

Raumreferentielle Ausdrücke in deutschsprachigen Romanen des 19. und 20. Jahrhunderts

Ein Werkstattbericht des Projekts CANSpIN

Die Räumlichkeit des Textes

Im DFG-Projekt „**Computational Approaches to Narrative Space in 19th and 20th Century Novels**“ (**CANSpIN**) des Schwerpunktprogramms 2207 „Computational Literary Studies“ beschäftigen wir uns mit der computergestützten Analyse von Raum in deutsch- und spanischsprachigen Romanen des 19. und 20. Jahrhunderts. Das Poster zeigt erste Annotationsdaten, die mit Hilfe der Annotationsrichtlinie CANSpIN.CS1 generiert worden sind und die die Räumlichkeit von ausgewählten deutschsprachigen Romanen sichtbar machen.

Die **Räumlichkeit eines Textes** erfassen wir mittels *raumreferentieller Ausdrücke*, worunter wir das gesamte räumliche Vokabular eines Textes verstehen, das zunächst unabhängig von der erzählten Welt betrachtet wird: Beides, das räumliche Vokabular und die erzählte Welt der Erzählung, sind zwei unterschiedliche Eigenschaften eines Erzähltextes. Während die erzählte Welt die räumliche Umgebung der Figuren meint, ist mit der Räumlichkeit eines Romans die Menge und Verteilung räumlicher Ausdrücke in einem Text beschrieben.

Annotationsrichtlinie CANSpIN.CS1

version 1.0.0 License CC BY 4.0 DOI 10.5281/zenodo.12706812

Um die Räumlichkeit der untersuchten Texte sichtbar zu machen, annotieren wir die Romane mit Hilfe der Annotationsrichtlinie CANSpIN.CS1 (*category set*).¹ Die Kategorien basieren auf einer Auswahl und Derivation von Kategorien der Systematiken von Mareike Schumacher² und Katrin Dennerlein³. Fünf Arten von raumreferentiellen Ausdrücken werden damit erfasst, untergliedert in insgesamt 21 Klassen: **Orte** ■, **Bewegungen** ■, **Dimensionierungen** ■, **Richtungen** ■ und **Positionierungen** ■.

Sie stiegen auf die Bergspitze. Dabei erinnerte sie sich, wie es damals in ihrem kleinen Heimatsdorf gewesen war. »Ein Berg heißt dich immer willkommen, im Gegensatz zur Stadt!«, hörte sie plötzlich jemanden sagen.

Erste Daten: Übersicht

Um große Korpora von Romanen annotieren zu können, trainieren wir Sprachmodelle für eine zukünftige automatisierte Annotation. Die bereits vorliegenden Trainingsdaten wurden manuell mit jeweils mindestens drei Annotator:innen erzeugt:

Jahr	Autor:in	Titel	Kapitel	Token	Annotationen	IAA (gamma)
1845	Ernst Adolf Willkomm	Weisse Sklaven oder die Leiden des Volkes	2-3-5	6539	982	Protoannotation
1864	Gustav Freytag	Die verlorene Handschrift	1-1-1	7179	1052	Protoannotation
1880	Marie von Ebner-Eschenbach	Lotti, die Uhrmacherin	14	3959	602	Protoannotation
1963	Heinrich Böll	Ansichten eines Clowns	1	2689	411	Protoannotation
1965	Uwe Johnson	Zwei Ansichten	1	744	154	0.88
—	—	—	—	21110	3201	—

TABELLE 1: Übersicht über manuell annotierte Romankapitel. Zur Berechnung des Inter-Annotator-Agreements diente die Gamma-Metrik unter gleichgewichteter Berücksichtigung von Segmentierung und Klassifikation⁴.

Räumlichkeit und Diegese

Die mit CANSpIN.CS1 erzeugten Annotationen können einerseits für rein quantitative Analysen verwendet werden. Andererseits eröffnen raumreferentielle Ausdrücke auch die Möglichkeit, die Beziehungen zwischen Räumlichkeit und Diegese in einem Mixed-Method-Ansatz in den Blick zu nehmen. Denn grundsätzlich gilt: Am Erzählen von Raum ist meist räumliches Vokabular beteiligt, während andererseits nicht jeder raumreferentielle Ausdruck zur Konstituierung der erzählten Welt beiträgt.

Die Beziehungen zwischen Räumlichkeit und Diegese sind daher ein vielversprechender Untersuchungsgegenstand. Das Verhältnis zwischen beidem könnte ein Signum spezifischer Schreibweisen sein: Wird im Text viel oder wenig räumliches Vokabular für die Konstruktion der erzählten Welt verwendet? Zudem könnten Muster der Räumlichkeit (in Menge, Verteilung, Auswahl und Korrelationen von Annotationen) Hinweise auf Eigenschaften der Diegese geben, was für deren computationale Erkennung folglich nutzbar wäre.

Annotationsmengen

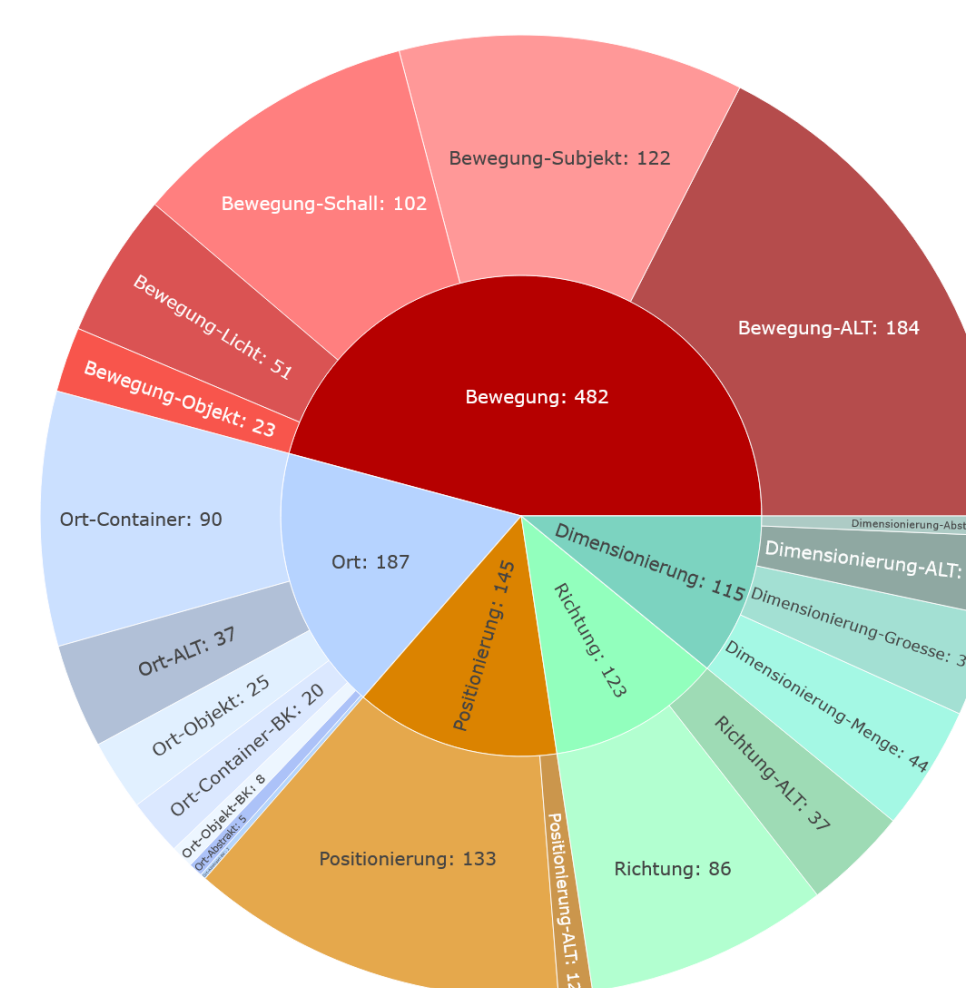


ABBILDUNG 1: Annotationsmengen im ersten Kapitel von **Gustav Freytag**: *Die verlorene Handschrift* [1864].⁵

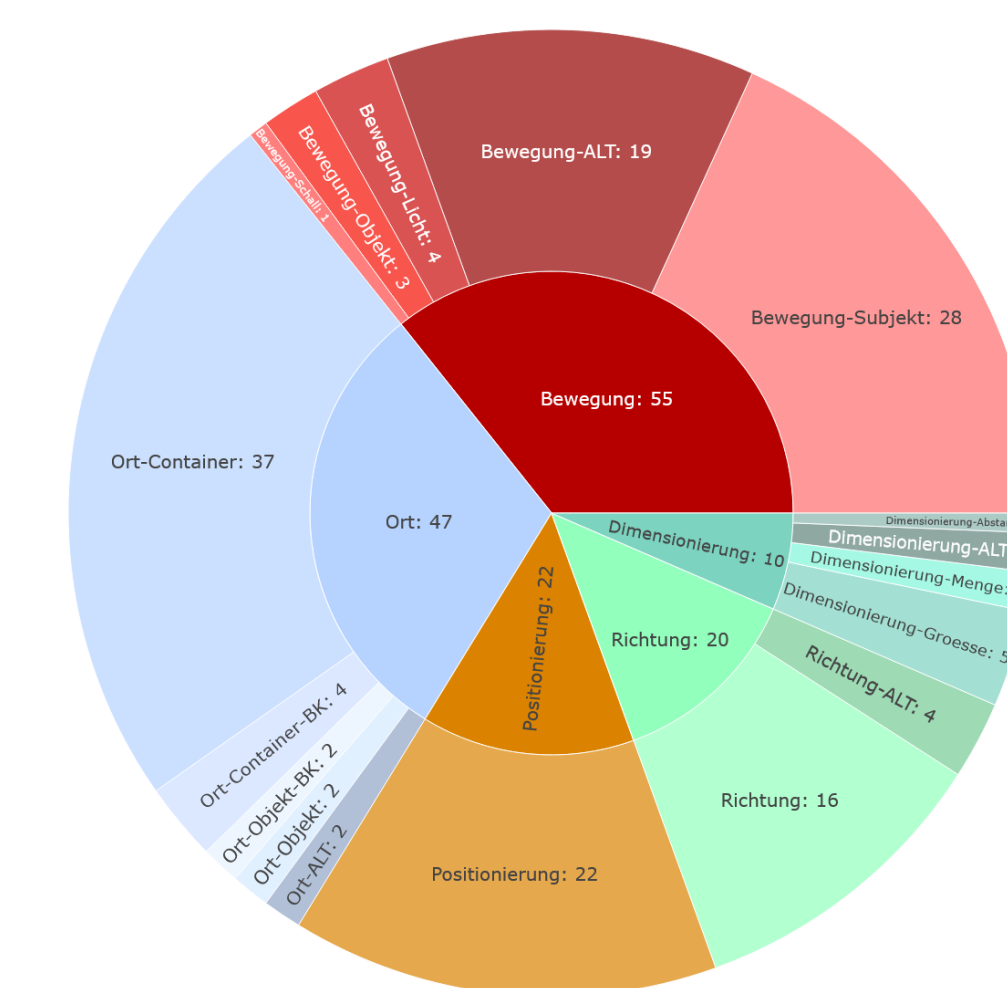


ABBILDUNG 2: Annotationsmengen im ersten Kapitel von **Uwe Johnson**: *Zwei Ansichten* [1965].

Im Mengendiagramm zu Freytags Roman (Abb. 1) ist zu erkennen, dass *Ort-Container*-Annotationen (**Substantive, die Bereiche und Räume bezeichnen, in denen sich prototypisch Menschen aufhalten können**) deutlich häufiger vorkommen als *Ort-Container-BK* (**Container, die im Satz Ausgangs- oder Zielpunkte von Bewegungen sind**). Dies korreliert mit dem Umstand, dass im ersten Kapitel Figuren-Bewegungen und Wahrnehmungen eher innerhalb von Containern stattfinden als zwischen ihnen.

Der Vergleich von Freytags (Abb. 1) und Johnsons Roman (Abb. 2.) offenbart Unterschiede in den Mengen der **Bewegungsklassen** (**Verben, die eine räumliche Distanz durch eine Bewegung von Subjekten, Objekten oder durch Wahrnehmungen produzieren**). Die häufige Darstellung von Dialogen hat bei Freytag eine große Zahl von *Bewegung-Schall* zur Folge, während bei Johnson vor allem Ereignisse erzählt werden und *Bewegung-Subjekt* dominiert.

Annotationsverteilungen

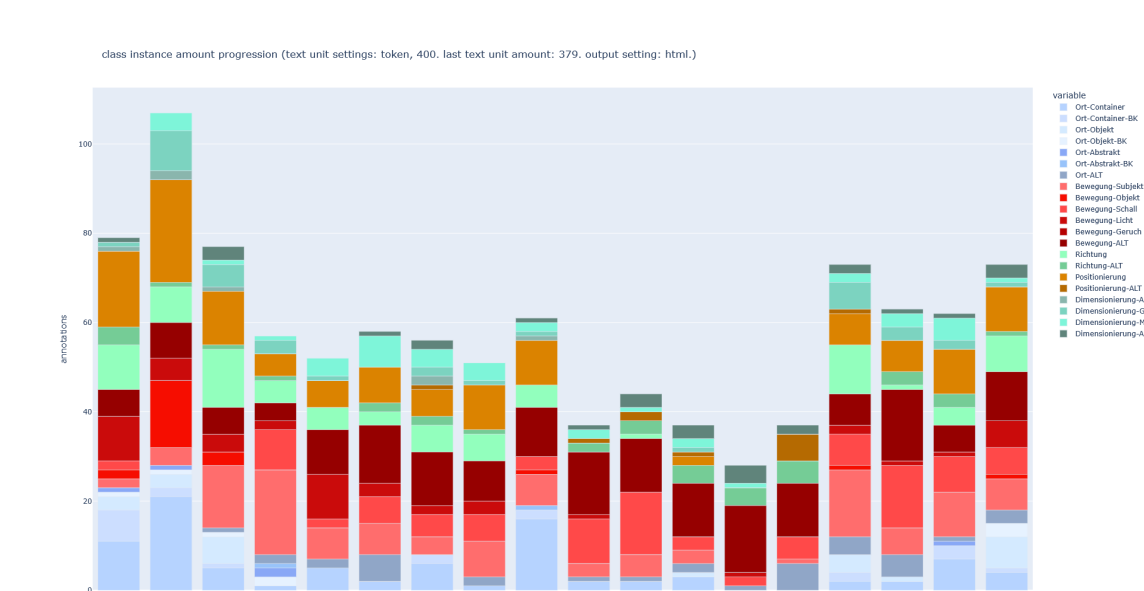


ABBILDUNG 3: Annotationsverteilung im ersten Kapitel von **Gustav Freytag**: *Die verlorene Handschrift* [1864].⁵

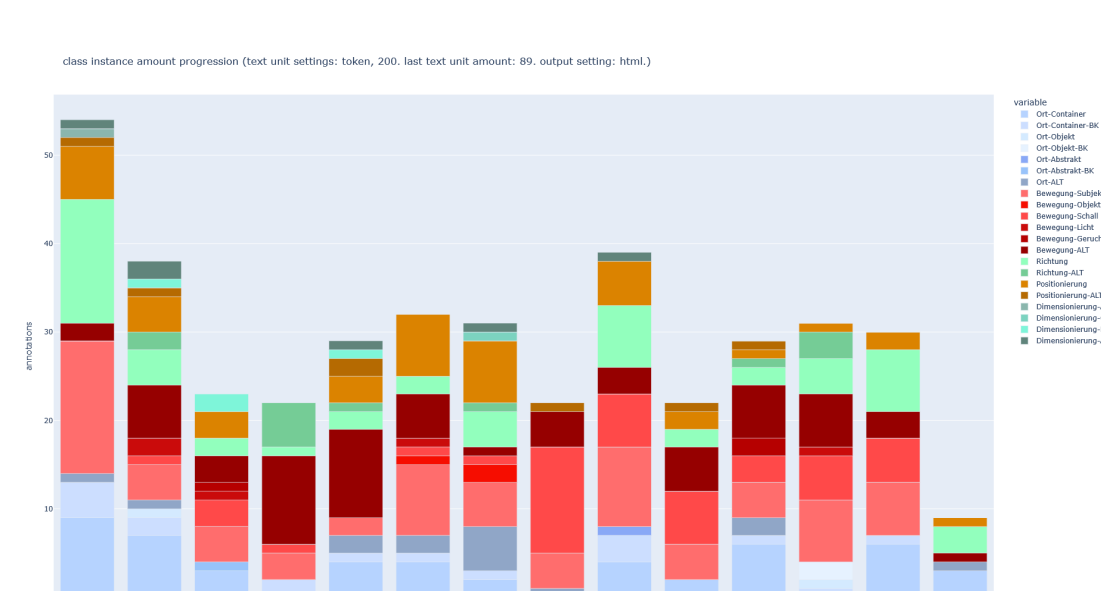


ABBILDUNG 4: Annotationsverteilung im ersten Kapitel von **Heinrich Böll**: *Ansichten eines Clowns* [1963].

Die Verteilung in Freytags Anfangskapitel (Abb. 3) zeigt zu Beginn viel *Bewegung-Licht* und *Bewegung-Objekt*, ab der Hälfte aber eher *Bewegung-Subjekt* und *Bewegung-Schall*. Auf die stark räumliche Exposition folgt die Darstellung von Dialogen. Bölls (Abb. 4) Exposition dagegen ist keine Raum-, sondern eine Routinenbeschreibung des Protagonisten. Dennoch fällt auch hier die Menge des räumlichen Vokabulars zu Beginn am höchsten aus.

Literatur

- [1] Henny-Krahmer, Ulrike, Nils Kellner und Marc Lemke: *CANSpIN.CS1*, Juli 2024. <https://doi.org/10.5281/zenodo.12706812>, Version 1.0.0.
- [2] Schumacher, Mareike: *Orte und Räume im Roman. Ein Beitrag zur digitalen Literaturwissenschaft*. Digitale Literaturwissenschaft. Metzler, Berlin und Heidelberg, 2023. <https://doi.org/10.1007/978-3-662-66035-5>.
- [3] Dennerlein, Katrin: *Narratologie des Raumes*. De Gruyter, Berlin und New York, 2009.
- [4] Reiter, Nils und Leonard Konle: *Messverfahren zum Inter-annotator-agreement (IAA)*. DARIAH-DE Working Papers, 44, 2022. <https://doi.org/10.47952/gro-publ-103>.
- [5] Konle, Leonard, Fotis Jannidis, Carolin Odebrecht und Lou Burnard: *German Novel Corpus (ELTeC-deu)*. <https://doi.org/10.5281/zenodo.4662482>, 2021 (Version 1.0.0) [Data set].