

**VERBVALENZUNTERSUCHUNGEN:**

**Eine Voraussetzung für die Monosemierung von Verbinhalten<sup>1</sup>**

**Inhalt:**

0. Vorbemerkungen
1. Monosemierung im Kommunikationsprozeß
2. Die Valenzuntersuchung
3. Die Monosemierung der Verbinhalte
4. Zusammenfassung

**0. Vorbemerkungen**

In den gebräuchlichen Wörterbüchern und Lexika sind die verschiedenen "Bedeutungen" der Lexeme<sup>2</sup> durch Paraphrasen mit anderen, meist auch nicht "eindeutigen" Wörtern gekennzeichnet. Diese Eingrenzung und Differenzierung von Bedeutungen durch Paraphrase ist für die maschinelle Monosemierung von Wortinhalten kein gangbarer Weg.<sup>3</sup> Soll nämlich die maschinelle Analyse eines Textes so differenziert sein, daß sie auch als Grundlage für die Synthese des Textinhaltes in einer anderen Sprache dienen könnte, dann muß sie die durch verschiedene Kontexte realisierten Inhalte von Wörtern von den möglichen realisierbaren unterscheiden. Das bedeutet, daß der Prozeß der Monosemierung, der von den Sprachteilhabern teils bewußt, teils unbewußt vollzogen wird, durchsichtig und in maschinenoperabler Weise nachvollziehbar gemacht werden muß.

**1. Monosemierung im Kommunikationsprozeß**

Im Kommunikationsprozeß wählt der Sprecher für seine Aussage Wörter, die seinen Intentionen der Vermittlung eines Sachverhalts oder Geschehens in ihrer Zusammenstellung nach grammatischen und syntaktischen Regeln entsprechen. Diese Wörter stehen dadurch in einem Kontext, der normalerweise so beschaffen ist, daß bei mehrdeutigen Wörtern nur der eine Inhalt aktiviert wird, der der Vermittlung dessen, was der Sprecher

sagen will, entspricht. Der Hörer vollzieht einen ähnlichen Prozeß, in dem er alle nicht relevanten Inhalte der ihm mitgeteilten Wortformen ausschaltet und die Mitteilung versteht.

Der für den Kommunikationsprozeß zwecks Informationsvermittlung oft unbewußte Vorgang des Ausschaltens nicht-relevanter, in der betreffenden Aussage durch den Kontext nicht aktivierter Wortinhalte wird beispielsweise in einer ausweichenden, absichtlich mehrdeutigen oder ironischen Aussage vom Sprecher oft bewußt eingesetzt. Dazu wählt er den Kontext so, daß mindestens bei einem Redeteil zwei oder mehr Inhalte potentiell aktiviert werden können.

Der Hörer wird sich entweder der Möglichkeit bewußt, bei einem oder mehreren Redeteilen der Aussage mehrere Inhalte aktivieren zu können und versteht die Aussage damit als mehrdeutig – oder er vollzieht diesen Prozeß nicht nach, was häufig in seinem Wunsch nach einer bestimmten Aussage des Sprechers begründet sein mag, und faßt damit die Aussage als eindeutig auf.

Der umgekehrte Fall, daß eine vom Sprecher eindeutig gemeinte Aussage vom Hörer als mehrdeutig verstanden wird, ist ebenso häufig wie der Fall, in dem eine mehrdeutige, vom Sprecher jedoch für eindeutig gehaltene Aussage vom Hörer in einer anderen als der vom Sprecher gemeinten Weise – jedoch ebenfalls eindeutig – verstanden wird. Es handelt sich dann um die klassische Form des Mißverständnisses, bei der sowohl Sprecher als auch Hörer verschiedene Inhalte eines Wortes oder eines Aussageteils aktivieren.

Dieses Problem, das eigentlich der Kommunikationsforschung angehört, stellt sich in ähnlicher Weise auch für die Monosemierung von Wortinhalten durch Kontexte bei der maschinellen Sprachbearbeitung. Dies soll im folgenden für die Verben kurz dargestellt werden.<sup>4</sup>

## **2. Die Valenzuntersuchung**

Voraussetzung für die maschinelle Analyse von Texten ist die genaue Untersuchung der einzelnen Lexeme auf ihre morphologischen, syntaktischen und semantisch-syntaktischen<sup>5</sup> Gegebenheiten hin. Die Untersuchung der Verben auf der semantisch-syntaktischen Ebene wird dabei vor allem Valenzuntersuchung sein.

2. 1 Ausgehend von den Valenzuntersuchungen Helbigs, dargestellt im “Wörterbuch zur Valenz und Distribution deutscher Verben” und den Valenzuntersuchungen, die am Institut für deutsche Sprache, Mannheim, an den Verben des ‘Grunddeutsch’ durchgeführt wurden, kam LIMAS, Bonn, zu einem anderen Ansatz, der sich in einem wesentlichen Punkt von den oben genannten unterscheidet.

Während in den Untersuchungen bei Helbig<sup>6</sup> und am IDS von der “grammatischen Valenz” der Verben (ob ein Verb z.B. an der ersten Valenzleerstelle einen Nominativ, an der zweiten einen Akkusativ oder einen Gliedsatz fordert) ausgegangen wird, geht LIMAS davon aus, daß das Verb keine “grammatischen Wertigkeiten” aufweist, sondern daß die Valenzforderungen der Verben auf der *s e m a n t i s c h - s y n t a k t i s c h e n* Ebene liegen, d.h., das Verb fordert an einer bestimmten Valenzleerstelle nicht einen bestimmten Kasus, einen Gliedsatz o.ä., es fordert einen bestimmten “Inhaltfaktor”<sup>7</sup>, der dann durch verschiedene grammatisch-syntaktische Gegebenheiten, wie etwa Kasus oder Gliedsatz realisiert werden kann. Die Realisation eines solchen semantisch-syntaktischen Inhaltfaktors, die durch das grammatische Feld an bestimmte grammatische und syntaktische Ausdrucksformen gebunden ist, ist jedoch erst an zweiter Stelle von Bedeutung.

Ein weiterer Unterschied zwischen den Valenzuntersuchungen von Helbig und dem IDS einerseits und denen der Forschungsgruppe LIMAS andererseits besteht darin, daß sowohl Helbig als das IDS zwischen obligatorischen und fakultativen Valenzgliedern unterscheiden, während LIMAS diese Unterscheidung nicht kennt. Das, was in den anderen Valenzuntersuchungen als fakultative Valenz angegeben wird, deckt sich nur zum Teil mit den ellipsierbaren valenten Inhaltfaktoren bei LIMAS<sup>8</sup>.

Sowohl bei Helbig als auch beim IDS wird häufig 2- und fakultativ 3-Wertigkeit angegeben. Dabei wird jedoch keine Aussage darüber gemacht, an welcher Valenzleerstelle dieses fakultative Valenzglied erscheint. Bei LIMAS erscheinen alle ellipsierbaren Valenzglieder als Inhaltfaktoren an der ihnen zukommenden Stelle in der Inhaltfaktorenformel. Ist z.B. das zweite Valenzglied der Inhaltfaktorenformel *GETR + ZWGR + GEZL* (Geschehensträger, Zuwendgröße, Geschehensziel) ellipsierbar bzw. fakultativ, so bleibt, auch wenn die Zuwendgröße nicht gesetzt wird, das Geschehensziel der semantisch-syntaktische

Substituent der d r i t t e n Valenzleerstelle.

2.2 Die Valenzuntersuchung bei LIMAS ist auf die Verwendbarkeit ihrer Ergebnisse bei der maschinellen Analyse und Synthese von Texten ausgerichtet. Schon für die maschinelle inhaltliche Analyse - und erst recht für die Synthese - eines Textes ist es notwendig, Kriterien zu entwickeln, mit denen es gelingt, die verschiedenen Wortinhalte voneinander abzugrenzen.

Diese Kriterien müssen unter dem Anspruch der Verwendung von EDV-Anlagen maschinell erkennbar sein.

Vorbedingung hierzu ist die genaue und gründliche Valenzuntersuchung eines jeden Verbs durch einen menschlichen Bearbeiter.

Das wiederum setzt als erstes die Niederlegung und Begründung von Kriterien, an denen sich der Bearbeiter orientieren kann, voraus, die im günstigsten Fall mit den maschinell erkennbaren Kriterien der zu schaffenden Lexikoneinträge übereinstimmen werden. Sie müssen derart beschaffen sein, daß eine beliebige Zahl verschiedener Bearbeiter eines Verbs mit großer Wahrscheinlichkeit (im Idealfall notwendigerweise) zu den gleichen Ergebnissen kommt.

Die Valenzuntersuchung durch den Bearbeiter vollzieht sich bei LIMAS in folgenden Schritten:

2.2.1 Es werden möglichst viele verschiedenartige Verwendungsweisen eines Verbs in Sätzen gesammelt.<sup>9</sup>

Dieses Beispielmaterial wird auf die für die Valenzuntersuchung definierte "Normalform" gebracht, d.h., der Satz soll, soweit möglich, im Aktiv erscheinen, das "Subjekt" soll im Singular stehen, die Zeitform soll Präsens sein und keines der Satzglieder durch seine Stellung besonders betont sein. Ist diese Normalform bei einer der Verwendungsweisen des Verbs nicht möglich, so etwa bei einem Wegfall der Präpositionalphrase bei einem Subjekt im Plural<sup>10</sup>, dann wird diese Unmöglichkeit als ein wesentlicher Gesichtspunkt notiert.

2.2.2 Aus den Sätzen des Beispielmaterials werden anschließend die freien Satzglieder ausgeschieden. Hierbei haben sich folgende Proben zur Feststellung der valenten Satzglieder bewährt:

2.2.2.1 Die *nicht*-Probe: Aussagesätze ohne Präpositionalphrasen werden verneint. Dabei steht das *nicht* im Normalfall nach dem letzten valenten Satzglied, vorausgesetzt, man verneint den ganzen Satz und nicht nur ein Satzglied. Z.B.:

1 *Peter schenkte seiner Freundin dieses Buch.*

Als mögliche Stellungen für *nicht* sind erlaubt:

2 *Nicht Peter schenkte seiner Freundin dieses Buch.*

3 *Peter schenkte nicht seiner Freundin dieses Buch.*

4 *Peter schenkte seiner Freundin nicht dieses Buch.*

Bei diesen Stellungen des *nicht* wird jedoch immer nur das dem *nicht* folgende Satzglied verneint und nicht zusammen mit der Verneinung des Verbs *schchenken* der ganze Satz.

Anders verhält es sich bei Endstellung des *nicht*.

5 *Peter schenkte seiner Freundin dieses Buch nicht.*

In diesem Fall wird zusammen mit dem Verb der ganze Satz verneint.

Damit sind alle Satzglieder, *Peter*, *seiner Freundin* und *dieses Buch* als valent erkannt.

Wiese der Ausgangssatz noch ein weiteres Glied, etwa die Präpositionalphrase *im Garten* auf, so würde die *nicht*-Probe zur Feststellung der Tatsache, ob es sich hierbei um ein valentes oder freies Satzglied handelt, versagen; denn Präpositionalglieder, soweit es sich um *selbständige* Satzglieder handelt, können durch diese Probe nicht in valente und freie Satzglieder geschieden werden.

Anders verhält es sich mit Präpositionalphrasen, die einem anderen Satzglied attribuiert sind. Hieße die Präpositionalphrase etwa *in rotem Leinen*, so würde die Zulässigkeit des verneinten Satzes

6 *Peter schenkte seiner Freundin dieses Buch in rotem Leinen nicht.*

erkennbar machen, daß die Präpositionalphrase in einer Merkmal-Relation zu *Buch* steht und nicht gesondert zu behandeln ist. Dies ist ein Beispiel dafür, daß Expandierungen von Satzgliedern, hier *dieses Buch* zu *dieses Buch in rotem Leinen* oder auch zu *dieses interessante Buch* für die Valenzuntersuchung keine Rolle spielen und deshalb solche Attribuierungen in diesem Bearbeitungsschritt ausgeschieden werden können.

2.2.2.2 Die Nominalphrasen - Probe: Aussagesätze mit Präpositionalphrasen werden in eine Nominal-Partizipial-Konstruktion transformiert, z.B.:

- 7 *Bonn liegt am Rhein.*
- 8 *das am Rhein gelegene Bonn*

Die Präpositionalphrase wird ausgelassen:

- 9 *\*das gelegene Bonn*

Diese Phrase ist abweichend. Damit ist die Präpositionalphrase als valent erkannt.

Die Probe führt bei etwa einem Drittel aller Fälle von Sätzen mit Präpositionalphrasen zu eindeutigen Ergebnissen.

2.2.2.3 Das Verb - Präpositions - Bindungslexikon. In diesem bei LIMAS von Burkhard Schaefer und Klaus Günther Schweisenthal erstellten Lexikon sind die deutschen Verben aufgeführt, die eine enge Bindung an eine oder mehrere Präpositionen aufweisen.

Für die Bindungsgrade ist eine Gradzahl von eins bis fünf eingeführt worden, die größere Zahl gibt dabei die geringere Bindung an.<sup>11</sup>

Bei Bindungsgrad eins und zwei handelt es sich immer um ein valentes Satzglied, bei Bindungsgrad drei und teilweise auch noch bei vier liegt dieser Verdacht nahe, ist aber noch durch die Nominalphrasen-Probe oder die Transformationsprobe zu überprüfen.

2.2.2.4 Die Transformationsprobe durch Beziehungsadverb: Hierbei wird die Präpositionalphrase in ein Adverb der Form *da+r+* Präposition transformiert, auf das dann ein *daß*-Satz oder ein Infinitiv + *zu* folgen muß, wobei immer zumindest ein Verb zu ergänzen sein wird. Z.B.:

- 10 *Er wartet auf dem Bahnhof*
- 11 *\*Er wartet darauf, daß . . .*
- 12 *\*Er wartet, (jdn/etw) zu . . .*

Satz 11 und 12 sind nicht allgemein abweichend, sondern nur als Paraphrase von 10.

- 13 *Er wartet auf den Freund / auf das Kommen des Freundes.*  
 14 *Er wartet darauf, daß der Freund kommt.*  
 15 *Er wartet darauf, den Freund zu sehen.*  
 16 *\*Er wartet, daß der Freund kommt.*  
 17 *\*Er wartet, den Freund zu sehen.*

Satz 17 ist hier insofern abweichend, als er keine Paraphrase von 13 darstellt, sondern das *um* der infinitivischen Zweckangabe ellipsiert ist.

Diese Probe hat als Ergebnis, daß *auf dem Bahnhof* als freies Satzglied und *auf den Freund* bzw. *auf das Kommen des Freundes* als valentes Satzglied erkannt werden. Die Differenzierung der beiden Sätze bezüglich ihrer Präpositionalphrasen wird in einem weiteren Schritt vorgenommen, wenn darüber entschieden wird, welche Nomeninhaltfunktionsklassen in der valenten Präpositionalphrase stehen müssen.

2.2.3 Wenn mit Hilfe dieser Proben die nicht-valenten Satzglieder aus den Sätzen des Beispielmaterials ausgeschieden worden sind, wird nochmals die Vollständigkeit des Beispielmaterials für das zu untersuchende Verb geprüft und notfalls Ergänzungen vorgenommen.

2.2.4 Im Anschluß daran beginnt die eigentliche Valenzcodierung, die in mindestens drei Schritten abläuft.

2.2.4.1 Die als valent erkannten Komponenten der Beispielsätze werden im Hinblick auf ihre Inhaltfaktoren analysiert. Dabei werden unterschiedliche grammatische Gegebenheiten (z.B. Nomen oder Pronomen im Akkusativ, *daß*-Satz, *was*-Satz, Quantifikatoren (*viel*, *nichts*)), die die gleiche semantisch-syntaktische Funktion im Satzgefüge erfüllen, zusammengeordnet und unter einem Oberbegriff, dem "Inhaltfaktor"<sup>12</sup>, zusammengefaßt. (Für die o.a. Konstellation von grammatischen Gegebenheiten würde das bedeuten: GEZL = Geschehensziel, als Substituenten der zweiten Valenzleerstelle bei einer Vielzahl von Verben.)

Durch diesen Arbeitsschritt entstehen für jedes Verb eine bzw. mehrere "Inhaltfaktorenformeln", die im Deutschen aus maximal vier Gliedern bestehen. In dieser Formel, der Inhaltfaktorenformel, taucht das Verb deshalb nicht auf, weil mit ihr die Valenzleerstellensubstituenten benannt werden, die vom Verb gefordert werden.

2.2.4.2 Der nächste Arbeitsschritt legt das grammatische Feld eines jeden Inhaltfaktors für das untersuchte Verb fest, d.h., die verschiedenen Möglichkeiten, die für die grammatische Realisation eines jeden Inhalt-

faktors für dieses Verb bestehen, werden angegeben.

Das "grammatische Feld" eines jeden Inhaltfaktors ist die Menge aller möglichen grammatischen Realisationen eines Inhaltfaktors ohne Berücksichtigung der Leerstellenzahl und ohne Berücksichtigung eines bestimmten Verbs. Dieses grammatische Feld wurde nach längeren Einzelarbeiten für jeden der 15 Inhaltfaktoren festgelegt, aber so flexibel gehalten, daß sich möglicherweise noch ergebende Erweiterungen keine Schwierigkeiten in ihrer Einbeziehung bieten.

Bei der Zusammenstellung der grammatischen Felder der Inhaltfaktoren bei dem einzelnen zu untersuchenden Verb wird darauf geachtet, daß jede grammatische Ausformung eines Inhaltfaktors mit jeder Realisierung der anderen Inhaltfaktoren kombinierbar ist. Schließt etwa ein *daß*-Satz an der ersten Valenzleerstelle einen *daß*-Satz an der zweiten Valenzleerstelle aus, obwohl alle anderen Elemente des grammatischen Feldes des ersten Inhaltfaktors mit dem Element *daß*-Satz des grammatischen Feldes des zweiten Inhaltfaktors kompatibel sind, so wird der *daß*-Satz aus dem grammatischen Feld des ersten Inhaltfaktors entfernt und mit dem um den *daß*-Satz verminderten grammatischen Feld des zweiten Inhaltfaktors gesondert aufgeführt, jedoch mit derselben Inhaltfaktorenformel.

2.2.4.3 In diesem dritten Bearbeitungsschritt sollen für die einzelnen Elemente des grammatischen Feldes eines jeden Inhaltfaktors Klassen angegeben werden. Hierbei handelt es sich für mit Nomen zu besetzende Stellen um Nomeninhaltfunktionsklassen, für die anderen Wortarten um entsprechende Inhaltfunktionsklassen sowie bei den Präpositionalphrasen um eine Klassifizierung des Phraseninhalts.

Da bisher bei LIMAS eine Klassifizierung der Nomen nur nach Ort, Zeit und Abstrakt im Zusammenhang mit der Analyse der deutschen Präpositionen<sup>13</sup> vorgenommen wurde, diese Klassifizierung jedoch für die Verbvalenzuntersuchung nicht ausreicht, werden in diesem Bearbeitungsschritt vorerst neben der o.a. Klassifizierung vom Bearbeiter weitere Angaben gemacht, die jedoch erst später, nach der Bearbeitung einer Vielzahl von Verben, systematisiert werden sollen. Denn erst dann wird überschaubar sein, welche Anforderungen an eine Nomenklassifikation von der Verbvalenzuntersuchung her gestellt werden müssen.

Ähnliches gilt für die Adjektive, deren Inhaltfunktionklassen jetzt erarbeitet werden.<sup>14</sup>

Bei abhängigen Sätzen als Valenzgliedern ist noch keine Klassifizierung vorgenommen worden; hier werden vom Bearbeiter nur notwendige Beschränkungen in Bezug auf die in ihnen verwendbaren Verben, Nomen usw. angegeben, falls solche Beschränkungen bestehen, ebenso wie in Bezug auf die Zeitrelation, in der der valente Satz zum Bezugsverb zu stehen hat. So verlangen einige Verben z.B. daß der valente Satz in der Zeitstufe der Nachzeitigkeit, Vorzeitigkeit oder auch Gleichzeitigkeit zu der des Verbs im Hauptsatz, für das die Valenzuntersuchung durchgeführt wird, steht, andere wiederum fordern keine solche Beschränkung.

2.2.4.4 Bei der Bearbeitung dieser drei Schritte wird schon darauf geachtet, ob möglicherweise ein Inhaltfaktor mit seinem gesamten grammatischen Feld ellipsierbar ist, d.h., ob die durch ihn gegebenen Informationen derart beschaffen sind, daß sie aus dem Vortext entnommen werden können, wenn man von dem Normalfall ausgeht, daß der zu analysierende Satz in einem Text steht. In diesem Fall wird die Möglichkeit der Ellipsierung angegeben. Weiterhin wird untersucht, ob die Ellipse eines Inhaltfaktors eventuell Einfluß auf die verbleibenden Valenzglieder ausübt. So erlangt die auf einen konkreten Sachverhalt bzw. eine konkrete Situation zugeschnittene Aussage durch die Ellipse eines Valenzgliedes häufig den Charakter der Allgemeingültigkeit.

Ellipsiert man in dem Satz

18 *Hans raucht eine Zigarette.*

mit der Inhaltfaktorenformel Geschehensträger + Geschehensziel (GETR + GEZL) das Geschehensziel *eine Zigarette*, so hat

19 *Hans raucht.*

als Opposition nicht mehr wie z.B. 18

20 *Hans raucht eine Zigarre.*

sondern

21 *Hans raucht nicht.*

Ließe sich aus dem vorangehenden Kontext etwa folgende Situation aufbauen

22 *Peter verteilte an alle, die rauchen wollten, Zigaretten. Auch Hans nahm eine. Er raucht jetzt.*

so bliebe ohne das Hinzufügen von *jetzt* dieser Satz auch in diesem Kon-

text ambigue, da 19 auch dann zu verstehen ist als

23 *Hans ist Raucher.*

Wichtig ist für die Sätze 18 und 20 außerdem die Eintragung, daß ein Artikel oder ein Zahlwort zu stehen hat. Steht nämlich das Geschehensziel (im Plural) ohne Artikel, so gilt als Paraphrase

24 *Hans ist Zigaretten- bzw. Zigarrenraucher.*

sowie für das artikellose Geschehensziel im Singular in

25 *Hans raucht Pfeife*

als Paraphrase gilt

26 *Hans ist Pfeifenraucher.*

Bei den Verben, die als Elemente des grammatischen Felds eines Inhaltfaktors abhängige Sätze aufweisen, wird darauf geachtet, ob der Gesamtsatz nach der Ellipsierung eines Inhaltfaktors dann ungrammatisch wird, wenn die abhängigen Sätze ihren speziellen Charakter behalten und dadurch der auf Allgemeingültigkeit abzielenden Aussage des Hauptsatzes widersprechen.

### 3. Die Monosemierung der Verbinhalte

Außerdem wird bei den oben beschriebenen Arbeitsschritten darauf geachtet, ob z.B. eine bestimmte Inhaltfaktorenformel schon eine Differenzierung verschiedener Verbinhalte bringt, oder ob diese Differenzierung sich erst durch die Festlegung des grammatischen Felds der einzelnen Inhaltfaktoren ergibt oder aber erst nach der Festlegung der Wortinhaltsklassen für die einzelnen Elemente des grammatischen Felds.

Bei dem Verb *bedeuten* ist schon eine Differenzierung auf der Ebene der Inhaltfaktoren gegeben. Die Inhaltfaktorenformel Nennträger + Nennziel (NTR + NZL) mit dem Beispielsatz

27 *eng. book bedeutet Buch.*

aktiviert einen anderen Verbinhalt als die Faktorenformel Geschehens-träger + Zuwendgröße + Geschehensziel (GETR + ZWGR + GEZL) wie in

28 *Er bedeutete ihr, den Mund zu balten.*

und wiederum einen anderen Verbinhalt die Faktorenformel NTR + ZWGR + NZL in

29 *Das bestandene Examen bedeutete ihm viel.*

Dies tritt noch deutlicher zutage, wenn man die unterschiedlichen Ausformungen des grammatischen Felds vergleicht. So ist die Zuwendgröße in 28 nur durch einen Dativ substituierbar, in 29 durch einen Dativ und eine Präpositionalphrase 'für+Nomen (= Lebewesen)'. Für die anderen Inhaltfaktoren gilt ähnliches.

3.1 Eine Schwierigkeit bildet häufig die Entscheidung, ob ein Verb in einer Bedeutung n- oder n-1wertig ist. In allen diesen Fällen handelt es sich bei einem der Valenzglieder um eine Zuwendgröße, die häufig als Possessivpronomen beim nachfolgenden Valenzglied auftritt. In der Valenzuntersuchung nach Inhaltfaktoren ist diese Schwierigkeit auf der semantisch-syntaktischen Ebene lösbar. Für Satz wie

30 *Der Sturm färbte ihm das Gesicht rot.*

und

31 *Der Sturm färbte sein Gesicht rot.*

wird eine vierstellige Inhaltfaktorenformel angegeben: Geschehensträger + Zuwendgröße + Geschehensziel + Merkmal (bezogen auf Geschehensziel). Im grammatischen Feld der Zuwendgröße sind dann zwei Einträge notwendig: einmal "Nomen oder Pronomen im Dativ" und zum anderen "Possessivpronomentransformation", wodurch die Möglichkeit des Zusammenfallens von zwei Inhaltfaktoren, nämlich Zuwendgröße und Geschehensziel in einem Satzglied, dem 'Akkusativobjekt' *sein Gesicht*, gegeben ist.

3.2 Die o.a. Kriterien, die von dem Bearbeiter bei der Valenzuntersuchung angewendet worden sind, schlagen sich in dem semantisch-syntaktischen Lexikoneintrag zu den Verben nieder.

Dieser Lexikoneintrag wird so organisiert, daß zu jedem Inhalt eines Verbs, also normalerweise mehrfach bei einer Grundform, als erstes die Inhaltfaktorenformel angegeben wird, dann zu jedem Inhaltfaktor der Inhaltfaktorenformel das grammatische Feld und als letztes zu jedem Element des grammatischen Felds die klassifizierten semantischen Substitutionsmöglichkeiten.

3.3 Bisher haben sich aus der Bearbeitung der Verben 15 verschiedene Inhaltfaktoren ergeben, die in 67 verschiedenen Inhaltfaktorenformeln auftreten. Die drei kleinsten grammatischen Felder haben jeweils nur einen Substituenten, das größte, das der Gleichsetzung, hat 18 Elemente.

Mit diesen Differenzierungen der Valenzglieder und der weiteren Unterdifferenzierung der Elemente des grammatischen Felds der Inhaltfaktoren ist in den meisten Fällen eine Monosemierung der Verben möglich.

3.3.1 Der Grad der Monosemierung wird für die Synthese größer sein als für die Analyse, da man bei der Synthese häufig solche Realisierungen der Elemente des grammatischen Felds sperren kann, die sich mit Realisierungen des grammatischen Felds eines anderen Inhaltfaktors überschneiden.

3.3.2 Bei der Analyse muß man jedoch von dem einen Satz des Textes ausgehen, der gegeben ist, ohne daß die für die Synthese geltenden Einschränkungen hier von Nutzen sind.

3.4 Für die Monosemierung von Verben bei der Textanalyse bedeutet das: Hat man das Verb des Hauptsatzes und die Verben möglicher abhängiger Sätze durch die syntaktische Analyse gefunden, so folgt aus einem Vergleich mit der Gesamtheit aller möglichen Valenzleerstellen-substituenten in grammatischer Sicht im Lexikon schon auf dieser Ebene eine Unterteilung der Satzglieder in solche, die möglicherweise valent sein können, und solche, die eindeutig als freie Satzglieder erkennbar sind.

Diese freien Satzglieder werden zu einer späteren Analyse zurückgestellt, jedoch so, daß jederzeit auf sie zurückgegriffen werden kann.

Die als möglicherweise valent erkannten Satzbestandteile werden mit den semantischen Anforderungen der Elemente der grammatischen Felder der Inhaltfaktoren verglichen. Hier lösen sich dann weitere Verb-mehrdeutigkeiten auf, dadurch, daß z.B. die einem Inhalt zugeordneten Inhaltfunktionsklassen durch die Anforderungen des Inhaltfaktors für das zu untersuchende Verb ausgeschlossen sind.

Im Anschluß daran werden alle verbleibenden, möglicherweise valenten Satzglieder mit den differenziertesten semantischen Valenzcodemasken verglichen, die für die zu analysierenden Verben entwickelt worden

sind. Dabei werden sich in den meisten Fällen Monosemierungen der Verben ergeben, dergestalt, daß der Verwendungsweise des Verbs in dem gegebenen Satz eindeutig eine Inhaltfaktorenformel zugeordnet werden kann.

3.5 Dieses Vorgehen soll an einigen Beispielsätzen erläutert werden.

- 32 *Die Mutter sieht nach dem Kind / auf das Kind.*
- 33 *Der Chef sieht auf Erledigung der Arbeit / auf Pünktlichkeit.*
- 34 *Der Mann sieht auf das Buch.*
- 35 *Die Fenster sehen auf den Hof / nach der Straße.*

Die Inhaltfaktorenformel für alle vier Sätze lautet: Geschehensträger + Beziehungsglied (GETR + BEZGL); das als GETR realisierte Element des grammatischen Felds ist jeweils ein Nomen im Nominativ, das des Beziehungsglieds eine Präpositionalphrase.

Realisiert ist der Nominativ in Satz 32 - 34 jeweils durch eine Größe, Ort, belebt, Person und in 35 durch eine Größe, Ort, unbelebt, Konkretum.

Für das Beziehungsglied ergibt ein Vergleich in dem Verb-Präpositions-Bindungslexikon, daß die Präpositionalphrase von Satz 32 und 33 in ihrem Phraseninhalte von dort her bestimmt sein kann, wobei die Bindung zwischen *sehen* und *auf* und *sehen* und *nach* den Phraseninhalt mit starker Wahrscheinlichkeit monosemiert. (Eine Mehrdeutigkeit zwischen der zweiten Möglichkeit von Satz 32 mit Satz 34, die auch von dem menschlichen Hörer des Satzes ohne Kontext nicht ausgeschlossen werden kann, bleibt bestehen.) Für eine etwaige Paraphrasierung dieses Verbinhaltes von *sehen* durch andere Verben, wie etwa *achten auf* oder *sich kümmern um* ist die differenzierteste Valenzcode-maske, in der bei der Präpositionalphrase zwischen Person und Abstraktum unterschieden wird, nur von sekundärer Bedeutung.

Bei Satz 34 ist die Beziehung, die durch die Präpositionalphrase angegeben wird, als Ort-Richtungs-Beziehung definierbar, und die Eintragung in dem grammatischen Feld gibt dies auch an. Es können entsprechend dieser Definition auch andere Ort-Richtungs-Präpositionen substituiert werden, etwa *in*, *neben*, *hinter* usw., ohne den durch die Valenzcodierung gesetzten Rahmen zu sprengen. Im obigen Beispielmateriale besteht eine Mehrdeutigkeit zwischen dem *auf*, bestimmt durch Verbbindung, und dem *auf* als Ort-Richtungs-Präposition, die

nur durch eine umfassende Kontextanalyse gelöst werden kann.

Für die Beziehungsgröße von Satz 35 gilt ähnliches wie bei Satz 34, nur daß hier zusammen mit der definierbaren Beziehung Ort-Richtung eine schwache Verbbindung fungiert, die den Austausch von *auf* und *nach* gegen andere Ort-Richtungs-Präpositionen ausschließt.

Für *sehen* wären, wenn nur obige vier Beispielsätze zu bearbeiten wären, folgende Verbinhalte zu unterscheiden:

- I. *sehen auf* oder *nach* bestimmt durch Verbbindung
- II. *sehen* + Ort-Richtung
- III. *sehen* + Ort-Richtung, eingeschränkt durch Bindung  
GETR = KKr<sup>15</sup>

#### 4. Zusammenfassung

Bei den bisherigen Valenzuntersuchungen an etwa 400 Verben hat sich gezeigt, daß der größte Teil der Verben bei einer maschinellen Analyse durch vorhergehende Valenzuntersuchungen in Beziehung auf die Verbinhalte zu monosemieren ist.<sup>16</sup> Davon unberührt bleiben nur die Fälle, in denen nur eine begrenzte Anzahl von Wörtern an einer bestimmten Valenzleerstelle als Substituenten fungieren können (vgl. Satz 35), um einen bestimmten Verbinhalt zu aktivieren.

Es scheint uns jedoch unmöglich, die Inhaltunktionsklassen derart fein zu subklassifizieren, daß z.B. für das Verb *sehen* Nomen wie *Haus*, *Fenster*, *Balkon* u.a. eine Inhaltunktionsklasse der Nomen bilden, da diese Nomen in Sätzen mit *sehen* als Geschehensträger möglich sind, so z.B. in Sätzen wie:

- 36 *Das Fenster / Der Balkon sieht nach der Straße / auf den Hof; das Haus sieht auf den See.*

Für das Verb *geben* müßte diese Nomeninhaltunktionsklasse jedoch um das Element *Haus* vermindert werden, wegen

- 37 *\*Das Haus gibt auf den See.*

obwohl folgende Sätze möglich scheinen:

- 38 *Das Fenster / der Balkon gibt nach der Straße / auf den Hof.*

Für andere Verben müßte dann diese Nomeninhaltunktionsklasse möglicherweise um einige Nomen vermehrt werden.

4.1 Eine andere Möglichkeit, die augenblicklich untersucht wird, geht von der Voraussetzung aus, daß Sätze wie Satz 35 dem metaphorischen Gebrauch eines Verbs zuzurechnen sind.

Der Satz 35 würde dann vollständig analysiert, die Inhaltfaktorenformel GETR + BEZGL würde gefunden, ebenfalls die grammatischen Substituenten als Nominativ und Präpositionalphrase Ort-Richtung erkannt. Als Maschinenkommentar würde im Anschluß an die Analyse jedoch zusätzlich ausgedrückt: metaphorischer Gebrauch, da GETR  $\neq$  LBW (= Lebewesen).

#### A n m e r k u n g e n

- 1 Dieser Aufsatz geht zurück auf einen Vortrag auf dem Kongreß der Gesellschaft für Angewandte Linguistik (GAL) im Oktober 1970 in Stuttgart. Weiterführende Literatur wurde, soweit möglich, berücksichtigt.
- 2 Hierunter werden sowohl Homographen wie *Schloß* = Gebäude und *Schloße* = Türschloß als auch Sememunterschiede (vgl. Stötzel, Ausdruckseite, S. 73 u.ö.) gefaßt.
- 3 Vgl. hierzu Meier-Vollmann, Monosemierungsalgorithmen, S. 95: "Ein relativ kleiner Teil von Arbeiten zeigt jedoch daß die von uns ausgewählte Methode zwar nicht die einzig mögliche ist, daß diese Arbeiten jedoch nur in dem Maße zum Ziel führen, in dem sie die volle Sememdefinition berücksichtigen. Der Hauptunterschied liegt also in der Art und Weise der Erzielung der Definition und dem Grad der Vollständigkeit, wobei sich die sog. Textwortsubstitutionsmethode aber als zu einseitig und ungenügend erwiesen hat, weil sie einerseits letztlich die Existenz gleichwertiger Synonyme voraussetzt, die in zahlreichen Fällen jedoch fehlt oder nur grob existiert und andererseits im allgemeinen wieder mehrdeutige Wörter als Substituente einsetzt und damit nur x durch y ersetzt." Verwiesen wurde im Vortext auf die Arbeiten von Katz-Fodor, S. Abraham und S. Kiefer.
- 4 Im weiteren werden die freien Satzglieder eines Satzes zur Monosemierung des Verbs nicht herangezogen, da auch in den aus der Valenzuntersuchung gewonnenen Lexikoneinträgen nur die valenten Glieder berücksichtigt werden können. Zur Scheidung von valenten und freien Satzgliedern vgl. S. 87 ff.
- 5 Dieser Begriff wird im folgenden so verwendet wie in der Kommunikativen Grammatik Alfred Hoppes.

- 6 So gibt Helbig, Wörterbuch zur Valenz und Distribution deutscher Verben, S. 149 für *erfreuen* an:

"I sich erfreuen<sub>2</sub> (V1 = Freude haben, genießen)

II. sich erfreuen → Sn [=Nominativ], Sg [=Genitiv] / NS daß  
[=Nebensatz] / Inf [=Infinitiv + zu]

III. Sn → keine Selektionsbeschränkungen (*Der Kollege, der Hund, das Institut, die Lampe, die Idee, das Schwimmen* erfreut sich großer Beliebtheit.)

Sg → Abstr (Beliebtheit, Respekt, Hochachtung, Vertrauen, Ruf, Achtung. . .) (Er erfreut sich *großer Beliebtheit*.)

p = an,

pSd → keine Selektionsbeschränkungen (Er erfreut sich *an dem Mädchen, an dem Hund, an dem Theater, an der Lampe, an seinem Ruf, an dem Schwimmen*.)

NS → Act(Er erfreut sich daran, *daß er helfen kann*.)

Inf → Act(Er erfreut sich daran, *den Kindern zu helfen*.)"

Hierbei wird auf Stufe I die Anzahl der Mitspieler als Index angegeben.

Auf Stufe II werden die obligatorischen und fakultativen Mitspieler als syntaktische Umgebungen des Verbs bestimmt. Auf Stufe III werden die obligatorischen und fakultativen Mitspieler des Verbs (vgl. ebd. S. 69 ff.) als semantische Umgebungen fixiert. Für unser Beispiel bedeutet das auf der semantischen Stufe III, das außer den oben gemachten Angaben folgende Beschränkungen für die Variante 1 angegeben werden (S. 150):

"Wenn bei V [ariante] 1 p [räpositionales] S [ubstantiv] [im] D [ativ] keine Selektionsbeschränkungen hat, muß Sn [=Nominativ] → Hum oder Abstr (als Hum) sein; wenn Sn keine Selektionsbeschränkungen hat, muß Sg [=Genitiv] → Abstr. folgen."

Bei dieser linearen Darstellung ist es jedoch schwierig, die sich ergebenden Selektionsbeschränkungen sichtbar zu machen.

So sind nach den o.a. Eintragungen für diese eine Verbbedeutung von *sich erfreuen* folgende Sätze zulässig:

a) \**Die Idee erfreut sich daran, daß es schneit.*

b) \**Das Schwimmen erfreut sich daran, daß es so gesund ist.*

c) \**Die Lampe erfreut sich daran, zu leuchten.*

usw.

(In [ ] stehen die Auflösungen der Abkürzungen.)

- 7 S. dazu u. S. 89 f.

- 8 Vgl. den Vortrag von U. Engel, Dependenzgrammatik in der Schule, gehalten auf der GAL-Tagung 1971 in Stuttgart, abgedr. im IRAL-Sonderband 1972 (i.V.)

- 9 Das Beispielmateriale wurde anfangs aus Lexika und der Kompetenz des Bearbeiters gewonnen. Seit einiger Zeit stehen jedoch auch Beispielsätze aus Texten zur Verfügung, die jedoch auch weiterhin aus der Kompetenz

des Bearbeiters und aus Lexika zu ergänzen sind, da seltene, jedoch mögliche Verwendungsweisen wegen des noch zu geringen Textumfangs nicht belegt sind.

- 10 Als Beispiel: *Die SPD vereinigte sich mit der DKP, - Die linken Parteien vereinigten sich.* (Wohlgemerkt, dies ist nur ein Beispiel für die Verwendung von *sich vereinigen*.)
- 11 Bindungsgrad 1 besagt, daß dieses Verb in dieser Bedeutung notwendig mit der fraglichen Präposition vorkommt, Bindungsgrad 2, daß es, wenn es mit einer Präposition steht, dann mit dieser einen. Bei Bindungsgrad 3 und 4 verhält es sich ähnlich, nur daß statt einer auch mehrere Präpositionen ähnlichen Inhalts vorkommen können.
- 12 Vgl. dazu den Aufsatz von B. Schaefer in diesem Band.
- 13 Vgl. dazu K.G. Schweisthal, Präpositionen in der maschinellen Sprachbearbeitung, - In: Schriftenreihe zur Kommunikativen Grammatik, Bd. 1, Bonn 1971.
- 14 Vgl. dazu den Aufsatz von Booss-Schweisthal im gleichen Band.
- 15 Alle weiteren Differenzierungen, die zu einer solchen Valenzuntersuchung als Grundlage für die Monosemierung von Verbinhalten gehören, sind eben ausgeführt worden.
- 16 Ein Algorithmus zur maschinellen Analyse ist dabei nur für die Valenzuntersuchung als ganzes und nicht, wie bei Meier-Volkman, für jedes Verb notwendig.

#### L i t e r a t u r

- Ballweg, Joachim; Hacker, Hans-Jürgen; Schumacher, Helmut: Satzbaupläne und Semantik. Vorüberlegungen zur semantischen Subkategorisierung deutscher Satzbaupläne mit gegebenem Verb. - In: Muttersprache 81, 1971, H. 4, S. 224-34.
- Helbig, Gerhard: Wörterbuch zur Valenz und Distribution deutscher Verben, Leipzig 1969.
- Helbig, Gerhard: Zum Problem der Stellung des Negationswortes "nicht". - In: Deutsch als Fremdsprache, Jg. 8, 1971, H. 2, S. 66-76.
- Hoppe, Alfred: Die semantische Syntax der Kommunikativen Grammatik auf EDV-Anlagen, Preprint Nr. 67, 1969, International Congress on Computational Linguistics, Stockholm/Sweden, Stockholm 1969.
- Meier, Barbara; Volkman, Eva: Monosemierungsalgorithmen polysemer deutscher Lexeme. - In: ZPSK, Bd. 24, 1971, H. 1-2, S. 91-120.
- Schweisthal, Klaus Günther: Präpositionen in der maschinellen Sprachbearbeitung. - In: Schriftenreihe zur Kommunikativen Grammatik, Bd. 1, Bonn 1971.
- Stötzel, Georg: Ausdrucksseite und Inhaltsseite der Sprache = Linguistische Reihe, Bd. 3, München 1970.