Arbeitspapiere des Sonderforschungsbereichs 340
"Sprachtheoretische Grundlagen der Computerlinguistik"
Bericht Nr. 95

Ein HPSG-Fragment des Deutschen
Teil 1: Theorie

Erhard Hinrichs, Detmar Meurers, Frank Richter,
Manfred Sailer, Heike Winhart

April 1997
Universität Stuttgart • Universität Tübingen
IBM Deutschland

Inhaltsverzeichnis

KAPITEL 1

Ein HPSG-Fragment des Deutschen: Vorwort

Erhard W. Hinrichs*

In der Projektphase 1995-97 des SFB 340 ist im Teilprojekt B8 ein Syntaxfragment des Deutschen im Grammatikformalismus der Head-Driven Phrase Structure Grammar (HPSG) spezifiziert und implementiert worden. Die Implementierung erfolgte im ConTroll System, das für die Verarbeitung von HPSG-Grammatiken vom Teilprojekt B4 in enger Kooperation mit dem B8-Projekt entwickelt worden ist.

Ziel des Fragmentprojekts war es, ein empirisch und theoretisch anspruchsvolles Syntaxfragment des Deutschen zu erstellen. Dabei sollten einerseits die in den Syntaxprojekten A1 und A3 im SFB 340 erarbeiteten empirischen Generalisierungen zum Aufbau des Verbalkomplexes und zur Vorfeldbesetzung einfließen. Zum anderen sollten die in den Projekten des A Bereichs verwendeten GB-theoretischen Annahmen, etwa zur Theorie langer Abhängigkeiten, auf ihre Tragfähigkeit und Integrierbarkeit in eine auf HPSG basierende Syntaxtheorie überprüft werden.

Das implementierte Syntaxfragment umfaßt ein sog. *Basisfragment*, das ein syntaktisches Grundgerüst für die Erschließung vertiefter Phänomenbereiche bereitstellt, und die detaillierte Ausarbeitung zweier Phänomenbereiche, die für das Arbeitsprogramm des SFB 340 insgesamt eine zentrale Rolle spielen: die syntaktische Analyse des Verbalkomplexes und der langen Abhängigkeiten.

* Email: eh@sfs.nphil.uni-tuebingen.de

Im einzelnen deckt das Fragment folgende syntaktische Phänomene ab:

- Basisfragment
  - Struktur einfacher Sätze: V1, V2, VL
  - Nominales System
    - DPs:
      - Personalpronomen
      - einfache Eigennamen (*Karl*)
      - artikellose Nominalphrasen (Plurale, Massennomina)
      - Determinatoren (definit, indefinit, interrogativ)
      - Quantoren (*wenigstens ein Fahrrad*)
      - pränominaler Genitiv (*Karls Schnupfen*)
      - Adjektive (*rosa, rot, dumm*)
    - PPs (*für ihn, in die Stadt*)
  - Verbales System
    - Adverbien (*dort, heute*)
  - (einfache) Koordination
    - DP (*Peter und Paul, die dumme Frau und der schöne Mann*)
    - S (*Maria singt und er lacht*)
    - ATB-Koordination (*das Buch kauft Peter und liest Maria*)
    - SLF-Koordination (*das Buch hat Peter gekauft und hat es Maria geschenkt*)
- Phänomenbereich I: Verbkomplex
  - kohärente Konstruktion
    - Linearisierung
      - Oberfeldkonstruktion
      - Zwischenstellung
    - Statusrektion
      - Ersatzinfinitiv
      - Ersatz-zu-Infinitiv
  - inkohärente und dritte Konstruktion
  - Partikelverben
  - Vorgangspassiv
- Phänomenbereich II: Lange Abhängigkeiten
  - Schnittstelle zu einer Illokutionstheorie
    - Aussagesätze
    - Fragesätze
    - Exklamativsätze

    - eingebettete vs. uneingebettete V2, V1 und VL Sätze
  - Komplexe Sätze
    - finite Komplementsätze
    - nichtfinite Komplementsätze
    - Adverbialsätze
    - Integrierte Verberstparenthesen
  - Topikalisierung
    - kurze Topikalisierung
    - lange Topikalisierung
  - Fragesätze
    - Entscheidungsfragen
    - W-Fragen
      - einfache w-Fragen
      - Mehrfach w-Fragen
      - Echo w-Fragen
      - eingebettete vs. Matrix w-Fragen
  - Inselbeschränkungen für lange Bewegung
    - lange Topikalisierung
    - lange w-Bewegung
    - Komplementsätze unter Brückenverben vs. Nichtbrückenverben
  - Wortstellung im Mittelfeld und Nachfeld
    - Linearisierungstheorie der Wortstellung im Mittelfeld (Scrambling)
    - Linearisierungstheorie von einfachen Extrapositionsdaten

Neben der empirischen Abdeckung ist das Syntaxfragment auch aus theoretischer Sicht interessant. Um Wortstellungsphänomene des Deutschen, speziell das Scrambling und die Extraposition, die sich in einer mono-stratalen Phrasenstrukturgrammatik nicht befriedigend behandeln lassen, adäquat analysieren zu können, basiert das Fragment auf einer sog. *Linearizierungsgrammatik*. Das Fragment greift damit jüngere Entwicklungen in der HPSG auf, in denen verschiedene Spielarten solcher Linearisierungsgrammatiken vorgeschlagen worden sind. Linearisierungsgrammatiken erlauben es, syntaktische Dominanz und lineare Abfolge als interagierende, aber repräsentationell unterschiedene Schichten syntaktischer Information zu spezifizieren. Das Fragment knüpft dabei im Besonderen an Richter und Sailer (1995) an, die gezeigt haben, wie sich auf der Grundlage der in SFB 340 entwickelten *Speciate Reentrant Logic* (King

1994), die auch dem von B4 für die Fragmentimplementierung bereitgestellten ConTroll System als logische Grundlage dient, ein Linearisierungsmodul für HPSG-Grammatiken spezifizieren läßt. Für die Bedürfnisse des Fragments ist die Linearisierungstheorie von Richter/Sailer hinsichtlich der für die Spezifikation von Wortstellungsprinzipien notwendigen (rekursiven) Relationen allerdings signifikant erweitert worden. Richter und Sailer (1997) haben dazu eine formale Sprache, RSRL, entwickelt, die die *Speciate Reentrant Logic* Kings um Variablen, Quantoren und Relationen erweitert und die es erlaubt, Linearisierungsprinzipien in direkter Entsprechung zu den zugrunde liegenden linguistischen Generalisierungen zu formulieren.

Neben der Behandlung von Wortstellungsphänomenen stellt die RSRL eine wesentliche Voraussetzung dafür dar, daß die Integration von einer an Einsichten des GB-Frameworks orientierten Theorie der langen Abhängigkeiten in sprachwissenschaftlich befriedigender Weise in der HPSG umgesetzt werden kann. Die Tatsache, daß eine Linearisierungsgrammatik auf der Basis allgemeiner formaler Grundlagen der HPSG definiert und in ConTroll implementiert werden konnte, bestätigt das Ergebnis aus (Richter und Sailer 1995), daß die Behandlung von Linearisierungsregeln nicht prinzipiell separate außerlogische Mechanismen erfordert.

Der vorgelegte Bericht hat zum Ziel, die empirischen und theoretischen Grundlagen zu erörtern, die für die Erstellung des Basisfragments und der beiden vertieften Phänomenbereiche von Bedeutung sind.

Das Kapitel “Die Satzstruktur des Deutschen und die Behandlung langer Abhängigkeiten in einer Linearisierungsgrammatik. Formale Grundlagen und Implementierung in einem HPSG-Fragment” von Frank Richter beschreibt die relationale Erweiterung der *Speciate Reentrant Logic*, SRL, wie sie von (Richter und Sailer 1997) vorgeschlagen worden ist. Das Papier zeigt, wie die RSRL als formales Ausdrucksmittel für die Behandlung der Satzstruktur und der langen Abhängigkeiten im Rahmen der Linearisierungsgrammatik eingesetzt wird, die dem HPSG-Fragment zugrunde liegt. Für den Phänomenbereich der langen Abhängigkeiten wird gezeigt, wie die im Rahmen der RSRL relational formulierte Theorie Inselbeschränkungen auszudrücken vermag, die in der GB-Tradition als Grammatikalitätsfilter für lange Abhängigkeiten definiert worden sind.

Das Kapitel “Statusrektion und Wortstellung in kohärenten Infinitkonstruktionen des Deutschen” von Walt Detmar Meurers stellt die Theorie des Verbalkomplexes dar, die im HPSG-Fragment erarbeitet worden ist. Das Hauptaugenmerk wird dabei auf Fragen der Wortstellung und

der Statuszuweisung gelegt. Auf der Grundlage einer detallierten empirischen Aufarbeitung dieser Phänomene wird eine HPSG Analyse vorgestellt, die Ausnahmen bei der Statusrektion und die Lizensierung von Wortstellungsvarianten durch lexikalische Klassenbildung von Verben repräsentiert.

Das Kapitel “Adjunkte in einer Linearisierungsgrammatik – Syntax, Semantik und Worstellung” von Manfred Sailer schlägt eine Adjunktanalyse vor, die die syntaktische Position und die semantische Kontribution von Modifikatoren als voneinander weitgehend unabhängige Ebenen grammatischer Information behandelt. Die daraus resultierende HPSG-Analyse hat die Eigenschaft, daß für Sätze mit durch Modifikatoren induzierten Ambiguitäten, wie sie bei Adjunktbewegung und PP-Anbindung auftreten, nur eine einzige syntaktische Struktur erzeugt wird.

Das Kapitel “Die Nominalphrase in einem HPSG-Fragment des Deutschen” von Heike Winhart schlägt eine Analyse für Nominalphrasen des Deutschen vor. Im Mittelpunkt des Interesses steht die Behandlung der Flexionszuweisung, der Nominalisierbarkeit pränominaler Elemente und die Analyse artikelloser Nominalphrasen. Die Analyse basiert auf den Arbeiten Gallmanns zur DP-Struktur im Deutschen (Gallmann 1990), die im Rahmen der GB-Theorie entwickelt worden sind. Es wird gezeigt, wie sich eine solche Analyse in eine HPSG-Grammatik übertragen läßt.

Der vorliegende Bericht beschränkt sich aussschließlich auf die Diskussion der empirischen und theoretischen Grundlagen, die für die Erstellung des HPSG Syntaxfragments von Bedeutung sind. Eine Beschreibung der im ConTroll System implementierten HPSG-Grammatik wird Gegenstand eines separaten SFB-Reports (Courbet, De Kuthy, Meurers, Richter, und Sailer 1997) sein.

1. Literaturverzeichnis

Courbet, E., K. De Kuthy, D. Meurers, F. Richter, und M. Sailer (1997). Ein HPSG-Fragment des Deutschen, Teil 2: Implementierung. Arbeitspapiere des SFB 340, Universität Tübingen.

Gallmann, P. (1990). *Kategoriell komplexe Wortformen. Das Zusammenwirken von Morphologie und Syntax bei der Flexion von Nomen und Adjektiv.* Reihe Germanistische Linguistik 108. Tübingen: Niemeyer.

King, P. J. (1994). An Expanded Logical Formalism for Head-Driven Phrase Structure Grammar. Arbeitspapiere des SFB 340, Nr. 59, Universität Tübingen.

Richter, F. und M. Sailer (1995). Remarks on Linearization. Reflections on the Treatment of LP-Rules in HPSG in a Typed Feature Logic. Magisterarbeit, Seminar für Sprachwissenschaft, Universität Tübingen.

Richter, F. und M. Sailer (1997). A formal interpretation of relations in HPSG. Ms. Universität Tübingen.