



## 言語情報処理の過去・現在・将来\*

長 尾 真\*\*

### 1. はじめに

現代は歴史的にみて大きな変革の時代であるといえるのではないだろうか。宇宙科学のすばらしい進歩はもちろんのこと、技術的にみても、計算機の誕生、それによる情報処理、特に言語情報の記憶や管理、処理、検索、さらには出版・印刷、一般事務への応用などの技術の進歩は、グーテンベルクの印刷術の発明以来これまでの時代を新しく根本的に変えるものとして位置づけることができるだろう。

情報の発生、伝達、消費の量と速度、これをささえなる技術を、100年前、20年前と現代とを比較してみればこれは明らかである。この変革はつい20年あまり前に始まり、現在我々はちょうどそのまっただ中にいること、この変革の将来の完成へむかって急速な進歩をとげていることを十分に認識し、これから的情報処理技術のあるべき姿をはっきりと見きわめる努力をする必要があるのではないだろうか。

言語情報処理を広くとらえると、これをささえなるハードウェア、処理システム、大量の文書情報のドキュメンテーションと情報検索、さらに言語そのものがどこまで計算機によって解析可能かという問題がある。これらの諸問題についての過去・現在を概観し、将来への展望をしてみたい。

自然言語はなぜ重要なか、こんなわかりきった質問をあらためて発する必要はないと考えられるかもしれないが、情報処理システムに自然言語が自由に入出力され、伝送・加工され、蓄積・検索される場面を考えると、その重要性がひときわ明確となってくる。自然言語はあいまいで、冗長度が大きく、工学の対象にならないと考えられてきたが、一方では、だれでもが使っているものであり、あいまいであると言いながら十分な情報伝達能力をもち、ときには何よりも詳細で微妙な情報表現能力をもっているという事実、さらに極端

な言い方をすれば、社会は言葉（法律、規則など）の上にきずかれたものであること、社会活動のほとんどが場面が言葉によるものであることなどから、これを機械によって便利にすることができるれば、社会に対して大きな影響を与えるし、またその機械化が現在ほど要求されている時代はかつてなかったと言うことができるだろう。このような観点から言語情報処理の諸問題を検討する必要があるのである。

### 2. 言語情報処理のためのハードウェア・ソフトウェアの進歩

言語の計算機による翻訳が盛んに試みられた1960年前後と現在とを比べてみると、計算機はその間に飛躍的な進歩をとげてきた。演算速度はもちろんのこと補助記憶装置の容量も非常に大きくなり、大きな辞書を一冊そっくりそのまま計算機に入れて活用するぐらいのことは簡単にできるようになってきた。

さらに最近数年間はワードプロセッサと称して、簡単な文字入出力、文字処理ができるシステムの研究が着実に進みつつある。英語におけるテキストエディタとしては、各地で種々のすぐれた機能のものが何年もの経験をつみ重ねつつ作られてきており、実にすばらしいものが多い。テキストの入力、追加、変更、削除はもちろん、出力の形式をととのえたり、文章の任意の単位での差し替え、テキスト全体を通じての用語の変更など多くのことが可能である。最近は任意の大きさの任意の字体を指定することができ、テキスト中への図面の挿入個所の指定、またその場所へ図面そのものをタブレットを用いて直接入力ができるといったシステムができている。要するに我々が紙と鉛筆でやっていたこと、活字による製版の作業の双方を同時にしかも自由に計算機で実行できるシステムを実現しつつあることである。

日本語についても種々の研究が精力的に行われてきている。日本語文字の入出力システムの発達も著しく、値段も比較的安くなってきたし、日本語文章のテキス

\* 情報処理学会第18回全国大会招待講演（昭和52年10月4日）

\*\* 京都大学工学部

トエディタの開発も、徐々に行われるようになってきた。日本語についての本格的なシステムとしては、いくつかの新聞社でこの2,3年前より行われるようになってきた新聞記事の収集から編集、割り付け、紙面の製版までの計算機による一貫システムがある<sup>1)</sup>。これは非常にすぐれたもので、今後さらに完成されたものに発展してゆくことを期待したい。このようなシステムがさらに改良され、コストも安くなければ、出版の世界にも本格的にとり入れられ、ほんとうにグーテンベルグ以来の印刷術の変革となるだろう。こういった将来を考え、言語情報処理に必要な機能は何かを明確にし、このような機能がどこでも日常使用できるようになることを目ざして、しっかりしたハードウェア、ソフトウェアが積極的に研究開発されるべきである。

日本語の場合、最大の問題はやはり漢字である<sup>2)</sup>。日常的には2,000字程度、住所・氏名など少し特殊なものを入れて4,000~5,000文字が必要であろう。しかし学術用語など非常に特殊なもの、あらゆる特殊分野のことを考えると、8,000文字あっても不足である。こういったいわゆる盤外字も適切に取り扱えるあまり値段の高くない漢字端末装置を設計することがまず必要となる。この場合最も重要なことは、漢字コードの標準化であり、種々の異体字までを考慮して2万字程度までの漢字に対して体系的なコードを与えることが必要となる。その際字体の指定、ルビの指定などが自由にでき、さらに英和辞典などのように、日本語と英語、フランス語など各國語の共存も可能となるよう制御コードの指定が漢字コードと矛盾なく行われるよう配慮することが重要である。

仮名から漢字への自動変換システムも長年の研究によってようやく良いものが出てきた<sup>2),3)</sup>。しかし日本語においては、漢字と仮名とはどうしても1対1対応とはならず、双方を並行して記しておく必要がある。特に人名、地名などでは仮名漢字変換の不可能なものが多い。

電子計算機で日本語が扱われにくい理由の第一は、漢字入出力が簡単にできないことであるが、漢字を扱うことのできるソフトウェアがほとんどまだ開発されていないことも進歩をさまたげる大きな理由である。英語アルファベットの場合は1バイトを単位とするソフトウェアであるが、漢字の場合には2バイトを1つの単位とするソフトウェアを作る必要がある。テキストエディタ、KWICシステム、情報検索システム、デ

ータベースシステム、その他数多くのソフトウェアをすべて2バイト単位で動くようにし、ユーザーに広く使えるようにすることはコンピュータメーカーの緊急の仕事である。

要するに、我々が日常紙と鉛筆を用いている場面からそれらをとりのぞいて、すべて機械化し、それがまたかも我々のノートブックであるかのように使うことのできるシステムを目指した研究を行うことが必要であり、これがオフィスオートメーションの中心課題の1つであることができる。現在IBMやXeroxなどではこれを重要テーマとして研究開発を積極的に進めていると聞く。むずかしいが重要な仕事であり、社会に対するインパクトも大きく、これから発展が楽しみな分野である。

### 3. 大量言語情報の管理と検索のシステム

大量の言語情報を記憶し、検索が能率よくできる管理システムはまだ未完成であると言わざるをえない。比較的形式のととのった文献情報の場合でも、抄録つきの不定長データで、特に日本語の場合には、ほとんどすべてこれからシステムを開発しなければならないといった状況である。

大量情報に対しては、まず分類ということを考えることになる。しかし從来に行われてきた10進分類法で処理できる時代は過ぎた。FIDなどでは100進分類法など種々の新しい分類法を研究しているが、現代は情報があまりにも多すぎ、分類だけではまったく不十分となり、個々の資料に直接アクセスできる方法が要求されている。現時点で大量の言語情報に対して効率的にアクセスしようすれば、どうしても種々の特徴的な用語をたよりにしてアクセスする必要があるがある分野において用語として何を選ぶのがよいかは非常にむずかしい問題である<sup>4),5)</sup>。用語は概念を代表し、用語相互間の関係を明確にすることはよりもなおさずその分野を体系的に明確化することと同じである。したがって完成された学問分野はよいが、現在発展途上にある分野においては、用語の選定は非常にむずかしい。しかしそのような変化の激しい分野においてこそ新しい情報に対する需要が多いのである。用語の選定はその分野の専門家が行うよりほかにないが、境界領域や新しい分野などでは用語をつねに更新する必要が生じ、用語のもつ概念も徐々に変化する。したがって情報システムをup-to-dateに保つためには大変な努力を必要とするのである。

したがってできればこのように用語の選定をあまり人手によらず計算機になるべく客観的にさせることができることなる<sup>6), 7)</sup>。さいわい日本語資料においては重要な概念を表す用語はほとんどの場合、漢字か片仮名によって書かれているという日本語の特殊性を活用することができる。漢字の連続した部分と片仮名の部分を区別して取り出すことは計算機を用いて容易にできるので、こうして取り出された漢字連続の中から重要な役目をする語を選びだすことができる。たとえば大量の文献資料をいくつかの分野にわけ、いくつかの特定の分野にはよく現れるが、その他の分野にはほとんど現れない語を抽出することによって重要語を決定する方法はかなり良い結果を与えることがわかつてきた<sup>8)</sup>。

ただ日本語の場合、特に科学技術の専門分野においては非常に長い漢字の連続が現れる。したがってこのような長い漢字の連続を適切に分割し、みじかい漢字の単語に分離することが必要である。我々は接頭語、接尾語として使われる漢字、さらに一般の国語辞典の見出し語などを計算機に記憶し、これを活用するとともに、2~3文字で1つの単語を構成することが多いことを考えて、長い漢字連続を単語単位に自動的に分割するプログラムを作り、93%程度の正しい分割結果をえている<sup>9)</sup>。それ以上の精度を得るために漢字のもつ意味を考えねばならないだろう。科学技術情報センターでは学術用語のKWIC表を作り、漢字連続の共通部分、異なる部分の現れ方を比較することによって漢字連続のどの部分を専門用語として取り上げるのがよいかを研究している。

漢字連続を単語に分割する問題は以上のような方法である程度解決できるが、片仮名单語の場合は多くの困難な問題がある。外国語の片仮名による表記法が一定していはず、発音をうつすとき、どう聞きとっているかの問題がある。またいくつかの連続した外国語単語の仮名表記のとき「・」を単語の間に挿入したりしなかったりするし、長音を表す「ー」を用いたり用いなかったりするなどがある。計算機で用語を規定するとき困難な問題が多い。

このように、同じ単語で異なった表現をする場合、同じ事柄をちがった単語で表現する場合などでは、同義語、類義語、反義語、関連語の関係などを明らかにすることがぜひとも必要である<sup>4), 5)</sup>。このような情報を体系的に整理した辞書を、シソーラスと呼んでいるが、このシソーラスをどれだけ精密に作ることができ

るか、情報検索のときにどのように活用することができるかが、情報システム全体、特に情報検索システムにとって最も重要なことである。オンライン対話型検索のときにシソーラスをうまく活用したシステムを作る必要があるだろう。シソーラス作成、重要語のきめ方におけるクラスタリングの手法は興味のある、今後まだまだ研究すべきテーマである<sup>10), 11)</sup>。

このように計算機での使用法を十分に考えた辞書・シソーラスを作ることが必要であるが、これは大変な仕事となる。一方世の中には古くから国語辞典や英和辞典などの辞書があり日常使用されている。したがってこのような辞書を計算機でも活用できないかを考えることも重要な問題である。我々は、既に Webster の英語辞典、新明解国語辞典を計算機で利用できるようになっている<sup>9)</sup>。特に国語辞典の場合には、読み方で引くほかに漢字でも引けるし、漢字仮名まじりやその他の種々の変形でも辞書引きができるようなシステムを作っており、漢字連続の分割、日本語文の単語単位への分割と品詞づけなどに活用している。しかし辞書の見出し語に対応する内容記述の部分を、どのように活用するか、どのように活用できるかは大きな問題である。この部分ができるだけ詳細に記述されていて、計算機でも利用できるような辞書の作成が望まれる。

情報検索のためのキーとして、各文献にキーワードをつけねばならないが、このいわゆるインデクシングを人手によらず、自動的に行おうという自動インデックスの研究もかなり進んできている。大量情報を扱うときは同じ質のインデクシングをすることができる、多数の人間が手で行うときの精粗のばらつきという問題をのがれることができる<sup>12)~16)</sup>。

情報検索システムは米国では全国的規模でかなり多くの大規模なシステムが動いている<sup>17), 18)</sup>。英語の場合はデータの作成、遠隔通信回線による対話的検索およびその端末など、いずれもあまり問題なく広く活用されるようになってきているが<sup>19)</sup>、日本語の本格的な情報検索システムは、まだほとんどないという状態である。

内容検索、質問応答などのシステムは現在研究段階にある<sup>20), 21)</sup>。しかし、情報システムへのアプローチとしては自然言語またはそれに類似の形式で行うのが最もよく、そのためにシステムによって程度の差はあるが自然言語文の解析が何らかの形で行われる必要がある。情報システムの使用法を知らない人をスムーズに誘導できること、また知っている人の場合でも、検

素、問い合わせをする時は最初から最適な質問をすることはまずありえず、これは本質的に試行錯誤的プロセスであり、対話的プロセスであることを十分に認識する必要がある。したがって対話とは何か、自由な対話をしながら自然にそれを1つの目標にむかって収斂させてゆく方法は何かについて研究する必要がある。これは現在、まだほとんど研究されていない分野である。

自動抄録の研究もこれから行わねばならない重要な研究分野である<sup>22)</sup>。意味に関係せず機械的に情報を圧縮する方法が試みられているが、文章の理解が計算機で満足にできないかぎり、抄録は現時点ではむりだといわざるをえないだろう。

#### 4. 自然言語の解析

自然言語の解析がむずかしい理由は何かを問うてみることは大切なことである。まず第1に1つの事柄を表現するための表現の可能性があまりにも多すぎるということがある。さらにそれらの表現にあいまいさが非常に多いことである。第1の問題は単語の選択だけでなく、文の構造の選択の可能性が非常に多いことからくる。これは表出したいことの人の頭脳の中における構造は何か、といふやる深層構造の問題と、それからの文章の発生過程の問題とであり、これらはいずれもほとんど解明されていない問題である。文のもつあいまいさとしては、構造的なものと意味的なものとがあるが、いずれにせよ別に外部からは付加情報を与えねばあいまいさを解消することができない。そこでどのような情報を与えれば言語のもつあいまいさを解消できるかが問題となる。文法情報、意味情報、さらには文脈情報や社会的知識などまでを活用しなければならない場合も多いが、それらの情報の具体的な内容、計算機内での蓄積の構造、その活用の仕方などについては今後の研究にまたねばならない。

文の構造解析というとき、たとえば主語、動詞、目的語などを明らかにするだけでなく、動作主格、経験者格、助格、目的格、起点格、目標格、位置格などをたてる、いわゆる格文法の考え方による解析が重要となる<sup>23)</sup>。しかしこの解析を行うためには単語のもつている意味と単語相互間の関係を明らかにすることが必要である。すなわち意味の問題が必然的に出てくるのである。そこで辞書が必要であるが、現在存在している通常の辞書はこういった立場からのものではないのでほとんど役にたたず、あらたに計算機による言語解析

に使いやすい情報をもった単語辞書を作ることが必要となる。これはぼう大な仕事であり、あらゆる場合について矛盾なく、くわしい記述をもった辞書を作ることは容易ではない。九州大学ではこの問題を10年以上こつこつとやってきて、かなりの程度まで到達しているといえよう<sup>24)、25)</sup>。

さて、言語の自動解析の研究は1960年前後、最も盛んに行われたが、1964年に至り米国ではALPACレポートと称する調査報告書が出され<sup>26)</sup>、機械翻訳はあらゆる観点からして、人間の行う翻訳にはおよばない。したがってそれまで数年にわたって出されて来た研究費はこれ以上出すのはむだであるという結論が示された。その結果言語の機械翻訳の研究は急速におとろえ新しい模索の時代にはいったといえる。

その後言語に対する新しいアプローチは人工知能研究の分野から出てきた。人工知能研究の分野においては既に1963年ごろから計算機と自然言語によって対話をする試みがいくつか研究され、BASEBALL<sup>27)</sup>、SIR<sup>28)</sup>などのプログラムが作られていた。SIRはいわゆる semantic network と称する知識の表現形式の最初のものであるということができよう。これらの研究の延長線上において、1971年、当時MITにいたT. Winogradは計算機と人間との間で非常に柔軟な対話のできるシステムを作った<sup>29)</sup>。彼はこの研究において文を解釈するためには知識が必要であること、知識は手続き的(procedural)にプログラムの中に表現することが便利なことを示した。またW. Woodsはデータベースに対して自然言語で質問を出し、單なる表引きの結果のみならず、表中の数値の総和、平均値などの計算結果をも要求に対する答として出すことのできるシステムを作った。このプログラムにおいて、WoodsはATN(Augmented Transition Network)という、有限オートマトンを重複的、回帰的に用いることのできる言語解析のモデルを提案し、文脈自由型構造解析よりも強力で見通しのよい文の解析システムを作った<sup>30)、31)</sup>。

我々はこのATNをさらに発展させ、構造解析の制御をより柔軟に強力にするとともに、日本語のように語順が自由にかえられるような言語に対してもうまく適用できるプログラミング言語PLATONを作った<sup>32)</sup>。そしてこれを用いて日本語文章の解析システムを作った。このプログラムの特徴は1つの文中の指示詞が何を指しているか、省略されている句は何かなどを、それまでに発話された文の情報を使って推定する能力を

もっている。すなわち知識としての long term memory の外に、発話された文脈を short term memory としても、そこに演繹・推論の機能を導入することが文の解析に必要であることを明確に示した<sup>33), 34)</sup>。1970 年以後のこれらいくつかの言語解析のシステムはいずれも質問応答システムとして取り上げられている。これは、文の解析を十分満足のゆくようにすることは、文の理解とは何であるかを明確にすることであるという立場にたち、それは自然言語による質問に対して適切な答を計算機が与えることができる、人間と計算機との間で自然な対話ができる、そのためのプロセスは何かをプログラムで明確に実現することができるとの観点に立っているからである<sup>35)</sup>。

さて、機械翻訳は現時点でのどのような状況にあるだろうか。1964 年の ALPAC レポート以後、米国における研究が衰退したことは先に述べたが、フランス、ドイツなどではその後も着実な研究が行われてきており、最近はカナダにおいても英仏翻訳が研究され天気予報という限られた領域において実際のシステムが動いている<sup>36)</sup>。これは英語による天気予報の文章をフランス語に翻訳して放送するものであり、pre-edit, post-edit のいずれもしない完全なものである。ただ、計算機による文の解析が不成功におわった場合のみ人間が介入し、フランス語訳を与える。最近ようやく米国においても翻訳の研究を再び始めようという機運がもり上りつつある。たとえばソールトレーク市におけるモルモン教会の仕事、XONICS, SYSTRAN システムなどいくつかのものが現れている<sup>37)</sup>。1964 年当時とくらべて言語解析の手法、プログラム技法が格段に進歩したこと、当時とくらべて計算機の処理能力、記憶能力が飛躍的に向上していることなどから、現時点で人力を投入し努力すれば、特定の分野の論文（たとえば分野を化学、計算機科学、プログラム・マニュアルなどに限った論文）はかなりの程度翻訳ができるものと考えられる。

## 5. 言語情報処理の将来

1976 年 3 月米国バージニア州ロスリンで FBIS Seminar on Machine Translation が開かれた<sup>37)</sup>。米国、カナダ、ヨーロッパからこの分野の研究者約 100 名が集り、現在の機械翻訳研究の展望を行い将来の可能性について討論を行った。1960 年初頭、機械翻訳に対する可能性を大きくうたいあげ、過大な期待を一般に与えた末、ALPAC レポートによりそれが否定されたこ

とがあつて、このセミナーの結論は非常に謙虚なものであった。すなわち機械翻訳を実用的に行うためにはもっと基礎研究をつみ重ねる必要があるというものである。しかしフランス、ドイツ、カナダなどはかなりしっかりした研究を押しすすめており、限られた分野で限られた用途に対しては決して機械翻訳が遠い先のことではないとの印象を与えており、上記の慎重な結論は“あつものにこりてなますをふく”の感なしとするのである。

しかし、研究は着実に進めねばならない。しかも機械翻訳以外にも自然言語処理技術で実際の社会において必要とされているものは多い。こういった広い立場に立って考えると次のような手順で問題を解決していくのが妥当であるだろう。

1. 文字入出力、編集、文章処理のシステムの実用化
2. 辞書を計算機にのせ使いこなすこと
3. 言語処理・解釈のための強力なソフトウェアの開発
4. 上記 1~3 を能率よく実行することのできるハードウェアの開発
5. 言語解釈、そのための日本語・外国語の比較对照研究

人間の知識の最高の表現手段である言語を機械で処理するためには、とうぜん機械にぼう大な知識を付与することができねばならない。そういう意味からも知識の凝集された形としての辞書をいかに作るかは大問題である。特定のせまい限られた分野については計算機に適した辞書を作ることもできるだろうが、分野を制限せず任意の文章を対象とすると考えると、あまりに詳細な知識の精密な記述を期待することはできない。また計算機にどこまで深く文章を理解させる必要があるかにも限度がある。一方人間は多くの場合書かれた本や辞書などから知識を得、ものごとの解釈の助けにしていることを考えると、計算機においても現在の人間のために作られた辞書や一般の本の中にもりこまれている情報をどのようにしたらうまく利用できるかということが問題となってくる。

多くの場合、辞書や本に書かれている情報は、計算機からみるとかなり局所的、断片的である。人間はこのような情報であっても、それらが使われる環境を頭に入れて適切な判断のもとに活用する能力をもっている。したがってこれから計算機は人間が記録して持っているぼう大な知識を、たといそれらが断片的、局

所的なものであっても、うまく取り入れて、計算機で活用できる形の知識の体系に組み立ててゆき、これを用いて与えられた問題（文の解析、理解と応答）をとくことのできるプロセスを作りあげる研究を行う必要があるだろう。

Stanford 大学の Feigenbaum らは、人間のもつ知識をうまく組織化し、計算機上に実現する方法を研究することが必要であるとし、知識工学（Knowledge Engineering）という名前の学問分野を提唱している<sup>38)</sup>。こういった考え方を言語処理、言語翻訳における知識の問題、翻訳プロセスに導入することを検討することも必要となってくるだろう。

言語はコミュニケーションの手段として最適のものであることは言うまでもない。従来、コミュニケーションにおける言語は確率的にしかとらえられてきていない。これを現代の言語理論、言語処理の技術の上にたってもう一度考えなおす必要があるだろう。コミュニケーションは送信者と受信者双方が共通の知識をもってはじめて成り立つものである。シャノンの場合、この知識は言語の一般的な確率的構造のみであった。しかし計算機が発達し、言語の構造的知識、社会常識などを記憶し、演繹・推論をすることができるようになると、このような環境モデル下における冗長でないコミュニケーションとは何か、効率のよい信頼度の高いコミュニケーションとは何かを、質問応答システムとの対比で研究することが重要なテーマとなってくる。これは今後積極的にとりくむべき興味あるテーマである。またこのような人間と機械との対話をを通じて、機械がかしこくなつてゆくプロセスも研究しなければならない。従来よりもはるかに高度な構造的なレベルでの学習機構を考え、刺激と反応のシステムを計算機と人間との間に作ることによって、機械が知識を増強しそれをうまく活用してゆくプロセスを考えるべき時代にきたといえよう。いずれにしても情報科学は一番むずかしい“ことば”の時代、頭脳活動の解明と計算機による実現化の時代に突入したということができる。

## 参考文献

- 1) 日本語情報処理の技術動向調査報告書、日本情報処理開発センター、昭和48年3月
- 2) 天野、河田、武田、カナ漢字変換機能を備えたワードプロセッサー、1977 信学会部門別会国大会、1977.8
- 3) 牧野、勝部、木沢、カナ漢字変換の一方法、情報処理、Vol. 18, No. 7 (1977)
- 4) シソーラスの現状、講座シソーラス 1回～5回、情報管理、Vol. 20, No. 1～No. 5 (1977)
- 5) 単言語シソーラスの決定と発展のための指針その1、その2、-ISO 2788、1974 年度版について、情報管理、Vol. 19, No. 5, No. 6 (1976)
- 6) G. Salton, C. S. Yang, C. T. Yu: Contribution to the Theory of Indexing, Proc. of IFIP '74 (1974)
- 7) G. Salton, A. Wong, On the Role of Words and Phrases in Automatic Text Analysis, Computers and the Humanities, Vol. 10, pp. 69～87
- 8) 長尾、水谷、池田、日本語文献における重要語の自動抽出、情報処理、Vol. 17, No. 2 (1976)
- 9) 長尾、辻井、山上、建部、言語情報処理のためのサポートシステム(II)、信学会オートマトンと言語研究会 AL 77-25, 1977. 7. ほぼ同様の内容は信学会部門別全国大会(1977)、漢字かな混り文処理のための辞書の構成、辞書を用いた複次結合語の自動分割、にある。
- 10) C. C. Gotlieb, S. Kumar, Semantic Clustering of Index Terms, JACM, Vol. 15, No. 4 (1968)
- 11) N. Sager, Computational Linguistics: Steps Towards Application, Workshop in Linguistics and Information Science, FID/LD, Stockholm (1976)
- 12) P. H. Klingbiel, Machine-Aided Indexing, Defence Documentation Center, Report DDC-TR-71-3 (1972)
- 13) C. R. Jacobs, Machine-Aided Indexing, Defence Documentation Center, Report DDC-TR-72-4 (1972)
- 14) 中井浩、機械補助索引(MAI)について(I), (II) 情報管理、Vol. 19, No. 4, No. 5 (1976)
- 15) C. T. Yu, G. Salton, Precision Weighting: An Effective Automatic Indexing, JACM, Vol. 23, No. 1 (1976)
- 16) M. E. Stevens, Automatic Indexing, A State-of-the-Art Report, NBS Monograph, 91 (1970)
- 17) 染谷、米国におけるオンラインサービスの現状と将来、情報管理、Vol. 17, No. 12 (1975)
- 18) R. K. Summit, O. Firschein, Document Retrieval Systems and Techniques, in Annual Review of Information Science and Technology, Vol. 9 (1974)
- 19) 福島、染谷、オンライン文献情報システムの会話機能について—各種システム会話言語の比較検討—、情報管理、Vol. 17, No. 8 (1974)
- 20) R. F. Simmons, Natural Language Question-Answering Systems: 1969, CACM, Vol. 13, No. 1 (1970)
- 21) 長尾、自然言語によるマン・マシン・システム、計測と制御、Vol. 13, No. 1 (1974)
- 22) 田中章夫、自動抄録処理におけるキーワードの

- 性質, 電子計算機による国語研究 V (1973)
- 23) C. J. Fillmore, The Case for Case, in E. Bach, R. Harms (Ed.) *Universals in Linguistic Theory*, Holt, Rinehart and Winston, N. Y. (1968)
- 24) 栗原, 自然言語の機械処理, 情報処理, Vol. 14, No. 4 (1973)
- 25) 藤田, 鶴丸, 吉田, 日本語の機械処理——日本語文の標準形分解——, 信学論 D, No. 7(1975)
- 26) National Academy of Sciences, National Research Council, *Language and Machines: Computers in Translation and Linguistics*, Washington D. C. (1966)
- 27) B. F. Green Jr. et al., *Baseball: An Automatic Question-Answerer*, Proc. WJCC, Vol. 19 (1961)
- 28) B. Raphael, SIR: A Computer Program for Semantic Information Retrieval, in M. Minsky (Ed.) *Semantic Information Processing*, MIT Press (1968)
- 29) T. Winograd: *Understanding Natural Language*, Academic Press. (1972)
- 30) W. A. Woods, R. M. Kaplan, The Lunar Sciences Natural Language Information System, BBN Report No. 2265, Bolt Beranek and Newman, Cambridge Mass. (1971)
- 31) W. A. Woods, Transition Network Grammars for Natural Language Analysis, CACM, Vol. 13, No. 10 (1970)
- 32) 長尾, 辻井, 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)
- 33) 長尾, 辻井, 田中, 意味および文脈情報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)
- 34) 長尾, 辻井, 田中, 意味および文脈情報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)
- 35) 長尾, 辻井, 自然言語処理プログラム, 情報処理, Vol. 18, No. 1 (1977)
- 36) J. Chandioux, METEO: an operational system for the translation of public weather forecasts, in 37)
- 37) D. G. Hays, J. Mathias, FBIS Seminar on Machine Translation, AJCL microfiche 46 (1976)
- 38) E. A. Feigenbaum, *The Art of Artificial Intelligence: Themes and case studies of knowledge engineering*, Proc. 5th IJCAI (Aug. 1977)