

THE UNIVERSITY OF IOWA



REPORT FOR QUALIFYING EXAM

---

# **Shifting Sands: Understanding Temporal Context Change**

---

*Author:*

Manisha KEIM

manisha-keim@uiowa.edu

*Advisor:*

Rishab NITHYANAND

rishab-nithyanand@uiowa.edu

Department of Computer Science

August 10, 2024

## 1 Introduction

Language is a dynamic, living entity that continually adapts to reflect the changing world around us. With the evolving language, words take on new meanings through technological advancements, cultural shifts, and political movements.

Particularly in political contexts, certain words or phrases can experience rapid shifts in meaning or become ambiguous. These are known as ‘data voids’ - situations where the meaning of a term has become unstable or contested, often due to deliberate efforts by political actors to redefine it.

Unlike more general language evolution, where new terms emerge to describe novel concepts, data voids in politics involve the strategic manipulation of existing vocabulary. Political actors may repurpose words or introduce new phrases to frame debates in their favor, creating confusion and a lack of shared understanding.

This phenomenon highlights how language can be weaponized as a political tool. As terms take on drastically different meanings for different groups, it can become challenging to maintain clear, authoritative information about their usage. The resulting ‘data voids’ can then be exploited to spread misinformation or propaganda.

My research goal is to find data voids in political context automatically and systematically study how word shifts their meaning over time. To achieve this, we have selected following research papers as a foundation for our research.

- Statistically Significant Detection of Linguistic Change. ACM WWW 2015 [8]
- Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. ACL 2016 [5]

The concept of context drift, as presented in these papers, refers to the change in the contextual usage of words over time. This can be particularly useful for detecting and analyzing data voids in political contexts.

Finding data voids is significantly important because political events or campaigns can cause sudden changes in how certain terms are used or interpreted. Political actors may deliberately redefine terms or introduce new phrases to frame debates in their favor. Words or phrases can take on drastically different meanings in different political contexts or for different groups. Some terms may develop coded political meanings that aren’t immediately apparent to all audiences. The lack of clear, authoritative information about a term’s new political meaning can be exploited to spread misinformation or propaganda.

These elements can rapidly change the basic meaning (denotation) and the feelings or ideas they evoke (connotation) of words, making them challenging to track and analyze.

**Example.** The term *crisis actor* has undergone a significant shift in meaning, exemplifying the phenomenon of context drift we’ve been discussing.

Originally this term was used in the entertainment and training industries, which refers to individuals hired to portray victims or other roles in simulated emergency scenarios. These simulations were used for training first responders, military personnel, or for other educational purposes.

However, in recent years, the term has been full of misinformation which refers to genuine victims of violent events (mass shooting) in an attempt to prove that an event was staged or did not happen. This term got prominent during Sandy Hook Elementary School shooting (2012), Los Angeles International Airport (LAX) shooting (2013), Pulse Nightclub shooting in Orlando (2016), Las Vegas shooting (2017).

This example clearly demonstrates how a term with a specific, legitimate meaning can be repurposed to spread misinformation and conspiracy theories. The shift creates a potential data void where searches for *crisis actor* might lead to a mix of factual information about emergency training and harmful conspiracy theories, making it difficult for users to distinguish accurate information.

## 2 Related Work

**Tracking how word meanings change over time.** The evolution of word meanings over time has been a subject of significant interest. Researchers employed a diverse array of methodologies, including *word embeddings* [8], *neural language models* [7], and *diachronic analysis* [5, 9], to track and examine semantic shifts across historical periods. Figure 1 presents a chronological overview of models and events in this domain. Much of this research utilized the *Google Books Ngram corpus* [3–5, 7–9, 11], a vast collection spanning from 1900 to 2009, encompassing over 500 billion books in seven languages. This dataset provides n-grams with corresponding yearly occurrences and frequencies.

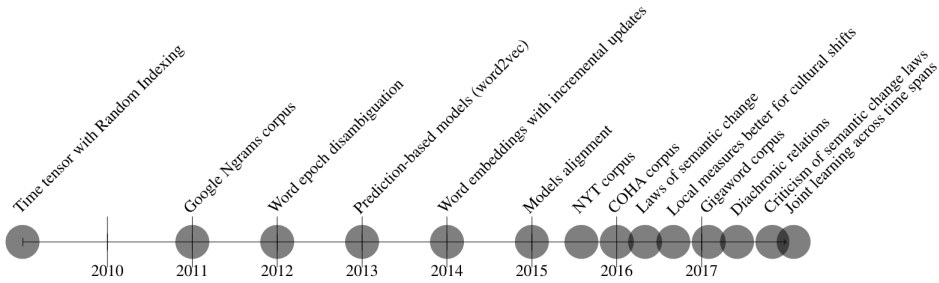


Fig. 1. Timeline of models used to study semantic shifts (taken from [9])

**Words can change meaning over time through several linguistic processes.** Words have the ability to undergo shifts in meaning over the course of time as a result of various linguistic mechanisms, including but not limited to:

- *semantic drift* [3, 7, 8]
  - Refers to the evolution of word meanings over time.
  - Example: The word ‘gay’ originally meant ‘happy’ or ‘carefree’, but now it predominantly means ‘homosexual’.
- *syntactic alterations* [1, 4, 8]
  - Syntax focuses on the structure of language. This type of change pertains to modifications in the arrangement of words in sentences.
  - Example: The word ‘apple’, which transitioned from being used as a ‘common noun’ (e.g., a fruit) to a ‘proper noun’ (referring to the Apple company) after the company’s rise in the 1980s.
- *broadening* [1, 4, 8]
  - A word’s meaning becomes more general than its earlier meaning, also known as generalization.
  - Example: ‘Holiday’ originally meant a religious festival, but now it can refer to any day of celebration or time off.
- *narrowing* [1, 8]
  - A word’s meaning becomes more specific than its earlier meaning, also known as specialization.
  - Example: ‘Literally’ used to mean ‘figuratively’ or ‘symbolically’. Now, it is used to emphasize the truthfulness of a statement.
- *amelioration*
  - A word takes on a more positive meaning over time.

- Example: The word ‘thrifty’ once meant ‘cheap’ but now suggests responsible use of resources.
- *pejoration* [3, 10]
  - A word takes on a more *negative* meaning over time.
  - Example: The word ‘awful’ meant ‘full of awe’ which transitioned to ‘terrible’ or ‘appalling’.

Hamilton et al. (2016) highlighted that *cultural* or *linguistic* factors could have driven these transformations. Certain studies concentrated on semantic changes, some worked on syntactic changes, [4, 8, 10] and others explored the broader evolution that words underwent [3, 5, 7, 8]. Table 1 outlines various types of linguistic changes and the approaches employed to identify them.

## 2.1 Semantic Drift

**Building on the extensive research on semantic change.** Semantic change refers to any change in the meaning(s) of a word over time or acquiring a new sense [3, 10]. Kutuzov et al. (2018) conducted a survey on semantic shifts, consolidating the existing academic research in this domain. Their work provides a comprehensive overview of the methodologies and findings related to tracking semantic changes over time using computational techniques.

**Investigating semantic shifts through word contexts.** Gulordava and Baroni (2011) detected semantic change by focusing on words used in the 1960s and 1990s. They compared the similarity of the surrounding words (words that co-occur with the target word) in these two time periods.

To assess similarity, the researchers calculated the *Local Mutual Information (LMI)* score between the central word and its surrounding words. A low LMI score between the target word and its surrounding words across the two time periods indicated a potential semantic shift. The study demonstrated that their distributional similarity models were effective in capturing *cultural shifts* in word meaning. For example, they found that the word ‘sleep’ acquired more negative connotations related to sleep disorders when comparing its contexts in the 1960s to those in the 1990s.

**Tracking meaning evolution through neural nets.** Kim et al. (2014) focused on analyzing how word meanings evolved between 1900 and 2009. They developed the first method that employed *prediction-based model* to trace semantic shifts. This involved training a model on data from a specific year  $y_i$  and then using the resulting word vectors as the starting point for training the model on the next year’s data  $y_i + 1$ .

Their method analyzed *global shifts* in a word’s vector semantics. Additionally, by plotting the time series of a word’s distance to its neighboring words in the model’s vector space, they visualized the period during which the semantic shift occurred. They demonstrated this for the word ‘cell’ compared to its early neighbors, ‘closet’ and ‘dungeon,’ and the more recent neighbors, ‘phone’ and ‘cordless.’ For ‘cell,’ the identified period of change (1985-2009), which interestingly coincides with the introduction and widespread adoption of cell phones by the public.

**Pinpointing significant shifts statistically.** Kulkarni et al. (2015) proposes a novel computational approach to identify and quantify the semantic and usage changes in words across various media (new products, movies and books). Building on the concept of distributed representations proposed by Hinton [6], they map words into a continuous vector space where words with similar meanings are positioned close together. The approach hinges on constructing *property time series* for each word. They propose three methods for constructing these time series (§4):

- (1) **Frequency:** This method analyzes changes in a word’s overall frequency of use, assuming a sudden shift in frequency might indicate a semantic shift.

| Linguistic<br>Change                  | Approach                      | TCD         | Example  |  |                                       |
|---------------------------------------|-------------------------------|-------------|----------|--|---------------------------------------|
|                                       |                               |             | Word     | Former Usage                               | New Usage                             |
| FREQUENCY [3, 5, 8]                   |                               |             |          |  |                                       |
| Semantic                              | Log Ratio                     | 1960 & 1990 | disk     | -  | -                                     |
|                                       | Word Frequency                | 1900–2005   | bitch    | female dog                                 | slang                                 |
| DISTRIBUTIONAL [1, 3–5, 7, 8, 10, 11] |                               |             |          |  |                                       |
|                                       | Local Mutual Information      | 1960 & 1990 | sleep    | deep sleep                                 | sleep disorder                        |
|                                       | Continuous Word Embeddings    | 1900–2009   | cell     | closet, dungeon                            | phone, cordless                       |
|                                       | Change Point Algorithm        | 1900–2005   | gay      | cheerful, dapper                           | lesbian, homosexual                   |
|                                       | Clustering (DBSCAN)           | 1900–2000   | mouse    | mice, rat                                  | cursor, pointer                       |
|                                       | Cultural and Linguistic Drift | 1800–2000   | virus    | infected with virus                        | spreading computer virus              |
|                                       | Diachronic Word Embeddings    | 1800–1999   | awful    | full of awe                                | terrible or appalling                 |
|                                       | Contextual Word Embeddings    | 1910 – 2009 | tenure   | short term leases, insecurity of tenure    | tenure of office, employment          |
|                                       | Semantic Similarity Analysis  | 2014–2022   | immunity | politics (‘legislator’, ‘representative’). | health-related (‘prevention’, ‘AIDS’) |
| PART OF SPEECH [4, 8]                 |                               |             |          |  |                                       |
| Syntactic                             | POS Tags                      | 1900–2000   | windows  | doors and windows of a house               | Microsoft Windows                     |
|                                       | Mixed-model Regressions       | 1800–2000   | actually | originally, nominally                      | presumed, believe                     |

Table 1. **Comparative overview of linguistic change detection approaches.** This table provides a comprehensive summary of various approaches and contributions to the study of linguistic changes, categorizing the type of changes (semantic or syntactic), and the methodologies used, including frequency, distributional and part of speech approaches. It highlights each study’s methods, such as word embeddings, statistical analysis, and clustering, along with the time periods of change detection (TCD) and examples of words.

- (2) **Syntactic:** This method examines the distribution of a word’s part-of-speech tags (e.g., noun, verb) across different time periods, aiming to capture changes in how the word functions grammatically.
- (3) **Distributional:** This approach leverages word embeddings which are created for each year, and then alignment is done to represent them in joint embedding space, and it’s utilized to construct distributional time series for a word’s displacement.

Finally, they employed statistically sound *change point detection algorithm* to identify significant moments in these time series, pinpointing the periods where word meaning or usage likely underwent a shift. Their results indicated that computational methods for the detection of semantic shifts can be robustly applied to time spans less than a decade.

**Word semantic modelling of polysemant.** Liao and Cheng (2016) explored the semantic changes of words. Their approach built on the understanding that word meaning is closely tied to its context. When a word's meaning changes, the surrounding words used with it (*context words*) are likely to change as well. They focused on *polysemous* words (words with multiple meanings) and aimed to detect when new meanings emerged.

They used the *skip-gram architecture with negative sampling* [12] to obtain word embeddings. This technique helped in capturing the contextual meaning of words by representing them in a continuous vector space. They employed *DBSCAN* to group word embeddings. *DBSCAN* helped in identifying clusters of similar word contexts and distinguishing them from noise, which could indicate semantic changes. To find similar words, they used a nearest neighbor search method called *Random Project Forest*. This method helped in identifying words that are contextually similar to a given word. Finally, they compared the stability of *similar words* with the stability of their *context words*.

**Distinguishing cultural shifts from linguistic drift.** Hamilton et al. (2016) addressed the challenge of distinguishing between *cultural shifts* and *linguistic drift*, both of which can contribute to semantic change. They proposed two distinct measures based on distributional representations to distinguish between these two types of semantic change:

- (1) **Local neighborhood measure:** This measure focused on the closest neighbors (most similar words) in a word's embedding. A drastic shift in these nearest neighbors suggests a significant change in core meaning, potentially driven by a *cultural shift* (e.g., 'gay' changing from carefree to referring to homosexuality).
- (2) **Global measure:** This measure considered the overall distribution of a word's surrounding words in a larger context window. Gradual changes in this broader distribution are more likely to reflect *linguistic drift*, the natural evolution of language due to regular processes (e.g., 'promise' expanding from a declaration to also suggesting a likelihood).

Prior research often treated semantic change as a single phenomenon. Hamilton et al. offered a novel approach by distinguishing between cultural shifts and linguistic drift using their two measures.

**Diachronic analysis on historical data:** The primary goal of Hamilton et al. (2016) is to track semantic changes and understand how the meanings of words shift in different historical contexts. They created word embeddings for different time periods using both the PPMI matrix with SVD and the SGNS model. These embeddings were generated from historical text corpora to capture the contextual usage of words in each period. After creating the embeddings, they aligned the embeddings. Once the embeddings are aligned, they calculate the *semantic displacement* of a word. This essentially measured how much a word's vector representation has moved in the embedding space between two time periods. The study demonstrates that their method effectively identified semantic change in words and uncovered two statistical 'laws' of semantic change:

- (1) **Law of Conformity**, which suggests that the rate of semantic change is inversely proportional to a word's frequency. High-frequency words tend to change meaning more slowly.
- (2) **Law of Innovation**, on the other hand, proposes that words with multiple meanings are more likely to undergo semantic transformations over time.

They leveraged the rich information within word embeddings to quantify the degree of semantic change (semantic displacement) and identified potential patterns.

**Semantic shifts with contextual embeddings:** Giulianelli et al. (2020) explored the phenomenon of lexical semantic change using an unsupervised approach using *BERT*. They utilized the *BERT* model to generate contextual word embeddings, which captured the meaning of a word based on

its surrounding context. The extracted word usage vectors are then clustered into different *usage types* using the *k-means clustering algorithm*. This helped to identify distinct senses or meanings of the words as they appear in various contexts. They analyzed these clusters for a specific word across different time periods. The study proposed three metrics to quantify semantic change:

- (1) **Entropy Difference (ED)**: Measures the change in uncertainty (entropy) of a word's usage distribution over time.
- (2) **Jensen-Shannon Divergence (JSD)**: Compares the similarity of word usage distributions across time intervals.
- (3) **Average Pairwise Distance (APD)**: Computes the average distance between word usage vectors from different periods, indicating shifts in word meaning.

The qualitative analysis indicated that the approach could capture various linguistic phenomena, including both synchronic (current usage) and diachronic (historical changes) aspects.

**Semantic similarity:** The impact of significant events like the COVID-19 pandemic on language and semantic change has also been a subject of study. Laurino et al. (2023) explores tracking fast semantic changes through a large-scale word association task, aiming to understand how the collective mental lexicon evolves in response to such global events. Their research highlights the dynamic nature of language and how it incorporates new senses. One key finding from their work is that words directly related to the pandemic exhibited a greater difference in semantic similarity between pre-pandemic and pandemic time periods. This suggests that these words underwent a more significant and rapid semantic shift compared to control words not associated with the pandemic. They also employ semantic similarity analysis to quantify the shifts in meaning for pandemic-related words and provides evidence that the COVID-19 pandemic acted as a catalyst for rapid semantic change. For instance, words like 'quarantine', 'mask', and 'social distancing' took on new and prominent meanings in everyday conversation.

## 2.2 Takeaways

Existing research on semantic change detection employs a diverse range of methods. Traditional approaches rely on frequency analysis, such as log ratio computations[3], while more recent studies utilized distributional representations[4, 7, 8], including word embeddings generated by techniques like SGNS, PPMI, and SVD[5]. Clustering algorithm like DBSCAN[11] was applied to detect semantic shifts in polysemous words. By clustering words based on their contextual similarity, they were able to identify when a word's meaning had shifted, as changes in clusters indicated emerging or evolving senses. BERT-based contextual embeddings[1] measured lexical semantic change by capturing how word meanings shift subtly across different contexts. The advent of word embeddings provided a quantitative framework for measuring semantic change, leading to the identification of laws like the *law of conformity* and *law of innovation* as proposed by Hamilton et al. (2016). These methods have successfully identified semantic shifts in corpora spanning from the 19th to the 21st centuries, revealing changes in word meanings and usage. Recent studies, such as Laurino et al. (2023), have begun to explore the rapid evolution of language in response to global events like the COVID-19 pandemic, highlighting the dynamic nature of semantic change.

## 2.3 Syntactic Alterations

**Linguistic change studies have shifted to include syntax alongside semantics.** The field of linguistic change has traditionally focused on *semantics*. However, recent studies have begun to explore the role of *syntax* in language evolution as well. The *syntactic* functionality of a word can evolve by transitioning into a new part-of-speech (POS) category. Nouns, due to their inherent



flexibility in meaning, exhibit a greater tendency to undergo these changes driven by cultural shifts. While verbs are more likely to participate in gradual semantic changes that follow established linguistic patterns. [4, 5]

**Acquiring a new part of speech.** Kulkarni et al. assigned part-of-speech (POS) tags to a large collection of text. They calculated the likelihood of a word appearing in specific grammatical contexts over time. To quantify temporal change, they compared the probability distributions of POS tags for a particular word across different time periods. This essentially measured the divergence between these distributions. An example they cited was the word ‘windows’. Its POS tag shifted from a common noun (referring to doors and windows of a house) to a proper noun (‘Microsoft Windows’). This highlights how a word’s grammatical function can change alongside its meaning.

**Capturing differences between nouns and verbs.** Hamilton et al. (2016) highlight how *cultural changes*, often influenced by new technologies, are closely tied to transformations within *local neighborhoods*, particularly sensitive to shifts in nouns. On the other hand, *linguistic changes* are more associated with *global measures* and are particularly responsive to variations in verbs.

To validate this hypothesis, the authors employed a statistical technique called a *linear mixed model* where word type (noun or verb) is a fixed effect, amount of change measured by each metric (local or global) treated as the dependent variable. By analyzing the model’s results, they could assess whether there’s a significant difference in the way nouns and verbs exhibit change. The evolution of words like ‘actually’, ‘must’, and ‘promise’ demonstrate these changes. For instance, ‘must’ has transitioned from expressing obligation to indicating necessity, showcasing a common pattern seen in modal verbs.

## 2.4 Takeaways

While research in semantic change has advanced, syntactic change remains relatively understudied. Kulkarni et al. (2015) utilized syntactic analysis through part-of-speech (POS) tagging to observe changes in word usage patterns, emphasizing shifts in syntactic roles over time. Hamilton et al. (2016) linked syntactic changes to broader cultural shifts, analyzing the evolution of nouns and verbs. Their study demonstrated how cultural shifts influenced language use, with certain syntactic changes correlating with broader societal trends. However, these studies only scratch the surface of syntactic change, with areas like broadening, narrowing, and amelioration largely unexplored.

## 2.5 Selected Papers

Several studies from our literature review provide valuable insights for addressing data voids in our analysis. The first being the work by Kulkarni et al. which introduces us to time series construction. Their distributional methods focuses on finding subtle semantic shifts to determine the context where a word is used. This concept aligns perfectly with our goal of uncovering shifts in word usage when encountering data voids.

Second paper by Hamilton et al. introduces us to word embeddings alignment. Since we want to compare word vectors from different time periods, vectors should be aligned in same coordinate axes. After aligning the embeddings for individual time periods, we can use the aligned word vectors to compute the semantic displacement that a word has undergone during a certain time-period. Next two section §3, §4 focusses on the selected papers in detail.



### 3 RP1: Statistically Significant Detection of Linguistic Change

#### 3.1 Introduction

Kulkarni et al. (2015) explores the problem of detecting changes in word usage patterns over time. They introduce methods to identify words that have undergone significant shifts in *meaning*, usage *frequency*, and *context*. This is achieved by analyzing historical corpora, which contain vast amounts of text data spanning multiple centuries. The study utilizes diverse datasets, including Twitter posts, Amazon product reviews, and Google Book Ngrams, to construct time series for individual words.

The research conducted by Kulkarni et al. exhibits a compelling alignment with our own research objectives, specifically focusing on the identification of *data voids* characterized by shifting semantic meanings over time. Their exploration into these linguistic shifts has revealed interesting insights, such as the evolution of the word 'gay,' (Figure 2) which transitioned from being cheerful or dapper to signifying homosexual or lesbian aspect around the mid-20th century.

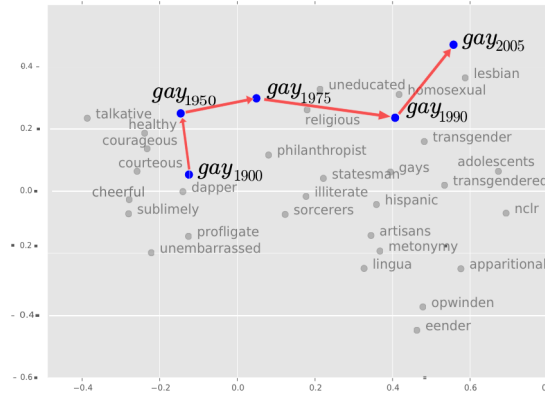


Fig. 2. Example of word 'gay' undergoing change in meaning over time (taken from [8]).

To achieve understanding of this phenomenon, they studied the following research questions.

- **RQ1.** *What methods can quantify the statistical relevance of observed changes in a word's usage across different time periods?*

Kulkarni et al.'s study uses three technical methods to understand how words change over time: *frequency analysis*, *syntactic analysis*, and *distributional analysis*. The frequency method tracks how often words are used, revealing spikes in usage when significant events occur, like how *breaking news* data voids can create a surge in specific keywords. The syntactic method looks at changes in the grammatical roles of words, while the distributional method examines word co-occurrence patterns to see how word meanings shift. Their approach not only uncovers the dynamic nature of language but also highlights "data voids," – areas where terms have *evolved*, become *outdated*, or *fragmented* in meaning.

- **RQ2.** *Given that a word's usage has changed, how can the precise moment or period of this shift be determined?*

Understanding data voids involves not only recognizing that a change is occurring but also analyzing the specifics of how that change unfolds over time. It's crucial to pinpoint not just the fact that a shift has happened, but also the exact moment when it took place. Kulkarni et al. tackle this challenge by implementing a *change point detection algorithm* based on the *Mean Shift* model. This method determines whether a word has experienced a significant shift and, if so, identifies the precise point at which this change occurred.

Next two sections, §3.2 and §3.4 focuses on the two research questions that this paper is studying, the approach they took to study and their results.

### 3.2 What methods can quantify the statistical relevance of observed changes in a word’s usage across different time periods?

**Overview.** The challenge is to identify how the meaning of words changes over time. To address this, they examine a corpus containing data from various domains like books, tweets, and reviews, known as the Temporal Corpus ( $C$ ) which spans over a time period ( $S$ ). This corpus is divided into smaller segments called snapshots  $C_t$ , each of length  $P$ . From these snapshots, a common vocabulary  $V$  is created having words which are common across all snapshots. To analyze changes in these words, a time series  $T(w)$  is constructed for each word  $w \in V$  (Figure 3).

Kulkarni et al. propose several methods to create this time series, aimed at understanding the evolution of words over time.

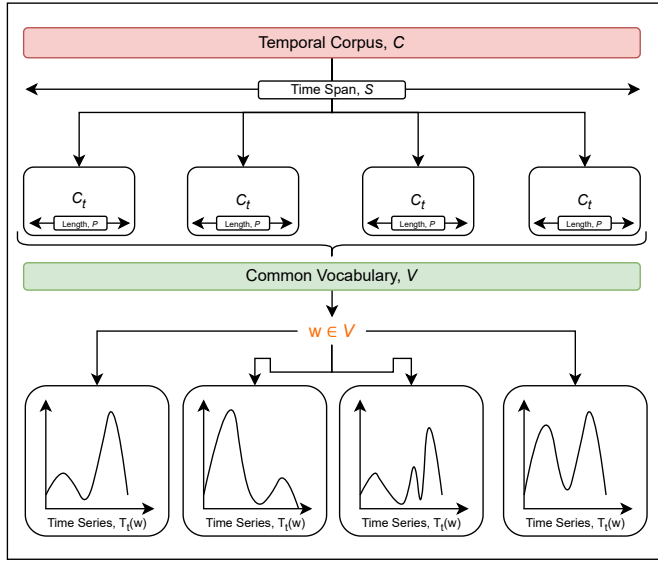


Fig. 3. Overview of time series construction.

**Construction of time series.** In this paper, three methods are explored for statistically modeling the evolution of words over time: frequency analysis, part-of-speech tagging, and word co-occurrence to create time series for each word.

**Frequency method.** The simplest and most direct method for detecting sudden changes in word usage is by analyzing frequency trends. This approach examines how often a word is used over a given period, providing insights into shifts in its popularity or relevance. Changes in frequency can indicate whether a word is gaining new meanings or losing old ones, reflecting broader cultural or societal trends. For example, the word “gay” in Figure 4 shows a noticeable spike in usage during the 1980s, signaling a shift in its meaning or cultural relevance.

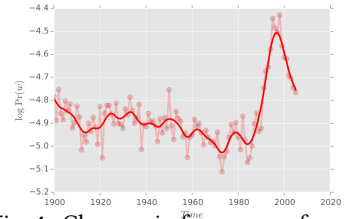


Fig. 4. Changes in frequency of ‘gay’ (taken from [8]).

Tools such as the Google Books Ngram Viewer and Google Trends are used for this purpose, as they provide large datasets and visual representations of word usage across different timeframes. They calculate the change in probability of a word appearing over time.

$$T_t(\mathbf{w}) = \log \frac{(w \in C_t)}{|C_t|} \quad (1)$$

**Syntactic method.** Frequency-based metrics, while simple to compute, are susceptible to errors caused by imbalances in the corpus’s domain and genre distribution. Fluctuations in word usage due to temporal events or the popularity of specific entities can obscure genuine shifts in word meaning. Moreover, a word’s grammatical role can evolve over time, such as acquiring a different part of speech category. To study this, the corpus is annotated with part-of-speech (POS) tags, and the probability distribution of these tags is calculated for each word across different time snapshots. This approach allows researchers to observe how a word’s syntactic role changes over time.

**Distributional method.** For detecting subtle semantic changes, which are not changed either due to frequency or through change in part of speech, this method was developed to understand in which context a word is used in and based on that understand the semantic changes. Distributional methods focusses on creating a semantic space that maps words to continuous vector space, where each word is represented by a vector. Once they created a temporal word embeddings for each word in each time snapshot, then they track the changes of the representations across the embedding space.

The researchers aimed to *learn word embeddings* by training neural language models on a corpus at different time snapshots. They initialized word vector representations randomly and optimized the model parameters using stochastic gradient descent. During training, the objective was to maximize the probability of context words appearing around a target word. After training, word embeddings were normalized by their L2 norm to ensure consistent representation.

The *alignment* process involved aligning word embeddings from different time snapshots into a unified coordinate system to characterize changes between them. Assumptions were made to aid the alignment process, including the equivalence of spaces under linear transformation and the preservation of local structure of most words over time. When the alignment model failed to align a word correctly, it indicated a potential linguistic shift, highlighting the importance of accurate alignment for tracking semantic changes over time. By aligning embedding spaces across various time snapshots into a joint embedding space, the distributional method constructs a *distributional time series* that captures the semantic evolution of words over time.

The paper shows the evolution of the word ‘tape’ over time. Initially, the word ‘tape’ referred to an ‘adhesive tape’ but underwent a semantic shift to also mean a ‘cassette tape’.

### 3.3 Takeaways

For the first research question, Kulkarni et al. (2015) explored multiple dimensions of word usage change over time. To address this, they constructed three types of time series: frequency-based, syntactic, and distributional.

Frequency-based method relies on analyzing the frequency of word usage over time. It provides immediate insights into trends and spikes in word usage. However, the frequency method is prone to sampling errors, it may indicate a change in frequency without a significant shift in meaning, as seen with the word ‘Hurricane’ during events like Hurricane Sandy, where frequency increased but meaning remained stable. Syntactic-based method analyzes the part of speech (POS) tags of words to detect changes in their syntactic roles over time. Distributional-based method is based on the fact that words appearing in similar contexts are semantically similar.

Among the three methods employed by Kulkarni et al. (2015) for constructing time series, the distributional time series proved most insightful, as it captures the contextual usage of word. This approach utilized word embeddings to capture the semantic context of a word within specific time periods.

The syntactic method, while reliable in detecting changes in part-of-speech tags, may overlook significant shifts due to its reliance on linguistic taggers. At the time of the study, this posed a challenge due to the limitations of tagging accuracy, which has since improved significantly with the advent of more advanced methods for annotating datasets with POS tags.

### 3.4 Given that a word’s usage has changed, how can the precise moment or period of this shift be determined?

The authors introduce a method for identifying significant changes in word usage over time using a *change point detection algorithm* based on the *Mean Shift model*. The process begins by constructing time series data for each word in the corpus—whether using frequency-based, syntactic, or distributional methods. The next step is to determine if the word has experienced a significant change. If a significant shift is detected, the algorithm then identifies and returns the *estimated change point (ECP)*, which marks the time when this change occurred.

Since language exhibits a stochastic drift. In this context, ‘stochastic’ refers to the randomness or probabilistic nature of the training process, where the models are trained on the same dataset but may produce different results each time due to this randomness. To resolve this, time series was normalized for each word by transforming the time series into a *Z-score series*.

The Mean Shift model is then applied, which involves calculating the mean of the time series data before and after each potential change point. By comparing these means, the algorithm determines if a statistically significant change has occurred at any given point in time.

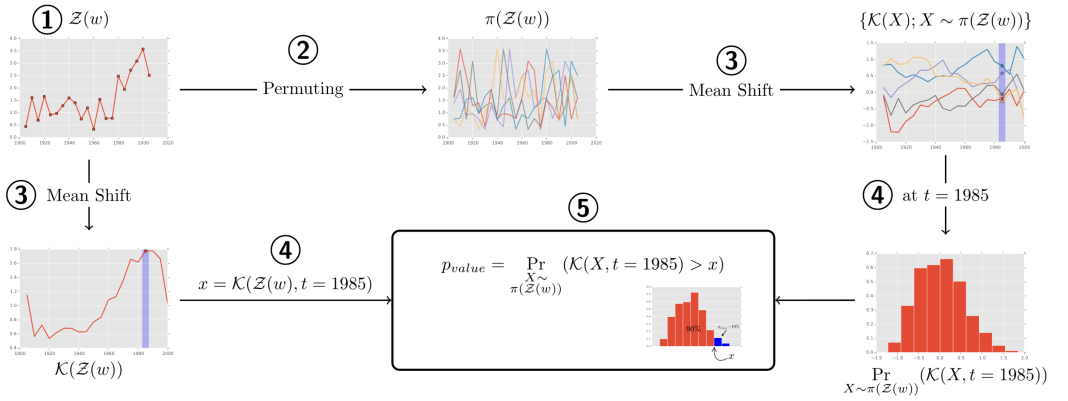


Fig. 5. Change Point Algorithm (taken from [8]).

**Change point algorithm.** After constructing time series data for each word in the corpus—whether using frequency-based, syntactic, or distributional methods (§4.2), the first step is to normalize the time series data for the word  $w$ . After normalization, the algorithm computes a mean shift series, denoted as  $K(Z(w))$ . The mean shift is defined as the difference in the mean of a time series at different time points. The algorithm calculates this shift by segmenting the time series around a pivot point, denoted as time point  $j$ , and comparing the means of the segments before and after this point. To assess the significance of the computed mean shifts, the algorithm employs

a bootstrapping method. This involves generating multiple bootstrap samples, denoted as  $BS$ . This is done by repeatedly drawing samples from the normalized data  $Z(w)$  until the number of samples equals  $B$ , which is predefined. For each sample, the algorithm calculates a p-value for each time point  $i$ . This p-value indicates how likely it is that the observed change at that time point is significant. The algorithm identifies potential change points by looking for times  $j$  where the Z-Score  $Z_j(w)$  is greater than or equal to the threshold  $\gamma$ . It then checks the p-values for these change points and selects the one with the smallest p-value, indicating the most significant change. The algorithm returns the p-value and the estimated change point (ECP). This information helps understand exactly when the meaning or usage of the word has shifted over time. Figure 5 illustrates the important aspects of the algorithm.

### 3.5 Takeaways

The change point detection algorithm (§3.4) presented in the paper detects significant linguistic shifts over time by analyzing word time series data. To detect changes, they first convert these time series into Z-score series, which normalize the data and make it easier to identify shifts. The Mean Shift model is then applied, which involves calculating the mean of the time series data before and after each potential change point. By comparing these means, the algorithm determines if a statistically significant change has occurred at any given point in time.

The paper illustrates this with the example of the word ‘tape,’ which transitioned from meaning adhesive tape to cassette tape, with the change detected in the 1970s. Similarly, the word ‘apple’ shifted from being used primarily as a common noun to a proper noun associated with Apple Inc., with this change occurring around 1984.

These examples illustrate the algorithm’s capability to pinpoint when a word’s usage and meaning undergo significant shifts. The paper emphasizes the importance of analyzing these changes to understand linguistic evolution better.

### 3.6 Results

**Time series analysis & historical analysis.** The results presented in Kulkarni et al. (2015) illustrate how words acquire new meanings and how these meanings change over time through time series and historical analysis.

**Frequency method.** For words like ‘transmitted,’ ‘bitch,’ ‘sex,’ and ‘her,’ the frequency and distributional methods reveal significant insights. For example, the sharp increase in the frequency of the word ‘her’ around the 1960s can be attributed to the concurrent rise of the feminist movement. However, the frequency method can produce many false positives due to the temporary popularity of specific social and political events.

| Method         | Examples                                       |
|----------------|--|
| Frequency      | transmitted, bitch, sex, her                   |
| Syntactic      | apple, hug, sink, click, handle, windows, bush |
| Distributional | diet, tape, plastic                            |

Table 2. Time Series Analysis

**Syntactic method.** The syntactic method uniquely detected the word ‘apple,’ which saw its most frequent part of speech tag shift significantly from ‘Noun’ to ‘Proper Noun.’ This method has a low

false positive rate but suffers from a high false negative rate, as evidenced by its detection of only two words in the study.

**Distributional method.** For words like ‘diet,’ ‘tape,’ and ‘plastic,’ the distributional method sheds light on significant changes in their usage. The popularity of dieting books, starting with the bestseller ‘Dr. Atkins’ Diet Revolution’ by Robert C. Atkins in 1972, shifted the meaning of ‘diet’ from simply referring to the food consumed to a lifestyle of food consumption behavior.

These findings demonstrate the nuanced ways in which word meanings evolve and the strengths and limitations of different methods in detecting these changes. Table 2 shows all the examples that the paper have shown for different methods. While frequency and distributional methods can highlight shifts in usage, they are prone to false positives. The syntactic method, though precise, may miss many significant changes. Therefore, a combined approach utilizing all methods may offer the most comprehensive insights into linguistic evolution

**Cross domain analysis.** The study extends its analysis to datasets like Amazon Reviews and Twitter, which cover much shorter time scales compared to the Google Books Ngram Corpus. This cross-domain analysis explores the impact of applying the distributional method on these more temporally constrained datasets. Table 3 presents words that exhibit semantic change across various datasets.

| Dataset Source | Examples                              |
|----------------|---------------------------------------|
| Amazon Reviews | streaming, ray, combo, rays, twilight |
| Twitter Tweets | candy, myster, rally, sandy           |

Table 3. Cross Domain Analysis

In the Twitter dataset, the word ‘sandy’ acquired a new sense following Hurricane Sandy hitting the East Coast of the USA. Similarly, the analysis on Amazon Reviews revealed changes in words associated with new products and technology trends.

These words demonstrate the method’s ability to capture emerging usages and meanings. The example of ‘sandy’ highlights how a natural disaster can influence linguistic shifts, with the word gaining a new context and meaning after the hurricane event.

These examples demonstrate that the method can effectively detect the introduction of new products, movies, and books, showcasing its adaptability to different domains and time scales. This cross-domain capability is particularly valuable for identifying and understanding contemporary linguistic changes in rapidly evolving contexts like social media and online reviews.

### 3.7 Takeaways

The paper by Kulkarni et al. (2015) presents an innovative approach to detecting significant linguistic changes over time using statistical methods and word embeddings. The goal behind their work is to systematically track and analyze how words evolve in meaning and usage across different temporal contexts. They construct time series for each word using three distinct methods: frequency, which captures the prevalence of word usage over time; syntactic, which examines shifts in part-of-speech tags; and distributional, which looks at changes in word co-occurrence patterns. They also employ a change point detection algorithm which identifies estimated change points, where significant linguistic shifts occur. Their analysis spans various datasets, including Google Books Ngram Viewer, Amazon Reviews, and Twitter, demonstrating the method’s adaptability and effectiveness in detecting both gradual and abrupt changes in word meanings across different domains and time scales.

## 4 RP2. Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change

### 4.1 Introduction

Hamilton et al. (2016) explores the dynamic nature of language by analyzing how word meanings evolve over time. Using *diachronic* (historical) word embeddings, a technique that maps words into vector spaces based on historical corpora, the study uncovers patterns and laws governing semantic shifts. Their approach provides a quantitative framework to examine how words change meaning, influenced by cultural, and linguistic factors.

The pace at which words undergo semantic change differs, with some words changing meaning more frequently compared to others. For example, the word ‘cat’ has remained relatively stable in its meaning, while ‘cast’ has evolved to have multiple meanings [5].

There are several hypotheses about the patterns in semantic change, including the increasing subjectification of meaning or the grammaticalization. This paper, however, focuses on following two specific questions about semantic change:

- **RQ1.** *What role does word frequency play in the evolution of word meanings?*

Frequency is a significant factor in linguistic changes, with high-frequency words often changing faster, while low-frequency words tend to be more resistant to change. The connection between word frequency and semantic change is a key unanswered question in the field of linguistics. The authors introduce the **law of conformity** to address this gap, demonstrating that frequent words change more slowly, and it clarifies the role of frequency in semantic change.

- **RQ2.** *How does polysemy relate to semantic change?*

Another unresolved question in linguistics is the relationship between semantic change and polysemy. Polysemous words, which have multiple meanings, appear in a variety of contexts. It remains unclear whether the diverse contextual use of these words makes them more or less prone to undergoing semantic change. The authors propose the **law of innovation**, which demonstrates that polysemous words are more likely to experience faster semantic changes. If a word has multiple meanings, it is more likely to lead to semantic change, especially in the case of rare senses.

In our work on data voids, we have recognized that word frequency significantly impacts their identification. For instance, the term ‘sandy’ saw a spike in usage during the Hurricane Sandy event. Similarly, terms associated with political agendas also experience increased usage during relevant events. This paper provides insights into how frequency influences the evolution of word meanings.

Additionally, our study identifies data voids involving polysemous terms, such as ‘migrant caravan.’ This term began to change in meaning when FOX News used it to describe people moving from the Mexico border towards the US, and it continued to evolve with new groups of migrants. Understanding these kinds of data voids is crucial, and this paper helps us identify subtle semantic changes in words with multiple meanings, enhancing our understanding of how these changes relate to frequency and context.

The paper aims to develop a robust methodology for quantifying semantic change using word embeddings and comparing different approaches to analyzing semantic change. This methodology is then applied in a large-scale cross-linguistic analysis spanning 200 years and four languages (English, German, French, and Chinese) to propose the above two statistical laws relating frequency and polysemy to semantic change.



## 4.2 Constructing Word Embeddings.

The authors employ three methods to construct word embeddings for different time periods. Initially, they create embeddings for each distinct period and then align these embeddings over time to maintain consistency. To quantify semantic change, they use various metrics. Specifically, they utilize Singular Value Decomposition (SVD), Positive Pointwise Mutual Information (PPMI), and Skip-Gram with Negative Sampling (SGNS). These distributional techniques represent each word by a vector that encapsulates information about the word's co-occurrence statistics.

**Positive Pointwise Mutual Information (PPMI).** PPMI is a statistical measure used to quantify the association between a word and its context within a corpus. In the paper, it is used to create word embeddings by capturing the strength of association between words and their co-occurring contexts over time.

First, a co-occurrence matrix is built, where each entry represents the frequency with which a word (target) and its context (typically a window of neighboring words) appear together in the corpus. The *joint probability* of a target word  $w_i$  and a context word  $c_j$  is calculated by dividing the co-occurrence frequency of  $w_i$  and  $c_j$  by the total number of word pairs in the corpus. The *marginal probability* of the target word  $w_i$  and the context word  $c_j$  are calculated based on the total occurrences of each word in the corpus. The Positive Pointwise Mutual Information (PPMI) between the target word  $w_i$  and the context word  $c_j$  is calculated using the formula:

$$M_{i,j}^{\text{PPMI}} = \max \left\{ \log \left( \frac{\hat{p}(w_i, c_j)}{\hat{p}(w_i)\hat{p}(c_j)} \right) - \alpha, 0 \right\}, \quad (2)$$

where:

- $\hat{p}(w_i, c_j)$  is the *joint probability* of the target word  $w_i$  and the context word  $c_j$  occurring together.
- $\hat{p}(w_i)$  and  $\hat{p}(c_j)$  are the *marginal probabilities* of the target word  $w_i$  and the context word  $c_j$ , respectively.
- $\alpha > 0$  is a positive constant used to smooth the values and prevent extreme positive values in the PPMI calculation.

This formula refines the PPMI measure by incorporating  $\alpha$  to adjust the log probability ratio. The inclusion of  $\alpha$  helps manage the sparseness of the data and the impact of low-frequency events, making the measure more robust and stable.

**Singular Value Decomposition (SVD).** SVD is a mathematical technique that factorizes a matrix into three other matrices, revealing important features of the original matrix. In the context of word embeddings, SVD helps to reduce the dimensionality of the data while preserving the essential structure and patterns in word co-occurrences. SVD decomposes the co-occurrence matrix into three matrices in the following manner:

$$\mathbf{w}_i^{\text{SVD}} = (\mathbf{U}\Sigma^\gamma)_i, \quad (3)$$

where:

- $\mathbf{w}_i^{\text{SVD}}$  is the vector representation of the word  $w_i$  in the SVD-based embedding space.
- $\mathbf{U}$  is the matrix of left singular vectors obtained from the decomposition of the co-occurrence matrix.
- $\Sigma$  is the diagonal matrix of singular values.
- $\gamma$  is a parameter that scales the singular values, allowing for tuning the influence of the different components.

In this formula, the word vector  $\mathbf{w}_i^{\text{SVD}}$  is derived by scaling the singular values in  $\Sigma$  by  $\gamma$  and then multiplying by the corresponding vectors in  $\mathbf{U}$ . This approach provides a flexible way to adjust the contribution of the components to the final word embeddings.

**Skip-Gram with Negative Sampling (SGNS).** SGNS is a technique used to learn high-quality word vectors by training on large corpora. The primary goal of SGNS is to learn word embeddings such that words that appear in similar contexts have similar vector representations. For a given target word  $w$  and a context word  $c$ , the model tries to maximize the probability  $P(c|w)$

$$\hat{p}(c_i | w_i) \propto \exp(\mathbf{w}_i^{\text{SGNS}} \cdot \mathbf{c}_j^{\text{SGNS}}), \quad (4)$$

where:

- $\hat{p}(c_i | w_i)$  is the estimated probability of a context word  $c_i$  given a target word  $w_i$ .
- $\mathbf{w}_i^{\text{SGNS}}$  is the vector representation of the target word  $w_i$  in the SGNS embedding space.
- $\mathbf{c}_j^{\text{SGNS}}$  is the vector representation of the context word  $c_j$  in the SGNS embedding space.

The probability  $\hat{p}(c_i | w_i)$  that a word  $c_i$  is a context word of  $w_i$  is proportional to the exponential of the dot product of their corresponding vector representations. The dot product  $\mathbf{w}_i^{\text{SGNS}} \cdot \mathbf{c}_j^{\text{SGNS}}$  measures the similarity between the target word and the context word in the vector space. A higher dot product indicates that the words are more likely to co-occur, leading to a higher estimated probability.

The model uses this probability estimation to distinguish true word-context pairs from randomly sampled negative pairs. The training objective is to maximize the probability of observed (target, context) pairs and minimize the probability of randomly sampled pairs, thereby learning embeddings that reflect the semantic relationships between words.

**Aligning Embeddings.** Since the embeddings are trained independently for each time period, their vector spaces may not be directly comparable. The paper addresses this by aligning the embeddings from different time periods. They aligned the embeddings across time periods, which transforms the embeddings to minimize the distances between corresponding words in different periods. This step ensures that similar meanings are represented similarly across time.

### 4.3 Takeaways

The authors created three types of word embeddings using PPMI, SVD, and SGNS, each capturing different information. PPMI measures how closely related two words are, resulting in sparse, high-dimensional vectors. SVD simplifies this information by reducing the dimensionality of the PPMI matrix into dense, low-dimensional vectors. Lastly, SGNS learns embeddings by predicting surrounding words in a given context, uncovering deeper semantic and contextual nuances that the other methods may not capture. Additionally, they employed alignment technique to align word embeddings across different time periods, ensuring consistency and enabling the analysis of semantic shifts over time.

### 4.4 Evaluation & Results

They compared the different approaches discussed in §4.2 on *synchronic accuracy* (similarity within time-period) and *diachronic validity* (semantic changes over time).

- **Synchronic Accuracy** Synchronic accuracy refers to the accuracy with which word embeddings capture the semantic relationships and meanings of words at a *specific point in time*. It assesses how well the models represent the semantic relationships among words in a given time period. SVD performed best here.

- **Diachronic Validity** Diachronic validity evaluates how effectively the methods can detect and quantify changes in *meaning over time*. It involves testing the methods against historical data to see if they can accurately identify known shifts in word meanings. The results indicated that SGNS was particularly effective in capturing these shifts, followed by SVD and PPMI.
  - **Detecting Known Shifts.** The goal in this task is for the methods to correctly capture whether pairs of words moved closer or further apart in semantic space during a pre-determined time-period.
  - **Discovering Shifts from data.** Tested whether the methods discover reasonable shifts by examining the top-10 words that changed the most from the 1900s to the 1990s according to the semantic displacement metric.

#### 4.5 Statistical Laws of Semantic Change

The authors establish two statistical laws of semantic change by analyzing a large dataset across multiple languages and time periods. They demonstrate these laws by involving the comparison of different word embedding models (PPMI, SVD, SGNS) against known historical semantic shifts and novel benchmarks.

How diachronic embeddings can be used to reveal statistical laws that relate frequency and polysemy to semantic change.

- **Law of conformity:** *Frequently used words change at slower rates.*  
Words with higher frequencies tend to experience slower rates of semantic change. This is statistically supported by the observation that high-frequency words have smaller semantic displacements over time.
- **Law of innovation:** *Polysemous words change at faster rates.*  
Polysemous words tend to change more rapidly. The study finds that these words show larger semantic displacements, indicating more significant changes in meaning over time.

#### 4.6 Takeaways

This paper explores how different types of word embedding models, such as PPMI, SVD, and SGNS, can be used to study diachronic shifts in word meanings. By analyzing these aligned embeddings, the authors identify two key statistical laws of semantic change. These findings provide a quantitative framework for understanding how and why languages evolve over time.

## References

- [1] Mario Giulianelli, Marco Del Tredici, and Raquel Fernández. 2020. Analysing Lexical Semantic Change with Contextualised Word Representations. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Dan Jurafsky, Joyce Chai, Natalie Schluter, and Joel Tetreault (Eds.). Association for Computational Linguistics, Online, 3960–3973. <https://doi.org/10.18653/v1/2020.acl-main.365>
- [2] Michael Golebiewski and danah boyd. 2019. Data Voids: Where Missing Data Can Easily Be Exploited.
- [3] Kristina Gulordava and Marco Baroni. 2011. A distributional similarity approach to the detection of semantic change in the Google Books Ngram corpus.. In *Proceedings of the GEMS 2011 Workshop on GEometrical Models of Natural Language Semantics*, Sebastian Pado and Yves Peirsman (Eds.). Association for Computational Linguistics, Edinburgh, UK, 67–71. <https://aclanthology.org/W11-2508>
- [4] William L. Hamilton, Jure Leskovec, and Dan Jurafsky. 2016. Cultural Shift or Linguistic Drift? Comparing Two Computational Measures of Semantic Change. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, Jian Su, Kevin Duh, and Xavier Carreras (Eds.). Association for Computational Linguistics, Austin, Texas, 2116–2121. <https://doi.org/10.18653/v1/D16-1229>
- [5] William L. Hamilton, Jure Leskovec, and Dan Jurafsky. 2016. Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, Katrin Erk and Noah A. Smith (Eds.). Association for Computational Linguistics, Berlin, Germany, 1489–1501. <https://doi.org/10.18653/v1/P16-1141>
- [6] Geoffrey E. Hinton. 1986. Learning distributed representations of concepts. In *Proceedings of the eighth annual conference of the cognitive science society*. Amherst, MA, Amherst, MA, 1–12.
- [7] Yoon Kim, Yi-I Chiu, Kentaro Hanaki, Darshan Hegde, and Slav Petrov. 2014. Temporal Analysis of Language through Neural Language Models. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, Cristian Danescu-Niculescu-Mizil, Jacob Eisenstein, Kathleen McKeown, and Noah A. Smith (Eds.). Association for Computational Linguistics, Baltimore, MD, USA, 61–65. <https://doi.org/10.3115/v1/W14-2517>
- [8] Vivek Kulkarni, Rami Al-Rfou, Bryan Perozzi, and Steven Skiena. 2014. Statistically Significant Detection of Linguistic Change. arXiv:1411.3315 [cs.CL] <https://arxiv.org/abs/1411.3315>
- [9] Andrey Kutuzov, Lilja Øvrelid, Terrence Szymanski, and Erik Velldal. 2018. Diachronic word embeddings and semantic shifts: a survey. In *Proceedings of the 27th International Conference on Computational Linguistics*, Emily M. Bender, Leon Derczynski, and Pierre Isabelle (Eds.). Association for Computational Linguistics, Santa Fe, New Mexico, USA, 1384–1397. <https://aclanthology.org/C18-1117>
- [10] Julieta Laurino, Simon De Deyne, Álvaro Cabana, and Laura Kaczer. 2023. The Pandemic in Words: Tracking Fast Semantic Changes via a Large-Scale Word Association Task. *Open Mind* 7 (06 2023), 221–239. [https://doi.org/10.1162/opmi\\_a\\_00081](https://doi.org/10.1162/opmi_a_00081) arXiv:[https://direct.mit.edu/opmi/article-pdf/doi/10.1162/opmi\\_a\\_00081/2133848/opmi\\_a\\_00081.pdf](https://direct.mit.edu/opmi/article-pdf/doi/10.1162/opmi_a_00081/2133848/opmi_a_00081.pdf)
- [11] Xuanyi Liao and Guang Cheng. 2016. Analysing the Semantic Change Based on Word Embedding. In *Natural Language Understanding and Intelligent Applications*, Chin-Yew Lin, Nianwen Xue, Dongyan Zhao, Xuanjing Huang, and Yansong Feng (Eds.). Springer International Publishing, Cham, 213–223.
- [12] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Distributed representations of words and phrases and their compositionality. In *Proceedings of the 26th International Conference on Neural Information Processing Systems - Volume 2* (Lake Tahoe, Nevada) (*NIPS’13*). Curran Associates Inc., Red Hook, NY, USA, 3111–3119.