

Zu einem Versuch empirischer Forschung
in der Literaturwissenschaft.

Kritische Überlegungen zu Werner Bauer, Renate Braunschweig-Ullmann, Helmtrud Brodmann, Monika Bühr, Brigitte Keisers, Wolfram Mauser, "Text und Rezeption".

Walther Kindt, S.J. Schmidt, Jan Wirrer, Reinhard Zobel

1. Kurzer Überblick über "Text und Rezeption"
2. Die rezeptionstheoretischen Voraussetzungen und ihre Probleme
3. Die impliziten ästhetischen und poetologischen Voraussetzungen
4. Textverstehen und Textverständlichkeit
5. Zum Begriff der "Interpretation"
6. Einige grundsätzliche Probleme empirischer Untersuchungen in der Literaturwissenschaft
7. Weiteres zu den Ergebnissen der Untersuchung
8. Abschließende kritische Bemerkungen

1. Kurzer Überblick über "Text und Rezeption"

- 1.1 In "Text und Rezeption" dokumentiert die Freiburger Forschergruppe W. Bauer, R. Braunschweig-Ullmann, H. Brodmann, M. Bühr, B. Keisers und W. Mauser eine in den Jahren 1969/70 durchgeführte Rezeptionsuntersuchung. Bei der gegenwärtigen Forschungslage darf diese Untersuchung wohl zu den wichtigsten Arbeiten gezählt werden, die zur Entwicklung der Literaturwissenschaft als einer empirischen Wissenschaft beitragen. Trotz aller generellen Vorbehalte, die gegen die Untersuchung anzumelden sind, erbringt sie viele interessante Resultate auf dem Gebiet des Rezeptionsverhaltens "literarischer" Texte und läßt erkennen, daß durch ähnliche empirische Untersuchungen

gen das rezeptive Verhalten gegenüber Texten weiter aufgeklärt und einige die Rezeption bestimmende Faktoren isoliert werden können. Demgemäß verstehen sich die folgenden Anmerkungen nicht als prinzipielle Infragestellung der Freiburger Untersuchung, sondern sie möchten auf einige Unzulänglichkeiten der Untersuchung hinweisen, die bei künftigen Forschungen dieser Art vermieden werden müssen. Es können hier allerdings nicht alle problematischen Punkte diskutiert werden, weil der Kritik nicht das gesamte Untersuchungsmaterial, sondern nur der in dem Buch "Text und Rezeption" abgedruckte Teil zugrundeliegt.

1.2 Im folgenden werden zunächst die wichtigsten Punkte aus der Freiburger Untersuchung referiert, Bauer et al. nennen ihre Untersuchung eine "literarische Wirkungsanalyse" und verstehen darunter eine "in Form einer schriftlichen Befragung durchgeführte Analyse der Direktwirkung eines zeitgenössischen literarischen Textes auf kontrastierende Lesergruppen" (vgl. Vorwort S. V); der Begriff "Wirkung" wird später erläutert als "das 'Einwirken' eines Textes auf den jeweiligen Leser im Rahmen des unmittelbaren Kommunikationsprozesses, eine Wirkung, erschließbar nur aus der Spontanantwort des Lesers, in der sich der Wirkungsprozeß manifestiert" (S. 5). Als Primärtext wurde für die Befragung das Gedicht "Fadensonnen" von Paul Celan ausgesucht:

Fadensonnen

Über der grauschwarzen Ödnis.
Ein baum -
hoher Gedanke -
Greift sich den Lichtton: Es sind
noch Lieder zu singen jenseits
der Menschen.¹

¹ Celan, P. (1967), S. 22

Die Auswahl dieses Gedichts wird damit gerechtfertigt, daß es folgenden beiden Kriterien genüge, durch die Bauer et al. den Anwendungsbereich der Wirkungsanalyse auf einen bestimmten Typ literarischer Texte hin eingrenzen:

"Er soll einen so begrenzten Umfang haben, daß es den Lesern möglich ist, den Text in kurzer Zeit und ohne Unterbrechung durchzulesen. (Eine zeitliche Unterbrechung der Lektüre würde den Rezeptionsvorgang insofern beeinflussen, als die spontane Reaktion des Lesers auf den Text als Ganzes dann nicht mehr möglich ist.)" (S. 6)

"Er soll Multivalenz besitzen, d.h. keine eindeutige Information enthalten, sondern eine Vielzahl kontrastierender Bedeutungsstränge, die eine "Aktivierung" des Lesers provozieren; dieser wird angesprochen und gibt eine Antwort (Reaktion) im Sinne einer "Interpretation"." (S. 6)

"Interpretation" wird bei Bauer et al. als "makrosemantische Verarbeitung des Primärtextes durch den Leser" verstanden (s. Fußnote 2 S. 6); auf den Begriff "makrosemantische Verarbeitung" kommen wir später zurück. Um nun die Wirkung des Celan-Gedichtes auf einen größeren Leserkreis zu ermitteln, wurde nach einigen Voruntersuchungen ein Fragebogen entwickelt, der neben 25 Fragen zu persönlichen Daten und Einstellungen weitere 19 Fragenkomplexe enthält, die direkt auf den Primärtext bezogen sind. Aus technischen Gründen konzentrierten sich die Freiburger Forscher auf die Befragung von Schülern und Studenten, eine Repräsentativermittlung konnten sie aus finanziellen Gründen aber auch innerhalb dieser Gruppe nicht durchführen. Insgesamt wurden die Fragebögen von 665 Versuchspersonen (= Vpn) ausgewertet.

1.3 Den theoretischen Hintergrund für die Entwicklung des Fragebogens bildete ein Kommunikationsmodell, das auf der Stufe der Rezeption von folgenden Voraussetzungen ausgeht (vgl. S. 5-13):

- (1) Der Primärtext wirkt als "Stimulus" auf den Leser und ist hierin "durch die Intention des Autors und die von ihm gewählte Art sprachlicher Codierung determiniert". (S. 7)
- (2) "Die vom Leser sprachlich codierte Antwort auf den "Stimulus" (Metatext) wird "manifeste Reaktion" genannt. Sie ist "durch folgende Faktoren determiniert: durch den Stimulus selbst, durch die Art der Decodierung des Textes sowie durch die vom Leser gewählte Art der Incodierung seiner Antwort, wobei der mögliche Einfluß subjektiver Wirkungsfaktoren unterschiedlicher Art zu berücksichtigen ist." (S. 7)
- (3) "Die Vorgänge der Stimulation und der Reaktion sind, als psychologische Faktoren, mit dem Arbeitsinstrument der literarischen Wirkungsanalyse nicht unmittelbar zu erfassen." (S. 7)
- (4) "Die beiden analysierbaren Faktoren des Kommunikationsprozesses im Rahmen der literarischen Wirkungsanalyse sind die beiden Arten von vorliegendem Text: Text₁ (Primärtext) und Text₂ (Metatext)." (S. 8)
- (5) Auf der bedeutungstheoretischen Ebene wird angenommen, daß Texte einen "Bedeutungsrahmen" (Semantischer Rahmen) haben, der ein "mikrosemantisches" und ein "makrosemantisches" Feld umfaßt. Unter ersterem wird verstanden: "Die Feinstruktur des Textes im semantischen Bereich, d.h. die Summe der Bedeutungskomponenten der einzelnen Wörter"; als makrosemantisches Feld wird bezeichnet: "Die vom Text vorgegebene Grobstruktur möglicher Interpretationsstränge, die vom Leser durch selektive Kombination von Bedeutungskomponenten realisiert werden."
(S. 5/6)

- (6) Aus der Perspektive der Wirkungsanalyse wird als spezifisches Kriterium für die Poetizität eines Textes seine Multivalenz angesehen:

"Unter Multivalenz eines Textes wird verstanden: eine Vielzahl variierender, kontrastierender, sich teilweise ausschließender Bedeutungsstränge im Text, die gleichberechtigt nebeneinanderstehen. Der Text ist daher nicht eindeutig zu verstehen."

"Der Multivalenz der semantischen Struktur eines Textes entspricht im Bereich der Rezeption die Vielfalt der durch Text₁ auslösbaren Aktivierungsprozesse bei verschiedenen Lesern. Die in Text₁ enthaltenen realen und potentiellen Bedeutungskomponenten geben den Lesern unterschiedliche, aber nicht unbegrenzte Möglichkeiten der Reaktion und der Aktualisierung des Textes. Sämtliche Vorgänge der Aktivierung des Lesers verlaufen innerhalb des durch die Textstruktur vorgegebenen Bedeutungsrahmens." (S. 12)

- (7) Der genaue Verlauf des Rezeptionsprozesses wird folgendermaßen beschrieben:

"Der Leser reagiert auf Text₁, indem er mittels seiner Sprachkompetenz versucht, die gegebene Information semantisch zu decodieren (mikrosemantische Verarbeitung). Dieser Vorgang wird als Verstehensprozeß (Perzeption) bezeichnet. Der Leser vollzieht in Ergänzung zu der semantischen Decodierung eine individuelle Decodierung von Text₁, in welcher dessen Zuordnung zum Erwartungsprogramm des Lesers erfolgt. Durch die von der Erwartung des Lesers gesteuerte selektive Kombination erkannter Bedeutungskomponenten wird gleichzeitig die Encodierung des Reaktionsprogrammes in Text₂ vorbereitet (makrosemantische Verarbeitung).² Dieser Vorgang wird als Reaktionsprozeß (Interpretation) im engeren Sinne bezeichnet." (S. 8)

- (8) "Art und Verlauf des Rezeptionsprozesses sowie der aus ihm resultierende Text₂ sind bedingt durch das Verhältnis von Text₁ und dem Erwartungskomplex des Lesers (bzw. der Lesergruppe) zum Zeitpunkt der Lektüre." (S. 9) Die "Erwartung" des Lesers setzt sich

nach Bauer et al. aus folgenden Komponenten zusammen:

- (a) Erwartung aus Erfahrungen sprachlicher Art,
- (b) Erwartung aus Erfahrungen im Umgang mit Texten, insbesondere literarischen Texten,
- (c) Erwartung aus individuellen Erfahrungen verschiedenster Art,
- (d) Erwartung, die sich aus der persönlichen Situation des Lesers im Augenblick der Lektüre ergibt. (Vgl. S. 9)

(9) Als Fähigkeiten des Lesers werden für den Rahmen der Wirkungsanalyse vorausgesetzt:

- (a) Eine durchschnittliche Sprachkompetenz für die Sprache, in der der Primärtext abgefaßt ist,
- (b) das Vorhandensein eines Mittelwerts semantischen Bewußtseins,
- (c) die Fähigkeit zur Decodierung von semantisch unbekanntem Verbindungen (z.B. Metaphern und Neologismen) aus ihren semantisch bekannten Bestandteilen. (Vgl. S. 10)

Die in (1) bis (9) aufgeführten Voraussetzungen bestimmen in entscheidender Weise den Aufbau und die Formulierungen des Fragebogens, der im folgenden Abschnitt näher beschrieben werden soll.

1.4 Der Fragebogen gliedert sich in vier Teile. In dem aus zehn Fragen bestehenden ersten Teil (Fragen 1 - 10) soll die Einstellung der Versuchspersonen zum Lesen und insbesondere zur "zeitgenössischen Lyrik" ermittelt werden. Anschließend werden die Versuchspersonen aufgefordert, das Gedicht von Celan zu lesen; der Name des Autors wird aber noch nicht ge-

nannt, "um eine vorzeitige Determinierung der Leserreaktion durch diese zusätzliche Information zu vermeiden." (S. 57) Die darauf folgenden 20 Fragen des Hauptteils (Fragen 12 - 31) gelten direkt dem Gedicht. Erst im dritten Teil (Fragen 32 - 40) wird der Name "Celan" angegeben und einige Fragen zur Person des Autors gestellt, aus deren Antworten der mögliche Einfluß der Kenntnis des Autors abgeschätzt werden soll. Der vierte und letzte Teil des Fragebogens (Fragen 41 - 46) enthält noch einige Fragen nach den persönlichen Daten der Vpn.

Für die folgende Diskussion ist insbesondere der Aufbau des Hauptteils wichtig, der daher genauer besprochen werden soll. Im Hauptteil sind drei Frageblöcke zu unterscheiden, die mit den drei von Bauer et al. theoretisch postulierten Rezeptionsstufen korrespondieren. Bei diesen drei Stufen handelt es sich um eine "Spontanstufe", eine "analytische Stufe (Mikrosemantik)" und eine "synthetische Stufe (Makrosemantik)"; es wird angenommen, daß die Vpn die Stufen "in variierender Reihenfolge und beliebiger Vollständigkeit realisieren können" (vgl. S. 63/64). Die der Spontanstufe zugeordneten Fragen fordern die Vpn zunächst zu einer Wertung des Gedichts und zu einer Wiedergabe des ersten Eindrucks auf (Frage 12 und 13). Danach werden Fragen zur Verständlichkeit bzw. zur Wichtigkeit der einzelnen Wörter gestellt (Frage 14-17). Durch diese Fragen soll den Versuchspersonen eine Hilfestellung in Richtung auf das Textverständnis bzw. die Möglichkeit zur persönlichen Akzentsetzung gegeben und zugleich soll ihre Aufmerksamkeit auf die "Metaphern" des Gedichts gelenkt werden. Der zweite Frageblock (Frage 18 - 25) korrespondiert mit der analytischen Rezeptionsstufe und in ihm werden die Versuchspersonen aufgefordert, freie Assoziationen zu den einzelnen Wörtern des Gedichts niederzuschreiben, und zwar zunächst

zu den von Bauer et al. als Metaphern eingestuften Wörtern/ Wortkomplexen "Fadensonnen", "baumhoher Gedanke" und "Lichtton" und anschließend zu den Wörtern "Ödnis", "Lieder" und "jenseits". "In diesen Assoziationen, die den Kern des Untersuchungsmaterials der Wirkungsanalyse bilden, soll der Leser - mehr oder weniger systematisch - den Prozeß der bewußten semantischen Decodierung von Text₁ [...] vollziehen." (Vgl. S. 59) Letzteres bedeutet nach Meinung der Autoren: Der Leser "erfaßt das semantische Gesamtpotential des jeweiligen Wortes, indem er im Verlauf des Assoziationsvorgangs dessen Einzelkomponenten [...] semantisch aktiviert und benennt." (S. 59) Die Autoren führen weiter aus, daß mit dem Assoziationsvorgang das mikrosemantische Feld des Primärtextes realisiert werde und der Leser über die semantische Komponentenanalyse der Einzelwörter die nötigen Voraussetzungen für ein umfassenderes Textverständnis (makrosemantische Realisierung) erwerbe. Die Vpn werden im zweiten Frageblock zunächst jeweils nach Spontanassoziationen befragt und anschließend nach Assoziationen aus dem Bereich von Kunstwerken (Dichtung, Bildende Kunst, Musik) sowie dem Farbbereich. Zur Ergänzung dieser Assoziationsfragen werden bestimmte Bedeutungskomponenten der drei Metaphern nach der Methode des semantischen Differentials ermittelt.¹

Der dritte Fragenkomplex (Frage 26 - 31) wird der synthetischen Rezeptionsstufe zugeordnet und soll die Versuchspersonen von der Beschäftigung mit einzelnen Wörtern wieder zum Gesamttext zurückführen. Zunächst werden die Vpn aufgefordert, das Gedicht noch einmal zu lesen und anschließend in wenigen Worten den Gesamteindruck wiederzugeben, den das Ge-

¹ Zu dieser Methode s. Osgood, Ch.E.; Soci, G.J.; Tannenbaum, Ph. (1957)

dicht auf sie gemacht hat (Frage 26). Bauer et al. sehen die Beantwortung dieser Frage als ersten Schritt der makrosemantischen Verarbeitung des Primärtextes an (S. 62). Mit der darauffolgenden Frage 27 wird den Vpn eine Entscheidung darüber abverlangt, welcher Satz des Gedichts am wichtigsten sei; Bauer et al. meinen, daß damit auch die notwendige Entscheidung für einen von drei im Gedicht potentiell enthaltenen makrosemantischen Rezeptionsstränge falle, die durch die Stichworte "Bildvorstellung", "dynamisches Vermittlungsprinzip" bzw. "anthropologisch/metaphysisch" gekennzeichnet werden (S. 62). In Frage 28/29 werden den Vpn 13 verschiedene, aus dem Material der Voruntersuchungen stammende Interpretationsansätze, die sich im Rahmen der drei genannten Grundtypen bewegen, zur Beurteilung vorgelegt: Die Vpn müssen angeben, welche der Ansätze ihnen als zutreffend bzw. unzutreffend erscheinen und welchem Ansatz sie sich ggf. anschließen würden. Auf diese Weise sollen die Vpn provoziert und zu einem eigenen Interpretationsversuch angeregt werden (S. 62). Frage 30 schließlich fordert dazu auf, den ausgewählten Interpretationsansatz mit eigenen Worten weiter auszuführen bzw. den eigenen Ansatz zu formulieren, sofern er von allen genannten Möglichkeiten abweicht. Damit soll also die Abbildung der makrosemantischen Strukturierung des Primärtextes in einen Metatext geleistet werden. Bauer et al. gehen davon aus, daß diese Formulierungsaufgabe auch für diejenigen Vpn, die Verständnisbarrieren bzw. unterschiedlich motivierte Rezeptionsschwierigkeiten zu überwinden haben, aufgrund der Beantwortung der vorangehenden Fragen keine Schwierigkeiten mehr bereitet (S. 63). Zum Abschluß des dritten Frageblocks wird noch eine Kontrollfrage zur Unverständlichkeit des Gedichts gestellt.

1.5 Im folgenden sollen in einem ersten Überblick einige der Ergebnisse aufgezählt werden, die nach Bauer et al. aus dem Untersuchungsmaterial resultieren.¹ Solche Ergebnisse sind z.B.:

- Bei der freien Assoziation seien überindividuelle Gesetzmässigkeiten festzustellen, die sich vor allem im Bereich der sog. Primärassoziation zeigen. Erst der hier anschliessende "potentielle mikrosemantische Bereich" wird vom subjektiven Erfahrungskontext beeinflusst (vgl. S. 218-219);
- die Multivalenz des jeweiligen Wortes bzw. Textes werde erkennbar "mit dem Einsetzen der affektiv-subjektiv gelenkten Wortstimuli im Bereich der Konnotationen" (vgl. S.219);
- bei der Wirkungsanalyse zeitgenössischer Texte sei nicht primär von der "literarischen Erfahrung", sondern vielmehr von den spezifischen Momenten eines Textes auszugehen (vgl. S. 220);
- die "literarische Erfahrung" habe bei den Untersuchungen nur insofern eine Rolle gespielt, als sie den Vpn das Verständnis des Textes erleichtere (vgl. S. 220-221);
- der traditionelle Versuch, durch Ermittlung der Autorintention eine eindeutige Interpretation zu finden, verstelle den Blick für die Multivalenz literarischer Texte (vgl. S. 222);
- bei Oberstufenschülern und Philologiestudenten aus den ersten Semestern sei es zu einer auffallend starken Häufung "klischeehafter Begriffe" gekommen (vgl. S. 222);
- durch das Fehlen signifikanter Unterschiede in der Rezeption von Lesern "mit ausgeprägtem literarischen Erfahrungshorizont" und Lesern mit geringer literarischer Erfahrung

¹ Eine genauere Diskussion dieser Punkte erfolgt in 8.

werde die Hypothese gestützt, derzufolge eine durchschnittliche Sprachkompetenz für das Verständnis eines multivalenten zeitgenössischen Textes ausreiche (vgl.S. 223).

2. Die rezeptionstheoretischen Voraussetzungen und ihre Probleme

2.1 Ein genereller Mangel der Untersuchung von Bauer et al., der allerdings in dem derzeit noch niedrigen Forschungsstand der Kommunikationswissenschaften begründet ist, hängt mit dem von den Autoren vorausgesetzten Kommunikations-/Rezeptionsmodell zusammen. Dieses Modell stellt mit seinen verschiedenen Rezeptionsstufen ein sehr komplexes System dar, über dessen empirischen Wert jedoch keinerlei Informationen vorliegen und das daher nicht ohne weiteres in eine empirische Untersuchung zugrundegelegt werden dürfte. Oder anders ausgedrückt: Der Wert einer empirischen Untersuchung, die sich maßgeblich auf ein empirisch nicht überprüfbares Rezeptionsmodell stützt, bleibt vorerst zweifelhaft. Bei der Untersuchung von Bauer et al. kommt erschwerend hinzu, daß sie in folgenden vier Punkten Schwächen aufweist:

- (1) Das zugrundegelegte Rezeptionsmodell arbeitet mit Begriffen, die nicht ausreichend geklärt sind (u.a. betrifft dies die Begriffe "Bedeutung", "mikrosemantisches Feld", "makrosemantisches Feld", "Erwartung"; (vgl. S. 5-13).
- (2) Bauer et al. erheben nicht nur global für das Rezeptionsmodell, sondern auch für die einzelnen Rezeptionsstufen den Anspruch der empirischen Realität, ohne diesen Anspruch näher zu begründen (vgl. S. 8/S. 57-64).
- (3) In der Untersuchung wird, ohne dies zu problematisieren, angenommen, daß bestimmte Typen von Fragen im Fragebogen eindeutig bestimmten Rezeptionsstufen zugeordnet werden

können (vgl. S. 57-64).

- (4) Bauer et al. sichern ihre Untersuchung nicht gegen die mögliche Gefahr ab, daß die Bestätigung ihrer Hypothesen ausschließlich oder teilweise auf der durch die Beantwortung des Fragebogens hervorgerufenen Aktivierung der Vpn beruht. (Diese Gefahr hätte durch Kontrollbefragungen ausgeschaltet werden müssen).

Um alle diese Schwächen vermeiden zu können, hätte es zusätzlicher und sehr viel weiter ausholenderer Untersuchungen bedurft: u.a. hätte das kommunikations- und bedeutungstheoretische Konzept näher ausgearbeitet und präzisiert werden müssen, es wäre notwendig gewesen, die einzelnen in diesem Konzept vorgesehenen Rezeptionsstufen unabhängig voneinander im Hinblick auf ihre Existenz und Beschaffenheit zu testen, und es hätte untersucht werden müssen, ob und auf welche Weise Einfluß auf das Rezeptionsverhalten von Vpn ausgeübt wird, wenn ihnen bestimmte Typen von Fragen zur Beantwortung vorgelegt werden. Es ist uns klar, daß Bauer et al. ein solches umfangreiches Untersuchungsprogramm nicht durchführen konnten, und es soll hier auch keinesfalls die Auffassung vertreten werden, daß in empirischen Untersuchungen immer schon alle Voraussetzungen selbst vollständig geklärt sein müssen. Diese Forderung wäre insbesondere in der Anfangsphase einer neuen Forschungsrichtung unangemessen, weil sie die Forschungsmotivation verringern und die Forschungsmöglichkeiten zu stark einschränken würde. Wir sind aber der Meinung, daß in einer Untersuchung, bei der bestimmte Punkte noch ungeklärt sind, auch entsprechende Vorsicht im Hinblick auf generalisierende Aussagen notwendig ist. Eine solche Vorsicht lassen die Ausführungen von Bauer et al. nicht erkennen. So behaupten die Autoren beispielsweise, daß die Vpn in den freien Assoziationen "den Prozeß der bewußten semantischen Decodierung" des Gedichttextes vollzögen, bzw. daß damit die

"Realisierung des mikrosemantischen Feldes" anhand der Assoziationstexte erfolge (vgl. S. 59). Erstens wäre es wieder notwendig, die hier verwendeten Begriffe zu klären, und zweitens gibt es keinerlei empirische Belege für die Richtigkeit dieser oder ähnlicher Behauptungen. Im Gegenteil, mit H. Clark (1970) kann man davon ausgehen, daß die freie Wortassoziation ein künstliches, im Laufe des Sozialisationsprozesses erst erlerntes Spiel ist, aus dem man allenfalls Rückschlüsse auf die semantische Decodierung der Vpn ziehen kann, das ansonsten aber durch eigenständige Operationen definiert ist. Auch die folgende These von Bauer et al. hätte ohne nähere Prüfung nicht der Entwicklung und Auswertung des Fragebogens zugrundegelegt werden dürfen:

"Mit der Beantwortung der Adjektivskalen des semantischen Differential leistet der Leser, insofern dies noch nicht in den freien Assoziationen geschehen ist, die bewußte linguistische Mikroanalyse der Einzelkomponenten der gegebenen Worte." (S. 61)

Diese Hypothese ist auch insofern problematisch, als mit dem semantischen Differential nur die emotiven Bedeutungskomponenten gemessen werden können (was Bauer et al. selbst ausdrücklich vermerken) und mit der Beantwortung der Adjektivskalen folglich eine Aktivierung denotativer Bedeutungskomponenten nicht notwendig einhergehen muß.

Wir wollen an dieser Stelle die generelle Diskussion über die rezeptionstheoretischen Voraussetzungen der Freiburger Untersuchung abbrechen, weil die mit ihnen verbundene Problematik hinreichend deutlich geworden sein dürfte. Im nächsten Abschnitt soll aber noch ein spezielles Problem angesprochen werden, das mit dem Aufbau und der Formulierung des Fragebogens zusammenhängt.

2.2 Die im Fragebogen unter 12-17 aufgeführten Fragen werden der hypothetisch vorausgesetzten ersten Rezeptionsstufe, der Spontanstufe, zugeordnet (vgl. S. 63). Diese Zuordnung scheint insbesondere für so detaillierte Fragen wie 16 und 17 (Einteilung der Wörter im Gedicht nach Wichtigkeit und Unverständlichkeit; vgl. S. 47) problematisch zu sein. Denn trotz der Aufforderung an die Vpn zu Beginn des Fragebogens, über einzelne Fragen nicht lange nachzudenken, sondern unmittelbar zu antworten, könnte in stärkerem Maße, als Bauer et al. dies annehmen, die Reflexion der Vpn einsetzen und die Spontanstufe überschritten werden. Um eine explizite Entscheidung über die Wichtigkeit und Unverständlichkeit eines Wortes treffen zu können, mußten die Vpn doch vermutlich erstens eine intensivere mikrosemantische Analyse des Wortes vornehmen und zweitens auch den Kontext, in dem das Wort steht, berücksichtigen, d.h. daß sie bereits in erheblichem Maße makrosemantische Beziehungen konstruierten. Nun brauchte die mögliche inadäquate Zuordnung der Fragen 16 und 17 zur Spontanstufe ja nicht unbedingt negative Auswirkungen auf das Untersuchungsergebnis gehabt zu haben. Tatsächlich sind solche Auswirkungen aber nicht auszuschließen und zwar deshalb, weil die in Frage 16 und 17 abverlangte Entscheidung möglicherweise die von den folgenden Fragen stimulierten Verarbeitungsprozesse und die Beantwortung der Fragen selbst vordeterminiert. Das hiermit angesprochene Problem der möglichen Vordeterminierung des Rezeptionsprozesses durch den Fragebogen selbst gewinnt noch durch den folgenden Umstand an Schärfe. Analysiert man die Formulierung der Frage im Hinblick auf eventuelle Vorweg-Beeinflussung des Rezeptionsverlaufs, dann sind insbesondere die Fragen 15, 16 und 17 problematisch:

- "Welches Wort fällt Ihnen am meisten auf?"

- "Bitte bewerten Sie die einzelnen Worte des Gedichts, indem Sie sie ihrer Wichtigkeit nach durchnummerieren. Das Wort, das Ihnen am wichtigsten erscheint, erhält also die 1, das nächstwichtige die 2, usw. [...]."
- "Welcher Satz des Gedichts erscheint Ihnen am wichtigsten?" (S. 47/52)

Diese Fragen haben Präsuppositionen, die nicht notwendigerweise wahr sein müssen¹; d.h. in diesen Fragen werden implizit Annahmen vorausgesetzt, die nach Einschätzung der Vpn nicht unbedingt zutreffen müssen. Hierbei handelt es sich um:

- Die Annahme, daß für die jeweilige Vpn ein Wort am auffälligsten ist,
- die Annahme, daß es für die Vpn möglich sein muß, die Wörter des Gedichts ihrer Wichtigkeit nach zu staffeln,
- die Annahme, daß für die Vpn ein Satz als am wichtigsten erscheint.

Die implizite Voraussetzung dieser Annahmen hätte sich ohne weiteres vermeiden lassen, nämlich dann, wenn man für die Fragen 15, 16 und 27 neutralere Formulierungen gewählt hätte, z.B. für die Frage 15 die Formulierung: "Sofern Ihnen ein Wort am meisten auffällt, welches Wort ist das?". Welchen Grund hat es also, daß Bauer et al. ihre und nicht eine neutralere Art der Formulierung wählten. Vielleicht waren sich die Autoren des Präsuppositionsproblems nicht be-

¹ Zu dem Problem von Fragepräsuppositionen vgl. auch Jan Wirrer (1976a)

wußt, vielleicht enthielten aber auch die aus den Voruntersuchungen stammenden Interpretationstexte jeweils bestimmte Schwerpunktsetzungen und diese Eigenschaft wurde von den Autoren generalisiert. Wie dem auch sei, die Art der Formulierung der Fragen 15, 16 und 27 kann den Effekt gehabt haben, daß sich die Vpn gezwungen sahen, die zugehörigen Fragepräsuppositionen zu akzeptieren und entsprechende Rangfolgen aufzustellen. In diesem Fall war der spätere Informationsspielraum aber von vornherein eingeengt und es ist dann auch nicht verwunderlich, daß die Interpretationsansätze der Vpn vielfach ebenfalls auf entsprechende Schwerpunktsetzungen zurückgehen (vgl. S. 205 ff. und S. 221).

3. Die impliziten ästhetischen und poetologischen Voraussetzungen

3.1 Im folgenden sollen einige Anmerkungen zu den impliziten ästhetischen Voraussetzungen der hier durchgeführten Wirkungsanalyse gemacht werden. Gerade im Interesse einer Weiterführung des insgesamt begrüßenswerten Ansatzes hin zu einer empirischen Literaturwissenschaft ist es erforderlich, den Versuch zu machen, diese Voraussetzungen zu explizieren, um so zu einer angemessenen Beurteilung von "Text und Rezeption" zu gelangen.

3.2 Der folgende Abschnitt beschäftigt sich mit den von den Autoren angegebenen Definienda der "Poetizität"¹. Nach Ansicht

1

Auf Seite 11 heißt es - leider ziemlich pauschal -: "Poetizität wird der Kunstcharakter eines literarischen Textes verstanden."

der Autoren soll ein literarischer Text "Multivalenz besitzen, d.h. keine eindeutige Information enthalten, sondern eine Vielzahl kontrastierender Bedeutungsstränge, die eine 'Aktivierung' des Lesers provozieren; dieser wird angesprochen und gibt eine Antwort (Reaktion) im Sinne einer 'Interpretation'." (S. 6)

Entscheidendes Indiz für die Poetizität eines Textes ist nach Ansicht der Autoren die "Wirkung im Sinne von 'Einwirken', als Auslösefaktor des Rezeptionsprozesses" (S. 11). Der Leser wird aktiviert sowohl durch Übereinstimmung als auch durch Divergenz seiner Erwartungen an und mit dem Text.

Diese Aktivierung kann nur in zweierlei Hinsicht erfolgen: entweder die Leserreaktion ist durch die Textstruktur in einer bestimmten Richtung determiniert, oder die Leserreaktion ist nicht determiniert. Texte, die ausschließlich (oder zumindest nach Erhalt von Zusatzinformationen) eine eindeutige Decodierung anbieten, sind im Sinne der Autoren nicht-poetisch. Poetisch sind allein multivalente Texte: "Das spezifische Kriterium für die Poetizität ist, aus der Perspektive der Wirkungsanalyse, die Multivalenz eines Textes." (S. 12)

Multivalente Texte lösen vielfältige Aktivierungsprozesse bei Lesern aus, die zwar unterschiedlich aber nicht unbegrenzt sind. Vielmehr verlaufen sämtliche Vorgänge der Aktivierung des Lesers "innerhalb des durch die Textstruktur vorgegebenen Bedeutungsrahmens". (S. 12) Poetisch ist also nur ein Text, der eine "Aktivierung des Lesers innerhalb verschiedener möglicher Bereiche" bewirkt. Als Kriterium für die Poetizität wird eindeutig die Multivalenz eines Textes angesetzt (vgl. S. 20).

Hierzu sind die folgenden kritischen Anmerkungen zu machen:

- (1) Die Autoren gehen davon aus, daß der Leser eines literarischen Textes über "durchschnittliche Sprachkompetenz" verfügt bzw. einen "sprachlichen Code" besitzt (S. 10). Neben diesem sprachlichen Code setzen die Autoren einen "ästhetischen Code" an, der nicht zum Besitz des Lesers gehört, sondern "der sich erst im Vorgang, des Lesens realisiert. Der ästhetische Code baut zwar auf der Kenntnis des sprachlichen Codes auf, konstituiert sich aber erst durch den fortschreitenden Prozeß der Rezeption. Der ästhetische Code ist entscheidend für die Individualität und Poetizität von Text₁." (S. 11)

Hier werden u.E. zwei Gesichtspunkte miteinander vermischt. Aus der Bestimmung des ästhetischen Codes als eines "potentiellen Faktors" (S. 11), der sich im Lesevorgang aufbaut, geht hervor, daß es sich ganz offensichtlich um einen Prozeß im Leser handelt. Wenn dann aber im nächsten Satz behauptet wird, der ästhetische Code sei "entscheidend für die Individualität und Poetizität von Text₁", so wird hier offenbar von einer Texteigenschaft gesprochen. Eine Vereinbarkeit beider Behauptungen ergäbe sich vielleicht, wenn der Leser komplizierte Interpretationshypothesen aufbauen würde, etwa derart, daß der Leser beim Rezipieren von T₁ über dem sprachlichen Code (in dem T₁ abgefaßt ist) einen ästhetischen Code aufbaut, um T₁ zu "verstehen", und dann diesen ästhetischen Code als Texteigenschaft interpretiert. Aber selbst bei solchen - von den Autoren nicht gelieferten - Deutungen bleibt der Begriff "ästhetischer Code" vage. Zumindest ist zu bemängeln, daß es sich beim ästhetischen Code sicher nicht um einen nach Termen aufzählbaren Code handelt, der mit dem sprachlichen Code vergleichbar wäre.

- (2) Aus der referierten Definition von Poetizität geht nicht eindeutig hervor, ob es sich bei "Poetizität" um einen heuristisch eingeführten Begriff oder um eine generelle ästhetische Hypothese mit allgemeinem Geltungsanspruch handeln soll. Als ein "entscheidendes Indiz für die Poetizität eines Textes" sehen die Autoren "Wirkung - im Sinne von "Einwirken"" an (S. 11).

Diese Bestimmung ist problematisch, da Wirkung wohl bei der Lektüre jedes sprachlichen Textes angesetzt werden muß, soll der Rezeptionsprozeß erfolgreich sein. Wirkung kann also nicht spezifisch sein für Poetizität. Das berücksichtigen - unausgesprochen - wohl auch die Autoren, wenn sie bei den Wirkungen (bzw. "Aktivierungen") ausschließlich vordeterminierte und offene Wirkung unterscheiden, und Texte, die "ausschließlich eine vordeterminierte Aktivierung bewirken" (S. 12) aus der Klasse der "poetischen Texte im eigentlichen Sinne" ausschließen¹. Aber auch diese Festlegung ist noch problematisch; denn die Behauptung der Autoren, "Alltagssprache, Werbetexte, Propagandatexte oder Tendenzdichtung" hätten gemeinsam, "daß die in ihnen enthaltene Information eine bestimmte und nur diese Reaktion im Leser auslösen soll" (S. 12) läßt sich empirisch wohl kaum bestätigen. Zumindest scheint mir aus persönlichen Erfahrungen ziemlich klar zu sein, daß Werbe- und Propagandatexte sowie Tendenzdichtungen auf unterschiedliche Rezipienten ganz unterschiedliche Wirkungen haben bzw. Akti-

¹ Was mit dem Ausdruck "poetische Texte im eigentlichen Sinn" gemeint ist, wird leider nirgendwo erläutert.

vierungen auslösen, und zwar nicht nur solche, die vom Autor intendiert waren.

- (3) Die Behauptung der Autoren, poetische Texte "im eigentlichen Sinne" aktivierten den Leser ohne Vordetermination, basiert auf der generell unterstellten Hypothese, Multivalenz sei das spezifische Kriterium für Poetizität (S. 12).

Diese Hypothese ist Grundpostulat bestimmter poetologischer (oder allgemeiner ästhetischer) Ansätze, die zu Beginn der 70er Jahre diskutiert wurden¹.

Aber die Multivalenzästhetik ist eine ästhetische Theorie neben anderen, deren Berechtigung erst noch in einem wissenschaftstheoretisch strengeren Sinne begründet werden muß. Bis dahin tut man gut daran, solche ästhetischen Voraussetzungen als heuristische Thesen explizit zu kennzeichnen, damit der Leser sie als explizit unterstellte Prämissen in Argumenten realisieren und auf ihre Relevanz hin abschätzen kann. Sicher wäre diese kritische Anmerkung zur impliziten Ästhetik müßig, handelte es sich nur um Prolegomena in einer Einleitung; u.E. spielen aber die vorausgesetzten ästhetischen Prämissen eine weitreichende Rolle bei der Anlage und Auswertung der Fragebögen, d.h. sie sind konstitutive Bestandteile der empirischen Gesamtuntersuchung.

¹cf. Schmidt (1974), Schmidt (1971)

3.3 Der folgende Abschnitt bringt einige Anmerkungen zu den ästhetischen Implikationen der Fragebogenformulierungen. Wie die Autoren berichten, wurden zunächst in einem Vorversuch semantische Strukturen bestimmter Assoziationsfelder (vor allem zu den Metaphern des Gedichts) ermittelt sowie allgemeine ästhetische Einstellungen zu sog. moderner Lyrik erhoben¹. Die im Vorversuch gewonnenen Ergebnisse haben dann die Formulierung des Fragebogens bestimmt: "Die Antworten der Germanistikstudenten auf die offenen Fragen, die speziell die Einstellung zur zeitgenössischen Lyrik betrafen, sollten im Hinblick auf weitere Versuche mit Nicht-Philologen die Vorlage für spätere Auswahlfragen bilden." (S. 36)

Diese Voruntersuchungen sollten der Ermittlung des "Erwartungshorizontes" dienen, in dem Leser mit einem Text wie P. Celans "Fadensonnen" konfrontiert werden sollten. Als primäre Charakteristika dieses Erwartungshorizontes für sog. moderne Lyrik ergaben sich die bekannten Vorurteile: Moderne Lyrik sei "ungewöhnlich" (S. 28), "exzeptionell" (S. 29), sie schockiere den Leser und sei unverständlich². Diese Vorbehalte werden nun keineswegs thematisiert (etwa in den Kontext

¹Die Definition moderne Lyrik = Lyrik nach 1945 muß wohl als eine Verlegenheitslösung angesehen werden. Wenn man sich ansieht, welche total divergierenden lyrischen Texte nach 1945 erschienen sind und wie weit sie oft hinter das Innovationsniveau der Literatur vor 1945 zurückfallen, dann wird selbst der heuristische Gebrauch einer solchen Definition mehr als fragwürdig.

²Man fragt sich, welche Lerngeschichte die befragten Germanistikstudenten durchlaufen haben mögen, wenn bei einer Befragung solche Erwartungshorizonte genannt werden.

einer Sozialisationsgeschichte gestellt bzw. in Abhängigkeit von zur Schulzeit gängigen Ästhetiktheorien gesetzt), sondern als unbefragtes Faktum in die Formulierung des Fragebogens übernommen.

Zu der Voruntersuchung, die leider im Buch nicht ausführlich vorgestellt wird, sind einige kritische Anmerkungen anzubringen:

- (a) Da die Voruntersuchung die außerordentlich wichtige Aufgabe erfüllen sollte, die Merkmale für die Fragebögen der Hauptuntersuchung zu liefern, ist es unverständlich, warum nur Germanistikstudenten nach ihrer Einstellung zur zeitgenössischen Lyrik befragt wurden. Wenn man - sei es aus finanziellen, zeitlichen oder organisatorischen Gründen - nur begrenzte Voruntersuchungen durchführen konnte, wäre zu erwarten gewesen, daß neben die signifikant negativen Merkmale, die aus der Befragung der Germanistikstudenten resultierten, von den Untersuchungsleitern neutrale Merkmale gesetzt würden, um eine Steuerung der Hauptuntersuchung durch eine nicht-repräsentative Voruntersuchung zu verhindern.
- (b) Daß eine möglicherweise unzulässige Steuerung der Hauptuntersuchung durch die Voruntersuchung vorliegt, läßt sich u.E. an den Fragebögen der Hauptuntersuchung zeigen.

Frage 5¹ beginnt mit den Merkmalen "unverständlich" und "befremdend" und führt unter den 13 wählbaren Merkmalen nicht weniger als 6 implizit oder explizit negativ wertende Merkmale auf (Merkmale 1, 2, 5, 8, 11, 13). Die als positiv zu betrachtenden Merkmale (3, 6, 9, 10, 12) sind außerordentlich vage ("faszinierend", "trägt zur Daseinsbewältigung bei" etc.). Unterrepräsentiert sind solche Merkmale, die man als (relativ neutrale) Beschreibungsmerkmale bezeichnen könnte; nur 4 "benutzt Alltagssprache" könnte in diese Rubrik eingeordnet werden.

Eine mögliche implizite Beeinflussung der Versuchspersonen in Richtung auf Übernahme des Präjudiz' "Moderne Lyrik ist unverständlich" wird potentiell verstärkt durch die Fragen 17 und 31, die wieder nach der Schwerverständlichkeit fragen².

¹ Die Frage lautet:
"Was sind, Ihrer Meinung nach, die wichtigsten Merkmale zeitgenössischer Lyrik?
(Mehrere Antworten sind möglich)

1... unverständlich, 2... befremdend, 3... bringt neue Wortbildungen, 4... benutzt Alltagssprache, 5... Sprachspielerei, 6... faszinierend, 7... Sprachmagie, 8... leeres Geschwätz, 9... gesellschaftskritisch, 10... politisch engagiert, 11... pessimistisch, 12... trägt zur Daseinsbewältigung bei, 13... nur Spezialisten zugänglich."
(S. 45)

² Wir werden darauf in 4.1 noch einmal zu sprechen kommen.

Auch bei Frage 6¹ liegt u.E. die Gefahr einer Beeinflussung vor; die in der Voruntersuchung erhobene negative Einschätzung zeitgenössischer Lyrik in der Erwartung zeitgenössischer Germanistikstudenten wird auf zweierlei Weise herausgestellt:

- (a) dadurch, daß genau diese Frage relativ früh in der Fragenliste und in unmittelbarer Nachbarschaft zu Frage 5 auftaucht und ein negatives Vorurteil in Form einer umfangreichen Fragenliste zur Debatte stellt;
- (b) durch die unklare Formulierung der Frage und der Antwortmöglichkeiten (vor allem Antwortmöglichkeit 4 und 5). Es fragt sich, warum hier keine positiven Kontrollfragen eingebaut worden sind, um Verzerrungen zu vermeiden.²

¹ Die Frage lautet:

"Wie stehen Sie zu dieser Aussage?

"Auch für den aufgeschlossenen Leser ist es schwierig, manchmal unmöglich, moderne Gedichte zu verstehen."

1... ich stimme völlig zu, 2... ich stimme im allgemeinen zu, 3... ich bin unentschieden, 4... ich möchte eher ablehnen, 5... ich lehne völlig ab." (S. 45)

²

R. Zobel hat in seinem Buch "Der Dramentext - ein kommunikatives Handlungsspiel." Göppingen 1975, 94, die Frage nach der Unverständlichkeit moderner Lyrik bei den Freiburger Autoren auch dahingehend kritisiert, daß es notwendig gewesen wäre, Fragebatterien aufzubauen, um unbewusste Hemmungen der Versuchspersonen bei der Beantwortung abzubauen.

Fragen 8, 9, 10 verstärken weiter die Gefahr, daß eine Klischeebildung auf die Vpn übertragen wird, da diese Fragen besonders triviale und in der Diskussion seit langem problematische Kategorien dadurch wieder aufwerten, daß sie explizit danach fragen: Kategorien wie "Schicksal", "Erlebnis", "Stimmung" und "Ausdruckswille" des Autors.

Nun mag man einwenden, daß solche Kategorien beim sog. Durchschnittsleser eben eine wichtige Rolle spielen und dabei auf die Voruntersuchungen verweisen, in denen sie angeblich gewonnen worden sind.

Dagegen wäre zu bemerken, daß die Voruntersuchung sicher nicht als repräsentativ angesehen werden kann (jedenfalls soweit dies aus der Arbeit zu entnehmen ist) und außerdem nirgendwo erwähnt wird, ob dies die einzigen Kategorien waren, die in der Voruntersuchung aufgetaucht sind. Offenbar ist hier die Auswahl der in den Fragen thematisierten Kategorien auch bei den Autoren gesteuert von einer bestimmten Poetik und allgemeinen Ästhetik, die aber leider nicht expliziert werden.

Das drückt sich u.E. auch darin aus, daß die als Kriterium der Poetizität eingeführte Multivalenz durchaus zwiespältig behandelt wird:

- (a) einmal wird sie positiv bewertet als Eröffnung einer Vielzahl möglicher Bedeutungen;
- (b) zum anderen aber wird die Wirkung der Multivalenz auf den Leser als "Unverständlichkeit" bezeichnet (cf. etwa S. 58), statt neutral von Mehr- bzw. Vielfachverständlichkeit zu sprechen.

Auch hier scheinen mir implizite Poetik und ihr Einfluß auf die praktische empirische Arbeit nicht genügend überdacht und expliziert worden zu sein, wodurch zumindest die Gefahr einer Verzerrung durch gerichtete Fragen gegeben ist. Hier hätten die angewandten Methoden viel genauer reflektiert und behutsamer eingesetzt werden müssen, um ein unbeeinflusstes Bild zu erlangen.

3.4 Die von den Autoren implizit vertretene Poetik geht an zentralen Punkten auch in die Auswertung ein. So wird im Kapitel III, 3 eine Auswertung nach folgenden Gruppen vorgenommen: Gruppe L_1 bilden die Vpn, die angeben, häufig zeitgenössische Lyrik zu lesen, Gruppe L_2 bilden die Vpn, die ihren Angaben gemäß gelegentlich, Gruppe L_3 diejenigen, die nie zeitgenössische Gedichte lesen. In Frage 12 werden die Vpn zur Bewertung des Celan-Gedichtes aufgefordert. Für die Antworten der Gruppe L_3 geben die Autoren folgende Interpretation:

"In L_3 , wo dieser Erwartungshorizont [der durch die Kenntnis zeitgenössischer Lyrik aufgebaute] fehlt, fällt die Beurteilung viel schlechter aus. Aufgrund des fehlenden Beurteilungsmaßstabes innerhalb dieser Gruppe weigern sich 39.1% der Vpn, ein Urteil abzugeben, während weitere 28.3%, deren Erwartung gegenüber dem Gedicht sich offenbar nicht an zeitgenössischen Texten ausgebildet hat, sondern am traditionellen, inhaltlich redundanten Gedicht, den vorgelegten Text als schlecht bezeichnen." (. 101)

Daß das sog. traditionelle Gedicht inhaltlich redundanter sei, ist wiederum ein gängiges Urteil, das legitimiert werden müßte. Darüberhinaus müßte der Geltungsbereich dieser Hypothese präzisiert werden, denn es kann kaum unterstellt werden, allé nichtzeitgenössischen Gedichte seien inhaltlich redundant.

Eine weitere Stelle, an der die in der Voruntersuchung ermittelten klischeehaften Vorstellungen ohne jede Stellungnahme zum Stellenwert dieser Kategorien in die Auswertung eingehen, ist die Auswertung der Antworten auf die Frage 5. Die Befragung ergibt hier folgende Ergebnisse:

"gesellschaftskritisch	274 Vpn	(57.8%)
bringt Neologismen	240 Vpn	(50.6%)
politisch engagiert	175 Vpn	(36.9%)
Sprachspielerei	156 Vpn	(32.9%)
befremdend	132 Vpn	(27.8%)
benutzt Alltagssprache	119 Vpn	(25.1%)
Sprachmagie	100 Vpn	(21.1%)

Etwas schwächer beantwortet wurden folgende Variablen:

pessimistisch	89 Vpn	(18.8%)
faszinierend	88 Vpn	(18.6%)
trägt zur Daseinsbewältigung bei	84 Vpn	(17.7%)
unverständlich	48 Vpn	(10.1%)
nur Spezialisten zugänglich	45 Vpn	(9.5%)
leeres Geschwätz	16 Vpn	(3.4%)

(S. 75)

Die Zahlen werden wie folgt interpretiert:

"Die Schwerpunktsetzung innerhalb der Gesamtbeantwortung dieser Frage, die teils formale, teils inhaltliche Aspekte hervorhebt, macht deutlich, daß offensichtlich von seiten der Leser zwei unterschiedliche Richtungen zeitgenössischer Lyrik gemeint sind: einerseits die sog. engagierte Lyrik (etwa im Sinne Enzensbergers), andererseits die "dunkle Lyrik" in der Nachfolge Mallarmés (etwa im Sinne von P. Celan).

Obwohl nun die für die letztere charakteristischen sprachlichen Merkmale in den Leserantworten deutlich hervorgehoben werden, scheut sich doch die Mehrzahl der Vpn - wie aus der schwachen Beantwortung bestimmter Variablen hervorgeht - die 'Unverständlichkeit' des hermetischen Typs zeitgenössischer Lyrik zuzugeben." (S. 75)

Offenbar werden hier der sog. engagierten Lyrik die Eigenschaften "gesellschaftskritisch", "politisch engagiert", "benutzt Alltagssprache" zugeordnet, wohingegen die übrigen Eigenschaften wohl eher als solche der sog. dunklen Lyrik angesehen werden. Außer "benutzt Alltagssprache" und "bringt Neologismen", die noch am ehesten als Beschreibungsmerkmale fungieren können, sind alle anderen Terme sehr vage, so daß eine eindeutige Zuordnung zu dem einen oder anderen Typ zeitgenössischer Lyrik schwieriger sein dürfte, als die Autoren anzunehmen scheinen. So sind z.B. Interpretationsmuster denkbar, unter denen die Lyrik Celans durchaus als "gesellschaftskritisch", Enzensbergers Lyrik hingegen als "befremdend" gelten könnte. Die Autoren jedoch entscheiden sich für die gängigen Interpretationsmuster, ohne eine solche Entscheidung zumindest zu problematisieren, ja, sie sind sich ihrer Sache derart sicher, daß sie es dem Leser überlassen, die ihrer Meinung nach gegebenen Zuordnungen zu treffen. Ebenso problematisch ist es, auf der Basis dieser gängigen Vorstellungen probabilistische Aussagen darüber zu machen, was die Vpn gemeint haben könnten.

Ästhetische Klischeevorstellungen werden auch bei der Interpretation der Antwortquoten auf die Frage 7 deutlich. Gefragt nach den Erwartungen an ein zeitgenössisches Gedicht antworteten:

"Anregung zum Nachdenken	385 Vpn	(80.2%)
Neue Sicht d. Wirklichkeit	231 Vpn	(48.1%)
Erkenntnisprozeß	157 Vpn	(32.7%)

Wesentlich schwächer (...) werden im Vergleich hierzu folgende Erwartungskomponenten angesprochen:

Belebung des vom Verstand Überdeckten	104 Vpn	(21.7%)
ästhetisches Vergnügen	99 Vpn	(20.6%)

Nur die wenigsten Vpn erwarten vom zeitgenössischen Gedicht:

innere Bereicherung	64 Vpn	(13.3%)
Bestätigung der eigenen Erfahrung	49 Vpn	(10.2%)
ein Gegengewicht zum grauen Alltag	30 Vpn	(6.4%)" (S.76)

Diese Ergebnisse interpretieren die Autoren folgendermaßen:

"Aus dieser Gesamtverteilung der Antworten läßt sich insgesamt das Zurücktreten einer intuitiv gefärbten, wesentlich subjektiv ausgerichteten Einstellung - wie sie vielleicht von dem an traditioneller Lyrik geschulten Leser auf die zeitgenössische Lyrik übertragen wird - zugunsten einer stärker rational ausgerichteten Leserperspektive beobachten, wie sie der Rezeption zeitgenössischer Texte angemessener zu sein scheint." (S. 76)

Hier müßte legitimiert werden, warum eine rational ausgerichtete Leserperspektive zeitgenössischen Texten "angemessener" ist.

Überdies bleibt vielfach unklar, welche der aufgezählten Erwartungen als rational, welche als weniger rational oder als irrational gelten sollen.

4. Textverstehen und Textverständlichkeit

4.1 Von besonderer Wichtigkeit für die hier besprochene Untersuchung ist die Frage nach der Verständlichkeit und dem Verstehen von Texten und Textbestandteilen. Nach der Verständlichkeit von "Fadensonnen" wird in den Fragen 14 und 31 gefragt:

"14) Finden Sie das Gedicht nach der ersten Lektüre unverständlich?
Ja mit Einschränkung Nein" (S. 47)

"31) Halten Sie das Gedicht, nachdem Sie längere Zeit darüber nachgedacht haben, für verständlich?
Ja Nein" (S. 53)

Auffällig ist, daß in Frage 14 genaugenommen nach dem Unverständnis, in Frage 31 dagegen nach dem Verständnis des Textes gefragt wird. Ganz abgesehen davon, daß sich die Ergebnisse auf die beiden Fragen bei paralleler Formulierung besser hätten miteinander vergleichen lassen, muß gefragt werden, ob sich die Autoren bei der Formulierung auch dieser beiden Fragen von ihrer implizit vorausgesetzten Poetik haben leiten lassen, dergemäß ein sog. hermetisches Gedicht (zunächst) unverständlich zu sein habe. Dies könnte auf die Ergebnisse insoweit einen Einfluß gehabt haben, als mit der Frage 14 möglicherweise ein Unverständnis gegenüber dem Celan-Text suggeriert wird. Sieht man die Ergebnisse im Lichte dieser möglicherweise suggestiven Fragestellung, so muß die Differenz zwischen den Antwortquoten auf die beiden Fragen entsprechend

vorsichtig interpretiert werden (bei Frage 14 finden nur 21% der Vpn den Text verständlich, bei Frage 31 hingegen 78.3%).

4.2 Frage 14 ist für die Ergebnisse der Untersuchung auch insofern wichtig, als die Vpn entsprechend der möglichen Antwort ("ja", "mit Einschränkung", "nein") auf die Frage 14 ("Finden Sie das Gedicht nach der ersten Lektüre unverständlich?") in drei Gruppen eingeteilt und diese Gruppen näher untersucht werden (die Gruppen werden im folgenden der Reihe nach mit G1, G2, G3 bezeichnet).

Auf S. 85 wird darauf hingewiesen, daß in der "Ja"-Gruppe G1 die Vpn im Alter von 10 - 17 Jahren überwiegen (54%). Die Autoren schließen daraus:

"Dies erlaubt einerseits die Annahme, jüngere Vpn hätten gegenüber Texten zeitgenössischer Lyrik besonders starke Verständnisschwierigkeiten im Vergleich zu älteren Vpn, andererseits ist es verständlich, daß die in dieser Altersgruppe erfaßten Gewerbeschüler und Gymnasiasten unterer Klassenstufen weniger Bedenken zeigen, ihr spontanes Unverständnis zum Ausdruck zu bringen als z.B. ein Student von ca. 25 Jahren."

Die erste dieser beiden Annahmen ist ohne zusätzliche Daten, aus denen genauere Rückschlüsse auf das tatsächliche Textverständnis der Vpn gezogen werden können, nicht zu rechtfertigen, zumal die zweite Annahme an dem Aussagewert der ermittelten Prozentverteilung für G1 zweifeln läßt. Sofern man aus dieser Verteilung überhaupt auf stärkere Verständnisschwierigkeiten bei der Gruppe der 10 - 17jährigen schließen kann, müßte man diesen Schluß durch eine sehr viel genauere und explizitere Argumentation begründen. Erstens wäre abzuschätzen, welche Korrektur sich für die Prozentverteilung ergibt, wenn man berücksichtigt, daß jüngere Vpn eher bereit sind, ein Unverständnis zuzugeben. Zweitens müßte die Hypothese gestützt werden, daß man von der Bejahung der Frage 14 für alle

Vpn in gleichem Maße auf Verständnisschwierigkeiten schließen kann (zunächst ist ja nicht auszuschließen, daß verschiedene Vpn mit objektiv gleicher Verständnisleistung ihre Leistung unterschiedlich bewerten); nur bei Gültigkeit dieser Hypothese kann die im ersten Schritt errechnete korrigierte Prozentverteilung auf das Vorliegen von "besonders starken Verständnisschwierigkeiten" übertragen werden. Drittens ist zu begründen, warum die aus den zwei Altersgruppen gebildeten Klassen - wenn sie schon nicht repräsentativ sind - überhaupt sinnvoll miteinander verglichen werden können (an einer solchen Vergleichbarkeit muß gezweifelt werden, weil in der Gruppe der 10 bis 17Jährigen mit den Gewerbeschülern vermutlich eine stärkere Inhomogenität besteht). Viertens schließlich müßte zunächst der korrigierte Prozentsatz der 10 bis 17Jährigen in G1 zu dem Gesamtanteil dieser Altersgruppe (39.2%; vgl. S. 71) in Beziehung gesetzt werden und erst dann kann errechnet werden, ob die altersmäßige Verteilung auf die Ja-Antworten bei Frage 14 als zufällig zu bewerten ist oder nicht. Schon an diesem einen Beispiel wird deutlich, daß Bauer et al. mit ihrem Datenmaterial ziemlich willkürlich umgehen - was vermutlich auf mangelnde Erfahrung der Autoren mit empirischen Untersuchungen zurückgeführt werden kann. Wir halten die Diskussion dieser Mängel deshalb für interessant, weil den Literaturwissenschaftlern bisher jegliche Erfahrung mit statistischen Methoden fehlt, sie diese Erfahrung aber vermutlich am leichtesten anhand der Erörterung von solchen Fragestellungen sammeln können, die aus dem Gegenstandsbereich der Literaturwissenschaft stammen.

Die Zusammensetzung der "Nein"-Gruppe G3 wird von Bauer et al. folgendermaßen beschrieben:

"G3 dagegen, in der die Vpn im Alter von 25 - 27 Jahren dominieren, besteht zum größten Teil aus Studenten. Unter diesen treten zahlenmäßig die Philologen und Theologen in den Vordergrund, die beiden Fachgruppen also, die aufgrund ihrer spezifischen Ausbildung Übung in der Interpretation von Texten besitzen und daher für das Verständnis eines schwierigen Textes besser vorbereitet sind als die übrigen Vpn." (s. 86)

In dieser Passage soll vermutlich die Aussage über die spezifische Ausbildung der Philologie- und Theologiestudenten als Erklärung dafür dienen, daß die Studenten dieser beiden Fachrichtungen innerhalb der Klasse aller Studenten aus Gruppe 3 zahlenmäßig am stärksten vertreten sind. Ohne die genaue Angabe der entsprechenden Zahlen ist dieser Erklärungsversuch von vornherein wertlos. Er ist aber auch deshalb problematisch, weil weder von der Nein-Antwort auf ein tatsächliches Verständnis geschlossen werden darf (s.o.), noch von der philologischen/theologischen Ausbildung auf eine größere Befähigung, "schwierige Texte" zu interpretieren, geschlossen werden darf. Einerseits ist die Relation zwischen Textverständnis und Antwort auf Frage 14 möglicherweise gerade bei den Philologie- und Theologiestudenten mit einem besonders starken Verzerrungsfaktor behaftet, der in einigen Fällen auf einer fachbedingten Überheblichkeit dieser Studenten der Frage 14 gegenüber beruhen könnte und in anderen Fällen auf der Angst, ein mit der philologischen Ausbildung für inkompatibel gehaltenes Unverständnis zuzugeben. Andererseits ist es bei dem z.Zt. herrschenden Defizit an Interpretationsmethoden (in strengem Sinne) sehr fraglich, ob und inwieweit das Philologie- oder Theologiestudium die Befähigung fördert, Texte zu verstehen. Generell ist in diesem Zusammenhang einzuwenden, daß der Begriff "Textverständnis" von Bauer et al. überhaupt nicht problematisiert und nicht überlegt wird, auf welche Weise Textverständnis objektiv gemessen werden kann.

Der unzulässige Schluß von der Beantwortung der Frage 14 auf das Textverständnis wiederholt sich auf S. 88-89, wo Bauer et al. eine Korrelation zwischen der "Qualität des Textverständnisses" und den "Wertmaßstäben" postulieren. Gleichwohl scheinen die dort veröffentlichten Zahlen, die angeben, wie sich die Antworten zur Frage 12 ("Wie finden Sie das Gedicht?") auf die Gruppen G1 und G3 verteilen, sehr interessant zu sein:

In der Gruppe G1 kreuzten an:

- Bei "gut"	11.7%
- bei "mittelmäßig"	ca. 23.0%
- bei "schlecht"	" 23.0%
- bei "kann ich nicht beurteilen"	42.0%

Demgegenüber waren es in der Gruppe G3:

- Bei "gut"	55.6%
- bei "mittelmäßig"	25.9%
- bei "schlecht"	7.4%
- bei "kann ich nicht beurteilen"	11.1%

Eine genauere Bewertung dieser Zahlen läßt sich allerdings erst nach dem Vorliegen weiterer Ergebnisse aus empirischen Untersuchungen vornehmen. In solchen Untersuchungen müßte beispielsweise folgende Frage geklärt werden:

- Ist eine Korrelation zwischen den von den Vpn gemachten Angaben zur Verständlichkeit und den Bewertungen der Vpn generell nachzuweisen oder gilt sie nur für bestimmte Autoren bzw. bestimmte Texttypen oder einzelne Texte?

Erst nach Klärung dieser Frage kann z.B. abgeschätzt werden, warum die Verneinung von Frage 14 mit einem relativ hohen Prozentsatz an positiven Bewertungen und einem relativ niedrigen

Prozentsatz an negativen Bewertungen einhergeht. Prinzipiell sind hierfür verschiedene Erklärungsmöglichkeiten denkbar, etwa die folgenden:

- Die Qualität des Celan-Gedichtes ist - gemessen an bestimmten für die Vpn gemeinsamen Maßstäben - relativ hoch.
- Sobald ein Text für verständlich gehalten wird, wird er in der Regel auch für gut gehalten.
- Eine Unsicherheit in der Beurteilung des Celan-Gedichtes, die man sich selbst oder anderen gegenüber aber nicht wahrhaben möchte, führt dazu, den Text zugleich als "verständlich" und als "gut" hinzustellen.

An der hier geführten Diskussion dürfte noch einmal exemplarisch deutlich werden, daß bei der Auswertung und Interpretation von empirisch erhobenen Daten eine besondere Vorsicht erforderlich ist. Zugleich deutet sich aber auch an, daß empirische Untersuchungen in der Literaturwissenschaft zu sehr interessanten Ergebnissen führen können und für einen Erkenntnisfortschritt in der Literaturwissenschaft auch notwendig sind.

4.3 Unklar bleibt ferner, welchen Begriff von "Verständlichkeit" und "Verstehen" die Autoren haben. Ohne dem Schluß dieser Besprechung, wo dieses Problem detailliert diskutiert wird, vorzugreifen, sei bereits hier auf folgende Probleme hingewiesen:

- Es wird nicht deutlich, ob "Verständlichkeit" eine isolierbare Texteigenschaft sein soll oder aber eine Eigenschaft, die einem Text bei seiner Rezeption zugesprochen wird;
- ist letzteres der Fall, dann bleibt zu fragen, hinsichtlich welcher Bezugsgruppe der Text verständlich sein soll: hinsichtlich des jeweils befragten Lesers, hinsichtlich aller Leser oder im Hinblick auf eine bestimmte Klasse von Lesern;

- Die Autoren machen sich keine Gedanken darüber, wie "Verstehen" zu operationalisieren ist.

5. Zum Begriff der "Interpretation"

5.1 Unklar bleibt, was die Autoren unter dem Begriff der "Interpretation" verstehen wollen und aufgrund welcher Maßstäbe sie unterschiedliche Interpretationen bewerten. Zwar wird "Interpretation" auf S. 6 Anm. 2 als "makrosemantische Verarbeitung" des Primärtextes durch den Leser verstanden - ein höchst fragwürdiger Definitionsversuch übrigens -, doch die Auswertung der Fragen 28 und 29 läßt den Verdacht aufkommen, daß sich die Autoren ihrem Definitionsversuch gegenüber selbst inkonsequent verhalten. - Mit diesen beiden Fragen werden die Vpn aufgefordert, eine Reihe von Interpretationsvorschlägen von "Fadensonnen" zu beurteilen. Der erste dieser Vorschläge lautet:

"Man stellt sich beim Lesen dieses Gedichtes bildlich eine ausgedörrte Landschaft vor, die von Sonnenstrahlen durchströmt wird." (S. 53)

Bei der Auswertung stellen die Autoren nun fest, daß diese Interpretation vor allem von jener Gruppe von Vpn als zutreffend bezeichnet wird, die angibt, der Text sei unverständlich (46.5% dieser Gruppe), wohingegen in der Gruppe der Vpn, die den Text ohne Einschränkung verständlich finden, nur 24.2% mit dieser Interpretation übereinstimmen. Diesen Befund interpretieren die Autoren folgendermaßen:

"Diese Konstellation scheint die schon in anderem Zusammenhang erwähnte Korrelation zwischen der Unverständlichkeit des Textes und einem von der visuellen Raumvorstellung ausgehenden Inter-

pretationsansatz erneut zu bekräftigen. Es muß angenommen werden, daß auf einer differenzierteren Stufe des Verständnisprozesses diese Vorstellung einer abstrakten Generalisierung weicht, die in einer Art Sekundärreaktion die sog. anthropologisch-metaphysische Interpretation des Textes vorbereitet." (S. 91) Das heißt: der oben zitierte Interpretationsvorschlag wird von den Autoren offenbar schlechter bewertet als z.B. die Interpretationen 7 und 9. Diese lauten:

"Trotz der Sinnlosigkeit des Lebens scheint es doch etwas Ewiges zu geben." (S. 52)

"Aus der Öde des Daseins rettet ein Hoffnungsgedanke, eine jenseitige Aufgabe." (S. 53)

Hier müßte zumindest verdeutlicht werden, auf der Basis welchen Wertsystems eine solche Wertzumessung berechtigt sein könnte.

Im Kapitel III,4.4 kommentieren und interpretieren die Autoren die Assoziationen zu den Metaphern "Fadensonnen", "baumhoher Gedanke" und "Lichtton". Diese Interpretationen sind wegen der lückenhaften rezeptionstheoretischen Vorstellungen und ebenso lückenhaften semantischen Theorien der Autoren z.T. fragwürdig.

Die Autoren unterscheiden drei "sich in den Assoziationen manifestierende Rezeptionsstufen":

- 1) "Eine fast wörtliche, unreflektierte Textproduktion (Synonym, Paraphrase), die bei den meisten Vpn erscheint und keine Analyse oder Verarbeitung der Metapher beinhaltet. Die Bedeutung dieser Stufe liegt darin, daß durch diese sprachliche Umsetzung der Prozeß der Aktivierung in Gang gesetzt wird und anfängliche Verständnisschwierigkeiten der Vpn offenbar werden." (S. 133-134)

- 2) "Die mehr oder weniger bewußte Analyse des Wortes, d.h. Einzelkomponenten des Wortes werden erkannt und sprachlich realisiert (Theorie, 6.2.1). An diesem Punkt des Rezeptionsvorgangs vollzieht der Leser also 'intuitiv' eine Semische Analyse. Hierbei sind zwei Assoziationsformen zu unterscheiden: Einerseits werden die Wörter assoziiert, die innerhalb des Sprachsystems der bestimmten Sache zugeordnet werden (exteriore Komponenten), andererseits kommen Adjektive hinzu, die ein subjektiv-affektives Moment mit diesem Gegenstand verbinden (Konnotationen)." (S.134)

Ad 1)

Es ist nicht einzusehen, warum die Paraphrasierung - die Ersetzung eines Wortes durch ein Synonym ist nichts anderes - nicht als Verarbeitungsprozeß aufgefaßt werden soll. Es gibt im Gegenteil gute Gründe dafür, die Paraphrasierung zu den Textverarbeitungsprozessen zu zählen.¹ Noch weniger einsichtig ist es, wieso dadurch, daß jemand eine Metapher paraphrasiert, Verständnisschwierigkeiten offenbar werden. Ebensogut könnte das Gegenteil der Fall sein. Durch eine manifeste Paraphrasierung könnte ein sehr vorsichtiges interpretatorisches Vorgehen offenbar werden im Gegensatz zu einem munteren Drauflosinterpretieren, mit dem ein Interpret lediglich vorgeben könnte, er meine zu verstehen.

Ad 2)

Hierzu muß festgestellt werden, daß ohne ausreichende theoretische Absicherung "exteriore Komponenten" und "Konnotationen" nicht leicht voneinander zu scheiden sind; denn was zum allgemeinen Sprachsystem, was zum idiolektalen System gehört, wird

¹ Vgl. Wienold und Rieser (1974), Petöfi (1974), Wirrer (1976b)

durch eine vorgängige Grammatik entschieden. Welche Grammatik hier zugrundeliegt bzw. welchen theoretischen und empirischen Status diese Grammatik haben soll, hätte von Bauer et al. zumindest angedeutet werden müssen.

Die mangelhafte zugrundegelegte Theorie wird auch an anderen Stellen der Ergebnisinterpretation deutlich. Dazu nur ein Beispiel:

"Auffallend bei den weiterführenden Assoziationen zu Faden ist eine Häufung der Assoziationen aus dem Bereich Alltag (Treppe, Schleier, Netz etc.). Unter dem Naturbereich sind Assoziationen wie: Schlangen, Krake, Regenwürmer aufgeführt." (S. 136)¹

Für welche Personengruppen soll diese Zuordnung gelten? Ist es selbstverständlich, daß für jedermann heutzutage ein Schleier zum Bereich "Alltag" gehört? Und vielleicht wird ein Regenwurmforscher "Regenwürmer" eher dem Bereich "Alltag" als "Schleier" zuordnen wollen. Jedenfalls gelten solche Zuordnungen empirisch stets nur für bestimmte Gruppen von Personen - und das dürfte in einer empirisch ausgerichteten Arbeit nicht übersehen werden.

Frage 20 lautet:

"An welche Werke aus Dichtung, Bildender Kunst oder der Musik denken Sie bei:

1. Fadensonnen 2. baumhoher Gedanke 3. Lichtton "

¹ Vgl. auch S. 123

Durch diese spezifische Frage sollen die freien Assoziationen im Hinblick auf die Analyse des Lesererwartungshorizonts erweitert werden.

"Erwartungshorizont" soll hier nicht streng gattungstypologisch verstanden werden, sondern als eine Art Rahmen der persönlichen Erfahrungen des Lesers (vgl. Theorie 3.3 und 3.4). Dieser umfaßt die Erwartungen, die beim jeweiligen Leser aus den individuellen kulturellen Erfahrungen im weitesten Sinne resultieren und durch die unterschiedlichen Bildungsvoraussetzungen der Vpn bedingt sind, sowie das spontane Assoziieren beliebiger Kunstwerke, welches sich aus der persönlichen Situation unmittelbar bei der Lektüre des vorliegenden Textes ergibt."¹

Hier fragt sich, warum die Autoren noch einmal speziell Assoziationen aus diesem Bereich erfragen und damit diese Bereiche vor anderen ausgezeichnet werden, die ebensogut zu den Bildungsvoraussetzungen und zur persönlichen Situation der Vpn "unmittelbar bei der Lektüre" zählen können. Präferieren die Autoren bestimmte Assoziationen zu den Metaphern oder ergibt sich das Erfordernis zu dieser Frage aus der Voruntersuchung? Über beides lassen die Autoren die Leser ihres Buches im unklaren.

6. Einige grundsätzliche Probleme empirischer Untersuchungen in der Literaturwissenschaft

6.1 Im folgenden Kapitel soll im Zusammenhang mit "Text und Rezeption" auf einige grundsätzliche Probleme hingewiesen

¹s. 155/56

werden, denen sich eine Literaturwissenschaft gegenübersteht, die empirisch forschen will.¹

Empirische Hypothesen sind zunächst weder wahr noch falsch, sie werden durch Experimente überprüft, aber diese Überprüfungen sind nicht endgültig, sind keine Wahrheitsbeweise. Häufig beinhalten empirische Hypothesen Wahrscheinlichkeitsaussagen über Kausalzusammenhänge, die man nach einer Überprüfung durch einen Versuch (ein Experiment) entweder weiter behaupten darf oder verwerfen muß. Um aber eine empirische Hypothese überprüfen zu können, ist es notwendig, vorher die in ihr verwendeten Begriffe zu operationalisieren.

Schauen wir uns daraufhin eine Hypothese, die in "Text und Rezeption" aufgestellt wird, einmal näher an.² Auf Seite 35 finden wir die Aussage: "Nur anhand des schon vorliegenden Materials waren die Differenzierungen der einzelnen Fragenkomplexe (...) und die Präzisierung der für den Versuch entscheidenden Arbeitshypothesen möglich."

Diese Aussage legt den Schluß nahe, die entscheidenden Hypothesen konnten in einem Versuch empirisch überprüft werden, da u.a. die in ihnen verwendeten Begriffe operationalisiert worden sind. Wie sieht das bei Bauer et al. in der Praxis aus? Sie schreiben (S. 29):

¹ Die Freiburger Forschungsgruppe spricht selber von 'einem ersten Versuch', der einen Diskussionsbeitrag zur Frage nach der Wirkung von Texten darstellt.

² Es wurde nur ein Beispiel, das symptomatisch scheint, hier ausgewählt; die Kritik trifft jedoch alle implizit oder explizit bei Bauer et al. aufgestellten Hypothesen.

"Die Exzeptionalität des Textes, die Besonderheit der einzelnen Metaphern, von der die Assoziationen des Lesers ausgehen, kann als Wirkungskonstante des Gedichts 'Fadensonnen' angenommen werden. Diese These wird durch die am Semantischen Differential durchgeführten Faktorenanalysen weitgehend bestätigt; sie geht von folgender Arbeitshypothese aus:

In dem Modellfall kühner Metaphorik, gekennzeichnet durch ihren Abstand von der bestehenden Metapherntradition und ihre Schockwirkung auf den Leser der Gegenwart, ist von seiten des Lesers ein dynamischer Assoziationsprozeß zu erwarten, in dem die gegebenen Wortstimuli (Neologismen) verarbeitet werden. Dabei gelangt der Leser durch die semantische Feinanalyse (...) zur hohen Aktivierung der Wortsensibilität und überwindet damit seine anfänglichen, in Einzelelementen des Textes begründeten Verständnisschwierigkeiten; andererseits vermag er durch die an die Metapher anknüpfenden Wortassoziationen in einen aktuellen Wirklichkeitsbezug vorzustoßen, womit bereits die Grundlage seiner individuellen Textdeutung (...) gegeben ist."

Diese Textstelle soll nun analysiert werden.

Bauer et al. stellen die Hypothese (H_0) auf:

"(H_0) Die Exzeptionalität des Textes, die Besonderheit der einzelnen Metaphern, von der die Assoziationen des Lesers ausgehen, kann als Wirkungskonstante des Gedichts 'Fadensonnen' angenommen werden."

Um einen besseren Überblick über (H_0) zu ermöglichen, soll diese Hypothese teilformalisiert werden.

(1) Es gibt eine Eigenschaft E^1 des Gedichttextes, die Ex-

¹ Die Eigenschaft E wird "letztlich begründet in der Abweichung der Metapher von der Sprachnorm." (S. 28)

tionalität genannt wird¹.

- (2) Die Besonderheit b der einzelnen Metaphern (m_1, m_2, m_3) des Gedichts 'Fadensonnen' begründet $E(g)$ also:

$$E(g) = b(m_1) \wedge b(m_2) \wedge b(m_3)$$

- (3) $E(g)$ kann angenommen werden als eine Eigenschaft, die konstant die Wirkung w auslöst.

$w = \text{def. Stimulation des Lesers zu Assoziationen}$

In(H_0) wird also ein Kausalzusammenhang behauptet: $E(g)$ verursacht Wirkung w der $E(g) \Rightarrow \text{Auftr.}(w)$

Es schließt sich die Behauptung an, diese Hypothese sei weitgehend bestätigt, und zwar durch die am Semantischen Differential (SD) durchgeführten Faktorenanalysen (FA_1, \dots, FA_n).

Das heißt:

- (4) Es gibt Faktorenanalysen, die die Gültigkeit der Hypothese (H_0) bestätigen können.
- (5) Es gibt ein SD (Semantisches Differential), an dem FA_1, \dots, FA_n durchgeführt werden können.
- (6) In der vorgelegten Untersuchung wurde (durch FA am SD durchgeführt) probabilistisch ("weitgehend") bestätigt, daß (H_0) weiter angenommen werden darf, bzw. daß (H_0) gültig ist.

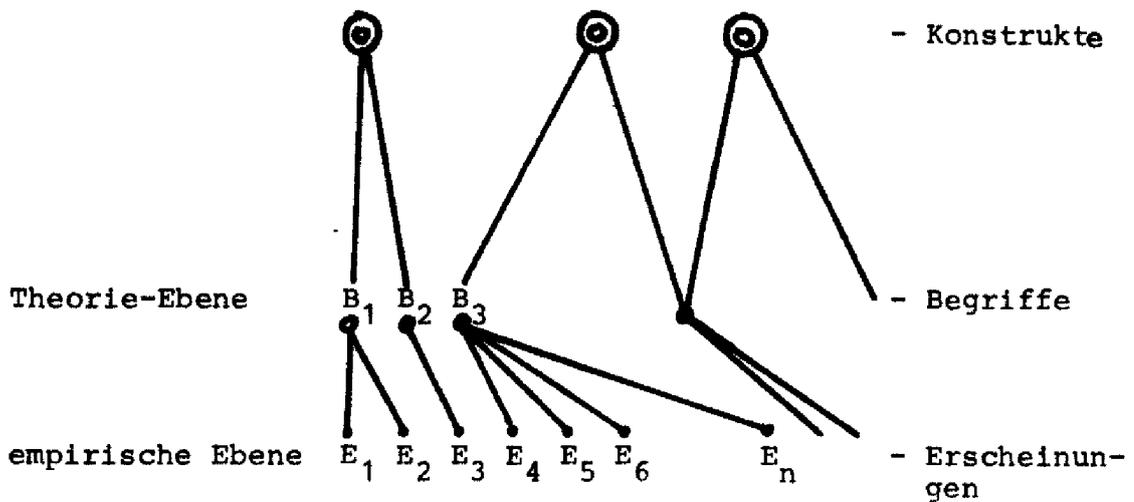
¹Einige Zeilen vorher wird von 'Ungewöhnlichkeit' des Textes in eben dem Sinn gesprochen (vgl. S. 28, S. 29).

Und nun folgt in dem zitierten Text - nachdem die Gültigkeit von (H_0) als 'weitgehend bestätigt' behauptet worden ist - die Einführung einer weiteren "Arbeitshypothese", von der (H_0) aber "ausgeht".

- (7) Es gibt Metaphern M_1, \dots, M_n
- (8) Es gibt eine Eigenschaft K von Metaphern, die Kühnheit genannt wird
- (9) Diese Eigenschaft ist gekennzeichnet durch die Merkmale x_1 und x_2 :
 - (9a) x_1 = def. Abstand von der bestehenden Metaphertradition,
 - x_2 = def. Schockwirkung auf den Leser der Gegenwart
- (9b) Es gibt dynamische Assoziationsprozesse
- (10) K Metaphern lösen in jedem Rezipienten einen dynamischen Assoziationsprozeß A aus
- (11) A besteht aus zwei verschiedenen Stufen St_1, St_j , die aufeinanderfolgen
- (12) St_1 = semantische Feinanalyse des Lesers (auf St_j folgt hohe Aktivierung der Wortsensibilität beim Leser, auf diese folgt: Leser überwindet seine ... Verständnisschwierigkeiten)
 - St_j = Vorstoßen des Lesers zu einem aktuellen Wirklichkeitsbezug durch Assoziationen (auf der Basis von St_j folgt die individuelle Textdeutung des Lesers).

Wir brechen an dieser Stelle die Rekonstruktion des Arguments ab. Um zu zeigen, daß mit der Argumentation der Autoren keine empirische Fragestellung bzw. Hypothese aufgestellt worden ist, erscheint es hilfreich, zunächst einmal einige stark vereinfachende, schematisierende wissenschaftstheoretische Vorbemerkungen zu machen.

Im Bereich empirischer Wissenschaften wird streng zwischen beobachtbaren Erscheinungen und gedanklichen Konstrukten unterschieden. Der Zusammenhang lässt sich an folgender Skizze erläutern:



Skizze 1

Die schwarzen Punkte stellen beobachtbare Erscheinungen auf der empirischen Ebene dar. Auf der untersten Ebene des theoretischen Aufbaus werden diese Erscheinungen durch exakt definierte Begriffe (einfache Kreise) repräsentiert. Über diesen Begriffen erheben sich weitere Konstrukte (Doppelkreise), die gedankliche Verbindungen zwischen den Begriffen herstellen usw. Eine Theorie in empirischen Wissenschaften setzt nun Erscheinungen, also: E_1, \dots, E_n zueinander in Beziehung. Ganz allgemein lassen sich daher wissenschaftliche Gesetzeshypothesen (vereinfacht für zwei Erscheinungen E_1, E_2) auf folgende Form bringen:

"Immer wenn E_1 , dann E_2 ".

Diese Gesetzmässigkeit kann am tatsächlichen Auftreten der Erscheinungen überprüft werden.

Ein Beispiel: "Immer wenn einer lange Haare hat, dann stinkt er."

E_1 : lange Haare haben

E_2 : stinken

Unter der Annahme, daß die Gesetzmässigkeit "Immer wenn für jemand gilt E_1 , dann gilt für ihn auch E_2 " gilt, folgt aus dem Auftreten der Erscheinung E_1 (ist mit dem Auftreten von E_1 immer verbunden) logisch notwendig das Auftreten von E_2 .

Unter Verwendung logischer Symbole:

$\forall x (E_1, x \supset E_2, x)$	Gesetz
E_1, y	Bedingung
<hr/>	
$\therefore E_2, y$	Folgerung

Die Gültigkeit der Gesetzeshypothesen wird bekräftigt, solange immer E_2 auf E_1 folgt, bzw. zusammen mit E_1 auftritt. Sobald Fälle auftreten, in denen E_1 nicht E_2 nach sich zieht, bzw. E_1 nicht in Zusammenhang mit E_2 auftritt, gilt die Gesetzeshypothese als widerlegt. Eine Theorie in den empirischen Wissenschaften baut sich aus einer Vielzahl solcher gesetzmässiger Hypothesen auf, d.h. jedes ihrer Teile ist an beobachtbaren Erscheinungen überprüfbar. Theorien, die nicht überprüfbar sind, werden von den empirischen Wissenschaften nicht anerkannt.

Neben deterministischen Hypothesen (aus E_1 folgt notwendig E_2), in denen die Gesetzmäßigkeiten in Form strikter Abhängigkeiten formuliert werden, gibt es probabilistische Hypothesen, die lediglich eine bestimmte Wahrscheinlichkeit (p) für einen gesetzmässigen Zusammenhang angeben. Folgt etwa bei 100maligem Auftreten von E_1 nur 60mal E_2 , so lautet die gesetzmässige Hypothese:

$$p (1 \times (E_1, x \supset E_2, x)) = 0,60$$

In der Literaturwissenschaft nun müssen häufig Aussagen über gesetzmässige Zusammenhänge gemacht werden, in denen ein ganzer Komplex von Bedingungen wirksam ist. Eine deterministische Theorie müßte die genaue Anzahl aller Bedingungen und deren Einfluß kennen, um jeden Einfluß genau vorhersagen zu können, da ja bereits ein signifikantes Gegenbeispiel die Theorie falsifiziert. Eine probabilistische Theorie kann dagegen auch Schlüsse ziehen aus Bedingungskomplexen, deren einzelne Elemente nicht oder noch nicht klar erfaßbar sind. Die Unschärfe in der Beobachtung der einzelnen wirkenden Bedingungen wird durch die Unschärfe der Wahrscheinlichkeitsangabe aufgefangen.

Man darf unter Berücksichtigung der vorangegangenen Ausführungen bei den zitierten Hypothesen von Bauer et al. davon ausgehen, daß Wahrscheinlichkeitsaussagen intendiert wurden. Auch wenn diese nicht explizit gemacht worden sind, deutet das "weitgehend bestätigt" in die vermutete Richtung. (H_0) lautet dann:

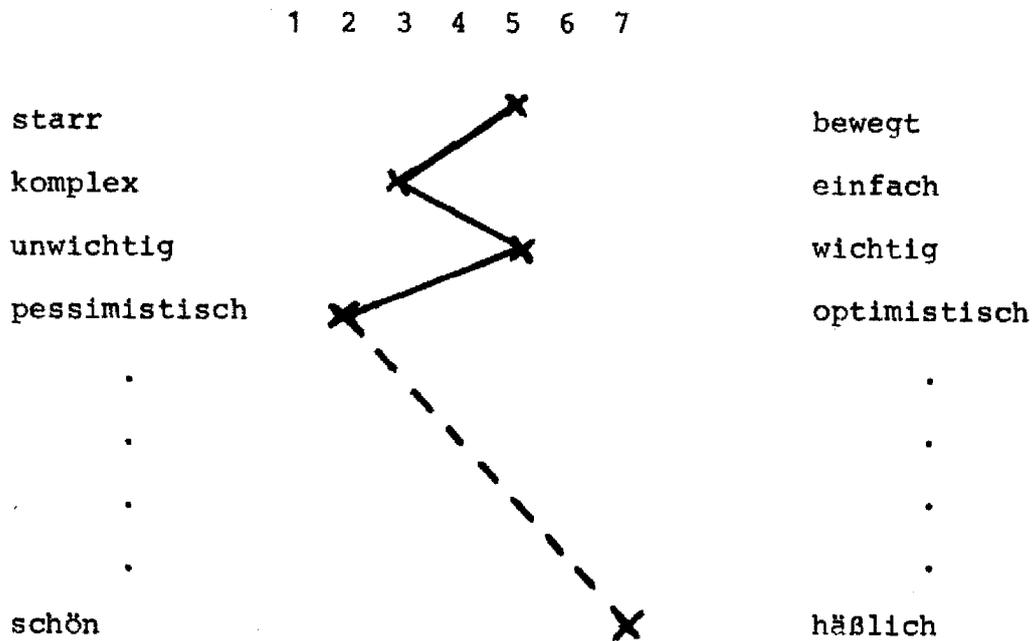
$$H_0 : p ((g) \supset \text{Auftr.}(w)) = ?$$

Da bei Bauer et al. kein statistisches Signifikanzniveau (1% Niveau s.s. = sehr signifikant, 5% Niveau s. = signifikant) angegeben wird, liegt im strengen Sinne mit (H_0) gar keine empirische Hypothese vor, die überprüft werden konnte. Die Behauptung, daß die These überprüft und weitgehend mit Hilfe der durchgeführten Faktorenanalysen bestätigt wurde, muß also problematisiert werden.

6.2 Da das Semantische Differential (SD) eines der wichtigsten Meßinstrumente für empirische Untersuchungen innerhalb der Literaturwissenschaft ist, zunächst einige generelle Bemerkungen zu diesem Themenkomplex und zur Faktorenanalyse (FA)¹. Das Semantische Differential ist ein von Osgood et al.² entwickeltes Meßinstrument um Bedeutungen (Bedeutungen ausschließlich verstanden als emotive Konnotationen) zu registrieren, die ein Individuum einem Objekt zumißt. Das Verfahren besteht darin, daß man den Vpn eine Reihe bipolarer Adjektivskalen vorlegt und sie zu einer Stellungnahme darüber veranlaßt, wo in etwa sie die Bedeutung des in Frage stehenden Objekts oder Stimulus' auf einer siebenstufigen Skala einordnen würden. Dazu ein stark vereinfachendes Beispiel: Die Vpn des Freiburger Versuchs sollten u.a. die Bedeutung einschätzen, die die Metapher "Fadensonnen" für sie persönlich hat, z.B. zwischen starr/bewegt, komplex/einfach, unwichtig/wichtig, pessimistisch/optimistisch, häßlich/schön usw.

¹Auf die theoretischen Implikationen dieses Meßinstruments und des Rechenverfahrens sowie auf praktische Probleme bei der Anwendung kann hier leider nicht eingegangen werden.

²Vgl. Osgood et al. (1957)



Skizze 2 (Polaritätsprofil)

Jede Vpn schätzt also die Bedeutung von "Fadensonnen" ein, indem sie entweder 'neutral' (Pos. 4) in Bezug auf jedes Adjektivpaar einstuft oder in drei Steigerungsstufen (3), (2), (1) und (5), (6), (7) sich für das eine oder andere Adjektiv entscheidet. Die aufgezeichneten Einstufungen jeder Vpn bezüglich aller Adjektivpaare nennt man Polaritätsprofile¹.

Man kann nun eine Gruppe von Bedeutungseinschätzungen zusammenfassen, z.B. die Polaritätsprofile der Schüler bis zu 15 Jahre, und sie mit denen anderer Gruppen, z.B. denen der Schü-

¹Der Ausdruck geht auf Peter R. Hofstätter zurück, der u.W. als erster das SD im deutschsprachigen Raum erprobte.

ler über 15 Jahre und denen der Studenten vergleichen. Nach Auswertung mit Hilfe statistischer Verfahren erhält man bestimmte Ähnlichkeitsrelationen zwischen den Bedeutungsdimensionen. Man kann aber auch die Bedeutungseinschätzungen verschiedener Objekte (bei Bauer et al. die Bedeutungseinschätzungen der verschiedenen Metaphern m_1 , m_2 , m_3) miteinander vergleichen.

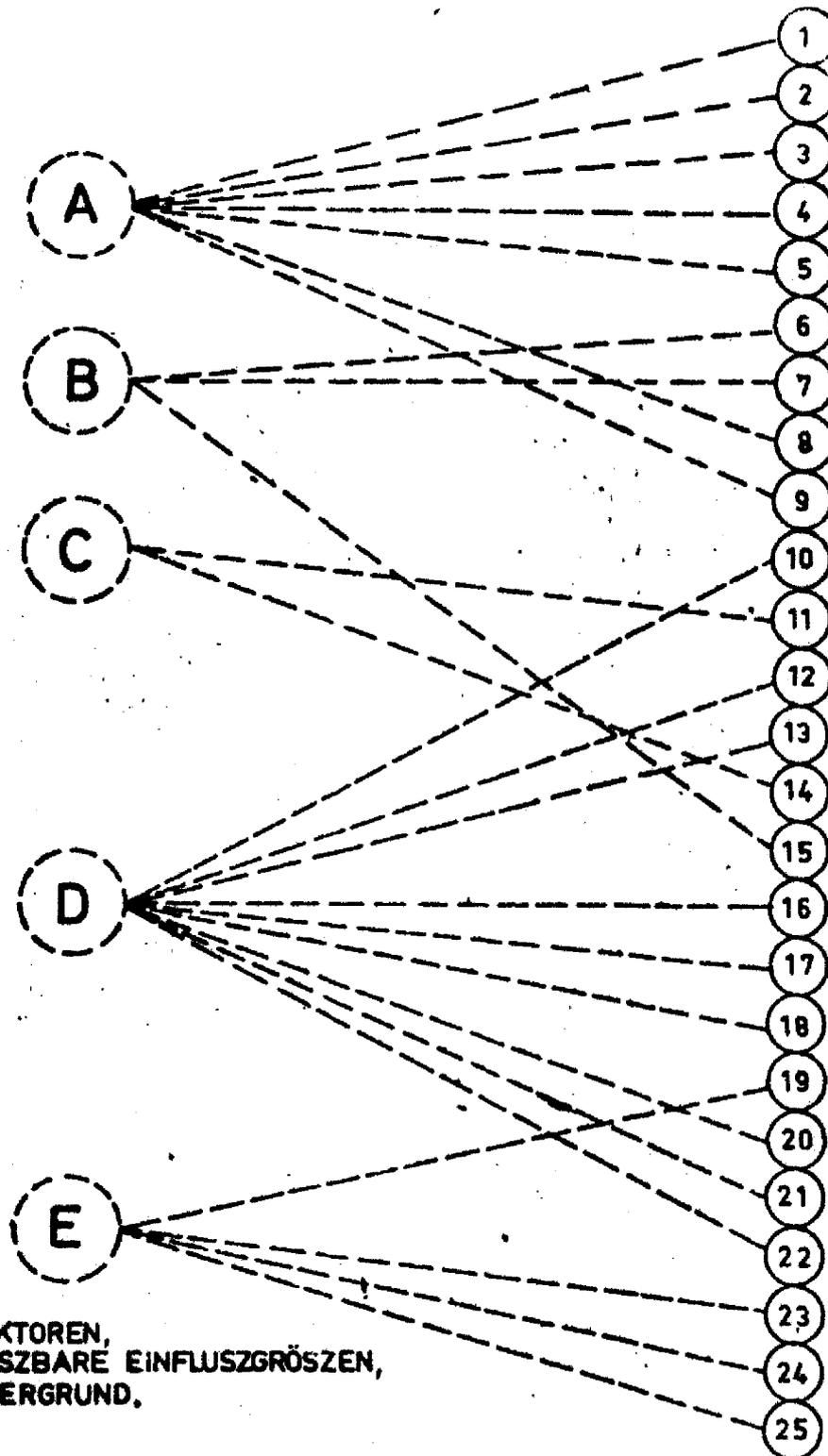
Soviel zunächst einmal über das von Bauer et al. verwendete Meßinstrument und nun einige Bemerkungen zum Verfahren der Faktorenanalyse.

"Hauptziel der Faktorenanalyse ist die Ableitung hypothetischer Größen oder Faktoren aus einer Menge beobachteter Variablen. Die Faktoren sollen möglichst einfach sein und die Beobachtungen hinreichend genau beschreiben und erklären."²

"Das faktorenanalytische Modell geht immer von der ... Annahme aus ..., daß das Meßbare nur eine Erscheinungsform von Größen ist, die im Hintergrund stehen und die man direkt nicht messen kann. Man möchte ... wissen, ob sich aus den Variablen, die man beobachtet hat, eine Größe isolieren läßt, ein sogenannter Faktor, der die beobachteten Zusammenhänge erklärt. Das Wort Faktor wird hier in einem anderen Sinne verwendet als im üblichen Sprachgebrauch. Es handelt sich dabei um eine mathematische Größe, die aus Beobachtungen abgeleitet ist."³

¹ Überla (1971), S. 3

² Überla (1971), S. 2f.



FAKTOREN,
NICHT MESZBARE EINFLUSZGRÖSSEN,
HINTERGRUND.

VARIABLEN,
MESZBARE BEOBACHTUNGSGRÖSSEN,
VORDERGRUND.

SKIZZE 1:

DAS ZUEINANDER VON VARIABLEN UND FAKTOREN.
DIE DURCH DIE FAKTORENANALYSE AUFGESTELLTE HYPOTHESE
IST GESTRICHELT GEKENNZEICHNET, DIE BEOBACHTUNGSDATEN
SIND AUSGEZOGEN.

(NACH ÜBERLA, KARL : FAKTORENANALYSE S. 4)

Faktoren sind, anders formuliert, gemeinsame Bedingungen für eine Reihe zusammengehöriger Merkmale. Skizze 3 veranschaulicht das hypothetisch für die Variablen (die Adjektivskalen) des bei Bauer et al. verwendeten SDs.

"Wenn also unter Faktorenanalyse die methodische Möglichkeit verstanden werden soll, aus den Zusammenhängen zwischen einer Gruppe zusammengehöriger Merkmale die ihnen gemeinsamen Bedingungen herauszuanalysieren, dann heißt das ...: Jeder Gegenstand, sofern er nur auf Merkmale zurückführbar ist, die eine geeignete Quantifizierung gestatten¹, kann einer Faktorenanalyse unterzogen werden"².

Dies gilt selbstverständlich auch für das komplexe Gefüge der Metaphernbedeutung bzw. der Textbedeutung (der Bedeutung des ganzen Gedichts "Fadensonnen"), das wir durch Aufdeckung der zugrundeliegenden Faktoren überschaubarer machen können.

¹ In bezug auf die Untersuchung von Bauer et al. gilt die Einschränkung von Ebner/Clauss: "In solchen Fällen (Anwendung des SDs, - Anm. der Verf.) pflegt man oft die mehrfach gestufte Variable als metrisch skaliert aufzufassen und zu behandeln (...). Das ist jedoch streng genommen nicht berechtigt. Man sollte sich darüber klar sein, daß derartige subjektive Graduierungen nach wie vor Nominalklassifikationen, allenfalls Ordinalskalierungen darstellen und keinesfalls im strengen Sinne 'Messungen' sind." (Clauss, Günther und Ebner, Heinz; Grundlagen der Statistik. S. 102).

² Roth (1965), S. 283

"Entscheidend ist, daß die Faktorenanalyse eine differenzierte Hypothese über die Struktur des Zueinanders der Variablen und Faktoren möglich macht, ohne daß man vorher eine bestimmte Struktur annehmen oder bereits kennen muß. Sie wird aus Beobachtungen abgeleitet. Die erhaltene Hypothese, . . . , läßt sich in weiteren Experimenten überprüfen."¹

6.3 Kommen wir auf die "Bestätigung" der These von Bauer et al. zurück. Mit dem SD werden Daten erhoben, die die emotiven Konnotationen (die Bedeutung) ausmachen, die die Vpn den Metaphern m_1 , m_2 und m_3 zumessen. Ob diese Bedeutungszuordnungen (bei Bauer et al. die Assoziationen) direkt durch m_1 , m_2 und m_3 oder erst durch die Aufforderung, das SD auszufüllen, ausgelöst wurden, bleibt zunächst einmal offen. Wenn nun mit den Daten, die die drei Metaphern betreffen (nicht das SD! Deshalb können die Faktorenanalysen auch nicht "am SD" durchgeführt worden sein) Faktorenanalysen durchgeführt werden, so führen diese erst zu einer Hypothesenbildung, nämlich darüber, welche und wieviele Faktoren die Bedeutung (die emotiven Konnotationen) der Metaphern erklären und beschreiben. Diese Hypothesen können dann durch weitere Faktorenanalysen mit den Daten von anderen Gegenständen (von anderen Metaphern, von dem ganzen Gedicht usw.) dahingehend überprüft werden, ob die Faktoren in Anzahl und Art bezüglich dieser anderen Objekte gleich oder ähnlich sind oder ob die Hypothesen nur für die zunächst untersuchten Gegenstände gelten.

Eine Überprüfung der Hypothese
(H_0): $p(e(g) \supset \text{Auftr.}(w)) = x$
ist mit Hilfe der FA nicht möglich.

¹Überla (1971), S. 3f.

Zusammenfassend läßt sich nun folgender Fragenkatalog aufstellen:

- a) Kann die behauptete Eigenschaft E durch Messung festgestellt werden?
- b) Wenn dies möglich ist, kommt die Eigenschaft E dann nur in Zusammenhang mit der Besonderheit der Metaphern m_1 , m_2 und m_3 des Gedichts g vor, oder kommt diese Eigenschaft E allen kühnen Metaphern zu?
- c) Kommt die Eigenschaft E auch sogenannten 'normalen' Metaphern zu, gilt also: $E(M)$.
- d) Kommt dieselbe Eigenschaft auch anderen Objekten zu?

Auf diese Fragen wird im nächsten Unterkapitel über die FA-Untersuchungen (6.4) eine vorläufige Antwort gegeben.

Auf die "Arbeitshypothese" braucht wohl nach den bisher gemachten Ausführungen nicht mehr weiter eingegangen zu werden; ähnliche Hypothesen finden sich an vielen Stellen in "Text und Rezeption", u.a. die Hypothesen von der "Dunkelheit" der Metapher als konstituierendem Element eines sog. hermetischen Gedichts (S. 26)¹, die Hypothese von den drei Rezeptionsstufen (S. 11, S. 58 ff., S. 63 f. u.a.), die Hypothese, daß ein Leser mit durchschnittlicher Sprachkompetenz in der Lage sei, zu einem Verständnis eines multivalenten zeitgenössischen Textes zu gelangen (S. 223), die Hypothese über Leser mit positivem, negativem oder mittlerem Textverständnis (S. 85) u.a.m.

¹ Diese Hypothese wurde aber von den Freiburgern nach den Vorversuchen nicht mehr weiter verfolgt.

Untersuchungen im Bereich der Literaturwissenschaft, die mit empirischen Methoden durchgeführt werden, sollten sich in Zukunft auf nur wenige Fragestellungen, die präzise gefaßt werden können und auf die Überprüfung nur weniger, aber exakt formulierter Hypothesen beschränken.

6.4 Mit den folgenden Erörterungen soll auf einige Schwierigkeiten hingewiesen werden, die sich bei der Durchführung eines Versuchs ergeben, wie ihn Bauer et al. unternommen haben; dabei sollen vor allem die Schwierigkeiten diskutiert werden, die die Strukturierung eines geeigneten Meßinstruments für empirische Fragestellungen im Bereich der Literaturwissenschaft mit sich bringen sowie Schwierigkeiten bei der Auswertung der mit Hilfe dieses Meßinstruments erhobenen Daten.

Ein genereller Vorbehalt wird gegenüber Untersuchungen wie der Freiburger noch einige Zeit gemacht werden müssen. Schlüsse aus den erhobenen Daten können nämlich nur bedingt verallgemeinert werden, denn aufgrund mangelnder finanzieller Unterstützung solcher Versuche wird die Literaturwissenschaft leider noch noch lange eine "Literaturwissenschaft der Philologiestudenten" bzw. in einer Variante eine "Literaturwissenschaft der Gymnasiasten und anderer Schüler" bleiben. Diese Einschränkung betrifft vor allem den empirisch-experimentellen Teil solcher Untersuchungen, denn Schüler und Studenten sind die einzigen Vpn, die noch in großer Zahl unbezahlt für Tests zur Verfügung stehen. Bei Schüler- und Studentenbefragungen darf aber nicht von vornherein davon ausgegangen werden, daß die Daten "objektiv", "gültig" (valide) und "zuverlässig" (reliabel) erhoben werden können. Gerade bei Schüler- und Studentenbefragungen schleichen sich immer wieder größere und kleinere, zufällige und systematische Fehler ein. Diese Fehler können aber zu einer groben Verzerrung der Ergebnisse und damit zu einer Verzerrung der Gesamtuntersuchung führen.

Dieser Gedankengang soll dadurch etwas vertieft werden, daß wir im folgenden die drei Gütekriterien empirischer Untersuchungen: "Objektivität", "Reliabilität" und "Validität" ganz kurz zu erläutern versuchen.

Die "Objektivität" eines Meßverfahrens bestimmt den Grad der Unabhängigkeit der Untersuchungsergebnisse von zufälligen oder systematischen Verhaltensvariationen des Untersuchers (während der Versuchsdurchführung), die ihrerseits zu Verhaltensvariationen der Vpn führen und deren Ergebnisse beeinflussen.¹ Objektivität muß aber nicht nur bei der Durchführung eines Experiments, sondern auch bei der Auswertung und Interpretation der erhobenen Daten nachgewiesen werden. Das Kriterium der "Reliabilität" (der Zuverlässigkeit) gilt dann als erfüllt, wenn die Messung keine zufälligen Fehler aufweist oder wie Manfred Stosberg ausführt:

"Reliabilität kann ... definiert werden als das Ausmaß, in dem eine Skala zu konsistenten Meßergebnissen führt, wenn sie mehrmals zur Messung einer gleichbleibenden Eigenschaft eingesetzt wird ... Wenn man die Reliabilität eines Meßinstruments bestimmen will, versucht man ... herauszufinden, in welchem Ausmaß die Ergebnisse, die mit Hilfe des betreffenden Instrumentes erzielt werden, durch zufällige Störfaktoren beeinflusst sind."²

¹ vgl. Lienert (1969) S. 13

² Stosberg (1972) S. 36

"Valide" (gültig) ist eine Messung, wenn sie keine systematischen Fehler aufweist, wenn alle Störvariablen unter Kontrolle gebracht worden sind. Da diese Forderung in der Praxis kaum erfüllbar ist, kann immer nur von vorläufiger Validität gesprochen werden, aber auch diese vorläufige Validität ist bei Schülerbefragungen selten gewährleistet.

Ob nun diese drei Gütekriterien empirischer Untersuchungen bei der Durchführung des Freiburger Versuchs berücksichtigt wurden (und ggf. wie weit eingeschränkt), darüber schweigen sich Bauer et al. leider aus. An einem Beispiel der Auswertungsobjektivität soll die sich aus diesem Vorgehen ergebende Problematik - allerdings erst am Ende dieses Unterkapitels erörtert werden.

Zunächst einige Anmerkungen zur Fragebogenstrukturierung bei Bauer et al. Es empfiehlt sich, bei der Entwicklung eines Fragebogens nicht nur inhaltliche, sondern auch formale Kriterien zu berücksichtigen. Diese Kriterien, z.B. bezüglich der Fragen- und Statement-Formulierung oder Intervieweranweisungen, kann man in der einschlägigen Literatur (u.a. Süllwold, 1969, Guilford 1954) nachlesen; sie gelten für alle empirischen Untersuchungen.

Gegen diese Kriterien verstößt u.E. die Doppelfrage 28/29 im Fragebogen von Bauer et al., die die Vpn in physischer und psychischer Hinsicht überfordert haben dürfte.

" 28) Beurteilen Sie bitte folgende Interpretationsansätze zu dem Gedicht "Fadensonnen" : (S. 29)

1. Man stellt sich beim Lesen dieses Gedichts bildlich eine ausgedörrte Landschaft vor, die von Sonnenstrahlen durchströmt wird.
2. Das Wichtigste in diesem Gedicht ist die Trennung in einen unteren und einen oberen Bereich und die vermittelnde Bewegung zwischen beiden.
3. Dem Leser kommt bei diesem Gedicht der Gedanke an die Erschaffung der Welt: Und die Erde war wüst und leer ... (Genesis)

Auch in anderer Hinsicht scheinen die Versuchsleiter ihren Vpn physisch und psychisch sehr viel zugemutet zu haben. Die Beantwortung von drei SDs (Fragen 23, 24, 25) hintereinander dürfte jedenfalls auch bei im Umgang mit derlei Meßinstrumenten geübten Vpn zu starken Ermüdungserscheinungen führen, noch dazu, wenn direkt danach - ohne entspannende Pufferfragen - eine Gedichtinterpretation (Frage 26) abgerufen wird und kurz darauf auch noch die eben zitierte Doppelfrage 28/29 beantwortet werden soll.

Auch bezüglich der Anforderungen, die an Vpn gestellt werden können, sollte bei zukünftigen Untersuchungen besser ein niedrigeres Niveau angesetzt werden, um auf keinen Fall durch eine Überforderung der Vpn die Versuchsergebnisse zu verfälschen.

Was die offenen Fragen des Fragebogens und ihre Auswertung betrifft, müssen erhebliche Bedenken gegenüber dem Freiburger Versuch angemeldet werden. Bauer et al. schreiben (S. 36): "Die offenen Fragen des zweiten Teils (Assoziation/Interpretation) wurden beibehalten, da sie für eine linguistische Analyse unentbehrlich sind." Unabhängig davon, daß natürlich auch die Linguistik (und nicht nur die Literaturwissenschaft) mit empirischen Methoden arbeiten kann, bleibt folgender Tatbestand festzuhalten: Antworten auf "offene Fragen" bieten zwar dem Untersucher mitunter reichhaltigere Informationen über den Untersuchungsgegenstand als Antworten auf geschlossene Fragestellungen, sie sind aber auf der anderen Seite nur äußerst schwierig zu gewichten, worunter ihre statistische Auswertung erheblich leidet, wenn nicht gar unmöglich gemacht wird. Es gibt bisher keine Textgrammatik mit einem entsprechend leistungsfähigen Lexikon, die eine explizite Darstellung des Parameters erlaubt. Der Bezugsrahmen für offene

Fragen nach der denotativen und/oder der konnotativen Bedeutung eines Textformulars¹ oder einer seiner Bestandteile (Metaphern z.B.) ist, selbst in einer "groben" Vermessung, zur Zeit noch nicht zu erstellen. Linguistische Analysen, wie sie den Freiburgern vorgeschwebt haben mögen, können aber den Bezugsrahmen eines Textformulars nicht ersetzen. Insofern hängt die Auswertung offener Fragen bisher noch generell im "luftleeren Raum". Hier klafft eine Lücke zwischen Theorie und empirischer Praxis, die vielleicht erst in einigen Jahren geschlossen werden kann².

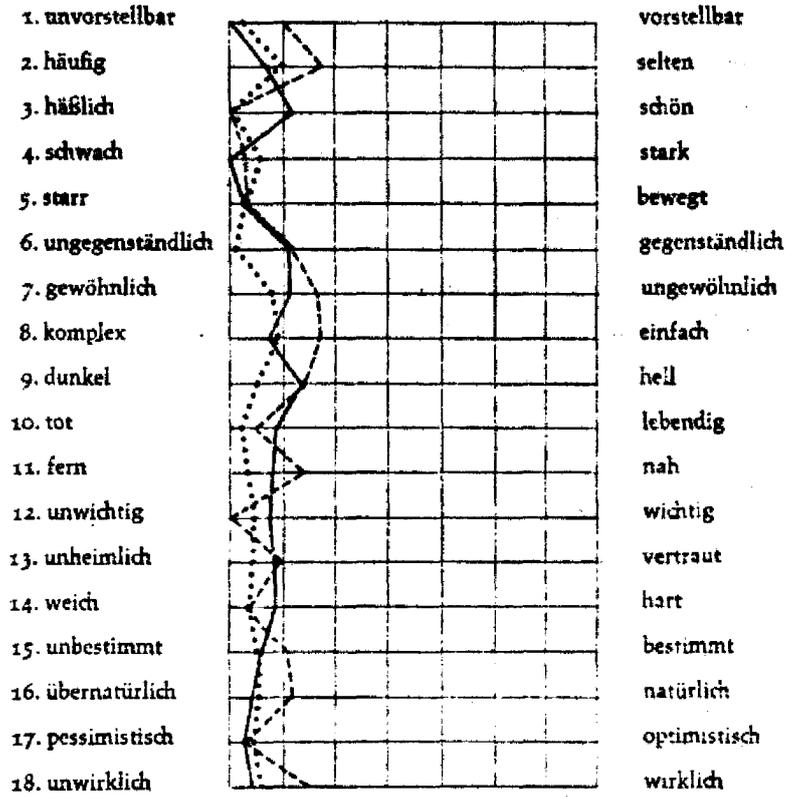
Im letzten Abschnitt soll nun die "Objektivität" der Auswertung der erhobenen Daten bei Bauer et al. erörtert werden, und zwar symptomatisch am Beispiel der Daten, die mit Hilfe des SDs gewonnen wurden. Dazu schauen wir uns zunächst die auf S. 81 abgebildete Tabelle der Distanzen-Mittelwerte an:

¹ Zu diesem Ausdruck cf. S.J. Schmidt (1974)

² Vgl. die weitere Diskussion der Problematik "offener Fragen" bei Zobel (1975) S. 128 ff.

DISTANZEN - MITTELWERTE

(0) (1) (2) (3) (4) (5) (6) (7)



----- Fadensonnen ——— Lichtton baumhoher Gedanke

Im Kapitel 1.5 wird von Bauer et al. davon gesprochen, daß die Bedeutungen (Bedeutungen wieder nur verstanden als emotive Konnotationen) der drei Metaphern hinsichtlich ihrer Ähnlichkeit bzw. ihrer Distanz analysiert werden sollen¹. Ähnlichkeit und Distanz können nun mathematisch-statistisch durch einen Korrelationskoeffizienten ausgedrückt werden, der zwischen + 1. (im Falle absoluter Gleichheit) und - 1. (im Falle äußerster Gegensätzlichkeit) variieren kann. Diese statistische Möglichkeit, Ähnlichkeit bzw. Distanz zu beschreiben, scheint in dem Satz gemeint zu sein: "Zusätzlich wurde ein quantitativer Vergleich des Ähnlichkeitsgrades der Assoziationsmuster vorgenommen." Welcher Korrelationskoeffizient errechnet wurde, wird nicht angegeben. Der von Hofstätter eingeführte Korrelationskoeffizient Q_{ab} z.B., der Distanzen zwischen den Gesamtdaten zweier SDs beschreibt, wäre wegen seiner allzu großen Ungenauigkeit nicht sehr gut geeignet².

Zu einer von Bauer et al. sicher nicht beabsichtigten Verwirrung des Sinnes der zitierten Passage trägt in diesem Zusammenhang die Abbildung der 'Distanzen - Mittelwerte' bei. Was wurde hier aufgetragen in dem SD? Was bedeutet hier die Skalierung zwischen 0 und 7?

¹ Diese zugegebenermaßen schon 'freie Übertragung' des Originaltextes bei Bauer erscheint klarer, weil ja nicht die Konzepte verglichen werden (bei Osgood: verbal concepts), sondern die Bedeutungseinschätzungen dieser 'verbal concepts', die die Vpn mit dem Ausfüllen der SDs, dem Ankreuzen auf den jeweils 18 Adjektivskalen vorgenommen haben.

² Vgl. dazu die Diskussion bei Hofstätter (1959), Kentler (1959), aber auch Ertel (1965 a), Ertel (1965 b) und anderen.

Noch eine weitere Frage drängt sich bei der Betrachtung des Diagramms auf: Handelt es sich bei den so aufgetragenen Mittelwerten der Distanzen vielleicht um ein Versehen? Müßten diese Werte nicht evtl. alle um Position 4 schwanken, anstatt zwischen den Positionen 0 und 2? In der veröffentlichten Form kann man das nicht entscheiden.

Darüberhinaus sagen Mittelwerte nur ganz bedingt etwas über eine bestimmte Datenverteilung aus, wenn man nicht die dazugehörige Streuung kennt. Dazu ein fiktives Beispiel: Wenn 600 Vpn für das Konzept "Fadensonnen" auf der Skala unvorstellbar/vorstellbar alle die Position 4 (weder - noch, neutral, weiß ich nicht) angekreuzt hätten, so würde der Mittelwert $\bar{X} = 4$ sein. Wenn von diesen Vpn beim Konzept "baumhoher Gedanke" auf derselben Skala 300 Vpn die Position 1 (absolut unvorstellbar) und die anderen 300 Vpn die Position 7 (absolut vorstellbar) angekreuzt hätten, so wäre der Mittelwert \bar{X}' ebenfalls = 4. Wer nun ohne weitere Informationen die beiden Mittelwerte \bar{X} und \bar{X}' miteinander vergleichen würde, der könnte meinen, die Konzepte "Fadensonnen" und "baumhoher Gedanke" seien in ihrer Bedeutungseinschätzung auf der Skala vorstellbar/unvorstellbar völlig identisch von den Vpn eingestuft worden, ein Schluß, den die Daten in unserem Beispiel widerlegen. Wären nun die Streuungen zu \bar{X} und \bar{X}' bekannt, nämlich $s = 0$ und $s' = 3$, dann hätte man die Mittelwerte sachgerecht beurteilen können.

In der veröffentlichten Form kann man das nicht entscheiden, daher: Bei einer empirischen Untersuchung muß explizit angegeben werden, was man wie aus welchen Daten errechnet hat. Die Berechnungen und die Daten müssen so umfangreich wie möglich mitveröffentlicht werden, damit andere Forscher die Auswertungsobjektivität überprüfen können.

Die gesamten Ausführungen in 1.5.2ff. sind aber nicht nur wegen des Fehlens einer Überprüfungsmöglichkeit mit einem großen Fragezeichen zu versehen. Auch der ausschließlich paarweise Vergleich von Mittelwerten mehrerer Stichproben zur Feststellung von statistisch signifikanten Unterschieden ist nicht korrekt.

- (a) Mehrere am gleichen Material durchgeführte Mittelwertvergleiche (t -Tests) sind nicht unabhängig voneinander. Wenn die erste Stichprobe mit der zweiten und der dritten verglichen wird, so gehen die Daten der ersten eben zweimal in die Berechnungen ein und der erste und der zweite t -Test haben zu 50% die gleichen Ausgangsdaten. Nimmt man an, man wollte die Mittelwerte von 10 Stichproben vergleichen, dann müßten insgesamt 45 einzelne t -Tests durchgeführt werden. Wenn aber so viele Tests durchgeführt werden, so ist schon aufgrund der Wahrscheinlichkeitsüberlegungen anzunehmen, daß irgendwo ein Unterschied signifikant wird, selbst dann, wenn alle 10 Stichproben aus derselben Population stammen. (Wer sehr oft an einem Glücksspiel teilnimmt, indem er viele Lose erwirbt, dessen Chance zu gewinnen steigt - auch wenn das Glücksspiel nach strengen Zufallsverfahren durchgeführt wird. Ebenso steigt bei mehreren t -Tests am gleichen Material die Chance für ein signifikantes Ergebnis.)
- (b) Weiter ist zu berücksichtigen, daß beim paarweisen Vergleich durch t -Tests die tatsächlich vorliegende Zahl der Beobachtungen nicht gut ausgenutzt würde. Wenn man fünf Stichproben zu je 20 Beobachtungen hat und nur jeweils zwei Stichproben mit einem t -Test vergleicht, so beträgt die Zahl der Freiheitsgrade immer $2 \times 20 - 2 = 38$ für jeden t -Test. In Wirklichkeit hat man aber insgesamt 100 Beobachtungen. Man würde also auf einen großen

Teil von Informationen verzichten, wenn man immer die restlichen 60 Fälle unberücksichtigt ließe. Dies erniedrigt die Chance eines signifikanten Ergebnisses. (Aber nicht im Sinne eines Ausgleichs zwischen den Erhöhungen von (a) und den niedrigeren Chancen aus (b)).

- (c) Diese Nachteile vermeidet man, wenn man eine Varianzanalyse durchführt. Mit dem Namen "Varianzanalyse" faßt man eine Gruppe von Prüfverfahren zusammen. Auf den ersten Blick ist ihnen allen gemeinsam, daß sie mehr als zwei Mittelwertsunterschiede überprüfen. Danach ist eine Varianzanalyse ein statistischer Test, der prüft, ob Mittelwerte mehrerer Stichproben aus ein- und derselben Population stammen¹. Ein Vorteil der Varianzanalyse liegt auch schon in ihrer Fragestellung. Es ist nämlich günstiger, zunächst zu prüfen, ob überhaupt zwischen irgendwelchen Mittelwerten Unterschiede bestehen und erst in einem zweiten Schritt ins Detail zu gehen und nachzusehen, wo sie möglicherweise liegen. Das Ganze rangiert hier vor dem Detail. Die möglicherweise existierenden Detailunterschiede werden erst nach einer Varianzanalyse in Nachtests untersucht.

Soviel zu den ungeklärten Fragen, die sich bei Bauer et al. aus einem nicht näher bezeichneten quantitativen Vergleich des Ähnlichkeitsgrades der Assoziationsmuster ergeben.

¹Zu diesem Begriff vgl. Ebner/Clauss (1975) S. 309ff., 327 ff. und S. 339ff. Die hier gegebene Definition ist allerdings zu eng, denn man kann varianzanalytisch auch prüfen, ob eine Regressionsbeziehung linear ist, und dies hat primär nicht unbedingt etwas mit Mittelwertsunterschieden zu tun.

Auf eine weitere Nuance des Textes muß aber noch eingegangen werden. Es heißt da: "Zusätzlich wurde ein quantitativer Vergleich des Ähnlichkeitsgrades der drei Assoziationsmuster vorgenommen." Zusätzlich! Bauer et al. haben also noch auf eine andere Art die Ähnlichkeit bzw. Distanz analysiert, ohne zu sagen, auf welche. Vielleicht vorwissenschaftlich, durch einen Sicht-Vergleich der aufgezeichneten Polaritätsprofile aller drei Metaphern? Oder durch einen Vergleich der Bedeutungszuordnungen zu allen drei Metaphern innerhalb einer Adjektivskala? "Die Analyse", so schreiben Bauer et al. "zeigt die diskriminierenden Elemente des Assoziationsmusters eines Konzepts auf und damit jene Assoziationen, in denen sich die Konzepte voneinander unterscheiden." "Wie unterscheiden sich die Konzepte? Statistisch signifikant? Oder nach vorwissenschaftlicher Betrachtung der Polaritätsprofile? Oder auf noch andere Art? Das sollte man schon genau wissen, wenn man daraus abgeleitete Kernelemente des Assoziationsmusters einer Metapher" verstehen und akzeptieren soll.

Kommen wir zur Objektivität der Auswertung jener Daten, die einer Faktorenanalyse unterzogen wurden. Auch hier sind, wie sich zeigen wird, erhebliche Mängel zu beanstanden.

Auf S. 144 schreiben Bauer et al.: "Methodische Probleme des Semantischen Differentials wurden in der Darstellung ausgeklammert, um den Text für den Leser zu vereinfachen." Obwohl man methodische Probleme des SDs ausgeklammert hat, ist die angestrebte Vereinfachung des Textes nicht geglückt. Vor allem besteht auch hier fast keine Überprüfungsmöglichkeit der Ausführungen von Bauer et al. Greifen wir ein wichtiges Beispiel dafür vorweg heraus. Bauer et al. beschreiben die Bedeutung, die die Faktorladungen einer Skala für den jeweiligen Faktor haben:

"Wesentlich ist, daß die Faktorenanalyse aber nicht nur Skalen entsprechend ihrer Abhängigkeit (Interkorrelation) ordnet, sondern daß auch der Abhängigkeitsgrad von Skala und Faktor quantitativ in der Faktorladung erfaßt wird. Die Faktorladung einer Skala ist ein Maß für die Nähe oder Distanz zu einem Faktor. Je höher die positive oder negative Faktorladung einer Assoziation ist, umso größer ist die Bedeutung für den Faktor ... die Interpretation eines Faktors orientiert sich an den Skalen mit hoher Faktorenladung."

Leider sind diese von Bauer et al. selbst als wichtig charakterisierten Angaben über die Faktorenladungen der einzelnen Variablen weder im Haupttext noch im Anhang mitveröffentlicht.

Auch die wichtige Darstellung der Varianzanteile eines Faktors vermißt man. Eine Überprüfungsmöglichkeit der Faktoreninterpretation ist also nicht gegeben. Diese wäre aber dringend erforderlich gewesen, um z.B. das Problem 'Mischfaktor' wissenschaftlich zu erörtern. Erst mit der Datenveröffentlichung hätte man entscheiden können, ob der Verdacht begründet ist, Bauer et al. hätten Osgood nicht richtig verstanden und deshalb auch ihre eigenen Ergebnisse ungenau interpretiert. Sie schreiben (S. 61): "Bei der Anwendung des Semantischen Differentials innerhalb des Fragebogens wurde gleicherweise von den in der Voruntersuchung gegebenen Antworten der Vpn als (!) auch von den von Osgood postulierten (Hervorhebung der Verf.) drei Faktoren des Semantischen Raumes ausgegangen." Osgood hat, soviel scheint sicher, nie "Faktoren postuliert"; er hat in dem von ihm und seinen Mitarbeitern erhobenen Datenmaterial, das sich hauptsächlich auf die Bedeutungsräume von Wörtern beschränkte, immer wieder drei Hauptfaktoren gefunden, den evaluative factor, den potency-factor und den activity-factor; daraus abzuleiten, er postuliere diese drei Faktoren (Dimensionen), erscheint zumindest mißverständlich. Ebenso scheint Vorsicht

geboten bei den beiden folgenden Passagen (S. 61):

"Die Anzahl der vergleichbaren Adjektivskalen beschränkt sich auf 18. Bei ihrer Zusammensetzung **m ü s s e n** (Hervorhebung der Verf.) die hypothetisch angenommenen Dimensionen innerhalb des multidimensionalen semantischen Raumes repräsentiert sein. Sie sind aus den in der Voruntersuchung hervortretenden alternativen verbalen Reaktionen der Vpn direkt ableitbar."

"Durch die Übersetzung aus dem englischen Skalensmaterial Osgoods wurden Adjektive gefunden, die seine mit den Begriffen <<Evaluation>> (Bewertung), <<Potency>> (Kraft) und <<Activity>> (Aktivität) gekennzeichneten Dimensionen repräsentieren. Dieses **G r u n d s c h e m a** (Hervorhebung der Verf.) semantischer Dimensionen wurde ergänzt durch die zwei neuen Dimensionen, <<Realität>> und <<Exzeptionalität>>. Diese Dimensionen, die für die hier gegebenen Konzepte spezifisch sind, wurden auf der Grundlage der Voruntersuchung gewonnen."

Das Problem konzeptspezifischer Adjektivskalen, das sich Bauer et al., aber auch Osgood stellte, darf inzwischen als nicht mehr aktuell bezeichnet werden. Suitbert Ertel widerlegt die Auffassung Osgoods, es müsse für jedes Objekt ein eigenes SD entwickelt werden, bei seinem Versuch ein SD zu standardisieren¹. Auf die interessante, aber sehr umfangreiche Beweisführung Ertels kann im Rahmen dieser Kritik leider nicht eingegangen werden².

Die Problematisierung der anderen Zitate (der S. 61) kann nach den schon gemachten Hinweisen über empirische Hypothesenbildung dem Leser überlassen werden.

¹Vgl. Ertel (1965a), (1965b)

²Vgl. Ertel (1965a)

Was kann man tun, um sich dennoch einige Klarheit über die Faktorenzusammenhänge zu verschaffen, die Bauer et al. errechnet haben? Man kann zunächst versuchen, die Faktorstruktur, in deren Rahmen die Bedeutungszuordnung für die drei Metaphern erfolgt, aus den bei Bauer et al. über mehrere Seiten verstreuten Angaben zu rekonstruieren.

Offensichtlich werden zunächst immer 3 Faktoren extrahiert (vgl. S. 143, 147, 149) - nach welchen Kriterien die Bestimmung von nur 3 zu extrahierenden Faktoren erfolgt, wird nicht erwähnt. Danach wird die Konsistenz der 3 Faktoren überprüft durch Berechnung von 4- bzw. 5-Faktor-Lösungen (vgl. S. 147, 149, 151).

Die aus den von Bauer et al. nur lückenhaft übermittelten Daten von uns rekonstruierte Faktorstruktur haben wir in der Skizze 4 (s. Seite 7) modellartig abgebildet. Für die Interpretation der so rekonstruierten Ergebnisse der Faktorenanalyse bei Bauer et al. bieten sich folgende Richtlinien an:

- (1) Die Faktoren werden zunächst als abstrakte Größen eingeführt.
- (2) Zur Kennzeichnung der Faktoren dürfen nur die Variablen (Adjektivskalen) zugelassen werden, die nicht bipolar auf mehrere Faktoren laden.
- (3) Solche Variablen, die nicht in allen Faktorenlösungen auf denselben Faktor laden, werden nur bedingt zur Interpretation herangezogen.
- (4) Die Gewichtigkeit der einzelnen Faktoren (auch untereinander) wird nicht interpretiert, weil nicht bekannt ist, welche Anteile an absoluter und an rotierter Varianz sie repräsentieren.

(5) Die 4- und 5-Faktor-Lösungen werden in die Interpretation mit einbezogen.

Nach diesen Richtlinien können wir anhand der Rekonstruktion folgende Befunde interpretieren:

In den 3-Faktor-Lösungen aller drei Faktorenanalysen (I, II, III) wird ein Faktor A von den Variablen:

tot - lebendig
starr - bewegt
pessimistisch - optimistisch
dunkel - hell

geladen. Auf einen Faktor B laden die Variablen:

unbestimmt - bestimmt
fern - nah
unvorstellbar - vorstellbar
ungegenständlich - gegenständlich
übernatürlich - natürlich

Auf einen Faktor C laden die Variablen:

häufig - selten
gewöhnlich - ungewöhnlich
schwach - stark
weich - hart

In höheren Faktorlösungen zeigen sich diese Faktoren als nur sehr bedingt stabil (im Gegensatz zu den Behauptungen von Bauer et al.). Die Zusammensetzung der Variablen, die auf Faktor B laden, zerfällt in allen drei Faktorenanalysen.

	I (Lichtton)		II (Fadensonnen)		III (bhG.)
Faktor A	tot	- lebendig	dto.		dto.
	starr	- bewegt	dto.		dto.
	pessimistisch	- optimistisch	dto.		dto.
	dunkel	- hell	dto.		dto.
	häßlich	- schön	dto. (bipolar)		dto.
	unheimlich	- vertraut (bipolar)	-		dto.
	unwichtig	- wichtig	-		dto.
		-	-		schwach - stark (bipolar)
		-	-		wirklich - unwirklich (")
		-	-		
Faktor B	bestimmt	- unbestimmt	dto.		dto.
		-	-		komplex - einfach
	fern	- nah	dto.		dto.
	unvorstellbar	- vorstellbar	dto.		dto.
	gegenständlich	- ungegenständlich	dto.		dto.
	natürlich	- übernatürlich	dto.		dto.
	wirklich	- unwirklich	dto.		dto. (bipolar)
	vertraut	- unheimlich (bipolar)	dto.		dto.
		-	unwichtig - wichtig		-
		-	häßlich - schön (bipolar)		-
Faktor C	häufig	- selten	dto.		dto.
	gewöhnlich	- ungewöhnlich	dto.		dto.
	schwach	- stark	dto.		dto.
	weich	- hart	dto.		dto.

(4-Faktorlösung: D = ein Teil des gespaltenen Faktors B D = pess.-optim. häßlich-schön dunkel-hell
 (C₃ + C₄ schwächen sich in der Ladung auf Faktor C ab) D = wichtig - unwichtig (C₃ + C₄ schwächen sich in der Ladung auf Faktor C ab)

(5-Faktorlösung: E = C₃ + C₄ E (Spaltung von Faktor B wie in I) E = "scheint ein Bewertungsfaktor zu sein" (Faktor B verliert einige Skalenladungen)

Zweimal bildet sich ein eigener Faktor D bzw. E heraus (in I und II). Näheres kann dazu nicht gesagt werden, weil Bauer et al. nicht mitteilen, welche Variablen auf D (I) und E (II) laden und welche Variablen in III nicht mehr auf B laden.

Die Zusammensetzung des Faktors A scheint ebenso instabil, denn in II wird von den Variablen

pessimistisch - optimistisch
dunkel - hell
häßlich - schön

ein eigener Faktor D geladen.

Die Zusammensetzung des Faktors C ist ähnlich instabil wie die von A und B. In I bilden die Variablen

schwach - stark
weich - hart

in der 4-Faktor-Lösung einen eigenen Faktor D, in II und III bröckeln die Ladungen dieser beiden Variablen auf C stark ab, wenn sie auch nicht ganz zerfallen.

Wie sollten die Faktoren nun benannt werden, die mit Hilfe der durchgeführten Faktorenanalysen ermittelt wurden? Für die Benennung einzelner Faktoren gibt es keine verbindlichen Richtlinien. Es bleibt mehr oder weniger dem Geschick des Untersuchers überlassen, eine treffende Benennung der Faktoren vorzunehmen; welchen Namen man einem Faktor geben sollte, hängt aber u.a. auch von der Höhe der einzelnen Faktorladungen ab, die im vorliegenden Fall nicht bekannt sind. Die Faktoren A, B und C - soweit kann man Bauer zustimmen - scheinen tatsächlich Mischfaktoren zu sein.

Faktor A: Ein Aktivitätsfaktor, der auch schwächere Ladungen von Bewertungsvariablen aufweist. (Diese Bewertungsvariablen laden in höheren Faktorlösungen einen eigenen Faktor)

Faktor B: Kann nicht eindeutig bezeichnet werden, da die präzisen Angaben über höhere Faktorlösungen fehlen. Aber eine Bezeichnung nach einer Realitäts-/Irrealitätskomponente erscheint problematisch, zumal andere Faktorenanalysen, denen Datenmaterialien über Textbedeutungen zugrundelagen, die ebenfalls mit Hilfe von SDs gewonnen wurden (cf. Zobel, 1975), einen Faktor auswiesen, der von den Variablen

einfach - kompliziert
bestimmt - unbestimmt
deutlich - verschwommen
gewöhnlich - ausgefallen

geladen wurde, Faktor B also sehr ähnlich war, ohne daß eine Realitäts-/Irrealitätsvariable auf ihn lud.

In diesem Faktorbereich bieten sich sinnvolle Überprüfungen der Faktorenhypothesen an.

Faktor C: Exzeptionalitäts-/Potenzfaktor. (In höheren Faktorlösungen laden die Potenzvariablen entweder einen eigenen Faktor oder ihre Ladungen auf Faktor C werden deutlich schwächer).

An dieser Stelle sei noch einmal auf die kritisierte These von Bauer et al. aus Abschnitt 1 verwiesen. Die Faktorenanalyse hat es ermöglicht, einen hypothetischen Faktor zu ermitteln, der von den Variablen:

häufig - selten
und gewöhnlich - ungewöhnlich

geladen wird. Die Hypothese, daß ein solcher Faktor existiert, den man mit Bauer et al. "Exzeptionalitätsfaktor" nennen könnte, muß durch weitere Versuche überprüft werden.

Nach dieser Interpretation der rekonstruierten Daten können wir auf die gestellten Fragen aus Kapitel 1. zurückkommen und sie vorläufig beantworten:

- (1) Die behauptete Eigenschaft E kann - wenn man nicht wie Bauer et al. 'Wirkung' im Sinne von Einwirken, sondern präzise im Sinne von 'konnotative Bedeutung zuordnen' versteht - durch Messung überprüft werden.
- (2) Man weiß nichts darüber, ob die Eigenschaft E nur im Zusammenhang mit der Besonderheit der Metaphern m_1 , m_2 und m_3 des Gedichts g vorkommt. Man darf aber vermuten - und das kann durch weitere Untersuchungen mit demselben Meßinstrument überprüft werden -, daß es eine Eigenschaft gibt, die nicht nur kühnen Metaphern zukommt, sondern jeder Metapher, jedem Gedicht, vielleicht jedem Text oder sogar jedem Objekt.

Erst jetzt nach den vorangegangenen Rekonstruktionen und der damit verbundenen besseren Übersicht über die generierten Faktorenhypthesen kann man die folgende Passage versuchen zu erschließen:

"Dimension: Aktivität - Bewertung

Das größte Gewicht kommt dabei der Dimension Aktivität-Bewertung zu.

Die Dimension der Aktivität wird durch die beiden Polaritäten tot - lebendig und starr - bewegt definiert, denen führende Funktion zukommt. Beide Skalen stehen in einem engeren Bezug zu der Dimension Aktivität-Bewertung als die Bewertungskomponenten. Während die Aktivitäts-Polaritäten einen größeren Einfluß auf die Charakterisierung der Dimension haben, treten die Bewertungspole insgesamt in ihrer Ausprägung etwas in den Hintergrund, sind aber in differenzierterer Weise vertreten. Die in dieser Dimension auftretenden Bewertungsskalen sind: pessimistisch - optimistisch, häßlich - schön und unheimlich - vertraut. Zu diesen drei wertenden Beschreibungskategorien treten die Skalen unwichtig - wichtig und dunkel - hell hinzu. Während beide Skalen ursprünglich anderen Bereichen zugeordnet waren, war in der Faktorenanalyse mehrmals zu beobachten, daß beide Polaritäten mit den Bewertungsskalen gemeinsam auftraten. Aus dieser Sachlage ist zu folgern, daß die beiden Skalen unwichtig - wichtig und dunkel - hell, sobald sie im Verband mit Bewertungsskalen auftreten, als Deskriptions- und Bewertungskomponenten anzusehen sind.

Innerhalb der für das Konzept Lichtton bedeutendsten Dimension, die eine Mischung aus Aktivitäts- und Bewertungsskalen darstellt, liegt der Akzent auf dem Aktivitäts-Dynamik-Aspekt, während die Deskriptions-Bewertungsskalen in ihrer Bedeutung zurücktreten. Aus der Mischung der Aktivitäts- und Bewertungsbereiche ist zu schließen, daß mit den Aktivitäts-Komponenten im Erlebnishorizont des Konzepts Lichtton Wertungen eng gekoppelt sind." (S. 145)

Das "größte Gewicht" heißt vermutlich: "der Faktor, der den Hauptteil an totaler Varianz erklärt, ist Faktor A". Wir übersetzen weiter: 'Die Variablen, die am stärksten auf ihn laden, sind: tot - lebendig und starr - bewegt. Alle Variablen aus dem Wertungsbereich haben dagegen nur geringere Ladungen. Trotz unserer Umformulierung bleibt aber der Tatbestand festzuhalten: Bei diesem Satz handelt es sich um eine überprüfbarere Behauptung. Daß die Bewertungsvariablen etwas in den Hintergrund treten, a b e r (dafür?) in differenzierterer Weise

vertreten sind, verwirrt wieder den Sinnzusammenhang, denn wenn zwei oder drei weitere Variablen aus dem Aktivitätsbereich, z.B. aktiv/passiv, gespannt/gelöst, lebhaft/träge, im benutzten SD aufgenommen worden wären, dann wären die Bewertungsvariablen nicht in "differenzierterer" Weise vertreten, so steht jedenfalls zu vermuten. Nur wenn diese Vermutung falsch wäre, hätten die diesbezüglichen Ausführungen von Bauer et al. einen Sinn.

Abschließend muß noch auf eine weitere Schwäche der Arbeit von Bauer et al. hingewiesen werden. Im folgenden Absatz sollen die wechselnden Interpretationen der Skala dunkel/hell im Kapitel 5 verfolgt werden, um zu zeigen, daß hier offensichtlich die Auswerter der faktorenanalytischen Befunde große Zuordnungsschwierigkeiten hatten.

Zunächst einmal erscheint es problematisch, die Bewertungsskalen im SD, das ja nur emotive Konnotationen erheben kann, als bewertende "Beschreibungsskalen" zu charakterisieren. Das Wort 'beschreiben' drückt auch umgangssprachlich einen so hohen Anteil denotativer Bedeutung eines Objekts aus, daß seine Verwendung in diesem Kontext fragwürdig bleiben muß. Fragwürdig bleibt auch die schon besprochene Postulierung von Faktoren vor Durchführung einer empirischen Untersuchung mit dem SD. Bauer et al. p o s t u l i e r t e n (im Gegensatz zu Osgood et al.) die 5 Faktoren: Evaluation, Potency, Activity, Irreality, Exceptionality und ordneten diesen Faktoren entsprechende Adjektivskalen zu, u.a. die Skala hell/dunkel dem Potenzialfaktor, die Skala wichtig/unwichtig dem Exzeptionalitätsfaktor. Das führte dann zu Schwierigkeiten wie der folgenden:

"Während beide Skalen ursprünglich anderen Bereichen zugeordnet waren, war in der Faktorenanalyse mehrmals zu beobachten, daß beide Polaritäten mit den Bewertungsskalen gemeinsam auftraten. Aus dieser Sachlage ist zu folgern, daß die beiden Skalen unwichtig/wichtig und dunkel/hell, sobald sie im Verband mit Bewertungsskalen auftreten, als Deskriptions- und Bewertungskomponenten anzusehen sind." (S. 145)

Weil nicht gefunden wurde, was man vorher angenommen hatte, sind die beiden Variablen nur "sobald sie im Verband mit Bewertungsskalen auftreten" als Bewertungsskalen interpretierbar. Wenige Absätze später liest man:

"Die Aktivitäts-Assoziationen sind vertreten durch die Skalen tot - lebendig, starr - bewegt und dunkel - hell. Gekoppelt mit diesem dynamischen Aspekt des Konzepts Fadensonnen ist eine Bewertung, die durch die Skalen häßlich - schön und pessimistisch - optimistisch repräsentiert wird." (S. 148)

Diese Zuordnung von hell/dunkel zu einem "dynamischen Aspekt" im Rahmen des Aktivitätsfaktors verwirrt umso mehr, weil eine Seite später wieder eine andere Zuordnung getroffen wird:

"Gegenüber der Drei-Faktor-Lösung tritt ein neuer Faktor auf, der aus den Assoziationspaaren häßlich - schön und pessimistisch - optimistisch des Bewertungsbereichs und der Polarität dunkel - hell aus dem ursprünglichen Potenzbereich besteht. Auch aus anderen Faktoren-Lösungen muß aber geschlossen werden, daß die Skala dunkel - hell eine große Nähe zum Bewertungsbereich hat. Aus diesem Grunde ist dieser vierte Faktor als fast reiner Bewertungsfaktor zu interpretieren." (S. 149)

Diese Rückkehr zur ursprünglichen Interpretation auf S. 145 wird aber zwei Absätze später wieder aufgehoben, denn nun ist die Skala hell/dunkel doch wieder eine Potenzkonnotation, die fast hermeneutisch charakterisiert wird.

"... während die Potenzkonnotationen durch die Polaritäten schwach - stark und dunkel - hell vertreten werden. Während der Aktivitäts-Bereich den dynamischen Aspekt repräsentiert, wird in den Potenzkategorien eher versucht, die ruhende Kraft, die Kraft vor der Bewegungsentfaltung, zu erfassen."
(S. 149-150)

Doch diese Aussage über die Skala hell/dunkel ist auch noch nicht endgültig, denn Bauer et al. fahren im nächsten Satz fort:

"Das Assoziationspaar dunkel - hell zeigt in verschiedenen Faktorenanalysen die Tendenz, zusammen mit Bewertungs-Skalen aufzutreten, so daß angenommen werden kann, daß diese Skala bereits dem Bewertungsbereich zuzuordnen ist. Zumindest aber ist daraus zu schließen, daß die Polarität dunkel - hell in beide Assoziationsbereiche hineinreicht, wodurch der zweite große Einflußbereich in dieser Dimension noch verstärkt wird." (S. 150)

Eine solche schwankende Interpretation läßt auf eine erhebliche Unsicherheit der Interpreten gegenüber ihrem Material und ihren Methoden schließen.

Diese Interpretations-Unsicherheit von Bauer et al. zeigt, worauf Philologen achten müssen, wenn sie sich daran machen, empirisch gesammelte Daten zu interpretieren. Ein analoges Beispiel liefert der Mittelteil des Abschnitts 5.2.2 (S. 145-146), wo im klassischen hermeneutischen Stil frei über die mit dem SD erhobenen Daten paraphrasiert wird.

Und eine letzte Anmerkung: Viele Schlußfolgerungen in den zitierten Kapiteln müssen u.E. kritisch in Zweifel gezogen werden, so, um ein Beispiel stellvertretend für alle anderen herauszugreifen:

"Aus der Mischung der Aktivitäts- und Bewertungsbereiche ist zu schließen, daß mit den Aktivitätskomponenten im Erlebnishorizont des Konzepts Lichtton Wertungen eng gekoppelt sind." (S. 145)

Da diese "Mischung" in allen Bedeutungszuordnungen zu den drei Metaphern zu beobachten war und außerdem die Ladung der Variablen aus dem Bewertungsbereich für diesen Faktor A sehr niedrig waren, kann man alle möglichen anderen Schlüsse aus dem Datenmaterial ziehen (u.a. daß man zu wenig Faktoren extrahiert hat), nur nicht den, den Bauer et al. gezogen haben.

Die Kritikansätze sollen an dieser Stelle abgebrochen werden. Dennoch empfehlen wir, sich unter Berücksichtigung der gemachten Einwände intensiv mit 'Text und Rezeption' auseinanderzusetzen, sich vor allem neben den hier nur symptomatisch angesprochenen Stellen auch mit der Zusammenfassung (Bauer et al., S. 217ff.) zu beschäftigen, denn dieses wichtige Buch berichtet über einen ersten Schritt auf dem Wege der Emanzipation der Literaturwissenschaft vom Zwang, "Geisteswissenschaft" sein zu müssen.

7. Weiteres zu den Ergebnissen der Untersuchung

7.1 Bereits im Unterkapitel 4.2 war davon die Rede, daß Bauer et al. die Vpn nach den Antworten auf die Frage 14 ("Finden Sie das Gedicht nach der ersten Lektüre unverständlich?") in drei Gruppen nach folgenden Kategorien aufschlüsseln: Ja, mit Einschränkung, Nein. In einer letzten Bemerkung zu diesem Komplex wollen wir auf eine Ungenauigkeit hinweisen, die Bauer et al. bei dem Vergleich der Antworten auf Frage 15 und 27 unterläuft ("Welches Wort fällt Ihnen am meisten auf?", "Welcher Satz des Gedichtes erscheint Ihnen am wichtigsten? Wo liegt, Ihrer Meinung nach, der Schwerpunkt?"). Unter Bezugnahme auf die Tatsache, daß in allen drei Gruppen G1 bis G3 das Wort "Fadensonnen" am häufigsten oder am zweithäufig-

sten als "das auffallendste Wort" angekreuzt wird (vgl. S. 89), schreiben Bauer et al. auf S. 90:

"In Frage 27 wird die Schwerpunktsetzung auf Fadensonnen in allen Gruppen der Vpn zugunsten des letzten Satzes des Textes aufgegeben, was die Vermutung aufkommen läßt, ein progressiver Rezeptionsprozeß impliziere eine Verlagerung der inhaltlichen Schwerpunkte."

Der hier angestellte Vergleich zwischen den Antworten auf Frage 15 und Frage 27 ist aber aus zwei Gründen problematisch:

Erstens braucht der Antwort auf Frage 15 keine Schwerpunktsetzung im Sinne von "Wichtigkeit" zu entsprechen, Bauer et al. hätten also eher die Antworten auf Frage 16 zum Vergleich heranziehen oder im Falle einer Übereinstimmung der Antworten auf Frage 15 und 16 auf diesen Umstand hinweisen müssen. Die Ergebnisse von Frage 16 werden hier aber gar nicht erwähnt (merkwürdigerweise sind sie nicht einmal in der Gesamtauszählung (vgl. S. 71-85) veröffentlicht, ohne daß diese Unterlassung begründet wird).

Zweitens dürfen die Ergebnisse der beiden Fragen nicht ohne weiteres als unterschiedliche Schwerpunktsetzungen gewertet werden, weil sie einmal einem isolierten Wort und das andere Mal einem Satz gelten. Konkreter gesagt: Es ist denkbar oder sogar wahrscheinlich, daß die Vpn zum selben Zeitpunkt zwar "Fadensonnen" für das auffallendste Wort, aber den dritten für den wichtigsten Satz halten, obwohl die in dem Satz enthaltenen Worte als wenig auffallend gelten können.

In einer weiteren Analyse schlüsseln Bauer et al. die Vpn nach ihrer Antwort auf Frage 2 auf ("Lesen Sie auch zeitgenössische Gedichte?"). Aus dieser Analyse möchten wir zwei Passagen als besonders krasse Beispiele von Überinterpretationen des Untersuchungsmaterials zitieren. Die erste Passage

bezieht sich auf die Gruppe derjenigen Vpn, die angeben, daß sie nie zeitgenössische Gedichte lesen (notiert als Gruppe L₃).

"Höchst interessant dagegen ist, daß in L₃ von 56.4% der Vpn die Schulzeit als Zeitpunkt der letzten Lyrik-Lektüre angegeben wird. Offenbar haben wir es hier mit Vpn zu tun, deren negativer Erwartungshorizont im Hinblick auf zeitgenössische Lyrik in unfreiwilliger Lektüre und Textinterpretation während der Schulzeit begründet ist. Die in der Schulzeit gebildeten Vorurteile gegenüber diesen Typen literarischer Texte erklären die Weigerung der Vpn, sich nach Abschluß der Schulausbildung noch freiwillig mit solchen Texten zu beschäftigen. Erstaunlich ist dagegen, daß auf Frage 3: "Wären Sie dafür, daß in der Schule zeitgenössische Lyrik behandelt wird?" Auch in dieser Gruppe [...] von der Mehrzahl der Vpn (73.8% mit "ja" geantwortet wird." (S. 97)

Zu dieser Passage ist zunächst anzumerken, daß in ihr die Annahmen, die mit den Stichworten "unfreiwillige Lektüre", "Vorurteile", "die Weigerung, sich freiwillig zu beschäftigen" und "methodisch gescheiterte Konfrontation" zusammenhängen, durch nichts abgedeckt sind. Bauer et al. können sich offensichtlich nicht vorstellen, daß es gesellschaftliche Gruppen gibt, die aufgrund des gesellschaftlichen Kontextes oder aufgrund von Zeitmangel zwar selbst kein Interesse an der Lektüre von zeitgenössischen Gedichten und wenig Berührungspunkte zu ihnen haben, diesen Gedichten aber trotzdem nicht prinzipiell ablehnend gegenüberstehen und mit ihnen auch nicht unbedingt negative Erfahrungen gemacht haben müssen. Sieht man sich des weiteren die Zusammensetzung von L₃ genauer an, dann braucht man sich nicht besonders darüber zu wundern, daß mehr als die Hälfte von diesen Vpn, die keine zeitgenössischen Gedichte lesen, zuletzt in der Schulzeit mit solchen Gedichten in Berührung gekommen

sind¹. L_3 setzt sich hauptsächlich aus Schülern (36.3%; vorwiegend aus der Mittelstufe), aus Gewerbeschülern (26.2%) und aus Studenten (34%) zusammen (vgl. S. 95); hinsichtlich der Verteilung der Studenten auf die verschiedenen Fachgruppen stellen Bauer et al. als "auffallend" fest, daß im Gegensatz zu den Nicht-Mathematikern, wo die zu L_3 gehörigen Studenten nur Minderheiten in der Fachgruppe bilden, 40% der Mathematiker zu L_3 gehören (weitere 50% der Mathematiker gaben übrigens an, nur gelegentlich zeitgenössische Gedichte zu lesen).

Was kann man aus diesen Zahlen ablesen? Vor einer endgültigen Bewertung der im obigen Zitat genannten 56.4% müßte als erstes geklärt werden, wieviele von den Schülern in L_3 auf Frage 3 mit "in der Schulzeit" geantwortet haben. Insofern als hiervon vorwiegend Schüler der Mittelstufe betroffen sind, ist es bei der Interessenlage solcher Schüler gut denkbar, daß sie sich hauptsächlich im Rahmen des Unterrichts mit zeitgenössischen Gedichten beschäftigen. Zweitens geht aus der Darstellung von Bauer et al. nicht eindeutig hervor, ob bei Frage 3 eine Mehrfach-Antwort zulässig ist, bzw. ob die Schüler nur jeweils eine der möglichen Alternativen angekreuzt haben. Falls beim Vorhandensein von Mehrfach-Antworten diese mit in die 56.4% eingegangen sein sollten, wäre das ein erheblicher Auswertungsfehler.

¹ Warum sieht Frage 3 eigentlich nicht die Möglichkeit vor, daß in der Schulzeit keine zeitgenössischen Gedichte gelesen worden sind? Bei einer Reihe von Vpn können jedenfalls aufgrund ihres Alters noch gar keine nach 1945 geschriebenen Gedichte im Unterricht behandelt worden sein. Für diejenigen von diesen Vpn, die außerdem der Gruppe L_3 angehören, gab es also bei Frage 3 keine Antwortmöglichkeit.

Sieht man einmal von den beiden eben besprochenen Unklarheiten ab, dann ist es bei Berücksichtigung der Ausbildungs- und Interessensituation von Schülern der Mittelstufe, Gewerbeschülern und literarisch desinteressierten Studenten (für die der Fall der Mathematikstudenten symptomatisch sein könnte) nicht besonders erstaunlich, daß ihr letzter Kontakt mit zeitgenössischen Gedichten in die Schulzeit fällt. Insbesondere muß diese Tatsache jedenfalls nicht derart spekulativ erklärt werden wie bei Bauer et al. und es muß schon gar nicht eine "methodisch gescheiterte Konfrontation mit zeitgenössischer Lyrik" angenommen werden.

In der zweiten Passage, die hier zitiert werden soll, geht es um einen Vergleich der Gruppe L₃ mit der Gruppe L₁, derjenigen Vpn, die angeben, daß sie häufig zeitgenössische Gedichte lesen. Dieser Vergleich wird im Hinblick auf die Frage 5 durchgeführt ("Was sind, Ihrer Meinung nach, die wichtigsten Merkmale zeitgenössischer Lyrik?") und bezieht sich auf 7 von 13 in Frage 5 vorgeschlagenen Merkmalen.

	L ₁	L ₃
unverständlich	13 0/0	26,2 0/0
befremdend	31,7 0/0	32,8 0/0
neue Wortbildungen	55,6 0/0	32,8 0/0
Sprachspielerei	33,3 0/0	36,1 0/0
Sprachmagie	41,5 0/0	14,8 0/0
faszinierend	30,9 0/0	9,8 0/0
nur Spezialisten zugänglich	6,5 0/0	16,4 0/0

Die sprachliche Exzeptionalität spiegelt sich in der Erfahrung beider Gruppen wider, jedoch unter verschiedenen Vorzeichen. In der Gruppe, die häufig zeitgenössische Lyrik liest (L₁), erfolgt diese Beurteilung sprachlich sehr bewußt und mit einer gewissen Sicherheit, die auf einer breiteren Textkenntnis basiert. Die generelle Wertung ist, wie aus dem stark beantworteten Attribut »faszinierend« hervorgeht, durchaus positiv ausgerichtet.

In Gruppe L₃ dagegen scheint diese sprachliche Exzeptionalität anders konnotiert zu sein, da sie mit der Unverständlichkeit des Textes korreliert und gleichzeitig eine Art Inferioritätsbewußtsein der Leser zum Vorschein bringt. Die Situation dieser Leser ließe sich daher so umreißen:

Der Leser hat durch einen ersten Kontakt mit zeitgenössischer Lyrik eine Art Schock erhalten; er fühlt sich solchen Texten gegenüber aufgrund ihrer Unverständlichkeit inkompetent und überläßt sie in Zukunft den »Spezialisten«. Dementsprechend stimmen in dieser Gruppe auch 46,2 0/0 uneingeschränkt die These zu, es sei auch für den aufgeschlossenen Leser praktisch unmöglich, moderne Gedichte zu verstehen.

Behauptungen wie die von der "sprachlich sehr bewußten Beurteilung in Gruppe L₁", von dem "Inferioritätsbewußtsein der Leser in Gruppe L₃" oder von dem angeblichen "Schock" lassen sich aus der angegebenen Tabelle nicht ableiten. Die Willkür in der Interpretation der Tabelle äußert sich z.B. auch darin, daß die mit dem höchsten Prozentsatz der Gruppe L₃ abgegebene Zustimmung zu "Sprachspielerei" nicht berücksichtigt wird. Angesichts dieser Auswertungspraktiken bleibt nur festzustellen, daß Bauer et al. die Reichweite und Aussagekraft ihrer Untersuchung erheblich überschätzt haben.

Die Hypothesen über die Gründe der Vpn, so und nicht anders als berichtet zu antworten, sind an vielen Stellen als gewagt zu bezeichnen.

56.4% der Vpn, die angeben, nie zeitgenössische Lyrik zu lesen, nennen als Zeitpunkt ihrer letzten Lektüre eines zeitgenössischen Gedichts ihre Schulzeit. Daraus schließen die Autoren:

"Offenbar haben wir es hier mit Vpn zu tun, deren negativer Erwartungshorizont im Hinblick auf zeitgenössische Lyrik in unfreiwilliger Lektüre und Textinterpretation während der Schulzeit begründet ist. Die in der Schulzeit gebildeten Vorurteile gegenüber diesem Typ literarischer Texte erklären die Weigerung der Vpn, sich nach Abschluß der Schulausbildung noch freiwillig mit solchen Texten zu beschäftigen." (S. 97)

Diese Erläuterungen mögen zwar vielen Lesern des Buches aufgrund eigener informeller Erfahrungen zutreffend erscheinen. Trotzdem sollte man von einer empirischen Untersuchung verlangen, daß, wenn überhaupt solche Hypothesen aufgestellt werden, diese durch zusätzliche Erhebungen gestützt werden. Abschließend soll noch ein besonders krasses Beispiel für eine

derartige kühne Ergebnisinterpretation gegeben werden. Im Kapitel III,7 werden die Resultate nach Interpretationsgruppen ausgewertet. Eine davon ist die Gruppe B, deren Interpretation auf eine Bildvorstellung hinauslaufen soll. Die Vpn dieser Gruppe weichen einer spontanen Bewertung des Celan-Textes am häufigsten aus. Dies versuchen die Autoren wie folgt zu begründen:

"Gruppe B weicht auf die unverbindliche Antwort 2 aus (kann ich nicht beurteilen), die ihre Hilflosigkeit dem Text gegenüber spiegelt, welche wesentlich in Verständnisschwierigkeiten begründet ist." (S. 183)

Warum hier eine Junktions zwischen Verständnisschwierigkeiten und Zurückhaltung bei der Bewertung hergestellt wird, bleibt unergründlich. Das müßte nämlich umgekehrt bedeuten, daß jemand, der meint, einen Text zu verstehen, eher zu einer spontanen Bewertung bereit ist. Das ist jedoch nicht notwendig so. Ebenso gut mag die Zurückhaltung gegenüber der Aufforderung, eine spontane Bewertung vorzunehmen, in der wohlbegründeten Motivation zu suchen sein, auch bewertende Hypothesen erst dann aufzustellen, wenn man sie für ausreichend abgesichert hält, unabhängig davon, ob man meint, den zu bewertenden Text zu verstehen oder nicht.

Eine genaue Bewertung der Analyse vorzunehmen, die sich mit den Antworten aus den Assoziationsfragen 18 - 22 beschäftigt (S. 115-142 und S. 155-166), ist u.a. deshalb schwierig, weil nicht alle Originaldaten veröffentlicht sind und weil z.B. nicht genau nachvollzogen werden kann, wie die Zuordnungen zu den Kategorien "affektiv n", "affektiv +" und "affektiv -" getroffen wurden (vgl. S. 119; die dort genannten Kriterien sind nicht ohne weiteres objektivierbar). Aufgrund dieser Schwierigkeiten soll hier auch gar nicht der Versuch

einer genauen Auseinandersetzung mit diesem Teil der Auswertung unternommen werden.

7.2 Ein weiterer genereller Kritikpunkt bezieht sich auf die Berechnung der χ^2 -Werte. Erstens reichen in vielen Fällen die veröffentlichten Daten nicht für eine Überprüfung der Rechnungen aus. Zweitens geht aus der Darstellung oft nicht eindeutig hervor, nach welchem Ansatz die Rechnungen vorgenommen wurden. Und drittens scheint bei den Berechnungen auch ein systematischer Fehler gemacht worden sein: An allen Stellen nämlich, wo wir die betreffenden χ^2 -Werte selbst nachgerechnet haben, ergaben sich andere Werte als bei Bauer et al. Hierzu wenigstens ein Beispiel:

Auf S. 85 werden die Vpn aufgrund ihrer Antworten auf Frage 14 ("Finden Sie das Gedicht nach der ersten Lektüre unverstündlich?") in 3 Gruppen eingeteilt. Die prozentuale Verteilung wird für die uns hier interessierenden Gruppen 1 und 3 mit folgenden Zahlen belegt:

Gruppe 1	Gruppe 3
35.5%	21%

Als Missing Data wird die Zahl 14 angegeben; folglich beträgt die Gesamtzahl der Antworten $665 - 14 = 651$ und es kann folgende absolute Verteilung ermittelt werden:

Gruppe 1	Gruppe 3
231	137

Auf S. 86 werden die beiden Gruppen in Hinblick darauf verglichen, zu welchem Anteil die zu Ihnen gehörenden Vpn Gedichte lesen (Frage 1), und dafür werden folgende Prozentzahlen angegeben:

Gruppe 1	Gruppe 3
9.7%	34.1%
22	47

In der letzten Zeile haben wir die zugehörigen absoluten Zahlen ermittelt¹.

Für den Unterschied zwischen Gruppe 1 und Gruppe 3 in Bezug auf das Merkmal "Gedichtlektüre" (im folgenden mit "G" abgekürzt) geben Bauer et al. 27,154 als χ^2 -Wert an (S. 86 Fußnote 6). Eine Berechnung dieses Wertes muß folgendermaßen vorgenommen werden².

Zunächst ergibt sich aus den bisherigen Zahlen, daß von den insgesamt 368 Vpn der Gruppen 1 und 3 69 Vpn Gedichte lesen. Damit erhält man folgende Erwartungswertverteilung für G:

	+ G	- G
Gruppe 1	43.3	187.7
Gruppe 3	25.7	111.3

¹ Da bei Bauer et al. keine Angaben über Missing Data bezüglich Frage 1 für die Gruppen 1 und 3 gemacht werden, sind die tatsächlichen absoluten Zahlen nicht zurückrechenbar. Der durch das Weiterrechnen mit den Zahlen 22 und 47 entstehende Fehler dürfte aber sehr klein sein.

² Vgl. etwa Clauss/Ebner(1975) S. 125

Die Abweichung zwischen erwarteten und beobachteten Werten beträgt absolut genommen jeweils $21,3$. Gemäß der allgemeinen Formel $\chi^2 = \frac{(o_i - e_i - 0,5)^2}{e_i}$ errechnet sich der ge-

suchte χ^2 -Wert also zu:

$$\begin{aligned}\chi^2 &= \frac{20,8^2}{43,3} + \frac{20,8^2}{187,7} + \frac{20,8^2}{25,7} + \frac{20,8^2}{111,3} \\ &= 33,01.\end{aligned}$$

Darüberhinaus erfahren die von Bauer et al. ermittelten χ^2 -Werte, legt man einmal ihre korrekte Berechnung zugrunde, mitunter zweifelhafte Interpretationen. Um ein Beispiel zu nennen: Auf S. 86 heißt es

"Auch die persönlichen Interessen weichen nicht signifikant* von den aus der Gesamtauszählung bekannten Schwerpunkten (Malerei, Musik, Sport) ab."

Die dazugehörige Fußnote gibt folgende χ^2 -Werte an:

*Malerei: $\chi^2_2 = 21,821$; $df = 3$, $p = 0,001$
Musik : $\chi^2_2 = 1,872$; $df = 3$, $p = 0,500$
Sport : $\chi^2_2 = 8,048$; $df = 3$, $p = 0,050$

Mit anderen Worten: Der χ^2 -Wert bei Malerei ist sehr signifikant auf dem 0,1% Niveau, der χ^2 -Wert bei Musik ist nicht signifikant, und der χ^2 -Wert bei Sport ist signifikant auf dem 5% Niveau.

Die Nullhypothese muß also zumindest im ersten Beispiel (Malerei) zurückgewiesen werden, denn χ^2 -Werte von 16,3 oder größer kommen nur in 0,1% aller Fälle vor (bei $df = 3$). Ob man eine Irrtumswahrscheinlichkeit von 5% wie im Beispiel 'Sport' für ausreichend hält, die Nullhypothese zu verwerfen

und die Alternativhypothese anzunehmen, ist weitgehend Ermessenssache, im Beispiel 'Malerei' muß die Alternativhypothese angenommen werden.

8. Abschließende kritische Bemerkungen

8.1 Zum Abschluß unserer Auseinandersetzung mit der Auswertung der Untersuchungsdaten sollen nur noch einige kritische Bemerkungen zu den in der Zusammenfassung (vgl. S. 218-226) aufgeführten Untersuchungsergebnissen formuliert werden.

Generell ist gegen diese Zusammenfassung einzuwenden, daß meistens nicht genau ersichtlich ist, auf welchen Ergebnissen der vorherigen Detailanalysen die gezogenen Schlußfolgerungen basieren. Zudem werden die Ergebnisse oft im Sinne des empirisch nicht überprüften Rezeptionsmodells interpretiert, mit anderen spekulativen und vagen Behauptungen vermengt wie z.B. der folgenden:

"Jeder Leser aktiviert weitere semantische Komponenten, womit ein persönlicher Arbeitsprozeß einsetzt, in dem die Vermischung von deskriptiven und konnotativen Elementen möglich wird." (S. 219)

8.2 Im folgenden wollen wir nun unter den Punkten (a) - (g) einzelne der nach Meinung von Bauer et al. aus dem Untersuchungsmaterial resultierenden Ergebnisse aufführen und jeweils kurz zu ihnen Stellung nehmen.

(a) Entgegen der Auffassung mancher Autoren, daß bei der freien Assoziation nur individuelle Komponenten ausschlaggebend seien, könnten "durch die Analyse der möglichen Rezeptionsverläufe im semantischen Bereich überindividuelle Gesetzmäßigkeiten festgestellt werden" (S. 218). Zunächst trete in der Mehrzahl der Assoziations-

texte eine Art gemeinsamer Nenner des jeweiligen Konzepts, die sog. "Primärassoziation" hervor, dessen Rezeptionskonstanten sich über sozial und bildungsmäßig stark differierende Gruppen hinweg nicht veränderten und in einem "mikrosemantischen Primärbereich" zusammenzufassen seien. An diesen Bereich schließe sich ein "potentieller mikrosemantischer Bereich", der durch Assoziationen gefüllt werde, die vom subjektiven Erfahrungskontext beeinflusst seien; dabei findet zugleich eine Vermischung von deskriptiven und kollotativen Elementen statt. (S. 219)

Aufgrund einer einzigen empirischen Untersuchung, die nur einen einzigen Text betrifft, bereits von Gesetzmäßigkeiten für die freien Assoziationen zu sprechen, dürfte verfrüht sein. Bei der Freiburger Untersuchung handelt es sich allenfalls um die Feststellung von Gemeinsamkeiten, die durchaus nicht gering zu schätzen sind.

- (b) "Mit dem Einsetzen der affektiv-subjektiv gelenkten Wortstimuli im Bereich der Konnotationen" werde die eigentliche Multivalenz des jeweiligen Wortes bzw. des Textes erkennbar. Während dieses Vorgangs, in dem der Leser den Konnotationsbereich aktiviere und schließlich einen bestimmten Konnotationskomplex als Schwerpunkt setze, erfolge die "makrosemantische Entscheidung" des Lesers. (S. 219)

Woher beziehen Bauer et al. ihre Informationen darüber, daß eine "makrosemantische Entscheidung" des Lesers erfolgt, während er den "Konnotationsbereich aktiviert"? Auf S. 217 führen die Autoren zu diesem Punkt aus:

"Bei einer Durchsicht des Interpretationsmaterials fällt auf, daß die Entscheidung für eine Interpretationsrichtung keineswegs willkürlich und plötzlich an einem bestimmten Punkt des Rezeptionsvorganges zustandekommt, sondern durch die vorangegangenen Assoziationen und das Semantische Differential vorbereitet ist. Es werden z.B. häufig wörtlich aus der Analyse des Assoziationsmaterials bekannte Syntagmen wiederholt. Auffällig ist auch das Wiederaufnehmen von ebenfalls in den Assoziationsfeldern wirksamen Konnotationen, die in den Interpretationstexten zu einer globalen Textdeutung verarbeitet werden und damit für die jeweilige Interpretationsvariante des jeweiligen Lesers bestimmend werden."

Die hier mitgeteilte Beobachtung, daß sich bestimmte in den Assoziationen vorkommende Elemente bei den Interpretationstexten wiederauffinden lassen, reicht nicht zur Stützung der Behauptung aus, die Entscheidung für eine Interpretationsrichtung werde durch die Assoziationen und das Semantische Differential vorbereitet. Es ist nämlich vorerst nicht auszuschließen, daß eine solche Entscheidung schon vor der Beantwortung der Assoziationsfragen fällt. Um über den möglichen Einfluß der Assoziationsfragen auf den Rezeptionsverlauf etwas aussagen zu können, müßten Kontrollversuche durchgeführt werden, in denen entweder solche Assoziationsfragen nicht gestellt werden dürften oder die Entscheidung für eine bestimmte Interpretationsrichtung schon vor Stellung dieser Fragen abverlangt werden müßte.

- (c) Die "literarische Erfahrung" des Lesers trete als Motivation für die Entscheidung im makrosemantischen Bereich zurück. Daraus folge, "daß bei der Wirkungsanalyse zeitgenössischer Texte nicht primär vom literaturhistorischen Kontext und vom allgemeinen ästhetischen Hintergrund des Lesers auszugehen ist, sondern von den textspezifischen Momenten im Rahmen des subjektiv kollotierten Erfahrungsbereichs." (S. 220)

Vorausgesetzt, daß die von Bauer et al. ermittelten Daten überhaupt eine solche Schlußfolgerung in Bezug auf den Celan-Text zulassen, ist dieses Ergebnis nicht ohne weiteres auf andere Gedichte von Celan übertragbar und schon gar nicht allgemein auf die Lyrik nach 1945 mit ihren vielen unterschiedlichen Richtungen.

- (d) Die Erfahrung mit literarischen Texten habe sich bei der Gruppe von Vpn, die im Hinblick auf die zeitgenössische Lyrik als "fortschrittlich" einzustufen seien (vgl. Fragen 1 - 10), nur insofern bemerkbar gemacht, als nach der Lektüre des Gedichts keine Verständnisschwierigkeiten aufgetreten und eine gewisse Routine in der Artikulation der Metatexte sichtbar geworden sei. (S. 220/221)

Insbesondere die letzte Folgerung ist problematisch. Die Anlage des Fragebogens zielt ja auf eine, von den Autoren gewollte Steuerung des Rezeptionsvorganges ab, deren tatsächliche Auswirkung allerdings ohne zusätzliche Untersuchungen nicht abgeschätzt werden kann. Genau betrachtet liefern Bauer et al. dem weniger versierten Leser im Fragebogen gerade bestimmte Vorinformationen implizit nach, die traditionellerweise als notwendig für das Verständnis von "dunkler Lyrik" erachtet werden. Zu diesen Informationen gehört insbesondere die von Bauer et al. für einschlägig gehaltene Auffassung, durch freie Wortassoziationen werde ein Zugang zu dem Gedicht von Celan ermöglicht (vgl. S. 25-27); diese Auffassung führt nämlich zu der Aufnahme der Assoziationsfragen in den Fragebogen und mit diesen Fragen wird den Vpn eine bestimmte Art der Betrachtung und Interpretation des Gedichtes nahegelegt, die sonst evtl. aufgrund der Kenntnis des Autors gewählt wird.

- (e) Innerhalb der Gesamtheit der Vpn könnten drei Gruppen unterschieden werden, die jeweils divergierende Interpretationsrichtungen realisiert hätten, was auf unterschiedliche Schwerpunktsetzungen zurückzuführen sei. (S. 221)

Zu diesem Ergebnis ist die kritische Frage zu stellen, ob es möglicherweise dadurch zustandekommt, daß die Vpn in Frage 27 zu einer auf diese Interpretationsrichtungen hinführenden Schwerpunktsetzung gezwungen werden (vgl. hierzu die Kritik in 2.2 an der Präsupposition in Frage 27).

- (f) Der durch die Ungewöhnlichkeit des Celan-Textes erschwerte Prozeß der semantischen Analyse werde gestört, wenn der Leser im Text eine bestimmte vom Autor verschlüsselte Intention suche, oder wenn er bereits bekannte Tendenzen des Autors in den Text projiziere. Die Multivalenz des Textes werde dadurch von außen eingeschränkt und der Leser nehme sich die Möglichkeit, den weitergespannten Rahmen des Textes zu erkennen. Weiterhin führe fachspezifische Interpretationsroutine häufig zu einer verallgemeinernden Textdeutung. Auffallend sei hier das Verhalten der Oberstufenschüler und der Philologiestudenten aus den ersten Semestern gewesen, in deren Metatexten es zu einer starken Häufung "klischeehafter Begriffe" ("Hoffnung", "Optimismus") gekommen sei, ohne daß vom Leser ein differenziertes semantisches Bewußtsein und ein eigenes Textverständnis erreicht werde. (S. 222)

Mit welcher Berechtigung sprechen Bauer et al. von einer "Störung" des Prozesses der semantischen Analyse in dem Fall, wo der Leser im Text eine vom Autor verschlüsselte Intention sucht? Es ist doch legitim, wenn ein Leser versucht, eine mutmaßliche Autorintention aus dem Text zu rekonstruieren, und wenn er dabei aufgrund seiner Kenntnis von anderen Texten des Autors bestimmte Bedeutungsmöglichkeiten des Textes als weniger wahrscheinlich ausschließt. Es handelt sich in einem solchen Fall um eine Beeinflussung des Rezeptionsprozesses, die Bauer et al. offensichtlich deshalb als negativ bewerten, weil sie aus dem Vorhandensein von "Multivalenz", das sie für das charakteristische Merkmal poetischer Texte halten, unberechtigterweise den normativen Anspruch herleiten, ein "multivalenter" Text müsse auch als solcher rezipiert, d.h. "der weiter gespannte Rahmen des Textes" müsse erkannt werden. Der von Bauer et al. abgelehnte "Absolutheitsanspruch, mit dem traditionelle Textinterpretationen auftreten" (vgl. S. 222), wird hier durch ein neues Dogma abgelöst, das aber den Widerspruch in sich trägt, daß einerseits dem Leser eine gewisse Freiheit bei der Interpretation zugestanden wird und andererseits bestimmte Formen der Interpretation als nicht adäquat ausgenommen werden. In diesem Sinne ist auch die Bemerkung über das Verhalten der Oberstufenschüler und Philologiestudenten aus den ersten Semestern zu kritisieren. Wenn in den Metatexten dieser Vpn bestimmte angelernte Klischeevorstellungen reproduziert werden, so gehört das eben - jedenfalls oberflächlich gesehen - zu ihrem "Textverständnis" und es kann ihnen nicht ein anderes "eigenes Textverständnis" aufgezwungen werden (in diesem Zusam-

menhang wäre allerdings genauer empirisch zu untersuchen, inwieweit der formulierte Metatext in verschiedenen Rezipientengruppen einschlägig für das theoretisch zu postulierende "Textverständnis" ist; gerade bei Schülern könnte man vermuten, daß das Anfertigen von Interpretationen mechanisch angelernt ist und diese Interpretationen nur teilweise mit der tatsächlichen Bedeutungszuordnung übereinstimmen).

- (g) Das Fehlen signifikanter Unterschiede in der Rezeption von Lesern mit "ausgeprägtem literarischem Erwartungshorizont" und Lesern mit geringer literarischer Erfahrung stütze die theoretisch postulierte Hypothese, daß ein Leser mit durchschnittlicher Sprachkompetenz in der Lage sei, zu einem Verständnis eines multivalenten zeitgenössischen Textes zu gelangen. (S. 223)

Ein Punkt, der schon mehrfach kritisiert wurde und der auch bei den Bemerkungen zu (f) hätte genauer problematisiert werden können, nämlich die unreflektierte und unklare Verwendungsweise des Begriffes "Textverständnis" muß im Zusammenhang mit der in (g) angesprochenen Hypothese noch einmal aufgegriffen werden. Spätestens bei der Formulierung dieser Hypothese hätten sich Bauer et al. näher mit der Frage auseinandersetzen müssen, was ein "Textverständnis" ist und wie im konkreten Fall gemessen werden kann, ob jemand ein solches "Textverständnis" erreicht hat oder nicht. Entweder ist ein Verständnis eines Textes jeweils das, was ein Rezipient unter dem Text versteht, d.h. ein Verständnis ist irgendeine Bedeutung, die ein Rezipient dem Text zuerkennt. In diesem Fall hat man für den Verständnisbegriff kaum Ausschlusskriterien und die genannte Hypothese wird nahezu trivial (zu ihrem Nachweis muß man die Leser nur dazu anregen, sich zu die-

sem Text irgendetwas vorzustellen, was für sie in irgendeiner Weise mit der Wortfolge im Text verträglich ist). Oder aber das, was als ein Verständnis des Textes zu gelten hat, ist durch den Text genau festgelegt und damit rezipientenunabhängig. In diesem Fall entsteht das Problem, wie objektiv darüber entschieden werden kann, ob ein Verständnis vorliegt oder nicht. Eine dritte Auffassung bestimmt "Verständnis" als den erfolgreichen Nachvollzug der Bedeutung, die der Autor mit seinem Text ausdrücken wollte. Bauer et al. scheinen in etwa die zweite Auffassung zu vertreten (wobei sie allerdings zulassen, daß ein Verständnis nur eine von mehreren möglichen Bedeutungen erfaßt), ohne sich des angedeuteten und u.E. nicht lösbaren Problems bewußt zu sein. Genauer gesagt stellt sich dieses Problem für Bauer et al. gar nicht und zwar deshalb, weil sie implizit das als Textverständnis voraussetzen, was sie selbst aufgrund ihrer präferierten ästhetischen Wertmaßstäbe und persönlichen literarischen Erfahrungen, bzw. was die literarisch versierten Leser für ein Textverständnis halten. Nur im Lichte einer solchen Voraussetzung ist wohl zu erklären, weshalb Bauer et al. behaupten, die genannte Hypothese werde dadurch gestützt, daß keine signifikanten Unterschiede in der Rezeption von Lesern mit ausgeprägtem literarischen Erwartungshorizont und Lesern mit geringerer literarischer Erfahrung vorliegen: dem ersten Typ von Lesern billigen Bauer et al. offensichtlich von vornherein die Befähigung zu, ein Verständnis eines zeitgenössischen multivalenten Textes zu erreichen.

LITERATURVERZEICHNIS

- Bauer, W.; Braunschweig-Ullmann, R.; Brodmann, H.; Bühr, M.;
Keisers, B.; Mauser, W. (1972), Text und Re-
zeption. Frankfurt.
- Celan, P. (1967), Atemwände. Frankfurt.
- Clark, H.H. (1970), Word Associations and Linguistic Theory.
In: Lyons, John (Hg.), New Horizons in Lin-
guistics. Harmondsworth. S. 271-286.
- Ebner, H.; Klaus, G. (1975)², Grundlagen der Statistik für
Psychologen, Pädagogen, Soziologen. Zürich
und Frankfurt.
- Ertel, S. (1965a), Standardisierung eines Eindrucksdifferen-
tials. In: Zeitschrift für experimentelle und
angewandte Psychologie XII, 1. S. 22-58.
- Ertel, S. (1965b), Weitere Untersuchungen zur Standardisierung
eines Eindrucksdifferentials. In: Zeitschrift
für experimentelle und angewandte Psychologie
XII, 2. S. 177-208.
- Hofstaetter, P.R. (1959), Zur Problematik der Profilmethode.
In: Diagnostica 1, V. S. 19-25.
- Kentler, H. (1959), Zur Problematik der Profilmethode. In:
Diagnostica 1, V. S. 5-18.
- Lienert, G.H. (1969)³, Testaufbau und Testanalyse. Berlin,
Basel, Weinheim.
- Osgood, Chr.E.; Suci, G.J.; Kannenbaum, Ph. (1957), The
Measures of Meaning. Urbana.
- Petöfi, J.S. (1974), Bemerkungen zu Wienold, G./H. Rieser
"Vorüberlegungen zur Rolle des Konzepts der
Textverarbeitung beim Aufbau einer empirischen
Sprachtheorie". (Erscheint demnächst in den Ak-
ten zum Kolloquium "Die Rolle der Grammatik in
der nichtautomatisierten und automatisierten
Textverarbeitung". Bielefeld. Februar 1974.)

- Roth, E. (1965), Die Faktorenanalyse in der Marktforschung. In: Bergler, R. (Hg.), Psychologische Marktanalyse. Bern und Stuttgart. S. 282-325.
- Schmidt, S.J. (1971), Ästhetizität. München.
- Schmidt, S.J. (1973), Texttheorie. München.
- Schmidt, S.J. (1974), Elemente einer Textpoetik. München.
- Stosberg, M. (1972), Analyse der Massenkommunikation: Einstellungsmessung. Düsseldorf.
- Überla, K. (1971)², Faktorenanalyse. Berlin, Heidelberg, New York.
- Wienold, G.; Rieser, H. (1974), Vorüberlegungen zur Rolle des Konzepts der Textverarbeitung beim Aufbau einer empirischen Sprachtheorie. (Erscheint demnächst in den Akten zum Kolloquium "Die Rolle der Grammatik in der nichtautomatisierten und automatisierten Textverarbeitung". Bielefeld. Februar 1974.)
- Wirrer, J. (1976a), Über Fragen zu Texten in Schulbüchern. In: Kindt, W.; Schmidt, S.J. (Hg.), Interpretationsanalysen. München.
- Wirrer, J. (1976b), Natürliche Textverarbeitung. Kleines Lexikon der Linguistik. Stichwort "Natürliche Textverarbeitung". In: Linguistik und Didaktik 27 (im Erscheinen).
- Zobel, R. (1975), Der Dramentext - Ein kommunikatives Handlungsspiel. Göppingen.