

ZEITENWENDE IM WORTFELD

*Semantischer Wandel im Wortfeld Militär/Verteidigung in deutschen Tageszeitungen:
Eine korpuslinguistische Diskursanalyse*

In der Folge aktueller geopolitischer Entwicklungen, allen voran des russischen Angriffskriegs auf die Ukraine, scheint sich seit einigen Jahren ein grundlegender gesellschaftlicher Richtungswechsel hin zu einer zunehmenden Akzeptanz von militärischer Aufrüstung, Verteidigungsfähigkeit, aber auch von Bundeswehr und Wehrpflicht abzuzeichnen. Dies spiegelt sich unter anderem in linguistischen Phänomenen wider, die beispielsweise konkrete Bedeutungsverschiebungen von Begriffen wie *Drohne* oder *Zeitenwende* oder ein tendenziell positiveres Sentiment im Allgemeinen umfassen können. In diesem Projekt soll untersucht werden, ob und in welcher Form sich solche Werte- bzw. Bewertungswandel sprachlich objektiv belegen und quantitativ erfassen lassen. Dafür eignen sich KI-basierte Verfahren zur semantischen Modellierung, insbesondere sogenannte diachrone semantische Embeddings die es erlauben, zeitliche Veränderungen in Bedeutungen und Konnotationen von Begriffen innerhalb des Wortfeldes Militär in der deutschen Tagespresse zu messen. Der erwartete Gewinn des Projekts besteht in einer linguistischen Bestandsaufnahme und datengetriebenen Ergänzung qualitativer, diskursanalytischer Analysen zu der sich aktuell vollziehenden "Zeitenwende".

1. Fragestellung

Hat sich die konnotative Bedeutung von Wörtern wie *Bundeswehr*, *Rüstung*, *Wehrpflicht*, *Landesverteidigung* oder *Rüstungsindustrie* in der deutschen Tagespresse nach 2022 messbar ins Positive (im Sinne einer Sentiment-Analyse) verschoben? Welche anderen, möglicherweise nicht primär aus diesem Wortfeld stammenden Konzepte erfahren darüber hinaus einen semantischen Wandel? Lässt sich die diskursive „Zeitenwende“ als quantifizierbarer Shift im Embedding-Raum nachweisen – und gab es Vorboten bereits ab 2014 (Krimkrise)?

Dies beinhaltet u.a. folgende, konkrete Untersuchungsfragen:

1. Welche semantischen Nachbarn haben die Zielwörter in verschiedenen Zeitscheiben? Verschiebt sich *Bundeswehr* beispielsweise von einem Cluster mit *Skandal*, *kaputtgespart* und *Berateraffäre* hin zu *Wehrhaftigkeit*, *Sicherheit* und *Friedenssicherung*?
2. Bewegen sich Konzepte wie *Rüstungsindustrie* oder *Wehrpflicht* von negativ oder historisch hin zu positiv oder aktuell konnotierten Kontexten?
3. Verdichtet sich das gesamte Wortfeld nach 2022 als Folge einer aktiven politischen Diskussion, welche die Bedeutung der Wehrhaftigkeit angesichts einer veränderten Weltordnung reflektiert oder gar aktiv stärken will? Rücken die Lexeme semantisch enger zusammen?

2. Forschungskontext

Während *qualitative* Diskursanalysen zur Zeitenwende-Rede bereits existieren (vgl. Fan & Chen, 2024), bieten Methoden aus dem Bereich der natürlichen Sprachverarbeitung (NLP) die Möglichkeit zur *quantitativen* Erfassung zeitlicher Bedeutungsverschiebungen in Korpusda-

ten, z.B. mithilfe diachroner Word Embeddings (Hamilton et al., 2016; Kutuzov et al., 2018), einem distributionellen semantischen Verfahren, mit dem die zeitliche Entwicklung von Begriffsbedeutungen basierend auf ihrem sprachlichen Kontext innerhalb eines Textkorpus modelliert werden kann.

Die spezifische Kombination aus Embedding-basierter Analyse des semantischen Wandels im deutschen Sicherheitsdiskurs und einer großen Menge sprachlicher Daten scheint bislang nicht untersucht worden zu sein. Dieses Projekt liefert somit ein *quantitatives Fundament* zur qualitativen Diskursforschung, indem eine Embedding-basierte Messung semantischen Wandels systematisch auf ein deutsches Zeitungskorpus angewendet wird. Neben computer- bzw. korpuslinguistischen Erkenntnissen ergeben sich wertvolle Impulse für weitere Forschungsgebiete, darunter die (kritische) Diskursanalyse, die politikwissenschaftliche Betrachtung oder die medienwissenschaftliche Aufarbeitung.

3. Daten und Methoden

3.1 Korpus und Zielwörter

Als Datenbasis soll die E-Paper-Sammlung der DNB dienen, die mit ihrer Verfügbarkeit von Zeitungen aus dem Zeitraum von ca. 2010 bis 2025 wichtige Schlüsselereignisse wie die Krim-Annexion, den russischen Einmarsch in die Ukraine und die Wehrpflichtdebatte einschließt. Dafür soll eine repräsentative Auswahl überregionaler Leitmedien (SZ, FAZ, WELT, taz, Handelsblatt, ZEIT, Spiegel u.a.) sowie ca. 10 lokaler Titel aus verschiedenen Regionen getroffen werden.

Als Zielwörter eignen sich Begriffe, die im semantischen Feld Militär und Verteidigung verortet oder eng mit diesem verwoben sind, wie z.B. **Krieg, Verteidigung, Bundeswehr, Aufrüstung, Wehrpflicht, Drohnen**, die nach statistischen Kriterien ausgewählt werden.

3.2 Vorgehen

Nach der **Vorbereitung der Daten** (Extraktion und Bereinigung der Texte, Tokenisierung, Satzsegmentierung, Lemmatisierung usw.) sollen **diachrone Embeddings der Zielwörter und ihrer Kontexte** trainiert werden, jeweils für unterschiedliche zeitliche Epochen. Angestrebt wird zunächst eine Auflösung von einem Jahr; daraus ergeben sich sechzehn Zeitscheiben. Abhängig von den technischen Möglichkeiten (Rechenleistung/-dauer) und dem realen Erkenntnisgewinn sollen auch granularere Auflösungen getestet werden. Auf dieser Basis können verschiedene Verfahren zur qualitativen und statistischen Auswertung von Trends in der semantischen Entwicklung der Zielwörter angesetzt werden, um sowohl lexikalischen Wandel als auch Änderungen im Sentiment über die Zeit zu untersuchen. Zu den in Frage kommenden Methoden zählen die **Projektion der Modelle verschiedener Epochen in einem gemeinsamen Raum**, die Berechnung geeigneter **Distanzmaße** für die Zielwörter in verschiedenen Epochen, **Nearest-Neighbour-Analysen, Sentiment- und Konnotationsanalysen**, sowie entsprechende **Visualisierungen** der identifizierten Verschiebungen. Aufgrund des potentiell hohen öffentlichen Interesses wird neben der fachlichen Aufbereitung eine zugängliche **Darstellung der Ergebnisse in einer moderner, interaktiven und ästhetisch ansprechenden Online-Anwendung** angestrebt.

3.3 Aufwandsschätzung

Benötigt wird maschineller Zugriff auf deutsche Zeitungseditionen aus dem Zeitraum 2010-2025 und Hardware mit GPU-Kapazitäten zum Training der Embeddings. Die Programmierung erfolgt zum Großteil in Python. Je nach Granularität der Periodisierung (z.B. wöchentlich, monatlich, quartalsweise oder jährlich) und (Vor-)Verarbeitungsstand der Daten kann der anzusetzende Zeitaufwand schwanken. Die Textextraktion und das Modelltraining erfolgen vor Ort und können voraussichtlich in einem Zeitraum von 1,5-2 Wochen abgeschlossen werden. Die Entwicklung der Verarbeitungspipeline mit Testdaten und Nachbereitung (Interpretation und Publikation eines Papers in einschlägigen Konferenzen im Bereich Computerlinguistik oder Digital Humanities) erfolgen nicht vor Ort. Als Zeitaufwand für das gesamte Projekt werden drei Monate angesetzt.

4. Ergebnisse

Im Rahmen des Projekts werden **epochenspezifische Embedding-Modelle** für den Zeitraum 2010–2025 erstellt, die den semantischen Raum der untersuchten Zielwörter in verschiedenen Zeitscheiben abbilden. Auf dieser Grundlage werden **quantifizierte Shift-Maße** berechnet, die die semantische Verschiebung ausgewählter Begriffe aus dem Wortfeld Militär/Verteidigung über einen Zeitraum von fünfzehn Jahren messbar und vergleichbar machen. Ergänzend entstehen **Nearest-Neighbour-Tabellen**, die für jedes Zielwort und jede Epoche die jeweils nächsten semantischen Nachbarn dokumentieren, sowie **Trajektorien-Visualisierungen**, die den Verlauf der Bedeutungsverschiebungen im Embedding-Raum anschaulich darstellen. Die Ergebnisse werden in einer **interaktiven Online-Anwendung** aufbereitet, die eine **explorative Auseinandersetzung mit den Daten** ermöglicht.

Methodisch leistet das Projekt einen Beitrag zur Evaluation von Embedding-Shift-Verfahren auf kurzfristigem diskursivem Wandel – einem Szenario, das in der bisherigen Forschung, die sich vorwiegend auf langfristigen lexikalischen Wandel konzentriert, noch wenig systematisch untersucht wurde. Inhaltlich liefert das Projekt eine datenbasierte, empirisch fundierte Quantifizierung der sprachlichen „Zeitenwende“ im öffentlichen Diskurs und schafft damit ein quantitatives Fundament, das qualitative diskursanalytische, politikwissenschaftliche und medienwissenschaftliche Arbeiten ergänzen und stützen kann. Sämtliche im Projekt entwickelten Methoden, Modelle und Visualisierungskomponenten werden so dokumentiert und bereitgestellt, dass eine **Nachnutzung** für weitere Untersuchungen vergleichbarer diskursiver Wandelprozesse in anderen Themenfeldern und Korpora leicht möglich ist.

5. Literatur

Fan, X., & Chen, Q. (2024). *Legitimizing the shift in German security discourse: A discourse analysis of Scholz's "Zeitenwende" (historic turning point) speech*. SSRN.
<http://dx.doi.org/10.2139/ssrn.4954113>

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 (Volume 1: Long Papers)* (S. 1489–1501). Association for Computational Linguistics.

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* (S. 1384–1397). Association for Computational Linguistics.