

Chancen und Probleme im Umgang mit automatisch ermittelten Neologismenkandidaten im Projekt „Neuer Wortschatz“

Annette Klosa-Kückelhaus

Abstract

Seit 2015 stellt das Projekt „Methoden der Korpusanalyse und -erschließung“ dem Projekt „Neuer Wortschatz“ Listen mit Neologismenkandidaten zur Verfügung.¹ Diese werden im Projekt redaktionell ausgewertet, wobei viele Kandidaten als für das Wörterbuch irrelevant aussortiert werden, einige aber auch als im Wörterbuch beschreibenswert eingeschätzt werden. Hierbei wird unterschieden zwischen solchen Kandidaten, zu denen klassische Wortartikel im „Neologismenwörterbuch“² erarbeitet werden, und solchen, zu denen neue sogenannte Kurzartikel erstellt werden, die ab Ende 2019 das Onlineangebot des „Neologismenwörterbuches“ ergänzen werden.

In diesem Vortrag werden zunächst die verschiedenen Kategorisierungen vorgestellt, nach denen die Neologismenkandidaten sortiert werden (z.B. Eigenname, Flexionsform, semantisch durchsichtiges Lexem). Während diese Kategorien hinreichend scharf voneinander unterschieden sind, ist die Zuordnung eines Neologismuskandidaten zu zwei Kategorien nicht immer einfach, und zwar zur Kategorie der semantisch durchsichtigen Lexeme und zur Kategorie der Eigennamen. Deshalb werden in einem zweiten Teil des Vortrags solche Fälle anhand von Korpusmaterial diskutiert, bei denen die redaktionelle Entscheidung über die semantische Durchsichtigkeit eines Kandidaten oder über seinen Status als Eigenname schwer fällt.

Eine weitere Problemgruppe stellen solche Neologismenkandidaten dar, bei denen es bestimmte Auffälligkeiten im Zeitverlauf gibt. Beispielsweise gibt es Kandidaten, die bis zu einem bestimmten Zeitpunkt nur als Eigenname verwendet werden, später aber eine Ausweitung ihres Denotatsbereiches erfahren und daher als Kandidaten für das Wörterbuch infrage kämen. Außerdem gibt es Kandidaten, bei denen nur ein sehr kurzes Hoch in der zeitlichen Verteilung der Gebrauchshäufigkeiten festzustellen ist (z.B. von 2012-2015). Sollen diese dennoch ins Wörterbuch integriert werden?

Abschließend sollen einige Zahlen illustrieren, mit welchen Mengen wir es bei der Auswertung der Listen mit Neologismenkandidaten generell zu tun haben. Wie viele Kandidaten wiederholen sich beispielweise, wenn wir die jährlich erstellten Listen miteinander vergleichen? Wie viele Kandidaten haben wir bereits auf anderem Weg ermittelt? Wie viele Kandidaten haben wir aus den korpuslinguistisch erstellten Listen gewonnen? Insgesamt soll der Vortrag so einen Einblick in die großen Chancen, aber auch in die Probleme geben, mit welchen das Projekt „Neuer Wortschatz“ im Umgang mit automatisch ermittelten Neologismenkandidaten zu tun hat.

¹ Siehe im Einzelnen hierzu Klosa, Annette/Lüngen, Harald (2018): New German Words: Detection and Description. In: Čibej, Jaka/Gorjanc, Vojko/Kosem, Iztok/Krek, Simon (Hg.): Proceedings of the XVIII EURALEX International Congress. Lexicography in Global Contexts, 17-21 July, Ljubljana. Ljubljana: Znanstvena založba, 2018. S. 559-569.

² Online unter <https://www.owid.de/docs/neo/start.jsp>.