On Lexical Semantic Change and Evaluation

Nina Tahmasebi, PhD
University of Gothenburg
Stuttgart, June 25th, 2019
Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019

33 papers at Lchange’19!
3 more papers at ACL
Outline

Lexical Semantic Change

Computational methods for LSC

Evaluation

Nina Tahmasebi, On Lexical Semantic Change and Evaluation,
Stuttgart, June 2019
Lexical Semantic Change
LiWA – Living Web Archives

- dealing with terminology evolution
- preparing for evolution aware access support

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Increasing amount of historical texts in digital format

Easy digital access for anyone! Not only scholars.

Possibility to digitally analyze historical documents at large scale.

Information from primary sources Not only modern interpretations.

Text-based Digital Humanities

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Spelling change

Teutschland

Deutschland

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Lexical replacement:
Named entity change

Nina Tahmasebi, On Lexical Semantic Change and Evaluation
Stuttgart, June 2019
Lexical replacement:

felitious

Petrograd

happy

St. Petersburg

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
awesome

He was an awesome leader!

He was an awesome leader!

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
What is the problem?
What is the problem?

Finding

Interpreting

Petrograd

St. Petersburg

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Sebastini’s benefit last night at the Opera House was overflowing with the fashionable and gay.
Nina Tahmasebi, On Lexical Semantic Change and Evaluation,
Stuttgart, June 2019
Sebastini’s benefit last night at the Opera House was overflowing with the fashionable and gay.

The Times, April 27th, 1787

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
What is the problem?

Finding

Interpreting
Wolf ‘varg’

girl

criminal
Semantic change
Same word, different sense

Lexical change
\( w(t_1) \rightarrow v(t_2) \)
- adjectives
- verbs
- ...

Named Entity Change
People, places, companies

Spelling variation
Same meaning, different spelling

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Aims

Find word sense changes **automatically** to find **what** changes, **how** it changed and **when** it changed
Vision

Given a word in a document at time $t$:

1. Mark words that are likely to have a changed meaning.

2. Find all changes to the word.

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Lexical Semantic Change

The (historical) linguistic perspective
Semasiological perspective

Word: Rock

Concept/meaning: Stone, Music

Semantic change

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Onomasiological perspective

word

concept/meaning

Lexical replacement

gay

happy

Happy

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Ono- and Semasiological are interlinked!

word

concept/meaning

red

Red
Ono- and Semasiological are interlinked!

Word: red, pink

Concept/meaning: Red
One more example

- **happy**
- **gay**

- **word**
- **concept/meaning**

- Happy
- Homosexual

- semasiological
- onomasiological

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Why?
A division of the semantic field ‘sharp’

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
A corpus comparison: [https://spraakbanken.gu.se/korp/](https://spraakbanken.gu.se/korp/)

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Methods for computational semantic change
Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Change type

- Novel word
- Novel word sense
- Novel related ws
- Novel unrelated ws
- Death

- Broadening
- Narrowing
- Join
- Split

Sense-differentiated
Single-sense

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
embeddings
neural embeddings
dynamic embeddings

Tahmasebi et al. 2008
Sagi et al. 2009
Gulordava & Baroni 2011
Tang et al. 2013
Kim et al. 2014
Kulkarni et al. 2015
Hamilton et al. 2016
Rodda et al. 2016
Costin-Gabriel & Rebedea Tjong Kim Sang 2016
Costin-Gabriel & Rebedea Tjong Kim Sang 2016
Eger and Mehler 2016
Azarbonyad et al. 2017
Takamura et al. 2017
Kahnmann & Heyer 2017
Bamler & Mandt 2017
Yao et al. 2018
Rudolph & Blei 2018
Rosenfeld & Erk 2018
Wijaya & Yentizerzi 2011
Lau et al. 2012
Cook et al. 2013
Cook et al. 2014
Mitra et al. 2014
Mitra et al. 2015
Tang et al. 2016
Kim et al. 2014
Kulkarni et al. 2015
Frerman & Lapata 2016
Tahmasebi & Risse 2017
Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019

topic models
word sense induction
Context-based method

Sagi et al.  
GEMS 2009

Data set split in approp. sets

- Broadening of sense
- Narrowing of sense
- With grouping: Added/removed sense

BUT:

1. No discrimination between senses
2. No alignment of senses over time!

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Word embedding-based models

Kulkarni et al. WWW’15

- Project a word onto a vector/point (POS, frequency and embeddings)
- Track vectors over time

Kim et al. LACSS 2014
Basile et al. CLiC-it 2016
Hamilton et al. ACL 2016

Image: Kulkarni et al. WWW’15

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Dynamic Embeddings

Share data across all time points
Avoids aligning

Bamler & Mandt:
  • Bayesian Skip-gram

Yao et al:
  • PPMI embeddings

Rudolph & Blei:
  • Exponential family embeddings (Beronoulli embeddings)

Sharing data is highly beneficial!

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Temporal Referencing

- Share contexts across all time points
- Individual vectors for words for each bin
- Avoids aligning

Sharing data is **highly beneficial!**

Dubossarsky et al
- SGNS
- PPMI embeddings

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Topic-based methods

1. Topic model (HDP)
2. Assign topics to all instances of a word.
   If a word sense $WS_i$ is assigned to collection 2 but not 1 then $WS_i$ is a **novel** word sense.
3. **BUT:**
   A. Only two time points (typically there is much noise!)
   B. **No alignment** of senses over time!

BNC

ukWaC

---

Lau et al.  Wijaya & Yeniterzi  EACL 2014  DETECT ’11

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Downsides topic models

Sense change \rightarrow Topic change

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Word sense induction

Step 1:
Word sense induction (curvature clustering) individual time slices

Step 2:
Detecting stable senses → units

Step 3:
Relating units → Paths

1. Stone
2. Rock
3. Music
4. Lifestyle

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Complexity

$O(|S|^T)$
LSC – individually trained embedding spaces

1. Embedding space

2. Alignment

3. Change degree/point

4. Differentiate between change types

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

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
LSC – dynamic embedding spaces

0. Embedding space

1. Smoothness

2. Change point

3. Differentiate between change types

Align while training

Track an individual word w over time

Change point/degree detection

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019

Bamler and Mandt, 2018
Sense-differentiated embedding spaces

1. Word sense induction
2. Word sense disambiguation

Image: Johansson and Nieto Pina, NODALIDA 2015

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Sense-differentiated dynamic embeddings

1. Word sense induction
2. Word sense disambiguation
3. Smoothness
4. Change point
5. Differentiate between change types

Align while training, with multiple senses

Track a word’s senses individually over time

Change point detection

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Evaluation
<table>
<thead>
<tr>
<th>collective text</th>
<th>minimum</th>
<th>optimum</th>
<th>medium</th>
</tr>
</thead>
<tbody>
<tr>
<td>signal change</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>signal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>topic, cluster, vector...</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>collective text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>individual text</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>individual</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Evaluation

signal change

minimum
 optimum
 medium

collective text

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Evaluation

signal change

minimum optimum medium

individual text

individual

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Evaluation

signal
topic, cluster, vector...

minimum

optimum

medium

collective text

individual text

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Evaluation

- Signal change
- Signal
- Topic, cluster, vector...
- Collective text
- Individual text
- Individual

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Evaluation

signal change

minimum

optimum

medium

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Evaluation

3 ways

Top/bottom results

Pre-determined list of
- Positive examples
- Negative examples
- Pairs

Controlled data

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Representativeness
Representativeness (2)
<table>
<thead>
<tr>
<th>Study</th>
<th># pos</th>
<th># neg</th>
<th>Top</th>
<th>Span</th>
<th># points</th>
<th># classes</th>
<th>Classes</th>
<th>Modes</th>
<th>Time / Sense Aware / Diff</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sagi, Kaufmann, and Clark (2009a)</td>
<td>4</td>
<td>0</td>
<td>S</td>
<td>M</td>
<td>56/9y</td>
<td>210</td>
<td>broad/narrow.</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Gulordava and Baroni (2011)</td>
<td>0</td>
<td>0</td>
<td>S</td>
<td>M</td>
<td>40y</td>
<td>2</td>
<td>change</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Tang, Qu, and Chen (2013)</td>
<td>33</td>
<td>12</td>
<td>S</td>
<td>M</td>
<td>59</td>
<td>110</td>
<td>B/N/novel/change</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Kim et al. (2014)</td>
<td>0</td>
<td>0</td>
<td>S/P</td>
<td>M</td>
<td>110</td>
<td>21/13/24</td>
<td>change</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Kulkarni et al. (2015)</td>
<td>20</td>
<td>0</td>
<td>S</td>
<td>M/A</td>
<td>105/y/12/y/2y</td>
<td>21/13/24</td>
<td>change</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Hamilton, Leskovec, and Jurafsky (2016b)</td>
<td>28</td>
<td>0</td>
<td>S/P</td>
<td>M</td>
<td>200/190</td>
<td>20/19</td>
<td>change</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Rodda, Senaldi, and Lenci (2016)</td>
<td>0</td>
<td>0</td>
<td>S</td>
<td>M</td>
<td>1200y</td>
<td>2</td>
<td>change</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Eger and Mehler (2016)</td>
<td>0</td>
<td>0</td>
<td>S/P</td>
<td>M</td>
<td>200/190</td>
<td>20/19</td>
<td>change</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Basile et al. (2016)</td>
<td>40</td>
<td>0</td>
<td>S</td>
<td>M</td>
<td>170</td>
<td>17</td>
<td>change</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Azarbayejad et al. (2017)</td>
<td>24</td>
<td>0</td>
<td>S</td>
<td>M</td>
<td>20/11</td>
<td>2/2</td>
<td>change</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Takamura, Nagata, and Kawasaki (2017)</td>
<td>10</td>
<td>0</td>
<td>S/P</td>
<td>M</td>
<td>2/64</td>
<td>21/64</td>
<td>change</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Kahmann, Niekler, and Heyer (2017)</td>
<td>4</td>
<td>0</td>
<td>S</td>
<td>M</td>
<td>≤ 165</td>
<td>48</td>
<td>change</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Bamler and Mandt (2017)</td>
<td>6</td>
<td>0</td>
<td>S/P</td>
<td>M/A</td>
<td>20/230/21</td>
<td>209/230/21</td>
<td>change</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Yao et al. (2018)</td>
<td>4/1888</td>
<td>0</td>
<td>S</td>
<td>M/A</td>
<td>27</td>
<td>27</td>
<td>change</td>
<td>no</td>
<td>no</td>
</tr>
<tr>
<td>Wijaya and Yeniterzi (2011)</td>
<td>4</td>
<td>2</td>
<td>S</td>
<td>M</td>
<td>500/69</td>
<td>500</td>
<td>change novel</td>
<td>yes</td>
<td>yes</td>
</tr>
<tr>
<td>Lau et al. (2012)</td>
<td>5</td>
<td>5</td>
<td>S</td>
<td>M</td>
<td>43 y</td>
<td>2</td>
<td>novel</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Cook et al. (2013)</td>
<td>0</td>
<td>0</td>
<td>S</td>
<td>M</td>
<td>14</td>
<td>2</td>
<td>novel</td>
<td>yes</td>
<td>no</td>
</tr>
<tr>
<td>Cook et al. (2014)</td>
<td>7/13</td>
<td>50/164</td>
<td>S</td>
<td>M</td>
<td>43y/17y</td>
<td>2/2</td>
<td>novel</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Mitra et al. (2015)</td>
<td>0</td>
<td>0</td>
<td>S</td>
<td>M/A</td>
<td>498/2</td>
<td>8/2</td>
<td>split/join/novel</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Fierrez and Lapata (2016)</td>
<td>4</td>
<td>0</td>
<td>S</td>
<td>M/A</td>
<td>311</td>
<td>16</td>
<td>change/novel</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Tang, Qu, and Chen (2016)</td>
<td>197</td>
<td>0</td>
<td>S</td>
<td>M</td>
<td>59</td>
<td>59</td>
<td>B/N/novel/change</td>
<td>no</td>
<td>yes</td>
</tr>
<tr>
<td>Tahmasebi and Risse (2017a)</td>
<td>35</td>
<td>25</td>
<td>S</td>
<td>M</td>
<td>222y</td>
<td>221</td>
<td>novel,B/N/stable</td>
<td>yes</td>
<td>yes</td>
</tr>
</tbody>
</table>


Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
<table>
<thead>
<tr>
<th>Table 3</th>
<th>Datasets used for diachronic conceptual change detection. Non-English</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sagi, Kaufmann, and Clark (2009a)</td>
</tr>
<tr>
<td></td>
<td>Gulordava and Baroni (2011)</td>
</tr>
<tr>
<td></td>
<td>Wijaya and Yeniterzi (2011)</td>
</tr>
<tr>
<td></td>
<td>Lau et al. (2012)</td>
</tr>
<tr>
<td></td>
<td>Cook et al. (2013)</td>
</tr>
<tr>
<td></td>
<td>Cook et al. (2014)</td>
</tr>
<tr>
<td></td>
<td>Mihalcea and Nastase (2012)</td>
</tr>
<tr>
<td></td>
<td>American National Corpus</td>
</tr>
<tr>
<td></td>
<td>Basile et al. (2016)</td>
</tr>
<tr>
<td></td>
<td>Tang, Qu, and Chen (2013, 2016)</td>
</tr>
<tr>
<td></td>
<td>Kim et al. (2014)</td>
</tr>
<tr>
<td></td>
<td>Kulkarni et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>Mitra et al. (2015)</td>
</tr>
<tr>
<td></td>
<td>Hamilton, Leskovec, and Jurafsky (2016b)</td>
</tr>
<tr>
<td></td>
<td>Eger and Mehler (2016)</td>
</tr>
<tr>
<td></td>
<td>Azarbayejani et al. (2017)</td>
</tr>
<tr>
<td></td>
<td>Rodda, Senaldi, and Lenci (2016)</td>
</tr>
<tr>
<td></td>
<td>Freymann and Lapata (2016)</td>
</tr>
<tr>
<td></td>
<td>Takamura, Nagata, and Kawasaki (2017)</td>
</tr>
<tr>
<td></td>
<td>Kahmann, Niekler, and Heyer (2017)</td>
</tr>
<tr>
<td></td>
<td>Tahmasebi and Risse (2017a)</td>
</tr>
<tr>
<td></td>
<td>Bamler and Mandt (2017)</td>
</tr>
<tr>
<td></td>
<td>Yao et al. (2018)</td>
</tr>
<tr>
<td></td>
<td>Rudolph and Blei (2018)</td>
</tr>
<tr>
<td></td>
<td>Helsinki corpus</td>
</tr>
<tr>
<td></td>
<td>Google Ngram</td>
</tr>
<tr>
<td></td>
<td>Google Ngram</td>
</tr>
<tr>
<td></td>
<td>British National Corpus (BNC), ukWaC</td>
</tr>
<tr>
<td></td>
<td>Gigawords corpus</td>
</tr>
<tr>
<td></td>
<td>BNC, ukWaC, Sibol/Port</td>
</tr>
<tr>
<td></td>
<td>Google books</td>
</tr>
<tr>
<td></td>
<td>Google Ngram (Italian)</td>
</tr>
<tr>
<td></td>
<td>Chinese People’s Daily</td>
</tr>
<tr>
<td></td>
<td>Google Ngram</td>
</tr>
<tr>
<td></td>
<td>Google Ngram, Twitter, Amazon movie reviews</td>
</tr>
<tr>
<td></td>
<td>Google Ngram, Twitter</td>
</tr>
<tr>
<td></td>
<td>COHA, Google Ngram</td>
</tr>
<tr>
<td></td>
<td>COHA, Süddeutsche Zeitung, PL76</td>
</tr>
<tr>
<td></td>
<td>New York Times Annotated Corpus, Hansard</td>
</tr>
<tr>
<td></td>
<td>Thesaurus Linguae Graecae</td>
</tr>
<tr>
<td></td>
<td>DATE corpus</td>
</tr>
<tr>
<td></td>
<td>Wikipedia (English and Japanese)</td>
</tr>
<tr>
<td></td>
<td>Guardian (non-public)</td>
</tr>
<tr>
<td></td>
<td>Times Archive, New York Times Annotated Corpus</td>
</tr>
<tr>
<td></td>
<td>Google Ngram, State of the Union addresses, Twitter</td>
</tr>
<tr>
<td></td>
<td>New York Times (non-public)</td>
</tr>
<tr>
<td></td>
<td>ACM abstracts, ML papers ArXiv, U.S. Senate speech</td>
</tr>
</tbody>
</table>

https://languagechange.org/publication/2018-surveypaper/
Towards computational lexical semantic change detection

VR funded
6 million sek (+ cofunding Språkbanken ~700k sek)
2019 – 2022
4 year project: https://languagechange.org/

Overall goal is to bridge the gap between the four of us and all that can benefit from the results.
Main goals

**Wp1: Swedish word sense induction**
- Using sense-differentiated dynamic embeddings

**Wp2: Semantic change**
- On the basis of Wp1

**Wp3: Lexical replacements**
- On the basis of Wp1
- Or using other textual clues

**Wp4: Applications**
- Applied sociology, historical linguistics, history of concepts, ... 

**WP*: Evaluation**
- Integrated in all work packages

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Activities

- News-list (news@languagechange.org)
- Introductory videos to LS change
- Workshops (next at ACL2019)
- Collaboration with other researchers: historians, sociologist, hist. linguists
- SemEval2020 task

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Project timeline

SC = Semantic Change
LR = Lexical Replacement
DH = Digital Humanities
SS = Social Science

Feb – Workshop Helsinki
August – ACL workshop
SemEval annotation

SLTC workshop?
DHSS conf?
SemEval workshop

*ACL workshop?
Final project event

Word sense induction
Evaluation of SC
Hypotheses from historical linguistics

Manual study of SC
Annotation for SC
Lexical Replacement

Manual LR-study
Collaboration with DH and SS

Semantic fields
LR-SC interchange

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Vision

Given a word in a document at time $t$

1. Mark words that are likely to have a changed meaning

2. Find all changes to the word

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Conclusions

Complexity in
• Multiple senses
• Many time points

Not all data are big data!

Evaluation
→ Common datasets and methods!
→ What is the result valid for?

Nina Tahmasebi, On Lexical Semantic Change and Evaluation, Stuttgart, June 2019
Thank you for listening!

Nina.tahmasebi@gu.se
nina@tahmasebi.se