

In: Beiträge zur Linguistik und Informationsverarbeitung Heft 20, 1970. München: Oldenbourg-Verlag, S. 7-31

Dokumentation eines Programms zur Analyse russischer Sätze

Von J. W. JANDA, A. ROTHKEGEL, H. ZIMMERMANN, Saarbrücken
TEIL 2

Pass 2

In Durchlauf 2 wird der Satz in einzelne Bearbeitungseinheiten eingeteilt. Ihre Grenzen entsprechen etwa den Subsatzgrenzen. Die Abgrenzungen werden jeweils durch Kommas und die Konjunktionen I und ILI bestimmt, wenn eine Reihe zusätzlicher Voraussetzungen erfüllt ist. Die Bearbeitung erfolgt von links nach rechts. Von dem finiten Verb ausgehend wird das Objekt - es muss ein Nomen sein - gesucht, das am weitesten nach rechts vom Verb entfernt ist. Hierbei werden Aufzählungen wie auch von diesem Objekt abhängige Nominalgruppen, z. B. Genitivattribute, die ebenfalls ein Substantiv enthalten müssen, berücksichtigt. Von dem am weitesten nach rechts stehenden Nomen wird das nächste freie Komma, I oder ILI gesucht und als Grenze bestimmt. Der Grenzanzeiger (Byte 51) wird entsprechend besetzt. Dieser Programmablauf wird wiederholt, bis der Satz vollständig, d. h. alle Wörter bearbeitet sind. Falls kein weiteres finites Verb vorhanden ist und keine Nominalgruppen mehr gefunden werden, die Objekte zu dem betreffenden Verb sein können, erfolgt der Rücksprung ins Hauptprogramm.

Pass 3

Teil I: Allgemeine Beschreibung

1. Subjekt- bzw. Prädikatbestimmung bei funktional mehrdeutigen Wörtern.

Im letzten ausgeführten Programmschritt des Analyseprogramms, Pass 3, werden Subjekt bzw. Prädikat innerhalb von komplizierteren Satzstrukturen (z. B. Verschachtelungen) bestimmt. In Pass 1 waren lediglich die eindeutigen Fälle behandelt worden, wobei der Begriff „eindeutige Fälle“ im Zusammenhang mit der Verfahrensweise in Pass 1 steht. Es handelt sich also um Subjekte oder Prädikate, die bei der Anwendung der Methode von Pass 1 eindeutig bestimmbar sind. Da dies bereits im Abschnitt 3, S. 28/29 (Heft 19; Teil 1) beschrieben wurde, ist hier nur summierend darauf einzugehen:

Als eindeutiges Subjekt werden in Pass 1 festgehalten:

1. Personalpronomen im Nominativ.
2. Substantive, die eindeutig im Nominativ stehen.
3. Relativpronomen im Nominativ.

Als Prädikat kommt in Pass 1 nur das finite Verb in Frage. Die bereits endgültigen Entscheidungen können jeweils durch Abfragen bei dem jeweiligen Wort getroffen werden. Es handelt sich also hier um formale Merkmale, die ohne den Kontext erkennbar sind. Dass diese Subjektbestimmung bzw. Prädikatbestimmung nicht ausreicht, liegt auf der Hand. So ist also auch in der Analyse vorgesehen, dass etwa Vertreter anderer Wortklassen⁶ wie auch Substantive, die im Kasus mehrdeutig sind, die Funktion eines Subjekts haben können⁷ und Partizipien oder Kurzformen die Funktion des Prädikats. Das Problem liegt darin, dass diese Wörter for-

mal nicht erkennen lassen, ob sie in der Funktion eines Subjekts oder in einer anderen Funktion stehen. Somit reicht auch die Abfrage formaler Kennzeichen wie z. B. die des Kasus nicht zur eindeutigen Bestimmung aus. Abfragen des Inventars und seiner Reihenfolge (wenn diese selbst Abhängigkeitsbeziehungen ausdrückt) werden auf dieser Stufe notwendig. Mit der veränderten Voraussetzung - funktionale Mehrdeutigkeit der Wörter - ändert sich konsequenterweise die Methode. War es in Pass 1 möglich, dass sich die Abfragen nur auf das gerade behandelte Wort bezogen, so erstrecken sich die Abfragen dann über alle Wörter des Satzes oder einer entsprechenden Einheit, wenn es um die Erkennung eines funktionalen Bezugs innerhalb des Satzes oder dieser Einheit geht, ohne dass dieser formal eindeutig erkennbar ist. Wir haben es also bei der Subjekt- bzw. Prädikatbestimmung von funktional mehrdeutigen Wörtern mit einer anderen Methode als in Pass 1 zu tun.

Der Wechsel in der Verfahrensweise ergibt sich auch aus der nunmehr etwas abgeänderten Zielsetzung. Sollen Funktionen erkannt werden, die sich auf den ganzen Satz beziehen, wie es bei Subjekt und Prädikat der Fall ist, und werden diese Funktionen von Wörtern getragen, die ihrer Form nach verschiedene Funktionen haben können - also funktional mehrdeutig sind -, so muss der Kontext die notwendigen Informationen liefern. Hierbei kann es sich um Informationen aus dem Wörterbuch wie auch um Ergebnisse der bis dahin erfolgten Analyse handeln. Da während des Programmablaufs nicht bekannt ist, wo die notwendigen Informationen erreichbar sind, müssen alle Wörter des Satzes auf bestimmte Merkmale hin abgefragt werden. Hierbei ist es unwahrscheinlich, dass *eine* Abfrage bereits ein endgültiges Ergebnis liefert, zumeist sind weitere Abfragen notwendig. Jede Abfrage leitet einen neuen Durchgang des Programms über alle Wörter des Satzes ein.

Da alle Wörter des Satzes bearbeitet werden müssen, ist es gleichgültig, ob sie von rechts nach links oder umgekehrt abgefragt werden. In Pass 3 wird der Satz von links nach rechts bearbeitet. Möglicherweise wurde dabei an die Stellung des Subjekts am Anfang des Satzes gedacht⁸. Nun zu den Abfragen selbst: Hier tritt das Problem der Abfragehierarchie (Reihenfolge) auf. Als Subjekt kommen generell in Frage: Substantive, Adjektive, Partizipien, Infinitive und Vertreter der Wortklasse Adverb vom Typ MNOGO.

Hierbei ist den Wortklassen „Substantiv“, „Adjektiv“ und „Partizip“ Priorität gegenüber den Wortklassen „Infinitiv“ und MNOGO gegeben. Für das Prädikat gilt: Finite Verben haben Priorität gegenüber Kurzformen, Adverbien und Adjektiven. Da jeweils die Bedingung erfüllt sein muss, dass das betreffende Wort noch nicht zugeordnet ist (vgl. dazu Zuordnungen in Pass 1), hängt das Ergebnis meist von den Ergebnissen der Zuordnungen ab. Konnte also in Pass 1 z. B. noch kein Subjekt bestimmt werden - da kein eindeutiges Subjekt vorhanden war - und liegt der Fall vor, dass ein Substantiv, Adjektiv, Partizip, Infinitiv oder MNOGO einem anderen Wort im Satz fälschlich nicht zugeordnet wurde, ist es sehr wahrscheinlich, dass auch die Subjektbestimmung im Ergebnis falsch ist.

Für das Prädikat gilt zweifellos die Priorität des finiten Verbs, da das finite Verb keine andere Funktion haben kann, also funktional eindeutig ist, während dies für Adjektiv und Kurzform nicht zutrifft. Die Verteilung der Prioritäten bei den Wortklassen, deren Elemente Subjekt sein können, ist problematisch.

Zwar gibt es Wortklassen, deren Elemente nicht die Funktion eines Subjekts ausüben können, ob man aber von einer Priorität unter den Wortklassen sprechen kann, deren Elemente als Subjekt fungieren können, ist zu bezweifeln.

Die Fragestellung als solche ist keine linguistische. Das Prinzip der Priorität, wie es hier gesehen ist, beruht auf der Häufigkeit des Vorkommens.

Zweckmäßig ist das Verfahren unter der Bedingung, dass in Pass 1 oder in einem zusätzlichen Pass alle Zuordnungsmöglichkeiten berücksichtigt und vermerkt werden. Wenn man sich darauf verlassen kann, dass für ein nicht zugeordnetes Wort alle Möglichkeiten - auch Korrekturmöglichkeiten - erschöpfend überprüft sind, wären wohl zufriedenstellende Ergebnisse zu erzielen. In diesem Fall beruhte ja letztlich die Erkennung der Fähigkeit, Subjekt zu sein, auf „subjektfreundlichen“ und „subjektfeindlichen“ Zuordnungen, d. h. auf grammatischen Regeln, die bereits in einem vorhergehenden Programmschritt angewendet worden sind. Die Subjektbestimmung wäre somit eine Entscheidung auf grammatischer Grundlage und nicht nach der Wahrscheinlichkeit. Da aber zum gegenwärtigen Stand des Analyseprogramms nicht alle grammatischen Beziehungen erschöpfend dargestellt werden (eine Reihe von syntaktischen Mehrdeutigkeiten wird gar nicht erst aufgelöst; vgl. Pass 1 der Satzanalyse), werden andererseits bei einer Subjektbestimmung auf Grund der Entscheidung nach Häufigkeit von vornherein Fehler in Kauf genommen.

Dieses Problem tritt an einer anderen Stelle - bei der Subjektbestimmung innerhalb von diskontinuierlichen Einheiten - nochmals auf (vgl. S. 12f.).

Bei einer erweiterten Analyse der Zuordnungsmöglichkeiten der Wörter untereinander könnte das hier verwendete Verfahren sowohl unter linguistischen Gesichtspunkten als auch im Hinblick auf die zu erwartenden Ergebnisse beibehalten werden. Es ginge dann zwar nicht mehr um die Abfrage einer bestimmten Wortklasse, die Priorität vor einer anderen hat, die dann erst im nächsten Durchgang behandelt würde, sondern die Abfragen würden sich auf bestimmte vorhandene Zuordnungsmöglichkeiten beziehen, innerhalb derer es eventuell Prioritäten geben kann. In diesem Fall wäre es zweckmäßig, für jede von ihnen einen getrennten Durchgang über den Satz zu machen.

Im übrigen ist darauf hinzuweisen, dass die Möglichkeit einer Aufzählung von Wörtern, die Subjekt oder Prädikat sein können, berücksichtigt ist. Diese werden als „multiple subject“ bzw. „multiple predicate“ bezeichnet⁹.

Auch ist es möglich, dass überhaupt keine Subjekt- bzw. Prädikatbestimmung erfolgt, wenn keine der vorgesehenen Bedingungen erfüllt ist.

2. Einteilung des Satzes in Einheiten

Da sich die Abfragen - wie im vorangehenden Abschnitt dargestellt - über den ganzen Satz, wenn es ein Einfachsatz ist, oder über eine entsprechende Einheit erstrecken, muss noch geklärt werden, welche Einheiten hier in Frage kommen. Es handelt sich dabei um Einheiten, die jeweils Subjekt und Prädikat enthalten können, im allgemeinen also um Hauptsätze und Nebensätze. Diese sind in einem vorherigen Programmabschnitt (vgl. Pass 2) bereits abgetrennt worden. Es sei daran erinnert, dass Komma und Konjunktionen, wenn sie nicht eine Aufzählung kennzeichnen, als Grenzen von Einheiten betrachtet werden und entsprechend gekennzeichnet sind. In Pass 3 wird noch eine zusätzliche Grenzmöglichkeit berücksichtigt: der Gedankenstrich, wenn er als Doppelzeichen¹⁰ auftritt.

Wenn die Einteilung in Einheiten nach den genannten Kriterien erfolgt, erhalten wir auch Einheiten, die möglicherweise kein Subjekt und/oder kein Prädikat enthalten, z. B. die Einheit vor und nach einem eingeschlossenen Nebensatz. Dazu ein deutsches Beispiel: Das Haus an

der Ecke, an dem gerade gebaut wird, gehört mir. Die erste Einheit „Das Haus an der Ecke“ enthält kein Prädikat, die dritte Einheit „gehört mir“ enthält kein Subjekt. Andererseits ergänzen sich die erste und die dritte Einheit.

Diese Fälle der Verschachtelung (Einbettung) werden in Pass 3 besonders berücksichtigt. Die Satzstruktur der entstandenen Einheiten wird als unbekannt vorausgesetzt, auch wenn bekannt ist, dass es sich bei einer Einheit um einen Relativsatz oder Adverbialsatz oder sonstigen Nebensatz handelt. Sobald mehr als zwei Einheiten im Satz festgestellt werden, wird die zweite Einheit automatisch als eine mögliche „eingebettete Satzeinheit“ im obigen Sinne betrachtet¹¹. Sie wird als „parenthetic clause“ bezeichnet, künftig als PC abgekürzt. Ob sie tatsächlich ein eingebetteter Satz ist, wird in einem speziellen Programm später festgestellt. Dies ist für die Subjekt- bzw. Prädikatbestimmung sehr wichtig. Wird ein Satz nämlich durch eine PC unterbrochen, muss sich die Subjekt- bzw. Prädikatbestimmung sowohl auf den Teil vor der PC als auch auf den Teil nach der PC erstrecken, d. h. bei den Abfragen müssen die Wörter einer PC übersprungen werden.

Bei der Feststellung, ob es sich um eine tatsächliche eingebettete Satzeinheit handelt oder nicht, gilt als Kriterium, dass sich die Einheit vor der PC und die Einheit nach der PC ergänzen: Die Einheit nach der PC muss entweder das Subjekt, das Prädikat oder ein Objekt zum Prädikat der Einheit vor der PC enthalten. Können Beziehungen dieser genannten Art zwischen den beiden diskontinuierlichen Einheiten nicht hergestellt werden, gelten sie als jeweils selbständige Einheiten. Im andern Fall, wenn es sich also herausgestellt hat, dass die erste Einheit durch die dritte ergänzt werden kann, werden diese beiden Einheiten zusammen als eine zu bearbeitende Einheit betrachtet, der eingeschobene Satz auch als solcher gekennzeichnet. Wie bereits erwähnt, wird seine eventuell vorhandene Nebensatzstruktur als Information nicht berücksichtigt. Zweckmäßiger wäre vielleicht, von der mittleren Einheit auszugehen. So bräuchte nur in den Fällen, in denen bestimmte - einbettungsfähige - Nebensätze oder der Gedankenstrich als Doppelzeichen auftreten, das spezielle Programm zur Erkennung der Ergänzungsmöglichkeiten von Einheiten durchlaufen zu werden. Beim jetzigen Stand des Programms wird dieser Programmteil und noch ein weiterer für jeden Satz, der aus mehr als zwei Einheiten besteht, durchlaufen. Wie leicht ersichtlich ist, muss dem Programm zur Erkennung der Ergänzungsmöglichkeiten einer Einheit der Versuch einer Subjekt- bzw. Prädikatbestimmung für die erste Einheit vorausgehen. Es muss ja schließlich bekannt sein, ob überhaupt ein Subjekt oder Prädikat oder auch Objekt fehlt, d. h. ob die Einheit ergänzungsfähig ist.

Hier ergibt sich wiederum - besonders für die Subjektbestimmung - eine Schwierigkeit durch die vorrangige Behandlung einzelner Wortklassen. Es ist nämlich vorgesehen, dass für die erste Subjektbestimmung nur die Wortklassen „Substantiv“, „Adjektiv“ und „Partizip“ abgefragt werden. Nur diese kommen auch in der ergänzenden Einheit in Frage, wenn in der ersten Einheit kein Subjekt bestimmt werden konnte. Für den Fall, dass auch hier kein Subjekt gefunden wird, sind in einer zweiten Subjektbestimmung noch andere Wortklassen als Subjekt zugelassen, die über die beiden Einheiten abgefragt werden. Dies bedeutet, dass es sozusagen zwei „Oberklassen“ von Wortklassen gibt, die als Subjekt fungieren können, wobei die eine „Oberklasse“ Priorität vor der anderen hat. Dies sei im folgenden kurz erläutert: Es wird angenommen,

1. a1 und b1 seien Vertreter je einer Oberklasse A und B, die als Subjekt in Frage kommen.
2. A hat Vorrang vor B.
3. In der Einheit vor PC kommt b1 vor, in der Einheit nach PC kommt a1 vor.

In diesem Fall verlief das Programm folgendermaßen:

1. Schritt: Da in einer ersten Subjektbestimmung b1 als Subjekt nicht zugelassen ist, wird festgestellt, dass die Einheit kein Subjekt aufweist und somit ergänzungsfähig ist.

2. Schritt: Bei den Abfragen der Ergänzungsmöglichkeiten würde a1 in der Einheit nach PC als Subjekt bestimmt werden. Ein 3. Schritt, die zweite Subjektbestimmung, ist überflüssig, da ein Subjekt gefunden ist. Es bleibt aber fraglich, welche Wortklasse „eher“ die Funktion des Subjekts übernehmen kann, wenn man von der Zulässigkeit dieser Fragestellung überhaupt absieht. Immerhin wäre durchaus denkbar, dass im obigen Beispiel b1 das Subjekt der ersten Einheit ist, a1 das Subjekt einer selbständigen dritten Einheit. Dieser Fall bleibt im Programm unbeachtet, könnte aber von der Anlage des Programms her leicht berücksichtigt werden. Im Einfachsatz (oder wenn die erste Einheit ein Subjekt enthält) ist nämlich die Möglichkeit eines „multiple subject“ (vgl. S. 11 f.) erfasst. Sie müsste auch für die Fälle gelten, in denen zuerst das Subjekt in der dritten Einheit bestimmt wird. Wie die Information weiter zu behandeln wäre - es muss ja schließlich die Entscheidung für *ein* Subjekt¹² getroffen werden -, bleibt zu untersuchen. Immerhin wäre bei einer solchen Notierung die Information „mögliches Subjekt“ noch vorhanden und könnte weiter verarbeitet werden, was bisher nicht der Fall ist. Letztlich kommt es wohl auch hier auf die Kenntnis aller Beziehungen an, für die die Subjektbestimmung im Einfachsatz erläutert wurde (vgl. S. 10f.).

Nachdem in einem groben Überblick dargelegt wurde, welchen Anteil Pass 3 an der Analyse hat, werden im folgenden die einzelnen Programmschritte von Pass 3 detailliert dargestellt. Es handelt sich um folgende Teile

1. Subjekt- bzw. Prädikatbestimmung für den Einfachsatz.
2. Erste Subjekt- bzw. Prädikatbestimmung für die erste Einheit, wenn der Satz mehr als zwei Einheiten enthält.
3. Erkennung der Ergänzungsmöglichkeiten der ersten Einheit durch die dritte Einheit und Bestimmung einer eingebetteten Satzeinheit.
4. Zweite Subjekt- bzw. Prädikatbestimmung für den Satz mit mehr als zwei Einheiten, wenn noch eine entsprechende Bestimmung fehlt. (Dieser Programmteil ist mit der Subjekt- bzw. Prädikatbestimmung für den Einfachsatz identisch, vgl. 1). Die Subjekt- bzw. Prädikatbestimmung erfolgt für alle Einheiten, also auch für die eingebettete Satzeinheit.

Teil II

Beschreibung des technischen Ablaufs der Programme in Pass 3.

Pass 3 gliedert sich, generell gesagt, in zwei Programmabschnitte. Es handelt sich hierbei:

1. um die Subjekt- bzw. Prädikatbestimmung¹³ und
2. um die Erkennung einer eingebetteten Satzeinheit¹⁴.

Das Programm für die Subjekt- bzw. Prädikatbestimmung ist nochmals in sich gegliedert, und zwar in der Weise, dass jeweils Einschränkungen in den Abfragen und Änderungen in der Reihenfolge der Abfragen vorgenommen werden können. Dieser Vorgang wird durch „Schalter“ reguliert.

Eigentlich handelt es sich hierbei um verschiedene Programmabschnitte, wie sie vorstehend zusammengestellt sind (vgl. Programmteile 1, 2 und 4). Da aber oftmals streckenweise die gleichen Abfragen vorkommen, erscheint allein aus Programmökonomie eine Zusammen-

legung vorteilhaft. Allerdings verkompliziert sich dadurch der gesamte Programmablauf, da auf richtige Schaltersetzung und -löschung wie auch auf Abfragen der Schalter an jeder Verzweigung geachtet werden muss. Um die Beschreibung möglichst verständlich zu halten, werden im folgenden bei Hinweisen auf bestimmte Programmteile die Bezeichnungen der genannten Zusammenstellung verwendet.

Was die Schalter selbst angeht, sei auf ihre gegenüber Pass 1 verschiedene Verwendungsweise besonders hingewiesen. Es gibt u. a. zwei Schalter, die nicht wie in Pass 1 Merkstellen für das Vorhandensein bestimmter Wörter darstellen, sondern die den Programmforgang regulieren.

Hierbei handelt es sich in erster Linie um die Unterscheidung einer Subjekt- bzw. Prädikatbestimmung im Einfachsatz (vgl. Programmteil 1) von der Subjekt- bzw. Prädikatbestimmung in einem Satz mit mehr als zwei Einheiten, der also möglicherweise eine eingebettete Satzeinheit enthält (vgl. Programmteile 2, 3 und 4). So gibt der Schalter SUPRASW¹⁵, mit 1 besetzt, darüber Auskunft, dass es sich im folgenden um die *erste* Subjekt- bzw. Prädikatbestimmung für die erste Einheit in einem Satz mit mehr als zwei Einheiten handelt, oder (wenn er 0 ist), dass es sich um eine Subjekt- bzw. Prädikatbestimmung ohne Einschränkungen handelt.

Auf die Zusammenstellung bezogen müsste folgende Schalterstellung gegeben sein, d. h. am Eingang des Programms gesetzt worden sein:

für Programmteil 1: SUPRASW = 0,
für Programmteil 2: SUPRASW = 1,
für Programmteil 4: SUPRASW = 0.

Programmteil 3 stellt einen von der Subjekt- bzw. Prädikatbestimmung, wie sie in 1, 2 und 4 durchgeführt wird, unabhängigen Programmteil dar, für den SUPRASW keine Rolle spielt.

Der andere Steuerschalter, SUPRPCAN¹⁷, wird im Programmteil 4 verwendet¹⁸. Er gibt an, ob im folgenden, wenn mit 1 besetzt, eine eingebettete Satzeinheit auf das Vorhandensein von Subjekt und Prädikat untersucht wird, oder ob es sich, wenn der Schalter 0 gesetzt ist, um eine Subjekt- bzw. Prädikatbestimmung im übrigen Satz handelt. Dieser Schalter ist besonders wichtig, da bei der Bearbeitung der diskontinuierlichen Einheiten die Wörter der eingebetteten Einheit übersprungen werden müssen. Dieser Schalter wird deswegen jeweils im Verein mit Byte 34 eines jeden Wortes abgefragt (Byte 34 gibt an, ob das betreffende Wort innerhalb einer PC steht). Soll nämlich z. B. eine eingebettete Einheit bearbeitet werden, dürfen sich die Abfragen nur auf Wörter beziehen, die in dieser Einheit stehen. Die übrigen Schalter entsprechen der Art der Schalter von Pass 1; wie dort ist jedem Schalter auch eine Merkstelle für die Wortnummer, ein Pointer oder Zeiger, zugeordnet.

Folgende Schalter und Zeiger werden zusätzlich in Pass 3 verwendet:

SUPRSUSW ¹⁹	= Schalter für „Subjekt vorhanden“
SUPRSUP ²⁰	= Zeiger für Wortnummer des Subjekts
SUPRPRSW ²¹	= Schalter für „Prädikat vorhanden“
SUPRPRP ²²	= Zeiger für Wortnummer des Prädikats
DASHSW ²³	= Schalter für „Gedankenstrich vorhanden“
DASHP ²⁴	= Zeiger für Wortnummer des Gedankenstrichs
NODASHBRWS ²⁵	= Schalter für „Gedankenstrich ist keine Grenze einer Einheit“

NLMOD5, 6, 7 ²⁶	= Schalter für „Erste Einheit enthält kein Subjekt und kein Prädikat, aber ein Substantiv, das kein Objekt ist“
ILMOD8 ²⁶	= Schalter für „Erste Einheit enthält kein Subjekt und kein Prädikat, aber einen nicht zugeordneten Infinitiv“
ALMOD5, 6, 7 ²⁶	= Schalter für „Erste Einheit enthält kein Subjekt und kein Prädikat, aber ein nicht zugeordnetes Adjektiv“
PFMODEL ²⁷	= Zeiger für die Wortnummer des nicht zugeordneten Wortes in der ersten Einheit, wenn sie kein Subjekt und kein Prädikat enthält
PCESTABL ²⁸	= Schalter für „Zweite Einheit des Satzes ist eine eingebettete Satzeinheit“

Obwohl es dem System entspräche - und auch der Übersichtlichkeit dienen könnte -, wird der Sachverhalt, ob ein Satz aus mehr als zwei Einheiten besteht, nicht auf einem Schalter notiert. Statt dessen finden am Anfang von Pass 3 in Abhängigkeit von den erfragten (Einheiten-) Grenzen Verzweigungen statt (die Grenzen waren bereits in Pass 2 festgestellt worden). Drei verschiedene Möglichkeiten sind dabei zunächst vorgesehen:

1. Die erste Grenze ist gleichzeitig Satzgrenze (SPPC 186, SPPC 110)²⁹.
2. Die zweite Grenze ist gleichzeitig Satzgrenze (SPPC 116).
3. Die zweite Grenze ist nicht die Satzgrenze (SPPC 117).

Für die Fälle 1 und 2 wird die Subjekt- bzw. Prädikatbestimmung ohne Einschränkungen durchgeführt, d. h. das Programm SUPRSUBR30 (vgl. S. 11) wird aufgerufen, der Schalter SUPRASW wird auf 0 gesetzt, was besagt, dass alle Abfragen des Programms SUPRSUBR durchlaufen werden können. Für die Subjektbestimmung bedeutet dies, dass außer Substantiven, Adjektiven und Partizipien auch Infinitive und „MNOGO“ als Subjekt in Frage kommen (vgl. S. 9). Bei der Bearbeitung von Fall 3 wird die Möglichkeit vorgesehen, dass die zweite Einheit eine eingebettete Satzeinheit, eine PC, ist. Deswegen wird zunächst eine Subjekt- bzw. Prädikatbestimmung für die erste Einheit versucht (vgl. S. 12).

Das Programm SUPRSUBR wird mit Einschränkungen durchlaufen. Dies wird dadurch kenntlich gemacht, dass der eigens dafür eingerichtete Schalter SUPRASW auf 1 gesetzt wird. Jetzt kommen als Subjekt nur noch Substantive, Adjektive und Partizipien in Frage. Daraufhin wird das Programm zur Erkennung der Ergänzungsmöglichkeiten der 1. Einheit, PCSUBR, durchlaufen. In Abhängigkeit von dem Ergebnis dieses Programmteils erfolgt nochmals eine Subjekt- bzw. Prädikatbestimmung für die einzelnen Einheiten, soweit sie jetzt noch fehlen. Das Programm SUPRSUBR wird nochmals durchlaufen, und zwar jetzt ohne Einschränkungen (SUPRPCAN ist also = 0). Außerdem wird hier der Schalter SUPRPCAN besetzt, je nachdem, welche Einheit bearbeitet werden soll³¹. Als letzter Hinweis zum technischen Ablauf der Programmteile sei noch gesagt, dass jeder Programmteil in Abhängigkeit von der Information durchlaufen wird, ob bereits ein Subjekt bzw. ein Prädikat in einem vorhergehenden Schritt, sei es innerhalb von Pass 3 oder in Pass 1, bestimmt worden ist. Diese Information liefern Subjektschalter (SUPRSUSW) und Prädikatschalter (SUPRPRSW). An diese allgemeinen Erläuterungen schließt sich nun die detaillierte Beschreibung der Programmteile, der Reihenfolge der genannten Zusammenstellungen folgend, an.

A. Subjekt- bzw. Prädikatbestimmung im Einfachsatz:

Verwendetes Programm: SUPRSUBR
Schalterstellung: SUPRASW = 0
(SUPRPCAN = 0, bleibt aber unberücksichtigt)

Im folgenden die Verzweigungen in Abhängigkeit von den bisherigen Ergebnissen

1. Fall: Subjekt und Prädikat sind bereits bestimmt:

Für das als Subjekt bestimmte Wort werden die „Subjekt-Informationen“ festgehalten (vgl. Speicherbeschreibung). Ebenfalls besetzt wird der „Subjektschalter“ (SUPRSUSW) wie auch die Merkstelle für die Wortnummer des Subjekts, der Subjekt-Pointer (SUPRSP) (vgl. SUPR 318). Für das Prädikat gilt dies in analoger Weise. Hier werden entsprechend der Prädikatschalter (SUPRPRSW) und Prädikat-Pointer (SUPRPRP) besetzt (SUPR 390).

Außerdem werden die übrigen Wörter des Satzes daraufhin geprüft, ob ein weiteres Wort als Subjekt oder Prädikat möglich ist. Als „multiple subject“ (vgl. S. 11) kommt ein Substantiv, Adjektiv oder Partizip in Frage, das noch nicht zugeordnet ist und als Kasus Nominativ haben kann (SUPR 310). (Das Substantiv darf nicht auf KAK oder 3T0 folgen.) Erfüllt ein Wort diese Bedingungen, gilt es als „multiple subject“. Es wird als solches im Arbeitsspeicher gekennzeichnet. Außerdem erhalten die beiden Wörter gegenseitige Verweise. Mehr als ein weiteres Wort, das Subjekt sein kann, wird nicht festgehalten. Das zuletzt gefundene erhält die Kennzeichnung, die übrigen bleiben ungekennzeichnet. Entgegen der sonstigen Priorität der Objekt-Zuordnung wird hier nicht geprüft, ob das Wort, wenn es auch noch nicht „satisfier“ ist, evtl. auch „satisfier“ sein könnte und ob eine Rektion des finiten Verbs noch unbesetzt ist.

Auch ein Infinitiv ist als Teil eines mehrteiligen Subjekts möglich (SUPR 370), wenn er kein „satisfier“ ist. Dasselbe gilt für ein Wort der Klasse MNOGO (SUPR 380); in diesem Fall darf es zwar ein „governor“, aber kein „satisfier“ sein.

Unberücksichtigt bleibt bei diesen Abfragen eine Priorität der Wortklassen. Es kann vorkommen, daß zunächst ein nicht zugeordnetes Substantiv im Nominativ gefunden wird. Die Suche nach einem weiteren „zweiten“ Subjekt geht aber weiter. Tritt nun ein Infinitiv auf, der nicht „satisfier“ ist, gilt dieser automatisch als „zweites Subjekt“, d. h. das zuletzt gefundene Wort, das die Bedingungen erfüllt, sei es ein Substantiv, Adjektiv, Partizip, Infinitiv, MNOGO, wird zum zweiten Subjekt erklärt. Zwar bleibt die Information „multiple subject“ bei den übrigen, wenn vorhanden, erhalten, der Hinweis bei dem ersten Subjekt existiert aber nicht mehr.

Ähnlich wird bei dem Prädikat verfahren (SUPR 319). Als zusätzliche Prädikate kommen aber nur solche in Frage, die schon in einem vorhergehenden Programmschritt, d. h. in Pass 1 als „primary predicate“ oder „secondary predicate“, bestimmt worden waren. Sind alle Wörter des Satzes auf die Möglichkeit eines zweiten Subjekts oder zweiten Prädikats überprüft, erfolgt der Rücksprung zum Eingang des Programms, von da der Rücksprung ins Hauptprogramm, wenn das Satzende erreicht ist (sonst beginnt das Programm bei dem nächsten Wort mit der Einteilung in Einheiten [vgl. S. 10f. und 14], worauf ein Sprung in das dem ermittelten Typ gemäße Programm erfolgt).

2. Fall: Prädikat bereits bestimmt, Subjekt nicht bestimmt:

Für das Prädikat wird so verfahren wie bei Fall 1. Wir betrachten hier nur den Versuch, das Subjekt zu bestimmen. Die Bedingungen für ein zweites Subjekt bei Fall 1 gelten hier als Bedingungen für das Subjekt überhaupt, also in Frage kommen: Substantiv, Adjektiv, Partizip, die nicht einem anderen Wort des Satzes zugeordnet sind und als Kasus einen Nominativ aufweisen, oder ein Infinitiv, der kein „satisfier“ ist, oder MNOGO, das kein „satisfier“, aber „governor“ ist. Wird bei diesen Abfragen ein Wort gefunden, das diese Bedingungen erfüllt, gilt das Subjekt als bestimmt. In diesem Fall erfolgt dann sogleich der Rücksprung. Konnte aber im ganzen Satz kein Wort gefunden werden, das diese Bedingungen erfüllt, erfolgt der Rücksprung, ohne dass ein Subjekt bestimmt worden ist.

3. Fall: Subjekt ist bestimmt, Prädikat nicht:

Die Schaltersetzung für das Wort, das Subjekt ist, erfolgt wie bei Fall 1 (SUPR 300). Zur Bestimmung des Prädikats werden weitere Abfragen durchlaufen (SUPR 700), wobei folgende Bedingungen gelten:

a) Das Subjekt ist ein Wort im Sg. neutrum (SUPR 766): als Prädikat können bestimmt werden: Wörter der Klassen Kurzform Adj./ Partizip, die ebenfalls Sg. neutr. haben, unregelmäßiger und regelmäßiger Komparativ (SUPR 418). Die Abfragen erfolgen vom Satzende aus nach links, d. h. das am weitesten nach rechts stehende Wort, das diese Bedingungen erfüllt, wird als Prädikat bestimmt. Fehlt ein solches Wort im ganzen Satz (SUPR 440), wird zwischen Subjekt und dem nächsten, nicht aufzählenden Komma (bzw. Konjunktion) ein Adjektiv gesucht (SUPR 461). Ist dieses Adjektiv Sg. neutr. (Kurzform) und kann es Nominativ sein, wird es als Prädikat betrachtet (SUPR 448). Im anderen Fall erfolgt keine Prädikatbestimmung.

b) Das Subjekt ist ein Infinitiv (SUPR 640): dann wird zunächst über die ganze Einheit hinweg eine Kurzform von Adjektiv/Partizip gesucht, die möglicherweise Prädikat sein kann (d. h. die in Byte 7 diese Information hat) oder ein Infinitiv. Wird zuerst eine solche Kurzform angetroffen (SUPR 816), wird diese als Prädikat bestimmt. Hierbei wird eine Aufzählung von Kurzformen berücksichtigt (nach Aufzählungskomma bzw. nach Konjunktion, SUPR 816, 817), wobei alle in einer Aufzählung befindlichen Kurzformen als Prädikat gekennzeichnet werden. Wird zuerst ein Infinitiv angetroffen (SUPR 816), wird die dem Satzende am nächsten stehende Kurzform (d. h. Abfrage vom Satzende aus nach links) als Prädikat bestimmt, wobei wiederum eine Aufzählung von Kurzformen in der genannten Weise berücksichtigt ist.

Bei den Abfragen nach einer Kurzform von Adjektiv/Partizip wird zusätzlich ein evtl. auftretender Gedankenstrich berücksichtigt (SUPR 869). Kommt er im Satz vor, wird zunächst abgefragt, ob noch ein zweiter vorhanden ist. In diesem Fall werden die Wörter zwischen den beiden Gedankenstrichen bei der Suche nach einem möglichen Prädikat übersprungen.

c) Das Subjekt ist ein Wort der Klasse MNOGO (SUPR 799): In diesem Fall wird ein unregelmäßiger Komparativ - wenn vorhanden - als Prädikat bezeichnet. Findet sich im ganzen Satz kein solcher, kann auch noch eine Kurzform Adj./Part. (mit Aufzählung) (SUPR 788) oder (SUPR 730) ein Wort der Klasse BOLEE oder ein regelmäßiger Komparativ Prädikat sein. Fehlen solche Wörter (SUPR 748), kommen auch Adverbien (nicht MNOGO-Typ) oder Kurzformen neutr. Sg. als Prädikat in Betracht, wenn sie nicht bereits zugeordnet sind³².

d) Das Subjekt ist kein Wort im Neutr. Sg., kein Infinitiv, kein MNOGO: Der erste unregelmäßige Komparativ wird zum Prädikat (SUPR 776). Wird kein solcher gefunden, kann auch ein Wort der Klasse BOLEE (SUPR 798) oder ein Adjektiv im Komparativ Prädikat sein. Kommt ein solches Wort im ganzen Satz nicht vor (alle Wörter werden danach abgefragt, bevor eine weitere Wortklasse überprüft wird) (SUPR 670), wird bis zum nächsten Komma (bzw. Konjunktion), das nicht Aufzählungskomma ist, ein Vollform-Adjektiv gesucht (SUPR 681, 676). Wenn es nach rechts zugeordnet ist, wird es übersprungen, ebenfalls wenn es nicht (u. a.) Nominativ sein kann. Außerdem darf es nicht als „satisfier“ gekennzeichnet sein. Wenn es zudem mit dem Subjekt in einem Kasus übereinstimmt, gilt es als Prädikat. Im anderen Fall bleibt es ebenfalls unberücksichtigt. Wird kein entsprechendes Vollform-Adjektiv gefunden (SUPR 680), kann noch ein Infinitiv, der nicht zugeordnet ist, Prädikat sein. Ein Komma sowie eine Konjunktion, die nicht eine Aufzählung verbinden, gelten als Grenze für die Prädikatsuche bei Adjektiv und Infinitiv. Wenn die genannten Bedingungen nicht erfüllt werden, erfolgt der Rücksprung ohne Prädikatbestimmung.

4. Fall: Kein Subjekt, kein Prädikat bestimmt (SUPR 719):

Etwas anders verläuft die Prädikatsuche, wenn außerdem noch eine Subjektbestimmung fehlt. Ein Subjekt wird im übrigen auch nicht bestimmt. Es handelt sich hier nur um eine Prädikatbestimmung. Zuerst wird über den ganzen Satz hinweg ein unregelmäßiger Komparativ gesucht. Tritt ein solcher im Satz auf (SUPR 718), gilt er als Prädikat. Der Lauf ist nach der entsprechenden Kennzeichnung beendet. Gibt es keinen unregelmäßigen Komparativ, werden alle Kurzform-Adjektive als Prädikate bestimmt (SUPR 786). Ist keine Kurzform vorhanden (SUPR 730), können Wörter der Klasse BOLEE oder ein regelmäßiger Komparativ das Prädikat bilden (SUPR 734), ebenso ein Adverb der Klasse MNOGO oder eine Kurzform neutr. Sg., wenn diese Wörter nicht bereits als „satisfier“ gekennzeichnet sind. Erfüllt keines der Wörter im Satz die genannten Bedingungen, erfolgt ein Rücksprung ohne Prädikatangabe.

Zusammenfassend sei die Subjekt-Prädikat-Bestimmung für den Einfaßsatz schematisch dargestellt:

1. Fall: Subj./Präd. bestimmt (evtl. 2 Subj. und 2 Präd.):

- als 2. Subj.: Substantiv, Adjektiv, Partizip (wenn nicht zugeordnet, wenn Nominativ möglich)³³
 Infinitiv (wenn kein satisfier)
 MNOGO (wenn kein satisfier, aber governor)
- als 2. Präd.: Wort, das bereits als Prädikat gekennzeichnet ist.

2. Fall: Präd. bestimmt, Subj. nicht:

- als Subj.: Substantiv, Adjektiv, Partizip (wenn nicht zugeordnet, wenn Nominativ möglich)
 Infinitiv (kein satisfier)
 MNOGO (kein satisfier, governor)

3. Fall: Subjekt bestimmt, Prädikat nicht:

- wenn 1. Subj. = Sg. neutr.
- dann als Präd. = Kurzform Adj/Part. (Sg. neutr.)
 BOLEE, regelmäÙ. Komp., unregelmäÙ. Komp.
 (Kurz- oder Langform, Sg. neutr., Nominativ)
- Adjektiv = (Kurz- oder Langform, Sg. neutr., Nominativ)
- 2. Subj. = Infinitiv

als Präd.	=	Kurzform Adj./Part. (mögl. Prädikat)
3. Subj.	=	MNOGO
als Präd.	=	unregelm. Komp. Kurzform Adj./Part. BOLEE, regelm. Komp. Adverb Kurzform neutr. Sg.
4. Subj.	=	nicht Sg. neutr., Inf., MNOGO
als Präd.	=	unregelm. Komp. BOLEE regelm. Komp. Vollform-Adj. (nicht zugeordnet, Nominativ, kein satisfier, Übereinstimmung mit Kasus von Subj., Abfrage nur bis Komma bzw. Konj.) Infinitiv (kein satisfier)

4. Fall: kein Subj., kein Präd.:

als Präd.	=	unregelm. Komp. Kurzform Adj./Part. regelm. Komp. (BOLEE) MNOGO Kurzform neutr. Sg. (kein satisfier)
keine Subjektbestimmung		

B. Subjekt- und Prädikatbestimmung bei „Schachtelungen“:

Anders verhält es sich bei dem Programmablauf für einen Satz mit mehr als zwei Einheiten. Hier wird zunächst eine erste Subjekt-Prädikat-Bestimmung für die 1. Einheit allein durchgeführt. Dieser Programmteil enthält lediglich einen Teil der Abfragen des Programms für die zweite Subjekt-Prädikat-Bestimmung. Die Auswahl der Abfragen wird durch den Schalter SUPRASW reguliert (vgl. S. 14). In einem weiteren Programmteil wird festgestellt, ob die 1. Einheit über eine mittlere hinweg in einer dritten weitergeführt wird (PCSUBR) (vgl. Programmteil 3, S. 14). Das Ergebnis dieses Laufs bestimmt im Weiteren, ob sich die anschließende nochmalige Subjekt-Prädikat-Bestimmung über die drei Einheiten getrennt (wenn PC keine clause ist) oder jeweils über die erste und dritte Einheit zusammen und getrennt davon über die mittlere Einheit, nämlich die PC, erstreckt (wiederum SUPRASW = 0, vgl. S. 48).

1. Erste Bestimmung von Subjekt und Prädikat in der ersten Einheit (*SUPRSUPR*).

(SUPRASW wird eingangs auf 1 gesetzt).

a) Wenn Subjekt und Prädikat bereits in einem vorangegangenen Schritt eindeutig bestimmt werden konnten, so erfolgt hier noch die entsprechende Schaltersetzung, für Subjekt wird SUPRSUSW auf 1 gesetzt, für das Prädikat entsprechend der Schalter SUPRPRSW. Die Nummer des jeweiligen Wortes wird in SUPRSP bzw. SUPRPRP festgehalten (vgl. SUPR 300, 301 für Subjekt, vgl. SUPR 390 für Prädikat).

Hierbei ist auch vorgesehen, dass mehrere Wörter als Subjekt und Prädikat in Frage kommen. Dies wird ebenfalls festgehalten, wobei das so genannte 1. Subjekt einen Verweis auf das 2. erhält und umgekehrt (SUPR 319), analog gilt dies für zwei Prädikate (SUPR 391). Notiert im

„Pointer“ wird jeweils das Subjekt bzw. Prädikat, das zuletzt gefunden wurde, d. h. das am weitesten nach rechts steht. (Die Wortnummer des anderen ist stets bei dem notierten als Verweis zu finden.)

b) Wenn noch kein Prädikat bestimmt worden ist, wird in diesem Lauf auch nicht versucht, es zu bestimmen. Dies bleibt in diesem Fall der 2. Subjekt-Prädikat-Bestimmung vorbehalten.

c) Konnte noch kein Subjekt bestimmt werden, wird ein Substantiv, Adjektiv oder Partizip gesucht (SUPR 310), das nicht bereits als Objekt bestimmt ist, nicht nach rechts zugeordnet ist (SUPR 314) und möglicherweise im Nominativ steht (SUPR 316). Sind diese Bedingungen erfüllt, wird das Wort als Subjekt bestimmt (SUPR 318). (Infinitiv und MNOGO gelten bei der ersten Subjektbestimmung nicht als mögliches Subjekt, vgl. 2. Subjektbestimmung.)

Zusammengefasst sind die Möglichkeiten für die erste Subjekt- bzw. Prädikatbestimmung folgende:

1. Prädikat ist noch nicht bestimmt:
keine Prädikatbestimmung

2. Subjekt ist noch nicht bestimmt:
als Subjekt kommen in Frage:

Substantiv	wenn nicht als Objekt bestimmt,
Adjektiv	nicht nach rechts zugeordnet,
Partizip	Nominativ möglich

2. Programm zur Erkennung der Ergänzungsmöglichkeiten der ersten Einheit durch die dritte Einheit und Bestimmung einer eingebetteten Satzeinheit (PC SUBR).

Dieses Programm basiert auf den Ergebnissen der ersten Subjekt- bzw. Prädikatbestimmung in der ersten Einheit. Danach werden wiederum vier verschiedene Möglichkeiten unterschieden.

1. *Fall*: Subjekt und Prädikat in der ersten Einheit sind bestimmt (PC 600).

a) Wenn dem letzten Wort der PC ein I oder ILI folgt (PC 616), (PC 486), gilt die zweite Einheit als PC. Sie wird als solche gekennzeichnet, indem jedes Wort dieser Einheit in Byte 34 die Information 'FF' erhält; außerdem wird ein Schalter gesetzt, PCESTABL = 1, der besagt, dass die als mögliche PC angesehene zweite Einheit tatsächlich eine eingebettete Satzeinheit ist.

b) Wenn die oben genannten Wörter nicht folgen, wird untersucht, ob nach der PC evtl. ein Objekt stehen kann, das sich auf das Prädikat der ersten Einheit bezieht (PC 618), wie es beispielsweise bei einem eingeschobenen Relativsatz wie im Deutschen auftreten kann: ER GIBT DAS BUCH, DAS ER AUSGELIEHEN HAT, SEINEM FREUND.

Diese Untersuchung erstreckt sich auf Wörter bis zur nächsten Grenze. Zunächst wird dabei festgestellt, ob das Prädikat der ersten Einheit eine noch nicht durch ein Objekt ausgefüllte Rektion hat. Im obigen deutschen Beispiel wäre es die Dativ-Rektion von „gibt“. Ist dies der Fall und gibt es in der Einheit nach der PC ein Substantiv oder Adjektiv bzw. Partizip (PC 640), das noch nicht als Objekt gekennzeichnet ist (PC 646), wird in einem Unterprogramm,

SPRGOVNS (nested subroutine), festgestellt, ob ein Kasus des betreffenden Nomens mit der noch freien Rektion des Verbs übereinstimmt. Bei positivem Ausgang wird, je nachdem die Stelle des 1. Objekts oder die des 2. Objekts (vgl. Arbeitsspeicher) unbesetzt ist, dieses Wort entsprechend gekennzeichnet. Damit ist über eine PC hinaus eine Prädikat-Objekt-Beziehung festgestellt, die im Ergebnisspeicher festgehalten wird. Auf Grund dieses Ergebnisses gilt die mittlere Einheit als eingebettete Satzeinheit, als tatsächliche PC. Hier ist anzumerken, dass vom Programm her nur die Zuordnung eines Objekts vorgesehen ist. Wenn kein Objekt gefunden werden kann, gilt die zweite Einheit nicht als eingebettete Satzeinheit, die erste und dritte Einheit werden als selbständig betrachtet (PC 494).

2. *Fall*: Subjekt der ersten Einheit ist bestimmt, das Prädikat nicht (PC 660).

In diesem Fall wird die Suche nach einem Prädikat auf Wörter der Einheit nach der PC ausgedehnt.

a) Wird bis zur nächsten Grenze oder bis zum Satzende kein Wort gefunden, das bereits als Prädikat bestimmt ist, gilt die zweite Einheit nicht als eingebettete Satzeinheit (PCE STABL wird auf 0 gesetzt, PC 494).

b) Bis zur nächsten Grenze wird ein Wort gefunden, das bereits als Prädikat bestimmt ist. Dieses Wort kann Prädikat zum vorhandenen Subjekt sein, wenn folgende Bedingungen erfüllt sind (PC 6701):

1. Es darf bis zur nächsten Grenze kein bereits als Subjekt bestimmtes Wort auftreten (sonst keine Kennzeichnung als PC).
2. Außerdem darf es kein Substantiv sein, das evtl. Subjekt sein könnte, d. h. es darf nicht im Nominativ stehen und muss als Objekt bereits gekennzeichnet sein.
3. Wenn diese Bedingungen erfüllt sind, werden Geschlecht und Numerus des Subjekts und des gefundenen Prädikats auf Übereinstimmung geprüft (PC 740). Ist eine Übereinstimmung möglich, wird PC als solche gesetzt (PC 486: PC E STABL = 1, Byte 34 = FF). Im anderen Fall gilt PC als nicht gesetzt (PC 494).

3. *Fall*: Das Prädikat in der ersten Einheit ist bestimmt, das Subjekt nicht (PC 300).

Ebenfalls über PC nach rechts hinaus wird ein mögliches Subjekt gesucht.

a) Bis zur nächsten Grenze oder zum Satzende werden zunächst die Substantive daraufhin geprüft. Wird ein Substantiv gefunden, das bereits als Objekt bestimmt ist, wird es übergangen und ein weiteres gesucht (PC 310); ebenso wird verfahren, wenn es kein Nominativ sein kann (PC 316). Wenn mehrere Kasus möglich sind (PC 314), muss zunächst festgestellt werden, ob das betreffende Substantiv evtl. ein Objekt zum Prädikat sein kann. Dies geschieht im Unterprogramm SPRGOVNS. Hierbei hat die Möglichkeit eines Objekts Vorrang vor der Möglichkeit, Subjekt zu sein. Wenn das Substantiv Objekt sein kann, wird PC als solche gesetzt, das Objekt als solches gekennzeichnet. Die mittlere Einheit wird ebenfalls als PC gesetzt, wenn sich das Substantiv als eindeutiges Subjekt bestimmen lässt (PC 316). Dazu ist aber Übereinstimmung des betreffenden Substantivs und des Prädikats im Genus und Numerus erforderlich. Diese wird im Unterprogramm CGN91INS abgefragt. Ist keine Übereinstimmung vorhanden, kann das Substantiv auch nicht Subjekt sein; dann wird ein weiteres Substantiv gesucht und analog bearbeitet.

b) Wenn kein Substantiv mehr vorhanden ist (PC 340), werden dieselben Abfragen wie unter a) für ein evtl. vorhandenes Adjektiv ausgeführt.

4. *Fall*: Subjekt und Prädikat der ersten Einheit sind nicht bestimmt (PC 410). In diesem Fall wird auch keine Bestimmung über PC hinaus versucht. Statt dessen werden bestimmte Schalter gesetzt.

a) Wenn in der ersten Einheit ein Substantiv auftritt (PC 400), das nicht als Objekt bestimmt ist (PC 406), wird der Schalter NLMOD5, 6,7 gesetzt. Die Wortnummer des Substantivs wird auf PFMODEL gespeichert, PC als „tatsächliche clause“ gesetzt (PC 486).

b) Anstelle eines Substantivs kann auch ein Infinitiv eine Schaltersetzung bewirken: ILMOD8, wenn er nicht zugeordnet ist (PC 410). Auch in diesem Fall wird PC gekennzeichnet.

c) Wenn in der ersten Einheit kein Substantiv oder Infinitiv gefunden werden kann, kann auf Grund eines nicht zugeordneten Adjektivs der Schalter ALMOD 5, 6, 7 gesetzt werden. PC wird ebenfalls gekennzeichnet.

d) Treffen die Möglichkeiten 1, 2 und 3 nicht zu, gilt PC nicht als „clause“.

Zusammenfassend kann über PCSUBR gesagt werden: eine PC wird als „tatsächliche clause“ (eine den Hauptsatz unterbrechende Einheit) betrachtet, wenn festgestellt werden kann, dass über die mittlere Einheit hinweg ein syntaktischer Bezug der ersten zur dritten Einheit besteht, der folgendermaßen geartet sein kann:

1. in der dritten Einheit steht ein Objekt zum Prädikat der 1. Einheit,
2. die 3. Einheit erhält das Prädikat, das auch für die 1. Einheit gilt,
3. die 3. Einheit enthält das Subjekt, das ebenfalls auch für die 1. Einheit gilt.

Die Bedingungen und Ergebnisse von PCSUBR sind zusammengefasst folgende:

1. Fall: Subjekt und Prädikat in der 1. Einheit sind bestimmt:

PC wird gesetzt³⁴; wenn³⁵ gilt

1. auf PC folgen I, ILI
2. nach PC ist Objekt zu Prädikat der ersten Einheit möglich

2. Fall: Subjekt in erster Einheit bestimmt, Prädikat nicht:

PC wird gesetzt, wenn gilt

1. nach PC ist ein Wort als Prädikat bestimmt und es gibt nach PC kein Subjekt oder ein Wort, das Subjekt sein könnte: das Prädikat stimmt in Genus und Numerus mit dem Subjekt überein

3. Fall: Prädikat ist in erster Einheit bestimmt, Subjekt nicht:

PC wird gesetzt, wenn gilt

1. nach PC kann ein Wort Objekt zum Prädikat sein
2. nach PC kann ein Wort Subjekt sein, das in Genus und Numerus mit dem Prädikat übereinstimmt
3. nach PC kann ein Adjektiv Subjekt sein, das in Genus und Numerus mit dem Prädikat übereinstimmt

4. Fall: Kein Subjekt ist bestimmt, kein Prädikat ist bestimmt:

Hier erfolgt keine Bestimmung von Subjekt oder Prädikat, lediglich werden einige Schalter gesetzt, wenn gilt

1. in der ersten Einheit steht ein Substantiv, das nicht Objekt ist
2. in der ersten Einheit steht ein Infinitiv, der nicht zugeordnet ist
3. in der ersten Einheit steht ein Adjektiv, das nicht zugeordnet ist

Die Ergebnisse des Programms PCSUBR werden bei der anschließenden nochmaligen Subjekt- bzw. Prädikatbestimmung verwendet. Bereits als Subjekt oder Prädikat bestimmte Wörter behalten ihre Kennzeichnung, lediglich in den Einheiten, in denen noch eine Subjekt- oder Prädikatbestimmung fehlt, wird eine solche durchgeführt. Dies geschieht wiederum mit SUPRSUBR, wobei jetzt SUPRASW wieder auf 0 gesetzt wird. Je nachdem, welche Einheit bearbeitet werden soll, wird SUPRPCAN gesetzt (vgl. S. 15/16).

Ist der nochmalige Programmdurchlauf der Subjekt- bzw. Prädikatbestimmung beendet, erfolgt der Rücksprung ins Hauptprogramm. Die Beschreibung von Pass 3 bezog sich in erster Linie auf die Abfolge der einzelnen Abfragen. Ob diese sachlich, d. h. linguistisch relevant sind, konnte an dieser Stelle nicht überprüft werden. Es bedarf dazu noch einer weiteren Untersuchung, in der u. a. etwa festgestellt wird, ob in dem Fall, wenn das Subjekt ein Infinitiv ist, als Prädikat nur noch Kurzformen in Frage kommen, um hier nur ein Beispiel zu nennen. Selbst der technische Ablauf der Programmschritte konnte nicht bis in alle Einzelheiten auf seine Verwertbarkeit hin überprüft werden, zumal das Programm zur Synthese, das wohl noch Teile der Analyse enthält, nicht zur Verfügung steht. Allerdings kann doch generell zu Pass 3 gesagt werden, dass es sich hier um einen Programmabschnitt handelt, der für ein Übersetzungsprogramm notwendig ist. Dass er in den einzelnen Abfragen linguistisch zu überprüfen, dass er auch, technisch gesehen, ausgebaut werden muss - hier sei nur an die Möglichkeit einer fehlenden Subjekt- oder Prädikatbestimmung erinnert -, ist eine Forderung, die sich nicht allein von der Kritik der analysierten Sätze her stellt, sondern auch vom Programm selbst.

Pass 4: Gedankenstrich-Mehrdeutigkeit

Der 5. Durchgang ist bisher noch nicht programmiert, er liegt nur in Flussdiagrammform vor. In ihm sollte lediglich der Gedankenstrich auf seine Funktion im Satz hin untersucht werden, soweit dies nicht in den vorausgehenden Durchgängen geschehen ist. Das ist für das Russische besonders wichtig, da der Gedankenstrich hier auch die Funktion der Kopula übernehmen kann.

Schlussbemerkung

Wie aus der Beschreibung des gesamten Verfahrens zu ersehen ist, befand sich das Analyseprogramm bei der Übernahme noch im Teststadium. Bei der Übergabe konnten lediglich 38 (fortlaufende) Sätze eines meteorologischen Fachtextes automatisch analysiert werden. Die Ergebnisse der Analyse lassen keine endgültige Stellungnahme (insbesondere zu den Ausbaumöglichkeiten dieses Verfahrens) zu³⁶.

An dieser kleinen Textmenge ließen sich bereits mehrere grammatische und technisch-formale Fehler oder Unzulänglichkeiten nachweisen. Es ist beabsichtigt, das Programm nach einigen Ergänzungen an einer größeren Textmenge zu überprüfen, ehe entschieden werden kann, inwieweit Teile dieses Verfahrens bei dem Saarbrücker Russisch-Deutsch-Übersetzungsprogramm zu verwenden sind.

Abb. 1970_a_1

Beispiel für den Ausdruck des Ergebnisses einer Satzanalyse

JOB	INITCALL	STEP GO				TIME 115607	DATE 68347			
STORAGE	0	4	8	12	16	20	24	28		
	32	36	40	44	48	52	56	60		
015240	01A00000	00000000	18000000	00016780	03000000	00000000	00000000	00000000		
015260	00000000	00000800	00000000	00000000	00000000	00000000	00000000	00000000		
015280	02200054	E2000000	20000000	41016795	03000000	00000000	00000000	00000000		
0152A0	00000000	00000800	00004040	40000000	00000000	00000000	00000000	00000000		
0152C0	03100040	42000000	10000020	800167A4	00000000	00000000	00000106	00000000		
0152E0	00000000	80000000	00000050	40000009	00000000	00000000	00000000	00000000		
015300	04640000	00000001	62000000	040167B5	00000000	80000000	00000000	00000000		
015320	00000000	00000100	00000000	00000000	00000000	00000000	00000000	00000000		
015340	05A00000	E2000000	18000000	000167CE	00000000	00000000	00000000	00000000		
015360	00000000	00000000	00000000	00000000	00000000	00040000	00000000	00000000		
015380	06200054	E2000000	20000000	410167E1	00030000	00000000	00000000	00000000		
0153A0	00000000	00000400	00005440	40000000	00000000	00000000	00000000	00000000		
0153C0	07040480	02100080	41000000	100167E1	00000000	00000000	00000000	00000000		
0153E0	00000000	00800081	00008404	02000000	00000000	00000000	00000000	00000000		
015400	08100000	21000000	10000020	80016806	00000000	00000000	00000000	07000000		
015420	00000000	00008040	00018010	29000000	00000000	00000000	00000000	00000000		
015440	09104090	43000000	10000020	8001681B	00000000	00000000	00000000	00000000		
015460	03000000	00004000	00020040	47000000	00000000	00000000	00000000	00000000		
015480	0A000000	00000001	05000000	0001682F	00000000	00000000	00000000	00000000		
0154A0	00000000	00000000	00000000	0F070000	00000000	00000000	00000000	00000000		
0154C0	00000000									

Erläuterung:

Die Ergebnisse der Satzanalyse werden stets und ausschließlich in diesem Format ausgegeben. Jedem Wort sind 64 Bytes zugeordnet (= zwei Druckzeilen). Der Inhalt eines Bytes wird durch zwei Zeichen (Ziffern oder Buchstaben) dargestellt. Zur Orientierung wurden von Hand die Nummern einiger Bytes eingetragen. Der Wortlaut des Eingabetextes erscheint nicht bei der Satzausgabe. Der hier analysierte Satz lautet im Russischen:

760 NAZEMNYX STANQII 1 13 OKEANICESKIX PROIZVOD4T ZONDIROVANIE
ATMOSFERY.

Anmerkungen (1-5: siehe Teil 1 dieser Dokumentation)

6 Vgl. S. 9.

7 Nebensätze in der Funktion eines Subjekts werden nicht berücksichtigt.

8 Als Grund wäre eine Verminderung der Analysezeit anzuführen.

9 Wie diese Information weiterhin verwendet wird (in der Übersetzung), kann dem Programm wie auch dem Ergebnisspeicher nicht entnommen werden.

Immerhin wären eventuell mit diesen Notierungen auch Möglichkeiten zu Korrekturen gegeben, wenn z. B. ein mehrteiliges Subjekt aus Vertretern verschiedener Wortklassen besteht, was vom Programm her zugelassen ist.

10 Zum Problem der Mehrdeutigkeit von Satzzeichen vgl. Zimmermann, Zur

Leistung der Satzzeichen, Dudenbeiträge 36, 1969, S. 33f.

11 Die Anzahl der Einheiten wird nicht unmittelbar notiert.

12 Unter Berücksichtigung seiner möglichen diskontinuierlichen Struktur.

13 SUPRSUBR = SUBjekt-PRedicate-SUBRoutine.

14 PCSUBR = PC - SUBRoutine. Unter Satzeinheit verstehen wir hier „vollständige Haupt- oder Nebensätze“.

15 SUPRASW = SUBject-PRedicate-Analyse-SWitch.

17 SUPRPCAN = SUBjekt bzw. PRädikat wird in einer PC ANalysiert (Wörtl. Übersetzung).

18 Programmteil 4 ist zwar seinem Ablauf nach mit Programmteil 1 identisch, der Schalter SUPRPCAN spielt aber für Programmteil 1 keine Rolle.

19 SUPRSUSW = SUBject-PRedicate-SUBject-SWitch (Schalter für das in der Subjekt- bzw. Prädikatbestimmung gefundene Subjekt).

20 SUPRSUP = SUBject-PRedicate-SUBject-Pointer (Zeiger für das in der Subjekt- bzw. Prädikatbestimmung gefundene Subjekt).

21 SUPRPRSW = SUBject-PRedicate-PRedicate-SWitch (Schalter für das in der Subjekt- bzw. Prädikatbestimmung gefundene Prädikat).

22 SUPRPRP = SUBject-PRedicate-PRedicate-Pointer (Zeiger für die Wortnummer das in der Subjekt- bzw. Prädikatbestimmung gefundene Prädikat).

23 DA SHSW = DASH-SWitch.

24 DA SHP = DASH-Pointer.

25 NODASHBRWSW = NO-DASH-BRanch- SWitch.

26 Die Schalter NLMOD 5, 6, 7 wie auch ILMOD 8 und ALMOD 5, 6, 7 werden im Programm PCSUBR gesetzt, wenn in der ersten Einheit kein Subjekt und kein Prädikat bestimmt werden konnten. Welche Bedeutung sie eigentlich haben, außer dass sie auf ein nicht zugeordnetes Substantiv, Adjektiv oder einen nicht zugeordneten Infinitiv hinweisen, ist dem Programm nicht zu entnehmen. Es ist denkbar, dass sie für einen späteren Schritt oder die Synthese vorgesehen sind.

27 PFMODEL = Pointer For MODEL.

28 PCE STABL = PC ist ESTABLished.

29 Die Ausdrücke in runden Klammern stellen Symboladressen des Programms dar. Mit ihrer Hilfe ist ein Vergleich mit den Fußdiagrammen möglich. (vgl. Anmerkung 2).

30 SUPRSUBR = SUBject-PRedicate-SUBRoutine.

31 Es sei wiederholt: SUPRPCAN = 1: Bearbeitung der PC, SUPRPCAN = 0: Bearbeitung der Haupteinheit.

32 Hier scheint eine Sprungadresse falsch gesetzt: Adverb und Kurzform kommen nämlich nur in Frage - laut Programmtext -, wenn es sich um das erste Wort handelt. Zwar wird dieser Programmschritt noch von anderen Stellen des Programms aus benutzt, wofür die Sprungadresse durchaus geeignet ist. Hier müssten die einzelnen Fälle gründlicher auseinandergehalten, die Ausgänge der einzelnen Bearbeitungsschritte differenzierter gehandhabt werden.

33 In runden Klammern sind die Bedingungen angegeben.

34 Dies bedeutet, dass die 2. Einheit als eingebettete Satzeinheit gilt. Als Kennzeichnung dient der Schalter PCESTABL und der Vermerk FF in Byte 34 eines jeden Wortes dieser Einheit.

35 Die nummerierten Bedingungen schließen sich jeweils aus.

36 Eine ins einzelne gehende Untersuchung der Satzanalyse liegt inzwischen vor.