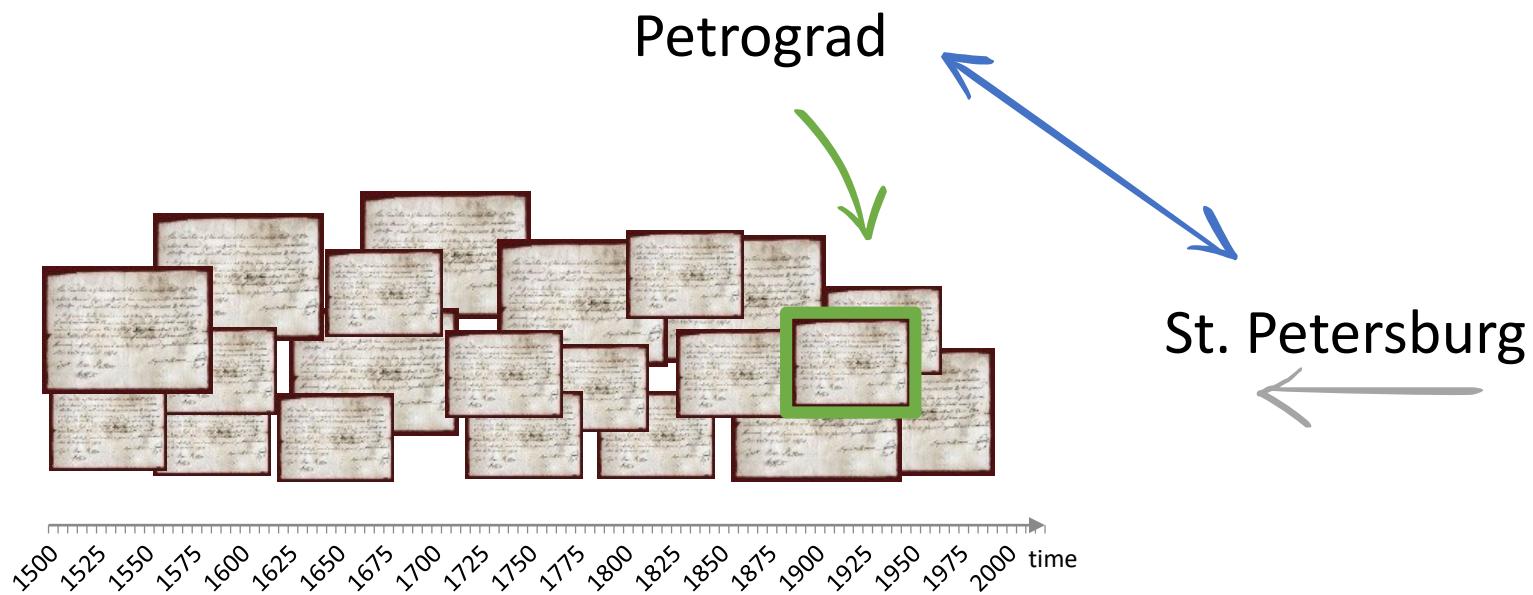
Given a word in a document at time t



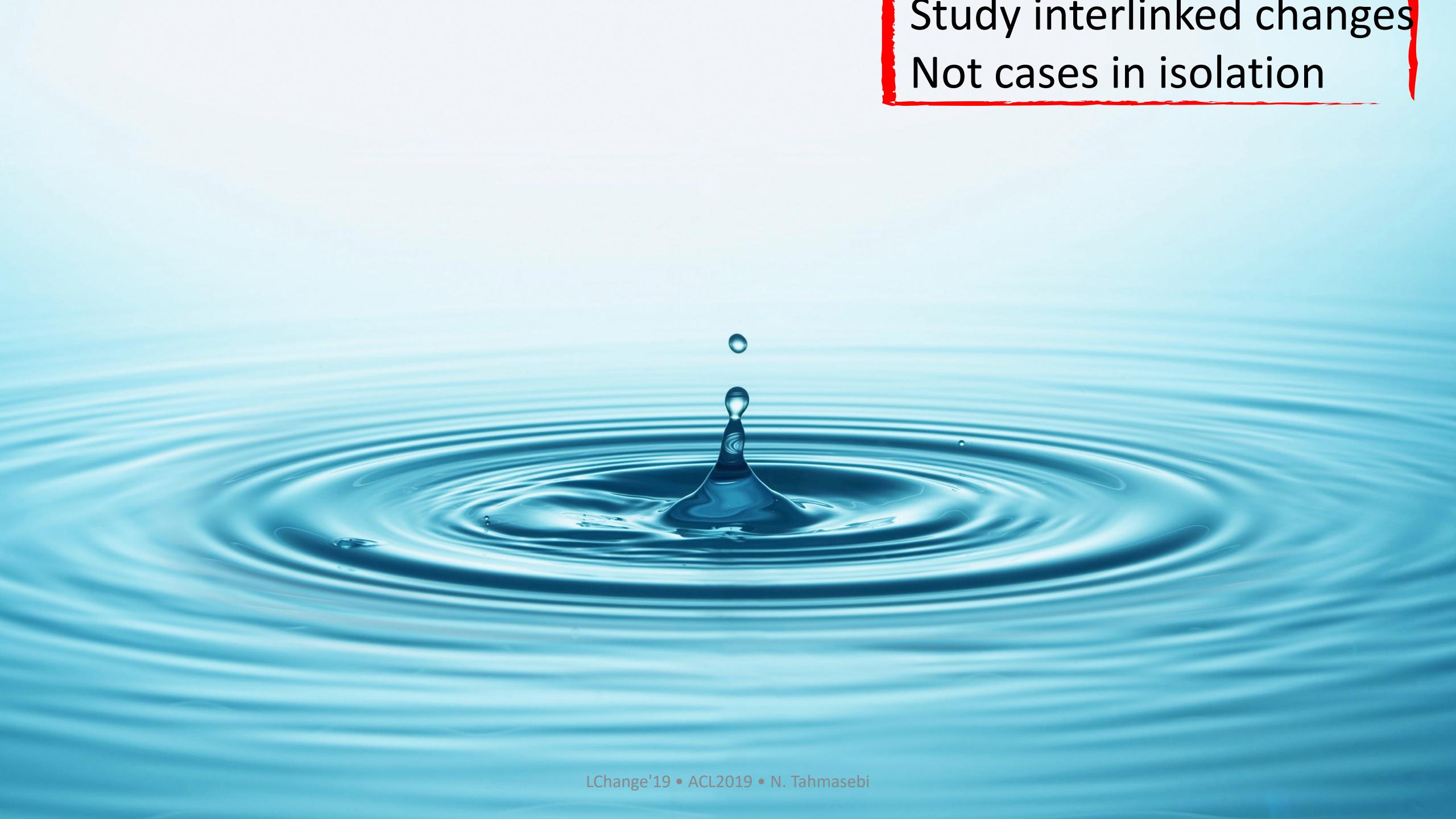
Who is the end-user?



What is the problem?



Study interlinked changes
Not cases in isolation



Solve problems that exist!



Our second keynote:
Claire Bowern

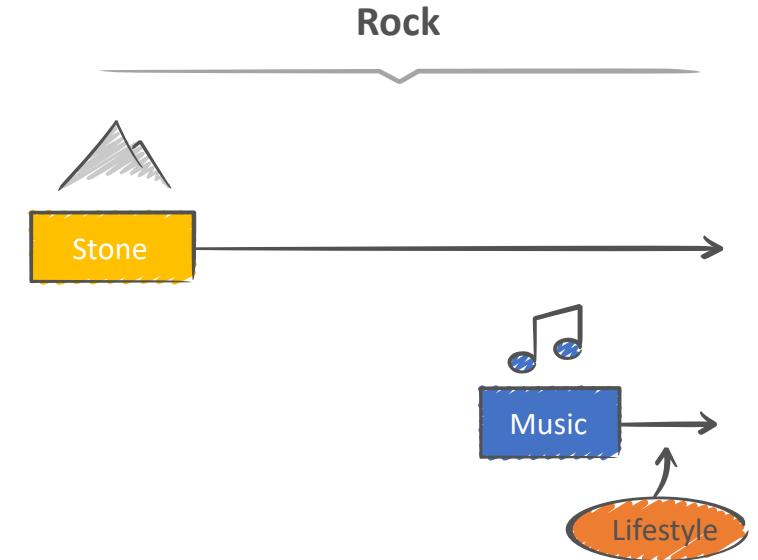
Not only
reconstruct change
but moving forward:

- Construct archives, store data to avoid problems in the future?
- Temporal IR?



Open challenges

- Sense-differentiated embeddings
- Methods for smaller-scale data
- More languages
- Methods for testing & evaluation
- Evaluating usefulness for follow-up analysis
- What is computational meaning and how does it affect change detection?





Thank you for attending!
www.languagechange.org