

第5世代の電子計算機に関する 調査研究報告書

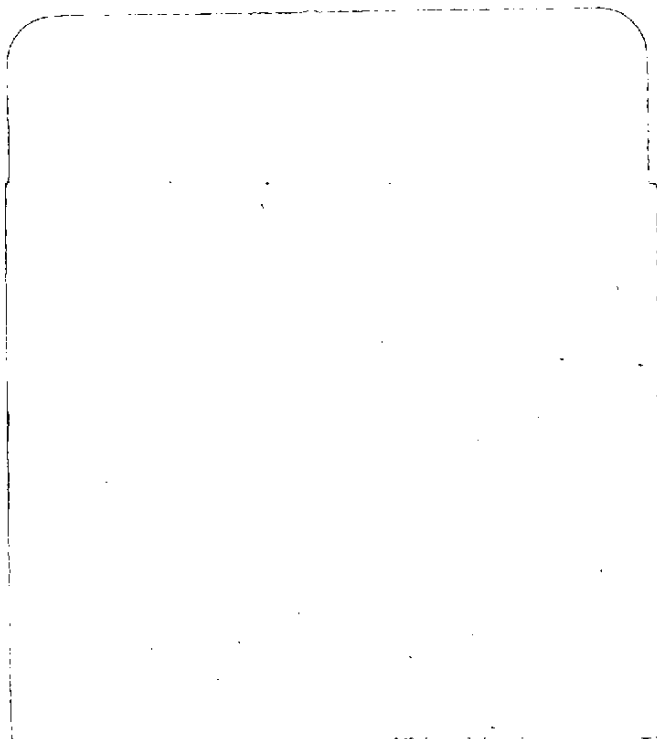
—— 自然言語処理 ——

昭和 55 年 3 月

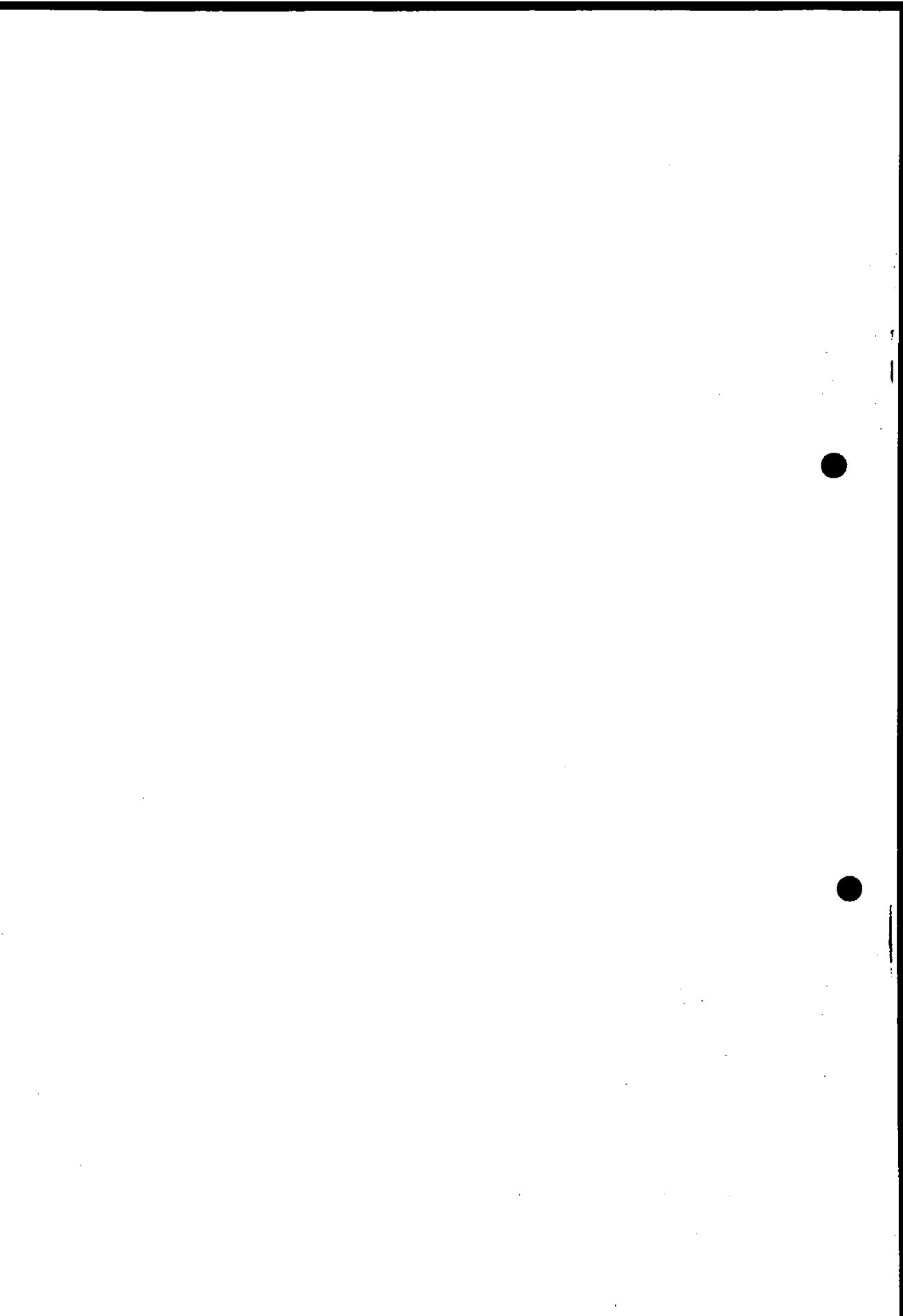


財団法人 日本情報処理開発協会

この報告書は、日本自転車振興会から競輪収益の一部である機械工業振興資金の補助を受けて昭和54年度に実施した「第5世代の電子計算機に関する調査研究」の成果をとりまとめたものであります。







自然言語処理・執筆者

(順不同)

氏 名	所 属
田 中 穂 積	電子技術総合研究所 パターン情報部 推論機構研究室室長
内 田 ユリ子	同 上 主任研究官
佐 藤 泰 介	同 上 技官
松 本 裕 治	同 上
横 山 晶 一	同 上
坂 橋 秀 一	電子技術総合研究所 パターン情報部 音声認識研究室主任研究官
三 国 一 郎	同 上 技官
大谷木 重 夫	電子技術総合研究所 制御部 論理システム研究室技官
池 田 尚 志	電子技術総合研究所 制御部 情報制御研究室主任研究官
寺 津 典 子	富山大学人文学部講師
山 梨 正 明	大阪大学言語文化部講師

自然言語処理報告書

目 次

調査の概要	1
I 自然言語処理システム	1 0
1. はじめに	1 0
2. 構文解析	3 1
3. 知識表現 } 次年度調査	
4. 文の生成 }	
5. 問題解決と推論	5 2
6. 談話理解	6 4
7. 音声理解	1 0 1
8. 応用システム	1 4 4
II 最近の言語学の潮流	2 2 3
1. 生成文法理論の展開	2 2 4
2. 生成文法理論における意味論の展開	2 3 5
3. ヨーロッパにおける言語学の展開	2 3 5
4. Montague 文法	2 4 8

調 査 の 概 要

調 査 の 概 要

本ワーキング・グループでは、大きく二つにテーマを分けて調査を行なった。一つは、計算機による自然言語処理技術の現状がどのようなものであるかを知るための調査である。もう一つは、現代の言語学の潮流を探るための調査である。これには、電総研の外部から、山梨正明大阪大学講師、寺津典子富山大学講師の協力をあおいだ。

(A) 第 I 部 自然言語処理システム

自然言語処理ワーキング・グループ報告書の第一部は、自然言語処理システムの研究の現状の調査と分析で、8章から構成されている。しかし、本調査が2年にわたることを想定して、第3章と第4章は、今回の報告書に含めなかった。また、調査を進める段階で、本年の報告書では、特定のテーマの重要性が認識され、それを集中的に調査することにした。第I部の構成と担当者は次のとおり。

第1章 はじめに (田中穂積・電総研)

第2章 構文解析 (佐藤泰介・電総研)

第3章 知識表現 } 次年度調査
第4章 文の生成 }

第5章 問題解決と推論——Default Reasoning (松本裕治・電総研)

第6章 談話理解 (内田ユリ子・電総研)

第7章 音声理解 (板橋秀一・電総研)
(三国一郎)

第8章 応用システム

8.1 機械翻訳 (池田尚志・電総研)

8.2 質問応答システム (田中穂積・電総研)

第2章では、構文解析技術を述べる。自然言語処理技術のなかで、形態素処理、構文解析については、ほぼ問題点が明らかになり、しかも技術的な進歩がみられる。問題点の多くは、構文解析が、意味処理や、文脈レベルの処理と完全に分離できないことに起因していると思われる。それらを除けば、技術的に確立されてきていると考えてよいだろう。構文解析用のアルゴリズムとシステムについては、人工言語(形式言語)の構文解析手

法を自然言語に応用するために、さまざまな拡張がなされている。これは、文脈自由文法を骨子として、各規則に何らかの手続き（プログラム）を付加し強化することによって、柔軟性をもたせ、文脈依存的な処理を行なわせようとするものである。我が国での研究状況と、日本語の特質が、構文解析にどのような影響を及ぼすかについても調査した。

第5章では、最近、不完全な知識の下で行なわれる推論の研究が進んでいるため、それに焦点をあてて調査した。これは default reasoning とよばれ、質問応答システムの問題解決を行なう場面で、ますます重要になってきている技術である。調査によればこれは、(i)変数割当てにおける default, (ii)閉じた世界における default, (iii)階層表現における default, (iv) Frame問題, などに分類整理される。(i)は変数に値が陽に割当てられていない時、変数にあらかじめセットされている常識値を用いる推論である。(ii)の閉じた世界では、ある事実の成立が証明不可能な時には、その事実の否定が成立するというものである。(iii)は、階層構造における同レベルの概念間に disjoint性が成立するという default reasoning である。(iv)は、箱の移動により、箱の位置は変化するが、箱の色は変化しないこと等をどのように扱うかという問題である。以上をどのようにして形式化するかは default reasoning の最近の研究により、部分的に明らかになってきている。これらを説明する。

第6章の談話理解では、主として、Halliday 等の最近の研究を調査することにした。今回は、Halliday の用語の説明と cohesion (結合関係)に関連する基本的な考え方を説明する。談話理解の問題は、今後増々重要になると思われるが、明確な理論とモデルが模索されている段階にある。しかも語用論に属す事柄が極めて多いのが特徴である。今後、現実の言語現象に幅広くあたり、分析する必要があると思われる。

第7章の音声理解の項では、特に70年代の米国における音声理解プロジェクトの結果を中心に調査した。70年前半の自然言語理解システムの研究の成果は、音声認識の研究者に希望を与えたように思われる。音声認識に、言語情報が必要であるとの認識は基本的には正しいといえるが、米国における音声理解プロジェクトの結果は、言語情報の利用が十分有効で

あるという段階にまだないということを改めて多くの研究者に認識させる結果になった。そして、むしろ、もう一度音響、音声学的レベルの処理を見直す方向が示唆されていることが明らかにされた。

第8章では、自然言語の応用システムについての調査を行なった。8.1では機械翻訳システム開発の歴史と現在の研究動向が概観される。質問応答システムと同様に、対象分野を限定することにより、実用化数歩のシステムがすでに実現可能な時期にきている。80年代は、そうした方向の研究が、進展すると思われる。しかし、機械翻訳に関していえば、かな漢字変換と同様に、当面、人の介入による post editing をどのようにスマートに行なうかが、短期的なレベルでは重要になるとと思われる。

8.2の質問応答システムに関する調査では、自然言語による質問応答システムの実例検討と、利害得失が考察されている。

(B) 第Ⅱ部 最近の言語学の潮流

第Ⅱ部では、現代言語学の潮流を調査した。構成は、つぎのようである。

第1章 生成文法理論の展開 (寺津典子・富山大)

第2章 生成文法理論における意味論の展開 (山梨正明・大阪大)

第3章 ヨーロッパにおける言語学の展開 (横山晶一・電総研)

第4章 Montague 文法 (大谷木重男・電総研)

第1章では、チョムスキー文法が、初期の標準理論から拡大標準理論を経て最近の痕跡理論の導入に伴う修正拡大標準理論へ進展していく過程を追ひ、それがどのように理論的に進化したかを調査した。またこれらの理論的な進化が、どのような理由によりなされたかを歴史に沿って説明した。特に修正拡大標準理論の基本的な考え方を詳しく説明し、痕跡が導入された意義と動機をまとめている。修正拡大標準理論の大略は下図のようである。

(意味解釈規則)

(1) 文文法： $\xrightarrow{B}$ 基底構造 $\xrightarrow{T}$ 表層構造 $\xrightarrow{SI-1}$ 論理形式
意味解釈規則 (SI-2)
他の認知体系 } : 論理形式 $\rightarrow$ '意味'

第2章では、生成文法理論における意味論の展開を調査した。第1章は、

主として統語論が中心であった。生成文法理論は1970年に入って、意味が重視され、それは、解釈意味論、生成意味論、格文法理論等の諸派に分岐してきている。チョムスキー文法の進展も、統語的側面のみならず、意味のとらえ方の進展という視点でみなおすことができる。第1章でも触れた解釈意味論に対置する形で生成意味論が生まれた動機、理由についても調査した。解釈意味論では、意味解釈規則の入力となる深層レベルを次第に表層レベルに近づける。生成意味論では、深層レベルを一層深化させ、意味表示レベルと同一視して行く方向をとる。両者の利害損失は一概にはいえないが、後者は、最近では語用論、発話行為論とも接点をもちつつある。いずれにしても、両者の意味論に、従来軽視されてきた語用論の問題をどう取り込むかが今後の課題であろう。

第3章では、ヨーロッパにおける言語学の展開を調査した。ともすれば、日本では、チョムスキー文法の隆盛の陰に隠れて、ヨーロッパの言語学の現状を知る機会が少ない。しかし、テニエールらの結合価の理論は、動詞を出発点とした文の構造分析として、注目すべきものであろう。また機能文法との関連でプラグ学派の先導的な役割も無視することができないだろう。今回は、その概略を説明する。

第4章では、最近注目をあびているモンタギュー文法について調査した。モンタギュー文法は、統語論と意味論を融合する道を与え、入力文を極めて体系的な方法で内包論理式に翻訳する枠組みを与えた。内包論理は、様相論理の一種であるが、モンタギューは、trivialでない程度の英語文について、これまで言語哲学の分野で問題とされていたFregeの関数原則、すなわち全体の意味は部分の意味から計算可能であるという原則が、整合性をもって成立する方式を提案した。今回の調査では、その概要を報告するにとどめたが、このモンタギュー文法の枠組みは、生成意味論の学者により注目され、それとの等価性が議論されるに及び、非常に活発に研究が進められている。また、計算機科学者の側でも、モンタギュー文法の考え方を取り入れたシステム作りが試みられるなど、今後の研究課題として、実に豊かな土壌を与えてくれたと評価されている。

以上で、自然言語処理ワーキング・グループ報告の大略説明を終えるが、

自然言語処理の研究は長期のものであることを認識するとともに、第5世代の計算機に対しても少なからぬインパクトを与えうる可能性をもっている、ということを指摘したい。研究開発の輪が、我が国でもっと広がることを期待したい。そうした輪の拡がりの中で、実用化を目指す試みも行なわれなければならないと思われる。

第 I 部 自然言語処理システム

I 自然言語処理システム

目 次

1. はじめに	10
1.1 序 論	10
1.2 自然言語処理システムの研究概観	10
1.3 自然言語処理システムと他の学問分野との関連	15
参 考 文 献	20
2. 構文解析	31
2.1 はじめに	31
2.2 自然言語の特徴と構文解析	31
2.3 構文解析のアルゴリズムとシステム	35
2.4 構文解析システム	38
2.5 日本語の構文解析	41
2.6 日本に於ける構文解析システム	44
2.7 おわりに	48
参 考 文 献	49
3. 知識表現 } 次年度調査	
4. 文の生成 }	
5. 問題解決と推論 Default Reasoning	52
5.1 変数割当てにおける default	52
5.2 論理データベースにおける default	53
5.3 階層表現における default	53
5.4 例外表現と default	54
5.5 Frame Problem	55
5.6 Default Reasoning	56
5.7 おわりに	62
参 考 文 献	63

6. 談話理解	6 4
6.1 はじめに	6 4
6.1.1 意味関係による談話内文結合現象の概念	6 4
6.1.2 談話理解における意味関係による 結合現象研究の意義	6 7
6.1.3 指示現象研究の位置付	6 9
6.2 指示現象	7 0
6.2.1 指示現象における照応性	7 0
6.2.2 指示表現のタイプ	7 2
6.2.3 名詞の役をする語群の構造	7 5
6.2.4 人称代名詞による指示	7 6
6.2.5 指示詞による指示	8 0
6.2.6 比較による指示	8 9
6.3 指示現象と代入・省略現象の比較	9 3
6.3.1 指示と代入・省略の対比の概観	9 3
6.3.2 指示と代入との相異	9 4
6.3.3 代入と省略との相異	9 5
参 考 文 献	9 6
7. 音声理解	1 0 1
7.1 はじめに	1 0 1
7.2 音声理解研究	1 0 2
7.2.1 音声理解の問題	1 0 2
7.2.2 初期の音声理解システム	1 0 4
7.3 ARPAの音声理解研究	1 1 8
7.3.1 プロジェクトの成果の概要	1 1 8
7.3.2 System Development Corporation	1 1 9
7.3.3 Bolt, Beranek and Newman Inc	1 2 1
7.3.4 CMU - Hearsay II	1 2 4
7.3.5 CMU - Harpy	1 2 5
7.3.6 その他の機関	1 2 6

7.4	ヨーロッパと日本の音声理解研究	127
7.4.1	ヨーロッパの音声理解研究	127
7.4.2	日本の音声理解研究	129
7.5	音声理解研究の貢献	131
7.5.1	音声認識技術開発の現状	131
7.5.2	音声知覚モデルとの関係	132
7.5.3	音声認識技術の問題点	133
7.6	音声認識の実用化	133
7.7	おわりに	135
	参考文献	136
8.	応用システム	144
8.1	機械翻訳	144
8.1.1	歴史的概観	144
8.1.2	現状と課題	164
8.1.3	機械翻訳の今後	180
	付 録	181
	参考文献	189
8.2	質問応答システム	197
8.2.1	はじめに	197
8.2.2	質問応答システムのモデル	198
8.2.3	質問応答と自然言語, QBE, DEDUCE	200
8.2.4	自然言語による質問応答の実例検討	204
8.2.5	意味文法	215
8.2.6	今後の課題	216
8.2.7	おわりに	220

1. はじめに

1.1 序 論

計算機の高度化には、i) 計算機速度の向上、ii) 知的機能の向上が考えられるだろう。ii) は、究極的には計算機のもつ知的機能を人間並みにしたいということである。自然言語を計算機で理解するという試みには、特にii) が関係している。

現在の計算機の理解できる言語は、マイクロ命令にはじまり、機械語、アセンブリ言語を経て高級プログラミング言語に至るまで、さまざまなレベルがある。しかし、これらは全て人工言語とよばれる範疇の言語である。これに対し我々人間は、日常生活の場において、自然言語、音声、イメージ等を媒介にして相互の意志疎通を行なっている。ところがこの、自然言語や音声、イメージ等の理解は、現在の計算機が最も不得意とするものの一つである。

しかし、計算機にこれらを理解させようとする(ある意味で野心的ともいえる)試みは、1970年代に入って、解決に向けて一定の前進がみられた、というのも事実である。このような自然言語と計算機をめぐる研究開発の輪は、計算機科学者のみならず、言語学者、心理学者、哲学者、論理学者をも巻き込み、拡大の一途をたどり、学際的な研究交流の場さえ生れようとしている。これは、そもそも自然言語そのものに内在する特質から生れた当然の帰結であるといみなされようが、極めて健全な動きであるといえる。これらは、1.2と1.3で説明する。

1.2 自然言語処理システムの研究概観

世界最初の電子計算機が誕生したのは、1945年のことであった。(電子)計算機が、その後の社会に及ぼした影響の大きさを考える時、計算機の出現は、第二次世界大戦後最大の技術革新の一つであったといえる。しかもその機械は、人間の肉体労働の一部を代行する機械でなく、人間の知的活動の一部を代行するものであった。これは従来の機械のイメージを書き換えるものであった。

計算機の当初の目的は、数値計算を高速に行なうことにあった。しかし、計算機の本質が、記号を処理する能力にあることは、Post, Turing等の数学者

により、計算機が出現する以前から、すでに明らかにされていた。計算機が本来備えていたこのような記号処理能力と、人間の知的な活動の多くが記号処理能力に依存している事実とを重ね合わせてみると、1950年代の後半になり、定理を証明したり、ゲームやパズルをするといった、より知的な人間の行為を計算機に代行させようとする計画、すなわち人工知能研究プロジェクトが発足するのは、当然の成り行きであったといえる。Minsky はつぎのように述べている〔Minsky, 1976〕。^{*)} *"I consider 'thinking' to be within the scope of effective computation"*。心理学の分野から人工知能の研究を推進しようとする試みもあった。Newell, Simon等がその代表であった。彼等は、人間の知的活動を、情報処理装置 (information processor)に見立てて解明しようとした〔Newell, 1973〕。そこに人間機械論的な「ものの見方」を認めることができる。これは、古典的な人間機械論ではなく、計算機という新しい機械を媒介にした新しい意味での人間機械論であることに注目したい。

計算機が記号を操作できることから、1960年に入って、計算機に言語を理解させようとする試みが活発化してくる。特に、当時の言語理論に新しい大きな流れを創り出したChomskyの生成変形文法は、機械的な記号操作をベースにしており、言語の機械処理の立場からも注目すべき理論であった。多くの人が、この理論をもとにした機械翻訳システムの実現を夢想し、計画を具体化しようとした。人間の操る言語は、高度に知的な人間活動の産物である。言語の使用と理解は、人間の豊富な知識と経験とをベースにしている。したがって、自然言語処理システムの実現は、高度に知的な機械の実現と密接に関連しており、それは本来的に人工知能の大きな研究テーマの一つであったといえる。しかし、1960年代の人工知能の研究は、ゲーム、パズル、ロボットの研究を中心に展開されていた。

1960年代の自然言語処理システムの研究に関していえば、機械翻訳システムの研究が中心であった。それに対して、人工知能の研究者は、機械翻訳システムの実現には、言語の意味 (semantics)の解明が必要であると主張していた。このような主張は、当時MITのMinsky〔Minsky, 1968〕の提唱した^{*)} *"Semantic Information Processing"*という考え方に集約されているといえる。当時の人工知能の研究者の多くは、ゲーム、パズル、定理証明、もしくは

知能ロボットの研究に対して楽観的であったが、機械翻訳システムの研究に対しては慎重であった。

それとは全く対照的に、当時の翻訳の研究者の多くは、機械翻訳システムの実現を楽観視していた。そして彼等の研究は、人工知能の研究と独立に進められていた。当時の言語理論は統語論(syntax)中心であった。彼等の多くは、統語論によって、かなりの程度の機械翻訳システムが実現可能であると考えていた。彼等の楽観論の背景には、計算機技術の進歩、それに Chomsky 等による新しい言語理論の出現があったと考えられる。しかし、統語論をベースにした機械翻訳システムの実現が困難なものであることが次第に明らかになってくる。そして米国で ALPAC (Automatic Language Processing Advisory Committee) の報告書がまとめられるに及び、最も活発な研究が行なわれていた米国での機械翻訳の研究は、急速に収束していった [Hutchins, 1978] [Petrick, 1976]。

このように、当時の機械翻訳の研究は、言語(ことば)における意味の重要性の再認という、今日からみると当然ともいえる結末を迎えた。しかも、そこから一定の成果として後世に受け継がれたものは、意外にも少なかったように思われる。ただ、Harvard 大学の久野 等が作成した Predictive Analyser [Kuno, 1965] は、その後の統語処理の研究に、大きな影響を与えている。Chomsky の生成変形文法に関していえば、その形式的な側面は、自然言語ではなくむしろ人工言語(プログラミング言語)の分野で有用な理論であることが実証された。

このような、機械翻訳システムの研究の挫折は、おそらく、言語理論の構築を目ざす言語学者にも、ある種の影響を与えたものと推察される。その頃から、生成変形文法の枠内で、Chomsky の標準理論に立脚した解釈意味論と、Lakoff, McCawley 等の主張する生成意味論との間に論争がはじまった。それは現在でも続いている [月刊言語, 1977] [Winograd, 1971]。

1960年代後半になると、自然言語処理システムの研究は、明らかに大きな壁で前途を立ち塞がれていた。この手詰り状態を開く成果は、1960年代後半から1970年にかけての人工知能の研究から生れた。1970年代の自然言語処理システムの研究状況を図 1.1 に示す。

とりわけ M I T の Winograd [Winograd, 1971] の作成した SHRDLU システムは、1970年代の自然言語処理システムの研究に大きな影響を与えた。これは積木の世界を操作するロボットに対して、英語の指令を出して、積木を積み上げたり移動する仕事を行なわせるものであった。SHRDLUは、人間との会話の場面や、文脈を用いた文の意味解釈を行なうことができる。SHRDLUの成功は、以前とは比較にならない程の計算機パワーが利用できるようになってきたこと、対象世界を極めて狭く限定したことなどが間接的な要因であるにしても、Winograd が、意味、文脈、場面、常識など、それまで困難で手がつけられないと見なされていた諸問題に、まじめに取り組んだからだと考えられる。部分的であるにせよ、それらの問題を解決し、動作しうる一つのシステムとして統合化した SHRDLU は、確実に自然言語処理システムの研究を前進させ、それに対する手がかりを与えた。SHRDLU 以後の研究は、SHRDLU の限界に挑戦することから開始されたとしても過言ではなからう [Wilks, 1974]。

1970年に入り、自然言語処理システムだとか、音声理解システムだとか、ことさら理解という名称をもつシステムが目につくようになった。これは、意味の問題を回避しようとしていた1960年代のシステムとの差を、言外に強調しようとしているからである。

1970年半ばに入ると、自然言語処理システムの研究は、人工知能の中核的な研究分野に成長する。ちなみに、1977年秋に米国のボストンで開催された第5回人工知能国際会議では、自然言語処理システムに関連する研究発表がおよそ1/3を占めている。電総研の淵一博は、その一年以上も前にこうした傾向を予測して、“ゲームとパズルの時代から、知能ロボットの時代を経て、言語と知識の時代へ”と述べている [淵, 1976]。

最近の自然言語処理システムの研究は、i) 対象世界の拡大、ii) 知識表現形式の検討、iii) 実用化可能性の検討、の方向で研究が進められている。

i), ii) は、自然言語処理システムの実現を長期的な視野におき、蓄積性のある研究を進める立場である。また、最近目立つことは、i), ii) に関して、他学問分野との接触、協同作業がみられることである。それについては次節で簡単に説明する。iii) の実用化可能性を検討する方向は、現在の計算機技術の進歩と無縁ではない。特に、データ・ベースに対する問い合わせ用言語と

して自然言語を使用する試みが、米国のIBM, SRI, BBN, イリノイ大などでプロジェクト化され、実現にむけて活発に研究されている。

この節を終えるにあたり、我が国の様子について簡単にふれておく。我が国でも、京都大学の長尾、辻井等〔長尾, 1976〕により、文脈を考慮した日本語文（理科の教科書文）の処理が試みられた。格文法の考え方と、WoodsのA T N〔Woods, 1971〕とをベースにしたパーザ〔長尾, 1974〕を用いており、我が国の自然言語処理システムの先駆けをなすものであった。最近では、機械翻訳システムの研究に主眼が注がれてきている。

電々公社武蔵野通研の雨宮、島津〔雨宮, 1977〕は、積み木の世界を対象にした日本語質問応答システムを作成しているが、SHRDLUほど高度なものではない。

九州大学を中心に吉田〔吉田, 1978〕等は、真の意味の自在な自然言語の処理は現時点では不可能であるとして、構文的な制限を課し、しかも自然で標準的な文は何かという問題を追求している。これとは別に、長期的な観点からすれば、栗原、吉田〔栗原, 1973〕等の膨大な言語データの意味分析の研究が注目される。

電子技術総合研究所では、小規模な日本語質問応答システムを作成し、理解した結果を図示するシステムを作成した〔佐藤, 1976〕。その後、パーザの開発〔田中, 1977〕を経て、意味抽出の問題に取り組んでいる〔田中, 1978〕。さらに、抽出した意味構造から文を再構成する研究などが行なわれている〔佐藤, 1979〕, 〔横山, 1979〕。

IBMの東京サイエンティフィック・センターでは、実用化をめざした（日本語による）データ・ベースに対する質問応答システムが開発された。質問文型に強い制限があるが、我が国唯一の本格的な質問応答システム作成の試みであるといえる。詳細は、本文8章8.2節で説明されるだろう。〔鷹尾, 1976〕。

自然言語の処理を進めるためには、言語データの計算機入力（たとえば、電総研の渕による新明解国語辞典、京都大学の長尾によるコンサイス英和辞典の磁気テープ化など）による生データの整備と共に、計算機を用いた大量データ処理の研究を進め、手作業の一部を自動化することも必要になろう。

1.3 自然言語処理システムと他の学問分野との関連

1970年に入り、自然言語処理システムの研究者と、言語学者、心理学者との接触は、ますます活発化したことを1.1で述べた。

言語学に関していえば、最近の自然言語処理システムの研究に対して大きな影響を与えた文法は、Fillmoreの格文法であろう〔Fillmore, 1968〕。特に、格フレームの考え方が大きな影響を与えている。格文法では、文の意味は、動詞を中心に幾つかの深層格から構成された格フレームを用いて決定される。たとえば「切る」という動詞は、動作主格、道具格、対象格などの深層格から構成される格フレームを持っていると考えるのである。^{*}「花子ハ ハサミデ 紙ヲ 切る」という文の場合には、動作主格＝花子、道具格＝ハサミ、対象格＝紙のような深層格によって、文の意味が決定される。格フレームの考え方は、動詞を中心にした意味構造抽出法に対する手がかりを与え、多くの自然言語処理システムで採用された。

このことは、1960年代の機械翻訳の研究者が、Chomskyの言語理論（標準理論ともよばれる）〔Chomsky, 1965〕に強く影響されたのと、著しい対照をなしている。Fillmoreの格文法が生成意味論の立場をとっていることを考えると自然言語処理システムの研究は、標準理論に立脚した解釈意味論より、むしろそれと対立する生成意味論の側から大きな影響を受けていることは興味深い。しかしこのことを、解釈意味論が否定されたなどと短絡化してとらえるべきではないだろう。両者の長所短所を慎重に見極めなければならない。解釈意味論には解釈意味論の良さがある。

Chomsky自身、自己の理論が進展し、最近では、言語が文字化（もしくは音声化）されたレベルに近い表層構造上で、さまざまな言語事象を説明しようとする痕跡理論（trace theory）を提案している。このように、最近の言語学の分野では、Chomsky言語学派の内部対立（生成意味論と解釈意味論の対立）や、関係文法などの新しい言語理論建設に向けての息吹きが観察される〔青木, 1976〕。

^{*} 佐久間〔佐久間, 1940〕は、格文法が提案されるかなり以前に、既に同様な考え方を提案している。

最近、Chomsky学派から、自然言語処理システムの研究を進めている人工知能の研究者に向けての厳しい批判と、それに対する反論が話題になった〔Dresher, 1976〕。これは、彼等にとって、人工知能の立場からの言語の研究が、無視できない存在になってきていることを示すものとして興味深い。一方それとやや異なる立場から、(過去における)人工知能の研究者が、人工知能の研究そのものに対して鋭い内部批判を行なっている。そこでは、Chomskyの言語理論に対する擁護がなされている。Chomskyの最新の言語論を計算機プログラム化しようとする試みもある〔Marcus, 1978〕。

Chomsky学派からの批判に関していえば、彼等は、つぎの問題を見逃しているように思われる。それは、たとえば言語学者が、さまざまな言語事象を説明しようと称する理論を建設したとしても、それが誤りを含まずに動作するシステムに直結するという保証は無い、ということである。さらに、たとえその理論に誤りがないとしても、それをどのようにして動作しうるシステムとして実現したら良いか、具体的な手法が問題になる。

言語の研究をめぐる最近の世界情勢は流動的で新理論誕生の胎動が感じられ、非常に興味ある時期にさしかかってきている。一応の結着がつくのは、少なくとも数年先のことになるだろう。

ここで、意味の問題について考察しておく。意味の問題は、古来さまざまな議論がなされてきた。にもかかわらず、依然としてその本質が科学的に解明されたとはいえない。最近では、言語学、心理学、哲学、論理学の各分野で、意味の本質を解明する努力が続けられている。特に、言語学者による深層構造と表層構造の設定、心理学者による意味ネットワークの考え方、哲学者や論理学者の使用する内包と外延の区別、可能的世界の導入などが注目される。

言語学者による深層構造と表層構造の設定は、言語の文字化(音声化)されたレベルに近い構造(表層構造)と、意味のレベルに近い構造(深層構造)との間に、構造的な差があるとする考え方である。先の痕跡理論では、その差が以前程際立たなくなっているが、それでもそれを意識することは、自然言語処理システムの作成者にとって必要であるといえる。意味に近いレベルは言語の種類によらない、すなわち普遍性をもっていると考えられている。この事実から、最近では機械翻訳システムの設計原理として、機械翻訳の対象となる二つ

の言語の中間に、そのような言語によらない普遍的なレベルを設定すべきであると考えられるようになってきた。

意味ネットワークの考え方は、心理学者の Quillian [Quillian, 1968] によって提案された。辞書項目の語の場合を考えてみる。辞書項目の各語の意味は孤立して存在しているのではない。語の意味は、周囲にある他の語と意味的な関連と緊張を保ちつつ存在している。これはちょうど網目（ネットワーク）のようなものを想像させる。網の結び目が個々の語に相当し、そこに語の中心的な意味が存在していると考えるのである。このような網目構造のことを意味ネットワークとよぶ。

辞書のような静的な意味ネットワークに対して、動的な意味記述として意味ネットワークを用いることがある。たとえば、文中の各語は、文中の他の語からの意味的な影響を受けながら、その語の（文中における）意味が決まると考えられる。先に述べた文の格構造は、与えられた文に依存して決まる動的な意味ネットワークの典型である。

意味ネットワークは、一般に、意味や知識をどのように表現すべきかという問題に直観的な枠組を与え、それを用いて様々な推論や連想を行なうための人間の記憶モデルとして提案された。モデルの素直さのために、意味ネットワークの考え方は自然言語処理システムの研究に大きな影響を与え、現在では、さまざまな角度から研究が進められている。

哲学者や論理学者の用いる内包と外延の区別は、意味の問題にとって重要である。内包と外延は、Frege の *sin* と *signifie* の区別とはほぼ一致している。「宵の明星」と「明けの明星」とは、外延的には同一物体（金星）を指すが、その内包としての意味が異なる。このような区別は意味を正確にとらえようとする時に問題になる。意味ネットワーク上に、内包と外延の区別をどのように表現すべきかを論じた Woods [Woods, 1975a] の考察があるが、完全な解決はまだ見い出されてはいない。

様相論理における可能的世界も、文脈や常識を処理する場合に有用な概念である。与えられた文の真偽は、可能的世界に依存して決まる。たとえば「日本ノ首相ハ大平デス」という文の真偽は、それが10年前のことなら正しくないだろう。10年前の世界は、明らかに一つの可能的世界である。時間は可能的

世界を決めるインデックスの一つである。談話の理解には、このような可能的世界の導入が必要になる。それに関連して、一種のラムダ計算である内包論理と様相論理とを武器に、統語論と意味論を統合的に扱おうとする Montague 文法の考え方が注目される [Allwood, 1977]。淵一博 [淵, 1977] が指摘しているように、本研究で用いるパーザの意味構造抽出法は、Montague 文法のそれと酷似している。

自然言語処理システムの研究から生れた成果が、最近言語学者や心理学者にも影響を与えている。意味の手続き的記述 [Winograd, 1971] や談話理解のためのフレーム応用 [Fillmore, 1977], [Minsky, 1975], [Schank, 1977], Schank の基本 ACT の実在性の検討 [Schank, 1973], 文パーズのための心理モデルとしての ATNG [Norman, 1975], [Woods, 1971] の応用等である。このように、自然言語処理システムそれ自体が、意味の理論を検証するための一つの実証可能なモデルになることも忘れてはならない。

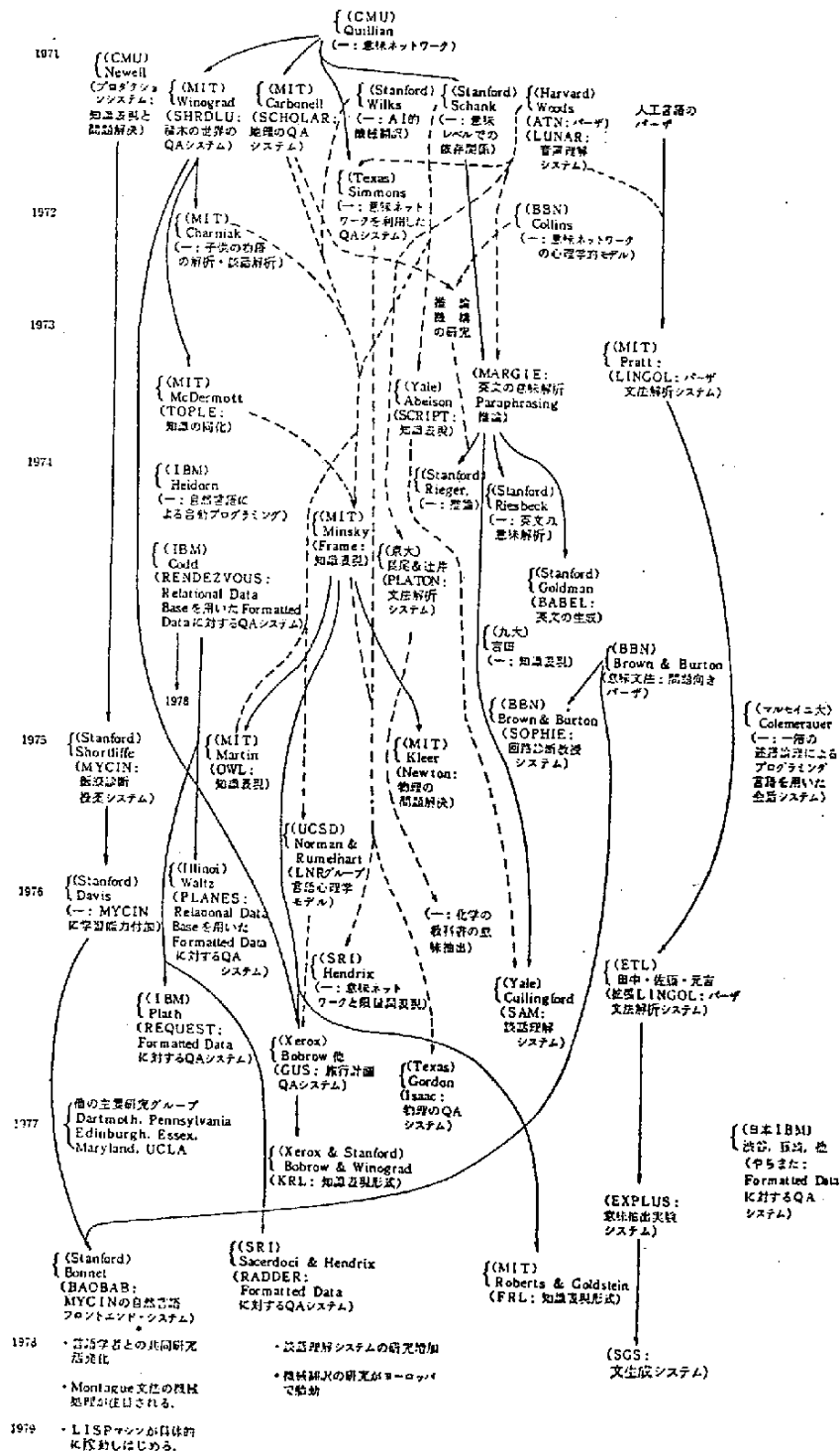


図1-1 1970年代の言語理解システムの研究状況

参 考 文 献*

- (略号 AJCL≡American Journal of Computational Linguistics
CACM≡Comm. of American Computing Machinery
IJCAI≡International Joint Conference of Artificial Intelligence)
- [Allwood, 1977] Allwood, J., Andersson, L. and Dahl, O.: "Logic in Linguistics", Cambridge Univ. Press, 1977.
- [Austine, 1962] Austine, J.L.: "How to Do Things with Words", Oxford Univ. Press, 1962.
- [Bach, 1968] Bach, E. and Harms R.T. (Eds.): "Universals in Linguistic Theory", Holt, Rinehart and Winston, 1968.
- [BBN, 1978] BBN Report 3878, "Research in Natural Language Understanding", Quaterly Progress Report No. 3, 1978.
- [Bever, 1971] Bever, T.G. and Rosenbaum P.S.: "Some Lexical Structures and Their Empirical Validity", in Steinberg and Jakobovits (Eds.): "Semantics", Cambridge Univ. Press, 1971.
- [Bobrow, 1975] Bobrow, D.G. and Collins, A.M. (Eds.): "Representation and Understanding", Academic Press, 1975.
- [Bobrow, 1977a] Bobrow, D.G. and Winograd, T.: "An Overview of KRL, a Knowledge Representation Language", Studies in Cognitive Science, Vol. 1, No. 1, 1977.
- [Bobrow, 1977b] Bobrow, D.G., Winograd, T. and the KRL Research Group: "Experience with KRL -0, One Cycle of a Knowledge Representation Language", IJCAI5, 1977, 213-222.
- [Bobrow, 1977c] Bobrow, D. G., Kaplan, R. M., Kay, M., Norman, D. A. and Winograd, T. "GUS, A Fame-Driven Dialog System", Artificial Intelligence, Spring, 1977, 155-173.
- [Bolc, 1978] Bolc L. (Ed.): "Natural Language Communication with Computers", Springer-Verlag, 1978.
- [Bonnet, 1978] Bonnet, A.: "BAOBAB, A Parser for a Rule-Based System Using a Semantic Grammar", STAN-CS-78-668, Stanford Univ., 1978.
- [Brown, 1975] Brown, J.S. and Burton, R.R.: "Multiple Representations of Knowledge", in Bobrow D.G. and Collins A. (Eds.): "Representation and Understanding", Academic Press, 1975, 331-349.

* 田中穂積: "計算機による自然言語の意味処理に関する研究", 電子技術総合研究所研究報告 797, 1979, の文献リストを補強したものである。各章末の文献リスト中には、この文献リストにのっていないものも多数収録されている。

- [Bruce, 1975] Bruce, B.: "Case Systems for Natural Language", *Artificial Intelligence*, 6, 1975, 327-360.
- [Carbonell, 1970] Carbonell, J.R.: "AI in CAI: An Artificial Intelligence Approach to Computer-Assisted Instruction", *IEEE, MMS-11*, 4, Dec. 1970, 190-202.
- [Carbonell, 1973] Carbonell, J.R. and Collins, A.M.: "Natural Semantics in Artificial Intelligence", *IJCAI3*, 1973.
- [Celce-Murcia, 1976] Celce-Murcia, M.: "Verb Paradigms for Sentence Recognition", *AJCL, Microfiche* 38, 1976.
- [Chafe, 1970] Chafe, W.L.: "Meaning and Structure of Language", The Univ. of Chicago Press, 1970.
- [Chafe, 1972] Chafe, W.L.: "Discourse Structure and Human Knowledge", in Carrol and Freedle (Eds.): "Language Comprehension and the Aquisition of Knowledge", John Wiley and Sons, 1972.
- [Chafe, 1973] Chafe, W.L.: "Language and Memory", *Language*, 49, 2, 1973, 261-281.
- [Chafe, 1974] Chafe, W.L.: "Language and Consciousness", *Language*, 50, 1, 1974, 111-133.
- [Chang, 1978] Chang, C. L.: "DEDUCE 2: Further Investigations of Deduction in Relational Data Bases", in Gallaire, H. and Minker, J. (Eds.): "Logic and Data Bases", Plenum, 1978.
- [Codd, 1978] Codd, E. F.: "Rendezvous Version 1: An Experimental English-Language Query Formulations System for Casual Users of Relational Data Bases", RJ2144(29407) Computer Science Research Report, IBJ Research Lab. San Jose, Calif., 1978.
- [Charniak, 1972] Charniak, E.: "Toward a Model of Children's Story Comprehension", AI TR-226, MIT, December 1972.
- [Charniak, 1976] Charniak, E. and Wilks, Y.: "Computational Semantics", North-Holland, 1976.
- [Chomsky, 1965] Chomsky, N.: "Aspects of the Theory of Syntax", The MIT Press, 1965.
- [Chomsky, 1971] Chomsky, N.: "Deep Structure, Surface Structure, and Semantic Interpretation", in Steinberg and Jakobovits (Eds.): "Semantics", Cambridge University Press, 1971.
- [Chomsky, 1975] Chomsky, N.: "Reflections on Language", Pantheon, 1975.
- [Collins, 1975] Collins, A., Warnock, E.H., Niello, N. and Miller, M.L.: "Reasoning From Incomplete Knowledge", in Bobrow, D. and Collins, A. (Eds.): "Representation and Understanding", Academic Press, 1975, 383-415.
- [Collins, 1975] Collins, A.M. and Loftud, E.F.: "A Spreading-Activation Theory of Semantic Processing", *Psychological Review*, Vol. 82, No. 6, 1975, 407-428.
- [Davis, 1975] Davis, R.R. and King, J.: "An Overview of Production Systems", Stanford AIM-271, Oct. 1975.
- [Dillon, 1977] Dillon, G.L.: "Introduction to Contemporary Linguistic Semantics", Prentice-Hall, 1977.

- [Dowty, 1979] Dowty, D.R.: "A Guide to Montague's PTQ", D. Reidel Publishing Co., 1979.
- [Dresher, 1976] Dresher, B.E. and Hornstein, N.: "On Some Supposed Contributions of Artificial Intelligence to the Scientific Study of Language", *Cognition*, 4, 1976, 321-398.
- [Finin, 1979] Finin, T., Goodman, B. and Tennant, H.: 'JETS: Achieving Completeness through Coverage and Closure', *IJCAIS*, 1979, 275-281.
- [Fillmore, 1968] Fillmore, C.J.: "The Case for Case", in Bach and Harms (Eds.): "Universals in Linguistic Theory", Holt, Rinehart, and Winston, 1968.
- [Fillmore, 1969] Fillmore, C.J.: "Types of Lexical Information", in Kieffer (Ed.): "Studies in Syntax and Semantics", D. Reidel Publishing Co., 1969.
- [Fillmore, 1977] Fillmore, C. J.: "Scenes-and-Frames Semantics", in Zampolli (Ed.): "Linguistic Structures Processing", North-Holland, 1977.
- [Fordor, 1977] Fordor, J.D.: "Semantics: Theories of Meaning in Generative Grammar", Thomas Y. Crowell Co., 1977.
- [Frederiksen, 1975] Frederiksen, C.H.: "Representing Logical and Semantic Structure of Knowledge Acquired from Discourse", *Cognitive Psychology*, 7, 1975, 371-458.
- [Goldman, 1975] Goldman, N.M.: "Sentence Paraphrasing from a Conceptual Base", *CACM*, 2, 18, 1975, 96-106.
- [Gregg, 1974] Gregg, L.W. (Ed.): "Knowledge and Cognition", Laurence Erlbaum Associates Publishers, Potomac, Maryland, 1974.
- [Grice, 1975] Grice, H.P.: "Logic and Conversation", in Cole, P. and Morgan, J.L. (Eds.): "Syntax and Semantics", Vol. 3, Academic Press, 1975, 41-58.
- [Grishman, 1976] Grishman, R.: "A Survey of Syntactic Analysis Procedures for Natural Language", *AJCL*, Microfiche 47, 1976.
- [Grosz, 1977] Grosz, B.J.: "The Representation and Use of Focus in a System for Understanding Dialogs", *IGCAI5*, 1977, 67-76.
- [Halle, 1978] Halle, M., Bresnan, J. and Miller, J.A.: "Linguistic Theory and Psychological Reality", The MIT Press, 1978.
- [Halliday, 1961] Halliday, M.A.K.: "Categories of the Theory of Grammar", *Word*, 17, 1961, 241-292.
- [Harada, 1977] Harada, S.I.: "Topics in Relational Grammar", *Ann. Bull. of RILP*, No. 11, University of Tokyo, 1977, 131-153.
- [Harris, 1979] Harris, L. R.: 'Experience with ROBOT in 12 Commercial Natural Language Data Base Query Application', *IJCAI6*, 1979, 365-368.
- [Harris, 1977] Harris, L. R.: 'User-Oriented Data Base Query with ROBOT Natural Language Query System', *Third International Conf. on Very Large Data Bases*, 1977, 303-311.
- [Hawkinson, 1975] Hawkinson, L.: "The Representation of Concepts in OWL", *IJCAI4*, 1975, 107-114.

- [Hays, 1976] Hays, D.G. and Mathias, J. (Eds.): "FBIS Seminar on Machine Translation", AJCL, Microfiche 46, 1976.
- [Hayes, 1979] Hayes, P. and Reddy, R.: "Graceful Interaction in Man-Machine Communication", IJCAI6, 1979, 372-374.
- [Heidorn, 1975] Heidorn, G.E.: "Augmented Phrase Structure Grammar", in Schank and Nash-Webber (Eds.): "Theoretical Issues in Natural Language Processing", Cambridge, Mass. 1975, 1-5.
- [Hendrix, 1977] Hendrix, G.G.: "Human Engineering for Applied Natural Language Processing", Stanford Research Institute AI Center, TR 139, Feb. 1977.
- [Hobbs, 1978] Hobbs, J.R. and Rosenschein, S.J.: "Making Computational Sense of Montague's Intensional Logic", Artificial Intelligence, 9, 1978, 287-306.
- [Hollan, 1975] Hollan, J.D.: "Features and Semantic Memory: Set-Theoretic or Network Model?", Psychological Review, Vol. 82, No. 2, 1975, 154-155.
- [Hutchins, 1978] Hutchins, W.J.: "Progress in Documentation-Machine Translation and Machine-Aided Translation", Journal of Documentation, 34, 2, June, 1978.
- [Kaplan, 1972] Kaplan, R.M.: "Augmented Transition Networks as Psychological Models of Sentence Comprehension", Artificial Intelligence, 3, 1972, 77-100.
- [Kintsch, 1972] Kintsch, W.: "Notes on the Structure of Semantic Memory", in Tulving and Donaldson (Eds.): "Organization of Memory", Academic Press, 1972.
- [de Kleer, 1975] de Kleer, J.: "Qualitative and Quantitative Knowledge", AI-TR-352, MIT, 1975.
- [Kuno, 1965] Kuno, S.: "The Predictive Analyzer and a Path Elimination Technique", CACM, 8, 1965, 687-698.
- [Kuno, 1977] Kuno, S. and Kaburaki, E.: "Empathy and Syntax", Linguistic Inquiry, Vol. 8, No. 4, 1977.
- [Lakoff, 1976] Lakoff, G.: "Pronouns and Reference", in McCawley, J.D. (Ed.): "Syntax and Semantics", Vol. 7, Academic Press, 1976, 275-335.
- [Lehnert, 1977] Lehnert, W.: "Human and Computational Question Answering", Cognitive Science, Vol. 1, 1977, 47-73.
- [Lehnert, 1978] Lehnert, W. G.: "The Process of Question Answering", Lawrence Erlbaum Associates, Publishers, 1978.
- [Lyons, 1968] Lyons, J.: "Introduction to Theoretical Linguistics", Cambridge Univ. Press, 1968.

- [Marcus, 1978] Marcus, M.: "A Computational Account of Some Constraints on Language", in Waltz (Ed.): "Theoretical Issues in Natural Language Processing-2", ACM, 1978, 236-246.
- [Martin, 1974] Martin, W.A.: "A System for Building Expert Problem Solving Verbal Reasoning", OWL-Notes MIT MAC, 1974.
- [Martin, 1978] Martin, W.A.: "Descriptions and Specialization of Concepts", LCS-TM-101, MIT, 1978.
- [McDermott, 1974] McDermott, D.V.: "Assimilation of New Information by a Natural Language Understanding System", AI TR-291, MIT, Feb. 1974.
- [Michie, 1977] Michie, D. (Ed.): "Machine Intelligence", 8, Edinburgh Univ. Press, 1977.
- [Miller, 1976] Miller, G.A. and Johnson-Laird, P.N.: "Language and Perception", Harvard Univ. Press, 1976.
- [Minsky, 1976] Minsky, M.: "Computation: Finite and Infinite Machines", Prentice-Hall, Inc., 1976.
- [Minsky, 1968] Minsky, M. (Ed.): "Semantic Information Processing", The MIT Press, 1968.
- [Minsky, 1975] Minsky, M.: "Framework for Representing Knowledge", in Winston (Ed.): "The Psychology of Computer Vision", McGraw-Hill, 1975.
- [Nagao, 1976] Nagao, M. and Tsujii, J.: "Analysis of Japanese Sentences Using Semantic and Syntactic Information", AJCL, Microfiche 41, 1976.
- [Nash-Webber, 1977] Nash-Webber, B. and Reiter, R.: "Anaphora and Logical Form: On Formal Meaning Representations for Natural Language", IJCAI5, 1977, 121-131.
- [Newell, 1973] Newell, A.: "Artificial Intelligence and the Concept of Mind", in Schank and Colby (Eds.): "Computer Models of Thought and Language", Freeman and Company, 1973.
- [Newell, 1973] Newell, A.: "Productions Systems: Models of Control Structures", in Chase (Ed.): "Visual Information Processing", Academic Press, 1973.
- [Nida, 1975] Nida, E.A.: "Componential Analysis of Meaning", Mouton & Co. N.V., 1975.
- [Nilson, 1974] Nilson, N. J.: "Artificial Intelligence", IFIP74, North-Holland, 1974, 773-801.
- [Norman, 1975] Norman, D.A. and Rumelhart, D.E.: "Explorations in Cognition", Freeman and Company, 1975.
- [Novak, 1976] Novak, G.S.: "Computer Understanding of Physics Problems Stated in Natural Language", AJCL, Microfiche 53, 1976.
- [Ogden, 1923] Ogden C.K. and Richards, I.A.: "The Meaning of Meaning", London: Kegan Paul, 1923.
- [Palmer, 1976] Palmer, F.R.: "Semantics", Cambridge Univ. Press, 1976.
- [Petrick, 1976] Petrick, S.R.: "On Natural Language Based Computer Systems", IBM J. Res. Develop., July, 1976, 315-325.
- [Petrick, 1976] Petrick, S.R.: "Summary of FBIS Seminar on Machine-Aided Translation", AJCL, Microfiche 46, 1976.

- [Plath, 1976] Plath, W.J. : "REQUEST : A Natural Language Question-Answering System", IBM J. Res. Develop., July, 1976, 326-335.
- [Pratt, 1973] Pratt, V.R. : "A Linguistic Oriented Programming Language", IJCAI3, 1973, 272-381.
- [Pratt, 1975] Pratt, V.R. : "LINGOL-A Progress Report", IJCAI4, 1975, 422-428.
- [Quillian, 1968] Quillian, M.R. : "Semantic Memory", In Minsky (Ed.) : "Semantic Information Processing", The MIT Press., 1968.
- [Raphael, 1975] Raphael, B. : "SIR : A Computer Program for Semantic Information Retrieval", in Minsky (Ed.) : "Semantic Information Processing", The MIT Press, 1968, 33-134.
- [Reddy, D.R. (Ed.) : "Speech Recognition", Academic Press, 1975.
- [Reiter, 1978] Reiter, R. : "On Reasoning by Default", in Waltz (Ed.) : "Theoretical Issues in Natural Language Processing-2", ACM, 1978, 210-218.
- [Rieger, 1974] Rieger, C. : "Conceptual Memory", Stanford AIM-247, Jan. 1974.
- [Rieger, 1979] Rieger, C. and Small, S. : "Word Expert Parsing", IJCAI6, 1979, 723-726.
- [Riesbeck, 1974] Riesbeck, C.K. : "Computational Understanding : Analysis of Sentences and Context", Stanford AIM-238, May 1974.
- [Roberts, 1977] Roberts, R.B. and Goldstein, I.P. : "The FRL Manual", MIT AI Memo 409, Sept. 1977.
- [Rumelhart, 1972] Rumelhart, D.E., Lindsay, P.H. and Norman, D.A. : "A Process Model for Long-Term Memory", in Tulving and Donaldson (Eds.) : "Organization of Memory", Academic Press, 1972.
- [Rumelhart, 1973] Rumelhart, D.E. and Norman, D.A. : "Active Semantic Networks as a Model of Human Memory", IJCAI3, 1973, 450-457.
- [Rustin, 1971] Rustin, R. (Ed.) : "Natural Language Processing", Algorithmic Press, 1971.
- [Schank, 1973a] Schank, R.C., and Colby K.M. (Eds.) : "Computer Models of Thought and Language", Freeman and Company, 1973.
- [Schank, 1973b] Schank, R.C. : "Identification of Conceptualizations underlying Natural Language", in Schank and Colby (Eds.) : "Computer Models of Thought and Language", Freeman and Company, 1973.
- [Schank, 1973c] Schank, R.C. et al : "MARGIE : Memory, Analysis, Response Generation, and Inference on English", IJCAI3, 1973, 255-261.
- [Schank, 1974] Schank, R.C. and Rieger, C. : "Inference and Computer Understanding of Natural Language", Artificial Intelligence, 5, 1974, 373-412.
- [Schank, 1975] Schank, R.C. : "Conceptual Information Processing", North-Holland/American Elsevier, 1975.
- [Schank, 1977] Schank, R.C. and Abelson, R. : "Scripts Plans Goals and Understanding", Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1977.

- [Shapiro, 1978] Shapiro, S.C. : "Path-Based and Node-Based Inference in Semantic Networks", in Waltz (Ed.) : "Theoretical Issues in Natural Language Processing-2", ACM, 1978, 219-225.
- [Shortliffe, 1976] Shortliffe, E.H. : "Computer-Based Medical Consultations : MYCIN", American Elsevier, 1976.
- [Simmons, 1973] Simmons, R.F. : "Semantic Networks : Their Computation and Use for Understanding English Sentences", in Schank and Colby (Eds.) : "Computer Models of Thought and Language", Freeman and Company, 1973.
- [Steedman, 1977] Steedman, M.J. : "Time and Modality", Cognitive Science, Vol. 1, No. 2, 1977, 216-234.
- [Steinberg, 1971] Steinberg, D.D. and Jakobovits, L.A. (Eds.) : "Semantics", Cambridge University Press, 1971.
- [Stenning, 1978] Stenning, K. : "Anaphora as an Approach to Pragmatics", in Halle, Bresnan and Miller (Eds.) : "Linguistic Theory and Psychological Reality", The MIT Press, 1978, 162-200.
- [Takahasi, 1972] Takahasi, H., Fujimura, O. and Kameda, H. : "A Behavioral Characterization of Topicalization : What is Topicalization? From a Computer System Point of View", Annual Bull. of RILP, No. 6, University of Tokyo, 1972, 93-110.
- [Tanaka, 1978] Tanaka, H., Sato, T. and Motoyoshi, F. : "EXPLUS—A Semantic Parsing System for Japanese Sentences", 3rd USA-JAPAN Computer Conference, 1978, 236-240.
- [Tanaka, 1979] Tanaka, H., Sato, T. and Motoyoshi, F. : "Predictive Control Parser : Extended LINGOL", IJCAI6, 1979.
- [Taylor, 1975] Taylor, B.H. and Rosenberg, R.S. : "A Case-Driven Parser for Natural Language", AJCL, 1975, Microfiche 31.
- [Tulving, 1972] Tulving, E. : "Episodic and Semantic Memory", in Tulving and Donaldson (Eds.) : "Organization of Memory", Academic Press, 1972.
- [Tulving, 1972] Tulving, E. and Donaldson, W.(Eds.) : "Organization of Memory", Academic Press, 1972.
- [Waltz, 1977] Waltz, D. L. : "Natural Language Interfaces", SIGART Newsletter, No. 61, Feb. 1977.
- [Waltz, 1976] Waltz, D. : "Natural Language Access to a Large Data Base", Computers and People, April, 1976, 19-26.

- [Wilks, 1973] Wilks, Y.: "An Artificial Intelligence Approach to Machine Translation", in Schank and Colby (Eds.): "Computer Models of Thought and Language", Freeman and Company, 1973.
- [Wilks, 1974] Wilks, Y.: "Natural Language Understanding Systems within AI Paradigm—A Survey and Some Comparisons", Stanford AIM-237, Dec. 1974.
- [Winograd, 1971] Winograd, T.: "Understanding Natural Language", Academic Press, 1971.
- [Winograd, 1975] Winograd, T.: "Frame Representations", in Bobrow and Collins (Eds.): "Representation and Understanding", Academic Press, 1975.
- [Winston, 1977] Winston, P.H.: "Artificial Intelligence", Addison Wesley, 1977.
- [Woods, 1971] Woods, W.A.: "Experimental Parsing System for Transition Network Grammar", in Rustin (Ed.): "Natural Language Processing", Algorithmic Press, 1971.
- [Woods, 1975a] Woods, W.A.: "WHAT'S IN A LINK: Foundations for Semantic Networks", in Bobrow and Collins (Eds.): "Representation and Understanding", Academic Press, 1975.
- [Woods, 1975b] Woods, W.A.: "Syntax, Semantics, and Speech", in Reddy (Ed.): "Speech Recognition", Academic Press, 1975.
- [Woods, 1977] Woods, W.A.: "Personal View of Natural Language Understanding" SIGART, No. 61, Feb., 1977, 17-20.
- [Yamanashi, 1972] Yamanashi, M.: "Lexical Decomposition and Implied Proposition", Chicago Linguistic Society, 8, 1972, 388-401.
- [Zampolli, 1977] Zampolli, A. (Ed.): "Linguistic Structures Processing", North-Holland, 1977.
- [雨宮, 1977] 雨宮真人, 島津 明: "日本語の意味解釈過程", 電子通信学会論文誌, J60-D, 8, 1977, 633-640.
- [青木, 1974a] 青木晴夫: "「意味」と言語の構造—文章論への動き", 月刊言語, 3, 9, 1974, 793-802.
- [青木, 1976b] 青木晴夫: "言語理論", 国語学, 105, 1976, 24-35.
- [電総研, 1978a] 電子技術総合研究所: "LISP User's Manual", EPICS-5-ON-4, Mar. 31, 1978.
- [電総研, 1978b] 電子技術総合研究所推論機構研究室: "拡張 LINGOL", Oct., 1978.
- [淵, 1977] 淵 一博, 田中穂積: "日本語理解システムへのアプローチ", パターン情報処理システム講演会論文集, 電総研 PIPS-R-No. 16, 1977, 145-161.
- [淵, 1979] 淵 一博: "自然言語と計算機—その内的関連を探る—" in 石綿敏雄(編): "計算機による日本語談話行動の総合モデル化", 昭和 53 年度文部省科学研究費研究報告書(課題番号 310702), 1979 年 1 月, 17-20.
- [淵, 1976] 淵 一博: "70年代の人工知能", 数理科学, 9, 1976, 14-18.
- [藤田, 1976] 藤田 毅, 鶴丸弘昭, 首藤公昭, 吉田 将: "日本語文の標準形について", 電子通信学会技報, AL 76-41, 1976.

- [月刊言語, 1977a] 月刊言語: “特集・チョムスキー理論の展開”, 3月号, 1977.
- [月刊言語, 1977b] 月刊言語: “特集・象は鼻が長い——日本語の分析”, 6月号, 1977.
- [原田, 1969] 原田信一: “構文論における変項の制限”, 英語教育, 8月号~12月号, 1969.
- [池上, 1975] 池上嘉彦: “意味論”, 大修館, 1975.
- [池田, 1977] 池田尚志: “一般化された格構造による意味表現を用いた構文解析法について” 信学論D, vol 60-D, 10, 1977.
- [井上, 1976] 井上和子: “変形文法と日本語(下)”, 大修館, 1976.
- [井上, 1978] 井上和子: “日本語の文法規則”, 大修館, 1978.
- [石綿, 1975] 石綿敏雄: “日本語の生成語彙論的記述と言語処理への応用”, 国研報告 54, 1975, 154-198.
- [石綿, 1979] 石綿敏雄(編): “計算機による日本語談話行動の総合モデル化”, 昭和53年度文部省科学研究費研究報告書(課題番号310702), 1979年1月.
- [神尾, 1977] 神尾昭雄: “数量詞のシンタックス”, 月刊言語, 6, 8, 1977, 83-91.
- [金田一, 1976] 金田一春彦(編): “日本語動詞のアスペクト”, むぎ書房, 1976.
- [北原, 1970] 北原保雄: “助動詞の相互承接についての構文論的考察”, 国語学, 83, 1970, 32-59.
- [国立国語研究所, 1951] 国立国語研究所: “現代語の助詞・助動詞一用法と実例”, 秀英出版, 1951.
- [久野, 1973] 久野 暉: “日本文法研究”, 大修館, 1973.
- [久野, 1978] 久野 暉: “談話の文法”, 大修館, 1978.
- [栗原, 1973] 栗原俊彦: “自然言語の機械処理”, 情報処理, 14, 4, 1973, 267-281.
- [松下, 1930] 松下大三郎: “標準日本口語法”, 白帝社, 1930.
- [三上, 1960] 三上 章: “象は鼻が長い”, くろしお出版, 1960.
- [水谷, 1973] 水谷静夫: “機械処理のための日本語文法”, 電気学会誌, 93, 11, 1973, 995-999.
- [桃内, 1978] 桃内佳雄: “指示連体詞「その」の文脈的機能に関する一考察”, 信学技報, AL78-31, 1978.
- [元吉, 1977] 元吉文男, 田中穂積: “未定義語を含む文の構文解析”, 情報処理学会18回全国大会予稿, No. 95, 1977.
- [長尾, 1974] 長尾 真, 辻井潤一: “自然言語処理のためのプログラミング言語 PLATON”, 情報処理, 9, 1974, 654-661.
- [長尾, 1976] 長尾 真, 辻井潤一, 田中一敏: “意味および文脈を考慮した日本語文の処理”, 情報処理, 1, 1976, 10-28.
- [長尾, 1977] 長尾 真, 辻井潤一: “自然言語処理プログラム”, 情報処理, 18, 1, 1977.
- [長尾, 1979] 長尾 真, 田中穂積編: “日本語情報処理特集号”, 情報処理, 20, 10, 1979.
- [奥津, 1974] 奥津敬一郎: “生成日本文法論”, 大修館, 1974.
- [奥津, 1978] 奥津敬一郎: “『僕ハウナギダ』の文法”, くろしお出版, 1978.

- [小野, 1978] 小野寛昭: “名前と名指し関係”, 計量国語学, 11, 4, 1978, 138-151.
- [佐久間, 1940] 佐久間鼎: “現代日本語法の研究”, 厚生閣, 1940.
- [佐藤, 1976] 佐藤泰介, 田中穂積, 淵 一博: “VISUALIZER—自然言語理解システムの立場からみた機械による空間の把握”, 電子通信学会技報, AL 76-11, 1976.
- [佐藤, 1978] 佐藤泰介: “意味構造からの日本語文生成プログラム”, 情報処理学会計算言語研究会17-1, 1978
- [白井, 1980] 白井英俊, 横尾英俊: “統語処理と意味解釈を同時に行うシステムについて”, 情報処理学会, 計算言語学研究会資料 21, 1980.
- [首藤, 1977] 首藤公昭, 鶴丸弘昭, 吉田 将: “日英機械翻訳のための述部処理システム”, 電子通信学会論文誌, J60-D, 10, 1977, 830-837.
- [鷹尾, 1976] 鷹尾洋一, 藤崎哲之助, 間下浩之, 諸橋正幸: “日本語によるデータ・ベースの照会”, IBM東京サイエンティフィック・センター, Dec., 1976.
- [高橋, 1977] 高橋太郎: “文中にあらわれる所属関係の種々相”, 国語学, 103, 12, 1975, 1-17.
- [滝浦, 1977] 滝浦静雄: “時間—その哲学的考察”, 岩波新書, 960, 1977.
- [田中, 1976] 田中穂積: “談話の理解とメモリ・モデル (その 1~3)”, 電総研彙報, 第 40 巻, 7, 9, 10, 1976.
- [田中, 1977] 田中穂積, 佐藤泰介, 元吉文男: “自然言語処理のためのプログラミング・システム—拡張 LINGOL について—”, 電子通信学会論文誌, J60-D, 12, 1977, 1061-1068.
- [田中, 1978a] 田中穂積: “日本語の意味構造を抽出するシステム EXPLUS について”, 電子通信学会論文誌, J61-D, 8, 1978, 548-556.
- [田中, 1978b] 田中穂積, 内田ユリ子: “意味抽出における概念階層の利用と手続き付加”, 情報処理学会人工知能と対話技法, 8-1, 1978.
- [田中, 1979] 田中穂積: “計算機による自然言語の意味処理に関する研究”, 電子技術総合研究所研究報告 46797, 1979.
- [鶴丸, 1976] 鶴丸弘昭, 藤田 毅, 首藤公昭, 吉田 将: “日本語の機械処理”, 電子通信学会技報, AL 76-43, 1976.
- [寺津, 1978a] 寺津典子: “T. Reinhart (1976): The Syntactic Domain of Anaphora”, メモ, 1978.
- [寺津, 1979b] 寺津典子, 山梨正明: “日本語における照応現象について (その一)”, in 石綿敏雄(編): “計算機による日本語談話行動の総合モデル化”, 昭和 54 年 1 月.
- [寺村, 1971] 寺村秀夫: “一タの意味と機能—アスペクト, テンス, ムードの構文的位置づけ”, in “言語学と日本語問題”, くろしお出版, 1971.
- [安井, 1978] 安井 稔: “言外の意味”, 研究社, 1978.
- [安井, 1979] 安井 稔: “新しい聞き手の文法”, 研究社, 1979

- [山田, 1978] 山田尚勇: “日本語テキスト入力法の人間工学的比較”, 日本語情報処理シンポジウム報告集, 1978.
- [山梨, 1979] 山梨正明: “自然言語の意味構造—含意関係からの再考察” in 長尾 真(編): “機械処理関係シンポジウム・辞書ファイル委員会活動報告書”, 昭和 54 年 3 月.
- [吉田, 1978] 吉田 将, 首藤公昭, 藤田 毅: “日本語文における曖昧さと関係表現について”, 電気四学会連合大会, 昭 53・10, 6-79~82.
- [横山, 1978] 横山晶一, 田中穂積: “意味構造からの日本語文の再構成”, 昭 53 年度情報処理学会 19 回全国大会, 昭 53, 473-474.

2. 構 文 解 析

2.1 はじめに

自然言語であれ、人工言語であれ、言語を処理するシステムにとって不可欠な構成要素として構文解析がある。構文解析とは“正しい文”と“正しくない文”を峻別し前者からその内部構造を取り出すことを言う。ここでこれから述べることはこの構文解析についてであるが、事は言語学、AI、計算機アルゴリズムに跨り簡単に概括することは仲々むずかしいので、まず「自然言語の構文解析」という観点から言語学的観点を含みつつ構文解析に光を当て、次に「構文解析のアルゴリズム」の観点から“技術としての構文解析”を探り、以下具体的な構文解析システムの紹介に移って行くことにする。

2.2 自然言語の特徴と構文解析

我々が日常使う言語 — 自然言語は、計算機言語等の人工言語にはない特徴を持っている。これらの特徴を構文解析について考えてみる。

- a) 適格文と非適格文との区別が定かでない。普段の会話を反省してみればわかるように、我々は文法的とは言い難い発話をする事が多々ある。それでも何とか用が足りているのは、聞き手が豊富な知識により話し手の言語表現のゆがみ脱落を訂正し、場面に沿って適切に発話を解釈しているからである。しかしそのような能力は機械には期待し難いので、機械にも解かるように“文”の範囲を制限する必要があるが出てくる。つまり機械にとっての適格文と非適格文を定めてやる必要があるわけだが、これは仲々むずかしいことである。例えば“言う”という語の使用を許すと、「『太郎に本を読ませたさせた』と花子が言った」などのように非適格文が補文として使われ、全体として適格文になるといったケースが出てくる。この文を構文解析するとしたら、非適格な補文も解析しなければならなくなってしまうわけである。又、出れる、見れる、食べれる、……のように適格文か非適格文か人によってハッキリしない例もある。

更に又普通の世界では「岩が話す」は非適格文であるが童話の世界では、立派な適格文であるというように場合（世界）によって非適格文が適格文に

なることもある。

従って単に適格文と言っても解析システムの目標を定め、解析されるべき文の全体、利用できる知識を想定して始めて1つの構文解析システムが扱う適格文のイメージを作ることができるのである。

b) 曖昧さ

自然言語では文の曖昧さはありふれた現象である。(計算機言語でも if …… then …… if …… then …… else …… のように else にかかる if に曖昧さが生じる場合があるが例外的である。)

曖昧さは色々な原因で生じる。構文解析レベルでは同音異義 (e.g. ニワニワトリガイルの“ニワ” — 形態素解析に関係する) , 多品詞語 (英語ではよくある。日本語でも“時”, “前”等は接続詞であると共に名詞でもある) , 修飾関係 (立派な息子の父) が代表的なものであろう。構文的曖昧さの扱い方には2通り考えられる。1つは曖昧さが生じた時点でもっともらしいものを1つ選び解析を続け、行き詰ったら他の選択肢を取るやり方 (serial 法) , もう1つはすべての可能な構文解析を同時に求めるやり方である (parallel 法) 。もちろん解析を進めて行く内に得た情報により曖昧さを取り除くこともできよう。

c) 辞書にない語, 入らない語

“ レンズ X 1 ” 等のラベル, 1, 2, 3, …… 等の自然数のように無限個ある語は辞書に入らない。又“ 第 20 回情報処理全国大会 ” のように非常に長い名詞連続も辞書には入らないとみた方が良い。このようなものについて解析システム設計者は取り扱いをあらかじめ考えておく必要がある。

構文解析の途中で辞書にない語にぶつかった時には今迄の解析結果から未知語の構文役割りを推定して多くの場合名詞) 解析を続けるというのが普通のやり方である。ついでに言えば, 未知語の構文的役割りを推定するという観点からは, 構文解析は Earley のアルゴリズムのような top-down, predictive type が有利であると考えられる。

d) 省 略

文中で語が省略されている場合, それを発見し補ってやらなければならない。文脈から明らかな語は省略されうるので, このような場合語を補う為には文

脈の理解が必要とされむずかしい問題となる。(e. g. 「僕ハウナギダ」)
簡単な場合には述語の格パターンを辞書に用意しておき、これによって必須格
が抜け落ちていたものを発見し補うことができる。

e) 指 示 詞

構文解析と直接の関係はないが、指示詞の対象と同定することは文を理解する上で大切なことである。指示されるものが指示詞の前にある(前方参照)後にある(後方参照)、稀には「コノ文ハ デス」の如く自分自身を指すものが陽に表現されていない(「2に3を加える。それに……」)場合も多く、問題は困難を極める。

さて以上述べてきた事は構文解析と自然言語の特徴一般との関係であるが、構文解析を言語理論的に捉えたとすれば、それは言語理論に依存するはずである。そこで幾つかの言語理論の枠組みを眺めておこう。

Chomsky の初期の定式化(標準理論60年代中頃)によるとCFG (Context Free Grammar)により基底部門で生成された句構造標識(Deep tree)は意味部門で意味解釈を受けて意味表示を得る一方、変形(transformation)を受けて表層のtreeとなり音韻部門で発話として具体化される。

この逆過程が初期Chomskyに沿った構文解析であるが、これを忠実にを行ったものとしてZwickyやPetrickの仕事がある。例えばZwickyは表層文を受け入れる文法(「surface grammar」)により文をparsingした後逆変形ruleを適用し、出て来た結果が基底部門のtreeになっているか確めた。途中で逆変形の過程が入っている為困難が倍加し処理速度は遅かった。

Chomskyの理論はその後拡大標準理論、拡大修正理論と進展し、次第に表層(に近い)構造の意味解釈の役割りが大になっている。従って構文解析的に言えば、表層文及び逆変形中の構造の意味解釈を採用しつつもとの構造へと文が変形されて行くことになる。

* 変形文法派を大きく分けると、解釈意味論派と生成意味論派に分けることができる。
解釈意味論派は表層文の意味表示が解釈によって得られるとし、一方生成意味論派は、
文の意味表示そのものから出発して、変形過程により表層文が生み出されるとする。

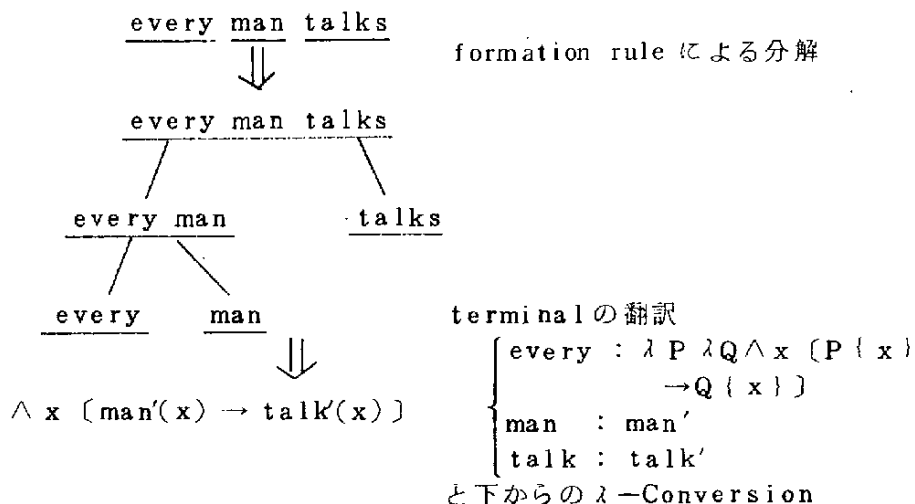
一方生成意味論派の Fillmore の考えによると文のものの構造は

Sentence $\rightarrow$ Modality + Proposition

の形をしている。Modality は様相, 時制, aspect, 否定等, Proposition は述語とその格範ちゅうから成る。述語は固有の格構造 (Case Structure) を持ち, 名格 (Case) は意味的役割りによって Objective Case, Agentive Case, Instrumental Case 等々に分類される。例えば, "Open" の格構造は, + [- O (I) (A)] であり - の位置に open が生じ Objective Case は義務的に, その他は任意に生じる。Modality と以上述べたような述語の格構造から変形によって表層文が生成されるわけであるから, Fillmore 流に言えば, 表層文から何らかの変形により, 文を Modality と格構造に分解することが文の構文解析である。述語の格構造の考えは Minsky の Frame, Schank の Primitives の理論と結びつき計算言語学の分野では広く使われている。

以上の変形文法論的アプローチと異なるものとして Montague の考えがある。Montague 文法は Categorical Grammar と Intentional Logic を融合させたものであり, 自然言語の文を論理式として捉える。Intentional Logic は高階の様相述語論理の一種であり, タイプ付き λ -Calculus を notation として採用している。Montague 流に言えば表層文を Intentional Logic の式に翻訳することがその構文解析である。翻訳によって得られた λ 式を可能多世界から成るモデルの中で評価することによりもとの文の真偽が決定される。

< Montague Grammar の構文解析例 >



この図でわかるように表層文を分解することにより tree が得られるが、これは末端に λ 式をぶら下げた Program の木になっている。結果的に LINGOL⁽¹⁰⁾の処理過程に似ている点が興味深い。Montague文法はまだ未発達(疑問文、命令文、morpheme……の扱い)であって今すぐ役に立つものではないが、その背後に厳密な semantics を持つという点で優れていると言えるだろう。

国文法に於ける構文解析の概念は、以上述べて来たような言語理論と比べるとはっきりしていないようである。ただ文の成分間の係り受け、辞と辞、詞と辞の接続という考えがあり、これがいわゆる構文解析に相当すると思われる。文の係り受け解析については後程触れる予定である。

2.3 構文解析用のアルゴリズムとシステム

計算機による構文解析を考えると、構文解析とは結局アルゴリズムである。構文解析用のアルゴリズムは自然言語の研究よりも形式言語の研究の中で発達をとげたので、ここで形式言語とその構文解析について少し基本事項を述べる。

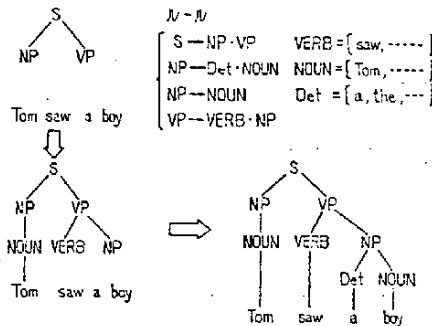
形式言語 (Formal Language) はオートマトン理論の分野で良く研究されている、一つの形式言語は使用する記号と書き換え規則とにより定められる。形式言語をクラス分けすると 3 型言語 (正則, regular), 2 型言語 (文脈自由, context-free), 1 型言語 (文脈依存, context-sensitive), 0 型言語の順に広い言語族を成し、各族に属する言語の受理機構として, Transition Net (TN), Push Down Automata (PDA), Linear Bounded Automata (LBA), Turing Machine (TM) が対応する。0 型言語を除いて各タイプに属する言語は帰納的集合 (recursive set) であるが、0 型言語は任意の帰納的に予算な集合 (recursively enumerable set) が生成できる。

C F G (Context Free Grammar, 文脈自由型文法) については良く研究され、また具体的システムへの応用も多いので以下 C F G に焦点を絞る。

C F G の構文解析の結果は構文解析木 (parsing tree) を使って表示される。図-1 に例が示されているように構文解析木は木の形で文の内部構造を表現したものである。構文解析アルゴリズムは木の作り方により、top-down / bottom-up serial / parallel の区別がある。top-down 法は木を上か

ら作り, bottom-up 法は木を下から作る (図 2-1 参照)。一度の構文解析により serial 法は一つの解析木を作ろうとし, parallel 法はすべての可能な解析木を作ろうとする。

(a) top-down 法: ルールは下記参照



(b) bottom-up 法: ルールは top-down 法に同じ

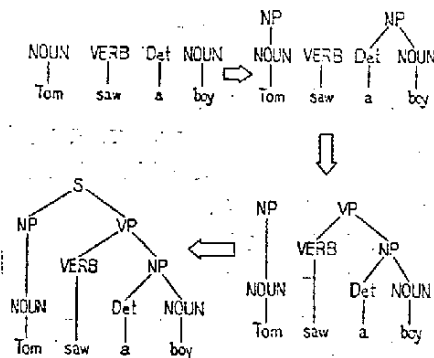


図 2-1 top-down 法と bottom-up 法構文解析

top-down & serial または bottom-up & parallel の組み合わせでパーザ (parser) が作られることが多い。

図-1 の(a)に示した top-down & serial 法を使う時は, 解析が行き詰った時のバックトラックメカニズムが必要である。また $S \rightarrow S \cdot NP$ などの再帰的ルールや, $A \rightarrow B$, $B \rightarrow A$ などサイクルを成すルール群については, 取り扱いに注意を払わないと無限に木を成長させ続け解析が終らない恐れがある。

Top-down法は通常入力文の長さに対し指数関数的に処理時間が伸びるが、WFST (Well Formed Substring Table)を利用することにより $O(n^3)$ に押さえることができる。WFSTは構文解析の途中で作られたPPT (Partial Parsing Tree)を貯めておくテーブルで、バックトラックが起きてもそれまでの部分的に正しい結果を残しておき、後で活用するためのものである。これにより構文解析の過程で同じPPTを2度作らなくなり、処理速度の改善につながる。

図2-1(b)に示したbottom-upなやり方では再帰的ルールについては心配しないが、サイクルを成すルール群に関しては無限に木を成長させてしまう恐れがあるので取り扱いに注意が要る。bottom-upな解析法で有名なのはYoung-Cocke-Kazamiの方法で長さ n の入力に対し $O(n^3)$ の時間で解析を終える。¹⁾

Bottom-up法の欠点は、木を下から上へと作って行く時、果してその木が最終的な構文解析木の一部に成りえるのか全く気にしない為、無駄な木を生成しがちなことである。この点を改良したのがEarleyのアルゴリズムである(図2-2参照)。このアルゴリズムはtop-down過程とbottom-up過程を持ち、

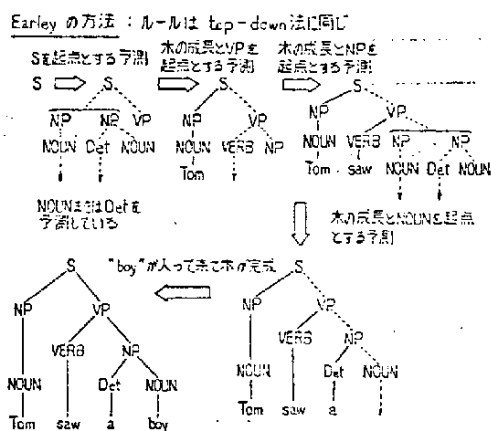


図2-2 top-down法とbottom-up法の融合

前者は成長すべきPPTのトップノードのカテゴリを予測し、後者はその予測

に従って P P T を実際に成長させる。長さ n の入力に対し、処理時間は文法に曖昧さがある時 $O(n^3)$ 、ない時 $O(n^2)$ である。

Earley の方法では top-down (予測) 過程は文法ルールの集合と起点のカテゴリが決まれば入力文に関係なく決ってしまう部分である。この点に着目して、ルールが与えられた時点で top-down 過程を "precompile" してビットテーブルの形で覚えおき、実際に入力文が与えられた時そのテーブルを参照することにより同じ top-down 過程を何度も繰り返さないようにすると Pratt のアルゴリズムになる。このアルゴリズムは言語処理システム LINGOL⁽¹⁰⁾ の基礎になっている。

Young-Cocke-Kazami, Earley, Pratt の方法は一般的アルゴリズムであるが、もし文法に曖昧さが無い場合にはさらに速く解析できる。LR(k) は Pratt の方法の中で上に P P T を成長させて行く際あらかじめ計算された先読みストリングセットの中に、実際の先読みストリングが入っている時のみ P P T を成長させるもので、"文法の precompile" を更に徹底したものと考えることができる。

以上紹介したアルゴリズムはいずれも純粋な構文解析法であり、自然言語の構文解析にそのまま使って適切かどうかは判断の分かれる所であるが、WFST Earley, Pratt, LR(k) 等の手法に見られる。"同じ木を2度作らない", "文法が与えられた時、そこから最大限の情報を引き出してコンパイラとしておく" 等の考え方自体は自然言語の構文解析システムを作る際参考になるものと思われる。

2.4 構文解析システム

ここでは形式言語の構文解析手法を背景とする自然言語構文解析システムを取りあげる。前節のアルゴリズムミックスな解析法と異なり、自然言語の多様な現象を取り扱う為いろいろな工夫が成されている。

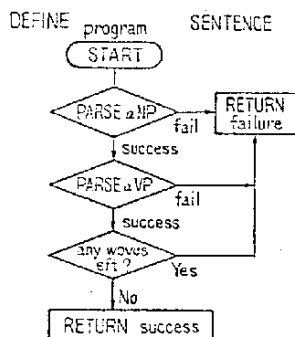
言語学的観点に立ったシステムは Harvard predictive analyzer [Kuno 1962] や MITRE [Zwicky et al. 1965] 等 1960 年代に試みられたが^{3), 4)} 1970 年代に入ると AI (Artificial Intelligence, 人工知能) の立場に立ったシステムが作られるようになった。想定されている言語は通例 CFG

+αである。

積木の世界に関するQ.A.システム SCHRDLU [Winograd 1971]の構文解析部分は PROGRAMMAR と呼ばれ,⁹⁾ Lisp プログラムの集まりである(図2-3(b)参照)。

(a) ルール $S \rightarrow NP \cdot VP$ -----
 $NP \rightarrow Det \cdot NOUN$ -----

(b) PROGRAMMAR



(c) ATN

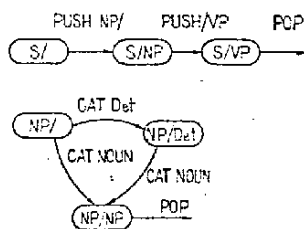


図 2-3 さまざまなコントロール方式

Hallidayのシステミック文法に基づいた定式化になっている。構文解析には積木の世界の知識を使い、top-down方式で曖昧さのある場合には良さそうな木を一つ作ろうとする。構文解析のすべてをプログラムの行うことによりシステム作製の柔軟性を得ている反面、システムの修正・変更・理解が困難になるという欠点が明らかになった。

このようなプログラムで何でもやろうという PROGRAMMAR の立場とは異なり、構文解析のコントロール部分を Transition Net (TN) の形にしたものが ATN (Augmented Transition Net) [Woods 1970] である。⁶⁾

ATN は TN を augment (強化) したものであり、状態の遷移は使用者によって書かれたプログラムをチェックする事により行なわれ、ついでにレジスタに情報をセットできる。状態間の情報交換はレジスタを介して行なわれる (図 2-3(c) 参照)。

ATN-フォーマリズムはさまざまなシステムに採用されているが、日本語の語順が比較的自由であるという特性を考慮して、パターンマッチングの機能を付け加え、また功妙なバックトラックメカニズムも加えたものが PLATON [長尾・江井 1974] である。²⁴⁾ パターンマッチングにより語順の自由度をうまく吸収してルールを適切に起動することができる。

ATN は基本的には top-down, depth-first であるが、逆に bottom-up breadth-first に処理を進めるのが Heidorn の APSG である [Heidorn 1978]⁷⁾。各ルールに対し適用条件、適用時の動作がプログラムで書け、処理結果は“レコード”と呼ばれるデータ構造にまとめられる。

APSG と同様に bottom-up であるが、よりアルゴリズム的に洗練されているものに LINGOL [Pratt 1973] がある。¹⁰⁾

これは前節に述べた Pratt のアルゴリズムに基づき、各ルールに COG 部, SEM 部の 2 つの使用者に解放された部分を付け加えたものである。使用者は COG 部に複数個の PPT ができた場合の優先順位を、SEM 部に構文解析後に行うべき処理のプログラムを書くことができる。

LINGOL を改良して、下から PPT を成長させる所で使用者の観点からして好しい木のみを選択的に成長させることができるようにし、又形態素解析に伴うバックトラックメカニズムを付け加えたものが拡張 LINGOL である。拡張 LINGOL は形態素解析と構文解析が一体化しており、左から右へと文を解析する途中で意味情報を利用することもできる。

以上述べて来たシステムとは別の発想からできたシステムもある。^{4), 46), 47)} 例えば Marcus [Marcus 79] によると、従来のシステムに見られる“無駄な木を作ってから後でそれを捨てる”という事態は、語义はより大きな単位を

最大3つ先読みしつつ、決定論的に解析を行うことにより防げる。この為には、*sift-reduce-parsing*を少し工夫し、大域的展望（*top-down*的要素）を導入し、更に又、スタックにできたPPTを入力文に戻す（この為先読み要素が名詞句など語以上の単位を含むことになる）操作も導入する。以上のような *parsing method* は人間の心理的プロセスに近いとも主張している。

又、構文解析と言っても余り *syntactic* な要素を含まないシステムもあるがここでは省略する。〔Schank et al. 79〕⁴⁶⁾、〔Rieger et al. 79〕⁴⁷⁾

最後に、*parsing algorithm* を定理証明として見る見方があることも付け加えておこう。^{42), 43)} これについては新しいプログラミング方法である *Logic Programming* との関連で見るべきであるが本題をはずれるので省略する。

2.5 日本語の構文解析

(1) 日本語の特徴と構文解析

日本語の持つ特徴がその構文解析に与える影響について幾つか取り上げてみたい。

日本語は漢字入力法に関する問題があるが、それはともかく構文解析で問題になるのは、日本語に分かち書きの習慣がないという事である。したがって構文解析のため、文から語を切り出す必要があるが、“／ニワ／トリ／ガイル”のように、切り方により文の意味が異なる場合もあれば、“ニワトリ”と“ニイク”と間違えて切り、失敗することもある。通常は字種の切れ目情報や最長一致（*longest match*）を用いるが、上の例で分かるように常にくまなくいくとは限らない。

構文解析を採用して語を切り出す方法もあり、この場合辞書に、ニワトリ、ニワ、トリが入っているとすれば、“ニワトリガイル”は二つの構文解析木を、“ニワトリニイク”は唯一の構文解析木を生じるであろう。切り方が複数ある時、正しい切り方を選択するためには意味、文脈情報の利用が不可欠である。

第2に日本語は語順が自由であると言われている。したがって2.1節で述べたような純アルゴリズム的方法は適さないとも考えられる。しかしながら良く考えてみると、日本語単文を、名詞句の並びに一述語一助動詞・補助用

言等一終助詞として捉えた場合、自由度があると言われているのは、主として格を成す名詞句の並びであり、名詞句内部や助動詞等を含む述語文節内部に自由度があるとは言えないであろう。3節で述べたアルゴリズムとの関連で言えば、述語から終助詞までは正則文法ではほぼ解析できると言われており、格を成す名詞句の並びの自由度もルールの作り方によって吸収できる。

副詞に関しても述語文節内部に割り込ませることはできず、特に終助詞の後には如何なる語も接続しない。総じて、文末に近づくほど語順が厳しくなる（この辺の事情については(13)参照）。更に基本的な格の並びには順序があると言われており、例えば存在文では場所格が先頭に来る例が多い。また談話の観点からしても、旧い情報を担う語が文の先頭の方に、新しい情報を担う語が文末の方に来ると言われ、談話レベルの情報構造が名詞句の並びを統制していることも考えられる。したがって日本語に於いては、係る文節が受けの文節より左にある限り基本的語順は確かに自由であるが、他の原因から決っている場合もあり、2.1節で述べたような方法が適用できる部分もあるとするのが妥当であろう。

第3に同音異義語が多い。英語には“Time flies like an arrow”の例が示すように名品詞語が多いが、(time, flies : 動詞, 名詞, like : 動詞, 前置詞)日本語には“キタ”(北, 着た, 来た)のように同音異義語が多い。そのため入力文がカナ表記・ローマ字表記の場合の曖昧さを生じる事が多く、構文解析上問題となる。

同音異義語による曖昧さを消すために、文の先頭に動詞は来ないから“キタへ行ク”の“キタ”は“北”であるというように統語情報を使う事もあるし、“葉”は意味素性が—animateなので、“葉ガキイタ”の“キイタ”は“聞イタ”ではなく“効いた”である、というように意味が使える事もある。勿論文脈や常識を使って初めて解決できる曖昧さもある。

第4に省略が多い。日本語では文脈から明らかな語は省略されるのが普通である。省略語を補うためには広く談話レベルの情報まで必要とされるので、非常に困難な問題である。

第5に日本語はSOV型言語であり、次節の係り受け解析の項で述べるように、修飾関係は原則として文の前方から後方へ向き、重要語は後ろに来る

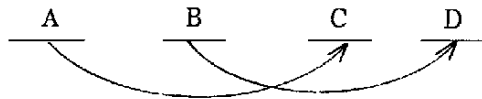
と言われている。このことは文全体の意味関係を捉えるため、まず重要語を文の後方から先頭へと探索して見つけ出し、この語を鍵として意味解釈を進めるという方法を示唆している。実際このやり方を採っているシステムもある。

その他、日本語は左枝分れ型言語であるとか、敬語表現が発達している等の特徴があるが紙面の都合で省略する。

(2) 日本語の係り受け解析

日本語文を文節の有限列として見た場合、文節間に成り立つ広義の修飾関係を係り受けと言い、大まかに連体修飾関係（“花しい花”）、連用修飾関係（“美しく咲く”）、格修飾関係（“花が咲く”）に分類できる。修飾する方を係りの文節、される方を受けの文節と言うことにすると次の原則が成立する。

- (i) 係り受け非交差……係りの文節から受けの文節へ矢印を書いて係り受け関係を示した場合、矢印は交差しない。



のようなことはない。

- (ii) 文末にない文節は、その右側のいずれかの文節に係る。

この原則をもとに係り受け関係を定めて行くわけであるが、アルゴリズム的に言うと、(i)の非交差条件からして bottom-up に構文解析を作っていることに等しい。これは係りと受けを括弧でくくれば明らかである。しかしながら、係り受け本来の意味からして、係り受け解析は意味的なものであり、意味情報主導型の構文解析であると言える。

具体化に当たっては(i), (ii)の条件を活かしスタックを使って、文の先頭から処理する（係る方を捉え、後で受ける方を探す）か、文の後方から処理する（受けとなる文節を先に捉え、後でそれに係るものを探す）方法が考えられる。

なお、日本語の係り受け解析に相通ずるものとして、Hays らの依存関係理論がある。この理論は文の構成要素の依存関係を重視する立場から、依存関係

規則により句構造規則では表示されない支配頂 (governor, 受ける方) と被支配頂 (dependent, 係る方) の依存関係を明示する文法である。

2.6 日本に於ける構文解析システム

日本に於ける構文解析システムは、係り受け解析を基礎としているものが多く、書き換え規則を使った統語規則 (syntactic rule) 主導型の解析システムは少ない。一般に日本語構文解析システムを設計する場合、係り受け解析等の意味主導型にするか、統語規則主導型にするか、二つの選択肢があるが、夫々異った特性を持つ。前者は積極的に意味・文脈情報を使いつつ統語・意味解析を同時に行うが、その為システムの理解のし易さ、モデュラリティが失なわれ易い危険性がある。後者は文法規則が明示され、アルゴリズムも良く研究されているが、解析過程のコントロールが統語規則に依存しがちで、コントロールの柔軟性が失われる危険性がある。しかし両者は決して排他的なものではない。係り受け解析によっても構文解析が作られ、また統語規則は意味情報によって強化・運用される必要があるのは当然である。統語解析と意味解析を同時に行うことについては Woods の意見も参考になるだろう。⁶⁾

次からは日本の構文解析システムを幾つか取り上げて概説する。

言語学者の作ったシステムとして IRISOA がある。²⁷⁾ 石綿の作ったこのシステムは依存関係文法を基礎として、変形と逆変形を探知する規則を含み、扱う言語と理論を特定しない述語中心の解析システムである。辞書と文法はプログラム本体と別になっている。処理は①辞書引き、②主要述語の発見、③文型をパターンマッチにより確定した後被支配頂を求める。これを繰り返して文にあるすべて支配頂を求める、④変形の逆探知などして解釈のつかない被支配頂の支配頂を求める、⑤結果の出力、の順に行なわれる。同音異義語に対してはすべての可能な解釈の組み合わせを作ってから処理する。日本語を扱う時、入力はいカタカナ分ち書きで入れている。

石綿と同じく依存関係文法に基づいたシステムに石原等の D-tree を使った英日機械翻訳システムがある。^{28), 29)} D-tree は文中の“語群”の依存関係をセマンティックネットとして表現した木構造で翻訳のための中間言語として使われる。英文入力は単語・熟語関係の辞書により各単語が識別され次に語群辞書

により各単語がまとめられて語群系列(WGC)に変換される。例えば"the dog", "have sent", "look for"は夫々一つの語群にまとめられる。次にWGCは規則によりブロックに分割され、ブロック中の述語に機能パターン(一種の文型)を割り当ててから、二つの語群間の文脈によらない依存関係を示してある第一種の依存規則を使って、各ブロックに対し部分D-treeを作る。次により大域的な依存関係を示した第二種の依存規則を使って、部分D-treeを二つずつまとめ(等位接続、関係節等の処理)、入力文全体に対応するD-treeを得る。翻訳は語群の訳出順序を決定した後、訳語辞書により語群ごとに対応する訳語を求め、適切な訳語の選択をして訳文を得る。依存関係の曖昧さに関してはすべての解を求めている。

日本語の係り受け解析については、吉田等のグループが実証的研究を進めている。^{21), 22)} 彼等は2文節間の係り受けを7つのタイプに整理する。そして各タイプの係り受けについて、与えられた係り受けを標準的係り受けの形に還元するための手法を提案している。辞書には標準的な係り受けの、"意味"が記述されているので、非標準的係り受けの意味も上述の手法により標準形に直してから求めることができる。更に、これを基盤として"文の標準型分解"の概念・方法が示され、意味処理への展望を開いている。また日本語文の述部についても検討しており、²³⁾ 活用、付属語の意味、アスペクト等を考慮した述部処理システムが作られている。ローマ字書きの入力を与えると最長一致により語・付属語をleft-to-rightにバックトラックをともしつつ切り出し、切り出された単位間の接続条件のチェックを行う。チェックがすむと次に例えば推量、断定等の意味表示を伴った分ち書きを求め、これから英語述部へと翻訳が行なわれる。

京大の長尾等はATNを改良したPLATON²⁴⁾を用いて中学校理化—化学の教科書を対象とした文理解システムを作っている。^{25), 26)} 統語処理・意味処理・文脈処理を同時に行う方針を採り、出力結果はセマティックネットである。単文の処理は、文中から動詞を取り出し(受け方の)、その動詞の前にある名詞句(係りの方)を格助詞を手懸りとして抽出する所から始まる。次に名詞のカテゴリや意味的役割り、上位—下位関係など、が記述されている名詞辞書その他により、必要ならバックトラックを伴いつつ名詞句内部の係り受け関係を確定し、次に格文法の考えに従って整理された動詞辞書を使うことにより、動詞の格構造から

動詞と解析ずみの名詞句との係り受け関係を確定する。もし埋まらない必須格があれば、Mスタック、KMスタック（それまでの解析結果から得た、または存在が推論できる名詞が置いてある）を探して埋めようとするが、埋まらない時には“トラップ”を設けて、後続文の処理により埋めてくれる名詞句ができてくるのを待つ。重文・複文は単文処理を再帰的に行うことにより処理している。M/KMスタックを使ってアナホーラ（指示・照応）の問題も扱っており、その後の我が国の言語解析の研究に影響を与えた。

電々公社の島津等は算術と地理に関するQ.A.システムMSSS78を作製している。^{30),31)}そこでの構文解析手法を見てみよう。

入力文はカタカナ分かち書き自由で与える。処理は語い辞書（辞と辞、詞と辞の接統関係を transition net の形で記述）を使って文から文節を切り出す所から始まり、次に左から右へ文節間の係り受けを調べつつ、より大きな係り受け関係にまとめる。その為の知識・手続きは、詞には付属している case patterns おかれており、処理対象となる文節と照合することにより起動される。case patterns は“名詞”に対しても与えられ、その詞が受ける文節の統語的・意味的条件、接統チェックの為の手続き、バックトラックの為の手続き、Q.A.の為の手続きなどが書かれている。解析は、スタックを使うことにより、shift reduce parsing¹⁾ 的に行なわれる。また Rlist と呼ばれるリストによりバックトラックの為の状態保存、可能な他のパスの保持が成される。曖昧さに関しては全部の可能な解を出すようにしている。

電総研の池田は格構造の考えを一般化し^{32),33)}、述語のみならず名詞等品詞一般も格構造（受けのあり方）を持つとして構文・意味解析システムを作製している。このシステムはすべての文法的知識を辞書の中に持ち、入力スタックに置かれた文節以下に分かち書きされた入力文を、先頭から shift reduce parsing 的に処理する。結果は句（一つの詞を中心とする意味的まとまり）から成る木構造として出力スタックに残る。辞書には、品詞、活用、辞の接統情報、格構造、係りの構造（語が他の語に係るあり方）が記述されている。補文構造を持つ使役・受身の助動詞、～たい、～もらう等の辞については、辞書に書かれた語の格構造を変形して処理している。連体修飾構造は処理ルーチンを再帰的に用いて処理し、また曖昧さに関してはすべての可能な解を求める。しかし、

係りの良さ（例えばより多くの依存関係を持つ句のあり方が良い）を求めることにより、句の解釈結果に優先度を設け、またそれにより不要な句の解釈を捨てることもある。指示語、省略語の同定も試みられている。

同じ電総研の田中等は今までの係り受け解析型の構文解析とは違い、拡張 LINGOL を使ったアルゴリズム的な構文解析を行っている。³⁴⁾

プログラムにより強化された CFG ルールと辞書を与えると文の構文解析が可能になる。入力ローマ字分ち書き自由で、話を最長一致で、構文情報を利用しつつ切り出し、left-to-right に高速に処理する。曖昧さがある時にはすべての構文解析木が得られる。さまざまなユーティリティプログラムが用意されており、会話的に構文解析システムを作ることができる。現在、拡張 LINGOL を使って日本語文の意味抽出システム EXPLUS が作製されている。³⁵⁾ EXPLUS の意味抽出過程はまずパーズングにより、各ノードに意味抽出のためのプログラムが付属した構文解析木を作る。次にこの木を手懸りに、各プログラムを上ノードから起動することにより係り受け等の意味抽出を行い、フレームの形にまとめるという 2 段階方式を採っている。

パーズングは left-to-right に行うが、意味抽出過程では、日本語文の重要語は文末側にあるという事情を考慮して right-to-left に行っている。意味構造の表現としてフレーム表現を使っているが、そこでは上位一下位関係、全体一部分関係、意味素性等が体系的に使われている。

最後に紹介するヤチマタは、日本語によるデータベース照会システムであり、実用的色彩が濃い。^{36), 37)} データモデルは関係形式に似た名詞表と動詞表から成り、夫々名詞句と動詞句に対応する。システムに対し＜名詞句＞は？ という形の文を問かけると、＜名詞句＞に対応した名詞表が得られる。

入力文はカタカナ分ち書き自由であり、REL¹²⁾ 等で用いられた Kay のチャート式アルゴリズムにより bottom-up に構文解析木を作る。木の終端の名詞にはデータベースの名詞表が対応する。ヤチマタの意味処理とは、名詞句と名詞句との関係を、基本的なデータベース検索演算とその引数の関係に変換する（意味の木を作る）事であるが、この作業は構文解析と同時に遂行される。文法は制限されており、例えば動詞句は名詞句を修飾する形のみ許される。意味処理を終えても曖昧さが終る場合（複数の検索コマンドが出きている）には使用者の指

示に従う。

以上採り上げたシステムは、1979年末迄に発表されたシステムであり、それ以降公開されたシステムについては調査をしていないことを断っておく。

2.7 おわりに

構文解析は自然言語処理にとって不可欠であり、また重要な位置を占るものである。ここに、構文解析システムを設計する際考慮すべき項目を思いつくまま列挙して結びとしたい。

- A) 背景となる言語学的枠組み：国文法，変形文法，Montague文法，自家製文法
- B) アルゴリズムの枠組み：係り受け解析，CFG+ α (α の中味)，その他
- C) 形態素解析・構文解析・意味解析の相互作用：密接にする，疎にする
- D) コントロール：スタック，ATN式，LINGOL式，その他
- E) 入力の形態：分ち書きの問題，字種
- F) 指示語，省略語，未知語，同音異語・多品詞語，並列表現，「の」による名詞連続，漢語による名詞連続，関係節，補文構造の取り扱い
- G) 文法的知識の表現：どこに置くか？ フレーム的表現，ネットワークによる表現，その他，上位一下位・部分一全体関係，意味素性，常識の利用

参考文献

- 1) Aho et al.: The theory of parsing, translation and compiling Vol. 1, Prentice-Hall, Inc. (1972).
- 2) Hopcroft et al.: Formal language and their relation to automata, Addison-Wesley (1969).
- 3) Kuno: The predictive analyzer and a path elimination technique CACM, 8 (1965).
- 4) Zwicky et al.: The MITRE syntactic analysis procedure for transformational grammars, AFIPS, FJCC, Vol. 27 (1965).
- 5) Woods: An experimental parsing system for transition network grammar, Natural language processing, in Rustin (ed.), Algorithmic Press (1970).
- 6) Woods: Semantics and quantification in natural language question answering, Advances in computers, Vol. 17, Academic press (1978).
- 7) Heidorn: Augmented phrase structure grammar, Theoretical issues in natural language processing, in Schank and Webber (eds.), Cambridge, Mass (1978).
- 8) Marcus: Diagnosis as a notion of a grammar, Theoretical issues in natural language processing, in Schank and Webber (ed.), Cambridge, Mass (1978).
- 9) Winograd: Procedures as a representation for data in computer program for understanding natural language, Academic Press (1972).
- 10) Pratt: LINGOL-A progress report AIJCAI (1973).
- 11) Grishman: A survey of syntactic analysis procedures for natural language, AJCL, microfiche 47 (1976).
- 12) Thompson et al.: Practical natural language processing: The REL system as a prototype, Advances in computers, Vol. 13, Academic Press (1975).
- 13) 久野 暉: 日本文法研究, 大修館 (1973).
- 14) 奥野敬一郎: 生成日本文法論, 大修館 (1974).
- 15) 井上和子: 変形文法と日本語, 大修館 (1976).
- 16) 柴谷方良: 日本語の分析, 大修館 (1978).
- 17) 安井 稔: 新しい聞き手の文法, 大修館 (1973).
- 18) 北保保雄: 助動詞の相互承接についての構文論的考察, 国語学, 83 (1970).
- 19) 西村他: 日本語基本文法単文篇, 電機研い報 No. 783 (1978).
- 20) 吉田他: 日本語の機械処理 日本語文の意味さと関係表現について, 電気四学会連合大会 (1978).
- 21) 吉田 将: 二文節間の係り受けを基礎とした日本語の構文分析, 信学論 D, Vol. 55-D, No. 4 (1972).
- 22) 吉田他: 日本語の機械処理——構文分析から意味分析へ——, 信学論 D, Vol. 55-D, No. 6 (1972).
- 23) 首藤他: 日英機械翻訳のための述部処理システム, 信学論 D, Vol. 60-D, No. 10 (1977).
- 24) 長尾他: 自然言語処理のためのプログラミング言語 PLATON, 情報処理, Vol. 15, No. 9 (1974).
- 25) 長尾他: 意味および文脈処理を用いた日本語文の解析——名詞句・単文の処理——, 情報処理, Vol. 17, No. 1 (1976).
- 26) 長尾他: 意味および文脈処理を用いた日本語文の解析——文脈を考慮した処理——, 情報処理, Vol. 17, No. 1 (1976).
- 27) 石綿敏雄: 変形とその逆推知を含む構文分析, 電子計算機による国語研究, 国研報告59 (1976).
- 28) 石原他: D-tree モデルとそれに基づく英日機械翻訳の為の言語分析について, 信学論 D, Vol. 57-D, No. 7 (1974).
- 29) 石原他: D-tree モデルに基づく一つの英日機械翻訳システムおよび実験, 信学論 D, Vol. 57-D, No. 7 (1974).
- 30) 雨宮他: 図形世界を話題とした質問解答システム, 情報処理, Vol. 18, No. 8 (1977).
- 31) 島津 明: 日本語の構文・意味解析——質問応答システム MSSS 78 の試み——, 信学会, AL 研究会, AL 78-94 (1978).
- 32) 池田尚志: 一般化された格構造による意味表現を用いた構文解析法について, 信学論 D, Vol. 60-D, No. 10 (1977).
- 33) 池田尚志: 語の構造・意味記述と文の構造・意味分析, 信学会, AL 研究会, AL 78-32 (1978).
- 34) 田中他: 自然言語のためのプログラミングシステム——拡張 LINGOL について——, 信学論 D, Vol. 60-D, No. 12 (1977).
- 35) 田中西積: 日本語の意味構造を抽出するシステム EXPLUS について, 信学論 D, Vol. 61-D, No. 8 (1978).
- 36) 猪俣他: ヤチマタにおける擬似日本語, 情報処理, CL 研究会, CL-8 (1976).
- 37) 藤崎他: データベース照会システム「ヤチマタ」と名詞句データ模型, 情報処理, Vol. 20, No. 1 (1979).
- 38) Fillmore 著, 田中春美他訳: 格文法の原理——言語の意味と構造——, 三省堂 (1975).
- 39) Hays: Dependency theory: a formalism and observations, Language, 40, 4 (1964).
- 40) Kay: The MIND system, Natural Language Processing, in Rustin (ed.) Algorithmic Press (1971).
- 41) Colmerauer: Metamorphosis Grammars, in Natural Language Communication with Computer, Lectures Notes in Computer Science, Springer Verlag (1978).
- 42) 洲 一博: 定理証明としてみた Earley/Pratt のパーシング・アルゴリズム, 情報処理学会 17 回全国大会 (1976).

- 43) 西岡, その他: 導出原理を用いた構文解析システム
信学会 AL研究会 AL76-74 (1976)
- 44) 首藤, その他: 日本語の機械処理のための文節構造モデル
信学論D Vol. J62-D, №12, (1979)
- 45) MARCUS : Ateory of Gntactctic Recognition for Natural Longuag,
Aritificial Intelligence : An MIT Perspective Vol. 1
(Winston & Brown ed.) MIT prass (1979)
- 46) Schank et al: Parsing Directly in to Knowledge Structures,
6th IJCAI (1979)
- 47) Rieger et al: Word Expert Parsing,
6th IJCAI (1979)

3. 知 識 表 現 } 次 年 度 調 査
4. 文 の 生 成 }

5. 問題解決と推論 Default Reasoning

不完全な知識のもとで行なわれる推論を default reasoning と呼ぶ。人間は、話題についての完全な知識を持たなくても対話をする事ができる。自然言語処理などの人工知能の分野やデータベースにおける default reasoning がどのように表現され、また、取り扱われるべきかをいくつかの例をもとにして説明する。

5.1 変数割当てにおける default

KRLやFRLなどの多くの知識表現向きの言語においては、変数値の欠如に対する割当て (default 値) を明に示す機構が組み込まれている [Bobrow 77]。ある状況や事柄に対してフレームを構成する際に、それに関するあらゆる情報を明示する事は、困難というよりは、不可能である。ところが、実際の会話などに現われる対象物のほとんどは曖昧性を持つ事が多い。例えば、会話の対象となっている物体の色であるとか、大きさであるとかは、たとえそれらが何らかの値を持っているにもかかわらず、明示されずに対話が続けられうる。これらの中には不完全なままの方が都合のよい場合もあるが、常識的な値を default 値として割当てる事も考えられる。例えば、KRLでは default は次のように表現される。

```
{Person UNIT Basic
```

```
:
```

```
<hometown ( a City ) Palo Alto ; DEFAULT >
```

```
:
```

```
}
```

すなわち、人を表現するフレームの出身地が不明の場合には default として Palo Alto とせよ、という事を意味する。

しかし、default としての適切さを考える時には、上のような表現法は好ましいとは言えない。例えば、「ある人の出身地が不明であるが、両親が A 市に住んでいる事がわかっている場合には、その人の出身地も A 市にすべきである。」というような推論規則は、もっともであるけれども、KRLなどの上では表現

する事が難しい。

5.2 論理データベースにおける default

例えば、飛行機の座席予約に関するデータベースについて考えてみる。A市とB市を直接結ぶ便があるかどうかを知りたいとすると、論理データベースに対する質問は、例えば、 $\text{DIRECT-CONNECT}(A, B)$ と表わすことができる。もしこの述語の証明に失敗したとする。この時は、そのような便は存在しなかったと仮定するのは妥当であろう。しかもこの仮定を置く事によって、データベースから述語の否定に関する情報をすべて取り除く事ができる。すなわち、データベースは、余分な知識を持たなくてもよいのである。その上、具体的なデータベースを見てもわかるように例外事項は、実際に必要なデータより遙かに多いのである。このような仮定を置いたデータベースを Reiter は closed world database と呼んでいる [Reiter 78b]。この仮定によって、無矛盾なデータベースに矛盾が生じる事がある。例えば次のような簡単な例を考える。

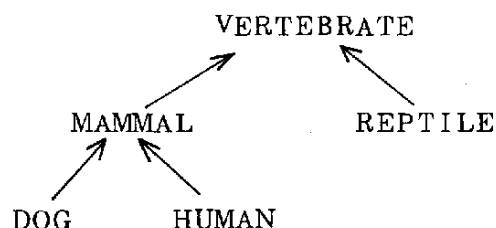
DB : $\{ P(a) \vee P(b) \}$

つまりこのデータベースには上のような1つの論理和しか含まれていない。この時、このデータベース内では $P(a)$ が正しいことは証明できない。すなわち $\text{DBVP}(a)$ である。したがって closed world database では、 $\neg P(a)$ が正しいと結論される。同様にして、 $\neg P(b)$ も結論されるが、これらの3つの論理式が1つのデータベースに存在するのは矛盾である。

Reiter はこれについて、もしデータベースに含まれる論理式がすべて Horn 節（高々1つの正項しか含まない節）ならば、上記のようなことが起こらない事を示している。

5.3 階層表現における default

ここでは、意味ネットワークにおける階層表現について考える。例えば、次のような階層表現があるとする。



この意味ネットワークは述語表現すると、次のような事実を表現している。

$$\left\{ \begin{array}{l} (\forall x) \{ \text{DOG}(x) \supset \text{MAMMAL}(x) \} \\ (\forall x) \{ \text{HUMAN}(x) \supset \text{MAMMAL}(x) \} \\ (\forall x) \{ \text{MAMMAL}(x) \supset \text{VERTEBRATE}(x) \} \\ (\forall x) \{ \text{REPTILE}(x) \supset \text{VERTEBRATE}(x) \} \end{array} \right.$$

意味ネットワークにおける I S - A 関係の推移律を述語論理における推移則 (modus ponens) と同様に見るならば、I S - A (DOG, MAMMAL) すなわち、犬が哺乳類である事を証明する過程はほとんど同じである。しかし、DOG の上位概念に REPTILE が存在しない事から、犬が爬虫類でない事を結論する意味ネットワークの推論方法は述語論理には存在しない。一階述語にこの能力を持たせる為には例えば、

$$(\forall x) \{ \text{DOG}(x) \supset \neg \text{REPTILE}(x) \}$$

のような公理を入れてやらなければならない。

意味ネットワークが備えているこのような推論規則を一階述語に持たせるにはどうすればよいだろうか。意味ネットワークが一階述語を用いて表現できる事を考えると、何らかの推論規則を述語論理に付け加えてやればよいという事がわかる。前節で述べた closed world database の考え方を採用するのも一つの解法であるが、意味ネットワーク自身が持っている default 能力をすべて述語論理上で表現するには、さらに他の規則が必要である。こうした論理系への試みは、5.6 節で示される。

5.4 例外表現と default

我々が一般に用いている知識には例外が含まれている事が多い。実際、ほとんどの一般的知識は例外を含んでいると言ってよい。(例外もあるが。) 例えば、「鳥は空を飛ぶ。」という言明はやはり例外を含んでいる。(ペンギン、

ダチョウなど)すなわち、鳥が飛ぶというのは確固たる事実である訳ではなく、このことすら default として扱う方が適切であると考えられる。これに対して、次のような述語によって、例外を明示する方法も考えられる。

$$(\forall x) \{ \text{BIRD}(x) \wedge \neg \text{PENGUIN}(x) \wedge \neg \text{OSTRICH}(x) \supset \text{FLY}(x) \}$$

しかし、この知識からだけだと、例えば、 $(\forall x) \{ \text{TORI}(x) \supset \text{BIRD}(x) \}$ という知識しかない TORI という鳥が飛べるかどうかは、一階述語によっては、何の結論も得られない。すなわち、TORI がペンギンであるかどうか、また、ダチョウであるかどうかの判定ができないからである。

上で述べた、「鳥は飛ぶ。」という知識自身を default として扱う方法について考えてみよう。これはつまりある動物が鳥である事がわかっている場合に、それが飛べないという知識が欠如している限りは、飛べるものと思っておこうという考え方である。

5.5 Frame Problem*

default reasoning が最も重要な働きをする問題に Frame Problem [McCarthy 69] [Raphael 71] [Hayes 73] がある。例えば、箱を移動させた時に、この箱の色や重さは変化するであろうか。また、この箱に色を塗った時に、箱の位置や形は変化するであろうか。箱を動かす時に、その箱の上に乗っている物体も同時に移動する事をどのように表現すればよいだろうか。Frame Problem は次のように述べる事ができる。

「ある状態変化により不変なもの、そうでないものをどのように表現すればよいか。」

不変なものを列挙する事によって、この問題を解決するのは非常に困難であろう。McCarthy らは、箱に色を塗っても位置は変化しないという事を次のように表わしている。 $(x, z/\text{OBJECT})(y/\text{POSITION})(s/\text{STATE})(c/\text{COLOR})$

$$\text{LOCATION}(x, y, s) \supset \text{LOCATION}(x, y, \text{paint}(z, c, s))$$

一般には、このような表現が数知れず必要になり、とても扱いきれるものでは

* Minsky のフレーム理論とは別の問題である。

ない。

Sandewall は Frame Problem を扱うために, UNLESS というオペレータを一階述語に導入した。^{*} UNLESS(P) が真になるのは, Pが公理から, 一般述語の演繹体系を用いて導びき得ない時である。例えば $A \wedge \text{UNLESS}(B) \supset C$ は「Aが成り立ち, Bが一階述語論理を用いて公理から導びくことができないならば, Cが成り立つ。」と読むことができる。彼は restricted wff (rwff) という, 条件付きの論理式を定義して, default reasoning を行なおうとしたが, 彼の提案した推論規則は満足なものではなかった。しかし, 次節で述べる default reasoning は主にこの考え方を基にしており, こうして発展した default reasoning の理論は, 前節までに述べた種々の default の形態をすべて包含するものである。

5.6 Default Reasoning

ここでは, 主に default reasoning に関する Reiter 及び McDermott らの仕事について説明する。[Reiter 79] [McDermott 78].5-2 節で述べた closed world assumption (CWA) は前節の UNLESS記号を用いて $\text{UNLESS}(P) \supset \neg P$ のように書ける。これを以後 $M \neg P \supset \neg P$ と書く事にする。

5.1 節における K R L の例は少し複雑であるが次のように書ける。

$$\begin{aligned} & \text{PERSON}(x) \wedge M(\exists y) \{ \text{hometown}(x)=y \} \\ & \supset \text{hometown}(x)=\text{Palo Alto}. \end{aligned}$$

5.4 節の例は次のように書けばよい。

$$\text{BIRD}(x) \wedge M \text{ FLY}(x) \supset \text{FLY}(x)$$

CWAについて考えてみよう。今, 論理データベースに P という述語が存在し, Pについて, $P(a) \leftarrow, P(b) \leftarrow$ の2つの節しか定義されていないとする。

この時, CWAは次の式

$$(\forall x) \{ P(x) \leftrightarrow x=a \vee x=b \}$$

が成り立つ事を要求するのに等しい。さらに, データベースに新しい節 $P(f(y)) \leftarrow P(y)$ が含まれ, しかもこれら3つの節以外に Pに関する節が含まれていな

* [Sandewall 72]

いと仮定すると、CWAは次の式を要求する。

$$(\forall x) \{ P(x) \leftrightarrow x=a \vee x=b \vee (\exists y) [x=f(y) \wedge P(y)] \}$$

Clark はこうした変形規則を用い、証明の失敗により、述語の否定形を証明するアルゴリズムを示している。〔Clark 78〕 上のような変形を彼はデータベースの完全化と呼んでいる。彼のアルゴリズムに忠実に従う事は、導出原理における unification が導出木のすべての枝で失敗に終るのを確める事に等しいので、特に質問に変数が含まれている場合には困難が生じる。すなわち、ここで注意しなければならないのは、このようなアルゴリズムは一般に部分的可能ですらないことである。導出木の unification がすべて失敗に終るかどうかを調べる事は、述語論理式の充足可能性を調べる事に等しく、充足可能な一階述語論理式の集合は帰納的可算ですらないからである。Clark は彼のアルゴリズムが質問のすべての解を見つける事ができるために、データベース内の節や質問が満たさなければならない条件についても述べている。

さて、一階述語論理は単調である。すなわち、新しい公理を付け加えたからといっても、公理系が無矛盾な限り、以前定理であったものが、定理でなくなる事はあり得ない。ところが、今まで議論した default 規則を伴う論理系は非単調である。すなわち、新しい公理の追加が系の定理の集合の単調非減少を保証しない。また、次のような2つの default 規則を持つ系を考える。

$$MA \supset \neg B, \quad MB \supset \neg A$$

$\neg A$ はこの系では導びく事はできないから、前者より $\neg B$ が導びかれる。ところが逆に、はじめに $\neg B$ の導出を試みると、その失敗により、後者から $\neg A$ が導びかれる事になる。すなわち A も B も、この系からは証明できるかどうかの判定がこのままではできない。さらに、5.2節の例のように矛盾を含む系すら存在しうる。

このように default 規則を含む一般的な default reasoning のための論理について議論するには、与えられた公理と default 規則によって閉じた論理式の集合を扱うのが妥当である。つまり、ある論理式の集合が存在して、それが与えられた公理を含み、一般述語の公理系及び default 規則によって閉じている時、それを1つの無矛盾な世界と見る訳である。

先に述べたことからわかるように、このような集合は一意に決まるとは限

らないし、存在すら保証されていない。実際に、 $\{MA \supset \neg B, MB \supset \neg A\}$ のみをもつ公理系には、2つの可能な世界、 $Th(\{\neg A\})$ 、 $Th(\{\neg B\})$ が存在する。ここに $Th(\{P\})$ は論理式 P と一階述語の公理系によって導びかれるすべての真なる論理式の集合となる。また、 $\{MA \supset \neg A\}$ には無矛盾な閉じた集合は存在しない。

Kramosil は、このような公理系の定義によって得られる集合は無意味であるか、役に立たないのどちらかだと言っているが [Kramosil 75]、これは公理系に単調性を強要したためである。言い方を改めると、彼は、default 規則を持つ公理系が単調ならばそれは default 規則を持たない公理系と同等であることを証明した。

さて、このような default 規則を持つ論理の具体的な定式化について述べよう。

5.6.1 Non Monotonic Logic (McDevmott 78)

オペレータ M を前に述べたように定義し、 A を M を含む一階述語のある公理の集合とする。次のようなオペレータ NM_A を定義する。

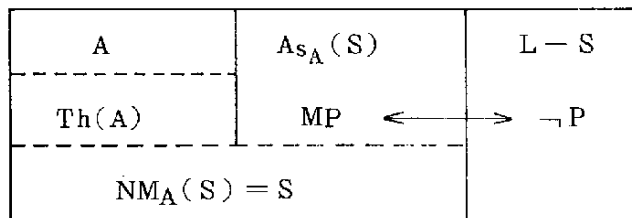
$$NM_A(S) \triangleq Th(A \cup As_A(S))$$

ここに S は wff の集合 L の任意の部分集合、

$$As_A(S) = \{Mq : q \in L, \neg q \notin S\} - Th(A)$$

によって As_A は定義される。

このオペレータによって $NM_A(S) = S$ となる論理式の集合 S が前に述べた閉じた集合である。 L を全体集合として、これらの関係を、 $NM_A(S) = S$ となる集合 S について図に示したのが下図である。



NM_A に関する不動点を $FP(A)$ で表わす。すなわち、 $FP(A) \triangleq \{S : NM_A(S) = S\}$ 。これによって彼らは、 A から証明可能な論理式の集合

TH(A) を次のように定義している。

$$TH(A) \triangleq \cap [\{L\} \cup \{S : NM_A(S) = S\}]$$

これは、 NM_A の不動点が存在する時は $TH(A)$ はそれらの共通部分、そうでない時には $TH(A)$ が L と一致することを意味する。

これらの定義より、もし、ある論理式 P が A から証明可能、すなわち $P \in TH(A)$ ならば P は NM_A のあらゆる不動点から導出可能であり、また逆も正しい事が証明される。

問題は $P \in TH(A)$ の判定法であるが、彼らは命題論理の上での完全なアルゴリズムを示しているが、このようなアルゴリズムの存在は命題論理の有限性の性質により明らかである。一階述語の上での $P \in TH(A)$ の判定はやはり、部分的可能ではないであろう。

5.6.2 Logic for Default Reasoning (Reiter 79)

Reiter の default theory は前節の McDermott の理論の若干定義が異なる。最大の相異点は、 M オペレータを併う default 規則を一階述語式としてではなく、演繹規則として扱う事である。default theory は次のような 2 組で定義される。(D, W)。ここに W は wff の集合、 D は default 規則の集合。default 規則は一般に次のような形で与えられる。

$$\frac{\alpha(\vec{x}) : M\beta_1(\vec{x}), \dots, M\beta_m(\vec{x})}{W(\vec{x})}$$

これは $\alpha(\vec{x})$ が成り立つという前提のもとで $\beta_1(\vec{x}), \dots, \beta_m(\vec{x})$ の否定が証明されないなら $W(\vec{x})$ が正しいとせよ、という意味である。

任意の wff の集合 S について $\Gamma(S)$ を、次の 3 つの性質を満たす最小の集合とする。

ある default theory (D, W) について、

$$D1. \quad W \subseteq \Gamma(S)$$

$$D2. \quad Th(\Gamma(S)) = \Gamma(S)$$

$$D3. \quad \frac{\alpha : M\beta_1, \dots, M\beta_m}{W} \in D$$

$$\text{かつ } \neg\beta_1, \dots, \neg\beta_m \notin S \text{ ならば } W \in \Gamma(S)$$

このオペレータ Γ を前節の NM_A と同様に考え、 Γ による不動点を “可能な世界” と考える。 $\Gamma(E) = E$ なる集合を (D, W) の extension と呼ぶ。

ここで、default theory と non monotonic logic の相異点を示す。
例えば次の default theory

$$\left(\left\{ \frac{A : MB}{B}, \frac{C : MD}{D} \right\}, \{ A \vee C \} \right)$$

は non monotonic logic では $\{ A \wedge MB \supset B, C \wedge MD \supset D, A \vee C \}$ と表現され、前者の不動点は $\text{Th}(\{ A \vee C \})$ 、後者は $\text{Th}(\{ A \vee C, B \vee D \})$ となる。

また、次のような default theory

$$D = \left\{ \frac{\alpha_1 : MW_1}{W_1}, \frac{\alpha_2 : MW_2}{W_2}, \frac{\alpha_3 : MW_3}{W_3} \right\}$$

$$W = \{ W_1 \supset \alpha_2, W_2 \supset \alpha_3, W_3 \supset \alpha_1, \\ W_1 \supset \neg W_2, W_2 \supset \neg W_3, W_3 \supset \neg W_1 \}$$

は唯一の extension として $\text{Th}(W)$ を取るが、対応する non-monotonic logic

$$WU \{ \alpha_1 \wedge MW_1 \supset W_1, \alpha_2 \wedge MW_2 \supset W_2, \alpha_3 \wedge MW_3 \supset W_3,$$

は NM_A に関する不動点を持たない。

default 規則の結論と M オペレータのついた論理式が等しいものを normal default 規則という。例えば、

$$\frac{\text{BIRD}(x) : \text{M FLY}(x)}{\text{FLY}(x)}$$

は normal default 規則である。normal default 規則だけからなる default theory を normal default theory とぶ。5 節までに述べた例の多くは normal default 規則のみを用いて表現する事ができる。

重要な点は normal default theory が少なくとも 1 つの extension を持つ事である。

non-monotonic logic ではすべての不動点に含まれる論理式を“証明可能”な論理式とし、 $P \in \text{TH}(A)$ について考察したが、default theory では、論理式 P に対して、 P を含むような extension があるかどうかを考察している。

彼は normal として一般の default theory の種々の性質および、証明手順

について述べているが、何度も繰り返すように、これは、部分的可解でもなく、いわば、証明の必要十分条件を述べたに過ぎない。

5.6.3 Truth Maintenance System [Doyle 78]

non monotonic logic を基にした証明システムに Truth Maintenance System (TMS) がある。

TMS は、個々の証明手順によって生成された定理の証明を保存し、 $\neg P$ と MP が同時に真にならないように管理する。また、矛盾を発見すれば、それを回避するために新しい公理を付け加えたりする。各式は justification という条件式の集合をもつ。P が 1 とラベル付けされるのは、ある justification 内の式すべてに 1 がラベル付けされた時に限る。MP は $\neg P$ が 0 の時、かつその時に限り 1 となる。

簡単な例を示す。

N-1 $\text{TIME}(T) = 1000$ (SL()(N-2))

N-2 $\text{TIME}(T) \neq 1000$

N-3 $\text{ROOM}(T) = 813$ (SL()(N-4))

N-4 $\text{ROOM}(T) = 801$

SL の第 2 番目の括弧内が否定的な justification, 第 1 番目が肯定的な justification を示す。したがって 1 行目は $M \rightarrow (\text{TIME} \neq 1000) \supset \text{TIME}(T) = 1000$ を示す。

TMS は 1 というラベル付けされた論理式だけが、次の証明段階に“見える”ようにする。このような論理式は in であり、そうでないものは out であるという。上の例では in である論理式にのみ justification が示されている。

TMS は non monotonic logic とは違って、1 つの不動点を捜すシステムである。例えば、前に挙げた次の例 $\{MC \supset \neg D, MD \supset \neg C\}$ について考えると、

N-1 $\neg D$ (SL()(N-2))

N-2 $\neg C$ (SL()(N-1))

と表現することができ、これが同時に in となる事はない。

5.7 おわりに

不完全な知識のもとで行なう推論を、主に述語論理の立場から述べた。具体的な言語やシステムについては余り述べなかったが、PLANNER なども否定として THNOT というオペレータを持ち、これは本章で述べた CWA と同様の扱いをしている。

述語論理を用いたこのようなアプローチは本章で何度も指摘したように本質的に部分的可解でない事は止むを得ず、今後は、可解な部分の解析、及び手続き型手法との融合等を検討する必要がある。

[References]

- [McCarthy 69] McCarthy, J. and Hayes, P. J., "Some Philosophical Problems from the Standpoint of Artificial Intelligence," in 'Machine Intelligence 4' Meltzer, B. and Michie, D. (eds.), Edinburgh University Press, pp.463-502, 1969.
- [Raphael 71] Raphael, B., "The Frame Problem in Problem Solving System," in 'Artificial Intelligence and Heuristic Programming' N. U. Findler and B. Meltzer (eds.), Edinburgh University Press, pp.159-169, 1971.
- [Sandewall 72] Sandewall, E., "An Approach to the Frame Problem and its Implementation," in 'Machine Intelligence' Ellis Horwood, pp.195-204, 1972.
- [Hayes 73] Hayes, P. J., "The Frame Problem and Related Problems in Artificial Intelligence," in 'Artificial and Human Thinking' Elithorn, A. and Jones, D. (eds.), Jossey-Bass Inc., pp.45-59, 1973.
- [Kramosil 75] Kramosil, I., "A Note on Deduction Rules With Negative Premises," 4th IJCAI, pp.53-56, 1975.
- [Moore 75] Moore, R. C., "Reasoning from Incomplete Knowledge in a Procedural Deduction System," MIT AI-TR-374, 1975.
- [Lipski 77] Lipski, W. Jr., "On the Logic of Incomplete Information," in Symposium on Mathematical Foundation of Computer Science, pp.374-381, 1977.
- [McDermott 78] McDermott, D. and Doyle, J., "Non-Monotonic Logic," MIT AI Memo 486a, 1978.
- [Doyle 78] Doyle, J., "Truth Maintenance System for Problem Solving," MIT AI-TR-419, 1978.
- [Clark 78] Clark, K., "Negation as Failure," in 'Logic and Data Base' Plenum Press, pp.293-322, 1978.
- [Reiter 78a] Reiter, R., "On Reasoning by Default," 2nd Symposium on Theoretical Issues in Natural Language Processing, pp.25-27, July, 1978.
- [Reiter 78b] Reiter, R., "On Closed World Data Base," in 'Logic and Data Base' Plenum Press, pp.55-76, 1978.
- [Reiter 79] Reiter, R., "A Logic for Default Reasoning," British Colombia TR-79-8, 1979.
- [Doyle 79] Doyle, J., "A Glimpse of Truth Maintenance," 6th IJCAI, pp.232-237, 1979.

6. 談 話 理 解

6.1. はじめに

6.1.1 意味関係による談話内文結合現象^{*1)}の概念

人間はいくつかの文を見たり聞いたりしたとき、それがひとまとまりのハナシを構成しているか、それとも単に脈絡のない文を集めたものにすぎないかを難なく区別することができる。現実には何ら脈絡のない文が次々に出てくるとはあり得ず、つながりを持って展開されるのがふつうであるが、その中で人間は、他とは独立してそれだけで閉じて内容の理解できるハナシのまとまりを識別している。このようなまとまりを、以下では談話ユニット^{*2)}とよぶ。談話ユニットには所定の長さは決まっておらず、短かいもの(例「禁煙」)から長いもの(例、一日中続いた議論)まで、広い範囲を持っている。

人間が、談話ユニットとそうでないものとを容易に区別できるのは、談話ユニットだけが持っていて他のものにはない何らかの特徴があるためと考えられる。これを客観的要因として識別すること、すなわち、人間の直観を支えているメカニズムを客観的対象として取出すことは、談話理解にとって重要である。

文は形式的なユニットであって、文の部分と文に統合する構造的な関係が存在する。しかし談話ユニットと文の間には文法構造の関係は存在せず、文を集めて談話を作りあげる構成規則というようなものがあるわけではない。談話ユニットは、見かけ上は文の集まりを成しているが、それは文により現実の発話や書かれた文章として実在化されており人間にわかるようにコード化されているにはすぎない。談話ユニットは、形式的なユニットではなく、意味のユニット(Semantic Unit)であり、文法構造的統合力とは別の力の働きでまとまりを成していると考えられる。

談話ユニットとそうでないものを区別しているもの、談話ユニットだけが持っている性質を、意味関係による結合性、あるいは意味結合性とよぶ。また、その性質が言語現象に表われて、2つ以上の文を結合しているとき、意味関係による談話内文結合現象、あるいは単に結合現象とよぶ。

このような現象のひとつに、照応指示(現象)が挙げられる。

*1 言語学では cohesion とよばれる。

*2 言語学では text とよばれることもある。

例. a. Wash and core six cooking apples.

b. Put them into a fireproof dish.

a. 6個の料理用のリンゴを洗い、しんを取れ。

b. それ(ら)を耐火皿に入れよ。

上の例では、"them" (それらを)が、"six cooking apples" (6個の料理用りんご)を振返って指示し、この前方照応機能が a. と b. の2文間に結合現象を生じさせ、2文を独立のものではなくしているということがわかる。この2項間の結合関係は、何を意味しているか。ここで前方照応項 "them" (それらを)は "six cooking apple" (6個の料理用りんご)と同一のものを指しているという "信号" である。そしてこの2項間の結合関係の意味は、"指示の同一性"^{*1)}であり、この指示の同一という意味こそが2文に意味結合力を与えたのであって、apple(りんご)ということば自体、あるいは them (それらを)ということば自体には何ら結合力はない。すなわち、意味結合は、適当な性質を満たす2項間に生じる現象であり、各項にあてはめられたことば自体は力を持たない。

上の例は、代名詞による同一物の指示という関係であるが、意味結合を与える関係はこれ以外にもいろいろあるし、代名詞が常に指示の同一性を表示する手段というわけではない。

意味結合関係にある一組の項同志を、意味結合帯^{*2)}とよぶ。前出の帯は、指示^{*3)}という帯である。具体的な言語データにあたって意味結合帯を分別し解析することは、意味結合機構の解明の基礎となる。英語の場合は、Haliday 等により、完全ではないがかなり詳しい分析がなされているので、本論では Haliday の考え方を中心にして報告する。日本語の場合は、統合的研究成果は見当たらないが、個々の事例研究が没興している最中であるので、いずれこの分野における総合的視野が開かれると思われる。

今まで述べて来たように、談話ユニット内に存在する意味関係が言語現象に表出している手がかりとして、談話の中にある要素の解釈^{*4)}がそれ自体では完

*1 coreferentiality

*2 cohesive tie

*3 reference

*4 interpretation

成せず、何か他の別の要素等の解釈に依存してはじめて完成する場合を対象として来た。

前の例で、"them" (それらを)は、解釈に際し、別の要素の存在を暗黙前提^{*1)}している。この前提の持つ要素は前出の文中の"six……"で満たされる。このように、自発的な前提項^{*2)}と、その被前提項^{*3)}の2つの要素の間に結合力が存在するが、この2つの要素は必ずしもこの順番で談話の中に表われるわけではない。基準となる最も単純なパターンは、前出例のような被前提項がはっきりとことばで示されていて、前提項のごく近くの前方に存在する場合であり、実際によく観察されるパターンでもある。(ごく近くということは、同一の文中ということではない。意味結合関係は、文の切れ目とは無関係に生じるものであるから、分析用データとしてはむしろ、被前提項と前提項が別の文の中にあって、両者が' 'や': 'で区切られている場合の方がより問題が明確になって、適当かもしれない。)

単純な前方照応(現象)^{*4)}では、被前提項となる単語が前提項のごく近くの前方に存在するが、かなり離れた所にある場合もある。また、単語のように短かい項ではなく、いくつもの文から成る文節をひとまとめにして指すこともある。1つの被前提項を何回もくり返して指したり、逆に、いくつもの被前提項を要約して1つの前提項で指して済ます場合もある。

このように、前方照応形の中にもいろいろのパターンがある。

前方照応(現象)とは逆に、被前提項が前提項の後に出てくる場合もある。英語では、this や here はしばしば後方を参照する。

例 This is how to get the best result.

.....やり方.....

.....

このような現象を後方照応(現象)^{*5)}とよぶ。

*1 presuppose

*2 presupposing item

*3 presupposed item

*4 ANAPHORA

ANAはupを表わすギリシヤ語からの借用の造語要素

PHORAは指示性を表わす

*5 CATAPHORA

CATAはdown, backの意味を表わすギリシヤ語から借用の造語要素

コロン(：)印はしばしば後方指示の信号である。

前方照応にせよ、後方照応にせよ、被前提項は談話の中に存在する。そこでこの2つを合せて、談話内照応現象^{*1)}とよぶ。

それに対して、談話中のある要素の解釈に必要な情報が談話内には全く見付からず、談話の周辺環境を見てはじめてわかる場合がある。談話の周辺環境は大変広く莫然としているが、問題となるのは談話のことばの選択に影響を与えた外部要因だけが対象であることはいうまでもない。このような場合を、談話外照応(現象)、あるいは、談話環境照応(現象)とよぶ。これは、文結合には関係しない。

以上のような照応現象は、英語にも独語にも日本語にも見出され、おそらくは他の言語においても通用する普遍性のある言語現象であると思われる。従って、この現象の背後にあるメカニズムは、言語に依存しない普遍的な存在と思われる。たとえば、記憶モデルの階層構造との関係、知識構造との関係は普遍性を探る立場の問題である。

6.1.2 談話理解における、意味関係による結合現象研究の意義

近年、言語学の分野でも談話研究の必要性が認識されてきているが、ここでも研究の手がかりとして照応現象のような、文の結合現象が注目されてきている。これは、このような現象の解明が、言語学的に有意義な一般性をとらえるのに重要な視点を提供するもののひとつであるからであって、言語学者にとって関心のある談話文法のある一面が照応現象の解明の研究の進展と共に明らかになることが指摘されている。

このように、談話における意味関係による結合現象の解明が、一部の特殊な興味にとどまらず、一般性のある知識体系の一部を明らかにするための有効な手がかりであることは、言語学においてのみならず、以下に述べるように、談話理解を含む認知科学および情報処理の分野でも同様である。

認知科学の分野では、談話にみられる意味結合現象は、知識表現(の体系)および世界理解モデルの構造の一部が言語現象に表出したものとしてとらえら

* 1 ENDOPHORA

* 2 EXOPHORA

れ、このような現象の解明は、それらの構造の一部を解明することにつながる
と考えられる。このような解明は、計算機による談話理解システムの合成プロ
セスと相互に作用しつつ進展するものである。

一般的に、意味関係に限らずとも、言語現象は人間に内在する、認識のメカ
ニズム、知識構造の投影、そして理解プロセスの軌跡と見なされるが、これら
を単発的な事象としてではなく、体系的モデルとして把握したときはじめてそ
れらに対する明るい視界が得られる。そして、計算機による談話理解システム
の作成はこのような把握を要求する。したがって、言語現象の事例分類におい
ても、体系的モデルとの関係を見渡せる視座にあることを必要とする。

現実のデータに密接した分析は不可欠であり、各々の現象を網羅的に追うこ
とも大切であるが、それにとどまらず、一般性のある知識表現および理解のモ
デルを形成し、それに体系的に組込んで行けるかどうか、これからの情報処
理にとって興味ある問題である。

談話理解システムは、先に述べたように、このようなモデル形成を促進する
と同時に、他では得られぬモデルの正当性に対する検証の手段を与えるもので
もある。モデルの適切性は、談話理解システムの知的能力の土台になるもので
あるからである。各分野の研究の成果 — 言語学的解明、自然言語の意味表現、
理解のモデル等について得られた知識 — の確かさに対する総合的な評価は、
計算機による談話理解システムを構成し、動作させてモデル検証するプロセス
を経ることによって客観的なものとすることができる。

このような立場で言語現象を見渡すとき、その中で、目的に近づくのに効果
的な現象に的を絞ることの重要性が浮かび上がってくるわけであるが、意味関係
による結合現象はそのような目的に沿う良性の素材といえよう。

たとえば、照応現象で、被前提項を同定する問題は、自然言語の形式的な意
味表現を追究するとき、その表現の適切性を検証するための手段とみなされ、
研究されている。すなわち、理論的に設定された形式的意味表現法が、照応の
問題に耐えられるかどうかという耐性テストを提供しているわけであるが、こ
れと同様のことが談話理解システムについても成立つ。すなわち、意味関係を
正しく理解し、照応などにも対処できるかどうかということは、談話理解シ
ステムに対する一種の耐性テストである。もしそのシステムが適切な知能を持っ

ているならば、そのような用法にも対処できる筈だからである。

そしてこのような機能の達成のためには、どのような知識構造を備えねばならないかが、一部分ではあるが、我々のグループにより解明されている。

6.1.3 指示現象研究の位置付

これまでに、意味関係による談話内文結合現象の概念と、談話理解における意味関係による文結合現象研究の意義について述べた。

この問題は、言語に依存しない普遍的な知識を意識するものであるが、研究のアプローチが現語現象を介するものである以上、特定の言語を扱わないでは済まされない。そして我々の関心の中心が日本語にあることは当然である。しかし、日本語については、この種の課題についてのまとまった報告は見あたらない。一方英語については、Halliday等により意味結合関係の研究が言語学的視点からまとまった成果を与えられているので、今回は、Hallidayの考えを中心に報告する。

英語に関する整理は、今後の日本語に関する整理に有効に利用できるものと思われる。

意味関係による文結合における重要な現象としては、第1に照応指示現象（以下単に指示現象とよぶこともある。）が挙げられる。その他、代入現象、省略現象による文結合、接続詞による文結合や同一語の繰返しなどによる語彙的意味による文結合などがある。

ここでは、これらのすべてを紹介するには紙数が限られているので、この中で特に重要であり、他の現象の解明の基礎となる視点を与える指示現象のみを集中的に扱い、最後に指示と代入・省略との相異にごく簡単に触れるにとどめた。

Hallidayの考えをなるべく忠実に紹介すると共に、我々の研究の関心との接点もいくつか指摘した。

6.2 指示現象

6.2.1 指示現象における照応性

どの言語にも、そのことば自体は具体的に何を指しているかはっきりしないが、文脈から解釈できるという項 (item) が存在する。たとえば、日本語では「それ」「彼」などである。これらは、談話内容あるいは周辺環境からその意味が聞き手^{*2)}に理解できることが明らかなので、話し手^{*3)}があえて直接には示さずにすませたものであり、談話内容の連続的であることが前提になっている。

これは、談話内の文結合現象の中で最も代表的な現象であり、指示 (現象) とよばれる。指示する項は、検索されるべき情報が周辺のどこかにあるという信号であり、その情報を使えば指示物^{*4)}がわかる (解釈できる) ことを前提に用いられる。

指示現象は、検索されるべき情報のあり場所 (談話内にあるかそれとも談話環境にあるか。また、談話内であれば、指示項より前に既出であるか、指示項の後になってあらわれるか) によって分類される。(図 6-1 参照)

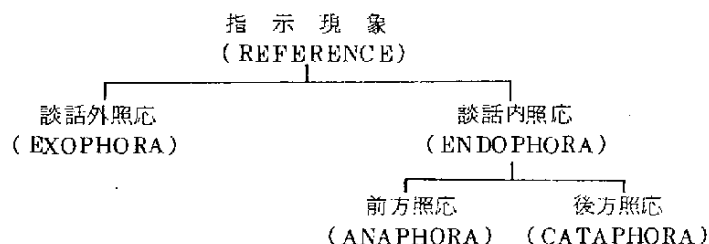


図 6-1 照応の分類

この中で、文間の意味結合に関係があるのは、談話内照応現象の方であり、指示する方の前提項に対応して、検索されるべき情報としての被前提項が談話内に存在する場合である。

談話外照応現象は、談話中のある要素の解釈に必要な情報が談話内には全く

*1) reference

*2) 以下特に述べなくても読み手も含むものとする。

*3) 以下特に述べなくても書き手も含むものとする。

*4) referent

見つからず談話の周辺環境を見てはじめてわかる場合である。実際の話しことばでは頻繁に用いられ、話し手と聞き手が親密な関係にあるほど、すなわち、両者の知識の共通部分の大きいほど、会話に「あれ」「それ」などが多くなる。^{*1)}

談話外照応現象は、文間意味結合には直接関係しない。それにもかかわらずこれも検討する価値があるのは、状況の指示の方が、談話にことばとして出て来たものを参照して解釈を行う指示よりも、原型だからである。すなわち、「あなたの目の前にあるもの」を指すことの方が「私がさっき言及したもの」を指すよりも、歴史的にみて先に会話にあらわれたことであろう。

有視界内にあるものを指すことから出発している指示現象は、意味の限定性または特定性を示すことがひとつの特徴である。

例 Three blind mice, three blind mice.

See how they run /

上の例で、“they”は、“three blind mice”を指すが、それは単なる three blind mice ではなく“the same three blind mice that we just been talking about.”という限定性がある。

この指示の限定性は、有視界にあるものでなくても適用される。

例 { a. Buy five red roses.
b. Give them to your mother.

上の例で、前方照応項“them”は、“five red roses”と同一のものを指しているという「信号」である。これら2項間の結合関係の意味は、「指示の同一性」であり、同一物を指しているという意味こそが2文間に意味結合を与えたのであって、“them”ということば自体には何ら結合性はないことは、既に6.1にて指摘した通りである。

「物事」の同一性は、指示現象の基礎となるものであるが、これが拡張され、「物事の類(class)」の同一という意味が2文間に結合性を与えている場合もある。

例 There's a problem with pollution even in this beautiful country. It seems that nothing can be done about this.

*1) 子供同士や夫婦の会話には、談話外照応が多いといわれている。

上の例で、this は、" this particular pollution"ではなくて、
" pollution in general"である。

このように、広義の'もの'、または'もの'の類の一致という意味のつながりさえあれば、文章構造の中で果たす機能は問題にならないことは、指示現象のひとつの特徴である。^{*1)}

ところで「太郎」とか「船」とか「走る」とかの語彙項目も指示する意味は持っている。これらは、個体としてのものそのものや、ものの類やプロセスに名付けする。それに対して、指示現象で扱う代名詞などの指示項は、指示物が文脈からわかるという信号であり、何も名付けない。この信号は意味を解釈するための信号であり、話し手が意味を回復できることを前提にして用いられている。すなわち、ここで話し手が聞き手に前提しているのは、事物と名前との1対1対応の知識ではなく、信号を解釈できる知識である。

6.2.2 指示表現のタイプ

指示の文法的な表現手段には、3つのタイプがある。

第1は、人称代名詞による指示で("it"による事物の指示もこれに含まれる)、人称のカテゴリを介して会話の状況の中で果している役割を指示する。(図6-2)

第2は、指示詞による指示で、近さの度合を示す物差上の位置によって指示する。(図6-3)

第3は、比較による指示で、同一性または相似性によって間接的な指示を行う。(図6-4)

6.2.4節以下で、それぞれについて述べる。

*1) 「代入」の場合は、文章構造の中で果たす機能が同等かどうか問題となる。

意味範疇
(Semantic category)

文法的機能
(Grammatical function)

クラス
(class)

人:

話し手(単数)
(speaker only)

聞き手, 第3者は有/無
(addressee, with/without
other person(s))

聞き手と第3者(達)
(speaker and other
person(s))

第3者(男)
(other person, male)

第3者(女)
(other person, female)

第3者達;物
(other persons; object)

物;談話の部分
(object; passage of
discourse)

一般的な人
(generalized person)

存在 (Existential)		所有 (Possessive)	
主要語 (Head)			修飾語 (Modifier)
名詞・代名詞 (noun・pronoun)		限定詞 (determiner)	
I	me	mine	my
you		yours	your
we	us	ours	our
he	him	his	his
she	her	hers	her
they	them	theirs	their
it		[its]	its
one			one's

図 6-2 人称代名詞による指示

意味範疇
(Semantic category)

文法的機能
(Grammatical function)

クラス
(Class)

接近度:
(Proximity)

近 い
(near)

遠 い
(far)

中立
(neutral)

選 択 的 (selective)		非 選 択 的 (non selective)
修飾語/主要語 (Modifier/Head)	付 加 語 (Adjunct)	修 飾 語 (Modifier)
限 定 詞 (determiner)	副 詞 (adverb)	限 定 詞 (determiner)
this these	here [now]	
that those	there then	
		the

図 6 - 3 指示詞による指示

文 法 的 機 能
(Grammatical function)

ク ラ ス
(Class)

一般の比較:
(General Comparison)

同 一 性
(identity)

一般的相似性
(general similarity)

相 異
(difference)

特定の比較:
(Particular comparison)

修飾語: (Modifier)	副 修 飾 語 (Submodifier)
指示的語/性質語 (Deictic / Epithet)	または 付 加 語 (Adjunct)
形 容 詞 (adjective)	副 詞 (adverb)
same identical equal	identically
similar additional	similarly likewise so such
other different else	differently otherwise
better more などの比較級と数量詞 (comparative / quantifier)	so more less equally

図 6 - 4 比較による指示

6.2.3 名詞の役をする語群^{*1)}の構造

文法的には、指示要素には名詞句の形をとっているものが多い。これは、他の意味結合現象に比較した場合の指示現象のひとつの特徴となっている。^{*2)} この名詞の役を果たす語群の構造は、意味結合の議論ではしばしば参照されるため、ここで簡単にそれについてまとめておく。

名詞的な語群の構造に対するひとつの見方は、主要語^{*3)}と修飾語(群)^{*4)}に分けて見る形式的構造による見方である。修飾語で主要語より先に来るものを前位修飾語^{*5)}、後に来るものを後位修飾語^{*6)}とよぶ。

例 The two high stone walls along the street

前位修飾語 主要語 後位修飾語

この他の見方には、実世界の現象を表現する役割による場合があり、これによりとり出される構造を経験的構造^{*7)}とよぶ。

経験的構造を通しての見方の場合、名詞的な語群の要素には、指示的語 (DEICTIC, 省略形 D), 助数語 (NUMERATIVE, 省略形 N), 性質語 (EPITHET, 省略形 E), 種別語 (CLASSIFIER, 省略形 C), 定限語 (QUALIFIER, 省略形 Q) そして事物語 (THING, 省略形 T) がある。この見方により上記の例の構造を見ると次のようになる。

例 the two high stone walls along the street

D N E C T Q

T以外の要素は何回使っても良い。"boys and girls" は1つのTであるし、また "hot or cold tea" には2つの形容詞があるが、1つのEとみなす。D, N, E, C, T, Qという順番は重要である。

これらの要素は対応する語のクラスは、Dは限定詞^{*8)}, Nは数詞^{*9)} かまたは数

*1) nominal group.

*2) たとえば、「代入」では、動詞の代入や文節の代入も多い。

*3) Head.

*4) Modifier. 以下単に修飾語とよぶ。

*5) Pre-modifier.

*6) Post-modifier.

*7) experimental structure

*8) determiner.

*9) numeral.

量詞,*1) Eは形容詞, Tは名詞であることが多いが, 必ずしも1対1対応があるわけではない。たとえば, 形容詞はDとしてもCとしても機能することがある。

Qは, たいていランクシフトした従属節か前置詞句である。

先にのべた形式的構造によれば, 名詞的語群には必ず主要語があるが, 主要語のクラスは何でもよいし, 経験的構造のどの要素でもよいことは, 以下の例が示している。

例 { these two customers
D N T
修飾語 主要語

{ these two
D N
修飾語 主要語

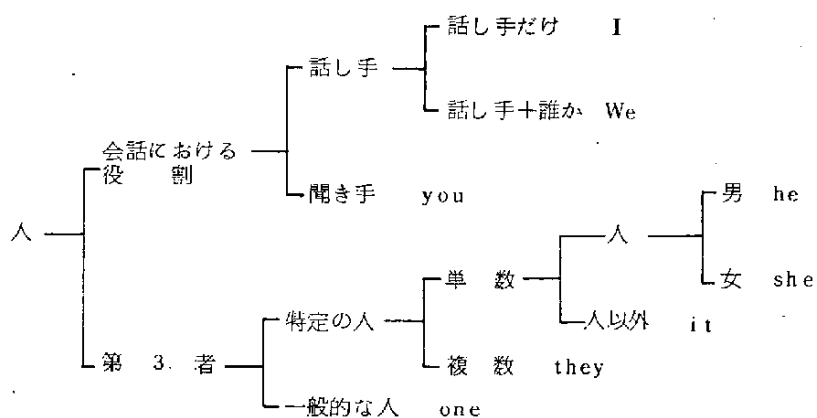
{ these
D
主要語

6.2.4 人称代名詞による指示

人称代名詞による指示は, しばしば前方照応現象を生じることが我々は経験的に知っている。しかし人称代名詞の存在が必ずしも結合帯を示しているわけではないし, 後方照応になっている場合もある。

指示の問題に入る前に, まず人称代名詞のカテゴリと人称代名詞のシステムをまとめておく。(図6-5参照)

*1) quantifier



	話 し 手	聞 き 手	特 定		一般的な人
			人	人 以 外	
単 数	I me mine my	you you yours your	he him his his	it it (its) its	one one - one's
			she her hers her		
複 数	we us ours our		they them theirs their		

図 6 - 5 人称のシステム

これらの指示要素は、会話の状況における役割を表わしている。

人称代名詞のシステムの中に、事物も含まれるのは、一見不適当のようであるが、中心的な意味が人称であるのでそれに代表させている。

人称代名詞を意味によって分類すると、交信プロセスにおける役割に従って、話し手^{*1)}、聞き手^{*2)}、第三者^{*3)}に分類される。(伝統的カテゴリに従えば、1人称、2人称、3人称。)

また、あるプロセスの関係者であるか、プロセスを受ける物の方の関係者であるかという分類もできる。1人称代名詞の例について見てみよう。

例	文	何の関係者	クラス	機能
(a)	I had a cat	プロセス	代名詞	主要語
(b)	the cat pleased me	プロセス	代名詞	主要語
(c)	tack mine	物	指示詞	主要語
(d)	my plate's empty	物	指示詞	修飾語

また、"we" は1人称複数といっても、単に"I"の集まりであるわけではない。次の3つの場合がある。

I + you

I + others

I + you + others

このように、見かけ上同一の代名詞も、文脈により、いくつかの意味に使われる。

人称代名詞の多くは指示要素として使われているが、その中で文の意味結合に関係しているものは、ほとんどが3人称の場合であって、1人称、2人称は談話外照応である場合がふつうである。但し、直接話法が含まれていると、1人称2人称でも、前照応指示がある。

例 There was a briefnote from Susan.

She just said, "I am not coming home this week end."

*1) speaker

*2) addressee

*3) The other role

3人称形で、談話中に指示物が見あたらないのは、談話が不完全であるか、談話外照応であるからである。そのとき指示物は、談話環境の中にある。それは必ずしも有視界内にあるとは限らない。話し手と聞き手が、共に同定できるものならば、3人称形で指示される。

すなわち、話し手と聞き手とが共有している知識の中にあるものは、代名詞で指示できる。視界に認識できるものは、知識の一部にすぎないわけである。

次に、特殊な例として、拡張された指示とテキスト指示について扱う。

“it”は他の代名詞と異なり、人や事物ばかりでなく、テキストの一部分を指示することができる。

The Queen said.

Curtsey while youre thinking what to say.

A
It saves time. Alice wondered a little time at this, but she was too much in awe of the Queen to disbelieve it. B

Aの指示は、“thing”を指示しているけれども、実在の人や物を指示しているのではなく、プロセス全体を指示していることがわかる。このように、1つのプロセスやいくつかのプロセスの連の全体、複雑な現象の全体を指示する拡張的な意味を持つ場合を、拡張された指示^{*1)}とよぶ。

文法的には、節または節のつながりの形をとる。

Aの指示はふつうの指示の拡張とみなされるのに対し、Bの指示は“thing”というよりは、概念的方向に変形されているといえる。これをテキスト指示^{*2)}とよぶ。

“it”以外にも、指示の“this”や“that”にはこれらの用法がある。このようにレベルの高い場合に聞き手が、談話の中でどこが指示に関係する部分なのかを理解するメカニズムは、未解明の分野である。

*1) extended reference

*2) text reference

6.2.5 指示詞による指示*1)

図6-6に指示詞による指示システムを示す。

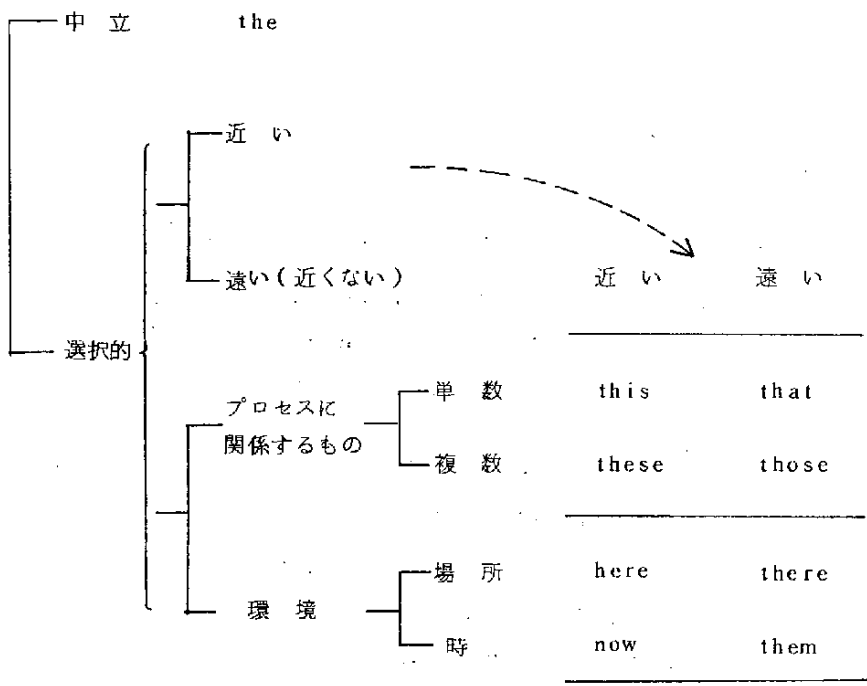


図6-6 指示詞による指示システム

*1) demonstrative Reference

指示詞による指示も談話に頻発するものである。この場合も、談話外照応があり、それらは文の意味結合には関係しないけれども、談話内照応の説明の基本になるので重要である。このことは、“the” についてもいえることである。

以下、次の3つに分けて説明する。

- 選択的で名詞の役をする指示詞 *1)
- “the”
- 指示副詞 *2)

6.2.5.1 選択的で名詞の役をする指示詞

(this, these, that, those)

システマティックな区別は次のようである。

- ① 近い／近くない (this, these / that, those)
- ② 単数／複数 (this, that / these, those)
- ③ 修飾語／主要語 ($\begin{smallmatrix} \text{this} \\ \text{that} \end{smallmatrix}$ + 名詞 / $\begin{smallmatrix} \text{this} \\ \text{that} \end{smallmatrix}$ だけ)

6.2.5.1.1 近い／近くない

〔内容的距離感〕

談話中に既出のことを指示する場合、指示物に対する話者の心理的距離感が指示詞の使い方に表われる。会話の場合、自分の発言したことは“this”や“those”，相手の発言したことは“that”や“those”を用いる傾向がある。

〔個人的な馴染みの度合〕

よく知っているものに“this”や“these”を使い、あまり馴染みのないものには“that”や“those”を使う。

例 I like the lions, and I like the polar bears.
These are my favourites. Those are my favourites,
too.

〔時間的距離感〕

現在、未来のときの指示には“this”を、過去のときの指示には“that”を使う。

*1) the selective nominal demonstratives

*2) demonstrative adverb

以上は一般論であって、これがあてはまらないこともある。次の例では、“this”と“that”の区別は、以上の議論のどれにもあてはまらない。

例 What about this exhibition ?

What about that exhibition ?

すなわち、照応指示的な用法で、this と that を何を根拠にして使いわけているか、という明確な区別は、現在の研究では、未だよくわからない部分もあるといえる。

6.2.5.1.2 単数／複数

複数の意味を持っていれば、複数形をとっていなくても“these / those”で指示することがある。

例 “Go … Wait … Return …” all these directions.

逆に、“this”や“that”でたくさんのもののリストを指すこともある。

例 turkey lamb ham …… , all that food.

6.2.5.1.3 修飾語／主要語

修飾語としての指示詞は、名詞と組合せて“この○○”，“あの○○”という意味に用いられる。主要語としての方は、指示詞だけで名詞は伴わず，“これ”，“あれ”の意味を持つ。通常は人間を指示しないが例外もある。

“指示詞＋名詞”の場合、その意味は常に前提項と同じである。（指示の同一性）

すなわち

$n_1, \dots, \text{this } n_2$ （ n は名詞）

において、 $n_1 \neq n_2$ であっても n_2 が n_1 の同意語^{*1)} か上位概念の語^{*2)}（より一般的な語）であれば指示の同一性が成立^{*3)}。この現象は、語彙的意味結合^{*4)}のタイプに属する。

*1) synonym

*2) superordinate

*3) これは日本語の「この、その、あの」についても成立し、人間の hyponymy という概念階層とのむすびつきにより明確に把握できることが、電総研の田中等によって指摘されている。

*4) lexical cohesion

例. There is a cat trying to get in,

Shall I open the window ?

Oh that cat / that animal / that trickstar's
always coming hear cadging.

指示詞が名詞を伴わず主要語として用いられる場合、上と同様に同一指示の場合もあり、また前提項が指示した範囲の制限を越えてより一般的な類へと意味が拡張することがある。次の例において、“that”は“this particular merger”ではなくて、“merger in general”である。

例 There's been another big industrtal merger.

It seems that nothing can be done about this.

複数形では、これについての差は明確ではなく、あいまいでどちらともとれる場合が多い。次の例は、2通りに解釈できる。

例 How did you like the recitations ?

I find those boring.

“that”は、ある特定の recitation ともとれるし、一般的なものともとれる。しかし

I found those boring.

では、特定の“recitation”を指す。

話しことばでは、イントネーションの強弱で区別する。特定の指示のときは強く、一般化されたときは弱く発音する。

6. 2. 5. 1. 4 指示詞による拡張された指示と事実に対する指示

指示詞の場合も、事物以外のものに対する指示があり、実際の用例としてはこれらの方が多い。

例 They broke a chinese base.

(i) That was valuable
= That vase was valuable

(物体の指示)

(ii) That was carless
= That breaking of the base

(事象の指示)

(i)と(ii)の相違は、いくつもの breakingがあったとき、(i)は複数になるが(ii)は単数のままであることにもあらわれる。

例 They broke a Chinese vase and damaged two chandeliers.

(i) Those were all very valuable.

That was all very careless.

指示詞による指示が談話中の名詞項そのものを指しているか、それ以上の拡張されたものを指しているかの区別は必ずしもはっきりしているとは限らないが、実際の用例としては拡張された指示の方が多く、これは“it”についてもいえることである。

「不思議の国のアリス」なる作品の最後の2章のデータによると、その中に51個の指示詞が含まれており、内訳は次のようである。

this	22
that	24
these	3
those	2

用法の内訳は次のようである。

拡張された指示	31
時間の指示	3
(拡張された指示の別形といえる)	
談話外指示 (会話中にあらわれる)	10
既出項への前方照応	5
不 明	2

このデータから、拡張された指示が用例の大半を占めていることがわかる。

〔情報の新旧と話しことばにおける強勢〕

拡張された指示における this, that を話しことばで用いるときは、新しい情報は、強勢を伴うという原則がある。前方照応では、既に出て来たものを指すので、新情報ではなく、強勢はない。一方、後方照応では this が先に出てくるから強勢があるのがふつうである。談話外照応では、常に強勢を伴う。

6.2.5.2 " the "

定冠詞" the " は、英文法では不定冠詞" a " と関係付けてのべられているが、ここでは限定詞^{*1)}として扱う。(図6-7参照)

	指 示 詞 (Demonstutive)	所 有 代 名 詞 (pocessive)
指 示 の (Referential)	選択的 (selective) { this that these those	話し手の立場 my your our
	非選択的 (non selective) the	他の立場 { his, her, their its one's
疑 問 の (Interogative)	which	whose

図6-7 限定詞のセットにおける" the " の位置

" the " は " that " の縮形であるとみなすことができ、特定の個体^{*2)}あるいはクラス中のサブクラス^{*3)}を指定する。" the " 以外の限定詞はそれ自体が指示要素を含んでいるが、それに対して" the " はそれ自体は内容を持たず、" the " を付けた要素が、特定の、識別可能のものであることを示しているにすぎない。すなわち、どこかでそれを識別するための情報を回復できるということである。それは状況の中でも、談話の中でもよい。すなわち" the " は談話外照応にも談話内照応にも用いられる。

6.2.5.2.1 談話外照応の" the "

指示項は、次の2つの方法で識別可能である。

*1) determiner

*2) a particular individul

*3) subclass in class

- (1) 特定の個体またはサブクラスが指示され、それらが特定の状況で識別可能

例. Mind the step.

Pass me the towel.

The snow's too deep.

- (2) 状況にかかわらず、指示物がことばを超えた所^{*1)}で識別可能^{*2)}

- ・クラス中にメンバーが1つしかない。

(例. the sun)

- ・メンバーはいくつもあるが、どのメンバーか言わずともわかる。

(例. the government)

- ・クラス全体 (例. the stars)

- ・個なるものがクラス全体を代表

(例. As the child grows, he learns to be independent)

6.2.5.2.2 談話内照応の "the"

- (1) 後方照応の "the"

後方照応における "the" は、同一の名詞的グループ内の修飾部だけを指示できる。

例. The best way to achieve
The longest stretch

このように "the" は, "this", "that" などのようにこれから出てくるテキストを指示することではなく、識別可能という信号を出しているにすぎない。

- (2) 前方照応の "the"

文の意味結合に関係する "the" は、前方照応の "the" のみである。

次のような場合がある。

- ④ 同一項が繰返される場合
- ⑤ 同意語または上位概念の語の場合

*1) on extra linguistic ground

*2) これを HOMOPHORIC reference という。

㉔ 内包により前方照応の標的が供給される場合

㉔, ㉔については、既にのべた“this”, “that”の場合と同様である。

㉔については、次の例を見ると、“the eye”は明らかに“the Cat”のものである*1)

例 “How are you getting on?” said

the Cat as soon as there was mouth enough for
it to speak with.

Alice waited till the eyes appeared, and ……

“the”の典型的な機能は、前方照応指示であると一般に信じられているし、“the”は2度目以降に出てくる名詞につける冠詞といわれている。しかし、実際には、同じ名詞が2度出て来なくても、はじめての出現でも“the”をつけることがあるし、前方照応指示の場合でさえ、同一の名詞であるとは限らないことは既にみて来た通りである。

実例では、純粹に前方照応指示である“the”が大多数をしめているというわけではない。実際の会話では、“the”は第1に談話外照応に機能するし、書かれた文では後方照応の方が多いのである。そして、これらのタイプは相互排他的ではなく、1つの“the”が2つないし3つの機能を同時に果たしていることもある。次の例を見よう。

例 Last year we went Dover for a holiday.

The holiday we had there was the best we have
ever had.

上の例で、“the holiday”は、“a holiday”に対する前方照応でもあり、“we had there”に対する後方照応(構造的)でもある。これをどちらかに決めようとすることは無意味である。次の例のように、より多様性を持つ場合もある。

例 Look at the moon. / The daytime moon always
seems so sad.

“The daytime moon”の“the”は、明らかに前方照応であると同時に daytime に対し後方照応であり、また状況指示という意味でもあり、月なる天体は1つであるという単体照応的な意味でもあり、また談話外照応指示でもあるのである。

*1) “the”と“the Cat”の間には、部分全体関係がある。この関係は partonymy とよばれ、日本語の“その”の使用についても成立し、階層構造で把握できることが電総研の田中等によって指摘されている。

6.2.5.3 指示副詞

指示詞的な副詞としては, "here", "there", "now", "then" がある。これらは, 指示詞的な役割以外にも用いられる。注意すべき同形異義の区別は, 次のようなものがある。

- (1) 指示詞的な "there" と代名詞の "there"

(例 There is a)

- (2) 指示詞的な "now" と接続詞の "now"

(例 now what we're going to do is this.)

- (3) 指示詞的な "then" と接続詞の "then"

(例 then you are quite made up your mind ?)

"here" と "there" は, "this" と "that" に似ており, 取り換えられることもある。位置の指示の "there" は前方照応に用いられるが, "here" は ("this" と同様に) 後方照応もできる。英語の後方照応の源はほとんど "this", "there", "here" の3つだけである。また, "here" も "there" も拡張されたテキストを指示することができる。

時の指示詞としての "then" は, 意味結合機能があり, "at the time just referred to" という意味を持ち, 前方照応指示的である。

例. In my young days we took these things more seriously. We had different ideas then.

文頭に付加詞としての指示詞を含む表現がくることがよくある。(例 in that case, that being so, after that, at this moment, under these circumstances) これらは, 接続詞的なものであるので, 指示詞の用法からは除外した。

6.2.6 比較による指示^{*1)}

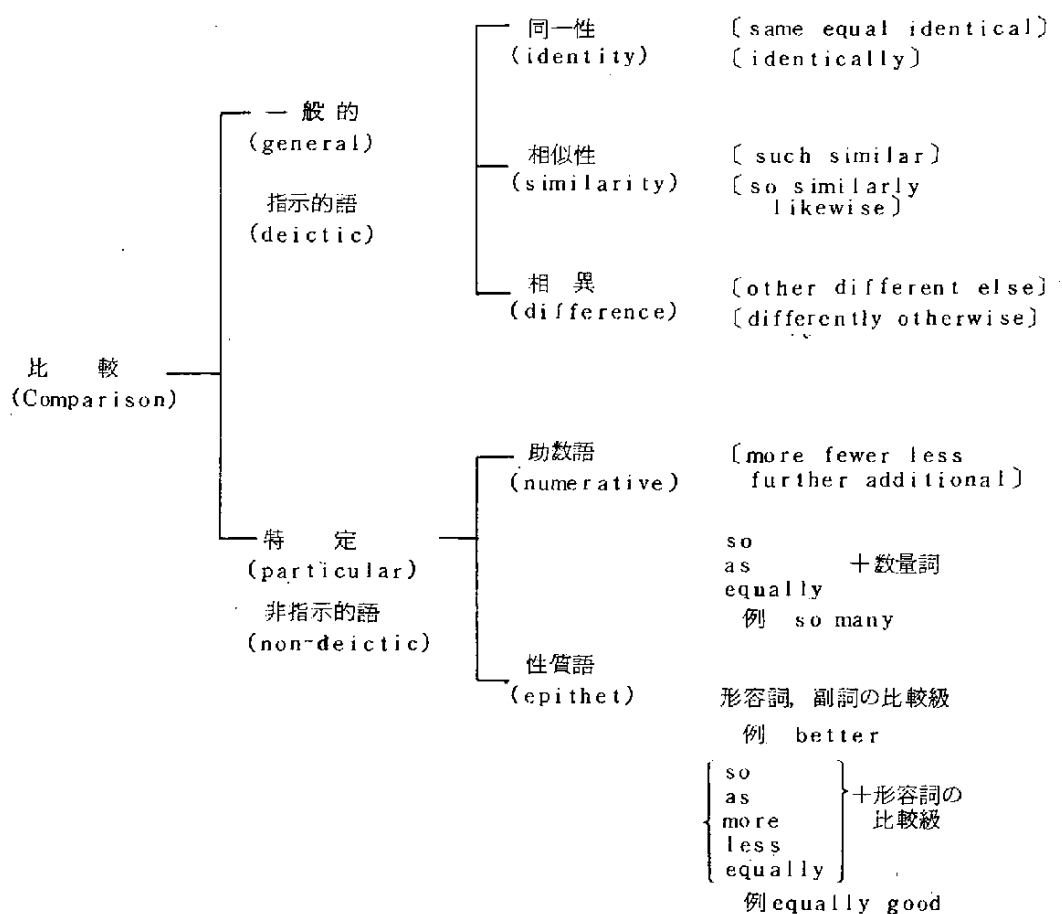


図 6-8 比較による指示のシステム

*1) comparative reference

6.2.6.1 一般的比較^{*1)}

一般的比較は2つのものが似ているか似ていないかを問題にする。詳しくは、2つのものが同一であるか、似ているか異なるか(同一でない、もしくは似ていない)を意味し、形容詞か副詞で表現される。以下にこの3つの場合の例を示す。

- 例 a. It's the same cat as the one we saw yesterday.
b. It's a similar cat to the one we saw yesterday.
c. It's a different cat from the one we saw yesterday.

上の3つは、文法構造で決定された後方照応指示の例であるが、以下の例は、セミコロンで分かれ、構造的には関係ない2要素間の後方照応で、意味結合を成している。

- 例 The other squirrels hunted up and down.
the nut bushes ; but Nuskine gathered robin's
pinecones off a beech bush, and stuck them
full of pine needle pins.

次の例で、文a.と文b.がこの順であれば、“the other seats”は、前方照応であるが、文bが単独であれば、談話外照応で、“other than those you see here”の意味になる。

- 例 a. They have given us special places in front row.
b. Would you prefer the other seats ?

前方照応する比較^{*2)}の場合、“so”や“such”では、文のかたまり^{*3)}を指示することがある。範囲はどんなに広くてもかまわない。前出の議論全体^{*4)}を指示することもある。(例 the same question arise ……)

“same”, “similar”, “identical”, “equal”, “different”などが出てきても、必ずしもはっきりした指示物がわかるとは限らない。次の次は、2通りに解釈し得るあいまいな例である。

- 例 Most people have the same breakfast everyday.

*1) general comparison
*2) anaphoric comparison
*3) passage
*4) fact とみなされる。

解釈 1 毎日同じものを食べる。

解釈 2 誰でも同じものを食べる。

名詞の役をする語群において、比較を表わす形容詞は、性質語として機能するから、助数語の次にくる。

例 There were two different cards.
= different from each other.

それに対し、指示的語として機能するときは、助数語の前にくる。

例 There were a different two cards.

"other", "additional", "else" は指示感覚で用いられるが、
"equal" にはそれがない。"other" には "different" の意味と両方がある。

6.2.6.2 特定の比較

特定の比較は、質または量の比較を意味し、ふつうの形容詞、副詞の比較級の形で表現する。形容詞は、名詞の役をする語群の中に位置し助数語または性質語である。

(1) 比較が量について行われる場合

(i) 数量詞^{*1)}の比較級 例 more mistakes.

(ii) 数量詞の修飾語としての副詞の比較級 例 as many mistakes.

(2) 比較が質について行われる場合

(i) 名詞の役をする語群の中の性質語の要素

(a) 形容詞の比較級 例 easier task.

(b) 形容詞の修飾語としての副詞の比較級

例 so difficult a task

(ii) 節の中の付加詞

(a) 副詞の比較級 例 Cambridge rowed faster.

(b) 副詞の修飾語としての副詞の比較級

例 she sang as sweetly.

特定の比較も指示的であり得る。比較の対象となって「優」、「劣」、「等」

*1) quantifier

の判定のもととなった規準としての指示物が存在する。

一方、特定の規準対象を持たぬ、一般的な状況指示の場合もある。

最上級は、最も一般化されたものである。これは自己限定的^{*1)}であるので、非指示的^{*2)}である。

後方照応の比較級には、意味結合に関係するものと、文法構造的結合であって意味結合には関係ないものがある。

比較による談話外照応は、日常生活ではよく用いられる。(例 Go slower /)

意味結合をもたらすのは前方照応指示の場合である。

6.2.6.3 "so", "such", "as" について

"so", "such", "as" は、3つとも同じ意味を持ち、"similar(ly)" を表わす。

"extremely" を意味する場合の"so", "such" は、"so that", "such that" において、後方照応するが、これらは文法構造的関係であり意味による結合ではない。

談話外照応には"so", "such" を用いることがあるが、"as" は用いられない。強い人を見て、

"I never thought he was so strong"

とほいうが、

"I never thought he was as strong"

とはいわない。

最後に"so", "such", "as" について、前方照応指示の例を挙げる。

例 a. He seemed most upset — I never knew he cared so.

b. He has a lean and hungry look.

Such men are dangerous.

c. "Are five nights warmer than one night, then?"

*1) self defining

*2) non-referential

Alice returned to ask. "Five times as warm,
of course"

"so", "such" は, 指示的な意味結合ばかりでなく, 「代入」^{*1)} や「接続」^{*2)} などの, 他の意味結合機能も持っている。

6.3 指示現象と代入・省略現象の比較

6.3.1 指示と代入・省略の対比の概観

次の例文の対比から直観的にこの3つの相異を知ることができるであろう。

文1. This is a fine hall you have here.

文2. { a. I am proud to be lecturing in it.
b. I have never lectured in a fine one.
c. I have never lectured in a finer. (φ)
φ : empty

a. の "it" は "hall" と同一指示的であるものに対し, b. の "one" は "hall" ということばに対する「代入」である。c. の φ は, b と同様であるが「ゼロ代入」, 「省略」である。

これらの共通点は, 3つとも "hall" を前提要素として, 聞き手がそれを知っており, はっきりと言わなくても回復できるものとしていることである。この例では, "hall" は前出の文の中にあり2つの関係が a, b, c 共, 文1. と文2. を結合している。

図6-9に指示と代入・省略との対比を概観する。

*1) substitution

*2) conjunction

	指 示 (Reference)	代 入 ・ 省 略 (Substitution・Ellipsis)
抽象化のレベル	意 味 的 (semantic)	語 文 法 的 (Lexicogrammatical)
前提するものの源	談話の状況全体	談 話 自 体
前提されるもの	意 味	項(語, 語群, 節)
語のクラスは保存されるか	必ずしもされない	さ れ る
前提された項による置換は可能か	必ずしもできない	で き る
文連結手段 (cohesive device) として使用できるか	で き る 前方照応と後方照応	で き る ほとんど前方照応

図 6-9 指示と代入・省略との対比の概覧

6.3.2 指示と代入との相異

指示と代入の相違のポイントは、指示が意味の関係を問題にするのに対し、代入は文法的関係であり、語や句などの語彙項目の関係であることである。

次の例で、“one”も“it”も共に“item”を前提項としているけれども“one”は代入で、“it”は指示である。

例. The relation between an item and
another one that presupposes it could
be explained as a ………

指示は、談話外照応にしろ談話内照応にしろ、周辺環境から意味が回復できることが前提になっていたが、代入はある要素をくり返す代りに使うカウンタ

である。したがって、代入の大半は談話内照応であり、かつその大半は前方照応である。その性質上、談話外照応の代入というものは、きわめて稀である。

代入は、ことばの形の上での関係であるから、文法構造上の機能も問題になる。次の例は、名詞代入と動詞代入の例であるが、“ax”も“one”も名詞の役をする語群の主要語であり、また“knows”も“does”も動詞の役をする語群の主要語である。

例 my (ax) is too blunt. I must get a sharper (one).

例 You think Joan already (knows) ?

I think every everyone (does).

このように、代入頂ともとの頂は同じ構造的機能を持つ。このような制限は、指示現象の場合には見あたらないものである。

6.3.3 代入と省略との相異

代入と省略とは同じようなものであるが、構造のメカニズムは異なるともいえる。省略された要素のある語群には、前文から得られたもので埋められるべきスロットが残されている。省略は、“one”、“do”などの代入の代りにゼロ代入が行われているといっても良い。

省略現象は、省略してわからなくなってしまうことではない。省略現象における省略は、「言わずともわかるもの」についてのみ行われる。

例 Hardly anyone left the country before the war.

この文を見るとわからないことがいっぱいあるが、これを知ることができる情報が前文に存在することを示す信号としての項の欠落、文構造の中で何か脱けているものがあるわけではない。このような「わからなさ」は省略ではない。それに対して、次の例では、欠けている要素が明確にわかる。

例 Joun brought some carnation, and Catherine some sweet pies.

このように、省略現象とは、何か構造的に必要なものが言われずに残されていることである。

代入と同様に、省略もまたその性質上、大多数は前方照応的である。

参考文献

— [本] —

- [Allwood 77] Allwood, J., Andersson, L. and Dahl, O.: "Logic in Linguistics", Cambridge Univ. Press, 1977.
- [Austin 62] Austin, J. L.: "How to Do things with Words", Oxford Univ. Press, 1962.
- [Bobrow 75] Bobrow, D. G. and Collins, A. M. (Eds.): "Representation and Understanding", Academic Press, 1975.
- [Boden 77] Boden, M.: "Artificial Intelligence and Natural Man", Basic Books, 1977.
- [Chafe 70] Chafe, W. L.: "Meaning and Structure of Language", The Univ. of Chicago Press, 1970.
- [Charniak 76] Charniak, E. and Wilks, Y.: "Computational Semantics", North-Holland, 1976.
- [Cofer 76] Cofer, C. N. (Ed.): "The Structure of Human Memory", Freeman and Company, 1976.
- [Halliday, 76] Halliday, M. and Hassan, R.: "Cohesion in English", Longman, London, 1976.
- [Lehnert 78] Lehnert, W. G.: "The Process of Question Answering", Lawrence Erlbaum Associates, Publishers, 1978.
- [Lyons 68] Lyons, J.: "Introduction to Theoretical Linguistics", Cambridge Univ. Press, 1968.
- [Miller 76] Miller, G. A. and Johnson-Laird, P. N.: "Language and Perception", Harvard Univ. Press, 1976.
- [Newell 73] Newell, A. et al.: "Speech Understanding Systems--Final Report of a Study Group", North Holland/American Elsevier, 1973.
- [Norman 75] Norman, D. A. and Rumelhart, D. E.: "Explorations in Cognition", Freeman and Company, 1975.
- [Ogden 23] Ogden, C. K. and Richards, I. A.: "The Meaning of meaning", London: Kegan Paul, 1923.
- [Schank 75] Schank, R. C.: "Conceptual Information Processing", North-Holland American Elsevier, 1975.
- [Schank 77] Schank, R. C. and Abelson, R.: "Scripts Plans Goals and Understanding", Lawrence Erlbaum Associates, Hillsdale, New Jersey, 1977.
- [Searle 69] Searle, J. R.: "Speech Acts: An Essay in the Philosophy of Language", Oxford Univ. Press, 1969.
- [Walker 78] Walker, D.: "Understanding Spoken Language", North-Holland, 1978.

- [Frederiksen 75] Frederiksen, C. H.: 'Representing Logical and Semantic Structure of Knowledge Acquired from Discourse', *Cognitive Psychology*, 7, 1975, 371-458.
- [Grice 75] Grice, H. P.: 'Logic and Conversation', in Cole, P. and Morgan, J. L. (Eds.), 'Syntax and Semantics', Vol. 3, Academic Press, 1975, 41-58.
- [Grosz 77] Grosz, B. J.: 'The Representation and Use of Focus in a System for Understanding Dialogs', *IJCAIS*, 1977, 67-76.
- [Hobbs 79] Hobbs, J. R.: 'Coherence and Coreference', *Cognitive Science* 3, 1979, 67-90.
- [Kahn 75] Kahn, K. M.: 'Mechanization of Temporal Knowledge', MIT MAC-TR-155, Aug. 1975.
- [Kintsch 72] Kintsch, W.: 'Notes on the Structure of Semantic Memory', in Tulving and Donaldson (Eds.), 'Organization of Memory', Academic Press, 1972.
- [Kuno 77] Kuno, S. and Kaburaki, E.: 'Empathy and Syntax', *Linguistic Inquiry*, Vol. 8, No. 4, 1977.
- [Lakoff 76] Lakoff, G.: 'Pronouns and Reference', in McCawley, J. D. (Ed.), 'Syntax and Semantics', Vol. 7, Academic Press, 1976, 275-335.
- [Lehnert 77] Lehnert, W.: 'Human and Computational Question Answering', *Cognitive Science*, Vol. 1, No. 1, 1977, 47-73.
- [McDermott 74] McDermott, D. V.: 'Assimilation of New Information by a Natural Language-Understanding System', AI TR-291, MIT, Feb. 1974.
- [Minsky 75] Minsky, M.: 'Framework for Representing Knowledge', in Winston (Ed.), 'The Psychology of Computer Vision', McGraw-Hill, 1975.
- [Moore 74] Moore, J. and Newell, A.: 'How Can Merlin Understand?', in Gregg (Ed.), 'Knowledge and Cognition', Lawrence Erlbaum Associates, Publishers, Potomac, Maryland, 1974.
- [Nagao 76] Nagao, M. and Tsujii, J.: 'Analysis of Japanese Sentences Using Semantic and Syntactic Information', *AJCL*, Microfiche 41, 1976.
- [Nash-Webber 77] Nash-Webber, B. and Reiter, R.: 'Anaphora and Logical Form: On Formal Meaning Representations for Natural Language', *IJCAIS*, 1977, 121-131.
- [Nash-Webber 78] Nash-Webber, B. L.: 'A Formal Approach to Discourse Anaphora', BBN Report No. 3761, May 1978.
- [Newell 73] Newell, A.: 'Artificial Intelligence and the Concept of Mind', in Schank and Colby (Eds.), 'Computer Models of Thought and Language', Freeman and Company, 1973.

[Winograd 77] Winograd, T.: "Understanding Natural Language", Academic Press, 1977.

[Winston 77] Winston, P. H.: "Artificial Intelligence", Addison Wesley, 1977.

— [論文] —

[Abelson 73] Abelson, R. P.: "The Structure of Belief Systems", in Schank and Colby (Eds.), "Computer Models of Thought and Language", Freeman and Company, 1973.

[BBN 78] BBN Report 3878: "Research in Natural Language Understanding", Quarterly Progress Report No. 3, 1978.

[Bien 76] Bien, J. S.: "Multiple Environments, Approach to Natural Language", AJCL, Microfiche 54, 1976.

[Bobrow 77a] Bobrow, D. G. and Winograd, T.: "An Overview of KRL, a Knowledge Representation Language", Studies in Cognitive Science, Vol 1, No 1., 1977.

[Bobrow 77b] Bobrow, D. G., Kaplan, R. M., Kay, M., Norman, D. A., and Winograd, T.: "GUS, A Fama-Driven Dialog System", Artificial Intelligence, Spring, 1977.

[Brown 75] Brown, J. S. and Burton, R. R.: "Multiple Representations of Knowledge", in Bobrow D. G. and Collins A. (eds.), "Representation and Understanding", Academic Press, 1975, 311-349.

[Bullwinkle 77] Bullwinkle, C.: "Levels of Complexity in Discourse Anaphora Disambiguation and Speech Act Interpretation", IJCAIS, 1977, 43-49.

[Chafe 72] Chafe, W. L.: "Discourse Structure and Human Knowledge", in Carroll and Freedle (Eds.), "Language Comprehension and the Acquisition of Knowledge", John Wiley and Sons, 1972.

[Chafe 73] Chafe, W. L.: "Language and Memory", Language, 49, 2, 1973, 261-281.

[Chafe 74] Chafe, W. L.: "Language and Consciousness", Language, 50, 1, 1974, 111-133.

[Charniak 72] Charniak, E.: "Toward a model of Children's Story Comprehension", AI TR-226, MIT, December 1972.

[Charniak 73] Charniak, E.: "Jack and Janet in Search of a Theory of Knowledge", IJCAIS, 1973, 337-343.

[Charniak 77] Charniak, E.: "Ms. MALAPROP, A Language Comprehension Program", IJCAIS, 1977, 1-7.

[Fillmore 77] Fillmore, C. J.: "Scenes-and-frames Semantics", in Zampolli, A. (Ed.), "Linguistic Structures Processing", North-Holland, 1977.

[Norman 73] Norman, D.: "Memory, Knowledge and the Answering of Questions" in Solso, R. L. (Ed.), "Contemporary Issues in Cognitive Psychology: The Logora Symposium", U. H. Winston and Sons, Washington, D. C., 1973.

[Petrick 76] Petrick, S. R.: "On Natural Language Based Computer Systems", IBM J. Res. Develop., July, 1976, 315-325.

[Phillips 79] Phillips, B.: "A Model for Knowledge and its application to Discourse Analysis", AJCL, Microfiche 82, 1979.

[Quillian 68] Quillian, M. R.: "Semantic Memory", In Minsky (Ed.), "Semantic Information Processing", The MIT Press., 1968.

[Rieger 74] Rieger, C. J.: "Conceptual Memory", Stanford AIM-247, 1974.

[Rumelhart 75] Rumelhart, D. E.: "Notes on a Schema for Stories", in Bobrow and Collins (Eds.), "Representation and Understanding", Academic Press, 1975.

[Schank 73] Schank, R. C. et al.: "MARGIE: MEMORY, ANALYSIS, RESPONSE GENERATION, and INFERENCE on ENGLISH", IJCAI3, 1973, 255-261.

[Schank 74] Schank, R. C. and Rieger, C.: "Inference and Computer Understanding of Natural Language", Artificial Intelligence, 5, 1974, 373-412.

[Schank 75] Schank, R. C. et al.: "SCRIPT, PLANS and KNOWLEDGE", IJCAI4, 1975, 255-261.

[Schank 76] Schank, R. C.: "The Role of Memory in Language Understanding", in Cofer, C. N. (Ed.), "The Structure of Human Memory", Freeman and Company, 1976, 162-189.

[Sidner 78] Sidner, C.: "A Progress Report on the Discourse and Reference Components of PAL", MIT AI Memo 468, 1978.

[Steedman 77] Steedman, M. J.: "Verbs, Time and Modality", Cognitive Science, Vol. 1, No 2, 1977, 216-234.

[Stenning 78] Stenning, K.: "Anaphora as an Approach to Pragmatics", in Halle, M., Bresnan, J. and Miller, G. A. (Eds.), "Linguistic Theory and Psychological Reality", The MIT Press, 1978, 162-200.

[Takahasi 72] Takahasi, H., Fujimura, O. and Kameda, H.: "A Behavioral Characterization of Topicalization: What is Topicalization? -- From a Computer System Point of View", Annual Bull. No. 6, Research Institute of Logopedics and Phoniatrics, University of Tokyo, 1972, 93-110.

[Tulving 72] Tulving, E.: "Episodic and Semantic Memory", in Tulving and Donaldson (Eds.), "Organization of Memory", Academic Press, 1972.

[Winograd 77] Winograd, T.: "A Framework for Understanding Discourse", Stanford AIM-297, April 1977.

[Winograd 75] Winograd, T.: "Frame Representations", in Bobrow and Collins (Eds.), "Representation and Understanding", Academic Press, 1975.

[Young 78] Young, R.: "Text Understanding: A Survey", AJCL, Microfiche 70, 1978.

〔Tanaka, 78〕 田中穂積, 内田ユリ子

“意味抽出における概念階層の利用と手続き付加”, 人工知能と対話技法 8-1, 1978.

〔Teratu, 79〕 寺津典子, 山梨正明

“日本語における照応現象について(その1)”,

文部省科学研究費特定研究(1)課題番号 310702

〔Teratu, 80〕 寺津典子, 稲田俊明, 山梨正明

“日本語における照応現象について(その2)”, 1979

文部省科学研究費特定研究(2)課題番号 410202, 1980

〔Uchida, 79〕 内田ユリ子, 田中穂積

“「ソノ」の分析と理解のモデル(その1)”, 情報処理学会第20回全国大会報告書, 1979.

7. 音 声 理 解

7.1 はじめに

音声は人間の用いる種々の意志伝達手段の中で最も根源的なものであり、日常生活において人間相互間の通信の大部分は音声言語によって行われている。このように、人間相互間の通信に最も重要な役割を果たしている音声の適用範囲を拡大して、人間と機械との間の交信にも利用したいと考えるのは極く自然な要求である〔Fujisaki 72〕。音声認識の研究は、この要求に応えるものとして1940年代から盛んに行われるようになった。現在では、特定話者の発声した100語程度の単語音声の認識装置は実用の域に達している。

しかし連続音声あるいは会話音声の認識においては多数の単語が連続して発声され、単語の境界が不明確になるとともに、単語境界附近の音が、先行または後続の単語の影響を受けて変化する（調音結合）ため、単語音声認識の手法をそのまま適用することはできない。音声は単なる音響信号ではなく、言語音声としてその意味内容等を伝達するために発せられているものであるから、音声自動認識の過程でもこのようなことを考慮しなければならないと考えられるようになって来た（図7-1参照）。

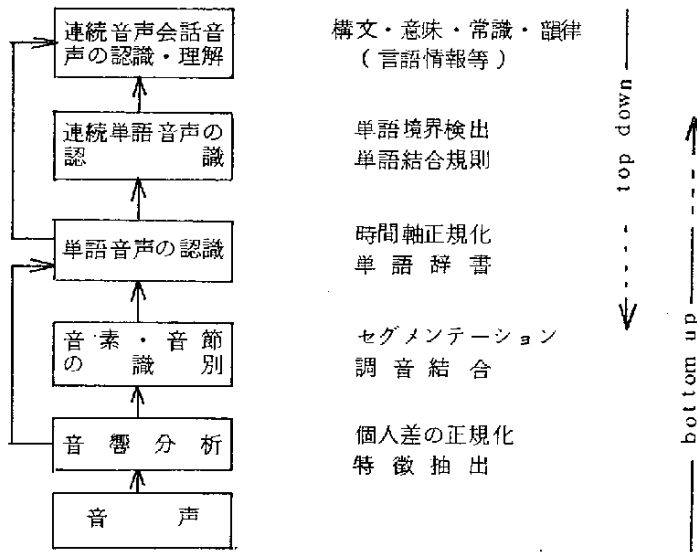


図7-1 音声認識の階層と関連する問題

7.2 音声理解研究

音声自動認識の過程に言語情報を導入する試みは古くからある(図7-2参照)。Denes による音素遷移確率の単語認識への利用〔Denes 59〕, 坂井・堂下による音素遷移確率の利用〔Sakai 63〕, Alter によるフォートラン語の認識における構文情報の利用〔Alter 68〕, Vicens による制限された自然言語文の音声認識〔Vicens 69〕等が先駆的な研究である。これらの研究を経て, Fant〔Fant 70〕やReddy〔Reddy 71〕によって言語情報を用いた自然言語音声の認識システムのモデルが提案された。最近になって人工知能研究, 特に自然言語処理研究の進展に伴って, 言語情報等を積極的に利用して会話音声の認識を行おうとする気運が出て来た〔Nakatsu 79〕。

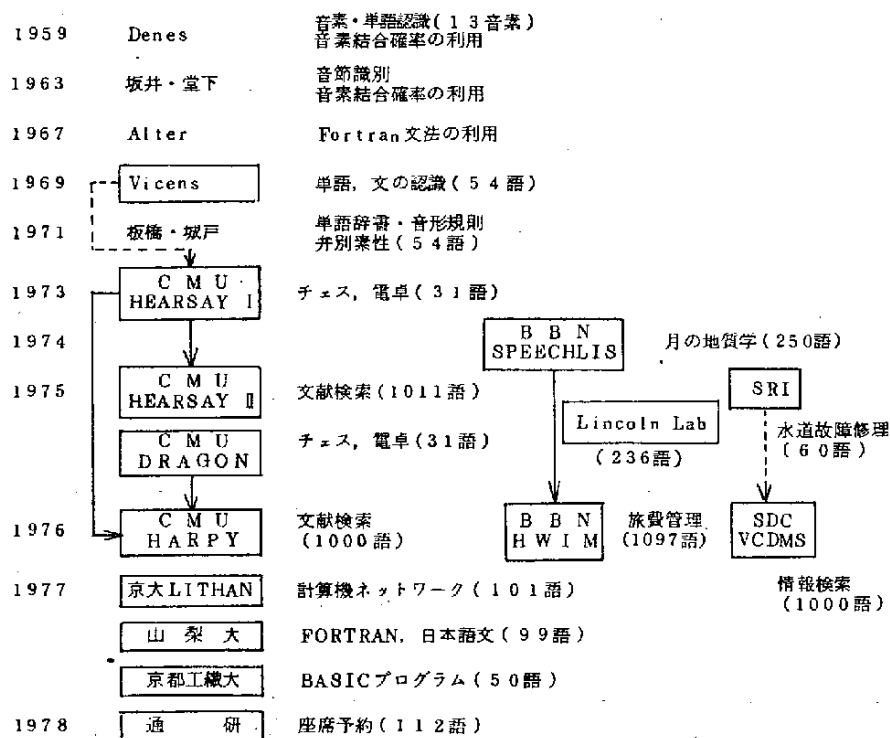


図7-2 音声理解研究の発展

7.2.1 音声理解の問題

ARPA (アメリカ国防省のAdvanced Research Projects Agency) は, 実用的なマン・マシン通信に使い得る音声認識システムを実現することを目的として, 1971年11月に5年間の研究開発計画を開始した〔Newell 73〕。

この計画で設定された目標の概要を表7-1の左側に示す。この計画では、従来からの音声研究を進展させることの他に、構文・意味・会話の内容等に関する言語情報をできるだけ有効に使って認識を行うことが強調された。また入力音声の一語一語を正確に認識することは必ずしも必要でなく、入力音声の意味を理解してシステムが正しい応答をできることが目標とされた。そこでこれを「音声理解システム」SUS (Speech Understanding System)と呼んでいる。さらにこの研究では、音声入力によってシステムに行わせようとする仕事(タスク)が明確に定められていることが特徴である。典型的な音声理解システムの構成を図7-3に示す〔Niimi 79〕。

音声理解の研究においては、具体的なタスクを設定することにより、単語音声の場合にはなかった種々の情報を利用することができる。すなわち、文の構造を規定する構文情報、タスクで使用される単語間の概念関係や属性関係などを表わす意味情報、タスクによって設定された世界における一般常識、音声認識システムとその利用者の間で交わされる一連の対話を通じて得られる文脈情報、音声のイントネーションや強調などの韻律情報などである。これら種々の異質の情報を統合して一つのシステムを構成することが、音声理解の課題である〔Niimi 79〕。

「音声理解システム」という考え方はARPA計画で提唱されたものであり、アメリカで最も盛んに行われて間もなく日本やヨーロッパでも取り上げられるようになった。従って以下ではARPA関係の記述が多くなっていることを予めお断りしておく。

表7-1 ARPAの5年間のプロジェクトの目標と
Harpy の性能との比較

目 的 (1971.11)	Harpy (1976.11)
連続発声を認識できる	YES
多数の	5人 (男性3人, 女性2人)
協力的な発声音を用いる	YES
静かな室で	計算機のターミナルルームで
高品質のマイクロホンを用いる	接話形のマイクを用いる
発声者ごとにシステムをわずかに	発声者ごとに 20 文章を用いて学
適応化してよい	習を行う
1,000 語の語彙を対象とする	1011
人工的な構文を用いて良い	平均の branching factor=33
制限されたタスクを対象とする	文献検索
意味的な誤りは 10% 以下	5%
実行時間の数倍以内で動作する	実行時間の 80 倍
100 MIPS の計算機を用いる	0.4 MIPS の計算機 (PDP-KA
(MIPS=million machine in-	10) 上で 36 bit 256 K のメモ
struction per second)	リを用いる。
	(1文章当りの処理コストは5ドル)

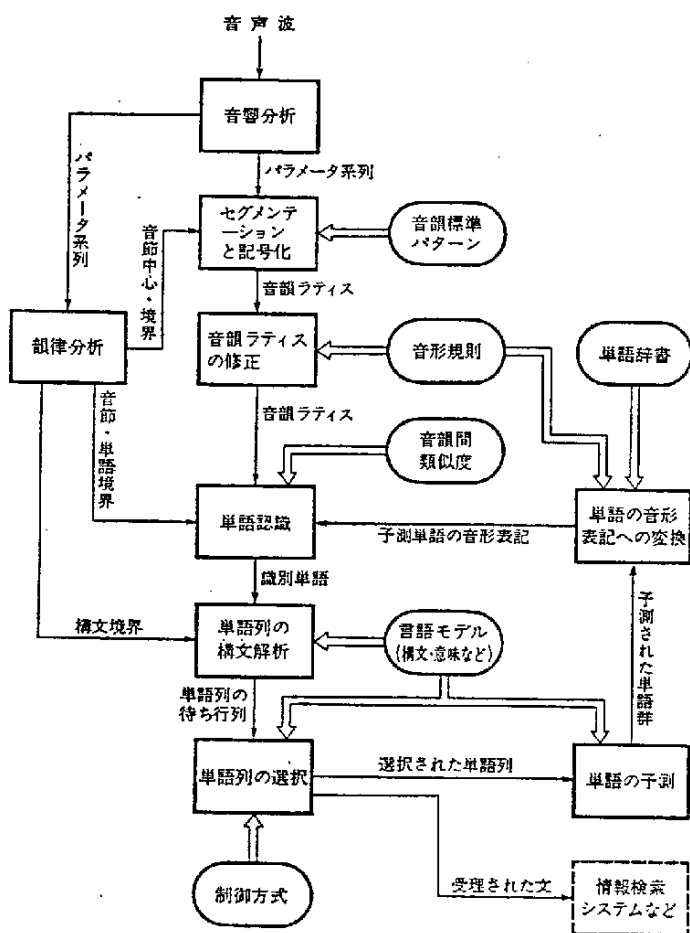


図 7-3 典型的音声理解システムの構成 [Niimi 79]

7.2.2 初期の音声理解システム

ARPA計画以前の言語情報を利用した音声認識システムおよびARPA計画の初期の頃のシステムの幾つかについて概説する。

(1) Alter のフォートラン認識装置〔Alter 68〕

音声自動認識システムは音響認識部(A)と言語認識部(L)に分かれている。Aとしてはタイプライターのキーインを利用しており、その正答率としては50～80%を仮定している。Aの出力の誤りには置換、脱落、附加の3種類ある。Lの機能はAの誤りを訂正し、単語のセグメンテーションを行うことである。

言語としてはフォートランを採り上げ、語彙は変数を除いて150語程度である。ただし変数名にはある種の制約を設けている。誤りの中には単語個有の誤りと単語によらない誤りとがあるが、まず後者のみを考えている。誤り訂正には順序復号化法(sequential decoding)を用いている。これは、入力信号が与えられたとき、復号化された推定値が真の値である条件付確率を最大にするものである。

構文はフォートランの制御文と割当文のすべてを含むが、入出力文・宣言文は含まない。そのため変数名に次のような制約をつけている。すなわち、変数はアルファベット4文字で構成され、初めの2文字が変数の型を表わしている。これは単語系列として処理される。

5～100文字から成る数百のステートメントでテストしたところ、処理は1秒以下で済んだものから30分要したものまでである。誤り訂正の例を表7-2、表7-3に示す。

表7-2 変数の誤り訂正例

INTENDED SENTENCE: RVA=Ø		
Actual Input	Decoded Sentence	Time (seconds)
RVA=Ø	RVA=Ø	0.48
QVA=Ø	RVA=Ø	0.83
RQA=Ø	RVA=Ø	1.00
RVA±Ø	RVA=Ø	1.45
RVAQØ	RVA=Ø	18.44
RVA=Q	RVA=Ø	3.65
QQA=Ø	STOP Ø	4.01

表 7-3 フォートラン語の誤り訂正例

INTENDED SENTENCE: CONTINUE		
Actual Input	Decoded Sentence	Time (seconds)
CONTINUE	CONTINUE	0.22
CONTANUE	CONTINUE	0.24
CONTANDE	CONTINUE	0.24
CONGANDE	CONTINUE	0.29
COI GANDE	CONTINUE	0.38
CONGAJGO	CONTINUE	0.60
CILGANDE	CALL SBDE	2.83
CONINUE	CONTINUE	0.57
CONGTINUE	CONTINUE	0.46
CONIGUE	CONTINUE	0.57

(2) Vicens の音声認識システム [Vicens 69]

実時間処理に近い音声認識システムを計算機によって実現している。このシステムは前処理、セグメンテーション（区分）、セグメント分類、単語認識、文認識の各段階から成る。前処理では音声信号を3周波数帯域（150～900, 900～2200, 2200～5000 Hz）に分け、各帯域出力の振幅と零交叉数（合計6個のパラメータ）を10 ms ごとに得ている。パラメータの時間的変化を見て、隣接する類似のものはまとめて定常区間とし、変化部分は過渡区間としてセグメンテーションを行う。各区間のパラメータの平均値により、セグメントに音素群（母音、摩擦音、鼻音、子音、バースト）のラベルづけを行なって分類する。セグメンテーションの結果は特徴行列として表わされる。6個のパラメータの他に、第1行には母音や無声摩擦音の数、セグメント数、母音セグメントのポインター、母音と摩擦音/s/の場所等の大まかな情報を置いてある。セグメンテーションには実時間の1.5倍を要する。

学習モードでは、上述のパラメータパターンが辞書に記憶される。認識モードでは、まず、発声の大まかな特徴を考慮して幾つかの候補を辞書から取り出す。各候補を入力パラメータと比較し、最も良いマッチングを示すものを選ぶ。比較操作では、最初にセグメントの間の対応を決定し、次いで対応するセグメントの間の類似度を、パラメータの値によって評価する。

ある文法によって定義される制限された言語の文を対象とするときは、特

微行列で表現された入力文を前後に走査して（辞書にある）単語を控し出す。文法からのフィードバックを利用して、構文的に正しくない単語表現をマッチングの過程から除く。結果は次の通りである。

1) 54単語, 話者1名, 学習4回, 98%

話者1名, 学習1回, 92%

処理時間2~3秒/語

2) 54単語, 話者10名, 学習9回, 85~90%

処理時間9~12秒/語

3) 70単語(フランス語), 話者1名, 97%

処理時間2~3秒/語

4) 561単語 } , 話者1名, 学習3回, 92%

短文 } 処理時間16~17秒

以上の結果から次のような結論を出している。

1) 3周波数帯域から得られた大まかなパラメータで上述のような満足すべき結果が得られたということは、前処理よりも後段のアルゴリズムの方が大事であることを示している。

2) 音声分析で用いる基本単位（音素、音節、単語）の中でどれが良いかということはあまり意味がない。この研究では処理の各段階ですべて使っている。

3) 限定された言語を正しく認識するには、音素（のような）単位への分類が正確に行われていなくても良い。

4) ヒューリスティックな方法で探索範囲を狭めるAI（人工知能）の手法が音声認識に非常に有効であると思われる。

5) 前処理部に費された努力よりも、構文を重視した文解析プログラムを作る方が実り多いようである。

(3) Speechlis

BBNでは、このプロジェクトが始まる前にLUNARという名の、質問応答システムで、Apollo 11の持ち帰った月の岩石の化学分析結果に関したものを開発していた。このシステムは、約3500語の語彙を持ち、入力される文章の例は、次の様なものである。

「高アルカリの岩石中の、ルビジウム平均濃度はいくらか？」等。

このシステムの入力を音声に、語彙を250語に、文法を縮小したものを課題にしたものがSpeechlisシステムである。

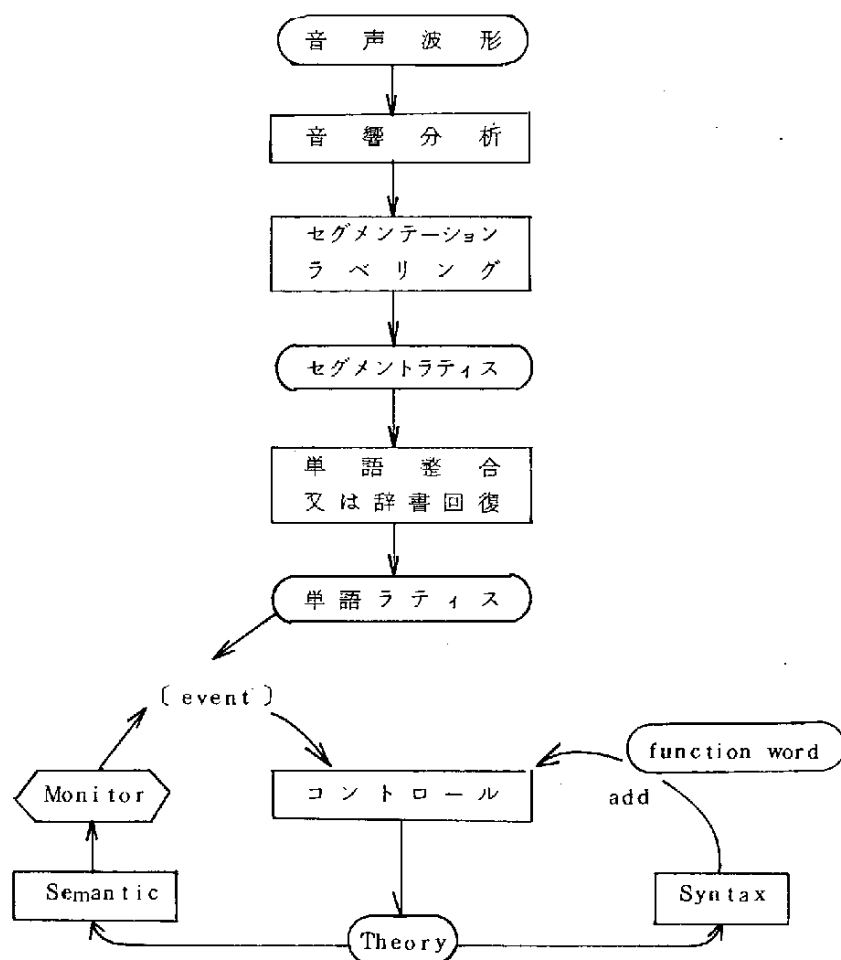


図7-4 Speechlisの構成

図7-4がデータの流に注目した場合の、システムのブロックダイアグラムである。このシステムの全体の構成は、Klatt と Stevens のサウンドスペクトログラムを読む実験 (Klatt 72) が、一つの基礎になっているので、人間が、スペクトルから文章まで、試行錯誤の末に到達する過程が十分に参考になっている。その為、このシステムには、音響分析から、意味、構文、それら全体の制御の部分まで、バランス良く取り入れられ、研究されている。各部の説明に入ろう。まず20KHzでサンプリングされた音声波形は、音響分析部へ送られる。ここで分析されて出力されるパラメータは、次の8種類である。

- 1) 原波形の全エネルギー
- 2) 微分波形の全エネルギー
- 3) 微分波形の低い周波数成分のエネルギー
- 4) 自己相関係数の第1項
- 5) 正規化予測誤差
- 6) 周波数スペクトルの変化分
- 7) 基本周波数
- 8) 2-pole モデルと14-pole モデル各々で線形予測を行って求めた pole 周波数

この特徴ベクトルは、セグメンテーションとラベリング部へ送られる。この部分では、音響→音韻変換規則等を用いて、各カテゴリとの適合度を調べ、閾値以上であったカテゴリ名を適合度の順に並べた、セグメントラティスを作る(図7-5(a))。このセグメントラティスから、単語辞書中の内容語 (content word) との整合度を計算し、その整合度の点数とともに、点数の順に候補単語を並べた、単語ラティスを作るのが、単語整合部である(図7-5(b))。この単語ラティスを用いて、意味、構文、プラグマティクス部が、制御部を中心として、発話内容の解釈を行なう。意味解析部には、Semantic Network と Case Frame Token が用いられ、構文解析には、Transition Network Grammar が用いられている。この上部処理の部分では、意味解析を主に行なう。単語ラティスから点数の高い単語を候補単語として取り上げると、意味の制約により、この単語の左又は右側に表われる可能性の高い単語を提案して、単語ラティスへ探しに行く。適当な点数の単

語が、適当位置に見つかり、構文情報やプラグマティクスを用いて、更に機能語（function word）等を付加したり、単語候補をしぼったりしながら、しだいに発話内容を推定していく。最後に、発話内容全体に渡って、意味のある、構文として正しい、発話と良く整合する、単語列（＝文）が決定される。

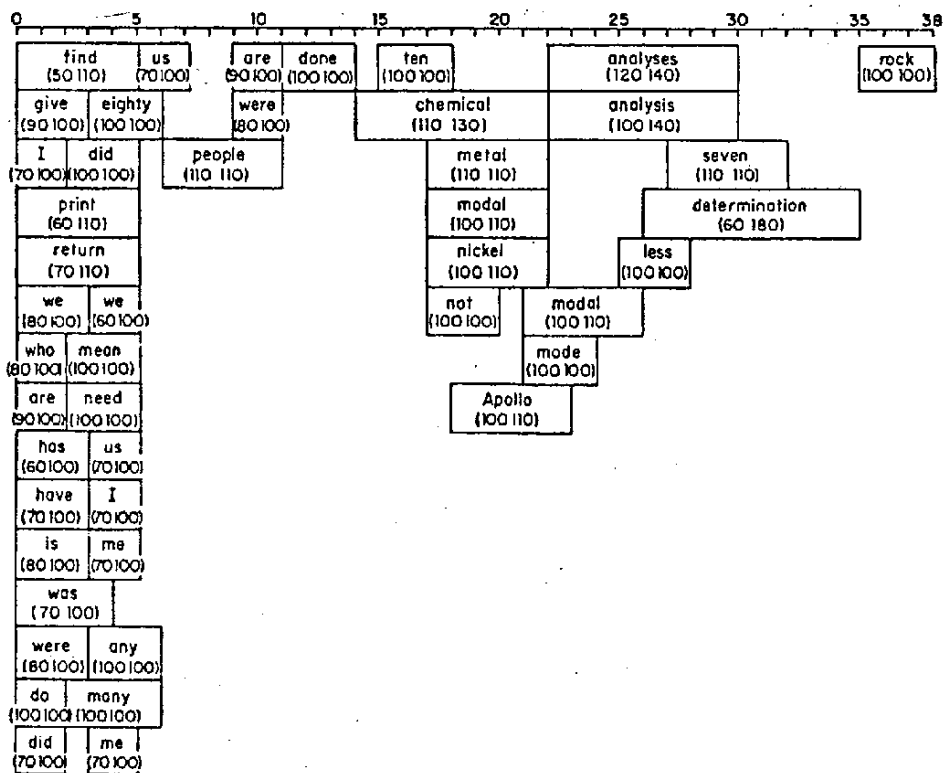
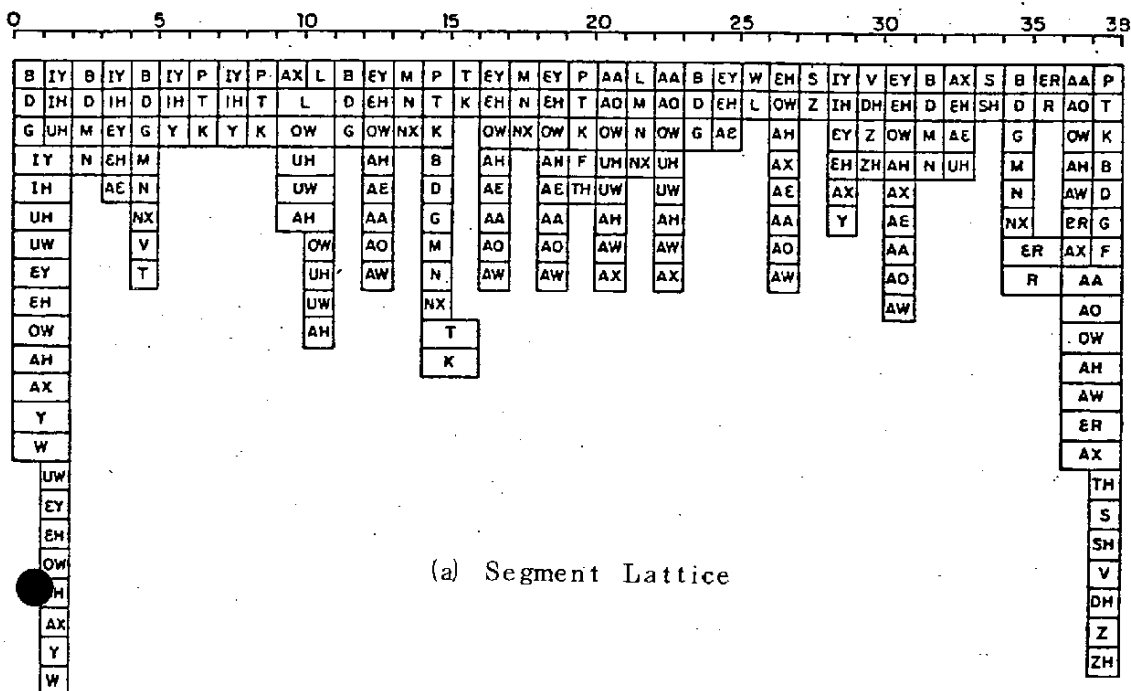


図 7-5 ラティス構造の例

(4) Hearsay I

このシステムは、Carnegie-Mellon 大学により開発されたもので、〔Reddy, Vicence, Neelly 69〕等による、500語の単語認識システムと、構文も考慮した、16単語の連続音声認識システムを音響分析レベルの基礎にし、又〔Gilligly 72〕のチェスのプログラム（ある状況下で、可能な駒の動きを、可能性の高い順に並べて出力する）を意味解析レベルの基礎にしたもので、31単語の語彙でチェスを指すことが初期の課題であった。

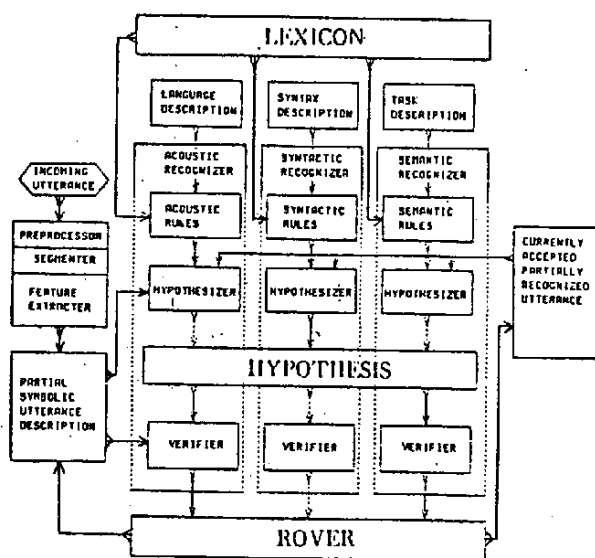


図 7-6 Hearsay I の認識プロセス

システムの構成を図 7-6 に示す。各部分の働きは、概ね次の通りである。

音響分析部では、10 msec 毎に次の 12 個の量を特徴ベクトルとして出力する。

- 1) フィルタを通さない原波形、2) 200Hz ~ 400Hz のフィルタ出力、
- 3) 400Hz ~ 800Hz のフィルタ出力、4) 800Hz ~ 1600Hz のフィルタ出力、
- 5) 1600Hz ~ 3200Hz のフィルタ出力、6) 3200Hz ~ 6400Hz のフィルタ出力の 6 チャンネルの出力各々につき、最大振幅と、

零交叉回数の2種の量を特徴とする。

各音韻に対して、複数のカテゴリ（標準パターン）を対応させ、話者変動や、背景雑音等の影響は、この部分である程度吸収する。つまり、標準パターンを話者毎に調整することはない。入力に対して、最小距離となるカテゴリを出力する。（nearest-neighbour 法の変形）

各フレーム毎に決定されたカテゴリ名を平滑化して、更に有声／無声の判定と摩擦音かどうかの判定結果を考慮に入れて、セグメンテーションを行なう。

この出力が認識部へ渡される。認識部は、音響、構文、意味の3つの部分と制御部に分かれている。認識制御は、ある状況（チェスの駒の配置）の下で、発話のある部分に関して、最も信頼のおける候補単語を提案できる部分を決めて、それを起動する。例えば、多くの場合は、チェスのプログラムに支えられている意味部が、可能な動きの中から、既にある候補単語を含むものを選び、その前後の単語を提案するが、音韻系列と辞書との整合度が高い単語を見つければ、音響部が提案する。次に残りの部分は、提案された単語に対して、自分の判断（検定）を行わなければいけない。こうして3つの部分全てが、その単語を受け入れると、次の新しい単語の仮説を立てる相に入る、という手順で1単語から始めて、候補単語系列（部分文）に1つずつ候補単語を加えて行き、最終的に発話全体を覆う単語の並びを決定する。

このシステムの特徴は、新しい候補単語を決定する段階では、音響、構文、意味の3つの部分が、相互の影響なしに独立に動けることである。この為に各部分を変更しても、他の部分を修正する必要がなく、課題を変更する事も比較的易しい。また、独立であるという事から、3つの部分が制御部によって起動される順番は、先見的には決っていない事も特徴の1つで、認識部全体としては、ある提案された仮説に対し、残りの部分が検定するという「仮説一検定」の方式を取っている事、この時、各部は候補単語に関する共通のデータベースを通して情報を交換する事などが特徴となる。

このシステムの性能が表7-4に示されている。これから、課題がチェスの他にニュースと医療に関してと卓上計算機の3つがあるが、それらには意味部分はなかった。実験例が少ないので十分な評価は難しいが、満足のゆくも

のではない様だ。

表 7-4 Hearsay I の実験結果 (1973 年秋)

1 DATA SET: TASK/ SPEAKER	2 WORDS IN LEX	3 %SEN	4 %COR	5 ACOUSTIC MODE			6 ACOUSTIC AND SYNTAX MODE				7 ACOUSTIC, SYNTAX, AND SEMANTIC MODE			
				X1N #1	X1N #1	X1N #1	XSEN #1	XWOR #1	AVE T/SEN	AVE T/S	XSEN MISS	XWOR MISS	AVE T/SEN	AVE T/S
1 CHESS/RM	31	14	62	61	78	85	63	87	11	8	100	100	8	8
2 CHESS/JB	31	19	66	44	89	78	74	93	12	8	100	100	8	8
3 CHESS/JB	31	21	105	39	88	73	15	89	15	8	48	88	13	7
4 CHESS/ITEL) BL	31	25	93	43	84	76	52	78	7	8	69	88	7	8
SUB-TOTALS		79	352	45	69	77	46	81	11	7	73	93	8	8
5 NEWS-RETRV/VL	28	19	104	38	53	67	21	83	13	7	XXX	XXX	XXX	XXX
6 MED-QUES/JA	26	21	92	28	38	49	24	48	13	11	XXX	XXX	XXX	XXX
7 DESK-CALC/JB	38	25	128	42	83	72	35	88	28	14	XXX	XXX	XXX	XXX
TOTALS		144	576	39	81	71	38	78	15	8	79	93	9	8

(5) DRAGON

Carnegie-Mellon 大学ではもう 1 つのシステムを開発している。DRAGON システムは Hearsay I と同じチェスを課題としているが、全体の構成が非常に簡単になっている。システム全体を、音響→音韻、音韻→単語、単語→文章、の 3 つのレベルに分けて表現し、各レベルを組合せて、音響→文章、への表現を作る。ここで「→」の意味するのは、「→」の左辺の変数が決った時に、右辺の変数がある値を取る確率が計算できる、ということである。

例えば、内部状態としての音韻系列 $X(t)$ が

$$\Pr (X(t+1) = j | X(t) = i) = a_{ij}$$

で表わされる様な、遷移行列 (a_{ij}) を持つ様な、マルコフ過程であると仮定し、観測される外部出力としての音響パラメータ系列 $Y(t)$ を

$$\Pr(Y(t) = k \mid X(t-1) = i, X(t) = j) = b_{ijk}$$

つまり1つ前の音韻と現在の音韻を与えられると、現在のパラメータが k となる確率が既に b_{ijk} で与えられているものとする。

この時、マルコフ過程の定義から、

$$\begin{aligned} & \Pr(X(t) = x(t) \mid X(i) = x(i) \ (i = 1, \dots, t-1)) \\ &= \Pr(X(t) = x(t) \mid X(t-1) = x(t-1)) \\ &= a_{x(t-1), x(t)} \end{aligned}$$

また

$$\begin{aligned} & \Pr(Y(i) = y(i) \ (i=1, \dots, t), X(j)=x(j) \ (j=1, \dots, t)) \\ &= \Pr(X(j) = x(j) \ (j=1, \dots, t)) \end{aligned}$$

$$\begin{aligned} & \cdot \prod_{i=1}^t \Pr(Y(i)=y(i) \mid X(i-1)=x(i-1), X(i)=x(i)) \\ &= \prod_{j=1}^t \Pr(X(j)=x(j) \mid X(j-1)=x(j-1)) \\ & \cdot \prod_{i=1}^t \Pr(Y(i)=y(i) \mid X(i-1)=x(i-1), X(i)=x(i)) \\ &= \prod_{i=1}^t a_{x(i-1)x(i)} \quad b_{x(i-1)x(i)y(i)} \end{aligned}$$

(ただし $x(0), a_{x(0)j} \quad b_{x(0)jk}$ は適当に与える)

となる。次に音響パラメータ $y(i) \ (i=1, \dots, t)$ が与えられた時に、音韻 $X(t)=j$ である確率を計算する為、特定の時刻 i に特定の状態(音韻) j である、全ての部分系列 $x(1) \dots x(t)$ の確率を加えたもの $\alpha(t, j)$ を定義する。

$$\begin{aligned} \alpha(t, j) &\equiv \Pr(X(t)=j, Y(i)=y(i) \ (i=1, \dots, t)) \\ &= \sum_{x(1), \dots, x(t-1)} \prod_{i=1}^t a_{x(i-1)x(i)} b_{x(i-1)x(i)y(i)} \end{aligned}$$

これから

$$\alpha(t, j) = \sum_i \alpha(t-1, i) a_{ij} b_{ij} y(t)$$

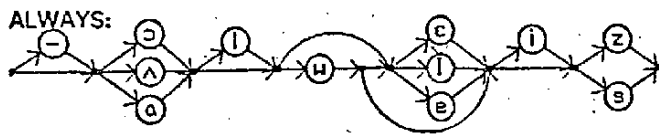
が成立する。さて終りに

$$\begin{aligned}
& \Pr (X(t)=j, Y(i)=y(i) \ (i=1, \dots, t)) \\
&= \frac{\Pr (X(t)=j, Y(i)=y(i) \ (i=1, \dots, t))}{\Pr (Y(i)=y(i) \ (i=1, \dots, t))} \\
&= \frac{\alpha (t, j)}{\sum_i \alpha (t, i)}
\end{aligned}$$

となり、音響パラメータ系列が与えられている時、各音韻が生ずる確率が計算できる。

この方式を下のレベルのものを上のレベルに埋め込んで行けば、各音響パラメータ列に対して、その時に生ずる文の確率が計算でき、その確率最大の文を、発話された文と考える。

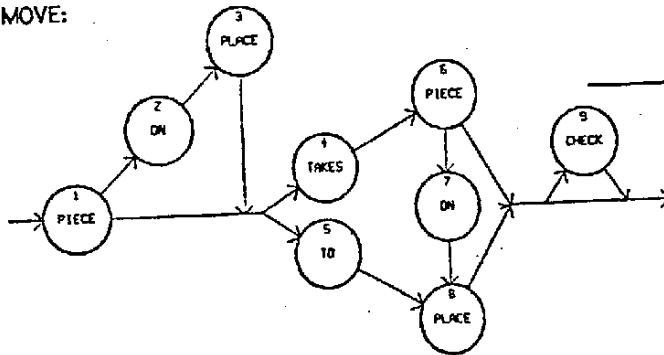
図 7-7 (a), (b), (c)は、マルコフ過程の遷移の状況を、network で表現したもので、図 7-7 (a)は、always という単語の音韻の遷移についての network で、図 7-7 (b)は、構文及び意味情報を用いた、文のレベルでの、構文的な情報の遷移 network で、図 7-7 (c)がそれらの subnetwork である。(a)が(c)に埋込まれ、(c)が(b)に埋め込まれることにより、音韻→文章が実現される。



PIECE:

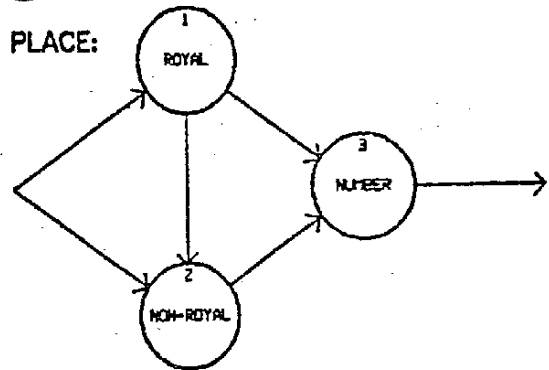
(a) 音韻レベル

MOVE:

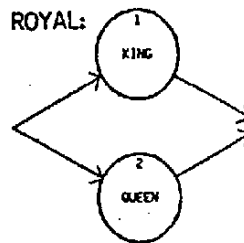


(b) 構文レベル

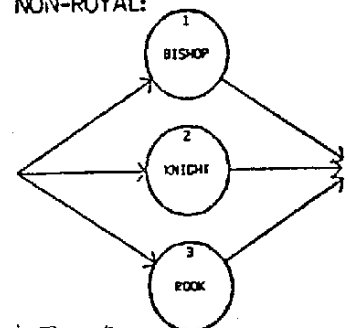
PLACE:



ROYAL:



NON-ROYAL:



(c) 構文レベルのサブネットワーク

図 7-7 DRAGON システムのネットワークの階層構造

7.3 ARPAの音声理解研究

Newellを議長とする9人からなる研究グループの報告を受けて、ARPAはアメリカ国内の大学・研究所に対し、1971年から5ヶ年（前期・後期各2年半）契約で資金援助（毎年約300万ドル）を行うことになった。主契約先はBBN（Bolt, Beranek and Newman Inc.）, CMU（Carnegie-Mellon Univ.）, Lincol Lab., SDC（System Development Corp.）, SRI（Stanford Research Institute）の5研究機関であり、補助契約先はSCRL（Speech Communications Research Lab.）, Speech Communication Lab., UNIVACの3研究機関である。

前期の契約が終了した1973年11月に各研究機関のデモンストレーションが行われ、その研究内容は1974年4月にCMUにおける音声認識シンポジウムで報告された。この時点で研究の進んでいたCMU, BBNおよびSRIとSDCの3グループに、契約先が選択統合されて研究が継続された。1976年8月から9月にかけて各グループの最終デモンストレーションが行われた。その運営委員会は、「本計画はCMUのHARPYシステムが初期の目標を達成して終了した」と報告している〔Medress 77〕。

以下では、主に上述の3グループの研究の概要を紹介し、他のグループについても簡単に触れることにする。

7.3.1 プロジェクトの成果概要

ARPA計画において最終的な評価を行なった4つのシステムの性能を表7-5に示す。表の中で「平均分岐数」（構文解析の各時点で予測される単語数）は、各システムが対象としているタスクの複雑さを示す一つの尺度である。語彙はどのシステムでも1000語程度である。誤り率の上でARPAの要求を満たしているのはCMU（カーネギーメロン大学）のHARPYのみである。HARPYの達成した性能は表7-1の右側に示されている。各システムが対象としているタスクおよび入力音声の例を表7-6に示す。また3種のシステムにおける単語認識部の性能をまとめて表7-7に示す。表7-5, 表7-6のシステムその他についてその概要を以下に述べる〔Klatt 77〕。

表 7-5 1976 年 11 月におけ各システムの性能

シ ス テ ム	文章認識率	branching factor (平均値)
CMU Harpy	95%	33
CMU Hearsay-2	91, 74	33, 46
BBN Hwin	44	195
SDC	24	105

表 7-6 4 つのシステムのタスクと入力音声の例

グ ル ー プ	タ ス ク	入 力 音 声 の 例
SDC	船についての 情報検索	"How fast is the Theodore Roosevelt?"
BBN Hwin	旅行の予算計 画の問合せ	"What is the plane fare to Ottawa?"
CMU Hearsay II, Harpy	文 献 検 索	"How many articles on psychology are there?"

表 7-7 3 つのシステムにおける単語認識部の性能

	SDC		BBN		CMU	
	ACCEPT	REJECT	ACCEPT	REJECT	ACCEPT	REJECT
正しい仮定に対する結果	92%	8%	84%	16%	94%	6%
誤った仮定に対する結果	3%	97%	34%	66%	49%	51%
1 単語当りの仮定の数	165		10		40	

7.3.2 Systems Development Corporation

ARPAの前半では、SDC が音響処理部を、SRI (Stanford Research Institute) が言語処理部を担当して、後半には、両者を合体させてシステムを作る予定であったが、種々の事情から不可能となり、ほんの数週間だけ SRI とつないだシステムが実現していた。その後は SDC にシステム全体が移り、言語の部分は、SRI のものをほとんど使用している。

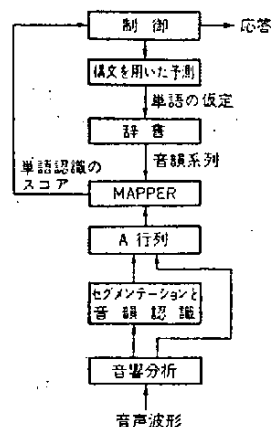


図 7-8 SDC システムの構成

最終的な SDC のシステムの構成を図 7-8 に示す。音響分析部では、20 KHz でサンプリングした波形を入力すると、10msec 毎に次のパラメータが出力される。

- 1) エネルギー
- 2) 零交叉回数
- 3) 基本周波数
- 4) 24 項の LPC を用いて求める Formant 4 つ

このパラメータを用いて、フレーム毎に上位 3 つの候補カテゴリ名を点数と いっしょにつけ、これらを局所的な連続性を仮定して平滑化する。この結果を用いて概そのセグメント分けを行い、その結果は A-matrix と呼ばれるデータアレイにしまわれる。発話の処理は、左から右へ行なわれ、まず全ての可能な文頭語を生成することから始まる。次に制御部が仮定した単語の音素表現を辞書から取り出し、予想される音声学的な変形を考慮する。この結果は図 7-9 に表わされる様な音韻グラフになる。このグラフは 1 つずつ mapper へ送られ、まだわからない発話の中の現在の位置と、どの程度良く音響的に整合がとれているかを調べられる。mapper は単語の音節構造に従って、予想される母音や、隣接した子音に特有の異音が存在するかを A-matrix を調べて決定する。完全に整合が取れることはほとんど無いので、mapper は A-matrix の

情報を用いて期待する単語がその部分に存在している確率を計算する。制御部は、mapper の得点を基にして、次に続く単語又は部分的な文を決定し、次に現われる可能性のある全単語のリストを作る。これをくり返すことにより、発話を解釈する。

このシステムの最大の欠点は、mapper が一旦辞書との整合度が悪い単語を捨てると、2 度と回復できない点にある。しかし、SRI の言語処理部が使えなくなってから SDC がその部分を設計し直す時間はあまりなかったのもしかたないことであろう。

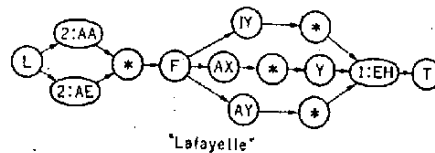


図 7-9 SDC のシステムの辞書表現の例

7.3.3 Bolt Beranek and Newman Inc.

BBN では、Speechlis の開発の延長線上に、HWIM (Hear What I Mean) システムの開発を計画した。これは、Speechlis の課題を、旅行運賃の問い合わせに変更し、語彙を 1000 語程度に増したもので、System 構成に関する基本的な態度はほとんど変っていない。システムの構成を図 7-10 に示す。

Speechlis との主な相違は、単語ラティスを、lexical decoding network (図 7-11(a)) による評価に置き換え、単語内と単語間の音声学的な音韻書き換え規則を network 中に埋め込んだことと、単語の認識 (確認) 部に、A-b-S 法により生成される音響パラメータを用いた、パラメータレベルでの DP マッチングを取り入れたこと、そして、Speechlis では意味解析部に Semantic Network 等が用いられていたが、それを ATN (Augmented Transition Network) 中に埋め込み、構文解析部と混在させたことである。(図 7-11(b)) このことにより上部処理の各部が、更に密接に働くことができるようになった。

システムの動作の概要は次の様である。音声波形は20KHzでサンプリングされ、音響処理部では、フォルマント周波数を推定する時、5KHzまでのSelective LPCを用いる以外はSpeechlisと同様にして、セグメントラティスを作る。

次に制御部が、lexical decoding networkを通してセグメントラティス上で最も良く一致する単語を探し（芽と呼ぶ）、単語確認部で、パラメータレベルの一致度を見ながら、点数が高ければ、構文予測部へ送られ、ATNを用いて文法上並びに意味上の制約を与えて、芽の単語の左右に現われる可能性のある単語を予測して、提案する。ここで提案された単語は、lexical decoding networkへ送られ、セグメントラティス上の探索が行なわれ、初めの状態に戻る。こうして単語列が少しずつ長くなって行き、発話全体を覆えば、理解したことになる。

BBNのシステムはSpeechlisの世代から、個々の構成部分に対しては、大変意欲的で注目すべき点が多かった。しかしシステム全体としての評価が低くなったのは、各部の仕事（計算）量が多く全体の処理時間が長く（実時間の1000倍程）なり過ぎ、システム全体の調整が十分に行なわれなかった事が一つの原因であろう。また、ブランディングファクタを比べると他のシステムに比べかなり多く、課題自体が難しかったこともある。残る課題は、システム全体がハード的に十分速かった時に、全体の調整を十分にやれば、どの位性能が上がったか、又課題の難易度の影響と音響分析のパラメータの使い方等の影響を明らかにする事であろう。

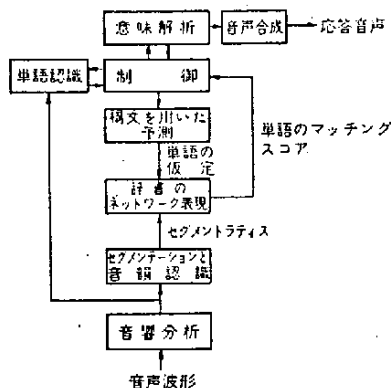
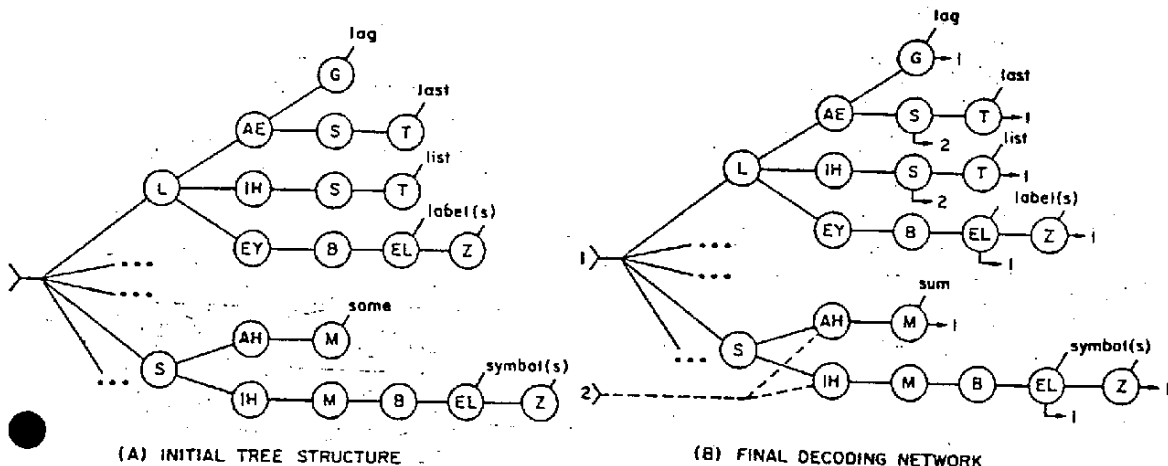


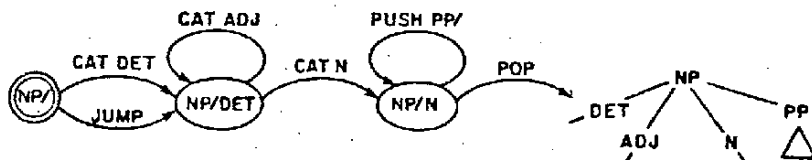
図 7-10 BBNのHwimのシステムの構成



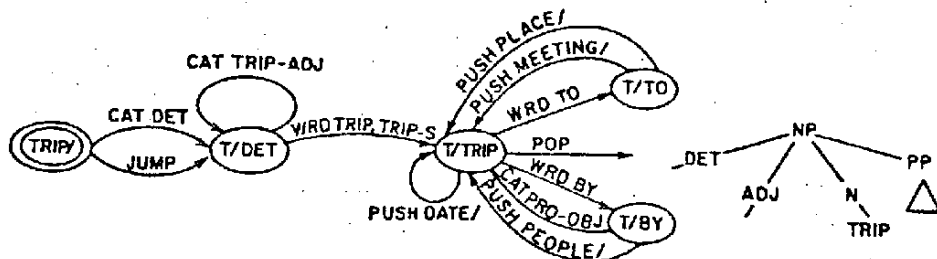
(A) 初期のネットワーク

(B) 単語間の音韻変化を考慮した
最終的なネットワーク

(a) Lexical decoding network の例



(A) 構文情報だけの例



(B) (A)に意味情報を埋め込んだA T N

(b) A T N の例

7.3.4 CMU-Hearsay II^{*}

Hearsay II システムは、Hearsay I の拡張版で主な改良点は次の3点である。

- 1) Hearsay I では、各部が仮説に対する検定を行なうという考え方を取り入れていたが、共通のデータ表現は、単語レベルで行なわれた為、他のレベルでの情報（音素や概念のレベル）は、単語レベルの形に表現しなければならなかったため、不便であったし、又、不十分な表現しかできない場合もあった。このことから、Hearsay II では、各知識源の情報は、レベル間でも関連づけられる様に、情報間にリンクを設けて、多層の表現を可能にした。
- 2) Hearsay I では、各仮説に対して、全ての部分の検定が終らなければ、次の仮説を提案する相に入れなかったが、これでは早くその仮説を棄てた部分があっても全体の計算が終るまで待っていなければいけない。そこで、Hearsay II では、その様な場合には、他の部分を止めて、次の相へ移れる様にした。

つまり各知識源は、起動に必要なデータが、「黒板」と呼ばれる、共通のデータベース上に総て存在することを確認すると、直ちにその知識源を駆動する。そして知識源は評価結果を「黒板」上に書いて評価を終る。制御部は常にこの「黒板」を監視していて、どの知識源が起動可能かを判断している。

- 3) Hearsay I では、棄てられた仮説に関する情報は全て捨てられていたため、他のその仮説と関連のある仮説が提案された場合にも、同じ様な評価を行なっていて、計算の能率が悪かった。これを避ける為に、Hearsay II では、棄却された仮説とその原因から、関連のある単語が、仮説に立てる必要が無くなったかどうかを判断できる様に、原因を調べて憶えておき、又各単語間の意味・構文上の関係を明らかにする様にした。

さて Hearsay II のシステムの構成を図 7-12 に示す。動作はある意味で BBN の HWIM と似た所があるが、システム構成上の基本思想が異っている。

音声波形は、音響分析部（Hearsay I と同じ）を経て、セグメントに分割され、

^{*}) Carnegie Mellon University

セグメントは音韻カテゴリに分類される(98種のカテゴリ)。これが「黒板」上のセグメント部に書き込まれると、各知識源が働き出し、各レベルでの仮説を作る。この部分の制御はデータ駆動形で、非同期に行なわれる。ここには、単語仮説提案部、単語検定部、構文解析部、構文予測部があり、最初に核となる単語を見つけて、構文情報を用いてその左右を延ばして行く。

Hearsay II は、BBN のシステムに比べて同じ様な解析手順を取りながらも、良い性能を示している。その理由は、branching factor を小さくしたことと、計算能率を上げる様に努力した点にあると思われる。また、各知識源が独立している為、個別に改良することが容易であるが、BBN のシステムは、意味と構文情報が混然一体となっているので一部の改良が他の部分への修正を引き起こす可能性がある。

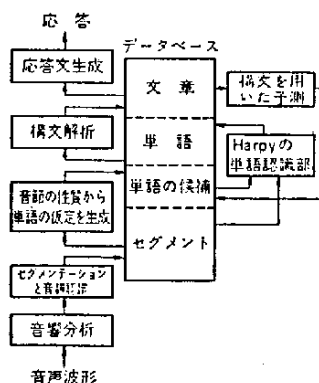


図 7-12 CMU の Hearsay II システムの構成

7.3.5 CMU-Harpy

Harpyの母体となったシステムは、Hearsay I と DRAGON であるが、システム構成は DRAGON の方式を受け継いでいる。DRAGON システムを改良した主な点は、1) 音響分析部に LPC 分析を採用した点、これにより各テンプレートとカテゴリの距離を最小予測誤差を用いて計算できるようになった。

2) DRAGON では各フレーム毎に状態を遷移させていたのに対し、セグメン

テーション（粗いものではあるが）を行ない遷移回数を減少させ、計算量を減らせた。3) network 中の探索において、DRAGON は全ての枝を breadth-first 法により探索していたが、beam search と呼ばれる、best-first と breadth-first を組み合わせた方法で、点数の低い枝は無視して、高い点数の周辺だけ並列に探す方法を取ったため、計算時間が激減した。この3点が主なものである。

図7-13がHarpyの構成である。音声波形は、10KHzでサンプリングされ、14項のLPC分析が行なわれる。そのパラメータの変化の程度の簡単な基準でセグメンテーションが行なわれて、15000の状態を持つnetworkに送られる。このnetworkは、1) 全ての可能な文章のグラフ表現、2) 音響セグメントに関する全辞書項目の表現、3) 単語境界における、音響セグメントの変化の規則、が全て埋め込まれている。

各入力セグメントは、98種のカテゴリと比較され、98個の最小予測残差を用いた距離が計算される。この距離の和が最小となる様なnetwork中の道を発話した文章と決定する。

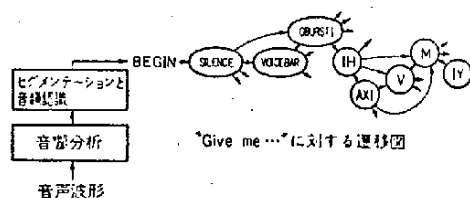


図7-13 CMUのHarpyシステムの構成
(状態遷移図の一部を示す)

7.3.6 その他の機関

(1) MITのLincoln Laboratory

Lincoln Laboratoryの研究者は、フォルマント追跡法や、破裂音の調音位置を分類する場合にフォルマント遷移の情報を用いる事など、音韻認識法の開発に貢献している。[Weinstein et al. 75] また、BBNのシス

テムに使われた lexical decoding network の基礎になった, lexical network 表現も開発している。〔Klovstad and Mondschein 65〕

(2) Stanford Reserch Institute

SRI では, SDC との協同でシステム開発を行なう予定であったが, 都合でそれが不可能になった時, 最後の数週間で, SDC の mapper を simulate する為に, 多くの入力に対する, mapper の出力を計算して残した。その後このデータを利用して, 与えられた語彙数に対して, かなり一般的な構文を許し, 独立した意味解析部と anaphoric reference や省略の情報を用いて, 発話の系列を, 文認識率 90% で認識できる様な条件を決定できる様になった。

(3) UNIVAC

UNIVAC の方式の特徴は, 韻律情報を積極的に利用している所である。このシステムでは, ピッチの時間パタンの概形を用いて句の切れ目や, 意味の切れ目, 又肯定か疑問か等がある程度判断できることを示している。更に, ストレスのある音節がどれであるかを検出するアルゴリズムを開発している。

〔Lea et al. 75〕

(4) IBM

IBM では ARPA とは独立して, 連続音声認識の研究をしている。初期には「音声タイプライター」を目標としていたこともあり, 音響-音韻変換に力を入れている。また最近では, DRAGON システムと同じ様に, マルコフ過程をモデルにした, DP マッチングによる, 全パスサーチを行っている〔Bahl et al. 78〕

7.4 ヨーロッパと日本の音声理解研究

7.4.1 ヨーロッパの音声理解研究

(1) フランスの音声理解研究

フランスでは CNET-Lannion と Nancy 大学等で 1974 年頃から音声理解に関する研究が進められている。ここでは Nancy 大学のシステムを簡単に紹介する。〔Haton 76〕

Haton と Pierrel のシステムは音響認識部, 構文解析部, 単語認識部 それに対話手続きから成っている。音響認識部は FFT または LPC (線形

予測符号化)により、入力音声进行分析してセグメンテーションを行ない、音素(程度の)系列に変換する。各時点で3つの音素候補を立てている。構文解析部は文法とすでに認識された単語にもとづいて、各時点で可能な単語の集合を予測するとともにシステム全体を制御する。Greibach文法にもとづいて上から下への処理を行なっている。対話手続は、結果にあいまいさが生じたときに、システムと使用者が対話することにより、正しい応答を出力するようにする方法である。構文解析部から音響分析部へのフィードバックはない。

タスクとしては電話交換業務を選んでこのシステムのテストをしている。文法規則を約20個、語彙40語としたとき、話者1名に対して95%の理解率が得られている(ただし、これは対話手続を含んでいる)。処理時間は文当たり0.8秒となっている。

(2) イタリアの音声理解研究〔De Mori 75〕

イタリアではトリノ大学のR. De Mori等が1974年頃から音声理解に関する研究を行なっている。彼らのシステムの特長は、(アメリカのシステムに比べると)音響処理部に重点が置かれていることである。音声信号は20KHzで標準化して、振幅・スペクトル・ピッチを抽出する。処理の最小単位としては音節またはそれより少し大きいものを考えている。一つの音響セグメントに対して、ホルマントの組を選択する仕方は複数あって、各々確率が割り当てられる。スペクトルとプロソディー(韻律)の特徴は、確率過程有限オートマンの推論で学習される。調音結合は入力について、すべての可能なホルマント軌跡を生成するパタン文法で記述されている。仮説は音節、単語または文が、入力文のスペクトルを生成し、構文構造がピッチコンターを生成できる確率にもとづいて評価される。文の記述は、入力文を再構成して意味が失われていないかどうかをチェックできる形で行われている。

タスクとしては旅行情報と予約業務を選び、200語の語彙と簡単なネットワークで記述される構文・意味部門を持ったシステムを構成している。しかしそのテスト結果については述べられていない。

7.4.2 日本の音声理解研究

(1) 電々公社通研〔Shikano 78, Nakatsu 78〕

武蔵野通研では1974年から76年にわたって列車の座席予約(28駅181列車を対象)をタスクとして(語彙112), オンライン会話音声認識システムの開発が行われた。認識系は音響処理部, 言語処理部, 音声応答部から構成されている。音響処理部ではLPC分析を行ない, VCV(母音・子音・母音)音節単位によるパターンマッチングと音声学的知识を利用したtop-down的音韻認識法, 母音標準パタンの学習による話者への適応などが特徴として上げられる。処理結果は, セグメント化のあいまいさと音韻認識のあいまいさを含む音韻ラティスの形に表現した。この音韻ラティスに正しい音韻が含まれる割合は78.6%であり, 音韻ラティスの持つ情報量から計算した音韻認識率は58.5%となっている。言語処理部はtop-down的に構成され, マッチングは左から右への順序で行われる。構文解析手順にはdepth-first法を用いている。文節ごとに区切って発声するという制限つきであるが, 実時間の5倍程度の処理時間で86%の文節認識率を得ている。これは男性8名が計算機室で40づつ計320の予約をした場合で, 予約の99%が正しく処理されている。このときの質問応答回数は(初回を除いて)平均3.2回である。発声者ごとの音韻認識率と文節認識率の関係から, 95%以上の文節認識率を得るには, 63%以上の音韻認識率が必要であると推定している。

(2) 京都大学〔Nakagawa 78, Sakai 77〕

京都大学では100単語の語彙から成る「計算機網への状態問合せと指令」をタスクとするシステムLITHANが開発された。このシステムは音響処理部, 音韻識別部, 単語同定部, 単語予測部, 解析指示部から構成されている。構文の記述には拡張文脈自由文法を用いている。解析指示部では, 述語が末尾に来るという日本語単文の性質と取り扱う世界の仕事の内容を利用する。単語予測部では文法だけでなく, タスクが持つ事実関係の知識も利用している。この場合, 予測・同定される単語群は, 一般には複数個となることにより, 多数の単語列が形成される。文解析の際にはbest-firstとbreadth-first法の両者の利点を取り入れて, 信頼性の高い単語列を並行して解析

を進める探索法を用いている。

男性10名が普通で発声した200文(1文は平均10単語から成る)の中128文が正しく認識された。このときの単語認識率は93%となっている。

このシステムは文法と辞書を変更するだけで他の新しいタスクに適用できる柔軟性をもっており、実際に「算術文」や「カレンダー」というタスクにも適用された。また話者の個人差に対処するための教師なし学習機能をもっている。

(3) 京都工繊大〔Niimi 77〕

京都工繊大ではプログラミング言語 BASIC に若干の制限を加えたものをタスク(語彙50)とするシステムが開発された。音響分析には1/4オクターブフィルタ20個を用いている。意味分析ではステートメントに付けられた行番号、キーワードの出現順序、変数の定義などの情報を抜かっている。解析は左から右に行われ、depth-first 法と breadth-first法の両者が、文認識率の面から比較検討されている。男性4名の発声した142文の中116文(81.7%)が正しく認識され、19文(13.4%)が誤認識され、残りはリジェクトされた。

また言語情報による単語予測能力を調べた結果、全辞書項目との照合と比較して、構文情報のみによる予測では約70%、これと意味情報とを併用した場合は約80%の労力を減少させ得ることが分った。一方、文認識率については言語情報を用いることにより約1.0%改善されることがシミュレーションにより分った。

(4) 山梨大学〔Sekiguchi 78〕

山梨大学では、最初単語ごとに区切って発声されたフォートランプログラムの音声認識の研究を進めていたが、その後、連続して発声された日本語文(語彙99語)の認識システムを開発した。音響レベルの処理に重点が置かれていてパラメータとしては振幅、零交叉数、偏自己相関係数、スペクトル、ホルマント周波数等各音素群に適したものをそれぞれ用いている。言語レベルでの特徴の一つは、一度認識された単語系列候補をあらためて識別音韻系列とマッチングすることによって、単語間にわたる調音結合の影響を再評価

している点である。男性4名の発声した36文の中28文(77.8%)が正しく認識されている。

7.5 音声理解研究の貢献

7.5.1 音声認識技術開発の現状〔Niimi 79, Klatt 77〕

音声理解システムにおいて、連続的な音声信号を離散的な言語記号に変換する過程の各レベルにおけるおおよその技術開発の現状を以下に述べる。しかし、これらは必ずしも音声理解研究ばかりでなく、音声研究全般の成果を含んでいる。

(1) 音響レベル

- 1) 線形予測法(LPC)の利用が普及した。
- 2) 調音領域での分析(声道断面積の推定)が進歩した〔Nakajima 78〕。
- 3) 調音結合の研究が必要である〔Lea 78〕。

(2) 音声学・音韻レベル

- 1) 対象が限られたタスクである場合には、音声波の音素系列への変換は必ずしも必要でない(Harpyの成功の例)。
- 2) 60個程度の音素カテゴリーに対して60~70%の識別率が得られている。
- 3) 個人差正規化のため、音素標準パターンを話者に適応させる手法が使われるようになった。
- 4) 音素識別のアルゴリズムは静的で、前後の影響を考慮したものは少ない。

(3) 単語レベル

- 1) 辞書中の単語の記述は、音形規則を適用したネットワーク形式が一般的となった。
- 2) ネットワーク表現による辞書項目とセグメントラティスの照合アルゴリズムが確立されたが、類似性の定量化には多くの問題が残っている。

(4) 韻律レベル

基本周波数や振幅包絡を用いた音節の識別、構文境界の検出などが試みられている程度である〔Lea 75〕。

(5) 構文レベル

- 1) 構文的制約の利用の有効性が定量的に評価された〔Nakagawa 76, Niimi 78, Erman 74〕。
- 2) 構文の記述と解析には拡張遷移網文法が便利である〔Woods 70〕。

(6) 意味・プラグマティックスレベル

- 1) セマンティックネットの使用が一般化した。しかしこのレベルの制約は構文記述の中に埋め込まれるのが普通である。
- 2) 複数の文の文脈情報を有効に利用するまでには至っていない。

(7) 制御レベル

- 1) システム構成：ネットワークモデルが提案され成功をおさめた。階層モデルの使用が一般化した。が、大規模システムでは成功していない。
- 2) 処理方向：英語では左から右へ順に処理する方法が有効である。
- 3) 仮説の評価と選択：best-first, breadth-first の他にビーム探索法が導入されたが、各方式の優劣を系統的に論じた研究は少ない。

(8) その他の問題

- 1) タスクの複雑さの尺度として、平均分岐数や等価語彙数が用いられるようになった。
- 2) 認識と理解について、文法の作り方によって、認識率が悪くても理解率（システムの応答）は良くなる例が示された（Hearsay II）。

7.5.2 音声知覚モデルとの関係〔Klatt 77〕

人間における音声知覚過程の心理学モデルとしてこれまで幾つかの理論やモデルが提案されている。音声理解研究は、これとは直接の関係はないが、その成果から次のようなことが推測される。

- (1) 単語の解読はフィードバックループを含まずに直接的に行われる。
- (2) 音響音声学および音韻論的知識をネットワーク構造に編集する過程で、言語の生成規則が利用される。
- (3) 文処理の過程で音声学的表現をすることは特に必要ない。
- (4) 構文解析もネットワーク等を利用することによって直接的に行われる。

7.5.3 音声認識技術の問題点〔Lea 78〕

- (1) 連続的な音声信号のセグメンテーション（区分）とラベリング。
- (2) 音韻規則（調音結合，単語境界の影響，音韻変動）の確立。
- (3) 韻律情報の構文解析，単語マッチング，音韻分析への利用法。
- (4) 意味論，談話分析，タスク等の拘束条件としての利用法。
- (5) 孤立単語認識と音声理解システムとの間の橋渡しが必要である。
 - 1) 認識能力はどの程度のものが必要か？
 - 2) タスクの難易の順序づけは可能か？
 - 3) 単語認識と音声理解の中間レベルのものは有用か？ もしそうならどんなものとなるか？
 - 4) HARPY 程度の学習訓練を要するシステムは使われ得るか？
 - 5) 殆ど学習を要しないシステムの開発は可能か？
- (6) 大量のデータを使用したシステムの試験評価が必要である。タスクの複雑さの測度とパフォーマンスの評価手続きの開発。
- (7) 変換器（マイクロホン），通信路（雑音），話者（男・女・子供・方言）の問題。HARPY 的システム（手動チューニング，標準パタン修正）の限界はどうか？
- (8) 新しい（未知の）単語の辞書への追加の仕方。

7.6 音声認識の実用化〔Chiba 78, Nakata 80〕

最近，特定話者を対象とした単語音声認識装置が相次いで商品化されている。これはその目的から大きく二つに分類することができる。一つは本格的な音声認識装置であり，性能が良い代わりに大型で価格も高い。もう一つはマイクロコンピュータの応用として家庭電気製品の音声制御をねらった簡易型のものであり，小型で価格も安くなっている。現在製品化されている音声認識装置はいずれも単語または音節を区切って発声したものを，使用者が予め登録した標準パタンとマッチングして認識するものである。ただし日本電気の製品だけは，連続単語を対象としている。また日立のHR150は不特定話者用となっている。

実用的な面からは，操作の確実性ということを考えると，利用者が長々と機

械に話しかけるよりも、単語音声に限定した方が、より適当と考えることもできる。しかしその場合でも、幾つかの単語が連続して入力できることが必要であろう。装置の小型化、語彙の増加、登録の簡単化、性能の向上、価格の低廉化が望まれる。現在市販されている代表的な音声認識装置を表7-8に示す。

表7-8 製品化された単語音声認識装置

メーカー	名 称	発 表	認識対象	時間正規化	登 録	認識率他	語 数
Threshold Technology	T500/600	1975	単 語	D P	～10回	99%以上	64
Interstate Electronics	VDES	"	"	非線形	～5	4チャンネルまで	250
高千穂交易	MIKE	"	"	—	3	—	16
日本電気	DP-100	1978	連続単語	D P	1	99%以上 2チャンネルまで	60(120)
富士通	—	1979	数 字 力 英 ナ 語	—	有	数字 99% カナ } 70% 英字 } 80%	12語 68 26
三洋電機	ボイス ズバコン	"	単 語	線形	有	—	32
東 芝	ボイス コマンダー	"	"	—	有	—	15
日 立	HR 150	"	"	—	不特定 話 者	90～95%	16～ 128
シャープ	—	"	数 字 そ の 他	—	有	—	10語 5

7.7 おわりに

ARPAの音声理解(SUS)計画の大きな特徴は、その研究の中心が音声そのものと言うよりも、構文、意味、プラグマティックスなどの各種の知識源を有機的に統合した一つの音声理解システムを構成することにある。計画に参加した研究者を見ても、純粋に音声の分野の人は少なく、言語や人工知能など計算機科学方面に従事していた人が多かったということである。また、このような方向で研究を進める前提として、音声波そのものの中には、話者の意図したものを解読するための情報がすべて含まれているわけではないので、他の種類の知識の利用が不可欠であると考えていた。この考え方は誤りではないが、このことを強調する余り、音響、音声学レベルの研究にはそれ程力が注がれなかったくらいがある。この点がこの計画の弱点となった〔Niimi 79〕。しかしタスクを限定すれば文音声の認識も不可能ではないことを示した功績は大きい。現在までに得られた成果を有効に生かして、実用に供し得る連続音声認識システムを開発することも重要である。Leaらの調査結果は、孤立単語や連続数字の認識とともに、Harpy 程度の連続音声認識システムにもかなりの需要があることを示している〔Lea 78〕。

SUS計画全体として、言語処理レベルでは種々の手法が提案され、その有効性が示されたが、音響処理レベルではあまり大きな貢献はなされなかったと言えよう。今後の連続音声認識の研究は、もう一度、音響、音声学的レベルの見直しをする方向へ進むものと思われる。

参考文献

- [Alter 68] Alter, R., "Utilization of Contextual Constraints in Automatic Speech Recognition," IEEE. Trans. vol. AU-16, NO.1, pp.6-11 (Mar. 1968)
- [Bahl 76] Bahl, L., Baker, J., Cohen, P., Dixon, N., Jelenek, F., Mercer, R., and Silverman, H., "Preliminary Results on the Performance of a System for the Automatic Recognition of Continuous Speech", IEEE ICASSP'76, pp.425-429 (Apr. 1976).
- [Bahl 78a] Bahl, L.R., Baker, J.K., Cohen, P.S., Cole, A.G., Jelinek, F., Lewis, B.L. and Mercer, R.L., "Automatic Recognition of Continuously Spoken Sentences from a Finite State Grammar", Proc. of IEEE ICASSP'78, pp.418-421 (Apr. 1978)
- [Bahl 78b] Bahl, L.R., Baker, J.K., Cohen, P.S., Jelinek, F., Lewis, B.L. and Mercer, R.L., "Recognition of a Continuously Read Natural Corpus", Proc. of IEEE ICASSP'78, pp.422-424 (Apr. 1978)
- [Bahl 79] Bahl, L.R., Bakis, R., Cohen, P.S., Cole, A.G., Jelinek, F., Lewis, B.L. and Mercer, R.L., "Recognition Results with Several Experimental Acoustic Processors", IEEE ICASSP'79, pp.249-251 (Apr. 1979)
- [Baker 75] Baker, J., "The Dragon System - An Overview", IEEE Trans. Acoust. Speech Signal Process., vol. ASSP-23, No.1, pp.24-29. (Feb. 1975)
- [Baker 79] Baker, J.M., "Performance Statistics of the HEAR Acoustic Processor", IEEE ICASSP'79, pp.262-265 (Apr. 1979)
- [Barnett 73] Barnett, J.A., "A Vocal Data Management System", IEEE, Trans. AU, vol. AU-21, No.3, pp.185-188 (Jun. 1973)
- [Barnett 74] Barnett, J.A., "A Phonological Rule Compiler", pp.188-192 in Erman (Apr. 1974).
- [Bates 74] Bates, M., "The Use of Syntax in a Speech Understanding System", IEEE Symp. Speech Recognition, pp.226-233 (Apr. 1974)
- [Bates 75] Bates, M., "The Use of Syntax in a Speech Understanding System", IEEE Trans. Acoust. Speech Signal Process., vol. ASSP-23, No.1, pp.112-117. (Feb. 1975)
- [Beek 77] Beek, B., Neuberg, E.P., and Hodge, D.C., "An Assessment of the Technology of Automatic Speech Recognition for Military Applications", IEEE Trans. ASSP, vol. ASSP-25, No.4, pp.310-322 (Aug. 1977)
- [Chiba 78] Chiba, S., "Spoken Word Recognition", Jour. Information Processing, vol.19, No.7, pp.667-674 (July 1978) (in Japanese)

[Christiansen 79] Christiansen, R.W. and Rushforth, C.K., "Word Spotting in Continuous Speech Using Linear Predictive Coding", IEEE ICASSP'79, pp.557-560 (Apr. 1979)

[Cohen 75] Cohen, P.S., and Mercer, R.L., "The Phonological Component of an Automatic Speech Recognition System", pp.275-320 in 'Speech Recognition' edited by Reddy (Academic Pr.)(1975).

[Cook 77] Cook, C.C. and Schwartz, R.M., "Advanced Acoustic Techniques in Automatic Speech Understanding", IEEE ICASSP'77, pp.663-666 (May 1977).

[Cook 79] Cook, C., "Word Verification in a Speech Understanding System", IEEE ICASSP'79, pp.553-556 (Apr. 1979).

[Denes 59] Denes, P. B., "The Design and Operation of the Mechanical Speech Recognizer at University College London", Journal of Brit. IRE. vol.19, No. 4, pp.219-234 (1959)

[Diller 79] Diller, T.C., "Phonetic Word Verification", IEEE ICASSP'79, pp.256-261 (Apr. 1979)

[Erman 74] Erman, L.D., 'Contributed Papers of the IEEE Symposium on Speech Recognition' (IEEE Catalogu No.74CH0878-9 AE)(Apr. 1974)

[Fant 70] Fant, G., "Automatic Recognition and Speech Research", STL-QPSR 1, No.1, pp.16-31 (1970)

[Fujisaki 72] Fujisaki, H., "Current Problems in Speech Recognition", JASJ. vol. 28, NO.1, pp.33-41 (Jan. 1972) (in Japanese)

[Fujisaki 78] Fujisaki, H., "Current Status and Future of Speech Research — A Review", JASJ, vol.34, pp.117-121 (1978) (in Japanese)

[Gilligly 72] Gillogly, J.J., "The TECHNOLOGY chess program," Artif. Intel., vol.3, pp.145-163 (1972)

[Glave 78] Glave, R.D. and van der Giet, G., "The DAWID Speech Recognition System", IEEE ICASSP'78, pp.429-432 (Apr. 1978)

[Goodman 76] Goodman, G., "Analysis of Languages for Man-Machine Voice Communication", Ph.D dissertation (CMU), (1976)

[Halle 62] Halle, M. and Stevens, K., "Speech Recognition: A Model and a Program for Reserch", IRE Trans. of Inf. Theory IT-8, pp.155-159 (Feb. 1962).

[Haton 76] Haton, J.P., and Pierrel, J.M., "Organization and Operation of a Connected Speech Understanding System at Lexical, Syntactic and Semantic Level", IEEE ICASSP'76, pp.430-433 (Apr. 1976).

[Haton 79] Haton, J. and Morel, O., "Automatic Recognition of Connected Digits Sequences by Means of Segmentation and Dynamic Programing", IEEE ICASSP'79, pp.245-248 (Apr. 1979)

[Hayes-Roth 76a] Hayes-Roth, F., and Lesser, V.R., "Focus of Attention in a Distributed-Logic Speech Understanding System", IEEE ICASSP'76, pp.416-420 (Apr. 1976).

[Hayes-Roth 76b] Hayes-Roth, F. and Mostow, D.J., "Syntax and Semantics in a Distributed Speech Understanding System", IEEE ICASSP'76, pp.421-424 (Apr. 1976)

[Hyde 72] Hyde, S.R., "Automatic Speech Recognition: A Critical Survey and Discussion of the Literature", in 'Human Communication: A Unified View', edited by E.E. David and P.B. Denes (McGraw-Hill, New York) (1972)

[Jaschul 79] Jaschule, J., "An Approach to Speaker Normalization for Automatic Speech Recognition", IEEE ICASSP'79, pp.235-238 (Apr. 1979)

[Klatt 72] Klatt D.H. and Stevens K.N., "Sentence recognition from visual examination of spectrograms and machine-aided lexical searching," in Proc. Int. Conf. Speech Commun. Processing, pp.315-318 (1972)

[Klatt 77] Klatt, D.H., "Review of the ARPA Speech Understanding Project", JASA, vol.62, No.6, pp.1345-1366, (Dec. 1977)

[Klovstad 74] Klovstad, J.W. and Mondschein, L.F., "The Caspers Linguistic Analysis System", IEEE Symp. Speech Recognition, pp.234-240 (Apr. 1974)

[Klovstad 75] Klovstad, J.W. and Mondschein, L.F., "The CASPERS Linguistic Analysis System", IEEE Trans. ASSP, vol.ASSP-23, No. 1, pp.118-123 (Feb. 1975)

[Kohda 78] Kohda, M., Nakatsu, R., Shikano, K. and Itoh, K., "On-Line Question-Answering System by Conversational Speech", JASJ, vol.34, pp.194-203 (1978) (in Japanese)

[Kohda 79] Kohda, M., Nakatsu, R. and Shikano, K., "Speech Recognition in the Question-Answering System Operated by Conversational Speech", IEEE ICASSP'79, pp.442-445 (Apr. 1979)

[Lea 73] Lea, W.A., "An Approach to Syntactic Recognition Without Phonemics", IEEE Trans. AU, vol.AU-21, No.3, pp.249-258 (Jun. 1973)

- [Lea 75] Lea, W.A., Medress, M.F. and Skinner, T.E., "A Prosodically Guided Speech Understanding Strategy", IEEE Trans. ASSP, vol.ASSP-23, No.1, pp.30-38 (Feb. 1975)
- [Lea 78a] Lea, W.A. and Shoup, J.E., "Recommendations for Advancing Recognition", 95th Meeting of ASA, (May 1978)
- [Lea 78b] Lea, W.A. and Shoup, J.E., "Gaps in the Technology of Speech Understanding", Proc. of IEEE ICASSP'78, pp.405-408 (Apr. 1978)
- [Lesser 75] Lesser, V.R., Fennell, R.D., Erman, L.D. and Reddy, D.R., "Organization of the Hearsay-II Speech Understanding System", IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-23, No.1, pp.11-23.(Feb. 1975).
- [Levinson 79] Levinson, S.E. and Rosenberg, A.E., "A New System for Continuous Speech Recognition - Preliminary Results", IEEE ICASSP'79, pp.239-244 (Apr. 1979)
- [Lowerre 76] Lowerre, B. T., "The HARPY Speech Recognition System", CMU WPSR (Apr. 1976)
- [Mariani 79] Mariani, J.J., Lienard, J.S. and Renard, G., "Speech Recognition in the Context of Two-Way Immediate Person-Machine Interaction", IEEE ICASSP'79, pp.269-272 (Apr. 1979)
- [Mercier 78] Mercier, G., Quinton, P. and Vives, R., "Dialogue Homme - Machine avec KEAL", Reserches/Acoustique CNET, vol.4, pp.187-206 (1977)
- [Miller 73] Miller, P.L., "A Locally-Organized Parser for Spoken Input", Ph.D thesis, MIT, (Mar. 1973)
- [Moor 77] Moor, R.K., "Evaluating Speech Recognizers", IEEE Trans. ASSP, vol.ASSP-25, No.2, pp.178-183 (Apr. 1977)
- [Nakagawa 76] Nakagawa, S., "A Machine Understanding System for Spoken Japanese Sentences", Ph.D thesis Kyoto Univ., (1976)
- [Nakagawa 78] Nakagawa, S., "Reserch Trends in Speech Recognition and Understanding by Machine", Journal of Information Processing, vol.19 pp.675-685(1978)(in Japanese).
- [Nakajima 78] Nakajima, T., Suzuki, T. et al., "Estimation of Vocal Tract Area Function by Adaptive Deconvolution and Adaptive Spech Analysis System" JASJ. vol.34, No.3, pp.157-166 (Mar. 1978) (in Japanese)
- [Nakata 80] Nakata, K., "Speech Input-Output Apparatus and Its Application", Jour. Soc. Instrument and Control Engrn. Japan, vol.19, No.3, pp.31-33 (Mar. 1980) (in Japanese)

[Nakatsu 78] Nakatsu, R. and Kohda, M., "An Acoustic Processor in a Conversational Speech Recognition System", Journal of IECEJ, vol.J61D, pp.261-268 (1978) (in Japanese)

[Nakatsu 79] Nakatsu, R., "Review of the ARPA Speech Understanding Project" Journal of IECEJ, vol.62, No.3, pp.334-337 (Mar. 1979) (in Japanese)

[Nash-Webber 74] Nash-Webber, B., "Semantic Support for a Speech Understanding System", IEEE Symp. Speech Recognition, pp.244-249 (Apr. 1974)

[Nash-Webber 75a] Nash-Webber, B., "Semantic Support for a Speech Understanding System", IEEE Trans. ASSP, vol.ASSP-23, No.1, pp.124-129 (Feb. 1975)

[Nash-Webber 75b] Nash-Webber, B., "The Role of Semantics in Automatic Speech Understanding", Representation and Understanding (Studies in Cognitive Science) Academic Press, pp.351-382 (1975)

[Newell 71] Newell, A., Barnett, J., Forgie, J., Green, C., Klatt, D., Licklider, J.C.R., Munson, J., Reddy, R. and Woods, W., "Speech-Understanding Systems : Final Report of a Study Group", (May 1971)

[Newell 75] Newell, A., Cooper, F.S., Forgie, J.W., Green, C.C., Klatt, D.H., Medress, M.F., Neuburg, E.P., O'Malley, M.H., Reddy, D.R., Ritea, B., Shoup, J.E., Walker, D.E. and Woods, W.A., "Consideration for a Follow-on ARPA Research Program for Speech Understanding Systems", (Aug. 1975)

[Newell 78] Newell, A., "HARPY Production Systems and Human Cognition", CMU WPSR (Sep. 1978)

[Niimi 75] Niimi, Y. and Asami, T., "The Use and Effects of Linguistic Knowledges in a Speech Recognition System", Journal of IECEJ, vol.58D, pp.741-747 (1975) (in Japanese)

[Niimi 77] Niimi, Y., Kobayashi, Y., Asami, T. and Miki, Y., "The Speech Recognition System of 'Spoken BASIC 1'", Journal of Information Processing, vol.18, No. 5, pp.453-459 (1977) (in Japanese)

[Niimi 78] Niimi, Y. and Kobayashi, Y., "A Voice Input Programing System Using BASIC-like Language", IEEE ICASSP'78, pp.425-428 (Apr. 1978)

[Niimi 79] Niimi, Y. : Speech Recognition, Kyoritsu Pub., Tokyo (1979) (in Japanese)

[O'Malley 73] O'Malley, M.H., Klover, D.R. and Dara-Abrams, B., "Recovering Parentheses from Spoken Algebraic Expressions", IEEE, Trans. AU, vol. AU-21, No.3, pp.217-225 (Jun. 1973)

[Oshika 75] Oshika, B., Zue, V.W., Weeks, R.V., Nue, H. and Aurbach, J., "The Role of Phonological Rules in Speech Understanding Research", IEEE Trans. Acoust. Speech Signal Process., vol.ASSP-23, No.1, pp.104-112(Feb. 1975).

[Paxton 76] Paxton, W.H., 'A Framework for Speech Understanding', Ph.D. Dissertation (Stanford University)(1976).

[Quinton 76] Quinton, P., "A Syntactic Analyzer Adapted to Speech Recognition", IEEE ICASSP'76, pp.453-456 (Apr. 1976)

[Reddy 71] Reddy, D. R., "Speech Recognition: Prospects for the Seventies", Proc. IFIP'71, pp.I-5-I-13 (1971)

[Reddy 72] Reddy, R. and Erman, R. and Neely, R. et al. "Working Papers in Speech Recognition -- I --", WPSR CMU, (Apr. 1972)

[Reddy 73] Reddy, D.R., Erman, L.D. and Neely, R.B., "A Model and a System for Machine Recognition of Speech", IEEE Trans. AU, vol.AU-21, No.3, pp.229-238 (Jun. 1973)

[Reddy 76] Reddy, D.R., "Speech Recognition by Machine: A Review", Proc.IEEE, pp.501-531 (Apr. 1976)

[Reddy 77] Reddy, R., et al., "Speech Understanding Systems : Summary of Results of the Five-Year Reserch Effort at Carnegie-Mellon University", CMU, (1977)

[Saitta 78] Saitta, L., "Fuzzy Semantic Network for a Speech Understanding System - An Experimental Study -", IEEE ICASSP'78, pp.433-436 (Apr. 1978)

[Sakai 63] Sakai, T. and Doshita, S., "Speech Recognition of Conversational Sounds", Jour. IECEJ, vol.46, pp.1696-1704 (Nov. 1963) (in Japanese)

[Sakai 76] Sakai, T. and Nakagawa, S., "A Classification Method of Spoken Words in Continuous Speech for Many Speakers", Journal of Information Processing, vol.17, pp.650-658 (1976) (in Japanese)

[Sakai 77] Sakai, T. and Nakagawa, S., "A Speech Understanding System of Simple Japanese Sentences in a Task Domain", Trans. IECE Japan, vol.E60, pp.13-20 (1977) (in Japanese)

[Sekiguchi 77] Sekiguchi, Y., Oowa, H., Aoki, K. and Shigenaga, M., "Speech Recognition System for FORTRAN Program", Journal of Information Processing vol.18, No.5, pp.445-452 (1977) (in Japanese)

[Sekiguchi 78] Sekiguchi, Y. and Shigenaga, M., "Speech Recognition System for Japanese Sentences", JASJ, vol.34, pp.204-213 (1978) (in Japanese)

[Shikano 78] Shikano, K. and Kohda, M., "A Linguistic Processor in a Conversational Speech Recognition System", Journal of IECEJ, vol.J61D, pp.253-260 (1978) (in Japanese)

[Skinner 76] Skinner, T.E., Kloker, D.R. and Medress, M.F., "A Speech Recognition System for Connected Word Sequences", IEEE ICASSP'76, pp.434-437 (Apr. 1976)

[Smith 79] Smith, A.R., "Word Hypothesization in the HEARSAY-II Speech System", IEEE ICASSP'79, pp.549-552 (Apr. 1979)

[Sondhi 78] Sondhi, M.M. and Levinson, S.E., "Computing Relative Redundancy to Measure Grammatical Constraint in Speech Recognition Tasks", IEEE ICASSP'78 pp.409-412 (Apr. 1978)

[Takeya 76] Takeya, S. and Tamati, T., "On a Syntax Analysis Method for Speech Recognition Using a Dependency Grammar", Journal of IECEJ, vol.J59-D, No.7, pp.475-482 (1976) (in Japanese)

[Tappert 73] Tappert, C.C., Dixon, N.R. and Rabinowitz, A.S., "Application of Sequential Decoding for Converting Phonetic to Graphic Representation in Automatic Recognition of Continuous Speech (ARCS)", IEEE Trans. AU, vol.AU-21, No.3, pp.225-228 (Jun. 1973)

[Tappert 74] Tappert, C.C. and Dixon, N.R., "A Procedure for Adaptive Control of the Interaction Between Acoustic Classification and Linguistic Decoding in Automatic Recognition of Continuous Speech", Artificial Intelligence, vol.5, pp.95-113 (1974)

[Vicens 69] Vicens, P., "Aspects of Speech Recognition by Computer", Stanford Univ. CS 127 MEMO AI-85, (Apr. 1969)

[Vintsiuk 76] Vintsiuk, T.K., Gavriljuk, O.N. and Shinkazh, A.G., "Phoneme-by-Phoneme Recognition of Speech Composed of the Words of Given Vocabulary", IEEE ICASSP'76, pp.450-452 (Apr. 1976)

[Vintsiuk 76] Vintsiuk, T.K., "Generative Grammars and Dynamic Programing in Speech Recognition with Learning", IEEE ICASSP'76, pp.446-449 (Apr. 1976)

[Walker 75] Walker, D.E., "The SRI Speech Understanding System", IEEE Trans. ASSP, vol.ASSP-23, No.5, pp.397-416 (Oct. 1975)

[Weisen 74] Weisen, R.A. and Forgie, J.W., "An Evaluation of the Lincoln Laboratory Speech Recognition System", J. Acoust. Soc. Am., vol.56, S27(A) (1974).

[White 78] White, G.M., "Dynamic Programing, The VITERBI Algorithm, and Low Cost Speech Recognition", IEEE ICASSP'78, pp.413-417 (Apr. 1978)

[Woods 70] Woods, W. A., "Transition Network Grammars for a Natural Language Analysis", Commun.ACM. vol.13, No.10 (1970)

[Woods 74] Woods, W.A. and Makhoul, J., "Mechanical Inference Problems in Continuous Speech Understanding", Artificial Intelligence, vol.5, pp.73-91 (1974)

[Woods 75] Woods, W.A., "Motivation and Overview of SPEECH-LIS: An Experimental Prototype for Speech Understanding Research", IEEE Trans. ASSP, vol.ASSP-23, No.1, pp.2-10 (Feb. 1975)

[Woods 76] Woods, W.A., Bates, M., Brown, G., Bruce, B., Klovstad, J.W. and Nash-Webber, B., "Uses of Higher Level Knowledge in a Speech Understanding System : A Progress Report", IEEE ICASSP'76, pp.438-441 (Apr. 1976)

[Woods 79] Woods, W.A. and Zue, V.W., "Dictionary Expansion via Phonological Rules for a Speech Understanding System", IEEE ICASSP'79, pp.561-564 (Apr. 1979)

[de Figueiredo 79] de Figueiredo, R.J.P. and Brzustowicz, J., "Techniques for Spectrogram Patterns Based on Dynamic Modeling", IEEE ICASSP'79, pp.266-268 (Apr. 1979)

[de Mori 75] de Mori, R. Rivoira, S. and Serra, S., "A Speech Understanding System with Learning Capability", Proc. 4th IJCAI, Tbilisi, USSR, pp.468-475 (1975)

[de Mori 78] de Mori, R., "Recent Advances in Automatic Speech Recognition", IJCPR'78, pp.106-124 (Nov. 1978)

[de Mori 79] de Mori, R. and Torasso, P., "Lexical Classification in a Speech Understanding System Using Fuzzy Relations", IEEE ICASSP'79, pp.565-568 (Apr. 1979)

[de Mori 79] de Mori, R., Lesmo, L. and Poncini, M., "The Structure of a Lexicon for a Speech Understanding System", IEEE ICASSP'79, pp.252-255 (Apr. 1979)

8. 応 用 シ ス テ ム

8.1 機 械 翻 訳

8.1.1 歴史的概観

8.1.1.1 はじめに

機械翻訳の研究は、1940年代の後半から今日に到るまでの長い歴史をもっている。初期の楽天的な期待、本質的な困難さの自覚と悲観主義の時代を経て、現在は、計算機技術の発展に支えられて再び見直しの気運にあるといえよう。

1950年代の後半から1960年代の前半にかけて、ソ連の様々な情報の入手に力を入れていたアメリカは、露→英自動翻訳の研究に期待し援助した。このことが世界的にも機械翻訳の研究を活発に促すことになった。しかしその後アメリカはALPACレポートによって機械翻訳システム開発の見直しを行い援助を大巾に後退させた。この影響で機械翻訳の研究は世界的にも急激に退潮することになる。

しかし、カナダのような二言語国家やEC (European Community) のような多言語社会では、言語の相互交換は避けられない課題であり、計算機をそのために役立てることは重要な研究課題となってきた。このような背景もあって、カナダやヨーロッパではALPACレポート以後も、用語データバンクなどの翻訳支援技術も含めて機械翻訳の研究が地味に続けられてきた。また当のアメリカにおいても、Georgetown 大学で開発された露→英自動翻訳システムは原子力関係の文献の翻訳に現実に用立てられてきたし、機械翻訳の研究が皆無になったわけでもなかった。

ところで機械翻訳の研究は、その困難さの故に、より整った言語理論を求めて言語学を刺激し、また操作的な意味論を求めてAI (Artificial Intelligence) に「理解」「知識表現」一般の問題を提起するなど、関連領域に大きな影響を与えてきた。

このように機械翻訳の歴史は、科学技術史的に見ても興味深いものがあるが、この章ではこれまでに作成された自動翻訳システムを中心に機械翻訳の歴史を概観する。

8.1.1.2 黎明期

機械で自動的に翻訳を行うという発想は、電子計算機がこの世に出現して未だ間もない1946年に、ロンドン大学のA. D. Boothとロックフェラー財団の自然科学部長W. Weaverが議論したことが嚆矢とされている〔340〕。

W. Weaverは第2次大戦中の暗号解読技術との関連で、A. D. Boothは電子計算機の記憶装置との関連で論じている〔350〕。

1949年にW. Weaverは、機械翻訳に関するメモランダム〔610〕を、200人程の知人に書き送った。このメモランダムは広範な興味を呼びおこし機械翻訳の研究を世に広める歴史的きっかけとなった。この後、ワシントン大学でE. Reiflerが、カリフォルニア大学でOswald, Fletcher等が、そしてMITでBar-Hillelが研究を開始するなど各地に機械翻訳の研究グループが誕生した。

1954年に、IBM社とジョージタウン大学の共同による公開実験が行われた〔170〕。これはIBM701を用いてロシア語から英語への翻訳を試みたものであった。ロシア語の単語250項目が登録され、対応する英単語は高口2個に限られ、語順の反転や選択法、英単語の省略、投入など簡単な6個の規則を用いての自動翻訳で、1文章あたり6～7秒かかるというものであった。この公開実験はごく初歩的なものであったが、機械翻訳に対する一般の注意をひくのに十分なものであった。〔附録1〕

当時の時点で機械翻訳の課題として、A. D. BoothとW. Lockeは次の5点をあげていた〔350〕。

- ① 入出力手段の発展。特にローマン・アルファベットを用いない言語の為に文字読取装置、更に音声認識装置を開発すること。遠方からの入出力の為に電話線を用いること。
- ② 辞書や文法を蓄える為の大規模で安価な記憶装置。
- ③ 辞書；おそらく対象分野毎に作ることになる。
- ④ 語尾変化；接頭辞の分類と分析。
- ⑤ 計算機向きの実際的な構文解析手法。

このように当時は未だかなり楽観的な見通しを持っていたことがうかがわれる。当時は逐語訳が基本であったが、複数の訳語をもつ多義語の問題が意味の

問題として取り上げられていた。構文解析は、目的言語での語順を整える為のほか、多義語の問題を解決するために必要だと考えられていた。

8.1.1.3 第 1 期

IBM社とジョージタウン大学による露→英公開実験の後、1960年代の前半に到るまで、アメリカでは政府の大きな援助を受けて各地で活発に研究開発が進められた。フランスでもグルノーブル大学が1961年にCETA(Centre d'Etudes pour la Traduction Automatique)を設立し、カナダではモントリオール大学が1962年に、また日本でも九州大学、京都大学、電気試験所などで1950年代後半に研究が開始されるなど、機械翻訳の研究は世界的に盛んになった。このほかソ連でもO. KulaginaやI. Merchuk, D. Panov等がソ連科学アカデミーの下で本格的に研究を開始していた。

しかしこの期間に開発されたシステムで、ともかく実際的な場面に供し得たシステムとして知られているのは、FTD(Foreign Technology Division of U. S. Air Force)の露→英システムとジョージタウン大学の露→英システムのみであった。多くの研究努力にもかかわらず、実際的な自動翻訳システムを実現するには程遠く、結局、自然言語の複雑さとその機械処理の困難さを浮き彫りにする結果となったのだった。

FTDのシステムは、IBM社のG. King等が、翻訳専用の計算機Mark IIを作ってその上に構成したシステムで、1962年に計画され1964年に稼動を始めている。これは方式上は逐語訳を基本とする初期のものと殆んど同じで、大巾に後編集に頼る結果となっている〔470〕。次節で述べるALPACレポートではこのシステムが検討の対象とされているが、同レポートは、人手による翻訳のほうが半分以上のコストで、しかもより速いと指摘している。しかしこの指摘にもかかわらず、このシステムは、1970年にSYSTRAN(後述)に置き換えられるまで、FTDで稼動し続けている。

ジョージタウン大学のシステムは、1964年に開発されたもので、AEC(Atomic Energy Commission)のORNL(Oak Ridge National Laboratory)とイタリアのイスブラにあるEURATOM(European Atomic Energy Commission)で現実に用立てられてきた。ORNLでは現在も稼動中であり、EURATOMでは1976年にSYSTRAN

に置き換えられるまで稼動していた。

これはロシア語から英語への翻訳システムで、主として専門の科学者にソ連の情報を提供しようとするものである〔200, 281, 600〕。たいていは後編集なしで使用されており要求があった場合のみ後編集している。入力はいきパンチャがロシア文字を英文字へ打ち換えて入力しているが、うまくいかなかった翻訳（約25%）の大部分は入力ミスによるものだという。1973年に出された報告〔281〕によれば、ユーザのうち92%がこのシステムを情報源として価値があり、かつreadableであるとして評価している。機械翻訳の利点は、何といってもスピードが速いこと（10ページ翻訳するのに数分、いきパンチャに2～3時間かかるが、とにかくその日のうちに翻訳結果がもらえるという。）に求められているようである。

このジョージタウン大学のシステムは、L. Dostertの指導の下に、M. Zarchnakが開発したもので、I/O部分はアセンブラで、その他は同じジョージタウン・グループのA. F. Brownが開発したSLC (Simulated Linguistic Computer) という言語で書かれている。このシステムは基本的には経験主義的なシステムで、明確にはモジュール化されたシステムとされている〔260, 470〕。このため計算機種の変更（IBM7090→IBM370）はあったものの、システム自体に大きな改訂は加えておらず、従って翻訳の質も1964年当時のものから変わっていない。スピードは2000語/9秒、辞書は50,000項目の規模である。〔附録2〕

自動翻訳が当初期待されていたようには簡単に出来るものでないことが認識され、自然言語の複雑さが自覚されてくるにつれ、機械翻訳に対する悲観主義が芽生えてくる。初期の頃からMITで機械翻訳の研究にとりこんでいたY. Bar-Hillelは、1960年にFAHQT (Full Automatic High Quality Translation) の不可能性について論じて〔40〕、この悲観主義を方向づけた。

Y. Bar-Hillelの議論は、次のようなものであった。翻訳は、やはり意味内容の理解がなければ不可能である。たとえば

The box was in the pen.

のpenの意味は、writing penでなくtoy boxであるという選択ができる

為には、この文章の意味が理解できていなければならない。しかし、このような意味の世界を ad hoc ではなく体系的に記述すること、形式化することは不可能である。このように意味理解の形式化の困難さの故に FAHQ T は不可能であるというのが Y. Bar-Hillel の議論の骨子であった。

ハーバード大学の久野, A. Oettinger 等は, ロシア語から英語への自動翻訳のプロジェクトにとりこんでいたが, 途中で自動翻訳の看板を降ろし, 自動英文法, 自動英語辞書の開発改良という道を経て, 数理言語学的研究からさらに純言語学的研究の色彩を濃くしていった。これは A. Oettinger と久野の「精度の高い完全に自動的な翻訳は不可能であるし, 精度が低くかつ理論的裏づけのない実用的機械翻訳は, 大学・研究機関が行うべきでない」という確信を反映したものであった〔670〕。

この間言語理論の分野では, 句構造文法の理論が発展し, N. Chomsky の変形生成文法理論が非常な影響力をもつようになった〔140, 150〕。書き換え規則に基づく文法の形成化・数学的研究は計算機言語の理論の発展をも促し, また構文解析の手法もいろいろ提案された。

最も進んだ構文解析法とされたのが, ハーバード大学の久野による Harvard Multiple-path Syntactic Analyser (Harvard Analyser 又 Predictive Analyser) であった。これは任意の文に対して, そのあらゆる可能な構文解析を自動的に導き出すシステムで, 次のような文法規則を用いる。

$$(Z, C) \mid Y_1, \dots, Y_{m-1}$$

〔300〕

$$(Z, C) \mid A$$

ここで, Z, Y_i は文, 名詞句, 動詞句等の統語的カテゴリー, C は品詞名であり, Z という統語カテゴリーが C で始まっているならば, C の次に $Y_1, \dots, Y_m$ という統語カテゴリーが来るはずだという予言的文法規則を表現している。たとえば

(文・名詞) | 動詞句, ピリオド

このような文法形式は久野と A. Oettinger が経験的に導いたものであるが, これは Chomsky の II 型文法 (文脈自由文法) と同等であることが証明されている〔210〕。久野等はこのような文法規則に対する自動構文分析システムを開発し, 約 3,000 の英文法規則と約 20,000 単語を含むシステムを作りあげ

た。このシステムで

Time flies like an arrow.

に対して6つの解釈を与えたことは有名である。

このように構文解析の手法にはかなりの進展が見られたが、しかし意味にかかわる問題に関しては殆ど手つかずであり、機械翻訳システムにも大きな進展は見られなかった。

8.1.1.4 ALPACレポート

ALPAC (Automated Language Processing Advisory Committee) は、米国のNAS (National Academy of Science) の下に作られた諮問委員会で、米国防省、米CIAおよび米国立科学財団に対して機械翻訳の研究開発のありように関する勧告をするために設けられたものであった。ALPACは1964年4月に設置され、翌1965年8月にALPACレポートとして知られる報告書〔10〕を提出した。

この報告書は、そもそもアメリカでの翻訳の事情に根ざしたものであったにもかかわらず、世界中の機械翻訳研究・自然言語処理研究の流れに強烈な影響を与えた。この報告書を根拠として「機械翻訳」は失敗であったとされ、機械翻訳研究に対する援助は削減され機械翻訳研究は、アメリカのみならず世界的に急速に退潮していったのである。

本節では、以下に、ALPACレポートの内容を紹介する。

ALPACレポートは、

- ① 英語への翻訳の必要度
- ② 機械翻訳システムの現状と問題点
- ③ 機械翻訳研究の意義
- ④ 何を為すべきか

の4点について述べているが、基調を成しているのは、

“米国では翻訳家は不足しておらず、機械に頼って翻訳をしなければならないような事情はない。”

ということである。レポートは次のように述べている。

「機械翻訳に対する強いニーズがもしあったなら、我々の態度は多分違ったものになっていたであろう。しかし我々は、そのようなニーズを見出すことが

できなかった。」

米国では機械翻訳に対するニーズが弱いということに関して、レポートは次の4点をあげている。

- ① 科学文献では、もともと英語で書かれたものが多く（ある統計では、英語76%、露語14%、仏語4%、独語4%、その他2%）英語への翻訳は需要が低い。
- ② 外国語文献をひんばんに利用する人は、その外国語を自ら学んだほうが安上りで手っとり早い。
- ③ 米国には、翻訳家、翻訳機関は多数あり不自由していない。
- ④ 全ての文献が翻訳するに価値あるものでないし、これ以上翻訳文献が増えても活用する人がいないだろう。

機械翻訳システムについては、1964年時点での翻訳例4つを示し(①)Banker-Ramo Corporation (②)Computer Concepts Inc. (③)FTD (④)EURATOM)、質、速さ、コストいづれの点でも人手による翻訳に劣るとしている。また10年前1954年のジョージタウン—IBMの公開実験のものと比較してもたいして進歩していないとしている。また将来に関しても、機械が"understand"しない限り機械翻訳の向上はあり得ないが、この"understanding"への挑戦は、科学的には重要な課題であることは認めるが、実際の機械翻訳を近い将来改善するための動機とはなり得ないとしている。

一方、計算機による用語辞書作りなど翻訳支援技術に関しては、ドイツ連邦軍翻訳局の作成した用語辞典TRG(Text-Related Glossary)と、ヨーロッパ石炭鉄鋼委員会で作成された同様の辞書システムを引用して、その有効性に期待を示している。人手による翻訳で最も時間を費やすのは用語上の問題であって、TRGを使った場合の比較実験によれば、翻訳速度は50~80%向上し、誤りも1/3に減少したとの報告が示されている。

このように機械翻訳そのものについては否定的な見解を示しているが、しかしこの約10年間の機械翻訳の研究から言語学上の知見が増えたこと、また計算機のプログラミング技術にも好影響を与えたことは評価している。さらに一般に言語研究の重要さと、言語研究への計算機側からの接近の有効性について、この諮問委員会の議長であったJ. R. Pierceは次のような補強的見解を示

している。

「計算機は言語学者に多くの新しい課題やヒントや可能性を示してくれた。これは原子物理学の課題や可能性と比較し得るものである。言語はその重要性において他の何物にも劣らない。計算言語学に必要な道具立ては、原子物理学に要する大加速機と比べれば安いものだ。新しい言語学は、極めて重要なまた非常に魅力的な課題を提起している。

この挑戦に立ち向かうことは、多くの分野で重要な貢献を為すであろうことは間違いない。言語について更に深く知ることは以下の課題の助けになるであろう。

- ① 外国語の教育をもっと効果的にする。
- ② 言語そのものの特性についての教育をもっと効果的にする。
- ③ 自然言語の使い方そのものをもっと効果的にする。
- ④ 特殊目的のための人工語の使用を可能にする。(たとえば航空管制用言語)
- ⑤ 言語の使い方や、伝達・思考に関する心理実験を有意義なものにする。
(言語が何であるかを知らなければ、何を説明すべきかも分からない。)
- ⑥ 計算機を翻訳や情報検索の助けに使う。」

最後にレポートは、結論として次の2点を勧告している。

- ① 計算言語学を大いに助勢すべきこと。これは、構文解析、文生成、構造論意味論などのほか翻訳の実験などを含む言語学の領域であって、あくまでも科学上の領域である。
- ② 実際の翻訳作業の向上のために助勢すべきこと。(速度を上げ、コストを下げ、質を確保すること。)具体的な検討課題としては、以下のようなことがある。
 - ① 翻訳の結果を評価する実際的な手法。
 - ② 人手による翻訳をスピード・アップする手だて。
 - ③ 翻訳のコストと質の評価法。
 - ④ 翻訳物の利用状況の調査研究…利用されないものを生産しない為に。
 - ⑤ 全体としての翻訳の遅れについて、遅れを除去する方法。
 - ⑥ 計算機による各種の翻訳支援のコストと速度に関する検討。

⑦ 機械化された編集・印刷の手法を翻訳作業と適合させること。

⑧ 電子辞書などを含めて、翻訳者のためのよりよい参照のシステムを作ること。

以上がALPACレポートの概要である。このレポートが機械翻訳に関して否定的な見解を示していることは間違いないが、しかし、

- ・ 言語理解の問題を含めて、言語研究、特に計算機サイドからの言語研究の重要性に言及していること。
- ・ 機械翻訳に関する否定的見解は、当時のアメリカの翻訳に関するニーズに根ざしたものであること。

などの点には留意しておくべきである。

8.1.1.5 第 2 期

ALPACレポートが提出されて以後、機械翻訳研究は急速に退潮していった。しかし、当のアメリカにおいても、その他の国でも、いくつかの機関では、その後も研究開発は続けられた。その主なものを、表8-1に示す。

これらのうちで実際に用立てられてきたシステムとしては、

- ① ジョージタウン大学のシステム
- ② SYSTRAN
- ③ TAUM-METEO
- ④ CULT

があげられる。①については8.1.1.3で述べた。以下②～④について概観する。

表 8-1 ALPAC レポート 以後のおもな機械翻訳研究

研究開発機関／者	国	プロジェクト名 ／システム名	時 期	参考文献等備考
ジョージタウン大学 (Z. Zarechnak ほか)	米		1950 頃～	8.1.1.3 参照
Latsec Inc. (P. Toma)	米	SYSTRAN	1964 頃～	8.1.1.5.1 参照
モントリオール大学 (J. Chandonx ほか)	加	TAUM	1960 頃～	8.1.1.5.2 } 参照 8.1.2.3.1.1 } METEO は TAUM のサブセット
グルノーブル大学 (B. Vanqnois ほか)	仏		1960 頃～	8.1.2.3.1.2 参照
テキサス大学 (W. A. Lehman ほか)	米	METALS	1965 ～1975	独→英 グルノーブル大学の中間言語 法と同様の手法 〔11〕
カリフォルニア大学 (W. S. Y. Wang)	米	POLA	1967 ～1975	中→英 TAUM と同様の手法 中国語特有の問題を扱っている 〔11, 54〕
Rand Corp. (M. Kay)	米	MIND	1970 頃	人間との対話を通じて 翻訳 を遂行するシステム 〔11, 55, 56〕
Xonics Inc. (B. Chalompka ほか)	米		1971 頃～	露→英 ジョージタウン大学のシステム の改作？ 〔57〕
Logos Development Corp.	米	LOGOS	1973 頃～	英→ベトナム ベトナム兵の為に英語の航空機 マニュアルを翻訳。SYSTRAN と同様の第1世代のシステム LOG III は英→ベトナム, 露 仏 〔11〕

Brigham Young Univ. (E. G. Lytle ほか)	米	ALP	1975 頃～	英→仏, 独, ポーランド, スペイン モルモン教会の教義書の翻訳 現在は人間が介在するシステム だが全自動をめざしている。 〔11, 58〕
香港大学 (S. C. Loh ほか)		CULT	1975 頃～	8.1.1.5.3 参照
	仏	TITUS	? 70年代後半～	フランス織物研究所で稼働中。 METEOと同様に入力言語を 制限している 〔27〕
ザール大学 (H. D. Maas ほか)	独	SUSY	? 70年代後半～	露, 仏, 英, エスペラント→独 SFB-100という言語処理 プロジェクトの一環 〔59, 60〕
Pearl Assurance Co. (W. Goshawke)	英	SLUNT	? 70年代後半～	入力言語を制限している。 Number Language という 中間言語を介している。 〔61〕
All-Union Centre for Transration (Yu. N. Marchuk ほか)	ソ		1974～	4つの機械翻訳プロジェクト が進行中 英, 独, 仏→露 〔62, 63, 64, 65〕

(注) 用語データベースなどの翻訳支援システムについては8.1.2表3参照

表 8-2 世界の機械翻訳研究 ([8 0] から引用)

Survey of machine translation systems

1 AMERICA

11 Canada

Lakehead University, Thunder Bay
University of Montreal

12 United States

Atomic Energy Commission, Oak Ridge/Georgetown University, Washington D.C.
USAF, Dayton
NASA, Houston
Brigham Young University, Provo
Massachusetts Institute of Technology, Cambridge *
University of California, Berkeley *
University of Texas, Austin *
University of Texas, Austin/Ramkhamkaeng University, Bangkok
Yale University, New Haven
Latsec, Inc./World Translation Center, Inc., La Jolla
Logos Development Corp., New Hampton
Smart Communications, Inc., New York
Xonics Inc., McLean/Tabor, Inc., Nokesville
Xyzyx Information Corp., Canoga Park

2 ASIA

21 British Crown Colony of Hongkong

Chinese University of Hongkong

22 Japan

University of Kyoto
Kyushu University, Fukuoka
Electrotechnical Laboratory, Tokyo

23 Lebanon

International Language Centre, Beirut

24 Malaysia

University of Science of Penang

3 EUROPE

31 Belgium

University of Antwerp

32 Bulgaria

Bulgarian Academy of Sciences, Sofia

33 Federal Republic of Germany

University of Heidelberg/University of Constance
University of Cologne
Ruhr University, Bochum
University of the Saarland, Saarbrücken
Centre for Textile Documentation and Information, Düsseldorf

34 France

University of Science and Medicine, Grenoble
French Textile Institute, Paris

35 Great Britain

University College, Cardiff
University of Essex, Colchester
Pearl Assurance Co., Ltd., London/Natural Language Translation Specialist Group

- 36 Italy
Euratom, Ispra *
- 37 Soviet Union
State University of Leningrad
Group for Language Statistics, Leningrad, Minsk,
Kishiniov, Machatshkala, Alma-Ata, Irkutsk
Central Research Institute for Patent Information,
Moscow *
Atominform, Information Centre for Nuclear Energy,
Moscow
Informelektro, Documentation Centre of the Institute
for Electrical Engineering, Moscow
Institute for Applied Mathematics, Moscow
State Pedagogic Institute for Foreign Languages,
Moscow
All-Union Centre for Translation of Scientific and
Technical Literature and Documentation, Moscow
Institute for Electronics, Automation and Telemech-
anics, Tiflis
- 38 Czechoslovakia
Karlovy University, Prague

Note

* Further development or use abandoned or interrupted.

Survey of machine-assisted translation systems

1 AMERICA

- 11 Canada
University of Montreal/Department of the Secretary of
State, Ottawa
- 12 United States
IBM, New York

2 EUROPE

- 21 Federal Republic of Germany
Bundessprachenamt, Hürth
DEMAG A.G., Duisburg
Siemens A.G., Munich
- 22 German Democratic Republic
Technological University of Dresden
- 23 Luxembourg
European Communities, Luxembourg
- 24 Netherlands
Netherlands Ministry of Foreign Affairs, The Hague
N.V. Philips Gloeilampenfabrieken, Eindhoven

Note

This list does not include the numerous multilingual
standardization data banks (dictionary-related terminology
data banks).

8.1.1.5.1 SYSTRAN

SYSTRANはジョージタウン大学のシステムをAECとEURATOMに装置した責任者であったP. Tomaがジョージタウン・システムを大巾に改良すべく1964年以来開発してきたシステムである〔550, 560〕。ドイツ科学財団との契約の下に最初のシステムを作り、後1967年には米空軍の援助を受け、さらに1968年にLatsec Inc.*が設立されてそこでSYSTRANの開発が続けられた。もともとロシア語から英語へのシステムであったが、Latsec Inc.で、英→露、英→仏、独→英などにも拡張されている。

1970年に、米空軍のFTDで稼動していた露英システムMarkIIが、また1976年にはEURATOMのジョージタウン・システムがSYSTRANに置き換えられて現在に至っている。またこの間にNASAにも納入されて、アポロ・ソユーズ計画の際に作業書の翻訳に使われたという。またEC(European Community)委員会では、ECで大きな問題となっている多言語間翻訳の問題への対処の一環として、1976年以来、SYSTRANの英-仏版を試験的に使っている。マニュアル類の翻訳作業の効率化を検討しているXerox社でも、SYSTRANを試験的に導入している〔26〕。

P. TomaはSYSTRANを、

「実際に動いている大規模なシステムとしては唯一のもので第3世代のシステムだ。1対の言語間の翻訳でなく、多言語間の翻訳システムに容易に拡張できる。」

と主張している〔560〕。

* 後、World Translation Inc.

SYSTRAN EXECUTIVE

Simplified 'sequence of events'

- A Input text;
- B Scan for high frequency words
and append information from HF dictionary;
- C Scan for idioms
and append information from ID;
- D Alphabetise, access Master Stem Dictionary,
and append information from MSD;
- E Return 'words' to original text order;
- F Disambiguate homographs;
- G Invoke Limited Semantics disambiguation routine, using Topical
Glossary parameters as necessary;
- H Invoke Structural Pass 0
- i.e., establish primary, 'global' syntactic dependencies, 'immediate
environments' etc.;
- I Invoke Structural Pass 1 ('right-to-left')
-i.e., recognise syntactic structures such as congruence, agreement,
apposition, partly by using punctuation marks as 'clues';
- J Invoke Structural Pass 2
-i.e., establish categorisation of subordinate clauses according to
'initiators';
- K Invoke Structural Pass 3
-i.e., establish Subject-Predicate relationships and embeddings;
- L Invoke Structural Pass 4
-i.e., clear out 'rag-bag' of current unsolved problems
(Human interaction possible, if needed.)
- M Insert 'articles', wherever necessary;
- N Resolve details of prepositional dependencies;
- O Access Limited Semantics File, if needed;
- P Translate prepositions;
- Q Access Lexical Programs, and switch on routines for semantic
disambiguation via conceptual calculus;
- R 'Invert' appended information records to produce target language units;
and execute accompanying instructions for TL Synthesis
- S Re-arrange elements to suit TL structuration;
- T Produce preliminary printout (only if desired for immediate analysis);
- U Printout results and channel to photo-composer.

図 8 - 1 SYSTRAN の処理手順〔290〕から引用

SYSTRAN DICTIONARY STRUCTURE

- A Simplified Part Record Structure in Master Stem Dictionary
(144,000+ stem entries reported at present in Russian MSD)
- I First Grammar Subpart
 - i) syntagmatic information, frequency characteristics, homograph flag etc.
 - ii) Part of speech codes
 - iii) Basic part of speech codes
 - iv) Paradigmatic set codes
 - v) Gender, number, case, person, tense info.
 - II Second Grammar Subpart
(‘Fillmore-type’ case information/‘valency’)
 - i) case requirements
 - ii) prepositional requirements
 - iii) syntactic function codes
 - III Third Grammar Subpart
(case/preposition codes)
 - i) instrumental codes
 - ii) dative case codes
 - iii) prepositional translation control bytes
 - IV Fourth Grammar Subpart
 - i) Topical glossary code
 - ii) English synthesis codes
 - iii) English meaning(s)
- B The Russian Idiom/Limited Semantics (LS) Dictionary is reported to contain at present:
- a) 3,300+ ‘declinable’ idioms and (technical) collocations.
 - b) 160,000+ ‘declinable’ idioms and (technical) collocations.

図 8 - 2 SYSTRAN の辞書構成 -〔290〕から引用

SYSTRANは、

- ① 入力
- ② 辞書引き
- ③ 構文解析
- ④ 翻訳

の4つの段階から成っている(図8-1)。辞書は、主記憶に入っている高頻度語辞とディスク内の非高頻度語辞書に分けてあり、更に語の長さによって3つに分割されている(図8-2)。高頻度語辞書は文章の入力時に照合し、その他の語は集めてソーティングした上で非高頻度語辞書と照合する。

構文解析は4パスから成り、ボトムアップ法による。

- ① 同形意義語の処理⁽⁺⁾
- ② 右から左に走査して、品詞情報から可能な構文グループを抽出する。

動詞+目的語、前置詞+目的語 など

- ③ 左から右に走査して、構文グループを更にまとめあげていく。節の境界を定め、主節か従節かなどを決める。
- ④ 各節の中を走査し、主語、述語を抽出する。

翻訳の段階は、たとえば冠詞を搜入するサブルーチンのような多くのサブルーチンから成っている。たとえば

result of treatment of burns
を、
result of treating the burns
にするというような構文の変形も行なわれる。

{ noun + of + verbal noun + of + noun
→ noun + of + gerundine + noun }

また、語の意味カテゴリー(約300用意してある)によって訳語を定めることもある。たとえば、前置詞の対応は、それにつく名詞によって次のように異ってくる。

+) 後World Translation Inc.

(+) 詳細は不明

$\Delta O + [+increase] \rightarrow up\ to$
 $\Delta O + [+decrease] \rightarrow down\ to$
 $\pi o + [+linear] \rightarrow along$
 $\pi o + [+noulinear] \rightarrow over$
 $\pi o + [+mental\ tool] \rightarrow using$

SYSTRANでは語の辞書記述の長さは任意であり、種々のプログラムを結合しておくこともできる。目標言語での語順の並べ換えなども含めて、かなりの部分が辞書に記述されるようになっている。このようにSYSTRANは言語理論なり体系的な方法論に根ざしたものではなく、ad hocで経験主義的な方式であるといえよう。このような批判に対してP. Tomaは、

「我々は言語理論には興味がない。とにかくよい翻訳ができればよいのだ。事実SYSTRANはこれまで現実に移動してきたし有効であった。」と反論している。

Xerox社でSYSTRANを実験的に使用した経験では、入力言語を制限して入力段階で負担を強いるか、または入力言語は自由にしておいて後編集に負担をかけるかのジレンマがあったが、入力言語を制限する(語・用法)方法をとった結果として、SYSTRANは十分に有用であるとの判断を示している〔190〕。今日までのところXerox社では、英語からフランス語への翻訳実験を行っており、今後、他の言語対の実験も行おうとしている。

EC委員会でも、1976年以来SYSTRANの英仏版を試用し改善を重ねてきた。EC委員会では、入力言語は制限せずに後編集に負担をかける方法をとっているが、それでもやはりSYSTRANは有用であるとの判断を示している〔27〕。ただし後編集の作業は心理的に非常にたいくつな、また苛だつ作業で、後編集を依頼された翻訳家には不評である〔20〕。この品質の低さは、辞書が未だ小さいことと、意味論上のknow-howが組み込まれていないことによるものとされている〔290〕。EC委員会では、英仏版に続いて、仏英版についても実験しており、更に英伊版も作成予定である。EUROTRA計画による新しいシステムができるまでは、SYSTRANの試用、改善を進めることになっている。〔附録3〕

8.1.1.5.2 METEO

METEOは、天気予報文の英仏翻訳という限られた目的の為のシステムである〔120, 130, 260〕。現実にカナダの新聞・テレビ・放送局で用いられており、このように公的に用立てられたシステムとしては最初で唯一のものである。これは、モントリオール大学のTAUM(Traduction Automatique de L'Universite de Montreal)プロジェクトチームが1974年にカナダ翻訳局から依頼されて作ったシステムで、TAUMを天気予報文用に大巾に特殊目的化したものである(TAUMについては§2.2で述べる)。1976年には本格的に稼動するようになり、毎分1,000程の速度で、1日約30,000語の天気予報文を翻訳している。システムは完全に自動的で、翻訳不能となって棄却されたものだけを人間が翻訳しており、システムが翻訳した文については、後編集は一切行われていない。棄却率は20%以下で、その原因は大部分が綴りや句読点などの入力ミスに基づくものだという。またコストは、人間によるこのバックアップを含めても、人手による場合の約1/3に下ったと報告されている。

METEOは次の4部分からなる。

- ① 慣用句辞書
- ② 一般辞書
- ③ 解析部門
- ④ 生成部門

慣用句辞書は約300項目から成り、本来のイディオムのほか、全く固定的な言いまわしであるため、その都度分析するようなことをしない表現、たとえば

forecast issued by the atmospheric environment service
や、また翻訳する必要のある地名、たとえば

Lake St Claire → lac Ste Claire
などを含む。

一般辞書は約1,200項目から成り、統語カテゴリー、翻訳、意味素姓などを記載している。また活用形などの異形態の語もすべて辞書に登録してある。

解析部門は約300項目の書き換え規則から成り、入力ストリングに対して文法カテゴリーと翻訳を与える。天気予報文の世界はT. Winogradの積木の

世界と同じように、十分に制約された小世界であり、文は単純で関係節や受身文はない。主として問題になるのは、電報文的な天気予報文の統語的境界を定めること、多用されている接続語の扱いなどである。

パーズングは、次の5つのパスから成る。

- ① 数字を含むストリングが、日・時刻・温度のいずれを表わすものかを分析する。
- ② 場所と時の表現を分析する。時の表現では期間を表わすものか、時間を表わすものかを区別できなければならない。

in the morning → daus la matinee

this morning → ce matin

接続語に関して、冠詞を補充しなければならないこともある。

this afternoon on evening

→ cet apres-midi ou ce soir

- ③ 残りのストリングを解析する。正しい翻訳を選択するためには、語の意味素姓を活用する必要がある。

heavy fog → brouillard generalise

harvy rain → forte pluie

接続語の勢力範囲を定めるのにも意味素姓を活用する必要がある。

snowflurries or rainshowers becoming intermittent

tonight... → (snowflurries or rainsnowers)becoming...

- ④ あいまいさ、整合性のチェック
- ⑤ パーズングできなかったものを棄却する。また若干の文体上の調整をする。

生成部門は約300項目の書き換え規則から成り、解析部門が作り出したパーズング木を分析して、必要な冠詞を挿入し、語順を整え、性・数の一致をとり、フランス語文を生成する。

METEOでは、文法規則の記述にQ-systemを用いている。Q-systemはColmerauerが開発したもので、木の列を別の木の列に変換する書き換え規則を扱うシステムである。この規則は、木のある部分に対して適用することも

できるし、ラベルやリストやトリーを変数として含むこともできる非常に強力なものである。

〔附録 5〕

8.1.1.5.3 CULT

CULT (Chinese University Language Translator) は、香港大学の S. C. Loh 等が開発した、中国語から英語への翻訳システムである〔360, 370, 380〕。S. C. Loh 等は、1969年に研究を開始して中国語の数学などの文章なら、人間が適当に介入すれば機械翻訳も可能であり有効だと判断し、1972年に最初の中→英翻訳システムを構成した。1975年からは、北京科学アカデミー発行の数学会誌 *Acta Mathematica Sinica* と、物理科学会誌 *Acta Physica Sinica* の公式の翻訳に用立てられている。

CULT は、意味のあいまいさや、非常に複雑な構造の文などについては、前編集によって解決をはかっており、後編集より前編集に重きをおいたシステムである。テキストの 5 割ぐらいの部分に事前に手を入れることによって、相当複雑な文章でも翻訳できるとしている。(+)

中国語（つまり漢字）の入力は、当初は人間が 5 桁の数字にコード化して入力していたが、最近のシステムでは、約 200 種の漢字の要素を配した独自のキーボードを開発し、入力作業の簡素化に努めている。またこの入力手法の改善と合わせて、バッチ・システムからオンライン・システムに改編し、前編集を対話的に行い、処理途中で人間が介入できるようにしている。

CULT は、殆ど品詞のみに頼った構文解析を行っており、何回かのパスで順次部分的に翻訳をすすめ、最後に並べかえて出力するという手法のシステムで、分析のレベルは浅く、人間の介入を志向したシステムである。

〔附録 6〕

8.1.2 現状と課題

8.1.2.1 はじめに

国際的なコミュニケーションの量が増え、地球が狭くなったと言われる今日、翻訳に対する需要はますます増大している。EC では EC 議会の職員の 3 人に

(+) 具体的にどのような前編集をするのかは不明である。

1人は翻訳業務に関わらねばならないという〔490〕。また例えば、Xerox社では、世界36ヶ国と取引があるが、そのうち英語を第1国語とするのは7ヶ国だけである。このためマニュアル等種々のドキュメンテーションの供給が大きな問題となってくるが、うち14ヶ国に対しては英語のままで、残り15ヶ国に対しては翻訳した上で供給せねばならないという〔190〕。

このように翻訳に対する需要は非常に増大している。しかも速いスピードと低いコストが求められる。一方計算機技術は、リードウェア技術、ソフトウェア技術ともに、格段の進歩を遂げてきている。また言語学も一定の進展を示している。

このような情勢のもとに、最近再び機械翻訳の研究に、はずみがつきかかってきている。ただしこの場合、機械翻訳といっても、自動翻訳をめざすものばかりでなく、人間の翻訳家のための用語データバンクや、翻訳作業のためのテキスト・プロセッサ、計算機と人間が対話的に翻訳を進めるシステムなどの翻訳支援システムを含めて、「翻訳作業のために電子計算機を活用すること」と広くとらえておくべきである。この章では、現在すすめられているいくつかの研究プロジェクトを中心に現状を概観する。

8.1.2.2 翻訳支援システム

8.1.2.2.1 テキスト・プロセッサ

自動翻訳や更には用語データバンクに疑問を呈する人でも、テキスト・プロセッサとして計算機を翻訳作業に活用することについては高い評価を下している〔20〕。テキスト・プロセッサは、紙と鉛筆（タイプライタ）を用いて、消したり書き加えたりしていた旧来の翻訳作業の煩瑣な部分を、機械でスマートに行わせる、いわば筆記革命である。

今日のほとんどの計算機システムは、程度の差はあるが、翻訳作業への応用とはかかわりなく、テキスト編集の機能をもっている。パーソナルコンピュータとして著名なXerox社のAltoマシンは、文章作成やプログラム開発のためのテキスト・プロセッサとしての応用にも留意したものである。同じくXerox社のTeitlemanは、同一画面にいくつものウィンドウを重ねられるなど高度の機能をもった編集およびプログラミング支援システムを開発している。

IBM社のE. Lippmanは、テキスト・プロセッサと用語データバンク等を

総合した翻訳支援システムを開発した〔320, 330〕。

8.1.2.2.2 用語データバンク

専門の翻訳家にとっても、専門用語、特に新しい技術や製品また新しい概念の発生などに伴う新造語や語義の変化に、常に正しく対応するのは不可能なことである。翻訳家が翻訳作業に費す時間のうち約60%は用語の問題であるという。常に最新の情報を著えた辞書が、しかも手近に使えれば、人間が翻訳作業をする場合に非常な助けとなる。このようなことから、ヨーロッパやカナダなどでは、用語データバンクの開発が進められてきた。表8-3に、その主なものを示す。用語データバンクは、あくまでも人間の翻訳者のためのものであって、自動翻訳のための機械辞書とは別個のものである。また用語データバンクの開発は、用語の統一、標準化の問題とも密接にからんでいる。

表8-3 おもな用語データバンク

開 発 機 関	国	システム名	時 期	文 献 等 備 考
ブリッセル自由大学 (M. L. Hirshberg ほか)		DICAUTOM	? 1950年代後半～	1963以後ECの石炭鉄鋼委員会で使われている。 EURODICAUTOMの前身
E C		EURO-DICAUTOM	1969～	(§2.2.2) 参照
モントリオール大学	加	TERMIUM	1970～	EURODICAUTOMと類似 〔11〕
Bundeswehr Language Service	西独	LEXLS	1965～	テキスト中にアンダラインをしてそのまま入力。(但し人手で原形に変換) 現在120万項目 〔11, 67〕
Siemens AG (J. Schultzy ほか)	西独	TEAM	? 1970頃～	(§2.2.2) 参照
AFNOR (フランス規格協会)	仏	NORMATERM	1973～	ISOにのっとった規格、標準のシソーラス(英, 仏) 〔11, 71〕
(J. Mathias ほか)	米	CETA	? 70年代後半～	中-英翻訳の為の辞書 現在60万項目 〔72〕

EURODICAUTOM は、EC の翻訳サービスのために、ルクセンブルグの計算センターで開発されたもので、DICAUTOM を発展させたものである〔20, 220, 230, 490〕。DICAUTOM (Dictionnaire automatique) は、1963 年以来、ブリュッセルの石炭鉄鋼委員会で使われていた用語データバンクである。

EURODICAUTOM は、ヨーロッパの各言語間での用語の対応や定義、使用例を著えたデータバンクであり、オンラインで対話的にアクセスできる。複合語に対しては、その部分列に関しても検索する。たとえば

relative cinematic viscosity

が検索できなかったなら、

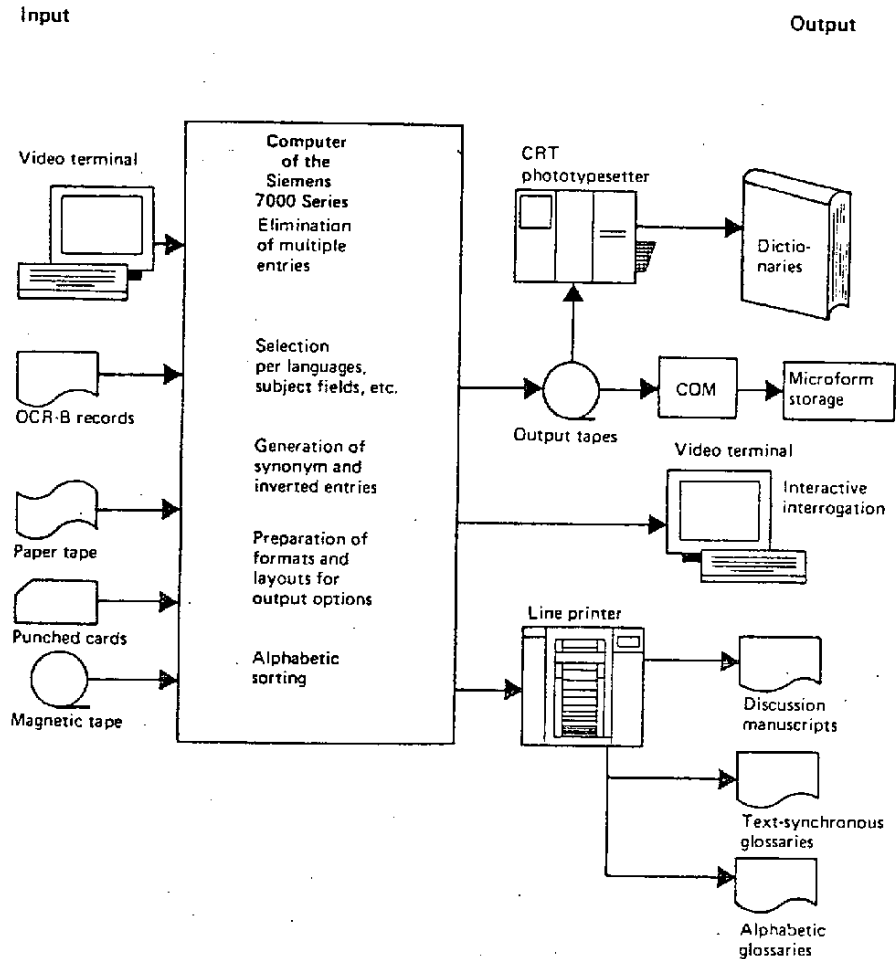
relative viscosity

cinematic viscosity

を自動的に検索する。入力語句が複数形であったり格変化をしているために検索が失敗するような場合には、自動的に、語尾部分を置き換えた形で再度検索するようになっている。

現在、約 130,000 項目登録されているが、本格的に活用されるためには、400,000 項目ぐらい必要だとみられている。満足のいく解答を、高い確率でしかも短時間に引き出せることが用語データバンクの有効性の指標であるが、EURODICAUTOM も未だ完成されたものではない。EURODICAUTOM は近い将来、EURONET を通じて利用できるようにすることが予定されている。

TEAM (Terminology Evaluation and Acquisition Method) は、Siemens 社が開発している翻訳支援システムで、用語や慣用句に対して訳語と使用例、定義を集成したものである〔60, 510, 520, 540〕。TEAM は入出力の形態として様々の手段を用意しており (図 8-3)、またテキスト・プロセッサと組み合わせて本格的な翻訳支援システムを構成することをねらっている (図 8-4)。



The TEAM Terminology Data Bank System

A characteristic feature of the system is the variety of its input and output facilities.

図 8 - 3 TEAMシステム - { 5 1 0 } から引用

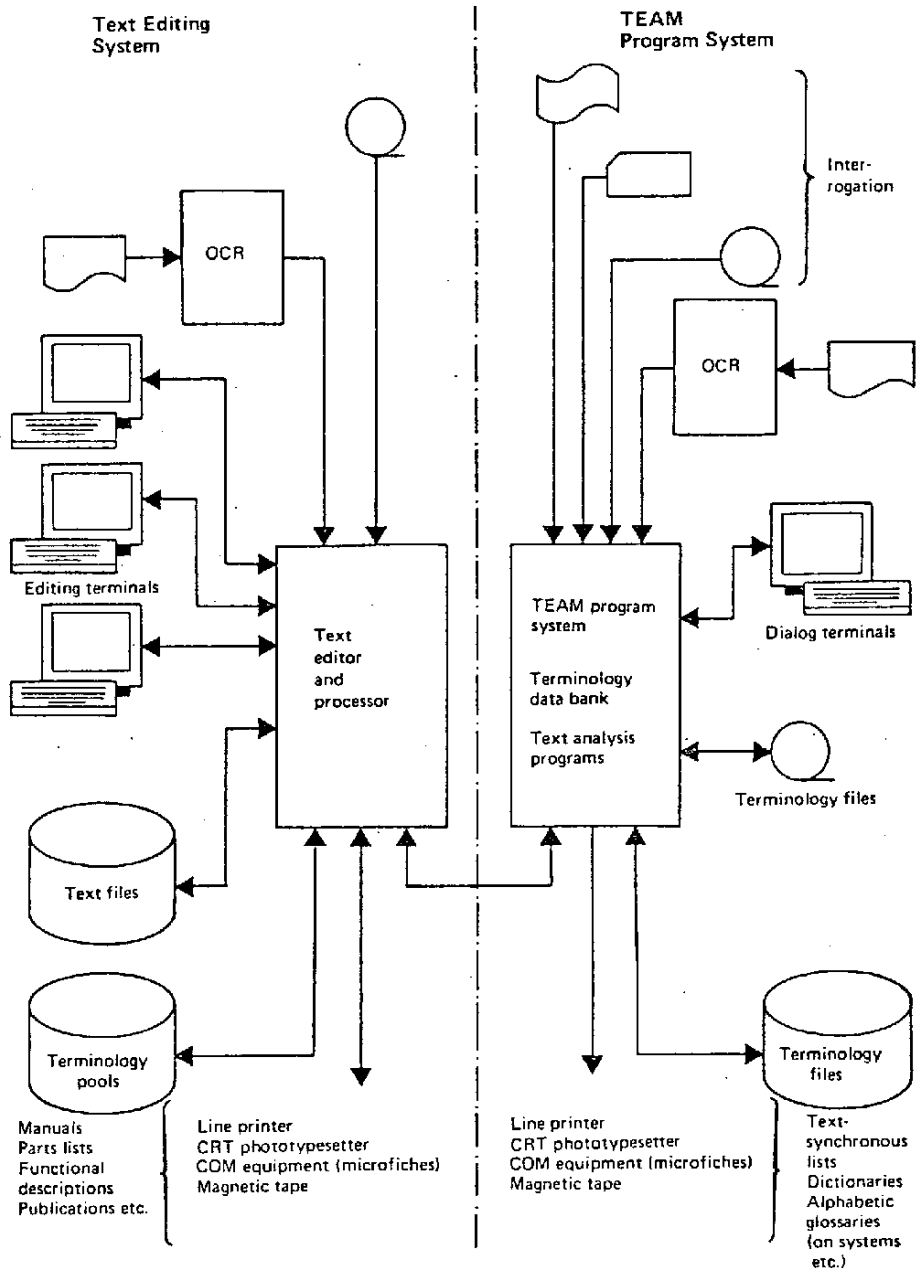


図 8 - 4 TEAMシステム - [510] から引用

8.1.2.3 自動翻訳

8.1.2.3.1 翻訳手法

自動翻訳は、その手法によって次の3つの世代に分けることができる。

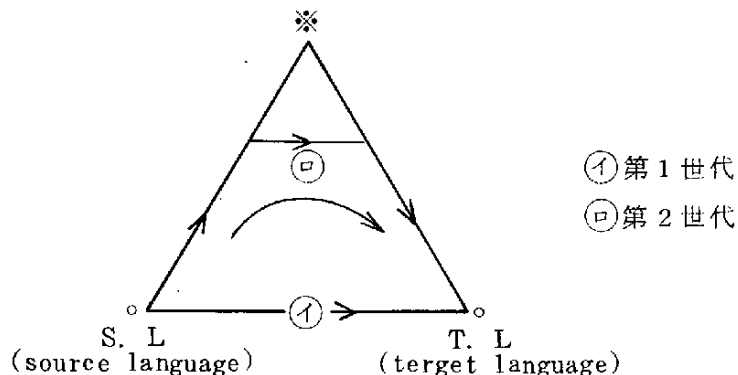
① 第1世代

特定の2ヶ国語間の翻訳を志向したもので、言語理論上の背景も乏しく基本的に経験主義で ad hoc なシステム。したがってプログラムのモジュラリティも低く、変更改善も施しにくい。

② 第2世代

- ・入力文を分析して中間的な表現におとす。
- ・目標言語での中間的な表現に変換する。
- ・出力文を生成する。

以上の、分析・変換・生成の3段階を明確に区別し、それぞれの言語の体系的な構文意味分析に基いたシステムで、多言語間の翻訳にも使える。プログラムもモジュラリティを高め見通しをよくし、変更・改善が容易にできるようになっている。



③ 第3世代

構文論よりむしろ意味論を重視し、言語理解の立場から自動翻訳を見直そうというもの。

今日までに開発され実際に用立てられたシステムは、METEOを除けば、全て第1世代の自動翻訳システムというべきであろう。第1世代のシステムは、翻訳の質が悪くても内容を読みとれる専門家に対してはある程度の価値はあっ

たが、そのまま出版物にできる程の質は得られていない。METEOは第2世代のシステムといわれているが、天気予報文という小世界に限定することによって成功を納めたシステムである。

EUROTRA など、当面するニーズに対応しようとしている現在の自動翻訳の研究は、第2世代のシステムをめざしている。

第3世代のシステムは、Y. Wilks などが先鞭をつけているが、未だこれからの課題である。第3世代のシステムは言語理解システム的一种であり、省略や指示語などによる参照関係を解くことなど、システムとしての自然言語のより深い把握をめざし、その上に構築されるシステムである。

以下、現在とりくまれている自動翻訳研究について概観する。

8.1.2.3.1.1 TAUMプロジェクト

カナダのモントリオール大学では、カナダ政府の援助を受けて、1962年以来一貫して自動翻訳の研究開発を続けている。これはTAUM (Traduction Automatique de L'Universite de Montreal) プロジェクトと呼ばれている。METEOはTAUMプロジェクトの一環としてできたシステムである。METEOは、天気予報という狭い世界に限定して構成されたシステムであったが、現在TAUMは、もっと一般に技術文献の翻訳を行うシステムを作成中である。これはAviationプロジェクトと呼ばれており、直接的にはカナダ空軍の為にアメリカの航空機マニュアルをフランス語に翻訳しようとするものである〔110, 260〕。

カナダ政府は、biculturalismの政策をとっており、政府の役人が英仏両語いづれに対しても同等に仕事が進められるようにしようとしている。またここ数年内に、文書管理や機械翻訳(用語データバンクを含む)のための計算センタを作ろうとしている。TAUMの動きはこのような動きに呼応するものである〔110〕。

Aviationシステムは、次の6つの部門から成る。

- ① 形態分析
- ② 辞書引き
- ③ 統語分析
- ④ 英→仏変換

⑤ 統語構造生成

⑥ 形態構造生成

形態分析は、辞書引きルーチンを使いながら次のようなステップで行なわれる。

① 不活用詞（前置詞，接続詞，冠詞など）の抽出

within → P(within)

several → QUANT(several)

in the process of → P(intheprocessof)

② 接頭辞の分離

understood → under+stood

③ 不規則形の分解

analyses → SW(analysis) + S

stood → SW(stand) + ED(pst)

④ 分離した接頭辞の復元

under + SW(stand) → SW(understand)

⑤ 未だ分析されていないもの（不変化詞でも不規則形でもないもの）の語幹と接辞への分解

・ -S, -ED, -ING, -LY, などの分離

tried → tri + ed

putting → putt + ing

series → serie + s

・ 終端の tt を t に変えたり，終端の i を y に変えたりして原形を想定する。

putt → put

tri → try

・ 以上のように分離した形，もとのままの形を両方とも辞書と照合する。

・ 辞書の情報を書き移す。

品詞 名詞の場合…

意味素姓 (ANI, CONC, ABST, …)

動詞の場合…

適合的な格のカテゴリ，必要な前置詞 (listen to, …)

など

辞書は、機能語など高頻度の語は主記憶に他はディスクに入っている。形態分析部門は Pasca-11 で直接にプログラムすることになっている。

統語分析部門は、次の 2 つの段階から成る。

- ① 名詞句や動詞句を解析して解析木を作りあげる。

$$\text{Art}(U^*) \rightarrow \text{Det}(\text{Art}(U^*))$$
$$\text{Det}(V^*) + \text{N}(X^*) \rightarrow \text{NP}(\text{N}(X^*), \text{Det}(V^*))$$

- ② 結果を標準形に変換する。すなわち；

- ・関係詞節の先行詞を復元する。
- ・動詞句を前に出し、名詞句をその引数の位置に配する。
- ・受動態を能動態に変える。
- ・形式主語を真主語に置き換える。

$\text{It be Adj that } S \rightarrow S \text{ be Adj}$

などの変形操作を施す。(図 8-5 参照)

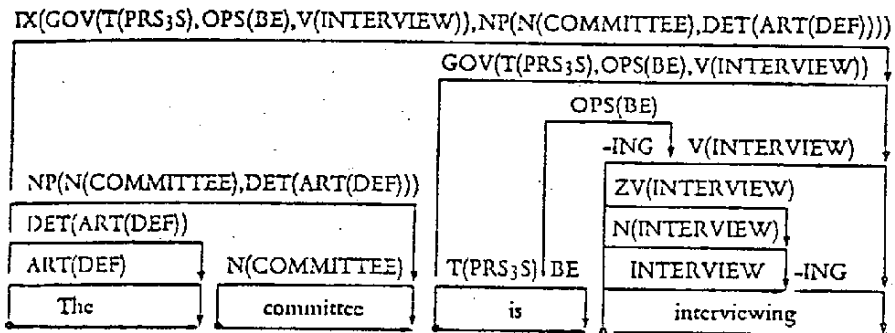


図 8-5 TAUMの分析例 ([260] から引用)

統語分析部門は、REZOという言語で書かれることになっている。REZOは、A. Woods の ATN (Augmented Transition Network) 文法に基づく言語であり、Pasca-11 で書かれている。REZOは、不成功に行きつく解析を早い時点で逃げたり、同一表現の分析をくり返すことを避けるなどの機能を

もっており、技術文献によく現れる長い名詞連続の分析を効率的に行なえるように配慮している。

変換部門は、各辞書項目に書かれている翻訳規則をインタープリットする形式をとっている。翻訳規則は一般に、解析された木構造に関連して記述される。

たとえば、“to replace”の翻訳規則は次のような記述になろう。

if : その目的語は「消耗品」という意味特徴をもっていない。

then : “remplacer”

else if : その目的語に「ひびわれた」「こわれた」などの形容詞がついていない。

then : “remettre en place”

else ...

統語構造生成部門は、複錯した木構造を次々と分解して木の列に変換する。

たとえば；

SN(N(Gens), Det(Les), GP(P(De), SN(N(Village), Det(Le))))

→ Det(Les) + SN(N(Gens), GP(...))

→ Det(Les) + N(Gens) + GP(...)

→ Det(Les) + N(Gens) + P(De) + SN(N(Village), Det(Le))

→ Det(Les) + N(Gens) + P(De) + Det(Le) + N(Village)

統語構造生成の操作は、このように完全に決定論的な操作である。この変換規則は、Q-systemで書くことになっている。

形態構造生成とは、最終的な表層上の記号列を生成することである。たとえば、

Det(LES) → les

P(DE) + DET(LE) → du

[附録 4]

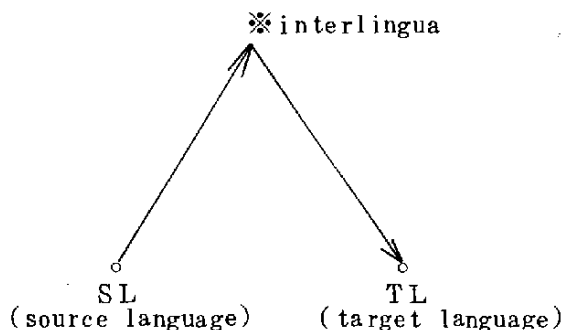
8.1.2.3.1.2 グルノーブル大学の自動翻訳研究

グルノーブル大学のCETA(Centre d'Etudes pour la Traduction Automatique)でも、1961年以来一貫して機械翻訳の研究開発をすすめてきた。

最初の約10年間は、中間言語法に基いて、ロシア語からフランス語への翻

訳システムを開発した。また日本語、ドイツ語からフランス語への翻訳も試みられた。

中間言語法というのは、翻訳しようとしている2ヶ国語間に人工的な中間言語（CETAでは pivot language と呼んだ）を設定して、中間言語への分析・変換、中間言語からの生成を通じて2ヶ国語間の翻訳を実現しようとする手法である〔260, 570〕。



しかし中間言語といっても意味表現にまで至ってわけではなく、語もそれぞれの言語でそれぞれに表現するのであって、統語構造が両言語の橋渡しになれるような形式を見出そうとしたものである。

CETAのこの方法は、中間言語に移すことによって原文にあった種々の情報を失ってしまつてうまく翻訳できない、システムが硬すぎて弾力性がない、などの欠点が明らかになった。1971年の段階でも正しく翻訳できるのは42%だったと報告されている〔260〕。

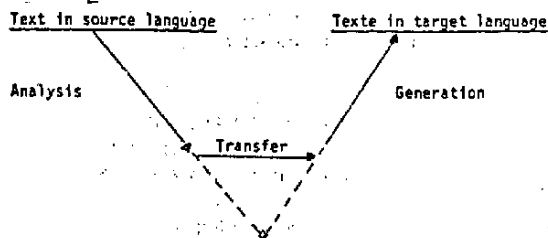
グルノーブル大学の計算機設備が1971年に更新されたのを機に、CETAはGETA（Groupe d'Etudes pour la Traduction Automatique）と改称し、また機械翻訳のアプローチを改めることになった。Pivot languageを介する方法をやめて、モントリオール大学のTAUMプロジェクトと同様の、分析・変換・生成の3段階から成る方法、すなわち第2世代の方法に切り換えた。これはpivot languageによる方法の行きづまりを打開するとともに、ヨーロッパのどの言語間の翻訳にも使えるシステムをめざしたものである。

GETAでは、処理の各段階に用いられるアルゴリズムは、それぞれの段階

に適合したものであるべきであって、必要以上に複雑であったり、強力であったりしてはならないという考えから、Q-systemやATNのような言語で一律に記述するのを排し、次のようないくつかの言語を開発した〔90,260,580,590〕。(図8-6, 図8-7参照)

- ① ATEF … 形態分析部門
- ② ROBRA … 統語分析部門および統語生成部門
- ③ TRANSF … 翻訳部門
- ④ SYGMOR … 形態生成部門

ATEFは、非決定性有限状態モデルに基いて、またSYGMORは、決定性有限状態モデルに基いて、記号列 木の変換を行う。ROBRAは、木 木の変換を行う。ATEFとROBRAには、無駄な木を生成を避ける為に、さまざまなヒューリスティックスを組み込めるようになっている。GETAのシステムは非常に柔軟で、人間が介入して対話的に分析、生成することも可能であるし、“real time knowledge”や“common sense knowledge”を組み入れることも不可能ではないとしている。



A translation system is a sequence of six models :

(1) morphological analysis, (2) multilevel analysis for the "analysis phase" (monolingual), (3) lexical transfer, (4) structural transfer for the "transfer phase" (bi-lingual), (5) syntactic generation, (6) morphological generation for the "generation phase" (monolingual).

図8-6 (〔590〕から引用)

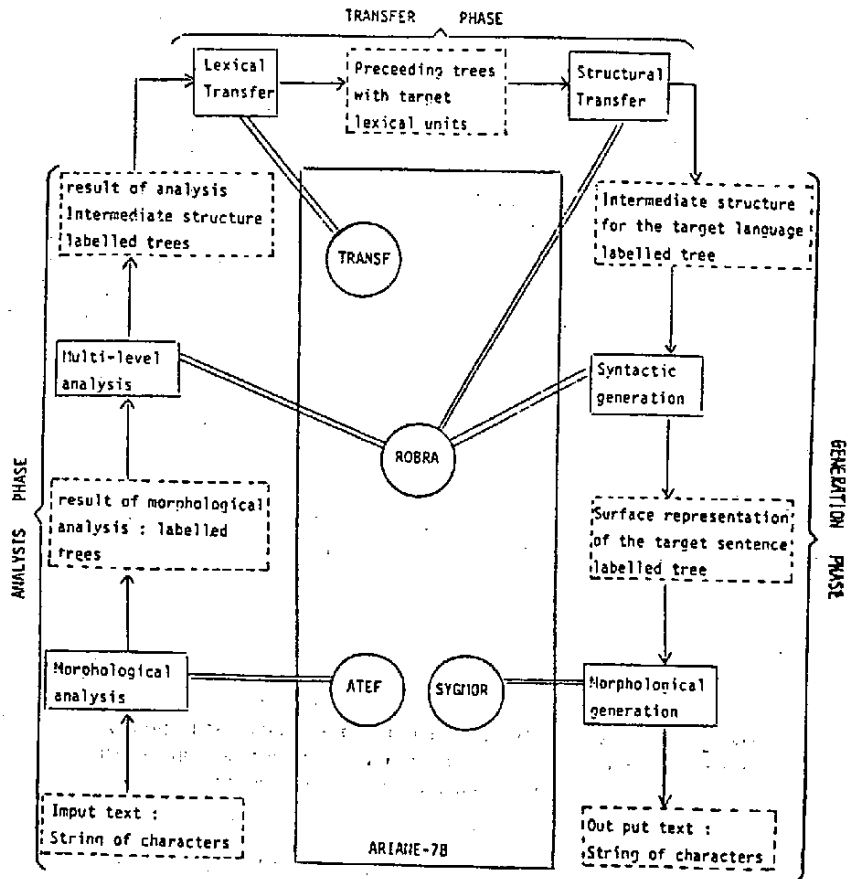


図 8 - 7 (〔590〕から引用)

8. 1. 2. 3. 2 意味理解の問題

かつて Bar-Hillel は、完全な自動翻訳の不可能性を主張して、意味理解の問題を提起した〔40〕。翻訳家は、背後にもっているいろいろな知識を駆使して文の意味を理解し翻訳という仕事を遂行している。しかし、人間のもっているこの知識というものは、形式化できない世界である。従って計算機で完全に自動的な翻訳をするということは不可能だという主張であった。

Bar-Hillelは次のような例を示している。

Little John was looking for his toy box.

Finally he found it. The box was in the pen.

John was very happy.

“ pen ” には 2 つの意味がある。

① play pen … おもちゃ箱

② writing pen … ペン (筆記具)

この場合には、もちろん①の意味に解すべきであるが、そのように解釈するためには現実の世界に関する「知識」を駆使しなければならない。これは toy box と pen の意味的な近さというような言語上の問題として扱うことはできない。なぜなら、たとえば inkstand と (writing) pen は 意味的に非常に近いが、

The inkstand is in the pen.

の場合には、やはり①の意味に解すべきである。

このようなことから、Bar-Hillel は「完全に自動的で高い品質の翻訳 (FAHQQT)」の不可能性を説いた。ところが一方で、「問題点が明確になった。『知識』を使って『理解する』ということが問題なのだ。この問題に挑戦しよう。」とする流れがあった。M. Minsky をはじめとする、いわゆる AI (Artificial Intelligence) 学派である。AI は言語の問題に限らず、広く人間の認識能力、問題解決能力、理解能力の問題、また『知識』の表現方法の問題など人間における情報処理に関心を寄せることになる。

機械翻訳の初期の「失敗」は、N. Chomsky 等の生成文法学派を中心とした統語理論の見直し・深化を促進した一方で、このように AI への関心を促進することになった。しかし AI における自然言語研究では、自然言語を人間と計算機とのインターフェイスとして用いようという応用とからめて、『理解』の証しを質問応答システムの構成に求める研究が多く、翻訳の問題がとりあげられることは少なかった。

しかし、翻訳できるためには、一定の『理解』が必要であることは明白であり、従って言語理解の研究が進めば、それは機械翻訳に影響を及ぼすであろうことは当然である。機械翻訳の第 3 世代というのは、このように、翻訳の背景

にある意味理解の問題に立入って自動翻訳のシステムを構想しようとするものである。

『理解する』という背景がなければどうしても解決できない事柄としては、Bar-Hillel が指摘したような多義語の問題のほか、代名詞などによる参照の関係を同定する問題、省略されている語を補う問題などが大きな問題である。

Y. Wilks は、彼のシステム NLUS (Natural Language Understanding System) で、英語からフランス語への翻訳を実験している〔630, 640, 650〕。Y. Wilks は、自動翻訳のためには『理解する』ことは確かに必要だが、しかし AI の分野で E. Charniak や R. Schank その他が展開しているほどに「深く理解する」必要はないだろうとみている。Y. Wilks の方法は一種の中間言語法であって、semantic template と呼ばれる意味表現を介して翻訳を実行している。

『理解』の問題に挑戦する AI 側からの様々のアプローチについては、別の章で詳しく論じられているので、ここでは述べない。

8.1.2.4 EC の多言語問題

EC では、EC 諸国語間の翻訳に対する需要が膨大なものになってきている〔490〕。EC 議会の予算の半分以上が、この多言語問題から発生する支出にあてられているという。EC 委員会の職員の 3 人に 1 人は、多言語問題の処理のための仕事に終日従事している。1977 年には委員会内部で行った翻訳が 50 万ページ外部に委託した分が 5 万ページもあったという。EC 参加国が増えれば、これはさらに組み合わせ論的に増えることになる。

EC 委員会では、1976 年に、EC 域内の情報流通を改善するための行動計画 (action plan) を策定した。行動計画の内容は、

① EURONET (European Information Network) を建設して、現在少数の人だけが使用可能なデータベースを多くの人に解放すること。

② 多言語問題を解決するためのさまざまな手段を見出すこと。

であった。もちろんデータベースも多言語的なものにしなければならず、①の課題は②の課題を惹起することになる。EC 委員会では、この行動計画を具体化し実行するための諮問委員会 CETIL を 1977 年に設立している。

具体的な動きとしては、

- ・EURODICAUTOMなど用語データバンクの整備
- ・SYSTRANの試用、改善
- ・新しい自動翻訳システム(EUOTRA)の計画

などが行なわれている〔490, 680〕。

8.1.3. 機械翻訳の今後

地球上の言語が統一されることはあり得ないし、また各言語地域がそれぞれに孤立化することもあり得ない。各地域内の社会的交流は確実に増大していく方向にあるが、言語のちがいはそれを妨げる障壁であり、従って言語の相互変換の必要性は確実に増大していく方向である。非常な発達を見せている現在の計算機技術を活用して、正しく、速く、しかも安く翻訳する技術を確立することは、依然として今後期待すべき課題である。

用語データバンクなどの翻訳支援技術は、有効性が確実な現実的技術的課題である。自動翻訳技術については、FAHQT(完全に自動的で高品質の翻訳)を期待することはできないが、言語理解研究のサブ・セットとして、言語をより深く体系的に把握することの上に成り立つ、今後期待すべき研究課題である。

附 録

① IBM-ジョージタウン・システム (露→英) の公開実験のサンプル ([1 0] から引用)

The quality of crude oil is determined by calory content.

The quality of saltpeter is determined by chemical methods.

TNT is produced from coal.

They obtain dynamite from nitroglycerine.

Ammonite is obtained from saltpeter.

Gasoline is prepared by chemical methods from crude oil.

They prepare ammonite.

Gasoline is produced by chemical methods from crude oil.

The price of crude oil is determined by the market.

Calory content determines the quality of crude oil.

TNT is prepared from coal.

② EURATOM のジョージタウン・システム (露→英) の出力例 ([1 0] から引用)

Biological experiments, which were conducted on different cosmic LETA-TEL6NYX APPARATI, the astrophysical investigations of cosmic space and the flights of Soviet and also American KOSMONAVTOV with the sufficient convincingsness showed; that the short-term orbital flights of below radiation belts of ground upon the absence of the increased solar activity in radiation relation are safe. Obtained by KOSMONAVTAMI of dose of radiation at the expense of initial cosmic radiation and the radiations of external radiation belt are so small, that cannot have harmful action on the organism of man.

In these experiments there were used mammals (dogs, mice, KRYSY, the maritime piglets), MUXI DROZOFILY, vegetable objects. The seeds of higher plants (wheat, the pea, LUK, pine, beans, REDIS, MORKOV6 etc.) MIKROSPORY TRADESKANQII, the culture of alga of chlorella on the different feed environments, numerous biological and QITOLOGICESKIE objects on TKANEVOM, cellular, SUBKLETOCNOM and molecular levels (Ghrdjan, 1962 and Antipov with Soavt 1962). In experiments on mammals special attention was devoted to the investigation of state of system of KROVOT-VORENI4, the determination of intermediate products the exchange of nucleinic acids (DEZOKSIQITIDINA and DIWEPOLOJITEL6NYX substances), the study of the state of natural IMMUNITETA The determination of content of SEROTONINA in blood. Besides this, there was conducted the check for the state or PIGMENTAQII the hair at black mice (the line C(57) Y) the Physiological shifts were studied also on the seeds of higher plants, microorganisms, the cells of the different tissues in culture and T D.

③ E CでのSYSTRANの試用例(英→仏)

……原文, なまの翻訳文, 後編集した文……([2 0] から引用)

THE ENGLISH ORIGINAL

Introduction

Before final implementation of Regulation (EEC) 2967/76 of 23rd November, 1976 on the water content of frozen and deep-frozen chickens, hens and cocks the question of the correlation between the two analytical methods for determining extraneous water in that regulation has been raised.

In order to compare the two analytical methods (annex III, the protein method, and annex IV the fat-free dry matter-method) the services of the EEC Commission asked two laboratories - Bundesanstalt für Fleischforschung, Kulmbach, Germany and the Danish Research Institute for Poultry Processing, Hillerød, Denmark - to carry out comparative analyses on deep-frozen chicken and hen carcasses with the two methods.

Assisting in planning and carrying out the study were also Dr. P. Stevens, Station de Recherches Avicoles - INRA, Nouzilly, France, and Dr. J. van Hoof, Laboratorium voor Hygiëne en Technologie van Eetwaren van dierlijke Oorsprong, Gent, Belgium.

The study was numbered P.200 by the EEC Commission services, and the actual work of the study was divided by the laboratories in the following way: All hens were slaughtered by the Gent Laboratory, and the hens were divided between Kulmbach and Hillerød. All chickens were slaughtered by the Hillerød institute and they were divided between Kulmbach and Hillerød.

This report describes the Danish part of the study for the chickens analyzed in Denmark. A report on the Danish analyses of the hens will form part II of this report.

Live chickens. The chickens for the study were taken from two small flocks of the same genetic origin reared in the Research Station for Poultry Breeding at Hillerød. They were at the time of slaughter (17th and 18th October, 1977) 39-40 days and 55-56 days respectively. They were White Plymouth Rocks. They have been given a feed with the following composition:

40% maize
20% barley
24% toasted soy bean cakes
3% meat and bone-meal
4% fish meal
3% oats
3% animal fat
3% minerals and vitamins

All had been starved for at least 12 hours before slaughter. On each of the two days 3 times 25 chickens (for three weight groups) were caught and transported to the near-by slaughterhouse (distance app. 500 meter). Each weightgroup should contain 24 chickens - the 25th being an extra, that was slaughtered but not used in the study.

Slaughtered chickens. All chickens were electrically stunned in a water stunner, killed and bled, but removed from the slaughter-line before the scalding tank. They were then hand-plucked and eviscerated. All organs and the neck were removed. They were dressed according to the definition: "plucked and drawn, without heads and feet, and without hearts, livers and gizzards, called 65% chickens", and with the removal of the neck.

The carcasses were then divided in two groups (12 from each weight class) and numbered by application of rubber rings around the thighs, the same number being applied to the two thighs of the same chickens. Those destined for immediate deep-freezing were given a number preceded by a Roman I, those destined for wet chilling with a number preceded by a Roman II.

THE RAW SYSTRAN OUTPUT

Introduction

Avant l'exécution finale du règlement (la CEE) 2967/76 de 23rd novembre 1976 sur la teneur en eau des poulets, ces poules et des coqs gelés et surgelés la question de la corrélation entre les deux méthodes analytiques pour déterminer de l'eau étrangère à ce règlement a été soulevée. Pour comparer les deux méthodes analytiques (annexe III, la méthode de protéines, et l'annexe IV la import-méthode sèche non grasse) les services de la Commission de la CEE ont demandé à deux laboratoires - la fourrerie Fleischforschung De Bundesanstalt, Kulmbach, l'Allemagne et l'institut de recherche danois pour la volaille traitant, Hillerød, Danemark - d'effectuer des analyses comparatives sur profond - des carcasses gelées de poule et de poulet avec les deux méthodes.

Aider à la planification et en effectuant l'étude étaient également dr. P. Stevens, poste De Recherches Avicoles - INRA, Nouzilly, France, et le sabot de dr. J. van, le camion anglais dierlijke Oorsprong, Gent, Belgique De Eetwaren de camion De Technologie de l'hygiène de Laboratorium voor.

L'étude a été numérotée P 200 par les services de la CEE de la Commission, et le travail réel de l'étude a été divisé par les laboratoires de la manière suivante : toutes les poules ont été abattues par le laboratoire De Gent, et les poules ont été divisées entre Kulmbach et Hillerod. Tous les poulets ont été abattus par l'institut De Hillerod et ils ont été divisés entre Kulmbach et Hillerod.

Ce rapport décrit la partie danoise de l'étude pour les poulets analysés au Danemark. Un rapport sur les analyses danoises des poules fera partie II de ce rapport.

Poulets vifs. Les poulets pour l'étude ont été pris de deux petits troupeaux de même origine génétique élevés dans le centre de recherches pour volaille élevant au Hillerod. Ils étaient lors octobre d'abattage (de dix-septième et dix-huitième, de 1977) 39-40 jours et de 55-56 jours respectivement. Ils étaient les rochers blancs Plymouth. Ils ont été donnés un fourrage avec la composition suivante :

maïs de 40 %

orge de 20 %

24 % ont grillé des gâteaux de haricot de soja

3 % viande et engrais d'os

farine de poissons de 4 %

avoine de 3 %

matière grasse animale de 3 %

minéraux et vitamines de 3 %

Tout avait été affamé pendant au moins 12 heures avant l'abattage. Sur chacun ces deux jours 3 fois 25 poulets (pour trois groupes de poids). Être recueilli et transporté à l'abattoir voisin (distance app. 500 compteurs). Chaque poids-groupe doit contenir 24 poulets - 25th être un figurant, qui a été abattu mais pas n'a pas été employé dans l'étude.

Poulets abattus. Tous les poulets ont été électriquement assommés dans une eau stunner, tués et bleded, mais enlevés de la abattage ligne avant le réservoir d'échaudage. Ils hand-plucked et ont été éviscérés alors. Tous les organes et le col ont été enlevés. Ils ont été habillés selon la définition "plume et tiré, sans ces têtes et des pieds, et sans des coeurs, foies et gésiers, appelés poulets de 65%, et avec l'élimination du col.

Les carcasses ont été alors divisées dans deux groupes (12 de chaque classe de poids) et numérotées par l'application des anneaux en caoutchouc autour des cuisses, le même nombre étant appliquée aux deux cuisses des mêmes poulets. On a donné ces destiné à la surgélation immédiate un nombre précédé par un Roman II.

THE SYSTRAN OUTPUT REVISED

Introduction

Avant la mise en oeuvre finale du règlement (CEE) n° 2967/76 du 23 novembre 1976 sur la teneur en eau des coqs, poules et poulets congelés ou surgelés, la question de la corrélation entre les deux méthodes d'analyse proposées dans ce règlement pour déterminer la teneur en eau étrangère a été soulevée.

Pour comparer les deux méthodes d'analyse (annexe III, méthode des protéines, et annexe IV, méthode des matières sèches non grasses), les services de la Commission de la CEE ont demandé à deux laboratoires - la Bundesanstalt für Fleischforschung de Kulmbach, Allemagne, et le Danish Research Institute for Poultry Processing de Hillerød, Danemark - d'effectuer des analyses comparatives sur des carcasses surgelées de poulets et de poules suivant les deux méthodes.

Ont également contribué à programmer et à réaliser cette étude le dr. P. Stevens, Station de Recherches Avicoles - INRA, Nouzilly, France et le dr. J. van Hoof, Laboratorium voor Hygiëne en Technologie van Eetwaren van dierlijke Oorsprong, Gand, Belgique.

Les services de la Commission de la CEE ont donné à cette étude le numéro P 200 et les laboratoires ont réparti de la manière suivante les travaux à effectuer pour réaliser celle-ci: toutes les poules ont été abattues au laboratoire de Gand et elles ont été réparties entre Kulmbach et Hillerød. Tous les poulets ont été abattus à l'institut de Hillerød et ils ont été répartis entre Kulmbach et Hillerød.

Le présent rapport décrit la partie danoise de l'étude effectuée sur les poulets analysés au Danemark. Un rapport sur les analyses réalisées au Danemark sur les poules fera l'objet de la partie II de ce rapport.

Poulets vivants. Les poulets qui ont fait l'objet de cette étude ont été choisis dans deux petites basses-cours de même origine génétique et ils avaient été élevés au centre de recherches pour l'élevage de la volaille de Hillerød. Lorsqu'ils ont été abattus (les 17 et 18 octobre 1977) ils avaient respectivement 39-40 jours et 55-56 jours. Il s'agissait de White Plymouth Rocks. On leur avait donné un fourrage dont la composition était la suivante :

maïs: 40%
orge: 20%
tourteaux de soja grillé : 24%
viande et engrais d'os: 3%
farine de poisson: 4%
avoine: 3%
graisses animales: 3%
minéraux et vitamines: 3%

Ils avaient tous été privés de fourrage pendant 12 heures au moins avant d'être abattus. Pendant chacun de ces deux jours, 3 lots de 25 poulets (correspondant à trois catégories de poids) ont été constitués et transportés à l'abattoir qui se trouvait à proximité 500 mètres environ. Chaque lot devait comporter 24 poulets; le 25ème étant en réserve, a été abattu, mais il n'en a pas été tenu compte dans l'étude.

Poulets abattus. Tous les poulets ont été assommés électriquement dans un réservoir d'eau, tués et saignés, mais retirés de la ligne d'abattage avant le bac d'échaudage. Ils ont été plumés à la main, puis éviscérés. Tous les organes et le cou ont été enlevés. Ils ont été préparés selon la formule "plumés et parés, sans les abats (tête, pattes, coeur, foie et gésier), appelés poulets 65%", et sans cou.

Les carcasses ont alors été réparties en deux groupes (12 de chaque catégorie de poids); par ailleurs, on les a numérotées au moyen d'anneaux en caoutchouc qui leur ont été passés autour des cuisses, le même nombre figurant sur les deux cuisses d'un même poulet. Pour celles destinées à la surgélation immédiate, ce nombre était précédé d'un I romain; pour celles destinées à la réfrigération humide, celui-ci était précédé d'un II romain.

④ T A U M — Project aviation の実験例 (英→仏)

([2 6 0] から引用)

[The utility system supplies half the hydraulic power to operate the ailerons, rudder, and horizontal tail flight controls, and all the power to operate the landing gear, speed brakes, wheel brake system, nose wheel steering, stability augmentor system, gun gas deflectors, gun gas purging system operation and the two-position nose gear strut. The flight control system supplies the remaining half of the hydraulic power to operate the ailerons, rudder and horizontal tail flight controls . . .

In addition to periodic replacement, the flight control and utility system filters in the fuselage aft section should be replaced whenever the hydraulic systems are suspected of being contaminated.]

Le circuit de servitudes fournit la moitié de l'énergie hydraulique pour actionner les ailerons, le gouvernail de direction, et les plans de l'empennage horizontal, et toute l'énergie pour actionner le train d'atterrissage, les aéro-freins, le circuit du servo-amortisseur, les déflecteurs de gaz du canon, le fonctionnement d'évacuation de gaz du canon et la contrefiche à deux positions du train avant. Le circuit de commande de vol fournit l'autre moitié de l'énergie hydraulique pour actionner les ailerons, le gouvernail de direction et les plans de l'empennage horizontal . . .

En plus du remplacement périodique, la commande de vol et le circuit de servitudes filtre dans la partie arrière du fuselage devrait être remplacé chaque fois que les circuits hydrauliques sont soupçonnés d'être contaminé.

⑤ T A U M - M E T E O の翻訳例 (英→仏) ([2 6 0] から引用)

[LOWER ST JOHN RIVER VALLEY
UPPER ST JOHN RIVER
WIND WARNING ENDED BOTH REGIONS.
SNOW AND BLOWING SNOW TONIGHT
BECOMING INTERMITTENT NEAR DAWN
FRIDAY. CLOUDY WITH PERIODS OF
LIGHT SNOW FRIDAY. STRONG GUSTY
NORTHEASTERLY WINDS TONIGHT
BECOMING NORTHWESTERLY WINDS
FRIDAY AFTERNOON.]

VALLÉE DU BAS ST JEAN
HAUT ST JEAN
FIN DE L'AVIS DE VENT POUR LES DEUX
RÉGIONS. CETTE NUIT NEIGE ET
POUDRERIE DEVENANT PASSAGÈRES
VENDREDI À L'AUBE. VENDREDI NUAGEUX
AVEC FAIBLES CHUTES DE NEIGE
PASSAGÈRES. CETTE NUIT VENTS FORTS
DU NORD EST SOUFFLANT EN RAFALES
DEVENANT VENTS FORTS DU NORD OUEST
VENDREDI APRÈS-MIDI.

⑥ C U L T の出力例 (中→英) ([4 3 0] から引用)

ON THE TRANSGRESSION OF CHARACTERISTIC FORMS

WU GUANG-LEI

(PEKING UNIVERSITY)

(RECEIVED DEC. 5, 1974)

FIBRE BUNDLE THIS KIND STRUCTURE, HAS THE FUNDAMENTAL IMPORTANT
SIGNIFICANCE IN DIFFERENTIAL GEOMETRY, WHICH NOW LOOKS VERY OBVIOUS.
THIS VIEWPOINT IS RAISED BY CHERN SHEENG-SHEN IN THE 40S, FIRST
EXPRESSED IN DETAIL IN A SHORT PAPER⁽¹⁾ WRITTEN BY HIM CONCERNING THE Gauss-
Bonnet FORMULA IN 1944; LET X BE A COMPACT RIEMANNIAN MANIFOLD, THE UNIT
VECTOR ON X FORM A SPHERE BUNDLE Y . THE "TOTAL CURVATURE" OF
RIEMANNIAN MANIFOLD X CAN BE EXPRESSED AS A CLOSED DIFFERENTIAL FORM Δ
ON X . THE COHOMOLOGY CLASS OF Δ CORRESPONDS TO Euler CHARACTERISTIC
NUMBER. CHERN FIRST POINTS OUT, THAT THERE EXISTS A DIFFERENTIAL FORM Π
ON THE SPHERE BUNDLE Y SUCH THAT $d\Pi = \Delta$, AND Π RESTRICTED ON THE
FIBER THEN SHOWS THE FUNDAMENTAL CLASS OF THIS FIBER. THIS PROPERTY OF
THE DIFFERENTIAL FORM Δ IS THEN CALLED THE TRANSGRESSION, Π IS CALLED
THE TRANSGRESSION FORM OF Δ . THE MAIN POINT IN THE PROOF OF CHERN
SHEENG-SHEN IS TO CONSTRUCT A TRANSGRESSION FORM IN DETAIL. HERE IT
FURTHER SHOWS ALSO THE APPLICATION AND POWER OF E. Cartan METHOD. IN 1959, Ellis
GENERALIZED Gauss-Bonnet FORMULA TO Stiefel-Whitney CHARACTERISTIC CLASS⁽²⁾.

IN 1932, ПОНТРИГИН GENERALIZED GAUSS SPHERICAL MAPPING, INTRODUCED ПОНТРИГИН CHARACTERISTIC CLASS^[6]. HE PROVED IN 1944 THAT ПОНТРИГИН CHARACTERISTIC CLASS OF REAL COEFFICIENTS CAN BE EXPRESSED BY THE POLYNOMIAL OF RIEMANNIAN CURVATURE^[7].

IN 1946 CHERN SHEENG-SHEN INTRODUCES CHERN CHARACTERISTIC CLASS^[8] ON THE COMPLEX MANIFOLD, AND ALSO PROVED THAT THE TRANSGRESSION IS ALSO THE CHARACTERIZATION OF THIS CHERN CLASS. LATER THIS THEORY IS GENERALIZED TO THE GENERAL FIBRE BUNDLE^[9].

参 考 文 献

- [10] ALPAC, "Language and Machines, Computers in Translation and Linguistics", National Academy of Sciences, 1966.
- [20] Arthern.P.J, "Machine Translation and Computerized Terminology Systems - A Translators View Point", in [530].
- [30] Bar-Hillel.Y, "Language and Information, Selected Essays on their Theory and Application", Addison-Wesley: Massachusetts, 1964.
- [40] Bar-Hillel.Y, "A Demonstration of the Nonfeasibility of Fully Automatic High Quality Translation", in [30].
- [50] Borko.H (Ed.), "Automated Language Processing", John Wiley and Sons: New York, 1968.
- [60] Brinkman.K.H and Tanke.E, "TEAM: A program System for the solution of Terminological and Lexicographical Problems", in [70].
- [70] Brinkman.K.H et., "Machine-assisted Translation in West Germany", JPRS 68726, 1977.
- [80] Bruderer.H.E, "The Present State of Machine and Machine-assisted Translation", in [160].
- [81] Bruderer.H.E, "Handbook of Machine Translation and Machine Aided Translation", North-Holland: Amsterdam, 1978. (?)
- [90] Cauche.J, "The ATEF and CETA System", Am. J. Comput. Linguist, microfiche 17, 1975.

[91] Chafe.W, "An Approach to Verbalization and Translation by Machine", Am. J. Comput. Linguist., microfiche 10, 1975.

[100] Chaloupka.B and Ghusnoli.G, "On Xonics System", in [250].

[110] Chandioix.J, "Creation of a Second Generation System for Machine Translation of

Technical Manuals", in [160].

[120] Chandioix.J, "METEO, An Operational System for the Translation of Public Weather

Forecasts", in [250].

[130] Chevalier.M, Danserau.J and Poulin .G, "TAUM-METEO, Description du systeme", Universite de Montreal, 1978.

[140] Chomsky.N, "Syntactic Structure", The Hague: Mouton & Co., 1957.

[150] Chomsky.N, "Aspects of the Theory of Syntax", M.I.T. Press, 1965.

[160] Commission of the EC, "Overcoming the Language Barrier, Third European Congress on Information Systems and Networks", 1977.

[161] Danserau.F.J, "Automated Language Processing", Annual Review of Information Science and technology, 1976.

[170] Dostert.L, "The Georgetown-I.B.M. Experiment", in [340].

[180] Dr.F.B, "The Lexicography Information System (LEXIS) of the Bundeswehr Language Service", in [180].

[190] Elliston.J, "Computer Aided Translation - A Buisness View Point", in [530].

[200] Gerald.M and Hutton.F, "Georgetown University MT System Usage, Nuclear Division, Union Carbide Corp., Oak Ridge", in [250].

[210] Gleibach.S, "A New Normal_form Th

eorom for Context_free Phrase Structure Grammer"
", J.Assoc.Comput.Mach., 12_1, 1965.

[220] Goetschalckx.J, "Terminological Activities in the European institutions, with Special Reference to EURODICAUTOM", in [160].

[230] Goetschalckx.J, "Eurodicautom",
in [530].

[240] Goshwake.W, "SLUNT (Spoken Languages Universal Numeric Translation)", in [160].
")

[250] Hays.D and Mathias.J (Eds.), "FBIS Seminar on Machine Translation", Am.J.Comput.

Linguist, microfiche 46, 1976.

[260] Hutchins.W.J, "Progress in Documentation, Machine Translation and Machine_aided Tr

anslation", J.Document., 34-2, 1978.

[261] Jordan.S.R, Brown.A.F.R and Hutton.F.C,
"Computerized Russian Translation at ORNL",
J. Am. Soc. Inf. Sci., 1, 1977.

[270] Kay.M and Bisbey.R, "The MIND Translation System --- A Study in man-machine Collaboration", NTIS AD 749000, 1972.

[280] Kay.M, "The MIND System", in [480]

[281] Kay.M, "The Proper Place of Man and Machines in Translation", in [250].

[281] Kertesz.F, "How to Cope with the Foreign language Problem: Experience Gained at a multi disciplinary Laboratory", J. Am. Soc. Inf. Sci.,

1-2, 1974.

[290] Knowels.F, "Error Analysis of System Output - A Suggested Criterion for the Internal Evaluation of Translation Quality and a Possible Corrective for System Design", in [530].

[291] Krollman.F, "User Aspects of an Automatic Aid to Translation as Employed in a Large Translation Service", in [160].

[300] Kuno.S, "Computer Analysis of Natural Languages", Proc.Symp.Appl. Math., 19, 1967.

[310] Laurent.J, "Utilization of the Technical Terminology Standardised at AFNOR (Association française de Normalization)", in [160].

"

[320] Lippman.E, "On-line Generation of Terminological Disests in Language Translation, An Aid in Terminology Processing", IEEE Trans. Prof. Commun., 18-4, 1975.

[330] Lippman.E, "An Approach to Computer-aided Translation", IEEE Trans. Eng. Writ. Speech, 14-1, 1971.

[340] Locke.W.N and Booth.A.D (Eds.), "Machine Translation of Languages", Greenwood Press: Connecticut, 1955.

[350] Locke.W.N and Booth.A.D, "Historical Introduction", in [340].

[360] Loh.S.C and Koss.L, "An Interactive On-line Machine Translation System (Chinese

into English)", in [530].

[370] Loh.S.C and Koss.L, "Computer Translation of Chinese Scientific Journals", in (22).

[380] Loh.S.C, "CULT (Chinese University Language Translator)", in [250].

[390] Lytle.E, "Automatic Language Processing Project", in [250].

[400] Maas.H.D, "The Saarbrücken Automatic Translation System (SUSY)", in [160].

[410] Marchuk.Yu.N, "Machine Translation in the All-Union Translation Center", in [160].

[420] Marten'janov.Yu.S, "Eliminating Formal Constraints of Language", in [160].

[430] Masterman.M, "The Essential Skills to be Acquired for Machine Translation", in (24).

[440] Mathias.J, "Chinese-English Translation Assistance Group", in [250].

[450] Moscovoj.V.A, "On the Elimination of Formal Word Order in the German Sentence", in [160].

[460] Dubine.I.I, "Machine Translation --- Moscow International Seminar", Am. J. Comput. Linguist., microfiche 51, 1976.

[470] Pendersraft.E.D, "Translating Languages", in [5].

[471] Petit.A.J, "Objective Evaluation

Criteria for Machine Translation", in [160].

[472] Raskin.V, "A Restricted Sublanguage Approach to High Quality Translation", Am. J. Comput. Linguist., microfiche 9, 1975.

[473] Rozencveis.V.J (Ed), "Machine Translation and Applied Linguistics", Athenaum: Frankfurt, 1974. (!)

[480] Rustin.R (Ed), "Natural Language Processing", Algorithmic Press: New York, 1973.

[490] Saser.J.C, "Multilingual Communication: Chairman's Introductory Review of Transl

ating and the Computer", in [530].

[500] Schank.R.C and Colby.K.M, "Computer Models of Thought and Languages", W.H. Freeman & co.: San Francisco, 1973.

[510] Schultz.J, "A Terminology Data Bank for Translators - Methods of Interosation in the TEAM System", in [160].

[520] Schultz.J, "Lexicography with TEA
M -- Automatic Dictionary Composition", in [7
01].

[530] Snell.B.M (Ed.), "Translating and
the Computer", North-Holland, 1979.

[540] Tanke.E.H, "Implementing Machine
Aids to Translation", in [530].

[550] Toma.P, "Systran", in [250].

[560] Toma.P, "Systran as a Multilingual
Machine Translation System", in [160].

[571] Troike.R.C, "The Future for MT",
Am. J. Comput. Linguist., microfiche
51, 1976.

[570] Vauquois.B, "Intermediate Language
e for Automatic Translation", in [480].

[580] Vauquois.B, "An Actual Large Scale
e Project of MT --- The European Community",
Comput. Linguist. (Inf. Process. Soc. Japan), 18
, 1979.

[590] Vauquois.B, "Discussion about Sof
tware Tools for MT and Natural Language Process
ing", Comput. Linguist. (Inf. Process. So
c. J
apan), 18, 1979.

[600] Wang.S.Y, "Chinese-English Machin
e Translation Project on Linguistic Analysis" "
, in [250].

[610] Weaver.W, "Translation", in [340].
")

[620] Weisenborn.J, "The Role and Form
of Analysis in Machine Translation --- The Auto
matic Analysis of French at Saarbrücken", in [160].

[630] Wilks.Y, "An Artificial Intelligence Approach to Machine Translation", in [500].

[640] Wilks.Y, "Machine Translation and Artificial Intelligence", in [530].

[650] Wilks.Y, "The Stanford Machine Translation Project", in [480].

[651] Young.M.E, "Machine Translation (A Bibliography with Abstracts)", NTIS Document PS-78/0448, 1978.

[651] Zachary.W, "Machine Aids to Translation --- A Concise State of the Art Bibliography", Am. J. Comput. Linguist., microfiche 77, 1978.

[660] Zarechnak.M, "Russian-English System, Georgetown University", in [250].

[6 7 0] 久野 暲, 「ハーバード大学の言語自動処理研究」, 計量国語学, 32, 1965

[6 8 0] 長尾 真, 「ヨーロッパにおける機械翻訳の現状」, 計量言語学研究会資料, 20, 1979

8.2 質問応答システム

8.2.1 はじめに

1970年に入り、自然言語理解システムの研究が各所でとり上げられたことは第1章の1.2節と1.3節で説明した。最近の言語理解システムの研究は、i) 対象世界の拡大、ii) 知識表現形式の検討、iii) 実用化可能性の検討、の方向で進められている。自然言語による質問応答システムは、iii) に関係している。

計算機に自然言語を理解させることが、困難なものであることは、よく知られている。そこで現在では、非常に狭い問題分野、たとえばデータ・ベース問い合わせシステムや、医療診断システム、またCAI(教授)システムのフロント・エンドとして自然言語を用いる方向が模索されている。欧州などで本格化しようとしている機械翻訳システムの実現も、当面、問題分野を限定して追求することになるだろう。問題分野を限定することにより、質問応答に用いる自然言語に対して、

(i) 使用語彙の制限

(ii) 構文パターンの制限

(iii) 意味的、語用論的知識の制限

などを課すことが可能であり、計算機による自然言語処理を容易にすることができると考えられている。^{*}一方、自然言語処理技術そのものをめぐって1970年代に、

(iv) 言語学の進展

(v) 計算機の性能向上とコスト・ダウン

(vi) 自然言語処理のための種々の技法の提案と試み

(vii) データ・ベース技術などのソフトウェア技術の進展

などがあり、米国、西欧、日本等で、自然言語による質問応答システムの実現が試みられた。(iv)、(vi)は、本ワーキング・グループ調査報告書中の他の章で、また(v)と(vii)については他のワーキング・グループ調査報告書中で、詳しく説明されるだろう。本節では、質問応答システムの動向を実例をまじえながら説明し、将来解決すべき問題点を探ることとする。

なお2節以下の本文中で〔Petrick, 1976〕のようにして引用してある文

* これに対する批判もあるが8.2.3で述べる。

献は、第1章の末尾にあげた参考文献リストに記載されているものを指す。

8.2.2 質問応答システムのモデル

質問応答システムは、次の4つのプログラムから構成されていると考えられる。

- (i) 与えられた文(入力文)を構文解析するプログラム
- (ii) 構文解析結果に意味解釈を施し、意味構造を抽出するプログラム
- (iii) 抽出された意味構造に推論を加えて問題解決を行うプログラム
- (iv) 問題解決結果、もしくは意味構造を線条化して、文(答)を生成するプログラム

ただし、(i)と(ii)の厳密な区分が難しい場合がある。たとえば、Schank等は、(i)と(ii)が融合し、しかも、そこでは構文レベルの情報よりも意味レベルの情報を優先させるべきであると主張している。一方、このような意味主導型の方式に対して、(i)と(ii)を完全に融合すべきでなく、むしろ、(i)と(ii)をゆるく結合した構成にすべきであると考え人もいる。図8-8は、後者の考えに従って、質問応答システムの概念図を示したものである。

⑤の語用論的解析は、主として文脈(context)に依存した処理を行うもので、まだその理論、手法が、言語学的にも十分解明されていない。しかしスムーズな質問応答を行うためには語用論的な現象が、さらに解明されなければならない。これについては、6.2で再びふれる。

図8-8のモデルでは、処理のレベルの深さにもよるが、⑥と⑦は、言語の種類によらない部分であると考えられる。

図8-8で、⑥、⑦、⑧をスキップして④と⑨を直結し、文の生成で、入力文と異種の言語の文を生成すれば、機械翻訳システムが実現することになる。このように、質問応答システムと機械翻訳システムの研究とは排反するものではない。

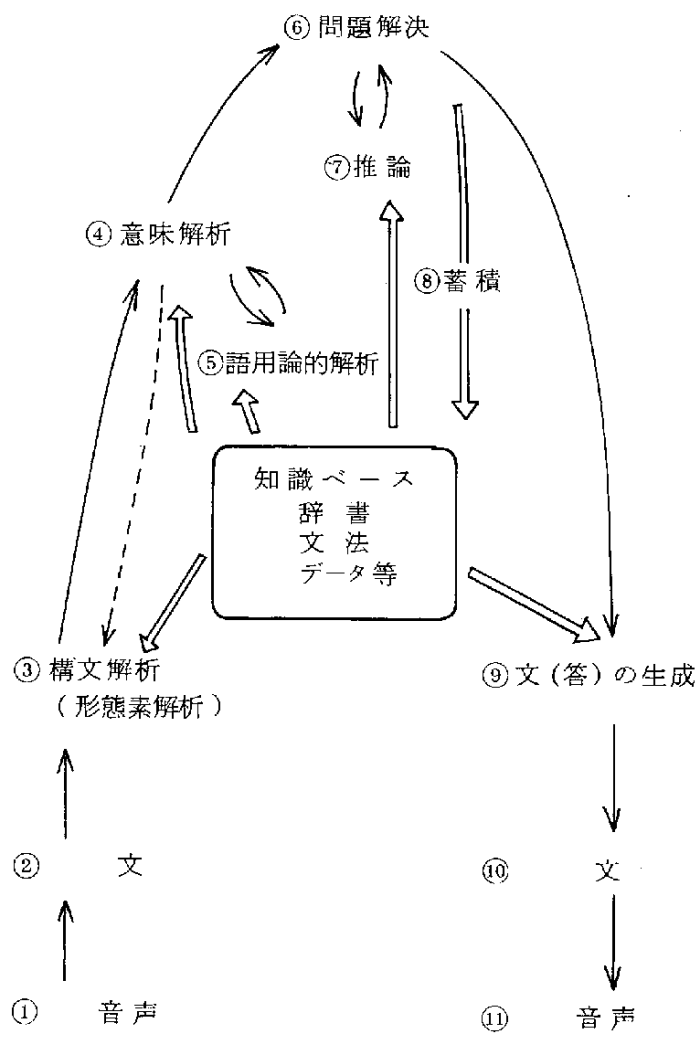


図 8 - 8 質問応答システムのモデル

8.2.3 質問応答と自然言語, QBE, DEDUCE

質問応答を行なうために、自然言語を用いることは、もしそれが可能であるなら、理想であろう。さらにいえば、音声によって質問応答が可能になることだろう。しかし、現在完全に実用の域に達した(自然言語による)質問応答システムはまだない。実験システムについての評価が昨年あたりから目につくようになってきた。これについては6節で説明する。ここでは、質問応答システムのインターフェイスとして自然言語を用いることが、現段階でどのような利点があるか、またそれに対してどのような反論があるかを〔Petrick, 1976〕に従って考えてみよう。

〔Petrick, 1976〕によれば、自然言語を質問応答システムのインターフェイスとして用いることの利点として、

- (i) 多くの(潜在的な)計算機ユーザは、人工言語を習得したり使用したがない。
- (ii) 少なくともある特定の計算機応用分野では、自然言語は通信手段として理想的なものである。
- (iii) 現在の自然言語処理技術をもう少し発展させれば、自然言語による質問応答システムが近い将来実現可能であろう。

したがって、

- (iv) 現在の自然言語処理技術で解決できない問題はあるが、それ以外の問題はできるだけ早く解決するよう努力すべきである。

(i)に対しては、次のような反論がある。

- (i)¹ 真の意味での完全な自然言語を計算機が扱えるようになるとは思えない。そこで、自然言語のサブセットを想定することになるだろうが、これは形式的な人工言語を習得し使用する以上に難しい。なぜなら、我々は、日常生活でサブセットではない自然言語の使用に慣れすぎているからである。

さらに、(ii)に対して、

- (ii)¹ 自然言語は本質的にきちんとしたものでなく、不明瞭で曖昧なもので、計算機言語として不適當である。そのため、計算機との通信に自然言語を用いると、処理の効率を低下させたり、誤りや誤解を引き起こす。

- (i)¹ は、確かにそうした側面があることは認めなければならないと思われる。

ただ筆者の乏しい経験によれば、しばらく特定の形式的な人工言語を使用しないと、その一部を忘れることがある。これは、使用する人工言語の数が増えれば増えるほど、よく経験することであろうと思われる。ところが、現在どのようなことを自分がしたいかは、日常の言葉（自然言語）で言い表わすことは可能である。

確かに、自然言語のサブセットの境界を明確に記憶しておくことは難しい。しかし、質問応答の対象問題領域を限定することにより、我々が質問応答に用いる自然言語は、自づと一定のサブセットに落ち着くのではなかろうか。無論これについては、人により、また同一人であっても時間により変動するので、質問文に対するデータを多量に収集蓄積しなければならない。我が国では、自然言語による実験的な質問応答システムが手軽に使用できる状態になく、経験を蓄積しにくい状況にあるので、特にそれを強調しておきたい。

(ii)¹ に関しては、“自然言語は、本質的にきちんとしたものではなく…”というくだりには同意しかねる。自然言語が不明瞭で曖昧なものである、という側面は確かにある。しかし政治家の国会答弁や外交書簡の解釈などを除けば、我々は、ほとんど曖昧さなく、意図していることを他人に伝達することができる。

(i)¹, (ii)¹ に関連した議論をもう少し精密化するために、データ・ベースに対する質問用言語として提案されている DEDUCE, QBE を自然言語と対比させて考察してみる。データ・ベースは、次の関係から構成されているものとする。

```
EMP (NAME, SAL, MGR, DEPT)
SALES (DEPT, ITEM, VOL)
SUPPLY (COMP, DEPT, ITEM, VOL)
LOC (DEPT, FLQQR)
CLASS (ITEM, TYPE)
```

これらに対して自然言語（英語）と DEDUCE による質問の例を図 8-9 に示す。

- (1) Find the volume of guns sold by the toy department.
- (1)' SALES(*vol, dept = toy, item = gun)
- (2) Find the salary of Anderson's manager.
- (2)' EMP(name = Anderson, mgr = x)
&
EMP(name = x, *sal)
- (3) How many type A items ?
- (3)' COUNT(CLASS(#item, type = A))
- (4) Find the names of those employees who make more than any employee in the shoe department.
- (4)' Vy(EMP(name = y, sal = x, dept = shoe)
→ EMP(*name, sal > x)

図 8 - 9 英文と DEDUCE による質問文
(〔Chang, 1978〕)

図 8 - 9 から次のことがわかる。

- (a) 英語の質問文の場合には、データ・ベースがどのような関係から構成されているかを気にする必要がない。
- (b) DEDUCE による質問では、データ・ベースがどのような関係から構成されているかを事前に知る必要がある。
- (c) DEDUCE による表現では、述語論理の知識が必要である。

(a) は、自然言語による質問の利点の一つにあげることができよう。しかし質問者がシステムに慣れて熟練化すると、自然言語ではなく、それらを簡略化した形式の質問文を用いたくなる。このあたりの事情はプログラミング技法のマクロ命令の使用と似たところがある。また、熟練した質問者は、使用しているデータ・ベースが、どのような構成になっているかを知りたがる傾向があるといわれている〔Harris, 1979〕。それにより、適切な質問文を発し、質問応答時間を短縮できる可能性があるからである。

データ・ベースがどのような関係から構成されているかが分かっている場合には、DEDUCEよりいまだ少しスマートな（図表言語とでもいう）言語が開発されている。これはQBE(Query By Example)とよばれるものである。たとえば、図8-9の質問文(1)（(1)）に対してQBEでは、下図のような枠を表示する。

これに対して

SALES	DEPT	ITEM	VOL

SALES	DEPT	ITEM	VOL
	TOY	GUN	P. VOL

とすると

VOL
17

が出力される。

QBEは、自然言語による問い合わせと同様に、マンマシン・インターフェイスをかなり改善する。QBEに対するコメントとしては、

(d) QBEによる質問ではデータ・ベースがどのような関係から構成されているかを事前に知る必要がある。

したがって、

(e) DEDUCE, QBE共、データ・ベースが多数の関係から構成されるに従って、熟練した使用者であっても使い勝手が悪くなる。

しかし、

(f) データ・ベースの構成が単純で、ある程度訓練を積んだ使用者ならQBEの方が自然言語よりも簡潔で使い易いかも知れない。

ここで、自然言語による質問による利点として(g)と(h)をあげておく。

(g) 自然言語の場合には、指示代名詞や連体詞等を用いて、既出のもの（文、名詞）を指しながら質問できる。

(h) 述語論理やそれに近いDEDUCEまたはQBEでは、既出の質問内容（もしくはその一部）を参照する表現法が貧困である。

(g)の機能は、文脈依存的(context sensitive)なもので、一文を越えた談話のレベルの機能の一つであり、我々の自然言語による対話を著しく柔軟なものにする。

以上をまとめると次のようになるだろう。

理想をいえば、自然言語(もしくは音声)を質問言語として用いることであろう。しかしこのことはDEDUCE, QBEを排除するものではない。熟練した使用者なら、これらを用いて効果的、効率的な質問応答を行なうことが可能な場面もあり得るからである。そのためには、データ・ベースがどのような関係から構成されているかを知ることと人工言語を習得することが前提になる。ただ、これらに熟知した使用者であっても、長期間使用しないと、その一部を忘却する可能性が十分にある。したがって、ここでも、自然言語によるインターフェイスが望ましい。

8.2.4 自然言語による質問応答の実例検討

本節で、いくつかの質問応答の実例を示し、自然言語による質問応答システムが、どのようなものを目ざしているかを調べてみよう。ただ、以下にあげる実例をみると、すでに相当複雑な自然言語が処理されているという気もするが、おそらく、それらは最善の模範応答例であり、この模範例から少しはずれた質問には、システム例が皆目理解を示さない、ということは十分ありうる。このような現在の自然言語による質問応答システムの問題点については8.2.6節でまとめて説明する。

(A) (Carbonell, 1970) (MIT)

SCHOLAR

(BBN)

南米の地理についての知識を学生に教授するシステム(CAI)*

アルゼンチンの言葉は何ですか？

恐らくスペイン語です。

非常によろしい。

アルゼンチンの広さは、およそどのくらいですか？

ちょっとペルーについて教えて下さい。

ペルーは国です。

南米にあります。

ペルーの首都はリマです。

ところで、あなたは前の質問にまだ答えていません。答えて下さい。

アルゼンチンの広さは1,000,000平方マイルです。

だいたい合っていますが、正確な答は1,200,000平方マイルです。

ペルーの東の国境に接している国は……

カトリックです。

違います。カトリックは国ではなくて宗教です。正しい答はブラジルです。

ペルーはどこにありますか？

アスンシオンは暑いですか？

アスンシオンはパラグアイの一部ですから暑いですよ。

ペルーとパラグアイで最も重要な相違点をあげて下さい。

基本的な相違点は、ペルーの人口はパラグアイの7.8倍だということです。

SCHOLAR は、高校程度の南米の地理に関する質問応答(CAI)システムである。このシステムでは、図8-8でいえば⑥と⑦の部分に重点をおいており、最近人間がどのような推論機構を用いて問題解決を行なっているかの研究に発展してきている。特徴をあげると、

- (a) 人間が会話の主導権を握ることが可能。(従来のものは、システムが一方向的に提示する質問に人間が答える方式であった。)
- (b) 知識の階層的構成法、知識のフレーム化。
- (c) 質問に対する答えが陽(explicit)に存在しない場合、適切な答を合成する推論能力をもつ。

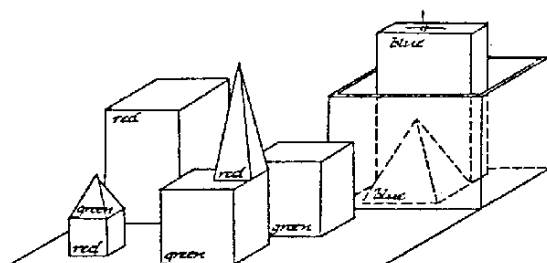
* 瀬一博監訳、ボブrow&コリンズ編：“人工知能の基礎” 近代科学社 “より抜粋

- (d) 形態素レベルでの誤入力を修正する機能をもつ。
- (e) 複雑な構造の文は扱えない。

(B) [Winograd, 1971] (MIT)

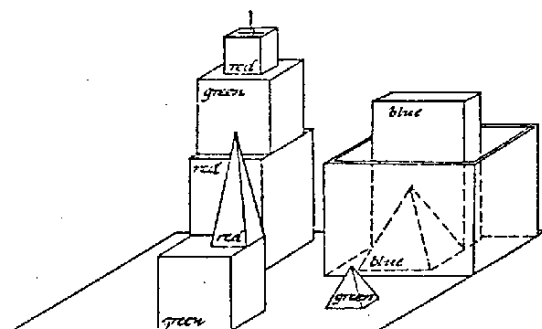
SHRDLU

ロボットに指令を与えて動作させるシステム



H1 will you please stack up both of the red blocks and either a green cube or pyramid?

M1 OK.



H2 which cube is sitting on the table?

M2 THE LARGE GREEN ONE WHICH SUPPORTS THE RED PYRAMID.

H3 is there a large block behind a pyramid?

M3 YES, THREE OF THEM: A LARGE RED ONE, A LARGE GREEN CUBE AND THE BLUE ONE.

H4 put a small one onto the green cube which supports a pyramid.

M4 OK.

(Hi は人, Mi はシステムからの応答)

SHRDLUは、コンピュータがかなり複雑な言語理解を行なうことを実例で示した。その後、SHRDLUの限界に挑戦することから、次の世代の（質問応答システムを含む）言語理解システムの建設が開始されたといえる。特徴をあげると、

- (a) 対象世界を積木の世界に限定。
- (b) 推論は人工知能向き言語 μ PLANNERを用いて記述。
- (c) 手続き的な知識表現により効率的な質問応答を達成。
- (d) Hallidayのシステミック文法を使用。
- (e) 答の理由を説明することが可能。

(C) [Brown, 1975] (BBN)

SOPHIE

回路図を与え、それに故障要因を付加し、学生に故障を発見させるプログラム(CAI)^{*}

WELCOME TO SOPHIE. A REACTIVE LEARNING ENVIRONMENT.
THE PIECE OF EQUIPMENT YOU ARE STUDYING IS A REGULATED POWER SUPPLY. WOULD YOU LIKE TO START WITH A FAULTED INSTRUMENT WHICH NEEDS FIXING?

>>YES

SOPHIEは、装置のモデルに故障を任意に選んで挿入すると、制御値がどのように設定されているかを学生に教える。

THE INSTRUMENT HAS A FAULT AND ITS PRESENT CONTROL SETTINGS ARE:

CC 1.0 CURRENT CONTROL (制御電流)
CR HIGH CURRENT RANGE SWITCH (電流切換スイッチ)
LOAD 20 LOAD RESISTANCE IN OHMS (負荷抵抗(オーム))
VC 1.0 VOLTAGE CONTROL (制御電圧)
VR HIGH VOLTAGE RANGE SWITCH (電圧切換スイッチ)

>>WHAT IS THE OUTPUT?

THE OUTPUT VOLTAGE IS 11.7 VOLTS.

>>WHAT IS IT IN A WORKING INSTRUMENT?

IN A WORKING CIRCUIT THE OUTPUT VOLTAGE IS 19.9 VOLTS.

>>WHAT IS THE VOLTAGE BETWEEN R8 AND Q1?

THE VOLTAGE AT N1 IS 47.6 VOLTS.

>>WHAT IS THE BSE EMITTER VOLTAGE OF THE VOLTAGE LIMITING TRANSISTOR?

* 測一博監訳：“人工知能の基礎”，近代科学社，1978。

回路診断システムであるが、使用者に提示する回路図には、誤動作を行なう部分が挿入されており、それを質問応答により発見させるもの。これにより使用者は、回路動作を習得することができる。主な特徴は、

- (a) 問題領域に即した意味文法 (8.2.5 参照) の導入。
- (b) 照応指示の決定と省略語の推定が可能。

(D) [Waltz, 1976] (Illinois 大)

PLANES

海軍の空軍機データ・ベースに対する質問応答

User: How many Phantoms required unscheduled maintenance in April 1975?
PLANES: Thirty Phantoms. Do you want to see them?
User: No. Did any of these require radar maintenance?
PLANES: Yes, three of them. Do you want to see them?
User: Yes.
PLANES: Phantom 159829 required a new radar screen. Phantom 159772 required a new power supply. Phantom 158336 crashed as a result of radar failure.
User: Did any other planes crash?
PLANES: I assume you mean during April 1975. Two other planes crashed: a Phantom and a P3.
User: Is it true that in both cases and April flights occurred while on the fuselage?
PLANES: I cannot understand your question. It is non-grammatical and meaningless.
User: Is it true that both Phantoms were from the same squadron?
PLANES: I assume you mean "from" instead of "fron." Yes both planes were from the same squadron.
User: Is there anything else in common between the two planes?
PLANES: Yes, both were serviced on April 24th by maintenance group HEBRON FMFPAC.

以上は、最終的な PLANES による質問応答のシナリオであり、実例とはいえないかも知れない。その後、PLANES の問題点を改良した JETS 計画 [Finin, 1979] が提案されている。特徴は、

- (a) 照応指示の決定と省略語の推定が可能。
- (b) SHRDLU同様、推論は μ PLANNERで記述。
- (c) 質問文に含まれる誤りや不適切さの指摘。

(E) [鷹尾, 1976](IBM)

やちまた

データ・ベースに対する質問応答

サッホ°ロヨ コ°コ° るシ°イセ°ソニ テ°ル トウキョウ ユキノ ヒ°ンハ?			
*フライト**	*****ヒニチ*****	**クウコウ**	**クウコウ**
JAL502	マイニチ	サッホ°ロ	トウキョウ
JAL504	マイニチ	サッホ°ロ	トウキョウ
JAL506	マイニチ	サッホ°ロ	トウキョウ
JAL508	2/11-2/14	サッホ°ロ	トウキョウ
JAL510	マイニチ	サッホ°ロ	トウキョウ
JAL514	ニチ, ク°ツ, 月, 24, 62 ヨウ	サッホ°ロ	トウキョウ

コレラノ ウチテ° キシュカ° シ°ヤツホ° SR ノ モノハ?			
*フライト**	*****ヒニチ*****	**クウコウ**	**クウコウ**
JAL502	マイニチ	サッホ°ロ	トウキョウ
JAL504	マイニチ	サッホ°ロ	トウキョウ
JAL510	マイニチ	サッホ°ロ	トウキョウ

特徴は

- (a) 日本語による実用レベルの質問応答システム。
- (b) 質問文は体言止めの文に限られる。
- (c) 質問文に用いる新しい言葉を必要に応じて定義可能。

(F) [Bobrow, 1977c] (Xerox)

G U S

フレームを用いた旅行計画作成システム

GUS: Hello. My name is Gus. I can help you plan a simple trip by air.
Where do you want to go?

(1) Client: I want to go to San Diego on May 28

GUS: What time do you want to leave?

(2) Client: I must be in San Diego before 10 am

GUS: Would you like P.S.A. flight 102 that arrives at 9.15 am?

(3) Client: What is the next flight?

GUS: Air California flight 310 that leaves at 8.30 am and arrives at 10.00 am

(4) Client: I'll take the first one

GUS: In what name should I make a reservation, please?

(5) Client: Dan Bobrow

GUS: I have confirmed the following flight:
P.S.A. flight 102 on Wednesday May 28 from San Jose to San Diego
that leaves at 7.30 am and arrives at 9.15 am.
What date do you want to return on?

(6) Client: On Friday in the evening.

GUS: Would you like the flight that leaves at 7.45 pm?

(7) Client: That's fine.

GUS: I have confirmed the following flight:
P.S.A. flight 307 on Friday May 30 from San Diego to San Jose
that leaves at 7.45 pm and arrives at 9.30 pm
Thank you for calling. Goodbye.

特徴は

- (a) フレームによる対話のガイド
- (b) 主導権混在型 (mixed initiative)
- (c) 知識表現形式の検討

(G) [Hendrix, 1977] (SRI)

L I F E R

データ・ベースに対する質問応答

(H1) What is the salary of Eric Johnson?

(M1) (ID 327-36-8172 SALARY 19250)

(H2) Position and date hired?

(Eric Johnson のことであるが、質問文にはそれが欠けており、補なわなければならない)

(M2) (ID 327-36-8172 POSITION Assoc-Prof DATE-HIRED 8/1/70)

.....
(Hi) What is the average for math department secretaries?

(Mi) AVAREGE<==spelling (綴りの誤り指摘)

SALARY 5 items accepted

AVERAGE IS 7631. 4999

(Hi+) Maximum (数学科秘書の俸給の最高額のことであると判断しなければならない)

(Mi+) SALARY 10730

特徴は

(a) 問題領域に即した意味文法の導入。

(b) 照応指示の決定と省略語の推定が可能。

(H) [Harris, 1978] (ダートマス大)

R O B O T

データ・ベースに対する質問応答

(H1) Give me the names of all employees who have
jobs working as a secretary in the city of Chicago.

(M1) The number of records to retrieve is 30.

REINSTEDT

LAWLER

MOOERS

SORIANO

LORIE

(H1) Print the salary of Smith and Lawler.

(M1) The number of records to retrieve is 269.

\$ 7,500	LAWLER
\$ 9,600	SMITH
\$ 10,800	LAWLER
\$ 5,100	LAWLER
\$ 12,000	SMITH

特徴は

- (a) カーペット会社で実際に利用されている。
- (b) 曖昧さの解除を使用者に求める方式。

(I) [Bonnet, 1978] (スタンフォード大)

BAOBAB

医療診断投薬システムに対するフロント・エンド

Would you like to enter a rule?

** YES

If: 1 - THE CULTURE SITE IS CSF
and 2 - THE PATIENT HAS NOT HAD RECENT NEUROLOGIC SYMPTOMS
and 3 - HE HAS NOT HAD RECENT NEUROLOGIC SIGNS
and 4 - THE WHITE BLOOD COUNT FROM THE SPINAL FLUID IS 10
and 5 - THE CSF PROTEIN IS LESS THAN 40
and 6 -

Then: 1 - IT IS LIKELY (.8) THAT THE INFECTION REQUIRING THERAPY
- MENINGITIS
and 2 -

the internal format of the rule is :

LISPRULE

PREMISE: (\$AND (SAME CNTXT SITE CSF)
(NOTSAME CNTXT NEUROSYP)
(NOTSAME CNTXT NEUROSIGN)
(LESSP* (VAL 1 CNTXT CSFCELLCOUNT)
-10)

(LESSP* (VAL 1 CNTXT PROTEIN)
40))

ACTION: (CONCLUDE CNTXT TREATINF MENINGITIS TALLY-800)

This is my understanding of your rule :

-
- If: 1) The site of the culture is csf, and
2) The patient has not had recent neurologic symptoms, and
3) The patient has not had recent neurologic signs, and
4) The WBC from the CSF is less than 10, and
5) The CSF protein is less than 40

Then: There is strongly suggestive evidence (.8) that the
infection which requires therapy is not meningitis
do you agree whth my interpretation?

** YES

good.

スタンフォード大で、Shortliffeは、MYCINとよばれる医療診断投薬システムを開発したが、MYCINに対して、医学的な新知識を自然言語（に近い形式）で付加することが可能。特徴は

- (a) 意味文法の導入。
- (b) より広い範囲の問題をカバーするための一般文法の導入。
- (c) 理解の結果を使用者に提示し、確認をとりつつ作業を進める。

(J) [Codd, 1978] (IBM)

RENDEVOUS

データ・ベースに対する質問応答

S 001 PLEASE TYPE IN YOUR QUERY:

U 001 HOW MANY WHITNEY SHIPMENTS HAVE A SHIPDATE 6/10/1975

S 002 IS 6/10/1975 INTENDED TO BE A DATE? (YES OR NO)

U 002 YES

S 003 WE ARE TRYING TO INTERPRET 6/10/1975 AS A DATE.
WHICH OF THE FOLLOWING DATES DO YOU INTEND:

- 1 JUNE 10
- 2 OCTOBER 6
- 3 NONE OF THE ABOVE

SELECT ONE ITEM OR NONE BY NUMBER

U 003 1

S 004 IS THE WORD 'WHITNEY' CORRECTLY TYPED? (YES OR NO)

U 004 YES

S 005 IS 'WHITNEY' THE NAME OF A PART? (YES OR NO)

U 005 NO

S 006 THE WORD 'WHITNEY' IS UNFAMILIAR. IS IT ONE OF THE
FOLLOWING:

- | | |
|---------------------|---------------------|
| 1 SUPPLIER NAME | 5 PROJECT LOCATION |
| 2 SUPPLIER LOCATION | 6 PART NAME |
| 3 SUPPLIER RATING | 7 PART TYPE |
| 4 PROJECT NAME | 8 NONE OF THE ABOVE |

SELECT BY NUMBER THE CLOSEST CATALOG ITEM

U 006 1

NOTE THE SYSTEM HAS 2 WAYS TO INTERPRET YOUR QUERY. PLEASE
INDICATE IF ANY OF THEM IS WHAT YOU INTENDED.

1 COUNT THE NUMBER OF SUPPLIERS NAMED WHITNEY WHO, ON
JUNE 10, 1975, SENT A SHIPMENT.

2 COUNT THE NUMBER OF SHIPMENTS SENT ON JUNE 10, 1975 BY
A SUPPLIER NAMED WHITNEY.

S 007 WHICH OF THE INTERPRETATIONS IS CORRECT:

- 1 FIRST
- 2 SECOND
- 3 NONE OF THE ABOVE

SELECT EXACTLY ONE ITEM BY NUMBER

U 007 2

* *** THERE IS NO SHIPMENT IN THE DATABASE SATISFYING YOUR QUERY.

特徴は

- (a) 曖昧さの解消を、メニューの選択により使用者に委ねる方式。
- (b) 質問文は、3節で説明した DEDUCE に変換される。
- (c) 文法的に少々誤った文の入力をも許す。

8.2.5 意味文法

前項でしばしば意味文法という用語が使われていた。意味文法は、BBNのBrownとBurton〔Brown, 1975〕がSOPHIEで用いて成功したことが発端になっている。対象領域を限定することにより、8.2.1節で説明したような制限(i), (ii), (iii)を課することができる。その制限により得られる利点をフルに活用するために、意味文法が考案された。

意味文法では、名詞、名詞句、動詞句、副詞句など、通常の文法で用いられる文法カテゴリーはさらに細分化されて、意味的な情報を反映した意味カテゴリーで置き換えられる。したがって、意味文法に従って構文解析した結果は、たとえば図8-10のようになる。

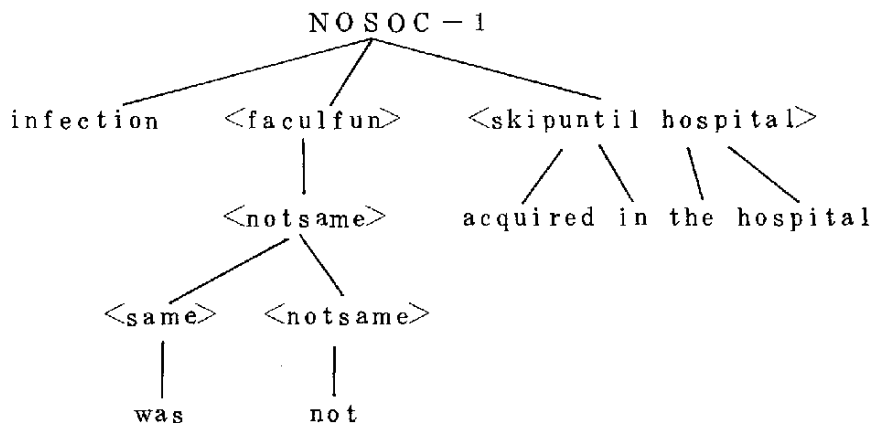


図8-10 意味文法による構文解析木
〔Bonnet, 1978〕

文法カテゴリー名に意味的な情報を組み込んでいるため、省略や代名詞の指すものを決める場合に、文法規則のもつ情報から、それらを予測することが可

能となる。

このような意味文法には、つぎの欠点がある。

- (a) 対象領域を絞らない限り、文法カテゴリー数が際限なく増加する。
- (b) 文法規則が対象領域に強く依存しているため一般性がなくなる。したがって、対象領域が変わる毎に、ほとんどすべての文法規則を書き直さなければならない。

8.2.6 今後の課題

8.2.6.1 概要

8.2.4で説明した質問応答システムに対して実際に作成し使用した経験からさまざまな問題点を抽出し、新しい世代の(自然言語による)質問応答システムを作成しようとする気運が高まっている。昨年我が国で行なわれた第6回人工知能国際会議には、反省も含めてその様な報告がいくつか見られた〔Finin, 1979〕,〔Harris, 1979〕,〔Hayes, 1979〕。それらも含めて、今後の課題をまとめてみよう〔BBN, 1978〕,〔Bobrow, 1977c〕,〔Bonnet, 1978〕,〔Codd, 1978〕,〔Hendrix, 1977〕,〔Lehnert, 1977〕,〔Petrick, 1976〕,〔Plath, 1976〕,〔Waltz, 1977〕,〔Woods, 1977〕,〔鷹尾, 1976〕,〔田中, 1979〕。以下に列举する課題は、相互に関連し合っている。したがって重複があるがとにかく気が付いた問題を羅列してみる。

- (a) 柔軟なパーザ
 - (b) 会話のモデル
 - (c) 応答理由の説明能力
 - (d) パラフレーズの処理
 - (e) 意味, 知識, データ表現法
 - (f) 推論機構
 - (g) システムの拡張性
 - (h) 誤入力の処理
 - (i) 多義語の表現と処理
 - (j) 限量詞, 否定詞の範囲の決定
- 日本語特有の問題として

(k) 日本語入力法

(a) のパーズの手法については、1970年代に、一定の前進があったというのは事実であろう(第2章)。しかし、依然として自然言語をパーズする問題に多くの課題がある。特に、構文レベルの情報と意味レベルの情報をどのようにパーズ過程で融合させるかが問題になっている。意味文法はきわめて限定した対象領域で、この問題を解決しようとしたものであるが、十分でないことは5節で説明した。しかし、このような問題は構文レベルにあるのではなく、意味、さらにいえば語用論のレベルにあり、そこに未解決の問題が山積していることによると思われる。筆者は現段階では、これら3つのレベルの情報をゆるく結合したパーズの方法が望ましいと考えている。(a)には(e), (h), (i), (j)の問題が関係している。

(b)については六章で説明した。(c)は、質問に対する答をシステムから受け取った時に、どのようにしてその答を計算したかをシステムが説明する機能である。これは、医療診断投薬システムMYCINでの使用経験から、説明機能の重要性が認識されたものである〔Shortliffe, 1976〕。

(d)は、ある事を自然言語で述べる時、さまざまな言い廻しがあることを問題にしている。たとえば、〔田中, 1979〕には、半径2cmの穴の空いた衝立を1cdの電球の前10cmの所に置く、という文には、少なくとも2万通り以上の言い廻しがあることが指摘されている。一文を複数の文の系列に分解することも可能であるから、システムがそれらのepistemicな意味の同一性を理解できるかどうかは、大きな問題になる。

(e), (f)の重要性については改めて説明を要しないだろう。今回の報告書には含めなかったが、(e)については第3章として、いずれ説明されるだろう。

(f)の一部は、本報告書の第5章で説明されている。

SCHOLARやBAOBAB, LIFERなどでは、Interlispの機能の一部を用いて、単語の綴りの修正を行なっている(h)。しかし、誤入力にはさまざまなレベルがあり、それらの検出と修正は、実用的なシステムでは大きな問題になる。これは人工言語についてもいえることである。入力プログラムが常に正しいと仮定すれば、それを処理するプログラム(コンパイラ、インタプリタなど)は容易に作成できる。しかし現実には、入力プログラムにはさまざま

な（予想もつかないような）誤りが含まれている。それらにどのように対処していくかにより、人工言語を処理するプログラムが巨大化、複雑化するのである。

(i)、(j)については、(a)に関してすでに述べた。(k)については8.2.6.3で説明する。

8.2.6.2 会話のモデル

使用者と質問応答システムとの会話は、2人の対話のモデルになぞらえることができる。通常の対話では、話し手と聞き手が相互にその役割りを入れ替え、また対話の主導権を時として交互に変えながら進行する。これらは、主導権混在型(mixed initiative)の会話とよばれ、知的な質問応答システムのもつべき機能の1つである。

このような機能を実現するためには、使用者の質問や答の意図を理解しなければならない。たとえば、現在の会話の主題は何か、焦点はどこにあるかを、会話の流れもしくは文脈から汲みとらなければならない。したがって、談話理解の問題が解決されねばならないが、談話の問題は、従来、語用論に属する問題として十分に議論がなされていなかった。

談話の理解には、指示代名詞や連体詞の指示するものの決定、言語行為論、焦点と主題、前提、推論、省略、理解のためのメモリ・モデル等の問題が含まれている。自然言語による質問応答のもつ利点の1つとして、8.2.3項の(g)で、指示代名詞や連体詞等を用いて、既出のもの(文、名詞)を指しながら質問できることをあげておいた。これはまさに、談話理解の問題に関係している。したがって、水準の高い質問応答システムを実現する場合には、談話をめぐる種々の問題を解決しなければならない。本報告書第1部第5章の調査と〔長尾、1979〕中の“談話理解の構造”をも参照されたい。

8.2.6.3 日本語入力法

漢字情報処理は、漢字の入出力に関係するもので、質問応答システムのフロント・エンドとなるものである。漢字情報処理は、漢字フォント・メモリの低廉化が1つの要因であろうが、レーザ・プリンタやインク・ジェット等に代表されるハード・コピー装置の高性能化、高速化、低廉化も、漢字情報処理を促す大きな要因であったと思われる。

このような出力装置の進歩に対して、漢字の入力法には依然として問題がある。漢字の数は英数字に比べて比較にならない程多数であるからである。これは、質問応答システムと日本語で対話しようとする時間問題になる。

質問応答システムと日本語で対話する時の漢字入力法として

- (i) 入力に必要な全文字を配置した鍵盤を用いる方法
- (ii) 漢字毎の連想コードを用いたタッチ法
- (iii) 同音漢字表示選択法
- (iv) カナ入力(のみ)
- (v) ローマ字入力(のみ)

(i), (iii) は素人でも使える利点はあるが、熟練しても、漢字入力速度がさほど向上しないという欠点がある。50字/分が限界であるとされている。(ii) は、英文タイプと同様に、鍵盤を目視することなく入力が可能であるが、各漢字に対応する連想コードを憶えなければならない。熟練後は、最高120字/分の入力が可能になるといわれている。

(iv), (v) は、同音異義語の判別を行なわねばならず、カナ漢字変換と同様な問題が生じるが、きわめて限定された対象領域では有用だろう。

筆者は(iv)と(v)を比較するなら、(v)が優れていると考えている。その理由は、

- (a) カナ・コード付きの特殊な鍵盤が不要
- (b) ローマ字鍵盤配列の方が、カナ鍵盤配列より人間工学的な配慮がなされている。
- (c) ローマ字表記の方が用言の形態素処理に適している。(たとえば「スル」と「シタ」には、共通のカナ文字が含まれない。これに対して、ローマ字表記のSURUとSITAには、共通のローマ字Sが含まれている)。

漢字入力法に関しては、昭和54年度より情報処理学会に委員会が設置され検討が進められている。多くの人の納得の行く結論が得られることを希望したい。

8.2.7 おわりに

自然言語による質問応答システムをめぐるさまざまな問題を（筆者の独断をも混えて）考察してきた。1980年代には、限定された対象領域で、自然言語により人と知的な会話を交す実用的な質問応答システムが実現するであろう。これは、計算機のもつ機能を高次化しようとする知識情報処理システムの実現と密接に関連する。米国と比較して、我が国での自然言語による質問応答システムの研究は立ち遅れているといわざるを得ない。しかし最近、多くの若い研究者によるこの分野への興味と関心が高まっていると聞く。自然言語処理の研究は本質的に長期のテーマであることを認識しつつ、今後の研究の進展に期待して結びとする。

第Ⅱ部 最近の言語学の潮流

Ⅱ 最近の言語学の潮流

目 次

1. 生成文法理論の展開	224
2. 生成文法理論における意味論の展開	231
3. ヨーロッパにおける言語学の展開	235
3.1 ヨーロッパにおける1960年代までの言語学の歴史	235
3.1.1 はじめに	235
3.1.2 フンボルトからドイツ青年文法学派まで	236
3.1.3 ド・ソシュールと構造主義	237
3.1.4 ブラーグ学派	238
3.1.5 その他の言語学派	239
3.2 テニエールと結合価理論	241
3.2.1 テニエールの結合価理論	241
3.2.2 結合価理論の完成	242
3.2.3 結合価辞典	242
3.3 テキスト言語学	243
3.3.1 テキスト言語学の歴史	243
3.3.2 テキスト言語学の論理的発展	244
3.3.3 テキスト言語学の構成	245
3.4 「おわりに」に代えて	246
参考文献	247
4. Montague 文法	248
4.1 Logic	250
4.1.1 Modal Logic	250
4.1.2 Kripke Model	251
4.1.3 Intentional Logic	255
4.2 Montague Grammar	262
4.2.1 Proper Treatment of Quantification	264
参考文献	274
今後の予定	275

1. 生成文法理論の展開

1. 1957年に、N. ChomskyがSyntactic Structuresにおいて生成文法理論の概要を発表して以来、生成文法理論は既存の理論をより精密にし、より優れた理論（可能な文法の範囲をより狭く定義する理論）に到達することをめざし、いくつかの重要な理論上の進展を経てきた。特に、初期の理論から標準理論への進展、次に標準理論から拡大標準理論への進展、更にこの拡大標準理論から痕跡理論(trace theory)の導入に伴う修正拡大標準理論への進展は、生成文法理論の展開の歴史の中で、画期的な飛躍といえる。¹⁾本稿では、まずこの理論進展の三つの段階の概要をたどり、次に最近の理論展開の成果である痕跡理論を中心とした修正拡大標準理論について紹介を行なう。

2. Chomsky(1957)は統語構造を明示的に記述するには句構造規則(Phrase Structure Rules)だけでは不十分であり、変形規則(Transformational Rules)を文法理論の中に導入する必要があることを明らかにした。²⁾これに対して、Chomskyは1965年のAspects of the Theory of Syntaxにおいて、この初期理論における変形規則の定義を一部分修正し、その機能を明確にした。また、基底部門を範疇部門(辞書)に分け、語彙項目の統語的性質に関する情報を統語素性の形で記述する必要があることを示し、更に変形規則の循環適用の原則などいくつかの新しい提案を行ない、標準理論を確立した。この第一の段階での主要な理論進展は、統語部門の中の二つの下位部門（即ち、基底部門と変形部門）に関して各々の職能とその相互関係を明確に規定した点である。

第二の段階の重要な理論進展は、Aspects以後のChomskyの三つの重要な論文（(Deep Structure; Surface Structure, and Semantic Interpretation(1970a); Remarks on Nominalization(1970b); Empirical

1) 詳しくは梶田(1976)『変形文法論の軌跡』（大修館）並びに太田・梶田(1974)『文法論Ⅱ』（英語学大系4，大修館）参照。（拡大標準理論に至る生成文法理論の細部に至るまでの動向が適格に捉えられよう。）『言語』第6巻，2号，3号の「特集チョムスキーの全体像」及び「特集チョムスキー理論の展開」も参照。

2) 変形規則の適用は構造依存の原理によることが明確に示された。

Issues in the Theory Transformational Grammar(1972a))を納めたStudies on Semantics in Generative Grammarは、深層構造のみが文の意味解釈に関与するという標準理論のひとつの基本的主張を修正し、文の意味解釈は、深層構造と表層構造の組によって決定されるという拡大標準理論を展開した。この拡大標準理論の枠組みにおいては、文の意味表示の決定に必要な文法関係は標準理論と同様に深層構造によって決定されるが、論理語の作用域、文の焦点や前提、照応関係などは、(音声的に解釈された)表層構造に適用される意味解釈規則によって決定されるとした。このように、この段階では、統語部門の内部構造の問題から、統語部門と意味部門の係わり合いの問題に研究の焦点が移り、表層構造が意味解釈に重要な役割を果たすことが明らかにされた。³⁾

現在、生成文法理論は理論進展の第三の段階のただ中にある。Chomskyの'Conditions on Transformations'(1973)にその萌芽がみられ、Reflections on Language(1975b)や'Conditions on Rules of Grammar'(1976)で具体的に論じられている痕跡理論に基づき、更に'On Wh-Movement'(1977b)並びにChomskyとLasnikによる"Filters and Control"(1977)やChomskyの未出版論文'On Binding'(1978)や'Markedness and Grammar'において展開されているいくつかの仮説に基づいて修正拡大標準理論が確立しつつある。⁴⁾この修正拡大標準理論の基礎を成す痕跡理論の基本的主張は、ある要素を移動する変形規則はその要素が移動された元

3) Chomsky(1970b)では、基底部門に焦点をあて句構造規則に課される制約に関する仮説として $\bar{X}$ 理論を提示している。この $\bar{X}$ 理論はJackendoff(1971, 1977a; 1977b), Berman(1974), Selkirk(1977)等で実証的研究を踏まえて展開されている。八木(1977)も参照。

4) 'Conditions on Transformations'(1973), 'Questions of Form and Interpretation'(1975a), 'Conditions on Rules of Grammar'(1976)等はChomsky(1977a)の論文集Essays on Form and Interpretationに再録されている。またRules and Representation,とLecture on Core Grammarがさらに準備中である。

の位置に痕跡 (trace) を残すというものである。修正拡大標準理論においては、この痕跡の導入により派生の最初または中間の段階に現われる句構造標識のある種の特徴が表層構造まで持ち越されることになる。これは、今まで (拡大) 標準理論の中で記述しようとする問題になると考えられていた現象を全体的派生制約という強力なしくみを文法に導入することなく説明することを可能とした。⁵⁾ このように、修正拡大標準理論は、拡大標準理論の枠組みの中へ、痕跡という理論的構成物を新しく導入し、それに伴って文法規則の定義及びそれが関連づける構造の定義を一部分一律に変更することにより、例えば、可能な変形規則の類を従来の理論よりもはるかに狭く定義し、更に変形規則の形式に関する条件を非常に簡潔にしたり、また、文の意味解釈は表層構造によってのみ決定されるとしたりして、従来の理論の記述能力を増大させることもまた減少させることもなく、可能な文法の範囲をできるだけ狭く定義しようとするものである。⁶⁾

以上、生成文法理論の展開の軌跡を Chomsky の主要な著作並びに論文を中心に概観した。ここで重要なことは、上述した理論進展は N. Chomsky ひとりによって推進されたわけではなく Chomsky の研究を中核とした統語論・音韻論・意味論の各分野における他の多くの研究の成果に拠っていることである。⁷⁾ 例えば、標準理論の確立には、Fillmore (1963), Katz and Fodor (1963), Katz and Postal (1964) 等による初期理論における問題点の指摘や意味部門と変形規則に関する考察がその原動力となっている。また、拡大標準理論は、Kuroda (1965, 1969), Jackendoff (1969, 1972), Akmajian (1970, 1973) 等の表層構造が文の意味解釈に関与することを示す言語事象を指摘した一

5) 例えば代名詞の交差現象 (Cross-Over phenomena) や縮約 (Contraction) に関する事象

6) 変形規則は Wh-移動規則と名詞句移動規則の二種類のみが核文法 (core grammar) に属するとされている。(核文法とは無標の規則の体系としての文法を指すものである。Chomsky and Lasnik (1977), Stillings (1977), Koster (1978a, 1978b) 等参照)

7) 各理論にはそれに基づく優れた英語に関する実証的研究が存在する。初期の理論内では、Lees (1960), Klima (1964), 標準理論内では Rosenbaum (1967), 拡大標準理論では Milsark (1974) 等が挙げられよう。

連の研究に基づいて誕生した。同様に痕跡理論及びそれに基づく修正拡大標準理論も、Wasow(1972, 1979), Selkirk(1972), Fiengo(1974, 1977), Sag(1976)等の諸研究の成果に多くを負っている。更に、標準理論に基づくRoss(1976)の変形規則の適用に課される諸制約の研究並びにEmonds(1970, 1976)の変形規則に関する構造保持の仮説(Structure-Hypothesis)は、この修正拡大標準理論に多大な影響を与えている。⁸⁾

3. Chomsky(1975b:105)は、痕跡理論に基づく広義の文法の組み立てを、概略、(1)のように提案している。(1)においてB:基底部門の規則、T:変形規則)

(1) 文法 (Sentence grammar) : $\xrightarrow{B}$ 基底構造 $\xrightarrow{T}$ 表層構造 $\xrightarrow[\text{規則(SI-1)}]{\text{意味解釈}}$ 論理形式

意味解釈規則 (SI-2) : 論理形式 $\longrightarrow$ 意味
他の認知体系

(1)に示した文法の体系は、(拡大)標準理論と次の点で著しく異なる。(1)では、意味解釈規則が大きく二種類(即ち、SI-1とSI-2)に分けられ、表層構造と最終的意味表示との間に、「論理形式」という新しい中間の意味表示のレベルが設定されている。Chomskyが従来の文法の体系とこのように著しく異なる(1)を提案した背景には、「形式文法の自律性に関する命題」という彼の考えが反映している。この命題は、文の形式に関する知識の体系としての形式文法(統語論、形態論、音韻論を含む)は、文の形式と意味との結び付きに関する知識の体系である意味論から独立した自律的体系としてできるだけ狭く定義されなければならないと主張する。痕跡理論の導入は(1)のような文法の体系を可能とするものであり、従って、この自律命題の達成に重要な役割を果たすと考えられる。⁹⁾

8) 本稿では音韻論について言及することができないので原口(1977)「生成音韻論の動向」(『英語学16』)参照

9) 痕跡理論について詳しくは、拙稿寺津(1976)「痕跡理論について(上)・(下)」(『英語文学世界』第11巻6号, 7号), 長原・稲田・池内(1976)「痕跡理論(1)-(3)」(『英語青年』第121巻10号-12号), 大島(1977)「痕跡理論をめぐって」(『言語』第6巻3号)池内(1979a)「文法理論の最近の動きに関する覚え書き」(Mulberry 28)等参照。

Chomsky and Lasnik(1977)は、(1)の体系における文文法に関して、更に(2)に示すようにその内部構造を明確にしている。¹⁰⁾

- (2)
1. 基底部規則
 2. 変形規則

〔表層構造〕

- | | |
|------------|-------------------|
| 3 a. 削除規則 | 3 b. Construal 規則 |
| 4 a. フィルター | 4 b. 数量詞解釈規則 |
| 5 a. 音韻規則 | ⋮ |
| 6 a. 文体規則 | ⋮ |
| | 〔論理形式〕 |

(Construal 規則には、相互代名詞規則 (Reciprocal Rule), 束縛照応規則 (Bound Anaphora), 分離指示規則 (Disjoint Reference), コントロール規則 (Control Rule) 等が含まれる。)

前述したように、(1)及び(2)のしくみに基づく修正拡大標準理論は、現在もその研究が進展中であり、痕跡の実体に関する定義や痕跡に係わる制約に関しても明確・精密化される方向にあるが、痕跡の基本的定義は(3)に示すようなものである。¹¹⁾

(3) ある変形規則が句 P を X の位置から Y の位置へ移動する時、その変形規則は X の位置に句 P によって束縛された痕跡 (a trace bound by P) を残す。(Chomsky(1975b : 95))

(3)から明らかなように、痕跡理論においては、移動された要素と痕跡との間に成り立つ関係は、照応関係 (Anaphora) における先行詞 (Antecedent) と束縛照応形 (Bound Anaphor) の間に成り立つそれと基本的に同じだとみなされている。このような痕跡を文法に導入する動機付けとして、(4)に示すよう

10) (2)において表層構造は痕跡や P RO を含み、削除規則及び文体規則の適用前の構造であることから従来のそれよりも抽象的なものである。(Cf. 浅い構造 (Shallow Structure))。

11) Chomsky(1978)では指標付与 (index assignment) の一般的な規約が明確にされ、P RO, 痕跡, 照応形, 名詞句が一律に扱われることを可能としている。

な各種の議論が提示されている。¹²⁾

- (4) a. 可能な変形規則の類についての原理的な説明に関する議論
b. 指定主語の条件 (Specified Subject Condition) 及び時制文の条件 (Tensed-S Condition) と変形規則及び意味解釈規則の相互作用に関する議論
c. 表層構造意味解釈の仮説に関する議論
d. 交差の現象と論理形式に関する議論
e. Wanna 形式に関する議論

修正拡大標準理論においては、痕跡理論を前提として、変形規則は単に「 α を移動せよ (Move α)」として定式化されている。この定式化により従来よりも過剰に生成される統語構造は、痕跡や PRO を含む分だけ従来よりも豊かな情報を持つ表層構造に適用される意味解釈規則並びにその意味解釈規則の適用(あるいは意味解釈規則の適用の結果得られる出力としての論理形式)に課される削除規則の出力に適用されるフィルターによって排除される。¹³⁾従って、理論全体としては、従来のどの理論よりも可能な文法の範囲を厳しく制限するものであり、生成文法理論の目標達成に近づくものであると考えられている。しかし、拙稿 寺津 (1977)「修正拡大標準理論における意味部門の内部構造について」及び寺津 (1979)「修正拡大標準理論における削除問題について」で指摘したように、¹⁴⁾修正拡大標準理論はその細部においてまだ精密化されるべき点を数多

12) (4d)に関してはWasow(1972, 1979)及びChomsky (1975a, 1976), (4e)に関してはChomsky and Lasnik(1977, 1978), Andrews(1978), Postal and Pullum(1978, 1979), Terazu(1977)等参照。Chomsky(1978)では指標付と並びに格に関する規約に基づき指定主語条件及び時制文条件は不透明の条件 (Opacity Condition) 及び主格島の条件 (Nominative Island Condition) として再定化が試みられている。

13) 修正拡大標準理論に基づく研究としてFreidin(1978), Dresher and Hornstein(1979), Koster(1978a, 1978b) Lightfoot(1976, 1977)等がある。またこれに対する批判としてはBresnan(1975, 1976, 1977), Ikeuchi(1979b, 1979c), Inoue(1979a, 1979b), Iwakura(1979), 寺津(1978)等参照。

14) 寺津(1977)は『富山大学人文学部概要』創刊号, 寺津(1979)は同3号。

く残しており、また根本的な修正を必要とする点もある。よって、この修正拡張標準理論の妥当性は、今後の研究によって更に検討される必要がある。

4. 以上、生成文法論における Chomsky を中心とした研究の動向に焦点をあてその理論進展を概観した。

ところで、初期の理論から標準理論確立に至る時期には理論内部における意見の対立はみられなかったが、標準理論以降、理論内部における論争が活発になり、Chomskyとは異なる方向に研究を進めるグループがいくつかでてきている。最後に、これらに関して簡略に言及しておく。¹⁵⁾ Aspects を契機に Fillmore によって格文法 (Case Grammar) が、また Gruber (1965, 1967) に端を発し、Lakoff, McCawley, Postal 等によって生成意味論 (Generative Semantics) が提唱され、拡張標準理論における解釈意味論 (Interpretive Semantics) に対立した。また、1970 年代に入り、Postal, Perlmutter, Johnson 等による関係文法 (Relational Grammar) や久野による機能文法 (Functionalism), 等の研究も活発に行われている。

15) 詳しくは原口 (1979) 「変形文法家の最近の理論展開」 (『言語』第8巻10号)

参照。

2. 生成文法理論における意味論の展開

生成文法理論は、その出現の背景として、(a)言語能力と言語運用の峻別、(b)規則に支配された言語の創造的側面の再認識、(c)言語の多様性と普遍性の認定に基づく深層レベル・表層レベルの設定、(d)両レベルを動的に統合する変形操作の概念の導入等の言語学的視点を合理論の哲学的基盤に基づいて明確な形で提示することにより、従来の伝統文法、記述文法、構造言語学の枠組に基づく意味論に対して新たな意味研究の座標を提示した点に、その意義があると言える。しかし、現時点での生成理論は、解釈意味論、生成意味論、格文法理論等の諸派に分岐し、厳密にはその対象とする意味領域の限定の仕方、最終目標への到達手段に関する理論上の対立が顕著になってきている。本稿では、これらの主要流派の成立過程とその相互関係を比較検討することにより、生成理論における意味論の理論的進展の一樣相を考察する。

生成文法理論は、公式には Chomsky (1955, 1957) に始まる訳であるが、この時点での表層構造と深層構造の区分及び両レベルに介在する変形規則の導入は、あくまで文法理論の中核をなす統語部門の定式化に際して提起されたものである。これに対し、Katz and Fodor (1963), Katz and Postal (1964) は、上記の深層構造に意味決定の中心的役割を認め、意味解釈規則の導入により、統語部門の統合を図り、ここに生成文法理論における意味論の基盤が確立されることになる。その後、Chomsky (1965) により、以上の枠組に基づく標準理論 (Standard Theory, ST) の定式化がなされる。ST の基本的特徴は、深層構造を意味表示の決定に必要な情報を過不足なく含み、すべての意味解釈規則の入力となるレベルとみなす点にある。ST はその後 Chomsky (1970), Jackendoff (1972) 等を起点とする拡大標準理論 (Extended Standard Theory, EST) として修正される。ST と EST の基本的相違は、深層構造と意味解釈規則の相互関係にある。上記の ST の概略から明らかなように、ST の深層構造は、すべての意味解釈規則の入力となるレベルであるのに対し、EST では、基底レベルの基本的文法関係に基づく意味は、ST と同様深層構造に意味解釈規則が適用して得られるが、変形規則の適用の結果生じる (a) 焦点と前提、(b) 数量詞・否定辞、(c) 照応関係などの意味規定は、表層構造

(ないしは、同構造に至る派生の間段階)に意味解釈規則が適用して得られることになる。ESTは更にChomsky(1973), Fiengo(1974)等を起点とし、痕跡理論(Trace Theory)を組み込んだ、修正拡大標準理論(Revised Extended Standard Theory, REST)として改訂される。RESTの基本的特徴は、ESTにおける上記(a)―(c)に係わる意味現象だけでなく、従来深層構造によって決定されるとみなされていた基本的文法関係に基づく意味情報も、派生過程を経て表層レベルに保持され、この後者のレベルに意味解釈規則が適用して、その最終的意味表示が得られるという点にある。以上のST, EST, RESTの経緯をみた場合、意味部門の解釈規則の入力となる言語レベルの位置づけに関し、理論的変容がみられるが、いずれの枠組においても、意味部門が深層構造、表層構造ないしは両者の中間段階のレベルをその意味規定に際しての入力とし、最終的意味表示をその出力としている点で、基本的に解釈意味論(Interpretative Semantics, IS)の下位理論に属するものと言える。

これに対置する形で登場したのが、Lakoff(1970, 1971), McCawley(1972), Postal(1970), Ross(1970)等によって提唱される生成意味論(Generative Semantics, GS)である。GSの基本的特徴は、ISにおける深層構造の概念を更に深化させ、変形規則の入力となる基底構造自体が、意味部門において、意味表示レベルとして規定される点にあると言える。形式的にみた場合、この意味表示レベルは、(a) S, V, NP等の限定された基本的範疇記号によって定義される句構造標識に基づいて表示されるレベルであり、(b) 選択制限・共起制限に関する意味規定、(c) 各種の推論規則に関する一般化が述べられるレベルである。この意味表示レベルから表層構造への派生過程に関する規則は、一連の変形規則(語彙変形規則・非語彙変形規則)に限定される。従って、上記ISが規定する意味での意味解釈規則は存在しないことになる。またISにおける深層構造は、すべての語彙変形規則の適用によって規定され、すべての非語彙変形規則の適用の入力となる自立的レベルとして仮定されるが、GSにおいては、意味表示から表層構造のレベルに至る派生過程に関しては、語彙変形規則が常に非語彙変形規則に先行して適用されるとは限らず、後者が前者に先行して適用される場合、あるいは、両者が交互に適用される可能性も認められる。従って、GSにおいては、ISで規定される自立的レベルとして

の深層構造のレベルは認定されないことになる。この語彙変形適用以前の抽象的意味表示から一連の変形派生を経て表層構造を規定していく方向は、Gruber (1965, 1967) 及び Chafe (1970) の意味モデルにもみられるが、広義に解するならば、これらのモデルは、基本的にはGSの路線に沿ったものと言える。

以上のIS, GSとの関連で独自の立場にある理論として、Fillmore (1966, 1968, 1971) に代表される格文法 (Case Grammar, CG) が考えられる。CGは広義には、IS, GSと同様生成理論の枠組に属するものであるが、その独自性は、格概念 (Case Notion) の導入にある。Fillmore (1966, 1968) では、この格概念を深層構造のレベルに導入している点でISに準ずる枠組に基づいている。しかし、CGはISが深層レベルで規定する主部／述部の統語的区分を認めず、同区分に中立的な述語と格の共起関係に基づく深層構造を仮定する。更に Fillmore (1971) ではこの深層レベルを深化させ、格設定を意味表示レベルにおいて導入するに至る。CGは、格という意味概念の導入に基づき、統語論と意味論の基底レベルにおける統合を図っている点で、やはりGSと基本的に同一の方向を目指した意味モデルを提示するものである。

以上の生成理論の意味論の展開をみた場合、概略 (i) 意味解釈規則の入力となる深層レベルを次第に表層構造に近づけていく方向と、(ii) 同レベルを更に深化させ、意味表示レベルと同一視していく方向が考えられる。前者の方向は、STからESTを経てRESTに改訂されていくISの修正過程を特徴づけるものであるが、この方向を推進する根拠としては、表層レベルに対する意味解釈上の役割の認定、統語論の自律性の仮説に基づく可能な変形規則の限定などが考えられる。また後者の方向は、GS, CG, Chafe, Gruberなどのモデルを特徴づけるものであるが、この方向を推進する根拠としては、ISの自律的な深層レベルからは一般的予測が不可能な統語事実、選択制限・推論規則に関する事実の認定が考えられる。が、いずれの方向をとるにせよ、少なくとも次の諸点は留意する必要がある。まずISは変形規則の役割を次第に軽減し、これに代って意味解釈規則の担う役割を増強しつつある。この方向は、一見統語論の自律性に基づき、可能な変形規則を規制する点では意味があるが、逆に意味解釈規則が不当に強力なものとなる危険性は免れない。特に最近のRESTの枠組では、痕跡理論の導入の下に、焦点・前提等の表層上の意味解釈だけでなく、

従来深層レベルで規定された文法関係、格関係等も一律に表層レベルで意味解釈されることになるが、この方向を押し進めるならば、知的意味関係の規定上の基盤としての深層レベル設定の根拠及びこの一般化を保证する変形規則の存在理由、ひいては深層構造と表層構造を区分する根拠が消失する可能性が当然予想される。他方、GSに代表される基底レベルの深化の方向は、ISの自律的深層レベルでは予測できない統語事実・意味事実の一般化を可能とするが、同時に意味表示レベルと表層レベルに介在する変形規則（特に、前者と従来の深層レベルに介在する語彙挿入以前の変形規則）の規制が問題となる。以上の方向の適否を決定していくためには、統語部門・意味部門の位置づけ（特に両部門の相互関係を規定する変形規則と意味解釈規則の位置づけ）、これらの規則の介在を通して得られる最終的意味表示の内部構造の明確化を図ると同時に、両部門と語用論の相互関係の明確化を図る必要がある。特に最近のGSの枠組は、単に基底レベルの深化にとどまらず、発話行為と統語論の関連づけ、会話の含意に基づく意味的・語用論的推論の明確化、変形規則の機能的側面からの再規定等の問題へとその研究領域を広げつつある。またRESTに至るISの枠組においても、照応化、移動規則に伴う省略などの問題を中心に、文レベル・文連結レベルの区分に基づく統語現象・意味現象の問題の再規定が行なわれてきている。但し、ISは形式文法の自律性の仮説に基づき、あくまで意味部門を含む文法を自律的部門として語用論から区分し、前者の形式的規定を主眼とする研究を押し進める方向にある。これに対し、GSはむしろ語用論の観点から統語部門・意味部門の諸現象を再規定する方向で各部門の機能的統合を図りつつある。この後者の方向は、Kunoの談話文法、Hallidayの体系文法に關しても方法論上の一致点が認められる。これらのいずれの方向がより有意義な知見を提示するか判断は現時点では下せないが、いずれにせよ生成理論における意味論の実質的な位置づけは、今後語用論を含む隣接部門との有機的な関連性を無視しては不可能となることは疑いない。

3. ヨーロッパにおける言語学の展開

3.1 ヨーロッパにおける1960年代までの言語学の歴史^十

3.1.1 はじめに

言語学の起源をたどると、古くはギリシアの哲学者や、サンスクリット語の有名なパーニニ文典(4世紀B. C.)までさかのぼることができる。またローマ帝国時代からルネサンスを通じても何人かの言語、文法研究者の名前が知られている。しかしながら現代に直接通じる本格的な言語学がおこってきたのは、19世紀におけるヨーロッパで、フンボルトらを中心とするロマン派言語学がその始まりであるとされている。

本稿では、ロマン派言語学以降の、いわゆる近代言語学について、まず1960年代までの歴史を概観する。この時期までの約150年間の歴史の中で、ヨーロッパにおいては、地味ではあるが、着実な歩みがしるされている。その言語観は、現在話題となっている、チョムスキーの理論にも大きな影響を与えているが、フェルディナン・ド・ソシュールなど一部を除いては、意外にその内容は知られていない。言語学史は一部の言語学者、ゲルマニストのみが知っていればいいという意見も確かに一理ある見方であろうが、1970年代になって特に盛んになってきた言語学の新しい潮流について理解するためには、そのバックグラウンドを知る必要があると思われる。

本稿では3.2において、70年代にはいって応用、展開の時期にさしかかったと思われる、結合価理論を取上げる。結合価理論は、構造主義言語学の特殊な形である依存関係文法の一タイプで、ヨーロッパでは変形生成文法とならんで最も重要な理論と見なされているが、これも我が国では余り知られていない。しかしながら既にドイツの日本語研究者によって、日本語への適用も試みられている。

3.3においては、ヨーロッパ言語学の最近(特に60年代以降)の成果である、テキスト言語学について述べる。テキスト言語学は、従来の文法では扱わなかった(あるいは扱えなかった)文以上の文法、すなわちテキストを解析しようという学問である。特に西独、スウェーデン、オランダなどで盛んに研究

十 本節の記述は、主として〔9〕〔14〕による。

が行われているが、ここではその一端を紹介する。

前述のように、ヨーロッパの現状は余り知られていない上に、文献も英語以外で書かれたものが少なくない。これらの文献は一応収集したものの、内容まで深く立ち入っていないものが多いので、本稿ではその一端を紹介するにとどめ、詳細な紹介は他日を期したいと考えている。

3.1.2 フンボルトからドイツ青年文法学派まで

ヴィルヘルム・フォン・フンボルトの名前は、弟で有名な地理学者、自然科学者であったアレクサンダー・フォン・フンボルトと共によく知られている。ヴィルヘルムの生涯（彼は政治家としても活躍した）については、[8]，[11]に詳しいのでここには述べないが、その言語学的な業績から、一般言語学の創始者とされている。

19世紀最大の言語理論家であった彼の思想が最もよく表されている研究は、インドネシア、ジャワ島のカーヴィ語に関する膨大な研究の序説として書かれた、「人間の言語構造の多様性と人類の精神的発展に及ぼすその影響について」という論文であると言われている。彼の主張によれば、言語とは、人間が心にもっている精神（これを彼は内部言語形式と呼ぶ）が外部に現われたものであり、従って言語は世界観を反映するものである。この理論によって彼は民族間の言語の相違、発音の違いなどを説明した。彼の言語理論は世界観理論と呼ばれ、今日に至るまで広く影響を与えている。たとえば、意味内容文法のグァイスゲルバーらは、20世紀にはいつからの後継者と見なされ、しばしば「新フンボルト学派」の名で呼ばれている。

フンボルトの研究がどちらかというと人間の精神的、内的なものに目を向けたものであったのに対し、言語学に自然科学的なものを取り入れて、言語を外的、歴史的観点から捕えようとする動きが19世紀後半のドイツでおこってきた。1870年代、ライプツィヒに本拠を置いた彼らの活動は、若い世代が保守的な世代に挑戦したという意味で、「青年文法学派」と呼ばれている。

青年文法学派の特徴は、言語を外的な、把握可能なものとして捕えようとしたところにあり、きわめて実証主義的である。中でも今日に至るまで特に注目すべき業績とされているのは、音韻の研究で、音声現象の生起規則に絶対的な

一貫性があることをはっきり主張したのは彼らが最初であるとされている。彼らの綿密、真摯な研究態度は後の言語の理論的研究に多くの影響を与えたが、しかしながら彼らの方法論的手法によって言語の全体像が見失われたことも否定できない。その後のドイツにおいては、この青年文法学派を克服しようとするさまざまな試みがなされたが、そのために言語外の要素がもちこまれ、言語体系そのものを対象とする研究に対して遅れをとることになった。ドイツでは1950～60年代くらいまで、言語学はある程度沈滞の時期を迎えるのである。

3.1.3 ド・ソシュールと構造主義

20世紀の初めにヨーロッパとアメリカではほぼ同じ頃始まった構造主義は、現代に至るまで言語学の大きな潮流となっている。アメリカでは、サピア、ブルームフィールドらを経て、ハリスの分布主義に至り（ハリスについては3.3でも多少触れる）、チョムスキーの登場となるが、本稿の範囲を逸脱するのでここでは述べない。一方、ヨーロッパにおいては、言語学のみならず、広く社会科学、人文科学の流れを巻きこんだ一大思想潮流となっている。^十

今日、構造主義の創始者と見なされているのはフェルディナン・ド・ソシュールである。青年文法学派（前述）などの指導を受けたド・ソシュールは、生涯に殆どまとまった著作を残さなかった（有名な「一般言語学講義」[12]は、生前の講義録に基づいて弟子達がまとめたものであるが、言語に対する彼の考え方を正確に反映しているかどうかについては議論のあるところである）が、彼の影響下になかった人々の理論的基盤をも含むような説を唱えたために構造主義の創始者と言われているのである。

彼の学説の最も基本的な点は、言語を体系として捕える立場である。この体系を彼は「ラング」と名づけ、それを具体化する「パロール」と区別する。そしてこの二つが合わさって人間の言語活動が成り立つとするものである。この、言語を体系として取り扱う立場こそ、構造主義言語学の基本的な考え方であり、今日までその底流となっているものである。

構造主義は、たとえば民俗学などの社会科学にも影響を及ぼしている（たとえばレヴィ＝ストロースなど）が、言語学の分野では、（アメリカを別として）

十 構造主義の全体的な流れを概観するものとしては[15]が詳しい。

大きく分けて3つの基本的な型があるとされている。それは、ド・ソシュール
の思想をそのまま受けついだジュネーヴ学派、音韻論の分野で特にすぐれた業
績をあげたブラーグ学派、言理学の名前で代表されるコペンハーゲン学派であ
る。このうち、ブラーグ学派は構造主義の一般的側面を最もよく表し、後のテ
キスト言語学（3.3 参照）への影響も大きいので、次節で取り上げることに
する。

3.1.4 ブラーグ学派

1926年、ブラーグに、チェコの言語学者マテジウスらを中心とした「ブラー
グ言語学サークル」という言語学者の集団が結成された、これには当初からロ
マン・ヤコブソン、1928年からトルベツコイが参加し、1929年にはブラー
グ言語学サークル紀要の創刊号に彼らの綱領を発表した。この学会そのものの
活動は10年ほどしか続かなかったが、その後各地に移り住んだこれらの学者
とその後継者たちを含めて「ブラーグ学派」と呼んでいる。

ブラーグ学派の思想を伺い知るにはその綱領の概要を見ればよい。それはお
よそ次の通りである〔9〕。

- ・言語は相互理解のための表現手段の体系であるから、言語学者は個々の発話の現実の機能を研究すべきである。

- ・言語は現実の物理現象であり、その形態は非言語的要因に大きく制約される。従って文化一般の言語と文学作品の言語等の間には、理論的にも実際的にも区別を認めることが是非とも必要である。

- ・言語は人格の知的表明と情的表明の両者を包括している。従って言語研究はこの両者を伝える言語形態の関係をも取り入れなければならない。

- ・書き言葉と話し言葉は同一ではなく、それぞれ固有の性質を持っているので、この両者の関係は科学的に研究すべきである。

- ・共時態と通時態の問題。⁺

- ・言語学における比較方法。⁺

- ・音韻論的研究は、まず第一に個々の言語における音韻論的対立類型の規定

⁺ いずれも詳細は略す。

に努めなければならない。形態論的現象は、音韻論的現象と切り離すべきではない。音韻論的対立はしばしば形態論のレベルで意味を持つからである。

以上のような綱領によって（それ自体は非常に基本的なものであるが）プラグ学派は豊富な業績を生み出した。その中で特筆すべきはトルベツコイ〔13〕、ヤコブソンの音韻論の研究と、機能的文パースペクティブであろう。前者は、いわゆる弁別的特徴で知られている。これは、ある音素を他の音素と弁別する音特性のことであるが、ヤコブソンは、実験音声学を取り入れることによって、「弁別的対立は二分法の原理に従って組織されている」と主張し、音素に対して非常に美しい体系を示した。

また、機能的文パースペクティブは、文を形式的構造ではなく、古い情報であるテーマと、新しい情報であるレーマという観点から分析するものである。この考え方が、文以上の文法というテキスト言語学に大きな影響を与えている（3.3参照）。

3.1.5 その他の言語学派

その他については種々あるが、ここでは触れる余裕がなかったので全体的な流れを示す図をここに入れておくことにする（図3-1）。これで前述のものだけでなく、フンボルト以来現代に至る流れがだいたいつかめるものと思う。

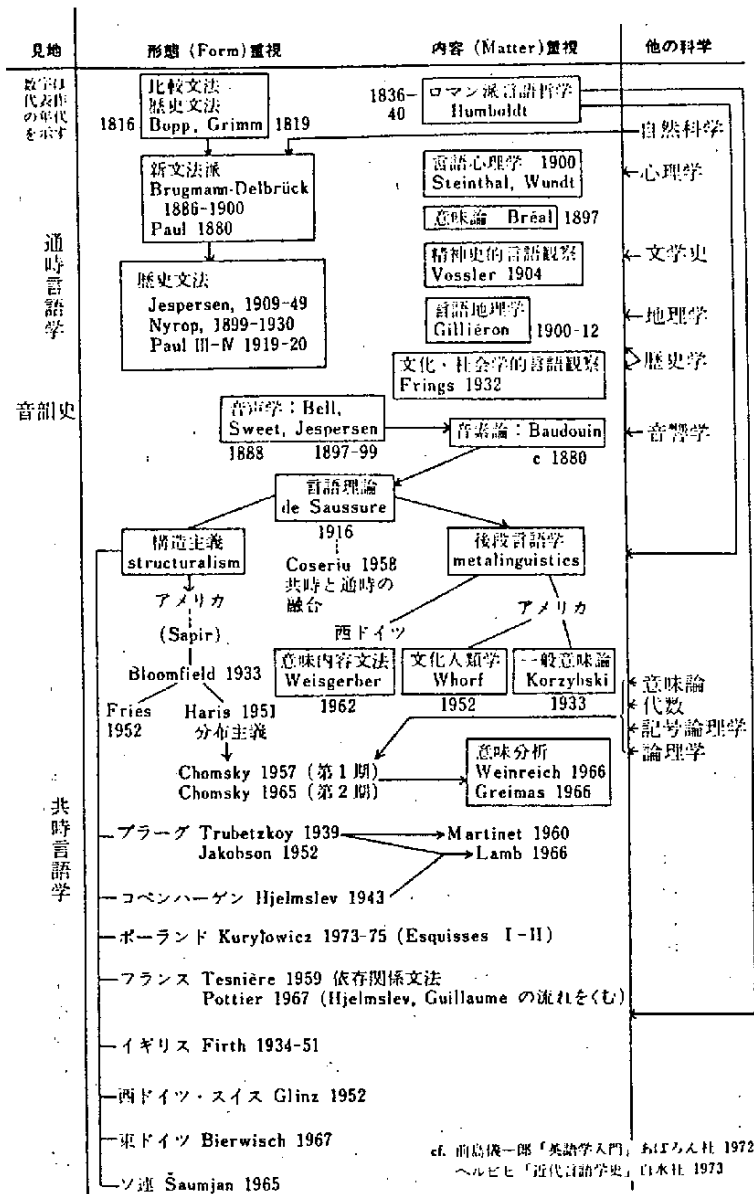


図 3-1 言語学派系統図 (「言語」第 6 巻 5 号 (1977) による。)

3.2. テニエールと結合価理論

3.2.1 テニエールの結合価理論

依存関係文法は、構造主義言語学の特殊な一形態と見られ、その中にも種々の形態があるが、その代表的な唱導者がテニエールである。以下に彼の理論の概略を紹介する〔10〕〔14〕。

依存関係文法において、句や文を生み出すいわゆる連結部門での連結は、部分と全体の関係を記述する（構成要素文法）のではなく、もっぱら部分に頼って、ある要素に別の要素より高い地位が与えられるという記述をとる（たとえば動詞は名詞を支配するなど）。テニエールは一連の発話連鎖を動詞を頂点とする依存関係の樹形図に分解し、文レベルとそれ以下の成分の区別を、樹形図上の位置で示す^十。従って文の構造的分析は動詞から出発する。動詞に直接従属する成分は、共演成分（actants）と状況成分（circonstants）に分類される。前者は「行為に参加する」動詞の従属成分で、フランス語では第一共演成分（伝統文法における主語）、第二共演成分（直接目的語）、第三共演成分（間接目的語）の3つがあり、いずれも動詞に直接依存し、樹形図上では同じレベルに置かれる。すなわち伝統的な主語と述語という対立は、第一と第二の共演成分という構造的な差になる。状況成分は「行為の状況」（時、所、仕方など）を表し、文の中で動詞に従属する成分である。状況成分の数は無制限である。

テニエールは一定数の共演成分を伴う動詞の能力を原子価にたとえて、動詞の結合価または動詞価（Valenz）と名づけた。テニエールは結合価の数によって動詞をゼロ価から四価に分類（ただしフランス語では四価の動詞は存在しない）した。これが今日の意味での結合価の概念であるが、テニエールの理論にはなお欠点があり、後にドイツを中心に彼の理論が修正された。70年代には結合価理論はほぼ完成の域に達し、業績をまとめたものも出版されている〔4〕。

十 この樹形図により外面的発話連鎖と内部的構造様式が区別されるが、これは後に多少修正されて生成文法の表層構造と深層構造の区別となって再登場する。

3.2.2 結合価理論の完成

依存関係文法は生成文法と並んで20世紀における最も重要な理論といわれ、その重要な中核をなす結合価の概念に対してさまざまな研究が行われた。そして結合価をどの範囲まで認めるかによって種々の主張があらわれた。今日東独のヘルビヒ、西独のエンゲルによって結合価理論はほぼ完成したと言われているが、その概要は次の通りである〔10〕。

- (1) 文成分の割りだしと動詞の結合価によるその区別。動詞に関係してくる文成分を、すべての動詞と結合しうる「自由規定成分」と動詞の結合価に依存する「補足成分」に区別する。
- (2) 補足成分のクラス分け。文の中の成分を入替えることによって(代名詞などで)統合機能が等しいものを見つけ出す(エンゲルは10のクラスに分けている)。
- (3) 補足成分のタイプ。文中に何個の、どのクラスの補足成分が生じるかは動詞の結合価に依存している。エンゲルは補足成分のこの組み合わせを基本文型と呼び、ドイツ語に関する最も重要なものとして30の型をあげた。しかしながらこの基本文型のあるものは、補足成分を取り除いても文法性の損われないものがある。このため補足成分をもう一段階区分し、補足成分を取り除いたことによって文が非文法になれば「必須補足成分」、そうでなければ「任意補足成分」という。
- (4) 文形式の補足成分。補足成分は常に語というわけではなく、副文や不定詞の形をとることもある。これらも結合価に係わると考える。

3.2.3 結合価辞典

動詞結合価の理論を外国人のためのドイツ語教育に応用しようという試みが70年代の初めから盛んになってきた(たとえば〔1〕)。これは、動詞を中心に示えることにより、文の構造を体系的に示すことができるようになり、外国人に文法的に正しい文を容易に作らせることができるという考え方に基づいている。

その目的のために、西独と東独から相次いで動詞の結合価辞典が刊行された〔3, 5〕。その詳しい内容については省略するが、いずれも上述の理論に基づ

き約500の動詞を並べたものである。また、注目されるのは日本語の結合価辞典が刊行されていることである〔6〕。これは「あがる」、「扱う」などの日本語動詞について、補足成分の種類を示し、それに対応するドイツ語動詞（複数）に対して同じことを示してある。この辞典の特徴は、日本の新聞、雑誌から選んだ豊富な文例に、それぞれ独語訳がついていることで、日本語を研究するものにとっても寄与するところが大きいと思われる。

なお、1978年に一般の辞書としては初めて基本文型を付した辞典が現われた〔7〕。語数は16,000と多くはないが、やはり外国人でドイツ語を学ぶものにとっては役に立つと思われる。

3.3 テキスト言語学^十

3.3.1 テキスト言語学の歴史

これまで述べてきた言語学上の種々な研究は、すべて文という範囲内にとどまるものであった。しかしながら、人間が対話したり本を読んだりする時には必然的に文のレベルを越えたつながりが問題になってくる。このような文以上の文法を補えようとする試み^{十一}の一つがテキスト言語学である。

テキスト言語学の歴史をたどって見ると、三つの精神的な流れがあげられる。最も古いものはギリシア時代までさかのぼることができる雄弁術で、考え方の整頓や朗読法に既にその萌芽が見られる。第二の先駆は文体論で、これがその当時の言語学では扱わなかった文以上の領域を扱っていた。第三の先駆は前にも少し触れたが、物語の構造分析をした民俗学者たち（レヴィ＝ストロースなど）である。

現在のテキスト言語学に対する最初の衝撃はZ. S. ハリスの「談話分析」についての研究によって与えられた（1952年）。この中で彼はアメリカ分布主義の構造的方法をテキストの範囲にまで拡張した。彼はテキストを部分文に対応する文列に分割し、これらの文列の等価性を確立した。

十 本節の記述は主として〔2〕による。

十一 文以上に対しては統語論的關係は成り立たないという説もあるが、ここではその立場はとらない。

一方、ヨーロッパでは、前述のように、プラグ学派において機能的文パースペクティブという形でテキストが扱われてきた。知識（古い情報）としてのテーマと、新しい情報としてのレーマは、テキストの枠内でのみ初めて可能になる。

上記の2つの研究や、生成理論などの影響を受けて、^十ヨーロッパ、特にドイツにおいてはテキスト言語学の研究が盛んに進められた。現在ドイツではミュンスター、コンスタンツ、ビーレフェルトなどの各市を中心に研究が進められている。また、ドイツ以外でも、オランダ、スウェーデンなどに意欲的な研究者たちがおり、今後もこれらの国、都市を中心に研究が進むものと思われる。

3.3.2 テキスト言語学の論理的発展

テキスト言語学の論理的（あるいは歴史的）発展は、三段階に分けることができる。第一の段階では、文が最高の言語単位であると見られていた。つまり、言語（ラング）は文の和であり、テキストは文がゆるく隣り合って集積したものとされていた。テキストの規則性は、もしそれが現われたならば、パロールまたは文体論の観点から調べられた。しかしながら文法を文文法にのみ限ると、文を越えた規則性を説明する際には、少なくとも次の2つの不都合な点が生じる。

- (1) あるテキストは、階層的に部分文に分れる一つの長い文と考えることができる。しかしながらこの考え方は対話には適用できない。
- (2) 文の使用や受入可能性に対して文脈の制約を確立することができる。それによって言語的、テキスト的環境は、文がその枠組の中で表されるような言語外の文脈的状况と等置される。これによって言語学の領域が制限されることになる。

第二段階では発話（英語 utterance, 独語 Äußerung, 仏語 énoncé, 露語 высказывание）の単位が認識された。これは一つまたはそれ以上の文から成っている。最小の単位としては、ある人間の連続した発言であり、この人間の沈黙か、他の人間への会話の移行で区切られる。しかしながら、発話を最高単位と考えた時には、多くの人々の対話は言語学的単位と見なされない。

十 アメリカにおいては、生成理論の研究者たちはテキストという単位には目を向けなかった。

第三段階ではテキスト（英語 discourse，独語 Text，仏語 discours，伊語 discorso，露語 tekst）が最高の言語単位と見なされる。人間はテキストの中で話し、書く。テキストは人間がその中で自分を表現しようとする主要な言語記号である。全体としてのテキストは完結した意味を持ち、従って翻訳に対してはまた適当な単位でもある。今日テキスト言語学ではこれを単位として研究が進められている。

3.3.3 テキスト言語学の構成

テキストを基本的な単位として用いるテキスト言語学は、通常の文文法と同じように、種々の部門に分けることができる。その分類を表3-1に示す。

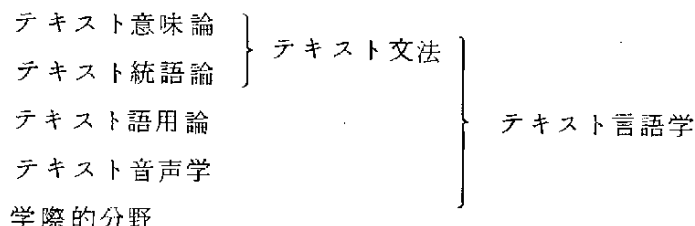


表3-1. テキスト言語学の構成

テキスト意味論は、文字通り、テキストの表す意味を研究する部門である。たとえばテーマとか意味連続などがこれに属する。この意味の統語的な表現を扱うのがテキスト統語論である。テキスト言語学においては意味論と統語論の区別は必ずしも明確ではないが、たとえば照応とか語構成などは統語論に属するであろう。文脈中のテキストの機能を扱うのが、テキスト語用論である。さらに、テキスト音声学では、統語論的なテキスト構成の音声学的な特徴について取り扱う。また、心理学、通信理論などとの関連を扱うのが学際的分野である。

テキスト言語学は目下発展しつつある学問領域であり、既にいくつかの叢書も刊行されつつあるが、それらの詳しい内容については他日に譲りたい。

3.4 「おわりに」に代えて

本来この節においては、ド・ソシュールの共時態とを統合したコセリウの言語理論(図3-1参照)、モンタギュー文法のドイツにおける受容、ロシアの記号理論などについて述べる予定であったが、種々の都合から他の機会に譲ることにする。

近代言語学の発祥の地でありながら、言語学者以外には殆ど知られていなかった、ヨーロッパの言語学の歴史とその発展の有様を、そのほんの一端を伺い知ることの役に立てれば幸いである。

参考文献

- [1] K. Adler, B. Steffens : Deutsch für die Mittelstufe
Texte und Übungen, Max Hueber Verlag, 1974.
- [2] W. Dressler : Einführung in die Textlinguistik, Max
Niemeyer Verlag, 1973.
- [3] U. Engel, H. Schumacher(Hrsg.) : Kleines Valenzlexi-
kon deutscher Verben, TBL Verlag Gunter Narr, 1978.
- [4] G. Helbig(Hrsg.) : Beiträge zur Valenztheorie, Mouton
1971.
- [5] G. Helbig, W. Schenkel(Hrsg.) : Wörterbuch zur Valenz
und Distribution, Leipzig, 1978.
- [6] J. Rickmeyer : Kleines japanisches Valenzlexikon, Hel-
mut Buske Verlag, 1977.
- [7] G. Wahrig(Hrsg.) : dtv Wörterbuch der deutschen Spra-
che, Deutscher Taschenbuch Verlag, 1978.
- [8] 泉井久之助 : 言語研究とフンボルト, 弘文堂, 1976.
- [9] M. イグニッチ, 早田, 井上訳 : 言語学の流れ, みすず書房, 1974.
- [10] 榎本久彦 : 結合価理論とその成果, 言語第8巻11号 pp. 46-53,
1979.
- [11] 亀山健吉 : フンボルト, 中公新書525, 中央公論社, 1978.
- [12] ド・ソシュール, 小林訳 : 一般言語学講義, 岩波書店, 1972.
- [13] トルベツコイ, 長嶋訳 : 音韻論の原理, 岩波書店, 1980.
- [14] G. ヘルビヒ, 岩崎他訳 : 近代言語学史, 白水社, 1973.
- [15] M. レイン編, 篠田監訳 : 構造主義, 研究社出版, 1978.

4. Montague 文法

Montague 文法については既にかんりの文献があり、この小論でその全てを紹介するのは不可能と思われる。そこで本論は、Montague 文法の理論的な枠組を理解することに主眼を置き、しかるのちに伝統的な文法理論の Formal な側面を如何に抽出するかについて、いくつかの文献をあたってみることにしたい。

私個人としては 20 世紀初頭以来の哲学的言語分析の努力が Montague 文法の出現によってようやく実際の自然言語に対する文法的枠組を与えることができたという点にまづ喜びと賞賛の意を表したい。

哲学者達の関心は当初から自然言語とその論理にあったと思われるが、自然科学の基礎を与える試みは多くのみのり豊かな成果を生み出したと思われる反面、自然言語それ自体乃至はその論理については歩みは極めて遅々としていたと思わざるを得ない。

最初に自然言語分析のための Formal system として Modal Logic が発明されたが、それは

◇ …… であるらしい

□ …… 絶対そうだ

という operator を propositional logic に付加したものであった。

Modal Logic の発明は比較的早かったが、述語論理のその後の発展に比較すると長い間、低迷していたように思われる。

その後 Modal Logic は第 1 階述語論理の上に拡張され Kripke によって model が作られるに及んで modality の内容が明確化され、それ以前の数学的論理学の成果を完全に吸収するに至る。Kripke による modality の解釈は possible world ω に関するものだと言えよう。

◇ $\leftrightarrow \exists \omega$

□ $\leftrightarrow \forall \omega$

Intentional Logic は文が話される状況を時間、場所、人等、様々な index で表わし

◇ $\leftrightarrow \exists (t, \text{place}, \text{person}, \dots)$

$$\square \leftrightarrow V(t, place, person, \dots)$$

と考えるという解釈をほどこすと同時に Logic が扱う対象として Predicate だけでなく, function を導入することとした。

Carnap は或る文の外延即ち $\square$ を成立せしめるような要素の集合は時の状況 $\omega = (t, place, person, \dots)$ に依存するものであるから

$$Ext(\square) = Ext(\square, \omega)$$

と考えるべきであるという提案を行なった。すると $\square$ の内包は

$$Int(\square) = \lambda \omega Ext(\square, \omega)$$

と定義される。

Montague と Scott はこの考え方に従って文の全ての構成要素が時の状況 ω に依存するものと考えた。いま個体の集合として人間全体をとるものとし、これを D としよう。すると自動詞 $verb$ は

$$P_{verb}(\omega, person) \in \{0, 1\}$$

即ち

$$verb = \lambda \omega \lambda person P_{verb}$$

また $person$ は

$$person(\omega) \in D$$

であるから

$$person = \lambda \omega person$$

文の構成要素をこのように ω に依存するものとし、またそれを関数ととらえることにより、文がその構成要素関数の合成で形成されるという Frege の原理が初めて整合性をもって成立することとなった。

Intentional Logic ではこのような文の断片についても語りうるために function を導入しているわけである。

Intentional Logic の成立の上に立って Montague は彼の文法理論を展開することができた。その特徴は、

- ① 文構成要素に Intentional Logic の function を assign する
- ② 文法要素は、その function の type と correspondence をとるというものである。Montague 文法は文法要素と関数 type との correspondence をとるという美的要請のために、Chomsky 流の文法体系とは著しく

異なる文法体系をとることになったが、Cooper が生成文法体系に Intensional Logic により semantics を導入し、Montague 文法との同等性を証明してより、両 system の違いが形式的なものであることが認識されるようになった。

4.1 Logic

4.1.1 Modal Logic

ここでは第1階論理を含む Modal Logic の system として Lewis による system $S4^*$ を紹介しよう。

system $S4^*$ の記号

- a 1 propositional variables $p_1, p_2, \dots$
- a 2 n 項 predicate variables $p_1(x_1, x_2, x_3), \dots$
- a 3 自由変数としばり変数
- a 4 論理記号 $\neg, \vee, \exists, \Box$

system $S4^*$ の WFF

- b 1 propositional variable は wFF
- b 2 A が WFF なら $\neg A, \Box A$ も wFF
- b 3 A と B が wFF なら $A \vee B$ も wFF
- b 4 $\mathcal{A}(a)$ が wFF なら $\exists x \mathcal{A}(x)$ も wFF

system $S4^*$ の Axiom

- c 1 $\mathcal{A}(a) \Rightarrow \exists x \mathcal{A}(x)$
- c 2 $\Box A \Rightarrow A$
- c 3 $\Box(A \Rightarrow B) \Rightarrow (\Box A \Rightarrow \Box B)$
 $\Box A \Rightarrow \Box \Box A$

system $S4^*$ の deduction

- d 1 $A, A \Rightarrow B \vdash B$
- d 2 $\mathcal{A}(a) \Rightarrow B \vdash \exists x \mathcal{A}(x) = B$
- d 3 $A \vdash \Box A$

但し $A \Rightarrow B \equiv \neg A \vee B$

様相論理に於いて

$$\Diamond A \equiv \neg \Box \neg A$$

と置き

$\Diamond A \sim A$ であるらしい

と読むと、自然言語の機能の一つ……であるようだ、という文体を吸収できる。しかし

$\Box A \sim$ 必然的に A である

という解釈は自然言語の語感としては、ちょっと異様なものである。歴史的には modal logic は propositional logic に $\Diamond$ を operator として導入することによって誕生した。命題論理は $\langle 0, 1, \geq, \neg, \Box \rangle$ のなす Boolean algebra であるが

$\Diamond$ は operator

$\Diamond : B \rightarrow B ; \Diamond P \leq P$

として導入されたわけである。すると

$\Box = \neg \Diamond \neg$

が B で定義されねばならないわけである。

ところで

$\neg P = \neg P$

$\Diamond \neg P \leq \neg P$

$\neg \Diamond \neg P \geq \neg \neg P = P$

$\Box P \geq P$

であるから $\Box P$ は「 P が成立する」という内容より強い内容を表わしている。おかげで $\Box P$ の解釈に悩むことになったものと思われる。

4.1.2 Kripke Model

Kripke は modal logic の体系に model を導入することに成功した。その主な点は

① possible world の collection を考え

② $\Diamond A \doteq \exists \alpha A$ 但し α は possible world

とすることである。

すると

$\Box A \doteq \forall \alpha A$

なる自然な解釈が得られるわけである。

形式的には次のような定義が与えられる

a. (M, R, V) が modal structure であるとは

a 1 $M : \emptyset$ でない set で possible world の全体

a 2 $R : M$ 上の反射的推移的二項関係

a 3 $V : \text{各 possible world } \alpha \text{ に対し通常の意味での model の基底集合 } V(\alpha) \text{ を対応させる map で}$

$$\alpha R \beta \Rightarrow V(\alpha) \subset V(\beta)$$

を満たす。

b. $S4^*$ と modal structure (M, R, V) が与えられたとき, $S4^*$ の (M, R, V) への解釈 W は

b 1 自由個体変数 a に対し

$$W(a) \in \bigcup_{\alpha \in M} V(\alpha)$$

b 2 命題変数 p と各 possible world α に対し

$$W(p, \alpha) \in \{1, 0\}$$

b 3 各 n -place predicate p と各 possible world α W に対し

$$W(p, \alpha) \in P(V(\alpha)^n)$$

即ち $V(\alpha)$ の要素の n -組からなる或る set を与える。

c. いま $S4^*$ の式 F に対しその W による値を次のように定義する。

F の自由個体変数を $\bigcup V(\alpha)$ の要素の名前につけ変えたものを F' とする。

c 1 F' が命題変数であれば $W(F', \alpha)$ も命題変数

$$\text{従って } W(F', \alpha) = 1 \Leftrightarrow W(F', \alpha) = 1$$

c 2 $F = p(\xi_1, \dots, \xi_n)$ で p が n -place predicate,

$\xi_1, \dots, \xi_n$ が $\bigcup_{\alpha} V(\alpha)$ の元とするとき

$$W(F', \alpha) = 1 \Leftrightarrow (\xi_1, \dots, \xi_n) \in W(p, \alpha)$$

c 3 $F = \neg A'$ のとき

$$W(F', \alpha) = 1 \Leftrightarrow W(A', \alpha) = 0$$

c 4 $F' = A' \vee B'$ のとき

$$W(F', \alpha) = 1 \Leftrightarrow W(A', \alpha) = 1 \text{ か } W(B', \alpha) = 1$$

c 5 $F' = \exists x \mathcal{A}'(x)$ のとき

$$W(F', \alpha) = 1 \Leftrightarrow \exists \xi \in V(\alpha) \quad W(\mathcal{A}'(\xi), \alpha) = 1$$

c 6 $F' = \Box A'$ のとき

$$W(F', \alpha) = 1 \Leftrightarrow \alpha R \beta \text{ なる任意の } \beta \text{ について}$$

$$W(A', \beta) = 1$$

c 7 $F' = \Diamond A'$ のとき

$$W(F', \alpha) = 1 \Leftrightarrow \alpha R \beta \text{ なる } \beta \text{ があって}$$

$$W(A', \beta) = 1$$

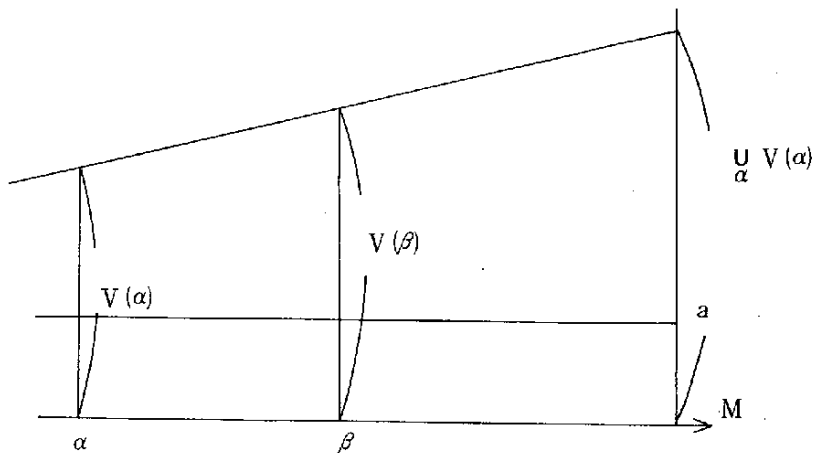
そこでもとの式 F に対しては、それに $\bigcup_{\alpha} V(\alpha)$ の元を代入したときの $W(F', \alpha)$ の値をそのときの値として採用することとする。

このとき (M, R, V) が W のもとで $S4^*$ の model であるとは

F が $S4^*$ の公理から導びかれるならば

$$\forall \alpha \quad W(F, \alpha) = 1$$

model を図に書くと



図の説明

R を例えば $\leq$ とすると possible world の全体 M は一直線上に並ぶ M 上

の点 α に対し possible world α での通常の意味での model の基底 $V(\alpha)$ が与えられるわけである。 $\alpha \leq \beta$ ならば $V(\alpha) \subset V(\beta)$ である。また個体変数 a は各 world で同一解釈を与えられる。

この model 解釈に於いて最も奇異に感ずるのは二項関係 R であろう。しかしこの二項関係 R こそ、modal logic を単なる predicate calculus から引き離すカギである。

いまもし

$$\Diamond A \equiv \exists \alpha A \text{ at } \alpha$$

とすると

$$\begin{aligned} \Diamond \exists x A &\equiv \exists \alpha [\exists x A] \text{ at } \alpha \\ &\equiv \exists x \exists \alpha [A] \text{ at } \alpha \\ &\equiv \exists x \Diamond A \end{aligned}$$

が成立することになる。

しかるに

$$\Diamond A \equiv \forall \alpha \exists \beta \alpha R \beta \wedge [A] \text{ at } \beta$$

であるために

$$\Diamond \exists x p(x) \wedge \neg \exists x \Diamond p(x)$$

とすることができる。

それには2つの異なる世界0と1をもち、世界0で P は恒偽、1なる世界では p は真となることもあるとする。 $\exists x p(x)$ は世界0で偽であるが世界1では真であるので $R = \leq$ とすると

$$\Diamond \exists x p(x) \text{ が成立つ。}$$

しかるに世界0の individual a は世界1の individual であり $p(a)$ がどちらの世界でもなり立たないようにできるから $\neg \Diamond p(a)$ が成立。世界0が individual として a だけしかもたないと決めると $\forall x \neg \Diamond p(x)$ が世界0で成立。これは $\neg \exists x \Diamond p(x)$ を意味する。

以上から Barcan formula と呼ばれる次の式

$$\Diamond \exists x p(x) \Rightarrow \exists x \Diamond p(x)$$

がこの model で成立しないことがわかる。

Kripke の model 解釈によって

$$[\Box A]_{\alpha} \equiv \forall \beta, \alpha R \beta [A]_{\beta}$$

即ち $\Box A$ が世界 α で成立するという事は $\alpha R \beta$ なる任意の世界で A が成立することであるという解釈が与えられた。

ここで R を時間関係、即ち

$$\alpha R \beta \quad \beta \text{ は } \alpha \text{ の過去}$$

という解釈をほどこすと

$$[\Box A]_{\alpha} \quad \alpha \text{ に至るまでずっと } A \text{ が成立してきた}$$

従って

$\Box A$ 現在ばかりでなく、これまでもずっと A が成立してきたと解釈できる。

4. 1. 3 Intentional Logic

possible world という概念は context という概念に密接な関連があると予想できる。そこで context を包含するような論理体系を設定するという考えが生まれることになる。

ここでは Montague の Proper Treatment of Quantification in Ordinary English (以下 P. T. Q と略す) に従って tensed Intentional Logic を紹介する。

A. Type

s, e, t をそれぞれ situation, entity, truth value を表わすものとして

A 1 e, t は type

A 2 α, β が type $\Rightarrow \langle \alpha, \beta \rangle$ も type

A 3 α が type $\Rightarrow \langle s, \alpha \rangle$ も type

B. Meaningful Expression (ME)

B 1 type a の constant と variable は $ME a$ に属す

B 2 $\alpha \in ME a$ で u が type b の変数 $\Rightarrow \lambda u \alpha \in ME \langle b, a \rangle$

B 3 $\alpha \in ME \langle a, b \rangle, \beta \in ME a \Rightarrow \alpha(\beta) \in ME b$

B 4 $\alpha, \beta \in ME a \Rightarrow \alpha = \beta \in ME t$

B 5 $\phi, \psi \in ME t$ で u が変数 $\Rightarrow \neg \phi, \phi \wedge \psi, \phi \vee \psi$

$\phi \rightarrow \psi, \phi \leftrightarrow \psi, \exists u \phi, \forall u \phi, \Box \phi, W \phi,$

$H \notin ME_t$

B 6 $\alpha \in ME_a \Rightarrow \wedge \alpha \in ME \langle s, a \rangle$

B 7 $\alpha \in ME \langle s, a \rangle \Rightarrow \vee \alpha \in ME_a$

とするとき $ME = \bigcup_a ME_a$

u が type b の変数であるとき u が type b の object を denote すれば $\lambda u \alpha$ は式 α に現われる u を全てその object で入れ換えたものを denote する関数である。

$\alpha \in ME \langle a, b \rangle$ であるときは α は type a の object から type b の object への関数と解釈され、 $\alpha(\beta)$ は α の argument に β を代入して得られた結果を表わす。

また

$\Box \sim$ it is necessary that

$W \sim$ it will be the case that

$H \sim$ it has been the case that

と読まれる operator である。

$\wedge \alpha$ は式の内包を表わす。即ち context が決まると denote するものが決まる関数である。

α がもし内包を表わしていれば $\vee \alpha$ はその外延を表わす。

α が context に依存していると考えれば

$\wedge \alpha \sim \lambda s \alpha$

と見なせる。すると $\beta = \lambda s \alpha$ の外延は

$\vee \beta \sim \alpha(s)$

とほぼ見なせるわけである。

Intentional Logic の Meaningful Expression に対し、その model 解釈を与えよう。

A, I, J をそれぞれ set of entities, set of possible worlds, set of moments of time とする。

このとき

C. Da, A, I, J : A, I, J に対する type a の object の possible denotation の全体を次のように recursive に定義

する。

$$C1 \quad D_e, A, I, J = A$$

$$C2 \quad D_t, A, I, J = \{0, 1\}$$

$$C3 \quad D_{\langle a, b \rangle}, A, I, J = D_{b, A, I, J}^{D_{a, A, I, J}}$$

$$C4 \quad D_{\langle s, a \rangle}, A, I, J = D_{a, A, I, J}^{I \times J}$$

ここで set A, B に対し A^B は B から A への関数の set $A \times B$ は A と B の元のペア $\langle a, b \rangle$ の set を表わす。

$C1$ は entity の possible denotation 全体が A であるということ, $C2$ は真理値の全体が $\{0, 1\}$ であることをいっている。type $\langle a, b \rangle$ は type a の object から type b の object への関数の type であったから possible denotation についても対応することが言えねばならない。これが $C3$, $C4$ は type a の expression は possible world I と moments of time に依存する type a の object への関数のことであることを表わしている。

Intentional Logic の constant symbo の structure $\langle A, I, J \rangle$ への解釈を次のように定義する。

d1 $\leq$ は J 上の大小関係

d2 F は constant に対してその denotation を定める関数で

$$\alpha \in \text{Con}_a \Rightarrow F(\alpha) \in D_{\langle s, a \rangle}, A, I, J$$

また g を Intentional Logic の variable に対して値を assign するもの即ち

$$d3 \quad u \in \text{Var}_a \Rightarrow F(u) \in D_{a, A, I, J}$$

とする。このとき Intentional Logic の式 α の structure $\langle A, I, J \rangle$ への $\langle \leq, F, g \rangle$ に於ける解釈は次のように定義される。

E. $\mathcal{U} = \langle A, I, J, \leq, F \rangle$ とし式 α の $\mathcal{U}_{i, j, g}$ に於ける値を $\alpha^{\mathcal{U}_{i, j, g}}$ を

$$e1 \quad \alpha \text{ が constant なら } \alpha^{\mathcal{U}_{i, j, g}} = F(\alpha) (\langle i, j \rangle)$$

$$e2 \quad \alpha \text{ が variable なら } \alpha^{\mathcal{U}_{i, j, g}} = g(\alpha)$$

$$e3 \quad \alpha = ME_a \text{ で } u \text{ が type } b \text{ の変数のとき}$$

$$[\lambda u \alpha]^{\mathcal{U}_{i, j, g}} : D_{b, A, I, J} \rightarrow D_{a, A, I, J}$$

$$\text{で } c \in D_{b, A, 1, J} \quad g'(y) = \begin{cases} g(y) & y \neq u \\ c & y = u \end{cases}$$

とすると

$$[\lambda_u \alpha] \mathcal{M}, i, j, g \quad (c) = [\alpha] \mathcal{M}, i, j, g'$$

$$\text{e 4} \quad \alpha \in \text{ME}_{\langle a, b \rangle} \text{ で } \beta \in \text{ME}_a \text{ のとき} \\ [\alpha(\beta)] \mathcal{M}, i, j, g = \alpha \mathcal{M}, i, j, g \quad (\beta \mathcal{M}, i, j, g)$$

$$\text{e 5} \quad \alpha, \beta \in \text{ME}_a \text{ なら} \\ [\alpha = \beta] \mathcal{M}, i, j, g = 1 \quad \stackrel{\text{fD}}{=} \quad \alpha \mathcal{M}, i, j, g = \beta \mathcal{M}, i, j, g$$

$$\text{e 6} \quad \emptyset \in \text{ME}_t \text{ なら} \\ [\neg \emptyset] \mathcal{M}, i, j, g = 1 \quad \stackrel{\text{fD}}{=} \quad [\emptyset] \mathcal{M}, i, j, g = 0$$

$\wedge, \vee, \rightarrow, \leftrightarrow$ についても同様

$$\text{e 7} \quad \emptyset \in \text{ME}_t \text{ で } u \text{ が type } a \text{ の変数とする。} \\ [\exists u \emptyset] \mathcal{M}, i, j, g = 1$$

$$\stackrel{\text{fD}}{=} \quad \exists c \in D_{a, A, 1, J} \quad g'(y) = \begin{cases} g(y) & y \neq u \\ c & y = u \end{cases}$$

$$\Rightarrow [\emptyset] \mathcal{M}, i, j, g' = 1$$

$$[\forall u \emptyset] \mathcal{M}, i, j, g \text{ についても同様}$$

$$\text{e 8} \quad \emptyset \in \text{ME}_t \text{ ならば} \\ [\Box \emptyset] \mathcal{M}, i, j, g = 1 \quad \stackrel{\text{fD}}{=} \quad \forall i', j' \quad \emptyset \mathcal{M}, i', j', g = 1 \\ [W \emptyset] \mathcal{M}, i, j, g = 1 \quad \stackrel{\text{fD}}{=} \quad \exists j' > j \quad \emptyset \mathcal{M}, i, j', g = 1 \\ [H \emptyset] \mathcal{M}, i, j, g = 1 \quad \stackrel{\text{fD}}{=} \quad \exists j' < j \quad \emptyset \mathcal{M}, i, j', g = 1$$

$$\text{e 9} \quad \alpha \in \text{ME}_a \text{ ならば} \\ [\wedge \alpha] \mathcal{M}, i, j, g = \lambda_i \lambda_j \quad \alpha \mathcal{M}, i, j, g$$

$$\text{e10} \quad \alpha \in \text{ME}_{\langle s, a \rangle} \text{ ならば} \\ [\vee \alpha] \mathcal{M}, i, j, g = \alpha \mathcal{M}, i, j, g \quad (\langle i, j \rangle)$$

$\emptyset$ を論理式即ち $\emptyset \in \text{ME}_t$ とすると

$$\emptyset = 1 \quad \text{at } \mathcal{M}, i, j \quad \stackrel{\text{fD}}{=} \quad \forall g \quad \emptyset \mathcal{M}, i, j, g = 1$$

さてここで公理と deduction rule を定めることにより tensed Intentional Logic を公理系として設定しなければならない。

1.L. の Meaningful Expression のうち modally closed な class MC を次

のように定義する。

F. MC の定義

- f 1 type a の変数 $x_a \in MC$
- f 2 $A_a \in ME_a \Rightarrow \wedge A_a \in MC$
- f 3 $A \langle a, b \rangle, B_a \in MC \Rightarrow A(B) \in MC$
- f 4 $A_a, B_a \in MC \Rightarrow A_a = B_a \in MC$
- f 5 $A_b \in MC \Rightarrow \lambda u_a A_b \in MC$

このように modally closed とは situation 乃至は context に依存して変化する式のことである。

$$\begin{aligned} \text{また} \quad T &\stackrel{fD}{=} \lambda x_t x_t = \lambda x_t x_t \\ F &\stackrel{fD}{=} \lambda x_t x_t = T \end{aligned}$$

とする。

このとき I.L. の公理は

G. I.L. の公理

- g 1 $g \Vdash T \wedge g \Vdash F \leftrightarrow \forall x_t g x$
- g 2 $x_a = y_a \rightarrow f_{at} x \leftrightarrow f_{at} y$
- g 3 $\forall x_a [f_{ab} x = g_{ab} x] \leftrightarrow [f = g]$
- g 4 x が type a の変数で $A \beta(x)$ なる式が「 $\lambda y C$ 乃至は $\wedge C$ を部分式としてもてば C が x を自由変数としてもつことはない。」という条件をみたし、更に B が type a の modally closed な式とすると

$$[\lambda x A](B) = A(B)$$

- g 5 $\Box [\vee f_{sa} = \vee g_{sa}] \leftrightarrow [f = g]$
- g 6 $[f_{sa} = g_{sa}] \rightarrow H[\vee f_{sa} = \vee g_{sa}]$
- g 7 $[f_{sa} = g_{sa}] \rightarrow W[\vee f_{sa} = \vee g_{sa}]$
- g 8 $\vee \wedge A = A$

さらに Deduction Rule は

$$h. A \leftrightarrow A', B \vdash B'$$

ここで B' は B に於いて A が現われている (但し λA 以外) ところではそれを全て A' におきかえたものとする。

structure $\langle A, I, J \rangle$ が I.L. の model であるとは

$$i. \exists \mathcal{A} = \langle A, I, J \leq, F \rangle \quad \forall \alpha \in M_{\mathcal{A}} t$$

$$\vdash \alpha \Rightarrow \forall i, j, g \quad \alpha \mathcal{A}_{i, j, g} = 1$$

が満たされることである。

以上で一応 I.L. の公理系と model 解釈の形式を述べた。

Intentional Logicでは通常の一階論理と異なって定数 c に対し M_e の元ではなく $M \langle s, e \rangle$ の元を割当ててゐる。その理由を以下に考察してみよう。

自然言語に於ける「morning star」は名詞句であり、従って一階論理では $D_{A, I, J}$ の元を assign されてしかるべきであろう。ところで

「morning star」=「金星」

「evening star」=「金星」

であるから

「morning star」=「evening star」

が成立。

ところで一階論理では

$$u = v \Rightarrow \varphi(u) \leftrightarrow \varphi(v)$$

が成立する。

φ として

$$\varphi(x) = \text{「} x \text{ は朝でる」}$$

をとると

「morning star は朝でる」

は正しい。

従って

「evening star は朝でる」

も正しくなければならない。これは明らかに我々の直感に合わない。

このことから我々は

「morning star」=「朝でる金星」

=「金星」at 夜あけの世界

= $\begin{cases} \text{金星} & \text{世界が夜あけのとき} \\ \text{無定義} & \text{その他の世界} \end{cases}$

と考えるのが妥当であることがわかる。

これが Intentional Logic で定数に $D\langle s, e \rangle A, I, J$ の元を assign する理由である。

次に

「former Johnes colleague」

という文を考えてみよう。

どの possible world $\langle i, j \rangle$ でも「Johnes colleague」は複数人いると考えられる。従ってそれは

$D_{e,A,I,J}$ の元のある set 即ち $D\langle e, t \rangle A, I, J$ の元

(「Johnes colleague」 $\langle i, j, x \rangle = 1$

$\Rightarrow x \in$ 「Johnes colleague」 $\langle i, j \rangle$ と見る)

「Johnes colleague」 $\langle i, j \rangle$ を指定する。

即ち

「Johnes colleague」 $\in D\langle s, \langle e, t \rangle \rangle A, I, J$

ところで「former」を名詞句を名詞句に移すもので

「former」 $\in D\langle \langle e, t \rangle, \langle e, t \rangle \rangle$

とし

(「former Johnes colleague」) $\langle i, j \rangle$

$=$ 「former」(「Johnes colleague」 $\langle i, j \rangle$)

と解釈することにする。

$\langle i_1, j_1 \rangle$ で Johnes が U.S. Senator

$\langle i_1, j_2 \rangle$ で Johnes が Lawyer

かつ $j_2 < j_1$ 即ち Johnes は Senator になる前は Lawyer だった

とするとき

「Johnes colleague」($\langle i_1, j_1 \rangle$) $\subset$ 「U.S. Senator」

「Johnes colleague」($\langle i_1, j_2 \rangle$) $\subset$ 「Lawyer」

であるから former は U.S. Senator の subset を Lawyer の subset へ移す関数となる。こういう解釈は former という語の意味の中から見い出すことの出来ないものであろう。例えば, Smith が他の職業から U.S. Senator に

なったとすると former は別の関数解釈を受けることになって矛盾する。

$$\alpha \in D\langle s, \langle e, t \rangle \rangle$$

に対し

$$(\text{「former」}(\alpha))(\langle i, j \rangle) = \bigvee_{j' < j} \alpha(i, j')$$

と定義する。

すると

$$\text{「former」}(\alpha)(a)$$

$$\Leftrightarrow \exists j' < j \quad \alpha(\langle i, j' \rangle, a)$$

$\alpha = \text{「Johnes Colleagues」}$ とすると

$$\Leftrightarrow \exists j' < j \quad a \text{ は } \langle i, j' \rangle \text{ で Johnes の colleague}$$

$$\Leftrightarrow a \text{ はかつて Johnes の colleague であった人}$$

となり

$$\text{「former Johnes colleague」}$$

の意味にあう。

またこのとき

$$\text{「former」} = \lambda \alpha \langle s, \langle e, t \rangle \rangle \lambda \langle i, j \rangle \lambda a_e \bigvee_{j' < j} \alpha(\langle i, j' \rangle, a_e)$$

であるから

$$\text{「former」} \in D\langle \langle s, \langle e, t \rangle \rangle \langle s, \langle e, t \rangle \rangle, A, I, J \rangle$$

このように自然言語の format を Logic の中に解釈していくためには、高階の intentional な object を Logic の meaningful expression に assign できねばならない。

Intentional Logic で初めてそのような試みが組織的に行なわれたわけである。

4.2 Montague Grammar

Montague の「On the nature of certain philosophical entities」に従って Montague の semantics に対する考えを述べよう。

いま

$$\text{“ The sun rose at eight ”}$$

という文を考える。

この文は " The sun rises at t "

+ " A certain event occured at eight "

という2つの文に分解可能である。

従って

event = " The sun rises at t "

が成立している。

従って①event とは " The sun rises at t " といった文そのものである。

という考え方が生まれる。

しかし人類が atom や particle という概念をもっていなかったときでも atom
や particle について様々な event が起きていたわけだから

反1 表現不可能な event が存在する。

という反証をあげて①を reject できる。

次に

② event とは $\{ t \mid \text{"The sun rises at t"} \text{ が実際におこった} \}$

といった set である。

と仮定してみよう。

すると

$\{ t \mid \text{"The sun rises at t"} \text{ が実際におこった} \}$
= $\{ t \mid \text{"Johne gets up at t"} \text{ が実際におこった} \}$

ということがありうるが②により " The sun rises at t " という文の表わす
event と " Johne gets up at t " という文の表わす event が一致してしまう。

即ち

反2 同時におこる異なる event を識別できない。

から②も reject される。

そこで

③ event とは $\lambda t \text{"The sun rises at t"}$ で表現されるような時刻 t に
関する或る property であり、時刻 t はあらゆる possible な時刻に
わたって変化し得るものとする。

としよう。

すると t を Johne が病気で入院したという仮定の世界の時刻にとれば

" John_e gets up at t " = 0

" The sun rises at t " = 1

が成立して2つの event を識別できる。

ここで event を meaning と見なすと命題 P で表わされる declarative sentence の意味は I.L. の式 $\wedge P$ であると思なすのがよいという結論が得られる。

以上から自然言語を intentional Logic に翻訳し、その意味を intentional Logic の semantics を用いて表明するという Montague の立場が納得できよう。

4.2.1 Proper treatment of quantification

本節では Montague の P.T.Q. の紹介をすることにより、実際の英語の文に対する Montague 流の文法と意味解釈システムを与えることにしたい。

Montague は P.T.Q. に於いて基本的には category 文法を採用した。

A. Category index Cat の定義

a 1 $e, t \in \text{Cat}$ ここでは e は entity, t は truth value を表わす。

a 2 $A, B \in \text{Cat} \Rightarrow A/B, A//B \in \text{Cat}$

A/B の意味は A/B なるカテゴリーに属する expression に対し、 B なるカテゴリーに属する expression を代入すると A なるカテゴリーに属する expression が得られるというものである。

例えば former は形容詞でタイプは $\langle s, \langle e, t \rangle \rangle, \langle s, \langle e, t \rangle \rangle$ であり、普通名詞句は $\langle s, \langle e, t \rangle \rangle$ のタイプと見ることができたが、

$$\langle s, \langle e, t \rangle \rangle = \text{CN}$$

$$\langle \langle s, \langle e, t \rangle \rangle, \langle s, \langle e, t \rangle \rangle \rangle = \text{Adj}$$

としてカテゴリー文法の記法では

$$\text{Adj} = \text{CN}/\text{CN}$$

とかくわけである。

/と//の違いは重要ではないが、基本的には semantical には同一の働きをするが syntactical には違う働きをするものと解釈する。

カテゴリー文法の枠内に従来の文法カテゴリー即ち品詞を位置付けてみよう。

I V 即ち intransitive verb は t/e に対応する。

これは e 即ち entity に対し t 即ち文を与えるものという意味である。

T : term - t/IV

TV : transitive verb - IV/T

IAV : IV modifying adverb - IV/IV

更に

t/t - sentence modifying adverb

IAV/T - IAV making proposition

IV/t - sentence を object にとる verb phrase

IV/IV - IV から verb phrase verb を作り出すもの

これらのカテゴリーに属する基本句としては

BIV = { run, walk, talk, rise, change }

BT = { John, Mary, Bill, ninety, he₀, he₁, he₂ …… }

BTV = { find, lose, eat, love, date, be, seek, conceive }

BIAV = { rapidly, slowly, voluntalily, allegedly }

BCN = { man, woman, park, fish, pen, unicorn, temperature }

Bt/t = { necessarily }

BIAV/T = { in, about }

BIV/t = { believe that, assert that }

BIV/IV = { try to, wish to }

BA = $\emptyset$ for other A

等が考えられる。

これらの基本句を組み合わせて新たな句を作り出す syntactic rule はカテゴリー A の集合を P_A として次のように定義される。

<基本規則>

S 1 $BA \subset P_A$

S 2 $\zeta \in P_{CN} \Rightarrow F_0(\zeta), F_1(\zeta), F_2(\zeta) \in P_T$

但し $F_0(\zeta) = \text{every } \zeta$

$F_1(\zeta) = \text{the } \zeta$

$F_2(\zeta) = \begin{cases} a \zeta & \zeta \text{ が子音から始まる時} \\ an \zeta & \zeta \text{ が母音から始まる時} \end{cases}$

S 3 $\zeta \in \text{PCN}$ かつ $\emptyset \in \text{Pt} \Rightarrow F_{3,n}(\zeta, \emptyset) \in \text{PCN}$

但し $F_{3,n}(\zeta, \emptyset) = \zeta$ such that $\emptyset'$ で $\emptyset'$ は $\emptyset$ に he_n ,

him_n が現われたときそれを ζ に現われる最初の BCN の性

が $\left\{ \begin{array}{c} \text{男} \\ \text{女} \\ \text{中} \end{array} \right\}$ 性に従ってそれぞれ $\left\{ \begin{array}{c} \text{he} \\ \text{she} \\ \text{it} \end{array} \right\}, \left\{ \begin{array}{c} \text{him} \\ \text{her} \\ \text{it} \end{array} \right\}$ に直したもの

<関数適用規則>

S 4 $\alpha \in \text{Pt}/\text{IV}, \delta \in \text{PIV} \Rightarrow F_4(\alpha, \delta) \in \text{Pt}$

但し $F_4(\alpha, \delta) = \alpha \delta'$ で δ' は δ の最初の verb 即ち (BIV,

BTV, BIV/t, BIV/IV) を三人称単数現在で置き換えたもの

S 5 $\delta \in \text{PIV}/\text{T}, \beta \in \text{PT} \Rightarrow F_5(\delta, \beta) \in \text{PIV}$

但し $F_5(\delta, \beta) = \begin{cases} \delta \beta & \beta \neq \text{he}_n \text{ のとき} \\ \delta \text{him}_n & \beta = \text{he}_n \text{ のとき} \end{cases}$

これは目的語に他動詞を適用すると自動詞が得られるという規則である。

S 6 $\delta \in \text{PIAV}/\text{T}, \beta \in \text{PT} \Rightarrow F_6(\delta, \beta) \in \text{PIAV}$

前置詞 + 目的語 = 副詞に対応

S 7 $\delta \in \text{PIV}/\text{t}, \beta \in \text{Pt} \Rightarrow F_8(\delta, \beta) \in \text{PIV}$

但し $F_8(\delta, \beta) = \delta \beta$

[IV that] + 文 = IV に対応

S 8 $\delta \in \text{PIV}/\text{IV}, \beta \in \text{PIV} \Rightarrow F_8(\delta, \beta) \in \text{PIV}$

副詞 + IV = IV に対応

S 9 $\delta \in \text{Pt}/\text{t}$ かつ $\beta \in \text{Pt} \Rightarrow F_8(\delta, \beta) \in \text{Pt}$

文副詞 + 文 = 文

S10 $\delta \in \text{PIV}/\text{IV}, \beta \in \text{PIV} \Rightarrow F_7(\delta, \beta) \in \text{PIV}$

IV + 副詞 = IV

但し $F_7(\delta, \beta) = \beta \delta$

<文結合規則>

S11 $\emptyset, \psi \in \text{Pt} \Rightarrow F_8(\emptyset, \psi), F_8(\psi, \emptyset) \in \text{Pt}$

但し $F_8(\emptyset, \psi) = \emptyset \text{ and } \psi$

$F_8(\psi, \emptyset) = \emptyset \text{ or } \psi$

$$S12 \quad r, \delta \in PIV \Rightarrow F_8(r, \delta), F_9(r, \delta) \in PIV$$

$$S13 \quad \alpha, \beta \in PT \Rightarrow F_9(\alpha, \beta) \in PT$$

< 量化規則 >

$$S14 \quad \alpha \in PT, \emptyset \in Pt \Rightarrow F_{10,n}(\alpha, \emptyset) \in Pt$$

$$\text{但し } F_{10,n}(\alpha, \emptyset) = \left[\begin{array}{l} \emptyset' \quad \alpha \neq \text{hek のとき } \emptyset' \text{ を } \alpha \text{ の最初の BCN} \\ \text{or BT の性に従って } \emptyset \text{ の } \text{he}_n \text{ を} \\ \left\{ \begin{array}{l} \text{he} \\ \text{she} \\ \text{it} \end{array} \right\}, \text{him}_n \text{ を } \left\{ \begin{array}{l} \text{him} \\ \text{her} \\ \text{it} \end{array} \right\} \text{ でおきかえ} \\ \text{たものとして} \\ \emptyset' \quad \alpha = \text{hek のとき } \emptyset' \text{ を } \emptyset \text{ の } \text{he}_n \text{ を} \\ \text{hek, him}_n \text{ を him}_n \text{ におきかえたものとして} \end{array} \right.$$

これは名詞の代入規則

$$S15 \quad \alpha \in PT, \zeta \in PCN \Rightarrow F_{10,n}(\alpha, \zeta) \in PCN$$

$$S16 \quad \alpha \in PT, \delta \in PIV \Rightarrow F_{10,n}(\alpha, \delta) \in PIV$$

< 時制と肯定, 否定 >

$$S17 \quad \alpha \in PT, \delta \in PIV \Rightarrow F_{11}(\alpha, \delta), F_{12}(\alpha, \delta), F_{13}(\alpha, \delta), F_{14}(\alpha, \delta), \\ F_{15}(\alpha, \delta) \in Pt$$

但し

$$F_{11}(\alpha, \delta) = \alpha \delta' \text{ で } \delta' \text{ は } \delta \text{ の最初の verb を三人称単数現在の否定形に}$$

$$F_{12}(\alpha, \delta) = \alpha \delta'' \text{ で } \delta'' \text{ は } \delta \text{ の最初の verb を三人称単数未来形に}$$

$$F_{13}(\alpha, \delta) = \alpha \delta''' \text{ で } \delta''' \text{ は } \delta \text{ の最初の verb を三人称単数未来否定形に}$$

$$F_{14}(\alpha, \delta) = \alpha \delta'''' \text{ で } \delta'''' \text{ は } \delta \text{ の最初の verb を三人称単数現在完了形に}$$

$$F_{15}(\alpha, \delta) = \alpha \delta'''' \text{ で } \delta'''' \text{ は } \delta \text{ の最初の verb を三人称単数現在完了否定形に}$$

英語の meaningful expression として $UPA_{A \in Cat}$ の要素をとろう。例えば

every man loves a woman such that she loves him.

という平叙文を考えると

これは

$\beta = \text{he}_0 \text{ loves a woman such that she loves him}_0. \epsilon P_t$

に $\alpha = \text{every man } \epsilon P_t$

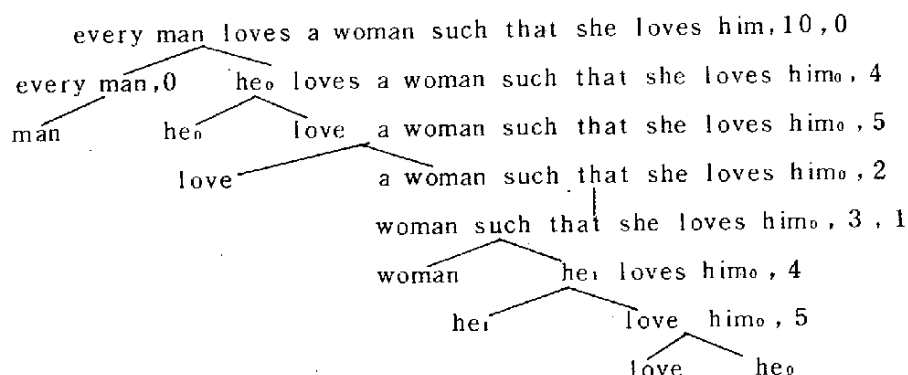
を S14 に従って代入した文と考えられる。この文はしたがって $F_{10,0}(\alpha, \beta)$

であるから、これを

every man loves a woman such that she loves him_{10,0}

every man he₀ loves a woman such that she loves him₀.

と表示することになると以下



のような解析図が得られる。

英文に対し、syntax が与えられたので、その meaningful expression を intentional logic の meaningful expression に翻訳する方法を syntax に従って帰納的に定義しよう。

翻訳規則

<基本規則>

- T 1 a) $\text{be} \Rightarrow \lambda P \lambda u P \{ \lambda v [\vee u = \vee v] \}$
 b) $\text{necessarily} \Rightarrow \lambda p [\Box \vee P]$
 c) $\text{John, Mary, Bill, ninety} \Rightarrow$
 $\lambda PP \{ \wedge_j \}, \lambda PP \{ \wedge_m \}, \lambda PP \{ \wedge_b \}, \lambda PP \{ \wedge_u \}$
 d) $\text{he}_n \Rightarrow \wedge \lambda PP \{ x_n \}$

- T 2 $\zeta \in \text{PCN}$ かつ $\zeta \xRightarrow{*} \zeta'$ のとき
 $\text{every } \zeta \Rightarrow \lambda P \forall x [\zeta'(x) \Rightarrow P(x)]$
 $\text{the } \zeta \Rightarrow \lambda P \exists y [\forall x [\zeta'(x) \equiv x = y] \wedge P(y)]$
 $F_2(\zeta) \Rightarrow \lambda P \forall x [\zeta'(x) \wedge P(x)]$
- T 3 $\zeta \in \text{PCN}, \emptyset \in \text{Pt}, \zeta \xRightarrow{*} \zeta', \emptyset \xRightarrow{*} \emptyset'$ のとき
 $F_3, n(\zeta, \emptyset) \Rightarrow \lambda x_n [\zeta'(x_n) \wedge \emptyset']$

<関数適用規則>

- T 4 $\delta \in \text{PIV}/\text{IV}, \beta \in \text{PIV}, \delta \xRightarrow{*} \delta', \beta \xRightarrow{*} \beta'$ のとき
 $F_4(\delta, \beta) \Rightarrow \delta'(\wedge \beta')$
- T 5 $\delta \in \text{PIV}/\text{T}, \beta \in \text{PT}, \delta \xRightarrow{*} \delta', \beta \xRightarrow{*} \beta'$ のとき
 $F_5(\delta, \beta) \Rightarrow \delta'(\wedge \beta')$
- T 6 $\delta \in \text{PIAV}/\text{T}, \beta \in \text{PT}, \delta \xRightarrow{*} \delta', \beta \xRightarrow{*} \beta'$ のとき
 $F_5(\delta, \beta) \Rightarrow \delta'(\wedge \beta')$
- T 7 $\delta \in \text{PIV}/\text{t}, \beta \in \text{Pt}, \delta \xRightarrow{*} \delta', \beta \xRightarrow{*} \beta'$ のとき
 $F_6(\delta, \beta) \Rightarrow \delta'(\wedge \beta')$
- T 8 $\delta \in \text{PIV}/\text{IV}, \beta \in \text{PIV}, \delta \xRightarrow{*} \delta', \beta \xRightarrow{*} \beta'$ のとき
 $F_6(\delta, \beta) \Rightarrow \delta'(\wedge \beta')$
- T 9 $\delta \in \text{Pt}/\text{t}, \beta \in \text{Pt}, \delta \xRightarrow{*} \delta', \beta \xRightarrow{*} \beta'$ のとき
 $F_6(\delta, \beta) \Rightarrow \delta'(\wedge \beta')$
- T10 $\delta \in \text{PIV}/\text{IV}, \beta \in \text{PIV}, \delta \xRightarrow{*} \delta', \beta \xRightarrow{*} \beta'$ のとき
 $F_7(\delta, \beta) \Rightarrow \delta'(\wedge \beta')$

<文結合規則>

- T11 $\emptyset, \phi \in \text{Pt}, \emptyset \xRightarrow{*} \emptyset', \phi \xRightarrow{*} \phi'$ のとき
 $\emptyset \text{ and } \phi \Rightarrow \emptyset' \wedge \phi'$
 $\emptyset \text{ or } \phi \Rightarrow \emptyset' \vee \phi'$
- T12 $r, \delta \in \text{PIV}, r \xRightarrow{*} r', \delta \xRightarrow{*} \delta'$ のとき
 $r \text{ and } \delta \Rightarrow \lambda x [r'(x) \wedge \delta'(x)]$
 $r \text{ or } \delta \Rightarrow \lambda x [r'(x) \vee \delta'(x)]$
- T13 $\alpha, \beta \in \text{PT}, \alpha \xRightarrow{*} \alpha', \beta \xRightarrow{*} \beta'$ のとき
 $\alpha \text{ or } \beta \Rightarrow \lambda P [\alpha'(P) \vee \beta'(P)]$

< 量化規則 >

T14 $\alpha \in P_T, \emptyset \in P_t, \alpha \stackrel{*}{\Rightarrow} \alpha', \emptyset \stackrel{*}{\Rightarrow} \emptyset'$ のとき

$$F_{10}, n(\alpha, \emptyset) \Rightarrow \alpha' (\wedge \lambda x_n \emptyset')$$

T15 $\alpha \in P_T, \zeta \in PCN, \alpha \stackrel{*}{\Rightarrow} \alpha', \zeta \stackrel{*}{\Rightarrow} \zeta'$ のとき

$$F_{10}, n(\alpha, \zeta) \Rightarrow \lambda y \alpha' (\wedge \lambda x_n [\zeta'(y)])$$

T16 $\alpha \in P_T, \delta \in P_{IV}, \alpha \stackrel{*}{\Rightarrow} \alpha', \delta \stackrel{*}{\Rightarrow} \delta'$ のとき

$$F_{10}, n(\alpha, \delta) \Rightarrow \lambda y \alpha' (\wedge \lambda x_n [\delta'(y)])$$

< 時制と肯定 >

T17 $\alpha \in P_T, \delta \in P_{IV}, \alpha \stackrel{*}{\Rightarrow} \alpha', \delta \stackrel{*}{\Rightarrow} \delta'$ のとき

$$F_{11}(\alpha, \delta) \Rightarrow \neg \alpha' (\wedge \delta')$$

$$F_{12}(\alpha, \delta) \Rightarrow W \alpha' (\wedge \delta')$$

$$F_{13}(\alpha, \delta) \Rightarrow \neg W \alpha' (\wedge \delta')$$

$$F_{14}(\alpha, \delta) \Rightarrow H \alpha' (\wedge \delta')$$

$$F_{15}(\alpha, \delta) \Rightarrow \neg H \alpha' (\wedge \delta')$$

T₄ ~ T₁₁ は文の意味をその構成要素の意味から構築しようという Frege 原理の表明になっている。

いま

$$\text{necessarily } P(n)$$

という文を考えよう。この文の翻訳結果は

$$\Box P(x)$$

であって欲しい。改定された Frege 原理では

$$[\text{necessarily}'] (\wedge P(x))$$

を変形した結果が $\Box P(x)$ であって欲しいことになるから

$$\text{necessarily}' = \lambda P \Box \vee P$$

これが T1b である。

次に

$$\text{he } n P$$

という文を考えよう。he n を変数 x_n と考えると

$$\wedge P(x_n)$$

がこの翻訳結果であって欲しい。

このことから

$$(\text{he}_n P)' = [\text{he}'_n](\wedge P)$$

が $\wedge P(x_n)$ とならねばならない。

故に
$$\text{he}'_n = \lambda P \wedge \forall P(x_n) = \lambda PP(x_n)$$

これが T_1, d である。

John P

という文を考える。

$$\begin{aligned} & [\text{John}'](\wedge(\text{he}_n P)') \\ &= \wedge P(\wedge_j) \end{aligned}$$

がその望ましい形である。というのも John には type e の entity j が対応するとして、それは文中では $\wedge_j$ の形、即ち situation に依存するものと考えるべきだからである。すると

$$[\text{John}'](\wedge P(x_n)) = \wedge P(\wedge_j)$$

が成立するから

$$\text{John}' = \lambda P P(\wedge_j)$$

が成立しなければならない。これが T_1, c である。

次に

Jackleen is Jacky

という文を考えてみよう。

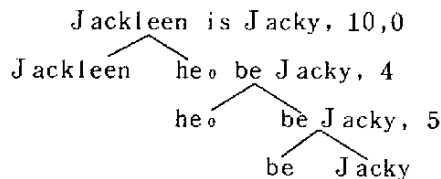
$$\text{Jackleen} \leftrightarrow j_1 \quad \text{Jacky} \leftrightarrow j_2$$

として

欲しい結果は

$$\wedge[j_1 = j_2]$$

である。



なる analysis tree に従って翻訳すると

$$\begin{aligned}
& [Jackleen'] (\wedge_{\lambda x_0} [he_0 \text{ be } Jacky']') \\
&= [Jackleen'] (\wedge_{\lambda x_0} \wedge [be \text{ Jacky}']' (x_0)) \\
&= [Jackleen'] (\wedge_{\lambda x_0} \wedge (([be]') (\wedge Jacky')') (x_0)) \\
Jackleen' &= \lambda PP (\wedge_{j_1}) \text{ より} \\
&= \wedge \wedge (([be]') (\wedge Jacky')') (\wedge_{j_1}) \\
&= \wedge (([be]') (\wedge Jacky')') (\wedge_{j_1}) \\
Jacky &= \lambda PP (\wedge_{j_2}) \text{ より} \\
&= \wedge (([be]') (\wedge \lambda PP (\wedge_{j_2}))) (\wedge_{j_1})
\end{aligned}$$

従って

$$\wedge (([be]') (\wedge \lambda PP (\wedge_{j_2}))) (\wedge_{j_1}) = \wedge [j_1 = j_2]$$

が得られる。

j_1, j_2 を free parameter と考えると $[be]'$ は 2 つの argument をもつ。
またその結果は $P(x, y)$ という形をもっているから

$$[be]' = \lambda P \lambda u P(\alpha(u))$$

なる形と仮定して代入してみると

$$\begin{aligned}
& \wedge ((\lambda u ((\wedge \lambda PP (\wedge_{j_2})) (\alpha(u)))) (\wedge_{j_1})) = \wedge [j_1 = j_2] \\
& \wedge ((\lambda PP (\wedge_{j_2})) (\alpha(\wedge_{j_1}))) = \wedge [j_1 = j_2] \\
& [\alpha(\wedge_{j_1})] (\wedge_{j_2}) = [j_1 = j_2]
\end{aligned}$$

そこで $\alpha = \lambda u \lambda v [\vee u = \vee v]$

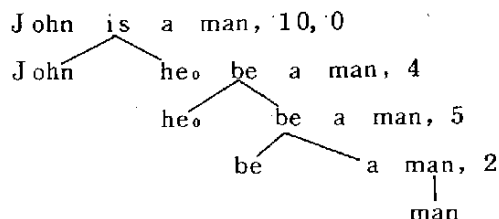
とすればよいことがわかる。

以上から

$$[be]' = \lambda P \lambda u P(\lambda v [\vee u = \vee v])$$

が得られる。これが Tb である。

T₂ 等についても方針は同様であろう。ここでは単に検証してみるにとどめる。



$$\begin{aligned}
& (\text{John is a man})' \\
&= (\text{John})' (\wedge_{\lambda x_0} (\text{he}_0 \text{ be a man})') \\
&= (\lambda PP (\wedge_j)) (\wedge_{\lambda x_0} (\text{he}_0 \text{ be a man})') \\
&= (\wedge_{\lambda x_0} (\text{he}_0 \text{ be a man})') (\wedge_j) \\
&= (\wedge_{\lambda x_0} ((\text{he}_0)' (\wedge \text{be a man})') (\wedge_j)) \\
&= (\wedge_{\lambda x_0} ((\wedge_{\lambda PP} (x_0)) (\wedge \text{be a man})') (\wedge_j)) \\
&= (\wedge_{\lambda x_0} \wedge (\text{be a man}) (x_0)) (\wedge_j) \\
&= \wedge (\text{be a man}) (\wedge_j) \\
&= \wedge ([\text{be}]' (\wedge \text{a man})) (\wedge_j) \\
&= \wedge ([\lambda P \lambda u P (\lambda v [\vee u = \vee v])] (\wedge \text{a man})') (\wedge_j) \\
&= \wedge ([\lambda u ([\wedge \text{a man}]' (\lambda v [\vee u = \vee v]))] (\wedge_j)) \\
&= \wedge ([\lambda u ([\wedge_{\lambda P \exists x [\text{man}'(x) \wedge P(x)]] (\lambda v [\vee u = \vee v]))] (\wedge_j)) \\
&= \wedge ([\lambda u ([\wedge \exists x [\text{man}'(x) \wedge [\vee x = \vee u]]]) (\wedge_j)) \\
&= \wedge \exists x [\text{man}'(x) \wedge [\vee x = j]] \\
&\equiv \wedge \text{man}' (\wedge_j)
\end{aligned}$$

参 考 文 献

Modal Logicについては

Schütte, K., "Vollständige System modaler und intuitionistischer
Logic" Springer, Berlin 1970

を参考とした。

Intentional Logicについては

Montague, R., "The Proper Treatment of Quantification in Ordinary
English" In J. Hintikka, J. Moravcsik, and P. Suppes
(Eds.) Approaches to Natural Language Dordrecht Reidel
1973 pp 221~242

に従ったが公理系については

Gallin, I. "Intentional Logic" North Holland

から採用した。

Montague Grammarについては上記 P.T.Q. をほとんど引写したが、Montague
の文の意味に対する考え方を知らるために

Montague, R., "On the nature of Certain Philosophical Entities" Monist
53: 159-94 (1969)

を参考とした。

Modal Logicに semanticsを導入することに最初に成功したのは

Kripke, S.A., "A Completeness Theorem in Modal Logic", Journal of
Symbolic Logic 24 (1959) 1-14

である。

Modal Logicの入門書としては

Hughes, G. & Cresswell, M.J., "An Introduction to Modal Logic"
Methuen and Co. Ltd. (1968)

Kripke Model を詳細に論じたものとして

Bowen, K.A., "Model Theory for Modal Logic" D. Reidel (1979)

がある。

Montague 文法入門としては

Dowty, D.R., "A Guide to Montague's P.T.Q." Reproduced by
Indiana University Linguistic Club (1978)

が良い。

今後の予定

今後の予定としては

Montague, R., "Universal Grammar" Theoria 36 373-98 (1970)
に従って一般文法に対する Montague の枠組を知る。

Universal Grammar の枠組に従って生成意味論, 解釈意味論と Montague
文法の間相互変換を論じた

Cooper, R., and Parsens, T., "Montague Grammar, Generative Semantics
and Interpretive Semantics" Montague Grammar Ed
Partee, B.H., pp 311~363 (1976)

について述べる。

以後は各論となるが,

Guenther, F., "Formal Semantics and Pragmatics for Natural
Languages" D.Reidel (1979)

Guenther, F., "Studies in Formal Semantics" North Holland (1978)

Partee, B.H., "Montague Grammar" Academic Press (1976)

から文献を選ぶ予定。

—— 禁 無 断 転 載 ——

昭和 5 5 年 3 月 発行

発行所 財団法人 日本情報処理開発協会

東京都港区芝公園 3 - 5 - 8

機 械 振 興 会 館 内

TEL(434)8211 (代表)

印刷所 株式会社 タ ケ ミ 印 刷

住所 東京都千代田区神田司町 2 - 16

TEL (254) 5840 (代表)

