



AAMT

The Asia-Pacific Association for Machine Translation

Journal

No.3

May 1993

アジア太平洋機械翻訳協会

第3回通常総会・研究成果発表会

日 時 平成 5 年 6 月 25 日

場 所 京橋会館 (東京都中央区銀座1丁目)

※第3回通常総会 (午後4時～5時)

議 題	第1号議案	92年度事業報告及び収支決算について
	第2号議案	93年度事業計画及び収支予算について
	第3号議案	理事の選任について
	第4号議案	その他

※研究成果発表会 (午後1時30分～4時)

①制限言語研究会 (委員長・内田裕士)

翻訳精度を向上させ、翻訳コストを低減させるため文書記述面から一定の制限を加える小日本語、サーバ利用で自動前処理技術の糸口を探る。

②システム利用技術研究会 (委員長・牧野武則)

翻訳を工学としてとらえ、プロセス管理、ユーザ辞書、文書管理面から問題提起と対応策を検討。ユーザ実態調査からも問題解決策を模索する。

③システム評価研究会 (副委員長・横山晶一)

日本語の翻訳困難例文を摘出し、その言い換え例の検討結果を提起する。機械翻訳システム、自然言語処理の双方からのアプローチで効率的翻訳手法を模索する。

④講 評

目 次

・システム利用技術研究会活動報告	(1)	使用されている分野別専門用語	(5)
・人間翻訳と機械翻訳 (ヒアリング)	(6)		
・研究所紹介 ATR 翻訳電話通信研究所	(10)	NECTEC (12)	KAIST (14)
・新製品紹介 パンセ GV (沖電気)	(20)	E TRAN J (ソフトウェア技術)	(22)
・技術早分り 機械翻訳の辞書	(16)		
・ニュー ス 機械翻訳ニュース速報	(13)	トピックス	(23)
・お 知 ら せ 通常総会/研究成果発表会 (表2)		KB&KS シンポジウム (19)	TMI'93 (5)

機械翻訳システムの利用状況調査

機械翻訳利用技術研究会では、効果的なシステム利用方法を探ることを目的として、ユーザ事例のヒアリングと共に、ユーザにおける機械翻訳システムの利用状況の調査と結果の分析を行なっている。

機械翻訳システムがユーザに提供され始めて数年になる。あるユーザはその翻訳品質を問題にし「利用できない」といった否定的な評価を与え、他のユーザは利用方法を工夫し、「利用目的を限れば十分に利用できる」といった肯定的な見方をしており、システムに対する評価は多様である。

いずれにしろ、現在の機械翻訳システムは、ワープロのように誰でも簡単に使え、高品質の翻訳が得られるというシステムではなく、システムの能力を理解し、効果的な利用方法を工夫することが必要だと認識されている。この状況は近い将来のシステムの能力の向上を考慮しても、翻訳が創造を伴う限り大きくは変わらないだろう。

機械翻訳システムは、忘れることがない大容量記憶能力と高速演算能力をもったコンピュータの特長を活かして、人間には実現できない程の大量高速翻訳と翻訳の低コスト化を実現することを目的としている。

これらの能力を積極的に活かす利用法が、人間翻訳者とは異なった新たな文書翻訳手法として確立されるだろうと期待されている。

ユーザ調査では以上のような観点に立ち、現在ユーザが機械翻訳システムをどのように利用し、効果を上げているか、あるいは問題を抱えているかを調べることを目的とした。

アンケート調査では、まだ30件程度（4/23現在）と回答数が少なく、統計的なまとめはできないが、丁寧な回答が多く寄せられているので、ユーザの利用状況の概観を知るには充分であった。今後も継続して分析を続け報告するが、ここではその概要のみを報告することにする。

効果的な利用のポイント

機械翻訳は、文書を工学的に処理するツールであ

る。工学的にというのは、翻訳者による手工芸的、芸術的翻訳ではなく、高速に低コストで、ある品質の翻訳を行なうということである。この観点に立てば、限られた性能の機械翻訳システムの有効な利用方法は、効果的な翻訳プロセス管理と適切な翻訳品質管理に帰着する。

文書の翻訳のプロセスは、およそ、原文作成から機械翻訳のための前編集、機械翻訳、後編集、さらに品質の向上のためのリライト、文書編集、版下作成といったプロセスからなる。このプロセスにさらに辞書作成・維持が加わる。このプロセスにおいて、文書処理の全体がコンピュータ上でなされるにしても、前編集、後編集、リライトといった編集作業は人手に頼らなければならない。これら人手の介入が、機械翻訳を用いた翻訳プロセスにおける時間とコストのボトルネックとなっている。

翻訳品質の管理は、プロセスの管理と表裏一体である。前編集に時間をかけ、後編集の作業を簡略化するか、後編集に力を入れ品質を確保するかは判断が難しい。日本人が日英を使うときには前編集に、英日を使うときには後編集に力を入れるのが常識的だが、果たしてそうだろうか。また、コンピュータ資源、機械翻訳の言語資源を有効に利用するには「利用する人」の問題も重要である。機械翻訳システムを利用する際、どのような資質を持った人がどのような体制でシステムを運用するのが有効なシステム利用の側面でもある。

[ユーザ調査項目]

このユーザ調査では、ユーザでの機械翻訳の利用の実態、特に機械翻訳の運用、品質の管理方法について調査することと、機械翻訳システムの利用面で大きな要素であるユーザ辞書、さらに業界、学界での用語辞書の状況に関して調査を行なっている。これら調査の内容と目的は次のとおり。

1) システム利用状況

機械翻訳を用いた翻訳のプロセスと品質の管理に重点を置き、どのような資質の人がどのようににシ

テムを利用しているかを調査すると共に、翻訳技術者の育成についても調査する。その結果から、機械翻訳システムの効果的な利用法、さらには機械翻訳を積極的に利用していく環境について検討する。

2) ユーザ辞書、専門用語辞書の作成・維持

辞書の作成・管理はユーザにとってかなりの負担になるといわれている。ユーザが開発している機械翻訳のための辞書の開発・利用状況を調査する。辞書に係わる課題を整理し、ユーザ間での将来の辞書共有の可能性と、方法について検討する。

3) 業界、学界での用語辞書の実態

用語の統一や管理の重要性は改めて云うまでもない。各種工業会や協会、学会では、その分野における用語をまとめ、管理しているところが少なくない。そうした用語集、用語辞書の状況について調査し、機械翻訳システムでの利用に関して検討をする。

調査結果の概要

回答が寄せられた中で、機械翻訳システムを組織的に利用しているユーザと個人で利用しているユーザがほぼ半々であった。

個人で利用しているユーザには、機械翻訳システムの利用に極めて積極的なユーザが多くシステムに対する期待が高い。半面、実際の利用については大量の文書の翻訳を行なっている人はほとんどなく、レターや文献の翻訳に利用している程度である。

しかし、個人のこうした利用は、機械翻訳システムが広く受け入れられたときの一般的な利用形態であり個人利用の形態と利用の際の課題を整理するための有用な情報が含まれている。

機械翻訳システムを大量の文書翻訳に用い商用ベースで利用している、あるいは利用しようとしているユーザは、圧倒的に翻訳会社が多い。その次に多いのが製造会社のマニュアル制作部門、翻訳部門である。

[システム利用調査の分析]

機械翻訳システムの利用状況に関するアンケートの回答をまとめたところ、現在では次のような実態が把握できた。

機械翻訳システムを組織的に利用している大規模ユーザは機械翻訳システム利用可能端末一台、一月当たり、数百ページから千ページの文書を翻訳している。マニュアル、技術文書、特許などを対象にしているところが多く、品種を限定して効率を上げよ

うとしている。これに対して、個人等の小規模なユーザは数ページから数十ページである。マニュアルの素翻訳や文献の流し読みに使っているケースが多い。

前編集にかける時間はA 4一ページ当たり数分から15分程度で、後編集に費やす時間は一般に前編集より長く、10分から30分の範囲に入るものがほとんどである。前編集は日本語に対して行なわれるのが多く、主に、長文を短く区切ることや、省略された要素を補って翻訳しやすくすることがおもな作業である。

後編集では訳語のチェック・修正のほか、機械翻訳では出力できない訳文への書き換え（清書）がおもなものである。これらの翻訳周辺作業は、システムの性質によっても違いがあるが、実績のあるユーザでは、編集や辞書登録に関するノウハウを蓄積しているところが多いようである。中には、機械翻訳システムの能力をサポートする機能を独自に開発・運用しているユーザもある。これらのノウハウやサポートツールは、後述の作業者人材の育成にも直接寄与するものである。

機械翻訳システムで実際に作業する人は一般に、前編集者・後編集者・リライト・オペレータに分類される。小規模ユーザでは当然、同一の人が複数の作業を行なっているが、大規模ユーザでも必ずしも分業しているわけではない。また、その英語力も、スペシャリスト（英検1級程度）ばかりではないことがわかった。これは、システムを使いこなす上では英語力はある程度あれば十分で、むしろシステムの特性を知ることが重要であることを意味すると思われる。

作業者の育成はほとんどのユーザが社内で独自で行なっている。機械翻訳システムを使用した経験をもとにした意見や要望に中では、まだ翻訳品質の向上を求めるものが、特に日英システムのユーザに多い。しかし、利用実績のある大規模ユーザでは、品質に対する要望より、システムの操作性や他システム・他ソフトウェアとデータの流通を容易にすることを求める声が目立った。これは、現在のシステムの翻訳能力を理解したうえで、いかにシステム全体あるいは翻訳業務全体の能率を向上させるかに注目しているからであり、現時点でのユーザの積極的なアプローチであると云えよう。

[辞書調査の分析]

調査対象の辞書は、独自に作成したある程度の大きさを持つ辞書/用語集であるが、加えて業界/分野標準と考えられる辞書/用語集があれば記入をお願いした。これらの辞書/用語集の各々について、以下の情報を収集させていただいた。

1. 辞書の概要：

分野・使用目的、言語、見出し数。

2. 見出しについて：

見出し語の長さの制限、複合語や慣用表現の登録可否、表記上の基準、見出し語設定の基準等。

3. 訳語について：

訳語の長さの制限、複合語や慣用表現の登録可否、複数の訳語の登録可否、訳し分けのための情報、表記上の基準、訳語設定の基準等。

4. その他の辞書情報について：

同義語や反対語、品詞等の文法的情報、意味マーカ等の意味的情報、読み・発音、語義の説明例文等が付与されているか。

5. 電子化および利用形態：

電子化の有無、電子化の形態（機械翻訳の辞書かどうか、ファイル形式等）、容量、利用形態（スタンドアロン/ネットワークで共有）等。

6. 辞書の作成管理：

登録形態（一括/対話）、登録手続き、辞書管理ツール、辞書作成コスト（用語・対訳の選定コスト/登録コスト）、辞書の改訂頻度、公開（提供）可能性、現状の問題点等。

今回は回収途中であるので、その傾向のみ概観する。

分野としては、情報、通信、電子、機械、航空等が多く見られ、日英・英日がほぼ半々である。語数は2,000～6,000語程度が多い。機械翻訳のユーザ辞書では当然品詞や意味マーカ等が付与されているが、語義や例文はない。RDBやテキストファイルで、管理されている辞書（用語集）には、各訳語が使われるケース、同義語、語義説明、例文が付与されているものが少なからずある。

訳語の優先順位は使用頻度順という回答が多いが、訳語の最終的な選択には人間の判断が必要とされ、今後の課題である。

辞書の利用形態としては、ネットワークで共有している例も多い。望まれている辞書関連ツールには文の登録、文章からの定型表現の抽出と登録、登録履歴管理等が挙げられている。

調査内容が細かったことおよび作成コスト等の統計的データを正確に把握しているユーザが少なかったこと等により、十分な回答記入を得ることができない場合もあったが、提供可能と回答された辞書（数千語の大きさ）も数件見られた。これらを参考に、将来の辞書流通システムに向けて、テキスト形式での辞書内容の標準化等も考えていきたい。

調査の分析に当たって

翻訳のコストを劇的に低減させる手段は、機械翻訳において他にはない。翻訳のコストを桁違いに低減し、1/10以下にする可能性を機械翻訳システムは持っている。

調査結果から、十分に機械翻訳システムの能力を利用しているユーザは、まだ多くは見られない。効果的に機械翻訳システムを利用しているユーザに見られる傾向は、

- ・翻訳プロセスの管理、人の管理が効果的になされている。
 - ・機械翻訳に対して長期にわたる経験があり、習熟曲線が立ち上がっている。
 - ・機械翻訳システムの翻訳品質を見切り、その範囲で利用している。
 - ・機械翻訳技術者を自社で育成している。
- といった点が上げられ、さらにそれぞれのユーザで特徴のあるシステム利用の工夫がなされている。それらの工夫には、
- ・複数の利用者による辞書などの資源の共有環境を構築している。
 - ・前編集の効率化を図っている。
 - ・版下作成でのマニュアル作成過程を統合化し、電子化を進めている。
- など、様々な試みがなされている。

こうした経験に関する情報が、機械翻訳システム利用者間で共有されることで、一層の機械翻訳システムの普及がなされ、また機械翻訳システムによる翻訳について社会的な認知が一層進むものと考えている。

この調査・分析は現在も継続中であり、今までに回答を寄せられた方々にお礼を申し上げますと共に、この調査に対する皆様の一層のご協力を期待しています。御意見・コメントを機械翻訳利用技術委員会にお寄せ下さるようお願い致します。

牧野武則(東邦大) 熊野 明(東芝)
松平正樹(沖電気) 平田幹人(日電)

使用されている専門用語

分野	用語名称	規格	頒布元	関係団体
工作機械	日本機械学会刊行物	JIS, ISO 工業標準	日本規格協会	日本工作機械工業会
通信機械	電気通信用語辞典		コロナ社	通信機械工業会
	TTC 標準用語集	TTC	電信電話技術委員会(93/9)	全
自動車部品		JIS	規格協会	日本自動車部品工業会
		JASO	自動車技術会	全
プリント回路	プリント配線盤用語	JIS	規格協会	プリント回路学会
	全	JPCA	プリント回路工業会	全
自転車	自転車関係基準集(93/6完成)		工業会	日本自転車工業会
	自転車用語	JIS ISO	規格協会	全
包装機械	包装用語	JIS	規格協会	日本包装機械工業会
	全	日本標準商品分類	政府刊行物	全
電気	学術用語集		文部省	日本電気協会
	電気専門用語集		コロナ社	全
電線	電線用語	IEC ANSI	規格協会	日本電線工業会
タイヤ	タイヤ用語	SAE	米国自動車協会	日本自動車タイヤ協会
		JIS, JATMA安全基準	規格協会	全
燃焼機器	家庭用暖機器用語	JIS	規格協会	日本ガス石油機器工業会
保安用品	保安用品用語	JIS	規格協会	日本保安用品協会
	ガス検知警報機用語		産業用ガス検知警報機工業会	全
	検知管式ガス測定量用語		全	全
ヒートポンプ	空気調和衛生用語集		空気調和衛生工業会	ヒートポンプ技術開発C
	冷凍空調食品用語集		日本冷凍協会	全
	学術用語集機械工学編		日本機械学会	全
	マックロヒル科学技術用語大辞典		日刊工業新聞社	全
	科学技術25万語大辞典		インタプレス	全
石油化学	プラスチック用語	JIS	規格協会	石油化学工業会
塩ビ	全	JIS	規格協会	塩化ビニル工業協会
塗料	塗料用語	JIS	規格協会	日本塗料工業会
	塗料用語辞典(色材協会編)		技報堂出版	全
植物油	油化学用語		油化学協会	全
特許	工業所有権法令集		発明協会	日本特許協会
	米国特許実務用語辞典		AIPPI Japan	全

用語集・辞典

[(財)日本規格協会……東京都港区赤坂4-1-24 ☎03-3583-8002]

JIS 工業用語大辞典(第3版) 工業技術用語70,000余語を収録 29,000円(税込み)

土木・建築 機械 電気 自動車 鉄道 船舶 鉄鋼 非鉄金属 化学 繊維 鉱山 紙窯業 医療・安全 航空
情報処理 その他(原子力 医療放射線 品質管理 音響 照明 自動制御 生産管理等)

JIS 用語辞典(工業材料編) 鉄鋼 化学 繊維 パルプ及び紙 窯業 3,800語収録 3,700円(税込み)

JIS 用語辞典(情報技術編) ワープロ、CAD等一般機械 情報処理 2,875語収録 2,400円(税込み)

各種団体で作成されている専門用語

名 称	区 分	語 数	国 語	追 廃	仕 様	頒布元	関係団体
鉄道車両用語	用語集	1,000	日, 英	5年毎	JIS	規格協会	日本車両工業会
事務機械用語	用語集		日, 英		JIS	規格協会	日本事務機械工業会
石油辞典	辞書	4,000	日, 英		任意	丸 善	石油学会
AI 専門用語	用語集	2,000	日, 英	2~3年	任意	会員のみ	日本航空宇宙工業会
航空宇宙略語集	用語集	2,000	日, 英	2~3年	任意	会員のみ	全
芳香族製品・タール	用語集	160	日, 英	10年	JIS	規格協会	日本芳香族工業会
鉱業用語辞典	用語集		日, 英独			石炭協会	日本石炭協会
ガス用語辞典	辞書	12,000	日, 英	適宜	その他	ガス協会	日本ガス協会
バルブ用語辞典	用語集	1,500	日, 英		JIS	オーム社	日本バルブ工業会
標準 DCST	用語		日, 英	5年	団体	協 会	日本ダイカスト協会
アルミニウム表面処理	用語	249	日英独仏	5年	JIS	規格協会	軽金属製品協会
工業計器性能用語集	用語集	500	日, 英	5年	JEMIS	工 業 会	日本電気計測器工業会
環境計測技術用語	用語集	500	日, 英	5年		全	全
電気通信技術用語集	辞書	2,000	日, 英	3~4年	任意	協 会	電気通信協会
教育用コンピュータ	用語		日, 英		任意	アスキー	コンピュータ教育開発セ ンタ

「TMI '93」 2nd Announcement

DATE July / 14 ~ 16 / 1993

Venue Kyoto International Community House

(2-1 Torii-cho, Awataguchi, Sakyou-ku, Kyoto, 606 Japan)

Program (Tentative)	7/14	10.30	Opening	
		11.00	Session 1	Corpus-based Approach
		14.00	Session 2	Example-based Approach
		16.00	Session 3	Preference
		18.00	Welcome Reception	
	7/15	9.00	Session 4	Rational Approach
		11.00	Session 5	New Methodology
		14.00	Session 6	Project
		15.30	Panel Discussion	
	7/16	9.00	Session 7	Source Language Analysis
		11.00	Session 8	Evaluation Mechanism
		14.00	Session 9	Multi-lingual Corpus

Host Association for Natural Language Processing

Co-operate Asia-Pacific Association for Machine Translation (AAMT)

International Association for Machine Translation (IAMT)

人間翻訳と機械翻訳

TCC代表 安藤 進

人間翻訳と機械翻訳について、お話をしたいと思っています。資料としまして、雑誌「bit」（共立出版）に寄稿した記事のコピー（1992年12月号および1993年4月号）をお配りします。これは日本語と英語の翻訳で、人間翻訳と機械翻訳のレベルはどの程度違うのかというテーマでまとめたものです。

私が翻訳の世界に入りましたのは、今から約15年ほど前になります。最初の5年間は、富士通研究所で機械翻訳の研究開発をやっておりました。その後は、翻訳の発注側で日英翻訳のチェックを英語のネイティブスピーカー10人ほどの仲間といっしょにしておりました。それから、㈱十印翻訳部の部長として、主として英語から日本語に翻訳する仕事をしてきました。

そこで、本日は、これまでの経験を踏まえて私なりの問題を提起しまして、後はみなさんのご質問に答える形で話しを進めて行こうと考えています。

◆ [日英翻訳について]

最初に、日本語から英語に翻訳する場合についてお話しいたします。

人間翻訳の例としては、(社)日本翻訳連盟の機関誌で、会員を対象にした翻訳講座を3年ばかりやっております。その講座の応募答案を人間翻訳の例として採用しました。

機械翻訳(MT)の例としては、現在商用化しているMTシステムによる翻訳の例を採用しました。MTシステムの出力結果そのまま(一次結果)、辞書登録や原文を書き換えるなどの前編集をした後の翻訳結果(二次結果)を用意して、人間翻訳と比較してみました。

さて、応募答案の数は40、答案作成者の年齢は20代から50歳代、国籍は日本がほとんどですが、米国から応募された人も数人おりました。原文はプリンタのマニュアルです。応募者の得意分野は必ずしも技術分野ではありませんし、翻訳経験もさまざまです。翻訳者を志して勉強中の人もあります。

翻訳の品質を評価する方法については、いろいろ

な観点があると思いますが、ここでは実務で通用する翻訳、つまり商売として通用する翻訳という観点で評価して見ました。

応募答案を、A、B、C、Dの4つのレベルに分けて評価しました。Aレベルの答案というのは、そのまま使えるというものです。残念ながら、Aレベルのものはありませんでした。Bレベルは、冠詞、前置詞など、英語のネイティブスピーカーが原文を見ずに簡単に修正できるレベルです。Bレベルの答案は、4例、全体の10%でした。Cレベルは、原文を参照すれば比較的簡単な修正で通用するという意味です。このレベルが8例、全体の20%になります。

したがって、実務的に通用するというのは全体の30%位ということになります。

次に、機械翻訳の出力を評価しますと、一次結果は、利用価値はほとんどありませんでした。しかし二次結果はマアマアの出来でした。ちょっと主観的かもしれませんが40~50%の出来とでも言いましょうか、その位のレベルになっていました。

人間翻訳でも、機械翻訳の二次結果のレベルに達していない答案が全体の30%ぐらいありました。この結果からみると、人間翻訳が必ずしもよいというわけではないことがわかります。また、機械翻訳だからといって必ずしもまったくダメだというわけでもないと言えらると思います。

◆ [英日翻訳について]

英語から日本語に翻訳する場合について、お話しいたします。

実例としましては、私が翻訳会社におりました頃今から一年ほど前になりますが、実施したトライアルを紹介します。新聞広告を出し、翻訳者と機械翻訳の後編集者の募集をしたのです。全国から600人近い応募者がありました。トライアル問題の題材はコンピュータ、通信、半導体でした。

応募者に問題を送りましたところ、約半数の人がギブアップしてしまいました。つまり、問題の内容を見て自信をなくしてしまったと思われれます。一般に、翻訳者になりたいと考えている人の数は多いのですが、技術的なバックグラウンドを持っている人

が少ないということがわかります。それでも、300人近い人が答案を返送してきました。その答案をA、B、C、D（不合格）という4つにわけて採点しました。その結果、Aレベルはなし、Bレベルが5名、Cレベルは15名となりました。合否すれすれの中から比較的良いと思われる人も含めて、約50名の人と面接をしました。最終的には、約20名の方を採用することになりました。

◆ [できる翻訳が足りない]

この数字を見ますと現在実務の世界で必要とされる翻訳力がある方は全体の10%もないということになります。英訳についてもほぼ同じことが言えます。これが、翻訳業界の現実だと思います。

もうひとつ例を紹介します。(社)日本翻訳連盟では、〈翻訳検定〉試験を実施していますが、この1、2、3級は先程のA、B、Cとほぼ同じレベルです。毎回、200人近い人が応募されますが、その中で、1級レベルになる人はほんの少数です。2級が15~20%位です。3級までで全体の30%ぐらいという数字になります。

先ほど、私がかつて翻訳会社にいたと申しましたが、当時、社内で見習い翻訳者を養成する制度がありまして、私はその指導を担当しておりました。当時の経験によりますと、少し頑張れば、また、適切な指導があれば伸びるということがわかりました。

残念ながら、今はその制度はなくなりました。最近、翻訳会社でもその余裕がなくなっているようです。当時の教え子は、今、第一線で活躍しています。

いずれにしても、即実践で使える人を探すのはなかなか難しい。また、そのような方がいても、その数は大変少ないのです。

◆ [機械翻訳の大衆化]

自然発生的に良い翻訳者が出てくるのを待ち望んでも現状を変えることは難しいのではないかと思います。

翻訳は特定のプロがやるものだという見方をする方が多いと思います。しかし、最近コンピュータ関連のマニュアルなど大量の技術文書の翻訳の仕事がありまして、従来の職人芸的な翻訳では、業界の要望に応えられないという事態になっております。

また、NIFTY-SERVEという商用ネットワークがあります。そこで、機械翻訳のサービスを提供しています。このサービスを利用すれば、簡単に機械

翻訳を一般の人でも体験できるようになってきました。

このような機械翻訳というツールを使って翻訳をやって行くというのがこれからの大きな流れではないかと思います。

(以下、出席者からの質問に対する応答)

◆機械翻訳における前編集の位置付けは？

□英語から日本語へ翻訳する場合は、後編集に重点を置き、日本語から英語に翻訳する場合は、前編集をするというのが一般的だろうと思います。

いずれにしても、翻訳の品質は、前編集者または後編集者の力量に依存します。

実務的には、翻訳の全工程の中で前編集と後編集をどのように位置付けるのが問題となります。翻訳工程を複雑にすると、それだけコストと時間がかかるからです。

前編集について見ますと、問題がたくさんあります。7~8年前、日英翻訳の前編集者の指導をしたことがあります。その経験からいいますと、作業をする人は比較的こり性な人が多かったと思います。

例えば、かかり受けの括弧をいろいろ試しても思いどおりの訳文が得られないので、何度もやり直している。その結果、1文に何時間もかかってしまうことがありました。

その結果、やればやるほど時間とコストがかかってしまうことになり、採算が合わなくなってしまいます。また、機械翻訳の出力が悪いために、前編集者が無理をしてしまう。しかし、そのわりに成果が上がらないというケースも多かったと思います。

機械翻訳は高速で翻訳するというメリットがあるのですが、人間が介在するために、逆に、ブレーキになってしまいます。

しかし、長い目で見れば、辞書に単語を登録したり、前編集をしながらシステムを使い込んでいくしかないと思います。

◆前処理者の要件は英語力なのか専門分野の経験なのかどちらが重要か？

□結論から言えば、両方ともある程度は必要だと思います。私の経験から言うと、英語力は、例えば、英検の準1級ぐらいがちょうどいいと思います。専門分野の知識としては、その分野の基礎知識があれば

十分だと思います。

例えば、コンピュータ関連では、CPUとかI/Oとかいう用語のイメージがわかる程度でいいと思うのです。でも、コンピュータというのを見たことも聞いたこともないという人では無理だと思います。

さて、翻訳業界には神話がたくさんあります。そのひとつが「専門知識」なのです。翻訳を発注する側、また翻訳会社でもそうですが、採用試験のときに経歴書を見ます。例えば、コンピュータ関係ですと、情報処理系の学部を出た人、プログラミングの経験者だと無条件に安心してしまう人が多い。

しかし、私の経験から言うと、なにか違うという気がしています。翻訳をするには、ある程度の専門知識は必要ですが、技術者指向の人と翻訳者志向の人とではやはりタイプが違うと思います。

専門知識があるか無いかでアプリオに裁断するのは実情に合っていないと思います。

むしろ、しっかりした明快な文章が書ける力のある人の方が適していると思います。英語力も専門知識もそこそこの人の方が原文をしっかり読んで明快な訳文を作成できるのです。

英語力のすぐれている人は、つい自分の英語力を過信して辞書を引く手間を惜しむ傾向があります。また、専門知識のある方は、自分の古い知識で判断しようとする傾向があります。

いわゆる産業翻訳というのは、ほとんど現代が対象です。執筆者も同じ時代に生きている人間です。翻訳の読者も同時代の人です。例えば、マニュアルの翻訳では、一般に対象読者は一般のユーザーですから、専門技術者ではない人の方がわかりやすい訳文が作れると思うのです。

また、専門用語について云いますと、例えば、科学技術関連では、何十万語という用語があります。医学分野でも何10万、いや何100万語とかの用語があると云います。人間はとても覚え切れない。

したがって、専門用語はコンピュータに任せて、人間は考えることに専念する。どう表現すれば、原文の意味が読者に明快に伝わるか、という方向に向かっていかなければいけないと思います。さて、MTシステムは、それぞれのメーカーによって多少違いますが、今のところ、人間が相当手直しをしなければ通用しません。ただ、人間の手が加わると人間のレベルを越えられないということに注意していただきたいのです。

とくに力の低い人が後編集した場合には総じて機械翻訳のレベルより悪くなる。MTシステムでできないところは手が出ないからそのままになります。一方、MTの方がよく知っていてよくできている部分を手直ししてしまうので、結果的には改悪してしまうのです。

ですから、機械翻訳の前編集や後編集をする人は翻訳ができる人でなければならないのです。翻訳のできない人には無理な仕事なのです。MTシステムを利用すれば、素人でもできるというのは、現代の間違った神話のひとつだと思います。

◆英検の資格保持者と翻訳品質との関連は？

□これも神話のひとつです。「おまえは英検の1級をもっているから、これを翻訳してくれ」とか、「あの人は英検1級をもっているから翻訳は安心できる」とかいう人が意外に多いのですが、ほとんど偏見と云っていいと思います。

むしろ、先ほどお話ししましたように、自称英語のできる人の場合、悪い面の方が目につくような気がします。確かに、英語力の一般的な力は相当なものなのですが、つい自分の力を過信してしまい、辞書を引く労力を省いてしまいがちなのです。

多分試験制度にも原因があるのかもしれませんが。従来の試験では、試験場という密閉された場所で、しかも辞書の持ち込みも禁止され、頭のなかにあるものが試されるのです。

ところが、例えば、(社)日本翻訳連盟が実施している〈ほんやく〉検定という試験がありますが、この試験では、パソコン通信でも受験できるように辞書、参考書、ノートなど何を見てもよいという制度になっています。

大切なのは、知識ではなく、理解力と表現力だということわけです。まだ、小規模な試験なので、一般人にはあまり知られていませんが、今後は広まっていくと思います。

◆これからの翻訳の理想像は？

□個人的には、執筆者と翻訳者とリライトが一堂に集まって話し合いながら訳していくのが理想だと思います。

私は、現在、フリーの翻訳者としてやっているわけですが、執筆者がわかる場合、電子メールで質問をしています。ほとんど、すぐ返事がきます。

質問の内容は、原文の解釈が複数ある場合、原文の誤りと思われる箇所、対象読者が違うので、こんな意味で訳したいがどうだろうか、...といったことです。

日英翻訳の場合、原文は日本人を対象にして書かれています。これを英語に翻訳する場合、対象読者が、例えば、米国人、英国人、オーストラリア人なのか、あるいはアジアの人なのかで表現が違ってくるからです。

翻訳者は、原文の最初の読者になります。また、リライタは2番目の読者ということになります。この人たちに理解できないようでは、一般読者はまず理解できないはずで。

英日翻訳でも同じことが言えます。英語の原文はあくまで、英語を母国語にしている人を対象にして書かれています。これを日本語にする場合、読者は日本人になります。

いわゆる直訳では、なかなか意味がわからない。逆説的な言い方をすれば、英語のわかるひととならわかるという訳文が多いのです。

さて、産業界全体から見ると、例えば、マニュアルの翻訳では、ダンボール何箱分という大量になります。通常は半年とか1年はかかっていたのですがそれでは間に合わない。

タイムリーな翻訳でないといふ情報としての価値がなくなります。また、翻訳の個性が強いと、全体としての統一がとれなくなります。

そこでどうしても、従来の職人芸的なやりかたでは、時代の要請に応えられなくなっています。やはり、言語処理技術を活用しなければ、問題は解決できないのです。

ただ、現在商用化されているシステムは、どうもこのような現実を踏まえていないようです。したがって、実用化が足踏みしているといった状態だと言えます。

◆大量翻訳のあるべき姿は？

□現状を踏まえると、ローカルな柔軟なネットワークでグループのコミュニケーションをとりながらやっていくという形態がいいと思います。

文体、用語、などは、一括して管理しない限り、大量の翻訳を短期間でしかも一定の品質を維持するのは困難です。また、最近の英文は必ずしも明快な文章ばかりではありませんので、意味不明な場合の

対処の仕方も大切になります。

実務の世界では、いつも時間とコストが問題になります。原文に忠実な翻訳を原点にして、読者にわかりやすい翻訳まで、さまざまなレベルがありますが、そのレベルに応じて翻訳料金も上がるというシステムのコンセンサスが必要だと思います。

翻訳者とチェッカーの役割分担、発注者と翻訳会社の役割分担などを明確にしていく必要もあります特に、発注側に翻訳の品質を客観的に判断できる方があまりいないことも大きな問題です。

◆翻訳は技術的バックグラウンドがなくても可能か？

□結論から言えば、可能だと思います。先ほどもいいましたが、この業界には困った神話があります。専門技術者と翻訳者とはタイプが違うと思います。

例えば自動車の翻訳をする人は、整備士の免許をもっている人でないといけないのか、あるいは、運転免許証をもっていなければいけないのか。そんなことはないと思います。

同じ車といえども、みなさんが利用しているマイカーから見ると、タクシーやトラックなどの職業として運転している人たちの世界はかなり違います。当然、仲間内でしゃべる言葉も違います。

ところが、翻訳という仕事は、自分たちの狭い世界から別の人々に情報を伝えることなのです。

極端な言い方をすれば、技術者は単眼、翻訳者は複眼、であると思います。

したがって、例えば、技術分野の翻訳者にとって必要な資質は、技術者と話しができることなのです。技術者がうまく表現できないこと、原文の表現のままでは読者に意味が十分伝わらないと思われる部分を質問できる力、あるいはそういう方向性をもっている人が適しているのです。

今後は、翻訳を発注する側も受注する側も、お互いに納得できる座標軸が必要になると思います。良い翻訳と悪い翻訳を区別する基準も必要です。そうしてはじめて今後の展望が開けると思うのですが、まだ、時間がかかると思います。そのためにも、機械翻訳協会とか、翻訳連盟のような公的な機関が率先して試案を世間に提示していかなければならないと、考えております。

(この原稿は、5月12日利用技術研究会で講演した記録テープをもとに筆者が書き直したものです)

A T R 音 声 翻 訳 通 信 研 究 所

「自動翻訳電話 日・米・独間の実験成功」
「A Translating Phone for Overseas Calls」

これらは1993年1月28日から29日にかけての国内外のテレビや新聞で大々的に報道されたニュースの見出しです。世界で初めての日本、アメリカ、ドイツを国際回線で結んで話し手の言葉を相手方の言葉に変えて実時間で伝える実験は、夢を現実に近づけるものとして一般の人々にも高く評価されました。

このような国際間の翻訳電話実験の大成功を花道にA T R 自動翻訳電話研究所は7年間の所期の目的を果たし、この度新たに、A T R 音声翻訳通信研究所が関西学術研究学園都市に発足することになりました。これもひとえに基盤技術研究促進センター及びN T Tをはじめ多数の民間出資法人のご協力とご支援の賜であります。

新しいA T R 音声翻訳通信研究所は「高度音声翻訳通信技術の基礎研究」を研究テーマに1993年から2000年の7年間を予定しています。新しい研究所での研究活動は前身のA T R 自動翻訳電話研究所の先駆的技術レベルを引継ぎ、個々の要素技術の完成度を高め、それらを統合化していくことにあります。

現在の音声翻訳技術では話す人が翻訳システムを相当意識する必要があります。例えば話し手はだれでもよいという訳にはいきません。また話す内容(話題)、言葉の種類、言葉の区切り、文法に則った文型、話す速度など翻訳システムに合わせる必要があります。この状態から脱皮し、話す人が翻訳システムを意識することなく、普段の話し言葉で喋り、それを自動的に即座に相手の言葉に変えて伝えることが理想です。特に言葉のハンディを強く意識している日本人が言葉の壁を全く意識せず、自由に外国の人々と話が出来たらどんなに素晴らしいでしょう。ビジネスに、観光に、日本人の活動範囲も益々広がるものと期待されます。

このような理想的な音声翻訳通信は、社会での期待が大きい反面、克服すべき困難な課題が山積しています。話し言葉には特有のくだけた言い回しがあり時には言葉が省略されたり、順序が倒置されたりし

ます。もう少し詳しく見てみますと、図1に示すように認識を難しくする音声現象として話者によって異なる発音、不明瞭な発音、抑揚・強調のある発音などがあげられます。また、翻訳を難しくする言語現象として、会話の状況に依存した多様な表現、意図や態度を表す表現、断片的・非文法的にだけた表現などがあります。コンピュータにとって苦しい処理なのです。これを解決するためには従来の手法にとらわれず、新しい手法が必要になります。そこで課題を次の4つに整理しその基礎技術の確立を図ります。

(1) 自然音声認識技術

通常の会話では、話す速さ、音の大きさ、スペクトルなどに音響的変形が生じたり、会話の途中に「えー」「んー」……といった間投詞や無意味な音が挿入されたりします。このため従来の静的な処理法だけでは認識精度を上げることはできません。この問題を克服するため、音素モデルや言語モデルを発話状況に合わせて動的に適応、再構成する手法を確立することが必要です。また人間は不特定話者の音声認識できるとともに新しい話者の音声の特徴に迅速に適応する能力があります。従来の音声認識では話者の音声特性を標準話者の特性とマッピングする方法が採られています。この機能を人間の認識能力に近づけるためには、マッピング自体の改良はもとよりマッピングをアダプティブに適用する手法の確立が必要で

(2) 発話韻律処理技術

人間が喋る音声には文字で表記できる狭義の言語情報以外にアクセント・イントネーションのような韻律情報や発話の意図などを暗に示す汎言語情報が含まれています。人間はこの両者を利用することにより内容伝達を正しく理解します。従って高度な音声翻訳通信システムでは狭義の言語情報に加え、この汎言語情報を抽出する手法の確立が必要で

にアクセント・イントネーションや話のテンポ・リズムなどの韻律的特性を制御する規則を構築し、自然で多様な音声を作成する音声合成技術の研究も進めています。

(3) 協調・融合翻訳技術

自然な対話では発話の状況や相手の知識を想定した断片的な表現、それらの間に不規則に挿入される別の意図を持った表現、更には倒置、言い直し表現などが現れます。このため規則に基づく翻訳（規則主導翻訳）だけでは限界があります。これを補完する技術として、句や節のパターンとその対訳用例を収集しておき、入力された音声を句や節のパターンに分解し、最も似た用例を抽出する手法（用例主導翻訳）も有望です。そこで、翻訳過程を規則主導翻訳、用例主導翻訳、さらには内容理解に基づく翻訳などの部分問題に分解し、その後、各部分問題の解を組み合わせることで最も確からしい翻訳を導出する基礎技術の確立を図ります。

(4) 音声・言語統合処理技術

自然な話し言葉を正しく理解するためには音声処理と言語処理の情報を互いに利用し、その相乗効果により精度を高める事が有効です。例えば、不用語とみられがちな間投詞や副詞の言語的機能から後続

の音声表現を予測し、音声認識率を高めることができます。また逆に音声の抑揚や強調などの韻律情報を用いることにより言語処理での構文や意味の曖昧さを解消したり特別な意図を抽出することが可能となります。この音声処理と言語処理の接点が「発話の状況」にあることに着目し、両者の結合について研究します。また音声関連の研究を進めるためには実際の音声言語データを収集した大規模データベースが極めて重要となりますのでその作成に取り組みます。

当研究所は上述した4つの課題の解決に積極的に取り組むため4つの研究室と企画課を配置しています。研究員は出資機関からの派遣職員に加え、欧米からの研究者も大勢いて、研究室では日本語、英語、仏語、独語が飛び交っています。音声翻訳通信の研究に必要な国際的な環境ができあがっているのです。

このような環境の中で、研究目標（夢）を高めに掲げまた内外の研究機関と協力して研究を進めたいと思っています。来るべき21世紀においては「異なる言語間のグローバルコミュニケーション」がますます現実的な要求になるものと思われます。本研究プロジェクトの高度音声翻訳通信技術がこのグローバルコミュニケーションにおいてお役に立てるよう全研究所員が一丸となって努力する所存です。今後ともご理解ご支援をよろしくお願ひしす。

(株)エー・ティ・アール音声翻訳通信研究所
社長 山崎 泰弘

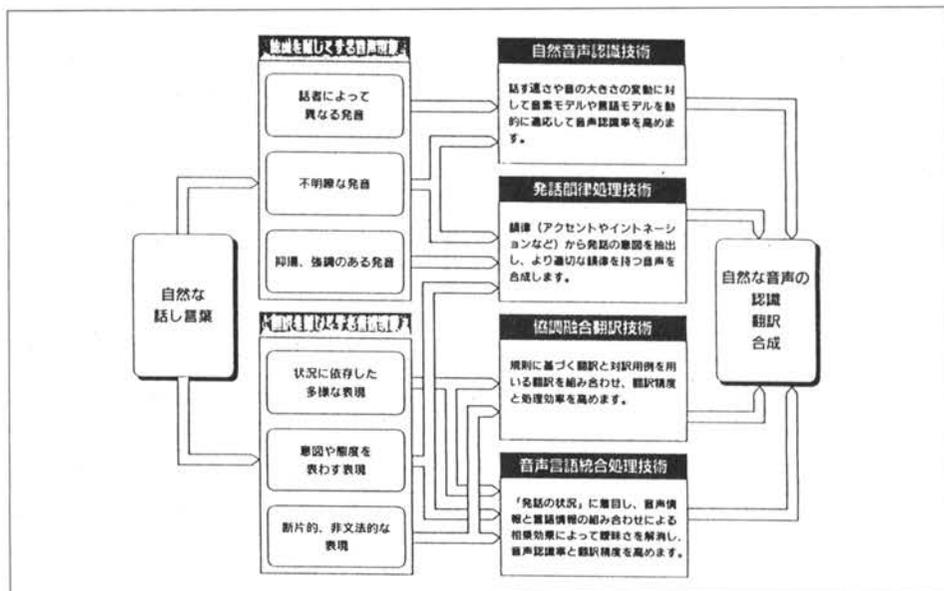


図1 音声翻訳通信における研究課題と解決技術

タイにおける機械翻訳と言語知能研究所の役割

タイの研究者がタイ語をコンピュータ上で処理できないか、と検討を始めてから40年以上が経とうとしている。これまで、この分野の研究はタイ語ワードプロセッサの開発とキャラクタ・コードの割りふりがメインフレーム上で始まり、さまざまな分野の研究者のお骨折りによって、今日ではコンピュータ端末でなに不自由なくタイ語の文章作成・読み取りを行うことが出来るようになった。

一方、英語タイ語の機械翻訳システム (MT) 開発は、1981年にフランスのグルノーブル大学から技術支援を受け、タイの研究者グループが始めた通称「ARIANE (アリアーヌ) プロジェクト」に端を発している。しかしながら、タイ語独特のあいまい性がネックとなってプロジェクトそのものの進行は遅れ、将来的な検討課題としてタイ語特有の問題点を多く示唆するに止まった。またこの時期、部分的にはあるが、他国語との比較を通して見たタイ語の言語学的な系統化モデルが示された。この点については、タイ語の完璧な言語モデル構築に至らないのが現状だが、それにもかかわらず、タイ語の文章を自由にパーシングし生成するシステムの開発に情熱を燃やす研究者は後を絶たないのも事実である。

もう一つ、タイ語のMTに関する研究は、アジア地域の多言語間機械翻訳システム (MMT) の研究開発の一環として、国立電子コンピュータ技術センター (NECTEC) と日本の国際情報化協力センター (CICC) との共同事業が、1987年以来進められて来た。これはタイ語を含む五カ国語 (中国語、マレーシア語、インドネシア語、日本語) の相互間の翻訳を中間言語「インターリングア」を利用して行うものである。多言語間のMTにはこのインターリングア方式が有効と思われ、1995年初頭の事業完成見込みである。

MT開発事業が開始される下地には、タイ語を母語としてさまざまな研究を進めると言語上のハンディキャップを負う、あるいは学術活動上タイ語を使うために起こる不利益などを憂慮する声がある。タイの研究者の間から上がったという背景がある。そこでコンピュータ技術者と言語学者を各大学から集い言

語のデータをどう整備し、構文分析アルゴリズムに反映するか、研究が始められた。研究開発プランは具体化し、タイ語の文章の分析・生成、またタイ語の基本語および専門用語の電子化辞書作成、翻訳支援システム、タイ語入出力システム、およびこれらの構成要素の統合システムと、それぞれ進められている。国からはモンクット国王ソーンブリ技術研究所 (KMITT) とモンクット国王ラドグラバンド技術研究所 (KMITL) から技術支援を受け、翻訳システムは定量的にも定性的にも向上しつつある。

NECTEC=CICCのMMT研究開発事業は国内の多くの研究機関や大学で自然言語処理分野への取り組みが意欲的に進められる状況が見られるようになった。研究課題としては、タイ語の文章の構文分析が最も多く取り上げられてきた。チュラロンコン大学のタイ語文章パーサCUPARSEは従属文法アプローチを採用、AITはGPSGによる文章パーサと、さまざまなアプローチが見られる。また、タイ語特有の言語的特徴についても、それぞれ課題を絞った研究が行われると聞いている。このように、言語学的な面からの視点を交え、大規模MTシステムの構築を実現する基礎固めが行われて来たことを評価したい。タイ語をソース言語としてもターゲット言語としても扱えるようなMTが、近い将来いろいろな機関から実用化される事を期待する。

言語知識研究所、通称 Links は1992年に国立の研究機関として設立されたNECTECの下部組織で、タイ政府の全面的助成を受け、自然言語処理 (NLP) とAIの分野で研究を進めるものである。また、国内の研究活動を奨励するためにも知的資源の提供を行うという役割を果たすため、Links自身が情報センターとなる構想である。

Linksの目標は以下のとおりである。

1. 情報処理、NLP、大規模データベースなどの研究に関する情報を集め、評価する。
2. NLP、AIに関する研究を通じて、システム開発の妥当性を評価する。
3. システム開発の共同事業を行い、国内の私企業関係団体における高度研究の知的資源・資金のポ

テンシャルを高める。

MT開発における Links は、NECTEC=CICCのMMTプロジェクトに参加し、タイ語システムの開発が他の四か国語とインターリングア経由で整合するよう指導する役割を負う。この点ではCICCの多大なご支援によって、現段階ではシステム開発におけるコーパスとして厳選したタイ語の文章を用いられるまでになった。今後の改良の成果として、実用段階へもって行く見通しが立ってきた。

MMTプロジェクトの所期の目的は、あくまでも参加各国が多国語間でデータ交換を効率的に行うことであり、そのために必要な情報を準備するツールとして、MMTの完成がぜひ待たれるところである。

ところで、開発段階では翻訳結果の精度を評価するコーパスセンテンスを用いたが、このコーパスベースに加えてテキストベース、語句辞書ベース、文法ルールベースを揃え、将来的な開発に向けて貴重な資源として活用し定期的にシステム評価を行い現実の文書翻訳ニーズへ対応させたいと検討中である。

コンピュータ技術者、言語学者の緊密な協調関係

により、ここ数年MT技術に関して双方の専門分野でさまざまな問題が解決されて来た。中でも最大検討課題はコンピュータ言語学の問題というよりタイ語の言語的な特徴そのものに起因するものであって、タイ語の文章表現では文単位に区切らないことから文の認識が大変困難となり、そのため、語・文の認識処理が難しく、未だに語の区分を完全に自動化する有効なアルゴリズムを得ていない。同様に、文章構造の表現も的確なモデルを見いだしていない。これら根本的な課題が解決され、MT開発の次の段階へ進むことが出来るよう、今後の研究成果に期待される。

科学技術環境省

国立電子コンピュータ技術センター

言語知識科学研究所

ウィラッチ・ソーンラートラムヴァニチ

(所在地) RAMAVJ ROAD RAJTHEVJ

BANGKOK 10400 THAILAND

AAMT法人会員

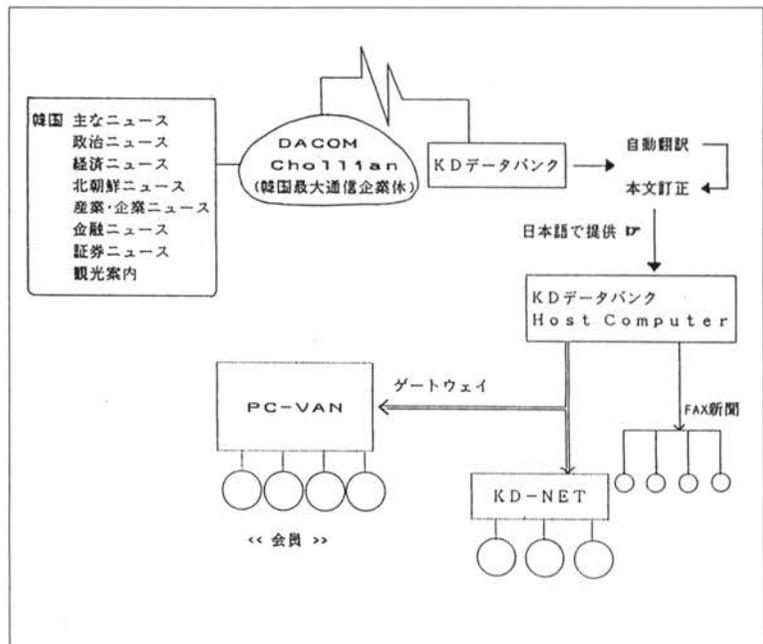
自動翻訳システムを利用した

「韓国 ニュース速報」

提供 (株) 高電社KDデータバンク

(株) 高電社(本社大阪市)では韓国の最大通信企業体DACOMが運営する「千里眼」から中央日報のニュースを去る2月24日から提供しています。これは高電社が韓日機械翻訳機で翻訳しPC-VANとKDデータバンクに日本語で提供しているものです。会員はパソコン通信かFAXを利用して最も早い時間帯で韓国のニュースを知る事ができます。内容は政治、経済、産業、金融、証券のニュースの他観光案内も行われています。PC会員はPC-VAN又はK OUDENSHA:KDバンクに加入する、料金は50円/分 FAX会員は3回/日送信、20,000円/月

(☎06-628-8880,03-3256-3061)



KAIST/CAIR (韓国高等科学技術院)

CAIRデジョン校コンピュータ科学科 ケイスン・チョイ教授

KAIST/CAIRの概要

KAIST/CAIRは1990年に設立され、KOSEFに認定された機関である。MT関連の過去3年間の研究実績として、MATES/EK(英/韓MTシステム)を開発した。韓英MTシステムは目下開発中で、自動通訳システムの一環として会話ベースの翻訳を目指す。

設立年=1990年

職員数=約377人

最近の大規模プロジェクト=

韓国語情報処理プロジェクト

知能知覚情報システム

マルチメディア・コンピュータシステム

主なプロジェクト提携先=

KOSEF基礎研究開発資金により運営される企業・研究機関との提携

将来的なゴール=高度情報処理技術の開発

英/韓MTシステム

MATES/EK

(機械翻訳環境システム/英語→韓国語)

プロジェクト概要

このプロジェクトは、UNIXワークステーションおよびパーソナルコンピュータ上で作動する機械翻訳システムのプロトタイプの開発を目的として、韓国高等科学技術院(KAIST)とシステム技術研究所(SERI)を中心とする開発グループが行ったものである。

実施期間は1988年8月より1992年5月までで、韓国科学技術省および民間2社(ダエウー・テレコミュニケーション社、サンスン電機社)の支援を受けた。

プロジェクト抄史

プロジェクトの初期段階では機械翻訳システムの開発の妥当性の評価を課題として、小規模のMTシステムを汎用リスプ上で開発、作業時間の面で多少課題を残したが、このシステムが十分評価され、これ

をベースとして規模が大きく実用性を持つものとして英/韓MTシステムMATES/EKをC言語で設計開発に取り組み、現在、MATES/EKは英文の韓国語訳を高速処理できるシステムとして完成した。

MATES/EKの設計思想

MTシステムは構成要素をそれぞれだんだん改良・拡張しながら開発するのが常識であるが、MATES/EKもその例にならない、まず環境の設定について文法(書き込み、改良、拡張)および辞書操作との兼ね合いで開発方針を決めるところから着手した。また、環境の設定に次いで文法書き込み用の言語CANNA(3種のtransformatonを使用)およびACF文法(コンテキストに縛られない拡張文法)を開発した。前者は意味分析、変換、文生成に、また後者は文章の分析に用いる。辞書操作を簡便化するため辞書ツールとしてエディタ、コンバータ、インタープリタを組み込んである。MATES/EKはトランスファ方式を採用、文生成に至る翻訳は3段階(英文分析、英語→韓国語変換、韓国語文生成)で行う。

MATES/EKの特徴

コアとして以下の特徴を持つ。

- ※文章の形態素解析を用いる副次的な語の把握と無条件のあいまい性排除
- ※ACF文法による一般化LRパーサーの採用
- ※分類規則と辞書規則の併用による辞書変換
- ※文法の平明性=文法書き込み言語CANNAを採用、3段階の変換を行う
- MATES/EKの開発に関しては、文法の開発に3タイプのサンプル文を用いた。
- ※文法のスケルトン構築=Quirkの英文法書から採用した文(259点)
- ※英文の文章構造のパターン化=ヒューレットパッカー社の英文スタンダードセットから採用(909点)
- ※電子/電気関係の文法の採用=IEEEコンピュ

ータ専門誌から例文の採用 (1708点)
 上記のサンプル文はシステム開発段階に引き続きシステム試験でも用いた。これらの翻訳結果は、MATES/EKのパフォーマンスが(翻訳スピード・成功率の両面で)文の長さに大きく依存すると評価された。サンプル文の98%を占める14ワード以下の文では翻訳がスムーズに行われたのに対し22ワード以上の文では成功率が80%を下回った。

MATES/EKのコア構成

辞書、文法書き込み言語CANNA、CANNAを用いる文法、英文形態素的分析アナライザ、韓国文形態素的ジェネレータより構成される。

- ※辞書 電子/電気分野2万語 一般分野3万語
 変換辞書に技術用語辞書、辞書規則辞書を持つ
- ※文法書き込み言語CANNA基礎オペレーションとして3段階の変換を行う文法ライターに対し以下の機能を与える。
 - ー変換辞書から訳語をダイナミックに抽出する(分類規則と辞書規則の併用)
 - ーあいまい性の処理を行うフィルター機能、採点機能
- ※英文形態素的分析アナライザ
 基本的な機能として語の把握と接辞の処理、複語の把握と慣用語句把握を行う。より効率的な分析

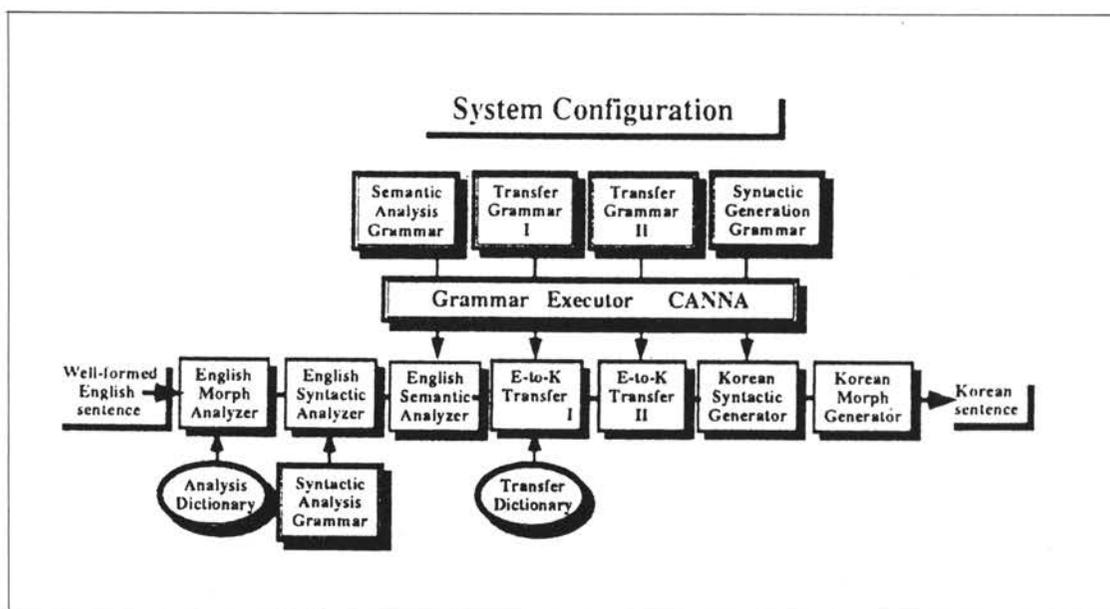
のため、分類上のあいまい性を解決する。

- ※英文文アナライザ
 ACF文法454項目
- ※英文意味アナライザ
 インプットとして英文文ツリーを取り、英文意味ツリーをアウトプットする。この過程で、フィルタ機能、採点機能により文のあいまい性を解決する。
- ※英文の韓国語への変換
 辞書変換と構文変換を行う。
 前者は変換辞書よりの確な訳語を抽出する。
 後者は英文の意味表現を韓国語の意味表現に変換する。
- ※韓国語文ジェネレータ
- ※韓国語形態素ジェネレータ

将来の展望

MATES/EKの現代バージョンは商業ベースのMTシステムのプロトタイプという性格上、今後は他分野の翻訳を行うべく辞書および文法の拡充など開発を進める。また翻訳品質を高めるためあいまい性の処理に関してハイブリッドなアルゴリズム開発に集中する所存である。

(開発品は機械翻訳サミット展示会に出品の予定)



機械翻訳用の辞書

今回は、機械翻訳用の辞書について紹介します。第1回から第6回まで機械翻訳を支える形態素解析構文解析、意味解析、文生成といった技術が紹介されてきましたが、これらの処理の全てのフェーズにわたり基礎的な言語データを提供する辞書についてのお話です。

一口に機械翻訳用の辞書と言っても、その処理方式の違いにより様々な形態の辞書が存在します。原文の中の語を訳語に置き換えて翻訳する直接翻訳方式では、訳語、品詞などの極めて単純な文法的機能の情報しか必要ありません。構文トランスファ方式においては、構文木を作成するための下位範疇化や変形といった情報が必要ですが、語義、格フレーム、などの細かい意味レベルの情報は必要ありません。言語独立の意味表現を介して翻訳する中間言語方式(ピボット方式)では、形態・統語・意味・文脈の全てのレベルの言語現象を扱った辞書が必要です。また、用例ベース翻訳方式では、文字列としての類似度とシソーラスを基本にしているため、形態素解析用の情報や構文解析用の情報を記述した辞書は必要ありません。

現在、世界中には30を越える機械翻訳システムがあり、また、自然言語処理システムまで範囲を広めれば、その数は100以上あります。これらのシステムは上に述べた方式や採用している文法に違いがありこれらの方式を可能にしている個々の辞書の差に焦点を当てても質的意味はあまりありません。従って観点を辞書中心に変えて、様々な機械翻訳方式や技術に役立てることのできる言語資源と辞書を見做しどのような情報の蓄積が必要とされているのか、また、開発の現状はどのようであるのかについてお話したいと思います。

1. 冊子体辞書と機械翻訳用の辞書との違い

シリーズの第1回に冊子体辞書との違いが簡単に説明されていますが、ここでもう少し詳しくその違

いを述べてみたいと思います。なお、上述のように、ここでは機械翻訳用の辞書に、方式に捕われない言語資源としての汎用辞書(つまり電子化辞書)を想定します。

人間が自然言語で書かれた文章を理解したり作ったりするためには、言葉の意味やその用法を知っていなければなりません。同様にコンピュータによる自然言語処理も、辞書の中に盛り込まれた情報をもとに行われます。つまり、コンピュータが自然言語を人間の様に処理していくためには、自然言語を理解するために必要な情報や、自然言語を使いこなすための情報などの知識を盛り込んだ非常に大規模な辞書が必要となります。

従来、電子辞書(機械可読辞書)という名前で、出版物としての辞書の内容をそのまま記憶媒体に蓄積し、機械的手段によって、その内容を参照、利用する辞書がいくつか開発されています。しかし、これらの電子辞書は人間が参照し、理解するためのものであり、機械翻訳用の辞書とは異なっています。

機械翻訳用の辞書は、人間用の辞書の単なる機械可読版ではなく、コンピュータが自然言語を理解するために必要な情報をすべて盛り込んだものでなければなりません。人間用の辞書は用例によって言葉の意味内容の理解を促すことが多く、スペース上の制約もあり、細かい記述がなされていない場合が多い様です。このため、膨大な常識や高度な推論能力を持つ人間には役立ちますが、機械処理用にそのまま流用することは不可能です。また、語釈文の中に現れる言葉がその言葉のどの意味で使用されているのかが明記されたものや、語と語の共起や使い方の違いが網羅的・体系的に記述されたものはありません。

機械翻訳用の辞書には、言葉の意味、つまり、単語が表すことのできる概念と、その単語がその概念を表すときの文法的特性、また、その概念をコンピュータが理解するために必要な知識が格納されていなければなりません。

従って、機械翻訳用の辞書は、冊子体の辞書と比較がして、次のような特徴を持つことが条件となります。

- (1)一般的な文章で用いられる語彙を網羅した大規模辞書
- (2)本格的な意味処理に必要な知識ベースを備えた辞書
- (3)大量のテキストに基づく高い客観性を持つ辞書
- (4)言語や分野等に対し高い一般性を持つ辞書

2. 辞書の内容と役割

次に、機械翻訳用の辞書に盛り込まれる内容とはどのようなものであり、どのように機械翻訳に役立つのかの実感をつかむためにもう少し詳しく説明しましょう。

(1)形態素レベルの情報

形態素レベルの情報としては、単語（形態素）の見出しおよび形態素の接続に関する情報があります。形態素の見出しは、文解析での最初のプロセスとして、入力文を構成する形態素を見つけ出すための形態素解析で用いられます。これは辞書引きのプロセスとも呼ばれます。また、文を合成していく形態素生成においても、出力文を得るためにこの見出しが用いられます。

(2)構文レベルの情報

構文レベルの情報は、一般に文法情報と呼ばれる情報で、名詞とか動詞、形容詞、副詞といった品詞やあるいはさらに細分化された品詞、その他表層格情報や構文支配情報などがあります。これらの情報は構文解析や構文生成で用いられ、この文法情報をもとにして解析規則や生成規則が作られます。

(3)意味レベルの情報

単語は複数個の概念を表すことができますし、また一般によく使われる単語ほど多くの概念を表しています。文を解析するときには、ある単語が文中に現れたときに、その単語がどの意味で使用されたかつまり、その単語がどの概念を表したのかを識別する必要があります。文解析において意味解析はこの目的のために行なわれるものでして、意味レベル

の情報は、単語の多様な概念を識別するための情報を与えるものです。このためには、あらかじめ単語が表すことのできる概念を辞書に記述しておかなければなりません。

意味レベルの情報は、単語の表す概念について種々の知識を記述した知識ベースとして用意されます。文解析においては、この知識ベースを用いて概念の親和性などをもとにした推論を行ない、多義性を解消することになります。

知識ベースはコンピュータに理解される形で概念の定義を与えるもので、常識として成り立つ事象やいろいろな知識を記述したものです。この知識の記述は、基本的には、概念の間に成立し得る関係をすべて数え上げることによってなされます。例えば“鳥が飛ぶ”という知識は“c#鳥”という概念（単語概念を区別するため以下、概念にはC#を付加します）が“c#飛ぶ”という概念の“動作主”になり得るといふ概念間の関係を示すことで表されます。また、“水は飲み物である”という知識は“c#水”という概念“c#飲む”という概念の“対象”になり得るといふ概念間の関係を示すことで表されます。

(4)語用レベルの情報

語用レベルの情報は、ごく自然な文を生成するために用いられる言い回しの情報です。例えば機械翻訳においては主に訳語選択を行なう際に用いられます。入力文に現れた単語がどの意味を表しているかが分かって、その意味を表すことのできる訳語は複数個存在する場合があります。例えば“c#直す”という概念を表す英語の単語は“modify”、“correct”“update”、“mend”など複数個存在します。その中から“c#直す”という概念の対象になるものが何かということによって適切な訳語を選ぶために用いられます。この情報は、ある単語とどういう単語が共起する（同じ文中にある関係で現れることができる）のかという共起関係を定義したものです。例えば、“c#直す”の対象が、“c#エラー”であれば“correct”という訳語を選ぶということは“correct”と“error”が“対象”という関係で共起するという情報を用いて行なわれます。

(5)対訳語に関する情報

対訳語に関する情報としては、表層的な対訳関係の情報と概念を通しての対訳に関する情報がありま

す。概念を通しての対訳情報には、ある概念とある概念は同義であるとか、類義であるとか、また上位-下位関係であるとかなどが定義されており、これらの関係を利用して訳語を決めていくことになります。例えば“犬”という単語が持っている概念と“dog”という単語が持っている概念が同じであるということとで訳語が選ばれたり、日本語の“麦”という単語が持っている概念を表す単語は英語にはないのため“麦”の下位概念である“大麦”あるいは“小麦”などに相当する概念を表す単語を持ってきて訳語にするというようなことが行なわれます。

3. 辞書開発の動向

このような特徴や内容を備え持つ言語データとしての辞書は、これからの自然言語処理にとって極めて重要であり、長期にわたり改良を続けていかなければならない性格のものであります。その開発・改良は一企業や一国家の事業の枠を越えたものであり、さらに、国家的規模で統一すべきだという共通の認識に立って、世界の主要地域で共同開発が進んでいます。これらの活動は、政府機関、学会、大学、出版社、企業に対し影響を及ぼしており、様々な新しい活動が展開されています。また、プロジェクト間の協力も始まろうとしています。

(1) 北アメリカ

米国では、辞書研究コンソーシアム(CLR: Consortium for Lexical Research)が、計算言語学会(ACL)に設けられた委員会の監督のもとに、DARPA(Defence Advanced Research Projects Agency)からの資金援助によって、ニューメキシコ州立大学に一昨年設立され、活動を始めています。CLRは、活動の初期段階であり、言語資源の共有という交通整理的色彩が強いようです。

新たに昨年4月、言語データコンソーシアム(LDC: Linguistic Data Consortium)が同じくDARPAにより設立され、ペンシルバニア大学(University of Pennsylvania)に設置されました。企業、大学及び政府機関により構成され、アメリカの国際協力を含め、センター的役割を担うものと思われます。将来はCLRと一体化する可能性もあるようです。話し言葉に対する比重が高いことが特徴であり、現在談話、文章、パラレル(対訳)・テキスト、構文解析済みテキスト(PennTreebank)、文章検索、そしてメッセ

ージ理解に関して試験的収集を行っています。また、語彙、文法、そして手書き書体の収集も計画しています。DARPAの辞書開発活動を引き継ぎニューヨーク大学の統語辞書、プリンストン大学の概念辞書(Word Net)を統合して、COMLEXという辞書を開発する予定です。

米国の第5世代コンピュータプロジェクトにあたるMCCでは、百科事典から人手により、CYCという識ベースを10年間プロジェクトで開発しています。現在、約200万程度の存在論的(ontological)知識が入っています。最終的には常識までもコンピュータで処理しようとする野心的な計画です。

カナダのウォータルー大学ではオックスフォード英語辞典(第2版)を辞書データベース化(LDB)する計画があります。これは未だ研究段階でして機械翻訳用の辞書開発にすぐ使える状態ではありません。

現在、客観的データの収集を目的としていくつのテキスト収集活動がありますが、互換性の問題に対処するために、規範作りの活動であるテキストコード化計画(TEI: Text Encoding Initiative)が1988年から学会レベルで進行中です。これは、テキスト・データを始めとする言語データを標準的資料とするために、テキスト構造を記述するコード化の標準を策定することを活動内容としています。米国のコンピュータと人文科学会や、計算言語学などが支援しています。米国が中心で欧州のメンバーが加わり、日本がごく最近参加し始めたところではあります。

(2) ヨーロッパ

ヨーロッパでは、ECのESPRIT計画内のACQUILEXプロジェクト及びMULTILEXプロジェクト、EUREKA計画内のGENELEXプロジェクトにおいて辞書開発が行われています。ヨーロッパのこれらの活動における傾向は、産学一体(産業主導型)となって辞書開発が推進されていることです。他の地域に較べ、辞書モデル、方法論、ツールの標準化が一步進んでいます。百家争鳴GPSG、LFGといった文法理論の中で、中立型辞書(Neutral Dictionary)を構築しようとしています。方法論は、既存の機械可読辞書(MRD)の利用に焦点があります。ヨーロッパにおけるこれらの活動の概略を以下に述べます。

ACQUILEXは、1ヵ国語に対する辞書開発技術の修正を目的としたものです。統合された多言語辞書知識ベースの構築に応用可能です。MULTILEXは

ヨーロッパ各国で利用できる多言語多機能辞書の仕様を策定し、辞書データを作成し、検索などの統合的なソフトウェアを開発し、応用システムによって辞書の有用性を検証しようというのが目的です。GENELEX は共通の辞書仕様の下でフランス語、イタリア語、スペイン語に対する包括的辞書を開発しようというものです。

CLR の発足に刺激されたケンブリッジ大学出版部がケンブリッジ・ランゲージ・サーベイ (CLS) という多言語電子化辞書・コーパス開発活動を発足させようとしています。これは、ACQUILEX を拡張する位置付けにあり、英語、イタリア語、オランダ語、スペイン語に加えて、フランス語、ドイツ語、日本語をカバーする計画です。

ヨーロッパに特筆すべきコーパス収集活動が2つありますので紹介します。

一つは、英国ナショナル・コーパス・プロジェクト (British National Corpus Project) で、これは英国英語を代表する言語データの収集を目的にスタートしたばかりのプロジェクトです。オックスフォード大学出版部、ロングマングループ、オックスフォード大学、ランカスター大学、大英図書館が協力しています。

もう一つは、Network of European Reference Corpora (NERC) で、EC による、共同体内主要言語化のテキスト収集の調整活動です。収集に共通の方法論を用い、コード化に TEI に基づく共通標準を設定しソフトウェアも共通のものを開発します。ピサ大が中心となりヨーロッパ各国の大学や出版社の協力の下で執り行われています。1千万語の収集が

現在の目標です。

専門用語に関して述べますと、EC の EURODI CAUTOM, そして UNESCO の下部機関 INFOTERM 等が大規模な専門用語データベースを開発しており、現在では、オンラインネットワークサービスにより世界中で利用できるようになっています。そして、これらのユーザの大半は翻訳者です。

(3) アジア

一方、アジアに目を向けますと、注目すべきプロジェクトが2つあります。一つは、EDR プロジェクトで、(株) 日本電子化辞書研究所 (EDR) が、日本語と英語を対象として大規模辞書・コーパス開発活動を行っています。開発の種類とサイズは、単語辞書 (基本語20万語、専門用語10万語)、対訳辞書 (30万語)、共起辞書 (30万語)、知識ベースである概念辞書 (概念体系40万、概念記述40万)、コーパス (25万文) です。概念辞書は日英共有ですが、他はそれぞれこのサイズの辞書が開発されています。

前述の機械翻訳用辞書としての特徴・内容を備えたものにする計画です。開発期間は未だ2年近く残していますが、開発の終了したものから公開してゆく予定です。

もう一つは、日本が主導で研究開発している近隣諸国間機械翻訳プロジェクトです。相手国は中国、マレーシア、インドネシア、タイの4ヵ国であり、EDR の概念辞書とリンクできるため、効率良い電子辞書開発が可能となった事例になっています。

[(株) 日本電子化辞書研究所 末松博]

KB & KS '93 (大規模知識ベースの構築と共有に関する国際会議1993)

国際会議 (12月1日~2日) 於 京王プラザホテル

- セッション [1] Academic and Social Demands for KB& KS
[2] Language Technology and Science
[3] Knowledge Technology and Science
[4] Sharable Knowledge Sources
[5] Panel Discussion

ワークショップ (12月3日~4日) 於 工学院大学 (予定)

主催 (財) 日本情報処理開発協会・問合せ先 [アイ.エヌ.エス (株) 03-3494-1869]

PENSÉE-GV (パンセ GV) は、沖電気が開発した日英/英日翻訳支援ソフトウェアです。富士ゼロックス社の統合文書処理ソフトウェア GlobalView 上で動作し、図表を含む文書をそのままのレイアウトで翻訳します。

従来の機械翻訳システムは、文字情報だけしか扱えなかったため、入力文書から文字情報を抽出して翻訳し、その結果をワープロやDTPシステムなどで整形する作業（レイアウトしたり図表を切り貼りするなど）を行なう必要がありました。

しかし、このような整形作業は機械的なもので、翻訳作業全体の効率を悪くする要因のひとつといえます。

沖電気は、この問題を解決するため、整形作業などの機械的な作業は自動化するという基本理念のもとに、入力文書のレイアウトや図表の情報をそのまま翻訳結果に反映する日英/英日翻訳支援ソフトウェア PENSÉE-GV を開発しました。

<特長>

●簡単な操作

操作は非常に簡単で、マウスで翻訳対象となる文書アイコンを指定し、プルダウンメニューから日英翻訳あるいは英日翻訳を選択するだけで自動的に翻訳結果の文書アイコンを作成します。また、複数の文書アイコンを納めたフォルダや文章、文章の一部といった指定も可能です。

●図表を含む文書の翻訳とそのレイアウト情報の再現

図表を含む原文のレイアウトや文字のフォント、修飾情報などはそのまま翻訳結果に反映されます。従って、入力文書から文字列部分を抽出し、翻訳後にレイアウトしたり図表を切り貼りするといった整形作業がほとんど不要となります。

●高品質な翻訳

翻訳処理は日英約9万語、英日約6万語の辞書と独自に開発した深層格文法を利用しており、高品質な訳文が得られます。特に長文の翻訳に威力を発揮します。

●高速翻訳

1時間に約15,000語の速度で翻訳しますので、1ページあたり数十秒で翻訳が完了します。

●翻訳オプション

日英翻訳では、マニュアルに頼出する主語なし文を、受動態に訳出する（標準）か命令文として訳出する（命令）かを指定するオプションがあります

英日翻訳では、「である」調で訳出する（標準）か「ですます」調で訳出する（敬体）かを指定するオプションがあります。

<仕様>

対象言語：日→英/英→日

DTP：Global View

翻訳：PENSÉE

ハードウェア：OKITAC-S

メモリ：16MB以上（32MB以上推奨）

ディスク：400MB以上

（OS及びGlobal Viewを含みます）

OSなど：SUN OS 4.1.1以上

+JLE 1.1.1以上

+OS/FX 拡張機能

+Global View

基本辞書：

日英：約9万語

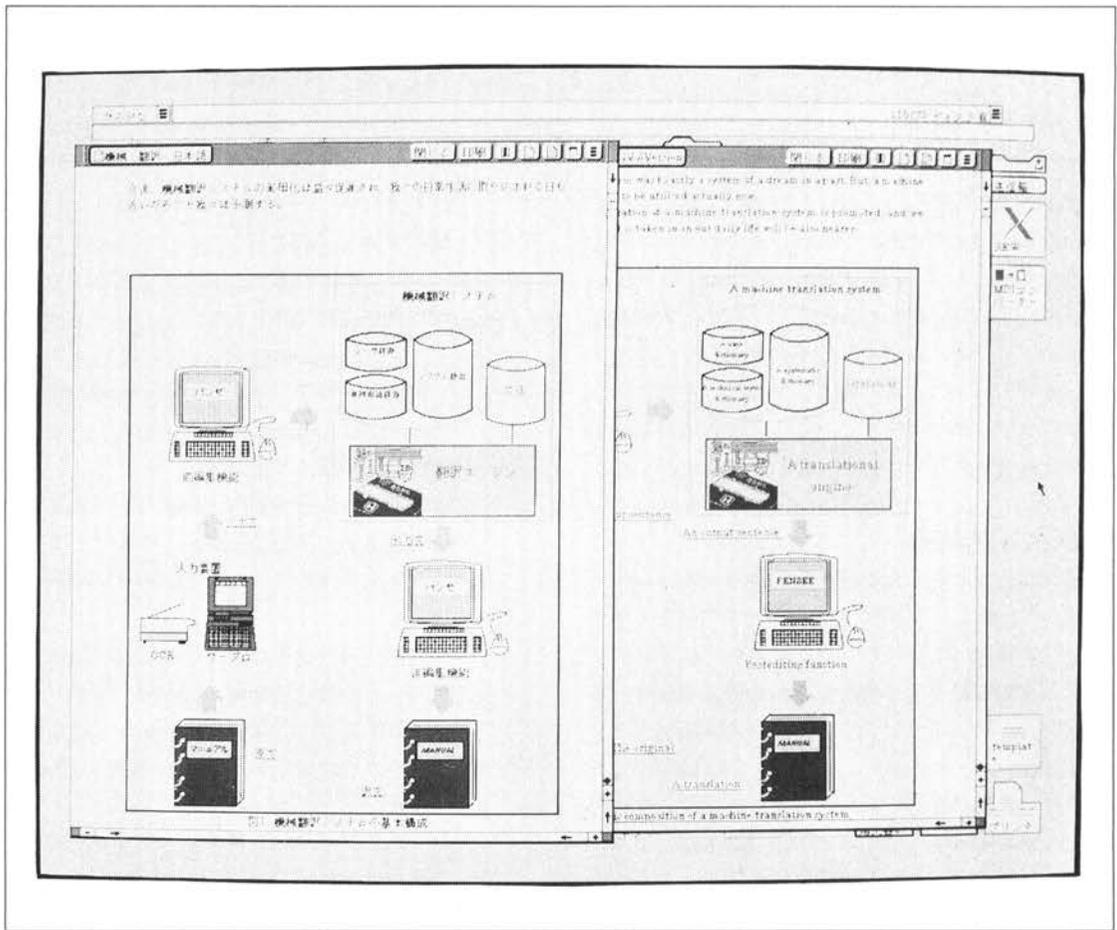
英日：約6万語

専門用語辞書（オプション）：

日英：16分野、約5万語

英日：16分野、約5万語

（情報処理、自動車、金属、数学など）



ユーザ辞書：語数無制限（ディスク範囲内）

処理速度：約15,000語/時間

（機種によって異なります）

翻訳オプション：

日英：標準/命令

英日：標準/敬体

OCR（オプション）：

日本語：SPARC Reader

（販売元：メディアドライブ研究所）

認識速度：約40文字/秒：

認識率：99%以上）

英語：Scan WorX

（販売元：サカティンクス株式会社）

認識速度：200~300文字：

認識率：99%以上）

外部文書の取り込み：

テキスト、JStar、一太郎 Ver3、4、OASYS、

Lotus1-2-3など

（テキスト、JStar 以外は変換ソフトが必要です。）

また、文書によって取り込めない図形があります。）

※ PENSÉE は沖電気工業(株)、大阪ガス(株)、オービス総研の登録商標です。

※ Global View は、富士ゼロックス(株)の登録商標です。

※ SUN はサンマイクロシステムズ社の登録商標です。

<問い合わせ先>

沖電気工業（株）

情報通信システム事業本部

オープンシステム事業部

市場開発部

TEL. 03-3454-2111

日本の国際化が急速に進み、海外の文献情報があふれるように国内に流入しておりますが、中には貴重な文献にもかかわらず目も通さず捨てられているものも多い、と思われま。そうした情報社会の中で、速やかに大事な情報を受け取るための道具として「E tran J」英日翻訳ソフトはお役に立つと思います。また低価格ながらその翻訳能力はさまざまな英文文献の読解に大きな力となります。

E tran J の特徴

・UNIX コマンドとして利用できる

E tran J は UNIX コンピュータで動作する事を意図しています。ワークステーションを仕事の道具として使用している方であれば、必要な時にコマンドをたたけば英日翻訳ができます。

大量の英文であればバッチ処理ができます。また UNIX の豊富なソフトウェア群と組み合わせることで多種多様な利用形態を可能にするために、E tran J の翻訳コマンドは、標準入出力を採用し、他のソフトウェアとパイプで結合したり、多重処理を容易に実現し、利用者が使い易い環境にユーザ・カスタマイズすることが可能です。

・ウインドウ・インタフェース

また翻訳コマンドに OS F / Motif を使用したウインドウ・インタフェースを用意し、対話的に翻訳したり、ユーザ辞書の更新が容易にできます。

・入出力テキストファイル

英文テキストは OCR 読取装置で入力したものでオンラインで電送されてきた英文テキストファイルでもすでにある周辺装置を活用して作成したテキスト・ファイルが利用できます。

翻訳結果は日本語 (EUCコード) あるいは英語・日本語対訳で出力されます。この翻訳結果を既存の清書ツールやエディタを利用して修正できます。

・翻訳方式

翻訳方式は句構造木を用いるトランスファー式、原文をボトムアップ方式で構造木を作っていく過程で、ノードが一つ見つかりと直ちにノードの日本語訳を生成してしまいます。

簡便なトランスファー方式であるため翻訳スピードは 5 万語 / 時間を実現しています。

・翻訳能力と翻訳カスタマイズ

E tran J は極力前編集を軽減し、後編集し易い翻訳結果を出力する事を目指しています。

翻訳は文単位で行い、最長 120 ワードの英文翻訳します。

翻訳能力は辞書の充実と「適訳規則」「文法規則」の学習に負う所が大きいです。

辞書は熟語辞書、基本単語辞書をもっています。その他に専門用語を集めた専門辞書がオプションとしてあります。

また、ユーザ辞書を作成し、有効に活用する事ができます。

「適用規則」「文法規則」等の翻訳ルールが翻訳プログラムと独立させてあるため適宜翻訳ルールをカスタマイズすることができます。

仕 様

辞 書 類	大 き さ
熟 語 辞 書	10,000 熟語
基 本 単 語 辞 書	50,000 語
適 訳 規 則	2,500 ルール
文 法 規 則	3,500 ルール
オプション辞書	
情 報 処 理 専 門 辞 書	25,000 語
医 学 ・ 医 療 専 門 辞 書	30,000 語
機 械 工 学 専 門 辞 書	20,000 語
電 気 ・ 電 子 工 学 専 門 辞 書	20,000 語
会 計 ・ 金 融 専 門 辞 書	25,000 語
登録に必要なディスク容量	約 3 メガバイト
実行のメモリーサイズ	約 3 メガバイト

トピックス

投資減税……政府は緊急総合景気対策として省力化税制を実施する。これは7月1日から1年間の時限立法であるが、この期間にユーザが導入した省力化機器にたいしては税制面での優遇措置(特別減税)の恩典が付与される。機械翻訳システムもその対象品目に含まれる。詳細は官報にて告知されるが、これにより機械翻訳システムの普及が一層促進されるものと期待される。

翻訳検定試験申し込み受付中……(社)日本翻訳連盟が認定する第17回翻訳検定試験(1~3級)が8月21日に実施される。試験科目は①英文和訳 ②和文英訳 ③独文和訳の3科目でこの中から自由に選択できる。また①と②は重複して受験できる。また集合試験の他全国の翻訳家志望者の便宜を考慮し、NIFTY-Serveの特設機能上でも受験できる。パソコン受験についての詳細は7月上旬にトップメニュー直前の「お知らせ」で案内される。この試験は辞書、参考書、ノート等何を見てもよい。試験結果は2ヵ月以内に通知される。受験料は1科目1万円。申し込み締切日は8月7日。申込先は(社)日本翻訳連盟(東京都中央区八丁堀2-8-10、☎03-3555-6365)まで。

第5回TCシンポジウム開催……テクニカルコミュニケーション協会主催第5回シンポジウムが8月26日に東京、新宿の工学院大学で開催される。テ-

マは「伝える技術、もっと優しく、ユーザに近く」となっている。詳細は近く広報されるが、当協会もこのシンポジウムに協力している。

絵本翻訳コンテスト……(株)バベル・プレスでは絵本翻訳コンテストを実施している。絵本にかかれた英文(翻訳の世界7月号に掲載)を課題とし、訳文を募集。応募締め切りは6月末。入選発表は同誌の12月号に掲載される。問合せ先は☎03-3295-2306

新書「技術翻訳のテクニック」(富井篤著)……英文の意味を誤解なく的確にとらえ自然な日本語とするためにどのような配慮が必要か、無生物構文の処理を体系的に具体例を豊富に盛り込んで解説されている。また機械翻訳システムについても記述されている。出版は丸善、全国の書店で販売中

AAMTジャーナル英文版……当協会は92年6月の総会で現在の協会名に名称変更し、会員組織もアジア太平洋地域に拡大し、機械翻訳システムの普及啓蒙に努めている。その一環としてAAMTジャーナルの英文版を発行、海外の会員に送付している他国際情報化協力センターのカウンターパートナーである各地域の情報化関連機関に設置されている「情報化ライブラリー」にも継続縦覧出来るようにしている。80年代以降アジア各地域への日本企業の生産シフトや情報拠点の設置が急増しているが、それだけに翻訳ニーズも高く、海外会員も次第に増加して来ている。

新会員の紹介

(個人会員)

貝塚 泉
末松 博
立田 種宏
八巻 修一

Shravan Vasishtn
Leighton K Chong
Dong Zhendong
Wang Qixiang
Jungyun Seo
Keh Yih Su
John Burton

(法人会員)

日揮株式会社
NECTEC

横浜市南区別所1-14-1 (日揮/業務第2部文書サービス)
Ramavi Road Rajthevj Bangkok 10400 Thai land

☎045-712-1111

☎662-248-8078-84

委員会活動

- 理事会 92年第2回 (3/19) ①93年度事業計画②93年度収支予算③事務所移転④通常総会開催案
93年第2回 (5/14) ①92年度事業報告②92年度決算報告③理事の交替④IAMT 対応策
- 運営委員会 92年度11回 (3/11) ①93年度事業計画②93年度収支予算③事務所移転
93年度1回 (4/13) ①92年度収支決算②特別会計報告③協賛案件
2回 (5/12) ①監査報告②研究成果発表会開催案③IAMT 対策
- 財務委員会 (4/21) ①会員入退会現況確認②実行予算確認③監査対応
- 編集委員会 (3/24) ①No.3 編集方針②No.2 英文版確認③ニュースレター対応
- 企画広報委員会 (3/26) ①翻訳フェア収支結果報告
- システム評価研究会 (4/8) ①翻訳困難例文取りまとめ②学者コメントと対応策③成果発表会対策
- システム利用研究会 (3/12) ①ヒアリング②アンケート実施要綱
(4/15) ①ヒアリング②アンケート調査実施方法確認③取りまとめ日程検討
(4/23) ①アンケート集計集中作業②集計概要取りまとめ
(5/12) ①ヒアリング②アンケート集計中間報告③業界用語調査結果まとめ
- 制限言語研究会 (3/29) ①研究方向再確認②研究スケジュール
(4/14) ①小日本語研究②項目別検討 (宿題)
(5/11) ①小日本語研究②項目別検討③サーバ化検討
- MTサミット実行委員会 (3/2) ①テクニカルツアー②IAMT 総会対応
(4/13) ①実行予算案見直し②各委員会動向調整
(5/14) ①実行予算案②申込状況確認③補助金状況確認
- プログラム委員会 (3/24) ①プログラム修正確認②プロシーディングフォーマット検討
- 展示委員会 (4/12) ①出展社確認②技術展示対策③レイアウト検討④小間割抽選会開催案
(5/28) ①展示小間割り抽選会②出展詳細説明
- 広報委員会 (3/16) ①動員目標検討②協賛団体対応③2nd アナウンスメント送付先検討
(5/18) ①参加申込状況確認②海外動員策検討

《お詫びと訂正》

前号に誤りがありましたので訂正します

頁行	(誤)	(正)
14 2	言語処理部	削除
15 44	鳴海武司	鳴海武史
23 44	伊東昌高	伊藤昌高
25 13	7/20パンケット	7/21

AAMT ジャーナルNo.3 (May 1993)

発行	アジア太平洋機械翻訳協会 (略称 AAMT)
所在地	〒107 東京都港区赤坂7-2-17-305 ☎ 03-3479-4396/4398 FAX 03-3479-4895
編集委員会	野村浩郷 (委員長) 鳴海武史 永野文美 杉山健司 信田恵彦 亀井真一郎 宮城雅之

第 4 回 機 械 翻 訳 サ ミ ッ ト

会 期 平成 5 年 7 月 19 日～22 日

場 所 ホテル オークラ 神戸

※シンポジウム国際会議 (7/20～22)

第 1 日 招 待 講 演 「機械翻訳技術の最新動向」

プロジェクト 「CICC 多言語機械翻訳システム」「Verb Mobile」

パネル 討 論 「機械翻訳技術の将来」

歓迎レセプション (18:30～20:00)

第 2 日 招 待 講 演 「機械翻訳利用技術の現状」

プロジェクト 「音声自動翻訳電話」「EC Language Project」

パネル 討 論 「機械翻訳の評価法」

バンケット (18:30～20:00) 有料@15,000

第 3 日 招 待 講 演 「電子化辞書から大規模知識ベースへ」

プロジェクト 「コーパスプロジェクト」「電子化辞書」

パネル 討 論 「国際協力」

IAMT (機械翻訳世界連盟) 総会

◇ワークショップ (7/19 14:00～17:30)

テーマ 「機械翻訳のための知識構築」

◇チュートリアル (7/19 14:00～17:30)

テーマ 「機械翻訳の効果的な利用法」

①機械翻訳のメカニズム

②機械翻訳のケース・スタディ

◇テクニカルツアー (7/19 10:00～)

(株)ATR 音声翻訳通信研究所 見学

※機械翻訳展示会 《7/19 (13:00より)～7/22 (14:00まで)》

出展団体：KAIST/CAIR (韓国)・国際情報化協力センター・日本電信電話・
日本電気・日立製作所・東芝・富士通・シャープ・松下電器・ATR
音声翻訳通信研究所・日本電子化辞書研究所・リコー・カテナ・ロ
ゴヴィスタ・ノヴァ・CSK

主催・アジア太平洋機械翻訳協会(社)日本電子工業振興協会

