ISSN 1883-1818
No.75
December 2021

AAMT Journal

Asia-Pacific Association for Machine Translation

機 械 翻 訳

桟敷物訳

AAMT 30th ANNIVERSARY

AAMT創立30周年特別記念号

AAMT Journal
Asia-Pacific Association for Machine Translation 機械翻訳
No.75

目　次

C O N T E N T S

巻頭言

『機械翻訳』を訳した翻訳者が考えていること

高橋 聡
個人翻訳者（日本翻訳連盟・副会長）

「機械翻訳」とは何か？　どこから来て、どこへいくのか？

これは、2020 年 9 月に刊行された拙訳書『機械翻訳 歴史・技術・産業』（ティエリー・ポイボー[著]、中澤敏明[解説]、高橋聡[翻訳]）の帯に、出版元である森北出版の担当編集者が付けてくれたキャッチコピーである。そして、日本翻訳連盟（JTF）が今年 4 月、創立 40 周年を記念して開催した第 1 回関西セミナーのタイトルにも、このフレーズを使わせていただいた。

なお、この訳書については、『AAMT Journal』No.74 で、立教大学の山田優先生が、実に詳しい書評を書いてくださった。解説と論評は、さすが専門家と唸ってしまう鋭さで、ありがたいことに私の翻訳にも言及してくださっている。この場を借りて、心よりお礼の言葉をお伝えしたい。

さて、冒頭のキャッチコピーは、現在の翻訳業界が機械翻訳を捉えるヒントとして実に的確だったと思う。そこで、以下ではこのコピーをもとにして、いち個人翻訳者がいま機械翻訳について考えていることを綴ってみたい。

●機械翻訳とは何か？

手前味噌で恐縮ながら、この問いの答えを本書はとてもコンサイスにまとめている。いろいろな機械翻訳技術のしくみと歴史、それが必要になった理由と開発の経緯が時代ごとに整理されており、さらに産業としての現状まで述べられている。類書のなかでも初めてではないだろうか。しくみを知れば、おのずと限界もはっきりする。技術的（数学的）に難解な要素はできるだけ省いてあるので、業界関係者には、このくらいのレベルで機械翻訳の概要を抑えていてほしい。その土台があって初めて議論が成り立つ。

●機械翻訳はどこから来たのか？

特に同業者、つまり翻訳者や通訳者にいちばん知っていただきたいのはこの点だ。「翻訳」というものを考えるとき、ヨーロッパ（中近東までを含む）と日本では、ベースとなる地理的・文化的・言語的土壌があまりにも違う。

ヨーロッパは、いわば"翻訳の本場"だ。複数の民族、複数の言語と文化が陸続きで混在・共存する世界。他民族を「バルバロイ」つまり「わけの分からない言葉を話す者」と呼んだり、バベルの塔の逸話を聖書に記したり、そこにあるのは常に、複数言語の存在が当然という意識だった。聖書自体も翻訳と深い関係にある。早くからヘブライ語、アラム語、ギリシャ語、ラテン語など複数の言語間で当たり前のように翻訳の需要があった。そういう伝統の末にあるのが、23 言語を公用語とする現在の EU なのだ。

そういう言語文化世界だからこそ、ライプニッツやデカルトの時代から普遍言語という概念が生まれ、エスペラントに代表される人工言語が考案された。そうした土壌があれば、20 世紀早々から機械翻訳という発想が芽ばえたのは、当然と言えるだろう（このくだりは、『機械翻訳』第 4 章に詳しい）。

一方の日本は、江戸末期から今日に至るまで一貫して、堂々たる"翻訳大国"だ。だが、そこで行われてきた翻訳のほとんどは「他言語から日本語へ」、なかでも「英語から日本語へ」の翻訳だった。つまりターゲット言語は常に日本語であり、考えるべきはソース言語の特性とターゲット言語の特性だけでよかった。

しかも、翻訳という営みは、伝統的に属人的なもの

A perspective from the translator of Poibeau's *Machine Translation*
Akira Takahashi
Freelance Translator, Executive Vice President of JTF

と捉えられている。翻訳は個々人の職人技というわけで、翻訳ものは「誰々訳の何々」と語られることが多い（翻訳者・翻訳家の地位が決して低くないのは、こういう背景のおかげでもある）。文芸（出版）翻訳と産業翻訳は区別して考えなければいけないが、前者の文化が後者に大きく影響している面はあるだろう。

ご存じのように、機械翻訳や、それに先だって広がった翻訳支援ツールは、産業翻訳の現場でさえ black sheep 扱いだ。実はその根底には、ヨーロッパと日本に存在する上述のような「翻訳観」の違いがあるのではないか―『機械翻訳』を訳してからそう考えることが多くなった。思い切りざっくり言うと、ヨーロッパ的な翻訳観がめざすのは、必要に迫られた「量産のための翻訳」つまり「誰がやっても一定以上の品質になる翻訳」であり、かたや日本的な翻訳観の中心にあるのは、その属人性に基づく「一点物の翻訳」なのだ。

一点物というとまるで文芸翻訳の話のようだが、そうではない。産業翻訳にも一点物の翻訳はある。また、成果物の性質としては量産的であっても、翻訳者ひとりひとりは一点物という意識で翻訳に臨むことが多く、だからこそ産業翻訳の品質が一定以上に保たれていると言える。もちろん、ヨーロッパにも一点物の翻訳は存在するが、産業翻訳との線引きがはっきりしているように感じる。逆に日本では、産業翻訳が文芸翻訳から独立しきっていないのだろうか。

同業者と話をしていても、一点物の翻訳（という意識）を主とする方々と、量産的な翻訳を日常的に扱っている方とでは、機械翻訳の話が基本的にかみ合わないことが多い。もちろん、どちらがいい悪いという話ではなく、翻訳に対するスタンスの違いだ。機械翻訳や翻訳支援ツールが、量産的な翻訳を対象にしているということが、実はいまだに共通認識になっていないと感じることも少なくない。レベルの違いということではなく、一点物の翻訳と量産的な翻訳との違いは意識しておいたほうが、話が進みやすいのではないか。

●機械翻訳はどこへいくのか？

このフレーズにだけは、若干の違和感がある。「～はどこへいく」という自動詞表現ではなく、「～をどこへ向かわせるのか」という他動的な問いかけであるべきではないか。つまり、関係者一同が機械翻訳の扱い方と方向性を主体的に考え、変えていくべきではないかと思うからだ。個人的には、主に以下のように考えている。

・機械翻訳研究者と翻訳者との協力態勢

ここをもっと深めれば、双方にとって利は大きいはず。その意味では、私のような個人翻訳者が、中澤先生や山田先生たちと多少近しく接しているのは、いい傾向と言えるかもしれない。こういう形をもっと広げていきたい。

・ソリューションの活用に関するガイドライン

機械翻訳が、社会一般でも悪者になってしまう事例が後を絶たない。言うまでもなく、官公庁などで機械翻訳が誤って使われるケースだ。販売者もソリューション導入の時点でもっと指導すべきだが、事例が官公庁で目立つ現状を考えると、機械翻訳の使い方についてのガイドラインが早急に必要だろう。ここは JTF や AAMT の出番ではないか。

・個人翻訳者が機械翻訳を正しく評価できる場

Google 翻訳の結果を見て笑うのは、翻訳者としてはあまり建設的ではない。いつまでも「できないところ」を見ているのではなく、「できる」ところを研究・評価しておくべきだ。機械翻訳とどう付き合うかという個々人の判断は、その先にある。

AAMT 長尾賞と学生奨励賞の紹介

アジア太平洋機械翻訳協会（AAMT）

1. はじめに

アジア太平洋機械翻訳協会(AAMT)は、機械翻訳の学術的および産業的発展を促進するために 2006 年に AAMT 長尾賞を設立し、さらに 2014 年には機械翻訳研究に対する学生諸君の奮起を促すために AAMT 長尾賞学生奨励賞を設立しました。機械翻訳に対する優れた学術的および産業的貢献のあった個人およびグループを毎年継続して表彰しております。

2. AAMT 長尾賞について

AAMT 長尾賞は、AAMT 初代会長である長尾真先生の日本国際賞賞金のご寄付を受けて、機械翻訳システムの実用化の促進および開発に貢献した個人あるいはグループを表彰するために 2006 年に設立されました。いわゆる学会の論文賞や学術賞ではなく、高性能の機械翻訳システムの商品化、機械翻訳システムを使った新しいサービスの開始、あるいは機械翻訳システム利用の新たな普及基盤の構築などの貢献を対象としています。もちろん学術的に意味のある成果も対象としています。詳細についてはウェブページ（https://aamt.info/news/nagao-2/）をご覧ください。

歴代の AAMT 長尾賞受賞者を表 1 に掲載しております。機械翻訳の学術的貢献だけでなく、産業的貢献が AAAMT 長尾賞の大きな評価対象となっていることがわかるかと思います。近年の受賞では、ソースネクスト株式会社様が第 14 回 AAMT 長尾賞(2019 年度)を受賞されました。多くの方が御存知の通り、携帯型音声翻訳端末であるポケトークを開発し、機械翻訳技術を一般に広く普及、認知させたことで有名です。第 15 回(2020 年度)では、医療用音声翻訳アプリ「MELON」を開発したコニカミノルタ株式会社 BIC

表 1 AAMT 長尾賞受賞者

第 1 回 (2006 年)	• 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 • 日本電気株式会社 ＮＥＣメディア情報研究所
第 2 回 (2007 年)	• 株式会社国際電気通信基礎技術研究所 音声言語コミュニケーション研究所
第 3 回 (2008 年)	• シャープ株式会社 情報通信事業本部 要素技術開発センター 機械翻訳チーム • 日本電気株式会社 および 株式会社高電社
第 4 回 (2009 年)	• 東芝ソリューション株式会社 および 株式会社東芝 • AAMT インターネットワーキンググループ
第 5 回 (2010 年)	• 「みんなの翻訳」構築・運用グループ
第 6 回 (2011 年)	• AAMT 共有化・標準化ワーキンググループ
第 7 回 (2012 年)	• 株式会社富士通研究所 機械翻訳システム研究開発グループ
第 8 回 (2013 年)	• 株式会社バオバブ代表取締役社長 相良美織
第 9 回 (2014 年)	• （独）情報通信研究機構 ユニバーサルコミュニケーション研究所 多言語翻訳研究室 内山将夫、隅田英一郎
第 10 回 (2015 年)	• 株式会社 ATR-Trek 業務用途音声翻訳システム開発プロジェクト パウル・ミヒャエル、鈴木昌広、袋谷丈夫
第 11 回 (2016 年)	• 株式会社みらい翻訳
第 12 回 (2017 年)	• 八楽(やらく)株式会社
第 13 回 (2018 年)	• 凸版印刷株式会社情報コミュニケーション事業本部 ソーシャルイノベーションセンター情報インフラ本部コンテンツ企画部 • 日中・中日機械翻訳実用化プロジェクト（京都大学・黒橋禎夫、科学技術振興機構・中澤敏明）
第 14 回 (2019 年)	• ソースネクスト株式会社
第 15 回 (2020 年)	• コニカミノルタ株式会社 BIC Japan • 東芝デジタルソリューションズ株式会社
第 16 回 (2021 年)	• Team Tohoku-AIP-NTT at WMT-2020（メンバー：清野 舜, 伊藤 拓海, 今野 颯人, 森下 睦, 鈴木 潤） • Mantra 株式会社

Introduction of AAMT Nagao Award and Student Award
Asia-Pacific Association for Machine Translation (AAMT)

図 1 第 16 回 AAMT 長尾賞: Team Tohoku-AIP-NTT at WMT-2020（左から、伊藤拓海さん、鈴木潤さん、清野舜さん、森下睦さん、今野颯人さん）

図 2 第 16 回 AAMT 長尾賞: Mantra 株式会社（左から、山中武さん、石渡祥之佑さん、日並遼太さん、関野遼平さん）

Japan 様と、特許庁の審査業務のための日英機械翻訳システムが評価された東芝デジタルソリューションズ株式会社様が受賞されております。いずれも産業界に大きなインパクトを与えた機械翻訳に関する貢献が評価されての受賞となっております。

今年度(2021 年度)の第 16 回 AAMT 長尾賞は、Team Tohoku-AIP-NTT at WMT-2020（メンバー：清野 舜, 伊藤 拓海, 今野 颯人, 森下 睦, 鈴木 潤）様と Mantra 株式会社様が受賞しました(図 1 と図 2)。Team Tohoku-AIP-NTT at WMT-2020 様は、機械翻訳の分野で権威の高い国際会議 WMT2020 において、大学、研究機関、企業の産学連携により非常に優秀な成績(多くのタスクで同率を含む 1 位の成績)を収め、機械翻訳システムの実用化に向けた日本の研究開発力の国際的プレゼンス向上に大きく貢献したことが評価されました。Mantra 株式会社様は、学術的にもまだ萌芽的な段階にあるマルチモーダル・文脈翻訳を漫画翻訳という形で事業化をしており、今までに無い新しい機械翻訳の形に挑戦する姿勢が大きく評価されました。

AAMT 総会に伴って授賞式が行われるはずでしたが、コロナ禍のため今回は 2021 年 6 月 23 日にネット配信による発表と授与品の郵送をもって授賞式に代えさせていただきました。

3. AAMT 長尾賞学生奨励賞について

AAMT 長尾賞学生奨励賞は、同じく長尾真先生から

表 2 AAMT 長尾賞学生奨励賞受賞者

第 1 回 (2014 年)	東京大学大学院教育学研究科 宮田玲さん
第 2 回 (2015 年)	京都大学大学院情報学研究科 後藤功雄さん
第 3 回 (2016 年)	奈良先端科学技術大学院大学 三浦明波さん
第 4 回 (2017 年)	京都大学大学院情報学研究科 John Richardson さん
第 5 回 (2018 年)	奈良先端科学技術大学院大学 小田悠介さん
第 6 回 (2019 年)	東京大学大学院工学系研究科 江里口瑛子さん
第 7 回 (2020 年)	筑波大学大学院システム情報工学研究科 龍梓さん 東京大学大学院情報理工学系研究科 Raphael Shu（朱中元）さん
第 8 回 (2021 年)	東京大学大学院情報理工学系研究科 石渡祥之佑さん

ご寄付頂いた日本国際賞賞金を原資として、機械翻訳の発展に長年取り組んでこられた長尾真先生の機械翻訳への熱い思いを次世代を担う学生諸君に伝え、将来の機械翻訳の発展に資する研究を見出し、学生諸君の奮起を促すことを趣旨として、2014 年度に新設されました。詳細についてはウェブページ（https://aamt.info/news/nagao-student/）をご覧ください。

歴代の AAMT 長尾賞学生奨励賞受賞者を表 2 に掲載しております。長尾賞と趣旨が異なり、学生による学術的貢献が大きな評価対象となっています。近年の

図 3 第 6 回 AAMT 長尾賞学生奨励賞
（左から、長尾真先生、江里口瑛子さん）

図 4 第 8 回 AAMT 長尾賞学生奨励賞（石渡祥之佑さん）

受賞では、東京大学大学院工学系研究科に当時在籍していた江里口瑛子さんが第 6 回 AAMT 長尾賞学生奨励賞(2019 年度)を受賞されました(図 3)。江里口さんは、当時ニューラル機械翻訳に構文情報を導入する先駆的な研究をされており、その後の追随するニューラル機械翻訳の研究に大きな影響を与えたことが評価されました。第 7 回(2020 年度)では、筑波大学の龍梓さんと東京大学の Raphael Shu（朱中元）さんが受賞されております。龍さんは統計的機械翻訳(フレーズテーブル)の技術をニューラル機械翻訳に応用することで未知語の問題を改善したことが評価されました。Shu さんは、連続的な潜在変数を用いた変分オートエンコーダに基づく非自己回帰的機械翻訳の先駆的な研究、離散的な量子化に深層学習を用いた単語分散表現の圧縮、構文・意味的な特徴の離散的な符号化に基づく多様な訳文生成の研究を行っており、ニューラル機械翻訳の高速化、軽量化、文体制御に関する独創的かつ有用性の高い手法の提案が大きく評価されました。

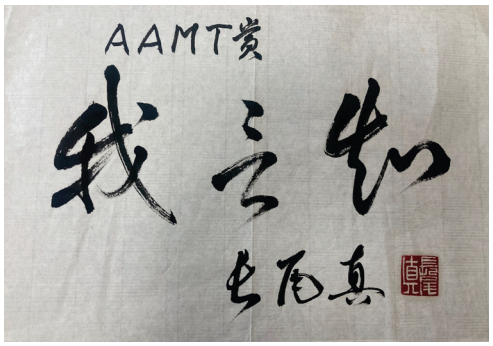
今年度(2021 年度)の第 8 回 AAMT 長尾賞学生奨励賞は、東京大学大学院情報理工学研究科に当時在籍していた石渡祥之佑さんが受賞しました(図 4)。石渡さんは Mantra 株式会社の代表を務めており、AAMT 長尾賞とダブルの受賞になります。受賞対象となった論文は、東京大学喜連川優教授の指導のもとに 2018 年度にまとめられた東京大学の博士論文「Translation and Description Methods for Multilingual Text Understanding」です。対象の論文では、機械翻訳のドメイン適応における未知語の扱いに関する研究、ニューラル機械翻訳にチャンク情報を取り込む手法の研究、局所的・大局的な文脈に応じて語句の語義を生成する手法を提案しています。未知語処理については、ベクトル空間の写像を用いて未知語処理を行っている点や、写像の学習に様々な工夫が用いられている点が大きく評価されました。チャンクベースのニューラル機械翻訳は単純にモデルを階層化するだけでなくモデルに様々な工夫を施すことで独自性と有用性のあるモデルを実現しています。用語説明生成は機械翻訳のタスクそのものではありませんが、大局的な文脈を用いている点に独創性があり、機械翻訳への適用可能性が十分に考えられ、この研究も大きく評価されました。

今年度受賞者による講演が 2021 年 12 月 8 日の AAMT 年次大会(オンライン)で行われる予定です。

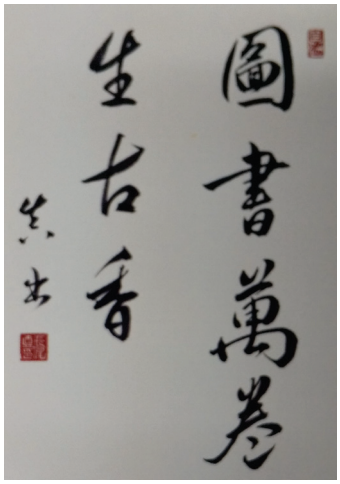
4. 長尾真先生を偲んで

AAMT 長尾賞および AAMT 長尾賞学生奨励賞を創設され、機械翻訳業界に数え切れない貢献をされ支えてくださった長尾真先生が今年 2021 年 5 月 23 日に逝去されました。長尾先生は書を嗜んでおられたことでも有名で、毎年 AAMT 長尾賞のために書を贈呈していただき、印刷されたものが授与品の一つとなってい

第 1 回（2006 年）「言不尽意」



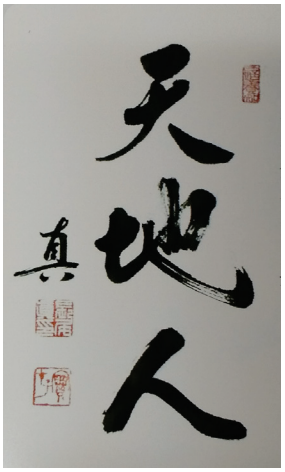
第 2 回（2007 年）「我不知」



第 3 回（2008 年）「図書萬巻生古香」

第 4 回（2009 年）「春風」

第 5 回（2010 年）「言心声」

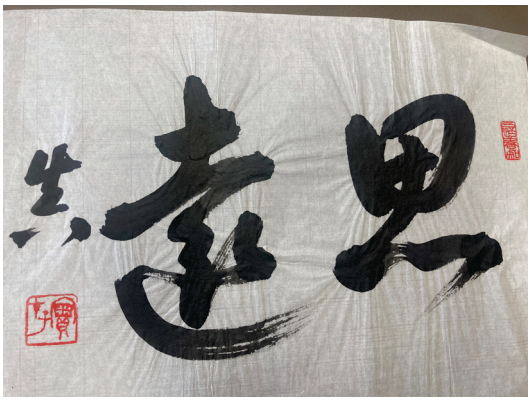


第 6 回（2011 年）「天地人」

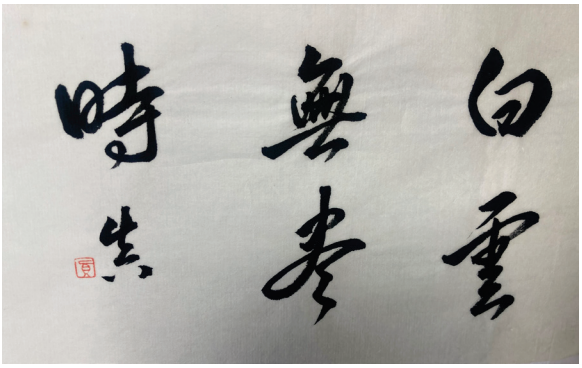
第 7 回（2012 年）「尚志」

図 5 長尾先生の書（第 1 回～第 7 回）

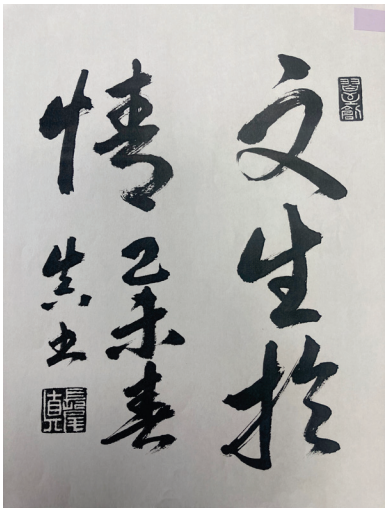
ました(図 5 と図 6)。授与式において長尾先生の書を拝見することを毎年楽しみにしている方は多かったのではないかと思います。今後長尾先生から書をいただくことは出来ませんが、長尾先生の遺志を引き継ぎ、機械翻訳業界の発展のために AAMT 長尾賞を継続していきたいと考えています。



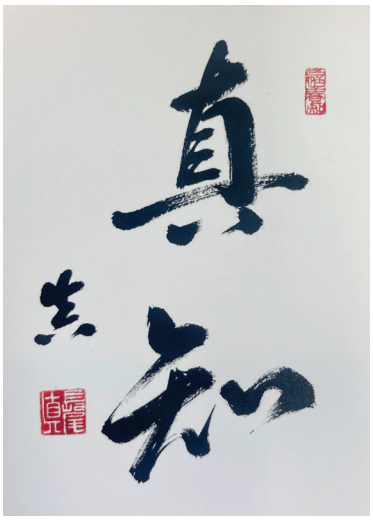
第 8 回（2013 年）「思遠」



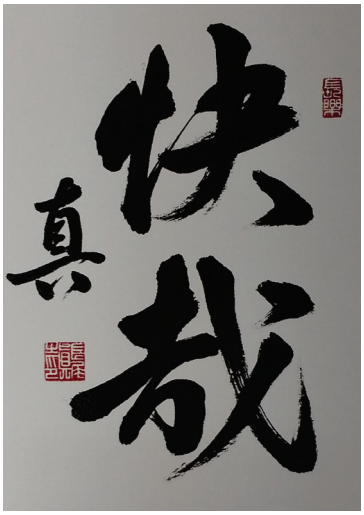
第 9 回（2014 年）「白雲時尽無」



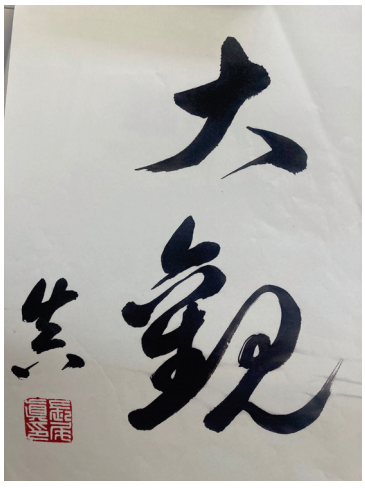
第 10 回（2015 年）「文生於情」



第 11 回（2016 年）「真知」

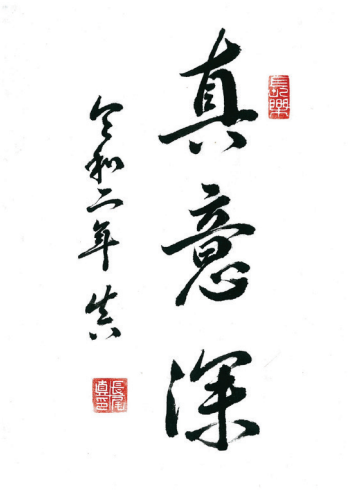


第 12 回（2017 年）「快哉」



第 13 回（2018 年）「大歓」

第 14 回（2019 年）「信念」



第 15 回（2020 年）「真意深」

図 6 長尾先生の書（第 8 回～第 15 回）

第 16 回 AAMT 長尾賞受賞記念

WMT-2020 ニュース翻訳タスクに参加して

清野 舜 [1,2] 伊藤 拓海 [2,1] 今野 颯人 [2,1] 森下 睦 [3,2] 鈴木 潤 [2,1]（代表執筆）
[1]理研 AIP [2]東北大学 [3]NTT コミュニケーション科学基礎研究所

1. はじめに

このたびは，WMT-2020 ニュース翻訳タスクに参加した「Team Tohoku-AIP-NTT（チームメンバーは本稿の著者全員）」が構築した機械翻訳システム，および，ニュース翻訳タスクにおける成績を高く評価していただき，第 16 回 AAMT 長尾賞に選出していただきましたこと，選考委員の方々をはじめ関係者の皆様に，この場をお借りして御礼申し上げます。

著者らの解釈では，オンラインの機械翻訳サービス開発を一般向け自動車開発とみなすと，WMT のコンペティション用の機械翻訳システム開発は，自動車競技 F1 のような事前に決められたレギュレーション下でどれだけ良いものを作れるか，という趣旨のシステム開発に相当すると考えています。その意味では，AAMT 長尾賞の一番の趣旨である「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の表彰」からはやや外れる部分もあると認識しておりますが，機械翻訳開発に関連する大きな成果と認定していただけたことを，著者一同非常に嬉しく思っています。

本稿では，機械翻訳における WMT の役割や，今回の成果の位置付けなどを説明させていただくとともに，WMT-2020 参加システムの要点や，良好な成績を得られた理由を簡単に紹介したいと思います。[1]

2. WMT の概要

WMT とは，元々は Workshop on Statistical Machine Translation の略称であり，2006 年に統計翻訳（Statistical Machine Translation）に特化したワークショップとして出発し，それ以来毎年開催されている。略称の WMT は，2006 年の第一回目から現在まで変わっていないが，第 11 回目の 2016 年よりワークショップから国際会議に格上げされ，WMT は Conference on Machine Translation の略称となった。つまり，Workshop から Conference に変更になったタイミングで「Statistical」が消去され，「統計翻訳に特化したワークショップ」から「機械翻訳全般の国際会議」へと変更された。これは，丁度，統計翻訳からニューラル翻訳へと時代が転換する時に合わせての名称変更となった [2]。 現在，機械翻訳を専門とする国際会議として広く知られている。

WMT の大きな特徴の一つとして，毎年機械翻訳に関連するコンペティション（通称 Shared task）が開催されることが挙げられる。コンペティションでは，通常，複数のタスクと言語毎に分けられたトラックが存在する。参加チームごとに参加タスクおよびトラックを選択し，レギュレーションに従ってシステムを構築する。例えば，翻訳タスクであれば，最終的に人手により翻訳品質の評価がおこなわれ，システムの順位付が決定する。

現在，WMT にて様々な翻訳関連タスクが実施されているが，年々タスクが増減する中で，ニュース翻訳タスク（News Translation Task）は，2006 年の第一回のワークショップからほぼ同じルールで毎年欠かさ

[1] 以降の節では基本的に常体により記述する．

[2] 2016 年と言えば，ちょうど Sennrich らの subword [7]や逆翻訳（back-translation）[8] が登場してニューラル翻訳が更に効果的であることが示された年である．WMT においてもニューラル翻訳が主流になる契機となった．

Participating in the WMT-2020 News Translation Task
Shun Kiyono, Takumi Ito, Ryuto Konno, Makoto Morishita, Jun Suzuki (Lead author of this article)
RIKEN, Tohoku University, and NTT

ず続けられている WMT コンペティションを代表する伝統的なタスクである。これまでに多くの参加チームが機械翻訳システムの翻訳品質を競い合い，その中で多くの翻訳技術が培われてきた。機械翻訳の研究開発の発展においては大きな役割と存在感をもって世界中の翻訳研究者に認知されている。実際, ACL, EMNLP といった自然言語処理の最難関国際会議はもちろんのこと，NeurIPS, ICML, ICLR, AAAI, IJCAI といった機械学習や人工知能分野の最難関国際会議に採録される機械翻訳関係の論文のほとんどが WMT の配布データを使って実験をおこなっている。コンペティション参加システムやその後続の研究成果と性能を比較することで，それぞれの提案法の効果を示してきた。

このように世界中の機械翻訳の研究者／開発者が直接的または間接的に関わる WMT ニュース翻訳タスクは，その性質上，良好な成績を得ることが一種のステータスとなり得る。そのこともあって，毎年多くのチームがコンペティションに参加している。

3. WMT-2020 参加トラックの結果

Team Tohoku-AIP-NTT は，2020 年 11 月に開催された WMT-2020 のニュース翻訳タスクにおいて，英語からドイツ語（En-De），ドイツ語から英語（De-En），日本語から英語（Ja-En），英語から日本語（En-Ja）の 4 言語対の翻訳トラックに参加し，その全てで翻訳品質の人手評価において一位（同率含む）を獲得した[1, 2]。また，ニュース翻訳において競争が激しい De-En および En-De の 2 トラックにおいては，自動評価指標（BLEU スコア）でも一位を獲得し[3]，後の言語学の専門家（professional linguists）による評価でも，他のシステムよりも本システムの結果が総じて好ましいという結果も得られた[4]。更に，言語学者による言語現象別の翻訳品質評価（複単語表現，固有名詞，機能語。動詞の時制など）においても，他のシステムを上回る最良の評価となった[5]。

このように，複数のトラックで弊チームが事実上の一位 [3]を獲得したことは，機械翻訳分野における日本のプレゼンス向上にも貢献したと考えている。

4. なぜ良好な結果が得られたのか

参加システムは，昨今広く用いられている，いわゆる Transformer をベースとしている。その意味においては独自の技術を使ったわけではなく，他のチームもほぼ同じ方法論やモデルを使っている。では，なぜ弊チームのシステムが他のチームより相対的によい結果が得られたのか？については，いくつかの要因が挙げられるが，総じて一言で言うと「頑張った」ということになるかもしれない。

最先端の技術をただ「知っている」ことと「実際にやる」ことには，大きな隔たりがあり，実際にやってみて初めて知る難しさや，そこから得られる経験的な知見がある。その観点で言うと，チームメンバー全員がそれまでの研究活動において大規模データに対する処理への耐性や，Transformer を効果的に使う知見を事前に得ていたことが大きな要因と考えている。そのため，最初から概ね全員が同じような感覚で作業を進めることができた。また，過去の実験で培った効果的な実験計画の方法論が確立しており，効率的に試行錯誤やハイパーパラメタの探索が行えたことも要因だと言える。

5. 参加システムの要点

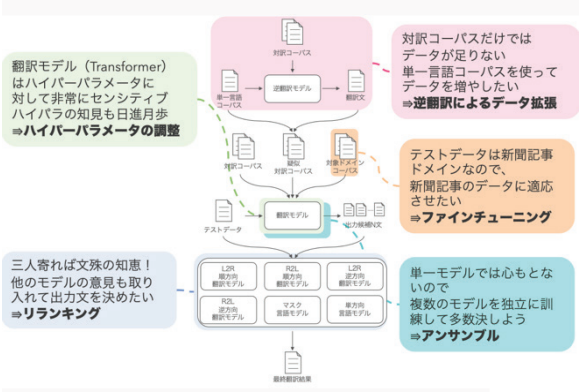
2020 年 2 月 29 日にデータが公開され，そこからシステム作成に入った。データの正規化やトークン区切りの話題から始まり，モデル構成を考えつつベースモデル（逆翻訳用）の学習をおこない，並行してハイパーパラメタ調整をする，といった感じにシステム開発を続けていき，最終的に 7 月 1 日に評価データの最終投稿をした。実質 4 ヶ月間の試行錯誤の結果になる。詳

[3] WMT では公式にシステムの順位を公表しているわけではないので，前述の評価結果の事実から「事実上の一位」という言い回しとしている．

細なシステム構成などは，弊チームの WMT におけるシステム論文を参考されたい[1]。また，AAMT/Japio 特許翻訳研究会が主催する「特許情報シンポジウム」にて本システムに関する講演を既に行っており，詳細なシステム構成に関する発表資料を一般公開している[6]。

現在の機械翻訳の主流は，いわゆる End-to-end 型のニューラルネットワーク（以下ニューラル翻訳モデルと表記）である。ニューラル言語モデルに相当するニューラル翻訳モデルのデコーダによる高い文章再現能力により，かなり流暢で自然さの高い文章が生成できる。これは，従来の統計翻訳と比べてニューラル翻訳モデルを用いる大きなメリットとして真っ先に挙げられる点である。もう一つ，従来の統計翻訳と比べて，一つのニューラルネットワークにより翻訳が実現できる点も，処理工程を簡略化する観点で大きなメリットと考えられている。

しかし，コンペティションに参加する場合は，その様相は一変する。ニューラル翻訳モデルを用いる場合でも，純粋に翻訳品質を向上するには，単一のニューラル翻訳モデルよりも，モデルを複数用意し，それらを組み合わせることや，補助モジュールを組み合わせることが非常に効果的で翻訳品質を飛躍的に高めることが経験的に知られている。つまり，コンペティションでよい成績を獲得するには，様々なモデルの組み合わせの知見を一つずつきっちり導入し，その少しずつの積み上げの結果得られる大幅な品質向上を実現する作業が必要になる。以下に，参加システムの構成図を示す。



この図が示すように，コンペティション参加システムは，データ拡張に必要な逆翻訳モデル，モデルアンサンブル用モデル，リランキング用の各種モデルを用いている。最終的に，一つのトラック（例えば，En-De）に対して，逆翻訳用の翻訳モデル，順方向（left-to-right）デコーダのニューラル翻訳モデルアンサンブル用に 4 モデル，逆方向（right-to-left）デコーダ（つまり文末から先頭に向けて予測するもデコーダ）をアンサンブル用に 4 モデル，および，リランキング用のマスク型言語モデルと単方向言語モデルの合計で 11 個のモデルを構築し，それらを組み合わせてシステムを構築した。参加した 4 トラック合計では，42 個のモデルを構築したことになる（英語の言語モデルは 2 つのトラックで共有できるため 2 モデル分少なくなる）。

6. 試行錯誤

実際に最終的な参加システムに導入されなかったが，システム開発の過程で試したいくつもの方法論がある。例えば，情報量指標に基づいたデータ選択を用いて，逆翻訳によって生成したデータも含めて，学習に効果の高そうなデータのみを選別し，それ以外の効果が薄い，或いは，悪影響のあると考えられるデータを学習データから除外するといった試みもおこなった。しかし，いくつかの取り組みは，検証段階で効果がでないことが確認され，やむなく参加システムからは外した。

実際には，こういったネガティブな結果に対する知見も実際にやってみることで初めて得られるものである。特に今回は，データが小規模である場合に効果的だと思われていた方法論も，扱うデータが大規模になるとほとんどその効果が見られなくなるといった現象が多く散見された。そのため，効果を期待して導入したいくつかの方法論（その中にはオリジナルの方法論も含む）が尽く利用できず，複雑なことを狙うよりも，データおよびモデルの大規模化と言った堅実かつ最も単純な方法論が最も効果的だった。ある意味，とても味気ない結果となった。

7. 今後の展望

コンペティションにはレギュレーションが存在し, 参加者が基本的に一定の条件下で競い合う。よって, 現在の自分たちの技術レベルを確認することができる。また, 良い成績を得れば, それが世界最先端の技術を持っていることの証明にもなる。そして, 全く無名の研究者や組織であったとしても, コンペティションで圧倒的によい結果を得れば, その分野で一目おかれる存在に一足跳びになることができる。この観点からも, 若手の研究者や開発者, または, 規模の小さい企業などは, 名を売るために挑戦するのも悪くはない選択と考えられる。しかし, 最も重要なのは, システム構築の過程で得られる経験や知見であり, 単に「知識として知っている」から実際に取り組んで「経験として知っている」へと昇華する良い場である点だと考えている。この経験が, その後の研究や開発の中で大きな武器となる。今後も組織横断的にモチベーションの高いチームを構築し, 定期的に WMT に参加していきたいと考えている。

参考文献

[1] Shun Kiyono, Takumi Ito, Ryuto Konno, Makoto Morishita, Jun Suzuki. “Tohoku-AIP-NTT at WMT 2020 News Translation Task”, In Proceedings of WMT-2020, pp. 145–155, 2020.

[2] Loïc Barrault, Magdalena Biesialska, Ondřej Bojar, Marta R. Costa-jussà, Christian Federmann, Yvette Graham, Roman Grundkiewicz, Barry Haddow, Matthias Huck, Eric Joanis, Tom Kocmi, Philipp Koehn, Chi-kiu Lo, Nikola Ljubešić, Christof Monz, Makoto Morishita, Masaaki Nagata, Toshiaki Nakazawa, Santanu Pal, Matt Post, Marcos Zampieri, “Findings of the 2020 Conference on Machine Translation (WMT20)”, In Proceedings of WMT-2020, pp. 1–55, 2020.

[3] 英語-ドイツ語, ドイツ語-英語間の翻訳品質の自動評価（BLEU スコア）で一位を獲得した根拠
http://wmt.ufal.cz/newstest2020-de-en/
http://wmt.ufal.cz/newstest2020-en-de/

[4] Nitika Mathur, Johnny Wei, Markus Freitag, Qingsong Ma, Ondřej Bojar. “Results of the WMT20 Metrics Shared Task”, In Proceedings of WMT-2020, pp. 688–725, 2020.

[5] Eleftherios Avramidis, Vivien Macketanz, Ursula Strohriegel, Aljoscha Burchardt, Sebastian Möller. "Fine-grained linguistic evaluation for state-of-the-art Machine Translation", In Proceedings of the WMT-2020, pp. 346–356, 2020.

[6] 第 6 回特許情報シンポジウム プログラム：
http://aamtjapio.com/symposium.html
スライド：
https://speakerdeck.com/butsugiri/ji-jie-fan-yi-konpeteisiyoncan-jia-bao-gao

[7] Rico Sennrich, Barry Haddow, Alexandra Birch. “Neural Machine Translation of Rare Words with Subword Units”, In Proceedings of ACL-2016, pp. 1715–1725, 2016.

[8] Rico Sennrich, Barry Haddow, Alexandra Birch. “Improving Neural Machine Translation Models with Monolingual Data”, In Proceedings of ACL-2016, pp. 86–96, 2016.

第 16 回 AAMT 長尾賞受賞記念

マンガ機械翻訳への招待

石渡 祥之佑
Mantra 株式会社

1. はじめに

このたびは AAMT 長尾賞という大変名誉ある賞を授与いただき誠にありがとうございます．Mantra 株式会社（以下，Mantra）は，マンガをはじめとする娯楽作品の多言語展開を加速することを目的とし 2020 年に設立されたスタートアップです．現在，Mantra では独自のマンガ機械翻訳エンジンを搭載した CAT ツール「Mantra Engine」や，多言語化されたマンガを活用した外国語学習サービス「Langaku」の開発を行っています．本稿では，マンガ翻訳の現状やその自動化に向けた技術的挑戦，当該領域における Mantra の取り組みを紹介します．

2. マンガ海外展開における課題

日本マンガの市場は国内外合わせて約 6,126 億円 [1] と大規模ですが，翻訳され日本国外で出版される作品点数は全体のわずか 10%程度と言われています．この理由の一つとして，マンガ翻訳に高い時間的・金銭的コストがかかる点が挙げられます．マンガ翻訳は作品の文化的背景に対する深い理解や高度な創造性を要するだけでなく，「吹き出し」領域の塗りつぶしや翻訳済テキストの配置といった煩雑な画像処理も含むクリエイティブなタスクであり，その遂行にはどうしても複数人の専門家（たとえば翻訳者，チェッカー，デザイナ）間の連携が必要となります．結果，翻訳コストは高くなり，限られた出版社の一部の雑誌，さらにその中でもごく一部の作品が，いくつかの言語に翻訳されるに留まっています．

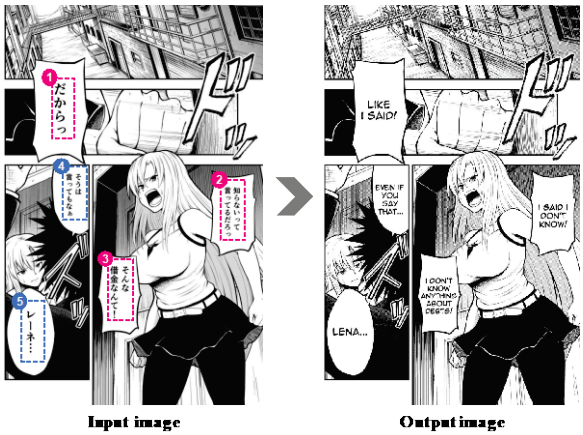


図 1　マンガ機械翻訳の入出力例．©Mitsuki Kuchitaka

限られたマンガしか正規に翻訳・出版されない状況の中で，正規に翻訳されていない作品を世界中のマンガファンが独自にスキャン・翻訳し，無償でオンライン頒布する「海賊版」の存在が問題視されています．こうしたファンは「自分が好きな作品をもっと多くの人に読んでほしい」という純粋な動機で翻訳・配信に取り組むケースがほとんどではあるものの，海賊版の存在によって正規に翻訳されるマンガの販売機会が毀損されるという問題点があります．たとえば，米国における日本マンガの海賊版被害額は年間 1.3 兆円に達する [2]といわれています．正規版がないから海賊版が作られ，海賊版があるために正規版が売れにくく，結果国内外の出版社は大規模なマンガの翻訳出版に乗り出しにくい，という悪循環に陥ってしまっています．

Mantra はこの問題を解決するためにマンガ翻訳の生産性向上を目指し，マンガに特化した機械翻訳エンジンの研究開発や，マンガ翻訳ツール「Mantra Engine」の提供を行っています．以降，3 節以降ではこれらの取り組みについて述べます．

Introduction to Machine Translation for Manga
Shonosuke Ishiwatari
Mantra Inc.

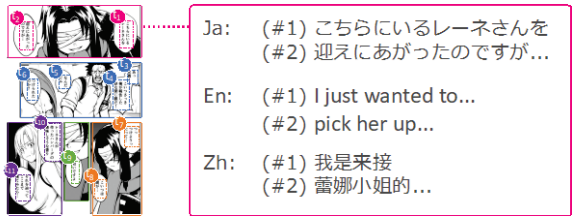


図 2　マンガ翻訳の評価用データセット「Open Mantra」

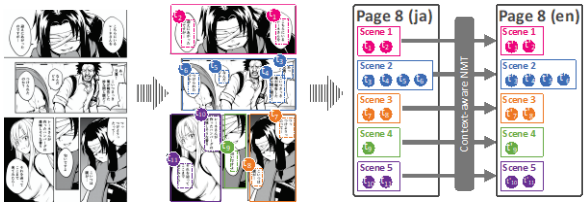


図 3　シーン内の文脈を考慮する機械翻訳

3.　マンガ翻訳自動化における技術的挑戦

近年，産業翻訳の領域において機械翻訳が翻訳作業の効率化に大きく寄与した一方で，マンガ翻訳への機械翻訳導入は未だ進んでいません．私たちは，これがマンガ翻訳の難しさに起因すると考えています．たとえばマンガでは（1）「コマ」や「吹き出し」の配置にしたがってテキストが非規則的に配置されるため文脈情報を抽出しにくい，（2）会話文のため省略される語が多い，（3）キャラクタや場面によって独特な言い回しが発生するといった性質から，高精度な機械翻訳の実現が困難です．たとえば図 1 に示すように，単一の文が複数の吹き出しに分割されている場合が多くあります．そのため，吹き出しごとに独立して翻訳するのではなく，図 1 ①〜③のように適宜語順を入れ替えながら日英翻訳を行う必要があります．また，吹き出し内のテキストは全て会話文であることから，原言語である日本語では，多くの場合主語や目的語が省略されます．さらに，倒置表現や語尾の「っ」など，マンガ特有の表現が多用されます．マンガのこうした性質は，どれも機械翻訳システムの精度を著しく低下させるものです．こうした問題を解決するために，私たちはマンガに特化した対訳コーパス構築と翻訳アルゴリズムの開発を同時に進めていく必要があると考えています．

4.　マンガ機械翻訳の実現に向けて

前節で述べたマンガ翻訳自動化の難しさをふまえ，本節では Mantra がこれまでに取り組んできた高精度なマンガ機械翻訳エンジンを実現するための取り組みを紹介します．

マンガ翻訳評価用データセット「Open Mantra」

翻訳エンジンの開発や性能改善において，定量的な評価を実施するためのデータセットの存在は必要不可欠です．前節で述べたように，多くのマンガは他ドメインのテキストと異なる性質をもつため，マンガ翻訳エンジンを評価するためにはマンガ専用の評価基盤が必要となります．著者らが本研究に着手した 2018 年当時，マンガ翻訳エンジンの性能を手軽に評価する方法は存在しませんでした．そこで，私たちは複数人の漫画家の協力のもと，5 ジャンル（恋愛，ミステリー，ファンタジー，日常，バトル）の作品を含む世界初のマンガ翻訳評価用データセット「Open Mantra」を研究用途に公開しました．図 2 に示すように，このデータセットでは日本語のマンガ画像にコマや吹き出しの座標が人手で付与され，各セリフは翻訳者によって英語と中国語に翻訳されています．このデータセットを公開することで，マンガ機械翻訳の研究開発を行う全ての研究者・技術者に翻訳エンジンの評価基盤を提供することができました．

シーンに基づくマンガ機械翻訳

3 節で述べたように，マンガでは単一の文が複数の吹き出しに分割されて表現されることが多くあります．プロの翻訳者らは吹き出し単位で独立して翻訳を行うのではなく，目的言語版のマンガとして自然になるよう，必要に応じて吹き出しのテキスト順序を入れ替えながら翻訳を行います．また，マンガの吹き出しは単純にページ内の上から下，右から左に配置されているわけではなく，コマ割り（コマの分割や配置）の制約を受けて配置されています．マンガに慣れ親しんだ読者にとっては，脳内でコマ割りを理解して吹き出しの

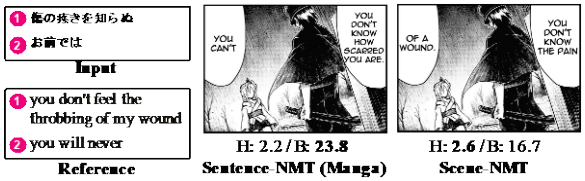


図 4　従来手法（左）とシーンに基づく翻訳（右）の比較

順序を推定し，適切な塊ごとにテキストを区切りながら読むことは無意識に行うことができる容易なタスクですが，計算機でこの処理を自動化する方法は自明ではありません．

私たちはこの問題に対し，コマを一つの「シーン」と捉え，シーンごとにテキストをまとめて翻訳する手法を提案しました．図３にこの手法の概要を示します．まず入力された画像からコマの領域を推定し，コマごとにその中に含まれるテキストを認識します．コマ内のテキストの順序を推定した後，この順序にしたがって各吹き出し内のテキストを特殊トークンで連結し，まとめて翻訳を行います．一般に文脈を考慮する機械翻訳 [3] [4] [5]手法では，文脈窓（考慮する周辺テキストの文数）が１文～２文などと固定されることが多いですが，マンガではコマによってセリフの多寡が大きく変化します．そこで文脈として用いる吹き出しの数は固定せず，コマごとに動的に変化させます．

図４に従来手法（吹き出しごとに独立に翻訳）と提案手法（シーンごとにまとめて翻訳）の例をそれぞれ示します．このように考慮すべき文脈の範囲を適切に設定することで，より自然な訳出が可能になります．

マンガ画像からの対訳コーパス自動抽出

マンガ独特の表現を正しく翻訳したり，前述した文脈範囲を動的に変化させる手法を実現したりするためには，プロの翻訳者が翻訳したマンガコーパスを用いて翻訳エンジンを訓練する必要があります．しかしながら，現状マンガに特化した対訳コーパスは存在せず，また一から人手で構築することもコスト的に現実的ではありません．そこで，Mantra は原言語・目的言語で描かれたマンガ画像から，自動的に文脈情報のついた対訳コーパスを抽出する手法を提案しました．

翻訳版のマンガは，原語版のマンガと比較して（１）表紙や目次，扉などが挿入・削除されてページ数が異なっている，（２）画像サイズやアスペクト比が調整されていることがあるため，単純なルールベース処理で対訳ページを見つけることはできません．さらに，ページ単位で対応関係が取れたとしても，縦書き・横書き等のレイアウトの違いが存在することで，文単位の対訳関係を抽出することは容易ではありません．そこで，Mantra は図 5 に示す対訳データ抽出パイプラインを設計し，この問題の解決を試みました．このパイプラインでは，入力される日本語のマンガ画像に対し，まず物体検出アルゴリズムにより吹き出し領域の矩形を検出（図 5b），これを吹き出し単位に分割（c~d）します．続いて画像検索手法を用いて当該ページと対応する目的言語のページを抽出（a）し，対応する原言語ページとピクセルレベルでアライメントを取ります（e）．最後に原言語，目的言語それぞれの文字認識エンジンによって，吹き出し領域ごとに対訳テキストを抽

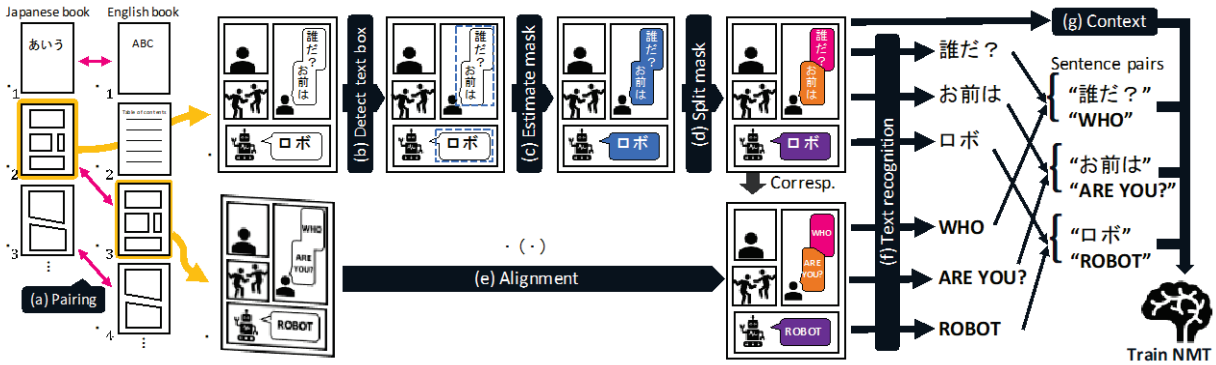


図 5　マンガ画像からの対訳コーパス自動抽出

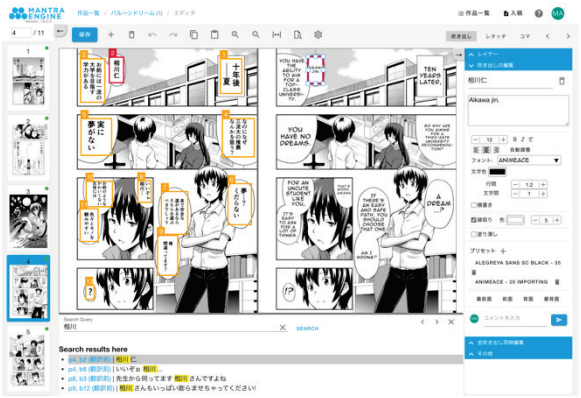


図 6　マンガ翻訳ツール「Mantra Engine」

出します．この仕組みによって，翻訳済マンガ画像の集合から全自動で対訳コーパスを構築する枠組みが完成し，大規模データで高精度なマンガ翻訳エンジンを訓練することが可能になりました．

5. マンガ翻訳ワークフローへの MT 導入

前節で述べた様々な取り組みにより，マンガ翻訳エンジンの開発・評価・改善を実施できるようになりました．本節では，このエンジンを翻訳現場で使用するためのクラウドツール「Mantra Engine」を紹介します．

図 6 に Mantra Engine の翻訳エディタ画面を示します．図中左では，入力された日本語画像から自動検出された吹き出し領域がオレンジ色の枠で表示されています．図中右には吹き出しごとに機械翻訳の結果が表示され，その内容を翻訳者やチェッカーが編集できる仕組みが用意されています．

また，Mantra Engine の利用者は訳文を編集するだけでなく，機械翻訳に反映される用語集の整備や写植作業（文字のレイアウトやフォントの切り替えなど），原文･翻訳文の全文検索，他の作業者とのコミュニケーションや編集権限の管理など，マンガ翻訳に係るあらゆる作業を当ツール内で完結することができます．

Mantra Engine は既に複数の出版社や翻訳業者，マンガ配信事業者に導入され，幾つかのマンガ作品については，実際に多言語同時配信の現場でも使われはじめています．

6. おわりに

「世界の言葉で、マンガを届ける。」これが，Mantra 設立時に著者らが掲げたスローガンです．しかしながら，この目標に向けて２人の大学院生がはじめたマンガ機械翻訳の研究は未だ萌芽的な段階にあります．より多くのマンガをより多くの人に，リアルタイムに届けられるようにするためには，まだ解決すべき技術的課題が多く残されています．技術が継続的に発展，普及していくためには，単一の企業や研究室が独自に研究開発を進めるだけではなく，多くの研究者や技術者が参加し，議論を重ねる研究コミュニティが形成されることが極めて重要です．本稿をきっかけに読者の皆さまがマンガ機械翻訳という未開の（そしてとても興味深い！）技術領域に興味を持ち，参入いただけることを私たちは強く期待しております．

また，Mantra のマンガ機械翻訳研究に惜しみない支援をくださった東京大学の吉永直樹先生，相澤清晴先生，松井勇佑先生，古田諒佑先生，東京理科大学の谷口行信先生，ならびに谷口研の学生の皆さま，誠にありがとうございました．このような名誉ある賞を授与くださった AAMT の皆さま，推薦いただいた先生方にも心より御礼申し上げます．

引用文献

[1] Roland Berger, 平成 25 年度 知的財産権ワーキング・グループ等 侵害対策強化事業（コンテンツ海賊版対策調査）最終報告書, 2014.

[2] 全国出版協会・出版科学研究所, 出版月報 2021 年 2 月号, 2021.

[3] S. Jean, S. Lauly, O. Firat , K. Cho, Does neural machine translation benefit from larger context?, arXiv, 2017.

[4] J. Tiedemann , Y. Scherrer, Neural Machine Translation with Extended Context, Proc. DiscoMT, 2017.

[5] L. Wang, Z. Tu, A. Way , Q. Liu, Exploiting Cross-Sentence Context for Neural Machine Translation, Proc. EMNLP, 2017.

第 8 回 AAMT 長尾賞学生奨励賞受賞記念

Translation and Description Methods for Multilingual Text Understanding

Shonosuke Ishiwatari

Mantra Inc.

1. Abstract

The development of Web technologies has been rapidly accelerating human communication and sharing of knowledge. The massive amount of text on the new communication platforms or knowledge sources often consists of documents in several domains and multiple languages. In order to obtain fresh and diverse information from multilingual and diverse text sources such as Twitter, Wikipedia, or arXiv, we need to cope with language barrier while also paying attention to domain differences.

Let us move on to the general topic of natural language processing. Machine translation, as one of the most promising applications of natural language processing, has been playing an essential role in overcoming the language barrier. The recent development of machine learning techniques and huge annotated corpora have considerably improved the performance of machine translation. In the face of the increasing use of multilingual platforms and knowledge sources, can machine translation help us understand the real text data in various domains? Can machine translation be applied to languages pairs whose vocabulary and grammar are significantly different (such as English vs. Japanese)? More generally, can machine directly help humans understand text written in unfamiliar domains/languages? These are the central topics of this thesis.

To answer the above questions, we propose (i) an instant domain adaptation method, (ii) an accurate translation method for English-to-Japanese translation, and (iii) an automatic description method for unknown phrases.

i) Instant domain adaptation for Statistical Machine Translation.

To translate text in various domains, the most basic method is domain adaptation. Most studies on domain adaptation require supervised in-domain resources such as parallel corpora or in-domain dictionaries. The necessity of supervised data has made such methods difficult to adapt to practical machine translation systems. In this thesis, we thus propose a method that adapts translation models without in-domain parallel corpora. Our method improves out-of-domain translation from Japanese to English by 0.5-1.5 BLEU score.

ii) Accurate translation method for English-to-Japanese translation.

English-to-Japanese translation is more difficult than other language pairs such as English-to-German or English-to-French translations. This is mainly because (1) Japanese sentence has much more words in a sentence compared to English, and (2) Japanese is a free-word-order language. To cope with these problems, we propose a chunk-based decoder for neural machine translation. Our method improves English-to- Japanese translation by 0.93 BLEU score and achieves state-of-the-art performance on the WAT '16 translation task.

iii) Automatic description generation for unknown phrases.

Even if a text is translated perfectly, or written in our familiar languages, it is still common for humans to become stuck on unfamiliar words or phrases. To help humans understand unknown phrases which are not included in hand-crafted dictionaries, we undertake a task of describing a given phrase in natural language based on its contexts. In contrast to the existing methods, our model appropriately takes important clues from contexts and achieves state-of-the-art performance in four description generation datasets.

The three methods are designed to help humans understand real multilingual text where (1) the target domain is unknown, (2) the source language may extremely differ from the users' languages, and (3) the users may be unfamiliar with the words/phrases in the text. Due to limitations of the paper length, in the rest of this article, we focus on describing the third part of the above contributions.

多言語テキスト理解のための翻訳および語義説明手法
石渡 祥之佑
Mantra 株式会社

2. Introduction

When we read news text with emerging entities, text in unfamiliar domains, or text in foreign languages, we often encounter expressions (words or phrases) whose senses we are unsure of. In such cases, we may first try to figure out the meanings of those expressions by reading the surrounding words (*local* context) carefully. Failing to do so, we may consult dictionaries, and in the case of polysemous words, choose an appropriate meaning based on the context. Learning novel word senses via dictionary definitions is known to be more effective than contextual guessing (Fraser, 1998; Chen, 2012). However, very often, hand-crafted dictionaries do not contain definitions of expressions that are rarely used or newly created. Ultimately, we may need to read through the entire document or even search the web to find other occurances of the expression (*global* context) so that we can guess its meaning.

Can machines help us do this work? Ni and Wang (2017) have proposed a task of generating a definition for a phrase given its local context. However, they follow the strict assumption that the target phrase is newly emerged and there is only a single local context available for the phrase, which makes the task of generating an accurate and coherent definition difficult (perhaps as difficult as a human comprehending the phrase itself). On the other hand, Noraset et al. (2017) attempted to generate a definition of a word from an embedding induced from massive text (which can be seen as global context). This is followed by Gadetsky et al. (2018) that refers to a local context to disambiguate polysemous words by choosing relevant dimensions of their word embeddings. Although these research efforts revealed that both local and global contexts are useful in generating definitions, none of these studies exploited both contexts directly to describe unknown phrases.

In this study, we tackle a task of describing (defining) a phrase when given its local and global contexts. We present LOG-CaD, a neural description generator (Figure 1) to directly solve this task. Given an unknown phrase without sense definitions, our model obtains a phrase embedding as its global context by composing word embeddings while also encoding the local context. The model therefore combines both pieces of information to generate a natural language description.

Considering various applications where we need definitions of expressions, we evaluated our method with four datasets including WordNet (Noraset et al., 2017) for general words, the Oxford dictionary (Gadetsky et al., 2018) for polysemous words, Urban Dictionary (Ni and Wang, 2017) for rare idioms or slangs, and a newly-created Wikipedia dataset for entities.

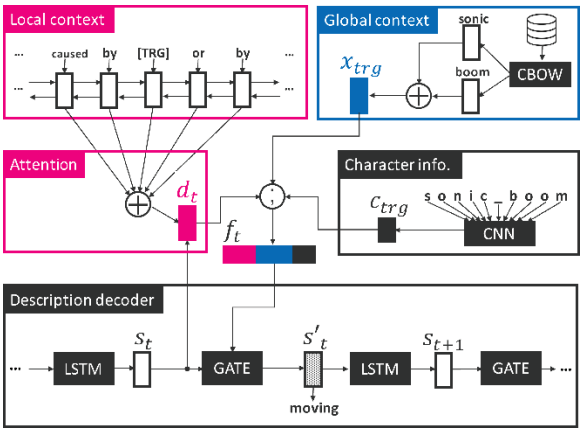


Figure 1: Local & Global Context-aware Description generator (LOG-CaD).

The contributions of this work are as follows:

- We setup a general task of defining phrases given their contexts. This task is a generalization of three related tasks (Noraset et al., 2017; Ni and Wang, 2017; Gadetsky et al., 2018) and involves various situations where we need definitions of unknown phrases (Section 3).
- We propose a method for generating natural language descriptions for unknown phrases with local and global contexts (Section 4).
- As a benchmark to evaluate the ability of the models to describe entities, we build a largescale dataset from Wikipedia and Wikidata for the proposed task. We release our dataset and the code to promote the reproducibility of the experiments (Section 5).
- The proposed method achieves the state-of-the-art performance on our new dataset and the three existing datasets used in the related studies (Noraset et al., 2017; Ni and Wang, 2017; Gadetsky et al., 2018) (Section 6).

3. Context-aware Phrase Description Generation

In this section, we define our task of describing a phrase in a specific context. Given an undefined phrase $X_{trg} = \{x_j, \cdots, x_k\}$ with its context $X = \{x_1, \cdots, x_I\}$ $(1 \leq j \leq k \leq I)$, our task is to output a description $Y = \{y_1, \cdots, y_T\}$. Here, X_{trg} can be a word or a short phrase and is included in X. Y is a

definition-like concrete and concise sentence that describes the X_{trg}.

For example, given a phrase "sonic boom" with its context "the shock wave may be caused by *sonic boom* or by explosion," the task is to generate a description such as "sound created by an object moving fast." If the given context has been changed to "this is the first official tour to support the band's latest studio effort, 2009's *Sonic Boom*," then the appropriate output would be "album by Kiss."

The process of description generation can be modeled with a conditional language model as

$$\boldsymbol{p}(\boldsymbol{Y}|\boldsymbol{X},\boldsymbol{X_{trg}}) = \prod_{t=1}^{T} \boldsymbol{p}(\boldsymbol{y_t}|\boldsymbol{y_{<t}},\boldsymbol{X},\boldsymbol{X_{trg}}). \quad (1)$$

4. LOG-CaD: Local & Global Context-aware Description Generator

In this section, we describe our idea of utilizing local and global contexts in the description generation task, and present the details of our model.

4.1 Local & global contexts

When we find an unfamiliar phrase in text and it is not defined in dictionaries, how can we humans come up with its meaning? As discussed in Section 2, we may first try to figure out the meaning of the phrase from the immediate context, and then read through the entire document or search the web to understand implicit information behind the text.

In this paper, we refer to the explicit contextual information included in a given sentence with the target phrase (i.e., the X in Eq. (1)) as "local context," and the implicit contextual information in massive text as "global context." While both local and global contexts are crucial for humans to understand unfamiliar phrases, are they also useful for machines to generate descriptions? To verify this idea, we propose to incorporate both local and global contexts to describe an unknown phrase.

4.2 Proposed model

Figure 1 shows an illustration of our **LOG-CaD** model. Similarly to the standard encoder-decoder model with attention (Bahdanau et al., 2015; Luong and Manning, 2016), it has a context encoder and a description decoder. The challenge here is that the decoder needs to be conditioned not only on the local context, but also on its global context. To incorporate the different types of contexts, we propose to use a gate function similar to Noraset et al. (2017) to dynamically control how the global and local contexts influence the description.

Local & global context encoders. We first describe how to model local and global contexts. Given a sentence X and a phrase X_{trg}, a bidirectional LSTM (Gers et al., 1999) encoder generates a sequence of continuous vectors $H = \{h_1 \cdots, h_I\}$ as

$$\boldsymbol{h}_i = \text{Bi-LSTM}(\boldsymbol{h}_{i-1}, \boldsymbol{h}_{i+1}, \boldsymbol{x}_i), \quad (2)$$

where $\boldsymbol{x}_i$ is the word embedding of word x_i. In addition to the local context, we also utilize the global context obtained from massive text. This can be achieved by feeding a phrase embedding $\boldsymbol{x_{trg}}$ to initialize the decoder (Noraset et al., 2017) as

$$\boldsymbol{y}_0 = \boldsymbol{x}_{trg}. \quad (3)$$

Here, the phrase embedding $\boldsymbol{x}_{trg}$ is calculated by simply summing up all the embeddings of words that consistute the phrase X_{trg}. Note that we use a randomly-initialized vector if no pre-trained embedding is available for the words in X_{trg}.

Description decoder. Using the local and global contexts, a description decoder computes the conditional probability of a description Y with Eq. (1), which can be approximated with another LSTM as

$$\boldsymbol{s}_t = \text{LSTM}(\boldsymbol{y}_{t-1}, \boldsymbol{s'}_{t-1}), \quad (4)$$
$$\boldsymbol{d}_t = \text{ATTENTION}(\boldsymbol{H}, \boldsymbol{s}_t), \quad (5)$$
$$\boldsymbol{c_{trg}} = \text{CNN}(X_{trg}), \quad (6)$$
$$\boldsymbol{s'_t} = \text{GATE}(\boldsymbol{s_t}, \boldsymbol{x_{trg}}, \boldsymbol{c_{trg}}, \boldsymbol{d_t}), \quad (7)$$
$$p(y_t|y_{<t}, X_{trg}) = \text{softmax}(\boldsymbol{W}_{s'}\boldsymbol{s'_t} + \boldsymbol{b}_{s'}), \quad (8)$$

where $\boldsymbol{s}_t$ is a hidden state of the decoder LSTM ($s_0 = \vec{0}$), and $\boldsymbol{y}_{t-1}$ is a jointly-trained word embedding of the previous output word y_{t-1}. In what follows, we explain each equation in detail.

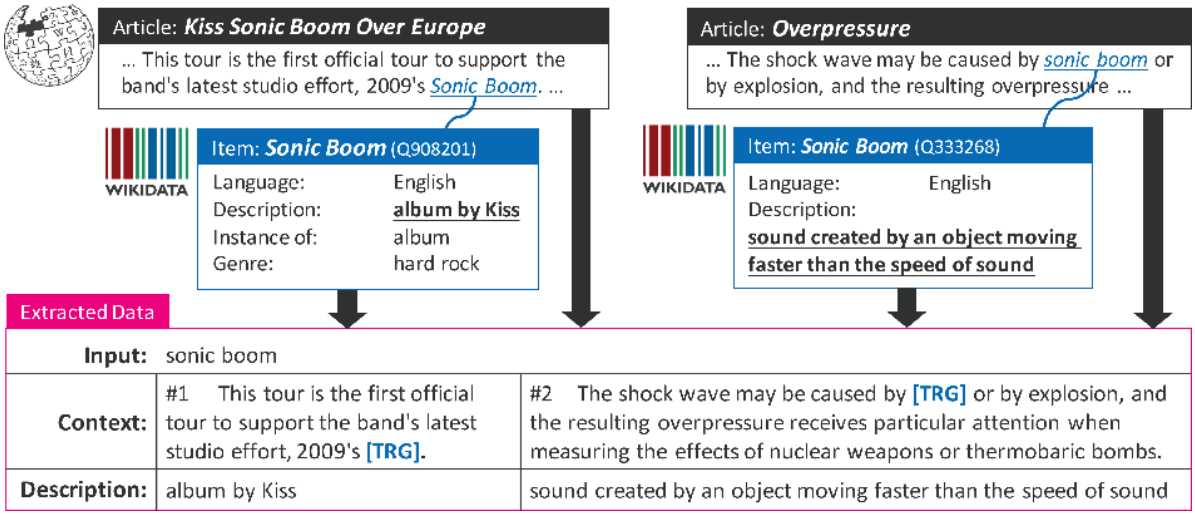


Figure 2: Context-aware description dataset extracted from Wikipedia and Wikidata.

Attention on local context. Considering the fact that the local context can be relatively long (e.g., around 20 words on average in our Wikipedia dataset introduced in Section 5), it is hard for the decoder to focus on important words in local contexts. In order to deal with this problem, the ATTENTION(·) function in Eq. (5) decides which words in the local context X to focus on at each time step. $\boldsymbol{d}_t$ is computed with an attention mechanism (Luong and Manning, 2016) as

$$\boldsymbol{d}_t = \sum_{i=1}^{T} \alpha_i \boldsymbol{h}_i, \tag{9}$$

$$\alpha_i = \text{softmax}(\boldsymbol{U}_h \boldsymbol{h}_i^T \boldsymbol{U}_s \boldsymbol{s}_t), \tag{10}$$

where $\boldsymbol{U}_h$ and $\boldsymbol{U}_s$ are matrices that map the encoder and decoder hidden states into a common space, respectively.

Use of character information. In order to capture the surface information of X_{trg}, we construct character-level CNNs (Eq. (6)) following (Noraset et al., 2017). Note that the input to the CNNs is a sequence of words in X_{trg}, which are concatenated with special character "_," such as "sonic_boom." Following Noraset et al. (2017), we set the CNN kernels of length 2-6 and size 10, 30, 40, 40, 40 respectively with a stride of 1 to obtain a 160-dimensional vector $\boldsymbol{c}_{trg}$.

Gate function to control local & global contexts. In order to capture the interaction between the local and global contexts, we adopt a GATE(·) function (Eq. (7)) which is similar to Noraset et al. (2017). The GATE(·) function updates the LSTM output $\boldsymbol{s}_t$ to $\boldsymbol{s'}_t$ depending on the global context $\boldsymbol{x}_{trg}$, local context $\boldsymbol{d}_t$, and character-level information $\boldsymbol{c}_{trg}$ as

$$\boldsymbol{f}_t = [\boldsymbol{x}_{trg}; \boldsymbol{d}_t; \boldsymbol{c}_{trg}] \tag{11}$$

$$\boldsymbol{z}_t = \sigma(\boldsymbol{W}_z[\boldsymbol{f}_t; \boldsymbol{s}_t] + \boldsymbol{b}_z), \tag{12}$$

$$\boldsymbol{r}_t = \sigma(\boldsymbol{W}_r[\boldsymbol{f}_t; \boldsymbol{s}_t] + \boldsymbol{b}_r), \tag{13}$$

$$\tilde{\boldsymbol{s}}_t = \tanh(\boldsymbol{W}_s[(\boldsymbol{r}_t \odot \boldsymbol{f}_t); \boldsymbol{s}_t] + \boldsymbol{b}_s), \tag{14}$$

$$\boldsymbol{s}'_t = (1 - \boldsymbol{z}_t) \odot \boldsymbol{s}_t + \boldsymbol{z}_t \odot \tilde{\boldsymbol{s}}_t, \tag{15}$$

Where $\sigma(\cdot)$, $\odot$ and ; denote the sigmoid function, element-wise multiplication, and vector concatenation, respectively. $\boldsymbol{W}_*$ and $\boldsymbol{b}_*$ are weight matrices and bias terms, respectively. Here, the update gate $\boldsymbol{z}_t$ controls how much the original hidden state $\boldsymbol{s}_t$ is to be changed, and the reset gate $\boldsymbol{r}_t$ controls how much the information from $\boldsymbol{f}_t$ contributes to word generation at each time step.

5. Wikipedia Dataset

Our goal is to let machines describe unfamiliar words and phrases, such as polysemous words, rarely used idioms, or emerging entities. Among the three existing datasets, WordNet and Oxford dictionary mainly target the words but not phrases, thus are not perfect test beds for this goal. On the other hand, although the Urban Dictionary dataset contains descriptions of rarely-used phrases, the domain of its targeted words and phrases is limited to Internet slang.

In order to confirm that our model can generate the description of entities as well as polysemous

words and slang, we constructed a new dataset for context-aware phrase description generation from Wikipedia and Wikidata which contain a wide variety of entity descriptions with contexts. The overview of the data extraction process is shown in Figure 2. Each entry in the dataset consists of (1) a phrase, (2) its description, and (3) context (a sentence).

For preprocessing, we applied Stanford Tokenizer to the descriptions of Wikidata items and the articles in Wikipedia. Next, we removed phrases in parentheses from the Wikipedia articles, since they tend to be paraphrasing in other languages and work as noise. To obtain the contexts of each item in Wikidata, we extracted the sentence which has a link referring to the item through all the first paragraphs of Wikipedia articles and replaced the phrase of the links with a special token **[TRG]**. Wikidata items with no description or no contexts are ignored. This utilization of links makes it possible to resolve the ambiguity of words and phrases in a sentence without human annotations, which is a major advantage of using Wikipedia. Note that we used only links whose anchor texts are identical to the title of the Wikipedia articles, since the users of Wikipedia sometimes link mentions to related articles.

6. Experiments

Datasets. We evaluate our method by applying it to describe words in WordNet (Miller, 1995) and Oxford Dictionary, phrases in Urban Dictionary and Wikipedia/Wikidata. For all of these datasets, a given word or phrase has an inventory of senses with corresponding definitions and usage examples. These definitions are regarded as groundtruth descriptions.

Models. We implemented four methods: (1) **Global** (Noraset et al., 2017) that access the pretrained embeddings of the phrase to be described, but has no ability to read the local context; (2) **Local** (Ni and Wang, 2017) which consists of the encoders to read the local context (i.e., usage examples of the terms); (3) **I-Attention** (Gadetsky et al., 2018), and the (4) **LOG-CaD**. Similar to our model, **I-Attention** utilizes both local and global contexts. Unlike our model, on the other hand, it does not use character information to predict descriptions. Also, it uses the local context only to disambiguate the phrase embedding instead of encoding the local context directly to predict descriptions. All the models are implemented with the PyTorch framework (Ver. 1.0.0).

Model	WordNet	Oxford	Urban	Wikipedia
Global	24.10	15.05	6.05	44.77
Local	22.34	17.90	9.03	52.94
I-Attention	23.77	17.25	10.40	44.71
LOG-CaD	**24.79**	**18.53**	**10.55**	**53.85**

Table 1: BLEU scores on four datasets.

Model	Annotated score
Local	2.717
LOG-CaD	**3.008**

Table 2: Averaged human annotated scores on Wikipedia dataset.

Automatic evaluation. Table 1 shows the BLEU (Papineni et al., 2002) scores of the output descriptions. We can see that the **LOG-CaD** consistently outperforms the three baselines in all four datasets. This result indicates that using both local and global contexts help describe the unknown words/phrases correctly. While the **I-Attention** also uses local and global contexts, its performance was always lower than the **LOG-CaD**. This result shows that using local context to predict description is more effective than using it to disambiguate the meanings in global context. In particular, the low BLEU scores of **Global** and **I-Attention** models on Wikipedia dataset suggest that it is necessary to learn to ignore the noisy information in global context if the coverage of pre-trained word embeddings is extremely low. We suspect that the Urban Dictionary task is too difficult and the results are unreliable considering its extremely low scores and high ratio of unknown tokens in generated descriptions.

Manual evaluation. To compare the proposed model and the strongest baseline in Table 1 (i.e., the **Local** model), we performed a human evaluation on our dataset. We randomly selected 100 samples from the test set of the Wikipedia dataset and asked three native English speakers to rate the output descriptions from 1 to 5 points as: 1) completely wrong or self-definition, 2) correct topic with wrong information, 3) correct but incomplete, 4) small details missing, 5) correct. The averaged scores are reported in Table 2. Pair-wise bootstrap resampling test (Koehn, 2004) for the annotated scores has shown that the superiority of **LOG-CaD** over the Local model is statistically significant ($p < 0.01$).

Input:	waste	
Context:	#1	#2
	if the effort brings no compensating gain it is a **waste**	We **waste** the dirty water by channeling it into the sewer
Reference:	useless or profitless activity	to get rid of
Global:	to give a liquid for a liquid	
Local:	a state of being assigned to a particular purpose	to make a break of a wooden instrument
I-Attention:	a person who makes something that can be be be done	to remove or remove the contents of
LOG-CaD:	a source of something that is done or done	to remove a liquid

Table 3: Descriptions for a word in WordNet.

Input:	daniel o'neill	
Context:	#1	#2
	after being enlarged by publisher **daniel o'neill** it was reportedly one of the largest and most prosperous newspapers in the united states.	in 1967 he returned to belfast where he met fellow belfast artist **daniel o'neill**.
Reference:	american journalist	irish artist
Global:	american musician	
Local:	american publisher	british musician
I-Attention:	american musician	american musician
LOG-CaD:	american writer	british musician

Table 4: Descriptions for a phrase in Wikipedia.

Qualitative Analysis. Table 3 shows a word in the WordNet, while Table 4 show the examples of the entities in Wikipedia. When comparing the two datasets, the quality of generated descriptions of Wikipedia dataset is significantly better than that of WordNet dataset. The main reason for this result is that the size of training data of the Wikipedia dataset is 64x larger than the WordNet dataset (14k entries vs. 887k entries).

For both examples in the two tables, the **Global** model can only generate a single description for each input word/phrase because it cannot access any local context. In the Wordnet dataset (Table 3), only the **I-Attention** and **LOG-CaD** models can successfully generate the concept of "remove" given the context #2. This result suggests that considering both local and global contexts are essential to generate correct descriptions. In our Wikipedia dataset (Table 4), both the **Local** and **LOG-CaD** models can describe the word/phrase considering its local context. As shown in the table, both the **Local** and **LOG-CaD** models could generate "american" in the description for "daniel o'neill" given "united states" in context #1, while they could generate "british" given "belfast" in context #2. On the other hand, the **I-Attention** model could not describe the two phrases, taking into account the local contexts.

7. Conclusions

This paper sets up a task of generating a natural language description for an unknown phrase with a specific context, aiming to help us acquire unknown word senses when reading text. We approached this task by using a variant of encoder-decoder models that capture the given local context with the encoder and global contexts with the decoder initialized by the target phrase's embedding induced from massive text. We performed experiments on three existing datasets and one newly built from Wikipedia and Wikidata. The experimental results confirmed that the local and global contexts complement one another and are both essential; global contexts are crucial when local contexts are short and vague, while the local context is important when the target phrase is polysemous, rare, or unseen.

As future work, we plan to modify our model to use multiple contexts in text to improve the quality of descriptions, considering the "one sense per discourse" hypothesis (Gale et al., 1992).

Acknowledgements

This article is a reconstruction of the author's PhD thesis "Translation and Description Methods for Multilingual Text Understanding." I am grateful to my PhD advisors at the UTokyo, Prof. Masaru Kitsuregawa, Prof. Masashi Toyoda, and Prof. Naoki Yoshinaga for guiding me during the long journey to the PhD. I would also like to express my gratitude to my mentors at CMU, Prof. Graham Neubig and Dr. Hiroaki Hayashi for their substantial contributions to the paper. Finally, I would also like to thank all the researchers who reviewed and recommended the thesis to AAMT Nagao Student Award.

References

Ion Androutsopoulos and Prodromos Malakasiotis. 2010. A survey of paraphrasing and textual entailment methods. *Journal of Artificial Intelligence Research*, 38:135–187.

Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural machine translation by jointly learning to align and translate. *In Proceedings of the Third International Conference on Learning Representations (ICLR)*.

Yuzhen Chen. 2012. Dictionary use and vocabulary learning in the context of reading. *International Journal of Lexicography*, 25(2):216–247.

Michael Connor and Dan Roth. 2007. Context sensitive paraphrasing with a global unsupervised classifier. In *Proceedings of the 18th European Conference on Machine Learning (ECML)*, pages 104–115.

Katrin Erk. 2006. Unknown word sense detection as outlier detection. In *Proceedings of the Human Language Technology Conference of the North American Chapter of the Association of Computational Linguistics (NAACL)*, pages 128–135.

Carol A. Fraser. 1998. The role of consulting a dictionary in reading and vocabulary learning. *Canadian Journal of Applied Linguistics*, 2(1-2):73–89.

Artyom Gadetsky, Ilya Yakubovskiy, and Dmitry Vetrov. 2018. Conditional generators of words definitions. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (ACL), Short Papers*, pages 266–271.

William A. Gale, Kenneth W. Church, and David Yarowsky. 1992. One sense per discourse. In Proceedings of the workshop on Speech and Natural Language, HLT, pages 233–237.

Felix A. Gers, Jurgen Schmidhuber, and Fred Cum-¨ mins. 1999. Learning to forget: Continual prediction with lstm. In *Proceedings of the Ninth International Conference on Artificial Neural Networks (ICANN)*, pages 850–855.

Philipp Koehn. 2004. Statistical significance tests for machine translation evaluation. In *Proceedings of the 2004 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 388–395.

Jey Han Lau, Paul Cook, Diana McCarthy, Spandana Gella, and Timothy Baldwin. 2014. Learning word sense distributions, detecting unattested senses and identifying novel senses using topic models. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 259–270.

Minh-Thang Luong and Christopher D. Manning. 2016. Achieving open vocabulary neural machine translation with hybrid word-character models. *In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 1054–1063.

Nitin Madnani and Bonnie J. Dorr. 2010. Generating phrasal and sentential paraphrases: A survey of data-driven methods. *Computational Linguistics*, 36(3):341–387.

Aurelien Max. 2009. Sub-sentencial paraphrasing by´ contextual pivot translation. In *Proceedings of the 2009 Workshop on Applied Textual Inference*, pages 18–26.

Aurelien Max, Houda Bouamor, and Anne Vilnat.´ 2012. Generalizing sub-sentential paraphrase acquisition across original signal type of text pairs. In *Proceedings of the 2012 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning*, pages 721–731.

George A Miller. 1995. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39– 41.

Roberto Navigli. 2009. Word sense disambiguation: A survey. *ACM COMPUTING SURVEYS*, 41(2):1–69.

Ke Ni and William Yang Wang. 2017. Learning to explain non-standard English words and phrases. In *Proceedings of the 8th International Joint Conference on Natural Language Processing (IJCNLP)*, pages 413–417.

Thanapon Noraset, Chen Liang, Larry Birnbaum, and Doug Downey. 2017. Definition modeling: Learning to define word embeddings in natural language. In *Proceedings of the 31st AAAI Conference on Artificial Intelligence (AAAI)*, pages 3259–3266.

Kishore Papineni, Salim Roukos, Todd Ward, and Wei-Jing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In *Proceedings of 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, pages 311–318.

Advaith Siddharthan. 2014. A survey of research on text simplification. *International Journal of Applied Linguistics*, 165(2):259–298.

多言語翻訳技術を活用した自動翻訳サービス

安西 健
凸版印刷株式会社

グローバル化が進み外国人との交流が増え、多言語化へのニーズが高まる中、IT 技術を活用したニューラル機械翻訳の登場によって、より自然な訳文を生成することが可能なった。この数年で、様々な分野において多言語翻訳技術を活用したサービスが普及してきている。

～音声翻訳サービス「VoiceBiz®」～

当社は、2018 年 5 月に音声翻訳サービス「VoiceBiz®」(※1)を販売開始し、現在、自治体、学校、医療施設、商業施設、刑務所など様々な分野で活用が広がっている。

在留外国人の対応において、ブラジル人（ポルトガル語）や中国人などのメジャーな言語を話す国の人は、各国のコミュニティーに属していることや通訳者による公共対応により周囲のサポートが受けられる環境もある。一方、希少言語に関しては集住していないことや公共や周囲のサポートが満足にできず全く話しが通じない場合もある。学校においては通訳者の確保も難しく、生徒やその父母の孤立につながることもあると聞いている。

音声翻訳サービス「VoiceBiz®」は、音声 12 言語、テキスト 30 言語の翻訳が可能で、専門用語や各分野で使用する定型文を搭載し、随時、機能のアップデートを図っている。また、機械翻訳で発生する誤訳に対して、「再翻訳機能」により翻訳結果が正しいかを確認できることで相手に正しく伝わっているか判断できる。一般的な会話の翻訳を目的とした音声翻訳機やアプリでは意思疎通が難しい専門的な分野でも，誤訳を少なくしコミュニケーションを取りやすくしている。

～AI 機械翻訳サービス「PharmaTra®(ﾌｧｰﾏﾄﾗ)」～

2020 年 4 月に AI 機械翻訳サービスとして、製薬業界の過去訳データを活用した「PharmaTra®(ﾌｧｰﾏﾄﾗ)」(※2）の販売を開始した。翻訳バンク®（※3）の一環として、世界の大手製薬会社の日本開発部門責任者の団体である R&D Head Club(RDHC)の 8 社から提供された 300 万文対を超えるコーパス（原文と訳文を対にして蓄積したデータベース）から深層学習で構築した AI 機械翻訳を活用。加えて、実用的な臨床開発用語集を搭載することにより、新薬開発関連文書を中心とした製薬文書に特化した翻訳を実現している。また、2021 年 3 月には、金融庁向けにもテキスト翻訳サービスの提供を開始した。今後は他の業界に特化したテキスト機械翻訳サービスの展開を進める予定。

～2025 年日本国際博覧会（大阪・関西万博）において「同時通訳」の利活用を推進～

当社は、総務省が 2020 年度より新規に実施する情報通信技術の研究開発課題「多言語翻訳技術の高度化に関する研究開発」の委託先 9 団体の 1 団体として選定され、「総務省委託・多言語翻訳技術高度化推進コンソーシアム」を 2020 年 8 月に設立した。(※4)

コンソーシアムは、グローバルコミュニケーション計画 2025(2020 年 3 月 31 日総務省より(※5))の推進のため、既に実用化されている『逐次通訳』の技術を『同時通訳』の技術にまで高度化し、ビジネス等の場面での利活用を可能にすることを目指している。

2025 年日本国際博覧会（大阪・関西万博）においては、様々な外国人来場者を想定し AI による高度な同時通訳技術により、「万博公式アプリ」、「バーチャル会場」、「Web 会場」での多言語同時通訳、会場内「多言

Automatic translation services utilizing multilingual translation technology
Takeshi Anzai
TOPPAN INC.

語案内端末・ロボット」など様々な自動翻訳システムの社会実装を検討している。(※6)

～今後のサービス展開～

新型コロナウイルス流行により、テレワークや遠隔会議が普及しオンライン環境よる多言語翻訳の必要性も高まっている。今後、コロナ禍の収束により、訪日外国人・在留外国人が増加することで社会経済活動において「言葉の壁」を感じさせないグローバルで自由な交流機会が拡大することが予想される。

当社は、多言語翻訳技術（機械翻訳・音声翻訳など）の高度化と社会実装の進展に向け、様々な分野において自動翻訳のニーズに応え、利用者の利便性を向上するサービスの実用化・普及を推進していく。

(※1)
https://www.toppan.co.jp/solution/service/VoiceBiz/
(※2)
https://www.toppan.co.jp/news/2020/03/newsrelease200309.html
(※3)https://h-bank.nict.go.jp/
(※4)
https://www.toppan.co.jp/news/2020/08/newsrelease_200828_2.html
(※5)
https://www.soumu.go.jp/menu_news/s-news/01tsushin03_02000298.html
(※6)
https://www.expo2025.or.jp/wp/wp-content/themes/expo2025orjp/assets/pdf/sponsorship/210819_explanatory_materials.pdf

長尾賞受賞者の今

8年間を振り返って

相良 美織
株式会社バオバブ

1. 変態と拡張と蓄積

2010年創業当初､統計機械翻訳全盛の夜明けを見るころ、バオバブは「留学生ネットワーク@みんなの翻訳」というプラットフォームにて機械翻訳のための学習データ（コーパス）を高品質、かつスピーディに構築し提供するサービスを展開していました｡実際には、機械翻訳で訳出した結果を留学生などの翻訳者が修正し（Post Edit）蓄積されたコーパスを機械翻訳に学習させ、その機械翻訳の結果をまた下訳として翻訳者が翻訳し、一定の品質に収斂させながら大量の学習データを構築する「人力と機械学習」を組み合わせた仕組みです。

その後､2015年に画像アノテーションサービスを開始、以降、音声の書き起こしや感情タグ付け、テキストアノテーションなど、現在では様々な分野に学習データ構築サービスの幅を広げています。

最近では、バオバブがデータセットの構築を手がけた東京大学による「ビジネスシーン対話対訳コーパスの構築と対話翻訳の課題」が言語処理学会第27回年次大会にて言語資源賞を受賞。クライアントに納品したデータがほんの少しでも役にたった、論文が学会で発表されたという知らせは何よりも嬉しく、いつも社内のチームやBaopart（注:バオバブにおけるアノテーターの呼称）に共有し、喜びを分かち合っています。

画像アノテーションでは、現在、４つの障害者就労支援施設にアノテーションの作業を委託しています。30ページ以上にも及ぶガイドラインを熟読し､細かなラベルの基準やルールをアノテーション作業に落とし込むこの作業は、何よりも正確性が要求されます。

このコツコツと大いに集中力を要する作業は自閉症を含む発達障害者の障害の適性とマッチし、そのスピードと緻密さは驚くばかりです。

興味深いのは、機械翻訳のための学習データを構築していた時の知見や失敗が大いに画像認識のための学習データ構築に資する結果になっていることです。

例えば､発達障害者とのコミュニケーションには｢機械翻訳に適した日本語文の書き方｣が非常に有用です。

下記に一部を紹介しましょう。

1. 主語をできるだけ省かない。
   特に主語が自分でない場合は省かない。
2. 目的語や助詞をできるだけ省かない。(明示する)
3. 1文あたりの文字数は長すぎないようにする。
4. 漢字をなるべく使う。
5. ことわざ、流行り言葉、擬音語、擬態語、オノマトペはできるだけ使わない。

例:

・ザザザッと仕上げてください。
・サクサク作業できない場合は教えてください。
・もうちょい左側に囲んでください。
・ぴえん

6. 適度に句読点を使う。
7. 明確な動詞を使う。
8. 無駄にカタカナをひらがなにしない。

2. 今、心にあるのは

「あなたはそれをどうやって、やるつもりなの？」

「『相良さんは将来何をやってみたいの』」と問われ「猫や動物と人間との自動翻訳機を作りたいです。」と回答した時、まっすぐに目を見て真摯に聴いてくださったあのお顔。かと思えば、2015年の関西MT勉

Looking back on the eight years
Miori Sagara
Baobab, Inc.

強会にて「人間を人工知能が超える日が来るのでしょうか？」という参加者からの質問に毅然と「人間とは、コンピュータに負けるような、そんなひ弱なものではない。人間は子、孫を残し、豊かな情感を持って生きる存在である。」と回答された時のあの会場の寄せては返す波のような静かな興奮。

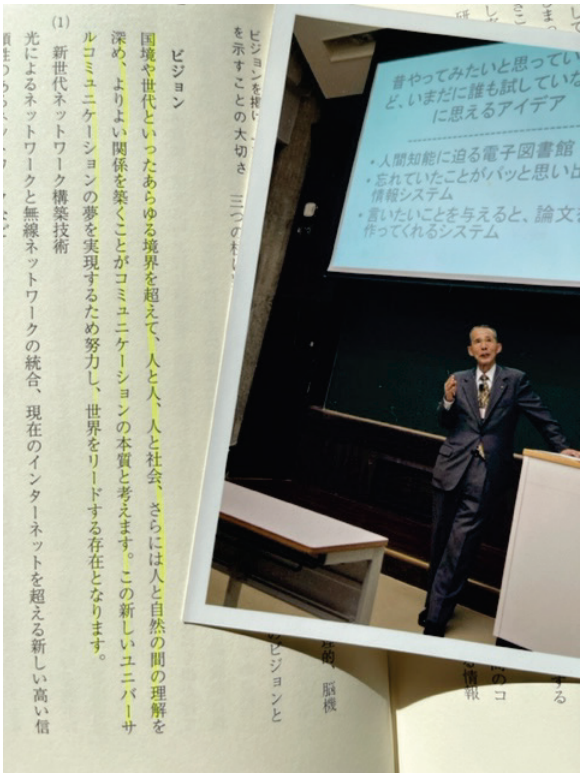
ただただその威光に圧倒されました。

「国境や世代といったあらゆる境界を超えて、人と人、人と社会、さらには人と自然の間の理解を深め、より良い関係を築くことがコミュニケーションの本質と考えます。この新しいユニバーサルコミュニケーションの夢を実現するため努力し、世界をリードする存在となります。」

2011 年情報通信研究機構に勤務した際に初めて当機構のビジョンを拝見し大変失礼ながら「役所の作ったビジョンにしてはかなりロックでイケている！」としびれました。勤務していた２年間、PC にずっと貼り付けていたこのビジョンは長尾先生が情報通信研究機構初代理事長に就任された際に作られたと知ったのはずっと後のこと。

今改めて、思う。なんてロックでイケているのだろう！

まだまだ精進が足りない。



「情報を読む力、学問する心」ミネルヴァ書房

写真：2015 年 3 月 16 日　関西 MT 勉強会

言語処理学会年次大会　於：京都大学

長尾賞受賞者の今

2017年長尾賞を受賞したヤラクゼンのその後とこれから

坂西 優
八楽株式会社

1. 受賞後のヤラクゼン

オンライン翻訳プラットフォーム「ヤラクゼン」は、シンプルな画面のオンライン翻訳ツールである。MSワードやエクセル、PDFなどの各種ファイルをそのままドラッグ＆ドロップして、複数の翻訳エンジンから最適なものを選び、機械翻訳にかけることができる。また、機械翻訳後は翻訳文の修正や文書共有、翻訳発注などのコラボレーション機能も利用できるのが特徴である。その使いやすいUI/UXと、その品質、スピードが評価されて延べ1,000社以上に導入されている。

「ヤラクゼン」は2017年、名誉ある長尾賞を頂いた。その際には、多くの翻訳関係者から祝福の声をいただき、ヤラクゼンの認知が向上し、結果、問い合わせが増加したように思われる。

その後、ニューラル機械翻訳のブレイクスルーの恩恵もあって、ヤラクゼンはファイル対応の拡充やセキュリティの強化、用語対応などの機能を強化し、結果、多くのユーザーを獲得することができた。また、ニューラルの出現時には、NICTの内山将夫先生にも様々なアドバイスを頂いた。

近年では、機械翻訳を効果的に活用するための「プリエディット」と「ポストエディット」のコツをまとめた著書「自動翻訳大全」をリリース。一時期は、アマゾンの「ビジネス英語一般カテゴリー」にてベストセラーになるなど好評をいただいた。

また、音声翻訳にも挑戦しており、ヤラクゼンに音声認識機能を追加したハードウェア「ヤラクスティック（β）」をリリースした。ヤラクスティック（β）は、6つのマイクを搭載し、筐体の周囲3メートルの音声のみを抽出、その後、音声認識、（ヤラクゼンによる）機械翻訳、文毎に表示する仕組みである。従来の音声翻訳機のように一文ずつ翻訳→発話するものではなく、筐体を手元に置いておくと、流れる会話を途切れることなく翻訳、字幕のようにテキスト表示するのが大きな特徴となっている。機械翻訳にはヤラクゼンを利用するため、それまでに蓄積した翻訳メモリ（フレーズ）や用語が反映されるようになっている。

2. ヤラクゼンのこれから

引き続き、世界で最も使いやすいCATツールを目指して、開発を続けていく。そのために、品質向上に注力しながらもモバイル対応やQA機能、コラボレーション機能などを拡充していく。

そして、機械翻訳と人の協業によるAugmented Translation(拡張翻訳)の世界を追求していきたい。

YarakuZen, a winner of the 2017 Nagao Prize, then and now
Suguru Sakanishi
Yaraku Inc.

長尾賞受賞者の今

みらい翻訳のこれまでとこれから

森 勇二
株式会社みらい翻訳

1. はじめに

みらい翻訳は「RBMT と SMT 等を統合した機械翻訳プラットフォームの開発と B2B ビジネスの実現」を評価頂き第 11 回（2016 年）AAMT 長尾賞を受賞しました。本稿では受賞から 5 年間の事業面、技術面の変遷を中心に紹介します。

2. 長尾賞受賞当時のみらい翻訳

我々が長尾賞を受賞した 2016 年、無料/有料問わず機械翻訳サービスは統計的機械翻訳(Statistical MT; SMT)が主流でした。みらい翻訳は国立研究開発法人情報通信研究機構(NICT)より SMT のエンジン供与を受け、ルールベース機械翻訳(Rule-Based MT; RBMT)と組み合わせたハイブリッド構成や事前並べ替え技術により日英のベース精度を高め、用途に特化したカスタマイズモデルを載せた機械翻訳エンジンの提供を行っていました。2014 年の創業以来初めての単年度黒字を達成したのもこの年でした。しかしながら、当時の単純な延長線上に今のみらい翻訳の姿はありませんでした。

現在の機械翻訳技術の主流であるニューラル機械翻訳(Neural Machine Translation; NMT)は 2014 年に原型となるエンコーダ・デコーダモデルが提案されましたが、実用的には我々が長尾賞を受賞した直後の 2016 年 9 月に発表された Google NMT(GNMT)のインパクトが大きく、この時期を境に我々もいつ NMT を出す予定なのかという問い合わせを頂くようになりました。それまでは NMT 固有の課題も多く実用化はまだ先のことだと考えていましたが、急ピッチでの実用化を余儀なくされました。

3. 機械翻訳 SaaS 企業への転身

2017 年、SMT ベースの機械翻訳エンジンの新規販売を停止し、NICT との共同研究の成果を実用化する形で NMT エンジンの自社開発を開始しました。同年、ビジネス向け文書翻訳 SaaS の Mirai Translator®の開発に着手、2017 年 12 月にリリースしました。現在、NTT コミュニケーションズ社への OEM 提供による COTOHA® Translator も含め、これらのサービスがみらい翻訳の主力事業となっています。Mirai Translator は我々の得意とする日本語を中心とした翻訳精度の高さを軸に、一般ビジネスユーザに使いやすいシンプルな UI、ISO27001、27017 を取得するなど高いセキュリティで運用していることを強みとして一定以上の規模の企業を中心に導入が進んでおり、数万人規模の大規模導入となるようなケースも増えています。翻訳機能としてはテキスト翻訳、ファイル翻訳の２種類を提供しており、ファイル翻訳はファイル形式として txt, docx, xlsx, pptx, pdf をサポートしています。また、ユーザ辞書、翻訳メモリによるカスタマイズが可能となっています。

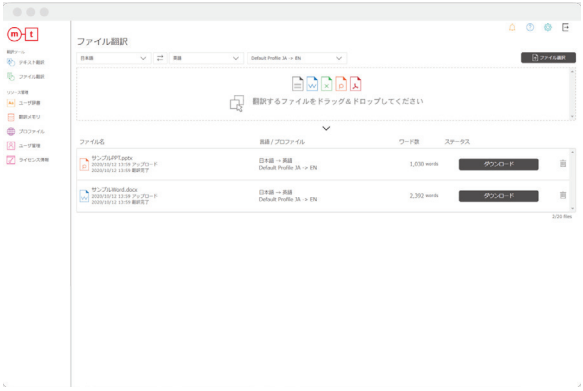


図 1 Mirai Translator®翻訳画面

The past and future of Mirai Translate, Inc.
Yuji Mori
Mirai Translate, Inc.

機械翻訳エンジンの内製化および自社サービスの開発・運用体制の構築により 2017-2018 年度は大幅な赤字を出すこととなりましたが、2019 年度に再度黒字化を達成しました。

創業以来、対外的に翻訳精度を伝える指標の一つとして TOIEC スコアを用いてきました。これは簡単に言うと TOIEC スコアが既知の被験者と機械翻訳とで日英翻訳の精度を競わせ、ネイティブ評価者による勝敗が均衡する箇所を機械翻訳のスコアとするもので、長尾賞受賞当時は約 650 点でした。SMT をベースとした漸近的な進化を想定した創業時の 2018 年における目標は 800 点でしたが、NMT により 2017 年に 900 点、2018 年には 960 点を達成しました。評価として意味を成さなくなってきていることと、評価自体にコストがかかることから最近はこの評価を行なっていません。前述の通り、創業当初のみらい翻訳は機械翻訳エンジンのカスタマイズを事業の中心としていましたが、ベース精度の向上により現在はカスタマイズサービスの引き合いはかなり少なくなっています。

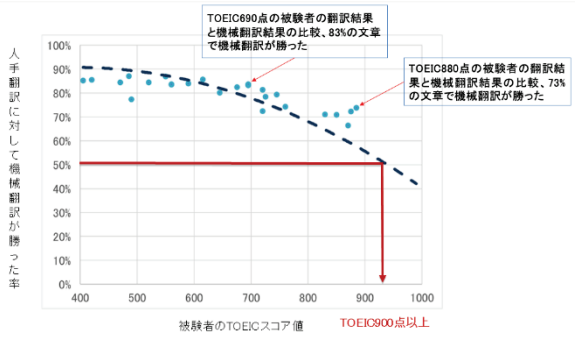


図 2 900 点を超える TOEIC スコアを達成

NMT により機械翻訳精度、とりわけ翻訳結果の流暢さは大幅に向上しましたが、仕組み上訳抜け/過剰生成や数値の間違い等のエラーは発生しやすくなりました。特にビジネスの現場では固有名詞や数値の間違いは深刻な問題を引き起こす可能性があります。そのため、みらい翻訳では全てをニューラルネットワークで解決するのではなく、言語に特化したルールの作り込みも行っています。多数の対応言語を誇る Web 翻訳サービスとの方向性の違いの一つと考えています。

4. ビジネスシーンでの DX、さらにその先へ

現在、我々は Mirai Translator®を多言語業務をこなす日系企業をメインターゲットとした生産性向上ソリューションと位置付けています。生産性向上のためには翻訳精度の向上以外にもやるべきことが数多くあります。一例を挙げると、ファイル翻訳の装飾保持です。RBMT や SMT と異なり NMT では原文と訳文の間の単語対応が明確でなく、当初のファイル翻訳では原文の装飾情報は翻訳の際に捨ててしまっていました。一方で翻訳して資料を仕上げるという観点では装飾も重要な情報で、仮に翻訳が完璧であったとしても失われた装飾情報の復元には後編集の工程でそれなりの稼働がかかります。そこで、翻訳と装飾保持を同時に解くモデルを開発しました。本機能により、ファイル翻訳の後編集工程にかかる時間を最大で 62%削減することに成功しました。

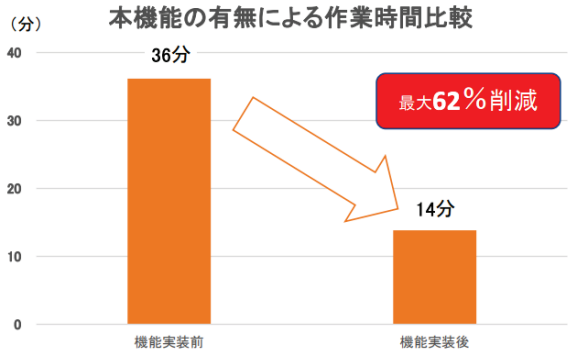


図 3 ファイル翻訳装飾保持の効果

NMT、特にアテンション付きエンコーダ・デコーダモデルの登場により機械翻訳の社会実装が大幅に進みました。これまで述べてきた通り、事業における当面の主戦場は企業における文書翻訳の生産性向上を如何に実現していくか、にあります。一方で我々が企業ビジョンとして掲げる「言語バリアフリーの世界の実現」に向けては個人向けサービスや音声翻訳サービスなど顧客セグメント、応用領域の両方を広げていく必要があります。みらい翻訳は今後も機械翻訳の社会実装により、言語バリアフリーの実現を推進していきます。

解説記事

MT 利用ガイドラインの必要性

山田 優

立教大学・AAMT

1. MT 利用ガイドラインの必要性

「自転車進入禁止」が「No Entrance Bicycles」、「授乳室」が「Suckle room」。千葉県浦安市で機械翻訳による誤訳が発覚した。その後、浦安市は検証委員会を設置。2021 年 3 月に同委員会は、誤訳の理由と対策を『浦安市多言語表記検証報告書』[1]にまとめた。必要だったのは翻訳確認体制の整備である[2]。

最近話題の「日本の英語を考える会」[3]は、日本国内に溢れる「変な英語」に関しての情報共有や意見交換を行っている。扱う事例は、機械翻訳による誤訳だけではないが、2020 年には同グループが目黒区のホームページに「Dark Procedure」という意味不明の表記を見つけ、それが「くらし手続き」の誤訳であることを確認。目黒区に連絡をして修正してもらった事が話題になった。機械翻訳の誤訳であった。上で挙げた例以外にも、数多くの MT 過誤の問題を、読者も知るところであろう。

さて、ここでのポイントは MT の実力不足では当然ない。考えてもらいたい。上述した問題はすべて人災ではないだろうか。具体的には、翻訳確認体制の整備の不足によるものかもしれない。そして、その根本原因は、人々の「翻訳」に対する不理解である。MT を使うのは人であり、MT ユーザへの啓蒙は、まさに喫緊の課題となっている。

上の状況に鑑み、AAMT では委員会を発足させ、健全かつ正しい MT（および翻訳・通訳）の利用を推進すべく「MT 利用ガイドライン（仮名）」を作成することを決定した。本稿では、その経緯と概要を示す[4]。

2. MT とは何か

ニューラル機械翻訳は、これまで越えることができなかった翻訳品質の天井を突き破る技術的な快挙を成し遂げた。それに伴い、MT ユーザ数は増え続け、MT 利用の場は拡大している。公共機関、ビジネスの利用にとどまらず、産業翻訳における MT+PE も、もはや当たり前になった。最近では、外国語教育の現場で、英語を学ぶために MT を活用するという議論もなされている。

この状況は、MT 利用を推進する AAMT にとっては喜ばしいことであるが、上述したような MT 過誤と誤訳による事故は社会問題に発展してきている。同協会は、健全かつ正しい MT 利用方法をユーザに啓蒙しなければならない立場にもある。

大切なのは、ユーザに「MT とは何か」を理解してもらうためには、「翻訳とは何か」を知ってもらう必要があるということだ。ユーザは、これまで翻訳自体についても、十分な説明を受けてこなかった。今の MT の勃興は、ユーザに「翻訳とは何か」を知る機会を与えてくれたとも言える。

ニューラル機械翻訳の勃興と AAMT の創立 30 周年との同期により「時は来たり」と喝破した隅田会長の言葉は正しく[5]、いま、このタイミングで「MT 利用ガイドライン」を作成することは理に適っているのである。

3. ガイドラインの内容

では、「MT 利用ガイドライン」の中身はどうなっているのか。実際には、AAMT 会員の志望者で構成する委員

All we need is “MT User Guidelines”
Masaru Yamada
Rikkyo University & AAMT

会で検討中であり、詳細はまだ決まっていない。以下では、「MT 利用ガイドライン」のイメージを理解してもらえるような参照冊子と、検討事項の概要を示す。

1 つ目。『翻訳で失敗しないために（Translation Getting it Right: A guide to buying translation）』[6]。2011 年に ATA（アメリカ翻訳協会）が作成し、JTF（日本翻訳連盟）の協賛で日本語版も発行された翻訳会社に（人手）翻訳を発注するユーザが翻訳で失敗しないための要点をまとめたガイドラインである。徹底的に翻訳ユーザの視点から書かれているのが良い。30 ページに満たない冊子で、内容も簡明である。「本当に翻訳する必要があるのか」、「言葉でなく絵や図で伝えたほうがよいのでは」、といった助言から始まり、原文の意味が伝わるだけでは駄目で、珍訳は笑いものにされ、ブランドや会社のイメージダウンになること等が書かれている。原文さえ用意すれば、翻訳者が適当に訳してくれるだろうという思い込みを変えるべく、ユーザ啓蒙に励む書である。「翻訳とは何か」をユーザ側の立場に立って、様々な視点とレベルで説明する。これを真似て「MT で失敗しないために」のような冊子を作れると良い。

2 つ目。『特許ライティングマニュアル』[7]。2013 年に産業日本語研究会が初版を発行した。特許文書を例に挙げつつ、人や機械にとって明瞭で翻訳もしやすい日本語文の書き方を紹介する。いわゆるプリエディット（前編集）のコツに似ている。文章の誤解防止が期待でき、翻訳者が翻訳しやすくなるのみならず、機械翻訳も容易になるので、高精度な機械翻訳文の低コストでの作成が狙える。大きな 7 つの基本ルールが明確に示されている。「短文にする」や、節・句レベルでは「省略しない」「読点を工夫する」などシンプルで具体的な説明がある。言語的な説明に限られるが、MT を使うという視点から参考になる情報が多い。

上記以外にも、現在、我々はあらゆる資料を参照し、ガイドラインに含めるべく情報と詳細化のレベルを検討している。例えば、JTF 品質ガイドラインにあるようなエラーカテゴリと対応づけて、「正確性」「流暢性」「用語」「スタイル適合」の観点から MT の実力を説明する方法もある。上述したプリエディットの要素と対応づければ、「短文にする」ことで「正確性」が向上するし、そうしないと正確性が低下するという説明ができる。

言語要素ではなく、利用状況に関する要素も重要だ。例えば、日本語（母語）と英語の組合せに MT を使う場合は、ユーザの英語力と訳出の方向性が重要な要素になる。英語力の高い人と、英語ができない人が MT を使う場合では、問題は異なるかもしれない。日本語への翻訳と英語への翻訳も、使い方やリスクの重篤度が異なるだろう。

他にも、オンライン MT を使うことに伴う、セキュリティ、倫理、人権、責任の所在などの諸問題を、ガイドラインにどこまで含めるのかも検討している。

4. 今後の予定

「MT 利用ガイドライン」は 2022 年 5 月末の完成を目指している。段階的に内容を充実化させていくこともあるが、その場合でも第 1 段階の内容を、その時までに発行する。

注・参考文献

[1] https://www.city.urayasu.lg.jp/_res/projects/default_project/_page_/001/032/102/tagengokenshohoukokusho.pdf

[2] https://mainichi.jp/articles/20210330/ddl/k12/010/087000c

[3] https://note.com/nihonnoeigo/

[4] Yamada, M. (2021) Translation process model and ethical issues – MT user guidelines, *eajs 2021 conference*, August 25, 2021.

[5] 隅田英一郎（2021）時は来たり, *AAMT Journal*, 74.

[6] https://www.jtf.jp/pdf/translation_order.pdf

[7] https://www.tech-jpn.jp/tokkyo-writing-manual/

解説記事

人間の翻訳と機械の翻訳（５）：翻訳者のコンピテンスとは何か

影浦 峡
東京大学・大学院教育学研究科

1. これまでのまとめと今回の話題

本連載では、第 1 回から第 3 回で、翻訳の対象、翻訳の対象である文書の性質、翻訳のプロセスについて論じました。第 4 回（前回）は少し趣向を変え、機械翻訳研究で語られている翻訳と翻訳論で語られている翻訳を計量書誌学的な観点から対比し、機械翻訳研究で捕らえられている翻訳の範囲を観察しました。

今回は、翻訳者のコンピテンスについて考えます[1]。話を始める前に、コンピテンス、そして関連する「知識」「スキル」の定義を与えておきます[2]。

知識：学習により情報を消化して得られるもの。

スキル：タスクの遂行や問題解決のために知識を適用しノウハウを活用する能力。

コンピテンス：仕事や学習の場で、また、専門性の開発や自己の啓発において、知識やスキル、個人的・社会的および／あるいは方法論的能力を使うための、確認された能力。

スキルが知識を使う能力であるのに対し、コンピテンスはさらに知識やスキルを使う能力として、一段高次のレベルで定義されていることがわかります。この区別は重要ですが、本稿の前半では、あまり定義にはこだわらず、ゆるやかに考えます。実際のところ、スキルとコンピテンスをはっきりと区別していない議論もありますし、また、どのようなコンピテンスが必要かをめぐる議論の展開を追う場合、時期によってコンピテンスの定義が同一でないこともあるからです。とはいえ、具体的な作業に対応して必要な知識、スキル、コンピテンスを定義し、MT がカバーしているのは知識なのかスキルなのかコンピテンスと言えるものなのかを検討する作業は、理論的に重要です。後半では少しそのことを意識します。

2. 翻訳コンピテンス論の展開

翻訳のスキルやコンピテンスをめぐる議論が盛んになった 1970 年代からの展開を簡単に追います[3]。

Wilss (1976)は、翻訳コンピテンス涵養の重要性を説き、そのコンピテンスを「技術的なテキスト、一般的なテキスト、文学テキストを目標言語で十分なかたちで再生産する能力」とし、そのために「起点言語と目標言語において文法、語、形態素の包括的な知識を有するだけでなく、起点言語と目標言語のテキスト世界に関する完全なスタイル上（テキスト上）の知識を必要とする」と述べています[4]。House (1980)は、翻訳コンピテンスを「読み、書き、聞き取り、話すことに加えて第5の基本的な**言語スキル**」と述べています[5]。

これらに対して Kiraly (1995)は、「翻訳コンピテンスではなく翻訳者コンピテンス」を翻訳教育の目標とすべきとし、この「用語を選ぶことで、プロ翻訳者のタスクが有する複雑な性格と、必要とされる**非言語的スキル**を強調する」ことができると述べています[5]。その後、コンピテンスといっても曖昧なままではないか、といった批判を経て[6]、コンピテンスのカテゴリが列挙されるようになります。

例えば、Kelly (2005)は、(a) コミュニケーションとテキストに関するコンピテンス、(b) 文化および間文化に関するコンピテンス、(c) 主題分野に関するコンピテンス、(d) 専門性および道具に関するコンピテンス、(e) 態度と心理＝生理的側面に関するコンピテンス、(f) 対人関係に関するコンピテンス、(g) 戦略に関するコンピテンス、という 7 種類のコンピテンス・

What do translators' competences consist of?
Kyo Kageura
Graduate School of Education, The University of Tokyo

カテゴリを挙げています[7]。

Wilss (1976)や House (1980)と比べると、コンピテンスの範囲が広くなっていることがはっきりとわかります。もちろんこれだけでは相変わらず、「で、結局、それはどういうもの？」という疑問が湧きますが、この点についても、翻訳教育や翻訳論のそれ以外のところで、一定程度、外在的な操作の手続きとして共有可能な具体性をもったかたちでコンピテンスを記述する試みは続いてきました。コンピテンスの定義が「確認された能力」であり、この「確認」は自己の内面報告ではあり得ないことを考えると、翻訳論におけるコンピテンスを巡る議論は、十分かどうかは別にして、(1) 翻訳の実態を反映し言語力以外に必要なコンピテンスを考慮する、(2) 具体的に共有可能な解像度での記述を目指す、という、適切な方向に向かってきたと言えるでしょう[8]。

3. ISO 17100 と EMT 2017

ISO (2015)は、翻訳者に求められるコンピテンスとして、以下の 6 つを挙げています[9]。

a) **翻訳コンピテンス**　［この規格に従って］内容を翻訳する能力で、言語コンテンツの理解と生成、クライアントと翻訳サービスプロバイダ（TSP）の取り決めをはじめとする翻訳の仕様に従って目標言語コンテンツを提供する能力を含む。

b) **起点言語と目標言語における言語とテキストに関するコンピテンス**　起点言語を理解する能力、目標言語の流暢さ、テキストタイプの約束事をめぐる知識。それを適用する能力を含む。

c) **調査、情報の収集と処理に関するコンピテンス**　起点言語コンテンツ理解や目標言語コンテンツ生成に必要な知識を効率的に入手する能力。ツールの利用や情報源の利用方法に関する能力を含む。

d) **文化に関するコンピテンス**　行動規範、用語法、価値体系、ロケールに関する情報を利用する能力。

e) **技術に関するコンピテンス**　ツールや IT システム等を用いて必要な技術的タスクを遂行する力。

f) **分野に関するコンピテンス**　起点言語文書のコンテンツを理解し、適切なスタイルと用語を使って目標言語を生成する能力。

a)の「この規格に従って」のところは、原文では「翻訳」を定めた項が具体的に参照されています[10]。そこでは、翻訳プロジェクトの目的に則して翻訳することと総論が書かれた上で、その際に、分野やクライアントの用語集等に従い一貫させること、意味の正確さを維持すること、適切な文法や句読法、正書法上の約束事に従うこと、表現の結束性や適切な言い回しを用いること、定められたスタイルガイドに従うこと、ロケール等基準に従うこと、適切なフォーマットを用いること、想定読者と翻訳目的に適したものにすること、が書かれています。

EMT (2017)は欧州共通翻訳修士枠組みにおけるコンピテンスの枠組みを定めたもので、ISO (2015)の a)-c)が「翻訳」の方略と方法に関するコンピテンス、f)が「翻訳」のテーマに関するコンピテンスと、分類が少し違いますが、大枠としては対応したコンピテンスが挙げられています[11]。他に人的・対人的なコンピテンス、サービス提供が挙げられています。教育の観点からこれらを含めるのは自然なことでしょう。

ISO (2015)も EMT (2017)も、コンピテンスを「言語的」なものとして（だけ）ではないものと捉えています。本連載第 1 回でも紹介しましたが、EMT (2017)は、少なくとも二つの言語の十分な力は、翻訳を学ぶための前提条件であると述べています。

4. 実証的観察から：調査のコンピテンス

ここでは、調査のコンピテンスに絞り、Onishi and Yamada (2020; 2021)の研究の一部を紹介します[12]。英日翻訳において、英語専攻の学生 5 名（翻訳授業の履修経験有）とプロの翻訳者 4 名を対象に、翻訳作業時の情報探索（調査）行動を量的・質的に分析したも

ので、次のようなことが観察されています。

1. 翻訳にかける時間はプロの翻訳者の方が長く、その差の大部分は情報探索に要した時間である。
2. プロの翻訳者は辞書の検索よりも辞書以外の検索に費やす時間がはるかに長いが学生は同じくらいの時間を費やす。辞書以外の情報探索が良質の翻訳を作る鍵でありそうである。
3. 目標言語の表現を検索する回数はプロの翻訳者が有意に多い。
4. 検索エンジン結果からのジャンプ数はプロの翻訳者の方が多い傾向がある。
5. プロ翻訳者は文書を理解するために下調べを行うが学生ではそうした意識は薄い。

全体の傾向として、学生が「言語」についての探索を行なっているのに対して、プロの翻訳者は文書の内容をめぐる探索を行なっていると言えそうです。このことは、翻訳が言語に関する行為ではなく文書に関する行為であることとよく対応しています。

この結果は、翻訳者の翻訳行動に広く一般化することはできません。例えば、特定領域・タイプの文書を専門に扱うプロの翻訳者は調査をほとんど必要としないことがあります。けれども、調査を要する翻訳タスクにおいて、良質の目標言語を作るためにプロの翻訳者がどのような行動を取るかについては、特徴を良く示しています。情報探索の能力が高い翻訳者の翻訳品質が高いことを示す研究もあります[13]。

5. コンピテンスの運用・運用のコンピテンス

ここまで、翻訳者に求められるコンピテンスについて枠組みと一部の実証結果を示してきました。実は、コンピテンスについては、概念的に少し曖昧なところが残っています。冒頭で挙げたコンピテンスの定義は「知識やスキル、・・・能力**を使う・・・能力**」となっています。ところが 2 節と 3 節でまとめたコンピテンスのカテゴリや枠組みは、基本的に、「知識やスキル、・・・能力」そのものです（ISO の b）に「それを適用する能力を含む」とありますが、これも、適用の判断を指しているのか個々の適用のスキルを指しているのかははっきりしません)。

4 節で見たプロの翻訳者と学生の翻訳者の違いは、単に言語レベルでの原文の理解力（それが何を意味するかについてここでは立ち入りませんが）や情報探索スキルの差だけを示しているのではなく、そもそも、どのようなときに情報探索を行う必要があるかをめぐる判断に関する差異を反映しています。誤訳は、起点言語文書がわからないときにではなく、わかっていないことを意識できないときに生まれます。トップレベルのプロの翻訳者に「なんとなく変だぞ」と感じる力があることはよく言われることです。

知識やスキルや能力を運用する力を運用のコンピテンスと呼ぶことにしましょう。冒頭のコンピテンスの定義は、第一義的には運用のコンピテンスを示しています。そうであるなら、「なんとなく変だぞ」のセンスは、コンピテンスの核にあると考えられます。

ここで**運用の**コンピテンスが問題になるのは、与えられた起点言語文書を前にしたとき、それを翻訳するために、既に有している知識やスキルを一般的に適用するだけでは十分ではないからです。敷衍しましょう。

第一に、3 節の ISO (2015)の紹介で述べたように、翻訳においては、(プロジェクト毎に) その都度、そこで定義された目的や条件、使用すべき用語集や句読法・正書法の規定等に従って訳出する必要があります。母語話者にとっては簡単なことと思われるかもしれませんが、学生のレポートを見るとそう簡単ではないことがわかりますし、大学教員でも、それを十分にできるわけでは必ずしもありません[14]。

第二に、本連載第 1 回で、言語処理には「知識がないとどうにも解けない問題がいっぱいある」という工藤拓さんの言葉を引用しましたが[15]、本質的に「言語処理」ではない翻訳では、知識があってもどうにもならなくて、というのも、与えられた文書を理解する必要があり、そこでは既往の知識を適用するのではなく知識を構築していく作業を伴い、それは知識の不足で

はなく文書を理解するということに必然的なことなので、すなわち、理念的に文書というものはそれを適用することで理解できる知識が事前に定義できないようなかたちであり、現実にもそのような文書は少なからずあって、翻訳者はそれに対応しなくてはならない、そのため、知識やスキルを身に付けていることは前提として、その都度、文書の理解に必要な知識を構築するために知識やスキルを発動する運用のコンピテンスが問われる、ということになります。

多少強引に「科学とは専門家の無知を信じることだ」というファインマンの言葉を借りると[16]、翻訳者は文書を前に自ら専門家だから専門家として自らの知識を適用することで翻訳ができるという前提ではなく、自らの無知を踏まえて翻訳に臨むという点で、未知の事象を前にした科学者に似ていると言えます。

少し話がずれますが、翻訳が完成するというのも不思議なもので、(工程としては定義できるとしても) 理論的にはどこで終了にしてよいのかよくわかりませんが、でも何故かしら、翻訳は一応、終わります。

6. おわりに

少し MT を意識してまとめます。第一に興味深いのは、翻訳者のコンピテンスをめぐる議論がコンピテンスを外在化して認識する方向に発展してきたことで、一般にこれは、民主的な共有、さらに機械化の方向なのですが、NMT の成功はむしろ処理のブラックボックス化と対応していること。NMT の出力が翻訳をよく知らないけれど言葉はある程度できる学生の訳に似ていることは、これに対応しています。MT の観点からは、スキル的なものとして翻訳者のコンピテンスのうちどの範囲が MT で扱われ、どの程度の能力を持っているかを、技術的な問題として操作可能性を維持した解像度で考えるのは重要な課題になるでしょう。

もう一つは、それぞれの文書に応じたその場での情報探索と理解と選択と決定が翻訳者のコンピテンスの中核にあることが示されたのに対し、これに対応する機構を MT は持っていないことで、喩えて言うと MT は事前の知識を適用することだけで全てを行おうとし、その都度、文書を未知のものとして理解しようとする科学的な態度を取れない、専門家のような存在です。分野適応をしたとしてもその分野の専門家になるに過ぎず、文書を前にした科学的態度をエミュレートできるわけではありません（技術においてそうしなくてはならないことは必ずしもないのですが)。

だから文書をその都度理解できる人間はすばらしいと言うことほど非科学的な態度もないので、ではどうすればよいかを問う必要があります。大きく二つの方向性がありそうです。第一は、MT の出力は翻訳結果ではないとすること。MTPE の研究はその方向を向いています。第二は、MT に翻訳者のコンピテンスに対応するコンピテンスを持たせること。品質推定（QE）が人間の翻訳の評価に対応するならば、与えられた文書に対する QE を起点にしてその文書の翻訳に必要な情報探索や学習を発動させる手法を考えることは、誰もが思いつく出発点の一つです。

ところで、「これで翻訳終了」の判断はどうすればよいのでしょうか。これは、人間にもわかっていません。MT が鳥のようにではなく飛ぶ飛行機のように実現できる可能性は、翻訳の全体が人間における行為としてしか定義されないことを考えると、あまり大きくはなさそうです。なので、MT が翻訳できるようになるために、翻訳についてまだ解明しなくてはならないところが、それを「解明」するとはどういうことかも含め、かなり残っていると考えるのが妥当なようです。

謝辞

翻訳・翻訳者のコンピテンスの整理は、朴惠氏と山田優氏（と）の研究（具体的な情報は「注・参考文献」に記載）に依拠しています。両氏に感謝します。本記事の一部は、科研研究費基盤(S)「翻訳規範とコンピテンスの可操作化を通した翻訳プロセス・モデルと統合環境の構築」（19H05660）に関わっています。

注・参考文献

[1] 本稿の英語タイトルでは ISO (2015) *ISO 17100: 2015 Translation Services – Requirements for Translation Services* を踏まえ“competences”を複数形にしました。ISO 日本語版では「力量」です。

[2] European Communities (2008) *The European Qualifications Framework for Lifelong Learning*. p. 11. の定義によっています。これらの定義は、European Master's in Translation (2017) *European Master's in Translation Competence Framework*. でも引用されています。コンピテンスの定義で「能力」が重なっていますが、元の定義でも“ability”と“abilities”が使われています。

[3] 次の文献に基づいています。朴惠, 影浦峡「翻訳コンピテンスとは何か、それはどのように規定されているか：翻訳教育カリキュラム開発に向けたレビュー」日本通訳翻訳学会第 19 回年次大会, 2018 年 9 月 9 日; Piao, H., Yamada, M. and Kageura, K. (to appear) “Metalanguages in translator education,” In Miyata, R., Kageura, K. and Yamada, M. (eds.) *Metalanguages for Dissecting Translation Processes: Theoretical Development and Practical Applications*. London: Routledge.

[4] Wilss, W. (1976) “Perspectives and limitations of a didactic framework for the teaching of translation,” In Brislin, R. W. (ed.), *Translation: Applications and Research*. New York: Gardner, pp. 117-137.

[5] House, J. (1980) “Übersetzen im Fremdsprachenunterricht,” In Poulsen, S. O. and Wilss, W. (eds). *Angewandte Übersetzungs-wissenschaft*. Arhus: Wirtschaftsuniversitzit Arhus, pp. 7-17. Kiraly (1995) *Pathways to Translation: Pedagogy and Process*. Kent: Kent State University Press.からの再引用。強調はいずれも引用者によるものです。

[6] 例えば、Lörscher, W. (1992) “Investigating the translation process,” *Meta*, 37(3), pp. 426-439.

[7] Kelly, D. (2005) *A Handbook of Translator Trainers*. London: Routledge.

[8] 外在的な操作の手続きとして共有可能な具体性をもった記述や理論はそもそも極めて重要で、障碍の社会化を経てユニバーサルアクセスを進めることができるのも、微積分の基本を私たちが高校で学ぶことができるのも、これによっています。思考力を補助する外在的な記号操作系の明確化とそれに基づく思考のテクノロジー化は、「主体的な学び」を強調する教育学研究者から否定的に評価されがちですが、これこそが科学と民主主義理念の重要な接合点です。

[9] ISO (2015) 書誌情報は[1]を参照してください。

[10] ISO (2015)の“5.3.1 Translation”の項です。

[11] European Master's in Translation (2017) [2]を参照。 [3]で挙げた Piao, Yamada and Kageura (to appear)に ISO (2015)との対応表があります。

[12] Onishi, N. and Yamada, M. (2020) “Why translator competence in information searching matters: An empirical investigation into differences in searching behavior between professionals and novice translators,” *Invitation to Interpreting & Translation Studies*, 22, pp. 1-23; Onishi, N. and Yamada, M. (2021) “Search competence as a key skill for survival in a translation ecosystem,” *7th IATIS Conference*.

[13] Enríquez Raído, V. (2013) *Translation and Web Searching*. London: Routledge.

[14] Grammaly がほぼ使えない理由の一部はこの点に関わります。

[15] 工 藤 拓 (2019) https://twitter.com/taku910/status/1100035871183536128

[16] Feynman, R. (2000) *Pleasure of Finding Things Out*. London: Viking. 関連して、影浦峡 (2012)「『専門家』と『科学者』: 科学的知見の限界を前に」『科学』 82(1), pp. 56-62 も参照。

イベント報告

【JTF40 周年特別企画】2021 年度第 1 回 JTF 関西セミナー

機械翻訳とは何か？ どこから来て、どこへいくのか？

石岡 映子
JTF 常務理事
株式会社アスカコーポレーション

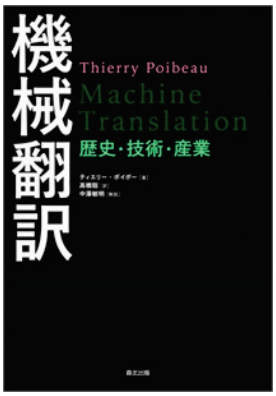
1. はじめに

2021 年 4 月 27 日（火）14:00～16:00、Zoom で開催された当イベントを報告する。一般社団法人日本翻訳連盟（JTF）40 周年を祝う最初のイベントとして、JTF 副会長でもある個人翻訳者の高橋聡氏と、東京大学大学院情報理工学系研究科客員研究員の中澤敏明氏にご登壇いただき、機械翻訳（MT）の進化を踏まえ、翻訳業界や翻訳者が MT とどう向き合い、付き合っていくのか、産業としての翻訳への MT の影響や意味、役割、今後の可能性についてお話を伺った。昨年出版された『機械翻訳:歴史・技術・産業』（森北出版社）に基づいて（セミナータイトルは、同書の帯から引用）翻訳された高橋氏と解説を執筆された中澤先生お二人のつながりでセミナーが実現した。

高橋氏からは、そもそも翻訳とは何か、翻訳者は今後 MT とどう向き合っていくのか、一部の仕事は MT に奪われていくという現実と、翻訳者の仕事の必要性と可能性など。

中澤氏からは、最新のニューラル機械翻訳の技術を、学習のデモを交えてわかりやすく説明いただいた。なぜ機械翻訳は間違えるのか、これから精度は上がっていくのか、など。お二人のご講演とパネルディスカッションの 3 部構成である。

2. 第1部　訳書を通じて個人翻訳者が考えた機械翻訳（高橋聡氏）

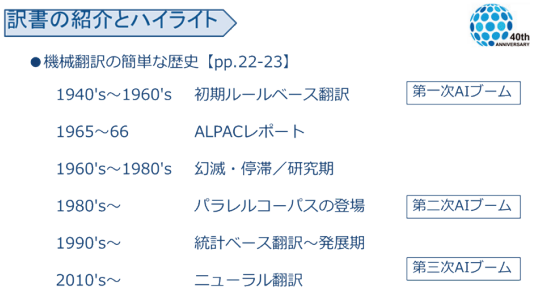


(1) 訳書の紹介とハイライト

第 1 章 はじめに
第 2 章 翻訳をめぐる諸問題
第 3 章 機械翻訳の歴史の概要
第 4 章 コンピューター登場以前
第 5 章 機械翻訳のはじまり
第 6 章 1966 年の ALPAC レポートとその影響
第 7 章 パラレルコーパスと文アラインメント
第 8 章 用例ベースの機械翻訳
第 9 章 統計的機械翻訳と単語アラインメント
第 10 章 セグメントベースの機械翻訳
第 11 章 統計的機械翻訳の課題と限界
第 12 章 ディープラーニングによる機械翻訳
第 13 章 機械翻訳の評価
第 14 章 産業としての機械翻訳
第 15 章 結論として：機械翻訳の未来

What is machine translation? Where does it come from, and where does it go?
Eiko Ishioka
Managing director, Japan Translation Federation
President, ASCA Corporation

解 説 2020 年時点でのニューラル機械翻訳（中澤敏明：東京大学特任講師）



(2) MT に対する多言語世界の要請

欧州議会内部では 2013 年時点で予算 3.3 億ユーロの翻訳の 93%が人力で行われているままで、増加する処理量と予算削減のためにさまざまな取り組みが進んでいるとのことである。カナダの天気予報では 1970 年から MT 訳が導入されており、信頼性も高いという事実は驚異的である。世界の特許市場での状況を考えると、多言語世界を知ることで、MT に対する要請の背景が見えてくる。2000 年代最初まで“日本は特殊”とされ、翻訳予算がついていたが、アジアの翻訳拠点は香港や上海に移り、今やグローバルの翻訳市場の一角でしかなくなり、特殊扱いは崩壊してしまった。

(3) MT に対する翻訳者の思い

翻訳者の MT に対するスタンスは人により異なる。MT の動向とは無縁に自分の翻訳を続けるという人がいるが、逆に、AI を積極的に導入したいという人もおり、大多数はその間で揺れている。

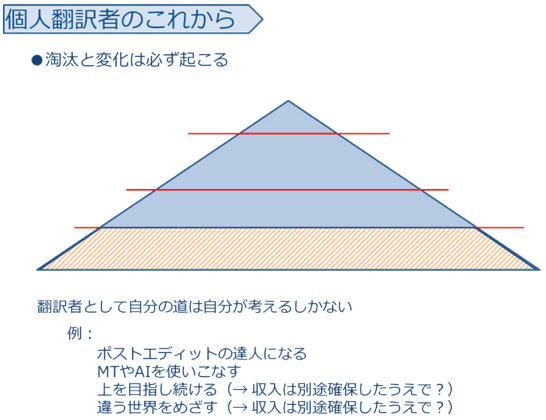
翻訳をする動機もさまざまである。好きだから、楽しいからという動機と、生活の手段として仕事にしているという動機があるが、大多数はその両端に振り切れるのではなく、中間にいる。このスタンスによって MT への考え方は異なる。

そもそも翻訳とは何か。人間の翻訳者は、何通りもの訳出パターンを頭に思い浮かべ、文種、文体、文脈などの条件に合わせて絞り込んでいく。MT に慣れたら何通りも翻訳案を考えることは必要なくなる。

情報としての翻訳は、TM 期を経て MT に移行していく。その中間にポストエディットが存在する。コンテンツとしての翻訳には、少なくとも当面は人間の翻訳が必要である。

(4) 個人翻訳者のこれから

淘汰と変化は必ず起こる。多くの人がやっている"裾野"の翻訳の仕事はなくなる可能性がある。翻訳者として自分の道は自分が考えるしかない。



どの道を選んでもそこに上下や貴賤はないが、進んだ先で見える世界はまったく違うものになる。それはどんな形で翻訳に関わりたいかによって決まるはずである。

(5) MT の扱われ方

検証を経ずに安易に MT を使ったり、売ったりすることは問題である。災害警報の誤訳の問題が過去に起こったが、ここでは「情報としての翻訳」としてすら十全に機能していない。これには、社会全体の取り組みが必要である。「MT なんて使いものにならない」ではなく、「使えるようにするにはどうすればいいのか」を考える、啓発することが必要で、翻訳者にも関わってほしい、と結んだ。

講演後の質疑応答では、ポストエディットの単価の妥当性、機械翻訳の台頭で翻訳者が搾取されているの

ではないか、などの質問が出た。一部の翻訳会社の「機械翻訳でコストダウンと納期短縮が可能！」といった売り方は場合にもよるとの回答。「機械翻訳は全然使えない。機械翻訳を使うと翻訳ができなくなる」などのSNSでの意見は極論だと思う、と答えた。

3. 第2部　機械翻訳の現状と課題、可能性（中澤敏明氏）

(1) NMTの訓練デモと最新の技術動向

リアルタイムでNMTの訓練デモが行われた。京都に関するデータが題材として使われた。対訳コーパス対して、Fairseq（Facebook AIが開発・公開しているフリーツール）を用い、SentencePieceでサブワードに分割した教師データで訓練。サブワードとは、未知語や複合語を単語と文字の中間的な単語に分割することで、あまりに低頻度な語を学習させることを回避し、語彙サイズを最適化する方法である。サブワードの手法にSentencePieceを使うことで、文から直接単語分割を学習することができるため、事前の手動での分割処理が不要になる。

例）

通常の単語分類

本山｜は｜、｜足利｜義満｜に｜より｜建立｜さ｜れた｜京都｜の｜相国｜寺

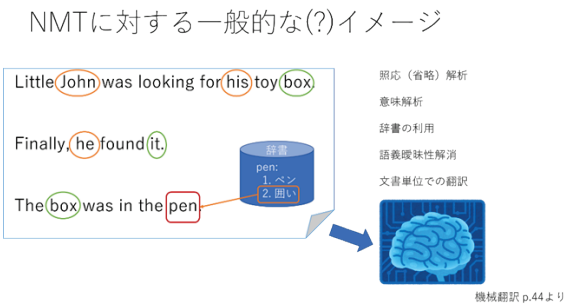
サブワード

本｜山｜は｜、｜足利義｜満｜により｜建立｜された｜京都｜の｜相｜国｜寺

（足利義〇の人がたくさんいることから、「足利義」までをひとまとまりとなる）

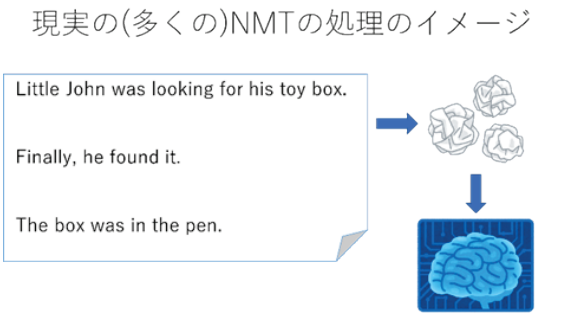
サブワードを用いると固有名詞や専門用語などが分割されるだけでなく、数字やURLなど翻訳が不要なものまでばらばらになってしまうため、一見翻訳する必要がないように見えるものであっても、誤訳が発生してしまう原因となってしまうという欠点がある。

次に、事前質問に対する解説をいただいた（一部抜粋）。



機械翻訳 p.44より

実際は、



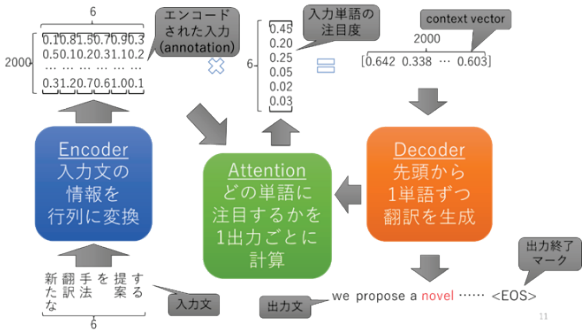
機械翻訳 P44 より

Q1：

辞書にはアノテーションが重要だと思われますが、多義語やアノテーション付きの辞書は存在するのでしょうか？

A：現在は使わない。

実際には、以下のような仕組みでNMTは処理される。



一般的なNMTの処理概要：スライド11P

NMTの特徴として、NMTでは、入力文を「置き換える」ことで翻訳するのではなく、入力文を見ながら、翻訳文を作り出すため、流暢な翻訳になる。また、ベクトル表現(word embedding)を使うおかげで柔軟な翻訳が可能である。一方で、統計的機械翻訳（SMT）のときのように単語単位での対訳をモデルに組み込む

ことは容易ではなく、翻訳完了マークが出力されるまで翻訳が続くので、入力文を過不足なく翻訳できていないことがある。また、全然違う訳が出てしまうことを防げないのも短所の一つ（例：I come from Tunisia。→ ノルウェーの出身です[Arthur et al., 2016]）。

Q2：
NMT の訳抜けの穴埋めをするために、ルールベース翻訳の結果と比較検証し、エラーを指摘するのはできないのか？
A：NMT では入力と出力の対応をとれないので、ルールベースの出力と合わせることは困難。

Q3：
良質の対訳データを組み込むために施策は取られているか？
A：既存の対訳コーパスの人手クリーニングもしくは自動クリーニング、Web クロールデータに対する文アライメントや翻訳メモリの蓄積など。

Q4：
文単位でなくドキュメントの文脈も加味して翻訳できる MT もあるようだが、最新の動向や、選ぶポイントを教えてほしい。
A： DeepL は照応・省略解析なども行っており、文単位ではなく複数文のまとまりでの翻訳が可能。Google 翻訳も最近、複数文のまとまりでの翻訳が可能になり、日進月歩なので実際に試すのが良い。

Q5：
敬語や口語の選択ができず、混ざった文章が出力させるのはなぜか？
A：訓練に使ったデータに両方が混じっているから。

Q6：人名や地名など、一般名詞と区別する技術はないのか？
A：語義曖昧性解消や固有表現抽出などの技術は独立して以前からあるが、NMT でそれらの技術を直接使うことはあまりなく、Transformer では入力文ごとに単語のベクトル表現を計算するので、文脈に応じてより適切に語義を推定することが可能。

Q7：
感情の読み取りが必要な小説などでは MT は使えないと言われているが、どうなのか？
A：数年以内にはできないが、10 年後の可能性はゼロではない。そもそも人間の翻訳でも「正しい」翻訳は存在しない？

(2) MT の現状と課題

できることは徐々に増えてきている。省略や照応解析の利用、文書単位の翻訳や資源が乏しい言語の翻訳、マルチモーダル翻訳など。
しかし課題はまだまだ山積。訳抜け・湧き出し、否定・肯定誤り、訳語統一、代名詞誤り、対訳辞書の利用、ドメインアダプテーション、翻訳速度など、さまざまな問題点を日々改善している。

(3) MT の可能性

深層学習の限界はまだよくわかっていないが、NMT が出た当初の 2014 年より成長スピードが落ちている印象がある。NMT は人間が一生かけて読む文書量よりもはるかに多くの文に触れているので、人間の翻訳より良い訳を出すこともある。しかし、いつもよいわけでないので、チェックが必要である。

人手が不要もしくは最低限でよいという翻訳の需要は必ず存在するし、その割合は多くなるはず。MT が活かせるところは積極的に活かすべきだと思う。翻訳されなかったものが翻訳されるようになり、仕事を奪うのではなくサポートとして活用し、翻訳全体の生産性を向上するものになってほしい。

4. 第 3 部　パネルディスカッション「機械翻訳とは何か、どこへいくのか？」

モデレーター：石岡映子氏（JTF 常務理事・関西委員長、株式会社アスカコーポレーション代表取締役）

石岡：弊社のクライアント対象のアンケートでは、8 割の企業が MT 導入済で、残りの 2 割の半分が検討中であった。JTF の最新の白書でも特許・医薬・工業が収入減、現場に MT が導入されたためと思われる。書籍のように人がやらないといけないところは伸びている。このような環境下でいくつかの課題を伺いたい。

まず、公共機関などでの間違った MT 表記についての問題はどうか？

高橋：社会全体で考えないといけない問題である。気になるのは、誤訳の情報を発信した後、当の公共機関がその後どうしたか、なぜか報道がない。反省し、改善しないといけないと思う。それには翻訳業界、JTF のような業界団体が先陣を切ってやっていくべきではないか？

石岡：受け止める側の観点はどうか？

中澤：オンラインのフリーソフトは自己責任において使用するのが普通であり、使用によって生じた損害は保障されない。そこに品質を求めることはナンセンスである。そういう教育を受けていないがためにリテラシーが低いことが問題であり、子どものころから教育するべきだ。また、フリーのものをどう使うかも考えるべき。

石岡：翻訳という仕事はなくなるのか。NMT になって導入が進み、実際収入減となっている。今後どうなっていくのか？

高橋：なくなっていく部分はある。その部分を担っていた人がこの先どうするかは、PE に移行する、人手でしかできない分野に挑戦するなど。全部なくなるわけではないが一部は確実になくなっていく。その前提で考えるべき。

石岡：ＭＴをハンドリングする方面の人材不足も感じている。

中澤:翻訳が細分化されると、新たな需要が出てくる。コーパス、辞書の整理、これらも MT で使いやすい形にするにはリテラシー、プログラミングなど、これまでの翻訳者に求められるものとは異なった新たな職業ができる。純粋な翻訳者は減っても新たな関連の仕事が出てくる。

高橋：翻訳に関わるデータ整備の仕事については、翻訳メモリが出現した時代から整理の必要があった。当初はやっていたのが、この 20 年、ほぼ放置状態。チューニングしていれば、翻訳メモリの精度も上がっていたはずである。その反省から、翻訳やレビューもできるリンギスト、いろんな方面に目を配り、レビュー結果をもとにクライアントと交渉できる人が求められる。翻訳者とは違うかもしれないが重要な仕事である。

石岡：業界は、クライアントから受け取った TM が適切でなかったとしても指摘してこなかった。TM の管理をできる人材は翻訳会社にとって欲しい人材である。翻訳者やチェッカーに、新しいキャリアパスという観点で挑戦してほしい。

石岡：グローバルの観点から、翻訳は未来への投資として重視されてきて、品質への期待は高い。一方、情報を早く出さなければならない時代、品質をどう考えたらいいか？

高橋：投資ならリターンが必要、リターンが見えづらいから企業も投資しない。翻訳者に十分な情報を提供せず、だから望むものが返ってこない。翻訳者の力不足でかたづけるのではなく、顧客と翻訳会社と翻訳者が品質を合意し、そこに結果が出れば投資につながるのでは。

中澤：そのとおり。マニュアルで売り上げが増えたかは計算できないし、将来がわからないからコストをかけられない。特にベンチャーなどでは予算もなく、製品のバージョンアップサイクルが早くて予算が取れな

いから機械翻訳になる。よりシビアにクライアントが必要とする品質をすりあわせコントロールすることへの LSP の役割は大きい。

石岡：目の前の納期ありきで、クライアントと必要な精度、品質をすりあわせていないことが多い。三者が合意をすることで、翻訳者、翻訳会社は商品として品質を担保し、クレームも減る。MT が浸透する中、改めて品質についてすりあわせするべき。

中澤：低い品質、安い見積もりにつけこむ人が出てくると、業界の首を絞めるので、適切に取り組むべきである。

質疑応答：

Q. 自分の翻訳メモリのような小規模データセットで、既成のエンジンの出力はどれくらい変化するのか。

中澤：数万文、十万文ぐらいないと変わらない、数千文でもニッチな業界ですべてを網羅しているのならあり。翻訳したい文のバリエーションによる。特許文全体を翻訳したいのに工業のコーパスしか持っていないなら、自分でバランスをとってみてもらうしかない。

高橋：TM 登場のときからそういう議論があって、メモリの扱いには注意が必要なのに、似たことが繰り返されそうな感じがする

Q. 将棋や囲碁では AI が人間に優っているのに、翻訳では MT が人間を越えられないのはなぜか？

中澤：将棋や囲碁は正解があるが、翻訳は明快なゴールがないから、ゴールが決まっているほうがやりやすい

高橋：文芸翻訳のような正解のない翻訳で、MT が初音ミクのように、1 つの個性になったらおもしろい。

中澤：そういう研究もある。たとえば太宰治風とか。人格を持たせるのはおもしろい。

5. 終わりに

今回のセミナーで MT と翻訳のことが結論付いたわけではない。継続的に正しい情報をもとに議論していくことが重要である。そこで中澤氏は、翻訳者と機械翻訳（の研究者）がお互いをもっと理解するための「翻訳と機械翻訳の座談会」という YouTube チャンネルを開設された。

https://www.youtube.com/channel/UC4fiKKrfcvQY1dcZkjxfnHQ

今後はこのチャンネルで、翻訳と機械翻訳の未来を語り合いたい。

参考資料

- Thierry Poibeau『機械翻訳：歴史・技術・産業』(森下出版)
  原題：“Machine Translation”（The MIT Press Essential Knowledge series）
- 機械翻訳はどこに行くのか？『機械翻訳：歴史・技術・産業』山田優（AAMT Journal No.74）

イベント報告

AAMT 招待講演『語学教育と MT』機械翻訳の問題 ～第二言語教育の立場から～

講演者：東京大学大学院総合文化研究科・教養学部教授トム・ガリー氏

出内 将夫

情報通信研究機構

1. はじめに

2021 年 6 月 23 日に開催されたアジア太平洋機械翻訳協会 第 2 回定時社員総会が開催されました。第 2 回定時社員総会は、新型コロナウィルスの感染拡大防止のため、オンライン開催となりました。参加者は約 100 名でした。東京大学のトム・ガリー教授から、言語教育への MT の影響に焦点を当てた招待講演がありましたので、以下に報告します。講演資料は https://aamt.info/wp-content/uploads/2021/06/AAMT_30th_PDF_Gally.pdf からアクセスできますので、ぜひ、これらの資料をご参照ください。

2. 招待講演

以下の 1 件の招待講演がありました。

◎『語学教育と MT』機械翻訳の問題 ～第二言語教育の立場から～：

トム・ガリー(Tom Gally)／東京大学大学院総合文化研究科・教養学部

前述の講演資料をもとに講演内容を次にまとめます。

◎『語学教育と MT』機械翻訳の問題 ～第二言語教育の立場から～：トム・ガリー(Tom Gally)氏

本日は機械翻訳(MT)関係者以外に、翻訳会社関係者、翻訳者、言語教育者の方なども参加されているということで、それを念頭に説明します。私は MT の専門家ではないので、MT の技術や現時点の精度については扱いません。

私は 1957 年に米国カリフォルニアで生まれ、大学と大学院で言語学と数学を専攻し、外国語としてロシア語と中国語を勉強しました。1983 年に来日し、日本語を勉強しました。1986 年にフリーランス翻訳者として日本語から英語の翻訳に携わりました。辞書の出版社からお声がけいただき、和英辞書や英和辞書の編集にも携わりました。2005 年から東京大学で英語教育にかかわる仕事に携わっています。学部生に向けた英語アカデミック・ライティングプログラムの開発と運営、大学院では第二言語教育に関する研究指導などを行ってきました。

AAMT にとっては失礼な話かもしれませんが、私は長年、実際の用途においては MT の利用が無理だと思っていました。人間の翻訳者は、頭の中で元の言語の文章の意味を理解して、翻訳先の言語で同じ意味の文を作成することを目指しますが、コンピュータは我々が理解する意味を扱えないと考えています。しかし 2016 年秋に、ニューラル機械翻訳(NMT)を応用した技術が Google から公開され、それを試して驚きました。それ以前の MT の出力は、主語や動詞の有無の判別すら難しい文がありましたが、正しい文の構造の持つ翻訳ができるようになっていました。その時から、私は MT と言語教育の問題を考えるようになりました。以下に、その問題について自分が思い付いた順番に説明します。

一つ目は「大学の英語アカデミック・ライティング授業を教える教員は、どのように学生の MT 利用に対

AAMT Invited Talk "Language Education and MT" Machine Translation Problems - From the Viewpoint of Second Language Education - Lecturer: Tom Gally, Professor, Graduate School of Arts and Sciences, College of Arts and Sciences, The University of Tokyo
Masao Ideuchi
NICT

応すべきか？」という問題です。英語教育プログラムでは、学生が自分たちで辞書を引き、英語で論文を書くことを想定していました。しかし、学生たちが日本語で文章を書き、MT で翻訳するケースが出てきました。授業の目的は、英語で論文を書くことだけではありません。研究者が MT を利用して論文を書くことは問題無いと考えますが、言語教育として MT の利用を許容するのかは別の問題です。

二つ目は「理科生がゼミの課題を遂行するために科学論文から情報を収集する必要があるときに、その学生に MT の利用を推薦すべきか？」という問題です。英語科目の先生は、英語文献を読む際は英語で読むべきだ、と考えていますが、研究を指導している先生が、MT の積極的な活用を推奨していた、というケースがありました。

三つ目は「企業や官公庁に就職する予定のある大学生にビジネス英語などを教えるときに MT の利用法を積極的に指導すべきか？」という問題です。大学で英語を教える目的の一つとして、学生が教わった内容を将来役立てられる、というものがあります。そのために MT のツールの使い方や工夫を教えた方が良い、という意見と、ツールを使うと英語ができるようにならないため避けた方が良い、という意見があります。

四つ目は「MT 利用は言語習得にどのように影響するか？」という問題です。学習者が MT を利用した場合、言語習得に悪影響があるのでは、という意見があります。例えば、30～40 年前に電卓が世に出た時、計算ができなくなるのではないか、また、ワープロが世に出た時に漢字を書けなくなるのではないか、という意見があったことと似ています。MT を効果的に使えば、学習に役立つのではないか、という意見もあります。現時点では、おそらく一部の学習者に良い影響、一部の学習者に悪い影響が出ていて、まだ良くわかっていない状況です。

五つ目は「実社会では MT がどのように使われているか？」という問題です。英語教育プログラムでは、三つ目の問題で述べたとおり、学生の将来を考慮します。学生が卒業後に実社会で MT を使う可能性があるため、MT がどのように使われているかを調査しようとしています。いろいろな人に聞き、Web の情報を調べていますが、今のところ、よく見えていません。

六つ目は「共通言語を持たない人たちが MT を使ったら、どのくらいコミュニケーションや協力をできるか？」という問題です。実用において、特にやり取りがある場合、MT は 100%の精度を求められません。私が 2 年前に見かけた例で、外国人観光客と日本人学生が、電車の中で MT のアプリを使って日本語と英語でコミュニケーションをしていた場面がありました。一部のコミュニケーションは、音声認識や翻訳が失敗して成り立たないことがありましたが、「どの駅で降りれば良いか」「どの国から来たか」「何を勉強しているのか」といったコミュニケーションには成功していました。しかし、こういった短い会話だけでなく、国際的な企業において様々な国出身の従業員の間で共通言語が無い場合に、MT で業務がうまく行くのかまでは、明らかになっていません。

七つ目は「中学生から『翻訳アプリがあるから英語を勉強しなくてもいいのでは？』と言われたらどのように返事すべきか？」という問題です。日本ではすべての子供たちに英語教育を行っています。小・中学校の先生によると、5 年前からこの質問がよく出ているとのことです。どのように返事をするか難しい問題です。

八つ目は「これからはすべての子供たちに英語を教える意味が果たしてあるのか？」という問題です。これは一番大きな問題であり、すぐに答えが出せません。現在、日本では義務教育において、すべての子供たちに外国語を教えています。学習指導要領で必修となったのは 2002 年ですが、実質的には 1970 年代からすべての中学校で英語を教えています。昨年から小学校でも教科として英語の授業が始まりました。MT でコミュニケーションできるようになった場合、英語教育の必要があるのか、という問題が生じます。

これらの問題を議論するために、言語習得

(language acquisition)、言語学習(language study)、言語教育(language education)の基本概念を整理します。言語習得は、母語でも第二言語でも、ある言語の利用が自動化すること、つまり考えずに話したり聞いたり書いたり読んだりできる状態になることを指します。意識的に文法などの規則を学び語彙を暗記しても、発話できない、あるいは発話できても意味を理解できていなければ、言語習得とは言えません。

言語学習は、一人で外国語を学ぶことも含む概念です。言語学習においては、MT のツールが役立つか悪影響を生むか、個人個人の判断に依ります。言語学習では適宜 MT のツールを使うのが良いと考えています。

一方、言語教育は特定の機関・組織・学校で行う言語に関する教育や学習のことです。本日の話では言語教育を対象とします。

言語教育についての様々な論争では、言語がどういうものか、という認識が異なっている場合があります。言語は、脳の中で処理をするような認知に関する営みであるという捉え方と、やり取りを伴い文化や社会的な現象との関係がある営みであるという捉え方の、二つの言語観があります。言語自体はその両面を持っており、どちらの考え方も正しいのですが、言語教育ではどちらの言語観を重視するのかという問題があります。今朝の新聞の一面記事に、文部科学省の有識者会議から、大学入試に英語の民間試験を導入することは困難である、という提言案が示され、夏には導入中止が決定される見込みである、というものがありました。民間試験導入については、数年前から様々な場所で議論があり、昨年の夏には日本学術会議からも提言がありました。

言語観の違いを、大学入試での試験問題を例に説明します。2020 年までのセンター試験英語科目の第 1 問目で問われていたのは、発音やアクセントなど、認知的な言語観に基づく問題となっていました。一方、2021 年に共通テストとなった英語科目の第 1 問目では、テキストのやり取りを読んで、その内容について問われる、社会的な言語観に基づく問題となっていました。センター試験から共通テストに変わったことで、問われる言語観が変わり、過去のテストを元に勉強してきた受験生は困ったと思われます。

言語教育を考える場合に、言語教育がどのような場面で行われるか、といった点も重要です。小学校、中学校、高校、大学（短大・専門学校を含む）、塾・予備校、語学学校について考えます。小学校から大学までは、国の法律に基づいて設立されており、国の規制や管理の影響を受けます。発表資料上で高校と大学の色を薄くしたのは、高校、大学と進むにつれ、その影響は弱くなるためです。小・中学校は、義務教育でもあるので、国の規制や管理を強く受けます。国と言っているのは、政府や文部科学省だけではなく、日本の中で英語を教えるべきというコンセンサスがあり、それが規制に含まれていると考えています。小・中学校では、教員免許があり、学習指導要領で細かく教える内容が決められており、教科書は検定というプロセスを経る必要があります。高校になると、教員免許や学習指導要領があるものの、規制は緩やかになります。大学では、外国語教育に関して、それぞれの大学でかなり自由に教育内容を決めることができます。もし公立の中学で、積極的に MT を教えようとして、検定教科書を使わずに MT の使い方を教えたら、問題になると思います。しかし、大学ではやろうと思えばできるのではないかと思います。ただし、大学で新しい学部を設立する際、MT があるからこの学部では英語を全く教えません、とした場合、設立が認められない可能性はあるかもしれません。塾・予備校と語学学校では、ほぼ自由で何を教えても構いません。語学学校では、学習者は顧客でもあるので、顧客が MT の使い方を学びたいというのであれば、誰も問題視しないことになります。

言語教育の目標も言語教育を考える上で重要です。様々な議論がありますが、言語教育の目標を要約すると、実用、教養、試験・資格の 3 つに大別されます。実用は、将来仕事で使えるようにという考え方です。教養は、外国語を学ぶことで外国の文化などに関する知識を得る、母語との対比で母語についても理解を深められる、外国語の文法などを学ぶことで頭の訓練になる、という考え方です。試験・資格は、大学入試や資格取得を指します。

学習者側の中には試験のために勉強するという人もいますが、国や教育者側は試験で、学習者の語学における実用や教養のレベルを測ろうとします。

言語教育の最大の問題は、学習者のモチベーションです。個人で学習するときも学校などで学ぶときも同様です。外国語の習得には時間がかかります。こつこつ続けて勉強するのは大変であるため、何をきっかけに言語の勉強を始め、どう継続して勉強するか、という問題があります。高校までは、進学するための試験に必要というモチベーションがありますが、大学の言語教育では、試験のモチベーションが無くなるため、特にこの問題の重要さを感じます。

言語教育にとっての MT の特徴を考えます。まず、使いやすく無料、といった特徴があります。30 年前であれば、メインフレームのコンピュータにアクセスできる企業のみが MT を扱える状況だったことから、MT が言語教育の現場に及ぼす影響は無かったでしょう。NMT が出てきたとき、スマートフォンなどの便利なデバイスが広く普及しており、すぐに教育の現場でも対応を決める必要がありました。別の特徴として、正確に見える、という特徴があります。学習済みの言語間で翻訳に MT を使った場合、間違いがあれば気づけますが、知らない言語への翻訳に MT を使うと、間違っていても気づけません。学習者が MT を使った場合、MT の間違いに気づけずに完璧だと思ってしまう、という問題があります。もう一つの特徴に、人間なら何年間も勉強して、やっと取得できる能力を持つように見える、というものがあります。私は過去に韓国で講演をしたことがあり、年数回韓国語で書かれたメールが来ます。私はハングルの文字が読めないので、まったく意味が分かりません。NMT が無いときは、韓国語が分かる学生に尋ね、自分に何かを依頼するメールではなく、シンポジウム開催の告知メールであることを確認していました。NMT が実用化された後は、多少不自然な訳はありましたが、NMT を利用して自分で内容を確認することができるようになりました。もし私がハングルで書かれたメールを自分で読めるようになるまで韓国語を学ぶとしたら、何年もかかると思われます。

これまでに挙げた問題を整理するために、言語教育に関係する言語観、場面、目的の要素をマトリックスに示しました。このマトリックスを用いて、問題を整理していきます。

一つ目のケースでは、大学受験を目指す高校生について考えます。彼らの目的は試験であり、高校や塾・予備校などの場面で、語彙を暗記したり文法を学んだりと認知的な言語観に沿った勉強をします。言語学習に MT を活用している学習者はいるかもしれませんが、私はまだそういった事例の報告は聞いたことがありません。今後も試験における MT の利用は認められないと考えられるため、このケースでは MT が役立つ機会は少ないと思われます。

二つ目のケースでは、英会話を学ぶ会社員について考えます。私は語学学校に勤めたことがありますが、会社員の方は商談や海外出張で会話できるようになるために学んでいることが多いです。実用が目的なので、言語観は社会的な側面が重視されます。会話ができるようになるには、会話そのものを練習する必要があるので、このケースでも MT が役立つ機会は限られると思われます。遠隔での会話では MT が使えるかもしれませんが、対面の会話で MT を使うと信用が得られない可能性があります。

三つ目のケースとして、語学学校でのシニア向けの文化講座を考えます。彼らの目的は教養で、言語観としては学ぶ国がどういう国であるかといった社会的な面が重視されます。このケースでは、その国の言語で書かれた Web ページや文書を、MT で翻訳して調べることにより、日本語で書かれた情報を調べるよりもその国のことを知れるため、MT は役立つと思います。

四つ目のケースは、国立大学の必修授業です。重視する言語観や目的は統一されているわけではなく、授業ごとに異なります。文法・語彙・読解力など認知的な側面を重視する授業、コミュニケーションを重視する授業の双方が提供されていることが多いです。この

ケースでは、コンセンサスが無いためMTの活用を議論することは難しいです。一部の私立大学では、学生が将来英語にかかわる仕事に就かないことが多く、英語を必修の授業としているものの、授業の目的・言語観が曖昧となっているケースもあると聞きます。

小・中学校の外国語授業についても大学と同様にコンセンサスがありません。学習指導要領や検定教科書を読むと、言語観について近年は社会的な面が重視されてきていますが、初めて学ぶ言語の音素やスペルを紹介する必要もあるため、認知的・社会的の両面を含んでいます。子供たち・保護者・先生のいずれもコンセンサスに至っていないのではないかと思います。

結論にならない結論となりますが、次のことが分からない限り、言語教育がMTにどのように対応すべきか分かるようにならないと考えています。一つ目はMTの社会的な実用性、二つ目はMTが言語習得への影響、三つ目は学習者のモチベーションへの影響です。学習者のモチベーションへの影響とは、MTがあるから外国語を勉強しなくても良い、となるのか、MTによって外国のことが分かるようになり、もっと学びたくなるのか、ということです。実用性、言語習得への影響、学習者のモチベーションへの影響は、第二言語教育専門家や応用言語学者が5年間あるいは10年間研究すれば、分かってくるのではないかと思います。あるいは学習者や教育者が実際にMTを試すことで分かる部分もあるため、今でも少しずつ分かってきている部分はあると思います。言語観や英語教育の目的に関するコンセンサスも必要です。コンセンサスについては、日本では長い間議論が続いている話であり、難しい問題です。問題が問題のままこの講演は終わりますが、皆様からの質疑を楽しみにしています。

◎質疑

質疑は、以下のとおりです。

隅田英一郎氏：

今日の話は日本人の義務教育で英語を教えるべきか、という話が重要な部分でしたが、外国の人が日本語を勉強することを考えると別の状況になりませんか。例えば、アメリカのビジネスマンがMTを使ってどんどん日本人と仕事の話をしたい場合に、どんな教育をすべきか、という話だとどうなりますか。

トム・ガリー氏：

その状況ではアメリカのビジネスマンが日本語を学ぶことになると思いますが、教育ではなく学習の範疇になると思います。個人の学習ではMTをどんどん使えばよいと思います。また、実際にMTを自習で使ったことがある人は、それを発信してほしいと思います。一つ例を挙げると、先日会った50代の日本人女性は、英会話の勉強に熱心で、英会話のレッスンに参加する前に、レッスンで話したいことを事前にMTで翻訳して覚えておく、と言っていました。これは従来の方法ではできなかったMTの効果的な利用方法だと思います。教育では、管理・統一して全員が勉強することになるため、勉強のペースを合わせ、テストも共通にする必要があるため、かなり大きな問題になると考えています。

会場：

小学校の義務教育において英語を教育するのは文部科学省を含めた国だけの判断ではなく、国民が英語を教えるべきだと考える風潮があって、それが影響しているとのお話がありました。だとすると文部科学省に影響をあたえるきっかけは何でしょうか。何らかの報告書でしょうか。それとも調査が実施されているのでしょうか。ご存じでしたらご教示ください。

トム・ガリー氏：

私が指導している大学院生がそのテーマで論文を書いています。文部科学省の中にさまざまな審議会があり、研究者、企業、政治家などいろいろな参加者がいて、提言がまとめられます。詳しくは分かりませんが、かなり複雑な要因が絡んでいるという認識です。その中でも、日本国民の中では英語を勉強すべき、というコンセンサスが強いと思われます。義務教育では外国語を教えることになっていますが、学習指導要領を読

むと、ほとんどの内容は英語についての記載です。英語以外の外国語を教える中学校は全国に数校しかありません。日本国内で外国語を使う機会があるとしたら、観光客などの統計を見るとアジアの国々の言葉を使う機会が多そうに見えます。それでも外国語として英語だけを教えているのは、日本国民の中でも外国語は英語である、という考え方が強いと考えています。

会場：

たとえば、ライティングに活用するなどの「MT を使いこなすための教育」についてどう思われますか。ただし、MT を使いこなすためには、高度な語学能力が必要になると思います。この点について、どう思われますか。ガリー先生の教育の経験も含め教えてください。

トム・ガリー氏：

最初に私がこの問題を発信したのは 2017 年のことだったと思います。隅田先生をお招きして MT を紹介いただいたシンポジウムで、テーマは「機械翻訳と第二言語ライティング」だったと記憶しています。教育では学生を採点する必要があるので MT の活用が難しい面があります。しかし理系の研究所で英語論文を書く場面を考えると、MT の使い方を教えることは効果的で、良い論文が書けるようになると考えています。例を挙げると、まず日本語で書いた文を英語に翻訳して英語の意味を確認するだけでなく、その英語を日本語に翻訳して日本語の意味を確認することを教えたり、専門用語の扱いを教えたり、MT を一度だけ使うのではなく何度も繰り返し使うことを教えたり、日本語のすべての文に主語を補うことを教えたりすることが有効だと考えています。翻訳会社でプリエディットとして行うことに近いので、プリエディットのコツ、ポストエディットのコツを教えることは MT を使う上でも有用です。ただしこれは完全に実用を考えた場合です。教育では、実用以外に教養も考える必要があり、言語観や目的にコンセンサスがありません。同じ職場・学校でも言語観や目的が異なっていることや、学習者間でも異なっていることがあります。私は先日、現在教えている学部一・二年生向けの「オンライン・ディスカッション」という授業で、MT を使っていますか、という匿名のアンケートをしたところ、学生の半分くらいから MT を使っているという回答がありました。MT を使っていない学生に理由を尋ねたところ、授業なので自分の力のみで行うべき、MT を使うのは不正である、といった回答がありました。教育の場面で MT を使うことの難しさはこのような部分にあると思います。

会場：

MT を利用すれば、外国に関する情報収集は容易になると思いますが、そのような MT の得意なことを外国語教育に応用することは考えられるでしょうか。

トム・ガリー氏：

これは外国語教育の目的に関する問題だと思います。数年前に知り合った教育者と、この問題について議論したことがありますが、小・中学校の先生が MT を積極的に活用して外国の様子を調べる授業を行えると考えているようです。教室でパソコンが使えれば、様々な国の子供たちの遊び、食事、服装などについて、その国々の人々が書いた Web サイトを調べることができます。MT を使って外国について知ることは、社会科の授業としてとても面白い授業になると思います。しかし、語学の授業として、どこまで取り入れるかは別の問題です。MT で外国について知る授業を体験した後に、語彙、文法、会話の授業を行うと、なぜ MT があるのに難しい英単語の勉強をするのか、という質問がより多く出ると思われるため、難しい問題だと思います。

会場：

学生のモチベーションや、コミュニケーションか、研究者養成かといった、大学の授業の目標設定に関しては、英語よりも必修の第二外国語の方が、課題が大きいかと思います。いろいろな先生と話す機会があると思いますが、第二外国語の先生方から MT をめぐる苦労話や利点は聞いてらっしゃいますか。

トム・ガリー氏：

東大では、すべての学生が第二外国語を学ぶ必要があります。入学した後、ゼロからフランス語、中国語、韓国朝鮮語、ロシア語など、今までに習ったことが無い言語を学びます。東大の教養課程では広く勉強させるという理念があるため、これからも第二外国語を必修とすると思います。学んだことのない言語では MT は使いにくいのではないかと思います。少なくとも、私が聞いたことがある範囲では、初級レベルの第二外国語教育では MT はまだ大きな問題と見做されていないようです。

会場：

MT を英語ライティングに応用すると、これまでの単語の使い方、文法の正しさ中心の作文教育から、アイデア・論理の展開などへの教育に有用ではないでしょうか。単語や文法の教育を軽視するわけではありませんが、それを超えた世界があるような気がします。

トム・ガリー氏：

東大の学部一年生向けのアカデミック・ライティングの授業を担当する教員の間で、MT に関する問題も含めて話し合っています。授業で重視しているのは、アイデアや論理の展開、読者の理解を考えた表現、先行研究の紹介の仕方などです。学生は大学に入るまでに語彙や文法を勉強してきているので、この授業では語彙や文法はそこまで重視していません。英語でまとまった文章を書くための技術の方が重要なので、MT を使って良いのではないか、と考える教員もいます。しかし、大学としては言語習得も重要であるため MT を許容すべきでないという意見もあり、コンセンサスに至っていません。

AI の発展で、自動運転や AI 倫理の問題が発生していることと同様に、MT を言語教育の世界でどうするか、という問題について議論が必要だと考えています。この問題に社会としてコンセンサスが出せるように、皆様も是非考えてみてください。

講演内に回答できなかった質問と回答については、https://aamt.info/aamt-soukai2021-seminar/ に掲載されていますので、ご参照ください。

3. おわりに

本講演は、オンライン開催となりましたが、多くの質疑があり活発な議論が行われました。言語教育における MT に関わる様々な問題と、それを考える上で必要となる観点を紹介いただき、社会におけるコンセンサスの重要性について説明いただきました。最後に、報告者の感想を記載します。講演で言及があった計算機を例に考えると、実用分野の資格試験では電卓使用可の試験も多いようで、MT によって外国語教育の応用や実用に近い部分は影響を受ける可能性があると感じました。一方で、電卓があっても算数や数学を学ぶ必要があるのと同様に、外国語教育の基礎は変わらない部分も多いと感じました。小・中学校の教育で ICT 活用が進んでいるため、MT を若いうちから当たり前に使う人が増えることで、社会のコンセンサスが生まれる日も近いのではないか、と感じました。

イベント報告

第 8 回アジア翻訳ワークショップ（WAT2021）開催報告

中澤 敏明
東京大学

1. はじめに

本稿では WAT2021[1]の開催報告を行う。アジア翻訳ワークショップ（Workshop on Asian Translation, WAT）はアジア言語を中心とした評価型機械翻訳ワークショップであり、2014 年に第 1 回（WAT2014）を開催して以降、毎年開催している。本稿の著者は初回からオーガナイザーの一人としてワークショップの運営を行っている。2016 年の第 3 回（WAT2016）以降は自然言語処理の国際会議との併設ワークショップとして開催しており、2021 年の第 8 回（WAT2021[1]）はタイのバンコクで開催された ACL-IJCNLP 2021 の併設ワークショップとして、2021 年 8 月 6 日にオンラインで開催された。

ワークショップは様々な機械翻訳の評価タスクの実施に加えて機械翻訳に関する研究論文の募集も行っており、WAT2021 では 5 件の研究論文と、ACL-IJCNLP2021 から 2 件の findings 論文を採択した。採択した研究論文のタイトルを以下に示す。

・BTS: Back TranScription for Speech-to-Text Post-Processor using Text-to-Speech-to-Text

・ Zero-pronoun Data Augmentation for Japanese-to-English Translation

・Evaluation Scheme of Focal Translation for Japanese Partially Amended Statutes

・Optimal Word Segmentation for Neural Machine Translation into Dravidian Languages

・Itihasa: A large-scale corpus for Sanskrit to English translation

また 2 件の招待講演も行われた。1 件目は Facebook AI の Francisco Guzmán 氏および Angela Fan 氏により“Massively Multilingual Translation and Evaluation”というタイトルで行われ、2 件目はカーネギーメロン大学の Graham Neubig 氏により“Understanding and Improving Context Usage in Context-aware Translation”というタイトルで行われた。

WAT2021 では 18 の言語を対象とした 14 の shared task が行われ、世界中から 26 のチームが参加した。表 1 に参加者のリストを示す。なお表中でアスタリスクがついているチームは、構築したシステムの解説論文を提出しなかったことを表している。

Team ID	Organization	Country
TMU	Tokyo Metropolitan University	Japan
NTT	NTT Corporation	Japan
NICT-2	NICT	Japan
NICT-5	NICT	Japan
NLPHut	Idiap Research Institute Switzerland, IIT BHU, BITS Pilani India, KIIT University India, Silicon Techlab pvt. Ltd India, University of Chicago	Switzerland, India, USA
TMEKU	Tokyo Metropolitan University, Ehime University, Kyoto University	Japan
*goodjob	Dalian University of Technology	China
YCC-MT1	University of Technology (Yatanarpon Cyber City)	Myanmar
YCC-MT2	University of Technology (Yatanarpon Cyber City)	Myanmar
NECTEC	National Electronics and Computer Technology Center (NECTEC)	Thailand
mcairt	CAIR	India
nictrb	NICT	Japan
sakura	Rakuten Institute of Technology Singapore, Rakuten Asia.	Singapore
IIT-H	International Institue of Information Technology	India
*gauvar	Amazon	Singapore
*JBJBJB	Indivisual participant	Korea
SRPOL	Samsung R&D Poland	Poland
NHK	NHK	Japan
CFILT	Computing for Indian Language Technology	India
iitp	IIT Patna	India
Volta	International Institute of Information Technology Hyderabad	India
coastal	University of Copenhagen	Denmark
CFILT-IITB	Indian Institute of Technology Bombay	India
CNLP-NITS-PP	NIT Silchar	India
Bering Lab	Bering Lab	South Korea
tpt_wat	Transperfect Translations	USA

表 1：WAT2021 Shared Task 参加者

昨年度に引き続き、海外はインドからの参加者が多かった。インド諸語を対象とした翻訳タスクにおいては扱う言語の数が年々増えており、インドの研究者を中心として活発に研究が進んでいるためであると思われる。他にも中国、韓国、シンガポール、タイ、ミャンマー、スイス、デンマークなどからの参加があり、国際化がより進んだワークショップとなっている。

WAT2021 で新たに追加されたタスクは以下の通りである：

・英語からマラヤーラム語（インドで話されている言語）へのマルチモーダル翻訳

Report of the 8th Workshop on Asian Translation (WAT2021)
Toshiaki Nakazawa
The University of Tokyo

・曖昧性のある動詞を対象とした日英のマルチモーダル翻訳

・10 言語のインド諸語と英語間のマルチリンガル翻訳

・使用するフレーズが指定された日英の制限翻訳タスク

表 2 に各タスクの参加チームリストおよび人手評価を行ったタスク（表中の human eval にチェックがついているタスク）を示す。

Team ID	ASPEC + ParaNatCom EJ	ASPEC Restricted EJ	ASPEC Restricted JE	ALT + UCSY En-My	ALT + UCSY My-En	NICT-SAP En-Hi/Id/Ms/Th IT	NICT-SAP En-Hi/Id/Ms/Th Wikinews	NICT-SAP Hi/Id/Ms/Th-En IT	NICT-SAP Hi/Id/Ms/Th-En Wikinews
TMU			✓						
NTT		✓	✓						
NICT-2						✓	✓	✓	✓
goodjob	✓								
YCC-MT1				✓					
YCC-MT2				✓					
NECTEC					✓				
nictrb		✓	✓						
sakura				✓	✓	✓	✓	✓	✓
NHK		✓	✓						
human eval	✓	✓	✓			✓		✓	

Team ID	JPC EJ	JPC JE	JPC CJ	JPC JC	JPC KJ	JPC JK	Multimodal En-Hi TX	Multimodal En-Hi HI	Multimodal En-Hi MM	Multimodal En-MI TX	Multimodal En-MI HI	Multimodal Flickr EJ	Multimodal Flickr JE	Multimodal MS COCO EJ
TMU	✓	✓			✓	✓								
NLPHut							✓	✓		✓	✓			
TMEKU												✓		✓
sakura												✓	✓	
iitp									✓					
Volta							✓		✓					
CNLP-NITS-PP							✓		✓					
Bering Lab	✓	✓	✓	✓	✓	✓								
tpt_wat	✓	✓	✓	✓	✓	✓								
human eval	✓	✓					✓	✓	✓	✓	✓	✓	✓	✓

Team ID	Indic21 En-X Bn	Kn	Ml	Mr	Or	Hi	Gu	Pa	Ta	Te	Indic21 X-En Bn	Kn	Ml	Mr	Or	Hi	Gu	Pa	Ta	Te
NICT-5	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
NLPHut	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
mcairt	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
sakura	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IIT-H	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
gauvar	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
JBJBJB	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
SRPOL	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CFILT	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
coastal	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
CFILT-IITB											✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
human eval	✓	✓	✓	✓	✓						✓	✓	✓	✓	✓					

表 2：各タスクの参加チームリスト

最も参加チームが多かったのはインド諸語を対象とした多言語翻訳タスクであり、その他のタスクは 1 から 4 チームの参加であった。人手評価の費用はスポンサー企業・組織により負担していただいているが、全てのタスクを対象とすることはできないため、いくつかのタスクを選んで行った。

次章ではいくつかの翻訳タスクについて、タスクの簡単な説明と評価結果、および得られた知見について述べる。

2. 評価結果と得られた知見

2.1 特許翻訳タスク

特許翻訳タスクは特許庁より提供された特許対訳コーパス JPC[1]を用いて行われ、日英・日中・日韓の双方向の翻訳タスクで構成されている。翻訳評価は自動評価と人手評価を行った。自動評価尺度としては BLEU[2]、RIBES[3]及び AM-FM[4]を用いている。AM-FM は正確さと流暢さの両方を考慮したような評価手法である。人手評価は特許庁が公開している「特許文献機械翻訳の品質評価手順」 [2]の中の「内容の伝達レベルの評価」に従って行った。これは機械翻訳結果が原文の実質的な内容をどの程度正確に伝達しているかを、人手翻訳の内容に照らして、下記 5 段階の評価基準で主観的に評価するものである。

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

表 3：JPO の機械翻訳評価基準

日英・日韓には 3 チーム(TMU, Bering Lab, tpt_wat)が、日中には 2 チーム(Bering Lab, tpt_wat)が自動評価サーバーに翻訳結果を提出したが、人手評価用の翻訳結果を提出したのは 1 チーム(TMU)のみであった。また予算の都合上、実際に人手評価を行ったのは日英、英日タスク 1 チーム分のみである。図 1 に自動評価結果を、図 2 および図 3 に人手評価結果を示す。

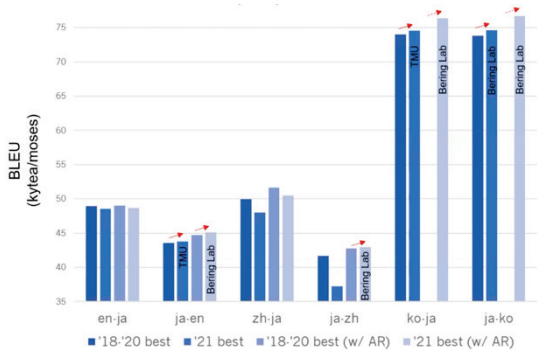


図 1：特許翻訳タスク自動評価結果(BLEU)

[1] http://lotus.kuee.kyoto-u.ac.jp/WAT/patent/

[2] https://www.jpo.go.jp/system/laws/sesaku/kikaihonyaku/tokkyohonyaku_hyouka.html

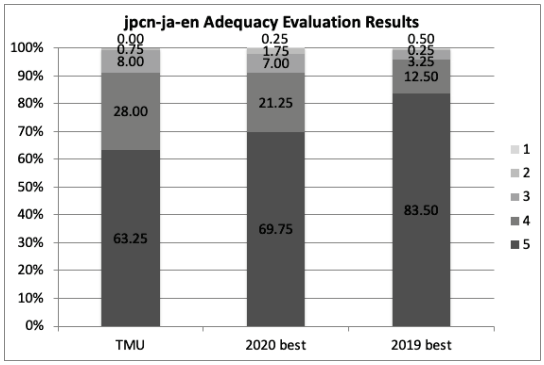


図 2：特許翻訳タスク人手評価結果（日英）

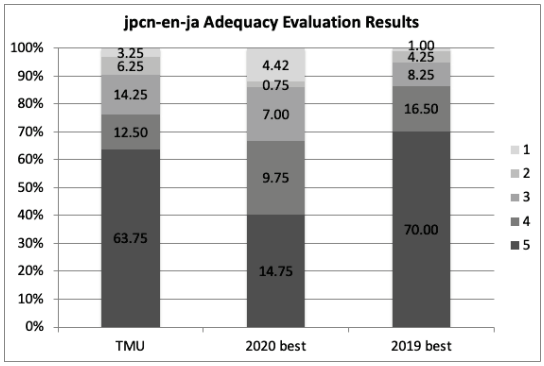


図 3：特許翻訳タスク人手評価結果（英日）

Bering Lab は JPC コーパスに加えて、自前で用意した 1300 万文からなる特許対訳コーパスも合わせて用いており、これのおかげで高い BLEU スコアを達成している。TMU は fine-tuning された日本語の BART[5]モデルを利用しており、韓日翻訳において最も良い AM-FM スコアを達成した。日英・英日の人手評価結果を見ると、残念ながら過去の WAT でのベストなシステムと比べるとやや低い精度という結果であった。しかしながら、BART という新たな NMT の枠組みを利用しても、過去のシステムに引けを取らない精度であることが示された。

2.2 日英・英日制限翻訳タスク

日英・英日制限翻訳タスクは WAT2021 で新たに追加されたタスクである。NMT は訳語統一が不得意であることはよく知られているが、専門用語や固有名詞を特定の用語に常に正しく翻訳することが求められるような文書の種類も多い。この問題に対して、現在の精度を知ることや解決方法を模索するために提案されたのがこのタスクである。

入力文とともに出力文で必ず使わなければならない訳語のリストが与えられるので、システムはこれらの訳語を必ず含むように翻訳を生成しなければならない。なお与えられるのは訳語のリストのみであり、入力文のどの語に対応する訳語なのかは与えられない。今回は ASPEC の日英データ(dev/devtest/test)を対象とし、10 人のバイリンガルに依頼して専門用語(制限用語)の抽出を行なった。表 2 に文単位の制限用語の平均数を示す。

	英→日 (フレーズ数、文字数)	日→英 (フレーズ数、単語数)
dev	(2.8, 16.4)	(2.8, 6.6)
devtest	(3.2, 18.2)	(3.2, 7.3)
test	(3.3, 18.1)	(3.2, 7.4)

表 4：文単位の制限用語の平均数

自動評価は BLEU スコアと一貫性スコアにより行い、最終的なシステムのランキングはこれらを組み合わせたスコアにより行なった(表 3 中の final)。一貫性スコアは全ての制限用語が出力できた文の割合であり、最終ランキングは全ての制限用語が出力できた文のみで計算した BLEU スコアにより決定した。人手評価は Direct Assessment および Contrastive Assessment[6]により行なった。

英日翻訳には 3 システムが参加し、日英翻訳には 4 システムが参加した。表 3 に自動評価と人手評価の結果を示す。

En-Ja		Human Eval.	
Team	final	src-based DA	src-based CA
NTT	**57.2**	**77.5**	**79.7**
NHK	33.9	74.1	77.2
NICTRB	28.8	73.6	77.1
(human ref.)	—	73.4	76.4
Ja-En		Human Eval.	
Team	final	src-based DA	src-based CA
NTT	**44.1**	**75.6**	**74.4**
NHK	37.5	73.9	73.5
NICTRB	31.8	72.1	71.8
TMU	22.6	50.2	48.3
(human ref.)	—	74.1	72.9

表 5：制限翻訳タスクの評価結果

全てのシステムが入力文に制限用語を付加して入力し、翻訳時には制限用語を可能な限り出力するように翻訳の探索を行うという方針をとっており、これらの手法は実際に指定された訳語を適切に出力するのに有効であることが示された。

タスクの仕様として訓練データには制限用語が付加されていないため、訓練時にはなんらかの方法で制限用語を準備する必要がある。各システムの違いを見ると、ここでの工夫の仕方が最終的な精度に影響を与えているようである。NHK は固有表現抽出技術により抽出された固有表現と、ベースラインとなる NMT において誤訳となった表現を制限用語として用いるという手法を提案している。一方で NTT は LeCA[7]と呼ばれる手法を用いている。結果を見ると NTT が提案した手法の方が精度良く制限用語を出力し、翻訳精度も高いということが示された。

興味深い点として、元の対訳文の精度(表 3 中の human ref)はそれほど高くないということが示された。今後対訳コーパス自体のクリーニングといったことが必要になる可能性がある。またほとんどのシステムが human ref よりも高精度を達成しており、現時点でも人間の翻訳に匹敵するような精度であることがわかる。

2.3 日英マルチモーダル翻訳タスク

Flicker30k の英日翻訳タスクには 2 チーム、日英翻訳タスクには 1 チームが参加した。訓練データのサイズが増加したことと、新たな翻訳技術が提案されたことにより、全ての翻訳結果が昨年度の結果よりも良い精度を達成した。英日方向ではテキストのみで翻訳を行なったシステムよりも、画像情報を用いたシステムの方が良い精度であったが、日英方向ではテキスト情報のみを用いたシステムの方が精度が高かった。これは用いたデータセットが英語から日本語へ、画像情報を参照しながら翻訳されたものであることが関係しているものと思われる。また単語の範囲と画像とのソフトアライメントを考慮したシステムを構築したチームが良い精度を達成しており、テキストと画像のグラウンディングが重要であることが示唆される。

Ambiguous MS COCO の英日翻訳タスクには 1 チームが参加したが、日英翻訳タスクへの参加はなかった。こちらのデータセットにおいても、テキスト情報だけでなく画像情報を適切に用いることで翻訳精度を改善させることができることが示唆された。

3. まとめ

本稿では WAT2021 における特許翻訳タスクと日英制限翻訳タスクおよび日英マルチモーダル翻訳タスクの結果を報告した。アジアの翻訳研究の活性化、データ整備等を目的として 2014 年に始めた WAT は、ドメイン数や言語数の増加、参加者数の増加など一定程度の成果を得ており、WAT を通じてアジア地域の機械翻訳研究コミュニティーの連携等が行えると良いと考えている。

NMT の翻訳精度は年々向上しているが、訳抜け、過剰訳、訳語の一貫性、長文の対応など未解決の問題はまだまだ多く残されている。WAT2021 で行われた制限翻訳タスクはこのうちの一貫性の問題に注目したものであり、実用的にも重要なタスクとなっている。今後も特定の現象に絞ったようなタスク設計が重要と考えており、引き続きデータの収集や評価指標の定義などを行なっていく予定である。

WAT は今後も継続して開催予定であるが、来年度の開催予定は未定である（11 月末辺りに決定予定）。ま

た WAT では翻訳評価にかかる費用等のためのスポンサーを募集しているため、興味のある方はご連絡いただければ幸いである。

参考文献

[1] Nakazawa, T., Nakayama, H., Ding, C., Dabre, R., Higashiyama, S., Mino, H., Goto, I., Pa Pa, W., Kunchukuttan, A., Parida, S., Bojar, O., Chu, C., Eriguchi, A., Abe, K., Oda, Y., Kurohashi, S., 2021. Overview of the 8th Workshop on Asian Translation, in: Proceedings of the 8th Workshop on Asian Translation (WAT2021). Association for Computational Linguistics, Online, pp. 1–45.

[2] Kishore Papineni, Salim Roukos, ToddWard, andWeiJing Zhu. 2002. Bleu: a method for automatic evaluation of machine translation. In Proceedings of ACL, pages 311–318.

[3] Hideki Isozaki, Tsutomu Hirao, Kevin Duh, Katsuhito Sudoh, and Hajime Tsukada. 2010. Automatic evaluation of translation quality for distant language pairs. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing, pages 944–952.

[4] Rafael E. Banchs, Luis F. D'Haro, and Haizhou Li. 2015. Adequacy-fluency metrics: Evaluating mt in the continuous space model framework. IEEE/ACM Trans. Audio, Speech and Lang. Proc., 23(3):472–482, March.

[5] Lewis, M., Liu, Y., Goyal, N., Ghazvininejad, M., Mohamed, A., Levy, O., Stoyanov, V., Zettlemoyer, L., 2019. BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension. arXiv:1910.13461 [cs, stat].

[6] Federmann, C., 2018. Appraise Evaluation Framework for Machine Translation, in: Proceedings of the 27th International Conference on Computational Linguistics: System Demonstrations. Association for Computational Linguistics, Santa Fe, New Mexico, pp. 86–88.

[7] Chen, G., Chen, Y., Wang, Y., Li, V.O.K., 2020. Lexical-Constraint-Aware Neural Machine Translation via Data Augmentation, in: Proceedings of the Twenty-Ninth International Joint Conference on Artificial Intelligence. Presented at the Twenty-Ninth International Joint Conference on Artificial Intelligence and Seventeenth Pacific Rim International Conference on Artificial Intelligence {IJCAI-PRICAI-20}, International Joint Conferences on Artificial Intelligence Organization, Yokohama, Japan, pp. 3587–3593. https://doi.org/10.24963/ijcai.2020/496

編集後記

内山 将夫
AAMT 編集委員会

本号は AAMT 創立 30 周年記念特集号です。皆様ご存じのように、ここ数年で NMT による機械翻訳が広く社会で実用されるようになりました。ここで改めて、機械翻訳とはどういうものか、どのように活用していくべきかを皆様と考えていきたいと思っています。

巻頭言では、高橋聡様より、「『機械翻訳』を訳した翻訳者が考えていること」をご寄稿いただきました。ご提言いただいた

- 機械翻訳研究者と翻訳者との協力態勢
- ソリューションの活用に関するガイドライン
- 個人翻訳者が機械翻訳を正しく評価できる場

はいずれも重要なものですので、今後、AAMT 策定の「MT 利用ガイドライン」等で、取り組んでいきたいと思います。

第 16 回 AAMT 長尾賞受賞者の Team Tohoku-AIP-NTT at WMT-2020 様からは、複数トラックで事実上の一位を獲得した、WMT-2020 ニュース翻訳タスクに参加した機械翻訳システムについてご寄稿いただきました。NMT の高精度化について、大変参考になると思います。同じく AAMT 長尾賞受賞の Mantra 株式会社様からは、「マンガ機械翻訳への招待」をご寄稿いただきました。ご紹介いただいた「Mantra Engine」は、機械翻訳を含むマンガ翻訳に係るあらゆる作業をツール内で完結できるものです。このように、ワークフローにしっかりと組み込まれた機械翻訳エンジンは、マンガ以外への機械翻訳の利用にも大きな示唆となると思います。

第 8 回 AAMT 長尾賞学生奨励賞は、AAMT 長尾賞とのダブル受賞の石渡祥之佑様から、「Translation and Description Methods for Multilingual Text Understanding」をご寄稿いただきました。ご提案の「未知フレーズに対する説明文」の生成は、新規性が高く、今後の発展性も期待できると思いました。

また、AAMT 創立 30 周年を記念して、過去の AAMT 長尾賞受賞者のなかから、凸版印刷株式会社様、株式会社バオバブ様、八楽株式会社様、株式会社みらい翻訳様に現在のご様子をご寄稿いただきました。いずれの方々も益々ご活躍されていることがお分かりかと思います。

現在 AAMT では「MT 利用ガイドライン」を策定中ですが、山田優様には、「MT 利用ガイドラインの必要性」をご寄稿いただきました。MT は、現在、広く社会的に使われていますが、それに伴う各種の問題も顕在化している状況ですので、MT 利用のガイドラインは正に必要です。本ガイドラインは 2022 年 5 月末の完成を目指しています。

影浦峡様には、「人間の翻訳と機械の翻訳（5）：翻訳者のコンピテンスとは何か」をご寄稿いただきました。学生（翻訳授業の履修経験有）とプロの翻訳者を対象とした、翻訳作業時の情報探索（調査）行動の量的・質的な分析の紹介など、MT の研究者にも示唆がある内容と思いました。

石岡英子様には、「【JTF40 周年特別企画】2021 年度第 1 回 JTF 関西セミナー：機械翻訳とは何か？ どこから来て、どこへいくのか？」のイベント報告をご寄稿いただきました。本号にもご寄稿いただいている高橋様と中澤様によるご講演・パネルディスカッションが大変興味深いです。

出内将夫様には、「AAMT 招待講演『語学教育と MT』機械翻訳の問題～第二言語教育の立場から～」のイベント報告をご寄稿いただきました。本講演では、トム・ガリー様から、「MT と言語教育の問題」をご紹介して

Editor's note
Masao Utiyama
AAMT Editorial Board

いただいています。MT の精度がどんどん良くなっていく現状で、言語教育と MT との関係について、大変参考になりました。

中澤敏明様には「第 8 回アジア翻訳ワークショップ（WAT2021）開催報告」をご寄稿いただきました。WAT2021 では 18 の言語を対象とした 14 の shared task が行われ、世界中から 26 のチームが参加しました。入力文とともに出力文で必ず使わなければならない訳語のリストが与えられる「日英・英日制限翻訳タスク」では、「人間の翻訳に匹敵するような精度」が達成されています。

AAMT 創立 30 周年の今年は、機械翻訳がこれから大いに使われていく初動の年になると思います。今後、10 年・20 年先に、機械翻訳がどのように世の中に広まるかが楽しみです。

AAMTジャーナル「機械翻訳」No. 75

【 発 行 日 】2021年12月15日
【 発 行 】アジア太平洋機械翻訳協会（AAMT）
ホームページ：https://aamt.info/
【 住 所 】〒619-0289
京都府相楽郡精華町光台3-5
国立研究開発法人情報通信研究機構　先進的翻訳技術研究室内
【編集委員会】内山将夫　後藤功雄　中澤敏明　新田順也　園尾聡　森口功造　隅田英一郎
仲山裕子　山田優　石川弘美　早川威士
【表紙デザイン】泉谷東十郎
【 題 字 】長尾真
【 事 務 局 】石川弘美
【 印 刷 所 】株式会社 プリントパック

Asia-Pacific Association for Machine Translation (AAMT)
c/o National Institute of Information and Communications Technology
3-5 Hikaridai Seika-cho Soraku-gun Kyoto 619-0289, Japan

AAMT

Asia-Pacific Association for Machine Translation

アジア太平洋機械翻訳協会