

Methoden der Syntaxanalyse

(Dieses Unterrichtsmaterial kann frei verwendet werden.)

1. Vorbemerkungen über die Linguistik als empirisch-theoretische Wissenschaft

Die heutige Linguistik erhebt den Anspruch, eine empirische und exakte Wissenschaft zu sein. Sie muß daher den Anforderungen genügen, die von der modernen Wissenschaftstheorie an jede empirisch-theoretische Wissenschaft gestellt werden. Dazu gehören z. B. die explizite Einführung der Termini, die Widerspruchsfreiheit der Aussagen, die Nachprüfbarkeit der Ergebnisse. Die Linguistik unterscheidet sich hierin in nichts von Naturwissenschaften wie z.B. der Physik.

Der Gegenstand der Linguistik ist die Sprache. Der Linguist möchte wissen, aus welchen Bestandteilen eine Sprache besteht und wie sie als Mitteilungssystem funktioniert. Dadurch unterscheidet sich die Linguistik von der Literaturwissenschaft, der es eher auf die durch Sprache mitgeteilten Sachverhalte ankommt, sowie von einem Teil der älteren Sprachwissenschaft, die anhand von sprachlichen Erscheinungen eher Aufschlüsse über historische und prähistorische soziale Zustände gewinnen wollte, statt über die Sprache selbst.

Gegenstand der Linguistik ist - genauer formuliert - das, was die Kommunikationspartner kennen müssen, wenn sie sich mit Hilfe von Sprache verständigen wollen. Obwohl es sich dabei zunächst um psychische Gegebenheiten handelt, um im Bewußtsein gespeicherte virtuelle Muster nämlich, nach welchen die Kommunikationspartner sprachliche Äußerungen produzieren bzw. verstehen, kann dieser Gegenstand durchaus empirisch-exakt untersucht werden. Informantenbefragung (wobei der Linguist, der die Sprache kennt, die er beschreiben will, zunächst sein eigener Informant ist), Tests und Analysen vorliegender sprachlicher Äußerungen sind die Mittel dazu.

Die Untersuchungen der Linguistik erstrecken sich über Ausdrucks- wie Inhaltsseite der sprachlichen Einheiten. Der Zusammenhang von Sprache als virtuellem Muster, das zu Zwecken der Kommunikation zur Verfügung steht, und den unterschiedlichen Verwendungen von Sprache in Kommunikationsakten kann nur durch interdisziplinäre Zusammenarbeit der Linguistik mit jenen Wissenschaften geklärt werden, die sich von Haus aus mit den Gebieten befassen, die mit Sprachverwendung in Zusammenhang stehen. Untersucht die Linguistik Sprachsysteme, die Soziologie soziale Systeme, die Psychologie psychische Vorgänge, die Medizin u.a.

neurologische Zusammenhänge, so wird eine Zusammenarbeit von Linguisten und Soziologen als interdisziplinäres Fach Soziolinguistik, jene von Linguisten und Psychologen als interdisziplinäres Fach Psycholinguistik und die von Linguisten und Medizinern z.B. als Sprachpathologie zur Hoffnung auf Ergebnisse berechtigen. Es sei aber betont: Der Gegenstand der Linguistik ist die Sprache *sui generis*. Soziolinguistik dagegen ist z.B. ein interdisziplinäres Fach, das so lange im argen liegen wird, wie einzelne Linguisten oder Soziologen nur auf Nachbargebieten dilettieren.

Jede empirische Wissenschaft untersucht ihren Gegenstand und beschreibt ihn. Die Beschreibungen werden unter den Fachangehörigen diskutiert und all jenen zur Verfügung gestellt, die für den Umgang mit dem betreffenden Gegenstand in der Praxis Orientierungshilfen benötigen. Nur der liebe Gott, wenn überhaupt jemand, hat direkten Zugang zu den Sachen! Wir können uns Gegenstände nur vorstellen. Wir machen uns ein Bild von ihnen, wir entwerfen ein Modell. Ein solches Modell nennt man eine **Theorie**. Ohne Theorie gibt es keinen Zugang zum Material. Empirische Wissenschaften sind daher stets auch theoretische Wissenschaften. Theorien werden so konzipiert, daß sie unter einer bestimmten Hinsicht generelle Aussagen über den betreffenden Gegenstand zulassen. Sie machen damit Vorhersagen über sein zukünftiges Verhalten unter eben jener Hinsicht möglich und sind so für praktisches Handeln verwertbar. Dafür zu sorgen, daß wissenschaftliche Theorien für jedermann zugänglich sind, ist eine politische Aufgabe.

Der Grad unserer Erkenntnis der Gegenstände hängt von der Güte der Theorie ab, die wir entwerfen. Darum sind Überlegungen wichtig, wie man Fehler beim Aufbau von Theorien vermeiden, sowie ihre Angemessenheit dem Gegenstandsbereich gegenüber gewährleisten kann. Die innere Konsistenz einer Theorie versucht man z.B. dadurch zu erreichen, daß man ihre Konstruktion den Gesetzen der Logik unterwirft. Ihre Adäquatheit wird dadurch angestrebt, daß man die theoretischen Aussagen experimenteller Kontrolle unterwirft und dabei nach intersubjektiver Übereinstimmung in der Beobachtung empirischer Daten trachtet.

Um eine konsistente Theorie konstruieren zu können, muß man mit möglichst wenigen Voraussetzungen beginnen. Die Voraussetzungen, ihr **Axiome**, müssen zudem explizit angegeben werden, damit sie von jedem, der sich auf eine Theorie stützt, durchschaut werden können, und damit es auch möglich ist, neue Theorien auf neuen Axiomen zu errichten.

2. Der strukturelle Ansatz der Linguistik

Die Linguistik betrachtet die Sprache unter dem Gesichtspunkt der Struktur. Eine Menge *M* heißt **mit einer Struktur versehen**, wenn zwischen den Elementen von *M* oder zwischen Elementen

und gewissen Teilmengen von M Relationen R bestehen. Der Linguist versucht also herauszufinden, welche Elemente ein Sprachsystem enthält und welche Relationen zwischen diesen existieren. Zunächst wird er von vorliegenden oder möglichen sprachlichen Äußerungen ausgehen. Diese muß er segmentieren, um zu den Elementen des Sprachsystems zu gelangen. Sodann wird er den Zusammenhang zwischen den Sprachelementen beschreiben. Man muß sich aber vor Augen halten, daß Segmentierung und Zusammenhang Kehrseiten derselben Medaille sind. Elemente lassen sich nur unterscheiden aufgrund von den Relationen, in denen sie stehen, und je nach Segmentierung wird man andere Relationen zwischen den Segmenten konstatieren. Teilt man einen Text z.B. in Phoneme, so hat man als Kriterium vorab bereits eine bestimmte Relation eingeführt, nämlich die der bedeutungsunterscheidenden Opposition von Phonemen. Teilt man einen Text dagegen in Silben, so macht man eine andere Relation zum Kriterium der Segmentierung und wird folglich auch eine andere Struktur des Textes ermitteln. Es gibt ein Prinzip, das sinnvollere von weniger sinnvollen Segmentierungen unterscheiden hilft. Es besteht darin, daß man nach Teilungen sucht, die möglichst begrenzte Elementeninventare ergeben. Bei einer Segmentierung nach Phonemen kommt man z.B. im Deutschen auf etwa 38 Elemente. Bei einer Einteilung nach Silben erhielte man ein viel größeres Elementeninventar. Wollte jemand z.B. die Zusammensetzung von größeren Einheiten wie Wörtern aus kleineren beschreiben, so würde er bei Zugrundelegung von Silben zwar wahrscheinlich weniger komplizierte Verkettungsregeln angeben, dafür aber ein sehr großes Inventar von Silben aufstellen müssen. Ökonomischer und den tatsächlichen Vorgängen bei der Sprachspeicherung und Sprachäußerung wahrscheinlich näher kommend, wird er daher lieber zunächst Regeln angeben, nach denen die große Zahl der Silben aus der kleinen Zahl der Phoneme konstruierbar ist.

Wenn man sprachliche Strukturen untersuchen will, so ist die Unterscheidung zweier Relationen, der syntagmatischen und der paradigmatischen, grundlegend. Gegeben sei z.B. eine Satzform A ist größer als B . Die Beziehung zwischen den Elementen A , ist größer als B ist eine syntagmatische Relation. Die durch sie hergestellte Struktur nennen wir ein **Syntagma**. Die syntagmatische Relation besteht also in der Verbindung von einzelnen Elementen zu einem Ganzen. Innerhalb des Syntagmas des obigen Beispiels können nun verschiedene Elemente an den Stellen A und B eingesetzt werden, etwa:

A		B
$\left\{ \begin{array}{l} \text{Karl} \\ \text{die Sonne} \\ \text{Amerika} \end{array} \right\}$	ist größer als	$\left\{ \begin{array}{l} \text{Gustav} \\ \text{der Mond} \\ \text{Europa} \end{array} \right\}$

Die Beziehung zwischen den einzelnen Elementen, die an der Stelle A bzw. B stehen können, nennt man **paradigmatische Relation**. Die Elemente, die in dieser Relation zueinander stehen, bilden ein **Paradigma**. Ein Paradigma ist also die Menge von Elementen, die an einer bestimmten Stelle eines Syntagmas gegeneinander ausgetauscht werden können. Natürlich gibt es auch für ist größer als paradigmatische Entsprechungen, für unser Beispiel etwa: *ist kleiner als* oder *ist älter als*.

Den syntagmatischen und paradigmatischen Relationen entsprechen zwei Analyseoperationen des Linguisten: die der Segmentierung und die der Klassifizierung. Erstens müssen die sprachlichen Äußerungen in Elemente zerlegt werden, so daß die Syntagmen erkennbar werden. Zweitens müssen die Segmente klassifiziert werden, indem man solche zusammenfaßt, die in derselben paradigmatischen Relation stehen. Zum Terminus Klasse: **Klassen** sind das Ergebnis einer Zerlegung einer Menge M in Teilmengen, sodaß jedes Element von M genau einer Teilmenge angehört.

Man kann Mengen auf dreierlei Weise definieren. Erstens kann man einfach alle Elemente, die zu einer Menge M gehören sollen, aufzählen:

$$M = \{a, e, i, o, u\}$$

Man nennt dies eine **extensionale Definition** der Menge M. Die Schreibweise mit geschweiften Klammern ist die übliche.

Zweitens kann man ein bestimmtes gemeinsames Merkmal der Elemente (oder auch mehrere Merkmale zusammen) zum Kriterium ihrer Zugehörigkeit zur Menge M erheben. Man nennt dies eine **intensionale Definition** der Menge M und schreibt:

$$M = \{x : P(x)\}$$

Dies ist zu lesen als *M ist die Menge aller x für die gilt P*. Wenn M z.B. die Menge der deutschen Vokale sein soll, so kann man sie einmal extensional definieren, indem man die Elemente wie oben aufzählt. Oder aber man definiert sie intensional, wobei P etwa lauten könnte: *sie können allein eine Silbe bilden* (vergleiche: a-bend, e-sel, i-gel, o-fen, u-fer). Wie wir sehen, können intensionale und extensionale Definitionen zu genau derselben Mengenbildungen führen. Da in der intensionalen Definition die Angabe einer Regelhaftigkeit enthalten ist und da man bei ihr nicht umfangreiche Inventare aufführen muß, wird man in der Linguistik häufig nach intensionalen Definitionen suchen. Andererseits wird durch die extensionale Aufführung der Elemente die Überprüfbarkeit erleichtert. Da Elemente und Regeln innerhalb der

Strukturbeschreibung zusammengehören, wird man häufig extensionale in intensionale Definitionen überführen und umgekehrt.

Die dritte Weise des Definierens besteht darin, daß man ein genaues Verfahren angibt, was man tun muß, um als Ergebnis genau die Elemente der zu definierenden Menge zu erhalten. Dies nennt man eine operationale Definition einer Menge. Beispiele dafür werden wir weiter unten bei der Besprechung heuristischer Techniken finden.

Eine vollständige Beschreibung eines Sprachsystems L wollen wir eine **Grammatik der Sprache L** nennen. Äußerungen, die dem Elementevorrat und den Regeln der Sprache L entsprechen, die also gültige Äußerungen in dieser Sprache sind, sollen **grammatische** Äußerungen heißen. Das Urteil darüber, ob eine gegebene Äußerung grammatisch in Bezug auf L ist, obliegt denjenigen, die diese Sprache beherrschen. Zunächst wird der Linguist sein eigenes Urteil darüber fällen. Sodann wird er es zur Diskussion stellen, um mit anderen Kennern der betreffenden Sprache darüber einen Konsens zu erreichen. Kommt eine solche Übereinstimmung nicht zustande, so wird man nach weiteren Kriterien suchen, unter denen die betreffende Äußerung von allen akzeptiert wird (z.B., durch Konstruktion einer besonderen Situation, in der sie vorkommen könnte). In den meisten Fällen führt dies zum Ziel. Die übrigen Fälle schließt man am besten vorläufig als zweifelhaft aus der Beschreibung aus.

3. Heuristische Techniken

Es sollen nun einige Analyseverfahren angegeben werden, mit deren Hilfe man versuchen kann, etwas über die Struktur sprachlicher Äußerungen herauszufinden, um daraus auf die Struktur des Sprachsystems zu schließen. Da es uns hier nur auf die Demonstration des Vorgehens an sich ankommt, wollen wir die vor-theoretische Entscheidung treffen, daß die sprachlichen Äußerungen zunächst in Wörter zu segmentieren seien. **Wörter** könnte man extensional definieren, indem man ein Wörterbuch zugrunde legt. Intensional definierend könnte man etwa sagen: Wörter sind Elemente einer Äußerung, die in geschriebener Form durch einen Zwischenraum von einander getrennt sind. Bei genauerer Beschäftigung mit der deutschen Sprache würde man erkennen, daß es nicht gut ist, als Grundelemente Wörter anzunehmen, sondern eher Teile von Wörtern. Doch soll uns diese Frage hier nicht beschäftigen. Prinzipiell kann man ähnliche Verfahren, wie die unten geschilderten, bei jeder Art von Segmentierung anwenden. Wir wollen hier auch nicht erörtern, wie man die größeren Einheiten, die aus Wörtern zusammengesetzt sind, abgrenzen kann, wie man also etwa **Text** oder **Satz** definieren könnte. Wir nennen im folgenden einfach jede Zusammensetzung von Wörtern eine **Kette**.

Beschäftigen wir uns als erstes mit syntagmatischen Relationen und betrachten wir Syntagmen zunächst unter dem Gesichtspunkt der linearen Abfolge ihrer Elemente, d.h. unter der Relation vorher/nachher! Syntagmen unterscheiden sich dann zunächst in der Anzahl ihrer Elemente. Wir nehmen einmal an, daß unser Syntagma aus vier Elementen besteht, und zwar aus vier Wörtern der deutschen Sprache. Es gibt dann 4! theoretisch mögliche Kombinationen (4! = vier Fakultät, d.i. $1 \times 2 \times 3 \times 4 = 24$). Nur einige davon sind jedoch nach den Regeln der deutschen Grammatik erlaubt und mit einer Gesamtbedeutung versehen.

3.1. Verschiebeprobe (Permutation)

Um syntagmatische Regeln zu finden, kann man die verschiedenen Kombinationsmöglichkeiten durchprobieren, indem man die Elemente untereinander verschiebt. Dabei überprüft man jedesmal:

- a) Ergibt die Verschiebung eine neue grammatische Kette?
- b) Hat die neue Kette dieselbe oder eine andere Bedeutung?

Gegeben sei folgendes Syntagma:

(1) A (diese) B (Nachricht) C (erschütterte) D (ihn)

Die Verschiebeprobe ergibt folgende grammatische Ketten:

(2) D (ihn) C (erschütterte) A (diese) B (Nachricht)

(3) A (diese) B (Nachricht) D (ihn) C (erschütterte)

(4) D (ihn) A (diese) B (Nachricht) C (erschütterte)

(5) C (erschütterte) D (ihn) A (diese) B (Nachricht)

(6) C (erschütterte) A (diese) B (Nachricht) D (ihn)

Die übrigen 18 Kombinationsmöglichkeiten sind nicht grammatisch. Der Leser möge dies bitte selbst überprüfen.

(1) und (2) dürften von der Mehrzahl der Kenner des Deutschen als bedeutungsgleich angesehen werden. Zumindest kann man sich eine Situation oder einen Kontext vorstellen, in welchen (1) und (2) ohne Bedeutungsveränderung gegeneinander austauschbar wären. Auf mögliche Bedeutungsunterschiede in Gestalt einer Hervorhebung (*ihn, und nicht jemanden anders, erschütterte diese Nachricht*) soll hier nicht eingegangen werden. ABCD und DCAB sind also freie Variationen der Wortstellung.

(3) und (4) sind für sich allein stehend ungrammatisch. Sie sind jedoch gültige Syntagmen, wenn sie in einem Kontext auftreten, wenn ihnen z.B. ein anderer Satz und ein *weil, während* o.ä. vorausgeht:

(7) ... weil ihn diese Nachricht erschütterte

In einem solchen Fall dürfte die Bedeutung von (3), (4) der von (1), (2) entsprechen.

(5) und (6) sind grammatische Ketten, doch haben sie eine andere Bedeutung als die übrigen. Allgemein umrissen, besteht der Unterschied in (1) bis (4) = *Aussage* (5) und (6) = *Frage*.

Die Stellung der Elemente in der Kette wirkt also in manchen Fällen bedeutungsdifferenzierend. Der Ausdrucksseite Wortstellung entspricht in manchen Fällen eine bestimmte Inhaltsseite, hier die Kennzeichnung als Aussage oder als Frage.

Die Verschiebungsprobe erbringt neben der Auffindung möglicher Syntagmen und etwaiger Bedeutungsunterschiede noch einen weiteren Aufschluß über die Struktur der betreffenden Kette: A und B können offenbar nur zusammen verschoben werden. Sobald man ihren Zusammenhang zerreißt, indem man sie einzeln verschiebt, ist eine ungrammatische Kette das Ergebnis. Daraus können wir schließen, daß diese beiden Elemente enger zueinander gehören als zu den übrigen, d.h. wir erhalten durch die Verschiebeprobe bereits einen ersten Hinweis auf eine besondere Relation zwischen diesen Elementen. Wir können daher A und B zu einem einzigen Element X zusammenfassen und sagen: *X besteht aus A und B*. Die Relation zwischen X und AB ist dann eine Menge-Teilmenge-Relation. Damit haben wir, zusätzlich zur Vorher/Nachher-Relation, eine erste hierarchische Relation im Syntagma konstatiert. **Hierarchisch** heißt eine Relation, wenn in ihr die Argumente von verschiedener Rangordnung sind wie z.B. die zwischen einer Menge und ihren Elementen.

Mit Hilfe der Verschiebeprobe können wir zu einer ersten Segmentierung von Syntagmen gelangen, die nicht mehr nur auf Wörtern beruht und sich bei näherem Zusehen auch als angemessener als jene erweisen würde. Wir könnten nämlich eine bestimmte Klasse von Konstituenten eines Satzes operational definieren als diejenigen Teile, die sich nur als ganze verschieben lassen. Die erste Teilung unserer Beispielskette sähe dann so aus:

(8) x (diese Nachricht) D (erschütterte) C (ihn)

Wenden wir darauf die Permutation vollständig an, ergibt sich, daß alle möglichen Kombinationen ($3! = 6$) grammatisch sind, denn das Ergebnis sind genau wieder die Ketten (1) bis (6).

Man kann die Permutation manchmal mehrfach hintereinander auf verschiedenen Ebenen durchführen und dadurch weitere hierarchische Relationen aufdecken. Nach obiger Operationsregel ergibt sich für eine Kette folgende erste Segmentierung:

(9) A (er) B (bat) C (mich) D (dir davon zu berichten)

Innerhalb der Gesamtkette kann D nur als ganzes verschoben werden (z.B.: DBAC). Innerhalb von D aber sind die Elemente wieder verschiebbar. Auf dieser zweiten Ebene ergibt die Verschiebeprobe:

(9i) a (dir) b (davon) c (zu berichten)

b (davon) a (dir) c (zu berichten)

Mit D und c liegen übrigens ähnliche Fälle vor wie schon mit A und B in (1) bis (9). Zwischen den Elementen von D bzw. c bestehen wieder besondere Relationen, denn sie lassen sich nur gemeinsam verschieben.

3.2. Austauschprobe (Kommutation)

Durch Anwendung der Verschiebeprobe enthüllt sich ein Teil der syntagmatischen Struktur von Ketten. Wie steht es nun mit den paradigmatischen Relationen der Elemente? Um diese herauszufinden, tauschen wir einzelne Elemente eines Syntagmas gegen andere aus und fragen uns, ob die neu gewonnenen Ketten grammatisch sind oder nicht. Bei folgendem Schaubild stelle man sich vor, daß sich die Spalten A bis D vertikal verschieben lassen. Dadurch werden die Elemente, die in der umrandeten Zeile in der Mitte stehen, ausgetauscht.

(10)	A	B	C	D

	über	gelb	Haus	Riese
	die	Gustav	während	schlafen
	mein	Hut	ist	grün
	dein	Papagei	war	neu
	Irenes	Auto	schien	weg

Zur besseren Demonstration sind die Wörter in den Kästchen schon geordnet. Verschiebt man A bis D nach unten, d.h. tauscht man die Elemente der umrandeten Zeile gegen darüber stehende aus, erhält man durchgehend eine ungrammatische Kette. Verschiebt man die Spalten jedoch nach oben, so daß die unter der Zeile stehenden Elemente eingesetzt werden, erhält man in allen Kombinationen eine neue grammatische Kette.

Die Austauschprobe ist ein erster Schritt dazu, die Elemente des Sprachsystems zu klassifizieren. Nur eine bestimmte Teilmenge aller Wörter kann an einer bestimmten Strukturstelle stehen. Die Gesamtmenge aller Elemente des Sprachsystems läßt sich somit zerlegen in paradigmatische Klassen. In unserem Falle gehören {*mein, dein, Irenes*} zur Klasse A, {*Hut, Papagei, Auto*} zur Klasse B und sofort. {*über, die*} gehören nicht zur Klasse A, so wenig wie {*Riese, schlafen*} zur Klasse D.

Weiter kommen durch dieses Verfahren gerade die Positionen, d.h. die Strukturstellen in einem Syntagma, als für die linguistische Analyse wichtige Einheiten in den Blick. Will man das vorliegende Syntagma generell beschreiben, geht es nicht um *mein* und *Hut*, sondern um die Positionen A und B, die in bestimmter Beziehung zueinander stehen. Dadurch, daß an bestimmter Stelle nicht beliebige Wörter kommutierbar sind, sondern nur solche einer bestimmten Klasse, erweisen sich die Positionen als Einheiten mit unterschiedlicher Funktion im Ganzen des gemeinten Zusammenhanges. Wir sind damit in der grammatischen Analyse ein erhebliches Stück weiter gekommen.

Neben der Aufdeckung der paradigmatischen Relationen gibt uns die Austauschprobe auch noch Hinweise auf besondere syntagmatische Relationen zwischen den Positionen, dann nämlich, wenn die Kommutation an einer Stelle begleitet werden muß durch eine Kommutation an einer oder mehrerer anderer Stellen.

(11)	A	B	C	D
	die	Frau	empörte	sich
	der	Frauen	empörten	
		Mann		

Wird in B *Frauen* gewählt, so muß in C *empörten* gewählt werden. Wird in B *Frau* gegen *Mann* ausgetauscht, so muß in A *die* gegen *der* ausgetauscht werden. Die Positionen B und C sowie A und B stehen folglich syntagmatisch nicht nur in der Relation des Vorher/Nachher, sondern offensichtlich noch in einer hierarchischen Relation, die hier jedoch nicht näher erörtert werden soll.

Die Positionen für die man Kommutationsversuche anstellt, sollte man nicht willkürlich festlegen. Für sich allein angewendet, ist die Austauschprobe ziemlich ungeeignet, bei der Segmentierung von Ketten zu helfen. Für nahezu jede Segmentierung findet man nämlich kommutierendes Material, wenn auch in unterschiedlicher Zahl. Man spiele das einmal an einem Satz mit n Elementen durch, indem man je 2,3,4 usw. bis n-1 Elemente zusammenfaßt und für sie nach Kommutationen sucht. Zu Anfang einer Analyse ermitteln wir Positionen besser mit Hilfe der Verschiebeprobe und versuchen erst dann, durch Austauschproben zu einer Klassifizierung der Positionen zu gelangen.

In (10) haben wir { *Hut, Papagei, Auto* } zu einer paradigmatischen Klasse zusammengefaßt. Es ist zu untersuchen, ob diese Elemente auch in allen anderen Syntagmen eine Klasse bilden. Das ist nicht der Fall. Vergleiche (10) mit

(12)				
	Günther	schläft	im	$\left. \begin{array}{l} \text{Hut} \\ \text{Papagei} \\ \text{Auto} \\ \text{Bett} \\ \text{Zimmer} \end{array} \right\}$

In (12) ist vielmehr { *Auto, Bett, Zimmer* } zu einer Klasse zusammenzufassen, *Hut* und *Papagei* aber auszuschließen. Und sogar in einem Gefüge mit denselben Positionen ergeben sich unterschiedliche paradigmatische Klassen, je nach den Paradigmen, die in den anderen Positionen stehen:

(13)	A	B	C
	der Papagei	spricht	auf der Stange
	das Auto	fährt	auf der Straße
	der Hut	liegt	auf Rädern

Man wird hier etwa die folgenden Paradigmenklassen aufstellen:

(14)	A1= {Papagei}	B1= {spricht}	C1= {auf der Stange}
	A2= {Auto}	B2= {fährt}	C2= {auf der Straße auf Rädern}
	A3= {Papagei Auto, Hut}	B3= {liegt}	C3= {auf der Straße}

Diese komplizierten Zusammenhänge versucht der Linguist zu berücksichtigen, indem er eine sogenannte Kreuzklassifikation vornimmt. Das Prinzip dabei ist, daß die Klassenzuweisungen der Elemente mit den Zuweisungen anderer Elemente zu anderen Klassen in Zusammenhang gebracht werden. Wie das formal geschehen kann, soll hier nicht näher erörtert werden. Im Beispiel (14) sind die Elemente durch die übereinstimmenden Buchstaben nach Positionen klassifiziert (A1,A2,A3) und zugleich durch die übereinstimmenden Ziffern nach möglichen Verbindungen mit anderen Klassen innerhalb des Syntagmas (A1,B1,C1).

Dem Leser mögen bei der Klassifizierung in (14) Zweifel gekommen sein, ob zur Klasse C1 nicht auch *auf der Straße* gehören müßte, oder zu C3 *auf Rädern*;

(15) Der Papagei spricht auf der Straße

(16) Der Hut liegt auf Rädern

Lassen sich nicht Situationen konstruieren, in denen Sätze wie (15) und (16) vorkommen könnten? Da in diesem Papier nur linguistische Methoden dargestellt werden sollen, nicht aber eine eigentliche Theorie aufgestellt werden soll, wollen wir es hier bei dem Hinweis auf die Problematik bewenden lassen. Die Lösung wird darin liegen, daß man unterscheidet zwischen dem, was in einer Grammatik als zugelassen gelten soll und was nicht, und dem, was in der Welt als möglich denkbar ist und was nicht. Die Grammatik darf sicher nicht so eng konzipiert werden, daß Sätze wie (15) und (16) ausgeschlossen werden. Deren Abweichung oder Nicht-Abweichung kann im Rahmen einer Theorie des Sprachgebrauchs erörtert werden, die die situativen Faktoren der Kommunikation einbezieht.

Es gibt aber eine große Zahl von Elementen, die man auf jeden Fall unterschiedlichen Paradigmenklassen zuordnen muß. In (13) kann der Klasse A auf keinen Fall *dem Auto* oder *aus* oder *während* zugewiesen werden. Wörter haben bestimmte formale Eigenschaften, die sie nur in ganz bestimmte Positionen von Syntagmen eintreten lassen. Viele Linguisten sehen daher im Anschluß an Noam Chomsky zwei verschiedene Klassifizierungen der Elemente des Sprachsystems vor:

- a) die *s t r i k t e* Subklassifizierung, die unter dem Gesichtspunkt erfolgt, in welchen syntagmatischen Positionen die Elemente überhaupt auftreten können,
- b) die *s e l e k t i o n a l e* Subklassifizierung, die berücksichtigt, welche Elemente sich gegenseitig zulassen, wenn sie in einem Syntagma an ihren jeweiligen Positionen stehen

Nach a) gehören *Papagei, Auto, Hut* zu einer Klasse, nach b) aber sind sie - möglicherweise durch Kreuzklassifikation - verschiedenen Klassen zuzuweisen.

3.3. Die Exklusionsprobe

Wenn man Zweifel hat, ob die Elemente, die an einer bestimmten Stelle in einem Syntagma austauschbar sind, alle demselben Paradigma zugerechnet werden können, kann man ausprobieren, ob sie auch zusammen im betreffenden Syntagma vorkommen können oder, ob sie sich dort gegenseitig ausschließen. Der Grundgedanke dabei ist, daß eine Position in einem Syntagma immer nur durch e i n Element der zugehörigen Klasse ausgefüllt werden kann. Können dagegen mehrere Elemente bestimmter Art gleichzeitig vorkommen, so sind für sie auch mehrere Positionen vorzusehen und damit wiederum mehrere Paradigmenklassen. Betrachte folgendes Beispiel:

(17)	A	B	C
	Die Ordinarien	proben	den Aufstand die Revolution den ganzen Abend

(17) ergibt eine grammatische Kette bei jedem Austausch an der Stelle C. Die Exklusionsprobe ergibt, daß sich *Aufstand* und *Revolution* ausschließen (Ein * kennzeichnet Ketten, die der Linguist als ungrammatisch betrachtet.):

(17i) * (Die Ordinarien) (proben) (den Aufstand) (die Revolution)

Demnach gehören Aufstand und Revolution hier zum selben Paradigma, von dem nur jeweils ein Element die Position C einnehmen kann.

(17ii) (Die Ordinarien) (proben) (den Aufstand) (den ganzen Abend)
 A B C D

Den ganzen Abend erweist sich als andere Position im Syntagma, denn es kann zusätzlich zu einer Füllung von C durch ein anderes Element auftreten. Daraus ist abzuleiten, daß den folgenden beiden Ketten, trotz äußerer Ähnlichkeit, zwei unterschiedliche Syntagmen zugrunde liegen:

(17iii) Die Ordinarien proben den Aufstand (= ABC)
 Die Ordinarien proben den ganzen Abend (= ABD)

3.4. Die Nektionsprobe

Die Exklusionsprobe kann ergänzt werden durch die Nektionsprobe. Unter **Nektion** versteht man die Verbindung zweier Elemente durch *und* bzw. *oder*. Da diese Verbindung eine generelle Möglichkeit für sehr viele sprachliche Einheiten ist, wäre es unökonomisch für durch *und/oder* verbundene Elemente jeweils zwei Positionen im Syntagma anzunehmen. Allzu viele Positionen müßte man dann verdoppeln. Es erscheint praktischer, eine generelle Regel vorzusehen, die die Verdoppelung von Elementen an einer bestimmten Position mit *und/oder* bewirkt. Dann ist aber die Möglichkeit der Nektion zweier Elemente gerade ein Beweis dafür, daß nur eine Position im Syntagma vorliegt. Vergleiche:

(18) (Die Ordinarien) (proben) (den Aufstand oder die Revolution)
 A B C

Dagegen:

(18i) * (Die Ordinarien) (proben) (die Revolution und den ganzen
 Abend)

Elemente, die auf bestimmter Ebene nicht mehr zur selben paradigmatischen Klasse gehören,
lassen sich nicht nektieren.

3.5. Die Abstrich- oder Tilgungsprobe

Eine spezielle Art der Kommutation ist die Tilgungsprobe. In (11) hatten wir gesehen, daß die Kommutation bestimmte hierarchische Relationen zwischen Positionen aufdecken kann. Das kann man noch deutlicher machen, indem man versucht, einzelne Elemente eines Syntagmas völlig zu streichen. Im folgenden Beispiel sind Tilgungen durch _ dargestellt:

(19) der sehr umstrittene Mayer

 der umstrittene Mayer

 der Mayer

 Mayer

(20) * Der sehr _ Mayer

 * sehr umstrittene Mayer

 * sehr umstrittene _

 * umstrittene _

Die Kette unter (19) werden wir akzeptieren, die unter (20) sowie noch andere Tilgungen dagegen nicht. Durch die Reihenfolge der möglichen Tilgungen erhalten wir nun Aufschluß über hierarchische Relationen zwischen den Elementen. Ein Element, von dem ein anderes abhängt, läßt sich erst dann tilgen, wenn jenes gestrichen wurde. Da der sehr Mayer ungrammatisch ist, der _ umstrittene Mayer dagegen nicht, folgern wir, daß sehr abhängig ist von umstrittene. Da Mayer auch allein vorkommen kann, die anderen Elemente aber nicht, ist eine Abhängigkeitsrelation zwischen jenen anderen Elementen und Mayer zu konstatieren.

Manche Elemente lassen sich nur gleichzeitig tilgen, wodurch ihre wechselseitige Abhängigkeit deutlich wird. Siehe z.B.:

(21) Der als Playboy bekannte Mayer

*Der als bekannte Mayer

*Der _Playboy bekannte Mayer

*Der _bekannte Mayer

als und *Playboy* sind wechselseitig voneinander abhängig. Sie bilden also eine besondere Gruppe.

Abhängigkeitsrelationen spielen eine große Rolle in der Sprachstruktur. Durch die Tilgungsprobe sind einige von ihnen zu entdecken. Abhängigkeitsbeziehungen sind ein Grund dafür, daß die Strukturen von sprachlichen Ketten nicht allein unter dem Gesichtspunkt der linearen Abfolgerelation beschrieben werden können, wie das Linguisten eine Zeitlang versucht haben. Abhängig von einander und somit besonders zusammenzufassen, sind manchmal auch Elemente, die im Vorher/Nachher durch andere Elemente getrennt sind, sogenannte **diskontinuierliche Konstituenten**:

(22) Karl hat heute nicht geschlafen

Karl schlief heute nicht

Die Kommutation von *schlief* gegen *hat* und die dann notwendige Tilgung von *geschlafen*, weist darauf hin, daß *hat* und *geschlafen* besonders eng zusammengehören.

Eine hervorragende Rolle innerhalb der hierarchisch geordneten Abhängigkeitsrelationen spielen die Verben. Mit Hilfe der Abstrichprobe kann man feststellen, wieviel Positionen bei gegebenem Verb maximal oder minimal in einem Satz besetzt sein müssen. Man bezeichnet diesen Sachverhalt mit dem Terminus **Wertigkeit** oder **Valenz des Verbs**. Vergleiche:

(23) Die Medizin hilft manchen Patienten gegen die Grippe

Die Medizin hilft manchem Patienten _____

Die Medizin hilft _____ gegen die Grippe

Die Medizin hilft _____

Für *helfen* ist also die Wertigkeit 1, 2, 3 anzugeben.

3.6. Die Transformationsprobe

Den Linguisten interessiert nicht nur, welche Syntagmen einer Sprache grammatisch sind und welche nicht, sondern auch, welche Bedeutungsbeziehungen zwischen verschiedenen grammatischen Syntagmen bestehen. Zwei Ketten, die trotz unterschiedlicher Zusammensetzung dieselbe Bedeutung haben, nennt man **Paraphrasen**. Paraphrasenbildung aufgrund rein formaler Operationen ist auch eine wichtige heuristische Methode der Linguistik. Das Prinzip dabei ist das folgende: Wenn man eine bestimmte, genau umrissene Umformung parallel an mehreren Ketten durchführt, müssen bei allen die Bedeutungen identisch bleiben oder sich auf dieselbe Art ändern. Ist das nicht der Fall, so handelt es sich bei jenen Teilen des Ausgangsmaterials, die nach der Umformung in ihrer Bedeutung divergieren oder für die die Umformung gar ungrammatische Ketten ergibt, um Elemente mit unterschiedlicher syntagmatischer Funktion. Gegeben seien folgende Sätze:

(24) Lydia ist leicht zu lieben

(25) Lydie ist bereit zu lieben

(24) läßt sich umformen zu

(24i) Es ist leicht, Lydia zu lieben

(24) und (24i) haben zumindest in einigen Verwendungen, die man sich für diese Sätze vorstellen kann, dieselbe Bedeutung. Die parallele Umformung von (25) ergibt jedoch eine Kette, die in allen Verwendungen entweder ungrammatisch ist oder aber eine von (25) zu unterscheidende Bedeutung hat.

(25i) * Es ist bereit, Lydia zu lieben.

Folglich sind *leicht* und *bereit* sowie die Syntagmen von (24) und (25) unterschiedlich zu beurteilen.

Man kann die Transformationsprobe häufig als operationales Verfahren einsetzen, nach dem man Kategorien definiert. Im Folgenden hat Satz (26) eine zumindest in manchen Verwendungen gleiche Bedeutung wie (26i):

(26) Der Brief wurde von Karl verlesen

(26i) Karl verlas den Brief

Wir können nun eine bestimmte Klasse von Fügungen mit *von* dadurch definieren, daß für sie genau die hier vorgenommene Umformungsmöglichkeit gilt. Zu dieser Klasse gehört dann auch *vom Vater* im folgenden Beispiel:

(27) Fritzchen wurde vom Vater gerügt

(27i) Der Vater rügte Fritzchen

Jedoch gibt es Fügungen mit *von*, die nicht auf diese Weise transformiert werden können:

(28) Der Garten von Herrn Schmidt wurde überflutet

(28i) * Herr Schmidt überflutet den Garten

Hier gilt nicht das Verhältnis *von Herrn Schmidt/Herr Schmidt*, sondern etwa *von Herrn Schmidt/das Herrn Schmidt* gehört. Fügungen mit *von* müssen also je nach Transformationsmöglichkeit verschiedenen paradigmatischen Klassen zugeordnet werden, und damit sind auch die betreffenden Syntagmen als unterschiedlich zu beurteilen.

Wesentlich ist, daß man dabei überhaupt nicht auf Interpretationen wie "Handelnder" oder "Eigentümer" zurückgreifen muß, sondern in Gestalt der Transformationsprobe über einen rein operationalen Nachweis für die intuitiv natürlich schon erahnten Unterschiede verfügt.

Die Mehrdeutigkeit von (29) kann man formal auflösen, indem man zeigt, daß *von Karl* in diesen Syntagma zwei verschiedenen Klassen zugewiesen werden kann:

(29) Es wurde die Briefftasche von Karl gefunden

(29i) Karl hat die Briefftasche gefunden

(29ii) Die Briefftasche, die Karl gehört, wurde gefunden

Allgemein unterscheidet man folgende Arten von Transformationen:

- a) die Permutation, = die lineare Abfolge der Elemente wird verändert,
- b) die Elimination, = bestimmte Elemente werden getilgt,
- c) die Insertion, = bestimmte Elemente werden hinzugefügt,
- d) die Substitution, = ein Element wird durch ein anderes gefunden.

Transformationen sind also spezielle Formen der Permutation und Kommutation. Das besondere daran ist, daß man die Bedeutungen der beteiligten Ketten konstant zu halten bemüht ist (Soweit sich Bedeutungsveränderungen ergeben, die einander genau entsprechen, kann man diese mit einbeziehen, z.B. Umwandlung von Aussagen in Fragen). Außerdem werden in den meisten Fällen der Paraphrasenbildung mehrere Transformationsarten zugleich zur Anwendung kommen. Vergleiche:

(30) Er bat mich, dir davon zu berichten

(30i) Er bat mich, daß ich dir davon berichte

In (30i) ist gegenüber (30) daß, ich insertiert, zu eliminiert und berichten durch berichte substituiert worden.

4. Methoden der Darstellung

Der Linguist muß die Ergebnisse seiner Untersuchung auf irgendeine Weise festhalten. Die Elemente, die Elementemengen, die Relationen müssen bezeichnet werden, damit die Struktur des Sprachsystems erkennbar wird und die Kenntnisse darüber in der Praxis verwertet werden können. Die Darstellung der ermittelten sprachlichen Strukturen erfolgt durch Symbole. Wir haben während unserer bisherigen Erörterungen schon wiederholt Symbole verwendet. In 3.1. und 3.2. haben wir Positionen durch die Buchstaben A,B,C,D bezeichnet. In 3.1. haben wir für die Zusammenfassung von A und B das Symbol X eingeführt. In 3.2. unterschieden wir die Segmente erster Stufe durch Großbuchstaben von denen der zweiten Stufe mit Kleinbuchstaben. Syntagmatische Relationen haben wir gelegentlich angedeutet durch das Hintereinanderschreiben der betreffenden Positionssymbole, z.B. in (17iii) ABC, ABD. Elemente in paradigmatischer Relation haben wir durch geschweifte Klammern zu Klassen zusammengefaßt.

Die verschiedenen linguistischen Symbolsysteme, die heute in Gebrauch sind, können hier nicht im einzelnen dargestellt werden. Sie stehen in engem Zusammenhang zu je speziellen linguistischen Theorien und müssen bei der Beschäftigung mit einer solchen Theorie gelernt werden. Es sollen im Folgenden nur einige Grundsätze erörtert werden.

In einer Strukturbeschreibung muß man Symbole für Elemente, Symbole für Klassen und Symbole für Relationen unterscheiden.

Sprachelemente, die als solche in Äußerungen vorkommen, werden in einer linguistischen Beschreibung meistens kursiv gesetzt. Vergleiche oben *Auto*, *Papagei*, *Hut* und viele andere! Durch die Unterstreichung wird angezeigt, daß es sich einerseits um Elemente der untersuchten Sprache selbst handelt und nicht um Elemente der Beschreibungssprache, zum anderen, daß über sie als Sprachelemente im Rahmen einer linguistischen Erörterung etwas ausgesagt wird und nicht etwa, daß über in der Realität vorkommende Gegenstände wie Autos, Papageien und Hüte etwas gesagt werden soll. Für die Sprache die den Gegenstand der linguistischen Untersuchung bildet, ist der Terminus **Objektsprache** üblich. Die Sprache, in der der Linguist seine Beobachtungen formuliert, nennt man **Metasprache**. Ein Beispiel möge den Unterschied verdeutlichen. In einer Zeitung steht der Satz:

(31) Eine entlaufene Kuh verursachte gestern auf der Autobahn
Stuttgart-Heilbronn einen Unfall

(31) ist eine Äußerung in deutscher Sprache, die vom Linguisten natürlich zum Untersuchungsobjekt gemacht werden kann, die aber in der Zeitung nicht als Beispiel für linguistische Zusammenhänge steht, sondern um über ein Geschehnis auf der Autobahn Mitteilung zu machen. *Kuh* darf in der Zeitung nicht kursiv gesetzt werden, denn nicht das Sprachelement *Kuh* hat ein Unglück verursacht, sondern eine leibhaftige Kuh.

Dagegen schreibt jemand in einem linguistischen Aufsatz:

(32) *Kuh* besteht aus drei Graphemen

(32) ist ein Satz in sprachwissenschaftlicher Metasprache. *Kuh* muß hier kursiv gesetzt werden, denn es steht für ein Sprachelement, über das eine linguistische Aussage gemacht wird. Es wäre unsinnig zu behaupten, daß eine leibhaftige Kuh aus drei Graphemen besteht.

Für Klassen von Elementen führt der Linguist Symbole ein, die die Klassen benennen. An irgendeiner Stelle innerhalb seiner Beschreibung muß er angeben, welche Elemente zu den jeweiligen Klassen gehören, d.h. er muß die Klassensymbole definieren. Das kann auf eine der drei in 2. beschriebenen Weisen geschehen. Das Klassensymbol ist eine Abkürzung für die Menge der Elemente dieser Klasse. Mit Klassensymbolen läßt sich bedeutend leichter umgehen, als wenn man stets alle Elemente der Klassen aufzählen muß. Wenn es auf die Relationen zwischen Klassen ankommt (bei der syntagmatischen Analyse), interessieren deren einzelne Elemente auch nicht unmittelbar, sondern nur die Klassen selbst.

Für die Relationen zwischen Elementen bzw. Klassen benötigt man weitere Symbole. Für Relationen gelten bestimmte mathematisch- logische Gesetzmäßigkeiten, die ihren Niederschlag in den Regeln für die symbolische Darstellung finden. Je nach Art der ermittelten Relation zwischen sprachlichen Einheiten sind verschiedene Darstellungen erforderlich. Aus der symbolischen Darstellung müssen die formalen Eigenschaften der sprachlichen Relationen hervorgehen.

In der gegenwärtigen Linguistik gibt es vor allem zwei Darstellungsweisen, die sich jedoch ineinander überführen lassen: eine eindimensionale und eine zweidimensionale. Bei ersterer werden die Zusammenhänge in Gestalt einer Formel d.h. durch eine lineare Folge von Symbolen nach wohldefinierten Regeln, angezeigt. Bei der zweiten wird in Gestalt eines Graphen auch die vertikale Verteilung der Symbole auf dem Papier zur Darstellung genutzt.

Betrachten wir das Syntagma, das dem Beispielsatz (1) zugrunde liegt:

(1) Diese Nachricht erschütterte ihn

Die Kette (1) ist nur ein Beispiel für das Syntagma. Mit Hilfe der Kommutationsprobe können wir paradigmatische Klassen ermitteln. Das Syntagma läßt sich dann unter Verwendung von Symbolen für die Paradigmenklassen genereller beschreiben. Wählen wir als Symbole für die Paradigmenklassen A,B,C,D und für das Symbol für das Syntagma von (1) S1! Die Abfolgerelation stellen wir durch das Symbol + dar und dadurch, daß wir die Klassensymbole in derselben Reihenfolge niederschreiben, in der die Elemente der betreffenden Klassen in sprachlichen Äußerungen erscheinen, also: A + B + C + D. Wir benötigen nun noch ein Symbol für die Relation, die zwischen S1 und A,B,C,D besteht. Wir wählen dafür das Symbol \rightarrow , das bedeuten soll **konstituiert sich aus**.

Die linguistische Beschreibung des Syntagmas lautet nun so:

- (R1) S1 A + B + C + D
- (R2) A {*diese, eine, jene...*}
- B {*Nachricht, Mitteilung, Beleidigung,...*}
- C {*erschütterte, empörte, verletzte,...*}
- D {*ihn, sie, den Chef,...*}

Die Regeln (R2), die die paradigmatischen Klassen auführen, nennt man **Lexikonregeln**, da sie Elemente aus dem Lexikon, dem Elementevorrat der Sprache, einführen. Geschweifte Klammern und Kommata haben wir oben schon als Symbole für die extensionale Klassifizierung kennengelernt. ...bedeutet, daß die Listen noch nicht abgeschlossen sind.

Wir hatten bei der Verschiebeprobe bemerkt, daß ein besonderer Zusammenhang zwischen A und B besteht, der über die reine lineare Abfolge hinausgeht. Dies können wir in unserer Beschreibung berücksichtigen, indem wir unsere kleine Grammatik wie folgt ändern:

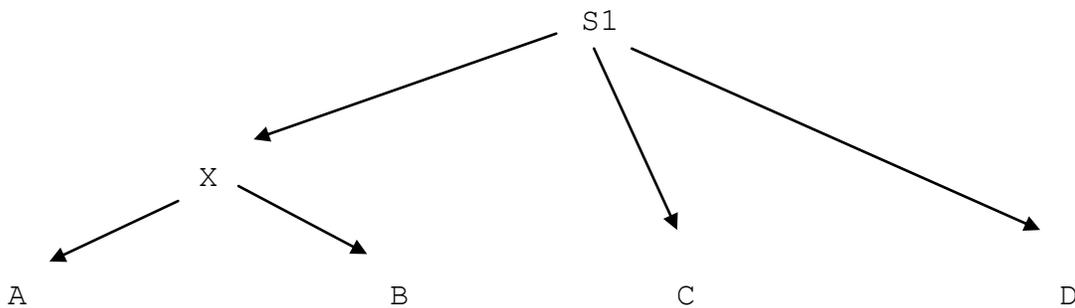
- (R1) S1 X + C + D
- (R2) X A + B
- (R3) wie oben (R2)

Es ist nun gewährleistet, daß - wo auch immer X auftritt - der enge Zusammenhang von A und B berücksichtigt ist.

Wir brechen unsere Erörterung zur Darstellung linguistischer Zusammenhänge durch Formeln hier ab. Es sollte nur das Prinzip klar gemacht werden. Verschiedene Darstellungen in der Literatur variieren darin, welche Konventionen für die verwendeten Symbole vorab getroffen werden.

Man kann denselben Zusammenhang wie oben auch durch einen Graphen beschreiben. Ein Graph ist allgemein ein Gebilde, in welchem Elemente und Elementeklassen durch Punkte (Terminus: **Knoten**) und die zwischen ihnen bestehenden Relationen durch Linien (Terminus: **Kanten**) dargestellt werden. Schreiben wir an die Stelle der Punkte die Symbole für die Klassen, so haben wir einen **interpretierten Graphen**. Schließlich vereinbaren wir noch, daß die Abfolgerelation durch die Richtung links- rechts und die Konstituiert-sich-aus-Relation durch die Richtung oben- unten auf dem Papier dargestellt werden sollen.

Die Struktur des obigen Syntagmas können wir nun durch folgenden Graphen angeben:



Man nennt diese Darstellung auch **Strukturbaum**. Strukturbäume haben den Vorteil, daß man sich mit ihrer Hilfe schnell und einprägsam über die Zusammenhänge in einem bestimmten Syntagma orientieren kann.

Die Verwendung von Formeln und Graphen gibt den grammatischen Beschreibungen der Linguistik ein mathematisches Aussehen, und in der Tat werden die zugrundeliegenden Symbolsysteme auch nach streng mathematischen und logischen Grundsätzen aufgestellt. Der Vorteil derartiger formaler Darstellungen sprachlicher Strukturen besteht darin, daß damit erstens eine abgekürzte Schreibweise gegeben ist, die bei den oft komplizierten Zusammenhängen die Übersicht erleichtert. Weiter können durch die Verwendung von Symbolen generelle Zusammenhänge dargestellt werden, die für eine Vielzahl von konkreten Fällen gelten. Schließlich lassen sich solche Symbolsysteme besonders gut auf ihre innere Widerspruchsfreiheit überprüfen. Gerade dieser letzte Punkt ist ein Grund dafür, daß eine exakte Linguistik ohne Formalisierung undenkbar ist.