



AAMT

The Asia-Pacific Association for Machine Translation

Journal

No.8

September 1994

アジア太平洋機械翻訳協会

MT-SUMMIT V

第5回機械翻訳国際会議

開催月日 平成7年7月10日～14日

開催場所 ヨーロッパセンター(ルクセンブルグ)

Prog. chair: Prof. Yorick Wilks (Dept. Computer Science, University of Sheffield)

主催 EAMT (European Association for Machine Translation)

共催 AAMT (Asia-Pacific Association for Machine Translation)

AMTA (Association for Machine Translation in the Americas)

IAMT (International Association for Machine Translation)

「TMI'95」

(第5回機械翻訳の理論的、実践的諸問題に関する国際会議)

開催月日 平成7年7月5日～7日

開催場所 University of Leuven, Belgium

論文募集 テーマ-MTとコンピュータ意味論・MTと音声言語・サブランゲージ・制限言語とMT

期限-1995年1月15日(コピー6部提出)

審査完了-1995年4月15日

使用言語-英語

参加料 1995年5月15日まで申し込み- 6,500 BF(ベルギーフラン) バンケット込 8,000 BF

5月16日以降申し込み-10,000 BF(ベルギーフラン) バンケット込 11,500 BF

問合せ先 University of Leuven, Center for Computational Linguistic

Maria-Theresiastraat 21, B-3000, Leuven, BELGIUM

(Tel. +32-16285088 E-Mail tmi95@et.kuleuven.ac.be)

目次

エッセイ	(AAMT 副会長 野々内 隆)	(1)
研究開発の動向(Future Directions of MT)		(2)
機械翻訳パーソナルユースの現状と対応策		(6)
研究機関の紹介(日立 中央研究所)		(13)
技術早分かり(機械翻訳の辞書)		(18)
アニュアル・レポート・ダイジェスト(JIPDC JEIDA-MT・PC)		(14)
新製品-HICATS(22):TransLand(24):ATLAS/Win(26):PC-Transer(28) 翻訳工房(30)		
トピック(翻訳サテライトオフィス・HHD3.1・世界の翻訳図書、世界の特許)		(17)
協会活動報告(第4回通常総会・機械翻訳セミナー)		(32)

日米経済協議で、ワシントンに日本の技術文書を翻訳するための機械センターを設置することが決まったとのニュースをみて、いよいよ機械翻訳の出番がやってきたと云う感慨でした。昔、知人が「これから機械翻訳が盛んになるので、翻訳の前処理と後処理の人材を教育するための学校を開設したい」と言っていたのを思い出しました。此の構想は時期尚早であったようですが、今から必要になるのか、或いは翻訳ソフトが進歩して、その様な人材が必要でなくなるのか、どうでしょうか？

中部太平洋に「ノノウチ島」という島があります。たまたま私と同じ名前なので興味を持って行って見ました。先の大戦で数千名の邦人が玉砕したマキンタラワから南へプロペラ機で一時間ほどのところにあります。人口2000人の珊瑚礁の島で電気／通信／ラジオ／新聞／工場／ホテル／レストランなど所謂文明と云われるものは何も無いと言って良いほどの社会です。英連邦を構成するキリバス共和国の一部の島です。首府タラワからタバイ大統領が私どもと一緒に来てくれました。歓迎会で、大統領がキリバス語で演説をして、それを同国の外務次官が英語に翻訳し、更に日本から同行した通訳がそれを日本語に翻訳すると云う回りくどい方法が採られました。私の挨拶も此の逆のルートをたどったわけです。音声認識の出来る機械翻訳機があれば有り難いと思つて思つたものでした。しかもポータブルであれば最高です。

此のような稀な例は別として、技術文書や小説が、主要な外国語であれば自由に翻訳が出来る時代がこないものでしょうか。たとえば、町の図書館にデジタル化された世界中の主要な文学作品が保有されていて、希望者はこれを機械翻訳機を通して、自由に耳からも目からも鑑賞することが出来たら素晴らしいと思います。目や耳の不自由な方にとっても大変便利なシステムとなるはずです。このシステムを更に通信回線を利用して、遠隔地でも利用できるようにすればもっと便利になります。

最近、海外旅行が増加してきましたので、日常会話程度は音声で表現出来る機械翻訳機が発売されています。さらに、音声入力が出来て、出力も音声で出来る時代が何時か来るのでしょうか。そうすると学校の英語の時間は何を目的にすることとなるのか、また新たな問題を発生することになりそうです。

(日立製作所 専務取締役)

Future Directions of MT

- Language, Meanings and Translation -

Jun-ichi Tsujii (Centre for Computational Linguistics, UMIST)

(1) はじめに

COLING94のチュートリアルが本会議前の8月2日～4日の3日間にわたって東京で開催された。チュートリアルは、解析、生成、プラグマティクス、機械翻訳、コーパス、マルチメディアUIと自然言語処理の要素技術や応用分野での最近のトピックをカバーする6つの講演から成り、それぞれ第一線の研究者により、現状・問題点から最新技術のアルゴリズムの紹介まで幅広くかつ丁寧な解説が行われた。参加者も100名以上で質疑も非常に活発であった。本報告ではUMISTの辻井教授の機械翻訳に関する講演の内容を紹介する。

辻井教授の解説は次の4つから成る。当日の講演時間の関係上配布資料の中で説明されなかった部分に関しても報告者の理解できる範囲で説明する。

- ・概要（機械翻訳システムの現状と問題点）
- ・伝統的パラダイムの紹介
- ・言語の意味と翻訳
- ・統合された枠組み

(2) 機械翻訳システムの現状と問題点

機械翻訳システムの研究開発は80年代から再び活発になった。利用者によっては商用システムをうまく使いこなし効果を得ているケースも多々あるが分野全体でこの10年を振り返れば、当初期待していたほどにはうまくいっていないという見方もある。特に日本での現状を分析してみると、立場の違いによって次のように言うことができる。

<背景>

開発者：膨大な投資 → 小さな市場
 利用者：高い期待 → 低いパフォーマンス
 研究者：大きな夢 → 重大なデッドロック

<失望>

これらの失望を教訓にして、今後の研究開発の方向を見つけなければならない。

<経験から学んだ教訓と今後の方向>

[1] (狭い意味での) 言語学は期待したほどには役立たない。

⇒ 翻訳ということをもっと経験的に研究をすること

[2] 当時考えたほど実際の市場の需要には出会わなかった。

⇒ 種々の形態のシステムの提供

[3] 汎用的な機械翻訳システムなど存在しない。

⇒ カスタマイズ技術

[1] (狭い意味での) 言語学は1つの言語内で閉じた研究であり、翻訳という2つの言語に係わる問題全てに解答を用意できるわけではない。むしろ、人間が行っている翻訳という行為に関する調査研究をもっと深める必要がある。言語学者が体系化した知識の代わりに、翻訳中に人間が経験的に得る知識

(翻訳例を含めた) を利用する。EBMT (事例ベース翻訳)、SBMT (統計ベース翻訳) の2つの方式が現在提案されている。EBMTとは、ある文書を翻訳する時に、それまでの翻訳結果をテンプレートとして再利用する方式で、1) 類似した翻訳事例の検索、2) 翻訳事例の分割、3) 翻訳事例の組み合わせの3つのステップから成る。翻訳結果の大量のデータベースをどれほど高速にアクセスできるか、類似した事例というものをいかにして見つけるかがボトルネックになる。また、部品を組み合わせ訳文を生成する時には、言語知識を利用する必要がある。SBMTは、大量の翻訳結果には「規則」が内在しているものと考え、抽出のために統計的手法を用いる。簡単なモデルでは2言語間の文書の単語の分布から単語間の対応関係を求めるだけだが、文中の位置やフレーズの対応関係を求めるモデルもある。例えば、英語で "starred" という単語は、フランス語で "marquees de un asterisque" (アスタ

リスクでマークされた)というフレーズに対応する。

S BMTの問題点としては、膨大な量のデータベースが必要なこと、抽出した規則が分野依存などの特殊なものなのかと、汎用性のあるものかの区別がつかないこと、などが挙げられる。

[2] 専門家向けと非専門家向けの2種類翻訳支援システムが考えられる。前者はターミノロジーのデータベースや、それまでの翻訳結果を翻訳中に随時参照できるツールが必要になる。後者では後編集不要な対話的翻訳支援ツールが望んでいる。同時に高品質の翻訳結果も望まれている。例えば、ビジネスライター作成システムなどである。

[3] あらゆる分野や文種の文書にでも利用できる汎用的なシステムが望めない以上、カスタマイズしながら翻訳システムを使っていく必要がある。しかしながら、カスタマイズの方法論が確立されているわけではない。利用者が思うような訳を得るにはユーザ辞書・専門辞書の構築・整備など時間とお金がかかる。できれば、ユーザが翻訳したい文書を大量に与えることで対象文書分野の知識が獲得されるのが望ましい。

以上3つの教訓と今後の方向を眺めると、いずれも“corpora”(コーパス)という言葉が浮かび上がってくる。

(3) 伝統的パラダイムの紹介

現在の機械翻訳システムの大部分は従来のパラダイムであるL BMT(言語理論ベース翻訳)に基づいている。基本的なアイデアを以下に紹介する。

翻訳対象となる言語表現の数は無限である。L BMTでは、ある翻訳の単位(translation units)を組み合わせ、全体の表現を構成するという立場(構成主義; compositionality)から、演繹的アプローチ(Reductionistic Approach)でこの無限性(infiniteness)に対応しようとする。

1) まず表現(expression)をより抽象化した形式(form)に変換する。この時、入力言語側の言語理論で体系化された有限の規則を用いる。関係代名詞(日本語では連体修飾)が現れる文では規則が再帰的に適用されることもある。

2) 次に、この抽象化された形式と等価な(equival

ent)出力言語側の形式に変換する。抽象化の階層レベルにより種々の対応が考えられる。

以下の例に挙げるように、この変換(形式の対応)はそれほど簡単なわけでない。

3) 最後に、出力言語側の言語理論に従って訳文を生成する。

例1) 引数が変わる場合

EN: John misses Mary.

FR: Mary manque a' John.

例2) カテゴリが変わる場合

EN: postwar {adj}

DE: nach dem krieg {pp}

例3) イディオム

EN: cut off smoothly

DE: glattschneiden

EN: give a cough

FR: tousser

例4) ヘッドスイッチ

EN: he HAPPENS to be ill

JPN: kare-ha TAMATAMA byokidearu.

例5) 構造の組み合わせが変わる場合

EN: John gave a WEAK cough.

FR: John toussa FAILBLEMENT.

例6) 同じ文からの訳文でも日本語とフランス語では subject の扱いが違う。

EN: John was given a book. {PASSIVE}

JPN: John-ni hon-wo ageta

[Subject Drop, ACTIVE]

FR: On a donn'e un livre 'a John.

{Neutral Subject, ACTIVE}

(4) 言語の意味と翻訳

L BMTで使っている言語理論は、あくまで単言語(mono-lingual)内で閉じた言語理論である。機械翻訳の全ての問題は翻訳対象となるそれぞれの言語の言語理論の問題に帰着できるのだろうか、また、翻訳の等価性は各言語の言語理論で抽象化された形式に基づき定義できるのだろうか、いずれの答えも「ノー」である。

例えば、「どのような手続きをすればよいですか」という日本語が“How can I apply?”という英語に翻訳できることは、局所的な構造だけで処理するL BMTでは説明できない。

また、次も同様の例である。

日本語：色及びサイズは部品表によるものとする。

直訳：Colour and size are supposed to depend
on the parts list.

自然訳：Find colour and size in the parts list.

表現とその意味の関係は複雑である。特に、翻訳の基本単位(unit)に何を据えるか、とどの表現を選択するかは、機械翻訳にとって重要な課題である。LBMTではこの2つでいずれも誤りを犯している。例えば、“wear”は日本語では次のように訳し分ける。

wear shoes → haku

wear a hat → kaburu

wear specs → kakeru

ところが「銀行強盗」の場面で“wear tights”は「タイツを履く」ではなく「タイツを被る」と訳す必要がある。このことは翻訳の単位を文の構造ではなく、情報参照の範囲とする必要があることを示す。つまり、理想的な機械翻訳システムでは、

- [1] テキストで記述される場面の中で何が起っているかを理解する
- [2] タイツが体のどの部分に着用されるかを推測する
- [3] 適切な日本語を選ぶ

といった処理を行うべきである。しかしながら、実際にこのようなことは可能なのだろうか、また本当に必要なのだろうか？

そこで次に、電車や映画館で「チケットを切る」行為の表現について考える。英語では「チケットを投入する(slot)」、日本語では「チケットにパンチを入れる(punch)」という『動作』を表す単語を用いるのに対し、ドイツ語では「チケットを有効にする(valid)」というように『状態』を表す単語を用いる。

更に興味深いのは、フランス語では全く逆の「チケットを無効にする」という単語を用いる。これらの違いは言語の意味だけでは決してわからない。

同様に、訳し分け(collocation)も慣習(convention)であることが多い。

Our firm {made|did} a deal with them.

The simmer {had|got} a cramp.

これらの例をみていくと、翻訳では、文脈に基づく動的に処理をしたからといって得られるわけではなく、領域の静的な知識に属する何かが必要であることがわかる。

この、言語表現と指示するもの(denotation)、2つの言語表現間を結ぶものを、慣習的結合(conventionalized link)と呼ぶことにする。

別の例として UNIXのマニュアル類によく現れる“match”という単語を考えてみる。市販の辞書では、「組み合わせる」「縁組みさせる」「取り組ませる」「競争させる」「匹敵する」「調和する」「一致する」「配合する」「整合する」という訳語が与えられているのに対し、UNIX マニュアルの訳本では「一致する」(全115箇所中 105)「対応する」(同 7)

「比較する、同じにする、認識する」(同各 1)というように辞書とは全く違う頻度で訳されている。更に「一致」「対応」の訳し分けに関しては、目的語が“string”に属するものか、“parenthesis”に属するもので決まっているのであって、動的な文脈で決まるものではない。このような対応は文書から抽出すべき領域に特化した規則である。UNIXマニュアルには、他にも、“new”の対義語は“old”ではなく“existing”である。

“described”と“discussed”は同義語である。といった規則も内在している。

結局、こういった慣習的結合は対象となるコーパスから得るしかない(知識獲得; knowledge acquisition)。今後このような領域固有の規則を(一般の言語規則と区別しながら)いかに得るかが重要な課題とある。

(5) 統合された枠組み

既に述べたように、翻訳の等価性は言語理論から得られるもの(合理主義者; rationalists)ではなく、人間の翻訳という行為そのものに内在する(経験主義者; empiricists)と考える。機械翻訳システムでは翻訳例がそれに相当する。前者が高度に抽象化されたレベルの表現なのに対して、後者では等価性が実例に基づき定義可能である。この場合、計算メカニズムとしては制約ベース(constraints-based)が有望である。

翻訳にとっての言語の意味とはテキストから動的に構築する文脈に基づき理解した解釈(interpretation)ではない。このことは、機械翻訳システムが翻訳文章を動的に処理しながらの「理解を介した翻訳」ではなく、理解をあくまで補助的なレベルでの表現と考え「理解を伴った翻訳」で十分であることを示唆する。

最後に、EBMTやSBMTと、LBMTが統合された機械翻訳システムを以下に挙げる。幾つかはColing 94の本会議での発表が予定されている。

CMU: Despatcher, Common Data Structure
AT&T: 定性的/定量的要素 (SBMT)
ET-6: LBMT and KBMT
Toshiba: customization
ATR: EBMT and LBMT

(この原稿はColing'94本会議に先立ち開催されたチュートリアルでの講演の要旨を記述したものである。)

「翻訳の日」の行事

9月30日は国際翻訳家連盟(FIT)の制定した「翻訳の日」であり、各国翻訳関連の団体では各種のイベントが展開されるが、「翻訳の多彩な局面」とのテーマの下に(社)日本翻訳連盟は次のイベントを行う。

《東京会場》

月 日 9月30日 13.30～

場 所 (大手町)サンケイ会館

行 事

第1部 基調講演 常盤信平氏(直木賞作家)

第2部 パネル討論「こんな翻訳が欲しい」

第3部 表彰式(翻訳作品、翻訳をテーマにした作品)

第4部 懇親会

併 設 翻訳支援ツール実演展示会

《大阪会場》

月 日 10月1日 10.00～17.30

会 場 (天満橋)日本経済新聞 大阪本社

行 事

第1部 基調講演 大澤水紀雄氏

報告 東京会場パネル討論

第2部 分科会

A ネイティブの英語表現

B 日本語訳文の品質をいかに高めるか

C 翻訳支援/システムの評価と利用法

第3部 翻訳者の就業環境を考えるセッション

参加登録料 2,000円(JTF会員) 一般 3,000円

主 催 (社)日本翻訳連盟

後 援 通産省(予定)、NHK、日経新聞、出版文化
産業振興財団

協 賛 アジア太平洋機械翻訳協会ほか

多言語機械翻訳プロジェクト

MMT' 94

(第4回実証試験・国際シンポジウム)

近隣諸国間の機械翻訳システムに関する研究協力プロジェクト(中国語、インドネシア語、マレーシア語、タイ語および日本語の多言語機械翻訳システム)は本年度を以て共同研究を終了するが、8年間に及ぶ研究成果を実証試験及び国際シンポジウムで公表する。

日 時 12月5日～6日

場 所 (紀尾井町)ホテル ニューオータニ

実 証 試 験 ・多言語機械翻訳システム

・多言語ワードプロセッサ

・多言語電子化辞書

・中国語OCRシステム

(見学自由)

シンポジウム (同時通訳有り)

・CICCプロジェクトの概要

・基調講演

・参加各国の研究成果報告

・招待講演

・研究成果の利用

(定 員)200名 参加料無料

主 催 (財)国際情報化協力センター

後 援 通産省

問合せ先 〒108 東京都港区芝5-30-9

(財)国際情報化協力センター

機械翻訳システム研究所

☎03-3798-5085 FAX03-3798-7294

機械翻訳パーソナルユースの現状と対応

機械翻訳システムが商品化されてから早くも10年を経過した。この間、機械翻訳の利用に関する各種調査が行われ、その結果に基づいて製品機能面での改善や改良も実施されてきた。コンピュータのダウンサイジングやライトサイジング化の進展にともないここ数年来、機械翻訳においてもパソコン利用による翻訳ソフトの需要が急成長を遂げている。このような状況下において、当協会利用技術研究会ではパソコン翻訳ソフトの利用実態調査を実施し、その集計結果については「6月22日の機械翻訳セミナー」で発表した。ここでは従来のW/S（ワークステーション）レベルとパソコンレベルでの利用面における相違について分析するとともに、今後の対応策についても考えてみたい。

パソコン翻訳ソフト利用調査の概要

調査月日 平成6年4月20日～5月10日

調査対象 PC Transer ユーザ

調査数 1,000名

回収数 110名(回収率11%)

調査方法 郵送によるアンケート調査

1. 集中化から分散化へ

パソコン翻訳ソフト利用ユーザのアンケート回答者の構成は企業人69%、大学10.9%、自営業10%、公務員10%の比率である。職種別にみると研究技術部門の43.6%をはじめ医学14.5%、設計、情報処理などいわゆる理系職種での利用が多いが、教育、経営、営業など文科系職種での利用が21%近くを占めてきていることである。

従来のワークステーションレベルの機械翻訳システムはその殆どと云っていいほど企業内の翻訳部やマニュアル制作部などで統合的に、あるいは集中的に使用されて来たものであるが、パソコン翻訳はさらに一歩踏み出し、大学や企業内の個人部門へと次第に拡大しつつある。これはコンピュータが汎用型から部門単位のオフコン、更に個人レベルへのパソコンへと浸透して行った経緯からみてもほぼ同じ経

過をたどっており、企業内での分散化傾向が顕著にうかがえる。

企業内での集中翻訳では、えてして大量翻訳が優先され、各部門からの翻訳依頼は、件数は多いが枚数の少ない原稿は頼みにくいし、個人の仕事のスケジュールと合致しにくい場合もある。また集中翻訳といっても各部門毎の専門辞書が揃っているとは限らないし、かえって個人が専門用語やユーザ用語を習得するほうが早く、気が楽だという点があるのかもしれない。

以前は企業内では貿易部門や海外事業部門がこれら翻訳など国際業務を担当していたものが、海外生産へのシフト、国際技術交流の活発化、輸出入の増大などによって、国際的業務は全部門でも日常業務化

	企業	大学	自営業	公務員	合計
研究	22	0	0	2	24
技術	18	0	1	0	19
医学	0	8	2	6	16
情報処理	6	1	0	0	7
営業	5	0	1	0	6
教育	4	0	2	0	6
設計	5	0	0	0	5
技術	5	0	0	0	5
事務	5	0	0	0	5
商品企画	2	0	0	0	2
社会	0	2	0	0	2
特許	1	0	1	0	2
品質管理	1	0	0	0	1
ソフト	1	0	0	0	1
会計士	0	0	1	0	1
生産管理	1	0	0	0	1
化学	0	1	0	0	1
機械	0	0	1	0	1
不動産	0	0	1	0	1
経済	0	0	0	1	1
建築	0	0	0	1	1
畜産	0	0	0	1	1
幼稚園長	0	0	1	0	1
合計	76	12	11	11	110

し、事務部門、管理部門など各部門へ拡大しつつある現状からすれば、今後この社内分散化傾向はさらに拍車がかかってくることは容易に想像できるのである。

2 英語能力と機械翻訳購入動機

利用者の英語能力について質問したところ「論文も十分に読みこなせ、英語は得意」とする人が29人(26.4%)を占めており、謙遜もあるかも知れないが不得意と答えた人は51人(46%)となっている。

この英語が得意と云う人の「購入動機」を見ると「仕事の上での必要性が増大したため」が最も多く、29人中17人を占め、次いで「仕事の能率を向上させるため」が12人(重複回答あり)となっている。さらにこの英語が得意とする人々の利用ソフトは英日翻訳が29人の内18人、日英翻訳が9人である。これは情報生産の構想から、作成、翻訳、加工、印刷、提供など、文書のトータル処理の中で、翻訳作業はその1プロセスに過ぎず、情報生産の効率化をはかる一環として、生産性向上策としての機械翻訳の導入が進展しつつあるとみて差し支えないだろう。

一方「英語を不得意」としている人々(51人)の購入動機を見てみると、やはり「仕事の上でその必要性が増大して来た」が51人中21人、「仕事の能率を高めるため」が17人、「翻訳ニーズが増大してきた」が11人となっている。利用ソフトは英日翻訳が38人、日英翻訳が6人、その他は英日、日英両方を利用している。

英語能力

	得意(論文可読)	日常会話 OK	不得意	その他	無回答
企業	16	18	40	2	0
大学	5	1	5	1	0
自営業	2	4	4	1	0
公務員	6	2	2	0	1
計	29	25	51	4	1
(%)	26.4	22.7	46.4	3.6	0.1

購入(利用) 動機

(重複回答有り)

	企業	大学	自営業	公務員	計	(%)
貸与品	0	0	0	0	0	0.0
業務の必要性の増大	40	5	4	5	54	50.0
翻訳ニーズの増大	16	2	1	4	23	20.9
趣味を活かすため	2	0	3	1	6	5.5
仕事の能率を向上	28	5	5	5	43	39.1
その他	0	0	0	0	0	0.0
計	86	12	13	15	126	100.0

そして翻訳の精度は「外部へ情報提供のため完全翻訳が必要」とする人が3人であるのに対し、文意が分かる程度の意識で良い」とする人と「流し読み、要約程度の翻訳」が合わせて25人で約半数を占めている。これらからすれば海外情報の収集に重点が置かれているようにも見受けられる。

従来のワークステーションレベルでの大量高速翻訳がマニュアル、技術文献翻訳などどちらかといえど広報的、指導説明的な目的のための翻訳が多かったが、パソコンレベルでは個人に情報武装のための目的とした翻訳という傾向も目立っている。

3 個人購入が20%を占める

購入形態としては業務上の購入が88人(80%)に対し、個人購入が21人(20%)を占めている。この個人購入者についても業務用として使用している人が5人(4.5%)もあり、それだけ翻訳ニーズが個人の業務の中に逼迫してきている事がうかがえる。

この個人購入者の内訳は企業人が9人、自営業4人、大学5人、公務員3人であり、翻訳内容はやはり論文作成や論文購読が13件、レポート作成、購読が6件、マニュアル、説明書、カタログの翻訳が14件(重複回答あり)などとなっている。

パソコン翻訳ソフトは現在市販されているもので英日、日英など片方向のもので7万5千円から19万8千円くらい、双方向で34万6千円くらいである。これに専門用語辞書が1分野について3万から10万円はするものが多い。これにハード本体や周辺機器としてのOCR、プリンタ等を加えれば、結構高価なものなる

	0%	1-39	40-60	61-99	100
全文完全翻訳	60	20	11	6	13
文意が判れば良い	24	22	27	19	17
要約程度の翻訳	53	26	21	6	3
辞書代わり	86	19	3	1	0
その他	98	14	0	2	1

	業務購入	個人購入	貸与品	その他	無回答
企業	70	5	0	0	1
大学	6	6	0	0	0
自営業	6	5	0	0	0
公務員	6	5	0	0	0
計	88	21	0	0	1
(%)	80.0	19.0	0	0	1.0

が、それでも、従来のW/Sレベルのものに比べれば約半額である。機能的な点からすればやはり処理スピードや容量、支援機能などの点で見劣りすることは否めない。しかし用法や用途を限定すれば十分利用できるものである。従って今後部門用、個人用としての比率が高まって来ることは間違いない。

利用形態としてはスタンドアロン(単体)使用が95.5%、システム化利用は3.6%、ネットワーク利用はわずか0.9%である。パーソナルユースという観点からすれば自然な形態と云えるかも知れないが、原文入力や辞書構築とその相互利用、例文データベース構築など利用領域の拡大などを考え合わせれば、今後企業内やグループにおいてはネットワーク化を推し進めて行く必要がある。

4 使用者は中堅サラリーマンが多い

利用者を年代別に見てみると、20代が12.7%、30代が42.7%、40代が20%、50代が16.4%、60代が4.5%という構成である。従来企業内での機械翻訳担当部門では20～30代前半のオペレータが多く作業に従事しているが、パソコン翻訳では30～40代が多いという特色がある。

個人購入では購入価格という面からすれば中高年層が多くなることは想像出来るが、業務購入が80%を占めている現状からすれば、やはり社内関係部門からの翻訳依頼や翻訳加工、あるいは中間的作業ではなく、中堅社員や技術者みづからが創造的な業務の一

環として、直接パソコン翻訳を利用していると見て良いのではないだろうか。ちなみに翻訳の内容をみても論文作成や論文購読が60%と圧倒的に多く、このほかレポート作成や購読が24%、またマニュアル、説明書、カタログ等から海外情報を収集するという利用形態をとっている。

5 利用システム

機械翻訳システムは大きく分けて入力装置、翻訳装置、出力装置から構成されている。翻訳装置としては専用機として、あるいは汎用機としてのW/Sが多く使用され、このハードに合致した翻訳ソフトが内蔵されている。入力装置としては、論文、技術資料など印刷媒体からの原文情報を光学的に読み取り、入力データを翻訳部へ記憶させなければならない。OCRがなければ、原文情報をFDやHDから移植入力するか、キーボードから入力することになる。

今回の調査対象者ではOCRを所有していないユーザは51名(46.4%、無回答も含め)となっており、原文入力環境としては必ずしも十分とはいえない状況にある。最も現行のOCRの認識精度も完全なものではなく、手入力のほうが早くて、確実という人も少なくないことも事実であろうが、パーソナルユースという観点からすれば、廉価で手軽なOCRや出力装置としてのプリンタの創出が望まれる。と同時にペーパレスの時流に沿ってFDやHDのやり取りや通信回線を通しての入力など翻訳プロセスの簡素化をはかる上においても当面の問題として解決されなければならない事であろう。

6 翻訳の精度は「文意が分かればいい」

専門部門としての翻訳部やマニュアル制作部などでは社内外に対して翻訳サービス(商品)を提供する必要上、完全翻訳が使命であり、云わばプロとして、またセミプロ的な業務を遂行する必要があった。それだけに分野別の専門用語は当然のこととして、そ

機械翻訳利用形態

	単体利用	システム利用	ネットワーク利用	その他
企業	72	3	1	0
大学	12	0	0	0
自営業	10	1	0	0
公務員	11	0	0	0
計	105	4	1	0
(%)	95.5	3.6	0.9	0

	20代	30代	40代	50代	60代	無回答
企業	13	35	14	11	3	0
大学	1	8	1	1	0	1
自営業	0	1	4	2	2	2
公務員	0	3	3	4	0	1
計	14	47	22	18	5	4
(%)	12.7	42.7	20.0	16.4	4.5	3.6

	0%	1-40%	41-70%	71-99%	100%
OCRで読み取り	46	11	7	13	30
通信回線入力	94	4	5	1	3
FD入力	78	20	3	2	4
キーボード入力	50	17	3	7	29
デジタル機器→FD入力	101	1	2	2	1
無回答	3	3	3	3	3

翻訳の種類

	企業	大学	自営業	公務員	合計
論文	41	11	3	11	66
マニュアル	36	1	4	2	43
説明書	31	1	5	2	39
レポート	16	0	6	4	26
カタログ	13	0	6	0	19
雑誌	9	1	5	3	18
連絡文書	10	0	5	1	16
手紙	11	0	3	0	14
特許書類	11	0	2	0	13
規格	11	0	2	0	13
契約書	10	0	2	0	12
新聞	2	0	2	2	6
法令	1	0	2	2	5
案内書	3	0	2	0	5
電子メール	3	0	1	1	5
オンライン	2	0	1	0	3
見積書	1	0	1	0	2
指図書	0	0	0	1	1
小説等	0	0	1	0	1
DB	0	0	0	0	0

の分野における文体、言い回し等もサービス商品としての付加価値を高めるためにも必要な事とされてきた。

しかしパーソナルユースでは勿論これらの要素は重要な問題であることには違いないが、翻訳というよりも海外情報の収集や、伝達の手段としての利用のウエイトが極めて高いので、情報の大意、概要さえ分かればいい、文意さえわかれば、多少言い回しや、テニオハがおかしくても構わないといった傾向が非常に強く出てきている。

調査結果から見ても翻訳目的として「完全翻訳」を必要としているものは11.8%であり、文意が分かればいい、要約程度の翻訳であればいいというのが旺

倒的である。この事からしても情報の収集手段としての機械翻訳の利用という傾向が強うかがい知れるのである。

また専門翻訳部門では業務の性格上 EJ/JEの双方向翻訳が絶対不可欠であったが、パーソナルユースとなると、個人のおかれている立場、翻訳の目的からして片方向の翻訳でもいいという面もある。

英文は機械や専門家の力を借りなくても何とか読めるが、外部へ出す文章だけは、正しく情報を伝えるためにも機械翻訳を使用したいという意向もうかがえる。

更に論文作成の下訳として、専門用語や例文データベースの蓄積、検索用としても利用するという使われ方もされている。つまり外文ワープロとしての利用という性格付けが目立っている。

7 翻訳内容は少量少品種

1件当たりの翻訳量はA4換算で3頁以下というのが25%、4~5頁が24.5%、6~10頁が17.2%である。

W/Sレベルは大量翻訳が主体であるのに対し、パソコン翻訳は少枚数翻訳用に利用されている。もっとも1件 500頁以上の翻訳を行っているというのも2人(1.8%)ある。

また月間の翻訳頁数にしても10頁以下が32%、20頁以下が15%、あわせて半数近くが20頁以下の少量翻訳である。翻訳サービス専門部門ではやはり月間2,000~3,000頁以上こなさないとシステムの償却にも問題をきたすことになるだろうが、パーソナルユースでは翻訳は単なる手段であり、翻訳によってえられる情報がむしろ重視され、みずからの業務の生産性を向上させる糧になるもので、やはり利用観点が違うということになるのであろう。

翻訳量

1件当たりの翻訳頁数(A4換算)(平均頁数)

	1-3	4-5	6-10	11-20	21-30	31-50	51-100	101-150	151-499	500-	無回答
企業	21	21	10	10	3	3	3	2	0	1	2
大学	1	1	5	3	0	1	0	1	0	0	0
自営業	4	2	2	0	0	1	1	0	0	1	0
公務員	2	3	2	1	1	1	0	0	0	0	1
計	28	27	19	14	4	6	4	3	0	2	3
(%)	25.5	24.5	17.2	12.7	3.6	5.5	3.6	2.7	0.0	1.8	2.7

8 編集の手法

今回の調査対象では英日という片方向のソフト利用者が52名(47.%)、日英翻訳ソフト利用者が22名(20%)その他が日英、英日の両方の翻訳ソフトを使用している。

機械翻訳システムは自動翻訳システムではない。現状では翻訳のためのツールである。したがって正しい翻訳のためには前編集や後編集が不可欠である。

編集のやり方については完全翻訳を目指す場合と文意さえ分かれば良いという考え方の違いによって編集のやり方も違ってくことは事実であるが、やはりなんらかの手を加えなければ十分に意味が通じる翻訳文にならないケースが多い。

英日翻訳の場合、全く前編集を行わない人が52人中12人(23%)に達している。一方短文化は必ず行うと云う人が6人、必要に応じて行う人が15人、時々行う人が4人であり、あわせて25人(48%)の人が何らかの形で前編集作業を行っている。スペルチェックは30人(58%)が行っているが、これはキーボード入力が多いことが原因していることもあろうが、機種によっては自動的にチェックできるものもある。このほか前処理項目としては構文補正が40%、品詞指定、適語選択がそれぞれ35%行われている。機械翻訳専門担当者の中には英日翻訳は、英文は「前処理は行わない。あるいはネイティブにチェックして貰う」「前処理よりも後処理に重点をおく事が秘訣」という人もいるが、やはり長文や複雑な構文をもつものは前処理も必要になってくるだろう。

次に日英翻訳ソフト利用者を見てみよう。利用者22人中全く前処理を行わないという人は2人(9%)であり、英日翻訳に比べれば前処理を行っている人が

多い。これも曖昧性が多い、日本文の現状を現わしているためであろう。前処理項目としては短文化や主語、述語、目的語の補充が多く、ついで構文補正、不用語の削除の順で、英文より和文の方が手間がかかっている様子がうかがえる。

9 辞書

パソコン翻訳の基本辞書では5万語から多いものは13万語も装備されている。新聞記事で使用される日常用語は2万語、週刊誌の場合でも2万語程度、但し多少扱い分野が異なるので、合わせて3万語の基本辞書があれば翻訳には先ず支障をきたさないと云われている。しかし年と共に言葉は変化していくもので、この追廃をおこなった上での使用語数であるから、メンテナンスを行わないとすれば必然的に必要な基本辞書語数も増大していくものである。さらにマスコミ関連では必要以上にカタカナ英語を生産したがる傾向がある。国際的に通用しない和製英語が多いことも事実で、素人には中々その見極めがつけにくいのも問題である。

また各専門分野においても独特の文体もあれば用語もある。更に技術の進歩とともに新語も生まれる。これは進歩の著しい分野ほど新語生産が多いが、中にはすぐ死語化するものも少なくない。科学技術情報センターの話によれば年間1万語の新語が誕生しているという。これら新語はまとめて電子化され、市場に専門辞書として商品化されるには2~3年はかかるものもあり、また需要の少ない分野では商品化される頻度も少ないという問題がある。

今回の調査で「専門用語辞書をいまだ揃えていない」と答えたユーザは110人中38人(35%)もある。この内7人はユーザ辞書を構築しているが、その分野は「原子力、獣医学、地理学等」である。業界によって用語の統一が積極的に実施されているところもあれば、小さな業界ではそこまで手が回らず、ユーザ用語が優先して使われている業界もある。

また「専門用語辞書を揃えている」と答えたユーザは110人中72人であるが、その内「1分野のみ」が34人、2分野の専門用語を揃えているのが21人、3分野が11人、4分野が4人、5分野が2人となっている。勿論利用者の職種によって必要な専門辞書の範囲も限定されてくるが、同じ分野、例えば電機分野といっても研究部門、生産部門、営業部門など職種によって使用される言葉やその内容、文体までも異なってくる。

前編集の手法

	◎ (必ず)	○ (必要に応じて)	△ (時々)	× (全くしない)	無回答
短文化	20	24	8	16	19
主語の補充	15	30	11	14	17
述語の補充	8	18	17	20	24
目的語の補充	3	22	17	15	30
構文補正	10	23	19	14	21
不用語削除	2	23	17	15	30
品詞指定	2	20	13	18	34
適語選択	13	27	6	13	28
スペルチェック	22	21	13	13	18
挿入句指定	3	11	11	25	37
誤字チェック	16	14	8	26	22
原文の書き直し	4	9	12	30	32
その他	0	0	0	10	77
全く行わない	23	0	0	0	0

未知語の処理

	企業	大学	自営業	公務員	合計
原文のまま	48	8	5	7	68
カタカナ表記する	6	1	1	0	8
独自解釈による表記	25	1	3	4	33
その他	0	0	0	0	0
無回答	12	2	2	0	16

使用中のパソコン翻訳ソフトの評価

項目	良い / 満足 (%)	普通 (%)	悪い / 不満 (%)	解らない / 無回答 (%)
操作性	21(19.1)	69(62.7)	16(14.5)	4(3.6)
翻訳品質	1(0.9)	4(3.6)	71(64.5)	4(3.6)
翻訳速度	21(19.1)	66(60.0)	20(18.2)	3(2.7)
機器互換性	5(4.5)	44(40.0)	32(29.0)	29(26.4)
前処理の容易さ	2(1.8)	47(42.7)	48(43.6)	13(11.8)
保存専門用語	9(8.2)	47(42.7)	34(30.9)	20(18.2)
システムの拡張性	3(2.7)	37(33.6)	41(37.2)	29(26.3)
価格(機能対比)	10(9.0)	37(33.6)	50(45.5)	4(3.6)

反面半導体など各分野にまたがって使用され、基本語化されても不思議ではないものもある。技術の融合化、業態化が進行して行く中で、具備されるべき専門用語も増加傾向をたどっていくことは明らかである。技術文献は構文的にはそれほど複雑なものは多くなく、むしろ専門用語が翻訳精度の70%近くを制すると云われていることからすれば、この辞書整備にはその管理方法も含めて標準化など留意されるべき問題も多い。

10 手間のかかる処理

機械翻訳を行う上において「手間のかかる処理について」自由記入方式で質問したところ「前編集」が12人である。先に述べたように前編集を行わない人が英日で12人、口英で2人、英口口英で9人である。従って前処理を行っている人が87人ということになるがその内12人(13.8%)が「前処理に手間がかかる」としている。これがため「自動編集機能、構文解析機能、高速編集機能、AI機能の充実」を要望している。しかし大半の人は自動翻訳システムでない以上前処理はやむを得ないものと理解されているものと思われる。

次いで多いのが「OCR入力」である。OCR所有は59人であるが、その内8人が入力に手間がかかるとしている。OCRの認識度は公表99.5%であるが、原稿の汚れ、記号、原稿の誤字脱字、印刷活字の大きさ、OCRの認識誤り等によってその修正作業に手間がかかることを意味しているものと思われる。反面OCRの内蔵を希望する人も3人いる。

その他「ユーザ辞書の登録」も手間のかかるものの1つに上げられており、自動辞書登録機能やCD-ROMによる廉価な辞書の提供などの要望が強い。

11 システムの評価

これは W/Sレベルとパソコンレベルの機械翻訳システムを対比しての評価ではなく、利用上の評価結果であるが、機器の操作性については「満足」、「普通」を合わせて82%、翻訳速度についても「満足」「普通」を合わせて79%であり、相応の評価が行われている。

機器システムとの互換性については「普通」が40%「不満」が29%である。パソコン翻訳ソフトは普及度の高い商品をベースにして設計されていることからすれば汎用性に不満がでるのも致し方ない事かも知れない。

前処理の容易さという点については「容易」とする人が2人、「普通」が47人であるのに対し「容易でない」とする人が48人とあいなかばしており、今後とも前処理支援機能の一層の充実が望まれる。

翻訳品質については64.5%の人が不満と答えている。これは今回の調査対象者の利用経験年数が6カ月未満というのが57.3%、6～12カ月という人が18.2%であわせて1年未満が四分の三を占めている事や、編集方法や辞書構築の利用実態から見ても「いまだに機械翻訳を完全に使いこなしていない」という理由もあるだろう。また英語が不得意という人が多く、それがため機械翻訳に期待したが、期待通りの翻訳精度が得られないという理由も考えられる。しかし翻訳品質は「満足」、あるいは「普通」と答えた人が5人(4.5%)いることも事実である。パソコン翻訳ソフトを使用し、月間500頁以上の翻訳をこなしている企業人も3人もいる現状からみても、一概に品質問題を論じることは出来ないのではないと思われる。

12 今後の対応策

さてこれまでの諸問題をまとめてみると

- 1 機械翻訳はさらにパーソナルユースの旺盛な需要に支えられて研究技術部門からオフィス部門への拡大傾向を強めて行くことが予想される
- 2 翻訳精度も完全翻訳から意識、要約など情報収集伝達のための翻訳という傾向も目立つ。
- 3 英語能力が得意という人も機械翻訳を利用しているが、翻訳が文書管理のトータル処理の一環としての位置付けにあることを物語っており、仕事の効率、生産性の向上のための利用と云う傾向が強い。
- 4 翻訳品質については問題が多いが、利用者の技術の向上、システム側の支援機能の強化、さらに利

PC 翻訳ソフトの需要の延び
(今後 3 年予測)

	0%	-10%	-20%	-30%	-50%	-75%	-100%	-150%	-200%	200% 以上	無回答
企業	3	3	8	11	11	1	12	2	3	2	20
大学	0	0	1	2	1	2	3	0	0	0	3
自営業	0	0	1	3	1	1	0	0	0	0	5
公務員	0	0	1	0	0	1	2	0	0	2	5

用環境の改善等が急務である。

以上のことから今後の対応について考えてみよう。

先ず利用技術についてであるが、機械翻訳専門部門の担当者でも一人前のオペレータになるには 6 か月から 1 年はかかると云われている。これは人間翻訳者が 1 人前になるには 10 年の経験が必要とされていることからすれば極めて効率が良いと云わねばならないが、しかしそのオペレータもグループ内で毎日専門的な業務として機械翻訳に取り組んでの上での話である。

パーソナル利用となると機械翻訳ばかりに取り組んでおれないという事情もある。それだけに利用技能の習熟には時間がかかるのは当然であろう。したがってパソコン翻訳に取り掛かる前には十分な事前準備が必要となる。つまり日英翻訳には前処理は欠かせないものであり、この技能を完全にマスターしてかからないと希望通りの翻訳精度は期待出来ないということである。翻訳マニュアルも製品によってその内容がまちまちで、必ずしも完全ではなく、統一性に欠ける点も多いが、少なくとも翻訳作業に入る前には翻訳マニュアルだけは完全にマスターしておくことが肝要である。

機械翻訳オペレーターの資質としては語学力と技術力が必要とされている。英語が苦手という人も改めて文法の本を読み返してみることも必要であろう。

忘れていた文法規則がよみがえればそれだけでも編集作業に大きくプラスする。単語は機械の方が抜群の記憶能力をもっているのも機械に依存すればよい。現在の機械翻訳が文法規則によって組み立てられていることからすれば事前準備としてこれを励行したいものである。

また専門辞書やユーザ辞書もこまめに整備して蓄積して行く必要がある。辞書を構築するにしても初めから数万語もある用語を作ろうとすれば大変な労力が必要となり、途中で挫折することも少なくないが、翻訳作業と平行して辞書を蓄積していくと、知らず知らずの内に蓄積されていくものである。市販の専門用語辞書といえども自分にとって必要でない単語もあるだけに、自分の蓄積した単語ほど利用頻度の高いものとなる。単語数が多すぎても検索するに

も手間がかかり過ぎることも有り得る。

ユーザ辞書の構築に際しては事後の相互利用をも考慮し、辞書管理ルールを明確に設定してから取りかからないと品詞区分が不明確になったり、重複登録したりで、かえって使用時に混乱をもたらしかねない。

また単語、用語ばかりではなく例文も大いに蓄積する必要があるだろう。定形的な文章、例えば契約書とか連絡文書などではその都度翻訳していたのでは時間もかかるし、効率的ではない。論文、レポートなど引用すべき文句も多いだけに、蓄積しておく必要があるし、また後処理された文書は以降の翻訳に再利用していくことが肝要である。

翻訳専門部門では OJT などで常に翻訳効率向上のための情報交換が行われているが、パーソナルユースにおいてもユーザが使用中に開発した利用ノウハウをお互いに情報交換でき、共用化できるような環境づくりも必要であろう。

またソフト供給側においても翻訳支援ソフトの充実に努めて行く必要があろう。パーソナルユースユーザは自ら支援ソフトを作成するのは難しいことから、文章推敲ソフトや自動編集に近づける各種支援ソフトを充実させていく事が急務であろう。

さらに翻訳ソフト購入者への利用技術講習会の開催や利用ノウハウの提供、情報交換の場の提供、的確なフォロー策が望まれる。

またこれ程までに普及が進展してくると翻訳インフラの整備も重要になってくる。例えばOA店がコピーサービスを実施しているように機器の互換性をカバーするような電子変換サービスや印刷媒体から電子媒体への変換を請け負うサービス、通信回線を通じて辞書の相互利用が出来るような会員制辞書センターなど、また個人が蓄積した辞書の公開や共同利用、ユーザ辞書の積極的提供などが整備されなければならない。翻訳インフラといっても大袈裟なものではなく、企業内ネットワークが整備されつつあり、また情報ハイウェイが叫ばれている今日、企業内、グループ内において、この実現にむけて取り組まなければならない時期に來ていると云えるだろう。

日立製作所中央研究所は、1942年に「今日の開発とともに、10年、20年先を目標とした研究をおこなう」という理念のもとに設立されました。

研究分野はインフォメーションとエレクトロニクスの広い分野に及びます。約1,000人の研究者が、新しい技術潮流の創出をめざして、基礎から応用に至るさまざまな研究に取り組んでいます。ここでは、翻訳関連の研究として、機械翻訳の研究と手話通訳の研究を紹介します。

1) 機械翻訳の研究

日立における機械翻訳の研究は、これまでシステム開発研究所で行われてきましたが、昨年中央研究所の研究テーマになりました。機械翻訳の研究グループが中央研究所に移り、次世代のシステムをめざして新たな展開をはかっています。

日立は、機械翻訳システムの研究を早くから続けてきたメーカーの一つです。80年代前半には、擬似的句構造に基づく構文トランスファ方式の英日機械翻訳システム、さらに概念依存図式に基づく意味トランスファ方式の日英機械翻訳システムのプロトタイプを開発しました。その成果に基づき、1986年にHICATS/JE¹⁾を、1987年にHICATS/EJ²⁾を事業部が製品化しています。

当時は大型コンピュータ上のシステムでしたが、その後ワークステーション版を開発し、現在に至っています。

HICATSなど、現在、市場に出ているシステムの多くは少しずつ控がってきています。しかし、多くのユーザが期待するレベルとのギャップは大きく、新たなブレークスルーが求められています。意味処理の高度化、文脈処理の導入などとともに、システムに学習能力をもたせることが重要であると思われます。従来のシステムでは、翻訳対象を限定してシステムをカスタマイズすることが実用化するうえで必須ですが、一方で文法規則の追加や修正をユ

ーザが行えないという大きな問題点があります。この問題を解決するため、現在、対訳コーパス、言い換えると翻訳の事例から翻訳知識を学習する技術に特に力を入れて、研究開発を進めています。

2) 手話通訳の研究

日立は福祉システムの研究にも力を入れています。その一つに手話通訳システムがあります。機械翻訳が外国人とのコミュニケーションにおける障壁の除去を目的としているのに対し、耳の不自由な方と健聴者とのコミュニケーションを支援するシステムです。開発したプロトタイプに対して、早期実用化の期待が多方面から寄せられています。

現在のプロトタイプは、語彙に制限があり、簡単な文しか翻訳できませんが、実用化に向けてさらに研究を進めています。

手話通訳システムは、手話を日本語の文に翻訳する手話認識システムと、日本語の文を手話に翻訳する手話生成システムの二つのサブシステムから構成されています。手話認識システムでは、手話の話者はデータグローブ³⁾をつけて手話の動作を行います。データグローブから入力される手の動きや位置のデータを解析し、パターン照合により手話単語を認識します。そのあと、手話には含まれていない助詞などを補って日本語の文を生成し、合成音声により音声を出します。手話単語と日本語の単語との対応は一つ対一つではなく、機械翻訳と同様に意味処理が必要です。

手話生成システムは、日本語の文章を解析して、手話単語の列に変換し、コンピュータグラフィックスによるアニメーションを生成するシステムです。

ここで紹介した機械翻訳、手話通訳のほか、翻訳関連の重要な技術として音声認識、音声合成、対話の研究があります。日立中央研究所はこの分野でも長期的な視野にたって研究を行っています。

(知能システム部 梶 博 幸)

1) = HICATS/JE: Hitachi Computer Aided Translation System/Japanese to English

2) = HICATS/EJ: HICATS/English to Japanese 3) = データグローブはVPL Research Inc.の商標です。

自然言語処理技術の動向

(社)日本電子工業振興協会自然言語技術委員会の93年度の調査研究に関する報告書である。

● 品質評価法の検討

この報告書の内容はテキスト技術、機械翻訳技術、機械翻訳サミット、海外技術動向、MTユーザ利用実態、自然言語技術に関するヒアリング調査という項建てになっている。

「機械翻訳技術に関する調査」では英日、日英翻訳システムの品質評価法開発の中間報告が行われている。英日システムではシステム開発者がシステムの不備をチェックする、ユーザが出力品質面での優劣を比較することを目的とした評価用コーパスを作成している。このコーパスは機械翻訳が取り扱うことが困難な言語現象を、主に曖昧性の解消に注目して収集分類しようとするものである。

平成4年度においては英語の短文中心に評価すべき項目を抽出し、評価基準を設定、把握部文型73文、時刻、助動詞143文、文種31文、選択部については構文60文、訳し分け88文のコーパスを選定した。

平成5年度ではこれに改良と拡充が行われ、またコーパスに付与する設問を中心に書式の改良も行われた。評価例文は英語の文法書、参考書等の文法項目一覧や翻訳困難例文集等から収集された。特に重文、復文、並列、比較文などが重点的に収集され、また単文レベルでの現象についても収集し、コーパスの網羅性についてもチェックされた。この評価基準に則って評価も行われた。今後は重要度、頻度などによるウエイト付けを行い、平成6年度には完成させる。

日英システムでは日英言語対依存の文法項目について設問の分析、修正、例文の付与等を行い、また汎用性、機能網羅度、使い勝手度等の評価軸の検討が行われた。また難易度と機能網羅度(基本機能)、緻密度(個別の詳細機能)を含め必要機能と高度な処理機能との2段階の評価レベルをつけて行くことが検討されている。

また機械翻訳の技術項目についてはシステムの基本機能と翻訳処理について検討されている。基本機能とは翻訳処理部分を除くシステムまわりの機能をさし、ユーザ側からの技術評価基準と開発者のため

の技術基準が検討されている。一方翻訳処理に関する評価については形態素解析、構文解析、意味解析、文脈解析、中間表現と構造変換、生成処理、非文法的な現象の処理、辞書、学習カスタマイズについての評価基準について検討が進められている。これらは平成6年度の完成を目指して引き続き検討される。

● 海外技術動向調査

海外調査については米仏を8～9月、11月は中国に赴き、国際会議や研究機関を訪問した。米国ではニューヨーク市立大学では並列処理による機械翻訳システムと確率に基づく日本語形態素解析について、ニューヨーク大学では自然言語処理研究の全容や電子化辞書、機械翻訳の知識自動獲得について情報交換を行い、またMCU-5(技術評価会)やIJCAI-93(人口知能国際会議)へ参加し、その概要が報告されている。また中国については清華大学で自然言語処理、音声認識と文字認識について、科智語言信息処理有限公司では米欧合作の機械翻訳システムについて、CS&CではCICCとの共同研究の中間言語方式機械翻訳システムや英漢の単言語対機械翻訳について、その他北京大学、方正集团新天地研究所、北京信息工程学院、中国科学院声楽所、北京語学院などの訪問調査結果が報告されている。

● ヒアリング調査

音声認識における言語処理の動向(NTT鹿野清宏氏)言語に関する制約の諸問題——日本語と英語の対応を例に——(名古屋大外池俊幸氏) 英語比較統語論の現況(明学大外池滋生氏)など14テーマについてのヒアリング調査の結果について報告されている。

● ユーザ利用事例訪問調査

日経印刷・機械翻訳センターと通商産業省通商産業研究所における機械翻訳システムの使用目的、利用状況、利用上の諸問題について調査されている。

日経印刷ではMTはドキュメント制作の一環であり、翻訳から版下制作まで行う。ユーザからは自社における基本語辞書についての資料を貰い、FD原稿での未知語を抽出、訳をつけてユーザへ戻して確認後ユーザ辞書に登録する。現在1万7千語のユーザ辞書をもっている。SGML化された文書からSGML部分をマスクして翻訳する。前編集は行っていないが、教育も含めリライターで修正後、後編集者を通して訳文を完成させるというプロセスをとっている。

[自然言語処理技術の動向に関する調査報告書]

(社)日本電子工業振興協会 編集 発行

A I 白書94年版

● 機械翻訳の利用状況

情報処理開発協会では93年7~8月にかけて「①コンピュータの高度利用に関する調査」及び「②人口知能(AI)の利用に関する調査」を実施し、その結果を分析解説し、「AI白書94年刊」を頒布している。

この調査はコンピュータ利用の4,022事業所(基礎資材産業、加工組立産業、生活関連産業、公共サービス、商業金融、教育公務、情報サービス産業等)を対象にし、①では952件(回収率23.7%)②では151件のデータが得られた。

個別AIシステムの導入状況のうち機械翻訳システムは時系列的には90年が3.4%、91年が3.1%、93年3.9%と微増している。1事業所当たりの平均導入率は90年においては1.2%、91年で1.13%であったが、93年では1.12%になっている。(92年は調査されていない)

未導入事業所の今後の導入意向としては導入する意向がないが74.7%であるのに対し、現在検討中が0.7%、将来検討したいが24.6%という高率を示しており、関心の高さを物語っている。

● 機械翻訳の利用分野

機械翻訳の利用事業所(47事業所)の93年における利用分野は、マニュアル翻訳が29.8%、特許等技術文献が23.4%、ビジネス文書12.8%、通信文メモ12.8%、新聞雑誌情報4.3%、カタログ2.1% 統計等経済資料2.1% その他21.3%で、91年の利用分野に比べれば通信文メモや統計等経済資料の翻訳率が増大しているのに対しこれ以外の利用分野は低下しており、翻訳分野が拡散していることがうかがえる。

また利用形態としては英日が日英を若干上回っているが、ほぼ半々で、翻訳結果をそのまま利用しているのは6.1%、人手で修正後利用しているのが39.4%、人間翻訳時の参考程度が27.3%、試験的運用が15.2%となっている。またあまり利用していないのが12.1%になっている。

機械翻訳の問題点としては「翻訳精度の低さ」が64%を占めており、次いで「辞書の開発、維持、メンテナンスが大変」が24%、「原文入力に手間がかかる」となっている。

● 機械翻訳の認知度

一方この報告書では中小企業(138社)を対象に「AI

技術の利用状況」について調査されている。それによると「汎用、オフコン、W/Sの導入状況は65.9%、パソコンの導入が59.4%、この他 NC旋盤、NCフライス盤、産業用ロボットなどの ME機器は35.5%導入されているが、機械翻訳の認知度は「内容を十分知っている」が11.6%「多少は知っている」が32.6%、「言葉は聞いたことがある」が41.3%、「全く知らない」が10.1%、その他4.3%で、未だ認知度が低いことを物語っている。

尤も「人工知能」を知らないが13.8%、エキスパートシステムを知らないというのが30.4%という事からすればある程度関心が持っている方かも知れない。

● 機械翻訳の利用事例

利用事例としてパソコン翻訳ソフト使用の髙テトラネット社が取り上げられている。同社では特殊なケースを除きほぼ100%機械翻訳を利用している。

導入動機としては①翻訳量が増えてくると人間の翻訳能力は急激に低下し、且つエラー率が急増する②翻訳を分担すると全体の整合性を保つことが困難でプログラムのモジュール分割とおなじ問題に遭遇することから機械翻訳を利用している。尤も①機械翻訳の出力を最終成果物として期待はしていない。

②翻訳を一連の工程としてとらえ、機械化の中心構成要素とみなしている。③機械翻訳はまだ発展途上にあるが、現時点でもメリットは多い。そのメリットとは①翻訳速度が早い②訳しもれなどのミスがない③翻訳結果が一定している④ワープロ入力が不要で、タイプミスがない⑤正しい訳語が使用できる⑥英語の出来ない専門家も使用できる⑦英語の出来ない日本語のプロが使用できる⑧辞書管理、成果物の管理が容易である事を上げている。

現状の問題点としては①英語が口語的になって来て構文解析の難しいものが増えている②英語のまま一括して残したいものを指示することが難しい③タグや制御文字が埋め込まれているものは機械翻訳の対象になりにくい④箇条書きの部分が弱い⑤コーディング例など翻訳対象外のものを一括指示できない⑥ビジネスの世界の言葉や表現に弱い⑦ピリオッド以外の区切り文字がうまく認識されない⑧補足的な文が文中に挿入されるとこの挿入部分がノイズとなって訳文がくずれることが多い点をあげている。

今後の問題としてはプレエディットの重要性とポストエディットでは推敲ロジックの充実が要請されている。

[AI白書1994・(財)情報処理開発協会 編集発行]

パソコンの調査研究

(社)日本電子工業振興協会パーソナルコンピュータ業務委員会の93年度調査研究活動の報告書である。

この報告書は3部構成となっているが、その中から関係のある部分を抜粋してみよう。

● パソコン出荷実績

1993年(暦年)におけるパソコン出荷は台数で2,913千台で前年に比べ134.5%も増大した。この国内出荷は2,153千台で122%の伸びであるのに対して、輸出は187%と驚異的な伸長を示した。金額的には総出荷で1兆2622億円で117%の伸び、国内向けは9452億円で117%、輸出は114%増となっている。単価は13.5%程度低下したことになる。この増勢は94年に入っても続いており1~3月では台数で前年同期比で186%、金額では168%増となっている。これはWindows3.1の登場が市場拡大の大きな布石になった。また国内輸出とも32ビット機へのシフトが完了した。

5カ年出荷予測では国内向けは年率113%の伸びで97年度は380万台(台数)であり、そのうちノートタイプは年率118%で160万台、また輸出も103%の伸びで97年度は90万台になると予測している。

● パソコンソフトの利用

使用されているソフトはワープロが82.1%、表計算が72%、データベース47.3%、ゲーム43.6%、通信30.7%、グラフィックス25.1%、言語22.2%の順となっている。

CD-ROMのアンケート調査(回答社121社)も実施されたが、CD-ROM内蔵パソコンは92年度は95千台であったものが93年度は265千台と265%の伸び、94年度238%、95年度は186%の伸びで100万台を突破すると予測している。

CD-ROMの本数は92年度191千本であったのに対し93年度は294千本と154%の伸び、94年度はさらに321%も伸びるとの見通しである。93年度の内訳はビジネス用が47本、家庭個人用191本、教育用が56本であるが95年にむけ家庭個人用が増大すると見込んでいる。

● 教育用パソコンの利用実態

パソコン教育利用の4中学、教師141、生徒2455、父兄2086人への大規模なアンケート調査結果が報告されている。これによると教師が授業で利用しているのは39%、授業以外で利用しているが34%となっている。

科目では技術家庭科が83%、理科78%、社会63%、数学56%、美術60%、保健体育と英語がそれぞれ50%の順である。

また教師のパソコン操作経験は経験ありが94%の高率を示している。授業以外でのパソコン利用では文書作成が81%、表計算グラフ作成が64%、データベースが28%である。教育用パソコンへの要望としてはソフトの内容、種類の充実(39%)、授業用パソコンの増設(36%)パソコン担当者の増員(36%)となっている。

一方生徒についてはパソコン操作経験が94%に達しており、さらに何らかのソフトを利用できる生徒は87%になっている。自宅でのパソコン利用は16%、その内勉強のために利用するが13%(生徒全体では3%となる)、文書作成が48%であり、ゲームが62%と圧倒的に多い。

父兄については自宅にパソコン所有が24%で年間3%の率で増加傾向にある。用途は仕事のためが60%文書作成33%、子供の教育のためは17%に留まっている。

改善要望としてはハード(42%)ソフト(27%)の低価格が目立っている。

● パソコンの将来動向

パソコンハード、ソフト、周辺機器メーカ74社へ将来の利用形態についてのアンケート調査を実施している。現在、5年後、10年後ともパソコンの使用形態は書類作成、定形業務と変わらないが、大きく利用されて行くものとしては電子メール、マルチメディアでの利用であり、利用率が下がるのは計測制御、設計/開発/解析となっている。

回答を基に1位を5点、2~3位を3点、4位2点、5位を1点とウエイト付けした見通しではあるが、書類作成は現在354点であるのに対し10年後は273点、定形業務は278点から264点となっており、機械翻訳については得点は低いものの現在6点、5年後10点、10年後16点と上昇傾向にある。

今後の使用形態としてはネットワーク化(動画伝送、国際ネット、オープン化、TV会議、共同体のネットワーク端末、LAN、WANをベースにしたマルチメディア通信、サーバの利用、汎用情報端末)と他の分野の製品との統合化(画像電話、FAX、映像などマルチメディア化)情報アクセスツール、デスクトップと移動体としての一体化が考えられる上、モバイル化として可搬性のマルチパーパス端末、データエントリー商品の出現も考えられると予測している。

[パソコンに関する調査研究]日本電子工業振興協会

実験翻訳サテライトオフィス

(社)日本翻訳連盟では平成5年度に翻訳サテライトオフィスの実現化についての調査研究を行った。

この目的は翻訳サービス業が首都圏に集中していること、また上級翻訳者が不足していること、地方に分散している翻訳経験者をネットワークを通じて有効活用していこうという狙いで、運用実験を行なった結果が報告されている。

この構想は「翻訳者が持っている翻訳情報」を大規模な翻訳知識データベースとしてセンターに蓄積すると共に翻訳情報の収集提供、例文検索や専門用語、術語の問い合わせにも応えられるようにしようというものである。実験センターは東京中野に設置され、翻訳文書の前後編集作業の分散化と、これら文書を一元管理するセンターシステムがイーサネットで結合された。実験に用いられた文書は中学高校の教科書、物語、情報処理解説書、科学一般、医薬本、サテライトオフィスでの作業は入力、前後編集であり、これを手作業での翻訳と生産性やコスト比較をした評価も行っている。今後の問題として専門辞書の充実、情報通信網の整備や運営管理面では情報の収集方法や情報リテラシー教育の強化が上げられている。

新刊書「オブジェクト指向設計の標準 HOOD 3.1」

ユーザインタフェースやデータベース等さまざまな分野に広く影響をおよぼしてきたオブジェクト指向技術、これはソフトウェア開発/分析の世界でも例外ではない。これまで多くの書籍が刊行されてきたが、どれが標準的存在になるのかとの議論もはなやかである。

HOODは欧州各国からえりすぐられたオブジェクト指向の研究者によって1987年に開発された記述法であり、欧州の数々の大規模制御系プロジェクトで使われている。

今回 HOOD技術グループの執筆による「オブジェクト指向設計標準のHOOD 3.1」を立田種宏氏が翻訳、出版される。(A5判、192頁 定価3,000円)

(問合せ先:海文堂出版 ☎03-3815-3953)

世界の翻訳図書

「世界の統計(国際統計要覧)94年版」が刊行されている。統計資料としてはやや古い、1986年に出版された翻訳図書(米国を除く)は61,531点、その内訳は文学が30,464点で59.3%を占める。ついで応用科学が7,215点で11.7%、社会科学が6,987点11.4%、宗教が3,899点で6.3%、地理歴史5.9%、哲学5.2%、自然科学4.9%、芸術4.2%の順となっている。

国別では日本で出版された翻訳図書は2,878点で世界の4.7%を占める。日本での翻訳部門別では文学物が44%と最も多いが、応用科学は15.5%、社会科学13.1%、芸術6.6%であり宗教は4.1%に留まっている。

翻訳図書出版の最も多い国はスペインの9,649点で世界の15.7%を占めついで東西ドイツ合せて8,936点(14.5%)、旧ソビエト8,202点と続いている。

原書の言語別では英語を原典とするものが61,531点中29,294点で最も多くて47.6%を占め、ロシア語が6,620点(10.6%)次いでフランス語が6,502点(10.6%)ドイツ語が5,079点(8.2%)、日本語はわずか232点で0.4%である。この日本語を原典とする翻訳図書は東西ドイツで45点、英国で30点、スペインで28点出版されたのに対し、米国では僅か1点のみである。(注:米国の統計は85年の数値)

日本で出版された翻訳図書(2,875点)も英語を原典とするものが79.8%を占め、ドイツ語7.8%、フランス語7.4%、イタリア語1%の順となっている。

世界の特許権数

1991年における特許権の動向は、出願件数で10万件を越えているのは日本が380,453件でトップ、米国177,388件、ドイツが109,187件であり、英国が95,533件、フランス79,075件と続いている。

登録件数では米国が96,514件、ドイツが43,190件、ついで日本36,100件、英国が34,074件となっている。

既存権利は米国が圧倒的に多くて1,154,204件(90年数値)で日本の579,695件を大きく上回っている。ドイツは244,488件、英国は222,558件(84年)である。

アジアの中では韓国の出願36,154件、登録8,691件中国は出願11,423件、登録は4,122件となっている。

機械翻訳の辞書

今回は機械翻訳システムで用いられる辞書について説明します。機械翻訳システムで用いられる辞書にも、機械翻訳システムの翻訳実行機能が参照する機械向けの辞書と、利用者が参照する人間向けの辞書とがありますが、今回は、機械向けの辞書に限定して説明します。

1. 機械翻訳用辞書の役割

機械翻訳システムにおける辞書の役割はおおよそ次のように整理することができます。

1) 原文を構成する語をシステムに教えます。

日本語原文の場合、語と語との間に空白がありませんから辞書の必要性はすぐわかりになると思います。英語原文の場合、単語の切れ目は空白によって明白なわけですが、辞書の重要な役割は複数の単語から成る複合語をシステムに知らせることです。

2) 原文における語の働き（機能）をシステムに教えます。

原文における語の働きを分類したものが品詞です。動詞ではさらに動詞パタンなどの詳細情報が必要となってきました。同時に、動詞では主語や目的語に来るものについて意味分類などの制約条件を教えます。

3) 訳文を構成するための訳語をシステムに教えます。

原文の語の持つ語義、さらに、それに対する訳語をシステムに知らせます。語義は、通常、識別名を持っていますが、中間言語の語彙（＝概念）が設定されているシステムでは、中間言語の語彙（＝概念）となります。また、各語義の情報として意味分類などをシステムに教えます。

4) 訳文の組立方をシステムに教えます。

訳語を用いて訳文を組み立てる際の訳語の機能を分類したものが訳語の品詞です。訳語の品詞は、通常、原語の品詞と一致します。さらに、原語の場合と同様、動詞では動詞パタンなどをシステムに教えます。また、最終的な表現を作り出す段階（形態素生成）では活用などの情報をシステムに教えます。

5) 選択のための情報をシステムに教えます。

複数の品詞を持つ語については、文法的な条件で1つに決まらなかった場合、デフォルトとして選択すべき品詞をシステムに教えます。また、多義語について、デフォルトとして選択すべき語義（訳語）をシステムに教えます。より進んだ情報としては、訳し分けを可能にする共起情報をシステムに知らせます。共起情報というのは、「トランプを切る」の「切る」の訳語を選択する際に、「トランプ」を手がかりとして、「shuffle」を選ぶという類の情報です。

2. 人間用の辞書との比較

1) 訳語

英日翻訳を例にすると、人間用の辞典は、英語の単語の日本語による解説になっているものがあります。これに対し、機械翻訳用の辞書の訳語は訳文を構成できるものです。

2) 原文解析・訳文生成を行うための情報

原文解析や訳文生成を行うために、通常、人間があまり意識しない情報が必要です。原文解析では、どの部分が主語でどれが動詞でどれが目的語かといった原文の構造を決定しなければなりません。そのために、各語が文の中でどのような役割を果たしているかをシステムは判断しなければなりません。そのための基となる各語の情報を辞書はシステムに提供します。逆に訳文生成では、訳語を組み合わせ、

訳文を作り出さなくてはなりません。そのために、各語を文の組み立てにどのように使うかという情報を辞書はシステムに提供します。これらの情報の種類や数は原文解析や訳文生成の方式に依存します。通常、設定が必要な情報には、まず品詞があり、さらに、名詞であれば意味分類、動詞であれば活用、動詞ボタンなどがあります。動詞ボタンというのは英語であれば、SV/SVO/SVC/…といった文を構成する型のことです。

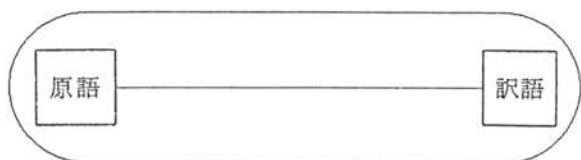
3. 辞書の構成

機械翻訳用の辞書の構成は、翻訳の過程で必要とする情報を何種類の辞書で構成するかによって次のように分類することができます。

1) 一体型

原言語から訳言語までが1種類の辞書で構成されます。

この辞書を構成する基本的な単位の持つ情報は、原言語の語とそれに対応する訳語です。われわれが通常用いる英和辞典や和英辞典はこのような辞書と同じ構成と言えます。



一体型は、一般利用者にとって、作成したり、管理したりすることが比較的容易であるという利点を持っています。

2) 3フェーズ型

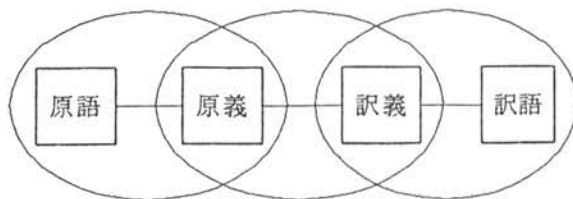
原言語辞書、対訳辞書、訳言語辞書の3種類の辞書で構成されます。

原言語辞書と訳言語辞書とは単言語の辞書になっており、対訳辞書に原言語と訳言語との語（語義）の対応関係の情報を格納します。

日英翻訳用の辞書を既に持っており、日仏翻訳用の辞書を新たに作る場合、日本語の原言語はそのま

ま利用でき日仏の対訳辞書と仏語の訳言語辞書のみを作れば済むという利点を持っています。また、逆に、仏英翻訳用の辞書を新たに作る場合には、英語の訳言語辞書を利用できます。

さらに、設計次第で、日本語の原言語辞書を、日本語の訳言語辞書としても利用できる可能性を持っています。

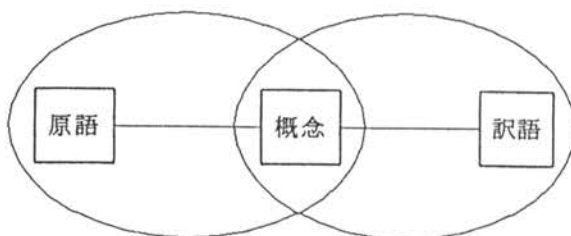


原言語辞書 対訳辞書 訳言語辞書

3) 2フェーズ型

原言語辞書、訳言語辞書の2種類の辞書から構成されます。

この場合には、明確に定義された概念のセットが設定されていることが必要です。各概念には、識別のためのコード、概念名等がアサインされます。原言語辞書は、原言語の語を中間言語の語に変換するための情報を与えます。また訳言語辞書は、中間言語の語を訳言語の表現に変換するための情報を与えます。



原言語辞書 訳言語辞書

日英翻訳用の辞書を既に持っており、日仏翻訳用の辞書を新たに作る場合、日本語の原言語はそのまま利用でき仏語の訳言語辞書のみを作れば済むという利点を持っています。また、逆に、仏英翻訳用の辞書を新たに作る場合には、英語の訳言語辞書を利

用できます。

さらに、設計次第で、日本語の原言語辞書を、日本語の訳言語辞書としても利用できる可能性を持っています。

辞書を作成して行く場合、膨大な概念セットの中から適切な概念を選択してアサインしなければなりません。そのためには、概念全体を分類整理したシソーラス（概念体系などと呼ばれます）が必要になってきます。このような概念セットは、将来、より高度な翻訳を行う際に必要になると予想される膨大な知識（百科辞典的な知識）とリンクされる可能性を開くものです。

4. 辞書の種類

辞書は登録されている語の種類によって、3種類に分けられます。

1) 基本語辞書

国語辞典に載っているような、分野にあまり依存しない汎用の語が登録されています。その語数は機械翻訳システムによって異なりますが通常、2万～20万程度です。基本語辞書に含まれている語の多くは複数の訳語を持っており、通常、デフォルトの優先順位がついています。

基本語辞書は、翻訳を実行する際、必ず使用されますが、通常、優先順位は最も低く設定されます。

2) 専門用語辞書

専門用語が登録されており、分野単位に別れているのが普通です。用意されている分野の数や分野毎の登録語数、また、登録されている語のセットの偏りはシステムによって、基本語よりも、大きな開きがあります。一般的に、基本語よりもより多くの語数を必要とします。

翻訳で使用する際には、基本語辞書よりは優先順位を高く、ユーザ辞書よりは優先順位を低く設定するのが普通です。また、システムによっては複数の専門用語辞書を同時に使用できます。

専門用語というものは日々増加しており、ローカルな造語もしばしば行なわれています。ですから、安定した訳語の決っていないものも存在します。専門用語辞書を導入するに際しては、自分達の翻訳したいと思っている文書をいくつか試してみて、ヒッ

ト率をある程度知っておいた方がよいと思います。

3) ユーザ辞書

ユーザの登録した語を格納します。その多くは専門語ですが、基本語でも、新規の訳語を設定する場合があります。また、基本語からなる複合語を登録する場合もあります。

翻訳で使用する際には、優先順位を最も高く設定するのが普通です。また、システムによっては複数のユーザ辞書を同時に使用できます。

5. ユーザ辞書の重要性

機械翻訳システムを本格的に利用して行こうとするならば、必ず、ユーザ辞書の作成という作業がともなってきます。また、ユーザ辞書を作成し、利用することによって機械翻訳システム導入の効果をより一層高めることができます。ユーザ辞書を作成して行くことによって、機械翻訳導入の最大効果の1つである用語の統一が促進されるからです。専門用語辞書を購入することによってユーザ辞書を作成する手間を軽減することはできますが、基本的な事情は変わりません。

辞書は共有化が比較的容易な資産です。辞書のファイル形式や情報の種類は機械翻訳システムによって異なります。しかし、辞書の持っている本質的な価値は機械翻訳システムにあまり依存しません。というのは、辞書の最も重要な情報は見出し語と訳語だからです。したがって、将来、他のシステムに乗り換えた場合でも、これまで作ってきた辞書が全く無駄になってしまうということはありません。

6. 辞書登録

辞書登録に必要な情報は、システムに依存します。また、同一システムにおいても、辞書登録の際に必要な情報は、語の品詞によって異なります。通常、最も情報が少なくて済むのは名詞です。逆に、最も多くの情報を必要とするのが動詞です。したがって辞書に登録を行う場合、最も容易なのが名詞です。

組織的に辞書登録を行う場合には、辞書登録の方法・手順・担当者を品詞別に考えた方がよいでしょう。特に、名詞とそれ以外の品詞は扱いを別にした方がよいと思います。名詞については、大量に登録

する場合、通常、効率よく登録する工夫が可能です。

辞書登録の際に注意すべき点

1) 見出し語

ユーザ辞書を作って行く際には、どの範囲を1語と捕らえて辞書に登録するのがよいかを考える必要があります。複合語を積極的に辞書に登録すると翻訳の品質に対してよい効果が得られます。これは、原文解析の際の構文構造のあいまい性を減少させると同時に、訳語の適切なアサインを可能にします。複合語をユーザ辞書に登録することは、原文を短くする前編集と並んで、機械翻訳の翻訳精度を高める有効な手段の1つです。

できる限り名詞の単位で辞書に登録するというのは、ユーザ辞書を作って行く際の1つのコツです。可能ならば、形容詞なども単独で登録せず、より大きく、複合名詞の一部として登録します。このようにすると辞書に登録しなければならない語の数は増えてしまいますが、原文解析の成功率を高めると同時に、用語の統一という面でよい効果が得られます。

2) 訳語

ユーザ辞書を作成する際に最も慎重に行う必要があるのは「訳語を何にするか」ということです。実際、登録を行う際に最も時間がかかり、難しいのが訳語の設定です。

英日翻訳を例にすると、人間用の辞典は、英語の単語の日本語による解説になっているものがあり、機械翻訳用の辞書の訳語としてはそのまま使えない場合があります。機械翻訳用の辞書の訳語は、訳文の一部となるということを意識して決めなければなりません。例えば、人間用の辞書では「osier」=>「やなぎの一種」となっていますが、このままでは訳文の一部として使うには不適当です。機械翻訳用の辞書の訳語としては「オジャ」または単に「やなぎ」とする工夫が必要です。

7. 訳語の優先順位

現在の機械翻訳システムが苦手としていることの1つが、適切な訳語を選択するということです。つまり、適切な訳語が辞書に含まれていてもそれをうまく選択することは現在の機械翻訳システムにとっ

てなかなか困難な問題です。したがって、あらかじめ設定した優先順位が重要となります。通常、ユーザ辞書に登録する語は訳語を1つしか持ちませんからこの問題は起こりません。基本語辞書に入っている語についてはリリースされた時点でデフォルトの優先順位がついていますが、基本語の訳語の優先順位はユーザの翻訳対象文書や翻訳目的によって異なってきます。したがって、使いながら訳語の優先順位を調整して行くことが必要です。

8. 辞書の管理

「辞書を分割して管理するのと1つにまとめて管理するのとどちらがよいか」というのは1つの問題です。単純に言えば、できる限りまとめて統合管理の方が管理という点では有利です。辞書が複数に別れている場合、ある見出しについてそれぞれの辞書をチェックしてみる必要が生じます。辞書の作り方、管理の仕方によっては同じ内容の見出し語と辞書情報がいくつもの辞書に重複して格納されているという事態も生じかねません。では、ユーザの持っている辞書を全て1つの辞書にまとめた方が必ずよいかというとそうとも言えない場合があります。同じ見出し語について訳語が違っているような辞書どうしは一緒にしない方がよいと思います。例えば専門用語では、分野がちがう辞書の間でこのようなことが起こります。このような場合には分野毎に辞書を分けておいた方が有利です。

9. 最後に

辞書というのは単純に言えば語をキーとする知識の集まりです。辞書の1つの発展方向は、1つの語をキーとする知識から、2つまたは複数の語をキーとする知識へと向かうと予想されます。1つの語をキーとする場合、基本語の数は単純には数万のオーダーですから、知識モジュールの数も数万のオーダーになります。。2つの語をキーとする場合、有効なキーの組合せは百万以上のオーダーになると予想されます。この位のオーダーになると常識のような知識を入れられるようになるかもしれません。

(〔財〕国際情報化協力センター 依田 透)

日英/英日機械翻訳システム H I C A T S (ハイキャッツ)

株式会社 日立製作所

日立機械翻訳システム「HICATS」は、日英・英日の双方向の翻訳機能を結合した総合的な翻訳支援システムです。

日立は、1970年代後半から機械翻訳システムの研究を進めており、1986年に「HICATS」として販売を始めました。当初メインフレーム上で稼働していましたが、1989年にワークステーション上で稼働するシステムの販売を始めました。その後も、各種機能の追加、文法データ・辞書データの増補・改良を重ねて、使いやすく、翻訳品質の良いシステムを目指して開発してきました。

この度、高性能ワークステーション「日立クリエイティブステーション 3050RXグループ」の開発に伴い、同ワークステーション上で稼働するシステムを開発しましたので、ご紹介致します。

[特 長]

★ 使いやすい操作

業界標準のOSF/Motifに準拠した X Window に対応しています。

マルチウィンドウ上で、機械翻訳結果の編集・校正作業が行えます。また、ポップアップメニューの採用、マウスを中心としたマンマシンインタフェースにより、単語選択・修正、コマンド選択、範囲指定、スクロールなどが簡単に行えます。

★ 使いやすい用語辞書作成機能

ユーザ固有の訳語をもつ単語は、ユーザ辞書を作成するとより良い翻訳結果が得られます。ユーザの辞書作成を支援するため、操作性の良い「辞書保守機能」をサポートしています。これにより、各ユーザの翻訳知識を辞書に登録することができます。

また、英日翻訳では、文書中の未知語、頻出する語、及び連語を検出し、一括処理でユーザ辞書に登録できます。

★ 簡単にユーザ辞書が作成できる学習機能

日英翻訳では、ユーザが修正した語を自動的に見つけだし、ユーザ辞書に登録します。また、後編集した動詞と、その動詞の目的語などの名詞を組みとしてユーザ辞書に登録します。

英日翻訳では、訳語を選択してユーザ辞書に登録します。

学習機能を使うと、ユーザ辞書への登録が簡単になります。

★ 高度な日本文診断機能

例えば、「HICATS/JEによって翻訳した論文とマニュアルを編集する。」という文では「HICATS/JEによって翻訳した」のは「論文」だけなのか「論文とマニュアル」なのかははっきりしません。また、HICATS/JEによって「翻訳する」なのか「編集する」なのかははっきりせず、システムはどちらの意味に解釈したら良いか判断できません。このような日本文のもつあいまい性をチェックするのが「日本文診断機能」です。この他にも単語の意味のあいまい性や複雑な文の指摘など、15種類の診断項目を備えています。

★ 訳出オプションにより、訳文の文型などを選択

例えば、主語のない日本文は通常受け身形の英文に訳されます。しかし、ドキュメントの種類によっては、命令文で訳す場合や主語"you"を補って訳した方がよい場合があります。訳出オプションを選択して翻訳することにより、ユーザは訳し方を指定でき、後編集作業を軽減できます。

このような訳出オプションを、23種類備えています。

★ 共起関係を利用した訳し分け

単語や単語のもつ属性の同時出現性(共起関係)を知識として利用することにより、原文を正確に解析します。それと同時に、各単語の適切な訳語を選択して、わかりやすい訳文を生成します。

●日英翻訳の例(「上げる」の訳し分け)

その車はスピードを上げた。

---> The car increased speed.

彼は、手を上げた。

---> He raised a hand.

●英日翻訳の例("bring"の語義決定)

He brought me the flowers on the table.

---> 彼は私にテーブルの上の花を持ってきた。

He will bring his daughter to the party tomorrow.

---> 彼は明日彼の娘をパーティーに連れてくるだろう。

★ 各種ファイルとの連動が可能

日立ワークステーション3050, 3050RX上で稼働するOAソフトウェア「OFIS-EXシリーズ」で作成した文書を、直接翻訳できます。これは、図入りの文書でも可能であり、図をそのまま活かした訳文の文書が作成できます。

また、UNIX, MS-DOSのファイルを入力し、翻訳結果をUNIX, MS-DOSのファイルに出力できます。

これらの機能により、入力の手間が少なくなり、編集作業も楽になります。

★ 電子辞書機能

(株)研究社 新和英中辞典、新英和中辞典を収録した電子辞書を提供しています。

辞書を参照したり、辞書の内容を二か国語エディタ上に複写したりでき、後編集の強力なサポートになります。

また、英文のスペルチェック機能もサポートしており、英文入力時のスペルミスを簡単にチェックし、修正できます。

★ 分野別専門用語辞書

専門的な用語を翻訳するために、8分野の分野別専門用語辞書を提供しています。

①情報処理 ②電気・電子・通信

③土木・建築 ④自動車・鉄道・船舶・航空

⑤自然科学 ⑥生物⑦機械⑧化学工業

[性能と価格]

ハードウェア、ソフトウェアを各種取り揃えており、ユーザの利用形態により選んで頂けます。

一例を紹介致します。

・3050RX-200 使用時

価格 約 350万円 ~

翻訳速度 日英:約 44,000語/時 73P(600語/P)

英日:約 57,000語/時 228P(250語/P)

・3050RX-330 使用時

価格 約 770万円 ~

翻訳速度 日英:約110,000語/時 183P(600語/P)

英日:約140,000語/時 560P(250語/P)

[他社商品名称に対する表示]

(1) OSF/Motifは、Open Software Foundation, Inc. の商標です。

(2) X Window Systemは、米国 the Massachusetts Institute of Technologyが開発したソフトウェアです。

(3) UNIXは、X/Open Company Limitedがライセンスしている米国ならびに他の国における登録商標です。

(4) MS-DOSは、米国Microsoft Corp. の登録商標です。

[お問い合わせ先]

株式会社 日立製作所

コンピュータ事業本部 製品企画本部

TEL (03)3763-2411 (大代表)

(内線 3280)

日英機械翻訳ソフト

T r a n s L a n d M a c 版

ブラザー工業株式会社

1. Macintoshに対応

当社は長年に渡り日英機械翻訳ソフトを研究開発し、社内でも利用してきました。そして、その成果として昨年（1993年）6月に、パーソナルユースを対象にした、PC98版日英機械翻訳ソフト「T r a n s L a n d」（MS-DOS対応）を発売しました。この製品は翻訳精度の良さと、お求め易い低価格によりユーザーから高い評価を受けています。

パソコンのユーザーインターフェイスは、グラフィカルなものに移りつつあります。当社は翻訳ソフトをより使いやすくするために、多様なOSに移植しています。

今回、その一環としてMacintoshへの移植・製品化をしましたので、ご紹介します。

2. 特長

2.1. 翻訳精度の改良

好評を頂いているPC98版機械翻訳ソフト（MS-DOS対応）に対して主に以下の改良を行いました。

- ・助動詞処理の強化
- ・助詞「は」があるときの英文の向上
- ・命令文の処理の強化
- ・名詞句の認識の精度向上

2.2. 翻訳設定

より望みの翻訳結果が出せるように以下の翻訳設定ができます。

- ・主語が無い時のデフォルト値
- ・冠詞のデフォルト値
- ・主語が無い時の英文の形
（受け身文か、命令文か）
- ・ユーザー辞書の使用の有無とその区別表示
- ・翻訳時間制限の設定

2.3. ユーザーインターフェイスの向上

・対訳画面

日本語と英語を左右対照で表示しますので日本語の修正、再翻訳、英文の編集がスムーズに行えます。また、最大9999文まで一括翻訳または選択した範囲の翻訳ができます。

・ユーザー辞書登録

独立のウインド画面上でユーザー辞書の登録、修正が自由に行えます。また、マウスカーソルを入力項目の上へ移動するだけでオンラインヘルプが表示されます。

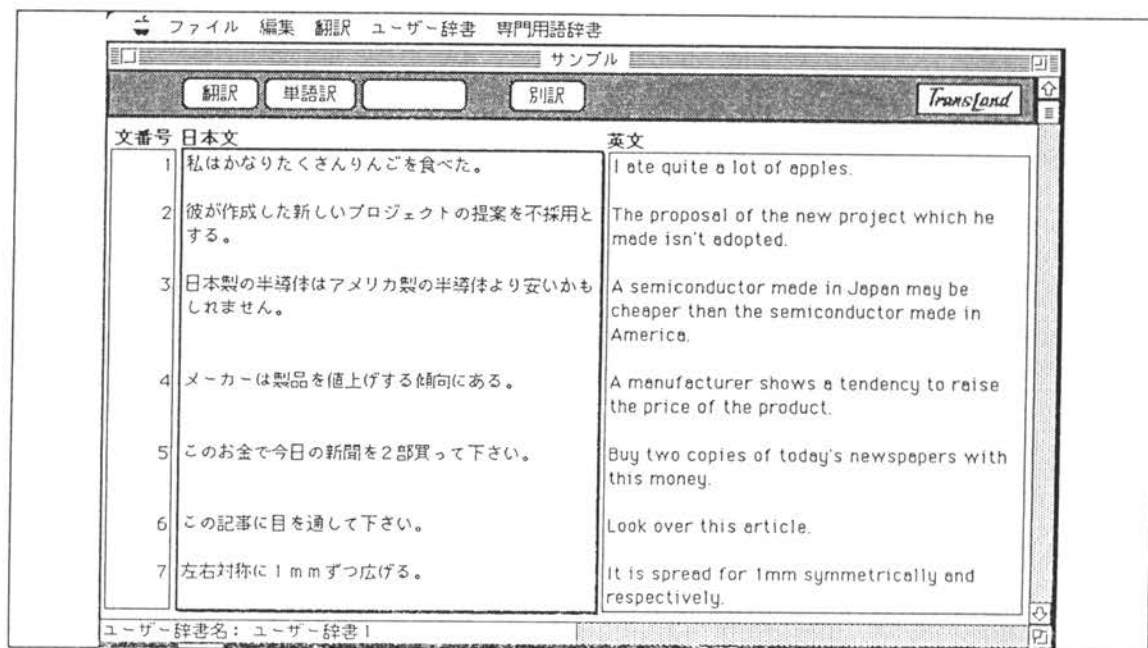
問い合わせ先

ブラザー販売株式会社 情報機器事業部

高 峰 052-263-5855

価 格 69,800円

動作環境	
対応機種	Macintosh CPU68030以上、Power Macintosh
O S	漢字 Talk7
メモリ	空きメモリ10MB以上（仮想メモリ使用可）
必要HDD	20MB以上（インストール時）



編集・翻訳画面例

翻訳設定

省略時の主語の選択 冠詞のデフォルト値

主語がないときの処理 ☐ 命令形 ☒ 受身形

ユーザー辞書 ☐ 使用しない ☒ 使用する

翻訳結果のユーザー辞書単語(※.....※) ☐ マークする ☒ マークしない

翻訳時間制限 (0~120秒) 秒を越えたと自動的に単語訳に切替

翻訳設定画面

ユーザー辞書 (名詞)

日本語

英語

主名詞 ☐ 「～する」で動詞になる

意味情報 ☒ 人・組織 ☐ 動植物 ☐ 部品 ☐ 道具 ☐ 食べ物
☐ 乗り物 ☐ 書類 ☐ 材料 ☐ 金銭 ☐ 発生物
☐ 現象 ☐ 行為 ☐ 単位 ☐ 性質 ☐ 知的産物
☐ 時間 ☐ 地名 ☒ 場所 ☐ イベント

英訳の冠詞 ☐ 制限なし ☐ the ☐ a ☒ につけない

可算/不可算 ☐ 可算 ☒ 不可算 ☐ 常に複数形

複数形のタイプ ☒ ☐

複数形

先頭の発音 ☒ 子音 ☐ 母音

人、組織、企業名をあらわします。
 「例」技術者、会社、委員会、業者、人事部

ユーザー辞書登録・修正画面

展示ショウ

データショウ'94

CALS Japan'94

会 期 1994年9月26日～29日（毎日10時～17時）
 会 場 東京国際見本市会場（晴海）B、C、南館
 入場無料

CALSとは情報社会の社会インフラ作り運動・世界
 共通のインフラ作り運動・政治、行政、経済
 をデジタル化する運動（東館で開催）

主 催 （社）日本電子工業振興協会・通信機械工業会 後 援 通産省 郵政省 NTT KDD

英日・日英機械翻訳システム 「ATLAS／Win」

富士通株式会社

1. はじめに

富士通株式会社では、Windows環境のパソコンで動作する英日・日英機械翻訳システム「ATLAS／Win V1.0」を1994年5月から発売を開始した。

当システムは、以下の特長を持っている。

- 1 高品質な翻訳
- 2 業務に対応した豊富な辞書
- 3 様々なニーズに対応した翻訳形態
- 4 OASYS／Win との連携
- 5 翻訳精度を高めるチューニング機構
- 6 Windows3.1で動作するマルチベンダ対応

2. 製品体系

ATLAS／Win は以下の4製品からなる。

- 1 「ATLAS／Win EJ V1.0」 英日機械翻訳システム
- 2 「ATLAS／Win JE V1.0」 日英機械翻訳システム
- 3 「ATLAS／Win 双方向翻訳オプション V1.0」
「ATLAS／Win EJ」または「ATLAS／Win JE」に対して、双方向翻訳を可能にする。
- 4 「ATLAS／Win 専門用語辞書 V1.0」
業種別24分野、合計120万語の辞書

3. 高品質な訳を追求するATLAS／Win

富士通では業界に先駆け1984年に機械翻訳市場に参入して以来、オープン化、ダウンサイジングを実現しながらも、常に翻訳品質の改善に取り組んで来た。

1. 意味処理方式

正確で自然な訳文を出力するため、意味処理方式を採用している。この方法は原文の意味を理解した上で、最適な訳語を正しい構文に出す方法である。構文パターンを置き換える構文トランスファ方式よりも、一歩進んでいる。以下に意味処理方式の具体例をあげる。

原文：日本語を英語に計算機で翻訳する。

訳文：The machine translates Japanese into English.

この原文には主語がないが、ATLAS／Win は「計算

機」を主語として意味解釈する。

原文：日本語を英語に手で翻訳する。

訳文：(*S) translates Japanese into English by hand.

同じ構文であっても、この場合、ATLAS／Win は「手」を方法と解釈して、「by hand」と副詞句で翻訳する。

2. 13万語の基本辞書

13万語を基本用語として標準装備している。この基本辞書には翻訳に必要な単語の読み、表記、意味カテゴリー、格パターン、訳語等の情報が格納されている。

3. 専門用語辞書

情報処理、電気・電子、医学、金融・経済、ビジネス分野等24種類、120万語の専門用語辞書を用意している。これらの専門用語辞書を業務内容に併せて使い分ければ、きめ細かい翻訳ができる。

原文：Open windows.

基本辞書のみ：「窓を開けて下さい」

情報処理用語辞書利用：

「ウィンドウをオープンして下さい」

また、これらの専門用語辞書は、英日・日英どちらの翻訳にも利用できる。

4. 翻訳精度を高めるチューニング機構

・利用者辞書の作成・単語登録 利用者の表現にあわせた形で単語を登録することができる。利用者辞書は、英日・日英どちらの翻訳にも利用できる。

・訳文スタイルの選択

目的に適した訳文スタイルの選択が可能

例：文体指定（英日）

原文：I go to Osaka.

訳文：である調「私は大阪に行く。」

です・ます調「私は大阪に行きます。」

例：主語のない平叙文の訳（日英）

原文：部品を取り外す。

訳文：能動態「(*S) detaches parts.」

受動態「Parts are detached.」

命令形「Detach parts.」

4. ATLAS/Win の翻訳形態

ATLAS/Win は、用途に応じて選択できる3つの翻訳形態を用意している。

1 対話翻訳

原文ファイル、訳文ファイルとも、画面の任意の場所に自由に開くことが可能。机上で辞書や例文を参照するように上下左右にファイルを開き、原文を入力・参照しながら訳文を作成できる。海外に送る通達文などを作成しながら翻訳するような時に使うと便利である。

2 一括翻訳

ファイルを指定するだけでMS-DOSテキストの内容を一括翻訳。またOASYS/Win で作成した文書をそのまま翻訳する。さらに英文の技術文書など紙に印刷

された大量資料の場合にはOCR を利用して翻訳することが可能である。

3 クリップボード翻訳

クリップボードのテキストを翻訳する。Windows 上の他のアプリケーションで作成した文をその場で翻訳できる。OASYS/Win の画面からワンタッチで翻訳させることも可能である。

〔お問い合わせ先〕

富士通株式会社 システム販売推進部

〒東京都千代田区丸の内1-6-1

(丸の内センタービル内)

TEL (03) 3216-3211 (代)

表1. ATLAS/Win の仕様

項目	主な機能概要
翻訳	英日翻訳 (ATLAS/Win EJ), 日英翻訳 (ATLAS/Win JE), 対話翻訳/一括翻訳/クリップボード翻訳, 文の翻訳/単語の翻訳
辞書	辞書作成/削除/切り換え, 単語検索/登録/削除
チューニング	訳文スタイル, 翻訳時間設定, 一文認識調整
ファイル変換	和文抽出, 英文抽出, 左右対訳変換
入力	MS-DOSテキスト/OASYS 2文書
出力	MS-DOSテキスト/OASYS 2文書 訳文のみ/上下対訳/左右対訳

動作環境

対応機種	日本語 Microsoft® Windows™ V3.1 が動作する機種 (i486™以上のCPU搭載機種推奨)
対応OS	日本語 Microsoft® Windows™ V3.1
必要メモリ	10MB以上
必要ディスク	35MB以上 (OS除く)

価格

ATLAS/Win EJ	19.8万円
ATLAS/Win JE	19.8万円
ATLAS/Win 双方向翻訳オプション	16.8万円

専門用語辞書一覧

辞書名	語数	価格	辞書名	語数	価格
(1)情報処理	96,000語	15万円	(13)農林水産	21,000語	5万円
(2)電気・電子	33,000語	10万円	(14)生物	21,000語	5万円
(3)物理・原子力	54,000語	15万円	(15)医学(生化学)	47,000語	10万円
(4)機械	57,000語	15万円	(16)医学(薬学)	21,000語	5万円
(5)工業化学	63,000語	15万円	(17)医学(解剖学)	48,000語	10万円
(6)プラント	53,000語	15万円	(18)医学(疾患・症状)	120,000語	15万円
(7)土木・建築	31,000語	10万円	(19)医学(精神医学)	25,000語	5万円
(8)金属	13,000語	5万円	(20)医学(医療機器)	48,000語	10万円
(9)地学	30,000語	10万円	(21)金融・経済	36,000語	10万円
(10)輸送	36,000語	10万円	(22)法律	6,000語	5万円
(11)自動車	32,000語	10万円	(23)ビジネス	104,000語	10万円
(12)軍事	24,000語	5万円	(24)住所・氏名	172,000語	10万円

① MS-DOS, Microsoftは米国Microsoft Corporationの登録商標です。

② Windowsは米国Microsoft Corporationの商標です。

③ i486は米国インテル社の商標です。

④ その他の製品は各社の商標または登録商標です。

英日／日英機械翻訳システム

PC-Transter for Windows

株式会社 ノ ヴ ァ

弊社は、本格的な機械翻訳システムをパソコン・ソフトとして提供、現在下記の製品群を中心に多くのユーザからの支持をいただいております。

- 英日機械翻訳システム PC-Transter/ej
Windows版、MAC版、DOS版(PC98、DOS/V)
- 日英機械翻訳システム PC-Transter/je
Windows版、MAC版、DOS版(PC98、DOS/V)
- 特許文対応英日機械翻訳システム
SuperTranster/ej DOS版(PC98、DOS/V)
- 上記システム対応専門語辞書
英日 18分野 日英 15分野

このうちWindows版は、本年6月に発売を開始した新製品で、従来のDOS、MAC版に新たな翻訳支援機能を搭載したものです。

今回はこのWindows版の新機能を中心に、以下に紹介をさせていただきます。

1) ステップ翻訳方式の採用

ステップ翻訳方式は、当社が開発したもので、修飾関係の難解な、また複雑な文に対し、句・節のまとまりを指示し、簡単な操作でより正確な構文解釈を実現しようとするものです。これまでの日英システムでのサポートに加え、今回英日システムにもサポートしました。

〔使用例〕

(原文) It outlines the minimum hardware and software requirements necessary to install and use

Visual C++, してとらえていますが、ここでステップ翻訳を利用し、正しい翻訳結果を導きます。

- 1) the minimum hardware and software requirements を名詞句として部分訳
- 2) the minimum hardware and software requirements necessary to install and use

Visual C++ を名詞句として再度、部分訳
3) 1)、2)を踏まえた全体訳の実行

(ステップ翻訳結果)

それは、Visual C++を設置して、そして使用するために必要な最小限のハードウェアとソフトウェア必要条件を概説する。

2) シソーラス辞書の標準搭載

日英システムに米国Info Soft社のシソーラス辞書(15万語)を標準搭載しました。これは、システムによる翻訳結果をベースとして、基本機能としてサポートされている訳語選択のほかに類義語の選択を可能とするものです。

例えば「翻訳する」の第1訳語であるtranslate に対しシソーラス辞書の意味

To express in another language,
while retaining the original sense

を参照すると、

convert、decipher、decode、interpret、
transcribe、transliterate

などの単語が同義語として表示され、ユーザは訳文への採用が可能です。

3) 訳例反映機能

英日・日英ともに「訳例反映機能」をサポートしました。これは、例えばマニュアルの改定版などの翻訳に際し、原文が変更されている文についてのみ翻訳を実行するものです。

参照用のファイルは Transterで使用されている対訳ファイルで、ユーザが作成することも可能です。したがって、良く使用される文を意識したい場合に使用することも考えられます。

4) 複数ファイルの翻訳・編集

同時に複数のファイルをオープンし、それぞれの

ファイルを翻訳したり、編集することが可能になりました。（最大4ファイル）

5) バックグラウンド翻訳

複数のファイルを画面に表示することなくバックグラウンドで翻訳が可能になりました。当然のことながら、より高速な翻訳が可能です。

6) サーバー・クライアント方式の採用

本Windows版より、システムの構成をサーバー・クライアント方式に変更しました。翻訳の実行部分を“翻訳サーバー”、ユーザインタフェースを“クライアント”として分離したもので、あたかもネットワーク上のサーバー・クライアントのよう

な動作が、1パソコン上で実現しています。

これは、開発サイドの機能分担のみならず、今後翻訳機能了他アプリケーション・ソフト、たとえばワープロ、表計算などからでも容易に利用できることを指向したものです。

まことに簡単ですがWindows版の新機能紹介とさせていただきます。

株式会社ノヴァ・システム営業部

〒160 東京都新宿区荒木町23 鈴商ビル301

電 話 03(3351)3356

F A X 03(3351)5766

PC - Transe/ej for Windows - ERIC.EDT	
ファイル(F) 編集(E) 検索(S) 翻訳(T) 単語(W) 辞書(D) 設定(M) ヘルプ(H)	
1 Iaro,	太郎、
2 How are you doing?	元気かい？
3 I read a newspaper article about Japanese machine translation systems.	私は、日本の機械翻訳システムについて新聞記事を読みました。
4 I was surprised that there were more than ten commercial systems in Japan.	10以上の商用のシステムが日本にあることに私は驚きました。
5 The article mentioned a system named PC-Transe/ej which runs on Windows computers and was developed by Nova Inc.	その記事は、Windowsコンピュータの上で動き、そして株式会社ノヴァによって開発されたPC-Transe/ejと名づけられたシステムに触れました
6 Can you make contact with the company and get some information about the system?	あなたは、その会社と接触することができて、そしてそのシステムについてのいくつかの情報を得ることができますか？
7 If it is possible, test the system with the following sentences and send me the result.	それが可能であるならば、そのシステムを以下の文でテストして、そしてその結果を私に送ってください。
8 1.	1.
9 Time flies like an arrow.	光陰矢の如し。
10 2.	2.

機械翻訳ユーザ講習会

開催時期 平成6年11月25日（午後1時間30分～5時）
 開催場所 名古屋商工会議所 大ホール（名古屋市中区栄町2丁目10番地）
 対 象 名古屋商工会議所会員・未経験者、初心者対象
 主 催 名古屋商工会議所・アジア太平洋機械翻訳協会

英日機械翻訳システム 「翻訳工房」

九州松下電器株式会社

1. 『翻訳工房』 誕生

九州松下電器株式会社の英日機械翻訳システム『翻訳工房』は、単に翻訳処理をするだけでなく、英文原稿の読み取りから翻訳結果のドキュメントを仕上げるまでを考えた、トータルドキュメンテーションシステムです。

つまり、英文原稿の読み込み、翻訳から仕上げまで、翻訳にかかわる処理を、トータルで行なうことを考えたシステムなのです。

それでは、『翻訳工房』の主な機能をご紹介しますしょう。

2. 『翻訳工房』の機能

(1) レイアウト機能

『翻訳工房』では、図・数式・表などの入った英文原稿そのままのレイアウトで翻訳結果（和訳）のドキュメントを仕上げ、印刷することができます。

これは、「レイアウト機能」というもので、『翻訳工房』の最大の特長となっています。この機能を使うことにより、従来大変手間がかかっていた処理、すなわち文字部分は機械翻訳し、図表部分は切っのりづけする、といったようなムダな作業を省くことができるようになりました。このように、『翻訳工房』は、翻訳する機能だけではなく、最終的に和訳されたドキュメントを仕上げることを目的としていますので、トータルでの翻訳作業時間を大幅に削減することができます。（図1 参照）

(2) 高速・高品質翻訳

『翻訳工房』では、1時間に約1.6万語の翻訳処理速度を実現しました。（Sparc Station 10使用時）

また、翻訳方式として、もっとも文法記述力の高い

「木構造変換方式」を採用することにより、高品質の翻訳を実現しました。

急ぐところはより速く、大事なところはさらにじっくりと処理することにより、必要なドキュメントを効率よく完成することができます。

(3) らくらく操作

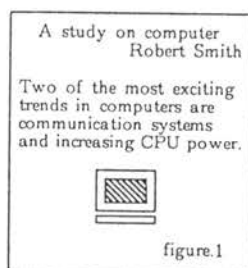
『翻訳工房』では、ユーザの使いやすさをまず第一に考えました。

すべての処理をマウス一つで行なえますから、とても簡単です。

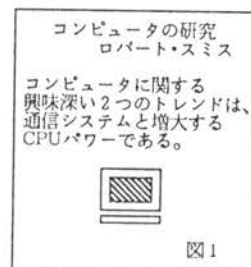
また、GUIに日本語Motifを採用し、分かりやすい画面でご利用いただけます。

さらに、翻訳後の訳語修正には、各種ツールを用意し、一層の使いやすさを追求しました。

『翻訳工房』のレイアウト機能



翻訳実行



【図1】

3. システム構成

最後に、『翻訳工房』のシステム構成をご紹介します。

(1) ハードウェア

本体	S p a r cワークステーション
CRT	カラーまたはモノクロCRT (19インチ程度)
スキャナ	I S 5 0 (ADF付き) リコー製
プリンタ	L B P 九松K X - P 4 6 3 0

(2) ソフトウェア

ユーザインターフェイス	M o t i f
O S	U N I X
ウィンドウシステム	Xウィンドウ
翻訳ソフト／辞書	
文字認識ソフト	

翻 訳 ソ フ ト	翻訳方式	意味トランスファ
	翻訳速度	16,000語／時
	基本辞書	7万語
		3,000～4万語／辞書
	専門用語辞書	情報工学, 電気工学, 機械工学, システム・制御工学 熱・応用熱工学, エネルギー工学, 運輸交通工学

文 字 認 識 ソ フ ト	原稿サイズ	A 4
	文字種	数字, 英字, 記号
	文字サイズ	6～60ポイント

なお、この製品に関するお問合せは、

長瀬産業株式会社

機器システム室／翻訳システムグループ

TEL : 03-3665-3060, 3396 (ダイヤルイン)

九州松下電器株式会社

担当：高橋

TEL : 092-477-1548 (ダイヤルイン)

協会活動報告

第4回通常総会開催

第4回通常総会は6月22日午後4時より機械振興会館で開催された。

会長挨拶、出欠の確認、議事録署名人の選出を行った後議事に入った。

第1号議案「93年度事業報告及び収支決算案」では①例文評価研究会、システム利用技術研究会、需要予測研究会などの研究活動状況、②研究成果発表会、MTサミット、翻訳現場見学会、専門家ヒアリングなどの普及啓蒙活動、③TMI'93、KB&KS、翻訳フェア94等の協賛事業、④AAMTジャーナル、MTニュースレターの発行状況や研究成果報告書、国際会議予稿集等文献頒布等の広報活動⑤IAMT総会やCICCとの連携業務など国際協力活動について等の報告が行われた。

また93年度決算報告では収入総額 3,039万、支出総額1,465万、差額1,574万は次期に繰り越すことが異議なく承認された。

第2号議案「94年度事業計画及び予算案」では①情報発信基盤の強化のために調査研究活動の積極化②ユーザ講習会、機械翻訳セミナーなどの普及啓蒙事業、他団体、学会との協賛事業などを一層推進する事業計画、及びこれを推進するための事業予算案(収入3,186万、支出3,186万)が異議なく承認された。

第3号議案「役員の改選」では理事4名の退任、留任12名、新任2名の新理事が選出された。また監事2名が留任で異議なく承認された。

「機械翻訳セミナー」開催

協会主催の「機械翻訳セミナー」は6月22日、芝公園の機械振興会館で110名の参加を得て開催された。

第1部は当協会の利用技術、文例評価、需要予測の各研究会の93年度の研究成果の発表を行った。

第2部はユーザ事例の発表である。ユーザ主導で開発された「Example-Based MT」(凸版印刷仁井主

第3期 協会役員

会長 長 尾 真 (京都大学教授)

副会長 野々内 隆 (日立製作所専務取締役)

副会長 小 谷 泰 造 (インターグループ社長)

理 事 浅 田 篤 (シャープ副社長)

吉 川 英 一 (日本電気取締役)

鈴 木 健 (日本電子工業振興協会専務)

田 中 穂 積 (東京工業大学教授)

岡 村 正 (東芝取締役)

西 岡 和比古 (CSK取締役)

長 澤 雅 浩 (松下電器産業取締役)

野 村 浩 郷 (九州工業大学教授)

渡 辺 学 (十印取締役)

平 栗 俊 男 (富士通専務取締役)

山 本 正 隆 (沖電気工業常務取締役)

監 事 勝 田 美保子 (日本翻訳連盟会長)

佐 藤 清 俊 (日本電子工業振興協会常務)

任研究員)のシステム解説、レイアウト、図式、表、数式を丸ごと翻訳する夢の工房を開発した長瀬産業(堂野前氏)前編集、文章推敲ソフトを駆使し、翻訳精度の改善や生産性向上に取り組むデータプロセス研究所(山下部長)機械翻訳システムを英語教育の教材として活用されている北海道教育大学(上山助教授)などバラエティに富んだ事例活用の報告に熱心な質疑応答が展開された。

第3部は「国際化の中で翻訳需要は、今…」と題し、日経国際ニュースセンター大澤社長の長年に及ぶ海外駐在や機械翻訳利用経験から翻訳の原点に立ち返ってユーモア溢れる講演があり、参加者に大きな感銘を与えた。

今回のセミナーには当協会会員はもとよりJTF会員、一般ユーザなど各方面からの参加があり、セミナーの合間をぬって活発な情報交換を行う光景が見受けられ、非常に好評を博する事が出来た。

新会員の紹介

金 城 功

竹 内 淑 子

川 口 不二夫

浜 口 宗 武

西 川 毅 一

文献頒布

研究成果報告書(94年版)

- 体 裁 B5版 230頁
内 容 ◆利用技術研究会
□パソコン翻訳ソフトのユーザ利用実態調査 □機械翻訳ビジネス
ユーザ訪問調査(編集ノウハウの収集) □機械翻訳の有効利用方法
◆例文評価研究会
□ユーザ例文の抽出とネイティブチェックの評価 □ユーザ例文の分
類と言い換え例 □科学技術例文の抽出と分類、言い換え例の作成
◆需要予測研究会
□90年の翻訳市場規模4700億、機械システム117億、□2000年、2010年
の市場規模は… □市場成長外部要因、内部要因を探る

頒 布 価 格 1部 3,000円(但し、個人会員1人1冊に限り特別価格1,500円にて提供)
申 込 先 アジア太平洋機械翻訳協会 事務局 (FAXでお申し込み下さい)

新刊書案内

こうすれば使える機械翻訳

- 体 裁 A 5判並製カバー装(264頁)
内 容 Part 1 機械翻訳の世界(基礎編)…機械翻訳の仕組み、周辺機能、技術展開他
Part 2 機械翻訳の有効活用法…翻訳プロセス、辞書の構成と活用法
Part 3 質問に答える… Q & A、ターミノロジー
Part 4 データファイル…年表、機種一覧、研究機関ほか
編・著者 編者・成 田 一(大阪大助教授) 著者・天野真家(東芝)村木一至(日電)
価 格 3,200円(消費税込)
申 込 先 バベル・プレス(〒101 東京都千代田区猿樂町2-2-3NSビル(〒03-3295-7128))

A A M T
ジャーナル
No. 8 (September/1994)

発 行 アジア太平洋機械翻訳協会
所 在 地 〒103東京都中央区日本橋小舟町4-10 大宮ビル
☎ 03-3664-5637/5638 FAX 03-3664-1352
E-Mail (PC-VAN) JRD08020 (NIFTY-SERVE) KGE00013
編集委員 野村浩郷(九州工業大) 杉山健司、亀井真一郎
事 務 局 星野禎男 西郷容子

