

# AAMT

Asia-Pacific Association for Machine Translation

# Journal



June 2015

No.59

アジア太平洋機械翻訳協会

## 目 次

巻頭言：	言葉の物質性が作動するとき.....影浦 峯..... 2
研究報告：	Challenges and Advances in Statistical Machine Translation of Japanese ..... Graham Neubig..... 4
レポート：	目的と目論み、存在と関係.....佐野 洋..... 7
連載：	機械翻訳事始め [ v ] .....坂本 義行..... 13
シンポジウム報告：	The Five-Year Race to Tokyo 2021 Started at the TAUS Executive Forum. The goal: Beat Google. ..... Jaap van der Meer..... 22
シンポジウム報告：	TAUS Tokyo Forum .....中村 哲三..... 27
プロジェクトレポート	多言語情報発信シンポジウム報告.....井佐原 均.....33
シンポジウム参加報告：	第 14 回日中自然言語処理共同研究促進会議（CJNLP2014）参加報告 .....中岩 浩巳，長瀬 友樹..... 40
レポート：	Japio 世界特許情報全文検索サービス（Japio-GPG/FX）の紹介 .....長谷川 雅昭,大塩 只明..... 42
レポート：	特許版・産業日本語の取組（特許ライティングマニュアルについて） .....早川 貴之..... 48
セミナー報告：	第 3 回 JTF 関西セミナーに参加して.....内田 真弓.....53
セミナー報告：	みらい翻訳のご紹介.....栄藤 稔.....56
製品紹介：	Transit <sup>NXT</sup> 最新サービスパック 8 について.....目次 由美子..... 57
委員会活動報告：	これまでの AAMT Forum メールマガジン Vol.5 ..... 機械翻訳課題調査委員会 WG1,2..... 59
委員会活動報告：	UTX 変換ツール正式版の公開 ..... 機械翻訳課題調査委員会 WG3..... 64
AAMT 会員のひろば：	..... GMO スピード翻訳株式会社.....66
AAMT 会員のひろば：	.....株式会社サン・フレア..... 68
AAMT 会員のひろば：	.....秋元 圭.....70
事務局からのお知らせ：	協会活動報告（2015 年 1 月～2015 年 4 月） .....AAMT 事務局.....72
編集後記	.....77

# C O N T E N T

Foreword:	Material Aspect of Language in Action: What NLP and Linguistics Have Overlooked ..... <i>K.Kageura</i> .....	2
Research Report:	Challenges and Advances in Statistical Machine Translation of Japanese ..... <i>G. Neubig</i> .....	4
Report:	Yet Another Approach to Understanding the Differences among the Languages..... <i>S. Hiroshi</i> .....	7
Serial lecture:	History of Machine Translation [ v ] ..... <i>Y.Sakamoto</i> .....	13
Symposium Report:	The Five-Year Race to Tokyo 2021 Started at the TAUS Executive Forum. The goal: Beat Google. ..... <i>Jaap van der Meer</i> .....	22
Symposium Report:	TAUS Tokyo Forum ..... <i>T. Nakamura</i> .....	27
Project Report	Symposium on Multilingual Information Sharing and Launch of TKUN (Translation Knowledgebase and User Network)..... <i>H.Isahara</i> .....	33
Symposium Report:	The 14th China-Japan Natural Language Processing Collaboration Promotion Conference (CJNLP) ..... <i>H.Nakaiwa,T.Nagase</i> .....	40
Report:	Overview of Japio Global Patent Gateway / Full Text (Japio-GPG/FX) Service ... <i>M.Hasegawa,T.Oshi</i> .....	42
Report:	Efforts for Technical Japanese for Patent(Patent Documents Writing Manual) ..... <i>T.Hayakawa</i> .....	48
Seminar Report:	How Machine Translation is used in the world of business ..... <i>M.Uchida</i> .....	53
Seminar Report:	Introduction of Mirai Translate, Inc. ......... <i>M. Etoh</i> .....	56
Products & Services:	Transit <sup>NXT</sup> About the latest Servicepack 8..... <i>Y.Metsugi</i> .....	57
Committee Report:	AAMT Forum mail magazine vol.5 ..... <i>WG1,WG2</i> .....	59
Committee Report:	Release of UTX Converter..... <i>WG3</i> .....	64
AAMT Members	..... <i>GMO Quicktranslate, inc.</i> .....	66
AAMT Members	..... <i>SunFlare Co., Ltd.</i> .....	68
AAMT Members	..... <i>K.Akimoto</i> .....	70
AAMT Activities:	AAMT Activities (from January 2015 to April 2015) ..... <i>T. Utsuro</i> .....	72
Editor's Note:	..... <i>T. Utsuro</i> .....	77

# 言葉の物質性が作動するとき

## Material Aspect of Language in Action

影浦 峯

東京大学・大学院情報学環／教育学研究科

翻訳では当たり前接するけれども言語「学」や言語処理「研究」ではあまり扱われていない、言葉の物質性について見てみたいと思います。話は少し拡散しますがご容赦下さい。

\* \* \*

例から入りましょう。

'Twasbrillig, and the slithytoves

Did gyre and gimble in the wabe:

All mimsy were the borogroves,

And the momerathsoutgrabe.

ルイス・キャロルの有名なナンセンス詩ジャバーウォッキの一節で、「翻訳せよ」と言われると困ってしまいます。詩は言葉の芸術なので、意味が取れたとしても「定義上翻訳不可能だ」という主張が出てきます。とはいえ、言葉の物質性に拘るならば朗読の声や書体などがどこまで違えば「その詩」でなくなるのかは微妙で、翻訳の可能性もそこから広がることになりましょう。もっと普通に、

「枝垂れさくら腐れ出し」

なども、回文であることに気づいてしまうと、意味だけでなく言葉の物質的属性の対応も維持しなくてはならなくなるため、翻訳が難しくなります。

\* \* \*

日常会話や語学教科書、文法書にも物質としての言葉は溢れています。例えば

He have a pen.

はいけない、というときのいけなさは、本来、単に音／形がいけないという以外に捉えようがないものですから、言葉の物質的側面に関わっています。

ですから「三人称単数主語で時制が現在ならば have は has にならなくてははいけない」という後付け

の説明は、言葉の物質的側面を扱っていることになります。つまり言語学も言語処理も (MSWord というソフトは上記 have に緑の波下線を引いてくれます) 言語の音／形という物質的な側面を扱っていると言えるのですが、だからこそ、ここのところで、翻訳が言葉のかたちや物質的側面を扱うと言ったときとの違いが際立ちます。

He have a pen.という文を前にしたとき、翻訳においては、少なくとも、

(a) 「彼がペンを持っている」という状況が話題化されているものであり、従って、have を些細な誤りとして扱い意味を取るか、

(b) He have a pen.という言葉そのものが話題化されているものとして、表現をそのまま残すかの判断が必要になります。

これに対して言語学 (や言語処理) では、He have a pen.という表現は基本的に、常に「彼がペンを持っている」という話題において扱われ、それを表現する媒体としての言語のかたちが正しいか誤っているかという限りで言葉の形が問題化されます。

言葉を換えると、翻訳で(b)の選択がなされるときには、表現を話題化している言語学者 (でなくてもかまいませんがいずれにせよ、「彼がペンを持っている」という話題において語られた He have a pen. という表現の適切さを論じる人) の言葉とその言語状況も同時に対象化されますが、言語学において自らの議論が同時に対象化されることはありません。つまりここでは、言葉の物質的側面だけではなく、それについて語られる言葉が、言葉を使うことにおいてどう位置づけられるかが、両者の違いを生んでいると言えます。

\* \* \*



さらに、言葉を使ってものを考える際に、言葉の物質的側面を利用することはごく普通のことです。少し話は広がってしましますが、次のような問題を解くことについて考えてみます。

ニコルとヴァレリーが森を散歩していたら広場に出ました。広場にはポプラの木が円形に植えられていたので二人は木の本数を数えることにしました。別々のところから数え始め、ニコルにとっての 20 番目の木はヴァレリーにとっての 7 番目の木で、ヴァレリーにとっての 94 番目の木はニコルにとっての 7 番目の木でした。木は全部で何本あったでしょう。(『ロジカルな思考を育てる数学問題集(上)』セルゲイ・ドリチェンコ著、岩波書店、2014 年)

文科系では、この問題に最初から拒否反応を起こし、解くことまでたどり着けない人がいます(20 人程の反応に基づいています)。興味深いことに、そうした人は、「この問題文の中で参照されている木は何本か」についての話ができません。答えは 4 本(数え始めを考慮するとき)ないし 2 本(考慮しないとき)で、ニコルの 20 番目=ヴァレリーの 7 番目に例えば A、ヴァレリーの 94 番目=ニコルの 7 番目に例えば B と名前を付けることが重要になります(模範解答もそこから始まっています)。

これによって小さな可能世界をめぐる思考実験を始めることができ(「A がニコルの 5 番目でヴァレリーの 10 番目だったら・・・」等々。「ニコルの 20 番目かつヴァレリーの 7 番目がニコルの 5 番目かつヴァレリーの 10 番目だったら・・・」と言うのは、クリプキ流の固有名論を持ち出すまでもなく苦しい話になります)、問題が扱っている世界の単純化や一般化が可能になります。

前の話と何の関係があるかと思われるかも知れませんが、A と B という記号は、話を展開するために使う以外、独自の意味はない、既に言語で表現された対象を改めてつなぎとめる、物質としての役割を担った記号である点が大切です。

He have a pen.とは物質性が問題になる所がいわば逆なので、大胆に議論を端折ることになります。この問題がほとんど意識せずに当たり前に出てしまう人は He have a pen.を前にした母語話者に、A、B と記号を付与して説明する人が「三単言」に言及する言語学者に対応すると考えると、A、B を付与する行為が問いにおいて(その中で)どのように作用するかを考える人は、言語表現に対して記号を重ねて行く行為を含めて言語表現の物質的側面を扱う翻訳者に対応すると考えることができます。

\* \* \*

かなり乱暴にはありますが、言葉の物質的側面を参照したりそれを活用したりする行為は言葉を使う中で遍在しており、それにもかかわらず言語学のようなかたちで言語を対象化すると対象化する言葉それ自体は対象から漏れ落ちるので、言葉の中で言葉の物質性がどう作用しているのか見落としがちであることを見てきました。

‘Dog’ consists of three characters.

機械翻訳は一般にこのような文---中世哲学の用語を使えば質量代表(suppositiomaterialis)に関わるもの---を苦手としています。コンピュータが、人間が漠然と考えるところの「意味」ではなく記号の列(したがって記号の物質的側面、といってもそれを形式化したものですが)を扱っていることを考えるならば、これは不思議なことです。本来、

jejustonedaybrillingtonwasjabberwockyglydegogohoure  
の字数を間違えずに数えることや、同一の記号を一貫して使い続けることは、コンピュータの方が遥かに得意ですから、MT が上記の文を苦手としているのは、多分、それを自分の仕事だと認識していないからではないでしょうか。

MT の進歩により、翻訳の世界は大きく変わっています。一方、人間が行っている翻訳という行為との出会いが、MT の課題設定をいっそう広げてさらに楽しいものにしていく可能性の側については、まだまだ開拓の余地があるかもしれません。

# Challenges and Advances in Statistical Machine Translation of Japanese

Graham Neubig

Nara Institute of Science and Technology (NAIST)

Japanese, despite its relatively large speakership and economic importance, has been one of the most elusive languages for statistical machine translation (SMT). Still, accuracy of SMT systems between Japanese and English lags far behind that of translation between other language pairs such as English and French, Spanish, or German.

There are a number of reasons for this, but by far the most prominent is that Japanese has a grammar that is greatly different from English, Chinese, and perhaps most other heavily researched languages (with the exception of Korean). However, in the past 5 years, there have been major advances in SMT of Japanese to/from English, leading to improvements to the point that SMT systems are posting results that exceed even strong rule based systems in manual evaluation. In this report, I will focus on three recent advances in SMT, with a focus on ones that are particularly applicable to Japanese:

## Incorporation of Syntax

One of the most prominent methods for SMT is *phrase-based machine translation*, which started in the early 2000s and remaining popular to this day, thanks in part to the widely used Moses decoder (<http://statmt.org/moses>). Phrase-based SMT is based on a translation model that looks something like a translation dictionary; it contains aligned multi-word phrases in the source and target language such as “went to” → “に行った”. When a new sentence to be translated is input by the user, phrase-based SMT translates between the two languages by translating and rearranging these short phrases into the target order, as shown in the figure on the right.



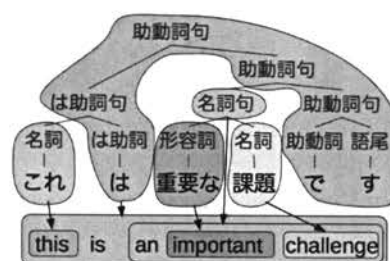
Phrase-based SMT has proven extremely effective for European languages, but it falls short on Japanese-English and other language pairs for which the languages have very different structures. This is mainly due to the fact that the simple idea of translating small chunks of words and testing all of the possible reorderings of them simply doesn't scale to language pairs that require extensive reordering. One of the main reasons for the great improvements in SMT accuracy for Japanese is methods to fix this reordering problem through the use of the *syntactic structure* of the source sentence. There are two main methods: pre-ordering, and syntax-based SMT.

*Pre-ordering* is based on the assumption that “phrase based SMT can do well if the word order of the two languages is similar.” In order to take advantage of this assumption, pre-ordering adds an extra pre-processing step where we first re-order the input sentence into a word order similar to the target language, then use standard phrase-based SMT techniques to perform translation of this reordered input. An example of this is shown on the right.



The advantage of this method is that it is possible to use tools in the pre-ordering step, such as syntactic parsers and hand-crafted rules, that are not easy to incorporate into phrase-based translation, and it has been shown that this is highly effective in improving the reordering phrase-based SMT output (Isozaki et al. 2010)

*Syntax-based SMT* is another method that makes it possible to overcome the difficulties posed by reordering for phrase-based SMT. Syntax-based SMT works by first performing syntactic parsing of the input sentence, then based on that syntactic parse, using translation rules that translate a part of the parse tree into words in the target language, as shown to the right. The advantage of this method is that it can, like pre-ordering, use the structure of the source sentence to achieve more accurate reordering. Compared to pre-ordering, it also has the advantage that reordering rules can be directly learned by the SMT system, and thus doesn't require splitting the MT process into two



separate steps. This method has proven quite effective (Neubig and Duh 2014), and recently achieved state-of-the-art results in a number of tasks, such as the English-Japanese, Japanese-English, Chinese-Japanese, and Japanese-Chinese translation tasks at the Workshop on Asian Translation 2014 (Nakazawa et al. 2014). At NAIST, we are currently developing an open-source toolkit Travatar (<http://phontron.com/travatar>) that implements recent advances in syntax-based translation.

### Generating Missing Words

Another problem often faced in translation to/from Japanese is the occurrence of words that only exist in one of the two languages. For example, in Japanese, grammatical particles indicating the topic, subject, or object of the sentence (“は”, “が”, and “を”) exist, but have no counterpart in English, and English determiners (“the” and “a”) have no counterpart in Japanese. Another problematic example results from the fact that in Japanese it is often permissible to omit the subject of the sentence when it is obvious from context, while this is usually not allowed by standard English grammar. Producing these words that don't exist in the source sentence is difficult with standard SMT techniques.

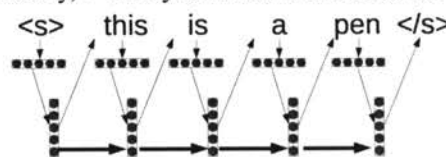
One way to fix this problem is to turn to pre-processing, specifically adding a “dummy” token into the source sentence. For example, in the example on the right, the pre-processing step examines the Japanese sentence, and guesses whether the sentence has an explicit subject or not, and if the sentence does not have a subject, it tries to guess the subject and insert the appropriate subject into the source sentence. This variety of pre-processing has been shown to be particularly effective for subject insertion (Kudo et al. 2014), but can also be used to insert Japanese particles into an English source sentence, or fix other similar divergences in syntax (Isozaki et al. 2010, Hatakoshi et al. 2014).



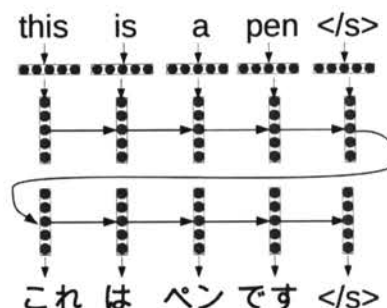
### Probabilistic Models using Neural Networks

The final topic I will cover in this report is probabilistic models using neural networks. Neural networks (popular recently for their use in “deep learning”) are a variety of machine learning that make it possible to learn very complicated probabilistic functions. Neural networks can be used in a number of ways in machine translation, but perhaps the most prominent are as “language models” and “translation models.”

*Language models*, in the context of SMT, are models that predict a probability for the output sentence, giving a high probability to natural sentences, and a low probability to un-natural sentences, helping the SMT system produce more natural output. Traditionally, SMT systems use a variety of language model called “n-grams,” which look at the past few words in the sentence to predict the probability of the next word, and multiply these probabilities together to get the probability of the whole sentence. Recently, a variety of neural network called “recurrent neural networks” have become a popular supplement to n-gram models (Mikolov et al. 2010). Recurrent neural networks work by reading in each word of the sentence, using an automatically learned function to “remember” the word in the form of a vector of numbers, then use this vector of numbers to predict the next word. These models have led to impressive gains in language modeling performance, and are starting to become a widely used element in state-of-the-art machine translation systems, particularly thanks to the open-source RNNLM toolkit (<http://rnnlm.org>).



*Neural translation models*, are a relatively new, but promising development. Neural translation models are, like language models, often based on recurrent neural networks. However, they differ from language models in that they first read in the source sentence one at a time using a recurrent neural network, essentially “remembering” the input sentence as a vector of numbers. They then predict the words in the output sentence one-by-one, selecting the word that has the highest probability, generating the output sentence. It is, in a way, hard to believe that such a simple process is powerful enough to learn all the intricacies involved in translation, but already neural translation models are posting impressive results that rival other methods in translation between European languages (Jean et al. 2015, Luong et al. 2015).



## Conclusion

This report has covered three advances in SMT technology that have proven particularly effective in improving accuracy of translation to or from Japanese in the past five years. All in all, this is an exciting time for Japanese SMT, as the accuracy is finally starting to approach a level that is actually acceptable for serious users.

It should be noted, however, that none of these three technologies are finished; in fact, research in some of these fields has just started. Unresolved problems include how to further improve the accuracy of syntactic parsing or deal with mistaken parse trees in translation, how to better estimate missing words by incorporating richer context and more varieties of information, and more refined neural network models with higher accuracy. In addition, there are still a number of other problems such as generating the correct conjugations of words in the target sentences, or choosing the correct word sense for ambiguous words. These will be among the challenges that researchers will have to tackle to make even more accurate, useful SMT systems in the near future.

## References

- Y. Hatakoshi, G. Neubig, S. Sakti, T. Toda, and S. Nakamura. Rule-based syntactic preprocessing for syntax-based machine translation. In Proc. SSST, 2014.
- H. Isozaki, K. Sudoh, H. Tsukada, and K. Duh. Head finalization: A simple reordering rule for SOV languages. In Proc. WMT and MetricsMATR, 2010.
- S. Jean, K. Cho, R. Memisevic, and Y. Bengio. On using very large target vocabulary for neural machine translation. In Proc. ACL, 2015.
- T. Kudo, H. Ichikawa, and H. Kazawa. A joint inference of deep case analysis and zero subject generation for japanese-to-english statistical machine translation. In Proc. ACL, pages 557–562, 2014.
- M.-T. Luong, I. Sutskever, Q. Le, O. Vinyals, and W. Zaremba. Addressing the rare word problem in neural machine translation. In Proc. ACL, 2015.
- T. Mikolov, M. Karafiat, L. Burget, J. Cernocký, and S. Khudanpur. Recurrent neural network based language model. In Proc. InterSpeech, pages 1045–1048, 2010.
- T. Nakazawa, H. Mino, I. Goto, S. Kurohashi, and E. Sumita. Overview of the 1st Workshop on Asian Translation. 2014.
- G. Neubig and K. Duh. On the elements of an accurate tree-to-string machine translation system. In Proc. ACL, pages 143–149, 2014.

## 目的と目論み、存在と関係

佐野 洋

東京外国語大学 大学院総合国際学研究院

### 1. はじめに

#### 1.1. 出来事（ものごと）の生起

「時間が流れる」という我々の常識的な意識は、現実（の物理）世界の事実ではありません。[1]は「われわれは時間が流れる、あるいは過ぎていくのを体験しない。われわれが体験しているのは、われわれの現在の知覚と、現在の記憶のなかにある過去の知覚との違いである。（231 頁）」と指摘します。同著では「時間の枠組みの正確な描像は、動く、あるいは変化する描像ではありえず、静的でなければならない」といい、時空とは単一の不変ブロックとなるそうです（[1]：235 頁から引用した図 1 を参照）。

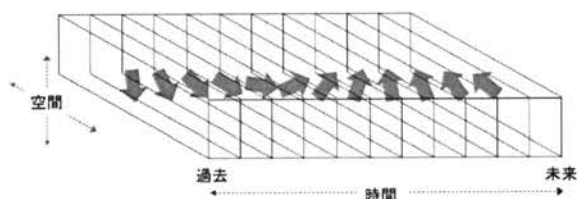


図 1

瞬間毎のスナップショットが、物理的に存在する空間全体にまたがり単一のブロックを構成します。図 1 では個々の時間（現在）のスナップショットを（本来は空間なので三次元ですが）二次元で表現しています。

「われわれが「瞬間」と呼ぶものは、時空を切った一片のスライスであって、こうしたスライスの内容がたがいに異なるとき、われわれはそれを空間を通る変化するなわち運動と呼ぶ」（[1]：236 頁）と説明しています。

物理学では今（現在）という概念は必要としません。なお、図 2 は一つのスナップショットを三次元の空間として表した図を図 1 に追加したものです。追加された三次元の空間の矢印は、動き（コト）を表します。空間内には当然ながらモノが含まれます。

我々が意識する常識的な認識は、現在が単一の瞬間

であると考えようとしています。主観的な時間の流れを感じているのです。言語学分野の多くの書籍でもこの主観的な時間の意識を使った説明がなされています。過去、現在、未来を点として直線上に配置する模式です。しかし、時間は瞬間の系列であって、すべての運動と変化は、異なる瞬間における、ある実体のバージョンの相違に他ならないのであり、それは瞬間そのものが不変であることを意味しています（[1]：233 頁）。そして現在は、客観的には単一の瞬間ではあり得ません。

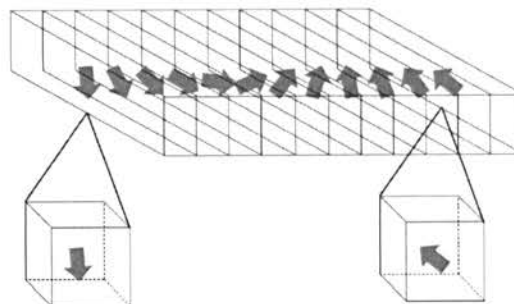


図 2

[1]では、出来事同士を「原因と結果」として説明するときには、動いている現在（主観的な時間の流れ）を用いて考えていると指摘しています。ですから、（我々が意識する）主観的な時間認識は、エピソード記憶と関係する意識内で創造された概念存在なのでしょう（[2]）。主観的な時間の流れは常に脳内のシミュレーションによって行われる意識の産物として扱う範疇のものです。

単一の動き、すなわち極度に限定され断片化された時空表現としての運動の把握（いわゆる出来事の意味）では、我々は恐らく無意識に不変の瞬間の系列を用いて考えて（言語学の表現形式と違い比較的現実に近い時間の捉え方をして）いるのでしょう。出来事の役割を考察する際には図 2 を用いるのがよいことになります。



## 1.2. 因果関係の捉え方

ヒトの思考パターンの特徴として、原因を見つけることに熱心であることが知られています。原因を見つける思考のプロセスが以下のように区分されています（[2]：64頁）。

- a 自分の行動の物理的な結果を予測することができる。
- b 他者の行動の結果を予測することができる
- c 他者の行動の原因を予測することができる
- d 物理的な出来事の原因を理解する

a～d への順序は、（生物の）進化の過程で現れる順序でもあるといえます。そして d は（不明な点が多いことを断りながらも）人間以外の動物には、ほとんど見られないそうです。原因の知覚（記憶との比較）によって過去概念が生まれます。さらに d は物理的な力が原因で起こる結果と、いわゆる行為者（の力）が引き起こした出来事を区別できる（因果関係に違いがある）ことになります。原因の知覚と予測する能力を組み合わせることで主観的な時間の流れが意識できます。

## 1.3. コトとモノ

諸事万端、森羅万象、すべてはものごと（物事）で成り立っています。ものごとをどのように表現（記号化して語と）するのかは言語によって違います。そして記号化の明示的な仕様として文法があります。文法は、形態論的、統語論的で意味論的な性質を説明する枠組みです。その中で品詞とは、枠組みの説明で使われる語彙カテゴリーです。個別言語で文法が異なるように品詞名（語彙カテゴリー名）も個別言語によって名称が違います。

本稿では、動詞と名詞、用言と体言に焦点をあてて、英語と日本語のものごとの捉え方と違いを、ヒトの時空間把握と、思考様式としての因果関係の捉え方の違いとして考察します。

## 2. 目的と目論み

### 2.1. 目的と動詞（英語）

英語は（従来から指摘されているとおり）、動詞の意味を「目的のある出来事」として表すことを好みます。動作・行為の背後に目的があると想定し、意図的であ

ると思えようとしています。意図は原因の特別なケース（[2]：145頁）で、因果関係理解のための要素の一つです。ですから典型的な出来事は、被動対象を伴い、動作主によって意図的に引き起こされ、完了する動作です。それは状態変化（比較的長い時間の経過）を伴いますから、三次元空間は連続体として時空間を作ります。時空間規模が大きいことが特徴です（図3）。なお、左右の点線のスナップショットは、動作・行為前の状態と、後の状態を示しています。

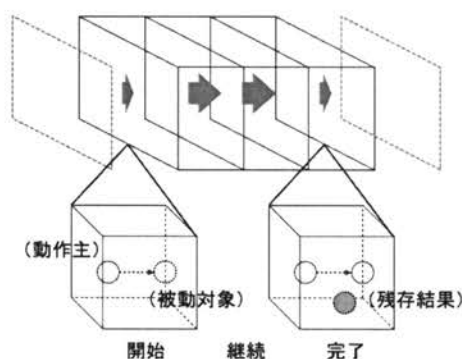


図 3

時空間が大きいのでモノが自律的に存在し、モノとモノが区別され、それらモノ同士が変化を伴う関係（一方から他方へ力が及ぶ他動性関係）で結ばれることが好まれます。そうして他動詞がよく使われます。

孤立語化が進む英語（[3]）は、位置関係を用いて動作主と被動対象を表します。つまり動詞の前の位置と後ろの位置で表します。自動詞と他動詞があり、同じ形をしている語彙がたくさんあります（自他同形）。動詞と名詞が同形の語彙も数多くあります。このことから（いわゆる）名詞句表現で動作や行為を表現することができます。中には形容詞も同じ格好をしている語もあります（日本語では考えられないことです）。

時空間の大きさは語が表す意味によって違うものの、比較的大きな時空間を使うので、（主観的な）時間推移の表現が発達し、そして文法的に分化し多用されます（完了表現や仮定表現等）。

時空間の規模が大きいので空間位置の関係性を表す形態素（in, on, at, to など）を語に前接させて出来事への関与の仕方や他の語との関係を表します。三次元の空間的な位置関係を担ったり、また、例えば、接

触するなど動的な関係表現を担わせたりします。形態素の種類が豊富で細かな違いを表現仕分けられます。つまり表現粒度として意識します。例えば「年 3.5%の融資」を"annual loans at 3.5%"と表現し at を使います。「債権の売買益」を"gains from bond trading"と言い換え、from を使います。「需要の回復」を"recovery in demand"で表すように in を用います。このように三次元時空間内の位置関係表現を明示します。

なお、ヒトは物理的な因果関係も理解できますから、  
 "The sun rises in the east and sets in the west." (「太陽は東から昇って西に沈む。」) などの自動詞表現もありますし、あまり変化の起こらないコトとして、be 動詞や自動性の高い動詞 (become、get 等) を使った状態表現も行います。

## 2.2. 目論みと用言 (日本語)

日本語は、用言の意味を「計画した出来事」として表すことを選好します。動作・行為の背後に目的があると想定はしますが、意図的であることを保留し、他者の心を理解しようとしします (目論み)。他者が知っていることを理解することは因果関係理解です ([2]: 145 頁)。ですから典型的な出来事は、主体によって目論まれ、(推論に拠って) 結果として起こるだろう出来事が引き起こされつつある動作です (取り掛かり表現、あるいは始動表現)。主体と動作主が一致する必然性も必要性もありません。目論む主体は、日本語が膠着型のことばであることから「日本人が」のように、形態素「が」を後続して表します。動作主も形態素「が」を後続して表します。

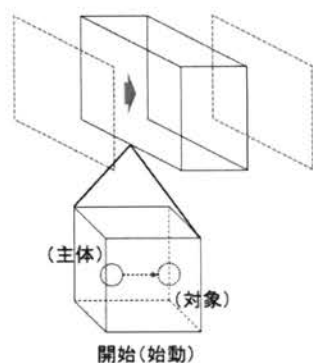


図 4

この場合の三次元空間も連続体として時空間を作り

ますが、その大きさは小さく、スナップショット (瞬間あるいは状態) に近いイメージです (図 4)。ですから、動詞の意味も状態性表現が多くなります。

時空間が小さいのでモノが関係的に存在し、それらモノ同士が在る関係 (一方と他方が繋がる静的な状態性の関係) で結ばれることが選好されます。

「が」をはじめ関係を表す形態素 (「が、を、に、へ、に、から」など) を体言に後接することで、出来事への関与の仕方や他の語との関係を表します。二次元の平面的な位置関係や静的な関係表現を担うので種類が少ないのが特徴です。助辞相当形態素 (「～について、～に関して、～に対して、～にとって、～によって、～において、～にかけて」など) が発達しているのは、人口が急激に増えた中世から、産業発展・振興に努めた近代にかけて複雑で多様な表現の弁別が必要になったからでしょう ([4])。

目論む主体と動作主を一致させ、そして被動対象を形態素「を」を後続させて表し、他動性表現を行います。ただ、コトの表現の大きさが小さいので、他動詞と称してみても動作・行為規模は自ずと小さく、それは、例えば「A 社は発電用設備を設置する。」の英語訳 "Company A plans to build a facility to generate." にみるように、日本語の他動詞の意味が、動作・行為がより明示される別の他動詞で表されていることでも分かります。同種の動詞、「改善を重ねる、推移を見守る、善処を検討する」などほとんど状態表現に近いでしょう。でもよく使われることばです。

このように自動性の用言と他動性の用言があります。ほとんどは字形によって区別されます (「上がる」「上げる」のように自他異形)。用言と体言が同形の語彙はサ変体言以外にほとんどありません (サ変体言は、「勉強する」のように「する」を後接して用言として用いることばです)。

さらに「震災の報道で悲しみが伝わった。」のように動作・行為が到達した意味のある事柄も自動性の用言で表現します (この例の場合、用言の意味は到達付近だけの有様を表しているのでしょう)。無論、他動性の用言を使った動作表現や行為表現もありますが、これまで述べたように用言の基本の意味は一貫して動作・

行為の始動部分を表します。

### 3. 存在と関係

#### 3.1. 存在と冠詞（英語）

英語は、大きな時空間内でモノが自律的に存在できるので、名詞の意味を「存在するモノ」として表すことを選好します。身近にあって外形（輪郭）がはっきりしている個物（man、woman、pen や apple）が、存在する典型的なモノです。people や machinery のように集合させて存在させるモノもあります。形容詞や形容詞句で性質や属性を限定することによって、他と区別させて存在させるモノもあります。時間範囲を区切ったり、領域を限定したりすることで外形を持たせ存在させるモノもあります（"a long history"/"a life"）。

英作文の参考書でしばしば「パラグラフ内で同じ単語（名詞）を繰り返すな」との指南がありますが、英語では、モノを表す表層語形が現れるたびに存在を想起するように解釈するからでしょう。文脈に間断なく存在するモノが現れるのは厄介です。すなわち言い換え表現は代名詞の役目を担っているのです。ですから関連語（シソーラス）辞書が発達するのも頷けますし、この手の指南書が教えるように代名詞が文中で重要な働きを担うのです。

存在物を示す形は、不定冠詞（"a/an"）＋単数形形態と複数形形態です。唯一物の存在を示す形は定冠詞（"the"）＋名詞形態です。

"The whale is a mammal, not a fish."（鯨は哺乳動物であり魚でない。）：永続的に続く仮想的な存在物を表現します。従って定義に近い意味表現ができます。さらに超時間的に無変化の性質は、モノの唯一性にも繋がっています。生起した事柄も唯一の存在物です。

"A whale is a mammal, not a fish."（鯨は哺乳動物であって魚ではないのだ。）：イルカやシャチなどと区別することで、概念体系内における存在物を表現します。他と弁別する特徴表現に近い意味が表せます。一時的な存在物なのです。ですから"He is an excellent singer."は「彼は歌が上手だ。」です。また、例えば、"a little farther along the river"のように時間経過も

ひとつの存在として解釈します。

"Whales are mammal, not fishes."（鯨は哺乳動物で魚でない。）：集合化した鯨に見られる特徴表現を表します。存在は永続的でなく、もっかのところ認識できるモノの集団に共通する特徴を表します。また、"teachings"（「教え」）のように事柄も（時制を排して）集合化して存在させます。

なお、（無冠詞＋）名詞形態を使ってモノの機能や働きを表現します。また、Britain のように生活世界の中の唯一物は固有名詞と呼ばれ、無冠詞であっても唯一物は語頭文字を大文字として表します。文法上、特別な扱いを受けるのです。英語においてモノは存在することが大切なのです。

#### 3.2. 関係と役割語（日本語）

日本語は、小さい時空間内で関係的にモノを示すので、体言の意味を「機能を表す」関係物として表すことを好みます。国語辞書で、幾つか言葉の定義を確認してみてください。機能表現が定義文として記載されていることが確認できます。例えば「机」は以下のようです。機能を中心に説明されています。台の形状や脚の数や大きさについての説明はありません。

机：(1) 本を読んだり、字を書いたりするのに用いる脚付きの台。(2) 飲食物を入れた器や供物をのせる台。（三省堂 大辞林）

また、「私、俺、僕、拙者、小生」「わたし、おれ、ぼく」や「あなた、君、きみ、お前、貴様」、「彼、彼女」、「私たち、手前、手前ども」、「お父さん、お母さん、(名前)、おじいさん、おばあさん」、「先生、社長、部長、課長」など役割を示す語が大変豊かです。これらはことばを用いる状況の中で、他との立場（機能）を明らかにします。例えば、[3]は he、she の翻訳の歴史を示し、「彼女」が翻訳語として近代に造語されたことを説明しています。he が代名詞（存在するモノを代行する役割り）であるのに対して、「彼」は「発言者と聞き手から外の、遠くのことを指す」として、（発話の場における関係を示す役割の）指示詞であるといえます（[3]：197 頁）。

さらに、いわゆる男言葉や女言葉のように、役割を

連想させる形態素も発達しています。ただし「行くわよ、食べるのじゃ、飲むぜ」のように用言の後続要素として現れます。そして日本語には挨拶表現が多い（様々な表現がある）ことも指摘されています。社会生活の中でお互いの関係性を判断する道具として、また確認する道具として使っているのでしょう。

勿論、私たちはモノが存在することを知っています。ご承知のように「いる、ある」を使って有性と無性の区別をした上でモノの存在を表現します。「いる、ある」の使い分けが当然の如くできることから分かるように、私たちはモノを、その存在性だけでなく有性（有情）と無性（無情）で区別し仕掛けています。例えば、筆筒（たんす）を一棹、二棹と数えるように、助数助辞（棹）を使うことで、存在（するモノ）の形状や性質まで区別します（たぶん中国語の量詞（類別詞）の影響でしょう）。

なお、固有の存在であっても文法上特別な扱い（文中のどこにあっても語頭文字を大文字にするなど）を受けることはありません。要するに、知っていることは大事なのですが、存在するということはさして重要でないのでしょう。

## 4. モノとコトの在り方

### 4.1. モノに対する感覚・知覚・情意の表現

モノの性質・属性の表現には大きく分けて二種類あります。感覚や知覚を通じて存在すると見なせる性質・属性です。「暑さ、寒さ、高さ、重さ」などです。これらは定量化可能な性質・属性で、客観的な扱いを受けることが多く、物質世界の中で常識的なスケールを想定することもできます。

もう一つは、推論、判断や評価を通じて解釈した上で存在すると見なせる性質・属性です。「美しさ、正しさ、汚さ、悪さ、綺麗さ」などです。これらは想像上の認識的な性質・属性であって主観的で、ですから共通の基準を見いだすことが困難です。但し、社会的な行動規範や規律規範、判断水準を設けることで、定量化可能な性質・属性であると見なし、例えば「良い、悪い」のように二値的な扱いをすることもあります（といっても「良い」程度と「悪い」程度が測れるわけもなく、尋常一様な水準からの乖離の程度が、両語で同

等である保障もないのです）。

また、「騒々しい、痛々しい」など動作・行為に価値付けを行いモノの性質・属性の表現に使うことも多く見られます。「嬉しい、悲しい、可笑しい」など認知主体が持つ感情をコト化して性質・属性の表現を行います。

#### 4.1.1. 存在を表すモノの有り方の説明（英語）

英語において形容詞は、名詞で表されるモノをより限定し、他と区別する働きがあります（存在の担保）。表したいモノの存在をより明瞭にする役割を担っています。典型的に他と弁別するため種類や固有の性質（一時属性）を説明する表現を担います。

例えば、"He is a clever man." は「男は多数いるが、彼はそのうち賢い部類に属する」もしくは「賢い男は多数いるが、彼はそのうちの一人である」という意味を表します。それに対して日本語の「彼は賢い男だ。」は「彼は賢い性質をもった男（だ）」という意味です。

指示形容詞（this、that、these、those）は同時に代名詞でもあります。each、every、either、neitherなどの配分形容詞も代名詞です（everyは除く）、数量形容詞（some、any、many）なども同じ機能を持ちます。「the+形容詞」で、その形容詞が表す人々全体を代表させることができることは皆さんもご存じのことです（"the Japanese"、"the accused"（被告人））。

存在するモノが比較の対象になるので、絶対基準が用いられます。形容詞は、比較のための語形変化を持っています。一時的な、あるいは主観的な比較基準を表す表現は as を用いて示します。関係代名詞による修飾表現は典型的な一時的な性質・属性の追加のための表現です。

（日本語と反対に）永続的であることと一時的であることを、客観的な表現と主観的な表現に対応させることもできます。

#### 4.1.2. 関係（機能）を表すモノの詳細な説明（日本語）

日本語で形容言や情態言（[4]）は、体言で表されるモノを詳しく説明する働きがあります。表したいモノの概念をより明瞭にする役割を担っています。例えば「今宵（1992年9月の夜に）見た金閣寺の美しい中秋の名月」と「今宵（1992年9月の夜に）見た惜

楽園の美しい中秋の名月」において、「月」は物理的な唯一物ですが、月を存在するモノとして特定することに関心がなく、二つの表現は話し手の内面世界で（外界の月を参照して）表象された「月」を詳しく説明する役割を担っています。

二種類のモノの性質・属性の表現に、形容言（「うすい、かたい、いたい、うれしい」など）と情態言（「しずか、はるか、ほのか」など）がある程度対応しています。この区分は活用形態の違いに基づきます（[6]）。意味的な違いから属性形容詞と情意形容詞に分けることがあります。属性形容詞は連体修飾や有題文を使い客観的な属性を表します（「柔らかいケーキ」、「ケーキは柔らかい」等）。情意形容詞は連体修飾や無題文を使い主観的な属性（話し手の判断や評価）を表します（「柔らかなケーキ」、「ケーキが柔らかだ」など）。

情態言は、推論、判断や評価を通じて解釈した上で存在すると見なせる性質・属性を表します。そして、モノの性質・属性として捉えるだけでなく、コトの様態として捉えることができるほか、性質を状況であるともみなす概念を表します。ですから「静か（な）」のように副詞（「静かに過ごす」）や名詞（「静かの海」）としても使います。

（英語と反対に）客観的であることと主観的であることを、永続的な表現と一時的な表現に対応させることもできます。ただし、形容言も状態言も活用しますから、一時的であるだけでなく判断や評価の意味が加わります。

## 4.2. コトに対する感覚・知覚・情意の表現

### 4.2.1. 評価を補う（英語）

時空間の表現の大きさが大きいので wind（「曲がrikuneru、曲がって進む」）のように動作や行為の有りさま（様態）を含む語彙も多くあります。

事態として、ある程度の表現の大きさがあるので出来事の生起の蓋然性を表す副詞（頻度副詞）が発達しています。話し手の判断や評価が動詞に表れないので文修飾の副詞が発達しています。様態や程度副詞にも話し手の評価や判断を表す語彙が多くあります（法助動詞も評価を補うためと蓋然性を表現するために使われていると考えられます）。

### 4.2.2. 有様を補う（日本語）

副詞の種類が多いことや擬態語/擬音語が発達していることも用言の表現規模が小さいことが影響しています。一つの出来事の表現内に動作や行為の有りさま（様態）を含めることができず、「くねくね進む」のように、擬態語/擬音語が副詞として、動作や行為の有りさま（様態）を取り立てて表す手段になっています。オノマトペが多い特徴も頷けます。

## 5. おわりに

本稿では、英語と日本語のものごとの捉え方と違いを、ヒトの時空間把握と思考様式としての因果関係の捉え方の違いとして考察しました。紙幅の関係から「文の骨格」、「基本の文型」、「コトの中のコト（事態の包摂、分詞と to 不定詞、「の」と「こと」等）」、「働き掛け（受動や使役）」などの話題は挙げていません。因果関係の捉え方を基本に、「文の情報構造」等も含めて両言語の違いを考察することができます。稿を改めて、ご報告したいと考えています。

### 参考文献

- [1] デイヴィット・ドイッチェ著 林一訳, 世界の究極理論は存在するか 多宇宙論から見た生命、進化、時間, 朝日新聞社, 1999.
- [2] ペーテル・ヤーデンフォッシュ著 井上逸兵訳, ヒトはいかにして知恵者となったのか 思考の進化論, 研究社, 2005.
- [3] 柳父章, 翻訳語成立事情, 岩波新書, 1982.
- [4] 沖森卓也, はじめて読む日本語の歴史, ベレ出版, 2014.
- [5] H・ブラッドリ(著) 寺澤芳雄(翻訳), 英語発達小史, 岩波文庫: 岩波書店, 1982.
- [6] リ・J.ウェイリー(著) 大堀壽夫他(訳), 言語類型論入門, 岩波書店, 2006.
- [7] 稲田倍穂, 日本列島の歴史と日本人, 文芸社, 2013.
- [8] 森岡健二, 要説日本文法体系論, 明治書院, 2001.



## 機械翻訳事始め[v]

坂本義行

アジア太平洋機械翻訳協会理事

1966年に米国からALPACレポートが勧告され、全世界に衝撃をあたえたことは、前回は報告したが、今回は、その後の状況について報告したいと思います。

## 1. MT から CL へ

情報処理学会では、1961年11月に機械翻訳研究会の第1回が開催された。原則として、奇数月に一回開催された。1962年1月から、研究会が研究委員会として、委員会制で発足することになった。

そして、1967年度から、機械翻訳研究委員会は、Computational Linguistics 研究委員会、通称「CL研究委員会」に改称された。

その結果と思われるが、1967年度以降の、この委員会等で発表された論文を私が調査した範囲では、1976年度までの、ほぼ10年間で、機械翻訳という言葉が用いられたタイトルは、全タイトル75件の中で、以下の4篇に過ぎない。

- ・1967.3.18、「ロシア語科学論文タイトルの自動翻訳」、岡本哲也
- ・1967.7.15、「Phrase Structure Languageの機械翻訳への応用」、田町常夫
- ・1967.12、「電子計算機による和文英訳」、坂井利之他
- ・1970.5.16、「機械翻訳のための基礎的プロジェクト」、野崎昭弘

その他のタイトルには、日本語の分析、英語の分析、辞書について、かな漢字変換、文字発生装置、漢字の符号化、記号処理、意味処理、KWIC、質問応答といった言語の

分析とシステム化といったCL研究が主流となったことが分かる。

## 2. 学会での発表状況

1966年以降、1980年までの約15年間に、学会誌「情報処理」<sup>1)</sup>に掲載された機械翻訳に関する論文は、次の6篇だが、この中で、研究発表としての論文は、1)、3)、5)の3件だけだ。それも、おそらく、ALPACレポートが、発表される前から研究なされてきたものの報告と考えられる。また、他の論文も言語処理、自然言語、といった言葉が中心となっている。機械翻訳を中心にした研究あるいは、プロジェクトは、皆無に等しいと言える。

- 1) 「機械翻訳の一模型」、西村恕彦、Vol.8, No.2, 1967.3, No.6.
- 2) 「最近の言語処理研究について」、坂井利之、長尾真、Vol.10, No.1, 1969. 1.
- 3) 「電子計算機による和文英訳」、坂井利之、杉田繁治、渡辺明正、Vol.10, No.6, 1969.11.
- 4) 「自然言語の機械処理」、栗原俊彦、Vol.14, No.4, 1973.4.
- 5) 「専門分野を対象とした日英機械翻訳について」、首藤公昭、Vol.14, No.9, 1973.9.
- 6) 解説「機械翻訳」、長尾真、Vol.20, No.10, 1979.10.

これに対して、日本語処理、漢字処理に関する論文は、毎年のように発表され、電子計算機のハードウェアの高速化、コンパクト化にともない、ソフトウェアの充実と

あいまって、日本語特有の処理技術が発表されてきた。

- 1) 「漢字の陰極線管表示方式」、仲摩雅年他、Vol.9, No.6, 1968.11.
- 2) 「データ自動圧縮編集機能をもつ KWIC 索引システム」、植村俊亮、Vol.9, No.6. 1968.11.
- 3) 「文字列の配列順序についての問題」、西村恕彦、Vol.10, No.1, 1969.1.
- 4) 「漢字情報処理特集について」、石田晴久、浦城恒雄、Vol.10, No.5, 1969.9.
- 5) 寄書「漢字テレタイプライタの使用経験」、西村恕彦、Vol.11, No.2, 1970.2.
- 6) 「和文入力に関する一考察」、保原信、山本稔、Vol.11, No.2, 1970.2.
- 7) 「計算機意味論の実験」、西村恕彦、岩坪秀一、Vol.11, No.3, 1970.3.
- 8) 報告「計算機による標準文字図形の発生」、Vol.12, No.3, 1971.3.
- 9) 「科学技術における情報検索の諸問題」、中井浩、Vol.13, No.6, 1972.6.

日本語特有の処理システムとして、ハード、ソフトの両面での処理に関しては、

- 1) 「漢字かな混り文変換システム」、松下温、山崎晴明、佐藤文和、Vol.15, No.1, 1974.1.
- 2) 資料「タッチ打法による漢字入力」、川上晃、川上義、Vol.15, No.11, 1974.11.
- 3) 解説「姓名のカナ漢字変換システム」、田中康仁、Vol.16, No.3, 1975.3.
- 4) 解説「漢字プリンターの縦横変換方式」、面谷和生、Vol.16, No.5, 1975.5.
- 5) 解説「漢字情報処理における特殊漢字」、大倉信治、Vol.16, No.6, 1975.6.
- 6) 解説「漢字処理システムの一例 (FCL)」、柿沼悌司、Vol.16, No.8, 1975.8.
- 7) 解説「高速漢字プリンタシステムにつ

いて」、黒崎悦明、Vol.16, No.9, 1975.9.

- 8) 解説「パターン合成による漢字入出力装置」、長谷川実郎、Vol.16, No.9, 1975.9.
- 9) 解説「S5000 漢字情報処理システムにおける文字コード」、菊池利徳、Vol.16, No.11, 1975.11.
- 10) 「日本語文入力用カナ漢字変換システムの試作」、木村健、遠藤安彦、小橋史彦、Vol.17, No.11, 1976.11.
- 11) 「カナ漢字変換の一方式」、牧野寛、勝部康人、木澤誠、Vol.18, No.7, 1977.7.
- 12) 解説「漢字処理装置」、長谷川実郎、Vol.10, No.4, 1978.4.
- 13) 「国語辞書の記憶と日本語文の自動分割」、長尾真他、Vol.19, No.6, 1978.6.
- 14) 「清書プログラム」、杉原厚吉、Vol.20, No.11, 1979.11.
- 15) 「漢字姓名入力システム」、野寄雅人他、Vol.19, No.3, 1978.3.
- 16) 展望「日本語のワード・プロセッシング」、河田勉、天野真家、Vol.21, No.8, 1980.8.
- 17) 解説「日本語文字発生方式」、野村仙一、小池博之、Vol.21, No.11, 1980.11.

自然言語処理を中心に、

- 1) 「自然言語処理のためのプログラミング言語 PLATON」、長尾真、辻井潤一、Vol.15, No.9, 1974.9.
- 2) 「速記反訳システム」、絹川博之他、Vol.16, No.6, 1975.6.
- 3) 談話室「外国語データ・バンクの構想」、披田野陽一、Vol.16, No.9, 1975.9.
- 4) 「新しい記号処理概念による SNOBOL インタプリタ」、三上和敬、豊田順一、田中幸吉、Vol.16, No.6, 1975.6.
- 5) 「ストリング処理用仮想計算機 VC/S とその上での SNOBOL 3 の実現、国立勉、

- 吉田雄二、福村晃夫、Vol.16, No.11, 1975.11,
- 6) 「語彙の一分割法の提案—意味及び文脈情報を用いた日本語文の解析—名詞句・単文の処理」、長尾真、辻井潤一、田中一敏、Vol.17, No.1, 1976.1.
  - 7) 「意味及び文脈情報を用いた日本語文の解析—文脈を考慮した処理」、長尾真、辻井潤一、田中一敏、Vol.17, No.1, 1976.1.
  - 8) 「日本語文献における重要語の自動抽出」、長尾真、水谷幹男、池田浩之、Vol.17, No.2, 1976.2.
  - 9) 報告「第6回計算言語学国際会議に出席して」、長尾真、辻井潤一、Vol.17, No.11, 1976.11.
  - 10) 「自然言語処理プログラム」、長尾真、辻井潤一、Vol.18, No.1, 1977.1.
  - 11) 「日本語情報処理の過去・現在・将来」、長尾真、Vol.19, No.2, 1978.2.
  - 12) 「日本語情報処理特集号の編集にあたって」、長尾真、田中穂積、Vol.20, No.10, 1979.10.
  - 13) 「情報科学と自然言語」、長尾真、Vol.20, No.3, 1979.3.
  - 14) 「べた書き文の分かち書きと仮名漢字変換」、牧野寛、木澤誠、Vol.20, No.4, 1979.7.
  - 15) 「英語論文の清書における英単語の自動文節に関する統計的方法、朝倉秀三、Vol.21, No.1, 1980.1.
  - 16) 「日本語文構造解析による自動インデクシング方式」、絹川博之、木村睦子、Vol.21, No.3, 1980.5.
  - 17) 「推論過程を Fuzzy 化した自然言語システム」、三重野博司、戸内順一、Vol.21, No.6, 1980.11.

さらに、言語理論から人工知能への問題として、

- 1) 講座「言語理論の最近の話題 I」、笠井琢美他、Vol.15, No.1, 1974.1.
- 2) 講演「人工知能研究についての二三の問題」、西野博二、Vol.15, No.6, 1974.6.

### 3. 国立国会図書館での文献

国立国会図書館のインターネットによるキーワード検索<sup>2)</sup>によると、

「機械翻訳」のキーワードで得られた 500 件の中から、1980 年までの文献を抽出すると、以下の 40 の文献が得られた。この文献の著者を調べてみると、九州大学関係者が 15 名、京都大学が 6 名、電気試験所が 5 名と 26 件の発表がなされている。すなわち、日本では、これらの機関が中心に機械翻訳の研究が進められてきたことが分かる。

- 1) 「言語の機械翻訳—1・2—」、栗原俊彦他、電気通信学会雑誌、46(11), 1963.11.
- 2) 「日米科学協力委員会の機械翻訳セミナーについて」、渡辺修、計量国語学会、1964.06.
- 3) 「機械翻訳の現況」、野崎昭弘、情報処理月例資料：4)、1965.
- 4) 「機械翻訳の意味」、大坪一夫、言語学論叢、1965.05.
- 5) 「機械翻訳の日本語に因んで」、井上増次郎、海星女子学院大学/短期大学紀要(通号4)、1965.12.
- 6) 「言語」、大河内一男、服部四郎、他、東京大学公開講座、1966.
- 7) 「機械翻訳における品詞分類と法則抽出」、清野武、電気通信学会雑誌、50(2)、1967.02.
- 8) 「A study on mechanical translation

- from English into Japanese」、杉田繁治、博士論文、1968.
- 9) 「機械翻訳の現状」、田町常夫、エレクトロニクス 13-05、1968.05.
  - 10) 「機械翻訳における諸文法について—Chomsky の文法論を中心にして」、小熊均、都留文科大学研究紀要、1968.06.
  - 11) 「機械翻訳システムについての予測」、西村恕彦、電気試験所彙報、1968.10.
  - 12) 「機械翻訳のための英和文法の研究」、西村恕彦、電気試験所研究報告、1969.01.
  - 13) 「D-tree モデルによる機械翻訳のアルゴリズムと諸問題について」、石原好宏、九州大学工学集報、1969.08.
  - 14) 「英語文の単語および熟語に含まれる構文情報の機械翻訳について」、石原好宏他、1969.10.
  - 15) 「ソ連の機械翻訳研究」、石綿敏雄、計量国語学会、1969.12.
  - 16) 「機械翻訳システムの開発」、西村恕彦、電気試験所研究報告、1970.01.
  - 17) 「機械翻訳プログラムの作成」、西村恕彦、電子技術総合研究所研究報告、1971.03.
  - 18) 「日英機械翻訳—1—」、栗原俊彦、九州大学工学部集報、44(4)、1971.08.
  - 19) 「日英機械翻訳について—2—述部の処理についての考察」、首藤公昭、福岡大学工学集報、1972.03.
  - 20) 「日英機械翻訳について—3—単語の意味分類と意味的基本構文」、首藤公昭、福岡大学工学集報、1972.08.
  - 21) 「トランジスタ回路に関する分野における日英機械翻訳-1-」、栗原俊彦、九州大学工学集報、1972.10.
  - 22) 「機械翻訳システムの研究」、西村恕彦、博士論文、1973.
  - 23) 「機械翻訳についての一考察」、奥村宣嘉、英米文学語学研究/関西学院編、(通号2)、1973.
  - 24) 「日常言語の機械翻訳(英文)」、猪俣靖、北九州工業専門学校研究報告、1973.01.
  - 25) 「機械翻訳の研究」、渡部満彦、言語生活、1973.02.
  - 26) 「日英機械翻訳について—4—P-P 関係および修飾構造の意味処理」、首藤公昭、福岡大学工学集報、1973.02.
  - 27) 「日英機械翻訳について—5—連語の処理および翻訳アルゴリズム」、首藤公昭、福岡大学工学集報、1973.05.
  - 28) 「D-tree モデルに基づく英日機械翻訳のための言語分析について」、石原好宏他、電子通信学会論文誌、D、1974.07.
  - 29) 「日英機械翻訳について—6—述部処理システム PPSI について」、首藤公昭、福岡大学工学集報、1976.03.
  - 30) 「日英機械翻訳のための述部処理システム」、首藤公昭、鶴丸弘昭、吉田将、電子通信学会論文誌、D、1977.10.
  - 31) 「機械翻訳および機械援助翻訳(抄訳)」、W.J.Hutchins 著、中井浩抄訳、情報管理、1979.03.
  - 32) 「機械翻訳(日本語情報処理〈特集〉)、長尾真、情報処理、1979.10.
  - 33) 「機械翻訳を理解するための基礎用語」、辻井潤一、エレクトロニクス、1979.12.
  - 34) 「機械翻訳システムに関する基礎的研究」、長尾真、京都大学、1980.
  - 35) 「気象情報理解システム(ISBAR)による日英機械翻訳」、田町常夫、九州大学、1980.
  - 36) 「意味表現を仲介とする科学技術文献機械翻訳システムに関する研究」、堂下修司、京都大学、1980.
  - 37) 「機械翻訳と機械援助翻訳」、石綿敏雄、

現代の図書館 18(3)、1980.09.

38)「機械翻訳論の試み」、石原好宏、山口  
大学工学部研究報告、31(1)、1980.10.

39)「着実に進展している機械翻訳(コンピ  
ュータ言語学(特集)、長尾真、科学朝  
日、1980.10.

40)「80年代の機械翻訳システム(計算機  
はどこまで言語を理解できるか)」、辻  
井潤一、自然/中央公論社、1980.10.

以上のような調査から、ALPAC レポート  
前後にわたって、この当時、我が国では、  
計算機技術の進歩とともに、これを基に、  
自然言語処理の技術が、進展してきたこと  
が分かる。

そして、トランジスタの進歩とともに、  
開発の中心が次に述べるような日本語ワー  
プの開発に向かった。

#### 4. 日本語ワードプロセッサ JW-10<sup>3)</sup>

1978年に東芝によって開発された。初の  
日本語ワードプロセッサ、1978年9月に発  
表。出荷開始は、1979年2月、当時、コン  
ピュータで取り扱える日本語の文字は、特  
殊な機器を用いる一部システムを除き、カ  
タカナが当たり前であった。一方、欧米で  
は、タイプライタから進化したワードプロ  
セッサが広く普及しており、日本語で利用  
できるワードプロセッサが望まれていた。

東芝では、総合研究所で森健一らによ  
って基礎研究を重ねた文節指定入力による  
かな漢字変換を、ミニコン TOSBAC-40L  
を母体にした事務机サイズのハードウェア  
に搭載し、日本語ワードプロセッサ JW-10  
として発売した。文節指定入力とは、文章  
の読みを文節ごとに区切って入力する方  
法で、たとえば、「漢字を簡単に入力する」  
という文章を入力する場合、「かんじを かん  
たん に にゅうりよくする」のように3つの文

節に区切って入力する。JW-10の入力方法  
には、文節指定入力のほかに、漢字指定入  
力があった。漢字指定入力は、漢字部分  
を指定して入力する方法で、先の例の場合  
には、「【かんじ】を【かんたん】に【にゅう  
りよく】する」のように入力する。かな漢  
字変換技術の開発にあたっての最大の課題  
は、同音異義語の選択であった。結局、文  
章の前後関係、使用頻度などを利用した方  
式とした。かな漢字変換で使用する辞書に  
は、最大 80,000 語まで登録が可能で、使  
用者別に頻度管理を行っていた。標準登録  
語として辞書に登録することができたが、  
その語数は、普通単語が 10,000 語、固  
有名詞が 8,000 語であった。JW-10 が  
発表された時の新聞では、かな漢字変換  
の実用性を、「例えば... ウラニワニニワ  
トリ→裏庭に 鶏」という表現を用いて  
いた。

	JW-10	JW-10 モデル 2	JW-5
発売	1979 年 2 月	1980 年 7 月	1981 年 1 月
価格	630 万円～	340 万円～	260 万円～

東芝の研究開発チームを率いていた森健  
一氏の談話として<sup>4)</sup>、

当初、手書きの書類を読み取る装置を作  
ろうとしていたが、役人から、「欧米の役所  
の生産性が高いのは、タイプライターを使  
って文書を作っているからだ」と指摘され、  
手書きの書類を作らなくて済むように、「ワ  
ードプロ」の開発を目指した。

そのコンセプトは、

- ・手書きよりも早く日本語の文章がで  
きる
- ・欧米のタイプライターのようにどこ  
でも持っていけるポータビリティ
- ・文書ファイルに、電話線を通じてど  
こからでもアクセスできる

ここで問題なのは、「かな」あるいはロー



マ字で入力して漢字に直す時に日本語に多い同音異義語が問題です。入力したい単語を、1回目の変換でどれだけ出せるか、つまり第一変換率をどれだけ高くできるかがカギです。実験した結果、手書きより早くしようと思ったら、第一変換率が95%以上にならなければ駄目だということが分かった。

## 5. ALPAC と私の研究

### ～学生から東京大学生産技術研究所へ

1959年に大学4年生になり、卒業研究の指導を電電公社、武蔵野通信研究所の半導体研究部の川口室長の指導を受け、「トランジスタの高周波特性」の研究を行なった。卒業後、就職を川口氏の推薦で、東京大学生産技術研究所、第三部、野村民也教授の研究室で働くことになり、1960年4月に大学卒業後は、ここで働かせていただくことになった。野村先生は、当時「東大のロケット」と呼ばれていたロケットのテレメータリングの専門家でした。この後、渡米するまでの1年半の間、研究室の先輩の助手の下で、トランジスタ回路の設計を行なった。

東大での研究内容は、自動雷検出装置の製作で、落雷の際のインパルス電流を24時間監視し、磁気テープに自動記録する装置で、トランジスタ300個からなる回路の作成を先輩のもとで、お手伝いする仕事でした。当時、半導体製品は、米国がリードしていて、トランジスタ回路の設計を進めるには、米国の大学院に進んでみたいという願望が心の中に膨らんできました。そこで、大学4年のとき、「フルブライト交換留学生」と、プリンストン大学へ奨学生の申請を行ない、幸いにもプリンストン大学大学院電気工学科に奨学生として合格した。

六本木での研究が半年を迎えたとき、留学のときが来た。

### ～フルブライト交換留学生として、プリンストン大学へ～

ここからは、私の人生を決めた出会いです。アメリカで迎えた春の初め、ロシアで3番目の宇宙飛行士であるチトフ氏が大学の講演会に来ました。その講演を聴いて、私の心を悩ませることになりました。当時、私は東大のロケット打ち上げの一員でした。まだ、日本では、弾道を描くロケットを追尾し、そこから情報を得るレベルで、人工衛星、まして有人宇宙船の打ち上げなど、何十年後しか実現できないレベルだと思っていました。そこで、この私の留学の経験を生かした仕事はないものかと、考えました。

この大学には、すでに、デジタル・コンピュータが設置され、そのプログラムの書き方を授業で習いました。「今の日本にはないが、必ず、デジタル・コンピュータの時代が来るはずだ」。このコンピュータを応用した研究はないものかと日々考え、コンピュータを利用して、言葉を翻訳する研究はどうか。自分は日本語と英語なら少しはできる、そこで、日英翻訳をコンピュータで行なえないのか。このことについていろいろと調べ、日本で、その研究が始まっている、その目的のコンピュータが電気試験所で、トランジスタを使って試作されていることが分かった。日本に帰って、これを私の仕事にしよう！

### ～公務員試験から電気試験所へ～

1962年10月、留学を終えて、帰国後、公務員試験を受け、合格、電気試験所の面接を経て、1965年4月、夢にまで見た電気

試験所に入所できました。研修期間の 3 ヶ月を終え、通産省工業技術院電気試験所電子計算機部情報処理特別研究室に配属されました。

私が入所したとき、電子計算機部長が和田弘氏、研究室長が蓼沼良一氏、研究室には西村恕彦氏が在室されていて、すでに英語から日本語へ翻訳する計算機「ヤマト」が研究室に置かれていました。これは、世界で最初に作られた翻訳をするための電子計算機でした。この装置は、トランジスタ 3,000 個を使い、筆笥のような大きさで、ブラウン管にそのプログラムの状態がビットで表示されるものでした。辞書を記憶するためのドラムは、当時の電気洗濯機位の大きさでした。これで、4K ビットの記憶容量です。今では、想像もつきません。正確ではありませんが、恐らく、400 語位の英和辞書が蓄積されていたと思われます。この計算機が「ヤマト」と銘々されたのは、「I love you.」と入力すると、「ワレ ナンジヲ アイス。」と「大和言葉」で出てくるような翻訳機だからです。この翻訳機の操作を学びました。

しかし、この当時の日本人の常識は、ある日、友達と会い、一杯やりながら、今度入所した研究室の名前を教えたら、情報処理とは、下水処理みたいなもの、それは、ソフトウェアの研究だということ、それって、アンダウェアのような下着のことか、言語処理特別研究室とは、暗号解読のような研究をやるどころかと質問を返されて、往生した記憶があります。今では、想像することもできない時代での、研究のスタートでした。

そして、機械翻訳の原理が理解でき始めた矢先に、ALPAC レポートが 1966 年に出されたわけです。そして、次の年から、電

気試験所の予算項目から機械翻訳という言葉は消えてしまいました。

そして、その後の 10 年間は、カナ漢字変換、KWIC、辞書の蓄積、検索、COBOL といったプログラミングといった自然言語処理の基礎研究と以下に記すような詰将棋や点字といった応用研究を行なった。

～「将棋ゲームの指し手の理解モデルとオンライン処理について」<sup>5)</sup>～

という論文をプログラム・シンポジウムで発表した。

このゲームを考えつくまでの経過を述べてみよう。日本語の計算機による処理を考えているとき、言語と将棋の比較の話を思い出した。確か次のような比較が述べられていたと思う。

形態論－駒の動き（駒の働き）

統辞論－駒組、戦法

意味論－詰める（必死、詰めの手順、終盤の技法）

音韻論－手順の分析、統合による形態論への導入

そこで思い出されるのが、スイスの言語学者ソシュール(F. De Saussure)の一般言語学論義(Cours de Linguistique Generale, 1916)の中で語られている言葉「想像しうるすべての比較のうちで、最も適切なのは、言語の営と将棋の勝負とのそれである」(小林英夫訳)である。もちろんここでの将棋とはチェスを指している。

自然言語における意味の定義と比較して、少なくとも処理結果(ゲーム終了時; 詰み)評価は明確であるという点に興味がかかれたのと同時に、チェスに比べて将棋についてのプログラム化は 2、3 の報告を除いて、その明確な手法の報告がなされていない。さらに、TSS 画像端末による日本語

(漢字) の表示が容易に行え、concurrent file 処理が可能なシステムが利用できるということから、人間が学習するための将棋 QA システムを試作してみようと思いついたのが、このプログラム開発の動機である。

画像の制御

TSS 画像端末により局面を表示するには、TOSBAC5600 FORTRAN-Y をホスト言語として TSS 環境下で働く PLOT-10 というグラフィック・ソフトウェアを用いている。駒を表示する漢字パターンは直線で分割された最大 62 個の座標で表わされている。1 文字は文字パターンのファイル (3,000 種の文字) から漢字番号で索引し、PLOT-10 により表示する。なお、玉方の駒は漢字パターンを座標変換したものを用いている。

後 手 持 駒	歩	歩	歩	歩	歩	歩	歩	歩	歩	一
	飛								馬	二
	歩	歩	歩	歩	歩	歩	歩	歩	歩	三
										四
										五
										六
	歩	歩	歩	歩	歩	歩	歩	歩	歩	七 先
	角								飛	八 手
	香	桂	銀	金	王	金	銀	桂	香	九 駒

第 1 図 初期局面の表示例

#### ～仮名点字自動代筆システム 6)～

視覚障害者が言葉を理解するには、耳を介するか、点字を通してである。現在視覚障害者が本を読むには、直接、人 (健常者) に読んでもらうか、またはこれを録音した磁気テープを利用するという方法と点字書・点訳書を読むといった方法がある。点字とテープは、その足らざるところをあい

補い、視覚障害者の読書において、欠くことのできない役割を果たしている。

訓練された視覚障害者は、点字を放送のアナウンサーがニュースの原稿を読む以上の速度で点字を読める。しかし、視覚障害者のために点字があっても、点字で書かれた点字書は、非常に少なく、かつ高価である。年間新しく発行される通常の単行本は、2 万点を越える。新聞・雑誌を含めると、2 倍に達するとおもわれるが、点字書は 200 点前後に過ぎない。これ以外は、点訳奉仕者による点訳書で、手作りで作られるので、1 冊だけである。点字書も多くて 300 部程度である。これは視覚障害者の人口が 25 万人、そのうち視力が完全でない者が 10 万人、さらに点字を読める者が 5 万人といった少ない面もあるが、出版される点字書は、原本である普通図書に比べ、数倍から数十倍の価格にもなる。この原因は、発行部数が少ないこと、一冊の分量が多くなり、高価となるといった悪循環によるものである。たとえば、夏目漱石の「坊ちゃん」の文庫本が点字では全 2 巻に、また三省堂の国語辞典が全 32 巻にもなる。

点字と普通文字の世界を結ぶ過程に、計算機を導入することにより、正確で、高能率、経済的な処理体系、点字情報処理の実用化が待たれている。とくに、自動代筆と自動点訳のシステム開発が求められている。

視覚障害者が、直接点字タイプライタを打鍵することにより、点字紙上に点字モニタが得られると同時に、穿孔された 8 単位の紙テープが得られる。これを計算機への入力とし、プログラムにより普通文字へ変換し、出力する方法を自動代筆と呼ぶ。

点字の世界でも、普通文字の世界と同様に日本と欧米とでは、その表記法が異なっ

ており、日本語は漢字・平仮名・片仮名といった多種の文字を使用している。しかし、先天的な視覚障害者に漢字を教えることは、非常に困難であるため、日本独自の Braille の 6 点点字に、仮名を当てはめた仮名点字表記法が一般に用いられている。既報の「Braille 符号と漢字の変換処理」において、長谷川氏考案の漢字点字符号系を用いる自動代筆システムについて述べているが、汎用性の面から、本システムは、仮名点字と普通文字の漢字仮名混じり文に変換するシステムである。

ここで採用した仮名点字体系は、できる限り日本点字表記法ならびに「点字毎日」の表記法に準ずるが、54 年度に改定の表記法及び表記法と矛盾する点は修正を加えたものである。

「点字毎日」の 3 部に当たる約 10 万字を標本として入力し、on-line の TSS 端末上に処理結果を出力するシステムを開発した。なお、入力を効率的に行うために、汎用の複合用語に対する速記処理機能が付加されている。

計算機により、仮名点字文を普通文（漢字かな混じり文）に変換する自動代筆システムでは、仮名から漢字への変換での同音異義語の処理である。

本システムでは、約 8,000 語の自立語と 200 語ほどの付属語の辞書を用い、その接続情報により解析を行い、同音異字の選択を行った。標本として「点字毎日」約 10 万字を用いて、変換実験を行い、90% 以上の変換率が得られた。ただしテキストは、文節レベルの分かち書きがなされている。さらに、詳細に分野別の変換辞書を付加することにより、95% 程度の実用的なシステムの開発目標が得られた。

1966 年頃になると、日本でも、商用の電子計算機が発売されました。そこで、研究

室では、この商用の計算機を用いて、機械翻訳のシステムの開発を始めました。プログラミング言語として、COBOL を用い、英語から日本語へ、入力文として、Bell Telephone Journal という技術文書を翻訳するシステムの開発を 5 年程かけ、5 名程の研究者で進め、新しい機械翻訳システム「やまと」を NEAC2200 の計算機上で COBOL による英日翻訳システムの開発をすすめ、これが完成したのは、1968 年でした。

#### ～研究室での機械翻訳システムの開発

つぎに、日本でのコンピュータの発展とともに、情報処理の時代がやってきました。英語で書かれた論文や技術レポートが増加する一方で、これを日本語に翻訳する技術翻訳家の数が圧倒的に不足する時代を迎えました。

そこで、科学技術庁が、国家プロジェクトとして、1982 年から、いわゆる「Mu プロジェクト」がスタートするわけですが、これは、次回に述べたいと思います。

#### 参考文献

- 1) 情報処理学会、電子図書館、「情報処理」。
- 2) 国立国会図書館、国立国会図書館サーチ。
- 3) 情報処理学会、コンピュータ博物館、「日本語ワードプロセッサ」。
- 4) 産学官連携ジャーナル、科学技術振興機構、2012.9.
- 5) 第 18 回プログラミングシンポジウム、1977.1.
- 6) 第 16 回情報科学技術研究集会発表論文集、1979.10.

## The Five-Year Race to Tokyo 2020 Started at the TAUS Executive Forum. The goal: Beat Google. ([www.taus.net](http://www.taus.net))

Jaap van der Meer

---

We have heard this story many times over the past sixty years: in five years from now, the computer will produce fully automatic high-quality translations. The story goes back to January 1954 when enthusiastic researchers in IBM's headquarters in New York told the press about the first experiment with MT that they just concluded in cooperation with scientists of Georgetown University. If all goes well, the electronic brain will be ready in five years to translate all of the Soviet Union's documents into the Queen's English in just a week's time.

Now Japan has made it its story. The inspiration is more peaceful: with the Olympic Games coming to Tokyo in 2020 Japan wants to surprise the world with high-quality real-time machine translation technology. Naïve or courageous? Well, this time the chances are definitely much more positive. Dr. Minoru Etoh, CEO of the newly formed company Mirai Translate (= Future Translate), is aware of the history of MT and its disillusion. However, in an interview in the last issue of TAUS Review he said: "Olympics are a good motivation for everyone to contribute their best work. We need to truly utilize this hype."

Minoru ("Mick") Etoh unveiled the plans for Mirai Translate at the TAUS Tokyo Executive Forum hosted by Oracle Japan on April 9 and 10. As Vice President of NTT DoCoMo, he learned to see the future when he was managing artificial intelligence and machine learning teams in California and Japan. He opened his presentation (<http://www.slideshare.net/TAUS/mirai-translate-taus-tokyo-2015>) with the credo: "Impossible only means that you have still screwed up the solution." Mirai Translate has all the bits and pieces to make it work. As a joint venture of NTT DoCoMo, FueTrek and SYSTRAN, and with strong ties to NICT, the company can avail itself of multiple translation technologies, good MT training tools, a large suite of natural language processing tools and – most importantly – large volumes of language and translation corpora. The Mirai Translate start-up team of fifteen people is hiding on the third floor of a hard-to-find office building, in one of the backstreets of Akasaka, just a few hundred meters from the prestigious NTT DoCoMo office tower.

Five years to beat Google. Nobody will say it in so many words – especially not in Japan – but Google Translate has become the benchmark for everyone who is set out



to develop automated translation. The hype for the Olympics of 2020 is a great driver of course, but there is more that can help to make this a success story. First, there is funding from the government. Two ministries offer financial support to various companies and projects helping them to speed up the development and deployment of machine translation. Second, as everyone who tried knows very well, automatic translation from and into Japanese is more complex and challenging than for most other languages. Third, the need for automatic translation in Japan is more urgent than for most other regions. Seventy-two percent of the Japanese don't speak English. The urgency that is felt now to open up communications with the rapidly increasing stream of visitors and business people is putting more focus on spoken translation (S2S). Moreover, as we have seen at the TAUS Tokyo Forum, the progress that is being made on this front is already impressive.

One example of the support from the Japanese government is the consortium that is formed by Toyohashi University of Technology. Professor Isahara opened the TAUS Tokyo Forum with a presentation of this multi-year project and the partnership with TAUS (<https://www.taus.net/think-tank/news/press-release/taus-and-toyohashi-university-of-technology-partner-to-support-japanese-industry-with-machine-translation>).

Japanese businesses – large corporations as well as language service providers –

can join this consortium and share their translation memory data helping each other to train and build better MT engines. The TAUS Data repository will be used as the common platform for the hosting and pooling of data. The project looks similar to the European Commission's initiative to collect data from regional and national governments in all member states to help expand and improve the MT@EC automated translation service.

Like the European Commission, the Japanese government is not betting on one horse only. Big hopes are given to NICT, the well-known research center based in Kyoto, which will push the already impressive VoiceTra4U application to a state-of-the-art real-time communications application with support from the Ministry of Internal Affairs. Dr. Eiichiro Sumita, the director of NICT, presented and demonstrated (<http://www.slideshare.net/TAUS/nict-tau-stokyo2015>) the accomplishment of NICT, in cooperation with 29 other institutes in 25 countries under the U-STAR consortium. If you have any doubts about whether this might work (we listened to the demos at the TAUS Forum), the researchers invite you to download the free apps from Google Play and the Apple App store. The Olympic hype (and the government support) allowed Dr. Sumita to increase the efforts. He doubled the number of researchers and formed the Global Communications Plan (GCP), again with the aim to collect as many as

translation data as possible from the large global corporations in Japan.

Michael Paul, working for ATR Trek – a division of FueTrek, one of the investors in Mirai Translate – made it clear that speech-to-speech translation is not just a gadget that people want to play with on their mobile phones. He presented a case story (<http://www.slideshare.net/TAUS/1430-michael-paul>) where ATR Trek has helped a Japanese global rubber company to establish much more efficient processes in its Chinese factories with the help of a custom-built domain-specific S2S system. To complete the picture of the S2S market sector Mark Seligman of Spoken Translation (<http://www.slideshare.net/TAUS/spoken-translation-taus-tokyo-forum-2015>) and the ‘TAUS house evangelist for S2S’ gave a lively show of over the phone translation apps.

Now the question is: does the introduction and adoption of S2S applications represent a disruptive or sustainable innovation for established language service providers like Lionbridge and Moravia? Paula Shannon, CSO, SVP of Lionbridge (<http://www.slideshare.net/TAUS/lionbridge-taus-tokyo-forum-2015>), assured the audience at the TAUS Tokyo Forum that machine translation is our industry’s single biggest innovation. Lionbridge looks at all scenarios for the use of MT: where it simply helps to add efficiency or reduce costs (because it usually does), but also where it opens completely new opportunities and may hurt existing

business, i.e. disruptive innovation. Ninety-three percent of the content on the internet is in English and seventy-two percent of users prefer it to be in another language, says Paula Shannon. Machine translation serves as a force multiplier, as we recorded in the TAUS MT Market Report.

Vladimir Reiff, who is heading up sales for Moravia in Asia (<http://www.slideshare.net/TAUS/1400-vladimir-reiff>), zoomed in on the case of deploying MT for after-sales support. A clear case, as it seems, and as it has been for a long time. Famous (for the insiders) are the slides that Microsoft presented already ten years ago showing that users in almost all language regions are almost as happy with the machine translated support articles as with the human translated articles. Chris Wendt (<http://www.slideshare.net/TAUS/1120-chris-wendt>), who is heading up MT Research at Microsoft, presented updated reports about the user experience on Microsoft’s sites. The delta in the satisfaction rate between human and machine translation is negligible, except – indeed – for Japanese.

The Japanese language barrier is just a little harder to crack, as witnessed by Nikon Precision Technology. The Japanese manufacturer of wafer steppers has a long history of using machine translation technology. There is simply no other way for the engineers and product groups internally to keep up with the communications and to respond effectively

to customer needs, which soon means real-time chat translation. Andy Jones (<http://www.slideshare.net/TAUS/1150-andrew-jones>) reported on the efforts and step-by-step successes he and his team have booked in narrowing the delta between human and machine translation. And which technology are they using these days? Microsoft Translator Hub!

The Microsoft Translator Hub popped up in many presentations at the TAUS Tokyo Forum. Think about Japanese legislation. There are about 8,000 laws and regulations in Japan and every year about 100 laws are amended. In the spirit of opening up Japan, these laws and regulations should somehow be translated into English. Crestec is participating in this ambitious MT project that Yasuhiro Sekine (<http://www.slideshare.net/TAUS/crestec-taus-tokyo-forum-2015>) reported on at the TAUS Tokyo Forum. The motor behind this ambition for now is the Microsoft Translator Hub.

Exact numbers were not shared for the conference recordings, but from the conversations over lunch, dinner and tea breaks, it became clear that MT has exceeded the 40% of total production output that TAUS estimated in its recent MT Market Report. In a joint presentation of Honyaku Center (the largest translation company in Japan) and Memsource, Hiroki Kawano (<http://www.slideshare.net/TAUS/hon>

[yaku-center-taus-tokyo-forum-2015](http://www.slideshare.net/TAUS/honyaku-center-taus-tokyo-forum-2015)) and David Canek talked about the rapid rise of Memsource to become the most popular translation technology in the company (<http://www.slideshare.net/TAUS/a-path-to-partnership-honyaku-centers-adoption-of-memsource-cloud>). The success can be contributed to a large extent to the power of the cloud, which allows Honyaku Center to measure and analyze everything, but perhaps also to Microsoft Translator Hub, which populates all the empty segments when no matches are found in the memory. All the meters point in a positive direction on the company dashboard when Memsource is the tool of choice. No wonder also that Chris Wendt called Memsource his favorite translation memory tool.

In summary, we can say that the TAUS Tokyo Forum 2015 was a milestone event. All the evidence is there: In five years, Japan will surprise the world with an unprecedented translation power and a population happily conversing in the languages of the visiting tourists and businessmen. But whether the winning platforms will be all Nippon powered technologies or be cruising also on Google and Microsoft clouds, remains to be seen. Microsoft has already gained a strong position and Google is not giving up. The fancy Google offices in Roppongi Hills are very attractive for the young and bright developers and computational linguists. In the race towards Tokyo 2020, we expect that many more investments will be

made, new products and platforms will be launched and more alliances will be formed. Just as we were wondering why Gengo, the crowdsourced translation service based in Tokyo, canceled their participation in the TAUS Forum at the last minute, we were surprised by the news that the company raised a new investment of \$5.4 million from Japanese firms. Matt Romaine, who took over the helm as CEO of Gengo, mentioned Tokyo 2020 as an important motivator for the Japanese investors to step in. For the new Gengo strategy of Open Data, Matt Romaine could very well have taken his inspiration from the TAUS Forum in Dublin in 2013, where he saw Aiman Coptly presenting the business intelligence dashboard of Oracle's Worldwide Product Translation Group. The datafication of the translation industry (<https://www.taus.net/think-tank/articles/event-articles/are-translation-industry-le>

[aders-up-to-the-challenge](https://www.taus.net/think-tank/articles/event-articles/are-translation-industry-leaders-up-to-the-challenge)) was announced there and then. In addition, it is obvious of course: if you can't measure and benchmark, you don't know anything... you certainly cannot claim that you beat Google.

If you want to know more, be inspired and find out how the future of the translation industry will play out, join TAUS at the next Industry Leaders Forum in Berlin (<https://events.taus.net/events/forums/taus-industry-leaders-forum-2015>) or the QE Summit in Dublin (<https://events.taus.net/events/conferences/taus-qe-summit-2015>).

なお、本原稿の日本語訳版「Googleに打ち勝つのは誰だ。」は以下のページでご覧になれます。

<https://ja.taus.net/blog/beat-google>



## TAUS Tokyo Forum

氏名： 中村 哲三

所属：(株) エレクトロスイス ジャパン

TAUS Tokyo Forum が今年も 4 月 9 日、10 日の両日、東京、青山のオラクルタワーで開催された(日本オラクル株式会社協賛)。2010 年から始まった TAUS Tokyo Forum は、早いもので、今年で 5 年目となる(2011 年は中止)。主催者はオランダの機械翻訳系のシンクタンク、TAUS (Translation Automation User Society) である。参加者は、ツールベンダー、LSP (Language Service Provider)、関連ソフトウェア企業、クライアント、新規ビジネス開発など約 60 名にのぼった。2020 年のオリンピックに向けて、機械翻訳 (MT) への期待が高まる中、官民それぞれにさまざまな動きが活発化していることもあって、それぞれの発表に対して白熱した議論が展開された。合せて 20 のプレゼンテーションが成された。その内訳は以下のとおりである：MT 動向 2、MT サービス 3、ケーススタディ 2、研究発表 6、音声 MT 翻訳 2、周辺技術/サービス 5



まず、オラクル(株)の千明(ちぎら)さんによるホストスピーチから始まり、TAUS の主宰者である Jaap van der Meer さんの今年度のプログラム紹介に引き継がれた。

それぞれのプレゼンテーションの概要は以下のとおりである。

### ■ 4 月 9 日(木)

#### ・豊橋技術科学大学と TAUS の提携

豊橋技術科学大学(TUT)の井佐原先生が、文部科学省の援助を受けて今年の 4 月に機械翻訳 (MT) の新規プロジェクトをスタートさせたことを発表した。これは、MT の品質を向上させるために翻訳メモリー(TM)を集積し、ドキュメント関連に活用していこうというものである。バイリンガルのテキストデータや用語辞書は、MT の品質向上には欠かせない。

TUT は TAUS と提携して、TAUS の TAUS Data Repository や評価システムである Dynamic Quality Framework を活用して、日本のクライアント企業や翻訳産業がそれぞれの TM などの言語リソースを共有し、翻訳品質を向上させていくことを促進する。このプロジェクトは、2 日前の「多言語情報発信シンポジウム(TKUN)」で一般に公開された。欧米の企業と異なり、日本の各企業からはなかなか TM データの共有という発想が出てこない。そういう意味でも、このプロジェクトが日本の TM データの共有を促進し、MT 発展に貢献するであろうことは予測でき、その動向は注目に値する。



#### ・MTの多角的な活用

世界の翻訳産業を牽引するトップLSPであるライオンブリッジ社の副社長、Paula Shannonさんの発表。ライオンブリッジ社がMTをどのようにとらえ、活用しているかを説明した。ライオンブリッジ社はMTを以下のように活用している：

- クライアントが翻訳できるコンテンツ量を飛躍的に増加させる
- 翻訳コストを低減する
- クライアントが素早く世界中のユーザーに情報を提供できるようにする
- 翻訳対象ごとに異なった品質で提供する

MTは翻訳会社にとって継続的革新なのか破壊的革新なのかというジレンマをもたらす。しかし、インターネットコンテンツの93%が英語で書かれていて、そのうち72%のユーザーが母国語でそれらの情報を入手したいと思っていることを考えれば、MTがその解決策になることは見逃せない。

#### ・日英における数字表現のMT翻訳

翻訳ベンチャーである八楽(Yaraku)のリングストMike Tian-Jian Jiangさんの発表。彼はこのフォーラムの常連でもある。自社内のEBMTを使って数字表現を含む文章のMT効率化を実現した。数字が含まれたTMのファジーマッチを合成して翻訳するものである。これを実現できているCATツールはなかなか少ない中、八楽では特にCJKについて良い成果を出している。

#### ・MTと検索エンジン

これまたこのフォーラムの常連であるマイクロソフト社のChris Wendtさんの発表。自社MTを使って自社の多言語カスタマーサポートサイト作成を効率化している。目標としては、翻訳をユーザーにとってどのようにすれば最適になるかということである。なにもすべて人力で手間暇かけて作ればいいというわけではない。第一、翻訳の量があまりにも膨大過ぎて、予算や時間制限がある中でなかなか

対応できるものでもない。チェコ語の例を使って説明。ナレッジベースの重要な部分2.5%は翻訳者が翻訳している。その2.5%のところのページビューが50%を占める。それ以外のところをMTを使って約10日で翻訳した。(ただし、日本語では難しい。)

#### ・日本語をMTで翻訳できるか

半導体関連のフォトリソグラフィ装置メーカーであるニコン プレシジョン テクノロジーのAndrew Jonesさんの日英MT苦労話。毎月300万文字の技術資料を翻訳する必要がある。Wordbee TMSにMicrosoft Translator Hubを連動させてMT化しているが、日英の翻訳品質はなかなか向上しない。現時点では、日本語のドキュメントを英語ネイティブが概要を理解するのに役立っている程度。現状での問題点は以下のとおり：

- 意思が明快でない表現
- 論理に一貫性がない
- 文章の構造が稚拙(やたらExcelを使う)
- 語数が多い

#### ・Women in Localizationの紹介

このフォーラムとは直接関係ないが、午後のセッションの初めに、Women in LocalizationというLinkedIn グループについて、日本支部代表の上田有佳子さん(ネットアップ株式会社)から説明があった。Women in Localizationは、2008年にカリフォルニアで設立された、ローカリゼーション業界で働く女性が交流を深めて知識や技術を共有し合うグループ。日本支部は現在22名。定期的に交流イベントを開催している。参加希望者は、LinkedInで「Women in Localization」とその日本支部「Women in Localization Japan」の両グループに入会申請をすること。以下参照：

<<http://womeninlocalization.com/>>

#### ・アフターサービスとしての MT

急速に成長しているチェコの LSP、モラビア社の Vladimir Reiff さんの発表。同社はアジア地域でも販売を大きく伸ばしている。同社は、アフターセールスサービスを拡売のチャンスとしてとらえている。企業はアフターセールスサービスを将来に向けての投資とは考えず、コストと感じているところがある。そういった中で MT の活躍の場が考えられる。コストを抑えて MT を活用してサポートマテリアルを効率的、効果的につくることによって、LSP はクライアントをつなぎとめることができる。効果的なサポートマテリアルによって、ユーザーの CS は向上する。そして、クライアントはそれを他社との差別化に利用できるからである。



#### ・日本発の MT エンジンの可能性

(株)ヒューマンサイエンスの徳田さん、中山さんによる発表。最近、ヨーロッパ発の MT エンジンが優勢になってきているが、それはあくまでも欧州言語に焦点を当てたものであることを認識すべきである。欧米言語とは異なる構造の言語ペアの場合、つまり日英などの場合は日本発の MT エンジンを使ったほうが品質の良い翻訳が得られるというもの。①日本発の SMT (Syntax あり)、②SMT (Syntax あり)、③SMT、④RBMT の 4 つのエンジンを使って、300 ページのマニュアルを英日で試した結果、①が最も品質が良かった。

#### ・日本法令の翻訳システムづくり

日本有数の LSP、(株)クレステックの関根さんのニッチ(だけど重要!)への挑戦。法務省が運営する「日本法令外国語訳データベースシステム」

<<http://www.japaneselawtranslation.go.jp/>>というものがある。ここに、これまでに翻訳された日本法令は 2015 年の 1 月で約 500 もある。しかし、この数字は満足できるものではない。日本で施行されている法令は約 8,000 もあるからである。しかも、翻訳されたものの多くは最新の改訂が反映されていない。その改訂を適用するのに時間がかかるからだ。毎年、100 もの法令に改訂が入るといふ。改訂のたびに翻訳を直すのは無理な話である。そこで登場するのが関根さんが開発、担当する MT である「Japanese Law Machine Translation」<[http://itrd.crestec.co.jp/jlmt/default\\_en.aspx](http://itrd.crestec.co.jp/jlmt/default_en.aspx)>。SMT に TM をコンバインしたもので、2014 年 7 月からテストを開始した。現在、TM を集積中。

#### ○ お茶の会

ホストであるオラクル社のご厚意により、両日の休憩時間に、トップフロアにあるお茶室でお茶の会が催された。海外からの参加者の関心を集めた。





#### ・ ポストエディット価格の決め方

これまた日本有数の LSP、(株)十印の川原さんによる興味深いトピック。MT が翻訳実務に大きくかわりだした昨今、しかもスタンダードな価格設定が定着していない現状、ポストエディット (PE) 価格の決め方は、当の翻訳会社のみならずクライアントや翻訳者にとっても非常に気になる場所である。十印は、①MT 翻訳の目的、②PE のレベル、③MT の品質の 3 つを PE 価格設定の要素とした。①では概要がつかめればいいのか、それとも対外的に使用したいのかということがある。③はコーパスの量や元原稿の品質レベルに影響される。この 3 つの要素をベースに、それぞれの評価に従って価格を設定した。当然のことだが、PE の価格は一定ではない。

#### ・ ドキュメント発行後のポストエディット

マイクロソフト社の Chris Wendt さんの再登場。SNS ネットワークをうまく使えば、スピードとコストと品質は共存できるというもの。マイクロソフト社は、しばらく前から外部リソースを活用して PE を行っている。ウェブ上でドキュメントを発行した後、そのドキュメントの表現に問題があった場合、外部のメインユーザーから SNS で指摘をしてもらい、その指摘を内部で評価して、必要であれば内容を変更するというもの。これにより、従来、スピードとコストだけの MT 結果に品質が加わるようになった。

#### ○ ネットワーキングディナー

一日目の終了後に、赤坂エクセルホテル東急で主催者によるネットワーキングパーティがあった。リラックスした雰囲気の中で、参加者たちは交流を深めあうことができた。



■ 4 月 10 日 (金)

#### ・ HT よりも MT

ドキュメント システム ソリューションの(株)ISE の若林さんと弊職によるプレゼンテーション。日本企業であっても、各企業が有する英仏などの大量の英語 TM データを活用すれば、MT 翻訳品質を大幅に向上できるというもの。言語によっては、翻訳者が翻訳したものと同等、もしくはそれ以上の成果をあげたものもあった。また日英については、MT 品質を向上させるために、その前段階として日本文を書きなおす作業(プリエディット)を提案した。弊職は、日本語をわかりやすく翻訳しやすくするために「Plain and Logical Japanese (PLJ) 77 Rules」というルールを設けている。これは既存のルールや、日本文/英文ライターである弊職が長年にわたって蓄積してきたものをまとめなおしたものである。PLJ をご希望の方は、中村まで：  
<tetsuzo.nakamura@electrosuisse.co.jp>



#### ・(株)みらい翻訳の紹介

設立したばかりのニューベンチャー、(株)みらい翻訳の栄藤さんから、自社とそのロードマップの紹介があった。みらい翻訳はNTTドコモ、フュートレック、シストラン3社の合併企業。NICTとも連携している。その出自から、「しゃべってコンシェル」などリアルタイムでの音声を絡めたMT展開、そしてビッグデータを活用しての新規分野の開拓など前途は果てしなく広がっている。日本人の72%が英語を話せない、理解できないという絶好の(?)環境下で、どのようにMTを活用して新規ビジネス分野を開拓していけるか、他人事ながら非常に気になるところである。

#### ・翻訳センター、Memsources と提携

日本一のLSP(株)翻訳センターの河野さんとクラウド型の翻訳プラットフォーム Memsources の David Canek さんによる発表。翻訳センターは2011年に Memsources を試験的に使い始めた。そして、現在では段階的に、言い換えると部門単位で Memsources を導入している。その成果は大きく表れてきており、昨年2014年には翻訳センターは売上で94億円、そして純利益2.7億円を達成した。部門によっては従来勢力の抵抗があるものの、収益の向上という観点から、次第に社内での地位が確立されてきた。

#### ・エニドアの翻訳サービス

クラウド翻訳サービス Conyac でおなじみの(株)エニドアの Una Softic さんによる発表。クラウド翻訳サービスとは、ネットを介して翻訳を手軽に依頼(ユーザー側)して、受注(翻訳者側)するもの。現在では、45,000 人ものバイリンガル、マルチリンガル翻訳者たちがネット上で、ローカリゼーション、コピーライティング、校正作業、PE などドキュメント関連の作業に従事している。

#### ・音声MT 翻訳のその先

音声ベンチャーSpoken Translation 社の Mark Seligman さんの発表。Converser 4.0、携帯電話を使ってリアルタイムMTのデモンストレーション。いわゆる Speech-to-speech (S2S) 翻訳。MTについては Pangeanic (Manuel さん)の協力を得ている。対象言語に翻訳したものを元言語に戻すことで、その翻訳の品質を評価する(Back translation)ことも行っている。



#### ・TM Town の紹介

栃木県在住の Kevin Dias さんが、ランチ前に飛び入りで発表。TM Town—A Translation Enablement Platform という TM 利用の翻訳者登録サイトを運営している。独自のアルゴリズムを使って、翻訳者と依頼翻訳業務のマッチングが判断される。翻訳者は TM Town に自由に登録できる。そして自分にマッチした仕事が来たら、それを割り当てられる。詳細は以下のサイトに : <<https://www.tm-town.com/>>

#### ・工場現場での音声MT翻訳ビジネスアプリ

音声翻訳ベンチャーの(株)ATR-TREK 社の Michael Paul さんによる発表。この会社は前出フュートレック社の子会社。スマートフォン向け音声翻訳サービスの「しゃべって翻訳」などをつくっている。この「しゃべって翻訳」は全国 300 か所の旅行案内所で海外からの観光客に活用されている。Michael さんは、この技術を日本の某ゴムメーカーの中国工場でのコミュニケーションツールに転用した。日本人技術者と中国人労働者の間のコミュニケーション促進に活用しているのである。騒音の大きな工場で、音声認識の問題をクリアしたりする苦労があった。

#### ・U-STAR コンソーシアム

NICT の隅田先生による S2S のプレゼンテーション。多言語音声 MT 翻訳を実現するために、世界中の音声 MT 翻訳研究をネットワーク化した。これが、「U-STAR コンソーシアム」(Universal Speech Translation Advanced Research consortium)である。U-STAR は、現在、世界 25 の国や地域の 30 の研究所から構成されている。U-STAR は世界の公用語の 95% をカバーしている。ここで使われている多言語音声 MT 翻訳アプリが「VoiceTra4U」であり、アップル、アンドロイド双方で無料でダウンロードできるようになっている。プレゼンでは、初期の音声 MT 翻訳(1992 年)を紹介しながら、それと現在のものを比較させた。まさに、音声 MT 翻訳の最先端。飛躍的な進化を遂げていると言える。現在、総務省の号令のもと、研究者を 2 倍に増やして、医療、ショッピング、交通、ホテルといったシーンでの活用の研究が成されている。

#### ・TAUS Quality Dashboard

TAUS の Jaap van der Meer さんの発表とまとめ。TAUS では会員相互の MT 品質の向上を促進していくために、Dynamic Quality Framework (DQF) を提供してきた。これをさらに使いやすいものにするために Quality Dashboard として進化させた。これにより、会員相互に MT の初歩から、MT アウトプットの評価、PE のトレーニングなどの分野まで幅広く MT に関する集合知としての MT 推進環境を整備した。



TAUS は MT のファシリテーターとしての活動により、世界中に参加メンバーを拡大してきている。さまざまな分野の TM データがつまった TAUS Data Repository は今や 550 億セグメントに達している。また、Dynamic Quality Framework システムは会員相互の MT 評価や MT の貴重なナレッジベースとして機能している。TAUS に関してのご質問は、主宰者である Jaap van der Meer <jaap@taus.net> または中村までどうぞ

<tetsuzo.nakamura@electrosuisse.co.jp>

予告： TAUS Tokyo Forum 2016

開催予定日 2016 年 4 月 19 日(火)、20 日(水)



## 多言語情報発信シンポジウム報告

井佐原均豊橋技術科学大学

---

### 1. シンポジウムの概要

名称：グローバル社会への対応を目指す多言語情報発信シンポジウム

～Translation Knowledgebase and User Network (TKUN) の提案～

主催：国立大学法人豊橋技術科学大学多言語翻訳研究本部

後援：経済産業省

一般財団法人テクニカルコミュニケーター協会

一般社団法人日本翻訳連盟

アジア太平洋機械翻訳協会

日時：2015 年 4 月 7 日（火）シンポジウム 10:00～17:00

情報交換会 17:00～19:00

場所：一橋講堂中会議場（東京都千代田区）

参加者：企業、翻訳会社を中心に 130 名

（内訳概数：産業界から 60 名、翻訳業界から 60 名、大学等から 10 名）

情報交換会 約 50 名

### 2. 開催趣旨

グローバル社会を迎え、国内製品の海外販売、国内企業の海外進出、国内観光地への海外からの集客など、様々な場面で海外への情報発信の重要性が高まっている。ヨーロッパにおいては欧米言語間の機械翻訳が実用化され、様々な分野で活用されはじめている。一方、日本では機械翻訳の性能は不十分と思う人が多く、機械翻訳をビジネスで活用するという考え方が十分に浸透していない。

この状況を打破するためには、我が国の国際競争力を強化につながる効率良い多言語情報発信を実現する枠組みが必須である。ここでは個々の組織の持つ言語資源を集約することによって機械翻訳の精度を向上することはもちろんであるが、適切な文書作成や後編集も含めた多言語情報発信プロセスを規定し、機械翻訳を単体のシステムではなく、このようなプロセスの一部として取り扱うことが必要である。これを実現する枠組みとして、Translation Knowledgebase and User Network (TKUN) を提案した。TKUN に業種の垣根を越えた様々な企業・組織が結集し、それぞれの持つ知識やデータを持ち寄ることにより、効率の良い多言語化が実現できる。今回、TKUN に関わる話題を議論し、協力者を募ることを目的に本シンポジウムを開催した。

### 3. シンポジウムの成果と今後の展開

このシンポジウムは TKUN の目指す方向性を、多言語化するべき文書を持つ産業界と、多言語化を実際に行う翻訳会社に直接語りかけることを目的に開催した。シンポジウム参加者は、おおよそ、ユーザ企業から 60 名、翻訳会社から 60 名、大学等から 10 名であった。学術集会ではなく、社会実装のための広報の場とする目的は十分に達せられた。産業界からの参加者として、様々な分野の著名な企業の参加があった。翻訳会社からは、国内外のトップクラスの企業からの参加を得た。また、産業界、翻訳業界の双方の業界団体の参加もあった。我が国における多言語情報発信への興味の高まりを実感するとともに、本シンポジウムにおいて、的確な対象に向けて TKUN の有効性を伝えることができた。

今後は、このシンポジウムを後援していただいた方々をはじめ、オールジャパンの体制での多言語情報発信環境の整備を推進したい。国際的に認知されている TAUS の枠組みをデータ共有のプラットフォームとして用いることにより、産業界、翻訳業界の双方の参加を容易にする予定である。

### 4. シンポジウムの詳細

#### 4. 1 オープニング

シンポジウムの開始にあたり、シンポジウムの主催者を代表して、大西隆豊橋技術科学大学学長／日本学術会議会長が、開催のご挨拶を行った。ここでは豊橋技術科学大学の概要を紹介し、TKUN が豊橋技術科学大学の取り組むべきテーマと考えていることが示された。

続けて、シンポジウムを後援していただいた組織からのご挨拶を頂いた。はじめに経済産業省商務情報政策局情報通信機器課広瀬健治課長補佐より、我が国の成長戦略における国際競争力確保の重要性を強調するとともに、データ共有が排他的なものではなく、さらなる成長に結びつくものであり、期待するとのご挨拶を頂いた。つづけて、山崎敏正テクニカルコミュニケーター協会専務理事から、日英翻訳に対する協会会員のニーズに合ったプロジェクトとして期待しており、協会としてバックアップしたいとのご挨拶を頂いた。東郁男日本翻訳連盟会長／株式会社翻訳センター社長より、翻訳業界へのクライアントのニーズを満たすためには機械翻訳の使用が考えられ、翻訳



大西隆学長（豊橋技術科学大学）



広瀬健治氏（経済産業省）

に関わる諸団体が協力することが必要とのご挨拶を頂いた。また、中岩浩巳アジア太平洋機械翻訳協会会長／名古屋大学教授から、今回のシンポジウムが機械翻訳協会の目指す方向と一致しており、機械翻訳を含めた翻訳ワークフローの実現に期待するとのご挨拶を頂いた。



山崎敏正氏（テクニカルコミュニケーター協会）



東郁男氏（日本翻訳連盟）



中岩浩巳氏（アジア太平洋機械翻訳協会）

#### 4. 2 基調講演

シンポジウムの初めに、京都大学名誉教授・元総長／日本学士院会員の長尾真先生による「グローバル社会における多言語機械翻訳（MT）のための相互協力の必要性」と題した基調講演が行われた。グローバル時代になり世界中の人々の交流、貿易、技術輸出、等々が盛んになり、多言語間の相互翻訳の必要性がますます高まってきている。特に 2020 年東京オリンピック・パラリンピックが開催されるという状況に対処するためには翻訳関係者と企業が相互に協力する体制を作ることが重要である。東京オリンピック・パラリンピックをスポーツのみならず、文化の祭典でもあると位置づけ、京都を対象とした動きを紹介した。機械翻訳システムの現状や性能向上の方法に触れ、産業界における翻訳システムへの要求がますます増大する言語産業時代の到来を示し、言語資源の作成と共有の重要性から、TKUN の必要性を論じて頂いた。



長尾真先生（京都大学）

#### 4. 3 TKUN の紹介

一般講演として、まず井佐原均（豊橋技術科学大学教授）が「国際競争力の強化に今、求められるもの。―TKUN の提案―」と題して講演を行った。我が国産業の国際競争力の強化に必須である効率の良い多言語情報発信を実現するために、ユーザの持つデータや知見の共有に基づく新しい翻訳プロセスを実現する枠組みとして、TKUN を提案した。TKUN では、機械翻訳を用いた多言語情報発信の性能向上のため、文書作成、機械翻訳（用語辞書の充実）、後編集の 3 つのチャンスを活用する。TKUN は、機械翻訳エンジンの開発プロジェクトではなく、企業の垣根を越えてのデータ共有を行うことにより、多言語情報発信プロセス全体の効率化を目指す、協働実証プロジェクトであることが強調された。



井佐原均（豊橋技術科学大学）

次に秋葉友良豊橋技術科学大学准教授が「開発と評価のための言語資源」と題して講演を行った。ここでは、現在の主流である統計的機械翻訳（SMT: Statistical Machine Translation）の手法を概説し、SMT の改良や評価に必要となる言語資源を紹介するとともに、言語資源の大きさと翻訳性能について論じた。また、多言語機械翻訳の実現において問題となる言語資源が不十分な言語に対して、他の言語の大規模資源を利用して機械翻訳を適用する手法について紹介した。



秋葉友良（豊橋技術科学大学）

#### 4. 4 海外の動向紹介

オランダに拠点を置く翻訳の国際的シンクタンクで、本活動と協力協定を結んでいる TAUS の紹介を TAUS のディレクターである Jaap van der Meer 氏が「Introduction of TAUS and insights in the global market for machine translation」というタイトルで講演を行った。TAUS が世界的な翻訳産業のリソースセンターであり、より良いグローバルコミュニケーションの実現を目指して、翻訳産業の規模や重要性の拡大を目指していることが示された。また、TKUN プロジェクトの支援によって日本語化された TAUS のウェブページも紹介された。



Jaap van der Meer 氏（TAUS）

Microsoft Research の Chris Wendt 氏は「Microsoft Translator and TAUS Data: The importance of in-domain training material for statistical MT」と題した講演が行った。Microsoft は 2003 年から自社の技術資料を機械翻訳し、後編集なしにウェブ上で公開している。このような精度の機械翻訳の実現には、分野ごとの学習データを使用することが重要であることを事例に沿って示し、分野ごとの翻訳メモリの共有の重要性を述べた。小規模でスタートし、成功例を示し、拡張するというサイクルの重要性を述べた。



Chris Wendt 氏  
(Microsoft Research)

#### 4. 5 産業界からの多言語情報発信への期待

パナソニック株式会社の大西浩司氏は「制作現場の一例と機械翻訳“深化”への期待」と題して、パナソニックでの使用説明書の開発体制を紹介し、電子マニュアル（e マニュアル）の多様化について触れた。またグローバル展開の加速により、マニュアルの多言語同時完成が必須となり、マニュアル作成の 30%以上の時間及びコストが翻訳に費やされることを示した。機械に出来ること、機械の方が得意なことは機械に任せたいと、機械翻訳への期待を寄せた。



大西浩司氏（パナソニック株式会社）

ヤマハ株式会社の遠藤幸夫氏は、翻訳メモリの導入により英語から欧州言語への翻訳で成果をあげたこと、しかし日英翻訳ではうまく使えなかったことを述べ、多くの部署で英訳のニーズがあることから機械翻訳への期待を述べた。ドキュメントのスタイルをルールとして覚えこませることで、あるいは大量のコーパスを統計的に分析させることで、それぞれの業界にとって自然な英文に翻訳が可能となり、いろいろな場面での翻訳スピードアップが可能となり、正確なメッセージを世界に発信できると、期待を述べた。



遠藤幸夫氏（ヤマハ株式会社）



バンドー化学株式会社の岩崎英丈氏は「多言語情報発信への期待～ビジネスの強化に Plus One の連携を!～」と題して、実務において直面した問題を例示し、その解決策となりうるデータベースの必要性を述べた。TKUN への期待として、語学ハンディを減らす緩衝材としての役割を求めた。具体的には、一般的分野では「何処でも利用出来るサービス」が、専門的分野では「完成まで連続的連携したサービス」が必要とされるであろうと述べた。



岩崎英丈氏（バンドー化学株式会社）

#### 4. 6 翻訳会社の立場から

株式会社翻訳センターの河野弘毅氏が「データを使った翻訳を成功させる方法」と題した講演を行った。Translation Management System を利用した翻訳において、機械翻訳が使われる場合はおよそ半分であり、そこではマイクロソフトと Google のシェアが圧倒的である。データを使った翻訳を成功させるには、大量かつ高品質の言語資産・優れた工学技術（エンジン改良）・上手な言語資産運用の三つの要素がある。言語資産（用語集・翻訳メモリ・コーパス）は翻訳会社による人力翻訳から取り込まれる。翻訳会社は言語資産を管理する役割を担うことになるとうの見解が示された。



河野弘毅氏（株式会社翻訳センター）

#### 4. 7 国際標準化の動き

機械翻訳に関わる国際標準の制定の現状について、株式会社川村インターナショナルの森口功造氏が「機械翻訳+ポストエディットに関する国際標準化の動き」と題して講演した。現在、ISO において、翻訳（ISO17100）、制限言語（ISO24620-1）、後編集（ISO18587）など、翻訳・通訳に関連する国際標準化について議論が進められている。講演では、機械翻訳処理をした後のポストエディット（後編集）に関わるプロセス、要員の資格などを国際標準化するため議論されている ISO18587 の現状を説明した。



森口功造氏  
（株式会社川村インターナショナル）

#### 4. 7 クロージング

シンポジウムの総括として、井佐原均豊橋技術科学大学教授から、TKUN の実働に向けて以下のようなメッセージが述べられた。

##### ユーザ企業の方に

可能な範囲のデータ共有から始めて、TKUN の有効性を実感していただきたい。

##### LSP (Language Service Provider) の方に

データを用いた翻訳プロセスの向上を実現し、ユーザ企業に有効性を示していただきたい。

##### 大学・研究機関・機械翻訳ベンダーの方に

ユーザ企業や LSP に実際に提供できるシステムの開発を進めていただきたい。

##### 国や業界団体の方に

皆様のリーダーシップによって、我が国企業による理解を促し、新しい枠組みの構築を支援していただきたい。

最後に、井上光輝豊橋技術科学大学理事副学長が、講演者、参加者にお礼を述べるとともに、機械翻訳の社会実装に向けて企業の皆様にご支援いただきたいとご挨拶した。



井上光輝理事副学長（豊橋技術科学大学）

#### 5. 終わりに

豊橋技術科学大学では多言語翻訳研究本部を設置し、TKUN の活動を推進してまいります。皆様のご支援をよろしくお願い申し上げます。

## 第 14 回日中自然言語処理共同研究促進会議(CJNLP2014)参加報告

アジア太平洋機械翻訳協会

中岩 浩巳, 長瀬 友樹

### 1. はじめに

日中自然言語処理共同研究促進会議は日本と中国の言語情報処理の研究者が情報交換する目的で 2001 年から毎年、日中交互で開催している招待ベースの国際会議です。2014 年は中国開催の年となり、2014 年 10 月 12 日、13 日に四川省の省都である成都にて開催されました。

本報告では、同会議の様子を発表内容を中心に紹介します。

### 2. 参加機関

中国での開催ということもあり、中国側の参加機関は中国の自然言語処理の研究開発を代表する大学・研究機関・企業からの参加がありました。具体的には、企業としては、マイクロソフトリサーチアジア、百度、富士通研究所北京、リコーソフトウェア研究所北京、大学からは、北京大、精華大、復旦大、哈爾濱工大、重慶大、東北大、厦門大、蘇州大、西華大、北京交通大、北京郵電大、北京言語大、大連理工大、江西師範大、研究機関からは中国科学院、中国科学技術情報研究所から 30 人の参加がありました。日本からは東北大、筑波大、豊橋技科大などの大学、JST、NICT の独立行政法人、富士通研、NTT の企業を中心に 19 人参加しました。

### 3. 会議概要

中国語情報処理学会会長の李生教授と、AAMT 初代会長の長尾真京大名誉教授からのご挨拶のあと、日中双方から様々な発表がありました。

招待講演としては、蘇州大の周国教授による中国語の談話構造アノテーションに関する講演と、マイクロソフトリサーチアジアの辻井先生によるバイオテキストの構造化に関する講演がありました。

一般公演は、オーラル 17 件とポスターセッション

22 件がありました。中国側の主な発表は、精華大における自然言語処理研究（語義曖昧性解消、中国語表現パターン抽出、言語横断情報検索等）、実時間 Web 検索を用いた機械翻訳（マイクロソフトリサーチアジア）、アノテーションの自動アダプテーション（中国科学院）、言語理解のためのエンティティ関係付与と関係抽出（中国科学院）、スケルトンベース翻訳（東北大(中国)）、百度自然言語処理グループにおける研究開発、統計翻訳とテンプレート翻訳を融合した中日機械翻訳（大連理工大）などで分野は多岐にわたります。

AAMT からは、AAMT の概要紹介と、最近の委員会活動の紹介を行いました。



オーラルセッション会場

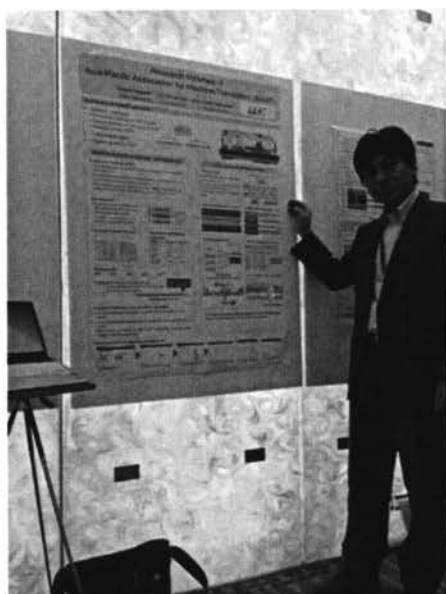
### 4. AAMT 委員会活動の紹介

AAMT からの発表は、10 月 12 日のオーラルセッションで、筆者（中岩）より、AAMT テストセットを用いた MT 出力の自動誤り分析、及び、用語集を簡単に作成・共有・再利用することを目的とした用語集形式 UTX の紹介を行いました。AAMT テストセットが中日翻訳エンジン評価のためのツール

であり、また、UTX が日本語、英語に加えて中国語をサポートしたことで、中国からの参加者にも興味を持って聞いていただけたようでした。

翌日のポスターセッションでは、AAMT テストセットによる自動評価サイトと、UTX 形式のチェックプログラムの実演デモを筆者（長瀬）が行いました。レンタルの Wifi ルータから日本のサーバへの接続が必要でしたが、問題なくデモが動作しました。会場に机の用意がありませんでしたが、前日のポスター発表者から PC スタンドを借りることができました。90 分のセッションで、7,8 名に対してデモ説明と質疑の対応をしました。

質疑の中で、エンジンベンダがテストセットにチューニングしてしまうことを危惧するコメントがありました。対処方法として、テストセットに大量のダミー文を混ぜて本物のテスト文がわからなくしていることを説明しました。今後の進め方として、特許や契約書など特定文書向けのテストセットの開発を検討中であることを伝えました。



ポスター発表の様子

UTX に関しては、複数の人から多訳語の扱いについての質問がありました。UTX は部門やプロジェクトといった比較的小さな組織を単位として、専

門性の高い用語を記述する枠組みであるため、多訳語の問題は起き難いことを説明しました。

## 5. 所感

筆者は本会議の初期から度々参加していますが、回を追うごとに、中国側の研究開発のレベルが向上している印象があります。これは、現在の自然言語処理に関連する主要国際会議における中国からの発表件数の多さを反映しているものであるともいえます。日本は、20 世紀までアジアにおける自然言語処理研究のリーダー的存在でしたが、残念ながら、今は中国に逆転されているといっても過言ではありません。これは現在の中国の経済状況を反映しているものでもあるのでしょうか。その意味でも、日本の研究開発は中国との連携が今まで以上に重要であり、本会議の意義がさらに重要になってきているのではないのでしょうか。

## 6. 次回の本会議

前述の通り、次回は日本開催の番となりますが、本会議の最後に青森で開催されることがアナウンスされています。

# Japio 世界特許情報全文検索サービス (Japio-GPG/FX) の紹介

長谷川雅昭、大塩只明

一般財団法人日本特許情報機構

## 1 はじめに

一般財団法人日本特許情報機構（以下「Japio」という）では、2012年7月に主要国・機関の特許文献を日本語と英語の両方で横断的に検索できる「Japio 世界特許情報検索サービス」（以下「Japio-GPG」という）の提供を開始し、AAMT ジャーナル No.55 (March 2014) でご紹介させていただきました。

また、2014年8月より、主要国・機関の特許明細書全文を蓄積データに加えた「Japio 世界特許情報全文検索サービス」（以下「Japio-GPG/FX」という）※1の提供を開始しています。今般、Japio-GPG/FX のご紹介のため誌面を頂きましたので、その特長、概要、使用している機械翻訳技術、今後の予定について、ご説明します。

※1 Japio-GPG/FX は、株式会社発明通信社と Japio が協同で提供するサービスです。



[Japio-GPG/FX トップページ]

## 2 Japio-GPG/FX の特長

### (1) 世界の特許情報を横断検索

Japio-GPG/FX は、日本、中国、米国、欧州と国際特許出願（以下「PCT」という）の特許明細書全文を日本語及び英語（機械翻訳含む）により、一括で横断検索できます。また、前記以外の国・地域・機関の特許出願も発明の名称、要約文の英語※2を日本語（機械翻訳）で検索できます。

※2 欧州特許庁 (EPO) が提供している全世界規模の特許誌情報データベース（以下「DOCDB」という）の英語情報を使用しています。

### (2) 日本語ダイレクト検索

あらかじめ機械翻訳された日本語を直接検索することにより、日本語の特許公報を検索する感覚で、世界の特許情報を検索できます。

また、機械翻訳された日本語は検索結果にも表示されますので、迅速に閲覧できます。この機械翻訳された日本語は、段落ごとに原文（英語、中国語）と対訳形式で表示しますので、機械翻訳文と原文を対比しながら閲覧できます。



[全文表示画面の一例]



### (3) 便利な検索支援機能

技術用語の異表記を知ることができる「用語検索支援機能」、出願人や発明者の表記バリエーション（表記ゆれ、省略形などの異表記）を確認できる「出願人検索支援機能」、「発明者検索支援機能」の各機能により、検索作業をサポートします。

#### ア. 用語検索支援機能

検索式に用いる技術用語の作成支援のため、入力したキーワードから日本語や英語の技術用語の関連語を抽出してリスト表示できる機能です。これは、Japio 独自の言語資源とノウハウを活用したもので、日本語・英語での技術用語の異表記語・関連語の把握と、漏れが少ない検索に役立ちます。

#### イ. 出願人検索支援機能

出願人名を検索し、日本語・英語・中国語での表記バリエーションを参照できます。その参照結果から自動生成される検索式を、コピー&貼付けだけで検索に利用できます。

#### ウ. 発明者検索支援機能

発明者名を検索し、日本と英字での表記バリエーションを参照できます。特に、外国人の発明者名をカタカナで検索して、その英語表記を調べることができます。

## 3 Japio・GPG/FX の概要

### (1) 蓄積データ

蓄積データは、各国・機関の特許公報、日本国特許庁（以下「JPO」という）及び独立行政法人工業所有権情報・研修館（以下「INPIT」という）が作成した和文抄録、EPO が提供している DOCDB 等を使用しています。

以下に、蓄積データの概要を説明します。

- ・米国、中国、欧州、PCT に特許公報全文とその英日機械翻訳した日本語を蓄積しています。また、日本の特許公報全文（日本語）とその機械翻訳した英語を蓄積しています。

- ・米国、欧州、中国の特許については、JPO 及び INPIT が作成した和文抄録（米国特許明細書と和文抄録、米国公開特許明細書と和文抄録、欧州公開特許明細書と和文抄録、中国公開特許和文抄録）も蓄積しています。

収録国	蓄積開始年 (公報発行年)	使用データ ※データ種別により蓄積期間は異なる
中国(CN)	1985年	中国公報データ、中国公開特許和文抄録テキストデータ、機械翻訳データ、代表図面、全図面、DOCDB
日本(JP)	1993年	日本公開特許公報データ、公表公報データ、日本特許公報データ、日本実用新案公報データ、機械翻訳データ、代表図面、全図面、PAJ、DOCDB、整理標準化データ
アメリカ(US)	1979年	米国出願公開特許データ、米国登録特許データ、米国公開特許和文抄録データ、米国特許和文抄録テキストデータ、機械翻訳データ、代表図面、全図面、DOCDB
欧州(EP)	1978年	欧州特許公開出願公報データ、欧州登録特許公報データ、欧州公開特許和文抄録テキストデータ、機械翻訳データ、代表図面、全図面、DOCDB
PCT(WO)	1993年	PCT国際出願データ、日本公表・再公表公報データ、機械翻訳データ、代表図面、全図面、DOCDB
韓国(KR)	1978年	DOCDB、代表図面、機械翻訳データ

#### [主要収録国データの提供範囲]

- ・前述の DOCDB は、主要国を含む 100 の国・地域と機関の特許情報を収録しています。これら特許情報には、特許公報の書誌情報（発行国、公報種別、出願番号、文献番号、出願日、文献発行日、出願人、発明者等）、特許分類（IPC、CPC、ECLA）、発明の名称、要約が含まれます。また、発明の名称、要約が英語である場合は、機械翻訳の日本語も合わせて蓄積しています。
- ・日本については、INPIT 作成した整理標準化データから FI、F タームを蓄積しています。

[検索対象国の設定画面]

## (2) 検索

### ア. 検索項目

検索項目は、書誌情報（文献発行国、公報種別、出願番号、公報番号、優先権主張番号、出願日、公報発行日、出願人、発明者）、特許分類（IPC、CPC、ECLA、FI、Fターム）、発明の名称（英語、日本語、中国語）、要約文（英語、日本語、中国語）、請求項（英語、日本語）、詳細な説明（英語、日本語）です。なお、英語と日本語は機械翻訳されたものも対象となります。

### イ. 検索方法

検索方法は、予め検索画面に検索項目が窓表示された「項目検索」と、質問式（クエリ）を入力する「詳細検索」の二種類があります。

項目検索では、項目間はAND条件（IPC、CPC、ECLA、FIのみOR検索）で、同一項目ではAND、OR、NOT検索を指定します。また、括弧も使用可能です。

詳細検索では、項目名、AND、OR、NOT、括弧を用いて質問式（クエリ）を作成・入力します。

## Japio世界特許情報全文検索サービス

Japio-GPG /FX

「検索履歴」 「マニュアル」 「ポータル」 「しおり表示」 「検索式表示」 「検索履歴表示」 「検索支援(用語)」 「出願人・発明者」
ログアウト

### a. 国・公報種別 選択エリア

**国・公報種別選択**

☒ 中国 (CN)  
☒ 特許公開  
☒ 特許公告・登録  
☒ 実用新案公告・登録

☒ 韓国 (KR)  
☒ 特許公開  
☒ 特許公告・登録  
☒ 実用新案公告  
☒ 実用新案公告・登録

☒ 日本 (JP)  
☒ 特許公開  
☒ 特許公告・登録  
☒ 実用新案公告  
☒ 実用新案公告・登録

☒ アメリカ (US)  
☒ 特許公開  
☒ 特許登録

☒ 欧州特許庁 (EP)  
☒ 特許公開  
☒ 特許登録

☒ PCT (WO)  
☒ 国際公開

☒ 5序+PCT  
 選択/選択解除

国指定選択 クリア

### b. 項目検索エリア

検索
クリア

検索項目	入力例
<input checked="" type="checkbox"/> 発明の名称 (TIE/TID)	AND検索 inkjet printer インクジェット プリンタ インクジェット inkjet
<input checked="" type="checkbox"/> 要約 (ABE/ABJ)	OR検索 laser+inkjet レーザー+インクジェット インクジェット+inkjet ※中国・日本クレームは英語デーク無し
<input checked="" type="checkbox"/> クレーム (CLE/CLJ)	
<input checked="" type="checkbox"/> 詳細な説明 (DSE/DSJ)	
出 願 人 (PAE/PAJ)	nippon patent "nippon patent"
発 明 者 (INE/INO)	yamada taro "yamada taro"
I P C (IC)	IPC FI ECLA CPC A01B1/00 A01B1/ A01B
F I (FI)	FI A01B13/02A
Fターム (FT)	Fターム 2B033AA01 2B033AA 2B033
E C L A (EC)	※IPC・FI・ECLA・CPC等はOR検索
C P C (CP)	
出 願 日 (AO)	20100101 201001 2010
公報発行日 (PO)	
出願番号 (AN)	
文獻番号 (PN)	US2008028494A1 US2008028494
優先権主張番号 (PR)	

検 索
ク リ ア

### c. 詳細検索エリア

検索クエリー

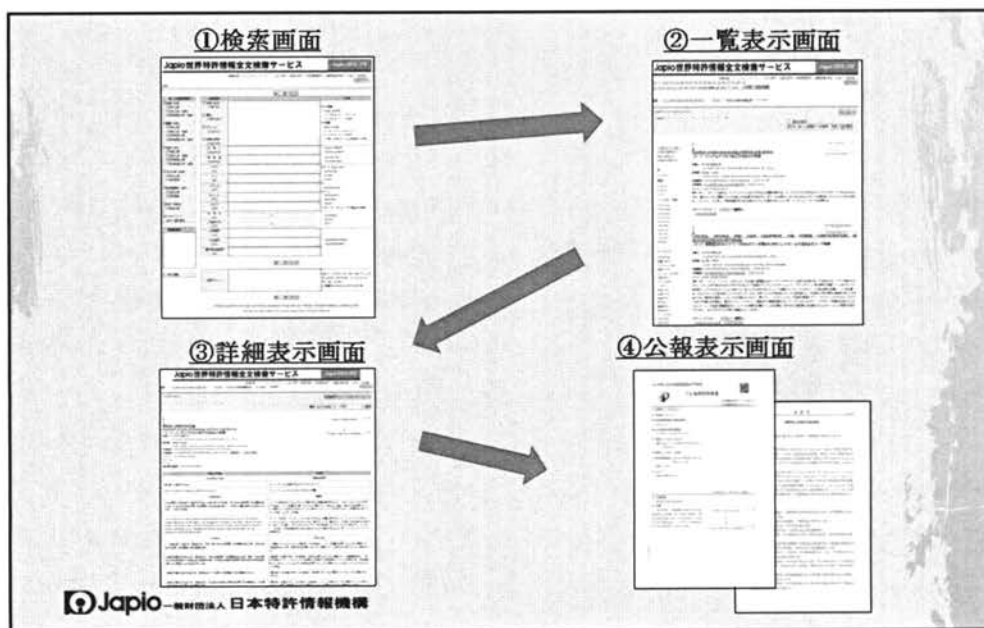
ABJ:インクジェット AND ABJ:プリンタ  
 IC:A01B\* AND (ABJ:インクジェット  
 OR ABE inkjet)  
 ※漢字(AND,OR,NOT)は大文字

検 索
ク リ ア

Japanese through the Japio Machine Translation System powered by NICT.

Copyright © Japan Patent Information Organization All Rights Reserved.

[検索画面]



[検索画面から各検索回答表示画面への操作のながれ]

#### ウ. 検索回答表示画面

検索回答表示画面は、「一覧表示」、「詳細表示」と「公報表示」の三種類があります。

- ①「一覧表示」画面は、一画面 10 件単位で、発明の名称、出願人、発明者、出願番号、出願日、文献番号、発行日、特許分類 (IPC、FI、F ターム)、要約、パテントファミリー、代表図面を一覧形式で表示します。また、この画面から、検索結果の文献情報に含まれる公報発行国、分類、用語等を用いた「絞込検索」や、検索式を手修正した「再検索」等を行えます。
- ②「詳細表示」画面は、一画面 1 件ずつ、全ての一覧表示項目及び CPC、優先権主張番号を表示します。日本、米国、中国、欧州、PCT については、全ての請求項、詳細な説明も表示します。また、この画面から、全図面、経過情報 (欧州のみ) を表示できます。
- ③「公報表示」画面は、特許公報のフロントページ又は全ページを PDF ファイルで表示します。

#### エ. 検索・表示支援機能

##### ①絞込インデックス

検索結果に含まれる公報発行国や公報種別、IPC、キーワード、発明者、出願人の各項目に対し、絞込を希望する項目をワンクリックすることにより、検索結果の絞込が瞬時にできます。検索用語等が文献頻度付きで表示されますので、容易に検索結果の絞込や分析ができます。

##### ②代表ファミリー表示

同じ発明が複数の国に出願されたいわゆる「パテントファミリー」を有する場合、検索回答が重複して表示されます。この重複表示を避けるため、「代表ファミリー表示」においては、パテントファミリーの中の代表的な言語 (英語、日本語等) を選択した表示ができます。この機能によりスクリーニングする文献数を削減し、検索効率を向上することができます。

#### オ. その他

Japio-GPG/FX は、以下の機能も備えています。

- ①しおり付与、表示
- ②検索履歴の保存、表示、再利用（セッション内有効）
- ③検索式の登録、表示、再利用
- ④ファミリー分析
- ⑤検索結果のソート
- ⑥ダウンロード（最大 1,000 件の検索結果の書誌情報及び要約文が対象）
- ⑦ファイル出力（全項目、1 件単位）
- ⑧翻訳 Web サービス（有料オプション）

テキストデータもしくは公報について、最新の Japio 辞書（約200万語の技術用語から厳選した 36分野の分野別機械翻訳辞書）を使用してリアルタイムに日英・英日の機械翻訳が行えるサービスです。

#### 4 Japio-GPG/FX の機械翻訳

##### （１）二系統の機械翻訳システム

Japio-GPG/FX は、英日・日英機械翻訳に「ルールベース方式機械翻訳（RBMT）システム」を、中日機械翻訳に「統計方式機械翻訳（SMT）システム」を使用しています。

##### （２）ルールベース機械翻訳（RBMT）システム

英日機械翻訳は、市販の特許文献用 RBMT エンジンに、Japio 独自の分野別機械翻訳辞書を組み合わせることで、翻訳精度の向上を図っています。この分野別機械翻訳辞書は、英語と日本語の対応特許文献から採取した技術用語対訳データベースを利用して作成し、各分野で最も使用頻度が高い訳語を採用することで的確な翻訳結果が得られるよう工夫しています。また、複合語を多数採用して単語レベルでは適訳語の特定が困難な用語への対応を行っています。

##### （３）統計方式機械翻訳（SMT）システム

独立行政法人情報通信研究機構と Japio との共同研究により、中国語と日本語の対応特許文献か

ら大量の中日対訳文対を作成しました。この大量の文対を機械学習させた SMT システムに用いることで、高精度な翻訳を実現しました。SMT システムは、学習データを統計的に処理して前後の語とのつながりの良い訳語を選定しますので、分野別機械翻訳辞書がなくても適切な技術用語が選定される確率が高いという利点があります。

##### （４）RBMT と SMT の併用効果

RBMT システムと SMT システムを併用することにより、情報提供までのタイムラグを短縮します。例えば中国文献では、SMT システムの中日翻訳で日本語に機械翻訳し、数ヶ月後に発行される英文抄録（CPA）を RBMT システムの英日翻訳で日本語に機械翻訳します。これら二種類の日本語情報を利用することにより、検索漏れが少ない日本語による検索が可能となります。

##### （５）訳質の改善

Japio では、機械翻訳文の訳質向上のため、機械翻訳辞書や中日対訳文対の充実や未知語対応など継続的な改善を図っています。

例えば、機械学習と翻訳に時間がかかる SMT の翻訳結果に残る未知語を解決するため、後処理用の辞書を作成して日本語に置き換えて表示しています。また、新たに公開された特許文献の対訳文を SMT 学習に加えて翻訳精度の向上を図っており近日中に中日 SMT システムを最新バージョンに入れ替える予定です。

##### （６）新たな言語対への対応

英日・日英・中日の話をしてきましたが、Japio-GPG/FX ではフランス語、ドイツ語、韓国語の特許に関しても、今後対応を予定しています。これは外部の機械翻訳資源を利用しているのですが、Japio はこれらの言語に関しても特許文献の対訳文を作成しますので、SMT 翻訳を利用することで、英語と中国語以外の言語に関しても、タイムラグの短縮と翻訳品質の向上を実現して、世界各

国の特許情報を英語と日本語で検索・参照できるように努めます。ご期待いただきたいと思います。

## 5 Japio-GPG/FX 今後の展開

Japio では、長年の特許情報事業を通じて培った機械翻訳技術を活かした Japio-GPG を 2012 年 7 月に提供開始して以来、利用者の皆様からの叱咤激励にお応えするべく蓄積データの拡大と機能向上を図ってまいりました。

公報全文データに対応した Japio-GPG/FX では、検索機能の充実にも取り組み、サービスの利便性の向上にも務めております。

また今後は、外部データベースとの連携機能や原語検索機能の充実なども予定しております。



## 特許版・産業日本語の取組（特許ライティングマニュアルについて）

早川 貴之

一般財団法人 日本特許情報機構

## 1. はじめに

我が国産業が持続的な発展・成長を実現するためには、絶え間なく新しい知的財産を創出し、その知的財産を無駄なく活用していくことが不可欠である。

研究開発や産業活動の過程で日々生まれる情報や知識を、本稿では「産業・技術情報」と総称するが、文書化されたものを例示すれば、研究論文・科学技術記事・特許明細書・設計書・技術仕様書・製品マニュアル類などが該当し、インターネット上で流通する膨大なコンテンツも含まれる。これらの情報は、日々大量に生み出されており、多種・多様かつ膨大なものである。

これら産業・技術情報は、社会ニュースや一般ビジネス文書と比べて、文法的に難解・用語定義が不明確・言い回しが分野特有などの理由により、その情報の受け手（利用者）に対して正確かつ円滑に伝達できないことが少なくない。また、コンピュータへの伝達がし得ず、効率的な情報処理を妨げる要因にもなっている。

これら産業・技術情報を、正確かつ円滑に伝達できるよう記述する手法が広く実現すれば、誰もがその情報へ容易にアクセスして活用でき、知的生産性の向上が期待できるのではないだろうか。

## 2. 産業日本語プロジェクト

これら産業・技術情報に、「産業日本語」（「産業・技術情報を人に理解しやすく、かつ、機械（コンピュータ）にも処理しやすく表現するための日本語」と定義される。）を用いる取組が「産業日本語

プロジェクト」として知られている。

「産業日本語研究会」を母体として、情報を正確かつ円滑に伝達できるような情報発信力の強化と、コンピュータ処理されやすいような知的生産性の向上に資する産業や科学技術の記述に用いられるべき日本語の枠組みのあり方について議論が行われている。この研究会は、平成 21 年度から毎年度に「産業日本語研究会・シンポジウム」を開催するなど、産業日本語に関わる提言や研究成果の公表・普及にも取り組んでいる。

[産業日本語研究会ウェブサイト]

<http://www.tech-jpn.jp/>

[産業日本語研究会・シンポジウム]

<http://www.tech-jpn.jp/symposium/>

産業日本語の基本的な考え方は、産業・技術情報に求められる機能や性格に応じて、その情報の記述に用いる日本語文の体裁・構造・使用する語句等を、あるルール（仕様）のもとで管理し標準化することにより、その日本語文が「伝えるべき情報を正確かつ円滑に伝達する」ことにある。また、これら産業・技術情報を、機械翻訳・情報探索をはじめとする言語処理技術を活用して機械処理しやすくすることも目指している。

## 3. 特許版・産業日本語

特許分野へ産業日本語を応用することを目標として、一般財団法人 日本特許情報機構（以下「Japio」という）では、平成 19 年度から「特許版・産業日本語」の活動を推進している。

特許明細書をはじめとする特許文書は、複雑な

係り受け構造の長文で記述したり、特許特有の言い回しで表記したり、一般化された抽象表現を多用したりするなど、他分野の文書とは明らかに異なる特徴を持っている。また、企業活動のグローバル化を反映し、欧米に加えて BRICs や ASEAN 諸国など新興国への特許出願が増加しており、特許文書が多言語へ翻訳されるニーズは年々高まってきている。

つまり、産業日本語の特許分野への応用は、産業・技術情報の利用性の高度化や効率化を図りやすい分野であると考えられる。特許文書を産業日本語で作成する出願人・弁理士は、諸外国への特許出願に伴う翻訳コストを軽減できるだけでなく、特許文書の内容が明晰・明瞭となることにより、諸外国における特許審査での拒絶を容易に回避し、早期の権利化や、権利化後の第三者との紛争予防といった大きな利益を得られると期待できる。

先述の「特許版・産業日本語」では、知的財産・特許翻訳・情報工学・言語処理等の専門家の方々のご協力を得て、特許文書の内容が明晰・明瞭となる日本語文のあり方、言い換えれば特許文書ライティング技術の向上を目指してきた。

また特許文書は、技術的観点と法的観点からも明晰・明瞭であることが求められる。実際の特許審査で通知された拒絶理由通知書を大量に収集し、拒絶理由の適用条文や通知内容の具体的態様の分析を試みている。これにより特許文書の作成で重点的に点検すべき課題等を発見・整理することを目標としている。

「特許版・産業日本語」の活動詳細については以下の報告書を参照いただきたい。

【特許版・産業日本語委員会報告書】

<http://japio-tjp.org/doc.html>

#### 4. 特許ライティングマニュアル（初版）

ここでは、「特許版・産業日本語」での研究成果

の一つである『特許ライティングマニュアル（初版）』（日本特許情報機構 特許情報研究所・平成 25 年 6 月刊行）をご紹介します。

このマニュアルは、特許文書を記述する日本語文を改善するための初期段階として、日本語文の言い換え規則を整理しルール化したものである。特許文書を作成する出願人・弁理士に限定されず、翻訳者・知財管理者・契約担当者・訟務担当者・知財研究者など広く関係者に利用いただけることを目指している。

この初版は、利用者の負担感を軽減するために、B5 版 32 頁の小冊子とした。言い換え規則を大きく 8 のカテゴリーに分け、【図 1】に示すように、それぞれキャッチコピーを付加している。

A	長文「分かりにくい」を避ける。 (ルール 第A条の1～4)	p 04
B	不整列「雑然とした並び」を避ける。 (ルール 第B条の1～4)	p 07
C	省略「落し穴」を避ける。 (ルール 第C条の1～5)	p 10
D	多義「どの意味なの？」を避ける。 (ルール 第D条の1～3)	p 13
E	非論理的「信号がない」を避ける。 (ルール 第E条の1～7)	p 16
F	非均質並立表現「ばらつき」を避ける。 (ルール 第F条の1～2)	p 20
G	冗長「蛇行しすぎ」を避ける。 (ルール 第G条の1～4)	p 23
H	難解「??」を避ける。 (ルール 第H条の1～2)	p 29

【図 1】8 つのカテゴリー

各カテゴリーは、より詳細なルールに細分化し、【図 2】に示す 31 のルールに整理した。これらルールは、単に列記するのではなく、改善を要する日本語文を例示し、その「改善例」を多く盛り込むよう心がけた。また、その「改善例」を一般的な文構造に置き換える「構造比較」を補うことにより、特許明細書の作成実務に関わる者にとっても、適用例が分かりやすいように工夫している。

【表 1】ルール一覧（31 ルール）

A	Aの1	長文の複文を連文（複数の短文）にし、各文の因果関係が明確になるよう言い換える。 前半（原因）と後半（結果）をそれぞれ独立した文にし、因果関係を示す接続詞で結ぶ。
	Aの2	長文で列挙される要素について、その説明を後続の文で行うように言い換える。
	Aの3	独立した複数事象の複文を連文（複数の短文）に言い換える。
	Aの4	複雑な条件設定表現を分割し、連文（複数の短文）に言い換える。
B	Bの1	主題成分を先頭に配置し明示する構文に言い換える。
	Bの2	修飾要素を被修飾要素の近くに置く表現に言い換える。
	Bの3	例示成分の配置を調整し、目的語と述語とが近くになるように言い換える。
	Bの4	主格（主語）成分を簡潔にするため、連体修飾節を運用節に言い換える。
C	Cの1	省略された主語を明示する表現に言い換える。
	Cの2	省略された目的語を明示するよう言い換える。
	Cの3	省略しすぎた表現を適正表現に言い換える。
	Cの4	一般的過ぎる用語を具体的な用語に言い換える。
	Cの5	指示語をそれが具体的に指すものに言い換える。
D	Dの1	格助詞「の」の多義性を解消するために、その意味を明示する表現に言い換える。
	Dの2	格助詞「で」の多義性を解消するために、その意味を明示する表現に言い換える。
	Dの3	格助詞「の」「で」の多義性を解消するために、その意味を明示する表現に言い換える。
E	Eの1	主題成分を明示するために、読点で区切る表現に言い換える。
	Eの2	条件節（運用節）を明示するために、読点で区切る表現に言い換える。
	Eの3	長い修飾成分を明示するために、読点で区切る表現に言い換える。
	Eの4	修飾成分が隣接する名詞と異なる名詞を修飾する場合は、読点で区切る表現に言い換える。
	Eの5	文修飾の副詞を明示するために、読点で区切る表現に言い換える。
	Eの6	複文表現において各要素の因果関係が明確になるよう言い換える。
	Eの7	一般的過ぎる動詞の使用を避ける表現に言い換える。
F	Fの1	並立表現の並立要素が同じ表現になるように整える。（語レベル）
	Fの2	並立表現の並立要素が同じ表現になるように整える。（文レベル）
G	Gの1	冗長な表現を簡潔な表現に言い換える。
	Gの2	語調を整える表現を削除し、冗長表現を適正な表現に言い換える。
	Gの3	過剰な補助動詞を削除するよう言い換える。
	Gの4	重複する表現要素を整理し適正な表現に言い換える。
H	Hの1	難解表現を平易表現に言い換える。
	Hの2	通常表現に言い換える。

各ルールの一例を【図 2】と【図 3】で紹介すると、カテゴリーA『長文「分かりにくい」を避ける。』はルール A の 1「長文は、連文（複数の短文）にする」を含み、カテゴリーB『不整列「雑然とした並び」を避ける。』はルール B の 1「修飾要素を被

修飾要素の近くに」を含んでいる。

A

**長文「分かりにくい」を避ける。**  
 ～文は短く！～

**ルール 第A条の1**  
 長文の複文を連文（複数の短文）にし、各文の因果関係が明確になるよう言い換える。  
 前半（原因）と後半（結果）をそれぞれ独立した文にし、因果関係を示す接続詞で結ぶ。  
 ＊複文：述語が2つ以上存在する文。述語を中心としたまとまりが2つ。

**改訂例 1**——長文は、連文（複数の短文）に
 

- 「このように、基板の表面処理工程を働いているので、凹凸を取り除くことができる。」
- ⇒「このように、基板の表面処理工程を働いている。そのため、凹凸を取り除くことができる。」

**修正比較**

- ・「（第1）ので、（第2）。」
- ⇒「（文1）。そのため、（文2）。」

【図 2】ルール A の 1

B

**不整列「雑然とした並び」を避ける。**  
 ～適切な順番に！～

**ルール 第B条の1**  
 主題成分を先頭に配置し明示する構文に言い換える。

**改訂例 1**——（主題成分）を先頭へ
 

- 「効果的に残すような構造を支持部は採用していない。」
- ⇒「支持部は、効果的に残すような構造を採用していない。」

**修正比較**

- ・「（目的部）（主題成分）は～していない。」
- ⇒「（主題成分）は、（目的部）～していない。」

【図 3】ルール B の 1

本マニュアルは平成 25 年 6 月から冊子頒布を開始し、平成 26 年 2 月の増刷版と合わせ約 800 部を配付した。また、電子データ提供のニーズが多かったことから、平成 25 年 8 月から PDF 版のダウンロードサービスを開始し、利用者情報を登録いただく仕組みとしたが、これまで約 600 回のダウンロードをいただいている。のべ 1,400 人へ頒布でき、本マニュアルの普及・啓発には、一定の成果が得られたものと認識している。

【特許ライティングマニュアル（初版）】

<http://japio-tjp.org/wmanual.html>

## 5. 特許ライティングマニュアルの改訂について

前節で紹介した『特許ライティングマニュアル（初版）』について、さらなる内容の充実を図るために、内容の再検討を進めて、ルールの組替えや

実例追加など大幅な改訂を予定している。

初版は言い換え規則というルールに記載が中心となっていたが、改訂版は特許文書の作成者がより明晰・明瞭である日本語文を記述できるような内容に充実したいと考えている。

例えば、実際の特許審査や訴訟で用語の明確性が争われた事件などの分析を行い、特許文書によく現れる用語や言い回しであっても、翻訳が難しい事例や多義的に解釈される事例、文法的・技術的・法律的に不明確となり易い事例等を掲載したい。

具体的には、以下のような内容を掲載することを予定している。

#### (1) 特許特有の用語の分析

特許文書では、【表 2】に例示する「二字熟語の動詞形」（「配設する」、「当設する」等）が多く用いることが特徴的である。これら用語は、他分野の文書ではほとんど用いられず、特許分野のみで慣用される用語が少なくない。

【表 2】特許文書で多用される二字熟語

配設	当接	連通	係合	載置	挿通
嵌合	出射	押圧	延在	係止	回動
付勢	立設	挟持	延出	突設	延設
担持	螺合	巻回	軸支	入球	貼着
摺動	介装	並設	連設	画成	捕集
混練	昇温	摺接	外嵌	開弁	拡張

(※2014 年に発行された約 30,000 件の公開特許公報の【要約】に出現する二次熟語の動詞形を抽出し、その出願頻度が高い用語を列挙した。)

これら二字熟語の動詞形は、特許文書で広く用いられるが、各文書における用語の定義が異なる場合もある。例えば「摺動する」は、英訳すると概ね“slide”であり、各文書による定義の違いは殆ど見受けられない。

しかし、例えば「係合する」は、『特許技術用語

集（第 3 版）』（日刊工業新聞社・平成 20 年 8 月刊行）では“engage”を推奨しているが、“fit”、“latch”、“connect”等と翻訳されることも多い。翻訳実務では、翻訳者が前後の文脈に照らして「係合する」の意味を具体的に解釈し、相応しい訳語を翻訳者が選択している。

多様な訳語が与えられる用語は、その意味や定義が翻訳で変わらないように、可能な限り別の表現を用いるなど、日本語文の記述を工夫する必要があると考えられる。実際の特許文書を収集し、そこで用いられる日本語文やその英訳文を分析することで明らかにできるのではないか。この分析結果から、注意すべき用語を特定・整理し、マニュアルの改訂作業に役立てたい。

#### (2) 拒絶理由通知書、裁判例等の分析

先述したように「特許版・産業日本語」活動では、特許審査において通知された拒絶理由通知書の分析を行っている。このように実際に審査や裁判で用語の解釈などで争われた事例を分析し、得られた知見等をマニュアルに盛り込む。例えば以下のような例が想定される。

##### <例 1>

特許審査において、特許明細書に記載された構成のうち、構成要素間の関係（位置、方向等）が明確でないと指摘される事例が少なくない。

日本語文では、一部記載を省略や抽象化しても文意理解に支障ないことがあるが、構成要素間の関係が明確でないと、翻訳した英語文で正確に訳されない原因ともなり得る。

技術常識に照らしたり、前後の文脈を考慮したり、図面を参照したりすると、特許発明の構成要素間の関係を理解できることもあるが、特許文書に用いる日本語文では、一文だけで十分に理解できることが重要である。

一例として、「遊技球毎に発射信号を送信し、そ

の発射信号に基づく始動入賞球が生じた場合には  
抽選結果信号前記発射信号と共に送信する信号送  
信手段」では、「信号送信手段」が何に対して送信  
するのか明確でない。

報の可能性の拡大につながることを期待したい。

#### <例2>

裁判等において、用語の明確性で争われた事例  
に基づいて、拒絶理由になりやすい用語、権利範  
囲に影響のある用語などを掲載し、このような用  
語を用いる際の注意を促していく。

一例として、「平均粒子径」の明確性が争われた  
ものを示す。

知財高裁平成20年（ネ）第10013号：粒子径の  
定義が複数存在するにも関わらず、明細書では測  
定方法等の具体的な記載がなかった。そのため、  
「平均粒子径」の示す意味が不明確であるとされ、  
特許無効の判断がなされた。

このように不明確と指摘されやすい用語を、拒  
絶理由や判例などを分析することにより抽出し、  
注意を要する用語として整理することを予定して  
いる。

## 6. おわりに

本稿では、「産業日本語プロジェクト」、「特許  
版・産業日本語」活動をご紹介します、主要な取組の  
一つである「特許ライティングマニュアル」につ  
いて紹介した。

このマニュアルが、特許文書に関わるより多く  
の関係者に利用いただけることが当面の目標であ  
る。これにより、明晰・明瞭である日本語文を用  
いる特許文書が増加することとなり、その産業・  
技術情報としての価値が一層高まる。そして、誰  
もが容易にアクセスして活用でき、知的生産性の  
向上が期待できる。

また、「特許版・産業日本語」の活動が、他の産  
業・技術分野へ横断的に広がることにより、他分  
野の文書の価値も高まり、我が国の産業・技術情



## 第3回 JTF 関西セミナーに参加して

内田真弓

関西外国語大学

第3回 JTF 関西セミナー(大阪)が2015年1月31日(金)、大阪大学中ノ島センターの「佐治敬三メモリアルホール」で開かれました。あいにくの天候にも関わらず、会場はほぼ満員。昼食には「ビリケン弁当」とお茶、さらにはコーヒーが提供されるなど、終始、和やかな雰囲気でおこなわれました。

今回は、「機械翻訳と向き合うときが来たーMTをもっと身近に、現実的に考えるー」をテーマに、企業、翻訳会社、翻訳者が一堂に会し、各立場からの現状や悩み、課題が報告され、フロアからの質疑応答を交えての議論が行われました。午前中は、3名のパネラーにモデレーターを交えてのパネルディスカッション、午後はさまざまな立場から機械翻訳に携わる5名による講演、そして講演終了後は懇親会と、盛りだくさんのセミナーでした。

報告者は、現在、どっぷりと翻訳業に使っているわけではなく、日常的に機械翻訳を利用することはありません。しかし、日々の翻訳業と翻訳教育を行う中、目下の関心は、「機械翻訳はどこでどのように使われているのか」、「大学の授業で取り入れられる何らかの情報はなにか」という点にあります。

セミナーの構成は次の通りでした。午前中は「医療分野で機械翻訳は本当に使えるのか?」というテーマでパネルディスカッションが行われました。パネラーは、武田薬品工業株式会社で翻訳やコーディネー

ター業務を担当されている生田明子氏、アジア太平洋機械翻訳協会会長の中岩浩巳氏、そして、株式会社みらい翻訳の栄藤稔氏、モデレーターは株式会社川村インターナショナルの森口功造氏でした。

生田氏からは、「機械翻訳に期待すること」というタイトルで、ここ最近の翻訳業務の事例を挙げながら、クライアントの立場から現状報告がありました。また、翻訳業務の発注時に留意している点についても指摘がありました。具体的には、「スピード」「ボリューム」「品質」「コスト」の4つの観点が翻訳業者選定時の重要な観点となるそうです。これらの観点に加えて、製薬会社という機密事項が多い業務上、「機密管理」が何よりも危惧されるとの指摘がありました。情報の機密性を担保できれば、翻訳業者を利用する機会がますます増加することが期待されとのことでした。そして、このように今後も翻訳業務が増加するなか、業務の効率化の有効な手段として機械翻訳システムの導入が果たす役割は大きいとも指摘がありました。

中岩氏からは、アジア太平洋機械翻訳協会がどのような活動をおこなってきたのか、また、機械翻訳の基礎と動向についての報告がありました。機械翻訳の代表的な手法に、「ルールベース機械翻訳」と「統計的機械翻訳」があります。前者は人間が作成した大量のルールを用いて翻訳する方法で、後者は既存の対訳データを統計的に処理し、翻訳ルールを自動的に構築して翻訳

を実現させるという手法です。中岩氏によれば、「統計的機械翻訳」の基盤として大量の対訳データの流通を確保できれば、こちらのほうが有望な手段と位置づけられるとのこと。というのも、「統計的機械翻訳」の場合、対訳データを確保できれば、機械翻訳システムを自動的に構築できるからです。それも、翻訳システムの多言語化も容易となるという利点もあるといえます。「ルールベース機械翻訳」と比較すると、ルール作成が不要となるため、ルール作成に必要な時間的、金銭的成本を削減できるという利点もあります。

栄藤氏からは、自身が代表取締役を務める「みらい翻訳」のビジョンと展望についてのご紹介がありました。「みらい翻訳」は、機械翻訳技術によって、世界中の誰もが国境を越えて交流ができる社会を実現することを目指しており、NTT DoCoMoをはじめとするさまざまな企業にサービスや技術の提供をされています。今後は、話し言葉の機械翻訳や Web サイトの翻訳ツール、クラウドソリューションなどの提供を予定しているとのこと。機械翻訳は、ともすればコスト削減を目的として導入される傾向がありますが、栄藤氏は、これまでに機械翻訳が使われていなかったところへ参入していく可能性についても考える必要がある、とおっしゃっていました。「機械翻訳の新しい常識をつくりたい」—そう話す「みらい翻訳」のこれからの動向から目が離せません。

午後の部は、翻訳会社や機械翻訳を開発する側の講演が主となりました。機械翻訳を用いる翻訳は、最終的に翻訳者が手直しをして完成させるのが一般的です。原文をより正確に翻訳できるよう、機械翻訳にかける前に TM (Translation memory)を用い

て事前翻訳をおこなったり、あるいは機械翻訳でよく見られる間違っただけの言い回しを一括で置換できるようなシステムを開発するなど、さまざまな取り組みがなされていることがわかりました。

翻訳は完成版を顧客に提供すると思い込んでいた私にとって意外だったのが、株式会社十印の翻訳の提供の仕方でした。十印では、顧客が求める品質と用途によって、機械翻訳を使う程度を変えるというのです。たとえば、なんとなく意味がわかればいいのであれば、翻訳者の作業は、機械翻訳の結果を確認する程度に留めます。そうすることで、コストが大幅に削減され、納期もぐんと短くなります。一方で、自然な文章レベルの翻訳が求められる場合、まずは機械翻訳で翻訳、その後、翻訳者により大幅な修正を加えていくことになります。こうすることで、高品質でありながら、すべてを人間が訳すよりも短時間で納品が可能となるわけです。

個人的に興味深かったのが、情報通信研究機構 (NICT) の内山将夫氏による「統計的自動翻訳 (SMT) の実用化」という発表でした。SMT は基本的に、集めた対訳データから正しいアルゴリズムを導き出し、翻訳を行うというもので、対訳データが多ければ多いほど、翻訳の性能はアップします。ちなみに、対訳コーパスの身近な例に、アルクの「英辞郎」があります。私も日常的に使用していますが、これは辞書には載っていない訳語を探すのに非常に便利です。

このような大規模な対訳データは 2000 年あたりから出現してきました。翻訳をおこなうのに必要なデータ量は、マニュアル翻訳であっても 10~30 万文程度、高性能な特許翻訳であれば 3000 万文程度は必要になるとのことで、これは個人経営の事務所

や会社が扱える分量ではありません。国家規模の取り組みにして、日本の翻訳技術を飛躍的に向上させるべきだ、というのが内山氏の見解でした。

データはそこら中に散在していますが、そのままではデータとして使用することはできません。そこで NICT が編み出したのが「みんなの自動翻訳@TexTra」です。「みんなの自動翻訳@TexTra」は誰でも無料で使用することが可能で、散在する対訳データをみんなで効率的に集めることを目的としています。対訳データの量が増えれば増えるほど自動翻訳の精度は上がっていきますから、参加が増えるほど、両者にとってプラスになるというわけです。これはありそうでなかった取り組みで、手軽に利用できますので、授業にも何らかの形で取り入れることができるのではないかと考えています。まずは自分で使ってみて、その使い心地を実感してみたいです。

今回の第3回 JTF 関西セミナーで、機械翻訳に関するイベントに参加するのは2回目となります。今回の講演を拝聴し、前回以上に、機械翻訳と翻訳者との棲み分けが明確であることがわかりました。機械によるデータの処理スピードは、とても翻訳者には真似できません。しかし、機械翻訳の処理において翻訳者の知識や経験を活用することで、機械翻訳であっても翻訳者と同等の翻訳を目指せそうです。また、機械翻訳の精度を高めるさまざまな工夫や、顧客の要望によって機械翻訳の使い方を変えるといった実務上の機械翻訳の活用法を伺い、今回のセミナーに参加したことで、機械翻訳を取り巻く環境がより明らかになりました。

## みらい翻訳のご紹介

栄藤(えとう) 稔

株式会社みらい翻訳社長 兼 NTT ドコモ 執行役員 イノベーション統括部長

### 1. みらい翻訳の生い立ち

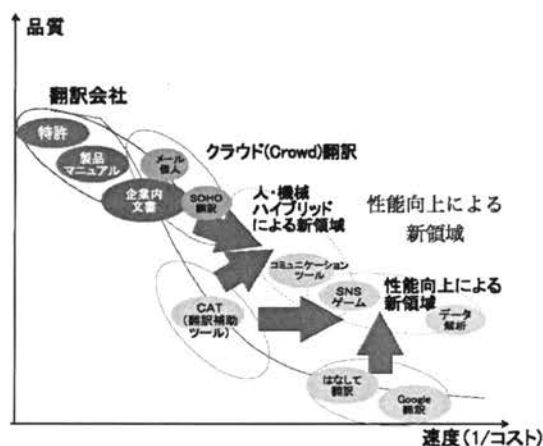
当社は NTT ドコモ、シストラン社、フュートレック社の合併会社として2014年10月に設立された。



フュートレック社は音声認識技術と情報通信研究機構(NICT)の翻訳技術を改良した実用化技術を持つ。NTTドコモはモバイル通信技術としゃべってコンシェルに代表される超大規模な自然言語サービス基盤構築技術を持つ。シストラン社は主に米国、フランス、韓国に開発拠点があり欧州言語を中核とする多言語翻訳技術、多種大量のコーパスとその管理学習システムを持っている。これらに併せて日本電信電話(NTT)とNICTの持つ自然言語処理、音声入力、多言語翻訳、コーパスに関する研究成果の提供を受けて、話し言葉から書き言葉に至る包括的な機械翻訳技術の産業応用(B2B事業)を目指すことを目的として同社を設立した。

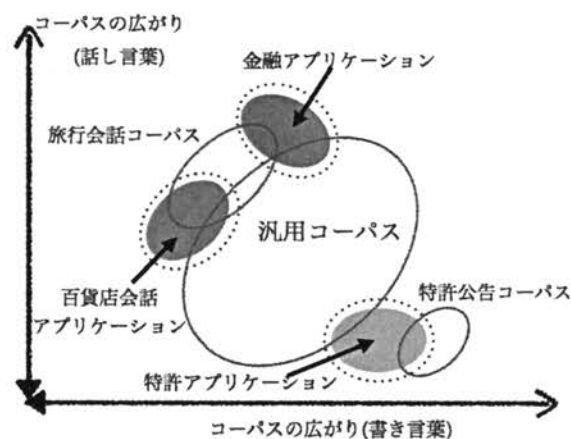
### 2. 目指すゴールとアプローチ

機械翻訳の応用分野として多言語コミュニケーション、データ解析(マーケティング、訴訟対応等)、産業文書翻訳(ローカライゼーション)を想定している。現在、文法が近い同じ語族間の翻訳、例えば英仏、英独、日韓の機械翻訳の品質は分野を限定するものの実用レベルに達している。日本語を中心としつつ、語族を超えた翻訳の品質をそのレベルまで高めたい。機械翻訳の品質がある閾値を超えることにより新たな市場(下図の矢印部分)が拓けるものと確信している。



当社の提供する基幹システムは以下である

1. 高度にカスタマイズ可能な産業文書・会話翻訳。  
SYSTRAN Enterprise Server を改良して提供。
2. 話し言葉実時間翻訳



カスタマイズは上図にあるようにドメイン単位で最適なコーパス(点線部)を切り出すGUI付きシステムにより実現している。また大規模クラウド利用技術<sup>1)</sup>を用いて高スループットソリューションも提供する。いずれも多言語対応で、これらが当社の強みとなっている。

<sup>1)</sup> (<https://www.39works.net/assets/dcm-cloudconsulting/>)

Transit<sup>NXT</sup> 最新サービスパック 8 について

目次 由美子

翻訳現場のニーズに則した実践的なツールとして定評を博している翻訳支援ツール Transit<sup>NXT</sup> のサービスパック (SP) 8 が 2015 年 2 月にリリースされた。この度のアップグレードではエディター操作における基本的な既存機能の強化に加え、改訂履歴を追跡するための新たな機能が実装されたばかりでなく、対応ファイル形式も追加され、機械翻訳機能においてはさらなる進化を遂げている。代表的な機能拡張をここに紹介したい。

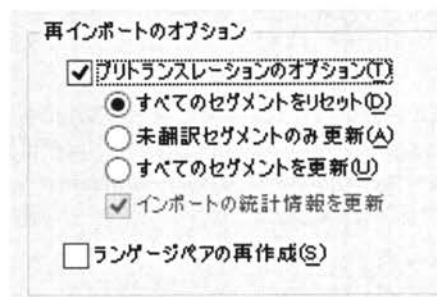
まず、作業用の言語が追加されたことを挙げたい。ビルマ語やタガログ語などのほか「ドイツ語 (標準)」も加えられ、対応言語はバリエーションを含め 200 を超えた。

対応ファイル形式も増えた。Microsoft Word、Excel、PowerPoint はバージョン 2013 に対応し、InDesign CC2014 と Adobe FrameMaker 12 に対応した。

Transit はグローバル企業のコーポレートビデオや研修ビデオなどを対象とした字幕を翻訳するための機能も装備されている。従来の SubRip、Text という形式に加えて、WebVTT にも対応した。

そして SDL 社の Trados で作成される翻訳パッケージ SDLPPX にも対応。Transit での翻訳後には返却パッケージ SDLRPX を作成できる。

Transit 搭載の再インポート機能を利用すれば、Trados で生成された翻訳プロジェクトであっても、ユーザー自身が蓄積してきた Transit 形式の翻訳メモリを効率的に併用することも可能。

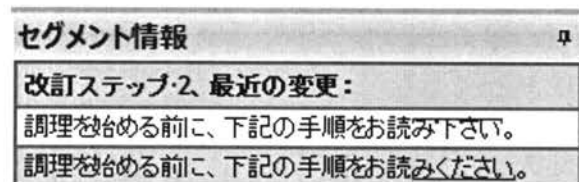


検索機能も拡張された。単純に用語を検索するだけでなく、カウントやマーキングが可能となった。さらに、セグメントフィルタ機能 (検索用語が含まれるセグメントのみを抽出する) との連携も容易になった。

Transit には「フォーマットチェック」と呼ばれる品質保証機能も装備されている。これによりソースファイル (翻訳元) とターゲットファイル (翻訳先) の内容が比較され、マークアップや数字が適切に保持されているか、句読点が置かれているかといったチェックを実行できる。このチェックオプションをセットとして保存できるようになったため、プロジェクトマネージャと翻訳者間でチェック項目を共有できるのみならず、特定の分野や顧客向けのチェックルーチンを確立できるようになった。



そして、改訂履歴の追跡も可能となった。セグメントごとに改訂内容が記録されるのである。直近の改訂は「セグメント情報」ウィンドウ内の上部に表示され、改訂箇所はマーキングにより明示される。



さらに、[改訂ステップを比較]ウィンドウに記録された改訂内容を一覧表示させて比較し、元の翻訳に戻すことも可能。



また、セグメントフィルタ機能を利用して改訂が記録されたセグメントのみを抽出表示させることも、有償オプションの品質レポート機能を利用して改訂が記録されたセグメントのみを Excel や HTML などのデータとして出力することも可能。

SP8 では対応 MT システムも追加された。

Transit には 2 種類の機械翻訳機能が装備されている。Transit での翻訳作業時にオンライン翻訳サービスを利用する「エディターMT」。対応済みの Google Translate と iTranslata4.eu に加え、My Memory と Microsoft Translator に新たに対応した。

もう 1 種の MT 機能は「インポート MT」と呼ばれる。翻訳対象の文書を Transit 形式に変換する際、翻訳メモリと新規文書の原文テキストをセグメント単位で比較し、100%一致する場合は翻訳メモリから訳文を自動的に挿入する。このプロセスにおいて機械翻訳システムに訳文を生成させて挿入させることが可能。旧来対応していた SYSTRAN と SmartMATE に加え、AsiaOnline

にも、さらには STAR グループで独自に開発した SMT エンジンである「STAR Moses」にも対応した。

インポート MT ではペナルティを与えるオプションの利用が可能であり、100%マッチとして機械翻訳により生成された訳文をターゲットファイルのすべてのセグメントには挿入しないようにできる。たとえばペナルティに 20%が付与されて生成された「MT 訳例」は 80%の類似率をもつ訳文としてファジーマッチウィンドウに表示されることにより、翻訳者・ポストエディターの手により適切な訳として完成させるプロセスが実現可能である。



Transit と連動する多言語多方向の用語管理ツールである TermStar においても SP8 での機能拡張が見受けられる。特に重複を制御するための機能が強化されており、データレコードフィルタのデフォルトに「重複」が追加されている。これにより、既存データレコード内の重複を抽出し、不要なレコードは削除するための操作を容易に実行できる。また、新規のみならず既存のデータレコードに対しても GUID (Globally Unique Identifier) を付与することが可能となった。さらに、新規用語を登録する際には既存のエントリおよびサブエントリを対象に重複が自動的にチェックされる機能も搭載された。

Transit は設立から 30 周年を迎えたシュタールグループの翻訳事業のスキルとノウハウが凝縮されている完成度の高い翻訳支援ツールであるが、今後もさらなる発展に期待したい。

## これまでの AAMT Forum メールマガジン Vol.5

機械翻訳課題調査委員会 WG1、WG2

### はじめに

AAMT Forum を通じて配信したメールマガジンのバックナンバーをお届けします。今回は第 38 号から第 46 号です。

今回は、2020 年夏季の東京オリンピック・パラリンピックに向けての様々なニュースがありました。翻訳を通じてより一層の「おもてなし」への貢献が期待されます。今後もメールマガジンを通じて様々な情報を皆様に配信させていただきます。

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

では、メールマガジン第 38 号から第 46 号を再びお届けします。

### AAMT Forum メールマガジン 38 号 (2015/01/23 配信)

こんにちは。AAMT Forum メールマガ担当です。大寒の厳しい寒さが身にしみるこの頃、いかがお過ごしでしょうか。ストーブやコタツなどの暖房器具にお世話になっているかと思いますが、火事にはならないように気をつけましょう。では、メールマガジン第 38 号をお届けします。

■グーグル翻訳アプリに「同時通訳」機能、カメラ翻訳も

米グーグルの翻訳アプリ「グーグル・トランスレート」がアップデートされます。会話にほぼ同時通訳のように対応できる「会話モード」機能と、画像で取り込んだ文字を翻訳する機能「カメラ翻訳」が加わります。

<http://jp.wsj.com/articles/SB11258286719794574597104580400433855858528>

■外国語の音声翻訳に 14 億円 15 年度予算案に計上

2015 年度予算案をめぐる 12 日の麻生太郎財務相と高市早苗総務相の折衝で、訪日外国人を支援するため、外国語の音声の高い精度で自動翻訳するシステム開発に 14 億円を計上することが決まりました。

<http://www.tokyo-np.co.jp/s/article/201501201001494.html>

■NTT ドコモ、日本語と外国語での会話を可能とするアプリ「はなして翻訳・JSpeak」の iPhone 向けを提供開始！

NTT ドコモは 19 日、スマートフォン（スマホ）やタブレットを利用して、日本語と外国語の間での会話を可能とする海外向け翻訳サービス「はなして翻訳・JSpeak」（以下、はなして翻訳）の iPhone 対応アプリを 2015 年 1 月 19 日（月）から提供するとお知らせしています。

<http://news.livedoor.com/article/detail/9690515/>

---

### AAMT Forum メールマガジン 39号 (2015/02/05 配信)

こんにちは。AAMT Forum メルマガ担当です。立春を過ぎ、これから徐々に暖くなるこの頃、いかがお過ごしでしょうか。今年もインフルエンザが猛威を振るってあります。外出したらうがい、手洗いを忘れずに。では、メールマガジン第39号をお届けします。

---

#### ■Twitter、外国語ツイートの翻訳機能

Twitterは、外国語のツイートの翻訳を表示できるようになったと発表しました。Microsoftの「Bing 翻訳」と提携して機械翻訳を表示するもので、40超の言語に対応しています。

[http://internet.watch.impress.co.jp/docs/news/20150123\\_685088.html](http://internet.watch.impress.co.jp/docs/news/20150123_685088.html)

---

#### ■NTT データ、技術文書を英文機械翻訳する法人向けクラウドサービスを発表

NTT データは2015年1月27日、製品マニュアルなどの技術文書を和英/英和で機械翻訳するクラウドサービスを4月から提供すると発表しました。

<http://itpro.nikkeibp.co.jp/atcl/news/15/012700313/>

### AAMT Forum メールマガジン 40号 (2015/02/20 配信)

こんにちは。AAMT Forum メルマガ担当です。春寒の候、いかがお過ごしでしょうか。まだ所によっては雪が降ることもありますが、それでも春の気配を少しずつ感じられる頃となりました。

では、メールマガジン第40号をお届けします。

---

#### ■パナソニックが東京五輪向け展示会 自動翻訳機など公開

パナソニックは10日、訪日外国人客に対応するための自動翻訳機や、決済機能を搭載したウェアラブル端末などの2020年の東京五輪向けに開発している製品や技術を公開しました。

[http://www.nikkei.com/article/DGXLASDZ10I2N\\_Q5A210C1TJ2000/](http://www.nikkei.com/article/DGXLASDZ10I2N_Q5A210C1TJ2000/)

---

#### ■フュートレックの子会社であるATR-Trekが、住友ゴム工業と

タイヤ製造業務向けに「日中音声翻訳システム」を開発

ATR-Trekでは、住友ゴム工業と共同で新規の利用分野に向けて効率的にカスタマイズを行い、タイヤ製造業務コミュニケーションに利用可能な「日中音声翻訳システム」を実用化し、同社中国工場（常熟、長沙）において本格的に利用が始まりました。

<http://japan.cnet.com/release/30093388/>

**AAMT Forum メールマガジン**  
**41号 (2015/03/04 配信)**

こんにちは。AAMT Forum メルマガ担当です。春眠あかつきを覚えずの快い季節になってまいりましたが、寝坊することのないように気をつけましょう。では、メールマガジン第41号をお届けします。

---

■東京マラソン:頑張れ和製翻訳アプリ 音声で14言語対応

2020年東京五輪・パラリンピックに向け、外国人との言葉の壁をなくするという取り組みの一環で、最先端の自動翻訳技術が大規模なスポーツ大会で初めて導入されました。

<http://mainichi.jp/select/news/20150222k000m040012000c.html>

---

■翻訳サイトの情報漏えい、悪いのは事業者? 利用者?

情報処理推進機構(IPA)は2月20日、ネット上の翻訳サービスで入力した文章が他ユーザーも閲覧できる状態になっていたことを受けて注意喚起を行いました。

<http://news.mynavi.jp/articles/2015/02/20/translation/>

---

■これさえあれば、外国人と自由に話せる!

パナソニックと九州旅客鉄道(JR 九州)は、3月3日~4日の2日間、JR 博多駅 3F の博多シティ改札口内スペースにおいて、

多言語翻訳ソリューションを通して、九州の魅力を発信するイベント「Wonder Japan Trip -2020年の観光を考える-」を開催しました。

[http://pc.watch.impress.co.jp/docs/news/20150303\\_690879.html](http://pc.watch.impress.co.jp/docs/news/20150303_690879.html)

---

**AAMT Forum メールマガジン**  
**42号 (2015/03/20 配信)**

こんにちは。AAMT Forum メルマガ担当です。春暖の候、いかがお過ごしでしょうか。近頃は気温も暖かく、過ごしやすくなってきました。一方で花粉症の季節でもありますので、対策はしっかりとしておきましょう。では、メールマガジン第42号をお届けします。

---

■NTT データ、技術文書を対象にした機械翻訳サービスを提供

株式会社 NTT データは、日本電信電話株式会社 (NTT) が開発した機械翻訳エンジン「多言語統計翻訳プラットフォーム」を採用した技術文書の翻訳サービスの提供を4月に開始します。

<http://www.keyman.or.jp/nw/20060736/>

---

**AAMT Forum メールマガジン**  
**43号 (2015/04/01 配信)**

こんにちは。AAMT Forum メルマガ担当です。桜の花も咲きはじめ、暖かな過ごし

やすい日が続くこの頃、いかがお過ごしでしょうか。新年度も始まり、新しい出会いが楽しみでもありますね。では、メールマガジン第 43 号をお届けします。

---

#### ■科学技術文献データベースの作成に「高精度自動翻訳システム」を導入

科学技術振興機構（JST）は、NICT と共同で開発した英日自動翻訳システムを科学技術文献データベースの作成・提供を行う文献情報事業に導入し、中国文献データベース（JSTChina）内の英語の標題、抄録がある文献記事の和訳に適用し、平成 27 年 3 月 30 日から公開しました。

<http://www.nict.go.jp/press/2015/03/30-1.html>

---

#### ■Apple Watch™向け翻訳アプリ「IHG™Translator」を提供

IHG はこのたび、無料翻訳アプリ「IHGTranslator」の Apple Watch™対応版をリリースすると発表しました。

<http://prtimes.jp/main/html/rd/p/000000283.000002364.html>

---

#### AAMT Forum メールマガジン 44 号（2015/04/16 配信）

こんにちは。AAMT Forum メルマガ担当です。若葉の緑が目にも鮮やかなこの頃、いかがお過ごしでしょうか。一時は冬戻ったかのように寒くなる時もあり、寒暖の差

が激しい時期でもあります。体調を崩さないように気をつけましょう。では、メールマガジン第 44 号をお届けします。

---

#### ■Skype 翻訳、中国語とイタリア語をサポート

Skype は 4 月 8 日（現地時間）、音声同時通訳ツール「Skype Translator（日本では「Skype 翻訳）」」（プレビュー）で中国語（北京語）とイタリア語をサポートしたと発表しました。

<http://www.itmedia.co.jp/enterprise/articles/1504/09/news101.html>

---

#### ■外国語の看板をその場で翻訳、東芝が基礎技術を開発

東芝は 2015 年 4 月 2 日、カメラで撮影した看板や標識、レストランのメニューなどの画像から文字列を検出する技術を開発したと発表しました。

<http://www.nikkei.com/article/DGXMZO85381930X00C15A4000000/>

---

#### AAMT Forum メールマガジン 45 号（2015/05/01 配信）

こんにちは。AAMT Forum メルマガ担当です。薫風の候、暖かな過ごしやすい日々が続いておりますね。既に大型連休を満喫されている方もいるのでしょうか。休んだ後は五月病などに罹らないように気をつけましょう。では、メールマガジン第 45 号をお



届けします。

---

■エキサイト翻訳 140 万語以上の専門用語を追加

エキサイト株式会社は、「エキサイト翻訳」に、2015 年 4 月 22 日（水）から 140 万語以上の専門用語の辞書を追加し、翻訳時に選択できる 8 系統 106 分野を設置しました。

<http://prtimes.jp/main/html/rd/p/000000450.000001346.html>

---

■「コリャ英和！一発翻訳 2016 for Win」シリーズ（CD-ROM 版）を新発売

ロゴヴィスタ株式会社は、Windows 用英日・日英翻訳ソフトの最新版「コリャ英和！一発翻訳 2016 for Win」シリーズを、2015 年 5 月 29 日（金）より発売します。

<http://www.logovista.co.jp/LVERP/information/news/2015-0424-kry2016win.html>

---

AAMT Forum メールマガジン  
46 号（2015/05/15 配信）

こんにちは。AAMT Forum メルマガ担当です。緑樹の候、いかがお過ごしでしょうか。先日は台風 6 号が来ましたが、大事なかったのでしょうか。今年の台風発生数は既に平年の 3 倍にも達しており、台風の当たり年になるようです。台風への備えは怠ら

ないようにしましょう。では、メールマガジン第 46 号をお届けします。

---

■車内案内を"自動翻訳"

ヤマハ「おもてなしガイド」アプリ使い  
東急バスで実証実験

ヤマハと東急バスは、代官山循環バスの車内アナウンスにヤマハが開発した「おもてなしガイド」を活用した実証実験「東急バス×ヤマハ Sound UD 化プロジェクト」を、2015 年 5 月 16 日から 9 月 30 日までの約 4 か月間実施します。

<http://www.j-cast.com/trend/2015/05/06234349.html>

## UTX 変換ツール正式版の公開

AAMT 機械翻訳課題調査委員会 WG3 (標準化・共有化)

### 1. UTX 変換ツールの公開

かねてよりオープンソースとして開発を行っていた UTX 変換ツールの正式版 Ver.1.11.0.1 を公開した。

<http://utxconv.sourceforge.net/ja/>

UTX 変換ツールは、用語集形式の標準規格である UTX (Universal Terminology eXchange) 形式<sup>1</sup>のフォーマットチェックや、各社翻訳ソフトウェアのユーザー辞書との変換を行うことができるソフトウェアで、誰でも無償で利用することができる。

### 2. UTX 変換ツールの画面

UTX 変換ツールは以下のような画面をもつソフトウェアである。



<sup>1</sup> UTX 仕様については以下を参照。  
<http://www.aamt.info/japanese/utx/>

今回、日本語 OS ではユーザーインターフェイスが日本語になるようにし、それ以外の言語の OS では英語のインターフェイスが表示されるようにした。また、日本語版と英語版のマニュアルを作成した。



### 3. UTX 変換ツールの機能

UTX 変換ツールは現在、以下の機能を持っている。

- ・ UTX に関する機能
  - ・ UTX ファイルの形式チェック (UTX 1.11)
  - ・ 禁止語の抽出
  - ・ 禁止語と承認語のペアの抽出
  - ・ 非標準語と承認語のペアの抽出
- ・ 変換機能
  - ・ UTX と ATLAS (富士通) のユーザー辞書 (\*.txt ファイル) との相互変換
  - ・ UTX と The 翻訳 (東芝) のユーザー辞書 (\*.txt

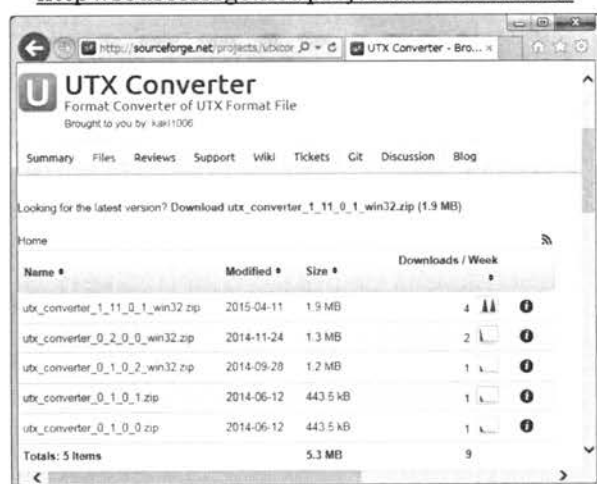
ファイル) との相互変換

- UTX と PC/MED/PAT/Legal Transer (クロスランゲージ) のユーザー辞書 (EJ は \*.opt ファイル、JE は \*.dic ファイル) との相互変換
- UTX から MultiTerm 用のインポート用テキストへの変換

#### 4. UTX 変換ツールの利用方法

UTX 変換ツールは以下の URL からダウンロードすることができる。

<http://sourceforge.net/projects/utxconv/files/>



本校執筆時点での最新版である Ver.1.11.0.1 をダウンロードする場合は、上記サイトから以下のファイルをダウンロードする。

utx\_converter\_1\_11\_0\_1\_win32.zip

UTX 変換ツールのインストール方法は、ダウンロードした zip ファイルを展開し、任意のフォルダに置くだけである。

UTX 変換ツールの利用方法は、UtxConverter.exe をダブルクリックで実行し、形式を変換したいファイルを画面にドラッグ&ドロップし、変換元と変換先のファイル形式を指定し、「実行」ボタンを押すだけというシンプルなものとなっている。

ダウンロード方法、インストール方法、利用方法に

ついては、オンラインで公開しているマニュアルにも記載しているので、詳細についてはそちらをご覧ください。

<http://utxconv.sourceforge.net/ja/>

#### 5. 今後の予定

UTX 変換ツールでは今後、以下のような機能の追加を予定している。

- UTX と J 北京 (高電社) のユーザー辞書との相互変換
- UTX と J ソウル (高電社) のユーザー辞書との相互変換
- UTX と Excel ファイルの相互変換
- UTX と TBX の相互変換
- UTX 1.20 への対応
- その他、ユーザーインターフェイスの改良

#### 6. 謝辞

UTX 変換ツールにおいて各社ユーザー辞書との変換機能を作成するにあたり、以下の各社から製品を AAMT に無償提供いただいた。ありがとうございました。

株式会社 クロスランゲージ

PC-Transer V21 プロフェッショナル

株式会社 高電社

J 北京 7 スタンダード 2

J ソウル 9 Expert

# AAMT会員のひろば

AAMT 会員の新たな交流の場を AAMT Journal 誌面上で提供するべくスタートいたしました「AAMT 会員のひろば」、会員の皆さまのご助力をいただきまして、第一回の No.41 のスタートから第七回を迎えることができました。今号では、法人会員二社、個人会員一名の皆さまからのご寄稿をいただいております。

独自のお取り組みのご紹介、機械翻訳研究への提言、AAMT の活動へのご要望など、今回も貴重なご意見をお寄せいただきました。

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

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

---

## 法人会員（敬称略・50 音順）

---

### 会員名

GMO スピード翻訳株式会社

<https://quicktranslate.com/>

〒150-8512 東京都渋谷区桜丘町 26-1 セルリアンタワー10F

---

### 自己紹介

GMO スピード翻訳株式会社が運営する「スピード翻訳 by GMO」は、24 時間 365 日いつでも発注できる翻訳サービスです。「スピード翻訳 by GMO」では、翻訳の「申込から納品まで」をワンストップで 24 時間 365 日いつでもご利用いただけます。通常の翻訳会社では味わえないスピード感のあるサービスをご提供いたします。

日本国内外を問わず、厳選された各分野の専門知識を持ったプロの翻訳者が直接翻訳をおこない、高品質な翻訳サービスを提供している LSP（Language Service Provider = 言語サービスプロバイダー）です。

翻訳したい原稿を自動見積りフォームに入力（コピー＆ペースト）するか、ファイルをそのままアップロードするだけで、お見積りと納期をその場で算出いたします。ご満足いただけるようであれば、すぐにご発注いただけます。

翻訳料はリーズナブルな従量制（最低価格：100 円・税別）からですので、お気軽にご利用いただけます。

また、お気に入りの翻訳者をブックマークしておけば、その翻訳者をダイレクトに指名して、翻訳が依頼できます（指名翻訳オプション）。企業、大学、研究機関、官公庁など幅広いお客さまからご発注いただいております。

---

## MT/翻訳とのかかわり

### MTおよび翻訳業界に期待すること

自己紹介でもご説明させていただきました通り、GMO スピード翻訳株式会社では翻訳コーディネーターが介在し役務を提供する旧来の翻訳会社のビジネスモデルではなく、翻訳に特化したクラウドソーシングをビジネスモデルとしているため、納品速度、翻訳料、ユーザビリティにおいて一般的な翻訳会社と比較すると優位性の高いサービスとして法人企業を中心に多くの方にご利用いただいております。しかしながら、もっとも顧客満足度に影響を及ぼす翻訳品質においては、従来の翻訳会社と比較すると不安定な要素が課題として残っており、翻訳品質を改善していくためのプロセスにおいてCATツールの導入やTM・MTの活用を重要視するようになりました。

弊社のサービスを実際に利用する顧客数は年間で2万件を超えており、その数は年々増加していますが、品質管理への取組みの過程においてCATツールの利用は大前提にあり、更に今後ポストエディティングの領域に進むことを前提とするのであれば、機械翻訳の導入は無視できないと感じたところから、AAMTへの加入を決意しました。インターネット（ウェブサイトや人力翻訳API「現在開発中」）を介して大量のコンテンツを翻訳できる当社のプラットフォームでは翻訳データの入稿から厳選された翻訳者による翻訳作業、納品までのプロセスを全て自動化していることから、従来の翻訳会社と比較して翻訳データが集まりやすい環境にあります。個別の企業から提供された翻訳資源を共有データとして運用し、再利用することは現在のところ法的な見解から困難ではありますが、セキュアな情報資産の管理を前提として固有の企業からお預かりする当該情報資産の効率化を行うことは出来ますので、対訳コーパスデータの収集に適した環境であることは言うまでもありません。

弊社には実務経験3年以上のトライアルに合格した厳選された翻訳者が在籍しています。専門分野を特定することや原文データの性質にもよりますが、当該対訳コーパスデータを教師あり学習をベースに磨きをかけていくことで、機械翻訳の精度向上に貢献できるのではないかと感じており、このような専門知識を持っているAAMTの加盟企業様や大学の研究室と積極的にコミュニケーションを行っていければと考えております。

法人名：GMO スピード翻訳株式会社

役職：代表取締役社長

氏名：古谷 祐一

### AAMTへの要望

GMO スピード翻訳株式会社は日本翻訳連盟の理事企業として、翻訳産業の発展のため活動を行っております。翻訳産業の発展及び活性化のためにAAMTと日本翻訳連盟の距離を更に縮め、日本国内のコンテンツの多言語化を急速に進め、来る2020年の東京オリンピックで訪日される外国人や年間1400万人を越える訪日外国人の方に良い体験をしていただく機会を創出していければと思っております。どうぞ、よろしくお願いいたします。



# AAMT会員のひろば

AAMT 会員の新たな交流の場を AAMT Journal 誌面上で提供するべくスタートいたしました「AAMT 会員のひろば」、会員の皆さまのご助力をいただきまして、第一回の No.41 のスタートから第第十六回を迎えることができました。今号では、法人会員二社、個人会員一名の皆さまからのご寄稿をいただいております。

独自のお取り組みのご紹介、機械翻訳研究への提言、AAMT の活動へのご要望など、今回も貴重なご意見をお寄せいただきました。

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

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

---

## 法人会員（敬称略・50 音順）

---

### 会員名

株式会社サン・フレア

<http://www.sunflare.com>

【本社】〒東京都新宿区四谷 4-7 新宿ヒロセビル 2F-10F

【西日本支店】〒541-0042 大阪府大阪市中央区今橋 4-3-22 淀屋橋山本ビル 10F

【北日本支店】〒039-2245 青森県八戸市北インター工業団地 1-4-43 八戸インテリジェントプラザ内

【海外】パリ駐在、大連事務所

---

### 自己紹介

当社は、「提案力」「言語力」「IT力」「デザイン力」「教育力」の5つの卓越した「人財力」を織りなしてオンリーワンソリューションを提供しています。これらの「人財力」で特許明細書、契約書、金融・マーケティング資料、医薬、医療機器など幅広い分野の翻訳を提供しています。また、CSR、CTD、SOP、論文などのメディカルライティングを提供しています。さらに、中国の治験実施支援、薬事申請文書の作成など中国進出のための様々なコンサルティングを提供しています。まさにサン・フレアは、ことばの壁・情報の壁・コミュニケーションの壁を乗り越え、世界に飛翔する企業をトータルランゲージサポートし、『ドキュメント総研®』へと進化しています。

---

## MT/翻訳とのかかわり

### MT および翻訳業界に期待すること

現在の機械翻訳（MT）が検索機の領域から抜け出せない限り、自動翻訳にあっては品質に期待するのは難しいと思われます。しかし、ルールベースの MT を翻訳支援ツールとして活用することで、かなりの効果のあることが実証されています。英文和訳を 8 時間で商品としての翻訳を 1000W しか翻訳できなかった翻訳者が、MT を使うことで 2500W 程度のスピードにまで達しました。一文一文を完成訳に仕上げ、ポストエディットはしていません。ただし、リンギストが未知語（新語）に訳語を付けて既知語としてユーザー辞書を作成する必要があります。また、MT は表現（文体）や用語を統一出来るので、分散翻訳に有効です。

問題は、文学者が執筆・編纂した辞書ではなく MT 専用の産業翻訳専用のコンパクトな辞書を作成することが必要なことです。MT は用語の選択肢の幅を広げるのではなく、狭めることで、高品質の翻訳を担保出来るからです。

株式会社サン・フレア ドキュメンテーション部門統括 齋藤 誠

### AAMT への要望

MT の研究者とソフトウェアベンダーだけでなく、翻訳会社、翻訳者が自由に意見を交換し、交流できる場を提供できるのが、AAMT の強みだと思います。これまで日本語と英語、日本語と中国語の言語対で評価システムを構築してきた成果もとても素晴らしいと感じています。これからは、MT 技術を組み込んだ翻訳ワークフローの標準化やテンプレート化について啓蒙的役割を果たしてリーダーシップを発揮していただくことを期待しています。また日本国内にとどまらず、MT 技術の先進国の成功事例の紹介と、MT 応用技術の啓蒙にも尽力していただきたいと思います。

# AAMT会員のひろば

AAMT 会員の新たな交流の場を AAMT Journal 誌面上で提供するべくスタートいたしました「AAMT 会員のひろば」、会員の皆さまのご助力をいただきまして、第一回の No.41 のスタートから第十六回を迎えることができました。今号では、法人会員一社、個人会員二名の皆さまからのご寄稿をいただいております。

独自のお取り組みのご紹介、機械翻訳研究への提言、AAMT の活動へのご要望など、今回も貴重なご意見をお寄せいただきました。

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

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

---

## 個人会員（敬称略・50 音順）

---

### 会員名

秋元 圭（あきもと けい）

合同会社ことばや 代表 / 株式会社みらい翻訳 技術顧問・アーキテクト /  
AAMT 機械翻訳課題調査委員会 WG3（標準化・共有化ワーキンググループ）委員

---

### 自己紹介

1976 年 東京都練馬区生まれ。獨協大学 外国語学部 英語学科 及び、獨協大学大学院 外国語学研究科 博士前期過程 英語学専攻において英語学（認知言語学）を学びながら、独学でプログラミングを学ぶ。

2003 年から 2014 年 6 月まで 11 年にわたり株式会社クロスランゲージ、株式会社クロスランゲージールアンドディにて機械翻訳ソフトウェアや機械翻訳エンジンの開発を行う。

2012 年 12 月より AAMT（アジア太平洋機械翻訳協会）の機械翻訳課題調査委員会に参加。

2015 年 1 月に合同会社ことばやを設立。

2015 年 2 月より株式会社みらい翻訳にて音声翻訳システムの開発を行う。

---

### MT/翻訳とのかかわり

今思えば、学生時代に音声翻訳ソフトウェアの試作品のようなものを作ったのが、機械翻訳との出会いであった。当時、独学でプログラミングを勉強しながら、様々なツールを作っていたのだが、その中のひとつに音声翻訳ソフトウェアがあった。簡単な機械翻訳プログラム（というか単なる対訳 DB）を作成し、OS に搭載されていた音声認識・音声合成の API と組み合わせて音声翻訳システムを作って遊んでいた。

大学院を卒業後、機械翻訳ソフトウェアの開発会社である株式会社クロスランゲージに入社し

た。入社当初はユーザーインターフェイス担当として Word/Excel/PowerPoint/Internet Explorer 用の多言語翻訳対応アドインや、PDF 翻訳ツール、OCR と機械翻訳を組み合わせた画面キャプチャ翻訳ツールなどの開発や、パッケージ製品の開発管理に従事した。その後は翻訳エンジン担当として、ルールベース機械翻訳エンジンと連携する多言語用例ベース翻訳エンジン（多機能翻訳メモリエンジン）の研究開発、特許における機械翻訳利用に関する調査、機械翻訳の翻訳精度を自動的に評価する手法の研究、機械翻訳エンジン SDK の開発等を担当しつつ、エンジングループの責任者として機械翻訳エンジン開発全体を統括しつつ、プロジェクトマネージャとして様々なプロジェクトの統括を行った。

2015 年 1 月に以下の事業を行う合同会社ことばやを設立した。

- ・ 言葉（自然言語）の研究
- ・ 自然言語処理技術の調査、研究、評価及び実験
- ・ 言語データの収集、解析、分析、評価、作成、提供及び販売
- ・ ソフトウェア製品の企画、設計、開発、製造、販売、導入及びサポート
- ・ ソフトウェアサービスの企画、設計、開発、製造、販売、導入及びサポート
- ・ 自然言語に関する業務のコンサルティング

現在は、株式会社みらい翻訳にて、音声認識と機械翻訳と音声合成を組み合わせた音声翻訳システムの開発を行っている。

また、2012 年より AAMT 機械翻訳課題調査委員会 WG3（標準化・共有化ワーキンググループ）委員に参加し、用語集（ユーザー辞書）の共通フォーマットである UTX（Universal Terminology Exchange）の仕様策定を行ってきた。2014 年より UTX と各社ユーザー辞書の変換を行うツールである UTX 変換ツール（UTX Converter）の開発をオープンソースで行っている。先日正式版を公開したのでぜひ利用していただければと思う。<http://utxconv.sourceforge.net/>よりダウンロードしていただくと、誰でも無償で使うことができる。

#### MT および翻訳業界に期待すること

2020 年の東京オリンピックが決まって以降、機械翻訳のニーズが急速に高まっていると感じる。現在はまだ機械翻訳の精度について懐疑的な意見を聞くことも多いし、機械翻訳の開発者としても実用化の難しさは実感している。2020 年まであと 5 年。この 5 年でどこまで機械翻訳の精度を向上できるかわからないが、5 年後には、少なくとも観光など特定ドメインの機械翻訳については実用化が達成でき、多くの人が機械翻訳を活用し、言葉の壁を意識せずにコミュニケーションがとれる社会が実現していると良いと思う。

#### AAMT への要望

私自身が AAMT の課題調査委員を務めているので、要望というより抱負となるが、機械翻訳の精度向上と普及のために役に立つ活動を行っていききたい。標準化・共有化のワーキンググループに属しているが、UTX だけでなくその他の規格の策定や、現在あまりおこなえていない「共有化」についても取り組んでいきたい。

# 協会活動報告

(2015 年 1 月～2015 年 4 月)

## 機械翻訳課題調査委員会

2015 年 1 月 20 日 (2014 年度 第 9 回)

- ① 前回委員会の議事録の確認
- ② 各 WG の活動について (各 WG に分かれて議論)  
(WG1、WG2)
  - ・ 中日基本文テストセットサイトの評価  
評価作業発注について (経緯の共有)  
今後のスケジュール  
サイト公開について
  - ・ AAMT/Japio 特許翻訳研究会 評価拡大部会について  
状況の共有  
今後の予定
  - ・ 来年度計画について  
オリンピックにかかわる翻訳関係テーマの検討(WG3)
  - ・ UTX 仕様に委員の名前を載せるか検討
  - ・ UTX ファイルの作り方の説明と Excel 対策  
Excel ファイルの読み書きのためにライブラリを買うか、VBA でやるか検討
  - ・ UTX Converter のマニュアル作成について  
費用、公開方法、ファイル形式について
  - ・ 仮の予算配分
  - ・ 用語データ  
用語集の追加について検討
  - ・ UTX 1.20 仕様
- ③ 全体会議
  - ・ 各 WG 検討内容報告
  - ・ 翻訳祭アンケート集計
  - ・ 総務省プロジェクトの状況

2015 年 2 月 27 日 (2014 年度 第 10 回)

- ① 前回委員会の議事録の確認



② 各 WG の活動について（各 WG に分かれて議論）

(WG1、WG2)

- ・ 中日基本文テストセットサイトの評価  
評価作業発注について（経緯の共有）  
今後のスケジュール  
サイト公開について
- ・ AAMT/Japio 特許翻訳研究会 評価拡大部会について  
2015 年度の報告内容  
特許用評価サイトの作成  
今後の予定
- ・ 来年度計画について  
利用者側からの評価について 具体的な活動  
オリンピックにかかわる翻訳関連テーマの検討

(WG3)

- ・ UTX Converter の翻訳  
UI メッセージとマニュアルの翻訳について検討
- ・ 松江 ISO 会議  
UTX を ISO 化するか否か検討
- ・ UTX 論点シートの検討
- ・ UTX のサンプル辞書作成について

③ 全体会議

- ・ 各 WG 検討内容報告
- ・ 総務省プロジェクトの状況

2015 年 3 月 27 日（2014 年度 第 11 回）

① 前回委員会の議事録の確認

② 各 WG の活動について（各 WG に分かれて議論）

(WG1、WG2)

- ・ 中日基本文テストセットサイトの評価  
納品物の確認  
簡単な集計結果  
設問評価と人間評価の比較  
今後の進め方
- ・ 利用者視点からの評価について（継続）  
利用者側からの評価について 具体的な活動
- ・ オリンピックにかかわる翻訳関連テーマの検討

(WG3)

- ・ UTX Converter マニュアルの翻訳  
懸案事項の検討
- ・ 次バージョン UTX Converter  
英日/日英以外の言語への対応（中国語、韓国語）、Systran 最新版について
- ・ 今年度活動報告と来年度活動計画  
昨年版をもとに文面案を作成
- ・ UTX 1.20 仕様（UTX 1.20 beta）  
二部構成への書き換えを検討
- ・ UTX 論点シートの検討

### ③全体会議

- ・ 各 WG 検討内容報告
- ・ 翻訳フェア、総会に向けた情報共有

2015 年 4 月 23 日（2015 年度 第 1 回）

- ① 前回委員会の議事録の確認
- ② 各 WG の活動について（各 WG に分かれて議論）  
（WG1、WG2）

- ・ 2014 年度成果・2015 年度目標の確認
- ・ 中日基本文テストセットサイトの評価  
評価サイトの不備修正  
翻訳祭までの作業  
発表スケジュール
- ・ 利用者視点からの評価について（継続）  
利用者側からの評価について 具体的な活動
- ・ 翻訳フェアの展示コーナー準備
- ・ オリンピックにかかわる翻訳関連テーマの検討

### （WG3）

- ・ UTX Converter 公開  
プレスリリースやサポートについて検討
- ・ UTX 論点シートの検討  
言語共通の品詞設定などについて検討
- ・ UTX 1.20 仕様（UTX Specification Version 1.20 (beta)）  
beta 版のままで、日本語訳を作成開始
- ・ その他  
変更点一覧「Summary of changes: UTX 1.11 to UTX 1.20beta」、総会用 WG3 の  
2014 年度活動報告・2015 年度活動計画を確認

### ③ 全体会議

- ・各 WG 検討内容報告
- ・翻訳フェア、総会に向けた情報共有

#### **インターネットWG**

- ①総会・MT フェアに対応した AAMT ホームページの更新
- ② 旧コンテンツから新コンテンツへの誘導の推進
- ③AAMT サイトのセキュリティー対応

#### **編集委員会**

2015 年 2 月 18 日（2014 年度 第 4 回）

- ① AAMT Journal No.58 の記事状況について
- ② AAMT Journal No.59 の執筆依頼計画について

## AAMT/Japio 特許翻訳研究会

2015 年 1 月 23 日(金)(2014 年度 第 6 回)

1. 議事録の確認
2. 園尾様の発表
3. MT Summit における Workshop について
4. 26 年度 AAMT/Japio 特許翻訳研究会報告書  
26 年度報告書の目次案とスケジュール
5. その他
6. 次回の開催について
  - ・ 開催の日時(場所) (3 月 13 日(金曜日) 18:00~20:00)
  - ・ 主な議題

2015 年 3 月 13 日(金)(2014 年度 第 7 回)

1. 議事録の確認
2. 26 年度 AAMT/Japio 特許翻訳研究会報告書 発表
3. その他
4. 次回の開催について
  - ・ 開催の日時(場所) (5 月 15 日(金曜日) 18:00~20:00)
  - ・ 主な議題

## AAMT/Japio 特許翻訳研究会 拡大評価部会

2015 年 1 月 23 日(金)(2014 年度 第 3 回)

1. 前回議事録の確認
2. 報告書内容概要(各グループより)
3. その他

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

---

AAMT ジャーナル 59 号をお送りします。

今号の巻頭言は、東京大学 大学院情報学環・教育学研究科 影浦峯先生より、御寄稿を頂きました。

理事の坂本義行先生からは、機械翻訳の歴史に関する貴重な御寄稿の第五弾を頂きました。

また、近年の統計的機械翻訳技術の動向に関する解説として、奈良先端科学技術大学院大学 Graham Neubig 先生より、特に日本語を対象とした統計的機械翻訳技術に特化した現象に関する御寄稿を頂きました。

その他、東京外国語大学 大学院総合国際学研究院 佐野洋先生からは、ものごとの捉え方における英語と日本語の違いに関する御寄稿を頂きました。

一方、今号に先立ちまして、4 月 9-10 日(木,金)に、TAUS (Translation Automation User Society) 東京フォーラムが開催されました。TAUS 代表の Jaap van der Meer 様には、翻訳業界をめぐる動向を、また、エレクトロスイス ジャパン 中村哲三様には、TAUS 東京フォーラムの参加報告を、それぞれ、御寄稿頂きました。

また、AAMT 前会長の豊橋技術科学大学 井佐原均先生からは、豊橋技術科学大学多言語翻訳研究本部主催のもと、経済産業省、テクニカルコミュニケーター協会、日本翻訳連盟、および、AAMT が後援して、4 月 7 日(火)に開催された多言語情報発信シンポジウムの開催報告を御寄稿頂きました。

加えて、関西外国語大学 内田真弓先生からは、本年 1 月に開催された第 3 回 JTF 関西セミナーの参加報告を御寄稿頂きました。同セミナーにおけるご講演の内容のうち、特に、株式会社みらい翻訳について、同社社長 栄藤稔様からの御寄稿を頂きました。

AAMT 会長の名古屋大学大学院 中岩浩巳先生、および、富士通研究所 長瀬友樹様からは、第 14 回日中自然言語処理共同研究促進会議参加報告を御寄稿頂き、AAMT 委員会活動の紹介の様子等についてご報告頂きました。

日本特許情報機構 早川貴之様からは、「特許版・産業日本語の取組について」を、および、長谷川雅昭様、大塩只明様からは、「世界特許情報全文検索サービス」をそれぞれ、御寄稿頂きました。

また、機械翻訳課題調査委員会 目次由美子様からは、「Transit 最新サービスパック 8 について」を御寄稿頂きました。



AAMT 内の活動報告として、機械翻訳課題調査委員会から、「UTX 変換ツール正式版の公開」、および、「AAMT フォーラムメールマガジンバックナンバー」を掲載いたしました。

その他、「AAMT 会員のひろば」の企画におきましては、法人会員 2 件、個人会員 1 件の紹介文を掲載しました。

# AAMT

Asia-Pacific Association for Machine Translation

## AAMT 入会のご案内

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

機械翻訳にご関心のあるすべての方にご入会をお勧めします。

### \* \* AAMT 会員の特典 \* \*

#### 1.AAMT Journal の購読ができます。

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

#### 2.機械翻訳関連の最新情報をメールでお届け

会員専用メーリングリストで、最新の機械翻訳関連の情報をお届けします。

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

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

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

#### 4.関連機関の主催する国際会議に参加できます。

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

AMTA\* : Association for Machine Translation in the Americas

EAMT\*\* : European Association for Machine Translation

#### 年会費は以下の通りです。

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

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

ご関心のある方は、事務局までお問い合わせください。

アジア太平洋機械翻訳協会（AAMT）

ホームページ：<http://www.aamt.info>

電子メール：[aamt-info@aamt.info](mailto:aamt-info@aamt.info)

## 入会申込書

以下の通り、アジア太平洋機械翻訳協会の会員申し込みを致します。

申込日 20 年 月 日

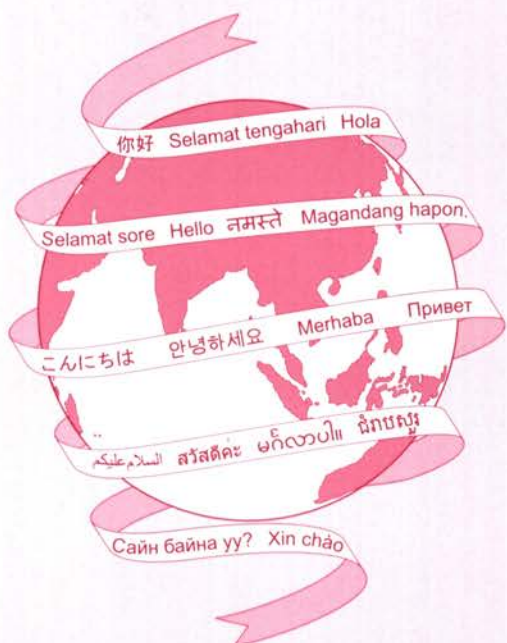
氏名（ローマ字）	
氏名（漢字）	
電話番号	
メールアドレス	
所属先	
所属先住所	〒
種別	<input type="checkbox"/> ユーザ <input type="checkbox"/> 研究開発者 <input type="checkbox"/> その他
機械翻訳に関するお知らせメールの配信	<input type="checkbox"/> 希望する <input type="checkbox"/> 希望しない

コメント

--



# AAMT



## AAMTジャーナル No.59

発行：アジア太平洋機械翻訳協会（AAMT）

ホームページ：<http://www.aamt.info>

住所：〒171-0014 東京都豊島区池袋2-55-2鈴木ビル3階  
(株)日本システムアプリケーション内

phone：03-5951-3961 fax：03-5951-3966

編集委員会：宇津呂 武仁 小谷 克則 大倉 清司

鈴木 博和 阿部 さつき 釜谷 聡史

表紙(図部分)デザイン：阿部 さつき

事務局：神崎 享子 荻野 孝野

印刷所：株式会社ユリクリエイト

Asia-Pacific Association for Machine Translation (AAMT)

c/o Japan System Application Co., Ltd.

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

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