

科学研究費助成事業（基盤研究（S））公表用資料 〔令和5（2023）年度 中間評価用〕

令和5年3月31日現在

研究期間：2021～2025
課題番号：21H05054
研究課題名：多元自動通訳システムと評価法に関する研究とその応用展開
研究代表者氏名（ローマ字）：中村 哲（NAKAMURA Satoshi）
所属研究機関・部局・職：奈良先端科学技術大学院大学・データ駆動型サイエンス創造センター・教授
研究者番号：30263429

研究の概要：

人間の通訳者が行う同時通訳を機械で実現することは格段に高度で困難である。人間の同時通訳の質を目指すには新たなフェーズをもたらす研究と実用場面でのデータ収集および性能改善の継続的エコシステムが必要である。よって本研究では、多元同時通訳方式、通訳品質の評価法とリアルタイム評価技術、コーパス構築とシステムについて研究を実施する。

研究分野：知覚情報処理、知能情報学

キーワード：音声翻訳、漸進的音声認識・合成、漸進的機械翻訳、通訳翻訳評価

1．研究開始当初の背景

コロナ禍により東京オリンピックにおいて4000万人の訪日外国人を受け入れる政府の目論見は達成が難しくなっている。一方で分担者の松下による調査では、コロナ禍以後、遠隔同時通訳(RSI)の普及が急速に進み、通訳がオンラインで1日24時間世界中のどこへでも提供できるようになったこと、また遠隔通訳では非言語情報が得にくい等の理由から通訳者に大きな負担がかかっていることが示されている。テレワークは自動同時通訳、字幕自動翻訳のニーズを飛躍的に高めており、オンライン時代に通訳者の支援、自動通訳を行う技術の開発は喫緊の課題である。代表者は、ATR、NICTで音声翻訳の研究を主導し、奈良先端大では世界に先駆けて英日間同時通訳の研究を開始した。これまでの研究で340時間の同時通訳コーパス、漸進的音声認識・機械翻訳・音声合成の基本方式の研究開発について多くの成果を上げ世界をリードしている。これらのニーズに答えるためには新たなフェーズをもたらす研究と、実用場面のデータ収集および性能改善の継続的エコシステムが必要である。本研究により漸進的音声認識・翻訳・合成の持続的な改良研究を越えた研究に取り組む。

2．研究の目的

本研究の目的は、通訳者の視点から「人間の同時通訳はどのように通訳しているのだろうか？よりよい通訳とはどのようなものだろうか？」、工学研究の視点から「機械でどこまで同時通訳ができるだろうか？通訳者のように文脈、事前知識、外部知識などの多元情報を使い、場面に合わせた通訳ができるだろうか？同時通訳技術は同時通訳者を支援できるだろうか？」、利用者の視点から、「人間にとってわかりやすい同時通訳とはどのようなものだろうか？内容が伝わるとはどのようなことだろうか？」を明らかにし、同時音声翻訳とその評価法の確立を行うべく次節に述べる研究開発を実施する。

3．研究の方法

上記の基本的問いに答える研究とシステムの構築、そして実運用システムの試作とデータ収集・改良のエコシステムの構築を行う。そのための具体的な課題は以下の通りである。

課題1：多元同時通訳方式：パラ言語音声翻訳およびビデオ・事前・外部知識の利用による多元同時通訳、通訳出力最適化、漸進的音声通訳方式高度化。

課題2：通訳品質の評価法とリアルタイム評価技術：通訳プロセス分析、通訳者支援技術、通訳者、自動通訳システム共通の通訳品質の評価法、脳活動を含むセンシングによる通訳品質客観的自動評価法の確立。

課題3：コーパス構築とシステム：通訳時間アライメント・品質アノテーション、コーパス増強、実運用システムの構築とデータ収集・改良のエコシステムの構築とアクティブラーニング、ライフロンクラーニング法の確立。

4. これまでの成果

これまでに収集した 340 時間の日英、英日の同時通訳データについてのアライメント、アノテーション、分析を行った。同時通訳者は、語順の違いなどにより、原発話者の文の区切りを保持せずに五月雨的に訳出するため文の対応関係が曖昧であるため、マニュアルのアライメント、自動アライメントを行った。このアライメントデータを用いて、機械翻訳モデルの学習を行い、より低遅延の通訳出力ができることを確認した。併せて本データを用いて通訳評価基準の開発を共同研究者とともに進めた。パラ言語についても、発話者の音声におけるフォーカスを言語的に言い換えて翻訳する方法を確立した。この研究は音声のパラ言語におけるフォーカスと書き言葉の変換可能性に関する音声言語研究の重要課題である。その他、音響信号処理、声質変換、発話表情変換、画像情報の利用に関する研究についてもおおよそ予定通り進んでいる。2022 年度には、IWSLT の同時通訳トラックに参加しシステムの構築、高度化を進めた。2023 年度はこれらの要素技術をリアルタイムシステムに統合して中間的な評価を進める。

5. 今後の計画

- パラ言語音声翻訳（感情・意図）：2023 年度以降は「強調（フォーカス）」に関する音声と言語表現の最適組み合わせ出力、2024 年度以降は感情の翻訳に取り組む。音声のパラ言語で表現されるフォーカス（意図）とテキスト表現のフォーカス（意図）の間を如何に等価変換し得るかについては音声言語研究における本質であり今後も基礎研究を継続する。
- 音声画像翻訳：2023 年度以降は、発話者の韻律同期もしくは感情表出時の顔動画の個性表現に関して検討を進め、静止画像としての個性の表現のみならず、動画生成時のキーフレーム補間時のアイデンティティ同期の方法について検討を進める。
- 事前情報、外部知識の利用：2021-2022 年度は、大規模コーパスを用いた翻訳モデル、インクリメンタル同時通訳アルゴリズムの高度化に集中したため 2023 年度からの開始としている。2023 年度は、分野やキーワード等の情報を明確に与える形での事前適応、2024 年度以降は講演原稿や関連 Web ページからの適応対象情報の自動抽出に基づく適応、Wikipedia や事前学習型モデルを活用した基盤モデル強化と事前適応の統合を試みる。
- 通訳プロセス分析：2023 年度以降は通訳プロセスのうち「順送り」や「省略」などの分析を深化させる。また進的翻訳技術との連携により応用技術に落とし込み、人間の通訳支援装置の開発も進め、オンライン通訳サービスへの通訳支援機能として実装、ユーザデータを収集した活用度や訳出への影響の分析を行う。
- 通訳者、通訳システムにおける共通の通訳品質の評価法：2023 年度以降は通訳者が重視する観点を考慮した通訳品質評価指標の確立、遅延や速度の面も考慮した総合的な評価の実施、自動評価方式を検討する。
- 通訳時間アライメント・品質アノテーション、コーパス増強：2023 年度以降は自動アライメントによる通訳対訳コーパスの増強と機械翻訳モデルの強化、また通訳品質評価への応用について検討する。
- 多元同時通訳コーパスの構築：同時通訳コーパスのチャンクのアライメントに加え、多元パラ言語アノテーション付きコーパス 50 時間、事前情報 50 時間の構築に着手する。
- 実運用システムの構築とデータ収集・改良のエコシステムの構築：2023 年度はモジュールの統合、評価を行い、エコシステムの設計、実装を進める。

6. これまでの発表論文等（受賞等も含む）

- T. Yanagita, S. Sakti, S. Nakamura, “Japanese Neural Incremental Text-to-Speech Synthesis Framework With an Accent Phrase Input,” IEEE Access, Vol. 11, 2023, pp. 22355-22363
- S. Novitasari, S. Sakti, S. Nakamura, “A Machine Speech Chain Approach for Dynamically Adaptive Lombard TTS in Static and Dynamic Noise Environments,” IEEE/ACM Transactions on Audio, Speech, and Language Processing, Vol. 30, 2022, pp. 2673-2688
- K. Matsushita, “How Remote Interpreting Changed the Japanese Interpreting Industry: Findings from an online survey conducted during the COVID-19 pandemic,” INContext: Studies in Translation and Interculturalism, Vol. 2, No. 2, 2022, pp. 167-185
- R. Fukuda, K. Sudoh, S. Nakamura, “Knowledge Distillation for Translating Erroneous Speech Transcriptions,” Journal of Natural Language Processing, Vol. 29, 2022, pp. 344-366
- H. Inaguma and T. Kawahara, “Alignment knowledge distillation for online streaming attention-based speech recognition,” IEEE/ACM Trans. Audio, Speech & Language Process, Vol.29, 2021

7. ホームページ等

<https://ahcweb02.naist.jp/page-4794/kakenhi-ngst/>