



# AAMT

The Asia-Pacific Association for Machine Translation

# Journal

No.18

march 1997

アジア太平洋機械翻訳協会

## 目 次

|        |                                                                             |    |
|--------|-----------------------------------------------------------------------------|----|
| 委員会報告  | 機械翻訳用電子化言語情報の共有に向けて                                                         | 1  |
| ヒアリング  | JICST 機械翻訳システムについて                                                          | 4  |
| 技術早分かり | 人間の気づかない曖昧さ                                                                 | 7  |
| 研究機関紹介 | 髙エイ・ティ・アール音声翻訳研究所 第三研究室                                                     | 11 |
| 留学生雑感  | 留学生から見た日本語                                                                  | 14 |
| 学会報告   | 電子情報通信学会 言語理解とコミュニケーション研究会<br>最近の研究発表内容の簡単なご紹介                              | 16 |
|        | 脳と用語シンボジュウム                                                                 | 17 |
| イベント報告 | CICC 多言語機械翻訳システムフォローアップ事業                                                   | 18 |
|        | ソフトウェア講演会と特許機械翻訳実験システムの公開                                                   | 19 |
|        | JTF 翻訳のつどい                                                                  | 6  |
| 交流会報告  | 第2、3回 JTF—AAMT 交流会                                                          | 20 |
| 委員会報告  | アンケート結果の報告                                                                  | 21 |
| 新製品紹介  | たちまち翻訳 英日翻訳支援システム for Windows95                                             | 22 |
|        | レターナビゲータ1.0                                                                 | 24 |
|        | シャープノートパソコン Mebius シリーズに搭載「英日翻訳支援ソフト（インター<br>ネット・英文メール作成支援対応）Power E/J 2.0」 | 26 |
| 新刊図書紹介 | インターネット英語の読み方&書き方&調べ方                                                       | 3  |
| 新聞雑誌記事 | 機械翻訳に関する新聞、雑誌記事                                                             | 6  |
| イベント紹介 | テクニカルコミュニケーションシンボジュウム'97のご案内                                                | 18 |
|        | 佐良木事務所・機械翻訳研究室の開設                                                           | 21 |
| 委員会だより | “会員の皆様からの投稿をお待ちしております”                                                      | 28 |
| 事務局だより | 協会活動報告                                                                      | 19 |
|        | “MT Summit に参加しよう！”                                                         | 28 |

# 機械翻訳用電子化言語情報の共有に向けて

技術動向調査委員会 村木一至 (NEC)、亀井真一郎 (NEC)

## 1. はじめに

WWWの浸透、あらゆる場面でのPCの活用など、様々な知識が電子化され、ネット環境下に分散して蓄えられつつある。ワープロを用いてネット上で分担共同して翻訳することも容易になった。翻訳ソフトを使って、気軽に英文ページをブラウズしたり、翻訳作業の効率化に役立てることも気楽にできる時代である。けれども？翻訳ということに話を限れば、依然としてコンピュータやネットワークが主役なのではなく、人間が、その知恵を出す頭脳が主役であることは変わっていない。

翻訳が人間主体の知的作業であるなら、人間の翻訳の知恵を互いに融通するために、PCやネットワークを大いに利用しようということは過去から言われてきた。ところが？融通するための、知恵や知識を電子化する手軽な手段がなく、更にその交換のためのソフトインフラも無いということが続いてきた。

翻訳ソフトウェアがPCの台数だけ出荷されるようになると、自分の知恵をユーザ辞書として登録して少しでも自分専用に賢く使い込むユーザも現れてくる。実際、なんらかのチューンアップをしなければ、翻訳ソフトもすぐお払い箱となりかねない。

翻訳に使用される辞書データの流通・相互利用の仕組みを作るべく、当協会に加盟している機械翻訳メーカーが中心となり、各社の機械翻訳システムのユーザ辞書と相互変換可能なユーザ辞書共通フォーマット (Universal Platform: UPF) の開発を情報処理振興事業協会 (IPA) 殿の支援をうけてスタートした。

そこで、本活動について概要を紹介する。

## 2. UPF 設定の背景

商用の機械翻訳 (MT) システムが製品化されて10年余が経過した。近年のインターネットの発達とパソコンの普及により、MTシステムは一般ユーザにも急速に浸透し始めている。

MTシステムを有効に活用するためには、各ユー

ザ毎に、ユーザ自身が頻繁に使用する単語対訳を「ユーザ辞書」に登録しておくことが望ましい。ところが、仮名漢字変換用ユーザ辞書登録でさえ面倒なように、翻訳用対訳辞書登録も面倒な事で、一人一人がユーザ辞書をそれぞれに充実させるには限界がある。

この問題に対して、個人が個別に持っている辞書データを流通させる機構を作ろうということになった。現在は国内の20数社が翻訳ソフトを商品化している。個人の興味の範囲と翻訳ソフトの違いを越えて、共通にユーザ辞書を交換できれば、知恵の足りない翻訳ソフトに皆から借りた知恵を詰め込んで、英語ともつき合える日本人が増えるのではと考えた。翻訳ソフトの利用が促進され、ひいては日本人の外国語文書受発信が促進される。

このような環境整備の具体的活動として、AAMTでは、今年度 (平成8年度) から来年度にかけ情報処理振興事業協会 (IPA) の創造的ソフトウェア育成事業の一つとして、万国共通なユーザ辞書記述フォーマットUPFの開発を開始した。この活動はAAMTに参加しているMTメーカー5社が知恵出しの中心となって進めている。

## 3. UPF 設定の基本方針

異種システム間で辞書データを交換できるようにするに、以下のような環境を想定した。

- (1) 共通フォーマットの設定
- (2) 共通フォーマットと各システムのユーザ辞書の間の双方向コンバータの提供
- (3) 共通フォーマットで記述された辞書を蓄え流通させるための、一般アクセス可能な電子環境の提供

上記 (1) のフォーマット開発は、現実に発売・利用されている複数のシステム間でそのユーザ辞書情報を比較することにより行なうこととした。この方針をとることで、共通フォーマットが現実のシステムから乖離してしまう危険を回避できると考えて

いる。このフォーマット開発活動は MT メーカー 5 社が中心となって UPF 実行連絡会を構成し、そこで原案を作成して各社に案の可否を諮るという手順で進めている。

上記 (2) の双方向コンバータは、各 MT メーカーがそれぞれ独自に開発するものである。共通フォーマット設定の際には、共通フォーマットから各システムのユーザ辞書への変換（ダウンロード）と、各システムのユーザ辞書から共通フォーマットへの変換（アップロード）との両方が可能となるよう考慮する必要がある。

上記 (3) の辞書共有環境としては、AAMT のホームページを中心に考えている。各ユーザはそこに置かれている UPF で記述された辞書データを自分の使っている MT システムのユーザ辞書のフォーマットに変換して使用することができる。また逆に各ユーザが自分の使用している MT システムで作成したユーザ辞書は、UPF 形式に変換してこの辞書共有環境に置き他ユーザ（異システムユーザも含む）と共有することができる。もちろん SGML のテキストとして、e-mail で直接ディレクトリリングすることを可能にしたい。

また、直接 UPF 形式で辞書を記述するための辞書エディタも提供する。

UPF の設定にあたっては、さしあたり日本語と英語の 2 カ国語を分析対象として具体的作業をすすめるが、UPF の形式は言語によらず多言語に対応できるものとする 것을 目指している。また一つの語の生成と解析の辞書はできる限り統一した形式で記述できるように仕様開発することとしている。具体的な記述形式は SGML に準拠したタグを用いる。

#### 4. UPF 設定の具体的活動

##### 4-1. 基本変換標準と拡張変換標準

上述したように UPF は、UPF から各システムのユーザ辞書への変換（ダウンロード）と各システムのユーザ辞書から UPF への変換（アップロード）との双方向が可能となるようにする必要がある。

現実に発売・利用されている複数のシステムがユーザ辞書として扱える語彙（品詞）には相違があるから、狭い意味で上記の双方向条件を満たすためには、各システムに共通して記述できる語彙（各システムに記述可能な語彙の「AND」）だけを対象範囲

とした仕様を決める必要がある。

一方、そのような「AND」仕様だけでは、詳細あるいは広範な語彙情報の記述を許すシステムが有効に活用されないという問題が生じる。そこで上記の両方の要求を満たすため、UPF では以下の 2 種類の対象範囲を設定することとした。

##### (a) 基本言語変換標準

全 MT システムのユーザ辞書で取り扱うことができ、UPF との間で相互変換（アップロード・ダウンロード）可能であることを推奨する語彙の情報を記述する形式

##### (b) 拡張言語変換標準

各 MT システムで記述する可能性のあるすべての語彙の情報を記述する形式

すでに商品化されユーザに使用されている 5 つの異なるシステムのユーザ辞書を比較検討して、現在上記 2 種の仕様作業を進めている。基本的には第 1 年度に基本言語標準を、第 2 年度に拡張言語変換標準を開発するが、基本言語標準の開発に際しても拡張言語標準を考慮する必要があるのは言うまでもない。また拡張言語標準の開発段階で再度基本言語標準に戻って改良する必要があるかも知れない。

##### 4-2. 基本言語変換標準の概要

基本言語標準に関して、まず各システム共通の「記述用語」の定義を行なう必要があった。つまり、品詞のセット、品詞の呼称から用語と定義の統一をする必要があった。特に日本語の場合、学校文法の品詞設定が機械翻訳にとってはなはだ不備であり、品詞設定に準拠すべき標準が存在しないからである。具体的には「形容動詞」という品詞を独立の品詞として立てているシステムと「形容詞」の下位として扱っているシステムが存在する。またその登録単位も語幹登録、終止形登録の二通りがあった。このような用語・形式の統一を行ない、基本言語変換標準としては、現在のところ以下を対象とすることで原案作りをすすめている。

日本語：名詞、固有名詞、動詞、サ変動詞、  
形容詞、形容動詞、副詞

英語：名詞、固有名詞、動詞、形容詞、副詞  
上記の品詞設定は、実際に現在までの 10 年間に MT ユーザが登録したユーザ辞書の約 9 割が名詞、固有名詞であるという各社に共通した経験的データに基

づいている。名詞の下位分類とすることも可能な固有名詞を独立の品詞として立てたのは、ユーザが高頻度で登録する可能性のある品詞であることがその理由である。和語動詞の登録の簡便さを考慮し、それとの一貫性をとって、活用語はすべて終止形で見出し登録する。このようにするとサ変動詞の語幹名詞用法は名詞として、活用語尾付き用言用法は動詞として登録すればよいが、固有名詞と同じように、サ変動詞が高頻度登録されるという理由から動詞の下位とはせず、独立した品詞として考えることとした。

#### 5. これから

翻訳用対訳辞書共通フォーマット (UPF) 開発活

動の概要について述べた。UPF の仕様としては、基本標準、拡張標準の2つを設定する予定である。第1年度末である現時点では、基本標準の仕様をほぼ固める目処がついた。今後は各システムのユーザ辞書との間の双方向変換の確認を行なうのと並行して、拡張標準の設定を行ない、来年度末までに基本・拡張の両仕様を FIX して一般に公開する予定である。最終案に至るまえに会員のみなさんや、識者からのご意見をいただく機会をなるべくはやい時期に設定する計画である。是非とも、ご支援を賜るよう御願います。

なお、97年度早々には、UPF 公開討論の為の Home Page も開設する事が決まっております、次号機関誌にてご案内したいと考えます。

#### 新刊図書紹介

### インターネット英語の読み方 & 書き方 & 調べ方

著 者： 安藤 進  
出版社： 共立出版  
価 格： 2,000円

発 行： 1997年3月10日  
ISBN4-320-07139-5C3050P3502

#### 本書の特色

インターネットの時代に即応した斬新な英語学習法を披露する。最新の英文表現辞典として Web を活用する手法、大手企業のホームページの読み方、サーチエンジンを利用した検索手法、メーリングリストやニュースグループの購読法などを紹介する。特に、電子メールの書き方はすぐ実際に応用できる表現集として役立つ。

#### 対象読者

研究者、技術者、学生、翻訳者、ビジネスマン、その他インターネット英語に興味を持つ一般読者。

#### 目次と概要

- 第1章 インターネット英語とは
- 第2章 インターネット英語に挑戦してみよう
- 第3章 インターネット英語の読み方と書き方のツボ
- 付 録 インターネット関連の参考情報

#### 連絡先

〒192-03 東京都八王子市上柚木3-5-6-301

Phone : (0426) 75-0182 Fax : (0426) 75-0183 Nifty-serve : GFG02221 E-Mail : sando@twics.com

# JICST 機械翻訳システムについて

科学技術振興事業団 科学技術情報事業本部 情報システム開発部 白木澤 佳子

## 1. はじめに

日本科学技術情報センター（JICST）は昨年10月に新技術事業団と統合し科学技術振興事業団（JST）として新しくスタートしました。

本稿では当事業団でJICST時代から引き続き開発、運用を行っているJICST機械翻訳システムについてご紹介します。

## 2. 開発の経緯

JSTでは、μプロジェクトの成果を元に昭和61年度から平成2年度にかけて日英機械翻訳システムの実用化を行いました。それより以前から日本の科学技術情報を海外に紹介するためにJICST-Eファイルという英語のデータベースを提供していましたが、平成2年夏以降、JICST-Eファイルに搭載する英語の標題や抄録を作成するために、日英機械翻訳システムを大型計算機で運用してきました（図1）。このシステムを元に平成5年度から7年度にかけて、ワークステーションおよびパソコン上で動くシステムを開発しました。ワークステーション版システムは英文標題・抄録作成のために利用している他、「翻訳ネットワークサービス」での翻訳処理に使用しています。パソコン版システムはWindows及び

Macintoshに対応したシステムを作成し、現在JST内部で利用しています。また、アメリカのワシントンに昨年5月に開設された日英科学技術文献機械翻訳センターでもMacintosh版システムを使用しておりアメリカ国内での日本情報入手の一助となっています。

## 3. 翻訳ネットワークサービス

JSTではワークステーション版システムを活用して平成8年3月より翻訳ネットワークサービスの提供を開始しました。これはインターネットのメール機能を利用したもので、利用者が日本語で書いたメールを送ると英語に翻訳された返信メールが返されてきます（図2）。

申し込みは無料で、翻訳料金は翻訳結果（英字）1文字につき1円となっています。前編集や後編集のサービスは行っていない。

サービスの詳細については以下のURLを参照してください。

<http://www-jmt.jst-c.go.jp/>

このホームページでは「お試し翻訳」も用意していますので、実際にどのような翻訳結果が返されてくるかホームページ上で試すことができます。翻訳ネ



図1 英文票題・抄録の作成



図2 翻訳ネットワークサービス



ットワークサービスではA4用紙1枚程度の文章を2～3分で処理しますが、実際に翻訳結果が利用者の手元に届くまでにはインターネット上での転送にかかる時間も考慮する必要があります。

#### 4. パソコン版システム

対象機種は、現状のパソコンの普及、発展動向を勘案してWindowsとMacintoshとしました。基本的な翻訳の質やシステム辞書の内容は大型計算機で運用していたシステムと同様とし、翻訳速度も市販のパソコン版翻訳ソフトウェアと同等以上であることを開発前提としました(図3)。

科学技術用語を中心とした40万語のシステム辞書と20万語の医学用語辞書(オプション)を最大の特徴としています。

Windows版、Macintosh版それぞれの動作環境は以下の通りです。

##### Windows版

- ・ Pentium 搭載の PC/AT 互換機
- ・ メモリ16MB 以上
- ・ ハードディスク240MB 以上
- ・ Windows95又は Windows NT

##### Macintosh版

- ・ Macintosh680X0 or Power PC
- ・ メモリ16MB 以上
- ・ ハードディスク250MB 以上
- ・ 漢字Talk7.1 or 7.5

パソコン版は次のような基本機能をサポートしています。

##### 1) 読み込み

日本語のテキストデータを読み込みます。読み込

まれた文章はシステムが自動的に句点(。)を認識して1文毎に区切ります。一つのファイルとして読み込める文は最大200文です。勿論、翻訳画面上で翻訳したい文を直接入力することもできます。

##### 2) 入力文の編集

読み込まれた文章を、機械翻訳に適するように人手で編集することができます。カット、コピー、ペースト、文の挿入・削除、文の結合・分割、検索/変更等の機能があります。

##### 3) 翻訳の実行

日本語の文章を英語に翻訳します。全文を一度に翻訳する他に、指定した範囲だけを翻訳することもできます。

##### 4) 翻訳文の編集

翻訳して得られた英文の編集を翻訳画面上で行うことができます。

##### 5) 翻訳文の印刷

翻訳結果をプリンタに出力します。画面イメージの印刷とテキスト形式の印刷が選択できます。

##### 6) 翻訳文の保存

翻訳結果をフロッピーディスクやハードディスク等に保存します。

JICST 機械翻訳システムの第1の特徴として、語数の豊富なシステム辞書を用いることにより、特に科学技術・医学関連の文献の翻訳に優れた結果を得ることができます。

またユーザ辞書には名詞、動詞、形容詞、形容動詞を登録することができます。英語に翻訳する際に必要な情報を入力する必要がありますが、この項目を最低限に絞って簡単に登録できるようにしています。

更に、科学技術文献以外の幅広い利用に対応するために、例えば、入力文の主語がない場合に、(a) 主語を補う (I, it...), (b) 受身文にする、(c) 命令文にする、というような文体の選択 (訳し分け) を可能にしました。これを訳し分け機能と呼んでいます。「訳し分け条件設定」画面から適した条件を選んで設定することができます。

翻訳を実行した後では、日本語に対応する英語または逆に英語に対応する日本語を翻訳画面上で表示させることができます (対訳対応)。また、システム辞書中に他の訳語が登録されている場合はその候補

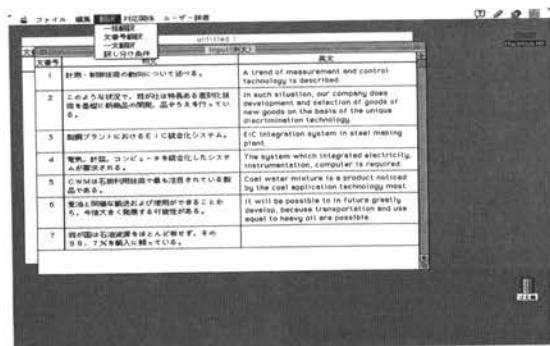


図3 パソコン版システム

語を表示させて、その中から適切な訳語を選択して変更することもできます（訳語変更）。

#### 5. おわりに

機械翻訳システムはまだ改良の余地がありますが、今後ますますその必要性が高まると考えられます。JSTではワークステーション版システムの運

用を行っている他、平成7年秋よりパソコン版システムをデータベース作成や翻訳サービスのサポートに使用しており、システム辞書の拡充を定期的に行う運用体制を作っています。ユーザの要望を積極的に取り入れて、更に使いやすく、翻訳の質が高い機械翻訳システムを開発していきたいと思います。

---

### イベント報告

## JTF 翻訳のつどい

日本翻訳連盟（JTF）主催の第4回翻訳のつどいが、'97年2月21日（金）、大阪日経本社で開催された。今回のテーマは「ネットワーク時代の翻訳ビジネス」で、講演、フォーラム、交流会の3つのセッションに約50名が参加、翻訳とビジネスに関するノウハウとアイデアの交換が行われた。

翻訳者宮代紘一氏は、「ネットワーク時代の翻訳者のライフスタイル」についての講演で、コンピューターを活用しなければ翻訳者は世間並の安定した収入は得られないと断言し、具体的な活用事例を披露された。

フォーラム「翻訳ビジネス最前線」では、4人の翻訳会社経営者が自社のビジネス戦略を披露し、参加者との討論を行った。4社とも異色の企業で、個性的で多彩なプレゼンテーションであった。

交流会は、翻訳会社と翻訳者の交流コーナー、新人カウンセリングコーナー、業界情報コーナーの3つに分かれて開催された。交流コーナーでは、翻訳者から「納品した翻訳に関するフィードバックがない」という注文が、翻訳会社からは「依頼した仕事に関してもっと質問してほしい」という注文が出た。新人カウンセリングコーナーには、就職カウンセリングのために女子大の先生が参加されており、翻訳産業の裾野の広がりつつある予兆が伺われた。

（日本翻訳ジャーナル2／1997より）

---

### 新聞雑誌記事

## 機械翻訳に関する新聞、雑誌記事

最近、新聞や雑誌に機械翻訳に関する記事が次々と出ている。インターネットの普及と多数の翻訳ソフトの出現により、機械翻訳に対する関心が高まってきた結果と思われる。いずれも、特定の文章を市販の機械翻訳ソフトで翻訳して比較したものである。

- |                                |                |                        |
|--------------------------------|----------------|------------------------|
| ・ 翻訳ソフトまだまだ 購入前に自分で試して         | 朝日新聞           | '96年11月22日夕刊           |
| ・ 英日翻訳ソフト14本のホントの実力            | 月刊アスキー         | '96年7月号                |
| ・ 翻訳ソフトの英語力を試す                 | 日経トレンディ        | '96年12月号臨時増刊号          |
| ・ 英文ホーム・ページを上手に訳すブラウザ翻訳ソフトはどれか | Internet User  | Vol. 2 No. 12 Dec 1996 |
| ・ 低価格英和翻訳ソフトの実力をみる             | DOS V magazine | '97-2.1                |



# 人間の気づかない曖昧さ

シャープ（株）情報商品開発研究所 吉見毅彦

## 1 はじめに

ある一つの言語表現に対して二通り以上の解釈が可能であるとき、その表現は曖昧であるといわれます。人間による翻訳と機械翻訳システムによる翻訳を比較するという観点からは、曖昧な表現を次の二種類に分類することができます。

（１）人間にとっても、機械翻訳システムにとっても曖昧な表現。

（２）人間にとっては曖昧ではないが、機械翻訳システムにとっては曖昧な表現。

（１）に分類される表現には、「豚は食べない。」という文などがあるでしょう。「豚は」が「食べない」の主語なのか目的語なのかを、この文だけでは人間でさえ決定することができないという曖昧さを含んでいるからです。これに対して、（２）に分類される表現には、「彼は好き嫌いが激しい。豚は食べない。」という二つの文から成るテキストなどがあるでしょう。このテキストでは、「豚は食べない。」の解釈を「一つ目の文で『彼』と表現されている人物が豚肉を食べない。」とすることは人間にとっては容易であるのに対して、現在の商用機械翻訳システムにとってはそうではありません。

今回の「技術早わかり」では、「彼は好き嫌いが激しい。豚は食べない。」というテキストの解釈を機械翻訳システムはなぜ一意に絞り込むことができないのかを、人間がこのテキストを翻訳するときの振舞いと比較しながら説明します。さらに、（２）に分類される他の表現の例をいくつか挙げます。

## 2 機械翻訳システムの現状

この節では、現在の商用機械翻訳システムがどのような枠組みで翻訳処理を行なっているのか、また、なぜそのような枠組みになっているのかを簡単に説明します。

### 2.1 一文ごとの翻訳

現在の商用機械翻訳システムでは、二つ以上の文

から成るテキストが入力として与えられると、テキスト全体を一括して処理するのではなく、まず一文切り出しと呼ばれる処理を行ない、この処理によってテキストから文を一つずつ取り出した後、各文毎に独立に翻訳を行ないます。例えば、「彼は好き嫌いが激しい。豚は食べない。」というテキストの二つ目の文を翻訳するとき、人間は一つ目の文が存在するという前提の下で処理を行ないますが、機械翻訳システムは一つ目の文は存在しないものとして処理を行ないます。

一文切り出しを行なう理由は、翻訳に要する時間を短縮するためです。あるテキストが  $m$  個の文から構成されており、各文について  $n$  通りの解釈が可能であるとします。テキスト全体として可能な解釈の数は、一文切り出しを行なわない場合、一般に  $n$  の  $m$  乗に比例します。これに対して、一文切り出しを行ない各文が独立であるとみなすと、 $n$  の  $m$  倍に比例します。例えば、 $m=5$ 、 $n=10$  の場合、一文切り出しを行なわないと 100000 通りになるのに対して、行なうと 50 通りにしかありません。このように、一文切り出しを行なうことによって考慮すべき解釈の数が減るので、翻訳時間を短縮することができます。

### 2.2 浅い知識に基づく翻訳

翻訳を行なう際には、日本語や英語に関する言語的な知識はいうまでもなく、翻訳対象分野に関する専門的な知識や一般常識などの非言語的な知識もまた重要な役割を果たします。しかし、現在の商用機械翻訳システムに与えられている知識は、言語に関する知識のごく一部に限られており、人間が持っている専門的な知識や一般常識などは与えられていません。従って、機械翻訳システムは、構文的な知識あるいは意味素や選択制限などの狭義の意味的な知識を拠り所として、文の解釈を一意に絞り込まなければなりません。

人間が持つ専門知識や一般常識を機械翻訳システ

ムに与えることは、現時点では、極めて難しい問題です。その理由として、まず、知識表現形式の設計に関する問題が挙げられます。専門知識や常識を機械翻訳システムに与え、翻訳の過程で有効に利用するためには、どのような知識をどのような形式で表現しておけばよいか、さらに、どの処理段階でどのように用いればよいかを明らかにしなければなりません。しかし、現時点では、そのような知識の定式化が十分には行なえていません。

もう一つの理由は、たとえ知識の表現形式が設計されたとしても、コンピュータ上に実装する際の問題が残されているという点にあります。商用機械翻訳システムの辞書には、数万から数十万の見出しが登録されているので、各見出しに与える個々の知識がたとえ単純なものであっても、辞書全体としては膨大な量となり、知識の実装には非常に多くの労力と長い時間を要します。また、新たな種類の知識を追加するとき、既に実装されている知識との間に矛盾が生じないように知識の整合性を維持することにも注意を払わなければなりません。

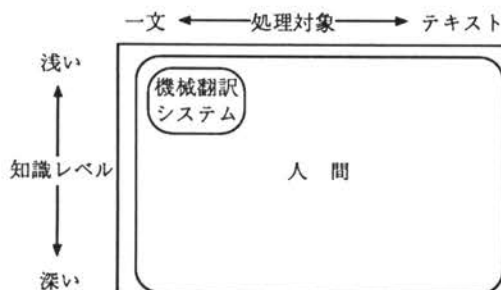
### 3 人間の気づかない曖昧さ

この節では、人間にとって曖昧でない表現が、2節で説明したような枠組みの機械翻訳システムにとってはなぜ曖昧となるのかを説明します。人間による解釈の曖昧さの解消も機械翻訳システムによる曖昧さの解消も、何らかの評価基準に従って可能な解釈に優劣を付け、その中から最も妥当な解釈を選択することによって達成されと考えられます。ある表現の最も妥当な解釈が一つだけ存在し、その解釈とそれ以外の解釈との優劣の差が大きい場合には、その表現は曖昧ではないとみなせ、それ以外の場合には、曖昧であるとみなせます。人間には解釈の優劣の差を付けることができるような表現に対して、機械翻訳システムには優劣の差を付けることができない主な原因は、次の二点にあります。

(1) 人間は文脈から得られる手がかりも解釈の評価基準として用いることができるのに対して、機械翻訳システムはできない。

(2) 人間は言語的知識だけでなく専門知識や一般常識などの非言語的知識も解釈の評価基準として用いることができるのに対して、機械翻訳システムはできない。

人間と機械翻訳システムの違いを概念図で示すと次のようになるでしょう。



#### 3. 1 一文ごとの処理に起因する曖昧さ

「彼は好き嫌いが激しい。豚は食べない。」というテキストを英語に翻訳することを考えます。話を簡単にするために、一つ目の文には曖昧さは含まれていないものとします。二つ目の文を翻訳する際には、次に挙げる四種類の曖昧さを解消しなければなりません。

(1) 「豚は」が「食べない」の主語かあるいは目的語かという構文的な曖昧さがあります。

(2) 日本語では、ある動詞の主語や目的語などは、それが何であるかが文脈からわかる場合には、明示的に表現されないのに対して、英語では、適切な代名詞として表現しなければなりません。従って、「he」や「they」などの代名詞のうち、どれを選ぶかを決定しなければなりません。

(3) 「豚」の語義には、「動物の豚」(英語では「pig」と「豚肉」(「pork」)があり、どちらの意味で用いられているのかという曖昧さがあります。

(4) さらに、「豚」という名詞に関して次のような曖昧さがあります。日本語では、名詞の数(単数形か複数形か)や定/不定性(名詞が特定のものを表しているのか、不特定のものを表しているのか)が明示されることは希であるのに対して、英語では、名詞の数や定/不定性(名詞に定冠詞を付けるか、不定冠詞とするか、あるいは何も付けないか)を明確に区別しなければなりません。

(1)～(4)の曖昧さは、独立ではなく、次のように関連しています。

[1と3] (1)で「豚は」を「食べない」の主語と解釈した場合には、(3)で「動物の豚」とい

う意味で用いられているとみなし、目的語と解釈した場合には、「豚肉」という意味で用いられているとみなすのが自然です。なぜなら、通常、「食べる」の主語になれる語句の語義は＜動物＞であり、目的語になれる語句の語義は、＜食べ物＞であるからです。

〔1と2〕(1)で「豚は」が「食べない」の主語であると解釈すると、「食べない」の目的語が表現されていないことになり、逆に、「豚は」が目的語であると解釈すると、主語が表現されていないことになります。〔1と3〕で説明したのと同じ理由で、「食べない」の主語と目的語のどちらが表現されていないと解釈するかによって、(2)で選択すべき適切な代名詞の候補が異なってきます。

〔3と4〕(3)で「豚」の意味を「豚肉」と解釈した場合には、“pork”は不加算名詞ですから、(4)で単数形か複数形かの区別を考慮する必要はありません。これに対して、「動物の豚」と解釈した場合には、最終的に“pig”を単数形とするか複数形とするかを決定しなければなりません。

人間が(1)～(4)の曖昧さを(ほとんど)意識せずに解消し、解釈を一意に絞り込むことができる理由は、可能な解釈に対して優劣を付けるための評価基準として、ある文で表現されるべき情報がそこで表現されていないければ、それは文脈で表現されている可能性が高いという経験則を用いているからだと考えられます。このような評価基準の下では、表現されるべき情報を文脈から復元することができる解釈は、そうでない解釈よりもより妥当であると判断できます。従って、「豚は食べない。」という文で表現されていない主語または目的語は、一つ目の文に現れる「彼」であろうということがまず推測できます。「彼」が「食べない」の主語であるか目的語であるかを決定するのは容易です。これで(2)の曖昧さは一意に絞り込まれます。「彼」を「食べない」の主語と解釈したことによって、(1)と(3)の曖昧さは、自然に解消されます。(3)で「豚」を“pork”と翻訳することから、(4)の名詞の単数形／複数形に関する曖昧さは生じません。このように、人間は、二つ目の文を翻訳するとき、テキストを構成する二つの文の間には何らかの関連があるとみなし、一つ目の文との整合性に基づいて文の可能な解釈に優劣を付け、最も妥当な解釈を選択します。

この結果、“He has strong likes and dislikes. He does not eat pork.”のように翻訳することができます。

これに対して、機械翻訳システムはどうでしょうか。2. 1節で、現在の商用機械翻訳システムでは一文切り出しという処理が行なわれることを説明しました。一文切り出しを行なうと、二つ目の文を翻訳するときに、一つ目の文から得られる情報を参照することができなくなるので、(1)～(4)の曖昧さを解消するための手がかりが得られなくなります。1節で説明したように、「彼は好き嫌いが激しい。」という文が存在しないものとする、「豚は食べない。」は人間にとっても曖昧な文です。この文に対して可能な翻訳としては、“A pig does not eat them.”や“He does not eat pork.”、“They do not eat pork.”など、十通り以上の候補を挙げることができます。これらの候補の中からどれが選ばれるかは、個々の機械翻訳システムに与えられている構文的な知識や狭義の意味的な知識に依存しますが、これらの知識だけでは、“He does not eat pork.”という翻訳が他の翻訳よりも妥当であると判断することはできません。従って、人間にとって妥当である翻訳が得られる可能性より、例えば、“He has strong likes and dislikes. A pig does not eat it.”のような、人間には思いもよらないような翻訳が得られる可能性のほうが高くなります。

### 3. 2 浅い知識に基づく処理に起因する曖昧さ

言語表現の解釈は、何らかの文脈が存在して初めて定まるものなので、文脈が存在しなければ、人間でさえ曖昧さを感じる場合があります。しかし、文脈が存在せず、曖昧さを解消するための手がかりが文脈から得られない場合でも、人間は、多くの場合、専門知識や一般常識など様々な知識を用いて、可能な解釈の中から（特別な文脈が存在しないという条件の下で）最も妥当な解釈を選択することができます。これに対して、機械翻訳システムは、2. 2節で説明したように、人間が持つ言語的知識のごく一部しか与えられていないので、人間から見ても妥当でない解釈を選択することもあります。ここでは、特別な文脈が存在しない場合に、最も妥当な解釈の一つに絞り込むことが人間には容易であるが、機械翻訳システムにはそうではないような語義の曖昧さを含む文の例と構文的な曖昧さを含む文の例を示しま

す。

語義の曖昧さの典型的な例の一つは、ある名詞が普通名詞として用いられているのか、それとも、固有名詞として用いられているのかという曖昧さでしょう。極端に言えば、あらゆる普通名詞が固有名詞としても使われます。例えば、次の例文1に現れる「両国」という名詞は、普通名詞としての語義(“both countries”) で用いられていますが、例文2に現れる「両国」は、固有名詞の語義(“Ryogoku”) で用いられています。

(例文1) 日本とイギリス両国の首相が共に記者会見に臨んだ。

(例文2) 大相撲初場所は12日から両国の国技館で始まる。

「両国」がどちらの語義で用いられているかを人間が判断できるのは、判断の根拠となるキーワードを発見することができ、さらに、そのキーワードと「両国」との構文的な関係などを解析できるからだと考えられます。例文1で「両国」の語義を“both countries”と解釈する根拠となったキーワードは、「日本」や「イギリス」、「首相」であり、例文2で「両国」の語義を“Ryogoku”と解釈する根拠となったキーワードは、「大相撲初場所」や「国技館」でしょう。「両国」の語義を最終的に決定するためには、キーワードを発見するだけではなく、「両国」とキーワードの構文的な関係などを解析しなければなりません。なぜならば、「両国」と「大相撲初場所」や「国技館」が共に一つの文に現れていても、「両国」の語義を普通名詞と解釈しなければならない場合もあり、キーワードの発見だけでは、語義の曖昧さの解消の確かな手がかりとはならないこともありえるからです。このような処理を機械翻訳システムで行なうためには、どのような知識をどのように表現しておき、その知識をどの処理段階でどのように用いればよいのかは、現時点では明らかではありません。今後の研究課題の一つです。

次に、構文的な曖昧さの例を示します。

(例文3) 警官が、人質にピストルを突き付けながら、逃げる指名手配犯を追跡する。

(例文4) 警官が、捜査本部に応援を要請しながら、逃げる指名手配犯を追跡する。

例文3では、構文的には、「人質にピストルを突き付けながら」の主語は「指名手配犯」であるとい

う解釈と、「警官」であるという解釈が可能です。しかし、特別な文脈が存在しないという条件の下では、一般常識から判断して、人間にとっては、前者と後者の優劣の差は十分大きく、後者は極めて不自然な解釈となります。同様に、例文4では、「捜査本部に応援を要請しながら」の主語は「指名手配犯」であるという解釈は、人間にとっては極めて妥当性の低い解釈となります。人間と同じように、機械翻訳システムが二つの解釈の間に優劣の差を付けることができるようにすることは、原理的には、意味素と選択制限を用いる現在の機械翻訳システムの枠組みで可能かもしれません。しかしながら、実現するためには、意味素や選択制限を十分詳細に記述する必要がありますが、このような精密な記述を、一貫性や一般性を失わずに、辞書に登録されている数万から数十万の見出しについて行なうことは、実際問題としては難しいでしょう。

#### 4 おわりに

解釈の曖昧さを解消し、人間にとって最も妥当な解釈を選択することができれば、機械翻訳システムによる翻訳は、人間による翻訳に大きく一歩近づくに違いありません。このような品質の高い翻訳を出力することができる機械翻訳システムを実現するために解決しなければならない課題は数多く残されていますが、大学や企業において研究への積極的な取り組みが行なわれています。

近年、コーパスや人間用の辞書などの多くの言語資源が電子化されるようになるにつれ、翻訳に役立つ知識を言語資源から自動的に獲得するための研究が注目されてきています。知識の獲得が自動化されれば、様々な知識を、一貫性と一般性を保ちつつ機械翻訳システムに与えることが容易になり、質の高い機械翻訳システムが実現できるでしょう。また、ハードウェアの進歩や処理アルゴリズムの高速化によって、様々な知識を用いた複雑な処理やテキストを一括して扱う処理が効率良く行なえるようになってきています。このように、様々な知識を用いて文脈処理を行なえる機械翻訳システムが研究レベルではいくつか存在しますので、将来、商用システムとして実用化されることが期待されます。

# (株)エイ・ティ・アール音声翻訳通信研究所 第三研究室

第三研究室 室長 飯田 仁

## 1. ATR 音声翻訳通信研究所の概要

(株)エイ・ティ・アール音声翻訳通信研究所は基盤技術研究促進センターから出資を受けて93年3月末に研究活動を開始しました。7年間の期限付きで、自然音声を対象とする多言語の音声翻訳通信技術の基礎研究に取り組んでいます。研究の目的は、音声認識、言語翻訳、音声合成等の要素技術とその統合化技術を確立し、異なる言語間のグローバルコミュニケーションの実現に資することにあります。研究対象とする自然な話し言葉は、発声のくずれ、韻律情報による意図の表現、文法の枠にとらわれない日常的表現、状況に依存した省略や断片的な表現などを含むため、これを認識し、相手方の言葉に翻訳して、そして自然性の高い合成音声を出力するためには高度な音声言語処理技術が必要となります。このため、(1)自然音声認識、(2)発話韻律処理、(3)協調融合翻訳、(4)音声言語統合処理の4つのサブテーマを設定し、研究所を構成する第一から第四の研究室がこれらの研究テーマを担当しています。さらに、各研究室が一体となったシステム統合化グループが音声翻訳システムとしての取りまとめを行い、統合化の方式の検討や、予備実験を行っています。また、音声翻訳に関する国際コンソーシアム「C-STARII」の中核メンバーとして、99年の秋に予定している日韓英独を中心とする国際共同実験のデモの準備を各研究室のメンバーで構成する検討グループが進めています。

このコーナーでは、協調融合翻訳というサブテーマの下、自然音声の多言語翻訳を中心に研究している第三研究室の研究活動を紹介致します。

## 2. 協調融合翻訳の概要とアプローチ

自然な話し言葉の特徴である多様で豊富な言い回しや文法的に不備な表現を取り扱う点で、用例主導翻訳手法の効果が大きいことを幾つかの比較実験を通して示してきました。たとえば、

・日本語における助詞が陽に示されない発話の解

釈と翻訳

- ・接続助詞や格助詞などで中止する表現の解釈と翻訳
- ・語の意味的結合の制約をゆるめることにより可能となる比喩的な表現の翻訳
- ・英語における前置詞句修飾先の最適解釈と翻訳などの実験を挙げることができます。

複雑な句や複文などに対する翻訳を行うためには、句や文を構成する語の依存関係をボトムアップに解析しつつ、同時に対訳用例を適用しながら、最も好ましい用例を選択するメカニズムが望まれます。そのような機構を、“constituent boundary parsing”と呼ぶ解析手法で実現しました。その翻訳機構は、用例と規則という両端に位置する情報を統合的に取り扱うと共に、両者を共通の枠組みで取り扱うことができる融合方式になっています。これにより、プロジェクト当初よりもくろんでいた翻訳方式、つまり文法に基づく言語解析手法と、対訳用例を活用する用例主導翻訳手法とを協調融合していくという翻訳問題解決手法の骨格を作ることができました。

さらに、このメカニズムを核として、翻訳のための知識である表現例の言語パターンや対訳例などを翻訳の言語対ごとに用意することにより、多言語翻訳のプロトタイプシステムを実現しました。実際に作成したシステムは、日英双方向、日韓双方向、ならびに日独方向の“chat translation”として位置付けられるものですが、4か国語の合成音声を出力します。旅行アレンジ全般に渡る各種サービス窓口と旅行(計画)客との間の対話が対象となります。自由入力による訳文評価も行い、協調融合翻訳方式の能力や問題点の分析を進めながら、能力拡充を図っています。評価結果などについては後述致します。

## 3. 用例翻訳のための基本技術とその応用

### 3.1 意味的類似性

用例翻訳の中心となる技術は同じ言語間での表現の類似性を計算する用例検索であります。その一つ



の手法として、単語間の意味的な類似度の計算があります。その類似性を保証する物差しとなるのがシソーラスであり、階層構造の違いなどから角川類語辞典やEDR 概念辞書を使った実験をいろいろと試みています。階層構造が平衡していないシソーラスによる一般的な類似度を提案し、翻訳や情報検索などの具体的な応用の観点から有効性を確認しています。

また、相互情報量を使った単語クラスタリング法の改良などを行なって、シソーラスを自動的に作成して、単語間の類似度計算の精度向上を検討しています。

### 3. 2 構文的類似性

文や構文を含めた大きな単位での類似性も重要な研究テーマの一つであり、このような類似性を表す指標の一つとして編集距離に基づく類似度について検討しています。二つのデータを一致させるために必要な最小編集操作コストとして定義する編集距離を使って、類似度を高速に計算する手法と、言語処理に利用する手法との研究を進めています。後者に関しては、入力文に類似した文を構文木付きコーパスから検索し、得られた文の構文木から入力文に対する構文木を類推する手法を提案し、既存のコーパスを用いた実験で効果を確認しています。今後、部分構造間の類似性の導入や意味的類似性を用いた解析技術との融合についてさらに研究を進めていく予定であります。

### 3. 3 並列自然言語処理

用例翻訳では、翻訳時間の大部分を用例検索が占め、残りの多くは構文的な曖昧性に起因する解析の時間であることが分かっています。そこで、連想メモリを利用したデータの照合や書換えを並列に処理する試みや、SIMD 計算機用のアルゴリズムを開発し、10万件の用例からの検索で1ミリ秒以下の処理時間(20,000MIPSの逐次計算機相当)を実現し、大規模な用例に対するスケーラビリティを確認しています。

また、MIMD 計算機上で、協調融合翻訳の解析のベースであるチャート法の並列処理実験を進めています。

今後は、確率文法や構文的類似度を使った解析、

ならびに対話構造などを使った省略補完、訳語選択、多義解消などの文脈処理、並列処理技術を活用し、音声認識の曖昧な出力を効率的に取り扱える手法の研究も進めていく予定です。

### 3. 4 情報検索

電子図書館が提供する膨大なテキストなどから情報をうまく検索する技術の必要性は高いので、用例検索を情報検索へ応用する試みも進めています。例えば、格助詞を含む表現ボタン「XがY」の具体例「政府が提案する」を入力すると、「運輸省が・・・提唱する」などの類似した表現を含む記事が検索できます。

また、コーパスの自動分類の必要性からテキスト検索の試みも進めています。

### 4. 協調融合翻訳の現状

日英・日韓双方向翻訳においては、自由入力による各種の翻訳実験ができる体制ができています。平均約0.5秒で最適翻訳結果を自動的に一意に示すことができます。各言語に対して、システムにとって初見の約70から80の対話、1,200-1,300文ずつを対象に、ネイティブ11人による評価を行いました。この評価作業には多くの時間を要しますが、改訂版についても評価作業を続けています。日英翻訳では約70%の自動翻訳率を示すものもあり、また日韓翻訳では、学習用の収録データが少ないにも拘わらず、日英より10%以上良い翻訳結果を得ています。併せて、文法上正しくない表現に対しても最適と判断する翻訳結果が得られますので、その場合韓国語としては文法的に正しくない表現であっても意味的には内容が分かる翻訳結果を出力することができます。

翻訳の対象となっている旅行アレンジの対話には、外国人が学ぶ日本語会話の初級および中級程度までをカバーする表現や言い回しが含まれます。また、JEIDAの翻訳評価文やビジネスレターなどを対象とする評価実験も行い、話し言葉のみならず書き言葉に対しても良好な翻訳結果が得られています。

対訳用例などの学習用に使う対話を収録するに当たっては、日常的な問合せや依頼、情報提示などの表現が豊富に得られる旅行アレンジ対話に的を絞る、通訳者を介した対話場面の設定や、収録方法を定めています。それにより、当面必要になると考えられ



る約80万語の品詞タグ付き日英バイリンガル対話を既に収録・加工しています。

## 5. 音声翻訳への適用

不特定話者で、かつ連続音声を対象とする音声認識結果は、キーボードからの入力ミスや、類似する音のつながりから連想できる程度の単語列の誤りなどとは大きく異なる結果を生じる場合が多々あります。そのような不完全な文や句を翻訳するために、つぎのような2段階構えのアプローチを追究しています。

### 5. 1 認識結果を修正する処理

話者特有の音響的な誤り特性を学習しておき、この誤り特性を用いて認識結果に含まれない正解音素または語彙を探索する機構を検討しています。

すでに、認識誤り特性の学習および認識結果から正解音素系列の探索を行う一つの機構を提案し、具体化してきました。さらに、誤り特性が誤り部分の前後の状況に依存していることを明らかにすることができたので、音素単位の音響モデル(音素 HMM)を形成する「状態」の系列として誤り特性を捉えることにより、正解音素の探索性能向上を確認しています。

### 5. 2 音声認識誤り文から了解できる部分を積極的に翻訳する処理

協調融合翻訳の機構では、文が完結しなくとも、単語列の入力にしたがって句レベルの用例の適用が漸次進めるため、部分構造に対する訳が常に作られます。そのため、語句の依存関係を用例に基づいて解析することで、認識結果文中の正解部分を特定しその部分のみ抽出することが可能となります。

すでに、正解部分の特定に、認識結果と用例との間の意味的距離、及び依存関係がカバーする範囲の語句の長さを利用することで、正解部分の特定が高性能で行なわれることを確認しています。そして、この特定法を音声翻訳システムに導入し、特定された正解部分のみを翻訳することで、従来の誤り文に対する音声翻訳性能を大幅に向上できることが分か

ってきました。

## 6. 詳細情報をもつコーパス構築と統計的な自然言語処理技術

自然言語処理で利用する用例、規則、辞書など様々な知識を自動的に構築したり、人手により作成された知識の妥当性や有効性を検証したりするために品詞情報や構文情報を付与したコーパスは大変有効な資源であります。

### 6. 1 コーパスの作成支援ツール

品詞情報に意味的なカテゴリを含めた詳細な体系に基づく文法を用いて、品詞情報、構文情報を付与したコーパスの作成支援ツールを作成し、そのツールを用いてコーパスを構築しています。英語のコーパスについては、約3000の品詞タグからなる文法による品詞情報と構文情報とを、人手による最終判断をしながら、100万語のコーパスを構築しました。日本語のコーパスについては、現在付与すべき品詞情報などの適切性を具体的にデータを作成しながら検討しています。

### 6. 2 統計的な自然言語処理

N-Gram による語の連鎖に関する統計的特徴の関係だけでなく、「文頭や文末の表現」や「単語の綴」などの多様な特徴を「語法の特徴」として利用しています。コーパスからそれらの統計を取り、品詞情報の付与や構文構造の決定に利用するようにしています。

さらに、「語法の特徴」の一般的な特性を捉えたうえで、新聞の社会面や旅行会話などの異なる分野における言葉の使われ方の偏りを明らかにしていきたいと考えています。そして、この偏り方に基づいて特定分野のデータを参照しながら、分野に応じて統計的な情報を調節することを考えています。具体的には、決定木といわれる学習手法を使って、「話法の特徴」を利用して、これまでに品詞情報の付与や構文情報の付与を行なうツールを構築してきました。それを使って、コーパスの作成を現在行っています。

## 留学生から見た日本語

東京工業大学情報理工学研究科 ティモシー・ボールドウィン

私は現在東京工業大学で修士の1年生として、日本語における連体修飾を研究している。ここでは、今までの日本語学習の長い道のりでの感想や直観について述べたいと思う。

日本語でも英語でも、1つの単語から様々な意味を生成することが可能である。英語では動詞と前置詞の組み合わせによって複数の格フレームが生成され、日本語の場合は動詞と格助詞の組み合わせによって複数の格フレームが生成される。しかし、日本語の格助詞と英語の前置詞は、ゼロ代名詞や省略において顕著に異なる。日本語では、いくつかの格要素が省略されても動詞の意味が分かる。「上げる」を述語とする以下の例文を考えてみよう。

- (1) 子供を大学に上げた。  
I saw my children through university.
- (2) 論文を明日までに上げる。  
I will finish the paper by tomorrow.
- (3) 値段を120円に上げた。  
I raised the price to 120yen.

例文(1)～(3)の「上げる」は全て異なる意味を持つが、いずれも「を」「に」格を格要素として持つ。以上の例文から「を格」を省略したものを以下に示す。但し、英語訳の括弧で囲まれた内容は日本語で省略された格要素のことを示す。

- (1a) 大学に上げた。  
(I) saw (them) through university.
- (2a) 明日までに上げる。  
(I) will finish (it) by tomorrow.
- (3a) 120円に上げた。  
(I) raised (the price) to 120yen.

いずれの例文も「に格」しか埋まっていないにも関わらず、「上げる」の意味はもとと同じである。

次に、もとの例文から「を格」ではなく「に格」

を省略するとどうなるか見てみよう。

- (1b) 子供を上げた。  
(I) lifted my children.
- (2b) 論文を上げる。  
(I) will finish the paper.
- (3b) 値段を上げた。  
(I) raised the price.

例文：(2)と(2b)、(3)と(3b)の「上げる」の意味はそれぞれ一致するが、例文(1b)と(1)の「上げる」は異なった意味になってしまう。これは例文(2)(3)の「に格」が任意要素であるのに対して、例文(1)では必須要素であるためである。

一方、英語の場合はどうなるか考察してみよう。多くの前置詞句が任意なので、上記の例文(1)～(3)の英語訳から下線部の前置詞句を削除するとどうなるかを試してみよう。例文(2)と(3)なら前置詞句を削除してももとの意味が保たれるが、例文(1)では前置詞が削除されることで動詞の意味が変わってしまう。大事なことは例文(1)とその英語訳で同じような言語現象が見られることなく、主語や目的語以外の要素で、しかも非慣用的な用法でも意味を絞り込めるということである。英語の中ではこのような前置詞句を補語として分類し、日本語だと「意味的に」必須な格要素(「意味的必須格」と呼ぶ)として扱う。

しかし、注目すべきところはいわゆる必須格(「を格」)の存在よりも、任意格(「に格」)の存在によって意味が限定される用法である。英語でいうと以下のような例文がある。

- (4) I served him for 3 years.  
「3年間彼に尽くした。」
- (5) I served him with tea.

「彼に紅茶を出してあげた。」

例文(1)と違って、例文(4)(5)の前置詞句‘for 3 years’と‘with tea’が、‘I served him’の意味をさらに限定するのであって、変容するのではない。例文(2)と違う点は、英語の文法によって補語が表層的に出現しなくてはならぬことで、英語の場合は付加語が補語によって決められた意味しか限定できないのに対して、日本語の場合は必須格が省略されても、任意格が文全体の意味を限定できる。すなわち、英語は補語が格フレームを決定した上で付加語が限定するが、日本語は「意味的必須格」が格フレームを決定して、その他の格要素が限定する。しかも日本語の場合に、「意味的必須格」以外の格要素は、文法機能ではなく、格フレームのデフォルト要素との類似性によって、全体の意味を限定する度合いが決まる。従って以上の例文(2)と(3)が極めて自然な用法となる。

さらに日本語について関心を持ったのが、他動詞と自動詞の相応性である。特に、その意味的關係に興味を持った。ひとつははっきりと記憶に残っているのは、「換気扇を回す」という表現を初めて読んだ時のことである。この言葉を英語に訳すとすると、‘turn on the exhaust fan’のようになって換気扇の動作を直接表さない表現になるが、このせいか、目的語になる物体自体をぐるぐる回すイメージしかなかった。考えてみても、実際に「回す」作業を実行しているのは回した本人よりむしろ電流であろう。但し「回る」と参照すると、「回るようにさせる」や「回る潜在的な力を放つ」という解釈ができて、別に例外的な表現でない。同じ用法として、「友達を部屋に入れる」、「風を通す」、「トイレを流す」などがある。だが「友達が部屋に入るようにさせる」、「風が通るようにさせる」、「トイレ(の水)が流れるようにさせる」と書き直すことによって、留学生としてごく普通の表現として解釈できるようになった。

そこで、「他動詞」と「自動詞」の漢字を見て自

覚したことに、日本語の動詞分類が英語のと根本的に違うということであった。英語だと動詞のことを‘transitive’や‘intransitive’と呼ぶが、動作を行なうことで何かが移されるか否かを表す。「殴る」だと殴る勢いが移されるし、「伝える」だと伝言などが渡ることになるが、「立つ」や「落ちる」だと何の移行も起こらない。さらに、‘intransitive’の形態素構造が「非-transitive」からなっていることより、あたかも‘transitive’の方が根源のようである。ところが、日本語だと「自動詞」が自らに起こることを表し、「他動詞」が他におよぶ動作などを表すが、漢字からいってもこの二つの概念が英語よりも独立している気がする。さらに、しかしこのような現象にもかからわず、「他動詞」と「自動詞」が密接な関係にあると安心していただころ、「預ける」と「預る」に出会った。一見、片方が他動詞で、もう片方が自動詞であるように見えるが、いずれも目的語をとるから両方が他動詞だということが分かる。しかも、いずれも同じような場面で使われ、同様な目的語をとるので、類似語としてしか思えなかった。周知の通り、「預ける」側がモノを渡し、「預る」側が受けとることになるので、意味が異なる。例文でいうと以下のようになる。

(6)「ボーイが荷物を預かってくれた。」

‘The porter looked after our luggage.’

(7)「荷物をボーイに預けた。」

‘We left our luggage with the porter.’

すなわち、「共感」動詞ペアのようになっていて、「預けてくれる」と「預る」ことが一致するため、他動詞と同意味の自動詞のような存在になる。だが、日本語教育の教科書の中でも「預る」のことを自動詞と記載するほど広く誤解されているようである。

以上、日本語のほんの一部を対象とした自分なりの日本語観であった。この内容から、私の日本語知識が不足していることも明らかであろうが、この先も多言語の日本語で色々と教わっていきたいと思う。

## 電子情報通信学会 言語理解とコミュニケーション研究会（略称 NLC） 最近の研究発表内容の簡単なお紹介

NLC 研究会幹事 杉村領一（松下電器産業株式会社）

電子情報通信学会「言語理解とコミュニケーション研究会」は、年5回の研究会開催、チュートリアル開催、合宿形式のシンポジウムなどを中心とした、会員の活発な研究交流の場となっています。

ここでは、昨年9月に開催されたチュートリアルと10月に開催された研究会の発表内容25件の中から、いくつかの内容を簡単に紹介いたします。

まず、9月に開催されたチュートリアルは、「自然言語処理におけるコーパスの利用」と題して、東京工業大学の徳永健伸先生、奈良先端科学技術大学院大学の松本裕治先生、弁護士吉田正夫先生に、大量の文例をベースに学習など種々の言語処理を行う最先端の技術や、文例を収集するにあたって研究者として知っておくべき著作権上の課題について、それぞれ講義をおこなっていただきました。

徳永先生からは、「RWC テキストデータベースについて」と題して、新情報処理開発機構（RWCP: Real World Computing Partnership）で進められている実世界に関するデータの収集と利用を目的としたテキスト、音声、画像、マルチモーダルの4種のデータベース作成について、紹介していただきました。また、これらに関するいくつかの研究テーマについて、深く掘り下げた議論をしました。

松本先生からは、「コーパスに基づく自然言語処理」と題し、自然言語処理研究者が利用可能な日本語データや、これらを用いた研究活動の鳥瞰図を示して頂いた上で、学習に関する先生の最近の研究内容について、初心者にも分かりやすい講義をしていただきました。

また、吉田先生には著作権上の一般事項の講義と、チュートリアル参加者から予め寄せられた質問事項について、どのように考え、また、行動すべきか分かりやすく説明をしていただきました。会場からも素朴な質問が数多く寄せられ、著作権について参加者は深く学ぶことができたのではないかと思います。

次に、10月の研究会はNLC 単独で開催され7件の発表がありました。発表内容は、機械翻訳関係が

2件、日本語文や英語文などの文法的な構造を解析する手法（構文解析）に関する発表が3件、機械と人間との対話に関するもの1件、言葉同士の連想関係を扱ったものが1件となっています。

機械翻訳関連の発表としては、日英翻訳において、動詞と名詞の意味的な共起関係を機械の支援を得ながら収集する方法についての報告がNTT コミュニケーション科学研究所の白井論氏からありました。報告では、例えば「彼女はさっと野菜の水を切った」という文に対して「She shook the water off the vegetables.」を訳出する際に、「さっと…切った」と「Shook…off」を対応させ、登録する方法を扱っています。そして、これらの関係を如何に効率良く作成、利用するかについて、実際の作成経験に基づいた報告がなされています。

構文解析については、まず「文中のある特定の句は、修飾句によって十分に意味が限定された場合、それ以上修飾句に係る確立は少なくなる」という仮説に基き、文解析を確率を用いて行う方法が、九州大学の田上敦士氏らによって発表されました。文を解析するためあらかじめ用意した文法規則を、確率によって順位付けして用いる方法を確率文法と呼んでいます。この手法は長い歴史を持ちますが、近年そのモデルの扱い易さや学習に対する柔軟性から、注目を集めています。

機械と人間の対話に関するものとしては、対話参加者同士が非同期に発話を行うシステム構成について、NTT 基礎研究所の堂坂浩二氏による発表がありました。従来一文一文が完全な形で入出力されることを前提として多くの対話研究が行われてきているのに対して、実際の人間同士の対話のように互いに相手の発話中に割り込むことを想定したシステムの作り方が考案されており、実際のデモなども通じて報告がありました。

以上のように研究会、チュートリアル等の活動を通じ、幅広いくかつ深い研究者の交流と研究活動の進展に役立てるよう努力をしています。

## 脳と用語シンポジウム

1996年12月12日、日本学術会議、情報処理学会、情報知識学会および専門用語研究会の主催による、「脳と用語シンポジウム」が日本学術会議講堂において開かれた。

招待講演の一つとして、当協会の前会長長尾真氏（京大教授）が「言葉のネットワーク」というテーマで講演された。

それぞれの単語の意味は、他の単語の意味との関係で全体的に決まていくものと見ることができる。これらの総体が言葉の意味を構成しているわけで、壮大な言葉のネットワークをなしているといえる。

このネットワークを構成している基本的な要素としては、同義語関係、上位・下位関係、原因・結果関係、行為・対象関係等いろいろ考えられる。人間はこれらを考慮して文を理解し文を作っている。自然言語処理においてもこのような人間の言語処理能力に迫る必要がある。

しかし、言葉の諸関係をすべて記述した辞書はまだ存在しない。最近ようやく IPAL 辞書や EDR 辞書のように機械処理用の電子化辞書が作られているが、さらに充実していくことが必要である。

このような言葉のネットワークを作るには、次のようなステップが必要であろう。

- ・大規模テキストコーパスの作成
- ・大規模テキストの自動構文解析、意味解析結果の蓄積
- ・語相互間の関係性の抽出

これを行うに最も大切なことは、人手によって作られた種々の辞書を上手に統合して活用することで、しかも自動的に工夫をすることであろう。クラスタリング手法の活用も必要であろう。また、専門用語の組織化が必要である。適切なテキストの選定、それに基づく用語の木を作成、用語の内容説明等のステップを踏んで行うのがよからう。

機械翻訳に関しては、この他に下記2つの一般講演でも言及されていた。

### (1)「情報流通と分類表の役割およびその体系」

図書館情報大学 石川徹也氏

国際間の情報流通における問題点の一つとして、EUの多言語問題がある。EU15カ国で使用される言語は45種ある。EU会議の公用語は英、仏、独の3カ国語であるが、公文書は最多45言語に翻訳しなければならない。現在翻訳経費の削減、正確性、即時化のために6言語対応の機械翻訳システム（SYSTLAN）が利用されているが、それでもなお、1,100人の翻訳者が存在する。

また、現在インターネットはほとんど英語記述であるが、民族のアイデンティティが崩壊する危機があるという議論が起こっている。

したがって、国際間情報流通の要件として、下記のことが言える。

- ①情報記述は各国の公用語で発信してよい。
- ②多言語対応情報検索システムが必要である。
- ③多言語対応機械翻訳システムが必要である。

### (2)「TKE'96にいたる最近のターミノロジー研究」

IRIS 情報学研究所 仲本秀四郎氏

機械翻訳に関して下記のように述べられた。

TKE'87において、ECが開発中の機械翻訳についての特別セッションが設けられ、EUROTRA（機械翻訳システム）プロジェクト、SYSTLANおよびSYSLING（EC9公式語の翻訳サービス）の報告が出席者の関心をひいた。

TKE'96後のヨーロッパ機械翻訳協会でのワークショップで、Dorothy Senezが「20年前研究段階であった機械翻訳は、今や実施への関心をひく状況になった」と発言した。



## CICC 多言語機械翻訳システムフォローアップ事業

(財) 国際情報化協力センター (CICC) 国際情報化研究所 部長代理 川島 允

財団法人国際情報化協力センター (CICC) では 1987年度から1994年度までの8年間にわたり中国、インドネシア、マレーシア、タイの海外4カ国と中間言語方式による多言語機械翻訳システムの研究協力を通商産業省の委託事業として実施した。1995、96年度の2年間はそのフォローアップ事業として、研究に参加した各国において辞書、文法の改良・追加、機械翻訳システム実用化に向けた追加研究を独自に実施出来るよう支援するための必要な技術指導を行った。フォローアップの実施機関としてCICC内に「機械翻訳フォローアップ推進委員会」を設置し、東京工業大学の田中穂積教授を委員長に迎え、日本及び海外4ヶ国の研究者を委員として活動を行ってきた。

主な活動は機械翻訳プロジェクトの成果の普及に

関する意見交換、中間言語仕様の改良、辞書の概念体系の検討等である。

1996年度は各国の研究者をCICCに招聘し、多言語機械翻訳システムの翻訳精度の向上をはかるため、基本動詞を中心としてケースフレームの検討や電子化辞書における概念体系等について、活発な討議を行う機会を2回設けた。

また、各国研究者から本システムを応用した技術開発の現状や各国における機械翻訳システムの動向についての報告があり、各国とも着実に自然言語処理技術を発展させていると言える。

本事業は通産省からの委託事業としては1996年度で終了する。CICCとしては今後ともこれら共同研究国の機関、研究者とのネットワークを維持し、情報交換等を続けていくこととしている。

### イベント紹介

### テクニカルコミュニケーション シンポジウム'97のご案内

テクニカルコミュニケーター協会は、下記のとおりテクニカルコミュニケーションシンポジウムを開催する。このシンポジウムは、テクニカルコミュニケーション技術の普及・向上、参加者の相互啓発を目的として、テクニカルコミュニケーションに関する各種テーマについて、発表・討論・意見交換の場を広く提供するものである。

日 時：1997年9月4日(木)・5日(金) 10:00～17:00

場 所：工学院大学 (東京都新宿区)

全体テーマ：「かたち、色、音、動き……電子空間でのTC(テクニカルコミュニケーション)」

基調講演、特別セミナー、分科会、事例・研究発表、展示会を企画中。

詳しい内容は6月に案内される予定。

問合せ先：TC協会事務局 (柳下、下田)

TEL 03-3368-4607 FAX 03-3368-4607-5087

## ソフトウェア講演会と特許機械翻訳実験システムの公開

1997年2月27日、名古屋市工業研究所において、下記の講演会と実験システムの公開が行われた。

主催 名古屋市工業研究所  
共催 名古屋市工業技術振興協会 システム技術研究会  
協賛 中部エレクトロニクス振興会  
名古屋市システムハウス協会 中部支部

### 1. ソフトウェア講演会

日米特許制度の動向と機械翻訳・インターネットを利用した特許システム

#### (1) 日米特許制度の動向…インターネットによる情報公開

弁理士会 業務対策委員会委員 弁理士 足立 勉氏

#### (2) 英文解析による英文特許の機械翻訳への応用

名古屋言語処理研究会代表 佐良木昌氏

#### (3) インターネットを利用した特許情報システム

名古屋市工業研究所 電子部 情報技術課長 加藤輝政氏  
電子部 情報技術課 小川 清氏

### 2. 特許機械翻訳の実験システムの公開

#### (1) インターネット上に公開されている米国特許公報（明細書・請求範囲・アブストラクト等全文）を直接ダウンロード

#### (2) これら英文を、市販の機械翻訳ソフトが翻訳可能なレベルまで、コンピュータによって自動編集

#### (3) さらに、翻訳後の訳文について、表現の是正も自動化

翻訳に適した技術分野は、自動車、電気電子、通信、制御、メカトロニクス、コンピュータ、光エレクトロニクス

以上

## 協会活動報告

|           |        |                                           |
|-----------|--------|-------------------------------------------|
| 運営委員会     | 12月7日  | IPA プロジェクト、展示会等の報告と討議。関連団体、MT 関連情報等の連絡。   |
|           | 1月29日  | '97年度事業計画の討議。シンポジウムについての討議。MT 関連情報等の連絡。   |
|           | 2月19日  | '97年度事業計画の討議。シンポジウムについての討議。MT 関連情報等の連絡。   |
| 市場動向調査委員会 | 12月17日 | ヒアリング：JST の機械翻訳システムについて                   |
|           | 2月27日  | ヒアリング：放送分野での機械翻訳の利用                       |
| 技術動向調査委員会 | 12月6日  | UPF 会議と共催で、UPF 仕様を検討。JTF との交流会。           |
| 例文評価研究会   | 1月24日  | 記号を含む翻訳困難文について。MT サミットでの発表について。JTF との交流会。 |
| 編集委員会     | 1月13日  | Journal No. 17および No. 18以降の編集について。        |

## 第2, 3回 JTF-AAMT 交流会

1996年10月14日の第1回JTF-AAMT交流会(AAMT Journal No. 17)に掲載)にひきつづき、1996年12月6日に第2回が、1997年1月24日に第3回が開催された。

第2回は技術動向調査委員会の委員を中心とするUPF会議のメンバーが出席した。JTFからは野村勇一専務理事のほか翻訳会社の代表が参加した。

出席者の自己紹介の後、村木一至委員長が技術動向調査委員会の活動の概要を紹介した。続いて、自由討論に入り、下記のような話題や意見がでた。

- (1) 技術文書の翻訳者は、その経歴から2種類に大別できる。一つは語学の専門家で翻訳は上手だが専門語に弱い。これに対しては専門の辞書が役立つ。もう一つは技術者上がりで、専門には強いが翻訳に弱い翻訳者である。  
また、「翻訳のできる人」より「翻訳の好きな人」の方が品質や納期の点で望ましい。
- (2) 現在、翻訳会社の仕事の95%は日英翻訳である。英日翻訳はほとんど自社内で行われているためと思われる。  
翻訳の納期はだんだん短くなって来ている。例えば、かつては40日くらいであったものが2週間程度になっている。MTの高速性を活用して対応している。翻訳は品質第一であり、MTだけに頼ることはできない。そのため多数の翻訳者が必要だ。
- (3) 日英翻訳の場合、日本語の原文の質を向上する必要がある。大分よくなってきたが、まだ改善の余地がある。  
マニュアルは、日本では製品ができてから技術者が書くが、欧米では開発と平行して文章の専門家が書く。
- (4) 新しい専門用語の訳を、例えばチャットシステム等を使って執筆者に教えてもらう手段は考えられないか。訳文の再利用は、新人は望むがベテランになると必要となくなる。
- (5) 悪い訳は修正に時間がかかり、しかも出来上がりが良くない。下訳でもある程度のレベルに達していないと役立たない。現状では簡単どころだけにしかMTを用いられない。MTの出現で失業する翻訳者が出ると思っていたが、まだまだ大丈夫のようだ。

第3回は、当協会からは例文評価研究会のメンバーが、JTFからは野村勇一専務理事のほか翻訳会社の代表ならびに個人翻訳者が出席した。

前回同様、自己紹介と横山座長の研究会の概要紹介の後、自由討論に入った。今回は事前にJTF側出席者から出された英文を、AAMTのメンバーが各社の英日翻訳機にかけた結果を配布し、話題の一つとした。これに関しては下記のような話が出た。

- ・ 2箇所には綴りミスがあった。ネイティブでも綴りミスが多い。ミスは人間には分かるがMTには分からない。
  - ・ MTとしては良い訳ができていると思われるが、請負翻訳のレベルには達していない。  
正直なだけの訳(直訳)ではだめで、こなれた訳が求められる。出題者の翻訳をAAMTに送る。
- このほかに、次のような話題がでた。
- (1) MTをうまく使うには、めんどくさくはあるが、カスタマイズすることが重要である。  
辞書データベースで訳し分けられるようにしたものもある。  
決まったユーザだけを手掛ける方が成功する。
  - (2) MTに大きな期待をしても無理である。なんとか実用になるレベルになった段階である。  
MTは英語の専門家でなくても使えるのが良い。
  - (3) 一般向けと産業翻訳向けでは、MTの開発方針が異なる。産業翻訳では、原文まで見なければならぬのでは2度手間で役立たない。超直訳でもよいから内容を掴めることが重要である。  
専門用語が正しく訳されていれば役立つ。医学のように専門語の多い分野ではかなり役立っている。
  - (4) アメリカでは、大学に産業翻訳の講座がある。  
日本でも検討されたが講師がいない。
  - (5) MTの開発には日本語の勉強が重要である。原文の悪いことが多い。特に日本語文は一般に長すぎる。副詞句等の修飾先が分からないことや主語が分からないことがよくある。MTにかかるような日本語に修正する必要がある。そうするとMTが更に成長する。英語では、一般に15～20単語の文が分かりやすい。関係代名詞は極力使わないようにすべきである。

## アンケート結果の報告

AAMT 編集委員会

AAMT Journal No.16で実施しました「AAMT Journal」に対するアンケート結果の報告をいたします。残念ながら回答は5件しかありませんでしたが、貴重なご意見をいただきありがとうございました。回答数が少ないので、会員全体の意向と考えるのは無理かと思いますが、今後の編集の参考にさせていただきます。

## (1) 読まれる度合いについて

どの記事も「いつも読む」または「ときどき読む」でした。「いつも読む」が一番多かったのは、技術早わかりと新製品紹介で、次いでヒアリングと翻訳の現場からでした。

## (2) 内容の良さについて

どの記事も「非常に良い」または「普通」でした。「非常に良い」が一番多かったのは、技術早わかりで、次いで翻訳の現場から、ヒアリングおよび研究機関紹介でした。

## (3) コメントについて

海外の研究機関の紹介もしてほしいとのコメントがありました。

国内の研究機関の紹介がほぼ終わりましたので、今後海外の紹介をしたいと考えています。

## (4) 記事の要望

## ① 海外での会議、学会などの情報および海外での機械翻訳の現状の記事

例えば、AAMT Journal No. 17の「AMTA-96参加報告」のように、主なものは載せてきました。

MT News International にも掲載されますのでご覧ください。今後さらに充実させていきたいと思えます。

## ② インターネット利用のMTソフトの内容の紹介

「新製品紹介」という形で紹介しておりますが、インターネットが脚光を浴びている現状から、この関係の記事を充実させていきたいと思えます。

## ③ 機械翻訳の使い方の事例

AAMT Journal No.13の「機械翻訳のための日本語の書き方」はその一例かと思えます。しかし、この件は関心が高いわりには記事が少ないことも事実です。検討いたします。

## ④ 個人向け翻訳関連市場調査

要望の強いことは分かっているのですが、なかなか実体がつかめないのが実状です。今後の課題いたします。

## イベント紹介

## 佐良木事務所・機械翻訳研究室の開設

このたび、名古屋の佐良木技術翻訳事務所が機械翻訳研究室を開設しました。言語学・自然言語処理・機械翻訳の研究室として運営されます。お立寄りください。

所在地：名古屋市東区葵2丁目2番21号（地下鉄 桜通り線 車道駅から徒歩200m）

電話：050-150-3755

E-Mail：saraki@st.rim.or.jp

# たちまち翻訳 英日翻訳支援システム for Windows95

株式会社 日立製作所

インターネットの普及の速度には、目を見張るものがあります。

その中で、英文ページの内容把握に手間取っている人も増えているのではないかな…手軽に買えて、簡単に使える翻訳ツールが必要なのではないかな…私共は、そう考え、このたび、「たちまち翻訳 英日翻訳支援システム for Windows95」を発売致しました。

一人でも多くの方に使っていただきたいという思いから、価格を7,800円に設定しました。

価格は7,800円ですが、文法・辞書には、日立が長年、取り組んできた、メインフレーム・ワークステーション用機械翻訳システムの開発で蓄積したノウハウがそのまま生かされています。

また、豊富な機能を持たせてありますので、いろいろな場面でお使いいただけると思います。

以下に、「たちまち翻訳」の機能をご説明します。

## ● WWW ブラウザ翻訳機能

WWW ブラウザで、「英日翻訳」ボタンをクリックするだけで、ホームページを翻訳できます。しかも、図表のレイアウトが保存されますから、英文のホームページの感覚そのままに日本語でホームページを読むことができます。翻訳結果の表示は「対訳表示」「訳文のみ」「原文+訳文」の3種類から選択できます。

また、ダウンロードしたローカルなページも翻訳可能ですから、接続料金の心配をせずに翻訳できます。



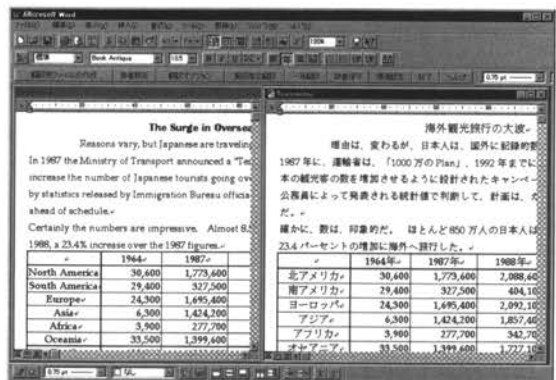
(WWW ブラウザ翻訳画面)

## ● Word 翻訳

Microsoft Word を起動して、翻訳したい文書を開けば、その場で翻訳できます。

翻訳専用のファイルを作って翻訳しますので、原文のファイルは元のまま残せます。

また、翻訳用ファイルは、Microsoft Word の文書と同じですから、日頃から使いなれた Word の機能を使って編集できます。もちろん、ブラウザ翻訳と同様に、図表のレイアウトは原文のものが生かされますから、図を貼り付け直すなどの操作は不要です。



(Word 翻訳画面)

## ● クリップボード翻訳

WWW ブラウザや Microsoft Word では開けない文書は、クリップボード経由で翻訳できます。英文



メールや英文ヘルプなどの翻訳に力を発揮します。

もちろん、WWW ホームページの部分訳にも活用していただけます。

### ● ユーザ辞書保守機能

基本辞書で、約6万語の見出し語をご提供していますが、ユーザ独自の訳語が必要な場合や固有名詞、新語などを登録するために、ユーザ辞書保守機能をご用意しました。

名詞・動詞・形容詞・副詞・熟語と、多彩な語が登録できます。意味分類コードや分野コードなどを用いて、1つの語を訳し分けることもできます。

### ● 翻訳オプション機能

文末の形を丁寧形にするかしないか、訳語はカタカナ訳と日本語訳のどちらを優先するか、文書の分野による訳し分けなどの指定が可能です。



#### 【基本仕様】

- ・ 翻訳の種類：WWW ブラウザ 翻訳、Microsoft Word の翻訳、クリップボード翻訳
- ・ 基本辞書：6万語

- ・ 辞書の指定：3つの辞書を同時に指定可能
- ・ ユーザ辞書登録：名詞・動詞・形容詞・副詞、およびそれぞれの熟語の登録・更新・削除・印刷
- ・ 翻訳オプション：「です・ます」調／「である・だ」調の選択。「日本語訳優先」／「カタカナ訳優先」の選択。分野コードによる訳し分け。

#### 【前提ソフトウェア】

- ・ Microsoft Windows95
- ・ Microsoft Word6.0または7.0 (Word 翻訳を行う場合)
- ・ Internet Explorer2.0または3.0, または Netscape Navigator2.0または3.0 (WWW ブラウザ 翻訳を行う場合)

#### 【メモリ】

16MB 以上 (OS 含む)

#### 【空きディスク容量】

25MB 以上

#### 【標準価格】

7,800円 (税別)

- ・ Windows, Microsoft Word は、米国 Microsoft Corporation の米国およびその他の国における登録商標です。
- ・ Netscape Navigator は米国、日本および他の国々における Netscape Communications Corp. の商標です。
- ・ Internet Explorer は、米国 Microsoft Corporation の米国およびその他の国における商標です。

#### 【ホームページ】

<http://www.hitachi.co.jp/soft/trans/>

#### 【この製品に関するお問い合わせ先】

日立オープンミドルウェア問い合わせセンター  
フリーダイヤル：0120-55-0504  
(9:00~17:00 土、日、祝日を除く)

# レターナビゲータ1.0

(株) インターグループ

通訳・翻訳・語学研修・国際イベント業務において圧倒的な実績を誇るインターグループは、電子化辞書や各種ソフト開発のノウハウを結集して、簡単に英文レターが作成できるパソコンソフト「レターナビゲータ1.0」を開発しました。実際の英文レター翻訳業務や専門的語学研修機関「インタースクール」での語学教育に裏打ちされた確かな英文ソフトとして、ビジネスシーンをはじめとよりパーソナルな分野でも、多くの一般ユーザの英文文書作成を支援するソフトとして、市場に投入する予定です。

## 1. 背景

昨今の国際交流はインターネット等の普及に伴い、以前にもまして幅広い層にまで拡大してきています。

海外通販等も電子決済システムの開発導入に伴い、今後ますますより多くの人々に利用されるようになるでしょう。

しかしながら、一番の壁となっているのは、言語の問題です。

今回開発しましたパソコンソフト「レターナビゲータ1.0」は、重要なビジネス上の伝達事項から、ちょっとしたプライベートな話題まで、英語が書けなくても日本語だけで、英文レターを作成し、そのまま電子メールで送れ、又、印字、宛名書きもできるソフトです。

## 2. 特長

### ①対訳表示

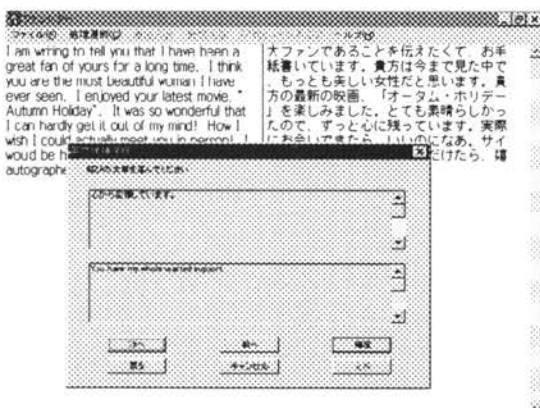
まず、書きたい内容を日本語で順に選びます。それだけで、日英対訳のレターが画面に表示されます。

後は、日付や固有名詞等の具体的な部分が、可変部分としてハイライト表示されますので、ユーザはその部分に特定のデータを入力するだけでレターを完成できます。

特定のデータはユーザ辞書にも登録できますので、次回からは、同様な情報を入れたい候補箇所には、その言葉が自動的に第一候補として表示されますの

で、もう一から入力する必要もありません。

「だいたいの内容さえ英語でできたら、後は自分で細部を作成するよ」というある程度英文の書けるユーザには、この英文レター下書き機能だけでも十分有用なものとなるでしょう。



### ②「電子情報辞書1.0」のサポート

固有名詞等以外の一般的な用語についても、変えた場合は簡単に入れ替えられます。

インターグループが電子媒体ソフトウェアのみで使用することを目的に開発した「電子情報辞書1.0」が基本辞書として、同ソフトには搭載されています。各可変部分がその辞書にリンクしているので、例えば、〈待ち合わせの場所〉を入力したい場面では、自動的に、〈待ち合わせの場所〉という意味カテゴリーをもつ言葉（公園—park、ホール—hall...等）が自動的に候補として表示され、ユーザはそれを選ぶだけで、レターの細部を完成してゆけます。

世界及び国内の地名も豊富に収録されていて、しかも、特定の州や県を選ぶと、次の都市名入力待ち部にはその州や県にある都市名だけが選択候補として表示されるなど、使って便利な機能満載の電子化基本辞書です。

### ③ユニーク！ カード文例

手紙までおかげさを書く必要はない——とりあえず季節の便りや、お祝い、お礼を簡単にカードで送りたい——という時の為に、挨拶状やクリスマス

カード等を簡単に作成できる機能を入れています。

かわいいイラストまで選んで同時に印刷できるので、オリジナルカードを作るのにもってこいです。



④住所一発印字——文書作成のみならず、“送付”まで支援！

差出人や受取人の住所・氏名をどこにどういう風にかきたいの？——と悩んだことはないですか。英文レターを出し渋っている人の多くが、アドレスデータを記載することの煩わしさを感じています。

「レターナビゲータ1.0」では、一度だけ住所・氏名を登録すれば、レター、電子メールの各フォーマットにあわせたレイアウト形式を自動的に選択してくれます。レターの文書作成を手助けしてくれるだけでなく、「送付」にまつわる煩わしさも解消してくれます。

⑤オリジナル「レタナビ」作成機能(オプション)

本来この種のソフトは、お仕着せのデータを使うしかないわけなのですが、「レターナビゲータ1.0」では、レターの分野から、細部の可変項目まで、自分で一から作って追加できます。

社内で既にモデル文書が存在する場合は、本ソフト内の、「レタナビ自動作成」機能のチュートリアルに従い、その文書をテンプレート化すれば、もう次からオリジナル「レタナビ」が使えます。

⑥「レターナビゲータ/データベース GL1.0」のサポート(オプション)

インターグループが開発している膨大な対訳文例辞書の中から、レター文例を1万文例以上日英対訳で収録しています。分野や使い方で分けされており、文字列検索も可能ですから、「レタナビ」の中で英文を追加したり、登録しておいたりする時に便

利なのは言うまでもなく、一般的に英文レター文例データベースとしても使えます。

### 3. 収録内容

①分野／パターン

ビジネス 39分野：64パターン

パーソナル 51分野：79パターン

②「電子情報辞書1.0」

約10,000 対訳語

③カード文例

8種類 48文例

④オリジナルイラスト

8種類 32点

### 4. 動作環境

OS : 日本語 Windows 95

CPU : Pentium 90MHZ 以上

メモリ：16MB 以上

### 5. 発売時期

97年4月1日

<オプション>については5月以降を予定。

### 6. お問い合わせ先

(株) インターグループ ソフト開発部

〒531 大阪市北区中津3-7-3

ビルコビル

TEL：06-372-8048 (代表)

FAX：06-372-3909



## シャープノートパソコン Mebius シリーズに搭載 「英日翻訳支援ソフト(インターネット・英文メール作成支援対応)Power E/J 2.0」

シャープ(株)

シャープでは、ノートパソコン Mebius シリーズに、従来よりインターネット WWW 情報や英文ファイル等の英文翻訳をサポートするソフトとして「英日翻訳支援ソフト」を搭載しておりましたが、この度「Power E/J 2.0」として、英文メールに関して、その作成に対する支援を行う「英文メール作成支援」機能と、届いた英文メールをバックグラウンドで自動的に英文翻訳を行ない、翻訳結果を受信した英文メールの添付文書として受け取ることができる「E/J for Mail」機能を追加することにより、仕様を更に強化致しました。尚、「英文メール作成支援」機能では、英文メールを和文から作成できるようにするため、日英翻訳機能を取り入れております。

英日翻訳支援ソフト(インターネット・英文メール作成支援対応)PowerE/J 2.0には(★は新機能)、

・E/J マネージャー

PowerE/J に含まれるソフトウェアをアイコンのクリックにより簡単に起動させることができるソフトウェア

・E/J for Text

ファイル上の英文などを簡単な操作で英日翻訳するソフトウェア

・E/J for WWW

Microsoft Internet ExplorerまたはNetscape Navigatorを使ってインターネットのWWW ページにアクセスしながら英日翻訳を行うソフトウェア

・英文メール作成支援(★→次頁図1)

英日/日英翻訳支援機能を使い、英文メールを簡単な操作で作成できるソフトウェア

・E/J for Mail(★→次頁図2)

Microsoft Exchangeを使って受信した英文メールを自動的に翻訳するソフトウェア

・大修館書店ジーニアス英和辞典(★)

・小学館プログレッシブ和英中辞典(★)

・学習研究社英和・和英辞典、国語辞典、漢和辞典の各ソフトで構成されています。

Power E/J 2.0は、インターネット WWW 翻訳支援、英文テキスト翻訳支援、英文メール作成支援、辞典(英日/日英)機能など、翻訳作業を必要とする場面をサポートする統合化仕様であり、また辞典として、国語/漢和辞典も備わっているため、翻訳から和文文章作成までをフルにサポートします。

今回は Power E/J 2.0で強化された新機能の中で、特に、英文電子メールの作成/発信が簡単に行えるようにするため、英日翻訳支援機能・日英翻訳支援機能を使用した「英文メール作成支援」機能の特長について説明します。

「英文メール作成支援」機能の特長は、以下のようになります。

<簡単新規作成機能>

挨拶、案内、お祝いなど定型のメールを作成する場合に、簡単新規作成機能を使えば、指定された項目を選択したり必要な項目に記入するだけで、自動的に英文メールが作成されます。作成されたメールは自由に編集できますので、追加したり変更したい部分を編集、検索・翻訳するだけで英文メールを完成させることができます。

簡単新規作成機能のデータは、(株)インターグループが開発した文書データベースおよび英文レター作成ソフト「レターナビゲータ」(Letter Navigator TM©1996)に基づいており、著作権はインターグループ社にあります。(Letter Navigator (TM) Copyright 1996 INTER GROUP CORPORATION All rights reserved.)

<豊富な用例>

原文の中からキーワードを抽出し、キーワードを含む用例やキーワードに関連のある用例を豊富な用例文の中から検索して表示します。このため、原文に近い英文を探すだけでなく、さまざまな用例を参考にして表現したい英文を作成することができます。

また、このとき原文を英訳した結果と、さらにそ

の英訳結果を和訳した結果も同時に表示しますので、用例以外にこれらの翻訳結果も英文作成の参考にできます。用例は、ドラッグアンドドロップやコピーなどの編集操作で簡単に取り込めます。

＜作成したメールをそのまま送信＞

英文メール作成支援機能で作成した英文をファイルに保存したり、クリップボードを経由することなく直接 Microsoft Exchange に渡して送信することができます。

＜直感的な編集操作＞

範囲選択されたテキストやファイルとして登録された文章であれば、英文メール作成支援のアイコンにドラッグアンドドロップするだけで、テキストを取り込むことができます。また、英文メール作成支援ウィンドウ内部のテキスト編集操作もドラッグアンドドロップによって行えます。

以上説明した機能により、次の手順で英文メールを作成し、電子メールを使って発信することができます。

ステップ1 簡単新規作成機能を使い英文メール文例集を基に、対話形式でメールのベースを作成します。

ステップ2 作成した文に追加／変更を行うため、

英文にしたい文を和文で入力します。

ステップ3 入力した和文に近い英文例を用例検索し、適切な用例を作成中のメールに取り込みます。

ステップ4 また、ステップ2で入力した和文を日英翻訳し、結果の英文をメールの文章として取り込めます（日英翻訳機能で得られた英文は、自動的に英日翻訳機能で和文になるため、英文の意味が確認できます）。

ステップ5 以上のステップ1～4により作成した英文メールを Microsoft Exchange 等を使って電子メール発信します。

＜Mebius シリーズに関するお問い合わせ先＞

シャープ株式会社コンシューマーセンター

西日本相談室 TEL (06)621-4649

東日本相談室 TEL (043)297-4649

・ Microsoft、Windows は、米国 Microsoft Corporation の登録商標です。

・ Netscape Navigatorおよび Netscapeは、Netscape Communications 社の商標です。

・ E/J for Text、E/J for WWW、英文メール作成支援などを使って翻訳を行う場合は、著作権法上、原著作者の許諾が必要になることがあります。

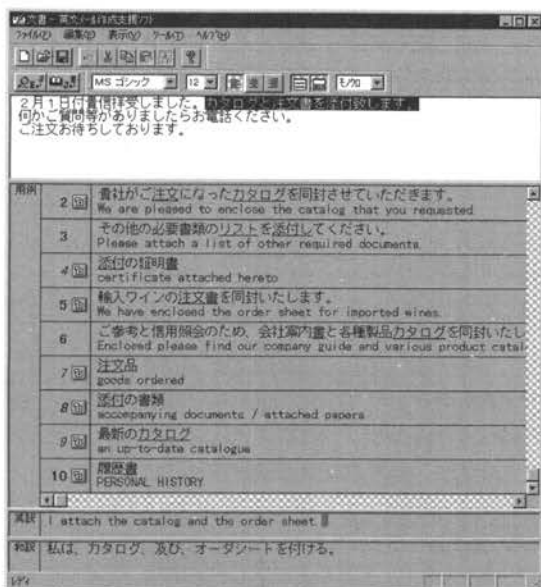


図1

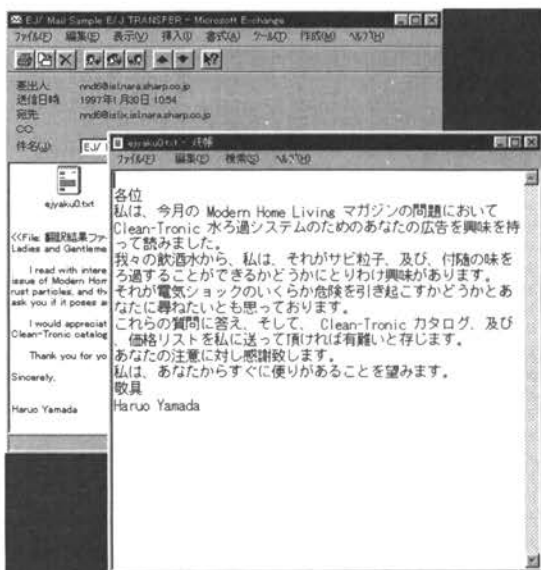


図2

事務局だより

## MT Summit に参加しよう！

すでにご承知のとおり、'97年10月29日から11月1日まで、米国の San Diego で MT Summit が開かれます。MT Summit は2年ごとに開かれる機械翻訳の世界大会です。今回はアメリカ機械翻訳協会 (AMTA) の担当で、着々と準備が進んでいます。MT News International No.15に状況が報告されていますのでご覧いただき、AAMT から多数参加してください。

委員会だより

## 会員の皆様からの投稿をお待ちしております

編集委員会

AAMT Journal は、アジア太平洋機械翻訳協会の機関誌です。協会から会員の皆様への情報伝達という役目のほかに、個々の会員から他の会員への情報発信の場としての役目も持っております。会員に有用な記事を投稿いただき、この Journal をより一層充実させたいと思います。

詳細は「AAMT および IAMT の機関誌について」AAMT Journal No.12 (Oct.'95) に掲載されています。

会員の皆様からの積極的な投稿をお待ちしております。

AAMT  
ジャーナル

No. 18  
(March 1997)

発行 アジア太平洋機械翻訳協会  
所在地 〒105 東京都港区芝公園3-5-12 芝公園真田ビル3階  
TEL : 03-5473-7135/7136 FAX : 03-5473-0569  
E-mail : KYN02317@niftyserve.or.jp (NIFTY)  
編集委員会 野村浩郷(委員長) 亀井真一郎 福持陽士 富士 秀  
事務局 田中 孝 神野千恵子  
印刷所 伸光写植印刷株式会社





