

《論文》

AI 同時通訳の未来と人間の同時通訳者

熊谷ユリヤ

1. はじめに

テクノロジーの進化により AI（人工知能）による MT（機械翻訳）が発達し、特に、ニューラル翻訳、ディープラーニングのお陰で飛躍的な進化を見せ、「自動『同時通訳技術』の社会実装実験も近い将来予定されている。その過程で、様々なレベルの膨大な種類のツールが日常的に使われている。「人間の通訳者は、将来的に AI に代替され淘汰されるのか」については、「そういう時代が来る」という意見と、「機械には対応しきれない人間的な要素があるため、そうはならない」とする意見が分かれている。

本稿では、既に自動文字起こし機能、自動翻訳、自動読み上げ機能を組み合わせて使用されている AI 音声翻訳の現在と通訳者、更には、起点言語の内容、表現、感情を、直訳ではなく目標言語の文化的文脈の中で自然な音声に変換するソリューションとしての AI 通訳の未来を考察する。また、人間の通訳者の未来、とりわけ、逐次通訳（CI）とは異なり、専門分野や学術分野も含む難易度の高い内容を、数秒遅れで通訳する同時通訳（SI）に焦点を当てる。英語会議（同時）通訳者である筆者の経験的知識も交えながら、人間の同時通訳者に未来はあるのかについて考察したい。

2. テクノロジーに代替される仕事

近年、この分野では、DeepL や Google Translate などの翻訳エンジン、Google Cloud Speech-to-Text、Watson Speech to Text、などの AI 音声認識・STT（Speech-to-Text 自動文字起こし）、VOICEVOX、VOICEPEAK、

AITalk などの TTS (Text-to Speech 音声合成変換) などのソリューションが一般化している。「翻訳機を超えた夢の AI 通訳機」「まるで専属の同時通訳」「AI が自動で翻訳・同時通訳、夢を叶える次世代のスマートイヤホン」と称する商品まで、進化が目覚ましい。そのため、「人間の通訳・翻訳家は AI に仕事を奪われる」という記述が特にインターネット上でも散見される。

WEF (世界経済フォーラム) が 2020 年に発表した報告書は、2025 年までに世界で約 8500 万の仕事が自動化と機械化により AI とロボットにとって替えられる一方、約 9700 万の新たな仕事が生まれると予測している。2025 年には世界の雇用の約 47% が機械とロボットが担当するという。パンデミックにより、この傾向はさらに加速している。「いつまでも人間が必要で AI が代替できない産業」に翻訳者・通訳者は言及されていないが、創造的、共感的な分野、複雑・戦略的な仕事まで多岐にわたるため、翻訳者、特に文学・文芸翻訳家と、通訳者、特に同時通訳者は、「いつまでも人間が必要な仕事」に該当すると考えられる。

3. 自動翻訳と自動通訳

翻訳者・翻訳家は起点言語を文化的文脈に沿って読み取り、訳すほかに、辞書や資料、ツールを駆使して目標言語の文化的文脈に書き換える。実務翻訳では CAT (コンピュータ支援翻訳)、MAHT (Machine-Aided Human Translation 機械が人を補助する人間による翻訳) も一般的となっている。一方、通訳者は予測不能な内容やアクセントの強い非母国語話者も多い聴解の後、リアルタイムで口頭による訳出を行う。文字と音声という違いのみならず、時間的猶予のある翻訳と即時性が求められる通訳は、本来、同一視して論じるべきではない。

しかし、現在「自動通訳」「自動同時通訳」ツールとされているものは、実際は、次の 3 つの段階を経た音声機械翻訳である。[1] 起点言語の発話音声を STT で文字起こしし、[2] AI がディープラーニングでの蓄積に基づいて文字から文字を翻訳した後、[3] TTS がそれを合成音声に戻す。近

年、格段に翻訳性能が進化した AI を用いた機械翻訳は、人間の脳を模した設計のニューラルネットワークで結果を予測し、訳出の質を自己評価しつつ学習ができるようになった。「起点言語の翻訳がリアルタイムに文字で表記され、それを更に合成音声で読み上げるシステム」が、人間による同時通訳に似ている、というのが「現時点での自動同時通訳」である。

翻訳業界では、すでに統計的機械翻訳とニューラル機械翻訳により、人間の翻訳者の役割が変化している。人間が最初から翻訳をする従来の翻訳者に代わって、機械が翻訳したものをポストエディターが Brush-up 編集する PEMT（ポストエディット機械翻訳）が主流である。文学・文芸翻訳の分野では人間の翻訳家の役割が現在は大きいものの、AI 性能の進化により、分業の比率が上がることは十分考えられる。それが、同時通訳に至るには即時性のハードルはあるものの、翻訳の世界と同様の傾向が、通訳の世界で更に加速して起ころうとしている。

4. 人間による同時通訳

起点言語の話者と目標言語に訳す通訳者が交互に話す形式の CI（逐次通訳）は、2倍の時間がかかるのに比べ、SI（同時通訳）は、話者の発言を聴きながら数秒遅れでほぼ同時に訳していく。国際会議や学会会場には、マイク、ヘッドセット、通訳操作機を備えた二人着席の遮音ブース、参加者が訳出を聞くためのレシーバーなどの機材、機材操作をするエンジニアが必要である。専門性の高い内容が多い上に、聴きながら訳を考えて話しながら次のチャンクを聴くという複数のタスクを同時に行うため、同時通訳をする会議通訳者には、高度な集中力が必要となる。疲労のため精度が低下しないように、2人から3人で15分から30毎に、通訳とサポート（桁の大きい数字や固有名詞のノートテイキング）を交代して担当する。また、専門分野や学術分野は毎回異なるため、会議毎に事前資料や周辺資料が必要で、多くの場合、予習後に発言者との打ち合わせがある。

同時通訳の経験やスキルにより各エージェントが A,B,C などのランクを

付けているが、パフォーマンスには個人差があることは否めない。「同時通訳はわからないことを飛ばすので訳出率は60から70%」「機械通訳もミスは多いが人間通訳もミスが多い」などインターネットで見られる誤った記述は、ランクが低い通訳者の訳出を聞いたか、優れた通訳者が、直訳や逐語訳を避けるため、目標言語では不要な部分を除いて自然な訳出を行ったと考えられる。経験豊富な通訳者は、アクセントや癖のある発話の訳出も精度が高く、ユーモア、皮肉など発言者の言外の意も理解し、紛らわしい表現も文脈で捉え、直訳ではなく自然で聴きやすく、憑依型の訳出をすることもできる。

新型コロナパンデミックで多くの会議や学会がリモートになったことに伴い、RSI（遠隔同時通訳）の需要も高まった。リモートと対面のハイブリッド会場か各同時通訳者宅のワークステーションから、Zoom 言語機能、KUDO や通訳会社独自の RSI システムにより、同時通訳を提供する。機材やエンジニアの費用を節約でき、遠隔の場合は通訳者の移動や宿泊費の節約にもつながるため、パンデミックが終息しても、この傾向は続くと思われる。

反面、現在のリモート環境下では、音声のテクニカルな問題もあり、本来、機械より人間の方が優れている発言者の感情や言外の意を聴き取ることが、ハードルの1つとなっている。更に、遠隔地のパートナーとの交替のタイミングが難しく、ノートテイキングに頼ることができないなど、通訳者の負荷が重くなっているため、テクノロジーがそれを補完している。

5. テクノロジーの支援を受けた人間通訳CAI とMAHI

人間の同時通訳者の多くは、テクノロジーの進化を脅威とは捉えておらず、上述の複雑で時間のかかる準備と、多重タスクとストレス度の高い現場の両方で、機械を補助として活用することは一般化しつつある。翻訳の場合と同様に、CAI(Computer-Assisted Interpreting コンピューター支援通訳)、MAHI (Machine-Aided Human Interpreting 機械が補助する人間による通

訳) も急速に普及している。

通訳業界では、「準備が7、8割」という考え方あり、AIが自動翻訳をするためにディープラーニングを行うのと同様に、分野・会議の重要度・難易度によっては長時間をかけて準備をする。同時通訳する会議のプログラムや式次第に基づき、プロフィール、過去の論文や資料、アクセントが予想される場合は動画を検索して耳慣れをする。発表資料や会議資料受け取り後は、必要に応じてファイル形式を変換・抜粋し、資料が膨大で高難易度の場合、機械翻訳にかけて対訳に目を通す。専門用語を抽出して対訳語彙リストや必要に応じて数字表を作成し、持込用のパソコンやタブレットでブース内でも参照できるようにするケースも少なくない。

テキスト資料ではなく動画や音声の資料しかないアフレコ同時通訳などの場合は、必要に応じてSTTにかけて語彙・表現リストを作成する。また、読み原稿がある場合は話者が早口になる傾向があるため、TTSの1.2倍や1.5倍速にかけて口慣らしをすることもある。2、3人で分担していても、質疑応答やディスカッションのためには、他の通訳者の担当する部分でも同様の準備が必要である。繁忙期には別件の同時通訳業務や準備と並行してこれらを行うため、テクノロジーを味方につける必要があるというのが実感である。

同時通訳の現場で、桁の大きい数字、初出の固有名詞、難解な語彙をノートテイキングする場合は、iPadとApple Pencilや、キャンバスディスプレイ搭載ペーパータブレット ReMarkble 2などのデジタルライティングシステムも使用されている。ブースメイトが隣にいないRSI（遠隔同時通訳）の場合は、LINEの同時通訳用スタンプや、2台目の端末をミュートのテレビ電話にして交代のタイミングをつかむ。また、ノートテイキングを補完するため、ワークステーションに数台のディスプレイを配して、語彙リスト、発表資料に通訳しながらアクセスしなければならない。

最近では、短時間で詳しい分野については、全くの一人で行う「ソロ同通」を打診されるケースもあり、ますますCAI, MAHIの需要が高まっている。InterpretBankは、いわば、人間のパートナーにかわる（Artificial Boothmate 人工ブースメイト）である。人自動音声認識機能を搭載し、検

索した専門用語や桁数の大きい数字を通訳者のワークステーションの画面にリアルタイムで表示してくれる。人間のブースメイトが書いたメモを見るため横を向いたり、手書き文字を解読するなど、通訳者のマルチタスクに更なる負荷をかけることなくサポートを受けることができる。

6. 機械による「(同時) 通訳」

機械は集中の結果疲労することもなく、一定の訳出精度を保ちつつ長時間通訳し続けることができる。学習した情報に瞬時にアクセスすることができるため、通訳者が求めるような意味の事前資料も打合せも不要である。開発の費用や初期費用は膨大であっても、利用料金自体は、人数分の通訳料、対面の場合の機材やエンジニア料とは比較にならない低額であることが多い。また、クライアントが選択するシステムやプログラムに応じて、パフォーマンスの精度が保証される。現在は音声同時翻訳に過ぎない機械翻訳は、意味のチャンク化、訳し下し、予測、待機、省略、言い換えなども人間の同時通訳者が用いる攻略を学習しつつある。

NICT（国立研究開発法人情報通信研究機構）によれば、2025年にはAIによる同時通訳が実用レベルに達するとして、大阪・関西万博を舞台として日本の自動同時通訳技術を世界に発信する計画である。これは、総務省の委託を受けた多言語自動翻訳技術のステップアップであり、文脈の理解や発言者の意図にも対応できるため、実装実験の場となる。しかし、その適応範囲は、万博会場と周辺、宿泊施設、来場者とのインタラクティブなプレゼンテーションが主である。

日本語と英語のような文型や語順が全く異なる言語間の同時通訳では、意味のチャンクに分割し、目標言語の語順に合わせて文頭から訳しおろしていくスキルが重要である。人間の通訳者は、スラッシュリーディングやサイト・トランスレーションなどの訓練により直感的にチャンク分割ができ、自然な訳出のため文末に来るべき語彙をリテンションできる。AIは、深層学習の過程で、同時通訳者のスキルを模していると言われる。

7. 音声機械翻訳の課題

Meta社は、2022年、テキストを介さず音声から音声へのリアルタイム音声翻訳をリリースした。AI搭載の福建語音声翻訳システムにより、文字体系を有しないためAIに学習させるソースが不十分で機械翻訳ツールを構築するのは不可能だった言語の、即時音声翻訳を実現させた。発話後、瞬時に翻訳された言語が合成音声で再生され、リアルタイムの意思疎通を実現させた。事前に学習させた音声エンコーダを用いることで、福建語の音声埋め込みを他の言語と同じ意味空間に符号化することが実験上は可能になっている。

音声翻訳のもう一つの課題は、起点言語の音声入力から目標言語のテキスト生成の過程で、方言やアクセントをいかに扱うかである。*Johns Hopkins*大学のJHU方言音声翻訳はASR（自動音声認識）、MT、ST（音声翻訳）を駆使してこの問題も克服しつつあり、人間的要素以外は前進している。

8. 通訳者 対 機械

Downieはその著書*Interpreters vs Machines*の中で、AIが支配する世界を通訳者が生き延びることができるのかを論じ、通訳者自身が主導権を握っているとしている。つまり「通訳者が評価の精度が高いものとして広く容認された絶対的基準となる」、通訳の新しい未来を選択することは可能だという楽観説を挙げている。同時通訳タスクの複雑さが一般に知られていないため、もっとPRが必要であるとも指摘している。

反面、人間の通訳が機械に取って代わられるまでの「応急処置」に過ぎない未来や、人間の通訳者が少数のニッチに限定される未来を、ある種の悲観的な選択肢として扱っている。「機械同時通訳」の大きな課題は、より人間的な翻訳をすることであるが、人間的な要素や感情を文脈変換して伝えられる高いレベルの同時通訳者がCAIツールを駆使するなら、完全に機械に取って代わられることはない。むしろ、テクノロジーにより苦行から解放され、より人間的なタスクに集中することが可能となる。

9. 現時点の結論

以上の考察と、CAIによるSI, RSIを実践する同時通訳者としての経験的結論としては、非常に長期的な視点では、従来のような人間の同時通訳者が不要になることは避けられないと考える。今後、一定の期間は過渡期として、同時通訳時のコンピュータ支援通訳ツールを用いて、共存する時代となる。過渡期が何年間、何十年間続くのかは現時点では正確に予測できないが、機械よりレベルの低い同時通訳者は淘汰される。

自動音声翻訳ではない真の自動同時通訳が完成するまでは、緊急避難的に、翻訳が辿った過程を通訳も辿ることになろう。その時点での単純作業は機械が担当し、人間的要素が重視される一部の仕事は、人間の通訳者が担当する事となる。過渡期を過ぎて長期的にも、少なくとも、AI同時通訳を駆使する、新たな役割が創出されると考えられる。次々に開発され目まぐるしく変化する言語コミュニケーション・テクノロジーに今後とも注視が必要である。

参考文献

- 内山将夫, 「NICTにおける自動『同時通訳技術』のご紹介」 AAMT 年次大会 2022 年 12 月 1 日 情報通信研究機構
- Defrancq, Bart (2020), Automatic Speech Recognition in the Booth Assessment of system performance, interpreters' performances and interactions in the context of numbers (John Benjamins Publishing Company)
- Desmet, Bart et. al (2018) "Simultaneous interpretation of numbers and the impact of technological support" Interpreting and technology, 29-59: LanguageScience Press
- Deysel, Elizabeth et. al (2019), Automatic speech recognition in the booth: Assessment of system performance, interpreters' performances and interactions in the context of numbers
- Downie, Jonathan (2019) *Interpreters vs Machines: Can Interpreters Survive in an*

AI-Dominated World? (2019 Routledge)

Fantinuoli, Claudio (ed. 2018), *Interpreting and technology*, 29-59: Language Science Press

Fantinuoli, Claudio (2018), *Interpreting and technology: The upcoming technological turn*, *Interpreting and technology* Edited by Claudio Fantinuoli Translation and Multilingual Natural Language Processing 11, Language Science Press

Meta, “RESEARCH Meta’s new AI-powered speech translation system for Hokkien pioneers a new approach for an unwritten language” 2022 <https://ai.facebook.com/blog/ai-translation-hokkien/> (2023/1/24 21:50)

Okoniewska, Alicja M. (2022), *Interpreters’ roles in a changing environment*, The Translator

Prandi, Bianca (2018), *An exploratory study on CAI tools in simultaneous interpreting: Theoretical framework and stimulus validation*

World Economic Forum “The Future of Jobs Report 2020”

Yang, Junyi et al (2022), *JHU IWSLT 2022 Dialect Speech Translation System Description*