

# Towards a Conversational Web? A Benchmark for Analysing Semantic Change with Conversational Bots and Linked Open Data

**Florentina Armaselu**

University of Luxembourg, Luxembourg, florentina.armaselu@uni.lu

**Christian Chiarcos**

Goethe Universität Frankfurt, Germany, chiarcos@cs.uni-frankfurt.de

**Barbara McGillivray**

King's College London, United Kingdom, barbara.mcgillivray@kcl.ac.uk

**Anas Fahad Khan**

Istituto di Linguistica Computazionale  
'A. Zampolli', Italy  
fahad.khan@ilc.cnr.it

**Ciprian-Octavian Truică**

University Politehnica of Bucharest,  
Romania  
ciprian.truica@upb.ro

**Giedrė Valūnaitė-Oleškevičienė**

Mykolas Romeris University,  
Lithuania  
gvalunaite@mruni.eu

**Chaya Liebeskind**

Jerusalem College of Technology,  
Israel  
liebchaya@gmail.com

**Elena-Simona Apostol**

University Politehnica of Bucharest,  
Romania  
elena.apostol@upb.ro

**Andrius Utkā**

Vytautas Magnus University,  
Lithuania  
andrius.utka@vdu.lt

## Abstract

The paper presents preliminary results from our experiments with large language models, linked data, and semantic change in multilingual diachronic contexts. It proposes the first steps towards a benchmark and aims at fostering discussion on the concept of conversational knowledge bots as emerging paradigms, and the use of linked open data in linguistic tasks.

## 1 Introduction

Developments in large language models (LLM) such as GPT-3, BLOOM and GPT-4 (Brown et al., 2020; Workshop BigScience, 2022; OpenAI, 2023) have drawn attention to the capabilities of deep learning technologies to support conversations between human and artificial agents using natural language. These types of conversation, spanning from question-answering to code generation, seem to indicate an emergent paradigm shift from current graphic- and keyword-based human-

computer interaction and search modes to a conversational way of interacting with machines and the World Wide Web. Although conversational agents such as ChatGPT and BLOOM have shown remarkable capabilities in generating human-like responses and ability to analyse and synthesise correct answers, the currently available versions may suffer from a few limitations, such as hallucinations, self-contradicting statements, or outdated information (Ji et al., 2023; Mündler et al., 2023).

The question that arises is, therefore, to what extent will this way of interacting affect present formalisms and concepts, in particular those related to the Semantic Web? Will the processing of large amounts of unstructured text and the availability of pre-trained language models with conversational abilities have an impact on the use of more structured forms of representing and accessing knowledge by means of vocabularies such as the Resource Description Framework (RDF), Web Ontology Language (OWL), Linked Open Data (LOD) or OntoLex? How might these two

paradigms influence each other and what possible forms of combining them might be imagined for applications in areas of research such as linguistics, data science and digital humanities?

Rather than providing direct answers to these questions, the aim of this paper is to discuss potential scenarios built on a use case that combines natural language processing (NLP) and linguistic linked open data (LLOD) to analyse semantic change in multilingual diachronic corpora. Sections 2 and 3 present related work and preliminary results from our experiments with ChatGPT (Brown et al., 2020), Bing (Mehdi, 2023), word2vec (Mikolov et al., 2013; Rehurek and Sojka, 2010), and OntoLex-FrAC (Chiarcos et al., 2022). Section 4 formulates questions based on these first-round observations and proposes a benchmark related to the concept of conversational knowledge bots and their application to linguistic tasks. Section 5 summarises our findings.

## 2 Related work

Research on semantic change, the phenomenon concerned with the change in the meaning of a lexical unit (word or expression) or of a concept over time, has seen significant progress in the natural language processing community in recent years (Tahmasebi et al., 2018; Tsakalidis et al., 2019; Schlechtweg et al., 2020). While the majority of these studies focus on corpus-driven embedding models covering different time intervals, some studies, e.g., Armaselu et al. (2022), have advocated for the integration of such distributional approaches with linked open data. Recent advances have also been reported in the area of linguistic linked data (Cimiano et al., 2020; Khan et al., 2021; McGillivray et al., 2023), which promotes the use of graph-based models to represent linguistic data, and in building AI-based conversational agents, such as OpenAI’s ChatGPT (Chat Generative Pre-trained Transformer), Microsoft Bing, and Google’s Bard. Studies on LLMs have drawn attention to both potential benefits and concerns (Maynez et al., 2020; Shuster et al., 2021; Talat et al., 2022; DIGHUM, 2023), to their ability to be trained on code, use external APIs (Chen et al., 2021; Schick et al., 2023) and integrate plugins.<sup>1</sup> However, to our knowledge, there have not been any enquiries on the opportunities and chal-

lenges of combining LLMs and LLOD in semantic change-related tasks. Given the trends in artificial intelligence (AI) possibly leading to a conversational Web paradigm, these forms of interaction and their impact should be considered within the linked data community. We will illustrate this point through examples from a use case in diachronic analysis.

## 3 Preliminary results

In this section, we describe preliminary results from semantic change experiments with ChatGPT and Bing, and word embedding methods.

### 3.1 ChatGPT and Bing

The experiments involved the use of the OpenAI platform that allows for interactions with the GPT conversational agent via a user account. At this stage, we carried out qualitative evaluations of GPT’s responses. Our first aim was to assess to what extent the agent can trace the semantic evolution of certain words, such as ‘revolution’, across several periods and languages or provide citations when explicitly asked to do so (Appendix A, Table 2).

Our second aim was to evaluate the agent’s ability to generate code based on specific relations, such as the sense of a word and its attestation, using OntoLex-Lemon (McCrae et al., 2017) and OntoLex(-FrAC) (Chiarcos et al., 2021, 2022), a model intended to complement its predecessor with a vocabulary necessary to represent information derived from corpora. The first results (Table 1) indicated that GPT seemed quite good at generating OntoLex. On the other hand, its responses related to OntoLex-FrAC, although sounding meaningful, were incorrect. This seemed to be due to a lack of training data using this formalism. This assumption was supported by GPT’s initial misreading of the acronym: *OntoLex-FrAC stands for OntoLex Lemon Extension for Frame-based Annotation of Corpora* (conversation 1); *OntoLex-FrAC (Ontology-Lexicon for Fine-Grained Annotations and Classification) is a model for representing fine-grained lexical and semantic information in natural language text* (conversation 7). Therefore, we tried to teach GPT about OntoLex-FrAC by asking it first to summarise a paper about the model (Chiarcos et al., 2022) by copying and pasting the full text. We tested what it learned by asking the ques-

<sup>1</sup><https://openai.com/blog/chatgpt-plugins>.

tion from Table 1. It proposed the same structure as `frac:attestation`, but with hallucinated vocabulary. In a third step, we provided the `le_falter_vi` example for attestation from the paper (p. 4020). The answer in Table 1 was obtained after asking the same question again in the fourth iteration.

---

**User:** Please express in OntoLex or OntoLex-FrAC the following sentence: The word 'révolution' in French, with the sense of 'mouvement circulaire effectué par un corps autour de son axe', was attested in 1727, according to Ortolang.

---

**Chat GPT:** Here is an example of how to express the sentence in OntoLex-FrAC:

```
css
:revolution_fr
  a ontolex:LexicalEntry ;
  rdfs:label "révolution"@fr ;
  ontolex:canonicalForm [
    a ontolex:Form ;
    ontolex:writtenRep "révolution"@fr
  ] ;
  ontolex:sense :revolution_fr_sense ;
  frac:attestation [
    a frac:Attestation ;
    rdf:value "révolution (sens de 'mouvement
circulaire effectué par un corps autour de
son axe') est attesté en 1727"@fr ;
    frac:attestationSource <https://www.ortolang.
fr/market/corpora/corpus-tcof>
  ] .
:revolution_fr_sense
  a ontolex:LexicalSense ;
  ontolex:definition [
    a ontolex:Definition ;
    rdfs:label "Définition de la révolution"@fr ;
    rdf:value "'Mouvement circulaire effectué par
un corps autour de son axe'"@fr
  ] .
```

---

Table 1: ChatGPT conversation 8 (23 March 2023). OntoLex-FrAC generation

To a certain extent, we managed to “teach” ChatGPT some OntoLex-FrAC, but corpus and date properties were missing, so ChatGPT made up something on its own. It still hallucinated a lot, influenced by the dialogue history. Remarkably, it learned that an attestation must contain an attestation string, and it made that up from the definition, so, with some more examples, it can produce valid OntoLex-FrAC. However, it failed to create a separate date property. Other GPT tests included similar conversations for words such as Lithuanian *ponas* (‘sir, lord’) and its equivalents in the other languages (conversation 10). Bing also misread the OntoLex-FrAC acronym. While correctly rendering OntoLex properties such as `ontolex:canonicalForm` and `ontolex:sense`, it included non-existing OntoLex-FrAC properties, e.g.,

`ontolexfrac:dataSource` and `ontolexfrac: dateOfAttestation` (Bing, conversation 1). Another aspect of the assessment referred to sources. For instance, when asked about the sources or methods used, the degree of detail of the GPT responses varied: from generic statements, *As an AI language model, I was trained on a large corpus of text data* (conversation 1); to recommendations, *I can suggest some resources [...]: National Library of Luxembourg [...], Corpus de Français Parlé à Bruxelles* (conversation 5); or to procedure descriptions, *In this example, we create a lexical entry [...] we include an attestation using the Frac vocabulary* (conversation 8).

### 3.2 Diachronic word embeddings

We compared the conversation results with the outcomes of our diachronic word embedding and LLOD modelling experiments using multilingual datasets (Appendix B, Table 3, 4). We trained standard word embedding techniques, such as word2vec (Mikolov et al., 2013; Rehurek and Sojka, 2010) and fastText (Bojanowski et al., 2017) on the datasets divided into time slices corresponding to centuries (LatinISE, Responsa) or smaller event-driven intervals (BnL Open Data). We extracted the neighbours of the target words in the different time slices via cosine similarity, following standard practice in semantic change detection. The goal was to query the models for similar terms expressing social, economic, cultural or historic facts, and compare them across several languages. We noted that whereas the time slice granularity of the order of centuries may point to meanings changing, emerging or fading out (LatinISE, SLIEKKAS, Responsa), the finer granularity seems to highlight polysemous usage in various contexts with no clear indication when a certain meaning has emerged or went out of use (BnL Open Data). In this respect, a combination of corpus- and dictionary-based knowledge may lead to richer contextual representations of semantic change.

## 4 Discussion

Section 3 experiments have shown that conversational agents such as GPT can provide information about the meanings of certain words or concepts and their evolution over time and across languages. However, to understand the mechanisms

that generated these changes, a deeper analysis of the sources providing evidence about them would be needed.

Metzler et al. (2021) consider that although state-of-the-art pre-trained language models are able to generate prose in response to an information need, they “do not have a true understanding of the world, they are prone to hallucinating, and crucially they are incapable of justifying their utterances by referring to supporting documents in the corpus they were trained over” (p. 2). In contrast, the models of the future should be able to leverage the “meta-information associated with documents like provenance, authorship, authoritativeness”, support “cross-lingual generalization”, integrate new data through “online” or “incremental” learning, and provide answers with a degree of detail close to those of a domain expert (pp. 2, 15, 16).

#### 4.1 LLOD aggregation

Before considering the different types of knowledge agents that may assist our task in the future, we will get back to our example of diachronic analysis. For instance, the uses and meanings of the French word *révolution* in a certain country would need to be informed by knowledge representations combining corpora and dictionaries to study the term occurrences in time and space and compare them against existing attestation evidence. Listing 1 shows an example of lexical entry for *révolution* and its attestation that we created using elements from the OntoLex-FrAC model (Chiarcos et al., 2021, 2022).

Listing 1: OntoLex-FrAC modelling example

```

:rev-fr_le_1 a ontolex:LexicalEntry ;
  ontolex:canonicalForm [
    ontolex:writtenRep "révolution"@fr ] ;
  ontolex:sense :rev-fr_s_1 .
:rev-fr_s_1 a ontolex:LexicalSense ;
  frac:attestation [
    a frac:Attestation ;
    frac_new:dictionary [
      dc:source
        <http://example.org/ortolang/révolution>;
      dc:definition
        "Mec. Mouvement circulaire...";
      dc:date "1727"^^xsd:gYear ] ;
    frac:corpus [
      dc:source
        <http://example.org/ark:70795/dqgfr3/
          pages/17/articles/DTL612>;
      dc:date "1789"^^xsd:gYear ;
      dc:title "L'art de conduire et régler
        les pendules et les montres";
      dc:publisher "A Luxembourg, Chez la Veuve
        de J. B. Kleber, Imprimeur de Sa Majesté";
      frac:quotation "La roue ...
        fait une révolution par heure ...";
      prov:agent [
        a prov:Organization ;
        foaf:name

```

```

        "National Library of Luxembourg";
      ] ;
    ] ;
  ] ;
frac:embedding [
  a frac:FixedSizeVector ;
  dc:extent "100"^^xsd:int ;
  dc:description "word2vec";
  rdf:value "[moyene, engrennat, tige ...]";
] .

```

We propose an extension of this formalism to include attestation both from dictionaries (provisionally marked by `frac_new:dictionary`) and corpora, by specifying as well the provenance and method used to obtain the corpus-based evidence. The `dc:source` identifies the dictionary entry and the document containing the corpus citation, while the `dc:date` refers to the attestation of the sense in the dictionary and the publication date of the corpus document. Complementary information may be added, such as title, publisher, author, etymology and translation relations, degree of certainty, agent identification, etc. While not all these categories of information can be available for the processed sources (especially, those from ancient times may be less complete or certain), this type of structured aggregation may provide more context and ground for possible inferences on the circulation of knowledge and the meaning of a term and its evolution across space, time, languages and cultures.

#### 4.2 Knowledge bots

Therefore, we imagine different forms of knowledge agents, from bots that provide outlines and connections between various themes, such as ChatGPT, to specialised agents able to focus on particular tasks and resources and return well documented responses. These responses can vary from answers to general questions, recommendations for reading or relevant resources, to dedicated search and processing of target datasets, code generation, and expert advice on a given topic. Such agents may also be taught to produce correct LLOD representations. This might lower the entry barrier for data providers, since the conversion can be automatised via GPT-like engines. For consumers, it may also lower the entry barrier, since it can help to explain turtle code in human language. In either way, it is not a substitute for having OntoLex/RDF data in the first place, but a complementary technology. LLMs lack semantic transparency and verifiability, and this is what LLOD can provide.

While transparency, interoperability, connectiv-



ity, unique identification, and ontological precision are chief assets of the Semantic Web technologies, the advances in AI-based unstructured data processing and content generation would probably imply changes in the way we create and interact with structured data on the Web. From this perspective, a series of questions should be addressed, such as: (1) What forms of knowledge agents can be foreseen to combine conversational abilities in natural language with search, processing and automatic generation of structured data in formats such as RDF, OWL and LLOD? (2) What is the role of the human agent and what types of task, interaction scenarios and potential threats can be envisaged within the human-bot interrelations? (3) How may the current Semantic Web formalisms evolve to accommodate these emerging modes of interaction and knowledge representation? (4) What new forms of collaboration between the LOD and NLP communities can be imagined to underpin the development of a conversational and more “content-aware” Web? To foster further discussion on these topics, we propose to create a shared repository of benchmarks related to combined LLM and LLOD scenarios within various use cases.

## 5 Conclusion and future work

We presented preliminary tests with language and linked data models in multilingual diachronic analysis. Taking into account the potential of AI-based agents, able of human-like conversations, and of an emerging conversational Web, we propose to create a benchmark repository shared within the (L)LOD community for use cases that combine conversational and linked data knowledge paradigms.

## Acknowledgment

This article is based upon work from COST Action *Nexus Linguarum*, *European network for Web-centred linguistic data science*, supported by COST (European Cooperation in Science and Technology). [www.cost.eu](http://www.cost.eu).

## Authors' contribution

FA wrote the manuscript, led the ChatGPT conversations 1-6, 9, 11, and contributed to the design of the semantic change experiments for French and LLOD modelling; CC led the ChatGPT conversations 7, 8, and contributed to Sections 3.1 and 4.2;

BM contributed to the revisions of the manuscript and to the design of the semantic change experiments, and provided the analysis of the Latin words; AFK contributed to the RDF modelling of the example in section 4.1 and proofreading; COT led the Bing conversation 1, and contributed to the revisions of the manuscript and Section 2; GVO contributed to the revisions of the manuscript and Section 3, and led the ChatGPT conversation 10; CL contributed to the revisions of the manuscript and Section 3, and led the ChatGPT conversation 12; ESA contributed to the revisions of the manuscript and Section 2; AU contributed to Section 1 and the overall revision of the manuscript. All authors reviewed the final manuscript.

## References

- Florentina Armaselu, Elena-Simona Apostol, Anas Fahad Khan, Chaya Liebeskind, Barbara McGillivray, Ciprian-Octavian Truică, Giedrė Valūnaitė-Oleškevičienė, and Marieke van Erp. 2022. [LL\(O\)D and NLP perspectives on semantic change for humanities research](#). *Semantic Web*, 13(6):1051–1080.
- Piotr Bojanowski, Edouard Grave, Armand Joulin, and Tomas Mikolov. 2017. [Enriching word vectors with subword information](#). *Transactions of the Association for Computational Linguistics*, 5:135–146.
- Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, Sandhini Agarwal, Ariel Herbert-Voss, Gretchen Krueger, Tom Henighan, Rewon Child, Aditya Ramesh, Daniel Ziegler, Jeffrey Wu, Clemens Winter, Chris Hesse, Mark Chen, Eric Sigler, Mateusz Litwin, Scott Gray, Benjamin Chess, Jack Clark, Christopher Berner, Sam McCandlish, Alec Radford, Ilya Sutskever, and Dario Amodei. 2020. [Language models are few-shot learners](#). In *34th Conference on Neural Information Processing Systems (NeurIPS 2020)*, Vancouver, Canada.
- Mark Chen, Jerry Tworek, Heewoo Jun, Qiming Yuan, Henrique Ponde de Oliveira Pinto, Jared Kaplan, Harri Edwards, Yuri Burda, Nicholas Joseph, Greg Brockman, Alex Ray, Raul Puri, Gretchen Krueger, Michael Petrov, Heidy Khlaaf, Girish Sastry, Pamela Mishkin, Brooke Chan, Scott Gray, Nick Ryder, Mikhail Pavlov, Alethea Power, Lukasz Kaiser, Mohammad Bavarian, Clemens Winter, Philippe Tillet, Felipe Petroski Such, Dave Cummings, Matthias Plappert, Fotios Chantzis, Elizabeth Barnes, Ariel Herbert-Voss, William Hebggen Guss, Alex Nichol, Alex Paino, Nikolas Tezak, Jie Tang, Igor Babuschkin, Suchir Balaji, Shantanu Jain, William Saunders, Christopher Hesse, Andrew N. Carr, Jan Leike, Josh Achiam, Vedant

- Misra, Evan Morikawa, Alec Radford, Matthew Knight, Miles Brundage, Mira Murati, Katie Mayer, Peter Welinder, Bob McGrew, Dario Amodei, Sam McCandlish, Ilya Sutskever, and Wojciech Zaremba. 2021. [Evaluating large language models trained on code](#). (arXiv:2107.03374). ArXiv:2107.03374 [cs].
- Christian Chiarcos, Elena-Simona Apostol, Besim Kabashi, and Ciprian-Octavian Truică. 2022. [Modelling frequency, attestation, and corpus-based information with OntoLex-FrAC](#). In *Proceedings of the 29th International Conference on Computational Linguistics*, pages 4018–4027.
- Christian Chiarcos, Thierry Declerck, and Maxim Ionov. 2021. [Embeddings for the lexicon: Modelling and representation](#). In *Workshop on Semantic Deep Learning*, page 1319.
- Philipp Cimiano, Christian Chiarcos, John P. McCrae, and Jorge Gracia. 2020. *Linguistic Linked Data. Representation, Generation and Applications*. Springer International Publishing.
- DIGHUM. 2023. [ChatGPT a catalyst for what kind of future? statement of the digital humanism initiative](#).
- Ziwei Ji, Nayeon Lee, Rita Frieske, Tiezheng Yu, Dan Su, Yan Xu, Etsuko Ishii, Ye Jin Bang, Andrea Madotto, and Pascale Fung. 2023. [Survey of hallucination in natural language generation](#). *ACM Computing Surveys*, 55(12):1–38.
- Anas Fahad Khan, Christian Chiarcos, Thierry Declerck, Daniela Gifu, Elena González-Blanco García, Jorge Gracia, Max Ionov, Penny Labropoulou, Francesco Mambrini, John McCrae, Émilie Pagé-Perron, Marco Passarotti, Salvador Ros, and Ciprian-Octavian Truică. 2021. [When linguistics meets web technologies. Recent advances in modelling linguistic linked open data](#).
- Joshua Maynez, Shashi Narayan, Bernd Bohnet, and Ryan McDonald. 2020. [On faithfulness and factuality in abstractive summarization](#). In *Annual Meeting of the Association for Computational Linguistics*, pages 1906–1919. ACL.
- John P McCrae, Julia Bosque-Gil, Jorge Gracia, and Paul Buitelaar. 2017. [The OntoLex-Lemon model: development and applications](#). In *Proceedings of eLex 2017 Conference*.
- Barbara McGillivray, Pierluigi Cassotti, Pierpaolo Basile, Davide Di Pierro, and Stefano Ferilli (in press). 2023. [Using graph databases for historical language data: Challenges and opportunities](#). In *Proceedings of the 19th Italian Research Conference on Digital Libraries, Bari, Italy, February 23-24, 2023*, CEUR Workshop Proceedings. CEUR-WS.org.
- Yusuf Mehdi. 2023. [Reinventing search with a new AI-powered microsoft bing and edge, your copilot for the web](#).
- Donald Metzler, Yi Tay, Dara Bahri, and Marc Najork. 2021. [Rethinking search: Making domain experts out of dilettantes](#). *SIGIR Forum*, 55(1):1–27.
- Tomas Mikolov, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. [Efficient estimation of word representations in vector space](#). pages 1–12.
- Niels Mündler, Jingxuan He, Slobodan Jenko, and Martin Vechev. 2023. [Self-contradictory hallucinations of large language models: Evaluation, detection and mitigation](#).
- OpenAI. 2023. *GPT-4 Technical Report*.
- Radim Rehurek and Petr Sojka. 2010. [Software framework for topic modelling with large corpora](#). In *LREC 2010 Workshop New Challenges for NLP Frameworks*, pages 45–50.
- Timo Schick, Jane Dwivedi-Yu, Roberto Dessì, Roberta Raileanu, Maria Lomeli, Luke Zettlemoyer, Nicola Cancedda, and Thomas Scialom. 2023. [Toolformer: Language models can teach themselves to use tools](#). (arXiv:2302.04761). ArXiv:2302.04761 [cs].
- Dominik Schlechtweg, Barbara McGillivray, Simon Hengchen, Haim Dubossarsky, and Nina Tahmasebi. 2020. [SemEval-2020 task 1: Unsupervised lexical semantic change detection](#). In *Workshop on Semantic Evaluation*, pages 1–23.
- Kurt Shuster, Spencer Poff, Moya Chen, Douwe Kiela, and Jason Weston. 2021. [Retrieval augmentation reduces hallucination in conversation](#). In *Findings of the Association for Computational Linguistics: EMNLP 2021*, pages 3784–3803.
- Nina Tahmasebi, L. Borin, and A. Jatowt. 2018. [Survey of computational approaches to lexical semantic change](#). *arXiv: Computation and Language*.
- Zeerak Talat, Aurélie Névéol, Stella Biderman, Miruna Clinciu, Manan Dey, Shayne Longpre, Sasha Luccioni, Maraim Masoud, Margaret Mitchell, Dragomir Radev, Shanya Sharma, Arjun Subramonian, Jaesung Tae, Samson Tan, Deepak Tunuguntla, and Oskar Van Der Wal. 2022. [You reap what you sow: On the challenges of bias evaluation under multilingual settings](#). In *BigScience Episode #5 – Workshop on Challenges & Perspectives in Creating Large Language Models*, pages 26–41, virtual+Dublin.
- Adam Tsakalidis, Marya Bazzi, Mihai Cucuringu, Pierpaolo Basile, and Barbara McGillivray. 2019. [Mining the UK Web Archive for semantic change detection](#). In *International Conference on Recent Advances in Natural Language Processing*, pages 1212–1221.
- Workshop BigScience. 2022. [BLOOM: A 176B-parameter open-access multilingual language model](#). (arXiv:2211.05100).

## Appendix A. ChatGPT

a) User	Chat GPT
What do you know about semantic change?	Semantic change is the process by which the meaning of a word or phrase changes over time. [...]
Can you trace the evolution of the word 'révolution' from 16 century to now?	[...] 16th century: The French word "révolution" first appeared to describe the motion of celestial bodies. [...] 17th century: The term began to be used [...] to describe any significant or sudden change [...] 18th century: [...] The term "révolution" came to be associated with the idea of overthrowing a despotic government or monarch [...]
Can you trace the evolution of the word "מהפכה" (mahapecha) from its first attestation to now?	[...] The earliest known use of the term "מהפכה" appears in the Hebrew Bible [...] the term refers to a sudden and dramatic reversal of fortunes [...]. In the medieval period, the term [...] was used to describe a military defeat or a natural disaster [...] During the late 19th and early 20th centuries, [...] the term [...] became associated with the idea of a Jewish national revolution [...]
Can you trace the evolution of the word "revoliucija" in Lithuanian [...]?	The Lithuanian word "revoliucija" first appeared in written texts in the 19th century, during the period of the Lithuanian National Revival [...] During the early 20th century, "revoliucija" continued to be used to describe political upheavals [...]
What is the equivalent of the word 'revolution' in Latin?	The equivalent [...] in Latin is "revolutio". It is derived from the verb "revolvere", which means "to roll back" or "to turn over". In Latin, "revolutio" was primarily used to describe the cyclic movement of celestial bodies [...]
Is there any equivalent of the word "επανástasi" (epanástasi) in Ancient Greek?	The Ancient Greek word that is most commonly used to refer to a "revolution" or "uprising" is "στάσις" (stasis). The word [...] is used in several works of Ancient Greek literature, including Thucydides' "History of the Peloponnesian War" and Aristotle's "Politics" [...] to describe the internal conflicts and power struggles [...]
b) User	Chat GPT
[...] please give me a citation from an earlier period that uses a related term, such as "revolte" or "revolucion".	[...] "Si comme la revolucion du ciel, qui jamais ne se peut arrester ne reposer, mais tous jours et toutes heures va, porte et conduit toutes choses." - Jean de Meun, Le Roman de la Rose, 1280-1290. This passage [...] uses the term "revolucion" in the sense of a continuous and unceasing movement or cycle [...]

Table 2: a) ChatGPT conversations 2, 6, 12 (11, 19, 28 March 2023). Semantic change over time and language; b) ChatGPT conversation 4 (18 March 2023). Citations

## Appendix B. Word embeddings

Dataset	Language	Time span	Size
LatinISE	Latin	2nd c. BCE - 20th c. CE	ca.10 mil. word tokens
Diorisis	Ancient Greek	8th c. BCE - 5th c. CE	10,206,421 word tokens
SLIEKKAS	Old Lithuanian	16th - 18th c.	10 texts, 350,000 words
BnL Open Data	French	1690 - 1918 (monographs)	15,966 documents, 6,514,941 word tokens
Responsa	Hebrew	11th - 21st c.	about 100 mil. word tokens

Table 3: Core datasets

LatinISE	SLIEKKAS	BnL Open Data	Responsa
450BCE-1BCE: <i>civitas</i> ('citizenship')	16th c.: <i>ponas</i> (rich person, title 'mister'; religious, 'lord', e.g., Jesus)	1690-1794: <i>révolution</i> (Mec. motion of a body around an axis)	11th-16th c.: מהפכה (revolution) (religious context, 'atheism', 'repentance')
1CE-450CE: <i>civitas</i> ('city')	18th c. <i>ponas</i> (rich person; independent person, 'master')	1831-1866: <i>révolution</i> (Geom. motion of a figure around an axis)	16th c.: מהפכה (frequency of the word declines)
451CE-900CE: <i>civitas</i> ('city')		1867-1889: <i>révolution</i> (Geol. natural phenomena)	17th-19th c.: מהפכה (context of war and tragedy)
		1890-1918: <i>révolution</i> (Pol. Hist. great political change)	20th c.-present: מהפכה (industrial, medical, ideological revolution)

Table 4: Word embedding results. Excerpts