Linguistische
Arbeiten 218

Herausgegeben von Hans Altmann, Herbert E. Brekle, Hans Jürgen Heringer, Christian Rohrer, Heinz Vater und Otmar Werner

Jürgen Rolshoven

Eine selbstlernende generativ-phonologische Grammatik

Max Niemeyer Verlag
Tübingen 1989

CIP-Titelaufnahme der Deutschen Bibliothek

Rolshoven, Jürgen : Eine selbstlernende generativ-phonologische Grammatik / Jürgen Rolshoven. – Tübingen : Niemeyer, 1989.
(Linguistische Arbeiten ; 218)
NE: GT

ISBN 3-484-30218-6 ISSN 0344-6727

© Max Niemeyer Verlag Tübingen 1989
Alle Rechte vorbehalten. Ohne Genehmigung des Verlages ist es nicht gestattet, dieses Buch oder Teile daraus photomechanisch zu vervielfältigen.
Printed in Germany. Druck: Weihert-Druck GmbH, Darmstadt.

Inhaltsverzeichnis

Vorwort VIII

1. Einleitung 1
1.1. Ziel der Arbeit 1
1.2. Zweck der Arbeit 1
1.3. Aufbau der Arbeit 3

2. Sprache 6

3. Das Steuerwerk 21
3.1. Globalfunktion 21
3.1.1. Format der parole (Eingabe) 22
3.1.2. Regelwerk (Generative Phonologie) 22
3.1.3. Kommunikativer Erfolg 23
3.1.4. Schrittweise Problemlösung 23
3.2. Lexikon und Prozeßmenge 32
3.2.1. Globalfunktion und Gliederung 32
3.2.2. Lernen 35
3.2.3. Parole einlesen und Lexikongraphenmenge erzeugen 37
3.2.4. Matrixverarbeitung 39
3.2.5. Erzeuge Hypothesen für Lexikongraphen 42
3.2.5.1. Überblick über die in einer Schleife eingebetteten Prozeduren 42
3.2.5.2. Erzeugung von Hypothesen für Lexikongraphen 47
3.2.6. Prozeßmenge und Hypothesenbewertungen 62
3.2.6.1. Überblick 62
3.2.6.2. Hypothesen für die Prozeßmenge 64
3.2.6.3. Hypothesenbewertung 70
3.2.7. Endbewertung 71
3.3. Prozeßmenge und generativ-phonologische Regeln 75
3.3.1. Übersicht 75
3.3.1.1. Regelerzeugung 75
3.3.1.2. Lernen 97
3.3.1.3. Die Erzeugung der Baumgraphenmenge 100
3.3.2. Hypothesenerzeugung 104
3.3.2.1. Übersicht 104
3.3.2.2. Klassifikation 111
3.3.2.3. Prozeßtypbestimmung 116
3.3.2.4. Strukturveränderung 117
3.3.2.5. Die Erzeugung der Strukturbeschreibung 118
3.3.2.5.1. Oberflächenkonstruktion und -rekonstruktion 118
3.3.2.5.2. Tiefenstrukturkonstruktion 121
3.3.2.5.3. Exklusives Oder 123
3.3.2.5.4. Auswertung 124
3.3.2.5.5. Hypothesenverschmelzung 129
3.3.2.6. Regelhypothesen- und Regelverschmelzung 135
3.3.3. Hypothesenauswahl 141
3.3.3.1. Hypothesenauswahl bei konkurrierenden (parallel erzeugten) Hypothesen 141
3.3.3.2. Hypothesenauswahl bei sich beeinflussenden Hypothesen (Hypothesenkollision) 145
3.3.4. Übertragen in das Regelwerk 148
3.4. Akzentuierungsregeln und Ausgabe der Ergebnisse 150

4. Ergebnisse 159
4.0. Transkriptionssystem 159
4.1. Generative Phonologie des Surselvischen 160
4.1.1. Einführung in das Surselvische 160
4.1.1.1. Das Surselvische in historischer Sicht 160
4.1.1.1.1. Äußere Sprachgeschichte 160
4.1.1.1.2. Innere Sprachgeschichte 162
4.1.1.2. Die aktuelle sprachliche Situation des Surselvischen 165
4.1.2. Textkorpus, Lexikon und generativ-phonologische Grammatik des Surselvischen 166
4.1.2.1. Das surselvische Textkorpus 166
4.1.2.2. Das Lexikon 169
4.1.2.3. Die generative Phonologie des Surselvischen 175
4.1.2.3.1. Die Akzentuierungsregeln 175
4.1.2.3.2. Die Prozeßregeln 176
4.2. Generative Dialektologie des Meglenorumänischen und des Aromunischen 189
4.2.1. Einführung in das Meglenorumänische und das Aromunische 189
4.2.1.0. Vorbemerkungen 189
4.2.1.1. Das Meglenorumänische und das Aromunische in historischer Sicht 195
4.2.1.2. Die sprachliche Situation des heutigen Meglenorumänischen und Aromunischen 200
4.2.1.3. Die Wahl der Aufnahmepunkte für die generative Dialektologie des Meglenorumänischen und des Aromunischen 205
4.2.1.4. Textkorpora und Sprachaufnahmen 206
4.2.2. Das Meglenorumänische 208
4.2.2.1. Der Aufnahmeort 208
4.2.2.2. Das Textkorpus 209
4.2.2.3. Das Lexikon des Meglenorumänischen 211
4.2.2.4. Die generativ-phonologische Grammatik des Meglenorumänischen 216
4.2.3. Die aromunischen Dialekte 225
4.2.3.1. Das Aromunische des Βέρμιον 225
4.2.3.1.1. Der Aufnahmeort 225
4.2.3.1.2. Das Textkorpus 226
4.2.3.1.3. Das Lexikon 228
4.2.3.1.4. Die generativ-phonologische Grammatik 232
4.2.3.2. Das Olympovlachische 242
4.2.3.2.1. Der Aufnahmeort 242
4.2.3.2.2. Das Textkorpus 242
4.2.3.2.3. Das Lexikon 244
4.2.3.2.4. Die generativ-phonologische Grammatik 249
4.2.3.3. Das Pindnerische 256
4.2.3.3.1. Der Aufnahmeort 256
4.2.3.3.2. Das Textkorpus 256
4.2.3.3.3. Das Lexikon 258
4.2.3.3.4. Die generativ-phonologische Grammatik 263
4.3. Generative Dialektologie 272
4.3.1. Unterschiede in den zugrundeliegenden Repräsentationen 272
4.3.2. Unterschiede in den Regeln 274
4.3.3 Unterschiede im Lexikon und in der Syntax 276
4.4. Sprachkontakt 278

5. Kritik und Ausblick 280
5.1. Kritik 280
5.2. Ausblick 281

6. Literatur 284

Vorwort

Die vorliegende Arbeit ist eine verkürzte und überarbeitete Fassung meiner im Sommersemester 1984 von der Philosophischen Fakultät der Universität zu Köln angenommenen Habilitationsschrift.

Es ist mir eine angenehme Pflicht, mich derer zu erinnern, die mich bei ihrer Anfertigung unterstützt haben. Es sind dies meine akademischen Lehrer H.D. Bork, A. Greive, H. Vater, der mir wiederum die Veröffentlichung in den Linguistischen Arbeiten ermöglichte, P.O. Samuelsdorff; es sind dies W. Dahmen, E. Kurschildgen, A. Ritter, B. Dresel, H. Kunzelmann und M. Rolshoven. W. Rolshoven erstellte das Typoskript auf einer Textverarbeitungsanlage; M. Jahn half mir beim Druck selbstgenerierter Zeichensätze; die Mitarbeiter des Rechenzentrums der Universität zu Köln berieten mich bei Programmierproblemen. Die Deutsche Forschungsgemeinschaft gewährte mir für diese Arbeit eine Kleinmittelhilfe und ein Habilitationsstipendium; sie stellte mir darüberhinaus für meine Forschung in Griechenland einen Kleinbus zur Verfügung.

Für alle diese Hilfen danke ich herzlich.

1. EINLEITUNG

1.1. Ziel der Arbeit

Diese Arbeit betrachtet Sprache als Algorithmus zur Lösung kommunikativer Probleme (vgl. dazu Kap. 2). Die Konstruktion des Algorithmus zur Simulation von Sprache und sprachlichen Aspekten ist dann wesentlicher Teil sprachwissenschaftlicher Tätigkeit. Die Eigenschaften des Algorithmus ergeben sich aus dem Spannungsfeld kommunikativer Aufgaben und in der Rede zu beobachtender Lösungen. Da Zahl und Struktur der sich in der Kommunikation stellenden Probleme nicht festzulegen sind, wird der Algorithmus nicht als fixes Programm zur Lösung einer endlichen Zahl strukturell definierbarer, typischer Probleme, sondern als adaptives, selbstlernendes System ausgelegt. Das selbstlernende System, das in dieser Arbeit beschrieben wird, simuliert einen Teilaspekt von Sprache: die Mengen morpho- und phonologischer Daten einiger bisher wenig erforschter romanischer Sprachen. Das selbstlernende System erhält als Eingabe phonische Ketten und zugehörige Strukturbeschreibungen (s. Kap. 4). Als Ausgabe erzeugt es generativ-phonologische Grammatiken. Die der Sprache eigentümlichen Aspekte des Selbstlernens und der Selbstorganisation werden vom Algorithmus im morphophonologischen Teilbereich simuliert. Der Algorithmus wird durch Computerprogramme realisiert und somit praktischer Bewährung ausgesetzt.

1.2. Zweck der Arbeit

Die Arbeit ist wegen der zugrundeliegenden sprachlichen Daten romanistisch[1]. In der Frage nach der Funktion von Sprache und den Bedingungen zur Erfüllung dieser Funktion ist sie allgemein sprachwissenschaftlich relevant. Abstrahiert von den zugrunde-

---

1) Die Arbeit ist romanistisch, weil ihr Gegenstand Daten aus romanischen Sprachen sind. Der Name der wissenschaftlichen Disziplin ergibt sich aus dem Gegenstand, nicht aus der Methode.

liegenden sprachlichen Daten berührt die Arbeit auch Fragen allgemeiner Problemlösestrategien in offenen Systemen. Die Realisierung des Algorithmus in einem Computerprogramm macht die Arbeit praktisch anwendbar.

Die hier verarbeiteten sprachlichen Daten stammen aus dem Surselvischen, dem Meglenorumänischen und aus drei Dialekten des Aromunischen. Die selbstlernende Grammatik erzeugt die ersten generativ-phonologischen Beschreibungen dieser Sprachen. Die Auswahl der Sprachen rechtfertigt sich jedoch nicht allein durch das Fehlen generativer Phonologien; vielmehr fordern ihr morphologischer Reichtum und ihre typologischen Unterschiede die selbstlernende Grammatik zu exemplarischer Bewährung heraus. Die weitmaschige Erfassung der süddanubischen Varietäten der Romania führt zu einer generativen Dialektologie, die die dialektale Ausgliederung als sprachgruppenindividuelle Lösung kommunikativer Probleme (im Kontakt mit nichtromanischen Sprachen) widerspiegelt. Damit betritt die Arbeit auch hinsichtlich der ausgewählten sprachlichen Da und in der Interpretation der Ergebnisse Neuland.

Die Fähigkeit zum Selbstlernen, die dieses System auszeichnet, beinhaltet, daß Teilaspekte von Sprache auch an anderen als den genannten romanischen Sprachen simuliert werden können. Die Voraussetzungen, an die die selbstlernende Grammatik gebunden ist, liegen nicht in der Struktur ausgewählter Einzelsprachen, sondern sind sprachwissenschaftlicher, nicht sprachlicher, Praxis eigen. Die Voraussetzungen ergeben sich allgemein aus den Bedingungen wissenschaftlicher Arbeit und speziell aus Grundsätzen und Annahmen, die sprachwissenschaftliche Arbeit leiten. Als Synthese dieser Bedingungen ist die selbstlernende Grammatik folglich allgemein sprachwissenschaftlich relevant.

Praktisch eignet sich die selbstlernende Grammatik - da losgelöst von einzelsprachlichen Daten und Strukturen konzipiert - als ein Instrument zur Erzeugung sprachwissenschaftlicher Kenntnisse, das zur morphophonologischen Beschreibung von Sprachen eingesetzt werden kann.

Die selbstlernende Grammatik ist eine Vorrichtung zur automatischen Lösung kommunikativer Probleme; ihr Ziel ist die Sicherung des kommunikativen Erfolgs. Betrachtet man die automatische Lösung von Problemen und die dafür angewandten Methoden unabhängig von sprachlichen Daten, so stellen sich die über die Sprachwissenschaft hinaus relevanten Probleme Selbstlernen, Selbstorganisation, automatische Problemlöseverfahren und die dazu notwendigen Regelungsmechanismen, die Frage der Erzeugung und Kontrolle von Hypothesen. Ähnliche Fragen finden sich in verschiedenen Bereichen, z.B. bei der Beschreibung biologischer, ökologischer oder sozialer Systeme. Die Untersuchung der Lern-, Erzeugungs- und Kontrollmechanismen in natürlicher Sprache gibt Aufschluß über deren Struktur; die selbstlernende Grammatik trägt dazu bei, diese Struktur aufzudecken.

Zur Beantwortung der gestellten Fragen wird eine praktische und folglich kontrollierbare Lösung präsentiert. Die Arbeit konstruiert einen Algorithmus, der durch ein Computerprogramm realisiert und getestet wird. Aspekte von Sprache werden auf einer Datenverarbeitungsanlage simuliert. Erkenntnisse über Sprachen werden durch einen Computer gewonnen.

1.3. Aufbau der Arbeit

Die Arbeit besteht aus zwei Hauptteilen. Der erste Teil - er umfaßt die Kapitel 2 und 3 - ist theoretisch ausgerichtet; er leitet aus den Annahmen über Sprache den Problemlösealgorithmus ab, der zweite Teil - Kapitel 4 - testet den Problemlösealgorithmus an sprachlichen Daten. Hier werden die Sprachgemeinschaften, die die sprachlichen Daten lieferten, die Datenaufnahme, die Daten und die Ergebnisse des Problemlösealgorithmus, die generativen Phonologien, beschrieben und kommentiert.

Der Aufbau des ersten Teils folgt dem Prinzip der schrittweisen Verfeinerung. Das Prinzip der schrittweisen Verfeinerung stellt die Formulierung des globalen Problems an den Anfang; das globale Problem wird durch Gliederung in Teilprobleme gelöst, die ihrerseits in ähnlicher Weise bis zu einer Ebene weiter analy-

siert werden, auf der einfache Probleme direkt gelöst werden können. Die Technik schrittweiser Verfeinerung erweist sich als effizientes Mittel zur Lösung komplexer Probleme; sie wird folglich für die Konstruktion von Algorithmen empfohlen (vgl. dazu Wirth (1978), Wirth (1979), Stetter (1981) und Kap. 2).

Bei der Anwendung des Prinzips der schrittweisen Verfeinerung wird das globale Problem in natürlicher Sprache formuliert. Die natürlichsprachige Beschreibung wird schrittweise durch eine formalere und die Struktur des Algorithmus widerspiegelnde Beschreibung ersetzt; der Algorithmus legt die dynamischen Prozesse fest, die sprachliches Lernen ermöglichen. Der Algorithmus wird nicht direkt als ein in einer Programmiersprache geschriebener Quelltext dargestellt, sondern in einer leichter lesbaren Beschreibungssprache.

Dieses Vorgehen bewahrt den Vorrang des sprachwissenschaftlichen Problems und seiner Lösung; dank strukturierten modularen Aufbaus bleiben auch komplexe Fragen durchsichtig; da das sprachwissenschaftliche Problem als vorrangig betrachtet wird, verzichtet die Arbeit auf die unmittelbare Beschreibung des Computerprogramms; dennoch sollte nicht vergessen werden, daß das Computerprogramm die Problemlösung in expliziter Weise an gegebenen Daten verifiziert.

Dem Prinzip schrittweiser Verfeinerung getreu findet sich im ersten Teil der Arbeit folgende Anordnung: Problembeschreibung, Problembedingungen, Problemlösung. Diese Anordnung wiederholt sich für jede der Verfeinerungsebenen.

Im Sinne des aufgestellten Prinzips beschreibt das zweite Kapitel das Globalproblem und spezifiziert die Bedingungen, denen die Problemlösung unterliegt.

Die Problemlösung wird im dritten Kapitel als Hierarchie von Teilproblemen dargestellt. Die hierarchisch tiefsten Teilprobleme finden ihre Entsprechungen in den Moduln des Computerprogramms, auf dessen Wiedergabe hier aus o.g. Grund und aus Platzgründen freilich verzichtet werden muß.

Der zweite Teil der Arbeit - das vierte Kapitel - enthält die Ergebnisse der selbstlernenden Grammatik. Dort werden die ausgewählten Sprachen, die aufgenommenen sprachlichen Daten, die Aufnahmepunkte, die Sprechergemeinschaften und die sprachliche Situation vorgestellt. Die Ergebnisse der selbstlernenden Grammatik werden wiedergegeben und im Sinne der Überwindung der allzu scharfen Dichotomien von Synchronie und Diachronie und von Sprachpunkt und Sprachraum (vgl. Kap. 2 dieser Arbeit) als dynamische Momentaufnahme interpretiert.

Gemäß der in Kapitel 2 eingeführten Terminologie stellt das vierte Kapitel parole und Regelwerk dar. Parole bezieht sich auf die aufgenommenen sprachlichen Daten, Regelwerk auf die durch die selbstlernende Grammatik erzeugten generativen Phonologien.

Das abschließende fünfte Kapitel faßt die Arbeit abwägend zusammen; es weist in einem Ausblick auf weitere Möglichkeiten selbstlernender Grammatiken hin.

2. SPRACHE

Ich betrachte Sprache als System mit drei Strukturebenen[1]. Ein System ist eine "abgegrenzte Anordnung von aufeinander einwirkenden Gebilden" (Stetter (1981:12)). Die aufeinander einwirkenden Gebilde nenne ich Strukturebene. Im folgenden werden die drei Strukturebenen und die zwischen ihnen bestehenden Einwirkungen beschrieben.

Die unterste Strukturebene nenne ich parole. Die parole wird von der mittleren Strukturebene, dem Regelwerk, erzeugt oder analysiert. Das Regelwerk enthält Regeln und das Lexikon. Die Einträge im Regelwerk werden durch das übergeordnete Steuerwerk erzeugt. Die Einträge im Regelwerk sind stets hypothetisch. Sie steuern Produktion oder Analyse der parole (vgl.a. Rolshoven (1982)). Die Erzeugung der Einträge im Regelwerk durch das Steuerwerk ist an jene Bedingungen gebunden, die für die Konstruktion von Hypothesen gelten. Die Einträge im Regelwerk müssen logisch widerspruchsfrei sein. Die vom Steuerwerk im Regelwerk produzierten Einträge müssen in der Lage sein, parole zu erzeugen und von anderen Systemen erzeugte parole zu analysieren[2].

---

1) Strukturebene ist von den niveaux d'articulation Martinets (1976:13 f) zu trennen. Dort geben niveaux d'articulation Möglichkeiten der Analyse der parole an. Hier jedoch sind Strukturebenen hierarchisch geordnete Stufen zur Erzeugung oder Analyse der parole. Bei Martinet beziehen sich die beiden niveaux d'articulation auf die parole, hier sind drei Strukturebenen einander jeweils nachgeordnet.

2) In der Vielfalt erzeugender und analysierender (natürlicher oder hier künstlicher) Systeme, die über die parole kommunizieren, zeigt sich der soziale Charakter der Sprache. Im Zusammenspiel von Systemen, die parole erzeugen oder analysieren, kann kommunikativer Mißerfolg zu Modifikationen im Regelwerk des erzeugenden oder des analysierenden Systems führen. Im letzteren Fall paßt sich das analysierende System an. Übernimmt es die Modifikation auch bei der Produktion von parole, so wäre das Ergebnis sprachlicher Wandel.

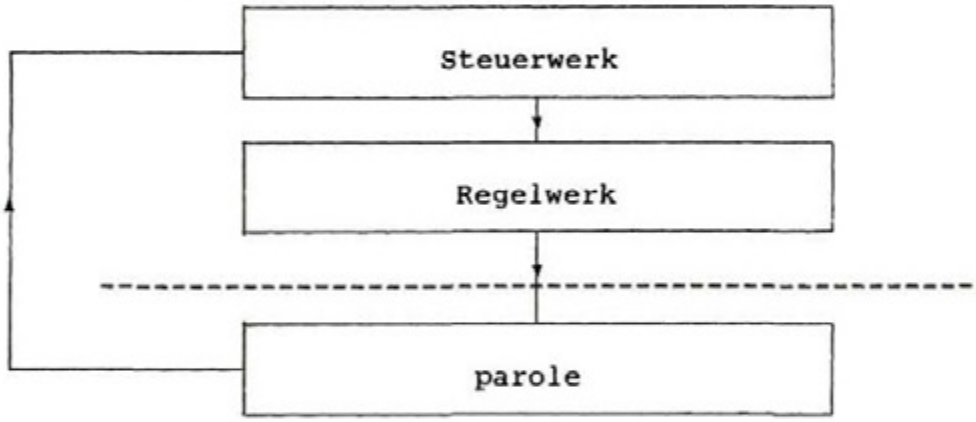
Die Produktionen oder Analysen des Regelwerks müssen adäquat sein. Sind sie es nicht, so hat das Steuerwerk Einträge des Regelwerks zu modifizieren, zu tilgen oder neu zu erzeugen. Enthält das Regelwerk Einträge gleicher erzeugender oder analysierender Kraft, so tilgt das Steuerwerk die notationell aufwendigeren. Das Steuerwerk begünstigt von Einträgen gleichen notationellen Aufwands den der größten analysierenden oder erzeugenden Fähigkeiten. Dies gilt auch für die Auswahl von Hypothesen; größte notationelle Einfachheit und weitester Anwendungsbereich sind ausschlaggebende Auswahlkriterien. Für den Fall konkurrierender Einträge, die entweder durch notationelle Einfachheit und kleinen Anwendungsbereich oder durch notationelle Komplexität und großen Anwendungsbereich gekennzeichnet sind, wägt das Steuerwerk Vor- und Nachteile der Auswahl ab. Entscheidend ist der notationelle Aufwand für die Beschreibung des gesamten Bereichs. Hier erweisen sich notationell komplexere Einträge oft als insgesamt vorteilhafter und genereller als mehrere spezifische notationell einfacherere Lösungen.

Die Bedingungen für die Konstruktion von Hypothesen werden in den einzelnen wissenschaftlichen Disziplinen beobachtet. Sie sind Normen für die automatische Erzeugung des Regelwerks (vgl. Popper (1971), Ströker (1973), für sprachwissenschaftliches Arbeiten Hjelmslev (1961:14,18)).

Die untere Strukturebene, die vom Regelwerk erzeugte oder analysierte parole, ist an das Steuerwerk rückgekoppelt. Diese Rückkoppelung ermöglicht dem Steuerwerk, den kommunikativen Erfolg der parole zu messen und bei Mißerfolg Modifikationen im Regelwerk zu veranlassen. Darüber hinaus ist das Steuerwerk - gemäß der Forderung nach Effizienz und notationeller Ökonomie - stets bestrebt, parole mit minimalem Aufwand im Regelwerk erfolgreich zu bearbeiten.

Durch die Rückkoppelung der parole an das Steuerwerk entsteht ein spezieller, sprachbezogener Regelkreis. Der Sollwert des Regelkreises ist kommunikativer Erfolg. Das Steuerwerk mißt Abweichungen vom Sollwert und verändert je nach Grad der Abweichung das Regelwerk.

In diesem Modell ist die untere Strukturebene der Beobachtung zugänglich. Die übergeordneten Ebenen sind " ...rekonstruierbar in einem induktiven Prozeß" (Seiler (1973:11)). Hilfe bei der (automatischen) Rekonstruktion leisten die sprachlichen Daten, die Bedingungen zur Gewinnung von Hypothesen und schließlich Erkenntnisse über Struktur und Arbeit der Systeme, die Steuer- und Regelwerk realisieren (vgl. Raible (1981), Creutzfeldt (1984), Osherson, Weinstein (1984:629), Berwick, Weinberg (1984)). Steuer- und Regelwerk können in natürlichen Systemen, menschlichen Sprechern bzw. Hörern, oder in künstlichen Systemen individuell aufgebaut und programmiert sein. Die Kommunikation verläuft über die intraindividuelle parole. Sie verbindet die individuellen Regel- und Steuerwerke. Dies verdeutlicht das Schema[3]:



Zur inhaltlichen Klärung wird die hier vorgelegte Auffassung von Sprache mit der strukturalistischen und der generativ-transformationellen auf Gemeinsamkeiten hin verglichen und dann eventuell dagegen abgegrenzt[4]. Der wesentliche Unterschied liegt in der neueingeführten Gliederungsebene Steuerwerk. Sie macht Sprache zum Problemlösealgorithmus.

---

3) Die Abfolge der Auffassungen über Sprache und der daraus folgenden Beschreibungen von Sprache wird in Arbeiten zur Geschichte der Sprachwissenschaft(z.B. bei Renzi (1976)) in Anschluß an Kuhn (1967) als Paradigmenwechsel bezeichnet. Zur Kritik am Begriff des wissenschaftlichen Paradigmas Shapere (1971) und zu seiner Verwendung in der Sprachwissenschaft ablehnend Percival(1976).

4) Nilson (1982:17) verwendet (in anderem Zusammenhang) die Begriffe global data base, production rules und control system. Sie finden sich hier als parole, Regelwerk und Steuerwerk.

Problemelösen beruht auf der Fähigkeit, Problemlösestrategien zu erlernen. Systeme mit dieser Fähigkeit heißen selbstlernend. Lernfähigkeit wird als Kriterium für Intelligenz betrachtet. Künstliche lernfähige Systeme führen zur Artificial-Intelligence-Forschung.

Das hier vorgeschlagene System wird darüber hinaus in Beziehung zur praktischen Arbeit des Sprachwissenschaftlers gesetzt. Außerdem wird es mit anderen Problemlösesystemen verglichen. Abschließend werden die Namen der Gliederungsebenen motiviert.

Strukturalistisch ist Sprache ein System von Zeichen. Die systemhafte langue wird in der parole aktualisiert (vgl. z.B. Helbig (1972:35)). Über die parole schließt der Sprachwissenschaftler auf die langue. Die sprachlichen Zeichen sind paradigmatisch gegeneinander abgesetzt. Syntagmatisch können sie gereiht werden. Paradigmatisch gegeneinander abgesetzte Zeichen stehen in Opposition. Das sprachliche Zeichen besteht aus signifiant und signifié.

Da Zeichen in Opposition stehen, ist es möglich, Aufdeckungsverfahren zu entwickeln, die Zeichen und Zeichenklassen aus syntagmatischen Kontinua herausarbeiten. Dank der Begriffe Opposition und paradigmatische Beziehung sind syntagmatische Kontinua auf diskrete sprach(wissenschaft)liche Einheiten abbildbar.

Im Rahmen dieser Arbeit sind die Möglichkeiten der Bildung diskreter Einheiten, der Bildung von Klassen diskreter Einheiten und ihrer syntagmatischen Reihungen wichtig. Mit dem Strukturalismus teilt die Arbeit die Ansicht, parole als Grundlage für Rückschlüsse auf ein sprachliches System zu nutzen. Dieses System ist freilich komplexer als die strukturalistische langue. Ist die langue ein eher fixes und statisches System von Oppositionen, so ist das Regelwerk im Gegensatz dazu dynamisch und dauernd veränderbar. Die Veränderungen hängen vom kommunikativen Erfolg der parole ab. Erfaßt das Steuerwerk einen unzureichenden kommunikativen Erfolg, so erzeugt es eine neue Hypothese im Regelwerk, deren Erfolg es durch die Rückkoppelung an die parole mißt. Das Regelwerk ist dank der Kontrolle des Steuerwerks ein dynamisches, permanent auf Anpassung an kommunikativen Erfolg und nota-

tionelle Minimierung gedrilltes System (von Subsystemen im Sinne der Prager Schule, vgl. Vachek (1966)).

Die strukturalistische Sprachauffassung verlegte den Schwerpunkt des sprachwissenschaftlichen Interesses von der parole auf die systemhafte langue. Diese Arbeit verlegt ihn von dem der langue am ehesten entsprechenden Regelwerk auf das Steuerwerk. Das Regelwerk ist weit weniger starr als die langue und somit wegen der Lernfähigkeit des umgebenden Systems der Problematik kommunikativen Mißerfolgs gewachsen. Das Steuerwerk enthält einen Mechanismus zur Konstruktion von Einträgen im Regelwerk.

Sprache und sprachwissenschaftliche Tätigkeit gewinnen bei der Betrachtung von Sprache als dreistufiges System gewisse Ähnlichkeiten: Die Aufgabe des Hypothesen konstruierenden Steuerwerks übernimmt in der strukturalistischen Linguistik der Sprachwissenschaftler. Er konstruiert - manuell - die Hypothesen der langue (= Regelwerk). In der Simulation des dreistufigen Systems wird diese (bei natürlichen Sprechern meist unbewußte, bei Sprachwissenschaftlern meist bewußte) Tätigkeit automatisiert. Wie sich aus der parole Rückschlüsse auf das Regelwerk gewinnen lassen, so läßt das Regelwerk auf die Struktur des Steuerwerks schließen.

Vielleicht entspricht das Steuerwerk dem langage Saussures (1976). Aber Saussure gibt keine Verfahren zur Konstruktion der langage und zur Erzeugung der parole an. Die strukturalistische Praxis beschäftigt sich dann auch mehr mit der langue als Instrument taxonomischer Unterscheidung. Damit steht dieser Arbeit das Werkzeug zur Kontrolle von Hypothesen zur Verfügung; die Kontrolle freilich dient nicht der (manuellen) Tätigkeit des Sprachwissenschaftlers, sondern dem Steuerwerk bei der Erstellung von Konstruktionsverfahren zur Erzeugung oder Analyse von parole.

Die generative Transformationsgrammatik betrachtet Sprache unter vorwiegend syntaktischen Gesichtspunkten als (autonomes) System sprachlicher Module (Chomsky(1982),(1984)); die Regeln dieser Module repräsentieren die Kompetenz eines Sprechers, das sprachliche Wissen. Sätze werden durch diese Regeln erzeugt oder analysiert. Die Regeln sind durch gewisse - möglicherweise gene-

tische - Vorgaben geprägt, die von der Grammatiktheorie als Prinzipien formuliert werden. Die Annahme solcher Vorgaben beruht u.a. auf folgenden Überlegungen (Fanselow, Felix(1987:106,126):

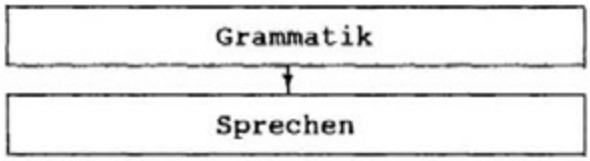
1. Lernbarkeit von Sprache ohne Vorgaben ist wegen
   a) quantitativer Undeterminierbarkeit vorgegebener Sprache,
   b) qualitativer Undeterminierbarkeit und
   c) Fehlen negativer Evidenz
   schwierig bzw. ausgeschlossen,
2. Spracherwerb ist nur in einem bestimmten zeitlichen Fenster möglich,
3. Sprachen sind infinit (sie enthalten unendlich viele Sätze),
4. vorgegebene Sprache ist oft nicht vollständig,

Gegen diese Argumente werden funktionale und psychologische Einwände vorgebracht. Das Konzept der Prinzipien wurde dahingehend verfeinert, daß diese Prinzipien parametrisiert wurden. Die Werte der Parameter werden erlernt. Die Prinzipien, die den Erwerb syntaktischen Wissens möglicherweise steuern, und die hier vorgeschlagene Sicht von Sprache sind kompatibel. Dank der Prinzipien wird die Suchtiefe beim Erwerb syntaktischen Wissens stark reduziert, da die Auswahl dessen, was zu lernen ist, eingeschränkt wird. Damit beeinflussen oder steuern diese Prinzipien das Verhalten des Steuerwerks. Für den Erwerb semantischen und morphologischen Wissens scheint es keine steuernden Prinzipien zu geben; folglich finden sich beim morphologischen Spracherwerb auch negative Evidenzen, z.B. durch Übergeneralisierung.

Die Strukturebene parole dieser Arbeit entspricht in gewisser Sicht der Performanz, die Ebene Regelwerk der Kompetenz. Die Arbeit teilt mit der Transformationsgrammatik das Format der Regeln im Regelwerk. Diese Regeln (s. Maurer (1969), Saloma (1973), Beeh (1981)) erzeugen oder analysieren (letzteres in grammatischen Beschreibungen freilich seltener; praktische Anwendung findet sich meist in parsing-Programmen zur automatischen Zuweisung von Strukturbeschreibungen an Sätze) durch umgekehrte Anwendung von Regeln parole. Die Regeln vermitteln hinreichend explizite Beschreibungen.

Diese Arbeit betrachtet Sprache als dreistufiges rückgekoppeltes System, die generative Transformationsgrammatik sieht Sprache eher als ein durch Prinzipien denn durch an kommunikativem Erfolg orientiertes System. Sprache ist darin weniger ein Problemlösealgorithmus. Möglicherweise erklärt sich daraus auch die Neigung zu starker Idealisierung eines Hörers oder Sprechers (Chomsky (1965:3)). Sicherlich ist einem dreistufigen System die Möglichkeit sprachlichen Wandels inhärent.

Die Auffassung von Sprache als System mit zwei Ebenen sieht Sprache als eher starren Code (vgl. das hier abgebildete Schema eines zweistufigen Systems Sprache:)[5].



Code ist eine (mathematische) Funktion, die Nachrichten N1 Nachrichten N2 zuordnet (vgl. Bauer, Goos (1973:28,52)). Der Code, der noematische Nachrichten in gesprochene (und umgekehrt) transformiert, wird in der Sprachwissenschaft Grammatik genannt. Hier kann Kommunikation nur möglich sein, wenn man ideale Sprecher (Hörer) mit identischen Codes postuliert, da unterschiedliche Grammatiken Verständigung unmöglich machen würden. Dabei stellen sich jedoch folgende Fragen:

1. Wie kann das Verständnis sprachlich falscher Sätze erklärt werden (vielleicht als Performanzangelegenheit?),
2. Sind eventuelle Prinzipien im Gehirn lokalisierbar, z.B. durch spezifische Aphasien,
3. Wie kann die Grenze zwischen frei erlerntem oder erlernbarem sprachlichen Wissen und genetisch vorgeprägtem festgelegt werden?

---

5) Diese Ansicht findet sich z.B. bei Nilson (1982:2): "This generating and understanding of language is an encoding and decoding problem of fantastic complexity".

Hier würde ich für eine wenn möglich weitgehende freie Erlernbarkeit und Modifizierbarkeit sprachlicher Regeln plädieren. Dies würde nämlich Kommunikation trotz gestörter Übertragung, trotz unterschiedlicher individueller Grammatiken und darüberhinaus Sprachwandel, Ausbildung von Dialekten (vgl. dazu Kap. 4) und Gruppensprachen und Formen des Spracherwerbs erfassen. Sprache ist sowohl ein Mittel zur (originellen, wie vor allem Kindersprache zeigt,) Lösung kommunikativer Probleme als auch Anwendung eines eher starren Programms. Diese beiden Faktoren machen die Plastizität natürlicher Sprachen aus.

Die Einführung der oberen Ebene Steuerwerk relativiert den Status der Grammatik. Sprache ist nicht nur ein Code, der parole erzeugt oder analysiert, sie ist mehr: sie ist dank des Steuerwerks auch eine Vorrichtung zur Codeerzeugung.

Dank der dritten Ebene Steuerwerk wird der kommunikative und soziale Charakter der Sprache hinreichend berücksichtigt. Sprache und Kommunikation bedingen einander. An dem kommunikativen Erfolg mißt sich die Effizienz der Grammatik und ihrer Verarbeitung.

Im dreistufigen Modell sind die Möglichkeiten für die Erzeugung des grammatischen Codes gegeben. Genetische Vorgaben für diesen Code können berücksichtigt werden. Freilich sind die Aussagen über das Verhältnis von genetischer Vorgabe und einzelsprachlicher Grammatik nicht leicht zu kontrollieren und nicht der Falsifikation offen (vgl. dazu Piattelli-Palmerini (1980), Beeh (1981)). Falsifikationsmöglichkeiten sollten bei wissenschaftlichen Arbeiten gegeben sein.

Die Generativität einer Grammatik und die daran anschließende Betrachtung von Sprache als ἐνέργεια nach Humboldt (1963) sollte sich nicht allein auf die (syntagmatische) Erzeugung

---

6) Sicherlich gibt es in natürlichen Systemen Vorrichtungen, die die hier Steuerwerk benannte Ebene und die Prinzipien ihres Funktionierens beherbergen. Durch Einführung der dritten Strukturebene Steuerwerk erweitert diese Arbeit den Bereich sprachwissenschaftlicher Tätigkeit von der Grammatik zu einer Metagrammatik.

mit Hilfe (rekursiver) Ersetzungsregeln beziehen (Rekursion alleine ist keine hinreichende Bedingung zur Charakterisierung natürlicher Sprachen), sondern auch die Möglichkeit der metasprachlichen Zeichenproduktion auf der Ebene des Regelwerks beinhalten (s. Bossong (1979), Coseriu (1975)).

Im dreistufigen System Sprache erzeugt das Steuerwerk im Regelwerk Hypothesen, mit deren Hilfe parole erzeugt oder analysiert wird. Kommunikativer Mißerfolg in der parole löst eine Intervention des Steuerwerks im Regelwerk aus. Die Hypothesen des Regelwerks werden reformuliert. Sind sie erfolgreicher, bleiben sie dem Regelwerk erhalten. Sind sie es nicht, werden sie getilgt. Neu erzeugte Hypothesen werden vom Steuerwerk auf Adäquatheit ihrer Produktionen oder Analysen in der parole und auf logische Korrektheit im Rahmen der Einträge des Regelwerks überprüft. Da Speicherplatz im Regelwerk beschränkt ist und Hypothesen geringeren notationellen Aufwands - sie erlauben dem System, alternative Problemlösestrategien zu speichern und sichern so eine höhere Verhaltensvariabilität - bevorzugt werden, ist das Steuerwerk darauf spezialisiert, im Regelwerk jene Regeln einzusparen, die für die Erzeugung oder Analyse der parole funktionslos geworden sind.

Das dreistufige Modell erklärt die Kommunikationsmöglichkeiten auch unterschiedlicher sprachlicher Systeme. Es sieht Sprachwandel als Anpassung an die Erfordernisse erfolgreicher Kommunikation und Spracherwerb als Phase intensiver Hypothesenerzeugung, die parole gelingen läßt. Gerade im morphophonologischen Spracherwerb (s.a.u.) kann das Spiel von Hypothesenbildung, -tilgung und -modifikation beobachtet werden. Rückkoppelung ist für den Spracherwerb notwenig. Ist sie bei natürlichen Sprechern im frühen Spracherwerb gestört, z.B. bei tauben Kindern, so wird die Lautproduktion abgebrochen.

Wenn aber Sprache ein Spiel von Hypothesenerzeugung, -tilgung und -modifikation ist, dann müßten in allen sprachlichen Bereichen beim Spracherwerb Fehler zu beobachten sein, die auf der Konstruktion zu falsifizierender Hypothesen beruhen. Dies ist in einigen Bereichen des syntaktischen Spracherwerbs nicht der Fall (s.a.o.), und daraus kann auf vorgegebene sprachliche Strukturen

geschlossen werden (vgl. Lenneberg (1967), Chomsky (1975), (1981), Gleitman (1984:556 ff)). Diese könnten in parametrisierter Form auftreten und werden im Spracherwerb mit Werten versehen (vgl. Berwick, Weinberger (1984:199 ff)).

Freilich setzt das dreistufige Modell Sprache keinesfalls ein sofortiges Testen einer konstruierten Hypothese durch Sprachproduktion voraus; die oben erwähnten Beobachtungen zum Ausbleiben von Fehlern könnten auch darauf beruhen, daß gewisse sprachliche Strukturen erst dann zur Sprachproduktion genutzt werden, wenn sie hinreichend sicher sind. Sie erwerben ihre Sicherheit durch über einen längeren Zeitraum bestehende erfolgreiche Analyse. Eine solche Spracherwerbstrategie fördert nicht nur korrekte Sprachproduktion, sondern verhindert auch eine Bildung von inkorrekten Hypothesen, die sich in der Spracherzeugung allzu stark verfestigen könnten. Daraus könnten dem sprachlernenden Individuum schwerwiegende Nachteile erwachsen. Dagegen wird freilich eingewendet, daß oberflächennahe Hypothesen im Spracherwerb nicht zwecks Härtung weggespeichert werden können. Fraglich ist in diesem Zusammenhang (beim syntaktischen Spracherwerb), in welcher Form Hypothesen aufbewahrt werden. Falls es hier dennoch Hypothesenspeicherung gibt, so könnte auch deren Hypothesenhärtung oder -bewährung auf Erfolg bei Analyse der parole beruhen.

Der Aufbau des Regelwerks setzt den Sprecher in die Lage, eine immer größere Zahl kommunikativer Situationen erfolgreich zu beherrschen. Verhaltensänderungen, die zu erfolgreichem Verhalten führen, werden Lernen genannt (Cube (1971: 71 ff)). Bei menschlichen Sprechern werden verbale Verhaltensänderungen besonders während der ersten Lebensjahre beobachtet.

Das Modell hier beschreibt nicht allein Sprache, sondern durch die Einführung der Strukturebene Steuerwerk auch die Praxis vieler Sprachwissenschaftler. Betrachtet man Abschnitte der Geschichte der Sprachwissenschaft, so stellt sich die Aufgabe des Sprachwissenschaftlers als Hypothesenerzeugung, -modifikation und -tilgung in der Grammatik, um deren Produktionen von in der parole natürlich erzeugter parole anzupassen.

Das Spiel von Hypothesenerzeugung und -falsifikation und seine Kontrolle sind im dreistufigen Modell Sprache integriert und in seiner expliziten Beschreibung durch ein Computerprogramm automatisiert.

Das Modell zeigt ein Verhalten, das die Wissenschaftstheorie in den wissenschaftlichen Einzeldisziplinen beobachtet und Wissenschaftler, Theorie und Fakten in einer Rückkoppelung vereint. Für die Ebene der Theorie gelten ähnliche Ökonomieprinzipien wie für das Regelwerk; wissenschaftliche Theorie und Sprache unterliegen dem gleichen Prinzip semiotischer Systeme.

In diesem Modell verlagert sich der Schwerpunkt sprachwissenschaftlicher Tätigkeit auf die obere Strukturebene, das Steuerwerk. Die Konstruktion eines adäquaten Steuerwerks wird zur Hauptaufgabe. Im dreigliedrigen System Sprache verwischen die scharfen Grenzen, auf der die Dichotomie von Kompetenz und Performanz von langue und parole beruht. Die Dichotomie reflektierte die Trennung einer flexiblen unteren Ebene, die von einer eher starren oberen Ebene gesteuert wurde. Im dreistufigen System verschwinden diese Unterschiede. Dank des Steuerwerks werden die Einträge im Regelwerk ähnlich generativ wie früher die untere sprachliche Ebene. Das System insgesamt wird sehr dynamisch, sehr anpassungsfähig, da selbstlernend.

Die Komplexität von Sprache macht es wünschenswert, aus der vergleichenden Beschreibung anderer komplexer und dynamischer Systeme Gewinn für die Beschreibung natürlicher Sprache zu ziehen (s. Jakobson (1969) und Seiler (1973:10,11)). Dazu wird der Gegenstandsbereich der Biochemie herangezogen. Die Biochemie versucht, die komplexe Vielfalt der Erscheinungsformen des Lebens auf die Steuerung durch einen überschaubaren genetischen Code zurückzuführen. Auch hier besteht eine Opposition zwischen der (relativen) Einfachheit der aus endlich vielen Elementen und ihren Kombinationen bestehenden Steuerung und der ungeheuer großen Vielfalt der Formen des Lebens (s. Hobohm (1977)).

Hinsichtlich eines individuellen Organismus handelt es sich bei der genetischen Steuerung um ein zweistufiges System. Es

ähnelt somit dem generativ-transformationellen Modell der Sprache und wurde folglich bereits mit ähnlichen notationellen Mitteln von Hermann, Rozenberg (1975) beschrieben.

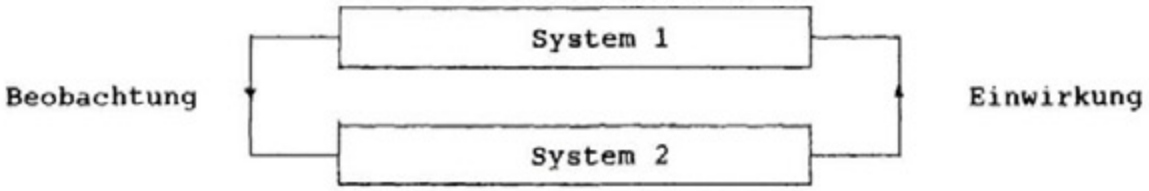
Die Möglichkeit der Änderung des genetischen Programms durch Mutation dynamisiert das System. Es erhält dadurch die Chance zur Evolution. Hier wird also eine dritte Ebene zur Programmänderung des genetischen Codes betrachtet. Freilich muß der Unterschied zum dreistufigen System Sprache beachtet werden: die übergeordnete Instanz, die die Änderung des genetischen Programms bewirkt, liegt außerhalb des Systems. Das System verändert sich nicht zielstrebig, sondern zufällig. Evolution ergibt sich für einen individuellen Organismus nicht kraft eigener Anstrengung, sondern in einer Art von trial and error erhält ein Organismus aufgrund von außen herangetragener Änderung bessere Überlebenschancen (s.a. Monod (1977)).

Erst höher entwickelte Organismem integrieren übergeordnete Instanz und Programmspeicher - im dreistufigen Modell Sprache Steuerwerk und Regelwerk - in einem zum Organismus gehörenden System. Der Programmspeicher ist in diesem System nicht mehr (relativ) starr wie in genetischen Programmen organisiert, sondern zumindest teilweise veränderbar aufgebaut. Während eine Änderung im genetischen Programm individuell katastrophale Folgen haben kann, wirkt sich eine falsche Hypothese mit negativen Folgen in einem flexiblen Programmspeicher für das Individuum meist nicht allzu schlimm aus, da die übergeordnete Instanz in diesem Fall eine neue Hypothese erzeugt und auf Erfolg oder Mißerfolg prüft. In der Zielgerichtetheit der Suche nach Erfolg zeigen sich die höher entwickelten Systeme mit drei Strukturebenen jenen mit zwei Ebenen überlegen (vgl. Hofstadter (1982)).

Dank der Rückkoppelung des Steuerwerks an die parole im dreistufigen Modell Sprache ist dem System Problemlösefähigkeit inhärent. Systeme mit der Möglichkeit abstrahierender Abspeicherung von Problemlösungen sind lernfähig. Unter Lernen wird eine relativ dauerhafte (hier:abgespeicherte) Veränderung des Verhaltens verstanden, die sich aus Übung ergibt (nach Manis (1974:20), vgl. Diskussion von Lerntheorien in Cube (1971), Kap.2). Systeme mit Lernfähigkeit sind intelligent. Künstliche Systeme mit Lernfähig-

keit besitzen künstliche Intelligenz[7]. Lernen läßt sich mit Cube (1971) als Akkomodationsprozeß mit Redundanzaufbau statistisch und mit Menzel (1970) automatentheoretisch beschreiben. Lernfähige Systeme umfassen zwei Subsysteme, im nachfolgendem Schema System 1 und System 2 genannt. Das System 1 beobachtet das System 2, das System 2 wirkt durch Rückkoppelung auf das System 1 ein. Dank dieser Verbindungen baut das System 1 Information im System 2 ab; es erzeugt Redundanz. Der Lernvorgang ist nach Cube (1971) beendet, wenn maximale Redundanz aufgebaut worden ist.

Das Schema nach Lipp (1967:1), Bild 1.1, hat folgendes Aussehen:



Das Schema von Lipp ist direkt auf das dreistufige Modell der Sprache übertragbar, wenn man sich gemäß Skizze o. p. 8 die oberen beiden Strukturebenen - wie durch die gestrichelte Linie angedeutet - im System 1 vereint denkt und parole System 2 zuordnet. Das System Sprache wird damit zu einem Spezialfall lernender Automaten. Es kann durch Automaten simuliert werden.

Das hier skizzierte Modell Sprache ist in sprachwissenschaftlichen Einzeldisziplinen einsetzbar; in dieser Arbeit wird es für morphophonologische Daten verwendet. Damit kann das Problem der genetischen Vorgabe im folgenden ausgeklammert werden. Mit der

7) Hier sei noch einmal ausdrücklich darauf verwiesen, daß Lernfähigkeit als Kriterium für Intelligenz gesehen wird (s.a. von Cube (1971:178). Der Begriff der Künstlichen Intelligenz wird oft vage als Fähigkeit zur Anwendung von Wissen aus Datenbanken in konkreten Situationen beschrieben, nicht jedoch als Erwerb von Wissen (so Newell (1973), Wilks (1973), Nilson (1982), Batori, Krause, Lutz (1982)).

Verschiebung des Schwerpunkts linguistischen Arbeitens auf die erste Strukturebene wird die Konstruktion des Steuerwerks bedeutsam[8]. Ihm kommen zwei Aufgaben zu: Kontrolle der parole einschließlich ihres kommunikativen Erfolgs und Erzeugung von Hypothesen für das Regelwerk. Der Name der höchsten Strukturebene erinnert daran, daß sie die Konstruktion von Symbolen und symbolischen Ausdrücken im Regelwerk steuert. Die Fähigkeit zur Erzeugung der Syntax für die Einträge im Regelwerk und zur Kontrolle der Wirkungen dieser Einträge in der parole macht das System selbstlernend; bei der Konstruktion des Steuerwerks handelt es sich also um die Beschreibung einer selbstlernenden Grammatik, deren Syntax die Erzeugung und Verkettung von Symbolen im Regelwerk steuert. Diese selbstlernende Grammatik gibt der Arbeit den Namen.

Das hier konstruierte Steuerwerk steht vor der gleichen Aufgabe, vor der ein generativer Phonologe steht. Es bildet ein Korpus von Sätzen auf generativ-phonologischen Regeln und ein Lexikon ab. Wie für generativ-phonologisch arbeitende Sprachwissenschaftler üblich, bestehen die Sätze aus Ketten in (broad phonetic) Transkription und zugehörigen semantisch-syntaktischen Strukturbeschreibungen (vgl. Kap. 4). Damit sind Ausgangs- und Zielsymbolfolgen festgelegt. Die Zielsymbolfolgen, Lexikon und generativ-phonologische Regeln, finden sich im Regelwerk.

Nach Ausarbeitung des Konzepts des dreistufigen Systems Sprache und mehrmonatiger Feldarbeit zur Aufnahme der sprachlichen Daten nahm die Konstruktion des Steuerwerks den größten Teil der Zeit, die auf diese Arbeit verwendet wurde, in Anspruch. Die Verschiebung des Schwerpunkts linguistischen Arbeitens auf die erste Strukturebene brachte erhebliche Schwierigkeiten mit sich. Vor allem die Umsetzung der Themenstellung in einen vollkommen expliziten Algorithmus, den ein Computerprogramm realisiert, wirkte

---

8) Für zweistufige Modelle der Sprache gibt es einige computerlinguistische Arbeiten generativ-phonologischer Provenienz, z.B. Friedmann, Morin (1971), Morin, Friedmann (1971), Morin (1971), Eastlack (1977). O.g. Arbeiten machen den Unterschied zu dieser Arbeit deutlich: Dort ist eine generativ-phonologische Grammatik Voraussetzung für ein Computerprogramm, hier ist sie sein Ergebnis.

zeitlich verzögernd. Das Computerprogramm ist unverzichtbarer Bestandteil einer Arbeit über das dreistufige System Sprache, da es eine theoretische, vollkommen explizite Formulierung einer Problemlösung vorlegt, die unter Ausschluß menschlicher Intuition Theorien testet und die Korrektheit einer wissenschaftlichen Ansicht auch dann nachweist, wenn sich eventuell explizit formulierte Teile der Theorie wegen ungewohnter Notation (daher wird hier ja auch von einer Präsentation des Programms abgesehen) direktem, verstehendem Zugriff entziehen. Gerade für das dynamische dreistufige Modell bedarf es der Kontrolle durch den Computer, da die Überprüfung sämtlicher sich gegenseitig beeinflussender Variablen manuell über einer offenen Strukturebene parole kaum mehr möglich ist.

3. DAS STEUERWERK

3.1. Globalfunktion

Im dreistufigen Modell Sprache ist die parole eine offene Systemkomponente. Sie ist nicht Ergebnis eines einzigen Regel- und Steuerwerks, sondern vieler anderer individueller Regel- und Steuerwerke und hat somit sozialen Charakter. Die Struktur eines individuellen Regelwerks wird durch die soziale parole mitbestimmt.

Folglich präsentiert diese Arbeit die Struktur des Regelwerks mit den Komponenten Lexikon und Regeln nur in einer Momentaufnahme. Die Momentaufnahme zeigt im Sinne des zweistufigen Modells der Sprache in der Transformationsgrammatik eine generative Phonologie. Das Regelwerk wird nach Bewältigung einer konkreten kommunikativen Situation, nach der jeweiligen Analyse der verschiedenen Textkorpora gezeigt. Das Steuerwerk bildet somit Textkorpora auf das Regelwerk ab. Die Abbildungsvorschrift wird allgemein durch Annahmen über Sprache und (sprachwissenschaftliche) Theoriebildung, konkret durch die Struktur der parole - d.h. hier des gerade zu verarbeitenden Textkorpus - bestimmt.

Diese Arbeit beschreibt die Interaktion der drei Ebenen des Systems Sprache. Die Textkorpora -parole(s)- bilden die Eingabedaten, das Regelwerk die Ausgabedaten; zu Beginn der Interaktion ist das Regelwerk leer. Das Steuerwerk, ein Algorithmus, der selbst veränderbar ist, übernimmt die Manipulation der Daten der parole. Der hier konstruierte Algorithmus ist ein Modell des Regelwerks im dreistufigen System Sprache.

Die Beschreibung des Steuerwerks setzt die Beschreibung der parole und des Regelwerks voraus. Da die Einwirkungen des Steuerwerks auf das Regelwerk im Hinblick auf kommunikativen Erfolg bewertet werden, müssen Kriterien für kommunikativen Erfolg bestimmt werden.

3.1.1. Format der parole (Eingabe)

Das Steuerwerk geht von einem Textkorpus aus. Das Textkorpus besteht aus Sätzen. Die Sätze des Textkorpus bilden die jeweilige parole. Zu jedem Satz gehört eine durchlaufende Zeichenkette in phonetischer Transkription (phonische Kette in broad transcription) und die zugehörige Strukturbeschreibung. Die Strukturbeschreibung besteht aus Klammerstrukturen und sogenannten Bedeutungseinträgen aus dem Lexikon einer generativen Syntax (vgl. Eingabekorpora in Kap. 4). Die Klammerstrukturen sind Strukturbäumen äquivalent. Klammerstrukturen wurden hier gewählt, da sie - im Gegensatz zu Strukturbäumen - maschinenlesbar sind.

3.1.2. Regelwerk (Generative Phonologie)

Das Steuerwerk bildet die parole, das Eingabetextkorpus, auf das Regelwerk, die generative Phonologie, ab. Das Regelwerk besteht aus Lexikon und Regeln. Das Lexikon enthält die in den Strukturbeschreibungen notierten Bedeutungen der Morpheme, die vom Steuerwerk erlernten Morphe(me) der phonischen Ketten und Anmerkungen über eventuell vom Steuerwerk gewonnene Erkenntnis über nicht reguläres morphophonologisches Verhalten, Distribution, Suppletivformen, Amalgamierung, Einbindung in den Regelapparat und Akzentuierung (s. Ausdruck der Lexika in Kap. 4).

Im Regelwerk gibt es zwei Regeltypen: Akzentuierungsregeln und Prozeßregeln. Die Notation der Regeln ist die der generativen Phonologie. Die Regeln geben auch Auskunft über positive oder negative Ausnahmen, Regelordnung, Suppletivformen und andere Spezialfälle.

3.1.3. Kommunikativer Erfolg

Kommunikativer Erfolg ist unter Annahme der in Kap. 2 und 3.1. gemachten Bedingungen die Fähigkeit von Steuer- und Regelwerk, Regeln aus dem Eingabetextkorpus - der parole - zu gewinnen und mit deren Hilfe auch wieder eventuell gleiche oder ähnliche - und somit wiederum analysierbare - parole zu erzeugen[1]. Regeln dieser Art sind adäquat. Neben Adäquatheit sind logische Richtigkeit und Einfachheit Kriterien der Bewertung. Die Bewertung des kommunikativen Erfolgs liegt beim Steuerwerk. Innerhalb der Bewertungskriterien besitzt logische Richtigkeit absolute Priorität.

3.1.4. Schrittweise Problemlösung

Dem Prinzip schrittweiser Problemlösung folgend wird die Konstruktion des Steuerwerks in eine Folge von Einzelproblemen zerlegt. Im ersten Schritt wird das Eingabetextkorpus - die parole - auf zwei Mengen ungerichteter Graphen[2] abgebildet. Die erste Graphenmenge heißt Lexikon. Die Elemente dieser Graphenmenge werden den types[3] der semantischen Einträge in den semantisch-syntaktischen Strukturbeschreibungen des Eingabetextkorpus emtnommen. Jedem type entspricht ein Graph. Die Knoten in den Graphen geben Auskunft über die Position (eventueller) (Allo)Morphe in den phonischen Ketten des Eingabetextkorpus, die Position in den semantisch-syntaktischen Strukturbeschreibungen und eventuelle Amalgamierung (vgl. Beispiele in Kap. 4 für ein Lexikon als Ergebnis des ersten Schritts der Problemlösung). Die Kanten der

---

1) Werden nur gleiche Ketten erzeugt, so ist das System hinsichtlich des Sprachwandels lediglich reaktiv, da es sich den in der parole vorgefundenen Bedingungen anpaßt, jedoch keine neuen in der parole setzt. Erzeugt es ähnliche Ketten, so nimmt es im sprachlichen Wandel eine aktive Rolle ein, ist jedoch andererseits einer größeren Gefahr kommunikativen Mißerfolgs ausgesetzt.

2) Zur Einführung in die Graphentheorie s. Brainerd (1971:38ff), Aho, Ullmann (1972:37ff). Ein Graph besteht aus Knoten und Kanten. Seien x und y Knoten (Elemente einer Knotenmenge), so ist das Paar (x,y) eine Kante des Graphen. Es kann sein, daß (x,y) eine Kante des Graphen ist, (y,x) jedoch nicht, z.B. in gerichteten Graphen (Beispiel: Phrasenstrukturbaum in der Syntaxtheorie). Ein Graph ist ungerichtet, wenn für alle Paare (x,y) gilt, daß auch (y,x) Kanten des Graphen sind.

3) zur type - token - Dichotomie Herdan (1966:332f).