

Farbe und Text – Quantitative Sondierungen am Beispiel von Hans Henny Jahnns*

*Toni Bernhart***

Zusammenfassung

Textstatistische Verfahren untersuchen die Farbetyma in der Prosa Jahnns. Das Ziel ist die transdisziplinäre Nutzung solcher Verfahren für die Literaturwissenschaft. Die Daten werden in Excel verwaltet und bearbeitet. Das Verhältnis zwischen Anzahl der Farbetyma und Anzahl der Textwörter (Farbdichte) steht in Zusammenhang mit der Textsorte und mit der poetologischen Anlage des Textes. Zwischen den Verteilungen einzelner Farben sind funktionale Abhängigkeiten sichtbar. Die Farbverteilung im Korpus korreliert deutlich mit der Farbverteilung in anderen Korpora. Die empirische Farbverteilung lässt sich angenähert durch eine Geradengleichung beschreiben. Diese Beobachtungen zeigen, dass Farbsysteme in hohem Maße geordnet sind und der individuellen Farbsetzung wenig Raum einräumen. Grau jedoch nimmt eine autorspezifische Sonderstellung ein. Es ist überdurchschnittlich häufig, und seine Verteilung zeigt eine relativ geringe funktionale Abhängigkeit von der Verteilung anderer Farben. Ein Ausblick deutet weitere Analysemöglichkeiten und Ergebnisse an.

17.1. Voraussetzungen

Der Beitrag schickt die Benennung der Farben, die Beschreibung des Korpus sowie die Offenlegung der Datenerfassung und der Methode voraus.

17.1.1. Farbe: Etymon und Wort

Die Referenz für das Farblexikon sind die elf *Basic color terms* von Berlin und Kay (1969). Fürs Deutsche adaptiert, ergibt sich die folgende Liste (alphabetisch geordnet): *blau, braun, gelb, grau, grün, lila, orange, purpur, rosa, rot, schwarz, violett* und *weiß*. Weil *purple* mit *purpur, violett* oder *lila* übersetzt wird (Bußmann, 1990, Duden-Oxford, 1991, Stoeva-Holm, 1996), ergeben sich im Deutschen 13 anstatt der elf Begriffe nach Berlin und Kay.

* Erschienen in: *Proceedings der GLDV-Frühjahrstagung 2001*, Henning Lobin (Hrsg.), Universität Gießen, 28.–30. März 2001, Seite 181–190. <http://www.uni-giessen.de/fb09/ascl/gldv2001/>

** Es handelt sich bei diesem Beitrag um Auskoppelungen aus meiner Dissertation, die am Institut für deutsche Literatur der Humboldt-Universität zu Berlin entsteht und den Arbeitstitel „*Adfection derer Körper*“. *Empirische Studie zu den Farben in der Prosa von Hans Henny Jahnns* trägt. Ich danke meinem Doktorvater, Herrn Prof. Dr. Lutz Danneberg, für seine anhaltende Unterstützung, eine in der Literaturwissenschaft nicht ganz tabufreie Methode anzuwenden, und den Herren Prof. Dr. Helmut Richter und Prof. Dr. Ewald Lang für die Beratung in statistischen und linguistischen Fragen.

Ich unterscheide zwischen Farbetymon und Farbwort. Das Farbetymon bezeichnet den chromatischen Wert des Stammes, das Farbwort seine Einbettung in eine lexikalische Kategorie. Farbetya sind in Majuskeln gesetzt (BLAU), Farbwörter kursiv (*blau*).

17.1.2. Korpus

Dem Beitrag liegen die zehn wichtigsten erzählenden Texte von Hans Henny Jahnn (1894–1959) zugrunde (gereiht nach ihrer Entstehungszeit; verwendete Siglen vorangestellt): U *Ugrino und Ingrabanien* (1916–18), P *Perrudja* (1926–29), P2 *Perrudja. Zweites Buch* (1929–33), A *Bornholmer Aufzeichnungen* (1934–35), H *Das Holzschiff* (1935–36), N1, N2 *Die Niederschrift des Gustav Anias Horn I und II* (1936–45), E *Epilog* (1945–48), J *Jeden ereilt es* (1951–52) und B *Die Nacht aus Blei* (1952–54, vgl. Jahnn, 1985–1994). Das Korpus im eigentlichen Sinne sind nicht Jahnn's Texte, sondern Tabellen in Microsoft Excel.

17.1.3. Datenerfassung

Die zehn Texte umfassen 3 341 Seiten und enthalten 3 607 Farbetyma. Diese werden aus ihrem wort- und satzsyntaktischen Kontext gelöst und in eine Tabelle eingegeben. Die Erfassung der Daten erfolgt aus praktischen Gründen „von Hand“.

	BLAU	BRAUN	GELB	GRAU	GRÜN	LILA	ORANGE	PURPUR	ROSA	ROT	SCHWARZ	VIOLETT	WEISS
U 1201				1	1				1				3
U 1202					1	2						1	4
U 1203													0
U 1204													0
U 1205													0

Tabelle 17.1.: Die ersten fünf Zeilen aus der Erfassungstabelle. Die Farbetyma sind durch Textsigle und Seitenzahl lokalisiert

17.1.4. Methode

Für die Darstellung und die Analyse des Gegenstandes verwende ich Methoden der Textstatistik (Sachs, 1992, Holmes, 1994, Altmann, 1995, Lenders und Willée, 1998, Biber et al., 1998), im Besonderen die Korrelation, die Regression und wahrscheinlichkeitsmathematische Verfahren. Die Anfänge der quantitativen Farblinguistik gehen auf Karl Groos¹ zurück.

¹ Leben und Wirken von Karl Theodor Groos sind eng mit Gießen, dem Tagungsort 2001 der GLDV, verbunden. Geboren 1861 in Heidelberg, habilitierte sich Groos 1889 in Gießen und war hier zwischen 1892 und 1898 sowie zwischen 1901 und 1911 als Professor der Philosophie tätig. Neben der Arbeit auf seinem Berufungsgebiet betrieb er vor allem Entwicklungs- und Biopsychologie. Groos starb 1946 in Tübingen.

17.2. Hypothesen, Datenauswertung, Folgerungen

Für eine vollständige Auswertung der Erfassungstabelle fehlt hier der Platz. Ich werde mich daher auf ausgewählte Sondierungen beschränken. Das erste Ergebnis aus der Erfassung zeigt die Verteilung der einzelnen Farbetyma in den zehn Texten.

	BLAU	BRAUN	GELB	GRAU	GRÜN	LILA	ORANGE	PURPUR	ROSA	ROT	SCHWARZ	VIOLETT	WEISS	
U	4	7	1	7	17	4	0	4	1	15	34	0	14	108
P	52	134	148	76	120	4	1	5	23	211	211	19	250	1254
P2	11	10	18	14	14	0	0	2	1	25	20	1	26	142
A	1	5	2	6	6	0	0	0	0	5	8	0	9	42
H	6	21	10	41	12	0	0	4	0	23	25	1	15	158
N1	35	74	49	78	89	1	0	4	7	140	198	7	160	842
N2	23	50	38	76	59	0	0	2	3	89	130	4	131	605
E	12	31	15	36	13	0	0	4	4	49	48	2	45	259
J	0	13	1	7	4	0	0	0	1	29	11	1	16	83
B	2	4	4	16	6	0	0	0	3	13	53	1	12	114
	146	349	286	357	340	9	1	25	43	599	738	36	678	3607

Tabelle 17.2.: Verteilung der Farbetyma in den Texten

17.2.1. Farbdichte

Die Farbdichte bezeichnet das Verhältnis der Anzahl der Farbetyma je 10 000 Textwörter. Die einzelnen Texte haben die folgende Farbdichte: U 30,2; P 55,1; P2 31,8; A 14,0; H 23,5; N1 33,5; N2 26,0; E 20,4; J 14,3; B 49,9. Das arithmetische Mittel dieser Werte beträgt 29,9. Die Hypothese, dass es zwischen der Chronologie der Textentstehung und der Farbdichte oder zwischen dem Textumfang und der Farbdichte einen Zusammenhang gibt, wird durch die Spearman-Rangkorrelation widerlegt. (Auf die Wiedergabe der Daten wird verzichtet.)

Sehr wahrscheinlich aber steht die Farbdichte mit der poetologischen Anlage des Textes und der Textsorte in Zusammenhang. Die geringste Farbdichte haben A (14,0), J (14,3) und E (20,4). Diese drei Texte sind posthum erschienen. Krey (1984) und Schweikert (1987) nehmen an, dass die realistische und offene Darstellung der Homosexualität in J und E mit ein Grund dafür war, dass diese Texte im Klima der fünfziger Jahre nicht erschienen sind. Ferner wollte Jahnn aus J ein Drehbuch entwickeln. A hat die Form eines Tagebuches und ist in hohem Maße dokumentarisch und autobiografisch.

Die zwei Texte mit der höchsten Farbdichte sind P (55,1) und B (49,9), ein frühes und ein spätes Werk. P ist expressionistisch, nicht-linear und sehr symbolisch. B erzählt einen aus J aus-

gelagerten Traum, die homosexuelle Thematik in B ist symbolisch verschleiert und metaphorisch verdunkelt. Nicht zufällig konnte B 1956 noch zu Jahnns Lebzeiten erscheinen.

Aus dieser Gegenüberstellung lässt sich abstrahieren: Geringe Farbdichte steht in Zusammenhang mit realistischer Darstellung und dialogischer Gestaltung. Hohe Farbdichte geht mit uneigentlichem Sprechen einher und lässt auf Metaphorik, Symbolik und Chiffrierung schließen.

A (tagebuchartige Anlage) und J (Grundlage für Film) indizieren die Textsortenabhängigkeit der Farbdichte. Diese wird auch durch andere Untersuchungen bestätigt (Groos, 1909, Lehmann, 1998). Tendenziell nimmt die Farbdichte ab in der Reihenfolge *Lyrik, erzählende Literatur, dialogische Texte, gesprochene Sprache*.

17.2.2. Regression

Der Korrelationskoeffizient ist die Steigung der Regressionsgeraden und daher ein Maß für die funktionale Abhängigkeit einer bivariaten Verteilung. Er vergleicht paarweise die Verteilung einzelner Farben in den zehn Texten.

Wenn die Verteilung der Farben einer mathematisch-funktionalen Abhängigkeit unterliegt, lässt sich daraus folgern, dass das Farbsystem in hohem Maße geordnet ist und wenig Raum für individuelle Gestaltung lässt, dass eine bestimmte Farbe in Abhängigkeit von einer anderen ihre Ausprägung erfährt und ohne diese gar nicht auftreten kann und dass die Häufigkeit einer Farbe prognostizierbar ist, wenn die Häufigkeit einer anderen bekannt ist.

Alle ermittelten Korrelationskoeffizienten (siehe Tab. 17.3) sind mindestens signifikant, mehr als die Hälfte sogar hoch signifikant verschieden von 0. Unterschiedliche funktionale Abhängigkeiten werden in den Größenrelationen sichtbar. Das arithmetische Mittel² der Koeffizienten ist bei BLAU, BRAUN, ROT und WEISS am größten (jeweils 0,95). Im Schnitt ist also die Häufigkeit dieser Farben am deutlichsten von der Häufigkeit anderer Farben abhängig. Der größte Einzelwert (0,99) beschreibt denn auch den Zusammenhang zwischen ebendiesen Farben BLAU und ROT, BLAU und WEISS, BRAUN und ROT, ROT und WEISS, aber auch zwischen GELB und VIOLETT sowie GRÜN und WEISS.

Am kleinsten ist der mittlere Korrelationskoeffizient von GRAU (0,81). Die Häufigkeit dieser Farbe hängt durchschnittlich am wenigsten von anderen Farben ab. Die drei kleinsten Einzelwerte beschreiben den Zusammenhang zwischen GRAU und ROSA (0,63), GRAU und VIOLETT (0,72), GRAU und GELB (0,73).

Die Korrelationsmatrix (Tab. 17.3) zeigt eine insgesamt hohe funktionale Abhängigkeit zwischen den Verteilungen einzelner Farben und damit einen hohen Grad an systemischer Ordnung. GRAU jedoch fügt sich am wenigsten in diese Ordnung. Wofür die Ordnung und die Sonderstellung von GRAU spezifisch sind, wird im Abschnitt 17.2.3 deutlich.

17.2.3. Verteilung der Farben im Kontext

Im evolutionären Modell von Berlin und Kay sind die elf *Basic color categories* hierarchisch geordnet. Zu den einzelnen Entwicklungsstadien einer Sprache gehören bestimmte Farben. Daraus leiten sich die folgenden Ränge ab (manche Farben sind gleichrangig): SCHWARZ (1,5); WEISS (1,5); ROT (3); GELB (4,5); GRÜN (4,5); BLAU (6); BRAUN (7); „purple“ (9,5); GRAU (9,5); ORANGE (9,5); ROSA (9,5) (siehe Berlin und Kay, 1969). Ich fasse VIOLETT, PURPUR und LILA zum berlin/kayschen „purple“ zusammen. Aus den absoluten Werten der Summen in

² Bei der Berechnung des arithmetischen Mittels bleibt die $r = 1$ -Diagonale unberücksichtigt.

	BLAU	BRAUN	GELB	GRAU	GRÜN	ROSA	ROT	SCHWARZ	VIOLETT	WEISS	Mittel
BLAU	—	0,98	0,94	0,87	0,98	0,89	0,99	0,95	0,94	0,99	0,95
BRAUN	0,98	—	0,97	0,84	0,96	0,94	0,99	0,92	0,97	0,98	0,95
GELB	0,94	0,97	—	0,73	0,92	0,97	0,95	0,84	0,99	0,94	0,92
GRAU	0,87	0,84	0,73	—	0,87	0,63	0,86	0,90	0,72	0,88	0,81
GRÜN	0,98	0,96	0,92	0,87	—	0,86	0,98	0,97	0,92	0,99	0,94
ROSA	0,89	0,94	0,97	0,63	0,86	—	0,91	0,81	0,98	0,89	0,88
ROT	0,99	0,99	0,95	0,86	0,98	0,91	—	0,95	0,96	0,99	0,95
SCHWARZ	0,95	0,92	0,84	0,90	0,97	0,81	0,95	—	0,86	0,96	0,91
VIOLETT	0,94	0,97	0,99	0,72	0,92	0,98	0,96	0,86	—	0,94	0,92
WEISS	0,99	0,98	0,94	0,88	0,99	0,89	0,99	0,96	0,94	—	0,95

Tabelle 17.3.: Pearson-Korrelationskoeffizienten der Farben zueinander. Die drei seltenen (siehe Tab. 17.1) LILA, ORANGE, PURPUR sind eliminiert. $n = 10$, $FG = n - 2 = 8$. Signifikanzschwellen für die Nullhypothese $r = 0$: $P = 0,05$ bei $r = 0,63$; $P = 0,01$ bei $r = 0,76$; $P = 0,001$ bei $r = 0,87$

Tab. 17.2 lassen sich die Farbränge bei Jahnn ableiten: SCHWARZ (1), WEISS (2), ROT (3), GRAU (4), BRAUN (5), GRÜN (6), GELB (7), BLAU (8), [VIOLETT + PURPUR + LILA] (9), ROSA (10), ORANGE (11). Die Spearman-Rangkorrelation zwischen den Farbrängen nach Berlin/Kay und den Farbrängen bei Jahnn beträgt 0,76. Tab. 17.4 zeigt, wie sich dieser Wert im Vergleich zu anderen Korpora, die zusammen mehrere Millionen Wörter umfassen, positioniert.

Umfangreiche Korpora (Stichproben) stimmen sehr deutlich mit dem universalsprachlichen Modell von Berlin und Kay überein. Die Farbverteilung erweist sich demnach als autor-, textsorten- und sprachunabhängig. Wenn nun also die Farbverteilung universalsprachlich ist, so ist auch die in Abschnitt 17.2.2 festgestellte funktionale Abhängigkeit zwischen der Verteilung einzelner Farben universalsprachlich und nicht autorspezifisch.

Dennoch steht das Jahnn-Korpus verhältnismäßig am Rande. Wesentlich Ausschlag gebend dafür ist GRAU, das bei Jahnn den vierten, bei Berlin/Kay den 9,5-ten Rang einnimmt. Die Gründe für die große Rangdifferenz sind die häufige leitmotivische Verwendung von *grau* in P und H (die „Grauen Herren“ Grigg und Lauffer) sowie die semantische Akzentuierung dieser „Zwischenfarbe“, die im Weltbild Jahnnns begründet liegt und das Vorhandensein und den Ausgleich von Gegensätzen (Schwarz vs. Weiß) betont. Die Forcierung von *grau* mag auch eine Erklärung dafür sein, warum sich GRAU am wenigsten deutlich in die funktionalen Abhängigkeiten zwischen einzelnen Farbverteilungen einfügt (siehe 17.2.2). Die statistisch messbare Sonderstellung von GRAU ist mithin autorspezifisch.

Artikellänge im Oxford English Dictionary (McManus, 1983)	0,51
Farbnennungen Spanisch (Bolton, 1978)	0,73
<i>Jahnn</i>	0,76
Französisch (Hays et al., 1972)	0,76
Amerikanische Erzählungen (McManus, 1983)	0,78
Spanisch (Hays et al., 1972)	0,81
<i>Summenkorpus Hebbel, Horaz, Hofmannsthal, George, Trakl, Rilke, Lasker-Schüler</i> (Gubelmann, 1922, Steiner, 1986, Thome, 1994)	0,83
Englische Romantik (McManus, 1983)	0,84
Amerikanisches Englisch (McManus, 1983)	0,84
<i>Deutsch</i> (Kaeding, 1898)	0,86
Englisch (Hays et al., 1972)	0,87
Russisch (Hays et al., 1972)	0,87
Deutsch (Hays et al., 1972)	0,89
<i>Zeitungssprache</i> (Rosengren, 1972–1977)	0,89
Chinesisch (McManus, 1983)	0,89
<i>Summenkorpus Bachmann, Becker, Böll, Dürrenmatt, Frisch, Handke, Kaschnitz, Lenz, Loetscher, Wolf</i> (Lehmann, 1998)	0,91
<i>Summenkorpus Goethe, Kleist, Schiller, Shakespeare, Spenser, Wagner, Schumann, Hoffmann, Tieck</i> (Groos, 1909, Groos und Netto, 1910, Groos et al., 1912, Schmidt, 1965)	0,91
<i>Gesprochene Sprache</i> (Ruoff, 1981)	0,93
Rumänisch (Hays et al., 1972)	0,96

Tabelle 17.4.: Steigende Spearman-Rangkorrelation zu Berlin/Kay. Korpora, deren Korrelationen zu Berlin und Kay von anderen Autoren berechnet wurden, in gerader Type, erstmals korrelierte Korpora kursiv. $n = 11$. $FG = n - 2 = 9$. Signifikanzschwellen für die Nullhypothese $r = 0$: $P = 0,05$ bei $r = 0,60$; $P = 0,01$ bei $r = 0,73$; $P = 0,001$ bei $r = 0,84$

17.2.4. Theoretische Farbverteilung

Die Beobachtungen in den Abschnitten 17.2.2 und 17.2.3 legen die Vermutung nahe, dass sich die empirische Farbverteilung f bei Jahn durch eine theoretische Verteilung f' beschreiben lässt. Dazu müssen der Merkmalwert x und die Häufigkeit f der Farben, bezogen auf eine Grundgröße, ermittelt werden. Als Grundgröße wird die bedruckte Seite verwendet. Die Gesamtzahl der Seiten im Korpus beträgt 3 341, die Gesamtanzahl der Wörter 1 099 104, die durchschnittliche Anzahl der Wörter je Seite beträgt 329. Zunächst werden einschlägige theoretische Verteilungen berechnet: Gleich-, Binominal-, Gauß- (Normal-) und Poissonverteilung. Der χ^2 -Test allerdings zeigt, dass die Abweichungen zwischen empirischer und theoretischer Verteilung hoch signifikant sind. Dies falsifiziert die Nullhypothese $f = f'$.

Um dennoch eine theoretische Verteilung zu erhalten, die sich annähernd mit der empirischen deckt, müssen zunächst die Häufigkeiten f der Merkmalwerte x logarithmiert werden. Dabei stellt sich das Problem, dass $f = 0$ nicht logarithmiert werden kann. Solche Werte werden daher manipuliert und durch 1 ersetzt (siehe Tab. 17.5). Die logarithmierten Werte $\ln f$ zeigen bei grafischer Darstellung im Bereich von $x = \{0, 1, 2, 3 \dots 10\}$ sehr deutlich einen linearen Trend. Bei $x > 10$ wird der lineare Trend verlassen. An die logarithmierte Verteilung bei $x < 11$ wird eine Regressionsgerade gelegt. Da jetzt eine Gerade vorliegt, können durch Differenzierung die Steigung a und der Schnittpunkt b für die Geradengleichung $y' = ax + b$ bestimmt werden. Für die Regressionsgerade zu den logarithmierten f -Werten gilt: $a = -0,5983$; $b = 7,2043$. Die Korrelation zwischen den logarithmierten f -Werten und der dazugehörigen Regressionsgeraden ist ausgesprochen hoch: $r = 0,9974$. Für die Farbverteilung in Jahnns Prosa gelten somit die folgenden Funktionen:

$$f' = e^{-0,5983x+7,2043} \quad (17.1)$$

$$\ln f' = -0,5983x + 7,2043 \quad (17.2)$$

Zieht man alle gemessenen Merkmalwerte $x = \{0, 1, 2, 3 \dots 22\}$ in Betracht, so ist die Abweichung zwischen empirischer und theoretischer Verteilung hoch signifikant (siehe χ^2 -Test in Tab. 17.5, S. 188). Wenn man aber die Extremwerte eliminiert, so sinkt das χ^2 , und die Wahrscheinlichkeit ($FG = n - 3$), dass sich die empirische und die theoretische Verteilung decken (Nullhypothese), steigt. Es gilt

$$\text{für } x = \{1, 2, 3, 4, 5, 6, 7, 8, 9\} \chi^2 = 9,60 \Rightarrow 30\% > P\chi^2 > 10\%$$

$$\text{für } x = \{2, 3, 4, 5, 6, 7, 8, 9\} \chi^2 = 6,67 \Rightarrow 30\% > P\chi^2 > 10\%$$

$$\text{für } x = \{3, 4, 5, 6, 7, 8, 9\} \chi^2 = 3,20 \Rightarrow 70\% > P\chi^2 > 50\%$$

$$\text{für } x = \{4, 5, 6, 7, 8, 9\} \chi^2 = 1,50 \Rightarrow 70\% > P\chi^2 > 50\%$$

Die theoretische Verteilung beschreibt ein *natürliches* Verhalten (hier eine *natürlich-sprachliche* Farbverteilung), die zufallsverteilt und mathematisch prognostizierbar ist, als Normverhalten beschrieben und als Ordnung innerhalb eines Systems interpretiert werden kann. Wenn sich empirische und theoretische Verteilung ohne signifikante Abweichung decken, besagt dies, dass das Verhalten des Individuums durch die Norm gesteuert wird und zufällig ist. Der Raum für individuelle, nicht zufallsgesteuerte Farbgestaltung ist mithin – quantitativ betrachtet – äußerst

x	f	$f_{\text{manipuliert}}$	lnf	f'	χ^2
0	1799	1799	7,49	1345,20	153,09
1	693	693	6,54	739,52	2,93
2	369	369	5,91	406,55	3,47
3	204	204	5,32	223,50	1,70
4	114	114	4,74	122,87	0,64
5	69	69	4,23	67,55	0,03
6	38	38	3,64	37,13	0,02
7	17	17	2,83	20,41	0,57
8	10	10	2,30	11,22	0,13
9	7	7	1,95	6,17	0,11
10	4	4	1,39	3,39	
11	6	6	1,79	1,86	
12	3	3	1,10	1,03	
13	3	3	1,10	0,56	
14	1	1	0,00	0,31	
15	0	1	0,00	0,17	
16	1	1	0,00	0,09	
17	0	1	0,00	0,05	
18	0	1	0,00	0,03	
19	1	1	0,00	0,02	
20	0	1	0,00	0,01	
21	1	1	0,00	0,00	
22	1	1	0,00	0,00	
	3341	3345		2987,65	186,79

24,10

Tabelle 17.5.: Logarithmierung der empirischen Verteilung, theoretische Verteilung nach Formel 17.1, χ^2 -Test

klein. Dies impliziert eine Relativierung qualitativer Interpretationen, die an vereinzelten Farbsetzungen festhalten. Hoch signifikante Abweichungen von der wahrscheinlichkeitstheoretischen Norm zeigen sich bei $x = 0$ und $x > 9$. Möglicherweise könnten ansteckende Verteilungen aus der Epidemiologie (Richter und Richter, 1980) auch solche Bereiche als zufallsverteilt erklären.

17.3. Ausblick

In den obigen vier Abschnitten werden ausschließlich ausgezählte lokalisierte Farbetyma verarbeitet. Im Anschluss daran gibt es die Möglichkeit, diese qualitativ zu kontextualisieren, die gewonnenen Informationen wieder zu quantifizieren und weiteren Berechnungen zuzuführen. Erweiterte Beschreibungsmerkmale für Farbwörter sind: Wortart, Einbettung in Satzglieder, semantische Differenzierungen (z. B. durch Komposition, Affigierung und Syntagmenbildung), Zuordnung zu Farbträgern und Klassifizierung der zugeordneten Farbträger.

Daraus lassen sich interessante Hypothesen und Ergebnisse ableiten. Einzelne Farben zeigen beispielsweise unterschiedliche Kombinationspotenziale und Kompatibilitäten, die sowohl in Komposita als auch in Aufzählungen festzustellen sind. GRAU etwa determiniert vorwiegend andere Farben und tritt äußerst selten als Determinatum auf. Die so genannten Gegenfarben BLAU und GELB sind inkompatibel. Dies ist nicht nur anhand von Komposita, sondern auch anhand von Aufzählungen messbar. GRÜN und SCHWARZ etwa tendieren im Korpus sehr deutlich zur Substantivierung, ROT hingegen dynamisiert, indem es überdurchschnittlich häufig verbalisiert ist (*röten*, *erröten*) und zusammen mit der Kopula „werden“ auftritt. Die Kookkurrenzmaxima bei Farbwörtern sind „Auge“, „Haar“ und „Gesicht“. Klassifiziert man die Farbträger, so bildet der menschliche Körper das Feld, das am häufigsten mit Farbe versehen wird. Der Mensch hat damit wesentlichen Anteil an der Konnotation der Farbe, die Farbsemantik ist anthropozentrisch. Das *Braun* des Körpers beispielsweise bedeutet bei Jahnn erotische Ausstrahlung. Das Kookkurrenzmaximum bei SCHWARZ ist nullwertig, d. h. es lässt sich nicht ausmachen, auf welchen Farbträger es sich bezieht. Diese Vagheit ist als Verselbständigung der Farbe deutbar, die bei Jahnn zu einem (positiven) Zustand und einem Raum der Geborgenheit werden kann: Die Musik ist „voll samtenen Schwarzes“ und „das Schwarze der Mutterschoß der Anfänge“ (Jahnn, 1997, 1987).

Literaturverzeichnis

- ALTMANN, G. (1995): *Statistik für Linguisten*. Nummer 55 in *Quantitative Linguistics*. Trier: Wissenschaftlicher Verlag.
- BERLIN, B. UND KAY, P. (1969): *Basic Color Terms. Their Universality and Evolution*. Berkeley, Los Angeles, Oxford: University of California Press.
- BIBER, D.; CONRAD, S. UND REPPEN, R. (1998): *Corpus Linguistics – Investigating Language Structure and Use*. Cambridge: University Press.
- BOLTON, R. (1978): „Black, white and red all over. The riddle of color term salience“. *Ethnology* 17.
- BUSSMANN, H. (1990): *Lexikon der Sprachwissenschaft*. Stuttgart: Kröner.
- DUDEN-OXFORD (1991): *Duden Oxford – Standardwörterbuch Englisch*. Mannheim, Wien, Zürich: Dudenverlag.
- GROOS, K. (1909): „Die optischen Qualitäten in der Lyrik Schillers“. *Zs. f. Ästh. u. allg. Kunstwiss* (4).

- GROOS, K. UND NETTO, I. (1910): "Psychologisch-statistische Untersuchungen über die visuellen Sinesindrücke in Shakespeares lyrischen und epischen Dichtungen". *Englische Studien* 43.
- GROOS, K.; NETTO, I. UND GROOS, M. (1912): "Die Sinnesdaten im „Ring des Nibelungen“. Optisches und akustisches Material". *Archiv f. d. ges. Psych.* (22).
- GUBELMANN, A. (1922): *Studies in the lyric poems of Friedrich Hebbel. The sensuous in Hebbel's lyric poetry.* New Haven: Yale University Press.
- HAYS, D. G.; MARGOLIS, E.; NAROLL, R. UND PEKINS, D. R. (1972): "Color term salience". *American Anthropologist* (74).
- HOLMES, D. I. (1994): "Authorship Attribution". *Computers and the Humanities* 28.
- JAHNN, H. H. (1985–1994): *Werke in Einzelbänden. Hamburger Ausgabe.* Hamburg: Hoffmann und Campe. U. Schweikert and U. Bitz (Herausgeber).
- JAHNN, H. H. (1987): *Späte Prosa.* Hamburg: Hoffmann und Campe.
- JAHNN, H. H. (1997): *Fluß ohne Ufer*, Band 1. Hamburg: Hoffmann und Campe.
- KAEDING, F. W. (Herausgeber) (1898): *Häufigkeitswörterbuch der deutschen Sprache.* Steglitz bei Berlin: Arbeitsausschuß der deutschen Stenographiesysteme.
- KREY, F. (1984): "Doppelleben". In: *Die Suche nach dem rechten Mann. Männerfreundschaft im literarischen Werk Hans Henny Jahnns*, herausgegeben von Popp, W., Berlin: Argument, Nummer 139 in Literatur im historischen Prozeß, Neue Folge.
- LEHMANN, B. (1998): *ROT ist nicht „rot“ ist nicht [rot]. Eine Bilanz und Neuinterpretation der linguistischen Relativitätstheorie.* Nummer 431 in Tübinger Beiträge zur Linguistik. Tübingen: Narr.
- LENDERS, W. UND WILLÉE, G. (1998): *Linguistische Datenverarbeitung – Ein Lehrbuch.* Opladen: Westdeutscher Verlag.
- MCMANUS, I. C. (1983): "Basic Colour Terms in Literature". *Language and Speech* 26.
- RICHTER, B. UND RICHTER, H. (1980): "Zitatausdrücke im Romandialog Stendhals". In: *Literatur und Konversation. Sprachsoziologie und Pragmatik in der Literaturwissenschaft*, herausgegeben von Hess-Lüttich, E. W. B., Wiesbaden: Athenaion.
- ROSENGREN, I. (1972–1977): *Ein Frequenzwörterbuch der deutschen Zeitungssprache. Die Welt. Süddeutsche Zeitung*, Band 2. Lund: CWK Gleerup.
- RUOFF, A. (1981): *Häufigkeitswörterbuch gesprochener Sprache, gesondert nach Wortarten, alphabetisch, rückläufig alphabetisch und nach Häufigkeit geordnet.* Tübingen: Niemeyer.
- SACHS, L. (1992): *Angewandte Statistik. Anwendung statistischer Methoden.* Berlin, Heidelberg, New York: Springer.
- SCHMIDT, P. (1965): *Goethes Farbensymbolik. Untersuchungen zur Verwendung und Bedeutung der Farben in den Dichtungen und Schriften Goethes.* Nummer 26 in Philologische Studien und Quellen. Berlin: Schmidt.
- SCHWEIKERT, U. (1987): "Nachwort". In: *H. H. Jahn: Späte Prosa*, Hamburg: Hoffmann und Campe.
- STEINER, J. (1986): *Rilke. Vorträge und Aufsätze.* Karlsruhe: Loeper.
- STOEVA-HOLM, D. (1996): *Farbbezeichnungen in deutschen Modetexten. Eine morphologisch-semantische Untersuchung.* Dissertation, Universität Uppsala. Acta Universitatis Upsaliensis, Studia Germanistica Upsaliensia, Bd. 34.
- THOME, G. (1994): "Die Funktion der Farben bei Horaz". *Acta Classica* 37.