Erschienen in: Texte gesprochener deutscher Standardsprache I. Erarbeitet im Institut für deutsche Sprache, Forschungsstelle Freiburg i. Br. - München: Hueber, 1971. S. 33-54.
(Heutiges Deutsch, II/1)

ZUR UMSCHRIFT GESPROCHENER HOCHSPRACHE

I. Vorbemerkungen

Gesprochene Sprache wird unter Voraussetzungen erzeugt, die mit denjenigen, unter denen geschriebene Sprache entsteht, nur sehr bedingt vergleichbar sind. Abgesehen davon, daß ein geschriebener Text mehrfach korrigiert werden kann, bevor er dem Adressaten geordnet zugänglich gemacht wird - im gesprochenen Text bleibt dagegen der Korrekturvorgang unverdeckt und bildet mit einen Teil der intendierten Information -, bestehen die größten Unterschiede wohl darin, daß 1. die gesprochene Sprache im Gegensatz zur geschriebenen nicht visuell sondern akustisch übermittelt wird, daß sie 2. wesentlich stärker situationsgebunden und mit außerlinguistischen Faktoren verknüpft ist als die geschriebene Sprache, und daß sie 3. primär dialogisch abläuft, während die geschriebene Sprache primär monologischen Charakter trägt. Diese Unterschiede werfen Probleme auf, die bei dem Versuch, gesprochene Sprache in schriftlicher Form festzuhalten, berücksichtigt werden müssen.

Wie im vorangegangenen Beitrag schon erwähnt, wird das Material auf Tonbänder gespeichert, d. h. nur die Sprache wird konserviert. Viele außersprachliche Elemente, wie z. B. die Situation, in der die Sprecher sich befinden und die in das Gespräch einbezogen sein kann, oder die Gestik und das Minenspiel der Sprecher untereinander, gehen dabei verloren, sofern nicht entsprechende Zusatzprotokolle geführt werden.

Dazu kommen außerdem rein aufnahmetechnisch bedingte Faktoren, die den konservierten Textablauf dem Original gegenüber verändern können: Ein Nebengeräusch oder simultanes Sprechen können einzelne Sequenzen im Text für den Hörer bzw. den Transkribenten unverständlich machen; daraus folgt jedoch nicht, daß die Kommunikation zwischen den einzelnen

Sprechern ebenfalls gestört gewesen sein muß. Wenn eine größere Anzahl von Sprechern an einem Gespräch oder an einer Diskussion beteiligt sind, und die einzelnen Sprecher jeweils nur kürzere Sequenzen sprechen, kann es vorkommen, daß allein aufgrund der Bandaufnahme eine eindeutige Identifizierung der Sprecher erschwert oder gar unmöglich ist.

Mit den heute zur Verfügung stehenden technischen Mitteln ist kein genaues Abbild einer Gesprächssituation herzustellen. Deshalb muß man sich der beim Aufnahmeverfahren eintretenden Störungen und Veränderungen bewußt sein und sie in den Textanalysen berücksichtigen.

Aus dieser Erkenntnis heraus wurde sowohl die Form der in den vorliegenden Texten angewandten Umschrift als auch das dabei benutzte Segmentierungsverfahren ausgerichtet auf die Fragestellungen, die an das Korpus herangetragen werden sollen. In erster Linie sollen an größeren Textmengen Fragen untersucht werden, die die Syntax und das Lexikon der gesprochenen Sprache betreffen. Da das Korpus auf datenverarbeitenden Anlagen gespeichert werden soll, müssen Voraussetzungen für eine Abrufbarkeit der Daten gegeben sein.

Allein die Möglichkeit, z.B. eine Worthäufigkeitsliste maschinell zu erstellen, ist ohne größere Schwierigkeiten nur möglich, wenn man auf eine phonetische Schreibweise verzichtet. Deshalb wurde die herkömmliche Rechtschreibung gewählt. Substantive sind durch Großschreibung indiziert, damit Substantivlisten erstellt werden können. Aus diesem Grund werden Satzanfänge - entgegen der Rechtschreibung - klein geschrieben, es sei denn ein Satz beginnt mit einem Substantiv. In einer geschlossenen phonetischen Liste werden einige von der Rechtschreibung abweichende phonetische Schreibweisen zugelassen. So kann das End-e schwinden, ebenso sind einige in der gesprochenen Spra-

che übliche Kontraktionen zugelassen. Mit dieser Regelung soll lediglich auf Ausspracheeigenschaften der Sprecher in diesem Bereich hingewiesen werden können.

Wie gesagt, versucht die Forschungsstelle zunächst Probleme zur Syntax der gesprochenen Sprache zu lösen. Mit aus diesem Grunde schien es angebracht, von einem syntaktischen Segmentierungsverfahren auszugehen und ein solches zu erarbeiten. Die Intonation der Texte wird unter besonderer Berücksichtigung der Kadenzen in einer Art Parallelcodierung unter Leitung von Professor Winkler an der Universität Marburg registriert[1]. Auf diese Weise können syntaktische Regularitäten mit der Intonation konfrontiert werden und umgekehrt.

Die Texte werden mit beiden Codierungen und datenverarbeitenden Anlagen am Institut für deutsche Sprache in Mannheim gespeichert und sollen für wissenschaftliche Untersuchungen verfügbar sein.

Auf den folgenden Seiten sind die Transkriptionsvorschriften - der Übersichtlichkeit wegen leicht gekürzt - aufgeführt wie sie die Forschungsstelle Freiburg[2] des Instituts für deutsche Sprache zur Textsegmentierung erarbeitet hat und wie sie auch in den hier vorliegenden Texten angewendet wurden. Daran schließt sich ein Kurzkommentar zu den einzelnen Zeichen an.

II. Transkriptionsvorschriften

1. Syntaktische Zeichen

[.] (Punkt) kennzeichnet das Ende eines Hauptsatzes bzw. eines Satzgefüges.

Als Hauptsatz bzw. Satzgefüge gelten Wortsequenzen, die von einem finiten Verb regiert werden; Subjekt ist in der Regel erforderlich.

Sondervorschriften:

1. Ein Imperativ gilt als Hauptsatz; dazu gehören auch Infinitive mit imperativer Bedeutung.
2. Subjektlose Passiv-Konstruktionen gelten als Hauptsatz.
3. Ist ein Satzglied des Hauptsatzes durch einen Nebensatz ersetzt, gilt der Hauptsatz als vollständig.
4. Wendungen wie: mich friert; mir ist kalt; mir war nicht geheuer zumute;
   gelten als Hauptsatz.

Syntaktische Bemerkungen:
Wenn bei zwei Hauptsätzen, die beide notwendig einen abhängigen Hauptsatz bzw. Nebensatz brauchen, nur ein abhängiger Hauptsatz bzw. Nebensatz steht, werden die beiden Hauptsätze wie ein Hauptsatz behandelt.

[.+ +.] kennzeichnet einen eingeschobenen Hauptsatz.
Ein in einen abhängigen Hauptsatz oder einen Nebensatz eingeschobener Hauptsatz wird durch [.+ +.] eingegrenzt.

Am Ende des gesamten Satzgefüges steht außerdem ein [.] (Punkt).

[s+ +s] grenzt einen von einem übergeordneten Hauptsatz abhängigen Hauptsatz ein.
Sind abhängige Hauptsätze oder Nebensätze eindeutig von einem anderen abhängigen Hauptsatz abhängig, werden sie mit den entsprechenden Segmentierungszeichen in die [s+ +s]-Sequenz mit hineingenommen.

Für den abhängigen Hauptsatz gilt die Hauptsatz-Definition mit den folgenden Einschränkungen:

Kriterien der Abhängigkeit:
Der abhängige Hauptsatz steht anstelle eines Satzgliedes des übergeordneten Satzes und ist erkennbar

1. an der Ersetzbarkeit des abhängigen Hauptsatzes durch einen mit daß eingeleiteten Nebensatz,
2. an Inversion in zwei aufeinanderfolgenden Hauptsätzen, wobei der eine durch einen Nebensatz mit der Konjunktion wenn ersetzt werden kann,
3. am Fragepronomen bzw. Frageadverb, wenn es nicht mit Verbendstellung gekoppelt ist,
4. an der Ersetzbarkeit des abhängigen Hauptsatzes durch einen Nebensatz mit der Einleitung ob oder als ob.

[,+ +,] schließt einen Nebensatz ein.

Kennzeichen des Nebensatzes ist das

Einleitewort, dem eine von einem finiten Verb regierte Wortsequenz folgt.

Einleitewörter sind: a) unterordnende Konjunktionen
b) Relativpronomen
c) Interrogativpronomen
d) Interrogativadverbien

Bei b, c und d ist Nicht-Verbzweitstellung obligatorisch, bei a ist sie optionell.

Nebensätze können ineinander geschoben sein. Der Abhängigkeitsgrad der Nebensätze untereinander wird nicht indiziert.

[i+ +i] Jeder Infinitiv mit zu wird mit den ihm zugehörigen Erweiterungen in [i+ +i] eingeschlossen.

Ausnahmen: Die Infinitive nach den Verben haben; pflegen; sein; scheinen; vermögen werden nicht in [i+ +i] gesetzt.

[n+ +n] schließt bei diskontinuierlicher Verschachtelung zweier Nebensätze alle Wortfolgen des zweiten Nebensatzes ein.

Falls es sich um einen diskontinuierlichen erweiterten Infinitiv handelt, steht [n+ +n] anstelle von [i+ +i].

[/] Der Schrägstrich steht am Anfang eines Satzes, in dem eine oder mehrere Wortsequenzen, die syntaktisch nicht angeschlossen werden können, vorkommen, sowie bei normwidrigen Kongruenzen und Rektionen.

[( )] schließt eine Paranthese ein. Paranthesen sind Einschübe in einen Satz,

die syntaktisch nicht angeschlossen sind, aber im semantischen Zusammenhang mit dem Satz stehen.

Parenthesen sind: a) Interjektionen
b) der Anredenominativ
c) ein eingeschobener unverbundener Hauptsatz
d) ein eingeschobenes unverbundenes satzwertiges Partizip
e) alle Affirmations- und Modalpartikel.

2. Die Syntax nicht betreffende Zeichen

a) den Textablauf betreffende Zeichen:

[?] Das Fragezeichen kennzeichnet das Ende einer Frage.
Fragen sind erkennbar an:
einem Fragepronomen und/oder der Wortstellung und/oder der Intonation.

Das Ende einer Frage muß nicht mit dem Satzende identisch sein.
Das [?] steht nicht bei sogenannten uneingeleiteten Konditionalsätzen.

[+g+] steht für lexikalisch nicht belegbare Artikulationen und für Wortfragmente.

Stehen mehrere dieser Erscheinungen bei einem Sprecher hintereinander, dann wird nur ein [+g+] für alle gesetzt.

[:] kennzeichnet eine Emphase. Als Emphase gelten nur die für den jeweiligen Sprecher extremen Akzentuierungen.

[+p+] zeigt eine Sprechpause bezogen auf die durchschnittliche Sprechgeschwindigkeit des Sprechers an.

b) das Lexikon betreffende Zeichen

[z+ +z] schließt ein Zitat und/oder einen Eigennamen ein.

Als Zitat gilt:

1. Jede Information, die als von fremder Seite stammend charakterisiert ist und die Form einer direkten Rede hat,
2. ein Eigenzitat aus einer geschriebenen Quelle,
3. ein Bibelvers, ein Literaturzitat, ein Sprichwort und "Geflügelte Worte".

Ausnahme: Selbstzitate aus mündlichen Quellen.

Eigennamen sind

Identifizierungen durch Herausheben eines Dinges unter gleichartigen; d.h. Namen sind dadurch gekennzeichnet, daß sie nicht mit unbestimmtem Artikel verbunden werden können (sonst werden sie zu Gattungsbezeichnungen). Entscheidend dabei ist das Fehlen jeglicher Differenzierung mit Hilfe des bestimmten Artikels.

[f+ +f] kennzeichnet Wörter, Sequenzen und Sätze in fremder Sprache oder im Dialekt.

Fremd- bzw. Dialektwörter sind:

a) solche fremdsprachliche und dialektale Wörter, die vom Sprecher selbst als Fremdwörter bzw. Dialektwörter apostrophiert werden.

b) Fremdwörter, die nicht im Großen Duden (Aufl. 1967) aufgeführt werden.

c) Dialektwörter, die nicht im Großen Duden (Aufl. 1967) aufgeführt oder solche, die darin als mundartlich gekennzeichnet sind.

d) formelhafte fremdsprachliche Wendungen, auch wenn sie im Duden verzeichnet sind.

[-] Der [-] (Bindestrich) steht

a) in flektierten Zusammensetzungen, die auf Namen zurückzuführen sind,

b) bei Dezimalzahlen und zerlegbaren Zahlen

c) bei zusammengesetzten fremdsprachlichen Wörtern,

d) bei zwei oder mehreren Komposita als Ergänzungsbindestrich für den allen Komposita gemeinsamen Teil der Wortbildung[3].

c) den Text oder die Transkription kommentierende Zeichen

[uuuuuuxxxxxx] leitet die Kennzeichnung der Sprecher ein, danach folgt die Namenssigle [aa; ab; ac; ... ,]

namentlich nicht identifizierbare männliche Sprecher werden mit [nm], weibliche mit [nw] gekennzeichnet.

[k+ +k] kennzeichnet simultan gesprochene Textstrecken; dabei wird nur die Sequenz jedes Sprechers in [k+ +k] eingeschlossen, die innerhalb des simultanen Bereichs liegt.

[...] (drei Punkte) kennzeichnen jede dem Transkribenten unverständliche Stelle im Text, seien es Wörter oder Wortsequenzen.

[ttttt] Hält der Transkribent eine erklärende Anmerkung zu einer Textstelle für angebracht, kann er sie nach [ttttt] im Klartext formulieren.

III. Kommentar zu den Transkriptionsvorschriften

Die Segmentierungsvorschriften, nach denen der Transkribent bei der Umschrift gesprochener Texte verfährt, können nicht beliebig differenziert sein; denn ein differenziertes Verfahren würde unverhältnismäßig viele Korrekturgänge erfordern, damit eine zu hohe Fehlerquote die Verwertbarkeit der Umschrift nicht beeinträchtigt.

Aus diesem Grunde wurde ein Verfahren gewählt, das auf wenige Zeichen beschränkt und übersichtlich ist. Außerdem hängt die Brauchbarkeit eines Segmentierungsverfahrens entscheidend von dem Grad der Formalisierung der Vorschriften ab; intuitives Vorgehen soll möglichst ausgeschlossen sein. Die am Text getroffenen Entscheidungen müssen von jedermann in gleicher Weise vollzogen und kontrolliert werden können. Obwohl diese Gesichtspunkte leitend waren, ließ es sich in einigen Fällen nicht vermeiden, auch nicht hinreichend formalisierbare Aspekte mit einzubeziehen. Das trifft besonders für die nicht-syntaktischen, den Textablauf betreffenden Zeichen (der Gruppe 2a der Transkriptionsvorschriften) zu.

Die Regeln, nach denen Programme in einer datenverarbeitenden Anlage operieren, haben die Organisation des Zeichensystems bestimmt. Jede Sequenz beginnt mit einem zusammengesetzten Zeichen, das am Ende der Sequenz in umgekehrter Zusammensetzung wiederkehrt. Dadurch ist gewährleistet, daß Sequenzen mit Hilfe von einfachen Suchprogrammen isoliert werden können[4].

Nur [.] (der Punkt) und [/] (der Schrägstrich) erscheinen nicht paarig, weil sie sich auf die größte vorkommende Sequenz, das Satzgefüge bzw. auf eine syntaktisch nicht anschließbare Sequenz, die innerhalb eines Satzgefüges steht, beziehen. Alle anderen nicht paarigen Zeichen kennzeichnen einzelne nicht-grammatische Ereignisse innerhalb eines Satzes.

Im folgenden Kommentar zu den einzelnen Zeichen wird neben einer Begründung der jeweiligen Segmentierungen auch eine Interpretationsanregung gegeben, soweit dies nach dem gegenwärtigen Stand der Untersuchungen vertretbar erscheint.

1. Syntaktische Zeichen

zu: [.]

Die regelmäßige Wiederkehr von Subjekt und Prädikat (Nominal- und Verbalphrase) ist auch im gesprochenen Deutsch in jedem Redeablauf zu beobachten. Sie kann als ein sicheres formales Kriterium für ein Segmentierungsverfahren an Texten eingesetzt werden, sofern primär syntaktische Untersuchungen an dem zusammenzustellenden Korpus durchgeführt werden sollen. Dementsprechend wurde das Satzgefüge als größte vorkommende Sequenz definiert. Es reicht von Punkt zu Punkt, zwischen denen u. a. jeweils nur e i n e Sequenz steht, die der Definition 'Hauptsatz' entspricht. Nur einmal wird diese Regel durchbrochen, nämlich dann, wenn es sich um zwei Hauptsätze handelt, denen gemeinsam e i n Nebensatz zugeordnet ist. In dem Fall kreuzen sich kontextuelle und syntaktische Beziehungen. Hier wurde dem übergreifenden Prinzip der Vorzug gegeben.

(Beispiel: er erwähnte nicht nur sondern er betonte auch ,+ daß die Angelegenheit entschieden werden müsse +, .)

Die unter den Sondervorschriften formulierten Fälle betreffen häufig auftretende Erscheinungen in den verschiedensten Texten, die als satzwertig angesehen werden müssen, obwohl sie kein Subjekt haben. Allerdings unterliegt der Ausfall des Subjekts bestimmten beschreibbaren Bedingungen.

(Beispiele:

zu 1) geht; nicht rauchen;
zu 2) darf geraucht werden?;
zu 3) ,+ daß er nicht raucht +, ist neu .)

zu: [.+ +.]

Durch das Eliminieren aller einem Hauptsatz zugeordneten Sequenzen, die ja durch paarige Zeichen eingeschlossen sind, kann ein Hauptsatz durch ein Suchprogramm isoliert werden. Er braucht deshalb nicht gesondert indiziert zu werden. Es sei denn, er ist in eine Sequenz eingebettet, dann muß er durch ein paariges Zeichen gesondert abgegrenzt werden, damit er aus der Sequenz, in die er eingebettet ist, isoliert werden kann. Demnach ist [.+ +.] eine Variante des [.] (Punktes).

(Beispiel:
s+ das wäre .+ hat er gesagt +. längst erledigt +s .)

Am Ende des Satzgefüges muß wieder ein [.] (Punkt) stehen, damit die den Hauptsatz umgebende Sequenz gegen den nächsten Hauptsatz abgegrenzt ist.

zu: [s+ +s]

Dieses Zeichen grenzt eine syntaktische Sequenz ein, die die Kriterien eines Hauptsatzes erfüllt, aber gleichzeitig und zusätzlich in enger kontextueller Beziehung zu einem anderen Hauptsatz steht. Die Abhängigkeit ist jedoch nicht formal, sondern inhaltlich bestimmt. Die aufgeführten Kriterien bestehen in einer Art Ersatzprobe, mit deren Hilfe der abhängige Hauptsatz in einen Nebensatz umgeformt werden kann. Die Frage, wieweit der übergeordnete Hauptsatz diese kontextuelle Abhängigkeit mit konstituiert, ist in der Transkriptionsvorschrift nicht berührt. Vor allem drei sprachliche Erscheinungen, die innerhalb dieses Bereichs auftreten, sprechen dafür, hier neben der syntaktischen auch die kontextuelle Ebene zu indizieren:

1. die Frage der Funktion des Verbs im übergeordneten Hauptsatz[5],
2. das Problem der direkten bzw. indirekten Rede in der gesprochenen Sprache und
3. damit verbunden der Gebrauch des Konjunktivs.

Ob Nebensatz oder abhängiger Hauptsatz als Konstruktion gewählt wird, hängt vom Texttyp ab. Wenn auch keine feste Verteilung festzustellen ist, so kann doch gesagt werden, daß bei der Verschriftlichung von gesprochenen Texten häufig die abhängige Hauptsatzkonstruktion in eine solche mit Nebensatz umgewandelt wird[6].

zu: [, + +,]

Eine formale Beschreibung des Nebensatzes allein aufgrund von syntaktischen Kriterien und nach Wortklassen wurde hier durchgeführt. Bei den unterordnenden Konjunktionen, die die Sequenz als Nebensatz definieren, konnte auf eine obligatorische Verbendstellung verzichtet werden; denn die Konjunktion selbst definiert die Abhängigkeit und ist folglich ein hinreichendes Kriterium. Übrigens wird die Verbendstellung bei einigen dieser unterordnenden Konjunktionen sowohl in der gesprochenen als auch in der geschriebenen Sprache nicht konsequent eingehalten. Beim Relativpronomen gibt es außer der mit ihm verbundenen Verbendstellung kein formales Kriterium zur Identifizierung des Nebensatzes. Das Relativpronomen allein ist nicht hinreichend, denn ein mit Verbzweitstellung gekoppeltes Relativpronomen kann ebenso als ein mit Demonstrativpronomen eingeleiteter Hauptsatz angesehen werden, zumindest kann das in der gesprochenen Sprache unter bestimmten Bedingungen zutreffen. Das "Relativpronomen" welcher, welche, welches wird, wenn es mit Verbzweitstellung verbunden ist, als Fragepronomen aufgefaßt.

Beispiel:
zu 1: sehen sie den Mann , + der einen Hut trägt + , ?.
dagegen: sehen sie den Mann ?. der trägt einen Hut .
zu 2: sehen sie den Mann , + welcher einen Hut trägt + , ?.
dagegen: *sehen sie den Mann . welcher trägt einen Hut ?.

zu: [i+ +i]

Der Infinitiv mit zu mit seinen ihm zugehörigen Erweiterungen wurde mit einbezogen, weil er einmal eine nebensatzähnliche Ergänzung darstellt, der Nebensatzdefinition aber nicht entspricht; und zum anderen eine Untersuchung der Frage ermöglicht, wie der Infinitiv heute gebraucht wird. Die genannten Ausnahmen haben obligatorisch den Infinitiv mit zu und werden deshalb nicht indiziert[7].

zu: [n+ +n]

Die diskontinuierliche Verschachtelung von Nebensätzen bzw. Infinitiven untereinander wurde theoretisch als möglich betrachtet und deshalb ein entsprechendes Zeichen vorgesehen. Bisher ist dieses Ereignis nur selten belegt (z.B. in dem Satz: aber z+ Libuda +z kann nur schlecht passen ,+ weil auch sofort +g+ die Briten wieder dazwischen sind und von der rechten Seite n+ den eigenen Angriff +n versuchen n+ wieder aufzuziehen +n +, . Die Auflösung der diskontinuierlichen Verschachtelung würde als Infinitiv [i+ +i] indiziert werden: z+ Libuda +z kann nur schlecht passen ,+ weil auch sofort +g+ die Briten wieder dazwischen sind und von der rechten Seite versuchen i+ den eigenen Angriff wieder aufzuziehen +i +, .). Das Zeichen muß nach Definition mindestens zweimal paarig auftreten. Es hat für [,+ +,] bzw. [i+ +i] die Funktion, die [.+ +.] für [.] (den Punkt) hat, nämlich eingebettete bzw. ineinander verzahnte Sequenzen gegeneinander abzugrenzen.

zu: [/]

Die syntaktisch nicht anschließbaren Sequenzen stehen an irgendeiner Stelle im Satzgefüge. Es wurde darauf verzichtet, sie durch ein paariges Zeichen einzugrenzen, weil sie in den meisten Fällen den Satz, in dem sie stehen, als Kontext zu ihrer Entschlüsselung benötigen. In einer großen Zahl der Fälle handelt es sich um Wortwiederho-

lungen, abgebrochene Satzplanungen und andere psychologisch bedingte Faktoren. Wollte man die verschiedenen unter diesem Zeichen vorkommenden Ereignisse systematisieren, müßte man eine Parallelcodierung vornehmen, die von verschiedenen Fragestellungen auszugehen hätte. Ein solches Unternehmen dürfte u.a. besonders für die Sprachpsychologie von Interesse sein.

zu: [( )]

Die Parenthese hat einmal die Funktion, bestimmte Partikel wie Interjektionen, Anreden, Affirmations-, Negations- und Modalpartikel zu erfassen. Darüber hinaus werden mit ihr Verschachtelungen bewältigt, die mit [.+ +.] allein nicht geleistet werden können. Folgendes Beispiel zeigt, daß ein eingeschobener unverbundener Hauptsatz auf diese Weise eingegrenzt werden kann: denn (sehen sie) ich habe keine guten Erfahrungen gemacht mit Menschen ,+ die man zu früh in einen Beruf zwingt +, . Die Entscheidung, ob und wo man eine Parenthese zu setzen hat, hängt manchmal allein von dem nur bedingt objektiv erfaßbaren Faktor 'Intonation' ab, obwohl die syntaktischen Kriterien vorliegen. In dem genannten Beispiel könnte aufgrund des syntaktischen Befundes folgendermaßen transkribiert werden: denn sehen sie s+ ich habe keine guten Erfahrungen gemacht mit Menschen +s , + die man zu früh in einen Beruf zwingt +, . Der Transkribent hat aus seiner eigenen Kompetenz heraus aufgrund der weniger objektiven Kriterien wie Sprechpause, Lautstärke oder Satzintonation zu entscheiden, welche der beiden Segmentierungsmöglichkeiten die vom Sprecher intendierte Struktur der Sequenz genauer festhält.

2. Die Syntax nicht betreffende Zeichen

a) den Textablauf betreffende Zeichen:

zu: [?]

Da jede Frage - sei sie nun allein durch die Wortstellung oder allein durch die Intonation oder durch beide Kriterien gemeinsam gekennzeichnet - mit einem Fragezeichen versehen wird, ist es möglich, alle Sätze, bei denen Inversion vorliegt, die jedoch keine Fragesätze sind, in der Transkription als Aussagesätze zu erkennen. Ebenso können alle Fragen, die nicht syntaktisch, sondern nur intonatorisch erkennbar sind, entsprechend identifiziert werden, ohne daß der über die Satzgrenze hinausgehende Kontext abgefragt werden müßte.

zu: [+g+]

Nur in der gesprochenen Sprache werden Ereignisse im Redeablauf produziert, die dem Lexikon nicht zugeordnet werden können, aber dennoch eine Funktion innerhalb der Rede haben können. So haben z.B. einige lexikalisch nicht belegbare Lautfolgen eine feste, von der Sprachgemeinschaft anerkannte, Funktion als Zeichen (besonders als Affirmation und Negation), wenn sie gezielt eingesetzt werden. Die gleichen oder ähnlichen Lautfolgen werden aber ebenso aus Routine mehr oder weniger gebraucht. Eine konsequente Scheidung nach bedeutungslosen und bedeutungstragenden Partikeln ist während des Transkriptionsvorganges nicht zu leisten. Deshalb wird zwar das Ereignis selbst notiert, eine Interpretation aber noch nicht gegeben. Tritt eine Häufung solcher Lautfolgen (z.B. mehrfaches murmelndes Zustimmen oder mehrfaches Hüsteln) bei einem Sprecher hintereinander auf, so wird nur ein [+g+] gesetzt.

zu: [+p+] und [:]

Sowohl bei der Sprechpause als auch bei der Emphase wurde kein physikalisches Meßverfahren zugrunde gelegt, weil ein solches Verfahren wenig über den individuell ablaufenden Redefluß eines Sprechers aussagen würde. Die hier gegebene Definition soll in der Anwendung nicht mehr leisten, als eine Indizierung der für den Sprecher extremen Haltungen im Redeablauf. Nicht die Abweichung von einer vorgegebenen Norm, sondern die auffällige Abweichung vom Normalverhalten des einzelnen Sprechers im zur Transkription vorliegenden Text ist entscheidend.

b) das Lexikon betreffende Zeichen

zu: [z+ +z]

Zitate und Eigennamen bilden eine besondere Gruppe in den Texten, weil beide für den Sprecher fest vorgegebene situationsgebundene Muster darstellen. Bei Zitaten muß neben der Lexik zusätzlich die syntaktische Komponente berücksichtigt werden; denn eine zitierte Sequenz hat nicht unbedingt eine dem Sprecher geläufige und von ihm regelmäßig gebrauchte syntaktische Struktur. Zitate und Eigennamen müssen deshalb von der übrigen Lexik und Syntax eines Sprechers getrennt werden können. Bei der Definition der Eigennamen wurde im wesentlichen H. Vater gefolgt[8].

zu: [f+ +f]

Auf eine formale Abgrenzung der fremdsprachlichen und dialektalen Wörter vom hochdeutschen Wortschatz wurde weitgehend verzichtet. Mit der Abgrenzung gegenüber dem Duden-Wörterbuch werden besonders die neu in der gesprochenen Sprache auftretenden Wörter gekenn-

zeichnet und jüngste Übernahmen und Einflüsse aus anderen Sprachschichten berücksichtigt.

zu: [-]

Flektierte Zusammensetzungen wie: Freier - Demokrat; oder Dezimalzahlen und zerlegte Zahlen wie: null-Komma-fünf, vier-vier-eins-sieben-sieben; oder Komposita wie: Ost- und Westwind; oder zusammengesetzte fremdsprachliche Wörter wie: teach-in, maitresse-de-maison, die nach den Rechtschreiberegeln nicht zusammengeschrieben werden, aber als ein Wort zu verstehen sind, müssen aus Computer-technischen Gründen mit Bindestrich versehen werden, weil andernfalls jedes Element der Fügung als selbständiges Wort gewertet werden würde.

c) den Text oder die Transkription kommentierende Zeichen

zu: [k+ +k]

In der gesprochenen Sprache können mehrere Sprecher zu gleicher Zeit sprechen, ohne daß dabei das verstehende Aufnehmen der einzelnen Sequenzen bei den Sprechern untereinander gestört sein müßte. In einer Diskussion z.B. kann das Thema durch fortgesetztes Ins-Wort-Fallen gesteuert werden. In der Umschrift jedoch ist dieser zwischen den Sprechern simultan ablaufende Vorgang nur in einem linearen Nacheinander darzustellen. Daß dieses Nacheinander in der Realität gleichzeitig abläuft, wird durch die Eingrenzung der simultan gesprochenen Sequenz in [k+ +k] angezeigt.

zu: [...]

Die Verständlichkeit eines Textes hängt einerseits von der Qualität der Bandaufnahme ab und zum andern von den Geräuscheinflüssen, denen die Gesprächspartner während eines Gesprächs ausgesetzt sind. Im letzten Fall kann ein Neben-Geräusch auch die Kommunikation zwischen den Sprechern gestört haben. Der Transkribent, der in beiden Fällen den Text akustisch nicht versteht, muß dieses Nicht-Verstehen an der entsprechenden Textstelle mit [...] anzeigen. Aus dem Zeichen kann demnach nicht geschlossen werden, ob die Kommunikation zwischen den Sprechern gestört war oder nicht.

zu: [ttttt]

Zu einer Aufnahme können nach diesem Zeichen mehrere informative, zum Textverständnis vielleicht notwendige Anmerkungen gegeben werden. Hier seien nur einige Ursachen genannt, die den Wunsch nach einer Anmerkung zur Textstelle auslösen können: Eine Aufnahmeunterbrechung, ein nicht ganz ausgesprochenes aber für den Textverlauf wichtiges Wort, ein Versprecher. Auch Anmerkungen zur Sprechsituation, zur Gestik u.ä. können auf diese Weise der entsprechenden Textstelle zugeordnet werden. Sie werden im Klartext formuliert und können beliebig lang sein.

Karl-Heinz Bausch

Anmerkungen

1 Sie ist in die in dieser Veröffentlichung vorliegenden Texte nicht eingetragen.

2 Diese Arbeit wurde schon begonnen, als sich die Arbeitsstelle noch in Kiel befand.

3 Im vorliegenden Textbuch ist aus technischen Gründen Worttrennung am Zeilenende ebenfalls mit Bindestrich durchgeführt. Dieses Zeichen darf nicht mit dem hier definierten verwechselt werden.

4 Das hier beschriebene Verfahren, bei dem die Maschine den Text nach Merkmalen absucht, die in Zeichen realisiert sind, bezeichnen wir als ein Vorgehen nach expliziten Kriterien. Wird beim Aussortieren auf im voraus gespeicherte grammatische Information rekurriert, sind die Kriterien implizit. Dieses Verfahren wird z. B. unter der Leitung von Prof. Eggers in Saarbrücken angewendet und ausführlich vorgestellt in: Elektronische Syntaxanalyse der deutschen Gegenwartssprache, ein Bericht von Hans Eggers und Mitarbeitern, Tübingen 1969.

5 Siehe dazu S. 60 die Bemerkungen über [s+ +s] im Kommentar des 'Schulklassengesprächs mit Günter Grass'.

6 Ute Elmauer (Quantitative Analyse und Interpretation des Verhältnisses von gesprochenen Texten und deren verschriftlichter Form, masch. Magisterarbeit, Freiburg i. Brsg. 1969) erwähnt neben der Umformung abhängier Sätze in Nebensätze mit der Einleitung daß zwei weitere Möglichkeiten:
"Umformung des abhängigen uneingeleiteten Satzes zu einem Hauptsatz durch Einbeziehung des Trägersatzes als Nominalgruppe in den Hauptsatz". (S. 51)
und
"Umformung des abhängigen uneingeleiteten Satzes

zu einem Hauptsatz durch Tilgung des Trägersatzes". (S. 50).

7 Vgl. z. B. Duden, Grammatik der deutschen Gegenwartssprache (= Der Große Duden, Band 4) bearbeitet von Paul Grebe, Mannheim 1966, unter § 5915.

8 Vater, Heinz, Eigennamen und Gattungsbezeichnungen, Versuch einer Abgrenzung, in: Muttersprache 75, 1965, S. 207 - 213.