

## 西ドイツにおける自然言語処理の現状

—シュトゥットガルト大の研究を中心として

横山 晶一

(電子技術総合研究所)

### 0. はじめに

1984年9月から1985年9月までの1年間、科学技術庁在外研究員としてシュトゥットガルト大学情報科学科 (Universitaet Stuttgart, Institut fuer Informatik) に滞在した。アメリカに比べるとヨーロッパの自然言語処理や人工知能の研究は紹介されることが比較的少ない [岡田 82, 辻井 82など]。本報告では、まず筆者の見た範囲内の西ドイツ全体の様子について述べ、その後で筆者の加わっていた日独機械翻訳プロジェクト SEMSYNの概要と評価について述べる。すでに帰国から半年が経過し、内容がいささか古くなったものもあるが、帰国後に得た情報もできるだけ取り入れることにする。なおシュトゥットガルト大学以外の情報は、いずれも筆者が見聞きした範囲内に限られているが、もし誤りがあれば是非ご指摘いただきたい。

### 1. 西ドイツ全体の様子

西ドイツも日本と同じように大学 (日本と違い、すべて国立 (実際には連邦制をとっているので州立と言うべきか) ) やメーカで自然言語処理や人工知能の研究が活発に行われている。公立の研究機関としてはGID (Gesellschaft fuer Information und Dokumentation 情報ドキュメンテーション研究所) とGMD (Gesellschaft fuer Mathematik und Datenverarbeitung 数理工学データ処理研究所) があるが、後者はどちらかという理論的な側面を研究する機関であり、また前者は研究部門を切り離して後者に統合しようという動きがあり、将来はプロジェクトのマネジメント機関となるものと思われる [Altenmueller 85]。従ってこれらの研究機関ではこの分野の研究がそれほど盛んに行われている訳ではない。また、中立的な機関としてマックスプランク研究所 (Max Planck Institut) やフラウンホーファー研究所 (Fraunhofer Gesellschaft) があるが、前者は物理学、生物学、半導体などの研究が中心で、情報処理とは直接関係がなく、後者はこの分野で

のプロジェクトをいくつか持っているが、いずれも大学との共同プロジェクトである。

その他日本と異なる点は、連邦研究技術省 (Bundesministerium fuer Forschung und Technik 以下BMFTと略記) の指導のもとに、メーカと大学との共同研究プロジェクト (Verbundprojektと呼ばれる。以下VPと略記) が数多く走っていることも日本とは異なる点であろう。以下では筆者の見聞した自然言語処理、機械翻訳のシステムについて個別に簡単に記す。

まず自然言語処理を用いた質問応答システムとして有名なハンブルク大学のHAM-ANSプロジェクトは、残念ながら概要だけの講演しか聞かず、実物に接することはできなかったが、すでにプロジェクトは解散し、現在は知識ベースシステムの研究と、論理型言語の研究プロジェクトがわずかに残っているにすぎないとのことである。図1.1にHAM-ANSプロジェクトの流れを示す。

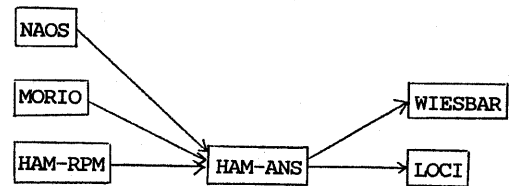


図1.1 HAM-ANSプロジェクトの流れ

機械翻訳のプロジェクトとして西ドイツ国内で筆者の知り得たのは、SEMSYN (シュトゥットガルト大学、日独)、SUSY-DJT (ザールラント大学 (Universitaet des Saarlandes)、日独、独日)、METALの実用化システム (Siemens、独英)、またそのシステムの流れをひくKonstanzでの研究などであるが、Konstanzの研究については実際に見聞きする機会がなかった。SEMSYNについては後に触れることとして、ここでは他の2つについて筆者が見聞したことをごく簡単に記す。

まずSUSY-DJTは、従来ザールラント大学で行われてきた機械翻訳システムSUSYなどを用いて、日本語とドイツ語との間の機械翻訳を行おうというプロジェクトである。筆者がザールラント大学を訪れた時点 (1984年12月) では、

SUSY-DJTについての本格的な研究概要やデモンストレーションについて見聞する機会や時間はなく、SUSY-Iの独英システムのデモンストレーションを見たにとどまるが、SUSY-IはFORTRANとアセンブラで書かれている巨大なシステムで、既にメンテナンスが不可能な段階になっているということであった。SUSY-DJTは、このSUSY-Iの改良版であるSUSY-IIを用いて、英語を中間言語として日本語、ドイツ語間の機械翻訳を行おうとするプロジェクトであるが、その全貌については、公式に発表されたものが(少なくとも筆者の知る限りでは)なく、システムの内容その他についてここで詳しく述べることはできない。なおこのプロジェクトは1985年一杯で終了し、その後はEUROTRA計画の一環として、次第にそちらへ移行していくそうである。

テキサス大学で開発された機械翻訳システムであるMETAL [首藤 81] は、もともとSiemensとの共同プロジェクトであり、実用化を目指した研究がMuenchenのNeu PerlachにあるSiemens中央研究所で行われている。筆者の訪問した時は技術的な話はあまり聞けなかったが、近日中にシステムを発売するとの事であった(「近日」は既に過ぎたが未だに発売の話は聞かない)。Siemensはその他にBMFTのVPで自然言語処理システムを作り始めている。その他のメーカとしては、Nixdorf(次節で少し触れる)、Volkswagenの子会社であるTriumph-Adler(主としてVP)、アメリカのATTの子会社であるSEL(Standard Elektrik Lorenz AG)、またIBM、Boschなどでこの分野の研究が行われている。

## 2. GWA I

ドイツでの人工知能研究の現状を知る一方法として、GWA I (German Workshop on Artificial Intelligence) がある。この会議は毎年一回5日間にわたってドイツ国内の各地で開かれており、ドイツにおける人工知能分野の研究の概要を把握できる。筆者は1984年10月8日~12日にWingst/Stadeで開かれた第8回会議に、途中の2日間のみであったが参加する機会があった。参加者は約120名である。図2.1に会議録からとった目次を示す [Laubsch 85]。

図2.1で (招) は招待講演、(独) とは会議録中でドイツ語のものであり、その他は英語で書かれている。ただし講演発表はすべてドイツ語で

行われた。主な講演内容については会議録を参照していただくこととして、ここでは予稿なしの講演のうち一つのみ取り上げる。図2.1のNixdorfのSavoryのシステムで、バスの故障診断と修理システムである。このプログラムはPrologで約800Kほどのものでデモをした訳ではないので具体的な内容についてはよく分からないものの、「一切自然言語処理を行っていない」

### 知識ベースシステム

- L. Steels: 知識表現システムのための設計の必要性(招)
- P. Maes: 知識ベースオフィスシステムの目標
- F. & B. Puppe: ワーキングメモリを持つ診断的推論

### 自然言語システム

- J. Kilbury: ID/LP文法の直接構文解析にたいするEarly-Shieberアルゴリズムの改良
- W. Frey: いくつかの名詞節の構文および意味
- R. Granger 他: ATLAST: 3レベルの言語解析システム
- M. Gehrke: 問題指向対話における省略の構造(独)

### 画像理解

- B. Radig: 画像理解と人工知能(独)(招)
- H. -J. Novak: 時間的な事象認識に対する関係マッチング戦略

### 認知

- C. Freksa 他: 認知と表現--認知科学における知識表現研究の外観(招)

- A. Kobsa: VIE-DPM: 自然言語対話システムにおけるユーザモデル

### AIプログラミング

- H. Stoyan: 人工知能におけるプログラミングスタイル(招)
- K. v. Luck 他: 制御構造の宣言的表現
- D. Bauer 他: STRUPPI: LISP構造に対するグラフィックブリティプリンタ(独)

### 学習と発見的手法

- Y. Kodratoff: 一般化した例に対する構造化マッチングの使用
- C. -H. Tzeng: 発見的ゲームプレイの数学モデル

### 論理

- W. Emde: SRLの新しいインプリメンテーション(独)
- B. Fronhoefer: ロジックプログラミングの観点に対する再帰的プログラムの改良

- C. Schwind: 人工知能における時間論理(招)

### パネル討論 言語学と人工知能

### 予稿なし

- P. Sturdza: 知識ベースとしてのデータベース: Alpha-Ariesシステム

- L. Luft: 知識表現の方法論的基礎に対する寄与

- K. -J. Engelberg: プロジェクトCon<sup>2</sup> Traiにおける知識表現の原理

- D. Schmedding 他: DORIAN: 意味指向型自然言語質問応答システム

- A. Guenther: マッチと証明による事実認識--fuzzyとPrologインプリメンテーションの比較

- J. Kreich 他: PrologとGKSによる画像認識および画像記述

- S. Savory: FF: 故障発見と修理計画のためのNIXDORFエキスパートシステム

- K. Hanakata: オブジェクト指向型言語CTALKにおける継承とメッセージ送信の制御に関する強力な動作プログラム

### その他

- IBM-PC上でのTLC-Lispのデモ

- SEMSYNのデモ

- HAM-ANSのデモ

- 討論: AI環境でのマスメディアとその困難さ

- LFGとGPSGに関する講義

- BMFT-VPの案内

- 自然言語システムに関する討論

図2.1 GWA I-84のプログラム

と述べたのが印象的であった。しかしながらシステム全体としてはユーザの使い易さをかなり考慮して作られているようであった。LFGとGPSGに関する講演は、これらの概念について基本的なことを解説するものであった。

なおGWA I-85も筆者の帰国後Dassel/Sollingで1985年9月23日~28日に開かれている。図2.2にGWA I-85のうちで自然言語システムのセッションで行われた講演題目のみを示す。ちなみにその他のセッションは、認知科学、知識表現、ロジックプログラミング、エキスパートシステム、自動演繹、ロボティクスとシーン解析である。

- H. Horacek他: データベース応用のための構文解析-意味解析変換
- R. Wirth: シソーラスの半自動拡張(独)
- S. Busemann他: HAM-ANSにおける連辞と完全動詞を持つ文の表現と処理
- G. Retz-Schmidt: 事象認識と言語表現におけるスクリプトと予測の役割
- J. Krems他: 言語能力の指標(独)
- J. Mueller他: STATESY-発見的探索アルゴリズムのためのテストシステム(独)
- J. Kilbury: KIT/NASEVにおけるGPSGに基づく構文解析(独)
- S. Knoepfler: Con<sup>3</sup>Traのアーキテクチャ(独)
- J. Kilbury: 言語のバリエーションと構文解析(独)
- 講演者不明: フランス語に対するLFGの構文断片(独)
- D. Stroele: 特定のデータベース応用のための処理階層を持つロジックプログラム(独)
- D. Heyer: Triumph-AdlerにおけるAIプロジェクト(独)
- R. Brueckel他: KOCOS-記号処理VLSIレイアウトのコンパクト化のためのエキスパートシステム計画
- R. A. Mueller: オブジェクトクラスにたいするエキスパートシステム

図2.2 GWA I-85の自然言語システムセッションのプログラム

### 3. シュトゥットガルト大のプロジェクト

シュトゥットガルト大学の情報科学科には、計算機理論、数値解析、プログラミング言語、データベース、マン-マシンコミュニケーション(MMC)、ハードウェアの6つの講座があるが、筆者の参加したのはMMCである。ここでは図3.1に示すような5つのプロジェクトが走っている。

- ・INFORM Triumph-AdlerとのVP
- ・SEMSYN BMFTプロジェクト
- ・盲人のためのプログラミング環境 DFGプロジェクト
- ・自然言語処理 IBM
- ・CTALK

図3.1 シュトゥットガルト大情報科学科MMCの主なプロジェクト

SEMSYN以外のものについてごく簡単に紹介する。まずINFORMは、人間にとって最適なプログラミング環境を研究するプロジェクトで、主としてLISPの環境下でグラフィックディスプレ

イ上にマルチウィンドウやプリティプリントなどの画面を供給するものである。この講座の中のプロジェクトとしては人数、金額ともに最も大きいものである。その中の成果の一つが図2.1のD. Bauer他の論文である。

盲人のためのプログラミング環境のプロジェクトは、文字通りの意味であるが、キーボードのところには点字を出力したり、点字プリンタを備えるなどの設備や、通信回線を通じて点字情報を送るなど、盲人がプログラミングを行ったり、コンピュータを使ったりできるような環境を整えつつある。

IBMの自然言語処理プロジェクトは1985年の後半にスタートしたばかりで、具体的な内容は明らかではないが、自然言語による質問応答システムを構築するものようである。

CTALKはCの上に開発したオブジェクト指向型言語で、従来のものに比べて種々の特徴を持っているが、より詳しい内容については[横山 85、横山 86]を参照されたい。

## 4. 日独翻訳プロジェクトSEMSYN

### 4.1 概要 [Laubsch 84、羽中田 85、Roesner 85] 図4.1にSEMSYNの処理の流れを示す。

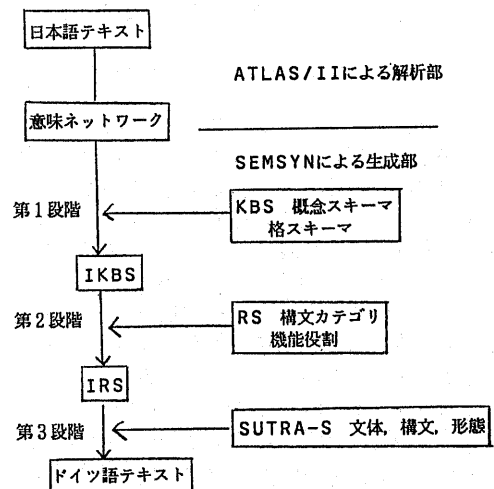


図4.1 SEMSYNの処理の流れ

この図から分かるようにこのプロジェクトの日本語解析部は富士通のATLAS/IIの日本語解析の出力を用いており(図4.1の上部参照)、プロジェクトで実際に行うのはその後の独文生成である。以下この図に沿ってSEMSYNの独文生成を解説する。まずATLAS/IIの上で日本語を

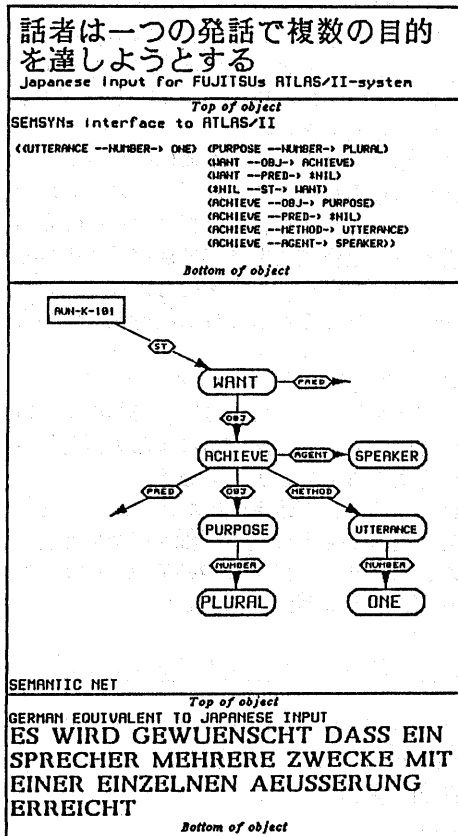


図4.2 日本語文からの意味ネットワークとドイツ語出力

解析して意味ネットワークを作る。これがSEMSYNの入力となる。図4.2に「話者は一つの発話で目的を達しようとする」という日本語文からの意味ネットワークとその図示を示す。この図から分かるように元の意味ネットワークはアークとノードからなる有向の形をしている。

ここから3段階でドイツ語を生成する。第1段階は意味構造から概念構造を作り出す過程である。そのためにあらかじめ用意された知識ベーススキーマ (knowledge base schema、KBS) を用いる。KBSには次の3種類がある。

**格スキーマ (case schema) :** 動詞概念や動作を表わす (図4.2ではWANT、ACHIEVE)。

**概念スキーマ (concept schema) :** 主として名詞概念を表わす (図4.2ではSPEAKER、UTTERANCEなど)。

**関係スキーマ (relation schema) :** アークの部分を取り扱う (図4.2ではPURPOSEなど)。

KBSの構造は、いずれのものも次の4つの部分からなっている。

- ・ **superc:** そのKBSの上位概念を示す。
- ・ **roles:** 格スキーマではFillmore流の格関係 (AGENT、OBJECT、INSTRUMENTなど) と、それから形成されるIKBS (instantiated knowledge base schema) が満たすべき条件を示す。概念スキーマでは、与えられた意味構造でその概念がどのように振る舞うかを示す。
- ・ **transformation rules:** 条件と動作のペアから成り、どのスキーマを適用すべきかということや、rolesの埋め方などを規定する。
- ・ **choices:** IRS (instantiated realization schema) で可能な構文パターンを提示する。

図4.3にACHIEVEとSPEAKERのKBS、ならびにそれから作られたACHIEVEのIKBSを示す。

第2段階ではこのIKBSからドイツ語の表層構造を導き出すための表現方法を示すIRSを作成する。これは次のような過程である。

- (1)ドイツ語動詞の辞書を検索してIKBSに対応した動詞を選び出す。
- (2)動詞以外の意味シンボルに対する辞書項目を決定する。
- (3)RS (realization schema) を検索してIKBSに対応するIRSを作る。

(1)ではIKBSのfillerをチェックして最も適当と思われる動詞を選択する。たとえば“REALISE”という意味シンボルに対しては、INSTRUMENTがPROGRAMという意味シンボルで満たされていれば動詞として“realisieren”でなく“implementieren”を選択するなどの操作を行う。これは(2)についても同様で、たとえば意味シンボル“COMPUTER”は、OBJECTの部分に教育関係の語で満たされている場合には“Informatik”が選ばれるが、その他のコンテキストにおいては“Computer”が選択される。(3)においても、どのrolesが満たされたか、それに対応するfillerは何か、どのようなタイプのテキストを生成するかということに基づいてIRSが作成される。たとえば図4.2の例では意味シンボルWANTのところにはAGENTというアークが出ていないので、受動態

の文とすることが決定され、“es”を仮主語とする文が生成される。RSにおいては一般的に格スキーマが節RS (CLAUSE-RS) に対応し、概念スキーマが名詞句RS (NP-RS) に対応する。図4.4に図4.3のIKBSから生成されたI

RSの例を示す。

第3段階は、このようにして作られたIRSから実際のドイツ語文を生成する過程であるが、この部分はHAM-ANSプロジェクトの言語生成部であるSUTRAを改良して用いている。

- ```
(a) (ACHIEVE
      (superc goal-oriented-act)
      (roles
        (Agent (class animate))
        (Goal)
        (Method (class abstract-object))
        (Instrument (class concrete-object)))
      (choices
        (if <the roles Agent, Goal & Method are filled>
          make
            (a Clause with
              (Subject <NP-realization of Agent-role>
                (verb "erreich_")
                (DirObj <NP--realization of Object-role>
                  (IndObjs
                    (a PP with
                      (Prep (One-of ("durch" "mit" "mitteils"))))
                      (Pobj
                        <NP-realization of Method-role>))))))
              ...))
        ...))

(b) (SPEAKER
      (superc animate-object)
      (specialization
        (Performs-act-for (class Organization)))
      (choices
        (if !Performs-act-for . { TV, Radio _
          ==> "Ansager" (announcer)
        }
        (if !Performs-act-for . { Government, Party _
          ==> "Sprecher" (spokesman)
        }

(c) (an ACHIEVE with
      (Agent (SPEAKER))
      (Object (PURPOSE (number (PLURAL))))
      (Method (UTTERANCE (number (SINGLE))))))
```

図4.3 (a)ACHIEVEのKBS (b)SPEAKERのKBS  
(c)ACHIEVEに対するIKBS

- ```
(a Clause with
  (Subject (a NP with
    (Head "Sprecher")))
  (Verb "erreich_")
  (DirObj (a NP with
    (Features (Numerus Plural))
    (Head { "Ziel" "Zweck" })
    (Adj "mehrere")))
  (IndObjs
    (a PP with
      (Prep (One-of ("durch" "mit" "mittels"))))
      (Pobj
        (a NP with
          (Features (Numerus Singular))
          (head "Aeusserung")
          (Adj "einzeln"))))))))
```

図4.4 IRSの例

このシステムはSYMBOLICSのLisp Machine上で稼動しており、IKBSからIRSへの変換過程は、この機械のオブジェクト指向型言語であるF LAVORsと、プロダクション方式であるAMORDの2方式で実現されている。IKBSからIRSへの変換過程を操作することによって、種々のドイツ語訳を生成することができる。図4.5に、「合成された暗号化関数を用いた共有データベースのアクセス」というタイトルの種々の訳し分けの例 (SEMSYNからの出力)を示す。

タイトル:

Die Verwendung von synthetisierten Kryptographiefunktionen zum Zugang zu gemeinsam benutzten Datenbanken.

文:

Eine synthetisierte Kryptographiefunktion wird zum Zugang zu gemeinsam benutzten Datenbanken verwendet.

焦点を考慮した文:

Zum Zugang zu gemeinsam benutzten Datenbanken wird eine synthetisierte Kryptographiefunktion verwendet.

仮主語を用いた文:

Man verwendet zum Zugang zu gemeinsam benutzten Datenbanken eine synthetisierte Kryptographiefunktion.

図4.5 タイトルの訳し分けの例

4.2 翻訳の際の問題点 このプロジェクトは1983年7月～1985年12月の2年半にわたって行われたが、その際種々の問題点が生じた。その中にはこのシステム固有の問題も含まれているが、その多くは一般の機械翻訳システム、さらには自然言語処理システムにも通じるものである。以下にこれらの問題点を列挙する [Hanakata 86]。なおこのプロジェクトが対象にしたのは情報処理関係の論文タイトルで、約1500について独文生成を行っている。その翻訳の質の評価については次節で述べる。

(1)日本語テキストそのもの問題

入力とした日本語論文タイトルの中には、おそらく英語論文タイトルの人間による直訳と思われるが、専門用語のおかしいもの(例:「第一次述語論理…」)、また人間の目で見ても日本語の意味のとりにくいものがあった。機械翻訳システムの評価を行う場合には入力の質というものも問題にすべきであろう。

(2)日本語意味ネットワーク抽出の誤り

ATLAS/IIの解析誤りである。

(3)日本語文での情報の欠如

日本語からヨーロッパ系言語への翻訳でよく指摘されるのは、日本語になくて相手言語にある情報をどのように扱うかということである。日本語は単数、複数の概念を陽に表わさないことが多いし、冠詞がないが、これらはドイツ語の場合いずれも必要である。

(4)接続詞の問題

日本語解析の場合にも言えることであるが、接続詞「および」や、接続助詞「と」のかかる範囲が問題になるが、これは翻訳の場合でも同じである。図4.6に「と」の意味ネットワークとそれによる接続のタイプを示す。実際のタイトルで調査したところでは、図4.6のうち約90%がタイプ(a)に、8%がタイプ(b)に、残りがタイプ(c)になるという結果が得られている。このシステムでは、実用上の観点から元の言語(すなわち日本語)の曖昧性を保存したまま訳文を出力することを行っている。たとえば「N-1通信網の実験と性能」という入力に対して、ドイツ語訳では“Die Leistung und Erprobung von den N-1-Kommunikationsnetzen”のように「実験」と「性能」を反転した訳文を生成しているが、これによって前半部「の」までの係り方の曖昧性を残したままの訳文を作ることができる。

(5)構文的な問題

このシステムでは意味ネットワークを中間表現としてとっているために、逆に構文的な情報が落ちてしまい、生成された訳文がやや不自然になることがある。

(6)文化の違いによる表現の違い

これには種々のレベルがある。まず品詞のレベルでは、日本語の品詞とドイツ語の品詞が異なる場合がよくある。たとえば「…に関する」というのは英語、ドイツ語ともに前置詞で表現されるし、「…を用いた」というのは英語では“using”というやや前置詞的な表現になるが、ドイツ語では関係代名詞プラス動詞の受け身形(… verwendet werden)にしたり、名詞(Verwendung)を用いたりする。

また、日本語や英語では、特に論文タイトルなどの場合、名詞連続を用いることが多いが、ドイツ語の場合には複合名詞として一語につながらない時には名詞を連続させることができない。従って日本語、英語ともにほとんど逐語訳でよい場合にもドイツ語では名詞間の関係を認定してその間を前置詞で結ばなければならないということが生じる。

時制の違いということも大きい。筆者が経験したもので最も印象的に感じたのは、次のような日本語テキストであった。

「計算機の専門知識のある教師を学校体系の内部で養成する問題」

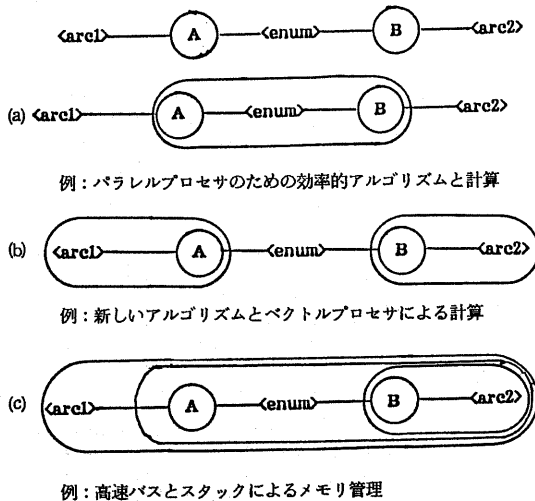


図4.6 「と」の意味ネットワークと接続タイプ

これに対する訳文は次のようになる。

Probleme bei der Ausbildung ueber Lehrer, die speziel les Computerwissen besitzen, innerhalb eines Schulsystems.

この出力に対するドイツ人のコメントは、すでに計算機の専門知識を持っている教師に対してなぜ教育を施さねばならないのかというものであった。これは文化の違いともいえるし、日本語の時制とドイツ語の時制の表現の仕方の違いによるということもできるであろう。

#### (7)代名詞化

日本語の代名詞「それ」や「それら」を英語に訳すときには主格、所有格、目的格のような格変化にのみ注意すればよいが、ドイツ語では性、数、格の一致が必要である。従ってその代名詞が何を受けているのかを同定しなければならない。その他、焦点のズレであるとかサブタイトルの扱いとか、タイトルのレベルでも問題は多い。

4.3 評価 プロジェクト終了にあたり、全体の5分の1にあたる300タイトルを任意に選んでそのドイツ語文の品質を評価することがチュービンゲン大学のゲルマニストグループによって行われた[Ezawa 85]。その評価基準を図4.7に示す。図の(a)と(b)について、それぞれ5段階評価で行っている。

この基準によって300タイトルを評価した結果を表4.1に示す。

[Ezawa 85]における評価は次のようなものであった。

#### (a)翻訳の正確さ

1. オリジナルと翻訳が完全に一致する。
2. オリジナルと翻訳がかなりの程度一致する。
3. オリジナルの情報は伝えている。
4. 一部だけが一致する。
5. オリジナルと一致していない(受け入れ難い)。

#### (b)可読性および文法性

1. スタイル、構文形態ともに完全に正しい。
2. 構文形態的には正しいが、スタイルに誤りがある。逆にスタイルは正しいが、構文形態的に誤りがある。
3. 可読であるが、構文形態的またはスタイル的に重大な誤りがある。
4. ドイツ語として可読とは言いがたい。
5. もはやドイツ語ではない。

図4.7 SEMSYNタイトルの評価基準

(1)形態的には、数少ない、偶然的なエラーを除いては、おおむね良好な翻訳がなされている。

(2)動詞と前置詞の結合に関しては動詞の結合価についてもっと研究することによりプログラムが改善できるのではないと思われる。特に所有格や"mit"の代わりに"von"や"durch"を用いるのはスタイルとしてあまり良いことではない。

(3)英語をそのまま写した語(たとえば"Management", "Scheduling"など)よりもドイツ語固有の語(たとえば"Verwaltung", "Zeitplanung"など)を用いるように辞書を変更した方がよい。

(4)関係代名詞節を用いる(たとえば"Zerlegung, die von einem Modell nicht abhaengig ist")よりもドイツ語特有の語構成を用いる(たとえば"model unabhaengige Zerlegung")方がよい。

(5)「と」の部分の反転による曖昧性保持への言及。

(6)日本語の冠詞の欠如などへの言及。

4.4 感想 SEMSYNプロジェクトは日本人とドイツ人ならびにポーランド人がメンバとなって行われていたが、ネイティブスピーカが複数いて常にディスカッションできるということはこのような機械翻訳プロジェクトの場合必要なことである。ネイティブスピーカ複数というのは母国語に対する感覚が個人個人で微妙に食い違うということからは是非とも実現すべきことである。そのような意味からもこのプロジェクトは興味ある実例を提供したものである。

表4.1 SEMSYNシステムの評価

評点	翻訳の正確さ		可読性、文法性	
	件数	割合	件数	割合
1	50	16.7%	204	68.0%
2	144	48.0%	89	29.7%
3	57	19.0%	7	2.3%
4	49	16.3%	0	0.0%
5	0	0.0%	0	0.0%

## 5. おわりに

西ドイツにおける自然言語処理の現状について、主としてシュトゥットガルト大学のSEMSYNプロジェクトを中心として解説した。筆者の滞独中種々お世話になったSEMSYNグループの羽中田賢治、Joachim Laubsch (現在アメリカHP)、Arkadiusz Lesniewski (現在SEL)、Dietmar Roesner、Claudia Braunの諸氏に感謝します。

### 参考文献

- [Altenmueller 85] Altenmueller, G.H.: Die erzwungene Zellteilung: GID: Service privat, Forschung bei der GMD?, DUZ 18, p.13, 1985
- [Ezawa 85] Ezawa, K., Okoma, S.: Bericht ueber die Evaluation des Testoutputs der SEMSYN-Titeluebersetzung, 1985.
- [Hanakata 86] Hanakata, K., Lesniewski, A., Yokoyama, S.: Semantic Based Generation of Japanese German Translation System -Result and Evaluation-, COLING-86投稿中.
- [Laubsch 84] Laubsch, J., Roesner, D., Hanakata, K., Lesniewski, A.: Language Generation from Conceptual Structure: Synthesis of German in a Japanese/German MT Project, COLING-84 Proceedings, pp. 491-494, 1984.
- [Laubsch 85] Laubsch J. (ed.): GWAI-84, 8th German Workshop on Artificial Intelligence, Wingst/Stade, October 1984, Informatik-Fachberichte 103, Springer Verlag 1985.
- [Roesner 85] Roesner, D.: Generierung von Deutsch aus einer Semantischen Repraesentation, GLDV-Jahrestagung 85, 1985.

[岡田 82] 岡田 直之: ヨーロッパにおける自然言語処理の現状 - グルノーブル大学における機械翻訳を中心として -、情報処理学会自然言語処理研究会資料30-2、1982。

[首藤 81] 首藤 公昭: テキサス大学における機械翻訳、情報処理学会自然言語処理研究会資料28-6、1982。

[辻井 82] 辻井 潤一: ヨーロッパにおける計算言語学の様子、情報処理学会自然言語処理研究会資料31-3、1982。

[羽中田 85] 羽中田 賢治: SEMSYN: 意味表現にもとづく独文生成プロジェクト、情報処理学会第30回全国大会7G-8、1985.3。

[横山 85] 横山 晶一: モノプロセス版CTALK、情報処理学会第31回全国大会1E-10、1985.9。

[横山 86] 横山 晶一、羽中田 賢治、Frank Selbmann: オブジェクト指向型言語CTALKを用いた知識表現の試み、情報処理学会第32回全国大会4M-1、1986.3。