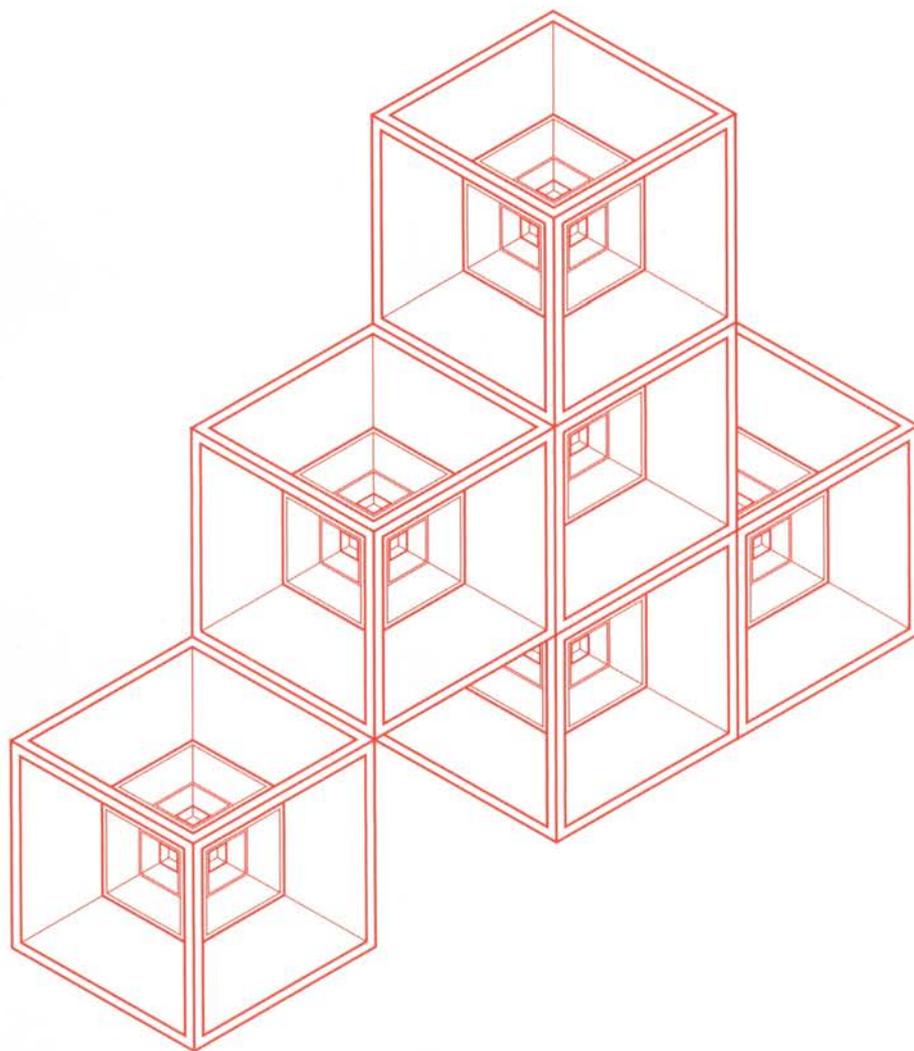


閱覽用

APMT

Asia-Pacific Association for Machine Translation

JOURNAL



March 2007 No.39

アジア太平洋機械翻訳協会

## 目 次

巻頭言：	AAMT 会長就任のご挨拶 .....	井佐原 均 .....	1
退任挨拶：	AAMT 会長退任にあたって .....	辻井 潤一 .....	3
海外報告：	アイルランドにおける MT 教育と MT 研究について (DCU の場合) ..	オヘイガン 統子、アンディ・ウェイ .....	5
シンポジウム報告：	第 16 回 JTF 翻訳祭—翻訳の原点、翻訳業界の共通認識— .....	野上 員生 .....	14
	『国際ワークショップ IWSLT 2006』報告 .....	隅田 英一郎 .....	16
	『GALA in Tokyo』参加報告 .....	村上 嘉陽 .....	18
業界展望：	今後の機械翻訳の目指す道 .....	(株) 富士通研究所 .....	23
役員リスト：	AAMT 新体制紹介 .....		24
新委員会紹介：	AAMT 機械翻訳課題調査委員会の活動について .....	内元 清貴 .....	26
新事務局紹介：	AAMT 新事務局紹介 .....	神崎 享子、和泉 絵美 .....	28
総会報告：	第 16 回通常総会および関連行事の報告 .....		29
翻訳システム：	2006 年 AAMT 長尾賞受賞に際して .....	日本電気 (株) .....	31
	OKI における機械翻訳技術開発 .....	沖電気工業 (株) .....	33
新製品紹介：	「The 翻訳サーバ Enterprise Edition V2.0」 .....	東芝ソリューション (株) .....	35
サービス紹介：	「My サイト翻訳」 .....	(株) 高電社 .....	38
MT 関連ソフトウェア：	インターネット機械翻訳サービスリスト .....	AAMT インターネット WG .....	40
活動報告：	協会活動報告 (2006 年 4 月～2007 年 2 月) .....		46
	MT SUMMIT XI 開催案内 .....		50
編集後記：	.....	宇津呂 武仁 .....	51

## CONTENTS

Foreword:	Message from the new president of AAMT .....	<i>H. Isahara</i> .....	1
Message:	Message from the former president of AAMT .....	<i>J. Tsujii</i> .....	3
Report:	Teaching and Researching MT in Ireland: The DCU Experience .....	<i>M. O'hagan, A. Way</i> .....	5
Symposium:	Report on "16th JTF Translation Fair" .....	<i>K. Nogami</i> .....	14
	Report on "IWSLT 2006" .....	<i>E. Sumita</i> .....	16
	Report on "GALA in Tokyo" .....	<i>Y. Murakami</i> .....	18
Future Outlook:	MT in the future .....	<i>Fujitsu Laboratories LTD.</i> .....	23
Board Members:	New AAMT organization .....		24
New Organization:	Introduction of the AAMT Committee for Seeking Future Direction of MT .....	<i>K. Uchimoto</i> .....	26
New Secretariat:	New AAMT Secretary-General .....	<i>K. Kanzaki, E. Izumi</i> .....	28
General Meeting:	Report on the 16th AAMT General Meeting and related events .....		29
MT System:	Report by the AAMT Nagao Award 2006 Winner: R&D toward practical use of speech translation system for hand-held devices .....	<i>NEC Corporation</i> .....	31
	Report by the AAMT Nagao Award 2006 Winner: R&D on MT technologies at OKI Electric Industry Co., Ltd. ....	<i>OKI Electric Industry Co., Ltd.</i> .....	33
New Products:	"The Honyaku Server Enterprise Edition V.2.0" .....	<i>Toshiba Solutions Corporation</i> .....	35
New Service:	"My Site Honyaku" .....	<i>Kodensha Co., Ltd.</i> .....	38
MT Web Links:	List of MT services on the Web (the Asia-Pacific Region) .....	<i>AAMT IWG</i> .....	40
AAMT Activities:	AAMT Activities (April 2006~February 2007) .....		46
	Announcement of MT SUMMIT XI .....		50
Editor's Note:	Message from the Chair of the AAMT Journal Editorial Committee .....	<i>T. Utsuro</i> .....	51

## AAMT 会長就任のご挨拶

独立行政法人情報通信研究機構 井佐原 均

本年度の総会において、辻井潤一前会長（東京大学教授）から、AAMT 会長を引き継がせていただきました。高名な諸先生方の務めてこられた職であり、身の引き締まる思いです。私自身、これまで技術動向調査委員会と編集委員会の委員長として AAMT の活動に関わってまいりましたが、これからは、小谷泰造副会長(株式会社インターグループ代表取締役社長)、飯田仁副会長(東京工科大学教授)を初め、理事・監事の皆様方、また、内元清貴機械翻訳課題調査委員会委員長(情報通信研究機構主任研究員)、宇津呂武仁編集委員会委員長(筑波大学助教授)、富士秀インターネットワーキンググループリーダー(富士通研究所)のご協力を得て、微力ながら、AAMT の一層の発展のお手伝いをさせていただきます。辻井潤一前会長にも引き続き、AAMT/Japio 研究会の委員長としてご尽力いただきます。

今年度から、従来の技術動向調査委員会と市場動向調査委員会を統合して、機械翻訳課題調査委員会として、新たな体制での活動を開始いたしました。長年、市場動向調査委員会の委員長を務めていただきました坂本義行先生に、この場を借りて御礼申し上げます。

インターネットで世界が繋がり、言語の異なるコミュニティとの情報交流も盛んになっています。しかし誰もが多くの言語を自在に操れるわけではありませんから、いわゆる言語障壁が大きな問題となります。言語障壁の解消を目指して、より高精度の機械翻訳システムの実現や、機械翻訳システムの活用法の確立が求められています。

機械翻訳システムを取り巻く環境としては、近年、大量の言語資源が利用可能となり、計算機のパワーも増大しました。これにより、より実用的な機械翻訳システムを開発できる可能性が高まっています。また、パッケージソフトをパソコンにインストールして使うといった使い方だけではなく、ホームページでの翻訳サービスのような新しい機械翻訳の使われ方も現れ、機械翻訳のユーザを広げています。

翻訳は言葉の違いを繋ぐ輪の役割を果たしますが、アジア太平洋機械翻訳協会（AAMT）は、さらに二つの輪を結びたいと思います。一つはアジア太平洋機械翻訳協会の名が示すように、アジア太平洋の国々における機械翻訳関連の活動を支援し、繋げていこうと思います。ご存知のように、アジア太平洋機械翻訳協会は1991年の設立以来、日本を中心にその活動を進めてきました。機械翻訳システムの開発のブレークスルーが予想される今、よりアジア圏全体に開かれた AAMT とするために、まず中国に China Branch の設置を目指します。

もう一つの輪として、AAMT は、さまざまな場面で機械翻訳に関わっている方々が、相互に交流できる場を目指します。開発された機械翻訳システムも使われなくては意味がありません。システムを開発する人とシステムを使う人との両輪で翻訳という知的作業が実現されます。AAMT は、この両輪を繋いで、機械翻訳システムが実際に広範囲に使われる環境の構築を目指します。

究極の機械翻訳システムは何でしょうか？そのひとつの方向は機械翻訳の存在を意識しない環境の実現でしょう。ネットワークや検索システムに組み込まれて、自然と多言語の世界を彷徨できるような環境の構築が考えられます。また、言語で記述された事柄だけを翻訳するのではなく、身振りや表情などの非言語情報を翻訳したり、さらには知識を直接変換するようなシステムの開発も新たな方向として考えられるでしょう。

人は、その知的活動や、感性の処理の多くを言葉によって行っています。AAMTは、このような人の活動の中核をなす言語を用いた情報交換を支援するシステムとしての機械翻訳システムの開発と活用を支援していきます。今後とも、皆様のご協力の下、言語障壁の克服に向けた活動を進めてまいりますので、よろしくお願いいたします。

## AAMT 会長退任にあたって

東京大学大学院 辻井 潤一

田中穂積前会長（元東工大教授、中京大教授）から AAMT 会長の任を引き継いだのは、1999 年の MT Summit(シンガポール)の翌年、2000 年 6 月でした。以後 6 年間、2005 年に再び MT Summit をアジア地区(プーケット島、タイ王国)で開催し、慣例にしたがって、本年度の総会で、井佐原均先生に会長を引き受けていただくことになりました。この間、2003 年のニューオーリンズでの Summit から 2005 年まで、国際機械翻訳連盟(IAMT)の会長も、つとめさせてもらいました。6 年間、機械翻訳に携わる方々、また、AAMT 事務局の人たちとともに、非常に楽しく、また、有意義な時間を過ごすことができました。

会員皆様のご尽力、ご協力のおかげと感謝しています。

この 6 年間は、非常に多くの出来事があった時期でもあります。シンガポールの次の Summit(2001 年)は、スペインのサンチャゴという、カソリック教の聖地、巡礼の町として栄えた町で開催されました。私が AAMT の会長として参加した最初の Summit です。

サンチャゴは、スペインの中でも、中世世界、キリスト教世界の象徴のような、古い町並みが印象的な町でした。ただ、この会議の直前に、9.11、ニューヨーク WTC のアタックがあり、日本からの参加者が激減しました。私がスペインに向かう前日のことで、会議中も、テレビでは 24 時間、事件の報道が続き、参加者全員が釘付けになるという具合でした。また、米国からの参加者は、空港で足止めされ遅れて到着するなど、事件の影響が非常に大きかった会議です。

多言語の情報収集は、それ以前から米国の研究プロジェクトとしてありましたが、この事件以後、

その必要性が強く認識され、機械翻訳に大きな研究投資が行われるようになります。振り返ってみると、機械翻訳分野にとっても、影響が非常に大きな事件でした。

統計翻訳の世界的なリーダーである Herman Ney 教授が、初めて MT Summit に出席したのも、このサンチャゴ会議で、2005 年のプーケット会議で会ったとき、彼が「スペインの会議では、統計翻訳はまだ新参者で好奇の目で見られていた。それが、プーケットでは、招待講演、パネル、一般発表にいたるまで、統計翻訳一色だ」と言っていたのが印象的でした。この 6 年間は、機械翻訳の手法に革命的な変化があった 6 年でもありました。

テロや局地紛争など、その時点での世界の政情にあわせて迅速に、それまでの研究蓄積のない言語を対象とする翻訳システムが必要となる。このことから、対象言語の文法を構築するといった、まどろっこしいシステムはだめで、大きな言語データさえあればとりあえずのシステムが構築できる統計翻訳がよい。こういう 9.11 以後の政治的背景が、研究を加速した、という側面があります。Ney 教授が初めて参加した会議、あるいは、9.11 とその後の翻訳技術の変革など、サンチャゴでの会議は、いろいろな意味で重要な会議となりました。

個人的には、この会議は、イタリアの Antonio Zampoli 教授と直接会う最後にもなりました。Zampoli 教授は、国際計算言語委員会(ICCL: International Committee of Computational Linguistics)の中核的なメンバーであり、また、言語資源と評価の国際会議(LREC)の創始者で、私が尊敬する先輩研究者の一人です。

Zampoli 教授には、AAMT が組織したシンガポールの Summit にも参加していただき、楽しい時間

を過しました。サンチャゴ会議にも出席していただいたのですが、夕食後、もう暗くなった中、聖ヤコブ大聖堂前の広場を足早に歩いているとき、教授が「最近、めまいが激しく、体調がよくない」と話しておられたのが、今も記憶に強く残っています。教授は、その翌年、私が出席できなかった LREC 会議で倒れられ、その後、国際会議などには出席されることなく、亡くなりました。教授は、統計的な機械翻訳の源流ともいえるデータに基づく言語学(コーパス言語学)を非常に早くから推し進められた方ですが、その業績をたたえた Antonio Zampoli 賞の第一回受賞者(2004 年 LREC, リスボン)は、IBM 統計翻訳の創始者である F.Jelinek 博士でした。

2003 年、2005 年の会議は、ニューオーリンズとプーケット島という、大きな自然災害に見舞われた町での開催となりました。2003 年 9 月当時のニューオーリンズは、夜も暖かく、深夜までジャズが流れる陽気な町でした。夜はスパイシーな Cajun Food とジャズ、怪しげな呪術のお店、Mardi Gras のカーニバルが想像できる喧騒、昼間は、「風と共に去りぬ」の舞台そのままのプランテーション屋敷と、北米とはかなり異なった文化に囲まれたくつろいだ会議となりましたが、その後、陽気でいかがわしい町ニューオーリンズは、台風カテリーナで様相が一変したようです。

プーケットは、逆に開催の前年に大津波があり、9ヶ月あまりの時間はあるものの、本当に予定通り開催してよいのかどうかなど、主催者としては相当に頭を悩ませました。実際には、予想以上に早く津波による被害から回復していること、また、会議を開催することが現地の人たちへの側面からの支援となるとの判断で、予定通り開催しました。

会議は、現会長の井佐原さん、また、Virach さん(Virach Sornlertlamvanich)をはじめとする現地の研究者のおかげで、非常な成功でした。特に、タイの人たち特有の、少しはにかむような笑みと相手

へのこまやかな心遣いなど、心理的・物理的に残っているであろう傷を感じさせない会議運営で、アジアとしての特徴がある会議となりました。

シンガポールからプーケットまでの6年間は、アジアの存在感が加速度的に増大した6年間でもありました。プーケットの会議では、中国・インド・パキスタン・インドネシア・シンガポールの研究者など、幅広い地域からの参加があり、これまでの日本・台湾・韓国に加えて、アジアの研究が占める割合が格段に増大した、という印象がありました。アジアは、研究の多様性、層の厚さなどの点で、姉妹組織 AMTA の北米、EAMT のヨーロッパと同等、あるいは、それらを凌駕するところまで成長したと、いってよいでしょう。

このアジアの存在感の増大は、これからも、さらに加速されるでしょう。それに伴って、AAMT の重要さも増すとおもいます。ただ、ヨーロッパ・北米と異なり、アジアの持つ言語の多様性、文化的な広がり、AAMT の組織運営にさまざまな難しさももたらすことでしょう。

機械翻訳技術の方も、サンチャゴからニューオーリンズ、プーケットへと一直線に発展してきた統計翻訳の枠組みが一段落し、次の飛躍を必要としているようです。AAMT 初代会長の長尾眞先生、IAMT 前会長の Ed Hovy 教授など、プーケット会議に参加した多くの研究者が、同じ印象を持たれたようです。

プーケット会議に参加した研究者の中に、サンチャゴの Ney 教授にあたる人たちがいたのかも知れません。次の6年も、この6年と同じ、あるいは、それ以上に、機械翻訳には大きな飛躍と変化があるでしょう。大いに期待していると同時に、その飛躍と変化、困難の解決に私なりの寄与をできれば、と思っています。

6年間、ありがとうございました。

## Teaching and Researching MT in Ireland: The DCU Experience

Minako O'Hagan, SALIS  
Andy Way, School of Computing  
[minako.ohagan@dcu.ie](mailto:minako.ohagan@dcu.ie)  
[away@computing.dcu.ie](mailto:away@computing.dcu.ie)

### 要旨

本稿では、著者の所属する国立ダブリン・シティ大学 (DCU) における MT 関連コースならびに研究の概要を紹介する。本学はバイオテクノロジーをはじめ、先端技術や技術革新をめざした教育、研究に力を入れており、コンピュータ学部における EBMT 研究、文学部でも MT ならびに翻訳支援技術の教育、研究が活発に行われている。翻訳専攻課程は学部、院レベルとも翻訳テクノロジーが必須科目となっており、機械翻訳を含む技術の理論、実践を取り入れた、日本ではまだみられない一環した翻訳者教育が行われている。さらに、本学では外部研究者の受け入れ体制も整っており、アイルランド国内で唯一日本語を専攻できることから、今後さまざまな形での日本との交流が期待される。

### 1. Overview

This article provides a snapshot of how MT and related technologies are taught and researched at Dublin City University (DCU), where both authors teach and conduct research in these areas. We will focus on translator training and research from the humanities perspective, with some reference to Japanese language students taking relevant courses, as well as MT research conducted in the School of Computing. DCU is the only Irish university offering 4-year undergraduate degree courses such as Language for International Communication (LIC) in which the students can choose Japanese as the major language and combine it with TS.

Together with Trinity College Dublin and University College Dublin, DCU forms part of the Irish national university network based in Dublin, Ireland. Having just celebrated its 25<sup>th</sup> year of existence, and being of moderate scale with about 10,000 students, DCU is known for its clear focus on technology-oriented disciplines where substantial external funding is drawn from the major funding bodies such as the European Union and Science Foundation Ireland.<sup>1</sup> The School of Computing at DCU has a large number of postdoctoral researchers and postgraduate students funded by such external sources, including 11 students actively engaged in MT research. In the context of translation, the School of Applied Language and Intercultural Studies (SALIS)<sup>2</sup> forms part of the joint Faculty of Humanities and Social Sciences at DCU and is known for internationally active TS research with the flagship postgraduate programmes being the Graduate Diploma in Translation Studies (GTS) and the Masters in Translation Studies (MTS). The year-long GTS programme typically attracts around 40 students from all over Europe and a small group of students from outside the continent. In keeping with the university's technology orientation, in addition to more traditional TS research areas, SALIS houses researchers working in the area of language and technology, including corpus linguistics, localisation, multimedia translation and CALL (Computer-aided Language Learning). Among some 20 PhD students currently registered at SALIS, nearly half are engaged in TS research of whom 5 are conducting translation research involving technology in some way.

### 2. Computer-aided Translation Training in Europe

As evident in the earlier initiatives such as the Language Engineering for Translators' Curricula (LETRAC) proposed in the late 1990s (Badia, 1999), academics in Europe have been concerned with the gap developing between professional requirements in the industry and translator training at tertiary institutions, chiefly in the

<sup>1</sup> <http://www.sfi.ie>

<sup>2</sup> The School is divided into three sections, comprising Language Studies, Intercultural Studies and Translation Studies where students can choose from German, French, Spanish, Japanese, Irish and introductory Chinese.

area of translation technology. The latter now forms an integral part of translator training in many universities in Europe with an increasing number of TS courses incorporating varying translation technology components. This situation contrasts with Japan where TS is only just emerging as an independent academic discipline with a handful of universities offering TS courses where the full inclusion of Computer-aided Translation (CAT) as part of translator training is something yet to happen. With the fundamental need for multilingualism and wider availability of CAT tools, the European education authorities in the area of translation appear to have a much greater awareness of the need for CAT than their counterparts in Japan. Because of the double-byte character requirements as well as other pertinent linguistic issues, there have tended to be delays for major CAT tools developed in Europe to be fully localised for the Japanese language environment, creating further gaps in the penetration of tools even in the translation industry, let alone in academia where a small number of institutions offer translator training. Having said that, there are Japanese-specific tools developed by a number of Japanese manufacturers<sup>3</sup> but these are almost unknown outside of Japan even among users who require Japanese language tools. One of the reasons for this is the fact that the tools developed in Japan tend to be designed to handle English and Japanese or Japanese and other Asian languages rather than European language pairs required by users in Europe.

### **3. Translation Technology Teaching at SALIS**

At SALIS the Translation Technology module is compulsory for students taking the TS stream for the undergraduate LIC degree as well as the postgraduate GTS programme. This is a double-weighted module which runs for 12 weeks (one semester) with 18 hours' plenary lectures for both courses together with 24 computer lab hours for the postgraduates and 12 for the undergraduates. Face-to-face in-class teaching is also facilitated by the use of a Virtual Learning Environment (VLE) called Moodle<sup>4</sup> embraced by DCU campus-wide since 2003. Via Moodle the students can access the course content ranging from lecture presentations to lab notes and articles for reading, from home as well as on campus. Using Moodle facilities, online discussion forums are set up from time to time for students to post their comments on certain selected topics. SALIS has a dedicated computer lab with 20 recently upgraded desktop computers running WindowsXP loaded with a number of CAT tools, including Translation Memory products such as Trados/SDLX and StarTransit as well as localisation software such as Déjà Vu. We have access to a Systran Enterprise translation server, including Japanese-English and SALIS is also a participant in academic programmes run by some of the major CAT software vendors.

The Translation Technology module is divided into Translation Memory (TM) and Machine Translation (MT) with the postgraduate students required to complete two main assignments for each topic whereas the undergraduate students are required to choose just one. The TM component focuses on one particular translation memory product and the students learn how to translate using TM. The lectures cover the principles of segmentation, different types of matches and alignment. The terminology component is covered by a separate dedicated terminology module (another compulsory double-credit module) for both courses, and thus is not covered by the Translation Technology module. The lab sessions take the students through the main features of TM with clearly laid-out exercises each week. At the end of the course, the students are expected to be able to handle any TM tool with relative ease. The MT component starts with the historical background and covers the principles of the main MT architectures before delving into well-known linguistic issues which pose challenges for MT (Kenny & Way, 2001). The lab hours provide practical hands-on sessions to test freely available online MT systems as well as some dictionary work to varying degrees.

For the 2006 academic year, there were about 60 students in total registered for this module with the numbers nearly equally divided between undergraduates and postgraduates. Their nationalities included German, French, Spanish and Japanese as well as the local Irish students with many of them competent in 3 languages. There were 6 undergraduate students and 1 postgraduate with Japanese as their main language. The module is designed as a language-independent course although it is inevitable that language-specific issues are discussed for MT with examples given in German, French, Spanish as well as Japanese where possible. The MT assignment topics chosen by these students with Japanese included the investigation of adversative passive, giving and receiving verbs, anaphora and homograph resolution by using test suites to systematically test chosen online MT systems. Given that these humanities students do not delve into programming, their assignments focus on 'black box' rather than 'glass box' evaluation, which nevertheless triggers a sufficiently deep investigation of the chosen

---

<sup>3</sup> AAMT's translation tool site (<http://www.aamt.info/localportal/japan/list2005.html>) lists a number of such products.

<sup>4</sup> <http://moodle.org>

linguistic features. The first author has also been involved since 2005 in translation technology training for teachers of translation at the external summer seminar series organised by the Consortium for Training Translation Teachers (CTTT)<sup>5</sup>. These experiences endorse the merit of including translation technologies to enrich translator training.

#### **4. Translation Technology Research at SALIS**

Two recently submitted PhDs undertaken at SALIS in an MT-related area concern the investigation of Controlled Language (CL) applied to technical documentation with one addressing post-editing effort in terms of technical, temporal and cognitive dimensions (O'Brien, 2006) and the other focusing on the comprehensibility, usefulness and acceptability of the use of MT at a multinational company when CL rules are applied in the authoring stage (Roturier, 2006). Both these areas of research were undertaken with close links to industry, in particular, with the latter forming the first doctoral research in SALIS wholly supported by a commercial company. This research addressed the effectiveness of the use of CL rules in the source English document which in turn was translated by MT into German and French. The effectiveness was measured by drawing on an end-user survey conducted on the company's website.

Another example is the recently completed MT research 'MovieTrans' project which was funded by the Enterprise Ireland Commercialisation Fund (Proof of Concept) and undertaken in SALIS and the School of Computing involving the authors.<sup>6</sup> This feasibility study focused on the development of an Example-based Machine Translation (EBMT) system (Armstrong et al., 2006b; Stroppa & Way, 2006) and its evaluation, undertaken as an interdisciplinary study, combining expertise in the areas of multimedia translation and corpus linguistics in addition to natural language processing. The project investigated the scope of the application of EBMT to translate subtitles of DVD film and bonus materials for English-German and English-Japanese. The project involved the creation of subtitle corpora, development of the EBMT system, testing of various MT evaluation methods and an assessment of the productiveness of different data types to seed EBMT (Armstrong et al., 2006a). The project was motivated by the fact that professional human translators working in the area of DVD subtitles are facing diminishing deadlines and lowering rates of pay due to fierce competition to cut costs, which pointed to the direction of a technology-based solution (O'Hagan, 2003). The underlying objective of this project was to test the feasibility of developing a computer-based subtitling tool incorporating EBMT in the face of the increasing availability of human produced subtitles on DVD for use as a homogeneous corpus of training material for the system.

The test results based on automatic evaluation metrics suggested that a homogeneous corpus consisting solely of subtitles worked better than a heterogeneous corpus made up of general corpus sources such as Europarl and the Japanese corpus by (Utiyama and Isahara, 2003). By training the system on 10K sentence pairs from the homogeneous corpus, we achieved almost 50% better results than training the system on 40K sentence pairs from the heterogeneous Europarl corpus (Koehn, 2005) for English—German system (Armstrong, 2007). For this project we were able to apply different evaluation approaches of both automatic metrics and human-based evaluation involving text-only and text in picture formats (Armstrong et al., 2006c). We also explored a small scale pilot web-based evaluation, using Moodle, with film clips containing EMBT subtitles, eliciting end-user feedback. The latter evaluation compared our raw EBMT output, raw Babelfish output and post-edited EBMT. The post-editing was performed under an extremely tight timeframe and by a non-native speaker of the target language, in part simulating industry contexts under which DVD subtitles are produced. The survey indicated a degree of willingness for the post-edited EBMT subtitles to be used on commercial DVD products. We are hoping to continue with the project at a later stage to pursue the option of actual integration of EBMT into an existing subtitling system to measure the improvement of human throughput in terms of quantity and quality as well as speed.

#### **5. Machine Translation Teaching in the School of Computing**

In 1991, DCU ran a BSc degree in Applied Computational Linguistics, jointly administered and taught by our two schools. However, in 2006 we produced our last graduates from this degree. Owing to a decrease in numbers applying for IT-related courses in Ireland as well as in the rest of the world, it proved no longer feasible to run the course with fewer than 20 incoming students. While the degree was running, the second author taught two

---

<sup>5</sup>For the 2006 programme, see: [http://isg.urv.es/cttt/istanbul\\_2006/program.html](http://isg.urv.es/cttt/istanbul_2006/program.html).

<sup>6</sup><http://www.nclt.dcu.ie/mt/dvd.html>

final year modules in MT: one on the more traditional rule-based (transfer and interlingua) approaches, and the other focussing on the leading corpus-based approaches of today, namely Statistical MT (SMT) and EBMT.

With the demise of the degree, no students in the School of Computing are exposed to MT. However, we have recently obtained a grant to examine the feasibility of a 4<sup>th</sup>-level graduate school in NLP with the other Dublin-located Universities as well as the University of Limerick, and if there is sufficient demand to run such a programme, then MT will return as a core component of any such course.

In the meantime, these courses have continued to be taught by the second author at the Swedish Graduate School in Language Technology, the Edinburgh-Saarbrücken Summer School in Computational Linguistics and Psychology, and at the University of the Basque Country in Donostia, Spain.

## 6. Machine Translation Research in the School of Computing

While we are perhaps best known for our work on Example-Based Machine Translation, due possibly to the second author having co-edited the leading book on the subject in our field (Carl & Way, 2003), we carry out an extremely wide range of research in MT in the National Centre for Language Technology,<sup>7</sup> including:

- EBMT
- SMT
- Hybrid MT
- Tree-Based Translation
- Discriminative MT
- Constraint-Based MT
- Extending Online Rule-Based MT Systems
- MT Evaluation
- Sign Language MT

### 6.1 EBMT

The MT group at DCU is probably best known for its work in the area of EBMT. Building on the Gaijin system (Veale & Way, 1997), we have developed perhaps the leading EBMT system around today and one of the best-performing corpus-based systems anywhere (cf. the performance of the system (Stroppa & Way, 2006) in the recent IWSLT-06 evaluation at ATR in Kyoto). This work uses marker (or closed-class word) tags to segment source and target sententially aligned data and in addition to word alignment probabilities and cognate information, derives high-quality source-target sub-sentential alignments.

The MaTrEx system currently translates between English and a number of languages, including French (Gough & Way, 2004; Way & Gough, 2003, 2005a/b), Spanish (Armstrong et al., 2006b), German and Japanese (Armstrong 2007; Armstrong et al., 2006a), Italian and Arabic (Stroppa & Way, 2006), Basque (Stroppa et al., 2006), Irish Sign Language (Morrissey & Way, 2005, 2006), and Dutch (van den Bosch et al., 2007).

### 6.2 SMT

SMT researchers have found it difficult to show improvements when syntax is incorporated into their systems (Koehn et al., 2003). However, recently we showed (Hassan et al., 2006) that adding supertags (Bangalore & Joshi, 1999) can improve a basic phrase-based SMT system by a significant margin on Arabic—English data. In ongoing work, we intend to provide even more contextual data by integrating CCG syntactic constraints into our PBSMT system.

### 6.3 Hybrid MT

As far as we are aware, we are the only researchers who have explicitly compared EBMT and SMT systems. In (Way & Gough, 2005b), we showed that our marker-based EBMT system was capable of outperforming a word-based SMT system when trained and tested on an English—French TM. In (Groves & Way, 2005, 2006), we

---

<sup>7</sup> <http://www.nclt.dcu.ie/mt/>

demonstrated that the same EBMT system produced better quality translations than a phrase-based SMT system on the same data sets. In contrast, on the Europarl data (Koehn, 2005), the PBSMT system outperformed the EBMT system. Perhaps more interesting, however, was that we showed that combining the EBMT and SMT sub-sentential alignments into novel hybrid MT systems resulted in improved translations over the baseline systems from which they were derived.

#### *6.4 Tree-Based Translation*

The merits of combining the positive elements of the rule-based and data-driven approaches to MT are clear: a combined model has the potential to be highly accurate, robust, cost-effective to build and adaptable. While the merits are clear, however, how best to combine these techniques into a model which retains the positive characteristics of each approach, while inheriting as few of the disadvantages as possible, remains an unsolved problem. One possible solution to this challenge is the Data-Oriented Translation (DOT) model (Hearne, 2005; Hearne & Way, 2003, 2006) based on Data-Oriented Parsing (DOP) (e.g. Bod et al., 2003) and combines examples, linguistic information and a statistical translation model.

In this work, we seek to establish how the DOT model of translation relates to the other main MT methodologies currently in use. We find that this model differs from other hybrid models of MT in that it inextricably interweaves the philosophies of the rule-based, example-based and statistical approaches in an integrated framework. We look to the innovative solutions developed to meet the challenges of implementing the DOP model, and investigate their application to DOT. This investigation culminates in the development of a DOT system; this system allows us to perform translation experiments which are on a larger scale and incorporate greater translational complexity than heretofore. Our evaluation indicates that the positive characteristics of the model identified on a theoretical level are also in evidence when it is subjected to empirical assessment. For example, in terms of exact match accuracy, the DOT model outperforms an SMT model trained and tested on the same data by up to 89.73% (Hearne, 2005).

#### *6.5 Discriminative MT*

In recent joint work with Antal van den Bosch, we describe a novel approach to EBMT where the sub-sentential alignments are marker-based chunks, and the decoder is a memory-based classifier (van den Bosch et al., 2007). While translation quality is less than that obtained via PBSMT, the system is extremely fast: compared to 23 hours' training (for 728K Dutch-English Europarl examples) and 30 minutes testing (1500 sentences) for PBSMT, the trigram version of our memory-based classifier needs just 29 minutes to train, and 7 seconds to test. Where speed is much more important than quality in the overall translation cost, therefore, this new research presents a reasonable alternative over mainstream methods.

#### *6.6. Constraint-Based MT*

Data-Oriented Parsing (DOP: Bod et al., 2003) is a hybrid, language-independent, parsing formalism. Combining rules, statistics and linguistics, all parsing knowledge is learned from existing texts. However, the expressive power of the DOP model is limited by the corpus representations it assumes. DOP makes use of context-free phrase-structure trees which characterise phrasal and sentential syntax, but cannot reflect linguistic phenomena at deeper levels. The integration of Lexical Functional Grammar (LFG: Bresnan, 2001), which is known to be beyond context-free, enables a development of a model which produces more linguistically detailed descriptions of language, known as LFG-DOP (Bod & Kaplan, 2003). Because of difficulties inherent in building an LFG-DOP system whose probability model corresponds to the probability distribution of derivations, there are currently no satisfactory implementations.

The GF-DOP model that we have recently developed (Finn et al., 2006) can be seen as an extension of the Tree-DOP model, and an approximation towards LFG-DOP. It combines the robustness of the DOP model with some of the linguistic competence of LFG. This model exploits a corpus of annotated c-structures: features are extracted from f-structures and appended to the c-structure category labels. We accurately identify constituent features and functions, and improve the quality of c-structures generated by modelling this grammatical information. The GF-DOP model improves over the Tree-DOP model in that it uses additional grammatical information to rule out derivations which the Tree-DOP model would consider valid. One weakness of the GF-DOP model is that it may be slightly less robust than the Tree-DOP model, but if necessary, we can back off to

unannotated fragments to generate a parse. Having shown initial promise for parsing, we are currently experimenting with GF-DOP for MT.

### *6.7 Extending Online Rule-Based MT Systems: TransBooster*

Most existing wide-coverage MT systems tend to handle simple or short sentences better than linguistically complex ones. The most important reason for this is that linguistic analyses are costly to develop and that shallow parses are often preferred to deep ones for reasons of robustness. TransBooster is a wrapper technology that operates on top of a baseline MT system. It does not perform automatic translation itself, but helps the baseline MT system improve its own possibilities through input simplification. TransBooster is a paradigm-independent approach to improve the output quality of wide-coverage MT through automatic sentence decomposition and subsequent chunk recombination. It includes the design, implementation (Java) and testing of a chunking algorithm that, based on high-quality linguistic analyses, has the potential to improve current state-of-the-art MT output.

TransBooster has been shown to improve online freely available MT systems (Mellebeek et al., 2005), EBMT (Owczarzak et al., 2006b) and SMT (Mellebeek et al., 2006a) systems, as well as being able to operate in a multi-engine architecture (Mellebeek et al., 2006b).

### *6.8 MT Evaluation*

Recently, more and more developers of MT systems have turned their hand to the problem of designing better automatic MT evaluation metrics to mirror the improvements seen in system performance. In (Owczarzak et al., 2006a), we add paraphrases to the BLEU MT evaluation metric (Papineni et al., 2006) to allow for paraphrases of material contained in the ‘gold standard’ human reference translations to boost MT evaluation scores.

### *6.9 Sign Language MT*

Sign languages are the first and preferred languages of the Deaf Community worldwide. As with other minority languages, they are often poorly resourced (cf. Morrissey & Way, 2006) and in many cases lack political and social recognition. As with speakers of minority languages, Deaf people are often required to access documentation or communicate in a language that is not natural to them. In an attempt to alleviate this problem we are developing an example-based machine translation (EBMT) system (Morrissey & Way, 2005) to allow Deaf people to access information in the language of their choice. While some research exists on translating between natural and sign languages, we believe ours is the first attempt to tackle this problem using an EBMT approach.

## **Conclusions**

In this paper, we have provided an overview of how translation technologies are integrated into teaching and research at DCU. In a climate where the number of students taking languages at third level is in general decline globally, the area of translation studies is increasing in popularity due to its practical nature. In particular, it is our experience that translation technology elements in translator training provide our graduates with not only immediately relevant professional skills but also more in-depth knowledge of natural language behaviour in the context of translation. Our experience in SALIS also suggests that MT provides ample scope for research from humanities perspectives as well as from an interdisciplinary viewpoint.

While there are no current MT courses available in the School of Computing, we have pointed out that modules which have been taught in the recent past may soon find themselves back on our books if the proposed Dublin Graduate School in NLP is allowed to run. Despite our current lack of activity in this area at undergraduate level, we have more than compensated for this by our wide variety of interests in MT research at postgraduate level and beyond.

In sum, MT and translation technology are alive and well in DCU. As well as teaching students at undergraduate level, and engaging in cutting-edge postgraduate research, we have a long history of hosting established researchers over varying periods in our labs. To this end, we would be keen to hear from prospective researchers at all levels.

## References:

- Armstrong, S.** 2007. *Using EBMT to Produce Foreign Language Subtitles*. MSc Thesis, Dublin City University, Dublin, Ireland.
- Armstrong, S., Caffrey, C., Flanagan, M., Kenny, D., O'Hagan, M., and Way, A.** 2006a. Improving the Quality of Automated DVD Subtitles via Example-Based Machine Translation. In *Translating and the Computer*. London: Aslib.
- Armstrong, S., M. Flanagan, Y. Graham, D. Groves, B. Mellebeek, S. Morrissey, N. Stroppa and A. Way.** 2006b. MaTrEx: Machine Translation Using Examples. *TC-STAR OpenLab Workshop on Speech Translation*. Trento, Italy.
- Armstrong, S., Caffrey, C., Flanagan, M., Kenny, D., O'Hagan, M., and Way, A.** 2006c. Leading by Example: Automatic Translation of Subtitles via EBMT. In *Perspectives: Studies in Translatology*, 14 (3), 163-184.
- Badia, T., Freigang, K-H., Haller, J., Horschmann, C., Huber, D., Maia, B., and Reuther, U.** 1999. LETRAC curriculum modules. Available from <http://www.iai.uni-sb.de/LETRAC/letrac-data.html> (last checked 15 December 2006).
- Bangalore, S. and Joshi, A.** 1999. Supertagging: an Approach to Almost Parsing. *Computational Linguistics* 25:237—265.
- Bresnan, J.** 2001. *Lexical-Functional Grammar*. Blackwell, Oxford, UK.
- Bod, R. and Kaplan, R.** 2003. A DOP Model for Lexical Functional Grammar. In Bod et al. (eds.) (2003) *Data-Oriented Parsing*. CSLI, Stanford, CA., pp.211—232.
- Bod, R., Scha, R. and Sima'an, K.** (eds.). 2003. *Data-Oriented Parsing*. CSLI, Stanford, CA.
- Carl, M. and Way, A.** (eds). 2003. *Recent Advances in Example-Based Machine Translation*. Kluwer Academic Publishers, Dordrecht, The Netherlands.
- Finn, R., M. Hearne, A. Way and J. Van Genabith.** 2006. GF-DOP: Grammatical Feature Data-Oriented Parsing. In *Proceedings of the Eleventh International Conference on LFG*, Konstanz, Germany (to appear).
- Gough, N., and Way, A.** 2004. Robust Large-Scale EBMT with Marker-Based Segmentation. In *Proceedings of the Tenth Conference on Theoretical and Methodological Issues in Machine Translation (TMI-04)*, Baltimore, MD., pp.95—104.
- Groves, D. and A. Way.** 2005. Hybrid Example-Based SMT: the Best of Both Worlds?. In *Proceedings of ACL 2005 Workshop on Building and Using Parallel Texts: Data-Driven Machine Translation and Beyond*, Ann Arbor, MI., pp.183—190.
- Groves, D. and A. Way.** 2006. Hybrid Data-Driven Models of Machine Translation. *Machine Translation: Special Issue on Example-Based Machine Translation* (in press).
- Hassan, H., M. Hearne, A. Way and K. Sima'an.** 2006. Syntactic Phrase-Based Statistical Machine Translation. In *Proceedings of the IEEE 2006 Workshop on Spoken Language Translation*, Palm Beach, Aruba.
- Hearne, M.** 2005. *Data-Oriented Models of Parsing and Translation*. PhD Thesis, Dublin City University, Dublin, Ireland.
- Hearne, M. and Way, A.** 2003. Seeing the Wood for the Trees: Data-Oriented Translation. *MT Summit IX*, New Orleans, LO., pp.165—172.

- Hearne, M.** and **A. Way.** 2006. Disambiguation Strategies for Data-Oriented Translation. In *Proceedings of the 11th Conference of the European Association for Machine Translation*, Oslo, Norway, pp.59—68.
- Kenny, D.** and **Way, A.** 2001. Teaching Machine Translation & Translation Technology: A Contrastive Study. In *Proceedings of the MT Summit VIII Workshop on Teaching Translation*, Santiago de Compostela, Spain (with D. Kenny), pp.13—17.
- Koehn, P.** 2005. Europarl: a Parallel Corpus for Statistical Machine Translation. In *MT Summit X*, Phuket, Thailand, pp.79—86.
- Koehn, P., Och, F.** and **Marcu, D.** 2003. Statistical Phrase-based Machine Translation. In *Proceedings of the Human Language Technology Conference (HLT-NAACL)*, Edmonton, Canada, pp.48—54.
- Mellebeek, B., A. Khasin, K. Owczarzak, J. Van Genabith** and **A. Way.** 2005. Improving Online Machine Translation Systems. In *Proceedings of the 10th Machine Translation Summit*, Phuket, Thailand, pp.290—297.
- Mellebeek, B., K. Owczarzak, D. Groves, J. Van Genabith** and **A. Way.** 2006a. A Syntactic Skeleton for Statistical Machine Translation. In *Proceedings of the 11th Conference of the European Association for Machine Translation*, Oslo, Norway, pp.195—202.
- Mellebeek, B., K. Owczarzak, J. Van Genabith** and **A. Way.** 2006b. Multi-Engine Machine Translation by Recursive Sentence Decomposition. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas*, Boston, MA., pp.110—118.
- Morrissey, S.** and **A. Way.** 2005. An Example-Based Approach to Translating Sign Language. In *Proceedings of the Second Workshop on Example-Based Machine Translation*, Phuket, Thailand, pp.109—116.
- Morrissey, S.** and **A. Way.** 2006. Lost in Translation: the Problems of Using Mainstream MT Evaluation Metrics for Sign Language Translation. In *Proceedings of the SALTMIL Workshop on Minority Languages, 5th International Conference on Language Resources and Evaluation (LREC 2006)*, Genoa, Italy, pp.91—98.
- O'Brien, S.** 2006. *Translatability Assessment in a Controlled Language and Machine Translation Environment*. PhD Thesis, Dublin City University, Dublin, Ireland.
- O'Hagan, M.** 2003. Can language technology respond to the subtitle's dilemma? – A preliminary study. In *Translating and the Computer 25*. London: Aslib.
- Owczarzak, K., D. Groves, J. Van Genabith** and **A. Way.** 2006a. Contextual Bitext-Derived Paraphrases in Automatic MT Evaluation. In *Proceedings of the SMT Workshop at HLT-NAACL 2006*, New York, NY.
- Owczarzak, K., B. Mellebeek, D. Groves, J. Van Genabith** and **A. Way.** 2006b. Wrapper Syntax for Example-based Machine Translation. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas*, Boston, MA., pp.148—155.
- Papineni, K., Roukos, S., Ward, Y.** and **Zhu, W-J.** 2002. BLEU: a Method for Automatic Evaluation of Machine Translation. In *Proceedings of the 40<sup>th</sup> Annual Meeting of the Association for Computational Linguistics*, Philadelphia, PA., pp.311—318.
- Roturier, J. (2006).** *An investigation into the impact of Controlled Language rules on the comprehensibility, usefulness, and acceptability of machine-translated English technical documentation for French and German users*. PhD Thesis, Dublin City University, Dublin, Ireland.
- Stroppa, N.** and **Way, A.** MaTrEx: DCU Machine Translation System for IWSLT 2006. In *Proceedings of the International Workshop on Spoken Language Translation*, Kyoto, Japan, November 2006.
- Stroppa, N., D. Groves, K. Sarasola** and **A. Way.** 2006. Example-based Machine Translation of the Basque Language. In *Proceedings of the 7th Conference of the Association for Machine Translation in the Americas*, Boston, MA., pp.232—241.

**Utiyama, M. and Isahara, H.** 2003. Reliable Measures for Aligning Japanese-English News Articles and Sentences. In *Proceedings of the 41<sup>st</sup> Annual Meeting of the Association for Computational Linguistics (ACL-03)*, Sapporo, Japan, pp.72—79.

**van den Bosch, A., N. Stroppa and A. Way.** 2007. A memory-based classification approach to marker-based EBMT. In *Proceedings of the METIS-II Workshop on New Approaches to Machine Translation*, Leuven, Belgium (to appear).

**Way, A. and Gough, N.** 2003. wEBMT: Developing and Validating an EBMT System using the World Wide Web. *Computational Linguistics* **29**(3): 421—457.

**Way, A. and Gough, N.** 2005a. Controlled Translation in an Example-Based Environment: What Do Automatic Evaluation Metrics Tell Us? *Machine Translation* **19**:1—36.

**Way, A. and Gough, N.** 2005b. Comparing Example-Based and Statistical machine Translation. *Natural Language Engineering* **11**(3):295—309.

## 第16回 JTF 翻訳祭 —翻訳の原点、翻訳業界の共通認識—

社団法人 日本翻訳連盟(JTF) 副会長  
株式会社ベストバージョン 代表取締役 野上 員生

### 翻訳祭開催

(社)日本翻訳連盟(JTF)の第16回翻訳祭が、「翻訳の原点、翻訳業界の共通認識 ～常識のように言われていることは正しいのか～」をテーマに、昨年10月12日、東京・八丁堀のマツダホールで開催された。メイン会場では、最近の翻訳界で大きい話題になったアーサー・ゴールデン著「さゆり」の翻訳者である小川高義氏の講演、次いで、井口耕二氏(JTF 常務理事、実務翻訳者)の司会により、「翻訳者のアタマの中」というテーマのもと、様々な分野で活躍するベテラン翻訳者に小川氏も加わっていただいてパネル・ディスカッションが行われた。最後に、マーク・アタウェイ氏(JTF 理事、ライオンブリッジ ジャパン(株)代表取締役)の講演が行われ、日本の翻訳業界でも大きい位置を占めるようになったローカリゼーション分野について詳しく解説していただいた。

### 講演1：小川高義氏 「翻訳は推理ゲームである」

翻訳は、訳者が捜査員となって作品の現場を推理していく「推理ゲーム」のようなものである。訳者は、冠詞、単数／複数形といった原文の言語面の情報や、訳出された日本語から見つかる常識的な矛盾点などを手がかりとしてあらゆる解釈の可能性を探っていく。文学作品においては、作品の世界が構築されていることが一番重要なことであり、訳者の仕事は、単に「訳す」ということではなく、日本語の別バージョンを「作る」ということなのである。



小川高義氏

### パネル・ディスカッション

「翻訳者の頭の中」～翻訳しているとき、何を考えているのか・考えるべきなのか～をテーマとして、4名のパネリストによる討論が行われた。

<パネリスト>

アンゼたかし氏(映像翻訳者)・小川高義氏(文芸翻訳家、横浜市立大学国際総合科学部準教授)・倉橋純子氏((有)知財工芸 代表取締役)・山田和子氏(実務翻訳者、文芸翻訳家)

司会：井口耕二氏(JTF 常務理事、実務翻訳者)



パネル・ディスカッションの様子

講演2：マーク・アタウェイ氏 「進化するグローバルローカリゼーションプロバイダと翻訳者への期待」

最近ではWeb2.0やSaaS（Software as a service）などへの流れの中、必要なときにネット上のシステムにアクセスすれば何でもできるという状況が実現されつつある。顧客からも、翻訳メモリを一括管理できるシステムや、各プロジェクトのさまざまな情報が確認できるポータルサイトなどが求められている。そのような要望にこたえるためのシステムやツール類の開発が行われ、自動化・効率化が図られつつある。



マーク・アタウェイ氏（向かって左）

交流パーティー

講演終了後には交流パーティーが催された。経済産業省の商務情報政策局サービス産業課の企画官・高野芳久氏から祝辞をいただいた。また JTF <ほんやく検定>の1級合格者8名が招待され、記念品が授与された。

翻訳プラザ

地下1階のサブ会場では翻訳プラザが開催され、各種の企業が翻訳支援ツールやCD版辞書などを出展した展示・デモコーナー、およびJTF法人会員の翻訳会社を中心とする書籍・翻訳相談コーナーなどが参加者の注目を集めていた。



一級合格者。（後列左から：佐藤禪氏・平崎肇氏  
前列左から：多々良順子氏・長田真理子氏・笠井暁子氏・田口有里恵氏・島津秋氏・堀内麻衣子氏）



東 郁男 JTF 会長

会長挨拶

仕事というのは同じやり方をしていてもつまらない。また、同じやり方ばかりしていると、そこから後退が始まる。翻訳の仕事も厳しいところが多いと思いますが、その中で少しでも楽しみを見つけながら仕事をやっていくことが必要でしょう。

「第16回JTF翻訳祭DVD」好評発売中！

【No.1】講演1：小川高義氏

【No.2】パネルディスカッション

【No.3】講演2：マーク・アタウェイ氏

詳細は <http://www.jtf.jp> をご覧ください。

## 『国際ワークショップ IWSLT 2006』報告



独立行政法人情報通信研究機構  
株式会社国際電気通信基礎技術研究所

隅田 英一郎



## 1. はじめに

IWSLT2006は11月27・28日に京都でC-STAR(音声翻訳に関するコンソーシアム)とATR((株)国際電気通信基礎技術研究所)の主催、NICT((独)情報通信研究機構)の共催で開催されました。IWSLTとはInternational Workshop on Spoken Language Translationの略であり、異なる言葉を話す人々の間での対話を実現するため音声翻訳技術の確立を目指し毎年開催されている会議です。

この時期、山々は紅葉し街は観光客に溢れ宿の手配や移動も大変ですが、同時に、外国人にとっては日本を満喫する千載一遇の好機でした。このことがあってかどうか分かりませんが、イタリア、独逸、仏国、スペイン、イギリス、アイルランド、米国、カナダ、韓国、中国、香港など海外より34人、さらに、国内より50人と盛況でした。

ワイベル教授(CMU)、ナイ教授(RWTH)、ウー教授(HKUST)、ゾン教授(CAS)、フェデリコ博士(IRST)、フォーゲル博士(CMU)、黒橋教授(京大)、永田博士(NTT)など有名研究者が一堂に会し、日本に居ながら国際的な研究交流できる機会でした。

## 2. ワークショップの概要

IWSLTは、共通の学習・テストデータを用いて各研究機関の音声翻訳手法の精度・性能を評価する「評価セッション」と音声翻訳に関する最新の研究成果を発表する「テクニカルセッション」の2つから構成されています。また、海外の著名な研究者の招待講演も魅力です。

### 2.1 評価セッション

まず評価セッションの意義を説明します。2000年前後に、自動翻訳の研究は一変しました。一つは、大量の対訳から翻訳知識を学習するコーパススペースの手法が世界中で研究されはじめたことです。

もう一つは、翻訳の品質を自動的に評価する手法が提唱され、自動翻訳の研究者・開発者に広く普及したことです。これらは、研究者が思いつく玉石混交のアイデアを次々に翻訳システムに組み込んで、その翻訳システムの「実力」を正確に手早く評価し、その「実力」に貢献した技術と貢献しなかった技術を判別することを可能にしました。現在の機械翻訳の研究者は、このように有効な技術を貪欲に探索しているのです。

この背景の下で、今年の評価セッションでは、日本語、中国語、イタリア語、アラビア語から英語への翻訳が行なわれました。世界各国の19グループによる21の翻訳システムが参加しました。また、今回で3度目となるIWSLTでは、新たに話者を制約しない自然な発話形態の「自由発話翻訳」のトラックが設けられました。

全体として以下の知見が得られました。

①英語に翻訳する場合に、原言語でみると、日本語、中国語が難しく(翻訳性能が低く)、アラビア語はより簡単で、イタリア語に至っては大幅に簡単であることが品質評価から確認できました。

②今回有効性が確認できた技術は以下です。

- 語順モデル(distortion model)の改良
- 構文の活用
- モデルの話題適応
- 単語グラフを直接処理する探索
- 新しいfeatureを使ったリスクアリング

③「自由発話翻訳」はすべてのシステムの翻訳品質が良くありませんでした。翻訳システムの学習用のデータと「自由発話」のデータのスタイルの相違が大きく、これを何処の研究機関もまだ克服できていないといえます。

また、最も有効性の高い技術群を示したRWTH

にベスト・ペーパー賞が与えられました。

## 2.2 テクニカルセッション

7件の発表がありました。注目を集めたアイデアを一つあげると、ウーフィン博士(NRC)の対訳コーパスのブートストラッピングがあります。

- モノリンガルデータを翻訳する。
- 信頼できる翻訳を同定する。
- 同定されたものを対訳として再学習する。

この方法はまだ緒についたばかりですが対訳コーパスが足りない場合の対策として今後の発展が期待されます。

## 2.3 招待講演

ラザリー博士(IRST)による「演説や講演などの独話」を対象とした EU 予算による音声翻訳プロジェクト TC-STAR の紹介とルーコス博士(IBM)による DARPA 予算による多言語多モーダル情報処理ブ



図 1 ラザリー博士(左)とルーコス博士(右)

プロジェクト GALE の紹介がありました(図 1)。

ともに、出口イメージをクッキリさせた上で、多額の予算と研究者を投じて実施され、結果、優れた技術革新を実現していることがよく分かり印象的でした。

## 2.4 IWSLT に参加したシステムの実力

最後に、最先端のシステムの翻訳品質性能を実感していただくために、好成績をあげた統計翻訳システムの訳文を下に示します。下線で示したように変なところもありますが平均的にはかなり高品質です。

## 3. 終わりに

国際ワークショップ IWSLT は来年イタリアの古都トレントで開催される予定です。日本からの参加を増やし日本の技術を世界に宣伝しましょう。

論文やスライドによる最新技術を含む全ての会議にかかわる情報が下記の URL からアクセスできます。是非参考にしてください。

参考

- [www.slt.atr.jp/IWSLT2006/index.html](http://www.slt.atr.jp/IWSLT2006/index.html)  
(IWSLT2006 のホームページ)
- [www.slt.atr.jp/IWSLT2006/archives/2005/11/proceedings.html](http://www.slt.atr.jp/IWSLT2006/archives/2005/11/proceedings.html) (IWSLT2006 の論文とスライド)
- [www.slc.atr.jp/IWSLT2006/archives/2005/11/references.html](http://www.slc.atr.jp/IWSLT2006/archives/2005/11/references.html) (過去の会議のページ)

### 原文

1. 航空券を家に忘れてしまいました。
2. オペラ座はどこですか。
3. このフィルムの現像と焼き付けをお願いできますか。
4. 背中マッサージはいかがですか。
5. 次のかたどうぞ。パスポートと申告用紙を出してください。何か申告する物はありますか。
6. 二つ目の角にあります。
7. 海側の部屋に替えてください。
8. お勘定をおねがいします。
9. 玉ネギをお願いします。
10. 搭乗開始時刻は何時ですか。

### 訳文

1. I left my ticket at home.
2. Where is the opera?
3. Can I develop and print this film?
4. Would you like a facial massage on my back?
5. Please give me your passport and next person form. Do you have anything to declare?
6. It 's on the second corner.
7. I 'd like to ocean view room.
8. I 'll be arriving. I have the bill please.
9. I 'd like onion.
10. What time does boarding?

## 『GALA in Tokyo』参加報告

株式会社ナビックス 村上 嘉陽

昨年の9月26日に、GALA (Globalization & Localization Association) の国際シンポジウム『GALA in Tokyo』が都内で開催され、一部のセッションで機械翻訳に期待する議論が展開されたため、ここに当日の状況をご報告させて頂く。

GALA はアメリカに本拠を置くローカライゼーションベンダーの国際団体で、合計すると世界39カ国から約200社が会員となっているが、そのうち日本企業の加入は10社強に止まっている(06年9月現在)。

GALA は、ローカライゼーションベンダー同士のコラボレーションを奨励し、ローカライゼーション業界の認知度を世界規模で高めることを目的として活動している団体である。

通常は年次総会を欧州と北米で各1回ずつ計2回開催するそうだが、日本でのイベント開催は今回が初めてである。

当日は午後2時からの伊藤彰彦 GALA 会長の開会挨拶に引き続き、基調講演に2つのパネルディスカッションが行われ、わずか半日とはいえ三部構成で大変内容の濃いシンポジウムとなった。

基調講演に立ったのは、アメリカ西海岸で翻訳・ローカライゼーション業界のリーサーやコンサルティングを行っているブラジル出身の Renato Beninatto 氏であった。

Beninatto 氏によると翻訳バイヤーの87%は翻訳作業を外注しており、こうした翻訳の市場規模は年間成長率が7.5%で、2006年には96億USドルにも達すると指摘していた。因みに世界の三大翻訳ベンダーは、ライオンブリッジテクノロジーズ(米)、L3(米)、SDL International(英)だそうである。また頻繁に翻訳される分野については、1位がテク

ノロジー分野で2位が医学とのことだが、3位は大接戦で広告・マーケティング分野と製造業、そして金融関係がほぼ横一線であることが分かった。

Beninatto 氏は、昨今の翻訳・ローカライゼーション業界のサービスモデルには3つのパターンがあると指摘する。1つめのパターンはテクノロジーで収益を上げる SDL International のビジネスモデルである。つまり自社の翻訳メモリ製品である SDL TRADOS を売ることによって収益を上げ、サービスは製品に付随するものと位置付ける考え方である。

これに対して2つめは、ライオンブリッジに代表される自社で開発したテクノロジーを無料とするビジネスモデルであり、3つめはそのハイブリッド版である。因みに3つめのハイブリッド版のビジネスモデルとは、実際にはワークフローを重視したアプローチであるとの指摘であった。

顧客が望むものという意味からは、現状では価格面での貢献を期待する声が目立ち、それに次いで納品物や試訳の品質、さらにはそこに至るまでの相互の信頼関係の構築が求められていることが明らかになった。

Beninatto 氏のデータを拝見して印象的だったのは、2000年と2005年では上位3社の顔ぶれが全て入れ替わっていたことである。恐らくM&Aの影響などもあつたこととは思われるものの、国際的には変化の早い業界であることが浮き彫りになったようにも感じられた。

そのひとつの事例として、人材派遣業やオフショア開発、さらにはIBMやXEROX、HPといった大手IT企業などから翻訳・ローカライゼーション業界に続々と参入している話も紹介され、海外では大きな動きがあることが分かり大変参考になった。

また国別に見ると、2005年現在の上位10社のうち、7社がアメリカかイギリスに本社を置く企業である一方で、日本企業は上位20社にもランクインしていなかったことから、日本の翻訳・ローカライゼーション業界は、その名に反して国際化、現地化が進む製造業などとは大きく異なる状況にあるように思われるとともに、日本のサービス業の国際化がさほど進展していないようにも思われて、妙に印象的であった。

第二部では Beninatto 氏に加えて(株)アークコミュニケーションズ社長の大里真理子氏、Moravia IT Japan 代表兼セールス・ディレクターの小西輝一氏をパネリストに迎え、(株)テック・インデックスの安田聡子氏がモデレーターとなってパネルディスカッション I が行われた。

テーマは、ローカライゼーション業界における営業・マーケティング活動であり、具体的には日本と海外のベンダーの活動手法の比較や、海外から見た日本のベンダーの評価、さらには営業・マーケティング活動の実態などについて意見が交わされた。

最初に Beninatto 氏からグローバルセールスに関するトピックスとして以下の3点が提示された。1つめは「我が社には〇〇ができます。」といった First person statement ではなく、「御社は△△をされたいのですね。」といった Second person statement を使って顧客ニーズを的確に掴むことであり、2つめは最早品質は他社との差別化要因にはならないこと、即ち品質はあって当然なのであり、その上で差別化できる要因をサービスに付加すべきということ、そして3つめは、もし国際的に成功したいのであれば、ビジネスに国境がない以上、自らを日本企業と位置付けるべきではないということである。

続いて大里氏からは、国際的に成功している製造業からヒントを探ることが有効との指摘があった。また効率化を求めて高品質を求める手法を目指すならば、専門化してリーズナブルな価格で大量に供給する手法を採用すべきとの指摘があり、他方リスク

を回避して変化への対応、即ち柔軟性を求めるのであれば、複数のサービスをプロジェクトベースで行うことが望ましいとの話があった。そして最後にビジネスコミュニケーションを加速させるソリューションを提供するためにも、個々の改善よりもプロセスを省略して利益を出したいということと、セールスとマーケティングとプロダクションの各指標の重要性が指摘され、小規模の会社が山のようにあるフラグメントされた業界だからこそ面白いと考えて、好きな仕事をゲリラ的に取って行きたいとの意欲的なコメントがあった。

小西氏からはチェコに本社があり、日本の他にも欧米や中国、中南米などにも拠点のある企業からの視点として、主なものをあげると以下のような指摘があった。

まず日本では ROI (投資収益率) を考えているのかが疑問であるということ、ローカライゼーションはイコール翻訳であるという感覚が根強いのではないかという疑問、業界としての広告媒体が限られてしまっている現状、そして営業とマーケティングが同一視されたり、同じ部門になっていたりする疑問などが指摘された。

その一方で日系企業と外資系企業の双方の良さを統合したいという話や、現在ある業界団体が協力関係を持つべきとする意見、さらにはメーカーへの働きかけは「翻訳売ります」ではなくコンサルティングにすべきとする提言などもあり、マーケティング手法の確立に向けて、調査・把握・分析・戦略・フィードバックなどを詳細に検討することの重要性も指摘された。

パネルディスカッション I の最後に行われた議論では、さらにいろいろなコメントが出された。例えば翻訳業界では職人気質の傾向があるが、職業翻訳の認知度は海外よりも日本の方が高いとする指摘や、日本企業では品質向上などの内発的なプロセスを重視する傾向にあり、ベンダーとしては顧客密着型の営業展開となるため、マーケティングと言っても実際にはあまりピンと来ないこと、そして海外で

顧客開拓をする際にはインターネットが唯一無二の手法だとする指摘などである。

またマーケティング予算の実例として、コカ・コーラ社が 2004 年度に収益の 10% (22 億 US ドル) を、テレビ、ラジオ、印刷物、インターネットといった諸媒体を通じて広告宣伝費として投入したことが紹介された。

議論の最後は営業・マーケティングの今後に関する話題となり、必要なことはイベントやフォーラムなどへ積極的に参加して情報収集とともに自社をアピールする場を持つことと、低コストでどのように生き抜くかを考えることの重要性が指摘された。とりわけ後者に関しては、もし自社のサービスが「低コストで高品質」であるならば、最早品質だけでは生き残ることができない時代に突入している以上、その理由を明確に説明しなければならないとする指摘であり、決して斬新な提言ではないものの、国際情勢を踏まえた話として大変納得の行く指摘であったように思う。

GALA in Tokyo の最後には、ローカライゼーションの自動化と省力化をテーマにパネルディスカッションⅡが行われ、パネリストとしてグローバルライゼーションコンサルタントの柳英夫氏、(株)十印の川原勝氏、そして(株)サン・フレアの渋谷淳氏が登壇した。

このセッションは、テクニカルセッションとして翻訳メモリを超えて次のステップへ向かうためには何が必要かといった議論が展開された。最初は渋谷氏が翻訳メモリに関するプレゼンテーションを行い、ランゲージベンダーは TRADOS の登場以降、テクノロジーに振り回されているとする指摘から始まった。そして翻訳メモリは既にベンダーにも顧客にも双方に定着した技術だが、同時に枯れてきた技術でもあり、そうしたこともあってか次世代の翻訳メモリを提唱する動きもあるとする話であった。例えば翻訳メモリは、デスクトップモデルから現在ではサーバーモデルでの Web ベースによる情報一元化が実現しているが、翻訳メモリ内でのノイズの

伝染にどう対処するかといった問題や、一次翻訳をサーバーベースの翻訳メモリに入れることへの疑問などには明確な回答が示されていないため、運用のベストプラクティスは試行中とは言うものの、技術としては既存のものであることには相違がない以上、技術的な限界に達している疑いがあるとする指摘である。

これに続いて生産性に関する説明もなされ、冒頭に翻訳メモリの登場は一種の産業革命であって、生産性の向上には大きく寄与したことが指摘された。ただ生産性の上昇曲線は昨今鈍化していることは間違いなく、一説によると踊り場状態にあるとも言われていることや、顧客の要望が高度化する中であってもアルゴリズムに変化がない事実や、翻訳メモリだけでは追いつけなくなっている現状などが次々と報告された。

こうした状況にあって将来的に明るい展望を持つことができるとしたら、それは人手ではなく、かつての翻訳メモリの登場のようなテクノロジーに期待すべきであるため、産業革命の次の波を模索する必要があるとの指摘があった。ただ翻訳メモリを完全に置換したものを模索するのは急進的且つ非現実的な話であるため、Transition Phase として求めるべきソリューションは、「次世代 TM」か「TM+X」のソリューションと考えられるとする指摘には大変共感することができた。

因みに前者は SDL TRADOS のような文単位だけではなく、フレーズ (sub-sentence) 単位でもマッチングをできるようにすることで旧訳の再利用率を向上させるものであり、MultiCorpora 社や Lingotek 社にその事例を見ることができるとする説明であったが、同時に日本語の言語としての特性と課題に話が及んだ際、形態素ごとの課金という提案があり、大変興味を持った。つまり「経済」という 2 文字の名詞も、「インプリメンテーション」という 11 文字の名詞も英語にするといずれも「economy」や「implementation」といった具合に 1 語になってしまう日英の非対称性を指して出

された提案だが、文字数やワード数ではなく、形態素数という提案は、万能ではないにしても確かに一理あると感じた次第である。

他方後者の場合は、従来にも増して「+X」の比重が大きくなってきているとの説明であり、その「+X」に機械翻訳が含まれていることが十分に予想されたため、この後の議論には大いに期待することができた。

続いてモデレーターを務めていた柳氏からワークフローに関するプレゼンテーションがあり、内容はGMS (Globalization Management System) についての説明であった。

通常翻訳代に1ドルかかるとすると、翻訳コンテンツの管理コストに0.7ドルかかると言われるが、柳氏はこうした現状を踏まえ、CMS (Contents Management System) との対比においてGMSを説明して見せた。

端的に言えば、プロジェクトマネジメントは往々にして情報のメンテナンスに時間を費やしてしまっていてプロアクティブな動きにならないため、ファイルをワークフローユニットとするCMSよりも、よりスケラブルな形での品質保証が可能なシステムとして、セグメントをワークフローユニットとするGMSを推奨するといった話であった。そしてプロジェクトマネジメントがLSP (ランゲージサービスプロバイダー) のサービス事業の中心である以上、テクノロジーに対してどのように取り組むかがキーファクターとなると強調し、テクノロジーとサービスを一体とする陣営と、テクノロジーとサービスを区分する陣営とに分かれるとの説明がなされた。

ここからは翻訳メモリ、ワークフローに続いて主題が機械翻訳となり、川原氏と再度渋谷氏によるプレゼンテーションとなった。

まず川原氏からは「機械翻訳の今と将来」と題したプレゼンテーションが行われ、ローカライゼーションベンダーの悩みが翻訳者の確保や品質と納期の矛盾 (短納期で高品質を求められる実態)、膨大な

翻訳量と限られた翻訳者数、そして用語やスタイルなどの統一性保持の3点であることが指摘され、顧客の悩みが情報量の増大と優先順位 (品質、コスト、納期のうち、Time-to-Market の観点から納期が最優先となる図式) であることも併せて指摘された。またローカライゼーションプロセス全体で見ると、システム的な解決や効果はある程度見えてはいるものの、翻訳作業自体にはまだまだ取り組みの余地があるとの見地から、機械翻訳に期待が寄せられていることも報告された。

さらに機械翻訳にとっての将来に向けての重要なポイントとして、ビジネスモデルと収益性に着目する必要があることや、機械翻訳に合ったドメインや品質の定義付けも必要であること、そして機械翻訳の利用形態を明確にすること、即ち翻訳作業そのものなのか翻訳支援にとどめるのかといったことも示された。

川原氏の指摘の中で頷くことができたのは、情報量が激増している昨今の状況と、それに伴ってLSPに対する要求度合いが変化してきていることから、機械翻訳には期待したいとする部分であった。

再びプレゼンテーションに立った渋谷氏からは最初にルールベース、用例ベース、そして統計ベースへと発展してきた機械翻訳の変遷の説明があったが、これに続いて統計翻訳がどこまで可能なのかという話になった。ここでは、一定の利点はあるものの、技術が未成熟であり、デファクトスタンダードが存在しない状況を指摘するなど、現状における問題点の説明もあった。しかし機械翻訳と人手による翻訳はアイデア次第で相互に補うことが可能であり、人手による翻訳で対応する所と機械翻訳で対応する所を明確に使い分ければ、いい結果を得ることも可能ではないかと問いかけていた。こうした意見以外にも、Google が近い将来翻訳会社のライバル的な存在となることも十分考えられるとの話も披露していた。

3人のパネリストの意見を総括すると、次のようになる。それはIT革命はインターネットの登場によ

り距離を無くすことができたが、翻訳メモリは例えば IT 革命のコンピュータに相当するものであって、社会に大きな貢献をするにはしたものの、結果として何か足りず、社会の根幹までは変革できなかった。だから今こそ新しいテクノロジーがどう影響するかを我々は真剣に考えるべきなのであって、翻訳メモリが登場した時と同様のインパクトが、もしかしたら機械翻訳でも起きる可能性がある以上、例え「月進年歩」であったとしても機械翻訳には取り組むべきであるとした見解である。

またより詳細な想定として、翻訳の自動化ができない大きな障害のひとつに、翻訳する原稿に手書きの赤字が大量に入っているなど、電子データではないことも指摘していた。このため、日本語から英語を通じての多言語ローカライゼーションの工程を想定すると、日本語から英語を制作する段階よりも、むしろ英語から多言語への段階でこそ **automated workflow** となることも考えられ得るとの興味深い指摘もあったことを付記しておきたい。

これまでにご紹介したことは、ほぼ当日の議論を網羅できているものと思われるが、例え 1 社ではできないことでも、団結すれば業界の動きも変えられるとする GALA の主旨には大いに共感することができた。そして **Join, Participate, Improve, Innovate** という 4 つの段階をそのままキーワードにしたかのようなキャッチフレーズを通して、業界の活性化を狙う意識にも同様に共感することができた。

翻訳メモリから機械翻訳までについて語られた最後のパネルディスカッションを思い起こしながら、決して GALA は技術者集団ではないものの、私には「次世代のテクノロジーを自分たちで作る」という思いが、最後の **Innovate** という言葉に込められているように思われた。

もしそうだとするならば、こうしてテクノロジーを生み出す同業者の気概と団結心に触れられたことこそが、**GALA in Tokyo** に参加した最大の収穫だったのかもしれないと、半年前の当日を振り返って今しみじみと感じている次第である。

## 今後の機械翻訳の目指す道

株式会社富士通研究所

## 普及期に入った自動翻訳ソフト

みなさんは、昨年マガジンハウス社から出版された「匂いをかかれる かぐや姫 ~日本昔話 Remix~」という本を御存じでしょうか？ これは、日本の昔話を日英自動翻訳にかけ、その結果を英日自動翻訳で日本語に戻して、「さあ、読んでみてくださいっ」という本です。

「一寸法師 → 少量法律助言者」「かぐや姫 → 匂いをかかれるとすぐに、プリンセス」「桃太郎 → 桃タロイモ」などと訳され、昔話がシュールな物語に変換されています。その評価(Amazon.co.jp のカスタマーレビュー)はと言うと、

「日本古来の昔話をいったん英語のフィルターを通し日本語に再変換することによって、誰もが知る昔話がとっっても奇天烈なお話になっています。」

「悪友からニヤニヤしながら渡され、ある日気分を変えようと読み始めた瞬間から大爆笑。嫌な気分もどこへやら、さわやかな気分で仕事に戻りました。」

「大意が把握できるだけに原作のすごさを実感できる一冊かと思います。」

と揶揄はしているものの、自動翻訳の存在には肯定的(?)です。実際、無料の翻訳サイトでは、一日500万件のアクセスがあり、宿題をこなす学生や技術書を読むビジネスマンに使われています。また、2006年のGoogle日本語検索のランキングでは、「地図」に1位は譲ったものの、「翻訳」は2位でした(Google年間検索ランキング)。

このような実態から見ると、「自動翻訳」は広く使われており、改善の余地はあるものの、「英語が読めなくても、書けなくてもそこそこのことはできる」という「翻訳ソフトの目標」はある程度の達成されているのではないのでしょうか。

## 今後の自動翻訳の目指す道

いわゆるフリーランスの翻訳家、翻訳会社の翻訳家には、まだまだ自動翻訳は使われていません。自動翻訳は、原文からいきなり訳文を出すので、翻訳家が頭の中で考える翻訳プロセスと合致せず、自動翻訳結果を追っていると却って時間がかかってしまうからです。

一方、翻訳業界は、昨今、クライアントからのコストダウン圧力が強く、労働集約的な翻訳作業のコストを如何に下げることが課題になっています。以前は、翻訳ソフトへ目もくれなかった翻訳業界ですが、翻訳メモリを始め、翻訳ツールへ熱い視線が注がれています。

今までの自動翻訳ソフトは、いわば「巨艦主義」で、原文が入ってくれば、ポンッと訳文を出す。細かいことはあまり気にしない。というものでした。翻訳家向けにはこれではいけません。翻訳作業は、わからない単語を辞書で調べたり、いい例文がないかインターネットで検索したり、既存の翻訳例から適切な訳を拾ってきたり、原文のシーンに合う背景に考慮したりと、高度な知的作業の集積です。自動翻訳は、ただ訳文を出すだけでなく、このような知的作業を助けるツールとして利用してもらえようになれば、翻訳専門家にも受け入れられるでしょう。「巨艦主義」から「いろいろなタイプの小回りの利くツールの集積」としての展開が望まれています。今後は、「巨艦主義」のWeb1.0からWeb2.0的な思想・手法へと変化させ、翻訳作業のコストダウンに貢献できるように進化して行きます。

## AAMT 新体制紹介

2006年7月より、当協会は、事務局の京都への移転とともに新体制へと移行しました。

### AAMT 役員名/AAMT Board Members

会長	井佐原均 (独立行政法人情報通信研究機構 自然言語グループリーダー)
President	Hitoshi Isahara (Group Leader, Computational Linguistics Group, NICT)
副会長	小谷泰造 (株式会社インターグループ 代表取締役社長)
Vice President	Tiazo Kotani (President, Inter Group Corporation) 飯田仁 (東京工科大学 教授) Hitoshi Iida (Professor, Tokyo University of Technology)
理事	辻井潤一 (東京大学大学院 教授)
Director	Jun-ichi Tsujii (Professor, University of Tokyo) 長尾眞 (独立行政法人情報通信研究機構 理事長) Makoto Nagao (President, NICT) 田中穂積 (中京大学 教授) Hozumi Tanaka (Professor, Chukyo University) 横山晶一 (山形大学 教授) Shoichi Yokoyama (Professor, Yamagata University) 坂本義行 Yoshiyuki Sakamoto 石崎俊 (慶應義塾大学 教授) Shun Ishizaki (Professor, Keio University) 崔杞鮮 (韓国 KAIST 教授) Key-Sun Choi (Professor, Computer Science Division, KAIST Korterm, Korea) ソンラートラムワーニッチ・ウィラット (タイ自然言語ラボラトリー ディレクター) Virach Somlertlamvanich (Director, Thai Computational Linguistics Laboratory) 吹譚正憲 (社団法人電子情報技術産業協会 専務理事) Masanori Fukiwake (President, Japan Electronics and Information Technology Industries Association) 青木正博 (東芝ソリューション株式会社 経営企画部 渉外担当部長) Masanori Aoki (Senior Manager, Public Affairs, Business Planning Div., Toshiba Solutions Corporation)

理事	富澤賢治（富士通株式会社 法務・知的財産権本部情報部 シニアエキスパート）
Director	Kenji Tomizawa (Senior Expert, Information Services Division, Law & Intellectual Property Group, Fujitsu Ltd.)
	村田稔樹（沖電気工業株式会社 研究開発本部技術マーケティング部 チームマネージャ）
	Toshiki Murata (Manager, Corporate Research and Development Center, Technology Marketing Department, Oki Electric Industry Co., Ltd.)
	福持陽士（シャープ株式会社 情報通信技術開発本部情報商品開発センター第一開発部 副参事）
	Yoji Fukumochi (Manager, Information & Communication R&D Group, Information Systems Products Development Center, Development Department I, Sharp Corporation)
監事	勝田美保子（株式会社十印 代表取締役会長）
Auditors	Mihoko Katsuta (CEO, Toin Corporation)
	種子田暁夫（日本電気株式会社 政策調査部 調査担当部長）
	Akio Taneda (Senior Manager, External Relations Division, NEC Corporation)

AAMT 組織図（2006 年 12 月現在）



## AAMT 機械翻訳課題調査委員会の活動について

独立行政法人情報通信研究機構

AAMT 機械翻訳課題調査委員長

内元 清貴

2006 年度より、これまでの技術動向委員会と市場動向調査委員会が統合され、「AAMT 機械翻訳課題調査委員会」が新設されました。当委員会では、機械翻訳に関する課題を調査し、アジアへの展開を視野に入れつつ、機械翻訳の研究や産業が健全な方向へ進むために中立機関として何をなすべきかを議論・検討します。

4 月から 6 月にかけての準備委員会を経て、現在、3 ワーキンググループ構成（WG①「評価」・WG②「調査・広報・啓蒙」・WG③「標準化・共有化」）で、月一回の委員会を基本に活動を行っています。本記事では、各ワーキンググループの活動のご説明を通して、この新しい委員会についてご紹介いたします。

## ワーキンググループ①「評価」

## (a) テストセットの構築と公開

旧技術動向委員会では、これまでベンダ向けに日英の言語ペアを対象としたテストセットを作成してきました。今後はその公開とともに多言語化を目指します。次の対象としては日中の言語ペアを予定しています。また、ベンダ向けだけでなくユーザ向けについても、テストセットとしてどのようなものがふさわしいかについて議論を重ね、方向性を打ち出していきたいと考えています。

## (b) 評価指標の確立

翻訳の質を適切に評価できるようにするために、学術的・技術的に新しい指標を提案し、その指標を用いた自動評価の方法を確立することを目指します。さらに、実用面を考慮し、低コストで高精度の評価ができるようにするために、自動で行なう部分と人手を加えるべき部分とをうまく分担することで効率化を図ることを考えています。また、翻訳の

質だけでなく、ある特定の分野へのチューニングがどのくらい効率良く行なえるかといった運用面も考慮した評価の方法も検討する予定です。

## ワーキンググループ②「調査・広報・啓蒙」

## (a) アンケート調査

旧市場動向調査委員会では、これまで日本での機械翻訳市場を調査してきました。今後は調査対象をアジア各国での市場へと拡張していく予定です。また、アンケートの内容については、これまでの調査結果や他 WG の活動内容を踏まえ、無料翻訳サイトの利用状況を追加するといった見直しも行なっていく予定です。アンケートは、インターネット WG と連携し、ホームページ上で行なう予定です。

## (b) ユーザ事例や FAQ の収集と公開

機械翻訳の訳質は、対象分野にあわせてユーザ辞書を充実させることにより向上します。しかし、具体的にどうすれば効率良く訳質を向上させられるかといったノウハウはあまり知られていないようです。本委員会では、専門家へのヒアリングやアンケート調査、関連する学会や講演会での情報収集などにより、機械翻訳の有効な利用方法を発掘し、ユーザ事例としてまとめることを考えています。また、会員から寄せられたご意見や質問等をもとに、ノウハウを FAQ としてまとめることも考えています。

## (c) 機械翻訳プロジェクトの調査と紹介

機械翻訳の学術的・技術的な動向を把握するために、機械翻訳に関する各国のプロジェクトや学会を広く調査し、解説等としてまとめることを考えています。

### ワーキンググループ③「標準化・共有化」

#### (a) 辞書・翻訳メモリ(TM)の共有

先にも述べましたように、一般に、辞書の充実によって機械翻訳の性能は向上します。しかし、各社の製品間で互換性がないためにユーザ間で共有するのが簡単ではなく、同様の辞書が個別に作られ、共通に使える辞書はほとんどないというのが現状です。当委員会では公的な機関として、共通に使える辞書を構築・管理し、会員に機械翻訳が効率良く行なえる環境を提供することを目指します。そのためには、標準的なフォーマットを定義し、普及させる必要があります。AAMTでは、過去にUPFという辞書記述の標準を提案しています。TMに関しては、TMXという標準が定められています。しかし、UPF、TMXともにまだ不十分な状態であるとの指摘があります。したがって、当委員会では、まず、既存仕様等についての実態調査を行なう予定です。その後、辞書とTMとの関係を考慮しつつ、仕様の標準化および共有化を検討します。最終的には、機械だけでなく人間も参照しやすいような辞書を目指します。また、将来的には、言語資源協会(GSK)といった言語資源コンソーシアムとも協力して言語資源を収集していくことも視野に入れたいと考えています。上記以外にも、共有化に伴い、資金面の問題、知財の扱い、信頼性の問題などにも対処していく必要があります。これらの問題を解決するためには、翻訳家の方々のご協力が不可欠であると考えています。

### 共通項目

#### (a) 白書の作成

インターネット WG と連携し、会員専用ページで各 WG の成果をタイムリーに発信していく予定です。そして、5年後を目処に活動内容を白書にまとめたいと考えています。

#### (b) 英語での発信

普段から活動内容を日本語と英語で定期的にまとめ、将来的にはその要約を多言語化することを検討しています。

## AAMT 新事務局紹介

—はじめまして! 新しい事務局よりごあいさつ—

AAMT 事務局長 神崎 享子

AAMT 事務局 和泉 絵美

2006年7月1日、AAMT事務局は、更なる業務効率化を目指し、東京・御茶ノ水のJEITA（(社)電子情報技術産業協会）事務所から、京都・精華町のNICT（(独)情報通信研究機構）知識創成コミュニケーション研究センター内へ移転いたしました。東京のオフィス街から一転、のどかな田園風景の中に対照的に先端技術の研究開発センターが立ち並ぶ新天地へ…大きな変化となりました。

精華町は、京都・大阪・奈良の三府県にまたがる京阪奈丘陵に位置する、文化・学術・研究の拠点・関西文化学術研究都市（通称：けいはんな学研都市）」の一角です。「関西のつくば」と言えば、東日本の方々にはその特徴がよりイメージしやすいでしょうか。したがって、たいていの方が持っておられる古都・京都のイメージとは異なるかもしれません。しかし、大阪・奈良とともに、日本の歴史の中でその豊かな文化的・自然的土壌の下で新しい価値を創造してきた精神が、先端科学技術を生み出す学研都市においても引き継がれていると言えます。NICT知識創成コミュニケーション研究センターは、本部を東京・小金井市に置くNICTのけいはんな支部です。当センターでは、情報通信システムと人をとりもつヒューマンインターフェイス技術とコンテンツ処理技術を軸として、豊かな情報生活を支える技術の研究開発が行われています。そのセンター内でAAMT事務局が入居する自然言語グループでは、機械翻訳をはじめとする自然言語処理の研究開発が行われており、井佐原均AAMT会長がリーダーとしてグループを率いています。

現在、井佐原会長の統括の下、神崎・和泉の二名で事務局業務を進めています。最後に、両名より会員の皆さまにごあいさつ申し上げます。

AAMT事務局長を仰せつかりました神崎享子です。AAMT会員の皆さまが機械翻訳の分野を盛り上げ牽引されていく、そのお手伝いができますことを、大変嬉しく思っております。新規の体制に変わりましたが、皆さまが安心してより一層アクティブにご活動できますよう、裏方からサポートして参る所存でございます。何卒よろしく願い申し上げます。

（神崎享子）

皆さまはじめまして。新事務局がスタートして約半年、AAMTは、個人・法人会員の皆さま、委員会活動にご協力いただいている皆さまなど、たくさんの方々のご厚意によって支えられているのだと実感しています。その方々への感謝の気持ちを持って業務を進め、「AAMT会員でよかった」と思っていただけのような、そして皆さまの日々のご活動に少しでも有益な情報をご提供できる組織でありたいと思います。どうぞよろしく願い申し上げます。

（和泉絵美）



「よろしく願いいたします！」

（左より）神崎事務局長・和泉

AAMT 事務局

〒619-0289 京都府相楽郡精華町光台 3-5

(独)情報通信研究機構 自然言語グループ内

電話：0774-93-4625 FAX：0774-93-4627

E-mail：[aamt-info@aamt.info](mailto:aamt-info@aamt.info)

## 第 16 回通常総会および関連行事

### (報告会・講演会・AAMT 長尾賞授与式・懇親会) の報告

当協会の第 16 回通常総会が、2006 年 6 月 16 日(金)午後 1 時 30 分より、東京・御茶ノ水の総評会館で開催されました。総会后、各委員会からの報告会と、2 名の講師による講演会、そして、本年度より設立された、「AAMT 長尾賞」の第一回授与式と受賞記念講演が盛況のうちに行われました。

#### 報告会

辻井潤一会長(当時)の開会挨拶の後、井佐原均技術動向調査委員長・坂本義行市場動向調査委員長、横山晶一 AAMT/Japio 特許翻訳研究会副委員長、富士秀インターネット WG リーダーより、それぞれの委員会における 2005 年度の活動が報告されました。

#### 講演会

はじめに、英語教育・第二言語習得研究における活発な活動で知られる、明海大学・投野由紀夫教授による、「人間の誤り、機械の誤り：外国語学習者のコーパスを見て考えること」の講演がありました。続いて、AAMT 初代会長で、MT の研究において第一人者である、(独)情報通信研究機構・長尾眞理事長より、「知識処理と学習・環境適合技術の時代」についてお話いただきました。

#### AAMT 長尾賞授与式・受賞記念講演

2005 年 4 月、AAMT 初代会長である長尾眞先生が日本国際賞を受賞され、その賞金の一部を AAMT の活動に資するようにとご寄付くださいました。これを受けて当協会では、「AAMT 長尾賞」を設けました。本賞においては、機械翻訳システムの実用化の促進および実用化のための研究開発に貢献した個人あるいはグループを表彰します。いわゆる学会の論文賞や学術賞ではなく、たとえば、高性能の機械翻訳システムを商品化した、機械翻訳システムを使った新しいサービスを開始した、といった貢献を対象とします。もちろん学術的にも意味のある成果を除外するものではありません。第一回の今回は、次の 2 グループに授与されました。

- ◆ 沖電気工業株式会社 研究開発本部 機械翻訳グループ
  - 受賞理由：エポックメイキングとなる機械翻訳システムやサービスの開発運用
- ◆ 日本電気株式会社 NEC メディア情報研究所
  - 受賞理由：携帯端末用自動通訳システムの実用化技術に関する研究・開発

授与式に続き、各社の代表者による受賞記念の講演が行われました。

#### 懇親会

本会后に行われた懇親会には、本年度も多数の参加があり、楽しいひとときを過ごしました。会の最後では、同月をもって AAMT 会長を退任された辻井潤一先生(東京大学大学院)よりご挨拶がありました。また、同じく同月をもって事務局業務を新事務局に引き継いだ中瀬真氏(事務局長、JEITA)、高田佳代子氏には、お二方に AAMT の活動を長年支えていただいた感謝の気持ちとともに、花束が贈られました。その後、井佐原均新会長より AAMT 新体制における意気込みが語られ、正式に新体制への引継ぎが行われました。



AAMT 長尾賞授与式の様子（長尾先生より、（左から）沖電気工業株式会社、NEC メディア情報研究所へ）

行事	内容	時間
総会	<p>第 16 回通常総会</p> <ol style="list-style-type: none"> <li>1. 開会の辞</li> <li>2. 会長挨拶 東京大学大学院教授 辻井 潤一</li> <li>3. 出席会員の確認</li> <li>4. 議事録署名人の選出</li> <li>5. 議案 <ul style="list-style-type: none"> <li>第 1 号議案 2005 年度事業報告（案）</li> <li>第 2 号議案 2005 年度決算報告（案）</li> <li>第 3 号議案 2006 年度事業計画（案）</li> <li>第 4 号議案 2006 年度収支予算（案）</li> <li>第 5 号議案 事務局移転の件</li> <li>第 6 号議案 役員改選（案）</li> </ul> </li> <li>その他、会員提案事項</li> <li>6. 報告事項 <ul style="list-style-type: none"> <li>総会後の組織・人事について（案）</li> </ul> </li> <li>7. 閉会の辞</li> </ol>	13:30-14:15
報告会	<p>会長挨拶</p> <p>各委員会からの 2005 年度活動報告：技術動向調査委員会、市場動向調査委員会、AAMT/Japio 特許翻訳研究会、インターネット WG</p>	14:30-15:30
講演会	<p>「人間の誤り、機械の誤り：外国語学習者のコーパスを見て考えること」 明海大学 教授 投野 由紀夫 氏</p> <p>「知識処理と学習・環境適合技術の時代」</p> <p>(独)情報通信研究機構 理事長／元 AAMT 会長 長尾 眞氏</p>	15:40-17:00
AAMT 長尾賞授与式 受賞記念講演会	<p>受賞グループ：</p> <p>沖電気工業株式会社 研究開発本部 機械翻訳グループ</p> <p>日本電気株式会社 NEC メディア情報研究所</p>	17:00-17:40
懇親会	<p>新お茶の水ビル・サンクレール 2F 「ピストロ備前」</p>	18:00-20:00

## 2006年 AAMT 長尾賞受賞に際して ～携帯端末用自動通訳システムの実用化～

日本電気株式会社 松川 淑子

### 1. 携帯端末用自動通訳システム

当社は「携帯端末用自動通訳システムの実用化技術に関する研究・開発」に対し、2006年にAAMT長尾賞をいただきました。この技術は、5万語という大規模な語彙を用いた旅行会話通訳機能を、携帯端末や携帯電話の製品へ搭載可能にしたものです。

当社は2002年、PDAでリアルタイムに動作する旅行会話向け日英双方向自動通訳システムを開発し、国土交通省の「e-エアポート」構想の一環として、成田空港において実証実験を重ねてきました(図1参照)[1]。また2005年12月、携帯電話用マルチコアプロセッサにおいてリアルタイム日英旅行会話通訳システムを実現しました(図2参照)[2]。この技術を多言語に展開することで、2006年1月、PDA単体で動作する日中旅行会話通訳ソフトウェアを開発しています。

そして2006年2月、日英旅行会話通訳機能を、NECの新製品であるモバイルマルチメディアプレーヤ VoToL という携帯端末上に、世界で初めて製品搭載しました(4月25日出荷、図3参照)[3]。

当社は、1983年のTelecom'83において小林宏治(当時NEC会長)が自動通訳電話のコンセプトモデルを発表して以来、一貫して自動通訳の技術開発を主導し開発を進めてきました。1991年、機械翻訳システムPIVOTと音声認識専用ハードウェアを開発・統合して、数百語規模の日英自動通訳システムINTERTALKERを試作し、Telecom'91に出展して国際的に高い評価を得ました。1999年には、日英旅行会話通訳をパソコン上のソフトウェアとして実現し、2001年にはパソコンのパッケージソフト「たび通」として製品化しました。このソフトは、5万語という豊富な語彙を持ち、旅行会話を



図1：PDA 日英旅行会話通訳



図2：携帯電話用プロセッサでの自動通訳



図3：モバイルマルチメディアプレーヤ VoToL

対象として通訳支援を行うものです。しかしこれら従来の自動通訳システムは、サーバやパソコンを用いたものであるか、もしくは携帯端末上のものであれば利用場面や語彙が限定されるものでした。これに対し我々は独自アルゴリズムの開発によって、

1) 5万語という豊富なボキャブラリにより表現の制約を意識せずに自然に話せる、2) 旅行会話の様々な会話を的確に訳すことができる、3) 携帯端末上でリアルタイムに動作可能という特徴を有する自動通訳システムを実現しました[4,5,6]。

## 2. 携帯電話向け翻訳

機械翻訳は、すでに携帯電話向け翻訳サイト「J-SERVER ポケット」(図4参照)[7]など、ASP事業として展開されていますが、音声翻訳機能が実現されることで、ストリーム情報を対象とすることができるようになり、ASP事業としても飛躍が期待されます。これらの事業は、2008年北京オリンピックに向けて急速に拡大することが見込まれ、今後ますます、機械翻訳が人々にとって身近なものになるでしょう。



図4:「J-SERVER ポケット」イメージ図

## 3. 今後の発展

自動通訳は、インターネットの発展とともにグローバル化する社会において、携帯端末や携帯電話を用いたユビキタスコミュニケーションツールとして注目を集めています。また、我々の開発した技術を利用することにより、グローバルなデータベース検索やオーダエントリ、対話エージェントなど様々なアプリケーションが組み込みシステムとして

大きく広がることが期待されます。この技術開発は、現在市場が拡大しつつある電子辞書や情報家電に翻訳機能を搭載可能とするもので、日本が世界に先行するものです。また、音声翻訳機能を携帯電話用プロセッサに搭載することも可能となり、半導体事業の拡大と競争力強化に大きく寄与するものです。本技術は、スケーラブルな拡張が可能であり、計算リソースの豊富なサーバに適用することにより、サーバ・クライアント型の高性能な自動通訳も可能となります。

携帯端末自動通訳は、自然言語処理研究者のみならず、世界の多くの人が望む夢です。我々は、その実現に向けて新たなステップを築くと同時に、従来にはない新しい分野において、機械翻訳技術の用途を拡大する道を開きました。これからも夢の実現に向けて進んでいきたいと考えています。

[1]<http://www.labs.nec.co.jp/Overview/soshiki/media/images/PDA1.mpg>

[2][http://www.labs.nec.co.jp/Overview/soshiki/media/images/translation\\_mp211.wmv](http://www.labs.nec.co.jp/Overview/soshiki/media/images/translation_mp211.wmv)

[3]<http://121ware.com/votol/>

[4] Okumura, et al.: An Automatic Speech Translation System for Travel Conversation, Human Language Technology Workshop, pp411-412,

[http://www.mt-archive.info/HLT-2002-Okumura.pdf\(2002/3/24\)](http://www.mt-archive.info/HLT-2002-Okumura.pdf(2002/3/24))

[5] Isotani, et all: An Automatic Speech Translation System on PDAs for Travel Conversation, Int Conf on Multimodal Interfaces (ICMI), pp211-216, (2002/10/15)

[6] 山端,他: PDA で動作する旅行会話向け日英双方向音声翻訳システム,情報処理学会自然言語処理研究会,Vol.2002 No.066, 2002-NL-150, (2002/07/15)

[7]<http://www.kodensha.jp/jservers/jsp/>

## OKIにおける機械翻訳技術開発

## ～第1回「AAMT長尾賞」を受賞して～

沖電気工業株式会社 村田 稔樹

私たち（沖電気工業株式会社研究開発本部機械翻訳研究グループ）は、昨年6月16日、アジア太平洋機械翻訳協会（以下AAMT）が主催する第1回「AAMT長尾賞」を受賞しました。これまでエポックメイキングとなる機械翻訳システムやサービスの開発運用を行ってきており、機械翻訳の実用化や普及に大きな貢献があったことを評価されたものです。

「AAMT長尾賞」は、機械翻訳の権威であり、またAAMTの創始者である長尾真氏（情報通信研究機構（NICT）理事長、京都大学元総長/名誉教授）が昨年「2005年日本国際賞」を受賞されたことを記念し設立され、機械翻訳の実用化の促進および研究開発に対して優れた貢献のあったと判断されるAAMT個人会員もしくは法人会員に属する個人またはグループに対して授与されるものであり、それを第1回目に受賞できたことは大変光栄に思っております。

以下では、受賞対象となった4つのシステム/サービスを順にご紹介し、最後に私たちが考える機械翻訳の今後について述べたいと思います。

まず最初は、1995年に世界で初めて考案および開発/商品化したWeb用機械翻訳システム

「PENSEE® for Internet」です。インターネットの登場により誰でも海外の膨大な情報に自由にアクセスすることが可能になった半面、それを活用するためには優れた英語能力が必要とされるという多くの日本人が持っていたジレンマを解消するための画期的なアプローチであり、クリック一つで図やリンクをもったWebページが見たままのレイアウトで翻訳されるさまは、驚きを持って迎えられました。また、機械翻訳を一般の人にとって身近で必要不可欠なものにするだけでなく、インターネット

そのものの普及に大きく貢献しました。現在でも、非常に多くのWeb翻訳関連製品やサービスが多く、企業から提供されており、機械翻訳業界全体の活性化をもたらしました。

2つ目は、1997年に開発した世界ではじめてJava™言語で書かれた機械翻訳エンジンです。このエンジンでは、単語、イディオム、簡単な言い回しや文全体を同じ翻訳パターンという形式で登録することができ、翻訳システム開発者のみならずユーザでも簡単に翻訳パターンを追加するだけで翻訳品質を向上させることのできる優れたカスタマイズ性を持っています。このエンジンではこの開発の過程では、まだ多くの課題を抱えていた登場間もないJavaの上で本格的な自然言語処理を構築した経験を活かし、開発会社のSun Microsystems社に積極的にフィードバックを行い、さらにJava Community Processなどの標準化活動にも参加し、Java言語の質の向上やUnicodeを用いた自然言語処理技術の確立に対しても貢献しました。

3つ目は、上記のJava版機械翻訳エンジンの優れたユーザカスタマイズ性を活かし、2004年からインターネット上において無償で公開しているコミュニティ型機械翻訳サイト「訳してねっと®」

（<http://www.yakushite.net/>）です。ユーザが興味を持つ分野の辞書をユーザが協力して登録することで、簡単に翻訳品質を向上させることができます。各人が入力した辞書は基本的に全ユーザが利用することができますので、一人一人が少しずつ様々な分野の辞書を登録することで、全体として様々な分野において高い翻訳品質の間で共有することができます。これは、翻訳の質の向上と辞書の充実という機械翻訳における本質的な問題を、ユーザ自身による評価・改善とユーザ同士の協調によって解決



しようという、まったく新しいコンセプトを世の中に提案するものです。英日/日英バージョンに加え、2006年1月には中日バージョンを追加しています。ユーザ参加型のウェブサイトとしても草分けであり、ユーザの幅広い興味やニーズにきめ細かく対応することができます。



最後は、2005年から行っている機械翻訳技術の特許検索への応用です。機械翻訳をクロスリンガル検索に応用することで、日本語で米国特許を検索できる機能をリコーテクノシステムズの「RIPWAY®」(http://www.ripway.net/)上にリコーグループと共同で作成し、機械翻訳の新たな応用先/ビジネス展開を見いだしました。国内特許公報とその英文抄録から約80の分野の専門用語辞書(のべ約140万語)を自動的に作成し、それらの辞書を適切に用いてすべての米国特許を日本語に翻訳することにより、日本語文章による検索が

可能となっています。実際の国内特許公報で使われている専門用語が訳出されているため、日本語文章による検索を高い精度で行えるのが特長です。また、米国特許データベースに対してあらかじめすべて機械翻訳してありますので、ユーザは機械翻訳を全く意識することなく検索することも大きな特長です。この日本語文章による米国特許検索サービスは、2006年7月から有償で「RIPWAY」にて提供されています。



今回の「AAMT 長尾賞」受賞は、これらの事業活動を通して、私たちの機械翻訳の実用化促進や研究開発への貢献が認められたものと考えています。今後、機械翻訳は様々なサービス、産業、場面で活躍することになると考えています。そのときは、前述の米国特許検索における機械翻訳のように、多くの場合、皆様からは直接見えないうちで機械翻訳システムが活躍していることと思います。私たちは、社会の必要不可欠なインフラとして発展する機械翻訳を支えるべく、幅広い研究開発に取り組んでまいります。

## 「The 翻訳サーバ Enterprise Edition V2.0」

東芝ソリューション株式会社

### 1. はじめに

東芝ソリューション株式会社は、「The 翻訳シリーズ」で定評ある翻訳エンジンを搭載し、ビジネスにおける様々な業種のニーズに応えたサーバ型翻訳ソフトウェア「The 翻訳サーバ Enterprise Edition V2.0」を発売しました。

The 翻訳サーバ Enterprise Edition V2.0 では、「意味トランスファ方式」に加え「CF 翻訳」と「2 段階翻訳」の採用により、以前にも増して文脈にふさわしい訳語を用いて翻訳を行えます。これらに加えて、メモリ辞書（翻訳メモリ）やユーザ辞書によって翻訳資産を蓄積・共有することで、統一された翻訳品質と高い翻訳精度を実現しています。

また、様々な利用環境で、多様な文書形態を扱えるようにするとともに、アプリケーション・インタフェースの提供によって、業務アプリケーションへの翻訳機能の組み込みを容易にしました。英語と日本語が混在した環境での業務を効率化し、ビジネスのグローバル化に対応できるようになっています。このグローバル化に伴い翻訳需要の急増している特許・医薬品業界における翻訳業務を支援するために、強力な専門の文法知識と辞書を備えた特許翻訳オプション、医療翻訳オプションも用意しています。

### 2. The 翻訳サーバ Enterprise Edition V2.0 の特長

#### (1) 自然言語処理技術を駆使し、高い翻訳精度を実現

最新版の翻訳エンジンは、文全体の意味に適した訳し分けを行う「意味トランスファ方式」、文書の文脈情報(Context)と書式情報(Format)から訳語を導き出す「CF 翻訳」、文書全体を先読みして得られた情報を利用する「2 段階翻訳」を採用し、より高度な翻訳精度を実現しています。

辞書の面では、190 万語の標準搭載辞書に加え、分野数・見出し語数ともに充実した 26 種類 482 万語の専門用語辞書の提供により、幅広い専門分野の翻訳要求に応じています。

#### (2) メモリ辞書やユーザ辞書で翻訳資産を蓄積・共有

過去の翻訳結果を原文とともに登録したメモリ辞書や業務用に作成したユーザ辞書などの翻訳資産を The 翻訳サーバ内に蓄積・共有できます。例えば、社内の翻訳部門が翻訳した社内資料、商談資料、ビジネスメール、契約書類など、日本語・英語の対となったデータから対訳や語彙を抽出し、対訳はメモリ辞書として、語彙はユーザ辞書として蓄積しておき、これらの翻訳資産を新たな翻訳に用いることで、全社的に統一された高品質な翻訳結果が得られます。また、付属の辞書メンテナンスツールを利用して辞書の編集を容易に行うことができ、翻訳資産の蓄積を支援します。

#### (3) 様々な利用環境で、多様なコンテンツに対応

ブラウザから URL やテキストを原文として指定して翻訳するホームページ翻訳やテキスト翻訳、原文を電子メールで送付して翻訳するメール翻訳、ホームページにボタンを埋め込んで翻訳する翻訳ボタンといった従来の機能に加え、本商品では、ブラウザで PC 上の文書ファイルを指定し、サーバ上で翻訳を行うファイル翻訳機能が利用できます。

HTML 文書や XML 文書、Microsoft®

Word/Excel/PowerPoint®の文書、PDF 文書などを指定できます。

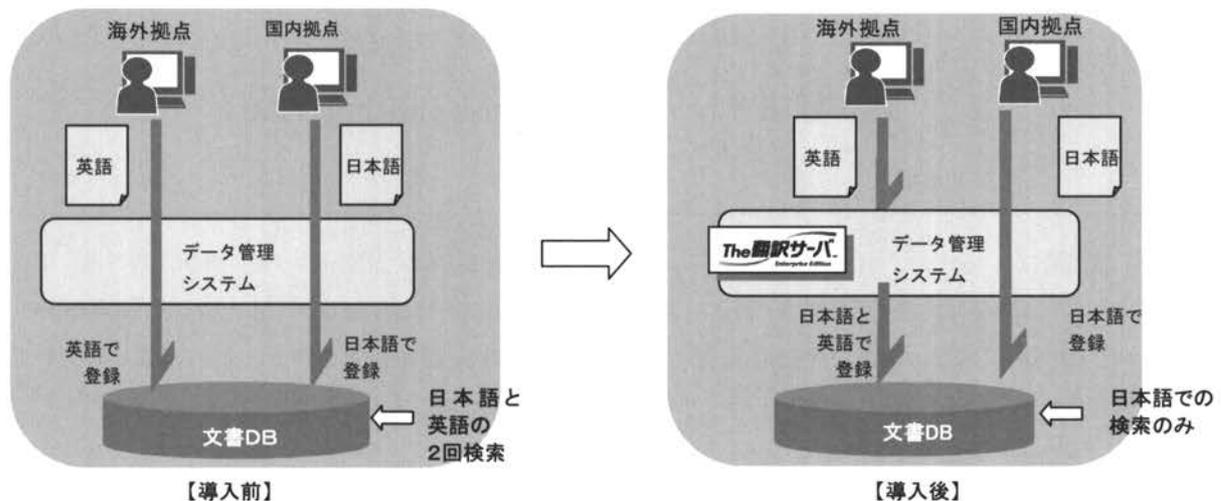


図 1: アプリケーションの組み込み例

#### (4) 業務アプリケーションへの組み込みが可能

アプリケーション・インタフェースの提供により業務アプリケーションに翻訳機能を組み込むことができるので、利用している環境やワークフローを変えることなく、日本語文書しか扱えなかったアプリケーションで英語文書が扱えるようになります。例えば、データ管理システムにおいて、登録データに日本語と英語が混在している場合、データベースから必要な情報を英語の情報もあわせて日本語で探し出し、参照・利用できるようにするためには、その都度翻訳作業が必要となります。データ管理システムに The 翻訳サーバを組み込めば、海外拠点からは従来通り英語で登録しても、本商品でこれを日本語に翻訳しておくことで、海外拠点からのデータも日本語によって検索が可能となります(図1)。このように、The 翻訳サーバにより日本語・英語の区別なく情報を一元的に管理・共有でき、業務効率の向上が図れます。

### 3. 特許翻訳オプション

The 翻訳サーバ Enterprise Edition V2.0 特許翻訳オプションでは、明細書や特許庁への意見書などの原文を翻訳前に解析して、翻訳しやすい文体へ自動変換します。これを特許専用文法・辞書で解析することによって、特許特有用語・表現や、請求項などの長文を適切に翻訳します。また、特許専門用

語辞書 16 万語、特許用例 2 万例は専門分野の翻訳が多い特許翻訳を強力にサポートします。

これを特許の外国出願や公知例調査等に利用すると、一連の作業を効率化できます。例えば、欧米の特許調査では、従来は英文でアブストラクトを読み重要度を判断していたのに対し、本商品と特許オプションを特許検索システムに組み込むことにより、日本語で重要度を判断できるようになります。このため、短時間でより広範囲な外国特許調査が可能となり、問題特許の見落としの削減が期待できます。

### 4. 医療翻訳オプション

The 翻訳サーバ Enterprise Edition V2.0 医療翻訳オプションでは、医学・薬学分野のうち、特に製薬業界向けに以下のような点を強化しました。

- (a) 医学・薬学分野専用の文法知識
- (b) 性状・品質適正化の基準となる「日本薬局方」に基づく薬学辞書、メモリ辞書、パターン辞書

これらにより、医学・薬学分野の文書に適した翻訳結果を得ることができます。例えば、「本品 0.05g を酢酸(100)5mL に溶かし、臭素試液 0.5mL を加えるとき、試液の色は直ちに消える。」を” Dissolve 0.05g of it in 5 mL of acetic acid (100), and add 0.5 mL of bromine TS: the color of a test solution disappears immediately.” のように製薬文書向けの文型で訳出できます。

さらに、本オプションでは、南山堂「医学英和大辞典（第11版）」/「プラクティカル医学略語辞典」に基づく医学辞書も提供しています。

The 翻訳サーバでは上記医療翻訳オプション以外にも、以下の医学・薬学向け専門用語辞書を提供し、医療文献の翻訳の助けとしています。(a) 医学・薬学分野（英日 279,000 語、日英 254,000 語）\*、(b) 日外アソシエーツ「医学用語大辞典」に基づく医学辞書（英日 197,000 語、日英 187,000 語）。\* MedDRA(日米欧間で定められた国際的に統一された医薬規制用語集)に配慮した用語を含む。

## 5. 動作環境

ハードウェア: PC/AT 互換機(MAGNIA シリーズ推奨)

OS(日本語版): Windows® Server 2003 Service Pack 1、Windows® Server 2003 R2、Windows® 2000 Server Service Pack 4

メモリ(The 翻訳使用分): 192MB

(翻訳を並列処理する場合は、処理を 1 件追加する毎に 128MB が追加が必要)

ディスク: 1GB 以上(インストール時)

(その他、キャッシュ/スプール領域として 1GB 以上推奨)

その他: JDK 5.0、

Internet Information Services 5.0/6.0、FTP、CGI が動作する環境

※ ファイル翻訳を利用する場合は、以下のアプリケーションが必要

doc ファイルの翻訳には Microsoft®Word 2003

xls ファイルの翻訳には Microsoft®Excel 2003

ppt ファイルの翻訳には PowerPoint® 2003

pdf ファイルの翻訳には Adobe®Acrobat®7.0

## ■ 製品情報 URL

<http://mt-server.toshiba-sol.co.jp/>

■ Microsoft、Windows、PowerPoint は、米国 Microsoft Corporation の米国およびその他の国における登録商標または商標です。

■ Adobe 及び Acrobat は、Adobe Systems Incorporated の登録商標または商標です。

■ The 翻訳、The 翻訳サーバは、東芝ソリューション株式会社の登録商標または商標です。

■ 記載されている商品の名称はそれぞれ各社が商標として使用している場合があります。

■ 本製品の仕様は予告なく変更される場合があります。

## My サイト翻訳

株式会社高電社

株式会社 高電社(本社：大阪市阿倍野区、代表取締役社長：岩城陽子)では、現在、法人向けにホームページ自動翻訳の ASP サービスを精力的に全国展開している。

この ASP サービスの中でも主力と位置づけているのが「ボタンひとつ自動翻訳」と称する「My サイト翻訳」サービスである。このサービスは、顧客の日本語ホームページ上に所定の外国語翻訳用リンクタグを貼付するだけの作業で、サイト訪問者がこのボタンを選択すれば、即座に外国語に自動翻訳して表示する。

このサービスに導入されている翻訳エンジンは、英語・中国語(繁体字/簡体字)・韓国語の最大4言語に対応し、特に同社が四半世紀をかけて自社開発してきた中国語、韓国語翻訳エンジンの翻訳精度は、国内外を問わず高く評価され、検索サイト excite や NTT ドコモ、KDDI、SoftBank、WILLCOM などの携帯電話公式サイトでも採用され、コンシューマ向け翻訳サービスを提供している。

またこの「My サイト翻訳」には、次のような特徴的な機能を搭載し、現在も成長を続けている。

### ■「My サイト翻訳」の特徴となる機能

- 翻訳結果をカスタマイズできる顧客専用オンラインユーザー辞書・文章登録サイトの提供
- 翻訳と同時に画像や動画、音声などサイト上に貼付のデータの差し替え機能
- 特殊タグによる翻訳非対象のページや部分指定機能
- 自動翻訳されたページも海外の検索サイトからヒットされる静的ページ自動保存機能
- SSL(暗号化された)ページも翻訳できる機能

高電社は、約3年半前に開始された外国人旅行者の訪日を促す「ビジット・ジャパン・キャンペーン」と共に、これら法人向け ASP サービスを構想、サービスの改良を手掛けながら企業や大学、官公庁、自治体等のあらゆる市場に、あらゆる用途で提供を行ってきたが、特に近年では中国・韓国との人的・経済的交流が活発になってきており、企業・観光・自治体関連への導入事例が増えてきているという。その背景としては、下記のような要因があると分析している。

### ■「My サイト翻訳」の導入の背景

- アジアと日本の往来が活発化している
- 日本企業の持つ技術・品質が海外から注目されている
- 中国・韓国の経済的発展と交流強化の動き

同社では、「My サイト翻訳」のサービス以外にも、掲示板やブログ、チャットなどのシステムと連動できる ASP サービス「翻訳ゲートウェイ」なども展開しており、人的翻訳による WEB ページと自動翻訳サービスを融合して提供するなど、外国語ページの総合的なコンサルタント事業へと拡大し、営業強化も図っている。

下記に代表的な導入事例を紹介する。

### ■「My サイト翻訳」の導入事例

<自治体>

岡山県倉敷市役所ホームページ

英語・中国語・韓国語で観光情報や生活情報を発信。観光客の招聘、市民向け情報サービス、災害時の情報インフラとしての効果に期待を寄せている。

<企業>

日本卓球株式会社ホームページ

英語で製品情報や試合速報を発信。海外のプレーヤー一層の拡大や、売上アップに期待している。



岡山県倉敷市役所 HP :

<http://www.city.kurashiki.okayama.jp/>



学校法人 亜細亜大学 HP :

<http://www.asia-u.ac.jp/>



日本卓球株式会社 HP :

<http://www.nittaku.com/>

## <大学>

学校法人 亜細亜大学ホームページ

英語・中国語(繁体字/简体字)・韓国語で新着情報や研究室の幅広い情報などを発信。留学生の受け入れと短期、長期の学生を国外に送り出している。

## ■ 本サービスに関するお問い合わせ先 株式会社 高電社

### ・大阪本社

大阪市阿倍野区昭和町 3-7-1

IT 事業部 J-SERVER 事務局

TEL : 06-6628-7730

### ・東京支店

東京都千代田区神田須田町 1-18

第 3 神田セントラルビル 2F

IT 事業部 J-SERVER 事務局

TEL : 03-3256-3061(代表)

E-Mail : [j-server@kodensha.jp](mailto:j-server@kodensha.jp)

## インターネット機械翻訳サービスリスト

AAMT インターネットワーキンググループ

- ★ 本リストでは、AAMT 独自調査による機械翻訳サービスを行なっている Web ページを記載しています。辞書引きサービスは原則として非掲載です。本リストに記載されていない機械翻訳サービス、あるいは記載されているサービスの URL 変更/停止などがございましたらご連絡頂けると幸いです。
- ★ 言語欄は、翻訳前あるいは翻訳後の言語として指定可能な言語を表し、翻訳方向は表していません。
- ★ 以下のサイトにアクセスしたことによる一切の損害を当協会は補償いたしません。
- ★ 順不同

サイト(サービス)名	サービス内容	対象言語
ajeab	Web/Text	英アラビア
	<a href="http://english.ajeab.com">http://english.ajeab.com</a>	
AltaVista Babel Fish Translation	Web/Text	英仏独西伊葡中日
	<a href="http://world.altavista.com/">http://world.altavista.com/</a>	
AMIKAI	Text	英仏独西葡中日韓
	<a href="http://www.amikai.com">http://www.amikai.com</a>	
@nifty	Web/Text	英日韓中
	<a href="http://www.nifty.com/globalgate/">http://www.nifty.com/globalgate/</a>	
ATLAS	Text	英日
	<a href="http://software.fujitsu.com/jp/atlas/sample.html">http://software.fujitsu.com/jp/atlas/sample.html</a>	
BIGLOBE	Web/Search	英中日韓
	<a href="http://honyaku.biglobe.ne.jp/">http://honyaku.biglobe.ne.jp/</a>	
clickq Japan	Web	日韓
	<a href="http://www.clickq.com">http://www.clickq.com</a>	
Cyber G·BOC	Web/Text	英日中
	<a href="http://www.g-boc.com/j/cyber/search/index.html">http://www.g-boc.com/j/cyber/search/index.html</a>	
英語ナビ	Web/Text	英日韓
	<a href="http://eigo-navi.com/">http://eigo-navi.com/</a>	
ePal.com	Text	英仏独西伊葡中
	<a href="http://www.epals.com/translation/">http://www.epals.com/translation/</a>	
Excite	Web/Text/MS Office/Search	英日中韓
	<a href="http://www.excite.co.jp/world/">http://www.excite.co.jp/world/</a>	
Foreignword.com	Text	25 言語
	<a href="http://www.foreignword.com/Tools/transnow.htm">http://www.foreignword.com/Tools/transnow.htm</a>	
Free Translation.com	Web/Text	英仏独西蘭葡日ノルウェイ
	<a href="http://www.FreeTranslation.com/">http://www.FreeTranslation.com/</a>	

Infoseek	Web/Text	英中日韓仏独伊西葡
	<a href="http://www.infoseek.co.jp/Honyaku">http://www.infoseek.co.jp/Honyaku</a>	
石川新情報書府	Web	英日
	<a href="http://shofu.pref.ishikawa.jp/">http://shofu.pref.ishikawa.jp/</a>	
J-Server	Web/Text/Chat/Mail	英中日韓
	<a href="http://www.j-server.com/index.shtml">http://www.j-server.com/index.shtml</a>	
金沢 100 万石まつり	Web	英日
	<a href="http://www.100mangoku.net/">http://www.100mangoku.net/</a>	
韓国 Web 旅行	Web/Text/Chat/BBS	日韓
	<a href="http://korea.hanmir.com/">http://korea.hanmir.com/</a>	
関西コンシェルジェ	Chat/Mail/BBS	日韓
	<a href="http://www.my-concierge.org/kansai/index.htm">http://www.my-concierge.org/kansai/index.htm</a>	
国立健康・栄養研究所	News	英中日
	<a href="http://humpty.nih.go.jp/">http://humpty.nih.go.jp/</a>	
Language Engineering	Web/Text/File/Mail	英仏独西伊葡中日
	<a href="http://www.lec.com/">http://www.lec.com/</a>	
LogoVista e-Trans	Web/Text	英仏独西伊葡日韓
	<a href="http://e-trans.logovista.co.jp/">http://e-trans.logovista.co.jp/</a>	
Microsoft Office Update Home	Text	英日
	<a href="http://software.fujitsu.com/jp/honyaku/eservice/index.html">http://software.fujitsu.com/jp/honyaku/eservice/index.html</a>	
Newstran.comm	Web/Text/News	英仏独伊西葡中日韓アラビア
	<a href="http://www.newstran.com/">http://www.newstran.com/</a>	
新居浜コンシェルジェ	Chat/Mail/BBS	日韓
	<a href="http://www.e-niihama.jp/">http://www.e-niihama.jp/</a>	
Paran Go Japan!	Web/Text	日韓
	<a href="http://gojapan.paran.com/index.html">http://gojapan.paran.com/index.html</a>	
OCN	Web/Text	日英韓
	<a href="http://www.ocn.ne.jp/translation/">http://www.ocn.ne.jp/translation/</a>	
ParSit	Web/Text	英泰
	<a href="http://suparsit.com">http://suparsit.com</a>	
Poltran.com	Text	英葡
	<a href="http://www.poltran.com/">http://www.poltran.com/</a>	
ProMT	Web/Text/Mail	英仏独西伊露
	<a href="http://www.online-translator.com/text.asp?lang=en">http://www.online-translator.com/text.asp?lang=en</a>	
Q 州マイコンシェルジェ	Web/Chat/Mail/BBS	日韓
	<a href="http://www.my-concierge.org/kyushu/">http://www.my-concierge.org/kyushu/</a>	

ReaD	Text	英日
	<a href="http://read.jst.go.jp/">http://read.jst.go.jp/</a>	
Reverso	Web/Text	英仏独西
	<a href="http://www.reverso.net/">http://www.reverso.net/</a>	
Rustran.com	Text	英露
	<a href="http://www.rustran.com/">http://www.rustran.com/</a>	
Sakhr	Text	英アラビア
	<a href="http://mason.gmu.edu/~aross2/sakhr.htm">http://mason.gmu.edu/~aross2/sakhr.htm</a>	
瀬戸内・夢海道コンシェルジェ	Chat/Mail/BBS	日韓
	<a href="http://www.my-concierge.org/shikoku/syodoshima/">http://www.my-concierge.org/shikoku/syodoshima/</a>	
Sister Cities International	Text	英仏独西伊葡日
	<a href="http://www.sister-cities.org/sci/resources/translation">http://www.sister-cities.org/sci/resources/translation</a>	
So-net	Web/Text	英日
	<a href="http://www.so-net.ne.jp/translation/">http://www.so-net.ne.jp/translation/</a>	
SPORTS CHANBARA	Web	英仏独西伊葡中日韓露泰
	<a href="http://www.internationalssportschanbara.net/entry.html">http://www.internationalssportschanbara.net/entry.html</a>	
sScience.com	Text	英日
	<a href="http://www.sscience.com/index_J.htm">http://www.sscience.com/index_J.htm</a>	
SYSTRANBox	Web/Text	英仏独西伊葡中日韓露
	<a href="http://trans.voila.fr/">http://trans.voila.fr/</a>	
Systranet.com	Web/Text/File	英仏独西伊葡希
	<a href="http://www.systranet.com/systran/net">http://www.systranet.com/systran/net</a>	
特許庁	特許公報	英日
	<a href="http://www.ipdl.ncipi.go.jp/homepg_e.ipdl">http://www.ipdl.ncipi.go.jp/homepg_e.ipdl</a>	
TransLand	Text	英日
	<a href="http://www.brother.co.jp/jp/honyaku/demo/index.html">http://www.brother.co.jp/jp/honyaku/demo/index.html</a>	
Transtar Web	Web/Text	英中日
	<a href="http://www.transtar.com.cn/en/transtech/transonline.asp">http://www.transtar.com.cn/en/transtech/transonline.asp</a>	
TransWalker JC	Text	日中
	<a href="http://www.kodensha.jp/jis/soft/pdamt/index.html">http://www.kodensha.jp/jis/soft/pdamt/index.html</a>	
TransWalker JK	Text	日韓
	<a href="http://www.kodensha.jp/jis/soft/pdamt/index.html">http://www.kodensha.jp/jis/soft/pdamt/index.html</a>	
TransWalker JET	Text	日英
	<a href="http://www.kodensha.jp/jis/soft/pdamt/index.html">http://www.kodensha.jp/jis/soft/pdamt/index.html</a>	
Transwiz	Web/Text	英日韓
	<a href="http://transwiz.mobico.com/">http://transwiz.mobico.com/</a>	

通連国際	Web/Text/Chat/File/Mail	英独中日露
	<a href="http://www.165net.com/jpindex.jsp">http://www.165net.com/jpindex.jsp</a>	
VIL Net	Web/Text	英日
	<a href="http://www.vil-net.jp/hon/">http://www.vil-net.jp/hon/</a>	
Yahoo Korea	Web	日韓
	<a href="http://kr.japan.yahoo.com/">http://kr.japan.yahoo.com/</a>	
ToggleText	Web/Text	英インドネシア
	<a href="http://www.toggletext.com/">http://www.toggletext.com/</a>	
金橋訳港世界通	Web/Text/Mail	英中日
	<a href="http://www.netat.net/">http://www.netat.net/</a>	
マジカルゲート	Web/Mail	英中日韓葡仏伊西独
	<a href="http://www.magicalgate.net/ja/index.html">http://www.magicalgate.net/ja/index.html</a>	
Al-Misbar	Web	英アラビア
	<a href="http://www.almisbar.com/salam_trans.html">http://www.almisbar.com/salam_trans.html</a>	
コネクト・インタナショナル	Mail	英日
	<a href="http://www.connect-i.co.jp/flextran/index.html">http://www.connect-i.co.jp/flextran/index.html</a>	
Tour2Korea	Web	日韓
	<a href="http://japanese.tour2korea.com/12Home/translator.asp?konum=7">http://japanese.tour2korea.com/12Home/translator.asp?konum=7</a>	
韓日文化交流	Web/Text/Mail	日韓
	<a href="http://www.kjclub.com/jp/">http://www.kjclub.com/jp/</a>	
alittera translation service	Text	英仏独伊西葡露
	<a href="http://www.alittera.com/Auto.asp">http://www.alittera.com/Auto.asp</a>	
livedoor Translate	Web/Text	英日中韓
	<a href="http://translate.livedoor.com/">http://translate.livedoor.com/</a>	
livedoor ギガメラー	Mail	英日中韓
	<a href="http://docs.livedoor.com/personal/">http://docs.livedoor.com/personal/</a>	
Worldlingo	Web/Text/Mail	日中蘭英仏独希伊韓露葡西
	<a href="http://www.worldlingo.com/ja/products_services/computer_translation.html">http://www.worldlingo.com/ja/products_services/computer_translation.html</a>	
Tango Town	Web	日英
	携帯電話専用	
NewsWatch	Text	英日
	<a href="http://www.newswatch.co.jp/translate/">http://www.newswatch.co.jp/translate/</a>	
LISA	Text	英独仏蘭希西伊葡中日韓露
	<a href="http://www.lisa.org/">http://www.lisa.org/</a>	
訳してねっと	Web/Text	英日中
	<a href="http://www.yakushite.net/">http://www.yakushite.net/</a>	

Google Language Tools	Web/Text	英独仏西葡
	<a href="http://www.google.com/language_tools?hl=en">http://www.google.com/language_tools?hl=en</a>	
CJT Survey	Mail	日英中韓仏伊葡
	<a href="http://www.cjt-survey.com/contents/mail-trans.htm">http://www.cjt-survey.com/contents/mail-trans.htm</a>	
宮崎市	Web/Text	日英韓中
	<a href="http://www.city.miyazaki.miyazaki.jp/gyousei/html/common/tran/">http://www.city.miyazaki.miyazaki.jp/gyousei/html/common/tran/</a>	
Dictionary.com	Text	日英韓中仏独蘭希伊葡露
	<a href="http://dictionary.reference.com/translate/text.html">http://dictionary.reference.com/translate/text.html</a>	
DreamWiz	Web/Text/Search	日韓
	<a href="http://japan.dreamwiz.com/">http://japan.dreamwiz.com/</a>	
Fresh eye	Web/Text	日英中
	<a href="http://www.fresheye.com/">http://www.fresheye.com/</a>	
Babelplex	Search	英中蘭仏独希伊日韓葡露
	<a href="http://www.babelplex.com/">http://www.babelplex.com/</a>	
Yahoo! Japan	Web/Text	英中日韓
	<a href="http://honyaku.yahoo.co.jp/">http://honyaku.yahoo.co.jp/</a>	
メールピア	Web 変更情報	英日
	<a href="http://www.asahi-net.or.jp/service/mail/mailpia/">http://www.asahi-net.or.jp/service/mail/mailpia/</a>	
xPost	Mail	英中日
	<a href="http://www.tisiki.net/">http://www.tisiki.net/</a>	
Mitaka - the big word	Web/Text	英仏独伊西葡日
	<a href="http://www.thebigword.com/MachineTranslation.aspx">http://www.thebigword.com/MachineTranslation.aspx</a>	
YAMATO 英和	Text	英日
	<a href="http://ogma.atnifty.com/ogma.htm">http://ogma.atnifty.com/ogma.htm</a>	
地球の歩き方	Text	英日韓中独葡
	<a href="http://www.arukikata.co.jp/translation/">http://www.arukikata.co.jp/translation/</a>	
GoKorea 日韓フレンズ	Web/Chat	韓日
	<a href="http://www.gokorea.jp/">http://www.gokorea.jp/</a>	
モバイル Yahoo!	Text	日英
	<a href="http://honyaku.mobile.yahoo.co.jp">http://honyaku.mobile.yahoo.co.jp</a>	
プラザトウキョウ・ドット・コム	Text	日英西伊葡仏露
	<a href="http://www.plazatokyo.com/translate.php">http://www.plazatokyo.com/translate.php</a>	
e 翻訳 ドット ネット	Search	日中韓英仏伊西独葡
	<a href="http://ehonyaku.net/">http://ehonyaku.net/</a>	
Yahoo! Babel Fish	Web/Text	13 言語
	<a href="http://babelfish.yahoo.com/">http://babelfish.yahoo.com/</a>	

宝贝魚	Web/Text	13 言語
	<a href="http://fanyi.cn.yahoo.com/">http://fanyi.cn.yahoo.com/</a>	
熟考	Web/Text	英日
	<a href="http://www.rozetta.jp/index.html">http://www.rozetta.jp/index.html</a>	

# 協会活動報告

(2006年4月～2007年2月)

## 総会

2006年6月16日

- |                       |                       |
|-----------------------|-----------------------|
| ① 第1号議案 2005年度事業報告(案) | ② 第2号議案 2005年度決算報告(案) |
| ③ 第3号議案 2006年度事業計画(案) | ④ 第4号議案 2006年度予算(案)   |
| ⑤ 第5号議案 役員改選(案)       | ⑥ 第6号議案 AAMT移転の件      |

## 報告会

2006年6月16日

- |                      |             |
|----------------------|-------------|
| ① 技術動向委員会            | ② 市場動向委員会   |
| ③ AAMT/Japio 特許翻訳研究会 | ④ インターネットWG |

## 講演会

2006年6月16日

「人間の誤り、機械の誤り：外国語学習者のコーパスを見て考えること」

明海大学 教授 投野 由紀夫 氏

「知識処理と学習・環境適合技術の時代」

情報通信研究機構 理事長／元 AAMT 会長 長尾 眞氏

## 懇親会

2006年6月16日 新お茶の水ビル・サンクレーン2階「ビストロ備前」

## 決算理事会

2006年6月12日

- |                       |                       |
|-----------------------|-----------------------|
| ① 第1号議案 2005年度事業報告(案) | ② 第2号議案 2005年度決算報告(案) |
| ③ 第3号議案 役員改選(案)       | ④ 第4号議案 AAMT移転の件      |
| ⑤ 第5号議案 役員改選(案)       | ⑥ その他                 |

## 監事会

2006年4月27日 平成17年度会計報告

## 運営委員会

2006年6月1日

- |                 |             |
|-----------------|-------------|
| ① 理事会、総会準備      | ② H17年度決算報告 |
| ③ 新 AAMT 体制について | ④ その他       |

## 市場動向調査委員会

2006年4月10日

- ① 17年度ベンダー調査報告
- ② 今後の活動について
- ③ その他

2006年5月12日

- ① 17年度ベンダー調査報告
- ② 今後の活動について
- ③ その他

2006年6月5日

- ① 今までの委員会活動を振り返って
- ② 今後の活動について
- ③ その他

## 機械翻訳課題調査委員会準備委員会

2006年4月27日

- ① 委員会趣旨説明
- ② 各委員会活動報告
- ③ 今後の活動に関する討議
- ④ その他

2006年6月1日

- ① 今後の活動に関する討議
- ② その他

2006年6月23日

- ① 活動内容について
- ② 新委員会、新メンバーの構成について
- ③ その他

## 機械翻訳課題調査委員会

2006年9月22日

- ① 新委員（長）の紹介
- ② 委員会の構成と活動内容について
- ② WGの構成
- ③ 各WGの活動についての議論
- ④ 活動計画の報告（各WGから）
- ⑤ 全体活動計画についての議論
- ⑥ その他

2006年10月30日

- ① MT評価に関する勉強会
- ② 各WGの活動についての議論
- ③ 活動計画の報告（各WGから）
- ④ 全体活動計画についての議論
- ⑤ その他

2006年11月27日（WG②「調査・広報・啓蒙」個別委員会）

- ① ユーザーアンケートについて
- ② MT動向調査について
- ③ 無料翻訳サイトの調査状況
- ④ その他

2006年12月18日

- ① MT 研究開発最新動向に関する勉強会
- ② 各 WG の活動についての議論
- ③ 活動内容の報告(各 WG から)
- ④ 全体活動内容についての議論
- ⑤ その他

2007年1月25日 (WG②「調査・広報・啓蒙」個別委員会)

MT ユーザアンケートのための最終調整

2007年1月26日～2月9日 AAMT Web ページ上にて「AAMT MT ユーザアンケート」の実施

2007年2月7日

- ① これまでの活動の総括
- ② 各 WG の活動についての議論
- ③ 活動内容の報告(各 WG から)
- ④ 全体活動内容についての議論
- ⑥ その他

## AAMT/Japio 特許翻訳研究会

2006年4月21日

- ① 新メンバー紹介
- ② H17 年度報告書について
- ③ 今年度の活動について
- ④ その他

2006年5月19日 Japio 訪問 (自由参加)

2006年6月8日

- ① Japio システム見学(5/19)報告
- ② 今後の活動について
- ③ その他

2006年7月13日

- ① 新メンバー紹介 (多摩美大 安藤氏)
- ② 報告1「特許対訳コーパスの規模別特性の研究」
- ③ 報告2「PAJ タイトルを用いた特許翻訳の自動評価」
- ④ 報告3「日英対訳を基に日本語部分の構文構造を求める手法の研究」
- ⑤ 今後の事務局運営について

2006年9月1日

- ① 研究会の運営体制について
- ② MT 評価について
- ③ 報告1「特許対訳コーパスの規模別特性の研究」
- ④ 報告2「Japio コーパスからの「一般用語の特殊訳例」のオミットについて」
- ⑤ 今年度の研究活動について

2006年10月6日

- ① 新任挨拶 (Japio 奥氏)
- ② 今年度の研究活動について
  - a. 自動評価
  - b. Japio コーパスの特殊訳例
  - c. 訳し分けの (自動) 評価
  - d. 次回サミットでの WS など

2006年12月1日

- ① 帰国挨拶 (Japio 大塩氏)
- ② 新メンバー紹介 (ATR 隅田氏)
- ③ 今年度の研究活動について
- ④ 報告1「BLEUの問題点について」
- ① 報告2「日英翻訳における対訳曖昧性に関する調査について」

2007年1月26日

- ① 今年度の研究活動について
- ② 今年度の報告書作成について
- ③ 次回サミットでの WS 開催について

## 編集委員会

2006年9月29日

- ① AAMT Journal No. 39 の企画について
- ② 次々号以降の企画について
- ③ 今後の AAMT Journal の形態について
- ④ その他

2006年10月30日

- ① AAMT Journal No. 39 の企画内容の確定
- ② 次々号以降の企画について
- ③ AAMT HP 上での公開について (インターネット WG リーダー参加)
- ④ その他

## インターネットワーキンググループ

- ① AAMT Web page の更新 (毎月)
- ② AAMT FORUM での情報発信 (月数回)
- ③ 委員会活動におけるメイリングリストの管理
- ④ その他ネットワーク関連のインフラ管理全般

# MT SUMMIT XI

September 10-14, 2007

Copenhagen Business School, Copenhagen, Denmark

The Eleventh Machine Translation Summit, organized by the European Association for Machine Translation (EAMT), will be held in Copenhagen, Denmark, 10-14 September 2007.

MT Summit XI will bring together scientists, developers and users of machine translation and other tools for multilingual processing. Cutting-edge research and demonstrations from leading academic, business and government organizations will showcase the latest results. During the main conference, MT Summit XI will have numerous sessions including research papers, reports on users' experiences, discussions of policy issues, invited talks and panel discussions. Before the main conference, the program includes two full days of tutorials and workshops. These wide-ranging activities in the machine translation will allow participants from research, business and government to get together to explore how to use machine translation technology to conquer language barriers and overcome the digital divide.

The EAMT cordially invites you all -- researchers, developers, providers, users, and observers -- to MT Summit XI, to contribute to and witness ongoing efforts in developing and using machine translation technology and tools for translation support.

For more details, please visit the EAMT website: <http://www.eamt.org/cfp/MTS-1stcall.htm>

2007年9月10日(月)～14日(金)

デンマーク、コペンハーゲン

European Association for Machine Translation (EAMT)主催による第11回機械翻訳サミット(MT Summit XI)が2007年9月、デンマークのコペンハーゲンにて開催されます。

MT Summit XIは、機械翻訳分野の研究者、開発者、ユーザ等の幅広い関係者を対象としています。主要な学術、民間、政府機関等による最新の研究発表やデモが予定されています。サミットの本会議では、研究論文発表、ユーザレポート、招待講演、パネル討論等が用意されています。更に、本会議前後の拡張プログラムとして、チュートリアル、ワークショップ、デモセッション等が予定されています。このように機械翻訳にかかわる幅広い層の参加者が一堂に会することにより、言語障壁やデジタルディバイドを克服するための議論が活発に行われるものと期待しております。

※ AAMT・Japio 委員会主催の特許翻訳に関するWSも予定されています。

MT Summit XIへの皆さまのご参加をお待ちしております。

発表申込み、参加申込み等の詳細に関しては、以下 EAMT のウェブページをご参照ください。

<http://www.eamt.org/cfp/MTS-1stcall.htm>

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会員の皆様、はじめまして。

昨年6月の総会より、辻井前会長から井佐原新会長へと体制の交代が行われ、井佐原新会長体制のもとで、AAMT ジャーナル編集委員会委員長を仰せつかりました。編集委員長という不慣れな役割ではございますが、経験豊富な編集委員、事務局、前編集委員長（新会長）の方々からの強力なサポートを受けて、会員の皆様のご期待に沿えるような紙面を作っていく所存です。

さて、今号は、井佐原均新委員長体制のもとでの最初の号であります。新体制のもとで、まずは表紙のデザインを変更するべし、との新会長命を受けまして、編集業務担当の(株)ナビックス様のご協力のもとで、複数案についての審議を経て、今回お手元に届きましたデザインとなりました。掲載記事につきましては、会長交代に伴います各委員会の新体制の紹介、新旧会長からのご挨拶をはじめとしまして、通常総会および関連行事の報告、翻訳関係催し物の報告、ベンダーからのレポート等を掲載しました。特に総会関連行事として、このたびAAMT長尾賞が創設されましたので、第一回受賞者の法人会員からのご寄稿を掲載いたしました。

次号以降の企画案につきましても、編集委員会で案を練っているところですが、新企画の芽として、翻訳サービス提供側の会員と翻訳サービス享受側の会員がもっと相互に知り合えるような企画、機械翻訳ビジネスの最新動向、海外機械翻訳動向といった案が出ていますので、順次取り組んでいきたいと考えています。会員の皆様からの企画案も積極的に取り上げたいと思いますので、案がございましたら、ぜひ編集委員会までお寄せ下さい。

*MEMO*



# AAMT

## AAMTジャーナル No.39

---

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 自然言語グループ内

phone：0774-93-4625 fax：0774-93-4627

編集委員会：宇津呂 武仁 大倉 清司 熊野 明

松川 淑子 村上 嘉陽

事務局：神崎 享子 和泉 絵美

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>

