Lexical Semantic Change Detection

Nicole Chen, Tammy Yang, Tzu-Yu Ko

UNIVERSITY OF MICHIGAN

 # SemanticChange
 # ComputationalModels
 # NLPResearch
 # AI application

Research Objectives

1. **Detect and quantify lexical semantic changes in Spanish.**
2. **Measure the evolution of Spanish word meanings.**

Introduction

Traditionally, linguists relied on manual hand-annotated word approaches for vocabulary semantic change evaluation, but recent advancements in computer science and computational linguistics have introduced self-driving language models. Leveraging digitized historical documentation and large-scale corpora, this research focuses on building a model for detecting lexical semantic change, aiming to contribute to information access systems in fields such as digital journalism and online chatbots.

Data

We utilized two distinct language corpora: the old corpus (1810-1906) and the modern corpus (1994-2020), and they were processed using spaCy and the target words were selected.
Lexical Semantic Change Detection (LSCD) was applied to identify words experiencing shifts in meaning over time. This approach streamlined the selection of target words and the creation of annotated usage samples.

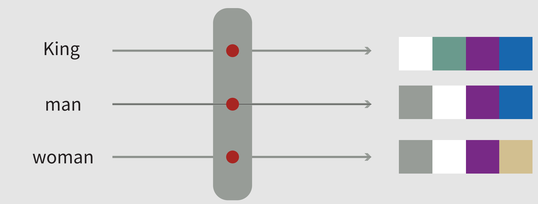
Corpus	Time Period	Tokens
Old Corpus	1810-1906	Around 13M
Modern Corpus	1994-2020	Around 22M

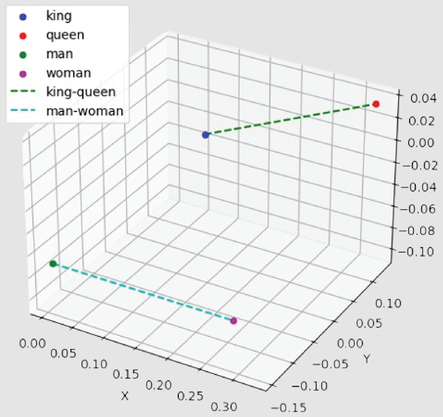
Discussion

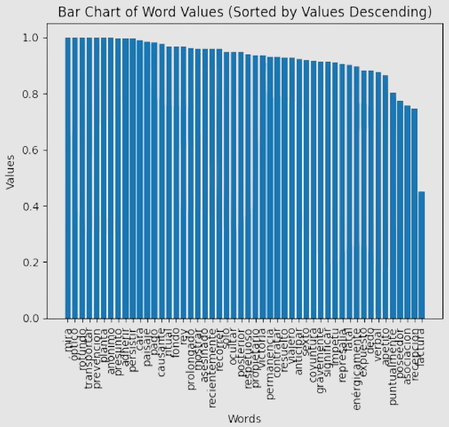
The study reveals semantic changes in words between old and modern eras, opening avenues for future research:
1. Cross-Linguistic Comparative Studies
2. Fine-Tuning for Specific Domains

Additionally, the research suggests considering diverse applications beyond textual data, such as audio and visual modalities, for diachronic language data. This opens the door for:
3. Multimodal Semantic Change Detection
4. Implementation in Information Retrieval Systems

Models







Time Now

*w*ord at Time *i* (old corpus)

- Same sense → No Evolution
- Sense different → **Word Sense Evolution**

{ Skip-Gram with Negative Sampling (SGNS)
Orthogonal Procrustes (OP)
Cosine Distance (CD)

Findings

Detect and quantify lexical semantic changes in Spanish

Using Skip-gram with Negative Sampling (SGNS), Orthogonal Procrustes (OP), and Cosine Distance (CD), our approach effectively quantified lexical semantic changes in Spanish. SGNS handled large vocabularies, and OP aligned vector representations, enabling the measurement of graded changes through cosine distance. The method identified shifts in word meanings by comparing vector representations in old and modern datasets. SGNS+OP+CD provided a robust framework for capturing subtle semantic nuances, offering a comprehensive understanding of how Spanish word meanings have evolved over time.

Measure the evolution of Spanish word meanings

We employed Spearman's Correlation Coefficient and COMPARE scores to measure the evolution of Spanish word meanings. Spearman's Correlation Coefficient assessed the strength and direction of the relationship between computed and golden scores, and COMPARE score predicted negated diachronic usage relatedness. With impressive scores of 0.543 and 0.561,respectively, our methods provide a thorough measurement of Spanish word meaning evolution. Beyond academic realms, the adaptability of our approach positions it for practical applications, highlighting its potential across diverse linguistic domains.

Measurement	Value
Spearman's Correlation Coefficient	**0.543**
COMPARE	**0.561**

Lexical Semantic Change Detection

1. Introduction

The origin and the history of our language semantic change have long been a topic of interest to social scientists. In the past, most of the research on vocabulary semantic change evaluation has been studied by linguists or scholars with manual hand-annotated word approaches. Recently, this complex and subjective research method has been noticed. Furthermore, computer scientists and computational linguistics proposed self-driving language models by computing and are looking forward to scaling up traditional semantic change research.

Historiography has been recognized as a critical resource based on its enormous language datasets spanning centuries. With digitized historical documentation developing exponentially, lexical semantic change on a single word in languages has been an essential research area for information access systems, such as digital journalism and online chatbots. As a result, large-scale corpora provide a chance to manipulate, test, and evaluate computational approaches to diachronic semantic shift. After reviewing SemEval, LSCDiscovery, and RuSemEval literature, our project is going to focus on building a model and implementing an experiment with Spanish corpus from LSCDiscovery research to practice models learned in class, including Word2Vec, SVD, Cosine Distance, and Cosine Similarity.

In this research endeavor, we aim to delve into the dynamic landscape of lexical semantics within the Spanish language, propelled by the overarching objectives of detecting and quantifying lexical semantic changes (RO1) and comprehensively measuring the evolution of Spanish words' meanings (RO2). Our research objectives direct our focus toward identifying and meticulously quantifying semantic shifts that may have occurred over time in the Spanish lexicon. By employing advanced natural language processing techniques and leveraging state-of-the-art language models, we intend to discern subtle variations in word meanings, contributing to a nuanced understanding of how semantics have transformed within the Spanish language.

Research Objectives:

- RO1: Detect and quantify lexical semantic changes in Spanish.
- RO2: Measure the evolution of Spanish words' meanings.

2. Background

2.1 The Concept of Lexical Semantic Change

Lexical semantic change is defined as “innovations which change the lexical meaning rather than the grammatical function of a form.” As time changes, vocabulary change is an interesting topic for linguists, regarding the role of language in human sphere activity. Since language is full of variation, it is extremely hard to identify the meaning of words from different materials. To solve this problem, linguists and computational linguists annotate words' meanings in specific periods

and then manually compare the meaning change in the past or build a language model by using linear approaches with vectors and evaluating meaning change.

2.2 Assessing Computational Systems for Semantic Analysis of Language

Semantic Evaluation (SemEval) originally stemmed from the Senseval word sense evaluation series. The primary goal of these evaluations is to understand how computational systems analyze and interpret the meaning of language. While humans easily grasp meaning, transferring that understanding to computers has proven challenging. The purpose of SemEval and Senseval is to assess the effectiveness of semantic analysis systems. "Semantic Analysis" involves a formal examination of meaning, and "computational" refers to approaches that can be practically implemented.

3. Related work

3.1 SemEval

SemEval serves as a platform to precisely define what aspects are crucial for computing meaning. It's a way to identify problems and solutions in the realm of meaning computation. Initially focused on identifying word senses, the evaluations have expanded to explore broader language dimensions. They now delve into areas like the relationships among sentence elements (semantic role labeling), connections between sentences (coreference), and the overall nature of language, including semantic relations and sentiment analysis.

SemEval team published “SemEval-2020 Task1: Unsupervised Lexical Semantic Change Detection” (Dominik S., Barbara M., et al., 2020), which employed 156 target words across various languages without splitting them into development and test sets. About half of these words were chosen from etymological dictionaries or research literature, and the other half from corpus vocabularies, matching the POS (Part-Of-Speech) tagging and frequency of the first group. Word occurrences in sentences were paired and annotated for semantic proximity. Exclusions were made for target words with numerous undecidable pairs or sparse annotations. Sense clusters were determined from the annotations, and binary and graded change scores were derived from these clusters. These scores were then used to assess participants in binary classification and ranking tasks. It provides researchers in this competition with all the pre-required information and datasets for multiple languages, including English, German, Latin, and Swedish.

3.2 LSCDiscovery

Lexical semantic change detection (LSCD) is the task of identifying words whose meaning changes over time. This research highlights shortcomings in previous LSCD-shared tasks, such as limited target words, inconsistent task formalizations, and a lack of focus on Spanish. To address these issues, researchers organized the first shared task on Spanish diachronic data, requiring participants to predict semantic change for the entire vocabulary (Discovery) and a subset of annotated target words (Detection). LSCD provides an extensive dataset of semantic judgments for 100 Spanish words spanning historical and modern corpora. The research aims to advance understanding and methodologies in detecting semantic changes in Spanish, contributing to future research and novel techniques in the field.

4. Method

4.1 Skip-gram Negative Sampling

Skip-gram with negative sampling is a technique used in natural language processing and word embedding models, such as Word2Vec, to learn word representations efficiently. It is an improvement over the original Skip-gram model, which tends to be computationally expensive when dealing with large vocabularies. Skip-gram with negative sampling addresses this issue by reducing the computational cost while still effectively learning word embeddings.

4.2 Orthogonal Procrustes

The main idea for the orthogonal matrix is mapping matrix A to matrix B. The optimization problem in the Orthogonal Procrustes Problem aims to find the best transformation matrix Q that minimizes the sum of squared differences between corresponding points in A and B after applying the transformation.

4.3 SGNS+OP+CD Method

The evaluation baseline we chose is Skip-Gram with Negative Sampling + Orthogonal Procrustes + Cosine Distance (SGNS+OP+CD). This approach interprets word vector representations in input corpora. After doing the same word2vec process, corpora are aligned using Orthogonal Procrustes to compute graded change as cosine distance among target words from old and modern word vectors.

4.4 Local Neighborhood Method

Generally, a local neighborhood algorithm typically focuses on the relationships or similarities between data points in their local vicinity. Instead of considering the entire dataset, the algorithm zooms in on a specific point or set of points and examines their relationships. This approach is often used in tasks where the local structure of the data is important, such as clustering or local feature extraction.

4.5 Spearman's Correlation Coefficient Evaluation

Spearman's rank correlation coefficient is a statistical measure used to assess the strength and direction of the monotonic relationship between two sets of data. Unlike the Pearson correlation coefficient, which measures the strength and direction of a linear relationship between two continuous variables, Spearman correlation is a non-parametric measure that works with both continuous and ordinal data. In this project, we plan to use the output score of each target word compared with the golden score, which is derived from the annotated data in a Word Usage Graph (WUG) generated through a comprehensive annotation and clustering process, to evaluate the performance of our trained model.

4.6 COMPARE Score Evaluation

COMPARE score is defined as the average of human semantic proximity judgments of usage pairs for certain words between corpus1 (*C1*) and corpus2 (*C2*) to predict the negated diachronic usage relatedness, where each pair consists of a use from *C1* and a use from the *C2* group). In our experiment, COMPARE directly measures the relatedness between the *C1* and the *C2* group and corresponds to the target word's mean in the COMPARE group (measuring the degree of semantic change).

5. Data

Our language corpus was assembled into two distinct corpora, each representing different time spans: the first spanning from 1810 to 1906 (referred to as the old corpus, *C1*), and the second covering the years 1994 to 2020 (referred to as the modern corpus, *C2*) (refer to Table 1 for details). The old corpus was compiled from various freely available sources obtained from Project Gutenberg, while the modern corpus was constructed using datasets from the OPUS project (Tiedemann, 2012). The old corpus involved concatenating all collected sources, while the modern corpus comprised datasets from TED2013, News-Commentary v16, MultiUN, and the Europarl corpus. TED2013 was utilized in its entirety, while 50 snippets, each consisting of 5000 lines, were randomly selected from the other datasets. We utilized spaCy to parse both corpora, resulting in four versions for each: the raw text, tokenized text, lemmatized text, and text annotated with part-of-speech tags.

Corpus	Time period	Tokens
Old Corpus (*C1*)	1810-1906	Around 13M
Modern Corpus (*C2*)	1994-2020	Around 22M

Table 1. Time periods and sizes of two corpora

After the preparation of the corpus, the next step was target word selection, which identified specific words within a corpus that were of particular interest for analysis. These words were chosen based on their potential to exhibit semantic changes over time. To study the semantic evolution of target words, we sampled instances of their usage from both old and modern corpora. These usage samples were then annotated, involving the assignment of labels or judgments by human annotators. This annotation process aimed to capture nuances in the meaning of the selected words across different periods.
LSCD provided the selected target words and their annotated usage samples. This presentation format allowed us to engage with the material and apply their computational models to detect and quantify semantic changes in the chosen words across the specified periods.

6. Models

Our models are outlined by two studies on Codalab, including Lexical Semantic Change Discovery in Spanish and RuShiftEval, and evaluation baselines are defined respectively to evaluate semantic change between corpus. Baseline models work for two functions; the first function is setting thresholds for participants in the evaluation phase. Second, baseline methods are functional training models for other researchers, such as our experiment, adapting baselines and data to understand how semantic change evaluation works.

6.1 SGNS+OP+CD Method in LSCDiscovey

One of the models we chose from LSCDiscovery is Skip-Gram with Negative Sampling + Orthogonal Procrustes + Cosine Distance (SGNS+OP+CD). This approach interprets word vector representations in input corpora.

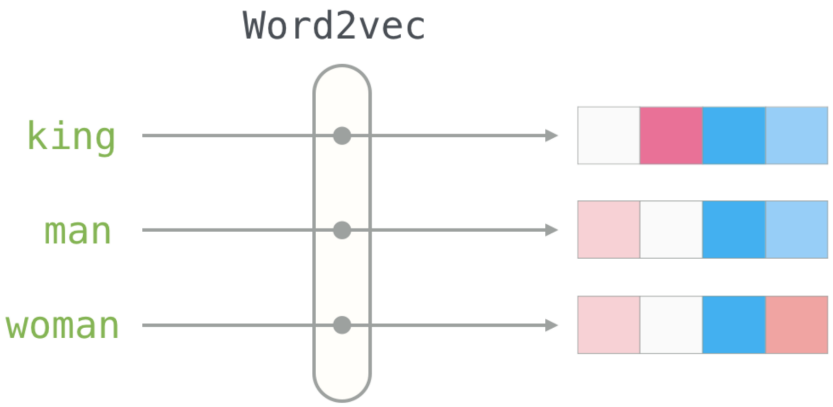


Figure 1. Word2Vec example cartoon in ‘king’, ‘man’, and ‘woman’ in English

Our model started by word embedding, such as Word2Vec(Figure 2). Word2Vec was to produce a vector space (Figure 2) to train a computer to understand the meaning of words. SGNS learns vectors, which are representations for each word in two corpora, by using a simple neural network model[1] producing with each word vector in the high-dimensional space. Aiming to compare the semantic change of the language, we only kept the words that were common in both corpora (both commonly used in the past and now) and aligned the vector representations trained in SGNS. The objective of SGNS is to maximize the probability of correctly predicting context words given a target word or vice versa.

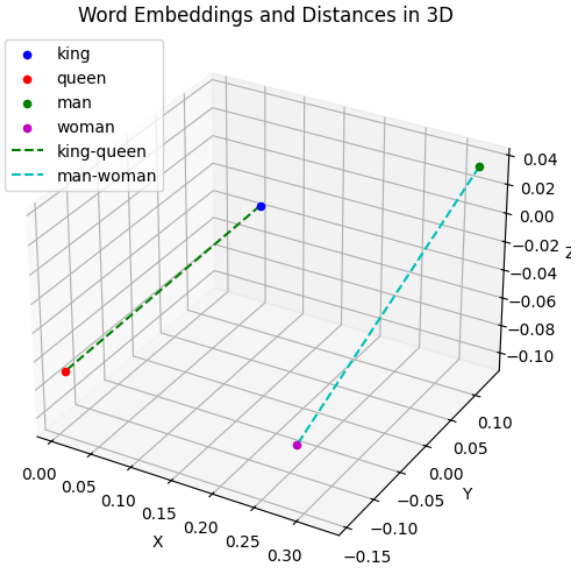


Figure 2. First of all, we tried to plot the king, queen, man, and woman-related positions in 3D space as a practice to understand how word2vec works. The 3D plot visually represents the word embeddings of "king," "queen," "man," and "woman" and highlights the relationships between these words in the semantic space, particularly focusing on gender relationships.

The following vector transformation, Orthogonal Procrustes, seeks to find a rotation or reflection on vectors that match these vectors in a space that makes them comparable. We initialize vectors

[1] As parameters we chose dim=100, window size=10, epochs=5, number of negative samples=5, subsampling threshold=0.001 (cf. Kaiser et al., 2020a).

by collecting the original vectors representing words in both corpora. Subsequently, subtracting the mean vector from each set of vectors ensures that the vectors are centered around the origin. In decomposing one set of vectors into three matrices, singular value decomposition (SVD) is the ideal model.

Besides cosine distance measurement, another scope we are going to focus on is the relationship between words by cosine similarity.

6.2 Local Neighborhood Approach

6.2.1 Cosine Similarity

Cosine similarity, on the other hand, is a specific metric used to measure the similarity between two vectors. In the context of natural language processing, cosine similarity is frequently employed to compare the similarity of two documents or word vectors. The cosine similarity between two vectors, A and B, is calculated as the cosine of the angle between them.

6.2.2 Local Neighborhood Method in LSCDiscovery

Our corpora were pre-trained by word2vec embeddings on measuring meaning change between centuries. At the word2vec stage, a skip-gram with a negative sampling (SGNS) algorithm was applied to vectors. Local Neighborhood measurement is based on the relevance of a word's nearest semantic neighbors. This measurement is firstly getting the target word *w's* set of nearest k nearest neighbors. According to cosine-similarity, the same word *w* will be extracted from different time periods (*t*, *t+1*) corpus and measured the vector change of the word separately between *t* and *t+1*(Figure 3). Finally, measure the extent of local neighborhood change with cosine distance, analyzing with the gold number to produce the Spearman correlation.

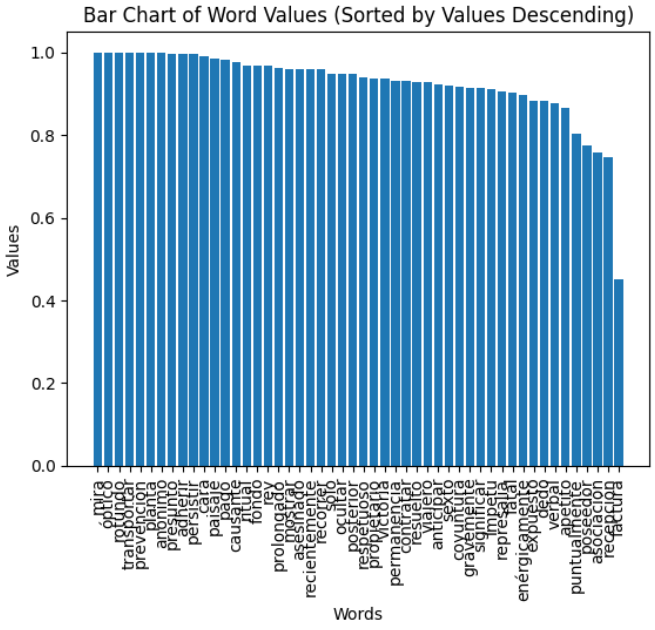


Figure 3. Bar chart of word values.

The bar chart (Figure 3) above shows the result of Local Neighborhood from LSCD Spanish Corpus. The histogram plot illustrates the distribution of cosine similarity values derived from the application of the local neighborhood method on the LSCDiscovery Spanish corpus from two eras.

The data frame presents word pairs with their corresponding cosine similarity values from our local neighborhood model, reflecting the semantic changes between time periods mentioned above. The height of each bar indicates the frequency of word pairs falling within that range. The bars are shaded in a gradient, emphasizing the varying degrees of local neighborhood changes.

7. Findings

- **RO1: Detect and quantify lexical semantic changes in Spanish.**
  Our approach, combining Skip-gram with Negative Sampling (SGNS), Orthogonal Procrustes (OP), and Cosine Distance (CD), proved effective in detecting and quantifying lexical semantic changes in Spanish. The SGNS model efficiently learned vector representations for words in both old and modern corpora, overcoming computational challenges associated with large vocabularies. The subsequent application of Orthogonal Procrustes facilitated the alignment of these representations, enabling the measurement of graded changes through cosine distance.

  The method successfully identified shifts in the meanings of target words by comparing their vector representations in the two datasets. The use of SGNS+OP+CD provided a robust framework for capturing subtle semantic nuances, offering a comprehensive understanding of how word meanings have evolved over time in the Spanish language.

- **RO2: Measure the evolution of Spanish word meanings**
  To measure the evolution of Spanish word meanings, we employed Spearman's Correlation Coefficient and COMPARE scores. Spearman's Correlation Coefficient allowed us to assess the strength and direction of the monotonic relationship between the computed scores of target words and their golden scores. This non-parametric measure proved effective in capturing the consistency of our model's predictions with the expected outcomes.

  The COMPARE score, which predicts the negated diachronic usage relatedness, provided valuable insights into the degree of semantic change. By directly measuring the relatedness between the old and modern groups of language usage, the COMPARE score facilitated a nuanced evaluation of how words have evolved in meaning over time. The inclusion of a mixed COMPARE group added depth to our analysis, considering pairs consisting of uses from both time periods.

  In summary, our evaluation methods, spearheaded by Spearman's Correlation Coefficient of 0.543 and COMPARE scores of 0.561 compared with the provided golden number (Table 2), contributed to a robust measurement of the evolution of Spanish word meanings. These findings provide a foundation for understanding the intricate semantic changes that have occurred in the Spanish language, offering valuable insights for applications in digital journalism, chatbots, and other information access systems. Moreover, our approach's adaptability to different language datasets positions it as a promising tool for exploring semantic changes across diverse linguistic domains.

Model	Change graded SPR(Spearman)	COMPARE SPR
SGNS + OP + CD	0.543	0.561

Table 2. Spearman rank coefficient correlation analysis on Change graded and COMPARE

8. Conclusion

We developed a model that efficiently detects and quantifies lexical semantic changes in Spanish by utilizing data mining techniques. This automated and comprehensive solution aims to enhance our understanding of the evolution of Spanish lexical semantics over time, addressing the limitations of traditional manual methods in scalability and objectivity. This research endeavors to track and measure the evolution of Spanish by identifying shifts in word meanings, offering potential applications in information access systems like digital journalism and chatbots, while also tackling the challenge of quantifying semantic changes in a language with rich variation. Research objectives set out to provide a holistic measure of the evolutionary trajectories of Spanish words, encapsulating the intricate journey of semantic modifications each term has undergone. Our multifaceted approach involves the utilization of methods to capture the semantic nuances, thereby facilitating a comprehensive analysis of the ever-evolving semantic landscape within the Spanish linguistic domain.

9. Future Study

Applying Skip-Gram Negative Sampling with Orthogonal Procrust adding on Cosine Distance as our first lexical semantic change detection model and Local Neighborhood as our second one, is experimental research to understand how to evaluate semantic change in words in practice; also, comparing the two model's performance on a language, here we chose Spanish. This research has provided changes in the meaning of words between old and modern eras. However, there might be several areas that could be explored in future studies.

From the language model perspective, there are two features to be the main implements as research topics.

9.1 Cross-Linguistic Comparative Studies

Future studies could extend the current research by conducting cross-linguistic comparative analyses. Exploring the effectiveness of the SGNS+OP+CD approach in detecting and quantifying lexical semantic changes in languages other than Spanish would contribute to a broader understanding of the method's applicability and potential variations across linguistic contexts.

For example, one of the implementations will choose a diverse set of languages with distinct linguistic characteristics. For example, consider English, Chinese, and Arabic, representing different language families and writing systems. Researchers can try to obtain diachronic corpora for each selected language, spanning significant historical periods. These corpora should capture changes in language usage over time.

9.2 Fine-Tuning for Specific Domains

Considering the adaptability of the SGNS+OP+CD approach, future studies might explore fine-tuning the methodology for specific domains or industries. Customizing the model to capture domain-specific semantic changes could enhance its applicability in fields such as legal documents, scientific literature, or specialized technical languages.

As for the diverse applications of semantic change detection, Language is not expressed only in text but audio and visual ways. In the future, the diachronic language data is going to be collected into audio or video streams and applied to machine learning models. On the other hand, our result has the ability to benefit the information retrieval system and evaluate the performance of language in the IR system.

9.3 Multimodal Semantic Change Detection

Considering the increasing availability of multimodal data, future research could explore the extension of semantic change detection to include visual and auditory modalities. Investigating how well the SGNS+OP+CD approach adapts to analyzing changes in meaning across multiple modalities would align with the evolving nature of contemporary data.

9.4 Implementation in Information Retrieval Systems

Evaluating the SGNS+OP+CD approach within the context of information retrieval systems could be a promising avenue. Assessing its effectiveness in enhancing search algorithms or information access systems by incorporating temporal semantic changes would have practical implications for digital journalism, chatbots, and other domains reliant on accurate language understanding.

10. Reference

[1] Frank D. Zamora-Reina, Felipe Bravo-Marquez, and Dominik Schlechtweg. 2022. LSCDiscovery: A shared task on semantic change discovery and detection in Spanish. In *Proceedings of the 3rd Workshop on Computational Approaches to Historical Language Change*, pages 149–164, Dublin, Ireland. Association for Computational Linguistics.

[2] Dominik Schlechtweg, Sabine Schulte im Walde, and Stefanie Eckmann. 2018. Diachronic Usage Relatedness (DURel): A Framework for the Annotation of Lexical Semantic Change. In Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 2 (Short Papers), pages 169–174, New Orleans, Louisiana. Association for Computational Linguistics.

[3] Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013a. Efficient estimation of word representations in vector space. In 1st International Conference on Learning Representations, ICLR 2013, Scottsdale, Arizona, USA, May 2-4, 2013, Workshop Track Proceedings.

[4] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013b. Distributed representations of words and phrases and their compositionality. In Proceedings of NIPS.

[5] Dominik Schlechtweg, Barbara McGillivray, Simon Hengchen, Haim Dubossarsky, and Nina Tahmasebi. 2020. SemEval-2020 Task 1: Unsupervised Lexical Semantic Change Detection. In

Proceedings of the Fourteenth Workshop on Semantic Evaluation, pages 1–23, Barcelona (online). International Committee for Computational Linguistics.

[6] 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)*, pages 1489–1501, Berlin, Germany. Association for Computational Linguistics.

[7] 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*, pages 2116–2121, Austin, Texas. Association for Computational Linguistics.

[8]Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.378.

[9]Matthew Honnibal, Ines Montani, Sofie Van Landeghem, and Adriane Boyd. 2020. spaCy: Industrial strength Natural Language Processing in Python.

[10]Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013b. Distributed representations of words and phrases and their compositionality. In *Proceedings of NIPS*.

Appendix

Link to the Semantic Change Score Excel: submission_LSCD

1. SGNS + OP + CD training results(selected top 30 rows as demo)

word	change_binary	change_graded	COMPARE
amenazado	0	0.4256628787	0.4256628787
molestar	0	0.3791331384	0.3791331384
huérfano	0	0.4631171247	0.4631171247
cola	0	0.4457427175	0.4457427175
local	0	0.5445406595	0.5445406595
formar	0	0.4021312713	0.4021312713
incierto	0	0.4506865398	0.4506865398
fuerza	0	0.377696036	0.377696036
instalar	0	0.4092893857	0.4092893857
número	0	0.3393246843	0.3393246843
despegar	1	0.6871293565	0.6871293565
agregado	1	0.836275899	0.836275899
aprobación	0	0.5268659978	0.5268659978
desconfianza	0	0.3822196125	0.3822196125
acento	1	0.5820776879	0.5820776879
atentar	0	0.546473447	0.546473447
consumado	1	0.5784434675	0.5784434675
modo	0	0.3400168414	0.3400168414
gracias	0	0.4114686511	0.4114686511
sensato	0	0.3480535052	0.3480535052
conquista	0	0.5280725221	0.5280725221
calificado	1	0.5933490118	0.5933490118
escaño	1	0.6000062121	0.6000062121
atentamente	1	0.6421759131	0.6421759131
tarjeta	0	0.3676950385	0.3676950385
culpa	0	0.3762681922	0.3762681922
arrogancia	0	0.3926167548	0.3926167548
mayo	0	0.3553640101	0.3553640101
colección	0	0.2874120658	0.2874120658