

Grammatikforschung als Naturwissenschaft?*

Ralf Vogel

Universität Bielefeld

Der Aufsatz versucht eine kritische Revision der gängigen Auffassung, dass Chomsky in seiner 50 Jahre zurückliegenden Skinner-Rezension den Behaviorismus im Bereich der Sprachwissenschaft erledigte und, darauf aufbauend, die damit zusammenhängende Frage, wie eine empirisch orientierte Grammatikforschung vorzugehen hat, zufriedenstellend beantwortete. Eine Diskussion grundsätzlicher Probleme, die in älteren und neueren Entgegnungen zum klassischen Paradigma der Universalgrammatik hervorgebracht wurden, mündet in ein Plädoyer für eine wissenschaftliche Grammatikforschung, die Sprachen als empirischen Untersuchungsgegenstand sui generis betrachtet und keiner Anlehnung an irgendeine Naturwissenschaft bedarf.

Keywords: verbal behavior, philosophy of linguistics

„The world has many aspects: mechanical, chemical, optical, electrical and so on. Among these are its mental aspects. The thesis is that all should be studied in the same way, whether we are considering the motion of the planets, fields of force, structural formulas for complex molecules or computational properties of the language faculty.“ (Chomsky, 1996, 31)

* Dies ist eine elaboriertere Version eines Vortrags, den ich anlässlich meiner Habilitation im Juli 2005 an der Universität Potsdam hielt. Ich danke allen TeilnehmerInnen der dortigen Diskussion für hilfreiche Anregungen, insbesondere Gisbert Fanselow, Peter Staudacher und Reinhold Kliegl. Meine Ansichten zu diesem Themenkomplex haben sich in der Zwischenzeit hier und da verändert, so dass dieser Aufsatz nur noch bedingt meiner damals präsentierten Sicht der Dinge entspricht. Meine Auffassung wurde wesentlich geprägt von zwei Seminaren zu sprachphilosophischen Fragestellungen, die ich an der Universität Potsdam zusammen mit Stefan Frisch und Heiner Drenhaus in den Jahren 2004 und 2005 hielt. Mein Dank geht auch an diese beiden Kollegen sowie die TeilnehmerInnen der beiden Seminare. Besonderer Dank gilt Markus Meyer für einige wertvolle Hinweise sowie den sehr fruchtbaren Gedankenaustausch, den wir jetzt schon einige Zeit führen. Alle Irrtümer, die dieser Aufsatz enthält, sind natürlich mir alleine anzulasten.

Inwiefern Linguistik als Teil des Studiums der Natur betrachtet werden kann oder soll, ist auch knapp fünfzig Jahre, nachdem dies von Noam Chomsky im Rahmen der generativen Grammatik postuliert wurde, noch umstritten.

Für zweifelsfrei naturwissenschaftliche Untersuchungsgegenstände wie beispielsweise die Anatomie der Artikulations- und Perzeptionsorgane, oder auch die Neurophysiologie des Sprechens und Verstehens, fällt die positive Antwort relativ leicht.

Ganz anders verhält es sich mit dem Studium der *Grammatik*, also der relativ abstrakten Regelsysteme, durch die unsere Sprachen charakterisiert sind. Einerseits ist eine Sprache ein Kulturprodukt, das in einer Sprachgemeinschaft geformt und weiterentwickelt wird, inklusive seiner Grammatik. Andererseits bringen wir offensichtlich mit unserer biologischen Ausstattung Voraussetzungen unserer Sprachfähigkeit mit, die die Gestalt einer Grammatik wesentlich vorbestimmen. Sonst wäre nicht zu verstehen, wieso Tiere keine Sprache (unserer Art) beherrschen.

Allgemeinplätze dieser Art sind die Ausgangsposition für die durch die generative Grammatik forcierte Positionierung der Grammatikforschung im Bereich der Kognitionspsychologie und -biologie. Die linguistische Forschung konnte so auf diesem Feld wesentliche Impulse liefern und es ist ihr als Disziplin dabei auch stets gelungen, sich als Trendsetterin in den modernen Wissenschaften vom Geiste darzustellen.

Als eines der wichtigsten frühen Dokumente dieser Neuausrichtung der Linguistik gilt Chomskys Rezension von B. F. Skinners „Verbal Behavior“ (Chomsky, 1959), deren Erscheinen sich 2009 zum 50ten Male jährt. Ich möchte dieses Datum zum Anlass nehmen für eine kritische Bestandsaufnahme.¹

Abschnitt 1 dieses Aufsatzes diskutiert den linguistischen Empirismus und

¹ Die Reichweite dieser Buchbesprechung lässt sich beispielsweise an der ungewöhnlich hohen Zitationsrate der Besprechung (etwa 750 Zitierungen auf *google scholar* am 1.9.2008) im Verhältnis zum besprochenen Werk (knapp 1900 Zitierungen) ermessen.

den Behaviorismus als seine kongeniale Ergänzung, vor allem im Hinblick auf Chomskys Kritik an Skinner. Ich werde dabei besonderes Augenmerk auf im Kreise der generativen Linguistik weniger bekannte Entgegnungen auf Chomsky aus dem Kreise des Behaviorismus legen, die zeigen, dass Chomskys Furor in mehrerer Hinsicht unbegründet und übertrieben vorgetragen war und auch an der Sache vorbei zielte, um die es in „Verbal Behavior“ eigentlich ging. Im Wesentlichen stützt Chomsky die Psychologie hier auf eine Sichtweise zurecht, in die sich die Linguistik, so wie Chomsky sie versteht, einordnen lässt, als zentraler Bestandteil einer sich zu der Zeit neu gründenden „Cognitive Science“.

Der eigentliche Kern der Chomskyschen Kritik ist eine postulierte Inadäquatheit des behavioristischen Lernmodells. Chomskys Gegenentwurf wird in Abschnitt 2 kurz dargestellt. In Abschnitt 3 werden zwei prinzipielle Probleme der Chomskyschen Linguistik-Konzeption und seiner Lerntheorie vorgestellt, das Argument von der Unzulänglichkeit des kindlichen Sprachinputs im Spracherwerb („Poverty of the stimulus argument“) für den Nativismus sowie das philosophische Problem von der Unbestimmtheit der Übersetzung, das Quine als besondere Schwierigkeit für eine mentalistisch aufgefasste Linguistik ausmachte. Beide Problemkreise lassen meines Erachtens erkennen, auch vor dem Hintergrund aktueller Erkenntnisse und Diskussionen, dass gegenwärtig viel mehr für die behavioristische Sicht spricht, vielleicht nicht in ihrer radikalsten Variante, denn für die nativistische/mentalistische Konzeption. Dies motiviert eine alternative Sichtweise, eigentlich eine Weiterentwicklung der vor-Chomskyschen, derzufolge Grammatikforschung als Teil einer eigenständigen wissenschaftlichen Linguistik aufzufassen ist. Ihr Gegenstand sind die sprachlichen Erscheinungen, eine „linguistische Realität“ (Devitt, 2006), die keiner Anleihen von oder Anlehnungen an andere Disziplinen bedarf. Im abschließenden Abschnitt wird diese Schlussfolgerung ein wenig ausgeführt werden.

1 Empirismus und Behaviorismus

Dass die Gestalt natürlicher Sprachen nicht zu erklären ist, wenn man nicht die *biologischen Voraussetzungen* derjenigen Organismen, die sie ausüben, einbezieht, kann gemeinhin als unstrittig gelten. Wie *sprachbezogen* diese biologischen Voraussetzungen sein müssen, ist allerdings nach wie vor eine offene und kontrovers diskutierte Frage. Für Chomsky, das wurde in seinen Schriften immer wieder deutlich, ist es diese *biologische* Perspektive – die Annahme sprachspezifischer biologischer Voraussetzungen der Sprache und deren Charakterisierung –, die im Zentrum der Grammatik-Forschung als Kerndisziplin der Linguistik stehen soll, und die die Linguistik zur *Naturwissenschaft*, einem Teilbereich der kognitiven Psychologie und der Humanbiologie, macht. Diese Position kennen wir als *linguistischen Nativismus*, der annimmt, dass grundlegende grammatische Prinzipien in Form der *Universalgrammatik* (UG) als Teil eines „Language Acquisition Device“ zu unserer angeborenen Ausstattung gehören. Wichtig zu betonen ist hier noch einmal, dass es um den Grammatikbegriff geht: ist eine Grammatik bloß ein Konstrukt im Rahmen linguistischer Modellbildung, ein soziales, in einer Sprachgemeinschaft verankertes Phänomen, oder biologische Voraussetzung von Sprache, der im Sinne eines weiten Organbegriffs *Existenz* zugesprochen wird? Letzteres ist Chomskys Position, wie sie auch in dem diesen Aufsatz einleitenden Zitat eindeutig hervor tritt.

Ein frühes Dokument für Chomskys biologisch-psychologische Positionierung ist die bereits erwähnte Besprechung von B.F. Skinners (1957) „Verbal Behavior“ in der Zeitschrift *Language* (Chomsky, 1959). Chomsky hatte schon in früheren Arbeiten seine Kritik an der empiristischen, am Behaviorismus orientierten seinerzeitigen Hauptströmung der amerikanischen Linguistik begründet.

Die „Naturwissenschaftlichkeit“ der Linguistik an sich war dabei nicht der Streitpunkt – denn gerade der Behaviorismus forcierte eine solche Orientierung für die Psychologie –, sondern die Frage, *wie* Linguistik als Naturwissenschaft

‚richtig‘ zu betreiben ist.²

1.1 Bloomfield

Die amerikanische Linguistik der 30er, 40er und 50er Jahre war wesentlich durch den sogenannten *taxonomischen Strukturalismus* Leonard Bloomfields (1887-1949) geprägt, für den die Ausrichtung am Ideal der empirischen Wissenschaften, insbesondere der beobachtenden Naturwissenschaften, in Methode und Gegenstandsauffassung kennzeichnend ist. Dabei ging es aber darum, die Linguistik als *eigenständige* Disziplin mit eigenen Methoden der Beobachtung zu begründen.³ Ob die Sprache Teil der Natur ist, ist hier aber zweitrangig. Wesentlich ist, dass Linguistik als empirische Disziplin mit den *methodischen* Standards einer quasi ‚ordentlichen‘, also einer Naturwissenschaft betrieben wird.

Ziel ist die adäquate *Beschreibung* der beobachtbaren Sprache. Ausgehend von der Analyse und Beschreibung des sprachlichen Signals, der phonetischen Beschreibungsebene, sollen Stufe für Stufe abstraktere Ebenen der strukturellen Beschreibung erreicht werden, die Silbe, das Morphem, das Wort, die Phrase, der Satz. Das Bloomfieldsche Programm kann so als konsequente Durchführung des Programms des logischen Empirismus verstanden werden.

Strukturelle Gesetzmäßigkeiten sollten dabei durchaus erfasst werden. Die Aufgabe der Sprachwissenschaft war es, die *Aufdeckungs-Prozeduren* zu definieren, mit denen es gelingt, vom sprachlichen Signal ausgehend die abstrakteren Struktur-Eigenschaften einer Sprache zu finden.⁴

² Anders gesagt: es ging darum, welche Naturwissenschaft der Linguistik als Vorbild dienen soll. Empiristische Sichtweisen zeigen eine starke Nähe zur Biologie und Psychologie. Chomskys Ideal ist eindeutig die Physik.

³ Eine umfassende Darstellung seines wissenschaftlichen Programms findet sich beispielsweise in (Bloomfield, 1933).

⁴ In der zeitgenössischen Spracherwerbsforschung ist diese Problemstellung präsent unter dem Terminus *Bootstrapping*: wie gelingt es dem lernenden Kind, die höheren abstrakteren Struktureigenschaften der Sprache, die es lernt, aus dem ‚stream of sound‘, dem es ausgesetzt ist, zu extrahieren? Welche phonetisch-prosodischen Hinweise signalisieren also beispielsweise

Die Semantik, das Studium der Bedeutung sprachlicher Ausdrücke, wird dabei bewusst ausgeblendet. Hier bedient sich Bloomfield einer psychologisch-behavioristischen Sichtweise. Die Bedeutung einer sprachlichen Form soll die Situation sein, in der ein Sprecher sie äußert, sowie die Reaktion, die sie im Hörer hervorruft. Semantik ist demnach eher eine psychologische als eine linguistische Fragestellung.

Getreu seinem empiristischen Programm *konnte* er sich der Frage der Bedeutung gar nicht anders als über Beobachtungskriterien nähern. Von dieser Warte aus kann man Skinners (1957) *Verbal Behavior* als Bloomfield ergänzendes semantisches Forschungsprogramm verstehen.

1.2 Behaviorismus

Die behavioristische Schule der Psychologie entstand zu Beginn des zwanzigsten Jahrhunderts und wurde für einige Jahrzehnte zum dominierenden Paradigma dieser Disziplin. An ihrem Beginn stand die Kritik an einigem, was in der noch jungen, sich gerade erst aus der Philosophie herauslösenden und als eigenständige Wissenschaftsdisziplin etablierenden Psychologie vertreten wurde. Als Gründungsdokument des Behaviorismus kann (Watson, 1913) gelten. Darin heißt es gleich zu Beginn:

„Psychology as the behaviorist views it is a purely objective experimental branch of natural science. Its theoretical goal is the prediction and control of behavior. Introspection forms no essential part of its methods, nor is the scientific value of its data dependent upon the readiness with which they lend themselves to interpretation in terms of consciousness.“ (Watson, 1913, 158)

grundlegende syntaktische Eigenschaften der Sprache und wie kann das Kind sie erkennen?

In seiner Betonung des objektiv Beobachtbaren und der Konzentration auf methodische Fragen kann der Behaviorismus ganz als Kind seiner Zeit verstanden werden, für die beispielsweise auch der logische Empirismus in der Philosophie steht, von dem wesentliche Impulse für die Philosophie und die Wissenschaftstheorie im 20. Jahrhundert ausgingen. Die Abgrenzung, die hier vorgenommen wird, bezieht sich auf zwei in dieser frühen Phase der Etablierung der Psychologie sehr wichtige Konzepte, das Bewusstsein und die Methode der Introspektion. Es geht um den spekulativen Charakter aller Diskussion um das Phänomen des Bewusstseins und die Unzuverlässigkeit und insbesondere Unüberprüfbarkeit und Nicht-Objektivität introspektiver Beobachtung.

Soll die Psychologie sich zu einer „harten“ Wissenschaft entwickeln, dann dürfen solche spekulativen Konzepte in ihrer Fundierung höchstens Nebenrollen spielen. Im Kern muss es um die Formulierung objektiv überprüfbarer Theoriesysteme gehen. Beobachtbar ist aber nicht das Psychische an sich, sondern nur seine Konsequenz, das menschliche Verhalten.

Ein weit verbreitetes Vorurteil über den Behaviorismus besteht in der Unterstellung, dass er die Existenz innerer Zustände abstreite. Das eigentliche Argument für die Verhaltensorientierung ist aber nicht eine solche (Nicht-)Existenzbehauptung, sondern der oben erwähnte notwendig spekulative Charakter allen Redens über diese inneren Zustände. Mit dem Behaviorismus in seiner ursprünglichen Variante ist die Hoffnung verbunden, dass die Gesetzmäßigkeiten menschlichen Verhaltens vollständig formulierbar sind, ohne auf solche Spekulationen über innere Zustände angewiesen zu sein. Das von Chomsky zum Hauptuntersuchungsgegenstand der Linguistik erklärte „computational system of human language“ ist als „mental module“ von genau dieser kritisierten Art.

In der Grammatikforschung und -schreibung finden wir bis heute ein allgemein akzeptiertes Nebeneinander von rein introspektiv oder impressionistisch abgesicherter grammatischer Analyse einerseits und methodisch reflektierter empirischer Grammatik-Forschung andererseits. Der Autor dieses Auf-

satzes nimmt sich von dieser Unentschiedenheit nicht aus.

Vor dem Hintergrund der Diskussion um die (Natur-)Wissenschaftlichkeit der Grammatikforschung ist es für mich aber umso erstaunlicher, dass insbesondere in der generativen Strömung Methodik-Fragen der Datenerhebung lange Zeit so gut wie keine Rolle spielten, und generative Forschungsarbeiten zu meist auf einer eher impressionistisch gewonnenen Empirie beruhen, wodurch ihnen ein spekulativer Charakter anhaftet. Ich werde später darauf zurückkommen.

Der Behaviorismus hat in seiner Geschichte verschiedene Stadien durchlaufen und verschiedene Ausprägungen angenommen. Eine einigermaßen faire Darstellung von Graham (2007) findet sich in der philosophischen Online-Enzyklopädie der Stanford University.

Graham spricht in dieser Darstellung von Behaviorismus im weiten Sinne („behaviorism loosely understood“) und Behaviorismus als Doktrin. Behaviorismus im weiten Sinne ist das Prinzip, dass psychologische Theorien durch Beobachtung menschlichen Verhaltens verifizierbar sein müssen, um den Status wissenschaftlicher Theorien beanspruchen zu können: „Arguably, there is nothing truly exciting about behaviorism loosely understood. It enthrones behavioral evidence, an arguably inescapable practice in psychological science.“ (Graham, 2007, Abschnitt 1). Behaviorismus als Doktrin verstanden zieht aus einer solchen allseits geteilten Haltung radikale Konsequenzen (Graham, 2007, Abschnitt 2):

1. Psychologie ist die Wissenschaft vom Verhalten, nicht die Wissenschaft vom Geist.
2. Verhalten kann beschrieben und erklärt werden, ohne auf mentale Ereignisse oder interne psychische Prozesse zurückzugreifen. Die Ursachen eines bestimmten Verhaltens sind in der Umwelt zu suchen, nicht innerhalb des Organismus.

3. Im Verlauf psychologischer Theoriebildung sollen mentale Begriffe oder Konzepte, sofern sie verwendet werden, entweder a) aus der Theorie entfernt und durch verhaltensbezogene Begriffe ersetzt werden oder b) in solche verhaltensbezogenen Konzepte übersetzt oder paraphrasiert werden.

Graham unterscheidet zwischen *methodologischem* Behaviorismus, für den Aussage (1) charakteristisch ist, *psychologischem* Behaviorismus, der mit Aussage (2) korreliert, und *analytischem* Behaviorismus, für den Aussage (3) steht. Skinner nennt seine Fassung, die alle drei Aussagen teilt, „radikalen Behaviorismus“. Insbesondere Aussage (2) ist hier zu beachten: gemeint ist mitnichten, dass Skinner die Existenz mentaler Prozesse bestreitet. Sie kommen sogar in seinem Modell menschlichen Verhaltens vor. Aber das Modell beinhaltet keine Aussagen über deren Struktur, und seine „Wahrheit“ hängt auch nicht von der Wahrheit solcher Aussagen ab. Dies ist letztlich eine Bedingung für die Verifizierbarkeit des gesamten Theoriegebäudes.

Die Konditionierung ist die wichtigste experimentelle Methode der Verhaltensforschung. Das Paradebeispiel für die *klassische Konditionierung* ist der berühmte Pawlowsche Hund. Der Verhaltensforscher Pawlow untersuchte, wie der Speichelfluss bei Hunden zunimmt, wenn sie Futter gezeigt bekommen. Parallel zur Präsentation von Futter läutete ein Glöckchen. Der Speichelflussreflex setzte irgendwann auch ein, wenn nur das Glöckchen läutete, ohne dass Futter präsentiert wurde. Ein Reflexverhalten ist also damit auf einen neuen Reiz konditioniert worden.

Diese Art von Konditionierung eines Reflexauslösers hat wenig mit dem zu tun, was wir beim Menschen unter Lernen verstehen. Der lernende Organismus wird im klassischen Behaviorismus als passiv-reaktiv, unter dem reflexauslösenden Einfluß äußerer Reize stehend, betrachtet – wobei allerdings auch innere organische Zustände wie eine volle Harnblase als „äußere“ Um-

weltreize in diesem Sinne betrachtet werden.

Skinner's *Radikaler Behaviorismus* geht weit über dieses recht primitive Bild hinaus.⁵ Seine wichtigste Innovation ist die *operante Konditionierung*. Er lenkte mit ihr das Augenmerk auf das *operante Verhalten*, also Verhalten, das darauf abzielt, die Umwelt zu beeinflussen. Verhalten ist hier keine passive Reaktion auf Reize, sondern es entsteht spontan und wird durch seine Konsequenzen geformt. Das operante Lernen wird auch als Lernen durch Belohnung bezeichnet. Skinner's Verhaltensbegriff schließt auch innere Vorgänge als „verdecktes Verhalten“ ein. Skinner sieht die operante Konditionierung in der Evolution der Art gleichermaßen am Werk wie in der Lerngeschichte des Einzelnen.⁶

Sprachliches Verhalten ist nun allgemein jedes Verhalten, das nicht direkt auf die Umwelt einwirkt, sondern nur durch das Verhalten eines anderen Menschen einen Effekt hat. Diese Definition ist so allgemein, dass auch tierisches Verhalten hier mit einbeziehbar ist.

Nehmen wir als Beispiel eine Ratte, die in einem Experiment in einem Käfig gehalten wird. In diesem Käfig gibt es eine Lampe und einen Schalter, den die Ratte betätigen kann. Der Schalter öffnet eine Klappe, hinter der bei leuchtender Lampe Futter zu finden ist. Mit der Zeit lernt die Ratte diesen Zusammenhang, so dass sie irgendwann nur bei angeschalteter Lampe den Schalter betätigt, um Futter zu bekommen. In gewissem Sinne tritt die Ratte mit dem Experimentator hier in eine Dialog-Situation. Zwischen ihnen ist mittels der operanten Konditionierung eine Sprache entstanden. Der dieser sehr allgemeinen Auffassung von Sprache angemessenere Begriff ist wohl eher „Kommunikation“. Um dies

⁵ Dies ist einer der Punkte, die Chomsky offenkundig nicht verstanden hatte, oder nicht verstehen wollte. Seine Rezension zeichnet ein Zerrbild des Behaviorismus, das so vermutlich nie vertreten wurde, und von Skinner schon gar nicht. Es entsteht der Eindruck, Behavioristen würden jegliches Verhalten auf simple Reiz-Reaktionsmuster nach Art der klassischen Konditionierung zurückführen.

⁶ Sehr gut recherchierte Überblicksartikel zu „Verbal Behavior“ und „Behaviorismus“ finden sich auf Wikipedia: <http://de.wikipedia.org/wiki/Behaviorismus> und http://de.wikipedia.org/wiki/Verbal_Behavior.

von Sprache im eigentlichen Sinne abzugrenzen, wurde für Sprache darüber hinaus das Kriterium der Austauschbarkeit der Rollen eingeführt, was in unserer Laborsituation nicht gegeben ist: der Experimentator kann die Ratte füttern, aber nicht umgekehrt die Ratte den Experimentator.

Sprachliches Verhalten ist gleichwohl mehr als das Sprechen einer Sprache, alle, auch gestischen oder schriftlichen, Handlungen fallen hierunter, wenn sie sich an eine andere Person richten. Sprachliches Verhalten ist behavioristisch betrachtet also im Grunde alles, was wir als Kommunikation zwischen gleichberechtigten Partnern, also Menschen, eventuell auch höhere Säugetiere, bezeichnen.

Unser Beispiel für operante Konditionierung zeigt noch etwas anderes, nämlich die Rolle von inneren Zuständen im behavioristischen Modell Skinnerscher Prägung.

Um sich das Verhalten der Ratte zu erklären, sind nämlich zwei Begriffe unabdingbar, die auf so etwas wie innere Zustände rekurrieren: *Motivation* und *Erfahrung*. Warum betätigt die Ratte überhaupt Schalter? Weil sie Hunger hat. Dies ist ihre *Motivation*. Wie kommt sie dazu, nach gewisser Zeit auf das Lampenlicht systematisch mit Schalterdrücken zu reagieren? Weil sie durch vorangehende Ereignisse konditioniert wurde. Diese *Erfahrung* hat sich der Ratte eingeprägt und die Assoziation von Licht und Schalterdrücken verstärkt. Die Gesetzmäßigkeiten dieses Konditionierungsmechanismus' bleiben erhalten, was auch immer künftige Forschung über das Gehirn und die neurophysiologischen Korrelate von Motivation und Erfahrung hervorbringen werden.

In diesem Sinne ist behavioristische Modellbildung zwar nicht frei von Bezügen zum Mentalen, aber doch unabhängig davon, wie wir es uns vorstellen. Das behavioristische Modell macht keine Aussage darüber, wie Motivation und Erfahrung neurobiologisch „implementiert“ sind. Es beschränkt sich darauf zu sagen: wenn man einen Organismus bestimmten Umweltbedingungen aussetzt, wird er gesetzmäßig ein bestimmtes Verhalten hervorbringen. Dazu muss er

über (die Fähigkeit zu) Motivation und Erfahrung verfügen.⁷

Hier setzt einer von mehreren teils sehr harschen Kritikpunkten Chomskys ein. Wie sehr er dabei dem behavioristischen Ansatz Unrecht tut und ihn möglicherweise bewusst missversteht, verdeutlichte MacCorquodale (1970) in einer der eher rar gesäten, dafür sehr ausführlichen Entgegnungen auf Chomsky aus dem Kreis des Behaviorismus:

„Chomsky does not refer to Skinner’s discussions of why he omits mediational constructs, but he was apparently little affected by them, finding the violation of his own preconceptions a sufficient reason for ignoring them: „One would naturally expect that the prediction of the behavior of a complex organism (or machine) would require, in addition to information about external stimulation, knowledge of the internal structure of the organism, the ways in which it processes information and organizes its own behavior“ (Chomsky, 1959, p. 27). Perhaps one would, but he need not. It is perfectly feasible and sufficient to know merely that the speaker’s „internal structure . . . processes information“ so as to generate lawful relations between the speaker’s circumstances (past and present) and his speech. Unless one is a neurophysiologist it is not necessary in the least to know how the internal structure goes about doing so nor which structures are involved. The psychologist’s knowing how it does so would not improve the precision of predicting behavior from knowledge of the speaker’s circumstances, nor would this knowledge make existing functional laws of behavior any more true, nor could it show them to be untrue, It is simply false, of course, that one cannot accurately predict behavior, even complex behavior, without knowing and taking into account the behavior’s

⁷ Allerdings haben Konditionierungs-Experimente durchaus Erkenntnisse über die Rolle von Motivation und Erfahrung beim Lernen hervorgebracht.

structure and internal processes; we do it all the time.“

(MacCorquodale, 1970)

Aufschlussreich sind an dieser Passage die unterschiedlichen Auffassungen davon, was es heißt, sprachliches Verhalten zu verstehen. Die behavioristische Sicht ist hier, dass die korrekte Vorhersage sprachlichen Verhaltens die Gültigkeit einer Theorie alleine untermauert. Chomskys Einwand, dass man doch gerne genauer wissen möchte, wie Gehirn und Geist in Bezug auf Sprache funktionieren, appelliert an die Intuition, dass Sprache etwas ist, das unser Gehirn hervorbringt. Zu verstehen, wie Sprache funktioniert, heißt dann, zu verstehen, wie das Gehirn das macht.

Man kann in diesen beiden Auffassungen auch erste Abgrenzungsversuche der sich herausbildenden Kognitionswissenschaft zur etablierten Experimentalpsychologie erkennen. Was die empirische Seite betrifft, so kann man Graham (2007) zustimmen, dass ein Behaviorismus im weiten Sinne unbestrittene methodische Maxime der Kognitionsforschung wie der Psychologie überhaupt ist. Natürlich hängt jede wissenschaftliche Hypothese über die Funktionsweise der menschlichen Kognition am empirischen Nachweis in Form von behavioraler Evidenz. Behaviorismus im engen Sinne beginnt damit, dass Psychologie auch nur als Wissenschaft vom menschlichen Verhalten gedacht wird.

Es stellt sich die Frage, welche Bedeutung dieser Schritt vom weiten zum engen Behaviorismus für die linguistische Praxis hat. Ein behavioristisches Modell unterscheidet sich vom mentalistischen nur im Verzicht auf die ohnehin spekulativen mentalistischen Aspekte. Und um dies an dieser Stelle auch deutlich zu machen: sprachliches Verhalten ist natürlich Verhalten *in einer bestimmten* Sprache. Insofern ist klar, dass jede Verifizierung oder Falsifizierung universalgrammatischer Hypothesen über das Studium konkreter Sprachen verläuft, allem Gerede von der universalen menschlichen Sprachfakultät und der postulierten Irrelevanz oder wenigstens Nachrangigkeit einzelsprachlicher Forschung

oder gar soziolinguistischer Erkenntnisse zum Trotz.

Das eigentlich Neue an der entstehenden Kognitionswissenschaft war die gestiegene Bedeutung der Modellierung, insbesondere im Rahmen der Künstlichen Intelligenzforschung. Chomskys Sprachauffassung kam diesem Zug der Zeit natürlich sehr entgegen. Der wesentliche Beitrag seiner frühen Arbeiten bestand darin, einer formalen, an die Kognitionsforschung anschlussfähigen Grammatikauffassung zum Durchbruch verholfen zu haben. Eine Grammatik muss algorithmisierbar und mithin auch implementierbar sein. Nur eine solche vollständig ausformulierte Grammatik kann überhaupt erst auf ihre empirische Adäquatheit überprüft werden. Eine moderne Grammatikforschung ist ohne eine solche Formalisierung nicht mehr denkbar. Dies ist das bleibende Verdienst der Chomskyschen Schule.

Dazu versprach man sich von der Computermodellierung des menschlichen Geistes, hier besonders der menschlichen Sprachfähigkeit, Einsichten in die menschliche Natur, zumindest was die komputationellen Aspekte der Kognition betrifft. Bis heute erklärt Chomsky C_{HL} , das „computational system of Human Language“ zu seinem Untersuchungsgegenstand.⁸ Die Frage der Lernbarkeit dient darüber hinaus als wichtiges Kriterium, um „erklärungsadäquate“ von bloß „beschreibungsadäquaten“ Grammatiken zu unterscheiden. Spracherwerb wird damit zu einem für die Postulate der generativen Grammatik essentiellen Forschungsgegenstand.⁹

⁸ Dazu gehört, und hierin steckt nun auch wieder die Abgrenzung zu einer bloß implementierenden KI-Forschung, dass das entworfene komputationelle Modell psychologisch realistisch sein soll. Das führt zu einer seltsamen Eigenart des Chomskyschen Denkens. Eine generative Grammatik ist einerseits eine Theorie über das Regelsystem einer Sprache oder sprachliche Regelsysteme allgemein. Andererseits wird durch die nativistische Annahme den in diesem Modell postulierten Regeln selbst Realität zugesprochen. Der nativistische Ansatz läuft also beständig Gefahr, Theorie und Untersuchungsgegenstand in eins zu setzen. Siehe dazu ausführlich (Meyer, 2006, Kapitel 1) sowie Devitt (2006).

⁹ In der Diskussion um den angeborenen Spracherwerbsmechanismus spielte anfangs auch die Frage eine Rolle, wie ein Lerner zwischen mehreren mit dem frühkindlichen sprachlichen Input verträglichen Grammatiken die „richtige“ Grammatik auswählen kann. Diese

Eine Psychologie, die die Beschäftigung mit dem Geist als unwissenschaftlich abtut, nämlich der Behaviorismus, steht einer solchen Sichtweise, mithin einer auch finanziell einträglichen Forschungsperspektive, diametral entgegen.¹⁰ Auch hieraus erklärt sich, wieso Chomsky sich in der Auseinandersetzung mit der behavioristischen Sprachpsychologie so eindeutig positionierte.

Völlig rätselhaft ist mir, wieso Chomsky nicht den Unterschied zwischen der psychologischen und der sprachwissenschaftlichen Sicht auf Sprache zur Kenntnis nimmt – ihm unterläuft eine geradezu groteske Fehlinterpretation dessen, worum es in „Verbal Behavior“ geht. MacCorquodale (1970) bringt es auf den Punkt, indem er aufdeckt, dass Chomsky Skinners Wahrscheinlichkeitsbegriff nicht versteht:

„Much more ominously for Skinner’s purposes, Chomsky seems not to grasp the difference between the overall probability of occurrence of an item in a speaker’s verbal repertoire, which is the frequency with which it occurs in his speech over time without regard to his momentary circumstances, and the momentary probability of a given response in some specified set of circumstances. (See, for example, Chomsky, 1959, p. 34.) The two probabilities are very different. The overall probability that any speaker will say,

Evaluierungsmetrik ist heutzutage aus der Diskussion so ziemlich verschwunden, obwohl das Problem im Prinzip nicht gelöst ist. Siehe dazu auch den Abschnitt 3.2.

¹⁰ Eine gewisse Pikanterie besteht darin, dass Forschung im Rahmen der generativen Grammatik in den Sechziger Jahren großzügig vom US-Militär gefördert wurde, so wie auch Chomskys Arbeit am MIT selbst praktisch von Anfang an – seiner linksradikalen politischen Haltung zum Trotz.

Ich entsinne mich gerne der Seminare in englischer Sprachwissenschaft während meines Studiums an der Universität Frankfurt/Main bei einem Professor, der sich ironisch damit rühmte, seine Dissertation einst Ende der Sechziger Jahre im „UCLA Air Force English Syntax Project“ verfasst zu haben. Neben einer ausführlichen Zusammenfassung der generativen Arbeiten zur Syntax des Englischen ging es in diesem Projekt auch um die computerlinguistische Umsetzung transformationsgrammatischer Modelle mit dem Ziel ihrer Anwendung in der inner-militärischen Kommunikation. Der Erfolg war wohl eher bescheiden.

for example, “mulct”, is very low; it occurs rarely in comparison with such responses as “the” or “of”. The probability that he will say “mulct” may become momentarily extremely high, as when he sees the printed word. Of the two, overall probability is a typically linguistic concern, while momentary probability shifts are, in a sense, the very heart of the psychologists’ problem, since they reflect the relation between speech and its controlling variables. [...] If Chomsky really did not, in fact, see this difference it is impossible to imagine what the rest of Verbal Behavior could have meant to him, and no wonder that he regarded it with such astonishment and dismay. “

Um ein Beispiel zu nehmen: Wir beobachten eine Situation, in der eine Person ein Musikstück hört und daraufhin „Mozart“ sagt. Skinners Fragestellung ist, welche Komponenten ein wissenschaftliches Modell sprachlichen Verhaltens beinhalten muss, das genau dieses Verhalten in dieser Situation vorhersagt. Dazu gehört die gesamte Lebensgeschichte der handelnden Person, inklusive ihrer sprachlichen Erfahrung, und dazu gehört natürlich auch der kontrollierende Reiz, wobei der Begriff „Kontrolle“ nicht so simplifizierend zu verstehen ist, wie Chomsky es in seiner Strohmann-Version von Behaviorismus tut: ohne das Hören des Musikstücks hätte die Person nicht „Mozart“ gesagt, aber eine andere Person, ungebildet in klassischer Musik, hätte völlig anders reagiert, oder den Reiz gar ignoriert. Der Reiz ist notwendig, aber nicht hinreichend, um das Verhalten zu erklären. Viel wichtiger für die Erklärung dieses Verhaltens ist aber natürlich, was im handelnden Individuum vorgeht. Skinners Hauptthese ist, dass die operante Konditionierung im Prinzip hinreichend ist, um das Verhalten des Sprechers, der „Mozart“ äußert, zu verstehen. Sein aktueller kognitiver Zustand, der dem Verhalten zugrundeliegt, ist das Produkt seiner vorgängigen Erfahrungen, die sich in Form der operanten Konditionierung eingepägt haben.

Eine schwer überprüfbare Hypothese.¹¹ Schon an diesem Hypothesencharakter stößt sich Chomsky. Rückblickend muss man ihm natürlich vorhalten, dass das Meiste, was von ihm selbst oder von ihm angestoßen in der Spracherwerbsdebatte vorgebracht wurde, nicht weniger hypothetisch ist. Im Übrigen ist das Aufstellen von Hypothesen ja nicht per se anrühlich.

Chomsky meint offenbar, seinen Sprachbegriff, der vom traditionellen linguistischen Sprachbegriff gar nicht so weit weg ist, verteidigen zu müssen, obwohl dieser gar nicht in Frage gestellt ist: die Grammatikforschung untersucht das gesamte Repertoire einer Sprache, nicht bloß sprachliches Verhalten in spezifischen Situationen.¹² Aber natürlich geht es um mehr: Grammatikforschung soll ja gerade zur kognitionswissenschaftlichen Disziplin, also zu einem Teil der Psychologie und Humanbiologie, werden – dazu ist nun aber eine Auffassung von Psychologie nötig, die nicht das tatsächliche menschliche Verhalten in den Vordergrund rückt, sondern davon gerade abstrahiert. Denn Grammatiken beschreiben ja das Gesamtrepertoire der *möglichen* Ausdrücke einer Sprache.

Das sprachphilosophisch Problematische an dieser Ineinssetzung wurde von vielen Seiten angesprochen. Chomsky und seine Anhängerschaft, die weit über den Kreis der generativen Grammatik hinausgeht, haben solchen Argumenten gegenüber eine bemerkenswerte Resistenz erkennen lassen. Ein aktuelles Beispiel für einen Philosophen, der die Auseinandersetzung trotzdem nicht scheut, ist der Sprachphilosoph Michael Devitt, insbesondere sein Buch „Ignorance of Language“ (Devitt, 2006) ist hier zu erwähnen. Devitt plädiert darin für einen nicht-mentalistischen Grammatikbegriff. Sprachliche Ausdrücke sollten nach

¹¹ MacCorquodale (1970) wies darauf hin, dass Skinner in „Verbal Behavior“ zwar den Begriff „Hypothese“ vermeidet, aber sich des Hypothesencharakters seiner Ausführungen wohl bewusst war.

¹² Chomsky lehnte die Bezugnahme auf jegliche probabilistische Information und somit jegliche quantitative empirische Analyse als irrelevant ab. Die allermeisten seiner SchülerInnen folgten ihm darin und dominierten damit zugleich die Entwicklung der theoretischen Linguistik über Jahrzehnte. Dass dies der Disziplin sehr schlecht bekommen ist, beklagte beispielsweise auch der Chomskyschüler Tomas Wasow (2002).

seiner Auffassung als reale Gegenstände aufgefasst werden und eine Grammatik als Theorie über deren Struktur, Bildungsgesetze, usw. Dies kann (oder sollte sogar) den Bezug auf Kognitives beinhalten, aber es wäre unangebracht, Grammatik darauf zu reduzieren. Und zwar deshalb, weil die linguistische Befassung mit sprachlichen Ausdrücken und die psychologisch-biologische Befassung damit, wie Menschen sprachliche Ausdrücke hervorbringen und verstehen, zwei im Grundsatz verschiedene Fragestellungen sind, die man nicht verwechseln darf.¹³

Devitts gern verwendetes Beispiel ist der Bientanz, dessen „Grammatik“ schon seit geraumer Zeit gut verstanden ist; die Frage aber, wie Bienen ihn hervorbringen, liegt noch ziemlich im Dunkeln. Ob Bienen dabei die Grammatik des Bientanzes intern mental repräsentieren oder nicht, ist eine offene Frage, denn für die neurobiologische „Implementierung“ dieses Verhaltens lassen sich ganz unterschiedliche Wege, auch emergentistische, denken. An der Richtigkeit der „Grammatik“ des Bientanzes ändert die irgendwann zu erfolgende Klärung dieser Frage nichts. Genauso verhält es sich nach Devitt mit dem Chomskyschen Kompetenzbegriff. Ob es sprachliches Wissen in der Form der Chomskyschen Kompetenz gibt, ist eine offene Frage – denkbar sind durchaus Szenarien, die ohne eine angeborene Sprachkompetenz auskommen. Jedenfalls ist es vor allem eine *andere* Frage.

In Chomskys oben geschilderter Verkennung des Skinnerschen Wahrscheinlichkeitsbegriffs, und somit des verhaltenspsychologischen Blickwinkels überhaupt, zeigt sich bereits die Verwechslung dieser beiden voneinander im Prinzip unabhängigen Fragestellungen:

1. Welche Gestalt hat der sprachliche Ausdruck (und warum)?

¹³ Devitt spricht hier vom „respect constraint“: eine Grammatik sollte so formuliert sein, dass sie die Erkenntnisse der Kognitionsforschung respektiert, so wie umgekehrt eine sprachpsychologische Theorie Erkenntnissen der Grammatikforschung nicht zuwiderlaufen sollte.

2. Wie bringt das biologische System den sprachlichen Ausdruck hervor?

Devitt legt darüber hinaus dar, dass die linguistische Fragestellung (1) primär ist, insofern als sie unabhängig von der zweiten beantwortbar ist, während Fragestellung (2) die Lösung von Fragestellung (1) erkenntnistheoretisch voraussetzt.

1.3 Weitere Kritikpunkte Chomskys an Skinner

Ein weiteres von Chomsky angesprochenes vermeintliches Problem ist, dass nach der behavioristischen Erklärung zum Lernen auch eine *Verstärkung* gehört, also eine Belohnung für das richtige Verhalten. Es ist aber nicht immer klar, worin eine solche Belohnung bestehen soll. Skinner spekuliert beispielsweise, dass dem lernenden Kind das gelungene Nachsprechen eines Wortes selbst schon Belohnung genug ist. Neuere Erkenntnisse über die Funktionsweise der Spiegelneuronen und ihre zentrale Funktion für die kognitiv-emotionale Entwicklung in den frühesten Lebensphasen liefern hier eine sehr einfache Antwort, die das behavioristische Modell im Prinzip bestätigt: Säuglinge sind sehr darauf erpicht, ihre Bezugspersonen zu imitieren, soweit sie das eben schon können. Das beginnt praktisch sofort nach der Geburt in Bezug auf Gesichtsausdrücke. Dieses Verhalten geht sehr schnell auf lautliche und gestische Imitationsversuche über. Damit einher gehen Glücks- und Verzückungsreaktionen, die von einem Ausstoß körpereigener Opiode begleitet werden. Das sollte behavioristisch gesprochen „Belohnung“ genug sein.¹⁴

Das Bild, das Skinner vom Spracherwerb zeichnet, ist für Chomsky in mehrerlei Hinsicht verzerrt. Die wichtigsten Einwände sind die folgenden:

- Kinder lernen eine Sprache so oder so, ob sie dafür belohnt werden oder nicht: Zwischen den Lernerfolgen eines Akademiker-Kindes, dessen El-

¹⁴ Eine sehr gut lesbare Darstellung des aktuellen Forschungsstands zu den Spiegelneuronen findet sich in Bauer (2005).

tern sich alle Mühe geben, ihm die Sprache korrekt beizubringen, und etwa einem kleinen Einwanderer-Kind, das vor allem von Gleichaltrigen „auf der Straße“ seine neue Sprache erwirbt, besteht kaum ein Unterschied.

- Besonders rätselhaft muss dem Behaviorismus die Tatsache erscheinen, dass die Eltern des Einwandererkindes sich noch so sehr anstrengen können, sie werden mit ihrem Kind nicht mithalten können beim Erlernen der neuen Sprache.
- Hier zieht Chomsky einen Vergleich zur Biologie anderer Spezies, bei denen sich bestimmte kritische frühe Lebensphasen identifizieren lassen, in denen eine das ganze Leben anhaltende Prägung auf bestimmte Gegenstände oder andere Tiere stattfindet, beispielsweise bei Vögeln. Analog sieht er den Erwerb der Muttersprache als einen Prozess frühkindlicher Prägung, der aufgrund biologischer Voraussetzungen in einem bestimmten Zeitfenster stattfinden muss.
- Ein dritter für Chomsky entscheidender Aspekt betrifft die sprachliche *Kreativität*: Das Phänomen, wie es sein kann, dass wir nicht die geringsten Probleme damit haben, beispielsweise in der Tageszeitung, Sätze zu lesen und zu verstehen, die wir vorher noch nie gehört oder gelesen haben, und, gleichermaßen, neue Sätze zu produzieren. Die sprachliche Kreativität bleibt für eine behavioristische Erklärung nach Chomsky völlig außer Reichweite.

Keiner dieser Einwände ist mit Skinners Sicht unverträglich. Chomsky lässt diesen Eindruck nur deshalb entstehen, weil er ein Zerrbild des Behaviorismus zeichnet, das in seiner Primitivität dem Skinnerschen Ansatz nicht gerecht wird. Für eine detaillierte Entgegnung verweise ich einmal mehr auf MacCorquodale (1970). Dass es kritische Phasen in der Reifung eines Organismus gibt, in denen

der Erwerb eines Verhaltens besonders gut vonstatten geht, ist natürlich auch den Behavioristen bekannt, auch aus der biologischen Verhaltensforschung bei Tieren. Natürlich ist dies in den behavioristischen Modellen reflektiert. Die Produktivität der menschlichen Sprache wiederum steht nicht im Widerspruch zu Skinners Konzept der „operanten Konditionierung“, das die Aufmerksamkeit auf die menschliche Fähigkeit zur Manipulation seiner Umwelt gemäß seiner Intentionen lenkt. Beiden wohnt ein Moment der Kreativität inne.

2 Chomskys Gegenentwurf

Der Mechanismus der operanten Konditionierung kann als allgemeines Schema für höhere Lernprozesse im Behaviorismus angesehen werden. Wichtige Kennzeichen sind eine assoziationsistische Sicht des Lernens und die wichtige Rolle, die Wiederholung, also Frequenz und, damit zusammenhängend, Wahrscheinlichkeit spielen.

Mit den oben geschilderten Einwänden sind die wesentlichen Ausgangspunkte für Chomskys Gegenentwurf bereits beschrieben. In den *Aspects of the Theory of Syntax* (Chomsky, 1965) beschreibt Chomsky diese wie folgt (zitiert nach der deutschen Ausgabe, 1969, S. 41):

„[...] Auf der Grundlage [der primären sprachlichen Daten] konstruiert das Kind eine Grammatik, d.h. eine Theorie der Sprache, in der die wohlgeformten Sätze aus den primären sprachlichen Daten nur einen kleinen Ausschnitt bilden. Um eine Sprache zu erlernen, muss das Kind demnach über eine Methode verfügen, wie eine geeignete Grammatik bei gegebenen primären sprachlichen Daten herzustellen sei. Als Vorbedingung für die Spracherlernung muss es

1. eine linguistische Theorie besitzen, die die Form der Grammatik einer möglichen menschlichen Sprache spezifiziert,
2. eine Strategie ausbilden, um eine Grammatik der passenden Form auszuwählen, die sich mit den primären sprachlichen Daten verträgt.

Als weitgesteckte Aufgabe für die allgemeine Sprachwissenschaft können wir das Problem stellen, eine Explikation für diese angeborene linguistische Theorie, die die Grundlage der Spracherlernung bildet, zu liefern.“

Die erste Vorbedingung, die „linguistische Theorie“ des Kindes, wurde von Chomsky mit dem Terminus „*Universalgrammatik*“ (UG) belegt. Die zweite Vorbedingung ist der ‚*Language Acquisition Device*‘ (LAD), der angeborene Spracherwerbsmechanismus. Beides zusammen bildet die ‚*Faculty of language*‘, oder auch das ‚*Sprachorgan*‘. Für Chomsky muss die wissenschaftliche Linguistik insbesondere den Inhalt der UG zum Forschungsgegenstand haben.

Natürliche Sprachen an sich sind demgegenüber nur insofern interessant, als sie uns Aufschluss geben können über die UG, und eine Theorie über die Grammatik einer einzelnen Sprache kann nur dann Chomskys Wissenschaftskriterien entsprechen, wenn sie im Einklang mit den Erkenntnissen über die UG steht.

2.1 Entwicklungsgeschichte des Chomskyschen Modells

Das „Aspects“-Modell: Der „kleine Linguist“ Zu Anfang wurde das lernende Kind noch als „kleiner Linguist“ verstanden. Das sogenannte Standard-Modell, das in den *Aspects* 1965 seine Ausformulierung fand, sieht für die Syntax eine universale *Basis-Komponente* vor, die eine *Tiefenstruktur* erzeugt, aus der dann in einer einzelsprachlichen *Transformationskomponente* die *Oberflächenstrukturen* von Sätzen erzeugt werden. Peters & Ritchie (1973) haben

den formalen Beweis geführt, dass eine solche Grammatik-Architektur in ihrer Beschreibungs-Mächtigkeit weit über die Menge der natürlichen Sprachen hinausgeht. Als Konsequenz für das lernende Kind ergibt sich hieraus, dass der Hypothesenraum des „kleinen Linguisten“ derart hoffnungslos riesig ist, dass das Lernen der Grammatik unmöglich ist. Dieses Modell erfüllt also Chomskys wesentliche Anforderung, die Lernbarkeit, nicht.

Prinzipien und Parameter Das Prinzipien- und Parameter-Modell, das ab Mitte der Siebziger bis etwa zum Ende der Achtziger Jahre entwickelt wurde, zeichnet demgegenüber ein sehr viel passiveres Bild vom lernenden Individuum. Hier wird die UG als eine Menge von grammatischen Prinzipien verstanden, die nur noch der einzelsprachlichen Parametrisierung bedürfen. Ich möchte als Beispiel für ein solches parametrisiertes Prinzip die Bildung von Informations-Fragen vorstellen. Im Deutschen wird das Frage-Pronomen in solchen Sätzen vorangestellt, in vielen anderen Sprachen, wie beispielsweise dem Chinesischen, geschieht dies nicht.

- (1) a. **Wen** hast du besucht? (Deutsch)
b. Du hast **wen** besucht? (Chinesisch)

Parameter: Fragepronomen werden vorangestellt JA NEIN

Das Setzen des Parameters erfolgt durch Beobachtung von Informationsfragen im sprachlichen Input. Das Prinzipien- und Parameter-Modell legt für eine Vielzahl sehr konkreter syntaktischer Problemstellungen den Rahmen der möglichen Lösungen fest. Die Menge der für die UG postulierten Prinzipien ist dementsprechend groß geworden. Mit der Chomskyschen Opposition gegen behavioristische Lernkonzeptionen im Hinterkopf muss die Primitivität des Prinzipien- und Parameter-Modells irritieren. Der Erwerb der Parameterbelegung für grammatische Prinzipien in einer Sprache wird in der P&P-Literatur oft metaphorisch als das Umlegen von Schaltern in Reaktion auf sogenannte

„Trigger“ beschrieben. Dieser Mechanismus entspricht ziemlich gut dem Reiz-Reaktions-Schema der klassischen Konditionierung für reflexhaftes Verhalten, das Chomsky am behavioristischen Denken so kritisierte. Skinner war mit ihm in dieser Hinsicht aber völlig einer Meinung, weshalb er ja auch die operante Konditionierung in den Mittelpunkt rückte – was Chomsky allerdings wie oben dargelegt nicht verstand oder ignorierte. In der Prinzipien- und Parameter-Theorie wird der Erwerb einer Parameterbelegung nun aber gerade nach Art der klassischen Konditionierung gedacht, als Reflex auf einen auslösenden Reiz (den „Trigger“). Chomskys Spracherwerbsmodell fällt hier also eigentlich noch hinter Skinner zurück.

Das minimalistische Programm: UG als rekursiver Kombinationsmechanismus In den Arbeiten im Rahmen des minimalistischen Programms ist seit Anfang der Neunziger Jahre die Zielstellung formuliert worden, die Theorie der UG auf das absolut Notwendige zu beschränken. In ihrer radikalsten Variante (Hauser, Chomsky, Fitch 2002, Chomsky 2005) enthält sie nur noch einen speziellen syntaktischen Kombinationsmechanismus, dem im Wesentlichen die Eigenschaft der *Rekursion* zueigen sein soll. Alles andere liegt im Lexikon der zu lernenden Sprache.

Die Vorstellungen über den Inhalt des Sprachorgans, die Universalgrammatik, haben sich also im Laufe der Jahre grundlegend gewandelt: vom kleinen Linguisten, der eine Vielzahl von Hypothesen über die Struktur der Grammatik, die er lernt, durchspielt, und im Laufe des Spracherwerbs allmählich zur „richtigen“ Lösung kommt, über eine Art Betriebsprüfer, der mit einer angeborenen Checkliste ins Leben entlassen wird, die den Hypothesenraum erheblich einschränkt und auch den Weg, der zum Ziel führt, weitgehend vorgibt, hin zu einem Lerner, der über Sprachen nur so viel mitgegeben bekommt, dass sie einem bestimmten Kombinationsmechanismus unterliegen und aus Lexemen mit bestimmten Eigenschaften bestehen, und der sich mit diesem Wissen das Lexi-

kon seiner Muttersprache und damit auch schon diese Sprache selbst, erschließt.

Das Modell der UG wurde so immer weiter reduziert. Ihr Kern aber, der syntaktische Kombinationsmechanismus, der allen Äußerungen zugrunde liegen soll, ist erhalten geblieben.

3 Probleme des Chomskyschen Programms

Das generative Paradigma forderte immer wieder teils heftige Kritik heraus. Zwei mir besonders substantiell erscheinende Problemfelder möchte ich im Folgenden näher beleuchten. Zum Einen ist es die Frage, wie sich die Notwendigkeit einer angeborenen UG empirisch nachweisen lässt, und zum Zweiten, damit zusammenhängend, die Frage, wie sich eine *bestimmte* Formulierung der UG empirisch bestätigen lässt.

3.1 Empirische Verifizierbarkeit – das Argument von der Unzulänglichkeit des kindlichen Sprachinputs („poverty of the stimulus“)

Schon in Chomskys frühen Arbeiten tauchte die folgende Denkfigur auf, die man in verschiedener Form bis heute wiederfindet, das Argument von der *Unzulänglichkeit des Reizes (Poverty of the Stimulus)*. Es hat zentrale Bedeutung in der Argumentation für den linguistischen Nativismus:¹⁵

- (i) Alle Kinder erlernen ihre Muttersprache, auch wenn ihr sprachlicher Input unvollständig ist.
- (ii) Kinder können Strukturen produzieren und als grammatisch erkennen, die nicht Teil ihres Inputs waren.
- (iii) Dies ist nur möglich, weil Kinder bei ihrer Geburt schon über Hypothesen über die Grammatik ihrer Zielsprache verfügen.

¹⁵ Die Geschichte des „poverty of the stimulus“-Arguments fasst Thomas (2002) zusammen.

Ein Beispiel, das Chomsky selbst anführte (Chomsky 1981, 30-35; Piattelli-Palmarini 1980, 39f), ist die These, dass Kinder nicht aus dem Input alleine lernen können, dass grammatische Regeln *strukturabhängig* sind. Dieses „Wissen“ ist Resultat des in UG enthaltenen syntaktischen Kombinationsmechanismus.

Nehmen wir an, dass Entscheidungsfragen aus den dazugehörenden Deklarativ-Sätzen gebildet werden, also (2-b) aus (2-a):

- (2) a. Der Mann hat geraucht.
b. Hat der Mann — geraucht?

Die Bildungsregel für Entscheidungsfragen könnte nun auf zwei Weisen formuliert werden, einmal unter Bezug auf die Struktur, und einmal unter Bezug auf die lineare Abfolge:

- (3) a. Stelle das finite Verb des Hauptsatzes voran!
b. Stelle das erste finite Verb voran!

Dass die zweite Regel inadäquat ist, lässt sich erst anhand von Beispielen erkennen, in denen ein eingebettetes finites Verb dem Finitum des Hauptsatzes vorangeht, wie das Finitum des Relativsatzes, der sich in (4) dem Subjekt des Satzes anschließt:

- (4) a. Hat der Mann, dem du böse bist, — geraucht? (Regel (3-a))
b. *Bist der Mann, dem du böse — , hat geraucht? (Regel (3-b))

Nach Regel (3-a) ist (4-a) korrekt gebildet, nach Regel (3-b) das nicht akzeptable Beispiel (4-b).

Daten vom Typ (4-a) sind nun, nach Chomsky, nicht Teil des kindlichen Inputs oder jedenfalls viel zu selten, um dem Kind bei der Entscheidung zwischen den beiden Regel-Alternativen zu helfen. Zu erwarten wäre dann also, keine Angeborenheit vorausgesetzt, dass Kinder öfter im frühen Stadium Strukturen wie

(4-b) produzieren oder als wohlgeformt akzeptieren. Dies ist nicht der Fall. Das konnte auch experimentell bestätigt werden (Crain und Nakayama 1987, Crain 1991).

Diese Argumentation geht implizit von folgenden Prämissen aus:

- Strukturabhängige Regeln können nicht aus dem Input, beispielsweise der stochastischen Verteilung von Wörtern und Wortsequenzen, erschlossen werden.
- Kinder machen keine Generalisierungen irgendwelcher Art über den Input, die auf solcher statistischen Information basierten.
- Kinder schließen nicht von anderen, im Korpus vorkommenden Strukturen auf die Struktur eines neuen Satztyps.

Sätze, die in nennenswerter Zahl im kindlichen Korpus vorkommen und die dieselbe Vorstellungsregel verwenden, sind beispielsweise Informationsfragen:¹⁶

(5) Was hat der Mann, dem du böse bist, — gesagt?

Lewis & Ellman (2001) zeigten in einer Simulation mit einem konnektionistischen Modell nicht nur, dass es möglich ist, die richtige Generalisierung für Strukturen dieses Typs alleine aus der stochastischen Information in einem Korpus zu lernen, sondern auch, ohne dass die zu testenden Sätze, also Entscheidungsfragen, bei denen sich ein Relativsatz an das Subjekt anschließt, im Trainings-Korpus vorkommen.

¹⁶ Im Unterschied zum Englischen, worauf die Diskussion sich bezieht, sind im Deutschen hier generell Vorstellungsstrukturen relevant:

(i) Morgen will der Mann, dem du böse bist, — uns besuchen

Das Experiment bestand aus einer Trainings- und einer Testphase. In der Trainingsphase wurde das Netzwerk unter anderem mit Sätzen der Art (6-a-f) konfrontiert, nicht aber mit den zu testenden Strukturen (6-g):

- (6) Trainingskorpus: (a-f), Teststruktur: (g).
- a. The little boy bites
 - b. The boy who is smiling smokes
 - c. Does the boy smoke?
 - d. Is the little girl happy?
 - e. Where did the man go?
 - f. Who is the boy who is smoking?
 - g. #Is the boy who is smoking silly?

Es wurde dabei darauf geachtet, dass das Trainingskorpus annähernd realistisch ist, dass also die relative Häufigkeit der Trainingssätze den tatsächlichen Verhältnissen entspricht, wie man sie anhand der CHILDES-Datenbank, einer großen Datenbank mit kindlicher und an Kinder gerichteter Sprache, errechnen kann. Die Testsätze wurden dem Netzwerk wortweise als Input eingegeben, wobei das Netzwerk seine Vorhersage über das folgende Wort als Output ausgab. Eine Korrekturfunktion gleicht den Output mit dem nächstfolgenden Input ab, die Diskrepanz geht ein in die numerische Adaption der Gewichte des Netzwerks. Über die Zeit bildet sich so im Netzwerk eine interne Repräsentation über die stochastische Verteilung der Wörter und ihre Übergangswahrscheinlichkeiten.

In der Testphase wurde das Netzwerk dann mit den Testsätzen der Form (6-g) konfrontiert. Chomskys Vorhersage, dass ein so trainiertes Netzwerk des öfteren Sätze vom Typ

- (7) *Is the boy who — smoking is silly?

produzieren würde, konnte klar widerlegt werden. Mehrheitlich wurden die korrekten Sätze vorhergesagt.

Das Netzwerk wurde in drei aufeinander aufbauenden Zyklen von Trainings- und Test-Phase getestet, um Stadien des Spracherwerbs zu simulieren. Es zeigte sich eine zunehmende Akkuratheit von Zyklus zu Zyklus.

Das Netzwerk hat allerdings mit gewisser Hartnäckigkeit eine andere Art von fehlerhafter Struktur vorhergesagt, nämlich eine Wiederholung des vorangestellten Verbs in seiner Ausgangsposition:

(8) Is the boy who is smoking is . . .

Fehler dieser Art werden interessanterweise auch von Kindern gemacht (Crain & Thornton, 1998).

Man kann aus der Studie von Lewis und Elman meines Erachtens Folgendes schließen:

1. Die in einem Korpus gegebene stochastische Information ermöglicht korrekte Generalisierungen über syntaktische Strukturen.
2. Aus solchen Generalisierungen lässt sich die Struktur neuer, im Korpus nicht enthaltener Satztypen korrekt ableiten.

Es gibt außerdem womöglich noch einen anderen Grund dafür, warum Sätze wie (7) nicht vorkommen: Der Relativsatz in diesem Beispiel verstößt gegen eine deskriptive Generalisierung, für deren Korrektheit praktisch jeder Satz im sprachlichen Input des Kindes Evidenz liefert:

- (9) a. Jeder Satz enthält höchstens ein finites Verb.
b. Kein finites Verb kann den Satz verlassen, zu dem es gehört.

Wieso sollte ein neuer Satztyp gegen diese Generalisierungen verstoßen, wenn kein anderer der bislang gelernten Satztypen dies tut?

Es ist meines Erachtens ein großer Makel des Chomskyschen Ansatzes, dass alternative Formulierungsmöglichkeiten wie diese offensichtliche Generalisierung zumeist völlig außer Betracht bleiben.

Wir lernen aus dieser Studie, dass ein syntaktisches Phänomen, an dem sich die Notwendigkeit der Angeborenheit zeigen lässt, hinreichend „exotisch“ sein muss, so dass sich seine strukturellen Eigenschaften nicht, oder zumindest nicht vollständig aus denen anderer bekannter Strukturen ableiten lassen. Es ist zumindest fraglich, ob Phänomene dieser Art überhaupt existieren.

Man könnte das Problem auch mit folgender Paradoxie beschreiben:

- (10) Das ideale universalgrammatische Prinzip wirkt so sehr im Verborgenen, dass sein Wirken sich nicht auf die statistischen Eigenschaften eines Korpus auswirkt.
- Das würde allerdings bedeuten, dass das Prinzip keinen empirischen Effekt hätte, und
 - dass es folglich nicht einmal von Linguisten gefunden werden könnte.

(10) beschreibt eine epistemologische Unmöglichkeit. Es handelt sich um eine *reductio ad absurdum*, die zeigt, wie sehr der Nativismus und der Chomskysche Kompetenzbegriff von der Ablehnung statistischer Verfahren abhängen. Man muss hier die Formulierung „keinen empirischen Effekt“ natürlich im Sinne der vorangegangenen Diskussion verstehen. Es genügt nicht, zu zeigen, dass eine bestimmte Struktur zu selten vorkommt, als dass sich ihre strukturellen Eigenschaften direkt erschließen ließen. Zu suchen ist vielmehr nach einer Struktur, die über derart exotische Eigenschaften verfügt, dass es unmöglich ist, diese aus den Eigenschaften häufiger vorkommender Strukturen zu erschließen. Die universalgrammatische Hypothese wäre, dass lernende Kinder mit einer solchen Struktur keine oder geringe Lernschwierigkeiten haben, weil ihnen ein angebo-

renes UG-Prinzip hilft, während ein statistisches Lernmodell für einen solchen Fall große Lernschwierigkeiten prognostiziert.

Was wäre ein denkbarer möglicher Fall? Zuerst könnte man an Phraseologismen denken. Aus dem Süddeutschen ist beispielsweise eine Wendung in den Sprachgebrauch übergegangen, in der wir eine Verb-Stellung vorfinden, die im Standard-Deutschen allgemein nicht akzeptabel ist:

(11) Ich habs/er,sie hats/... wollen wissen

Normalerweise wäre nur die Abfolge „wollen wollen“ möglich. Will man annehmen, dass „wollen wissen“ Teil des Standard-Deutschen ist, so könnte man postulieren, dass diese Wendung aus UG-Sicht leicht lernbar ist, weil das Kind um die Parametrisiertheit eines Wortstellungs-Prinzips weiß und das Vorkommen alternativer Abfolgen grundsätzlich mit einbezieht. Für den statistischen Lerner muss diese Struktur einfach als mit den sonstigen Regeln nicht zu vereinbarende Ausnahme behandelt werden. Nun sind solche Ausnahmen auf lexikalischer Ebene allerdings ohnehin nichts ungewöhnliches. Und so könnte man Lernmechanismen, die auf lexikalischer Ebene gelten, einfach auf die syntaktische Ebene übertragen, beispielsweise im Sinne der Konstruktionsgrammatik.

Nun ist aber bereits aus der Spracherwerbsforschung bekannt, dass der Erwerb seltener vorkommender Strukturen schwerer fällt als der Erwerb fundamentaler allgegenwärtiger Strukturfestlegungen einer Sprache. Auch UG-Verfechter bestreiten natürlich nicht die Relevanz des sprachlichen Inputs für den konkreten Verlauf des Erwerbsprozesses.

UG-Verfechter könnten sich auf den Standpunkt zurückziehen, dass eine deskriptive Generalisierung wie in (10) die Kategorien „Satz“ und „finites Verb“ voraussetzt, und somit auch deren Lernbarkeit aus dem sprachlichen Input. Man könnte nun kritisch fragen, ob dies legitim ist. Ist es überhaupt möglich, solche Kategorien aus dem Input zu lernen? Die Antwort muss hier lauten: Diese

Möglichkeit muss von beiden Ansätzen vorausgesetzt werden. Auch in der nativistischen Sicht muss das Kind ja in der Lage sein, die Strukturen und Kategorien, von denen es qua biologischer Ausstattung nur auf sehr abstrakte Weise „weiß“, in der Muttersprache, egal um welche es sich dabei handelt, wiederzufinden. Das kann gar nicht anders gehen als über die Interpretation der statistischen Information im muttersprachlichen Input – in derselben Weise wie der statistische Lerner dies tut.

Solchermaßen gewonnene „Beobachtungskategorien“ wie im Englischen, beispielsweise,

- (12) a. Wörter, wo häufig ein „the“ oder ein „a“ unmittelbar davor steht (Nomen)
- b. Wörter, wo häufig „-ed“ hinten dran hängt (reguläre Verben)

müssen im UG-Modell mit dem angeborenen Wissen abgeglichen werden. Der statistische Lerner braucht demgegenüber nicht über diese Art der Kategorisierung hinauszugehen. Universalität liegt hier lediglich in den allgemeinen Voraussetzungen des Lernprozesses. Dass die Sprachen der Welt bestimmte Charakteristika teilen, kann dann aber immer noch auf andere Faktoren zurückgeführt werden: die gemeinsamen allgemeinen kognitiv-biologischen Voraussetzungen, Gemeinsamkeiten der sozialen Umwelt, in der Spracherwerb stattfindet sowie allgemeine Gesetze des Sprachwandels.

Das UG-Modell ist also das komplexere, und steht deshalb nach wie vor in der *empirischen Bringschuld*, Nutzen und Notwendigkeit eines zusätzlichen Schrittes im Spracherwerbsprozess – Abgleich der aus der Beobachtung gewonnenen Generalisierungen und Kategorisierungen mit angeborenem sprachlichen Wissen – zu begründen. Lewis und Elman haben gezeigt, wie schwierig diese Aufgabe ist. Die nativistische Theorie des Spracherwerbs unterschätzt die Möglichkeiten statistischen Lernens in relativ drastischer Weise. Einer der

Hauptvorwürfe von Pullum & Scholz (2002) in ihrem Target-article zu einem Sonderheft der Zeitschrift „The Linguistic Review“ besteht denn auch darin, dass es nur eine sehr bescheidene empirische Forschung zu den Grenzen statistischen Lernens im Sinne der UG gibt.

Statistische Lernmodelle, und besonders konnektionistische wie das eben besprochene, sind der behavioristischen Denkweise artverwandt, weil beide auf die Tradition des Assoziationismus zurückgehen.¹⁷ In beiden Modellen beruht Lernen darauf, dass Assoziationen zwischen einem Input und einem Output durch Wiederholung verstärkt werden.¹⁸

3.2 Die Unbestimmtheit der Übersetzung

Der Philosoph W.V.O. Quine, ein konsequenter Vertreter des Behaviorismus, zählte zu den vehementesten Kritikern von Chomskys mentalistischer Konzep-

¹⁷ Hier sind vor allem David Hume zu nennen und der britische philosophische Empirismus des 18. Jahrhunderts.

¹⁸ Bei aller Skepsis gegenüber der Übertragung von Ergebnissen aus der KI-Forschung auf Kognitionspsychologie kann man vielleicht konnektionistische Modelle als mit dem Behaviorismus verträgliche Theorie der „Black box“ ansehen. Der Philosoph Daniel Dennett bringt diese verschiedenen Stränge derselben Denktradition unter dem Label „ABC Learning“ (associationism-behaviorism-connectionism) zusammen (siehe z.B. Dennett, 1996). In dem Zusammenhang ist auch aufschlussreich, dass mit der *Optimalitätstheorie* das gegenwärtig wichtigste Konkurrenzprojekt zur Chomskyschen Grammatikkonzeption innerhalb der generativen Grammatik aus konnektionistischen Versuchen der Grammatik-Modellierung hervorging (siehe Prince & Smolensky, 1993, 2004; Smolensky & Legendre, 2006a,b). Zur Rolle konnektionistischer Modellierung in der Grammatikforschung siehe auch Meyer (in diesem Band). Die Geschichte des Konnektionismus hat übrigens gewisse Parallelen zur Behaviorismus-Debatte. So, wie die Rezeption von Chomskys Skinner-Rezension das Ansehen des behavioristischen Paradigmas erheblich beschädigte, wurde der konnektionistische Ansatz in der KI-Forschung durch das Buch „Perceptrons“ (Minsky & Papert, 1969) für fast 20 Jahre ausgebremst. Eine ausführliche Darstellung dieser Geschichte gibt Olazaran (1996). In beiden Fällen muss die Argumentation im Nachhinein als viel weniger überzeugend eingeschätzt werden als sie Zeitgenossen erschienen haben muss. Die „Attacken“ gingen in beiden Fällen von Wissenschaftlern des MIT aus. Ihr Erfolg resultiert vermutlich aus der Allianz eines anti-assoziationistischen Zeitgeistes mit den ökonomischen Interessen einflussreicher Forschungsinstitutionen.

tion der Grammatik. Sein wichtigster Einwand ist, dass es unmöglich ist, eine bestimmte Formulierung der (angeborenen) Universalgrammatik als wahr auszuzeichnen, weil es keine Fakten gibt, anhand deren diese Entscheidung getroffen werden könnte.

Sein Ausgangspunkt ist das Argument von der *Unbestimmtheit der Übersetzung*, das er vor allem in seiner Schrift „*Word and Object*“ (Quine, 1960) entwickelte. Es beschreibt anhand des Gedankenexperiments der „radikalen Übersetzung“ die These, dass es viele Möglichkeiten gibt, eine Sprache angemessen zu übersetzen, und dass es keine Möglichkeit gibt, *anhand von Fakten* zwischen zwei verschiedenen solcher Übersetzungen zu entscheiden.

Quine verwendet das Beispiel eines Feldlinguisten, der die Sprache eines ihm völlig unbekanntes Dschungelvolkes beschreiben will. Da ihm auch die Kultur des Volkes fremd ist, hat er kaum externe Anhaltspunkte, um die Bedeutung der von ihnen verwendeten Wörter zu erschließen.

Sagt ein Dschungelbewohner zu ihm das Wort „*Gavagai*“, während er auf ein Kaninchen zeigt, kann dies ‚Kaninchen‘ bedeuten, genauso gut aber auch ‚Kaninchenheit‘, ‚Kaninchenteil‘ usw. Der Linguist wird sich zunächst behelfsweise auf eine der Möglichkeiten festlegen und mit seiner Übersetzungstätigkeit fortfahren. Dabei wird er auf andere Wörter stoßen, die ähnliche Unbestimmtheiten mit sich bringen.

Merkt er nun, dass er eine seiner Festlegungen korrigieren muss, dann bieten sich ihm oft mehrere Möglichkeiten. Stellen wir uns, in Erweiterung des Quineschen Gedankenexperiments, beispielsweise vor, dass unser Linguist die Äußerung „*Gavagai koko*“ mit „Dies ist ein Kaninchenteil“ übersetzt, dann gibt es für die Wort-für-Wort-Übersetzung zwei Möglichkeiten, die dann wiederum durch die Verwendung dieser Wörter in anderen Kombinationen die Übersetzung anderer Wörter beeinflussen:

- (13) *gavagai* *koko*
1. Möglichkeit: Kaninchenteil ist
2. Möglichkeit: Kaninchen ist-Teil-von
 „Dies ist ein Kaninchenteil“

Der eine Feldlinguist mag die eine Möglichkeit wählen, ein zweiter die andere. Im Ergebnis entstehen dabei verschiedene und *einander widersprechende* „Übersetzungshandbücher“ der Dschungelsprache, die sich trotzdem in ihrem Ergebnis nicht unterscheiden. Zwischen ihnen zu entscheiden, welche die ‚wahre‘ ist, ist, so Quine, unmöglich, da kein objektives Kriterium für eine solche Entscheidung existiert.

Dieses Argument ist auch auf die Grammatikforschung übertragbar: die Frage nach der richtigen Beschreibung des grammatischen Regelsystems einer Sprache ist ebenfalls keine Faktenfrage. Quines (1970) Beispiel in seinem Text „Methodological Remarks on Current Linguistic Theory“ sind zwei dänische Schulklassen, die auf verschiedene Weise Englisch lernen, einmal mit Otto Jespersens Grammatik des Englischen und die andere Klasse mit einer am Lateinischen orientierten Schulgrammatik. Nehmen wir an, dass die Schüler beider Klassen hinterher gleich gut Englisch sprechen. Es gibt nun keine Beobachtungstatsachen, die uns helfen könnten zu entscheiden, mit welcher Grammatik ein bestimmter Schüler Englisch gelernt hat. Darüber hinaus passen beide Grammatiken gleich gut auf das Englisch eines Muttersprachlers, aber wir wissen, dass dieser die Sprache weder mit der einen, noch mit der anderen Grammatik gelernt hat. Insofern, als es sich bei der UG um eine Grammatik im eigentlichen Sinne handelt, unterliegt sie derselben Unbestimmtheit.

Dass dieses Argument kein reines Gedankenspiel ist, lässt sich an einem aktuellen Beispiel illustrieren. Es gibt eine Kontroverse unter deutschen Linguisten darüber, ob das Deutsche im Dativ markierte Subjekte besitzt oder nicht. Ausgangspunkt ist das Isländische, das Dativ-Subjekte besitzt. Für das Verb ‚*lika*‘, das ein Dativ-Subjekt und ein Nominativ-Objekt hat, ergeben sich zwei

Übersetzungsmöglichkeiten. Nimmt man an, dass das Deutsche Dativ-Subjekte haben kann, dann wird man den Kasus-Rahmen konstant halten und das Verb ‚*gefallen*‘ vorziehen. Nimmt man an, dass im Deutschen Subjekte im Nominativ stehen, wird man ‚*mögen*‘ als angemessene Übersetzung ansehen, um wenigstens die grammatischen Funktionen Subjekt und Objekt in der Übersetzung beizubehalten. Dies ist in (14) illustriert.¹⁹

- (14) *Maríu líkar thessa bók*
 Maria-DAT gefällt dieses Buch-NOM **Deutsch mit Dativ-Subjekt**
 Maria-DAT mag dieses Buch-NOM **Deutsch ohne Dativ-Subjekt**
 „Maria gefällt/mag dieses Buch“

Wir haben hier zwei konkurrierende Beschreibungen der Grammatik des Deutschen, die sich in ihrer Auffassung über die grammatische Kategorie des Subjekts unterscheiden. Will man zwischen ihnen entscheiden, müsste es ein objektives Kriterium dafür geben, welcher Subjekt-Begriff für das Deutsche der richtige ist.

Ein solches Kriterium steht uns jedoch nicht zur Verfügung. ‚SUBJEKT‘ ist eine Kategorie, anhand deren Linguisten Konsituenten mit bestimmten Eigenschaften klassifizieren. Wir sind aber darin völlig frei, wie wir diese Klassifizierung vornehmen. Die Subjekt-Eigenschaft *an sich* lässt sich nicht beobachten. Ob wir eine bestimmte *andere* Eigenschaft als Subjekt-Eigenschaft gelten lassen oder nicht, ist alleine unsere Entscheidung. Selbst wenn wir Beobachtungen machen, die mit unseren Entscheidungen unverträglich sind, gibt es nicht nur eine Möglichkeit der Korrektur.

Man könnte beispielsweise meinen, dass die Tatsache, dass Satz (15) möglich ist, dafür spricht, dass ‚*es*‘ hier Subjekt ist, und nicht Objekt:

¹⁹ Für verschiedene Positionen in dieser Frage, siehe bspw. Fanselow (2000); Jónsson (2003); Wunderlich (2003).

(15) Es gefällt mir

Wenn ‚es‘ nämlich als Objekt fungiert, kann es nicht am Anfang des Satzes stehen:

(16) *Es mag ich

Wer die Annahme, dass ‚*gefallen*‘ ein Dativ-Subjekt hat, retten möchte, hat allerdings die Möglichkeit, darauf zu verweisen, dass ‚es‘ in (15) im Nominativ steht und in (16) im Akkusativ, und kann dann postulieren, dass dies ausreicht, um die satzinitiale Position zu legitimieren.

Schwerer wiegt das Problem im Hinblick auf die Theorie der Universalgrammatik. Nehmen wir an, dass die Kategorie ‚SUBJEKT‘ Teil der Universalgrammatik ist, dann darf es nur eine *wahre* Definition dieser Kategorie geben, und es kann dann auch nur *eine* der beiden Möglichkeiten, für das Deutsche ‚SUBJEKT‘ zu definieren, wahr sein.

Wie sollte man nun die wahre Definition für ‚SUBJEKT‘ herausfinden? Man kann sich ja immer nur auf konkret beobachtbare Eigenschaften der fraglichen Elemente beziehen, beispielsweise Kasus, Wortstellungspräferenzen und ähnliches. Wenn diese widersprüchlich sind, wie in unserem Fall (der Dativ ist ein typischer Objekt-Kasus, die satzinitiale Stellung im unmarkierten Fall ist eine typische Subjekteigenschaft), hilft uns das aber nicht weiter.

Wir entscheiden uns dann dafür, einer bestimmten Eigenschaft Priorität einzuräumen, und diese Entscheidung selbst ist nun wieder an andere nicht unabhängige zu rechtfertigende Entscheidungen auf der Grundlage von anderen Beobachtungen gebunden. Es ist unmöglich, der Unbestimmtheit der Übersetzung zu entkommen, da es unmöglich ist, empirische Kriterien anzugeben, anhand deren sich die wahre Definition der universalen Kategorie ‚SUBJEKT‘ herausfinden ließe, die Teil unserer biologischen Ausstattung sein soll. Uns fehlen schlicht die Fakten, anhand deren sich solche Entscheidungen treffen ließen.

Quines Ablehnung des Mentalismus liegt genau hierin begründet: Wer mentale Repräsentationen zum Untersuchungsgegenstand erhebt, muss ihnen Existenz zusprechen und davon ausgehen, dass diese eine bestimmte aufdeckbare Gestalt haben. Wegen der Unbestimmtheit der Übersetzung lässt sich aber keine Hypothese über die konkrete Gestalt einer mentalen Repräsentation empirisch belegen. Konsequenterweise kann man dann auch nicht mentale Repräsentationen zum Gegenstand eines wissenschaftlichen Forschungsprogramms machen. Quine liefert also eine philosophische Rechtfertigung dafür, dass in der behavioristischen Psychologie das Mentale als „black box“ behandelt wird. Sie ist einer strikten empirischen Forschung nicht zugänglich.

Freilich muss eine Beschreibung des Mentalen mit den Beobachtungen über menschliches Verhalten verträglich sein, aber eine solche Beschreibung lässt sich eben auf verschiedene Weisen bewerkstelligen.

Ich habe bereits angedeutet, dass Chomsky viele Behauptungen über in der UG enthaltene Prinzipien inzwischen zurückgenommen hat, und im Grunde nur noch einen allgemeinen rekursiven Kombinationsmechanismus in der UG verortet.

Auch für diesen Kombinationsmechanismus gibt es natürlich mehrere Definitionsmöglichkeiten und wir haben keinerlei Entscheidungs-Kriterien zur Verfügung, wenn es sich um empirisch gleichwertige aber einander ausschließende Definitionen handelt. Keine konkrete Theorie über den angeborenen Kombinationsmechanismus kann also für sich in Anspruch nehmen, wahr zu sein.

Während Chomsky die Gültigkeit von Quines Argument für den Vergleich zweier Sprachen akzeptiert, erkennt er seine Relevanz für die Grammatikforschung nicht an (Chomsky, 1968). Indem Chomsky andererseits ausdrücklich auf *biologische* Evidenz für die Universalgrammatik verweist, die es irgendwann einmal geben soll, scheint er Quines Einwand unausgesprochen sehr wohl anzuerkennen. Naturwissenschaftliche Erkenntnisse wären ja auch nach Quine Fakten, die eine Entscheidung zwischen Theorien ermöglichen. Quine schließt

die Möglichkeit solcher Evidenz allerdings prinzipiell aus.

Chomskys Hypothese eines menschlichen Sprachorgans macht die Grammatikforschung zu einer Teildisziplin der Biologie. Wichtigstes Argument für das Sprachorgan ist die ihm zugesprochene Rolle im Spracherwerb.

Wir haben jedoch gesehen, dass das wichtigste Argument für Chomskys Theorie des Spracherwerbs, die Unzulänglichkeit des kindlichen Sprach-Inputs, schwer oder möglicherweise gar nicht empirisch zu belegen ist.

Jüngste Arbeiten Chomskys, beispielsweise ein Beitrag in der Zeitschrift *Science* zusammen mit den Biologen Marc Hauser und Tecumseh Fitch (2002) deuten denn auch ein Forschungsprogramm an, das zum Ziel hat, nachzuweisen, wie der postulierte rekursive Kombinationsmechanismus der Sprache evolutionär entstanden ist, und dass sich genau darin der Mensch von anderen Spezies biologisch unterscheidet. Dieses Forschungsprogramm ist im Grunde nicht mehr linguistisch: es untersucht die kognitiven Fähigkeiten von Primaten, das Genom von Menschen und Primaten und ihre Neurophysiologie, und unterliegt auch nicht mehr unbedingt dem Quineschen Problem der Unbestimmtheit.

Mit der Verkürzung des Forschungsgegenstands der Grammatikforschung auf den syntaktischen Kombinationsmechanismus wird dabei aber sovieles an relevanten Phänomenen ausgeblendet, dass man die Frage stellen muss, ob dieses sehr enge biologische Dach noch die Grammatikforschung vollständig abdeckt. Ich denke, dies ist eher nicht der Fall.

4 Schluss

Ich möchte abschließend eine andere Möglichkeit der Begründung linguistischer Universalien beleuchten, eine Begründung über ontologische Voraussetzungen der Sprache. Hier spielt auch die Biologie eine Rolle, aber eben nur als *Voraussetzung* für Grammatiken.

Zu den ontologischen Determinanten natürlicher Sprachen gehört sicher-

lich, dass sie in Raum und Zeit stattfinden, von Menschen in automatisierter Weise gesprochen und verstanden werden, und von unserer Welt handeln.

Welche Eigenschaften muss eine Sprache haben, damit sie von Menschen unbewusst in hoher Geschwindigkeit *automatisiert* gesprochen und verstanden werden kann? Meines Erachtens ist die wichtigste Eigenschaft die *Regelmäßigkeit*, und zwar sowohl im *Gebrauch* der elementaren Einheiten (Phoneme, Morpheme, Lexeme) wie ihrer *Anordnung*. Ein Verhaltensmuster kann nur dann automatisiert werden, wenn es im wesentlichen invariant abläuft, also regelhaft ist.

Chomsky (1965) erklärte zum Untersuchungsgegenstand der formalen Linguistik einen „idealen Sprecher/Hörer“, dem unendlich viel Zeit und unendlich viel Speicherplatz zur Verfügung steht, um Sätze korrekt zu bilden und zu verstehen. Der Verweis auf die unendlichen Ressourcen ist in dieser Idealisierung vor allem dem Problem geschuldet, dass wir nicht-idealen Sprecher hin und wieder Fehler produzieren. Performanz in diesem Sinne ist also einfach potentiell fehlerbehaftetes Sprachverhalten, da es in Echtzeit mit beschränkten Ressourcen zustande kommt. Dem möchte ich hier entgegen halten, dass ein solcherart idealisierter Sprecher grammatischer Regeln gar nicht mehr bedarf. Im Gegenteil müssen wir davon ausgehen, dass Grammatiken im Laufe der diachronen Entwicklung im Hinblick darauf optimiert wurden, dass mit ihrer Hilfe unter beschränktem kognitiven Ressourceneinsatz in kurzer Zeit sprachlicher Austausch gelingt. Wenn es also einen inneren Zusammenhang zwischen Regelmäßigkeit und knappen Ressourcen gibt, dann kann ein Abstrahieren von den Auslösern der „Verregelung der Sprache“ keine Erkenntnisse über Sprache befördern. Vielmehr muss es sie verschleiern. Anders gesagt: Chomskys idealer Sprecher/Hörer spricht keine natürliche Sprache.²⁰

²⁰ Chomsky verteidigte in der Diskussion des Aspekte-Modells seinen idealen Sprecher/Hörer gegen Kritik mit dem Argument, dass Idealisierung in der Wissenschaft notwendig ist. Dem kann man natürlich nur zustimmen; trotzdem muss die Frage erlaubt sein, ob diese bestimmte Idealisierung hilfreich ist oder nicht.

Als Konsequenz aus der Forderung nach regelhaftem Gebrauch der lexikalischen Einheiten ergibt sich weiter eine Forderung an die *Größe des Lexikons*. Wörter dürfen nicht willkürlich für unterschiedliche Zwecke verwendet werden. Da die Menge der Sachverhalte, über die wir sprechen wollen, potentiell unendlich groß ist, präferieren wir ein eher großes Lexikon, dessen *Obergrenze* allerdings durch die Kapazitäten unseres *Gedächtnisses* bestimmt ist. Zur Automatisierbarkeit gehört natürlich auch, dass die Struktur unseres Gehirns gewisse Grenzen dafür setzt, wie komplex ein zu automatisierendes Verhalten sein darf.

Die *Artikulierbarkeit* einer Sprache verlangt eine *Sequenzierung* des sprachlichen Materials. Unser Artikulationsorgan ist so konstruiert, dass wir immer nur ein Wort auf einmal sprechen können. Aus der Sequenzierung ergibt sich im Verbund mit der Forderung nach Regelmäßigkeit die prominente Rolle von Wortstellungsregeln, von Präzedenz- und Adjazenz-Forderungen für die Grammatik natürlicher Sprachen.

Die wohl universale Aufspaltung des Lexikons zumindest in die *morpho-syntaktischen* Kategorien *Verb* und *Nomen* lässt sich möglicherweise auf die semantische Voraussetzung zurückführen, dass wir in Prädikat-Argument-Strukturen Bedeutung kodieren.

Inwiefern ist auch die syntaktische Rekursivität Konsequenz solcher ontologischer Voraussetzungen? Hier lohnt es sich, genauer hinzuschauen. Die allgemeine Rekursivität ist im Grunde nämlich eher eine Eigenschaft des Formalismus, den wir verwenden, um Sprache zu beschreiben, oder allgemein unseres Denkens, als Eigenschaft der Sprache selbst. Exzessive Rekursion führt zu einer unverständlichen Äußerung. Das folgende Beispiel ist aus dem Buch „Mathematical Methods in Linguistics“ (Partee et al., 1990, 477f) entnommen:

- (17) a. The cat died.
b. The cat the dog chased died.
c. The cat the dog the rat bit chased died.

- d. The cat the dog the rat the elephant admired bit chased died.

Spätestens mit (17-d) wird das Verständnis unmöglich. Die Autoren klassifizieren solche Sätze gleichwohl als „grammatical [...] although incomprehensible“. Dem liegt eine Charakterisierung der grammatischen Regeln zugrunde, die vom realen Sprecher abstrahiert. Wenn es zutrifft, dass die Regelmäßigkeit der Sprache selbst nur vor dem Hintergrund ihrer ontologischen Voraussetzungen, hier: dem Zwang zur Automatisierung, zu verstehen ist, ist eine solche Abstraktion illegitim.

Anstatt also ein abstraktes Konzept von Rekursivität hochzuhalten, das empirisch leer läuft, wäre es aus dieser Sicht eher angezeigt, die Beschränkungen, die den Rekursionsmöglichkeiten auferliegen, bei der Beschreibung von Grammatiken in den Blick zu nehmen.²¹

Schon diese skizzenhaften Überlegungen zeigen, dass die Bestimmung substantieller linguistischer Universalien auch ohne nativistische Annahmen möglich ist.

5 Schlussbemerkung

Mein Fazit zu fünfzig Jahren Streit um den naturwissenschaftlichen Status der Grammatikforschung fällt eher skeptisch aus. Die Verortung der Linguistik in der Biologie steht und fällt mit dem Nachweis des Chomskyschen Sprachorgans. Dieser Nachweis ist erstens mit biologischen Methoden zu führen und zweitens ist es darum eher schlecht bestellt. Dass der Nachweis die Annahmen und Methoden der Grammatikforschung nennenswert verändern würde, ist eher nicht zu erwarten. Es ist eben eine *andere* Fragestellung.

Aktuelle Tendenzen auch der generativ inspirierten Grammatikforschung räumen der empirischen Erforschung der Sprache einen wesentlich größeren

²¹ Sternefeld (2000) kam bereits zu ähnlichen Schlussfolgerungen, aufbauend auf Kolb (1997).

Raum ein. Damit einher geht auch eine eher unbewusst vollzogene Reformulierung des Sprachbegriffs, die selten eingestanden wird. Mit Chomskys Skinner-Rezension entstand eine eigenartige Kluft zu empirisch orientierten Wissenschaften wie der Psychologie, zu der man ja einerseits eine große Nähe suchte, da Linguistik als Teil der Psychologie verstanden werden sollte. Andererseits verweigerte sich aber gerade der Chomskysche Mainstream der generativen Grammatik einer damit eigentlich zwingend einhergehenden experimentell überprüfbareren Gegenstandsauffassung. So gerät die Rede von der Linguistik als Psychologie oder Biologie zum reinen Lippenbekenntnis nach dem Motto: „Wasch mir den Pelz, aber mach mich nicht nass!“

Die Grammatikforschung hat in erster Linie die Beschreibung und Analyse der natürlichen Sprachen und ihrer Grammatiken, also Devitts (2006) „Linguistic Reality“, zum Gegenstand. Dieser Gegenstand bleibt derselbe, ob es ein Sprachorgan gibt oder nicht. Die Linguistik, mit der Grammatikforschung als Kerndisziplin, hat sich längst ihren Platz als eigenständige empirische Wissenschaft erstritten, und der Beitrag der von Chomsky begründeten generativen Grammatik ist dabei nicht unerheblich. Wenn man Grammatikforschung *zusätzlich* als Biologie *interpretiert*, wird sich an ihrem eigenständigen Status nichts ändern.

Literaturverzeichnis

Bauer, Joachim (2005). *Warum ich fühle, was du fühlst. Intuitive Kommunikation und das Geheimnis der Spiegelneurone*. Hoffmann und Campe, Hamburg.

Bloomfield, Leonard (1933). *Language*. Holt, Rinehart, and Winston, New York.

Chomsky, Noam (1959). ‘A Review of B. F. Skinner’s Verbal Behavior.’ *Language* 35:26–58.

- (1965). *Aspects of the Theory of Syntax*. MIT Press, Cambridge, Massachusetts.
- (1968). ‘Quine’s Empirical Assumptions.’ *Synthese* 19:53–68.
- (1981). *Reflections On Language*. Foris, Dordrecht.
- (1996). *Powers and prospects. Reflections on human nature and the social order*. South End Press, Boston.
- Crain, Stephen & Rosalind Thornton (1998). *Investigations in Universal Grammar: A Guide to Experiments on the Acquisition of Syntax and Semantics*. MIT Press, Cambridge, MA.
- Dennett, Daniel (1996). *Kinds of Minds. Toward an Understanding of Consciousness*. Basic Books, New York.
- Devitt, Michael (2006). *Ignorance of Language*. Oxford University Press, New York.
- Fanselow, Gisbert (2000). ‘Optimal Exceptions.’ *In*: Barbara Stiebels & Dieter Wunderlich (eds.), ‘Lexicon in Focus,’ pp. 173–209. Akademie-Verlag, Berlin.
- Graham, George (2007). ‘Behaviorism.’ *Stanford Encyclopedia of Philosophy*. URL: <http://plato.stanford.edu/entries/behaviorism/>
- Jónsson, Jóhannes Gísli (2003). ‘Not so Quirky: On Subject Case in Icelandic.’ *In*: Ellen Brandner & Heike Zinsmeister (eds.), ‘New Perspectives on Case Theory,’ pp. 127–163. CSLI Publications, Stanford, CA.
- Kolb, Hans-Peter (1997). ‘Is I-Language a Generative Procedure?’ *In*: ‘GB-Blues: Two Essays,’ vol. 110 of *Arbeitspapiere des Sonderforschungsbereichs 340*, pp. 3–11. University of Tübingen.

- Lewis, John D. & Jeffrey L. Ellman (2001). 'Learnability and the Statistical Structure of Language: Poverty of Stimulus Arguments Revisited.' *In*: 'Proceedings of 26th Annual Boston University Conference on Language Development,' pp. 359–70. Cascadilla Press, Somerville, MA.
- MacCorquodale, Kenneth (1970). 'On Chomsky's Review of Skinner's VERBAL BEHAVIOR.' *Journal of the Experimental Analysis of Behavior* 13:83–99.
URL: <http://behavior.org/computer-modeling/maccorquodale/maccorquodale1.cfm>
- Meyer, Markus (2006). *Grammatische Praxis. Probleme der grammatischen Theoriebildung und der Grammatikschreibung*. Stauffenburg, Tübingen.
- Minsky, Marvin A. & Seymour A. Papert (1969). *Perceptrons. An Introduction to Computational Geometry*. MIT Press, Cambridge, Mass.
- Olazaran, Mikel (1996). 'A Sociological Study of the Official History of the Perceptrons Controversy.' *Social Studies of Science* 26(3):611–659.
URL: <http://www.jstor.org/stable/285702>
- Partee, Barbara H., Alice ter Meulen, & Robert E. Wall (1990). *Mathematical Methods in Linguistics*, vol. 30 of *Studies in Linguistics and Philosophy*. Kluwer der Abzocker, Dordrecht.
- Peters, Stanley & Robert W. Ritchie (1973). 'On the generative power of transformational grammars.' *Information Science* 6:49–83.
- Piattelli-Palmarini, M. (1980). *Language and Learning. The Debate between Jean Piaget and Noam Chomsky*. Harvard University Press, Cambridge Massachusetts.

- Prince, Alan & Paul Smolensky (1993). 'Optimality Theory. Constraint Interaction in Generative Grammar.' Rutgers Optimality Archive: ROA-537.
URL: <http://roa.rutgers.edu/view.php3?roa=537>
- (2004). *Optimality Theory. Constraint Interaction in Generative Grammar*. MIT Press, Cambridge, Mass.
- Pullum, Geoffrey & Barbara Scholz (2002). 'Empirical Assessment of Poverty Stimulus Arguments.' *The Linguistic Review* 19:9–50.
- Quine, William van Orman (1960). *Word and Object*. Studies in Communication. The Technology Press of The Massachusetts Institute of Technology/John Wiley & Sons, New York/London.
- (1970). 'Methodological Reflections on Current Linguistic Theory.' *Synthese* 21:386–398.
- Skinner, Burrhus Frederic (1957). *Verbal Behavior*. Appleton Century Crofts, New York.
- Smolensky, Paul & Géraldine Legendre (2006a). *The Harmonic Mind, Volume 1: From Neural Computation to Optimality-Theoretic Grammar: Cognitive Architecture*. MIT Press, Cambridge, Mass.
- (2006b). *The Harmonic Mind, Volume 2: Linguistic and Philosophical Implications*. MIT Press, Cambridge, Mass.
- Sternefeld, Wolfgang (2000). 'Grammatikalität und Sprachvermögen.' In: Josef Bayer & Christine Römer (eds.), 'Von der Philologie zur Grammatiktheorie,' pp. 15–44. Niemeyer, Tübingen.
- Thomas, Margaret (2002). 'Development of the concept of the "Poverty of the Stimulus".' *The Linguistic Review* 19:51–71.

Wasow, Tomas (2002). *Postverbal Behavior*. The University of Chicago Press, Chicago.

Watson, John B. (1913). 'Psychology as the Behaviorist Views It.' *Psychological Review* 20(3):158–177.

URL: <http://psychclassics.yorku.ca/Watson/views.htm>

Wunderlich, Dieter (2003). 'Optimal Case Patterns: German and Icelandic Compared.' *In*: Ellen Brandner & Heike Zinsmeister (eds.), 'New Perspectives on Case Theory,' pp. 331–367. CSLI Publications, Stanford, CA.