# みらい翻訳における機械翻訳サービス

企業リポート

栄藤　稔*

Machine Translation Services by Mirai Translate, Inc.

Key Words：Deep Neural Network, Natural Language Processing, Machine Translation, TOEIC

## はじめに

2014年10月株式会社みらい翻訳を設立した．爾来代表取締役社長を務めている．同社の株主はモバイル通信事業のNTTドコモ，韓仏米に開発拠点を持つシストラン・インターナショナル，総合家電メーカーであり産業応用事業を強化中のパナソニック，国内最大規模の産業翻訳サービス企業である翻訳センターからなり，国際的な合弁事業としてコンピュータによる自動翻訳（機械翻訳）サービスを国内企業中心に提供している．

みらい翻訳は従業員のほとんどが自然言語研究者，ソフトウェア技術者，翻訳者からなっている技術開発会社である．設立時からNTT研究所との技術交流があり，一方で情報通信研究機構（NICT）と共同研究を開始している．設立時に腐心したことは，グローバル展開の視点を持ちつつ，日本国における言語資源と技術を当社に集中させるエコシステム構築である．この準備はうまく行ったと自負している．会社が実質的にスタートした2015年春当時，最新の機械翻訳技術は統計的翻訳（Statistical Machine Translation, SMT）であった．これに後述する数千万規模のコーパスで学習すれば，国際的にも競争力のある性能があると期待していた．ところが，2016年秋に米国グーグル社がニューラル機械翻訳（Neural Machine Translation, NMT）を商用運用したことにより，景色が一変した． NMTにより性能が格段に向上したのである．

これを受けて，みらい翻訳はSMTに関する技術開発とサービス販売を全て停止させ，2017年の1年間をNMT開発に集中した．2018年の春を迎えて，ようやくみらい翻訳独自開発のNMTによる機械翻訳サービスを市場に出すこととなった．この会社の3年間の運営を通して得られた知見を紹介したい．

* Minoru ETOH

1960年11月生まれ
広島大学大学院　工学研究科　システム工学専攻博士前期課程（1985年）大阪大学博士（工学）（1993年）　松下電器産業（現パナソニック）、NTTドコモを経て現在、大阪大学　先導的学際研究機構　教授　兼　株式会社みらい翻訳　社長
TEL：03-6268-8080
E-mail：etoh@ieee.org

## 統計的機械翻訳（SMT）

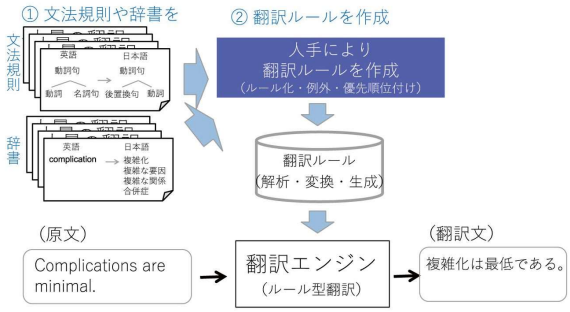


図1　ルール型機械翻訳（RBMT）．

統計的機械翻訳（SMT）について説明する前にルール型機械翻訳（Rule-Based Machine Translation, RBMT）について触れておく．2010年ごろまで主流だった方式である．文法規則や辞書から人手により翻訳ルールを登録し，ルールに基づき翻訳文を作成する．教科書的な文法ルール，構文が明確な文章であれば，翻訳しやすいという特長を持っている．何しろ統計的翻訳と違い機械学習用のデータが必要でない．必要なのは，専門用語辞書の登録で，新語の翻訳が可能となる．短所はルールのメンテナンスである．例外や優先順位付けは細かく設定はできるが，人手による登録コストがかかる．語順が違う言語間でも，分野を問わず長文を大崩れせず翻訳できるが，一定レベル以上の精度向上はコスト要因により困難で，

訳質に不自然さが残る．これらの欠点を克服するために2010年くらいから商用化されたのが，統計的機械翻訳である（図2）．対訳文をコーパスという．これを数百万から数億収集して機械学習することにより，確率的な組み合わせ計算により翻訳文を作成する． SMTは単語，句，構文に基づいたコーパス解析が基本となる．

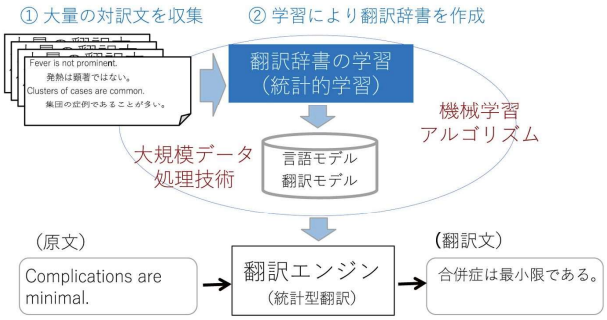


図2　統計的機械翻訳（SMT）．

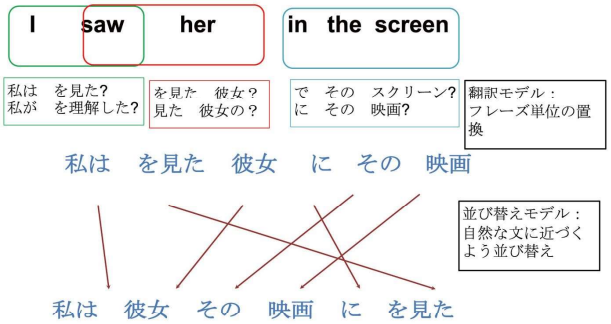


図3　統計的機械翻訳（SMT）の動作例．

図3にその例を示す．統計的な最適性から，文章を様々な場合分けをして区切り，それを並び替える．要はコピペを統計データに照らして最適に行う世界である．結果として，短い文章である程，自然な文章として翻訳でき，また口語等で構文がくずれた文章でも翻訳しやすいという特長を持つ．一方で，翻訳精度の向上には，コーパスの「質と量」が肝となる．こうなると，翻訳エンジンの性能を決めるのは，エンジンのアルゴリズムではなく，データだ．

### 顧客企業にデータがない事件

SMTの時代（2010-2016）からデータを持つことが機械翻訳の性能を決めた．特許や外国政府への各種申請書，マニュアル，議事録などの翻訳を産業文書翻訳という．機械翻訳を産業文書に適応する流れを図4に示す．顧客企業の話す言語にチューニングをかけること，すなわちカスタマイゼーションを行うことにより機械翻訳の性能を上げることが期待される．これが米国西海岸を中心としたインターネット企業の機械翻訳に対抗できる戦略となる．洋服屋に例えるなら，吊るしのスーツを売るのではなく，オーダーメイドの服を作るのである．

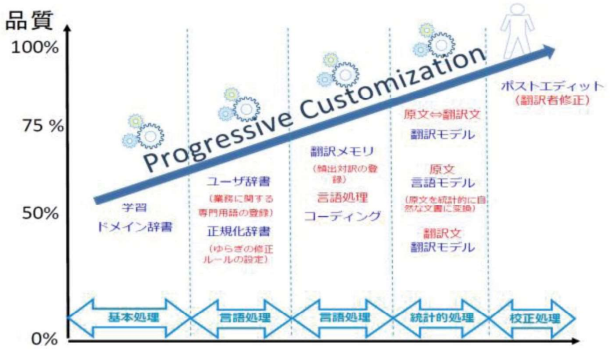


図4　産業文書機械翻訳の流れ．

みらい翻訳のパートナー企業であるシストラン・インターナショナルはこのカスタマイゼーション戦略により2015年に100ライセンスを超えるSMTによる機械翻訳システムを販売した．顧客企業には国防関連の政府機関，多国籍金融機関，自動車メーカー，航空機メーカーがある．みらい翻訳も2015年から，その戦略に従って販売を試みた．課題だったのは，シストランの欧米顧客企業とみらい翻訳の日本顧客企業の間に「デジタル化」のレベル差があることであった．例えばヨーロッパの航空機メーカーは英仏独の言語で従業員のコミュニケーションが行われている．グローバルに製品を売るために，産業文書が多言語で管理されている．一方で日本企業でのデジタル化は進んでいなかった．顧客企業にデータがないのである．SMTの性能を決めるのはコーパスの質と量であるから，上記の戦略は以下の単純な式にまとめることができる．

**機械翻訳性能＝みらい翻訳の持つデータによる基本性能＋顧客企業が持つデータにより向上する性能**

翻訳結果と正解翻訳例との一致率を測る自動指標としてBLEU [1)] がある．翻訳性能は最終的には目視による主観評価が必要となるが，BLEU値が自動指標としてよく使われる．経験的にSMTでは35を越えれば実用的である．図5は成功例の一つである．あ

る顧客企業がもつデータ，すなわち対訳コーパスを技術文章から抽出した．それが100万あれば，SMTの特性から当該企業の技術文書翻訳で最高の翻訳性能が出せることが示せた．そのような企業を増やしたい．残念ながら2016年時点で，そのような顧客企業は稀有であった．この経験で筆者は機械翻訳事業とは別に「日本企業のデジタル変革」を使命と感じるようになった．

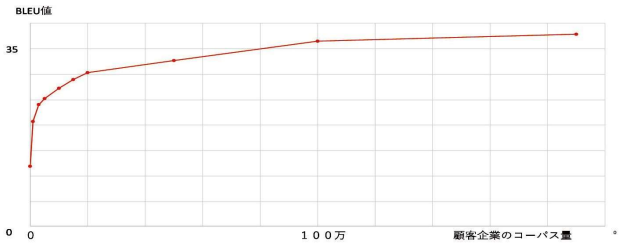


図5　顧客企業データ利用による翻訳性能の向上．

## スクランブル発進

2016年6月米国グーグル社がニューラル機械翻訳（NMT）の発表を行ってから，その性能向上に驚愕した．機械翻訳の主観評価は翻訳結果の妥当さ（adequacy）と流暢さ（fluency）の5段階評価で行うことが多い[2]．妥当さは，5：すべての情報，4：ほとんどの情報，3：多くの情報，4：少しの情報，1：情報なしが翻訳文に含まれているかどうかを主観で判断する．流暢さは5：全く問題ない，4：良い，3：非母国語的，2：不自然，1：理解不能というレベルで翻訳文の自然さを主観で判断する．これまで，グーグル社のSMTもみらい翻訳のSMTも評価値は妥当さも流暢さも3から3.5くらいであった．可でもなく不可でもなく，なんとか通じるレベルである．それがグーグルNMTの登場によって4のレベルに到達してしまった．流暢さが良く，伝えたい情報がほぼ含まれるということだ．この違いはどこから来るのか？

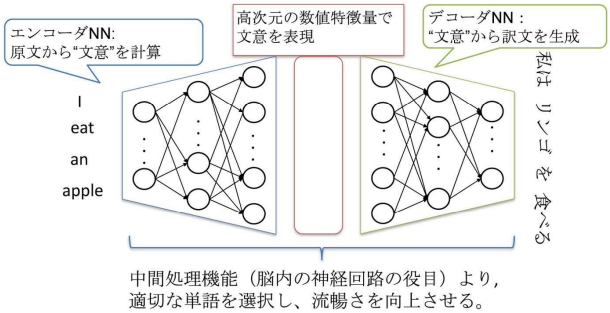


図6　NMTの直感的理解．

SMTもNMTも過去の対訳データを大量に必要とする．その点は同じだ．異なるのは，SMTがフレーズ単位の並び替えを学習するのに対して，NMTは文章を数値的特徴量（文意）へと直接変換する．基本的には，フレーズ（名詞，形容詞，動詞からなる文節）を抽出しない．入力（原文）と出力（訳文）の対応づけを数百万から数億行うことによって，入力から文意を表現する数値表現と数値表現から流暢な文章を出力とする構造を学習する（図6）．並び替えではないので，「私行学校した」と入力すると，“I went to school.”という訳文が出力される．

それをまとめると図7に示すようにSMTとNMTの言語処理は，一段レベルが違うものとなる．今のNMT技術の基本形は2014年ごろに発表された．みらい翻訳を設立した時，筆者はNMTの実用化は2020年ごろだろうと想定していたが，それからたった2年と少しで実用化された．見通しが甘かった．

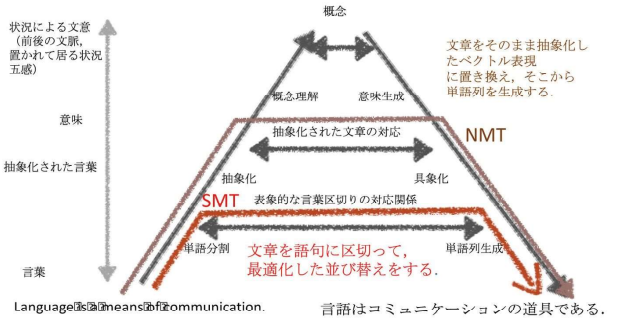


図7　Vauquoisの三角形[3]に準じたSMT, NMTの解釈．

みらい翻訳では，2016年末，SMTに関する開発と販売を中止し，全ての経営資源をNMT開発に投下した．一番の懸念は，SMT時代に有効だったカスタマイゼーション戦略がNMTでも通用するかどうかだった．巨大データによるNMT対専門データによるNMTの戦いで，後者が局地戦（言葉が専門的な業界向け機械翻訳）で勝つチャンスがあるのかどうか？　1年間の開発の結果，答えは「はい．」だった．結果として2017年1月にみらい翻訳は自社のNMTサービスを提供するところまでこぎつけた．

## 性能が臨界点を超えた時に見える景色

機械と翻訳者が100文のビジネス会話を日英の双方向で翻訳し，結果を専門家が5段階評価した．

NMTはレベル4を越えてSMTに日英，英日も優っている．注目すべきは，プロの翻訳者よりも日英で妥当さ，流暢さで勝ち，また英日で若干及ばないでいる．この結果は同品質で，人力で数時間かかる翻訳作業が数分に短縮されることになることを示唆している．

機械翻訳をTOEIC換算値で性能評価する手法が提案されている[4]．現時点で，機械が意訳をすることや長文読解で要約することはできない．　できることは，人と英作文の良さを競った時，TOEIC何点の人と五分の勝負となるか？　その点数を推定するというものである．このために，

1．過去半年にTOEICを受験した被験者を成績400点から900点までの間で数十人集め，300問の日本語原文を自力で英訳してもらう．
2．機械翻訳で同じ日本語原文を英訳する．
3．日英バイリンガルの評価者3名に被験者対機械翻訳の優劣を決判定してもらう．
4．勝率0.5（＝機械翻訳と人手翻訳が互角）となるところをTOEICスコア換算値とする．

を行なった．その結果を図8に示す．TOEIC 400点～700点の被験者との勝負では機械が勝率8割以上で勝つことが分かる．どこで勝率5割になるかということを検証するための直線を引くと結果はTOEIC 957点となった．実際，TOEIC 900点以上の被験者を集めていなかったため，外挿でその点を性能とするのは難しいため，TOEIC 900点以上の英作文能力という性能を結論とした．「モジュールの性能がある臨界点を超えた時，それまでは不可能だと思われていたサービスが急激に立ち上がる．」という経験を筆者はしてきた．2006年当時，筆者が，携帯

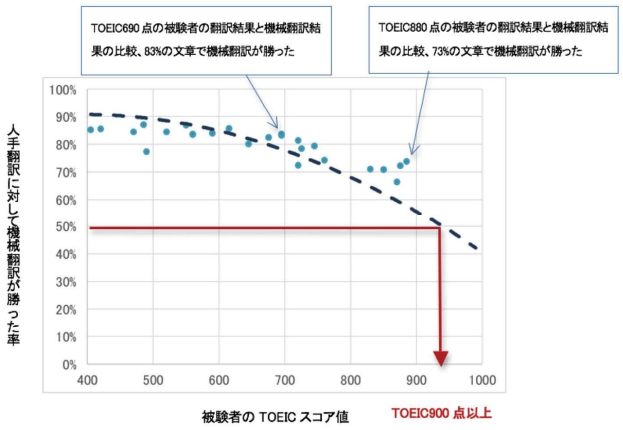


図8　英作文のTOEIC換算値．

電話に音声認識インタフェースを導入した時，周囲の反応は冷ややかだった．その後，認識率が95%を超えるあたりで，急激に実用的なサービスが立ち上がった．機械による画像認識はもはや人間の能力を超えて，画像検索や監視セキュリティなどのサービスが普通に行われている．機械翻訳はこれまで，文法構造の異なる言語族間での性能が低かった．NMTの登場により，3年で主観評価が「可」から「良」に変わった．これから見える景色は劇的に異なるだろう．

## おわりに

NMTの性能はSMTと同じように，データの質と量で決まる．米国グーグル社の機械翻訳に勝てるのですか？　という質問については，こう答えたい．「データがあるところで勝ち，ないところで負ける．」医療現場での会話や，産業文書の翻訳で，そこにしかないデータがあれば，オーダーメイドの仕立て屋のアナロジーで，高度にカスタマイズされた機械翻訳サービスが提供できる．現場のデータを観測し，それを機械翻訳に活かす．そのような仕組みを作っていきたい．それには日本企業のデジタル変革が必要である．そこに産業間の協働に生き残りのシナリオがあると感じる．

## 参考文献

1) K. Papineni, S. Roukos, T. Ward and W. Zhu. “BLEU: a method for automatic evaluation of machine translation.”, In ACL’02: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics, pp.311-318. (2002)
2) 安田圭志，隅田英一郎．“テキストの自動評価機械翻訳の研究・開発における翻訳自動評価技術とその応用”，人工知能学会誌　vol.23, pp.2-9. (2008)
3) Bernard Vauquois. "A survey of formal grammars and algorithms for recognition and transformation in mechanical translation." In *Ifip congress (2)*, vol. 68, pp.1114-1122. (1968)
4) 菅谷史昭，竹澤寿幸，横尾昭男，山本誠一．“音声翻訳システムと人間の比較による音声翻訳能力評価手法の提案と比較実験”，信学論，D-Ⅱ, Vol. J84-D-Ⅱ, No.11, pp.2362-2370. (2001)