

Material zur Vorlesung

Theo Herrmann: „Einführung in die Sprachpsychologie“

Die Vorlesung gliedert sich grob in die folgenden Themengebiete:

1. *Einleitung*
2. *Sprachproduktion*
3. *Sprachrezeption*

Das Material zur Einleitung entfällt, Material zur Sprachproduktion folgt anschließend.

Material zur Sprachrezeption:

Friderici, A. D. (1999) (Hrsg.), *Sprachrezeption*, Göttingen: Hogrefe; daraus:

Kap. 1: „Das Erkennen gesprochener Wörter“ von U. H. Frauenfelder & C. Floccia (S. 1-4),

Kap.3: „Gesprochene Wörter im Satzkontext“ von P. Zwitserlood (S. 85-108),

Kap. 5: „Die Architektur des mentalen Lexikons“ von J. Engelkamp & R. Rummer (S. 155-191),

Kap. 7: „Syntaktische Struktur und Sprachperzeption“ von B. Hemforth & G. Strube (S. 243-265).

3. Der Prozeß der Sprachproduktion (Text von Theo Herrmann)

Inhalt

3.1 Allgemeines zum Produktionsvorgang

Produzieren und Rezipieren im Vergleich

Drei Prozeßstufen

Erste Stufe: Die Erzeugung der kognitiven Äußerungsbasis

Zweite Stufe: Die sprachliche Enkodierung

Dritte Stufe: Die Artikulation

3.2 Teilprozesse bei der Erzeugung der kognitiven Äußerungsbasis

Fokussieren und Selektion am Beispiel der Handlungsaufforderung

Fokussieren und Linearisieren

Parameterfixierung von Teilprozessen

Formatierung der kognitiven Äußerungsbasis

3.3 Teilprozesse bei der einzelsprachlichen Enkodierung

Allgemeine Prinzipien

Teilprozesse der einzelsprachlichen Enkodierung

Grammatische Enkodierung

Einige empirische Belege

3.4 Zur Kontrolle der Sprachproduktion

Kontrolle nur am Ende des Sprachproduktionsprozesses?

Regulationsebenen der Sprachproduktion

3.1 Allgemeines zum Produktionsvorgang

Produzieren und Rezipieren im Vergleich

Beim sprachlichen Kommunizieren spricht man selbst, oder man hört dem Partner zu. Manchmal tut man auch beides zugleich; Sprachrezeption und Sprachproduktion können sich, beispielsweise beim „Durcheinanderreden“, zeitlich überlappen. Allerdings werden nach den hierzulande gültigen Diskurskonventionen sowohl das „Durcheinanderreden“ als aber auch die „peinliche Stille“, das „allseitige Schweigen“, möglichst vermieden; es soll zu jeder Zeit des Gesprächs möglichst genau eine Person sprechen. (Sacks, Schegloff & Jefferson [1974] haben diesbezügliche Prinzipien des „turn taking“ analysiert.) Das Verhalten der Kommunikationsteilnehmer ist eng aufeinander bezogen: Davon, wie wir das von einem Partner Gesagte und Gemeinte verstehen, hängen unsere eigenen Äußerungen ab, und unsere Äußerungen wirken sich auf das sprachliche Verhalten des Partners aus, das wiederum wir selbst aufnehmen und verarbeiten müssen, usf. (Im folgenden Text behandeln wir nur die *phonetische* („mündliche“) Sprachproduktion. Auf das *Schreiben* kann aus Raumgründen nicht Bezug genommen werden. Zum Schreiben vgl. u.a. Bereiter & Scardamalia, 1987.) – Die mentalen und neuronalen Systeme, in denen die Sprachrezeption und die Sprachproduktion ablaufen, haben eine unterschiedliche Architektur und Funktionsweise. Die Sprachproduktion und die Sprachrezeption können insbesondere nicht einfach als gegenläufige, reziproke Prozesse – etwa als „Senden“ und „Empfangen“ – verstanden werden. Wir erläutern das wie folgt:

Betrachten wir die Rezeption einer sprachlichen Äußerung: Wir können das vom Sprecher mit einer Äußerung Gemeinte auch bei relativ hohem Geräuschpegel oder selbst bei erheblicher Verzerrung der technischen Übermittlung (Rundfunk, Telefon usf.) oft erstaunlich gut *verstehen*. Dies, ohne jeden Laut, jedes Wort und jede im Gesprochenen realisierte grammatische Regel identifiziert bzw. kognitiv verarbeitet zu haben. Viele Lücken können wir schließen, indem wir auf der Basis unseres Sprach- und Weltwissens schlußfolgern bzw. rekonstruieren (Rickheit & Strohner, 1993). Andererseits können wir aber dieselbe Äußerung, soll sie grammatisch wohlgeformt und situationsangemessen formuliert sein, nur dann „frei“ *produzieren*, wenn wir *jedes* ihrer Wörter und *jeden* ihrer Laute erzeugen und *alle* anzuwendenden grammatischen Regeln korrekt verwenden. Insofern ist bei der Sprachproduktion im Unterschied zur Sprachrezeption kein kognitiver Teilprozeß verzichtbar

bzw. durch andere Teilprozesse kompensierbar. Auch wenn diese Kennzeichnung nicht für alle Teilprozesse und Prozeßvarianten der Produktion und Rezeption von Äußerungen in gleichem Maße gilt, so ist doch die *Verzichtbarkeit bzw. Kompensierbarkeit von Teilprozessen* für die Sprachproduktion sehr viel weniger charakteristisch als für die Sprachrezeption.

Die Sprachrezeption kann als ein Gesamtvorgang verstanden werden, bei dem auch *parallel* ablaufende Teilprozesse gemeinsam und mit wechselndem Anteil interaktiv zum Endergebnis beitragen. Bei der Sprachproduktion handelt es sich dagegen um ein Geschehen, bei dem das Vorliegen von Resultaten eines vorgeordneten Teilprozesses zur notwendigen Voraussetzung für den Start eines nachgeordneten Teilprozesses wird; Sprachproduktion erfolgt im wesentlichen *seriell* (Levelt, 1989). Auch wenn dies ebenfalls nur in Annäherung richtig ist und auch Rückwirkungen von nachgeordneten Teilprozessen der Sprachproduktion auf vorgeordnete unterstellt werden können (s. unten), so ist die Produktion doch viel mehr als die Rezeption ein *serielles Aufeinanderfolgen von Teilprozessen*.

Die *neuronalen Grundlagen* von Sprachrezeption und Sprachproduktion lassen sich kurz wie folgt kennzeichnen (vgl. zur Einführung Crystal [Hrsg.], 1995; S. 257 ff.; s. auch Höhle, 1995; Werani, 1999): Als neuronale Basis für die Sprachrezeption ist das *Wernicke-Zentrum* von großer Bedeutung. Es ist im oberen Hinterabschnitt des Temporallappens der dominanten (meist linken) Hälfte der Großhirnrinde lokalisiert und dehnt sich bis in den Parietallappen aus. Das Wernicke-Zentrum ist besonders eng mit der Hörinde verbunden. Die neuronale Basis der Sprachproduktion ist wesentlich im *Broca-Zentrum* angelegt, das im unteren Hinterabschnitt des Frontallappens der dominanten Hirnhemisphäre liegt und das funktional in die motorischen Steuerungsareale des Gyrus praecentralis integriert ist. – Die Zuordnung der Sprachrezeption zum Wernicke-Zentrum und der Sprachproduktion zum Broca-Zentrum ist mit einer Reihe von Einschränkungen zu versehen: Zunächst ist zu beachten, daß sowohl an der Sprachrezeption als auch an der Sprachproduktion neben dem Wernicke- und dem Broca-Zentrum in komplexer Weise andere (auch subkortikale) Bereiche des Gehirns beteiligt sind. Das Wernicke- und das Broca-Zentrum und andere an der Sprachrezeption und Sprachproduktion beteiligte Hirnareale sind – offensichtlich nicht bei allen Individuen in exakt gleicher Weise (vgl. Werani, 1999) – miteinander verschaltet. Bis heute ist es nicht gelungen, die bei der Sprachrezeption und Sprachproduktion beteiligten Gehirnbereiche und ihr Zusammenspiel völlig präzise einzugrenzen. Jedenfalls dokumentieren die bei der

Sprachrezeption und der Sprachproduktion jeweils aktivierten neuronalen Verschaltungen, daß es sich bei der Sprachrezeption und der Sprachproduktion um sehr unterschiedliche und insbesondere nicht um zwei lediglich „gegenläufige“ Prozeßstrukturen handelt. Dies erweist übrigens schon der triviale Sachverhalt, daß zwar das Innenohr die perzeptive Eingangsinstantz der rezeptiven Sprachverarbeitung, aber selbstverständlich nicht die Endinstanz der Sprachproduktion ist. Und ebenso dienen zwar die Sprechorgane (Kehlkopf, Ansatzrohr usw.) der abschließenden Exekutive der produktiven Sprachverarbeitung, sie bilden aber keineswegs die Eingangsinstantz der Sprachrezeption. Die fehlende Spiegelbildlichkeit der beiden Prozesse läßt sich mittels vieler neuropsychologischer Einzelbefunde demonstrieren. Das zeigt sich beispielsweise für das keineswegs nur „gegenläufige“ *Zusammenspiel* von Wernicke- und Broca-Zentrum bei Sprachrezeption und Sprachproduktion (vgl. Gainotti, Carlomagno, Craca & Silveri, 1986), wie es unter anderem beim pathologischen Ausfall eines der beiden Zentren sichtbar wird.

Verletzungen (Läsionen) des Wernicke- wie auch des Broca-Areals sind überwiegend durch Hirndurchblutungsstörungen (insbes. durch Schlaganfälle) verursacht (Höhle, 1995; Huber, 1981). Läsionen des Wernicke-Areals führen in der Regel (s. unten) zur *sensorischen Aphasie*: Schwere Ausfälle finden sich bei der Sprachrezeption; die Lautsequenzen des Partners werden zwar erkannt, doch mangelt es an der Entschlüsselung der Sprachbedeutung. Die Sprachproduktion der Patienten ist hingegen flüssig bis zur Geschwätzigkeit. Die Patienten überdecken ihre Unfähigkeit, den Sinn von Wörtern und größeren Äußerungsteilen zu verstehen, indem sie beim eigenen Sprechen zu stereotypen Floskeln und bizarren Umschreibungen greifen oder bloß unverständliches Kauderwelsch produzieren. – Läsionen des Broca-Zentrums führen in der Regel (s. unten) zur *motorischen Aphasie*: Hier ist das Sprechen des Patienten mühsam, langsam, stockend, unflüssig. Die Grammatikalität der Sätze geht verloren; die Sätze sind kurz und wie im Telegrammstil formuliert, jedoch weitgehend sinnvoll. Das Sprachverstehen ist im allgemeinen nicht oder nur wenig gestört. Der Vergleich des *Sprechens* von Wernicke- und Broca-Aphatikern verdeutlicht, daß es zu einfach wäre, die Sprachrezeption einfach dem Wernicke-Zentrum und die Sprachproduktion dem Broca-Zentrum zuzuordnen.

Die sensorische und die motorische Aphasie können übrigens in seltenen Fällen ausbleiben oder nur schwach ausgeprägt sein, obwohl das Wernicke- bzw. das Broca-Areal zerstört sind. Andererseits werden schwere Fälle sensorischer und motorischer Aphasie bisweilen auch bei völliger oder weitgehender Unversehrtheit des

Wernicke- bzw. Broca-Zentrums beobachtet. Mit den aphatischen Störungen können übrigens in interindividuell höchst unterschiedlicher Weise auch nonverbal-kognitive Defizite einhergehen (Werani, 1999).

Drei Prozeßstufen

Man kann seine situationsspezifischen Ziele dadurch verfolgen oder auch den geltenden sozialen Konventionen dadurch gehorchen, daß man *Sprachäußerungen* erzeugt. Die Sprachproduktion kann also zum einen der Verfolgung eigener *Handlungsziele* dienen. Beispiele: Ich möchte, daß Jutta das Fenster schließt; ich möchte den Stammtisch durch einen bestimmten Witz unterhalten; ich will eine bestimmte Prüfungsfrage beantworten. Oder es handelt sich um die *Einhaltung von Konventionen*. Beispiele: Ich muß den Nachbarn, der sich nähert, grüßen; ich sollte beim Small talk keine peinliche Pause entstehen lassen, also irgendetwas sagen; ich sollte mich am Telefon formgerecht verabschieden. – Der *mentale Prozeß* der Erzeugung von Sprachäußerungen erfolgt, verwendet man eine sehr pauschale Einteilung, in *drei Stufen* (vgl. u.a. Levelt, 1989):

1. *Erzeugung der kognitiven Äußerungsbasis*: Der Sprecher generiert eine *Mitteilung* bzw. eine *Botschaft* („message“), die den momentanen Sprecherabsichten und der soeben vorliegenden Kommunikationssituation möglichst adäquat ist. Diese Botschaft ist (von wenigen Ausnahmen abgesehen) *kognitiv-nonverbal*, d.h. sie stellt eine noch nicht in einer Einzelsprache formulierte „gedankliche Struktur“ dar.
2. *Enkodierung*: Die Botschaft wird einzelsprachlich *enkodiert* bzw. *formuliert*.
3. *Artikulation*: Auf der Basis der einzelsprachlichen Enkodierung wird sprechmotorisch eine *manifeste Sprachäußerung* bzw. ein *hörbares Sprechsignal* erzeugt.

Wir geben zunächst eine kurzgefaßte *Übersicht* über diese drei Stufen der Sprachproduktion.

Erste Stufe: Die Erzeugung der kognitiven Äußerungsbasis

Die Erzeugung bzw. Bereitstellung der kognitiven Äußerungsbasis läßt sich in drei Teilprozesse untergliedern: (i) Die *Fokussierung*, (ii) die *Parameterfixierung von Teilsystemen der Sprachproduktion* und (iii) die *Formatierung*:

(i) *Fokussierung*: Der Sprecher *fokussiert* dasjenige, von dem die Rede sein wird. Anders formuliert: Das Sprecher-System stellt diejenige Information bereit, die einzelsprachlich enkodiert werden soll. Schlesinger (1977, p. 81) spricht in diesem Zusammenhang von der Erzeugung einer nichtsprachlichen „kognitiven Struktur“; der Sprecher fokussiert dasjenige, „from what he is aware of and intends to talk about“. Dabei wählt er diejenigen kognitiven Inhalte aus, die er enkodieren wird (= *Selektion*) und bringt die für die Verbalisierung ausgewählten kognitiven Inhalte in eine bestimmte Reihenfolge (= *Linearisierung*). Beispiel: Der Sprecher erinnert sich an einen bestimmten Arztbesuch, um von diesem zu erzählen. Dabei spricht er aber nicht über eine besonders attraktive Arzthelferin, die ihm sehr gefiel. Auch an diese erinnert er sich zwar, doch *seligiert* er für seine Erzählung in Anbetracht seines Kommunikationspartners und der kommunikativen Gesamtsituation andere Aspekte des Arztbesuchs: Er beginnt mit dem langen, sorgenvollen Warten im Wartezimmer, berichtet dann über eine unangenehme ärztliche Untersuchung und endet mit dem erleichterten Verlassen der Arztpraxis; diese Äußerungssequenz entspricht der *Linearisierung* der kognitiven Äußerungsbasis.

Bisweilen gehen dem Sprechen komplexe *sprachliche Planungen* voraus: Wie zum Beispiel komme ich mit jemandem ins Geschäft, der zunächst kein Interesse an einem Kauf erkennen läßt? Die Bereitstellung der kognitiven Äußerungsbasis kann aber auch *automatisch*, ganz ohne Planungsprozesse erfolgen. So kann man in der Regel ohne jede explizite Planung Anreden verwenden, grüßen, zum wiederholten Mal einen Joke erzählen, alltägliche Bestellungen aufgeben, usf.

(ii) *Parameterfixierung von Teilsystemen der Sprachproduktion*: Bei der Bereitstellung der nichtsprachlichen Basis für das einzelsprachliche Formulieren werden mentale (und zugleich neuronale) Teilsysteme des Sprechersystems, in denen die Teilprozesse der Sprachproduktion ablaufen, in situationsangemessener Weise *eingestellt*: Ihre Parameter werden fixiert, sie

werden instantiiert (Herrmann & Grabowski, 1994; s. auch unten). So kann der Teilprozeß des Artikulierens einer sprachlichen Äußerung je nach der Situationsbeschaffenheit auf Flüstern, auf normale Lautstärke oder auf lautes Rufen eingestellt werden. Über eine kommunikative Situation hinweg pflegt eine dieser Varianten im allgemeinen konstant gehalten zu werden; so muß denn auch zum Beispiel nicht bei der Artikulation eines jeden Sprachlauts immer wieder neu „festgelegt“ werden, daß im gegebenen Fall geflüstert werden soll. In gleicher Weise kann sich der Sprecher situationspezifisch (etwa im Gespräch mit einem hörgeschädigten Kommunikationspartner) auf eine besonders ausgeprägte und überdeutliche Art des Akzentuierens einstellen. Oder er wählt seine Standardsprache oder aber einen Dialekt, vielleicht auch eine Fremdsprache; er stellt sich auf eine bestimmte Sprachschichthöhe (von der anspruchsvollen Hochsprache bis zur Intim- oder Vulgärsprache) ein, er wählt entweder die Sie- oder die Du-Form, er wählt Anreden, er blendet ganze Klassen von konventionell verbotenen, zum Beispiel „politisch inkorrekten“ Wörtern oder Wendungen aus, usf. Alle diese Vorgaben haben die Eigenschaft von *situationspezifischen Instantiierungen von Prozeßparametern*; sie werden nicht bei der Produktion eines jeden Äußerungselements immer wieder neu festgelegt, was höchst unökonomisch wäre. Ändert sich allerdings die Gesprächssituation, so ändern sich auch die Parameterfixierungen: Wenn sich der Partner langsam entfernt, so steigert man im Gegenzug die Lautstärke; man kann bei der Änderung des Gesprächsgegenstands (etwa beim Übergang vom geschäftlichen zum privaten Teil eines Gesprächs) von der Standardsprache in den Dialekt wechseln, usf. Und man kann auch im Wege expliziter Planung Einstellungen *ad hoc* „aufheben“: So mag jemand trotz der Einstellung auf eine gehobene und formelle Redeweise aus kommunikationstaktischen Gründen (zum Beispiel als Teil seiner „Selbstinszenierung“) in seine Rede plötzlich einen Vulgärausdruck einstreuen.

(iii) *Formatierung*: Wir nennen den kognitiven Teilapparat, der die einzelsprachliche Formulierung bzw. Enkodierung einer Sprachäußerung besorgt, den *Enkodiermechanismus* (s. unten Abschnitt 3.3). Dieser arbeitet außerordentlich schnell und im allgemeinen robust und zuverlässig. Diese hohe Leistungsfähigkeit wird damit erkaufte, daß die kognitive, nichtsprachliche Basis der Sprachproduktion in einem genau bestimmten *Format* vorliegen muß, um vom Enkodiermechanismus überhaupt als *Prozeßinput* aufgenommen werden zu können (DeSmedt & Kempen, 1991; Garrett, 1988; Levelt, 1989). – Es kommt hinzu, daß gleiche Gedanken, gleiche Äußerungsinhalte, auch gleiche Äußerungsziele zu höchst unterschiedlichen Formulierungen führen können; die jeweils vorliegenden Gedanken,

Vorstellungen, Gefühle sind im allgemeinen als Enkodierinput unterbestimmt und müssen entsprechend *vereindeutigt* werden. Zum Beispiel möge der Sprecher seinem Kommunikationspartner mitteilen wollen, daß ein bestimmtes Mädchen einen bestimmten Jungen liebt. Diese Sachlage kann man in einfacher Weise, so wie das in der Psychologie seit langem üblich ist (vgl. Kintsch, 1974; s. auch unten), als *Proposition* (d.h. als Prädikat-Argument-Struktur) schreiben:

[Prädikat: LIEBEN (Agent: MÄDCHEN; Patient: JUNGE)]

Der Sprecher kann nun diese immer gleiche Proposition im Deutschen *alternativ* – wir geben hier lediglich Beispiele in der Präsens-Indikativ-Form – unter anderem wie folgt formulieren:

Das Mädchen liebt den Jungen.

Das Mädchen mag den Jungen gern. [Das Prädikat ist anders verbalisiert.]

Den Jungen liebt das Mädchen. [Im Deutschen kann man die Wortreihenfolge ändern.]

Der Junge wird von dem Mädchen geliebt. [Ein Passivsatz.]

Es ist das Mädchen, das den Jungen liebt. [Ein sog. Spaltsatz.]

Es ist der Junge, der vom Mädchen geliebt wird. [Ein Spaltsatz im Passiv.]

Was das Mädchen betrifft, so liebt sie den Jungen. [Ein sog. Pseudo-Spaltsatz.]

Wen das Mädchen liebt, ist der Junge. [Ein anderer Pseudo-Spaltsatz.]

Sie liebt ihn. [Pronominalisierung: wenn schon zuvor von beiden die Rede war.]

Das Mädchen liebt ihn. [Hier setzt der Sprecher voraus, daß auf den Jungen Bezug genommen wird.]

Sie liebt den Jungen. [Diesmal ist das Mädchen „präsupponiert“.]

Das Mädchen den Jungen. [Ellipse: etwa als Antwort auf die Frage „Wer liebt wen?“]

Liebt nicht das Mädchen den Jungen?! [Poetisch-emphatisch.]

Usf.

Es bedarf komplexer psychischer Mechanismen, um, wie in diesem Beispiel, eine und dieselbe kognitive Äußerungsbasis (hier: Proposition) in *genau eine von vielen möglichen* situationsadäquaten, also semantisch und pragmatisch angemessenen und zugleich grammatisch korrekten Sprachäußerungen zu transformieren (Herrmann & Grabowski, 1994, S. 361 ff.).

Zur Bereitstellung der kognitiven Äußerungsbasis gehören – zusammengefaßt – (i) die kognitive *Fokussierung* dessen, wovon die Rede sein soll, wobei die *Selektion* und *Linearisierung* eine wichtige Rolle spielen, (ii) die situationsangemessene *Parameterfixierung* (*Einstellung*) von *Teilsystemen* des Sprechersystems, in denen Teilprozesse der

Sprachproduktion ablaufen, und (iii) die für den Enkodiermechanismus geeignete *Formatierung* der Äußerungsbasis.

Zweite Stufe: Die sprachliche Enkodierung

Die fokussierte und formatierte *kognitive Äußerungsbasis* wird sozusagen Stück für Stück an den Enkodiermechanismus weitergegeben. Sie bildet den *Enkodierinput*. Der Enkodiermechanismus erzeugt auf der Basis dieses Enkodierinputs eine *Folge von Phonemen*. Dabei sind der entstehenden Phonemfolge spezifische Information über Morphem- und Wortgrenzen, Informationen über die Segmentierung in Silben, über den Betonungs- und Tonhöhenverlauf und über andere Merkmale der „gebundenen Rede“ (= Prosodie) beigegeben. Der *Output* des Enkodiermechanismus besteht also aus *Phonemsequenzen*, die mittels Zusatzinformation *markiert* sind (s. auch unten). Die markierten Phonemsequenzen dienen wiederum als *Input* für den Prozeß der *Artikulation* (= dritte Stufe der Äußerungsproduktion).

Anmerkungen zu Phonemen: Phoneme sind die kleinsten bedeutungsunterscheidenden Einheiten einer Sprache; Phoneme sind in diesem Sinne „distinktiv“. So signalisieren beispielsweise die Phoneme /b/ und /f/ die unterschiedliche Bedeutung der Wörter „binden“ und „finden“. Die Anzahl der Phoneme variiert von Sprache zu Sprache zwischen etwa 65 und weniger als zehn. Das Deutsche hat etwa 40 Phoneme. Zum Beispiel gibt es im Deutschen nicht das im Englischen übliche Phonem /ð/ (wie in „father“). (Man darf die Phoneme nicht mit den Buchstaben verwechseln, von denen es im Deutschen 26 gibt. Pelz, 1992.) – Man kann Phoneme als *Klassen von Lauten* verstehen. Während die Phoneme linguistisch (d.h. als *phonologische* Einheiten) bestimmt sind, sind die ihnen subsumierten Laute durch physiologische (insbes. sprechmotorische) und physikalisch-akustische Merkmale (d.h. als *phonetische* Einheiten) bestimmt. Die Wissenschaft von den Phonemen ist die *Phonologie*, die Wissenschaft von den Lauten ist die *Phonetik*. Man kennzeichnet Phoneme durch Schrägstriche (/ /), Laute durch eckige Klammern ([]).

Einige Beispiele zum einzelsprachenpezifischen Zusammenhang von Phonemen und Lauten: Zum Phonem /r/ gehören im Deutschen Laute, die phonetisch besonders unterschiedlich sind; das /r/-Phonem des Deutschen kann, beispielsweise je nach Mundart, durch ganz verschiedene Varianten des Zungen- und Rachen-[r] realisiert sein. Diese große Unterschiedlichkeit ist etwa im Italienischen nicht vorhanden. Im Japanischen sind die [r]- und die [l]-Laute einem und demselben Phonem subsumiert, was übrigens zu den notorischen „L-r-Velwechslungen“ bei Japanern führt, die Deutsch lernen wollen (Helfrich, im Druck). Gehauchte (aspirierte) und nicht-gehauchte (nicht-aspirierte) [k]-Laute gehören im Deutschen zu einem und demselben Morphem /k/; im Chinesischen sind sie Exemplare zweier Morpheme /k^h/ (= aspiriert) vs. /k'/ (= nicht-aspiriert) und signalisieren so verschiedene Wortbedeutungen.

Die Funktionen des Enkodiermechanismus werden bis heute theoretisch sehr unterschiedlich konzipiert (vgl. u.a. Levelt, 1989; 1992; Dell, 1989; Dell, Burger, & Svec, 1997; Schade, 1992; im Druck). Unumstritten ist jedoch, daß die einzelsprachliche Enkodierung (allenfalls neben anderen Teilvorgängen) einerseits aus der adäquaten *Wortwahl* sowie *Wortflexion* („kam“, „gesehen“ usf.) und andererseits aus der syntaktischen Strukturierung von Satzteilen und Sätzen (= *Satzbau*) besteht: Die sprachliche Enkodierung gliedert sich also (unter anderem) in *lexikalisch-morphologische* und in *syntaktische* Teilvorgänge.

Bei der theoretischen Modellierung der *lexikalisch-morphologische Enkodierung* kann man (1.) Dreikomponententheorien von (2.) Zweikomponententheorien unterscheiden (vgl. dazu auch Herrmann & Graf, 1999):

1. Die *Dreikomponententheorien* (z.B. Garrett, 1988; Levelt, 1989) implizieren die theoretische Unterscheidung der folgenden drei Kategorien: (i) Es existieren nonverbale *Begriffe* („concepts“), also mentale Repräsentationen von Dingen, Ereignissen und Sachverhalten mit ihren jeweiligen Merkmalen und Funktionen, die die Menschen unabhängig von der von ihnen verwendeten Einzelsprache kognizieren können. (ii) Es gibt sprachgebundene Einträge in ein Internes bzw. Mentales Lexikon; diese Einträge werden *Lemmata* (engl.: „lemmas“) genannt und sind demjenigen ähnlich, was man üblicherweise als *Wortbedeutungen* bezeichnet. (iii) Es gibt *Wortformen*, die aus Morphemen, Silben und Phonemen bestehen. Die Dreikomponententheorien unterscheiden also unter anderem nichtsprachliche *Begriffe* von den in Lemmata realisierten einzelsprachlichen *Wortbedeutungen*.

Anmerkungen zu Lemmata: In der *Lexikographie* bedeutet „Lemma“ (dort meist gleichbedeutend mit „Lexem“) einen lexikalischen Eintrag in seiner konventionell festgelegten, stichwortartigen (meist schriftlichen) Zitierform. Ein lexikalischer Eintrag enthält auch grammatische Informationen. Lemmata im *psycholinguistischen* Sinne sind Elemente eines *Mentalen* (mental repräsentierten) bzw. *Internen* Lexikons. Auch sie haben neben ihren semantischen Merkmalen grammatische Merkmale. So sind sie als Wortarten (vgl. Nomen, Verb, Adjektiv usf.) spezifiziert, sie können ein grammatisches Geschlecht (vgl. Maskulinum, Femininum, Neutrum) haben, usf. In psycholinguistischen Zusammenhängen werden die Ausdrücke „Lemma“ und „Lexem“ nicht gleichbedeutend verwendet, „Lexem“ bedeutet hier meist die Wortform (vgl. dazu Levelt, 1989).

Anmerkungen zu Morphemen und Silben: Morpheme sind die kleinsten bedeutungstragenden Sprachbausteine. Das Wort „Lehrerin“ besteht aus drei Morphemen: „lehr“ verweist auf eine spezifische Tätigkeit, „er“ auf den

Akteur dieser Tätigkeit und „in“ auf das weibliche Geschlecht des Akteurs. Es gibt Morpheme, die nur in Verbindung mit anderen auftreten: Das Morphem „wirsch“ kommt in der deutschen Gegenwartssprache nur im Zusammenhang mit dem Negationsmorphem „un-“ vor: „un|wirsch“. Man unterscheidet unter anderem lexikalische und grammatische Morpheme: Die lexikalischen Morpheme vermitteln Bedeutung, die grammatischen Morpheme dienen der Flexion und sprachlichen Ableitung (Derivation): In der Verbform „gestellt“ ist „stell“ ein lexikalisches, „ge-“ und „-t“ sind grammatische Morpheme. – Morpheme sind nicht mit Silben gleichzusetzen. Das viersilbige Wort „Wasserleiche“ besteht aus zwei Morphemen (Wasser|leiche). *Silben* haben immer einen Silbenkern (Nukleus) und fakultativ einen Silbenanfang (Onset) und/oder ein Silbenende (Koda). Beispiele: Die Silbe „Bast“ hat „b“ als Silbenanfang, „a“ als Silbenkern und „st“ als Silbenende; die Silbe „Blei“ hat „bl“ als Silbenanfang und „ei“ als Silbenkern, das Silbenende ist hier nicht belegt.

Beim einzelsprachlichen Enkodieren werden *zunächst* im Mentalen Lexikon für jeden zu verbalisierenden (nichtsprachlichen) Begriff ein oder auch mehrere zu ihm passende Lemmata gesucht, zwischen denen gegebenenfalls ausgewählt werden muß (= „lexikalische Suche“). Beispiel: Der Enkodiermechanismus möge auf die deutsche Umgangssprache eingestellt sein, und das Sprechersystem möge die Mitteilung enkodieren, daß der Sprecher seinem Sohn keine Träne nachweint, wenn dieser *weggeht*. Dabei muß unter anderem der *Ereignisbegriff* WEGGEHEN enkodiert werden (vgl. auch Klix, 1992). Das Mentale Lexikon mag die *Lemmata* WEGGEHEN, WEGLAUFEN, ABHAUEN, LEINE ZIEHEN und DIE PLATTE PUTZEN enthalten. Aus hier nicht interessierenden Gründen mag für die Weiterverarbeitung das Lemma ABHAUEN selektiert werden. Auch dieses Lemma hat, wie schon erwähnt, grammatische Merkmale; es handelt sich um ein Verb, das zum Beispiel nicht mit einem direkten Objekt kombiniert werden kann. (Man kann nicht sagen: „*Ich haue dich ab.“)

Ein *späterer* Teilprozeß des Formulierens, die *morphologische Enkodierung*, transformiert das Lemma ABHAUEN angesichts des zu produzierenden Satzes in die flektierte *Wortform* „ab|hau|t“, die aus den drei Morphemen „ab“, „hau“ und „-t“ besteht.

Vertreter der Dreikomponententheorien haben große Mühe drauf verwendet, besonders den Unterschied zwischen Lemma und Wortform und die strikte Aufeinanderfolge von zunächst der Generierung von Lemmata und dann erst der Wortformen experimentell nachzuweisen (vgl. Meyer & Schriefers, im Druck; Roelofs, 1997; s. auch unten). – Wir werden die heute bekannteste Dreikomponententheorie, das serielle Modell von Levelt (1989), im Abschnitt 3.3 ausführlich darstellen.

2. Die *Zweikomponententheorien* (Herrmann, 1994; Herrmann & Graf, 1999; vgl. auch Caramazza, 1997; Klabunde & v. Stutterheim [eds.], 1999) unterscheiden nur (i) *Begriffe* und (ii) *Wörter* bzw. *Morpheme*. Bei den Zweikomponententheorien wie auch in anderen Teilen der Sprachpsychologie und Linguistik wird diskutiert, ob Sprachproduktionsmodelle eher auf der Basis von Morphemen oder von (flektierten) Wörtern aufgebaut werden sollten (Miller, 1993). Je nach gewählter Position gelten die nachfolgenden Erläuterungen (a) für flektierte Wörter, die aus Morphemen zusammengesetzt sind, oder (b) für Morpheme, deren Sequenz flektierte Wörter ergeben kann (vgl. auch Herrmann & Grabowski, 1994, S. 35 ff.; Lyons, 1989).

Zweikomponententheorien gehören zu den konnektionistischen oder auch subsymbolischen Netzwerk- bzw. Aktivationsausbreitungsmodellen (vgl. Eikmeyer, Schade & Kupietz, 1995; Goschke & Koppelberg, 1990; Helm, 1991; Schade, 1992; im Druck). – Das Sprechersystem verfügt neben seinem deklarativen Langzeitspeicher („Begriffsnetzwerk“, „Wissensnetzwerk“) über ein geschichtetes *Wortgenerierungsnetzwerk*, in dem die Morpheme bzw. flektierten Wörter einer Einzelsprache *multipl repräsentiert* sind. (Auch sprachliche Wendungen, d.h. feste Wortfolgen, wie zum Beispiel „guten Morgen“, „auf Teufel komm raus“, „denk’ ich mal“, „küß’ die Hand“, können *als ganze* im Wortgenerierungsnetzwerk multipl repräsentiert sein.)

Multiple Repräsentation bedeutet folgendes: Ein Morphem bzw. ein Wort und gegebenenfalls eine ganze Wendung sind jeweils durch ein *Bündel* („*cluster*“) von *Netzwerkknoten* (Knotenbündel) repräsentiert. Die Knoten eines jeden Bündels sind untereinander eng verknüpft: Ihr jeweiliger Aktivationsbetrag wirkt auf die Aktivierungsbeträge aller anderen Knoten desselben Bündels in starkem Maße steigernd bzw. erregend (= *exzitorisch*) oder senkend bzw. hemmend (= *inhibitorisch*) ein. Zu einem *Knotenbündel* gehören die folgenden *Klassen* von untereinander eng verknüpften Knoten, die in unterschiedlichen Schichten des Wortgenerierungsnetzwerks lokalisiert sind (Schade, 1992):

- Knoten für das gesamte *Wort* (bzw. für das Morphem oder die gesamte feste Wortfolge),
- Knoten für jede ihrer *Silben*,
- Knoten für jeden *Silbenteil* (Silbenanfang, Silbenkern, Silbenende)
- Knoten für jedes ihrer *Phoneme*.

Daneben werden weitere wortspezifische Knotenarten angenommen, die sich unter anderem auf das Wissen über emotiv-bewertende Wortmerkmale [„Klangphysiognomie“] und gegebenenfalls über nichtgrammatisch-abstrakte Merkmale [z.B. französisches Fremdwort] beziehen. Auf diese letzteren Komponenten der Knotenbündel wird im gegenwärtigen Zusammenhang nicht Bezug genommen. (Vgl. dazu Herrmann, 1994.)

Die Wörter, Morpheme bzw. Wendungen sind insofern also im Wortgenerierungsnetzwerk *multipl* repräsentiert. Zum Beispiel möge der *Begriff* WEGGEHEN im Wortgenerierungsnetzwerk ein *Knotenbündel* aktivieren, durch das das flektierte Wort „abhaut“ multipl repräsentiert ist: Es sind das gesamte Wort („abhaut“), die Silben „ab“ und „haut“, die betreffenden Silbenteile (Silbenkern „a“ und Silbenende „b“ sowie Silbenanfang „h“, Silbenkern „au“ und Silbenende „t“) sowie alle Phoneme (/a/, /b/, /h/ ...) aktiviert.

Im Wortgenerierungsnetzwerk ändert sich der *Aktivationsbetrag der Netzwerkknoten* auf dreierlei Weise (Schade, 1992): (i) *Knoten gleicher Art*, zum Beispiel die Knoten, in denen Silben repräsentiert sind, beeinflussen sich untereinander *inhibitorisch*: Sind die Knoten gleicher Art nach ihrer Aktivierungsfolge „benachbart“, so hemmen sie sich besonders stark. Erhöht sich also zum Beispiel beim Wort „abhaut“ das Aktivationsniveau der ersten Silbe „ab“, so *senkt* dies das Aktivationsniveau der unmittelbar nachfolgend aktivierten zweiten Silbe „haut“. Dasselbe gilt für die Morpheme, für Wörter, Wendungen, Silbenteile und Phoneme. (ii) Andererseits besteht zwischen den *unterschiedlichen Knotenarten* eines Knotenbündels eine *exzitorische* Beziehung: So aktiviert zum Beispiel das Gesamtwort „abhaut“ die Silbe „ab“, deren Silbenteile „a“ und „b“ und die Phoneme /a/ und /b/; das Phonem /h/ aktiviert den Silbenanfang „h“, die Silbe „haut“ sowie das ganze Wort „abhaut“, und so weiter. Die Aktivationen aller Einzelknoten, die in multipler Weise jeweils ein Morphem, ein Wort bzw. eine Wendung repräsentieren, verstärken sich also gegenseitig. (iii) Außerdem desaktiviert sich jeder Knoten sofort, nachdem er sein Aktivationsmaximum erreicht hat (= *Selbstinhibition*). – Alle diese Aktivationsänderungen führen gemeinsam dazu, daß – vereinfacht formuliert – das *Aktivationsmaximum* im Wortgenerierungsnetzwerk von Wendung zu Wendung, von Wort zu Wort, von Morphem zu Morphem, von Silbe zu Silbe, von Silbenteil zu Silbenteil und von Phonem zu Phonem *wandert*. Es entsteht eine *Aktivationswelle*. Die *Aktivationswelle auf der Phonemebene des Wortgenerierungsnetzwerks* bildet zusammen mit spezifischen Zusatzinformationen aus anderen Netzwerkschichten (Silbenstruktur usw.) den Input des Artikulationsgenerators. – Es zeigt sich, daß die Zweikomponententheorien keinen klassifikatorischen Unterschied zwischen Wortbedeutungen und Wortformen machen.

Von großer Wichtigkeit ist neben der lexikalisch-morphologischen Enkodierung die *grammatisch-syntaktische Enkodierung*. Auch deren Beschaffenheit wird theoretisch unterschiedlich konzeptualisiert. Im wesentlichen findet man drei Grundauffassungen:

1. Bei der Formulierung eines Satzes wird bereits zu Beginn der Satzplanung ein *fertiges syntaktisches Satzschema* (ein grammatischer Strukturbaum) bereitgestellt, in dessen Leerstellen („slots“) die nacheinander im Mentalen Lexikon gefundenen Wörter bzw. Wortformen eingebaut werden (z.B. Bock, 1982). Von der Beschaffenheit des vorgängigen Satzschemas hängt also die Auswahl von flektierten Wortformen, von grammatischen Funktionswörtern usw. ab.
2. Das syntaktische Satzschema wird erst *allmählich*, je nach der Beschaffenheit derjenigen Lemmata entwickelt, die (zufolge der sukzessiv vorliegenden Elemente einer „message“) in einer bestimmten Reihenfolge verfügbar werden (= *inkrementelle Sprachproduktion*; vgl. Kempen & Hoenkamp, 1987; s. unten 3.3). Etwas vereinfacht formuliert: Das genaue Satzschema, das die Wortreihenfolge, die Flexionen, die Funktionswörter usw. determiniert, liegt erst fest, wenn die Auffindung aller Lemmata beendet ist (z.B. Levelt, 1989).
3. Die konnektionistischen Netzwerktheorien, zu denen die soeben dargestellten Zweikomponententheorien gehören, berücksichtigen die grammatische Steuerung der Saterzeugung wie folgt: Die grammatische Steuerung greift massiv auf die *Aktivierungsbeträge* aller im Wortgenerierungsnetzwerk enthaltenen Knotenbündel zu, die in multipler Weise Wörter, Morpheme und Wendungen repräsentieren (s. oben). Diese grammatische Ko-Determination der Knotenbündel-Aktivierung ist durch ein spezifisches Teilnetz realisiert, das aus *Plan-Knoten* aufgebaut ist (Jordan, 1986): Je nach dem gewählten grammatischen Satzschema und dem Status seiner sequentiellen Abarbeitung (bzw. nach dem jeweiligen Zustand von Mustern von Planknoten-Aktivierungen) sind viele Morpheme, Wörter oder Wendungen im Wortgenerierungsnetz völlig *inhibiert*, andere sind hingegen stark *voraktiviert*. Welche der beschriebenen *Wellen von Aktivationsmaxima* im Wortgenerierungsnetzwerk überhaupt ablaufen können, unterliegt also grammatischen Vorgaben; die Aktivationswellen, die den Output des Enkodiermechanismus bilden, sind *grammatisch ko-determiniert* (Herrmann & Grabowski, 1994, S. 401 ff.).

Dritte Stufe: Die Artikulation

Wie erwähnt, kann diese letzte Stufe der Sprachproduktion aus Raumgründen hier nicht explizit erörtert werden. Eine ausführliche Übersicht findet man bei Levelt (1989, pp. 285 ff., 413 ff.) Zur Artikulation einer sprachlichen Äußerung gehören in erster Linie die sehr komplexe Sprechmotorik und deren neuronale Steuerung sowie das hörbare *Sprechsignal* als ihr Resultat. Man beachte dabei, daß beim Sprechen mehrmals pro Sekunde größenordnungsmäßig einhundert koordinierte Nervenimpulse erzeugt werden müssen, die die Phonation in der Kehlkopfregion (Vokale und Hauchlaute), die Einstellung von Zunge, Lippen und anderen Artikulatoren (vgl. z.B. die Konsonanten [b] vs. [p]) sowie die Bereitstellung der jeweils passenden Resonanzräume (z.B. Einbeziehung der Nasenhöhle beim Laut [m]) besorgen. Zudem sind die zu erzeugenden Muster von Nervenimpulsen unter anderem davon abhängig, welche Laute und Lautgruppen beim Sprechen als *nächste* erzeugt werden sollen. (Beispiel: Das Rachen-[r] in „Rum“ und in „Rahmen“ werden sprechmotorisch unterschiedlich erzeugt.) – Neben der Phonation, der Tätigkeit der Artikulatoren und der Änderung der Resonanzräume können auch die folgenden Phänomene in einen entsprechend weit aufgefaßten Sachverhaltsbereich des Artikulierens einbezogen werden: die Lautkomposition einzelner Silben, Wörter und Wortfolgen, die phonetische Gestaltung der „gebundenen Rede“ (= die Prosodie), also der wortübergreifende Sprachrhythmus, die Betonungs- und Tonhöhenabfolgen, der Einbau von Pausen, von sprachlichen „Füllseln“ („ä“, „hm“), usf.

Wir halten fest: (a) Die Sprachproduktion läßt sich grob in drei Hauptstufen gliedern: (i) die Erzeugung der kognitiven *Äußerungsbasis*, (ii) die einzelsprachliche *Enkodierung* und (iii) die *Artikulation*. (b) Bei der Erzeugung der kognitiven (nichtsprachlichen) Äußerungsbasis sind drei Gesichtspunkte zu unterscheiden: die *Fokussierung* der zur sprachlichen Enkodierung anstehenden kognitiven Inhalte, die situationsspezifische *Parameterfixierung* von Teilsystemen des Sprachproduktionssystems und die *Formatierung* des Enkodierinputs. (c) Die einzelsprachliche Enkodierung enthält bei grober Unterscheidung (s. dazu Abschnitt 3.3) als Komponenten (i) die *lexikalisch-morphologische* („Wortwahl“) und (ii) die *grammatisch-syntaktische Enkodierung* („Satzbau“). Für beide Prozesse gibt es zur Zeit unterschiedliche Theorien: Unterscheiden lassen sich die Drei- und Zweikomponententheorien der lexikalisch-morphologischen Enkodierung sowie die drei

Hauptauffassungen zur grammatischen Steuerung der Satzproduktion: Bereitstellung eines „fertigen“ grammatischen Satzplanes zu Beginn der Satzerzeugung; inkrementelle Satzproduktion; grammatische Steuerung des Aktivationsverlaufs im Wortgenerierungsnetz durch Aktivationsmuster von Plan-Knoten. – Wir behandeln nach dieser kurzen Übersicht über den Prozeß der Sprachproduktion im folgenden (3.2) die Bereitstellung der kognitiven *Äußerungsbasis* und (3.3) die einzelsprachliche *Enkodierung* im Detail. Wir schließen (3.4) mit einer knappen Darstellung der *Kontrolle* über das eigene Sprechen.

3.2 Teilprozesse bei der Erzeugung der kognitiven Äußerungsbasis

Fokussieren und Selektion am Beispiel der Handlungsaufforderung

Die Bereitstellung der kognitiven Basis für das einzelsprachliche Enkodieren besteht wesentlich darin, sich bestimmten kommunikativen Anforderungen und zugleich bestimmten Sachverhaltsbereichen kognitiv zuzuwenden (= *Fokussieren*). Dabei ist zu beachten: Fast nie spricht man über irgendetwas „nur so“, sondern man beabsichtigt etwas mit seinem Sprechen, und sei es nur die Einhaltung einer sozialen Regel (s. oben). Wir werden diese Sachlage anschließend anhand eines Beispiels veranschaulichen und behandeln dabei zugleich einen wichtigen theoretischen Gesichtspunkt, den der *Selektion*.

Als *Selektion* ist im hier interessierenden Zusammenhang folgendes zu verstehen: Man *sagt* immer nur einen Teil dessen, was man *meint* (vgl. auch Hörmann, 1976). Anders formuliert: Das tatsächlich Enkodierte und Artikulierte entspricht immer nur einem *Teil* der jeweiligen kognitiven Äußerungsbasis. Dies nennt man das *Pars-pro-toto-Prinzip* (lat.: Teil-für-das-Ganze-Prinzip) der Sprachproduktion (vgl. Laucht, 1979; Herrmann, 1995). Der Sprecher wählt also im Zuge der Bereitstellung der Äußerungsbasis einen bestimmten *Teil* derselben zur Enkodierung und Artikulation aus. Und der Kommunikationspartner rekonstruiert aufgrund der von ihm gehörten Sprachäußerung, was der Sprecher meint und beabsichtigt (vgl. auch Rickheit & Strohner, 1993). Für solche Rekonstruktionen benötigt der Partner spezifisches *Sprach- und Sachwissen*. Über dieses partnerseitige Wissen muß auch der Sprecher verfügen, und er muß es auch seinem Partner zuschreiben und es bei ihm

einkalkulieren, wenn er genau den situationsspezifisch angemessenen Teil seiner kognitiven Äußerungsbasis für die Enkodierung und Artikulation auswählen will. Daraus ergibt sich: Erfolgreiches Kommunizieren erfordert beim Sprecher und Partner *gemeinsames Wissen* („shared knowledge“; vgl. Clark, 1996, pp. 92 ff.). Sprachproduktion und Sprachrezeption sind also wie folgt zu charakterisieren: (i) Der Sprecher wählt aus dem von ihm Gemeinten das Gesagte aus. (ii) Der Partner schlußfolgert bzw. rekonstruiert aus dem von ihm rezipierten Ergebnis dieser Auswahl das sprecherseitig Gemeinte. (iii) Dies alles erfordert beim Sprecher und Partner *gemeinsames Wissen* über Sachverhalte und über die Sprache (und vor allem über den Zusammenhang zwischen Meinen und Sagen).

Daß der Sprecher nicht alles sagt, was er meint, kann zum einen mit der für jede erfolgreiche Kommunikation erforderlichen *Sprachökonomie* erklärt werden: Beispiel: Wenn der Sprecher, voraussetzen kann, daß der Partner bereits weiß, daß sich der Sprecher beim Reden auf seinen angeheirateten Cousin Heinz bezieht bzw. daß er sich wiederholt auf diesen bezieht, muß der Sprecher nicht immer wieder „mein angeheirateter Cousin Heinz“ sagen; es genügt auch das Pronomen „er“ (= Pronominalisierung).

Zum anderen gewährt das Pars-pro-toto-Prinzip *zusätzliche Information* für den Partner: Der Sprecher wählt einen Teil der kognitiven Äußerungsbasis aus, den er verbalisieren wird (= *Selektion*). Und das Verbalisieren jeweils *genau eines von mehreren Teilen* des Gemeinten gibt dem Partner, über die Vermittlung des Gemeinten hinaus, *zusätzliche Information*. Der Partner rekonstruiert aus dem Gesagten nicht nur das vom Sprecher Gemeinte, sondern er erfährt dadurch, welcher Teil desselben verbalisiert wird, noch etwas darüber hinaus. – Betrachten wir in diesem Zusammenhang die sprachliche *Aufforderung* (vgl. dazu Blum-Kulka, House & Kasper (Eds.), 1989; Herrmann, im Druck).

Als *Auffordern* (engl.: *to request*) versteht man eine Klasse von Handlungsweisen, mit denen Personen versuchen, einen oder mehrere Kommunikationspartner zu einer Handlung zu verpflichten (Kasper, 1990). Hierbei liegt die vom Auffordernden gewünschte Handlung des Partners stets im Interesse des Auffordernden, aber nicht notwendig im Interesse des Partners (Edmonson & House, 1981). Die beim Partner veranlaßte Handlung kann in aktivem Tun, im Unterlassen oder auch in einer Erlaubnis bestehen: Bei der Aufforderung: „Trage bitte die Biotonne vors Haus!“ versucht der Sprecher, den Partner zu einem bestimmten aktiven Tun zu veranlassen; bei der Aufforderung: „Spiel nicht mit der Sicherungsdiskette!“ intendiert der Sprecher eine partnerseitige Unterlassung; „Laß mich bitte mit Fred ins Kino gehen!“ bezieht

sich auf eine Erlaubnis. (Der Einfachheit halber beschränken wir uns im folgenden auf Aufforderungen zum aktiven Tun.)

Sprachliche Aufforderungen können auf dreierlei Weise erfolgen:

- (i) Man fordert in direkter Weise durch den Gebrauch des *Imperativs* auf: „Mähe bitte den Rasen!“
- (ii) Man fordert ebenso direkt durch Verwendung sogenannter *Performativkonstruktionen* auf: Man *sagt*, daß man „auffordert“, „bittet“ u. dgl.; man verwendet also entsprechende performative Verben (in der Ersten Person Indikativ Präsens): „Ich *fordere Sie auf*, den Rasen zu mähen.“
- (iii) Man *nennt Bedingungen* unterschiedlicher Art oder man *fragt* nach Bedingungen, die nach den gültigen Konventionen einer Sprachgemeinschaft eine Aufforderung rechtfertigen bzw. begründen (Gordon & Lakoff, 1971):
 - (a) Man kann auffordern, indem man das Vorliegen der für jede Aufforderung notwendigen Bedingung behauptet, nämlich daß man den Partner überhaupt zu einer Tätigkeit o. dgl. verpflichten *will*. Dann kann die Aufforderung lauten: „Ich *will*, daß du den Rasen mähest.“ –
 - (b) Oder es kann sich um ebenfalls notwendige Bedingungen handeln, die sich auf Abmachungen, auf Sitten und Bräuche oder auch auf explizite Regeln und Gesetze beziehen (= sog. *deontische* Bedingungen): „Bist du nicht heute mit Rasenmähen dran?“ Oder: „Ich kann von dir verlangen, daß du den Rasen mähest.“ –
 - (c) Oder es handelt sich um ebenso notwendige *partnerseitige* Bedingungen der *Bereitschaft* oder des *Könnens*. Nur wenn diese Bedingungen vorliegen, sind Aufforderungen konventional gerechtfertigt und erfolgversprechend. Man kann auch auffordern, indem man das Vorliegen dieser Bedingungen behauptet oder danach fragt: „Du *willst* doch bestimmt den Rasen mähen.“ Oder: „*Kannst* du gerade mal den Rasen mähen?“ –
 - (d) Es gibt auch indirekte Aufforderungen, bei denen sich der Sprecher auf einen *defizitären Zustand* bezieht, der durch das Handeln des Partners beseitigt werden soll: „Der Rasen sieht wieder schlimm aus.“ Oder man referiert auf das *Ziel* oder den *Sollzustand*, das oder den man durch das partnerseitige Handeln erreichen möchte: „Ach, es geht doch nichts über einen gut gemähten Rasen.“ –
 - (e) Schließlich kann es sich um ganz versteckte *Fingerzeige* (engl.: „hints“) handeln: So mag jemand den Partner zum Rasenmähen veranlassen wollen, indem er sagt: „Vergiß nicht: Morgen will sich Mutti die Rosen ansehen.“ –

Alle diese *Aufforderungsvarianten* kann man in drei Gruppen zusammenfassen: (1) Die Varianten (i), (ii) und (iii/a) bilden die *direkten Aufforderungen* (*D-Aufforderungen*). (2) Die

Varianten (iii/b) und (iii/c) werden als *mäßige* (= mäßig direkte) *Aufforderungen* zusammengefaßt (*M-Aufforderungen*). Sie sind nicht so massiv wie die D-Aufforderungen, nehmen aber noch explizit auf den Partner bzw. darauf Bezug, was dieser tun soll. (3) Bei den *indirekten Aufforderungen* (*I-Aufforderungen*) (iii/d) und (iii/e) wird weder auf den Partner noch auf das gewünschte Tun des Partners explizit Bezug genommen.

Zur kognitiven Äußerungsbasis des Sprechers, der jemanden auffordert, gehören *alle* Sachverhalte von (i) bis (iii/e). Der Sprecher weiß, daß man nach unseren sozialen Regeln nur ernsthaft auffordert, wenn man eine bestimmte partnerseitige Handlung überhaupt initiieren will, wenn man hinreichend dazu befugt ist, wenn man den Partner dafür zureichend bereit und in der Lage findet und wenn man einen bestimmten defizitären Zustand beseitigen bzw. ein einschlägiges Ziel oder einen bestimmten Sollzustand erreichen will. Der Sprecher muß also immer kalkulieren, ob *alle* diese Bedingungen für sein Aufforderungshandeln vorliegen; andernfalls fordert er nicht auf. In der Regel stellt er in diesem Falle die für das Auffordern notwendigen Bedingungen in irgendeiner geeigneten Weise her und fordert erst dann auf. (Zum Beispiel schildert er in detaillierter Weise seine schwierige Situation und äußert allerlei Begründungen dafür, daß er diese Situation nicht selbst verschuldet hat, um mit dieser Taktik die noch fehlende, aber für sein Auffordern erforderliche Bereitschaft des Partners erst zu erzeugen.)

Alle genannten Bedingungen gehören zu einer *Wissensstruktur*, die man im Laufe seiner kognitiven Entwicklung als die Wissensstruktur AUFFORDERN erworben hat (Herrmann, 1983). Der Sprecher muß diese *gesamte* Wissensstruktur in Kommunikationssituationen, in denen das sprachliche Auffordern eine naheliegende Handlungsalternative darstellt, *aktivieren* und das Vorliegen aller Bedingungen *abprüfen* – sei es im Wege expliziten Überlegens und Planens oder aber automatisiert und implizit. Mit anderen Worten: Wer eine Situation so einschätzt, daß er sein Ziel am ehesten durch eine Handlungsaufforderung erreicht, *fokussiert* die gesamte Wissensstruktur AUFFORDERN bzw. stellt sie als kognitive Äußerungsbasis bereit (vgl. auch Herrmann, im Druck).

Niemand sagt aber beim Auffordern alles, was er meint; vielmehr wählt er aus der kognitiven Gesamtstruktur AUFFORDERN einen Teil aus: in der Regel *eine* der Aufforderungsvarianten (i) bis (iii/e) bzw. entweder eine D-, M- oder I-Aufforderung. Die ausgewählte und verbalisierte Aufforderungsvariante steht dann – *pars pro toto* – für die Gesamtstruktur AUFFORDERN. Und der *Partner*, der eben diese *Variante* rezipiert, rekonstruiert aus ihr die Gesamtstruktur

AUFFORDERN. Wenn der Partner also hört: „Kannst du mal den Rasen mähen?“, dann schlußfolgert er nicht, daß sich der Sprecher bloß darüber informieren will, ob er zum Rasenmähen imstande sei, sondern er versteht, daß er den Rasen mähen *soll*. Und hört der Partner: „Mäh mal den Rasen!“, so schlußfolgert er unter anderem, daß der Sprecher den Rasen ernsthaft gemäht haben will, daß der Sprecher sich zum Auffordern hinreichend befugt fühlt, daß er unterstellt, der Partner könne den Rasen mähen, usf. Falls jedoch der Partner aus dem Gesagten das Gemeinte nicht rekonstruieren kann, *mißversteht* er die Aufforderung und fühlt sich nicht aufgefordert; die Kommunikation mißglückt. Freilich kann ein Partner eine Aufforderung auch sehr wohl als solche verstehen, aber sozusagen aus taktischen Gründen vorgeben, sie mißzuverstehen:

A: „Der Rasen ist viel zu lang“ [= Variante (iii/d)]

B: „Ja, finde ich auch. Tschüß.“

Welche Aufforderungsvariante als Ergebnis einer *Selektion* verbalisiert wird, ist von vielerlei Bedingungen abhängig: Zum Beispiel darf man in Israel, verglichen mit angelsächsischen Konventionen, sehr direkte Aufforderungen manifestieren, ohne für unhöflich gehalten zu werden (Blum-Kulka, 1985). Für den deutschsprachigen Raum gelten zudem die folgenden Voraussetzungen, wobei *Standardsituationen* von *Nichtstandardsituationen* des Aufforderns unterschieden werden (Grabowski-Gellert & Winterhoff-Spurk, 1988):

- *Standardsituationen* sind häufige, alltägliche Aufforderungssituationen, in denen die Ansprüche des Sprechers an den Partner eher gering sind und der Sprecher die *Bereitschaft* des Partners und die eigene *Legitimation* als *hoch und unzweifelhaft* eingeschätzt. Hier dominieren in unserem Sprachraum I-Aufforderungen („Mein Bier ist alle.“ Oder: „Ich bräuchte noch ein Bier.“ [= (iii/d)]) sowie auch sprachliche Ellipsen („Ein Bier bitte!“). Besteht in Standardsituationen jedoch ein gewisser Zweifel am *Können* des Partners, so kommen gehäuft M-Aufforderungen von der Art vor: „Können Sie mir schnell noch ein Bier bringen?“ D-Aufforderungen fehlen hier ganz.
- *Nichtstandardsituationen* sind nicht häufig und nicht üblich; sie stellen höhere Anforderungen an den Partner (Grabowski-Gellert & Winterhoff-Spurk, 1988; Meyer, 1992.). Entsprechend ist die *Bereitschaft* des Partners nunmehr *zweifelhaft*. Ist zugleich die eigene Legitimation hoch und über jeden Zweifel erhaben, so wählt der Sprecher massive D-Aufforderungen. („Gib mir sofort meine hundert Mark zurück!“). Kogniziert er

seine Legitimation hingegen ebenfalls als schwach, so dominieren besonders höflich formulierte M-Aufforderungen (sogenannte *Wunschfragen*): „Könntest du mir vielleicht bitte mal hundert Mark leihen?“.

Diese Befunde haben sich aus einer großen Anzahl von experimentellen Untersuchungen ergeben (vgl. zusammenfassend Herrmann & Grabowski, 1994, S. 153 ff.), von denen wir eine zur Veranschaulichung kurz darstellen (Herrmann, 1983, p. 139 f.):

An einem Spiel, das Ähnlichkeit mit dem bekannten Monopoli-Spiel hat, nahmen jeweils drei Personen teil: Der Experimentator war der Spielleiter, ein Konfident des Versuchsleiters und eine Versuchsperson spielten gegeneinander. Die Versuchsperson hielt den Konfidenten für eine zweite „naive“ Versuchsperson. Die beiden Spieler befanden sich in der Rolle von Detektiven, die Aufträge ausführen müssen. In das Spiel wurde eine Situation eingebaut, in der die Versuchsperson eine Pistole dringend benötigt. Die Versuchsperson mußte den Konfidenten auffordern, ihm diese Pistole zu überlassen. Der Konfident war in einer Situation, die ihn *nicht zweifelsfrei bereit* sein ließ, die Pistole herauszurücken; er brauchte sie nämlich dringend selbst. Es handelte sich also um eine Nichtstandardsituation. In einer ersten von zwei Versuchsvarianten war die Versuchsperson *hoch legitimiert*, die Pistole vom Mitspieler zu verlangen; die Pistole war Eigentum der Versuchsperson. In einer zweiten Versuchsvariante war die Legitimation der Versuchsperson *zweifelhaft*; sie mußte sich die Pistole beim Mitspieler, der ihr Eigentümer war, ausleihen. (Das partnerseitige Können, s. oben, wurde in diesem Experiment nicht explizit variiert.) Das Spiel wurde mit 144 männlichen Versuchspersonen (Studierenden und Oberstufenschülern) durchgeführt. Das experimentelle Spiel erbrachte das folgende *Ergebnis*: Bei der Kombination *geringe partnerseitige Bereitschaft & hohe sprecherseitige Legitimation* verwendeten 68 % der Versuchspersonen D-Aufforderungen, die übrigen 32 % wählten vor allem M-Aufforderungen, daneben aber auch einige I-Aufforderungen. Beim Vorliegen der Kombination *geringe partnerseitige Bereitschaft & geringe sprecherseitige Legitimation* verwendeten 68% der Versuchspersonen M-Aufforderungen, vor allem „Wunschfragen“ (s. oben), die übrigen 32 % verwendeten I-Aufforderungen. D-Aufforderungen fehlten hier ganz.

Bei *zweifelhafter partnerseitiger Bereitschaft* (Nichtstandardsituation) beeinflußt also in hohem Maße die eigene *Legitimation* die Variantenwahl: Fühlt man sich zweifelsfrei legitimiert, so erzeugt man bevorzugt direkte Aufforderungen. Ist man jedoch in Hinsicht auf die eigene Legitimation unsicher, so produziert man vor allem „Wunschfragen“ (M-Aufforderungen). – Das Experiment wurde unter der Bedingung einer Standardsituation wiederholt: Der Partner war diesmal zur Hergabe der Pistole *ohne allen Zweifel bereit*, denn er benötigte die Pistole nicht. Jetzt dominierten sowohl bei zweifelsfreier als auch bei zweifelhafter sprecherseitiger Legitimation die M-Varianten, vor allem (iii/c). Auch hier gab es keine D-Aufforderungen.

Man beachte, daß der Sprecher auch beim Auffordern mittels der jeweils gewählten Aufforderungsvariante, d.h. mit Hilfe des jeweils verbalisierten Teils der Gesamtstruktur AUFFORDERN, dem Partner *zusätzliche Informationen* gibt. Zum Beispiel gibt der Sprecher

dem Partner zu erkennen, wie er die partnerseitige Bereitschaft oder die eigene Legitimation einschätzt.

Fokussieren und Linearisieren

Zur Bereitstellung der kognitiven Äußerungsbasis gehört es im Regelfall, die Bestandteile dessen, wovon die Rede sein soll, in eine bestimmte Reihenfolge zu bringen, also eine sequentiell geordnete „message“ zu erzeugen, aus der dann eine gleichfalls sequenzierte sprachliche Äußerung resultiert. Die Sequenz einer manifestierten Äußerung wird durch mehrere Determinanten beeinflusst: (i) Die beim sprachlichen Enkodieren entstehenden Phonem-, Morphem- und Wortfolgen sind immer auch das Resultat der *grammatischen Steuerung* (s. oben): So gibt es (von Sprache zu Sprache verschiedene) grammatische *Wortfolgeregeln*. Man darf nicht sagen: „*Du am Nachmittag heute kommst.“ Oder: „*Bert her the milk gave.“ (ii) Die Sequenz der resultierenden Äußerungsbestandteile richtet sich daneben nach der kognitiven Äußerungsbasis; diese besitzt von vornherein oder sie erhält durch entsprechende mentale Operationen eine sequentielle Ordnung. Die Herstellung dieser sequentiellen Ordnung nennt man das (konzeptuelle) *Linearisieren*; das Linearisieren ist eine operative Komponente des *Fokussierens*. (Zum Linearisieren vgl. Levelt, 1981; 1982; 1989; Ehrich, 1989.)

Die kognitive Basis einer Äußerung kann leicht und fast ohne kognitiven Aufwand linearisiert werden, wenn die zur Verbalisierung anstehende Wissensstruktur *zeitlich geordnete Sachverhalte* repräsentiert. Der Sportreporter kommentiert ein Fußballspiel im allgemeinen so, wie es abläuft: vom Anstoß bis zum Abpfiff. Berichtet man über einen Film, so beginnt man seine Erzählung in der Regel beim Filmanfang und endet vor dem Abspann. Allerdings kann man sich – unter besonderer Zielsetzung und unter kognitivem Aufwand – aus dieser von der Sache her gegebenen Sequenz lösen, man kann markante Episoden an den Anfang stellen, man kann die Geschichte „von hinten her aufrollen“, usf. Generell richtet sich die Linearisierung der kognitiven Äußerungsbasis aber nach der *immanenten zeitlichen Ordnung des Geschehens* – soweit es eine solche gibt. Vieles, worüber man spricht, besitzt indes keine immanente temporale Ordnung. Das gilt zum Beispiel für *räumliche Anordnungen*. Eine

Landschaft, eine Stadt, eine Straße, ein Haus und ein Zimmer besitzen keine „natürliche“ zeitliche Strukturierung. Wie linearisiert man solche Gegebenheiten?

Levelt (1989, pp. 138 ff.) unterscheidet generell zwei Gruppen von Gesichtspunkten, nach denen Gegebenheiten linearisiert werden: (1.) *inhaltsbezogene* und (2.) *prozeßbezogene* Determinanten.

1. *Inhaltsbezogene Determinanten*: Hierbei handelt es sich zunächst um die soeben genannten Gegebenheiten mit eigener Zeitstruktur; die zeitliche Ordnung des *Äußerungsinhalts* determiniert hier die Linearisierung. Die Linearisierung folgt, Levelt zufolge, sozusagen einer „natürlichen Ordnung“. Andere Gegebenheiten sind selbst zwar nicht immanent chronologisch strukturiert wie etwa ein Film, doch ergeben sich *Anhaltspunkte* für die Linearisierung unter anderem aus dem üblichen *Umgang* mit ihnen. Beispiel: In einem Selbstbedienungsrestaurant manifestiert man eine bestimmte *übliche Reihenfolge von Teilhandlungen*; diese übliche Handlungsfolge verwendet man dann bei einer Beschreibung des Restaurants für die Linearisierung der Räumlichkeiten, die *per se* nicht zeitlich strukturiert sind. Anhaltspunkte können sich auch aus der *Kommunikationssituation* ergeben, in der über die Gegebenheiten gesprochen wird. Beispiel: Gibt man jemandem eine Wegauskunft (Klein, 1982), so bezieht sich das Sprechen zwar auf ein Raumareal, das wiederum selbst keine immanente zeitliche Strukturierung besitzt. Doch beschreibt man einerseits nur solche Teile dieses Areal, die für die *Zielfindung des Partners* relevant sind (= Selektion, s. oben; vgl. auch Denis, 1996). Andererseits bringt man die sprachliche Bezugnahme auf die relevanten Teile des Areal derart in eine Reihenfolge, daß man zunächst über den Startpunkt, dann sukzessiv über Kreuzungen, Abzweigungen und andere wichtige nacheinander auf der Route wahrnehmbare Orientierungspunkte („Wegmarken“) und zum Schluß über dasjenige Objekt o. dgl. spricht, das den Zielpunkt bildet oder anzeigt. Dies ist sozusagen die „natürliche Weise“, bei Wegauskünften zu linearisieren. Wenn man Raumareale (zum Beispiel eine Stadt) beschreiben will, ohne aber damit eine Wegauskunft geben zu wollen, benutzt man häufig den Kunstgriff, so zu beschreiben, als wandere jemand durch das Areal. Indem der Sprecher einen solchen „imaginären Wanderer“ und dessen Weg durch das Areal als Vorgabe für die Linearisierung wählt, löst er sein Problem, das *per se* nicht zeitlich strukturierte Areal in einer für den Partner nachvollziehbaren Reihenfolge zu beschreiben (Carroll, 1993; Herrmann & Schweizer, 1998, S. 169 ff.).

Menschen pflegen häufig wiederkehrende und für sie wichtige Objekte, Objektkonstellationen und andere Umweltgegebenheiten in Form *kognitiver Schemata* zu speichern. (Vgl. allgemein zu kognitiven Schemata [auch „frames“, „scripts“] Brewer, 1987; Dutke, 1996; Graesser, Gordon & Sawyer, 1979; Grafman, 1995; Mandler & Johnson, 1977; Reimer, 1989; Rumelhart & Ortony, 1977; Schank & Abelson, 1977.) Für Kirchen, Restaurants, Küchen, Arztpraxen, Besteckschubladen, Computer und unzählige andere Gegebenheiten besitzen wir solche kognitiven Schemata. Zur Unterscheidung von den weiter unten diskutierten „Wie-Schemata“ können diese kognitiven Schemata auch als „*Was-Schemata*“ bezeichnet werden. Soweit man über solche Schemata verfügt, kennt man die *invariante* Architektur und Funktionsweise von *Klassen* von Objekten oder Objektkonstellationen, beispielsweise von Küchen. In den Schemata sind für die Besonderheiten singulärer *Schema-Exemplare*, zum Beispiel singulärer Küchen, *Leerstellen* („slots“) vorgesehen. Die Leerstellen des Schemas werden bei der kognitiven Verarbeitung eines Schema-Exemplars ausgefüllt (instantiiert). Die besondere Charakteristik des jeweiligen Schema-Exemplars (d.h. die jeweilige Ausfüllung der Leerstellen) wird zusätzlich zum Schema gespeichert (Graesser, Gordon & Sawyer, 1979). Das hat den Vorteil, daß man beispielsweise zum Zwecke der adäquaten Erinnerung an eine bestimmte Küche nicht sämtliche Details dieser Küche vollständig speichern muß. Es genügt vielmehr (a) zu wissen, daß es sich überhaupt um eine Küche handelt (= kognitives Küchen-Schema). Auf diese Weise braucht man sich zum Beispiel nicht separat zu merken, daß es einen Herd und eine Besteckschublade gibt; denn alle Küchen haben, soviel man weiß, Herde und Besteckschubladen. Außerdem (b) muß man alles dasjenige separat speichern, was man bei just dieser Küche für charakteristisch hält, daß es also zum Beispiel in dieser Küche einen Vogelbauer und auf dem Küchenschrank eine griechische Vase gibt. – Leerstellen von Schemata bleiben übrigens nie „leer“, sondern werden im Zweifelsfall zufolge erworbener Erwartungen bzw. Voreinstellungen „auf Verdacht“ instantiiert. So mag jemand aufgrund seiner vielfältigen Erfahrungen mit Küchen „auf Verdacht“ annehmen, auch die Küche, über die er soeben spricht, habe *ein* Küchenfenster.

Kognitive Schemata können nach entsprechender Ausfüllung ihrer Leerstellen als *kognitive Äußerungsbasis* fokussiert werden. Dabei besitzen die schematisch repräsentierten Gegebenheiten ihre jeweilige invariante Architektur und Funktionsweise, und sie erfordern eine bestimmte Art und Weise, mit ihnen umzugehen und haben damit so etwas wie eine „natürliche Ordnung“. Diese inhaltliche Bestimmtheit leitet die *Linearisierung* oder legt sie zumindest nahe.

Linde und Labov (1975) haben in New York City untersucht, wie Menschen ihre Wohnungen beschreiben und wie sie diese Beschreibung linearisieren. Sie fanden eine erstaunliche interindividuelle Übereinstimmung darin, daß die Beschreibung an der Eingangstür beginnt und daß dann ein virtueller Rundgang durch die Wohnung im Uhrzeigersinn oder entgegen dem Uhrzeigersinn folgt. (So gut wie niemand stellt sich zum Beispiel den gezeichneten Grundriß seiner Wohnung vor und arbeitet ihn von links nach rechts ab.) Einzelne Räume, die vom Flur oder vom Eingangsbereich abgehen, werden so beschrieben, als wenn man lediglich von der Tür aus in sie hineinschaut. Führt ein Raum aber zu einem oder mehreren weiteren Räumen, so wird so beschrieben, als gehe man durch den ersten Raum hindurch. Nach der verbalen Abarbeitung einer solchen „Zimmerflucht“ kehrt der Sprecher in den Flur oder Eingangsbereich zurück und nimmt auf seinem Rundgang den nächsten Raum in Angriff (s. unter (2.): Stapelprinzip). Auch hier sind es die Merkmale der beschriebenen Gegebenheit (Wohnung) und des Umgangs mit ihr, welche bestimmte, interindividuell hoch übereinstimmende Linearisierungen vorgeben oder nahelegen.

2. *Prozeßbezogene Determinanten*: Hier handelt es sich um vom „Inhalt“ des Geäußerten unabhängige Determinanten. Levelt (1982) unterscheidet drei Prinzipien, die unabhängig vom Äußerungsinhalt die Linearisierungsprozesse steuern: (a) Nach dem *Prinzip der Verbundenheit* werden räumlich *benachbarte* Gegebenheiten unmittelbar nacheinander fokussiert und verbalisiert. (b) Das *Stapelprinzip* wird angewendet, wenn Raumkonstellationen *Verzweigungspunkte* besitzen, wenn man also beispielsweise eine Y-förmige Anordnung beschreibt. Nach dem Stapelprinzip wird vom Verzweigungspunkt aus eine der alternativen Wegstrecken bis zum Ende abgearbeitet, bevor man *zum Verzweigungspunkt zurückkehrt* und von dort aus nunmehr die andere Wegstrecke bis zum Ende beschreibt. (c) Nach dem *Ökonomieprinzip* wird (im Falle von (b)), vom Verzweigungspunkt ausgehend, *zuerst* der *kürzere* Ast abgearbeitet. Das erklärt Levelt so, daß bei dieser Strategie sowohl der Sprecher als auch der Partner den Verzweigungspunkt weniger lange im Gedächtnis behalten müssen, als wenn zuerst der längere Ast beschrieben würde. Robin und Denis (1991) untersuchten Raumanordnungen, bei denen von einem Verzweigungspunkt gleichlange Äste abgehen. In diesem Fall wurde derjenige Ast zuerst fokussiert, der (in variabler Weise) weniger komplex und damit *schneller zu beschreiben* war. Das ist eine überzeugende Bestätigung von Levelts Ökonomieprinzip.

Eine weitere Determinante der Linearisierung von Raumkonstellationen folgt aus dem *Genese- und Ankerprinzip* (Buhl, Katz, Schweizer & Herrmann, 2000; vgl. auch Herrmann & Schweizer, 1998). Sprecher beschreiben Objekte auf Wegen (Routen) in derjenigen Reihenfolge, in der sie sie nacheinander wahrgenommen und gespeichert haben. Die

Linearisierung ist also durch die *Erwerbsreihenfolge* determiniert (= Genese-Effekt) Wenn man einen Weg während des Kennenlernens in beiden Richtungen gegangen oder gefahren ist, wenn also entgegengesetzte Erwerbsrichtungen vorliegen, so beschreibt man den Weg im allgemeinen entsprechend der *zuerst* benutzten Erwerbsrichtung; die Linearisierung entspricht der *Ersterfahrung* (= Anker-Effekt). Diese Effekte lassen sich unter anderem auch dann nachweisen, wenn zwischen der Ersterfahrung und der „freien“ Wegbeschreibung eine „erzwungene“ *andersartige* Beschreibung erfolgt. Das ist zum Beispiel dann der Fall, wenn der Sprecher zwischenzeitlich einem Kommunikationspartner den Weg von einem Startpunkt bis zu einem Zielpunkt in derjenigen Richtung beschreiben muß, die nicht mit der Wegerichtung bei der Ersterfahrung übereinstimmt. Auch dann erfolgt bei späterer „freier Beschreibung“ die Linearisierung wieder nach dem Ankerprinzip (Engelbert, 1992).

Denis (1996) gab seinen Versuchspersonen quadratische Karten vor, auf denen neun Objekte gleichabständig so abgebildet sind, daß sie drei Zeilen und drei Spalten bilden; die Karten lassen zudem keine Wege oder dergleichen erkennen. Die Objekte sollten anschließend möglichst vollzählig in freier Abfolge reproduziert werden. Schon weil die Anordnung keine Wegstruktur besaß, ließ sich hier keine Linearisierung nach dem Geneseprinzip erkennen. In 60 % der Fälle wurden die Objekte vielmehr zeilenweise genannt, 6 % der Angaben erfolgten spaltenweise, 13 % enthielten Linearisierungen im Uhrzeigersinn, der Rest entsprach keiner erkennbaren Systematik.

Parameterfixierung von Teilprozessen

Wie schon unter 3.1 beschrieben, geht die Bereitstellung der kognitiven Äußerungsbasis mit der *Einstellung (Parameterfixierung)* derjenigen mentalen (und zugleich neuronalen) Teilsysteme einher, in denen die Teilprozesse der Sprachproduktion ablaufen. Das betrifft zum Beispiel die Einstellung der Artikulationsorgane auf Flüstern, Rufen usw. und die Einstellung des Sprechersystems auf Sprachen, Dialekte, Sprachschichten und andere sprachliche „Register“ (Biber, 1995). Über diese Parameterfixierungen bzw. Einstellungen des Sprechersystems ist bis heute wenig bekannt. Im folgenden wird auf einige Gesichtspunkte der Instantiierung von Prozeßparametern hingewiesen, die auf *globale Steuerungsvarianten* des Sprechersystems zurückzuführen sind (vgl. zum folgenden Herrmann & Grabowski, 1994, S. 254 ff., 278 ff.).

Reizsteuerung der Sprachproduktion: Die Bereitstellung der kognitiven Äußerungsbasis kann in der erlernten, hochautomatisierten Bezugnahme auf eine soeben rezipierte Partneräußerung bestehen (vgl. auch die Beispiele bei Vukovich, 1999, S. 112 ff.). Überaus häufig findet man folgende „Doubletten“:

Partner: „Guten Morgen!“ – Sprecher: „Guten Morgen!“

Partner: „Bitte!“ – Sprecher: „Danke!“

Partner: „Gestatten Sie?“ – Sprecher: „Gerne.“

Entsprechend gibt es auch „Dreischritte“ oder längere Sequenzen der folgenden Art (Malone, 1995):

Partner: „Wie finden Sie das Wetter?“ – Sprecher: „Na, furchtbar.“ – Partner: „Nicht?!
Finde ich auch.“

Auch diese werden weitgehend auf der Basis erlernter Reiz-Reaktions-Verknüpfungen realisiert. Zumal die „Doubletten“ können ganz überwiegend als *konventional festgelegte* „adjacency pairs“ (etwa: „Paare von Nachbarn“; vgl. Tsui, 1991) interpretiert werden. Doch sind diese Sequenzen wie auch jeder „Small talk“ oft nicht vollständig reizgesteuert und damit geradezu „gedankenlos“; überwiegend enthalten sie auch einige Anteile, deren Erzeugung höhere kognitive Anforderungen stellt (Clark & Wilkes-Gibbs, 1986; Ungerer, Liebertz & Morgenroth, 1998).

Sprachproduktion als bloße Reproduktion: Was gesagt werden soll, kann gegebenenfalls als ganzes sozusagen wortwörtlich aus dem Langzeitspeicher in den Arbeitsspeicher kopiert werden und dient so als Grundlage für die Sprachproduktion. (Man kann selbstverständlich auch Sprachäußerungen im Langzeitgedächtnis speichern.) So produziert man auch auswendig gelernte Sprüche, Lieder, Gedichte und ähnliches:

„Einigkeit und Recht und Freiheit ...“

Schema-Steuerung der Sprachproduktion: Die Sprachproduktion kann (i) durch „*Was-Schemata*“ und (ii) durch „*Wie-Schemata*“ gesteuert sein:

(i) „*Was-Schemata*“: Von dieser Steuerungsvariante war bereits die Rede. Kognitive Schemata von der Art des Küchen- oder Wohnungsschemas können, wie ausgeführt, wegen der jeweiligen Architektur der betreffenden Gegebenheiten, ihrer Funktionsweise, aber auch wegen des üblichen *Umgangs* mit diesen Gegebenheiten bestimmte Linearisierungen, aber auch bestimmte Selektionen und andere konzeptuelle Aufbereitungen der kognitiven Äußerungsbasis *nahelegen*. Der Sprecher kann sie „als ganze“ im Langzeitspeicher aufrufen, und ihre Leerstellen können unterschiedlich instantiiert sein. Diese Schemata ko-determinieren dann die Äußerungsproduktion. Ein Beispiel sind die Wohnungsbeschreibungen nach Linde & Labov (1975), über die wir berichtet haben.

(ii) „*Wie-Schemata*“: Der Sprecher stellt ein im Langzeitspeicher verfügbares *Prozedurales Schema* („*Wie-Schema*“) bereit, also beispielsweise das erlernte Schema zum standardisierten *Erzählen* von Märchen (Märchen-Schema; Stein & Glenn, 1979), das prozedurale Wissen über das *Anfertigen* von Erlebnisaufsätzen im Deutschunterricht oder aber von Bauplänen und von Packungsbeilagen, über die *Produktion* von Zeugenberichten vor Gericht o. dgl. oder etwa das prozedurale Wissen darüber, wie man ein Kochrezept *konzeptualisiert* und *verbalisiert* (Koch, 1992):

„Es war einmal ...“

„Man nehme ...“

Die Beispiele zeigen bereits, daß die „*Wie-Schemata*“ im allgemeinen nur von begrenzten Personengruppen im Wege spezifischer Lernvorgänge erworben werden: Nicht jede(r) verfügt über etablierte kognitive Schemata für das Generieren von Packungsbeilagen und Bauplänen oder auch nur von Kochrezepten und Märchen. Die „*Wie-Schemata*“ können theoretisch den weiter oben beschriebenen *kognitiven Schemata* subsumiert werden (Herrmann & Schweizer, 1998, S. 159 ff.). Auch sie sind als standardisierte Wissensstrukturen zu verstehen, die der Sprecher „als ganze“ aus dem Langzeitspeicher abrufen und deren Leerstellen („slots“) er mit den jeweiligen Inhalten füllt, die er kommunizieren will. Man beachte aber den Unterschied beispielsweise zwischen dem kognitiven Schema von Küchen (= „*Was-Schema*“), und von Kochrezepten (= „*Wie-Schema*“). Die Prozeduralen Schemata („*Wie-Schemata*“) sind

erlernte Strategien für das *standardisierte Sprechen* über Gegebenheiten. Diese prozeduralen Schemata sind meist nur wenig inhaltlich bestimmt; ihre Leerstellen sind vielfältig instantiierbar. So kann der Sprecher einmal das Märchen von „Hänsel und Gretel“, dann wieder das Märchen vom „Sternthaler“ oder sogar ein selbsterfundenes Märchen erzählen. Oder der Sprecher mag einmal das Rezept für ein Saumagengericht und dann wieder das (geheime) Rezept für die echte Sacher-Torte mitteilen wollen. Das die Sprachproduktion steuernde Prozedurale Schema ist dann jeweils tentativ gleich; nur die Instantiierung der Schema-Slots variiert stark. – „Wie-Schemata“ können gegebenenfalls mittels „Was-Schemata“ instantiiert werden. Beispiel: Bezieht sich eine Zeugenaussage, für deren Produktion ein Sprecher (zum Beispiel ein Polizist) über ein etabliertes „Wie-Schema“ verfügt, auf eine bestimmte Küche, die er mit Hilfe eines gespeicherten Küchen-Schemas (= „Was-Schema“) gespeichert hat, so werden die „slots“ der Zeugenaussage (als einer besonderen Diskursform) mit Sachverhaltsinformation aufgefüllt, die unter anderem dem schematisierten Wissen über Küchen entnommen sind. Das Sprechersystem erzeugt hier die resultierende Äußerung im Wege des spezifischen Zusammenwirkens eines „Wie-“ und eines „Was-Schemas“ (vgl. Herrmann, Kilian, Dittrich & Dreyer, 1992).

Ad hoc-Steuerung der Sprachproduktion: Nicht immer erfolgt die Sprachproduktion reiz- oder schemagesteuert oder besteht lediglich aus der Reproduktion von Auswendiggelerntem. Wir verfügen ersichtlich nicht für alles, was wir äußern wollen, über bereits fertige oder routiniert zu realisierende Konzeptualisierungen. In unsere Rede mischen sich ständig Ad-hoc-Produktionen. Oder das Zu-Äußernde besteht ausschließlich aus einer ad hoc produzierten „message“, bei der weder die kognitive Bezugnahme auf unmittelbar vorausgehende Partneräußerungen noch auf Auswendiggelerntes noch die Mobilisierung von „Was“- oder „Wie-Schemata“ eine Rolle spielen.

Die Ad-hoc-Steuerung der Sprachproduktion – mag sie sich lediglich auf Teile von Äußerungen oder auf die jeweils gesamte Äußerung eines Sprechers beziehen – kann unter dem Gesichtspunkt beschrieben werden, daß hier die *Planung* der zu verschlüsselnden Botschaft ihrer sprachlichen Enkodierung vorausgeht.

Planendes Verhalten beruht auf oder geht einher mit *kontrollierten* (auch: „bewußten“ oder auch: „willkürlichen“, „intentionalen“, „absichtlichen“) kognitiven Prozessen. Das Planen besteht aus kognitiven Vorgängen, die in einer gewissen „Bewußtseinshelligkeit“ und auch in

bestimmter Weise „absichtsvoll“ oder auch „explizit“ ablaufen. Kontrollierte kognitive Prozesse und damit auch Planungsvorgänge enthalten die Antizipation bzw. die mentale Vorwegnahme von Verhaltenskonsequenzen (Oesterreich, 1981). Kontrollierte Prozesse erfordern Aufmerksamkeit; sie benötigen „kognitive Ressourcen“. (Die Ressourcentheorie der Aufmerksamkeit ist von Neumann [1992] sowie von Rummer, Mohr und Zimmer [1998] kritisiert worden.) Bei der Verarbeitung von Information bildet das sprachliche Planen einen „Flaschenhals“, der den schnellen und komplexen Informationsfluß behindert. Kontrollierte Prozesse laufen relativ langsam ab und sind relativ leicht störbar. – Man kann den kontrollierten Prozessen die *automatischen* bzw. *impliziten* Prozesse gegenüberstellen (Berry & Broadbent, 1984; Norman & Shallice, 1986; Reber, 1989; Perrig, Wippich & Perrig-Chiello, 1993). Automatische bzw. implizite Prozesse gehen kaum mit psychischer Anspannung („effort“) einher, sie verlaufen schnell und sind robust gegenüber Störungen (s. auch Posner & Snyder, 1975; Bosshardt, 1999; Power, 1985; Rummer, 1996).

Falls zuvor häufig erfolgreich ausgeführte *Fokussierungsroutinen* in einer neuartigen, unbestimmten, komplexen Kommunikationssituation nicht mehr zu den erwünschten Konsequenzen führen, kann ein *reflexiver Prozeß* einsetzen: Der Sprecher unterbricht den Prozeß der Fokussierung und rekapituliert oder rekonstruiert, wie er die zu verschlüsselnde Botschaft soeben konzipieren wollte und welche Schwierigkeiten dabei vorliegen. Er kann den erfolglosen Versuch bewerten, kann ihn allenfalls verwerfen und eine neue Strategie für die Bereitstellung der zu kommunizierenden Botschaft starten (vgl. Norman & Shallice, 1986; Tisdale, 1998). Die Sprechplanung kann so verstanden werden, daß das Sprechersystem seine eigenen kognitiven Prozesse „verdatet“, d.h. daß diese Prozesse und auch ihre Ergebnisse und Bedingungen im Arbeitsspeicher kurzfristig deklarativ repräsentiert werden. Über diesen Daten von kognitiven Prozessen, ihren Ergebnissen und Bedingungen laufen dann andere kognitive, vor allem bewertende und optimierende Prozesse ab. Das Planen setzt also ein, wenn die übliche Reizsteuerung oder Schema-Steuerung der Sprachproduktion (s. oben) versagen bzw. auf Hindernisse stoßen. Das Planen besteht dann in der Revision des im vorliegenden Fall erfolglosen Ablaufs automatisierter Fokussierungsprozesse. Sprachliches Planen ist dem Problemlösen strukturäquivalent, weil auch dieses ins Spiel tritt, sobald die routinemäßige Lösung einer (Routine-) Aufgabe versagt (Dörner, 1987; 1999).

Sprachliches Planen kann sich auf ganz unterschiedliche Ziele richten: Bisweilen plant man den Ablauf bzw. die Beschaffenheit kommunikativer *Gesamtsituationen*, zum Beispiel von

Sitzungen, Verhandlungen oder anderen institutionell normierten Gesprächen. Oder die Planung bezieht sich lediglich auf *Diskurssegmente*, zum Beispiel auf eine in ein zwangloses Gespräch eingestreute Erzählung. (Das impliziert, daß der Sprecher in diesem Fall kein „Wie-Schema“ für solche Erzählungen sein eigen nennt; sonst müßte er nicht planen.) In gleicher Weise kann auch eine einzelne *Teiläußerung*, ein einzelner verbalisierter *Gedanke*, ein bestimmtes *Argument* o. dgl. oder eine einzelne *sprachliche Wendung*, eine *grammatische Konstruktion* oder sogar ein einzelnes *Wort* zum Gegenstand expliziter Planung werden. (Beispiel: Wie benennt man heutzutage „politisch korrekt“ einen strohdummen Menschen?)

Bei der Reizsteuerung, der Schema-Steuerung, der Ad hoc-Steuerung und bei der bloßen Reproduktion von Auswendiggelerntem ist das Sprechersystem *unterschiedlich eingestellt*. Zum Beispiel muß der Sprecher bei der Ad hoc-Steuerung von Gesprächsbeiträgen ein viel genaueres mentales Protokoll über dasjenige führen, was er und sein Partner bereits gesagt haben, vergleicht man dies mit der Reiz-Steuerung, zum Beispiel mit einem inhaltlich beliebigen und weitgehend inkohärenten Small talk. Oder wenn man aus kulinarischem Anlaß bei der Beschreibung eines Gerichts sein erlerntes Kochrezept-Schema abarbeitet, muß man sich nicht so sorgfältig merken, was man bereits gesagt hat und was nicht, als wenn man jemandem einen komplexen Gesprächsgegenstand in freier Weise erläutert, unvorbereitet einen Vortrag darüber hält oder mit mehreren Partnern darüber diskutiert: Das mentale Teilsystem, das das mentale *Kommunikationsprotokoll* fortschreibt und speichert, ist also bei Ad hoc-Steuerung stärker aktiviert als bei Reiz- oder Schema-Steuerung (Herrmann & Grabowski, 1994, S. 336 f.). Allgemein erlaubt und erfordert die Ad hoc-Steuerung eine bei weitem größere Variabilität der Wortwahl (vgl. Herrmann & Deutsch, 1976: Benennungsflexibilität) als die Schema-Steuerung; die Leerstellen von Schemata (vgl. etwa Kochrezepte, Gebrauchsanweisungen u. dgl.) werden überwiegend mit konventionellen Begriffen gefüllt, die nur in beschränkter Weise variabel enkodiert werden können. Bei Schema-Steuerung sind häufig bestimmte Varianten der Sprachschichthöhe und generell die Wahl sprachlicher Register festgelegt: Das Sprachregister, in dem Märchen erzählt werden, ist weitgehend vorgegeben und unterscheidet sich stark vom Sprachduktus einer Bauanleitung. Dies unabhängig von der konkreten Kommunikationssituation, in der gesprochen wird. Ganz anders bei Reiz-Steuerung: Hier ist die Registerwahl extrem situationsgebunden; der Sprecher ahmt das Register des Partners nach, und dies tut der Partner ebenso. Die Register der Kommunikationspartner gleichen sich also an; es kommt, zum Beispiel in Hinsicht auf die gewählte Sprachschichthöhe, zu einem kommunikativen Gleichgewicht. Bei Ad hoc-

Steuerung ist die Registerwahl ebenfalls situationsspezifisch außerordentlich variabel. Auch die Registerwahl ist hier häufig das Ergebnis eines Planungsprozesses. – House (1996) findet bei großen Stichproben deutscher Texte generell mehr sprachliche Ad hoc-Konstruktionen als bei Textstichproben im britischen Englisch; im britischen Englisch verwendet man dagegen häufiger „Sprachroutinen“ als im Deutschen. Diese Sachlage wird von House mit verschiedenen „Mentalitäten“ der beiden Völker interpretiert: Die Deutschen neigen generell und auch in ihrem Sprachverhalten mehr zur Unsicherheitsvermeidung als die auch kommunikativ sorgloseren Briten.

Erzählungen über Selbsterlebtes, die in informellen Kommunikationssituationen zum Zwecke der Unterhaltung des Partners und zur Herstellung oder Sicherung der sozialen Beziehung zwischen Sprecher und Partner produziert werden, erfordern vom Sprecher, soweit es sich bei diesem nicht gerade um einen professionellen Entertainer handelt, in erheblichem Maße Planungsprozesse. Solche Erzählungen, besonders wenn man über irgendein zum ersten Mal erlebtes Ereignis zum ersten Mal erzählt, sind zum großen Teil *ad hoc-gesteuert*. Dagegen sind die zum Beispiel bei Zeugenvernehmungen erzeugten *Berichte* über Selbsterlebtes wesentlich *schema-gesteuert*: Der Zeugenbericht ist ein sprachliches Muster, die sich speziell für die Zwecke von Institutionen (Polizei usw.) ausgebildet hat und das auch nur im Zusammenhang mit institutionellen Vorgängen am Platze ist (Braun, 1985; Hoffmann, 1991).

Rummer, Grabowski, Hauschildt und Vorweg (1993) haben zur Produktion von Erzählungen und Berichten das folgende Experiment durchgeführt: Zwei Versuchspersonengruppen sehen einen Film über einen Warendiebstahl im Optikergeschäft. In einem anschließenden Rollenspiel machen die Versuchspersonen der einen Gruppe eine Augenzeugenaussage gegenüber einem sie vernehmenden *Polizisten*; die Versuchspersonen der anderen Gruppe erzählen ihrem *Nachbarn* in ungezwungener Unterhaltungsatmosphäre von den erlebten Ereignissen im Optikergeschäft. Es wird untersucht, ob (und wie) die Zeugenaussage eher als schema-gesteuert und die unterhaltende Erzählung eher als ad hoc-gesteuert verstanden werden können: Welche prozeduralen Unterschiede treten auf; wie wird über „dasselbe“ Geschehen bei unterschiedlicher Einstellung des Sprachproduktionssystems spezifisch *verschieden* gesprochen?

Neben einer Reihe von hier nicht interessierenden Befunden zeigt sich, daß der sachbezogene *Zeugenbericht* in Hinsicht auf die meisten Textmerkmale interindividuell bei weitem weniger stark variiert als die unterhaltende Erzählung. Die interindividuell wenig variierenden Zeugenaussagen gegenüber dem Polizisten entsprechen in ihrer Struktur demjenigen, was man in der Textlinguistik generell als Bericht bezeichnet (vgl. Gülich, 1980). Charakteristisch ist hier die Beschränkung auf die zu schildernden Ereignisse. Diese Berichtstruktur und damit das zufolge der theoretischen Voraussetzungen aktivierte Bericht-Schema werden beibehalten, wenn sich die Versuchspersonen unter *Zeitdruck* äußern.

Die *unterhaltenden Erzählungen* gegenüber dem Nachbarn gehen dagegen über die eigentliche Ereignisschilderung weit hinaus. Immer wieder werden *ad hoc* (von Sprecher zu Sprecher stark variierende) sprachliche *Selbstinszenierungen, subjektive Bewertungen, expressive Stellungnahmen, Schlußfolgerungen* verschiedener Art und auf unterschiedliche Ereignisphasen bezogene spezielle *Detaillierungen* produziert. Diese heterogene Fülle der zur eigentlichen Ereignisschilderung hinzukommenden Beigaben wird interessanterweise auch unter *Zeitdruck* beibehalten; unter *Zeitdruck* werden hier vielmehr die ereignisbezogenen Teiläußerungen der Erzählungen deutlich kürzer und lückenhaft.

Zwischen unterhaltender Erzählung und Zeugenbericht bestehen also Unterschiede in Hinsicht auf die interindividuelle Übereinstimmung der Probanden: Die Steuerung der Sprachproduktion durch ein in unserer Kultur konsensuelles Schema, das darüber hinaus sogar nur in bestimmten institutionellen Zusammenhängen manifest wird, resultiert verständlicherweise in relativ geringen Unterschieden der Sprachäußerung zwischen Individuen. Freie Erzählungen dagegen sind weitgehend *ad hoc*-gesteuert und lassen viel mehr individuelle Freiheit zu jeder Art von Ausschmückung. Diese Ausschmückungen bilden im Verlauf der jeweiligen Gesamterzählung immer wieder Neben- oder Seitenthemen; der Sprecher muß danach immer wieder zum Hauptthema, zur „eigentlichen“ Sache (hier: Verlauf des Geschehens im Optikergeschäft), zurückfinden. Dies erfordert spezielle sprachliche Maßnahmen, damit weder Sprecher noch Partner „den Faden verlieren“ (Levelt, 1989, p. 136). Schon deshalb ist die Diskursstruktur und damit auch die kognitive Basis der Sprachproduktion beim Erzählen viel komplexer als beim Zeugenbericht. Und daß es sich bei alledem um über den Gesamtdiskurs fortdauernde *Einstellungen* des Sprechersystems handelt, zeigt sich unter anderem darin, (1) daß sich der *Zeitdruck* über die gesamte Erzählung bzw. den gesamten Bericht hinweg jeweils gleichmäßig auswirkt, (2) daß diese Auswirkungen aber ganz verschieden sind. Die vielen Ausschmückungen der Erzählung sind denn auch keineswegs fakultative Zugaben, die etwa unter *Zeitdruck* entfallen könnten; sie werden auch unter *Zeitdruck* beibehalten – auf Kosten der Ausführlichkeit der bloßen Ereignisschilderung. Andernfalls würde der Sprecher den Partner nämlich nicht gut unterhalten, die Sprecher-Partner-Beziehung würde nicht mit eben diesem diskursiven Mittel gepflegt. Und darauf kommt es beim Erzählen in erster Linie an.

Formatierung der kognitiven Äußerungsbasis

Die kognitive, nichtsprachliche Basis der Sprachproduktion muß, wie bereits erwähnt, in einem genau bestimmten *Format* vorliegen, um vom Enkodiermechanismus als Prozeßinput

angenommen zu werden (DeSmedt & Kempen, 1991; Garrett, 1988; Levelt, 1989). Es kommt hinzu, daß gleiche Gedanken, gleiche Äußerungsinhalte, zu höchst unterschiedlichen Formulierungen führen können und daß diese *Unterbestimmtheit* nicht durch unterschiedliche Prozeßvarianten im Enkodiermechanismus beseitigt werden kann (s. oben). Die erforderliche Formatierung der Äußerungsbasis erfolgt durch spezifische Vorgänge in einem Teilsystem des Sprechersystems, das die nichtsprachliche kognitive Basis unter Verwendung von einzelsprachlicher Information, auf deren Nutzung das System jeweils eingestellt ist, zu einem für den Enkodiermechanismus geeigneten *Enkodierinput* aufbereitet. Diese Formatierung wird in der Sprachpsychologie unter verschiedenen theoretischen Gesichtspunkten behandelt. Wir erläutern kurz zwei theoretische Vorschläge:

(1) Nach dem *Regulationsmodell der Sprachproduktion* (Herrmann & Grabowski, 1994) besitzt das Teilsystem des Sprechersystems, in dem die kognitive Basis der Sprachäußerung bereitgestellt wird (= „Zentrale Kontrolle“) ihm funktional nachgeordnete *Hilfssysteme*, die weitgehend automatisiert, also mit geringem kognitiven Aufwand arbeiten, die untereinander vernetzt und die an die „Zentrale Kontrolle“ rückgekoppelt sind. Die Hilfssysteme erzeugen aus dem Teil der kognitiven Äußerungsbasis, der jeweils für die einzelsprachliche Enkodierung vorgesehen ist (= „Protoinput“), den Enkodierinput: Sie vereindeutigen die jeweilige Äußerungsbasis so, daß im Enkodiermechanismus *genau eine* situativ angemessene und zugleich grammatische Äußerung entstehen kann. Die theoretische Unterstellung von Hilfssystemen ergibt sich zum Beispiel aus der schon erwähnten Tatsache, daß selbst so einfache Propositionen (Protoinputs) wie [Prädikat: LIEBEN (Agent: MÄDCHEN, Patient: JUNGE)] etwa im Deutschen in jeweils eine *Vielzahl* von grammatisch korrekten Äußerungen transformiert werden können (s. die Beispiele im Abschnitt 3.1). Der zwar schnell und robust, aber unflexibel arbeitende Enkodiermechanismus ist selbst nicht in der Lage, zu einem und demselben Protoinput *genau eine* situationsadäquate wie auch grammatische Äußerungsvariante auszuwählen. – Man kann die folgenden *Hilfssysteme* unterscheiden:

- *STM-Generator*: Eines der Hilfssysteme gibt, aufgrund des Protoinputs und der *Ergebnisse der anderen, parallel arbeitenden Hilfssysteme*, dem Enkodiermechanismus die Satzart, das Tempus und den Modus des zu erzeugenden Satzes vor: So muß zum Beispiel festgelegt werden, ob gesagt werden soll: „Das Mädchen liebt den Jungen.“ oder „Wird das Mädchen den Jungen lieben?“ oder: „Könnte es das Mädchen sein, das den Jungen liebt?“ oder: „Der Junge ist von dem

Mädchen geliebt worden.“, usf. Mit dieser Festlegung werden zugleich die grammatischen Regeln prädeterniert, die jeweils bei der Generierung des betreffenden Satzes im Enkodiermechanismus angewendet werden müssen.

- *Transformationsgenerator*: Ein anderes Hilfssystem sorgt dafür, daß die zu erzeugende Äußerung auf die vorgängigen Äußerungen des Partners und die kommunikative Gesamtsituation abgestimmt wird. Auf die in einer bestimmten Gesprächssituation geäußerten Frage des Partners: „Liebt das Mädchen den Jungen?“ antwortet der Sprecher nach den hierzulande gebräuchlichen Sprachverwendungsregeln kaum mit der Äußerung: „Das Mädchen liebt den Jungen.“, sondern sehr viel eher mit: „Ja.“ Oder: „Ja, das tut es.“ Fragt der Partner zuvor: „Wer liebt wen?“, so lautet die Antwort wiederum nicht: „Das Mädchen liebt den Jungen.“, sondern diesmal: „Das Mädchen den Jungen.“ (vgl. dazu Klein, 1984). Fragt der Partner aber zum Beispiel: „Was ist passiert?“, so kann der Sprecher nunmehr durchaus antworten: „Das Mädchen liebt den Jungen.“ (Alle drei Versionen beruhen auf einer und derselben kognitiven Äußerungsbasis, die wie folgt als Proposition geschrieben werden kann: [Prädikat: LIEBEN (Agent: MÄDCHEN, Patient: JUNGE)].) – Nach unseren Höflichkeitsnormen antwortet man auf die Frage: „Könnten Sie mir bitte Ihren Namen nennen?“ weder einfach mit „Ja.“ noch in der Regel auch nicht mit: „Otto Schulze.“, sondern mit: „Ja, Otto Schulze.“ (Clark & Schunk, 1980).

- *Kohärenzgenerator*: Wieder ein anderes Hilfssystem realisiert die situationsadäquate Kohärenz und Konsistenz der Äußerung. Völlig inkonsistent sind Äußerungen, wenn man sie etwa wie folgt formuliert: „Er hatte drei Söhne. Der König war schon sehr alt. Ein König rief seine Söhne zu sich. ...“ Vielmehr wird der Zusammenhang sogleich deutlich, wenn es heißt: „Ein König hatte drei Söhne. Der König / Er war schon sehr alt. Er rief seine Söhne zu sich. ...“ Es gibt hier also eine konsistenzgarantierende Verwendungssequenz: (i) unbestimmter Artikel – (ii) bestimmter Artikel / Pronomen – (3) Pronomen. – Auch die folgende Sequenz ist nicht optimal: „Müller treibt den Ball die Linie entlang. Maier fault ihn.“ Besser ist: „Müller treibt den Ball die Linie entlang. Er wird von Maier gefault.“ Durch die Passivierung des zweiten Satzes bleibt die semantische Belegung des Satzsubjekts erhalten; es wird in beiden Sätzen etwas über Müller ausgesagt. Man spricht hier von

einer „Topic-comment-Struktur“ (Engelkamp & Zimmer, 1983). Eine andere Kohärenzstrategie besteht darin, daß man sich in einem Satz zunächst auf etwas bezieht, wovon bereits die Rede war oder was der Sprecher aus einem anderen Grund als mit dem Partner geteiltes Wissen voraussetzen kann. An das Bekannte hängt man dann etwas Neues an. Im nächsten Satz ist dieses Neue nunmehr das Alte, an das wiederum etwas Neues angehängt wird, usf.: „Hier vom Bahnhof aus gehst du zuerst zum Wasserturm. Vom Wasserturm aus gehst du zum Kaufhof. Im Kaufhof fährst du in die dritte Etage. In der dritten Etage ist hinten die Teppichabteilung. In der Teppichabteilung liegen vorn links die Berberteppiche ...“ Dies ist die „Alt-neu-Strategie“ der Konsistenzherstellung (vgl. Clark & Haviland, 1977).

- *Emphasengenerator*: Ein weiteres Hilfssystem verleiht, wieder mit den jeweils verfügbaren einzelsprachlichen Mitteln, der Äußerung die spezifische Emphase, die unter anderem der Aufmerksamkeitslenkung des Partners dient. So kann man sagen: „Es ist das Mädchen, das den Jungen liebt.“ Dieser Spaltsatz lenkt die Aufmerksamkeit des Partners speziell auf den Begriff MÄDCHEN.

Ein Experiment von Hornby (1974) demonstriert, daß mit Hilfe von *Spaltsätzen* von der Art: (a) „Es sind Rosen, die ihr ehemaliger Freund ihr geschenkt hat.“ vs. (b) „Es ist ihr ehemaliger Freund, der ihr Rosen geschenkt hat.“ die Aufmerksamkeit des Partners unterschiedlich gelenkt werden kann (vgl. Engelkamp & Zimmer, 1981). Den Versuchspersonen wurden Diapositive, von denen eines zum Beispiel einen jungen Mann mit Blumen und ein anderes ein junges Mädchen mit Blumen darstellte, kurzzeitig dargeboten. Sie hatten die Aufgabe, diese Bilder mit dem Inhalt von gleichzeitig dargebotenen Spaltsätzen von der Art (a) und (b) auf *inhaltliche Übereinstimmung* vs. *Nichtübereinstimmung* zu vergleichen. Die Antwortalternativen waren also „stimmt überein“ vs. „stimmt nicht überein“ (= Verifikationsaufgabe). Die Versuchspersonen gaben bei den Objekten, die im jeweiligen *Relativsatz* verbalisiert waren, signifikant mehr *falsche* Übereinstimmungs- und Nichtübereinstimmungsurteile ab, als bei denjenigen, auf die sich der „Es-ist“-Teil der Spaltsätze bezog. So sind es im obigen Satz (a) die Rosen, und im Satz (b) ist es der Freund, bei denen die Versuchspersonen weniger Zuordnungsfehler machten als beim jeweils anderen Objekt. Die Information, auf die der Partner besonders achtet, ist also im „Es-ist“-Teil von Spaltsätzen und nicht im Relativsatz untergebracht.

- Die Verwaltung des mentalen *Kommunikationsprotokolls*, auf das bereits hingewiesen wurde, ist die Aufgabe eines weiteren Hilfssystems.

Die Hilfssysteme arbeiten parallel und interaktiv. Der resultierende *Enkodierinput* ist das Ergebnis eines regulatorischen Abgleichs der Outputs aller beteiligten Hilfssysteme. Dieser Abgleich erfolgt nach dem Prinzip des „constraint satisfaction“ („Erfüllung von Randbedingungen“; vgl. Güsgen & Hertzberg, 1992): Jedes der parallel arbeitenden Hilfssysteme hat die Aufgabe, spezifische Ist-Werte an einen Soll-Wert anzugleichen (Beispiel Kohärenzgenerator: Bearbeitung einer vorliegenden Struktur von Propositionen nach dem Topic-comment-Prinzip; s. oben). Der Output jeder Stelloperation eines Hilfssystems ist eine relevante Einflußgröße („Störgröße“) für die Ist-Werte aller anderen Hilfssysteme. Das Sprechersystem hat also das Problem, zum Zwecke der Erzeugung eines adäquaten und in sich konsistenten Enkodierinputs bei *jedem* Hilfssystem bzw. im *gesamten* Netzwerk von Hilfssystemen die Bedingung optimaler Ist-Soll-Angleichung zu erfüllen. – Die Hilfssysteme benötigen für ihre Arbeit bereits einzelsprachliche Information (vgl. zum Beispiel den STM-Generator), ohne jedoch schon die eigentliche Enkodierung zu leisten. Der Protoinput bzw. die kognitive Äußerungsbasis werden vielmehr mittels spezifischer Informationen aus der Hilfssystemebene „markiert“ (= markierter Enkodierinput) und ergeben so den eindeutigen Input des Enkodiermechanismus (vgl. Herrmann & Grabowski, 1994, S. 361 ff.).

(2) Chafe (1976) nimmt an, daß Menschen zum Zweck des Sprechens zunächst einen spezifischen Teil ihres Weltwissens für die Verbalisation bereitstellen (s. oben: Selektion) und die jeweils ausgewählten Wissensstrukturen „schematisieren“, also sozusagen vorstrukturieren: Der ausgewählte Teil des Weltwissens wird in *Propositionen* (Prädikat-Argument-Strukturen) gegliedert, und es werden für die mit Wörtern zu bezeichnenden Begriffe propositionale Rollen (Prädikat, Agent-Argument, Patient-Argument usw.) festgelegt (vgl. dazu auch Grabowski, 1991). Auch nach Bock (1982) werden zur Vorbereitung der sprachlichen Verschlüsselung gedankliche Inhalte in ein Format gebracht, welches diese überhaupt erst zur einzelsprachlichen Enkodierung geeignet macht: Die Glieder eines Gedankengefüges werden in einer bestimmten Reihenfolge angeordnet, und es wird so zerlegt, daß wiederum propositionale Strukturen von Begriffen entstehen, welche jeweils eine bestimmte propositionale Rolle einnehmen. Das Ergebnis dieses Vorgangs, der nach Bock in einer von ihr so genannten „referentiellen Arena“ abläuft, ist noch strikt nonverbal.

Bei Levelt (1989, pp. 75 ff., pp. 161 ff.) generiert das Sprechersystem „lexikalische“ Konzepte. Diese Konzepte werden ebenfalls als Elemente von Propositionen aufgefaßt. Doch

berücksichtigt auch dieser Autor, daß die propositionale Schreibweise, so wie sie traditionell verwendet wird, nicht alle Informationen berücksichtigt, die der Enkodiermechanismus als Input braucht. (Das wurde soeben unter (1) ausgeführt.) Levelt verwendet im Anschluß an Bresnan (1982), Jackendoff (1983) sowie Kempen und Hoenkamp (1987) ein sehr viel detaillierteres propositionales Beschreibungsmodell für „messages“ und vermeidet damit die *Unterbestimmtheit* des Protoinputs (s. oben (1)). Levelt beschreibt konzeptuelle Strukturen mittels Graphen, wobei die lexikalischen Begriffe Glieder dieser Strukturen sind (vgl. *Abbildung x1*). Es ist hier nicht möglich, Levelts komplexes Beschreibungsmodell auch nur annähernd vollständig darzustellen. Wir können nur an einem Beispiel den folgenden Einblick in Levelts Konzeption geben:

Bei üblicher propositionaler Schreibung (vgl. oben Chafe, 1976) wird zum Beispiel die Sachlage, daß das Kind der Mutter die Katze gab, wie folgt dargestellt:

[Prädikat: GEBEN (Agent: KIND, Rezipient: MUTTER, Objekt: KATZE)].

Die so repräsentierte Information reicht nun aber für die einzelsprachliche Generierung einer Äußerung nicht aus. Zum Beispiel ist nicht dargestellt, daß sich das Ereignis in der Vergangenheit abgespielt hat (vgl. unter (1) den STM-Generator).

Levelt (1989, pp. 161 ff. u.a.O.) rekonstruiert den Sachverhalt, daß das Kind der Mutter die Katze gab, wie folgt (vgl. zum folgenden Abb. x1): Das Sprechersystem verfügt während der Sprachproduktion über eine *propositionale Teilstruktur*, die zur semantischen Kategorie der Ereignisse (EVENT) gehört und die den lexikalischen Begriff KIND enthält. Das Ereignis, zu dem KIND gehört, hat im einzelnen die folgenden Merkmale: (1) Das Ereignis geschah in der *Vergangenheit* (PAST). (2) Zum Ereignis gehört das Prädikat *Verursachen* (CAUSE); d.h. es wird etwas verursacht. (3) Kern des Ereignisses ist eine Person (PERSON). (4) Diese Person-Kategorie ist mit dem *Begriff* KIND belegt. (5) Das Ereignis hat eine semantische Verbindung zu einem *anderen Ereignis* (EVENT); dieses Ereignis impliziert, daß eine Katze vom Kind zur Mutter transferiert wird.

Bei Levelt (p. 164) ist dieses *zweite Ereignis* wie folgt konzipiert: (1) Eine semantische Kategorie DING (THING) ist mit dem Begriff *Katze* belegt. (2) Es existiert eine semantische Kategorie GO_{poss}, die man als BESITZERWECHSEL bezeichnen kann: Die Katze wechselt den

Besitzer. (3) Außerdem besteht ein PFAD (PATH) von einer PERSON₁ (KIND) zu einer PERSON₂ (MUTTER). – Es ergibt sich die folgende propositionale Gesamtstruktur: Indem das Kind in der Vergangenheit (PAST) etwas verursacht (s. oben: CAUSE), wird KATZE possessiv von KIND zu MUTTER transferiert (GO_{poss}).

Man kann Levelts propositionale Darstellung des hier als Beispiel herangezogenen Sachverhalts mit der obigen traditionellen Schreibweise [Prädikat: GEBEN (Agent: KIND, Rezipient: MUTTER, Objekt: KATZE)] vergleichen und mit Levelt unterstellen, daß diese konzeptuelle Struktur nicht nur eine *vollständigere Beschreibung* eines Sachverhalts liefert, sondern daß sie die *tatsächlich existierende kognitive Äußerungsbasis* einer Sprachäußerung wiedergibt. Dann liegt ein hinreichend *vereindeutigter* Enkodierinput vor (s. unten 3.3).

— hier etwa Abbildung x1 —

Solche propositionalen Strukturen enthalten, wie man leicht sieht, bereits grammatische Information. Sie ist nach Levelt Bestandteil des Enkodierinputs, also des Inputs eines mehrgliedrigen Formulierungsprozesses. Man kann diese Information als *proto-grammatische* Information bezeichnen. Ob zum Beispiel einem lexikalischen Konzept wie KIND in einer Einzelsprache ein Nomen bzw. der Kern einer Nominalphrase entspricht, wird nicht erst während der einzelsprachlichen Enkodierung festgelegt, sondern ist schon durch die proto-grammatische Information der kognitiven Äußerungsbasis vorgegeben: Der Begriff KIND gehört nämlich zur Kategorie PERSON; das KIND verursacht etwas (CAUSE).

Auch in Levelts Konzeption bleibt ungeklärt, wie die konzeptuellen *Teilprozesse* auf der Ebene der Bereitstellung der kognitiven Äußerungsbasis genau beschaffen sind, die aus irgendwelchen gedanklichen Inhalten so detailliert beschreibbare „messages“ entstehen lassen: Wie hat man sich also die Teilprozesse vorzustellen, die irgendwelche Gedanken, bildhafte Vorstellungen, auch Gefühlsanmutungen und dergleichen zu so präzise und hochdetailliert beschriebenen Begriffsstrukturen machen? Welcher kognitive Vorgang führt zum Beispiel dazu, daß in unserem Beispiel die Kategorien GO_{poss}, PATH oder CAUSE auftreten?

Gemeinsam ist den dargestellten wissenschaftlichen Positionen (1) und (2), daß kognitive Inhalte, wie auch immer sie zuvor intern repräsentiert sein mögen, zum Zwecke der

Formatierung in einzelne *Begriffe bzw. Konzepte* zerlegt und zugleich in die Form *propositionaler Strukturen* gebracht werden, wodurch die einzelnen Begriffe ihre *propositionale Rolle* erhalten. – Die theoretische Konzeptualisierung gedanklicher Inhalte als (Strukturen von) *Propositionen* ist nicht ohne Kritik geblieben (so u.a. Dörner, 1999; Johnson-Laird, 1980). Zumindest hat es sich aber schon aus Gründen der theoretischen „Sparsamkeit“ als nützlich erwiesen, die kognitive Äußerungsbasis und ihre Formatierung mit Hilfe von Propositionen (einschließlich der detaillierten Leveltschen Graphen) zu *interpretieren*. Propositionen haben dann die wissenschaftstheoretische Funktion sogenannter „hypothetischer Konstrukte“, die sich in der Forschungspraxis bewähren können, bei denen aber auch nicht auszuschließen ist, daß sie in Zukunft durch überlegene theoretische Konstruktionen ersetzt werden (vgl. Gardenne, 1984).

Kempen und Hoenkamp (1987) sowie Levelt (1989; vgl. auch DeSmedt & Kempen, 1991) legen Wert auf den Gesichtspunkt, daß die Formatierung wie alle übrigen Teilprozesse der Sprachproduktion nach dem Prinzip der *inkrementellen Verarbeitung* („incremental processing“) erfolgt: Jede nachgeordnete Prozeßstufe startet ihre Arbeit erst dann, wenn der Input der vorgeordneten Prozeßstufe vorliegt. Genauer: Die in zeitlicher Sequenz verfügbaren Input-Elemente der vorgeordneten Stufe stoßen auf der nachgeordneten Prozeßebene sogleich Stück für Stück die ihnen zugeordneten spezifischen Arbeitsschritte an. Es gibt keine Rückkopplung von der nachgeordneten zur vorgeordneten Stufe. – Demnach beginnt, Levelt zufolge, die Suche nach dem *ersten Wort* (Lemma), das für die Verbalisierung einer „message“ geeignet ist (s. unten Abschnitt 3.3: lexikalische Enkodierung), erst bei Verfügbarkeit der *ersten* („linken“) *Teilstruktur* des Propositions-Graphen (s. Abb. x1: linke Teilstruktur mit dem Begriff KIND). Sie beginnt jedoch sogleich mit der Suche nach dem *ersten* Lemma, bevor bereits die *gesamte* „message“ vorliegt (vgl. dazu auch Caramazza, 1997).

Dies schließt folgendes nicht aus: Die Stufen bzw. Ebenen der Sprachproduktion arbeiten in der Regel *zeitlich parallel*. So möge auf der Stufe der Wortsuche soeben ein Lemma L_i aufgesucht werden, mit dem ein Input-Begriff B_j enkodiert werden soll. Dabei stand der Begriff B_j zwar als Teil der „message“ schon etwas früher zur Verfügung als das zugehörige Lemma L_i – eben weil B_j der Input ist, der erst zum Aufsuchen von L_i führte. Doch während auf der Stufe der Wortsuche die Suche nach dem Lemma L_i erfolgt, wird auf der Stufe der „message“ nunmehr bereits *gleichzeitig* der nächste Begriff B_{j+1} generiert bzw. spezifiziert. Alle Stufen der Sprachverarbeitung sind also in der Regel gleichzeitig in Betrieb; dabei beginnt aber jede nachgeordnete Stufe ihre Stück-für-Stück-Produktion, sobald ihr die vorgeordnete Stufe das erste Ergebnis ihrer Stück-für-Stück-Produktion liefert.

3.3 Teilprozesse bei der einzelsprachlichen Enkodierung

Wie schon unter 3.1 berichtet, wird der Vorgang der einzelsprachlichen Enkodierung bzw. der Formulierung mittels unterschiedlicher theoretischer Modelle beschrieben (vgl. zu theoretischen Modellen der Sprachproduktion Herrmann, im Druck). Im folgenden soll der mehrstufige Enkodierprozeß im Anschluß an das *Sprachproduktionsmodell von Levelt* (1989; vgl. auch Levelt, 1992; sowie Bock & Levelt, 1994; Roelofs, 1997) erläutert werden. Dieses Modell gehört zu den im Abschnitt 3.1 beschriebenen *Dreikomponententheorien* (vgl. auch Garrett, 1988). Diese Modellklasse unterscheidet (i) nichtsprachliche *Begriffe*, also mentale Repräsentationen von Dingen, Ereignissen und Sachverhalten mit ihren jeweiligen Merkmalen und Funktionen („concepts“), (ii) Einträge in ein Internes bzw. Mentales Lexikon, die *Lemmata* („lemmas“) genannt werden, sowie (iii) *Wortformen* (Morpheme und Morphemgruppen, die u.a. Silbenstruktur besitzen). Nichtsprachliche *Begriffe* sind hier also theoretisch von den in den Lemmata realisierten *Wortbedeutungen* getrennt.

Allgemeine Prinzipien

Wie auch Garrett (1988) vertritt Levelt die folgenden theoretischen Grundpositionen:

- *Serialität*: Die an der Sprachproduktion beteiligten Teilprozesse arbeiten strikt *seriell*. Eine *serielle* Verschaltung von Teilprozessen liegt vor, wenn spezifische Informationselemente aus der jeweils vorgeordneten Prozeßebene als Input für die jeweils nachgeordneten Prozeßebene vorliegen müssen, bevor eine Weiterverarbeitung dieser Elemente auf der nachgeordneten Ebene beginnen kann (s. oben: inkrementelle Verarbeitung). So muß erst ein bestimmter Begriff auf der Ebene der „message“ vorliegen und als Input an die nächstfolgende Ebene der lexikalischen Enkodierung (s. unten) weitergegeben werden, bevor auf dieser Ebene ein passendes Lemma gesucht werden kann. Von den jeweils niedrigeren Ebenen gibt es zu den jeweils höheren Ebenen keine Informationsrückmeldung.

- *Modularität*: Die für jeden Teilprozeß benötigte Information ist *spezifisch*, d. h. sie steht nur dem jeweiligen Teilprozeß zur Verfügung. So kann auf Lemmata lediglich während des Teilprozesses der lexikalischen Enkodierung zurückgegriffen werden; Information über die phonologische Form von Wörtern steht nur auf der Ebene des phonologischen Teilprozesses zur Verfügung, usf. In diesem Sinne ist Levelts Modell durch das Prinzip der Modularität (Abgekapseltheit) der beteiligten prozeduralen Teilstrukturen gekennzeichnet (vgl. Fodor, 1983).

Teilprozesse der einzelsprachlichen Enkodierung

Levelt zerlegt den Gesamtprozeß der Sprachproduktion in *drei autonome Instanzen*, die er als „*conceptualizer*“ „*formulator*“ und „*articulator*“ bezeichnet (Levelt, 1989). Nach einer detaillierteren Einteilung zerfällt der Sprachproduktionsprozeß in *sechs strikt seriell verschaltete Prozesse* (Prozeßebenen). Der „*conceptualizer*“ entspricht der Ebene 1, der „*formulator*“ bzw. die einzelsprachliche Enkodierung umfaßt die Ebenen 2 bis 4, und der „*articulator*“ umfaßt die Ebenen 5 und 6. (Wir werden die *grammatische Enkodierung* im Anschluß an die folgenden sechs Prozeßebenen separat behandeln.)

Ebene 1: Aus einem *nichtsprachlich-konzeptuellen Prozeß* resultieren die im Abschnitt 3.2 behandelten „*messages*“. Diese bestehen aus Graphenstrukturen, zu deren Knoten *lexikalische Konzepte* gehören. Die Graphenstrukturen haben *proto-grammatische Eigenschaften*, sie besitzen Merkmale wie prädikative Relationen (z.B. GO_{poss}), Person, Ding, Vergangenheit, Pfad, Numerus, usf. (s. oben).

Ebene 2: Die auf Ebene 1 erzeugte „*message*“ stößt die *lexikalische Enkodierung* an. Mit diesem Teilprozeß werden den lexikalischen Konzepten Stück für Stück Lemmata zugeordnet, die in einem Mentalen Lexikon gespeichert sind. Die Lemmata sind interne Repräsentationen, die ebenfalls grammatische Merkmale (Wortart, grammatisches Geschlecht usf.) besitzen. Meist wird einem „*lexikalischen*“ Konzept nur ein „*lemma*“ zugeordnet, bisweilen aber auch mehrere. (Dem lexikalischen Konzept WEGGEHEN können zum Beispiel in der „*Ganovensprache*“ die *drei* Lemmata DIE | PLATTE | PUTZEN entsprechen.)

Aus den Lemmata werden im Wege der nächsten beiden Teilprozesse die *Wortformen* erzeugt:

Ebene 3: Zunächst werden die Lemmata *morphologisch* enkodiert; es entstehen Morpheme. Auf der Basis eines Lemmas und einiger zusätzlicher Kriterien, die sich aus der „message“ (Ebene 1) herleiten (z.B. Tempus, Numerus), werden ein oder mehrere Morpheme generiert (z.B. „kam“, „Hund|e“). Mehreren Lemmata (mit ganz unterschiedlicher Bedeutung) können auch dieselben Morpheme zugeordnet sein (z.B. „Absatz“).

Ebene 4: Auf die Morpheme greift der *phonologische Enkodierungsprozeß* zu, der auch *metrische* Wortmerkmale (Silbenstruktur) generiert. Nicht für jedes einzelne Morphem ist ein separates phonologisches Muster gespeichert: vgl. „lçbìn“, “libéndig“, „lçbhaft“ usf. Die phonologische Enkodierung erfolgt vielmehr meist per Anwendung von einzelsprachlichen Regeln; dabei treten auch Silbenbildungs- und Betonungsregeln der jeweiligen Einzelsprache ins Spiel. Die phonologische Enkodierung ist also *regelbasiert*. Das Resultat ist eine „abstrakte“ Repräsentation von Wortformen, die noch nicht phonetisch-artikulatorisch spezifiziert sind.

Ebene 5: Die *phonetische* Enkodierung erzeugt aus den Ergebnissen der phonologischen Enkodierung phonetische Pläne für einzelne Wörter und für die „verbundene Rede“.

Ebene 6: Durch den nach dem phonetischen Plan gesteuerten sprechmotorischen Vorgang der *Artikulation* entsteht das hörbare Sprechsignal.

Grammatischen Enkodierung

Die *grammatische Enkodierung* einer „message“ erfolgt bei Levelt auf der Ebene der lexikalischen Enkodierung (Ebene 2). Die proto-grammatischen Strukturen der „message“ (Ebene 1) bilden die Input-Information für einen Prozeß, der einerseits zum Aufruf von passenden Lemmata (s. oben: lexikalische Suche) und andererseits zur inkrementellen Generierung einer *syntaktischen Satzoberfläche* führt. Bei der Generierung der syntaktischen Satzoberfläche interagieren die Auffindung von Lemmata und spezielle grammatische Operationen in komplexer Weise. Die Satzstruktur wirkt sich auf die Ebenen 4 und 5 aus: Die

phonologische und phonetische Enkodierung sind in starkem Maße von der zuvor erzeugten Satzstruktur abhängig. Der Prozeß der grammatischen Enkodierung wird von Levelt im Anschluß an Kempen und Hoenkamp (1987; vgl. auch die spätere Erweiterung bei DeSmedt & Kempen, 1991) als ein *formaler prozeßgrammatischer Algorithmus* beschrieben. Dieser Formalismus kann im gegenwärtigen Zusammenhang nicht expliziert werden (vgl. ausführlich Levelt, 1989, pp. 236 ff.).

Zum generellen Verständnis genügt die folgende Erläuterung: Der Algorithmus von Kempen und Hoenkamp (1987) setzt voraus, daß die auf Ebene 1 plazierten Elemente bzw. Teilstrukturen einer „message“ in einem *strikten Nacheinander* als Input für die Ebene 2 verfügbar werden. Das hat zur Folge, daß auch die den einzelnen lexikalischen Begriffen zugeordneten Lemmata in strikter Reihenfolge im Mentalen Lexikon aufgesucht werden (vgl. oben: inkrementelle Verarbeitung). Die grammatischen Merkmale des jeweils aufgerufenen Lemmas (z.B. Wortart, grammatisches Geschlecht) lösen auf Ebene 2 Operationen (insbes. kategoriale Prozeduren verschiedener Ordnung) aus, die zur Herstellung vollständiger Satzphrasen und zum Aneinanderbauen mehrerer Satzphrasen zum Gesamtsatz führen (Levelt, 1989, p. 237). Die sukzessiv generierten Lemmata stoßen den Aufruf weiterer Lemmata an (z.B. kann im Deutschen ein Nomen je nach Kasus und Genus einen bestimmten definiten Artikel aufrufen) und erzeugen so Stück für Stück die gesamte syntaktische Satzstruktur. Die Satzstruktur ist also, im Unterschied etwa bei Garrett (1988; vgl. oben Abschnitt 3.1), nicht als ein Satzrahmen zu verstehen, der bereits zu Beginn der Satzproduktion als ganzer zur Verfügung steht und in dessen „slots“ die aus dem Mentalen Lexikon abgerufenen Lemmata nacheinander eingefüllt werden. Vielmehr entsteht die Satzstruktur erst *peu à peu* zusammen mit dem Nacheinander des Auffindens von Lemmata und der Erkennung ihrer semantischen und grammatischen Merkmale.

Wir kommen zur Veranschaulichung auf unser obiges Beispiel zurück, das auf traditionelle Weise durch die propositionale Schreibung [Prädikat: GEBEN (Agent: KIND, Rezipient: MUTTER, Objekt: KATZE)] dargestellt werden kann. Die Ebene 2 (Lexikalische Enkodierung) erhält als ersten Input die „linke“ propositionale Teilstruktur (EVENT), in der KIND enthalten ist (s. Abb. x1). Auf Ebene 2 wird zu KIND der passende Eintrag im Mentalen Lexikon, also ein Lemma, gesucht. Bei Einstellung auf die englische Sprache möge das Lemma CHILD gefunden werden. Durch bestimmte kategoriale Operationen werden grammatische Merkmale von CHILD spezifiziert: CHILD kann (als Nomen) nur Kern einer Nominalphrase sein;

desgleichen hat CHILD das Merkmal *Singular*. Auch haben Nomina (im Englischen) *Artikel*; so möge als Lemma der definite Artikel THE ausgewählt werden. THE wird im Englischen als Artikel „links“ von CHILD angeordnet: Es entsteht die Nominalphrase THE CHILD. Da CHILD das erste gefundene Lemma war, erhält diese Nominalphrase die Anfangsposition des Satzes. (Wäre in einem anderen Falle der Begriff KATZE der erste Input gewesen, so bestünde der Satzanfang aus THE CAT.) Nach hier nicht im einzelnen darzustellenden Operationen wird THE CHILD als *Satzsubjekt* deklariert. KIND als PERSON verursacht etwas (CAUSE). In der „message“ wird nun eine zu der gefundenen Nominalphrase THE CHILD passende Teilstruktur gesucht, der unter anderem das propositionale Prädikat GO_{poss} entnommen wird. (GO_{poss} bedeutet, wie erinnerlich, daß ein Ding auf einem „besitzspezifischen Pfad“ von einer Person₁ zu einer Person₂ transferiert wird.) Für dieses Prädikat wird im Mentalen Lexikon, auch unter Beachtung der Kategorie CAUSE, ein Verb-Lemma gefunden: GIVE. – Der Satzbildungsprozeß wird in dieser Weise fortgesetzt, bis die gesamte syntaktische Oberflächenstruktur festliegt. Nach Durchlaufung aller sechs Produktionsebenen resultiert schließlich die artikulierte Sprachäußerung: „The child gave the mother the cat.“

Die Wichtigkeit der *Reihenfolge* der Inputs aus Ebene 1 verdeutlicht sich übrigens, wenn man in derselben Weise die Enkodierung von Propositionen wie [LIEBEN (Agent: OTTO, Patient: ANNA)] analysiert: Hier hängt es entscheidend von der Input-Reihenfolge ab, ob die Äußerungen „Otto loves Anna.“ oder „Anna loves Otto.“ resultieren. Im Deutschen kann übrigens die Proposition [LIEBEN (Agent: OTTO, Patient: ANNA)] bei entsprechender Satzmelodie (starke Betonung von „Anna“) korrekt in die Äußerung überführt werden: „Anna liebt Otto.“ Dies etwa als eine etwas erstaunte Antwort auf die Frage: „Wen liebt Otto?“ (Im Deutschen ist die Wortfolge, im Unterschied etwa zum Englischen, so flexibel, daß das direkte Objekt auch vor dem Satzprädikat und das Satzsubjekt dahinter stehen können [= OPS-Struktur]).

Die bei Levelt *inkrementell* entstehende syntaktische Satzstruktur entspricht im Prinzip der Phrasenstrukturgrammatik von Chomsky (1965): Die Satzstruktur kann als ein sozusagen von links nach rechts generierter grammatischer Satzstrukturbaum aufgefaßt werden, dessen terminale Konstituenten die Lemmata sind.

Kempen und Hoenkamp (1987) wie auch Levelt (p. 240 u.a.O.) konzipieren den Produktionsalgorithmus, der die „message“ nach den Maßgaben der Phrasenstrukturgrammatik in eine syntaktische Oberflächenstruktur verwandelt, als ein System von einzelnen sog. *Produktionen*, das heißt von bedingten Befehlen: „Wenn x, tue

y!“ Der Kempen-Hoenkamp-Algorithmus, mit dessen Hilfe das Sprechersystem, Levelt zufolge, den Aufbau der syntaktischen Satzstruktur realisiert, enthält also eine Vielzahl von Prozeßkomponenten, von denen jede eine spezifische Tätigkeit y ausübt, sobald sie in einen spezifischen Zustand x gerät. Diese Vorstellung (vgl. auch Anderson, 1983) schließt ein, daß das Ergebnis der Tätigkeit y_i einer im Zustand x_i befindlichen Produktion P_i den Zustand x_j einer anderen Produktion P_j so ändert, daß diese die Tätigkeit y_j ausführt, usf. So entsteht eine hochkomplexe, interaktive (zeitlich parallele) Dynamik der multiplen Ausführung von bedingten Befehlen, wobei diese Dynamik jedoch streng regelgeleitet (algorithmisch) ist, ohne daß jedoch eine einzelne zentrale Entscheidungsinstanz den Gesamtprozeß steuert.

Levelts Sequenzmodell ist in *Abbildung x2* zusammengefaßt.

— hier etwa *Abbildung x2* —

Einige empirische Belege

Die Unterscheidung spezifischer Teilprozesse, die an der einzelsprachlichen Enkodierung beteiligt sind, und die Serialität der Sprachverarbeitung werden auf unterschiedliche Weise empirisch zu erhärten versucht. Dabei standen für lange Zeit Analysen spontaner Sprechfehler nebst Analysen der Art und Weise, wie Sprecher solche Fehler korrigieren, im Vordergrund (vgl. u.a. Dell, 1984; Fromkin [ed.], 1973; Levelt, 1983). Viele sprachpsychologische Experimente sind inzwischen hinzugekommen. Ebenso wie in den vorherigen Abschnitten dieses Kapitels müssen wir uns auf einige Beispiele beschränken.

Daß sprachliche Äußerungen nicht Wort für Wort, sondern in Form von grammatischen Satzteilen (Satzphrasen; Satzkonstituenten) erzeugt werden, kann man wie folgt nachweisen (Levelt, 1983): Wenn sich jemand verspricht, unterbricht er sich zumeist sofort, wenn er den Fehler bemerkt. Dann aber fährt er überwiegend nicht – diesmal richtig – genau an derjenigen Stelle fort, an der er sich unterbrochen hat. Vielmehr geht er ganz überwiegend an den *Beginn der Satzphrase* zurück, innerhalb welcher der Fehler auftrat. Jemand korrigiert sich beispielsweise oft wie folgt: „Gib mir mal den groben *hm* den großen Stift!“ Man geht hier an den Beginn der Nominalphrase („den großen Stift“) zurück. Man korrigiert sich nie: „Gib mir

mal den groben *hm* Ben Stift!“ Solche Befunde verdeutlichen die *Phrasenstrukturierung der Sprachproduktion* und damit den Einfluß syntaktischer Regeln auf die Satzbildung.

Die *Unterscheidung verschiedener Teilprozesse* ergibt sich aus Plausibilitätsüberlegungen der folgenden Art, die ebenfalls auf der Analyse von Sprechfehlern beruhen. Die Unterscheidung von drei Prozeßebenen – Erzeugung der „*message*“, Erzeugung der *Lemmata* und Erzeugung der *Wortformen* – hat unter anderem ihr Korrelat in der Unterscheidung *verschiedener Arten* von Sprechfehlern. Das folgende Beispiel stammt von Dell (1989; vgl. dazu auch Bock, 1990). (Dell vertritt nicht das soeben dargestellte, von Levelt vorgeschlagene Sequenzmodell der Sprachproduktion und bevorzugt auch eine andere Terminologie. Das kann aber für die gegenwärtige Argumentation außer Betracht bleiben.) – Jemand will sagen: „The squeaky wheel gets the grease“ (etwa: „Geschmiert wird das quietschende Rad.“) und sagt irrtümlich: „The skreaky gwease gets the weel“:

- (1) *Ebene der „message“*: Die generierte „message“ ist eine metaphorische Realisierung der Idee, daß das, was am meisten Krach macht, am meisten Aufmerksamkeit erregt.
- (2) *Ebene der Lemmata*: Die „message“ ruft im Mentalen Lexikon unter anderem die beiden Lemmata WHEEL und GREASE auf. Die beiden *Lemmata* werden jedoch den falschen „slots“ einer grammatischen Satzstruktur (Garrett, 1988) zugeordnet, so daß GREASE in die Subjektrolle und WHEEL in die Objektrolle geraten.
- (3) *Ebene der Wortformen*: Hier wird das *Phonem* /ɪ/ mit dem *Phonem* /w/ der nun unmittelbar nebeneinander angeordneten Wortformen „squeaky“ und „grease“ vertauscht: Es entsteht „skreaky gwease“.

Um diesen Sprechfehler sachgerecht analysieren zu können, muß man also *drei Ebenen* der Sprachproduktion unterscheiden.

Daß *zunächst* interne Lexikoneinträge und *dann erst* die Wortformen generiert werden, läßt sich anhand einer Klasse von Experimenten erhärten, die in großer Zahl im Zusammenhang mit Levelts seriellem Modell (1989) durchgeführt wurden. Ein bekannt gewordenes Beispiel

stammt von Schriefers, Meyer und Levelt (1990): Man zeigt den Versuchspersonen auf dem Bildschirm einfache Objekte, die so schnell wie möglich benannt werden sollen. Dieser Benennungsvorgang wird wie folgt gestört: In variablem zeitlichem Abstand zum Beginn der Objektexposition auf dem Bildschirm hören die Versuchspersonen über Kopfhörer sogenannte Distraktor-Wörter. Diese sind entweder nach ihrer *Bedeutung* (semantisch) oder nach ihrer *Wortform* (phonologisch) den Wörtern ähnlich, mit denen die Objekte zu benennen sind. Daneben werden Wörter dargeboten, die dem Zielwort weder semantisch noch phonologisch ähnlich sind. Die Distraktor-Wörter werden entweder 150 ms vor oder 150 ms nach dem Beginn der Exposition des Bildes dargeboten. Bei Verwendung der deutschen Sprache könnte zum Beispiel auf dem Bildschirm eine Maus abgebildet sein; die Versuchspersonen haben also möglichst schnell „Maus“ zu sagen. Kurz nach oder vor dem Erscheinen des Bildes hören die Versuchspersonen über Kopfhörer entweder das Wort „Ratte“ (= semantische Ähnlichkeit) oder das Wort „Haus“ (= phonologische Ähnlichkeit) oder ein semantisch und phonologisch unähnliches Wort (z.B. „Kamm“). Vorausgesetzt werden dabei die beiden folgenden gut gesicherten Sachverhalte: (i) Reaktionen auf einen Reiz S_1 können verlangsamt werden, wenn annähernd gleichzeitig ein zu S_1 *ähnlicher* Reiz S_2 exponiert wird (= Interferenzeffekt). (ii) *Ähnliche* Reize S_2 können auch die Reaktion auf S_1 beschleunigen (= Erleichterungseffekt). Welcher der beiden Effekte jeweils auftritt, hängt von relativ komplexen Bedingungen ab, die hier nicht erörtert werden müssen (vgl. dazu Glaser & Dünkelhoff, 1984).

Es stellte sich heraus, daß die *semantisch* ähnlichen Distraktorwörter eine auf die Objektbenennung verlangsamende Wirkung hatten, wenn sie 150 ms *vor* Beginn des Erscheinens des Objekts auf dem Bildschirm exponiert wurden (= semantischer Interferenzeffekt); wenn sie jedoch 150 ms *nach* der Bildexposition dargeboten wurden, blieben sie wirkungslos. Bei den *phonologisch* ähnlichen Distraktor-Wörtern verhielt es sich anders: Diese beschleunigten den Benennungsprozeß, wenn sie 150 ms *nach* dem Beginn der Bildexposition dargeboten wurden (= phonologischer Erleichterungseffekt); wenn sie aber 150 ms *vor* der Bildexposition exponiert wurden, zeigten sie keine Wirkung. Daraus schließen die Autoren, daß bei der Erzeugung der Objektbenennung *zuerst* ein lexikalisch-semantischer Teilprozeß (s. oben: Ebene 2) und *dann erst* ein phonologischer Teilprozeß (s. oben: Ebene 4) ablaufen. Die beiden Ebenen sind also im Sinne des Levelt-Modells streng unterscheidbar und in modellkonformer Weise seriell geordnet.

3.4 Zur Kontrolle der Sprachproduktion

Kontrolle nur am Ende des Sprachproduktionsprozesses?

Es wurde im Abschnitt 3.3 verdeutlicht, daß Sprechfehler und ihre Korrektur für die Theorie der Sprachproduktion eine erhebliche Rolle spielen (vgl. u.a. Garrett, 1988). Die Korrektur von Sprechfehlern der verschiedensten Art setzen die *Kontrolle* der eigenen Sprachproduktion voraus. Generell kann die Kontrolle als der Vergleich von Ist-Werten mit Soll-Werten verstanden werden, wobei Ist-Soll-Abweichungen zu Korrekturoperationen (Stelloperationen) führen. Statt von Kontrolle kann auch von *Regulation* gesprochen werden. Was diese Kontrolle bzw. Regulation des eigenen Sprechens betrifft, so vertritt zum Beispiel Levelt (1983) die Auffassung, daß der Sprecher seine Sprachproduktion nur dadurch kontrolliert, daß er das *Endergebnis* der Sprachproduktion mehrstufig rezipiert (vgl. oben **xx**) und auf dem Wege dieser Rezeption Fehler und Mängel bemerken und beheben kann. Der Autor setzt damit voraus, daß eine Überwachungsinstanz *außerhalb* des Sprachproduktionssystems existiert, die das *per Sprachrezeption* rückgemeldete Endergebnis der eigenen Sprachproduktion bewertet und gegebenenfalls Korrekturen startet (= „Editoren-Theorien“; vgl. Levelt, 1983). Eine andere Modellklasse setzt nicht nur eine externe Überwachungsinstanz voraus, vielmehr erfolgt die Regulation hier auch *innerhalb* des Sprachproduktionssystems („Konnektionistische Theorien“; vgl. Dell, 1986). Zugleich wird die Auffassung vertreten, daß Kontrollvorgänge bereits zeitlich vor der „Endfertigung“ einer Äußerung stattfinden bzw. daß auch das Ergebnis *einzelner Teilprozesse* der Sprachproduktion sofort kontrolliert und gegebenenfalls korrigiert wird (Berg, 1986; Schade, 1990; vgl. auch Herrmann & Grabowski, 1994, S. 412 ff.).

Für die zuletzt genannten Auffassungen werden unter anderem die folgenden Argumente genannt: Fehler bei der Phonemerzeugung werden generell sehr schnell erkannt und korrigiert (Schade, 1990). Diese außerordentliche Schnelligkeit der Korrektur kann kaum damit erklärt werden, daß der Sprecher erst seine „fertige“ Äußerung „mithört“, sie mehrstufig rezipiert, sie dann Phonem für Phonem prüft und sie, falls ein Phonem falsch ist, dieses „repariert“ (s. oben: „Editoren-Theorien“). Die Aufmerksamkeit des Sprechers richtet sich zudem im

allgemeinen auf den Inhalt seiner Rede. Trotzdem vermag er, obwohl die Phonemerzeugung im allgemeinen automatisch verläuft, bei weitem die meisten seiner Phonemfehler zu korrigieren. Er korrigiert seine inhaltlichen bzw. semantischen Fehler in der Regel sogar seltener als seine Phonemfehler (Nooteboom, 1980).

Regulationsebenen der Sprachproduktion

Nach dem Regulationsmodell der Sprachproduktion (Herrmann & Grabowski, 1994) wird das Sprechen als eine komplexe Stelloperation des Sprechersystems konzeptualisiert: Liegt bei einem Sprechersystem eine Ist-Soll-Differenz bestimmter Art vor, so führt das unter spezifischen Bedingungen zu einer Stelloperation, deren Ergebnis eine Sprachäußerung ist. Jemand möge wünschen, daß ein offenes Fenster geschlossen ist. Er kann dann aufstehen und das Fenster schließen, er kann seinen Kommunikationspartner anblicken, mit dem Finger schnippen und auf das Fenster zeigen, er kann dem Partner sagen: „Kannst du mal gerade das Fenster zumachen?“, usf. In allen diesen Fällen impliziert der Soll-Zustand des Sprechersystems ein geschlossenes Fenster, während der Ist-Zustand mit einem geöffneten Fenster einhergeht. Aus einem Inventar verschiedener Handlungen, zu dem auch verbale Aufforderungen gehören, wählt das System eine Stelloperation zur Ist-Soll-Angleichung aus. Unter bestimmten Bedingungen wird eine *sprachliche Aufforderungsäußerung* produziert (vgl. oben Abschnitt 3.2: Auffordern). Unter anderen Bedingungen wird das Fenster durch eine nicht-kommunikative instrumentelle Handlung geschlossen, oder es wird eine nonverbale Aufforderung (z.B. Fingerschnippen und Deuten) erzeugt. – Diese Sachlage läßt sich leicht auf jegliches Sprechen generalisieren: Das Sprechen ist eine Klasse von Stelloperationen unter anderen.

Falls Menschen sprechen, ist auch ihre Sprachproduktion vielfältig kontrolliert bzw. reguliert. Man kann die *Regulationsebenen der Sprachproduktion* wie folgt unterteilen (vgl. auch Herrmann & Grabowski, 1994, S. 347 f.):

- *Generelle Handlungsregulation*: Sprecher überwachen ihre eigenen sprachlichen Äußerungen unter personrelevanten Kriterien von der Art der diskursiven Rationalität (inhaltliche und logische Richtigkeit der Äußerung u. dgl.), des Selbstbilds und der sozialen „Gesichtswahrung“, der Wahrung der Beziehung zum Partner, der

Übereinstimmung mit generellen oder spezifischen ethisch-moralischen, aber auch ökonomischen und anderen lebenspraktischen Grundsätzen (Standards), und sie korrigieren allenfalls die eigenen Äußerungen (vgl. auch Carver & Scheier, 1982). Diese personrelevanten Kriterien werden selbstverständlich nicht nur bei der Regulation des eigenen sprachlichen Handelns verwendet, sondern sie gehören zu den Grundlagen der individuellen Handlungskontrolle überhaupt.

- *Situationsbezogene Regulation*: Die aus der eigenen Sprachproduktion resultierenden Äußerungen bewirken in der Regel in der situativen Umgebung des Sprechers Veränderungen, die von diesem registriert werden und so als neue Ist-Werte zu neuen Ist-Soll-Vergleichen und (meist) zu neuen Stelloperationen führen. So stellt sich der Sprecher beim Reden laufend auf die Reaktionen seines Gesprächspartners ein. – Da dies auch für den Gesprächspartner gilt, handelt es sich bei Sprecher und Partner um zwei gekoppelte (regulierte) Systeme, deren Stelloperationen als Störgrößen des jeweils anderen betrachtet werden können. Solche Interaktionen führen üblicherweise (im Falle „geglückter Kommunikation“) nach einigen Regelungszyklen bei beiden Systemen zu einem (vorübergehenden) Zustand der Ist-Soll-Minimierung und somit zu einem beide Systeme umfassenden (momentanen) Gleichgewichtszustand. Dieser Gleichgewichtszustand kann zum Beispiel als gelungene „Konstitution von Bedeutung im Gespräch“ oder als erreichtes Einverständnis über die Rollenzuteilung von Gesprächsteilnehmern interpretiert werden (vgl. auch Clark, 1996; Clark & Wilkes-Gibbs, 1986; Marková & Foppa [eds.], 1990).

- *Regulation mit Hilfe des Kommunikationsprotokolls* (s. auch oben Abschnitt 3.2): Der Sprecher speichert für eine kurze Zeit den Wortlaut dessen, was er und sein Partner zuvor gesagt haben (Sachs, 1967; Luther & Fenk, 1984; vgl. auch Gathercole & Baddeley, 1993). Außerdem speichert er in rekodierter Form für längere Zeit so etwas wie das „Sinnprotokoll“ dessen, was bisher gesagt wurde. In der Regel wird die aktuelle Sprachproduktion anhand dieser Protokoll-Informationen, zum Teil automatisch, bisweilen aber auch unter erheblichem Aufmerksamkeitsverbrauch, überwacht und gegebenenfalls korrigiert (vgl. dazu Horton & Keysar, 1996; Rosen & Engle, 1997; Rossnagel, 1995). So werden sprachliche Bezugnahmen auf Dinge, Ereignisse und Sachverhalte im Normalfall nicht zweckwidrig repetiert, nur weil man vergessen hat, daß man etwas schon einmal gesagt hat. Man kann sich aber während eines Redebeitrags

absichtsvoll und in geeigneter Weise wiederholt auf dieselben Redegegenstände beziehen. Oder man kann den Hauptgedankengang seiner Rede unterbrechen, ihn später wieder aufnehmen und diese Wiederaufnahme sprachlich markieren (v. Stutterheim & Klein, 1989). Diese und andere Merkmale kohärenten Sprechens erfordern die Bewertung und allenfalls die Korrektur dessen, was man sagen will, im Lichte dessen, was bereits gesagt wurde.

- *Auf das Sprecherziel bezogene Regulation:* Sprecher prüfen laufend, ob sie ihre zufolge ihres Kommunikationsziels vorgesehene gedankliche Botschaft, also die für die Sprachproduktion bereitgestellte kognitive Äußerungsbasis (s. oben 3.2), in angemessener Weise in eine Sprachäußerung umsetzen. Dabei investieren sie in variablem Ausmaß, abhängig vom Partner und von der kommunikativen Gesamtsituation, Aufmerksamkeitsressourcen (vgl. auch Rosnagel, 1995). Gegebenenfalls bringen sie Korrekturen an: „Das Fell war schwarz - nein eigentlich dunkelbraun.“ Oder: „Ihr Bekannter - genauer gesagt: ihr alter Freund ...“. Oder: „Gib mir - könntest du mir mal bitte das Salz reichen?“
- *Elementare Fehlerregulation:* Fehler können während der einzelsprachlichen Enkodierung auftreten. Man kann bei solchen elementaren Fehlern unter anderem die folgenden Fallklassen unterscheiden (vgl. auch Levelt, 1989, pp. 460 ff.):

Lexikalische Fehler: „... rechts hm links vom Rathaus ...“. Oder: „Da kam der Franz nein der Karl.“

Grammatische Fehler: „Die Kinder der ganzen Klasse ist - sind verspätet zurückgekommen.“ Oder: „...ist der ganze Kuchen aufgegessen geworden hm aufgegessen worden.“

Phonetisch-metrische Fehler: „Das ist völlig imponkatibel hm inkompatibel ...“ Oder: „... der retzte - letzte Rest ...“.

Die Regulation dieser elementaren Fehler erfolgt überwiegend automatisch. Sprecher sind sich der Erkennung und auch der Korrektur zum Beispiel von phonetisch-metrischen Fehlern meist nicht bewußt, d.h. sie verfügen dann über keine explizite deklarative

Repräsentation dieser Vorgänge (vgl. auch Schade, 1990; s. aber auch Jou & Harris, 1992).

Das Problem der Kontrolle der Sprachproduktion kann nicht auf die Erkennung und Korrektur von *elementaren Fehlern* lexikalischer, grammatischer oder phonetisch-metrischer Art reduziert werden. Doch sei abschließend noch einmal daran erinnert, daß gerade die elementaren Sprechfehler seit langem ein wesentliches Forschungsthema der Sprachproduktionspsychologie sind (Levelt, 1989, pp. 462 ff.; vgl. auch Dell, 1980; 1986; Fromkin, 1973; Levelt, 1983; Schade, 1992; Shattuck-Hufnagel, 1983). Ihre Analyse dient insbesondere dem Aufschluß über die generelle mentale „Architektur“ der Sprachproduktion und insbesondere dem Nachweis der Unterscheidung von Teilprozessen der einzelsprachlichen Enkodierung (s. oben: Abschnitt 3.3). Da das Sprechen aber eine Art des menschlichen Handelns ist, kann man die Sprechfehler auch als *Handlungsfehler* verstehen, wie andere (nonverbale) Handlungsfehler auch (vgl. Clark, 1996, pp. 284 ff.). Dabei mag es sich um spontane Patzer bei der Ausführung von Handlungen handeln oder um wissens- oder regelbasierte Fehler, die durch situationsspezifisch unzureichende Pläne und Vornahmen entstehen (Rasmussen, 1982). Sprechfehler werden von der Sprachpsychologie bis heute fast nie als Handlungsfehler interpretiert. In anderen Forschungstraditionen der Psychologie haben Analysen von Handlungsfehlern gleichwohl erhebliche Aufschlüsse über die Struktur und über die Funktionen des menschlichen Handelns erbracht (Reason, 1990).

Literaturverzeichnis

- Anderson, J. R. (1983). *The Architecture of Cognition*. Cambridge: Harvard University Press.
- Berg, T. (1986). The problems of language control: Editing, monitoring, and feedback. *Psychological Research*, 48, 133-144.
- Berry, D. C. & Broadbent, D. E. (1984). On the relationship between task performance and associated verbalizable knowledge. *Quarterly Journal of Experimental Psychology*, 36 A, 209-231.
- Bereiter, C. & Scardamalia, M. (1987). *The Psychology of Written Composition*. London:

Erlbaum.

Biber, D. (1995). *Dimensions of register variation: a cross linguistic comparison*. Cambridge: Cambridge University Press.

Blum-Kulka, S. (1985). The multifunctionality of discourse markers: The case of requests. *Theoretical Linguistics*, 12 (213), 213-229.

Blum-Kulka, S., House, J. & Kasper, G. (Hrsg.) (1989). *Cross-Cultural Pragmatics: Requests and Apologies* (Advances in Discourse Processes, Vol. 31). Norwood: Ablex.

Bock, J. K. (1982). Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. *Psychological Review*, 89, 1-47.

Bock, J. K. (1990). Structure in language: Creating form in talk. *American Psychologist*, 45, 1221-1236.

Bock, J. K. & Levelt, W. J. M. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (ed.), *Handbook of psycholinguistics* (pp. 945-984). San Diego, CA: Academic Press.

Bosshardt, H.-G. (1999). Effects of concurrent mental calculation on stuttering, inhalation, and speech timing. *Journal of Fluency Disorders*, 24, 43-72.

Braun, K.-F. (1985). *Der Einfluß visueller Informationen auf die Sprachverarbeitung von Kindern*. Unveröffentlichte Dissertation, Universität Saarbrücken.

Bresnan, J. (1982). *The Mental Representation of Grammatical Relations*. Cambridge: MIT Press.

Brewer, W. F. (1987). Schemas versus mental models in human memory. In P. Morris (ed.), *Modelling cognition* (pp. 187-197). New York: Wiley.

Buhl, H. M., Katz, S., Schweizer, K. & Herrmann, Th. (2000). Einflüsse des Wissenserwerbs auf die Linearisierung beim Sprechen über räumliche Anordnungen. *Zeitschrift für Experimentelle Psychologie*, 47 (1), 17-33.

Caramazza, A. (1997). *Access of phonological and orthographic lexical forms: evidence from*

- dissociations in reading and spelling*. Hove: Psychology Press.
- Carroll, M. (1993). Deictic and Intrinsic Orientation in Spatial Descriptions: A Comparison Between English and German. In J. Altarriba (Hrsg.), *Cognition and Culture: A Cross-Cultural Approach to Cognitive Psychology*. Amsterdam: Elsevier.
- Carver, C. S. & Scheier, M. F. (1982). Control theory: A useful conceptual framework for personality-social, clinical, and health psychology. *Psychological Bulletin*, 92, 111-135.
- Chafe, W. L. (1976). *Bedeutung und Sprachstruktur*. München: Hueber.
- Chomsky, N. (1965). *Aspects of a theory of syntax*. Cambridge, MA: MIT Press.
- Clark, H. H. (1996). *Using language*. Cambridge, MA: Cambridge University Press.
- Clark, H. H. & Haviland, S. E. (1977). Comprehension and the Given-New Contract. In R. O. Freedle (Hrsg.), *Discourse Production and Comprehension*. (Discourse Processes: Advances in Research and Theory, Vol. 1) (S. 1-40). Norwood: Ablex.
- Clark, H. H. & Schunk, D. H. (1980). Polite Responses to Polite Requests. *Cognition*, 8, 111-143.
- Clark, H. H. & Wilkes-Gibbs, D. (1986). Referring as a Collaborative Process. *Cognition*, 22, 1-39.
- Crystal, D. (Hrsg.) (1995). *Die Cambridge-Enzyklopädie der Sprache*. Zürich: Verlag Neue Zürcher Zeitung.
- Dell, G. S. (1984). Representation of serial order in speech: Evidence from the repeated phoneme effect in speech errors. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10 (2), 222-233.
- Dell, G. S. (1986). A spreading-activation theory of retrieval in sentence production. *Psychological Review*, 93, 283-321.
- Dell, G. S. (1989). *Language production and connectionist models of the mind*. Colloquium

presented to the Department of Psychology, Michigan State University, East Lansing, MI.

- Dell, G. S. Burger, L. K. & Svec, W. R. (1997). Language production and serial order: A functional analysis and a model. *Psychological Review*, 104 (1), 123-147.
- Dell, G. S. & Newman, J. E. (1980). Detecting phonemes in fluent speech. *Journal of Verbal Learning and Verbal Behavior*, 19 (5), 608-623.
- Denis, M. (1996b). *The description of routes: a cognitive approach to the production of spatial discourse* (Notes et documents LIMSI, Nr. 96-18). Orsay: Centre National de la Recherche Scientifique.
- DeSmedt, K. & Kempen, G. (1991). Segment grammar: A formalism for incremental sentence generation. In C. L. Paris, W. R. Swartout & W. C. Mann (Eds.), *Natural language generation in artificial intelligence and computational linguistics* (pp. 329-349). Boston, MA: Kluwer.
- Dörner, D. (1987). *Problemlösen als Informationsverarbeitung* (3. Aufl.). Stuttgart: Kohlhammer.
- Dörner, D. (1999). *Bauplan für eine Seele*. Reinbek bei Hamburg: Rowohlt.
- Dutke, S. (1996). Generic and generative knowledge: Memory schemata in the construction of mental models. In W. Battman & S. Dutke (eds.), *Processes of the molar regulation of behavior* (pp. 35-52). Lengerich: Pabst Sciences Publishers.
- Edmondson, W. & House, J. (1981). *Let's talk and talk about it: a pedagogic interactional grammar of English*. München: Urban und Schwarzenberg.
- Ehrich, V. (1989). Die temporale Festlegung lokaler Referenz. In Ch. Habel, M. Herweg & K. Rehkämper (Hrsg.), *Raumkonzepte in Verstehensprozessen. Interdisziplinäre Beiträge zur Sprache und Raum* (Linguistische Arbeiten 233) (S. 1-16). Tübingen: Niemeyer.
- Eikmeyer, H.-J., Schade, U. & Kupietz, M. (1995). Ein konnektionistisches Modell für die Produktion von Objektbenennungen. *Kognitionswissenschaft*, 4, 108-117.

- Engelbert, H. M. (1992). *Die Aktualgenese von Raummodellen und ihr Einfluß auf die Linearisierung*. Unveröff. Diplomarbeit. Universität Mannheim: Lehrstuhl Psychologie III.
- Engelkamp, J. & Zimmer, H. D. (1981). Clause relations and picture viewing. *Archiv für Psychologie*, 133, 129-138.
- Engelkamp, J. & Zimmer, H. D. (1983). *Dynamic Aspects of Language Processing. Focus and Presupposition*. Berlin: Springer.
- Fodor, J. A. (1983). *The modularity of mind*. Cambridge: MIT Press.
- Fromkin, V. (1973). *Speech errors as linguistic evidence*. The Hague: Mouton.
- Gainotti, G., Carlomagno, S., Craca, A. & Silveri, M. C. (1986). Disorders of classificatory activities in aphasia. *Brain and Language*, 28, 181-195.
- Gardenne, V. (1984). *Theorie und Erfahrung in der psychologischen Forschung*. Tübingen: Mohr.
- Garrett, M. F. (1988). Processes in language production. In F. J. Newmeyer (ed.), *Language: Psychological and biological aspects* (pp. 69-97). Cambridge, MA: Cambridge University Press.
- Gathercole, S. & Baddeley, A. D. (1993). *Working memory and language*. Hove: Erlbaum.
- Glaser, W. R. & Dünkelhoff, F. J. (1984). The Time Course of Picture-Word Interference. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 640-654.
- Gordon, D. & Lakoff, G. (1971). *Conversational postulates*. Papers from the Seventh Regional Meeting (pp. 63-84). Chicago: Chicago Linguistic Society.
- Goschke, Th. & Koppelberg, D. (1990). Connectionist Representation, Semantic Compositionality, and the Instability of Concept Structure. *Psychological Research*, 52, 253-270.
- Grabowski, J. (1991). *Der propositionale Ansatz der Textverständlichkeit: Kohärenz, Interessantheit und Behalten*. Münster: Aschendorff.

- Grabowski-Gellert, J. & Winterhoff-Spurk, P. (1988). Your Smile Is My Command: Interaction Between Verbal and Nonverbal Components of Requesting Specific to Situational Characteristics. *Journal of Language and Social Psychology*, 7, 229-242.
- Graesser, A. C., Gordon, S. G. & Sawyer, J. D. (1979). Recognition Memory of Typical and Atypical Actions in Scripted Activities: Tests of a Script Pointer Tag Hypothesis. *Journal of Verbal Learning and Verbal Behavior*, 18, 319-332.
- Grafman, J. (1995). Similarities and distinctions among current models of prefrontal cortical functions. In J. Grafman, K. J. Holyoak & F. Boller (eds.), *Annals of the New York Academy of Science*, Vol. 769, 337-368.
- Gülich, E. (1980). Konventionelle Muster und kommunikative Funktionen von Alltagserzählungen. In K. Ehlich (Hrsg.), *Erzählen im Alltag* (S. 335-384). Frankfurt/M.: Suhrkamp.
- Güsgen, H.-W. & Hertzberg, J. (1992). *A perspective of constraint-based reasoning: An introductory tutorial*. Berlin: Springer.
- Helfrich, H. (im Druck). Kulturelle und konventionale Bedingungen der Sprachproduktion. In Th. Herrmann & J. Grabowski (Hrsg.), *Psychologie der Sprachproduktion* (Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie III: Sprache, Bd. 1). Göttingen: Verlag für Psychologie.
- Helm, G. (1991). *Symbolische und konnektionistische Modelle der menschlichen Informationsverarbeitung: eine kritische Gegenüberstellung*. Heidelberg: Springer.
- Herrmann, Th. (1983). *Speech and situation: A psychological conception of situated speaking*. Berlin: Springer.
- Herrmann, Th. (1994). Forschungsprogramme. In Th. Herrmann & W. Tack (Hrsg.), *Enzyklopädie der Psychologie* (Themenbereich B, Serie 1, Band 1: Methodologische Grundlagen der Psychologie) (S. 251-294). Göttingen: Verlag für Psychologie, Dr. C. J. Hogrefe.

- Herrmann, Th. (1995). *Allgemeine Sprachpsychologie. Grundlagen und Probleme* (2. Auflage). Weinheim: Beltz.
- Herrmann, Th. (im Druck). Theorien und Modelle der Sprachproduktion. In W. Deutsch, Th. Herrmann & G. Rickheit (Hrsg.), *Handbuch der Psycholinguistik* (Handbücher der Sprach- und Kommunikationswissenschaft). Berlin – New York: Walter de Gruyter.
- Herrmann, Th. & Deutsch, W. (1976). *Psychologie der Objektbenennung* (Studien zur Sprachpsychologie: 5). Bern: Huber.
- Herrmann, Th. & Grabowski, J. (1994). *Sprechen. Psychologie der Sprachproduktion*. Heidelberg; Berlin; Oxford: Spektrum, Akademischer Verlag.
- Herrmann, Th. & Graf, R. (1999). Conceptual and semantic knowledge: A psychological perspective. In R. Klabunde & Ch. von Stutterheim (eds.), *Representations and Processes in Language Production* (pp. 17-41). Wiesbaden: Deutscher Universitäts-Verlag.
- Herrmann, Th., Kilian, E., Dittrich, S. & Dreyer, P. (1992). Was- und Wie-Schemata beim Erzählen. In H. P. Krings & G. Antos (Hrsg.), *Textproduktion. Neue Wege der Forschung* (Fokus Bd. 7) (S. 147-158). Trier: WVT Wissenschaftlicher Verlag Trier.
- Herrmann, Th. & Schweizer, K. (1998). *Sprechen über Raum. Lokalisieren und seine kognitiven Grundlagen*. Bern: Huber.
- Hoffmann, L. (1991). Vom Ereignis zum Fall. Sprachliche Muster zur Darstellung und Überprüfung von Sachverhalten vor Gericht. In J. Schönert (Hrsg.), *Erzählte Kriminalität. Zur Typologie und Fuktion von narrativen Darstellungen in Strafrechtspflege, Publizistik und Literatur zwischen 1770 und 1920* (S. 87-113). Tübingen: Narr.
- Höhle, B. (1995). *Aphasie und Sprachproduktion: Sprachstörungen bei Broca- und Wernicke-Aphasikern*. Opladen: Westdeutscher Verlag.
- Hörmann, H. (1976). *Meinen und Verstehen. Grundzüge einer psychologischen Semantik*.

Frankfurt/M.: Suhrkamp.

Hornby, P. A. (1974). Surface Structure and Presupposition. *Journal of Verbal Learning and Verbal Behavior*, 13, 530-583.

Horton, W. S. & Keysar, B. (1996). When do speakers take into account common ground? *Cognition*, 59, 91-117.

House, J. (1996). Contrastive discourse analysis and misunderstanding. In M. Hellinger & U. Ammon (eds.), *Contrastive sociolinguistics* (pp. 345-361). Berlin: Mouton de Gruyter.

Huber, W. (1981). Aphasien. Klinisch-neurolinguistische Beschreibung und Erklärungsversuche. *Studium Linguistik, Heft 11*, 1-21.

Jackendoff, R. S. (1983). *Semantics and Cognition* (Current Studies in Linguistics Series, 8). Cambridge: MIT Press.

Johnson-Laird, P. N. (1980). *Thinking: readings in cognitive science*. Cambridge: Cambridge University Press.

Jordan, M. I. (1986). *Serial Order: A Parallel Distributed Processing Approach* (Technical Report, NO. 8604). San Diego: University of California, Institute for Cognitive Science.

Jou, J. & Harris, R. J. (1992). The effect of divided attention on speech production. *Bulletin of the Psychonomic Society*, 30, 301-304.

Kasper, G. (1990). Linguistic politeness: Current research issues. *Journal of Pragmatics*, 14 (2), 193-218.

Kempen, G. & Hoenkamp, E. (1987). An Incremental Procedural Grammar for Sentence Formulation. *Cognitive Science*, 11, 201-258.

Kintsch, W. (1974). *The Representation of Meaning in Memory*. Hilldale: Erlbaum.

Klabunde, R. & Stutterheim, C. von (Hrsg.) (1999). *Representations and processes in language production*. Wiesbaden: Deutscher Universitäts-Verlag.

Klein, W. (1982). Einige Bemerkungen zur Fragemelodie. *Deutsche Sprache*, 10, 289-310.

- Klein, W. (1984). Bühler Ellipse. In C. F. Graumann & Th. Herrmann (Hrsg.), *Karls Bühlers Axiomatik* (S. 117-143). Frankfurt/M.: Klostermann.
- Klix, F. (1992). *Die Natur der Verstandes*. Göttingen: Hogrefe.
- Koch, W. (1992). Automatische Generierung von Kochrezepten. In H. P. Krings & G. Antos (Hrsg.), *Textproduktion. Neue Wege der Forschung* (Fokus Bd. 7) (S. 311-338). Trier: WVT Wissenschaftlicher Verlag Trier.
- Laucht, M. (1979). Untersuchungen zur sprachlichen Form des Aufforderns. In W. Tack (Hrsg.), *Bericht über den 31. Kongreß der Deutschen Gesellschaft für Psychologie* (S. 89-91). Göttingen: Hogrefe.
- Levelt, W. J. M. (1981). The Speaker's Linearization Problem. In D. E. Broadbent, J. Lyons & S. Longuet-Higgins (Hrsg.), *Psychological Mechanisms of Language* (Philosophical Transactions of the Royal Society London, B 295) (S. 305-315).
- Levelt, W. J. M. (1982). Linearization in Describing Spatial Networks. In S. Peters & E. Saarinen (Hrsg.), *Processes, Beliefs, and Questions. Essays on Formal Semantics of Natural Language and Natural Language Processing* (S. 199-220). Dordrecht: Reidel.
- Levelt, W. J. M. (1983). Monitoring and Self Repairs in Speech. *Cognition*, 14, 41-104.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge/London: A Bradford Book/MIT Press.
- Levelt, W. J. M. (1992). Accessing Words in Speech Production: Stages, Processes and Representations. *Cognition*, 42, 1-22.
- Linde, C. & Labov, W. (1985). Die Erforschung von Sprache und Denken anhand von Raumkonfigurationen. In H. Schweizer (Hrsg.), *Sprache und Raum: Psychologische und linguistische Aspekte der Aneignung und Verarbeitung von Räumlichkeit. Ein Arbeitsbuch für das Lehren von Forschung* (S. 44-65). Stuttgart: Metzler. (Original erschienen 1975: Spatial Networks as a Site for the Study of Language and Thought. *Language*, 51, 924-939.)

- Luther, P. & Fenk, A. (1984). Wird der Wortlaut von Sätzen zwangsläufig schneller vergessen als ihr Inhalt? *Zeitschrift für Experimentelle und Angewandte Psychologie*, 31, 101-123.
- Lyons, J. (1989). *Einführung in die moderne Linguistik* (7., unveränderte Aufl.). München: Beck. (Original erschienen 1968: *Introduction to Theoretical Linguistics*. Cambridge: Cambridge University Press.)
- Malone, M. (1995). How to do things with friends: Altercasting and Recipient Design. *Research on Language & Social Interaction*, 28 (2), 147-170.
- Mandler, J. M. & Johnson, N. S. (1977). Remembrance of Things Parsed: Story Structure and Recall. *Cognitive Psychology*, 9, 111-151.
- Marková, I. & Foppa, K. (Eds.). (1990). *The dynamics of dialogue*. Hemel Hempstead, Hertfordshire: Harvester Wheatsheaf.
- Meyer, A. S. & Schriefers, H. (im Druck). Die lexikalische Kodierung. In Th. Herrmann & J. Grabowski (Hrsg.), *Enzyklopädie der Psychologie (Bereich C, Serie III, Band 1: Sprachproduktion)*. Göttingen: Verlag für Psychologie, Dr. C. J. Hogrefe.
- Meyer, J. R. (1992). Fluency in the production of requests: Effects of degree of imposition, schematicity and instruction set. *Journal of Language and Social Psychology*, 11 (4), 233-251.
- Miller, G. A. (1993). *Wörter – Streifzüge durch die Psycholinguistik*. Heidelberg: Spektrum Akademischer Verlag. (Original erschienen 1991: *The Science of Words*. New York: Scientific American Library.)
- Neumann, O. (1992). Theorien der Aufmerksamkeit: Von Metaphern zu Mechanismen. *Psychologische Rundschau*, 43, 83-101.
- Nooteboom, S. G. (1980). Speaking and unspeaking: Detection and correction of phonological and lexical errors in spontaneous speech. In V.A. Fromkin (Ed.), *Errors in linguistic performance* (pp. 87-95). New York: Academic Press.

- Norman, D. A. & Shallice, T. (1986). *Attention to action: Willed and automatic control of behavior*. In R. J. Davidson, G. E. Schwartz & D. Shapiro (Eds.), *Consciousness and self-regulation. Advances in research and theory (Vol. 4, pp. 1-18)*. New York: Plenum Press.
- Oesterreich, R. (1981). *Handlungsregulation und Kontrolle*. München: Urban und Schwarzenberg.
- Pelz, H. (1992). *Linguistik für Anfänger* (10. Aufl.). Hamburg: Hoffmann und Campe.
- Perrig, W. J., Wippig, W. & Perrig-Ciello, P. (1993). *Unbewußte Informationsverarbeitung*. Bern; Göttingen: Huber.
- Posner, M. I. & Snyder, C. R. R. (1975). Attention and cognitive control. In R. L. Solso (Ed.), *Information processing and cognition: The Loyola Symposium*. Hillsdale: Erlbaum.
- Power, M. J. (1985). Sentence production and working memory. *Quarterly Journal of Experimental Psychology*, 37A, 367-386.
- Rasmussen, O. E. (1982). Reflection and reflectivity in the concept of activity. *Psyke and Logos*, 2, 311-322.
- Reason, J. (1990). *Human Error*. Cambridge: Cambridge University Press.
- Reber, A. S. (1989). Implicit learning and tacit knowledge. *Journal of Experimental Psychology: General*, 118, 219-235.
- Reimer, (1989). *FRM: Ein Frame-Repräsentationsmodell und seine formale Semantik: zur Integration von Datenbank- und Wissensrepräsentationsansätzen*. Heidelberg: Springer.
- Rickheit, G. & Strohner, H. (1993). *Grundlagen der kognitiven Sprachverarbeitung: Modelle, Methoden, Ergebnisse*. Tübingen: Francke.
- Robin, F. & Denis, M. (1991). Description of perceived or imagined spatial networks. In R. H. Logie & M. Denis (Eds.), *Mental images in human cognition* (pp. 141-152). Amsterdam: North-Holland.

- Roelofs, A. (1997). The WEAVER Model of word-form encoding in speech production. *Cognition*, 64 (3), 249-284.
- Rosen, V. M. & Engle, R. W. (1997). The role of working memory capacity in retrieval. *Journal of Experimental Psychology: General*, 126, 211-227.
- Rosnagel, C. (1995). Kognitive Belastung und Hörerorientierung beim monologischen Instruieren. *Zeitschrift für Experimentelle Psychologie*, XLII, 94-110.
- Rumelhart, D. E. & Ortony, A. (1977). The Representation of Knowledge in Memory. In R. C. Anderson, R. J. Spiro & W. E. Montague (Hrsg.), *Schooling and the Acquisition of Knowledge* (S. 99-135). Hillsdale: Erlbaum.
- Rummer, R. (1996). *Kognitive Beanspruchung beim Sprechen*. Weinheim: Beltz.
- Rummer, R., Grabowski, J., Hauschildt, A. & Vorweg, C. (1993). *Reden über Ereignisse: Der Einfluß von Sprecherzielen, sozialer Nähe und Institutionalisiertheitsgrad auf Sprachproduktionsprozesse* (Arbeiten aus dem Sonderforschungsbereich 245 "Sprache und Situation" Heidelberg/Mannheim, Bericht Nr. 56). Universität Mannheim: Lehrstuhl Psychologie III.
- Rummer, R., Mohr, G. & Zimmer, H. (1998). Leistungsbegrenzungen im verbalen Arbeitsgedächtnis: Argumente gegen energetische Ressourcenkonzepte. *Kognitionswissenschaft*, 7, 134-140.
- Sachs, J. (1967). Recognition memory for syntactic and semantic aspects of connected discourse. *Perception and Psychophysics*, 2, 437-442.
- Sacks, H., Schegloff, E. & Jefferson, G. (1974). A Simplest Systematics for the Organization of Turn-Taking for Conversation. *Language*, 50, 696-735.
- Schade, U. (1990). Kohärenz und Monitor in konnektionistischen Sprachproduktionsmodellen. In G. Dorffner (Hrsg.), *Konnektionismus in Artificial Intelligence und Kognitionsforschung. 6. Österreichische Artificial Intelligence-Tagung (KONNAI), Salzburg, 18.-21. Sept. 1990: Proceedings* (S. 18-27). Berlin:

Springer.

Schade, U. (1992). *Konnektionismus. Zur Modellierung der Sprachproduktion*. Opladen: Westdeutscher Verlag.

Schade, U. (im Druck). Repräsentationale Grundlagen der Sprachproduktion. In Th. Herrmann & J. Grabowski (Hrsg.), *Psychologie der Sprachproduktion* (Enzyklopädie der Psychologie, Themenbereich C: Theorie und Forschung, Serie III: Sprache, Bd. 1). Göttingen: Verlag für Psychologie.

Schank, R. C. & Abelson, R. P. (1977). *Scripts, Plans, Goals, and Understanding*. Hillsdale: Erlbaum.

Schlesinger, I. M. (1977). *Production and Comprehension of Utterances*. Hillsdale: Erlbaum.

Schriefers, H., Meyer, A. S. & Levelt, W. J. M. (1990). Exploring the time course of lexical access in language production. *Journal of Memory and Language*, 29, 86-102.

Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In P. F. MacNeilage (Ed.), *The production of speech* (pp. 109-136). New York: Springer.

Stein, G. L. & Glenn, C. G. (1979). An Analysis of Story Comprehension in Elementary School Children. In R. O. Freedle (Hrsg.), *New Directions in Discourse Processing* (Advances in Discourse Processes, Vol. 2) (S. 53-120). Norwood: Ablex.

Stutterheim, C. von & Klein, W. (1989). Textstructure and referential movement. In R. Dietrich & C. F. Graumann (Hrsg.), *Language processing in social context* (S. 39-76). Amsterdam: North-Holland.

Tisdale, T. (1998). *Selbstreflexion, Bewußtsein und Handlungsregulation*. Weinheim: Beltz.

Tsui, A. B. M. (1991). Sequencing rules and coherence in discourse. *Journal of Pragmatics*, 15, 111-129.

Ungerer, D., Liebertz, P. & Morgenroth, U. (1998). *Telefonieren im Fahrzeug – Fahrerbelastung und Fahrfehler*. Universität Bremen. Sensomotorik-Labor/ADAC-

Zentrale München.

Vukovich, A. (1999). *Kovariation und Ebenmaß: Grundformen der Symmetrie in natürlichsprachlicher Rede*. Regensburg: Roderer.

Werani, A. (1999). Sprach- und Kognitionsforschung in der Aphasologie – Ein Überblick. *Sprache & Kognition*, 18, 20-29.

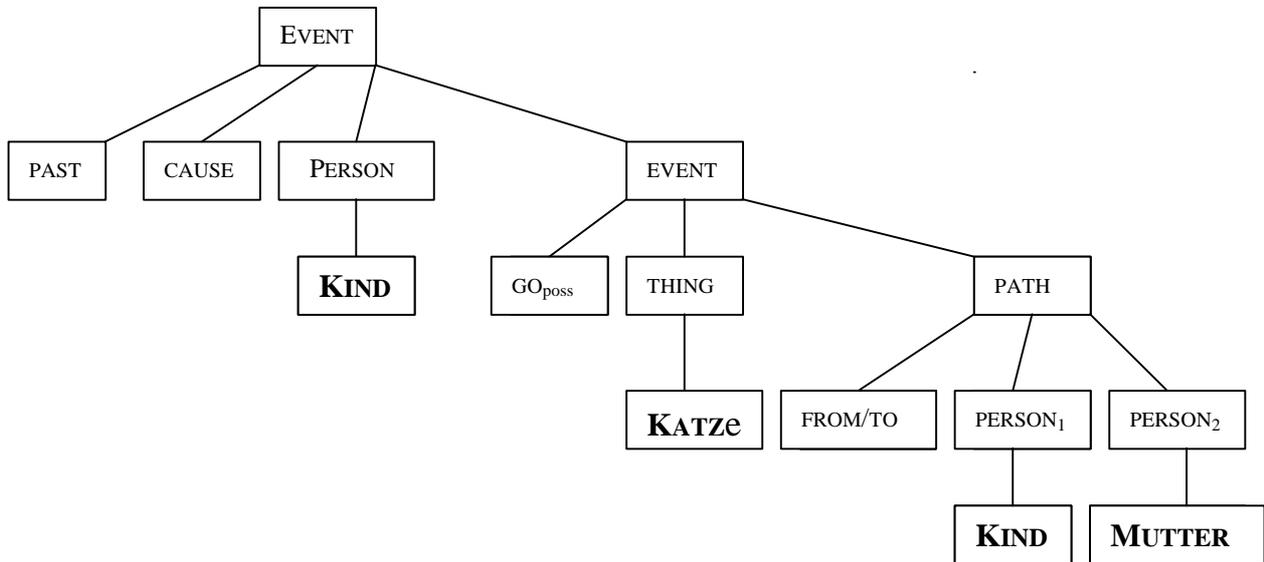


Abb. X1: Propositionale Darstellung einer kognitiven Äußerungsbasis nach Levelt (1989, p. 164; s.Text).

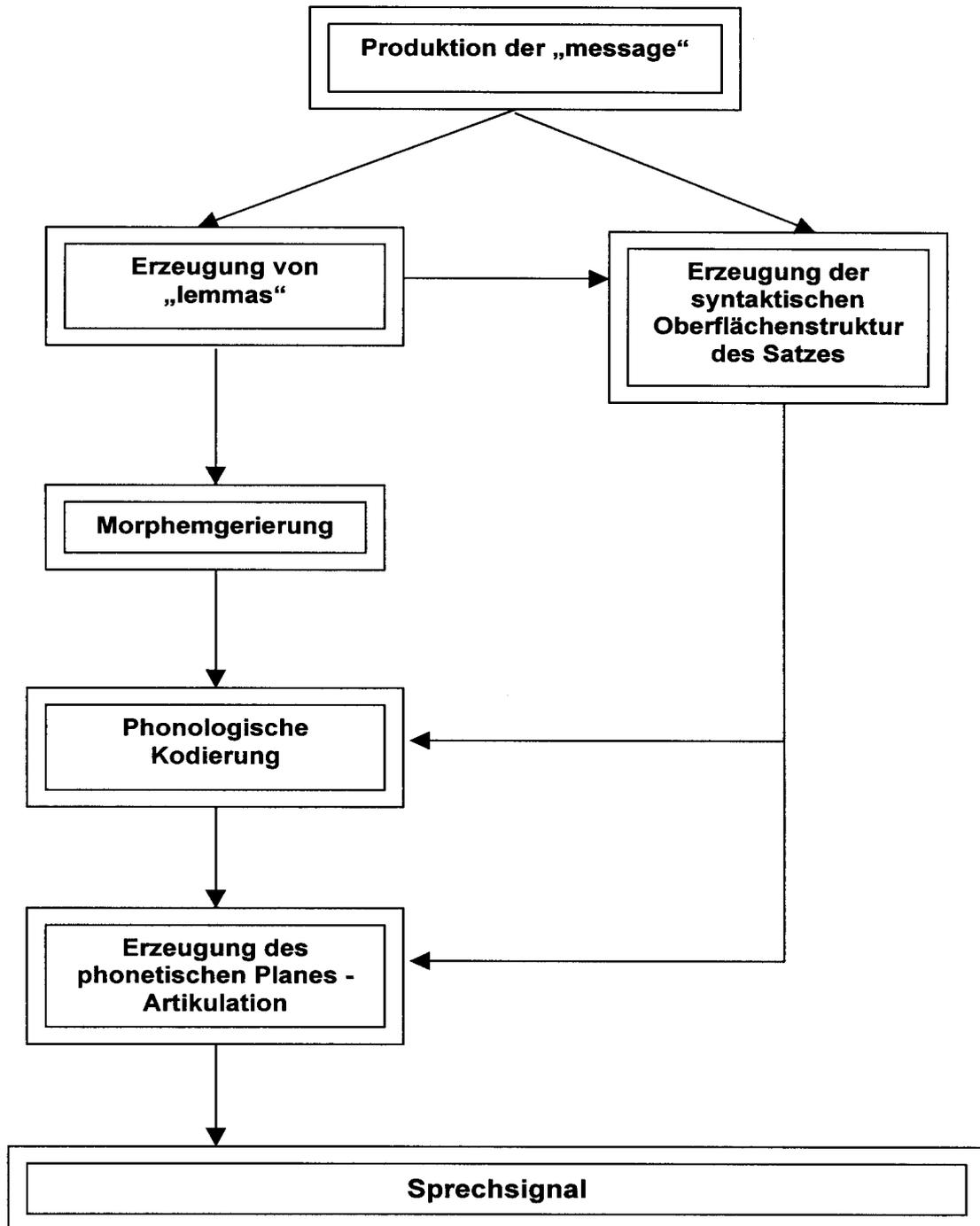


Abb. X2: Levelts Modell serieller Teilprozesse der Sprachproduktion (s.Text).