Paradigmatic vs. Syntagmatic

...

The refugee is a person.
The refugee fled her home.
The migrant cannot return safely.

A migrant is a person.
The displaced person fled his home.
The displaced person cannot return.

...

...

The refugee is a person.
The refugee fled her home.
The migrant cannot return safely.

A migrant is a person.
The displaced person fled his home.
The displaced person cannot return.

...

Evaluating Language Change Detection

Detecting Different Forms of Semantic Shift in Word Embeddings via Paradigmatic and Syntagmatic Association Changes

Anna Wegmann, Florian Lemmerich, Markus Strohmaier
RWTH Aachen University & Utrecht University

November 2020, **International Semantic Web Conference 2020**

Word Embeddings change

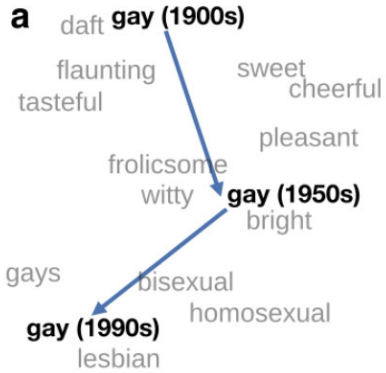


Figure from [W.L. Hamilton et al., Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change, 2016]

Word Embeddings change

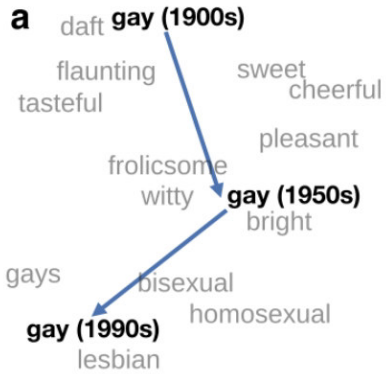


Figure from [W.L. Hamilton et al., Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change, 2016]

Semantic Shift Problem:

Given word w and texts $T_1,...,T_k$ in time-sensitive order

$\Rightarrow$ (How) did w shift "in meaning" over time?

Semantic Shift Problem

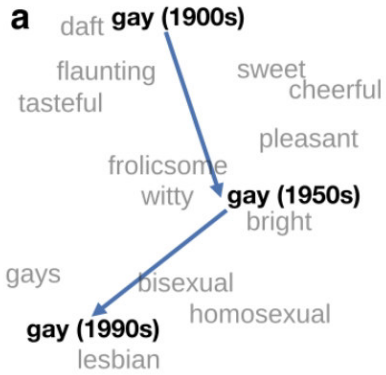


Figure from [W.L. Hamilton et al., Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change, 2016]

Why is this interesting?

- linguistic/societal analysis

Semantic Shift Problem

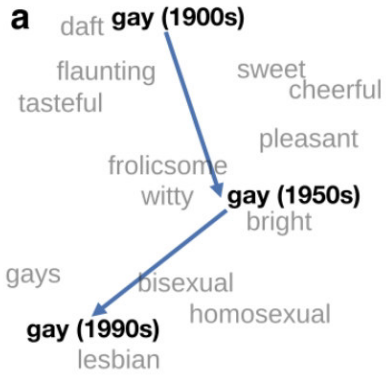
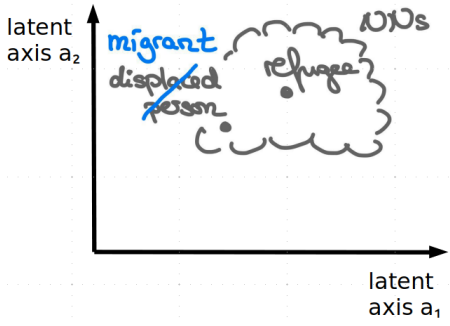


Figure from [W.L. Hamilton et al., Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change, 2016]

Why is this interesting?

- linguistic/societal analysis



Semantic Shift Problem

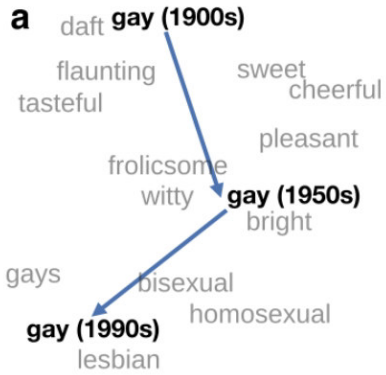


Figure from [W.L. Hamilton et al., Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change, 2016]

Why is this interesting?

- linguistic/societal analysis
- practical algorithmic questions: E.g., when should we update embeddings?
  $\rightarrow$ e.g., RDF2Vec[1]

[1] [Cochez et al., Global RDF vector space embeddings, 2017]

Related Work: Paradig. & Syntagmatic Associations

1. Only types of relations
between words are paradigmatic
& syntagmatic[1]

[1], e.g., in [F. de Saussure, Cours de linguistique generale, 1916];

Related Work: Paradig. & Syntagmatic Associations

1. Only types of relations between words are paradigmatic & syntagmatic[1]

2. Paradigmatic & syntagmatic relations capture different semantic properties[2]

[1], e.g., in [F. de Saussure, Cours de linguistique generale, 1916]; [2] in [M. Sahlgren, The word-space model: using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces, 2006] and [Sun et al., Learning word representations by jointly modeling syntagmatic and paradigmatic relations, 2015]

Related Work: Measures of Semantic Shift

1. Only types of relations between words are paradigmatic & syntagmatic[1]

2. Paradigmatic & syntagmatic relations capture different semantic properties[2]

3. Several evaluation approaches for the same concept of semantic shift[3]

[1], e.g., in [F. de Saussure, Cours de linguistique generale, 1916]; [2] in [M. Sahlgren, The word-space model: using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces, 2006] and [Sun et al., Learning word representations by jointly modeling syntagmatic and paradigmatic relations, 2015]

[3] in [P. Shoemark et al., Room to Glo: A Systematic Comparison of Semantic Change Detection Approaches with Word Embeddings, 2019],[V. Kulkarni et al., Statistically Significant Detection of Linguistic Change, 2015], [A. Rosenfeld et al., Deep Neural Models of Semantic Shift, 2018]

Related Work: Measures of Semantic Shift

1. Only types of relations between words are paradigmatic & syntagmatic[1]

2. Paradigmatic & syntagmatic relations capture different semantic properties[2]

3. Several evaluation approaches for the same concept of semantic shift[3]

4. Different measures for semantic shift are different[4]

[1], e.g., in [F. de Saussure, Cours de linguistique generale, 1916]; [2] in [M. Sahlgren, The word-space model: using distributional analysis to represent syntagmatic and paradigmatic relations between words in high-dimensional vector spaces, 2006] and [Sun et al., Learning word representations by jointly modeling syntagmatic and paradigmatic relations, 2015]

[3] in [P. Shoemark et al., Room to Glo: A Systematic Comparison of Semantic Change Detection Approaches with Word Embeddings, 2019],[V. Kulkarni et al., Statistically Significant Detection of Linguistic Change, 2015], [A. Rosenfeld et al., Deep Neural Models of Semantic Shift, 2018] ;[4], e.g., in [Hamilton et al., Cultural shift or linguistic drift? Comparing two computational measures of semantic change, 2016], [Hamilton et al., Diachronic word embeddings reveal statistical laws of semantic change, 2016]

Research Question

How can we evaluate the sensitivity of measures to paradigmatic & syntagmatic shift?

Experimental Setup

Quick Concept: Word Embeddings

Input

...

The refugee is a person.
The refugee fled her home.
The migrant cannot return safely.

A migrant is a person.
The displaced person fled his home.
The displaced person cannot return.

...

Text Corpus

Quick Concept: Word Embeddings

Input

...

The refugee is a person.
The refugee fled her home.
The migrant cannot return safely.

A migrant is a person.
The displaced person fled his home.
The displaced person cannot return.

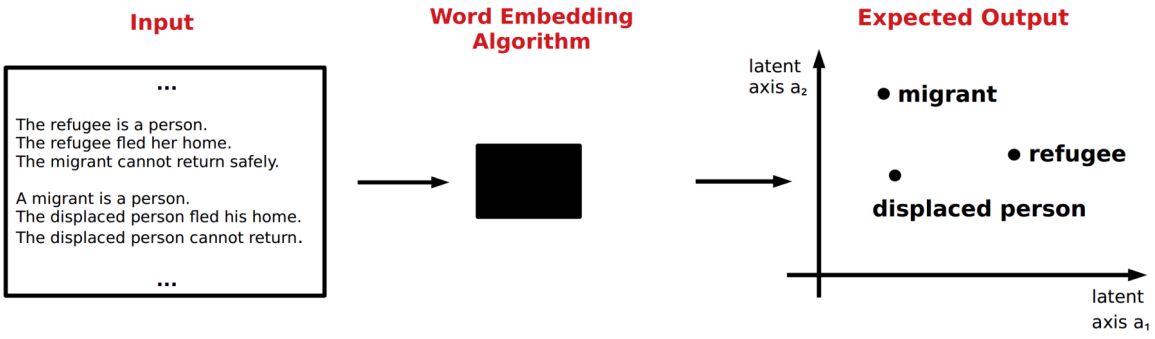
...

Text Corpus

Word Embedding Algorithm

E.g.: Word2Vec

Quick Concept: Word Embeddings



Datasets

1) **Amazon** reviews: 2005 - 2014 with $\approx$ six billion words

2) **Reddit**: 2012 - 2018 with $\approx$ 170 billion words.

3) **Wikipedia**: 2014 - 2018 with $\approx$ 13 billion words

made available by 1) [J. McAuley et al., Image-based Recommendations on Styles and Substitutes, 2015] on `jmcauley.ucsd.edu/data/amazon/`, 2) J. Baumgartner on `https://files.pushshift.io/reddit/` and 3) Wikimedia on archive.org

Synthetic Attacks

NNs - nearest neighbors or closest words to "refugee"

Corpus Change

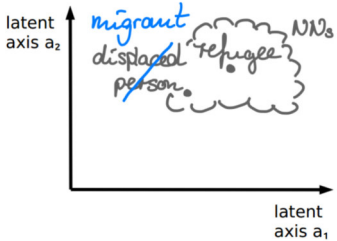
The refugee is a person.
The refugee fled her home.
The refugee cannot return safely.

migrant

A displaced person is a person.
The displaced person fled his home.
The displaced person cannot return.

⇒

Expected Embedding Change



Paradigmatic Attack

similar to [Kulkarni et al., **Statistically significant detection of linguistic change**, 2015]

Synthetic Attacks

NNs - nearest neighbors or closest words to "refugee"

Corpus Change

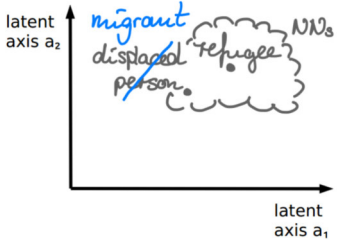
The refugee is a person.
The refugee fled her home.
The refugee cannot return safely.

migrant

A displaced person is a person.
The displaced person fled his home.
The displaced person cannot return.

⇒

Expected Embedding Change



Paradigmatic Attack

Corpus Change

Syrian

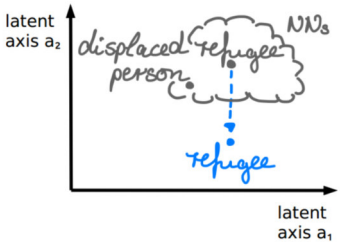
The refugee is a person.
The refugee fled her home.
The refugee cannot return safely.

the war to Syria

A displaced person is a person.
The displaced person fled his home.
The displaced person cannot return.

⇒

Expected Embedding Change



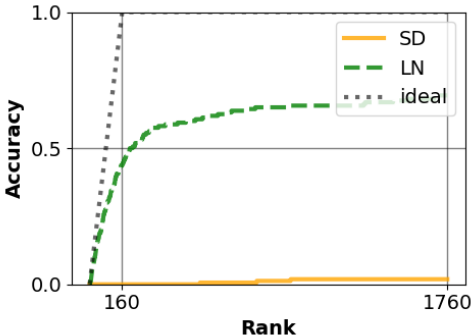
Paradigmatic & Syntagmatic Attack

Experimental Results

Results

Measures of Semantic Shift:
- Local Neighborhood (LN)[1]
- Global Semantic Displacement (SD)[2]

Paradigmatic Attack

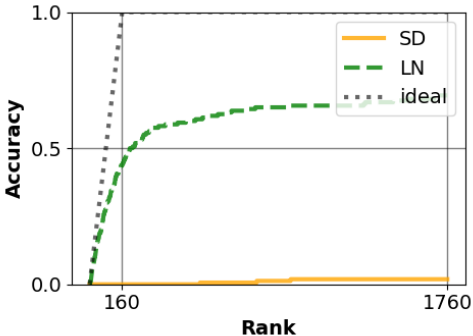


[1] [Hamilton et al., Cultural shift or linguistic drift? Comparing two computational measures of semantic change, 2016];
[2] [Hamilton et al., Diachronic word embeddings reveal statistical laws of semantic change, 2016]

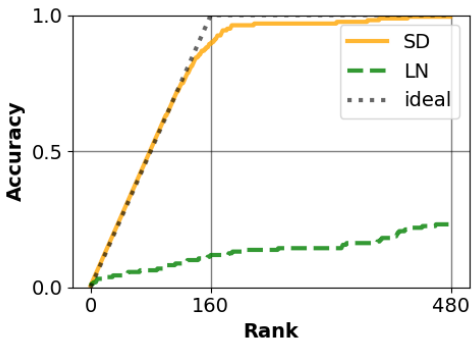
Results

Measures of Semantic Shift:
- Local Neighborhood (LN)[1]
- Global Semantic Displacement (SD)[2]

Paradigmatic Attack



Para. & Syntagmatic

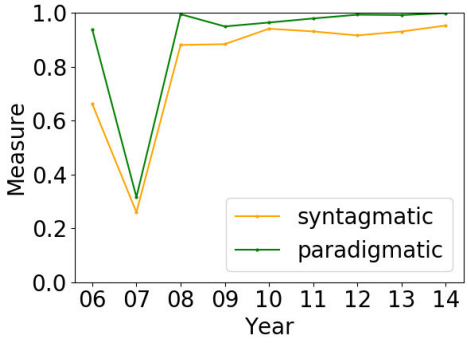


[1] [Hamilton et al., Cultural shift or linguistic drift? Comparing two computational measures of semantic change, 2016];
[2] [Hamilton et al., Diachronic word embeddings reveal statistical laws of semantic change, 2016]

Empirical Results

with the best paradigmatic
and syntagmatic measure

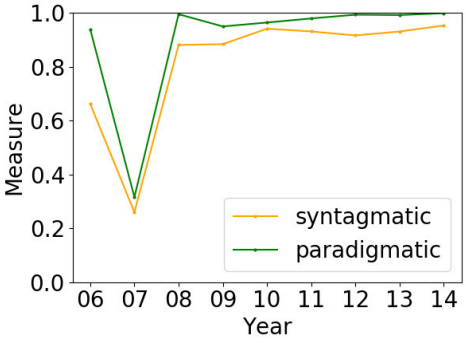
Synchronous Paradigmatic and Syntagmatic Shift



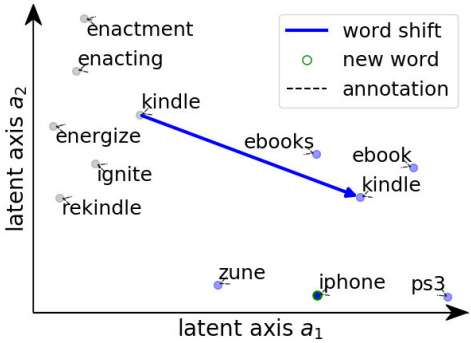
"kindle" on Amazon

"kindle" - Amazon ebook reader introduced in 2007

Synchronous Paradigmatic and Syntagmatic Shift



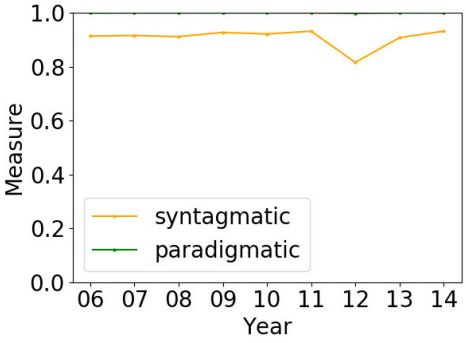
"kindle" on Amazon



Embedding Shift from 2006 to 2007

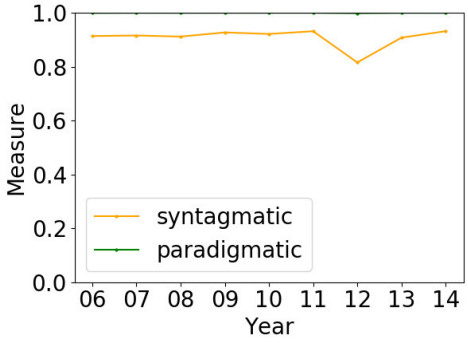
"kindle" - Amazon ebook reader introduced in 2007

Syntagmatic without Paradigmatic Shift

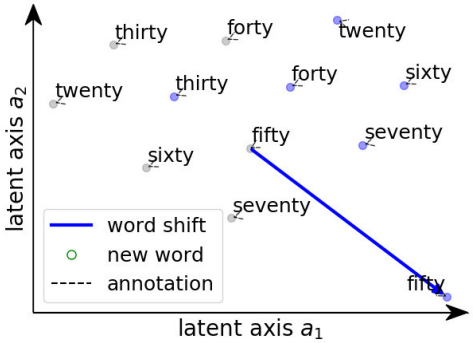


"fifty" on Amazon

Syntagmatic without Paradigmatic Shift



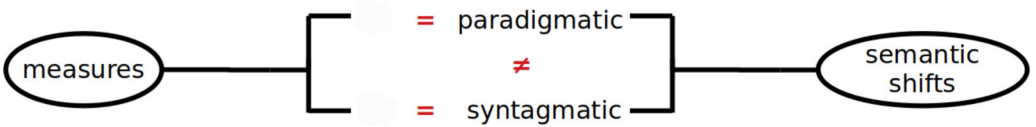
"fifty" on Amazon



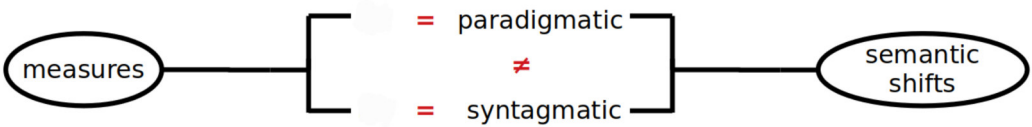
Embedding Shift from 2011 to 2012

In Conclusion

Results and Contributions



Results and Contributions



$\Rightarrow$

i. operationalization of paradigmatic and syntagmatic shift

ii. more nuanced understanding of semantic shift

Limitations

- semantic shift $\neq$ paradig.-/syntagmatic shift?

Limitations

- semantic shift $\neq$ paradig.-/syntagmatic shift?
- measure shift $\Rightarrow$ paradigmatic/syntagmatic shift?

Limitations

- semantic shift $\neq$ paradig.-/syntagmatic shift?
- measure shift $\Rightarrow$ paradigmatic/syntagmatic shift?

and Future Work

- thresholding for (RDF) embeddings

Limitations and Future Work

- semantic shift $\neq$ paradig.-/syntagmatic shift?
- measure shift $\Rightarrow$ paradigmatic/syntagmatic shift?

- thresholding for (RDF) embeddings
- inferring the reason for semantic shifts

See you at the *virtual* ISWC 2020 Q&A session

or online under `https://annawegmann.github.io/`

Paper link: https://annawegmann.github.io/pdf/Detecting-Different-Forms-of-Semantic-Shift.pdf