



Using BERT to Explore Lexical Semantic Change of Prepositions

LIUDMILA RADCHANKAVA
(JOHANNES GUTENBERG UNIVERSITY MAINZ)


VASILY KONOVALOV
(AIRI, MIRAI)

Prepositions and the Challenge of Semantic Change

Why Study Prepositions? Prepositions are highly polysemous They combine lexical and grammatical functions Their semantic change is often gradual and difficult to detect

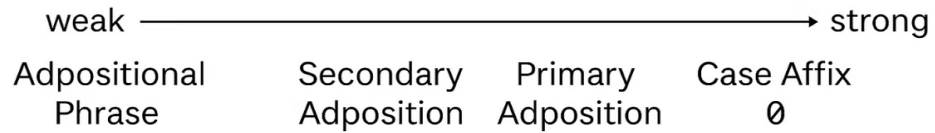


Challenges Limited attention to prepositions in NLP Word Sense Disambiguation (WSD) for prepositions is complex Diachronic semantic change is hard to measure automatically



Research Question Can contextualized BERT embeddings help: Disambiguate preposition senses? Detect diachronic semantic change?

Theoretical Background



(Lehmann 1986)

The instrumental/comitative suffix *-(y)lA* derives from the verb *birle-* 'to unite', passing through the stages *bile* and *ile* before becoming a fully integrated case suffix.

birle- → *bile* → *ile* → *-(y)lA*

(Koefoed & Marle 2004: 1583)

Syntactic-semantic continuum

From frozen expressions to transparent syntactic sequences

Integration levels vary

Peripheral elements less integrated than central ones

Functional load differs

Central elements carry greater functional weight

Frequency distinction

Central elements appear more frequently in usage

Lexical vs. Grammatical Status

(after Boye & Harder 2012)

Core Idea:

Grammatical expressions have ancillary (discursively secondary) status.

Lexical expressions can be discursively primary.

◆ Lexical (Primary)

Can carry assertive focus

Can be independently addressed

Semantically contentful

Example:

*He walked **behind** the house.*

→ *Behind* expresses spatial content and can be contrastively focused.

◆ Grammatical (Secondary)

Cannot carry independent focus

Cannot be directly questioned

Structurally dependent

Example:

*This will be the death **of** me.*

→ **Why (of)?* ✗

The preposition cannot be focused → grammatical status.

Method Approach

Workflow

BERT → Contextual embeddings → kNN
labeling → Distance to nearest neighbor →
Diachronic analysis

Contextualized embeddings from **BERT**

Preposition Sense Disambiguation via **kNN
classifier**

Distance to nearest labeled exemplar

Spearman correlation between distance and
time

Datasets

- Training Data

The Preposition Project (SemEval 2007) 34 English prepositions 16,557 training sentences 8,096 test sentences

Sense labels based on ODE

- Diachronic Data Corpus of Historical American English (COHA) 1820s-2010s 475+ million words Balanced by genre 2,000 sampled sentences per preposition

- Target Prepositions

Simple: *in, on, at, with, of*

Compound: *into, onto, beside, above, about*

Preposition Sense Disambiguation

- 1 BERT + kNN achieves 79.8% accuracy Comparable to previous supervised systems No external lexical resources used
- 2 Diachronic Trends Simple prepositions → weak or inconsistent correlation

Compound prepositions → clearer negative correlation Earlier usages often show larger distances Key Insight

Compound prepositions show more detectable contextual shifts over time than simple prepositions.

Interpretation & Limitations

➤ Interpretation

Distance \neq direct measure of new sense

It reflects **contextual divergence** from modern labeled senses

Captures gradual usage shifts and stylistic change

➤ Why Compound > Simple?

Compound prepositions are less grammaticalized

More discrete contextual shifts

Simple prepositions change more gradually

➤ Limitations

Sense inventory is fixed (SemEval 2007)

Distance influenced by frequency imbalance

Measures contextual shift, not lexicographic innovation

Conclusion

BERT embeddings effectively model preposition senses

kNN provides a simple and interpretable PSD approach

Distance-based proxy reveals diachronic contextual shifts

Compound prepositions show clearer detectable change

Take-Home Message

Contextual embeddings can capture not only sense distinctions, but also gradual patterns of language change.

References (Selected)

Boye & Harder (2012). *A Usage-Based Theory of Grammatical Status and Grammaticalization*

Devlin et al. (2019). *BERT: Pre-training of Deep Bidirectional Transformers*. NAACL.

Litkowski & Hargraves (2007). *SemEval-2007 Task 6: Word-Sense Disambiguation of Prepositions*.

Davies (2012). *Corpus of Historical American English (COHA)*.

Lehmann (2015). *Thoughts on Grammaticalization*.