

Das kleine Wörterbuch der Redeeinleiter

Brunner, Annelen

brunner@ids-mannheim.de
Leibniz-Institut für Deutsche Sprache, Deutschland

Tu, Ngoc Duyen Tanja

tu@ids-mannheim.de
Leibniz-Institut für Deutsche Sprache, Deutschland

Weimer, Lukas

weimer@sub.uni-goettingen.de
Niedersächsische Staats- und Universitätsbibliothek
Göttingen, Deutschland
ORCID: 0000-0001-6919-3646

Das Poster¹ präsentiert eine Online-Ressource, die es erlaubt, 523 Redeeinleiter dynamisch zu durchsuchen. Sie bietet damit niedrigschwelligem Zugang zu empirischen Daten aus einem umfangreich manuell annotierten Korpus von fiktionalen und nicht-fiktionalen Texten, die zwischen 1850 und 1919 erschienen sind. Durch die Einbindung in eine große Forschungsinfrastruktur wird Zugänglichkeit garantiert und Nachnutzung vereinfacht.

Gegenstand

Redeeinleiter sind sprachliche Ausdrücke, die relativ zu einer direkten oder indirekten Rede- oder Gedankenwiedergabe in Voran-, Mittel- oder Nachstellung stehen und diese einleiten (Breslauer, 1996; Michel, 1966; Jäger, 1968). Auch wenn Verben wie *sagen* und *denken* sehr häufig sind, können Redeeinleiter auch kreative Ausdrücke sein und unterschiedliche Wortarten umfassen, z.B.

(1) *Der Untersuchungsrichter legte ihm ans Herz, daß, wenn er nicht angeben wolle, wo er den Einbruch verübt habe, sein Kopf sich schon als abgetan betrachten könne.* [aus: Hermann Sudermann: Miks Bumbullis (1917)]

(2) *»Kratzt der Alte einmal wieder«, brummte er, »und stört die ordentlichen Leute in ihrer Nachtruhe.«* [aus: Franz Grillparzer: Der arme Spielmann (1840)]

(3) *[...] daß ihnen aber nicht einmal immer das geliefert wird, was sie zu fordern ein Recht haben, bezeugen häufige Klagen von Auswanderern [...].* [aus: Deutsche Auswanderer-Zeitung, Unbekannte:r Autor:in (1852)]

Durch ihre Vielfältigkeit sind Redeeinleiter ein interessanter Untersuchungsgegenstand, sowohl aus linguistischer als auch literaturwissenschaftlicher Perspektive. Sie eignen sich, um dynamische Prozesse im lexikalischen Inventar der Sprache zu untersuchen (Tu, erscheint 2024) und spielen eine wichtige Rolle bei der Einbettung von Figurenrede in den narrativen Kontext (McHale 2014, Abschnitt 2).

Datengrundlage und Funktionalitäten

Grundlage für die Ressource ist das Kernkorpus „Redewiedergabe“ (RW-Korpus; Brunner et al., 2020a), welches im Rahmen eines DFG-Projekts entstand. Das Korpus umfasst ca. 49.000 Tokens und enthält Ausschnitte aus Erzählungen sowie Zeitungs- und Zeitschriftenartikeln aus dem Zeitraum 1850-1919. Das Textmaterial ist balanciert nach Dekaden sowie dem Merkmal fiktional vs. nicht-fiktional und wurde aufwendig manuell nach Formen von Rede-, Gedanken- und Schriftwiedergabe annotiert (Brunner et al., 2020b): Eine Konsensannotation wurde auf Basis zweier unabhängiger Annotationen erstellt. Zwar ist das RW-Korpus sowohl in TEI-konformem XML-Format als auch in einem spaltenbasierten Textformat vollständig frei verfügbar (<https://zenodo.org/records/3739239>), es bedarf jedoch technischer Kenntnisse, spezialisierte Informationen wie die über Redeeinleiter zu extrahieren. Die vorgestellte Ressource bietet hierfür einen bequemen und niedrigschwelligem Zugang.

Alle 3059 Einleiter-Vorkommen von direkter und indirekter Rede- oder Gedankenwiedergabe wurden mit ihren Attributen extrahiert und zu einer Häufigkeitsliste zusammengefasst. Für jeden der 523 Redeeinleiter-Typen bietet die Ressource einen Überblick über die Vorkommensverteilung nach den Dimensionen „Medium“ (Rede- oder Gedankenwiedergabe), „Wiedergabetyp“ (direkt oder indirekt), „Position“ (initial, medial oder final) und „Textsorte“ (fiktional oder nicht-fiktional). Abbildung 1 illustriert dies für den Einleiter *sagen*. Abbildung 2 gibt einen Überblick über die Attributverteilung des gesamten Inventars.

Abbildung 1: Vorkommensverteilung für den Redeeinleiter *sagen*.

Medium	Redewiedergabe:	535
	Gedankenwiedergabe:	1
Wiedergabetyp	direkt:	398
	indirekt:	138
Position	initial:	291
	medial:	123
	final:	122
Textsorte	fiktional:	435
	nicht-fiktional:	101

Für jede mögliche Attributkombination, mit der ein Redeeinleiter vorkommt, wurde ein zufällig gewählter Korpusbeleg extrahiert. Dieser besteht aus dem Satz, in dem der Redeeinleiter vorkommt, sowie dem vorangehenden und dem nachfolgenden Satz. Eine Verlinkung verknüpft ihn mit dem entsprechenden Dokument im RW-Korpus. Die Liste der Redeeinleiter kann zudem nach Attributwerten gefiltert werden und Redeeinleiter sowie die zugehörigen Be-

lege können als Excel- oder CSV-Tabellen exportiert werden.

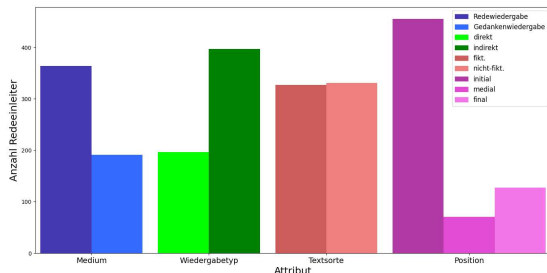


Abbildung 2: Verteilung des gesamten Einleiter-Inventars über die Attribute; gezählt wurden Redeeinleiter-Typen, wobei solche, für die mehrere Merkmale zutreffen, mehrfach gezählt wurden (z.B. bei Vorkommen in allen drei Positionen Zählung für jede der Positionen einmal).

Mit Hilfe der Ressource bekommt man nicht nur einen grundsätzlichen Überblick darüber, welche Einleiter in welchen quantitativen Verhältnissen verwendet werden, sondern sie erlaubt auch Kontrastierungen in unterschiedlichen Dimensionen. Die Filterfunktion ermöglicht zudem, gezielt solche Einleiter anzuzeigen, die in mehreren Kontexten vorkommen können (z.B. als Rede- und als Gedankeneinleitung). Neben den interessanten Beobachtungen, die die Ressource zu dem Datenbestand des RW-Korpus selbst erlaubt, bietet es sich an, diese mit Daten aus anderen Textgrundlagen zu vergleichen (z.B. moderne Literatur).

Zugänglichkeit und Nachhaltigkeit

Das „Kleine Wörterbuch der Redeeinleiter“ ist unter der Adresse <https://www.owid.de/plus/redereinleiter> über OWIDplus abrufbar. OWIDplus ist eine Plattform für multilinguale lexikalisch-lexikografische Daten, für quantitative lexikalische Auswertungen und für interaktive lexikalische Anwendungen und zudem Teil des Portfolios, welches das Leibniz-Institut für Deutsche Sprache in das NFDI-Konsortium Text+ einbringt. Als Teil des breiten Tool-, Kompetenz- und Datenportfolios von Text+ erhöhen sich Auffindbarkeit und Verbreitung (z. B. durch die Text+ Registry, Genêt et al., 2023) der Ressource, wodurch es einer breiteren Wissenschaftscommunity entlang der FAIR-Kriterien (Wilkinson et al., 2016) zur Verwendung und Nachnutzung bereitgestellt werden kann. Gleichzeitig sind durch die institutionelle Anbindung an das IDS als zertifiziertes Datenzentrum nachhaltige Zugänglichkeit und Archivierung gesichert.

Fußnoten

1. Contributor Roles: Annelen Brunner (Data Curation, Writing – original draft, Writing - review & editing), Ngoc Duyen Tanja Tu (Data Curation, Visualization,

Software, Writing – original draft, Writing - review & editing), Lukas Weimer (Data curation, Writing - review & editing)

Bibliographie

Breslauer, Christine. 1996. "Formen der Redewiedergabe im Deutschen und Italienischen." *Sammlung Groos* 60. Heidelberg: Groos.

Brunner, Annelen, Stefan Engelberg, Fotis Jannidis, Ngoc Duyen Tanja Tu und Lukas Weimer. 2020a. "Corpus REDEWIEDERGABE". In *Proceedings of the 12th International Conference on Language Resources and Evaluation (LREC)*, 803-12.

Brunner, Annelen, Lukas Weimer, Stefan Engelberg, Fotis Jannidis und Ngoc Duyen Tanja Tu. 2020b. "Annotationsrichtlinien des Projekts ‚Redewiedergabe. Eine literatur- und sprachwissenschaftliche Korpusanalyse‘". <https://zenodo.org/record/3547594>.

Genêt, Philippe, Tobias Gradl, Kilian Hensen, Christoph Kudella und Daniela Schulz. 2023. "F wie Registry. Die Text+ Registry als Hilfsmittel zur Auffindbarkeit von Ressourcen". In *FORGE23: Forschungsdaten in den Geisteswissenschaften – kritisch betrachtet*, 168-172.

Jäger, Siegfried. 1968. "Die Einleitungen indirekter Reden in der Zeitungssprache und in anderen Texten der deutschen Gegenwartssprache. Ein Diskussionsbeitrag". *Muttersprache* 78: 236-49.

McHale, Brian. 2014. "Speech Representation". In *The living handbook of narratology*, hg. Peter Hühn, John Pier, Wolf Schmid und Jörg Schönert, 812-824. Hamburg: Hamburg University Press. [http://hup.sub.uni-hamburg.de/lnh/index.php?title=Speech Representation&oldid=891](http://hup.sub.uni-hamburg.de/lnh/index.php?title=Speech+Representation&oldid=891).

Michel, Georg. 1966. "Sprachliche Bedingungen der Wortwahl. Eine Untersuchung an Ausdrücken der Redeeinführung (Erster Teil)". *Zeitschrift für Phonetik* 19: 103-129.

Tu, Ngoc Duyen Tanja. Erscheint 2024. "Eine korpuslinguistische Untersuchung zur lexikalischen Vielfalt von direkten und indirekten Redeeinleitern." *Online-only Publikationen des Leibniz-Instituts für Deutsche Sprache*.