

多言語翻訳技術の高度化に関する研究開発

基本計画書

1. 目的

総務省では、世界の「言葉の壁」の解消に向けて、2014年に「グローバルコミュニケーション計画」を策定し、東京2020オリンピック・パラリンピック競技大会も見据えながら、国立研究開発法人情報通信研究機構（以下「NICT」という。）の多言語自動翻訳技術（逐次翻訳技術）の研究開発及び社会実装を推進してきた。これまでの取組により、NICTの多言語自動翻訳技術は、AI技術、特に深層学習を導入して短文の逐次翻訳を12言語¹で実用レベルの精度を実現し、NICTからの技術移転も進め、民間企業による多様な翻訳サービスの実用化・普及につながっている。

近年、訪日外国人及び在留外国人は共に増加傾向にあり、2018年の訪日外国人は3000万人を超え、在留外国人は住民の50人に1人の割合に達している²。特に在留外国人については今般の入管法改正（2019年4月施行）による特定技能の在留資格の創設も背景として今後も増加・多国籍化が見込まれる。地方も含む日本全体で、外国人との交流の機会が今後ますます増えていき、日常生活の様々な分野（行政手続・医療・交通・観光等）や仕事などの場面で「言葉の壁」と向き合わなければならない状況が生じることが見込まれる。

また、日本で開催される国際会議の件数は年々増加傾向にあり、企業活動ではグローバル化の進展や人材確保に対応するため、海外拠点の拡大や外国企業との連携、外国人材の雇用を進めている。ビジネスや国際会議での議論の場面も含め、多言語での会議対応や十分なコミュニケーションが可能な「通訳」へのニーズが高まっている。

さらに、2025年には日本国際博覧会（大阪・関西万博）を控えており、自動翻訳技術を飛躍的に発展させ、パビリオンや講演会場での利活用を念頭に世界中の人々とリアルタイム会話ができる多言語通訳システムの実現が期待されている。

こうした課題に対応するため、本研究開発では、NICTの多言語自動翻訳技術の更なる高度化により、従来の短文の逐次翻訳から、会話の文脈（文章の流れ）や話者の意図も補完した「同時通訳」へと飛躍させるとともに、開発成果を通訳サービスの提供又は利活用の主体となる企業等が共通的に利用可能となる仕組みを構築し、社会実装することを目指す。

¹当初目標としていた訪日外国人対応の10言語（日本語、英語、中国語、韓国語、タイ語、インドネシア語、ベトナム語、ミャンマー語、フランス語、スペイン語）に加え、在留外国人対応の2言語（ブラジルポルトガル語、フィリピン語）も含めた12言語を実用レベルの精度で実現。

²訪日外国人数（年間）：約3,119万人（2018年）、在留外国人数：約273万人（2018年12月時点）。

2. 政策的位置付け

「グローバルコミュニケーション開発推進協議会 第6回総会」(2019年7月)において、総務大臣から、大阪・関西万博の誘致成功も踏まえ、「2025年度までに同時通訳の実現」などNICTの多言語自動翻訳技術の更なる高度化に向けた「グローバルコミュニケーション計画」の次期計画を検討・策定する考えが表明された。これを踏まえ、同協議会において現在、本研究開発の方向性を含む「次期グローバルコミュニケーション計画(仮称)」に向けた検討が進められており、今後、総務省において同計画を策定・公表する予定である。

「統合イノベーション戦略2019」(2019年6月閣議決定)及び「AI戦略2019」(2019年6月統合イノベーション戦略推進会議決定)において、2025年度までに、「周囲の状況や文化的背景も考慮し、話者の意図を補足しながら利用可能(議論レベル)な翻訳技術を段階的に実現」、「高精度と遅延の最小化を両立する実用レベルの同時通訳の実現」とされている。

「経済財政運営と改革の基本方針2019」(2019年6月閣議決定)において、「2025年大阪・関西万博の開催等に向け、深層学習(ディープラーニング)による同時通訳を含む自動翻訳システムの開発・普及など、新しい日本の創造に関する取組を地方自治体や民間企業と連携しながら進め、レガシーを創出し、将来の開催国等に示していく」とされている。

「外国人材の受入れ・共生のための総合的対応策」(2019年12月外国人材の受入れ・共生に関する関係閣僚会議決定)において、「多言語自動音声翻訳技術については、2025年大阪・関西万博も見据え、日常生活・行政手続・観光等の場面に加え、ビジネスや国際会議等での議論の場面も含め、日本人と外国人及び外国人同士でストレスなく十分なコミュニケーションを可能とするため、AIによる同時通訳の実現に取り組むとともに、今般の入管法の改正も踏まえ、特定技能外国人を含め、在留外国人に対応する観点から強化対象言語を追加し、併せて翻訳精度の向上を図る」とされている。

3. 目 標

(1) 政策目標 (アウトカム目標)

現在のNICTの多言語自動翻訳技術は、発話開始から発話終了までを一区切りとした文章(短文のまとまり)をそれぞれ単独に翻訳することが可能であり、1対1でFACE-TO-FACEの短い対話の場面で有効に利用されている。他方、現在の技術では、発話した内容が翻訳されて相手に伝わるまでの時間が元の発話時間の倍以上の時間を要し、連続した文章(長文)を短文のまとまりごとの区切りを検知しながら随時翻訳することや、会話における文章の流れ(前後の文脈)や状況を意味的に理解し、補完した上で翻訳する(例えば省略の多い日本語を外国語に翻訳する際に文章の流れを考慮して省略を補う)ことまでは対応できない。このため、発話に連動し同時並行で訳出し続けていく、いわゆる「同時通訳」については実現していない状況にある。

社会経済活動のグローバル化の進展等を背景として、ビジネスや国際会議での議論・交渉等の場面、国際的な各種イベント(セミナー・シンポジウム・展示会)での説明・応答等の場面など、1対NやN対Nの状況も含め、多言語による発話と連動したリアルタイムな意思疎通や相互理解を可能とする「通訳」へのニーズが高まっている。また、日本国内では、緊急時の対応も含め、日常生活の様々な分野で外国人とのコミュニケーション手段の確保が必要となる場面が一層増えていくことが見込まれる。そのような中、人の通訳者の確保・養成には大きなコストや時間を要し、言語によっては極めて困難な状況にある。

こうした課題を踏まえ、本研究開発により、自動同時通訳を実現し、低遅延かつ最適な精度の実用レベルの自動同時通訳システムを実社会に実装するための技術を確立する。さらに、開発成果を通訳サービスの提供又は利活用の主体となる企業等が共通的に利用可能となるプラットフォームを構築する。

この取組が、国内外の様々な場面での「通訳」ニーズに応え、社会経済活動における「言葉の壁」の解消、企業等の労働生産性の向上や「働き方改革」の推進、利用者利便の一層の増進に寄与する。これによって、外国人との共生社会の実現、我が国のAI自動翻訳技術の国際的な競争力の強化にもつながっていく。

(2) 研究開発目標 (アウトプット目標)

NICTの多言語自動翻訳技術を発展させ、自動同時通訳を実現するための基盤となる技術(入力分割・要約・翻訳出力最適化技術、多様な情報源を活用した通訳精度向上技術、自動通訳性能評価尺度の確立)及び実用レベルの自動同時通訳システムを実社会に実装するための技術(最適なユーザインターフェース技術、入力音源分離技術、プラットフォーム技術)を確立する。その実施に当たっては、通訳ニーズが高い実社会の複数の分野・場面での実証を通じて得られた知見をフィードバックしながら開発を行う。

さらに、これらの開発成果を通訳サービスの提供又は利活用の主体となる企業等が共通的に利用可能となる自動同時通訳プラットフォームを構築する。

4. 研究開発内容

(1) 自動同時通訳基盤技術

① 概要

自動同時通訳を低遅延かつ実用的な精度で実現するためには、翻訳単位(翻訳可能な短い意味的まとまり)ごとに発話に連動し同時並行で訳出し続けていく基本的なメカニズムを構築するとともに、入力発話に対して、現在の逐次翻訳における厳密で直訳的な表現の訳出とは異なり、厳しい時間的制約の中で聞き手が理解しやすい適切な内容・表現・タイミングで訳出されることが必要である。

また、自動同時通訳では、各翻訳単位が前後の会話(文章)等を考慮せずそれぞれ単独に訳されることにより十分な通訳精度を確保できない可能性があることから、会話

(文章)の流れや周囲の状況などの多様な情報源を活用して通訳に反映し、通訳精度を向上させることも必要である。

このため、ア) 入力発話から翻訳単位の分割検出、要点の絞り込み、単語変換を適切に行って訳出する入力分割・要約・翻訳出力最適化技術、イ) 翻訳単位を超えた多様な情報源を活用した通訳精度向上技術を確立するための研究開発を行う。

さらに、これらの技術によって実現する自動同時通訳の性能を的確に評価するためには、従来の評価尺度では対応できず、通訳の利用が想定される実環境で求められる様々な要件を考慮した新たな評価尺度が必要となることから、ウ) 自動通訳評価尺度を確立するための研究開発を行う。

これらア)～ウ)の技術の研究開発の実施に当たっては、自動同時通訳に必須となるAI学習、検証・評価用のデータベースを整備し、これを各技術にフィードバックしながら行うものとする。

② 技術課題

ア) 入力分割・要約・翻訳出力最適化技術

低遅延かつ実用的な精度の自動同時通訳を実現するため、長い入力発話(連続した文章)から瞬時に音声と言語に関わる情報を読み込み、翻訳単位(翻訳可能な短い意味的まとまり)を限定した上でこれを切り出し検出する技術、入力発話情報から内容の薄い部分を削除して要点に絞り込む(要約する)技術、単言語内で聞き手が理解しやすい単語や表現に変換する技術、及びこれらの技術を組み合わせてAI学習により翻訳単位ごとに自動で聞き手にわかりやすい最適な通訳結果を出力し続ける技術の研究開発を行う。

イ) 多様な情報源を活用した通訳精度向上技術

自動通訳において同時性を確保するためには短い翻訳単位ごとに訳出することが必要となるところ、翻訳単位が単語又は1文に満たないような極めて少ない情報の場合に、各翻訳単位が前後の会話(文章)等を考慮せずそれぞれ単独に訳されることにより、通訳結果において十分な精度を確保できない可能性がある。

自動同時通訳の有効性を高めるため、入力発話に係る前後の会話(文章)や会話(文章)全体の流れの情報、話者の発言の傾向や特性を示す情報(話者による講演・著書など)、会話や話者を取り巻くシーンや状況を示す情報(静止画・動画)、直近の多言語のニュースやWEB掲載の情報など、翻訳単位を超えた多様な情報源と紐づけて、AI学習により文脈理解、状況認識、話者の意図などを的確に通訳に反映する技術の研究開発を行う。その実施に当たっては、自動同時通訳の精度向上に最適な情報源を比較検討する。

ウ) 自動通訳性能評価尺度の確立

自動同時通訳の性能を的確に評価するためには、従来の逐次翻訳技術で用いて

きた客観的なデータ (BLEU 値³等) だけではなく、通訳の利用が想定される実環境で求められる様々な要件 (内容の正確さや専門性、使用する語彙や表現の的確さ、発話速度など) を考慮して、利用者の充足度に関する評価尺度が必要となる。

このため、人の通訳とも比較考慮しながら、各要件に対する利用者の充足度を分析・検証し、自動同時通訳の性能を判定・クラス分けするための評価尺度を確立するとともに、各要件に対する自動通訳性能の充足度を半自動的に評価できる技術を確認するための研究開発を行う。

③ 到達目標

ア) 入力分割・要約・翻訳出力最適化技術

自動同時通訳技術と人 (通訳者の場合も含む) の連携及び自動同時通訳技術の単独活用の場合において、「ウ) 自動通訳性能評価尺度の確立」で確立した評価尺度に基づき、実利用可能な有用性を確認すること。

イ) 多様な情報源を活用した通訳精度向上技術

本技術を「ア) 入力分割・要約・翻訳出力最適化技術」と併せて自動同時通訳に適用することにより、「ア) 入力分割・要約・翻訳出力最適化技術」のみを適用した場合よりも、「ウ) 自動通訳性能評価尺度の確立」で確立した評価尺度に基づき、実利用可能な有用性が改善することを確認すること。

ウ) 自動通訳性能評価尺度の確立

通訳形態 (1 対 1、1 対 N、N 対 N など) ごとに求められる各要件に対する利用者の充足度について、人による評価と高い相関を持つような、自動同時通訳性能の評価尺度を確立する。

(2) 自動同時通訳システム技術

① 概要

低遅延かつ実用的な精度の自動同時通訳技術を通訳ニーズが高い実社会の環境に適用してシステム化するとともに、自動同時通訳システムが実社会で広く利用可能となる仕組みを構築し、社会実装できるようにする必要がある。

このため、ア) ビジネス・国際会議での議論・交渉等の場面など 1 対 N 又は N 対 N や片方向又は双方向通訳の状況での自動通訳に最適なユーザインターフェース技術、イ) 自動同時通訳システムの共通基盤となるプラットフォーム技術の研究開発を行う。

その際、通訳ニーズが高い実社会の複数の分野・場面での実証を通じて得られた知見等を各技術にフィードバックしながら研究開発を実施するとともに、(1) 及び (2) の各技術の開発成果を通訳サービスの提供又は利活用の主体となる企業等が共

³ 「参照訳」と呼ばれる正解に近い翻訳と自動翻訳システムによる翻訳の類似性を表す評価尺度。BLEU 値が大きいほど良い翻訳結果 (参照訳に類似している) と判断できる。

通的に利用可能となる自動同時通訳プラットフォームを構築する。

② 技術課題

ア) 自動同時通訳ユーザインタフェース技術

a) システム利活用要件に応じた統合検証技術

ビジネスや国際会議での議論・交渉等の場面、国際的な各種イベント(セミナー・シンポジウム・展示会)での説明・応答等の場面、緊急時の対応も含む日常生活の様々な分野(日常生活・行政手続・医療・交通・観光等)での外国人とのコミュニケーション手段が必要となる場面を想定し、話者・聴者の数が1対N又はN対N、話者から聴者への片方向通訳又は複数話者間での双方向通訳といった状況における自動同時通訳システムの実用シーンを念頭に置いて、それぞれの状況ごとに必要な要件を検証・評価し、自動同時通訳システムの社会実装に必要な最適なユーザインタフェース技術の研究開発を行う。

その実施に当たっては、自動同時通訳システムの利活用シーンに応じた要件(利用者の使い勝手や利便性、システムの機能・性能や実用面での仕様など)の精査が重要となるため、通訳ニーズが高い実社会の複数の分野・場面での実証を行い、実証を通じて得られた知見等をフィードバックしながら繰り返し改善を図っていくアジャイル型の手法により研究開発を実施する。

b) 入力音源分離技術

ビジネスや国際会議での議論・交渉等の場面で、複数の話者が交互に発話するようなケース(例:司会者と講演者の発言)、複数の話者が同時に又は重複して発話したり、交互発話の間隔が短いようなケース(例:パネルディスカッションや電話会議)を想定しながら、入力発話や背景音が混在する状況下でも、それぞれの話者の音声や背景音(話者以外の音声も含む)を分離し、話者の入力マイクへの背景音の混入を抑制することにより、適切な通訳処理を可能とする技術の研究開発を行う。

イ) 自動同時通訳プラットフォーム技術

自動同時通訳システムを実社会で広く利用できるようにするため、自動同時通訳に必須となるAI学習、検証・評価用のデータベースに基づく自動同時通訳プラットフォーム技術の研究開発を行う。

自動同時通訳プラットフォームは、バックエンドとなる自動同時通訳基盤技術を実装した自動同時通訳コアエンジンを内包し、外部APIによって自動同時通訳機能を提供するPaaS基盤として、それを利用する企業等の利活用性を最大化するために、研究開発期間中の各時点における最新のパブリッククラウド(IaaS)基盤技術を最大限活用して構築する。

③ 到達目標

ア) 自動同時通訳ユーザインタフェース技術

a) システム利活用要件に応じた統合検証技術

講演等において話者が伝達しようとした内容の意味が聴者に理解されているか、会議等において複数話者間で意思疎通ができており会議として成立しているかなど、用務達成率を80%以上とすること。

b) 入力音源分離技術

複数の話者に入力マイクが1つずつ割り当てられ、それぞれの入力マイクから独立した音声認識を行う条件において、主となる話者を特定し、適切に音声認識を行うことができるようになること。

イ) 自動同時通訳プラットフォーム技術

自動同時通訳プラットフォームは、多様な情報源を活用した通訳精度向上技術に対応するため、端末等から複数の音声と動画のストリームを取り込み自動同時通訳を補完することによる高性能化に資する有効な前段処理を行って翻訳エンジンに届ける仕組みを実装すること、自動同時通訳機能を提供するPaaS基盤が構築される物理データセンターのうち1箇所が全停止した場合にもサービス提供を継続できることなどの条件を満たし、実運用のための技術移転が可能なレベルで実現すること。

5. 研究開発期間

令和2年度から令和6年度までの 5年間

6. その他 特記事項

(1) 特記事項

① 提案者は、下記課題(1)-ア)、(1)-イ)、(1)-ウ)、(2)-ア)、(2)-イ)のいずれか又は複数の課題に提案することができる。ただし、課題(2)-ア)についてはさらに小さい課題単位(例:課題(2)-ア)-a)の提案も可能とする。

(1) 自動同時通訳基盤技術

- ア) 入力分割・要約・翻訳出力最適化技術
- イ) 多様な情報源を活用した通訳精度向上技術
- ウ) 自動通訳性能評価尺度の確立

(2) 自動同時通訳システム技術

- ア) 自動同時通訳ユーザインタフェース技術
 - a) システム利活用要件に応じた統合検証技術

- b) 入力音源分離技術
- イ) 自動同時通訳プラットフォーム技術

② 課題（２）-ア）-a) の受託者が、本研究開発課題全体のとりまとめを行うことを想定する。提案に当たっては、3.（２）に記載された自動同時通訳プラットフォームの構築が全体の研究開発目標となっていることに留意すること。

（２）提案および研究開発に当たっての留意点

- ① 提案に当たっては、国際的な研究開発の状況を踏まえた上で、基本計画書に記されているアウトプット目標に対する達成度を評価することが可能な具体的な評価項目を設定し、各評価項目に対して可能な限り数値目標を定めること。
- ② 目標を達成するための研究方法、実用的な成果を導出するための共同研究体制又は研究協力体制、及び達成度を客観的に評価するための実験方法について、具体的に提案書に記載すること。
- ③ 複数機関による共同研究を提案する際には、研究開発全体を整合的かつ一体的に行えるよう参加機関の役割分担及び分担する技術間の連携を明確にし、インターフェースを確保するとともに、研究開発期間を通じて継続的に連携するための方法について具体的に提案書に記載すること。
- ④ アウトカム目標の達成に向けた適切な研究成果（アウトプット等）の取扱方策（研究開発課題の分野の特性を踏まえたオープン・クローズ戦略を含む）について提案書に具体的に記載すること。
- ⑤ 本研究開発成果を確実に成果展開し、アウトカム目標を達成するため、事業化目標年度、事業化に至るまでの実効的な取組計画（研究開発成果の事業化に向けた具体的な計画、体制、資金等）についても具体的に提案書に記載すること。
- ⑥ 研究開発の実施に当たっては、関連する要素技術間の調整、成果の取りまとめ方等、研究開発全体の方針について幅広い観点から助言を頂くと共に、実際の研究開発の進め方について適宜指導を頂くため、学識経験者、有識者等を含んだ研究開発運営委員会等を開催する等、外部の学識経験者、有識者等を参画させること。
- ⑦ 本研究開発は総務省施策の一環として取り組むものであることから、総務省が受託者に対して指示する、研究開発に関する情報及び研究開発成果の開示、関係研究開発プロジェクトとのミーティングへの出席、シンポジウム等での研究発表、共同実証実験への参加等に可能な限り応じること。
- ⑧ 1. ～ 3. に記載のとおり、NICTの多言語自動翻訳技術の更なる高度化によりAI

による多言語同時通訳等を実現するための研究開発を実施し、その開発成果の社会実装を推進していくため、提案書の内容作成及び研究開発の実施にあたっては、NICTが開発した多言語翻訳技術との互換性が確保されていること。

(3) 人材の確保・育成への配慮

- ① 研究開発によって十分な成果が創出されるためには、優れた人材の確保が必要である。このため、本研究開発の実施に際し、人事、施設、予算等のあらゆる面で、優れた人材が確保される環境整備に関して具体的に提案書に記載すること。
- ② 若手の人材育成の観点から行う部外研究員受け入れや招へい制度、インターンシップ制度等による人員の活用を推奨する。また、可能な限り本研究開発の概要を学会誌の解説論文で公表するなどの将来の人材育成に向けた啓発活動についても十分に配慮すること。これらの取組予定の有無や計画について提案書に記載すること。

(4) 研究開発成果の情報発信

- ① 本研究開発で確立した技術の普及啓発活動を実施すると共に、その活動計画・方策については具体的に提案書に記載すること。
- ② 研究開発成果については、原則として、総務省としてインターネット等により発信を行うとともに、マスコミを通じた研究開発成果の発表、講演会での発表等により、広く一般国民へ研究開発成果を分かりやすく伝える予定であることから、当該提案書には、研究成果に関する分かりやすい説明資料や図表等の素材、英訳文書等を作成し、研究成果報告書の一部として報告する旨の活動が含まれていること。さらに、総務省が別途指定する成果発表会等の場において研究開発の進捗状況や成果について説明等を行う旨を提案書に記載すること。
- ③ 本研究開発終了後に成果を論文発表、プレス発表、製品化、ウェブサイト掲載等を行う際には「本技術は、総務省の『多言語翻訳技術の高度化に関する研究開発』による委託を受けて実施した研究開発による成果です。」という内容の注記を発表資料等に都度付すこととする旨を提案書に明記すること。