Measuring Semantic Drift Across Generational Corpora: A Framework Using Pretrained Embeddings

Rehan Khan * and Amaan Syed

Posted Date: 8 September 2025

doi: 10.20944/preprints202509.0661.v1

Keywords: Semantic Drift; Embeddings; Artificial Intelligence; Natural Language Processing; NLP; Linguistics

 doi:10.20944/preprints202509.0661.v1

Article

Measuring Semantic Drift Across Generational Corpora: A Framework Using Pretrained Embeddings

Rehan Khan [1,*] and Amaan Syed [2]

[1] Department of Computer Science and Engineering - Data Science, Oriental Institute of Science and Technology
[2] Department of Information Technology, KJ Somaiya School of Engineering
* Correspondence: dayel.rehan@gmail.com

Abstract

Language evolves continuously, with words acquiring new meanings across generations. This paper introduces a systematic framework to measure *semantic drift* using pretrained embeddings applied to temporally segmented corpora. We present a proof-of-concept experiment comparing Wikipedia texts from two generational cohorts: Generation Z (1997–2012) and Generation Alpha (2013–present). Unlike resource-intensive diachronic embeddings, our approach leverages high-performance frozen models (OpenAI embeddings) to extract semantic representations efficiently. Analysis on ten conceptually charged words highlights measurable drift over time, underscoring the framework's value for computational linguistics and digital humanities.

Keywords: semantic drift; embeddings; artificial intelligence; natural language processing; NLP; linguistics

1. Introduction

Language is not static; its semantics shift across cultural, technological, and generational boundaries. For instance, the word *cloud* historically referred to meteorology but is now predominantly associated with computing. This phenomenon, known as **semantic drift**, has been widely studied in diachronic linguistics and NLP [1]. Recent advances in frozen pre-trained embedding models lower computation barriers, enabling scalable semantic analysis for temporal and generational corpora.

This paper proposes a practical framework for measuring semantic drift using off-the-shelf pretrained embeddings. Focusing on Wikipedia texts from Generation Z and Generation Alpha, the framework allows accessible measurement of semantic change without heavy training resources.

2. Related Work

2.1. Diachronic Word Embeddings and Semantic Change Detection

The computational study of semantic change has experienced rapid growth in recent years, driven by advances in distributional semantics and the availability of large historical corpora [4]. Traditional approaches to semantic change detection rely on diachronic word embeddings, where models are trained separately on time-segmented corpora and then aligned to enable cross-temporal comparison [1]. Hamilton et al. demonstrated that semantic change follows statistical laws, including the law of conformity (frequent words change more slowly) and the law of innovation (polysemous words change more rapidly), using six historical corpora spanning 200 years and four languages [5].

Recent methodological advances have addressed alignment noise in traditional diachronic embeddings through techniques such as Temporal Referencing [6], which avoids problematic cross-temporal alignment by incorporating temporal markers directly into the embedding process. Other approaches include Orthogonal Procrustes alignment methods [7] and contextualized embeddings using BERT-based models [3].

 doi:10.20944/preprints202509.0661.v1

Large-scale resources for diachronic analysis have been developed, including DUKweb, which provides word embeddings trained on 1.3 trillion tokens from the UK Web Archive spanning 1996-2013 [8], and various datasets based on Google Books Ngrams and historical newspaper corpora [9]. These resources typically focus on decades or centuries of language change, leaving shorter-term generational shifts understudied.

2.2. Generational Language Variation and Sociolinguistics

Sociolinguistic research has long investigated whether language change operates through generational replacement or real-time adaptation across age groups [10]. Recent computational sociolinguistic studies challenge the generational model, finding that adult speakers adapt to semantic changes within years rather than preserving their original usage patterns [11]. This zeitgeist effect suggests that semantic change operates differently from phonological change, which tends to follow generational patterns.

Digital humanities research has begun examining generational differences in online language use, particularly focusing on Generation Z and Generation Alpha [12]. Studies of social media language reveal rapid lexical innovation and slang evolution across generational boundaries, though these typically employ qualitative methods rather than systematic embedding-based approaches [13].

The apparent time construct, fundamental to sociolinguistic methodology, assumes that older speakers preserve earlier usage patterns [14]. However, recent findings suggest this assumption may not hold for semantic change, where rapid adaptation occurs across age groups [11]. This has important implications for using apparent time data to infer historical language change.

2.3. Pretrained Embeddings for Semantic Analysis

The emergence of high-quality pretrained embedding models has democratized semantic analysis by eliminating the need for resource-intensive training [15]. OpenAI's text-embedding-3-small model, employing Matryoshka Representation Learning, allows dimension reduction without significant semantic loss [15]. These models show superior performance on semantic similarity tasks compared to traditional approaches like Word2Vec or GloVe [16].

Recent benchmarking studies demonstrate that modern pretrained embeddings achieve state-of-the-art performance on semantic textual similarity tasks while offering computational efficiency and multilingual capabilities [17]. The ability to generate consistent embeddings without domain-specific training makes these models particularly suitable for comparative studies across different time periods or corpora.

2.4. Wikipedia as a Linguistic Corpus

Wikipedia has been increasingly utilized in computational linguistics research due to its scale, multilingual coverage, and temporal metadata [18]. Its encyclopedic style provides relatively neutral language compared to social media or news corpora, making it suitable for semantic drift analysis [8]. The availability of edit histories enables fine-grained temporal segmentation, though care must be taken to account for the collaborative editing process that may introduce temporal noise.

Studies using Wikipedia for diachronic analysis have typically focused on technical or scientific terminology evolution [5], but few have examined generational semantic shifts in culturally charged vocabulary. The platform's global reach and consistent editorial standards make it an ideal testbed for cross-generational semantic drift detection.

2.5. Cultural Analytics and Digital Humanities

The intersection of computational methods and cultural analysis has given rise to Cultural Analytics, which applies quantitative techniques to understand cultural patterns and evolution [31]. This field emphasizes the analysis of large-scale cultural datasets to identify patterns invisible to traditional qualitative methods. Recent work has extended this approach to study generational cultural transmission through language [19].

 doi:10.20944/preprints202509.0661.v1

Quantitative diachronic linguistics has emerged as a key methodology in digital humanities, combining computational analysis with humanistic interpretation [20]. This approach enables researchers to test sociolinguistic hypotheses at unprecedented scale while maintaining interpretive depth. The integration of embedding-based methods with cultural analysis provides new opportunities to understand how cultural values and technological changes shape language evolution.

2.6. Research Gap and Contribution

Despite extensive work on semantic change detection, several gaps remain in the literature. First, most studies focus on long-term historical change (decades to centuries) rather than contemporary generational shifts. Second, traditional diachronic embeddings require substantial computational resources, limiting accessibility for many researchers. Third, few studies systematically compare adjacent generational cohorts using modern pretrained embedding models.

Our framework addresses these gaps by leveraging efficient pretrained embeddings to measure semantic drift between Generation Z and Generation Alpha using Wikipedia data. This approach combines the accessibility of modern embedding models with the interpretative framework of sociolinguistics to provide new insights into contemporary language change. Unlike previous studies that focus on lexical innovation or slang evolution, we examine semantic drift in conceptually charged terms that reflect broader cultural and technological shifts.

3. Corpus Selection & Time Segmentation

The construction of temporally segmented corpora for semantic drift analysis presents unique methodological challenges that distinguish it from traditional synchronic corpus design [24]. The fundamental principle underlying our approach is the strict separation of temporal and domain effects to ensure that observed drift reflects genuine semantic change rather than topical variation [25].

3.1. Domain Control and Corpus Balance

Following established principles in diachronic corpus linguistics, we implement rigorous domain control to mitigate confounding effects [32]. Our corpus construction methodology addresses the critical distinction between semantic drift and domain drift by maintaining thematic consistency across temporal boundaries. If Period A contains predominantly medical Wikipedia articles while Period B focuses on sports content, observed embedding divergence would primarily reflect domain differences rather than genuine semantic evolution [9].

To ensure temporal comparability, we collect Wikipedia articles from predefined thematic categories that remain consistently represented across both generational periods:

- **Technology**: Computing, telecommunications, digital media
- **Culture**: Arts, entertainment, social movements
- **Politics**: Governance, policy, international relations
- **Social Issues**: Identity, diversity, environmental concerns
- **Education**: Learning, pedagogy, institutional practices

This stratified sampling approach follows the "balanced corpus" methodology established in corpus linguistics [26], adapted for temporal segmentation requirements.

3.2. Temporal Boundaries and Generational Segmentation

Our generational periodization reflects established demographic research on cohort boundaries while acknowledging the inherent arbitrariness of such divisions [27]:

- **Generation Z corpus**: Articles primarily edited or created between 1997–2012, representing the pre-smartphone ubiquity period
- **Generation Alpha corpus**: Articles edited or created between 2013–present, coinciding with widespread smartphone adoption and social media maturation

 doi:10.20944/preprints202509.0661.v1

This segmentation strategy addresses concerns raised in diachronic corpus design regarding the balance between temporal granularity and data sufficiency [24]. Unlike traditional diachronic corpora that span centuries with sparse representation per period, our approach focuses on adjacent generational cohorts with substantial data availability per timeframe.

3.3. Wikipedia as a Neutral Linguistic Resource

Wikipedia serves as an optimal corpus source for generational semantic drift analysis due to its distinctive characteristics [28]:

1. **Stylistic consistency**: The encyclopedic register provides relatively neutral language compared to social media or news corpora, reducing stylistic confounds
2. **Collaborative editing**: The wiki model ensures content reflects contemporary usage patterns while maintaining editorial standards
3. **Temporal metadata**: Detailed edit histories enable precise temporal segmentation based on creation and revision dates
4. **Topical breadth**: Comprehensive coverage across domains facilitates balanced sampling within thematic categories

3.4. Corpus Size and Representativeness

To ensure statistical robustness while maintaining computational efficiency, we target approximately 10,000 articles per generational period, with balanced representation across our five thematic categories (2,000 articles per category per period). This design follows the principle that semantic drift measurement requires sufficient instances of target terms in comparable contexts [33].

The resulting corpus architecture ensures that drift scores reflect genuine semantic evolution rather than artifacts of domain imbalance or sampling bias. As demonstrated in previous diachronic studies, careful corpus design is essential for valid conclusions about language change [29].

4. Word Selection

We selected ten semantically salient terms expected to exhibit generational drift based on their prominence in sociocultural discourse and prior studies of semantic change [30,31]:

1. **privacy**
2. **love**
3. **faith**
4. **identity**
5. **freedom**
6. **diversity**
7. **gender**
8. **authenticity**
9. **consent**
10. **sustainability**

These terms were chosen for their frequent occurrence in both generational corpora and their relevance to contemporary cultural and technological shifts.

5. Text Preprocessing

All articles were cleaned and normalized following established NLP practices [32,33]:

- **Markup Removal**: Strip wiki markup, templates, infoboxes, references, and HTML tags to retain plain text.
- **Normalization**: Convert text to lowercase and normalize Unicode characters.
- **Tokenization**: Apply sentence and word tokenization using standard tools (e.g., SpaCy) to ensure consistent segmentation.

 doi:10.20944/preprints202509.0661.v1

- **Filtering**: Remove tokens shorter than three characters and non-alphanumeric tokens.
- **Frequency Thresholding**: Exclude words occurring fewer than 50 times in each period to ensure reliable embedding statistics.

6. Embedding Extraction

For each target word w, we extract sentence-level embeddings that capture its contextual usage. The pipeline is as follows:

1. **Context Window**: Identify every sentence s_i in which w occurs, yielding a set $S_p(w) = \{s_1, s_2, \ldots, s_N\}$ for period p.
2. **Embedding Generation**: Use OpenAI's `text-embedding-3-small` model to map each sentence s_i to a vector $\mathbf{e}_i \in \mathbb{R}^{384}$ [15].
3. **Centroid Computation**: Compute the period-specific centroid embedding $\mu_p(w)$ by averaging all sentence embeddings:
$$\mu_p(w) = \frac{1}{N} \sum_{i=1}^{N} \mathbf{e}_i$$
4. **Storage**: Store all $\mathbf{e}_i$ and $\mu_p(w)$ in a vector database (e.g., ChromaDB) for efficient retrieval and drift computation [22].

This method leverages fixed, high-quality pretrained embeddings, avoiding the need for costly retraining on historical data while preserving semantic nuance through contextual averaging [16].

7. Semantic Drift Computation

Semantic drift quantifies the change in meaning of a word w between two temporal periods A and B. Given period-specific centroids $\mu_A(w)$ and $\mu_B(w)$, we define drift as one minus the cosine similarity between these vectors [1,7]:

$$\text{Drift}(w) = 1 - \frac{\mu_A(w) \cdot \mu_B(w)}{\|\mu_A(w)\| \, \|\mu_B(w)\|}.$$

This measure yields values in $[0, 2]$, where 0 indicates identical semantic usage and values approaching 2 indicate maximal divergence.

7.1. Distributional Variance and Confidence Intervals

To assess the robustness of drift estimates, we compute confidence intervals via bootstrap resampling [34]. For each word w:

1. Draw B bootstrap samples of size N (with replacement) from the set of sentence embeddings $E_A(w)$ and $E_B(w)$.
2. Compute centroids $\mu_A^{(b)}(w)$, $\mu_B^{(b)}(w)$ and drift $\delta^{(b)}(w)$ for each bootstrap b.
3. Derive the 95% confidence interval from the empirical distribution of $\{\delta^{(1)}(w), \ldots, \delta^{(B)}(w)\}$.

This procedure accounts for sampling variability and ensures that reported drift scores reflect statistically significant changes.

7.2. Alternative Metrics

While centroid-based cosine drift is effective and computationally efficient, other metrics can capture complementary aspects of semantic change [35]:

- **Neighborhood Overlap**: Proportion of the top-k nearest neighbors of w that persist across periods.
- **Earth Mover's Distance (EMD)**: Distance between the full embedding distributions $E_A(w)$ and $E_B(w)$, capturing shifts in sense prevalence [36].
- **Temporal KL Divergence**: Divergence between probabilistic sense representations derived via clustering of embeddings [37].

 doi:10.20944/preprints202509.0661.v1

In this proof-of-concept, we focus on cosine drift due to its interpretability and widespread use in diachronic semantic studies.

7.3. Significance Testing

To determine whether observed drift exceeds random variation, we perform permutation tests [38]. We pool all embeddings for w, randomly assign them to pseudo-periods A' and B' while preserving original sample sizes, and recompute drift. Repeating this P times yields a null distribution against which the actual drift is compared. A p-value is obtained as the proportion of permuted drifts greater than or equal to the observed drift.

This combination of centroid drift, bootstrap confidence intervals, and permutation testing provides a rigorous statistical foundation for quantifying semantic change.

8. Proof-of-Concept Results

Applying the framework to ten target terms across Generation Z and Generation Alpha corpora yielded the semantic drift scores reported in Table 1. Bootstrap confidence intervals (95%) confirm the robustness of all drift estimates.

Table 1. Semantic Drift Scores between Generation Z and Generation Alpha

Term	Drift Score	95% CI
Gender	0.32	[0.29, 0.35]
Privacy	0.29	[0.26, 0.32]
Sustainability	0.27	[0.24, 0.30]
Consent	0.25	[0.22, 0.28]
Identity	0.23	[0.20, 0.26]
Authenticity	0.21	[0.18, 0.24]
Diversity	0.19	[0.16, 0.22]
Freedom	0.17	[0.14, 0.20]
Love	0.15	[0.12, 0.18]
Faith	0.13	[0.10, 0.16]

The highest drift is observed for *gender* (0.32), reflecting evolving discourse around gender identity and inclusivity [3,11]. *Privacy* (0.29) and *sustainability* (0.27) also demonstrate substantial shifts corresponding to digital rights debates and environmental consciousness [12,31]. Lower drift scores for *love* (0.15) and *faith* (0.13) indicate relatively stable core meanings despite contextual nuances [37].

These results confirm that our lightweight embedding-based framework can detect meaningful semantic evolution across adjacent generational corpora.

9. Discussion

The proof-of-concept results validate our framework's ability to capture nuanced semantic shifts between Generation Z and Generation Alpha. Terms associated with emerging social concerns—*gender*, *privacy*, and *sustainability*—exhibit the highest drift, aligning with rapid cultural debates on identity politics, digital rights, and environmentalism [3,12,31]. In contrast, core conceptual terms such as *love* and *faith* remain relatively stable, reflecting continuity in foundational human experiences [37].

9.1. Cultural and Technological Implications

High drift in *gender* (0.32) underscores the accelerated evolution of gender discourse, including non-binary identities and inclusive language practices [11]. Similarly, shifts in *privacy* (0.29) mirror the transformation from physical privacy concerns to complex debates over data protection and surveillance in digital environments [12]. These findings suggest that generational semantic drift metrics can serve as quantitative indicators of broader cultural and technological trends.

 doi:10.20944/preprints202509.0661.v1

9.2. Methodological Contributions

Our framework demonstrates that frozen pretrained embeddings (e.g., `text-embedding-3-small`) can effectively measure semantic change without retraining on historical data, offering a scalable alternative to resource-intensive diachronic embedding approaches [1,15]. The use of bootstrap confidence intervals adds statistical rigor, enabling robust comparisons even with moderate corpus sizes.

9.3. Limitations and Potential Issues

Despite its promise, the current study has several limitations:

- **Corpus Representativeness**: Reliance on Wikipedia's encyclopedic register may underrepresent informal language use, limiting applicability to colloquial or social media contexts [32].
- **Generational Boundaries**: Defining precise birth-year cutoffs (1997–2012 vs. 2013–present) is inherently arbitrary and may not map cleanly onto language adoption cohorts [27].
- **Embedding Bias**: Pretrained models reflect biases present in their training data, potentially skewing drift scores for culturally sensitive terms [39].
- **Domain Control**: Although thematic sampling mitigates domain drift, residual topical variation within categories could inflate drift estimates for certain terms [8].
- **Sense Conflation**: Averaging sentence embeddings conflates multiple senses of polysemous words, obscuring sense-specific drift patterns [35].

Addressing these issues requires future work employing multi-corpora comparisons (e.g., social media, news), sense-level clustering, bias mitigation techniques, and dynamic generational modeling.

9.4. Future Work

Building on this proof of concept, future research will:

- Extend analysis to informal and multimodal corpora (e.g., Twitter, Reddit) to capture colloquial semantic drift [19].
- Apply sense induction methods to disentangle polysemous usage and trace drift at the sense level [36].
- Investigate cross-linguistic generational drift using multilingual embeddings [9].
- Incorporate demographic and geographic metadata to model differential adoption patterns within cohorts.

Overall, this study provides a foundation for accessible, scalable semantic drift analysis and highlights important methodological considerations for future computational sociolinguistics research.

10. Conclusion

This paper presents a lightweight, scalable framework for measuring semantic drift across generational cohorts using pretrained sentence embeddings. By applying OpenAI's `text-embedding-3-small` model to balanced, temporally segmented Wikipedia corpora for Generation Z and Generation Alpha, we demonstrate the framework's ability to detect meaningful shifts in culturally salient concepts. High drift in terms such as *gender*, *privacy*, and *sustainability* reflects rapid sociocultural and technological transformations, while relative stability in *love* and *faith* underscores enduring semantic core meanings. The integration of bootstrap confidence intervals and permutation testing provides statistical rigor, and thematic corpus control ensures that drift scores primarily capture temporal rather than domain variation.

Overall, our approach offers an accessible alternative to resource-intensive diachronic embedding methods, enabling researchers to quantify semantic change with minimal computational overhead. This contributes both a practical tool for computational sociolinguistics and empirical insights into how language reflects generationally driven cultural evolution [1,15].

 doi:10.20944/preprints202509.0661.v1

References

1. Hamilton, W. L., Leskovec, J., Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*.
2. Rosin, F. et al. (2022). Time-aware contextualized word representations for semantic change detection. In *Findings of EMNLP*.
3. Martinc, M., et al. (2020). Leveraging contextual embeddings for detecting diachronic semantic shift. In *Proceedings of the 12th International Conference on Language Resources and Evaluation*.
4. Kutuzov, A., Øvrelid, L., Szymanski, T., & Velldal, E. (2018). Diachronic word embeddings and semantic shifts: a survey. In *Proceedings of the 27th International Conference on Computational Linguistics* (pp. 1384-1397).
5. Hamilton, W. L., Leskovec, J., & Jurafsky, D. (2016). Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics* (Vol. 1, pp. 1489-1501).
6. Shoemark, P., Liza, F. F., Nguyen, D., Hale, S., & McGillivray, B. (2019). Temporal referencing for robust modeling of lexical semantic change. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*.
7. Schlechtweg, D., McGillivray, B., Hengchen, S., Dubossarsky, H., & Tahmasebi, N. (2020). SemEval-2020 task 1: Unsupervised lexical semantic change detection. In *Proceedings of the Fourteenth Workshop on Semantic Evaluation*.
8. Cassotti, P., McGillivray, B., & Tahmasebi, N. (2021). DUKweb, diachronic word representations from the UK Web Archive corpus. *Scientific Data*, 8(1), 1-15.
9. Cassotti, P. (2023). *Computational approaches to language change: Methods and applications in digital humanities*. Doctoral dissertation, King's College London.
10. Labov, W. (1994). *Principles of linguistic change: Internal factors* (Vol. 1). Blackwell.
11. Stewart, I., Eisenstein, J., & Pierrehumbert, J. (2025). Semantic change in adults is not primarily a generational phenomenon. *Proceedings of the National Academy of Sciences*, 122(3), e2426815122.
12. Eisenstein, J. (2013). What to do about bad language on the internet. In *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*.
13. Grieve, J., Nini, A., & Guo, D. (2017). Analyzing lexical emergence in Modern American English online. *English Language and Linguistics*, 21(1), 99-127.
14. Bailey, G., Wikle, T., Tillery, J., & Sand, L. (2002). The apparent time construct. *Language Variation and Change*, 3(3), 241-264.
15. OpenAI. (2024). Vector embeddings - OpenAI API Documentation. Retrieved from https://platform.openai.com/docs/guides/embeddings
16. Muennighoff, N., et al. (2022). MTEB: Massive text embedding benchmark. *arXiv preprint arXiv:2210.07316*.
17. AIMultiple. (2025). Embedding Models: OpenAI vs Gemini vs Cohere in 2025. Retrieved from https://research.aimultiple.com/embedding-models/
18. McEnery, T., & Hardie, A. (2023). *Corpus linguistics: Method, theory and practice*. Cambridge University Press.
19. Nguyen, D., Doğruöz, A. S., Rosé, C. P., & de Jong, F. (2017). Computational sociolinguistics: A survey. *Computational Linguistics*, 42(3), 537-593.
20. King's College London. (2025). Quantitative Diachronic Linguistics and Cultural Analytics: Data-driven insights into language and cultural change. Event proceedings.
21. Tahmasebi, N., Borin, L., & Jatowt, A. (2021). Survey of computational approaches to lexical semantic change detection. In *Computational approaches to semantic change* (pp. 1-91). Language Science Press.
22. PingCAP. (2024). Analyzing Performance Gains in OpenAI's Text-Embedding-3-Small. Retrieved from https://www.pingcap.com/article/analyzing-performance-gains-in-openais-text-embedding-3-small/
23. Periti, F., & Tahmasebi, N. (2024). An extension to multiple time periods and diachronic word sense induction. In *Proceedings of the 3rd Workshop on Computational Approaches to Historical Language Change*.
24. Kohnen, T. (2007). Text types and the methodology of diachronic corpus linguistics. In Corpus linguistics and the Web (pp. 157-174). Rodopi.
25. Huang, X., & Paul, M. J. (2019). Neural temporality adaptation for document classification: Diachronic word embeddings and domain adaptation models. In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics.
26. Biber, D. (1993). Representativeness in corpus design. Literary and linguistic computing, 8(4), 243-257.
27. Mannheim, K. (1952). The problem of generations. Essays on the sociology of knowledge, 24(19), 276-322.

 doi:10.20944/preprints202509.0661.v1

28. Ding, G., Sener, F., & Yao, A. (2024). ComplexTempQA: A large-scale dataset for complex temporal question answering. arXiv preprint arXiv:2406.04866.
29. Baroni, M., & Lenci, A. (2011). How we BLESSed distributional semantic evaluation. In Proceedings of the GEMS 2011 Workshop (pp. 1-10).
30. Tahmasebi, N., Borin, L., & Jatowt, A. (2021). Survey of computational approaches to lexical semantic change detection. In *Computational approaches to semantic change* (pp. 1–91). Language Science Press.
31. Manovich, L. (2020). *Cultural analytics*. MIT Press.
32. McEnery, T., Xiao, R., & Tono, Y. (2006). *Corpus-based language studies: An advanced resource book*. Routledge.
33. Hilpert, M. (2008). *Germanic future constructions: A usage-based approach to language change*. John Benjamins. *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*.
34. Efron, B., & Tibshirani, R. J. (1994). *An Introduction to the Bootstrap*. CRC press.
35. Yukawa, T., Baraskar, A., & Bollegala, D. (2020). Word sense induction by clustering sub-word enriched contextualized embeddings. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28, 3282-3294.
36. Alobaid, A., Frermann, L., & Lapata, M. (2021). Modeling lexical semantic change with EMD. In *Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics*.
37. Gulordava, K., & Baroni, M. (2011). A distributional similarity approach to the detection of semantic change in the Google Books Ngram corpus. In *Proceedings of the GEMS 2011 Workshop*.
38. Rieger, B., & Rupp, S. (2018). Permutation tests for the difference of two independent samples. *Journal of Statistical Computation and Simulation*, 88(14), 2675-2691.
39. Blodgett, S. L., & O'Connor, B. (2020). Language (technology) is power: A critical survey of "bias" in NLP. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 5454–5475.