Temporal Analysis of Language through Neural Language Models

Yoon Kim* Yi-I Chiu* Kentaro Hanaki* Darshan Hegde* Slav Petrov◇

*New York University ◇Google Inc.

Summary

1. Detect semantic change of English words across time through a chronologically trained neural language model.
2. Model is trained on the Google Ngrams corpus from 1850–2009.
3. Words such as *gay* and *cell* are identified as having changed during that time period.
4. Model further identifies the specific periods during which such words underwent change.
5. For some words, the identified periods of change mirror real historical events.

Introduction

- Language changes across time:
  - Existing words adopt new senses (*gay*), or lose existing senses (*guy*)
  - New words are created (*internet*)
  - Some words die out (*doth*)
- Given diachronic corpora, can we automatically detect *which* words changed meaning in a given time period?
- If so, can we also identify *when* such words underwent change?

Method

- Obtain word vectors specific to each year using a Skipgram neural language model (Figure 1).
- Model is chronologically trained so that the obtained word vectors are in the same vector space.
- Compare cosine distance of same words from different years.
- Words that changed will have moved in the vector space and thus will have large cosine distance.
- We can also identify when words changed by looking at the time period during which words moved 'rapidly' in the vector space.

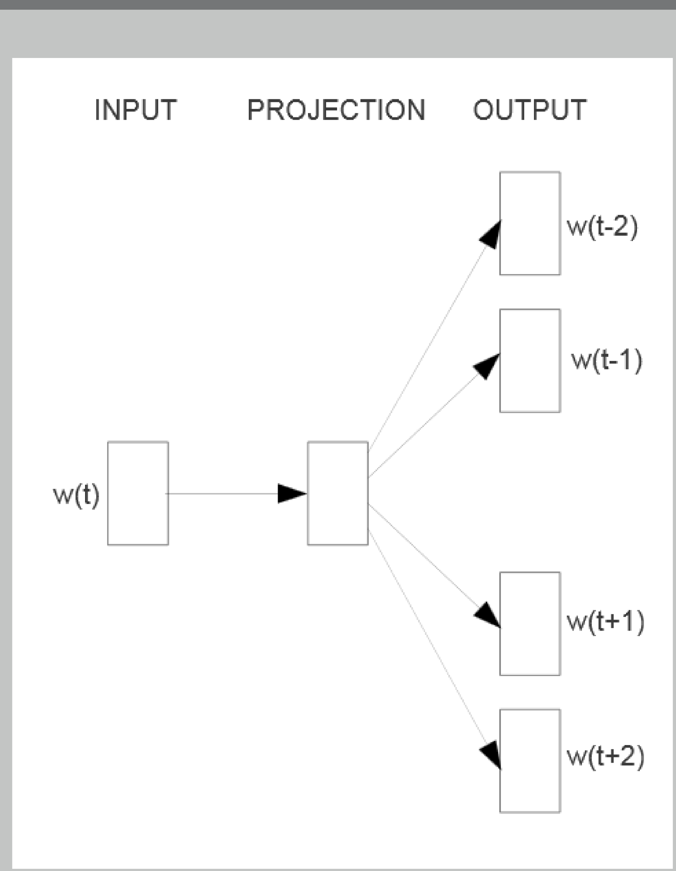

Figure 1: Skipgram architecture (Mikolov et al., 2013)

Data and Training

- For training data, we choose 10 million words each year from 1850 to 2009 from the Google 5-grams English fiction corpus.
- **Training procedure**
  - For each year in the corpus:
    - Train the Skipgram model (via SGD) on the yearly corpus over multiple epochs until convergence.
    - Initialize next year's word vectors with previous year's word vectors (if there are new words, initialize them randomly).
- Training takes approximately 4 days on a single machine.

Results: Most/Least Changed Words from 1900–2009

- After training, we have word vectors (of dimension 200) specific to each year.
- We can compare cosine distance of same words from different years (e.g. *cell:1900* vs *cell:2009*) to see how they have changed.
- We can identify which words have changed the most (or least) during a given time period by ranking the cosine distance of all words against their starting points (Table 1).

Most Changed		Least Changed	
Word	Similarity	Word	Similarity
checked	0.3831	*by*	0.9331
check	0.4073	*than*	0.9327
gay	0.4079	*for*	0.9313
actually	0.4086	*more*	0.9274
supposed	0.4232	*other*	0.9272
guess	0.4233	*an*	0.9268
cell	0.4413	*own*	0.9259
headed	0.4453	*with*	0.9257
ass	0.4549	*down*	0.9252
mail	0.4573	*very*	0.9239

Table 1: Top 10 most/least changed words between 1900–2009.

Results: Neighboring Words

- Some words in the *most changed* category are obvious (e.g. *gay*, *cell*).
- Some are not (e.g. *headed*, *actually*).
- Most of the *least changed* words are function words.
- We can look at neighboring words from different years to get better context as to how these words changed (Table 2).
- For example, it seems like *checked* changed meaning from "to hold in restraint" to "to verify by consulting an authority" or "to inspect so as to determine accuracy".
- Example sentences:
  - 1900: "She was about to say something further, but she *checked* herself."
  - 2009: "I *checked* out the house before I let them go inside."

Word	Neighboring Words in 1900	Neighboring Words in 2009
gay	*cheerful* *pleasant* *brilliant*	*lesbian* *bisexual* *lesbians*
cell	*closet* *dungeon* *tent*	*phone* *cordless* *cellular*
checked	*checking* *recollecting* *straightened*	*checking* *consulted* *check*
headed	*haired* *faced* *skinned*	*heading* *sprinted* *marched*
actually	*evidently* *accidentally* *already*	*really* *obviously* *nonetheless*

Table 2: Top 3 neighboring words (based on cosine similarity) specific to each time period for the words identified as having changed.

Results: Identifying Periods of Change

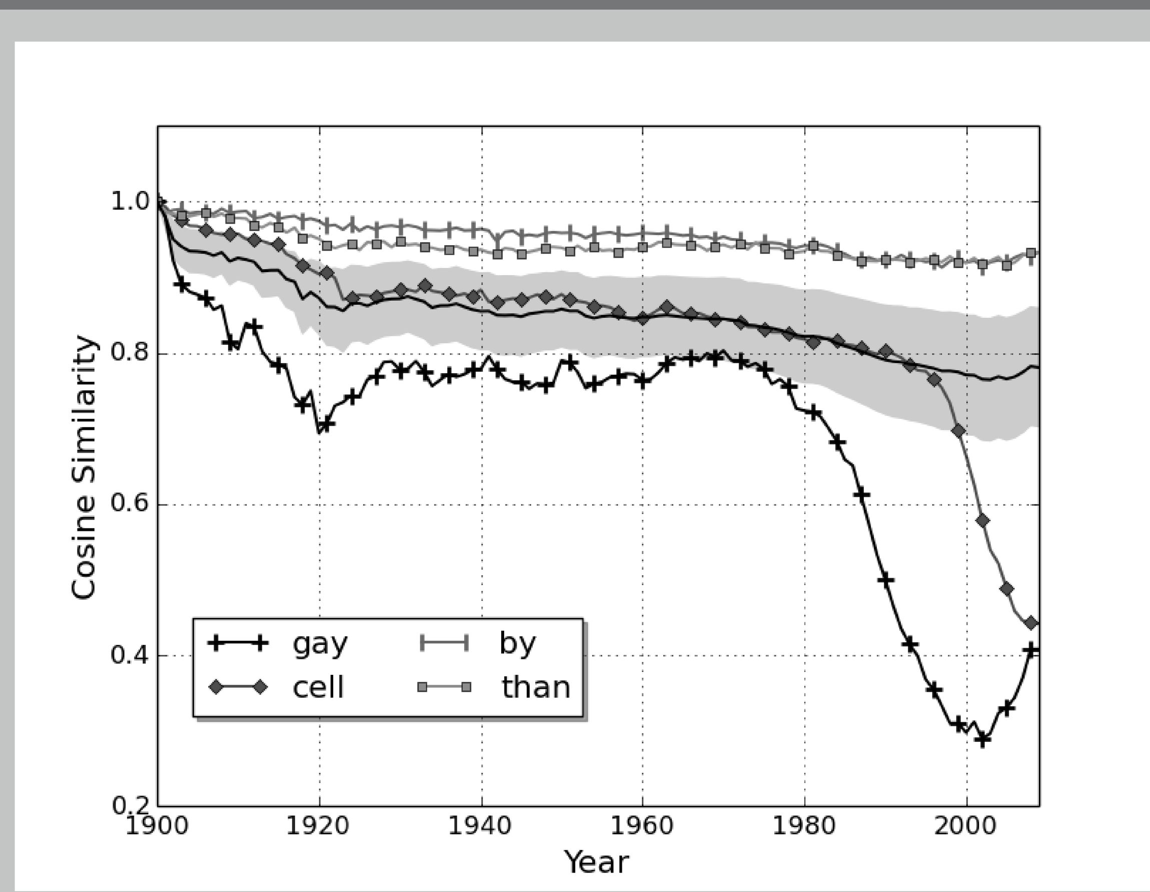

Figure 2: Cosine similarity of changed (*cell*, *gay*) and unchanged (*by*, *than*) words against their starting points in 1900. Middle line is the average cosine similarity of all words. Shaded region corresponds to one standard deviation of errors.

- To look at *when* such words changed, we can plot the cosine similarity of a word against itself from a reference year (in our case, 1900). Words undergoing change will have 'moved' rapidly in the vector space.
- For example, we can see that *gay* changed meaning most rapidly during 1970–2000, while *by* remained steady from 1900–2009.
- For some words, the identified period of change mirrors real historical events (as in Wijaya and Yeniterzi, 2011):
  - Period of change for *gay* mirrors the gay movement in the 1970s.
  - Period of change for *cell* agrees with the introduction of the cell phone to the general public in 1984.

Conclusions and Future Work

- We have introduced a method for analyzing change in language across time through a chronologically trained neural language model.
- The model simultaneously identifies the words that changed as well as the periods during which they changed.
- An interesting direction of work could involve inferring the type of change (pejoration vs. amelioration, broadening vs. narrowing, etc.) from movements in the vector space.

References

T. Mikolov, K. Chen, G. Corrado, J.Dean. 2013. Efficient Estimation of Word Representations in Vector Space. *arXiv:1301.3781.*

D.T. Wijaya, R. Yeniterzi. 2011. Understanding semantic change of words over centuries. *Proceedings of the 2011 Workshop on Detecting and Exploiting Cultural Diversity on the Social Web.*