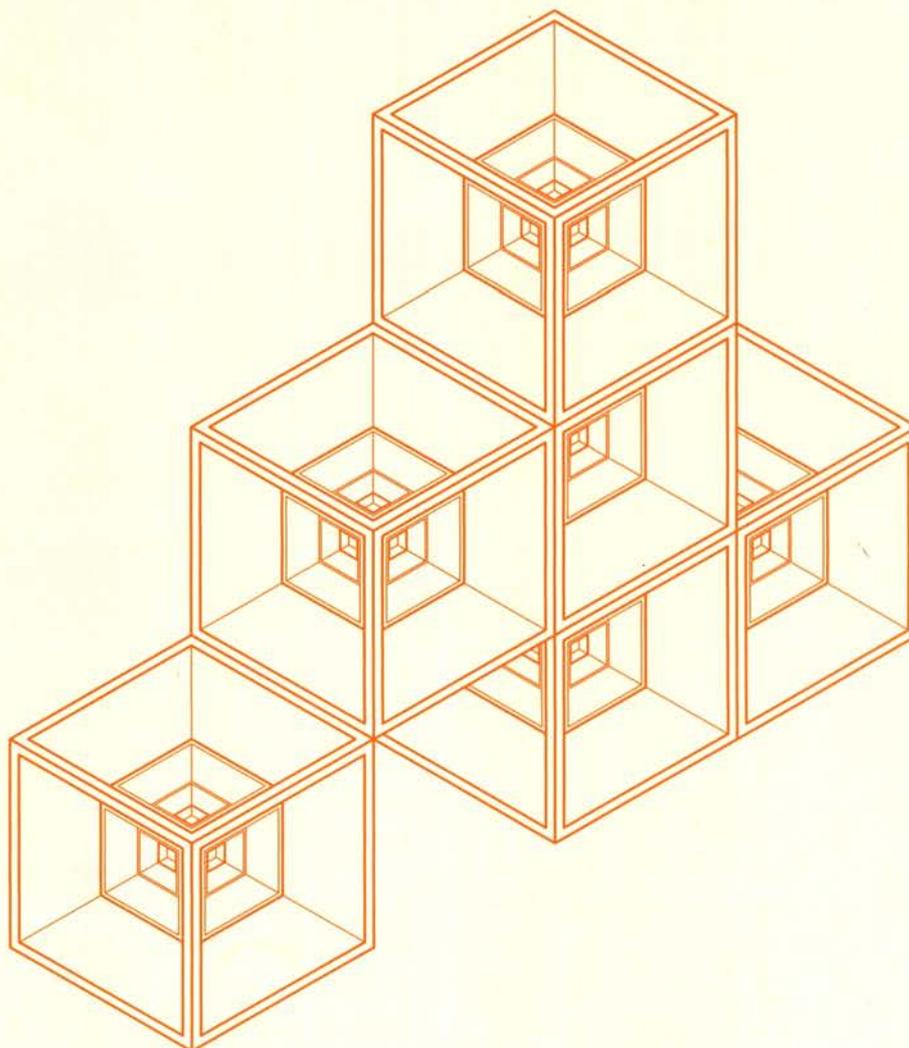


# AAMT

Asia-Pacific Association for Machine Translation

# Journal



December 2009 **No.46**

アジア太平洋機械翻訳協会

# 目 次

巻頭言：	青い鳥を探して .....	伊藤 悦雄..... 1
追悼田中穂積先生：	追悼 田中穂積先生 .....	井佐原 均..... 2
	追悼 田中穂積先生 .....	長尾 眞 ..... 3
	田中先生を偲ぶ.....	辻井 潤一..... 5
MT Summit 報告：	MT Summit XII 参加報告ー特許翻訳ワークショップを中心にー .....	江原 暉将..... 6
	MT Summit XII 参加報告 .....	宮澤 信一郎..... 8
	第 12 回機械翻訳国際会議 (MT Summit XII) および第 3 回特許機械翻訳ワークショップ (The 3 <sup>rd</sup> Workshop on Patent Translation) 参加報告 .....	横山 晶一..... 11
	MT Summit 2009 参加報告.....	内山 将夫..... 14
ACL2009：	Report on ACL-IJCNLP 2009 .....	Canasai Kruengkrai .. 15
	Conference Report Joint conference of the 47 <sup>th</sup> ACL and the 4 <sup>th</sup> IJCNLP .....	..... Wenliang Chen... 18
	Report on ACL 2009 at Singapore .....	Hailong Cao..... 20
AAMT 長尾賞：	AAMT インターネットワーキンググループの長尾賞受賞に際して .....	AAMT IWG... 22
委員会活動報告：	英日・日英機械翻訳に関するアンケート結果報告 .... AAMT 機械翻訳課題調査委員会 WG2 ..	25
	ユーザー辞書共通フォーマット UTX (英語版) .... AAMT 機械翻訳課題調査委員会 WG3 ..	39
会員紹介：	AAMT 会員のひろば (第 6 回) .....	44
総会報告：	第 19 回通常総会および関連行事の報告.....	54
活動報告：	協会活動報告 (2009 年 6 月～2009 年 10 月).....	56
編集後記：	.....	宇津呂 武仁... 59

# CONTENTS

Foreword:	Where is "blue bird" ? .....	<i>E. Ito</i> ..... 1
In memory of Professor Hozumi TANAKA:	.....	<i>H. Isahara</i> ..... 2
	.....	<i>M. Nagao</i> ..... 3
	.....	<i>J. Tsujii</i> ..... 5
MT Summit Report :	MT Summit XII Report -Focus on the Workshop on Patent Translation-.....	<i>T. Ehara</i> ..... 6
	Report on MT Summit XII .....	<i>S. Miyazawa</i> ..... 8
	Report on MT Summit XII .....	<i>S. Yokoyama</i> ..... 11
	A Report of MT Summit 2009 .....	<i>M. Uchiyama</i> ..... 14
ACL 2009:	Report on ACL-IJCNLP 2009 .....	<i>Canasai Kruengkrai</i> .. 15
	Conference Report Joint conference of the 47 <sup>th</sup> ACL and the 4 <sup>th</sup> IJCNLP .....	..... <i>Wenliang Chen</i> ... 18
	Report on ACL 2009 at Singapore .....	<i>Hailong Cao</i> ..... 20
AAMT Nagao Award:	Reflections on Receiving the Nagao Award 2009 .....	<i>AAMT Internet Working Group</i> ..... 22
Committee Report:	Questionnaire on English-Japanese / Japanese-English MT .....	..... <i>MT Committee for seeking future direction of MT</i> ..... 25
	Common Format for User Dictionary UTX.....	..... <i>MT Committee for seeking future direction of MT</i> ..... 39
AAMT Members:	AAMT Agora.....	44
General Meeting:	Report on 19 <sup>th</sup> AAMT General Meeting and related events .....	54
AAMT Activities	AAMT Activities (June 2009～October 2009) .....	56
Editor's Note:	Message from the Chair of the Editorial Committee.....	<i>T. Utsuro</i> ..... 59

\*連載企画「言葉と人材ーグローバル時代における企業内の言語とコミュニケーション」は都合により休載いたします。

## 青い鳥を探して

東芝ソリューション株式会社

プラットフォームソリューション事業部 伊藤 悦雄

「きわめて非論理的です」。これは、SF「スタートレック」のミスタースポックの有名なセリフなのでご存知の方も多いでしょう。人間の感情に基づく行動が非論理的であることを揶揄するセリフです。その人間の感情や行動を源としているため非論理的であって当然な「言語」を、論理的に処理する「自然言語処理」の困難さは、改めて述べるまでもないでしょう。上記のスタートレックの世界ですら、22世紀になっても不完全な翻訳システムしか存在せず言語学者が調整を加えながら通訳をしています。無論、我々は22世紀まで待つことはできません。それ以前に「使える」機械翻訳システムを実現するために何をすればよいかを考えるのが我々の役目だと考え、日々を過ごしています。散文調になりますが、その日々の考えを記させていただきます。

「使える」機械翻訳システムと書いたのは、「完璧な」機械翻訳システムである必要はないと感じているからです。かな漢字変換が現在も完璧な変換を行っていないのに日々の生活から切っても切れないものになっている。そのような存在になることは機械翻訳システムの一つの中間ゴールではないでしょうか。そのためには、万人の文書を正しく翻訳できることが必要ではなく自分の身の回りにある文章の翻訳ができれば、それは「使える」機械翻訳システムになるのです。どうすれば、「使える」機械翻訳ができるのか、幸せの青い鳥はどこにいるのかを探さなくてはいけないときでしょう。

万人の文書を翻訳するためには、万人に共通する翻訳知識が必要で、北風のように力まかせでルールを書くルールベース方式の翻訳には限界を感じ始めます。一時、翻訳ソフトは「辞書語数」という森の中で光る「語数」というお菓子の家にとらわれてしまい、語数競争に明け暮れたこともありました。しかし、ふと気づき、落としてきた小石を振り返ると、どこで路を誤ったかがわかります。

語数を増やしても自分に関係ない言葉であれば存在しないのと同じであることに、ロバになる前に気づいたのは幸いです。それよりも、自分の身近にある文書の内容を参考にして訳を決めたり、その文書から辞書を作成したりするといった、自分に役立つ辞書をいかに簡単に作成するかがポイントだったのです。また、翻訳する文書の中にちりばめられたヒントを読み取って翻訳する観察眼も重要です。

身近にある文書を正しく翻訳するためには、身近な情報を使用すればよかったのです。機械翻訳の世界でも、「使える」翻訳システムを作るための青い鳥は身近なところにいたのです。

余談ですが、当社の翻訳ソフト「The 翻訳」のパッケージにも青い鳥がかかかれていました。この商品に搭載されている技術によって、第4回 AAMT 長尾賞を受賞することができました。(決して童話を意識したものではなかったのですが、) 私の近くにも青い鳥がいて、私にも幸せを運んでくれたのでした。

## 追悼 田中穂積先生



=====  
大変に悲しいことですが、1993年から1999年までアジア太平洋機械翻訳協会（AAMT）の会長を務められました田中穂積先生が2009年7月27日にお亡くなりになりました。田中先生は1997年にシンガポールで開催された機械翻訳サミットの組織委員長を務められ、1997年から99年までは機械翻訳国際連盟（IAMT）の会長を務められるなど、AAMTに関わる活動にご尽力いただいております。心よりお悔やみ申し上げます。

田中先生はアジアや世界での自然言語処理・機械翻訳の研究の推進に重要な役割を果たされました。また多くのAAMTメンバーにとって、掛け替えのない先輩であり、友人でありました。

このたび、特に田中先生とお親しかった長尾眞先生と辻井潤一先生から、田中穂積先生を偲んでの文章をお寄せいただきました。ここに掲載させていただきます。田中先生、長い間、ありがとうございました。

アジア太平洋機械翻訳協会会長 井佐原 均

=====

## 追悼 田中穂積先生

国立国会図書館

館長 長尾 眞

田中穂積先生が去る 2009 年 7 月 27 日に亡くなった。67 歳であった。先生の突然の、しかもまだまだこれからという年齢での逝去にショックを受けない人はいなかっただろう。聞いた時は信じられなかった。先生の葬儀は 8 月 1 日、お住まいの近くの葬場で多数の参列者に見送られながら行われた。私も悲痛な心境で弔辞を読んだ。

先生は 1941 年生まれ、1964 年東京工業大学工学部情報工学科卒、66 年同修士修了、ただちに当時の通産省電気試験所（現在の産業技術総合研究所）に入られた。そして故淵一博氏のもとでソフトウェア、自然言語処理、音声処理などの研究に従事されたのち、母校工学部に助教授として戻られ、すぐに教授に昇進された。以後 2005 年に定年退職されるまでの 20 年余の間、情報工学科の教授として多くの優れた研究成果をあげるとともに、優秀な人材を数多く輩出された。また教育・研究だけでなく大学行政にも多くの貢献をされた。

東京工業大学をご退職のあとは名古屋にある中京大学の情報科学部認知科学科の教授を 2009 年まで務められた。その後、請われて北陸先端科学技術大学院大学の教授になられ、これから面白い研究をしようとしておられた矢先の逝去であった。その間 2006 年には東京工業大学に新しく作られた総合研究院先進研究機構長に就任しておられる。

先生の御研究は、一口でいえば人間が言語をどのように理解し発話しているかをコンピュータで実現することにあつたといえるだろう。言語の形態素解析と構文解析に種々の工夫をし、効率の良い速い処理方法を考案するとともに、人間がきっと行っているだろうように形態素解析と構文解析とを有機

的に統合して解析を進める方法を開発するなどの努力をされた。

音声研究においても認識とともに合成についても多くの研究の成果を出された。そして音声研究と言語処理研究とを統合して、人とロボットとの間の対話システムの研究開発にまで進まれた。

こういった研究の根底には人間頭脳の働きについての先生の尽きない興味があつたことは確かである。認知科学という概念が出てきた 1980 年代から先生はこの分野に関心を払うとともに種々の形で積極的にかかわってこられた。先生は言語という専門分野のほかこういった学際的な研究に常に関心を持ち、積極的に参加してこられた。

田中穂積先生は学外の活動にも積極的にかかわってこられた。たとえば淵氏がはじめられた第 5 世代コンピュータ研究開発にも参加されたし、国際情報化協力センター（CICC）が企画したアジア諸言語間の機械翻訳プロジェクトにも中心人物の一人として研究を推進された。また日本電子化辞書研究所（EDR）の辞書作成にも種々のアドバイスをし てこられた。

学会活動で先生が最も貢献されたのは言語処理学会と言語資源協会（GSK）の設立であろう。言語処理研究は情報処理学会や電子情報通信学会の研究会などで行われていたが、一つの場に集まって皆が切磋琢磨するのが良いとして、先生と私が中心となって言語処理学会を 1994 年に発足させた。そして第 2 代の会長を務められた。その後デジタルテキストが増加し、またインターネット上に多くのテキストが流通するようになって、欧米でテキストや音声のデータを集めたコーパスが作り始められ、

また辞書や各種の言語処理ツールなども集められて、研究者が自由に使えるような環境が作り始められたので、日本においてもぜひこれを作る必要があるとして、関係者に呼び掛け、言語資源協会を 2003 年に設立され、日本の自然言語処理研究者のために大きな貢献をされた。そのほかには人工知能学会会長、アジア太平洋機械翻訳協会会長なども歴任された。

先生のお人柄は非常に謙虚で、いつも接する人を自然に引き込み、その時々に関心を持っている

事柄、研究について率直にお話をされていた。大学における学生や職場における後輩などに対しては常に寛大な態度で接し、決して「これをしてください」といった強制をすることのない人で、多くの人に慕われる存在であった。

このような優れた人を早く亡くしてしまったことは日本の言語情報処理分野にとって大きな損失である。まことに残念なことで、心からご冥福をお祈りいたします。

## 田中先生を偲ぶ

東京大学／マンチェスター大学

教授 辻井 潤一

田中先生の急逝の報に接し大きな驚きと悲しみを感じています。

田中先生には30年以上の長きにわたって公私ともにご指導いただきました。先生が、言語処理、言語理解、機械翻訳という分野で果たしてこられた指導的な役割、これらの分野で活躍される先生の大勢のお弟子さんなど、先生が亡くなられた今、その存在の大きさをあらためて感じています。

1966年に出たALPAC報告でいったん下火になった機械翻訳が80年代に再び活発化し、日本では、第5世代コンピュータプロジェクトや電子化辞書(EDR)のプロジェクト、アジア諸言語の翻訳を目指したCICCプロジェクトという具合に、言語処理、機械翻訳の研究が急速に拡大していくのですが、その中心に田中先生がいつもおられたという印象があります。また、日本電子工業振興協会(現在の電子情報技術産業協会)では、長尾眞先生(現、国会図書館館長)や企業の機械翻訳研究者の方々と海外視察を企画され、それに一緒させていただいたことを昨日のこのように思い出します。

先生がAAMT会長、私が副会長として1999年のシンガポールMT Summitを組織しました。日本での機械翻訳の開発が一段落していたこともあり、企業からの参加の減少など、スポンサー確保が大変で、海外からの講演者の招待などにかかる費用の工面に先生と奔走しました。国際会議を、日本ではなくシンガポールという遠隔の地で開催するということで、先生と一緒にご苦労したのも、今ではよい思い出になっています。

また、私事ですが、私の母親がやはり64歳の

若さで亡くなりました。田中先生はそのことを折にふれ私に「若くして亡くなられて残念だね」といってくださっていました。母が亡くなってのち、7年間英国に滞在して帰国しましたから、母の死からはもう随分と時が経っていた、帰国してしばらくの頃、「辻井君も随分といろんなことを経験して、いま、お母さんが生きておられたら、たくましくなったと喜ばれると思う」といってくださったのを、鮮明に覚えています。その先生が、こんなに早くに亡くなられるとは思っていませんでした。

先生の急逝の報を海外の知己に知らせたのですが、口々に先生の温かい人柄を語り、その急逝を残念がっていました。その中に、「先生の温かい微笑みは、われわれの心からなくなることはないでしょう」というメッセージがありました。また、*Computational Linguistics* という論文誌に、田中先生の教え子である徳永・東工大教授、T. Baldwin博士(メルボルン大学)と一緒に追悼文を書きました。その中で、Baldwin博士が「先生は学生の個人的悩みにも真剣に耳を傾けてくださり、留学生にとって心の休まる避難所であった」といった意味のことを書いていました。先生の人柄を示すものだと思います。

本協会をはじめとする学術界への先生の貢献はわざわざ申すまでもないかと思います。私にとってはそれにもまして個人的な心の師であるとともに、僭越な言い方ですが、機械翻訳研究の同志である親友を失ってしまった、という無念の気持ちでいっぱいです。

先生のご冥福、心よりお祈りいたします。

## MT Summit XII 参加報告 —特許翻訳ワークショップを中心に—

山梨英和大学 人間文化学部 人間文化学科

教授 江原 暉将

### 1. はじめに

機械翻訳関係の国際学会である MT Summit XII に参加したので、主観を交えて報告する。なお、この会議の論文や一部の発表資料は、ワークショップも含めて Machine Translation Archive に掲載されている。

会議は 8 月 26 日から 30 日の 4 日間カナダの首都オタワの Hotel Château Laurier で開かれた。26 日は 6 件のチュートリアルが行われ、本会議は 27 日から 3 日間行われた。5 件のワークショップは会議期間中に分散して行われた。Technology Showcase が 27 日に行われた。

大会委員長の Laurie Gerber 氏(写真 1)によると、投稿論文数は 89 件、採択は 48 件(内ポスター 27 件)であった。事前登録者数は 278 名であり、実参加者は 300 名を越えた模様である。南北アメリカから 166 名、ヨーロッパから 74 名、アジアからは 31 名の参加があった(いずれも事前登録者数)。



写真 1 大会委員長の Laurie Gerber 氏

### 2. 本会議

本会議に関する概要は、招待講演の一つである Pierre Isabelle and Roland Kuhn : MT: The current research landscape にスマートにまとめられており、上記 MT Archive に掲載されているので、そちらを参照していただくとして、本文では筆者の印象を中心に述べる。

一般講演は MT Research、Government Users、

Commercial Users の 3 つのトラックに分かれて行われた場合が多い。また、5 件の招待講演と 3 件のパネルが行われ、ポスターセッションが休憩をはさんで 2 回行われた。

全体を通じて、あいかわらず統計方式の発表が多いものの、規則方式の MT に自動後編集を付加する手法が印象に残った。MT を実用にする場合、前編集または後編集が必要な場合が多いが、前編集では訳の改善が思うようにできず、いきおい後編集に頼ることが多い。しかし、同じような後編集を繰り返すのは翻訳者にとって苦痛であり、自動後編集が望まれている。自動後編集の手法としては、統計的機械翻訳の手法を応用するもの(SPE: Statistical Post Edit)や、後編集のための規則を自動学習する方法などがあつた。

Government Users のトラックでは、各国の政府機関での MT の利用状況について報告があり、特許庁の番井進氏から日本の特許庁での MT の利用や世界の特許庁の間での MT の相互利用について発表があつた。

### 3. 特許翻訳ワークショップ

筆者らが主催した The 3rd workshop on patent translation が大会最終日の 30 日に行われた(写真 2)。本 workshop は AAMT/Japio 特許翻訳研究会の支援を受けている。2 件の招待講演、5 件の技術発表、3 件のユーザー報告、1 件のパネル討論が行われた。日本、韓国、中国、欧州、米国、カナダの特許庁あるいは類似機関からの参加があり、活発に議論された。参加者は事前登録で 26 名、実際には 40 名くらいであった。

特許文書は膨大であり、しかも各国が自国語で運用している。科学技術文献のように英語が標準言語というわけではない。ここに特許分野での MT 利用の背景がある。実際、特許文献の調査や各国特許庁での出願審査に MT が利用されている。このような状況の中で、更なる MT の技術進歩と利用促進を目的として本 workshop が開かれた。Phuket、Copenhagen に続いて第 3 回となる。



写真2 特許翻訳ワークショップ

招待講演の一つは、EPO (European Patent Office)の Sophie Mangin 氏からで、EUにおける特許流通と機械翻訳の利用について報告した。26言語が翻訳に関係し、特に、英語、仏語、独語の3つの公式言語が重要である。また15カ国が自らの国語での特許情報へのアクセスを望んでいる。それらを支援するのがEPOでのMTの役割である。現在は、英語をpivotとしているが、将来は英仏独の3言語をpivotとして考えたいと述べていた。日中韓の3言語から英語へのMTについても触れた。

もう一つの招待講演は、SIPO (State Intellectual Property Office of China)の Dan Wang 氏からで、SIPOにおける中英MTの利用と改良について報告した。上記MTは中国の特許を英語で検索するためと、特許文献のfull text翻訳に利用されている。MT方式は意味を利用したtransfer方式であり、文を57種類に分類している。また、辞書の充実に統計方式MTの技術を応用している。

KIPO (Korean Intellectual Property Office)からは、KIPOでのMT開発の経緯と技術的特長およびIP5(日韓中欧米の5極特許庁)での情報交換のためのMTの開発計画について報告された。前者についてはSIPOと同様、文パターンの利用が特徴である。特許文は文形が類似しているので文パターンの利用が有効であろう。後者については、5極でMTに関する情報を共有しMTのノウハウを蓄積するとともに、翻訳品質を向上させる計画とのことである。

JPO (Japan Patent Office)からは、JPOでのMT利用と改良について報告された。一般ユーザー向けのIPDL (Industrial Property Digital Library)と

海外の特許庁職員向けのAIPN (Advanced Industrial Property Network)での日英MT利用について述べた。また将来計画として、特許文献と科学技術論文を統合した多言語の言語横断情報検索などについて触れた。

その他、USPTO (The United States Patent and Trademark Office)から他庁のMTサービスの利用状況について報告があり、CIPO (Canadian Patent Office)からも簡単なコメントがあった。

一般講演では、(1) NTCIR-7での特許文書翻訳コンテスト、(2) MT出力の自動評価基準のメタ評価、(3) 信頼性尺度を用いた中英特許平行コーパスの高精度化、(4) 1ドキュメントからでも抽出できることを特徴とする専門用語の抽出ツール、(5) 格フレームを用いた特許文書向けの翻訳あいまい性の解消、の5件の発表があった。

最後に、ユーザーに研究者・開発者を加えたReal World Challenges of Patent Translationと題したパネルが行われた。ここでは、以下のテーマで議論を行なった。(1) MT利用の現状、(2) MT利用の利点、(3) 今後のMT利用計画、(4) MT関連のツール、(5) 将来展望。この中で研究者側から参加したPhilipp Koehn氏は、特許文書の定型性に着目し、機械学習の応用の可能性を示唆した。また、特許は多言語平行コーパスとして規模が大きく、研究者としても魅力ある分野であると述べていた。

#### 4. おわりに

特許翻訳ワークショップを中心にしてMT Summit XIIの状況を解説した。データベース、コンテンツ、知財などと呼ばれる知識インフラが、ますます重要視される中で、機械翻訳の役割も広がって行くであろう。その中心分野の一つに特許があることは間違いない。次回のMT Summitは、2011年に中国の厦門(Xiamen、アモイ)で開かれる。その時には、特許翻訳の分野でも技術と利用が一段と進んでいることと思う。

#### 参考文献(サイト)

- [1] Machine Translation Archive,  
<http://www.mt-archive.info/>
- [2] Japio Year Book,  
<http://www.japio.or.jp/00yearbook/index.html>

## MT Summit XII 参加報告

秀明大学 英語情報マネジメント学部  
教授 宮澤 信一郎

## はじめに

12回目のMT Summitが8月26日から30日までカナダのオタワでAMTA（アメリカ機械翻訳協会）の主催で開催された。私の主たる目的は、前々回のブーケットでのMT Summit Xで初めて開催され今回3回目となる、日本が中心となって開催する特許翻訳のワークショップ「3rd Workshop on Patent Translation」への参加である。私の主な役割は、ワークショップのプロシーディングやホームページの作成・編集である。

オタワは日本からの直行便は無く、トロントやモントリオールなどに比べると日本人に馴染みが薄い都市ではないかと思うが、れっきとした首都であり、美しい街並みが続く街である。人口は約81万人、そのため全体的にゆったりした感じをうけた。運河や河川が多く水上交通が盛んで、会場と宿泊場所になったホテルに隣接する運河では、パナマ運河の小型版の閘門（水のエレベータの役割の水門）で水位を調節して、小型船を行き来させる光景がめずらしかった。前回のMT Summitの開催地であったコペンハーゲンと同様に、車は日中でもエンジンを始動させると自動的に点灯する灯火が義務付けられていた。全車が点灯していてもさほどまぶしくない。EU域内では、2011年以降販売される自動車に「昼用ヘッドライト」の設置を義務付けることが決定している。対向車や歩行者が車の接近を認識しやすく、死亡事故を3-5%減少する効果があるとされている。死亡に至らない事故も含めると相当大きな効果があるのではないかと痛感させられた。

## 会議概要

会場はHotel Chateau Laurierである。国会議事堂

に隣接しており、エリザベス女王の定宿にもなっている100年ほどの歴史を有したシャトー型のホテルである。創設者はタイタニックの事故でホテル開業直前に遭難死しており、ホテルの名前は代わりに開業の式典を行った時の首相の名前に由来するそうである。タイタニックは海に浮かぶホテルと言われたが、何か因縁めいた話である。参加者はアメリカ地域を中心に世界各国から約300名であった（日本からは15名）。前回のコペンハーゲンでの参加者は約270名であったが、今回は、AAMT主催であるので参加者が300を越えてもらいたいものである。26日には6つのチュートリアルが開かれたが、私は26日にオタワに到着したので、参加していない。17:30からはレセプションが行われたが、ホテル到着が夕刻になり、後半しか参加できなかったのは残念であった。



写真1 会場風景

27日は、Welcome - Conference Overview に続いて、招待講演が3件あった。始めはウィルス対策ソフトウェアSymantecのトップの招待講演「Deploying novel MT technology to raise the bar for quality at Symantec: Key advantages and challenges」であり、

Symantec の各国のユーザに向けた対応を迅速かつ低コストで行うために MT を採用し推進していることの紹介である。次が翻訳会社 Translated.net の CEO による講演「Getting a share of the human translation market with the world's largest Translation Memory」で、Web を通じて多くの翻訳者を使い、多数のユーザに、多言語で大量の翻訳を提供しており、これらのサービスを行うために高速でメンテナンスが容易な TM（翻訳メモリー）の構築を行っていることが紹介された。最後に主催者側とカナダの NRC の研究者による講演「MT: The Current Research Landscape」であり、現在の MT の研究の全体像と課題についての解説である。その後「Research Papers」、「Government Users」、「Commercial Users」の 3 つのトラックに分かれて一般発表が行われた。今回の MT Summit の特徴は「Government Users」のトラックが独立して設けられたことである。これは政府や国による MT の需要が増している現れと思われる。全期間を通じて、一般発表は、Research Papers 21 件、Government Users 8 件、Commercial Users 9 件、Special Session 6 件、総計 44 件の発表があった。

28 日は一般発表の他に、全体のパネルディスカッションが開かれた。これは「Converging Technologies: What are the benefits for MT users? 」というテーマで行われ、MT といろいろな技術の結合が行われているが、それがユーザにとってどのような利益をもたらすかという内容であった。午後からは、Special Session と並行して、27 件のポスターセッションと 30 にのぼる企業や機関による Technology Showcase が開かれた。



写真2 ポスターセッション

夜はバンケットが近くの博物館 Museum of Civilization のレストランを借りて行われた。川沿いに歩いて十数分の距離にあり、往復の途上で三々五々、知り合い同士が談笑しながら、気持ちの良い夜風に吹かれて歩いたのは愉快であった。

29 日は「Research Papers」、「Government Users」の 2 つのトラックが開かれ、同時にワークショップ「Beyond Translation Memories」が行われた。途中、全体の招待講演「What is the Meaning of Reaching 94% Success?」、「Can You Score Higher with MT?」の 2 件が行われた。最後に Closing Session が開かれ、AAMT 会長の井佐原均氏が IAMT 会長となり、次回の MT Summit が 2011 年に中国のアモイで開催されることが発表された。

30 日には 2 つのワークショップが並行して開催された（会期を通じてでは 5 つのワークショップ）。私は「The 3rd Workshop on Patent Translation」に参加していたので、他のワークショップの様子は分からない。

このワークショップでは前回、前々回に引き続いて **Proceeding** の編集作成を担当した。表紙と印刷は **AMTA** 側が行ってくれるので、何度か **AMTA** の担当者とやり取りを行って仕上げについて注文をつけた。そのため現地に行くまでどのような仕上がりになっているか心配であったが、綺麗な仕上げになっていて安心した。現地スタッフに感謝したい。ワークショップには約 40 名の参加者があり会場はほぼ満杯でこの分野への関心の深さがうかがわれた。**Chair** は山梨英和大学の江原先生と山形大学の横山先生が共同でつとめた。午前中は江原先生の開会挨拶に続き、**Japio** (財団法人 日本特許情報機構) 理事長からの挨拶、招待講演 2 件、一般発表 3 件が行われ、午後はユーザからの講演 3 件、一般発表 2 件が行われた。その後、パネルディスカッション「Real World Challenges of Patent

Translation」が 4 カ国特許庁と大学 1 名、司会の 6 名のパネリストでディスカッションや質疑応答が行われた。ワークショップは盛況に終了し安堵した。次回の **MT Summit** は中国アモイになったが、中国における知的所有権の課題の上でも引き続き特許翻訳のワークショップを開催する意義は十分あると感じた。

#### おわりに

この会議で機械翻訳の研究の最前線や問題点、ユーザのニーズなどが明らかになり、世界の機械翻訳研究者、開発者、ユーザが一同に会して情報を共有しあつたことは大変意義深い。今後も **MT Summit** がより発展して機械翻訳の発展に大きく貢献することを願っている。最後になるが、いろいろ御支援いただいた **AAMT** や **Japio** の方々に感謝申し上げる次第です。

## 第12回機械翻訳国際会議 (Machine Translation Summit XII) および

## 第3回特許機械翻訳ワークショップ (The 3rd Workshop on Patent Translation) 参加報告

山形大学

教授 横山 晶一

## 1. はじめに

第12回機械翻訳国際会議 (Machine Translation Summit XII、以下 MT Summit XII と略称) は、2009年8月27～29日、カナダの首都オタワの、Hotel Chateau Laurier で開催された。会場は、オタワでも最も伝統ある、首都の中心に位置する文字通り城のような外観のホテルで、ホテル名は、タイタニックで遭難死した設立者に代わって開設式典を行った、当時のカナダ首相の名にちなむ。MT Summit にはこれで3回連続の出席になるが、私個人としては AMTA 内での会議への出席は初めてである。また、前回のコペンハーゲンの会議では、AAMT への報告書ははやばやと書いたのに、ジャーナルへの投稿をぼんやりして忘れていたので、久しぶりの報告となる。

参加者は pre-registration の段階 (8月17日現在) で、27の国から279名で、この大会の主催者である Laurie Gerber の話では、最終的には300名くらいであったということである。場所的な面から、アメリカ地域からの参加者が最も多く166名。うち118名がアメリカから、48名がカナダからであった。アジアからは31名で、日本が15名 (私が知る限りでは、AAMT の井原会長、神崎亨子事務局長、AAMT/Japio 特許翻訳研究会から江原暉将山梨英和大教授、宮澤信一郎秀明大教授と私、番井進氏 (特許庁)、Japio から林昭彦理事長、渡邊豊英部長、柿田剛史係長、潮田明氏 (富士通研)、奥村明俊氏 (NEC)、釜谷聡史氏 (東芝)、内山将夫氏 (NICT)、山本和英氏 (長岡技科大) と学生、山形大の私の研究室の学生)、中国が4名、香港が1名となっている。ヨーロッパからは、アイルラン

ドからの16名を筆頭として、74名が参加した。5つの招待講演 (英、伊、加、米2) と3つのパネル、一般セッションは89編の応募論文のうち48編が採択され、うち21件が通常の発表、27件がポスターでの発表であった。

また、これと前後して同じホテルで6つのチュートリアル (26日)、5つのワークショップ (本会議前、会議中、会議後) が開催された。筆者らが参加したのは、そのうちの第3回特許翻訳ワークショップ (The 3rd Workshop on Patent Translation) (共同座長江原暉将山梨英和大教授、横山晶一) である。このワークショップは、日本、デンマーク、カナダ、アメリカ、オランダ、ドイツ、中国、韓国などから pre-registration 段階で37名の参加者があり、招待講演2つと、ユーザからの講演3つ、Japio の林昭彦理事長からの挨拶、一般発表5つ、パネルディスカッション (Moderator: 潮田明氏 (富士通研)) が行われた。

本会議では、前日に歓迎晩餐会、2日目にバンケットも開かれ、各講演では活発な質問が行われた。以下に本会議、ワークショップの概要を報告する。

## 2. 本会議の概要と特徴

今回の本会議の特徴は、研究 (MT Research)、政府機関による取り組み (Government Users)、商用 (Commercial Users) という3つの並行セッションが設けられたことである。ごく初期の MT Summit が、やや政策的な会合であった当時に戻ったという印象を受けた。

大会初日の幕開けには、本会議開催の主催者である

Laurie Gerber による Welcome and Conference Overview が行われた。講演の途中で、井佐原 AAMT 会長が、7 月下旬に逝去した田中穂積元会長 (IAMT やシンガポールの MT Summit VII の主催者でもあった) の追悼を述べた。かつての電総研において、田中穂積氏の下で井佐原氏とともに研究した筆者としては、井佐原氏の個人的な思いがよく理解でき、なかなか感動的な心温まるスピーチだったと思う。

招待講演では、Symantec の Johann Roturier 氏が、Systran との共同プロジェクトとして、Post Editing を統計的に行う試みや、きちんとした TM を作る試みを紹介した。現在、日、中、仏、独、伊の言語でこれらの試みが進行中であるということも述べた。また、Translated.net の CEO である Marco Trombetti 氏は、多量の翻訳サービスを行うために高速でメンテナンスが容易な 10 億語の TM を目指していることを講演した。さらに、主催者側、カナダ NRC の Pierre Isabelle 氏と Roland Kuhn 氏の講演では、現在の MT の研究と、本会議の発表 (ポスターも含む) の概説を結びつけて要領よく解説していた。この講演のみを聞くと、全体像が分かるという大変お得なものである。彼らが特に強調したのは、ユーザの要求と現在の研究とのギャップで、その要素として、7つをあげていた。

パネルディスカッションは、全体のパネルとして、Covering Technologies: What are the benefits for MT users? というものが行われた。内容は、RB + SMT, TM + cloud computing, advanced leveraging + shared TM data, speech recognition + MT など、いろいろな技術の結合が行われているが、それがユーザにとって、質や性能、使い勝手、経済性、安全性などからどのような利益をもたらすかというものであった。

もう一つのパネルで、special session として設けられた Preparing Translators for the current technology landscape は、4 名の女性パネリストによる「翻訳者を志す文系の学生にどうやって技術的なことを教えたらいいか」という面白いパネルであった。参加者はほんの数名であったが、コンピュータに素養

のない学生への教育というテーマで議論が行われた。

一般講演は、上記のように、主として、研究 (MT Research)、政府機関による取り組み (Government Users)、商用 (Commercial Users) という 3 つの並行セッションに分かれて進められた。また、ポスターセッションはオープンスペースで 2 回、それぞれ 2 時間にわたって行われた。私個人の印象としては、後編集に統計的手法を使う研究、多言語システム、マイナーな言語の扱い、RBMT, SMT, EBMT のいくつかを融合する試みなどが多かったと思う。また、政府機関では、軍による安全保障に MT を使う試みが印象に残った。

### 3. 第 3 回特許翻訳ワークショップ

このワークショップは、AAMT/Japio 特許翻訳研究会 (委員長: 辻井潤一) における種々の議論をもとに、MT Summit でワークショップを開いて、この分野の研究者の意見交換ができればよいという趣旨の下で開催された。第 1 回は Phuket (Chair: 横山晶一)、第 2 回は Copenhagen (Co-Chair: 辻井潤一、横山晶一) である。

今回は第 3 回のワークショップで、Chair は江原暉将、横山晶一が共同でつとめた。今回は文字通り共同座長であったので、前 2 回に比べると個人的には非常に楽であった。参加者は約 40 名で、そのうち 2 名が招待講演者、また、user report やパネリストも招待の形をとった。前回の主催者である Bente Maegaard 氏、今回の主催者である Laurie Gerber 氏も参加された。以下に概要を簡単に述べる。

招待講演 (司会: 江原暉将) では、江原氏によるイントロと、Japio 林昭彦理事長の挨拶の後、EPO (The European Patent Office) の Sophie Mangin による活動内容の紹介があった。EPO においては、特許に用いられる基本言語として、英語、フランス語、ドイツ語の 3 つがあり、esp@cenet というサービス (日本の IPDL に相当する特許 DB サービス) の一部として、いくつかのヨーロッパの言語との MT が試みられている。ヨーロッパ以外では、米、日、韓、中による IP5

cooperation という協力の枠組みが進んでおり、この下に MT を含む 10 の foundation が組織されている。この枠組みが将来に向けての情報交換や MT の中心になることが期待されている。

もう一つの招待講演は、中国 SIPO (The State Intellectual Property Office of P.R.C) の王丹氏による講演で、2005 年 4 月に立ち上げ、2007 年 4 月から Single-user version、2008 年 4 月から改良テストを行っている CPMT (China Patent Machine Translation) のプロジェクトについて述べた。興味深いのは、RBMT において、Hierarchical Network of Concepts (HNC) の中で、文を 57 カテゴリに分けて、性能を上げているということであった。なお、RBMT, SMT, EBMT などの統合も将来的には試みる予定である。ただ、王氏は、ここでパーナード・ショーの有名なエピソード (ある女性が彼に言い寄り、自分と結婚したら彼の知性と自分の美貌を兼ね備えた子供が生まれると言ったところ、ショーが逆なら大変と言って断った) を引用してユーモラスに締めくくったのが印象的であった。

一般講演は午前と午後 2 つ (午前の司会: Svetlana Sheremetyeva、午後の司会: 梶博行) あり、合計 5 件の発表が行われた。江原氏や筆者が発表した、詳細は略す。

ユーザによる講演 (司会: 横山晶一) では、アメリカ (Arti Shah 氏)、韓国 (Young Pyo Kim 氏)、日本 (番井進氏) による発表があり、各国特許庁の現状と計画が紹介された。最後に潮田明氏がモデレーターとなったパネル (各国から 5 人のパネリスト) があり、各国特許庁における特許の MT の利用計画などを討論

した。カナダの特許庁の人間も会場から参加し、各国特許庁が会する貴重な機会となった。

最後に横山が締めくくりの挨拶を行い、第 4 回のワークショップも是非行いたい旨を述べた。

#### 4. おわりに

前回に比べると、本会議では、中国語やアラビア語に関する発表がやや減り、研究面でも SMT と RBMT や EBMT との融合を図る動きが本格化した印象がある。本会議も、ワークショップも、現在の機械翻訳のかかえる問題点が明確になり、大変有意義であった。機械翻訳関係の研究者、ユーザ、メーカー、政府関係者等が世界レベルで一同に会する機会はこの会議以外には非常に少ない。そのためにも今後ともこの会議の継続と発展を祈っている。ワークショップをさらに発展させ、情報を共有しようという気運が盛り上がりつつあるのは心強い。特に今回、政府サイドの発表が多かったことは、安全保障などに MT を利用する機運が高まっていることを示している。日本政府からは特許庁以外に発表がなかったのは残念である。

会議の内容は、2 で述べた Isabelle, Kuhn の概要も含めて MT archive に載せる予定とのことであるし、ワークショップも含め、ビデオ撮影を行っていて、これも利用できるとのことである。

なお、次回の会議は、AAMT 会長の井佐原均氏が IAMT 会長となって、中国のアモイで 2011 年に開催の予定である。第 4 回のワークショップも、今後の議論の中で詰めてゆきたい。秋の気配のただよう気持ちのよいオタワで開催された会議の関係者に感謝する。

## MT Summit 2009 参加報告

独立行政法人 情報通信研究機構

内山 将夫

MT Summit 2009は、300人以上の参加者による盛大なものでした。参加者は、大学、企業、政府からまんべんなく来ており、機械翻訳の研究が盛んなことを示していました。

バンケットのときに、企業参加者の一人から、参加理由として、米国政府が大いなる顧客のため、政府にアピールするためにも参加しているという話を聞きました。

これに関してですが、日本政府においても、翻訳に関する需要は相当多いはずですから、政府が機械翻訳に関して、もっと予算を割けばよいのと思いました。世界的には、翻訳が再びブームとなっていますが、日本では、翻訳のブームは再来していないように感じます。

このところの機械翻訳の進歩は急激ですので、この分野にもっと予算と人員を振り向ける必要があるように思います。

さて、会議の内容ですが、私は主にテクニカルセッションに出ていたもので、そこで聴講した講演のいくつかについて書きたいと思います。

まず、フレーズベースのSMTでは、フレーズテーブルが非常に重要な言語情報ですので、これを改良する研究として以下の2つがありました。(1) Complexity-Based Phrase-Table Filtering for Statistical Machine Translation. (2) Phrase Translation Model Enhanced With Association based Features. (1)は、ある種の統計的な尺度により、テーブル中のフレーズをフィルタリングするという話です。

(2)は、色々な統計的な尺度をフレーズのスコアとして追加することにより、翻訳精度が少し向上するというものです。

コーパスに関連しては、United Nations General Assembly Resolutions: A Six-Language Parallel Corpus がありました。これは、国連文書のパラレルコーパスで、<http://www.uncorpora.org/>より利用できます。また、Development of a Japanese-English Software Manual Parallel Corpus は、日英のオープンソースソフトウェアのマニュアルから作られた対訳コーパスです。これは<http://mastarpj.nict.go.jp/~mutiyama/manual/index.html> から利用可能です。

翻訳支援にMTを利用するものとしては、Interactive Assistance to Human Translators using SMT methods や Phrase-based MT in a Computer-assisted Translation environment がありました。

MTを組み込んでいない翻訳支援としては、Hosting Volunteer Translators がありました。この論文では「みんなの翻訳」(<http://trans-aid.jp/>)という翻訳者支援サイトについて報告しています。

そのほか、併設のThe 3rd Workshop on Patent Translation では、各国の特許庁における機械翻訳に関する取り組みの話をお聞きしました。使われている機械翻訳システムはルールベースが主なようです。コーパスベースの機械翻訳システムが特許翻訳に使われるのはまだ先のことのようです。簡単ですが、以上で、報告を終わります。

---

## Report on ACL-IJCNLP 2009

Canasai Kruengkrai, NICT

---

ACL-IJCNLP 2009 is the first joint conference between the Association for Computational Linguistics (ACL) and the International Joint Conference on Natural Language Processing (IJCNLP) of Asian Federation of Natural Language Processing (AFNLP), held on 3-5 August 2009 in Singapore [1]. The conference covers a broad range of technical areas related to natural language and computation. This paper briefly highlights particular papers, published in proceedings of ACL-IJCNLP 2009 [2], that hopefully reflect recent trends in NLP research.

We start with an interesting research direction that focuses on exploiting *bilingual information* to improve performance in various areas. Zhao *et al.* in *Cross Language Dependency Parsing using a Bilingual Lexicon* propose a method to enhance dependency parsing in one language by using a translated treebank from another language. A bilingual lexicon and a simple machine translation technique are used to generate the translated treebank. Resulting word pairs with dependency relations in the translated treebank are then used as additional features.

Snyder *et al.* in *Unsupervised Multilingual Grammar Induction* assume that ambiguous syntactic structures in one language may correspond to less uncertain structures in the other language. They propose an unsupervised Bayesian model for learning bilingual syntactic structure using parallel corpora. Oh *et al.* in *Bilingual Co-Training for Monolingual Hyponymy-Relation Acquisition* propose a bilingual co-training framework that makes use of machine-translated training data from another language. Their assumption is similar to that of Snyder *et al.* in which the reliable part in one language may be overlapped by an unreliable part in another language. They present the framework to enlarge the training data by translating only the reliable part of classification results in another language. Ganchev *et al.* in *Dependency Grammar Induction via Bitext Projection Constraints* propose a learning method for transferring dependency parses across bitext. Their idea is to use word-level alignments and a source language parser to constrain the space of possible target trees.

Wan in *Co-Training for Cross-Lingual Sentiment Classification* proposes a co-training approach to improve the classification accuracy of polarity identification of Chinese product reviews. Machine translation services are used to translate English training reviews into Chinese reviews, and translate Chinese test reviews and additional unlabeled reviews into English reviews. Thus, the problem can be viewed in two independent views where the co-training approach can exploit the two redundant views of features. Gao *et al.* in *Exploiting Bilingual Information to Improve Web Search*

propose to use bilingual queries for learning a linear ranking function. The bilingual queries, which appear in query logs of different languages but represent equivalent search interests, can be used to generate bilingual document pairs for improving ranking.

Next, we explore another research direction related to applying semi-supervised and unsupervised machine learning in NLP. Liang *et al.* in *Learning Semantic Correspondences with Less Supervision* propose a probabilistic generative model that simultaneously segments the text into utterances and maps each utterance to a meaning representation. The Expectation Maximization (EM) algorithm is used for parameter estimation. Persing and Ng in *Semi-Supervised Cause Identification from Aviation Safety Reports* introduce a cause identification system that identifies causes explaining why the aviation incident described in a given report occurred. They present a bootstrapping algorithm that automatically augments a training set by learning from a small amount of labeled data and a large amount of unlabeled data.

Mochihashi *et. al* in *Bayesian Unsupervised Word Segmentation with Nested Pitman-Yor Language Modeling* propose a model for unsupervised word segmentation. Their model is based on a nested hierarchical Pitman-Yor language model where the Pitman-Yor spelling model is embedded in the word model. Ravi and Knight in *Minimized Models for Unsupervised Part-of-Speech Tagging* propose an unsupervised learning method for POS tagging. They apply integer programming to explicitly search for the smallest model that explains the data and then use the EM algorithm to fit the model to the data.

Then, we move on to using the Web as a linguistic resource. Sauper and Barzilay in *Automatically Generating Wikipedia Articles: A Structure-Aware Approach* propose a method for creating a multi-paragraph overview article by selecting relevant material from the web and organizing it into a single coherent text. Shnarch *et al.* in *Extracting Lexical Reference Rules from Wikipedia* present construction of a large-scale resource of lexical reference rules from Wikipedia. Mausam *et al.* in *Compiling a Massive, Multilingual Dictionary via Probabilistic Inference* propose a probabilistic method, based on graph sampling, for automatically constructing a multilingual dictionary.

Finally, Best Paper Awards are given to the following papers: *Concise Integer Linear Programming Formulations for Dependency Parsing* by Martins *et al.*, *Reinforcement Learning for Mapping Instructions to Actions* by Branavan *et al.*, and *K-Best A\* Parsing* by Pauls and Klein.

In conclusion, this paper presents some recent NLP studies appearing in ACL-IJCNLP 2009. We emphasize on three research directions: 1) exploiting bilingual information, 2) applying semi-supervised and unsupervised machine learning, and 3) using the Web as a linguistic resource. The next ACL conference will take place on July 11–16 2010 at Uppsala University Campus in

Sweden.

## References

- [1] The Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP (ACL-IJCNLP 2009), <http://www.acl-ijcnlp-2009.org>.
- [2] Proceedings of ACL-IJCNLP 2009, <http://www.aclweb.org/anthology/P/P09>.

## Conference Report

### Joint conference of the 47<sup>th</sup> ACL and the 4<sup>th</sup> IJCNLP

Wenliang Chen, NICT

ACL-IJCNLP 2009 was held at the Suntec Singapore International Convention & Exhibition Centre, Singapore from August 2 to 7. It was collocated with the Conference on Empirical Methods in Natural Language Processing (EMNLP) 2009. There were 6 tutorials and 12 workshops besides main conference and EMNLP.

#### 1. Papers

In this year, main conference included full papers and short papers. There were 925 total valid paper submissions, a record number of submissions, from more than 40 countries. 21%(120) of full paper submissions were accepted and 26%(93) of short paper submissions were accepted.

Presentations were divided into 40 sessions, normally 4 run in parallel. There were 3 “syntax and parsing” sessions, which I was interested in, including 14 papers. Among them, there were 3 papers that tried to use bilingual data to improve syntactic parsing.

Zhao et al., 2009 presented an approach of using English Treebank for Chinese parsing. First, they translated English Treebank into Chinese by a simple word-by-word translation model. Then new features were represented based on the translated treebank for a shift-reduce parser. The experimental results showed that their method was quite useful.

Ganchev et al., 2009 presented an approach of using bitext projection constraints to induce dependency grammar. They used a bilingual-corpus (E-x) in which the English-side sentences were parsed by a trained parser and performed word alignment. Then they obtained auto-annotated trees for language x. Finally, they trained a parser by adding some specific rules for x.

Snyder et al., 2009 proposed a method to use bilingual cues to perform unsupervised constituency parsing. They adapted an unordered tree alignment to the probabilistic setting for the models. The models loosely bound parallel trees while allowing language-specific syntactic structure. They performed inference under the model using MCMC and dynamic programming.

#### 2. Tutorials

I attended a tutorial “Topics in Statistical Machine Translation” lectured by Kevin Knight, and Philipp Koehn. This tutorial went more deeply into selected topics of intense current interest. They introduced the algorithms of efficient decoding for phrase-based and syntax-based MT, Minimum-Bayes risk, and system combinations. How to scale to large data is an important issue in MT. For this, they listed the efficient methods for phrase table pruning, storage, suffix

arrays, and large language models including distributed LMs and noisy LMs. In recent years, new models for word alignment, rule selection, and syntax-based language models were proposed by many researchers.

### 3. EMNLP

In this year, EMNLP also received a new record number of submissions: 475. There were 163 papers accepted in total. Of these, 96 were accepted for oral presentations, and 67 for poster presentations. Presentations were scheduled in 29 sessions. There were 2 oral sessions and poster session related to parsing, including 11 papers. 2 papers proposed the methods to use bilingual data to improve syntactic parsing.

Huang et al., 2009 presented a method to improve dependency parsing by using a Chinese treebank with its human translation. They performed word alignment first. And then represented new features based on the bilingual constraints. The results showed that their method was simple yet effective.

Smith and Eisner, 2009 proposed an approach by using a parallel corpus to improve parsing. The English-side sentences were parsed by a trained parser and word alignment was performed. Then they obtained auto-annotated trees for another language. Finally, they trained a parser by using Quasi-Synchronous Grammar features.

### 4. More information

More Information can be found at the conference website: <http://www.acl-ijcnlp-2009.org/>. It contains the full program and other information.

### References

1. Hai Zhao, Yan Song, Chunyu Kit and Guodong Zhou. 2009. Cross Language Dependency Parsing using a Bilingual Lexicon, In Proceedings of ACL-IJCNLP.
2. Kuzman Ganchev, Jennifer Gillenwater and Ben Taskar. 2009. Dependency Grammar Induction via Bitext Projection Constraints, In Proceedings of ACL-IJCNLP.
3. Benjamin Snyder, Tahira Naseem and Regina Barzilay. 2009. Unsupervised Multilingual Grammar Induction, In Proceedings of ACL-IJCNLP.
4. Liang Huang, Wenbin Jiang and Qun Liu. 2009. Bilingually-Constrained (Monolingual) Shift-Reduce Parsing, In Proceedings of EMNLP.
5. David A. Smith and Jason Eisner. 2009. Parser Adaptation and Projection with Quasi-Synchronous Grammar Features, In Proceedings of EMNLP.

---

## Report on ACL 2009 at Singapore

Hailong Cao, NICT

---

### Outline of conference

The joint conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing (ACL-IJCNLP 2009) took place on August 2-7th, 2009 in Singapore, the culturally and technologically vibrant garden city of Asia.

Tutorials and workshops were scheduled on 2 and 7 August, respectively. The main conference was held from 3 to 6. Presentations were divided into four streams, normally run in parallel: 120 full papers (oral), 42 short papers (oral), 51 short papers (poster), 5 student research workshop papers (oral), and 7 student research workshop papers (poster). There are also two invited talks:

- Qiang Yang: Heterogeneous Transfer Learning with Real-world Applications
- Bonnie Webber: Discourse - Early problems, current successes, future challenges

In the closing session, Professor Frederick Jelinek of Johns Hopkins University was presented with the lifetime achievement award. And three papers received Best Paper Awards in ACL-IJCNLP 2009:

- Concise Integer Linear Programming Formulations for Dependency Parsing  
Andre Martins, Noah Smith and Eric Xing
- Reinforcement Learning for Mapping Instructions to Actions  
S.R.K. Branavan, Harr Chen, Luke Zettlemoyer and Regina Barzilay
- K-Best A\* Parsing  
Adam Pauls and Dan Klein

The next ACL conference will be held in Uppsala, Sweden, July 11–16, 2010.

### Papers about machine translation

Machine translation (MT) is definitely one of the hottest research topics in this conference, and there are as many as 5 sessions about MT among the full paper presentations. The papers presented in the MT sessions covered syntax-based model, system combination, decoding algorithm and word alignment.

#### \*Syntax-based model

(Haitao Mi et al.) applied forest based method in a tree-string MT model in 2008. This year, (Yang

Liu et al.) demonstrated the power of forest in a tree-tree model.

The strength of forest is also leveraged by (Hui Zhang et al.) to improve the tree sequence to string model.

(Jun Sun et al.) improved the tree sequence model by using another technology: non-contiguous tree sequence.

(Galley and Manning) applied quadratic-time MST parsing to compute language model scores for phrase-based MT.

(Deyi Xiong et al.) proposed a syntax-driven bracketing model for phrase-based translation.

(Phil Blunsom) presented a phrasal synchronous grammar model of translation equivalence. Without resorting to heuristics or constraints from a word-alignment model, they directly induced a synchronous grammar from parallel sentence-aligned corpora.

#### **\*System combination**

Hypothesis alignment is a key step in system combination. (Boxing Chen et al.) first compared four commonly used word alignment methods, namely GIZA++, TER, CLA and IHMM, for hypothesis alignment. Then they proposed a method to build the confusion network from intersection word alignment, which utilizes both direct and inverse word alignment between the backbone and hypothesis to improve the reliability of hypothesis alignment.

(Chi-Ho Li et al.) re-designed the Indirect HMM(IHMM) alignment, which is one of the best hypothesis alignment methods for conventional MT system combination, in an incremental manner. They proposed three different solutions to define state transitions on a gradually expanding confusion network.

#### **\*Decoding algorithm**

(Mu Li et al.) and (Yang Liu et al.) went a step further than system combination by proposing collaborative decoding and joint decoding algorithm respectively.

(John DeNero et al.) presented a consensus decoding algorithm which can improve performance as well as MBR but run 80 times faster than MBR.

(Zhifei Li et al) proposed a decoding algorithm for MT which can measure the goodness of a translation using all of its derivation. The paper was nominated for the best paper award.

#### **\*Word alignment**

(Aria Haghighi et al.) proposed supervised word alignment methods that exploit ITG constrains. They showed that through the combination of relaxed learning objectives, many-to-one block alignment potential, and efficient pruning, ITG models can yield state-of-the-art word alignments.

(Fei Huang) presented a confidence measure for word alignment based on the posterior probability of alignment links. They keep high confidence alignment links and remove low confidence alignment links. The idea is quite simple but it is very effective in terms of both alignment F-score and BLEU score.

## AAMT インターネットワーキンググループの長尾賞受賞に際して

インターネットワーキンググループ

富士 秀、伊藤 悦雄、介弘 達哉、小谷 克則

インターネットワーキンググループ（以下、「WG」）は、1996年に当協会内の組織として発足し、協会の情報インフラを運用・管理する「黒子」的な役割を担ってきました。発足以来13年の間に、法人会員企業からのWGメンバーを中心に、実に数多くの方の協力をいただいて活動を進めてきました。

この度、2009年度長尾賞という形で、これまでの私達の活動に対する評価をいただいたことは大変に感慨深いことです。この喜びを、これまでWG活動にかかわってこられた多くの方々と分かち合いたいと思います。またこの機に、あらためて当WGの活動を紹介させていただきたいと思います。

### 1. 現在の活動

#### 1.1. 協会インフラの運用

##### ・ホームページ

協会独自のサイトを立ち上げ、公開ページにおいて対外的な情報発信を行っています。一方、協会会員を対象にした「会員専用ページ」を同サイト上に構築して運用しています。

##### ・メールマガジン

協会発の機械翻訳関連情報を協会会員にタイムリーに発信するために、メールマガジンを立ち上げています。

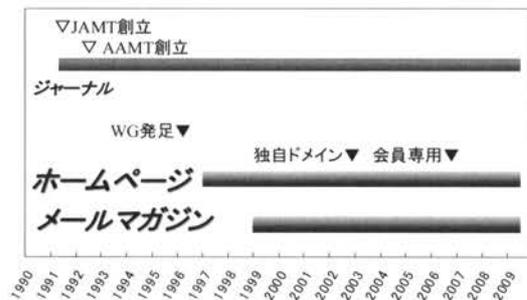


図 1. WG 発足と担当するインフラ

(ジャーナルは編集委員会が担当)

#### 1.2. その他の活動

ホームページおよびメールマガジンの運用に加え、以下の活動を推進しています。

##### ・委員会活動支援

近年、協会の各委員会では、従来のコンテンツ掲載に加えて、より積極的なウェブ活用を始めており、WGもこれをサポートしています。委員会のウェブ活用の例としては、ダウンロード可能コンテンツの公開、ウェブ上アンケート、等があります。

##### ・アジア太平洋向け情報流通

協会発足当初から、日本語中心になりがちな協会の情報流通において、アジア太平洋地域に向けた情報流通の枠組み作りを模索しています。

### 2. WG の活動の歴史

#### ・1996年 WG 発足

WG はジャーナル編集委員会の一環としてスタートしました。ジャーナルに掲載しているコンテンツのうち、特に定期的な更新が必要となるコンテンツを、当時急速に普及し始めていたインターネット上で公開する、というのが当初の最大の目的でした。

#### ・1997年 初期ホームページ作成

当協会初のホームページを作成して公開しました。コンテンツはごく限定的なもので、協会の全体概要とMTシステム一覧のみという内容でした。サイトとしては、電子情報技術産業協会 (JEITA) のサーバの一角に間借りさせていただきました。

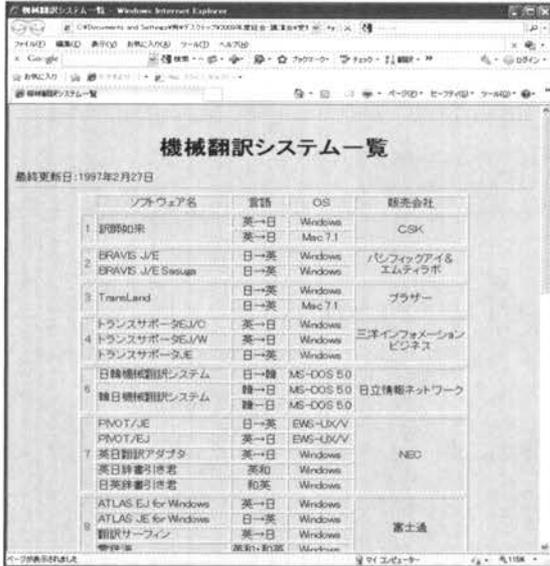


図 2. MT システム一覧 (1997 年掲載)

・1999 年 メールマガジン立ち上げ

ホームページよりさらに即時性が求められる情報の発信を目的として、協会メールマガジンを立ち上げました。配信対象のコンテンツは、会議・行事案内、新製品紹介、事務局連絡等で、基本的に同様の内容が現在まで引き継がれています。

・2000 年 総合的ホームページの立ち上げ

それまで限定的な内容だった協会ホームページを全面的に設計し直し、協会の活動内容を網羅的に紹介するサイトとして再スタートしました。主要なコンテンツに関しては、英語を併記するようにしました。また、一つのウェブサイトとして統一感のあるデザインとしました。これによってコンテンツは充実しましたが、コンテンツは依然として間借りのサーバ上に搭載していました。



図 3. 総合的ホームページ (2000 年公開)

・2002 年 独自ドメイン取得・デザイン一新

この年初めて、協会専用ディスクスペースをレンタルサーバ上に取得しました。併せて協会専用ドメイン "aamt.info" を取得し、コンテンツをこのドメイン上に移行しました。

さらに、コンテンツの大幅な拡充と最構成を行い、グローバル対応を行いました。協会基本情報は、アジア太平洋地域の 4 カ国語で表示できるようにしました。また、アジア太平洋各地域のローカル情報を日英両言語で掲載しました。また、サーバの提供するアクティブ機能を利用して、入会フォームを作成し、アクセスログの解析を開始しました。



図 4. 協会ドメイン取得とデザイン一新 (2002 年)

### ・2003年 委員会活動のサポート

協会各委員会のウェブ上での活動をサポートし始めました。

当時の市場動向調査委員会で実施していた MT ユーザのアンケート調査を協会サイト上で行えるようにサポートを行いました。委員会活動で必要となるアクティブコンテンツが利用できるよう支援を行い、この年から現在にいたるまで、アンケートをサイト上で行っています。

また、他の委員会の活動についても順次サポートを行っています。

### ・2004年 アジア太平洋展開

これまで WG で作成してきた各地域のローカル情報を、各地域で独自に更新していただくべく、手始めにタイローカルページの運用移管を行いました。タイローカルページについて、タイのサーバ上でコンテンツを作成・更新していただき、日本の AAMT ページからリンクを張るようにしました。



図5. タイサーバ上のタイローカルページ  
(統一デザインのページをタイ側で更新)

### ・2007年 会員専用ページの立ち上げ

サーバ上にパスワード認証付きの区域を作成して会員専用ページを立ち上げ、会員向けのウェブコンテンツの掲載を開始しました。ジャーナルの電子データや、各委員会の議事録等の掲載を開始しました。

またこの年、従来利用してきた2種類のプロバイダを一本化しました。これによって、ホームページアドレスとして使われていた"aaamt.info"は、すべての協会関連メールアドレスでも使われるようになりました。

#### メーリングリスト

AAMT  
FORUM

当協会ではMT関連情報をメーリングリストでタイムリーに会員に提供しています。  
このページにはその内容を整理して掲載しています。



#### AAMTジャーナル

AAMTジャーナルの記事をPDF形式で掲載しています。  
ご利用は、著作権法の範囲内に留め、他への転載はご遠慮下さい。



#### MT News International

MT News Internationalの記事をPDF形式で掲載しています。  
ご利用は、著作権法の範囲内に留め、他への転載はご遠慮下さい。



#### 各委員会の議事録

各委員会の議事録を掲載しています。  
ご利用は、著作権法の範囲内に留め、他への転載はご遠慮下さい。

図6. 会員専用ページのトップページ

### まとめ

当 WG の活動に携わった多くの方たちの継続的なご尽力により、協会のインフラを段階的に、しかし確実に充実させてきました。

今後、機械翻訳の利用は、ますますインターネットと密接に結びついたものとなると考えられます。当 WG で構築してきた環境は、単なるインフラとしてだけではなく、協会の活動とともに発展していくような存在となればと考えています。

## 英日・日英機械翻訳に関するアンケート結果報告

機械翻訳課題調査委員会ワーキンググループ 2

## 1. はじめに

AAMT 機械翻訳課題調査委員会 WG2（調査・広報・啓蒙）では、その前身である市場動向調査委員会の時代から調査活動の一環として、年一回インターネット上で機械翻訳に関する一般ユーザを対象としたアンケートを行っている。AAMT の Web サイトに特設ページを設け、設問に回答してもらうという形式である。本稿では、2009 年（2008 年度）のアンケート調査結果について報告する。

2009 年のアンケートでは、新たに携帯電話による機械翻訳の利用調査、および翻訳メモリの利用調査を行った。さらに、従来から継続している PC による機械翻訳の利用調査も内容を整理して行った。

近年、携帯電話による通信サービスが広く普及し、携帯電話によるテキストや音声の機械翻訳サービスが始まった。今後、機械翻訳の利用形態として携帯電話のニーズが高くなると思われる。そこで、携帯電話による機械翻訳サービスに関する利用状況を調査することにした。また、翻訳に関する技術に翻訳支援技術があり、この技術として翻訳メモリが実際に利用されている。そこで、翻訳メモリの利用状況を調査することにした。

今回の調査では、アンケートの回答は PC だけでなく、携帯電話からも回答できるようにした。ただし、携帯電話からの場合は、携帯電話による機械翻訳の利用と PC による機械翻訳の利用についての調査のみとした。また、自動翻訳や PC による翻訳支援の必要性を問い、必要としている回答者のみで分析を行った。

以下に 2009 年のアンケート実施概要を示す。

- アンケート期間
  - 2009 年 1 月 28 日～2 月 10 日
- アンケート進呈品
  - 抽選で 10 名に機械翻訳ソフト、20 名にクオカード（3,000 円分）を進呈。
  - 回答者は進呈品として機械翻訳ソフトかクオカードかを指定可能。また、進呈品の機械翻訳ソフトは 4 種類で、それらの商品名は回答前に確認可能だが、商品の指定は不可。
- アンケート回答者数
  - 1,811 人（うち、自動翻訳や PC による翻訳支援を必要とすると回答した人は 1,059 人）
  - なお、過去のアンケート回答者数は、2003 年 1,413 人、2004 年 1,720 人、2005 年 1,018 人、2006 年 1,061 人、2007 年 835 人、2008 年 1,495 人（946 人）。括弧内は自動翻訳を必要とする人数。

## 2. 設問項目

主に以下に示す項目について質問した。

- 携帯電話による機械翻訳の利用について（知名度、使用経験、利用したい場面、利用した場面、入出力の種類、満足度、良い点、悪い点）
- 翻訳メモリの利用について（知名度、使用経験、使用理由、翻訳方向、利用分野、コンテンツ、機械翻訳ソフトの併用、満足度等）
- PC による機械翻訳の利用について（最も良く使用する機械翻訳の種類、利用頻度、満足度）

- 回答者の基本属性について（年齢層、性別、職業、英語力）

### 3. 携帯電話の機械翻訳サービスに関する調査結果

本章では、携帯電話を端末とした機械翻訳サービスについて実施した調査結果を示す。

アンケート回答者の内、携帯電話向け翻訳サービスサイトを知らないとの回答は39%、1種類知っているが36%であった（図1）。2種類以上知っている回答者を含め約6割が知っていることとなり、携帯電話向けの翻訳サービスがかなり知られていることが分かる。知っている携帯電話向け翻訳サイトは「モバイル版Yahoo!翻訳」が回答者の60%を超え、続いて携帯版エキサイトと続いている（図2）。また、実際に使用したことがあるサイトもこの順と変わらない（図3）。これは、昨年実施したインターネット翻訳サイトに対する調査結果と一致している（AAMT ジャーナル45号参照）。知っているサイトと使ったことがあるサイトの関係を見ると、存在を知っている人の内、約半数はそのサイトを使用したことがあることが分かる（図4）。

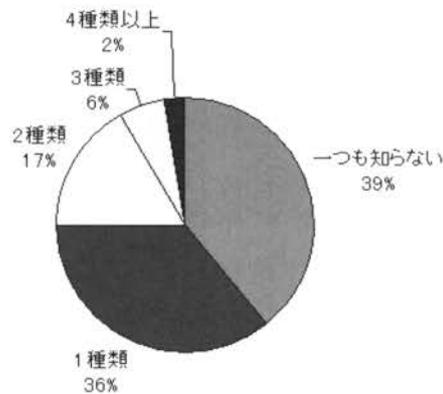


図1 知っている携帯電話向け翻訳サイト数

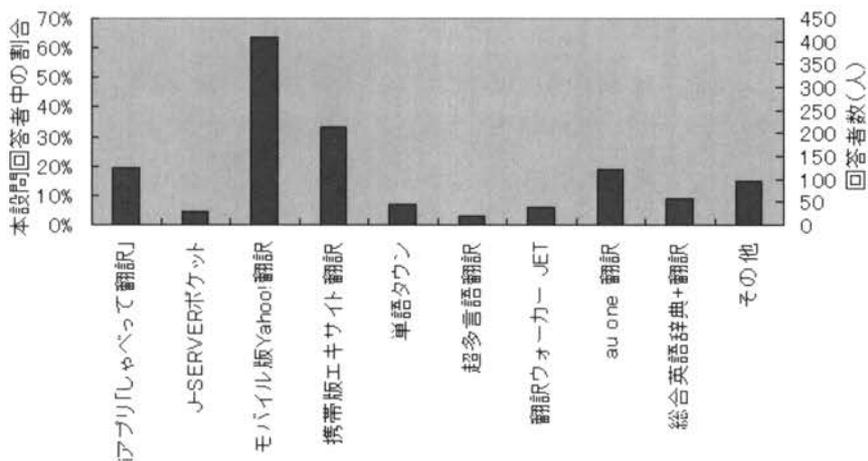


図2 知っている携帯電話向け翻訳サイト名

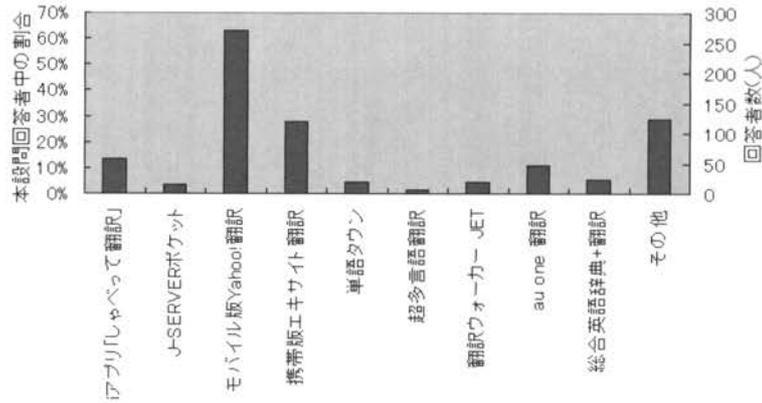


図 3 使ったことがある携帯電話向け機翻訳サイト名

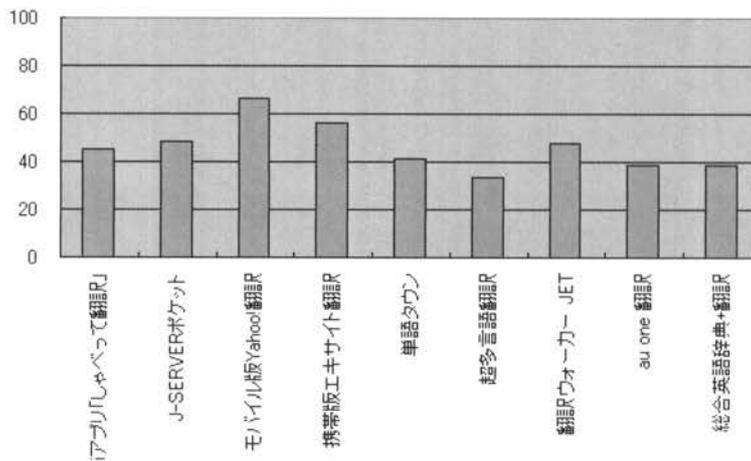


図 4 それぞれのサイトを知っている人の内、そのサイトを使ったことがある人の割合

携帯電話向け翻訳サービスを使用したい場面に関する調査結果を図 5 に示す。これによると、「海外旅行での会話に」使用したいが 60%を超えている。しかし、実際には20%未満しか、「海外旅行での会話に」使用しておらず、「英語のテキストの和訳に」使用している人が約半数であることがわかる(図 6)。

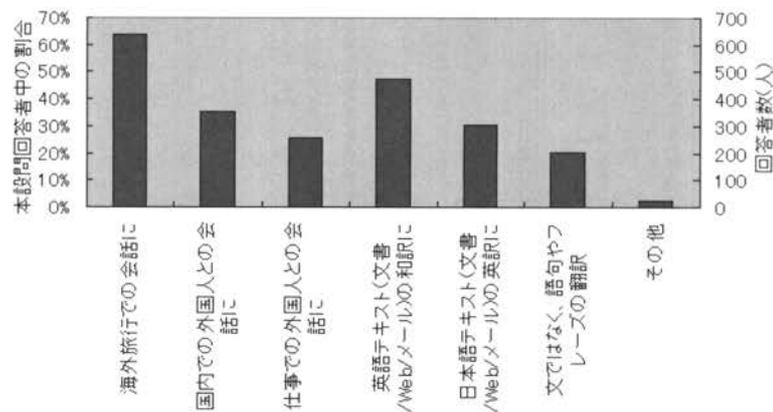


図 5 携帯電話向け翻訳サービスを使いたい場面

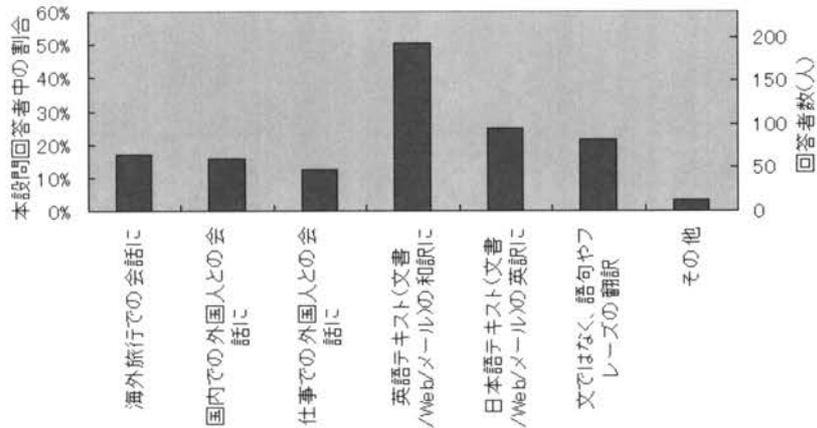


図 6 携帯電話向け翻訳サービスを実際に使用した場面

翻訳サービスの入出力の割合を図 7 に示す。ボタンによる文字入力とテキストによる出力の割合が高いが、音声入力も 20%程度利用されていることが分かる。

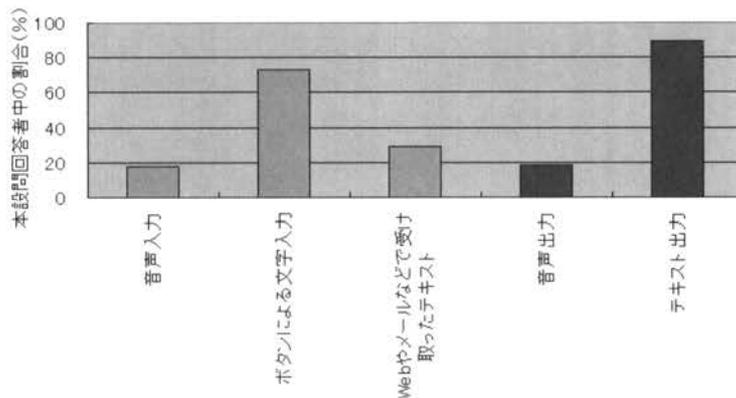


図 7 翻訳サービスの入出力

携帯電話向け翻訳サービスに対する満足度調査結果を示す(図 8)。調査は、「大変不満足」から「大変満足」までの 6 段階評価であり、「音声認識精度」「翻訳品質」「翻訳速度」「音声合成(読み上げ)の品質」という個々の機能の項目と、「使いやすさ」「全体の満足度」「役に立ったか」「今後も使いたいか」という全体に対する項目がある。

個々の項目に関してはあまり満足度が高くないが、その中でも「翻訳の速さ」は他の項目と比べ満足度が高いことがわかる。また、「満足度」「役に立ったか」「今後も使いたいか」の項目より、「あまり満足できなかったが、そこそこ役に立つので、また使いたい」という傾向がわかる。

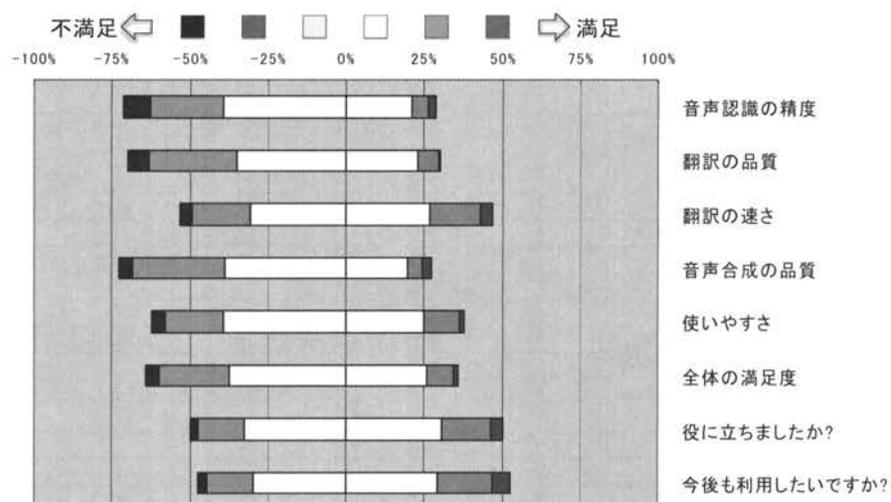


図 8 携帯電話向け翻訳サービスの満足度

携帯電話の機械翻訳サービスの良かった点・悪かった点についての調査結果を示す。これは、携帯電話の機械翻訳サービスを利用したことがある人に、良かった点・悪かった点をそれぞれ自由記述で記載してもらったもので、それぞれ 327 件（回答者数 321 人）、324 件（回答者数 316 人）のコメントを得た。表 1・表 2 にそれぞれ内容による分類と要望数の内訳、記載内容の代表例を示す。

表 1 携帯電話の機械翻訳サービスの良かった点（機械翻訳不要と回答した人のコメントも含む）

大分類	小分類	計	代表的なコメント
[利便性]	[一般]	120	いつでもどこでも翻訳できる。
	[辞書機能]	75	辞書が手元にはないときは使える
[レスポンス]	-	27	スピードがはやい。
[概要を理解できる]	-	23	概要がわかる
[翻訳できること]	-	22	翻訳できたこと
[操作性]	-	17	使いやすい
[翻訳品質]	-	5	精度が思ったより良かったです。
[コスト]	-	3	無料で使える
[音声]	[音声認識ができること]	1	実環境で音声認識が動作したこと
	[音声合成品質]	1	音声思ったより鮮明だったこと
[長所なし]	-	26	
他	-	7	
		327	

表 2 携帯電話の機械翻訳サービスの悪かった点（機械翻訳不要と回答した人のコメントも含む）

大分類	小分類	計	
[翻訳品質]	[一般]	110	もう少し精度を高めて欲しい。
	[語彙不足]	15	語彙が少ない
	[会話・口語]	7	翻訳されてる言葉と会話したときの言葉がかみ合わなかった
	[辞書情報量]	7	辞書と比べると内容（質）には、満足できない。
	[長文]	4	長いと翻訳文が意味不明になる。
	[分野限定]	1	1タスクが旅行会話限定であること
[入力の手間]	-	42	一文字ずつボタンで入力するのが面倒だった
[レスポンス]	-	21	時間がかかる
[操作性]	-	20	使いにくい
[音声]	[音声認識精度]	8	音声認識が非常に悪く、使い物にならない。
	[音声合成品質]	7	音声がやや聞き取りにくい
	[音声機能がない]	4	音声認識や音声合成機能が自分の携帯にもあったらいいと思う。
[見づらい]	-	9	画面が小さい。
[コスト]	-	6	高かった
[機能・完成度]	-	3	もっと完成度が上がってくるといいと思います。
[短所なし]	-	50	
他	[役に立たない]	2	役に立たない
	[文字がわからないと翻訳できない]	2	文字がわからないと翻訳できない
	[セキュリティ]	1	携帯電話には多くの個人情報が登録されているほか、決済機能も併せ持つなど紛失時のリスク。「安全・安心」に関するサービス。
	他	5	限界を感じた。
		324	

携帯電話の機械翻訳サービスについては、まず携帯電話で翻訳できるという利便性自体が長所として捉えられている一方、翻訳品質、音声認識・合成の品質については厳しい評価がなされている。また、携帯電話特有の問題として、操作性の悪さ（入力の手間・画面の狭さなど）があげられている。今回の結果では、従来の PC 向け機械翻訳システム・サービスに対する結果とは異なり、紙の辞書・電子辞書を比較対象としたコメントが多いが、これは、上記のような問題から、携帯電話の機械翻訳サービスは辞書引きもしくはその延長として利用されていることが多いことを裏付けるものと思われる。

#### 4. 翻訳メモリに関する調査結果

本節では、翻訳メモリに関する調査結果を示す。

図 9 に、翻訳メモリの認知度・利用度を示す。そもそも 76% の回答者が翻訳メモリというものの自体を知らない。残る 24% のうち、無回答が 8%、翻訳メモリを知っていて使う人が 8%、翻訳メモリを知っていて使わない人が 8% である。全体から見ると 10% 弱である。翻訳メモリの利用層が限られたものであることが窺える。

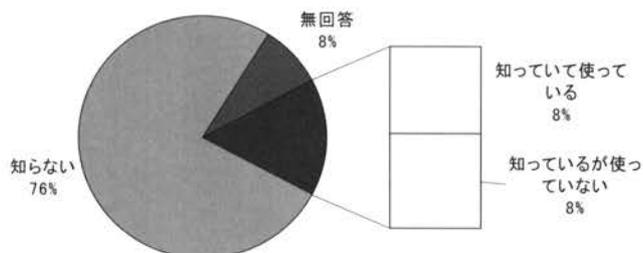


図 9 翻訳メモリの認知度・利用度

図 10 に使用している翻訳メモリ（複数選択可）を示す。Trados のみが 50%超の利用率で、それ以外の全てのソフトウェアが 20%未満の水準にある。翻訳メモリの定番が Trados であることがアンケートからも裏付けられる。

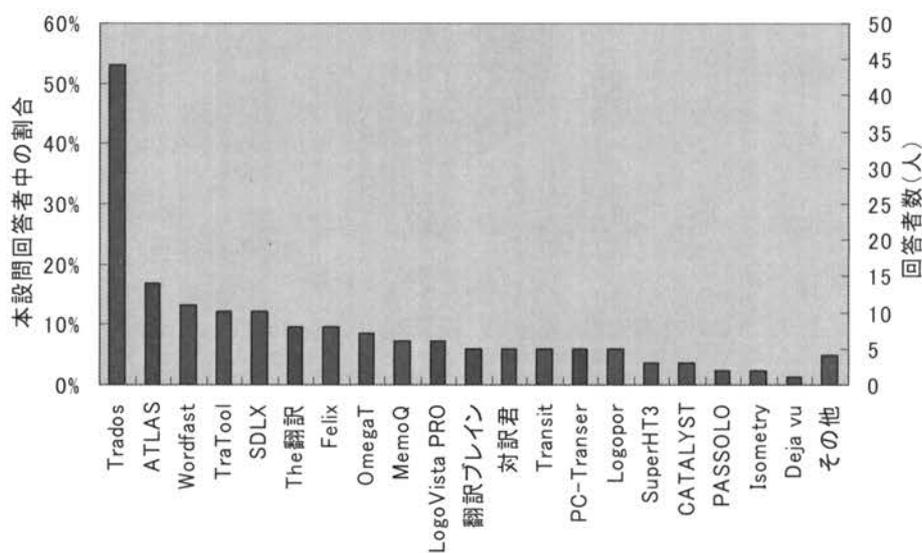


図 10 使用している翻訳メモリツール

図 11 に翻訳メモリの使用理由を示す。翻訳メモリは、産業翻訳において、翻訳発注者から使用を義務づける形で支給されることが多いと言われているが、その割合が 3 割程度であることがわかる。また、翻訳メモリ使用者の大半も、義務だから翻訳メモリを利用するだけでなく、一般にその有用性を認めていることがわかる。

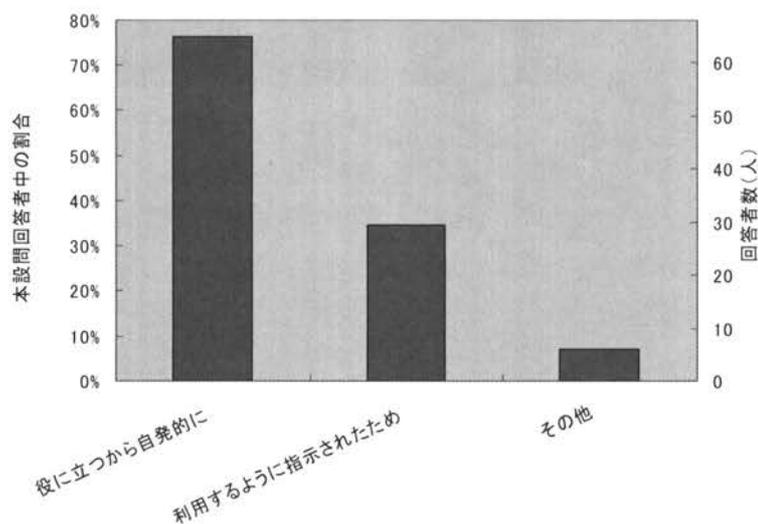


図 11 翻訳メモリの使用理由

図 12 には翻訳メモリ利用時の翻訳方向を示す。英日のみが 50%を超過、次いで両方向が 30%強、日英のみが 10%強となっている。

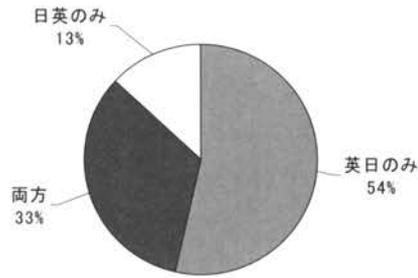


図 12 翻訳メモリ利用時の翻訳方向

図 13 には翻訳メモリコンテンツの調達方法を示す。複数選択可能な設問で、「自分で作成する」が 60%弱、次いで「翻訳依頼者から支給される」が約 40%、「翻訳メモリツールや翻訳ソフト付属のものを使う」が 30%弱、「翻訳依頼者以外から既存のもの入手」が 20%強など、支給される翻訳メモリのみならず、自分でも工夫して可能な限り良質の翻訳メモリを調達している人が多いことが窺える。

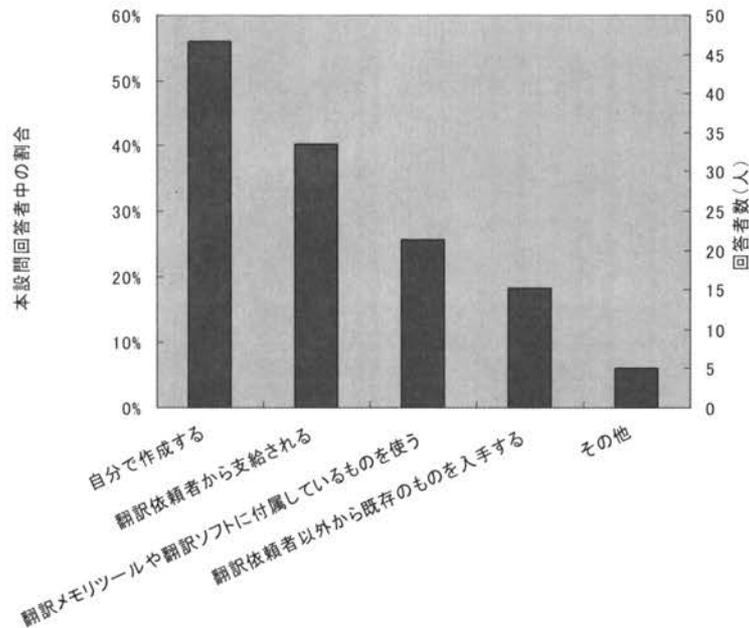


図 13 翻訳メモリコンテンツの調達方法

図 14 には翻訳ソフト購入時に翻訳メモリ機能を考慮する人の割合、図 15 には翻訳メモリと機械翻訳を併用する人の割合を示す。これらの結果から、70~80%は機械翻訳ソフトと翻訳メモリの併用を考えていることがわかる。

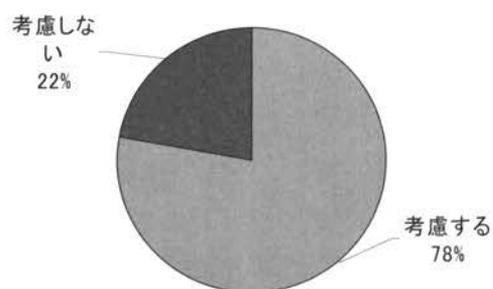


図 14 翻訳ソフト購入時に翻訳メモリを考慮するか

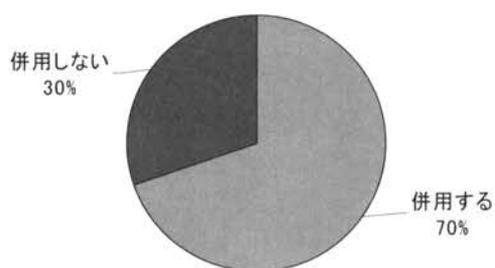


図 15 翻訳メモリと機械翻訳を併用するか

図 16 に翻訳メモリの満足度を示す。コンテンツと価格には比較的不満が高い。価格についての不満は、Trados の利用率が圧倒的に高いことから、Trados の価格への不満が反映している可能性がある。

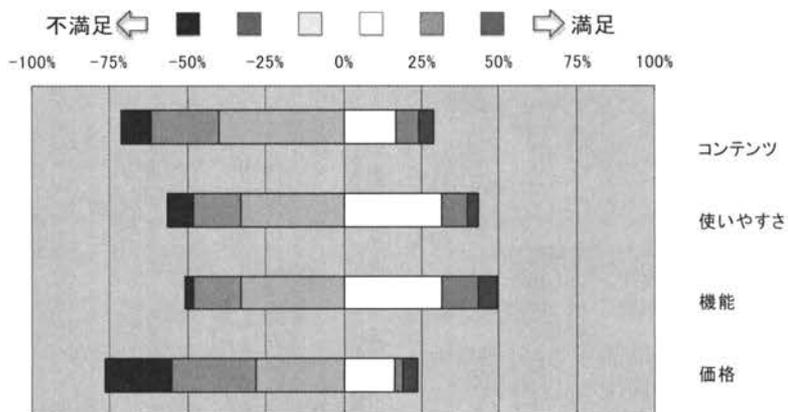


図 16 翻訳メモリの満足度

図 17 には翻訳メモリの改善点についての結果を示す。ここに示した選択肢の理由の他、自由記述による記述も可能な設問だったが、特定の機能や不具合の多さについてのものなど、全部で数人のコメントがあったのみであり、この選択肢の内容で翻訳メモリの改善点はほぼ網羅していると思われる。

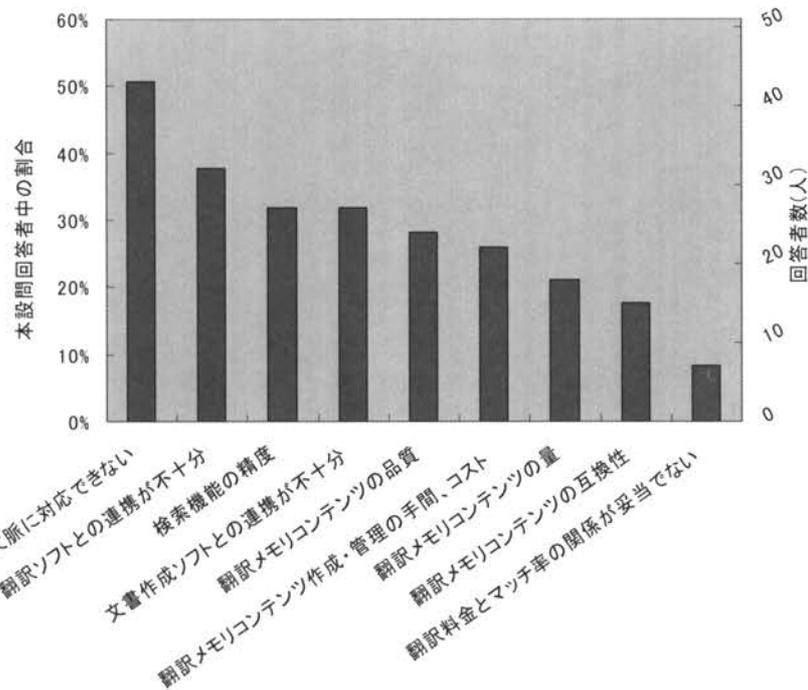


図 17 翻訳メモリの改善点

## 5. PC向け機械翻訳

本章では、PC での機械翻訳の使用に関する調査結果を示す。今回の調査は、「無償の翻訳サイト（以下「無償サイト）」「有償の翻訳サイト（以下「有償サイト）」「パソコンに標準搭載されていた翻訳ソフト（以下「バンドルソフト）」「市販されている PC 用翻訳ソフト（以下「購入ソフト）」に分け、分析を行った。また、満足度に関する調査は、「大変不満足」から「大変満足」までの 6 段階評価で実施した。

回答者数としては、無償サイトの使用者が圧倒的に多い（図 18）。

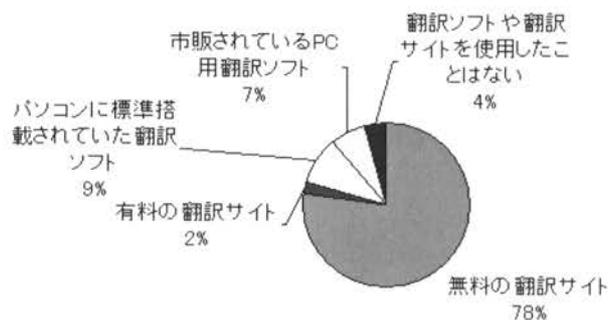


図 18 最も良く使用する機械翻訳の種類

使用頻度は、無償サイトおよびバンドルソフトでは全体的に低く、有償サイト並びに購入ソフトでは高い傾向がある。使用頻度が高い場合には対価を支払っても購入する傾向があるといえる（図 19）。

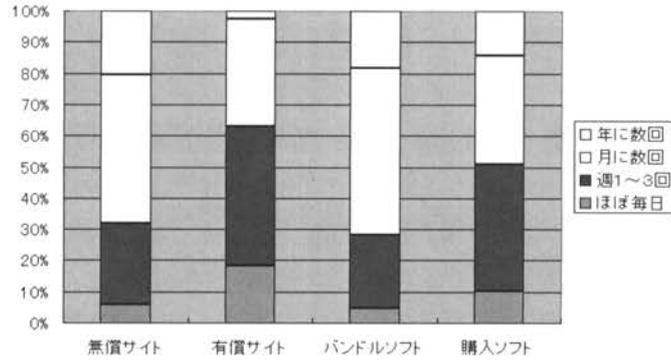


図 19 機械翻訳の種類別の使用頻度

翻訳精度の満足度に関しては、無償サイトは低いのではないかと予想を裏切り、上記4区分では、若干「有償サイト」と「購入ソフト」の満足度が高いものの、あまり差が無い結果となった(図 20)。これは、無償サイトの場合は「無料なのでこれくらいでも OK」という割りきりが、購入ソフトの場合は「支払った額に見合った精度かどうか」という対投資効果という判断基準が生じてくるためだと考えられる。

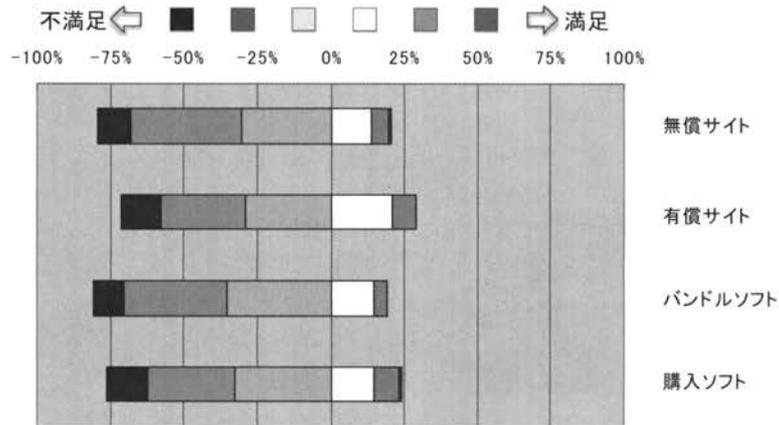


図 20 機械翻訳の種類別の翻訳精度満足度

その他、「使いやすさ」に関しては種類ごとに大きな差は見られない(図 21)。

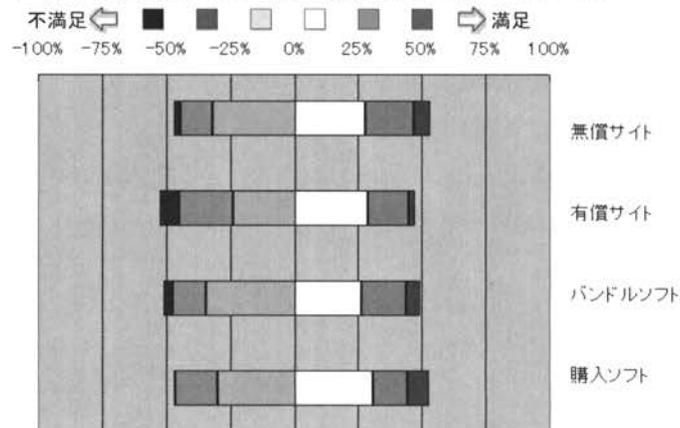


図 21 機械翻訳の種類別の使いやすさ満足度

また、機能に関しては、無償サイトと購入ソフトでは明らかな違いがあるにもかかわらず、満足度には差が見られない(図 22)。

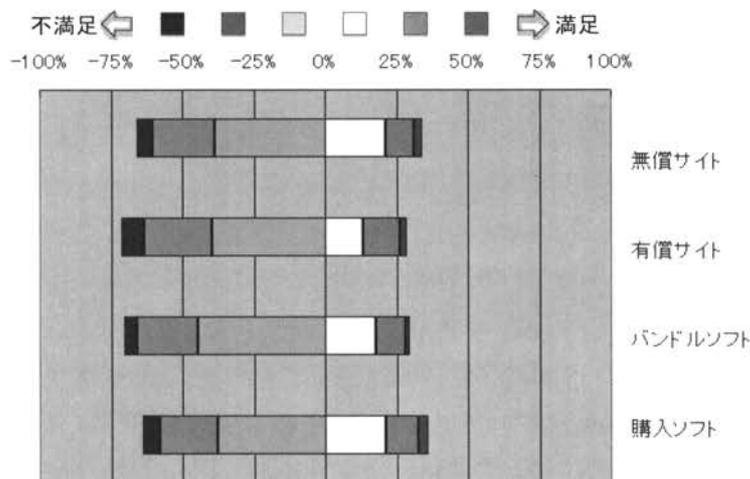


図 22 機械翻訳の種類別の機能満足度

価格の満足度に関しては、「有償サイト」と「購入ソフト」の満足度が、「無償サイト」と「バンドルソフト」に比べて低いが、半数以上が満足側となっている。

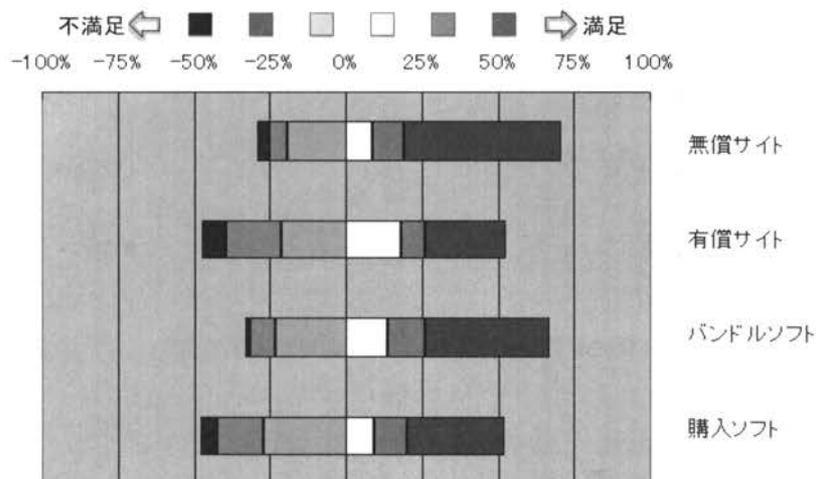


図 23 機械翻訳の種類別の価格満足度

一方、翻訳精度に関する満足度を英語能力別に分析した結果を図 24 に示す。英語力あがると満足度が下がるのではないかという予想に反し、TOEIC 750 点相当未満には差がなく、750 点以上で「大変不満足」者の割合が増加していることが分かる。これは、2000 年に AAMT で行った調査の「TOEIC 690 点以下の人には機械翻訳が役立つ」という結果に関連すると考えられる (AAMT ジャーナル 31 号参照)。

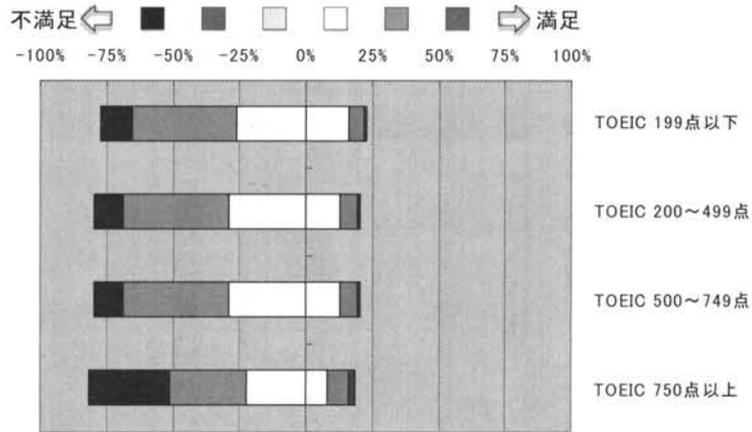


図 24 英語力別の翻訳精度満足度

## 6. 回答者のプロフィール

本節では、回答者の基本属性（年齢、性別、職業、英語力）を示す。

図 25 に年齢、図 26 に性別、図 27 に職業、図 28 に英語能力（横軸は TOEIC スコア）毎の回答者の割合を示す。

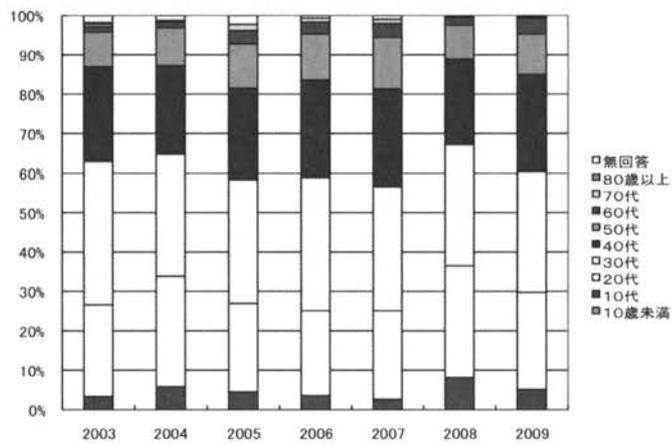


図 25 年齢

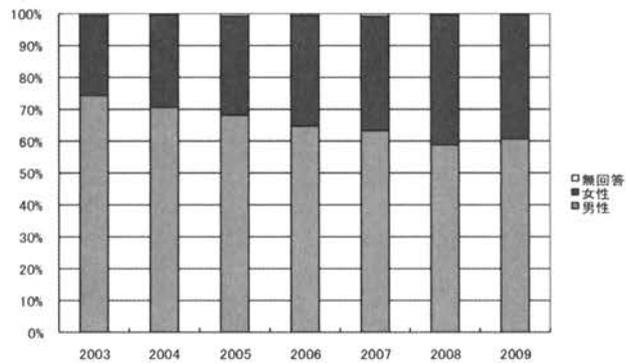


図 26 性別

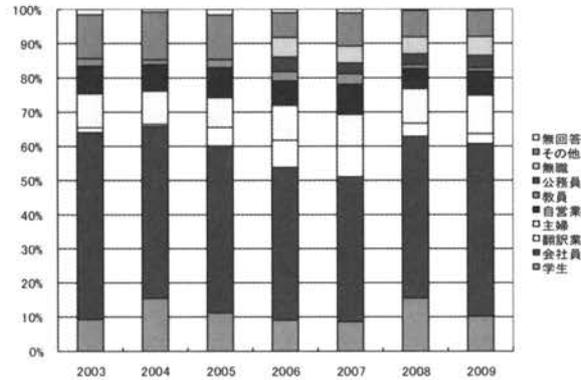


図 27 職業

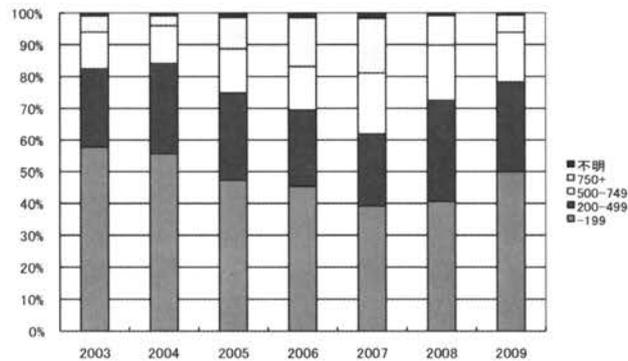


図 28 英語能力

## 7. まとめ

機械翻訳に関するアンケートを実施した。本年は新しい設問項目により、下記のような新しい知見を得ることができた。

- 携帯翻訳サービス
  - 一般的なサービスになりつつある
  - 入出力はテキスト中心だが、音声もある程度使われている
  - 利用者の約半数が、役に立った／今後も利用したいと回答
- 翻訳メモリ
  - 利用者の 76%は役に立つから自発的に使っている
  - 70%の人が MT を併用している
- PC 向け機械翻訳サービス
  - 78%の人が無料翻訳サイトを現在最もよく利用する自動翻訳と回答
  - 翻訳精度に満足・どちらかといえば満足な人は 20-25%

今後も機械翻訳に関する技術の利用状況やニーズに係る調査を続け、機械翻訳に関する技術をより役立つものとするために参考となる情報を提供していきたい。

-以上-

## Common Format for User Dictionary UTX (Universal Terminology eXchange)

WG3, MT Research Committee

---

"Translation software is useless" Are you sure?

### Advantages of UTX

- Significantly improve the translation accuracy
- Accumulate, share, and reuse translation assets
- Cut the time and cost to check up terminology

### Introduction

UTX (universal terminology eXchange) is a common format for user dictionaries established by the AAMT (Asia-Pacific Association for Machine Translation). It is an open specification, and the AAMT is working towards its standardization. The AAMT is an organization consisting of researchers, developers, manufacturers, and users of machine translation (MT). MT is the core technology of translation software.

### Characteristics

UTX greatly improves the accuracy of translation software by sharing the knowledge of terminology in a bilingual format. The goal of UTX is to create a simple, easy to make, easy to use dictionary from a user's viewpoint, not a developer's.

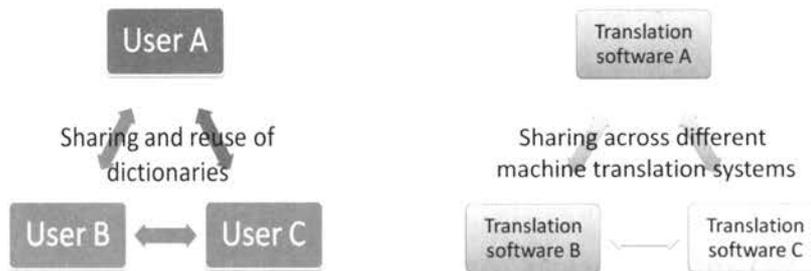
Different manufactures' translation software can use the same UTX dictionary. Also, a UTX dictionary is human-readable, and it can be used as a glossary which does not involve translation software at all.

### Why UTX?

With UTX, a user can easily share and reuse user dictionaries of translation software. When translation software fails to translate correctly, the problem is often that it doesn't have sufficient translation knowledge of certain words and phrases that should be translated. UTX can greatly improve the accuracy of translation software by supplementing translation knowledge as user dictionaries.

Until now, even if a user of translation software makes the effort to prepare user dictionaries, the knowledge is fragmented and dispersed, and thus not effective. Also, even a simple plain text file is difficult to share or to reuse, if its format is not standardized as a dictionary. However, if a single standard like UTX is adopted, dictionaries can be widely shared among various tools, such

as translation software from different manufactures, and they can also be reused quickly.



## Types

There are two types of UTX:

UTX-Simple ... A simple tab-delimited format.

UTX-XML ... A XML format which can contain many data (specification under development).

## Who uses UTX?

UTX is specifically designed to be used by end users of translation software, or translators. It does not require any advanced technical knowledge of linguistics, grammars, or machine translation software etc. to create or use it. It can be made only from minimum data such as basic parts of speech (noun, verb, etc.), and the plural form, if the entry is a noun.

## In which domains?

UTX can be used in any domain, if it is specialized to a certain extent, such as ICT, medicine, law, engineering translation etc. It may not be suitable for translation of non-specialized, general contents.

## What kind of words should we include?

Basically, a UTX dictionary should contain **only technical terms** of a specific domain. In most cases, entries are nouns, especially compound nouns. For example, a word like "XML declaration" can be correctly translated into its Japanese equivalent, "XML 宣言" by just registering it in a user dictionary. Basic vocabulary like "window" should not be included, because such words are already contained in the system dictionaries of translation software. Translation accuracy can be improved by collecting, sharing, and reusing the data of fine-tuned bilingual translations which are not included in translation software out-of-box.

Sentences should not be included, except when it is appropriate to treat them as a kind of "words." As a rule, UTX should be separated from translation memory, which is a bilingual database of sentences, rather than words.

## Is it multilingual?

Yes. The character code of UTX is Unicode (UTF-8) and all the languages that can be handled by UTF-8 are accepted. However, a UTX-Simple dictionary includes only entries of the single source

language A, and their translations in the single target language B. The dictionary from B to A comprises another dictionary.

### How do we make UTX?

UTX-Simple dictionaries can be easily created and edited with any text editor or spread sheet application. With a conversion tool, a UTX-Simple format dictionary can be converted into some proprietary formats for translation software.

### URL for the conversion tool

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

A tool is also under development to convert various formats to UTX format.

### To open source developers and translators - Why not release your glossary in UTX format, and share it with others?

By releasing and sharing glossaries, you can localize your software more quickly, and allow users to access it from all over the world.

### UTX mailing list

Anyone can participate in the discussions on UTX through a mailing list.

Please search the Web with the keywords: "UTX mailing list."

### Reference materials

Okura Seiji et al. (2008) "Introduction to UTX, a specification for a shared user dictionary" 13th Annual Meeting of Association for Natural Language Processing

Francis Bond et al. (2009) "Sharing User Dictionaries across Multiple Systems with UTX-S" in Second International Workshop on Intercultural Collaboration (IWIC2009), Stanford

### Example

(based on UTX-Simple 1.0)

#UTX-S 1.0; en-US/ja-JP;			
2008-03-15T10:00:00Z+09:00; copyright: Francis Bond (2008); license: CC-by 3.0			
#src	tgt	src:pos	src:plural
new	新規の	adj	
fast	高速な	adj	
prosody	韻律	noun	prosodies
save	保存する	verb	

# denotes a comment line.

- 1st line (comment line): Basic information on the dictionary. Each item is divided by a semicolon and a space. "UTX-S" means UTX-Simple.

```
# UTX-S <version number>; <source language>/<target language>; <last update
timestamp>; copyright: <Copyright holder>; license: <License>; <Additional data>
(optional)
```

2nd line (comment line): The names of properties. Each item is tab-delimited.

In the above example:

```
# <Source word> <Target word> <Part of speech of the source word> <Plural form of the source
word>
```

The 3rd and subsequent lines contain actual entries. Each item is tab-delimited.

The character-encoding method is UTF-8 (with no BOM).

### Tips for making a UTX dictionary (glossary)

- Define the domain of the dictionary clearly.
- Only use upper case for proper nouns.
- The basic form of word should be entered (singular form for a noun, root form for a verb - as you would see in a commercial dictionary).
- Any comments should be noted separately in the comment field, not as a part of the entry.
- Choose only the single, most appropriate translation corresponding to a source word. If it has multiple DISTINCTLY different meanings, they can be treated as separate entries.

### Disclaimer

The AAMT and its members offer the specifications of UTX, UTX-Simple, and UTX-XML (henceforth, UTX specifications) "AS IS," and give no guarantee on anything related to UTX, and have no responsibility to any results from the use of UTX dictionaries. The authors of the various tools related to UTX do not give assurance on the accuracy of the conversion. Please use UTX specifications at your own responsibility.

Although the AAMT has not abandoned the copyright of UTX specifications, UTX specifications are made public, and anyone can use them. The author of a dictionary can specify the licenses of the dictionary with or without a commercial activity. However, a user may need to check the origin of data individually about the legitimacy of the copyright of the data used in a dictionary. The AAMT and its members do not check or guarantee the legitimacy of the copyright of dictionaries created by individual users except for the dictionaries created by the AAMT.

A user can use a dictionary created based on UTX specifications according to the license of the dictionary. Since the licenses of the distributed dictionaries differ, please confirm the license and terms indicated in the UTX file header.

**MT Research Committee Members of Sharing/Standardization Working Group**

(not in a particular order)

YAMAMOTO Yuji	CosmosHouse
ITOU Hajime	Inter Group Corp.
MURATA Toshiki	Oki Electric Industry Co., Ltd.
Francis BOND	Nanyang Technological University
SHIMAZU Miwako	Toshiba Solutions Corporation
OKURA Seiji	Fujitsu Laboratories Limited
Michael Konin KATO	Learning Consultant

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

Contact: [aamt-info@aamt.info](mailto:aamt-info@aamt.info)

2009 July edition

# AAMT会員のひろば

AAMT 会員の新たな交流の場を AAMT Journal 誌面上で提供するべくスタートいたしました「AAMT 会員のひろば」、会員の皆さまのご助力をいただきまして、第一回の No.41 のスタートから第六回を迎えることができました。今号では、個人会員 5 名の皆さまからのご寄稿をいただいております。

国内外で、学生、研究者、学識経験者など様々な分野でご活躍されている方々から、かつての AAMT 事務局、AAMT Journal の懐かしいエピソード、機械翻訳研究への提言、AAMT の活動へのご要望など、貴重なご意見をお寄せいただきました。

AAMT Journal では今後も引き続き、会員の皆さまからのご寄稿を心よりお待ちしております。

ご寄稿・お問い合わせは AAMT 事務局(E-mail: AAMT-info@AAMT.info)まで宜しくお願いいたします。

---

## 個人会員（敬称略・50 音順）

---

### 会員名

立見みどり (School of Applied Language and Intercultural Studies  
Dublin City University) / TATSUMI Midori

---

### 自己紹介

わたしは現在、アイルランドのダブリンにある Dublin City University (DCU) という大学の博士課程に在籍し、機械翻訳に関する研究を行っている学生です。

翻訳との最初のかかわりは、途上国支援団体のボランティア翻訳でした。その後並行して、米国に本社を置くソフトウェア開発会社で 5 年間、さらにフリーランスとして 10 年間、主に IT 関連文書を翻訳していました。さすがに「紙とペン」という時代は経験していませんが、最初に翻訳を始めた頃はまだパソコンが普及しておらず、ワープロ専用機を使っていました。ソフトウェア開発会社でコンピュータを使い始め、モデム通信からインターネットになり、フリーになってからは翻訳メモリソフトも導入し、翻訳環境が Web に移行されたり機械翻訳が普及し始めたりと、翻訳とのかかわりの中でさまざまなテクノロジーの進歩を目の当たりにしてきました。

そんな頃、大学で翻訳を勉強し直してみたくなり、英国 Imperial College London で、MSc in Scientific, Technical & Medical Translation with Translation Technology という修士号を取得しました。当時はこんな学位がこの世に存在するとは思いませんでしたが、技術・産業翻訳や翻訳技術に関する教育や研究はヨーロッパの大学では結構盛んに行われていることがわかりました。その後、DCU とセキュリティおよびデータアーカイブソフトを開発・販売する Symantec Corporation による産学協同研究プロジェクトに参加する機会を得、機械翻訳の利用者側から、機械翻訳技術をより有効に活用する方法についての研究を行っています。

わたしのテーマは、機械翻訳の人手によるポストエディット（後編集）です。原文の特徴と訳文のエラーの関連性などに基づき、どのようなポストエディティングが必要となるのか、どのような要素が困難や負担の原因となっているのかなどを分析しています。機械翻訳の導入率が上がるにつれ、今後ポストエディタは重要な役割を

担う職業になるでしょう。翻訳の現場で、人とテクノロジーがよりよい形で役割を分担するにはどうすればよいのか。機械翻訳ワークフローの改善、翻訳自動修正技術の向上、ポストエディティングのためのガイドラインやトレーニングの開発など、課題はたくさんあると思います。何らかの形で翻訳者や翻訳技術の利用者の役に立つような研究成果をあげられればと考えています。

---

### AAMT へのご要望

研究者という立場で翻訳に関わり始めたのがつい最近であり、またアイルランドに在住しているため、日本の研究事情、業界事情に関する情報が入手しにくいと感じています。AAMT に参加させていただいたのはその点を克服するためでもあります。今年の9月には Ottawa での MT Summit XII でポスターを発表する機会があり、最新の状況を知るとともに、日本からの研究者の方々ともお知り合いになることができたのですが、なかなかそのような機会はまれです。遠方からでも、AAMT のサイトを介してより幅広い情報交換や交流の機会があれば、と思います。

また、産業翻訳、技術翻訳など、マーケット規模の大きな業界では機械翻訳を取り入れようとする動きが強まっているように思われます。しかし、業界内でその動向を調査して発表するような仕組みがあまりなく、あっても公開されることがあまりなく、情報が不足しているのではないのでしょうか。AAMT で、ぜひ機械翻訳を取り巻く状況全体に関する調査のイニシアティブをとっていただけないでしょうか...

---

### 会員名・所属

田中孝/TANAKA Takashi

---

### 言語の壁を越えて

銃を持った若い米兵の前を通過して家に入ると、二人の兵士が靴を脱いで座敷に上がり込み、銃を傍らに置いて日本人と飲食欲談していた。近所の大学生が、近くの飛行場に進駐してきた米兵と親しくなり、自宅に招いたのだ。私も二度程父に連れられて行った。話の内容は勿論判らず、話の仲間に入れないもどかしさはあったが、楽しい雰囲気だった。ほんの数ヶ月前まで鬼畜米英と叫んで戦っていた仲とは到底思えない。これが最初に出会った英語の壁であり、その高い壁をただ見上げるだけであった。

中学に入り英語を学ぶ。面白くてそれなりに進歩はするが、なかなか体得できない。高校、大学と学習年月はどんどん増すが、満足できるレベルには到達しない。もどかしさと能力の無さを感じながら、高い壁を見上げてよちよちと登っていた。

高校や大学ではフランス語とドイツ語を学習したが、英語よりさらに低いレベルでいた。漢文を習う。これは日本語であると思うが、中国語として理解できればもっと興味が増すのではないかと考える。世界中にはこの他にも沢山の言語がある。それぞれの言語を話す人達や文化に直接接したい欲望が強まる。私の前には無数の言語の壁が聳えていた。

現状では英語が国際語の役割を担っている。したがって、英語はぜひともマスターすべきであると思う。しかし、すべて英語で、というのは納得できない。中国人と英語で会話する、韓国の文献を英語で読む、等々、何か受け入れ難い。とは言え沢山の言語を習得する能力は無いし、時間も無い。そんな折、エスペラントの存在を知った。言語の壁を無くそうと、いくつかの国際語や世界語が考え出されているが、エスペラントが最も実用に近

いように思え、学習した。世界中の人が母国語の他にエスペラントを習得すれば、言語の壁を越えて交流できる。

エスペラントで文通をした。アジア、アフリカ、中南米、欧州大陸など、英語圏外の人々と楽しく付き合えた。米国人や英国人ともエスペラントで文通した。英語での文通経験は無いが、英語よりも平等な気持ちで交流出来たと思う。習い始めて数年後、大阪でエスペラント世界大会が開催され、大勢のエスペランティスト（エスペラントを話す人）が世界中から集まった。エスペラントで狂言を演ずることになり、私は太郎冠者をやった。猛練習の甲斐あって、狂言の面白さが全員に通じ、皆大声で笑ってくれた。一方通行ながら、沢山の言語の壁を同時に越えたことを実感した。しかし会話となるとまだまだ難しい。自然言語に比べて文法の規則性が強く学習しやすいが、一つの言語には違いなく、納得のいくレベルへの到達には苦勞する。人造語とは言えヨーロッパの言語を基に造られているので、日本人には欧米人以上の努力が必要である。それでもこの大会で、言語の壁を越えての交流の楽しみを味わえた。

一方その頃、機械翻訳の研究が盛んになって、その成果を新聞などで見かけるようになった。コンピュータや人工知能の研究の進展と、米ソ間の冷戦による技術情報の翻訳需要の増大が、機械翻訳の研究を加速していたように思う。私は多少仕事に関連があったため、興味をもって機械翻訳の進展を見守っていた。「これが完成すれば母語のみで世界中の人達と付き合える」「世界中の書物を母語で読める」「外国語を学ばなくても言語の壁を越えられる」と大きな期待を抱いた。エスペラントと機械翻訳のどちらが勝つか、エスペランティストと議論したこともあった。

企業を定年退職後、AAMT事務局の仕事を紹介された。機械翻訳にはずっと関心を持ち続けていたので、うれしかった。人類にとって非常に重要な機械翻訳の発展に少しでもお役に立てればと思った。また、機械翻訳関連分野における一流の方々とお付き合いも大変魅力であった。

最初の仕事は、小舟町から芝公園への事務所の引越であった。小舟町では仕事をしてないので断定はできないが、芝公園の事務所は狭くはなったが隣に大きなきれいな会議室があり、また機械振興会館がすぐ近くで、会議や打合せには便利になったように思う。同会館へお越しの折に立ち寄って下さる会員の方も多かった。

当初は京大教授の長尾真先生が会長であった。先生との連絡は機械振興会館等へお越しの折を利用したが、多くは郵便や電話であった。後半は東工大教授の田中穂積先生が会長であった。先生の研究室は大岡山にあったので、私が通勤途中に時々立ち寄った。先日田中先生のご逝去の報に接し、大変驚いた。東工大を定年退職されたとはいえお元気でご活躍と聞いていたのに、非常に残念なことである。心からご冥福をお祈りする次第である。横道に逸れるが、あの世には言語の壁はあるのだろうか。霊は言語を介さずに通じ合えるように思うのだが。それとも宗教間の壁があるのだろうか。あの世では、少なくとも天国では、壁がないか、あるいは壁を越えて通じ合える手段があると信じたい。

機関誌 AAMT Journal の発行には力を入れた。当時はインターネットがまだ整備されておらず、機関誌が会員との接触手段として大変重要なものであったからである。事務局として、会員への情報伝達が遅れないよう、発行期限は厳守した。また、表紙の色で号数が判るようにした。この雑誌は 40 ページ程度の薄いものなので背表紙に文字が書けない。それで、電子部品などに使われていたカラーコードを使って号数が判るようにした。1 は茶色、2 は赤、等々。重ねてあっても色で号数が判り、すぐ取り出せる。

協会は当時財政的に苦しかった。会員増による収入増は絶えず検討されたが、利益の得られる事業への目途がなかなか立たない状況下では、むしろ会員の減少を止めるのに苦勞した。会長や運営委員の皆様方等のご尽力で

持ち堪えていたと思う。また、預金金利の急激な引き下げがあり、利子収入が激減した。その結果、支出の削減に頭を悩ませたのも今や懐かしい思い出である。

AAMT の会則には、AAMT の使用言語は日本語と英語と決められている。しかし、実際にはほとんど日本語だけであった。また、AAMT の守備範囲からみると英語は範囲外と言えるが、国際語としての役目を持っている現実からみて使用言語と認めざるを得まい。しかし、AAMT の守備範囲には多種多様の言語がある。これらが使用言語になってないのはおかしい気がした。機械翻訳の現状からみて止むを得まいが、不完全でもよいから、機械翻訳を使っていずれの言語も使用言語となるよう願っていた。そうでないと機械翻訳の専門の協会としてはお尻が落ち着かなかった。開発途上にあるのだから不完全でもよいから、機械翻訳を用いて多言語での会議や機関誌の発行が出来ないか、と日々願っていた。

かつて電気試験所の電子部長をしておられた和田弘氏のお名前を聞いた。和田氏には昭和 30 年代後半に、或る委員会や研究会でお世話になった。当時日本の電気・電子技術の向上のために尽力されていたが、機械翻訳でも活躍されていたことは AAMT の仕事をしてはじめて知った。調べてみると、当時氏が発表された論文も見付かった。まことに懐かしい出来事であった。

AAMT 事務局を辞した後、特許調査の仕事をした。この時、機械翻訳を特許翻訳に使えないか少々検討した。技術分野でありかつ特定分野であるので特許翻訳専用の翻訳機もいくつか開発されていた。いずれの装置でも、うまく訳せる場合も、そうでない場合もある。目的言語の判る人であれば間違いは見抜けるが、そうでない人には理解できない訳文とか誤解を招く訳文になる。そのため、翻訳の補助としては使えるが、翻訳機としてはまだ信頼して使える状態ではないと判断した。その後大分年月が経ったので、いまではかなりレベルが上がったものと思う。

AAMT には、研究者から利用者まで各分野の専門家、しかも各分野の第一人者がおられる。これは、AAMT の大きな財産である。通常、研究者の研究結果に基づき製造者が製品を開発製造し、これをユーザーが利用する。この逆な流れはなかなか得難い。

AAMT ではこれが出来る。ユーザーから出た使用上の問題点を製造者に伝えて改良する。製造者での問題点は研究者に伝えられて解決する。AAMT はこれが出来ると思う。AAMT の大きな強みだと思う。いろいろ困難があろうかと思うが、この強みを生かして機械翻訳のレベルを引き上げてほしい。

機械翻訳が完成し世界中の言語の壁が自由に越えられるようになれば、人類はまた神の怒りをかうのだろうか。人間はバベルの塔を建てて神の怒りをかい、相互に通じない多数の言語を持つようになった。その結果人類は各地にそれぞれ固有の多彩な文化を作り上げた。神の目的はここにあったのではないかと思う。しかし一方、人類は互いに争い戦い合うようになった。言語の壁は争いの大きな原因であることは否定できない。これは神の意志だろうか。いや、これは神の予想外の結果なのではないか。

神はかわいい子を苦しめはしないと信じよう。人類が努力して争いの種になる言語の壁を乗り越える手段を開発することを、神は喜ばないことはない。関係者のなお一層の努力と、機械翻訳の発展を祈る。

十分に調べもせず、思い付くままに書いたため、まとまりの悪い読みにくい文章になった。また、誤りがあるかも知れない。ご容赦を願う。

---

---

会員名・所属

長瀬久明（兵庫教育大学）／NAGASE Hisaaki

---

### 機械翻訳後の未習得外国語文の評価ツールの提案

ターゲット言語を全く知らない人が機械翻訳ソフトを用いて誤訳のない情報を発信できるツールは出来ないだろうか。これは機械翻訳門外漢の筆者が昨年から抱いている期待である。このことについて（無知を顧みず）述べてみたい。具体的には、「原文の意味から程遠い誤訳は無いことを確認できる」ツールである。このようなツールについて、専門家の方は答をお持ちで、私が知らないだけとしたら、それは「いま既にインターネット上にあり誰でも無料で利用可能です。」から「ありません。」までの間のどこに位置しているのだろう。

何をしたいのかを白状しておこう。筆者が勤務する兵庫教育大学は小さな国立大学法人、単科大学である。ところが、驚くことに（研究生も含めての話だが）、学生の約1割が留学生である。つまり「より優秀な留学生を集める」ことは重要事項である。加えて、中国と韓国からの留学生が大半を占めている。このことから、本学の Web ページが中国語や韓国語に機械翻訳されたとき、日本語を知らない中国人や韓国人に理解してほしい、というのが本音である。すべての Web ページの中文版とハングル版を作っても済むのだが、個人的に機械翻訳に活躍してほしい。また、日本語で発信できれば内容に責任が持てる。

しかし、これは望めない話なのかもしれない。Web ページを作る教職員は、ターゲット言語（中国語、韓国語）を全く知らない。閲覧者になるのは容易だが、発信者になれるだろうか。黒田らは、機械翻訳の利用者の英語に関する知識レベルを3段階（H、M、L）に分け、機械翻訳システムに対して「これらの異なるタイプの利用者は確実に異なるタイプの要求を持っている。」と述べている[黒田・加藤, 2009]。しかし、英語以外、例えば中国語、韓国語などではどうだろう。多くの人には「知識なし」のレベル（L以下）、言わば第4のタイプであり、黒田らの利用者のなかに入っていないのかも知れない。では、このタイプの人が機械翻訳システムを用いて発信、情報を伝達することは不可能なのだろうか。

現在のシステムに次の機能をもつツールを追加すれば可能ではないかと筆者は推測している。それは「日本語原文の意味から程遠い誤訳は無いことが確認できる」機能である。この機能さえあれば私は使ってみよう。もっと具体的に言うと、次の2つの機能がほしい。

- (1) 単語についての機能：ある日本語に複数の翻訳候補が見つかったとき、各候補の対訳（逆翻訳された日本語）を見て選択できる機能。
- (2) 翻訳後の各文節についての機能：翻訳後の中文（あるいはハングル）の各文節を逆翻訳した日本語と、ならびに、文節間の関係（主語—述語、など）を見ることが出来る機能。これがあれば、日本語原文のどこがどこに（主語に、述語に、目的語に、などと）翻訳されたのかが分かる。

上の2つの機能があれば、ふつうの日本語 Web ページに加え、中文に機械翻訳して読む人向けの日本語 Web ページや、ハングルに機械翻訳して読む人向けの日本語 Web ページができると期待される（実際に使ってみなければ確認はできないが）。この考えに至った元は次の2つの実験である。

## 実験 1：折返し翻訳

折返し翻訳は、ハングルを例に採れば、

原文（日本語）→日韓翻訳→訳文（ハングル）→韓日翻訳→折返し翻訳文（日本語）

と翻訳し、折返し翻訳文の意味が原文と一致するまで、原文（日本語）の意味を保ちながら、字面だけを別の表現に変更する。利用者が翻訳システムに適応するこの方法を宮部らは翻訳リペアと呼んでいる[宮部、吉野、重信, 2007]。ところが、この方法は訳文（ハングル）が誤訳されていても分からないという欠点がある。したがって、この欠点を当事者間で補うことが出来る場合に限り、この方法は適用できる。すなわち、大学の Web ページのような発信に用いることは躊躇（ためら）われる。とはいえ、訳文がハングルの場合、翻訳精度は高い。したがって、残る課題は、誤訳が無いことが『確認できる』ことだけと思われる。

## 実験 2：中国人留学生と日本人学生との共同作業

材料は本学の Web ページから、あちこちの段落をピックアップした。1 段落（日本語 100～200 文字）に 90 分をかけた。留学生への「丸投げ」ではなく、まず日本語原文を日中翻訳し、留学生が中文の誤訳箇所を指摘した。次に日本人学生が原文の対応箇所を修正した（1 文を 2 文に分けたこともあり、留学生が提案することもあった）。この修正作業を数回、繰り返すと、90 分が費やされる。いくつかの段落について、このような共同作業を行った。修正作業の開始時、すなわち最初の日本語原文はほとんど、留学生に全く理解できない中文に翻訳された。しかし、修正作業の終了時、すなわち最終の日本語原文からはほとんど、「文法の誤訳は無い。用語は所々不自然だが通じるレベル。」と留学生が評価する中文が得られた。

これら 2 つの実験から次のことが考えられる。すなわち、実験 2 において留学生が「誤訳箇所を指摘」することによって得られる情報は、前述の機能(1),(2)によって得られる情報で間に合うのではないか。また、前述の機能(1),(2)は素人目にはツール化できそうである。もしツール化できれば、実験 1 で残った課題（誤訳が無いことが確認できること）がクリアできたことになろう。すると、留学生を頼んで共同作業しなくても、現在の機械翻訳システムとこのツールを併用した翻訳リペアにより、誰でも 1 人で、『〇〇語に機械翻訳して読む人向けの日本語 Web ページ』を作成でき、機械翻訳の適用範囲がメール等だけでなく Web ページに広がる。言い換えると、文法の誤訳は無く、用語は所々不自然だが通じるレベルの〇〇語に機械翻訳されるような日本語 Web ページを誰でも作ることができる。

以上、推測も含む粗野な議論だが論文ではないのでお許し願いたい。もっと多くのデータを蓄積していくことは必要である。その過程でうまく行かない事例が得られれば、追加の機能を考えることになる。ただし、仕様を満たすプログラムが作成できるレベルのコミュニケーションが翻訳リペアにより可能であることは実証されており[宮部、吉野、重信, 2007]、翻訳リペアとツールの併用により、このレベルを下回ることはないだろう。また、タイトルでは「評価」という厳めしい言葉を使ったが、この意味は、Web ページの作者が「これで通じるだろう。アップロードしよう。」と判断できる、という程度の意味である。

## 参考文献

[黒田・加藤, 2009] 黒田航、加藤絃三：「今の機械翻訳に利用者が望めること、望めないこと」（『日本語学』2009 年 10 月号、pp. 30-41.）

[宮部、吉野、重信, 2007] 宮部真衣、吉野孝、重信智宏：「折返し翻訳を用いた翻訳リペアの効果」（『電子情

#### 会員名・所属

平田 周 (静岡大学大学院工学研究科客員教授 事業開発マネジメント専攻

三田教育研究所 LLP 所長) / HIRATA Schu

---

#### 機械翻訳の閉塞性 — 出口から考える

##### 機械翻訳と私

昨年、AAMT に参加させていただいた新米ですが、『日本語で英語を書く』(東洋経済新報社 1984) という奇妙なタイトルの本を書いたのは、機械翻訳ソフトの商品化第1号となったブラビスがまだ開発中の頃のことでした。この本は、機械翻訳ソフトが実用化された時のことを想定して思考実験したものです。

もし機械翻訳ソフトが完成して、和文英訳の粗訳原稿が大量にコンピュータからアウトプットされるようになった場合、誰がその英文粗訳原稿を読み、訂正する(ポストエディティング)をするのだろうかと考えました。英語と日本語の両方に精通していなければチェックはできませんが、そのような能力を持っている人はごくわずかです。コストも高いでしょう。そもそもプロの翻訳家は、コンピュータが出した間違いだらけの英文原稿に手を入れる仕事を好むはずがありません。英語がわからない人が、コンピュータ翻訳の英文の手紙を、そのまま送ってしまうという笑いたくなるような悲劇も起きるに違いありません。

私は、コンピュータのアウトプットが最終原稿になるようなモデルでなければ、いくら機械翻訳の精度が向上しても、問題は解決しないと考えました。それには、間違いを少なくするためのプリエディティングではなく、コンピュータが100%正しく翻訳できるように、リライトが必要だと結論しました。

コンピュータが間違わずに翻訳できるような日本語とはどのようなものか。それについて考えたのが、冒頭に紹介した私の本でした。1万部以上買って下さったのですが、その後の機械翻訳の開発に影響することはありませんでした。機械翻訳ソフトを開発中の大手企業のご担当者にも声をかけたのですが、難しいから挑戦するのだという意気揚々たるものがありましたから、関心と呼ぶことはできませんでした。私が作った英語的な日本語(Englic Japanese)をそれらのソフトに翻訳させてみたのですが、自然な日本語をどう処理するか改良の手が加えられ、うまく英語になりませんでした。いうなれば、さんまの塩焼きを食べたいのに、フランス料理のレシピになっていたようなものでした。けっきょく、私は機械翻訳のことから興味を失ってしまいました。

#### 海外から非難ごうごうの日本語情報の鎖国状態

英語になっていない日本語の文献や情報は膨大です。というよりは、英語に翻訳される日本語文献は、まさに大海の一滴のようなものです。外国人にとって有用な日本語情報は膨大にあります。学術研究でも、英文のアブストラクトがつけられる程度であって、英文で書かれる論文はきわめてわずかです。

明治以来、海外の情報や作品は次々と日本語化されてきました。学者の中心的な役割は、こうした海外の作品を邦訳して日本に紹介することでした。文学作品でも、これほど広い世界から翻訳された国はなく、英語になっていない文学作品を研究したい外国人は、日本語を勉強して読むということさえ現実です。

しかし、あまりに情報が多くなってどれを翻訳すればよいかの評価が難しくなり、一方で、難しい内容のものは売れないから出版社が取り上げないという事態になり、日本に海外の作品を紹介するという学者の仕事はなく

なりました。本来なら、ここで日本語の文献や作品を英語にして海外の役に立つというのが学者の任務なのですが、英文を書くことに長けていない日本人学者にはとうていできないことです。

とはいえ、グローバリゼーションの時代、放置しておいてよいわけはありません。わが国のことが理解されず、価値ある研究成果も世界に知られることなくお蔵入りです。日本に貴重な文献や情報があることを知っている外国人は、これらのものが日本語にならず放置されていることを強く非難します。日本にとっては、海外で日本のことが正しく理解されないという弊害を生んでいます。孤立状態です。

これを解決できるものとして、機械翻訳は期待されたわけです。しかし、その現実はあまりに悲惨な状況にあります。いま、私は25年ぶりに、機械翻訳の世界に戻ってきました。機械翻訳プログラムの開発に興味があるからではありません。わが国の海外に対する情報鎖国状況を打破しなければ、グローバリゼーションに日本が取り残されることがわかっているからです。

### 機械翻訳開発エンジニアとの悲しい会話

先日、ある大手コンピュータ会社で翻訳ソフト開発を担当しているエンジニアの一人に会って、話をする機会がありました。その会社の機械翻訳プログラムの正解率は70%程度ということで、毎年1%でも誤訳が減れば、評価されるということでした。毎年1%なら、30年かかることになりませんが、ものごとの原則から、90%を超えるとほとんど改良ができなくなります。それでも10人くらいの専門家が所属するその部門が存続できているのは、自社のパソコンに搭載されている翻訳ソフトの改良とメンテナンスが必要だからでしょう。不評ではあっても、止めるわけにはいきません。

私は言いました。70%の正解率であれば、正解できる日本語の種類をまとめて、その範囲で日本語原稿を書いたらよいのではないかと。上司に相談してくれたようですが、忙しいからそんなことをしている余裕はないということでした。

現在という一点から、将来に夢を馳せるというのはロマンチックではありますが、科学的ではありません。科学を持ち出さなくても、合理性を欠きます。なぜ出口から考えないのだろうと思うのですが、それは私一人なのでしょう。では、機械翻訳の出口とは何か。それについて、長くなってしまいましたので、簡単に述べておきます。またいずれこの問題について発表できる機会があることを願っています。

### 機械翻訳のリバース思考

誤訳率の小さい機械翻訳ソフトができれば、海外に向けてもっと日本の文献や情報を海外に伝えることができるだろう。英語が苦手な人も簡単に英文の手紙やメールを海外の知人に送ることができる。それは日本を大いにグローバル化する。そう考えるのが一般的な思考です。

しかし、逆から考えれば、膨大に存在する日本語情報や文献を、安く、大量に、すばやく英語にして海外に紹介する必要に迫られています。それにはどうしなければならないか。そこから、現在の機械翻訳の閉塞を打ち破ることができるかと信じています。それは、25年前に私が書いた「英語もどきの日本語 (Englic Japanese)」です。これはプリエディティंगではなく、リライトイングです。では、それを誰がやるのか。私はこの仕事を高齢者の人たちにしてもらえば、収入を得る新しい道を用意できていると思っています。英語を書く能力を身につけてもらうことは至難のことですが、Englic Japanese ならちょっとした講習で習得可能です。

問題は、Englic Japanese を完全に近い状態で翻訳できるソフトがないことです。しかし、これは現在正解率

70%でしかない機械翻訳ソフトの実用化を図ることよりもずっと簡単なことです。正解できる範囲の能力を使えばよいのですから。

この私の意見がどなたかの目に止まることを祈っています。もし共鳴いただける方がおられれば、ご一緒にこの問題を考えたいと思います。経済産業省なども関心を寄せています。レスポンスをお待ちします。

2009/10/19 平田 周

---

## 会員名・所属

吉次基宣（東京ドイツ文化センター図書館）／YOSHITSUGU Motonori

---

## 自己紹介

私は、東京ドイツ文化センターの図書館で司書として働いています。東京ドイツ文化センターは、ドイツ語の普及とドイツの文化の紹介を目的とする国際文化交流機関です。東京にいながら、ドイツ語が飛び交い、ドイツが体験できるドイツへと開かれた窓というような場所です。私はここでドイツ語とドイツ文化への愛情をもって日々過ごしています。そういうわけで、私は、翻訳とか通訳に深い関心を持っています。ただ機械翻訳の開発となると、大学での専攻がヨーロッパの現代哲学だったので、関心はあるものの全く無縁の存在でした。また、機械翻訳というものは、大学や企業の大きなプロジェクトとして開発されるものと思っていたのです。ところが、個人でコツコツと開発に取り組んでいる人たちがいるのです。それを知ったのは、柴田勝征著『C言語による英和翻訳システム』1990 ラッセル社、を読んだからです。これにはソースコードも載せられていました。これを読んで「ひょっとしたら私でも何かできるのかも知れない」と考えたのです。そんな折、信州大学でインターネットで情報工学を学べる大学院があることを知りました。こうした流れに勇気を得て、このインターネット大学院で情報工学を学んで、独日機械翻訳システムを開発してみようと決意したのです。その後4年をかけて大学院の修士課程の科目を修了し、修士論文の課題として、「独日機械翻訳システムの開発」を取り上げました。初めは上記の本のソースコードを分析しそれをもとに翻訳システムを構築しようとしたのですが、人の書いたソースコードを読み取り、それをもとにプログラムを書くのは容易ではなく、結局、4年間で学んだ知識に基づいて独自のプログラムを書き、指導教授の親身の指導に助けられながら、何とかシステムの一部を実現し、これが、修士論文となりました。それは、PHPで書かれたネット上に公開できる簡単なドイツ語を日本語に翻訳できる一応の機械翻訳システムと言えるものでした。こうして本来文系の人間であった私が、情報工学の修士資格を取り、また大学院で学ぶ前には、考えられないような翻訳システムを部分的ではあるとはいえ実際に実現してしまったのです。それは私にとって大きな喜びであり、また、驚きでもありました。これは、今から二年前のことでした。その時AAMTに入会をしたのです。それは、これまで学んだことを深め、翻訳システムの開発にさらに取り組んでいくためです。AAMTからは、多くの貴重な情報をいただいておりますが、今後はもっと積極的にAAMTの活動に参加させていただきたいと思っています。

---

## 機械翻訳および翻訳業界に期待すること

機械翻訳システムの開発は、非常に難しい課題です。それは、翻訳に係わる言語というものがかなり厄介な代物だからです。現代の哲学でも言語が一つの中心問題として受け止められ、我々に最も身近なものでありながら、その謎はますます深まっているという状況です。機械翻訳システムは、このような言語を操作し、処理しようと

するものなので、元来我々が予想する以上に難しい課題なのです。ですから、これをプロジェクトとして開発しようということ自体が、言語の厄介な性格と合わないと言えるかも知れないのです。翻訳システムの開発には、今後も多くの努力と時間が必要となるでしょう。機械翻訳の開発に関わっている方々には、今後も息の長い粘り強い努力を期待します。また、言語をめぐる現代哲学の展開にも注目していただければと思います。

---

#### AAMT へのご要望

AAMT からは、日ごろ多くの貴重な情報をいただき、感謝しております。AAMT の会合に出席させていただきますと、なんらかのプロジェクトに参加されている方々のお話を伺います。それは大変興味深いものですが、私は、個人として翻訳システムを作ることを試みているので、そのような個人が自由に参加できるフォーラムのようなものがあればありがたいと思っています。

---

## 第 19 回通常総会および関連行事の報告

AAMT 事務局

当協会の第 19 回通常総会が 2009 年 6 月 22 日（月）13 時 30 分より東京・竹芝のホテルアジュール竹芝にて開催されました。総会后、各委員会からの報告会、講演会、そして第 4 回 AAMT 長尾賞授与式と受賞者による記念講演会が盛況のうちに行われました。

### 第 19 回通常総会

1. 開会の辞
2. 会長挨拶（独）情報通信研究機構 上席研究員 井佐原 均
3. 来賓挨拶
4. 出席会員の確認
5. 議事録署名人の選出
6. 議案
  - 第 1 号議案 2007 年度事業報告（案）
  - 第 2 号議案 2007 年度決算報告（案）
  - 第 3 号議案 2008 年度事業計画（案）
  - 第 4 号議案 2008 年度収支予算（案）
  - 第 5 号議案 役員改選について（案）
  - 第 6 号議案 会則修正案について（案）
  - その他・会員提案事項
7. 閉会の辞

### 報告会

- |                       |                        |
|-----------------------|------------------------|
| 1. 機械翻訳課題調査委員会        | 委員長 後藤 功雄（(独)情報通信研究機構） |
| 2. AAMT/Japio 特許翻訳研究会 | 委員 下畑 さより（沖電気工業(株)）    |
| 3. 編集委員会              | 委員 熊野 明（(株)東芝）         |

### 講演会

講演 I：「常に時代の一步先を見つめて」

小谷 泰造氏（株式会社インターグループ 代表取締役）

講演 II：「Language Resources and Machine Translation in Korea」

Key-sun Choi 氏（KAIST）

（講演予定の Key-sun Choi 氏が欠席のため、井佐原 AAMT 会長が代理で講演）

#### 第4回 AAMT 長尾賞授与式・記念講演会

受賞者：(1) 東芝ソリューション株式会社プラットフォームソリューション  
株式会社東芝研究開発センター 知識メディアラボラトリー

受賞理由：ユーザ辞書の自動抽出から辞書の自動選択までの優れた技術開発により  
機械翻訳の実用度を大きく押し上げる功績が顕著なため。

受賞者：(2) AAMT インターネットワーキンググループ

受賞理由：AAMT 会員ならびに機械翻訳ユーザに最新の機械翻訳システム／サービス  
一覧を長年にわたり提供し続け、機械翻訳ソフト／サービスの普及に努  
め、機械翻訳の発展／啓蒙に寄与した功績が顕著なため。

選考委員長：飯田仁（東京工科大学）

選考委員：横山 晶一（山形大学）

Virach Sornlertlamvanich（Thai Computational Linguistics Laboratory）

村田 稔樹（沖電気工業(株)）



授賞式



記念講演会

AAMT インターネットグループリーダー

(株)富士通研究所 富士 秀氏



東芝ソリューション（株）伊藤悦雄氏

#### 懇親会

本会後の懇親会は、多数の参加者にお集りいただき、学識経験者、研究者、翻訳家等、幅広い分野の方々が活発な意見交換を行い有意義な交流の場となりました。

## 協会活動報告

(2009年6月～2009年10月)

### 第19回通常総会

2009年6月22日

- 第1号議案 2008年度事業報告(案) 第2号議案 2008年度決算報告(案)  
第3号議案 2009年度事業計画(案) 第4号議案 2009年度予算(案)  
第5号議案 役員選任について(案) 第6号議案 会則修正案について(案)  
その他・会員提案事項

### 報告会

2009年6月22日

- ①機械翻訳課題調査委員会 ②AAMT-Japio 特許翻訳研究会  
③編集委員会

### 講演会

2009年6月22日

講演Ⅰ：「常に時代の一步先を見つめて」

小谷 泰造氏 (株式会社インターグループ 代表取締役)

講演Ⅱ：「Language Resources and Machine Translation in Korea」

Key-sun Choi 氏 (KAIST)

(講演予定の Key-sun Choi 氏が欠席のため、井佐原 AAMT 会長が代理で講演)

### 第4回 AAMT 長尾賞授与式・記念講演会

受賞者：(1) 東芝ソリューション株式会社プラットフォームソリューション  
株式会社東芝研究開発センター 知識メディアラボラトリー

受賞理由：ユーザー辞書の自動抽出から辞書の自動選択までの優れた技術開発により  
機械翻訳の実用度を大きく押し上げる功績が顕著なため。

受賞者：(2) AAMT インターネットワーキンググループ

受賞理由：AAMT 会員ならびに機械翻訳ユーザに最新の機械翻訳システム/サービス一  
覧を長年にわたり提供し続け、機械翻訳ソフト/サービスの普及に努め、  
機械翻訳の発展/啓蒙に寄与した功績が顕著なため。

### 懇親会

2009年6月22日 ホテルアジュール竹芝 白鳳

## 決算理事会

2009年6月22日

- |       |               |       |               |
|-------|---------------|-------|---------------|
| 第1号議案 | 2008年度事業報告(案) | 第2号議案 | 2008年度決算報告(案) |
| 第3号議案 | 2009年度事業計画(案) | 第4号議案 | 2009年度予算(案)   |
| 第5号議案 | 役員選任について(案)   | 第6号議案 | 会則修正案について(案)  |
- その他・会員提案事項

## 機械翻訳課題調査委員会

2009年6月23日(2009年度 第3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) |
| ③活動内容についての議論   | ④活動内容の報告(各WGから)         |
| ⑤まとめと次回委員会について |                         |

2009年7月28日(2009年度 第4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) |
| ③活動内容についての議論   | ④活動内容の報告(各WGから)         |
| ⑤まとめと次回委員会について |                         |

2009年8月27日(2009年度 第5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) |
| ③活動内容についての議論   | ④活動内容の報告(各WGから)         |
| ⑤まとめと次回委員会について |                         |

2009年9月28日(2009年度 第6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) |
| ③活動内容についての議論   | ④活動内容の報告(各WGから)         |
| ⑤まとめと次回委員会について |                         |

## AAMT/Japio 特許翻訳研究会

2009年7月3日(2009年度 第3回)

- |            |                 |
|------------|-----------------|
| ①前回議事録の確認  | ②今年度の活動についてI~IV |
| ③次回の開催について |                 |

2009年7月30日

NTCIR・AAMT/Japio 特許翻訳研究会合同会合開催

- |                        |                  |
|------------------------|------------------|
| ①AAMT/JAPIO 特許翻訳研究会の紹介 |                  |
| ②NTCIRの活動について          | ③MTサミット・特許翻訳WS関連 |
| ④次回の開催について             |                  |

2009年9月7日(2009年度 第4回)

- |           |           |
|-----------|-----------|
| ①前回議事録の確認 | ②MTサミット報告 |
|-----------|-----------|

③今年度の活動について I～II

④次回の開催について

2009年10月9日(2009年度 第5回)

①前回議事録の確認

②今年度の研究内容について

③次回の開催について

#### 編集委員会

2009年6月22日 AAMT Journal No.45 発行

MLにて最新号の進捗打合せ及び次回委員会の開催日調整。

2009年10月27日(2009年度 第2回)

①AAMT Journal No.46 進捗確認

②No.47以降の企画について

③その他

#### インターネットWG

①AAMT ホームページの更新(毎月)

②AAMT Forum メイリングリストの管理

③AAMT Forum への情報発信(月数回)

④委員会活動におけるメイリングリストの管理

⑤会員専用ホームページ開設にむけた準備

⑥その他事務局ネットワークのインフラ管理全般

#### 【AAMT Journal No.45 訂正のご案内】

AAMT Journal No.45 の目次ページに誤植がございました。

お詫びとともに、下記の通り、訂正させていただきます。

(誤) AAMT 会員のひろば(第4回) → (正) AAMT 会員のひろば(第5回)

AAMT ジャーナル編集委員会委員長  
筑波大学大学院システム情報工学研究科  
知能機能システム専攻  
宇津呂 武仁

AAMT ジャーナル 46 号をお送りします。

当協会元会長の田中穂積先生が7月27日に急逝されました。今号では、AAMT初代会長長尾眞先生、前会長辻井潤一先生、現会長井佐原均先生よりご寄稿いただきました追悼文を掲載させていただきました。編集委員一同、先生のご冥福を心よりお祈りいたします。

今号では、第四回のAAMT長尾賞の選考結果を受けまして、受賞者である二組織よりご寄稿していただきました。まず、これまで長年に渡り、ユーザー辞書の自動抽出・辞書の自動選択の技術開発を進めて来られました東芝ソリューション株式会社の伊藤悦雄様に巻頭言をご執筆頂きました。また、インターネットワーキンググループのメンバーとして、長年に渡り機械翻訳の発展・啓蒙に寄与してきた富士通研究所の富士秀様よりご寄稿いただきました。

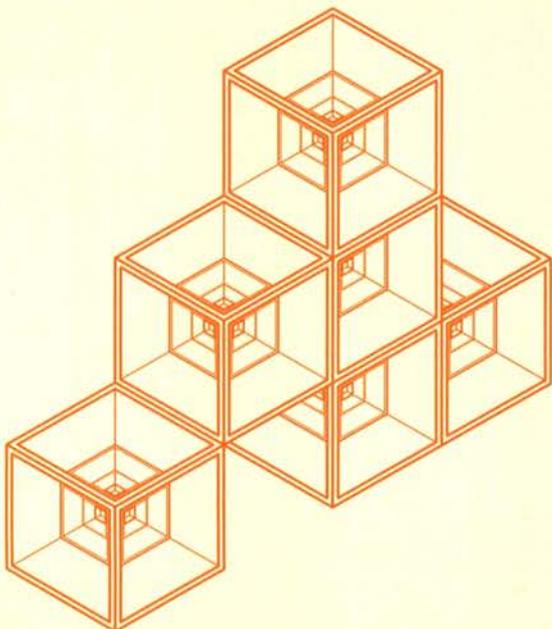
また、8月にオタワ(カナダ)におきまして開催されました第12回機械翻訳サミットへの参加報告を掲載しました。このサミットの併設ワークショップでは、AAMT/Japio特許翻訳研究会のメンバーが中心となって特許翻訳ワークショップを開催しましたので、このワークショップの報告も掲載しました。あわせて、同じく8月にシンガポールにおきまして開催されましたACL-IJCNLP (Joint conference of the 47th Annual Meeting of the Association for Computational Linguistics and the 4th International Joint Conference on Natural Language Processing of the Asian Federation of Natural Language Processing) への参加報告を掲載しました。

一方、AAMT内の活動報告として、機械翻訳課題調査委員会からご寄稿していただきました。その他、「AAMT 会員のひろば」の企画におきましては、個人会員の紹介文5件を掲載しました。

***MEMO***



# AAMT



## AAMTジャーナル No.46

---

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：〒619-0289 京都府相楽郡精華町光台3丁目5番地

独立行政法人 情報通信研究機構 言語翻訳グループ内

phone：0774-93-4625 fax：0774-93-4627

編集委員会：宇津呂 武仁 後藤功雄 大倉 清司

熊野 明 三浦 貢 村上 嘉陽

事務局：神崎 享子 喜田 あゆみ

印刷所：株式会社ナビックス

---

Asia-Pacific Association for Machine Translation

c/o Computational Linguistics Group, NICT

3-5 Hikaridai, Seika-cho, Soraku-gun, Kyoto, Japan 619-0289

Phone:+81-774-93-4625 FAX:+81-774-93-4627

URL:<http://www.aamt.info>