

Ontologische Semantik

–

Eine korpusgestützte Untersuchung ihrer Verbalisierung
in deutschen Lexikonartikeln

Inaugural-Dissertation
zur Erlangung der Doktorwürde
der
Philosophischen Fakultät
der
Rheinischen Friedrich-Wilhelms-Universität
zu Bonn

vorgelegt von

Michael Johann Dembach

aus

Brampton, Kanada

Bonn 2024

Gedruckt mit der Genehmigung der Philosophischen Fakultät
der Rheinischen Friedrich-Wilhelms-Universität Bonn

Zusammensetzung der Prüfungskommission:

Prof. Dr. Kristian Berg

(Vorsitzender)

Prof. Dr. Claudia Wich-Reif

(Betreuerin und Gutachterin)

Prof. Dr. Ulrich Schade

(Gutachter)

Prof. Dr. Thomas Klein

(weiteres prüfungsberechtigtes Mitglied)

Tag der mündlichen Prüfung: 21.06.2023

Im Leben ist es ja nie der mathematische Satz, den wir brauchen, sondern wir benützen den mathematischen Satz nur, um aus Sätzen, welche nicht der Mathematik angehören, auf andere zu schließen, welche gleichfalls nicht der Mathematik angehören.

Wittgenstein, Tractatus logico-philosophicus, Satz 6.211

Für meine Frau und Lektorin Bianca

Inhaltsverzeichnis

Abbildungsverzeichnis	6
1 Einleitung	8
2 Forschungsstand	13
3 Formale und sprachwissenschaftliche Grundlagen	18
3.1 Relationen und Mengenlehre	19
3.2 Semantische Rollen und Pertinenzrelationen	22
3.2.1 Semantische Rollen	24
3.2.2 Pertinenzrelationen	31
3.3 Aussagen- und Prädikatenlogik	32
3.4 Description Logics	41
3.5 Intensionale Logik	49
4 Ontologiebegriff	60
5 Ontologien und Semantik	68
5.1 Logische Merkmalssemantik	69
5.2 Prototypensemantik	76
5.3 Frame-Semantik	80
6 Korpus und Annotationstypen	83
6.1 Erkenntnisse, die durch das Korpus gewonnen werden	83
6.2 Textauswahl	85
6.3 Annotationstypologie	90
7 Mengentheoretische Relationen	97
7.1 Taxonomische Relation	100
7.2 Äquivalenz	114
7.3 Komplement	122
7.4 Instanz-Klasse	131
7.5 Instanz-Metaklasse	136
7.6 Zusammenfassung	140
8 Semantische Rollen als ontologische Rollen	141
8.1 Agent	144
8.2 Force	152
8.3 Producer	158
8.4 Depiction	166
8.5 Shape	171
8.6 Experiencer	174
8.7 Sender	180
8.8 Place	182
8.9 Zusammenfassung	183

9	Pertinenzrelationen	185
9.1	Meronym – Holonym	186
9.2	Possessor – Possession	202
9.3	Zusammenfassung	206
10	Numerische Funktionen	207
11	<i>False Friends</i>	213
11.1	Metasprache	214
11.2	Formale Ambiguität	218
11.3	Wissensgebiet	219
11.4	Zusammenfassung	221
12	Fazit	221
	Literaturverzeichnis	226
	Quellenverzeichnis	226
	Forschungsliteratur	255

Abbildungsverzeichnis

1	Verortung einiger semantischer Rollen in einem zweidimensionalen Koordinatensystem	29
2	Einfache Wein-Ontologie	60
3	Spektrum formaler Ausdrucksmöglichkeiten	63
4	Die mengentheoretischen Relationen im Gesamtkorpus	97
5	Vergleich zwischen Wikipedia und Duden: Mengentheoretische Relationen	99
6	Die Differenz zwischen WK und WoIK: Mengentheoretische Annotationen	99
7	Die Verteilung des Annotationstyps Oberklasse über die Wissensgebiete	100
8	Die Verteilung des Annotationstyps Äquivalent über die Wissensgebiete	115
9	Die Verteilung des Annotationstyps Komplement über die Wissensgebiete	123
10	Die Verteilung der Instanzrelation über die Wissensgebiete	131
11	Die Verteilung von Instanz und Metaklasse über die Wissensgebiete	137
12	Der Vergleich zwischen Wikipedia und Duden: Semantische Rollen	142
13	Die Differenz zwischen WK und WoIK	143
14	Das Gesamtkorpus mit klassischer Rollenannotation	143
15	Die Verteilung von <i>Agent</i> über die Teilkorpora	145
16	Die Verteilung von <i>Force</i> über die Teilkorpora	152
17	Die Verteilung von <i>Producer</i> über die Teilkorpora	159
18	Die Verteilung von <i>Depicting</i> über die Teilkorpora	166
19	Die Verteilung von <i>Shape</i> über die Teilkorpora	171

20	Die Verteilung von <i>Experiencer</i> über die Teilkorpora	175
21	Die Verteilung von <i>Sender</i> über die Teilkorpora	180
22	Die Verteilung von <i>Place</i> über die Teilkorpora	182
23	Der Vergleich zwischen Wikipedia und Duden: Pertinenzre- lationen	186
24	Der Vergleich zwischen WK und WoIK: Pertinenz-Annotationen	186
25	Meronym und Holonym in den verschiedenen Wissensgebieten	187
26	Die <i>Possessor-Possession</i> -Relation in den Teilkorpora	203
27	Der Vergleich zwischen WK und WoIK: <i>Datatype Property</i> . .	209
28	Der Vergleich zwischen Wikipedia und Duden: <i>False Friends</i> .	214
29	Die Verteilung von <i>False Friends</i> über die Wissensgebiete . .	214

1 Einleitung

Wie sich die Bedeutung von Wörtern darstellen lässt, ist eine der grundlegenden Fragen der Sprachwissenschaft. Die Technologie der Ontologien bietet eine mögliche Antwort auf diese Frage: Ontologien erfassen Bedeutung durch eine Mischung aus logischen Ausdrucksmitteln und Relationen, die zwischen Lexemen angesetzt werden. Diese Vorstellung von Bedeutung soll im Folgenden ontologische Semantik genannt werden. Als Verfahren ist diese insofern erfolgreich, als sie es ermöglicht, Bedeutungsaspekte eines Wortes in einer maschinenlesbaren Form darzustellen. Das heißt konkret, dass ein Computer bspw. in der Lage ist, aus der Information, dass Hunde Säugetiere sind, abzuleiten, dass Hunde lebend gebären – weil Säugetiere lebend gebären.

Die beiden Relationen, die in diesem Beispiel impliziert sind, bestehen einmal zwischen den Konzepten *Hund* und *Säugetier* und zwischen *Säugetier* und *Lebewesen, die lebend gebären*. Genau genommen, besteht eine dritte Relation, und zwar zwischen *Hund* und *Lebewesen, die lebend gebären*; diese Relation wird aus den ersten beiden geschlussfolgert.

Wie werden solche Relationen sprachlich ausgedrückt? Sind sie implizit in Sätzen versteckt oder werden sie explizit realisiert? Auf den ersten Blick liegt die Vermutung nahe, dass Verben Relationen ausdrücken. Ein Blick in die folgenden Beispiele offenbart jedoch, dass die Zusammenhänge komplexer sind.

- (1) a. *Hunde sind Säugetiere.*
- b. *Der Hund ist ein Säugetier.*
- c. *Der Begriff Hund bezeichnet ein Säugetier.*
- d. *Hunde sind eine Familie in der Klasse der Säugetiere.*
- e. *Lassie ist ein Hund.*

In (1a) und (1b) wird die Relation zwischen *Hund* und *Säugetier* durch das Verb *sein* hergestellt – warum ist es zulässig, für denselben Sinnzusammenhang einmal den Plural und einmal den Singular zu verwenden? In (1c) wird

ein anderes Verb, *bezeichnen*, verwendet, um die Relation zu realisieren; außerdem wird hervorgehoben, dass es sich bei *Hund* zunächst nur um einen Begriff handelt. Worin liegt die kommunikative Funktion solcher Formulierungen? In (1d) wird erneut *sein* verwendet, aber in diesem Fall stellt das Verb eine Relation zwischen *Hund* und *Familie* dar. *Säugetier*, das Oberkonzept, wird in diesem Fall als Genitivattribut zum Substantiv realisiert. In (1e) wird wieder eine Form von *sein* verwendet, diesmal stellt *sein* aber nicht eine Relation zwischen zwei Konzepten her, sondern zwischen einer konkreten Entität und einem Konzept. Wie lassen sich diese beiden Verwendungen von *sein* unterscheiden? Diese nicht abgeschlossene Liste an möglichen Formulierungsmustern für eine Relation zeigt, dass es sinnvoll ist, mehr als das Verb zu betrachten, um eine Relation zu identifizieren.

Mit dieser Arbeit soll untersucht werden, wie ontologische Semantik in natürlich-sprachlichen Texten realisiert wird. Ziel ist es, ein Verständnis von der Realisierung ontologischer Informationen zu erhalten, das aus einer generellen sprachwissenschaftlichen Perspektive hervorgeht und möglichst frei von technischen Vorüberlegungen ist. In dieser Arbeit soll unter einem Konzept eine abstrakte Einheit verstanden werden, die in der Kognition von Sprachproduzent*innen in irgendeiner Form vorliegt (vgl. Spalek 2012: 62). *Hund* ist ein Konzept, ebenso *Lebewesen, die lebend gebären*. Mit *Begriff* ist ein lexikalisiertes Konzept gemeint, also ein Konzept, das im mentalen Lexikon durch ein Lemma vertreten ist (vgl. Spalek 2012: 56). Mit *Konstrukt* meine ich ein Konzept, das aus einer wissenschaftlichen Beschäftigung hervorgegangen ist.

Kapitel 2 soll das Thema zunächst im aktuellen Stand der Forschung vororten. Dabei wird kurz das Verständnis von Semantik im Kontext des *Semantic Web* angerissen, um zu zeigen, wie sich das Konzept Ontologie aus diesem entwickelt hat. Daraufhin wird aufgezeigt, dass diese Arbeit Grundlagenforschung im Bereich des *Ontology Learning*, des automatischen Erzeugens von Ontologien aus natürlich-sprachlichen Texten, leistet, indem sie Aufschluss

darüber gibt, welche Formen von semantischen Informationen für Ontologien relevant sind und wie diese realisiert werden.

Kapitel 3 führt in die formalen Grundlagen ein und soll zwei Zwecke erfüllen: Zunächst werden hier die grundlegenden Konstrukte der formalen Semantik, die im Laufe der Arbeit verwendet werden, eingeführt und definiert. Dieses Kapitel wird außerdem die enge Verwandtschaft des Konstruktes Ontologie mit der Sprachwissenschaft verdeutlichen, indem es aufzeigt, dass Ontologien zwar im Kontext des *Semantic Web* entstanden sind, dass die Grundlagen aber in die Anfänge der formalen Semantik zurückreichen. Das Kapitel greift zunächst das linguistische Konstrukt der semantischen Relation auf und zeigt, wie dieses mit den Konstrukten der Mengenlehre korrespondiert. Diese Korrespondenz ist der Grundstein von Ontologien. Auf diesem Sachverhalt aufbauend werden andere Formen von Relationen, die zwischen Begriffen bestehen können, in den Blick genommen, u.a. wird gezeigt, dass sich das Konstrukt der semantischen Rolle ebenfalls als Relation begreifen lässt und damit mit den Konstrukten der Mengenlehre dargestellt werden kann. Diese Darstellung bietet die Basis, um andere Relationen zu formalisieren. Im Anschluss werden die Grundlagen der Logik und mit diesen der Wahrheitsbegriff eingeführt, welcher die Basis zur Formalisierung natürlich-sprachlicher Aussagen bildet. Außerdem sollen grundsätzliche Verfahren des logischen Schlussfolgerns, die im späteren Verlauf der Arbeit Anwendung finden, vorgestellt werden. Kapitel 3.4 sei hier besonders herausgestellt, da dort die Sprache $\mathcal{SHOIN}(\mathcal{D})$ eingeführt werden soll, die im Hauptteil der Arbeit verwendet wird, um natürlich-sprachliche Sätze zu formalisieren. Den Abschluss des Kapitels bildet eine Reflexion der intensionalen Semantik, die aufzeigt, inwiefern diese für die anstehenden Untersuchungen in Betracht gezogen werden muss.

Kapitel 4 widmet sich dem Konstrukt der Ontologie. Die verschiedenen Versuche, Ontologien zu definieren, werden vorgestellt und miteinander verglichen. Anhand einer Beispielsontologie wird außerdem ein eigener Onto-

logiebegriff entwickelt. Kapitel 5 widmet sich einigen Semantiktheorien, die mit Merkmalen operieren. Obwohl Ontologien ein semantisches Konstrukt darstellen, wurde die Stellung, die dieses Konstrukt innerhalb der semantischen Theorien einnimmt, noch nicht ausgiebig reflektiert. Das Kapitel soll daher herausarbeiten, dass Ontologien maßgeblich den Vorstellungen der logischen Merkmalssemantik entsprechen. Diese wird daraufhin mit zwei konkurrierenden Theorien, der Frame- und der Prototypensemantik, verglichen, die als Gegenentwürfe zur Merkmalssemantik verstanden werden können; dabei wird allerdings dargelegt, dass sich diese Theorien zumindest im Kontext dieser Arbeit nicht unvereinbar gegenüberstehen.

Kapitel 6 widmet sich dem Korpus, das für die Untersuchungen dieser Arbeit erstellt und annotiert wurde. Das Korpus besteht mit 39.496 Worttoken zum größten Teil aus den einleitenden Sätzen von Wikipedia-Artikeln, weil in diesen zur Definition eines Begriffes ontologische Zusammenhänge ausformuliert werden. Diese Texte wurden um ein Korpus von Duden-Artikeln, das 8.454 Worttoken fasst, ergänzt, um zu untersuchen, wie textsortenspezifisch bestimmte Phänomene eventuell sind. Insgesamt fasst das Korpus 47.950 Worttoken. In dem Kapitel wird erläutert, wie bei der Erstellung vorgegangen wurde und welche Erkenntnisse mit einem solchen Korpus gewonnen werden können. Das Kapitel soll aber ebenfalls ausführen, zu welchen Zwecken das Korpus nicht erhoben wurde und welchen Anforderungen es deswegen evtl. nicht gerecht wird. Das Kapitel endet mit einer Vorstellung der verschiedenen Annotationstypen, mit denen das Korpus versehen ist.

Mit Kapitel 7 beginnt die Darlegung der Forschungsergebnisse. In diesem Kapitel wird dargestellt, wie die basalen Konstrukte der Mengenlehre im Korpus realisiert werden. Dabei wird detailliert auf deren Formulierung eingegangen, um den Variationsreichtum herauszuarbeiten. Es wird außerdem die Frage gestellt, ob tatsächlich alle Relationen, die sich mengentheoretisch bilden lassen, eigene Formulierungsmuster aufweisen. Wenn dem nicht so ist, wenn z.B. zwei Relationen die gleichen Muster aufweisen, wäre dies ein Hin-

weis darauf, dass sie von Sprachproduzent*innen nicht als unterschiedliche Relationen wahrgenommen werden.

In Kapitel 8 werden schließlich weitere Relationen, die zwischen Begriffen bestehen können, untersucht. Dabei soll herausgearbeitet werden, wie sich diese Relationen als semantische Rollen darstellen und dadurch auch typologisieren lassen. Dass dies möglich ist, stellt eine der zentralen Thesen dieser Arbeit dar. Das in diesem Kapitel vorgestellte Verfahren versteht sich als eine konzeptionelle Möglichkeit, Relationen jenseits der Mengenlehre in eine Ontologie zu überführen. Kapitel 10 befasst sich mit Funktionen. Es gilt herauszuarbeiten, welche Formulierungen der natürlichen Sprache sich aus welchen Gründen am ehesten als Funktionen formalisieren lassen. Kapitel 11 bespricht bestimmte Formulierungen, deren Inhalt nicht formalisiert werden sollen, deren Identifikation aber dennoch wichtig ist, da sie mit ontologisch-relevanten Informationen verwechselt werden könnten.

Kapitel 9 widmet sich den sogenannten Pertinenzrelationen. Dabei handelt es sich um Relationen, bei denen ein räumliches oder possessives Verhältnis vorliegt. Auch hier werden die unterschiedlichen Realisierungsmöglichkeiten in den Vordergrund gestellt, um erneut herauszuarbeiten, wie unterschiedlich diese Relationen realisiert werden können. Das Kapitel wird außerdem ausarbeiten, welche verschiedenen Unterklassen dieser Relation angesetzt und wie diese miteinander kombiniert werden können.

Kapitel 10 behandelt schließlich Wortgruppen, die in Relation zu einem numerischen Wert gesetzt werden. Bei diesen Relationen ergeben sich verschiedene Möglichkeiten der Realisierung, wobei sich die Frage stellt, welche Realisierung am ehesten dem sprachlichen Empfinden entspricht und ob dies überhaupt relevant ist.

Kapitel 11 schließt den Hauptteil mit einer Betrachtung von Formulierungen ab, deren Form ihren Inhalt verschleiert. Die Betrachtung dieser Formulierungen ist wichtig, da sie leicht falsch interpretiert und damit formalisiert werden können.

Kapitel 12 bildet den Schlussteil und fasst die Ergebnisse noch einmal zusammen. Die Ergebnisse der Arbeit werden zusammengefasst und ein Ausblick wird gegeben, wie die Ergebnisse im Bereich des *Ontology Learning* verwendet werden könnten.

2 Forschungsstand

Ontologien sind eine Technologie, die im Kontext des *Semantic Web* entstanden ist, welches eine bestimmte Vorstellung von Semantik mit sich bringt, die zunächst angerissen werden soll, bevor das Vorhaben dieser Arbeit in diesen Kontext eingegliedert werden kann. Ein Antrieb hinter dem *Semantic Web* ist es, dass die Kommunikation zwischen Maschinen untereinander oder zwischen Mensch und Maschine funktionieren sollte wie natürliche Sprache, dazu müssen Konzepte eine Bedeutung über die Grenzen eines einzelnen Systems hinaus haben (vgl. Berners-Lee u.a. 2001: 3 und Segaran u.a. 2009: 3). Bspw. soll eine Software des *Semantic Web* in der Lage sein, verschiedene Homepages nach dem Konzept *Auto* zu durchsuchen, auch wenn dieses in den Homepages durch andere Begriffe vertreten ist. Im Zentrum der Methodik steht dabei insbesondere die Graphentheorie, die den theoretischen Rahmen bietet, um Daten und Datensätze miteinander in Beziehung zu setzen (vgl. Hebler u.a. 2009: 24-25 und Wood u.a. 2014: 27). Mit Ontologien kam schließlich der Anspruch auf, dass man nicht nur Daten, in denen bestimmte Konzepte genannt werden, miteinander in Beziehung setzen kann, sondern die Konzepte selber über Beziehungen definieren kann, wenn die verwendeten Relationen in einer logischen Sprache ausgedrückt werden (vgl. Eskridge und Hoffman 2012: 59-61). Dieser Ansatz ist eng mit der Frage verknüpft, ob sich strukturierte Daten, also in einer logischen Sprache formulierte Daten, nicht aus unstrukturierten Daten, worunter natürlich-sprachliche Texte fallen, erzeugen lassen. Das Fachgebiet, das sich mit dieser Frage beschäftigt, wird *Knowledge Extraction* genannt, in dessen Teilgebiet *Ontology Learning* sich

diese Arbeit verorten lässt (vgl. Corcoglioniti u.a. 2016: 3263).

Ontology Learning besteht einerseits aus *Ontology Population*, bei dem eine bestehende Ontologie mit Informationen über konkrete Entitäten angereichert wird. Bspw. könnte eine Ontologie über politische Ämter mit den entsprechenden Amtsträger*innen angereichert werden (vgl. Faria u.a. 2014: 27). Auf der anderen Seite steht *Ontology Enrichment* (vgl. Tissaoui u.a. 2022: 2), bei dem eine Ontologie um weitere Konzepte erweitert wird. Bspw. könnte die Ontologie über politische Ämter um weitere Ämter ergänzt werden. Diese Arbeit will zu beiden Themen einen Beitrag leisten. *Ontology Learning* befasst sich zu großen Teilen mit natürlicher Sprache. Es wird versucht, die Aussagen, die in Texten getätigt werden, in eine ontologische Struktur zu überführen, d.h. sie als Relationen zu begreifen und logisch auszudrücken. Die zwei zentralen Fragen, die sich dabei aufdrängen, lauten: Wie viele unterschiedliche Relationen sind anzusetzen und wie werden diese Relationen in Texten realisiert?

Diese Fragen werden in der Forschungsliteratur eingehend diskutiert: Szostak (vgl. 2014: 46-49) erkennt den Wert einer ausgearbeiteten Typologie von Relationen; sein Vorschlag mit fünf verschiedenen Typen, die sich teilweise überlagern, lässt allerdings noch Raum für Spezifizierungen. Faria u.a. (vgl. 2014: 32) konzentrieren sich auf die Realisierung dieser Relationen in Texten und schreiben von bestimmten *Triggern*, die in Texten identifiziert werden müssen und als eine entsprechende Relation in eine Ontologie übertragen werden können. Bspw. ist das Verb *heiraten* in seinen verschiedenen Wortformen ein *Trigger* dafür, dass zwischen zwei Entitäten die Relation des Verheiratet-Seins besteht. Dieser Ansatz steht vor der Herausforderung, eine konsistente und eindeutige Zuordnung von Relationen und *Triggern* auszuarbeiten, was bei Faria u.a. (2014) – zumindest aus einer rein sprachwissenschaftlichen Perspektive – nur in Ansätzen geschieht. Bspw. kann die Relation des Verheiratet-Seins auch durch ein Genitivattribut ausgedrückt werden: *die Frau des Bauern*.

Ein solche Zuordnung versuchen Corcoglioni u.a. (vgl. 2016: 3261) dadurch zu erlangen, dass sie die in den Texten aufgefundenen Relationen auf eine bestehende Datenbank von *Frames* abbilden. Der Ansatz ähnelt in manchen Teilen der Vorgehensweise dieser Arbeit, allerdings lässt er die Frage offen, wie die Relationen in den Texten identifiziert werden können.

Viele methodische Ansätze des *Ontology Learning* konzentrieren sich bewusst nur auf einen bestimmten Teil von Informationen: Ganino u.a. (vgl. 2018: 2311) extrahieren nur Individuen aus Texten und Ayadi u.a. (vgl. 2019: 575) haben einen Ansatz erarbeitet, mit dem die Instanzen der *Biomolecular Network Ontology* um weitere Informationen ergänzt werden können. Selbst ganz aktuelle Ansätze mit *Deep Neural Networks*, wie die Arbeit von Muscetti u.a. (vgl. 2022: 1288-1289), oder mit *Latent Dirichlet Allocation*, wie die Arbeit von Tissaoui u.a. (vgl. 2022: 5-9), versuchen zunächst nur, die Hierarchie der Konzepte in einer Ontologie automatisch aus Texten zu erweitern.

Es wird deutlich, dass das Teilgebiet des *Ontology Learning* viele Fragestellungen der Sprachwissenschaft und dabei vor allem der Semantik streift. Systematische Gegenüberstellungen der beiden Fachgebiete sind allerdings selten, was vermutlich in den unterschiedlichen Anwendungsbereichen begründet liegt. Zu erwähnen ist an dieser Stelle die Arbeit von Nirenburg und Raskin (2004), die aber mit einem sehr weiten Ontologiebegriff operieren, der pragmatische, diskurstheoretische und psychologische Aspekte involviert (vgl. Nirenburg und Raskin 2004: 25) und damit weit von dem semantischen Fokus dieser Arbeit wegführt. Das *tertium comparationis*, auf dem ein Vergleich von Ontologien und Semantik aufbauen könnte, ist der Umstand, dass beide Konzepte sich mit Begriffen beschäftigen. Die Semantik beschäftigt sich generell mit Bedeutung und somit auch mit der Bedeutung von Begriffen. In Ontologien werden Begriffe definiert, indem sie als Klassen formalisiert werden – eine Vorgehensweise, die in Kapitel 4 angedeutet wurde. Ein*e Ontolog*in stellt sich demnach die Frage, wie sich die Bedeutung von Begriffen in eine logische Sprache überführen lassen. Ähnliche Bemühungen

markieren auch die Anfänge der formalen Semantik. Beschäftigte sich die logische Semantik zunächst mit den künstlichen Sprachen der Logik – wobei vor allem Frege (1879) zu nennen ist – wurden später unter dem Begriff auch Ansätze zusammengefasst, die versuchten, Probleme der Semantik mit dem Instrumentarium der Logik anzugehen. Der Ansatz der logischen Semantik lässt sich folgendermaßen zusammenfassen: Ein Ausdruck einer natürlichen Sprache wird durch einen Ausdruck einer Konstruktssprache so ersetzt, dass der Wahrheitswert des Satzes unverändert bleibt, um den Ausdruck dann mit den Mitteln der Logik zu evaluieren (vgl. Busse 2009: 36). Dies beschreibt in weiten Teilen das Teilgebiet der Ontologie (vgl. Herre u.a. 2006: 2) und die modelltheoretische Semantik, die in Kapitel 3.3 vorgestellt wurde, lässt sich als eine Methode begreifen, logische Sätze zu evaluieren.

Wissenschaftsgeschichtlich schloss sich an die logische Semantik die sogenannte logische Merkmals- oder Komponenten-Semantik an, die sich einem Problem stellt, mit dem man sich auch bei der Erstellung einer Ontologie konfrontiert sieht: Welche Aspekte eines Begriffes sind relevant und sollten in die logische Sprache überführt werden? Ist bspw. alles, was man über eine Avocado wissen kann, Teil der Bedeutung des Wortes *Avocado* (vgl. Löbner 2015: 19, Ziem 2008: 38 und Busse 2012: 9-10)? Dies kann kaum der Fall sein, da man sehr wenig über eine Avocado wissen muss, um das Wort richtig zu gebrauchen. In merkmalssemantischen Theorien wird versucht, Begriffe zu bestimmen und von anderen zu unterscheiden, indem systematisch ermittelt wird, welche Merkmale jeweils auf die einzelnen Begriffe zutreffen und welche nicht (vgl. schon Katz und Fodor 1963: 187). Besonders im ontologischen Kontext sind dabei zwei Konzepte von Bedeutung, die schon aus der antiken Philosophie als die relevanten Merkmale eines Konzeptes benannt werden: *genus proximum* und *differentia specifica* (vgl. Katz 1972: 182). Dem Verfahren liegt die Vorstellung zugrunde, dass ein Begriff durch seine nächsthöhere Klasse und das, was ihn von anderen Vertretern dieser Klasse unterscheidet, bestimmt werden kann; die beiden Merkmale bzw. Merkmalsbündel

sollen zusammen notwendig und hinreichend sein. Dieses Verfahren zur Bestimmung von Begriffen stellt einen zentralen Aspekt von Ontologien dar.

Die Sprachwissenschaft kann einen Grundlagenbeitrag zum *Ontology Learning* leisten: Sie kann zunächst eine konsistente Typologie von ontologisch-relevanten Relationen entwerfen, die unabhängig von verschiedenen Wissensdomänen ist und sich maßgeblich daran orientiert, wie Bedeutung in natürlicher Sprache strukturiert ist – denn, wie eingangs erwähnt, ist eine Orientierung an natürlicher Sprache eines der Grundprinzipien des *Semantic Web* (vgl. Berners-Lee u.a. 2001: 3). Der zweite Beitrag, den die Sprachwissenschaft leisten kann, besteht darin, systematisch zu erforschen, wie die relevanten Relationen in Texten realisiert werden. Wie groß dieser Variationsreichtum sein kann, wurde schon durch die Beispiele (1) im Einleitungskapitel angedeutet. Diese Komplexität und dieser Variationsreichtum, den natürliche Sprache aufweisen kann, wird von anderen Wissenschaften häufig unterschätzt. Bspw. gehen Lubani u.a. (vgl. 2019: 11) in ihrer Studie zu verschiedenen Aspekten von *Ontology Population* nur kurz darauf ein, dass Relationen durch Verben oder evtl. auch innerhalb einer Nominalphrase (NP) ausgedrückt werden können.

Diese Arbeit schafft eine sprachwissenschaftliche Grundlage für das Gebiet des *Ontology Learning*, indem aufgezeigt wird, wie eine semantisch fundierte Typologie von Relationen entworfen werden kann – u.a. durch Hinzuziehen des Konstruktes der semantischen Rollen –, und erforscht wird, wie diese Relationen realisiert werden. Neben Sprachwissenschaftler*innen ist sie daher für Forscher*innen im Bereich der Informatik und Computerlinguistik, welche Methoden im *Ontology Learning* entwickeln, relevant.

3 Formale und sprachwissenschaftliche Grundlagen

Dieses Kapitel soll darstellen, welche semantischen Zusammenhänge sich wie formalisieren und damit in eine Ontologie übertragen lassen. Der Hauptzweck liegt demnach darin, die formalen Konstrukte, die für diese Arbeit relevant sind, einzuführen. Um eine möglichst allgemeine sprachwissenschaftliche Perspektive einzunehmen, werden informatische Details weitestgehend vermieden, indem keine konkreten Ontologiesprachen dargestellt oder verwendet werden.¹ Stattdessen soll eine *Description Logic* (DL) eingeführt werden, die im weiteren Verlauf zur Formalisierung von natürlich-sprachlichen Aussagen verwendet wird. Ein zweites Ziel dieses Kapitels besteht darin, den theoretischen Zusammenhang zwischen Ontologien und formaler Semantik aufzuzeigen. Ferner soll das Konstrukt der semantischen Rollen vorgestellt werden, auf dessen Grundlage in Kapitel 8 bestimmte Relationen formalisiert werden sollen. Das Kapitel schließt mit einer Reflexion über die Bedeutung der intensionalen Logik für die Ausführungen dieser Arbeit.

Zunächst soll jedoch das Wort *formal* kurz konkretisiert werden. Ein formales System zeichnet sich allein durch seine Struktur aus. Formeln in einer logischen Sprache sagen zunächst nichts über die Welt aus. Man kann mit ihnen lediglich von als wahr angenommenen Sätzen weitere Sätze ableiten (vgl. Zoglauer 2016: 9). Dies trifft auch auf eine Ontologie zu: Ob das Wissen, das in einer Ontologie hinterlegt ist, mit der Wirklichkeit übereinstimmt, ist eine Frage der jeweiligen Wissensdomäne; die Ontologie ist lediglich eine Struktur, mit der neue Inhalte auf der Basis von als wahr akzeptierten Inhalten erschlossen werden können. Eine Aussage wird außerdem als formal bezeichnet, wenn sie in einer logischen Sprache formuliert ist (vgl. Lampert

¹Zu F-Logic vgl. Angele u.a. (2009), zum *Resource Description Framework* (RDF) vgl. Pan (2009), zur *Web Ontology Language* (OWL) vgl. Antoniou und van Harmelen (2009) und Uschold (2018). Relevant bzgl. der Verwendung einer Prädikatenlogik zweiter Stufe ist außerdem Giraud (2014).

2003: 15).

3.1 Relationen und Mengenlehre

Dieses Unterkapitel erläutert Begriffe der Mengenlehre, um aufzuzeigen, wie sich mit diesen der zentrale Begriff des *Semantic Web*, die Relation, formalisieren lässt. An den entsprechenden Stellen werden Parallelen zu dem Konstrukt der semantischen Relationen aufgezeigt, das die Verbindungen zwischen Wortbedeutungen beschreibt (vgl. Schwarz-Friesel und Chur 2014: 60). Durch diese Darstellungen werden die nötigen Grundlagen geschaffen, um den Ausführungen speziell in Kapitel (3.4), aber auch generell in dieser Arbeit folgen zu können. Darüber hinaus deuten die Darstellungen einen Zusammenhang zwischen Ontologien und Sprachwissenschaft an, der über die Grundlagen des *Semantic Web* hinausgeht.

Die Menge ist das grundlegendste Konstrukt der Mathematik. Eine Menge ist eine abstrakte Ansammlung von Objekten, die die Elemente der Menge sind, wobei deren Reihenfolge und Häufigkeit irrelevant ist (vgl. Partee u.a. 1990: 3-8). Man schreibt $a \in A$, um zu sagen, dass a Element der Menge A ist. Im Kontext von Ontologien wird meistens von Klassen und nicht von Mengen gesprochen, ein Unterschied, dem in der Semantik oft keine Bedeutung zugesprochen wird (vgl. Lyons 1977: 154). Hier soll eine Menge als *Klasse* bezeichnet werden, wenn ihre Elemente eine strukturelle Eigenschaft teilen, die sie als Elemente der Menge definiert. Die Menge der Säugetiere ist eine Klasse, da ihre Elemente alle die Eigenschaft teilen, lebend zu gebären. Die Elemente einer Klasse werden als Instanzen bezeichnet. Ab Kapitel 4 werden meistens die Begriffe *Klasse* und *Instanz* gebraucht; nur wenn über Inhalte gesprochen wird, die sich explizit mit der Mengenlehre beschäftigen, werden die Begriffe *Menge* und *Element* verwendet.

Die Bedeutung von Wörtern oder Begriffen lässt sich als Klasse darstellen. Das Wort *Säugetier* lässt sich als die Klasse begreifen, deren Elemente alle Säugetiere sind. Ein Adjektiv wie *sterblich* bezeichnet die Klasse der

sterblichen Entitäten und ein Verb wie *laufen* die Klasse der laufenden Entitäten. NPn bilden ebenfalls Klassen: *Sterbliche Hunde*; *Deutschlands Hunde*; *Hunde, mit schwarzem Fell* oder *Hunde, die gerne auf Wiesen spielen* sind Ausdrücke, die sich als Klassen darstellen lassen. Es ist zunächst nebensächlich, dass nicht unbedingt alle Elemente einer Klasse aufgezählt werden können.

Zwischen Mengen lassen sich Relationen definieren. Bei den grundlegenden Relationen, die im Folgenden beschrieben werden, handelt es sich um gleiche Mengen, die Teilmenge, die Vereinigungsmenge, die Schnittmenge und die Differenzmenge (vgl. Partee u.a. 1990: 9-17 und Lohnstein 2011: 10-17). Die folgenden Konstrukte werden zur Veranschaulichung immer für den Fall von zwei Mengen dargestellt, lassen sich aber auf beliebig viele Mengen erweitern.

Zunächst sind zwei Mengen gleich, wenn sie dieselben Elemente enthalten. Die Menge der Farben auf der polnischen Flagge und die Menge der Farben auf der indonesischen Flagge sind gleich (die Reihenfolge der Elemente ist zur Bestimmung einer Menge nicht von Bedeutung). Die korrespondierende semantische Relation ist die Synonymie: Zwei Begriffe sind gleich, wenn sie dieselben Dinge denotieren (vgl. Busse 2009: 104 und Stuckenschmidt 2011: 21).²

Eine Menge A ist eine Teilmenge einer Menge B, wenn alle Elemente von A auch in B sind; man schreibt dafür $A \subseteq B$. A ist eine echte Teilmenge von B, wenn alle Elemente von A auch in B sind, die Mengen aber nicht gleich sind; hierfür schreibt man $A \subset B$. Umgekehrt schreibt man $A \supset B$ bzw. $A \supseteq B$, wenn A die (echte) Obermenge von B ist. Diese Relation ist vielleicht die grundlegendste nicht nur in der Mengenlehre, sondern auch in der Semantik. Sie entspricht der semantischen Relation der Hyperonymie bzw. Hyponomie, die zwischen einem Über- und einem Unterbegriff besteht (vgl. Busse 2009:

²Die Diskussion, ob so etwas wie echte Synonymie existiert oder sich vermeintliche Synonyme über Konnotationen unterscheiden lassen, kann im formalen Kontext ausgeklammert werden.

104 und Stuckenschmidt 2011: 21). Der Begriff *Hund* bezeichnet bspw. eine Teilmenge dessen, was der Begriff *Lebewesen* bezeichnet.

Die Vereinigungsmenge von zwei Mengen A und B ist definiert durch $A \cup B = \{x \mid x \in A \text{ oder (nichtausschließendes oder) } x \in B\}$; es ist die Menge von Elementen, die entweder in A, in B oder in beiden Mengen enthalten sind. Der Begriff *Eltern* lässt sich als Vereinigungsmenge der Begriffe *Mutter* und *Vater* darstellen. Ein analoges Konstrukt bei den semantischen Rollen steht aus.

Die Schnittmenge von A und B ist definiert durch: $A \cap B = \{x \mid x \in A \text{ und } x \in B\}$; es ist die Menge der Elemente, die sich sowohl in A als auch in B befinden. Der Begriff *Amphibienfahrzeug* bezeichnet die Schnittmenge aus *Land- und Wasservehikel*. Auch hier steht ein analoges Konstrukt bei den semantischen Rollen aus.

Die Differenzmenge von A und B ist definiert durch: $A \setminus B = \{x \mid x \in A \text{ und } x \notin B\}$; also die Menge der Elemente, die Teil von A, aber nicht von B sind. Eine Differenzmenge wird immer gebildet, wenn eine Gruppe von einer anderen unterschieden wird. Es soll hier Vohidova (vgl. 2016: 93-99) gefolgt werden, die zwei semantische Relationen ansetzt, die mittels der Differenzmenge dargestellt werden können: Komplementarität und Kohyponymie. Komplementarität liegt vor, wenn ein Begriff die Verneinung des anderen impliziert und umgekehrt. Als Beispiel sollen *tot* und *lebendig* dienen. Wenn etwas tot ist, dann ist es nicht lebendig und umgekehrt. Neben der semantischen Relation der Komplementarität steht die Kohyponymie: Zwei Begriffe sind Kohyponyme, wenn sie komplementär sind, aber ein gemeinsames Hyperonym haben. Bspw. sind *Hund* und *Katze* Kohyponyme in Bezug zu *Säugetier*. Es handelt sich formal um Differenzmengen, die innerhalb einer Obermenge gebildet werden.

Abschließend sei eine spezielle Art von Relationen eingeführt: Funktionen. Es seien zwei Mengen X und Y – dann ordnet eine Funktion f jedem Element aus X genau ein Element aus Y zu. Man schreibt $f : X \rightarrow Y$ und nennt

X die Definitionsmenge und Y die Zielmenge (vgl. Partee u.a. 1990: 30 und Lohnstein 2011: 28). Auch in der natürlichen Sprache lassen sich Beispiele für Aussagen finden, die sich als Funktionen formalisieren lassen, bspw. Prädikate, die das Alter einer Person angeben: In *Peter ist 29 Jahre alt* wird der Person *Peter* das Alter 29 zugewiesen.

Das Unterkapitel hat den Begriff der *Relation* mengentheoretisch konkretisiert und aufgezeigt, wie natürlich-sprachliche Ausdrücke als Relationen im mengentheoretischen Sinne dargestellt werden können. Außerdem wurden formale Eigenschaften von Relationen besprochen, die im Kontext von Ontologien nutzbar gemacht werden können.

3.2 Semantische Rollen und Pertinenzrelationen

Diese Arbeit wird sich nicht nur mit Relationen beschäftigen, die zwischen Mengen bestehen, sondern auch mit solchen, die zwischen den Elementen zweier Mengen bestehen. Da solche Relationen nicht direkt zwischen Begriffen bestehen, sind sie strenggenommen nicht zu den semantischen Relationen zu zählen und finden sich daher auch seltener in semantischen Darstellungen. Auch mengentheoretisch unterscheiden sie sich in ihrer Formalisierung von den semantischen Relationen. Ontologisch sind sie dennoch von Bedeutung. Zur Veranschaulichung soll die Relation zwischen einem Auto und seinen Reifen dienen. Um eine solche Relation zu formalisieren, werden zwei weitere Konstrukte der Mengenlehre benötigt: das kartesische Produkt und das Tupel. Das kartesische Produkt der Mengen A und B, geschrieben $A \times B$, ist eine Menge, die entsteht, wenn die beiden Mengen derart miteinander kombiniert werden, dass jedes Element der einen Menge, die Definitionsmenge genannt wird, einmal jedem Element der anderen Menge, die Zielmenge genannt wird, zugeordnet wird. Die Elemente des kartesischen Produktes, die bei dieser Operation entstehen, nennt man Tupel, geschrieben $\langle a, b \rangle$ für das Tupel der Elemente a und b. Bei Tupeln ist im Gegensatz zu Mengen die Häufigkeit und Reihenfolge der Elemente von Bedeutung, weswegen sie auch

geordnete Paare genannt werden. Formal wird das kartesische Produkt folgendermaßen definiert:

$$A \times B := \{ \langle a, b \rangle \mid a \in A, b \in B \}$$

Die Relation zwischen Autos und Reifen lässt sich nun als Untermenge des kartesischen Produktes beschreiben, das sich aus den Mengen Auto und Reifen ergibt – Untermenge, weil evtl. nicht jede Paarung von Auto und Reifen in dieser Relation enthalten ist (vgl. Partee u.a. 1990: 28-30). Alle Relationen, die nicht direkt zwischen Mengen bestehen und keine Funktionen sind, können auf diese Weise dargestellt werden.

Relationen können bestimmte formal-beschreibbare Eigenschaften haben, die in Ontologien nutzbar gemacht werden, indem durch sie Aussagen abgeleitet werden können. Diese Eigenschaften sollen an der Menge der Menschen beispielhaft dargestellt werden: Eine Relation ist symmetrisch, wenn für jedes geordnete Paar $\langle x, y \rangle$ in M das Paar $\langle y, x \rangle$ ebenfalls existiert. Ein natürlich-sprachliches Beispiel wäre *istVerwandtMit*; wenn Person A verwandt mit Person B ist, dann ist Person B ebenfalls verwandt mit Person A; existiert für kein Paar $\langle x, y \rangle$ das Paar $\langle y, x \rangle$ ist die Relation antisymmetrisch (vgl. Partee u.a. 1990: 40 und Lohnstein 2011: 23). Wird in einer Ontologie angegeben, dass eine Relation symmetrisch ist, kann ein Computer aus der Aussage, dass A mit B verwandt ist, die Aussage, dass B mit A verwandt ist, schlussfolgern. Eine Relation ist reflexiv in einer Menge M , wenn sie alle geordneten Paare der Form $\langle x, x \rangle$ enthält. Ein natürlich-sprachliches Beispiel wäre *hatDenSelbenGeburtstagWie*; Person A hat am selben Tag Geburtstag wie Person B (vgl. Partee u.a. 1990: 39-40 und Lohnstein 2011: 23). Eine Relation ist transitiv, wenn für alle geordneten Paare $\langle x, y \rangle$ und $\langle y, z \rangle$ das Paar $\langle x, z \rangle$ ebenfalls in M existiert. Ein Beispiel wäre *istEinVorfahreVon*; wenn Person A der Vorfahre von Person B und Person B der Vorfahre von Person C ist, dann ist Person A der Vorfahre von Person C (vgl. Partee u.a. 1990: 41 und Lohnstein 2011: 24). Zu einer zweistelligen Relation kann des Weiteren ein inverses Gegenstück definiert werden, das alle Tupel umkehrt. Besteht eine Relation aus

geordneten Paaren der Form $\langle x,y \rangle$, besteht ihr Inverses aus den geordneten Paaren $\langle y,x \rangle$ (vgl. Partee u.a. 1990: 29).

In dieser Arbeit werden Relationen, die zwischen den Elementen zweier Mengen bestehen, in zwei Gruppen aufgeteilt, den semantischen Rollen und den Pertinenzrelationen, die im Folgenden besprochen werden.

3.2.1 Semantische Rollen

Eine der zentralen Thesen dieser Arbeit lautet, dass ein wesentlicher Anteil der in einer Ontologie hinterlegten Informationen sich als semantische Rollen darstellen und formalisieren lässt. Dieses Kapitel will das Konstrukt der semantischen Rollen vorstellen, während in Kapitel 6.3 die Annotationstypen vorgestellt werden, die sich aus den Überlegungen ergeben, und in Kapitel 8 die empirische Überprüfung dieser These erfolgt.

Das Konstrukt der semantischen Rolle weist Gemeinsamkeiten mit dem der semantischen Relation auf: Bei beiden wird eine Verbindung zwischen zwei sprachlichen Einheiten postuliert (vgl. Engel 1996: 225-227). Ein wesentlicher Unterschied besteht darin, dass semantische Relationen unabhängig von einer Verbalisierung bestehen, während semantische Rollen sich nur bzgl. eines Prädikats bestimmen lassen (vgl. Primus 2012: 5). Die semantische Relation, die in (2a) verbalisiert wird, besteht auch unabhängig von diesem Satz, während die Relation zwischen *Hund* und *Mann* in (2b) keinen grundsätzlichen Zusammenhang zwischen den beiden Begriffen darstellt:

- (2) a. *Ein Hund ist ein Säugetier.*
- b. *Ein Hund beißt einen Mann.*

Semantische Rollen sind ein Konstrukt, das die semantischen Verhältnisse zwischen den Partizipanten eines Satzes beschreibt, indem es Kategorien wie Agens für handelnde Entitäten, Patiens für von der Handlung unmittelbar betroffene Entitäten oder Benefaktiv für Entitäten, denen durch die Handlung ein Vor- oder Nachteil entsteht, aufstellt (vgl. Blake 1930: 34 und Gruber

1965: 7-8).³ Die folgenden Beispiele sollen das Konstrukt veranschaulichen:

- (3) a. *Peter zerbricht die Vase.*
- b. *Die Vase zerbricht.*
- c. *Peter zerbricht die Vase zu Marias Unmut.*

In den Beispielen (3) stellt *Peter* jeweils den Agens, *die Vase* den Patiens, und *Maria* den Benefaktiv dar. Dass in (3a) und (3c) *Peter* das Subjekt bildet, während in (3b) *Die Vase* das Subjekt bildet, zeigt, dass die semantischen Rollen nicht mit den Kategorien von Subjekt und Objekt zusammenfallen.

Semantische Rollen stellen wie die semantischen Relationen nun einmal Relationen dar (vgl. Primus 2012: 2). Dies wird bspw. daran deutlich, dass nicht nach der semantischen Rolle eines isolierten Elementes gefragt werden kann: Nach der Rolle von *Vase* kann nicht gefragt werden, ohne einen Satz als Kontext anzugeben. Das Verb *zerbrechen* aus (3a) ließe sich als eine Menge von Tupeln darstellen, bei denen das erste Element derjenige, der etwas zerbricht, ist und das zweite Element das, was zerbrochen wird. Die intransitive Variante in Beispiel (3b) kann als einstellige Relation dargestellt werden (vgl. Kapitel 3.1).⁴

Das Verb gibt maßgeblich Auskunft darüber, welche semantischen Rollen in einem Satz anzusetzen sind. Es legt bestimmte Eigenschaften seiner Argumente nahe, anhand derer man versucht, semantische Rollen zu definieren (vgl. Dowty 1989: 77). Das Verb *kaufen* legt bspw. einen willentlich handelnden Akteur, den Käufer, neben dem Verkäufer, dem Geld und der Ware nahe. Dieses Merkmal, dass jemand willentlich handelt, wird gemeinhin als ein notwendiges Merkmal für die Rolle des Agens gesehen (vgl. Dowty 1991: 550). Das Kaufereignis kann aus unterschiedlichen Perspektiven betrachtet werden, die jeweils durch ein bestimmtes Verb ausgedrückt werden (vgl. Fillmore 1977b: 12, 74). Hieran zeigt sich, dass semantische Rollen nicht diskret,

³Semantische Rollen sind konzeptionell eng verflochten mit Frames. Frames werden kurz in Kapitel 5.3 als Alternative zur Merkmalssemantik vorgestellt.

⁴Eine systematische Darstellung, wie sich Verben und andere Wortarten mengentheoretisch interpretieren lassen, folgt in Kapitel 3.3.

sondern skalar oder prototypisch zu verstehen sind (vgl. Rissmann und Majid 2019: 1853): So handelt der Verkäufer auch willentlich, allerdings etwas weniger, da zumindest aus der Perspektive von *kaufen* im Gegensatz zu *verkaufen* der Käufer die Handlung initiiert. Ein Argument kann also bis zu einem gewissen Grad agentivisch, d.h. willentlich handelnd, oder affiziert, d.h. von der Handlung betroffen, sein, oder allgemeiner: Ein Argument kann einer bestimmten semantischen Rolle mehr oder weniger entsprechen.

Dies bedeutet, dass die Forderung, dass jeder semantischen Rolle genau ein syntaktisches Argument und jedem syntaktischen Argument genau eine semantische Rolle entspricht (vgl. Chomsky 1993: 36), nicht aufrecht erhalten werden kann (vgl. Jackendoff 1987: 381-383). Wenn zwei Argumenten dieselbe Rolle zukommt, wird von *Rollendispersion* gesprochen; wenn mehr als eine Rolle auf ein syntaktisches Argument kommt, wird von *Rollenkumulation* gesprochen (vgl. Primus 2012: 5).

Welche semantischen Rollen in dieser Arbeit schließlich angesetzt wurden und wie diese Typologie erarbeitet wurde, soll in Kapitel 6.3 erläutert werden, da für diese Typologie zunächst der Ontologiebegriff diskutiert und semantik-theoretisch verortet werden muss. Dieses Kapitel widmet sich im Weiteren der Frage, welche Bedeutung semantische Rollen für die Erstellung von Ontologien haben können. Dafür soll das Kaufereignis aus ontologischer Sicht betrachtet werden. Ein *Käufer* kann als eine Person, die für Geld Waren erlangt, definiert werden. Ein genauerer Blick auf diese Definition offenbart, dass zwischen den Begriffen *Käufer* und *Person* die semantische Relation der Hyperonymie angesetzt werden kann. Das Verhältnis zwischen *Käufer* und *Geld*, sowie zwischen *Käufer* und *Ware* kann durch semantische Rollen beschrieben werden: Bzgl. *Geld* ist *Käufer* ein Agens, da er es willentlich abgibt, und bzgl. *Ware* ist er ein Benefaktiv, da er diese zu seinem Vorteil erhält. Die Beobachtung, dass Teile einer Definition sich als semantische Rollen identifizieren lassen, ist für die automatische Textverarbeitung oder Ontologierstellung von großer Bedeutung. Unabhängig davon, welche Art von Algorithmus

genutzt werden soll, um aus einem natürlich-sprachlichen Text ontologisch-relevante Informationen zu extrahieren, muss Einigkeit darüber herrschen, welche Informationen überhaupt relevant sind und wie diese benannt werden sollen. Dass dabei die Relationen der Mengenlehre bzw. semantische Relationen relevant sind, ist hinlänglich bekannt (vgl. Stuckenschmidt 2011: 21), wie aber bspw. das Verhältnis zwischen Käufer und Geld im Kaufereignis zu benennen ist, steht zur Diskussion. Das Konstrukt der semantischen Rollen bietet hier m.E. einen vielversprechenden Ansatz.

Bestimmte Teile einer Definition lassen sich also als semantische Relationen identifizieren und andere Teile als semantische Rollen. Wohlgermerkt wird hier nicht behauptet, dass sich alle Definitionen auf diese Art auflösen lassen, und schon gar nicht, dass sich alle Begriffe auf diese Art definieren lassen; in diesem Einführungskapitel genügt die Feststellung, dass sich bestimmte Definitionsstrategien mit den Konstrukten der semantischen Relationen und Rollen beschreiben lassen. Ebenfalls sei an dieser Stelle erwähnt, dass semantische Rollen nicht ausschließlich durch Verben, sondern auch durch andere Konstruktionen evoziert werden können (vgl. Philipp u.a. 2017: 1), worauf im Hauptteil zurückzukommen sein wird.

Es sei angemerkt, dass Ontologien sich teils erheblich bzgl. der Informationen, die für sie relevant sind, unterscheiden. Genügt es in einer Ontologie zu definieren, dass ein Auto u.a. aus vier Reifen besteht, ist für eine andere Ontologie wichtig, um was für eine Art von Bestandteil es sich bei Reifen handelt: eine Komponente, ein Mitglied, ein Substrat etc. Diese Arbeit befasst sich mit sprachlichen Mustern, wie diese bezeichnet werden können und mit welchen ontologischen Relationen sie korrelieren. Für die Erstellung einer Ontologie können diese Erkenntnisse eine Unterstützung sein, sie werden aber im Einzelnen nicht individuelle Entscheidungen ersetzen können, wie welche Informationen in eine Ontologie zu übertragen sind.

Im Gegensatz zu den semantischen Relationen herrscht bei den semantischen Rollen große Uneinigkeit darüber, wie viele semantische Rollen anzu-

setzen sind (Levin und Rappaport 2005: 35-50 und Jurafsky und Martin 2021: 2-3). Manche Autor*innen gehen so weit, zu sagen, dass semantische Rollen das diffuseste Konstrukt der Linguistik seien (vgl. Newmeyer 2010: 689). Semantische Rollen zu definieren oder ein Verfahren zu entwickeln, das eine Typologie von semantischen Rollen produziert, ist in jedem Fall ein mehr als schwieriges Unterfangen, zu dem verschiedene Ansätze existieren, die kurz angerissen werden, um zu verdeutlichen, warum ein eigener Ansatz für notwendig befunden wurde:

Eine der gängigsten Methoden sieht vor, die Rollen über Implikationen zu definieren (vgl. Dowty 1991: 552). Bspw. impliziert die Rolle des Agens, dass die Entität, die diese Rolle zugewiesen bekommt, willentlich handelt. Aber auch diese Methode führt zu keiner abgeschlossenen Typologie von Rollen, sondern verlagert die Problematik auf die Frage, welche Implikationen für eine Rolle anzusetzen sind (vgl. Dowty 1991: 553-554).

Einige Autor*innen stellen Argumentstrukturen mit der *Lexical Decomposition Grammar* (LDG) dar, einer formalen, um den Lambda-Operator erweiterten Sprache (vgl. Joppen und Wunderlich 1995: 123-128 und Wunderlich und Lakämper 2001: 382). Dieser Ansatz ist aber ebenfalls nicht imstande, eine abschließende Typologie vorzuschlagen, da sogenannte Basisverben wie CAUSE und BECOME in der LDG verwendet werden, ohne dass geklärt ist, wie viele von diesen anzusetzen sind. Die Frage wird also lediglich verlagert. In der Psycholinguistik gibt es Ansätze, das Verfahren des *primings* zu nutzen, um die Menge der semantischen Rollen einzugrenzen (vgl. Ziegler und Snedeker: 2018). Die Ergebnisse sind vielversprechend, aber eine systematische Untersuchung über eine große Anzahl an Rollen steht noch aus.

In der vorliegenden Studie wurden die semantischen Rollen auf der Grundlage des Korpus bestimmt. Dabei wurden zwei heuristische Kriterien hinzugezogen. Als erstes Kriterium werden dabei die beiden Dimensionen Agentivität und Affiziertheit verwendet. Die beiden Rollen *Agent* und *Patient*⁵ lassen sich

⁵Im Folgenden werden alle semantischen Rollen der Einheitlichkeit wegen auf Englisch und

als Endpunkte zweier Achsen auf einem Koordinatensystem denken (vgl. Abbildung 1): *Agent* liegt am Endpunkt der Achse agentivisch, aber im Ursprung der Achse affiziert. *Patient* ist das genaue Gegenteil. Die anderen Rollen lassen sich in dem Koordinatensystem, das sich aus diesen Dimensionen ergibt, verorten: Der *Beneficiant* liegt bspw. bei agentivisch im Ursprung und bzgl. der Affiziertheit unter dem *Patient*, da er von der Handlung etwas weniger betroffen ist. *Force*, eine Rolle, die noch eingehender besprochen wird, liegt bzgl. der Affiziertheit im Ursprung und bzgl. der Agentivität knapp vor dem *Agent*. *Experiencer* und *Stimmulus* sind beide zu ungefähr gleichen Teilen agentivisch und affiziert.

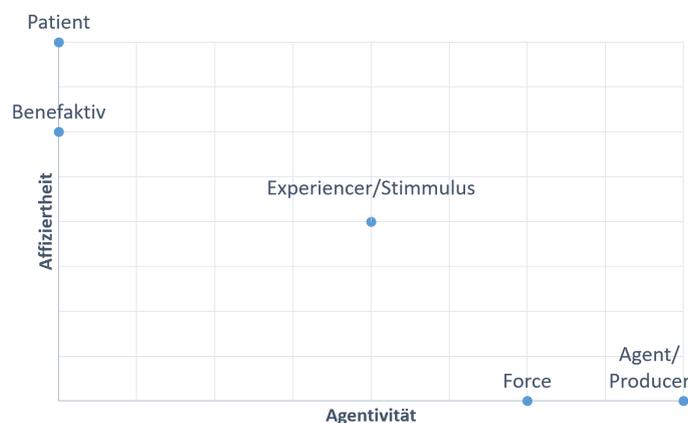


Abbildung 1: Verortung einiger semantischer Rollen in einem zweidimensionalen Koordinatensystem

Abbildung 1 visualisiert erneut, dass semantische Rollen nicht als sich gegenseitig ausschließende Konstrukte zu begreifen sind (vgl. Rozwadowska 1989 und Dowty 1991: 551), da bspw. *Experiencer* und *Stimmulus* an gleicher Stelle verortet sind. Um diesen Aspekt etwas stärker auszuführen, sei noch ein weiteres Beispiel gegeben.

- (4) a. *Der Mann riecht die Blume.*
 b. *Der Mann betrachtet die Blume.*

In den Beispielen (4a) und (4b) würde *Der Mann* jeweils mit der Rolle des *Experiencer* und *die Blume* jeweils mit der Rolle *Stimmulus* belegt werden.

kursiv geschrieben. Dadurch wird auch ausgedrückt, dass es sich um einen Typen der in dieser Arbeit entworfenen Annotationstypologie handelt.

Beide Rollen beinhalten agentivische Anteile. Etwas zu riechen, stellt eine Handlung dar, auch wenn diese vom Handelnden nicht unbedingt bewusst initiiert wird. Aber auch einen Sinnesreiz verursachen, in diesem Fall gerochen werden, kann agentivisch interpretiert werden, wie die psycholinguistische Forschung belegt (vgl. Primus 2012: 39). Dabei können die Anteile unterschiedlich stark ausgeprägt sein: Der agentivische Anteil in Beispiel (4b) liegt stärker beim *Experiencer*. Als zweites Kriterium dient der Satzrahmen. Dazu folgendes Beispiel: Abbildung 1 verortet *Agent* und *Producer* an gleicher Stelle; über den Satzrahmen lassen sie sich voneinander differenzieren:

- (5) a. *Ein KFZ-Mechaniker repariert Autos.*
b. *Ein KFZ-Konstrukteur produziert Autos.*
c. *Ein Bäcker backt Brötchen.*

Bei *KFZ-Mechaniker* und *KFZ-Konstrukteur* handelt es sich um zwei verschiedene Konzepte (was nicht bedeuten soll, dass nicht eine Person beide Aufgaben erfüllen kann). Die Beispiele (5a) und (5b) legen nahe, dass es sinnvoll ist, zwei unterschiedliche Rollen, *Agents* und *Producer*, anzunehmen, um diesen Unterschied zu fassen: Während das Handeln eines *Agents* auf etwas Bestehendes einwirkt (den *Patient*), bringt der *Producer* etwas überhaupt erst in die Existenz (das *Product*). Das Beispiel (5c) hingegen zeigt, dass nicht für alle denkbaren Konstellationen unterschiedliche semantische Rollen angenommen werden müssen, da das Subjekt in diesem Satz ebenfalls als *Producer* und das Objekt als *Product* annotiert werden kann. Das Konstrukt der semantischen Rollen bewahrt sich demnach eine Generalität, die notwendig ist, um Gemeinsamkeiten zwischen verschiedenen semantischen Zusammenhängen aufzudecken.

Nicht alle Argumente, die in einer Relation stehen können, müssen realisiert werden. Bspw. lässt sich ein Bäcker als jemand definieren, der Brötchen herstellt, wobei das Material, aus dem Brötchen hergestellt werden (*Ressource*), nicht zwingend genannt werden muss, obwohl dieses ontologisch zu dem Prozess des Brötchenbackens gehört. Es gilt außerdem festzuhalten, dass Ar-

gumente nicht immer in der gleichen Relation stehen. So kann *Agent* mit *Patient*, aber auch mit *Beneficient* in Relation stehen. Die verschiedenen Kombinationsmöglichkeiten auszuloten, ist eine der Aufgaben, die in Kapitel 8 angegangen wird. Welche semantischen Rollen in dieser Arbeit angesetzt wurden, d.h. als Annotationstypen verwendet werden, wird in Kapitel 6.3 dargelegt.

3.2.2 Pertinenzrelationen

Meronymie bezeichnet die Relation zwischen einem Teil, dem Meronym, und dem Ganzen, dem es angehört, dem Holonym (vgl. Smith 2004: 8-9, Ziem 2008: 41 und Stuckenschmidt 2011: 60-63). In dieser Relation stehen bspw. die Begriffe *Auto* und *Reifen*. Das Inverse dieser Relation wird Holonymie und beide Relationen zusammen werden Teil-Ganzes-Relation genannt. Bei der Arbeit mit dem Korpus wurden sechs verschiedene Untertypen dieser Relation identifiziert, die sich – abgesehen von einer fehlenden Entsprechung für *feature-activity* – grundsätzlich an der Typologie von Stuckenschmidt (2011: 61) anlehnen. Die einzelnen Untertypen seien kurz mit Beispiel aufgeführt:

1. Komponente-Objekt (Reifen – Auto)
2. Objekt-Ansammlung (Baum – Wald)
3. Material-Material (Zucker – Kuchenteig)
4. Material-Objekt (Aluminium – Flugzeug)
5. Ort-Gebiet (Bonn – Deutschland)
6. Teilprozess-Prozess (Abheben – Fliegen)

Eine eingehende Besprechung der einzelnen Typen erfolgt in Kapitel 9 anhand der Belege aus dem Korpus. An dieser Stelle seien sie nur aufgeführt, um zu verdeutlichen, dass die Relation vielfältig ist und sich nicht allein auf materielle Verhältnisse beschränkt, wie das Beispiel der Begriffe *Auto* und *Reifen* evtl. vermuten ließ. Die Teil-Ganzes-Relation nimmt einen Platz zwischen den semantischen Relationen und den semantischen Rollen ein. Sie ist keine semantische Relation, da sie nicht zwischen den Klassen an sich besteht,

sondern ein tatsächliches Verhältnis zwischen den Instanzen dieser Klassen darstellt. Die Relation lässt sich aber auch nicht zu den semantischen Rollen zählen, da sie unabhängig von einer Verbalisierung besteht. Die Klasse, der das Holonym angehört, scheint im Übrigen, einen Einfluss auf das Meronym zu haben. So bezeichnet man einen bestimmten Teil eines Pinguins als *Flügel*, weil der Pinguin ein Vogel ist, obwohl *Flosse* unter bestimmten Aspekten eine passendere Bezeichnung wäre, da ein Pinguin diese Extremität nicht zum Fliegen, aber sehr wohl und sehr gewandt zum Schwimmen und Tauchen benutzt. Flossen sind aber wiederum Teile eines Fisches. Hieran zeigt sich, dass zwar verschiedene Typen von Relationen angesetzt werden können, diese aber evtl. nicht isoliert voneinander im Sprachbewusstsein repräsentiert sind.

Possessor und *Possession* stehen genau auf der Grenze zwischen semantischen Rollen und Pertinenzrelationen. Sie können als semantische Rollen begriffen werden, da einem *Possessor*, also jemandem, der über etwas verfügt, unzweifelhaft ein agentivisches Potential zuzusprechen ist und dementsprechend der *Possession* ein bestimmter Grad an Affiziertheit. Bestimmte Zusammenhänge lassen sich allerdings aus beiden Perspektiven betrachten. *Arm* ist zweifelsohne Teil von *Körper*, jedoch verfügt eine Person auch über ihre Arme. Wegen solcher Fälle wurde entschieden, *Possessor-Possession* zu den Pertinenzrelationen zu zählen.

3.3 Aussagen- und Prädikatenlogik

Bisher wurde ausgeführt, dass zwischen Begriffen Relationen verschiedener Art bestehen und dass diese Relationen durch die Konstrukte der Mengenlehre beschrieben und damit formalisiert werden können. Der Sinn und Nutzen einer Ontologie erschöpft sich aber im besten Fall nicht darin, Klassen und ihre Relationen darzustellen, sondern die dabei entstehende Struktur auf Konsistenz zu prüfen und gegebenenfalls neue Relationen abzuleiten (vgl. Horridge 2005: 57). Damit dies möglich ist, müssen Relationen als Aussagen dargestellt werden, wozu ein Wahrheitsbegriff benötigt wird, den die Logik be-

reitstellt. Dieses Unterkapitel soll in die grundlegenden Konstrukte der Logik einführen, namentlich die Aussage, Quantoren und Prädikate. Eine formale Definition der einzelnen Konstrukte erübrigt sich, weil am Ende des Oberkapitels alle Konstrukte in der Sprache $SHOIN(D)$ zusammengeführt werden, welche formal ausformuliert wird.

Die Logik ergänzt das Konstrukt der Ontologie dadurch, dass sich die Wahrheitswerte der Logik in 0 und 1 übersetzen lassen, wodurch sich Inhalte in maschinenlesbarer Form darstellen lassen. Eine Ontologie ist eine Form elektronisch implementierter Logik (vgl. Barwise und Etchemendy 2005: 2). Der ursprüngliche Zweck der Logik liegt allerdings im logischen Schlussfolgern, also im Ableiten von Wahrheitswerten aus als wahr oder falsch akzeptierten Aussagen. Ob ein bestimmter Sachverhalt wahr oder falsch ist, hat streng genommen keine Relevanz für die Logik; sie beschäftigt sich damit, was folgt, wenn bestimmte Aussagen als wahr angenommen werden (vgl. Kamp und Reyle 1993: 13-14 und Beckermann 2014: 34). Auch für diese Arbeit ist die „Folgerungsbeziehung“ zwischen Prämisse und Konklusion von besonderem Interesse.

Die Bausteine der Aussagenlogik sind atomare Aussagen und die beiden Wahrheitswerte. Atomare Aussagen sind Aussagen, die sich nicht weiter zerlegen lassen und denen jeweils nur ein Wahrheitswert zugeordnet werden kann; eine Aussage kann nur wahr oder falsch sein (*tertium non datur*). Als Beispiele mögen zunächst folgende Sätze dienen, die semantische Relationen realisieren:

- (6) a. *Ein Hund ist ein Säugetier.*
- b. *Der Mond besteht aus Käse.*
- c. *Menschen sind sterblich.*

Solche Aussagen werden formalisiert, indem sie in Objekte transferiert werden, die keine innere Struktur mehr besitzen (vgl. Urchs 1993: 3) und denen Wahrheitswerte zugeordnet werden (vgl. Carstensen u.a. 2010: 31-32). Die

Beispiele (6a), (6b) und (6c) ließen sich z. B. als die Aussagen A, B und C darstellen, wobei – nach allen derzeitigen Kenntnissen der Wissenschaft – B den Wahrheitswert *falsch* und A und C den Wert *wahr* zugewiesen bekämen. Dies ist jedoch, wie bereits erwähnt, weniger eine Frage der Logik. Für Logiker*innen von Interesse ist die Frage, welche Wahrheitswerte sich ergeben, wenn man Aussagen miteinander kombiniert. Dazu stehen die Junktoren zur Verfügung, die in (1) dargestellt sind. Streng genommen ist die Negation kein Junktor, da sie keine Aussagen miteinander verknüpft; einfachheitshalber wird sie aber meist – und so auch hier – wie ein Junktor behandelt (vgl. Carstensen u.a. 2010: 35).

Junktor	Sprechweise	Bezeichnung	Form	Beispiel
\neg	nicht	Negation	$\neg A$	Der Mond besteht nicht aus Käse.
\wedge	und	Konjunktion	$A \wedge B$	Der Mond besteht aus Käse, und alle Menschen sind sterblich.
\vee	oder	Disjunktion	$A \vee B$	Der Mond besteht aus Käse, oder alle Menschen sind sterblich.
\rightarrow	wenn, dann	Implikation	$A \rightarrow B$	Wenn der Mond aus Käse besteht, dann sind alle Menschen sterblich.
\leftrightarrow	genau dann, wenn	Äquivalenz	$A \leftrightarrow B$	Genau dann wenn der Mond aus Käse besteht, sind alle Menschen sterblich.

Tabelle 1: Die Junktoren der Aussagenlogik

Tabelle 2 ist eine Wahrheitstafel, an der abgelesen werden kann, in welchen Konstellationen die Junktoren zu welchen Wahrheitswerten führen. Es ist ersichtlich, dass die Junktoren bzgl. ihrer Bedeutung nicht in einem 1:1-Verhältnis mit ihrem natürlich-sprachlichen Gegenstück stehen:

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftrightarrow B$
W	W	F	W	W	W	W
W	F	F	F	W	F	F
F	W	W	F	W	W	F
F	F	W	F	F	W	W

Tabelle 2: Wahrheitstabelle

Logische Aussagen stellen die Grundbausteine einer Ontologie dar und semantische Relationen – wie die Beispiele zeigen – lassen sich auch als lo-

gische Aussagen darstellen. Die Aussagenlogik bildet die Grundlage für das logische Schlussfolgern, das Inferieren. Im Kontext von Ontologien wird der Begriff *Reasoning* verwendet, der jedoch nichts anderes als *automatisiertes Inferieren* bzw. *Schlussfolgern* meint.⁶

Letztlich werden beim Inferieren Äquivalenzen gesucht, also zwei Aussagen, die unter den gleichen Bedingungen wahr sind. Tabelle 3 zeigt die wichtigsten Äquivalenzen der Aussagenlogik (vgl. Partee u.a. 1990: 112). In der rechten Spalte stehen die äquivalenten Aussagen: Pro Zeile stehen zwei Aussagen, die über Äquijunktion miteinander verbunden sind.

Name	Form
Idempotenzgesetze	a) $(A \vee A) \leftrightarrow A$ b) $(A \wedge A) \leftrightarrow A$
Assoziativgesetze	a) $((A \vee B) \vee C) \leftrightarrow (A \vee (B \vee C))$ b) $((A \wedge B) \wedge C) \leftrightarrow (A \wedge (B \wedge C))$
Kommutativgesetze	a) $(A \vee B) \leftrightarrow (B \vee A)$ b) $(A \wedge B) \leftrightarrow (B \wedge A)$
Distributivgesetze	a) $(A \vee (B \wedge C)) \leftrightarrow ((A \vee B) \wedge (A \vee C))$ b) $(A \wedge (B \vee C)) \leftrightarrow ((A \wedge B) \vee (A \wedge C))$
Identitätsgesetze	a) $(A \vee \text{Falsch}) \leftrightarrow A$ b) $(A \vee \text{Wahr}) \leftrightarrow \text{Wahr}$ c) $(A \wedge \text{Falsch}) \leftrightarrow \text{Falsch}$ d) $(A \wedge \text{Wahr}) \leftrightarrow A$
Komplementgesetze	a) $(A \vee \neg A) \leftrightarrow \text{Wahr}$ b) $\neg \neg A \leftrightarrow A$ (doppelte Negation) c) $(A \wedge \neg A) \leftrightarrow \text{Falsch}$
DeMorgans Gesetze	a) $\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$ b) $\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B)$
Konditionale Gesetze	a) $(A \rightarrow B) \leftrightarrow (\neg A \vee B)$ b) $(A \rightarrow B) \leftrightarrow (\neg A \rightarrow \neg B)$ (Kontraposition) c) $(A \rightarrow B) \leftrightarrow \neg(A \wedge \neg B)$
Bikonditionale Gesetze	a) $(A \leftrightarrow B) \leftrightarrow ((A \rightarrow B) \wedge (B \rightarrow A))$ b) $(A \leftrightarrow B) \leftrightarrow ((\neg A \wedge \neg B) \vee (A \wedge B))$

Tabelle 3: Wichtige Äquivalenzen der Aussagenlogik

Im engeren Sinne wird beim Inferieren aus einer bestimmten Anzahl von Prämissen eine Konklusion abgeleitet, die unter jeder Bedingung wahr ist, unter

⁶Da sich die Arbeit nicht mit den algorithmischen Aspekten des logischen Schlussfolgerns beschäftigt, wird im weiteren Verlauf der Begriff *Inferieren* statt *Reasoning* verwendet.

der auch die Prämissen wahr sind (vgl. Partee u.a. 1990: 115). Die folgenden Inferenzregeln sind die herkömmlichen Regeln der Aussagenlogik (vgl. Partee u.a. 1990: 118):

Name und Abkürzung	Form	Beispiel
Modus Ponens (M.P.)	$A \rightarrow B$ A <hr/> B	Wenn ein Hund ein Säugetier ist, gebärt ein Hund lebend. Ein Hund ist ein Säugetier. <hr/> Ein Hund gebärt lebend.
Modus Tollens (M.T.)	$A \rightarrow B$ $\neg B$ <hr/> $\neg A$	Wenn ein Reptil ein Säugetier ist, gebärt ein Reptil lebend. Ein Reptil gebärt nicht lebend. <hr/> Ein Reptil ist kein Säugetier.
Hypothetischer Syllogismus (H.S.)	$A \rightarrow B$ $B \rightarrow C$ <hr/> $A \rightarrow C$	Wenn ein Hund ein Säugetier ist, ist er ein Wirbeltier. Wenn ein Hund ein Wirbeltier ist, hat er eine Wirbelsäure. <hr/> Wenn ein Hund ein Säugetier ist, hat er eine Wirbelsäure.
Disjunkter Syllogismus (D.S.)	$A \vee B$ $\neg A$ <hr/> B	Eine Gitarre hat ein Schallloch oder einen elektrischen Tonabnehmer. Peters Gitarre hat keinen elektrischen Tonabnehmer. <hr/> Peters Gitarre hat ein Schallloch.
Simplification (Simp.)	$A \wedge B$ <hr/> A	Hunde haben Fell, und Vögel haben Federn. <hr/> Hunde haben Fell.
Konjunktion (Conj.)	A B <hr/> $A \wedge B$	Hunde haben Fell. Vögel haben Federn. <hr/> Hunde haben Fell, und Vögel haben Federn.
Addition (Add.)	A <hr/> $A \vee B$	Gitarren haben ein Schallloch. <hr/> Gitarren haben ein Schallloch oder einen elektrischen Tonabnehmer.

Tabelle 4: Allgemeine Inferenzregeln der Aussagenlogik

Mit diesen Regeln kann geprüft werden, ob eine bestimmte Aussage aus anderen Aussagen abgeleitet werden kann. Dies soll an einem rein formalen Beispiel demonstriert werden: Es soll gezeigt werden, dass $(A \rightarrow B)$ aus den Prämissen $(A \rightarrow (B \vee C))$ und $\neg C$ folgt. Im Folgenden werden die Aussagen durchnummeriert. Rechts von den abgeleiteten Aussagen steht die Nummer der Aussage, aus der abgeleitet wurde, zusammen mit der jeweiligen ange-

wendeten Regel.

- (7) a. $(A \rightarrow (B \vee C))$
- b. $\neg C$
- c. $\neg A \vee (B \vee C)$ (a: Konditionales Gesetz)
- d. $(\neg A \vee B) \vee C$ (c: Assoziativgesetz)
- e. $\neg A \vee B$ (b, d: D.S.)
- f. $A \rightarrow B$ (e: Konditionales Gesetz)

Am Anfang des Beweises stehen die zwei Prämissen (a., b.). Dann wird aus Aussage a mit Hilfe des Konditionalen Gesetzes Aussage c abgeleitet. Auf das Ergebnis wird das Assoziativgesetz angewendet. Dann kann aus den Aussagen b. und d. Aussage e. durch einen disjunkten Syllogismus und schließlich durch das konditionale Gesetz die zu beweisende Aussage f. abgeleitet werden.

Für Ontologien gibt es viele verschiedene *Reasoner*, d. h. in unterschiedlichen Programmiersprachen verfasste Algorithmen, die genutzt werden können, um aus einer Ontologie neues Wissen zu generieren. Im Kontext dieser Arbeit ist jedoch nur wichtig, dass diese *Reasoner* nichts anderes versuchen, als die genannten Äquivalenzen und Inferenzregeln anzuwenden.

Die Prädikatenlogik kann als eine Erweiterung der Aussagenlogik um die Konstrukte der Quantoren und des Prädikates verstanden werden. Die ersten bzw. maßgeblichen Versionen der Prädikatenlogik wurden bei Frege (1879) und Carnap (1934) formuliert, wobei deren Notation sich noch bedeutend von den heute üblichen unterscheidet. Ein Prädikat ist eine Einheit, die eine bestimmte Anzahl von Freistellen eröffnet; man spricht auch von der Stelligkeit des Prädikats (vgl. Carstensen u.a. 2010: 43). Das Prädikat *schlafen* eröffnet z. B. nur eine Stelle – die des Schlafenden –, während *schlagen* zwei Stellen eröffnet – die des Schlagenden und die des Geschlagenen. Diese Stellen können durch Konstanten oder Variablen gefüllt werden. Die Verwandtschaft zum sprachwissenschaftlichen Prädikat und die konzeptionelle Nähe zur Valenz-

grammatik sind unverkennbar. In diesem Kapitel wurde gesagt, dass semantische Relationen auch als Aussagen formalisiert werden können. Die Prädikatenlogik erlaubt eine detaillierte Formalisierung, bei der die semantische Relation durch ein Prädikat und die Argumente, zwischen denen die Relation besteht, ausgedrückt wird.

Quantoren binden Variablen. Damit ist gemeint, dass sie eine Angabe darüber machen, für welche der durch die Variable vertretenen Entitäten die Aussage gelten soll (vgl. Partee 1990: 141). In der Prädikatenlogik existieren der Allquantor \forall , der besagt, dass die Aussage für alle durch die Variable vertretenen Entitäten gelten soll, und der Existenzquantor \exists , der besagt, dass die Aussage für mindestens eine durch die Variable vertretene Entität gelten soll.

Als Beispiel soll auf Satz (6c) zurückgekommen werden. Es ist evtl. nicht eindeutig, wie der Satz zu verstehen ist. Ist *sterblichSein* eine Eigenschaft, die einem Menschen möglicherweise oder notwendigerweise zukommt? Die Aussagenlogik bietet keine Möglichkeit, diese Frage zu beantworten; mit der Prädikatenlogik kann man den Satz jedoch folgendermaßen ausdrücken:

$$(8) \quad \forall x(\text{istMensch}(x) \rightarrow \text{istSterblich}(x))$$

Es existieren zwei Prädikate, *istMensch* und *istSterblich*, die jeweils dieselbe Variable als Argument haben und die mit der Implikation aus der Aussagenlogik verbunden sind. Die Variable ist durch den Allquantor gebunden. In natürliche Sprache übersetzt würde der Satz lauten: *Wenn eine Entität ein Mensch ist, dann ist sie sterblich* bzw. *Alle Menschen sind sterblich*.

Die Prädikatenlogik unterscheidet zwischen Konstanten und Variablen. In (8) wird eine Variable verwendet; die gleichen Prädikate können sich allerdings auch auf eine Konstante beziehen wie in (9). Die Individuenkonstanten bezeichnen eine bestimmte Entität und können nicht durch Quantoren gebunden werden, während die Entitäten, die durch die Variablen vertreten werden, erst durch die Kombination von Quantoren und Prädikaten beschrieben werden (vgl. Partee u.a. 1990: 141):

(9) $(\text{istMensch}(\text{Peter}) \wedge \text{istSterblich}(\text{Peter}))$

Die formale Semantik hat sich ausgiebig mit der Frage beschäftigt, wie prädikatenlogische Formeln zu ihrem Wahrheitswert gelangen. Dabei wurden verschiedene Verfahren entwickelt, die genau diesen Prozess beschreiben sollen. Das prominenteste Verfahren dürfte die Modelltheorie sein, deren wichtigste Konstrukte im Folgenden angerissen werden sollen, da die Modelltheorie eine der Grundlagen des *Semantic Web* darstellt (vgl. Allemang und Hendler 2011: 14-16) und da im folgenden Unterkapitel die Sprache $\mathcal{SHOIN}(\mathcal{D})$ modelltheoretisch interpretiert werden soll.⁷

In der modelltheoretischen Semantik werden die Sätze einer Objektsprache semantisch untersucht, wobei diese Objektsprache selber keine Wahrheitswerte enthält. Diese werden durch eine Metasprache eingeführt. Mittels geeigneter Namen kann in der Metasprache auf Sätze der Objektsprache Bezug genommen und das Prädikat *wahr* zu- oder abgesprochen werden (vgl. Link 1976: 3-4 und Lepore 1983: 177). Die modelltheoretische Semantik vermeidet auf diese Weise die Diskussion um einen absoluten Wahrheitsbegriff und arbeitet stattdessen mit einem Wahrheitsbegriff, der relativ zu einem Modell steht, welches durch die Metasprache erzeugt wird.

Ein Modell besteht zunächst aus der Diskursdomäne D . Dies ist die Menge, die die sogenannten Individuenkonstanten vorhält. Als Beispiel soll eine einfache Diskursdomäne erstellt werden, die nur aus drei Personenbezeichnungen besteht:

$$D = \{\text{Person1}, \text{Person2}, \text{Person3}\}$$

Der zweite Teil eines Modells ist die Interpretationsfunktion, die jeder Individuenbezeichnung ein Element aus D zuweist, und jedem n -stelligen Prädikat eine Untermenge aus dem n -stelligen kartesischen Produkt aus D . Diese Funktion könnte folgendermaßen aussehen:⁸

⁷An dieser Stelle soll nur die Grundidee der Modelltheorie vorgestellt werden. Für eine vollständige Darstellung, die u.a. die Notation erläutert, sei auf die entsprechenden Kapitel bei Lohnstein (2011) und Partee u.a. (1990) verwiesen. Bzgl. des Konstruktes der Variablenbelegung sei auf den Originaltext Tarski (1935) verwiesen.

⁸Ein Problem der Darstellung ist in solchen Fällen, dass sprachliche Ausdrücke gefunden

$$F(\textit{Peter}') = \textit{Person1}$$

$$F(\textit{Maria}') = \textit{Person2}$$

$$F(\textit{Charlotte}') = \textit{Person3}$$

$$F(\textit{schlafen}') = \textit{Person1}$$

$$F(\textit{lieben}') = \langle \textit{Person2}, \textit{Person3} \rangle$$

In diesem Modell gibt es also drei Personen, Peter, Maria und Charlotte, wobei Peter schläft und Maria Charlotte liebt. Anders ausgedrückt: Die Aussage $\textit{schlafen}(\textit{Peter})$ und $\textit{lieben}(\textit{Maria}, \textit{Charlotte})$ sind wahr, während bspw. die Aussage $\textit{lieben}(\textit{Peter}, \textit{Maria})$ falsch ist, weil das Tupel $\langle \textit{Peter}, \textit{Maria} \rangle$ sich nicht in der Menge der Tupel befindet, die F dem Prädikat \textit{lieben}' zuweist.

Zusammenfassend ist ein Modell ein Tupel $\langle D, F \rangle$, wobei D die Diskursdomäne darstellt und F eine Funktion ist, die

1. jeder Bezeichnung für eine Individuenkonstante ein Element aus D zuweist.
2. jedem n-stelligen Prädikat eine Untermenge von $\underbrace{Dx Dx \dots xD}_n$ zuweist.

In diesem Kapitel wurde die Prädikatenlogik erster Stufe eingeführt, die ihre Bezeichnung dadurch erhält, dass die Quantoren nur die Argumente von Prädikaten binden können.⁹ Außerdem wurde das Konstrukt der Modelltheorie als mögliches Verfahren, prädikatenlogische Formeln zu interpretieren, angerissen, um im folgenden Kapitel auf dieses zurückkommen zu können. Die Prädikatenlogik ist in der formalen Semantik von großer Bedeutung, da alle weiteren Logiken, u.a. auch die im folgenden Unterkapitel besprochenen *Description Logic*, auf ihr aufbauen. Grundsätzlich ließen sich alle Aussagen einer OWL-Ontologie auch in Prädikatenlogik erster Stufe darstellen. Aber es ist üblich, im Kontext von Ontologien mit *Description Logics* zu arbeiten.

werden müssen, die explizit nicht-sprachliche Entitäten bezeichnen sollen. Hier wurde das Wort *Person* mit einer Nummer versehen, um zu verdeutlichen, dass es sich um wohlunterschiedene Personen handelt. Das Apostroph kennzeichnet hingegen Ausdrücke als Teil der logischen Sprache.

⁹Eine Prädikatenlogik zweiter Stufe, die hier nicht behandelt wurde, würde das Quantifizieren über Prädikate ermöglichen.

3.4 Description Logics

Nachdem die relevanten Inhalte der formalen Semantik erörtert wurden, wird nun die Sprache eingeführt, mit der ontologische Relationen in dieser Arbeit formuliert werden sollen. Dabei handelt es sich um die *Description Logic SHOIN(D)*. Vorher sollen aber noch einige Worte geäußert werden, die den Dschungel von Bezeichnungen für formale Sprachen, der sich mittlerweile andeutet, etwas lichten werden. Ontologien sind das Ergebnis eines Übersetzungsprozesses, der den natürlich-sprachlichen Diskurs über Bedeutung in ein künstliches Notationssystem überträgt (vgl. Weber 1996: 10). Das mit Abstand gängigste Notationssystem dieser Art ist die *Web Ontology Language* (OWL).

Die verschiedenen Versionen von OWL realisieren verschiedene *Description Logics* (DL), wobei es sich wiederum um Fragmente einer Prädikatenlogik erster Stufe handelt (vgl. Hitzler u.a. 2008: 67 und Hitzler u.a. 2009: 159). Der Grundgedanke hinter dem Wort *Fragment* ist, dass nur ein Ausschnitt einer anderen Sprache realisiert wird. Bei DLs werden nur Ausdrücke zugelassen, welche die Sprache entscheidbar halten. Eine logische Sprache ist entscheidbar, wenn ein Algorithmus existiert, der in endlicher Zeit bestimmen kann, ob eine Entität zu einer Klasse gehört oder nicht (vgl. Hitzler u.a. 2008: 67).¹⁰ Daneben gibt es in DLs aber auch Sprachkonstrukte, die die herkömmliche Prädikatenlogik erweitern, indem sie in einem Ausdruck mehrere Ausdrücke zusammenfassen. Viele solcher Beispiele werden in diesem Unterkapitel folgen.

Die Bezeichnungen für DLs sind Akronyme und sollen für die Sprachkonstrukte stehen, die die jeweilige DL enthält; dabei ist die Zuordnung zwischen Buchstaben und Sprachkonstrukt häufig unintuitiv. Die einfachste DL heißt *ALC*, was für *Attributive Concept Language with Complements* steht. In *SHOIN(D)* steht das *S* für *ALC* plus Rollentransitivität, das *H* für Hierar-

¹⁰Auch dieser Aspekt wird im weiteren Verlauf nicht von großer Bedeutung sein, wurde hier aber erwähnt, um die Motivation hinter DLs zu beleuchten.

chie, das O für Klassen von Objekten, das I für Inverse, das N für numerische Restriktionen und (D) für die Verwendung von Datentypen (vgl. Hitzler u.a. 2008: 172). Wie diese Konstrukte eingebunden sind, wird im Folgenden erläutert, wobei sich die meisten Zusammenhänge aus den bereits angestellten Überlegungen ergeben.

Zunächst sind DLs in zwei Bereiche unterteilt: die *Terminology Box* (TBox) und die *Assertional Box* (ABox). Dabei stehen in der TBox die Formeln, die nur Klassen betreffen, während in der ABox Formeln stehen, die Instanzen beinhalten (vgl. Hitzler u.a. 2009: 162).

Grundsätzlich besteht eine ABox aus Aussagen der Form $C(a)$, $R(a,b)$ oder $a = b$, bei der a und b Instanzen sind, C eine Klasse und R eine Rolle (vgl. Hitzler u.a. 2009: 162). Die Formeln der ABox unterscheiden sich abgesehen von der Erweiterung um das Gleichheitszeichen in der Notation also nicht von den prädikatenlogischen Formeln (vgl. Hitzler u.a. 2008: 169). Die Aussagen der TBox betreffen ausschließlich Klassen, weswegen die aus der Mengenlehre bekannten Operationen und Symbole verwendet werden, allerdings in einer eckigen Schreibweise, um eine Verwechslung mit der herkömmlichen Mengenlehre zu vermeiden. Es sei an dieser Stelle darauf hingewiesen, dass die kleinen Buchstaben a , b etc. für Instanzen und nicht zwingend für Individuen stehen. Individuen sind die Elemente der Diskursdomäne; Instanzen sind die Elemente einer Menge, damit aber nicht zwingend Individuen. Neben Individuen können also auch Klassen wiederum als Instanzen auftreten. In (10) sind eine ABox und eine TBox-Formel aufgeführt:

$$(10) \quad \text{a. } A(x)$$

$$\text{b. } A \sqsubseteq B$$

$$\text{Übersetzt in } L_1: \forall x.A(x) \rightarrow B(x)$$

Der zweiten Formel in (10) ist eine Übersetzung in Prädikatenlogik beige-fügt, die einer genaueren Betrachtung unterzogen werden soll: Dass die prädikatenlogische Übersetzung die Beziehung zwischen Unter- und Oberklasse

adäquat ausdrückt, ist zunächst evident. Darüber sollte aber nicht vergessen werden, dass Mengenlehre und Prädikatenlogik zwei verschiedene Konstrukte sind. In der Mengenlehre werden Mengen gebildet und in der Prädikatenlogik Aussagen, denen ein Wahrheitswert zukommt. Die Formel in (10b) ist ebenfalls eine Aussage und hat einen Wahrheitswert. In $\mathcal{SHOIN}(\mathcal{D})$ wird also das Vokabular der Mengenlehre leicht verändert genutzt, um Zusammenhänge auszudrücken, die in Prädikatenlogik komplizierter auszudrücken sind. Eine Übersetzung ist grundsätzlich möglich.

In dieser Arbeit werden folgende Konventionen befolgt: Wenn ein natürlicher Begriff in eine Formel als Klasse oder Instanz übertragen wird, dann wird er in der Normalform angegeben: d.h. bei Substantiven im Nominativ Singular, bei Adjektiven unflektiert und bei Verben im Infinitiv. Dies gilt auch für komplexere NPn, welche außerdem zusammen und mit Großbuchstaben geschrieben werden. Umlaute und ß werden durch Doppelbuchstaben ausgedrückt. In (11) sind einige Beispiele aufgelistet:

- (11) a. $Hund(Lassie)$
 b. $Hund \sqsubseteq Saeugetier$
 c. $HundOhneFell \sqsubseteq Hund$
 d. $Riesengross \sqsubseteq Gross$

Die anderen aus der Mengenlehre bekannten Relationen werden folgendermaßen in $\mathcal{SHOIN}(\mathcal{D})$ ausgedrückt:

- (12) a. $A \sqcap B$
 b. $A \sqcup B$
 c. $\neg A$

Komplexe Klassen können durch Kombination verschiedener Relationen erzeugt werden, wie in dem folgenden Beispiel (vgl. Hitzler u.a. 2008: 169):

- (13) $A \sqsubseteq (B \sqcap C) \sqcup (D \sqcap \neg E)$.
 Übersetzt in L_1 : $\forall x.A(x) \rightarrow (B(x) \wedge C(x)) \vee (D(x) \wedge \neg E(x))$

DLs nutzen feste Ausdrücke für die leere Menge und für die Menge, die alle Dinge enthält (zumindest alle Dinge, über die gesprochen wird). Sie sind folgendermaßen definiert (vgl. Hitzler u.a. 2008: 161):

- (14) a. $\perp \equiv C \sqcap \neg C(x)$
 Übersetzt in L_1 : $\forall x.C(x) \wedge \neg C(x)$
- b. $\top \equiv C \sqcup \neg C(x)$
 Übersetzt in L_1 : $\forall x.C(x) \vee \neg C(x)$

Die grundlegende Logik ist die folgende: Die leere Menge (\perp) ist die Schnittmenge einer Menge mit ihrem Komplement; die Vereinigungsmenge einer Menge mit ihrem Komplement enthält dagegen Dinge bzw. die Diskursdomäne (\top).

DLs bieten außerdem die Möglichkeit, eine Klasse schlichtweg durch Aufzählen ihrer Elemente zu definieren; auf die Art definierte Klassen werden geschlossene Klassen genannt (vgl. Hitzler u.a. 2009: 163):

- (15) $C \equiv a, b, c$
 Übersetzt in L_1 : $C \equiv \forall x.(C(x) \rightarrow (x = a \vee x = b \vee x = c))$

Eine Besonderheit von DL ist der Umgang mit Relationen. In DL existieren nur zweistellige Relationen, die als Rollen bezeichnet werden (nicht zu verwechseln mit den semantischen Rollen; wenn Verwechslungsgefahr besteht, werden die Rollen einer DL auch ontologische Rollen genannt). In Kapitel 3.3 wurde ausgeführt, dass es sich bei einer zweistelligen Relation auf einer Menge um eine Menge von Tupeln handelt. Im Kontext der Wissensrepräsentation ist häufig das zweite Element des Tupels von besonderem Interesse. Als Beispiel diene die Teil-Ganzes-Relation. Für eine Definition ist weniger von Interesse, welche Tupel sich in dieser Relation befinden, also was woraus besteht, sondern es ist von Interesse, woraus etwas Bestimmtes besteht, bspw. woraus ein Auto besteht. Um diesem Sachverhalt gerecht zu werden, können Rollen in DL folgendermaßen verwendet werden (vgl. Hitzler u.a. 2009: 173-174):

- (16) a. $\exists R.C$
 Übersetzt in L_1 : $\exists y.R(x,y) \wedge C(y)$
- b. $\forall R.C$
 Übersetzt in L_1 : $\forall y.R(x,y) \rightarrow C(y)$

(16a) sagt aus, dass es mindestens ein y gibt, das mit einem x in einer Rolle, also in einer Relation, steht und einer Klasse C angehört; (16b) sagt aus, dass alle y , wenn sie mit einem x in einer Rolle stehen, einer Klasse C angehören. Eine Rolle wird also mit einem Quantor und einer Klasse, die das Objekt bildet, zu einem Ausdruck, der dann wiederum eine Klasse bildet, kombiniert. Da der Sinn dieser Ausdrücke sich vielleicht nicht auf den ersten Blick erschließt, seien hier zwei Beispiele gegeben:

- (17) a. $Pruefung \sqsubseteq \exists hatPruefer.Professor$
 Übersetzt in L_1 :
 $\forall x.Pruefung(x) \rightarrow (\exists y.hatPruefer(x,y) \wedge Professor(y))$
- b. $Pruefung \sqsubseteq \forall hatPruefer.Professor$
 Übersetzt in L_1 :
 $\forall x.Pruefung(x) \rightarrow (\forall y.hatPruefer(x,y) \rightarrow Professor(y))$

(17a) sagt demnach aus, dass eine Prüfung einen Professor als Prüfer haben muss. Übersetzt in Prädikatenlogik: Für alle x gilt, wenn sie der Klasse *Pruefung* angehören, dann gibt es mindestens ein y , das mit x in der Relation *hatPruefer* steht und der Klasse *Professor* angehört. (17b) sagt aus, dass eine Prüfung ausschließlich Professoren als Prüfer hat. In Prädikatenlogik: Für alle x gilt, wenn sie der Klasse *Pruefung* angehören, dann gilt für alle y , wenn sie mit x in der Relation *hatPruefer* stehen, dann müssen sie der Klasse *Professor* angehören.¹¹

Die Beispiele 17a und 17b zeigen außerdem, wie zentral die Unterklassenrelation für Ontologien und DL ist: Die Aussage *Eine Prüfung hat einen Prüfer* wird umformuliert in *Eine Prüfung ist eine Unterklasse der Dinge, die einen Prüfer hat*.

¹¹Es sei angemerkt, dass eine Formel mit Allquantor, wie (17b), selten alleine steht, da sie nicht voraussetzt, dass die Rolle existiert.

In der Kombination mit Rollen sind in DL außerdem numerische Restriktionen möglich (vgl. Baader u.a. 2009: 23).

- (18) a. $\leq nR.C$
 Übersetzt in L_1 :
 $\forall x. \neg(\exists x_1 \exists x_2 \dots \exists x_n ((x_1 \neq x_2 \neq \dots \neq x_n) \wedge R(x, x_1) \wedge R(x, x_2) \wedge \dots \wedge R(x, x_n)) \wedge C(x_1) \dots C(x_n))$
- b. $\geq nR.C$
 Übersetzt in L_1 :
 $\forall x. (\exists x_1 \exists x_2 \dots \exists x_n ((x_1 \neq x_2 \neq \dots \neq x_n) \wedge R(x, x_1) \wedge R(x, x_2) \wedge \dots \wedge R(x, x_n)) \wedge C(x_1) \dots C(x_n))$

(18a) beschreibt die Dinge, die mit höchstens n Dingen in C bzgl. der Rolle R verbunden sind; (18b) beschreibt die Dinge, die mit mindestens n Dingen in C bzgl. der Rolle R verbunden sind (vgl. Hitzler u.a. 2009: 173-174). Die logische Übersetzung sieht kompliziert aus, ist jedoch einfach erklärt. (19) sagt bspw. aus, dass eine Prüfung eine Unterklasse der Dinge ist, die höchstens zwei Prüfer haben; es gibt also nicht mehr als zwei, d.h. keine drei, wohlunterschiedenen Entitäten, die mit *Pruefung* in der Relation *hatPruefer* stehen.

- (19) $Pruefung \sqsubseteq \leq 2hatPruefer.Professor$
 Übersetzt in L_1 :
 $\forall x (Pruefung(x) \rightarrow \neg(\exists x_1 \exists x_2 \exists x_3 ((x_1 \neq x_2 \neq x_3) \wedge hatPruefer(x, x_1) \wedge hatPruefer(x, x_2) \wedge hatPruefer(x, x_3) \wedge Professor(x_1) \wedge Professor(x_2) \wedge Professor(x_3)))$

Funktionen lassen sich in DLs über eine numerische Restriktion ausdrücken, da es sich bei ihnen letztlich um Relationen mit der Restriktion 1 handelt (vgl. Calvanese und de Giacomo 2007: 193).

Wie eingangs erwähnt wurde, steht das D in $SHOIN(D)$ für Datentypen, worunter Maßeinheiten, aber auch Einheiten aus der Informatik wie *Boolean* oder *String* zu verstehen sind. Wenn bspw. die Höhe des Mount Everest angegeben werden soll, reicht es nicht aus, nur eine Zahl anzugeben, sondern es muss ein Datentyp, bspw. Meter, angegeben werden. Im modelltheoretischen Sinne handelt es sich hierbei um eine Erweiterung des Modells, die aber nach

allgemeiner Auffassung unkompliziert und eher von technischem Interesse ist; wichtig ist lediglich, dass die Diskursdomäne und die Menge der Datentypen strikt voneinander getrennt sind (vgl. Horrocks u.a. 2007: 456-457). Solche Relationen, deren Range ein Datentyp ist, werden *Datatype Properties* genannt. Es handelt sich bei diesen also um eine Menge von Tupeln bestehend aus einem Element der Diskursdomäne und einem Datentyp.

$\mathcal{SHOIN}(\mathcal{D})$ stellt außerdem Ausdrücke bereit, um Aussagen über Rollen zu tätigen und diese so weiter zu spezifizieren. Diese Aussagen betreffen die Eigenschaften, die eine Rolle haben kann (vgl. Kapitel 3.3): Seien R und S zwei Rollen, dann existieren folgende Ausdrücke (vgl. Hitzler u.a. 2008: 170):

- (20) a. $R \sqsubseteq S$
 Übersetzt in $L_1: \forall x. \forall y. R(x, y) \rightarrow S(x, y)$
 b. $R \equiv S^-$
 Übersetzt in $L_1: \forall x. \forall y. R(x, y) \leftrightarrow S(y, x)$
 c. $\text{Trans}(R)$
 Übersetzt in $L_1: \forall x. \forall y. \forall z. (R(x, y) \wedge R(y, z)) \rightarrow R(x, z)$
 d. $R \equiv R^-$
 Übersetzt in $L_1: \forall x. \forall y. R(x, y) \rightarrow R(y, x)$

Formel (20a) besagt, dass R eine Unterrolle von S ist, also alle Tupel, die in R sind, sind auch in S . (20b) besagt, dass R und S invers zueinander sind; für jedes Tupel in R existiert das umgekehrte Tupel in S . (20c) drückt Transitivität aus: Wenn sich die Tupel $\langle x, y \rangle$ und $\langle y, z \rangle$ in R befinden, dann auch das Tupel $\langle x, z \rangle$. Formel (20d) drückt Symmetrie aus, wenn sich das Tupel $\langle x, y \rangle$ in R befindet, dann auch das umgekehrte Tupel.

Abschließend seien die Ausdrücke der DL $\mathcal{SHOIN}(\mathcal{D})$ zusammengefasst:

1. \perp bezeichnet die leere Menge.
2. \top bezeichnet alles in der Diskursdomäne.
3. Wenn A und B Klassen sind, R eine Rolle und a ein Individuum ist, dann sind auch $A \sqsubseteq B$, $A \sqcap B$, $A \sqcup B$, $A \equiv B$, $\neg A$, $\forall R.A$, $\exists R.A$, $\leq nR.C$, $\geq nR.C$, $A(a)$, a, b Ausdrücke in $\mathcal{SHOIN}(\mathcal{D})$.

4. Wenn a und b Individuen sind, dann sind $A(a)$, $a \equiv b$ und $a \neq b$ Ausdrücke in $\mathcal{SHOIN}(\mathcal{D})$.
5. Wenn R und S Rollen sind, dann sind $Trans(R)$, $Sym(R)$, $Ref(R)$, $Irr(R)$ und $Dis(R, S)$ Ausdrücke in $\mathcal{SHOIN}(\mathcal{D})$.
6. Wenn D ein Datentyp, d ein Element von D und R eine Rolle sind, dann sind $\exists R.D$, $\forall R.D$ und $R.d$ Ausdrücke in $\mathcal{SHOIN}(\mathcal{D})$.

Die Semantik von $\mathcal{SHOIN}(\mathcal{D})$ soll modelltheoretisch definiert werden (vgl. Hitzler u.a. 2009: 172-175). Die Diskursdomäne soll mit Δ und die Interpretationsfunktion mit I bezeichnet werden. Für I gilt:

1. Sie weist jeder Bezeichnung für eine Individuenkonstante ein Element aus Δ zu.
2. Sie weist jedem Klassennamen eine Untermenge der Diskursdomäne zu.
3. Sie weist jedem Rollennamen eine Untermenge des kartesischen Produktes der Diskursdomäne $\Delta \times \Delta$ zu.

Die Semantik der komplexen Ausdrücke lässt sich nun in Relation zum Modell ausdrücken:

1. Die Interpretation von \perp denotiert die leere Menge: $I(\perp) = \emptyset$.
2. Die Interpretation von \top denotiert die Diskursdomäne: $I(\top) = \Delta$.
3. Die Interpretation eines Komplements denotiert die Diskursdomäne ohne die entsprechende Klasse: $I(\neg C) = \Delta \setminus I(C)$.
4. Die Interpretation einer Schnittmenge denotiert die Schnittmenge der jeweiligen Denotate: $I(C \sqcap D) = I(C) \sqcap I(D)$.
5. Die Interpretation einer Vereinigungsmenge denotiert die Vereinigungsmenge der jeweiligen Denotate: $I(C \sqcup D) = I(C) \sqcup I(D)$.
6. Die Interpretation einer Existenzrolle denotiert die Menge von Individuen x , die mit mindestens einem anderen Individuum y , das einer Klasse C angehört, in einer Relation R stehen: $I(\exists R.C) = \{x \mid \text{es existiert ein } y \text{ mit } (x, y) \in I(R) \text{ und } y \in I(C)\}$.
7. Die Interpretation einer Rolle mit Allquantor denotiert die Menge von Individuen x für die gilt, wenn sie mit einem Individuum y in einer Relation R stehen, dann muss y der Klasse C angehören: $I(\forall R.C) = \{x \mid \text{für alle } y \text{ mit } (x, y) \in I(R) \text{ und } y \in I(C)\}$.

8. Die Interpretation einer numeral nach oben beschränkten Rolle denotiert die Menge von Individuen x , die mit höchstens n Individuen der Klasse C in einer Relation R stehen: $I(\leq nR.C) = \{x \mid |\{y \mid (x,y) \in I(R) \text{ und } y \in I(C)\}| \leq n\}$.
9. Die Interpretation einer numeral nach unten beschränkten Rolle denotiert die Menge von Individuen x , die mit mindestens n Individuen der Klasse C in einer Relation R stehen: $I(\geq nR.C) = \{x \mid |\{y \mid (x,y) \in I(R) \text{ und } y \in I(C)\}| \geq n\}$.

Damit ist die Sprache $SHOIN(\mathcal{D})$ vollständig definiert (vgl. Hitzler u.a. 2009: 175).

3.5 Intensionale Logik

In diesem Kapitel sollen die grundlegenden Konzepte der intensionalen Logik erläutert und ihre Relevanz für diese Arbeit ermittelt werden. Der Grund für diese Auseinandersetzung liegt darin, dass $SHOIN(\mathcal{D})$ eine extensionale Sprache ist – die Gegenüberstellung *extensional* und *intensional* wird im Folgenden ausgeführt – und sich daher die Frage stellt, wie mit intensionalen Phänomenen umzugehen sei bzw. ob solche im Kontext dieser Arbeit evtl. gar nicht zu erwarten sind.

Einige der grundlegenden Überlegungen zur intensionalen Semantik gehen auf Gottlob Frege zurück, der allerdings noch die Begriffe *Bedeutung* und *Sinn* verwendete, wobei *Bedeutung* das sein soll, was durch ein Zeichen bezeichnet wird, und *Sinn* die „Art des Gegebenseins“ (Frege 1962: 39). Heute sind die Begriffe *Extension* und *Intension* üblich. Die Extension lässt sich als die Menge der durch einen Begriff denotierten Entitäten definieren. Die Semantik von $SHOIN(\mathcal{D})$, wie sie in Kapitel 3.4 definiert wurde, ist ein Beispiel für ein extensionales Verfahren: Es wird eine feste Menge von Entitäten, die Diskursdomäne, definiert, und die Extension eines Begriffes wird in Relation zu dieser Menge bestimmt.

Dieses Verfahren birgt einige Unzulänglichkeiten. Es ist offensichtlich, dass es kein Modell gibt, dass alle existierenden Zusammenhänge der realen Welt abbilden kann; ein Modell kann immer nur einen Ausschnitt der Welt

abbilden. Es ist außerdem ersichtlich, dass dieses Verfahren dem kognitiven Prozess der Bedeutungsbestimmung nicht sehr nahe kommen kann: Es ist ohne große Probleme möglich, den Begriff *Hund* richtig zu verwenden, ohne auch nur ansatzweise alle Hunde auf der Welt zu kennen. Schließlich stellt sich auch die Frage nach dem Umgang mit fiktionalen Begriffen: In einem Modell, das in irgendeiner Form an die Wirklichkeit angelehnt ist, existieren weder Einhörner noch Vampire, d.h., diese Begriffe denotieren bzgl. eines solchen Modells beide die leere Menge. Die Schlussfolgerung, dass ein Vampir ein Einhorn ist, bzw. dass alle fiktionalen Entitäten gleich sind, ergibt ontologisch allerdings wenig Sinn. In einer Ontologie sollen diese Klassen voneinander unterscheidbar sein.

Diesen Fragestellungen kann sich mit dem Begriff der Intension genähert werden. Im Folgenden soll dieser anhand eines Beispiels, das sich in ähnlicher Weise schon bei Frege findet, erläutert werden: Warum ist Beispiel (21a) informativ und Beispiel (21b) nicht?

- (21) a. *Der Morgenstern ist der Abendstern.*
b. *Der Morgenstern ist der Morgenstern.*

Die Frage soll zunächst mit den Instrumenten der Modelltheorie analysiert werden. Angenommen sei eine Diskursdomäne Δ , die aus den Planeten unseres Sonnensystems besteht. Eine Interpretationsfunktion I , die die Verhältnisse unseres Sonnensystems richtig abbildet, müsste dann die Begriffe *Morgenstern* und *Abendstern* – und die Venus – jeweils auf dasselbe Individuum abbilden, also $I(\text{Morgenstern}) = I(\text{Abendstern}) = I(\text{Venus})$. Diese Modellierung gibt schon einen Hinweis auf die Frage, warum der eine Satz informativ ist und der andere nicht: Dass der gleiche Ausdruck, die gleiche Sache bezeichnet, ist *a priori* wahr (vgl. Frege 1962: 38), und deshalb ist Beispiel (21b) redundant. Dass die Interpretationsfunktion zwei Ausdrücke auf dasselbe Individuum abbildet, ist dagegen nicht selbstverständlich und deswegen informativ.

Es stellt sich die Frage, ob es neben den sprachhistorischen und soziolinguistischen Gründen, warum zwei Begriffe das Gleiche bezeichnen können, auch einen rein formalen Grund geben kann. Die Bedeutung von *Morgenstern* könnte man folgendermaßen paraphrasieren: *Das hellste Gestirn vor Sonnenaufgang*; den Sinn von *Abendstern* dementsprechend: *Das hellste Gestirn nach Sonnenuntergang*. Beide Ausdrücke bezeichnen in den meisten Fällen die Venus, da sie der Erde am nächsten ist und dementsprechend am hellsten leuchtet.¹² Es gibt aber auch Konstellationen, bei denen die Venus nicht sichtbar ist, da sie sich hinter der Sonne befindet; dann ist der Merkur möglicherweise der Morgen- oder Abendstern. Dies führt zu einer interessanten Beobachtung: Der Wahrheitswert der Aussage *Die Venus ist der Morgenstern* hängt davon ab, wann sie geäußert wird. Außerdem ist die Frage, was der hellste Stern vor Sonnenaufgang ist, für jeden Planeten unterschiedlich zu beantworten: Vom Neptun aus gesehen wird die Venus wohl nie den Morgenstern darstellen, da sie zu weit weg ist. Die Paraphrasierungen stellen die Intension dar, während die *Venus* die Extension der beiden Begriffe darstellt. Wenn die Begriffe Intension und Extension so gefasst werden, kann gut erklärt werden, warum die Sätze (21a) und (21b) einen unterschiedlichen Informationswert haben: Die Begriffe haben schlichtweg eine andere Bedeutung, bezeichnen aber (in den meisten Fällen) dieselbe Entität.

Ein Begriff kann also mehrere, situationsabhängige Extensionen haben, während die Intension immer dieselbe bleibt; genauso können zwei Begriffe trotz unterschiedlicher Intensionen dieselbe Extension haben. Deswegen ist die Aussage *Der Morgenstern ist der Abendstern* informativ; man könnte sie folgendermaßen explizieren: *Diese Welt ist so beschaffen, dass der hellste Stern nach Sonnenuntergang, der Abendstern, und der hellste Stern vor Sonnenaufgang, der Morgenstern, derselbe Planet sind.*

Eine prominente Definition von *Intension* geht auf Montague (vgl. 1970: 70-71) zurück, der Intension als eine Funktion sieht, die einen Begriff als Ar-

¹²Die Umstände, warum die Venus meistens als Morgen- oder Abendstern wahrgenommen wird, sind zu Zwecken des Beispiels vereinfacht.

gument nimmt und bzgl. einer Referenz angibt, was der Begriff denotiert. Um das Beispiel des Morgensterns wieder aufzugreifen: Eine Referenz kann ein Zeitpunkt sein, zu dem eine Aussage getätigt wird, und dementsprechend denotiert *Morgenstern* abhängig vom Zeitpunkt die Venus oder den Merkur. Als Referenz wäre aber auch ein anderer Planet denkbar. Dieses Verfahren bietet einen Umgang mit der Ausgangsfrage: Dass der Morgenstern der Abendstern ist, ist eine informative Aussage, da beide Begriffe eine unterschiedliche Intension aufweisen und deswegen potentiell auch eine unterschiedliche Extension aufweisen könnten.

Diese Ausführungen haben Problemstellungen der Temporal- und Modalsemantik angeschnitten. Die Temporalsemantik erweitert ein Modell um die total-geordnete Menge der Zeitpunkte, die verschiedene Zustände des Modells darstellen; bspw. Zeitpunkt T_1 , an dem die Venus der Morgenstern ist, und T_2 , an dem der Merkur der Morgenstern ist (vgl. Lohnstein 2011: 232-233). Die Modalsemantik versucht, das Konzept *möglicherweise wahr* zu formalisieren, indem sie eine Menge von Welten entwirft. Diese sind wie in der Temporalsemantik verschiedene Zustände eines Modells, die Alternativen zu einem bestimmten Zustand darstellen. Diese Welten können sich nur marginal unterscheiden oder gänzlich anders konstruiert sein und sollten als Hilfsmittel gedacht werden, um verschiedene Zustände zu formalisieren; bspw. könnte man den Unterschied zwischen Erde und Neptun durch zwei verschiedene Welten ausdrücken, in denen unterschiedliche Individuen das Prädikat *Morgenstern* besetzen (vgl. Lohnstein 2011: 286-287; Lyons 1977: 164-165). Die Menge der möglichen Welten kann dabei folgendermaßen definiert werden: Eine Welt ist eine Gruppe von Faktoren, die jeweils zu einer eindeutigen Extension führen (vgl. Link 1976: 30).

All diese Ansätze können in einer formalen Sprache zusammengeführt werden, die verschiedene Operatoren für modale und temporale Aspekte vorsieht und jede Aussage mit zwei Indizes versieht, welche die Welt und den Zeitpunkt angeben, auf den sich die Aussage bezieht (vgl. Lohnstein 2011:

295).

$$(22) \quad [[(\text{istMorgenstern}(x))]]^{M,w,t,g}$$

Der Ausdruck in (22) wäre ein Ausdruck einer solchen Sprache. Er ist wahr oder falsch bzgl. eines Modells M , einer Welt w , eines Zeitpunkts t , und einer Variablenbelegung g .¹³ Diese etwas längere Ausführung schien nötig, da das Verständnis von Intension in der Literatur sehr divergent ist. $SHOIN(\mathcal{D})$ ist eine extensionale Sprache: Die Bedeutung eines Ausdrucks ergibt sich allein dadurch, welche Individuen in der Diskursdomäne ihm zugeteilt werden (vgl. Kapitel 3.4).

Bisher wurde mehrfach behauptet, dass Logik ein von den Tatsachen der Welt letztlich unabhängiges System sei (vgl. Kapitel 3.3). Eine Verbindung bestehe nur insofern, als man versuche, die Welt bzw. einen Teilausschnitt richtig zu formalisieren, damit man neue Aussagen über die Welt ableiten kann. Die intensionale Logik scheint diese Aussage zu unterlaufen, indem sie den Wahrheitswert einer Aussage bzw. die Extension eines Begriffes relativ zu einer Faktenlage bestimmt. Es ist sinnvoll, zum vorherigen Verständnis von Logik zurückzukehren, da es aus verschiedenen Gründen (vgl. Kapitel 4) der Vorstellung hinter Ontologien entgegenkommt. Im Folgenden wird diskutiert, wie dies zumindest im Kontext dieser Arbeit möglich sein könnte.

Es ist evident, dass es Aussagen geben muss, denen unabhängig von einer Welt und einem Zeitpunkt ein Wahrheitswert zugewiesen werden kann. $2 + 2 = 4$ wäre solch eine Aussage, die immer wahr ist. Es stellt sich die Frage, ob Definitionen oder Teile von Definitionen nicht auch notwendig wahr sind. Folgendes Beispiel ist ein *genus proximum*:

$$(23) \quad \text{Hund} \sqsubseteq \text{Säugetier}$$

In natürlicher Sprache bedeutet (23), dass etwas ein Säugetier ist, wenn es ein Hund ist, oder einfach, dass jeder Hund ein Säugetier ist. Diese Aussage ist

¹³Eine solche Darstellung wirft viele Fragen auf, die hier nicht weiter behandelt werden sollen (vgl. dazu von Kutschera (1976) und Lohnstein (2011)).

extensional wahr, wenn in einem Modell alle Individuen, die Elemente der Menge der Hunde sind, auch Elemente der Menge der Säugetiere sind. Sie ist aber auch intensional wahr: Sie ist in jeder Welt zu jedem Zeitpunkt wahr; für diesen Umstand wird auch die Bezeichnung *notwendig wahr* verwendet (vgl. Lohnstein 2011: 286-289). Wenn eine fiktive Welt erdacht würde, in der es Tiere gibt, die Hunden in vielen Eigenschaften gleich sind, außer dass sie ihre Jungen nicht säugen, dann würden diese Tiere aus diesem einen Grund nicht als Hunde bezeichnet werden, weil es schlichtweg zur Definition eines Hundes gehört, seine Jungen zu säugen, sprich, ein Säugetier zu sein. Als weiteres Beispiel soll eine Realisierung der Meronymie als möglicher Teil einer Definition diskutiert werden: (24) drückt aus, dass eine Gitarre (mindestens) ein Schalloch besitzt.

(24) Gitarre $\sqsubseteq \exists \text{hatSchalloch.Schalloch}$

Bis zu einem gewissen Zeitpunkt in der Musikgeschichte war das Schalloch ein notwendiger Bestandteil der Gitarre. Die in (24) formalisierte Aussage schien wie (23) unabhängig von einem Zeitpunkt und einer Welt wahr zu sein. Was ist dann aber eine E-Gitarre? Es gibt zwei Möglichkeiten: 1. Eine E-Gitarre ist keine Gitarre, da sie sich in der Klangerzeugung grundsätzlich von einer Gitarre unterscheidet, und die Bezeichnung *E-Gitarre* ist letztlich ungenau und nur einer äußerlichen Ähnlichkeit geschuldet. 2. Eine E-Gitarre ist eine Gitarre und die Aussage in (24) muss dahingehend erweitert werden, dass eine Gitarre entweder ein Schalloch oder einen elektrischen Tonabnehmer besitzt. Beide Varianten sind legitim; entscheidend ist, dass eine festgelegte Variante unabhängig von Zeitpunkt und Welt gilt.

Einen Begriff wie *Morgenstern* zu formalisieren, ist herausfordernd, da zu diskutieren wäre, wie das Attribut, *das hellste Gestirn vor Sonnenaufgang zu sein*, darzustellen ist. In (25) wurde schlichtweg eine Datatypeproperty, *amHellstenVorSonnenaufgang* eingeführt, die im Falle von *Morgenstern* auf *wahr* abbildet. Dies ist etwas vereinfacht, genügt aber den Zwecken der Dar-

stellung.

(25) $\text{Morgenstern} \sqsubseteq (\text{Planet} \sqcup \exists \text{amHellstenVorSonnenaufgang.true})$

Eine solche Definition wäre intensional wahr, was letztlich bedeutet, dass sie eine Funktion im Sinne von Montague (vgl. 1970: 70-71) bildet, die zu jedem Index (Zeitpunkt, Weltsituation) einer Diskursdomäne auf die jeweilige Extension des Begriffes abbildet. Solche Aussagen, die unabhängig von einem Index wahr sind, sind diejenigen, die in eine Ontologie überführt werden sollen. Allerdings sind nicht nur Definitionen, sondern auch die semantischen Relationen intensional: Sie bestehen zwischen den Bedeutungen von Begriffen unabhängig von einer Welt und einem Zeitpunkt.

Aus modelltheoretischer Sicht führt dieser Ansatz zu einer interessanten Schlussfolgerung: Es kann Aussagen geben, die notwendig wahr sind, aber je nach Modell keine Extension haben. Die Aussage (23) ist wahr, auch wenn sie bzgl. eines Modells geäußert wird, das keine Individuen der Klasse Hund enthält (dies folgt schon aus der Formel allein). Letztlich lassen sich sogar notwendig falsche Klassen bilden, also solche, die nie eine Extension haben können, wie bspw. die Klasse der runden Quadrate.¹⁴ Dieser Zusammenhang ist wichtig, wenn z. B. eine Ontologie über fiktionale Dinge erstellt werden soll, also Begriffe, die bzgl. der realen Welt die leere Menge denotieren. Über ein intensionales Verfahren lassen sie sich dennoch unterscheiden.

An dieser Stelle sei eine kurze Zusammenfassung gegeben: Definitionen sind Aussagen, die notwendig wahr sind; sie können als intensionale Funktionen, also Funktionen, die für jede Weltsituation und jeden Zeitpunkt die Extension eines Begriffes angeben, aufgefasst werden. Im Folgenden stellt sich die Frage, welcher Art die Aussagen sind, die nicht notwendig wahr sind, und ob sie für Ontolog*innen relevant sein können. Ein Beispiel für eine Aussage, deren Wahrheitswert temporal abhängig ist, wurde in diesem Kapitel schon ansatzweise diskutiert:

¹⁴Als weiterführende Literatur in die Erforschung solcher Klassen seien folgende Arbeiten genannt: Parsons (1980), Priest (2005) und Giraud (2016).

(26) istMorgenstern(Venus)

Ein grundsätzlicher Unterschied zwischen den notwendig wahren Aussagen und (26) besteht darin, dass *Venus* ein Individuum bezeichnet, während *Morgenstern* und *Abendstern* Klassen, mit jeweils nur einer Instanz, darstellen. Tatsächlich besteht ein sehr wichtiger Zusammenhang zwischen Aussagen, die nicht notwendig wahr sind, und solchen, die Individuen bzw. Individuenkonstanten betreffen: Weil Klassen und ihre Definitionen abstrakte Konzepte sind, erhalten Aussagen über Klassen ihren Wahrheitswert aufgrund von logischer Konsistenz. Aussagen über Individuen erhalten ihren Wahrheitswert allein aufgrund eines Zustandes – darstellbar durch ein Modell –, der sich ändern kann bzw. zu dem Alternativen bestehen können.

Die Unterscheidung zwischen Individuen und Klassen ist eine grundsätzliche, die sich in der Geschichte ontologischer Betrachtungen wiederholt finden lässt (vgl. Sowa 2000: 4-5). An dieser Stelle sei Charles Sanders Peirces Unterteilung von Zeichen nach drei Typen hervorzuheben: Ein Erstes ist laut Peirce etwas, das ohne Beziehung zu etwas anderem, ganz für sich selbst existiert, während ein Zweites durch seine Relation zu etwas anderem existiert. Schließlich setzt er noch ein Drittes an, das ein Erstes und ein Zweites verbindet (vgl. Peirce 1986: 431). Klassen lassen sich unter die Kategorie einordnen, die Peirce Zweites nennt, weil sie wie in (25) durch Relationen definiert werden können, wodurch die zentrale Stellung, die Relationen in dieser Arbeit haben, hervorgehoben wird. Individuen jedoch fallen unter das Konzept, das Peirce als Erstheit bezeichnet; sie sind unabhängig von allem anderen und existieren nur für sich selbst. An dieser Stelle könnte die legitime Frage gestellt werden, ob ein Individuum nicht auch durch Relationen definiert werden kann, also nach Peirce ein Zweites ist. Der Unterschied zwischen Individuen und Klassen soll durch einige Beispiele verdeutlicht werden: Ein konkreter Kontext, in dem versucht wird, ein Individuum, nämlich eine Person, eindeutig zu bestimmen, liegt vor, wenn eine Behörde ein Ausweisdokument ausstellt. Dann wird versucht, hinreichend viele Daten von der Person, die sich

ausweisen will, festzuhalten (Geburtsort, Eltern etc.). Es soll ausgeschlossen werden, dass eine andere Person den Ausweis benutzen kann. Da allerdings kein absolutes Maß dafür existiert, wie viele Daten hinreichend sind, um eine Person eindeutig zu bestimmen – letztlich ist es möglich, dass eine Person am selben Tag, am selben Ort, mit selben Namen etc. geboren wird –, ist es nicht möglich, eine Person abschließend zu definieren. Ein Individuum, wie eine Person, ist ein Erstes und existiert nur für sich. Ein weiteres Beispiel kommt aus der Geographie. Ein konkretes geographisches Objekt (ein Land, eine Stadt, ein See, ein Berg) stellt modelltheoretisch ein Individuum dar. Geographische Objekte lassen sich durch ihre Koordinaten wie im folgenden Beispiel eindeutig identifizieren:

(27) Der Mount Everest liegt bei $27^{\circ}59' 16''$ N°, $86 55' 29''$ O°.

Dass aber selbst eine Aussage wie (27) nicht eindeutig im Sinne von notwendig wahr ist, wird daran deutlich, dass geographische Koordinaten sich im Laufe der Zeit ändern können: Städte, Länder, Berge, Seen breiten sich aus oder ziehen sich zusammen (vgl. Sowa 2000: 323). Die Beispiele zeigen, dass Individuen zwar Merkmale haben, dass sich aber nicht formal ermitteln lässt, welche Kombination von welchen Merkmalen benötigt wird, um ein Individuum eindeutig zu bestimmen. Für Individuen existiert demnach kein Prinzip wie *genus proximus et differentia specifica*, mit welchem diese sich eindeutig bestimmen lassen.

An diese Überlegungen knüpft die Frage an, welche Bedeutung Individuen für Ontologien haben. Prinzipiell ist es möglich, Ontologien zu erstellen, die ohne Individuen auskommen, die also nur Klassen definieren und darstellen, in welchen Relationen die Klassen zueinanderstehen (dazu noch Kapitel 5). Individuen können jedoch integraler Bestandteil einer Ontologie sein; bspw. ist eine Ontologie über politische Ämter in der Bundesrepublik Deutschland informativer, wenn sie um die verschiedenen Amtsträger*innen ergänzt wird, auch wenn diese Ergänzung formal nicht notwendig ist. Auch

wenn Individuen die Komplexität des Forschungsgegenstandes erhöhen, sollten sie m.E. nicht komplett außen vor gelassen werden, weswegen zu überlegen gilt, wie sie und die mit ihnen einhergehenden intensionalen Zusammenhänge integriert werden können.

Ein klassisches Beispiel soll illustrieren, welche Aspekte bei Aussagen über Individuenkonstanten beachtet werden müssen (vgl. Russell 1905: 483-484):

(28) *Der derzeitige König von Frankreich hat eine Glatze.*

Zunächst wirkt die Aussage schlichtweg falsch. Aus dem Prinzip *tertium non datur* (vgl. Kapitel 3.3) folgt aber, dass eine verneinte Aussage den gegenteiligen Wahrheitswert der unverneinten Aussage haben muss; wenn (28) demnach falsch ist, müsste die Verneinung, dass der König keine Glatze hat, wahr sein, was jedoch ebenfalls nicht zutrifft. Die Problematik in (28) rührt daher, dass die Existenz eines derzeitigen Königs von Frankreich präsupponiert wird, d.h. diese Existenz wird angenommen, aber nicht ausformuliert, weswegen sie auch nicht negiert werden kann. Diesem Umstand muss eine Formalisierung in irgendeiner Form gerecht werden (vgl. Russell 1905: 490). Eine Ontologie wird diesem Umstand gerecht, indem sie schlichtweg die Existenz aller ihrer Individuen präsupponiert: Wenn in einer Ontologie eine Aussage über ein Individuum getroffen wird, dann existiert dieses auch. Eine Ontologie hat die Aufgabe, einen bestimmten Sachstand korrekt darzustellen; es ist nicht von Bedeutung, was möglicherweise sein kann, sondern nur, was tatsächlich (bzgl. eines Modells) ist. (28) ließe sich folgendermaßen in $SHOIN(\mathcal{D})$ übertragen, wobei der Ausdruck *Individuenkonstante1* für eine spezielle Individuenkonstante stehen soll, die durch ihre Nummer von den anderen Individuenkonstanten in der Ontologie unterschieden wird.

- (29) a. *KoenigVonFrankreich(Individuenkonstante1)*
b. *Glatzentraeger(Individuenkonstante1)*

Im Kontext von Individuen stellt sich außerdem die Frage nach dem Umgang

mit temporalen Aspekten, die auch für Ontologien relevant sein kann. (28) mag bspw. zu bestimmten Zeitpunkten wahr gewesen sein. Die Diskussion, wie mit temporalen Aspekten im Zusammenhang von Ontologien umzugehen ist, hat unterschiedliche Regelsysteme hervorgebracht (vgl. Baader u.a. 2007: 229). In dieser Arbeit sollen auch temporale Aspekte mit der Sprache $SHOIN(\mathcal{D})$ dargestellt werden. Die Diskursdomäne wird dabei nicht um die Menge der Zeitpunkte erweitert, wie es sonst für eine temporale Logik üblich wäre (vgl. Lutz u.a. 2008: 1), da $SHOIN(\mathcal{D})$ dies nicht vorsieht; stattdessen werden Zeitpunkte als Datentypen begriffen. Die Aussage, dass ein bestimmter ehemaliger König von Frankreich eine Glatze hatte (falls diese Aussage tatsächlich als relevant erachtet wird), ließe sich dann folgendermaßen darstellen:

- (30) a. $KoenigVonFrankreich.Zeitpunkt1-Zeitpunkt4(Individuenkonstante1)$
 b. $Glatzentraeger.Zeitpunkt2-Zeitpunkt4(Individuenkonstante1)$

KoenigVonFrankreich und *Glatzentraeger* sind in dieser Formel Funktionen, die auf einen Datentyp, in diesem Fall einen Zeitraum, abbilden. Funktion und Funktionswert bilden zusammen eine Klasse. Die Formeln sagen also aus, dass ein bestimmtes Individuum von Zeitpunkt 1 bis Zeitpunkt 4 König von Frankreich war und von Zeitpunkt 2 bis Zeitpunkt 4 eine Glatze hatte.

Ziel dieses Unterkapitels war es, die Intensionale Logik als grundlegendes Forschungsgebiet der formalen Semantik vorzustellen und aufzuzeigen, welche Relevanz sie im Kontext von Ontologien haben kann und wie mit den mit ihr verbundenen Phänomenen umgegangen werden kann. Dabei zeigte sich, dass der Unterschied zwischen Klassen und Individuen ein erheblicher ist, als Aussagen über Klassen notwendig wahr oder falsch sind, wohingegen Aussagen über Individuen nur möglicherweise wahr oder falsch sind. Es konnte aber gezeigt werden, dass diese Zusammenhänge für Ontologien von geringerer Bedeutung sind, da Ontologien ohnehin nur aus wahren Aussagen bzw. aus als wahr angenommenen Aussagen bestehen. Einzig der Umgang mit

temporalen Zusammenhängen musste eingehender reflektiert werden, wobei ausgeführt wurde, dass es ausreicht, Zeitpunkte als Datentypen darzustellen, auf die dann Bezug genommen werden kann.

4 Ontologiebegriff

Abbildung 2 zeigt die grafische Darstellung einer Beispielsontologie. Sie soll im Folgenden genutzt werden, um die Ausführungen zum Ontologiebegriff zu illustrieren.

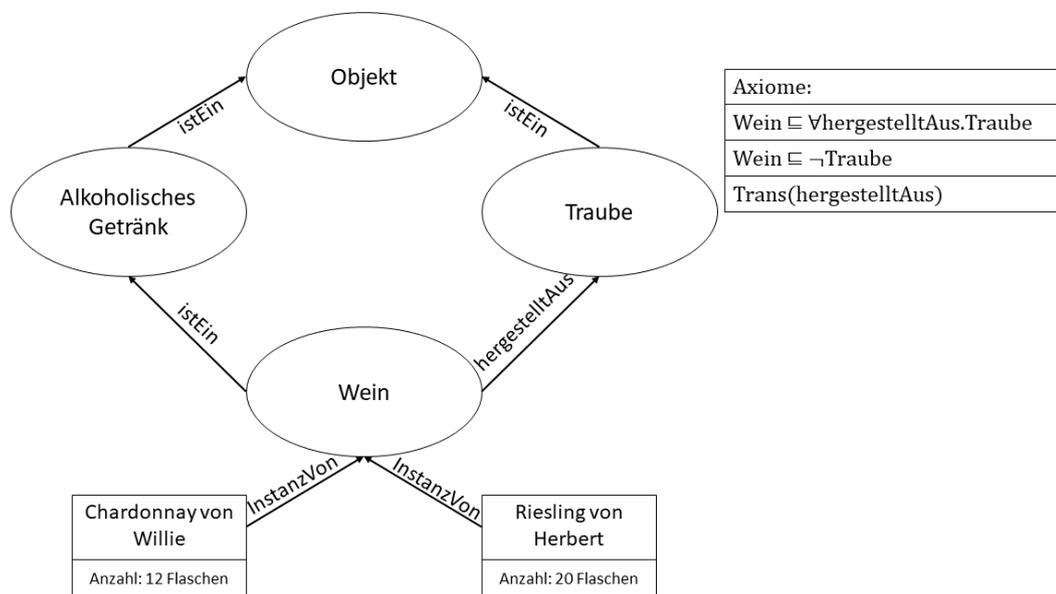


Abbildung 2: Einfache Wein-Ontologie

Dieses Kapitel vergleicht verschiedene Definitionen des Konstrukts Ontologie. Zunächst soll das Spektrum an Konstrukten, die Komponenten einer Ontologie sein können, vorgestellt werden; dabei wird herausgearbeitet, was die notwendigen Bestandteile einer Ontologie sind, um zu einer Definition zu gelangen, die exhaustiv ihre Bestandteile aufzählt. Danach soll die prominenteste Ontologiedefinition vorgestellt werden, da diese die bisherigen Ausführungen durch einen anderen definatorischen Ansatz ergänzt. Das Kapitel schließt mit einer einfachen Arbeitsdefinition, die vor allem geeignet ist, den Ausführungen in den weiteren Kapiteln zu folgen. Es sei erwähnt, dass ein

philosophischer und ein informatischer Ontologiebegriff existieren. Der philosophische Begriff bezeichnet die Lehre vom Seienden, von allem, was ist (vgl. Busse u.a. 2014: 286). Der informatische Ontologiebegriff ist im Gegensatz zum philosophischen zählbar: Ist von *einer Ontologie* oder von *Ontologien* die Rede, wird demnach der informatische Begriff verwendet (vgl. Weller 2013: 209).

Zunächst muss der Begriff der *Ordnungsrelation* eingeführt werden: Eine Ordnungsrelation ist reflexiv, antisymmetrisch und transitiv und kann als Verallgemeinerung der Kleiner-gleich-Relation (\leq) verstanden werden, mit deren Symbol sie auch bezeichnet wird (vgl. Ganter und Wille 1996: 1-2 und Partee u.a. 1990: 278). Liegen alle Elemente einer Menge in einer Ordnungsrelation, dann nennt man diese Menge geordnet. Bei einer geordneten Menge lassen sich bestimmte Elemente definieren: Gegeben sei eine geordnete Menge (M, \leq) und A sei eine Teilmenge von M , dann ist ein Element $s \in M$ eine untere Schranke von A , wenn $s \leq a$ für alle $a \in A$; eine obere Schranke ist analog ein Element $s \geq a$ für alle $a \in A$. Existiert in der Menge der unteren Schranken von A eine größte Schranke, so wird diese Infimum genannt und mit $\bigwedge A$ bezeichnet; existiert in der Menge der oberen Schranken von A eine kleinste Schranke, so wird diese Supremum genannt und mit $\bigvee A$ bezeichnet (vgl. Ganter und Wille 1996: 5; Partee u.a. 1990: 281). Das oberste Element einer geordneten Menge M , definiert als $\bigvee M$, wird Einselement genannt und mit 1_V bezeichnet; analog kann ein Nullelement 0_V existieren (vgl. Ganter und Wille 1996: 5).

Eine Ontologie kann als Versuch betrachtet werden, die Begriffe eines Themengebietes und die Relationen zwischen diesen Begriffen darzustellen (vgl. Neches u.a. 1991: 40). Eine solche Definition mag vor dem inneren Auge ein Bild ähnlich Abbildung 2 evozieren, in dem beschriftete geometrische Figuren, die für die Begriffe stehen, über Linien, die Relationen anzeigen, miteinander verbunden sind. Diese Vorstellung kann hilfreich beim Verständnis sein und ist wahrscheinlich auch der Grund, warum Ontologien häufig mit den

semantischen Netzen von Quillian (1967) verglichen werden; sie drückt allerdings nicht aus, dass Ontologien ein formales Konstrukt sind und dass die Relationen logische Aussagen repräsentieren. Semantische Netze bedienen sich zwar je nach Ausformung ebenfalls einer formalen Sprache, die Verbindungen zwischen den Konzepten innerhalb eines Netzes sind aber eher assoziativer Natur (vgl. Quillian 1969: 462), während die Relationen einer Ontologie definierender Natur sind. In semantischen Netzen würde das Konzept *Vogel* bspw. mit dem Konzept *Fliegen* assoziiert werden, während diese Assoziation bei Ontologien nur eine untergeordnete Rolle spielt, da zwar die meisten, aber nicht alle Vögel fliegen und außerdem Lebewesen existieren, die fliegen und keine Vögel sind. Dass Vögel fliegen, ist also weder hinreichend noch notwendig und daher müsste für eine Ontologie eine andere Relation gefunden werden, um das Konzept *Vogel* zu definieren.

Die wahrscheinlich am häufigsten zitierte Definition stammt von Gruber (1993: 199): „An ontology is an explicit specification of a conceptualization.“ Diese Definition mag allerdings die Frage aufwerfen, was eine Konzeptualisierung und eine explizite Spezifikation sein sollen. Guarino u.a. (2009) versuchen diese Definition auszuformulieren. Unter einer Konzeptualisierung versteht Guarino (vgl. 1997: 298) letztlich eine Diskursdomäne, auf der von vornherein bestimmte Relationen definiert sind, Individuen bspw. einer bestimmten Klasse zugeordnet und dadurch konzeptionalisiert sind. In diesem Sinne ist auch die Ontologie in Abbildung 2 eine Konzeptualisierung. Als explizite Spezifikation einer solchen Konzeptualisierung verstehen Guarino u.a. (vgl. 2009: 9-10) die Verbindung dieser mit einer logischen Sprache, welche mit L bezeichnet werden soll, durch eine Interpretationsfunktion, die die Bezeichnungen der Sprache für Individuen, Klassen und Relationen auf die entsprechenden Instanzen und Untermengen der Konzeptualisierung abbildet. Was sie unter einer Ontologie verstehen, kommt also letztlich dem sehr nahe, was in Kapitel 3.3 als Modell definiert wurde, mit dem Unterschied, dass dort die Relationen erst durch die Interpretationsfunktion erzeugt wur-

den, während sie bei Guarino u.a. (2009) direkt bestehen. Der entscheidende Unterschied liegt aber darin, dass Guarino u.a. (2009) dieses Modell als Ideal betrachten, dem sich eine Ontologie immer nur nähern kann (vgl. Guarino u.a. 2009: 10-11).

Eine Ontologie bedient sich des Instrumentariums formal-logischer Aussagen. Einen guten Überblick über das Spektrum solcher Aussagen und damit die verschiedenen Typen von Ontologien und ontologie-ähnlichen Konstrukten bietet die folgende Abbildung, die Ora und McGuinness (2001: 4) entnommen ist:

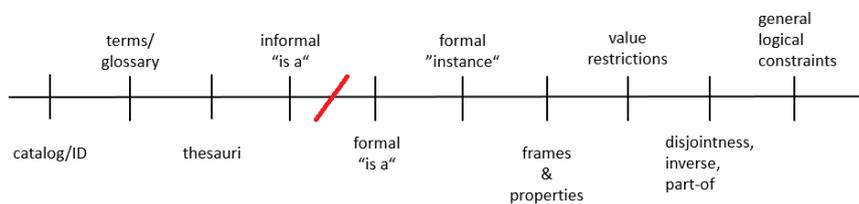


Abbildung 3: Spektrum formaler Ausdrucksmöglichkeiten

Abbildung 3 ist so gestaltet, dass die Dichte der Formalisierung von links nach rechts zunimmt, bzw. die Typen von logischen Aussagen, die formuliert werden können, mehr werden. Ganz links in Abbildung 3 steht ein einfacher Katalog – eine Liste von Wörtern, bspw. zu einem bestimmten Thema, ohne weitere Erläuterungen. Die Beispielontologie aus Abbildung 2 bestünde einfach aus den Wörtern *Objekt*, *Alkoholisches Getränk*, *Traube*, *Wein*, *Chardonnay von Willie* und *Riesling von Herbert*, wenn sie als Katalog dargestellt würde. Ein Schritt weiter rechts steht ein Glossar, worunter eine Liste von Wörtern mit einer jeweiligen Bedeutungsangabe verstanden wird. Ein Eintrag in so einem Glossar könnte bspw. folgendermaßen aussehen: *Wein* – *Alkoholisches Getränk*. Ein Thesaurus enthält zusätzlich Relationen zwischen den Wörtern, z. B. Synonymie: *Traube* – *Weinbeere*. Mit *informal* „is a“ ist eine Hierarchie gemeint, die als Ordnungsstruktur (\leq) auf den Begriffen liegt. Eine Darstellung, die solch eine Hierarchie integriert, kommt der graphischen Darstellung in Abbildung 2 schon sehr nahe. *informal* meint in diesem Zusammenhang allerdings, dass bspw. *istEin* und *hergestelltAus* nicht wie in (2) unterschieden

werden; dadurch wäre *Wein* den beiden Begriffen *Alkoholisches Getränk* und *Traube* gleichermaßen untergeordnet, obwohl die beiden zugrundeliegenden Relationen ganz anderer Art sind. All diese Konstrukte können und werden teilweise als Ontologien aufgefasst, wogegen unter bestimmten Perspektiven auch nichts einzuwenden sein mag. Sie sind allerdings nicht formal streng genug, um logische Inferenzen zuzulassen. Daher ist der Ontologiebegriff dieser Arbeit enger gefasst.

Der Ausdruck *formal „is a“* meint eine Ordnungsrelation, die auf den Begriffen liegt. Dies entspricht dem Prädikat *istEin* in Abbildung 2 bzw. \sqsubseteq in dem Vokabular einer *DL*. *Formal „instance“* meint die Existenz von Instanzen und bedingt damit die Unterscheidung von Instanzen und Klassen. In der Beispielontologie werden Instanzen und Klassen durch die Form der Darstellung, Quadrate und Ovale, unterschieden. Die Zugehörigkeit einer Instanz zu einer Klasse wird durch die Relation *InstanzVon* ausgedrückt. *frames* sind Eigenschaften, die für Klassen angegeben werden können. Eine Eigenschaft von Wein ist bspw., dass er aus Trauben hergestellt wird, was in Abbildung 2 als Relation dargestellt ist.

Die letzten drei Punkte in Abbildung 3 werden in Abbildung 2 als Formeln aufgegriffen. Der Begriff *value Restrictions* bezeichnet die Möglichkeit, die Eigenschaften von Klassen einzuschränken und dadurch genauer zu beschreiben; dies geschieht bspw. im ersten Axiom, wenn festgelegt wird, dass Wein ausschließlich aus Trauben hergestellt wird. Die Beschränkung besteht darin, dass so nur bestimmte Klassen die zweite Stelle der Relation besetzen können. Man bemerke, dass der Pfeil zwischen *Wein* und *Traube* in Abbildung 2 diesen Sachverhalt nur teilweise ausdrückt: Er macht nicht explizit, dass Wein nicht auch aus etwas anderem hergestellt werden kann. Dass diese Aussage und die anderen Axiome explizit in $SHOIN(\mathcal{D})$ ausgedrückt werden müssen, unterstreicht, dass die graphische Realisierung von Ontologien ihre Grenzen hat. Die Begriffe *disjointness*, *inverse*, *part-of* bezeichnen die Möglichkeit, Aussagen über die Beziehungen zwischen den Klassen über die

Taxonomie hinaus treffen zu können, bspw. dass *Wein* und *Traube* komplementär sind – nichts, was ein Wein ist, ist eine Traube. Wenn zusätzliche, logische Beschränkungen formuliert werden können, bspw. dass eine Relation wie *hergestelltAus* transitiv ist (vgl. Kapitel 3.1), wird dies in der Darstellung *general logical constraints* genannt.

Je nach Autor*in gibt es unterschiedliche Auffassungen, welche formalen Mittel Teil einer Ontologie sind. Für Gómez-Pérez u.a. (vgl. 2004: 11) besteht eine Ontologie aus Klassen, Relationen, Funktionen, formalen Axiomen und Instanzen, während Cimiano (vgl. 2006: 13) nicht zwischen Klassen und Instanzen, dafür aber zwischen Attributen und Datentypen, die diesen Attributen zugewiesen werden können, unterscheidet. Bei Cimiano (vgl. 2006: 13) ist vor allem hervorzuheben, dass er ein *upper semilattice* als Ordnungsstruktur für die Klassen der Ontologie postuliert. Das heißt, dass es eine Ordnungsrelation – im Falle von Abbildung 2 wäre dies *istEin* – und daneben ein oberstes Element geben muss – im Falle von Abbildung 2 wäre das die Klasse *Objekt*. Diese Ergänzung ist aus folgendem Grund wichtig: Dass eine Ordnung von Elementen in Ober- und Unterklassen einen semantischen Mehrwert generiert, ist intuitiv verständlich (trotzdem wird auf diesen Punkt noch einzugehen sein); dass es aber zusätzlich ein oberstes Element geben muss, ist ein elaborierterer Gedanke. Angenommen die Ontologie in Abbildung 2 hätte kein oberstes Element, dann wären *Wein* und *Alkoholisches Getränk* geordnet, es wäre aber unklar, wie sich *Traube* hierarchisch zu *Wein* verhält. Nur mit einem obersten Element stehen alle Klassen in einem hierarchischen Verhältnis.

Um die verschiedenen Konstrukte zu ordnen, soll im Folgenden gefragt werden, welche die notwendigen Bestandteile einer Ontologie wenigstens im Sinne dieser Arbeit sind. Eine Ontologie hat als notwendigen Bestandteil Klassen (C). Klassen werden als notwendig erachtet, da sie das Wissen, bzw. die Semantik tragen. Eine Ontologie ohne Klassen, die also nur aus Individuen und den ihnen zugehörigen Daten besteht, ist von einer herkömmlichen Datenbank nicht zu unterscheiden. Wenn die Struktur in Abbildung 2 nur aus

den Aussagen bestände, dass es zwölf Flaschen vom Chardonnay von Willie gibt und 20 Flaschen vom Riesling von Herbert, dann wäre dies nur eine Datenbank. Eine Ontologie hat als notwendigen Bestandteil Rollen (R). Diese bestehen zwischen den Klassen, zwischen Klassen und der Menge der Datentypen oder zwischen Individuen und Datentypen. Rollen sind notwendiger Bestandteil einer Ontologie, da sie herangezogen werden, um die Klassen zu definieren. Die Klassen einer Ontologie sind durch eine Ordnungsrelation (\leq) geordnet. Es handelt sich dabei um die „ist-ein“-Relation. Diese Relation ist notwendig, da nur sie eine formalisierbare Hierarchie zwischen den Klassen schafft. Erst eine solche Hierarchie ermöglicht logische Schlussfolgerungen wie *Modus Ponens* oder *Modus Tollens* (vgl. Kapitel 3.3).

Eine Ontologie enthält notwendigerweise genau ein Einselement ($\bigvee 1$ bzw. \top) – eine Klasse, die alle anderen Klassen als Unterklasse hat. Die Notwendigkeit des Einselementes ist eng mit der Ordnungsrelation verbunden. Generell wäre auch eine Ordnungsrelation ohne Einselement denkbar; dann wäre allerdings nur für Teilmengen der Klassen definiert, wie sie zueinander in hierarchischer Beziehung stehen; durch das Einselement ist die Hierarchie abgeschlossen. In Abbildung 2 ist *Objekt* das Einselement. Würde es fehlen, wäre zwar definiert, dass Wein ein alkoholisches Getränk ist, das aus Trauben hergestellt wird; es wäre aber unklar, inwiefern sich Wein und Traube voneinander unterscheiden. Durch das Einselement kann ermittelt werden, dass sie – in dieser kleinen Beispielontologie – taxonomisch maximal voneinander entfernt sind, da ihr Supremum gleich dem Einselement ist. Eine Ontologie enthält notwendigerweise genau ein Nullelement ($\bigwedge 0$, die leere Klasse bzw. \perp). So, wie die leere Menge fundamental für die Mengenlehre ist (vgl. Kapitel 3.1), benötigt eine Ontologie ein Nullelement bzw. eine leere Klasse, u.a. um auszudrücken, dass eine Klasse unerfüllbar ist. Bspw. ist die Klasse der runden Quadrate gleich diesem Nullelement.

Individuen sind kein notwendiger Bestandteil einer Ontologie (vgl. Kapitel 3.5). Dies mag verwundern, da so gut wie jede Ontologie Individuen

enthält und auch in vielen Beispielen mit Instanzen gearbeitet wurde, aber es ist auch möglich, eine Ontologie zu erstellen, die nur aus Klassen besteht und lediglich die Wissenszusammenhänge eines Themengebietes darstellt. Ließe man die Individuen in Abbildung 2 weg, hätte man eine solche Ontologie. Es wäre auch denkbar, eine Ontologie über abstrakte Klassen wie *Liebe* und *Hass* zu erstellen; diese müsste evtl. sogar notwendigerweise ohne Individuen erstellt werden – die ABox wäre also leer.

Datentypen sind kein notwendiger Bestandteil einer Ontologie. Ähnlich wie Instanzen sind Datentypen ein Bestandteil der meisten Ontologien. Es ist in der Beispielontologie ersichtlich, dass es nützlich sein dürfte, die Anzahl der erhältlichen Flaschen eines Weines anzugeben. Trotzdem wäre die Ontologie in Abbildung 2 eine gültige Ontologie, würde man diese Information weglassen.

Mengentheoretisch ist eine Ontologie also eine logische Struktur $O = \{C, R, \leq, \top, \perp\}$, die mit einer logischen Sprache L ausformuliert wurde. Diese Definition befindet sich zwischen den Ontologietypen aus Abbildung 3 insofern, als diese Definition nicht alle dort aufgeführten formalen Konstrukte als notwendig erachtet.

Zum Abschluss soll eine Arbeitsdefinition von Ontologien gegeben werden, die zwar nicht spezifisch genug ist, um als vollwertige Definition zu bestehen, da sie z.B. notwendige Elemente wie das Einselement unterschlägt, die aber besonders geeignet ist, um den Ausführungen in den folgenden Kapiteln zu folgen: Eine Ontologie ist eine Menge als wahr erachteter logischer Aussagen. Diese Definition mag nach den vorangegangenen Ausführungen nahezu banal wirken, es kann aber nicht bestritten werden, dass bei der Erstellung einer Ontologie wenig anderes getan wird, als logische Aussagen zu erstellen, die für wahr erachtet werden.

Um dieser Definition Rechnung zu tragen, wird die Ontologie aus Abbildung 2 in $\mathcal{SHOIN}(\mathcal{D})$ dargestellt.

(31) 1. *Wein(ChardonnayVonWillie)*

2. $Wein(RieslingVonHerbert)$
3. $AlkoholischesGetraenk \sqsubset Objekt$
4. $Traube \sqsubset Objekt$
5. $Wein \sqsubset \neg Traube$
6. $Wein \sqsubset AlkoholischesGetraenk$
7. $Wein \sqsubset \exists hergestelltAus.Traube$
8. $Trans(hergestelltAus)$

Aus den Formeln in (31) können weitere abgeleitet werden: Dadurch, dass \sqsubset transitiv ist, lassen sich hypothetische Syllogismen bilden (vgl. Tabelle 4): Bspw. dass der Chardonnay von Willie, dadurch dass er ein Wein ist, ein alkoholisches Getränk ist und aus Trauben hergestellt sein muss. Es ist der Nutzen einer (guten) Ontologie, dass solche Aussagen nicht explizit formuliert werden müssen, sondern hergeleitet werden können.

5 Ontologien und Semantik

In Kapitel 3 wurde ausgeführt, dass Ontologien als Konstrukt maßgeblich auf Erkenntnissen aus der formalen Semantik aufbauen. Kapitel 5.1 wird im Folgenden ausführen, dass Ontologien aus einer semantiktheoretischen Perspektive größtenteils dem Konstrukt der logischen Merkmalssemantik entsprechen. Diese steht allerdings von Seiten anderer semantischer Theorien in der Kritik (für eine besonders drastische Formulierung dieser Kritik vgl. Busse 2012: 12-19), was in Kapitel 5.2 und 5.3 zum Anlass genommen wird, sich mit diesen Kritikpunkten auseinanderzusetzen, indem die zwei maßgeblichen Konkurrenztheorien vorgestellt und auf ihre Tauglichkeit aus ontologischer Sicht untersucht werden. Dabei soll vorweggenommen werden, dass ontologische Semantik nicht als Theorie verstanden werden will, die alle semantischen Probleme lösen kann, sondern eine Theorie neben anderen darstellt (vgl. auch

Sowa 2010: 24). Ontologische Semantik ist eine Strategie von vielen, die Bedeutung von Wörtern und deren Beziehungen untereinander zu formalisieren und darzustellen.

5.1 Logische Merkmalssemantik

Die logischen Sprachen wurden bisher als eine Möglichkeit eingeführt, semantische Zusammenhänge zu formalisieren. Welchen Stellenwert logische Sprachen, und damit logische Semantik, im Spektrum der semantischen Theorien natürlicher Sprache hat, wurde bisher nicht erörtert. Der logischen Merkmalssemantik liegt die Vorstellung zugrunde, dass ein Begriff durch seine nächsthöhere Klasse und das, was ihn von anderen Vertretern dieser Klasse unterscheidet, bestimmt werden kann (vgl. Kapitel 2). Die Ontologie in Abbildung 2 enthält einen Begriff, der im Sinne der logischen Merkmalssemantik definiert ist: *Wein*. Sein *genus proximum*, also die nächsthöhere Klasse, ist *Alkoholisches Getränk*, seine *differentia specifica* der Umstand, dass Wein aus Trauben hergestellt wird. Diese Definition ließe sich folgendermaßen in die Sprache $SHOIN(\mathcal{D})$ übertragen:

$$Wein \equiv (AlkoholischesGetraenk \sqcap \exists hergestelltAus.Trauben)$$

Wein ist äquivalent zu der Schnittmenge der alkoholischen Getränke und der Dinge, die aus Trauben hergestellt werden. Diese Definition kann außerdem als eine intensionale Funktion (vgl. Kapitel 3.5) begriffen werden, da sie für jeden Weltzustand und Zeitpunkt auf die Individuen der Klasse Wein abbilden würde.

Nicht alle möglichen Merkmale kommen als *genus proximum* oder *differentia specifica* in Frage. Beispiel (32a) ist eher eine gültige Definition als (32b), weil in dem zweiten das *Stuhl* durch viele andere Möbelstücke, bspw. *Couch*, *Tisch* etc., ersetzt werden kann, und der Satz wahr bliebe (vgl. Weinrich 1966: 446, Minsky 1988: 245 und Ziem 2008: 24). Anders ausgedrückt, das Merkmal, dass ein Stuhl im Wohnzimmer stehen kann, ist höchstens notwendig, nicht aber hinreichend.

- (32) a. *Ein Stuhl ist ein Möbelstück, das nicht gepolstert und zum Sitzen gedacht ist.*
b. *Ein Stuhl ist ein Möbelstück, das im Wohnzimmer stehen kann.*

Was die Merkmalssemantik unter einem Begriff versteht, deckt sich weitgehend mit dem, was in Kapitel 3 als Klasse definiert wurde, und dementsprechend scheinen die Methoden dieser Semantiktheorie den ontologischen Methoden in vielerlei Hinsicht zu gleichen: Ein Begriff wird in eine Hierarchie von Begriffen eingegliedert und lässt sich durch eine bestimmte Eigenschaft, die in der Ontologie als Relation dargestellt wird, von der übergeordneten Klasse unterscheiden.

Wenn Klassen in einer Ontologie definiert werden, wird ebenfalls nach Merkmalen gesucht, die mit Hilfe einer logisch-fundierten Sprache ausformuliert werden (vgl. Hitzler u.a. 2009: 111-157); insofern verkörpert die logische Merkmalssemantik die Vorstellung von Bedeutung, die auch Ontologien zugrunde liegt. In der Literatur zur Merkmalssemantik lässt sich allerdings wenig zu der Frage finden, wie Individuen zu definieren seien (vgl. Kapitel 3.5). Dieser Umstand liegt darin begründet, dass Individuen einer anderen Zeichenebene im Peirce'schen Sinne angehören: Ein Individuum lässt sich nur identifizieren, nicht wie ein Begriff definieren; insofern ist es auch nicht verwunderlich, wenn die Merkmalssemantik über Individuen schweigt. Es scheint aber wichtig, auf diesen Umstand zumindest hinzuweisen, da die Frage, welche Informationen bzgl. Individuen relevant sind, für die Erstellung von Ontologien von Bedeutung ist (vgl. Kapitel 4).

Als theoretisches Konstrukt scheint die logische Merkmalssemantik, das ideale Verfahren bereitzustellen, die Bedeutung von Begriffen zu bestimmen. Beispielanalysen haben allerdings gezeigt, dass es sehr schwierig oder gar unmöglich ist, für alle Ausdrücke der Sprache notwendige Merkmale zu finden. Die Kritik kann in zwei Perspektiven aufgeteilt werden: Die erste Perspektive bezieht sich auf das Kriterium des notwendigen und hinreichenden Merkmals und hinterfragt, ob dieses ein sinnvolles Konstrukt ist. Die zweite Perspektive

hinterfragt die Denotate an sich. Ich möchte zunächst auf die erste Perspektive eingehen, indem ich sie in Bezug auf vier Einwände diskutiere.

Der erste Einwand fragt, wie Objekte zu behandeln seien, die eigentlich eindeutig einer Klasse angehören, aber kein notwendiges Merkmal mit dieser teilen. Als Beispiel soll der Begriff *Tasse* dienen: Er ließe sich als ein Trinkgefäß mit Henkel definieren. Aber wie wären dann Tassen zu bezeichnen, deren Henkel abgebrochen ist? Ist eine Tasse, deren Henkel abgebrochen ist, noch eine Tasse? Diese Frage ließe sich unter einer bestimmten Perspektive nach folgenden Überlegungen mit Nein beantworten: Zunächst muss zwischen der Klasse der Tassen und einer bestimmten Tasse, einem bestimmten Objekt, unterschieden werden. Dieses bestimmte Objekt kann nun für ein bestimmtes Zeitintervall der Klasse der Tassen und für ein anderes Zeitintervall – nachdem der Henkel abgebrochen ist – der Klasse der Becher zugeordnet werden (so wie die Venus mal der Morgenstern sein kann und mal nicht, vgl. Kapitel 3.5). Dieser Ansatz ist formal konsistent, wirkt aber etwas unintuitiv, da ein*e Sprachteilnehmer*in eine Tasse mit abgebrochenem Henkel wohl nicht direkt als Becher bezeichnen würde. Eine zweite Möglichkeit wäre, eine Unterklasse von Tasse zu postulieren, der die Tassen mit abgebrochenem Henkel angehören. Diese Modellierung wird der Intuition gerecht, dass etwas eine Tasse sein kann, das zwar keinen Henkel mehr hat, aber mal einen hatte. Die Existenz oder eher die Möglichkeit eines Henkels ist in dieser Definition gegeben. Auch wenn die zweite Alternative dem natürlich-sprachlichen Empfinden vermutlich näher kommt, hat die erste durchaus ihre Berechtigung, wie folgende Überlegung verdeutlicht: Man stelle sich vor, die Bruchstellen, an denen ein Henkel abgebrochen ist, würden keramisch bearbeitet, um sie unkenntlich zu machen; dann wäre aus der Tasse letztlich ein Becher geworden. Beide Alternativen führen zu einem sinnvollen Umgang mit den Begriffen, ohne das Konstrukt des notwendigen Merkmals grundsätzlich in Frage zu stellen.

Der zweite Einwand geht in eine ähnliche Richtung, rückt aber die Pragmatik in den Fokus: Wie ist es möglich, Begriffe – wie *Tasse ohne Henkel*

– sinnvoll zu verwenden? Wenn ein Henkel ein notwendiges Merkmal einer Tasse ist, wie kann eine Tasse noch als solche identifiziert werden, wenn ihr Henkel abgebrochen ist? Diese Frage ist sicherlich relevant, sie scheint aber auch zwei Ebenen der sprachwissenschaftlichen Beschreibung zu vermischen: Ein Objekt richtig zu identifizieren und mit dem richtigen Begriff darauf zu verweisen, ist ein komplexer Prozess, der vieles involviert, das von der konkreten Situation abhängt und nicht allein durch semantische Überlegungen beschrieben werden kann (vgl. Barclay u.a. 1974, McKoon und Ratcliff 1988, Hauser 2000: 80, Zwaan u.a. 2002 und Helbig 2006: 28-35). Dieser Prozess soll an einem einfachen Beispiel reflektiert werden: Eine Person besitzt eine mentale Repräsentation einer Tasse, die höchstwahrscheinlich einen Henkel involviert. Wenn auf einem Tisch nun ein Glas und ein Trinkgefäß ohne Henkel, das jedoch aus Keramik besteht, ständen, und diese Person aufgefordert würde, die Tasse zu nehmen, dann würde sie mit sehr hoher Wahrscheinlichkeit das Gefäß aus Keramik nehmen, da es trotz des fehlenden Henkels wegen seines Materials der mentalen Repräsentation einer Tasse am nächsten kommt. Auch wenn dieses Beispiel den Prozess des Verstehens stark vereinfacht, verdeutlicht es, dass die Fähigkeit, eine Tasse ohne Henkel in einer konkreten Situation als Tasse zu identifizieren, zwar eine gewisse semantische Flexibilität voraussetzt, aber nicht das Konzept der Tasse als ein Trinkgefäß mit Henkel grundsätzlich in Frage stellt. Die Bedeutung eines Begriffes und seine Verwendung sind zwei verschiedene Dinge, die sich durchaus gegenseitig beeinflussen – das soll nicht bestritten werden –; der Versuch, einen Begriff unabhängig von seiner Verwendung zu definieren, ist dennoch eine legitime Aufgabe, da Begriffe auch ohne einen Kontext mentale Repräsentationen evozieren.

Der dritte Einwand besagt, dass alle notwendigen Merkmale eines Begriffes nicht ermittelt werden könnten (vgl. Wolski 1980: 50, Lutzeier 1985: 99 und Kleiber 1993: 18-19). Da ein notwendiges Merkmal sich dadurch auszeichnet, dass alle Mitglieder der Klasse es teilen, ist es unmöglich, ein sol-

ches zu ermitteln, da niemand von sich sagen kann, dass er alle Mitglieder der Klasse – auch zukünftige etc. – überhaupt kennt. Ein Beispiel wurde schon in Kapitel 3.5 besprochen: Vor dem 20. Jahrhundert hätte man das Schallloch als notwendiges Merkmal einer Gitarre angesehen, bis schließlich E-Gitarren erfunden wurden, die stattdessen einen elektrischen Tonabnehmer haben. Ob nun der Begriff der Gitarre revidiert werden muss oder E-Gitarren uneigentliche Gitarren sind, ist eine Frage des semantischen Wandels – oder für Ontolog*innen der Modellierung – und widerlegt nicht die Tauglichkeit der Merkmalssemantik zur Beschreibung von Begriffsinhalten. Es gilt zu verstehen, dass eine Klasse etwas Abstraktes ist und es daher möglich ist, eine Klasse zu definieren, ohne alle eventuellen Mitglieder dieser zu kennen (wie sinnvoll sie definiert ist, ist eine andere Frage). Die Klasse der Gitarren ist definiert, obwohl es möglich ist, dass in Zukunft noch weitere Instrumente entwickelt werden, die dann zu den Gitarren gezählt werden können. Ebenfalls könnte es passieren, dass ein Instrument entwickelt wird, das einer Gitarre sehr ähnlich ist, aber ein notwendiges Merkmal nicht teilt; dann wäre es evtl. notwendig, eine weitere Klasse zu definieren, die hierarchisch direkt über der Gitarre und diesem neuen Instrument anzusiedeln wäre, um beide unter dieser neuen Klasse zu vereinen. Das ändert jedoch nichts daran, dass sich ein notwendiges Merkmal für eine Gitarre angeben lässt, ohne alle Mitglieder dieser Klasse zu kennen. Ganz im Gegenteil: Mit Hilfe eines solchen Merkmals lassen sich noch unbekannte Mitglieder der Klasse als solche identifizieren.

Der vierte Einwand besagt, dass es Wörter gibt, für die sich keine Merkmale finden ließen. Farben seien hier als extremes Beispiel angeführt: Für den Begriff *rot* lasse sich bspw. die Oberklasse *Farbe* aber kein hinreichendes Merkmal finden (vgl. Jackendoff 1983: 113). Rein naturwissenschaftlich ist das hinreichende Merkmal, dass bei Rot die Wellenlängen oberhalb von 600nm und unterhalb von 750-800nm dominieren. Dies ist in der Tat kein Wissen, dass jede*r Sprachteilnehmer*in des Deutschen sofort verbalisieren kann – es ist in dem Sinne kein Sprachwissen. Sprachteilnehmer*innen wür-

den auf die Frage, wie diese Farbe zu definieren sei, wahrscheinlich mit ansatzweise metaphorischen Beschreibungen wie *dunkler als Gelb und wärmer als Blau* operieren. Solche Beschreibungen sind aus einer wissenschaftlichen Sicht nicht tauglich, deuten aber dennoch darauf hin, dass die Ausformulierung einer *differentia specifica* eine im Sprachbewusstsein verankerte Strategie ist.

Alle bisher genannten Einwände deuten auf ein Spannungsfeld zwischen natürlich-sprachlicher Semantik und wissenschaftlichen Definitionen hin: Es existiert eine Opposition zwischen enzyklopädischem Wissen, das wissenschaftlich erarbeitet wurde, und Sprachwissen, das ein*e Sprachteilnehmer*in zum Teil implizit mit der Sprache erlernt (vgl. Lakoff 1973: 475). Ontolog*innen könnten es sich einfach machen und darauf verweisen, dass nur enzyklopädisches Wissen interessant sei; es lohnt sich allerdings auch aus semantischer Sicht, einen Moment über diese vermeintliche Opposition nachzudenken, da sich letztlich alle bisher genannten Kritikpunkte an der Merkmalssemantik aus diesem Spannungsfeld ergeben.

Das Beispiel der Farben ist ein Extremfall, da sich die hinreichende Bedingung dem Wissen vieler Sprachteilnehmer*innen nahezu komplett entziehen dürfte, diese das Wort *Rot* aber dennoch richtig verwenden können. Es gibt aber auch Fälle, in denen die enzyklopädische und die Alltagssprachliche Perspektive aufeinandertreffen. Bspw. würde jedes Kind einen Delfin wohl als *Fisch* bezeichnen, wenn es noch nicht gelernt hat, dass Delfine Säugetiere sind. Sprachwissen – ein Fisch hat Flossen und lebt im Wasser – und enzyklopädisches Wissen – ein Delfin ist ein Säugetier, und Säugetiere sind keine Fische – kollidieren an dieser Stelle. Das Beispiel zeigt, dass Sprachwissen aus einer kognitiven Perspektive tatsächlich nicht allein nach dem Modell von notwendigen und hinreichenden Bedingungen strukturiert ist, dass sich sprachliches Wissen aber durchaus auf diese Weise ergänzen lässt. Ein*e Sprachteilnehmer*in lernt, dass ein Delfin kein Fisch ist.

Dieser Zusammenhang wird schon seit längerem in der Semantikfor-

schung diskutiert, wobei es Stimmen gibt bzw. gab, die verneinen, dass enzyklopädisches Wissen sprachliches Wissen ergänzt (vgl. Bierwisch 1983: 97-98). Letztlich besteht aber bei den meisten Autor*innen die Einsicht, dass die Grenzen zwischen beiden Bereichen nicht klar zu definieren sind, wie das Beispiel *Delfin* illustriert (vgl. Palmer 1976: 46 und Bergenholtz und Kaufmann 1996: 170-171).

Mitunter wird in Frage gestellt, ob die Begriffsbestimmungen, die sich aus der Merkmalssemantik ergeben, den Inhalt der jeweiligen Begriffe überhaupt richtig fassen. Die folgenden Beispiele aus Busse (2014: 14) sollen aufzeigen, dass in den entsprechenden Begriffen mehr Bedeutung liegt, als durch die Merkmale abgedeckt wird: Warum wird ein 40-jähriger Mann, dessen Erzeuger verstorben sind, nicht eine Waise genannt? – Bei genauerer Betrachtung stellt sich die Frage als irreführend heraus, da der Begriff *Waise* als Kind (*genus proximum*), das seine Eltern verloren hat (*differentia specifica*), definiert werden kann – eine Waise ist schlichtweg kein Erwachsener. Die Frage, warum das *genus proximum* nun *Kind* und nicht allgemein *Mensch* ist, mag durchaus interessant sein (wahrscheinlich weil Konnotationen wie *Hilflosigkeit* mitschwingen), sie stellt aber nicht die grundsätzliche Definition in Frage. Warum zögern wir, den Papst einen Junggesellen zu nennen? – Auf diese Fragen sind verschiedene Antworten möglich. Evtl. wäre es redundant, da das Merkmal *Junggeselle* ein notwendiges für den Begriff *Papst* ist. Vielleicht werden als Junggesellen aber auch nur jene verstanden, die noch heiraten wollen. Auffällig ist jedenfalls, dass sich die Begriffe nicht widersprechen: Wenn jemand den Papst als Junggesellen bezeichnen würde, würde dies evtl. humoristisch verstanden werden, der Humor würde aber darauf beruhen, dass die Aussage wahr und nicht absurd ist. Wie ist das Kompositum *Ersatzkaffee* [im Original: imitation coffee] zu verstehen, das einen logischen Widerspruch auszudrücken scheint: nämlich, dass das Bezeichnete zugleich Kaffee ist und eben kein Kaffee ist? Hier liegt nur ein Widerspruch vor, wenn restriktiv davon ausgegangen wird, dass der zweite Teil eines Kompositums

immer das *genus proximum* ausdrückt. Dass dies aber nicht der Fall ist, legen diverse Komposita nahe (vgl. *Fledermaus*, *Seepferdchen* etc.). Dass Ersatzkaffee kein Kaffee ist, sondern ein Ersatz für Kaffee, ist ein interessanter Fall von Komposition, stellt die logische Merkmalssemantik jedoch vor keinerlei Herausforderungen.

Allen hier gestellten Fragen und formulierten Einwänden ist gemeinsam, dass sie bestimmte Konnotationen hervorheben, die durch die Merkmale nicht abgedeckt werden. Die logische Merkmalssemantik mag in der Tat nicht geeignet sein, Konnotationen zu untersuchen, da sie versucht, den Kern eines Begriffes zu bestimmen, also per Definition Konnotationen ausblendet. Es soll hervorgehoben werden, dass dieses Kapitel die Kritik an der Merkmalssemantik nicht als unbegründet darstellen möchte, sondern lediglich herausarbeiten will, dass diese Kritik die Methoden der Merkmalssemantik nicht im Kern widerlegt, sondern ergänzt. Eine Schwäche der Merkmalssemantik liegt darin, dass sie mit der Dynamik natürlicher Sprachen nur unzureichend umgeht: Die Denotation eines Begriffes ist ständig im Wandel; Semantik ist wahrscheinlich die Ebene der Sprache, die sich am schnellsten ändert. Dies kann auch für Ontologien eine Herausforderung darstellen und bspw. Aktualisierungsarbeiten nach sich ziehen. Dass die Merkmalssemantik dennoch relevant ist, belegen im Übrigen auch aktuellere, psycholinguistische Studien, die aufzeigen, dass die kognitive Organisation von Konzepten durchaus merkmalsbasiert ist und innerhalb dieser Organisation die logischen Merkmale prominent vertreten sind (vgl. Vigliocco u.a. 2004 und McRae u.a. 2005).

5.2 Prototypensemantik

Die Prototypensemantik stellt einen Gegenentwurf zur logischen Merkmalssemantik dar und soll hier als ein solcher besprochen werden. Die Frage lautet, ob sich Ontologien nach den Erkenntnissen dieser Theorie erstellen lassen. Im Folgenden wird gezeigt, dass die Prototypensemantik aus einer formalen Perspektive die logische Merkmalssemantik nicht ausschließt, dass also

beide Ansätze miteinander vereint werden können und der Frage, auf welcher Theorie eine Ontologie aufgebaut werden sollte, somit ihre Dringlichkeit genommen wird.

Auf der Arbeit von Berlin und Kay (1969) aufbauend untersuchte Eleanor Rosch Farbbezeichnungen in verschiedenen Sprachen und kam zu dem Ergebnis, dass es Farben gibt, die eine bestimmte Farbkategorie prominenter vertreten als es andere Vertreter dieser Kategorie tun; d.h. unter anderem, dass sie von Sprachteilnehmer*innen schneller als Vertreter der jeweiligen Kategorie identifiziert und mit kürzeren Bezeichnungen belegt werden (vgl. Rosch 1973: 329). Eine Formalisierung erfolgte durch Lakoff (1973) mittels der *Fuzzy Logic*.¹⁵ Diese Erkenntnisse weitet Rosch auf Begriffe generell aus und schlussfolgert, dass es ideale Vertreter einer Kategorie gebe, die als kognitive Ankerpunkte für Sprachteilnehmer*innen dienen und Prototypen genannt werden (vgl. Rosch 1975: 544). Dies ist der Kern der Prototypensemantik.

Die Prototypensemantik und ihre Methodik kommen aus der Psychologie,¹⁶ was zur Folge hat, dass sie als Beschreibungsmodell der Flexibilität, die natürlicher Sprache innewohnt, gerechter wird als eine starre Merkmalssemantik. Sie orientiert sich näher an der mentalen Repräsentation, die bei Sprachteilnehmer*innen für einen Begriff hinterlegt ist, als daran, eine Menge von Merkmalen zu suchen, die einen Begriff zwar eindeutig definieren und die logisch konsistent sind, dafür evtl. nicht der mentalen Repräsentation entsprechen.

Prototypen sind insofern ein Gegenentwurf zur logischen Merkmalssemantik, als die Merkmale, die diese Prototypen definieren, nicht von allen Vertretern der Klasse geteilt werden müssen, wie ein notwendiges Merkmal, sondern davon ausgegangen wird, dass es überlappende Merkmale gibt, die in dieser Theorie Familienähnlichkeiten genannt werden (vgl. Rosch und Mervis 1975: 575). Dieses Konstrukt kann folgendermaßen veranschaulicht werden:

¹⁵*Fuzzy Logic* kann auch in OWL umgesetzt werden (vgl. Calegari und Ciucci 2007).

¹⁶Die Methodik wird bei Schmid (2000) ausführlich dargestellt.

Ein idealer Vertreter einer Kategorie hat die Merkmale A, B und C; weniger ideale Vertreter haben nur zwei dieser Merkmale und Außenseiter nur eines. Eine prototypische Tasse, um das Beispiel wieder aufzugreifen, hätte bspw. einen Henkel, bestünde aus Keramik und ist zum Trinken von Heißgetränken gedacht. Eine chinesische Tasse ist weniger prototypisch, da sie keinen Henkel aufweist. Ein Glas ist kaum noch eine Tasse, aber immer noch mehr als bspw. ein Stuhl, da ein Glas zumindest für Getränke gedacht ist.

Dass die Prototypensemantik besser geeignet ist, die mentale Repräsentation eines Begriffes zu beschreiben, verleiht ihr Relevanz. Es darf allerdings gefragt werden, ob diese Relevanz für kognitive Modelle das einzige Kriterium ist, an dem sich eine semantische Theorie messen kann. Speziell im Kontext von Ontologien erweist sich die Prototypensemantik als problematisch. Ein wichtiger Bestandteil einer Ontologie ist die Taxonomie (vgl. Kapitel 4). Ontolog*innen müssen sich also fragen, was die Oberklasse von *Tasse* ist und wie diese prototypisch aussieht. Als Oberklasse ließe sich *Trinkgefäß* angeben, aber wie könnte ein sinnvoller Prototyp eines Trinkgefäßes, im Sinne eines idealen Vertreters, aussehen?

Die Prototypensemantik ist zwar empirisch begründet, trotzdem lässt sich darüber, was z.B. ein idealer Vertreter von *Trinkgefäß* ist, auch theoretisch nachdenken: Als Gedankenexperiment kann diese Frage bspw. darauf zugepunktet werden, ob der Prototyp eines Trinkgefäßes aus Glas oder aus Keramik besteht. Es ist vorstellbar, dass eine Testperson in einem experimentellen Umfeld sich aus sozialen, kulturellen und situativen Umständen für eine der beiden Möglichkeiten entscheidet – genau dies führen Rosch und Mervis (1975) aus. Dies würde aber nur belegen, dass eine Person bei dem Begriff *Trinkgefäß* evtl. eher an ein Glas als an eine Tasse denkt, es wäre aber abwegig deswegen anzunehmen, dass eine Tasse ein weniger idealer Vertreter der Klasse *Trinkgefäß* wäre: Beides sind Gefäße, die primär für Getränke gedacht sind. Die Fragestellung lässt sich auch untersuchen, indem die Taxonomie in die andere Richtung verfolgt wird: Eine Unterklasse von Tasse ist z.B. Kaffeetasse.

Dass Sprachteilnehmer*innen eine prototypische Vorstellung von einer Kaffeetasse haben, darf an dieser Stelle angenommen werden. Die Frage lautet vielmehr, wie eine weniger prototypische Kaffeetasse aussehen könnte. Das Bild von einer Kaffeetasse scheint relativ fix zu sein. Natürlich ließe sich sagen, dass eine Espressotasse eine weniger prototypische Kaffeetasse ist, aber auf dieser spezifischeren Ebene ist dies eine triviale Feststellung. Natürlich ist eine Espressotasse eine weniger prototypische Kaffeetasse – es handelt sich schlichtweg um eine andere Art Tasse. Die beiden Gedankenexperimente zeigen, dass Prototypen an Applikabilität einbüßen können, wenn die taxonomische Relation in beiden Richtungen verfolgt wird. Dies ist ein Umstand, den auch Rosch und Mervis (vgl. 1975: 587) beschrieben haben: Sie sprechen von einer Basisebene, auf der das Konstrukt des Prototypen am besten funktioniert, und von einer abstrakteren sowie einer spezifischeren Ebene, auf denen es weniger gut funktioniert.

Aus der Perspektive der Merkmalssemantik ließe sich kritisieren, dass die Prototypensemantik dem Visuellen viel Raum zuspricht. Bei manchen Begriffen ist der visuelle Aspekt allerdings untergeordnet. Bspw. ist ein Schraubenzieher maßgeblich durch seine Funktion bestimmt. Ein*e Sprachteilnehmer*in mag eine bestimmte Vorstellung von einem Schraubenzieher haben, bspw. mit einem roten Plastikgriff; letztlich wäre es aber unsinnig, einem Werkzeug, das den primären Zweck erfüllt, Schrauben ein- und auszudrehen, abzusprechen, ein weniger guter Vertreter dieser Klasse zu sein, weil sein Griff bspw. nicht aus Plastik ist.

Prototypen haben die Funktion einer kognitiven Abkürzung und stellen als solche einen relevanten und interessanten Forschungsbereich für die Semantik dar (vgl. Mangasser-Wahl 2000: 26). Es wäre aber zu weit gegriffen, zu behaupten, dass sie die Merkmalssemantik, die ihre eigenen Unzulänglichkeiten hat, komplett ersetzen könnte. Dies wäre auch ein unzulässiges Vermischen der Perspektiven: Die Merkmalssemantik versucht, sich den Begriffen logisch-formal zu nähern, während die Prototypensemantik auf einem kogni-

tionswissenschaftlichen Fundament aufbaut.

Ein Ansatz aus der Wissensrepräsentation, den man für die Darstellung von Prototypen in Ontologien nutzen kann, sind die sogenannten *Default Rules* bzw. *Default Logic*. Dies sind Inferenzregeln, die immer mit einer zusätzlichen Bedingung formuliert werden, nämlich, dass „herkömmliche Bedingungen“ gelten, wobei im Einzelnen definiert werden muss, wie solche Bedingungen zu verstehen sind (vgl. Delgrade und Schaub 2000: 108-109 und Sowa 2000: 373). Als *Default Rule* ließe sich bspw. formulieren, dass eine Tasse, wenn sie ein herkömmlicher Vertreter ihrer Art ist (*Default Rule*), einen Henkel hat. Auf diese Weise würde Tassen ohne Henkel nicht grundsätzlich die Klassenzugehörigkeit verweigert werden.

Der Ansatz dieses Unterkapitels war, auszuloten, inwiefern die Prototypensemantik als semantische Theorie für das Erstellen von Ontologien genutzt werden kann. Dabei wurde ausgeführt, dass sie der mentalen Repräsentation, die ein Begriff bei Sprachrezipient*innen evoziert, wahrscheinlich näher kommt, als die starre Merkmalssemantik, dass sie aber uneindeutige Ergebnisse liefert, wenn sie auf abstraktere oder spezifischere Begriffe angewendet wird. Zum Schluss wurde angedeutet, dass die Prototypensemantik den Methoden der Logik nicht so grundsätzlich entgegensteht, wie diese und andere Ausführungen es evtl. wirken lassen.

5.3 Frame-Semantik

Die Frame-Semantik stellt eine weitere semantische Theorie dar, die als Gegenentwurf zur logischen Merkmalssemantik verstanden werden kann. Frames lassen sich als Ordnungsstrukturen von Wissen verstehen (vgl. Busse 2012: 55-56, Fraas 2013: 2 und Ziem 2018: 7), bzw. als eine Menge von Begriffen – dies sind die Merkmale in dieser Theorie –, die mit einem bestimmten anderen Begriff assoziiert werden und diesen prototypisch beschreiben. Das Wort *prototypisch* ist hier nicht ohne Bedacht gewählt – die Frame Semantik vereint Vorstellungen aus der Prototypensemantik mit solchen der Va-

lenztheorie von Tesnière (2015), indem sie postuliert, dass ein Begriff einen „Satzrahmen“ evoziert, eine Menge von Wörtern, die mit dem Begriff in semantischer und syntaktischer Relation stehen und die durch grammatische Konzepte wie Subjekt und Objekt nicht hinreichend markiert werden (vgl. Fillmore 1968: 17 und Fillmore 1982: 117-119). Zwei Punkte sind von besonderer Bedeutung: Die Frame Semantik operiert mit Relationen, und diese Relationen sind prototypischer Natur, d.h. sie bestehen nicht notwendigerweise für jeden Vertreter einer Klasse. Frame-Semantik ist insofern ein Gegenentwurf zur logischen Merkmalssemantik, als sie postuliert, dass Sprachteilnehmer*innen nicht eine feste Menge von Merkmalen im Bewusstsein haben, wenn sie einen Begriff verwenden, sondern dass ein Begriff nur mit einem bestimmten Hintergrundwissen, das sich als Menge von Relationen zu anderen Begriffen darstellen lässt, richtig zu verstehen ist (vgl. Fillmore 1982: 118 und Fillmore und Atkins 1992: 76-77). Dieses Wissen ist dabei das Resultat aus Erfahrungen (vgl. Fillmore 1977a: 75). Wie schon die Prototypensemantik, kommt die Frame-Semantik der mentalen Repräsentation eines Begriffes näher als die logische Merkmalssemantik und ist flexibler, da sie keine der Relationen als notwendig ansieht.

Ein Standardbeispiel für einen Frame ist das des Kaufhandlungsereignisses mit den Partizipanten Käufer, Verkäufer, Geld und Ware, das schon in Kapitel 3.2 vorgestellt wurde, um das Konstrukt der semantischen Rollen zu diskutieren. Frames und semantische Rollen scheinen sich auf den ersten Blick sehr ähnlich, ein entscheidender Unterschied besteht aber darin, dass semantische Rollen Bestandteile eines Satzes beschreiben, während ein Frame ein kognitives Muster darstellt, das zunächst unabhängig von einer Verbalisierung existiert.

Ein Frame des Begriffes *Tasse* könnte folgendermaßen aussehen: *Trinkgefäß, Keramik, Kaffee, Tee, Schreibtisch* etc. Es ist leicht verständlich, dass das Wort *Tasse* all diese Assoziationen bei Sprachrezipient*innen aufrufen mag. Gegenüber der Merkmalssemantik erweitert die Frame Semantik also

den Bedeutungsumfang eines Begriffes; nicht nur das *genus proximum* und die *differentia specifica*, sondern viele andere Begriffe, die assoziativ mit dem zu bestimmenden Begriff verbunden sind, zählen zur Bedeutung und sind für das richtige Verständnis eines Begriffes wichtig.

- (33) a. *Der Journalist kontrollierte die Rechtschreibung.*
b. *Der Mechaniker kontrollierte die Rechtschreibung.*

Dass Frames tatsächlich eine kognitive Relevanz haben, lässt sich psycholinguistisch nachweisen, bspw. dadurch, dass ein Satz wie (33a) schneller verarbeitet wird als einer wie (33b) (vgl. Bicknell u.a. 2010: 493-494).

Es stellt sich allerdings die Frage, wann ein Frame hinreichend oder vollständig bestimmt ist (vgl. Ziem 2008: 15). Um zu bestimmen, ob etwas eine Tasse ist oder nicht, muss ein Maß existieren, welche und/oder wie viele Merkmale zur Bestimmung herangezogen werden müssen. Es existiert aber kaum Einigkeit darüber, wie ein solches Maß aussehen könnte, und in einem gewissen Sinne ist es das Ziel der Frame Semantik, sich eine maximale Flexibilität zu bewahren (vgl. Barsalou 1992: 29-30). Für Ontolog*innen stellt diese Vagheit dagegen ein Problem dar. Die Sprache $SHOIN(\mathcal{D})$ könnte zwar auch genutzt werden, um Frames auszuformulieren. Wenn ein solcher Frame aber lediglich aus nicht weiter strukturierten Assoziationen besteht und kein Verfahren existiert, das festlegt, ob der Frame in irgendeiner Form abgeschlossen ist, der Begriff also hinreichend bestimmt ist, dann ist es auch nicht möglich, die formalen Ansprüche zu erfüllen, die eine Ontologie stellt, da es z.B. nicht möglich ist, ein Individuum einer Klasse zuzuordnen, wenn für diese Klasse kein festes Merkmal existiert. Hier besteht der Vorteil der logischen Merkmalssemantik gegenüber den anderen hier vorgestellten Theorien: Sie vermeidet Vagheit, indem sie mit fest definierten Begriffen operiert (vgl. Johnson-Laird 1988: 26).

6 Korpus und Annotationstypen

Im Folgenden wird erläutert, welche Erkenntnisse das Korpus bringen soll und wie dieses erstellt wurde. Dann sollen die Annotationstypen vorgestellt und begründet werden.

6.1 Erkenntnisse, die durch das Korpus gewonnen werden

Die Ausgangsfrage dieser Arbeit lautete: Inwiefern wird ontologische Semantik in natürlicher Sprache verbalisiert? Dabei soll zunächst die Perspektive der Ontolog*innen eingenommen werden, die mit einer aus ihrer Arbeit resultierenden Erwartungshaltung an natürlich-sprachliche Texte herantritt, in denen Begriffe erläutert werden. Welche Formulierungen kennzeichnen welche ontologisch-relevante Information? Gibt es alternative Formulierungen? Sollten hinreichende Merkmale eher als Genitiv-, Präpositional- oder Relativsatz-Attribute formuliert werden? Diese und weitere Fragen ließen sich auch introspektiv untersuchen. Die Vermutung liegt allerdings nahe – und dies wird sich im Hauptteil bestätigen –, dass der Variationsreichtum in den Realisierungen sehr hoch sein dürfte und eine rein introspektive Untersuchung einen zu kleinen Ausschnitt des Phänomenbereichs darstellen würde. Diese Vermutung soll an einem Beispiel demonstriert werden. Die prominenteste ontologische Relation ist wahrscheinlich die taxonomische Relation. Es liegt nahe, zu sagen, sie würde durch das Verb *sein* markiert:

(34) *Ein Hund ist ein Säugetier.*

Allerdings wird diese Relation auch in den folgenden Formulierungen realisiert (die Unter- und Oberklasse werden zur Verdeutlichung teilweise hervorgehoben):

- (35)
- a. Das Wort **Hund** bezeichnet *ein Säugetier*, das...
 - b. Als **Hund** bezeichnet man *ein Säugetier*, das...
 - c. **Das Säugetier Hund** ist gekennzeichnet durch...
 - d. Die **Hunderasse** des Huskys ist gekennzeichnet durch...

e. *Ein Schäferhund ist gekennzeichnet durch...*

Die Beispiele zeigen, dass schon bei dieser Relation ein großer Variationsreichtum an Formulierungen vorliegen kann. Bei anderen ontologischen Relationen, die weniger standardisiert sind, darf evtl. ein noch weiteres Spektrum an Formulierungen erwartet werden.

Jedoch ergibt sich auch in die andere Richtung kein einheitliches Bild: Das Verb *sein* kann auch andere Relationen ausdrücken:

- (36) a. *Es ist Mitte August.*
b. *Die Erde ist ein Planet.*

In (36a) bezeichnet *sein* keine ontologische Relation. Das Beispiel wirkt banal, aber es zeigt, dass nicht jeder Satz ontologische Relationen enthält, während so gut wie jeder Satz bspw. semantische Rollen enthält. Ontologische Relationen werden nur dann realisiert, wenn ein Satz oder Text Wissensinhalte vermitteln will. (36b) realisiert zwar eine ontologische Relation, aber nicht die taxonomische, sondern die Instanzrelation.

Die Erkenntnis, dass diese kurzen Überlegungen zur taxonomischen Relation bereits ein relativ komplexes Bild ergeben haben, hat dazu geführt, die Arbeit korpusgestützt auszurichten. Es soll verhindert werden, dass der Phänomenbereich aufgrund von rein introspektiven Überlegungen unzureichend dargestellt wird. Die Arbeit versteht sich jedoch nicht als korpusbasiert, was bedeuten würde, dass letztlich eine statistische Untersuchung des Korpus die alleinige Quelle der Erkenntnis darstellen würde (vgl. Lemnitzer und Zinsmeister 2015: 22). Statistische Fragestellungen, wie bspw., welche Relationen am häufigsten realisiert werden, sollen nur am Rand von Bedeutung sein (siehe Kapitel 6.3). Um die Arbeit korpusbasiert auszurichten, müsste das Korpus bedeutend größer sein, was aber mit der komplexen Annotationstypologie, die in dieser Arbeit zugrundegelegt wird, nicht zu bewältigen wäre.

Eine wichtige Erkenntnisquelle stellt das Korpus bzgl. der semantischen Rollen dar. In Kapitel 3.2 wurde ausgeführt, dass die Menge der semantischen

Rollen kaum festgelegt werden kann, und es darf angenommen werden, dass die Menge der semantischen Rollen nicht rationalistisch bestimmbar ist. Mit Hilfe des Korpus und der in Kapitel 3.2 vorgestellten Kriterien zur Unterscheidung von semantischen Rollen lässt sich aber eine Typologie entwickeln, die zwar nicht den Anspruch erheben kann, alle denkbaren semantischen Rollen zu umfassen, aber als Heuristik zumindest auf einem empirischen Fundament aufbaut.

Für eine korpusgestützte Untersuchung eignet sich außerdem die Frage, wie mit den eventuellen Unzulänglichkeiten der Merkmalssemantik, die einen Aspekt von ontologischer Semantik bilden und in Kapitel 5 ausgeführt wurden, umgegangen wird: Wie werden Begriffe definiert, für die sich keine hinreichenden Merkmale finden lassen bzw. wo treten zirkuläre Definitionen auf? Werden solche Zirkel von den Artikel-Verfasser*innen zumindest ansatzweise antizipiert, und lassen sich die Begriffe, die sich einer solchen Einteilung entziehen, genauer charakterisieren? In eine ähnliche Richtung geht die Frage, wie Relationen verbalisiert werden, die weit oben in der Taxonomie angesiedelt sind. Finden die Autor*innen noch einen adäquaten Oberbegriff für *Materie*?

Ein ganz anderer Nutzen für das erstellte Korpus ergibt sich für die Computerlinguistik: Ein annotiertes Korpus kann verwendet werden, um Algorithmen aus dem Bereich des maschinellen Lernens zu trainieren, oder es kann als Goldstandard dienen, an dem solche Algorithmen oder auch regelbasierte Systeme evaluiert werden.

6.2 Textauswahl

Es wurde ein Korpus aus 717 Texten bzw. 47.750 Worttoken erstellt,¹⁷ die sich auf neun Teilkorpora aufteilen, die bzgl. ontologischer Relationen händisch annotiert wurden. Bei den Texten handelt es sich jeweils um den einleitenden

¹⁷Die Zahlen weichen von den automatisch annotierten Worttoken ab, da bei diesen die Belegangabe der Duden-Texte mitgezählt wurde.

den Satz von Wikipedia-Artikeln, da diese immer Definitionen beinhalten und dadurch ontologische Relationen realisieren, sowie Artikel des Bedeutungs- oder des Fremdwörterbuchs der Duden-Reihe. In einigen Fällen wurde bei den Wikipedia-Artikeln mehr als der erste Satz betrachtet, da die Definition sich über die Satzgrenze hinauszog. Der Begriff, der in einem Artikel definiert wird, soll im Folgenden *definiendum* genannt werden – das, was zu definieren ist.

Das maßgebliche Kriterium der Textauswahl waren die ontologischen Relationen selber. Die Teilkorpora sollten nicht zwingend in einem statistischen Sinne ausgeglichen sein, sondern vielmehr die Struktur einer Ontologie widerspiegeln. Wenn in einem Text bspw. etwas als Oberklasse identifiziert wird, stellt sich die Frage, wie diese Oberklasse in die Ontologie zu überführen ist. Solche Fragen lassen sich schwerlich mit einem Korpus aus zufällig ausgesuchten Texten untersuchen. Wird also in einem Text über Delfine bspw. erwähnt, dass ein Delfin ein Säugetier mit Flossen sei, so ist es aus ontologischer Sicht zielführend, die Texte zu *Säugetier* und *Flosse* ebenfalls ins Korpus aufzunehmen, sofern sie existieren.

Folgendermaßen wurde vorgegangen: Es wurde zunächst ein Wikipedia-Artikel gesucht, der in einer taxonomischen Einordnung so weit unten angesiedelt ist, dass zu ihm auf Wikipedia kein mit einem Artikel versehenes Hyponym gefunden wurde, bspw. Wildrose (Stand: 20.06.2022). Ein solcher Artikel diene als Einstiegsartikel: Von ihm ausgehend wurde dann die taxonomische Relation nach oben verfolgt, indem geschaut wurde, ob eine zu einem Begriff genannte Oberklasse ebenfalls durch einen Artikel in Wikipedia vertreten war. War dies der Fall, wurde der Artikel ebenfalls ins Korpus aufgenommen und die nächste Oberklasse ermittelt. Dies wurde so lange durchgeführt, wie Wikipedia es zuließ. Von dem *definiendum* Wildrose ausgehend wurde über *Rose*, *Rosengewächs*, *Pflanze* schließlich ein Text zu *Lebewesen* erreicht, zu dem zwar noch ein Hyperonym im Text angegeben wird, *organisierte Einheit*, welches aber keinen eigenen Artikel bei Wikipedia be-

sitzt. Die taxonomische Relation endet in diesem Sinne bei Lebewesen. Wenn nach diesem Algorithmus vorgegangen wird, entsteht schließlich eine taxonomisch geordnete Menge von Artikeln. In jedem dieser Artikel werden allerdings auch andere ontologische Relationen verbalisiert, die ebenfalls zu Artikeln führen, bspw. durch den Sachverhalt, dass eine Rose Stamm, Äste und Zweige besitzt. Wenn zu solchen Begriffen Artikel existierten, würden diese ebenfalls ins Korpus aufgenommen. Bei ihnen wurde dann nochmals die taxonomische Relation so weit wie möglich verfolgt. Dieses Verfahren wurde für ein Teilkorpus z.T. mehrfach durchgeführt.

Ein auf diese Weise erstelltes Korpus kann, da die Texte miteinander in ontologischen Relationen stehen, in eine Ontologie überführt werden. Ausgehend von *Wildrose* entstand bspw. ein Textkorpus zur Botanik. Existierte zu einem Artikel im Korpus ebenfalls ein Artikel im Bedeutungs- oder Fremdwörterbuch der Duden-Reihe, wurde dieser ebenfalls ins Korpus aufgenommen. Im Korpus existieren also Wikipedia-Artikel ohne analogen Duden-Artikel, aber nicht umgekehrt.

Die Duden-Artikel sind als Ergänzung zu verstehen. Aus einer konzeptionellen Perspektive lassen sich der Duden und Wikipedia kaum vergleichen. Wikipedia-Artikel werden von einer offenen Gemeinschaft geschrieben und ständig revidiert, während die Duden-Artikel von einer relativ festen Redaktion erarbeitet werden und an Auflagen gebunden sind. Ferner versteht sich Wikipedia als Enzyklopädie, die einen Begriff aus verschiedenen wissenschaftlichen Perspektiven beleuchtet, während der Duden ein Wörterbuch darstellt und sich damit auf sprachliche Aspekte konzentriert. Aber gerade weil es sich um zwei verschiedene Herangehensweisen handelt, ist es interessant zu fragen, ob bei beiden tendenziell die gleichen Relationen realisiert werden und wie sich diese Realisierungen unterscheiden.

Die neun Teilkorpora unterscheiden sich bzgl. des Wissensgebietes, dem die enthaltenen Texte zugeordnet werden können. Die Unterteilung in verschiedene Wissensgebiete wurde vorgenommen, um die Hypothese zu unter-

suchen, dass in unterschiedlichen Wissensgebieten zum Teil unterschiedliche ontologische Relationen realisiert werden. Die Definition der Wissensgebiete wurde dabei der *Information Coding Classification* (ICC) von Dahlberg (2014) entnommen, die an dieser Stelle kurz vorgestellt sei:

Die ICC ist ein komplexes Klassifikationssystem, das mit 729 verschiedenen Sachgebieten operiert, die in neun Objektbereiche eingeteilt werden (vgl. Dahlberg 2014: 101). Das hier vorgestellte Korpus enthält zu jedem dieser neun Objektbereiche ein Teilkorpus, da sich Dahlbergs Unterteilung an den Prinzipien der philosophischen Ontologie orientiert (vgl. Dahlberg 2014: 101) und damit der Perspektive dieser Arbeit besonders entgegenkommt. Dahlbergs Objektbereiche sind: Form- und Strukturbereich, Materie- und Energiebereich, Kosmos- und Geobereich, Biobereich, Humanbereich, Gesellschaftsbereich, Wirtschafts- und Produktionsbereich, Wissenschafts- und Informationsbereich, Geisteswissenschaften und Kulturbereich.

Für jeden der neun Objektbereiche wurde also nach dem vorgestellten Auswahlverfahren ein (Teil-)Korpus erstellt. Das Korpus, das aus dem Begriff *Wildrose* entstand, vertritt bspw. den Biobereich.

Ein letzter Aspekt bleibt zu erwähnen: Das beschriebene Verfahren sammelt Texte zu Begriffen, die aus formaler Perspektive Klassen beschreiben. In einer Ontologie können jedoch auch Individuen formal abgebildet werden; dies wären konkrete Entitäten wie *Michael Schumacher*, *Mount Everest* etc. Für Individuen gelten aus bestimmten Gründen, die in Kapitel 3.5 behandelt wurden, andere Bedingungen. Bei der Korpuserstellung schlägt sich dies darin nieder, dass die in den Korpustexten verwendeten Relationen, die sich auf Individuen beziehen, oft aus dem Wissensgebiet herausführen oder nur auf einen Datentyp verweisen. Bspw. liegt es nahe, für das Wissensgebiet Humanbereich einen Wikipedia-Artikel zu einem bestimmten Menschen ins Korpus aufzunehmen. Im einleitenden Satz wird dann zu diesem Menschen der Geburtsort und der Geburtstag angegeben. Der Geburtsort passt nur noch bedingt zum Wissensgebiet Humanbereich, während sich ein Datum zunächst

gar keinem Wissensbereich zuordnen lässt. Deswegen wurden für jedes Wissensgebiet Texte zu zehn Instanzen gesammelt, ohne dass zwischen diesen Texten Relationen bestehen müssen.

Die so entstandenen Teilkorpora enthalten durchaus vereinzelt Texte, die mehreren Wissensbereichen zugeordnet werden könnten, was aber durchaus im Sinne einer ontologischen Repräsentation ist. Dass das Verfahren generell aufging, zeigt sich daran, dass keine Dubletten auftraten, d.h. kein Text befindet sich in zwei der Teilkorpora.

Abgesehen von den Wissensgebieten wurde das Gesamt-Korpus (GK) auch unter anderen Aspekten in Teilkorpora aufgeteilt: Das Duden-Korpus (DK) enthält nur die Duden-Texte, und das Wikipedia-Duden-Korrespondenz-Korpus (WDK) enthält nur die Wikipedia-Texte, die einen korrespondierenden Text im DK haben. Das WDK wurde herangezogen, wenn es explizit um den statistischen Vergleich der Duden- und der Wikipedia-Texte geht. Ferner wurde ein Wikipedia-Korpus (WK) erstellt, das alle Wikipedia-Texte enthält, und ein Wikipedia-Korpus ohne die Texte zu Individuen: das Wikipedia-ohne-Individuen-Korpus (WoIK).

Korpus	Worttoken
Form- und Strukturbereich	7044
Materie und Energiebereich	6054
Kosmos und Erdbereich	5892
Biobereich	6056
Humanbereich	5462
Gesellschaftsbereich	5204
Wirtschaft und Technik	3978
Wissenschaft und Information	4726
Geisteswissenschaften	5072
Duden-Korpus	8.454
Wikipedia-Korpus	39.496
Wikipedia-Duden-Korrespondenz-Korpus	16.254
Wikipedia-ohne-Individuen-Korpus	33.584
Gesamtkorpus	47.950

Tabelle 5: Die verschiedenen Teilkorpora

6.3 Annotationstypologie

Bevor ein Blick auf das Korpus geworfen wird, müssen noch ein paar Worte über die Annotationstypen verloren werden. In den Kapitel 3.1 und 3.2 wurden vier Typen von ontologischen Rollen herausgearbeitet, die an dieser Stelle zusammengetragen seien:

1. formalisierte semantische Relationen bzw. mengentheoretische Relationen.
2. formalisierte Pertinenzrelationen.
3. formalisierte semantische Rollen, wenn sie dazu beitragen, ein Konzept zu definieren.
4. bestimmte Funktionen, wenn sie zur Definition einer Entität beitragen.

Zu diesen vier Gruppen seien im Folgenden verschiedene Annotationstypen vorgestellt. Zusätzlich wurde eine fünfte Gruppe von Annotationstypen angesetzt, die keine ontologischen Relationen, aber dennoch ontologisch relevante Informationen markieren. Diese Gruppe wurde lexikographische Formulierungen genannt.

Zunächst sei betont, dass die annotierten Einheiten immer in einer Relation stehen, d.h. zu jeder Annotation existiert noch eine weitere, die ebenfalls in der Relation steht. Das Standardbeispiel sei die taxonomische Relation: Wenn eine Einheit als Unterklasse annotiert wird, muss eine andere Einheit als Oberklasse annotiert werden. Bei den semantischen Rollen verhält sich das ebenso: Wenn etwas als Producer annotiert wird, muss eine weitere Annotation *Product* oder zumindest *Process* existieren. Die einzigen semantischen Rollen, die in einer einstelligen Relation stehen können, sind *Agent* und *Force*, wobei für *Force* kein Fall im Korpus auftrat (vgl. Kapitel 3.2).

Zunächst seien die mengentheoretischen Annotationstypen vorgestellt.

Unterklasse: Dieser Annotationstyp entspricht der Untermenge aus der Mengenlehre und steht in fester Relation mit einer Oberklasse.

Oberklasse: siehe *Unterklasse*.

Klasse: Dieser Annotationstyp wird für Entitäten angesetzt, die mit einem Individuum in Relation stehen, welches als *Instanz* annotiert wird.

Instanz: siehe *Klasse*.

Metaklasse: Hierbei handelt es sich um Klassen, deren Instanzen wiederum Klassen sind (vgl. Unterkapitel 7.5).

Für die Pertinenzrelationen (vgl. Kapitel 3.2.2) wurden folgende Annotationstypen angesetzt:

Meronym: Dieser Annotationstyp markiert eine Entität, die Bestandteil einer anderen Entität, dem Holonym, ist.

Holonym: siehe *Meronym*.

Possessor: Dieser Annotationstyp wird für Entitäten angesetzt, die über eine andere Entität, der Possession, verfügen können.

Possession: siehe *Possessor*.

Die Kriterien zur Typologisierung semantischer Rollen (vgl. Kapitel 3.2.1) hat zu folgenden Annotationen geführt:

Agent: Der *Agent* bezeichnet eine willentlich handelnde Entität. Sie stellt eine übergeordnete Rolle in dem Sinne dar, dass andere Rollen wie *Producer* oder *Experiencer* in unterschiedlicher Menge agentivische Anteile aufweisen. *Agent* selber wurde im Korpus annotiert, wenn die Rolle als Handelnder nicht weiter spezifiziert werden konnte, also wenn kein *Producer*, *Experiencer* oder eine andere Rolle mit agentivischem Anteil als Annotation in Frage kam. Was im Korpus als *Agent* annotiert wurde, entspricht dem, was in der Literatur oft Proto-Agens genannt wird (vgl. z.B. Primus 2012: 25).

Beneficient: *Beneficient* bezeichnet eine Entität, der durch eine Handlung ein Vor- oder Nachteil entsteht.

Content: *Content* bezeichnet den Inhalt einer Nachricht bzw. eines Datenträgers.

Depicted: *Depicted* bezeichnet eine Entität, auf die durch etwas anderes durch Form und Gestalt oder Konvention verwiesen wird.

Depicting: Ein als *Depicting* annotiertes Argument verweist durch seine Form

und Gestalt oder durch Konvention auf etwas anderes. Diese Rolle hat eigentlich keinen agentivischen Anteil, wird aber oft auf eine metaphorische Weise verwendet, die einen Anschein von Agentivität erweckt. Sie steht in Relation mit *Depicted*.

Experiencer: siehe Stimulus.

Force: *Force* bezeichnet eine Entität, die etwas bewirkt, dabei aber nicht willentlich agiert. So würde ein Sturm, der ein Haus abdeckt, mit *Force* annotiert werden.

Goal: Diese Rolle bezeichnet etwas, das durch eine Handlung erreicht werden soll. Dabei kann es sich um jegliche Art von Zustandsänderung handeln.

Instrument: *Instrument* bezeichnet eine Entität, die für eine Handlung genutzt werden kann. Im Gegensatz zu Ressource besteht sie in der Regel nach Beendigung der Handlung in selber Form fort.

Manner: *Manner* bezeichnet die Art und Weise, mit der eine Handlung ausgeführt wird.

Medium: *Medium* bezeichnet die Entität, über die ein *Content* transportiert wird. Siehe *Sender*.

Patient: *Patient* ist erneut eine übergeordnete Rolle. Sie bezeichnet eine Entität, die durch eine Handlung physisch verändert wird. Wie *Agent* wird diese Annotation angesetzt, wenn eine weitere Differenzierung nicht möglich ist. Diese Annotation entspricht dem Proto-Patiens (vgl. Primus 2012: 37).

Place: *Place* bezeichnet einen Ort, an dem etwas stattfindet oder sich befindet.

Process: *Process* bezeichnet die Handlung, bei der etwas entsteht. Die annotierten Texteinheiten sind meistens substantivierte Verben. Sie steht in einem Rahmen mit *Producer* und *Product*.

Producer: Ein *Producer* erzeugt ein *Product*. Sie unterscheidet sich von *Agent* und *Force* also u.a. durch den Rahmen, in dem sie auftritt. Eine weitere Unterscheidung, etwa in einen willentlichen und einen unwillentlichen *Producer*, wäre denkbar, wurde aber nicht angesetzt.

Product: *Product* bezeichnet eine Entität, die durch eine Handlung erst ent-

steht. Diese Rolle steht zusammen mit *Producer* oder *Process*.

Receiver: *Receiver* bezeichnet den Empfänger einer Sache. Diese Rolle hat Ähnlichkeit mit dem *Beneficient*, kann aber durch ihren Rollenrahmen, der *Sender*, *Content* und *Medium* enthält, von diesem unterschieden werden.

Ressource: *Ressource* bezeichnet eine Entität, die benötigt wird, um etwas herzustellen, und die, im Gegensatz zu *Instrument*, dabei aufgebraucht wird.

Role: *Role* bezeichnet eine bestimmte Position oder einen Aufgabenbereich, aus dem heraus ein *Agent* handelt. Sie steht ausschließlich mit diesem in einer Relation.

Sender: *Sender* ist hier im informationstheoretischen Sinne gemeint. Die Rolle bezeichnet eine Entität, die einen *Content*, über ein *Medium* an einen *Receiver* verschickt.

Shape: Diese Rolle bezeichnet die äußere Form, die eine Entität haben kann.

Stimulus: Als *Stimulus* wird eine Entität annotiert, die bei einer anderen Entität, dem *Experiencer*, eine Sinneswahrnehmung hervorruft. Dabei sind *Stimulus* und *Experiencer* zu nahezu gleichen Teilen agentivisch.

Substance: *Substance* bezeichnet etwas, das eine bestimmte Form hat oder annehmen kann. Sie steht zusammen mit *Shape*.

Auf den ersten Blick mag die hier angenommene Anzahl von Rollen sehr hoch wirken, was die Gefahr bergen könnte, dass die einzelnen Rollen ihre Generalität verlieren. Wie sich aber in Kapitel 8 zeigen wird, ist diese Anzahl notwendig, um dem Gegenstand gerecht zu werden. Außerdem wirkt die hier angesetzte Menge im Vergleich mit bspw. Frame-Net, das mit über 800 verschiedenen Frames, die jeweils eigene Rollen beinhalten, arbeitet (vgl. Ruppenhofer u.a. 2010: 5), klein. Bei den semantischen Rollen wurde außerdem eine spezielle Annotation eingeführt, für den Fall, dass eine semantische Rolle negiert wurde. In *Ein Exoplanet, ist ein Planet, der nicht Teil des Sonnensystems ist.* würde bspw. *Ein Exoplanet* als No-Meronym annotiert werden.

Funktionen in einer angemessenen Form zu annotieren, stellt eine Herausforderung dar. Da Funktionen ein Argument nehmen, und diesem etwas

zuweisen, im Fall dieser Arbeit immer einen numerischen Datentyp, wurde entschieden, jeweils das Argument und den zugewiesenen Datentyp zu annotieren. Dabei wurde innerhalb der Datentypen nicht weiter unterschieden und der Annotationstyp *Datatype* angesetzt, da eine Unterscheidung von Datentypen eine informatische Fragestellung und sprachwissenschaftlich zunächst von geringerer Bedeutung ist.

Argument: Diese Annotation bezeichnet eine Entität, der durch eine Funktion ein Wert von einem bestimmten Datentyp zugewiesen wird. Die Annotation steht mit *Datatype* in Relation.

Datatype: Ein Wert, der einem *Argument* durch eine Funktion zugewiesen wird.

Bei der letzten Gruppe von Annotationen, den *False Friends*, handelt es sich um gängige Formulierungsmuster, deren Form auf eine ontologische Relation hindeutet, die jedoch nicht zutreffend ist. Dies sei kurz anhand folgender Beispiele erläutert:

- (37) a. **Das Wort** Wachstum bezeichnet die Zunahme von Größe.
b. Nonverbale Kommunikation bezeichnet **einen Teil** der menschlichen Kommunikation.
c. **In der Biologie** bezeichnet Wachstum eine Zunahme von Größe.

Metasprache: Dieser Annotationstyp wird verwendet, wenn eine sprachliche Einheit in einem Text selber als solche bezeichnet wird. In dem Satz (37a) wird *Das Wort* als *Metasprache* annotiert. Die entsprechende Einheit könnte als Oberklasse missinterpretiert werden.

Formale Ambiguität: Dieser Annotationstyp ist auf den ersten Blick evtl. verwirrend. In (37b) wird *einen Teil* als *Formale Ambiguität* annotiert. Die Syntax mag zunächst darauf hindeuten, dass es sich hierbei um eine Oberklasse handelt; erst bei genauer Betrachtung wird deutlich, dass diese NP die Relation der Meronymie zwischen *Nonverbale Kommunikation* und *menschlichen Kommunikation* herstellt. Als *Formale Ambiguität* werden also Wortgruppen annotiert, die eine Relation zwischen anderen Einheiten im Satz an-

deuten.

Wissensgebiet: Dieser Typ wird für Formulierungen angesetzt, die das Wissensgebiet explizit benennen, auf das sich eine Definition bezieht, und können als Unterklasse interpretiert werden. In (37c) wird *In der Biologie* als *Wissensgebiet* annotiert. Die mögliche Unterklasse wäre dann *biologisches Wachstum*. Der ontologische Informationsgehalt dieser Formulierungen kann schnell übersehen werden.

Eine Bemerkung zur Terminologie ist noch notwendig: Häufig werden die Begriffe *Token* und *Worttoken* synonym verwendet, was in dieser Studie ausdrücklich nicht der Fall ist: *Token* meint ein konkretes Vorkommen. Eine konkrete Wortgruppe, die als Oberklasse annotiert wurde, ist bspw. ein Token des Annotationstyps Oberklasse. Ein einzelnes durch Leerzeichen von anderen abgetrenntes Wort ist ein Worttoken.

Die folgende Tabelle zeigt die Verteilung der Rollen im Gesamtkorpus:

Annotationstyp	Anzahl	Annotationstyp	Anzahl
Worttoken	47950	Possessor	36
Aequivalent	617	Possession	35
Unterklasse	574	Shape	33
Oberklasse	545	Resource	30
Meronym	338	Stimulus	28
Holonym	272	Substance	27
Instanz	203	Role	25
Komplement	151	Beneficient	23
Wissensgebiet	137	Medium	16
Klasse	131	Experiencer	15
Product	110	Content	12
Process	97	Mathematische Formulierung	7
Instrument	84	Receiver	6
Metaklasse	75	Sender	5
Datatype	65	Connected	4
Place	64	No-Meronym	4
Depicted	63	Connecting	3
Patient	61	Manner	3
Depicting	58	No-Agent	3
Argument	58	No-Instrument	2
Producer	54	No-Patient	2
Metasprache	53	No-Possessor	2
Agent	51	No-Holonym	1
Force	44	No-Producer	1
Rollenindikator	44	No-Product	1
Goal	41	No-Resource	1

Tabelle 6: Die Verteilung der Rollen im Gesamtkorpus

Die Gesamtansicht ist interessant bzgl. der Frage, wie die drei Kategorien, in die ontologische Relationen eingeteilt wurden – semantische Relationen, semantische Rollen und Funktionen –, statistisch verteilt sind. Es zeigt sich bzgl. des Type-Token-Verhältnisses (vgl. Scherer 2014: 38), dass die mengentheoretischen Annotationen mit großem Abstand am häufigsten im Korpus vertreten sind. Dies darf nicht verwundern, da sie vom Wissensgebiet unabhängig sind, während z.B. erwartet werden kann, dass bestimmte semantische Rollen eher in bestimmten Wissensgebieten auftreten. Bspw. wird eine semantische Rolle wie *Agent*, die also eine willentlich handelnde Entität bezeichnet, in einem Korpus mit Texten zur Erde und zum Kosmos nicht so häufig vertreten sein, wie in einem Korpus über Menschen.

Abschließend sei darauf hingewiesen, dass sprachliche Einheiten nicht immer eindeutig annotiert werden können. Die Problematik kann einerseits entstehen, wenn sich zwei Rollen sehr ähnlich sind, wie das bspw. bei *Possessor-Possession* und *Meronym-Holonym* der Fall ist (vgl. Kapitel 3.2). Manchmal ist die Formulierung irreführend. Wenn bspw. gesagt wird, dass ein Holzofen eben Holz verwende, legt das Verb nahe, es handle sich bei *Holz* um ein *Instrument*, wobei sachlich *Ressource* der angemessenere Annotationstyp wäre. Es kann aber auch der Fall sein, dass der Wissenszusammenhang eine eindeutige Annotation nicht zulässt. Wenn bspw. ein Bakterium aus einer Zelle besteht, bedeutet dies, dass ein Bakterium eine Zelle ist (vgl. Kapitel 9.1). Es ist also unklar, ob in diesem Fall die taxonomische oder die Teil-Ganzes-Relation anzusetzen ist. Die folgenden Kapitel widmen sich der Auswertung des Korpus, wobei viele solcher Zweifelsfälle diskutiert werden.

7 Mengentheoretische Relationen

Zunächst seien die mengentheoretischen Relationen im Gesamtkorpus entsprechend der Häufigkeit ihres Vorkommens dargestellt:

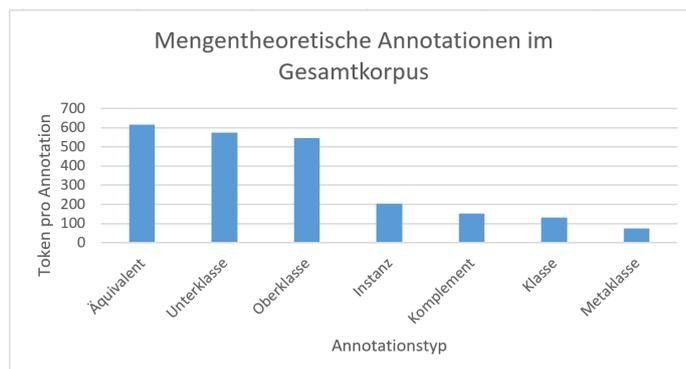


Abbildung 4: Die mengentheoretischen Relationen im Gesamtkorpus

Die Grafik zeigt deutlich, dass es zwei Gruppen gibt: Äquivalent sowie Ober- und Unterklasse sind mit viel mehr Token vertreten als die restlichen Annotationstypen. Eine Leitfrage des Kapitels soll sein, wie diese Verteilung zu begründen ist. Der Grund, dass Äquivalent so häufig vertreten ist, liegt daran,

dass Entitäten in der Relation gleich bezeichnet werden, d.h., immer wenn die Relation vorliegt, liegen mindestens zwei Token vom Typ Äquivalent vor. Wird die Tokenzahl von Äquivalent halbiert, liegt diese Annotation auf Platz drei zwischen Oberklasse und Instanz.

Um einen Vergleich zwischen Wikipedia und dem Duden anstellen zu können, wurde das Teilkorpus Wikipedia-Duden-Korrespondenz (WDK) erstellt, das nur die Wikipedia-Texte enthält, zu denen sich ein Text im Duden-Korpus befindet. Dies war ein erster Schritt der Normalisierung, da im Gesamtkorpus (GK) ungefähr fünfmal so viele Wikipedia-Texte wie Duden-Texte vertreten sind. Zu vielen Wikipedia-Artikeln existiert kein Pendant im Duden, da der Duden als Wörterbuch bspw. viele Komposita nicht erläutert: Wenn im Duden ein Artikel zu *wild* und ein Artikel zu *Rose* existiert, gibt es keinen Artikel zu *Wildrose*, während Wikipedia als Enzyklopädie eher einen solchen Artikel führt, da eine Enzyklopädie Konzepte erläutert und u.a. die Wörter selbst.

Da die Wikipedia-Artikel allerdings auch länger sind als die Duden-Artikel, ist das WDK immer noch bedeutend größer: Es umfasst 16.254 Worttoken, während das DK nur 9.992 Worttoken umfasst. Daher könnte die Idee naheliegen, das Korpus ein zweites Mal zu normalisieren, indem die jeweilige Tokenzahl des Duden-Korpus durch die eigene Worttokenanzahl geteilt und mal die des WDK genommen wird. Davon wurde aber abgesehen, da die geringere Größe des DK darin begründet liegt, dass Duden-Artikel sich eines sehr verkürzten Stils bedienen, der ohne komplette Sätze, teilweise nur mit einem Wort pro Eintrag, auskommt, während Wikipedia-Artikel sich einer komplexeren Syntax bedienen. Dies bedeutet aber nicht zwingend, dass im DK weniger ontologische Relationen an sich realisiert werden müssen. Wenn trotzdem weniger ontologische Relationen realisiert werden, dann handelt es sich dabei um eine redaktionelle Entscheidung, die durch eine weitere Normalisierung weniger sichtbar wäre.

Der Vergleich zeigt zunächst, dass im DK tatsächlich generell weniger on-

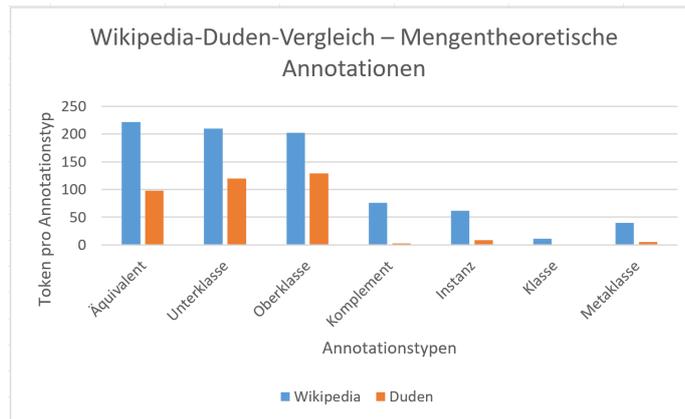


Abbildung 5: Vergleich zwischen Wikipedia und Duden: Mengentheoretische Relationen

tologische Relationen realisiert werden als im WDK – unabhängig vom Annotationstyp. Weiterhin zeigt sich, dass die Annotationen *Äquivalent*, *Unter-* und *Oberklasse* im WDK ungefähr doppelt so oft vertreten sind wie im DK, während die anderen mengentheoretischen Relationen im DK kaum vertreten sind. Die Vermutung liegt nahe, dass in den Wikipedia-Texten insgesamt mehr Relationen realisiert werden, um mehr Aspekte abzudecken, während die Duden-Artikel ökonomisch aufgebaut sind und nur so viele Relationen realisieren, wie notwendig ist, um die entsprechenden Konzepte zu definieren. Diese Hypothese soll in den folgenden Unterkapiteln überprüft werden.

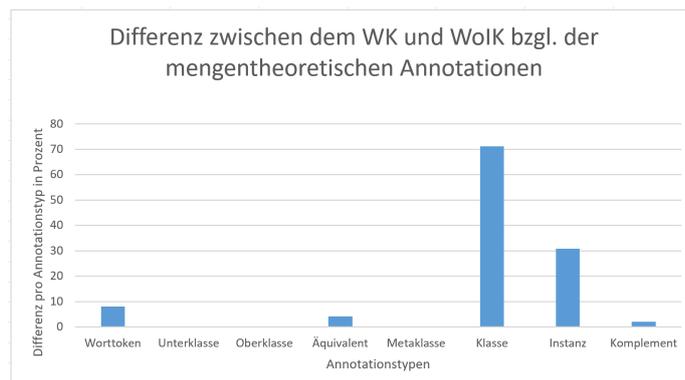


Abbildung 6: Die Differenz zwischen WK und WoIK: Mengentheoretische Annotationen

Abbildung 6 zeigt die Differenz zwischen dem WK und WoIK in Prozent; sie illustriert also welche Annotationen speziell durch die Texte hinzukommen, die Individuen zum Thema haben. Als Referenzwert ist die Differenz der To-

ken angegeben. Das Ergebnis ist nicht verwunderlich: Die mit Abstand größte Differenz besteht bei Klasse und Instanz. Dass Klasse eine noch höhere Differenz als Instanz aufweist, erklärt sich dadurch, dass eine Instanz zu mehreren Klassen gehören kann; es können also pro Instanz mehrere Klassen genannt werden. Bei allen anderen Annotationen liegt eine geringere Differenz vor als die Differenz der Worttoken; hier liegen also keine ungewöhnlichen Werte vor.

7.1 Taxonomische Relation

Die Taxonomie ist das Grundgerüst einer Ontologie. Da die taxonomische Relation durch das *genus proximum*, die nächste Oberklasse, angegeben wird, ist sie fundamentaler Bestandteil vieler Definitionen. Es handelt sich um eine zweistellige Relation, wobei die Entitäten Unterklasse und Oberklasse genannt werden. Da die Relation allerdings transitiv ist, können konkret mehrere Unter- oder Oberklassen genannt werden.

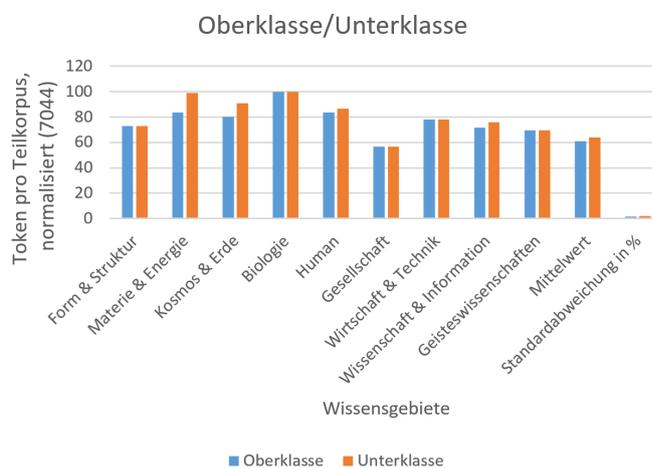


Abbildung 7: Die Verteilung des Annotationstyps Oberklasse über die Wissensgebiete

Die Statistik (Abbildung 7) zeigt wie die Annotationstypen Unter- und Oberklasse in den Wissensgebieten vertreten sind. Die Werte solcher Statistiken, die die Verteilung über die Wissensgebiete darstellen, sind bzgl. des größten Teilkorpus, Form und Struktur, das 7.044 Worttoken aufweist, nor-

malisiert. D.h., die anderen Teilkorpora wurden durch ihre jeweilige Anzahl an Worttoken geteilt und dann mal 7.044 genommen. Um die Standardabweichungen der verschiedenen Statistiken untereinander vergleichbar zu machen, werden diese in Prozent angegeben. Die Standardabweichung von Unter- und Oberklasse ist mit knapp unter zwei Prozent äußerst gering; die Typen sind also über alle Wissensgebiete gleichmäßig vertreten.

Die Relation kann mit dem Verb *sein* realisiert werden, was den Standardfall darstellt, da diese Formulierung den logischen Aspekt dieser Relation betont. Tatsächlich lassen sich viele Beispiele bei Wikipedia finden, die diesem Muster folgen:

- (38) a. ***Chladnische Klangfiguren sind Muster, die auf einer mit Sand bestreuten dünnen Platte (am besten aus Metall) entstehen, wenn diese in Schwingungen versetzt wird.*** (Wikipedia, Chladnische Figur)
- b. ***Energie ist eine fundamentale physikalische Größe, die in allen Teilgebieten der Physik sowie in der Technik, Chemie, Biologie und der Wirtschaft eine zentrale Rolle spielt. Ihre SI-Einheit ist das Joule.*** (Wikipedia, Energie)

Im Duden wird die taxonomische Relation, wenn sie genutzt wird, anders realisiert. Duden-Artikel sind sehr ökonomisch konzipiert, was bzgl. der taxonomischen Relation zur Folge hat, dass sie elliptisch realisiert wird: Die Unterklasse bildet den Titel des Artikels und nach eventuellen linguistischen Informationen wie Aussprache und Genus folgt nach einem Doppelpunkt die Oberklasse:

- (39) a. ***Quadrat: Rechteck mit vier gleich langen Seiten*** (Der Duden, Das Bedeutungswörterbuch: 763)
- b. ***Flüssigkeit: Ein Stoff in flüssigem Zustand.*** (Der Duden, Das Bedeutungswörterbuch: 398)

Dass diese syntaktischen Strukturen von Leser*innen angemessen interpretiert werden können, scheint ein Indiz dafür zu sein, dass die taxonomische Relation ein universelles Prinzip darstellt.

In den Wikipedia-Artikeln existieren noch viele andere Formulierungen, die diese Relation ausdrücken.

- (40) a. Als **Auslenkung** oder *Elongation* bezeichnet man bei einer Schwingung **die momentane Entfernung** x eines Punktes P (auf einer Kurve oder einer Geraden) von seiner Ruhelage. (Wikipedia, Auslenkung)
- b. Unter **Schifffahrt** (bis zur Rechtschreibreform 1996 und z. B. bei traditionellen Firmennamen weiterhin: *Schiffahrt*) versteht man **die Benutzung** von Wasserfahrzeugen auf Binnengewässern und Meeren zu unterschiedlichen Zwecken, vor allem zur gewerbsmäßigen Beförderung von Personen und Gütern. (Wikipedia, Schifffahrt)

Den beiden Beispielen liegt ein gewisses sprachliches Bewusstsein zugrunde: (40a) betont, dass es sich bei *Auslenkung* zunächst um ein Wort handelt, das eine gewisse Sache denotiert. Ebenso scheint (40b) durch das Verb *verstehen* zu betonen, dass es sich bei *Schifffahrt* um etwas handelt, das interpretiert werden muss, das also nicht per se für etwas anderes steht.

Generell scheint die Möglichkeit zu bestehen, sowohl Ober- als auch Unterklassen im Plural auszudrücken, wie Beispiel (41) belegt:

- (41) Als **Tontechniker** werden umgangssprachlich **Personen** bezeichnet, die im Umfeld der Tonaufnahme aktiv sind. (Wikipedia, Tontechniker)

Dies ist aber nur möglich, wenn der Plural nicht ein anderes Konzept nahelegt. Würde in (40b) für *Schifffahrt* der Plural verwendet werden, würde damit nicht mehr das generelle Konzept der Schifffahrt denotiert werden, sondern Fahrten mit einem Schiff; die NP *Schifffahrt* im Plural meint ein zählbares Konzept, die Singular-Verwendung in (40b) ein unzählbares. Dieser Unterschied wäre auch im Singular sichtbar: Das unzählbare Konzept steht dementsprechend ohne Artikel, während das zählbare mit Artikel auftritt.

Die Belege erwecken den Anschein, dass der bestimmte Artikel für die taxonomische Relation nicht in Frage kommt. Intuitiv mag man auch versucht sein, dem zuzustimmen, weil der bestimmte Artikel eine NP als partikulär kennzeichnet, während der unbestimmte Artikel eine NP als generisch kennzeichnet (vgl. von Heusinger 1997: 10-11). In dem Vokabular der formalen Semantik bedeutet dies, dass der bestimmte Artikel für Individuen und der unbestimmte für Mengen gedacht ist.

Dies ist aber nicht der Fall, wie Beispiel (42) belegt:

- (42) **Die Compact Disc Digital Audio** (kurz *CD-DA*, auch *Audio Compact Disc* oder *Audio-CD*) ist **ein optischer Massenspeicher**, der seit 1979 von *Philips* und *Sony* zur digitalen Speicherung von Musik entwickelt wurde. (Wikipedia, Compact-Disc)

Solche Fälle sind aus formaler Sicht widersprüchlich, obwohl sprachlich an ihnen kaum Anstoß zu nehmen ist. Zunächst ist zu bemerken, dass eine Realisierung der Unterklasse mit dem unbestimmten Artikel auch möglich wäre, ohne dass sich etwas an der Bedeutung des Satzes ändert. Als zweites fällt auf, dass in diesem Beispiel nur die Unterklasse mit dem bestimmten Artikel realisiert wird, während die Oberklasse mit dem unbestimmten Artikel steht, wobei es aber durchaus möglich wäre, beide NPn mit dem bestimmten Artikel zu bilden. Es kann vermutet werden, dass diese Formulierung die Compact Disc als ein ganz spezielles Speichermedium herausstellen will; aber dies muss an dieser Stelle Spekulation bleiben. Es gilt festzuhalten, dass Unter- und Oberklasse auch mit dem bestimmten Artikel realisiert werden können.

Bisher wurden Fälle betrachtet, in denen Unter- und Oberklasse als Ergänzungen des Verbs realisiert wurden, sie können aber auch als Attribut realisiert werden:

- (43) a. **Trittschall** ist eine besondere Form **von Körperschall**, der durch die Bewegung von Menschen und Tieren auf einem Fußboden entsteht. (Wikipedia, Trittschall)
- b. Als **parlamentarisches Regierungssystem** bezeichnet man jene Ausformungen **parlamentarischer Systeme westlicher Demokratien**, in denen die Regierung zu ihrer Wahl und in ihrer Amtsausübung auf die direkte oder indirekte Unterstützung durch das Parlament angewiesen ist. (Wikipedia, Parlamentarisches Regierungssystem)

In beiden Fällen wird die Oberklasse durch Attribute realisiert: einmal durch ein PP-Attribut (*von Körperschall*) und im anderen Fall durch ein Genitivattribut (*parlamentarischer Systeme*).

Besonders bei (43b) lohnt sich eine genauere Betrachtung. Die taxonomische Relation in diesem Beispiel ist folgende:

ParlamentarischesRegierungssystem \sqsubseteq
ParlamentarischesSystemWestlicherDemokratien

Die NP *jene Ausformungen*, von der die Oberklasse Genitivattribut ist, zeigt an, dass es sich bei *parlamentarischer Systeme westlicher Demokratien* um eine Oberklasse handelt. Bemerkenswert ist außerdem, dass die Oberklasse selber ebenfalls ein Genitivattribut besitzt, das von ihr abhängt. Dadurch wird indirekt eine zweite Oberklasse aufgemacht und mit wenigen Worten eine relativ komplexe Taxonomie erzeugt:

ParlamentarischesRegierungssystem \sqsubseteq
ParlamentarischesSystemWestlicherDemokratien \sqsubseteq
ParlamentarischesSystem

Zur Ableitung dieser Taxonomie ist eine Form von linguistischem Inferrieren notwendig: Wenn es parlamentarische Systeme westlicher Demokratien gibt, dann muss es andere parlamentarische Systeme geben, sonst wäre das Genitivattribut überflüssig, und dies bedeutet wiederum, dass eine Oberklasse parlamentarisches System existiert.

In all diesen NPn kommt das Wort *System* vor, es handelt sich also bei allen Begriffen um Systeme, womit dieses Konzept als oberste Oberklasse in die Taxonomie aufgenommen werden kann:

ParlamentarischesRegierungssystem \sqsubseteq
ParlamentarischesSystemWestlicherDemokratien \sqsubseteq
ParlamentarischesSystem \sqsubseteq
System

Es ist möglich, noch einen Schritt weiterzugehen. Die NP *parlamentarisches Regierungssystem* beinhaltet in sich ebenfalls eine taxonomische Relation:

ParlamentarischesRegierungssystem \sqsubseteq *Regierungssystem*

Es ist formal allerdings nicht ganz eindeutig, an welcher Stelle der Begriff *Regierungssystem* in der Taxonomie anzusetzen ist, weil das Lemma *Regierung* nicht wiederholt wird. Am naheliegendsten ist sicherlich folgende Einordnung:

ParlamentarischesRegierungssystem \sqsubseteq
ParlamentarischesSystemWestlicherDemokratien \sqsubseteq

ParlamentarischesSystem ⊆
Regierungssystem ⊆
System

Aus dem Hauptsatz in (43b) lässt sich also eine fünfstufige Taxonomie ableiten, wobei die einzelnen Relationen unterschiedlich zustande kommen: Durch das Prädikat, durch Komposition, durch zwei adjektivische Attribute und durch ein Genitivattribut.

Beleg (43b) hat gezeigt, dass Komposition ebenfalls eine Möglichkeit sein kann, die taxonomische Relation zu realisieren, dabei ist der Kopf des Kompositums die Oberklasse. Dieses gilt natürlich nur für endozentrische Komposita, während bei exozentrischen Komposita der Kopf nicht als Oberklasse zu verstehen ist. Bspw. ist ein Seehund kein Hund und eine Fledermaus keine Maus.

Die häufige Verwendung von Komposita im Deutschen führt dazu, dass manche Definitionen sich redundant lesen. In (44) ist dies z.B. der Fall, da die Oberklasse als Kopf des jeweiligen Kompositums schon realisiert ist und im Text daher wiederholt wird.

- (44) a. **Lebewesen: Wesen** mit organischem Leben, besonders Mensch oder Tier. (Der Duden, Das Bedeutungswörterbuch: 621)
- b. Unter **Unternehmenszielen** versteht man in der Betriebswirtschaftslehre **Ziele**, die der unternehmerischen Betätigung zugrunde liegen und der Förderung des Betriebszwecks in der Zukunft dienen; sie repräsentieren den vom Unternehmen angestrebten Soll-Zustand. (Wikipedia, Unternehmensziel)
- c. **Die Abrundungsfunktion** [...] und **die Aufrundungsfunktion** sind **Funktionen**, die jeder reellen Zahl die nächstliegende nicht größere bzw. nicht kleinere ganze Zahl zuordnen. (Wikipedia, Abrundungsfunktion und Aufrundungsfunktion)

Dieser Zusammenhang wird allerdings häufig antizipiert, sodass die Oberklasse nicht genannt wird, wenn sie bereits im Kompositum steckt (45a) oder zumindest den Kopf der NP bildet (45b).

- (45) a. **Ein Kommunalunternehmen** (abgekürzt *KU*, im Falle gemeinsamer Kommunalunternehmen auch *gKU*) umfasst Aufgabenbereiche einer kommunalen Körperschaft (Gemeinden, Landkreise und

Zweckverbände), die sich nach Art und Umfang für eine selbständige Wirtschaftsführung eignen und der öffentliche Zweck dies rechtfertigt.) (Wikipedia, Kommunalunternehmen)

- b. **Bei computervermittelter Kommunikation** nutzen Menschen Rechner zum Aufbau einer Datenverbindung sowie zum Austausch von Nachrichten und/oder weiteren Mitteilungen. (Wikipedia, Computervermittelte Kommunikation)

Diese Beispiele zeigen, dass Komposita als taxonomische Relationen interpretiert werden können. Wenn dies nicht der Fall wäre, würden die Belege (45a) und (45b) unvollständige Definitionen darstellen. Es zeigt aber auch, besonders im Zusammenhang mit Oberklassen, die durch Attribute realisiert werden, dass keinesfalls immer ein Verb diese Relation ausdrückt. Die Formalisierung solcher Fälle ist allerdings wieder redundant.

Kommunalunternehmen \sqsubseteq *Unternehmen*

Bisher wurden immer Beispiele angegeben, bei denen die erstgenannte NP die Unterklasse und die zweitgenannte NP die Oberklasse bildet. Dass dies aber nicht immer so sein muss, belegt das folgende Beispiel:

- (46) **Eine Substanz ist dann ein Gas**, wenn sich ihre Teilchen in großem Abstand voneinander frei bewegen und den verfügbaren Raum gleichmäßig ausfüllen. (Wikipedia, Gas)

Beleg (46) ist bemerkenswert, da er zeigt, wie subtil die Syntax des ganzen Satzgefüges ineinandergreift, um zur richtigen Interpretation zu führen: Würde (46) nur aus dem Hauptsatz bestehen, wäre ohne Weltwissen eine richtige Interpretation unmöglich: *Eine Substanz ist ein Gas*. Die Tatsache, dass der angeschlossene Nebensatz ein Konditionalsatz ist, ist ausschlaggebend dafür, die taxonomische Relation in der richtigen Richtung zu verstehen. Wäre ein Relativsatz angeschlossen, wäre eine richtige Interpretation ohne Weltwissen abermals unmöglich: *Eine Substanz ist ein Gas, das...* Die Partikel *dann* scheint den Zweck zu haben, den konditionalen Charakter der Aussage und damit auch die richtige Interpretation zu verdeutlichen.

Die Umkehrung der Reihenfolge von Unter- und Oberklasse ist auch bei anderen Formulierungen zu finden:

- (47) a. ***Unter Sprache*** versteht man im allgemeinen Sinn **alle komplexen Systeme der Kommunikation**. (Wikipedia, Sprache)
- b. ***Parasprache*** [...] bezeichnet **sämtliche, die Sprache begleitende, vokale, das heißt an Laute gebundene Mittel, die für die Kommunikation von Bedeutung sind**. (Wikipedia, Parasprache)
- c. ***Datenerfassung*** bezeichnet **alle Methoden der zeitgleichen oder zeitfolgerichtigen Messungen und Zählungen, gegebenenfalls einschließlich Zeitstempel für messbare oder zählbare Daten und Gruppen von zusammenhängenden Daten**. (Wikipedia, Datenerfassung)

In den Beispielen (47) wird die Umkehrung durch die Pronomen *alle* und *sämtlich* eindeutig. Würde dieses Artikelwort fehlen, wäre die Reihenfolge Unterklasse > Oberklasse wieder die naheliegende Interpretation, wobei die Sätze in diesem Fall unvollständig wirkten; ein angeschlossener Nebensatz wäre zu erwarten.

Die Beispiele (46) und (47) legen nahe, dass die Reihenfolge Unterklasse > Oberklasse als der unmarkierte Fall angesehen werden darf, wenn es um die Realisierung der taxonomischen Relation geht, da die andere Reihenfolge mit zusätzlichen Wörtern, bspw. *dann* und *alle*, einhergehen.

Beleg (48) ist interessant, weil eine Oberklasse, *Wechselgröße*, erst im zweiten Satz genannt wird, und die Relation über eine anaphorische Aufnahme hergestellt wird. Es stellt sich die Frage, ob es sinnvoll ist, zwei Oberklassen, *Maximalwert* und *Wechselgröße*, anzunehmen, oder ob nur *Wechselgröße* die Oberklasse darstellt. Rein sprachlich ist diese Frage nicht zu klären. *Maximalwert* als Oberklasse anzusetzen, ist deswegen kritikwürdig, weil viele Dinge in einem bestimmten Kontext der Maximalwert sein können. Maximalwert scheint keine Klasse zu sein, die sich über eine bestimmte Eigenschaft sinnvoll definieren lässt. Dies ist aber ein editorisches und kein formales Argument.

- (48) ***Die Schallschnelleamplitude*** ist der **Maximalwert der Schallschnelle**. Da diese **eine Wechselgröße** ist, erfolgt ihre Angabe meistens als **Effektivwert oder auch als Scheitelwert**. (Wikipedia, Schallschnelleamplitude)

Beispiel (49) bedient sich noch stärker der Strategie, weitere Oberklassen

durch anaphorische Aufnahme einzuführen. Wie die Oberklassen sich in diesem Beispiel zueinander verhalten, ist allerdings sehr schwer zu rekonstruieren. Es scheint in diesem Fall so zu sein, dass alle drei Oberklassen in sonst keinem spezifischeren Verhältnis zueinander stehen.

- (49) ***Der Schalldruckpegel** (englisch *Sound Pressure Level* und oft mit *SPL* abgekürzt) ist **eine logarithmische Größe** zur Beschreibung der Stärke eines Schallereignisses. Der Schalldruckpegel ist eine berechnete Größe, basierend auf dem Effektivwert des Schalldrucks. Er gehört zu **den Schallfeldgrößen**.* (Wikipedia, Schalldruckpegel)

Das folgende Beispiel ist außergewöhnlich, da die Oberklasse hier durch ein Pronomen realisiert wird, ohne dass dieses anaphorisch eine andere Einheit im Text aufgreift. Die Konstruktion ist interpretierbar, da das Pronomen *jemand* nur für Personen verwendet werden kann. Man könnte den Satz also folgendermaßen umformulieren: *Ein Regisseur ist eine Person, die Regie führt.*, und damit die Oberklasse deutlicher machen. Die Konstruktion ist aber sicherlich auch in dieser Form möglich, da es zum Weltwissen eines Menschen gehört, dass Regisseure Menschen bzw. Personen sind. Die Oberklasse zu realisieren ist in diesem Fall also nicht notwendig.

- (50) ***Regisseur: Jmd.**, der [berufsmäßig] Regie (1) führt, die Regie hat.* (Der Duden, Das Fremdwörterbuch: 963)

In der Formalisierung muss das Pronomen allerdings in eine sinnvolle Oberklasse umgewandelt werden:

Regisseur \sqsubseteq *Person*

In den Belegen (51) und (52) ergeben sich taxonomische Relationen dadurch, dass durch das Wort *wie* Beispiele eingeführt werden.

- (51) *Ein Gasplanet oder Gasriese („planetarer Gasriese“) ist in der Astronomie ein Riesenplanet, der überwiegend aus **leichten Gasen** wie **Wasserstoff** und **Helium** besteht.* (Wikipedia, Gasplanet)
- (52) *Der Nutzschaall oder das Nutzschaal ist das Gegenteil von Störschaall bzw. von Störschaal. Man unterscheidet **den Nutzschaall**, wie **Musik** oder **die Stimme** beim Gespräch, und **den Störschaall**, wie **Baustellen- oder Verkehrslärm**.* (Wikipedia, Nutzschaall)

Semantisch sind diese Formulierungen mit *wie* allerdings ambig, da es sich bei den Unterklassen auch um Instanzen handeln könnte, wie in folgender Formulierung: *Menschen wie Peter und Hans*. Die Eindeutigkeit ergibt sich hier daraus, dass es sich bei den entsprechenden Substantiven nicht um Eigennamen handelt.

Bzgl. der taxonomischen Relation lohnt sich eine gesonderte Betrachtung des Wissensgebietes der Biologie. Taxonomie ist ein Teilgebiet der Biologie, das sich mit der Ausarbeitung einer selbigen für Lebewesen beschäftigt. Die diversen Ober- und Unterklassen, die dabei ermittelt werden, füllen auch die Texte aus diesem Wissensgebiet in einer großen Anzahl. Als Beispiel soll folgender Text dienen:

(53) **Die Himbeere** (*Rubus idaeus*) ist eine **Pflanzenart** aus der Untergattung **Idaeobatus** in der Gattung **Rubus** innerhalb der Familie der **Rosengewächse** (*Rosaceae*). (Wikipedia, Himbeere)

In (53) werden, neben verschiedenen anderen mengentheoretischen Relationen, vier Oberklassen realisiert: *Pflanze*, *Idaeobatus*, *Rubus* und *Rosengewächse*. Es ergibt sich folgende Hierarchie:

$$\text{Himbeere} \sqsubseteq \text{Idaeobatus} \sqsubseteq \text{Rubus} \sqsubseteq \text{Rosengewächse} \sqsubseteq \text{Pflanze}$$

Die Hierarchie ergibt sich dabei teilweise aus einem Zusammenspiel der Metaklassen *Untergattung*, *Gattung* und *Familie* sowie der Präposition *innerhalb*. Die taxonomische Struktur des Satzes soll genauer betrachtet werden: Zuerst wird die unterste Unterklasse und dann die oberste Oberklasse realisiert; die anderen Klassen werden dann zwischen diesen beiden in der Reihenfolge, in der sie genannt werden, eingeführt. Dass *Pflanze* allerdings als höchste Oberklasse anzusetzen ist, erscheint zwar unmittelbar sinnvoll, wie zu dieser Interpretation gelangt werden kann, ist allerdings nicht ohne Weiteres zu erklären.¹⁸ Dies ist durch das Zusammenspiel der Metaklassen und der Präpositionen möglich, da das Wissen, dass eine Himbeere eine Pflanze ist, ohnehin vorhanden ist. Ohne dieses, kann die taxonomische Struktur des

¹⁸Die Tatsache, dass hier ein Kompositum *Pflanzenart* und nicht das Simplex *Pflanze* vorliegt, wird in Kapitel 7.5 behandelt.

Satzes kaum richtig entschlüsselt werden. Es wird ein gewisses Vorwissen benötigt, um den Satz richtig zu interpretieren.

Einen ähnlich exzessiven Gebrauch der taxonomischen Relation findet sich in der Mathematik bzw. in der Geometrie. In (54) werden verschiedene Ober- und Unterklassen des Begriffes Parallelogramm aufgezählt, deren Verhältnis zueinander wiederum nicht spezifiziert ist. Auffällig ist im Gegensatz zum Duden-Text (55), dass *Viereck* noch durch zwei weitere Adjektive spezifiziert wird. Dies ist insofern interessant, als das Verhältnis der beiden Adjektive zur *differentia specifica* nicht eindeutig ist. Ist ein Parallelogramm ein konvexes Viereck, bei dem die gegenüberliegenden Seiten parallel sind, oder jedes Viereck, bei dem die gegenüberliegenden Seiten parallel sind. Es verhält sich so, dass die Adjektive sich aus der Definition ergeben; sie sind also strenggenommen redundant.

(54) **Ein Parallelogramm** [...] ist **ein konvexes ebenes Viereck**, bei dem gegenüberliegende Seiten parallel sind. Parallelogramme sind **spezielle Trapeze und zweidimensionale Parallelepipede. Rechteck, Raute (Rhombus) und Quadrat** sind Spezialfälle des Parallelogramms. (Wikipedia, Parallelogramm)

Das Verhältnis der Oberklassen und auch der Unterklassen zueinander ist dabei nicht weiter spezifiziert. Der zweite und dritte Satz der Textes lassen sich folgendermaßen formalisieren:

$$\begin{aligned} (\text{Rechteck} \sqcup \text{Raute} \sqcup \text{Quadrat}) &\sqsubseteq \\ \text{Parallelogramm} &\sqsubseteq \\ (\text{Trapez} \sqcup \text{Parallelepipede}) &\sqsubseteq \text{Viereck} \end{aligned}$$

Die Vereinigungsmenge von Rechteck, Raute und Quadrat ist eine Unterklasse von Parallelogramm (was impliziert, dass die einzelnen Klassen ebenfalls Unterklassen von Parallelogramm sind), welche wiederum eine Unterklasse der Vereinigungsmenge von Trapez und zweidimensionalem Parallelepipede ist. Dass ein Parallelogramm eine Unterklasse von Trapez und Parallelepipede ist, wird durch die beiden Adjektive *speziell* und *zweidimensional* ausgedrückt. Genauso drückt das Wort *Spezialfälle* aus, dass *Rechteck*, *Raute*

und *Quadrat* Unterklassen von *Parallelogramm* sind. Ohne das entsprechende Wissen lässt sich aber nicht ermitteln, ob *Viereck* die höchste Oberklasse darstellt oder auf einer Ebene der Hierarchie mit *Trapez* und *Parallelepiped* steht.

An dieser Stelle lohnt sich ein direkter Vergleich mit dem Artikel im Duden:

- (55) ***Parallelogramm***: ***Viereck*** mit parallelen gegenüberliegenden Seiten (Der Duden, Das Fremdwörterbuch, 826)

Dieser nennt nur die Oberklasse *Viereck* und die *differentia specifica*. Dies ist ein deutlicher Beleg dafür, dass in den Wikipedia-Texten ontologische Relationen viel häufiger realisiert sind, als dies zur Definition des Konzeptes notwendig ist. So erklären sich die statistischen Unterschiede zwischen DK und WDK bzgl. der mengentheoretischen Relationen.

Beispiel (56a) nennt ebenfalls mehrere Oberklassen, erneut in einer umklammernden Reihenfolge: Zunächst wird eine Oberklasse genannt, die relativ hoch in der Taxonomie anzusetzen ist, von der aus sich über Unterklassen dem ursprünglichen Begriffe wieder genähert wird. Die eigentliche Reihenfolge wird dabei durch die Phrase *aus der Stoffgruppe* und die Formulierung *zählt dort* gesteuert.

- (56) a. ***Amphetamin*** (auch: *Phenylisopropylamin* oder *Amfetamin*) ist eine synthetische chemische Verbindung aus der Stoffgruppe der ***Phenylethylamine*** und zählt dort zu den ***Weckaminen*** (Amine mit „aufweckender“ Wirkung). (Wikipedia, Amphetamin)
- b. ***Amphetamin***: Indirekt wirkendes ***Sympathikomimetikum*** mit hohem Suchtpotenzial). (Der Duden, Das Fremdwörterbuch: 85)

Die Taxonomie, die in diesem Beispiel entsteht, ist also folgende:

Amphetamin \subseteq
Weckamin \subseteq
Phenylethylamin \subseteq *ChemischeVerbindung*

Interessanterweise nennt der Duden-Artikel (56b) eine andere Oberklasse.

In den folgenden Belegen werden besonders viele Unterklassen aufgezählt, da anscheinend der Anspruch einer (nahezu) vollständigen Aufzählung bestand:

- (57) *Als **Schwingungen** oder Oszillationen (lateinisch oscillare ‚schaukeln‘) werden wiederholte zeitliche Schwankungen von Zustandsgrößen eines Systems bezeichnet. [...] Man unterscheidet:*
- ***periodische** und **nichtperiodische** (quasiperiodische oder chaotische) **Schwingungen***
 - ***ungedämpfte, gedämpfte** und **aperiodische Schwingungen***
 - ***freie, erzwungene** (fremderregte), **selbsterregte** und **parametererregte Schwingungen***
 - ***lineare** und **nichtlineare Schwingungen***
 - ***Schwingungen mit einem Freiheitsgrad, mit mehreren Freiheitsgraden** und **mit unendlich vielen Freiheitsgraden** (Schwingungen eines Kontinuums)*
 - ***kontinuierliche Schwingungen** und **Oszillation zwischen diskreten Zuständen** (Wikipedia, Schwingung)*

Letztlich ließen sich die Unterklassen unter sich noch einmal in unbenannte Klassen ordnen: *Periodische* und *nichtperiodische Schwingungen* würden bspw. eine Klasse bilden, die eine Einteilung bzgl. der Periodik vorsieht.

Je allgemeiner ein Begriff ist, desto schwerer wird es, eine Oberklasse für ihn anzugeben. Dass trotzdem ein gewisser Zwang besteht, eine solche Oberklasse zu finden, verdeutlichen die folgenden beiden Beispiele:

- (58) a. ***Ein Winkel** ist in der Geometrie ein Teil der Ebene, der von zwei in der Ebene liegenden Strahlen (Halbgeraden) mit gemeinsamem Anfangspunkt begrenzt wird.* (Wikipedia, Winkel)
- b. ***Winkel: Geometrisches Gebilde** aus zwei Geraden, die von einem Punkt ausgehen* (Der Duden, Das Bedeutungswörterbuch, 1130)

In (58b) ist der Ausdruck *Gebilde* trotz der Ergänzung *Geometrisches* ein recht unspezifisches Wort. Im korrespondierenden Wikipedia-Artikel (58a) wird gar keine echte Oberklasse angegeben: Die NP *ein Teil* scheint zunächst eine zu sein, entpuppt sich aber als etwas, das in dieser Arbeit als Rollenindikator bezeichnet wird – es ist nicht sinnvoll eine Oberklasse *Teil* anzusetzen (siehe Kapitel 11.2).

Auch der folgende Beleg verdeutlicht, dass der Drang besteht, das Muster

ein x ist ein y in irgendeiner Weise zu erfüllen, da *Beziehung* letztlich nur ein anderes Wort für *Relation* ist. Auch der Relativsatz fasst nur in andere Worte, was sowohl für *Relation* und *Beziehung* gilt – dass sie zwischen Dingen bestehen. Der Satz folgt einem Muster, das für viele andere Konzepte funktioniert, bei dem allgemeinen Konzept *Relation* allerdings zu Redundanzen führt.

- (59) **Eine Relation** [...] ist allgemein **eine Beziehung**, die zwischen Dingen bestehen kann. (Wikipedia, Relation)

Eine Formalisierung wie die folgende ist nicht sinnvoll, da es sich eigentlich um äquivalente Begriffe handelt:

Relation \sqsubseteq *Beziehung*

Eine weitere Strategie mit sehr allgemeinen Konzepten umzugehen, besteht darin, mehrere Oberklassen zu nennen, die zueinander synonym sind:

- (60) **Eine Menge** ist **ein Verbund**, eine Zusammenfassung von einzelnen Elementen. (Wikipedia, Menge)

Auch hier lässt der Text offen, ob nicht eher alle drei Begriffe, *Menge*, *Verbund*, *Zusammenfassung*, zueinander synonym sind. Erneut ist die Spezifizierung von *einzelnen Elementen* wenig aussagekräftig, da ein Verbund oder eine Zusammenfassung immer aus irgendwelchen Elementen bestehen muss.

Die Strategie, mehrere Oberklassen zu nennen, findet sich auch im Duden. So werden in (61a) zwei Oberklassen genannt, da eine anscheinend nicht spezifisch genug ist, während der korrespondierende Wikipedia-Artikel (61b) mit *Begriff* gar keine echte Oberklasse nennt, sondern nur das syntaktische Muster der taxonomischen Relation erfüllt.

- (61) a. **Parameter: Klangeigenschaft der Musik, eine der Dimensionen des musikalischen Wahrnehmungsbereichs.** (Duden, Das Fremdwörterbuch: 828)
b. **Parameter** [...] ist ein aus den Naturwissenschaften entlehnter Begriff und bezeichnet in der Musik elementare Aspekte wie *Tondauer, Tonhöhe und Tonstärke bzw. Lautstärke*. (Wikipedia, Parameter)

Ein besonderer Fall liegt vor, wenn das im Text besprochene Konzept ein Adjektiv ist:

- (62) a. **rational**: *Alle Zahlen, die sich durch Brüche ganzer Zahlen ausdrücken lassen* (Der Duden, Das Fremdwörterbuch, 952)
- b. **Freiäugig** oder *freisichtig* werden in *Naturwissenschaft und Technik jene Beobachtungen und Messungen genannt, die ohne optisches Hilfsmittel (ausgenommen Korrekturen von Fehlsichtigkeiten der Augen) gewonnen werden, d. h. ohne Verwendung von zum Beispiel Fernrohr, Lupe, Mikroskop oder Kamera.* (Wikipedia, Freiäugig)

In (62a) lautet der Artikel – und damit die Unterklasse – nur *rational*, als Oberklasse wird dann ein Substantiv genannt. Auch wenn dies natürlich-sprachlich leicht verständlich ist, stellt sich die Frage, wie so eine Formulierung zu formalisieren ist. Eine Lösung besteht darin, das Substantiv der Oberklasse für die Unterklasse zu ergänzen. Eine solche Formalisierung sieht folgendermaßen aus:

RationaleZahlen \sqsubseteq *Zahlen*

7.2 Äquivalenz

Äquivalenz ist eine reflexive, transitive, symmetrische Relation. Zwei Begriffe sind äquivalent zueinander, wenn sie die gleiche Menge von Objekten denotieren. Das Standardbeispiel einer Äquivalenzrelation in natürlicher Sprache ist die Synonymie, wobei hier ein einfacher, extensionaler Synonymiebegriff im Sinne der vorausgegangenen Definition angesetzt wird, der Aspekte wie Konnotation vernachlässigt. Als Äquivalenzen werden im Folgenden aber auch Akronyme und Fremdwörter verstanden. Die Relation ist zweistellig, wobei beide Annotationen in der Relation Äquivalent heißen. In einem Text können allerdings mehr als zwei Begriffe als äquivalent zueinander ausgezeichnet werden. Wenn bspw. drei Begriffe A, B und C als äquivalent ausgezeichnet werden, liegt die Äquivalenz als Relation dreimal vor: Von A zu B, von A zu C und von B zu C.

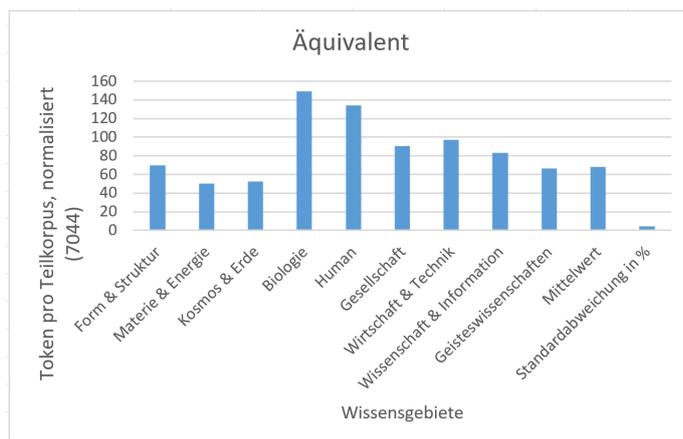


Abbildung 8: Die Verteilung des Annotationstyps Äquivalent über die Wissensgebiete

Die Statistik (Abbildung 8) zeigt, dass dieser Typ weniger gleichmäßig verteilt als Unter- und Oberklasse ist. Die Standardabweichung liegt bei vier Prozent. Dabei sind Materie und Energie, Biologie sowie der Humanbereich die Wissensgebiete, die die Standardabweichung ein wenig nach oben drücken.

Besonders in den Duden-Artikeln sind Äquivalenzen sehr häufig vertreten. (63) soll dabei als Beispiel dienen, wie in Duden-Artikeln ein Begriff durch Synonyme umrissen wird:

- (63) **Gravitation: *Schwerkraft, Anziehungskraft*, bes. die zwischen der Erde u. den in ihrer Nähe befindlichen Körpern.** (Der Duden, Das Fremdwörterbuch: 427)

Da $SHOIN(\mathcal{D})$ ein Zeichen für Äquivalenz vorhält, fällt eine Formalisierung leicht:

$$Gravitation \equiv Schwerkraft \equiv Anziehungskraft$$

Die häufige Verwendung von Äquivalenzen in den Duden-Artikeln führt oft zu Artikeln, bei denen letztlich nur Wortbildungsprozesse durchgeführt wurden:

- (64) **Wahrnehmung: *das Wahrnehmen*** (Der Duden, Das Bedeutungswörterbuch: 1104)

Die Äquivalenz in (64) ergibt sich daraus, dass man aus dem Verb *wahrnehmen* einmal durch Derivation das Substantiv *Wahrnehmung* und einmal durch

Konversion das Substantiv (*das*) *Wahrnehmen* ableiten kann. Es steht sicherlich zur Diskussion, ob die beiden Substantive wirklich äquivalent sind, ob *Wahrnehmung* nicht eher den gesamten Sinnesapparat, während *Wahrnehmen* den Prozess bezeichnet. Solche Fälle wurden dennoch als Äquivalenzen annotiert, da dies offensichtlich die Intention solcher Artikel ist: Ein Wort wird durch ein äquivalentes beschrieben.

Eine Form, äquivalente Klassen einzuführen, besteht in einer Auflistung.

- (65) a. **Ein regelmäßiges Polygon, reguläres Polygon, regelmäßiges Vieleck, reguläres Vieleck oder Isogon** [...] ist in der Geometrie ein ebenes Polygon, das sowohl gleichseitig als auch gleichwinklig ist. (Wikipedia, Regelmäßiges Polygon)
- b. **Ein Polygonzug oder Streckenzug** ist in der Mathematik die Vereinigung der Verbindungsstrecken einer Folge von Punkten. (Wikipedia, Polygonzug)
- c. [...] **Dieser sogenannte absolute Betrag, Absolutbetrag, Absolutwert** oder auch schlicht **Betrag** ist immer eine nichtnegative reelle Zahl. (Wikipedia, Betragsfunktion)

Es soll kurz reflektiert werden, was die Relation syntaktisch auszeichnet: Zunächst bilden die Begriffe in (65) zusammen das Subjekt des Satzes und damit auch eine NP. Ihre Einheit wird noch dadurch verstärkt, dass sie sich in beiden Beispielen den unbestimmten Artikel sozusagen teilen: Dieser wird nur einmal am jeweiligen Anfang des Satzes realisiert, bezieht sich aber auf alle NPn. Besonders in (65b) scheint dies ein wichtiger Zusammenhang zu sein: Wenn der unbestimmte Artikel in diesem Satz wiederholt würde, könnte der Eindruck entstehen, dass es sich um zwei verschiedene Klassen handle, die bestimmte Eigenschaften teilen *Ein Polygonzug oder ein Streckenzug ist...* Die Bedeutung von *oder* wäre in diesem Fall nicht eindeutig.

Im folgenden Beleg werden die äquivalenten Begriffe nicht durch eine Konjunktion, sondern durch die Abkürzung *bzw.* miteinander verbunden und als Äquivalente ausgezeichnet:

- (66) a. **Sedimentation** bzw. **Sedimentierung** [...] ist das Ablagern von Teilchen aus Flüssigkeiten oder Gasen unter dem Einfluss der Gewichtskraft oder der Zentrifugalkraft. (Wikipedia, Sedimentation)

Dieser Fall ist interessant, da hier, im Gegensatz zu (65a) und (65b), kein unbestimmter Artikel die beiden Substantive verbinden kann. Ein *oder* wäre in diesem Fall wieder ambig; die Abkürzung *bzw.* ist hingegen eindeutig.

Eine andere, naheliegende und eindeutige Strategie besteht darin, ein Synonym oder eine Aufzählung von Synonymen in Klammern hinter eine NP zu setzen, wie in den folgenden Beispielen:

- (67) a. **Die Abrundungsfunktion** (auch **Gaußklammer**, **Ganzzahl-Funktion**, **Ganzteilmfunktion** oder **Entier-Klammer**) und die **Aufrundungsfunktion** sind Funktionen, [...]. (Wikipedia, Abrundungsfunktion und Aufrundungsfunktion)
- b. Ein Viereck (auch **Tetragon**, **Quadrangel** oder **Quadrilateral**) ist eine Figur der ebenen Geometrie, nämlich ein Vieleck mit vier Ecken und vier Seiten. Analog zu Dreiecken ist auch eine Verallgemeinerung des Vierecksbegriffes auf nichteuklidische Geometrien (gekrümmte Vierecke) möglich. (Wikipedia, Viereck)
- c. **Mit trigonometrischen Funktionen** oder auch **Winkelfunktionen** (seltener: **Kreisfunktionen** oder **goniometrische Funktionen**) bezeichnet man rechnerische Zusammenhänge zwischen Winkel und Seitenverhältnissen (ursprünglich in rechtwinkligen Dreiecken). (Wikipedia, Trigonometrische Funktion)

Beispiel (67c) zeigt zudem, dass innerhalb der runden Klammern verschiedene, erklärende Wörter stehen können.

Abkürzungen wurden ebenfalls als Äquivalent annotiert. Dies mag auf den ersten Blick verwundern, da es sich bei Abkürzungen zunächst um etwas ganz anderes zu handeln scheint als bei Synonymen. Aus einer formalen Sicht sind es aber ebenfalls alternative Bezeichnungen für dieselbe Menge.

- (68) **Der Verwaltungsakt**, abgekürzt **VA**, stellt im deutschen Verwaltungsrecht eine Handlungsform der öffentlichen Verwaltung dar. (Wikipedia, Verwaltungsakt)

Zu diesen Fällen lässt sich allerdings nicht viel sagen. Selbst wenn nicht wie in (68) nochmal verdeutlicht wird, dass es sich um eine Abkürzung handelt, wird sie sofort als solche erkannt und auf die vorausgehende NP bezogen.

Eine weitere Form der Äquivalenz wird durch Fachbegriffe gebildet. Systematisch werden Fachbegriffe im Wissensgebiet der Biologie verwendet:

- (69) **Die Birnen (*Pyrus*) bilden eine Pflanzengattung, die zu den Kernobstgewächsen (*Pyrinae*) in der Familie der Rosengewächse (*Rosaceae*) gehört.** (Wikipedia, Birnen)

Bei Fachbegriffen scheint die Angabe in runden Klammern typisch zu sein. Würde die Relation bspw. mit *oder* realisiert werden (*Die Birnen oder *Pyrus* bilden eine Pflanzengattung*), läge nahe, dass es sich um gleichwertige Begriffe handelt, was in dem Sinne aber nicht der Fall ist, da der eine alltags-sprachlich und der andere fachsprachlich genutzt wird. Diese Unterscheidung ist formal jedoch irrelevant.

Die folgenden Beispiele nutzen beide Strategien:

- (70) a. **Als *Folge* oder *Sequenz* wird in der Mathematik eine *Aufzählung* (*Familie*) von endlich oder unendlich vielen fortlaufend nummerierten Objekten (beispielsweise Zahlen) bezeichnet.** (Wikipedia, Folge)
- b. **In der Mathematik ist eine *Funktion* (lateinisch *functio*) oder *Abbildung* eine *Beziehung* (*Relation*) zwischen zwei Mengen [...]** (Wikipedia, Funktion)

Die Vermischung beider Strategien (70a) könnte verwirren. Zunächst wird *Sequenz* als Äquivalent zu *Folge* unter der Verwendung von *oder* eingeführt. Dann wird *Familie* als Äquivalent zu *Aufzählung* in runden Klammern aufgeführt. *Familie* wird hier wahrscheinlich in Klammern gesetzt, um es als Fachbegriff zu kennzeichnen und einer Verwechslung mit der herkömmlichen Bedeutung des Wortes vorzubeugen.

In (70b) werden ebenfalls beide Strategien genutzt, wobei in diesem Beleg die runden Klammern das Fremdwort kennzeichnen und das deutsche Wort mit *oder* angeschlossen ist. Der Inhalt der runden Klammer hinter *Funktion* kann ebenfalls als Fachbegriff interpretiert werden, wobei es sich aber mehr um eine etymologische Erklärung handelt, die im Übrigen genauso auch für *Relation* hätte angegeben werden können.

Eine weitere Möglichkeit, äquivalente Ausdrücke einzuführen, bestände in der Verwendung von Appositionen. Es wäre evtl. zu diskutieren, ob von den genannten Fällen nicht schon einige als Appositionen verstanden werden

könnten, aber an dieser Stelle sollen unter (strengen) Appositionen vollständige NPn, in Abgrenzung zu bspw. (65a), verstanden werden, die nur durch Kommata getrennt hinter ihrem Bezugswort stehen. Diese Möglichkeit wird selten genutzt.

- (71) *Fixstern [...] ist eine auf die Antike zurückgehende Bezeichnung derjenigen Sterne, die am Himmel ihre Position nicht verändern und stets dieselbe Stellung zueinander einnehmen, im Unterschied zu den Wandelsternen, den Planeten.* (Wikipedia, Fixstern)

Die NP *den Planeten* in (71) bildet eine solche Apposition, und es liegt nahe, die beiden NPn deswegen als äquivalent zu interpretieren. Dieser Beleg ist jedoch verwirrend, da Wandelsterne und Planeten keine äquivalenten Konzepte sind – die Planeten sind abgesehen von der Erde eine Teilmenge der Wandelsterne. Fixsterne sollen im Gegensatz zu Wandelsternen und Planeten gedacht werden, was dann aber nicht hieße, dass die letzten beiden äquivalent sind. Die Syntax dieses Beleges ist missverständlich.

Kurzbezeichnungen sind ebenfalls eine Form von Äquivalenz:

- (72) *Das Ausbreitungsmedium oder kurz Medium bezeichnet in Physik und Technik einen Träger für die Ausbreitung eines Signals oder einer physikalischen Erscheinung.* (Wikipedia, Ausbreitungsmedium)

Dieser Fall bedarf besonderer Beachtung, da Komposita, wie im vorigen Kapitel ausgeführt wurde, häufig eine taxonomische Relation realisieren. Dies ist in (72) der Fall: Ein Ausbreitungsmedium ist ein Medium. Wenn die Begriffe *Ausbreitungsmedium* und *Medium* als äquivalent ausgezeichnet werden, bedeutet dies jedoch im Umkehrschluss, dass der Modifikator des Kompositums überflüssig ist, da er nicht weiter spezifiziert. Aus diesem Grund ist es fraglich, ob solche Abkürzungen als Äquivalenzen in eine Ontologie übernommen werden sollten.

In manchen Fällen besteht der Bedarf, die Äquivalenzrelation explizit zu benennen, d.h für sie einen eigenen Satz zu formulieren. Dies ist in folgenden Beispielen der Fall:

- (73) a. *Gelegentlich wird für **das kartesische Produkt** auch die mehrdeutige Bezeichnung „**Kreuzprodukt**“ verwendet.* (Wikipedia, Kartesisches Produkt)
- b. *Es gibt **die vier fundamentalen Wechselwirkungen** Gravitation, Elektromagnetismus, schwache Wechselwirkung und starke Wechselwirkung. Sie werden auch als **die vier Grundkräfte der Physik** oder als **Naturkräfte** bezeichnet.* (Wikipedia, Fundamentale Wechselwirkung)
- c. ***Willenskraft** gilt als Äquivalent für charakterliche Merkmale wie **Ausdauer (Beharrlichkeit), Zähigkeit, Entschlossenheit, Tatkraft, Robustheit oder Zielstrebigkeit.*** (Wikipedia, Willenskraft)

In (73a) liegt die Motivation, einen ganzen Satz zur Realisierung der Äquivalenzrelation aufzuwenden, wahrscheinlich darin, den vermeintlich äquivalenten Begriff als mehrdeutig zu kennzeichnen. In (73a) und (73c) lässt sich zumindest feststellen, dass die entsprechenden Begriffe Wörter sind, deren Form keinen Hinweis zu ihrer semantischen Verwandtschaft liefert. Die Begriffe *Abrundungsfunktion* und *Ganzzahl-Funktion* aus Beleg (67a) teilen dagegen als Kompositum denselben Kopf, wodurch eine Verwandtschaft der Begriffe angedeutet wird. Weil solche morphologischen Hinweise in (73a) und (73c) nicht vorhanden sind, besteht ein höheres Bedürfnis, die Äquivalenz sichtbar zu machen. Es kommt hinzu, dass es sich bei den Begriffen nicht um Fachbegriffe handelt, also um Wörter, die aus einem anderen Kontext bekannt und mit einer Bedeutung belegt sind; dass sie in diesem Kontext äquivalent sein können, ist daher erklärungsbedürftig. Bei (73c) stellt sich allerdings noch die Frage, ob die angeführten Begriffe nicht ausschließlich zusammengenommen äquivalent zu *Willenskraft* sind.

In dem folgenden Beleg werden verschiedene Äquivalenzen für den zentralen Begriff auf verschiedene Weise präsentiert:

- (74) ***Schallstrahlungsdruck** (auch: **Akustischer Strahlungsdruck**, engl.: **Acoustic radiation pressure**) ist ein physikalisches Phänomen, das seit den 1850er Jahren bekannt ist und erstmals von Rayleigh theoretisch beschrieben wird. [...] Da aufgrund des Druckes eine messbare Kraft auf das Hindernis ausgeübt wird, spricht man auch von **Schallstrahlungskraft**.* (Wikipedia, Schallstrahlungsdruck)

Bei diesem Beispiel werden zwei Äquivalenzen in runden Klammern genannt.

Interessant ist, dass hier das Adverb *auch* genutzt wird, um die Äquivalenz zu verdeutlichen. Ferner ist interessant, dass die englische Bezeichnung eingeführt wird, obwohl diese letztlich nur eine wörtliche Übersetzung der einzelnen Segmente des Kompositums ist. Wirklich heraus sticht aber, dass am Ende des relativ langen Textes nochmal eine Äquivalenz realisiert wird, die nicht nur eingeführt, sondern sogar begründet wird: Da es sich bei Schallstrahlungsdruck um eine Kraft handle, spreche man auch von *Schallstrahlungskraft*. Es liegen zwei äquivalente Komposita vor, die sich nur bzgl. ihres äußersten Kopfes unterscheiden, was – wenn man streng syntaktisch-formal schlussfolgerte – nahelegen könnte, dass *Kraft* und *Druck* äquivalent sind. Eine solche Schlussfolgerung würde die logische Konsistenz von natürlicher Sprache allerdings überschätzen.

Äquivalenz kann auch zwischen Individuen bestehen:

- (75) ***Die Andromedagalaxie, auch Andromedanebel oder Großer Andromeda-Nebel, ist die der Milchstraße nächstgelegene Spiralgalaxie, rund 2,5 Millionen Lichtjahre entfernt.*** (Wikipedia, Andromedagalaxie)

Dieser Fall ist syntaktisch schwer von dem zu unterscheiden, bei dem für eine Instanz mehrere Klassen aufgezählt werden, die zueinander äquivalent sind. Das folgende kreierte Beispiel wäre ein solcher Fall:

- (76) ***Angela Merkel, Bundeskanzlerin und Geschäftsfrau, residiert in einer Eigentumswohnung.***

In diesem Fall handelt es sich bei *Bundeskanzlerin* und *Geschäftsfrau* um Klassen und nicht um eine äquivalente Bezeichnung für Angela Merkel. Syntaktisch ist der Unterschied kaum zu erkennen; höchstens das Wort *auch* in (75) kennzeichnet die ihm folgenden Begriffe als Äquivalente zu dem vorausgehenden. Außerdem kann lexikalisch eine Unterscheidung getroffen werden, da *Angela Merkel* zwei Eigennamen sind, während es sich bei *Bundeskanzlerin* und *Geschäftsfrau* um Appellativa handelt. Allerdings spielt auch Weltwissen eine wichtige Rolle bei dieser Unterscheidung.

7.3 Komplement

Zwei Begriffe sind komplementär zueinander, wenn es keine Individuen gibt, die Instanz beider durch die Begriffe denotierten Klassen sind. Daraus folgt, dass die Relation symmetrisch ist. Sie korrekt zu formalisieren ist diffiziler, als es auf den ersten Blick scheinen mag, wie folgendes Beispiel demonstriert:

(77) *Fledermäuse sind keine Mäuse.*

Würden Fledermäuse als Komplement von Mäusen oder als Unterklasse dieses Komplements formalisiert werden, bedeutete dies modelltheoretisch zunächst nur, dass alles was keine Maus ist, eine Fledermaus sein könnte.

$Fledermaus \sqsubseteq \neg Maus$

Diese Aussage ist natürlich sehr unspezifisch. Das folgende Beispiel lässt zudem vermuten, dass in (77) zusätzliche Informationen präsupponiert werden.

(78) *Fledermäuse sind keine Autos.*

Beispiel (78) wirkt im Gegensatz zu (77) irrelevant oder selbstverständlich. Der Grund liegt darin, dass Sprachrezipient*innen bei der Bewertung einer Aussage nicht ihr gesamtes Weltwissen aktivieren. Natürlich-sprachliche Aussagen quantifizieren nicht über die gesamte Diskursdomäne (vgl. Barwise und Cooper 1981: 161 und Westerståhl 2019: 14), d.h., es wird nicht für jedes Individuum in der Diskursdomäne geprüft, ob es, wenn es eine Fledermaus ist, keine Maus ist, sondern nur für den Teil der Diskursdomäne, der für die jeweilige Aussage als relevant erachtet wird. Autos hingegen werden bei Aussagen über die Oberklasse von Fledermäusen von vornherein nicht als relevant erachtet. Es mag auf den ersten Blick naheliegen, dass nur für die Klasse der Fledermäuse geprüft werden muss, ob alle Instanzen nicht auch Instanzen der Klasse der Mäuse sind. Dieses Vorgehen würde aber nur funktionieren, wenn die entsprechende Klasse schon fest definiert ist, d.h. wenn schon bekannt ist, was alles eine Fledermaus ist; die Klassen, die in den Texten behandelt werden, werden allerdings erst definiert.

Ein beträchtlicher Teil dieses Kapitels wird darauf verwendet, zu zeigen, dass das *genus proximum*, also die Oberklasse, den Hintergrund bildet, vor dem Komplemente interpretiert werden, was in den Texten auf unterschiedliche Weise antizipiert wird.

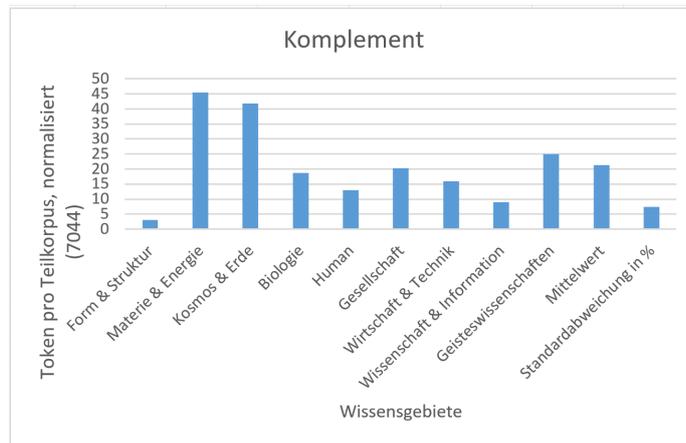


Abbildung 9: Die Verteilung des Annotationstyps Komplement über die Wissensgebiete

Der Annotationstyp ist in allen Wissensgebieten außer Form und Struktur moderat vertreten. Viele Token finden sich allerdings in Materie und Energie sowie Erde und Kosmos. Dies resultiert in einer Standardabweichung, die mit über fünf Prozent höher als die von Äquivalent und bedeutend höher als die von Unter- und Oberklasse ist.

Eine häufige Formulierung besteht darin, das Komplement als Präpositionalphrase mit dem Wort *neben* einzuführen.

- (79) a. **Obertöne** (auch *Partial-*, *Teil-*, *Aliquot-*, *Neben-* oder *Beitöne*) sind die *neben dem Grundton* mitklingenden Bestandteile eines *musikalisch instrumental oder vokal erzeugten Tones*. (Wikipedia, Oberton)
- b. **Die Tonhöhe** ist *neben der Tondauer, der Lautheit und der Klangfarbe* eine wichtige Eigenschaft *musikalischer Töne und der Vokale gesprochener Sprache*. (Wikipedia, Tonhöhe)

Eine sinnvolle Formalisierung ist nur in Kombination mit einer Oberklasse möglich:

$$\text{Oberton} \sqsubseteq (\text{Ton} \sqcap \neg \text{Grundton})$$

Es muss nur in der Klasse der Töne kontrolliert werden, ob alles was ein Ober-ton ist, kein Grundton ist. Um also ein Komplement sinnvoll formalisieren zu können, muss eine Oberklasse gefunden werden, die dem *definiendum* und seinem Komplement gemein und dabei nicht die Diskursdomäne ist. In den meisten Fällen erfüllt das *genus proximum* diese Funktion:

- (80) a. **Soziales System** ist ein zentraler Begriff der soziologischen Systemtheorie, der eine Grenze zieht zum **Ökosystem**, zum **biologischen Organismus**, zum **psychischen System** sowie zum **technischen System**. (Wikipedia, Soziales System)
- b. Eine **Platte** ist in der Technischen Mechanik bzw. in der Bautechnik ein in der Ebene ausgebreitetes Bauteil, das aus steifem Material besteht (ebenes Flächentragwerk) und durch senkrecht auf sie wirkende Kräfte und durch Momente um Achsen, die in der Plattenebene liegen, belastet wird. Im Gegensatz zur **Platte** wird **eine Scheibe** durch Kräfte in ihrer Ebene belastet. (Wikipedia, Platte)
- c. **Staub** [...] ist die Sammelbezeichnung für feinste feste Partikel verschiedener Größe und verschiedenen Ursprungs, die einen gewissen Zeitraum in Gasen, insbesondere in der Luft, suspendiert bleiben können (siehe Schwebstoffe). Es werden **Schwebstaub** und **Staubniederschlag** unterschieden. (Wikipedia, Staub)

In (80a) werden die Begriffe *soziales System*, *Ökosystem*, *biologischer Organismus*, *psychisches System* und *technisches System* als Komplemente realisiert. Sie haben die gemeinsame Oberklasse *System*, dabei muss jedoch geschlussfolgert werden, dass ein biologischer Organismus zumindest in diesem Fall auch als System zu verstehen ist, während die Schlussfolgerung für die anderen Komplemente zumindest sehr naheliegt, da das Wort *System* immer Kopf der entsprechenden NP oder des Kompositums ist.

In (80b) wird *Scheibe* im zweiten Satz als Komplement eingeführt. Auch hier muss die Schlussfolgerung erbracht werden, dass die Oberklasse von *Platte*, *Bauteil*, auch die Oberklasse von *Scheibe* ist, was in diesem Fall auf nahezu pragmatischer Basis geschehen muss, da es keine syntaktischen Hinweise gibt: Aufgrund der Maxime der Relevanz (vgl. Grice 1975: 47) kann ein*e Sprachrezipient*in schlussfolgern, dass eine *Scheibe* auch ein *Bauteil* ist. Das Weltwissen ist dabei selbstverständlich hilfreich.

In (80c) werden schließlich *Schwebstaub* und *Staubniederschlag* als Komplemente und Unterklassen von *Staub* eingeführt. Interessant ist hierbei, dass *Staubniederschlag* aufgrund seiner morphologischen Struktur eher als Unterklasse von *Niederschlag* verstanden werden könnte. Die taxonomische Struktur steht im Widerspruch mit der morphologisch-formalen Struktur; der Kopf des Kompositums ist nicht die Oberklasse.

Im folgenden Beleg wird ein Komplement in runden Klammern durch die Abkürzung *Ggs.* für *Gegensatz* eingeführt.

(81) **Variable:** *Veränderliche Größe (Ggs.: **Konstante**)*. (Der Duden, Das Fremdwörterbuch: 1158)

Das *genus proximum*, (mathematische) Größen, lässt sich in veränderliche und konstante Größen restlos aufteilen, was der Text allerdings nicht explizit aussagt. Aus ontologischer Sicht ist dies allerdings eine wertvolle Information. Klassen, die sich durch restloses Aufzählen ihrer Unterklassen bestimmen lassen, werden auch geschlossene Klassen genannt.

Auch im folgenden Beleg wird neben dem *definiendum* eine geschlossene Klasse definiert.

(82) **Die Schallkennimpedanz** [...], *auch akustische Feldimpedanz oder spezifische akustische Impedanz genannt, ist zusammen mit **der akustischen Flussimpedanz** und **der mechanischen Impedanz** eine der drei in der Akustik benutzten Impedanzdefinitionen*. (Wikipedia, Schallkennimpedanz)

Die Klasse *Impedanzdefinition* wird durch die Aufzählung ihrer Unterklassen definiert. In diesem Beleg wird allerdings explizit ausgedrückt, dass es sich bei den drei Unterklassen um eine vollständige Aufzählung handelt. Eine Formalisierung dieser Aussage könnte folgendermaßen aussehen:

Impedanzdefinition \equiv
(*Schallkennimpedanz* \sqcup
Akustische Flussimpedanz \sqcup
Mechanische Impedanz)

Die Klasse der Impedanzdefinitionen ist äquivalent zu der Vereinigungsmenge von Schallkennimpedanz, akustischer Flussimpedanz und mechanischer

Impedanz, d.h. wenn ein Individuum Instanz der Klasse Impedanzdefinition wäre, müsste es ebenfalls Instanz einer der drei Unterklassen sein.¹⁹ Bei dieser Formalisierung steht die Schallkennimpedanz nicht mehr im Zentrum, sie ist aber elegant, da sie Impedanzdefinition definiert und implizit die Information verarbeitet, dass Schallkennimpedanz, akustische Flussimpedanz und mechanische Impedanz Unterklassen von Impedanzdefinition sind. Dass diese drei Unterklassen zueinander komplementär sind, also die Schnittmengen leer sind, könnte durch folgende Formel noch ergänzt werden.

$$\perp \equiv (\text{Schallkennimpedanz} \sqcup \text{AkustischeFlussimpedanz} \sqcup \text{MechanischeImpedanz})$$

Die Definition von Schallkennimpedanz, die der Text eigentlich treffen will, ließe sich, unter Auslassung der Äquivalenzen, auf zwei Arten formalisieren, da sich hier die de-morganschen Gesetze anwenden lassen (vgl. Kapitel (3.3)):

$$\begin{aligned} \text{Schallkennimpedanz} &\sqsubseteq \\ &(\text{Impedanzdefinition} \sqcup \\ &\neg(\text{AkustischeFlussimpedanz} \sqcup \text{MechanischeImpedanz})) \end{aligned}$$

Oder:

$$\begin{aligned} \text{Schallkennimpedanz} &\sqsubseteq \\ &(\text{Impedanzdefinition} \sqcup \\ &(\neg \text{AkustischeFlussimpedanz} \sqcap \neg \text{MechanischeImpedanz})) \end{aligned}$$

Beide Formalisierungen sagen aus, dass die Schallkennimpedanz eine Impedanzdefinition ist und dabei weder die akustische Flussimpedanz noch die mechanische Impedanz ist.

Der folgende Beleg führt implizit zwei Komplemente ein:

- (83) *Als **Pflanzen** (lateinisch *Plantae*) werden *Lebewesen* bezeichnet, die sich nicht fortbewegen können und Photosynthese betreiben. **Pilze** und **Bakterien**, die früher auch als zum Pflanzenreich gehörig betrachtet wurden, sind heute ausgeschlossen.* (Wikipedia, Pflanze)

Der Text sagt aus, dass Pilze und Bakterien früher als Unterklassen von Pflanzen betrachtet wurden, nun aber von dieser Klasse ausgeschlossen sind. Dies

¹⁹Ob es überhaupt eine Instanz dieser Klasse geben kann oder ob diese abstrakt ist, ist dabei nicht relevant.

legt nahe, dass sie innerhalb der Klasse der Lebewesen Komplemente von Pflanzen sind. Diese Schlussfolgerung soll am Beispiel von Pilz etwas genauer betrachtet werden: Pilze wurden einst zu den Pflanzen gezählt, folgender Syllogismus war also gültig:

1. $Pilz \sqsubseteq Pflanze$
2. $Pflanze \sqsubseteq Lebewesen$
3. $Pilz \sqsubseteq Lebewesen$ (1, 2: Hypothetischer Syllogismus)

Aussage 1 wird im Text allerdings zurückgenommen, wodurch der ganze Syllogismus in sich zusammenfällt. Rein logisch kann also nicht geschlussfolgert werden, dass Pilze Lebewesen sind. Die Schlussfolgerung lässt sich nur aufrechterhalten, wenn davon ausgegangen wird, dass die Aussage, Pilze seien Pflanzen, präsupponiert, dass Pilze Lebewesen seien. Die Einteilung zu den Pflanzen ist dann eine weitere Spezifizierung – wird diese zurückgenommen, bleibt zumindest die Aussage stehen, dass Pilze Lebewesen sind.

Im folgenden Fall wird ein Komplement angegeben, für das es schwer ist, eine Oberklasse zu finden. Dies liegt aber u.a. daran, dass es auch schwer ist, eine Oberklasse für den in Frage stehenden Begriff zu finden:

- (84) **Risiko** weist je nach Fachgebiet einen unterschiedlichen Begriffsinhalt auf, allgemein wird hierunter die Möglichkeit des Eintritts künftiger Ereignisse, die nachteilige Auswirkungen wie Verlustgefahren in sich bergen, verstanden. Komplementärbegriff ist **die Sicherheit**. (Wikipedia, Risiko)

Als Oberklasse für *Risiko* wird *Möglichkeit* angegeben, und diese könnte auch als Oberklasse für den Komplementärbegriff *Sicherheit* angesetzt werden: So wie das Risiko die Möglichkeit des Eintritts eines Ereignisses ist, ist die Sicherheit, die Möglichkeit, dass es nicht eintritt. Diese Modellierung wäre im mathematischen Sinne komplementär. Es muss aber festgehalten werden, dass *Risiko* und *Sicherheit* sehr abstrakte Begriffe sind und sich diese Modellierung deswegen evtl. vom Sprachgefühl der Sprachproduzent*innen entfernt.

Bei den folgenden Belegen ergibt sich das Komplement dadurch, dass explizit Klassen genannt werden, die nicht die Eigenschaften des *definiendums* teilen:

- (85) a. **Fixstern** [...] ist eine auf die Antike zurückgehende Bezeichnung derjenigen Sterne, die am Himmel ihre Position nicht verändern und stets dieselbe Stellung zueinander einnehmen, im Unterschied zu **den Wandelsternen**, den Planeten. (Wikipedia, Fixstern)
- b. Ein **Blauer Riese** ist ein Riesenstern der Spektralklasse O oder B mit der 10- bis 50-fachen Sonnenmasse. Die Leuchtkraft **Blauer Riesen** ist höher als die **der Hauptreihensterne**. (Wikipedia, Blauer Riese)
- c. **Die Gravitation** (von lateinisch *gravitas* für „Schwere“), auch Massenanziehung oder Gravitationskraft, ist eine der vier Grundkräfte der Physik. Sie äußert sich in der gegenseitigen Anziehung von Massen. Sie nimmt mit zunehmender Entfernung der Massen ab, besitzt aber unbegrenzte Reichweite. Im Gegensatz zu **elektrischen** oder **magnetischen Kräften** lässt sie sich nicht abschirmen. (Wikipedia, Gravitation)

In (86a) und (86b) ergeben sich die Komplemente zusätzlich dadurch, dass ihre Position in einer auf Größe basierenden Hierarchie erläutert wird, wobei die Positionen vor und nach dem entsprechenden Ausdruck ebenfalls genannt werden: Dies sind die Komplemente. Viele Belege für dieses Muster finden sich im Bereich Erde und Kosmos. Während in der Biologie eher über Taxonomien klassifiziert wird, bilden im Bereich Erde und Kosmos detaillierte Hierarchien häufig die Grundlage für Klassifikationen. Dadurch ergeben sich in diesem Bereich viele Komplemente. Anbei weitere Belege:

- (86) a. **Zwerggalaxien** sind extragalaktische Sternsysteme, die zwischen **den kleineren Kugelsternhaufen** und **den größeren „normalen“ Galaxien** liegen. (Wikipedia, Zwerggalaxie)
- b. Als **Asteroiden** [...], Kleinplaneten oder Planetoiden werden astronomische Kleinkörper bezeichnet, die sich auf keplerschen Umlaufbahnen um die Sonne bewegen und größer als **Meteoroiden** (Millimeter bis Meter), aber kleiner als **Zwergplaneten** (ca. tausend Kilometer) sind. (Wikipedia, Asteroid)
- c. **Braune Zwerge** sind Himmelskörper, die eine Sonderstellung zwischen **Sternen** und **Planeten** einnehmen. (Wikipedia, Brauner Zwerg)

Das folgende Beispiel zählt alle Mitglieder einer Art auf und formuliert dadurch Komplemente:

- (87) *Man unterscheidet fünf Arten von Riesen.*
- **Unterriesen**, Sterne der Leuchtkraftklasse IV. Sie befinden sich im HRD zwischen dem Riesenast und der Hauptreihe.
 - **(Normale) Riesen** der Leuchtkraftklasse III. Sie bilden im HRD den Riesenast.
 - **Helle Riesen**, Sterne der Leuchtkraftklasse II. Sie finden sich im HRD oberhalb der normalen Riesen.
 - **Überriesen**, Sterne der Leuchtkraftklasse I. Aufgrund ihrer noch höheren Leuchtkraft liegen sie im HRD noch über den hellen Riesen.
 - **Hyperriesen**, Sterne der Leuchtkraftklasse 0. (Wikipedia, Riesenstern)

Die Klassen dieser Aufzählung lassen sich als Komplemente verstehen, da es der Anspruch einer solchen Aufzählung ist, dass die Klassen keine Schnittmenge haben. Dass ein solcher Anspruch besteht bzw. zu einer sinnvollen Kategorisierung gehört, ist wiederum Weltwissen. Ein ähnlicher Fall wurde schon im Zusammenhang der taxonomischen Relation besprochen (vgl. Beleg 57).

Das Komplement in (88) wird aus der Intention heraus formuliert, zwei ähnlich anmutende Konzepte voneinander abzugrenzen:

- (88) **Die zwanghafte Persönlichkeitsstörung** ist eine von **der Zwangsstörung** völlig verschiedene psychische Störung – trotz mancher Ähnlichkeiten in der sichtbaren Symptomatik. (Wikipedia, Zwanghafte Persönlichkeitsstörung)

Bei den folgenden Sätzen ergibt sich das Komplement dadurch, dass die jeweilige Klasse darüber definiert wird, was sie nicht ist. Hier ist gut zu sehen, wie die Definitionen sich tatsächlich ergänzen:

- (89) a. **Eine Republik** [...] ist eine Sammelbezeichnung für **alle nicht-monarchischen Staatsformen**. (Wikipedia, Republik)
- b. **Der Begriff Monarchie** [...] bezeichnet eine Staatsform mit einer Person, dem Monarchen, welcher das Amt des Staatsoberhauptes auf Lebenszeit oder bis zu seiner Abdankung innehat. **Die Monarchie** bildet somit das Gegenstück zum **modernen Republikbegriff**. (Wikipedia, Monarchie)

Der folgende Beleg veranschaulicht, wie verschiedene logische Relationen ineinandergreifen können, wodurch weitere logische Relationen offensichtlich werden:

- (90) *Unter Lernen versteht man **den absichtlichen** (intentionales Lernen) und **den beiläufigen** (inzidentelles und implizites Lernen) Erwerb von Fertigkeiten.* (Wikipedia, Lernen)

Zunächst gibt der Text nur an, dass absichtliches Lernen ein Komplement zu beiläufigem Lernen ist. Wenn *absichtliches Lernen* aber äquivalent zu *intentionalem Lernen* ist und *beiläufiges Lernen* äquivalent zu *inzidentellem bzw. implizitem Lernen* und außerdem *absichtliches Lernen* ein Komplement von *beiläufigem Lernen* ist, dann ist *intentionales Lernen* ein Komplement zu *inzidentellem bzw. implizitem Lernen*.

1. *AbsichtlichesLernen* \equiv *IntentionalesLernen*
2. *BeilaeufigesLernen* \equiv *InzidentellesLernern* \equiv *ImplizitesLernen*
3. *AbsichtlichesLernen* \neg *BeilaeufigesLernen*
4. *IntentionalesLernen* \neg *InzidentellesLernen*
5. *IntentionalesLernen* \neg *ImplizitesLernen*

Der folgende Beleg ist interessant, da die Komplemente *Zuschauerraum* und *Bühne* sich hier durch ihre räumliche Verteilung ergeben.

- (91) ***Bühne**: Vom **Zuschauerraum** abgegrenzte, meist erhöhte Fläche im Theater, auf der gespielt wird.* (Der Duden, Das Bedeutungswörterbuch: 250)

Was die beiden Begriffe zu Komplementen macht, sind nicht die ihnen inhärenten Eigenschaften, sondern dass sie unterschiedliche Bestandteile des Theaters sind.

Der letzte Text behandelt eine Instanz: das Theaterstück „Das Leben des Galilei“.

- (92) *Bertolt Brechts Theaterstück **Leben des Galilei**, von ihm meist als **episches Stück**, nicht als **Drama** bezeichnet, wurde 1939 im dänischen Exil verfasst und am 9. September 1943 in Zürich uraufgeführt.* (Wikipedia, Das Leben des Galilei)

In diesem Text werden allerdings zwei Klassen als Komplemente dargestellt. Das Besondere an diesem Fall ist der Hinweis, dass diese Unterscheidung

zunächst nur durch den Autor des Stückes getroffen wurde. Es bleibt demnach offen, ob einmal die Einteilung des Stückes zutrifft, aber auch, ob die beiden Klassen überhaupt als Komplemente zu verstehen sind.

7.4 Instanz-Klasse

Die Relation zwischen einer Instanz und ihrer Klasse wird prototypisch mit *sein* formuliert, also ähnlich der taxonomischen Relation. Ein Indiz für eine Instanz ist der bestimmte Artikel, während eine Klasse als generische NP normalerweise mit dem unbestimmten Artikel markiert wird. Dass dies nicht so sein muss, wurde schon in Unterkapitel 7.1 gezeigt. In den meisten Fällen sind Instanzen auch Individuen, weshalb davon auszugehen ist, dass die meisten Instanzen in den Texten annotiert wurden, die sich Individuen widmen.

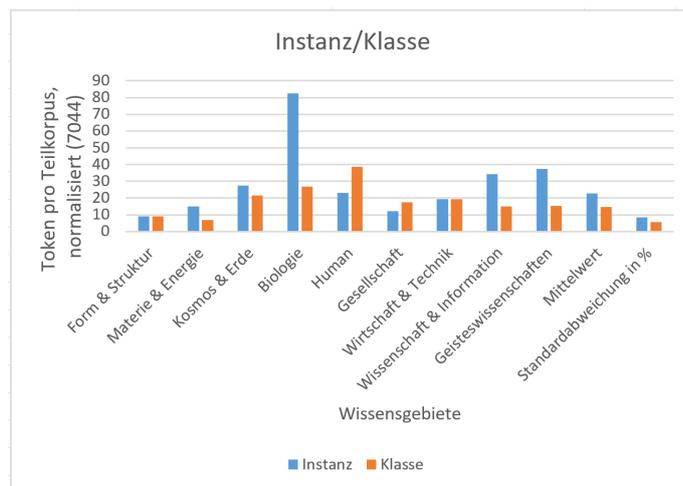


Abbildung 10: Die Verteilung der Instanzrelation über die Wissensgebiete

Zur Statistik muss zunächst gesagt werden, dass der Annotationstyp Instanz in zwei Relationen stehen kann: Er steht in Relation mit einer Klasse oder mit einer Meta-Klasse. Die Häufigkeit der Relation zwischen Instanz und Klasse lässt sich demnach nur an den Zahlen bzgl. Klasse ablesen, da der Typ Klasse nur in dieser Relation auftritt. Die Zahlen für Instanz sind in Abbildung 10 dennoch aufgeführt, weil der Vergleich zwischen den Wissensgebieten interessant ist. Die Statistik zeigt eine relativ hohe Standardabweichung bei Instanz (mit knapp unter zehn Prozent höher als bei Unter-, Oberklas-

se, Äquivalent und Komplement), die hauptsächlich durch das Wissensgebiet der Biologie verursacht wird. Der Typ Klasse ist hingegen sehr regelmäßig über alle Wissensgebiete verteilt. Auffällig ist zudem, dass im Humanbereich signifikant mehr Annotationen des Typen Klasse als des Typen Instanz vorliegen.

Zunächst sei gesagt, dass sich einige Beispiele finden lassen, die dem vermeintlichen Standard dieser Relation entsprechen, d.h., die Instanz ist eine NP mit bestimmtem Artikel und die Klasse eine NP mit unbestimmtem Artikel:

- (93) a. **Die Dreiecksungleichung** ist in der Geometrie **ein Satz**, der besagt, dass eine Dreiecksseite höchstens so lang wie die Summe der beiden anderen Seiten ist. (Wikipedia, Dreiecksungleichung)
- b. **Die Russellsche Antinomie** ist ein von **Bertrand Russell und Ernst Zermelo entdecktes Paradoxon** der naiven Mengenlehre, das Russell 1903 publizierte und das daher seinen Namen trägt. (Wikipedia, Russellsche Antinomie)

Die Klasse kann bei dieser Relation allerdings auch als eine NP mit bestimmten Artikel auftreten. Dies ist grammatisch der Fall, wenn ein Superlativ oder Zahladjektive verwendet werden. Folgendes Beispiel beinhaltet gleich beide Varianten.

- (94) **Die Venus** ist mit einer durchschnittlichen Sonnenentfernung von 108 Millionen Kilometern der zweitinnerste und mit einem Durchmesser von ca. 12.100 Kilometern **der drittkleinste Planet** des Sonnensystems. Sie zählt zu **den vier erdähnlichen Planeten**, die auch terrestrische oder Gesteinsplaneten genannt werden. (Wikipedia, Venus)

Der zweite Satz in (94) kann auf zwei verschiedene Arten formalisiert werden. Zunächst sagt er aus, dass die Venus ein Planet ist: *Planet(Venus)*. Es ließe sich aber auch die Klasse erdähnlicher Planet ansetzen: *ErdaehnlicherPlanet(Venus)*. Das Besondere an dieser Klasse ist, dass sie über eine Instanz, die Erde, definiert ist.

In Kapitel 3.5 wurde ausgeführt, dass sich Klassen über Relationen zu anderen Klassen definieren lassen. Beispiel (94) zeigt aber, dass sich Klassen auch in Relation zu einer Instanz definieren lassen. Solcher Beispiele gibt es

tatsächlich mehrere, wobei nicht immer die Instanzrelation vorliegt: In (95a) wird eine anonyme Klasse von Asteroiden und Kometen in Relation zu den Instanzen Jupiter und Neptun definiert, und in (95b) wird die Klasse der Exoplaneten in Relation zur Sonne definiert:

- (95) a. ***Ein Zentaur**, benannt nach den Kentauren der griechischen Mythologie, ist ein Mitglied einer Klasse von **Asteroiden** und **Kometen**, die sich zwischen den Umlaufbahnen von **Jupiter** und **Neptun** um die Sonne bewegen (siehe auch Liste der Asteroiden – Zentauren).* (Wikipedia, Zentaur)
- b. ***Ein Exoplanet**, auch extrasolarer Planet, ist ein planetarer Himmelskörper außerhalb [...] des vorherrschenden gravitativen Einflusses **der Sonne**, aber innerhalb des gravitativen Einflusses eines anderen Sterns oder Braunen Zwergs.* (Wikipedia, Exoplanet)

Ebenso gibt es Beispiele von Instanzen, die in Relation zu anderen Instanzen definiert bzw. bestimmt werden:

- (96) ***Die Sonne** ist **der Stern**, der der Erde am nächsten ist und das Zentrum des Sonnensystems bildet. Sie ist ein durchschnittlich großer Stern im äußeren Drittel der Milchstraße. Die Sonne ist ein Zwergstern, der sich im Entwicklungsstadium der Hauptreihe befindet.* (Wikipedia, Sonne)

Beispiel (96) sticht im doppelten Sinne hervor: Zunächst wird hier die Klasse, der Sonne als Instanz angehört, wieder als NP mit bestimmtem Artikel realisiert; dies liegt an dem Superlativ, der sich im Relativsatz anschließt. Weiterhin wird die Sonne in Relation zu der Instanz Erde definiert.

Ähnlich wie die taxonomische Relation durch ein Kompositum oder eine NP mit Adjektiv ausgedrückt werden kann, kann die Relation zwischen Instanz und ihrer Klasse durch ein Kompositum ausgedrückt werden. Die Klasse ist dabei der Kopf und die Instanz der Modifikator. Dass es sich um eine Instanz handelt, lässt sich erneut am bestimmten Artikel erkennen:

- (97) a. *Der **Kogelbaum** befindet sich in der niederösterreichischen Katastralgemeinde Senning (Gemeinde Kirchstetten).* (Wikipedia, Kogelbaum)
- b. *Die **germanische** (im germanistischen Zusammenhang „erste“) **Lautverschiebung** (fachsprachlich kurz „Grimm“, englisch auch*

Grimm's law genannt) kennzeichnet den Übergang vom (ur)indogermanischen zum (ur)germanischen Konsonantensystem. (Wikipedia, Erste Lautverschiebung)

- c. Die **Eulerschreibweise** ist nach dem Mathematiker Leonhard Euler benannt, der diese für die Feinanalyse von Musikstücken in reiner Stimmung einführte. (Wikipedia, Eulersches Tonnetz)

Die Tatsache, dass die Klasse schon im Namen steckt, scheint in diesen Belegen antizipiert zu werden, da sie nicht noch einmal angegeben wird, bspw.:
Der Kogelbaum ist ein Baum, der..

Die Abgrenzung zur taxonomischen Relation fällt manchmal schwer:

- (98) *Louis de Broglie entdeckte, dass alle Teilchen durch Materiewellen beschrieben werden können. **Die Wellenlänge** einer solchen Materiewelle wird **De-Broglie-Wellenlänge** genannt und hängt vom Impuls p des Teilchens ab.* (Wikipedia, De-Broglie-Wellenlänge)

Der Name *De-Broglie-Wellenlänge* legt nahe, dass es sich um eine Instanz handelt, da der Modifikator ein Eigenname ist. Inhaltlich handelt es sich aber um eine Klasse von Wellenlängen.

Auch bei der folgenden Aufzählung ist nicht klar, wie die einzelnen Elemente zu bewerten sind.

- (99) ***Schallfeldgrößen** sind physikalische Wechselgrößen, die zur Beschreibung des Zustands eines Schallfeldes verwendet werden. Zu den Schallfeldgrößen gehören:*
- ***der Schalldruck** und die abgeleitete logarithmische Größe **Schalldruckpegel***
 - ***die Schallschnelle** und die abgeleitete logarithmische Größe **Schallschnellepegel***
 - ***die Schallauslenkung***
 - ***die Schallbeschleunigung** (Partikelbeschleunigung)*
 - ***der Schallfluss** (Wikipedia, Schallfeldgröße)*

Der bestimmte Artikel deutet in diesen Beispielen daraufhin, dass es sich bei den Elementen der Aufzählung um Instanzen von der Klasse Schallfeldgröße handelt. Inhaltlich ist die Frage nicht eindeutig zu entscheiden. Wenn bspw. der Schalldruck als eine bestimmte aber abstrakte mathematische Größe gedacht wird, dann handelt es sich um eine Instanz von Schallfeldgröße; wenn der Schalldruck als das Ergebnis einer konkreten Messung gedacht wird, dann

ist ein bestimmtes Ergebnis eine Instanz, das zur Klasse Schalldruck gehört, welche dann eine Unterklasse von Schallfeldgröße ist. Das Beispiel zeigt, dass die Unterscheidung von Instanz und Klasse schwierig wird, wenn es sich um abstrakte Einheiten handelt.

Eine besondere Art von Instanzen stellen Personen dar. Aus formaler Sicht läge es erst einmal nahe, festzuhalten, dass es sich bei einer bestimmten Person eben um eine Person oder einen Mensch handelt. Sätze wie „*Keira Knightley ist ein Mensch, der...*“ oder „*Jackie Chan ist eine Person, die...*“ lassen sich allerdings kaum finden. Der Grund liegt auf der Hand: Der Eigenname kennzeichnet diese Instanzen schon als Personen. Als Klassen, die bzgl. einer Person angegeben werden, eignen sich besonders Berufsbezeichnungen:

- (100) a. *Keira Christina Knightley OBE [...] ist eine britische Schauspielerin.* (Wikipedia, Keira Knightley)
b. *Regina Halmich [...] ist eine ehemalige deutsche Boxsportlerin.* (Wikipedia, Regina Halmich)
c. *Jackie Chan, MBE [...] ist ein hongkong-chinesischer Schauspieler, Filmproduzent, Drehbuchautor, Filmregisseur, Stuntman und Sänger.* (Wikipedia, Jackie Chan)

Ein besonderes Augenmerk soll auf Beleg (100c) gelenkt werden, da Belege wie dieser erklären, warum die Annotation Klasse im Teilkorpus zum Humanbereich so oft vertreten ist.

Das folgende Beispiel zeigt indirekt, dass eine mit einem Eigennamen belegte Instanz zunächst als Person interpretiert wird, da in diesem Beispiel im Gegensatz zu (100) die Klasse explizit erwähnt wird.

- (101) *Yamataka Jindai Sakura [...] ist ein 1800 bis 2000 Jahre alter Kirschbaum auf dem Gelände des japanischen Tempels Jissō-ji in der Gemeinde Hokuto in der Präfektur Yamanashi.* (Wikipedia, Yamataka Jindai Sakura)

Beleg (102) ist interessant, da die Klasse, der die Instanz angehört, nie explizit genannt wird; sie wird implizit durch den Begriff *Ausgabe* eingeführt.

- (102) *Spektrum der Wissenschaft (SdW) wurde 1978 als deutschsprachige Ausgabe des seit 1845 in den USA erscheinenden Scientific American*

begründet, hat aber im Laufe der Zeit einen zunehmend eigenständigen Charakter gegenüber dem amerikanischen Original gewonnen.
(Wikipedia, Spektrum der Wissenschaft)

Der Text setzt bei den Leser*innen demnach ein gewisses Weltwissen voraus:
Es gilt zu wissen, dass Magazine als Ausgaben veröffentlicht werden.

7.5 Instanz-Metaklasse

Diese Relation ist analog zu der Relation Instanz-Klasse mit dem Unterschied, dass die Instanz kein Individuum, sondern selber eine Klasse ist. Ein Beispiel wäre der Begriff *Hunderasse*: Eine Instanz dieses Begriffes wäre Husky, wobei der Begriff *Husky* kein Individuum denotiert, sondern eine Klasse. Dieses Phänomen wird Metamodellierung genannt, weswegen für die entsprechenden Klassen der Ausdruck *Metaklasse* gewählt wurde (vgl. Hitzler u.a. 2008: 152). Wenn Klassen auch Instanzen sein können, stellt sich die Frage, worin der Unterschied zwischen der taxonomischen Relation und der Relation zwischen Instanz und zugehöriger Klasse liegt. Ein wichtiger Unterschied ist, dass die taxonomische Relation transitiv ist, die Relation zwischen Instanz und Klasse nicht, wie folgendes Beispiel illustriert: Hunde sind Säugetiere, und damit ist jede Instanz der Klasse Husky ein Säugetier. Eine Hunderasse dagegen ist eine unterscheidbare Züchtung von Hunden – diese Eigenschaft wird an die Rasse Husky vererbt. Ein bestimmter Husky, also eine Instanz der Klasse Husky, ist jedoch keine unterscheidbare Züchtung von Hunden.

Innerhalb der mengentheoretischen Annotationstypen bestehen bzgl. Metaklasse die größten Unterschiede im Korpus. In Form und Struktur, Materie und Energie sowie Kosmos und Erde ist dieser Typ gar nicht vertreten; im Bereich Biologie existieren ungemein viele Token, und in den restlichen Wissensgebieten ist er nur vereinzelt vertreten. Die Standardabweichung beträgt fast zwanzig Prozent. Die Zahlen zu den Instanzen sind dieselben, die schon aus Abbildung 10 bekannt sind. Die beiden Abbildungen zusammen erklären, warum im Wissensbereich der Biologie so viele Annotationen vom Typ

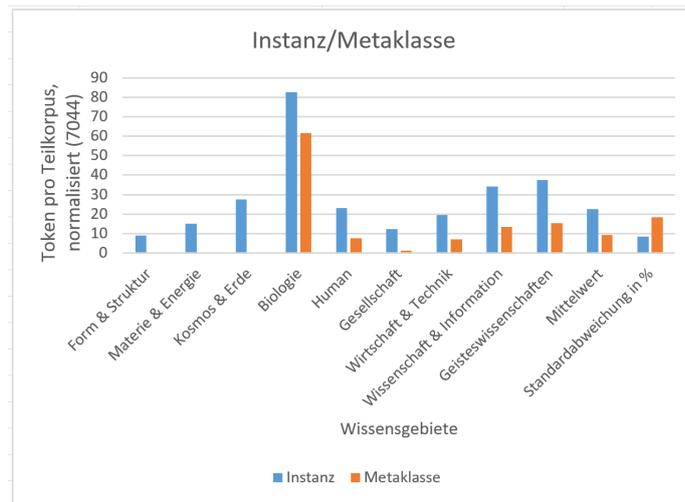


Abbildung 11: Die Verteilung von Instanz und Metaklasse über die Wissensgebiete

Instanz vorliegen: weil in diesem Korpus viele Metaklassen existieren, zu denen eine Instanz zusätzlich in Relation steht.

Es ist nicht zu erwarten, dass für diese Relation andere Formulierungen verwendet werden, als für die taxonomische Relation oder die herkömmliche Relation zwischen Instanz und Klasse, wodurch sich aber die Frage ergibt, ob diese Relation grammatisch von den anderen überhaupt unterscheidbar ist.

Zunächst lassen sich Instanz und Metaklasse ebenfalls durch *sein* verbinden:

- (103) *Die Erdbeeren (Fragaria) sind eine Gattung in der Unterfamilie der Rosoideae innerhalb der Familie der Rosengewächse (Rosaceae).* (Wikipedia, Erdbeeren)

Erdbeere wird in diesem Beispiel einmal als Unterklasse von *Rosoideae* und *Rosengewächse* und als Instanz von *Gattung* ausgewiesen. Zusätzlich stehen die Metaklassen in einer taxonomischen Relation: *Gattung* ist eine Unterklasse von *Unterfamilie*. Dass es sich bei *Rosengewächse* und *Rosoideae* um Klassen handelt, wird abgesehen von den Metaklassen *Gattung* und *Unterfamilie* zusätzlich durch den Plural angedeutet. Diese Informationen müssen durch drei separate Formeln ausgedrückt werden:

Erdbeere \sqsubseteq *Rosengewächse* \sqsubseteq *Rosoideae*
Gattung(*Erdbeere*)
Gattung \sqsubseteq *Unterfamilie*

Die zweite Formel stellt die Relation zwischen Klasse und Metaklasse her.

Die Relation wird häufig durch das Verhältnis eines Modifikators zu seinem Kopf ausgedrückt. In dem folgenden Beleg ist die Instanz als Genitivattribut realisiert.

- (104) **Einige Gattungen der Rosengewächse bilden Sammelnussfrüchte aus.** (Wikipedia, Sammelnussfrucht)

Die folgenden Texte führen alle, wie schon Beleg (103), vor, wie komplex verflochten die doppelten taxonomischen Verhältnisse durch den Gebrauch von Metaklassen werden können:

- (105) a. **Prunus ist eine Pflanzengattung innerhalb der Familie der Rosengewächse (Rosaceae).** (Wikipedia, Prunus)
b. **Die Rosoideae sind eine Unterfamilie innerhalb der Familie der Rosengewächse (Rosaceae).** (Wikipedia, Rosoideae)
c. **Amphetamin (auch: Phenylisopropylamin oder Amfetamin) ist eine synthetische chemische Verbindung aus der Stoffgruppe der Phenylethylamine und zählt dort zu den Weckaminen (Amine mit „aufweckender“ Wirkung).** (Wikipedia, Amphetamin)
d. **Die Rosengewächse (Rosaceae) sind eine Pflanzenfamilie in der Ordnung der Rosenartigen (Rosales) innerhalb der Kerneudikotyledonen.** (Wikipedia, Rosengewächse)

Prunus ist bspw. eine Unterklasse von *Rosengewächs* und eine Instanz der (Meta-)Klasse *Pflanzengattung*, welche wiederum eine Unterklasse von *Familie* ist. Es ergibt sich in diesen Beispielen eine doppelte Taxonomie, einmal die der Klassen und „daneben“ die der Metaklassen.

Diese doppelte Taxonomie scheint aber typisch für die Naturwissenschaften, und dort allen voran für die Biologie, zu sein. In anderen Wissensbereichen werden Metaklassen nur selten eingeführt:

- (106) **Die doppelte Buchführung bzw. Doppik – auch kaufmännische Buchführung genannt – ist die in der privaten Wirtschaft vorherrschende Art der Finanzbuchhaltung.** (Wikipedia, Doppelte Buchführung)

In Beispiel (106) wird durch das Schlüsselwort *Art* strenggenommen eine Metaklasse eingeführt. In diesem Wissensbereich wird sich aber keine Hierarchie

im Sinne von Art, Gattung, Familie und Ähnlichem ergeben, deswegen wäre es in solchen Fällen überzogen, in einer Ontologie tatsächlich eine Metaklasse anzusetzen. Es genügt, doppelte Buchführung als Unterklasse von Finanzbuchhaltung zu modellieren.

Ähnlich verhält sich das folgende Beispiel:

- (107) *Epik [...] ist neben Dramatik und Lyrik eine **der drei großen Gattungen** der Literatur und umfasst erzählende Literatur in Vers- oder Prosaform.* (Wikipedia, Epik)

Bei Gattung handelt es sich um eine Metaklasse, ihre Erwähnung liegt aber evtl. darin begründet, dass es schwer ist, eine Oberklasse für Dramatik anzugeben. Die Oberklasse ist das Genitivattribut *Literatur*, diese aber direkt anzugeben, bspw. „*Epik ist Literatur*“, klingt ungenau. Dies scheint der Grund zu sein, hier auf die Formulierung mit der Metaklasse zurückzugreifen.

Abschließend soll noch ein Blick auf die Definition von Begriffen, die Metaklassen bezeichnen, geworfen werden:

- (108) *Seit den Anfängen der **Rosenzüchtung** im 18. Jahrhundert bis heute sind weltweit über 30.000 **Rosensorten** entstanden. Sie werden gärtnerisch verschiedenen **Rosenklassen** zugeordnet, die in Gruppen zusammengefasst werden können.* (Wikipedia, Rosenklasse)

Der Text erläutert die Metataxonomie bzgl. Rosen:

Rosensorte \sqsubseteq *Rosenklasse* \sqsubseteq *Gruppe*

Der Text vermeidet jegliche Formulierungen, die sonst die taxonomische Relation ausdrücken. Durch die Verben *zuordnen* und *zusammenfassen* wird stattdessen explizit die Hierarchie dieser Metataxonomie ausgeführt.

Die Definitionen der Begriffe *Gruppe* und *Klasse*, die sich im Duden finden lassen, beschreiben das Konzept der formalen Klasse. Die beiden Definitionen sind effektiv identisch.

- (109) a. ***Gattung: Gruppe** von Dingen, Lebewesen, die wichtige Merkmale oder Eigenschaften gemeinsam haben.* (Der Duden, Das Bedeutungswörterbuch: 422)

- b. **Klasse: Gruppe** von Lebewesen, Dingen, die durch gemeinsame Merkmale, Eigenschaften, Fähigkeiten o.Ä. gekennzeichnet sind. (Der Duden, Das Bedeutungswörterbuch: 568)

Die Beispiele (108) und (109) deuten daraufhin, dass das Konstrukt der Metaklasse in der Sprachkompetenz der Sprachproduzent*innen nicht stark vertreten ist, da keine prototypischen Formulierungen existieren und die Bedeutung undifferenziert erläutert wird. Metaklassen ergeben sich aus dem stringenten Fortführen des mengentheoretischen Ansatzes, der sich zu Teilen eignet, die Semantik der natürlichen Sprache zu beschreiben, der mit Metaklassen aber in einen Bereich führt, der sich dem Sprachgefühl entzieht.

7.6 Zusammenfassung

Die Betrachtung der mengentheoretischen Relationen hat zunächst gezeigt, dass der Variationsreichtum natürlicher Sprache nicht unterschätzt werden darf: Es existieren diverse Formulierungen für die jeweiligen Relationen. Bzgl. der taxonomischen Relation gilt es hervorzuheben, dass diese sowohl durch Verben, als auch innerhalb einer NP realisiert werden kann, dabei stellt der Kopf die Oberklasse und der Modifikator die Unterklasse dar. Wird diese Relation jedoch durch ein Verb realisiert, besteht die Tendenz, dass das Subjekt die Unterklasse bildet. Ebenso konnte gezeigt werden, dass es Muster gibt, die formal-syntaktisch die taxonomische Relation ausdrücken, aber nicht als solche interpretiert werden sollten. Äquivalenzen sind in den meisten Fällen als Appositionen oder als Einschübe in Klammern realisiert. Im Duden werden sie gelegentlich alleine stehend für eine Definition eingesetzt. Beim Komplement wurde herausgearbeitet, dass eine Oberklasse der in Frage stehenden Begriffe immer mitgedacht und teilweise ergänzt werden musste, da dieses den Rahmen, in dem die Komplemente zu verstehen sind, eingrenzt. Bzgl. der Relation zwischen einer Instanz und ihrer Klasse galt es zunächst, herauszuarbeiten, wie diese Relation von der taxonomischen unterschieden werden kann. Hier konnten Tendenzen ausgemacht werden, wie die Realisie-

rung von Instanzen als NPn mit bestimmtem Artikel, die aber keinesfalls als ausnahmslos zu betrachten sind. Ferner konnte gezeigt werden, dass die Unterscheidung zwischen einer Klasse und einer Instanz nicht immer eindeutig ist. Das Verhältnis von Instanz und Metaklasse war der diffizilste Teil dieses Kapitels. Es entpuppte sich schließlich als ein Konstrukt, das sich zwar formal-logisch definieren lässt und besonders im Wissensgebiet der Biologie auftritt, das sprachlich aber keine eigenen Formulierungen hervorbringt, sondern sich derselben Formulierungen wie die taxonomische Relation und die Instanzrelation bedient.

8 Semantische Rollen als ontologische Rollen

Dieses Kapitel widmet sich den semantischen Rollen, wobei weniger Fokus auf die konkreten Formulierungen gelegt wird, als dies im vorausgegangen Kapitel geschehen ist. Der Grund ist folgender: Für die mengentheoretischen Relationen existieren Standardformulierungen, und es war ein Ziel des entsprechenden Kapitels, zu zeigen, wie groß der Variationsreichtum neben diesen Standardformulierungen ist. Ähnliches galt für die Pertinenzrelationen. Im Gegensatz dazu wurden die semantischen Rollen schon als ein Phänomen eingeführt, das sich unter Umständen weit von der sprachlichen Oberfläche entfernen kann (vgl. Kapitel 3.2). Daher soll in dieser Arbeit nicht in aller Ausführlichkeit diskutiert werden, wie bspw. ein *Agent* im Deutschen realisiert wird. Dies ist ohnehin ein Thema, dem sich andere Arbeiten ausführlich gewidmet haben (vgl. Primus 2012). In diesem Kapitel steht vielmehr im Vordergrund, wie sich die einzelnen semantischen Rollen formalisieren lassen. Hier steht Ontolog*innen ein größerer Spielraum zur Verfügung, als dies bei den mengentheoretischen Relationen der Fall war.

In diesem Kapitel wird folgenden Fragen nachgegangen: Wie werden die semantischen Rollen in den Texten realisiert? Wie lassen sich diese, wenn sie identifiziert wurden, in ontologische Rollen übertragen? Welche Details

können bei diesem Prozess evtl. vernachlässigt werden? Dabei muss bedacht werden, dass diese Arbeit ein generelles Konzept vorstellt, mit dem Informationen aus einem Text in eine Ontologie überführt werden können. Dieses Konzept muss aber in einem konkreten Fall noch den Ansprüchen der jeweiligen Ontologie angepasst werden. Es ist z.B. denkbar, dass die Unterscheidung von *Force* und *Agent* für manche Ontologien unerheblich und für andere noch nicht differenziert genug ist.

Um einen Überblick über die Verteilung der semantischen Rollen zu gewinnen, lohnt es sich, zunächst das DK mit dem WDK zu vergleichen. Abbildung 12 zeigt zunächst, dass in Wikipedia anscheinend viel mehr Rollen realisiert werden, als im Duden. Das verwundert nicht, weil die Wikipedia-Texte ausführlicher sind. Einige Ausnahmen existieren allerdings: *Product*, *Role*, *Shape* und *Substance* kommen häufiger im DK vor.

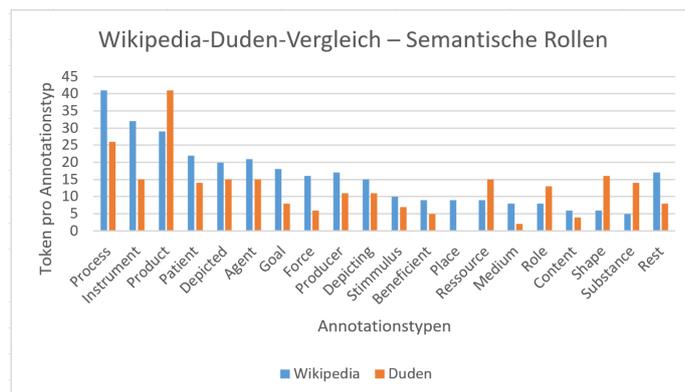


Abbildung 12: Der Vergleich zwischen Wikipedia und Duden: Semantische Rollen

Abbildung 13 zeigt die Differenz in den Annotationstoken zwischen dem WK und dem WoIK. Als Vergleichswert werden die Worttoken mit angeführt:

Der einzige erklärungsbedürftige Wert ist der von *Place*. Alle anderen liegen bei dem Vergleichswert oder bei Null. Bzgl. *Place* lässt sich in Abbildung 12 ablesen, dass dieser Annotationstyp im DK gar nicht vertreten ist. Woran das liegen kann, wird im entsprechenden Unterkapitel (8.8) erläutert.

Bei der Betrachtung von (12) fällt auf, dass die klassischen semantischen Rollen, *Agent*, *Patient*, *Theme*, nicht so weit oben in der Statistik auftauchen,

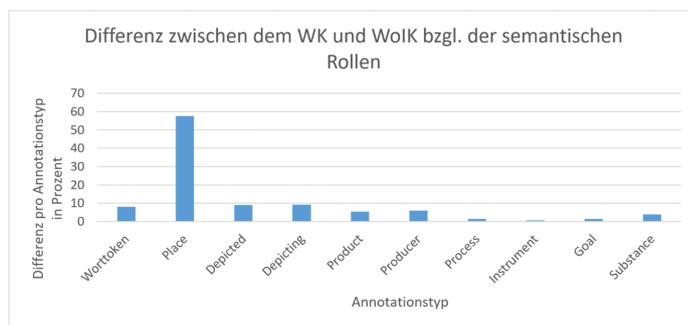


Abbildung 13: Die Differenz zwischen WK und WoIK

wie man es vielleicht erwartet hätte. Dies liegt daran, dass die Rollen spezifiziert wurden, und eine allgemeinere Rolle wie bspw. *Agent* nur als Annotation angesetzt wurde, wenn keine spezifischere angegeben werden konnte. Dieses Vorgehen führt dazu, dass bspw. die Rolle *Theme*, die eigentlich zum Standardrepertoire gehört, gar nicht aufgeführt wird. Abbildung 14 zeigt die Verteilung über das Gesamtkorpus, wenn ein klassischeres Set von semantischen Rollen angesetzt worden wäre (vgl. Sowa 2000: 508):

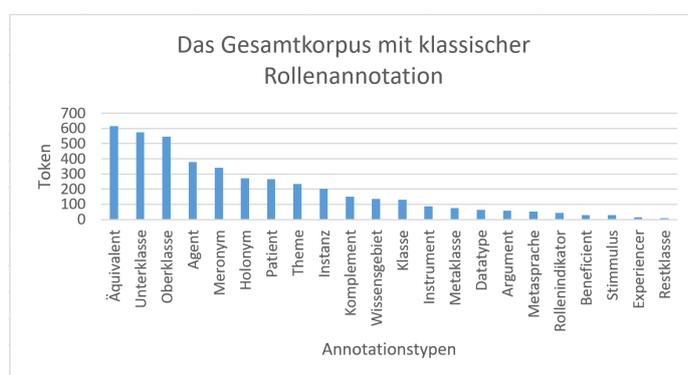


Abbildung 14: Das Gesamtkorpus mit klassischer Rollenannotation

Für diese Grafik wurden die Rollen *Agent*, *Force*, *Depicting*, *Sender*, *Producer*, *Process*, *Possessor* und *Role* unter *Agent* zusammengefasst; die Rollen *Depicted*, *Medium*, *Content* und *Shape* unter *Theme*; die Rollen *Patient*, *Resource*, *Connected* und *Product* unter *Patient*; und schließlich die Rollen *Beneficent* und *Receiver* unter *Beneficent*. Außerdem wurden alle Annotationen der Form No-x unter der entsprechenden Annotation zusammengefasst, also *No-Agent* wurde zu *Agent* gerechnet. Das Ergebnis entspricht den Erwartungen in dem Sinne, dass *Agent* die häufigste semantische Rolle darstellt und

auf dem vierten Platz direkt hinter Äquivalent, Unter- und Oberklasse steht.

Die Motivation, Tabelle 14 zu erstellen, bestand darin, einen Anknüpfungspunkt zu klassischeren Darstellungen zu schaffen. Die Arbeit mit dem Korpus hat hingegen schnell gezeigt, dass ein solches Set an semantischen Rollen zu unterspezifiziert für die Fragestellungen dieser Arbeit ist. Viele der Unterscheidungen, die im Annotationsprozess aufgrund der in Kapitel 6.3 besprochenen Kriterien erstellt wurden, sind auch ohne einen Blick ins Korpus gut begründbar. Als Beispiel diene die in der Literatur eher selten erwähnte Rolle *Force*. Sie wurde in Abbildung 14 unter *Agent* gefasst, aber ob bspw. ein Waldarbeiter bewusst einen Baum fällt oder ob der Wind diesen zum Fallen bringt, sind völlig unterschiedliche Vorgänge. Auch die Annotation *Process* musste unter *Agent* gefasst werden. Hierbei handelt es sich meist um substantivierte Verben. Dass es sich hierbei nicht um willentlich handelnde Entitäten handelt, ist offensichtlich.

Bei den bisherigen Annotationstypen war immer deutlich, in welcher Relation sie stehen: Zum Beispiel stehen Unter- und Oberklasse in der taxonomischen Relation und Meronym und Holonym stehen in der Teil-Ganzes Relation. Bei den semantischen Rollen ist dies anders. *Agent* kann in einer Relation mit *Patient*, mit *Beneficient* oder mit beiden stehen. Die verschiedenen Kombinationen, die im Korpus vertreten waren, werden im Folgenden vorgestellt und besprochen, wobei pro Unterkapitel immer die agentivischste Rolle, bzw. die Rolle, auf die sich die Prädikation bezieht, im Fokus steht. Eine Ausnahme stellt hier *Place* dar. Diese Rolle ist nicht agentivisch, erhält aber trotzdem ein eigenes Unterkapitel, da sie einige Eigenheiten aufweist, die gesondert erläutert werden sollen.

8.1 Agent

Das erste Unterkapitel soll sich dem Annotationstyp *Agent* widmen. Die entsprechende Rolle ist im Kontext dieser Arbeit von besonderer Relevanz: An verschiedenen Stellen war schon die Sprache von agentivischen Aspekten,

worunter diverse Formen von willentlichem Handeln verstanden wurden. Agentivisch wurde dabei als eine Skala begriffen: Ein Argument in einem Satz kann mehr oder weniger agentivisch sein. Der Annotationstyp *Agent* soll die Argumente bezeichnen, die am äußersten Ende dieser Skala stehen, die sozusagen maximal agentivisch sind, bei denen das willentliche Handeln im Fokus steht. Dieser Annotationstyp wurde angesetzt, wenn keine andere agentivische Rolle als Annotation in Frage kam. Bei anderen Rollen steht das willentliche Handeln nicht allein im Fokus; bei *Producer* steht bspw. neben dem willentlichen Handeln die Tatsache, dass bei dem Handeln etwas entsteht, das *Product*, im Fokus.



Abbildung 15: Die Verteilung von *Agent* über die Teilkorpora

Abbildung 15 zeigt, wie die Rolle über die verschiedenen Teilkorpora verteilt ist. Sie ist mit fast vierzig Token am häufigsten in Wirtschaft und Technik vertreten und damit doppelt so häufig wie in den Geisteswissenschaften, wo sie am zweithäufigsten vertreten ist. In Form und Struktur, Kosmos und Energie sowie Biologie ist sie gar nicht vertreten. Diese Umstände schlagen sich in einer Standardabweichung von knapp über fünfzehn Prozent nieder. Die naheliegende Erklärung ist, dass in den naturwissenschaftlichen Wissensgebieten wenig über willentlich handelnde Akteure geschrieben steht. Dafür müsste in diesen Gebieten der Annotationstyp *Force* häufig auftreten.

Als erstes soll *Agent* in Kombination mit *Patient* betrachtet werden. Hierbei handelt es sich um das allgemeine Verständnis von Subjekt und Objekt: Es

existiert ein handelndes Individuum und etwas, worauf die Handlung abzielt. Ein solcher Fall liegt in (110) vor, wobei das Subjekt hier anaphorisch durch das Relativpronomen realisiert wird.

(110) *Absender*: Männliche Person, die *etwas* abschickt. (Der Duden, Das Bedeutungswörterbuch: 61)

Die in Frage stehenden Textteile werden im Folgenden immer hervorgehoben. Diese Hervorhebung entspricht im Allgemeinen auch den tatsächlich getätigten Annotationen. Anaphorische Ausdrücke wie Pronomen, wurden nicht annotiert, wenn sie im Text aufgelöst werden können; dann wurden die Einheiten, auf die sich der jeweilige Ausdruck bezieht, annotiert.

Eine Formalisierung könnte folgendermaßen aussehen:

$Absender \sqsubseteq (MaennlichePerson \sqcap \exists abschicken.Objekt)$

An dieser Formalisierung müssen zwei Dinge erklärt werden: Erstens wurde das Indefinitpronomen *etwas* mit *Objekt* übersetzt. Diese Entscheidung liegt darin begründet, dass *Objekt* in den meisten Ontologien die allgemeinste Klasse – d.h. am höchsten in der Taxonomie einer Ontologie verortet – für konkrete Entitäten ist. Diese Entscheidung muss also von Ontologie zu Ontologie evtl. unterschiedlich getroffen werden. Zweitens wurde die ontologische Rolle *abschicken* angesetzt. Ein Absender ist also die Untermenge der Schnittmenge von männlichen Personen (laut Duden!) und der Klasse von Dingen, die mindestens einmal über die Relation *abschicken* mit Objekt verbunden sind. Die Annotation *Agent* wurde also in dem Sinne verarbeitet, dass das Verb des Textes in eine Rolle übersetzt wurde, dessen Objekt der *Patient* des Satzes ist.

Dies ist nur eine Möglichkeit, *Agent* zu formalisieren. Sie hat den Vorteil, dass sie sehr nah am eigentlichen Text ist, aber evtl. den Nachteil, dass auf diese Weise jedes Verb, das mit einem *Agent* steht, in eine ontologische Rolle überführt wird – unter Umständen könnte ein generalisierender Umgang mit Verben sinnvoller sein. Ein maximal generalisierender Umgang würde für einen *Agent*, der mit einem *Patient* steht, die Rolle *actsOn* ansetzen.

$Absender \sqsubseteq (MaennlichePerson \sqcap \exists actsOn. Objekt)$

Diese Formalisierungsstrategie würde nicht zwischen Verben wie bspw. *verschicken* und *verbrennen* differenzieren, was für die meisten Ontologien sehr wahrscheinlich zu allgemein wäre. Ein Mittelweg zwischen beiden Strategien, der mit einer vorgefertigten Verbontologie fungiert, die nur bestimmte als wichtig erachtete Verben enthält, auf welche die Verbtoken im jeweiligen Text abgebildet werden, wäre ebenfalls denkbar.

In beiden Formalisierungsstrategien geht es darum, eine ontologische Rolle zu finden, die eine Relation zwischen zwei Entitäten ausdrückt. Die erste Strategie ist dabei sehr spezifisch, die zweite sehr allgemein. Wie die Rollen dabei benannt werden, ist formal letztlich irrelevant, solange sie wohlunterschieden sind. Im Folgenden werden die ontologischen Rollen auf Englisch benannt, u.a. um sie von den konkreten Einheiten in den Texten zu unterscheiden. Davon wird nur abgewichen, wenn ein konkretes Verb als ontologische Rolle dargestellt werden soll. Diese Strategie bietet sich aber ohnehin nur für bestimmte semantische Rollen an, wie sich im Folgenden noch zeigen wird.

Die Aussage, dass ein *Agent* immer eine willentlich handelnde Entität sei, legt den Schluss nahe, dass es sich immer um Menschen oder zumindest um Lebewesen handelt. Unter anderem vor dem Hintergrund der KI-Forschung der letzten Jahrzehnte ist diese These allerdings nicht mehr haltbar. Letztlich handelt es sich dabei aber um eine philosophische Diskussion und nicht eine sprachwissenschaftliche Diskussion (vgl. Libet 2004 und Morsella 2008). Im Korpus wurden teilweise auch Einheiten mit *Agent* annotiert, die keine Lebewesen denotieren, wie in (111):

(111) *Kognition ist die von einem verhaltenssteuernden System ausgeführte Umgestaltung von Informationen.* (Wikipedia, Kognition)

Dieser Text ist ein bisschen anders konstruiert als (110) und könnte folgendermaßen formalisiert werden:

$Kognition \sqsubseteq (Umgestaltung \sqcap \exists gestaltetUm.Information \sqcap \exists ausgefuehrtVon.VerhaltenssteuerndesSystem)$

Der Unterschied liegt darin, dass der *Agent* hier als PP realisiert wird, die als Attribut an einem Partizip hängt, und der *Patient* ebenfalls eine PP ist, die als Attribut an einem Substantiv hängt. Das Subjekt des Satzes ist *Kognition*. Das abgeleitete Substantiv *Umgestaltung* wird in zwei Rollen überführt: Eine repräsentiert den Aktiv von *umgestalten* und führt daher als Objekt den *Patient*, und die andere Rolle ist als Oberklasse von *Kognition* formalisiert. *Kognition* ist also die Unterklasse der Schnittmenge von Dingen, die Informationen umgestalten, und Dingen, die durch verhaltenssteuernde Systeme ausgeführt werden.

Auch hier sei zum Vergleich nochmal eine allgemeinere Formalisierung vorgestellt:

$$Kognition \sqsubseteq (Umgestaltung \sqcap \exists actsOn.Information \sqcap \exists actedBy.VerhaltenssteuerndesSystem)$$

Kognition ist also eine Form der *Umgestaltung*, bei der auf *Information* gewirkt wird und bei der durch ein verhaltenssteuerndes System gewirkt wird. Die ontologische Rolle *actedBy* stellt das passivische Pendant zu *actsOn* dar.

Das folgende Beispiel demonstriert, wie viel Information in einem natürlich-sprachlichen Text stecken kann:

- (112) *Republik: Staatsform, bei der die oberste Gewalt durch Personen ausgeübt wird, die für eine bestimmte Zeit vom Volk oder dessen Vertretern gewählt werden.* (Der Duden, Das Bedeutungswörterbuch: 788)

Zunächst ist *Personen* sowohl *Agent* als auch *Patient*. Diese Rollenkumulation kommt durch einen Relativsatz zustande; im Text sind die beiden Rollen auf das Substantiv und das Relativpronomen verteilt. Zusätzlich sind diese Personen Teil der *Republik*, was ebenfalls durch einen Relativsatz ausgedrückt wird:

Die Herausforderung für eine Formalisierung besteht nun darin, zu explizieren, dass dieselbe Klasse das Meronym ist und diese beiden Rollen vertritt. Dies kann nur dadurch erreicht werden, dass das Konzept einer gewählten Person an einer anderen Stelle in der Ontologie definiert wird, um dann in der

Formel zu Republik verwendet zu werden:

$$\text{Republik} \sqsubseteq (\text{Staatsform} \sqcap \exists \text{consistsOf.GewahltePerson})$$

Bei den folgenden Beispielen steht ein *Agent* in Relation mit einem *Beneficiant*.

- (113) **Korrespondent**: *Person, die für eine Zeitung, einen Fernsehsender o.Ä. aus einer bestimmten Stadt, einem bestimmten Land Bericht erstattet.* (Der Duden, Das Bedeutungswörterbuch: 592)

Auch für diese Kombination kann sich die Formalisierung direkt am Prädikat orientieren.

$$\text{Korrespondent} \sqsubseteq (\text{Person} \sqcap (\exists \text{erstattetBerichtFor.Zeitung} \sqcup \exists \text{erstattetBerichtFor.Fernsehsender}))$$

Aus dem Verb wird eine Rolle abgeleitet, wobei durch das Suffix *For* angezeigt wird, dass es sich bei dem Objekt der Rolle um einen *Beneficiant* handelt. (113) ist ein gutes Beispiel dafür, dass zumindest eine leichte Generalisierung des Prädikats sinnvoll sein kann, da das Funktionsverbgefüge *Bericht erstatten* gleichbedeutend zu *berichten* ist. Es wäre gewünscht, dass beide Varianten in der Ontologie nur durch eine Rolle vertreten würden, da Redundanzen zu vermeiden sind.

Im folgenden Beispiel steht der *Agent*, die öffentliche Verwaltung, in einer Relation mit einem *Instrument*, dem Verwaltungsakt:

- (114) **Der Verwaltungsakt**, abgekürzt VA, stellt im deutschen Verwaltungsrecht eine Handlungsform **der öffentlichen Verwaltung** dar. (Wikipedia, Verwaltungsakt)

Für das Verhältnis von *Agent* und *Instrument* lässt sich eine komplett neue Rolle ansetzen:

$$\text{Verwaltungsakt} \sqsubseteq (\text{Handlungsform} \sqcap \exists \text{usedBy.OeffentlicheVerwaltung})$$

Ein Verwaltungsakt ist eine Handlungsform, die von der öffentlichen Verwaltung (als Instrument) genutzt wird. Diese Rolle ist in dem Sinne passivisch, als sie die handelnde Entität, in diesem Fall die öffentliche Verwaltung, zum Objekt macht. Eine aktivische Variante, *uses*, die letztlich das Inverse die-

ser Relation ist, kann ebenfalls angesetzt werden. Diese würde für folgendes Beispiel (115) gebraucht werden, in dem das *definiendum* der *Agent* ist:

- (115) **Ein Korrepetitor** (auch *Repetitor* oder *Solorepetitor*) spielt **am Klavier** statt des Orchesters, wenn Sänger, Chöre, Instrumentalisten, Tänzer oder Schauspieler ein Stück neu lernen oder wiederholen, ihre Rollen einstudieren oder Szenen geprobt werden (Oper, Operette, Ballett usw.), und gibt darüber hinaus in Einzelproben korrigierende musikalische Hinweise. (Wikipedia, Korrepetitor)

Die folgende Formalisierung nutzt die aktive Rolle:

$Korrepetitor \sqsubseteq (\exists uses.Klavier)$

Diese Formalisierung nennt keine Oberklasse im Sinne von Person oder Mensch, da der Text dies auch nicht tut. Es könnte aber diskutiert werden, ob eine solche Klasse nicht trotzdem ergänzt werden sollte oder inwieweit sie durch Schlussfolgerungen ermittelt werden könnte.

Besonders im Wissensbereich Wirtschaft steht ein *Agent* häufig in Relation mit einem *Goal*:

- (116) a. **Buchhaltung**: Abteilung eines Betriebes, die für die **Buchführung** verantwortlich ist. (Der Duden, Das Bedeutungswörterbuch: 249)
b. **Ein Unternehmen** ist eine wirtschaftlich selbstständige Organisationseinheit, die mit Hilfe von Planungs- und Entscheidungsinstrumenten Markt- und Kapitalrisiken eingeht und sich zur **Verfolgung des Unternehmenszweckes und der Unternehmensziele** eines oder mehrerer Betriebe bedient. (Wikipedia, Unternehmen)

Eine Formalisierung erfolgt wieder über eine speziell definierte Rolle:

$Buchhaltung \sqsubseteq (Abteilung \sqcap \exists hasGoal.Buchfuehrung)$

Auch für diese Relation lässt sich ein inverses Gegenstück anlegen: *goalOf*.

In dem folgenden Beispiel wird dieses gebraucht:

- (117) **Der Betriebszweck** ist das dauerhaft verfolgte **Arbeits- und Produktionsziel** eines Betriebes. (Wikipedia, Betriebszweck)

Die folgende Formalisierung konzentriert sich nur auf diese Relation:

$Betriebszweck \sqsubseteq (\exists goalOf.Betrieb)$

Für das Verhältnis zwischen *Agent* und *Role* kann eine Relation, *performsAs*, und ihr Inverses, *performedBy*, angesetzt werden. Der folgende Beleg stammt aus dem Bereich Wirtschaft und Technik:

- (118) **Die Kommunalverwaltung** ist der Teil der öffentlichen Verwaltung, der als Stadt-, Gemeinde- und Kreisverwaltung sowie als Zweckverband in ihrer Eigenschaft als **Behörde bzw. Verwaltungsträger** die kommunale Hoheitsgewalt als grundrechtsähnliches Recht ausübt. (Wikipedia, Kommunalverwaltung)

Anbei die entsprechende Formalisierung:

$$\text{Kommunalverwaltung} \sqsubseteq (\exists \text{isPartOf.OeffentlicheVerwaltung} \sqcap \exists \text{performsAs.Behoerde})$$

Role ist eine ungewöhnliche Rolle, die nur zusammen mit *Agent* auftritt. Es wurde überlegt, ob sich in solchen Fällen nicht ein doppelter *Agent* ansetzen lässt (Kommunalverwaltung und Behörde) oder ob es sich nicht letztlich um die taxonomische Relation handelt (Kommunalverwaltung als Unterklasse von Behörde), aber letztlich wurde *Role* angesetzt, um auszudrücken, dass es sich hierbei um ein ganz bestimmtes Verhältnis von einem *Agent* zu einem Konzept handelt, aus dem heraus dieser agiert.

Zuletzt sollen die Fälle betrachtet werden, bei denen *Agent* alleine eine Relation bildet. Da in diesem Fall *Agent* eine einstellige Relation darstellt, kann sie nicht mit dem Konzept der ontologischen Rolle formalisiert werden. Unterschiedliche Strategien sind hier möglich.

Die naheliegendste Möglichkeit formalisiert die Verben als einfache Klassen. Das Verb *schwimmen* bezeichnet bspw. die Klasse der Entitäten, die schwimmen. Im folgenden Beispiel liegt ein Funktionsverbgefüge vor, weshalb man versucht sein könnte, *Regie* als *Patient* zu annotieren. Letztlich liegt hier aber nur ein *Agent* vor:

- (119) **Regisseur**: Jmd., der [berufsmäßig] Regie führt, die Regie hat. (Der Duden, Das Fremdwörterbuch: 963)

Eine Formalisierung, die das Verb als Klasse verwendet, sähe folgendermaßen aus:

Regisseur \sqsubseteq (*Person* \sqcap *RegieFuehrend*)

Das folgende Beispiel versucht den Begriff *Tätigkeit* zu definieren und nutzt dabei den Begriff *Mensch* als *Agent*.

(120) *Tätigkeit (Aktivität) bezeichnet ein Handeln des Menschen und kann sowohl körperliche als auch geistige Verrichtungen umfassen.* (Wikipedia, Tätigkeit)

Die Definition ist zirkulär und daher auch schwer zu formalisieren. Folgende Formalisierung wäre möglich:

Taetigkeit \sqsubseteq (\exists *handeln.Mensch*)

Eine Tätigkeit ist eine Unterklasse der Dinge, bei der der Mensch handelt. *Handeln* steht hier in einer ganz allgemeinen Bedeutung, sodass sich im Vokabular der semantischen Rollen auch formulieren ließe, dass Tätigkeit eine Unterklasse der Dinge sei, bei denen der Mensch ein *Agent* ist.

8.2 Force

Als *Force* werden Entitäten annotiert, von denen ein Vorgang ausgeht, die dabei aber nicht willentlich handeln. Es ist nicht davon auszugehen, dass diese Rolle im Deutschen grammatisch anders realisiert wird als *Agent*; die Unterscheidung wird daher oft lexikalisch getroffen werden.

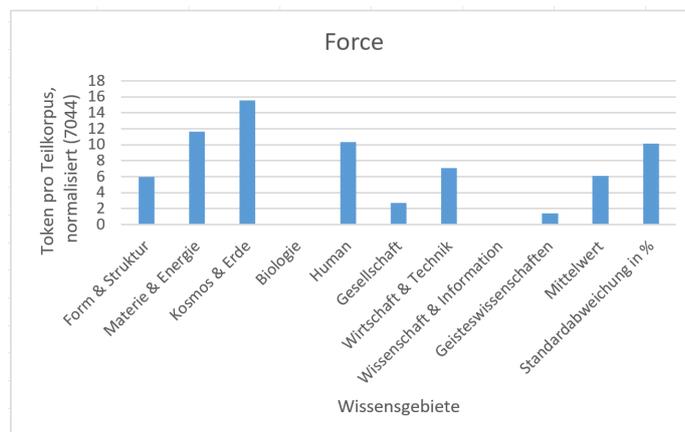


Abbildung 16: Die Verteilung von *Force* über die Teilkorpora

Abbildung 16 zeigt im Vergleich mit Abbildung 15, dass diese Rolle über die Teilkorpora nahezu komplementär zur Rolle *Agent* verteilt ist: Sie wurde am

häufigsten in den Naturwissenschaften angesetzt, wo *Agent* nahezu gar nicht angesetzt wurde, und ist bei den Gesellschafts- und Geisteswissenschaften weniger vertreten. Auffällig ist allerdings, dass sie keinmal im Teilkorpus zur Biologie oder zu Wissenschaft und Information angesetzt wurde. Die Standardabweichung ist mit zehn Prozent geringer als die von *Agent*.

Force steht ebenso wie *Agent* häufig in Relation mit einem *Patient*. Dabei wird typischerweise *Force* als Subjekt und *Patient* als Akkusativobjekt realisiert (121a); es gibt aber diverse Alternativen (121b-121d).

- (121)
- a. **Ein Geschäftsvorfall** (oder Geschäftsfall) ist in der Buchführung eine Transaktion, die **den Jahresabschluss oder den Haushalt von Wirtschaftssubjekten beeinflusst**. (Wikipedia, Geschäftsvorfall)
 - b. **Meteorit: In die Erdatmosphäre** eindringender kosmischer Körper. (Der Duden, Das Fremdwörterbuch: 714)
 - c. **Eine Platte** ist in der Technischen Mechanik bzw. in der Bautechnik ein in der Ebene ausgebreitetes Bauteil, das aus steifem Material besteht (ebenes Flächentragwerk) und **durch senkrecht auf sie wirkende Kräfte und durch Momente um Achsen**, die in der Plattenebene liegen, belastet wird. (Wikipedia, Platte)
 - d. **Weckamin: Der körperlich-geistigen** **Abspannung** entgegenwirkendes, stimulierendes Kreislaufmittel. (Der Duden, Das Fremdwörterbuch: 1176)

Eine mögliche Formalisierung soll an Beleg (121a) gezeigt werden:

Geschäftsvorfall \sqsubseteq (*Transaktion* \sqcap \exists beeinflusst.*Jahresabschluss*)

Die Relation zwischen *Force* und *Patient* wird also auf die gleiche Weise vorgenommen, wie die zwischen *Agent* und *Patient* – das Verb wurde als Relation interpretiert und zusammen mit dem *Patient* in eine Rolle überführt. Die Beispiele werfen die Frage auf, ob nicht doch *Agent* und *Force* zusammengeführt werden sollten. Unter bestimmten Umständen mag dies vertretbar sein. Es dürfte aber auch kaum angezweifelt werden, dass ein Unterschied zwischen Vorgängen besteht, die aufgrund eines Entscheidungsprozesses erfolgen, und solchen, bei denen dies nicht der Fall ist. Der Unterschied zwischen beiden Rollen hat auch Konsequenzen für eine Ontologie: In der Taxonomie kann eine Klasse existieren, die potentiell handelnde Entitäten bzw. im weiteren

Sinne komplexe, entscheidungstreffende Systeme darstellt. Entitäten, die als *Agent* in Frage kommen, müssen Unterklasse dieser einen Klasse sein. Dies gilt natürlich nur, wenn diese Unterscheidung für die jeweilige Ontologie von Bedeutung ist.

Dass *Force* eine andere Rolle als *Agent* ist, zeigt sich aber auch daran, dass sie mit anderen Rollen korreliert. Im folgenden Beispiel steht sie in Relation mit *Medium*:

(122) ***Körperschall*** ist *Schall*, der sich **in einem Festkörper** ausbreitet. (Wikipedia, Körperschall)

Dieses Verhältnis lässt sich erneut am besten mit einer eigenen Relation in der Ontologie ausdrücken:

Koerperschall \sqsubseteq (*Schall* \sqcap \exists *movesIn.Festkoerper*)

Die Relation *movesIn* drückt also explizit das Verhältnis zwischen einer Kraft und einem Medium aus, wobei sich die eine in der anderen ausbreitet. Ein ähnliches Verhältnis zwischen *Agent* und *Medium* existiert nicht, was einen tiefen ontologischen Grund zu haben scheint: Ein willentlich handelnder Akteur ist nicht in dem Sinne an ein Medium gebunden, wie dies eine physikalische Kraft ist.

Das folgende Beispiel ist hervorzuheben, da es einerseits typisch für den Annotationstypus *Force* ist, aber bzgl. der Annotation von *über die Gravitation* nicht ganz eindeutig ist:

(123) ***Dunkle Materie*** ist eine postulierte Form von *Materie*, die nicht direkt sichtbar ist, aber **über die Gravitation** wechselwirkt. (Wikipedia, Dunkle Materie)

Der kausale Zusammenhang in diesem Beispiel ist typisch für Vorgänge, die nicht aus willentlichem Handeln entstehen: Dunkle Materie beeinflusst die Gravitation, die dann weiterwirkt. Dies wirft aber die Frage auf, wie *über die Gravitation* zu bewerten ist. Einerseits könnte diese Wortgruppe auch als *Force* annotiert werden, wobei das Objekt des Wirkens nicht weiter genannt wird. Sie könnte auch als *Patient* annotiert werden, da auf sie durch die Dunkle

Materie eingewirkt wird. Die Formulierung selber lässt zunächst an die Annotation *Instrument* denken; hier könnte aber eingewendet werden, dass *Instrument* nur mit *Agent* stehen kann, da das Verwenden eines Instruments ein willentliches Handeln voraussetzt. Gravitation in diesem Fall als *Patient* zu betrachten, ist m.E. am sinnvollsten, bildet aber nur einen Teil der im Text genannten Informationen ab:

DunkleMaterie \sqsubseteq (*Materie* \sqcap \exists *wechselWirkt.Gravitation*)

Force wurde auch als Annotationstyp für viele Subjekte im Bereich Form und Struktur gebraucht. Letztlich beruhte diese Entscheidung einzig auf der Einsicht, dass es sich in diesen Fällen nicht um *Agent* handeln kann.

- (124)
- a. *In der Mathematik ordnet **die Betragsfunktion** einer reellen oder komplexen Zahl ihren Abstand zur Null zu. Dieser sogenannte absolute Betrag, Absolutbetrag, Absolutwert oder auch schlicht Betrag ist immer eine nichtnegative reelle Zahl.* (Wikipedia, Betragsfunktion)
 - b. ***Ein Polynom** summiert die Vielfachen von Potenzen einer Variablen bzw. Unbestimmten.* (Wikipedia, Polynom)
 - c. ***Funktion:** Veränderliche Größe, die in ihrem Wert von einer anderen abhängig ist.* (Duden, Das Fremdwörterbuch: 392)
 - d. *In der Mathematik ist **eine Funktion** (lateinisch *functio*) oder Abbildung eine Beziehung (Relation) zwischen zwei Mengen, die jedem Element der einen Menge (Funktionsargument, unabhängige Variable, *x*, *x*-Wert) genau ein Element der anderen Menge (Funktionswert, abhängige Variable, *y*, *y*-Wert) zuordnet.* (Wikipedia, Funktion)
 - e. ***Als Ortsvektor** (auch Radiusvektor oder Positionsvektor) eines Punktes bezeichnet man in der Mathematik und in der Physik einen Vektor, der von einem festen Bezugspunkt zu diesem Punkt (Ort) zeigt.* (Wikipedia, Ortsvektor)

Form und Struktur stellt in dem Sinne einen besonderen Bereich dar, da dort spezielle Definitionen vorliegen, die nur behelfsmäßig in natürlicher Sprache ausgedrückt werden können. Dass bspw. eine Betragsfunktion einer Zahl ihren Abstand zu Null *zuordnet*, ist letztlich eine Metapher. Nicht umsonst hat die Mathematik ein eigenes System aus Regeln und Zeichen entwickelt, das auch als Sprache bezeichnet werden kann. Sollte auf der Basis des hier vorgestellten Annotationssystems eine konkrete Ontologie erstellt werden, dann

dürfen Texte aus diesem Bereich nur ein Randphänomen darstellen und nicht den Fokus. Eine explizit mathematische Ontologie sollte sich an den exakten Definitionen dieser Wissenschaft orientieren.

Der folgende Text ist besonders im Vergleich mit (121d) bemerkenswert. Dort wurden Weckamine als *Force* annotiert, da sie wirken. In (125) wird letztlich dieselbe Aussage getätigt, aber die Formulierung, eine PP die vom Verb *bezeichnen* abhängt, erinnert an Homonymie.

(125) *Als **Weckamine** werden bisweilen unscharf Amine mit stimulierender (aufweckender) Wirkung bezeichnet. Bekannte Vertreter dieser Gruppe sind das Amphetamin, Methamphetamin und Ephedrin.* (Wikipedia, Weckamin)

Dies ist ein Beispiel dafür, dass bestimmte sprachliche Muster zwar typisch für bestimmte semantische Zusammenhänge sind, bspw. PPn mit *mit* für Homonymie, aber auch für andere semantische Zusammenhänge verwendet werden können.

Die folgenden Belege sind im Passiv verfasst.

- (126) a. ***Das Völkerrecht** (Lehnübersetzung zu lateinisch *ius gentium*, 'Recht der Völker') ist eine überstaatliche, aus Prinzipien und Regeln bestehende Rechtsordnung, durch die **die Beziehungen zwischen den Völkerrechtssubjekten** (meist Staaten) auf der Grundlage der Gleichrangigkeit geregelt werden.* (Wikipedia, Völkerrecht)
- b. *Eine konstitutionelle Monarchie ist als Variante der Monarchie eine Staatsform, in der **die Macht des Monarchen** durch **eine Verfassung** geregelt und beschränkt wird.* (Wikipedia, Konstitutionelle Monarchie)
- c. ***Eine substanzinduzierte Psychose** (auch bekannt als Drogenpsychose oder drogeninduzierte Psychose) ist eine psychotische Störung, die mit oder ohne ärztlicher Verordnung von **einer oder mehreren psychotropen Substanzen** (wie z. B. Alkohol, Cannabinoide, Kokain, Amphetamine, Halluzinogene, neue psychoaktive Substanzen etc.) ausgelöst wurde.* (Wikipedia, Substanzinduzierte Psychose)

Es können erneut ontologische Rollen angesetzt und genutzt werden, die den passivischen Verhältnissen gerecht werden. Nachfolgend ist bspw. eine Formalisierung von (126a), bei der die Beziehung zwischen einem *Instrument* und einem *Goal* mit der ontologischen Relation *usedFor* übersetzt wird.

Voelkerrecht \sqsubseteq (*Rechtsordnung* \sqcap
 \exists *consistsOf.Prinzip* \sqcap \exists *usedFor.BeziehungZwischenVoelkerrechtssubjekten*)

Bei (127a) sind die Verhältnisse undurchsichtiger:

- (127) a. *Der Begriff der Gattung oder Textgattung ordnet literarische Werke in inhaltlich oder formal bestimmte Gruppen.* (Wikipedia, Gattung)
b. *Der Begriff der Gattung oder Textgattung wird genutzt, um literarische Werke in inhaltlich oder formal bestimmte Gruppen einzuordnen.*

Die Formulierung legt zunächst nahe, dass *Gattung* als *Force* und *literarische Werke* als *Patient* annotiert werden sollten. *Gattung* zählt aber nicht zu Begriffen, die genug agentivisches Potenzial haben, um die Rolle *Force* zu vertreten. Es ist sinnvoller, *Gattung* als Instrument zu verstehen, wie sich auch an der Paraphrasierung (127b) erkennen lässt:

Diese Interpretation verlangt zur Formalisierung eine ontologische Rolle für das Verhältnis von *Instrument* und *Patient*. Da dies außerdem ein Beispiel ist, in dem ein *Agent* – derjenige, der einordnet – nicht genannt wird, liegt es nahe, die Strategie zur Formalisierung von *Agent* zu verfolgen und das Verb in eine ontologische Rolle zu überführen und mit *usedFor* zur Kennzeichnung von *Instrument* auszuzeichnen.

Gattung \sqsubseteq \exists *usedForOrdnen.LiterarischesWerk*

In den folgenden beiden Belegen werden Dinge dadurch definiert, dass nicht auf sie eingewirkt werden kann. Es existiert demnach eine Kraft, aber sie wirkt auf den *Patient* nicht ein – es handelt sich demnach nur gedacht um einen *Patient* bzw. *No-Patient*.

- (128) a. *Ein Körper, der sich im mechanischen Gleichgewicht befindet, erfährt keine Beschleunigung; er verharrt folglich in Ruhe oder bewegt sich mit konstanter Geschwindigkeit.* (Wikipedia, Mechanisches Gleichgewicht)
b. *CD-ROM: Speicherplatte [für PCs mit entsprechendem Laufwerk], deren Programminhalt abgerufen, aber nicht verändert werden kann.* (Der Duden, Das Fremdwörterbuch: 199)

Die folgende Formalisierung greift dementsprechend das Verb auf, um es in

eine Rolle zu überführen, wie es als Formalisierungsstrategie für die Relation zwischen *Force* und *Patient* bereits vorgestellt wurde, verneint aber das zweite Prädikat.

$$CD\text{-}Rom \sqsubseteq (Speicherplatte \sqcap \exists ruftab.Programminhalt \\ \sqcap \neg \exists veraendert.Programminhalt)$$

Bei (128a) findet sich indes mit *mit konstanter Geschwindigkeit* ein Beleg für den Annotationstyp *Manner*. *Manner* ist eine Rolle, die die Art und Weise, in der eine Handlung ausgeführt wird oder ein Geschehen sich abspielt, bezeichnet. Eine Formalisierung könnte folgendermaßen aussehen:

$$KoerperImMechanischenGleichgewicht \sqsubseteq \\ (\exists MannerOfAction.KonstanteGeschwindigkeit)$$

Ein Körper im mechanischen Gleichgewicht ist eine Unterklasse der Dinge, die mit konstanter Geschwindigkeit handeln. Diese Formalisierung ist unspezifischer als der eigentliche Text, da sie nicht spezifiziert, welche Handlung mit konstanter Geschwindigkeit ausgeführt wird; sollte dies von Bedeutung sein, bestände eine zweite Möglichkeit der Formalisierung darin, das entsprechende Verb mit einem Suffix, bspw. *WithManner*, zu versehen:

$$KoerperImMechanischenGleichgewicht \sqsubseteq \\ (\exists bewegenWithManner.KonstanteGeschwindigkeit)$$

Welche Variante zu wählen ist, lässt sich nur von Ontologie zu Ontologie entscheiden.

8.3 Producer

Producer definiert sich darüber, dass das zweite Element in dieser Relation, das *Product*, durch die Handlung erst entsteht. Prinzipiell wäre es möglich, die Unterscheidung von *Agent* und *Force* auch für *Producer* anzusetzen; also zwischen einem willentlichen und nicht willentlichen *Producer* zu unterscheiden. Dies erschien für die folgenden Ausführungen allerdings zu ambitioniert.

Die Annotation ist für die semantischen Rollen relativ gleichmäßig verteilt, was sich an einer Standardabweichung von knapp unter zehn Prozent ablesen

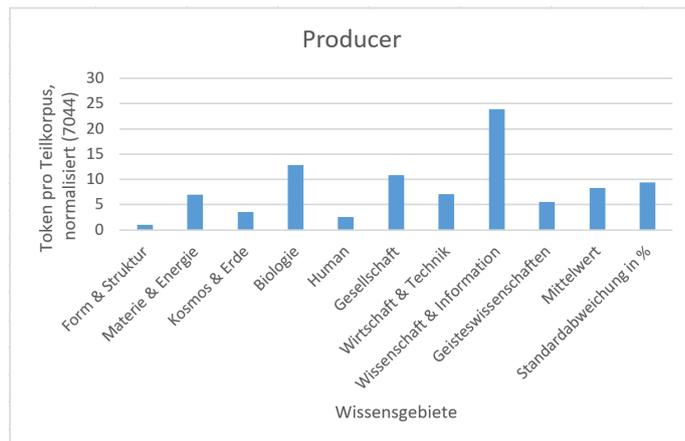


Abbildung 17: Die Verteilung von *Producer* über die Teilkorpora

lässt.

Der Prototyp dieser semantischen Rolle sind Unternehmen, da es explizit deren Aufgabe ist, etwas zu produzieren. Unternehmen sind dabei, obwohl sie Kollektive sind, in einem hohen Grad agentivisch; sie bzw. ihr Management treffen bewusst Entscheidungen, oft nach einem längeren Entscheidungsprozess.

- (129) *Tesla, Inc.* (bis 1. Februar 2017 *Tesla Motors*) ist ein US-amerikanisches Unternehmen, das **Elektroautos** sowie **Stromspeicher und Photovoltaikanlagen** produziert und vertreibt. (Wikipedia, Tesla Inc.)

Es ist naheliegend, diese semantische Rolle wieder durch eine definierte ontologische Relation auszudrücken, die sowohl aktivisch als auch passivisch vorliegt.

$\exists \textit{produces.Elektroauto}(\textit{Tesla})$

Die Formalisierung mag auf den ersten Blick seltsam wirken, da eine solche bisher noch nicht vorkam. Sie ist der Tatsache geschuldet, dass Tesla ein Individuum und keine Klasse ist. Sie sagt also aus, dass Tesla eine Instanz – keine Unterklasse! – der Dinge ist, die Elektroautos produzieren.

Im Korpus lassen sich relativ viele Individuen finden, die als *Producer* annotiert wurden. Sie stammen aus den Teilkorpora Wirtschaft und Information sowie Wissenschaft und Technik. Die folgenden Belege zeigen, dass *Producer* an bestimmten Verben identifiziert werden kann oder auch nur durch Weltwis-

sen. In (130a) wird diese Rolle durch das Partizip *herausgegeben* explizit gemacht. In (130b) hingegen ist die Rolle nur als Genitivattribut realisiert; dass es sich in diesem Fall um *Producer* und nicht bspw. *Possessor* handelt, kann nur geschlussfolgert werden, wenn man weiß, dass *Illustrierte* herausgegeben werden.

- (130) a. *Gala* ist eine wöchentlich von **Gruner + Jahr** herausgegebene *Illustrierte*. (Wikipedia, Gala)
 b. *Die Welt* [...] ist eine überregionale deutsche Tageszeitung **der Axel Springer SE**. (Wikipedia, Die Welt)

In beiden Fällen stellt das *Product* das *definiendum* dar, weswegen die Rolle in ihrer passivischen Form zur Formalisierung angesetzt wird:

$$\exists \text{producedBy.Gruner} + \text{Jahr}(\text{Gala})$$

Auch Personen können als *Producer* annotiert werden:

- (131) *Die Eulerschreibweise* ist nach dem Mathematiker **Leonhard Euler** benannt, der diese für die Feinanalyse von Musikstücken in reiner Stimmung einführte. (Wikipedia, Eulersches Tonnetz)

In diesem Fall ist das *definiendum* als Klasse formalisiert, weswegen sich folgende Formalisierung ergibt:

$$\text{Eulerschreibweise} \sqsubseteq \exists \text{producedBy.LeonardEuler}$$

Die folgenden Belege sind Beispiele für die Rolle in der Variante, in der *Producer* nicht willentlich vorliegt:

- (132) a. In der klassischen Mechanik ist **das Gravitationsfeld** (auch *Schwerkraftfeld*) das *Kraftfeld*, das **durch die Gravitation von Massen** hervorgerufen wird. (Wikipedia, Gravitation)
 b. In der Gruppentheorie ist **eine zyklische Gruppe** eine Gruppe, die **von einem einzelnen Element a** erzeugt wird. (Wikipedia, Zyklische Gruppe)

Da es sich hierbei wieder um Klassen handelt, werden sie in gewohnter Weise formalisiert:

$$\text{Gravitationsfeld} \sqsubseteq (\text{Kraftfeld} \sqcap \exists \text{producedBy.GravitationVonMassen})$$

An (132b) ist zu beachten, dass das Verb *erzeugen* ein sehr deutliches Indiz dafür ist, dass *Producer* hier angesetzt werden sollte; letztlich handelt es sich

aber wieder um einen Fall aus dem mathematischen Kontext und *erzeugen* wird hier in einem metaphorischen Sinne verwendet. Das kognitive Muster, dass etwas in einem Prozess entsteht, wird als Blaupause für einen Sachverhalt verwendet, mit dem es eigentlich eine andere Bewandnis hat, für den aber keine natürlich-sprachlichen Ausdrucksmittel bereitstehen.

Neben *Producer* wurde noch die Rolle *Process* angesetzt, die ebenfalls mit *Product* in Relation steht. Bei den Token handelt es sich immer um substantivierte Verben. Dies ist die einzige Rolle, die explizit für Vorgänge angesetzt wurde.

- (133) a. **Signalisierung** (oder auch Zeichengabe) ist ein Überbegriff für die Abgabe von **Zeichen** (Signalen oder Nachrichten) in verschiedenen Bereichen der Kommunikation und Technik. (Wikipedia, Signalisierung)
- b. **Unter Denken** werden alle (psychologischen) Vorgänge zusammengefasst, die aus einer inneren Beschäftigung mit Vorstellungen, Erinnerungen und Begriffen **eine Erkenntnis** zu formen versuchen. (Wikipedia, Denken)

Aus formaler Sicht besteht für eine Ontologie kein relevanter Unterschied zwischen *Producer* und *Process*, weswegen sich die Relation genauso formalisieren lässt wie die zwischen *Producer* und *Product*.

Signalisierung \sqsubseteq (*Abgabe* \sqcap \exists *produces*.*Zeichen*)

Der Grund, *Process* einzuführen, liegt nur darin, dass es sich bei den Argumenten um eine andere Art von Entitäten handelt als bei *Producer*. Weitere ontologische Relationen müssten allerdings eingeführt werden, um das Verhältnis zwischen *Producer* und *Process* und dem entsprechenden Inversen auszudrücken, also zwischen dem, der einen Prozess ausführt, und jenem Prozess. Diese sollen *carriesOut* und *carriedOutBy* genannt werden.

Es kann vorkommen, dass ein Begriff sowohl einen Prozess als auch das Ergebnis dieses Prozesses bezeichnet. Dies ist ein Fall von Metonymie:

- (134) a. **Kreuzung** ist ein Begriff aus der Genetik. Bei Pflanzen und Tieren wird einerseits die **geschlechtliche Fortpflanzung** zwischen zwei genetisch verschiedenen, aber relativ nahe verwandten Arten, Unterarten oder Sorten (bei Pflanzen) beziehungsweise Ras-

sen (bei Tieren) bezeichnet, andererseits auch deren Ergebnis Kreuzung genannt. (Wikipedia, Kreuzung)

- b. **Wahrnehmung** (auch *Perzeption* genannt) ist bei Lebewesen **der Prozess und das subjektive Ergebnis** der Informationsgewinnung (*Rezeption*) und -verarbeitung von Reizen aus der Umwelt und aus dem Körperinneren. (Wikipedia, Wahrnehmung)

Diese Texte lassen sich nicht in einer Aussage formalisieren, da aus ontologischer Sicht zwei verschiedene Klassen vorliegen. Dies ist in einer konkreten Ontologie kein größeres Problem, da die Namen der Klasse nur als Anzeigenamen fungieren; identifiziert werden die Klassen über eine *Uniform Resource Identifier* (URI). Das heißt, es kann zwei verschiedene Klassen mit demselben Namen geben, die aber über ihre URI wohlunterschieden sind. Dieser Mechanismus kann in $\mathcal{SHOIN}(\mathcal{D})$ nur imitiert werden. Im Folgenden werden die beiden Klassen über Zahlen voneinander unterschieden.

$$\begin{aligned} \text{Kreuzung}_1 &\sqsubseteq (\text{GeschlechtlicheFortpflanzung} \sqcap \exists \text{carriedOutBy}.\text{Lebewesen}) \\ \text{Kreuzung}_2 &\sqsubseteq (\exists \text{producedBy}.\text{Kreuzung}_1) \end{aligned}$$

Im folgenden Beleg werden zu einem *Product* sowohl ein *Producer* als auch ein *Process* angegeben:

- (135) **Als natürliche Sprache** bezeichnet man in der Sprachwissenschaft eine **von Menschen** gesprochene Sprache oder eine **Gebärdensprache**, die **aus einer ungesteuerten historischen Entwicklung** entstanden ist. (Wikipedia, Natürliche Sprache)

Da beide semantischen Rollen bzgl. der Formalisierung keinen Unterschied ergeben, sähe sie folgendermaßen aus:

$$\begin{aligned} \text{NatuerlicheSprache} &\sqsubseteq \\ &(\exists \text{producedBy}.\text{Mensch} \sqcap \exists \text{producedBy}.\text{HistorischeEntwicklung}) \end{aligned}$$

An dieser Stelle kommt die Bedeutung der Quantoren zu tragen: \exists bedeutet, dass eine Relation dieser Art existieren muss; es sagt nicht aus, dass dies die einzige ist. Deswegen kann *producedBy* zweimal verwendet werden.

Producer und *Product* können auch mit einem *Beneficient* zusammenstehen, wie der folgende Text belegt:

- (136) *Pressefotografen stellen für Printmedien wie Zeitungen, Zeitschriften, Magazine, aber auch für andere Medien wie Fernsehsender oder Onlinemedien **aktuelle, reproduktionsfähige Bilder von Personen, Vorgängen oder Ereignissen** her.* (Wikipedia, Pressefotograf)

Eine etwas der Anschaulichkeit wegen verkürzte Formalisierung sähe folgendermaßen aus:

$$\text{Pressefotograf} \sqsubseteq (\text{Fotograf} \sqcap \exists \text{produces.Fotos} \sqcap \exists \text{producesFor.Printmedium})$$

Product kann im Übrigen auch mit *Instrument* alleine in einer Relation stehen:

- (137) *Ein Saiteninstrument, auch Chordophon [...], ist ein Musikinstrument, bei dem zur **Tonerzeugung eine oder mehrere Saiten** verwendet werden, die zwischen zwei Punkten gespannt sind.* (Wikipedia, Saiteninstrument)

Diese Kombination ist interessant, da ein *Producer* indirekt dadurch vorhanden ist, dass die semantische Rolle *Instrument* mit einer agentivischen Rolle in Relation stehen muss. Es muss etwas geben, das sich dem Instrument als solches bedient. Es ist aber nicht immer wichtig, was für eine Entität den *Producer* besetzt, deswegen muss er, wie in (137), nicht zwangsweise realisiert werden. Bzgl. der Formalisierung ergeben sich hier allerdings zwei unterschiedliche Möglichkeiten: Zunächst kann eine ontologische Relation *usedFor* definiert werden, in der *Instrument* und der *Process* stehen, für welchen dieses gebraucht wird. Zunächst soll nur diese Aussage und nicht die ganze Definition formalisiert werden:

$$\text{Saiten} \sqsubseteq \exists \text{usedFor.Tonerzeugung}$$

Die andere Möglichkeit besteht darin, die Saiten nicht als *Instrument*, sondern als *Producer* zu begreifen und in Relation mit dem *Product* zu setzen, das als Teil des Kompositums *Tonerzeugung* realisiert wird.

$$\text{Saiten} \sqsubseteq \exists \text{produces.Ton}$$

Die erste Variante ist m.E. treffender, da bei der Tonerzeugung ein willentlich handelnder *Producer* vorliegt, der die Saiten zur Tonerzeugung nutzt und der im Text nur aufgrund des Passivs nicht genannt wird.

Das *definiendum* des Textes war allerdings *Saiteninstrument* und nicht *Saiten*. Die Relation zwischen diesen beiden Klassen ist die Meronymie, was im Text unspezifisch durch *bei dem* ausgedrückt wird. Der Text müsste also durch zwei Formeln ausgedrückt werden:

$$\begin{aligned} \text{Saiteninstrument} &\sqsubseteq (\exists \text{consists.Saiten}) \\ \text{Saiten} &\sqsubseteq \exists \text{usedFor.Tonerzeugung} \end{aligned}$$

Es existiert eine Rolle, die exklusiv mit *Product* und *Producer* oder *Process* in einer Relation steht: *Ressource*. Sie bezeichnet im weitesten Sinne ein Material, das durch einen Vorgang, bei dem etwas anderes entsteht, verändert oder ganz aufgelöst wird. Im folgenden Beleg stehen bspw. *Product*, *Process* und *Ressource* in Relation:

- (138) **Alkoholische Getränke** werden aus **kohlenhydrathaltigen Flüssigkeiten** durch **alkoholische Gärung** erzeugt. (Wikipedia, Alkoholische Gärung)

Hier lassen sich erneut zwei ontologische Relationen *convertedFrom* und *convertedTo* ansetzen.

$$\begin{aligned} \text{AlkoholischeGaerung} &\sqsubseteq \\ &(\exists \text{convertedFrom.KohlenhydrathaltigeFluessigkeit} \sqcap \\ &\exists \text{producedBy.AlkoholischeGaerung}) \end{aligned}$$

Ressource ist eine Rolle, die häufig in einer Kette steht: Ein *Product*, das aus einer *Ressource* hergestellt wurde, kann selbst wieder als *Ressource* weiterverarbeitet werden. Es handelt sich hierbei allerdings nicht zwingend um Transitivität: Nur weil B aus A hergestellt wird, und C aus B, wird C nicht zwingend aus A hergestellt. Dies wäre nur der Fall, wenn B ausschließlich aus A und C ausschließlich aus B hergestellt wird.

Im Folgenden sind zwei Belege, bei denen solch eine Kette beschrieben wird:

- (139) a. **Samen**: Aus **der Blüte** einer Pflanze sich entwickelndes Gebilde, aus dem **eine neue Pflanze** entstehen kann. (Der Duden, Das Bedeutungswörterbuch: 811)
- b. Als **Stoffwechsel** oder **Metabolismus** [...] bezeichnet man die gesamten chemischen und physikalischen Vorgänge der Umwand-

lung **chemischer Stoffe bzw. Substrate** (z. B. Nahrungsmittel und Sauerstoff) in **Zwischenprodukte** (Metaboliten) und **Endprodukte** im Organismus von Lebewesen. (Wikipedia, Stoffwechsel)

Solche Texte sollten am besten in mehrere Formeln aufgeteilt werden. Bspw.:

$\text{Samen} \sqsubseteq (\text{Gebildet} \sqcap \exists \text{convertedFrom.Blute})$
 $\text{Blute} \sqsubseteq (\exists \text{isPartOf.Pflanze} \sqcap \exists \text{convertedTo.Pflanze})$

Der folgende Beleg enthält ein *Product*, welches durch das Verb *entstehen* angedeutet wird, ohne dass der *Producer* dabei genannt wird. Stattdessen wird ein Argument realisiert, das sich am ehesten mit *Goal* annotieren lässt:

(140) **Die griechische Tragödie** [...] entstand im Rahmen der offiziellen **Feierlichkeiten in Athen zu Ehren des Weingottes Dionysos**. (Wikipedia, Griechische Tragödie).

Um die Relation zwischen einem *Product* und *Goal* auszudrücken, soll die ontologische Rolle *producedFor* angesetzt werden:

$\text{GriechischeTragoedie} \sqsubseteq (\text{Tragoedie} \sqcap \exists \text{producedFor.FestDesDionysos})$

Der folgende Text ist ein Beleg für *No-Producer*:

(141) **Die Pilze** (Fungi) bilden das dritte große Reich eukaryotischer Lebewesen neben den Tieren (Animalia) und den Pflanzen (Plantae). Sie sind wie die Pflanzen, zu denen sie lange gerechnet wurden, sesshaft, können jedoch keine **Photosynthese** treiben. (Wikipedia, Pilz)

Da es sich hierbei um die Relation zwischen *Producer* und *Process* handelt, die verneint wird, sähe eine Formalisierung folgendermaßen aus:

$\text{Pilz} \sqsubseteq (\text{Lebewesen} \sqcap \exists \neg \text{carriesOut.Photosynthese})$

Ebenso wurde ein Beleg gefunden, in dem etwas als *No-Product* ausgewiesen wird:

(142) **Gewohnheitsrecht** ist ungeschriebenes Recht, das nicht **durch Gesetzgebung** zustande kommt, sondern auf lange andauernder Anwendung von Rechtsvorstellungen und Regeln beruht, die die Beteiligten im Rechtsverkehr als verbindlich akzeptieren. (Wikipedia, Gewohnheitsrecht)

Eine mögliche Formalisierung sähe folgendermaßen aus:

$\text{Gewohnheitsrecht} \sqsubseteq (\text{UngeschriebenesRecht} \sqcap \exists \neg \text{producedBy.Gesetzgebung})$

8.4 Depiction

Bei der Relation zwischen *Depicting* und *Depicted* handelt es sich um das Verhältnis zwischen einer Sache, die etwas darstellt, und dem, was dargestellt wird. Die beiden Rollen stehen in einer festen Verbindung. Das Besondere an dieser Relation ist, dass keine der beiden Rollen agentivisch ist. Die Relation ist weiterhin weder transitiv noch reflexiv oder symmetrisch. Der Aspekt der Symmetrie mag weiter ausgeführt werden: Bspw. kann ein Bild, *Depicting*, eine Person, *Depicted*, darstellen. Dieser Umstand beruht auf einer Form von Ähnlichkeit. Verweist dann nicht die Person auch auf das Bild von ihr? Natürlich ähnelt die Person auch dem Bild von ihr, aber dies zu tun, ist nicht ihre Intention. Das Bild ist sekundär und orientiert sich an dem, was es darstellt, nicht umgekehrt.

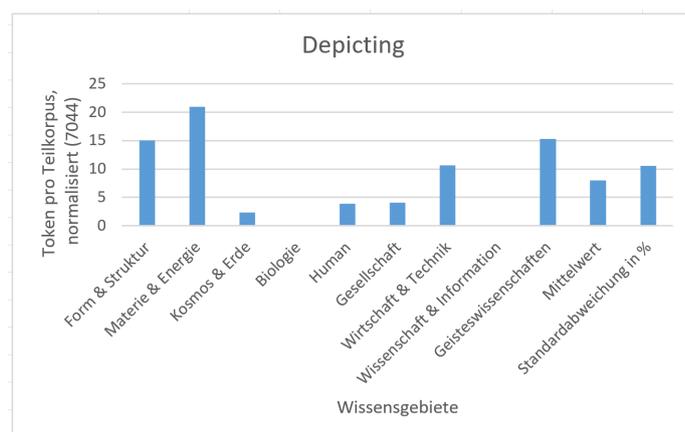


Abbildung 18: Die Verteilung von *Depicting* über die Teilkorpora

Der Annotationstyp ist im Bereich Biologie sowie Wissenschaft und Information gar nicht vertreten und am häufigsten in Materie und Energie. Die Standardabweichung befindet sich mit knapp über zehn Prozent im Mittelbereich der Standardabweichungen bei semantischen Rollen.

Das Beispiel hat womöglich den Eindruck erweckt, die Relation beruhe allein auf dem Prinzip der Ähnlichkeit. Dem ist aber nicht so. Es gibt diverse Arten, wie eine Sache auf eine andere verweisen kann, wobei viele typisch für ein bestimmtes Wissensgebiet sind.

In der Mathematik gibt es viele Belege für diese Relation, wobei diese im

Gegensatz zu vielen anderen Belegen aus diesem Wissensbereich durchaus wörtlich und nicht metaphorisch zu verstehen sind:

- (143) a. **Skalar**: *Mathematische Größe, die allein durch einen Zahlenwert bestimmt wird.* (Der Duden, Das Fremdwörterbuch: 1036)
 b. **Ein Skalar** ist eine *mathematische Größe, die allein durch die Angabe eines Zahlenwertes charakterisiert ist (in der Physik gegebenenfalls mit Einheit).* (Wikipedia, Skalar)

Zur Benennung der ontologischen Relationen bietet es sich an, die Namen der semantischen Rollen zu nutzen, hier *depicts* und *depictedBy*:

$$\text{Skalar} \sqsubseteq (\text{MathematischeGröße} \sqcap \leq 1 \text{depictedBy.Zahlenwert} \sqcap \geq 1 \text{depictedBy.Zahlenwert})$$

Dies ist im Übrigen erneut ein Beleg, für den es notwendig ist, zwei numerische Quantoren zu kombinieren, um auszudrücken, dass eine ganz bestimmte Zahl von Relationen dieser Art vorliegt.

Für die Physik ist es typisch, Kennwerte zu definieren, welche ein bestimmtes physikalisches Verhalten kennzeichnen.

- (144) a. **Ein Schallereignis** ist ein *physikalisch-akustischer Vorgang, der räumlich, zeitlich und eigenschaftlich durch physikalische Parameter bestimmt ist und objektiv existiert, unabhängig davon, ob oder wie es wahrgenommen wird. Beschrieben wird ein Schallereignis z. B. über Schallfeldgrößen, Schallquellenentfernung, Schallereignisrichtung und Frequenzspektrum.* (Wikipedia, Schallereignis)
 b. *In der Physik beschreibt ein Feld die räumliche Verteilung einer physikalischen Größe.* (Wikipedia, Feld)
 c. **Die Schallschnelle** [...] gibt an, mit welcher *Wechselgeschwindigkeit die Luftteilchen (bzw. Teilchen des Schallübertragungsmediums) um ihre Ruhelage schwingen; also die Momentangeschwindigkeit eines schwingenden Teilchens.* (Wikipedia, Schallschnelle)

Als Beispiel sei Beleg (144c) formalisiert:

$$\text{Schallschnelle} \sqsubseteq$$

$$\exists \text{depicts.MomentangeschwindigkeitSchwingendenTeilchens}$$

Überhaupt lassen sich im wissenschaftlichen Vokabular viele Beispiele für die Relation *Depicting-Depicted* finden: Modelle, Thesen, Hypothesen, Grafiken

etc. sind alles potenzielle Träger einer dieser Rollen. Folgend ein entsprechender Beleg aus dem Gesellschaftsbereich:

- (145) **Lerntheorien** sind Modelle und Hypothesen, anhand derer **Lernvorgänge** psychologisch beschrieben und erklärt werden sollen. (Wikipedia, Lerntheorie)

Lerntheorie \sqsubseteq (*Modell* \sqcap *Hypothese* \sqcap \exists *depicts.Lernvorgang*)

Lerntheorien leisten selbstverständlich mehr, als Lernvorgänge darzustellen, und die Formalisierung erlaubt sich an dieser Stelle eine Verallgemeinerung, die auch mit einer Ungenauigkeit einhergeht. Es ist aber anzunehmen, dass diese Formalisierung in den meisten Ontologien hinreichend genau sein sollte. Dies ist sie, wenn sie geeignet ist, die Lerntheorie von anderen Theorien und Modellen zu unterscheiden, was hauptsächlich dadurch geleistet werden kann, dass sich unterschiedliche Klassen von Theorien und Modelle dadurch unterscheiden, was sie darstellen. Eine Lerntheorie und eine physikalische Theorie unterscheiden sich dadurch, dass sie unterschiedliche Dinge darstellen wollen. Diesen Unterschied darzustellen, ist *depicts* hinreichend spezifisch.

Ein weiteres Anwendungsgebiet aus der Wissenschaftssprache stellen gewisse Werte, Vorgänge oder Entitäten dar, die als Marker betrachtet werden und als solche auf etwas anderes verweisen:

- (146) **Die germanische (im germanistischen Zusammenhang „erste“) Lautverschiebung** (fachsprachlich kurz „Grimm“, englisch auch Grimm's law genannt) kennzeichnet **den Übergang vom (ur)indogermanischen zum (ur)germanischen Konsonantensystem**. (Wikipedia, Erste Lautverschiebung)

Der folgende Beleg ist ein Beispiel, bei dem die Relation tatsächlich reflexiv ist: Wenn alles, was direktorial ist, einem Direktor/einer Direktorin ähnelt, ähnelt ein Direktor/eine Direktorin allem Direktorialem:

- (147) **Direktorial**: [...] c) **einem Direktor / einer Direktorin** [in der Art des Benehmens] *ähnelnd, entsprechend*. (Der Duden, Das Fremdwörterbuch: 273)

Wenn dies für wichtig erachtet wird, könnte es in diesem Fall über Äquivalenz ausgedrückt werden.

Direktorial $\equiv (\exists \text{depicts}.\text{Direktor})$

In die Relation kann ein *Instrument* als zusätzliches Argument aufgenommen werden:

- (148) **Lyrik**: *Dichtungsgattung, in der subjektives Erleben, Gefühle, Stimmungen usw. od. Reflexionen mit den Formmitteln von Reim, Rhythmus, Metrik, Takt, Vers, Strophe u.a. ausgedrückt werden.* (Der Duden, Das Fremdwörterbuch: 676)

Diese Kombination bedarf zwei neuer Rollen: *depictsWith* und *depictedWith*.

Die folgende Formalisierung ist auf das Wesentliche verkürzt:

Lyrik $\sqsubseteq (\text{Dichtung} \sqcap \exists \text{depicts}.\text{Stimmung} \sqcap \exists \text{depictsWith}.\text{Reim})$

Außerdem fand sich ein Beleg, bei dem *Place* und *Beneficient* zusätzliche Argumente darstellen:

- (149) **Konzert** [...] *im Sinne einer Musikveranstaltung nennt man den Vortrag von Musik vor einem eigens zu diesem Zweck versammelten Publikum („Hörerschaft“), meist auf einer dafür geeigneten oder eigens errichteten Bühne.* (Wikipedia, Konzert)

Der *Beneficient*, *Publikum*, wird wegen der Präposition *vor* in einer Form realisiert, die den Annotationstyp *Place* nahelegt, und in der Tat könnte dieses Argument als ein solcher interpretiert werden, dennoch wird ein Konzert auch für ein Publikum gegeben, wie die Phrase *zu diesem Zweck* nahelegt. Ontologische Rollen, die ein *Place* als Objekt haben, sollen mit dem Suffix *At* belegt werden. In der folgenden Formalisierung wird *Publikum* als *Beneficient* und *Bühne* als *Place* verstanden:

Konzert $\sqsubseteq (\text{Vortrag} \sqcap \exists \text{depicts}.\text{Musik} \sqcap \exists \text{depictsFor}.\text{Publikum} \sqcap \exists \text{depictsAt}.\text{Buehne})$

Einen etwas verwirrenden Fall stellt das Verhältnis zwischen Schauspieler*innen und deren Rollen dar. Es liegt nahe, dies als *Agent* und *Role* zu annotieren; dies scheint nahezu der prototypische Fall dieser Relation zu sein. Bei genauerer Betrachtung scheint aber die Relation zwischen *Depicting* und *De-*

pictured angebrachter. Dafür sprechen auch die Belege (150a) und (150b), in denen die Verben *darstellen* und *verkörpern* verwendet werden, die beide starke Indikatoren für *Depicted* sind. Die Relation zwischen *Agent* und *Rolle* beschreibt jemanden, der aus einer bestimmten Rolle oder Funktion heraus handelt; Schauspieler*innen hingegen handeln nicht aus einer Rolle heraus, sondern die Rolle darzustellen, ist die Handlung.

- (150) a. **Schauspieler**: *Person, die (nach entsprechender Ausbildung) bestimmte Rollen auf der Bühne, im Film o.Ä. künstlerisch gestaltet, darstellt.* (Der Duden, Das Bedeutungswörterbuch: 824)
- b. *Als **Schauspieler** (genannt auch *Mime*) wird ein Akteur bestimmter künstlerischer und kultureller Praktiken bezeichnet, der mit Sprache, Mimik und Gestik **eine Rolle** verkörpert oder als (Kunst-) Figur mit dem Publikum interagiert.* (Wikipedia, Schauspieler)

Zum Abschluss dieses Unterkapitels soll ein Beleg diskutiert werden, der in mehrfacher Hinsicht eine Herausforderung darstellt:

- (151) **Der Satz des Pythagoras** (*auch Hypotenusensatz*) *ist einer der fundamentalen Sätze der euklidischen Geometrie. Er besagt, dass in allen ebenen rechtwinkligen Dreiecken die Summe der Flächeninhalte der Kathetenquadrate gleich dem Flächeninhalt des Hypotenusenquadrates ist.* (Wikipedia, Satz des Pythagoras)

Zunächst stellt sich die Frage, ob das Verhältnis zwischen *Satz des Pythagoras* und dem, was dieser aussagt, überhaupt als *Depicting-Depicted* klassifiziert werden sollte. Dies schien am naheliegendsten: Der Satz oder die Formel stellt einen bestimmten Sachverhalt dar. Es soll aber nicht bestritten werden, dass diese Einteilung zur Diskussion steht. Weiterhin stellt sich die Frage, wie der komplexe Sachverhalt, der dann als *Depicted* annotiert würde, zu formalisieren sei. Eine Möglichkeit bestände darin, die Formel $a^2 + b^2 = c^2$ als Text, also als einen Datentyp zu begreifen, dann könnte über *depicts* auf ihn verwiesen werden. Dieses Vorgehen ist aber sehr umständlich. Naheliegender ist es, in solchen Fällen nicht in den Kategorien der Sprachwissenschaft zu denken und den Satz des Pythagoras einfach in einer Liste von Formeln zu speichern.

8.5 Shape

Die Relation zwischen *Shape* und *Substance* beschreibt das Verhältnis zwischen einer Entität und einer Form, die dieser Entität eigen ist und über die sie sich definieren lässt. Beide Argumente sind nicht agentivisch.

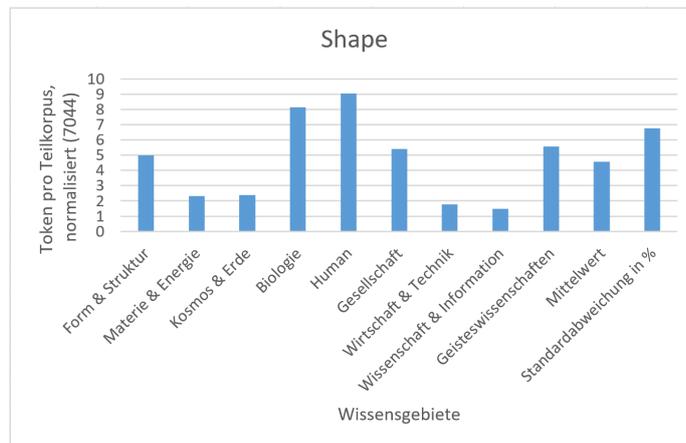


Abbildung 19: Die Verteilung von *Shape* über die Teilkorpora

Für diesen Annotationstyp gibt es nicht viele Belege. Die meisten lassen sich im Humanbereich mit normalisierten neun Belegen finden. Der Typ ist allerdings in allen Wissensgebieten vertreten und weist deswegen mit unter sieben Prozent eine für die semantischen Rollen geringe Standardabweichung auf.

Beleg (152a) ist ein sehr prototypischer Vertreter dieser Annotation: Die PP *in der Form* legt die Annotation *Shape* nahe, und die NP *unsichtbarer Stoff* ist nicht nur die Oberklasse von *Gas*, sondern ist auch ein Indiz für *Substance*. *Substance* umfasst als Annotationstyp aber mehr als nur Materie und kann auch konkrete Objekte bezeichnen, wie in (152b):

- (152) a. **Gas:** *Unsichtbarer Stoff in der Form wie Luft.* (Der Duden, Das Bedeutungswörterbuch: 420)
- b. **Sonne:** *Als gelb bis glutrot leuchtende Scheibe am Himmel erscheinender, der Erde Licht und Wärme spendender Himmelskörper.* (Der Duden, Das Bedeutungswörterbuch: 892)

Eine entsprechende Relation in der Ontologie könnte *shapedAs* genannt und folgendermaßen verwendet werden:

Sonne \sqsubset (*Himmelskoerper* \sqcap \exists *shapedAs.Scheibe*)

Viele Belege für diese Relation lassen sich im Wissensbereich der Biologie finden:

- (153) a. **Rose:** (*als Strauch wachsende*) Stacheln tragende Pflanze mit meist glänzenden Blättern und viel blättrigen, oft angenehm und stark duftenden, gefüllten Blüten in verschiedenen Farben. (Der Duden, Das Bedeutungswörterbuch: 798)
- b. **Erdbeere:** Wild und im Garten wachsende, niedrige Pflanze, deren Blätter in Rosetten stehen und die Erdbeeren als Früchte hervorbringen. (Der Duden, Das Bedeutungswörterbuch: 347)
- c. **Pflaume:** Dunkelblaue, eiförmige Frucht des Pflaumenbaums mit gelblich grünem, aromatischem Fruchtfleisch und einem länglichen Stein. (Der Duden, Das Bedeutungswörterbuch: 737)
- d. **Als Strauch** (auch Busch oder Kleines Gehölz) wird eine Wuchsform von Pflanzen bezeichnet. (Wikipedia, Strauch)

Der Name *shapedAs* suggeriert ein statisches Merkmal, wobei der Beleg (153a) hier einen Prozess nahelegt. Ob dieser Unterschied – etwas hat beständig eine Form oder wächst in diese hinein – von Bedeutung ist, muss von Ontologie zu Ontologie entschieden werden. In dieser Arbeit sollen alle Fälle durch *shapedAs* abgedeckt werden.

Die Rolle *Shape* wird oft explizit durch ein Wort wie *Form* angesprochen; dies ist in (152a) der Fall und geschieht in verschiedenen Varianten in (153): In (153c) wird das adjektivische Kompositum *eiförmige* verwendet und in (153d) das substantivische Kompositum *Wuchsform*. Im zweiten Fall muss im Übrigen das Inverse von *shapedAs*, das *shapeOf* genannt werden soll, angesetzt werden:

$Strauch \sqsubset (\exists shapeOf.Pflanze)$

(154) ist ein Beleg aus dem Teilkorpus zu Materie und Energie:

- (154) **Schall:** [...] **Wellenförmig** sich ausbreitende Schwingungen, die vom menschlichen Gehör wahrgenommen werden können. (Der Duden, Das Bedeutungswörterbuch: 819)

Shape ist in diesem nicht die Form der *Substance* an sich, sondern die Form, in der sie sich ausbreitet, was in diesem Fall effektiv das Gleiche ist, da Schall eine Entität ist, die sich permanent ausbreitet, die also nicht in statischer Form

existiert. Die Formulierung des Duden ist somit spezifisch, es könnte auch ebenso geschrieben werden, dass Schall wellenförmig sei.

Schall \sqsubseteq
(*Schwingung* \sqcap *existsshapedAs.Welle*
 \sqcap \exists *experiencedBy.MenschlichesGehoer*)

Auch bestimmte Strukturformen lassen sich unter dieser Relation fassen. Hierfür lassen sich besonders im Gesellschaftsbereich Belege finden:

- (155) a. **Das Regierungssystem** beschreibt die formale Ausgestaltung und die Funktionsweise **einer Regierung**. (Wikipedia, Regierungssystem)
- b. **Das Direktorialsystem** ist eine Organisationsform **der Führungsspitze** und zugleich im politischen Kontext ein Regierungssystem. Alternativbezeichnung sind für letzteres **Direktoriales Regierungssystem** oder **Direktorialverfassung**. (Wikipedia, Direktorialsystem)

Im folgenden Beleg ist wieder das Wort *Form* Indikator für *Shape*. In diesem Fall ist es allerdings Kopf einer NP, dessen Adjektivattribut letztlich die Rolle trägt:

- (156) **Die geschriebene Sprache**, auch *Schriftsprache* genannt, bezeichnet ein Zeichensystem, das formal nicht festgelegt ist, jedoch speziellen Normen gehorcht und **in schriftlicher Form** vorliegt. (Wikipedia, Geschriebene Sprache)

Eine mögliche Formalisierung, die das Adjektiv als Klasse übersetzt, sähe folgendermaßen aus:

GeschriebeneSprache \sqsubseteq (*Zeichensystem* \sqcap \exists *shapedAs.Schriftlich*)

Schließlich wurden auch in den Geisteswissenschaften Belege für diese Relation gefunden. In (157a) wird statt *Form* mit dem Wort *Gestaltung* die Rolle angezeigt. In (157b) werden erneut substantivische Komposita mit *Form* als Kopf verwendet:

- (157) a. **Bühnenbild** bezeichnet entweder die visuelle Gestaltung **eines szenischen Raumes** oder dessen Material, also sämtliche Einrichtungen, Malereien und Kulissen sowie die Bühnenmaschinerie, aus denen Bühnenbilder bestehen. (Wikipedia, Bühnenbild)
- b. **Epik** [...] ist neben Dramatik und Lyrik eine der drei großen Gattungen der Literatur und umfasst erzählende Literatur in **Vers- oder Prosaform**. (Wikipedia, Epik)

Zum Abschluss zwei Belege aus dem Duden, bei denen das mit *Substance* annotierte Argument ein Adjektiv ist:

- (158) a. **Quadratisch**: *in der Form eines Quadrats.*
(Der Duden, Das Fremdwörterbuch, 938)
b. **Konvex**: *Erhaben, nach außen gewölbt.*
(Der Duden, Das Fremdwörterbuch, 620)

Der verkürzende Schreibstil des Dudens hat in (158a) einen etwas redundanten Effekt, und dementsprechend ist auch eine Formalisierung des Textes zirkulär:

Quadratisch $\sqsubseteq \exists \text{shapeOf} . \text{Quadrat}$

In (158b) ist die Definition zwar nicht zirkulär, da es sich aber um zwei Adjektive, eines davon mit Attribut, handelt, die in der Relation stehen, stellt sich die Frage, ob sie mittels einer ontologischen Rolle oder einfach über Äquivalenz realisiert werden sollten:

Konvex $\sqsubseteq \exists \text{shapeOf} . \text{NachAussenGewoelbt}$
Konvex $\equiv \exists \text{shapeOf} . \text{NachAussenGewoelbt}$

8.6 Experiencer

Experiencer bezeichnet eine Entität, der eine Sinneswahrnehmung widerfährt. Die Rolle scheint zu gleichen Teilen agentivisch als auch affiziert zu sein. So kann sich eine Sinneswahrnehmung aufdrängen, aber auch explizit gesucht werden. Im Gegensatz zu *Agent* und *Force* hat *Experiencer* ein festes Gegenstück, *Stimmulus*, das ebenso zu gleichen Teilen agentivisch und affiziert ist.

Die Zahlen zu diesem Annotationstyp sind zu klein, um aussagekräftig zu sein. Die Standardabweichung erscheint in der Abbildung hoch, dies ist aber den wenigen Belegen geschuldet. Statistisch befindet sie sich mit knapp unter zehn Prozent in einem für die semantischen Rollen herkömmlichen Bereich.

- (159) *Der Mann sieht die Blume.*

Bei *Experiencer* handelt es sich also um eine Rolle, die wie *Agent* und *Force* ganz regulär das Subjekt des Satzes bilden kann, wodurch die Vermutung na-



Abbildung 20: Die Verteilung von *Experiencer* über die Teilkorpora

heliegt, dass sich diese Rolle ebenfalls wenig an grammatikalischen Zusammenhängen identifizieren lassen wird. Ähnlich wie *Force* über die Oberklassen der entsprechenden Entität identifiziert werden kann, kann *Experiencer* über das Verb identifiziert werden; dieses muss einer Klasse von Wahrnehmungsverben, wie *sehen*, *schmecken*, *erkennen* etc., angehören.

Interessanterweise tritt dieser vermeintlich prototypische Fall, *Experiencer* als Subjekt, im Korpus kaum auf. Am häufigsten wird *Experiencer* als PP mit *durch* als Kopf realisiert.

- (160) a. *Überwachung ist die zielgerichtete Beobachtung und Informationserhebung von Objekten, Personen, Personenvereinigungen oder Gegenständen durch am Geschehen unbeteiligte Dritte.* (Wikipedia, Überwachung)
- b. *Als auditive, aurale oder akustische Wahrnehmung oder Hören bezeichnet man die Sinneswahrnehmung von Schall durch Lebewesen, genauer z. B. die Hörereignisse, die bei bestimmten Schallereignissen entstehen.* (Wikipedia, Auditive Wahrnehmung)

In diesen Texten handelt es sich bei den *definienda* um substantivierte Verben, und der *Experiencer* hängt als PP-Attribut an der jeweiligen Oberklasse, die ebenfalls ein substantiviertes Verb ist. Der *Stimmulus* ist in beiden Fällen als PP mit *von* als Kopf realisiert.

Ueberwachung \sqsubseteq (*Beobachtung* \sqcap
 \exists *experiencedBy.UnbeteiligteDritte* \sqcap \exists *experiences.Objekte*)

Der folgende Beleg zeigt, dass *Experiencer* ähnlich wie *Agent* und *Force* in

einem metaphorischen Sinne verwendet werden kann:

- (161) *Der Ausdruck juristische Person ist mehrdeutig: [...] Herkömmlich wird der Ausdruck juristische Person nur für solche Rechtspersonen verwendet, die der Gesetzgeber ausdrücklich anerkannt hat.* (Wikipedia, Juristische Person)

Das Verb *anerkennen* hat im juristischen Sinne eine speziell definierte Bedeutung, die nur im übertragenen Sinne etwas mit Sinneswahrnehmung zu tun hat. Die folgende Formalisierung ist demnach nur für eine Ontologie geeignet, die den Fokus nicht auf juristische Begriffe legt, sondern diese nur am Rande behandelt.

$JuristischePerson \sqsubseteq (Rechtsperson \sqcap experiencedBy.Gesetzgeber)$

Da in diesem Fall der *Stimulus* auch den zu definierenden Begriff darstellt, wird die passivische Variante der ontologischen Relation verwendet.

Eine ähnlich metaphorische Verwendung, die als Metonymie bezeichnet werden kann, liegt in den folgenden Belegen vor, in denen es um ein Betätigungsgebiet und die jeweilige Tätigkeit geht. Durch solche Formulierungen kommt die Tokenanzahl für die naturwissenschaftlichen Wissensbereiche zustande:

- (162) a. *Modejournalismus ist als Teilbereich des Kulturjournalismus ein journalistisches Special-Interest-Themengebiet, das sich mit den Themen Mode, Lifestyle, Fotografie und Styling befasst.* (Wikipedia, Modejournalismus)
- b. *Die Astrophysik befasst sich mit den physikalischen Grundlagen der Erforschung von Himmelserscheinungen und ist ein Teilgebiet der Astronomie.* (Wikipedia, Astrophysik)
- c. *Trigonometrie: Dreiecksmessung; Zweig der Mathematik, der sich mit der Berechnung von Dreiecken unter Benutzung der trigonometrischen Funktionen befasst.* (Der Duden, Das Fremdwörterbuch: 1139)

Es ist selbstverständlich nicht das jeweilige Themengebiet, das sich mit diesen Dingen beschäftigt, sondern die jeweiligen Journalist*innen oder Wissenschaftler*innen – ein Detail, das für die meisten Ontologien zu vernachlässigen sein dürfte. Da der *Experiencer* in diesen Belegen das Subjekt ist, wird

bei der Formalisierung nur die aktivische Variante der ontologischen Relation verwendet. Bspw.:

Mode journalismus \sqsubseteq (*Kultur journalismus* \sqcap \exists *experiences.Mode*)

Der folgende Beleg stellt einen Randfall dar, der zu diskutieren ist:

(163) **Ein Unternehmen** ist eine wirtschaftlich selbstständige Organisationseinheit, die mit Hilfe von Planungs- und Entscheidungsinstrumenten **Markt- und Kapitalrisiken eingeht** und sich zur Verfolgung des Unternehmenszweckes und der Unternehmensziele eines oder mehrerer Betriebe bedient. (Wikipedia, Unternehmen)

Allgemein geht es um die Formulierung *ein Risiko eingehen*. Zunächst sieht dies nach einer Formulierung aus, die einen *Agent* fordert, wobei das *Risiko* den *Patient* bildet. An dieser Herangehensweise könnte man aussetzen, dass mit dem Risiko nicht wirklich etwas passiert und dies ohnehin eine abstrakte Sache ist. Die Formulierung könnte auch als Funktionsverbgefüge begriffen und durch das Verb *riskieren* paraphrasiert werden; dann wäre der *Patient* dasjenige, was riskiert wird. Ein Unternehmen riskiert bspw. sein Kapital. Es ließe sich aber auch argumentieren, dass *Risiko* ein *Stimmulus* ist, etwas das betrachtet und bewertet wird, und dass das Unternehmen in diesem Sinne der *Experiencer* ist.

Solche Diskussionen müssen von Fall zu Fall entschieden werden, wobei die Intention oder Perspektive der Ontologie eine große Rolle spielt. Der Beleg zeigt aber, wie viele Details und Konnotationen in natürlich-sprachlichen Ausdrücken stecken können.

Am Anfang des Unterkapitels wurde gesagt, dass *Experiencer* und *Stimmulus* in einer festen Relation zueinander stehen. Dies heißt aber nicht, dass immer beide realisiert werden. Es gibt Fälle, in denen eine der Rollen weggelassen werden kann, da sie ohne Weiteres ergänzt werden kann. Bei den folgenden Belegen fehlt der *Experiencer*. In (164a) kann aus dem Weltwissen ergänzt werden, dass Menschen oder Lebewesen Sterne am Nachthimmel sehen können bzw. überhaupt Beobachtungen anstellen. In (164b) ergibt sich die Auslassung aus der Substantivierung des Verbes *beobachten*; auch hier

kann aus dem Weltwissen ergänzt werden, dass speziell unter Verwendung technischer Hilfsmittel Menschen Dinge beobachten.

- (164) a. **Stern:** *Als silbrig-weißer, funkelnder Punkt besonders am nächtlichen Himmel sichtbares Gestirn.* (Der Duden, Das Bedeutungswörterbuch: 919)
- b. *Die Beobachtung ist die zielgerichtete, aufmerksame Wahrnehmung von **Objekten**, Phänomenen oder Vorgängen, gegebenenfalls unter Verwendung technischer Hilfsmittel.* (Wikipedia, Beobachtung)

Bzgl. der Formalisierung muss eine editoriale Entscheidung getroffen werden, was als *Experiencer* eingesetzt wird. In der folgenden Formel wurde *Mensch* eingesetzt.

$$\text{Stern} \sqsubseteq (\text{Gestirn} \sqcap \exists \text{experiencedBy.Mensch})$$

Der umgekehrte Fall, dass der *Stimmulus* nicht realisiert wird, ist ebenfalls anzutreffen. Im folgenden Beleg ist diese Auslassung dem ökonomischen Schreibstil des Dudens geschuldet:

- (165) **Auditiv:** *Das **Gehör** betreffend, zum Gehörsinn od. -organ gehörend.* (Der Duden, Das Fremdwörterbuch: 134)

In diesem Fall ist es auch in der Formalisierung nicht notwendig, den *Stimmulus* zu nennen. Auditiv ist schlichtweg alles, was über das Gehör wahrgenommen werden kann:

$$\text{Auditiv} \sqsubseteq \exists \text{experiencedBy.Gehoer}$$

Dass es sich überhaupt um die Relation zwischen *Experiencer* und *Stimmulus* handeln muss, kann nur aus dem Wort *Gehör* geschlussfolgert werden. Das Partizip *betreffend* ist in dieser Hinsicht nicht spezifisch genug.

Der folgende Beleg, der den fast gleichen Sachverhalt erläutert, ist dabei so formuliert, als handle es sich um *Force* und *Patient*.

- (166) **Ton:** *Auf das **Gehör** wirkende gleichmäßige Schwingung der Luft.* (Der Duden, Das Bedeutungswörterbuch: 960)

Ohne Weltwissen kann hier eine Einordnung als *Experiencer* und *Stimmulus* nicht erfolgen. Die Problematik wird noch dadurch verstärkt, dass aus physi-

kalischer Sicht eine Annotation als *Force* und *Patient* plausibel ist: Letztlich wirkt etwas auf etwas anderes ein. Die Rolle *Experiencer* betont allerdings, dass Wahrnehmung ein nicht komplett passiver Akt ist. Dieser Aspekt ist wiederum nur von Bedeutung, wenn das Gebiet der Kognition für die Ontologie eine Rolle spielt.

Im Korpus wurde ein Beleg gefunden, bei dem ein *Stimulus* mit *No-Instrument* in einer Relation steht. Der *Experiencer* muss bei diesem Beleg wieder ergänzt werden.

- (167) *Freiäugig oder freisichtig werden in Naturwissenschaft und Technik jene **Beobachtungen und Messungen** genannt, die **ohne optisches Hilfsmittel** (ausgenommen Korrekturen von Fehlsichtigkeiten der Augen) gewonnen werden, d. h. ohne Verwendung von zum Beispiel Fernrohr, Lupe, Mikroskop oder Kamera.* (Wikipedia, Freiäugig)

Eine Formalisierung könnte folgendermaßen aussehen:

$$\text{Freiaeuig} \sqsubseteq \exists \neg \text{experiencedWith.Hilfsmittel}$$

Die ontologische Relation *experiencedWith* stellt dabei die Verbindung zwischen einem *Stimulus* und einem *Instrument* her, mit dem der *Stimulus* wahrgenommen werden kann.

Bei den folgenden Belegen könnte diskutiert werden, ob es sich um *No-Experiencer* handelt, wobei dieser in (169) zusätzlich ergänzt werden müsste:

- (168) *Realitätsverlust beschreibt den geistigen Zustand **einer Person**, welche nicht (mehr) in der Lage ist, **die Situation**, in der sie sich befindet, zu begreifen.* (Wikipedia, Realitätsverlust)
- (169) *Selbstüberschätzung, manchmal auch Vermessenheitsverzerrung, ist eine Form der systematischen Fehleinschätzung **eigenen Könnens und eigener Kompetenzen**.* (Wikipedia, Selbstüberschätzung)

Eine solche Einteilung ist allerdings ungenau, da es sich nicht so verhält, dass die eigenen Kompetenzen bzw. die eigene Situation gar nicht wahrgenommen werden, sondern nur fehlerhaft wahrgenommen werden. Es wäre auch denkbar, eine Unterrelation von *experiences* einzuführen, die speziell dafür verwendet wird, wenn etwas fehlerhaft wahrgenommen wird: *experiences-Faulty*. Ein solcher Ansatz ist formal konsistent, da ein adverbial spezifiziertes

Verb sich als Unterrelation der Relation darstellen lässt, die das unspezifizierte Verb formalisiert (vgl. Lohnstein 2011: 143-144). Es ist aber zweifelhaft, ob eine solche Spezifizierung in einer Ontologie einen Mehrwert bringt, da sie vermutlich die Komplexität der ontologischen Struktur deutlich erhöht.

8.7 Sender

Sender ist eine Rolle mit starkem agentivem Anteil. Sie steht in einer relativ festen Relation mit zwei anderen Rollen: *Receiver* und *Content*. Die Relation beschreibt explizit den Austausch von Nachrichten und Informationen.

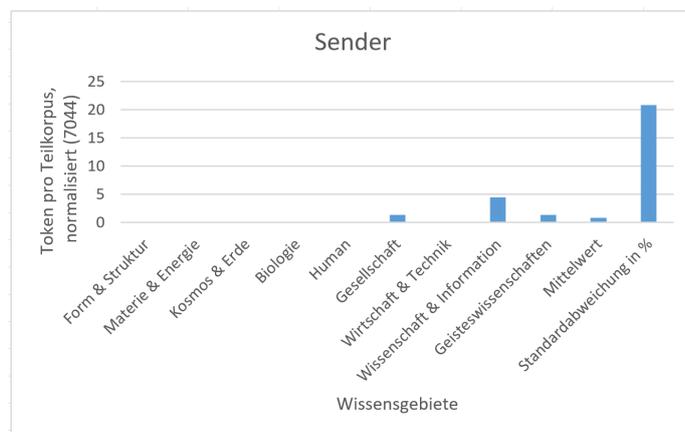


Abbildung 21: Die Verteilung von *Sender* über die Teilkorpora

Auch in diesem Fall liegen nicht genug Fälle vor, als dass die Statistik wirklich aussagekräftig sein könnte. In mehr als der Hälfte der Wissensgebiete ist der Annotationstyp nicht vertreten, weswegen sich mit knapp über zwanzig Prozent eine sehr hohe Standardabweichung ergibt.

Der Prototyp dieser Annotation wird im Sender-Empfänger-Modell beschrieben, welches gleichzeitig einen Beleg für die entsprechenden Rollen darstellt.

- (170) a. **Ein Empfänger** (auch Adressat oder Rezipient) ist im informationstheoretischen Sinn eines Sender-Empfänger-Modells eine Person oder Institution, die **eine Nachricht** oder eine andere Information **durch ein Medium vom Absender** (Adressant) übermittelt bekommt. (Wikipedia, Empfänger)

- b. *Information ist in der Informationstheorie das Wissen, das ein Absender einem Empfänger über einen Informationskanal vermittelt.* (Wikipedia, Information)

Eine Formalisierung könnte folgendermaßen aussehen:

$$\text{Empfaenger} \sqsubseteq ((\text{Person} \sqcup \text{Institution}) \sqcap \exists \text{receivesFrom.Absender} \sqcap \exists \text{getsSent.Nachricht} \sqcap \exists \text{sentBy.Medium})$$

Die Relation zwischen *Receiver* und *Sender* wird *receivesFrom* genannt, das Inverse wäre *sendsTo*; die Relation zwischen *Receiver* und *Content* wird *getsSent* genannt, dessen Inverses *sentTo* hieße. Zusätzlich wird im Text noch das *Medium* genannt, über das der *Content* übermittelt wird. Die entsprechenden ontologischen Relationen wurden *sendsBy* und *getsSentBy* genannt.

Content und *Medium* können alleine in einer Relation zueinander stehen.

Der folgende Text ist ein Beleg dafür:

- (171) *Compact Disc Read-Only Memory (CD-ROM) bezeichnet ein Permanentspeichermedium für digitale Daten.* (Wikipedia, CD-ROM)

Diese Relation wurde *stores* mit dem Inversen *storedOn* genannt.

$$\text{CD-ROM} \sqsubseteq (\text{Permanentspeichermedium} \sqcap \exists \text{stores.DigitaleDaten})$$

Die Relation wurde u.a. im Gesellschaftsbereich häufig annotiert, da auch Gesetze, Anweisungen etc. ihr zugehörig sind:

- (172) *Unter einem Hoheitsakt (staatlicher Hoheitsakt) versteht man eine Anordnung, die der Staat von oben herab (hoheitlich) beschließt, bei der somit Staat und Bürger in einem Über-Unterordnungsverhältnis (Subordinationsverhältnis) zueinander stehen.* (Wikipedia, Hoheitsakt)

In der folgenden Formalisierung wird außerdem die ontologische Relation *sentFrom* zwischen *Content* und *Sender* eingeführt. Dass der Bürger der *Receiver* ist, wird nur indirekt dadurch ausgedrückt, dass das Verhältnis zwischen Staat und Bürger als Unterordnungsverhältnis ausgedrückt wird.

$$\text{Hoheitsakt} \sqsubseteq (\text{Anordnung} \sqcap \exists \text{sentFrom.Staat} \sqcap \exists \text{sentTo.Buerger})$$

8.8 Place

Der Annotationstyp *Place* bedarf einer besonderen Betrachtung, da er unabhängig von einer Handlung bestehen kann. Dies unterscheidet ihn von allen anderen Annotationstypen, die bisher besprochen wurden. Letztlich werden mit diesem Typus verschiedene lokale Rollen wie *Path*, *Goal* (als Ziel einer Bewegung) oder *Source* (als Ursprung einer Bewegung) zusammengefasst, die sich über das Verhältnis einer Entität zu dem entsprechenden Ort unterscheiden (vgl. Sowa 2000: 508-510). Eine solche Differenzierung kann durchaus sinnvoll sein, wurde hier aber unterlassen, da sich für diese Arbeit keine weiteren Erkenntnisse durch solch eine Spezifizierung ergäben.

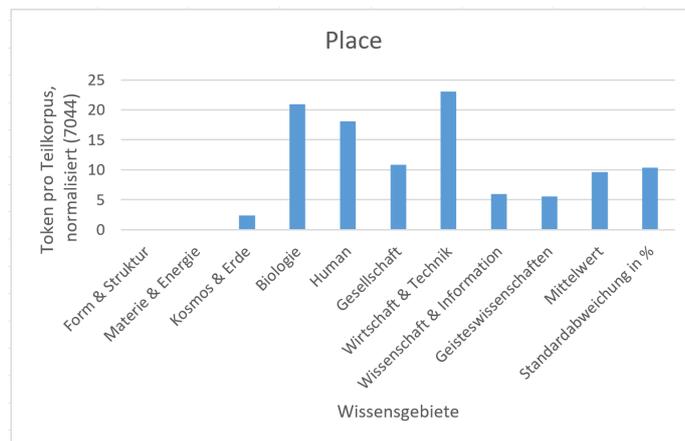


Abbildung 22: Die Verteilung von *Place* über die Teilkorpora

Abbildung 22 zeigt, dass der Typ mit einer Standardabweichung bei zehn Prozent recht gleich über alle Wissensgebiete verteilt ist. Nur bei Form und Struktur sowie Materie und Energie liegen keine Token vor. Der Annotationstyp wurde im Text bisher im Zusammenhang mit Beleg (149) angesprochen. Dies ist einer von zwei Belegen, bei dem *Place* in Relation mit einer Klasse steht. In den anderen Fällen steht *Place* in Relation mit einer Instanz. Dies ist auch der Grund, warum der Annotationstyp im DK nicht vertreten ist. Die folgenden Belege sind Beispiele für solche Relationen:

- (173) a. **BlackRock Inc.** (englisch ‚schwarzer Fels‘) ist eine international tätige US-amerikanische Fondsgesellschaft mit Sitz **in New York City**. (Wikipedia, BlackRock)

- b. **Kongeegen** (deutsch die Königseiche) ist eine Stieleiche, die mit einem Alter von 1400 bis 2000 Jahren nicht nur als ältester Baum Dänemarks gilt, sondern vermutlich auch die älteste Eiche Nord-europas ist. (Wikipedia, Kongeegen)

Instanzen werden in diesen Belegen über den Ort, an dem sie sich befinden, definiert. An dieser Stelle könnte gefragt werden, worin der Unterschied dieser Annotation zu dem Typ Ort-Gebiet der Teil-Ganzes-Relation bestehe, wie sie bspw. in Beleg (185) angesetzt wurde. In der Tat ist der Unterschied nicht immer leicht auszumachen. Wenn etwas als *Place* annotiert wurde, dann bedeutet dies, dass es sich um einen Ort handelt, mit dem eine andere Entität, die kein Ort ist, in Relation steht. Bei Bewegungsverben führt dies zu einer eindeutigen Unterscheidung: Wenn sich eine Entität zu einem Ort bewegt, dann ist diese Entität ein *Agent* und der Ort *Place*, und es liegt keine Form von Teil-Ganzes vor. In den Belegen (173) mag diese Unterscheidung weniger eindeutig sein, dennoch werden *BlackRock* und *Kongeegen* nicht als Bestandteile der jeweiligen Orte, und damit selber als Orte, wahrgenommen, sondern als Entitäten, die sich an einem Ort befinden.

Eine Formalisierung könnte folgendermaßen gestaltet sein:

$BlackRockInc. \sqsubseteq (\exists situatedAt. NewYorkCity)$

Diese Formel ist ungenau, da es sich streng genommen nur um den Hauptsitz von *BlackRock* handelt, der sich in *New York City* befindet; diese Ungenauigkeit dürfte aber in den meisten Ontologien vertretbar sein. Ansonsten könnte eine spezielle Rolle für Unternehmen angesetzt werden, bspw. *hatHauptsitzIn*, die auf den entsprechenden Ort verweist.

8.9 Zusammenfassung

Dieses Kapitel hat sich damit beschäftigt, wie die ontologisch relevanten Informationen, die nicht auf die grundlegenden Relationen der Mengenlehre abgebildet werden können, formalisiert werden können. Der Ansatz sah dabei vor, die relevanten Wortgruppen vor dem Konzept der semantischen Rollen zu analysieren. In einem zweiten Schritt, der den Kern dieses Kapitels bilde-

te, galt es dann, zu prüfen, welche Rollen miteinander in Relation stehen, und diese Relationen in ontologische Rollen zu überführen. Tabelle 7 stellt alle dabei ermittelten Rollen zusammen. Es sei noch einmal ausdrücklich darauf hingewiesen, dass dies die ontologischen Rollen sind, die anhand des Korpus ermittelt wurden; dies heißt nicht, dass nicht andere Kombinationen von semantischen Rollen möglich sind. Eine Besonderheit in Tabelle 7 stellten *Agent* und *Force* dar. Die Relation zwischen *Agent* und *Patient* sowie *Force* und *Patient* wird durch das Verb des entsprechenden Satzes dargestellt: Dieses wird in eine ontologische Relation überführt. Bei der Relation zwischen *Agent* und *Beneficient* wird ebenfalls das Verb überführt und um das Suffix *For* ergänzt. Steht ein *Agent* allein, wird gar keine ontologische Rolle angesetzt. Stattdessen wird das Verb in eine Klasse überführt und der *Agent* wird als Unterklasse dieser formalisiert.

Semantische Rolle	Ontologische Rolle	Semantische Rolle	Ontologische Rolle
Agent	VERB als Oberklasse	Instrument-Goal	usedFor
Agent-Beneficient	VERB + For	Instrument-Patient	usedFor + VERB
Agent-Goal	hasGoal	Process-Producer	carriedOutBy
Agent-Instrument	uses	Producer-Beneficient	producesFor
Agent-Patient	VERB	Producer-Process	carriesOut
Agent-Role	performsAs	Producer-Product	produces
Agent/Force-Manner	VERB + WithManner	Product-Beneficient	producedFor
Content-Medium	storedOn	Product-Goal	producedFor
Content-Receiver	sentTo	Product-Process	producedBy
Content-Sender	sentFrom	Product-Producer	producedBy
Depicted-Depicting	depictedBy	Product-Ressource	convertedFrom
Depicted-Instrument	depictedWith	Receiver-Medium	getsSentBy
Depicted-Place	depictedAt	Receiver-Sender	receivesFrom
Depicting-Depicted	depicts	Reiceiver-Content	getsSent
Depicting-Instrument	depictsWith	Role-Agent	performedBy
Depicting-Place	depictsAt	Sender-Content	sends
Experiencer-Instrument	experiencesWith	Sender-Medium	sendsBy
Experiencer-Stimulus	experiences	Sender-Receiver	sendsTo
Force-Instrument	uses	Shape-Substance	shapeOf
Force-Medium	movesIn	Stimulus-Experiencer	experiencedBy
Force-Patient	VERB	Substance-Shape	shapedAs
Goal-Agent	goalOf	Stimulus-Instrument	experiencedBy
Instrument-Agent	usedBy		

Tabelle 7: Die ontologischen Rollen für die Relationen zwischen den semantischen Rollen

Während in den Kapiteln zu den mengentheoretischen Relationen und den Pertinenzrelationen die konkrete Realisierung dieser Relationen einen großen Raum einnahm, war dies in diesem Unterkapitel weniger der Fall. Der Grund liegt darin, dass semantische Rollen für sich ein Konzept sind, das relativ stark

von konkreten Realisierungen abstrahiert. Bei der Anzahl der Rollen, die angesetzt wurden, und den komplexen Verflechtungen, die in diesem Kapitel ausgeführt wurden, stellt es daher keine aussichtsreiche Arbeit dar, alle syntaktischen Möglichkeiten, eine Rolle zu realisieren, im Detail zu beschreiben. Stattdessen wurden prototypische Formulierungen einzelner Rollen identifiziert, um dann gegebenenfalls festzustellen, dass diese Formulierungen auch genutzt werden können, um semantisch von diesen Prototypen abweichende Zusammenhänge zu beschreiben. Außerdem wurde ausgeführt, dass zur Identifikation einer Rolle die Formulierung alleine nicht hinreichend sein kann, sondern auch die semantische Klasse des Arguments betrachtet werden muss. Bspw. unterscheiden sich *Force* und *Agent* nur bzgl. dieser Klasse, die eine enthält unwillentlich-handelnde, die andere willentlich-handelnde Entitäten.

Abschließend lässt sich sagen, dass der Ansatz, ontologisch-relevante Informationen als semantische Rollen zu begreifen, vielversprechend ist, da er sich auf nahezu alle Texte des Korpus anwenden ließ. Wenn ein System entwickelt würde, das auf Basis dieses Konzeptes versucht, aus natürlichsprachlichen Texten automatisch Ontologien zu erzeugen, würden weitere Herausforderungen zu bewältigen sein. Es ist bspw. zu erwarten, dass die Anzahl der semantischen Rollen reduziert werden müsste, da die hier dargestellte Menge sehr differenziert ist. Dennoch wäre die Entwicklung eines solchen Systems ein lohnenswertes Unterfangen.

9 Pertinenzrelationen

Die Pertinenzrelationen nehmen eine besondere Stellung ein, da sie nicht rein mengentheoretischer Natur sind, sondern reale Gegebenheiten ausdrücken (vgl. Kapitel 3.2.2).

Der Vergleich zwischen DK und WDK (Abbildung 23) zeigt, dass die Teil-Ganzes-Relation in Wikipedia eher vertreten ist, während im Duden-Korpus *Possessor-Possession* häufiger vertreten ist. In beiden Korpora ist die Teil-

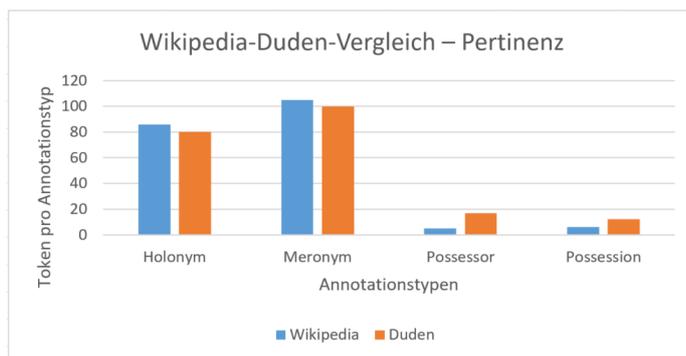


Abbildung 23: Der Vergleich zwischen Wikipedia und Duden: Pertinenzrelationen

Ganzes-Relation dabei viel häufiger vertreten als *Possessor-Possession*.

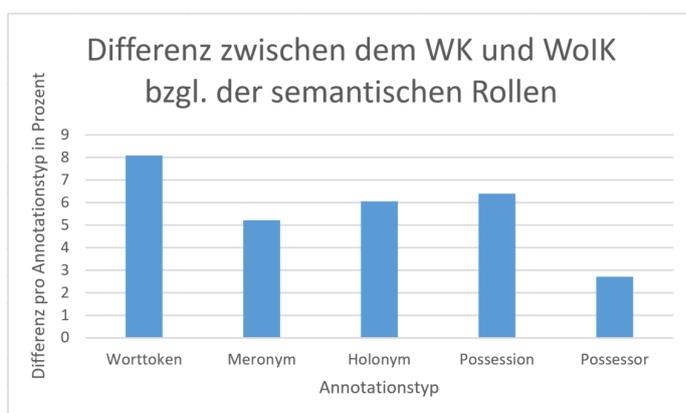


Abbildung 24: Der Vergleich zwischen WK und WoIK: Pertinenz-Annotationen

Der Vergleich zwischen WK und WoIK (Abbildung 24) bringt wenig Erklärungsbedürftiges zutage. Bei keinem Annotationstypus liegt die Differenz über der Worttokendifferenz.

9.1 Meronym – Holonym

Meronymie bezeichnet die Relation zwischen einem Teil, dem Meronym, und dem Ganzen, dem es angehört, dem Holonym (vgl. Kapitel 3.2.2). Das Inverse dieser Relation wird Holonymie und beide Relationen zusammen werden Teil-Ganzes-Relation genannt.

Die Statistik zeigt, dass die beiden Annotationstypen und damit die Teil-Ganzes-Relation auf alle Wissensgebiete, abgesehen von Materie und Energie, wo die Relation selten vertreten ist, gleichmäßig verteilt ist. Die Stan-

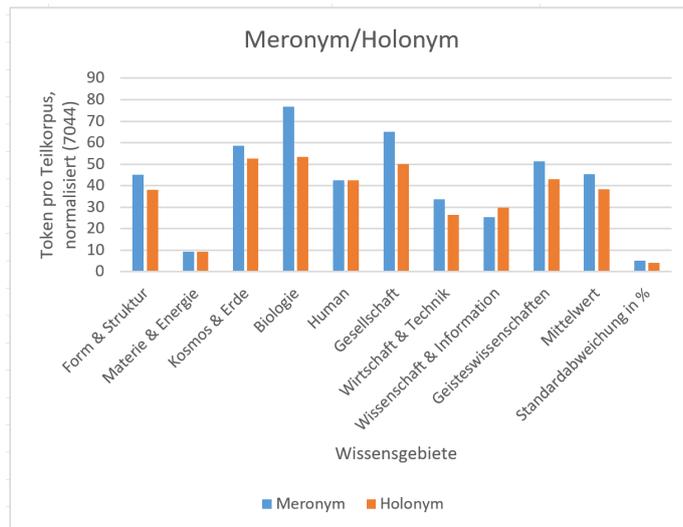


Abbildung 25: Meronym und Holonym in den verschiedenen Wissensgebieten

Standardabweichung ist dementsprechend bei beiden Typen mit unter fünf Prozent gering. Außerdem zeigt sich, dass Meronyme gegenüber Holonymen leicht in der Überzahl sind.

Für die mengentheoretischen Relationen hält $SHOIN(\mathcal{D})$ eigene Zeichen vor: \sqsubseteq , \sqsupset etc.; für andere Relationen, wie die Teil-Ganzes-Relation, müssen selber ontologische Rolle definiert werden. Die ontologischen Rollen sollen alle auf Englisch benannt werden, z.B. *isPartOf* für die ontologische Rolle, die Meronymie vertritt, und *consistsOf* für die ontologische Rolle, die Holonymie vertritt.

Es existieren mehrere Verben, die Holonymie ausdrücken; da im Duden meistens keine vollständigen Sätze formuliert werden, wird das Verb hier häufig als Partizip realisiert:

- (174)
- a. **Sand** ist ein natürlich vorkommendes, unverfestigtes Sediment, das sich überwiegend aus **Mineralkörnern** mit einer Korngröße von 0,063 bis 2 Millimeter zusammensetzt. (Wikipedia, Sand)
 - b. Allgemeiner besteht **das kartesische Produkt** mehrerer Mengen aus der Menge aller **Tupel von Elementen** der Mengen. (Wikipedia, Kartesisches Produkt)
 - c. **Die Hauptreihe** wird in der Astronomie durch **die Sterne** gebildet, die ihre Strahlungsenergie durch Wasserstoffbrennen im Kern freisetzen. (Wikipedia, Hauptreihe)
 - d. **Polynom**: Aus mehr als zwei **Gliedern** bestehender, durch Plus-

od. *Minuszeichen verbundener mathematischer Ausdruck*. (Der Duden, Das Bedeutungswörterbuch: 888)

- e. **Eine zweistellige Verknüpfung**, auch binäre Verknüpfung genannt, ist in der Mathematik eine Verknüpfung, die genau **zwei Operanden** besitzt. (Wikipedia, Zweistellige Verknüpfung)

Ein große Herausforderung bei der Formalisierung der nicht mengentheoretischen Relationen besteht darin, die entsprechenden Aussagen mit dem richtigen Quantor zu belegen (vgl. Kapitel 3.4).

Für (174a) genügt der Existenzquantor, da das Wort *überwiegend* anzeigt, dass Mineralkörner nicht der einzige Bestandteil von Sand sind.²⁰ Eine Formalisierung in $\mathcal{SHOIN}(\mathcal{D})$ sähe folgendermaßen aus:²¹

$$\text{Sand} \sqsubseteq \exists \text{consistsOf} . \text{Mineralkorn}$$

Sand ist Unterklasse der Dinge, die aus Mineralkörnern bestehen (und evtl. aus anderen Dingen). Die Nuance, dass Sand zwar nicht ausschließlich, aber zu einem sehr großen Anteil aus Mineralkörnern besteht, die genau genommen in *überwiegend* steckt, kann in $\mathcal{SHOIN}(\mathcal{D})$ nicht ausgedrückt werden.

Beispiele (174b) und (174c) müssten z.B. mit dem Allquantor formalisiert werden, da das kartesische Produkt ausschließlich aus Tupeln und nicht aus anderen Dingen und die Hauptreihe nur aus Sternen mit einer bestimmten Eigenschaft besteht:²²

$$\begin{aligned} \text{KartesischesProdukt} &\sqsubseteq (\forall \text{consistsOf} . \text{Tupel} \sqcap \exists \text{consistsOf} . \text{Tupel}) \\ \text{Hauptreihe} &\sqsubseteq (\forall \text{consistsOf} . \text{SternDWIKV} \sqcap \exists \text{consistsOf} . \text{SternDWIKV}) \end{aligned}$$

Es sei hier noch einmal wiederholt: Ein kartesisches Produkt ist eine Untermenge der Dinge, die, wenn sie aus etwas bestehen, dann ist dieses etwas ein Tupel – und sie bestehen aus Tupeln. Der letzte Teilsatz ist notwendig, da die Aussage sonst offenlassen würde, ob das kartesische Produkt überhaupt aus etwas besteht. Etwas natürlich-sprachlicher könnte man zusammenfassen: Ein kartesisches Produkt besteht ausschließlich aus Tupeln.

²⁰In dem Sinne meint *consistsOf* keine vollständige Aufzählung an Bestandteilen.

²¹Dass die Mineralkörner eine bestimmte Größe haben, soll hier aus Gründen der Anschaulichkeit nicht formalisiert werden.

²²Diese Eigenschaft, Wasserstoff im Kern zu verbrennen, wird in der Formalisierung aus Gründen der Lesbarkeit DWIKV abgekürzt. Diese spezielle Klasse von Sternen muss natürlich an anderer Stelle definiert sein, damit die Formeln so formuliert werden können.

Beispiel (174d) nennt explizit eine Zahl, die übertroffen werden muss, und müsste daher mit einem numeralen Quantor formalisiert werden:

$Polynom \sqsubseteq_{\geq 3} consistsOf.Glied$

Ein Polynom ist Unterklasse der Dinge, die mindestens dreimal (größer zwei) über die Relation *consistsOf* mit Instanzen der Klasse *Glied* verbunden sind.

Beispiel (174e) nennt hingegen eine Zahl, die genau getroffen werden muss. Dies fordert die Kombination von zwei numeralen Quantoren:

$ZweistelligeVerknuepfung \sqsubseteq$
 $(\geq 2 consistsOf.Operand \sqcap \leq 2 consistsOf.Operand)$

Eine zweistellige Verknüpfung besteht aus größer-gleich und kleiner-gleich zwei Operanden, also aus effektiv zwei Operanden. Bemerkenswert an diesem Beispiel ist außerdem, dass die Relation mit dem Verb *besitzen* realisiert wird. Dieses Verb ist eigentlich ein Indiz für eine andere Relation, die im folgenden Unterkapitel besprochen werden soll: *Possessor-Possession*. Ohne diese Relation vorwegzunehmen, kann gesagt werden, dass in diesem Beispiel trotz des Verbs Holonymie vorliegt, was im Vergleich mit (174d) deutlich wird: Beide Texte behandeln mathematische Ausdrücke und woraus diese bestehen; warum sollten hier zwei verschiedene Relationen vorliegen? Das Beispiel zeigt dementsprechend erneut, dass keine 1:1-Zuordnung zwischen sprachlichem Muster und semantischem Zusammenhang existiert.

Die Belege (174) beinhalten verschiedene Indizien für Quantoren. Die alleinige Verwendung des Existenzquantors wird durch ein Signalwort, *überwiegend*, nahegelegt, während die Verwendung des numerischen Quantors durch die explizite Nennung der Zahl eindeutig ist. Die Verwendung des Allquantors in Kombination mit dem Existenzquantor, wie in den Formalisierungen von (174b) und (174c), wird nicht markiert: Dadurch, dass keine andere Relation genannt oder angedeutet wird, könnte geschlossen werden, dass in diesen Fällen der Allquantor zu verwenden sei.

Häufig wird die Teil-Ganzes-Relation nicht durch ein Verb ausgedrückt, sondern durch ein Attribut und seinen Kopf. Solche Fälle sind bzgl. des Typs

der Relation weniger eindeutig. Die folgenden Beispiele enthalten jeweils ein PP-Attribut:

- (175) a. **Viereck**: *Von vier geraden Linien begrenzte Fläche mit vier Ecken.* (Der Duden, Das Bedeutungswörterbuch: 1083)
b. **Winkel**: *Geometrisches Gebilde aus zwei Geraden, die von einem Punkt ausgehen.* (Der Duden, Das Bedeutungswörterbuch: 1130)
c. Als **Spross** (Plural *Sprosse* oder *Sprossen*) bezeichnet man den aus **Sprossachse** und **Blättern** bestehenden, meist oberirdisch wachsenden, Teil **des Kormus** (Organisationsform der Gefäßpflanzen). (Wikipedia, Spross)

Bei den Beispielen (175) handelt es sich um Ellipsen, bei denen das Verb *bestehen* ausgelassen wird – dies wird im Vergleich zu Beispiel (175c) deutlich, bei dem *bestehen* noch als Partizip an die PP angehängt ist.

Die Möglichkeit, Meronyme als Genitivattribute zu realisieren, ist ein Stück weniger eindeutig, wozu passt, dass in (176) das Meronym *Innenwinkel* so formuliert ist, als sei die Tatsache, dass ein Rechteck Innenwinkel hat, schon bekannt:

- (176) *In der Geometrie ist ein **Rechteck** (ein Orthogon) ein ebenes Viereck, dessen **Innenwinkel** alle rechte Winkel sind.* (Wikipedia, Rechteck)

Am wenigsten eindeutig, gewissermaßen versteckt, sind Meronyme in adjektivischen Attributen. Im folgenden Beispiel stecken in den Adjektiven *gleichseitig* und *gleichwinklig* die Meronyme *Seite* und *Winkel*:

- (177) **Ein regelmäßiges Polygon**, reguläres Polygon, regelmäßiges Vieleck, reguläres Vieleck oder Isogon [...] ist in der Geometrie ein ebenes Polygon, das sowohl **gleichseitig** als auch **gleichwinklig** ist. (Wikipedia, Regelmäßiges Polygon)

Solche Meronyme sind evtl. nicht sofort ersichtlich, da Adjektive seltener als Klassen gedacht werden. Formal lassen sich Adjektive jedoch als die Klasse der Dinge begreifen, auf die das Adjektiv zutrifft. Daher lässt sich Beispiel (177) auf mindestens zwei Arten formalisieren. Zunächst lässt sich die Holonymie ausdrücken, indem die Klassen *GleicheSeite* und *GleicherWinkel* genutzt werden, die an anderer Stelle in der Ontologie definiert sind:

RegelmaessigesPolygon \sqsubseteq
 (*Polygon* \sqcap \exists *consistsOf.GleicheSeite* \sqcap \exists *consistsOf.GleicherWinkel*)

Eine andere Möglichkeit wäre aber, die Adjektive als Klassen zu formalisieren:

RegelmaessigesPolygon \sqsubseteq (*Polygon* \sqcap *GleichSeitig* \sqcap *GleichWinklig*)

Dies wirft die Frage auf, ob nicht jede Relation, unabhängig davon, ob sie ursprünglich adjektivisch oder nicht formuliert ist, einfach als Hyponymie zu formalisieren wäre. So könnte bspw. für (174a) auch einfach die Klasse Mineralkörnig angesetzt werden, um dann folgende Formalisierung vorzunehmen:

Sand \sqsubseteq *Mineralkoernig*

Der fundamentale Unterschied bei einer solchen, prinzipiell möglichen, Formalisierung wäre, dass die Option der Quantifizierung verloren ginge. Es ist nicht mehr möglich, auszudrücken, ob Sand ausschließlich oder nur u.a. aus Mineralkörnern besteht. Es muss allerdings von Fall zu Fall unterschieden werden, ob von dieser Möglichkeit Gebrauch gemacht werden soll.

Auch in den folgenden Belegen wird die Teil-Ganzes-Relation in dem Sinn implizit realisiert, als die Meronyme selber präsupponiert werden, indem ihre Eigenschaften thematisiert werden:

- (178) a. **Ein gleichseitiges Polygon** ist in der Geometrie ein Polygon der euklidischen Ebene, bei dem **alle Seiten** gleich lang sind. (Wikipedia, Gleichseitiges Polygon)
- b. **Ein Parallelogramm** [...] oder Rhomboid (rautenähnlich) ist ein konvexes ebenes Viereck, bei dem **gegenüberliegende Seiten** parallel sind. (Wikipedia, Parallelogramm)

Der Grund für diese Präsupposition liegt darin, dass das jeweilige Meronym auch schon zur Oberklasse gehört. Gewissermaßen ist die *differentia specifica* dieser Klassen eine Spezifizierung der Meronyme. Eine Formalisierung wird diesen Texten am ehesten gerecht, wenn wieder Klassen verwendet werden, die an anderer Stelle definiert sind. Bspw.:

GleichseitigesPolygon \sqsubseteq (\exists *consistsOf.GleicheSeite* \sqcap *Polygon*)

In dem folgenden Beispiel bedarf es Weltwissen, um die Meronyme als solche

zu erkennen, da sie im Text nur als Merkmale bezeichnet werden, wodurch die Mengen der möglichen Relationen kaum eingegrenzt wird.

- (179) a. **Die Rosen** (*Rosa*) sind die namensgebende Pflanzengattung der Familie der Rosengewächse (*Rosaceae*). [...] Diese bilden durch ihre typischen Merkmale **Stacheln**, **Hagebutten** und **unpaarig gefiederte Blätter** eine sehr gut abgegrenzte Gattung. (Wikipedia, Rosen)

Bei der taxonomischen Relation sprach einiges dafür, es als den unmarkierten Fall zu betrachten, wenn die Unterklasse das Subjekt bildete. Bei der Teil-Ganzes-Beziehung lässt sich zumindest feststellen, dass das Holonym häufiger das Subjekt bildet als das Meronym. In den Beispielen (174) verhält es sich immer so, dass das Ganze durch seine Teile definiert ist, die Teile aber nicht durch ihre Zugehörigkeit zum Ganzen definiert werden. Das kartesische Produkt bspw. besteht aus Tupeln; Tupel selber können aber auch als Bestandteil anderer Dinge vorkommen. Mineralkörner einer bestimmten Größe werden nicht darüber definiert, dass sie ab einer bestimmten Anzahl Sand bilden. Bei (174c) ist der Fall nicht so offensichtlich. Die Hauptreihe scheint über die ihr angehörenden Sterne definiert, die Sterne lassen sich aber auch darüber definieren, dass sie der Hauptreihe angehören. Dennoch verhält es sich auch hier so, dass die Hauptreihe eine Form der Klassifikation ist, die sich aus ihren Mitgliedern zusammensetzt, ohne diese also nicht bestände, während die in Frage stehenden Sterne auch ohne die Klassifikation bestehen.

Es scheint, sich ein allgemeines Prinzip formulieren zu lassen: Während ein Ganzes sich häufig als eine spezifische Kombination von Einzelteilen definieren lässt, können die Einzelteile in verschiedenen Kombinationen vorkommen und lassen sich deswegen tendenziell nicht über ihre Eigenschaft als Bestandteil definieren. Es existieren allerdings systematische Ausnahmen: So gehört es zur Definition eines Fahrradreifens, dass er Teil eines Fahrrades ist oder zumindest als solches gedacht wird. Es lässt sich verallgemeinern, dass immer eine Unterklasse von Bestandteilen gebildet werden kann, deren *differentia specifica* die Zugehörigkeit zu einem bestimmten Holonym ist: Der

Fahrradreifen ist Teil des Fahrrads, der Autoreifen Teil des Autos, die Kellertür Teil des Kellers.

Doch werden für solche Begriffe beide Klassen, bspw. *Fahrradreifen* und *Fahrrad*, entlang der taxonomischen Relation nach oben verfolgt, verliert das Homonym an Bedeutung: Ein Reifen ist nicht notwendiger Bestandteil eines Vehikels; Schiffe bestehen z.B. nicht aus Reifen. Türen sind nicht notwendiger Bestandteil eines Gebäudes; es existieren auch Gartentüren, die dann bspw. in einem Zaun eingefügt sind.

Einen solchen Fall stellt Beispiel (180) dar. Das Beispiel definiert genau genommen nicht die Klasse Planet, wie der Titel des Artikels vorgibt, sondern die Klasse Planet des Sonnensystems. Der Begriff *Sonnensystem* ist also das Holonym.

(180) **Planet:** *(nicht selbst leuchtender) Himmelskörper im Sonnensystem.*
(Der Duden, Das Bedeutungswörterbuch: 742)

Der Artikel sagt aus, dass ein Planet ein Himmelskörper sei, der nur im Sonnensystem vorkomme, was letztlich nicht richtig ist, deswegen wird die Formalisierung der gesamten Definition vorgenommen:

$$\text{PlanetDesSonnensystems} \sqsubseteq (\text{NichtSelbstLeuchtenderHimmelskoerper} \sqcap \forall \text{isPartOf.Sonnensystem} \sqcap \exists \text{isPartOf.Sonnensystem})$$

Der folgende Beleg ist ebenfalls einer der wenigen Fälle, bei dem letztlich etwas als Bestandteil definiert wird.

(181) **Stern:** *Als silbrig-weißer, funkelnder Punkt besonders am nächtlichen Himmel sichtbares Gestirn.* (Der Duden, Das Bedeutungswörterbuch: 919)

Der Artikel formuliert die allgemeine Definition von Stern als ein Gestirn, das Teil des (sichtbaren) Nachthimmels ist. Diese Erklärung ist in sich konsistent, wenn man sie aber verwissenschaftlichen wollte, müsste man auch an dieser Stelle sagen, dass nur Sterne des Nachthimmels Teil dieses sind, während es auch Sterne gibt, vermutlich sogar die Mehrzahl, auf die dies nicht zutrifft.

Dass diese Definition nicht wissenschaftlich ist (was im Übrigen nicht der Anspruch einer Wortklärung sein muss), liegt letztlich daran, dass das Konzept des Himmels unbestimmt ist. Die Belege (180) und (181) zeigen dennoch, dass Dinge nur in spezifischen Fällen als Teil von etwas definiert werden.

Es steht zur Diskussion, ob die Teil-Ganzes-Relation transitiv ist, wofür zunächst vieles zu sprechen scheint: Wenn bspw. ein Dach Teil eines Hauses ist, und ein Ziegel Teil eines Daches, dann ist ein Ziegel Teil eines Hauses. Es gibt aber auch Fälle, bei denen die Transitivität befremdlich wirkt: Ein Dirigent ist Teil eines Orchesters, und ein Finger ist Teil des Dirigenten – also ist der Finger Teil des Orchesters? Die Schlussfolgerung mag nicht komplett abwegig sein, aber es scheinen doch zwei verschiedene Dinge gemeint zu sein, wenn vom Teil des Orchesters und vom Teil des Dirigenten gesprochen wird. Der Umstand deutet an, dass es sinnvoll ist, verschiedene Unterarten der Teil-Ganzes-Relation anzunehmen, welche dann für sich transitiv sind. In der Arbeit mit dem Korpus wurden sechs Untertypen von Meronymie identifiziert, die in Kapitel 3.2.2 bereits kurz vorgestellt wurden. Im Folgenden wird anhand der Belege untersucht, wie sich diese Typen unterscheiden und wie sie sich evtl. miteinander kombinieren lassen.

Der erste Typ ist Komponente-Objekt, wozu als Beispiel die Relation zwischen einem Ast und einem Baum dienen kann. Es handelt sich demnach um die Beziehung zwischen zwei Entitäten unterschiedlicher Klassen, die in einer festen Teil-Ganzes-Beziehung zueinander stehen. Dieser Typ darf als Prototyp der Teil-Ganzes-Relation verstanden werden. (175c) kann als ein konkretes und (174b) als ein abstraktes Beispiel aus dem Korpus dienen. Diese Relation ist transitiv: Wenn ein Blatt ein notwendiger Bestandteil von einem Ast ist, und ein Ast wiederum notwendiger Bestandteil eines Baumes, dann ist ein Blatt notwendiger Bestandteil eines Baumes.

Der zweite Typ ist Objekt-Ansammlung, womit bspw. die Relation zwischen einem Baum und einem Wald gemeint ist. Der Unterschied dieses Typen zu Komponente-Objekt besteht darin, dass die Meronyme in dieser Rela-

tion alle der gleichen Klasse angehören und das Holonym ab einer gewissen Anzahl darstellen. (174c) stellt ein passendes Beispiel aus dem Korpus dar.

Bzgl. der Transitivität ist dieser Typ bedingt kombinierbar mit dem Typ Komponente-Objekt. Dieser Sachverhalt verdient eine genauere Betrachtung. Wenn ein Ast Teil eines Baums ist (Komponente-Objekt) und ein Baum Teil des Waldes (Objekt-Ansammlung), warum ist dann ein Ast nicht Teil eines Waldes? Als Grund könnte angeführt werden, dass Äste sich zwar so gut wie immer in einem Wald finden lassen, weil sie Teil der Bäume sind, dass eine bestimmte Menge an Ästen aber keinen Wald ergeben, so wie dies eine bestimmte Anzahl von Bäumen tut. Die Frage, ob hier Transitivität anzusetzen ist, kann nicht endgültig beantwortet werden und hängt davon ab, wie streng die Relation in einem logischen Sinne aufgefasst wird. Rein kausal besteht eine Transitivität zwischen den Typen; für das Sprachgefühl ist dieses allerdings oft seltsam.

Die Teil-Ganzes-Relation vom Typ Objekt-Ansammlung involviert oft Instanzen, wobei Meronym und Holonym Instanzen sind. So ist Angela Merkel Teil der CDU (182a) oder verschiedene Länder sind Teil der Europäischen Union (182b).

- (182) a. **Angela Dorothea Merkel** (geb. Kasner; * 17. Juli 1954 in Hamburg) ist eine deutsche Politikerin (CDU). (Wikipedia, Angela Merkel)
- b. **Die Europäische Union (EU)** ist ein Staatenverbund aus 27 europäischen Ländern. Außerhalb des geographischen Europas umfasst die EU Zypern und einige Überseegebiete. Sie hat insgesamt etwa 450 Millionen Einwohner. (Wikipedia, Europäische Union)

Der dritte Typ ist Material-Material genannt werden. Wie die Bezeichnung schon andeutet, sind bei dieser Relation sowohl Meronym als auch Holonym unzählbare Begriffe. Ein Beispiel aus dem Korpus wäre (174a). Grundsätzlich ist dieser Typ transitiv: Wenn Sand aus Mineralkörnern besteht, und eine Sandburg nun einmal aus Sand, dann besteht eine Sandburg aus Mineralkörnern. Obwohl dieses Beispiel sehr eindeutig ist, entspricht es nicht ganz dem Sprachempfinden, zu sagen, dass eine Sandburg aus Mineralkörnern bestehe.

Es zeigt sich, dass Transitivität bei der Teil-Ganzes-Relation nicht so sehr im Sprachbewusstsein verankert ist, wie dies bei der taxonomischen Relation der Fall ist.

Der vierte Typ ist Material-Objekt, womit bspw. das Verhältnis zwischen Holz und einem Baum gemeint ist. Der Unterschied zu Komponente-Objekt liegt darin, dass eine Komponente zählbar ist, während Material einen unzählbaren Begriff meint. Ein Beispiel aus dem Korpus wäre die Wikipedia Definition von Stern:

- (183) a. *Unter **einem Stern** [...] versteht man in der Astronomie einen massereichen, selbstleuchtenden Himmelskörper aus **sehr heißem Gas und Plasma**, wie zum Beispiel die Sonne.* (Wikipedia, Stern)

Bzgl. der Transitivität ist dieser Typ bedingt mit Komponente-Objekt kombinierbar. Wenn ein Ziegelstein aus gebranntem Lehm besteht, und ein Haus aus Ziegelsteinen, dann besteht ein Haus zum Teil aus gebranntem Lehm. Dies mag sprachlich etwas ungewöhnlich wirken, ist aber sachlich und logisch richtig. Auch mit dem Typ Objekt-Ansammlung scheint eine bedingte Kombinierbarkeit zu bestehen. Ein Wald besteht zum Teil aus Holz, weil ein Baum aus Holz besteht.

Der fünfte Typ heißt Ort-Gebiet und beschreibt das Verhältnis von einem Objekt, das als Ort interpretiert wird und in einem anderen Ort lokalisiert werden kann. Ein Beispiel aus dem Korpus wäre (181), in welchem ein Stern als etwas definiert wird, das im Sonnensystem zu verorten ist. Auch dieser Typ scheint mit den anderen nur bedingt kombinierbar.

Im folgenden Beleg werden zwei Meronyme miteinander kombiniert, wobei eines vom Typ Material-Objekt und eines sowohl vom Typ Ort-Gebiet als auch Komponente-Objekt ist:

- (184) ***Eine Platte** ist in der Technischen Mechanik bzw. in der Bautechnik ein in **der Ebene** ausgebreitetes Bauteil, das aus **steifem Material** besteht (ebenes Flächentragwerk) und durch senkrecht auf sie wirkende Kräfte und durch Momente um Achsen, die in der Plattenebene liegen, belastet wird.* (Wikipedia, Platte)

Es lässt sich zusammenfassen, dass eine Platte eine Komponente ist, die sich in der Ebene befindet und aus steifem Material besteht. Dieses Beispiel deutet eine bedingte Transitivität an. Es könnte bspw. gesagt werden, dass steifes Material sich in der Ebene in Form einer Platte befindet.

Der Typ Ort-Gebiet besteht ebenfalls häufig zwischen Instanzen:

- (185) *Deutschland [...] ist ein Bundesstaat in Mitteleuropa.* (Wikipedia, Deutschland)

Als sechster und letzter Typ soll Teilprozess-Prozess behandelt werden, bei dem der Name schon verrät, dass es sich bei Meronym und Holonym um Prozesse handelt, wobei das Meronym einen Teilprozess darstellt. Bspw. ist Bezahlen ein Teil von Einkaufen. Ein passendes Beispiel aus dem Korpus ist folgendes:

- (186) *Investigativer Journalismus (von lateinisch investigare ‚aufspüren‘, ‚genauestens untersuchen‘) setzt eine langwierige, genaue und umfassende Recherche vor Veröffentlichung voraus.* (Wikipedia, Investigativer Journalismus)

Im Korpus ließen sich viele randständige Beispiele für diesen Typ finden. Das Verhältnis zwischen Krankheiten und ihren Symptomen könnte als Beispiel herangezogen werden. Wenn eine Krankheit als ein Prozess angesehen wird, was durchaus plausibel ist, dann können die Symptome als Teilprozesse angesehen werden, die evtl. parallel ablaufen:

- (187) a. *Psychose ist ein Grundbegriff aus der Psychiatrie. Früher stand er für alle Arten von psychischen Erkrankungen. Heute bezeichnet der Begriff einen unscharf definierten Symptomenkomplex (Syndrom), der durch Halluzinationen, Wahn, Realitätsverlust oder Ich-Störungen gekennzeichnet ist.* (Wikipedia, Psychose)
- b. *Die zwanghafte Persönlichkeitsstörung oder anankastische Persönlichkeitsstörung [...] auch Zwangspersönlichkeitsstörung, gehört zur Gruppe der Persönlichkeitsstörungen (Cluster C). Typisch für sie sind Rigidität, Perfektionismus, ständige Kontrollen, Gefühle von Zweifel sowie ängstliche Vorsicht, keine Fehler zu machen.* (Wikipedia, Zwanghafte Persönlichkeitsstörung)

Ob es sich bei diesen Fällen aber wirklich um Teilprozess-Prozess handelt, mag diskutiert werden. Im Korpus wurden die Teil-Ganzes-Annotationen ohnehin nicht bzgl. des Typen differenziert, weil zu viele Zweifelsfälle existieren. Teilprozess-Prozess ist mit den anderen nicht kombinierbar, da es sich bei allen Elementen in der Relation um einen Prozess handeln muss.

Der Hauptgrund, die Teil-Ganzes-Relation in Untertypen aufzuteilen, lag darin, das Verhalten der einzelnen Typen bzgl. der Transitivität zu untersuchen, wenn sie miteinander kombiniert werden. Dabei ergab sich aber kein genaues Ergebnis. Aus einer logischen Perspektive erschienen alle Typen abgesehen von Teilprozess-Prozess unter Beibehaltung der Transitivität miteinander kombinierbar. Dem Sprachempfinden standen aber viele Beispiele entgegen.

So wie es möglich ist, eine ganze Reihe von Unterklassen aufzuzählen, ist es auch möglich, eine ganze Reihe von Meronymen aufzuzählen:

- (188) ***Eine Galaxie** ist eine durch Gravitation gebundene große Ansammlung von **Sternen, Planetensystemen, Gasnebeln, Staubwolken, Dunkler Materie** und sonstigen astronomischen Objekten mit einer Gesamtmasse von typischerweise 109 bis 1013 Sonnenmassen (M_{\odot}).* (Wikipedia, Galaxie)

Beispiel (188) zeigt, dass die Meronyme dabei verschiedenen Typen angehören können, so ist ein Stern eine Komponente und Dunkle Materie eben Materie.

Auch in den folgenden Belegen wird mehr als ein Meronym realisiert, wobei die Verhältnisse teilweise verwoben sind:

- (189) a. *Im mathematischen Teilgebiet der linearen Algebra bezeichnet **Skalar** ein Element des **Grundkörpers eines Vektorraumes**, meist also eine reelle Zahl.* (Wikipedia, Skalar)
- b. ***Der interstellare Staub** ist derjenige Anteil am **kosmischen Staub**, der Teil **der interstellaren Materie** ist.* (Wikipedia, Interstellarer Staub)

Beispiel (189a) lässt sich folgendermaßen zusammenfassen: Skalar ist Teil

von Grundkörper ist Teil von Vektorraum.²³ Eine solche Kette von Teil-Ganzes-Relationen lässt sich so in $\mathcal{SHOIN}(\mathcal{D})$ nicht ausdrücken, da innerhalb von Rollen nur bereits existierende Klassen oder anonyme Klassen verwendet werden können. Es ist nicht möglich, eine Klasse innerhalb einer Rolle zu verwenden und zu definieren. Es besteht also die Möglichkeit, mit einer anonymen Klasse zu arbeiten oder den Text in zwei Aussagen aufzuteilen:

$$\text{Skalar} \sqsubseteq \exists \text{isPartOf} . (\exists \text{isPartOf} . \text{Vektorraum})$$

Diese Formalisierung anonymisiert Grundkörper, indem sie aussagt, dass ein Skalar Teil einer Sache ist, die wiederum Teil des Vektorraumes ist. Ein erheblicher Nachteil dieser Formel besteht darin, dass sie Grundkörper nicht nur anonymisiert, sondern auch unbestimmt lässt. Es existieren evtl. viele Dinge, die Teil des Vektorraumes sind, welchem gehört ein Skalar an? Deswegen ist die Aufteilung in zwei Aussagen wohl vorzuziehen, da sie es ermöglicht, Grundkörper an anderer Stelle ausführlich zu definieren:

$$\begin{aligned} \text{Skalar} &\sqsubseteq \exists \text{isPartOf} . \text{Grundkoerper} \\ \text{Grundkoerper} &\sqsubseteq \exists \text{isPartOf} . \text{Vektorraum} \end{aligned}$$

(189b) ist ein bemerkenswerter Beleg dafür, wie viel Information in einem natürlich-sprachlichen Satz stecken kann. Der Satz müsste in zwei Aussagen formalisiert werden:

$$\begin{aligned} \text{InterstellarerStaub} &\sqsubseteq \\ &(\exists \text{isPartOf} . \text{KosmischerStaub} \sqcap \exists \text{isPartOf} . \text{InterstellarerMaterie}) \end{aligned}$$

$$\text{InterstellareMaterie} \sqsubseteq \exists \text{consistsOf} . \text{KosmischerStaub}$$

Der erste Teil sagt aus, dass interstellarer Staub Teil der Schnittmenge des kosmischen Staubs und Teil der interstellaren Materie ist; der zweite Teil spezifiziert, dass interstellare Materie zum Teil aus kosmischen Staub besteht. Dass so viel Information in diesem Text steckt, liegt letztlich an dem Demonstrativpronomen *derjenige*, dadurch ist möglich zu spezifizieren, dass nicht nur interstellarer Staub Teil von kosmischem Staub und damit Teil von interstella-

²³Dies ist mathematisch nicht ganz korrekt, sprachlich jedoch die naheliegendste Interpretation.

rer Materie ist, sondern dass es sich um einen speziellen Teil des kosmischen Staubes handelt, und interstellare Materie auch aus anderen Dingen besteht.

Die Grenze zwischen der Teil-Ganzes-Relation und der taxonomischen Relation kann verschwimmen.

- (190) a. *Wasserschall ist Schall, der im Wasser übertragen wird. Das zugehörige Fachgebiet **der Akustik** ist **die Hydroakustik**.* (Wikipedia, Wasserschall)
- b. ***Controlling** ist ein Sammelbegriff, der **Planung, Koordination und Kontrolle** des Unternehmens einschließt.* (Wikipedia, Controlling)

Inhaltlich ist eindeutig, dass Hydroakustik in (190a) eine Unterklasse von Akustik ist, aber das Wort *zugehörig* legt ein räumliches Verhältnis nahe. In (190b) ist nicht eindeutig, ob es sich bei Planung, Koordination und Kontrolle um Meronyme oder Unterklassen handelt. Ist Planung ein Teil von oder ein Art von Controlling? Das Wort *einschließen* bspw. legt erneut ein räumliches Verhältnis nahe.

Die Beispiele belegen eine grundsätzliche Tendenz, die mengentheoretischen Relationen, allen voran die taxonomische Relation, räumlich zu denken, wodurch die taxonomische Relation und die Teil-Ganzes-Relation vermischt werden können. Diese Tendenz zeigt sich schon bei Venndiagrammen, die oft zur Illustration von Mengen und ihrer Relationen herangezogen werden (vgl. Partee 1996: 13-20); sie bedienen sich einer räumlichen Darstellung, die genau genommen nicht das Wesen dieser formalen Relationen beschreibt.

In (191) kommt diese Tendenz besonders stark zu Vorschein.

- (191) ***Ordnung**: Größere Einheit, die aus mehreren **verwandten Tier- oder Pflanzenfamilien** besteht.* (Der Duden, Das Bedeutungswörterbuch: 723)

Das Verb *bestehen* zeichnet die Relation eigentlich eindeutig als Meronymie aus; es handelt sich aber um die Metaklasse Ordnung und ihre Instanzen, die Familien. In Kapitel 7.5 wurde bereits ausgeführt, dass sich die Relation zwischen einer Instanz und einer Metaklasse dem sprachlichen Bewusstsein ent-

zieht. Auf diese Weise lässt sich erklären, dass sich die Formulierung an der Teil-Ganzes-Relation orientiert.

Ein besonderer Fall, in dem sich die taxonomische Relation und die Teil-Ganzes-Relation vermischen, liegt in (192) vor:

(192) **Bakterie**: *Aus nur **einer Zelle** bestehender, kleinster Organismus, der Fäulnis, Krankheit, Gärung hervorrufen kann.* (Der Duden, Das Bedeutungswörterbuch: 176)

Die Vermischung liegt in dem Sinne vor, dass, wenn ein Bakterium nur aus einer Zelle besteht, gesagt werden könnte, dass ein Bakterium eine Zelle ist. In diesem Fall sind die beiden Relationen äquivalent.

In den Geisteswissenschaften existieren viele Zusammenhänge, die man als Meronymie interpretieren kann, bei denen jedoch auch andere Interpretationen zulässig wären. Dies liegt mitunter daran, dass in den Geisteswissenschaften viele Themen behandelt werden, bei denen keine echte Materialität vorliegt. Wenn in solchen Fällen etwas als Teil von etwas anderen definiert wird, liegt von vornherein eine gewisse Metaphorik vor.

(193) **Komödie**: *Bühnenstück mit **heiterem Inhalt** (in dem menschliche Schwächen dargestellt und Konflikte heiter gelöst werden.* (Der Duden, Das Bedeutungswörterbuch: 584)

In (193) kann *heiterer Inhalt* als Meronym interpretiert werden. In dieser Arbeit wurde aber auch die semantische Rolle *Content* definiert, die an dieser Stelle auch angesetzt werden könnte. Letztlich muss in solchen Fällen eine redaktionelle Entscheidung getroffen werden.

Das Unterkapitel soll mit einer Betrachtung der Fälle abgeschlossen werden, bei denen die Relation in irgendeiner Form negiert wird, bei denen also etwas explizit nicht als Teil von etwas anderem definiert wird. Dies passiert äußerst selten; die hier besprochenen Fälle sind alle, die im Korpus zu finden waren:

(194) a. **Ein Exoplanet**, *auch extrasolarer Planet, ist ein planetarer Himmelskörper außerhalb des **vorherrschenden gravitativen Einflusses der Sonne**, aber innerhalb des gravitativen Einflusses eines*

anderen Sterns oder Braunen Zwergs. Extrasolare Planeten gehören also nicht dem Sonnensystem, sondern anderen Planetensystemen an. (Wikipedia, Exoplanet)

- b. **Galaxie:** *Großes Sternensystem außerhalb der Milchstraße.* (Der Duden, Das Fremdwörterbuch: 386)
- c. **Pilz:** *Blatt- und blütenlose Pflanze, meist aus fleischigem Stiel und unterschiedlich geformten Hut bestehend.* (Der Duden, Das Bedeutungswörterbuch: 740)

In (194a) wird bspw. ein Exoplanet als ein Planet definiert, der sich im Einfluss eines anderen Sterns als der Sonne befindet, der also nicht Teil des Sonnensystems ist. Die Aussage ließe sich auf zwei Arten formalisieren:

$$\begin{aligned} \text{Exoplanet} &\sqsubseteq \neg \exists \text{isPartOf.Sonnensystem} \\ \text{Exoplanet} &\sqsubseteq \exists \text{isPartOf}.\neg \text{Sonnensystem} \end{aligned}$$

Entweder wird ein Exoplanet als ein Planet definiert, der zu den Dingen gehört, die nicht Teil des Sonnensystems sind, oder als ein Planet, der Teil des Nicht-Sonnensystems, also allem was nicht zum Sonnensystem gehört, ist. Die erste Variante kommt der natürlich-sprachlichen Intuition wahrscheinlich näher; beide Formalisierungen sind allerdings äquivalent. Beleg (194b) funktioniert analog zu (194a). Interessanterweise stellt das Holonym in beiden Fällen eine Instanz dar.

9.2 Possessor – Possession

Possessor-Possession bezeichnet die Relation zwischen der Entität, die etwas besitzt, und dem, was sie besitzt. In vielen Fällen ähnelt sie der Teil-Ganzes-Relation, und es existiert ein Grenzbereich, in dem die Relationen kaum auseinander gehalten werden können. Besitzt bspw. ein Mensch einen Arm oder ist der Arm Teil eines Menschen?

Die Verteilung der entsprechenden Annotationstypen über die Teilkorpora zeigt, dass eine deutliche Häufung bei den nicht-naturwissenschaftlichen Wissensgebieten besteht. In Form und Struktur sowie Biologie wurden gar keine Fälle annotiert und in Kosmos und Erde selbst normalisiert weniger als fünf. Einziger Ausreißer aus diesem Trend ist Materie und Energie, der sich durch

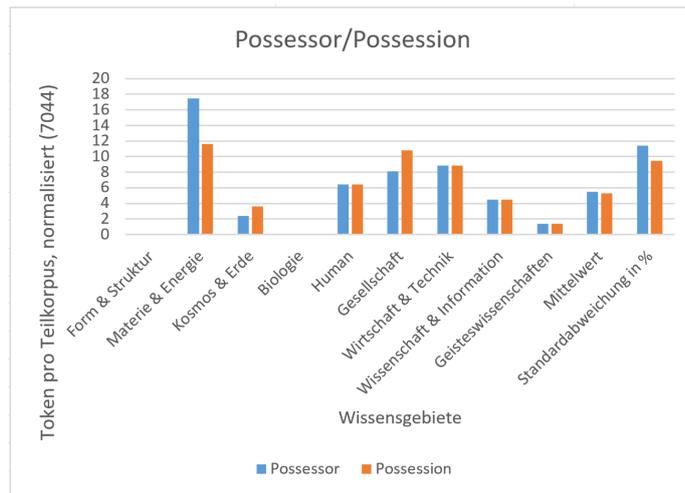


Abbildung 26: Die *Possessor-Possession*-Relation in den Teilkorpora

Fälle wie (200) erklärt. Diese sollen an entsprechender Stelle besprochen werden. Die Standardabweichung ist bei beiden Typen mit um die zehn Prozent relativ hoch und bedeutend höher als die von Meronym und Holonym.

Die prototypischen Fälle dieser Relation sind jene, in denen etwas explizit als Besitz deklariert wird, wie in (195) als PP mit dem Kopf *Eigentum*:

- (195) **Öffentliche Unternehmen** (*auch Staatsunternehmen genannt*) sind öffentlich-rechtliche Unternehmen oder privat-rechtlich organisierte Unternehmen im mehrheitlichen oder vollen Eigentum **des Staates** oder **seiner Untergliederungen**. (Wikipedia, Öffentliches Unternehmen)

Dieses Beispiel bietet sich an, um die Grenzen zur Teil-Ganzes-Definition ausführlicher zu erörtern. Dazu zwei Beispiele:

- (196) a. *Öffentliche Unternehmen sind Teil des Staates.*
 b. *Öffentliche Unternehmen sind Teil des Eigentum des Staates.*

Beispiel (196a) ist zwar etwas unspezifisch, grundsätzlich scheint dies aber eine zulässige Aussage zu sein, was zeigt, wie nah *Possessor-Possession* und Teil-Ganzes einander sind. Beispiel (196b) vermischt die prototypischen Formulierungen der jeweiligen Relation, wodurch ebenfalls kein inhaltlicher Widerspruch oder anderweitig zu beanstandender Inhalt entsteht. Diese Beobachtungen legen nahe, dass es sich evtl. um zwei Untertypen einer Relation handelt.

Das folgende Beispiel verdeutlicht jedoch, dass *Possessor-Possession* Aspekte abdecken kann, die durch Teil-Ganzes nicht abgedeckt werden:

- (197) **Ein Einparteiensystem** beziehungsweise eigentlich Einparteiensystem ist ein politisches System, bei dem **die alleinige Herrschaft einer Partei gesetzlich festgeschrieben ist.** (Wikipedia, Einparteiensystem)

Die Possession in diesem Beispiel ist *die alleinige Herrschaft*, während eine Partei der Possessor ist. Die folgenden Beispiele verdeutlichen, dass dieses Verhältnis offensichtlich nicht als Teil-Ganzes verstanden wird:

- (198) a. *Bei einem Einparteiensystem besitzt die Partei die alleinige Herrschaft.*
b. *?Bei einem Einparteiensystem ist die alleinige Herrschaft Teil der Partei.*
c. *?Bei einem Einparteiensystem ist der Staat Teil der Partei.*
d. *Bei einem Einparteiensystem übt eine Partei die alleinige Herrschaft aus.*

Beispiel (198a) soll zunächst zeigen, dass es sich bei diesem Sachverhalt tatsächlich um *Possessor-Possession* handelt; dies scheint der Fall zu sein, da eine Formulierung mit *besitzen* keinerlei Anstoß erregt. Beispiel (198b) ist dagegen sehr fragwürdig; die Aussage scheint nicht richtig zu sein. Bei Beispiel (198c) wurde der Inhalt anders aufgearbeitet: Dass die Partei die alleinige Herrschaft hat, bedeutet effektiv, dass sie den Staat regiert. Dieses Beispiel wurde gewählt, um auszuschließen, dass (198b) nur inakzeptabel erscheint, weil Herrschaft ein Abstraktum darstellt. Auch dieses Beispiel scheint nicht akzeptabel. Ein Grund, warum die Teil-Ganzes-Relation diese Fälle nicht zu fassen scheint, liegt evtl. im agentivischen Aspekt der Possessor eigen ist. Dies deutet sich dadurch an, dass die Formulierung in (198d) wiederum vollkommen akzeptabel ist.

Die ontologischen Rollen sollen *possesses* und *possessedBy* heißen und sind zueinander invers. Das Inverse, also die Umkehrung der Tupel, drückt in diesem Fall das Passiv aus. Die Aussage in (195) könnte also folgendermaßen formalisiert werden:

OeffentlichesUnternehmen $\sqsubseteq \exists$ *possessedBy.Staat*

Fähigkeiten und Wissen, welche Entitäten besitzen, fällt ebenfalls unter die Relation *Possessor-Possession*.

- (199) a. **Mensch**: *Mit Vernunft und Sprache* ausgestattetes höchstentwickeltes Lebewesen. (Der Duden, Das Bedeutungswörterbuch: 662)
- b. **Wissen**: *Gesamtheit der Kenntnisse, die jmd. [auf einem bestimmten Gebiet] hat.* (Der Duden, Das Bedeutungswörterbuch: 1133)
- c. **Juristische Personen** *des öffentlichen Rechts sind Rechtssubjekte, die auf öffentlich-rechtlichem und privatrechtlichem Gebiet Rechtsfähigkeit kraft Gesetzes besitzen.* (Wikipedia, Juristische Person des öffentlichen Rechts)

Auch hierbei stellt die als *Possession* annotierte Entität etwas Abstraktes dar, wobei in diesem Fall aber deutlich ist, dass über Wissen und Fähigkeiten verfügt werden kann, wie es für Besitz charakteristisch ist. Bei diesen Belegen zeigen sich verschiedene sprachliche Muster, die diese Relation realisieren: In (199a) bildet *ausgestattet* eine um eine PP erweiterte Adjektiv-Phrase, während in (199b) das Verb *haben* die Relation herstellt.

Ein anderer Phänomenbereich, bei dem *Possessor-Possession* angesetzt wurde, liegt vor, wenn ein Merkmal einer Entität wie in den folgenden Beispielen angegeben wird:

- (200) a. **Compact-Disk**: *Aus metallisiertem Kunststoff bestehende kleine, durch Laserstrahl abtastbare Speicherplatte von hoher Ton- bzw. Bildqualität.* (Der Duden, Das Fremdwörterbuch: 224)
- b. **Die Staatsform** *charakterisiert die Organisationsform, „Verfassung“ und äußere Herrschaftsordnung eines Staates und ist damit ein wichtiges Merkmal der staatlichen Grundordnung.* (Wikipedia, Staatsform)
- c. **Die Schallschnelle** [...] *gibt an, mit welcher Wechselgeschwindigkeit die Luftteilchen (bzw. Teilchen des Schallübertragungsmediums) um ihre Ruhelage schwingen; also die Momentangeschwindigkeit eines schwingenden Teilchens.* (Wikipedia, Schallschnelle)
- d. **Frequenz**: *Schwingungszahl von Wellen (pro Sekunde).* (Der Duden, Das Fremdwörterbuch: 408)

Viele weitere dieser Fälle stammen aus dem Teilkorpus Erde und Kosmos, wie bspw. (200d). Dass solche Fälle zumindest dem Sprachgefühl nach Be-

sitzverhältnisse wenn auch einer abstrakteren Art sind, zeigt sich daran, dass Formulierungen mit *besitzen* oder *verfügen über* möglich sind. Dies ist der Hauptgrund, in diesen Fällen *Possessor-Possession* anzusetzen; es wäre aber durchaus vertretbar gewesen, in diesen Fällen eine weitere Pertinenz-Relation anzusetzen. An dieser Stelle sei der Hinweis erlaubt, dass bei der Erstellung einer Ontologie ohnehin abgewogen werden muss, wie die Informationen in die Ontologie übertragen werden (vgl. Kapitel 3.2). Bspw. kann (200b) auf die folgenden Weisen formalisiert werden:

$$\begin{aligned} \text{Staatsform} &\sqsubseteq \exists \text{possessedBy}.\text{StaatlicheGrundordnung} \\ \text{StaatlicheGrundordnung} &\sqsubseteq \exists \text{hatStaatsform}.\text{true} \end{aligned}$$

Im ersten Fall wird staatliche Grundordnung mit Staatsform über die Relation *Possessor-Possession* verbunden analog zu den Beispielen der Teil-Ganzes-Relation, wobei Staatsform *Possession* ist. Diese Formalisierungen könnten für Ontolog*innen aber zu nah am sprachlichen Muster sein, bzw. dieses Muster ist für sie irrelevant, weswegen sie die Relation als Funktion *hatStaatsform* mit dem Wertebereich wahr und falsch ausdrücken könnten.

Im Korpus wurde ein Fall als No-Possessor annotiert, also eine Entität, die darüber definiert wird, etwas nicht zu besitzen. Dabei handelt es sich um den Begriff *Laiendarsteller*:

(201) *Laiendarsteller* (auch *Laienschauspieler*) sind *Darsteller*, die das Schauspiel meistens nicht hauptberuflich ausüben und **keine schauspielerische Ausbildung abgeschlossen haben**. (Wikipedia, Laiendarsteller)

Die Relation ist an der Formulierung nicht direkt zu erkennen, sondern bedarf einer Inferenz: Wenn Laiendarsteller keine Ausbildung abgeschlossen haben, dann besitzen sie keinen Abschluss.

9.3 Zusammenfassung

Die Betrachtung der Pertinenzrelationen hat im Besonderen bei der Teil-Ganzes-Relation gezeigt, dass zwar Standardformulierungen für die Relation existieren, daneben aber eine beträchtliche Anzahl weiterer Formulierungen steht.

Außerdem wurde gezeigt, dass die Standardformulierungen auch genutzt werden, um die taxonomische Relation auszudrücken. Um eine Relation richtig zu identifizieren, müssen daher Inhalt und Form in einem reziproken Verhältnis zueinander begriffen werden. Die Betrachtung der *Possessor-Possession*-Relation war vor allem typologisch wertvoll. Es konnte gezeigt werden, dass diese Relation eine notwendige Erweiterung der Pertinenzrelationen darstellt, da es zwar Fälle gibt, in denen sie sich mit der Teil-Ganzes-Relation überlagert, aber auch Fälle, die diese nicht abdecken kann. Es konnte außerdem gezeigt werden, dass ein Besitzverhältnis eine Art kognitive Blaupause darstellt, da auch Zusammenhänge, bei denen es sich eigentlich nicht um ein solches Verhältnis handelt, nach diesem Muster verbalisiert werden.

10 Numerische Funktionen

Dieses Kapitel beschäftigt sich mit Funktionen, die auf eine Zahl abbilden. In *Description Logics* können Rollen speziell als funktional gekennzeichnet oder axiomatisch eingeführt werden (vgl. Hitzler u.a. 2009: 164). Häufig werden Funktionen im Zusammenhang mit Datentypen verwendet. Bspw. könnte man Größenangaben als Funktionen formalisieren, indem man eine Funktion definiert, die für eine Person oder ein Gebäude die jeweilige Größe bzw. Höhe in Metern angibt. Letztlich lässt sich aber jede Relation, die in dieser Arbeit bisher besprochen wurde, auch als Funktion darstellen. Dies soll an einem möglichst allgemeinen Fall, der Unterklassen-Relation, dargestellt werden. Eine Aussage wie $Hund \sqsubseteq Wirbeltier$ ließe sich auch über eine Funktion, *unterklasseVonWirbeltier*, darstellen, deren Definitionsmenge die Menge der Unterklassen der Tiere und deren Zielmenge die Menge der Wahrheitswerte wahr und falsch ist. Sie bildet dann *Hund* auf wahr ab: $unterklasseVonWirbeltier(Hund) = wahr$. Dies ist zwar keine effiziente Formalisierung, aber sie rückt eine wichtige Frage in den Fokus: Wann sollten Funktionen angesetzt werden und wann nicht?

Die Frage lässt sich im Kontext dieser Arbeit umformulieren: Gibt es natürlich-sprachliche Ausdrücke, die eine Formalisierung als Funktion nahelegen? Wie könnten solche Ausdrücke aussehen? Bei welchen natürlich-sprachlichen Relationen besteht ein 1:1-Verhältnis? Bspw. hat jeder Mensch genau einen Vater; dieser Umstand könnte als Funktion dargestellt werden. Dass sich in diesem Fall eine Formalisierung als Funktion anbietet, bedarf allerdings einer genaueren Betrachtung dieser Relation, und es darf angezweifelt werden, dass dieses Wissen Sprachproduzent*innen sehr präsent ist. In anderen Worten: Diese Relation als Funktion zu formalisieren, ist möglich und evtl. auch sinnvoll, die Begründung fundiert aber nicht auf einem Zusammenhang im Sprachbewusstsein der Sprachproduzent*innen.

Der Gegenstandsbereich, auf den sich dieses Kapitel konzentriert, behandelt Funktionen, die auf Zahlenwerte abbilden. In diesem Fall scheint mir eine Darstellung als Funktion am effizientesten und auch bzgl. des natürlichen Sprachbewusstseins am naheliegendsten. Dies sei kurz am Beispiel Größe erklärt. Die Größe einer Person bspw. ließe sich auch über die Relation zwischen Instanz und Klasse formalisieren. Ist eine Person 195cm groß, könnte die Klasse der 195cm-großen Personen angesetzt werden und die in Frage stehende Person wäre eine Instanz dieser Klasse. Dieses Verfahren ist allerdings äußerst ineffizient und aus natürlich-sprachlicher Sicht unintuitiv. Eine Person hat eine bestimmte Größe, nicht mehrere, und diese lässt sich als Zahlenwert angeben. Eine Formalisierung als Funktion liegt nahe.

Zwei Annotationen wurden im Korpus für diese Relation angesetzt: *Argument* und *Datatype*. Das *Argument* ist die Entität, der ein *Datentyp*, bzw. ein Wert, zugeordnet wird. Obwohl es sich in diesem Kapitel nur um numerale Werte handeln soll, wurde der Ausdruck *Datentyp* gewählt, weil das D in *SHOIN*(\mathcal{D}) für Datentyp steht.

Die Abbildung 27 zeigt die Differenz zwischen dem WK und WoIK bzgl. den Annotationen *Datatype* und *Argument* in Prozent. Zum Vergleich ist zusätzlich die Differenz zwischen den Token aufgeführt. Das WoIK hat etwas

unter 10% weniger Token als das WK, dabei aber fast 40% weniger Annotationen vom Typ *Argument* und über 40% weniger Annotationen vom Typ *Datatype*. Dies ist ein klarer Hinweis darauf, dass Funktionen, die auf Zahlenwerte abbilden, weitaus häufiger im Kontext von Individuen vorkommen.

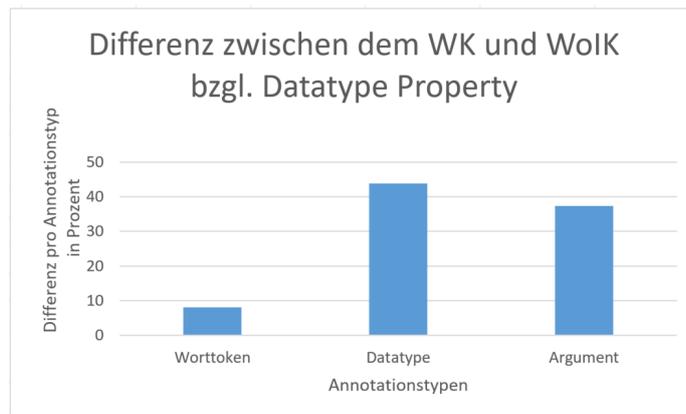


Abbildung 27: Der Vergleich zwischen WK und WoIK: *Datatype Property*

Diese Differenz deckt sich mit der Intuition. Individuen lassen sich über einen bestimmten Wert identifizieren, während sich Klassen selten über einen bestimmten Wert definieren lassen. So hat ein bestimmter Berg, bspw. der Mount Everest, eine ganz bestimmte Höhe, während sich die Höhe, die ein Individuum haben muss, um Instanz der Klasse Berg zu sein, nicht in einer konkreten Zahl ausdrücken lässt. Selbstverständlich gibt es auch Klassen, die sich über konkrete Zahlenwerte definieren lassen. Bspw. hat ein Mittelgebirge Berge zwischen 500 und 1500 Metern Höhe, aber es besteht eine starke Tendenz, konkrete numerische Werte eher für Individuen anzusetzen.

Die meisten Belege stammen aus dem Teilkorpus zu Erde und Kosmos.

- (202) a. **Braune Zwerge** sind Himmelskörper, die eine Sonderstellung zwischen Sternen und Planeten einnehmen. Ihre Massen reichen mit **unter 75 Jupitermassen** nicht aus, um wie in den leichtesten Sternen, den Roten Zwergen, eine Wasserstofffusion in ihrem Inneren in Gang zu setzen. (Wikipedia, Brauner Zwerg)
- b. **Ein Riesenstern** (oder einfach nur Riese) ist ein Stern mit extrem großem Durchmesser und extrem großer Leuchtkraft. [...] In der Regel haben Riesen einen Radius **zwischen 10 und 100 Sonnenradien** bei einer Helligkeit, die zwischen dem 10- und 1000-fachen unserer Sonne liegt. (Wikipedia, Riesenstern)

- c. **Die Andromedagalaxie**, auch *Andromedanebel* oder *Großer Andromeda-Nebel*, ist die der Milchstraße mit **rund 2,5 Millionen Lichtjahre** Entfernung nächstgelegene Spiralgalaxie. (Wikipedia, Andromedagalaxie)

In (202a) kann die im Text genannte Einheit auch als Funktionsname verwendet werden. Eine Formalisierung sähe dementsprechend folgendermaßen aus:

BraunerZwerg \sqsubseteq (*Himmelskoerper* \sqcap \leq *Jupitermasse.75*)

Braune Zwerge sind Unterklassen von Himmelskörpern und den Dingen, die eine Jupitermasse von 75 haben. Man bemerke, dass die Rolle hier ohne Quantor steht, da ein konkreter Wert eingesetzt wird.

In (202b) wird ein Intervall von Werten genannt. Wie ein solches Intervall genau dargestellt wird, ist eine informatische Frage und unterscheidet sich von System zu System, deswegen soll an dieser Stelle eine einfache Lösung verwendet werden:

Riesensterne \sqsubseteq (*Sterne* \sqcap *Radius.10-100Sonnen*)

Bei dem *definiendum* in (202c) handelt es sich um ein Individuum, weswegen die Formalisierung auf den ersten Blick etwas ungewöhnlich aussieht:

EntfernungMilchstrasse.2,5MillionenLichtjahre(*Andromedagalaxie*)

In den Belegen (202a) und (202c) sind die Funktionswerte in einer von *mit* regierten PP enthalten. Diese Formulierung erinnert an Meronymie, wozu passt, dass in (202b) der Funktionswert Teil eines Objekts ist, das von *haben* abhängt. Es besteht die Tendenz, Funktionswerte nach dem Muster der Meronymie zu realisieren.

Auch in anderen Wissensgebieten lässt sich diese Tendenz erkennen:

- (203) a. **Rechtwinklig**: *Einen Winkel von 90° habend*. (Der Duden, Das Fremdwörterbuch: 778)
- b. **Sand** ist ein natürlich vorkommendes, unverfestigtes Sediment, das sich überwiegend aus Mineralkörnern mit **einer Korngröße von 0,063 bis 2 Millimeter** zusammensetzt. (Wikipedia, Sand)

Die folgenden beiden Belege zeigen, dass es sich dabei um mehr als nur ein Muster handelt. Beide ließen sich sowohl als Meronyme oder als Funktionswerte formalisieren.

- (204) a. **Das Great Barrier Reef** (wörtlich übersetzt: (Großes) Barriere-riff, Großes Barrierriff oder Great-Barrier-Riff) vor der Nordostküste Australiens ist die größte zusammenhängende Ansammlung von **über 2.900 einzelnen** Korallenriffen der Erde. (Wikipedia, Great Barrier Reef)
- b. **Die Bundesstadt Bonn** (lateinisch Bonna) ist eine kreisfreie Großstadt im Regierungsbezirk Köln im Süden des Landes Nordrhein-Westfalen und Zweitregierungssitz der Bundesrepublik Deutschland. **Mit 329.673 Einwohnern** (Stand 31. Dezember 2019) zählt Bonn zu den zwanzig größten Städten Deutschlands. (Wikipedia, Bonn)

Die Formalisierung für (204a) könnte folgendermaßen lauten:

$$\geq 2900 \text{ besteht Aus. Korallenriffen}(\text{Great-Barrier-Riff})$$

$$\text{AnzahlDerKorallenriffe.}2900(\text{Great-Barrier-Riff})$$

Anbei die Formalisierung für (204b), wobei hier zwei Formeln benötigt werden, um die Variante mit Meronymen abzubilden:

$$\geq 329.673 \text{ besteht Aus. Einwohner}(\text{Bonn})$$

$$\leq 329.673 \text{ besteht Aus. Einwohner}(\text{Bonn})$$

$$\text{AnzahlDerEinwohner.}329.673(\text{Bonn})$$

Diese Belege verdeutlichen noch einmal, wie schwer es ist, zu entscheiden, wann eine Funktion und wann eine herkömmliche Rolle angesetzt werden sollte.

Selbst innerhalb der numerischen Werte lassen sich verschiedene Typen definieren. Bei den folgenden Belegen handelt es sich um Daten im Sinne von spezifischen Tagen:

- (205) a. **New Horizons** (englisch für Neue Horizonte) ist eine Raumsonde der NASA, die im Rahmen des New-Frontiers-Programmes **am 19. Januar 2006** startete, um das Pluto-System und den Kuipergürtel zu erforschen. (Wikipedia, New Horizons)
- b. **Thomas Alan „Tom“ Waits** (* **7. Dezember 1949** in Pomona, Kalifornien) ist ein US-amerikanischer Sänger, Komponist, Schauspieler und Autor. (Wikipedia, Tom Waits)
- c. **Die Wirecard AG** mit Sitz in Aschheim bei München ist ein **1999** gegründetes börsennotiertes Zahlungsdienstleistungsunternehmen. (Wikipedia, Wirecard)
- d. **Bertolt Brechts Theaterstück Leben des Galilei**, von ihm meist als episches Stück, nicht als Drama bezeichnet, wurde 1939 im dänischen Exil verfasst und **am 9. September 1943** in Zürich uraufgeführt. (Wikipedia, Das Leben des Galilei)

- e. *Yamatoka Jindai Sakura [...] ist ein 1800 bis 2000 Jahre alter Kirschbaum auf dem Gelände des japanischen Tempels Jissō-ji in der Gemeinde Hokuto in der Präfektur Yamanashi.* (Wikipedia, Yamataka Jindai Sakura)

Auch diese ließen sich wieder über herkömmliche Rollen definieren: Man könnte eine Klasse mit Daten erstellen und bspw. über eine Rolle *hatGeburts-tag* jede Person mit ihrem Geburtstag verbinden. Jedoch wäre dies wieder der umständlichere Weg und, auch wenn dies kein starkes Argument ist, so stände ein solches Vorgehen der sprachlichen Intuition entgegen. Ein Geburtstag wird nicht als Relation zu der Entität eines Tages wahrgenommen, sondern, zumindest in den meisten Fällen dürfte dies so sein, als ein Wert, wie er bspw. im Personalausweis steht.

Zum Abschluss des Unterkapitels soll noch ein schwieriger Fall angerissen werden, bei dem evtl. eine numerale Funktion implizit ist.

(206) *Der Panda ist vom Aussterben bedroht.*

Zunächst kann man feststellen, dass es sich bei *Der Panda* um eine generische NP handelt, gemeint ist die Klasse des Pandas. Wie lässt sich die Eigenschaft, vom Aussterben bedroht zu sein, formal greifen? Natürlich ließe sich eine Funktion *vomAussterbenBedroht* anlegen, die eine Klasse von Tieren auf einen Wahrheitswert abbildet. Dieses Vorgehen würde aber nicht die innere semantische Struktur des Prädikats *vom Aussterben bedroht sein* abbilden. Dies bedeutet, dass so wenige Individuen dieser Klasse existieren, dass bald keine Fortpflanzung mehr stattfinden kann und die Klasse daher in naher Zukunft keine Individuen mehr enthält. Wichtig scheint mir der erste Teil dieser Paraphrasierung, dass nur noch wenige Individuen existieren. Dies ist letztlich eine Aussage über die Kardinalität der Klasse. Es könnte eine Funktion *Cardinality* angelegt werden, die eine Klasse auf ihre Kardinalität (zu einem bestimmten Zeitpunkt) abbildet. Dann könnte eine zweite Funktion *vomAussterbenBedroht* angelegt werden, die das Ergebnis von *Cardinality* auf einen Wahrheitswert abbildet, und zwar *true*, wenn die Kardinalität einen bestimm-

ten Wert unterschreitet, ansonsten *false*.

Es lässt sich festhalten, dass es zumindest aus einer sprachwissenschaftlichen Perspektive schwierig ist, auszumachen, wann eine Funktion angesetzt werden sollte. Schließlich wurden numerale Werte, die mindestens größer als zehn sind, als ein sinnvoller Typus von Informationen ausgemacht, bei dem eine Formalisierung als Funktion anderen Formalisierungen vorzuziehen ist. Es wurde weiterhin festgestellt, dass dieser Typus besonders häufig im Zusammenhang mit Individuen auftritt.

11 *False Friends*

Das letzte Kapitel des Hauptteils widmet sich einem ganz anderen Typus von Annotationen. Es handelt sich um Formulierungen, die auf den ersten Blick auf eine bestimmte ontologische Relation hinzudeuten scheinen, bei genauerer Betrachtung aber einen anderen Zusammenhang kodifizieren – oder um Formulierungen, deren ontologischer Inhalt zunächst nicht ersichtlich ist, wie im Fall des Annotationstypus Wissensgebiet. Diese Formulierungen sollten von Rezipient*innen ohne größere Schwierigkeiten verstanden werden können. Das Potenzial zur Verwechslung mit anderen Annotationen besteht hauptsächlich auf einer isolierten, formal-syntaktischen Ebene, die im Bereich der Computerlinguistik aber umso wichtiger ist. Zwei dieser Annotationstypen, Metasprache und Rollenindikatoren, sind außerdem ein Indiz dafür, dass ein tendenzieller Zwang zur Realisierung eines *genus proximum* besteht. Die Benennung dieser Annotationstypen wurde für diese Arbeit entwickelt und soll am Anfang jedes Unterkapitels kurz motiviert werden.

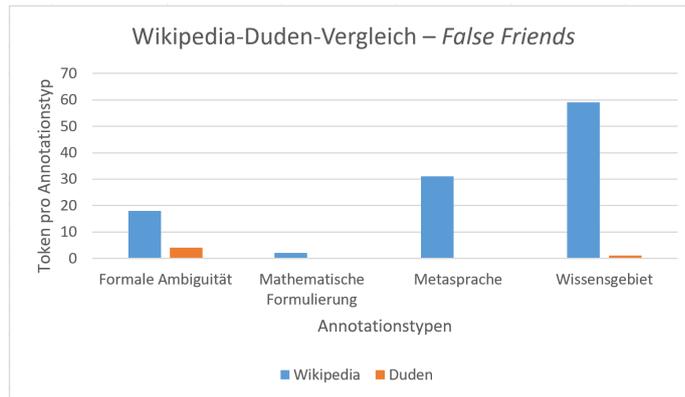


Abbildung 28: Der Vergleich zwischen Wikipedia und Duden: *False Friends*

Der Blick in die Statistik offenbart ein klares Bild: Diese Annotationen treten nahezu ausschließlich in den Wikipedia-Texten und nicht in den Duden-Texten auf. Der Grund hierfür ist erneut die Tatsache, dass die Duden-Texte viel ökonomischer verfasst sind. Lexikographische Formulierungen sind ein Phänomen, das sich nur ergibt, wenn genug Raum für Text besteht.

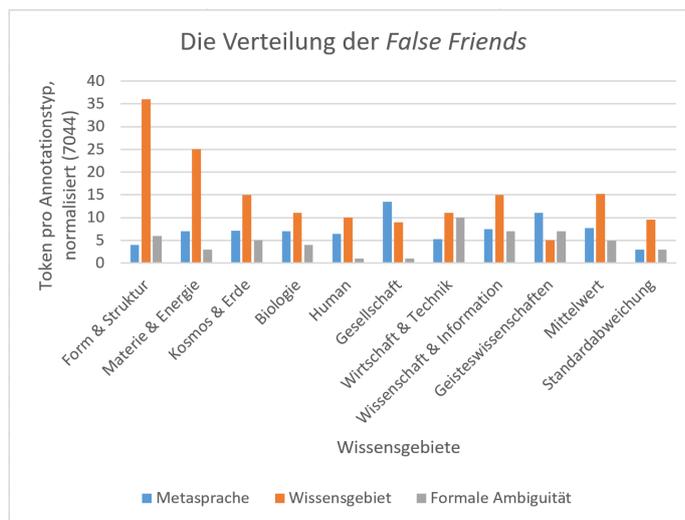


Abbildung 29: Die Verteilung von *False Friends* über die Wissensgebiete

11.1 Metasprache

Unter Metasprache werden Formulierungen verstanden, die ihre eigene Sprachlichkeit reflektieren, indem sie das *definiendum* explizit als Begriff und nicht als Sache bezeichnen.²⁴ Dabei werden Formulierungen verwendet, die zunächst den Eindruck erwecken, es handle sich um die taxonomische Relation,

²⁴Zum Begriff der Metasprache vgl. Wedekind u.a. (2004: 459).

weswegen diese Formulierungen hier den *False Friends* zugerechnet werden.

Es folgen diverse Belege aus verschiedenen Wissensgebieten:

- (207)
- a. **Störschall** ist die **Bezeichnung** für ein Geräusch, das eine gewollte Schallaufnahme oder die Wahrnehmung eines Schallereignisses, z. B. eines Musikvortrags, stört. (Wikipedia, Störschall)
 - b. **Das Wort Wissenschaft** [...] bezeichnet die Gesamtheit des menschlichen Wissens, der Erkenntnisse und der Erfahrungen einer Zeitepoche, welches systematisch erweitert, gesammelt, aufbewahrt, gelehrt und tradiert wird. (Wikipedia, Wissenschaft)
 - c. **Diaspore** oder auch **Propagule** ist **der botanische Sammelbegriff** für alle verbreitungsfördernden Einheiten wie Samen, Teilfrüchte, Früchte, Fruchtsände, Sporen, Brutknospen (Bulbillen) oder sonstige vegetative Brutkörper, die der Ausbreitung von Pflanzen oder Pilzen dienen. (Wikipedia, Diaspore)
 - d. **Drama** [...] ist **ein Oberbegriff** für Texte mit verteilten Rollen. (Wikipedia, Drama)

Syntaktisch werden diese Fälle genauso formuliert wie die taxonomische Relation oder die Relation zwischen Instanz und Klasse, um die es sich gewissermaßen auch handelt. Für Ontolog*innen ist die Information, dass *Störschall* zunächst nur eine Bezeichnung und nicht die Sache an sich ist, allerdings wenig relevant. Eine Formalisierung wie die folgende wäre in jeder Ontologie unsinnig:

Bezeichnung(Stoerschall)

Die Funktion solcher Formulierungen besteht vermutlich darin, eine gewisse Distanz zu den Begriffen aufzubauen, indem sie als solche identifiziert werden. Solche Formulierungen deuten an, dass es sich immer um Konzepte handelt, die die Realität evtl. nur bedingt widerspiegeln.

Der folgende Beleg ist in diesem Zusammenhang interessant:

- (208) *Die Monarchie bildet somit das Gegenstück zum modernen Republikbegriff.* (Wikipedia, Monarchie)

Zunächst ist die Metasprache hier Teil eines Kompositums. Wollte man kleinlich sein, müsste man anmerken, dass nicht die Monarchie Gegenstück zum Republikbegriff sei, sondern der Monarchiebegriff – allerdings wird nur das eine Wort als Begriff gekennzeichnet. Der Beleg scheint zu zeigen, dass die

Verwendung von Metasprache eine fakultative Strategie darstellt, die nicht streng angewandt werden muss.

Ein außergewöhnliches Beispiel für Metasprache stellt der folgende Beleg dar, da hier ein sehr spezifisches Wort, *Homonym*, verwendet wird:

(209) *Organisation [...] ist ein Homonym, das sich allgemein mit „Bewerbstellung“ übersetzen lässt.* (Wikipedia, Organisation)

Die Verwendung von Metasprache erfüllt hier allerdings den ausdrücklichen Zweck, ein Wort als mehrdeutig zu kennzeichnen.

In den folgenden Belegen erfüllt die Metasprache den Zweck, verschiedene Begriffe voneinander abzugrenzen bzw. miteinander in Beziehung zu setzen:

- (210) a. *Der Begriff „**Schiffahrt**“ ist nicht mit Schiffsfahrt, also einer Fahrt mit einem Schiff, zu verwechseln.* (Wikipedia, Schiffahrt)
- b. ***Willenskraft** oder **Willensstärke** (englisch willpower) ist eine alltagssprachliche Bezeichnung für den Fachbegriff Volition aus der Psychologie und Managementlehre.* (Wikipedia, Willenskraft)

Eine wenig ausgeprägte Form von Metasprache liegt vor, wenn das Verb *bezeichnen* verwendet wird.

- (211) a. *Ton **bezeichnet** in der Musik ein Schallereignis (sowie dessen Hör-eindruck und gedankliche Abstraktion), das von Musikinstrumenten, der menschlichen Stimme oder anderweitig mittels eines elastischen Körpers erzeugt wird und dem eine mehr oder weniger exakte Tonhöhe zugeordnet werden kann.* (Wikipedia, Ton)
- b. *Als lineare Funktion wird oft (insbesondere in der Schulmathematik) eine Funktion $f: \mathbb{R} \rightarrow \mathbb{R}$ der Form $f(x) = m \cdot x + n$ $m, n \in \mathbb{R}$, also eine Polynomfunktion höchstens ersten Grades, **bezeichnet**.* (Wikipedia, Lineare Funktion)
- c. *Interpretation oder (musikalischer) Vortrag **bezeichnet** in seiner musikpraktischen Bedeutung zunächst die Ausführung einer Komposition durch einen oder mehrere Musiker in öffentlicher Aufführung oder als Tonaufzeichnung.* (Wikipedia, Interpretation)

In all diesen Fällen wird die taxonomische Relation realisiert. Das Verb *bezeichnen* deutet allerdings an, dass es sich bei dem jeweiligen *definiendum*

nicht um die Sache selber, sondern um einen Begriff handelt. Würde man diese Strategie streng verfolgen, müssten die Sätze eigentlich folgendermaßen lauten:

(212) *Das Wort Ton bezeichnet eine Untermenge der Dinge, die durch das Wort Schallereignis bezeichnet werden.*

Selbstverständlich ist eine solche Formulierung zu umständlich und bietet keinen informativen Mehrwert, da Sprachrezipient*innen die Verhältnisse ohne größere Schwierigkeiten richtig interpretieren können. Die Belege (211) verdeutlichen, dass Metasprache ein Formulierungsmuster darstellt, das ergänzend gewählt werden kann.

Die Verwendung von Metasprache lässt sich allerdings nicht nur aus dem Bedürfnis, die Konzeptionalität von Begriffen zu markieren, erklären, sondern kann ergänzend auch schlichtweg als ein Formulierungsmuster für Definitionen verstanden werden. Es scheint, ein gewisser Zwang zu bestehen, ein *genus proximum* zu benennen, und wenn dies inhaltlich nicht ohne Weiteres möglich ist, bzw. wenn sich nur ein sehr allgemeiner Kandidat finden lässt, stellt Metasprache eine Ausweichstrategie dar. Für diese These stellt der folgende Text einen starken Beleg dar:

(213) *Eine Geometrische Figur ist **ein Begriff** aus der Geometrie, der uneinheitlich verwendet wird und häufig undefiniert bleibt.* (Wikipedia, Geometrische Figur)

Zunächst lässt sich feststellen, dass der Text etwas unsauber formuliert ist: Korrekterweise müsste es heißen: *Geometrische Figur ist ein Begriff...* Die NP *Eine Geometrische Figur* kann nicht mehr metasprachlich verwendet werden, da der unbestimmte Artikel es unmöglich macht, unter der NP noch den anderen Begriff zu verstehen. Der Satz wechselt nach dem Verb also regelrecht in den Modus von Metasprache. Diese Entscheidung, zu wechseln, könnte dadurch ausgelöst worden sein, dass *Figur* als *genus proximum* zu redundant wirkte, da dieses Wort schon in der NP realisiert wird. Diese Erklärung gilt

natürlich nur vor dem Hintergrund, dass zumindest ein tendenzieller Zwang zur Realisierung eines *genus proximum* angenommen wird.

11.2 Formale Ambiguität

Als formale Ambiguitäten wurden Wörter und Wortgruppen annotiert, die sich auf zwei verschiedene Arten formalisieren lassen: Entweder als Oberklasse oder als Teil einer anderen Relation. Aus formaler Sicht sind beide Varianten gleichwertig, wobei die zweite als informativer gewertet werden kann. Ein typisches Beispiel findet sich in Beleg (214):

- (214) *Algebra: **Teilgebiet** der Mathematik, das sich besonders mit Gleichungen und mit den Verknüpfungen mathematischer Strukturen befasst: eine Eins in Algebra haben.* (Duden, Das Bedeutungswörterbuch: 75)

Formal kann nichts gegen eine Formalisierung wie die folgende eingewandt werden:

Algebra \sqsubseteq Teilgebiet

Algebra ist eine Unterklasse der Teilgebiete. Die Aussage ist richtig, bietet aber wenig informativen Mehrwert. Die wichtige Information ist, dass Algebra ein Teil der Mathematik ist, dass also ein Fall von Meronymie vorliegt (vgl. Kapitel 3.2.2). Die folgende Formalisierung ist daher informativer:

Algebra \sqsubseteq \exists isPartOf.Mathematik

Der überwiegende Teil solcher Ambiguitäten zeigt ein meronymisches Verhältnis an. Dabei kann die ambige Wortgruppe sowohl das Meronym (215a) als auch das Holonym (215b) und 215c) realisieren:

- (215) a. *Ein rechter Winkel, kurz auch Rechter, ist ein Winkel von 90° und damit **der vierte Teil** eines Vollwinkels zu 360° .* (Wikipedia, Rechter Winkel)
- b. *Eine Menge ist **ein Verbund, eine Zusammenfassung** von einzelnen Elementen.* (Der Duden, Das Bedeutungswörterbuch: 249)
- c. *Als Rechtsordnung (oder Rechtssystem) bezeichnet man **die Gesamtheit** des gültigen objektiven Rechts in dessen Anwendungsbereich, beispielsweise das Recht eines Staates. Neben dem durch*

die Legislative gesetzten Recht (Rechtsetzung) gehört zur staatlichen Rechtsordnung auch das Gewohnheitsrecht und die Gesetzesauslegung durch die Judikative (Rechtspflege), sodann die zwangsweise Durchsetzung des Rechts mit Hilfe der Staatsgewalt (insbesondere durch die Polizei oder den Gerichtsvollzieher, der auch privatrechtliche Ansprüche durchsetzt). (Wikipedia, Rechtsordnung)

In manchen Fällen beinhalten die formal ambigen Ausdrücke noch Informationen über die einfachen Relationen hinaus:

- (216) a. *Die Sonne ist der Stern, der der Erde am nächsten ist und **das Zentrum** des Sonnensystems bildet. (Wikipedia, Sonne)*
b. *Quadrate sind **die Seitenflächen** eines platonischen Körpers, nämlich des Würfels. (Wikipedia, Quadrat)*

In (216a) zeichnet die NP *das Zentrum* die Sonne zwar als Teil des Sonnensystems aus, sie beinhaltet aber selbstverständlich auch noch Informationen über die Anordnung des Sonnensystems. Die NP *die Seitenflächen* in (216b) sagen ebenfalls nicht nur aus, dass Quadrate Teile eines Würfels sind, sondern bestimmte Teile. Ob man solche Einheiten nur als Rollenindikator betrachtet oder versucht mehr Informationen zu formalisieren, ist letztlich eine Frage der Granularität.

Der letzte Beleg soll noch einmal das Verhältnis dieser Ausdrücke zum *genus proximum* problematisieren:

- (217) *Buchhaltung: **Abteilung** eines Betriebes, die für die Buchführung verantwortlich ist. (Duden, Das Bedeutungswörterbuch: 249)*

Einerseits kann *Abteilung* als Meronym interpretiert werden und Buchhaltung damit als Teil eines Betriebes. Andererseits könnte Buchhaltung als Unterklasse von Abteilung und Abteilung wiederum als Teil eines Betriebes formalisiert werden. Beide Varianten wären effektiv gleich.

11.3 Wissensgebiet

Diese Annotation wird für Wörter und Wortgruppen angesetzt, die das Wissensgebiet, in dessen Kontext ein Text zu verstehen ist, angibt. Als Beispiel

sei der folgende Beleg gegeben:

- (218) *In der euklidischen Geometrie definiert man: Zwei Geraden sind parallel, wenn sie in einer Ebene liegen und einander nicht schneiden.* (Wikipedia, Parallelität)

Es sind nicht irgendwelche Geraden unter den beschriebenen Bedingungen parallel, sondern euklidische Geraden. Das Wissensgebiet könnte also genutzt werden, um eine Klasse zu spezifizieren. Ob dies gewollt ist, muss im Einzelfall entschieden werden; es ist in jedem Fall konsistent. Im folgenden Beleg wird diese Strategie verfolgt, um die Unterklasse zu spezifizieren.

- (219) *Ein astronomisches Objekt (auch Himmelsobjekt oder spezieller Himmelskörper) ist ein Objekt im Weltall, das von der Astronomie und der Astrophysik untersucht wird.* (Wikipedia, Astronomisches Objekt)

Der Annotationstypus Wissensgebiet wird zu den *False Friends* gerechnet, da diese Möglichkeit der Spezifizierung nicht auf den ersten Blick ersichtlich ist, da das Wissensgebiet meist als Adverbial, wie bspw. in (218), realisiert wird, und damit als Modifizierung des ganzen Satzes und nicht einer einzelnen Klasse erscheint. Wissensgebiet und Metasprache sind Annotationstypen, die oft zusammen auftreten. Etwas wird als Begriff gekennzeichnet, und anschließend wird spezifiziert, in welchem Bereich dieser Begriff verwendet wird:

- (220) a. *Auslautverhärtung ist ein Fachbegriff aus der Sprachwissenschaft, im Speziellen aus der Phonetik und Phonologie, und bezeichnet den Vorgang, dass Geräuschkonsonanten (d.h. Plosive, Affrikate und Frikative) am Ende einer Silbe (also in ihrem Auslaut) ihre Stimmhaftigkeit verlieren und stimmlos werden.* (Wikipedia, Auslautverhärtung)
- b. *Kreuzung ist ein Begriff aus der Genetik.* (Wikipedia, Kreuzung)
- c. *Herrschaftsform bezeichnet zentrale Begriffe der Politikwissenschaft, der Soziologie und der politischen Philosophie. Diese betreffen die spezifischen Ausprägungen politischer Herrschaft, d. h. die zentralen Aspekte der Herrschaftsausübung (Umfang, Struktur, Herrschaftsweise) und des Herrschaftszugangs.* (Wikipedia, Herrschaftsform)

- d. *Parameter [...] ist ein aus den Naturwissenschaften entlehnter Begriff und bezeichnet in der Musik elementare Aspekte wie Tondauer, Tonhöhe und Tonstärke bzw. Lautstärke.* (Wikipedia, Parameter)
- e. *Bildredakteur ist eine Berufsbezeichnung aus dem Medienbereich.* (Wikipedia, Bildredakteur)

Dass Wissensgebiete in Unterklassen überführt werden können, zeigt sich besonders deutlich an folgendem Beleg:

- (221) *Beispiele für Schwingungen sind in der Mechanik, in der Elektrotechnik, der Biologie, in der Wirtschaft und in vielen anderen Bereichen anzutreffen.* (Wikipedia, Schwingung)

Hier ließe sich eine Reihe von Unterklassen zu dem Begriff *Schwingung* bilden: elektrotechnische Schwingung, biologische Schwingung, wirtschaftliche Schwingung.

11.4 Zusammenfassung

Der Blick auf die lexikographischen Formulierungen hat nochmal zurück zu der taxonomischen Relation geführt. Während in Kapitel 7.1 gezeigt wurde, auf wie viele unterschiedliche Weisen diese Relation realisiert werden kann, wurde in diesem Kapitel verdeutlicht, dass Formulierungen existieren, die formal-syntaktisch wie diese Relation auftreten, semantisch aber anders interpretiert werden sollten. Es zeigt sich, dass weder syntaktische Muster fest auf ontologische Rollen abgebildet werden können noch umgekehrt.

12 Fazit

Diese Arbeit hatte die Aufgabe, das Konstrukt der ontologischen Semantik korpusgestützt zu untersuchen, um einen Grundlagenbeitrag für das Forschungsgebiet des *Ontology Learning* zu leisten. Dabei standen zwei Fragen im Fokus: Auf welcher Basis kann eine Typologie von ontologisch-relevanten Informationen erstellt werden und wie werden diese Informationen in Texten verbalisiert? In Kapitel 2 wurde diese Arbeit thematisch positioniert. In

diesem Kontext wurde auch das Verständnis von Semantik vorgestellt, das hinter dem Begriff *Semantic Web* steht. In Kapitel 3 wurden die formalen und sprachwissenschaftlichen Konstrukte, die im weiteren Verlauf der Arbeit Verwendung fanden, eingeführt und definiert. Generell wurde in diesem Kapitel ausgeführt, welche formalen Möglichkeiten Logik und Mengenlehre semantischen Untersuchungen zur Verfügung stellen. Im abschließenden Unterkapitel wurde erörtert, welche Relevanz die Aspekte der intensionalen Logik für die Arbeit haben, mit dem Ergebnis, dass besonders Individuen mit den Instrumenten der intensionalen Logik beschrieben werden können. Es wurde daraufhin ausgeführt, wie dies auch mit dem extensionalen Ansatz der Modelltheorie geschehen kann, wenn diese um die Domäne der Datentypen erweitert wird. Kapitel 4 hat eine Ontologiedefinition erarbeitet, indem sie die notwendigen Komponenten einer Ontologie aufführte und außerdem ausführte, warum Individuen und Datentypen kein notwendiger Bestandteil einer Ontologie sind. Kapitel 5 führte aus, warum die Vorstellung von Bedeutung, die Ontologien zu Grunde liegt, im Wesentlichen der logischen Merkmalssemantik entspricht. Daraufhin wurden zwei alternative Merkmalssemantiken, die Prototypensemantik und die Framesemantik, vorgestellt, um deren Tauglichkeit für ontologische Untersuchungen zu ermitteln. Es wurde festgestellt, dass nach den Prinzipien dieser alternativen Theorien letztlich auch Ontologien erstellt werden können, dass die logische Merkmalssemantik aber besser geeignet ist, da ihr eine eindeutigere Vorstellung von Definitionen zu Grunde liegt. In Kapitel 6 wurde das Korpus, das zur Unterstützung der Untersuchungen erstellt wurde, das Verfahren der Korpuserstellung und die Annotationstypologie vorgestellt. Aufgrund der komplexen Annotationstypologie war es notwendig, das Korpus händisch zu erstellen. Die Auswertung des Korpus orientierte sich an der Frage, wie ontologisch relevante Informationen in Texten identifiziert und in eine Ontologie überführt werden können. Dabei zeigt sich zunächst, dass in den Wikipediatexten mehr ontologisch relevante Informationen realisiert werden. Die Kapitel 7 und 9 widmeten sich den mengen-

theoretischen Relationen und den Pertinenzrelationen. In diesen Kapiteln war es ein Hauptanliegen, zu zeigen, dass kein 1:1-Verhältnis zwischen natürlich-sprachlichen Mustern und bestimmten Relationen vorliegt. Es ließen sich bestimmte Standardfälle finden, aber daneben auch ein großer Bereich von Ausnahmen und Zweifelsfällen. Eine weitere Erkenntnis dieses Kapitels war, dass Metaklassen sich zwar formal definieren lassen, dass sie allerdings kein Konstrukt sind, das in der Kognition von Sprachproduzent*innen verankert ist, da sie sich in ihrer Verbalisierung nicht von herkömmlichen Klassen unterscheiden. In Kapitel 8 konnte gezeigt werden, dass sich viele ontologisch-relevante Informationen tatsächlich als semantische Rollen begreifen lassen und dass dieser Ansatz die Basis für eine Übersetzung von natürlich-sprachlichen Texten in formale Konstrukte bilden kann. Ziel dieser Arbeit war es dabei, diesen Ansatz aus einer sprachwissenschaftlichen Perspektive vorzustellen und zu untersuchen. Wenn in Zukunft ein konkretes System auf Basis dieses Ansatzes gebaut würde, müssten sicherlich Anpassungen vorgenommen werden. Eine mögliche Anpassung wäre dabei, die Annotationstypologie auf bestimmte Rollen zu beschränken. In Kapitel 10 wurde ausgeführt, welche semantischen Zusammenhänge am ehesten als Funktionen, die auf eine Zahl abbilden, formalisiert werden sollten. Dabei wurden vor allem Zusammenhänge, die sehr große Zahlen involvieren, identifiziert. In diesem Kapitel konnte außerdem gezeigt werden, dass vor allem Individuen über numerische Werte definiert werden. Kapitel 11 zeigte erneut, wie komplex das Verhältnis von Syntax und Semantik ist, indem es Formulierungen vorstellt, die zwar wie semantische Relationen aussehen und unter Umständen sogar als solche begriffen werden können, in den meisten Fällen aber die semantische Bedeutung anderer Argumente im Satz hervorheben. Spätestens hier wurde deutlich, dass sprachliche Strukturen und ontologische Zusammenhänge nicht eins zu eins aufeinander abgebildet werden können (vgl. von Kutschera 1975: 37). Gerade deswegen ist es lohnenswert, sich mit ontologischen Relationen und ihrer Verbalisierung in natürlich-sprachlichen Texten auseinanderzusetzen. Es

offenbart die Vielfalt, mit der die Zusammenhänge der Wirklichkeit ausgedrückt werden, und lässt evtl. Rückschlüsse darauf zu, wie der Mensch seine Wirklichkeit wahrnimmt. Konkret könnten die Erkenntnisse genutzt werden, um automatisch ontologisch relevante Informationen aus Texten zu gewinnen und in Ontologien zu überführen. Es wäre gewinnbringend, im Bereich der *Ontology Learning* einen Algorithmus zu entwickeln, der eine größere Textmenge bzgl. der hier vorgestellten Typologie annotiert. Es scheint ebenso lohnenswert, den hier vorgestellten Ansatz mit einer weniger differenzierten Typologie zu automatisieren. Da diese Arbeit aber im Kern sprachwissenschaftlich konzipiert ist, wurde die Annotationstypologie nicht vereinfacht, sondern hinreichend spezifisch ausgearbeitet, um den unterschiedlichen Texten im Korpus gerecht zu werden.

Die wachsende Bedeutung künstlicher Intelligenz (KI) betrifft auch die in dieser Arbeit erforschten Konzepte. Die Methoden dieser Arbeit, sprachliche Muster zu identifizieren und zu formalisieren, ließen sich in Methoden der regelbasierten KI und symbolischem KI übersetzen. Daneben stehen Methoden des *Machine Learning* bei denen ein Computer in die Lage versetzt wird, neue Daten zu bewerten, indem er auf bekannten Datenmengen trainiert wird (vgl. Hastie u.a. 2009: 1-2). Diese Methoden produzieren derzeit erstaunliche Ergebnisse, und es wäre zu erwarten, dass mit ihrer Hilfe auch im Bereich des *Ontology Learning* viele Probleme gelöst werden können. In dem Kontext von *Machine Learning* stehen auch *Large Language Models*. Dabei handelt es sich Modelle, die probabilistisch auf großen Datenmengen trainiert werden und eine Wahrscheinlichkeit für eine Sequenz von Wörtern errechnen (vgl. Teubner u.a. 2023: 96). Praktisch bedeutet dies, dass mit solchen Modellen bspw. die wahrscheinlichste Fortführung eines Textes errechnet und ausgegeben werden kann. Dabei steht die Frage im Raum, inwieweit ontologisches Wissen in diesen Modellen enthalten ist und abgerufen werden kann. Ein Schwachpunkt, den die Methoden des *Machine Learning* für die Arbeit mit Ontologien aufweisen, liegt im Anspruch, dass das Wissen in einer

Ontologie diskret sein sollte, d.h. die hinterlegten Axiome sollten bestimmt wahr und nicht nur wahrscheinlich wahr sein. Es gilt zu untersuchen, welche Auswirkungen dieser Umstand hat. Umgekehrt stellt sich die Frage, ob Ontologien als diskrete Wissensbasis nicht eine Ergänzung zu probabilistischen Ansätzen bilden können. Abschließend lässt sich sagen, dass die Verbindung von natürlich-sprachlicher und formaler Semantik in Form einer Ontologie ein spannendes Forschungsfeld darstellt, das in Zukunft noch von Bedeutung sein wird, da es eine Brücke zwischen menschlicher und maschineller Kognition darstellt.

Literaturverzeichnis

Quellenverzeichnis

Duden-Ausgaben

Duden (2020): *Das Fremdwörterbuch*. 12., vollständig überarbeitete und erweiterte Auflage. Herausgegeben von der Dudenredaktion, Berlin: Dudenverlag.

Duden (2018): *Das Bedeutungswörterbuch*. 5., neu bearbeitete und erweiterte Auflage. Herausgegeben von der Dudenredaktion, Berlin: Dudenverlag.

Wikipedia-Artikel

(2060) Chiron:

[https://de.wikipedia.org/wiki/\(2060\)_Chiron](https://de.wikipedia.org/wiki/(2060)_Chiron), letzter Zugriff: 30.11.2020.

A24 (Unternehmen):

[https://de.wikipedia.org/wiki/A24_\(Unternehmen\)](https://de.wikipedia.org/wiki/A24_(Unternehmen)), letzter Zugriff: 02.12.2020.

Abrundungsfunktion und Aufrundungsfunktion:

https://de.wikipedia.org/wiki/Abrundungsfunktion_und_Aufrundungsfunktion, letzter Zugriff: 11.11.2020.

Absender:

<https://de.wikipedia.org/wiki/Absender>, letzter Zugriff: 19.11.2020.

Absolute Monarchie:

https://de.wikipedia.org/wiki/Monarchie#Absolute_Monarchie, letzter Zugriff: 30.10.2020.

Ada Lovelace:

https://de.wikipedia.org/wiki/Ada_Lovelace, letzter Zugriff: 30.11.2020.

AGB-Stern:

<https://de.wikipedia.org/wiki/AGB-Stern>, letzter Zugriff: 18.11.2020.

Akustische Impedanz:

https://de.wikipedia.org/wiki/Akustische_Impedanz, letzter Zugriff: 12.11.2020.

Albert Einstein:

https://de.wikipedia.org/wiki/Albert_Einstein, letzter Zugriff: 30.11.2020.

Algebraische Struktur:

https://de.wikipedia.org/wiki/Algebraische_Struktur, letzter Zugriff: 11.11.2020.

Alkoholische Getränke:

https://de.wikipedia.org/wiki/Getränk#Alkoholische_Getränke, letzter Zugriff: 29.10.2020.

Amine:

<https://de.wikipedia.org/wiki/Amine>, letzter Zugriff: 20.10.2020.

Amphetamin:

<https://de.wikipedia.org/wiki/Amphetamin>, letzter Zugriff: 29.10.2020.

Analogon (Chemie):

[https://de.wikipedia.org/wiki/Analogon_\(Chemie\)](https://de.wikipedia.org/wiki/Analogon_(Chemie)), letzter Zugriff: 20.10.2020.

Andromedagalaxie:

<https://de.wikipedia.org/wiki/Andromedagalaxie>, letzter Zugriff: 30.11.2020.

Anforderung:

<https://de.wikipedia.org/wiki/Anforderung>, letzter Zugriff: 24.11.2020.

Angela Merkel:

https://de.wikipedia.org/wiki/Angela_Merkel, letzter Zugriff: 30.11.2020.

Anordnung (Recht):

[https://de.wikipedia.org/wiki/Anordnung_\(Recht\)](https://de.wikipedia.org/wiki/Anordnung_(Recht)), letzter Zugriff: 24.11.2020.

Anstalt des öffentlichen Rechts:

https://de.wikipedia.org/wiki/Anstalt_des_öffentlichen_Rechts, letzter Zugriff: 09.11.2020.

Antigone (Sophokles):

[https://de.wikipedia.org/wiki/Antigone_\(Sophokles\)](https://de.wikipedia.org/wiki/Antigone_(Sophokles)), letzter Zugriff: 02.12.2020.

Antischall:

<https://de.wikipedia.org/wiki/Antischall>, letzter Zugriff: 12.11.2020.

Äpfel:

<https://de.wikipedia.org/wiki/Äpfel>, letzter Zugriff: 20.10.2020.

Arbre de Diane:

https://de.wikipedia.org/wiki/Arbre_de_Diane, letzter Zugriff: 30.11.2020.

Art (Biologie) :

[https://de.wikipedia.org/wiki/Art_\(Biologie\)](https://de.wikipedia.org/wiki/Art_(Biologie)), letzter Zugriff: 20.10.2020.

Asteroid:

<https://de.wikipedia.org/wiki/Asteroid>, letzter Zugriff: 17.11.2020.

Astronomie:

<https://de.wikipedia.org/wiki/Astronomie>, letzter Zugriff: 28.10.2020.

Astronomisches Objekt:

https://de.wikipedia.org/wiki/Astronomisches_Objekt, letzter Zugriff: 17.11.2020.

Astrophysik:

<https://de.wikipedia.org/wiki/Astrophysik>, letzter Zugriff: 28.10.2020.

Auditive Wahrnehmung:

https://de.wikipedia.org/wiki/Auditive_Wahrnehmung, letzter Zugriff: 12.11.2020.

Aufführung:

<https://de.wikipedia.org/wiki/Aufführung>, letzter Zugriff: 19.11.2020.

Aufgabe (Pflicht):

[https://de.wikipedia.org/wiki/Aufgabe_\(Pflicht\)](https://de.wikipedia.org/wiki/Aufgabe_(Pflicht)), letzter Zugriff: 24.11.2020.

Ausbreitungsmedium:

<https://de.wikipedia.org/wiki/Ausbreitungsmedium>, letzter Zugriff: 27.10.2020.

Ausdauer (Psychologie):

[https://de.wikipedia.org/wiki/Ausdauer_\(Psychologie\)](https://de.wikipedia.org/wiki/Ausdauer_(Psychologie)), letzter Zugriff: 18.11.2020.

Auslautverhärtung:

<https://de.wikipedia.org/wiki/Auslautverhärtung>, letzter Zugriff: 27.11.2020.

Auslenkung:

<https://de.wikipedia.org/wiki/Auslenkung>, letzter Zugriff: 12.11.2020.

Bakterien:

<https://de.wikipedia.org/wiki/Bakterien>, letzter Zugriff: 20.10.2020.

Bedecktsamer:

<https://de.wikipedia.org/wiki/Bedecktsamer>, letzter Zugriff: 20.10.2020.

Beleuchtung:

<https://de.wikipedia.org/wiki/Beleuchtung>, letzter Zugriff: 20.11.2020.

Beobachtung:

<https://de.wikipedia.org/wiki/Beobachtung>, letzter Zugriff: 24.11.2020.

Bernoulli Ungleichung:

https://de.wikipedia.org/wiki/Bernoullische_Ungleichung, letzter Zugriff: 26.11.2020.

Betragsfunktion:

<https://de.wikipedia.org/wiki/Betragsfunktion>, letzter Zugriff: 11.11.2020.

Betrieb:

<https://de.wikipedia.org/wiki/Betrieb>, letzter Zugriff: 09.11.2020.

Betriebszweck:

<https://de.wikipedia.org/wiki/Betriebszweck>, letzter Zugriff: 09.11.2020.

Bettlereiche:

<https://de.wikipedia.org/wiki/Bettlereiche>, letzter Zugriff: 30.11.2020.

Bild (Zeitung):

[https://de.wikipedia.org/wiki/Bild_\(Zeitung\)](https://de.wikipedia.org/wiki/Bild_(Zeitung)), letzter Zugriff: 02.12.2020.

Bild der Frau:

https://de.wikipedia.org/wiki/Bild_der_Frau, letzter Zugriff: 02.12.2020.

Bildredakteur:

<https://de.wikipedia.org/wiki/Bildredakteur>, letzter Zugriff: 19.11.2020.

Birnen:

<https://de.wikipedia.org/wiki/Birnen>, letzter Zugriff: 20.10.2020.

BlackRock:

<https://de.wikipedia.org/wiki/BlackRock>, letzter Zugriff: 02.12.2020.

Blauer Riese:

https://de.wikipedia.org/wiki/Blauer_Riese, letzter Zugriff: 18.11.2020.

Blüte:

<https://de.wikipedia.org/wiki/Blüte>, letzter Zugriff: 20.10.2020.

Bonn:

<https://de.wikipedia.org/wiki/Bonn>, letzter Zugriff: 02.12.2020.

Brauner Zwerg:

https://de.wikipedia.org/wiki/Brauner_Zwerg, letzter Zugriff: 28.10.2020.

Breite Föhre:

https://de.wikipedia.org/wiki/Breite_Föhre, letzter Zugriff: 30.11.2020.

Brombeeren:

<https://de.wikipedia.org/wiki/Brombeeren>, letzter Zugriff: 20.10.2020.

Buchhaltung:

<https://de.wikipedia.org/wiki/Buchhaltung>, letzter Zugriff: 24.11.2020.

Bühne (Theater):

[https://de.wikipedia.org/wiki/Bühne_\(Theater\)](https://de.wikipedia.org/wiki/Bühne_(Theater)), letzter Zugriff: 19.11.2020.

Bühnenbild:

<https://de.wikipedia.org/wiki/Bühnenbild>, letzter Zugriff: 20.11.2020.

Cannabinoide:

<https://de.wikipedia.org/wiki/Cannabinoide>, letzter Zugriff: 29.10.2020.

CD-ROM:

<https://de.wikipedia.org/wiki/CD-ROM>, letzter Zugriff: 18.11.2020.

Chaldnische Klangfigur:

https://de.wikipedia.org/wiki/Chaldnische_Klangfigur, letzter Zugriff: 27.10.2020.

Chef vom Dienst:

https://de.wikipedia.org/wiki/Chef_vom_Dienst, letzter Zugriff: 19.11.2020.

Chemische Verbindung:

https://de.wikipedia.org/wiki/Chemische_Verbindung, letzter Zugriff: 20.10.2020.

Chemischer Stoff:

https://de.wikipedia.org/wiki/Chemischer_Stoff, letzter Zugriff: 20.10.2020.

Chemisches Element:

https://de.wikipedia.org/wiki/Chemisches_Element, letzter Zugriff: 20.10.2020.

Cicero (Zeitschrift):

[https://de.wikipedia.org/wiki/Cicero_\(Zeitschrift\)](https://de.wikipedia.org/wiki/Cicero_(Zeitschrift)), letzter Zugriff: 02.12.2020.

Compact Disc Digital Audio:

https://de.wikipedia.org/wiki/Compact_Disc_Digital_Audio, letzter Zugriff: 18.11.2020.

Computervermittelte Kommunikation:

https://de.wikipedia.org/wiki/Computervermittelte_Kommunikation, letzter Zugriff: 10.11.2020.

Controlling:

<https://de.wikipedia.org/wiki/Controlling>, letzter Zugriff: 24.11.2020.

Cotton Tree:

https://de.wikipedia.org/wiki/Cotton_Tree, letzter Zugriff: 30.11.2020.

Darstellende Kunst:

https://de.wikipedia.org/wiki/Darstellende_Kunst, letzter Zugriff: 10.11.2020.

Darsteller :

<https://de.wikipedia.org/wiki/Darsteller>, letzter Zugriff: 19.11.2020.

Das Leben des Galilei:

https://de.wikipedia.org/wiki/Leben_des_Galilei, letzter Zugriff: 02.12.2020.

Daten:

<https://de.wikipedia.org/wiki/Daten>, letzter Zugriff: 18.11.2020.

Datenerfassung:

<https://de.wikipedia.org/wiki/Datenerfassung>, letzter Zugriff: 24.11.2020.

Datenlink:

<https://de.wikipedia.org/wiki/Datenlink>, letzter Zugriff: 10.11.2020.

Datenspeicher:

<https://de.wikipedia.org/wiki/Datenspeicher>, letzter Zugriff: 18.11.2020.

Datenübertragung:

<https://de.wikipedia.org/wiki/Datenübertragung>, letzter Zugriff: 10.11.2019.

De Broglie Wellenlänge:

<https://de.wikipedia.org/wiki/Wellenlänge#De-Broglie-Wellenlänge>, letzter Zugriff: 27.11.2020.

Denken:

<https://de.wikipedia.org/wiki/Denken>, letzter Zugriff: 18.11.2020.

Der gestiefelte Kater:

[https://de.wikipedia.org/wiki/Der_gestiefelte_Kater_\(Tieck\)](https://de.wikipedia.org/wiki/Der_gestiefelte_Kater_(Tieck)), letzter Zugriff: 02.12.2020.

Deutsche Bank:

https://de.wikipedia.org/wiki/Deutsche_Bank, letzter Zugriff: 02.12.2020.

Deutschland:

<https://de.wikipedia.org/wiki/Deutschland>, letzter Zugriff: 02.12.2020.

Dezibel (Einheit):

[https://de.wikipedia.org/wiki/Bel_\(Einheit\)](https://de.wikipedia.org/wiki/Bel_(Einheit)), letzter Zugriff: 27.11.2020.

Diaspore:

<https://de.wikipedia.org/wiki/Diaspore>, letzter Zugriff: 20.10.2020.

Die Ratten:

https://de.wikipedia.org/wiki/Die_Ratten, letzter Zugriff: 02.12.2020.

Die Räuber:

https://de.wikipedia.org/wiki/Die_Räuber, letzter Zugriff: 02.12.2020.

Die Tageszeitung:

https://de.wikipedia.org/wiki/Die_Tageszeitung, letzter Zugriff: 02.12.2020.

Die Welt:

https://de.wikipedia.org/wiki/Die_Welt, letzter Zugriff: 02.12.2020.

Diktatur:

<https://de.wikipedia.org/wiki/Diktatur>, letzter Zugriff: 30.10.2020.

Direktorialsystem:

<https://de.wikipedia.org/wiki/Direktorialsystem>, letzter Zugriff: 30.10.2020.

Doppelte Buchführung:

https://de.wikipedia.org/wiki/Buchführung#Doppelte_Buchführung_-_Grundlagen,
letzter Zugriff: 24.11.2020.

Doppler-Effekt:

<https://de.wikipedia.org/wiki/Doppler-Effekt>, letzter Zugriff: 27.11.2020.

Drachenviereck:

<https://de.wikipedia.org/wiki/Drachenviereck>, letzter Zugriff: 11.11.2020.

Drama:

<https://de.wikipedia.org/wiki/Drama>, letzter Zugriff: 20.11.2020.

Dreiecksungleichung:

<https://de.wikipedia.org/wiki/Dreiecksungleichung>, letzter Zugriff: 26.11.2020.

Dunkle Materie:

https://de.wikipedia.org/wiki/Dunkle_Materie, letzter Zugriff: 17.11.2020.

Echolot:

<https://de.wikipedia.org/wiki/Echolot>, letzter Zugriff: 27.10.2020.

Einparteiensystem:

<https://de.wikipedia.org/wiki/Einparteiensystem>, letzter Zugriff: 30.10.2020.

Einsame Pinie:

https://de.wikipedia.org/wiki/Einsame_Pinie, letzter Zugriff: 30.11.2020.

Element (Mathematik) :

[https://de.wikipedia.org/wiki/Element_\(Mathematik\)](https://de.wikipedia.org/wiki/Element_(Mathematik)), letzter Zugriff: 20.10.2020.

Emergenz:

<https://de.wikipedia.org/wiki/Emergenz>, letzter Zugriff: 20.10.2020.

Empfänger:

[https://de.wikipedia.org/wiki/Empfänger_\(Information\)](https://de.wikipedia.org/wiki/Empfänger_(Information)), letzter Zugriff: 19.11.2020.

Energie:

<https://de.wikipedia.org/wiki/Energie>, letzter Zugriff: 12.11.2020.

Epik:

<https://de.wikipedia.org/wiki/Epik>, letzter Zugriff: 20.11.2020.

Erdbeeren:

<https://de.wikipedia.org/wiki/Erdbeeren> , letzter Zugriff: 20.10.2020.

Erste Lautverschiebung:

https://de.wikipedia.org/wiki/Erste_Lautverschiebung, letzter Zugriff: 27.11.2020.

Esther (Racine):

[https://de.wikipedia.org/wiki/Esther_\(Racine\)](https://de.wikipedia.org/wiki/Esther_(Racine)), letzter Zugriff: 02.12.2020.

Eukaryoten:

<https://de.wikipedia.org/wiki/Eukaryoten>, letzter Zugriff: 20.10.2020.

Eulersche Zahl:

https://de.wikipedia.org/wiki/Eulersche_Zahl, letzter Zugriff: 26.11.2020.

Eulersches Tonnetz:

https://de.wikipedia.org/wiki/Eulersches_Tonnetz, letzter Zugriff: 27.11.2020.

Europäische Union:

https://de.wikipedia.org/wiki/Europäische_Union, letzter Zugriff: 02.12.2020.

Europäischer Rat:

https://de.wikipedia.org/wiki/Europäischer_Rat, letzter Zugriff: 02.12.2020.

Evolution:

<https://de.wikipedia.org/wiki/Evolution>, letzter Zugriff: 20.10.2020.

Exoplanet:

<https://de.wikipedia.org/wiki/Exoplanet>, letzter Zugriff: 28.10.2020.

Exponentialfunktion:

<https://de.wikipedia.org/wiki/Exponentialfunktion>, letzter Zugriff: 11.11.2020.

Feld (Physik):

[https://de.wikipedia.org/wiki/Feld_\(Physik\)](https://de.wikipedia.org/wiki/Feld_(Physik)), letzter Zugriff: 28.10.2020.

Filmeditor:

<https://de.wikipedia.org/wiki/Filmeditor>, letzter Zugriff: 19.11.2020.

Filmschauspieler:

<https://de.wikipedia.org/wiki/Filmschauspieler>, letzter Zugriff: 19.11.2020.

Finanzbuchhaltung:

<https://de.wikipedia.org/wiki/Finanzbuchhaltung>, letzter Zugriff: 24.11.2020.

Finanzrisiko:

<https://de.wikipedia.org/wiki/Finanzrisiko>, letzter Zugriff: 09.11.2020.

Fixstern:

<https://de.wikipedia.org/wiki/Fixstern>, letzter Zugriff: 18.11.2020.

Folge (Mathematik):

[https://de.wikipedia.org/wiki/Folge_\(Mathematik\)](https://de.wikipedia.org/wiki/Folge_(Mathematik)), letzter Zugriff: 20.10.2020.

Fortpflanzung:

<https://de.wikipedia.org/wiki/Fortpflanzung>, letzter Zugriff: 20.10.2020.

Frankfurter Allgemeine Zeitung:

https://de.wikipedia.org/wiki/Frankfurter_Allgemeine_Zeitung, letzter Zugriff: 02.12.2020.

Freiäugig:

<https://de.wikipedia.org/wiki/Freiäugig>, letzter Zugriff: 18.11.2020.

Frequenz:

<https://de.wikipedia.org/wiki/Frequenz>, letzter Zugriff: 12.11.2020.

Fundamentale Wechselwirkung:

https://de.wikipedia.org/wiki/Fundamentale_Wechselwirkung, letzter Zugriff: 17.11.2020.

Funktion (Mathematik):

[https://de.wikipedia.org/wiki/Funktion_\(Mathematik\)](https://de.wikipedia.org/wiki/Funktion_(Mathematik)), letzter Zugriff: 11.11.2020.

Funktion (Organisation):

[https://de.wikipedia.org/wiki/Funktion_\(Organisation\)](https://de.wikipedia.org/wiki/Funktion_(Organisation)), letzter Zugriff: 24.11.2020.

Füssigkeit:

<https://de.wikipedia.org/wiki/Flüssigkeit>, letzter Zugriff: 20.10.2020.

Gala (Zeitschrift):

[https://de.wikipedia.org/wiki/Gala_\(Zeitschrift\)](https://de.wikipedia.org/wiki/Gala_(Zeitschrift)), letzter Zugriff: 02.12.2020.

Galaxie:

<https://de.wikipedia.org/wiki/Galaxie>, letzter Zugriff: 17.11.2020.

GameStar:

<https://de.wikipedia.org/wiki/GameStar>, letzter Zugriff: 02.12.2020.

Ganzrationale Funktion:

https://de.wikipedia.org/wiki/Ganzrationale_Funktion, letzter Zugriff: 11.11.2020.

Gas :

<https://de.wikipedia.org/wiki/Gas>, letzter Zugriff: 17.11.2020.

Gasplanet:

<https://de.wikipedia.org/wiki/Gasplanet>, letzter Zugriff: 17.11.2020.

Gattung (Biologie):

[https://de.wikipedia.org/wiki/Gattung_\(Biologie\)](https://de.wikipedia.org/wiki/Gattung_(Biologie)), letzter Zugriff: 20.10.2020.

Gattung (Literatur):

[https://de.wikipedia.org/wiki/Gattung_\(Literatur\)](https://de.wikipedia.org/wiki/Gattung_(Literatur)), letzter Zugriff: 20.11.2020.

Gaußsche Summenformel:

https://de.wikipedia.org/wiki/Gaußsche_Summenformel, letzter Zugriff: 26.11.2020.

Gebärdensprache:

<https://de.wikipedia.org/wiki/Gebärdensprache>, letzter Zugriff: 19.11.2020.

Gelber Riese:

https://de.wikipedia.org/wiki/Gelber_Riese, letzter Zugriff: 18.11.2020.

Gemisch:

<https://de.wikipedia.org/wiki/Gemisch>, letzter Zugriff: 20.10.2020.

Geometrische Figur:

https://de.wikipedia.org/wiki/Geometrische_Figur, letzter Zugriff: 11.11.2020.

Geräusch:

<https://de.wikipedia.org/wiki/Geräusch>, letzter Zugriff: 12.11.2020.

Gerhard Schröder:

https://de.wikipedia.org/wiki/Gerhard_Schröder, letzter Zugriff: 30.11.2020.

Geschäftsvorfall:

<https://de.wikipedia.org/wiki/Geschäftsvorfall>, letzter Zugriff: 24.11.2020.

Geschriebene Sprache:

https://de.wikipedia.org/wiki/Geschriebene_Sprache, letzter Zugriff: 19.11.2020.

Geschwindigkeit:

<https://de.wikipedia.org/wiki/Geschwindigkeit>, letzter Zugriff: 12.11.2020.

Gesetz der großen Zahlen:

https://de.wikipedia.org/wiki/Gesetz_der_großen_Zahlen, letzter Zugriff: 26.11.2020.

Gesetz:

<https://de.wikipedia.org/wiki/Gesetz>, letzter Zugriff: 24.11.2020.

Gesetzgebung:

<https://de.wikipedia.org/wiki/Gesetzgebung>, letzter Zugriff: 20.10.2020.

Gespenster (Ibsen):

[https://de.wikipedia.org/wiki/Gespenster_\(Ibsen\)](https://de.wikipedia.org/wiki/Gespenster_(Ibsen)), letzter Zugriff: 02.12.2020.

Gesprochene Sprache:

https://de.wikipedia.org/wiki/Gesprochene_Sprache, letzter Zugriff: 19.11.2020.

Gestik:

<https://de.wikipedia.org/wiki/Gestik>, letzter Zugriff: 19.11.2020.

Getränk:

<https://de.wikipedia.org/wiki/Getränk#Getränk>, letzter Zugriff: 20.10.2020.

Gewaltenteilung:

<https://de.wikipedia.org/wiki/Gewaltenteilung>, letzter Zugriff: 30.10.2020.

Gewinnchance:

<https://de.wikipedia.org/wiki/Gewinnchance>, letzter Zugriff: 09.11.2020.

Gewohnheitsrecht:

<https://de.wikipedia.org/wiki/Gewohnheitsrecht>, letzter Zugriff: 20.10.2020.

Gleichseitiges Polygon:

https://de.wikipedia.org/wiki/Gleichseitiges_Polygon, letzter Zugriff: 11.11.2020.

Goethes Faust:

https://de.wikipedia.org/wiki/Goethes_Faust, letzter Zugriff: 02.12.2020.

Gravitation :

<https://de.wikipedia.org/wiki/Gravitation>, letzter Zugriff: 17.11.2020.

Gravitationsfeld:

<https://de.wikipedia.org/wiki/Gravitationsfeld>, letzter Zugriff: 28.10.2020.

Great Barrier Reef:

https://de.wikipedia.org/wiki/Great_Barrier_Reef, letzter Zugriff: 30.11.2020.

Griechische Tragödie:

https://de.wikipedia.org/wiki/Griechische_Tragedie, letzter Zugriff: 20.11.2020.

Größe (Mathematik):

[https://de.wikipedia.org/wiki/Größe_\(Mathematik\)](https://de.wikipedia.org/wiki/Größe_(Mathematik)), letzter Zugriff: 24.11.2020.

Grundfrequenz:

<https://de.wikipedia.org/wiki/Grundfrequenz>, letzter Zugriff: 12.11.2020.

Gruppe (Mathematik):

[https://de.wikipedia.org/wiki/Gruppe_\(Mathematik\)](https://de.wikipedia.org/wiki/Gruppe_(Mathematik)), letzter Zugriff: 20.10.2020.

Hagebutten:

<https://de.wikipedia.org/wiki/Hagebutte>, letzter Zugriff: 20.10.2020.

Halluzination:

<https://de.wikipedia.org/wiki/Halluzination>, letzter Zugriff: 20.10.2020.

Halluzinogen:

<https://de.wikipedia.org/wiki/Halluzinogen>, letzter Zugriff: 29.10.2020.

Hamlet:

<https://de.wikipedia.org/wiki/Hamlet>, letzter Zugriff: 02.12.2020.

Handeln:

<https://de.wikipedia.org/wiki/Handeln>, letzter Zugriff: 09.11.2020.

Handlungsformen der Verwaltung:

https://de.wikipedia.org/wiki/Handlungsformen_der_Verwaltung, letzter Zugriff: 24.11.2020.

Hauptreihe:

<https://de.wikipedia.org/wiki/Hauptreihe>, letzter Zugriff: 18.11.2020.

Herrschaftsform:

<https://de.wikipedia.org/wiki/Herrschaftsform>, letzter Zugriff: 30.10.2020.

Himbeere:

<https://de.wikipedia.org/wiki/Himbeere>, letzter Zugriff: 20.10.2020.

Hoheitsakt:

<https://de.wikipedia.org/wiki/Hoheitsakt>, letzter Zugriff: 24.11.2020.

Humankompetenz:

<https://de.wikipedia.org/wiki/Humankompetenz>, letzter Zugriff: 18.11.2020.

Hybris:

<https://de.wikipedia.org/wiki/Hybris>, letzter Zugriff: 20.11.2020.

Hydrostatisches Gleichgewicht:

https://de.wikipedia.org/wiki/Hydrostatisches_Gleichgewicht#Astrophysik, letzter Zugriff: 28.10.2020.

Ich-Störung:

<https://de.wikipedia.org/wiki/Ich-Störung>, letzter Zugriff: 20.10.2020.

Information:

<https://de.wikipedia.org/wiki/Information>, letzter Zugriff: 10.11.2020.

Instrumentelle und operante Konditionierung:

https://de.wikipedia.org/wiki/Instrumentelle_und_operante_Konditionierung, letzter Zugriff: 18.11.2020.

Interplanetarer Staub:

https://de.wikipedia.org/wiki/Interplanetarer_Staub, letzter Zugriff: 17.11.2020.

Interpretation (Musik):

[https://de.wikipedia.org/wiki/Interpretation_\(Musik\)](https://de.wikipedia.org/wiki/Interpretation_(Musik)), letzter Zugriff: 19.11.2020.

Interstellarer Staub:

https://de.wikipedia.org/wiki/Interstellarer_Staub, letzter Zugriff: 17.11.2020.

Investigativer Journalismus:

https://de.wikipedia.org/wiki/Investigativer_Journalismus, letzter Zugriff: 19.11.2020.

Jackie Chan:

https://de.wikipedia.org/wiki/Jackie_Chan, letzter Zugriff: 30.11.2020.

Journalist:

<https://de.wikipedia.org/wiki/Journalist>, letzter Zugriff: 19.11.2020.

Juristische Person:

https://de.wikipedia.org/wiki/Juristische_Person, letzter Zugriff: 24.11.2020.

Juristische Person des öffentlichen Rechts:

https://de.wikipedia.org/wiki/Juristische_Person#Juristische_Person_des_öffentlichen_Rechts, letzter Zugriff: 09.11.2020.

Kameralistik:

<https://de.wikipedia.org/wiki/Kameralistik>, letzter Zugriff: 24.11.2020.

Kartesisches Produkt:

https://de.wikipedia.org/wiki/Kartesisches_Produkt, letzter Zugriff: 11.11.2020.

Keira Knightley:

https://de.wikipedia.org/wiki/Keira_Knightley, letzter Zugriff: 30.11.2020.

Kernobstgewächse:

<https://de.wikipedia.org/wiki/Kernobstgewächse>, letzter Zugriff: 20.10.2020.

Klangfarbe:

<https://de.wikipedia.org/wiki/Klangfarbe>, letzter Zugriff: 12.11.2020.

Klasse (Biologie):

[https://de.wikipedia.org/wiki/Klasse_\(Biologie\)](https://de.wikipedia.org/wiki/Klasse_(Biologie)), letzter Zugriff: 20.10.2020.

Klassische Konditionierung:

https://de.wikipedia.org/wiki/Klassische_Konditionierung, letzter Zugriff: 18.11.2020.

Klavier:

<https://de.wikipedia.org/wiki/Klavier>, letzter Zugriff: 20.11.2020.

Kleinkörper (Astronomie):

[https://de.wikipedia.org/wiki/Kleinkörper_\(Astronomie\)](https://de.wikipedia.org/wiki/Kleinkörper_(Astronomie)), letzter Zugriff: 17.11.2020.

Kugelbaum:

<https://de.wikipedia.org/wiki/Kugelbaum>, letzter Zugriff: 30.11.2020.

Kognition:

<https://de.wikipedia.org/wiki/Kognition>, letzter Zugriff: 18.11.2020.

Kokain:

<https://de.wikipedia.org/wiki/Kokain>, letzter Zugriff: 29.10.2020.

Kolumne:

<https://de.wikipedia.org/wiki/Kolumne>, letzter Zugriff: 19.11.2020.

Komet:

<https://de.wikipedia.org/wiki/Komet>, letzter Zugriff: 17.11.2020.

Kommunalunternehmen:

<https://de.wikipedia.org/wiki/Kommunalunternehmen>, letzter Zugriff: 09.11.2020.

Kommunalverwaltung:

<https://de.wikipedia.org/wiki/Kommunalverwaltung>, letzter Zugriff: 09.11.2020.

Kommunikation:

<https://de.wikipedia.org/wiki/Kommunikation>, letzter Zugriff: 10.11.2020.

Kommunikationsmittel:

<https://de.wikipedia.org/wiki/Kommunikationsmittel>, letzter Zugriff: 18.11.2020.

Komödie:

<https://de.wikipedia.org/wiki/Komödie>, letzter Zugriff: 20.11.2020.

Kompetenz (Psychologie):

[https://de.wikipedia.org/wiki/Kompetenz_\(Psychologie\)](https://de.wikipedia.org/wiki/Kompetenz_(Psychologie)), letzter Zugriff: 18.11.2020.

Konditionierung:

<https://de.wikipedia.org/wiki/Konditionierung>, letzter Zugriff: 18.11.2020.

Kongeeegen:

<https://de.wikipedia.org/wiki/Kongeeegen>, letzter Zugriff: 30.11.2020.

Königreich England:

https://de.wikipedia.org/wiki/Königreich_England, letzter Zugriff: 02.12.2020.

Konstitutionelle Monarchie:

https://de.wikipedia.org/wiki/Konstitutionelle_Monarchie, letzter Zugriff: 30.10.2020.

Kontrolle:

<https://de.wikipedia.org/wiki/Kontrolle>, letzter Zugriff: 24.11.2020.

Konvexe Menge:

https://de.wikipedia.org/wiki/Konvexe_Menge, letzter Zugriff: 11.11.2020.

Konzert (Veranstaltung):

[https://de.wikipedia.org/wiki/Konzert_\(Veranstaltung\)](https://de.wikipedia.org/wiki/Konzert_(Veranstaltung)), letzter Zugriff: 19.11.2020.

Körperschaft des privaten Rechts:

https://de.wikipedia.org/wiki/Körperschaft_des_privaten_Rechts, letzter Zugriff: 09.11.2020.

Körperschall:

<https://de.wikipedia.org/wiki/Körperschall>, letzter Zugriff: 22.10.2020.

Korrepitor:

<https://de.wikipedia.org/wiki/Korrepetitor>, letzter Zugriff: 20.11.2020.

Korrespondent:

<https://de.wikipedia.org/wiki/Korrespondent>, letzter Zugriff: 19.11.2020.

Kosmischer Staub:

https://de.wikipedia.org/wiki/Kosmischer_Staub, letzter Zugriff: 17.11.2020.

Kraftfeld:

<https://de.wikipedia.org/wiki/Kraftfeld>, letzter Zugriff: 28.10.2020.

Kreiszahl:

<https://de.wikipedia.org/wiki/Kreiszahl>, letzter Zugriff: 26.11.2020.

Kreuzung (Genetik) :

[https://de.wikipedia.org/wiki/Kreuzung_\(Genetik\)](https://de.wikipedia.org/wiki/Kreuzung_(Genetik)), letzter Zugriff: 20.10.2020.

Kuipergürtel:

<https://de.wikipedia.org/wiki/Kuipergürtel>, letzter Zugriff: 30.11.2020.

Kulturapfel:

<https://de.wikipedia.org/wiki/Kulturapfel>, letzter Zugriff: 20.10.2020.

Kulturrosen:

<https://de.wikipedia.org/wiki/Kulturrosen>, letzter Zugriff: 20.10.2020.

Kunst:

<https://de.wikipedia.org/wiki/Kunst>, letzter Zugriff: 10.11.2020.

Laiendarsteller:

<https://de.wikipedia.org/wiki/Laiendarsteller>, letzter Zugriff: 19.11.2020.

Lautstärke:

<https://de.wikipedia.org/wiki/Lautstärke>, letzter Zugriff: 12.11.2020.

Lebewesen:

<https://de.wikipedia.org/wiki/Lebewesen> , letzter Zugriff: 20.10.2020.

Legislative:

<https://de.wikipedia.org/wiki/Legislative>, letzter Zugriff: 24.11.2020.

Lernen :

<https://de.wikipedia.org/wiki/Lernen>, letzter Zugriff: 18.11.2020 .

Lerntheorie:

<https://de.wikipedia.org/wiki/Lerntheorie>, letzter Zugriff: 18.11.2020.

Leuchte:

<https://de.wikipedia.org/wiki/Leuchte>, letzter Zugriff: 20.11.2020.

Lichttechnik:

<https://de.wikipedia.org/wiki/Lichttechnik>, letzter Zugriff: 20.11.2020.

Lineare Funktion:

https://de.wikipedia.org/wiki/Lineare_Funktion, letzter Zugriff: 11.11.2020.

Logarithmus:

<https://de.wikipedia.org/wiki/Logarithmus>, letzter Zugriff: 11.11.2020.

London:

<https://de.wikipedia.org/wiki/London>, letzter Zugriff: 02.12.2020.

Luftschall:

<https://de.wikipedia.org/wiki/Luftschall>, letzter Zugriff: 22.10.2020.

Lyrik:

<https://de.wikipedia.org/wiki/Lyrik>, letzter Zugriff: 20.11.2020.

Mandelbaum:

<https://de.wikipedia.org/wiki/Mandelbaum>, letzter Zugriff: 20.10.2020.

Marktrisiko:

<https://de.wikipedia.org/wiki/Marktrisiko>, letzter Zugriff: 09.11.2020.

Mars (Planet):

[https://de.wikipedia.org/wiki/Mars_\(Planet\)](https://de.wikipedia.org/wiki/Mars_(Planet)), letzter Zugriff: 30.11.2020.

Masse (Physik):

[https://de.wikipedia.org/wiki/Masse_\(Physik\)](https://de.wikipedia.org/wiki/Masse_(Physik)), letzter Zugriff: 18.11.2020.

Massenspeicher:

<https://de.wikipedia.org/wiki/Massenspeicher>, letzter Zugriff: 18.11.2020.

Materie (Physik):

[https://de.wikipedia.org/wiki/Materie_\(Physik\)](https://de.wikipedia.org/wiki/Materie_(Physik)), letzter Zugriff: 20.10.2020.

Mechanisches Gleichgewicht:
https://de.wikipedia.org/wiki/Mechanisches_Gleichgewicht, letzter Zugriff: 28.10.2020.

Medium (Kommunikation):
[https://de.wikipedia.org/wiki/Medium_\(Kommunikation\)](https://de.wikipedia.org/wiki/Medium_(Kommunikation)), letzter Zugriff: 18.11.2020.

Menge (Mathematik):
[https://de.wikipedia.org/wiki/Menge_\(Mathematik\)](https://de.wikipedia.org/wiki/Menge_(Mathematik)), letzter Zugriff: 20.10.2020.

Mensch :
<https://de.wikipedia.org/wiki/Mensch>, letzter Zugriff: 19.11.2020.

Menschliche Stimme:
https://de.wikipedia.org/wiki/Menschliche_Stimme, letzter Zugriff: 12.11.2020.

Meteorit:
<https://de.wikipedia.org/wiki/Meteoroid>, letzter Zugriff: 17.11.2020.

Michelin:
<https://de.wikipedia.org/wiki/Michelin>, letzter Zugriff: 02.12.2020.

Milchstraße:
<https://de.wikipedia.org/wiki/MilchstraÙe>, letzter Zugriff: 30.11.2020.

Militärdiktatur:
<https://de.wikipedia.org/wiki/Militärdiktatur>, letzter Zugriff: 30.10.2020.

Mimik:
<https://de.wikipedia.org/wiki/Mimik>, letzter Zugriff: 19.11.2020.

Modejournalismus:
<https://de.wikipedia.org/wiki/Modejournalismus>, letzter Zugriff: 19.11.2020.

Moderator:
<https://de.wikipedia.org/wiki/Moderator>, letzter Zugriff: 19.11.2020.

Monarchie:
<https://de.wikipedia.org/wiki/Monarchie>, letzter Zugriff: 30.10.2020.

Musikinstrument:
<https://de.wikipedia.org/wiki/Musikinstrument>, letzter Zugriff: 20.11.2020.

Musiktheater:
<https://de.wikipedia.org/wiki/Musiktheater>, letzter Zugriff: 10.11.2020.

Muster (Struktur):
[https://de.wikipedia.org/wiki/Muster_\(Struktur\)](https://de.wikipedia.org/wiki/Muster_(Struktur)), letzter Zugriff: 27.10.2020.

Nachricht:
<https://de.wikipedia.org/wiki/Nachricht>, letzter Zugriff: 19.11.2020.

NATO:
<https://de.wikipedia.org/wiki/NATO>, letzter Zugriff: 02.12.2020.

Natürliche Sprache:
https://de.wikipedia.org/wiki/Natürliche_Sprache, letzter Zugriff: 19.11.2020.

Naturrecht:

<https://de.wikipedia.org/wiki/Naturrecht>, letzter Zugriff: 20.10.2020.

Nebel (Astronomie):

[https://de.wikipedia.org/wiki/Nebel_\(Astronomie\)](https://de.wikipedia.org/wiki/Nebel_(Astronomie)), letzter Zugriff: 17.11.2020.

Neue psychoaktive Substanzen:

https://de.wikipedia.org/wiki/Neue_psychoaktive_Substanzen, letzter Zugriff: 29.10.2020.

New Horizons:

https://de.wikipedia.org/wiki/New_Horizons, letzter Zugriff: 30.11.2020.

Nomura Holdings:

https://de.wikipedia.org/wiki/Nomura_Holdings, letzter Zugriff: 02.12.2020.

Nonverbale Kommunikation:

https://de.wikipedia.org/wiki/Nonverbale_Kommunikation, letzter Zugriff: 10.11.2020.

Norah Jones:

https://de.wikipedia.org/wiki/Norah_Jones, letzter Zugriff: 30.11.2020.

Nordrhein-Westfalen:

<https://de.wikipedia.org/wiki/Nordrhein-Westfalen>, letzter Zugriff: 02.12.2020.

Nutzschall:

<https://de.wikipedia.org/wiki/Nutzschall>, letzter Zugriff: 22.10.2020.

Oberton:

<https://de.wikipedia.org/wiki/Oberton>, letzter Zugriff: 12.11.2020.

Objekt planetarer Masse:

https://de.wikipedia.org/wiki/Objekt_planetarer_Masse, letzter Zugriff: 28.10.2020.

Öffentliches Unternehmen:

https://de.wikipedia.org/wiki/Öffentliches_Unternehmen, letzter Zugriff: 09.11.2020.

Old Tjiko:

https://de.wikipedia.org/wiki/Old_Tjikko, letzter Zugriff: 30.11.2020.

Optischer Massenspeicher:

https://de.wikipedia.org/wiki/Optischer_Datenspeicher, letzter Zugriff: 18.11.2020.

Ordnung (Biologie):

[https://de.wikipedia.org/wiki/Ordnung_\(Biologie\)](https://de.wikipedia.org/wiki/Ordnung_(Biologie)), letzter Zugriff: 20.10.2020.

Organisation:

<https://de.wikipedia.org/wiki/Organisation>, letzter Zugriff: 30.10.2020.

Organisationseinheit:

<https://de.wikipedia.org/wiki/Organisationseinheit>, letzter Zugriff: 09.11.2020.

Ortsvektor:

<https://de.wikipedia.org/wiki/Ortsvektor>, letzter Zugriff: 11.11.2020.

Pando (Baum):

[https://de.wikipedia.org/wiki/Pando_\(Baum\)](https://de.wikipedia.org/wiki/Pando_(Baum)), letzter Zugriff: 30.11.2020.

Parallelität:

[https://de.wikipedia.org/wiki/Parallelität_\(Geometrie\)](https://de.wikipedia.org/wiki/Parallelität_(Geometrie)), letzter Zugriff: 11.11.2020.

Parallelogram:

<https://de.wikipedia.org/wiki/Parallelogramm>, letzter Zugriff: 11.11.2020.

Parameter (Musik):

[https://de.wikipedia.org/wiki/Parameter_\(Musik\)](https://de.wikipedia.org/wiki/Parameter_(Musik)), letzter Zugriff: 12.11.2020.

Parasprache:

<https://de.wikipedia.org/wiki/Parasprache>, letzter Zugriff: 19.11.2020.

Paraverbale Kommunikation:

https://de.wikipedia.org/wiki/Paraverbale_Kommunikation, letzter Zugriff: 10.11.2017.

Parlamentarisches Regierungssystem:

https://de.wikipedia.org/wiki/Parlamentarisches_Regierungssystem, letzter Zugriff: 30.10.2020.

Pascal (Einheit):

[https://de.wikipedia.org/wiki/Pascal_\(Einheit\)](https://de.wikipedia.org/wiki/Pascal_(Einheit)), letzter Zugriff: 27.11.2020.

Permanentspeichermedium:

<https://de.wikipedia.org/wiki/Permanentspeichermedium>, letzter Zugriff: 18.11.2020.

Person :

<https://de.wikipedia.org/wiki/Person>, letzter Zugriff: 19.11.2020.

Personenvereinigung:

<https://de.wikipedia.org/wiki/Personenvereinigung>, letzter Zugriff: 09.11.2020.

Persönlichkeitsstörung:

<https://de.wikipedia.org/wiki/Pers>

Pflanze:

<https://de.wikipedia.org/wiki/Pflanze>, letzter Zugriff: 20.10.2020.

Pflanzenwuchsform:

<https://de.wikipedia.org/wiki/Pflanzenwuchsform>, letzter Zugriff: 20.10.2020.

Pflaume:

<https://de.wikipedia.org/wiki/Pflaume>, letzter Zugriff: 20.10.2020.

Phenylethylamin:

<https://de.wikipedia.org/wiki/Phenylethylamine>, letzter Zugriff: 20.10.2020.

Photosynthese:

<https://de.wikipedia.org/wiki/Photosynthese>, letzter Zugriff: 20.10.2020.

Physikalische Größe:

https://de.wikipedia.org/wiki/Physikalische_Größe, letzter Zugriff: 20.10.2020.

Pilz:

<https://de.wikipedia.org/wiki/Pilze>, letzter Zugriff: 20.10.2020.

Planet:

<https://de.wikipedia.org/wiki/Planet>, letzter Zugriff: 20.10.2020.

Planetensystem:

<https://de.wikipedia.org/wiki/Planetensystem>, letzter Zugriff: 17.11.2020.

Planung:

<https://de.wikipedia.org/wiki/Planung>, letzter Zugriff: 24.11.2020.

Platte :

[https://de.wikipedia.org/wiki/Platte_\(Technische_Mechanik\)](https://de.wikipedia.org/wiki/Platte_(Technische_Mechanik)), letzter Zugriff: 27.10.2020.

Pluto:

<https://de.wikipedia.org/wiki/Pluto>, letzter Zugriff: 30.11.2020.

Poincaré-Vermutung:

<https://de.wikipedia.org/wiki/Poincaré-Vermutung>, letzter Zugriff: 26.11.2020.

Politische Partei:

https://de.wikipedia.org/wiki/Politische_Partei, letzter Zugriff: 24.11.2020.

Politisches System:

https://de.wikipedia.org/wiki/Politisches_System, letzter Zugriff: 30.10.2020.

Polygon:

<https://de.wikipedia.org/wiki/Polygon>, letzter Zugriff: 11.11.2020.

Polygonzug:

[https://de.wikipedia.org/wiki/Polygonzug_\(Mathematik\)](https://de.wikipedia.org/wiki/Polygonzug_(Mathematik)), letzter Zugriff: 11.11.2020.

Polynom:

<https://de.wikipedia.org/wiki/Polynom>, letzter Zugriff: 20.10.2020.

Positives Recht:

https://de.wikipedia.org/wiki/Positives_Recht, letzter Zugriff: 20.10.2020.

Potenz (Mathematik):

[https://de.wikipedia.org/wiki/Potenz_\(Mathematik\)](https://de.wikipedia.org/wiki/Potenz_(Mathematik)), letzter Zugriff: 20.10.2020.

Potenzfunktion:

<https://de.wikipedia.org/wiki/Potenzfunktion>, letzter Zugriff: 11.11.2020.

Präsidentielles Regierungssystem:

https://de.wikipedia.org/wiki/Präsidentielles_Regierungssystem, letzter Zugriff: 30.10.2020.

Pressefotograf:

<https://de.wikipedia.org/wiki/Pressefotograf>, letzter Zugriff: 19.11.2020.

Problemlösen:

<https://de.wikipedia.org/wiki/Problemlösen>, letzter Zugriff: 18.11.2020.

Protokollarische Rangordnung:

https://de.wikipedia.org/wiki/Protokollarische_Rangordnung, letzter Zugriff: 24.11.2020.

Prozess:

<https://de.wikipedia.org/wiki/Prozess> , letzter Zugriff: 20.10.2020.

Prüfen:

<https://de.wikipedia.org/wiki/Pr>

Prunus:

<https://de.wikipedia.org/wiki/Prunus> , letzter Zugriff: 20.10.2020.

Psychische Störung:

https://de.wikipedia.org/wiki/Psychische_Störung, letzter Zugriff: 18.11.2020.

Psychose:

<https://de.wikipedia.org/wiki/Psychose>, letzter Zugriff: 20.10.2020.

Psychotherapie:

<https://de.wikipedia.org/wiki/Psychotherapie>, letzter Zugriff: 18.11.2020.

Psychotrope Substanz:

https://de.wikipedia.org/wiki/Psychotrope_Substanz, letzter Zugriff: 29.10.2020.

Publikum:

<https://de.wikipedia.org/wiki/Publikum>, letzter Zugriff: 10.11.2020.

Puppentheater:

<https://de.wikipedia.org/wiki/Puppentheater>, letzter Zugriff: 10.11.2020.

Quadrat :

<https://de.wikipedia.org/wiki/Quadrat>, letzter Zugriff: 11.11.2020.

Quadratitsche Funktion:

https://de.wikipedia.org/wiki/Quadratische_Funktion, letzter Zugriff: 11.11.2020.

Räterepublik:

<https://de.wikipedia.org/wiki/Räterepublik>, letzter Zugriff: 30.10.2020.

Rationale Funktion:

https://de.wikipedia.org/wiki/Rationale_Funktion, letzter Zugriff: 11.11.2020.

Raute :

<https://de.wikipedia.org/wiki/Raute>, letzter Zugriff: 11.11.2020.

Realitätsverlust:

<https://de.wikipedia.org/wiki/Realitätsverlust>, letzter Zugriff: 20.10.2020.

Recherche:

<https://de.wikipedia.org/wiki/Recherche>, letzter Zugriff: 19.11.2020.

Rechnungswesen:

<https://de.wikipedia.org/wiki/Rechnungswesen>, letzter Zugriff: 24.11.2020.

Recht:

<https://de.wikipedia.org/wiki/Recht>, letzter Zugriff: 20.10.2020.

Rechteck:

<https://de.wikipedia.org/wiki/Rechteck>, letzter Zugriff: 11.11.2020.

Rechter Winkel:

https://de.wikipedia.org/wiki/Rechter_Winkel, letzter Zugriff: 11.11.2020.

Rechtsfolge:

<https://de.wikipedia.org/wiki/Rechtsfolge>, letzter Zugriff: 24.11.2020.

Rechtsordnung:

<https://de.wikipedia.org/wiki/Rechtsordnung>, letzter Zugriff: 24.11.2020.

Redakteur:

<https://de.wikipedia.org/wiki/Redakteur>, letzter Zugriff: 19.11.2020.

Reelwertige Funktion:

https://de.wikipedia.org/wiki/Reellwertige_Funktion, letzter Zugriff: 11.11.2020.

Regelmäßiges Polygon:

https://de.wikipedia.org/wiki/Regelmäßiges_Polygon, letzter Zugriff: 11.11.2020.

Regierungschef:

<https://de.wikipedia.org/wiki/Regierungschef>, letzter Zugriff: 30.10.2020.

Regierungssystem:

<https://de.wikipedia.org/wiki/Regierungssystem>, letzter Zugriff: 30.10.2020.

Regina Halmich:

https://de.wikipedia.org/wiki/Regina_Halmich, letzter Zugriff: 30.11.2020.

Regisseur:

<https://de.wikipedia.org/wiki/Regisseur>, letzter Zugriff: 20.11.2020.

Reine Stimmung:

https://de.wikipedia.org/wiki/Reine_Stimmung, letzter Zugriff: 27.11.2020.

Reinstoff:

<https://de.wikipedia.org/wiki/Reinstoff>, letzter Zugriff: 20.10.2020.

Reizbarkeit:

<https://de.wikipedia.org/wiki/Reiz>, letzter Zugriff: 20.10.2020.

Relation (Mathematik):

[https://de.wikipedia.org/wiki/Relation_\(Mathematik\)](https://de.wikipedia.org/wiki/Relation_(Mathematik)), letzter Zugriff: 11.11.2020.

Reporter:

<https://de.wikipedia.org/wiki/Reporter>, letzter Zugriff: 19.11.2020.

Repräsentation:

[https://de.wikipedia.org/wiki/Repräsentation_\(Politik\)](https://de.wikipedia.org/wiki/Repräsentation_(Politik)), letzter Zugriff: 24.11.2020.

Republik:

<https://de.wikipedia.org/wiki/Republik>, letzter Zugriff: 24.11.2020.

Requisit:

<https://de.wikipedia.org/wiki/Requisit>, letzter Zugriff: 20.11.2020.

Rewe (Markt):

[https://de.wikipedia.org/wiki/Rewe_\(Markt\)](https://de.wikipedia.org/wiki/Rewe_(Markt)), letzter Zugriff: 02.12.2020.

Riesenplanet:

<https://de.wikipedia.org/wiki/Riesenplanet>, letzter Zugriff: 17.11.2020.

Riesensterne:

<https://de.wikipedia.org/wiki/Riesensterne>, letzter Zugriff: 18.11.2020.

Risiko:

<https://de.wikipedia.org/wiki/Risiko>, letzter Zugriff: 09.11.2020.

Rolle (Theater):

[https://de.wikipedia.org/wiki/Rolle_\(Theater\)](https://de.wikipedia.org/wiki/Rolle_(Theater)), letzter Zugriff: 19.11.2020.

Rosen:

<https://de.wikipedia.org/wiki/Rosen>, letzter Zugriff: 20.10.2020.

Rosenartige:

<https://de.wikipedia.org/wiki/Rosenartige>, letzter Zugriff: 20.10.2020.

Rosengewächse:

<https://de.wikipedia.org/wiki/Rosengewächse>, letzter Zugriff: 20.10.2020.

Rosenklasse:

<https://de.wikipedia.org/wiki/Rosenklassen>, letzter Zugriff: 20.10.2020.

Rosoideae:

<https://de.wikipedia.org/wiki/Rosoideae>, letzter Zugriff: 20.10.2020.

Roter Riese:

https://de.wikipedia.org/wiki/Roter_Riese, letzter Zugriff: 18.11.2020.

Roter Überriese:

https://de.wikipedia.org/wiki/Roter_Überriese, letzter Zugriff: 18.11.2020.

Roter Zwerg:

https://de.wikipedia.org/wiki/Roter_Zwerg, letzter Zugriff: 18.11.2020.

Rubus:

<https://de.wikipedia.org/wiki/Rubus>, letzter Zugriff: 20.10.2020.

Russellsche Antinomie:

https://de.wikipedia.org/wiki/Russellsche_Antinomie, letzter Zugriff: 26.11.2020.

Saiteninstrument:

<https://de.wikipedia.org/wiki/Saiteninstrument>, letzter Zugriff: 20.11.2020.

Same (Pflanze):

[https://de.wikipedia.org/wiki/Same_\(Pflanze\)](https://de.wikipedia.org/wiki/Same_(Pflanze)), letzter Zugriff: 20.10.2020.

Samenpflanzen:

<https://de.wikipedia.org/wiki/Samenpflanzen>, letzter Zugriff: 20.10.2020.

Sammelnussfrucht:

<https://de.wikipedia.org/wiki/Sammelnussfrucht>, letzter Zugriff: 20.10.2020.

Sand:

<https://de.wikipedia.org/wiki/Sand>, letzter Zugriff: 27.10.2020.

Satellit (Astronomie):

[https://de.wikipedia.org/wiki/Satellit_\(Astronomie\)](https://de.wikipedia.org/wiki/Satellit_(Astronomie)), letzter Zugriff: 17.11.2020.

Satz des Pythagoras:

https://de.wikipedia.org/wiki/Satz_des_Pythagoras, letzter Zugriff: 26.11.2020.

Schall:

<https://de.wikipedia.org/wiki/Schall>, letzter Zugriff: 22.10.2020.

Schalldruck:

<https://de.wikipedia.org/wiki/Schalldruck>, letzter Zugriff: 12.11.2020.

Schalldruckpegel:

<https://de.wikipedia.org/wiki/Schalldruckpegel>, letzter Zugriff: 12.11.2020.

Schallenergie:

<https://de.wikipedia.org/wiki/Schallenergie>, letzter Zugriff: 12.11.2020.

Schallereignis:

<https://de.wikipedia.org/wiki/Schallereignis>, letzter Zugriff: 12.11.2020.

Schallfeld :

<https://de.wikipedia.org/wiki/Schallfeld>, letzter Zugriff: 12.11.2020.

Schallfeldgröße:

<https://de.wikipedia.org/wiki/Schallfeldgröße>, letzter Zugriff: 12.11.2020.

Schallfluss:

<https://de.wikipedia.org/wiki/Schallfluss>, letzter Zugriff: 12.11.2020.

Schallgeschwindigkeit:

<https://de.wikipedia.org/wiki/Schallgeschwindigkeit>, letzter Zugriff: 12.11.2020.

Schallintensität:

<https://de.wikipedia.org/wiki/Schallintensität>, letzter Zugriff: 12.11.2020.

Schallkennimpedanz:

<https://de.wikipedia.org/wiki/Schallkennimpedanz>, letzter Zugriff: 12.11.2020.

Schallleistung:

<https://de.wikipedia.org/wiki/Schallleistung>, letzter Zugriff: 12.11.2020.

Schallschnelle:

<https://de.wikipedia.org/wiki/Schallschnelle>, letzter Zugriff: 12.11.2020.

Schallschnelleamplitude:

<https://de.wikipedia.org/wiki/Schallschnelleamplitude>, letzter Zugriff: 12.11.2020.

Schallstrahlungsdruck:

<https://de.wikipedia.org/wiki/Schallstrahlungsdruck>, letzter Zugriff: 12.11.2020.

Schauspiel:

<https://de.wikipedia.org/wiki/Schauspiel>, letzter Zugriff: 10.11.2020.

Schauspieler:

<https://de.wikipedia.org/wiki/Schauspieler>, letzter Zugriff: 19.11.2020.

Schauspielschule:

<https://de.wikipedia.org/wiki/Schauspielschule>, letzter Zugriff: 19.11.2020.

Scheinfrucht:

<https://de.wikipedia.org/wiki/Scheinfrucht>, letzter Zugriff: 20.10.2020.

Schild (Sternbild):

[https://de.wikipedia.org/wiki/Schild_\(Sternbild\)](https://de.wikipedia.org/wiki/Schild_(Sternbild)), letzter Zugriff: 30.11.2020.

Schwingung:

<https://de.wikipedia.org/wiki/Schwingung>, letzter Zugriff: 27.10.2020.

Sedimentation:

<https://de.wikipedia.org/wiki/Sedimentation>, letzter Zugriff: 27.10.2020.

Selbstlernkompetenz:

<https://de.wikipedia.org/wiki/Selbstlernkompetenz>, letzter Zugriff: 18.11.2020.

Selbstmotivation:

<https://de.wikipedia.org/wiki/Selbstmotivation>, letzter Zugriff: 18.11.2020.

Selbstreguliertes Lernen:

https://de.wikipedia.org/wiki/Selbstreguliertes_Lernen, letzter Zugriff: 18.11.2020.

Selbstüberschätzung:

<https://de.wikipedia.org/wiki/Selbstüberschätzung>, letzter Zugriff: 20.11.2020.

Semipräsidentielles Regierungssystem:

https://de.wikipedia.org/wiki/Semipräsidentielles_Regierungssystem, letzter Zugriff: 30.10.2020.

Signalisierung:

<https://de.wikipedia.org/wiki/Signalisierung>, letzter Zugriff: 20.11.2020.

Sinuston:

<https://de.wikipedia.org/wiki/Sinuston>, letzter Zugriff: 12.11.2020.

Skalar (Mathematik):

[https://de.wikipedia.org/wiki/Skalar_\(Mathematik\)](https://de.wikipedia.org/wiki/Skalar_(Mathematik)), letzter Zugriff: 11.11.2020.

Soffleur:

<https://de.wikipedia.org/wiki/Souffleur>, letzter Zugriff: 20.11.2020.

Sone:

<https://de.wikipedia.org/wiki/Sone>, letzter Zugriff: 27.11.2020.

Sonne:

<https://de.wikipedia.org/wiki/Sonne>, letzter Zugriff: 20.10.2020.

Sonnensystem:

<https://de.wikipedia.org/wiki/Sonnensystem>, letzter Zugriff: 17.11.2020.

Soziales System:

https://de.wikipedia.org/wiki/Soziales_System, letzter Zugriff: 30.10.2020.

Soziotechnisches System:

https://de.wikipedia.org/wiki/Soziotechnisches_System, letzter Zugriff: 30.10.2020.

Sparkasse KölnBonn:

https://de.wikipedia.org/wiki/Sparkasse_KölnBonn, letzter Zugriff: 02.12.2020.

Spektrum der Wissenschaft:

https://de.wikipedia.org/wiki/Spektrum_der_Wissenschaft, letzter Zugriff: 02.12.2020.

Sprache :

<https://de.wikipedia.org/wiki/Sprache>, letzter Zugriff: 19.11.2020.

Sprechen:

<https://de.wikipedia.org/wiki/Sprechen>, letzter Zugriff: 19.11.2020.

Spross:

<https://de.wikipedia.org/wiki/Spross>, letzter Zugriff: 20.10.2020.

Staat:

<https://de.wikipedia.org/wiki/Staat>, letzter Zugriff: 30.10.2020.

Staatsform:

<https://de.wikipedia.org/wiki/Staatsform>, letzter Zugriff: 24.11.2020.

Staatsoberhaupt:

<https://de.wikipedia.org/wiki/Staatsoberhaupt>, letzter Zugriff: 30.10.2020.

Staatspräsident:

<https://de.wikipedia.org/wiki/Staatspräsident>, letzter Zugriff: 24.11.2020.

Stachel (Botanik):

[https://de.wikipedia.org/wiki/Stachel_\(Botanik\)](https://de.wikipedia.org/wiki/Stachel_(Botanik)), letzter Zugriff: 20.10.2020.

Staub:

<https://de.wikipedia.org/wiki/Staub>, letzter Zugriff: 17.11.2020.

Stefan Raab:

https://de.wikipedia.org/wiki/Stefan_Raab, letzter Zugriff: 30.11.2020.

Steinobst:

<https://de.wikipedia.org/wiki/Steinobst> , letzter Zugriff: 20.10.2020.

Stephenson 2-18:

https://de.wikipedia.org/wiki/Stephenson_2-18, letzter Zugriff: 30.11.2020.

Stern:

<https://de.wikipedia.org/wiki/Stern>, letzter Zugriff: 17.11.2020.

Stern 1. Größe:

https://de.wikipedia.org/wiki/Stern_1._Größe, letzter Zugriff: 18.11.2020.

Steuerung (Systemtheorie):

[https://de.wikipedia.org/wiki/Steuern_\(Systemtheorie\)](https://de.wikipedia.org/wiki/Steuern_(Systemtheorie)), letzter Zugriff: 24.11.2020.

Stimmloser bilabialer Plosiv:

https://de.wikipedia.org/wiki/Stimmloser_bilabialer_Plosiv, letzter Zugriff: 27.11.2020.

Stimulans:

<https://de.wikipedia.org/wiki/Stimulans>, letzter Zugriff: 20.10.2020.

Stoffeigenschaft:

https://de.wikipedia.org/wiki/Stoffeigenschaft#Physikalische_Stoffeigenschaften, letzter Zugriff: 20.10.2020.

Störschall:

<https://de.wikipedia.org/wiki/Störschall>, letzter Zugriff: 22.10.2020.

Strauch:

<https://de.wikipedia.org/wiki/Strauch>, letzter Zugriff: 20.10.2020.

Stromgröße:

<https://de.wikipedia.org/wiki/Stromgröße>, letzter Zugriff: 24.11.2020.

Substanz :

<https://de.wikipedia.org/wiki/Substanz>, letzter Zugriff: 20.10.2020.

Substanzinduzierte Psychose:

https://de.wikipedia.org/wiki/Substanzinduzierte_Psychose, letzter Zugriff: 29.10.2020.

Sussex:

<https://de.wikipedia.org/wiki/Sussex>, letzter Zugriff: 02.12.2020.

Szene (Theater):

[https://de.wikipedia.org/wiki/Szene_\(Theater\)](https://de.wikipedia.org/wiki/Szene_(Theater)), letzter Zugriff: 20.11.2020.

Tanz:

<https://de.wikipedia.org/wiki/Tanz>, letzter Zugriff: 10.11.2020.

Tanztheater:

<https://de.wikipedia.org/wiki/Tanztheater>, letzter Zugriff: 10.11.2020.

Tätigkeit:

<https://de.wikipedia.org/wiki/Tätigkeit>, letzter Zugriff: 09.11.2020.

Tausch (Soziologie):

[https://de.wikipedia.org/wiki/Tausch_\(Soziologie\)](https://de.wikipedia.org/wiki/Tausch_(Soziologie)), letzter Zugriff: 10.11.2018.

Tesla Inc.:

https://de.wikipedia.org/wiki/Tesla,_Inc., letzter Zugriff: 02.12.2020.

Theater:

<https://de.wikipedia.org/wiki/Theater>, letzter Zugriff: 10.11.2020.

Ton (Musik):

[https://de.wikipedia.org/wiki/Ton_\(Musik\)](https://de.wikipedia.org/wiki/Ton_(Musik)), letzter Zugriff: 12.11.2020.

Tondauer:

<https://de.wikipedia.org/wiki/Tondauer>, letzter Zugriff: 12.11.2020.

Tonhöhe:

<https://de.wikipedia.org/wiki/Tonhöhe>, letzter Zugriff: 12.11.2020.

Tontechniker:

<https://de.wikipedia.org/wiki/Tontechniker>, letzter Zugriff: 19.11.2020.

Tragödie:

<https://de.wikipedia.org/wiki/Tragödie>, letzter Zugriff: 20.11.2020.

Transaktion (Wirtschaft):

[https://de.wikipedia.org/wiki/Transaktion_\(Wirtschaft\)](https://de.wikipedia.org/wiki/Transaktion_(Wirtschaft)), letzter Zugriff: 24.11.2020.

Trapez:

[https://de.wikipedia.org/wiki/Trapez_\(Geometrie\)](https://de.wikipedia.org/wiki/Trapez_(Geometrie)), letzter Zugriff: 11.11.2020.

Trigonometrische Funktion :

https://de.wikipedia.org/wiki/Trigonometrische_Funktion, letzter Zugriff: 11.11.2020.

Trittschall:

<https://de.wikipedia.org/wiki/Trittschall>, letzter Zugriff: 12.11.2020.

Tyrannis:

<https://de.wikipedia.org/wiki/Tyrannis>, letzter Zugriff: 30.10.2020.

Überwachung:

<https://de.wikipedia.org/wiki/Überwachung>, letzter Zugriff: 24.11.2020.

Umlaufbahn:

<https://de.wikipedia.org/wiki/Umlaufbahn>, letzter Zugriff: 20.10.2020.

Unternehmen:

<https://de.wikipedia.org/wiki/Unternehmen>, letzter Zugriff: 09.11.2020.

Unternehmensführung:

<https://de.wikipedia.org/wiki/Unternehmensführung>, letzter Zugriff: 24.11.2020.

Unternehmensziel:

<https://de.wikipedia.org/wiki/Unternehmensziel>, letzter Zugriff: 09.11.2020.

Unternehmerrisiko:

<https://de.wikipedia.org/wiki/Unternehmerrisiko>, letzter Zugriff: 09.11.2020.

Urteil (Deutschland):

[https://de.wikipedia.org/wiki/Urteil_\(Deutschland\)](https://de.wikipedia.org/wiki/Urteil_(Deutschland)), letzter Zugriff: 24.11.2020.

Variable (Mathematik) :

[https://de.wikipedia.org/wiki/Variable_\(Mathematik\)](https://de.wikipedia.org/wiki/Variable_(Mathematik)), letzter Zugriff: 20.10.2020.

Vektor:

<https://de.wikipedia.org/wiki/Vektor>, letzter Zugriff: 11.11.2020.

Vektorraum:

<https://de.wikipedia.org/wiki/Vektorraum>, letzter Zugriff: 11.11.2020.

Venus (Planet):

[https://de.wikipedia.org/wiki/Venus_\(Planet\)](https://de.wikipedia.org/wiki/Venus_(Planet)), letzter Zugriff: 30.11.2020.

Vereinigtet Königreich:

https://de.wikipedia.org/wiki/Vereinigtes_Königreich, letzter Zugriff: 02.12.2020.

Verhaltenstherapie:

<https://de.wikipedia.org/wiki/Verhaltenstherapie>, letzter Zugriff: 18.11.2020.

Verknüpfung (Mathematik) :

[https://de.wikipedia.org/wiki/Verknüpfung_\(Mathematik\)](https://de.wikipedia.org/wiki/Verknüpfung_(Mathematik)), letzter Zugriff: 20.10.2020.

Vermutung von Hodge:

https://de.wikipedia.org/wiki/Vermutung_von_Hodge, letzter Zugriff: 26.11.2020.

Verwaltung:

<https://de.wikipedia.org/wiki/Verwaltung>, letzter Zugriff: 09.11.2020.

Verwaltungsakt (Deutschland):

[https://de.wikipedia.org/wiki/Verwaltungsakt_\(Deutschland\)](https://de.wikipedia.org/wiki/Verwaltungsakt_(Deutschland)), letzter Zugriff: 24.11.2020.

Videojournalist:

<https://de.wikipedia.org/wiki/Videojournalist>, letzter Zugriff: 19.11.2020.

Viereck:

<https://de.wikipedia.org/wiki/Viereck>, letzter Zugriff: 11.11.2020.

Vietnam:

<https://de.wikipedia.org/wiki/Vietnam>, letzter Zugriff: 10.11.2020.

Vogelkirsche:

<https://de.wikipedia.org/wiki/Vogel-Kirsche>, letzter Zugriff: 20.10.2020.

Völkerrecht:

<https://de.wikipedia.org/wiki/Völkerrecht>, letzter Zugriff: 24.11.2020.

Vonovia:

<https://de.wikipedia.org/wiki/Vonovia>, letzter Zugriff: 02.12.2020.

Wachstum (Biologie):

[https://de.wikipedia.org/wiki/Wachstum_\(Biologie\)](https://de.wikipedia.org/wiki/Wachstum_(Biologie)), letzter Zugriff: 20.10.2020.

Wahl:

<https://de.wikipedia.org/wiki/Wahl>, letzter Zugriff: 24.11.2020.

Wahrnehmung:

<https://de.wikipedia.org/wiki/Wahrnehmung>, letzter Zugriff: 20.10.2020.

Wahrscheinlichkeit:

<https://de.wikipedia.org/wiki/Wahrscheinlichkeit>, letzter Zugriff: 09.11.2020.

Warten auf Godot:

https://de.wikipedia.org/wiki/Warten_auf_Godot, letzter Zugriff: 02.12.2020.

Wasserpuppentheater:

<https://de.wikipedia.org/wiki/Wasserpuppentheater>, letzter Zugriff: 10.11.2020.

Wasserschall:

<https://de.wikipedia.org/wiki/Wasserschall>, letzter Zugriff: 12.11.2020.

Weckamine:

<https://de.wikipedia.org/wiki/Weckamine>, letzter Zugriff: 20.10.2020.

Weißer Zwerg:

https://de.wikipedia.org/wiki/Weißer_Zwerg, letzter Zugriff: 18.11.2020.

Weisung (Deutschland):

[https://de.wikipedia.org/wiki/Weisung_\(Deutschland\)](https://de.wikipedia.org/wiki/Weisung_(Deutschland)), letzter Zugriff: 24.11.2020.

Westdeutsche Allgemeine Zeitung:

https://de.wikipedia.org/wiki/Westdeutsche_Allgemeine_Zeitung, letzter Zugriff: 02.12.2020.

Wildrosen:

<https://de.wikipedia.org/wiki/Wildrosen>, letzter Zugriff: 20.10.2020.

Willenskraft:

<https://de.wikipedia.org/wiki/Willenskraft>, letzter Zugriff: 18.11.2020.

Winkel:

<https://de.wikipedia.org/wiki/Winkel>, letzter Zugriff: 11.11.2020.

Wirecard:

<https://de.wikipedia.org/wiki/Wirecard>, letzter Zugriff: 02.12.2020.

Wirkstoff:

<https://de.wikipedia.org/wiki/Wirkstoff>, letzter Zugriff: 29.10.2020.

Wissen:

<https://de.wikipedia.org/wiki/Wissen>, letzter Zugriff: 10.11.2020.

Wissenschaft:

<https://de.wikipedia.org/wiki/Wissenschaft>, letzter Zugriff: 28.10.2020.

Yamataka Jindai Sakura:

https://de.wikipedia.org/wiki/Yamataka_Jindai_Sakura, letzter Zugriff: 30.11.2020.

Zahlungsstrom:

<https://de.wikipedia.org/wiki/Zahlungsstrom>, letzter Zugriff: 24.11.2020.

Zentaur (Asteroid):

[https://de.wikipedia.org/wiki/Zentaur_\(Asteroid\)](https://de.wikipedia.org/wiki/Zentaur_(Asteroid)), letzter Zugriff: 30.11.2020.

Ziel:

<https://de.wikipedia.org/wiki/Ziel>, letzter Zugriff: 09.11.2020.

Zwanghafte Persönlichkeitsstörung:

https://de.wikipedia.org/wiki/Zwanghafte_Persönlichkeitsstörung, letzter Zugriff: 18.11.2020.

Zwangsstörung:

<https://de.wikipedia.org/wiki/Zwangsstörung>, letzter Zugriff: 18.11.2020.

Zweistellige Verknüpfung:

https://de.wikipedia.org/wiki/Zweistellige_Verknüpfung, letzter Zugriff: 20.10.2020.

Zwerggalaxie:

<https://de.wikipedia.org/wiki/Zwerggalaxie>, letzter Zugriff: 17.11.2020.

Zwergplanet:

<https://de.wikipedia.org/wiki/Zwergplanet>, letzter Zugriff: 28.10.2020.

Zwetschge:

<https://de.wikipedia.org/wiki/Zwetschge>, letzter Zugriff: 20.10.2020.

Zwischenmenschliche Kommunikation:

https://de.wikipedia.org/wiki/Zwischenmenschliche_Kommunikation, letzter Zugriff: 10.11.2016.

Zyklische Gruppe:

https://de.wikipedia.org/wiki/Zyklische_Gruppe, letzter Zugriff: 20.10.2020.

Forschungsliteratur

- Allemang, Dean / Hendler, Jim. 2011. *Semantic Web for the Working Ontologist*. Waltham, MA: Morgan Kaufman.
- Angele, Jürgen / Kifer, Michael / Lausen, Georg. 2009. Ontologies in F-Logic. In: Staab, Steffen / Studer, Rudi (Hgg.), *Handbook on Ontologies*, 45–70. Heidelberg: Springer.
- Antoniou, Grigoris / Harmelen, Frank van. 2009. Web Ontology Language: OWL. In: Staab, Steffen / Studer, Rudi (Hgg.), *Handbook on Ontologies*, 91–110. Heidelberg: Springer.
- Ayadi, Ali / Sameta, Ahmed / Beuvron, François de Bertrand de / Zanni-Merk, Cecilia. 2019. Ontology population with deep learning-based NLP: a case study on the Biomolecular Network Ontology. In: *Procedia Computer Science* 159. 572–581. <https://doi.org/10.1016/j.procs.2019.09.212> (aufgerufen: 27.07.2023).
- Baader, Franz / Küsters, Ralf / Wolter, Frank. 2007. Extensions to Description Logics. In: Baader, Franz / Calvanese, Diego / Nardi, Deborah McGuinness Daniele / Patel-Schneider, Peter (Hgg.), *The Description Logic Handbook: Theory, Implementation, and Applications*, 226–268. Cambridge: Cambridge University Press.
- Baader, Franz / Horrocks, Ian / Sattler, Ulrike. 2009. Description Logics. In: Staab, Steffen / Studer, Rudi (Hgg.), *Handbook on Ontologies*, 21–43. Heidelberg: Springer.
- Barclay, J. R. / Bransford, John / Franks, Jeffrey / McCarrel, Nancy / Nitsch, Kathy. 1974. Comprehension and Semantic Flexibility. In: *Journal of Verbal Learning and Verbal Behavior* 13. 471–481.
- Barsalou, Lawrence. 1992. Frames, concepts, and conceptual fields. In: Lehrer, Adrienne / Kittay, Eva (Hgg.), *Frames, Fields, and Contrasts. New Essays in Semantic and Lexical Organization*, 21–74. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Barwise, Jon / Cooper, Robin. 1981. Generalized Quantifiers and Natural Language. In: *Linguistics and Philosophy* (4). 159–219.
- Barwise, Jon / Etchemendy, John. 2005. *Sprache, Beweis und Logik. Aussagen- und Prädikatenlogik*. Paderborn: mentis Verlag.
- Beckermann, Ansgar. 2014. *Einführung in die Logik*. Berlin/Boston: De Gruyter.
- Bergenholtz, Henning / Kaufmann, Uwe. 1996. Enzyklopädische Informationen in Wörterbüchern. In: Weber, Nico (Hgg.), *Semantik, Lexikographie und Computeranwendungen*, 167–180. Tübingen: Max Niemeyer Verlag.
- Berlin, Brent / Kay, Paul. 1969. *Basic color terms. Their universality and evolution*. Berkeley/Los Angeles: University of California Press.

- Berners-Lee, Tim / Hendler, James / Lassila, Ora. 2001. The Semantic Web. A new form of Web content that is meaningful to computers will unleash a revolution of new possibilities. In: *Scientific American* 1–4.
- Bicknell, Klinton / Elman, Jeffrey / Hare, Mary / McRae, Ken / Kutas, Marta. 2010. Effects of event knowledge in processing verbal arguments. In: *Journal of Memory and Language* 64 (4). 489–505.
- Bierwisch, Manfred. 1983. Semantische und konzeptuelle Repräsentation lexikalischer Einheiten. In: Ružičzka, Rudolf / Motsch, Wolfgang (Hgg.), *Untersuchungen zur Semantik*, 61–99. Berlin: Akademie Verlag.
- Blake, Frank R. 1930. A Semantic Analysis of Case. In: *Language* 6. 34–49. <https://doi.org/10.2307/521984> (aufgerufen: 27.07.2023).
- Busse, Dietrich. 2009. *Semantik*. Paderborn: W. Fink.
- Busse, Dietrich. 2012. *Frame-Semantik. Ein Kompendium*. Berlin/Boston: De Gruyter.
- Busse, Dietrich. 2014. Begriffsstrukturen und die Beschreibung von Begriffswissen. Analysemodelle und -verfahren einer wissensanalytisch ausgerichteten Semantik (am Beispiel von Begriffen aus der Domäne Recht). In: *Archiv für Begriffsgeschichte* 56. 1–35. <https://www.jstor.org/stable/24361917> (aufgerufen: 20.07.2021).
- Busse, Johannes / Humm, Bernhard / Lübbert, Christoph / Moelter, Frank / Reibold, Anatol / Rewald, Matthias / Schlüter, Veronika / Seiler, Bernhard / Tegtmeier, Erwin / Zeh, Thomas. 2014. Was bedeutet eigentlich Ontologie. In: *Informatik Spektrum* (37). 286–297.
- Calegari, Silvia / Ciucci, Davide. 2007. Fuzzy Ontology, Fuzzy Description Logics and Fuzzy-OWL. In: Masulli, Francesco / Mitra, Sushmita / Pasi, Gabriella (Hgg.), *Applications of Fuzzy Sets Theory. 7th International Workshop on Fuzzy Logic and Applications, WILF 2007 Camogli, Italy, July 7-10, 2007 Proceedings*, 118–126. Berlin/Heidelberg: Springer.
- Calvanese, Diege / Giacomo, Giuseppe de. 2007. Expressive Description Logics. In: Baader, Franz / Calvanese, Diego / Nardi, Deborah McGuinness Daniele / Patel-Schneider, Peter (Hgg.), *The Description Logic Handbook: Theory, Implementation and Applications*, 178–218. Cambridge: Cambridge University Press.
- Carnap, Rudolf. 1934. *Logische Syntax der Sprache*. Schriften zur wissenschaftlichen Weltauffassung. Berlin/Heidelberg: Springer. 10.1007/978-3-662-25375-5 (aufgerufen: 27.07.2023).
- Carstensen, Kai-Uwe / Ebert, Christian / Endriss, Cornelia / Jekat, Susanne / Klabunde, Ralf / Langer, Hagen (Hgg.). 2010. *Computerlinguistik und Sprachtechnologie. Eine Einführung*. 3. Auflage. Heidelberg: Springer.
- Chomsky, Noam. 1993. *Lectures on Government and Binding. The Pisa Lectures*. 7. Auflage. Berlin/New York: De Gruyter.

- Cimiano, Philipp. 2006. *Ontology Learning and Population from Text. Algorithms, Evaluation and Applications*. New York: Springer US.
- Corcoglioniti, Francesco / Rospocher, Marco / Aproso, Alessio Palmero. 2016. Frame-Based Ontology Population with PIKES. In: *IEEE Transactions on Knowledge and Data Engineering* 28 (12). 3261–3275.
- Dahlberg, Ingetraut. 2014. *Wissensorganisation. Entwicklung, Aufgabe, Anwendung, Zukunft*. Würzburg: Ergon Verlag.
- Delgrade, James P. / Schaub, Torsten. 2000. The Role of Default Logic in Knowledge Representation. In: Minker, Jack (Hgg.), *Logic-Based Artificial Intelligence. The Springer International Series in Engineering and Computer Science*, 107–126. Boston: Springer.
- Dowty, David. 1989. On the Semantic Content of the Notion of ‘Thematic Role’. In: Chierchia, Gennaro / Partee, Barbara / Turner, Raymond (Hgg.), *Properties, types and meaning*, 69–129. Dordrecht: Kluwer.
- Dowty, David. 1991. Thematic proto-roles and argument selection. In: *Language* 67. 547–619.
- Engel, Ulrich. 1996. Semantische Relatoren. In: Weber, Nico (Hgg.), *Semantik, Lexikographie und Computeranwendungen*, 223–234. Tübingen: Max Niemeyer Verlag.
- Eskridge, Thomas C. / Hoffman, Robert. 2012. Ontology Creation as a Sense-making Activity. In: *IEEE Intelligent Systems* 27. 58–65.
- Faria, Carla / Serra, Ivo / Girardi, Rosario. 2014. A domain-independent process for automatic ontology population from text. In: *Science of Computer Programming* 95. 26–43.
- Fillmore, Charles. 1968. The Case for Case. In: Bach, Emmon / Harms, Robert (Hgg.), *Universals in Linguistic Theory*, 1–88: Holt, Rinehart and Winston.
- Fillmore, Charles. 1977a. Scene-and-frames Semantics. In: Zampolli, Antonio (Hgg.), *Linguistic Structures Processing*, 55–81. Amsterdam/New York/Oxford: North Holland Publishing Company.
- Fillmore, Charles. 1977b. The Case for Case Reopened. In: Cole, Peter / Sadock, Jerrold (Hgg.), *Syntax and Semantics. Bd. 8. Grammatical Relations*, 59–81. New York/London: Academic Press.
- Fillmore, Charles. 1982. Frame Semantics. In: Cole, Peter / Sadock, Jerrold (Hgg.), *Linguistics in The Morning Calm*, 55–81. Seoul: Hanshin Publishing Corp.
- Fillmore, Charles / Atkins, Beryl. 1992. Towards a Frame-Based Lexicon: The Semantics of RISK and its Neighbors. In: Lehrer, Adrienne / Kittay, Eva (Hgg.), *Frames, Fields and Contrasts. New Essays in Semantic and Lexical Organization*, 75–101. Hillsdale, NJ: Routledge.

- Fraas, Claudia. 2013. Frames - ein qualitativer Zugang zur Analyse von Sinnstrukturen in der Online Kommunikation. In: Frank-Job, Barbara / Mehler, Alexander / Sutter, Tilmann (Hgg.), *Die Dynamik sozialer und sprachlicher Netzwerke. Konzepte, Methoden und empirische Untersuchungen an Beispielen des WWW*, 259–283. Wiesbaden: VS-Verlag.
- Frege, Gottlob. 1879. *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. <https://gallica.bnf.fr/ark:/12148/bpt6k65658c> (aufgerufen: 17.04.2019).
- Frege, Gottlob. 1962. Sinn und Bedeutung. In: Patzig, Günther (Hgg.), *Funktion, Begriff, Bedeutung*, 38–63. Göttingen: Vandenhoeck & Ruprecht.
- Ganino, Giulio / Lembo, Domenico / Mecella, Massimo / Scafoglieri, Federico. 2018. Ontology population for open-source intelligence: A GATE-based solution. In: *Software: Practice and Experience* 48 (12). <https://doi.org/10.1002/spe.2640> (aufgerufen: 27.07.2023).
- Ganter, Bernhard / Wille, Rudolf. 1996. *Formale Begriffsanalyse*. Berlin/Heidelberg: Springer.
- Giraud, Thibaut. 2014. Constructing Formal Semantics from an Ontological Perspective. The Case of Second-Order Logics. In: *Synthese* 191. 2115–2145. <https://link.springer.com/article/10.1007/s11229-013-0387-9> (aufgerufen: 29.04.2019).
- Giraud, Thibaut. 2016. On Modal Meinongianism. In: *Synthese* 193. 1–17. <https://philpapers.org/rec/GIROMM> (aufgerufen: 29.04.2019).
- Grice, Herbert Paul. 1975. Logic and Conversation. In: Cole, Peter / Morgan, Jarry L (Hgg.), *Speech acts*, 41–58. New York, NY: Academic Press.
- Gruber, Jeffrey. *Studies in Lexical Relations*. PhD thesis. Massachusetts Institute of Technology. 1965. <http://www.ai.mit.edu/projects/dm/theses/gruber65.pdf> (aufgerufen: 09.03.2022).
- Gruber, Thomas. 1993. A translation approach to portable ontology specification. In: *Knowledge Acquisition* 5. 199–220.
- Guarino, Nicola. 1997. Understanding, building and using ontologies. In: *International Journal of Human-Computer Studies* 46. 293–310.
- Guarino, Nicola / Oberle, Daniel / Staab, Steffen. 2009. What is an Ontology? In: Staab, Steffen / Studer, Rudi (Hgg.), *Handbook on Ontologies*, 1–17. Heidelberg: Springer.
- Gómez-Pérez, Asunción / Fernández-López, Mariano / Corcho, Oscar. 2004. *Ontological Engineering with examples form the areas of Knowledge Management, e-Commerce and the Semantic Web*. London.
- Hastie, Trevor / Tibshirani, Robert / Friedman, Jerome. 2009. *The Elements of Statistical Learning. Data Mining, Inference, and Prediction. Second Edition*. Springer Series in Statistics. New York: Springer.

- Hauser, Roland. 2000. *Grundlagen der Computerlinguistik*. Berlin/Heidelberg: Springer.
- Hebeler, John / Fischer, Matthew / Blace, Ryan / Perez-Lopez, Andrew. 2009. *Semantic Web Programming*. Indianapolis, IN: Wiley.
- Helbig, Hermann. 2006. *Knowledge Representation and the Semantics of Natural Language*. Berlin/Heidelberg: Springer.
- Herre, Heinrich / Heller, Barbara / Burek, Patryk / Hoehndorf, Robert / Loebe, Frank / Michalek, Hannes. General Formal Ontology (GFO). A Foundational Ontology Integrating Objects and Processes. Technischer Bericht. Universität Leipzig. 2006.
- Heusinger, Klaus von. 1997. *Salienz und Referenz. Der Epsilonoperator in der Semantik der Nominalphrase und anaphorischer Pronomen*. Berlin: Akademie-Verlag.
- Hitzler, Pascal / Krötzsch, Markus / Rudolph, Sebastian / Sure, York. 2008. *Semantic Web. Grundlagen*. Berlin/Heidelberg: Springer.
- Hitzler, Pascal / Krötzsch, Markus / Rudolph, Sebastian. 2009. *Foundations of Semantic Web Technologies*. Boca Raton, FL: Chapman & Hall/CRC.
- Horridge, Matthew. 2005. *A Practical Guide To Building OWL Ontologies using Protégé 4 and CO-ODE Tools*. 1.3. Auflage. Manchester. <http://mowl-power.cs.man.ac.uk/protegeowltutorial/resources> (aufgerufen: 29.07.2021).
- Horrocks, Ian / McGuinness, Deborah L. / Welty, Christopher. 2007. Digital Libraries and Web-Based Information Systems. In: Baader, Franz / Calvanese, Diego / Nardi, Deborah McGuinness Daniele / Patel-Schneider, Peter (Hgg.), *The Description Logic Handbook: Theory, Implementation and Applications*, 436–457. Cambridge: Cambridge University Press.
- Jackendoff, Ray. 1983. *Semantics and Cognition*. Cambridge: MIT Press.
- Jackendoff, Ray. 1987. The Status of Thematic Relations in Linguistic Theory. In: *Linguistic Inquiry* 18. 369–411. <https://www.jstor.org/stable/4178548> (aufgerufen: 08.06.2022).
- Johnson-Laird, P. N. 1988. *The Computer and the mind. An introduction to cognitive science*. London: Fontana Press.
- Joppen, Sandra / Wunderlich, Dieter. 1995. Argument linking in Basque. In: *Lingua* 97 (97). 123–169.
- Jurafsky, Daniel / Martin, James H. 2021. *Speech and Language Processing. Draft of December 29, 2021*. <https://web.stanford.edu/jurafsky/slp3/19.pdf> (aufgerufen: 08.03.2022).
- Kamp, Hans / Reyle, Uwe. 1993. *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht/Boston/London: Springer.

- Katz, Jerrold. 1972. *Semantic Theory*. New York: Harper and Row.
- Katz, Jerrold / Fodor, Jerry. 1963. The Structure of a Semantic Theory. In: *Language* 39. 170–210.
- Kleiber, Georges. 1993. *Prototypensemantik. Eine Einführung*. Tübingen: Gunter Narr Verlag.
- Kutschera, Franz von. 1975. *Sprachphilosophie*. München: Wilhelm Fink Verlag.
- Kutschera, Franz von. 1976. *Einführung in die intensionale Semantik*. Berlin/New York: De Gruyter.
- Lakoff, George. 1973. Hedges: A Study in Meaning Criteria and the Logic of Fuzzy Concepts. In: *Journal of Philosophical Logic* 2. 458–508.
- Lampert, Timm. 2003. *Klassische Logik. Einführung mit interaktiven Übungen*. Frankfurt a. M.: Ontos.
- Lemnitzer, Lothar / Zinsmeister, Heike. 2015. *Korpuslinguistik. Eine Einführung*. 3. Auflage. Narr Studienbücher. Tübingen: Narr.
- Lepore, Ernest. 1983. What model theoretic semantics cannot do? In: *Synthese* 54 (2). 167–187. <https://philpapers.org/rec/LEPWMT> (aufgerufen: 17.04.2019).
- Levin, Beth / Rappaport, Hovav. 2005. *Argument realization*. Cambridge: Cambridge University Press.
- Libet, Benjamin. 2004. Haben wir einen freien Willen? In: Geyer, Christian (Hgg.), *Hirnforschung und Willensfreiheit*, 268–289. Frankfurt a. M.: Suhrkamp.
- Link, Godehard. 1976. *Intensionale Semantik*. München: Fink.
- Löbner, Sebastian. 2015. *Semantik. Eine Einführung*. 2. Auflage. Berlin/Boston: De Gruyter.
- Lohnstein, Horst. 2011. *Formale Semantik und natürliche Sprache*. 2. Auflage. Berlin/New York: De Gruyter.
- Lubani, Mohammed / Noah, Shahrul Azman Mohd / Mahmud, Rohana. 2019. Ontology population: Approaches and design aspects. In: *Journal of Information Science* 45 (4). <https://doi.org/10.1177/0165551518801819> (aufgerufen: 27.07.2023).
- Lutz, Carsten / Wolter, Frank / Zakharyashev, Michael. Temporal Description Logics: A Survey. In: *2008 15th International Symposium on Temporal Representation and Reasoning*. 3–14. 2008.
- Lutzeier, Peter. 1985. *Linguistische Semantik*. Stuttgart: J. B. Metzler.
- Lyons, John. 1977. *Semantics*. Band 1. Cambridge: Cambridge University Press.

- Mangasser-Wahl, Martina. 2000. Roschs Prototypentheorie - Eine Entwicklung in drei Phasen. In: Mangasser-Wahl, Martina (Hgg.), *Prototypentheorie in der Linguistik. Anwendungsbeispiele - Methodenreflexion - Perspektiven*, 15–31. Tübingen: Stauffenburg.
- McKoon, Gail / Ratcliff, Roger. 1988. Contextually relevant aspects of meaning. In: *Journal of Experimental Psychology: Learning, Memory, and Cognition* 14 (2). 331–341.
- McRae, Ken / Cree, George Scott / Seidenberg, Mark / Mcnorgan, Chris. 2005. Semantic feature production norms for a large set of living and non-living things. In: *Behavior Research* 37 (4). 547–559.
- Minsky, Marvin. 1988. *The Society of Mind*. New York: Touchstone.
- Montague, Richard. 1970. Pragmatics and intensional logic. In: *Synthese* 22. 68–94.
- Morsella, Ezequiel. 2008. The mechanisms of human action: Introduction and background. In: Morsella, Ezequiel / Barg, John A. / Gollwitzer, Peter M. (Hgg.), *The Oxford handbook of human action*. 2. Auflage, 1–34. Oxford: Oxford University Press.
- Muscetti, Michela / Rinaldi, Antonio M. / Russo, Cristiano / Tommasino, Cristian. 2022. Multimedia ontology population through semantic analysis and hierarchical deep features extraction techniques. In: *Knowledge and Information Systems* 64. 1283–1303. <https://doi.org/10.1007/s10115-022-01669-6> (aufgerufen: 27.07.2023).
- Neches, Robert / Fikes, Richard / Finin, Tim / Gruber, Thomas / Patil, Ramesh / Senator, Ted / Swartout, William. 1991. Enabling Technology for Knowledge Sharing. In: *AI Magazine* 12 (3). 37–56.
- Newmeyer, Frederick. 2010. On comparative concepts and descriptive categories: A reply to Haspelmath. In: *Language* 86 (3). 688–695.
- Nirenburg, Sergei / Raskin, Victor. 2004. *Ontological Semantics*. Massachusetts.
- Ora, Lassila / McGuinness, Deborah. The Role of Frame-Based Representation on the Semantic Web. Technischer Bericht. Knowledge Systems Laboratory, Stanford University. 2001.
- Palmer, Frank. 1976. *Semantics. A new outline*. Cambridge: Cambridge University Press.
- Pan, Jeff. 2009. Resource Description Framework. In: Staab, Steffen / Studer, Rudi (Hgg.), *Handbook on Ontologies*, 71–90. Heidelberg: Springer.
- Parsons, Terence. 1980. Modifiers and quantifiers in natural language. In: *Canadian Journal of Philosophy* 10. 29–60.
- Partee, Barbara. 1996. The Development of Formal Semantics in Linguistic Theory. In: Lappin, Shalom (Hgg.), *Handbook of Contemporary Semantic Theory*, 1–38.

- Partee, Barbara / Meulen, Alice ter / Wall, Robert. 1990. *Mathematical Methods in Linguistics*. Dordrecht/Boston/London: Kluwer Academic Publishers.
- Peirce, Charles. 1986. Dritte Vorlesung über den Pragmatismus. In: Kloesel, Christian / Pape, Helmut (Hgg.), *Semiotische Schriften*, 431–462. Frankfurt am Main: Suhrkamp Verlag.
- Philipp, Markus / Graf, Tim / Kretschmar, Franziska / Primus, Beatrice. 2017. Beyond Verb Meaning: Experimental Evidence for Incremental Processing of Semantic Roles and Event Structure. In: *Frontiers in Psychology* 1–16. <https://doi.org/10.3389/fpsyg.2017.01806> (aufgerufen: 27.07.2023).
- Priest, Graham. 2005. *Towards Non-Being. The Logic and Metaphysics of Intentionality*. Oxford: OUP Oxford.
- Primus, Beatrice. 2012. *Semantische Rollen*. Kurze Einführungen in die germanistische Linguistik, Band 12. Heidelberg: Universitätsverlag Winter.
- Quillian, Ross. 1967. Word Concepts: A theory and simulation of some basic semantic capabilities. In: *Behavioral Science* 12(5). 410–430.
- Quillian, Ross. 1969. The teachable language comprehender: a simulation program and theory of language. In: *Communications of the ACM* 12 (8). 459–476.
- Rissmann, Lilia / Majid, Asifa. 2019. Thematic roles: Core Knowledge or linguistic construct? In: *Psychonomic Bulletin & Review* 26. 1850–1869.
- Rosch, Eleanor. 1973. Natural Categories. In: *Cognitive Psychology* 4. 328–350.
- Rosch, Eleanor. 1975. Cognitive Reference Points. In: *Cognitive Psychology* 7. 532–547.
- Rosch, Eleanor / Mervis, Carolyn. 1975. Family Resemblances: Studies in the Internal Structure of Categories. In: *Cognitive Psychology* 7. 573–605.
- Rozwadowska, Bożena. 1989. Are thematic relations discrete? In: Corrigan, Roberta / Eckman, Fred / Noonan, Michael (Hgg.), *Linguistic categorizations*, 115–130. Amsterdam: John Benjamins.
- Ruppenhofer, Josef / Ellsworth, Michael / Petruck, Miriam R. L. / Johnson, Christopher R. / Scheffczyk, Jan. 2010. *FrameNet II: Extended Theory and Practice*. <https://framenet2.icsi.berkeley.edu/docs/r1.5/book.pdf> (aufgerufen: 20.05.2022).
- Russell, Bertrand. 1905. On Denoting. In: *Mind* 14 (4). 479–493. <http://www.jstor.org/stable/2248281> (aufgerufen: 05.08.2019).
- Scherer, Carmen. 2014. *Korpuslinguistik*. 2. Auflage. Kurze Einführungen in die germanistische Linguistik, Band 2. Heidelberg: Universitätsverlag Winter.

- Schmid, Hans-Jörg. 2000. Methodik der Prototypentheorie. In: Mangasser-Wahl, Martina (Hgg.), *Prototypentheorie in der Linguistik. Anwendungsbeispiele - Methodenreflexion - Perspektiven*, 33–54. Tübingen: Stauffenburg.
- Schwarz-Friesel, Monika / Chur, Jeanette. 2014. *Semantik. Ein Arbeitsbuch*. 6. Auflage. Tübingen: Narr.
- Segaran, Toby / Evans, Colin / Taylor, Jamie. 2009. *Programming the Semantic Web*. Sebastopol, CA: O'Reilly.
- Smith, Barry. Beyond Concepts: Ontology as Reality Representation. In: *Proceedings of FOIS 2004. International Conference on Formal Ontology and Information Systems*. 1–12. 2004. <https://philpapers.org/archive/SMIBCO.pdf> (aufgerufen: 06.10.2023).
- Sowa, John. 2000. *Knowledge Representation. Logical, Philosophical and Computational Foundations*. Pacific Grove: Course Technology Inc.
- Sowa, John. 2010. The Role of Logic and Ontology in Language and Reasoning. Slightly revisited reprint. In: Poli, Roberto / Seibt, Johanna (Hgg.), *Theory and Applications of Ontology: Philosophical Perspectives*, 1–28. Berlin: Springer. <https://link.springer.com/book/10.1007/978-90-481-8845-1> (aufgerufen: 11.08.2020).
- Spalek, Katharina. 2012. Wortproduktion. In: Höhle, Barbara (Hgg.), *Psycholinguistik 2., unveränderte Auflage*. Akademie Studienbücher - Sprachwissenschaft, 53–66. Berlin: Akademie Verlag.
- Stuckenschmidt, Heiner. 2011. *Ontologien. Konzepte, Technologien und Anwendungen*. 2. Auflage. Berlin/Heidelberg: Springer.
- Szostak, Rich. 2014. The Basic Concepts Classification as a Bottom-Up Strategy for the Semantic Web. In: *International Journal of Knowledge Content Development & Technology* 1. 39–51. <http://dx.doi.org/10.5865/IJKCT.2014.4.1.039> (aufgerufen: 27.07.2023).
- Tarski, Alfred. 1935. Der Wahrheitsbegriff in den formalisierten Sprachen. In: *Studia Philosophica* 1. 261–405.
- Tesnière, Lucien. 2015. *Elements of Structural Syntax. Translated by Timothy Osborne and Sylvain Kahane*. Amsterdam: John Benjamins Publishing Company.
- Teubner, Timm / Flath, Christoph / Winhardt, Christof / Aalst, Wil van der / Hinz, Oliver. 2023. Welcome to the Era of ChatGPT et al. In: *Business and Information Systems Engineering* 65. 95–101.
- Tissaoui, Anis / Sassi, Salma / Chbeir, Richard / Mechergui, Ameni. 2022. A top-down enriching approach for ontology learning from text. In: *Concurrency and Computation. Practice and Experience* 34 (19). 1–19. <https://doi.org/10.1002/cpe.7036> (aufgerufen: 27.07.2023).

- Urchs, Max. 1993. *Klassische Logik. Eine Einführung*. Berlin: Akademie Verlag.
- Uschold, Michael. 2018. *Demystifying OWL for the Enterprise*. San Rafael, CA: Morgan & Claypool Publishers.
- Vigliocco, Gabriella / Vinson, David / Lewis, William / Garrett, Merrill. 2004. Representing the meaning of object and action words: The featural and unitary semantic space (FUSS) hypothesis. In: *Cognitive Psychology* 48. 422–488.
- Vohidova, Nofiza. 2016. *Lexikalisch-semantische Graduonymie. Eine empirisch basierte Arbeit zur lexikalischen Semantik*. (Studien zur deutschen Sprache. Forschungen des Instituts für deutsche Sprache). Tübingen: Narr.
- Weber, Nico. 1996. Formen und Inhalte der Bedeutungsbeschreibung: Definition, Explikation, Repräsentation, Simulation. In: Weber, Nico (Hgg.), *Semantik, Lexikographie und Computeranwendungen*, 1–42. Tübingen: Max Niemeyer Verlag.
- Wedekind, Hartmut / Ortner, Erich / Inhetveen, Rüdiger. 2004. Informatik als Grundbildung. Teil IV: Objektsprache/Metasprache. In: *Informatik-Spektrum* 27. 459–466.
- Weinrich, Uriel. 1966. Explorations in Semantic Theory. In: Sebeok, Thomas (Hgg.), *Current Trends in Linguistics: Vol.3: Theoretical Foundations*, 395–477. Den Haag/Paris: De Gruyter.
- Weller, Karin. 2013. Ontologien. In: Kuhlen, Rainer / Semar, Wolfgang / Strauch, Dietmar (Hgg.), *Grundlagen der praktischen Information und Dokumentation*. 6. Auflage, 207–218. Berlin: De Gruyter.
- Westerståhl, Dag. 2019. Generalized Quantifiers in Natural Language Semantics. In: Lapin, Shalom / Fox, Chris (Hgg.), *The Handbook of Contemporary Semantic Theory, Second Edition*, 9–39. UK: Wiley Blackwell.
- Wittgenstein, Ludwig. 2003. *Tractatus logico-philosophicus. Logisch-philosophische Abhandlung*. Frankfurt a. M.: Suhrkamp.
- Wolski, Werner. 1980. *Schlechtbestimmtheit und Vagheit - Tendenzen und Perspektiven. Methodologische Untersuchungen zur Semantik*. Tübingen: De Gruyter.
- Wood, David / Zaidman, Marsha / Ruth, Luke. 2014. *Linked Data. Structured data on the Web*. Shelter Island, NY: Manning.
- Wunderlich, Dieter / Lakämper, Renate. 2001. On the interaction of structural and semantic case. In: *Lingua* (111). 277–418. [https://doi.org/10.1016/S0024-3841\(00\)00040-1](https://doi.org/10.1016/S0024-3841(00)00040-1) (aufgerufen: 27.07.2023).
- Ziegler, Jayden / Snedeker, Jesse. 2018. How Broad Are Thematic Roles? Evidence From Structural Priming. In: *Cognition* 179. 221–240.

- Ziem, Alexander. 2008. *Frames und sprachliches Wissen. Kognitive Aspekte der semantischen Kompetenz*. Sprache und Wissen, Band 2. Berlin/New York: Walter de Gruyter.
- Ziem, Alexander. 2018. Frames interdisziplinär: zur Einleitung. In: Ziem, Alexander / Inderelst, Lars / Wulf, Detmer (Hgg.), *Frame-Theorien im Vergleich: Modelle Anwendungsfelder, Methoden*, 7–22. Düsseldorf: düsseldorf university press.
- Zoglauer, Thomas. 2016. *Einführung in die formale Logik für Philosophen*. 5. Auflage. Stuttgart: Vandenhoeck & Ruprecht.
- Zwaan, Rolf / Stanfield, Robert / Yaxley, Richard. 2002. Language Comprehenders Mentally Represent the Shape of Objects. In: *Psychological Science* 13. 168–171.