



# Das maschinelle Übersetzungssystem EUROTRA

Markus Ruffing

## MÜ im Aufwind

*Obwohl die maschinelle Übersetzung (MÜ) fast so alt ist wie der Computer selbst, wurde sie von der Wissenschaft in vergangenen Jahren oft als Stiefkind behandelt. Erst in jüngerer Zeit wird der Technologie in verstärktem Maße Aufmerksamkeit gewidmet.*

*Die Ursachen für eine solche globale Trendänderung liegen in der wachsenden Komplexität unserer modernen Welt: nationale (und kontinentale) Grenzen überschreitende Kontakte bedingen ein jährliches Anwachsen des Bedarfs für Übersetzungen um durchschnittlich etwa 15 %; dies gilt nicht nur für die europäischen Sprachen, sondern ebenso für Japanisch, Koreanisch, Arabisch und Chinesisch. Die Politik – und insbesondere die europäische Politik – hat neue Märkte eröffnet für eine immer differenzierter werdende Palette von Waren und Dienstleistungen. Dies erfordert eine multilinguale Kommunikation und Dokumentation. Eine Änderung hat sich auch für den Gegenstand der Übersetzungen ergeben: während in den fünfziger und sechziger Jahren noch literarische Übersetzungen im Vordergrund standen, dominieren seit den siebziger Jahren technische Informationen das Anwendungsfeld.*

*Im Einsatz befindliche Übersetzungssysteme konnten durch ihre Leistungsfähigkeit zur Anerkennung der Disziplin MÜ beitragen. Erwähnt seien hier nur einige der komplexeren Systeme wie LOGOS, SPANAM und SYSTRAN. Sie wurden in den sechziger und siebziger Jahren in den USA entwickelt.*

*Die Qualität des Outputs eines MÜ-Systems hängt in entscheidendem Maße von den gewählten Sprachpaaren ab, dem Dokumententyp und der relevanten Terminologie. Die meisten operationalen Systeme unserer Tage sind in der Lage, Französisch und Englisch in beiden Richtungen zu übersetzen; die Mehrzahl verfügt ebenfalls über Deutsch und Spanisch als Quell- oder Zielsprache. Daß Englisch die am weitesten entwickelte Quellsprache ist, ergibt sich fast von selbst.*

*Die Vorteile, welche ein maschinelles Übersetzungssystem dem Nutzer bieten kann, sind weitgehend anerkannt: größere Geschwindigkeit bei geringeren Kosten sowie eine erhebliche Konsistenz. Trotz dieses Umstandes muß noch sehr viel Arbeit in der MÜ geleistet werden; dies betrifft sowohl die linguistischen Grundlagen als auch die Software und Benutzerfreundlichkeit.*

*Ein stetiges Anwachsen der Forschungsbemühungen in den letzten paar Jahren ist insbesondere in Japan zu verzeichnen. Jeder größere Computerhersteller entwickelt Systeme zur Übersetzung vom Englischen ins Japanische und umgekehrt. Das wichtigste Forschungsprojekt in Europa ist das EUROPEAN TRANSLATION SYSTEM (EUROTRA). Es unterscheidet sich von bisherigen Systemen, die hauptsächlich einen wörterbuchbasierten Ansatz verfolgen, durch differenzierte linguistische Analyse, Synthese und Transfer in alle Amtssprachen der europäischen Gemeinschaft.*

## EUROTRA – ein Beispiel europäischer Zusammenarbeit

EUROTRA ist ein vom Rat der europäischen Gemeinschaft beschlossenes Forschungs- und Entwicklungsprogramm zur Realisierung eines prototypischen Systems, das die maschinelle Übersetzung zwischen den Amtssprachen der EG ermöglichen soll. Es ist das einzige Forschungsprojekt, bei dem alle 12 Mitgliedsstaaten aktiv mitarbeiten. Europaweit sind etwa 200 Forscher tätig. Die Organisation liegt in den Händen der Kommission. Diese finanziert auch die zentralen Kosten in den Bereichen Linguistik und Software; sie leistet einen Beitrag zu den Kosten der nationalen Komponenten.

Die mit dem Projekt verfolgten Ziele sind allgemein die Förderung der Linguistik und im besonderen die Förderung des Wissenstransfers in der EG. 1990 soll ein vorindustrieller Prototyp zur Verfügung stehen, mit dessen Hilfe alle 9 Amtssprachen analysiert und generiert werden können. Dabei gilt es den Transfer für 72 (!) Sprachpaare auszuarbeiten.

Der Wortschatz und die Fachgebiete, mit denen EUROTRA arbeitet, sind eingeschränkt (Verwaltungs- und technische Texte, Informationstechnologie, Satellitenkommunikation). Als Hardware dienen Microrechner (SUN, HP, VAX), auf denen das Betriebssystem UNIX installiert ist. Die Grundlage für das Vorgehen bilden aktuelle linguistische Formalismen. Dabei bietet EUROTRA den Vorteil, daß neue Erkenntnisse in der Linguistik unmittelbar eingearbeitet werden können. Die Compilierung linguistischer Regeln erfolgt mit der KI-Programmiersprache PROLOG.

*EUROTRA: Das einzige Forschungsprojekt aller EG-Länder.*



*Die Phasen  
von EUROTRA*

In der Vorbereitungsphase (1978-1984) konnten Fortschritte bezüglich der Stimulierung der Forschung auf dem Gebiet der Computerlinguistik und des innereuropäischen Wissenstransfers erzielt werden. Die Forschungsphase (1985-1988) brachte zwar eine „explizite Definition“ des weiteren Ziels, nämlich der Entwicklung eines Prototyps. Insgesamt ging es aber aus organisatorischen Gründen nur in sehr kleinen Schritten voran (einige Länder haben sich erst in jüngerer Zeit dem Projekt angeschlossen, das Team der Kommission verfügte teilweise nicht über hinreichend qualifiziertes Personal).

### Die deutsche Komponente EUROTRA-D

*EUROTRA-D: Entwicklungs-  
schwerpunkt Saarbrücken*

Die Planung und der Entwurf der deutschen Komponente von EUROTRA geschieht am Institut der Gesellschaft zur Förderung der angewandten Informatioforschung e.V. (IAI) in Saarbrücken, Martin-Luther-Str. 14. Das Saarland und der Bundesminister für Forschung und Technologie tragen 52% der Kosten. Aufgabe der Saarbrücker Forschungsgruppe (etwa 20 wissenschaftliche Mitarbeiter) ist die Entwicklung der deutschen Analyse- und Synthesekomponenten sowie der Wörterbücher. Zudem hat das IAI zentrale Aufgaben zur Software-Konzeption und -Entwicklung im Auftrag der EGK übernommen. Die Transferwörterbücher (von anderen offiziellen Sprachen ins Deutsche) werden in Zusammenarbeit mit der Universität Bonn erstellt. Unterstützende Forschungsprojekte finden an den Universitäten Berlin, Bielefeld, und Stuttgart statt.

### Die Arbeitsweise von EUROTRA

#### Das Prinzip der linguistischen Repräsentationsebenen

*Die Grundidee: Aufspaltung der  
Übertragungsvorgänge in kleinere  
Zwischenschritte*

Die Grundidee von EUROTRA besteht darin, die Übersetzung von der Quell- in eine Zielsprache in mehrere kleinere Transfervorgänge aufzuspalten. Zu diesem Zwecke nimmt man verschiedene linguistisch motivierte Repräsentationsebenen an. In EUROTRA heißen diese

EMS – EUROTRA Morphological Structure,  
ECS – EUROTRA Configurational Structure,  
ERS – EUROTRA Relational Structure,  
IS – EUROTRA Interface Structure.

Die EMS ist eine nur theoretisch angenommene Ebene, die zu einem späteren Zeitpunkt realisiert werden soll. Den gesamten Übersetzungsvorgang muß man sich dann als Folge von einfachen Übersetzungsvorgängen zwischen diesen Ebenen vorstellen. Formal läßt sich dies so ausdrücken:

$$S1 \dots ECS(Q) \dots ERS(Q) \dots IS(Q) \dots IS(Z) \dots \\ ERS(Z) \dots ECS(Z) \dots S2$$

(wobei S=Satz, Q=Quellsprache, Z=Zielsprache)

Diese Vorgehensweise macht EUROTRA zu einem multilingualen Übersetzungssystem, bei dem für jeweils 8 Quellsprachen Transferkomponenten erstellt werden müssen. So für das Deutsche etwa:

IS(Q) → IS(Deutsch) für

(Q) {Engl, Frz, Span, Port, Niederl, Griech, Dän, Ital.}

#### Regelformalismus und Syntax

*Ein Grammatikformalismus für  
alle Gemeinschaftssprachen*

Von allen Sprachgruppen wird ein einheitlicher Grammatikformalismus verwendet, der von einem zentralen Team entwickelt wurde. Er folgt den Prinzipien der sogenannten Unifikationsgrammatiken.

Die Repräsentationssprache für die erwähnten Ebenen von Eurotra ist der C,A,T-Formalismus (Abk. für Constructor, Atom, T-rule). Dieser bedient sich einer Reihe unterschiedlicher Regeltypen. Sie haben allesamt die für die PROLOG-Darstellung charakteristische Form



Die Repräsentationen werden durch Grammatikregeln bestimmt, welche Generatoren heißen. Ein Generator besteht aus 2 Arten von Regeln: a-Regeln und b-Regeln. b-Regeln haben die Aufgabe, grammatikalische Strukturen aufzubauen. Angenommen, wir wollen den folgenden Satz repräsentieren:

„Der Eigentümer verlangt die Herausgabe der Sache“.

Aus der Schulzeit kennt man noch traditionelle „Satzteile“ wie „Subjekt“, „Prädikat“, „Objekt“. Man würde sagen „Eigentümer“ ist das Subjekt, „verlangt“ das Prädikat, „Herausgabe“ ein Objekt.

Die Linguisten geben sich nun nicht mit dieser Einzelbetrachtung zufrieden. Sie sagen: Will man wirklich funktionale Einheiten im Satz identifizieren, die die syntaktische Struktur in differenzierter Weise widerspiegeln, so muß es sich dabei schon um größere und flexiblere Kategorien handeln.

Solche Kategorien heißen „Phrasen“. Eine Phrase, in der das Substantiv die wichtigste Einheit bildet, heißt „Nominalphrase“ (np); „der Eigentümer“ ist eine solche Phrase. Diese Phrase besteht aus einem Artikel und einem Substantiv, im Englischen „determiner“ (det) und „noun“ (n).

Die Nominalphrase könnte aber auch durchaus anders aussehen; es leuchtet unmittelbar ein, daß es sich funktional um die gleiche Phrase handeln muß, wenn diese lautet „ein Eigentümer“,

oder

„ein vermeintlicher Eigentümer“,

oder

„ein vermeintlicher Eigentümer der Sache“.

Die Nominalphrase muß in solchen Fällen eben etwas abgeändert bzw. um zusätzliche Elemente erweitert werden; für die genannten Beispielphrasen wären dies ein Adjektiv und eine weitere Nominalphrase.

Der Gedanke funktionaler Kategorien, die unterschiedlichste Elemente oder sogar weitere komplexe Kategorien in sich aufnehmen, kann weitergesponnen werden: So gibt es auch Verbalphrasen (vp), Adverbialphrasen (ap), Präpositionsphrasen (pp), etc. Die Verbalphrase in unserem eingangs genannten Beispielsatz lautet „verlangt die Herausgabe“.

Sie besteht aus einem Verb (v) und einer Nominalphrase.

Wenn „s“ Satz bedeutet, dann können wir jetzt schon mit einigen b-Regeln eine kleine Grammatik aufbauen, die in der Lage ist, den Beispielsatz zu repräsentieren:

(1) s. [np, vp]

(2) np. [det, n]

(3) vp. [v, np]

Die b-Regeln (1)-(3) sind in PROLOG-Notation geschrieben; da sie beschreiben, wie eine Phrase (bzw. ein Satz) aufgebaut ist, kann man sie auch als „Phrasenstrukturregeln“ bezeichnen.

a-Regeln sind dazu da, für bereits vorhandene b-Regeln Bedingungen zu formulieren, die bei ihrer Anwendung gelten sollen. Für die Regel (2) könnte man beispielsweise festlegen, daß der Artikel (det) und das Substantiv (n) in Kasus, Numerus und Genus übereinstimmen müssen:

np. [ ( { det, gen=G, kas=K, num=N } ), ( n, { gen=G, kas=K, num=N } ) ]

Die Überführung von Repräsentationen einer Ebene in solche einer anderen vollziehen schließlich t-Regeln, genannt translator. In ihrer einfachsten Form drückt eine solche t-Regel die Beziehung zwischen zwei atomaren Objekten aus. Ein Deutsch-Englisch-Translator könnte z.B. so aussehen:

(gov, {lu=gehen}). [] => (gov, {lu=go}). []

*Die „Translator“ führen von einer Repräsentation zur anderen*



Komplexere t-Relationen werden rekursiv auf Subobjekten definiert. Sie erlauben die Zerlegung eines Quellobjekts in seine Bestandteile und die Rekonstruktion eines Zielobjekts.

## Die Repräsentationsebenen im einzelnen

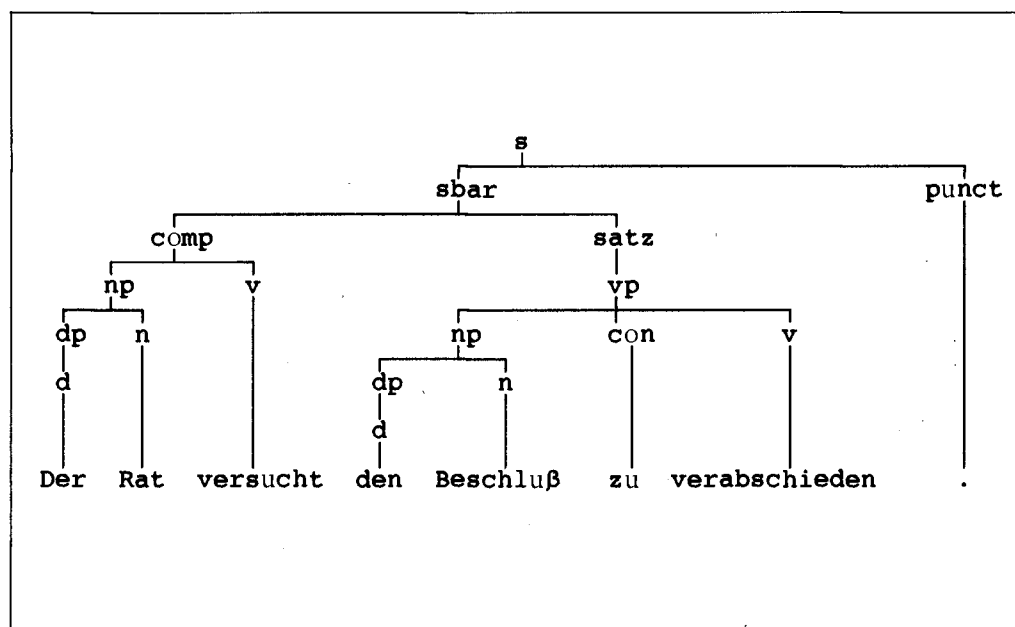
### Die Konstituentenstruktur (ECS)

Die ECS beschreibt die Oberflächensyntax einer Sprache. Sie folgt den Konventionen der sog. x-bar Syntax (auch x oder x') nach Jackendoff (1977) und Gazdar et al. (1985). Als Teiltheorie gehört sie in die Gruppe der Weiterentwicklungen von Noam Chomskys Transformationsgrammatik. Die x-bar Syntax geht davon aus, daß in Phrasenstrukturregeln die rechte Seite einen lexikalischen Kopf enthalten muß, der zusammen mit anderen Elementen im Satz eine niedrigere kategoriale Ebene einnimmt als die linke Seite:

$$x_n \rightarrow \dots x_{n-1} \dots$$

*Der „Darstellungsbaum“ in der  
deutseben Komponente von  
EUROTRA*

Die verschiedenen Ebenen nennt man Projektionen des Kopfes der Konstruktion. Die höchste Ebene ist die maximale Projektion  $x_{max}$ . Die deutsche Komponente von EUROTRA verwendet eine etwas vereinfachte Form dieser x-bar Konvention. Ebenfalls eingearbeitet in die Konstituentenstrukturbeschreibung ist die sogenannte Doppelkopfanalyse des deutschen Satzes (Reis, 1985; Lernerz, 1984; Netter, 1986). Sie nimmt eine einheitliche Struktur des Satzes an (Haupt- und Nebensatz) mit einem initialen Knoten namens „comp“, der 2 Positionen enthält. Knoten heißen die Abzweigungen in einem Baum, der zur graphischen Darstellung der Analyse und Synthese eines Satzes dient. Die EUROTRA-Software erzeugt diese Bäume zur Veranschaulichung der Prolog-Regeln. Das folgende Beispiel stellt die Repräsentation eines deutschen Satzes auf der ECS-Ebene dar:



Dabei geht das Programm von unten nach oben („bottom up“) und von links nach rechts vor.

Zunächst werden die ersten lexikalischen Einheiten im Wörterbuch identifiziert, das in einer eigenen Datei abgelegt ist. Die Wörterbucheinträge sind ebenfalls PROLOG-Regeln, bei denen der Rumpf von der leeren Menge (man spricht in diesem Zusammenhang von „leerer Liste“) gebildet wird. In einfacher Form sehen die Einträge für die ersten beiden lexikalischen Einheiten so aus:

```
der = (d, {lu=der, lex=der}). [].
```

```
rat = (n, {lu=rat, lex=rat}). [].
```



Mit „lex“ ist die Möglichkeit gegeben, auszudrücken, daß im Satz eine flektierte Wortform verwendet wird, die von der Grundform abweicht. Bei dem Verb unseres Beispielsatzes würde der Eintrag für *lu* der Infinitiv „versuchen“ sein und für *lex* die finite Verbform „versucht“.

„rat“ ist in der zweiten Regel klein geschrieben, weil PROLOG Großbuchstaben, oder Zeichenfolgen, die mit einem Großbuchstaben beginnen, als Variablen ansieht.

Den obigen Regeln ist ebenfalls zu entnehmen, daß es sich bei den lexikalischen Einheiten um einen Artikel (*d*) und ein Nomen (*n*) handelt. Das Programm sucht nun in der Datei mit der ECS-Grammatik nach einer Regel, in der *d* und *n* vorkommen. Zunächst zeigt sich, daß der Artikel *d* Bestandteil einer „Determiner Phrase“ (*dp*) ist. Dann findet man eine Regel, die besagt, daß *dp* und *n* zusammen eine Nominalphrase (*np*) bilden.

Auf die gleiche Weise wird nach Aufsuchen des Lexikoneintrags „versucht“ herausgefunden, daß das Verb (*v*) und die Nominalphrase unter einem Knoten namens „comp“ stehen.

Entsprechend muß man sich die Vorgehensweise des Programms für den restlichen Satz vorstellen, und im Prinzip geht es bei den anderen Ebenen genauso. Der Unterschied liegt eben darin, daß es sich jeweils um eine andere Grammatik und ein anderes Lexikon handelt. (Anzumerken bleibt noch, daß die Grammatikregeln bei EUROTRA im Endeffekt doch recht kompliziert aussehen, da sie schließlich für jede Art von syntaktischer Realisierung Vorsorge treffen müssen. Dies kann beispielsweise dadurch geschehen, daß eine ganze Reihe von fakultativ oder alternativ auftretenden Elementen eingebaut wird.)

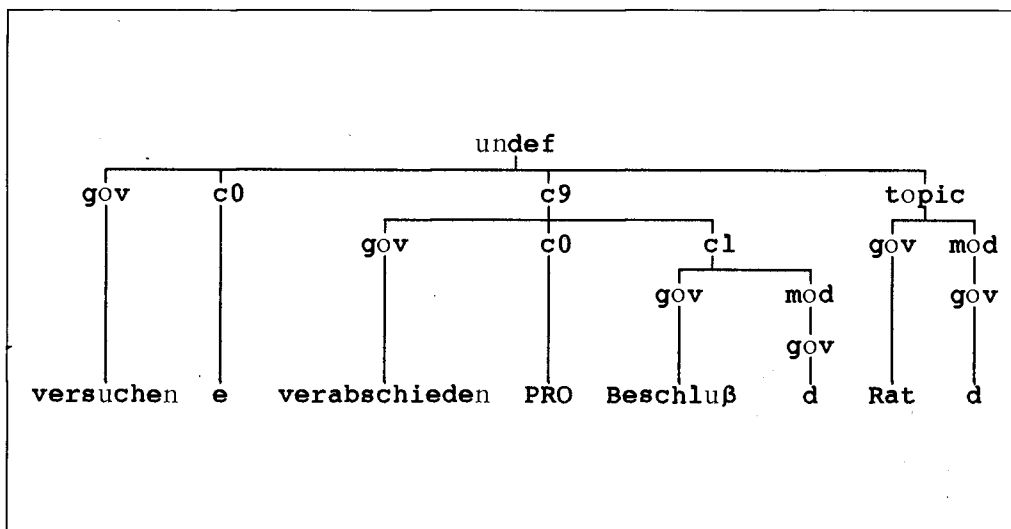
#### Die relationale Struktur (ERS)

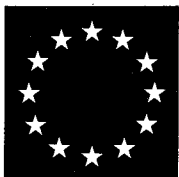
Die relationale Struktur spiegelt die Eigenschaft lexikalischer Einheiten wider, andere Syntagmen zu binden, oder – wie Linguisten auch sagen – zu subkategorisieren. In traditionellen Theorien wird dieses Phänomen Valenz genannt. Die subkategorisierenden Einheiten heißen im Englischen „governor“, ihre valenzgebundenen Glieder „complement“ und die nicht valenzgebundenen Elemente „modifier“. Die Unterscheidung zwischen letzteren beiden geschieht mit Hilfe einer Reihe linguistischer Tests. Im einzelnen subsumiert die Literatur folgende Erscheinungen unter den Begriff der Valenz:

*Was versteht man unter „Valenz“?*

- Die Anzahl der Komplemente, die eine lexikalische Einheit binden kann (quantitative Valenz);
- die Art der gebundenen Elemente (qualitative Valenz).

Die qualitative Valenz unterscheidet noch einmal eine Bindung aufgrund syntaktischer Eigenschaften und eine solche aufgrund semantischer Eigenschaften. Für die ERS-Ebene ist nur erstere relevant, da sie lediglich die syntaktische Dependenzstruktur aufzeigt. Für unterschiedliche Komplemente wurden bei EUROTRA Komplementklassen geschaffen, die mit den Ziffern 0-9 angesprochen werden können.





*Ein Satz ist dann „euroversal“  
abgebildet, wenn nur noch minima-  
le Übertragungsschritte zur Ziel-  
sprache anfallen*

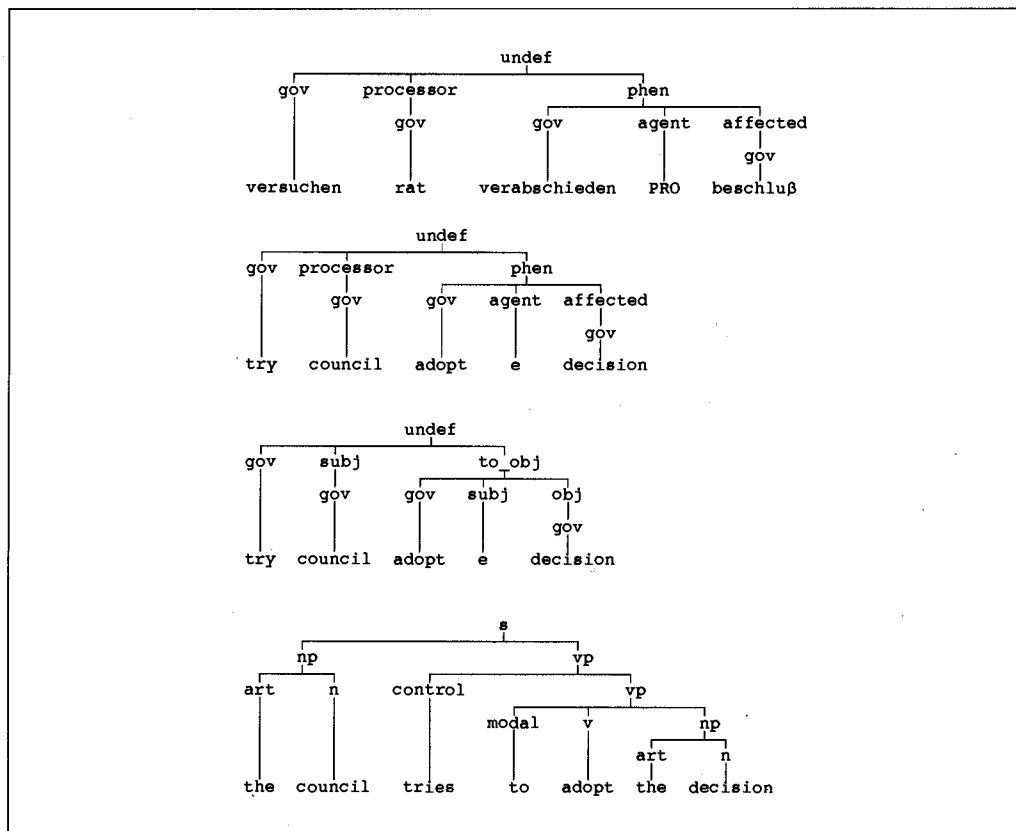
Der vorherige Baum veranschaulicht die ERS-Repräsentation für unseren Beispielsatz. Der oberste Knoten heißt „undef“, da keine bestimmte Textumgebung vorhanden ist. „e“ (für „empty“) ist ein struktureller Platzhalter, der anzeigt, daß an entsprechender Position eine bestimmte syntaktische oder semantische Kategorie, welche das System eigentlich erwartet hat, nicht realisiert wurde.

### Die Interface-Struktur (IS)

Die Interface-Struktur ist nun die Ebene, auf der der Transfer zur anderssprachigen Ziel-IS durchgeführt wird. Sie ist definiert als „Ebene des minimalen Transfers zwischen Quell- und Zielsprache“; dies bedeutet, daß die Transferschritte zwischen den Interface-Strukturen der verschiedenen Sprachen so klein wie möglich sein sollen. Zu diesem Zwecke ist es sinnvoll, die Repräsentation des Satzes so „euroversal“ wie möglich zu gestalten.

Linguistisch kommt der IS die Aufgabe zu, die semantischen Beziehungen zwischen dem Governor der Konstruktion und den Complements darzustellen. Es geht also hier um die andere Art qualitativer Valenz, nämlich die, welche eine Bindung aufgrund semantischer Eigenschaften zum Gegenstand hat. Der Sinn einer solchen semantischen Betrachtung liegt darin, unterschiedliche Lesarten lexikalischer Einheiten voneinander unterscheiden zu können. Dazu werden diesen Einheiten bestimmte Valenzrollen oder „participant roles“ zugeordnet (in den Beispielen heißen diese agent, affected, phenomenon, processor), welche sie semantisch charakterisieren. Die Rollen werden durch Paraphrasierungstests festgelegt und voneinander abgegrenzt.

Die nun folgenden Bäume zeigen nacheinander die Repräsentation des Beispielsatzes auf der deutschen IS-Ebene, der englischen IS-Ebene sowie auf den englischen ERS- und ECS-Ebenen.



### Blick in die Zukunft

*Was wurde bisher erreicht?*

Neben den Analyse- und Synthesekomponenten für die 9 Amtssprachen standen Ende 1988 etwa 50% der Transferkomponenten (mit unterschiedlichem Ausbau) zur Verfügung. Für die IS-Ebene sind ca. 4000 Einträge geschrieben worden; etwa 300 Wörter werden davon zu Testzwecken verwendet.

Mit dem heutigen Labormodell von EUROTRA lassen sich bereits eine ganze Reihe von Sätzen mit schwierigeren linguistischen Problemen übersetzen. Dazu zählen z.B.



Wortstellungsprobleme, Passivbildung, Präfixe, Relativsätze, komplexe Nominalphrasen oder semantische Mehrdeutigkeiten. Können auch juristische Texte wie Urteile oder Vorschriften übersetzt werden? – Die Frage ist prinzipiell mit ja zu beantworten: Man kann die Wörterbücher jederzeit um weitere Einträge ergänzen, und der C,A,T-Formalismus ist so flexibel, daß er unterschiedlichste Korpora repräsentieren kann.

In einer Übergangsphase (1991-92) sollen ausgewählte Sprachpaare weiterentwickelt und linguistische Regeln wie auch die Wörterbücher überarbeitet werden. Ab Beginn der 90er Jahre ist geplant, daß die Industrie die vorhandenen Konzepte und Forschungsergebnisse in praktischen Anwendungen umsetzt. Wie dies im einzelnen aussehen soll vermag im Augenblick noch niemand zu sagen.

Als langfristig angelegtes, grundlagenorientiertes Projekt wird EUROTRA frühestens Mitte der 90er Jahre praktisch wirksame Einsätze finden können. Wie bei jedem Forschungsprojekt, so sind eben auch diesem bestimmte Risiken immanent. Aber trotz solch ungewisser Zukunftsaussichten wird die maschinelle und computergestützte Übersetzung in den nächsten Jahren eine wichtige Rolle bei den Anstrengungen der EG zur Überwindung der Sprachbarrieren spielen. Das Projekt EUROTRA ist ein Beweis für den politischen Willen der Mitgliedsstaaten, ein zentrales Problem der Kommunikation gemeinsam in Angriff zu nehmen.

## Literatur

- Abraham, W. (Hrsg.):* Erklärende Syntax des Deutschen; Tübingen 1985; (Studien zur deutschen Grammatik 25);
- Arnold, D. et al.:* The Eurotra Reference Manual, Release 6.0; edited by Ewin Valentini and Karsten Stroerup, Utrecht 1989;
- Chomsky, N.:* Lectures on Government and Binding; Foris Publications, Dordrecht-Holland/Cinnaminson-USA 1981
- Fanselow, G., Felix, S., W.:* Sprachtheorie; UTB-Taschenbuch, Band 1 und 2, Tübingen 1987;
- Gazdar, G., Klein, E., Pullum, G., Sag, I.:* Generalized Phrase Structure Grammar; Oxford (Engl.) and Cambridge (Mass.) 1985;
- Jackendoff, R.:* X-bar Syntax, A Study of Phrase Structure; The MIT Press, Cambridge, Mass. and London, Engl. 1977;
- Lenerz, J.:* Diachronie Syntax: Verb Position and COMP in German; in: Toman 1984;
- Lyons, J.:* Semantics, Band 1 und 2; Cambridge University Press, Cambridge, London, New York, Melbourne 1977;
- Netter, K.:* Getting Things out of Order. An LFG Proposal for the Treatment of German Word Order; Coling Proceedings 1986;
- Polenz, P. von:* Deutsche Satzsemantik; Verlag Walter de Gruyter, Berlin/New York 1985;
- Reis, M.:* Satzeinleitende Strukturen. Über COMP, Haupt- und Nebensätze, w-Bewegung und Doppelkopf-analyse; in: Abraham 1985.
- Schmidt, P.:* Valency Theory in a Stratificational MT-System; Coling Proceedings 1986;
- Schmidt, P.:* Konstituenz und Dependenz in der deutschen Komponente des MÜ-Systems Eurotra-D; in: Sprache und Datenverarbeitung, 11. Jahrgang 1987, Heft 1;
- Sharp, R.:* CAT-2 – Implementing a formalism for multi-lingual machine translation; in: Proceedings of the second international conference on theoretical and methodological issues in MT, Carnegie Mellon University, Pittsburgh 1988;
- Steiner, E.:* Generating Semantic Structures in Eurotra, in: Coling Proceedings 1986;
- Steiner, E., Schmidt, P., Zelinsky-Wibbelt, C.:* From Syntax to Semantics – Insights from Maschine Translation; Pinter Publishers, London 1988;
- Steiner, E.:* Zur Zuweisung satzsemantischer Rollen im Rahmen des MÜ-Projekts EUROTRA; in: Klenk/Scherber/Thaller, Computerlinguistik und philosophische Datenverarbeitung; Georg Olms Verlag, Hildesheim/Zürich/New York 1987
- Toman, J.:* Studies in German Grammar; Dordrecht-Holland/Cinnaminson-USA 1984 (Studies in Generative Grammar 21) Steiner et al., 1988, S.2