

座談会



機械翻訳システム開発における諸問題†

出席者 (五十音順)

田中 穂積¹⁾, 野村 浩郷²⁾, 溝口 文雄³⁾吉田 将⁴⁾, 司会 長尾 真⁵⁾

翻訳とその自動化の必要性

司会 (長尾) それじゃ、機械翻訳システムの開発における諸問題というテーマで、座談会をやりたいと思います。機械翻訳システムは、最近随分いろいろなところで、話題になってきましたし、すでにアメリカなどで、質はよくないが幾つかの実用システムが出てきています。一方では、ヨーロッパ共同体とか、そのほかの国においてはこういったシステムでは満足できず、言語の障壁、ランゲージバリアというのがひしひしと感じられていて、言葉を翻訳するという問題は、非常に重要な問題になってきているわけですね。おそらく、日本語なんかもっとどんどん翻訳して日本の事情をくわしく海外に知らせれば、通商摩擦なんかでも、あるいは少しはましになるかもしれないわけですね。

野村 ヨーロッパでは、機械翻訳というのは、重要視されているわけですが、むしろ日本のほうが、そのような要求は非常に強いんじゃないかと思うん

ですね。日本というのは、いわゆる単一言語民族で、特にまた外国語が苦手なようですから、機械翻訳というのは、実は非常に大切なことですね。

ですからこれはほんとうは、国をあげて大きく取り組んでいくべき問題だ、というふうに考えていますけれども。

吉田 ほんとうは、非常に大事なことです。外からの要求も強く、企業とか、そういうところになりますと、これはまた必要性に迫られていると思いますけれども、個人レベルでは、どうもことの重要性を感じていない人が多いんじゃないかなという気がするんですが。

田中 これからは日本という狭い殻に閉じ込もっていたんじゃないけなくて、いろんな情報を外から吸収したり、こちらからもいろんな情報を外に流さなくちゃいけない時代になっていくと思います。そういう意味で、夢物語としてでなく、機械翻訳はそのための格好のテーマじゃないかと思えますけれども。

溝口 確かに個人のレベルだと、それほど必要性はないと考えている反面、1回海外へ出ると、初めて言語の重要性を実感しますね。それが最近になって、やっと貿易摩擦とか何かで、初めて認識されてきて、外国との相互理解のためにも、機械翻訳は重要だとい

†日時 昭和57年3月24日 17:00~20:00

場所 神田 YMCA

1) 電子技術総合研究所 2) 日電公社武蔵野通研
3) 東京理科大学 4) 九州大学 5) 京都大学

感じがするんですね。スタンフォード大学の人達も、とにかく通産のレポートが欲しいとよくいますよ。向こうからは論文などをすぐ、ただで送ってくるわけですね。こちらは取りっ放して、こちらから送るというのは、日本語で書いたものか、あるいはめんどうくさいから、英語のアブストラクトぐらいになってしまう。

司会 確かに、日本の研究レベルは非常に高く、ものすごく何かやっているらしいけれども、実態は全然わからないというのが、欧米の連中が非常に歯がゆく思っているところらしいですね。

だから、そういう意味でも日本の文献とか、研究活動の内容がどどん外国語に直って、彼らにわからせることが必要で、そういうことは、やっぱりなかなか日本にいちゃ、わからないんですね。

野村 最近、国内だけを相手にしている大企業は少ないと思いますが、たとえ国内でのみ企業活動を行っている場合でも、情報の交流は国際的になってきていますので、「外国語は関係ない」というわけにはゆかなくなってきましたね。ましてその企業の規模が大きい場合には、国際的に見ても、企業責任とでも申しますか、そのようなものが要求されてきているように思いますので、その意味でも、機械翻訳は非常に重要だと思えます。

田中 翻訳に関していえば、英語とか、フランス語とか、ドイツ語とか、そういうメジャの言葉に翻訳するのは、確かに非常に重要なんですが、アラビア語とか、そういうマイナな言葉も大切で、そういうところにこそ大きな言語のバリアがあるのではないのでしょうか。そのあたりは、もし機械化できれば、大いにやらなければいけないことだと思いますね。

吉田 あるメーカーの方に聞いたんですが、ある1つの計算機の機種をつくりますと、それに関連したマニュアル類を、英語に翻訳するのに、最低、億という単位の金はかかると。しかも、国によっては、もう英語は受け付けられないところもあるようですし、ですから、企業にとっては、これは大変な問題ですね。

司会 それはどうも事実らしいですね。たとえば、開発途上国に、いろいろな物を売るときに、アメリカあたりだと、英語のドキュメントしか持っていないのに、日本からは、その国の言葉のドキュメントを持っていくらしいんですね。そうすると、やっぱりそっちのほうがいいですね、何といっても(笑)。

日本語の特殊性と翻訳の困難さ

司会 そういう点で、機械翻訳というのは、いろいろクローズアップされてこざるを得ないし、技術的にも随分進歩してきたわけですがけれども、どの辺まで行きそうでしょうか。また日本語の特殊性という問題とも絡めて考えるときに、どういうふうになっているのでしょうかね。

田中 機械翻訳そのものを、ほんとうに人間と同じレベルでやるというのは、ちょっと先かと思うんですけども、間に人間が入って、機械翻訳システムと協調しながら、翻訳の文章をつくり出すというのは、意外に5年とか、そのあたりでできるんじゃないかと思えます。確かに日本語の特殊性もあるんですが、一番の問題は、いろいろな言葉に共通な意味の問題を解決する努力をするということで、中期5年ぐらいの目標というのと、10年ぐらいの目標と、2つぐらいに分けて、機械翻訳というのを考えるのがいいのではないかと考えています。



野村 機械翻訳がどこまでゆくかということですが、話題や分野を限定し、最終的には人間がチェックするということを考えると、実用的な観点からは、ずいぶん役に立つものができるのではないかと思います。もちろん、それは完全な自動翻訳とはだいぶ違っていて、人間の翻訳を手助けするようなものですが、しかし、このようなレベルの機械翻訳でも、言語学的な研究を含めて基礎的な研究を非常に多く、かつ深く行わなければならないと思います。

たとえば、これは日本語の特殊性にもなりますが、「の」による接続と申しますか、あるいは「の」による埋め込み文の代用と申しますか、これらは表層の文の形をいかに訳すかということではなく、その文が何をいわんとしているかの意味をくみとって、しかるべき変形を行った後に、相手言語の構造へ変換することが必要だと思います。このような処理ができないと、人間の手助けをする翻訳システムとしても不足ですし、またこのような研究は、まだまだこれからだと思います。

吉田 確かに日本語と英語を較べてみましても、たとえば、冠詞1つにしても、どう付けていいのか私などが考えてもわからないですね。

司会 ですから、機械翻訳というのを、理想形で追

い求めるということは、工学的な立場からは、やっぱり無理なんです。源氏物語あたりを翻訳することと機械翻訳とは、全然違うんだと、そこをはっきりさせないと。機械翻訳ができる、できるというるのはけしからん話だ、そんなことはできるはずがないという議論が、すぐ出てくるわけですね。

吉田 それから、日本語の特殊性ということではもって回ったようないい回しとか、特殊な用語とか、「前向きに対処します」というような、翻訳するのに困るような、日本人の社会だけで通用するような言葉といますか、外国人にもわかってもらおうという努力の全くないものが、非常に多いですね。この態度は話し言葉だけでなく、技術的な文書にまで持ちこまれている。

やはり日本語そのものを、ある程度コントロールする必要があるんじゃないかと、私は思っているんですけど。

田中 日本語の特殊性でいえば、省略が非常に多いということですね。たとえば、日本語を英語に翻訳する場合には、主語を必ず立てないと、英語の文章として成り立たないということがあります。省略が日本語には多いということは、1つの問題ではあると思います。逆にほかの言語から日本語に直す場合には、主語をいつでも訳すような翻訳の文章は、必ずしも翻訳文として、適切でないということがあり得るでしょう。

野村 それに、慣用的な表現が言語によって大きく異なっていますし、さらに、それがかなり個別で、したがって個別的な言語データを多く用意しておかなければならないことになります。しかも、それらが言語ごとに違った構造を持っている場合は対応付けだけでも大変なことだと思います。

司会 この間グルノーブルから帰ってきた辻井君(京大工学部)もいっていたけれども、機械翻訳を考えるとときには、日本語と英語だけでは物事がわからないんで、少なくともフランス語とかドイツ語とかロシア語とか、構造の違う言語がいくつもあるんだから、そういうものも頭の中に入れて、システム作りをしないと、だめなんだということですね。そういう言語学的な諸問題をどういうふうにして工学的に明確化していくかというのはどうやったらいいんでしょう。

田中 そういうしっかりしたシステム作りは、やはり国がはっきりした意識をもって引っぱってゆかない

といけないのではないのでしょうか。

溝口 文部省では、やっぱり英語教育が中心になってしまう。

司会 やっぱり、プラクティカル言語というか、実用言語というものの意識を、しっかり持たないとね。

まあ、それにしても、あまりむずかしいことはいわないで、逐語訳でもいいから、意味が少なくともあまり落ちないで、意味が通じるような文章であれば、機械翻訳の結果としてはよしとする、というような線で、まずはいかないと、何事も事は進まないような気がしますね。

辞書の重要性

野村 一番低いレベルで考えますと、いわゆる辞書引きで、単語を置き換えてくれるだけでも、非常に役に立つ場合があるんですね。

イディオムなんかも含めて、辞書引きしてくれるだけで十分役に立つ場合がある。特に自分の知らない言語で書かれた文献を読まなければならない時など、痛切に感じますね。これはやはり言語処理において辞書が基本的な役割りを果たしていることを示すものだと思いますが。

司会 そうですね。人間のやらなければならんことと、機械のやってくれることをうまく組み合わせたトータルシステムですね。そこでほんとうにエコノミカルなシステムがつくれるのかどうかというのが、問題ですね。そこでの辞書の位置づけですね。専門用語なんか、翻訳者だって知らないんですから、それは絶対に集めておかないとだめですね。

吉田 やっぱり専門用語の問題ですね。非常に新しい用語などに簡単な解説がついているという、ターミノロジー・データバンクですか、そういうものが翻訳者にとっては、まず必要ですね。

田中 メモリはどんどん安くなっているわけで、そういうものが非常に小型で、簡便に使えるようになりますからね。

司会 動詞などでは、機械が使えるような動詞の辞書構造が必要でしょうけれども、専門用語は訳語をきちんとつけるということをやれば、機械にも人間にも、両方の役に立つ用語集ができるでしょう。

田中 ところがそれは、日本できちんとやっているかどうかという(笑)。いまからでも遅くないから、どんどんやっていかないと、いけないですね。

吉田 ある機関で責任をもってやらないといけない

んでしょうね。息の長い仕事ですから。

司会 ヨーロッパなどでは、100万語ぐらい集めて、計算機の中に入れて、定義をちゃんとつけて使っている。

溝口 言語というのは、オープンプロブレムですから、解決つかないような問題の要素を含んでいるわけですね。だからこそ、次のステップにそれまでの成果が生かされていくというようなことを考えなきゃいけない。

田中 そういう意味では、専門用語辞典のようなものは、これからでもいいから、きちんとしたものをつくれば、それは次の世代に財産として残っていくわけですね。

野村 専門用語をまとめるときに、今後は、いわゆる言語を処理するという立場から、用語というものを整理していくということも、必要じゃないかと思うんですね。それによって、未登録の複合的な用語が複合語処理のようにして近似的に処理できることになるでしょうから。

そして、そういうものをまとめることによって、逆に新しい言葉をつくるときに、何らかの指針を与えるというような、そういう効果も出てくると思うんですね。

司会 それは非常に大事なことだと思いますね。

野村 外国では、こういった仕事のための政府関係のしかるべきしっかりした機関があるようですね。

日本でも、いろいろあるとは思いますが、外国ほど明確でないようですね。そういったところが、もう少しはっきりした方針を出してもらえたら、非常にありがたい気がしますね。

司会 JICSTなんか、相当やっておられるんだから、その辺の知識、うんちくを傾けて、いろいろ指導をしてくれると、いいんじゃないかと思うんですね。

田中 かなり大きなプロジェクトになるはずですね。専門分野といっても非常に多いんだから。

野村 専門用語以外の、ふつうの辞書も、もちろん重要ですね。現在までに電総研とか長尾先生の所の研究とかがありますが、いろんな言語の辞書が早く計算機の中に入って欲しい、しかも自由に使えるようになって欲しいですね。さらに解析などに使うための辞書構造も統一的な方向を出していただけるとありがたいですね。また分野や対象に応じた標準的な見出し項目のセットなどができると非常にありがたいですね。

文法と解析

司会 いままで言語学的に、大体機械翻訳のレベルでわかってきたのは、どんなことであって、これからやらなければいかんというのは、どんなことでしょうか。

田中 形態素解析と構文解析は、まあまあわかってきた。1970年代にずいぶん進歩したと思います。

吉田 確かにそうなのですが、やはり複合語の問題とか、接頭、接尾語の問題とか、まだ残っていますね。

田中 複合語とかになると、形態素解析ということもさることながら、意味の問題が絡まってくるような形態素解析になる。

ですから、これからは意味を重要視したものでしょうね。

司会 構文解析まではずいぶん研究され成果があったけれど、翻訳となるとね。たとえば、英語とフランス語、ドイツ語の間だと、どんな表現法で物事を発表していくかという構図は似ているけれども、日本語とヨーロッパ言語だと、思考の形態も違うし、発語していく順序が、全然違うでしょう。

そうしたときに、かなり構造の違う翻訳を考える必要がありますね。そうすると、やっぱり構文レベルでまだまだやらなければいかんこと、たくさんあるんじゃないでしょうか。

野村 構文解析という面では、解析メカニズムは、はっきりしてきていると思うんですが、しかし文法規則という面では、まだ十分な評価がされていないのではないかと思います。かなり大規模な文法規則も作られているようですので、特定の文についてはなく、一般的な文について、それを十分カバーしているか否かを評価し、必要ならば追加・修正などを行って、だれでもが信頼して使える標準的な文法規則でもいえるものが提供されるとありがたいですね。そうすると、それに語の接続などの制限条件として、分野や対象に応じた意味的なものも取りこめるようになるのではないのでしょうか。

吉田 自然言語の問題はどれ1つ取ってみても難しく、昔から問題とされていて、今日でもいぜんとして解決されないでいることだらけのように見えます。機械翻訳にしても、20年前のものと較べて現在も本質的に進歩していないのではないかという意見を耳にすることがあります。しかし、その中でも言語理論の成果



は着実に取り入れられてきており、形態素や構文の解析力が格段に強化されたこと、文の意味処理の機構がかなり明確になってきたことは否定できないと思います。ただ、意味処理には非常に時間がかかるので、形態素解析、構文解析の段階で、あいまいさの落とせるものは、しっかり落として解析し、それから意味を入れるということを考えないと、能率の非常に悪いものになるんじゃないかと思えます。

野村 アルゴリズムあるいはメカニズムという観点から考えると、効率をあげることは非常に大切ですが、あいまいさや多義性はさけられない問題ですので、可能性をさぐるという意味では、メカニズムだけでなく、文法を含んだ能率を中心に検討する必要がありますと思います。そのへんはまだ十分ではないように思いますね。

溝口 意味の問題が大切だということですが、それをどういうふうに取り込むかということと同時に、どの意味がほんとうに意味を持つかというようなメタレベルの問題が、あるんじゃないかなと思います。

司会 だから研究としては、構文解析で明確になる問題、それで何が解けなくて残るか、次に意味の問題で、どういうところまで解けるか、そのほかに文脈の要素は入れなくちゃいけないとか、あるいは知識のことはどうするかとか、研究というレベルで考えれば、やっぱり1つ1つ段階をきちっと分けて、その段階で、どれだけのことが解けてきたか、明確になってきたかということ、確かめながら進めることが必要ですね。

田中 たとえば、take という動詞は、環境によってもすごく意味が変わるわけですね。主に目的語によって、I take a cold. というと、かぜをひく、I take a bus. というと、バスに乗ると訳さなければなりません。この場合はさらに目的語に「に」をつけなければいけない。そういうふうな問題は、意味のことをきちんとやっていけば、ある程度は解決できるということですから、訳語の選択は意味の問題になってくるんですね。

野村 意味だけでもまだ不十分で、次のステップとなると、今度は文脈がものすごく入ってきます。その中で、要するに、何をしゃべろうとしているか、何を伝えようとしているかということが非常に重要な役割を果たしてくるわけですね。

そうすると、言語から一步出ちゃって、いわゆる知識の問題とか、そういうところに広がってしまいますので……。

司会 武蔵野通研ではそこまで含んだ大きな構想で研究しているんじゃないですか。

野村 そうですね。なかなか難しいですがね。

溝口 要するに、日本語の構文解析として、集約すれば幾つぐらいあって、こうすれば大体の構文解析ができるといった目安は。

司会 まあ、立つと思えますね。

ソフトウェア

司会 機械翻訳システムを作ってゆく上のもう1つの大きなファクタはソフトウェアだと思いますね。たとえば、論文表題ならそれに合うような文法は何かとか、あるいは、アブストラクトならアブストラクトに合うような文法は何かとか、あるいはまた、契約文書に合うような文法は何かとか、そういうサブグラマという考え方をいろいろやって、テストしなければわからないというとき、そのツールとなる文法記述システムみたいなもの、これがぜひとも必要ですね。それを、何かうまくつくる方法はないですか。

野村 外国では、まずそれからつくっていくという感じがしますからね。日本でも、どこかいい物をつくってもらいたいなと思いますね。

司会 そうするのは、電総研とか、通研でできないですかね。

野村 私の所では、言語の分析を少しやっています。部分的にしる、まとまるまでには、もう少し時間が必要です。言語学の分野の方々のおちえを借りたいことが山ほどあります。

田中 私どもは、拡張 LINGOL というのをつくって、結構あれでうまく、いろいろなことができると思っていたわけですが、最近 PROLOG があちらこちらで注目されています。PROLOG というのは自然言語処理のための一種のツールとして動くような、そういうおもしろいところがあるんですね。

溝口 田中さんが拡張 LINGOL から PROLOG に移る前に、拡張 LINGOL でつくられている、いろいろな文法がありますね。そういう成果を何かうまく整理するというのが貴重だと思うのですが。

吉田 先ほどのターミノロジ・データバンクみたいに、ある人はこのシステムで、こういうヒューリスティックな情報を考え出したとか、こういう文法をつくったとか、そういうものがデータとして欲しいなと思うときがありますね。自分のシステム用に、研究として使うのにいいし、手元があれば、いつか使おうと思

いますからね。

司会 アメリカなんかではお互いのツールを相互交換して、みんなが使い合いますね。そういうことを通じて、そのツールというのは客観的な物としてでき上がる。そしてその上に何かつくる、それをまたみんなが使い合って、これは確かなものであるというようにところに落ちついていくでしょう。

田中さんがつくられた拡張 LINGOL なんか、随分いろいろなところで使われて、あれはあれとして、いろいろと効用があったんだから、そういうソフトウェアの流通性、あるいは文法も、いろいろな人がつくった文法を、お互いにテストしてみるとか、コメントをいうとか、そういう場が必要ですね。

田中 それは非常に大切ですね。アメリカの場合、それが非常にうまくいっている1つの理由は、ARPA ネットワークを通して主要な研究機関が協力し合いますからね。

司会 作った人だけわかっているんじゃない、ほんとうの意味の知識じゃないですね。いろんな人が試みることでできて、わかるというような構造になっていないと。共通の財産になって、ほんとうの知識なんだから。

田中 そこで非常に問題になるのは、あるプログラムをつくったときに、それをほかの人に使ってもらえるような説明書をつくらないといけない。

司会 日本の研究は、その辺がある種の貧困ですね。

吉田 ATN とか、拡張 LINGOL などは、非常にベーシックなもので、概念としてもわかりやすいし、ですから、よく使われる要素を持っていますが、ちょっと複雑になってくると、やっぱりむずかしいという面があるような気がしますね。

野村 ドキュメントがしっかりしていて、持ってきたものが、すぐ自分のところで動けば、すぐ使ってみると思うんですね。

ところが、やっぱり動くマシンがないとか、あるいは、国産の計算機で LISP が動いていなかったり、仕様が違っていたりして、そういった別の面での障害があるような気がしますね。

溝口 だから、機械翻訳のときの機械というとき、どういう性格をつけるかというのは、やっぱり検討する必要があると思うんですね。

田中 PROLOG がいいか、LISP がいいか、いろいろありますね。

司会 研究という意味じゃ、いろいろな試みがあったらいいと思うんですね。

ただ、そういうふうによったものが、ほかの人も使って検証できるとか、追試ができるとか、あるいはその上にまた、自分の新しいものを積み上げていけるといような研究環境は、やっぱり必要ですね。

今後の開発体制

司会 今後の開発体制という意味じゃ、そういうことを、相当しっかりしたところ、たとえば、国とか、電総研とか、いろいろなところが音頭をとってやっていただかないと、いけないんじゃないですか。

溝口 第5世代コンピュータプロジェクトというのは、そうした方向の基礎づけじゃないかという気がするんですね。いままでの蓄積を整理統合して、どうするかという問題。



要するに、各研究所ごとの問題を離れて、共通の難問は何かという意識に立っていかなければならないんじゃないですか。そうしないと、繰り返しになってしまう。

田中 そういう意味じゃ、開かれた研究所というか、そういう場が機械翻訳にも必要でしょう。京都大学とか九州大学に、機械翻訳の研究所か何かつくって(笑)。

司会 国として、1つの研究所をつくるだけの値打はあると、ほくは思いますね。

吉田 私は、これはどうしても必要なものだと思いますね。機械翻訳研究所というのかどうか知りませんが、情報学という立場から言語の問題を研究するところですね。それこそ、先ほどのターミノロジーデータバンクも含めてですね。

野村 私もほんとうに必要なだと思います。しかもそれが真の意味で開かれた研究所であることですね。さらに外国の研究機関などとの交流が絶対必要ですから、国のバックアップが不可欠だと思います。ヨーロッパでは国の機関として言語局とか翻訳局といったものがあると聞いていますが、言語障壁の大きい日本でこそそういったものを真剣に考える必要があるのではないのでしょうか。

田中 不思議なことに、日本には情報に関する研究所はないですね。ニュートラルな中立機関としての研究所は。

吉田 今後の開発体制ということで考えますと、たとえば、機械翻訳のための辞書を一生懸命つくっている人がいるわけですね。まあ、辞書をつくるとなる

と、相当な年月がかかります。ところが、そういうものは、なかなか論文なりにくいわけですね。論文の評価の仕方もあるべきではないと思います。たとえば、単なる謝辞ではなく、参考文献欄に記載し、論文1つ書いたのと同程度に評価するとか。

野村 また情報交換の場と申しますか、研究発表の場も広げてゆく必要があるでしょうね。情報処理学会の自然言語処理研究会でも発表の件数が増えてきていますし、潜在的なものもかなりあるような気がします。

司会 機械翻訳は、いろんなところで、いろんなことをやってきたから、この辺で一度大いにディスカッションして、ある種の方向性を見つけたり、辞書なんか特にそうでしょうけれども、分担的な協調のことも、そろそろ考えなければいかんでしょう。

野村 最近、ヨーロッパなどから、非公式ですけど共同研究と申しますか、あるいは研究分担と申しますか、そういうさそいがあるという話を耳にしますね。もちろん研究者レベルでの話の段階ですが、我々もそのような希望をもっているわけですから、実現できればすばらしいと思います。

田中 そういうテーマだと思えますね。奥行きが深いし、横幅も広い、若い人の関心が、非常に最近高まっている感じもしますね。

司会 それと、言語学プロバの人を、いかに呼び込んできて、そういう人との交流の下で、こういう研究を進めるかということが大事なんじゃないか。言語学者が簡単に参加して、いろいろなことが書けるといった、そういうソフトウェア体制を我々がつくって、やっぱり言語学者が直接書いて実験することですね。現在だと、そういうツールづくりができるように、大体なってきたわけですよ。これを研究して、ソフトづくりをすれば、ぼくはやれるんじゃないかと思うんですがね。

田中 早い話が、ぼくが言語のミニ解析みたいなことをやっていたんだけれども、ああいうのは、ほんとうにいい道具ができれば、中学生とか高校生でもつくっちゃうんじゃないでしょうか。

司会 それは事実だと思いますね。かなり簡単なカンパセーションとか、そのぐらいのシステムは、つくれるから。

野村 だれでも使えるようなシステムでないと、だめなんですね。計算機のことをよく知っていないと使えないようなソフトウェアシステムは、やっぱり使いづらくて使われないでしょう。したがって、たとえば

言語学者の人が、計算機のことをほとんど知らなくても、単なる道具として、紙と鉛筆の代わりとして使って、文法を変えてみるとか、あるいは辞書をつくってみるとか、そういうことが自由にできるような環境はつくらなければいけないし、いまやろうと思えば、ある程度そういうシステムはつくれるわけですね。外国ではもうある程度あるわけだし。

田中 そういう協力にはもうちょっと時間がかかるという気がしますね。

司会 ヨーロッパなどで議論していると、お前の機械翻訳のプロジェクトに、トランスレータは何人入っているかと必ず聞くんですね。彼らは、言語学者と翻訳者とを区別している。いや、コンピュータサイエンティストだけで、ごりごりやっていますという、じゃ、成功しないんじゃないかと。

言語学者がいても、この仕事はなかなかむずかしい。翻訳者というのを抱き込んで、実践的にやっていくというのを誇らしげにしている。

野村 翻訳専門家の人達は、翻訳についてのノウハウを非常に多く持っているのですから、それが十分生かせるような体制をつくって、それを取り込むことも必要だと思います。そのためにはエキスパートとしての翻訳者を研究するとか、いわゆる人間を対象とした心理学的な研究も必要ではないかと思っています。

溝口 医療診断システムをつくったときの経験では、全部がコンピュータサイエンティストだと、やっぱりうまくゆかない。

波及効果

司会 機械翻訳システムの開発は、いろいろやらなければいけないんですけども、それができたとしたら、波及効果とか、応用とか、そういう分野については、いかがでしょうか。

野村 応用分野は、限りなく広いと私は思うんですね。それは、単に翻訳だけでなく、一般にドキュメント、それから抄録の作成とか、そういったいろいろなものに、翻訳の基礎技術というのは、そのものの形で応用できると思うんですね。

たとえば、文を書くときに、よい文を書く、よい日本語を書くということに対して、非常に影響を及ぼしてくるんじゃないかと考えているんです。

吉田 私も実は、日本語の間の翻訳システムという



のをやりたいと思っています。日本語から、もうちょっといい日本語に翻訳する。これは日本語を解析し、さらにゼネレートするという、両方の段階を取るわけですね。

ですからそれは、実は機械翻訳の問題の1つでもありまして、これから力を入れようと思っているところなんです。技術文書の書き方などの日本語教育システムにも役立つと思っているわけです。

田中 そういう技術的な面での実用的なメリットがありますね。たとえば、マニュアルを翻訳する。最初に、ランゲージバリアのところ、いろいろ問題がありましたけれども、そういうバリアも解消されていく、そういう非常に大きな影響を、今後持つわけですけれども、機械翻訳の研究を通じて、言葉そのものが一体何だろうとか、そういう問題が、明らかになってきやしないか、そういう気がします。

そういう意味で、言語学をやっている人も、参加してほしいし、もっといえば、我々は一体どうして言葉を理解して、翻訳したり、いろいろコミュニケーションしているのかという、そういうふうな方面のもの、すなわち、ある意味で心理プロセスというか、そっちのほうの研究にも、何らかの影響があるんじゃないかと、考えているんですがね。

司会 おそらく、応用分野で一番興味があるのは、対話システムじゃないかと思うんですね。

これは、非常にいろいろなところに応用されるでしょうし、たとえば、オフィスオートメーションとかね。今後の情報システムの中では、オフィスのことだと思っただけですね。

そういったことで、いまおっしゃったような、人間の心理のプロセスまで考えたような研究が必要であると思うんですが、その辺はどうですか。

溝口 知的会話システムの設計には、絶対に機械翻訳の成果が必要です。たとえば、1つは計算機との対話に必要なコマンド言語の開発です。具体的には、人間の言語を計算機の人工言語に翻訳してくれるようなものですね。

もう1つは、ドキュメンテーション分野ですね。それも単に、マニュアルを計算機にストアしておくのではなく、使い方だとかの目的に応じてドキュメント化されている必要がある。こうした利用者の使いやすさを考える分野はエディタの心理学と呼ばれています。

野村 機械翻訳というのは、ちゃんと書かれた文章を、他の言語のちゃんとした形の文章として訳すだけ

ではないと思うんです。たとえば、われわれが、論文とか、解説書を外国語で書きたい場合に、自分の思っていることを、計算機に伝えて、しかも計算機がそれを文章としてアレンジしてくれる、というようなことをしてくれれば、一番いいわけですね。

つまり、われわれが普通何か書こうかなと、最初に思うときには、まだ不完全な形でしか、考え方がまとまっていない。

それを断片的にいえば、ちゃんとした1つのまとまった文章として、つくり上げてくれると非常にありがたいですね。しかも、それがたとえば英語で、あるいはアラビア語でできれば、これは最高にすばらしいことなんで、機械翻訳の解析で、意味を理解する云々ということは、そういうことを十分に含んでいると思うんですね。

もちろん、そのようなことができるようになるのは、もう少し先でしょうが、人間的な作業をサポートするシステムというのは、ある程度作れるわけですね。人間と計算機のインタラクションを幾つかのレベルで考えて、徐々に高度化させてゆくシステムの例としても、機械翻訳は非常に良いテーマであり、このような基礎技術がさきほどのOAの基本となるような気がします。

溝口 野村さんがいったように、計算機を通じての、われわれの知的な能力を増幅してゆくという、いわばシンバイオス（共生）という考え方がありますね。それを通じて計算機を活動的なパートナーにしてゆくという方向が考えられるわけです。

田中 自然言語というのはシンボルを扱うわけで、それに都合のいいような計算機という方向がありますね。

そうすると、第5世代計算機が目ざしているような知識情報処理システムに、機械翻訳なり、自然言語処理の研究からの寄与があり得るんじゃないかと思っています。

司会 いろいろと多面的な議論ありがとうございます。いずれにしても、これからの世界的規模の通信とかコミュニケーションの世界の中で、人間が自由に情報システムとつき合ってゆくためには、言葉の機械処理や機械翻訳がどうしてもはいらねばシステムが成り立たないという時代にきつつあることは確かで、そういったことをできるだけ多くの人達に認識してもらって、機械翻訳の研究が進み、実用システムが1日も早く実現することを期待したいと思います。

出席者紹介(五十音順)

田中 穂積 (正会員)

昭和16年生。昭和39年東京工業大学制御工学科卒業，昭和41年同大学院修士課程修了。同年通産省電気試験所(現電子技術総合研究所)入所。以来，オペレーティングシステムの研究，自然言語理解システムの研究に従事。現在電子技術総合研究所パターン情報部推論機構研究室長，工学博士。

長尾 真 (正会員)

1936年生。1961年京都大学工学部電子工学科修士課程修了。京都大学工学部助手，助教授を経て，1973年京都大学工学部教授，現在に至る。国立民族学博物館併任教授。自然言語の機械処理，機械翻訳，パターン認識，画像処理などの研究を行って来た。

野村 浩郷 (正会員)

昭和19年生。昭和42年大阪大学工学部通信工学科卒業。昭和44年同大学院修士課程修了。同年日本電

信電話公社武蔵野電気通信研究所に入所，昭和47年工学博士。同研究所基礎研究部で一貫して知能処理の研究を行ってきた。特に，しきい値論理代数，学習理論，知識表現と知識習得，自然言語処理，日本語処理，高機能パーソナルコンピュータ，マンマシンインタフェースなどを研究し，現在は，知識ベースと意味理解を中心とする知能的機械翻訳の研究に従事。

溝口 文雄 (正会員)

昭和16年生。昭和41年東京理科大学工業化学科卒，昭和43年同大学院修士課程修了。同年東京理科大学理工学部経営工学科勤務，以来認知科学，知識工学の研究に従事。現在東京理科大学理工学部経営工学科助教授。工学博士，共著「知的情報システムの設計」，共著「知識工学入門」。

吉田 将 (23巻2号参照)