

Wittgensteins Nachlass: „Semantische Annotation von Adjektiven im Big Typescript.“

Angela Krey

Centrum für Informations- und Sprachverarbeitung (CIS)

Ludwig-Maximilians-Universität München

angela.krey@campus.lmu.de

1. Einführung

Im vorliegenden Abstract soll die Posterpräsentation für semantische Annotation von Adjektiven im Big Typescript von Ludwig Wittgenstein vorgestellt werden. Die Annotation wurde im Rahmen des „Wittgenstein in Co-Text“ Projektes am Centrum für Informations- und Sprachverarbeitung (CIS) als Bachelorarbeit realisiert. Die Arbeit beschäftigt sich mit der Erstellung und Einführung semantischer Adjektivklassen in einem für das Big Typescript (Ts-213) erstellte Lexikon. Bezeichnet wird es als wiTTLex und ist im deutschen DELAF-Format gehalten. Es soll alphabetisch sortiert alle Wörter umfassen, die im gesamten Nachlass von Ludwig Wittgenstein vorkommen. Das Format ist speziell für die Arbeit mit lokalen Grammatiken und unserer Suchmaschine „wiTTFind“ (siehe <http://wittfind.cis.uni-muenchen.de>) gewählt. Formen eines Wortparadigmas, grammatische und syntaktische Informationen sind in einer bestimmten Kodierung enthalten. Dieses Lexikon ist an unsere Suchmaschine „wiTTFind“ gebunden und ermöglicht, dass mit der Suchmaschine nicht wie üblich nur nach Wörtern, sondern mithilfe der Klassifikation nach semantischen Wortklassen gesucht werden kann. An allen erfassten Stellen der durchsuchten Werke können

dadurch semantisch annotierte Adjektive gefunden, beziehungsweise entdeckt werden.

2. Klassifikation

Als Hauptkriterium für die Klassifikation dient nicht nur die Häufigkeit der Adjektive und deren Abdeckung, auch die Ansichten Wittgensteins zum Thema Klassifikation und Bedeutung sollen beachtet werden. Hierbei stellt sich schnell heraus, dass ein allgemeines Klassifikationsmodell für das Ts-213 nicht möglich ist und sogar eine autoren-spezifische Klasse, die Stilistika, erstellt werden muss. Außerdem spielen Farbadjektive eine große Rolle. Farbadjektive sind nicht nur die häufigsten Adjektive, Wittgenstein widmet Farben und Farbenmischung im Kapitel Phänomenologie des Ts-213 ein eigenes Unterkapitel. Aus diesem Grund wurde entschlossen, für Farbadjektive noch eine Unterklassifikation durchzuführen. Im Laufe der Klassifikation stellt sich immer wieder heraus, dass manche Wörter nicht eindeutig differenziert werden können und in mehreren Klassen vertreten sind. Zum Beispiel das Wort „lauter“ kann als Numeralia oder als die Steigerungsform von laut, die zur Klasse Komparativa gehört, auftreten. Das Wort „gut“ kann einer Eigenschaft einer Person entsprechen oder einer Evaluation.

Klasse	Kodierung	Kriterien	Beispiele
Farben	COL	Farben nach Farbenlehre Wittgensteins, Unterkategorie: Grundfarbe, Zwischenfarbe, Transparenz, Glanz, Farbigkeit	rot, klar, blau, gelb, schwarz, rein, klar
Numeralia	NUM	quantitativ	zwei, erst, einzig, viel, weniger, zweite
Relation	REL	„auf ein Wort verweisend oder sich darauf beziehend“/ Fachsprache/ „wie“/ „aus“	physikalisch, wesentlich, kausal, grammatisch, sinnlos
Eigennamen	EN	von Eigennamen abgeleitete Adjektive	Fermatschen, Skolemschen, Sheffersche, Eulerscher
Temporalia	TEMP	Zeitangaben	unmittelbar, ehe, lang, andermal, vorig, meist
Evaluation	EVAL	Wertung des Betrachters/ Konnotation/ Subjektivität	falsch, richtig, wahr, einfach, genau
Zustände	ZU	abgeschlossener/ definierbarer Zustand, Lokativa	ganz, wirklich, gegeben, gebraucht, gesagt, allgemein
Komparativa	KOMP	Ähnlichkeit/ Gleichheit/ Steigerung	anders, gleich, verschieden, analog, besser
Stilistika	STIL	autorenspezifische Abstufungen	wohl, wirklich, bestimmt, klar, gewiß
Eigenschaft	EIG	Eigenschaften die Personen zugeschrieben werden	gebildet, ungeschickt, peinlich, selbstständig, gutwillig, naiv
Ereignisse	ER	nicht abgeschlossener Verlauf/ Handlung	unendlich, hinweisend, entsprechend, vorbereitend, folgend

Abbildung 1: Semantische Klassifikation der Adjektive

3. Benutzerinterface

Um in unserer Suchmaschine „wiTTFind“ (siehe <http://wittfind.cis.uni-muenchen.de/semantik.php>) die neuen Suchoptionen benutzerfreundlicher zu gestalten, wurde für diesen Zweck eine „semantische Suche“ mit spezieller Eingabemaske für die neuen semantischen Klassen erstellt. Das Hauptkriterium soll die manuelle Auswahl der semantischen Kategorien sein. Die Eingabezeile der Suchmaschine wurde zu einem Eingabefeld umfunktio- niert. Je nachdem, welche semantische Klasse mit einem Häkchen gekennzeichnet wird, erscheint diese in der Suchzeile. Des Weiteren sollen auch mehrere Kategorien miteinander kombiniert werden können, die Suche sich auf bestimmte Wortar-

ten beschränken, und auch den Kontext miteinbe- ziehen können. Mehrere Häkchen können gesetzt und somit mehrere semantischen Klassen kombi- niert werden. Der Kontext der Suche kann zusätz- lich durch Satzzeichen eingeschränkt werden, hierbei sind vor allem in Anführungszeichen ge- setzte Wörter wie 'rot' oder „rot“ gemeint. Außer- dem müssen die Kategorien durch Beispiele ver- anschaulicht werden. Diese Beispiele werden in Tooltips für jede Klasse gespeichert, die angezeigt werden, sobald per Mouseover der Tooltip akti- viert wird. Für einen noch besseren Überblick können bereits erstellte Frequenzlisten durch Anklicken der Beispiele dargestellt werden.

WiTTFind

Suche einschränken

Semantische Klassen für Adjektive und Nomen

- ☒ **Farben** Beispiele für <ADJ> ← anklicken für Frequenzliste
- ☒ **Numeralia** Beispiele für <ADJ>
- ☐ **Eigennamen** Beispiele für <ADJ>
- ☐ **Temporalia** Beispiele für <ADJ>
- ☐ **Zustände** Beispiele für <ADJ>
- ☐ **Eigenschaft** Beispiele für <ADJ>
- ☐ **Ereignisse** Beispiele für <ADJ>

Semantische Klassen nur für Adjektive

- ☐ **Evaluation** Beispiele für <ADJ>
- ☐ **Relation** Beispiele für <ADJ>
- ☐ **Komparativa** Beispiele für <ADJ>
- ☐ **Stilistika** Beispiele für <ADJ>

Abbildung 2: Beispielanfrage im Benutzerinterface

4. Evaluation

Die statistische Auswertung der semantischen Klassen und ihren Konkordanzen am Ende der Ar-

beit zeigt, dass die Adjektive jeder Klasse mit ähnlichen Nomen kombiniert werden können und somit ein ähnliches Verhalten aufzeigen.