

AAMT

Asia-Pacific Association for Machine Translation

Journal



February 2016 **No.61**

アジア太平洋機械翻訳協会

目 次

巻頭言：	MT サミットの 30 年.....	梶 博行.....	01
シンポジウム参加報告：	MT Summit XV 参加報告・ワークショップ開催報告.....	須藤 克仁.....	03
シンポジウム参加報告：	第 2 回アジア翻訳ワークショップ (WAT2015) 開催報告.....	中澤 敏明、美野 秀弥、後藤 功雄、Graham Neubig、黒橋 禎夫、隅田 英一郎.....	07
連載：	機械翻訳事始め[VI].....	坂本 義行.....	13
プロジェクト紹介：	音声翻訳の国家プロジェクト.....	隅田 英一郎.....	19
プロジェクト紹介：	特許庁における機械翻訳の取り組み.....	加藤 啓.....	25
レポート：	映像翻訳の授業活用例：字幕制作を通じたメディア理解.....	魚住 真司.....	35
シンポジウム参加報告：	CJNLP2015 参加報告.....	山本 ゆうじ.....	41
セッション参加報告：	JTF 翻訳祭「MT Live ～機械翻訳の担うべき役割～」.....	熊野 明.....	44
製品・サービス紹介：	特許公報・技術文献にフォーカスした npat 翻訳サービス.....	本間 奨.....	46
委員会活動報告：	これまでの AAMT Forum メールマガジン Vol.7.....	機械翻訳課題調査委員会 WG1,2.....	50
委員会活動報告：	実務翻訳における機械翻訳の利用に関するアンケート結果報告.....	機械翻訳課題調査委員会 WG3.....	55
事務局からのお知らせ：	協会活動報告 (2015 年 9 月～2016 年 1 月).....	AAMT 事務局.....	60
編集後記	65

CONTENT

Foreword:	30 Years of MT Summit.....	H. Kaji.....	01
Symposium Report:	Report --- MT Summit XV and PSLT 2015.....	K. Sudoh.....	03
Symposium Report:	Report of the 2nd Workshop on Asian Translation (WAT2015).....	T. Nakazawa, H. Mino, I. Goto, G. Neubig, S. Kurohashi, E. Sumita.....	07
Series:	History of Machine Translation [VI].....	Y.Sakamoto.....	13
Project Report:	National project of speech translation toward 2020 Olympic games in Japan.....	E. Sumita.....	19
Project Report:	Initiatives of the Japan Patent Office on Machine Translation.....	K.Kato.....	25
Report:	An Alternative Method for Media Studies: Understanding News/Documentaries through Subtitling.....	S. Uozumi.....	35
Symposium Report:	Report: CJNLP2015 (China-Japan Natural Language Joint Research Promotion Conference).....	WG3(Y.Yamamoto).....	41
Session Report:	JTF Translation Festival "MT Live - Role of Machine Translation -".....	A. Kumano.....	44
Products & Services:	npat translation services with a focus on patent publications and technical literature.....	S.Honma.....	46
Committee Report:	AAMT Forum mail magazine vol.7.....	WG1, WG2.....	50
Committee Report:	Report on questionnaire results of using machine translation for business translations.....	WG3.....	55
AAMT Activities:	AAMT Activities (from September 2015 to January 2016).....	60
Editor's Note:	T. Utsuro.....	65

MT サミットの 30 年

梶 博行

静岡大学大学院情報学領域 教授

私ごとで恐縮だが、この 3 月に大学の定年を迎えることになり、編集委員会から MT 研究の思い出でも書くようにとのご指示をいただいた。AAMT 関係ではこの 10 年あまり AAMT/Japio 特許翻訳研究会の委員として活動してきたが、AAMT の最も重要なイベントの一つである機械翻訳サミットとの関わりも長いので、その 30 年を振り返ってみたい。

周知のとおり、第 1 回 MT サミットは 1987 年 9 月に箱根で開催された。IAMT の傘下に AAMT、EAMT、AMTA を設置し、これら 3 地域の協会が持ち回りで MT サミットを開催することが決まったのはワシントン DC での第 3 回 MT サミットにおいてであり、第 1 回の会議は長尾真先生のリーダーシップの下、当時の日本電子工業振興協会の機械翻訳システム調査専門委員会が中心となって開催した。筆者は、当時、この委員会の委員を務めており、先輩委員に教えていただきながら準備に加わった。いわゆる学会とは異なり、MT システムのユーザ、研究者、開発者、さらに政策関係者が意見を交換し議論する場を提供するという会議の目的もこの準備委員会で決められた。サミットという名称にもそのことが込められている。

第 1 回 MT サミットは文字どおり機械翻訳分野の主要なメンバーが一堂に会した会議で、ALPAC 報告にも関わった Martin Kay 博士、MT 研究の歴史と技術に関する大部の書物を著作した W. John Hutchins 博士、当時、世界の注目を集めた第 5 世代コンピュータプロジェクトのリーダー淵一博博士が基調講演を行った。日米欧のリーダーたちによるパネル討論やプロジェクトの報告に加えて、商用システムの開発者とユーザがペアで講演する MT サミットらしいセッションが設けられた。当時、メーカーで日英機械翻訳システム

の研究開発を担当していた筆者はそのセッションの講演者の一人として発表した。

第 2 回 MT サミットは 2 年後にミュンヘンで開催されたが、箱根サミットと同様に日本から多くの方が参加した。筆者も自社のシステムのその後の改良や実用化への取り組みについて講演した。日本の各社がシステムのデモのために大挙して参加したことも、各社の機械翻訳に対する期待の大きさを物語っている。AAMT が初めて主催した第 4 回 MT サミットは 1993 年に神戸で開催された。機械翻訳に対する期待にかけりがみられる時代になっていたが、EDR の電子化プロジェクト、CICC の近隣諸国間の機械翻訳プロジェクト、ATR の自動翻訳電話プロジェクトなど世界からも注目されるプロジェクトが活発に行われていた時代で、グローバルコミュニケーションのための国際協力をテーマにした意義深い会議になったことを記憶している。

その後、MT サミットは 3 地域の持ち回りによる隔年開催の会議として定着していく。筆者は少し疎遠になっていたが、AAMT 第 2 代会長・田中穂積先生の下で 1999 年に開催された第 7 回シンガポールサミットにプログラム委員として関わることになった。前々回あたりの会議から投稿ベースの一般論文のセッションも設けられるようになっていたが、機械翻訳が再び冬の時代を迎えていたことや開催地の地理的条件から、論文が集まるのか懸念された。プログラム委員長・辻井潤一先生の発案で、新世紀の MT、音声翻訳、制限言語、ローライゼーションなど多数の特別セッションを設け、それぞれの分野のリーダーに招待講演をお願いした。こうして魅力のある会議にすることで、一般論文の投稿数も増え、日本での開催とほぼ同数の 200 人余りの参加者を得て、成功裏に会議を終えるこ

とができた。

AAMT/Japio 特許翻訳研究会では、その活動の一環として、MT サミット併設の特許翻訳ワークショップを開催してきた。第 10 回ブーケットサミットから横山晶一先生を中心に回を重ねてきたが、今年の第 15 回マイアミサミットでの第 6 回ワークショップ（対象範囲を広げ、特許・科学技術文献翻訳ワークショップに改名）に筆者はオーガナイザの一人として参加した。特許や科学技術文献の機械翻訳に関わっている人は世界的にはそれほど多くなく、論文や参加者を集めるのに毎回苦労しているが、それだけにワークショップを発展させ、AAMT/Japio 研究会の活動を発信していくことが重要であるといえる。

マイアミサミットでは 30 年の長さを痛感した。基調講演ではニューラルネットワーク MT など最新の話題が取り上げられ、スピーカも若手が起用されていた。SMT の発展に貢献した Philipp Koehn 博士が IAMT 賞を受賞したことからも、MT が新しい時代を迎えているといえる。同時に、ユーザ、研究開発者、政策関係者が集う MT サミットの伝統が受け継がれており、MT サミットの草創期に関わった一人として喜ばしく感じた。さて、次の第 16 回 MT サミットは AAMT の主催により 2017 年 9 月に名古屋で開催される。24 年ぶりの日本での MT サミットがどんな会議になるのか楽しみである。

MT Summit XV 参加報告・ワークショップ開催報告

須藤克仁

日本電信電話（株）コミュニケーション科学基礎研究所

1. 概要

本稿では、2015 年秋に開催された MT Summit XV、及び筆者がオーガナイザを努めた併設ワークショップ PSLT 2015 について報告する。

2. MT Summit XV 参加報告

MT Summit XV（第 15 回機械翻訳サミット）は、2015 年 10 月 30 日～11 月 3 日に米国マイアミの Hyatt Regency Miami にて開催された。会期初日及び最終日はチュートリアルとワークショップ、本会議は中 3 日間であった。本会議には約 200 名の機械翻訳関係者（研究者、機械翻訳ソフト業者、翻訳業者、翻訳者等）が参加した。開催地の関係もあり米州からの参加者が多数を占め、日本から 20 名程度の参加があったもののその他アジア諸国からの参加は少数であった。

本会議は、研究者・商用ユーザ及び翻訳者・政府ユーザの 3 つの平行セッションと、計 5 件の招待講演、パネルディスカッション、及び技術展示を行う Technology showcase で構成されていた。

招待講演は以下の 5 件であった。

・ニューヨーク大学の Kyunghyun Cho 氏によるニューラルネットワーク機械翻訳に関する講演

・ジョンズ・ホプキンス大学（元 NAIST 助教）の Kevin Duh 氏によるニューラルネット機械翻訳のための自然言語のモデル化に関する講演

・欧州委員会(EC) の Spyridon Pilos 氏による EC での機械翻訳プロジェクトに関する講演

・Google の Macduff Hughes 氏による今後 10 年の機械翻訳の展望に関する講演



招待講演時 満席の聴衆

・ジョンズ・ホプキンス大学の Matt Post 氏による話言葉翻訳サマータワーワークショップに関する講演

全体的に研究者寄りの講演が多かったが、現時点での最先端技術であるニューラルネットワーク機械翻訳に関する話題は研究者以外の参加者からも注目を集めており、活発に議論されていた。

パネルディスカッションは David Rumsey (ATA)、Jost Zetsche (翻訳者)、Jose Palomares (Venga) の 3 氏により、翻訳者と機械翻訳の関係、主に機械翻訳が翻訳者にとって有用であるかどうかについての議論が行われた。従来の翻訳メモリの利用に加え、後編集やコンピュータ補助翻訳(computer aided translation)の導入によって今後翻訳者の仕事のやり方には変化が訪れるだろうという認識が示された。筆者にとって興味深かった議論としては、機械翻訳を利用した人手の翻訳の生産性の評価のあり方として単純に時間だけで測ることは望ましくないのではないか、人手翻訳と機械翻訳が協調していけるはずだが秘匿すべき情報の管理については課題がある、といったものがあった。

一般講演では、筆者は大部分研究トラックを聴講していた。ACL や EMNLP 等の国際会議においてはニ

ニューラルネットワーク機械翻訳に関する論文が急速にその数を増やしているが、機械翻訳サミットでは件数に比して内容が多岐に渡る印象を受けた。また、翻訳後編集に関する研究発表が聴衆からの質疑の活発さという面で目立っていたことが印象に残っている。以下、筆者が目にした論文をいくつか簡単に紹介する。

まず、機械翻訳のフレーズの分散表現に関する発表が2件（蘇州大学の Wang 氏ら、NICT の美野氏ら）あり、ニューラルネットワークを利用した句の類似度を利用して句に基づく統計翻訳を改善している。対応する両言語の句の分散表現を一致させるよう学習する、という意味では両者は非常に似通っているが、Wang 氏らが再帰型(recursive) ニューラルネットワークに基づく手法であるのに対し、美野氏らは回帰結合型(recurrent) ニューラルネットワークを利用している。単に表層のみで句を区別するとデータの不足やノイズの混入に弱くなってしまうため、似た意味の句を分散表現で近い領域に射影し、その情報を利用して頑健性を向上させるアプローチとして興味深い。

また、複雑な形態素変化の起こる言語に対する機械翻訳に関する発表が2件（コロンビア大学の Kholy 氏ら、アムステルダム大学の Tran 氏ら）あった。詳細は割愛するが、性・数・時制による形態素変化は日英の間でも正しく対応付けることが容易でなく、単に内容語が一致していればよいという段階からさらに進んでいくためにこうした技術が必要であろう。

今回の研究論文の中で異色と感じたのはニューヨーク市立大学の Zhai 氏らによる、従来のパイプライン型処理ではなく end-to-end の処理によって統計翻訳のモデル化を行う研究である。句に基づく統計翻訳では通常複数のモデルを段階を踏んで学習するが、Zhai 氏らの研究では、簡単な初期モデルから始めて、対訳データに対する強制(forced) デコーディング結果からの対訳句抽出と構造化パーセプトロンによるモデル更新の繰り返しによって一挙にモデルを学習する方法を提案している。現在は従来型の手法と同程度の結果が得られたに過ぎないが、統計翻訳の複雑さを解決する



中岩会長による次回 MT Summit の告知

一つの試みとして興味深かった。

技術展示では、機械翻訳システムのデモ (Systran や Microsoft 等)、機械翻訳を自社サービスに組み込んでいる企業の技術紹介 (IBM、SDL、eBay 等)、など計 20 ブースなどが技術や製品の紹介を行っていた。多くの参加者を集めていたのは Microsoft の Skype Translator であり、接話マイクを利用してはいたものの、騒がしい会場で英語・ドイツ語の音声翻訳がかなりの精度で動作していたことが印象的であった。

本会議最終日には IAMT による IAMT Award of Honor の表彰があり、フレーズベース機械翻訳を牽引し、Moses の開発を主導しているジョンズ・ホプキンス大学の Philipp Koehn 教授に授与された。また、今回の MT Summit は、2017 年 9 月に名古屋大学で開催されることが AAMT 中岩浩巳会長から告知された。

会議を通じて筆者が強く感じたのは、機械翻訳は産業上の応用が明確に存在する技術であって、応用によって重視されるポイントが大きく異なる、ということである。特に機械翻訳サミットや AMTA、EAMT 等の会議では機械翻訳の研究者・開発者だけでなく翻訳者や翻訳業者からの参加者も多く、応用の視点から機械翻訳を見つめ直せることを再確認した。研究トラックだけに注目すると必ずしも完成度の高い研究が並んでいるというわけではないのだが、応用、特に翻訳後編集やコンピュータ補助翻訳の利便性向上のための研究やユーザからの問題提起といった内容は機械翻訳専門の会議であるからこそのものであると感じた。

3. ワークショップ開催報告

AAMT/Japio 特許翻訳研究会の活動の一環として、2005 年の第 1 回から数えて 6 回目のワークショップを、機械翻訳サミットに併設する形で会議初日の 10 月 30 日に開催した。今回は特許翻訳に限定せず幅広く技術文書に関する機械翻訳の課題を扱うという観点から、特許・技術文書翻訳ワークショップ(Workshop on Patent and Scientific Literature Translation: PSLT 2015) と題した。ワークショップ co-chair は梶博行先生(静岡大学大学院教授)と須藤の 2 名が担当した。参加者数は 20 名強であった。

セッション 1: MT in Patent Organizations

本セッションでは、公的な知財担当機関における機械翻訳の活用についての 2 件の招待講演をお願いした。公的機関において機械翻訳が有効に活用され、必要な知的財産の情報に簡単にアクセスできるようになることは技術や産業の発展に重要であり、この数年での大きな進展を知ることができ非常に有益であった。

1 件目の招待講演は世界知的財産機構(World Intellectual Property Organization: WIPO) の Bruno Pouliquen 氏による、WIPO の機械翻訳についての講演であった。WIPO では多言語の特許文書から自動的に対訳コーパスを構築し、統計的機械翻訳によって発明の名称と概要の機械翻訳を行い、他言語での特許検索を可能にするサービスを提供している。多言語データの規模が膨大であることから、学習はすべて自動化している、また高速・省メモリな翻訳を実現するためのデータ選択や、ドイツ語・日本語等一部言語での事前並べ替え等も行っている、といった、システムを運用する上での様々な工夫が紹介された。2 件目は特許庁の加藤啓氏による、特許庁での機械翻訳に関する取り組みについての講演であった。特許庁では機械翻訳の評価基準の策定、また中国語特許への対応を見据えた日中対訳用語辞書の整備などを進めている。また 2015 年からは中国・韓国の特許に対応した特許

検索システムを公開していること、米国・欧州・中国・韓国の各特許庁との情報共有を行う OPD(Open Portal Dossier) において日本語から英語への機械翻訳が活用されていることなどが紹介された。

セッション 2: Effective Use of Patent MT

本セッションでは機械翻訳の活用による翻訳者支援をテーマとし、Iconic Translation Machines Ltd. の John Tinsley 氏に招待講演をお願いした。実際のビジネスとして考えた場合には生産性やユーザにとっての使いやすさが重要であり、後編集に渡す機械翻訳がいかにあるべきか、ということについて考えさせられた。講演では、Google のようにターゲットを絞らない機械翻訳とは対極的に、分野適応をすることで特許のような特殊な分野の機械翻訳の精度改善が可能で、後編集による翻訳の生産性の大きな向上につながる、と実例を示しながら説明があった。彼らのシステムでは、用語辞書やキーワード抽出、ルールベース翻訳、統計翻訳などの様々なモジュールを組み合わせ、後編集を含めたワークフローの生産性が高まるように調整が行われている。生産性の評価には TAUS の Dynamic Quality Framework (DQF) を利用し、後編集が直接翻訳よりも生産性が高まることを確認している。また、参照訳に対する翻訳編集率(Translation Edit Rate: TER) が 40% を下回るくらいになると後編集の効率がよくなるなどの分析結果が示された。

セッション 3: Challenges for Advanced Patent MT

本セッションでは、特許機械翻訳のさらなる改善について、1 件の招待講演と 2 件の技術論文の講演が行われた。これまでのセッションでは機械翻訳の使われ方が議論の対象であったが、本セッションは機械翻訳の改善のための技術が焦点となった。

最初に、独ハイデルベルク大学の Stefan Riezler 教授による、ユーザフィードバックに基づく統計翻訳の改善についての講演をお願いした。翻訳後編集結果を利用して統計翻訳を改善する研究はこれまでも様々

行われているが、彼らの研究では、従来研究が前提としていた専門翻訳者の後編集結果ではなく、ユーザの局所的なフィードバック（例えば翻訳結果に対する品質推定値、実験では参照訳に対する文単位 BLEU 値）のみを利用する方法を提案している。基本的な考え方は構造化パーセプトロンと同様で、フィードバックで与えられる損失が小さくなるようにモデルパラメータを更新するというものである。この手法により、議会議事録からニュースへの分野適応を行った場合の翻訳精度が改善することが示された。

技術論文の 1 件目は、京都大学の John Richardson 氏らによる、機能語の翻訳誤り訂正の研究であった。機械翻訳において前置詞や関係詞といった機能語はしばしば正しく翻訳できないことがあるため、構文木から構文木への翻訳における出力構文木に対する編集操作によって機能語の誤り訂正を行う手法が提案された。

技術論文の 2 件目は、筑波大学の Zi Long 氏らによる、日中パテントファミリーからの専門用語対訳知識獲得の研究であった。日本語と中国語の間で対訳用語獲得を行う場合の、それぞれを単語分割して対応する用語を探すという既存手法に対し、本研究では中国語側を文字単位で分割し、単語分割の誤りや曖昧性によって対訳用語の検出漏れを減らす手法を提案した。

セッション 4: Beyond Patent Translation

最終セッションは特許以外の技術文書の機械翻訳に向けての取り組みをテーマとして、1 件の招待講演と 2 件の技術論文の講演が行われた。

招待講演は、科学技術振興機構(JST)の中澤敏明氏による、中国と日本の間での機械翻訳を通じた科学技術情報交換の実現に向けた取り組みについての紹介であった。JST と京都大学、中国科学技術情報研究所(ISTIC)が共同で進めている日中間の機械翻訳のための言語資源や言語解析エンジン、機械翻訳エンジンの整備・開発を行うプロジェクトの計画と現状についての説明と、2015 年 10 月に開催された第 2 回アジア言語ワークショップ(WAT)の結果紹介が行われた。

技術論文の 1 件目は、愛媛大学の野口敬輔氏による、大量の特許分野対訳データを利用した他分野への分野適応についての研究であった。特許分野では比較的容易に非常に大量の対訳データが得られるため、共変量シフトを用いたデータ重み付けを利用して新聞分野に近い特許対訳データ中の文を選択し、新聞分野の翻訳を改善する手法が提案された。

技術論文の 2 件目は、静岡大学の綱川隆司先生による、特許文書中の専門用語の Wikification (Wikipedia エントリへの対応付け)についての研究であった。本研究では、日本語の特許文書中の専門用語に対して適切な日本語版 Wikipedia エントリが存在しない場合でも対応する英語版 Wikipedia エントリへ対応付けできるように、日本語の専門用語の英訳候補を統計翻訳のフレーズテーブルから生成して対応する英語版 Wikipedia エントリを探す、という手法が提案された。

本ワークショップでは様々な立場の招待講演者からの講演があり、特許・技術文書翻訳の現状や課題について議論された。Pouliquen 氏の講演で指摘されたように、現状の特許機械翻訳は assimilation (同化:他国語を母語に翻訳し自らが理解できるようにすること)の目的であれば、いくつかの言語対で実用に耐える水準の翻訳が可能になりつつある一方、dissemination (異化: 母語を他国語に翻訳し、他国語話者に理解してもらえるようにすること)の目的では、読者が特定されないゆえに比較的高水準の翻訳が求められることから、まだ少し先の未来の話であろう。国際出願や詳細な技術調査の観点では高水準の翻訳が求められることは間違いなく、人手での翻訳は欠かせない。昨今の機械翻訳、統計翻訳の急速な進化によって、特許のような対訳資源が豊富かつ文体が整っている分野においては翻訳後編集やコンピュータ補助翻訳等が人手での翻訳の生産性を大きく向上できる可能性が高く、また各種技術文書も同様の期待がある。そのため、本ワークショップの領域は非常に有用なものであり、本ワークショップは今後も継続して開催すべきものとする。

第2回アジア翻訳ワークショップ (WAT2015) 開催報告

中澤 敏明	美野 秀弥	後藤 功雄	Graham Neubig	黒橋 禎夫	隅田 英一郎
JST	NICT	NHK	NAIST	京都大学	NICT

1. はじめに

アジア翻訳ワークショップ (Workshop on Asian Translation, WAT) はアジア言語を中心とした評価型機械翻訳ワークショップであり、2014年に第1回 (WAT2014) が開催され[1]、2015年10月16日に第2回 (WAT2015) が開催された[2]。WAT2015では機械翻訳評価に参加した各チームによる翻訳システムの説明のほか、Institute for Infocomm Research in SingaporeのResearch DirectorであるHaizhou Li氏、及び情報通信研究機構ユニバーサルコミュニケーション研究所多言語翻訳研究室長である隅田英一郎氏による招待講演も行われ、50名を超える参加者が集まり盛況であった。

WAT2015で新たに行ったことは以下のことである。

- ・中→日、韓→日の特許文翻訳サブタスクの追加
- ・特許庁による特許文献機械翻訳の品質評価手順に基づいた人手翻訳

本稿ではWAT2015の概要や結果などをまとめて報告する。なお評価結果の詳細や各参加チームの翻訳システムの説明などはWAT2015のウェブサイトを確認することができる¹。また投稿されたすべての論文はACL Anthologyに掲載されているので、参考にさせていただきたい²。

2. データセット

WAT2015ではWAT2014に引き続き、アジア学術論文抜粋コーパスASPEC³を用いた日⇔英、日⇔中の科学技術論文翻訳のほか、新たに特許庁より提供されたJPO特許対訳コーパス (JPO Patent Corpus, JPC)

を用いた中→日、韓→日の特許文翻訳サブタスクが追加された。JPCは、2012年と2013年に公開された中国語と日本語の特許文書の明細書(CJ)、及び2011年、2012年、2013年に公開された韓国語と日本語の特許文書の明細書(KJ)の対訳文の一部である。各コーパスに含まれる分野の割合は、化学、電気、機械、物理がそれぞれ25%ずつであり、データ量は訓練データが100万文、開発データ、開発試験データ、試験データがそれぞれ2000文ずつである。各対訳文は対訳特許文書から内山・井佐原の方法[3]により対訳文としての類似度が高い順に自動的に抽出されたものである。ただし試験データのみ人手でスクリーニングを行い、対訳コーパスとしての質を高めている。

3. ベースラインシステム

比較対象となるベースラインシステムとして、フリーソフトウェアの統計的機械翻訳ツール Moses、市販のルールベース翻訳ソフトウェア、オンライン翻訳サービスを用いた。統計的機械翻訳手法はフレーズベース、階層フレーズベース、String-to-Tree (日英、日中のみ)、Tree-to-String (英日、中日のみ) の4つを用意した。各翻訳手法のパラメータ設定などは、WATのページで公開されている⁴。ルールベース機械翻訳としてThe 翻訳 V15 (日英、英日)、ATLAS V14 (日英、英日)、PAT-Transer 2009 (日英、英日)、J北京7 (日中、中日)、蓬萊 2011 (日中、中日)、Jソウル9 (韓日)、高麗 (韓日) を利用した。オンライン翻訳サービスとしてはGoogle翻訳及びBing翻訳を利用した。なおルールベース機械翻訳及びオンライン翻訳サービスのシステムIDは匿名化されており、特定できないよ

¹ <http://lotus.kuee.kyoto-u.ac.jp/WAT/>

² <http://aclanthology.info/events/ws-2015#W15-50>

³ <http://lotus.kuee.kyoto-u.ac.jp/ASPEC/>

⁴

<http://lotus.kuee.kyoto-u.ac.jp/WAT/baseline/baselineSystems.html>

うにしている。

4. 翻訳評価

翻訳評価は自動評価と人手評価を行い、人手評価はさらに2つの方法で行った。

4.1 自動評価

WATでは自動評価サーバーを用意し、参加チームが機械翻訳結果の自動評価結果をいつでも確認できるようにしている。自動評価システムはワークショップ終了後も利用可能であり、引き続き翻訳結果の提出・修正を受け付けている。

自動評価尺度としてはBLEU[4]及びRIBES[5]を用いている。日本語と中国語に関しては単語分割基準の違いによりスコアが変化するため、いくつかの単語分割ツールを使って翻訳結果を単語に分割し、それぞれの基準でのスコアを算出する（詳細はWATのウェブサイト⁵参照）。

4.2 人手評価

人手評価はWAT2014と同様、クラウドソーシングを利用することでコストを削減した（Pairwise Crowdsourcing Evaluation）。クラウドソーシング評価はフレーズベース統計的機械翻訳システムをベースラインとして、各翻訳システムの結果を文ごとにベースラインと一対比較し、その勝敗数をスコア化することで行う。WAT2014では各文ペアにつき3名の評価者で評価を行い、評価対象システムの翻訳がベースラインよりも良いという判断を+1、悪いという判断を-1、同程度を0としたとき、全ての評価者の判断を足し合わせて正の値となれば最終判断を勝ち、負の値ならば負け、0ならば同程度とした。WAT2015では5名の評価者で評価を行い、全ての評価者の判断を足し合わせて+2以上となれば最終判断を勝ち、-2以下ならば負け、それ以外（-1以上1以下）ならば同程度とすることで評価の信頼性を高めた。この基準の変更により同程度と判断される割合が増えることになり、WAT2014とWAT2015の結果を直接比較することは

できないが、[2]ではクラウドソーシング評価データをサンプリングすることで、WAT2014と同じ設定とした比較も行っている。

評価者を増やしたものの、クラウドソーシングによる評価では評価者が翻訳の専門家ではないため、依然として評価結果が不安定であるなどの問題がある。そこでWAT2015ではサブタスクごとにクラウドソーシング評価上位3チーム（韓→日のみ特別に4チーム）の翻訳結果に対して、専門家による評価（JPO Adequacy Evaluation）も行った。専門家による評価の基準として、特許庁が公開している「特許文献機械翻訳の品質評価手順」⁶の中の「内容の伝達レベルの評価」に従って行った。これは機械翻訳結果が原文の実質的な内容をどの程度正確に伝達しているかを、人手翻訳の内容に照らして、下記5段階の評価基準で主観的に評価するものである。

- 5: すべての重要情報が正確に伝達されている。（100%）
- 4: ほとんどの重要情報は正確に伝達されている。（80%~）
- 3: 半分以上の重要情報は正確に伝達されている。（50%~）
- 2: いくつかの重要情報は正確に伝達されている。（20%~）
- 1: 文意がわからない、もしくは正確に伝達されている重要情報はほとんどない。（~20%）

クラウドソーシング評価はテストセットの中からランダムに選択された400文を対象としているが、内容の伝達レベルの評価ではそのうちのさらに200文を対象を絞って行った。

5. 参加チームリスト

表1にWAT2015に翻訳結果を提出したチームと、各チームが参加したサブタスクの一覧を示す。国外から3チームが参加し、企業から4チームが参加した。

⁵ <http://lotus.kuee.kyoto-u.ac.jp/WAT/evaluation/>

⁶ https://www.jpo.go.jp/shiryoutoushin/chousa/tokkyohonyaku_hyouka.htm

6. 評価結果

図 1 に内容の伝達レベルの評価結果 (Adequacy) を示す。全体的に見ると、以下のような傾向があることがわかる。

- ・日本語「から」の翻訳よりも、日本語「へ」の翻訳の方が精度が高い

- ・日⇔中よりも、日⇔英の方が精度が高い

- ・韓→日は非常に精度が高く、実用レベルと言える

また図 2 に 2 種類の自動評価 (BLEU、RIBES) と 2 種類の手評価 (Adequacy、Crowd) の結果をまとめたグラフを示す。Crowd は日⇔英評価においては Adequacy と正の相関を持っているが、日⇔中評価においては必ずしも正の相関を持ってはいない。これはクラウドソーシング評価者の多くが日本人であり、日⇔英よりも日⇔中評価が難しいためである可能性がある。またクラウドソーシング評価者は文字の全角半角の違いなど、文意に影響しないような点を評価している可能性があり、これが結果に反映されている (特に韓→日の System C) とも考えられる。

2 つの自動評価手法は、内容の伝達レベルの評価と概ね相関していると言えるが、韓→日の System D の評価において大きく乖離している。この現象の正確な原因はまだ特定できていないが、可能性の一つとして挙げられるのが、不必要な語順変更である。韓国語と日本語は語順の変更がほとんど不要である上に、そもそもの翻訳精度が高いため、ちょっとした語順変更誤りが人手評価に悪影響を与える可能性が高い。つまり文の途中の翻訳は参照訳と完全に一致しているが、先頭に来るべきものが末尾に来てしまったりすると、全体的な意味が不明となり、自動評価では高いスコアとなるが人手評価では誤った翻訳と判断されるのである。一般的に「自動評価スコアは高ければ高いほど人手で見た時の翻訳精度も高いが、自動評価スコアが低いからといって悪い翻訳とは限らない」と言われるが、今回得られた結果は「自動評価スコアが高くても良い翻訳とは限らない」というものであり、非常に興味深い。今後さらに分析を進める予定である。

なお本稿では述べていない他の評価指標や、全参加チームの全評価結果などは[2]にまとめられているので、参考にさせていただきたい。

7. まとめと今後の展望

本稿では WAT2015 の概要や結果などについて概説した。12 チームが参加し、様々な手法での翻訳結果が集まり、これらを分析することで様々な知見が得られた。提出された翻訳結果は近日中に一般公開する予定である。

WAT は今後も毎年開催する予定であり、WAT2016 では新たな言語対を追加できる見通しが立っている。なお今回は行えなかったが、文脈を考慮した翻訳評価も今後行っていきたいと考えている。

参考文献

- [1] Toshiaki Nakazawa, Hideya Mino, Isao Goto, Sadao Kurohashi, and Eiichiro Sumita. 2014. Overview of the 1st Workshop on Asian Translation. In Proceedings of the 1st Workshop on Asian Translation (WAT2014), pages 1-19.
- [2] Toshiaki Nakazawa, Hideya Mino, Isao Goto, Graham Neubig, Sadao Kurohashi, and Eiichiro Sumita. 2015. Overview of the 2nd Workshop on Asian Translation. In Proceedings of the 2nd Workshop on Asian Translation (WAT2015), pages 1-28.
- [3] Masao Utiyama and Hitoshi Isahara. 2007. A Japanese-English Patent Parallel Corpus. MT summit XI.
- [4] Kishore Papineni, Salim Roukos, Todd Ward, and Wei Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In Proceedings of ACL, pages 311-318.
- [5] Hideki Isozaki, Tsutomu Hirao, Kevin Duh, Katsuhito Sudoh, and Hajime Tsukada. 2010. Automatic evaluation of translation quality for distant language pairs. In Proceedings of the 2010

表 1 : WAT2015 に翻訳結果を提出したチームと参加タスク一覧

チーム ID	組織名	科学技術論文				特許	
		JE	EJ	JC	CJ	CJ	KJ
BJTUNLP	北京交通大学				✓		
EHR	Ehara NLP Research Laboratory		✓		✓	✓	✓
Kyoto-U	京都大学	✓	✓	✓	✓	✓	
NAIST	奈良先端科学技術大学院大学	✓	✓	✓	✓		
naver	NAVER		✓				✓
NICT	情報通信研究機構	✓					✓
ntt	NTT コミュニケーション科学基礎研究所					✓	
Sense	ザールラント大学 & 南洋理工大学	✓	✓				✓
TMU	首都大学東京	✓					
TOSHIBA	東芝	✓	✓	✓	✓	✓	✓
WASUIPS	早稲田大学					✓	
WEBLIO-EJ1	ウェブリオ		✓				

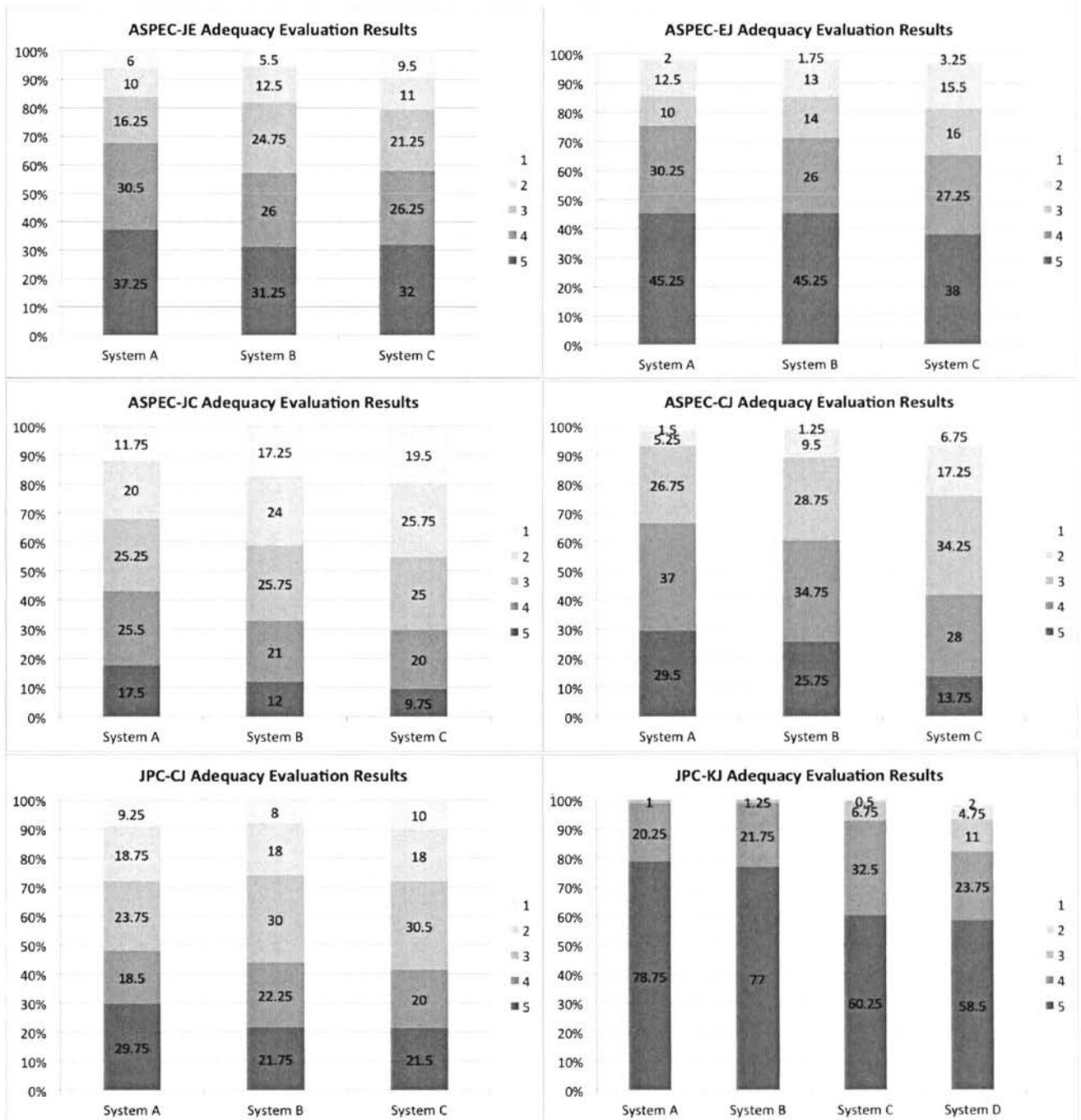


図 1 : JPO 基準による「内容の伝達レベルの評価」結果 (チーム名は匿名化してある)

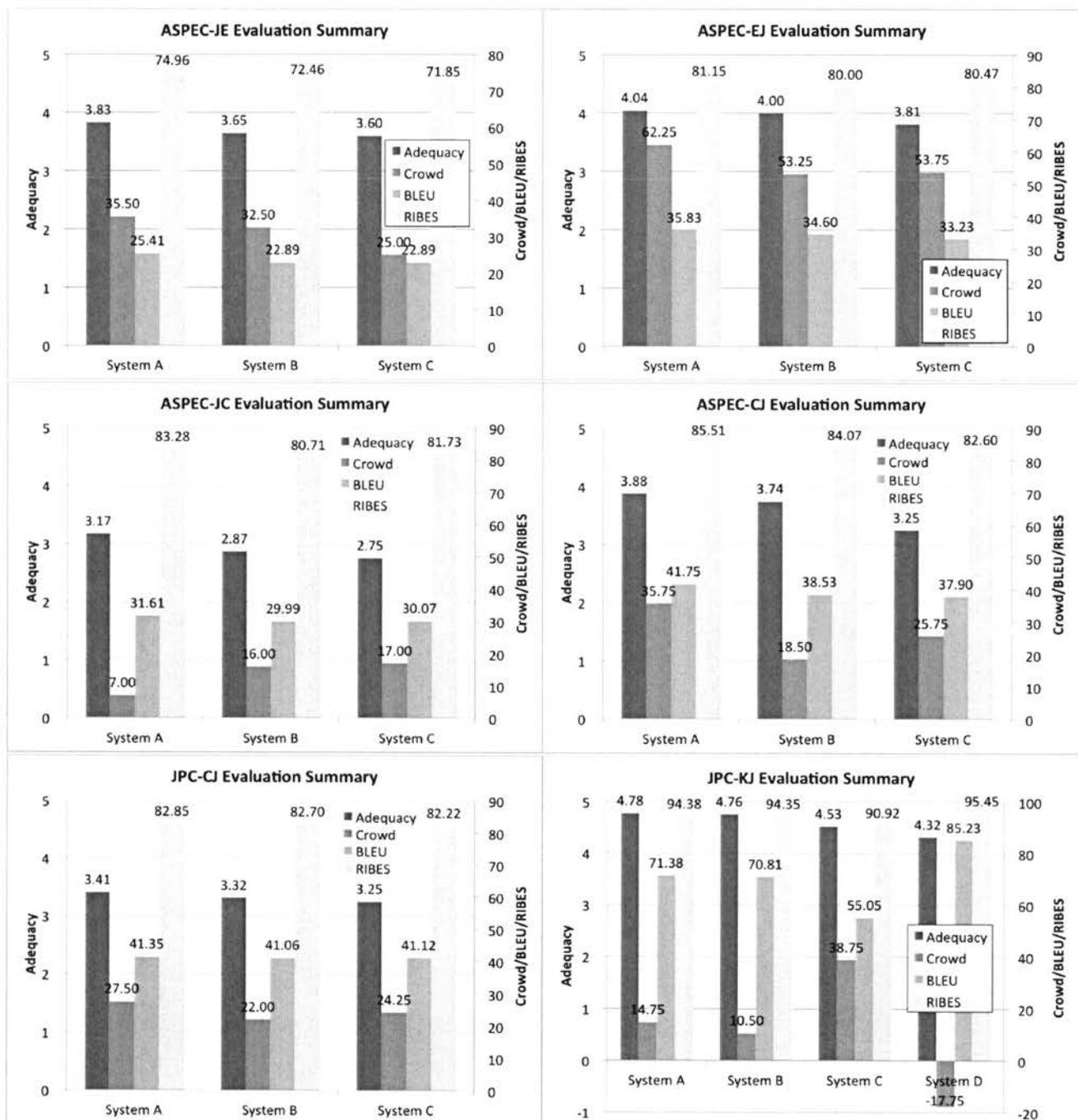


図 2 : 上位 3 チームの評価結果サマリー (チーム名は匿名化してある)

機械翻訳事始め[VI]

坂本義行

アジア太平洋機械翻訳協会理事

今回は、日本で初めて、機械翻訳の国家プロジェクトとして、1982年度から90年度にかけて、科学技術庁の科学技術振興調整費で行われた、「日英科学技術文献の速報システムに関する研究」について、文献¹⁾と文献²⁾をもとに、報告します。

1. はじめに

言語の機械翻訳は長年の夢であったし、現在もそうである。言語は人間の頭脳活動の最も具体的な表れであり、言語活動は人間の頭脳活動のメカニズムが解明されない限り、十分に解明されることはないだろう。したがって、言語の機械翻訳という問題に直接取り組むことは無謀であるといえることができる。しかし1950年代の終わりごろから始められた機械翻訳の研究は、途中う余曲折はあったものの、多くのことを明らかにしてきた。したがって、これを総合化してしっかりしたシステムを作れば、ある程度使えるシステムが作れるだろうという期待をここ十年間くらい前から持っていたのである。しかしそれを具体的に示すためには、いくつかの条件が必要であった。それは、強力な計算機処理能力を使えること、かなりの人数の研究者による数年間の集中した研究開発作業が必要なこと、そのための研究費と研究体制をどうするかという問題であり、これは1大学の1研究室ではできないものではなかった。

一方、機械翻訳に対する社会的要請は、ここ数年非常に大きくなってきており、また、日本の科学技術活動の実態を諸外国に知らせることが、国としても必要となってきた。幸いにも国はこのような状況をよく認識し、科学技術会議は科学技術振興調整費の一部を用

いて昭和57会計年度(1982年度)から3年間(その後予算などの都合で4年間に修正された)にわたって、「日本科学技術文献速報システムに関する研究」(通称:MUプロジェクト)を発足させることを決定した。

2. 科学庁の研究の背景と成果

科学技術の高度化に伴い、研究者がより創造的かつ効率的な研究活動を進めていく上で、膨大な量の最新の科学技術情報を随時利用することが不可欠となってきた。このような科学技術情報のうち外国語で書かれた文献(海外科学技術文献)が全体の70%を占めており、これらの文献を利用する際には翻訳の需要が非常に強く、文献量の増大によりこれら文献の翻訳の機械化が求められている。

また、わが国の科学技術の発展に伴い、またこれらが国際的にも認められてきたことから、わが国の科学技術文献の海外への提供が強く求められてきているが、言語の障害により十分にこれに対応できない状況である。

これらの問題を解決するために、近年進歩の著しい情報技術を利用した日英科学技術文献の速報システムの実現が望まれている。

このため、科学技術庁は科学技術会議の方針に沿って昭和57年～60年度の4年間にわたり、科学技術振興調整費により、「日英科学技術文献の速報システムに関する研究」を実施しました。本研究は、諸外国との科学技術文献交流促進の必要性に鑑み、それら文献(抄録)の翻訳を効率的に行うために、科学技術用語辞書データベースを用いて科学技術文献を翻訳するための構文パターン方式による日英(英日)言語処理ソフト

ウェアの開発、科学技術分野の専門用語を収集、整理し、コンピュータで利用可能な形に編集した日英科学技術用語辞書データベースの開発、そして、これらを統合して利用する翻訳システムの開発、を目標に進められました。

そのうち、構文パターン変換方式による日英(英日)言語処理ソフトウェアの開発においては、強力な機械翻訳用ソフトウェアシステム GRADE を開発し、これによって日英・英日翻訳文法を作成し、また、形態素解析と形態素生成の各サブシステムを開発しました。日英科学技術用語辞書データベースの開発においては、名詞辞書として7万語の日英翻訳辞書、2万語の英日翻訳辞書、動詞辞書として5千語に及ぶ翻訳辞書を作成した。また、総合システムの開発では言語処理システムと翻訳辞書を統合して日英・英日翻訳を行う総合システムの開発を行った。

昭和62年3月
科学技術庁科学技術振興局長 藤咲浩二

3. 研究の趣旨

- (1) 科学技術は、ますます高度なものとなり、その知識の質・量ともに加速度的に増大している。研究者がより創造的かつ広範な分野にわたる膨大な量の最新の科学技術情報を随時利用することが不可欠となっている。
- (2) わが国においては、科学技術文献サービスは、日本科学技術情報センター、通商産業省工業技術院情報計算センター、農林水産省農林水産研究情報センター、大学等で行われているが、こうした機関では近年の科学技術の高度化、研究活動の活発化により取り扱う文献の数が極めて膨大なものとなっている。
そして、外国語で書かれた文献(海外科学技術文献)が、全体の70%を占めるという状況にある。
- (3) 従来このような海外科学技術文献を国内で利用するには翻訳の需要は非常に強く、多くの文献が

翻訳され、利用されている。またこれらの文献をデータベース化するには専門の翻訳技術者によりアブストラクト等の翻訳を行ってきた。

しかし、量の増大によりこうした作業の機械化、省力化が求められている。

- (4) 一方、わが国の産業活動、科学技術活動が益々高度化し、国際化するに伴い、わが国の科学技術情報は益々国際的に利用されつつあり、科学技術情報活動の国際化が必須の問題となってきた。
- (5) とくに、近年わが国の高度な科学技術水準に注目した欧米先進国や、わが国が研究協力等で重要な役割を演じている発展途上国等から、わが国の科学技術文献の提供が強く要請されているが、言語が障壁となって十分これに対応できない状況にあり、対応策が求められている。
- (6) こうした問題を解決するために近年進歩の著しい情報技術を活用した日英科学技術文献の速報システムの実現が望まれる。

本研究は、現在実用化されているコンピュータ技術を前提に構文パターン変換方式による速報システムを構築することにより最終的には、科学技術文献アブストラクトの翻訳作業を支援することをねらいとする。

4. 研究概要

諸外国との科学技術文献交流促進の必要性に鑑み、それら文献(抄録)の翻訳を効率的に行うため

- ① 科学技術用語辞書データベースを用いて科学技術文献を翻訳するための構文パターン変換方式による日英(英日も含む。以下同じ)言語処理ソフトウェアの開発。
- ② 科学技術分野の専門用語を収集、整理し、コンピュータで利用可能な形に編集した日英科学技術用語辞書データベースの開発。
- ③ これらを統合して利用しうる翻訳システムの開発、を行う。

(1) 言語処理システムの開発

① 構文パターン変換方式の開発

既存のコンピュータ技術を前提に、科学技術文献を言語処理する構文パターン変換方式の手法を開発する。

② 日英言語処理ソフトウェアの開発

構文パターン変換方式の手法を利用して、日本語から英語へ、また英語から日本語へ科学技術文献を言語処理するためのソフトウェアを開発する。

(2) 日英科学技術用語辞書データベースの開発

① 日英専門用語集の分析、対訳シソーラスの作成、既存文献からの対訳語の抽出及びこれらの辞書データへの編成

科学技術分野全般の日英専門用語集の分析、対訳シソーラスの作成及び既存文献からの対訳語抽出を行い、これらのデータを日英科学技術用語辞書作成のための入力データに編成し、データベース化する。

② 科学技術用語辞書作成・管理システムの開発

日英科学技術用語辞書データベースの設計を行い、当該辞書データベースを作成・管理するシステムを開発する。

③ 翻訳のための動詞辞書の作成

文章に現れる動詞、形容詞等の翻訳辞書を作成する。

(3) 総合システムの開発

① ファイル管理システム及び編集ソフトウェアの開発

ファイル管理、更新を行うソフトウェア及び入出力時の編集を行うソフトウェアを開発する。

② 総合運用管理システムの開発

以上を総合的にシステム化した利用、運用のための管理システムを開発する。

(4) 研究推進

上記の研究の効率的推進のため研究実施計画、実施状況等必要な事項について連絡調整を行うとともに研

究結果の検討を行う。

この目的にそって、科学技術文献の日英翻訳システム、英日翻訳システムの2つのシステムを作ることになった。文献の分野としては種々検討の末、われわれシステム開発者が理解しやすい電気工学分野をとることにし、日本語文献としては日本科学技術情報センターの電気工学分野の文献速報のタイトルおよび抄録文、英語文献としては、INSPECの同じ分野の抄録をとることにした。研究開発グループとしては、京都大学、電子技術総合研究所(ETL)、日本科学技術情報センター(JICST)、

工業技術院筑波情報計算センター(RIPS)の4者が協力して研究開発することになった。研究開発の分担は次のようである。京都大学は機械翻訳用ソフトウェアの開発、解析、変換、生成の文法規則の開発、解析・変換・生成の文法規則の開発・解析・変換・生成に用いる辞書情報の種類の決定、各単語に対してそれらの情報を与える作業基準の作成、変換部分の辞書データの一部の作成などを担当した。ETLは形態素解析、形態素生成のソフトウェアとその辞書の作成、解析のための用言辞書の作成を担当した。JICSTは主として体言辞書の作成を担当した。体言は一般用語のほかにも大量の電気工学分野の専門用語を含む。RIPSは京都大学が作る機械翻訳システムに入力部分、テキスト出力部分、対話的修正部分、辞書の保守に関するソフトウェアなどを付け加え、機械翻訳のための総合システムを作ることになっている。

【参考文献】

1) 長尾真、辻井潤一、中村順一、坂本義行、鳥海剛、佐藤雅之、「科学技術庁機械翻訳プロジェクトの概要」、「情報処理」、Vol. 26, No. 10, 1985.

2) 科学技術庁科学技術振興局、昭和57年度～60年度 科学技術振興調整費「日英科学技術文献速報システムに関する研究」。

表1 年次計画

研究項目	1982年度	1983年度	1984年度	1985年度
(1) 言語処理システムの開発				
① 構文パターン変換方式の開発	基本・詳細設計	システム開発		性能試験 運用試験
② 日英言語処理ソフトウェアの開発	基本・詳細設計	システム開発		
(2) 日-英科学技術用語辞書				
① 日-英専門用語集の分析、対訳ソースの作成、既存文献からの対訳後の抽出及びこれらの辞書データの編成	データの収集・分析・データベース化			
② 科学技術用語辞書作成・管理システムの開発	基本・詳細設計	システム開発		
③ 翻訳のための動詞辞書の作成	データの収集・分析・データベース化			
(3) 総合システムの開発				
① ファイル管理システム及び編集ソフトウェアの開発	基本・詳細設計	システム開発		
② 総合運用管理システムの開発	基本・詳細設計	システム開発		
(4) 研究推進				
所要経費(合計)	132百万円	163百万円	162百万円	161百万円

表 2 研究実施体制及び所要経費

(単位：千円)

研究項目	担当機関	1982 年度	1983 年度	1984 年度	1985 年度
(1) 言語処理システムの開発	通商産業省工業技術院電子技術総合研究所	78,934	57,513	67,310	24,921
① 構文パターン変換方式の開発	院電子技術総合研究所(一部を京都大学に委託)	38,267 (30,587)	46,748 (43,148)	56,541 (48,901)	20,150 (19,139)
② 日英言語処理ソフトウェアの開発	通商産業省工業技術院電子技術総合研究所	40,667	10,765	10,769	4,771
(2) 日-英科学技術用語辞書データベースの開発	科学技術庁研究開発局	46,761	79,262	72,313	111,061
① 日-英専門用語集の分析、対訳ソーラスの作成、既存文献からの対訳語の抽出及びこれらの辞書データへの編成	科学技術庁研究開発局(日本科学技術情報センターに委託)	4,377	21,234	39,601	64,917
② 科学技術用語辞書作成・管理システムの開発	科学技術庁研究開発局(日本科学技術情報センターに委託)	38,544	29,700	13,200	5,280
③ 翻訳のための動詞辞書の作成	通商産業省工業技術院電子技術総合研究所	3,840	28,328	19,512	40,864
(3) 総合システムの開発	通商産業省工業技術院計画課(情報計算センター)	4,960	25,736	21,199	23,628
① ファイル管理システム及び編集ソフトウェアの開発	通商産業省工業技術院計画課(情報計算センター)	1,920	18,501	16,103	10,704
② 総合運用管理システムの開発	通商産業省工業技術院計画課(情報計算センター)	3,040	7,235	5,096	12,924
(4) 研究推進	科学技術庁研究調整局、振興局	1,023	703	869	892
合 計		131,678	163,214	161,691	160,502

表3 研究推進委員会

委 員	所 属
○長 尾 真	京都大学 工学部教授
相 沢 輝 昭	NHK 放送技術研究所画像研究部主任研究員
石 綿 敏 男	茨城大学 教養部教授
猪 瀬 博	東京大学 工学部教授
内 山 英 夫	(財)日本特許情報機構 常務理事
柏 木 寛	通商産業省 工業技術院電子技術総合研究所電子計算機部長
草 薙 裕	筑波大学 文芸・言語学系教授
坂 本 義 行	通商産業省 工業技術院電子技術総合研究所ソフトウェア部主任研究官
佐 藤 征 夫	科学技術庁 振興局管理課情報室長
鳥 海 剛	日本科学技術情報センター 参事
中 井 浩	常盤大学 人間科学部コミュニケーション学科教授
中 島 邦 雄	通商産業省 工業技術院計画課長
野 村 浩 郷	日本電信電話公社(株) 基礎研究所情報通信基礎研究部言語処理研究グループリーダー
藤 原 鎮 雄	千葉大学 理学部教授
淵 一 博	(財)新世代コンピュータ技術開発機構 常務理事・研究所長
水 谷 静 夫	東京女子大学 文理学部教授
村 田 賢 一	情報処理振興事業協会 技術センター特別研究員
渡 辺 定 久	通商産業省 工業技術院計画課研究情報管理企画室運営班長

(注：○は推進委員長、1986年3月現在)

音声翻訳の国家プロジェクト

隅田 英一郎

情報通信研究機構 (NICT)

国家プロジェクト「グローバルコミュニケーション計画」¹では、2020年の東京オリンピック・パラリンピックを「言葉の壁」のない社会のショーケースとして、多言語音声翻訳システムについて世界に発信することを目指している。本稿では同計画における多言語音声翻訳システムの到達点のイメージとそこへの道筋について説明する。

1. 背景

2015年は多くの外国人が日本を訪れ秋に「1,500万人超え」を記録し、年内に2,000万人²をうかがえる状況である。

一方、訪日外国人と日本人の間には依然として「言葉の壁」が立ちはだかつており、コミュニケーション不足による様々な問題や機会損失が多数存在している。そこで、「言葉の壁」をなくし自由でグローバルなコミュニケーションを実現するため、多言語音声翻訳技術を高度化するとともに社会へ実装するための技術の確立が必要となる。

「グローバルコミュニケーション計画 ～多言語音声翻訳システムの社会実装～」(平成26年4月11日 総務大臣発表)においては、世界の「言葉の壁」をなくしグローバルで自由な交流を実現するため、情報通信研究機構(以下、NICT)が開発した多言語音声翻訳技術を高度化し、社会実装を推進していくこととしている。

総合科学技術・イノベーション会議が取りまとめた「科学技術イノベーション総合戦略 2014 ～未来創造

に向けたイノベーションの懸け橋～」(平成26年6月24日 改定 閣議決定)において、「科学技術イノベーションが取り組むべき課題」として「個々人が言語や文化の壁を超えるための多言語音声認識や翻訳技術」が明記されている。この他にも2020年開催の東京オリンピック・パラリンピックをマイルストーンに多くの施策で多言語対策の加速がうたわれている。

2. 推進体制

グローバルコミュニケーション計画を遂行するために、NICTは2014年9月に先進的音声翻訳研究開発推進センター³をNICT内に設置し、民間企業研究者の加入によりオールジャパンの研究開発体制を確立した。

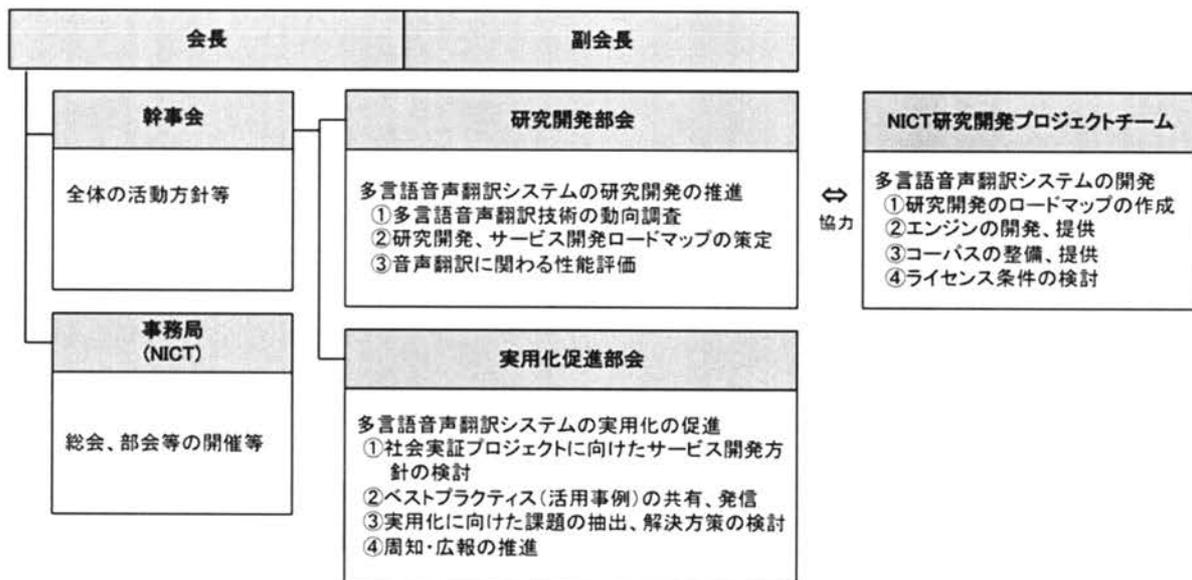
さらに、2014年12月には、産学官の力を結集して多言語音声翻訳技術の精度を高めるとともに、その成果を様々な分野のアプリケーションに適用して社会実装していくために必要な検討を行い、グローバルコミュニケーション計画の推進に資することを目的として、グローバルコミュニケーション開発推進協議会⁴が設立された。

¹ http://www.soumu.go.jp/main_content/000285578.pdf

² 「観光立国実現に向けたアクション・プログラム2014」(平成26年6月17日 観光立国推進閣僚会議)においては、2020年までに訪日外国人旅行者数2000万人にすることを目標に掲げた。

³ <http://www2.nict.go.jp/astrec/>

⁴ <http://gcp.nict.go.jp/>



(各部会の下に適宜ワーキンググループを設置)

図 1 協議会の体制

本推進協議会の設立には、経済界、情報通信事業者、学識経験者等、産学官の幅広い分野から参加があり、研究者・事業者・利用者等が結集した⁵。図 1 の体制で連携しており、研究開発部会と実用化促進部会の 2 部会で、長期的な技術開発というシーズの側面と実際の利用場面で必要となる改良や周辺技術等ニーズの側面の両面からの検討を進めている。

3. 対象分野

実用化促進部会で、有識者にご意見をいただき、多言語音声翻訳技術を社会実装する際の具体的なサービスイメージを検討している。主な対象分野は、①医療分野、②ショッピング分野、③鉄道、タクシー、宿泊施設を含む観光分野となっている。この他、分野共通の必須機能を提供するプラットフォームのあり方についても議論を行っている。

3.1. 医療

近年、外国人の旅行者・定住者が急増しており、医療機関での外国人対応が課題となっている。医師によ

る診断、看護師によるケア、薬剤師による服薬指導、等、病院内の様々な場面での利活用が期待されている。

一方、医療は「命に係わる」分野であり、誤認識、誤訳の問題への懸念がある。これに対しては二つの回答がある。一つは、本当にシビアな条件では、人間の医療通訳者に対面で対応してもらい、より深刻でない場合は、医療通訳者に電話を介して対応してもらい、通常の場合では自動通訳を使うなど、状況により棲み分けた運用を行うことが不可欠となる。逆に、例えば手術室の中は、資格の問題で人間の通訳者が入りにくいところがあり、シビアな条件でも機械の翻訳が期待される場合もある。もう一つは、医療のコミュニケーションでは、通訳は一つのチャンネルにすぎず、表情、動作、バイタル等を見てトータルで判断がなされるので、過度な心配は不要である。むしろ、言葉が一言も分からず、言葉のチャンネルが閉じた状態ではコミュニケーションが成立しないことが多く、ホスピタリティの低下や医療過誤が懸念され大きな課題となっている。

3.2. ショッピング

近年、百貨店、量販店からコンビニまで、「爆買」と

⁵ 2015.11.20 現在 133 会員

言われる勢いで海外からの旅行者による大量購入が急増している。ショッピングをしない旅行者はいないわけで、ここでのコミュニケーションの壁の解消の効果は絶大である。これまでも、電話経由の人間の通訳を利用したり、指差し会話帳を活用したり、様々な手段が試みられている。しかしながら、電話通訳に連絡するまでもなく、指差し会話帳の利用が必ずしもベストでない場合（例えば、「トイレはどこですか？」）、指差し会話帳になかったりするような場合（例えば、「〇〇の製品はどこですか？」）などに、自動音声翻訳への期待が大きい。その場で即応するスピード感が必要なことも多く、両手の自由になるウェアラブル端末への期待も大きい分野である。また、商品説明のパネルやタグの活用も重要である。音声でなく文字認識やQRコードなどと自動翻訳の連携も必要となる。さらに、リコメンデーションのシステムや在庫管理システムとの連携など統合的なサービスが期待される分野である。また、免税制度の改正への対応、レストランや食品売り場ではハラルなどタブーへの対応も関連してくる。

3.3. 観光

普通の人々にとって、最重要な分野が観光である。言葉の壁の克服は、観光体験を一段上の快適なものに変えることが期待されるし、結果、リピータを増やして、日本の観光立国に多いに貢献する。

人間の観光ガイドであれば、例えば英語等の一つの言語で説明することになるが、自動通訳システムであれば、言語の異なる観光客の個々の言語に合わせて、説明を同時に多言語で配信することも可能になるし、ガイドは日本語さえできればよいので、観光ガイドの様々な制約がとれ、ガイドの利用拡大が期待される。

訪日外国人が、日本国内で観光等のために交通機関を利用して移動する場合、タクシー・電車等の利用が想定される。鉄道では、乗換案内、切符・ICカードの払い戻し、忘れ物対応、車内・駅内での遅延等の案内放送などの場面で音声翻訳の活用が期待される。タクシーの中では、ドライバーは前方を見ている必要が

あるので、翻訳結果の表示や翻訳に係るやりとりについて工夫が必要となる。鉄道やタクシーで共通の課題となるのは騒音である。大きい騒音の中で音声認識を成立させるために、マイクなどハードウェアの工夫やソフトウェアによる雑音抑圧が重要課題となる。

また、個人営業の多いタクシーや小規模な宿泊施設では運用費用も問題となる。このようなビジネスモデルの検討も必要である。

4. 技術課題

NICT が開発してきた多言語音声翻訳システムは、短い旅行会話の翻訳を比較的精度よく実現している。他方で、この多言語音声翻訳システムを実際の社会において実用レベルで使うためには、多言語音声翻訳システムの性能向上に加えて、実際の社会における実証実験を通じて得られた知見をフィードバックしながら、研究開発に取り組む必要がある。ここで2つの技術を例に説明する。

4.1. 雑音抑圧技術

① 概要

医療やショッピングや観光地など屋内外の様々な場面、タクシーや鉄道等の公共交通機関等の騒音下での多言語音声翻訳システムの利活用では、様々な擾乱がマイクに入力され、その影響で音声認識精度が劣化し、翻訳精度も劣化を引き起こす。

その悪影響を軽減するため、発話者がマイクに入力した音声情報から雑音成分を抑圧し、発話者の音声をより明瞭化した上で、当該音声情報を多言語音声翻訳システムへと入力する雑音抑圧技術の研究開発が不可欠になる。例えば、小売店、量販店、百貨店、交差点、乗車券窓口、鉄道、バス、タクシーの走行中の会話ができることを目指す必要がある。

② 技術課題

雑音抑圧技術は環境（例：雑音の種類、大きさ、音声入力デバイス、認識エンジンの音響モデル）に大きく依存する。高雑音環境下では、雑音抑圧を行わなけ

れば精度高く認識できないが、一方で雑音抑圧量を上げ過ぎると利用者の音声自体も抑圧してしまい、認識結果の誤りを増加させる場合もある。このため、利用環境によって適切な動作パラメータを設定する必要がある。音声入力デバイスに内蔵した複数のマイクにより発話者の音声を指向性集音することで雑音下での認識精度の改善が期待されるが、これを有効にするためには指向性の向きと範囲を最適化する必要がある。利用者の音声自体を抑圧せずに、十分な雑音抑圧を達成するためには、発話・騒音を識別して分離する音源分離技術、及び発話者の位置と周囲の騒音の位置・音量に基づき、指向性の向きと範囲を制御する適応ビームフォーム技術を、利用者の状況に応じてリアルタイム処理する必要があると考えられるが、現状は音声入力デバイス毎にチューニングした固定方向・固定範囲の指向性となっている。発話者の位置と周囲騒音の位置・音量の変化に追従して雑音抑圧を実現可能とする適応集音技術を確立する必要がある。

4.2. 効率的に翻訳精度を上げる半自動学習技術

① 概要

対訳コーパスに基づく自動翻訳技術が進展したが、高精度が実現できている専門分野向けシステムにおいても、モデルの不完全性等に起因して一定量の翻訳の誤りは避けられないのが現状である。自動翻訳技術は仮名漢字変換システムのように自動的にチューニングされエラーが削減されることはなく、現状ではログのオフラインでの分析とそれに基づく改良が必要であり迅速には成し得ないという点に限界がある。こうした状況において、誤訳と正訳の識別技術、及び、誤訳を正訳に変える技術の研究開発が不可欠である。初期の翻訳率が70%の場合で、10%改善して80%となるような効果が期待される。

② 技術課題

ア) 自動翻訳結果に信頼度を付与する技術 複数の機械翻訳結果の出力や翻訳の信頼度に関する

自動学習機能を活用することで、自動翻訳結果に信頼度を付与する技術を確立する。また翻訳結果に対してユーザーが評価をつけるなどの外部入力を用いて信頼度を付与する技術を創出する。

イ) 自動翻訳結果の信頼度を利用してシステムを改良する技術 入力された翻訳対象の文章と、それに対する信頼度の高い翻訳結果の集合からモデルを生成し、最初に作られた翻訳モデルと混合し、自動的に改良する技術を確立する。翻訳結果の信頼度の低い文を自動選択し、人手翻訳することで、翻訳量を削減する。ある。こうした状況において、誤訳と正訳の識別技術、及び、誤訳を正訳に変える技術の研究開発が不可欠である。初期の翻訳率が70%の場合で、10%改善して80%となるような効果が期待される。

ウ) フォールバック機能を利用したシステムの改良に関する技術 翻訳の信頼性が低い時に、ユーザーが人手翻訳を呼び出す機能（フォールバック）を実現することで、フォールバックによって得られる翻訳と原文の対を機械翻訳の改良に自動的に利用できるような技術を開発する。

エ) 複数フィードバックを利用したシステムの改良に関する技術 同じ文章について複数の自動翻訳結果を生成し、どの翻訳結果のどの部分を集めれば正しい意味の翻訳となるかに関して、自動改良する技術の研究。また、複数の評価者の評価結果の利活用も検討する。

5. 実証実験

初めに述べたように、政府は、訪日外国人を増やすことを目標としており、これを加速する手段の一つとして多言語音声翻訳システムの研究開発及び社会実装に取り組むこととしている。

都市のインフラとして多言語サポート整備の必要性

が急速に高まっていて、特に災害が発生した時や病気になった時などでも、日本語を話すことができない外国人の安心・安全を確保するためには「言葉の壁」を越えた意思疎通が不可欠であるが、訪日外国人の使う多言語の通訳・翻訳サービスを人手で提供するのには限界がある。このため、多言語音声翻訳技術を用いた翻訳サービスが病院、ショッピングセンター、観光地、公共交通機関等の生活拠点に導入され、日本語を理解できない外国人でも日本国内で「言葉の壁」を感じることなく、我が国の生活に必要なサービスを利用できるようにすることが必要となる。

NICT が開発してきた多言語音声翻訳システムは、日英中韓の4か国語間の短い旅行会話の翻訳を比較的精度よく実現している。他方で、この多言語音声翻訳システムを実際の社会において実用レベルで使うためには、多言語音声翻訳システムの性能向上に加えて、実際の社会における実証実験を通じて得られた知見をフィードバックしながら、研究開発に取り組む必要がある。

このため、平成27年度5月に総務省委託研究の平成27年度情報通信技術の研究開発である「グローバルコミュニケーション計画の推進ー多言語音声翻訳技術の研究開発及び社会実証ーⅠ.多言語音声翻訳技術の研究開発」⁶への提案が公募され、パナソニック株式会

社、日本電信電話株式会社、パナソニック ソリューションテクノロジー株式会社、株式会社 KDDI 研究所、株式会社みらい翻訳、NICT が委託先として選定された。さらに、上記団体に東日本電信電話株式会社、株式会社日立製作所、パナソニック システムネットワークス株式会社、KDDI 株式会社、富士通株式会社、株式会社富士通研究所、エヌ・ティ・ティ・アドバンステクノロジー株式会社、株式会社 エヌ・ティ・ティ・データ経営研究所を加えた14団体で、「総務省委託研究開発・多言語音声翻訳技術推進コンソーシアム」⁷（以下「本コンソーシアム」）を2015年10月26日に設立した。本委託によって、実際の社会で利用可能であり、かつ、翻訳サービスを提供する企業が共通して利用可能な多言語音声翻訳プラットフォームが実現されることが期待される場所である。

6. 関連の動き

NICT の音声翻訳のスマートフォン用アプリ VoiceTra⁸が大幅に改善された。VoiceTra に対する関心は全国で高まっており、各地・各分野で利用実験が始まっている。例えば、沖縄観光協会、八王子観光協会、名古屋テレビ塔をはじめとする多言化対応に積極的な団体、京急、東京メトロ、りんかい線、上信鉄道をはじめとする鉄道事業会社、交番業務や警備などの多言

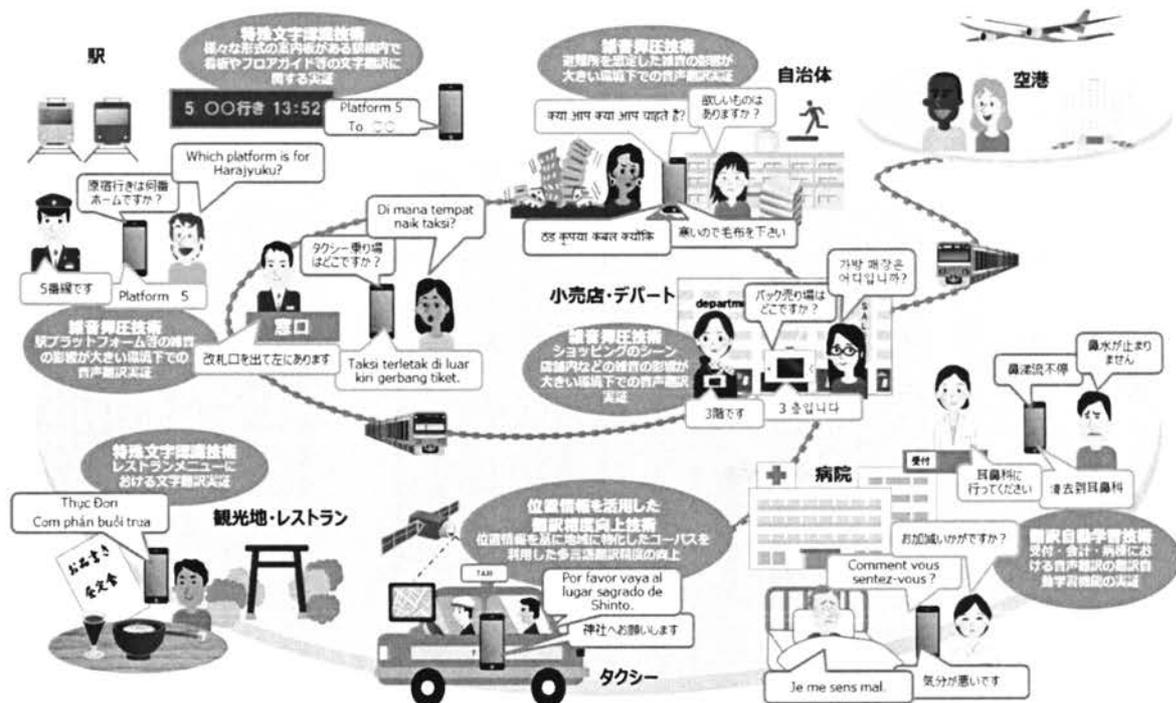


図 2 多言語音声翻訳技術イメージ

語化課題を抱える岡山県警をはじめとする警察、東京マラソンや東京国際ユースサッカー大会のような外国人が多数参加するスポーツ大会などがある。App StoreあるいはGoogle Playで“VoiceTra”を検索してダウンロードし無料で利用できるのも、VoiceTraを実際を使って、音声翻訳技術の現在の実力を知り、2020年のゴールを想像していただければ幸甚である。

音声翻訳ソフトウェアを搭載するハードウェアの研究開発が活発化しつつある。例えば、パナソニック株式会社が研究開発を進めている、日本語を多言語に翻訳して繰り返し拡声することができるメガホン型翻訳機（以下、「メガホンヤク®」という）がある。災害時などに外国人も多数いる混雑した現場で、放送型の多言語音声翻訳をすることを目的としており、成田空港で実証実験を行っている⁹。外国人対応をする多言語ロボットの研究開発も複数^{10,11}、同時に進んでいる。ヤマハは音声アナウンスに透かし技術で情報を載せ、多言語情報を、ユーザーのスマートフォンに提供する技術とNICTの音声翻訳の連携の研究開発を始めた¹²。音で情報を送るため、電波を要せず配信できる点が特徴になっている。

また、畑違いの株式会社NTTドコモが、自動翻訳技術の研究・開発・販売を目的とした株式会社みらい翻訳¹³が設立したことも新たなダイナミックな展開が始まった証拠の一つと言える。

7. まとめ

2020年のオリンピック・パラリンピック東京開催を契機として、NICTの技術をベースに次世代の音声翻訳エンジンを開発し広く社会実装することを目的とした国家プロジェクト「グローバルコミュニケーション計画」が始まった。NICTの内外の体制が強化・確立され、国費も措置された。民間独自の活動も活発に進み、「言葉の壁がない世界」の実現に向けて異次元の活

動へと昇華している。

⁹ <http://www.naa.jp/jp/press/pdf/20151112-megahonyaku2.pdf>

¹⁰ http://www.gizmodo.jp/2015/07/hello_3.html

¹¹ <http://www.travelvision.jp/news-jpn/detail.php?id=69223>

¹² <http://www.nict.go.jp/press/2015/07/13-1.html>

¹³ https://www.nttdocomo.co.jp/info/news_release/2014/09/29_01.html

Initiatives of the Japan Patent Office on Machine Translation

Kei Kato

Patent Information Policy Planning Office, Policy Planning and Coordination Department, Japan Patent Office

1. Introduction

The number of patent applications filed worldwide greatly increased to 2.57 million in 2013, compared with 1.58 million in 2004. Especially, the number of patent applications filed in China has dramatically increased and grew to 32.1% of the total applications in 2013 from 8.3% in 2004. From a viewpoint of searching patent information, this situation shows that the needs for searching foreign documents including those of the Chinese languages have increased. That is, as the examiners judge the novelty of inventions, inventive step, etc. based on the result of prior art search in the patent examination, it is necessary to search prior art documents including these foreign patent documents precisely as well as efficiently in order to grant a stable right recognized in the world. Actually, as the rate at which patent documents of the Chinese languages are cited as prior art in the examination has gradually accelerated at each Intellectual Property (IP) Office, it has become a challenge common to all IP Offices to improve an access environment such as understanding and searching foreign patent documents written in languages other than its native language.

On the other hand, it is not easy to understand and to exhaustively search foreign patent documents around the world. Needless to say, while translation enables us to understand them in our native languages, translating quite a large number of documents only by human has some limitations in terms of costs and resources, as the number of patent applications filed worldwide has increased year by year. Under this situation, the expectation of improvement of the environment for searching foreign documents by utilizing machine translation is getting higher.

It is hoped that machine translation will be utilized in order to disseminate patent information to other IP Offices. Currently, while intellectual property related activities are getting globalized, applications for a patent of one invention are more often filed in several countries or regions of the world. In case applications for a patent of the same invention are filed in several countries or regions, examiners in each IP Office need to share examination information (dossier information) owned by each IP Office in order to eliminate inefficiencies of conducting duplicate examinations or searches. But if each IP Office provides its examination information in its native language, it is difficult for other IP Offices to utilize such examination information. In order to promote the international work sharing, it is important to disseminate not only information in its native language but also one translated into other languages such as English.

In view of this situation where attention is getting focused on utilizing machine translation in the area of patent information, the Japan Patent Office (JPO) has taken measures to actually utilize machine translation. In this paper, I will introduce a search system for the Chinese and Korean documents translation and a service to provide examination information which the JPO has provided.

2. Improvement of the Environment for Searching Patent Documents of the Chinese and Korean languages

2.1 The Chinese and Korean Gazette Translation and Search System (CKGS)

As shown in Figure 1, in 1998 the documents to be understood and searched in the Japanese language accounted for 55% of the total patent documents issued in the world. But recently as the number of patent documents of the Chinese language has rapidly increased, documents to be understood only in the Chinese and Korean languages have accounted for 65% of the total. In January 2015, the JPO launched a new service of the “Chinese and Korean Gazette Translation and Search System” (CKGS) which enables users to search in the Japanese language patent and utility model documents of the Chinese and Korean languages in order to enhance the convenience of the environment for searching them.

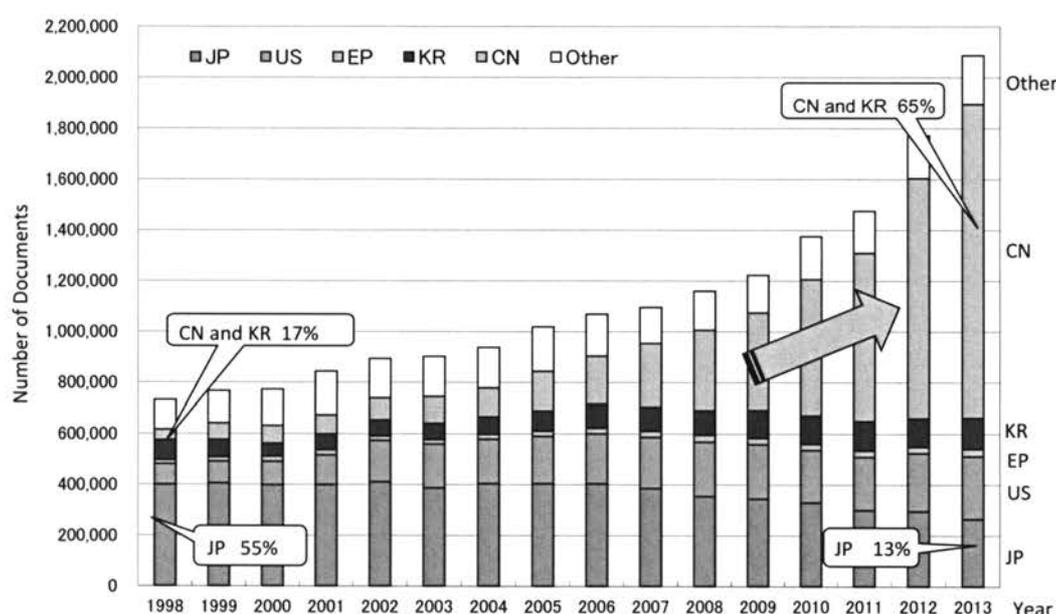


Figure 1. Rapid increase in Chinese patent documents. Note: Patent (incl. utility model) documents issued worldwide are categorized by language and duplicated data are eliminated. Regarding patent documents for the same invention which have been filed and published at multiple Offices, those documents published in Japanese are counted as JP. In case of no Japanese publications, such documents are counted first as US (English), secondly as EP (English, French, German), thirdly as KR (Korean) and last as CN (Chinese), if each language is applicable.

Through the Internet, the CKGS has become available not only to examiners but also to general users. A rule-based translation method is adopted for this system so that important technical terms can be translated and searched properly. The system stores patent and utility model documents machine translated from Chinese and Korean to Japanese, so that Japanese key words can be used for the full-text search of these documents. As of the end of March 2015, the system can be used to search about 12 million Chinese and Korean patent and utility model documents and it is planned to increase the number of such searchable documents gradually in the future.

The JPO has been conducting projects since Fiscal 2012 to develop a specialized dictionary for patent terms which will be used for the system’s machine translation [1-3], as its initiatives to improve machine translation quality from Chinese to Japanese. Specifically, corresponding Japanese and Chinese sentences are extracted from the patent family of a same

invention applied to Japan and China in order to prepare Chinese-Japanese translation corpuses as well as to develop a specialized dictionary. Currently, about 153 million Chinese-Japanese translation corpuses and the specialized user dictionary containing more than 2.2 million words have been developed. The user dictionary is installed to the CKGS to improve the system's translation quality.

Furthermore, measures are taken to enable the system to handle new technical terms as they appear. Specifically, based on information concerning unknown words which are detected as those not registered in the machine translation dictionary during the system's machine translation and that concerning incorrect translations and other errors reported by the system's users, efforts are made to update the dictionary, make additional registrations to the system's translation memory, and tune the translation engine parameters. With these efforts, the dictionary and the translation engine will be improved so that new technical terms can be handled as they appear.

As described above, continued efforts are being made to improve the system's machine translations so that practical application environment for uses to search Chinese and Korean documents can be provided, and the system's users appreciate it very much.

2.2 Quality Evaluation of Machine Translations from Chinese to Japanese and Those from Korean to Japanese

For realizing machine translations to be widely accepted and used, the quality of such translations plays a very important role. As a prerequisite to deciding whether to introduce such machine translations, the quality of such translations must be evaluated properly. A "Survey on How to Evaluate Quality of Patent Document Machine Translations" was conducted in FY 2013 to investigate how the quality evaluation method of machine translated patent documents should be [4]. Based on the survey's results, "Quality Evaluation Procedures for Patent Document Machine Translations" was developed in Fiscal 2014 as a guideline to evaluate the quality of machine translation results properly [5].

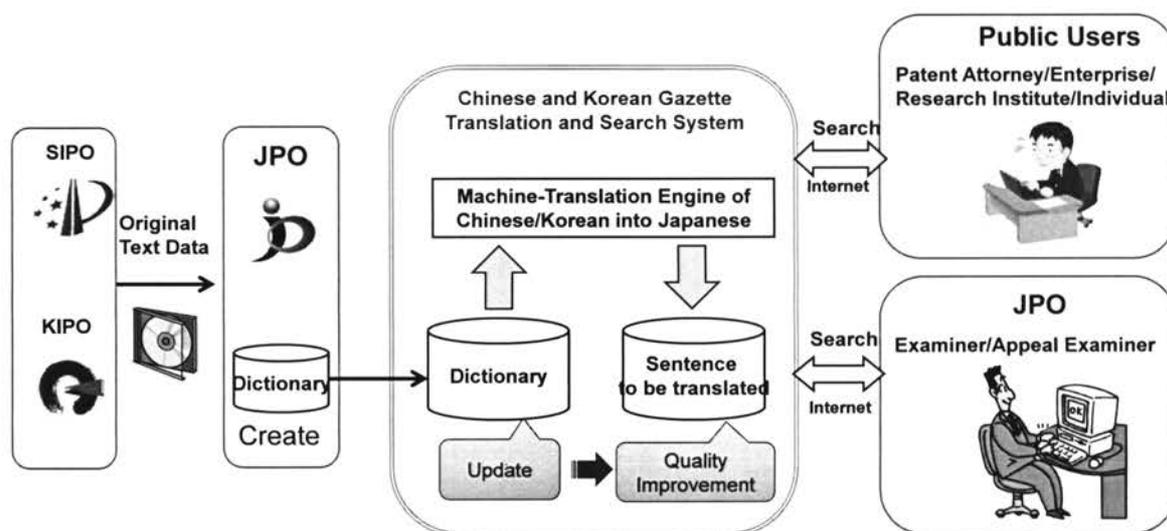


Figure 2. Conceptual diagram of Chinese and Korean Gazette Translation and Search System.

In order to verify the results of the JPO's efforts to improve the system's machine translation quality in accordance with the "Quality Evaluation Procedures for Patent Document Machine Translations," machine translations provided by the Chinese and Korean Gazette Translation and Search System were evaluated in the end of 2014 with regard to how accurately the system could translate technical terms (or the system's translation accuracy of technical terms) and how well the system could communicate original contents (or the system's original contents communication level).

With regard to the system's translation accuracy of technical terms, the system's machine translations of Chinese documents into Japanese were surveyed. A total of 196 technical terms were selected from various fields, and evaluated and classified into the following 4 grades by human evaluators.

- A (Properly Translated Word): Compared with one translated by a human, it is a word translated to a technically same or similar meaning and generally used.
- B (Acceptably Translated Word): It is not a translation word generally used as a technical word, but its meaning is almost correct.
- C (Mistranslated Word): It is a mistranslated word.
- D (Untranslated Word): It is a word not translated or whose translation was erroneously omitted.

With regard to the system's original contents communication level, the system's machine translations from Chinese and Korean to Japanese were surveyed. 100 machine translated sentences of each of the languages were selected from various fields, and evaluated and classified into the following 5 grades by human evaluators.

- 5: All contents of its important information are communicated correctly (100%).
- 4: Almost all contents of its important information are communicated correctly (80% or more).
- 3: No less than half of the contents of its important information are communicated correctly (50% or more).
- 2: Several contents of its important information are communicated correctly (20% or more).
- 1: Its translations cannot be understood, or almost no contents of its important information are communicated correctly (less than 20%).

In order to verify how effective it was to install a Chinese-Japanese dictionary developed by JPO to the system, the system's machine translations from Chinese to Japanese were evaluated before and after that dictionary with approximately 1 million words was installed to it for both surveys on the system's translation accuracy of technical terms and on the system's original contents communication level. In addition, for the purpose of comparison, the translated sentence using statistic machine translation (SMT) with 100 million Chinese-Japanese translation corpuses were also evaluated in the same manner.

Figure 3 shows the evaluation results of the system's original contents communication level through machine translations from Chinese and Korean to Japanese. While the system's average original contents communication level for Korean-Japanese translations is above 4 points, one for Chinese-Japanese translations is merely above 2 points. In addition, in all the technical fields, the system's original contents communication levels for Korean-Japanese translations are higher than those for Chinese-Japanese translations. It was concluded that this was due to the fact that the grammatical characteristics of Korean were closer to those of Japanese than those of Chinese, and the results indicate that continued efforts must be made to improve the system's translation accuracy of Chinese documents. Furthermore, the system's translation accuracy of either Chinese or Korean documents varies from one field to another, and measures must also be taken for individual fields.

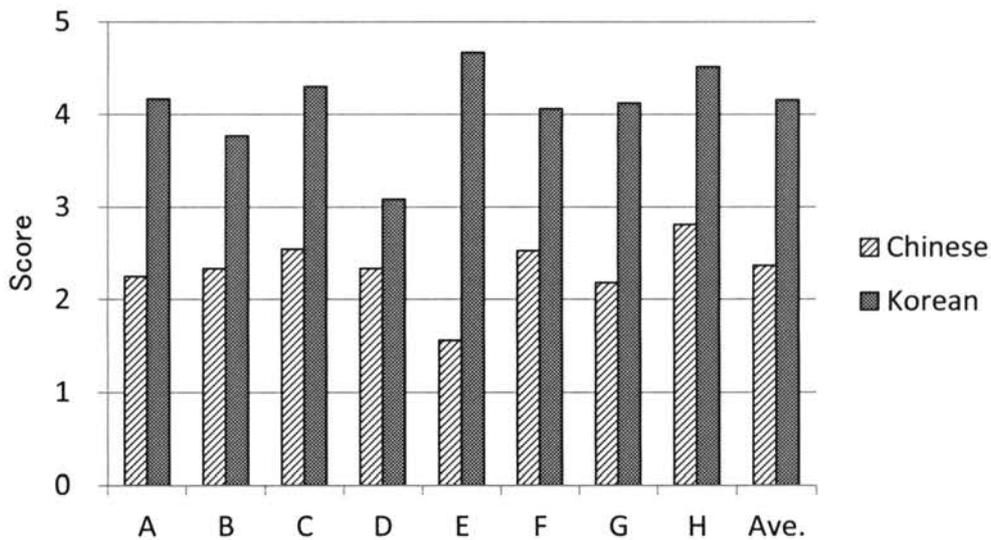


Figure 3. Original contents communication levels through machine translations from Chinese and Korean to Japanese in various fields. The figures show average evaluation points for respective fields averaging points of individual machine translations in the respective fields which are evaluated according to the 5 ranks of 1 to 5 points. "A" represents human necessities, "B" performing operations/transporting, "C" chemistry/metallurgy, "D" textiles/paper, "E" fixed constructions, "F" mechanical engineering/lighting/heating/weapons/blasting, "G" physics and "H" electricity.

Figures 4 and 5 show how accurately the system can translate Chinese documents' technical terms and how well the system can communicate their original contents, before and after the Chinese-Japanese dictionary developed by JPO was installed, respectively. That dictionary being added to the system's translation engine, the system's translation accuracy of technical terms increased by about 4%, and the system's original contents communication level by about 13%. Specifically, the translation accuracy of technical terms by the system's machine translation based on a rule-based method is higher than one by the SMT.

On the other hand, the original contents communication level by the SMT is evaluated to be higher. Specialized technical terms are frequently used as key words to search patent documents, and the translation accuracy of such technical words plays an important role for such searches. However, original contents communication level is also important to understand documents' contents, and thus efforts must be made to improve the system's translation quality.

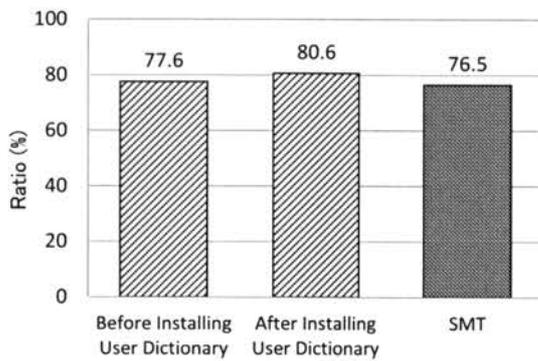


Figure 4. Evaluation of translation accuracy of technical terms in Chinese-Japanese machine translations: The above figures represent the sums of evaluation A (Properly Translated Word) and B (Acceptably Translated Word) ratios averaged over the 196 words.

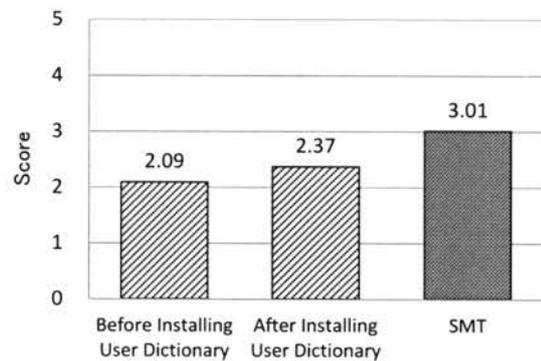


Figure 5. Evaluation of original contents communication Level: The above figures represent 5 level evaluations of 1 to 5 points on machine translations averaged over the 100 sentences.

3. Providing Information on Examination in English

3.1 Advanced Industrial Property Network (AIPN)/One Portal Dossier (OPD)

The number of yearly patent applications in the world increased 1.6 times in 10 years from 2004 to 2013, the 80%, which showed 2.08 million applications in 2013, were filed with five intellectual property offices in Japan, U.S., Europe, China and the Republic of Korea (or the IP5 Offices). Out of these, not a few applications for single inventions were filed in multiple countries and regions.

In order to facilitate examination work sharing among examiners in respective IP offices, the JPO has been providing foreign IP offices with its information on patent applications and examinations (dossier information) translated from Japanese to English by machine translation since October 2004 through a network called the Advanced Industrial Property Network (AIPN). The AIPN enables individual IP offices' examiners to cross check the individual IP offices' information on patent applications and examinations (dossier information) mutually so that if a same invention is filed to multiple countries and regions, duplicate searches and examinations on the same invention can be excluded. The AIPN enables Japan's dossier information to be understood in English, and individual IP offices' examination results to be shared among the individual IP offices.

Furthermore, an IT service called the "One Portal Dossier" (OPD) was started in 2013 for the IP5 offices' examiners, which collects the IP5 offices' information on examinations together, including the JPO's information on examinations machine translated by the AIPN, and provides such information in an easy-to-see format. Currently, this OPD's service is enhanced, and the JPO's information on examinations is machine translated and provided to general users together with other IP5 offices' dossier information. In 2014 alone, through this service, the JPO's examiners browsed about 160,000 cases in the dossier information of the EPO, the USTPO, the SIPO and the KIPO, and the examiners of the EPO, the USTPO, the SIPO and the KIPO browsed about 210,000 cases in the dossier information of the JPO. Such dossier information helps to make examinations to be more efficient.

The JPO's information on examinations, which is provided through the AIPN/OPD, contains texts machine translated by a translation engine based on a rule based method. In general, machine translations from Japanese to English tend to show low translation accuracy when compared with those from English to Japanese. Notwithstanding, the JPO is constantly enhancing its machine translation dictionary to improve its service's translation accuracy. Specifically, the JPO collects untranslated words (unknown words) and registers their translations to its dictionary, as well as analyze feedbacks about mistranslations in the AIPN from foreign IP offices and incorporate their corrections into its dictionary. Thereby, the JPO is making efforts to enhance its dictionary.

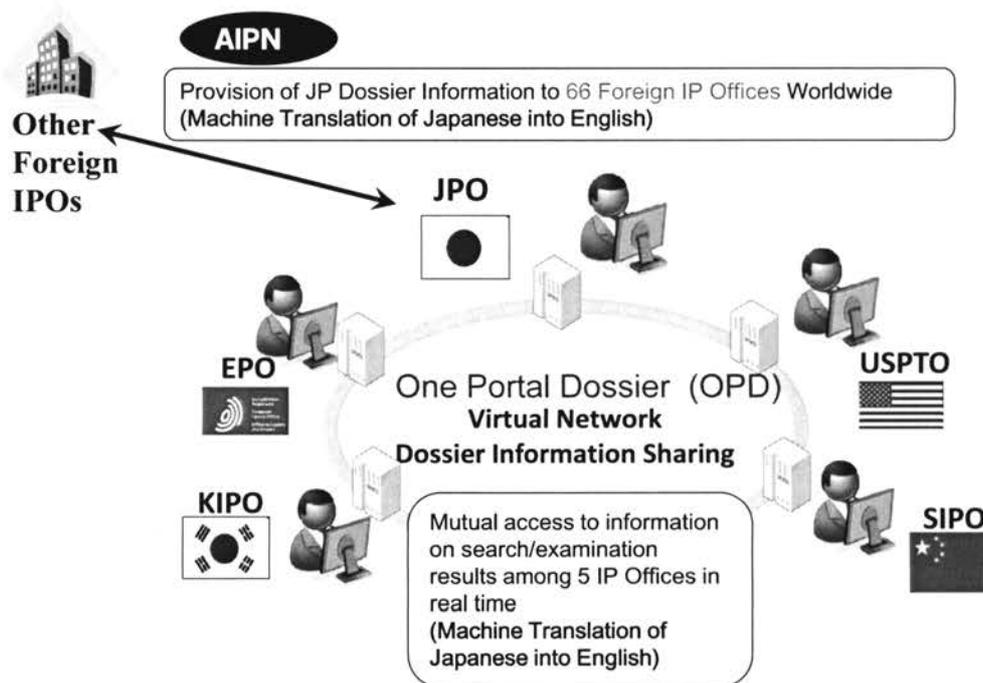


Figure 6. Conceptual diagram of AIPN/OPD.

3.2 Quality Evaluation of Japanese-English Machine Translations

In order to verify the translation accuracy of machine translations from Japanese to English which the JPO provides, in 2011 the JPO conducted a survey on translation accuracy evaluation for providing English texts of machine translations of information related to patent examinations [6]. This survey evaluated 511 sentences extracted from patent gazettes and reasons for refusal in the fields of electricity, physics and chemistry. Two types of translation methods, that is, the JPO's machine translation system (AIPN) and commercially available Japanese-English translation software for patent, were applied to the sentences to obtain their machine translations, and the 2 types of machine translation methods were evaluated automatically as well as by human evaluators. The human evaluators evaluated the individual machine translations from the view point of whether the machine translations correctly reflect their originals' meanings, and assigned 0 to 4 points to each of them according to a 5 grade evaluation scheme. Three kinds of evaluation methods, that is, BLEU, NIS and IMPACT were used for the automatic evaluation.

Figure 7 shows the results of the human evaluators' evaluation. While the AIPN's average point is 2.5 points, the commercial product's is 2.0 points. In addition, the results indicate that the AIPN's points are higher than the commercial

product's in all the technical fields. Results obtained by the automatic evaluation methods (BLEU) show trends almost similar to one obtained by the human evaluator's evaluation (Figure 8). Because of the results mentioned above, it can be concluded that the JPO's efforts to improve the accuracy of its Japanese-English machine translations have made some progress. However, it is noticed that the translation accuracy varies from one technical field to another. In addition, different evaluation methods produced different results, and a method to evaluate translation accuracy must be reviewed and examined continually. The JPO is still continuing to make its efforts to improve the quality of its Japanese-English machine translations even after the 2011 survey, and also plans to conduct another survey on the quality evaluation of its Japanese-English machine translations including the analysis of evaluation methods in FY 2015 in order to verify the results of this effort.

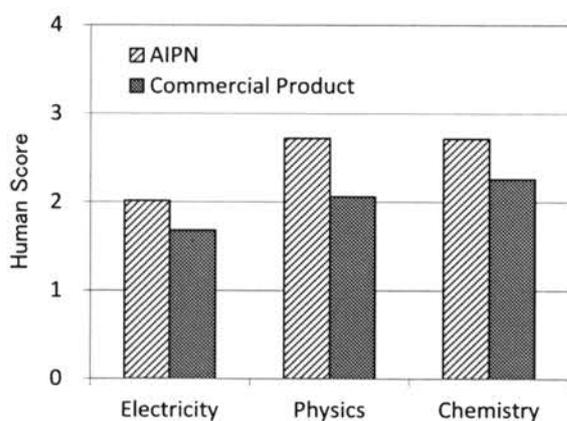


Figure 7. Evaluation of Japanese-English machine translations' accuracy by human evaluators.

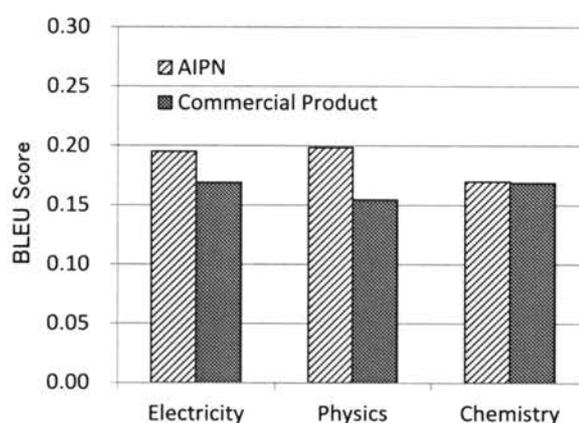


Figure 8. Evaluation of Japanese-English machine Translations' accuracy by BLEU.

3.3 Launch of External Provision of English Translations of Appeal/Trial Decisions

In order to enhance and strengthen global dissemination of information on appeals/trials, the JPO started to provide English translations of appeal/trial decisions etc. (appeal/trial decisions, decisions on oppositions and Hantei-advisory opinions on the technical scope of intellectual property rights) that help in the understanding of the legal interpretation and its operation by human translation on the JPO website (in Japanese and English) from January 2016. On the other hand, it is difficult to translate all of appeal/trial decisions only by human translation in terms of costs etc. for the dissemination. Thus, in order to provide more English translations of appeal/trial decisions to users globally in addition to those by human translation, the JPO began providing Japanese-English machine translations of published appeal/trial decisions on the Japan Platform for Patent Information (J-PlatPat) (in English version) from January 2016 (Figure 9).

As written appeal/trial decisions include unfamiliar words, that is, technical terms and phrases specific to appeal/trial decisions as well as a large number of long sentences, translating sentences in appeal/trial decisions by machine translation is very challenging. However, the JPO is promoting registrations of terms and phrases specific to appeal/trial decisions in the system's translation dictionary. For example, it is confirmed that English translations of decisions on appeals against examiner's decision of refusal reach a satisfactory level to be understood in outline in its top part where the conclusion is stated, its latter part where how appeal examiners judge the case is stated, etc. In the future, the JPO will make continued

efforts to improve the quality of Japanese-English machine translations of appeal/trial decisions through increasing registrations of technical terms and fixed phrases specific to appeal/trial decisions in the system's translation dictionary.



Figure 9. Display screen of English translations of appeal/trial decisions on J-PlatPat.

4. Conclusion

Demands for machine translations of patent information are increasing year by year from the view point of patent information searches and dissemination. While machine translation is a developing technology and is being improved constantly, its accuracy has not become satisfactory yet, and one of its important issues is how to improve its accuracy. The JPO is evaluating the quality of Chinese-Japanese and English-Japanese machine translations in Fiscal 2015, and henceforth it may be necessary in the future, for example, to consider a new machine translation system including a machine translation engine based on quality evaluation results to improve machine translation quality. In addition, to improve translation quality, methods and criteria to evaluate individual machine translation schemes' qualities are also important because they are preconditions to evaluating translation quality. With regard to this, it is a common issue among the IP5 offices how to improve machine translation quality, and the offices plan to have further discussions on making methods and criteria for machine translation quality evaluations common among the IP5 offices.

References

1. JPO, Survey on Chinese-Japanese Dictionary Development and Machine Translation Performance Improvement for Machine Translations of Chinese Patent Documents, 2013.
2. JPO, Survey on New Words for Machine Translations of Chinese Patent Documents, 2014.

3. JPO, Survey on Dictionary Development and Machine Translation Quality Evaluation for Machine Translations of Chinese Patent Documents, 2015.
4. JPO, Survey on How to Evaluate Quality of Patent Document Machine Translations, 2014.
5. JPO, Quality Evaluation Procedures for Patent Document Machine Translations, 2014.
6. JPO, Survey Report on Translation Accuracy Evaluation for Providing English Texts of Machine Translations of Information Related to Patent Examinations, 2011.

映像翻訳の授業活用例：

字幕制作を通じたメディア理解

魚住真司

関西外国語大学 外国語学部

1. はじめに

以下の画像(1)は、筆者がNHK報道カメラマンだった当時に取材した米国人カメラマンのインタビュー映像である。



画像(1) NHK「イブニングネットワークきんき」
(1990)より

筆者自身が企画したレポートだったので、翻訳作業も自分で行うことになった。ところが、慣れない作業でできあがった日本語字幕は3行にわたり、主人公の口元まで字幕が覆いかぶさってしまった。テレビ放送では、一度に見せることのできる日本語字幕は、どんなに込み入った話でも意識して2行に収めなければならない。そんな映像翻訳のルールを知るようになるのは後年、教育者に転じてからの話である。

2. 映像翻訳を授業に

映像翻訳には「字幕翻訳」と「吹き替え翻訳」の2

種類ある。どちらも映画なりテレビ番組なりの英語音声で、前者は日本語字幕を制作して映像に合成させ、後者は日本語音声に吹き替える(子供向け作品の場合は後者が多い)。本稿では筆者の授業実践例として、前者の「字幕翻訳」を「映像翻訳」とほぼ同義で話を進めていくことにする。

映像翻訳には様々なルールがあるが、その基本原則をあげておこならば、①映像1秒につき字幕は4文字まで(それより多いと大人の目でも追いきれない)、②1行あたり13~14文字まで、多くても一度に見せる文字列は2行(14文字×2行=28文字)にとどめる、③句読点は不要、かわりに半スペースを使う。

たとえば、画像(1)には計32文字ある。逆算すると、このインタビュー映像は、長さが8秒(32文字=4文字×8秒)以上ないと、視聴者は字幕を最後まで読み切ることができない、ということになる。

さて、筆者が勤務する関西外国語大学で担当した授業「メディア・スタディーズII」(メディア学の上級クラス)は、当初、英語能力の向上を直接の目的としていたわけではなかった。その授業は社会学的見地に立って、メディアが抱える現代的な諸問題を理解しようというものであった。その際、外国語大学だからという理由で、英語教材の使用を思い立ったわけである。

教科書は米国の大学で実際に使用されている「マスメディア」(Opposing View Pointsシリーズ, MASS MEDIA, Greenhaven, 2009.)。これと併用させて、米国のテレビ局が制作した優れたドキュメンタリー番組を視聴させることが、学生のメディアへの興味関心につながると筆者は確信していた。しかし問題は、それ

を「どのような方法で『視聴』させる」かであった。

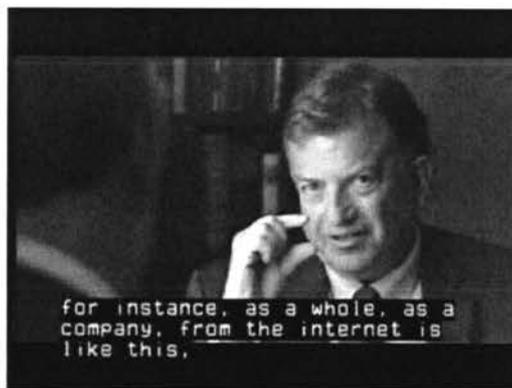
そこで、筆者自身が過去に苦心したことのある「映像翻訳」を、受講生にも実践させてみてはどうだろうか—英語音声のテレビ番組を、日本語字幕の制作を通して「噛みしめる」ことで、メディアに対する深い理解が得られるに違いない、と考えた。また、受講生の中からプロの映像翻訳家を目指す学生が出てくることを願って、実際に放送現場でも使用されている映像翻訳補助ソフトウェア（カンパス社の SST）の導入を大学に提言し協力を得ることができた。

ところで、このような IT 環境での英語授業を早くから実現させている藤野（2005）によると、映像翻訳を授業に採り入れるねらいは次の 3 つであるという。

- ①豊かな表現力の育成
- ②自主的、意欲的学習態度の育成
- ③AV 機器や PC の積極的活用。

外国語大生はいえ、将来ほとんどの卒業生が日本社会で活躍することになるわけであるから、筆者には藤野の指摘した①が特に重要であると思われた。藤野が述べるように、映像翻訳の授業では「数多くの翻訳作業をこなすことによって、生徒が多彩な表現を学び取り、『固い』専門用語を翻訳することから、感情や状況を『柔軟に』表現できる能力の伸長を目指す」のに自然とつながるからだ。

「感情や状況を柔軟に表現」—これに関して、画像 (2) を見てほしい。米公共放送 PBS が 2007 年に放送したドキュメンタリー番組 "News War" の 1 シーンである。実際に筆者の授業で使用しているこのビデオクリップの、画面下に表示されている英文はクローズド・キャプションの文字を表示させたものである。クローズド・キャプションとは、1970 年代に聴覚障がい者向けに米国で開発され、80 年代に広く普及し、90 年代には 13 インチ以上のテレビ受信機に搭載が義務化された字幕表示システムである。字幕は番組音声を正確に文字化したものであり、番組の進行に沿ってどんどん更新されてゆく仕組みで、今では英語が母語でない者の英語習得にも活用されている。



画像 (2) PBS 「News War」(2007) より

さて、上のシーンでは、男性が右手指で「ちょっとだけ」といったような仕草を見せながら、次のようにインタビュー相手に語りかけている。

Our reporting shows that the revenue coming to CBS, for instance, as a whole, as a company, from the internet is like this, compared to the revenue that's coming from normal advertising. 直訳するならば、次のようになるだろう。

「我々のレポートが示すところでは 会社全体としての CBS へのインターネットからの収入はこれくらい 通常の広告収入と比較すれば」

このように、キャプションで示された英文に忠実な、機械的な翻訳だと、「like this」は「これくらい」にならざるをえない。しかし、それでは字幕を目で追いかけている視聴者にとって、親切な訳とは言えない。ここは、たとえば「ほんの少し」といったように、重要な情報が一目で伝わるような「意識」の方が良い。

同様に、「\$ 1300 万」と字幕が出て、日本の視聴者はそれが一体どれくらいの金額なのか一瞬ではわからない。為替レートの問題はあるにせよ、やはり「13 億円」と、おおよそでも良いから換算した数値を字幕で見せるべきであろう。

つまり映像翻訳とは、限られた時間内にできるだけ少ない文字数で、視聴者の番組理解を促進させる作業なのである。そのためには英語読解力はもちろんだが、むしろ日本語表現力が試されると言っても良いだろう。

3. その実際

「メディア・スタディーズ II」は2010年度から半期週2回の授業として開講した。毎週1回を教科書の日、別の1回を映像翻訳に専念させることにした。当初は映像翻訳補助ソフト SST を受講生に使用させることでプロ並みの字幕制作環境が整ったが、使用契約が切れた2014年度からはMicrosoft Wordで翻訳作業を実施させた。ただし、SSTで可能だった字幕合成まではWord単体ではできないので、字幕合成まで完遂させるにはMovie Makerなどの映像編集ソフトを併用することになる。

さて、授業で使用したドキュメンタリー番組は、米国の公共放送ネットワークPBSの"Frontline"シリーズで2008年に放送された"News War" (60分間)である。テレビニュースの歴史と、メディア全般に共通する現代的な問題について深く掘り下げた、各方面からの評価が極めて高い秀逸な報道番組である。DVDが発売されており、授業での利用を促進すべくTeachers' Guideファイルも収録されており、教育者にとって大変に便利なものである。

教員の授業準備としてはまず、①この60分の中から1分~1分30秒程度のビデオクリップを、授業回数分選定することである。通常、授業初期は受講生たちが翻訳に手間取り、思うように授業が進まないのが、授業回数分とは言っても全部で10ファイル分程度で間に合う。

そして、②その授業回数分の映像ファイルをwmv(ウィンドウズ・メディア)などの「身軽な」ファイル形式に変換して、あらかじめ受講生人数分のUSBメモリーに入れておく。さらに、③各回の映像ファイルの、英文スクリプト(英語台本)を用意しておく。

筆者はクローズド・キャプションの文字を画面に表示したままの映像をダビングして映像ファイル化したので、キャプションの英文字はパソコンでのファイル再生時にもテレビと同じように表示される。映像翻訳のプロならば、これだけでも字幕制作は可能であろう

が、やはり授業となると紙に印刷された英文スクリプトが欲しいところである。受講生たちはその紙に、知らない英単語の意味を調べて書き込んだり、難解な表現に対して筆者が与えるヒントをメモするのに役立っている。

なお、英文スクリプトの作成にはフューテック社のFA-400デコーダを利用した。この機種は、クローズド・キャプションの文字データを、テキストファイルとしてパソコンに取り出せる。このような機器を利用することで、授業準備にかかる時間を節約した。

授業は、まずその回で翻訳に取り組む映像ファイルをクラス全員でざっと視聴することから始める。そして、難解な表現やマスコミ業界の専門用語ならびに時代背景などについて、筆者がホワイトボードを使って説明する。その上で、各受講生は自分のUSBメモリーの中に収められた映像ファイルをそれぞれのペースで再視聴しつつ、また辞書なども引きながら字幕を考えていくのである。



画像(3) 字幕を考える受講生たち

およそ20~30分ほど経過したところで、受講生一人ずつに、おおよそどのようなことが述べられているのか大意を発表させる。大きな流れをつかんだことを確認した上で、さらに20~30分ほど時間をかけて、字幕を完成させる。

完成した字幕は印刷させて、それをクラスで鑑賞したり、教員が講評を加えたりする。最後に、教員があらかじめ制作しておいた模範字幕の合成された映像フ

ファイルを皆で視聴する。ただしこの最後の作業は、特に授業回数の浅い段階においては、授業時間（90分）が足りなくなつて、次回の授業冒頭にまわさざるを得なかったのが実情である。したがって授業初期は、なるべく短い映像ファイル（1分未満）でのぞんだ方が良好だろう。ネイティブが放送で話す英語は、日本の学生が予想する以上にスピードが速く、また英語を母語としない者にとっては「不思議な」言い回しが散見されるのであるから。

例えば、実際に授業で使用した英文スクリプトの中に、次のような表現があり、受講生たちは困惑した。

People that are consuming news or other information online do so in... in a mode that we refer to as "info snacking."

「情報をお菓子化する」とは一体どういうことだろうか。その趣旨はおおよそ次のようであろう。

「ネットで情報を得る人は『情報のつまみ食い』をしているんだ」

また、以下のような英文を日本語字幕で短く表現するにはどうしたら良いだろう。

They're going to be online reporters, bloggers, man on the street, that sort of thing, and they're now part of the whole news calculation around the world.

ある受講生は以下のように字幕を付けた。

「ブロガー その辺の人 そんな人々こそが 今や世界中のニュースに欠かせない存在なのだ」

悪くはない。筆者は次のように模範を示した。

「ネット記者 ブロガー 普通の人・・・世界中で 今やこういった人々がニュースには欠かせないのだ」

いかがだろうか。かなり短く、思い切った訳となっているが、「1秒=4文字」ルールを厳守するためには、これくらい大胆な意識が必要となる。つまり、翻訳する際に心がけることは、原文の趣旨を外さない範囲で、持てる日本語表現力を駆使して、なんとか短くなるよう工夫するのである。

ところで著作権については授業開始時に、受講生に対して次のように説明してある。すなわち、① "News War" の著作権は米公共放送 PBS（厳密には "Frontline" シリーズを制作しているボストンの PBS 系列局 WGBH）に帰属する。ただし、②授業で番組映像を利用するのは日本の著作権法（第35条）でも米国の著作権法（フェアユース規定）でも認められている。しかしながら、③受講生が番組映像を教室外に持ち出すことは認められない。

従って、映像ファイルの入った USB メモリーは持ち帰りを禁じており、毎回の授業終了時に返却させている（つまり、これら USB メモリーは教員側が用意してやることになる）。それゆえ映像翻訳の予復習はできないが、教科書の方で宿題を課している。

4. 受講生たちからのフィードバック

この授業の受講生たちは、映像翻訳をどのように思っているのだろうか。直近（2015年度秋学期）の授業は、15名の登録者の内、1名が登録削除、1名が途中脱落となった。最後まで出席した13名の内、本人の承諾を得たものの中から教員へのフィードバックを、以下に紹介してみる。

（3年生・女子）

「・・・事前に配布された英文を時間をかけて訳したものを発表するのではなく、一度映像をみて、少ない時間で、訳していくことがよかった。映像を見てすぐ訳すので、普通の文字だけの記事を訳すより楽しかった。また実際にその言葉を喋っている人の表情をみて訳せるので、文字だけではわからない感情も感じながら訳すことができたと思う。これが宿題や事前に予習できるものだったら、だらけてしまう気がする。見たものをすぐ訳すという新鮮さが良いんだと思った・・・」

まさに映像翻訳を授業内で実践した意義が、ストレートにフィードバックされたと言えよう。興味深いのは、「短期決戦」を肯定的にとらえている部分である。

テレビ放送という、時間に厳密なメディア (のビデオ) を教材に使用したことで、授業に緊迫感が漂い、通常の英語授業とは異質の体験をさせることができたようである。

しかし一方で、次のような感想も寄せられた。

(4年生・女子)

「・・・翻訳しなければいけない文章が長く、その分時間も長くなってしまい、先生が授業中に話している内容を聞くのがおそろかになってしまうのが残念です。先生の話をもっとしっかり聞きたいのに、翻訳もしなくてはいけないので、どちらもしっかり両立できるように、工夫した時間配分をしていただきたいです。」

これは受講生たちが必死になって翻訳している最中に、教員が思いついたように番組の内容について解説を始めたり、政治的・経済的背景に話がおよんだりするものであるから、翻訳に集中できなくなってしまったのであろう。各受講生の英語能力に差異があることも関係するのであろうが、このようなことが起こらないように、この種の授業では、教員は話をする時と黙って翻訳に集中させる時間とを、意識的に区切ることが大事であることが良くわかった。

さらに次のような感想もある。

(4年生女子)

「メディアについて学べると同時に英語力の維持につながったと感じた・・・短時間で翻訳する難しさを知りました。また、そのまま訳すのと翻訳することは全く違うこと、相手に伝わりやすく訳することが課題でした。専門用語が多すぎて、時々ついていけなくなり授業に必死でした。」

本授業は、メディア学と、(結果として) 英語授業の2つを同時に受講するようなものなので、もちろん容易ではないだろう。やはり、メディア界特有の英語表現や専門用語については、より詳細な事前解説が必要であった。あるいは、英語授業に特化させる場合は、教材にもっとやさしい内容の映像を選択すべきと思われる。

(4年男子)

「特に映像を見なくても、訳せてしまうところが気になりました。」

これは、映像ファイルとはまた別に、英文スクリプトを渡しているのだから、これを見れば訳せるのではないか、ということである。確かに、SSTの契約期間が終わり、字幕制作をWordに切り替えてからは、あえて「映像を見ながら」「番組出演者の表情をチェックしながら」、という「手間」をかけない受講生がいたことも事実である。SSTでは、まさに放送現場で行われているような、映像に直接字幕を「打ち込んでいく」感覚の新鮮な作業となったが、Wordはそのおもしろさが伴わない。また、映像メディアそのものに対する興味や関心度は、受講生によってまちまちなので、仕方の無いところかもしれない。

(3年男子)

「・・・授業中に先生が他の学生をしている事、言った事から自分で『やってみる』事がリアルタイムでできる映像翻訳の授業は大学生に合っていると感じた。」

自分だったらどんな字幕を付けてこの番組を視聴者に見てもらうか、擬似的ではあっても一種の「ライブ感」を持ちながら授業参加できるのが映像翻訳である。

5. おわりに

米国におけるメディア状況を理解させるための手段として採用した映像翻訳が、授業回数を重ねるうちに、映像翻訳自体が英語授業の目的に値することを、複数の文献と自らの実体験から確認することができた。

ところで、日本でいち早く映像翻訳を英語教育で実践してきた染谷(2009)によれば、「字幕翻訳」授業の目的は「原文の意図を正確に理解し、これを簡潔かつ確かな日本語で再表現する訓練を通じて、教養ある社会人に求められる高度な言語運用力(英語・日本語)を養成する」ことにあるという。

日本人学生が将来、英語と日本語の橋渡し役を果た

すためには、結局は母語である日本語表現力が問われる。その鍛錬がおろそかだと日本社会の要請に十分応えることはできない。ネイティブ・スピーカーによる英語によるディスカッション中心の授業を否定するわけではもちろんないが、翻訳・反訳（＝翻訳した日本語から再び原文の英語に戻す）の英語授業が、いま、再評価されていることに共感を覚えるのである。

6. 参考文献

太田直子『字幕屋は銀幕の片隅で日本語が変だと叫ぶ』
光文社, 2007.

染谷泰正「大学における『字幕翻訳』授業のための指導モデル」『日本通訳翻訳学会第10回大会』研究発表スライド, 2009.

http://someya-net.com/84-Subtitle_TeachingModel/Jimaku_Handout_TsuyakuGakkai09.pdf (last visited: Jan. 30, 2016)

染谷泰正『字幕付き動画の作成とウェブページへの埋め込みに関するテクニカルノート』未刊行論文集, 2008-9.

http://kakonacl.xsrv.jp/subtitle/allinone_small.pdf
(last visited: Jan. 30, 2016)

日本映像翻訳アカデミー『はじめての映像翻訳』アルク, 2011.

藤野智子「オーラルコミュニケーションにおける『映像翻訳』の試み—字幕制作ソフトとアフレコ用機材を利用して」『東京学芸大学附属高等学校大泉校舎研究紀要』30: 59-74, 2005.

紅粉芳恵「CaptionTubeを使った字幕作成と授業デザイン: インタビュー動画を利用して」『外国語教育フォーラム』14: 61-75, 2015.

付記・本稿に登場する受講生たちのコメントや画像は、本人の承諾を得た上で使用しています。

CJNLP2015（日中自然言語処理共同研究促進会議）参加報告

秋桜舎・山本ゆうじ

機械翻訳課題調査委員会ワーキンググループ3（UTX チーム）

1. はじめに

本稿では、第15回CJNLP2015にAAMT WG3(UTX チーム) リーダー山本ゆうじ(筆者)が参加し発表を行った結果を報告する。

2015年10月18日から19日にかけて、CJNLP2015 (China-Japan Natural Language Joint Research Promotion Conference、日中自然言語処理共同研究促進会議)が、青森市で開催された(写真1)。本会議は、自然言語処理技術での研究協力や共同研究の促進を目的に日本と中国で交互に開催されているもので、今回は15回目である。今回は、AAMT から中岩会長と筆者が参加した。筆者は、機械翻訳課題調査委員会のメンバーとしての参加である。

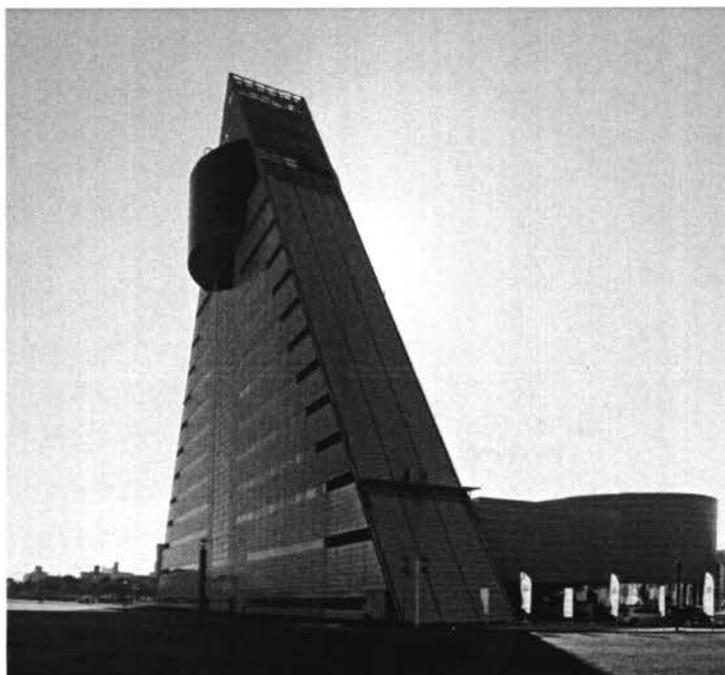
会議参加者総数は30数名で、多くは中国からの自然言語処理研究者であった。会議での使用言語はすべて英語である。

2. 発表と会議

筆者は、聴衆の多くが自然言語処理研究者であることを踏まえた上で、UTX 用語集形式 <<http://www.aamt.info/japanese/utx/>>の意義を説明する発表を行った。発表の他に、中国での機械翻訳の現状について情報収集も行った。筆者の関心は、主に機械翻訳の実用面での活用であるため、以下で取り上げる内容には偏りがあることはご了承ください。

筆者は、会議初日の午前中のセッションで発表した。筆者の発表の要点は、機械翻訳の品質を改善するには、コーパス(言語データ)の元になる人間翻訳の品質を改善する必要があり、そのためにはよい用語データが必要で、またそのためにUTXが利用できる、ということである。英語では、Chicago Manual、APA、MLAなどの表記基準が、学会や出版物で広く使われており、句読点や引用方法が定められている。だが日本語では標準的な表記基準がなく、また文字

写真1 会議会場(初日)・青森県観光物産館アスパ



種、送りがななどで多様な表記が可能である。このため、用語データなどを参照しないと、メガネ、めがね、眼鏡などの多様な表記が脈絡なく混在することがある。表記基準がなく、多様な表記や句読点が混在していると、言語資源としての品質が落ちることもある。表記基準に沿って表記統一され、また用語集で用語統一された翻訳文から作成された文章のコーパスは、用語と表記の面で一貫性がある。このようなコーパスを使えば、統計翻訳の精度はある程度上げられるものとする。参加者が研究対象としている文書で表記基準が使われているか質問したが、表記基準という観点には意識されていないようであった。

なお筆者の発表後に、参加者から「だれが用語ステータスの基準を決めるのか」という質問が一件あった。「用語集の管理者が決定する」と回答した。

会議全体の傾向として、機械翻訳の精度に関する議論が多く、機械翻訳を人間がどう使用するかについての議論は比較的少なかったように思われた。だが、中国科学院の Chengqing Zong 氏による、機械翻訳、翻訳メモリー、用語集、言語入力システムを統合したアプリケーション（非公開）を紹介する発表もあった。その他の発表では、中英翻訳や英文解析の研究が多く、日本語に関連する研究がやや少ないと感じた。

上記以外には、（中国にある）東北大学 Xiao Tong 氏は、氏らが開発した統計機械翻訳システム NiuTrans¹について発表されたが、構成がよく、英語も極めて流暢であったので、分かりやすかった。著者は、機械翻訳システムの精度を向上するために用語データを利用することを検討しているかと質問したところ、検討しているとのことであった（この質問は、UTX のような用語データが統計機械翻訳システムで活用されるかということ念頭に置いていた）。また、NiuTrans は、翻訳業界で広く使われてい

¹ Xiao 氏のチームが開発したオープンソースの統計翻訳システム NiuTrans
<<http://www.nlplab.com/NiuPlan/NiuTrans.html>>

る翻訳メモリー ツール SDL Trados との連携をしているとのことなので、どのようなかものか質問した。SDL Trados 専用のプラグインとして提供されるものではなく、ホットキー、そしておそらくクリップボード経由で情報をやり取りするシンプルな連携のようであった。そうではあっても、筆者の経験では、日本の機械翻訳関連の会議で翻訳メモリーやポストエディットに言及することはまれなので、興味深かった。

初日の夜には交流の宴席が開かれたが、そこでもいろいろと興味深い話が出た。発表での質疑の時間は限られているので、参加者一人一人から細かい話や本音を聞けるのはこのような場である。たとえば、中国では言語資源公開やオープンソースのような考え方はまだ少ないとの意見があった。日本でもそうであるのだから、これは意外ではないが残念なことである。

2 日目には、専門家を発見するアルゴリズム、ウェブ検索を改善する方法などについて発表があった。中岩 AAMT 会長は、日本語でのゼロ代名詞の自動検出について発表された。ポスター セッションも開かれ、用語データに関係する発表もいくつかあった。ポスター セッションのほうが、口頭発表よりも説明をじっくり聞くことができ、深く理解できた。

また、中国からの参加者が Google について言及していたのが興味深かった。Google は自然言語処理では無視できない存在だが、中国では基本的に Google のサービスは使用できないはずだからである。

3. まとめ

今回の会議は、自然言語処理研究のさまざまな視点を知ることができた非常に貴重な経験であった。UTX が実現する用語管理は、実用翻訳の観点からのものである。特定の技術や手法ではなく、広い視点からの総合的な提案である。本会議では、実務翻訳とは異なる研究者の視点を垣間見ることができた。また、体系的翻訳の改善に関するいくつかのアイデ

アが得られた。たとえば、静岡大学の綱川隆司先生が Wikipedia 記事を元データとして活用するポスター発表をされていたが、Wikipedia 記事の品質自動判定は、文書・翻訳品質の自動判定に発展できるかもしれない。さらに、中国のさまざまな年代の研究者の熱意に触発される面もあった。

本会議では、研究者と私のような実務者の意識の違いも確認できた。UTX 用語集形式は、のようなタグを使う XML 形式ではなく、原語や訳語の情報をタブ文字で区切ったシンプルなタブ区切り形式である。UTX 形式がタブ区切りとなったのは、実用面からこの形式で必要十分だからである。しかし、今回、研究者と話して、研究者にとってもシンプルなタブ区切り形式の利点があることを確認した。実務者は、用語データを不特定多数と頻繁にやり取りするが、相手の技能や知識はさまざまである。たとえば、翻訳会社が翻訳者に使用する用語を指定する場合、また翻訳者が使用した用語を翻訳会社に知らせる場合である。また実務者には、データを調整したり加工したりする時間的余裕はない。用語データは、なるべくそのまま用語ツールでスムーズに使えることが重要である。これらの理由により、要素が明確化され標準化された XML で用語データをやり取りしたほうが確実に便利なおことがある。しかし、研究者の立場は異なる。研究者は、自分でデータを加工できる知識があるので、決められた形として XML を使う必要がなく、むしろシンプルなタブ区切り形式のほうが扱いやすい。ただ、UTX のように構造が標準化されていることで、研究で収集したデータの再利用や共有がしやすくなるのは大きな利点と言える。今後、研究者に対しても「標準化されたタブ区切り」である UTX の利点をアピールしてゆきたい。

JTF 翻訳祭「MT Live ～機械翻訳の担うべき役割～」

熊野明

東芝ソリューション(株)

1. JTF 翻訳祭の新企画

JTF 翻訳祭は、日本翻訳連盟(JTF)の年次イベントで、毎年興味深い企画を行っている。2015年のJTF翻訳祭は日本翻訳連盟の25周年企画として、11月26日、都内のアルカディア市ヶ谷で開催された。「四半世紀の時を超えて、そして次なる未来へ」というテーマのもと、例年より多い30セッションに多くの翻訳関係者が参加した。

筆者は、90分枠を2つ使って行われた午後のセッション「MT Live ～機械翻訳の担うべき役割～」に講演者として参加した。機械翻訳を紹介した後、実際に操作して、その使い方や翻訳結果を会場の皆さんに見ていただく、今回初めての企画である。

この企画を聞いたのは2015年6月。JTF理事でもある(株)十印の渡邊社長から相談を受けた。渡邊社長は機械翻訳に関して理解がある方で、翻訳業界での利活用が重要と考えておられる。

「翻訳家の集まりであるJTFでも、機械翻訳に関する関心は高まっている。既に活用している人もいるが、決して多くない。話を聞いたことはあるが、実際の翻訳を見た人は少ない。そこで、セッションの聴講者の前で実際に機械翻訳ソフトを操作し、生の機械翻訳を見てもらいたい。当日翻訳する課題文は、当日までベンダーに公開しない。事前準備なしでエンジン本来の実力を披露したい。」

これが渡邊氏による「MT Live」発案の動機である。これを聞いて「面白い」と思い、すぐに参加を決めた。

2. 主旨の説明と各社システムの紹介

当日の会場は6トラック中で最大の240人収容の大きな部屋であった。どのくらいの聴講者が集まるか心

配したが、講演開始までに満席になり、その後立ち見も出るほどに増えた。

前半セッションの冒頭、モデレータである渡邊氏がこの企画の主旨を説明した。続いて機械翻訳に関する基本的な情報として、RBMT(ルールベース機械翻訳)とSMT(統計的機械翻訳)の特徴、プレエディットとポストエディットの役割などを紹介し、多くの聴講者と機械翻訳に関する情報を共有した。ここで渡邊氏がフロアの聴講者に対して、機械翻訳を使っているか質問したところ、約2割が使っていることが分かった。

引き続き、今回参加した機械翻訳システムを、各社の担当者が紹介した。機械翻訳ベンダーとして参加したのは3社。講演の順に示す。

(1) 東芝ソリューション(株) 熊野「The 翻訳プロフェッショナル」

翻訳精度を高める知識、Microsoft Office などとの連携、特許文書処理などを紹介。

(2) (株)クロスランゲージ専務取締役玉井康裕氏「WEB-Transer」

Web ページの情報を、簡単に多言語翻訳できる機能を、実際の利用サイトの情報を含めて紹介。

(3) (株)NTT データ技術開発本部松永務氏「技術文書を対象にした法人向け機械翻訳サービス」

NTT 研究所の統計機械翻訳エンジンで、顧客保有の対訳データを学習して翻訳するサービスを紹介。

3. 現場で渡された課題文を翻訳

各社のシステムで課題文の英日翻訳を行った。課題文は、後半で登壇する翻訳者3人が用意したもので、特許文24文、IT分野43文、一般文29文。この内容はこの時点まで非公開であった。各社の機械翻訳は、

ユーザ辞書のカスタマイズなどで準備することはできない。会場で調整できるのは、入力文書を概観して、適切な専門用語辞書を選択することくらいである。

なお、NTTデータのシステムは技術文書専用であり、「一般」文書を翻訳するための学習データがないため、一般文は翻訳しなかった。

各システムでの原文入力から翻訳までの操作手順は、大型スクリーンに投影され、講演者と聴講者の間で、ライブ感を共有することができた。

この課題文翻訳で各社のシステムが出力した訳文は、事務局が取りまとめ、後半のセッションまでに印刷されて聴講者全員に配られた。

4. 専門翻訳者が訳文を評価

後半 90 分は、各分野の専門翻訳者が自分の用意した課題文の翻訳結果に対して、各自が講評することから始まった。聴講者は、直前に印刷された翻訳結果を手にしているの、すべての訳文を見ることができる。

一人目は、久遠特許事務所代表で弁理士の奥山尚一氏。特許文書翻訳のエキスパートで、特許の課題文に対する講評を行った。

特許文の翻訳の難しさを説明した後、機械翻訳の有効性を述べた。「技術用語が正しく翻訳されており、特許請求範囲のような特殊な文を除けば、下訳には十分な精度。現在の機械翻訳は十分利用可能なレベル」と評価した。予想以上に好意的なコメントだった。

二人目は、エイブス技術翻訳スクール校長・チーフインストラクタの疋田正俊氏。技術文書翻訳の専門家で、IT 分野テキストの翻訳結果に講評した。

「representative は、組織の代表者を示す場合が多いが、担当者を指すこともあり、正しい翻訳には文脈や背景知識が必要。主語の you や助動詞の may は、日本語に訳さないほうが多い場合が多い。」など、英文技術文書に対する日本語訳のノウハウを紹介した。奥山氏同様、機械翻訳に対して活用可能であるとコメントした。

最後は、日英翻訳者で翻訳学校講師の遠田和子氏。

出題した「一般」課題文には、機械的な判断が難しい多義語を含むものや、通常の技術文を口語的に言い換えた文などが含まれていた。

“He sat behind the wheel and ...” 「この”wheel”を車輪ではなくハンドルと判断するには、句としての登録か、共起語を参照した知識が必要である。」とコメントした。“My cell is running out of juice.” は携帯電話の充電が切れそうな状況を表している口語表現だが、The 翻訳プロフェッショナル、WEB-Transter とともに正しく訳すことはできなかった。もちろん、The battery of my cell phone is running low. という通常の表現は、どちらも正しく翻訳できた。

「同義の言い換え文を翻訳するのは、機械では無理だろう。比喩の表現も機械では無理。人間が活躍する部分だ。」と人間の翻訳知識の重要性を強調して講評を締めくくった。

5. ディスカッション

機械翻訳に関して、ベンダー、翻訳者だけでなく、フロアの聴講者を交えてディスカッションを行った。

翻訳業界での課題の一つは、多言語翻訳らしい。英語、中国語、韓国語などの主要な言語だけでなく、欧州各言語、アジア各言語、などの少量多品種の翻訳をサポートするツールが必要という意見が述べられた。

モデレータから「テクノロジーは翻訳の世界を変えるか」というテーマが示され、翻訳者の一人は「翻訳メモリが発案された時には関心を持つ人は少なかったが、今では翻訳に必須のツールになっている。機械翻訳も、このようになるだろう。」と述べた。私は、既に必須のツールになっていると考えているが。

6. おわりに

JTF 翻訳祭に初めて講演参加したが、技術や製品を直接紹介できるだけでなく、翻訳家の講評からも機械翻訳の実用性を示すことができ、有意義であった。

企画準備段階から支援していただいた日本翻訳連盟関係者の皆さんに感謝します。

特許公報・技術文献にフォーカスした npat 翻訳サービス

本間 奨

日本特許翻訳株式会社 <http://npat.co.jp>

1.はじめに

これからは、TPP で代表される国と国との様々な障壁が取り除かれ、同じ仕組みで運営されていくこととなります。いろいろな障壁が取り除かれていくとしても、最後に残る最大の障壁は言語の壁が残ります。知的財産分野でも海外、とりわけ米国、欧州、中国、韓国そしてこれからはアジアの外国特許も日本の特許同様に調査することが必要となってきます。このときに、言語の壁が大きく立ちはだかります。人手で1件の外国公報を翻訳すると十万円をこえる多大な費用と時間が発生します。一方現時点では機械翻訳では精度の面で課題があります。

さらに、海外の特許公報（明細書）を技術文献として利用しようとする、国・公報種別で段落の順序がことなり、図と図の説明と本文が別々の頁にレイアウトされるため特許公報を技術文献として利用する場合、必ずしも使いやすいものではありませんでした。

日本特許翻訳株式会社が独自に提供する PatSpread では、重要な海外の特許公報（明細書）の全文をわかりやすい正確な日本語で、さらに見やすい2頁見開き形式(スプレッド形式)で横長表示され、PC ディスプレイの横長画面にマッチするため見易いというメリットを持っています。あわせてどの国のどの種別の公報でもフロントページは発明の名称、要約、代表図、請求の範囲が1画面に集約されており、発明特定事項の理解がしやすいというメリットがあります。この PatSpread は以下に説明する翻訳システムを用いて、外国公報の翻訳を行っています。

2.英語・中国語・ドイツ語・フランス語・韓国語から日本語翻訳に対応可能な npat 翻訳サービス

欧米圏の言語では、S(主語)-V(動詞)-O(目的語)の語順が同一であり、このような言語間の機械翻訳は精度がよく広く実用化されています。一方日本語ではS-O-Vとその語順が大きく異なり、英日の統計的機械翻訳(SMT)の精度はルールベース翻訳を上回ることができませんでした。

2010年頃を境に、あらかじめ訳文対となる英文をS-O-V形式の語順変換を行ってからSMTを適用する方法を用いることで、英日でも「わからない」⇒「わかる」へ質的に精度が向上し、2013年のNTCIR-10タスクで、SMTがルールベースの精度を上回ることが初めて立証されました。

日本特許翻訳株式会社では、このような語順変換SMTをコアにその前処理のプリエディット工程と後処理のポストエディット工程を膨大な事例分析から学習したMT Plus 翻訳システムを開発しました。その成果は、特許庁実用新案抄録の翻訳に採用されすでに100万件以上が特許庁検索システムJ-PlatPatで使用されています。npat 翻訳サービスは、このMT Plusを用いて、外国公報の翻訳公報PatSpreadと技術文献や知財文書を対象としたドキュメント翻訳、テキスト翻訳からなるクラウドサービスで、図1の翻訳フローで提供されます。

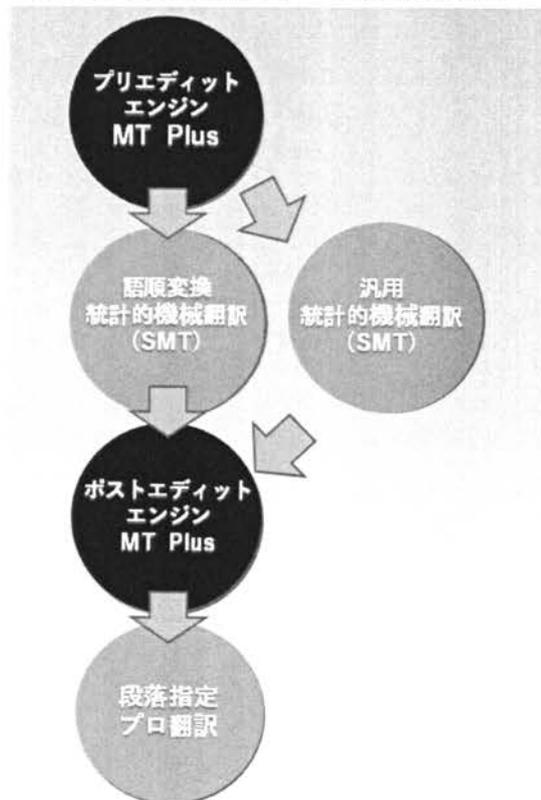


図1 翻訳フロー（プロ翻訳は必要に応じて提供される）

3. 翻訳精度

npat 翻訳で用いる MT Plus の英日・中日翻訳精度については、競合サービス比較で高精度であることを報告しました。^{注1)}

以下に中日翻訳精度について、特許庁(JPO)「中韓文献翻訳・検索システム」との比較をおこない図2に結果をまとめました。図2は、日本企業からの特許出願や翻訳業務をおこなっている中国の大手特許事務所の日本部責任者 雙田飛鳥氏にお願いして評価していただいたものです。

図2の評価は、100文字あたりの誤訳合計スコアを100から差し引いた値です。翻訳精度評価は、

- ①未訳
 - ②文法誤り(主語、動詞、目的語、状語、定語)
 - ③単語の誤訳
- の観点で評価されます。

npat 翻訳(図中 MT Plus)が青バー、特許庁が赤バーで示しました。

全体評価スコアは npat 翻訳が 70 に対して JPO「中韓文献翻訳・検索システム」では 14 のように npat 翻訳が高いスコアを示しています。詳細にみると、JPO は同じ請求項 1 でも特許間のばらつきが大きいことがわかります。

このようなばらつきについては、2014 年度 中国 IPG 特別テーマ調査委員会の報告^{注2)}でも指摘されています。

この報告では、JPO 翻訳「中韓文献翻訳・検索システム」について

『(1) 中国語の先行文献調査

発明・実用新案の出願前先行文献調査をする分には、中国語の解らない発明者・知財スタッフであっても概ね発明の内容を理解できるため、有効なツールであると考えます。

(2) クリアランス時のノイズ取り

例えば、母集団が 1000 件ある中からノイズを取って数百件に件数を絞ることは可能であると考えます。ただし、発明の内容をほとんど理解できない翻訳文が含まれるため、完全にノイズを除去するのは困難である。この場合、より詳細な知財検討を行いたい場合、別途人の手による翻訳等を検討する必要があると考えます。』

のように結論づけています。

(2) の下線部の理由として、「誤訳が多くほとんど理解できない」割合が全体の 17.5%あるためとしています。

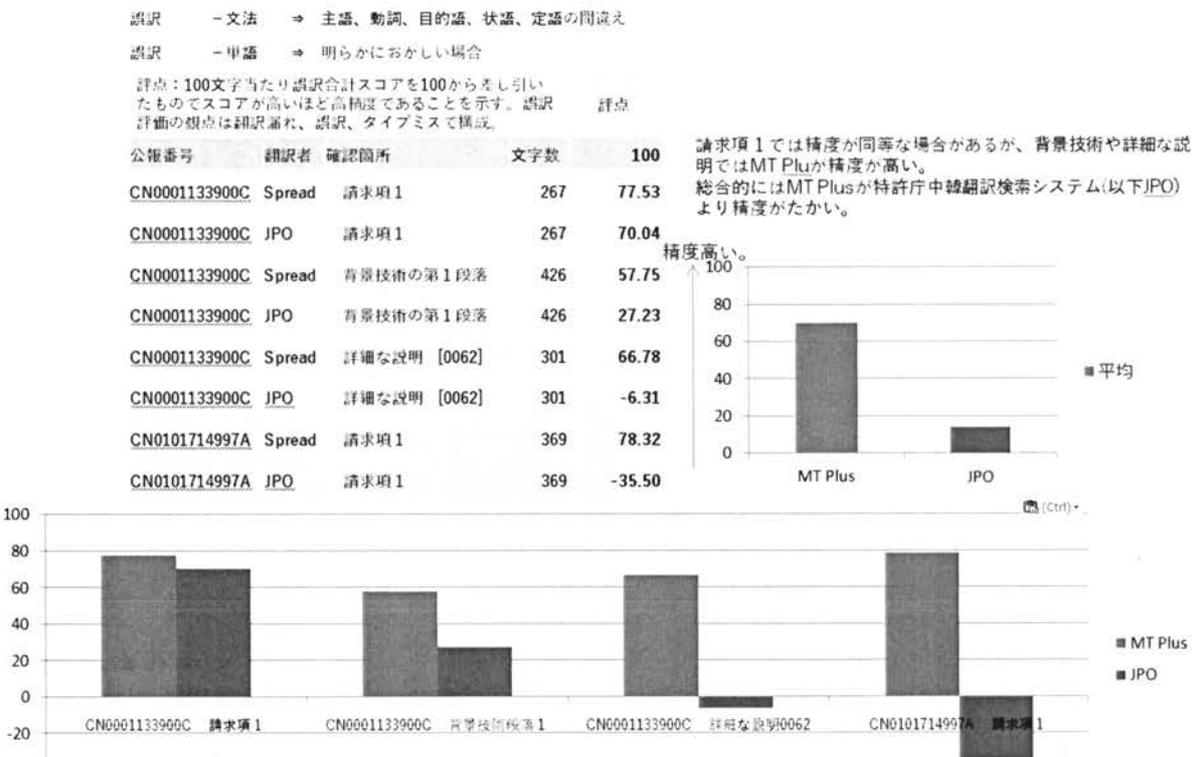


図2 npat 翻訳システムの中日翻訳と特許庁「中韓文献翻訳・検索システム」との比較結果

4.外国公報の日本語翻訳公報:PatSpread

PatSpread は、米国特許、中国特許・実案、ヨーロッパ特許、WO 特許の五大特許庁に加え韓国、DE、FR、GB、CA などの外国特許庁から発行された外国公報の日本語翻訳公報のことです。PatSpread では図 3 のように PC ディスプレイにマッチした 2 頁見開き形式で提供されます。

PatSpread の特徴は npat 翻訳システムにより高精度に翻訳されることに加え、以下の①～③の工夫により見やすく表示されます。

①「図」と「図の説明」と「本文」は公報中で別々のページに掲載されているため技術文献として利用しにくいという課題がありました。figref タグまたはテキスト参照をおこない「図」と「図の説明」を本文に挿入して技術的な内容を理解しやすくしました。以下は EP のドイツ語特許の図挿入例です。

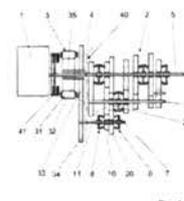


図1は、本発明の構成要素を有する機械的装置を示す。本図に示すものは、図1に示す構成要素の組立状態を示す。本図に示すものは、図1に示す構成要素の組立状態を示す。本図に示すものは、図1に示す構成要素の組立状態を示す。

②mathML に対応した複雑な数式表示が可能

USP/EP/CN では複雑な数式に mathML という特殊な xml を用いています。「中韓文献翻訳・検索システム」では、中国特許 mathML はタグのみで表示されるため、数式内容は原文公報イメージを見る必要がありますが、PatSpread では mathML レンダリングを可能としており、数式が請求の範囲の構成要素となるケースや技術文献として利用する場合、数式内容の理解が可能となっています。

以下に USP の数式表示例を示しました。

[00019]
閉ループ又は二個の同一の負荷に対して使用される出力の場合には、元の式(1)で偏微分方程式の問題は、固有関数に直交条件に基づいて複数の高階微分方程式に変換される。そこにおいては、 q_0 (t) 式(7)式(6)の二次高階微分関数の駆動項目である。
[00020]

$$\frac{d^2 A_m}{dx^2} + \sum_{n=1}^{\infty} A_n(t) = q_0(t) \quad (6)$$

$$q_0(t) = \int_0^L \left(-\frac{\partial \rho}{\partial x} \frac{\partial \phi}{\partial x} \right) \phi_n dx \quad (7)$$

[00021]
なお、固有関数 $\phi_n(x)$ の直交性に基づいて設計された入力電極は圧電トランスのモードル・フィルタリング機能を行い得る式(7)から導出する。このため、入力電極構成 A(x) 式(8)として表される第3モードに動作モード用に設計されている。
[00022]

$$A(x) = \sin\left(\frac{2\pi x}{L}\right) \left\{ H\left[x - \frac{L}{4}\right] - H\left[x - \frac{3L}{4}\right] \right\} \quad (8)$$

[00023]
なお、図2に示すように、入力電極の希望の構成式(8)から導出する。そこにおいては、圧電トランスの長さは、対応する電極構成2)に対して正規化される。式(6)の支配方程式において駆動項目は(7)に式(8)を代入することにより以下のように表示される。
[00024]

$$q_0(t) = \frac{2\pi}{L} \sqrt{\frac{2}{L}} \int_0^L \left(\frac{\sin\left(\frac{2\pi}{L}\left(x - \frac{L}{4}\right)\right)}{2L - 3L} + \frac{\sin\left(\frac{2\pi}{L}\left(x - \frac{3L}{4}\right)\right)}{2L - 3L} \right) E_m^m(t) \quad (9)$$



図3 PatSpread MT 中国特許サンプル例。原文 XML に書誌的事項・要約・代表図を組み合わせて作成される。

③中国語・韓国語の公報中の表については表内の見出し語も翻訳され、大きな表も見やすく表示されます。(以下中国特許例)

[00107]
[テーブル1]
[00108]

	比較例1である	比較例2である	実施例1	実施例2
絶縁性無機材料の比[vol%]	40	40	40	40
導電動(電)性無機材料の比[vol%]	30	30	30	30
微細孔の平均直径[μm]	1	3	1	1
微細孔の含有比[vol%]	30	30	30	30
中空の有無	—	—	2箇の場所	1箇の場所

[00109]

	比較例1である	比較例2である	実施例1	実施例2
絶縁性無機材料の比[vol%]	40	40	40	40
導電動(電)性無機材料の比[vol%]	30	30	30	30
微細孔の平均直径[μm]	1	3	1	1
微細孔の含有比[vol%]	30	30	30	30
中空の有無	—	—	2箇の場所	1箇の場所

5.ドキュメント翻訳

ドキュメント翻訳とは、科学技術文献のもつ複雑なレイアウトや図・表・上付き下付き文字、欧文フォント、ハイパーリンクなどのドキュメント情報を正確にレイアウトして翻訳ドキュメントに反映する独自技術によるドキュメントの翻訳サービスです。お客様の技術文献や PatSpread でカバーされない pdf 公報/審決書/拒絶理由通知/補正クレームなどの知財関連文書をアップロードするだけで同じ

レイアウトの翻訳結果を取得することが可能です。原文ドキュメントはスキャンイメージファイル (.jpeg, .tif, .png) または adobe 社の Acrobat か Distiller で作成された pdf またはリッチテキストを用います。翻訳ドキュメントは、pdf とリッチテキストの2つの形式でダウンロード可能となります。ユーザーは、リッチテキスト形式の翻訳ドキュメントを利用して、必要に応じて、ユーザーの MS office ワード環境で翻訳編集が可能となります。

ドキュメント翻訳は科学技術文献のほか、審決書や拒絶理由通知、補正クレームなどの知財関連文書の迅速な翻訳サービスをご利用いただけます。

6.おわりに

特許検索サービス分野では、中国・韓国を含む五大特許庁から発行される外国特許の日本語対応が急務の課題であり、正確で見易い特許翻訳が強く望まれています。あわせて、非特許文献である科学技術文献についても、英文以外の文献調査も必要になってきており、npat 翻訳サービスが外国特許調査の生産性向上、ひいては研究開発の一助となることを願ってやみません。

注 1) AAMT ジャーナル 56 号 p39-42

注 2) 「中韓文献翻訳・検索システムに関する調査研究」

<http://www.jetro-kpik.org/html/ipg/report/20150414202101.pdf>

Green and Sustainable Chemistry, 2013, 1, 17-42
http://dx.doi.org/10.4236/gsc.2013.11017 Published Online May 2013 (http://www.scirp.org/journal/gsc)

**有害な液体廃棄物のリサイクルに用いられる
Ag, Hg, Cr沈澱**

Silvia C. Gutierrez-Gutierrez^{1,2}, Guadalupe Riancho-Oviedo¹, Carlos Vega-Lobato¹, Juan Manuel Lopez¹, Pedro Arida Perez¹, Susana M. Martinez Galeano¹, Daniel Cruz Gonzalez¹
¹Research Department, Universidad Politécnica del Estado de Puebla (UPAEP), Puebla, México; ²Technological Research Division, Instituto Nacional de Investigaciones Científicas (CONICYT), Sistema de Estudios Avanzados, Instituto Tecnológico de Toluca (Toluca, México); ³Chemical Engineering, Instituto Tecnológico y de Estudios Superiores de Puebla (ITESP), Mexico, México; ⁴Instituto Tecnológico y de Estudios Superiores de Puebla (ITESP), Mexico, México

Received February 11, 2013; revised March 12, 2013; accepted March 20, 2013

著作権 © 2013 Silvia C. Gutierrez-Gutierrez. これはオープンアクセスの記事です。著作権者に帰属する権利を認め、無制限の使用、配布、および複製を許可する。この作業はCreative Commons Attribution Licenseの下で提供されています。

ABSTRACT
有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱は、環境への汚染を防止するために重要である。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。

Keywords: Hazardous Liquid Waste, Treatment for Recycling, Silver, Mercury, Chromium

1. Introduction
有害な液体廃棄物のリサイクルは、環境への汚染を防止するために重要である。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。

18 S. C. GUTIERREZ-GUTIERREZ ET AL.

有害な液体廃棄物のリサイクルは、環境への汚染を防止するために重要である。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。

2. Experimental
有害な液体廃棄物のリサイクルは、環境への汚染を防止するために重要である。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。

2.1. Ag Precipitation, AgCl
有害な液体廃棄物のリサイクルは、環境への汚染を防止するために重要である。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。

2.2. Hg Precipitation, HgS
有害な液体廃棄物のリサイクルは、環境への汚染を防止するために重要である。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。

2.3. Cr Precipitation, Cr(OH)₃
有害な液体廃棄物のリサイクルは、環境への汚染を防止するために重要である。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。この研究は、有害な液体廃棄物のリサイクルに用いられる有害な液体廃棄物 (Ag, Hg, Cr) の沈澱の最適化を目的として行われた。

図 4 npat 翻訳サービスで提供されるドキュメント翻訳による化学論文の翻訳例。

これまでの AAMT Forum メールマガジン Vol.7

小谷克則

関西外国語大学、機械翻訳課題調査委員会 WG1、WG2

はじめに

AAMT Forum を通じて配信したメールマガジンのバックナンバーをお届けします。今回は第 56 号から第 65 号に加えて、号外が一号あります。

今回は、東京オリンピックや移民問題といった社会情勢に関連したニュースなどがあり、世界中の様々な場面で機械翻訳がコミュニケーションの潤滑油となっていることが伺えました。また、法人会員の皆様方に関連するニュースもありました。今後もメールマガジンを通じて様々な情報を皆様へ配信させていただきます。

尚、メールマガジンとして配信すべき情報がありましたら、是非お知らせください。また、その他、お気づきの点なども AAMT 事務局までお寄せください。よろしくお願いいたします。

では、メールマガジン第 56 号から第 65 号、そして号外を再びお届けします。

AAMT Forum メールマガジン

2015/09/17 配信 vol.056

こんにちは。AAMT Forum メールマガジン担当です。朝夕はめっきり涼しく、過ごしやすくなりました。今年は秋の大型連休、シルバーウィークがあり、いつもよりも秋の魅力をじっくり楽しめそうですね。では、メールマガジン第 56 号をお届けします。

■救急隊の翻訳アプリ開発 消防庁、訪日客に迅速対応

総務省消防庁は、2020 年東京五輪・パラリンピックに向け増加が見込まれる外国人観光客らに対応するため、救急隊員の業務用スマートフォンに導入する多言

語音声翻訳アプリの開発に乗り出します。

<http://www.sankei.com/life/news/150903/lif150903029-n1.html>

■「てがき翻訳」のトライアルサービスを提供開始

株式会社 NTT ドコモは、手書き文字やイラストで意思を伝える翻訳サービス「ひつだん翻訳 TM」の名称を「てがき翻訳 TM」へ変更し、9 日より、Android™ スマートフォンやタブレット向けのトライアルサービスを提供しました。

https://www.nttdocomo.co.jp/info/news_release/2015/09/09_00.html

AAMT Forum メールマガジン

2015/10/02 配信 vol.057

こんにちは。AAMT Forum メールマガジン担当です。仲秋の候、いかがお過ごしでしょうか。衣替えの季節となりましたが、天気の移り変わりの激しいこの頃、その時々に適した装いに気をつけましょう。では、メールマガジン第 57 号をお届けします。

■300 万ダウンロード突破の翻訳アプリ 33 言語対応で大幅リニューアル

エキサイト株式会社は、翻訳アプリ「エキサイト英語翻訳」のフルリニューアルを行い、「エキサイト翻訳」として配信を開始しました。

<http://prtimes.jp/main/html/rd/p/000000474.000001346.html>

■「独語－アラビア語」のグーグル翻訳が利用急増、移民流入で

米グーグルは 25 日、自社が提供する翻訳サービス「グーグル翻訳」について、中東からの移民が殺到しているドイツで、ドイツ語とアラビア語間の利用需要が従来の 5 倍に急増しているとして、これらを第一言語とする人々にサービス改善への協力を呼び掛けました。

<http://www.afpbb.com/articles/-/3061330>

AAMT Forum メールマガジン
2015/10/16 配信 vol.058

こんにちは。AAMT Forum メルマガ担当です。朝晩の冷え込みが日増しに感じられるこの頃、いかがお過ごしでしょうか。紅葉で山が染まり始める季節、今年の色付きはまずまず綺麗になるそうですが、どのように色づくか楽しみです。では、メールマガジン第 58 号をお届けします。

■総務省が多言語音声翻訳、リクルートら 3 社に委託研究開発

リクルートライフスタイルは 10 月 9 日、同社とリクルートコミュニケーションズ、およびフュートレックの子会社の ATR-Trek による共同提案が、総務省の委託研究開発に採択されたと発表しました。

<http://news.mynavi.jp/news/2015/10/12/028/>

■みずほ銀行、タブレット端末を使った自動音声翻訳サービスの実証実験を開始

みずほ銀行は 10 月 1 日、NTT ドコモと連携して、タブレット端末を使った自動音声翻訳サービスの実証実験を開始しました。このサービスを利用すると、ボタンを押してタブレット端末に話しかけるだけで、外国語と日本語を自動的に翻訳できます。両替などで銀行を訪れる海外からの旅行者などへのサービス支援策として、その有効性が期待されます。

<http://news.mynavi.jp/news/2015/10/05/122/>

AAMT Forum メールマガジン
2015/10/30 配信 vol.059

こんにちは。AAMT Forum メールマガ担当です。霜降の候、だんだんと寒さを強く感じるようになって来ました。本格的な冬を前に体が慣れないこの時期、体調の変化には気をつけましょう。では、メールマガジン第 59 号をお届けします。

■パナソニックや NTT、多言語音声翻訳技術を推進するコンソーシアムを設立

パナソニック、NTT、国立研究開発法人情報通信研究機構、パナソニック ソリューションテクノロジー、KDDI 研究所、みらい翻訳らは、10 月 26 日に「総務省委託研究開発・多言語音声翻訳技術推進コンソーシアム」を設立したと発表しました。

<http://japan.cnet.com/news/business/35072511/>

■翻訳エディタに新機能搭載！ベストセラー翻訳ソフトの最新版！「コリヤ英和！一発翻訳 2016 for Mac」シリーズ (DVD-ROM 版) を新発売

ロゴヴィスタ株式会社は、最新の Mac OS 環境に対応した英日・日英翻訳ソフト『コリヤ英和！一発翻訳 2016 for Mac』(DVD-ROM) を 2015 年 11 月 20 日 (金)

より発売します。

<http://www.logovista.co.jp/LVERP/information/news/2015-1023-kry2016mac.html>

AAMT Forum メールマガジン

2015/11/20(11/6)配信 vol.060

こんにちは。AAMT Forum メルマガ担当です。晩秋の候、いかがお過ごしでしょうか。秋の行楽シーズンを迎え、名所は今が紅葉狩りの真っ盛りでしょうか。暖かい日にはそうした所へ足を運ぶのも良いかもしれませんね。では、メールマガジン第 60 号をお届けします。

■多言語音声翻訳アプリ“VoiceTra”(ボイストラ)が進化

NICT は、これまでの研究開発成果を反映させた多言語音声翻訳アプリ“VoiceTra”(ボイストラ)の新バージョンを 10 月 22 日(木)に公開しました。

<http://www.nict.go.jp/press/2015/10/22-1.html>

■東芝、日英/日中の通訳機能を備えた会議支援システムを開発

東芝は 10 月 29 日、外国人とのビジネス会議などにおいて、翻訳結果をリアルタイムでパソコンやタブレットなどに字幕表示し、外国語コミュニケーションを支援する日英・日中の通訳機能を備えた会議支援システムを開発したことを発表しました。

<http://news.mynavi.jp/news/2015/10/30/147/>

AAMT Forum メールマガジン

2015/11/20 配信 vol.061

こんにちは。AAMT Forum メルマガ担当です。秋の日は釣瓶落としといいますが、この頃は本当に日が落ちるのが早くなりました。師走となれば何かと忙しくなりますので、準備できることは早めに準備をしましょう。では、メールマガジン第 61 号をお届けします。

■日本語を自動で翻訳し拡声するメガホン型翻訳機を成田空港が試験配備

成田国際空港は 12 日、空港での緊急時案内用として、日本語を多言語に翻訳して繰り返し拡声することができるメガホン型翻訳機「メガホンヤク」を試験配備することを発表しました。

<http://www.rbbtoday.com/article/2015/11/13/137007.html>

■翻訳サービスを使うコツは、日本語脳と英語脳を使い分けること

機械翻訳で上手に日英翻訳を行うために考えるべき日本語と英語の違いと、中間日本語についての記事です。

http://www.lifehacker.jp/2015/11/151106book_to_read.html

AAMT Forum メールマガジン

2015/12/04 配信 vol.062

こんにちは。AAMT Forum メルマガ担当です。師走の候、いかがお過ごしでしょうか。今年も残りあとわずかとなりました。何かと忙しい時期ではありますが、怪我や病気などなく新年を迎えることができるように気をつけましょう。では、メールマガジン第 62 号をお届けします。

■クロスランゲージ、業務用翻訳ソフトの新バージョンを2種類発売！

ビジネス／科学技術分野向けの業務用翻訳ソフト「PC-Transer 翻訳スタジオ V22 for Windows」は12月4日（金）、医学分野に特化した翻訳ソフト「MED-Transer V14 for Windows」は12月18日（金）より販売が開始されます。

<http://prtimes.jp/main/html/rd/p/000000077.000002804.html>

■NTT ドコモ、「てがき翻訳」の正式版を提供開始、Android アプリをアップデート

NTT ドコモは11月30日、9月よりトライアルサービスとして提供している翻訳アプリ「てがき翻訳」を正式版に昇格させて改めてリリースしました。

<http://getnews.jp/archives/1278966>

■鳥取市の「1000円タクシー」が位置情報で翻訳精度上げる社会実証

11月18日より、KDDI株式会社と株式会社 KDDI 研究所による多言語音声翻訳システムを活用した社会実証が、鳥取市を訪れる訪日外国人向けの周遊観光タクシー（1000円タクシー）で、実施されることになりました。

http://qzss.go.jp/usage/userreport/tottori_151127.html

AAMT Forum メールマガジン

2015/12/18 配信 vol.063

こんにちは。AAMT Forum メールマガ担当です。一年が経つのは、ほんとうにあつという間ですね。来年

も皆様にとって良き年でありますように心よりお祈り申し上げます。では、メールマガジン第63号をお届けします。

■自動音声翻訳、実用化へ加速 京都で外国人観光アシスト

訪日外国人観光客が増加し、2020年東京五輪も控える中、京都府内で音声翻訳技術の開発が加速しています。

<http://www.kyoto-np.co.jp/environment/article/20151204000063>

■音声翻訳のロボットコンシェルジュ、名古屋駅で実証実験を開始-KNT-CT

近畿日本ツーリストは2015年12月10日と11日、JR名古屋駅コンコースでの「名古屋・東海観光PRイベント」で、音声翻訳システムを搭載した卓上ロボット「Sota」による「ロボットコンシェルジュ」のデモンストレーションを実施しました。

<http://www.travelvoice.jp/20151211-56663>

AAMT Forum メールマガジン

2016/01/08 配信 vol.064

こんにちは。AAMT Forum メールマガ担当です。新年あけましておめでとうございます。今年一年もまた、自分で決めた目標へ向かって頑張りましょう。では、メールマガジン第64号をお届けします。

■クロスランゲージの多言語自動翻訳サービスを「駅パラ」サイトに導入

東日本旅客鉄道株式会社が運営する駅ビルポータル

サイト「駅パラ」に、多言語ホームページ自動翻訳サービス「WEB-Transer@ホームページ」が導入され、英語、中国語繁体字、中国語簡体字、韓国語の各言語による情報の多言語発信が可能になりました。

<http://prtmes.jp/main/html/rd/p/000000079.000002804.html>

■リアル『ほんやくコンニャク』？オフラインでも動作可能な小型音声翻訳機「ili」、日英中に対応

ボタンをプッシュして翻訳したい文章を話すだけの使い方が簡単、オフライン環境で使えるリアルタイム翻訳ガジェット「ili」の紹介記事です。

<http://japanese.engadget.com/2016/01/05/ili-ces-2016/>

AAMT Forum メールマガジン

2016/01/22 配信 vol.065

こんにちは。AAMT Forum メールマガ担当です。厳寒の候、いかがお過ごしでしょうか。今週は全国的に大荒れの天気でしたね。雪の降った後は滑りやすいですが、降っていなくても路面の凍結には気をつけましょう。では、メールマガジン第65号をお届けします。

■沖縄の企業が挑む 日本語音声→ミャンマー語に自動翻訳

システム開発のOCCは、日本語の音声をミャンマー語に自動翻訳するシステムを開発しています。経済開放が進むミャンマーへの進出も視野に2020年までの完成を目指しています。

<https://www.okinawatimes.co.jp/article.php?id=149498>

■ロゴヴィスタ、経済・金融文書向け翻訳ソフトの新

版を発売

ロゴヴィスタ株式会社は、経済・金融ビジネスに関する文書の翻訳に適した英日・日英翻訳ソフトのWindows版「LogoVista 経済・金融ビジネス V2」を2016年1月22日に発売します。

<http://www.keyman.or.jp/nw/20065366/>

AAMT Forum メールマガジン

2016年1月号外

AAMT が策定した用語集形式 UTX の新バージョン 1.20 のベータ版仕様書が日本語と英語で公開されました。

UTX は無料で使える用語集形式の標準規格です。UTX は、Excel で簡単に作成・一覧編集でき、翻訳支援ツール、翻訳ソフトなどさまざまな翻訳ツールで活用できます。

UTX1.20 ベータ版仕様は2016/2/29まで公開され、その間、利用者の方の意見を募集します。多数のご意見をお待ちしております。

<http://www.aamt.info/japanese/utx/>

UTX 1.20 の用語集データも順次公開予定です。また、UTX の活用法、UTX 1.11 からの移行方法なども紹介していきます。

最新情報は翻訳工学ページをご覧ください。

<https://www.facebook.com/translation.engineering>

実務翻訳における機械翻訳の利用に関するアンケート結果報告

小谷克則

関西外国語大学、AAMT 機械翻訳課題調査委員会

1. はじめに

AAMT 機械翻訳課題調査委員会では機械翻訳に関する調査・広報・啓蒙活動を担当するワーキンググループ (WG2) を中心に、産業翻訳業界における機械翻訳システム活用の実態を調査してきた。この調査活動の一環として、2012 年以降、JTF 翻訳祭においてアンケートを実施してきた。JTF 翻訳祭は、翻訳者、翻訳会社、翻訳発注企業などの翻訳業界団体である日本翻訳連盟 (JTF) が主催するイベントであり、講演会、展示会により構成される。

本稿はそのアンケートによる調査結果を報告する。

2. アンケート調査の概要

- 調査実施日

第 25 回 JTF 翻訳祭当日 (2015 年 11 月 26 日) に調査を実施した。

- 調査方法

アンケート用紙 (A4 表裏) 配布による無記名の調査とした。

JTF 翻訳祭の展示コーナーに AAMT がブースを確保し、ブースを訪れた人に加え、ブース前を通りかかっている人に対して直接協力を依頼した。ブースにアンケート用紙と筆記用具を用意しておき、その場でアンケートに回答いただき、回収した。

なお、回答者に抽選で当たる景品として、クオカード、翻訳ソフトを用意した。翻訳ソフトは、中国語⇄日本語翻訳ソフト「J 北京 7 スタンダード 2 定価：48,600 円 (税込)」と英語⇄日本語 翻訳ソフト「PC・Transer 翻訳スタジオ V22 for Windows 定価：128,000 円 (税別)」であった。これらの翻訳ソフトは、本協会法人会員の株式会社高電社様、株式会社クロスランゲージ様のご厚意により提供いただいたものである。この場を借りて心より御礼申し上げます。

- 回答状況

回答者数は 64 名であり、翻訳祭参加者 (1000 名と仮定) 全体に対する回答率はおよそ 6%であった。回答者数は過去の調査と同程度であった。2012 年度は 60 名、2013 年度は 64 名、2014 年度は 62 名であった。



写真 1 翻訳プラザ (展示会) の様子

(出展企業・団体は 40 を数えた。)

3. 回答者

● 回答者の職種

翻訳業務の従事者といっても様々な仕事がある。以下は全回答者 64 名の職種の内訳である。

- | | |
|--------------------------------------|------|
| (1) フリーランスの翻訳者 | 13 人 |
| (2) 翻訳会社の社内翻訳者 | 5 人 |
| (3) 企業（翻訳会社以外）・官公庁等の翻訳者 | 10 人 |
| (4) 翻訳会社の翻訳業務管理者、コーディネーター | 13 人 |
| (5) 企業（翻訳会社以外）・官公庁等の翻訳業務管理者、コーディネーター | 5 人 |
| (6) その他・無回答 | 14 人 |

計 64 名のうち 28 名（44%）が翻訳を生業としている人であり、18 名（28%）が翻訳の管理を仕事としている人であった。

● 性別・年齢

年齢は 20 代から 60 代まで広く分布しているが、30 代（19 人）と 40 代（20 人）が多かった。以下、20 代（8 人）、50 代（7 人）、60 代（4 人）と続く。

性別は女性 34 人に対して男性 17 人（無回答が 13 人）であり、女性の回答者の割合が多かった。



写真 2 AAMT ブースの様子

（左から山本委員と秋元委員、アンケート調査と併行して、AAMT 策定の用語集形式 UTX の紹介を行った。）

4. 調査結果

本節では、回答者のうち翻訳作業そのものを生業とする 28 人を集計の対象とし、項目別に調査結果を紹介する。但し、複数回答や未回答により、回答の内訳の合計が 28 人とならない項目がある。

4.1 機械翻訳の利用状況

● 機械翻訳の利用頻度

翻訳者が普段どのくらいの頻度で機械翻訳を使っているかについて、翻訳者のタイプ別に集計した結果を表 1 に示す。

機械翻訳を「ほぼ毎日使っている」という回答は 20%（4 人）にとどまり、「使っていない」という回答は 25%（5 人）であった。

表 1 機械翻訳の利用頻度

	フリー ランス	翻訳会 社	企業・官 公庁
ほぼ毎日	1	1	2
週 1~3 回	2	1	1
月に数回	2	0	0
年に数回	4	0	1
使っていない	1	1	3

質問：翻訳ソフト/サイトの利用頻度は？

● 利用している機械翻訳の種類

機械翻訳は、インターネット上の無料サービスから、一本が 10 万円以上する高価なソフトまで、様々な形態で提供されている。翻訳者はどのようなソフトあるいはサービスを使っているのだろうか。表 2 は翻訳者がよく利用する機械翻訳の種類についてまとめたものである。

翻訳者がよく使う機械翻訳のトップはこれまでと同様、「無料の翻訳サイト」という結果となった。翻訳会社等から翻訳ソフトが提供される例は少数に留まり、翻訳者が個人的に利用するケースが大半を占めるようである。

近年、翻訳ソフトがスマートフォンや携帯電話のアプリケーションとして搭載されるようになった。そこで、今回、スマホ・携帯アプリを回答項目として追加した。しかし、予想に反し、今回の回答者における利用者は皆無であった。

表2 よく利用する機械翻訳の種類

	フリー ランス	翻訳会 社	企業・ 官公庁
無料の翻訳サイト	4	2	4
有料の翻訳サイト	1	0	0
自社または翻訳会社などで提供される翻訳システム（サービス）	0	0	1
市販のPC翻訳ソフト	1	0	0
スマホ・携帯アプリ	0	0	0

質問：現在最もよく使う翻訳ソフトや翻訳サイトは？

● 機械翻訳の満足度

表3は機械翻訳の満足度を5段階で評価した結果をまとめたものである。「とても満足」と回答した人は皆無だったものの、24%（4人）が「やや満足」と答えており、現在の翻訳システムでも使い方で翻訳者の役に立つ場面があるようである。

表3 機械翻訳全般の満足度

	フリー ランス	翻訳会 社	企業・ 官公庁
とても不満	1	0	0
やや不満	1	2	1
どちらでもない	6	0	2
やや満足	2	0	2
とても満足	0	0	0

質問：翻訳ソフト/サイトの満足度は？
(5段階で評価)

満足度の根拠を具体的に知るために、「翻訳品質」と

「翻訳機能（使い勝手）」について別々に満足度を調査した。以下にその結果を紹介する。

● 翻訳品質

表4を見ると、翻訳品質に満足していると答えた人はわずか1人であり、不満であると答えた人（11人）が65%と大半を占める。したがって、翻訳者を満足させるためには、抜本的な翻訳品質の向上が不可欠といえる。

表4 翻訳品質の満足度

	フリー ランス	翻訳会 社	企業・ 官公庁
とても不満	1	0	1
やや不満	5	2	2
どちらでもない	4	0	1
やや満足	0	0	1
とても満足	0	0	0

質問：翻訳ソフト/サイトの品質の満足度は？
(5段階で評価)

● 翻訳機能（使い勝手）

表5から、機械翻訳は機能としては「満足」と答えた人が41%（7人）に対して、「不満」と答えた人は18%（3人）であることから、機能への満足の高さがうかがえる。翻訳ソフト/サイトの使い勝手については品質ほど不満に感じている翻訳者はいないようである。

表5 翻訳機能の満足度

	フリー ランス	翻訳会 社	企業・ 官公庁
とても不満	1	0	0
やや不満	1	0	1
どちらでもない	5	2	0
やや満足	3	0	4
とても満足	0	0	0

質問：翻訳ソフト/サイトの機能の満足度は？
(5段階で評価)

4.2 実務翻訳の現状

● 言語対と翻訳方向

近年、様々な言語対への対応が進む機械翻訳であるが、日本語を扱う機械翻訳の主たる対象言語は英語、中国語である。今後、どのような言語への対応が必要となるかを検討するため、実際の翻訳業務においてどのような言語対、そして翻訳方向が多いのかを調べた。表6は翻訳業務における主な言語対と翻訳方向についてまとめたものである。

その他の言語対と翻訳方向として、中国語⇒英語、英語⇒多言語、日本語⇒インドネシア語が挙げられた。

表6 言語対と翻訳方向

	フリー ランス	翻訳会 社	企業・ 官公庁
日本語⇒英語	6	3	6
日本語⇒中国語	1	0	5
英語⇒日本語	10	1	5
中国語⇒日本語	0	2	3
その他	1	1	2

質問：現在の翻訳業務における主な言語対と翻訳方向は？

● 対象分野

表7 翻訳対象の言語と分野

	日英	英日	日中	中日	その他
社会科学	4	7	2	1	2
自然科学	1	2	2	1	2
医療	2	2	0	0	1
IT産業	6	4	3	2	2
非IT産業	4	4	2	2	3
文化・芸術	3	2	1	2	1
スポーツ	2	1	1	2	0
旅行・観光	2	3	1	1	1
生活情報	3	2	2	2	0
その他	4	1	0	0	0

表7は、言語対と翻訳方向を翻訳文書の分野に応じてクロス集計した結果である。これを見ると、翻訳対象の言語と分野の関連性がある程度推定できる。

日英翻訳はIT産業での利用が多く、英日翻訳は社会科学での利用が多い。一方、日中・中日翻訳では分野による違いは小さい。

● 言語対・翻訳方向と対象分野

表8は、言語対と翻訳方向を翻訳文書の種類に応じてクロス集計した結果である。これを見ると、翻訳対象の言語と種類の関連性がある程度推定できる。

多くの文書において日英翻訳と英日翻訳が主として用いられていることから共通言語としての英語の優位さがうかがえる。一方、ホームページ、仕様書・設計書、報告書では英語以外の翻訳の有用性がうかがえる。

表8 翻訳対象の言語と種類

	日英	英日	日中	中日	その他
ホームページ	4	5	1	2	1
取扱説明書	3	4	0	1	2
仕様書・設計書	5	2	3	3	3
報告書	5	5	2	2	2
論文	3	1	1	0	1
契約書	5	4	1	1	2
特許	2	3	0	0	1
メール	4	2	1	2	0
その他	5	6	0	0	1

5. まとめと今後の課題

2015年度の第25回JTF翻訳祭で実施したアンケート結果を中心に、実務翻訳における機械翻訳の利用実態について報告した。

今回の調査により、日英翻訳の活用が目立った分野はIT産業、種類では仕様書・設計書、報告書、契約

書が多かった。また、英日翻訳の活用が目立った分野は社会科学、種類ではホームページ、報告書、取扱説明書が多かった。このように業務翻訳においては、翻訳の言語対や方向により対象分野や文書の種類に違いがあることがわかった。

過去の調査結果から、翻訳者の多くが機械翻訳を活用していることが期待されたが、今回は利用率の低下がみられ、今後、機械翻訳の更なる普及活動が課題となることが判明した。また、機械翻訳の品質を翻訳の言語対・翻訳方向、そして分野や種類別に調査し、より一層、業務翻訳への機械翻訳の普及を図る必要がある。

今後も、翻訳者へのアンケート調査を通じて、機械翻訳の利用者の視点から機械翻訳の課題を見い出す必要がある。同様の活動は、本協会が参画した **Machine Translation (MT) Summit XV** (2015年10月30日から11月3日、米国フロリダ州マイアミで開催)における機械翻訳利用者と翻訳者のセッション (the Commercial MT Users & Translators' Track) の設置などが挙げられる。

その他、機械翻訳のセミナーなどを通じての普及活動も課題として挙げられる。例えば、昨年は **JTF** により本協会会長中岩浩巳らをパネリストとして「機械翻訳と向き合うときが来た—MTをもっと身近に、現実的に考える—」というテーマのセミナーが開催された。

同様のセミナーや講演は、全米最大の翻訳者及び通訳者の団体である **American Translators Association (ATA)** の年次大会においても行われた。その一例として、**Hands-on Machine Translation: Do-It Yourself Tools** (Dilinger, Mike)、**Demystifying Machine Translation** (Marciano, Jay)、**Ready to Surf the Machine Translation Post-Editing Wave?** (Palomares, Jose)、**Post-Editing: How to Future-Proof Your Career in Translation** (Camargo, Paulo)が挙げられる。

協会活動報告

(2015年9月～2016年1月)

機械翻訳課題調査委員会

2015年9月18日(金)(2015年度 第6回)

① 各WGの活動について(各WGに分かれて議論)

(WG1、WG2)

- ・MT Summit について
10/30～11/3 アメリカで開催。AAMTは3日目に発表する。
- ・翻訳フェアの展示各社反応について
- ・AAMT/Japio 特許翻訳委員会の活動状況の紹介(中間報告)
拡大評価部会のテーマとして特許文テストセットを作成している。
- ・中日テストセット評価について
これから評価の相関を分析する予定。
- ・JTF 翻訳祭での展示について

(WG3)

- ・ISO/TC37 Liaison 申請の是非について検討した結果、とりあえずは様子見とする。
- ・UTX 1.20 に関する事項について検討した。
- ・UTX 変換ツールの開発経過について報告があった。
- ・JTF 翻訳祭でのWG3の活動について検討した。
- ・UTX1.20 公開に向けての計画を立てた。
- ・宿題

② 全体会議

- ・JTF 翻訳祭の翻訳プラザ(展示会) 機材の使用について
- ・JTF 翻訳祭のプログラム確定に際して、機械翻訳関連のプログラムの確認

2015年10月16日(金)(2015年度 第7回)

① 各WGの活動について(各WGに分かれて議論)

(WG1、WG2)

- ・市場動向調査委員会の報告書の確認をした。
十分な資料が存在しないことがわかった。
- ・翻訳祭で行う予定のアンケート内容をレビューした。

(WG3)

- ・青森で行われたCJNLP発表(山本氏)に関する報告があった。
- ・ISO/TC37 Liaison 申請の是非について再検討した。

- ・ UTX 1.20 の記述に関する検討を行った。
 - ・ UTX 変換ツール/UTX サイトへのリンクを増やすことについて検討した。
 - ・ 宿題
- ② 全体会議
- ・ 各 WG の検討事項共有
 - ・ JTF 翻訳祭の計画を練った。
翻訳プラザ配布物、展示内容の確認など。
 - ・ 宿題
JTF 翻訳祭における翻訳プラザのための機器の手配など。

2015 年 11 月 11 日 (水) (2015 年度 第 8 回)

- ① 全体会議
- ・ 日中会議報告
 - ・ MT Summit 報告
 - ・ JTF 翻訳祭の準備
JTF 祭準備のための分担などについて検討した。

2015 年 12 月 18 日 (金) (2015 年度 第 9 回)

- ① 前回委員会の議事録の確認
- ② 各 WG の活動について (各 WG に分かれて議論)
(WG3)
- ・ UTX 1.20 β の Web 掲載画面の確認
β 仕様の公開期間は 2016/2/29 まで。2 月末から 3 週間はフィードバックを
検討し、その後正式公開とする。
 - ・ 言語処理学会第 22 回年次大会 (NLP2016 仙台)での発表について検討した。
 - ・ UTX 変換ツールの開発経過の報告があった。
 - ・ その他 UTX 関連
- ③ 全体会議
- JTF 翻訳祭に関する感想および反省事項などについてまとめた。

2015 年 1 月 22 日 (金) (2015 年度 第 10 回)

- ① 前回委員会の議事録の確認
- ② 各 WG の活動について (各 WG に分かれて議論)
(WG1,2)
- ・ 中日テストセットメタ評価について

- ・ AAMT/Japio 特許翻訳研究会評価部会との連携について
- ・ 言語処理学会の発表について
- ・ 今年度の成果、来年度の計画
- ・ その他

(WG3)

- ・ UTX1.20 への既存用語変換の試みの報告
- ・ UTX 1.20 と TBX との互換性に関する説明
- ・ UTX1.20 説明資料作成について
- ・ 3月の言語処理学会での発表申し込みに関する報告

③全体会議

- ・ 各 WG の検討事項共有
- ・ 言語処理学会のテーマセッションについて
- ・ その他

・ インターネットWG

以下の活動をオンラインで実施した。

- ① 法人会員情報等の総会・AAMT 機械翻訳フェアに関連したコンテンツの更新
- ② AAMT ジャーナル目次情報等の通年コンテンツの掲載・更新
- ③ 利便性・プレゼンス向上に向けたサイト構成見直しの検討
- ④ 次回MTサミットに向けたホームページ活用の検討を開始

編集委員会

2015年11月27日（2015年度 第3回）

- ① AAMT Journal No.61 の執筆依頼計画について

AAMT/Japio 特許翻訳研究会

2015年9月25日(金)(2015年度 第4回)

1. 議事録の確認
2. 静岡大学大学院 綱川先生からのご発表
3. 拡大評価部会の報告
4. MT サミットの Workshop の説明
5. その他
6. 次回の開催について
 - ・ 開催の日時(場所) 12月4日(金)
 - ・ 次回研究会後の忘年会のお知らせ
 - ・ 予定
 - ・ 主な議題

AAMT/Japio 特許翻訳研究会 拡大評価部会

2015年9月25日(金)(2015年度 第2回)

1. 2015年度の拡大評価部会の活動計画
 2. 各グループの活動発表
 - ・ 自動評価グループ
 - ・ 人手評価グループ
 - ・ テストセットグループ
 3. その他
- 次回開催日予定 1月下旬予定

AAMT/Japio 特許翻訳研究会 MT サミット参加

2015年10月30日(金)から11月3日(火)

マイアミ、フロリダ州のハイアットリージェンシーマイアミホテルでMT サミット参加。

AAMT/Japio 特許翻訳研究会

2015年12月4日(金)(2015年度 第5回)

1. 議事録の確認
2. 筑波大学宇津呂研究室、龍様のご発表
3. MT サミットの報告 須藤様 (NTT コミュニケーション科学基礎研究所) のご発表
加藤様 (特許庁) PSLT ワークショップでの発表内容のご紹介

4. 27年度 AAMT/Japio 特許翻訳研究会報告書作成スケジュールなど
5. その他
6. 次回の開催について
 - ・開催の日時（場所）2016年1月29日（金）
第3回拡大評価部会も開催（開始時間は16：00～に変更されました。）
 - ・主な議題

AAMT/Japio 特許翻訳研究会

2016年1月29日（金）（2015年度 第6回）

1. 議事録の確認
2. 越前谷先生のご発表
3. 27年度 AAMT/Japio 特許翻訳研究会報告書
27年度報告書の目次案とスケジュール
4. その他
5. 次回の開催について
 - ・開催の日時（場所）2016年3月11日（金）
奈良先端大・ニュービック先生をお招きし「ニューラル機械翻訳(仮)」講演決定
 - ・予定
 - ・主な議題

AAMT/Japio 特許翻訳研究会 拡大評価部会

2016年1月29日（金）（2015年度 第3回）

1. 各グループの活動発表
 - ・自動評価グループ
 - ・人手評価グループ
 - ・テストセットグループ
2. その他
3. 次回開催日予定 5月予定

AAMT ジャーナル編集委員会委員長
筑波大学 システム情報系 知能機能工学域
宇津呂 武仁

AAMT ジャーナル 61 号をお送りします。

今号の巻頭言は、本年をもって定年をお迎えになる静岡大学の梶博行先生より、御寄稿を頂きました。

理事の坂本義行先生からは、機械翻訳の歴史に関する貴重な御寄稿の第六弾を頂きました。

プロジェクト報告として、特許庁の加藤啓様より、特許庁における機械翻訳に関する取り組みについて、貴重な御寄稿を頂きました。また、情報通信研究機構隅田 英一郎様より、2020 年東京オリンピックに向けた国家プロジェクト「グローバルコミュニケーション計画」の活動について、貴重な御寄稿を頂きました。

その他、関西外国語大学 外国語学部 魚住真司先生からは、映像翻訳に関する御寄稿を頂きました。

一方、今号に先立ちまして、昨年 11 月 26 日(月)に、AAMT が後援をします第 25 回 JTF 翻訳祭が開催され、「MT Live ～機械翻訳の担うべき役割～」が企画されました。講演者の東芝ソリューション熊野明様には、この企画の開催報告を御寄稿頂きました。

また、日本電信電話 コミュニケーション科学基礎研究所 須藤克仁様より、昨年 10-11 月に米国マイアミにおきまして開催されました第 15 回機械翻訳サミットへの参加報告を御寄稿頂きました。このサミットの併設ワークショップでは、AAMT/Japio 特許翻訳研究会のメンバーが中心となって特許・技術文書翻訳ワークショップを開催しましたので、このワークショップ開催の様子につきましても、あわせて御報告頂きました。

さらに、昨年 10 月 16 日(金)に開催されました、第 2 回アジア翻訳ワークショップの報告を、科学技術振興機構 中澤敏明様、情報通信研究機構 美野秀弥様、隅田英一郎様、NHK 後藤功雄様、奈良先端科学技術大学院大学 Graham Neubig 先生、京都大学 黒橋禎夫先生より御寄稿頂きました。

その他、秋桜舎 山本ゆうじ様より、昨年 10 月に青森市で開催されました第 15 回日中自然言語処理共同研究促進会議の参加報告を御寄稿頂きました。

日本特許翻訳株式会社 本間奨様からは、「特許公報・技術文献にフォーカスした npat 翻訳サービス」を御寄稿頂きました。

AAMT 内の活動報告として、機械翻訳課題調査委員会から、昨年の JTF 翻訳祭において実施された「実務翻訳における機械翻訳の利用に関するアンケート」の結果報告、および、「AAMT フォーラムメールマガジンバックナンバー」を掲載いたしました。

AAMT

Asia-Pacific Association for Machine Translation

AAMT 入会のご案内

AAMT は、機械翻訳の発展を目的として、機械翻訳の研究者、開発者、製造者、利用者が集まった任意の組織です。委員会による定期的な調査研究をはじめ、機関誌の発行、シンポジウム、セミナー等各イベントの開催など幅広く活動を行っています。

機械翻訳にご関心のあるすべての方にご入会をお勧めします。

** AAMT 会員の特典 **

1. AAMT Journal の購読ができます。

会員には、機関誌である AAMT Journal（年 2～3 回発刊予定）が送付されます。購読料は年会費に含まれています。

2. 機械翻訳関連の最新情報をメールでお届け

会員専用メーリングリストで、最新の機械翻訳関連の情報をお届けします。

MT 新製品、新サービスの紹介、国際会議、シンポジウムのお知らせ、WEB での MT 関連記事の紹介など盛りだくさんです。

3. AAMT が組織する委員会や調査活動に参加し、機械翻訳や翻訳に関心のある方との交流を深め、知見を広めることができます。

機械翻訳に関する言語資料の調査、広報、標準化活動に参加したり、AAMT Journal や会員専用メーリングリストで、自社製品、サービスの紹介を行うことができます。

4. 関連機関の主催する国際会議に参加できます。

IAMT の主催で隔年開催される MT Summit をはじめ、AAMT、AMTA*、EAMT** の主催する会議やワークショップに参加できます。

AMTA* : Association for Machine Translation in the Americas

EAMT** : European Association for Machine Translation

年会費は以下の通りです。

法人会員：入会金 1 口 10,000 円 年会費 1 口 50,000 円

個人会員：入会金 1,000 円 年会費 5,000 円（学生は学生会費 1,000 円）

ご関心のある方は、事務局までお問い合わせください。

アジア太平洋機械翻訳協会（AAMT）

ホームページ：<http://www.aamt.info>

電子メール：aamt-info@aamt.info

AAMT



AAMTジャーナル No.61

発行：アジア太平洋機械翻訳協会（AAMT）

ホームページ：<http://www.aamt.info>

住所：〒171-0014 東京都豊島区池袋2-55-2鈴木ビル3階

（株）日本システムアプリケーション内

phone：03-5951-3961 fax：03-5951-3966

編集委員会：宇津呂 武仁 小谷 克則 大倉 清司

阿部 さつき 釜谷 聡史

表紙(図部分)デザイン：阿部 さつき

事務局：神崎 享子 荻野 孝野

印刷所：株式会社ユリクリエイト

Asia-Pacific Association for Machine Translation (AAMT)

c/o Japan System Application Co., Ltd.

Suzuki Building 3F 2-55-2, Ikebukuro, Toshima-ku Tokyo 171-0014, JAPAN

aamt-info@aamt.info