

### Hintergrund: Merkmale in der Linguistik

Das Wort 'Merkmal' bedeutet im Prinzip soviel wie 'Eigenschaft' und bezieht sich auf die individuellen Attribute, die ein bestimmtes Objekt aufweist. Objekte können über Mengen von Merkmalen beschrieben werden, und wenn bestimmte Objekte dieselben Merkmale aufweisen, kann man sie ggf. in einer Klasse zusammenfassen. Die Verwendung von Merkmalen erlaubt es, Objekte miteinander zu vergleichen und sie dabei voneinander zu differenzieren oder Klassen von Objekten mit denselben Merkmalen zu bilden.

#### Aufbau und Notation von Merkmalstrukturen

Eine sehr gebräuchliche Notationsform für Merkmalstrukturen bzw. -bündel sind die aus der Wissensrepräsentation bekannten 'Objekt-Attribut-Wert'-Tripel.

Unter 'Objekt' sind physische oder begriffliche Entitäten zu verstehen, Attribute sind die allgemeinen Charakteristika oder Eigenschaften dieser Objekte, Werte sind die spezifischen Beschaffenheiten oder Ausprägungen eines Attributes. Einfaches Beispiel:

Objekt	Merkmal	
	Attribut	Wert
□	Form	quadratisch
	Farbe	weiß
○	Form	kreisförmig
	Farbe	weiß
■	Form	quadratisch
	Farbe	schwarz
●	Form	kreisförmig
	Farbe	schwarz

Der Wertebereich der Attribute 'Form' und 'Farbe' in umfasst in diesem Beispiel jeweils zwei Ausprägungen, nämlich Form: {quadratisch, kreisförmig} und Farbe: {schwarz, weiß}. Die Zuweisung eines konkreten Wertes an ein Attribut ergibt ein sog. 'Attribut-Wert-Paar', z.B. [Farbe weiß] oder [Form quadratisch].

Auf der Basis der Zuordnung von Attribut-Wert-Paaren zu Objekten können Klassen gebildet werden, beispielsweise die Klasse von Objekten, die kreisförmig sind, die Klasse von Objekten, die schwarz sind usw. Die Menge aller einem Objekt zugeordneten Attribut-Wert-Paare ergibt dessen Merkmalstruktur, die in eckigen Klammern notiert wird, für die oa. Objekte also beispielsweise



Ein interessanter Aspekt dabei ist die Möglichkeit, auch verschachtelte oder **komplexe Merkmalsstrukturen** zu bilden derart, dass der Wert eines Attributes selber wieder eine

Merkmalstruktur ist, wie z.B. in dem folgenden Beispiel, in dem Attribut X ein Wert zugeordnet ist, der selber ein Merkmalsbündel ist:



Auch in der Linguistik dienen Merkmale dazu, Objekte zu charakterisieren und Klassen zu bilden. In den folgenden Abschnitten soll dargestellt werden, wie solche Merkmale aufgebaut sind und welche Funktionen sie für die Repräsentation von Information auf den verschiedenen sprachlichen Ebenen (Phonetik, Morphologie, Syntax, Semantik) haben können.

#### Phonetische Merkmale

Ein anschauliches Beispiel für die Klassifizierung von Objekten über Merkmale stammt aus dem Bereich der Phonetik. Hier spielen Merkmale traditionell eine wichtige Rolle und die Versuche, die Laute und Lautsysteme der menschlichen Sprache(n) auf der Ebene von Merkmalen zu repräsentieren, reichen bereits in die 30er Jahre des letzten Jh. zurück.

In der artikulatorischen Phonetik geht es um die physiologischen Prozesse, die sich bei der Artikulation vollziehen: Sprachlaute werden beschrieben mit Bezug auf die Stellung der Artikulationsorgane bei der Lautproduktion. So basiert z.B. die im internationalen phonetischen Alphabet (IPA) verwendete Klassifikation von Lauten auf Artikulationsmerkmalen. Als Attribute für diese Merkmale können z.B. die an die Produktion von Lauten beteiligten Prozesse (Lufstromprozess, Phonationsprozeß, oro-nasaler Prozeß, Artikulationsprozeß) dienen, die für einen spezifischen Laut jeweils unterschiedliche Werte aufweisen.

Als Beispiel nehmen wir zunächst [p] (wie in *Pein*) und [b] (wie in *Bein*). Sie haben bedeutungsunterscheidende Funktion (*Pein* und *Bein* bedeuten jeweils etwas anderes), sind sich allerdings ähnlich: es sind beides Bilabiale, d.h. in beiden Fällen sind bei der Lautbildung die Lippen geschlossen. Die Luft staut sich zunächst hinter diesem oralen Verschuß, um dann, wenn die Lippen geöffnet werden und der Verschuß gelöst wird, quasi explosionsartig zu entweichen. Es gibt aber einen wesentlichen Unterschied zwischen [b] und [p], der mit der Stellung der Glottis (Stimmritze) zu tun hat: beim [b] geraten die Stimmlippen in Schwingung, es ist also ein stimmhafter Laut, während das [p] ein stimmloser Laut ist. Der Unterschied zwischen den beiden liegt also in der Ab- oder Anwesenheit der Eigenschaft, stimmhaft zu sein, was wie folgt dargestellt werden kann:

STIMMHAFT	STIMMLOS
[b]	[p]

Über die Eigenschaft bzw. das Merkmal, stimmhaft zu sein oder nicht, können [p] und [b] voneinander unterschieden werden, und da diese beiden Laute im Deutschen eine bedeutungsunterscheidende Funktion haben, wird dieses Merkmal als 'distinktives Merkmal' bezeichnet.

Wäre die Stimmhaftigkeit kein distinktives Merkmal, so wären *Bein* und *Pein* nur verschiedene Aussprachevarianten ein- und desselben Wortes (was sie in manchen Dialekten ja auch sind).

Nehmen wir einen weiteren Laut: das [m], wie z.B. in *Main*. Wie bei [b] und [p] auch sind bei der Bildung dieses Lautes die Lippen geschlossen. Er unterscheidet sich dadurch, dass beim [m] das Velum derart gesenkt ist, dass die Luft durch die Nase entweichen kann.

Laute, die diesbezüglich dieselbe Charakteristik aufweisen wie das [m], werden 'Nasale' genannt. Laute, die dieselbe Charakteristik aufweisen wie [p] und [b] sind 'orale Laute'. Über die Merkmale Stimmhaft/Stimmlos und Nasal/Oral können die bilabialen Laute [p], [b] und [m] wie folgt in der folgenden Merkmalsmatrix angeordnet werden.

	STIMMHAFT	STIMMLOS
ORAL	[b]	[p]
NASAL	[m]	

Hier kann man erneut sehen, wie Merkmale eingesetzt werden, um Klassen zu bilden. So gehören [b] und [m] zur Klasse der stimmhaften Laute, [b] und [p] zu den oralen Lauten usw. Auf den einzelnen Laut bezogen, können diesen jeweils Merkmalsbündel wie folgt zugeordnet werden:

$$[m]: \begin{bmatrix} \text{STIMMHAFT} \\ \text{NASAL} \end{bmatrix} \quad [b]: \begin{bmatrix} \text{STIMMHAFT} \\ \text{ORAL} \end{bmatrix} \quad [p]: \begin{bmatrix} \text{STIMMLOS} \\ \text{ORAL} \end{bmatrix}$$

Dieses ist wie folgt zu lesen: Der Laut [m] hat die Merkmale [STIMMHAFT] und [NASAL]. Der Laut [b] hat die Merkmale [STIMMHAFT] und [ORAL] usw.

Wenn wir davon ausgehen, dass die an die Produktion von Lauten beteiligten Prozesse (Luftstromprozess, oro-nasaler Prozeß usw.) für einen spezifischen Laut Attributstatus haben, hätten wir mit Bezug auf Eigenschaften wie 'Stimmhaftigkeit' und 'Nasalität' die folgenden Attribute samt Wertebereichen:

Glottisstellung {stimmhaft, stimmlos}

Velumstellung {oral, nasal}

Merkmale, bei denen der Wertebereich des Attributes nicht mehr als zwei Elemente umfasst, werden auch als binäre Merkmale bezeichnet. Bei den durch die Werte STIMMHAFT/STIMMLOS und ORAL/NASAL gekennzeichneten Klassen handelt es sich um binäre Oppositionsklassen. Das heißt, dass ein bestimmter Laut entweder der einen Klasse angehört oder der anderen: die Klasse der nasalen Laute enthält alle Laute, die nicht zur Klasse der oralen Laute gehören und umgekehrt, und die Klasse der stimmhaften Laute enthält alle Laute, die nicht zur Klasse der stimmlosen Laute gehören und umgekehrt. Logisch kann dieser Sachverhalt wie folgt ausgedrückt werden:

$$\neg \text{NASAL} \Leftrightarrow \text{ORAL} \wedge \neg \text{ORAL} \Leftrightarrow \text{NASAL}$$

$$\neg \text{STIMMHAFT} \Leftrightarrow \text{STIMMLOS} \wedge \neg \text{STIMMLOS} \Leftrightarrow \text{STIMMHAFT}$$

Bekanntermaßen wird in solchen Fällen häufig ein Element aus dem Wertebereich herausgepickt und entweder positiv oder negativ spezifiziert, statt [NASAL] / [ORAL] also beispielsweise [+NASAL] / [-NASAL]. Umgesetzt für die Beispiellaute sieht das wie folgt aus:

$$[m]: \begin{bmatrix} + \text{STIMMHAFT} \\ + \text{NASAL} \end{bmatrix} \quad [b]: \begin{bmatrix} + \text{STIMMHAFT} \\ - \text{NASAL} \end{bmatrix} \quad [p]: \begin{bmatrix} - \text{STIMMHAFT} \\ - \text{NASAL} \end{bmatrix}$$

Hier wurde von einem Oppositionspaar ein Merkmal stellvertretend für beide verwendet und durch ein vorangestelltes Plus- oder Minuszeichen entweder positiv oder negativ

gekennzeichnet.<sup>1</sup> Soll ein Merkmal un spezifiziert bleiben, kann es mit einem '±' versehen werden: [±NASAL], [±STIMMHAFT].

Über [p], [b] und [m] ist weiter oben ausgesagt worden, dass sie zur Klasse der bilabialen Laute gehören, d.h. dass bei ihrer Bildung als aktiver Artikulator die Unterlippe; als passiver Artikulator die Oberlippe beteiligt sind. Aktiver und passiver Artikulator sind ebenfalls Eigenschaften, die zur Charakterisierung von Lauten dienen. Im Unterschied zur Stellung des Velums bzw. zur Stellung der Glottis aber umfassen die jeweiligen Wertebereiche mehr als nur zwei Elemente, z.B:

passiver Artikulator {labial, dental, alveolar, velar} Im Unterschied zu einem Merkmal wie [±NASAL] gibt es mehr als nur eine Alternative: ein Laut, der nicht labial ist, kann dental sein (d.h. er wird mit der Unterlippe und den Zähnen gebildet, wie die Laute [f] und [v]), alveolar (d.h. er wird mit dem Zungenblatt und dem Zahndamm gebildet, wie z.B. [t] und [d]), velar (d.h. er wird mit dem Zungenrücken und dem Gaumen gebildet, wie beispielsweise [k] und [g]) usw. Ohne an dieser Stelle zu stark ins Detail zu gehen, können verschiedene Laute wie folgt in eine 'passive-Artikulator-Tabelle' eingeordnet werden:

LABIAL	DENTAL	ALVEOLAR	PALATAL	VELAR
p,b,m	f,v	t,d,n	ʃ,ʒ	k,g,ŋ

Ein Laut ist entweder nasal oder nicht-nasal. Natürlich kann man eine analoge Aussage auch für einen Artikulationsort formulieren: ein Laut ist entweder labial oder nicht-labial, entweder dental oder nicht-dental usw.

Während aber Elemente mit dem Merkmal [-NASAL] eine Oppositionsklasse zu den [+NASAL]-Elementen aufweisen, ist dieses bei den negativ markierten Merkmalen den passiven Artikulator betreffend nicht der Fall. Außer [p], [b] und [m] alle aufgeführten Konsonanten (also f, v, t, d, n, ʃ, ʒ, k, g und ŋ das Merkmal [-LABIAL], doch sie bilden selber keine eigene Klasse, sondern unterteilen sich in die Klassen der labiodentalen Laute, der alveolaren Laute usw.

Dieser Sachverhalt kann ebenfalls logisch dargestellt werden, zur Rekapitulation nochmal die Darstellung für [±NASAL] und [±STIMMHAFT]

$$\neg \text{NASAL} \Leftrightarrow \text{ORAL} \quad , \quad \neg \text{STIMMHAFT} \Leftrightarrow \text{STIMMLOS}$$

Zum Vergleich nun

$$\neg \text{LABIAL} \Rightarrow \text{DENTAL} \vee \text{ALVEOLAR} \vee \text{PALATAL USW.},$$

was soviel heißt wie 'nicht-labial impliziert dental oder alveolar oder palatal usw.' Während also die Negation von NASAL bzw. STIMMHAFT entweder ORAL oder STIMMLOS impliziert und umgekehrt, impliziert die Negation von LABIAL die Disjunktion der anderen Artikulations-Merkmale.

Soviel zunächst zu Merkmalen und Merkmalsbündeln in der Phonetik. Die Frage nach der Binarität von Merkmalen wird im Zusammenhang mit der Klassenbildung bei den semantischen Merkmalen wieder auftreten.

<sup>1</sup> Anstatt für die NASAL/ORAL Opposition [±NASAL] zu verwenden und [ORAL] aufzugeben, hätte man vielleicht auch [±ORAL] nehmen und [NASAL] aufgeben können. In manchen Fällen ist die Entscheidung zwar begründet, in manchen aber auch umstritten (was ist 'besser': [±MÄNNLICH] oder [±WEIBLICH]?)

**Morphologische Merkmale**

Im Bereich der Flexionsmorphologie werden Merkmale bekanntermaßen verwendet, um Wortformen über deren (sekundäre) grammatische Kategorien zu beschreiben. Eine Wortform wie *füttert* kann exemplarisch wie folgt charakterisiert werden:

füttert :	KATEGORIE : VERB	Dieses Merkmalsbündel besteht aus Angaben über die lexikalische Kategorie der entsprechenden Wortform, und über deren Markierung hinsichtlich Person, Numerus, Tempus und Modus.
	PERSON : 3	
	NUMERUS : SINGULAR	
	TEMPUS : PRÄSENS	
	MODUS : IINDIKATIV	

Auch hier bestehen Merkmale aus der bestimmten Wertes für dieses Attribut. Dieser Wert entstammt einem festgelegten Wertebereich oder Wertevorrat. Für das Attribut KASUS umfaßt der Wertebereich in der deutschen Sprache die Elemente {NOMINATIV, AKKUSATIV, GENITIV , DATIV}, für das Attribut NUMERUS umfaßt der Wertevorrat die Elemente {SINGULAR, PLURAL} usw.

**Merkmale in der Wortartenklassifikation**

Nicht nur die Wortformen, sondern, auf einer etwas theoretischeren Ebene, auch die Wortarten selber können durch abstrakte Merkmale charakterisiert werden. Traditionelle Wortklassen wie 'Nomen' oder 'Verb' oder 'Adjektiv' sind monolithische Konstrukte, über die bestimmte Gemeinsamkeiten nicht direkt erfaßt werden – beispielsweise Gemeinsamkeiten zwischen Nomina und Adjektiven bezüglich ihrer Flexionseigenschaften oder Gemeinsamkeiten zwischen Verben und Adjektiven hinsichtlich ihrer prädzierenden Funktion.<sup>2</sup> Darüberhinaus eignet sich diese Art von Klassifizierung häufig nicht, um im Sprachvergleich sinnvoll eingesetzt zu werden. In der modernen Sprachwissenschaft gibt es entsprechend Bestrebungen, diese monolithischen Konstrukte quasi aufzubrechen und in Form von Merkmalsbündeln zu erfassen derart, dass eine genauere Spezifizierung ermöglicht wird, die die Gemeinsamkeiten und Unterschiede der Kategorien besser beschreibt. Einige Versuche sind ziemlich widersprechend, andere weniger, auf die genauen Hintergründe gehen wir hier nicht ein sondern zeichnen diesen Punkt nur exemplarisch an einem informellen Beispiel nach (das sich allerdings an einem nicht so vielversprechenden Ansatz orientiert, es geht hier um's Prinzip).

Wir verwenden zwei binäre Merkmale, nämlich N {+,-} und V {+,-}, wobei 'N' für Nominalität, 'V' für Verbalität steht. Damit können wir vier Merkmalsbündel bilden:

- 1.  $\begin{bmatrix} +N \\ -V \end{bmatrix}$     2.  $\begin{bmatrix} -N \\ +V \end{bmatrix}$     3.  $\begin{bmatrix} +N \\ +V \end{bmatrix}$     4.  $\begin{bmatrix} -N \\ -V \end{bmatrix}$

Diesen Merkmalsbündeln könnten wir wie folgt die 'traditionellen' Lexemklassen zuordnen: 1: Nomina, 2: Verben, 3: Adjektive 4: Partikel. Das ermöglicht es z.B., Aussagen wie die folgenden über das Deutsche zu treffen:

- Diejenigen Klassen, deren Attribut N positiv spezifiziert ist, werden bezüglich Genus flektiert (trifft zu auf 1 (Nomina) und 3 (Adjektive))

- Diejenigen Klassen, deren Attribut V positiv spezifiziert ist, können prädikativ verwendet werden (trifft zu auf 2 (Verben) und 3 (Adjektive)<sup>3</sup>).

Wenn wir die Präpositionen unter Partikel zählen, können wir z.B. für das Englische eine Aussage treffen wie

- Nur diejenigen Klassen, deren Attribut N negativ spezifiziert ist, weisen Kasus zu (trifft zu auf 2 (Verben) und 4 (Partikel, hier: Präpositionen)).

**Merkmale in der Syntax**

**A. Valenzmerkmale**

Ein weiteres Einsatzfeld von Merkmalen im Zusammenhang mit Lexemen besteht in Angaben darüber, in welchem Kontext dieses Lexem auftreten kann oder – anders ausgedrückt – inwiefern dieses Lexem die Umgebung, in der es auftritt, vorstrukturiert. Ein Satz wie \* *Sie schenkte ihrer Schwester* ist deshalb nicht ok, weil etwas fehlt: das Verb *schenken* tritt, von Ausnahmen wie in *Wir schenken gerne* abgesehen, in dem Kontext [NP \_\_ NP NP] auf, im Beispielsatz fehlt also eine NP. Hier geht es also um die Valenzeigenschaften einzelner Lexeme, die in der Linguistik diese häufig in Form von Prädikat-Argument-Strukturen repräsentiert sind. Die Verben *run*, *kick* und *put* hätten jeweils die folgenden Prädikat-Argument-Strukturen: *run(x)*, *kick(x,y)* und *put(x,y,z)*. d. h. *run* hat ein Argument, *kick* hat zwei Argumente und *put* hat drei. Die in Form von Prädikat-Argument-Strukturen notierten Valenz- oder Subkategorisierungseigenschaften einzelner Lexeme nehmen in vielen zeitgenössischen Grammatikmodellen eine zentrale Position ein. Wenn wir uns nur auf die Zahl der Argumente konzentrieren, können wir diese Eigenschaften Bezug auf die Beispiele wie folgt notieren:

*run* [Valenz 1], *kick* [Valenz 2], *put* [Valenz 3]

Diese Merkmale können dazu dienen, Unterklassen von Lexemklassen zu bilden. So entsprechen Verben mit dem Merkmal [Valenz 2] den transitiven Verben (wie *treten*, *küssen*, *lesen*), Verben mit dem Merkmal [Valenz 1] den intransitiven Verben (*schlafen*, *tanzen*, *husten*) usw. Merkmale dieser Art werden demgemäß auch als 'Subkategorisierungsmerkmale' (s.u. im Abschnitt über Subkategorisierung) bezeichnet.

Darüberhinaus lässt sich aber noch genauer klassifizieren, um was für Argumente es sich dabei handelt, beispielsweise mit Bezug auf syntaktische Kategorien, grammatische Funktionen oder semantische Relationen. An dieser Stelle sei daran erinnert, dass der Wert eines Merkmals selber wieder eine Merkmalsstruktur sein kann. Wenn wir die Verben *kick* und *live* nehmen, hätten beide eine zweistellige Prädikat-Argumentstruktur: *kick(x,y)* und *live(x,y)*. Sie unterscheiden sich unter anderem darin, dass das zweite Argument unterschiedlichen kategorialen Status hat, nämlich eine NP in bei *kick* und eine PP bei *live*. Diese Information kann wie folgt in einer komplexen Merkmalsstruktur erfasst werden:



**B. Merkmalbasierte Grammatikformalismen**

Im Bereich der neueren Grammatikforschung spielen Merkmale als Notationsform für die Repräsentation syntaktischer und semantischer Daten eine immer gewichtigere Rolle. So

<sup>2</sup> Das ist nicht ganz astrein, insofern auch appellative Nomina prädzierende Funktion haben, da sie Zugehörigkeit zu einer Klasse von Objekten (Hunde, Tisch usw.) denotieren.

<sup>3</sup> über die Kopula vermittelt

basieren Grammatikmodelle aus der Familie der Unifikationsgrammatiken, wie z.B. die LFG (= Lexikalisch–Funktionale Grammatik) oder die HPSG (= Head Driven Phrase Structure Grammar), auf der Verwendung komplexer Merkmalsstrukturen für die Darstellung der in einem Satz enthaltenen syntaktisch/ semantischen Information.

Der Einsatz von Merkmalsstrukturen ist ein essentielles Moment in einer Unifikationsgrammatik (fortan UG). In einer UG geht es ua. darum, über die Kombination von einzelnen Merkmalsstrukturen eine Gesamtstruktur z.B. eines Satzes zu erzeugen. Die Kombinationsmöglichkeit für Merkmalsstrukturen unterliegt dabei ganz bestimmten Regularien: es können nur solche Strukturen unifiziert, also vereinigt werden, die nach bestimmten formalen Voraussetzungen miteinander kompatibel sind. Ohne zu detailliert auf diese Voraussetzung einzugehen, soll ein ganz einfaches Beispiel kurz erläutern, wie in diesem Rahmen gearbeitet wird. Nehmen wir zunächst die Merkmalstrukturen für die Wortformen *das* und *Haus*, die problemlos miteinander vereinigt werden können:

$$\text{das: } \left[ \begin{array}{l} \text{DEF : +} \\ \text{NUM : SG} \\ \text{KAS : \{NOM, AKK\}} \\ \text{GEN : NEUT} \end{array} \right] \quad \text{Haus: } \left[ \begin{array}{l} \text{NUM : SG} \\ \text{KAS : \{NOM, AKK\}} \\ \text{GEN : NEUT} \end{array} \right] \quad \text{das} \sqcup \text{ Haus} \left[ \begin{array}{l} \text{DEF : +} \\ \text{NUM : SG} \\ \text{KAS : \{NOM, AKK\}} \\ \text{GEN : NEUT} \end{array} \right]$$

Die Attribut–Wert–Paare haben jeweils die gleiche Form, die Frage, ob Kasus nominativ oder akkusativ ist, kann hier nicht entschieden werden. Der einzige Unterschied zwischen den Merkmalsbündeln besteht im Merkmal [DEF:+] beim Determinator. Dieses wird in das Bündel für die NP *das Haus* übernommen, die definite Referenz hat.

Im Gegensatz dazu die Merkmalsbündel für *ein* und *Hunde*:

$$\text{ein: } \left[ \begin{array}{l} \text{KAT : DET} \\ \text{DEF : -} \\ \text{NUM : SG} \\ \text{KAS : NOM} \end{array} \right] \quad \text{Hunde: } \left[ \begin{array}{l} \text{KAT : N} \\ \text{NUM : PL} \\ \text{KAS : \{NOM, GEN\}} \end{array} \right]$$

Hier gibt es Probleme, denn der Wert des Attributs NUMERUS ist in den beiden Strukturen unterschiedlich. Das bedeutet, dass sie nicht miteinander unifiziert werden können.

In einer ausgearbeiteten Form sind die Merkmalstrukturen in den oa. Grammatiktypen äußerst komplex, da sie Information aus den verschiedensten sprachlichen Ebenen (Phonologie, Syntax, Semantik) umfassen können.

Zum Abschluß dieses Abschnittes und zur Illustration des letzten Punktes wollen wir eine Merkmalsstruktur betrachten, wie sie im Rahmen der LFG zu finden ist.

Es geht um die Repräsentation des Satzes *A girl handed the baby a toy*. Der Merkmalsstruktur zu entnehmen sind unterschiedliche Daten, z.B. dass 'a girl' die Subjektsfunktion hat, dass das Tempus des Satzes Präteritum ist, dass die NP 'the baby' definite Referenz hat usw

$$\left[ \begin{array}{l} \text{SUBJ} = \left[ \begin{array}{l} \text{SPEC} = [-\text{DEF}] \\ \text{NUM} = \text{SG} \\ \text{PRED} = \text{'GIRL'} \end{array} \right] \\ \text{TENSE} = \text{PAST} \\ \text{PRED} = \text{'HAND'} < (\uparrow \text{SUBJ}) (\uparrow \text{OBJ2}) (\uparrow \text{OBJ}) > \\ \text{OBJ} = \left[ \begin{array}{l} \text{SPEC} = [+ \text{DEF}] \\ \text{NUM} = \text{SG} \\ \text{PRED} = \text{'BABY'} \end{array} \right] \\ \text{OBJ2} = \left[ \begin{array}{l} \text{SPEC} = [-\text{DEF}] \\ \text{NUM} = \text{SG} \\ \text{PRED} = \text{'BALL'} \end{array} \right] \end{array} \right]$$

Ein wichtiger Aspekt von Merkmalsstrukturen in der Syntax liegt in dem Umstand, dass bestimmte Attribute ihre Werte teilen können, d.h. dass die Information darüber, welchen Wert ein bestimmtes Merkmal aufweist, ggf. erst über die Unifikation mit einem anderen Merkmal geliefert wird. Dazu gleich, im 'praktischen' Teil, noch mehr Information.

### Semantische Merkmale

Zum guter Letzt kurz einige Merkmale aus dem Bereich der lexikalischen Semantik. Zum Einstieg die folgenden Beispiele.

?*Diese Bücher wachsen schnell.*                      ?*Hunde lesen in der Regel nur sehr ungern.*

Beide Sätze sind mit Fragezeichen markiert, da sie, wenn sie wörtlich verstanden werden sollen, semantisch anomal sind. In beiden Fällen ist eigentlich klar, was das Problem ist: Bücher wachsen genausowenig, wie Hunde lesen. Die Kombinationen |*Bücher—wachsen*| und |*Hunde—lesen*| müssen also irgendwie als besonders gekennzeichnet werden. Betrachten wir die weiteren Beispiele:

$$\left. \begin{array}{l} * \text{Lautsprecher} \\ * \text{Zahnspangen} \\ * \text{Tische} \\ \text{Kühe} \\ \text{Amöben} \\ \text{Buchen} \end{array} \right\} \text{wachsen sehr schnell} \quad \left. \begin{array}{l} * \text{Tischdecken} \\ * \text{Bienen} \\ * \text{Augenringe} \\ \text{Kinder} \\ \text{Männer} \end{array} \right\} \text{lesen ungern}$$

Die Gemeinsamkeit der Substantive, die mit *wachsen* kombiniert werden können, besteht in der Tatsache, dass sie zu einer Klasse gehören, die belebte Objekte bezeichnen, die anderen Substantive dagegen nicht. Im Falle von *lesen* reicht aber das Kriterium belebt–unbelebt nicht aus; denn auch Hunde und Bienen sind belebt, können aber nichtsdestotrotz nicht lesen. Nicht anomal sind die Formen nur dann, wenn das entsprechende Subjekts–Substantiv zur Klasse der Menschen gehört. Über die folgende Merkmalsmatrix können die Objekte entsprechend klassifiziert werden:

	BELEBT	MENSCHLICH
<i>Lautsprecher</i>	–	–
<i>Tische</i>	–	–
<i>Kühe</i>	+	–
<i>Buchen</i>	+	–
<i>Kinder</i>	+	+
<i>Männer</i>	+	+

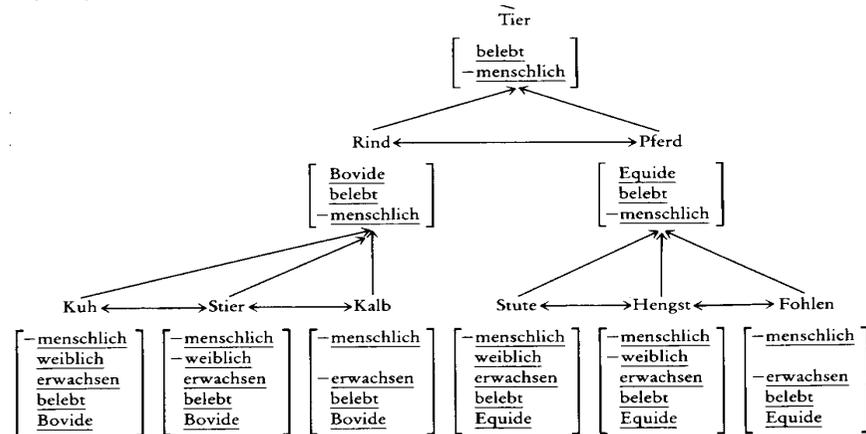
Auch hier handelt es sich jeweils um binäre Oppositionsklassen, d.h. wir können die Merkmale [±BELEBT] und [±MENSCHLICH] herausbilden. Diese stehen in einem interessanten Abhängigkeitsverhältnis zueinander: jedes Objekt mit dem Merkmal [+MENSCHLICH] wird auch das Merkmal [+BELEBT] haben.

Ein Merkmalsbündel wie \*  $\begin{bmatrix} - \text{BELEBT} \\ + \text{MENSCHLICH} \end{bmatrix}$  ist ausgeschlossen. Das Merkmal [+BELEBT] wird deshalb im Zusammenhang mit dem Merkmal [+MENSCHLICH] auch als 'redundantes Merkmal' bezeichnet.

Um auf die oa. Beispielsätze zurückzukommen: es müssen für *wachsen* bzw. *lesen* Regeln formuliert werden mit Angaben, die – für eine nicht-anomale Lesart – gewährleisten, dass sie nur mit Objekten mit entsprechenden Merkmalen verknüpft werden. Solche Angaben werden 'Selektionsbeschränkung' genannt und können beispielsweise die folgende Form haben:

*lesen*: [NP<sub>[+menschlich]</sub> \_\_\_ NP] oder *wachsen*: [NP<sub>[+belebt]</sub> \_\_\_]

Einzelne Lexeme können über diese Art von Merkmalen semantisch (intensional) klassifiziert und voneinander differenziert werden. Ein Beispiel dafür ist die Klassifikation bestimmter Tierbezeichnungen (Quelle: Grewendorf u.a. *Sprachliches Wissen*. 1989 Frankfurt: Suhrkamp), in der binäre (menschlich, weiblich) und nicht-binäre (Bovide, Equide) Merkmale vertreten sind:



Attraktiv an solchen Merkmalsstrukturen ist z.B. die Tatsache, dass man sie in der Beschreibung von semantischen Relationen wie z.B. Hyponymie, Synonymie, Komplementarität usw. einsetzen kann (beispielsweise 'erbt' die Merkmalsstruktur eines Hyponyms alle Merkmale seines Hyperonyms, wie die oa. Graphik zeigt) und somit zu generalisierenden Aussagen über semantische Beziehungen kommt. Problematisch daran ist unter anderem, dass die 'Grenzen' zwischen einzelnen Begriffen häufig nicht so scharf sind, wie es eine derartige Klassifikation suggeriert, als Beispiel diene die begriffliche Differenzierung zwischen 'Tasse' und 'Becher' – was könnte hier das differenzierende Merkmal sein?

### PS-Grammatiken: Übergenerierung, Subkategorien und Merkmale

Einer traditionelle, kontextfreie Mini-PS-Grammatik wie die folgende

#### Grammatik A

- S → NP VP
- NP → Det N
- VP → V NP
- Det → {the, some}
- N → {dog, cat}
- V → {bit}

beschreibt genau 16 wohlgeformte Sätze des Englischen (*the dog bit the cat, some cat bit the dog, the dog bit some cat* usw.) Wenn das Lexikon um Einträge erweitert wird, z.B. um das Nomen *horse*, den Det *my* und und die Verben *loved* und *hated* erhöht sich diese Zahl drastisch, allerdings nur um solche Formen, die genau dieselbe Struktur aufweisen wie die genannten Beispiele.

Um auch andere Strukturen zu integrieren, müssen die Regeln für die Konstituentenklassen modifiziert werden – die Klasse VP hat auch andere Realisierungen als V–NP; die Klasse NP hat auch andere Realisierungen als Det–N; z.B. in *some rather sweet cat, the dog with bad breath* usw. Für die Integration solcher Formen kann die NP-Regel wie folgt modifiziert werden:

NP → Det (AP) N (PP)

Allerdings muss die Grammatik dann auch noch um Information bezüglich der Struktur von Adjektiv- bzw. Präpositionalphrasen erweitert werden und Adjektive, Adverbien und Präpositionen müssen im Lexikon aufgeführt sein:

- AP → (Adv) A
- PP → P NP
- A → {ugly, sweet}
- Adv → {very}
- P → {with}

Um auch andere Formen von Verbalphrasen, z.B. solche mit keiner oder zwei nominalen Ergänzungen und andere mit der Grammatik beschreiben zu können, kann auch die VP-Regel geändert werden, z.B.

VP → V (  $\left\{ \begin{array}{l} \text{NP} \\ \text{NP PP} \\ \text{eS} \end{array} \right\}$  )

Dieser Regel entsprechen ausbuchstabiert 4 VP-Regeln:

- VP → V
- VP → V NP
- VP → V NP PP
- VP → V eS

'eS' steht für 'eingebetteter Satz', z.B. in einem Satz wie *The cat knew that some dog bit the horse*. Eine Regel für diese Kategorie und der Lexikoneintrag für die Konjunktion könnten so aussehen:

ES → Conj S

Conj → *that*

Das Resultat solcher Modifikationen ist eine Erweiterung der oa. Grammatik A (inklusive Lexikon) wie folgt:

**Grammatik B**

S → NP VP

NP → Det (AP) N (PP)

VP → V (  $\left\{ \begin{array}{l} \text{NP} \\ \text{NP PP} \end{array} \right\}$  )  
 eS

AP → (Adv) A

PP → P NP

eS → Conj S

Det → {*the, these, my, some, a ...*}

N → {*girl, cat, dog, boy, tree, garden ...*}

V → {*bit, kicked, knew, put, slept...*}

Adv → {*rather, very...*}

A → {*ugly, boring, sweet, nice...*}

P → {*in, with, on, near ...*}

Conj → {*that, whether...*}

Ein fundamentales Problem solcher kontextfreier PS-Grammatiken liegt nun darin, dass sie *übergenerieren* – das heißt, dass nicht nur wohlgeformte, sondern auch nicht-wohlgeformte Sätze aus ihnen abgeleitet werden können, aus der oa. Grammatik z.B. *\*the boy kicked that the cat bit* oder *\*these dog slept* und viele mehr. Die Frage ist, wie man einem solchen Problem begegnet. Ein traditionelles Verfahren, um dieses Problem in den Griff zu bekommen, ist die sogenannte **Subkategorisierung**. Bei den Verben und den diversen Subklassen Vi, Vt, Vt2 usw. haben wir diese Methode bereits in ihren Grundzügen kennengelernt. Zu diesem Punkt noch ein weiteres Beispiel auf Deutsch:

Die folgende Mini-Grammatik lässt nicht nur die NP unter a) sondern auch die unter b) als wohlgeformt zu, da die notwendige Kongruenz zwischen Det und N nicht ausgedrückt ist:

NP → Det N

Det → {*der, die, das*}

N → {*Frau, Kind, Mann*}

a) *die Frau*

*der Mann*

*das Kind*

b) *\*der Frau, \*das Frau*

*\*die Mann, \*das Mann*

*\*die Kind, \*der Kind*

Wenn man in Lexikon und in der Grammatik mit monolithischen Kategorien wie N oder VP oder Det operiert, benötigt man, um dieses Problem in den Griff zu bekommen, eine immens große Zahl von Subkategorien und darauf bezogene kontextfreie Regeln. Für das vorliegende Beispiel könnte die Grammatik in dieser Weise wie folgt modifiziert werden, damit die nicht-wohlgeformten Syntagmen ausgefiltert werden:

Det<sub>1</sub> → *der*      N<sub>1</sub> → *Mann*      NP → Det<sub>1</sub> N<sub>1</sub>

Det<sub>2</sub> → *die*      N<sub>2</sub> → *Frau*      NP → Det<sub>2</sub> N<sub>2</sub>

Det<sub>3</sub> → *das*      N<sub>3</sub> → *Kind*      NP → Det<sub>3</sub> N<sub>3</sub>

Eine solche Analyse ist sehr unbefriedigend.

Hier wird suggeriert, dass wir es bei den Artikeln mit distinkten Kategorien zu tun haben – anstelle von Det<sub>1</sub>, Det<sub>2</sub> und Det<sub>3</sub> hätten wir als Kategorienbezeichnungen ja genau so gut X, Y und Z nehmen können, d.h. das Gemeinsame dieser Elemente bezüglich ihrer formalen und semantisch/funktionalen Eigenschaften wird in der Grammatik nicht mehr erfasst: drei unterschiedliche Klassen, drei unterschiedliche Sorten Maus.

Der Effekt davon ist, dass die Anzahl der NP- Regeln unnötig aufgebläht wird. Die oben unter a) aufgeführten NP haben durchgehend dieselbe Struktur, in der ein Nomen auf einen Determinator folgt. Auch hier verursachen die Subkategorien einen Generalisierungsverlust, da wir jetzt für jede wohlgeformte Konstruktion eine eigene Regel aufstellen müssen. Worauf es hier also ankäme, wäre, herauszustellen, dass in einer wohlgeformten NP auf die Übereinstimmung des Genusmerkmals von Artikel und Substantiv zu achten ist, ohne dafür zig neue Kategorien und Regeln einzuführen.

Eine Möglichkeit, dieses zu bewerkstelligen, besteht darin, die Elemente der Grammatik um **Merkmale** zu erweitern und dann restringierende Angaben darüber zu machen, welche Merkmalskombinationen zulässig sind (und welche nicht).

Derartige Angaben werden auch mit dem englischen Terminus *Feature Constraints* bezeichnet.

Anstelle also im oa. Beispiel neue Klassen einzuführen, behalten wir Det und N bei, erweitern aber das Lexikon um ein Genusmerkmal und machen dann in der NP-Regel klar, dass nur Elemente mit demselben Genusmerkmal eine wohlgeformte Konstruktion eingehen können.

Exemplarisch könnte das so aussehen (G = Genus; M,F,N = Maskulin, Feminin, Neutrum):

Det<sub>[G]</sub> → {*der<sub>[M]</sub>, die<sub>[F]</sub>, das<sub>[N]</sub>*}

N<sub>[G]</sub> → {*Mann<sub>[M]</sub>, Frau<sub>[F]</sub>, Kind<sub>[N]</sub>*}

NP → Det<sub>[Gx]</sub> N<sub>[Gx]</sub>

Den Kategorien Det und N wird ein Attribut [G] für Genus zugeordnet, das in Abhängigkeit davon, wodurch die Kategorie jeweils instanziiert ist, drei unterschiedliche Werte annehmen kann.

In der NP-Regel wird auf das Attribut Bezug genommen und durch die Variable X ausgedrückt, dass zwar nicht festgelegt werden soll, ob Genus nun M oder F oder N ist, dass aber in beiden Fällen derselbe Wert vertreten sein muss. Auf diese Art ist sichergestellt, dass Det und N miteinander kongruieren.

**Merkmale im Lexikon**

Unter 'Lexikon' haben wir bisher nicht mehr verstanden als die bloße Zuordnung von Wortformen zu Lexemklassen – spezifische und auch für die Syntax relevante Eigenschaften wurden dabei nicht explizit ausgedrückt, weder was die morphologische Form betrifft noch was die Valenzeigenschaften einzelner Lexeme angeht. In den nachfolgenden Abschnitten soll es darum gehen, das Lexikon genau um diese Art von Information anzureichern. Das wesentliche Ziel dieser Herangehensweise besteht darin, in kleinen Schritten von einer derivationalen Grammatik, die einen umfangreichen Regelkomplex und ein informationsarmes Lexikon aufweist, zu einer Grammatik zu gelangen, in der das Lexikon sehr viel mehr Information enthält und die syntaktischen Regeln oder Operationen in entsprechend abgespeckter Form daherkommen können.

Als Notation verwenden wir Merkmalsbündel und beginnen zunächst mit Merkmalen für die Kategorie und den Kongruenzmerkmalen. 'Kongruenz' heißt im englischen Sprachraum **'Agreement'** und diesen Ausdruck wollen wir hier auch verwenden, unsere Attribute nennen wir in Kurzform **Kat** und **Agr**.

Ein Grundtemplate für einen Lexikoneintrag sieht danach so aus

$$\text{Wortform} \begin{bmatrix} \text{Kat} & x \\ \text{Agr} & y \end{bmatrix}$$

Der Wertebereich des Attributes Kat soll zunächst wieder auf den traditionellen Lexemklassen basieren: Kat {N, Vi, Vt, Vt<sub>2</sub>, A, Adv, Det, P...} (um die Subkategorisierung der Verben kommen wir so lange nicht herum, wie wir noch keine Valenzmerkmale im Lexikon integriert haben).

Das Attribut Agr hat als Wert selber wieder eine Merkmalstruktur. Da wir uns am Englischen orientieren, umfasst diese nur die Attribut-Wert-Paare für Numerus (kurz Num) und Person (kurz Pers), die ihrerseits die folgenden Wertebereiche haben: Num {sg, pl} und Pers {1,2,3}.

$$\text{Wortform} \begin{bmatrix} \text{Kat} & x \\ \text{Agr} & \begin{bmatrix} \text{Num} & a \\ \text{Pers} & b \end{bmatrix} \end{bmatrix}$$

Ein wichtiger Aspekt dabei ist, dass nicht alle Kategorien die gleiche Merkmalstruktur aufweisen müssen. Das Agr-Merkmal ist nicht für alle Wortformen gleich relevant – Präpositionen oder Adverbien beispielsweise werden nicht bezüglich Numerus oder Person flektiert. In diesen Fällen taucht das Merkmal schlichtweg nicht auf:

$$\text{in} \begin{bmatrix} \text{Kat} & \text{P} \end{bmatrix} \text{ oder } \text{very} \begin{bmatrix} \text{Kat} & \text{Adv} \end{bmatrix}$$

Bei bestimmten Kategorien können aber auch Merkmale hinzukommen, im Falle der Determinatoren<sup>4</sup> z.B. für Definitheit: Def {+, -}, bei den Verben für Tempus: Tns {pres, past}, bei den Pronomina für Kasus: Kas {subj, obj}

$$a \begin{bmatrix} \text{Kat} & \text{Det} \\ \text{Agr} & \begin{bmatrix} \text{Num} & \text{sg} \\ \text{Pers} & 3 \end{bmatrix} \\ \text{Def} & - \end{bmatrix} \quad \text{kicks} \begin{bmatrix} \text{Kat} & \text{V} \\ \text{Agr} & \begin{bmatrix} \text{Num} & \text{sg} \\ \text{Pers} & 3 \end{bmatrix} \\ \text{Tns} & \text{Pres} \end{bmatrix}$$

In manchen Fällen ist der Wert eines Attributes nicht festgelegt, beim Artikel *the* z.B. der Wert für Numerus, der sowohl sg als auch pl sein kann. In diesem Fall führen wir das fragliche Attribut nicht in der Merkmalsmatrix auf:

$$\text{the} \begin{bmatrix} \text{Kat} & \text{Det} \\ \text{Agr} & [\text{Pers} 3] \\ \text{Def} & + \end{bmatrix}$$

Zur Illustration betrachten wir ein paar konkrete Beispiele, bestimmte Feinheiten (wie z.B. der genaue Unterschied zwischen *some* und *a* außer der Numerusfrage) werden nicht berücksichtigt:

$$\text{dog} \begin{bmatrix} \text{Kat} & \text{N} \\ \text{Agr} & \begin{bmatrix} \text{Num} & \text{sg} \\ \text{Pers} & 3 \end{bmatrix} \end{bmatrix} \quad \text{dogs} \begin{bmatrix} \text{Kat} & \text{N} \\ \text{Agr} & \begin{bmatrix} \text{Num} & \text{pl} \\ \text{Pers} & 3 \end{bmatrix} \end{bmatrix}$$

$$\text{bites} \begin{bmatrix} \text{Kat} & \text{V} \\ \text{Agr} & \begin{bmatrix} \text{Num} & \text{sg} \\ \text{Pers} & 3 \end{bmatrix} \\ \text{Tns} & \text{Pres} \end{bmatrix} \quad \text{bit} \begin{bmatrix} \text{Kat} & \text{V} \\ \text{Tns} & \text{Past} \end{bmatrix}$$

$$\text{bite} \begin{bmatrix} \text{Kat} & \text{V} \\ \text{Agr} & [\text{Num} \text{pl}] \\ \text{Tns} & \text{Pres} \end{bmatrix} \quad \text{bite} \begin{bmatrix} \text{Kat} & \text{V} \\ \text{Agr} & \begin{bmatrix} \text{Num} & \text{sg} \\ \text{Pers} & \{1,2\} \end{bmatrix} \\ \text{Tns} & \text{Pres} \end{bmatrix}$$

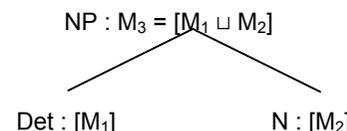
$$I \begin{bmatrix} \text{Kat} & \text{Pron} \\ \text{Agr} & \begin{bmatrix} \text{Num} & \text{sg} \\ \text{Pers} & 1 \end{bmatrix} \\ \text{Kas} & \text{subj} \end{bmatrix} \quad \text{them} \begin{bmatrix} \text{Kat} & \text{Pron} \\ \text{Agr} & \begin{bmatrix} \text{Num} & \text{pl} \\ \text{Pers} & 3 \end{bmatrix} \\ \text{Kas} & \text{obj} \end{bmatrix}$$

$$\text{some} \begin{bmatrix} \text{Kat} & \text{Det} \\ \text{Agr} & [\text{Pers} 3] \\ \text{Def} & - \end{bmatrix} \quad \text{a} \begin{bmatrix} \text{Kat} & \text{Det} \\ \text{Agr} & \begin{bmatrix} \text{Num} & \text{sg} \\ \text{Pers} & 3 \end{bmatrix} \\ \text{Def} & - \end{bmatrix}$$

### Merkmale in der Regelkomponente

Die Modifikation des Lexikons wird es uns ermöglichen, bestimmte der Formen, die mit den Grammatiken A und B übergeneriert wurden, von vorneherein zu verhindern. Dazu müssen wir allerdings auch die Regelkomponente modifizieren – in den Regeln muss natürlich ein Bezug zu den Merkmalsstrukturen bestehen. Diese Modifikation läuft darauf hinaus, dass wir die Kategorien NP, VP, PP usw. – ähnlich wie die Einträge im Lexikon – ebenfalls durch Merkmalsstrukturen charakterisieren. Der entscheidende Punkt dabei ist, dass sich diese Merkmalsstrukturen ergeben aus der Vereinigung, also der Unifikation der Merkmalsstrukturen der abhängigen Elemente, und dass dabei ganz bestimmte Restriktionen (die oa. *feature constraints*) zu berücksichtigen sind.

Die nachfolgende Graphik soll diese Herangehensweise illustrieren. 'M' steht für Merkmalstruktur:



Zu lesen ist das wie folgt: die Merkmalstruktur M<sub>3</sub> der NP stellt die Unifikation der Merkmalstrukturen M<sub>1</sub> (vom Determinator) und M<sub>2</sub> (vom Nomen) dar.

<sup>4</sup> auf bestimmte Abhängigkeiten oder Redundanzen zwischen Merkmalen, wie z.B. die Tatsache, dass das Person-Merkmal für Det immer 3 sein wird, solange wir die Pronomina nicht zu den Determinatoren zählen, sei hier nur hingewiesen.

Entscheidend für die Unifikation von Merkmalsstrukturen ist, dass sie zueinander kompatibel sein müssen. Wir gehen auf den genauen Unifikationsalgorithmus hier nicht ein, sondern beschreiben die Bedingungen nur informell:

- wenn in den zu unifizierenden Merkmalstrukturen ein Merkmal mit demselben Attribut vertreten ist, kann Unifikation nur stattfinden, wenn das Attribut den gleichen Wert hat. Beispiel: ein Merkmal [Num:sg] ist nicht mit einem Merkmal [Num:pl] unifizierbar.
- wenn in der einen der zu unifizierenden Merkmalsstrukturen ein Merkmal auftritt, das in der anderen nicht vertreten ist, so wird dieses Merkmal in die Unifikation übernommen.

Beispiel:

$\left[ \begin{array}{l} \text{Agr} \text{ [Pers 3]} \\ \text{Def} \text{ -} \end{array} \right]$  kann mit  $\left[ \begin{array}{l} \text{Agr} \text{ [Num sg]} \\ \text{Pers} \text{ 3} \end{array} \right]$  zu  $\left[ \begin{array}{l} \text{Agr} \text{ [Num sg]} \\ \text{Pers} \text{ 3} \\ \text{Def} \text{ -} \end{array} \right]$  unifiziert werden.

Wir berücksichtigen zunächst drei unterschiedliche **Typen von Restriktion**, nämlich

- (a) ob bestimmte Konstituenten in der Regel ein konkretes Merkmal aufweisen und welches das ist oder
- (b) ob Konstituenten bezüglich bestimmter, im Lexikon aufgeführter Merkmale kongruieren müssen oder
- (c) ob eine anderweitig nicht spezifizierte Konstituente die Merkmale einer bestimmten Tochter aufweist.

**Die formale Repräsentation**

In der Notation der Regeln und der dazugehörigen Restriktionen orientieren wir uns an dem PATR Formalismus. Im Einzelnen gilt dabei, dass die Restriktionen unterhalb der 'eigentlichen' Regel aufgeführt werden.

Wir führen uns zunächst einmal umgangssprachlich den Sachverhalt vor Augen, den es formal darzustellen gilt, und beginnen exemplarisch mit einer NP-Regel, die eine Restriktion von dem oa. Typ (b) enthält.

Zunächst machen wir eine Angabe über die Elemente, aus denen sich eine mögliche, wohlgeformte NP konstituiert. Die Form ist Ihnen wohlbekannt.

Umgangssprachlich	Formale Repräsentation
Eine mögliche NP konstituiert sich aus einem Det, einer fakultativen AP und einem N in dieser Reihenfolge	NP → Det (AP) N

Als nächstes wollen wir angeben, dass der Determinator und das Nomen in ihren Kongruenzmerkmalen übereinstimmen müssen. Wir müssen in der formalen Darstellung also angeben, um welches Element des Regelrumpfes es geht (Det bzw. N) und dann, welche Bedingung erfüllt sein muss. In der PATR-Notation wird diese Information so dargestellt, dass die 'Bezugsgröße' (Det bzw N) und das entsprechende Merkmal in Spitzklammern notiert werden. Das Gleichheitszeichen zwischen beiden drückt aus, dass die Merkmale kompatibel sein müssen:

Umgangssprachlich	Formale Repräsentation
Der Det und das N müssen in ihren Kongruenzmerkmalen übereinstimmen	<Det Agr> = <N Agr>

Die nächste Restriktion, die wir formulieren wollen, ist vom o.a. Typ (c). Erinnern Sie sich in diesem Zusammenhang bitte an die Aussagen über Dependenz, genauer gesagt die folgende Aussage

Der Kopf determiniert die formalen Eigenschaften der Konstituente, deren Kopf er ist.

Umgangssprachlich	Formale Repräsentation
Die Kongruenzmerkmale der NP sind die des Kopfes N	<NP Agr> = <N Agr>

Die formale Repräsentation dieser Restriktion ist mit der vorherigen fast identisch, und doch gibt es einen großen Unterschied in der Interpretation: im vorherigen Fall werden die Merkmalstrukturen, die für den Det und das N im Lexikon eingetragen sind, auf ihre Unifizierbarkeit überprüft. Hier aber geht es darum, die Kongruenzmerkmale der NP auf den Wert des Kopfes zu setzen. Wir brauchen diese diese Eigenschaften der NP dann, wir Aussagen über die Kongruenz zwischen NP und VP machen wollen.

Somit hätte die Beispiels-NP-Regel die folgende Form:

Umgangssprachlich	Formale Repräsentation	Als PATR-Regel
Eine mögliche NP konstituiert sich aus einem Det, einer fakultativen AP und einem N in dieser Reihenfolge	NP → Det (AP) N	Rule {Nominalphrase} NP -> Det (AP) N: <Det Agr> = <N Agr> <NP Agr> = <N Agr>.
Det und N müssen in ihren Kongruenzmerkmalen übereinstimmen	<Det Agr> = <N Agr>	<b>Wichtig:</b> zwischen der 'eigentlichen' Regel und den Restriktionen muss ein Doppelpunkt stehen.
Die Kongruenzmerkmale der NP sind die des Kopfes N	<NP Agr> = <N Agr>	

Im Unterschied zur mangelnden Kongruenz ist das Problem im Falle der folgenden Beispiele anders gelagert: *\*him kicked the dog* oder *\*the boy kissed she*. Hier wurde nicht berücksichtigt, dass durch Personalpronomina realisierte NP in Abhängigkeit von ihrer Position (also als Subjekts- bzw. Objekts-NP) unterschiedlichen Kasus aufweisen müssen. Dieser Punkt ist ein Fall von Restriktionstyp (a) und sehr leicht erfaßt:

- S → NP VP  
    <NP Agr> = <VP Agr>  
    **<NP Kas> = subj** → Eine NP in dieser Position muss das Kasusmerkmal 'subjective' aufweisen
- VP → Vt NP  
    <VP Agr> = <Vt Agr>  
    **<NP Kas> = obj** → Eine NP in dieser Position muss das Kasusmerkmal 'objective' aufweisen
- NP → Pron  
    **<NP Kas> = <Pron Kas>** → Eine NP, die als Kopf ein Pronomen hat, erhält ihr Kasusmerkmal von genau diesem Kopf

Auf den nächsten Seiten geht es darum, die hier getroffenen Einsichten an konkretem Material zu illustrieren. Wir werden dabei mit einen kleinen Datensatz beginnen, für diesen eine Grammatik schreiben und dieses sukzessive über die Hinzunahme weiterer Daten erweitern.

**Datensatz I**

- 1) This cat loves her.      3) \*Her loves this cat      5) \*These dogs loves her  
 2) These dogs love this cat.      4) \*These dogs love she      6) \*This dogs love her

**Lexikon für Datensatz I:**

Neben der Kategorie ist Folgendes für die in Datensatz I gegebenen Syntagmen relevant:

- a) das Kongruenzmerkmal 'Numerus' bei allen Wörtern  
 b) das Kasusmerkmal bei Pronomina.

<i>this</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{Det} \\ [\text{Num} \text{ sg}] \end{array} \right]$	<i>dogs</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{N} \\ [\text{Num} \text{ pl}] \end{array} \right]$	<i>she</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \\ \text{Kas} \end{array} \begin{array}{l} \text{Pron} \\ [\text{Num} \text{ sg}] \\ \text{subj} \end{array} \right]$
<i>these</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{Det} \\ [\text{Num} \text{ pl}] \end{array} \right]$	<i>cat</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{N} \\ [\text{Num} \text{ sg}] \end{array} \right]$	<i>her</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \\ \text{Kas} \end{array} \begin{array}{l} \text{Pron} \\ [\text{Num} \text{ sg}] \\ \text{obj} \end{array} \right]$
<i>love</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{V} \\ [\text{Num} \text{ pl}] \end{array} \right]$	<i>loves</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{V} \\ [\text{Num} \text{ sg}] \end{array} \right]$	

**Regelkomponente für Datensatz I:**

	Umgangssprachlich	Formale Notation	Kommentar
1.	Ein möglicher Satz konstituiert sich aus einer NP und einer VP in dieser Folge NP und VP stimmen in ihren Kongruenzmerkmalen überein Die NP weist das Kasusmerkmal 'subjective' auf	$S \rightarrow NP VP$  <NP Agr> = <VP Agr>  <NP Kas> = subj	verhindert Syntagma 5 verhindert Syntagma 3
2.	Eine mögliche NP konstituiert sich aus einem Det und einem N in dieser Folge Det und N stimmen in ihren Kongruenzmerkmalen überein Die Kongruenzmerkmale der NP sind die des Kopfes N	$NP \rightarrow Det N$  <Det Agr> = <N Agr>  <NP Agr> = <N Agr>	verhindert Syntagma 6
3.	Eine mögliche NP konstituiert sich aus einem Pronomen Die Kongruenzmerkmale der Gesamt-NP sind die des Pron Das Kasusmerkmal der Gesamt- NP ist das des Pron	$NP \rightarrow Pron$  <NP Agr> = <Pron Agr>  <NP Kas> = <Pron Kas>	
4.	Eine mögliche VP konstituiert sich aus einem V und einer NP in dieser Folge Die Kongruenzmerkmale der Gesamt-VP sind die des V Die NP muss das Kasusmerkmal 'objective' aufweisen	$VP \rightarrow V NP$  <VP Agr> = <V Agr>  <NP Kas> = obj	verhindert Syntagma 4

**Datensatz II**

- 7) I love this cat      8) \*I loves this cat  
 Phänomen/Problem: Der Wert des Kongruenz-Merkmals 'Person' von Subjekts-NP und VP stimmt in Syntagma 8 nicht überein: *loves* ist 3. Person, *I* ist 1. Person.

Lösung: das entsprechende Merkmal muss im Lexikon aufgenommen werden. An der Regelkomponente muss nichts modifiziert werden, da 'Person' zum Merkmalskomplex 'Agr' gehört und die Kongruenzbedingungen in den Restriktionen bereits ausgedrückt sind.

**Lexikon für Datensätze I & II:**

(*this* und *these* wie in Lexikon I)

<i>dogs</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{N} \\ [\text{Num} \text{ pl}] \\ [\text{Pers} \ 3] \end{array} \right]$	<i>she</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \\ \text{Kas} \end{array} \begin{array}{l} \text{Pron} \\ [\text{Num} \text{ sg}] \\ [\text{Pers} \ 3] \\ \text{subj} \end{array} \right]$	<i>I</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \\ \text{Kas} \end{array} \begin{array}{l} \text{Pron} \\ [\text{Num} \text{ sg}] \\ [\text{Pers} \ 1] \\ \text{subj} \end{array} \right]$
<i>cat</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{N} \\ [\text{Num} \text{ sg}] \\ [\text{Pers} \ 3] \end{array} \right]$	<i>her</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \\ \text{Kas} \end{array} \begin{array}{l} \text{Pron} \\ [\text{Num} \text{ sg}] \\ [\text{Pers} \ 3] \\ \text{subj} \end{array} \right]$	<i>love</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{V} \\ [\text{Num} \text{ sg}] \\ [\text{Pers} \ 1] \end{array} \right]$
<i>loves</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{V} \\ [\text{Num} \text{ sg}] \\ [\text{Pers} \ 3] \end{array} \right]$	<i>love</i> $\left[ \begin{array}{l} \text{Kat} \\ \text{Agr} \end{array} \begin{array}{l} \text{V} \\ [\text{Num} \text{ pl}] \\ [\text{Pers} \ 3] \end{array} \right]$	

**Regelkomponente für Datensätze I & II:**

Identisch mit Regelkomponente für Datensatz I

**Datensatz III**

- 9) These very ugly dogs love this cat.      10) These dogs love this beautiful cat.  
 Phänomen/Problem: innerhalb der Nominalphrase tauchen die Syntagmen 'very ugly' und 'beautiful' auf.

Lösung:

- (a) das Lexikon muss um die entsprechenden Elemente erweitert werden  
 (b) die PS-Regel für nicht-pronominale NP muss modifiziert werden.

Bezüglich Punkt (b) können zwei Möglichkeiten diskutiert werden, nämlich :

- 1) NP → Det Adv A N  
 NP → Det A N

oder

- 2) NP → Det (AP) N  
 AP → (Adv) A

Beide Regelblöcke erfassen die oa. Daten. Wie kann begründet werden, dass Möglichkeit 2) sinnvoller ist? Um diesen Punkt zu klären, nehmen wir für ein informelles Beispiel ein paar deutsche Daten:

- a) Ein froher Mann
- b) Ein sehr froher Mann
- c) Ein über das Wetter froher Mann
- d) Ein über das Wetter sehr froher Mann
- e) Er ist froh
- f) Er ist sehr froh
- g) Er ist über das Wetter froh
- h) Er ist über das Wetter sehr froh

Die unterstrichenen Syntagmen gehören jeweils zur Konstituentenklasse Adjektivphrase (AP): einmal ist diese Bestandteil einer NP (a-d), einmal Bestandteil einer VP (e-h). Wenn wir auf den Klassentyp AP verzichten – so wie es die Möglichkeit 1) weiter oben tut – bliebe uns nichts anderes übrig, als für jede der Formen a)-h) eine spezifische Regel auszuformulieren:

- a) NP → Det A N
- b) NP → Det Adv A N
- c) NP → Det PP A N
- d) NP → Det PP Adv A N
- e) VP → Vc A
- f) VP → Vc Adv A
- g) VP → Vc PP A
- h) VP → Vc PP Adv A

Man sieht, was passiert, wenn die Zugehörigkeit zur Klasse AP von jeweils a) und e) (*froher / froh*), b) und f) (*sehr froher / froh*), c) und g) (*über das Wetter froher / froh*) und d) und h) (*über das Wetter sehr froher / froh*) nicht erfasst wird.

Mit dem Konstrukt 'AP' und durch entsprechende Angaben darüber, wie die es realisiert sein kann (nur mit einem Adjektiv, mit einem Adjektiv und einem Adverb, mit einer PP und einem Adjektiv oder mit allen dreien) kann man nicht nur diese Zugehörigkeit erfassen, sondern dann entsprechend auch die Regelflut für NP und VP eindämmen:

NP → Det (AP) N  
 VP → V AP

Der entscheidende Punkt ist, dass es sich bei den Unterschieden in den Sätzen a)-h) um Unterschiede der Realisierung der Klasse AP handelt – nicht der Klassen NP und VP. Rufen Sie sich diesbezüglich die Ableitung von PS-Strukturen aus Dependenzstrukturen ins Gedächtnis – auch hier wird deutlich, dass die Dependenzbeziehung nur zwischen jeweils Nomen bzw. Verb und Adjektiv besteht, und die anderen Elemente, sprich Adverbien und PP, dann in einer Dependenzrelation zu eben diesem Adjektiv stehen.

**Lexikon für Datensätze I – III:**

(*this, these, dogs, cats, I, she, her, love* und *loves* wie in Lexikon für Datensätze I & II)  
*ugly* [Kat A]      *beautiful* [Kat A]      *very* [Kat Adv]

**Regelkomponente für Datensätze I – III:**

(Regeln 1, 3 und 4 wie in Regelkomponente für Datensätze I & II)

	Umgangssprachlich	Formale Notation	Kommentar
2.	Eine mögliche NP konstituiert sich aus Det, einer fakultativen AP und N in dieser Folge Det und N stimmen in ihren Kongruenzmerkmalen überein Die Kongruenzmerkmale der Gesamt-NP sind die des Kopfes N	NP → Det (AP) N  <Det Agr> = <N Agr>  <NP Agr> = <N Agr>	} beschreibt Syntagmen 9 und 10
5.	Eine mögliche AP konstituiert sich aus einem fakultativen Adv und A in dieser Folge	AP → (Adv) A	

**Datensatz IV**

11) Dogs love this cat.      12) \*Dog loves this cat.

Phänomen/Problem: Eine NP kann durch ein N realisiert sein. Dies haut aber nur dann hin, wenn dieses N im Plural steht.

Lösung: es muss eine weitere Regel für die NP formuliert werden. Dabei muss die Restriktion bezüglich des Numerusmerkmals ausgedrückt sein.

**Lexikon für Datensätze I – IV:**

identisch mit Lexikon für Datensätze I – III

**Regelkomponente für Datensätze I – IV:**

(Regeln 1–5 wie in Regelkomponente für Datensätze I–III)

	Umgangssprachlich	Formale Notation	Kommentar
6.	Eine mögliche NP konstituiert sich aus N Die Kongruenzmerkmale der Gesamt-NP sind die des Kopfes N  Das Nomen muss im Plural stehen <sup>5</sup>	NP → N  <NP Agr> = <N Agr>  <N Num> = pl	verhindert Syntagma 12

<sup>5</sup> Auf NP, die ein nicht-zählbares N als Kopf haben wie in *Sincerity frightened the boy* oder *Trouble is brewing* gehen wir hier nicht ein.