

**科学研究費助成事業 研究成果報告書**

平成 29 年 8 月 1 日現在

機関番号：14603

研究種目：基盤研究(A) (一般)

研究期間：2012～2016

課題番号：24240032

研究課題名(和文) 外国語ニュース・講演の音声同時通訳方式に関する研究

研究課題名(英文) Simultaneous speech translation methods for news and lectures in foreign languages

研究代表者

中村 哲 (Nakamura, Satoshi)

奈良先端科学技術大学院大学・情報科学研究科・教授

研究者番号：30263429

交付決定額(研究期間全体)：(直接経費) 35,800,000円

研究成果の概要(和文)：同時通訳基本方式研究として、フレーズベース統計翻訳における右確率を用いた同時通訳方法により、翻訳単位を短くする方法、翻訳単位の長さを調整する手法を提案した。さらに、形態素情報を使って文を分割する方法、Tree-to-string翻訳での部分構文構造を考慮して分割する方法を提案して高精度化を実現。さらなる精度改善のため、訳文に単語順序の入れ替えが発生するかを予測するモデルを構築すると共に、ニューラル翻訳の実装、統計翻訳のリランキング、注意型ニューラル翻訳の研究を進めた。また、同時通訳コーパスとして日英合計約80時間、講義データの書き起こし約50時間、うち約22時間分の日英翻訳を完了した。

研究成果の概要(英文)：In this project new simultaneous speech-to-speech translation algorithms are proposed. First algorithm has a mechanism to decide to output or hold the phrases to the machine translation module until the current time based on the right probability in the phrase-based statistical machine translation. Second algorithm is able to segment the input phrase sequence based on greedy search according to POS bigram information. Third algorithm predicts next phrase or local parse tree element based on SVM with the incremental bottom-up parser. Here, the algorithm decides to output or hold the phrases again. The experiments showed that the proposed algorithms successfully realized the simultaneous speech translation. Furthermore neural machine translation algorithms with attention mechanisms are investigated. The 80 hours of J-E interpretation data, 50 hours of JP lecture transcription data, and 22 hours of J-E translation data are collected to be used for simultaneous speech translation research.

研究分野：知能コミュニケーション

キーワード：音声情報処理 音声翻訳 音声認識 機械翻訳

### 1. 研究開始当初の背景

日本の社会、企業の国際化に伴い、外国語の情報の取得、多言語会議の必要性が高まり、これらを支援する音声同時通訳技術の研究開発が喫緊の課題であった。研究開始当初の時点で、すでに旅行会話を対象とした音声通訳技術はすでに産業化できるレベルに到達していたが、ニュース・講演や、会議の音声同時通訳には対応できていなかったため、基礎研究の視点から再検討し同時音声通訳システムの構築を目標とした。

### 2. 研究の目的

本課題では、言語の壁を打ち破る、話した文をリアルタイムで翻訳する音声同時通訳基本方式の研究に取り組んだ。旅行会話を対象に一文毎に発話後に翻訳する音声通訳は実用化が進んでいるが、ニュース・講演や、会議の音声同時通訳は、一文単位の統計的翻訳の研究の枠組みでは、翻訳単位、順序、五月雨処理など同時性達成にあたっての様々な問題のため困難でほとんど未着手であった。本研究では、英語・日本語間で中級程度の通訳者と同等の音声同時通訳性能を実現し、外国語ニュース・講演の音声同時通訳の基礎研究、プロトタイプシステムの構築を行うことを目標とした。

### 3. 研究の方法

本課題では、全期間における基盤研究に加えて、H24～H26に音声同時通訳対訳コーパスの構築、H27～H28に音声同時通訳システムの構築と実験を計画した。具体的には、H26～H28に英日の講演のコーパスを合計60時間収集し、初、中、上級の同時通訳者による通訳と、元音声、通訳音声の書き起こしとタグ付けを行った。このコーパスに基づき、同時通訳の基礎研究を行い、外国語ニュース・講演の音声同時通訳の基本方式研究、プロトタイプシステムの構築を行い世界初の英日・日英音声同時通訳システムの実現を目指した。

### 4. 研究成果

#### (1) 同時通訳基本方式研究：

初年度(H24)は、同時通訳を実現するため、音声通訳システムの遅延の原因として考えられる機械翻訳モジュールを同時処理化する手法の基礎検討を行った。具体的にはフレーズベース機械翻訳のフレーズテーブルを用いて、翻訳単位を短くする方法を提案した。まず、原言語のフレーズパターンを利用することで、翻訳単位を決定する。さらに、翻訳単位が短くなりすぎて翻訳精度が劣化することを防ぐために、両言語の語順が同等である確率(right 確率)を利用して、翻訳単位の長さの調整を行う新しい手法を提案した。日英双方向の同時通訳に並行して、英仏の同時通訳実験も行い、同じ文構造(SVO)の言語間の場合に、本研究で提案した右確率による方法を用いれば、短く区切りながら同時通訳

を行っても比較的性能劣化がないことも明らかとなった。

この同時通訳用機械翻訳モジュールを、ポーズ区切りで認識結果を出力する音声認識と音声合成と統合し、音声同時通訳プロトタイプを構築した。

平成25年度は、五月雨に翻訳結果を出力する新たな方法として、フレーズベース統計翻訳における右確率を用いた同時通訳方法を提案した。右確率は原言語と対象言語で語順の入れ替えが起こりやすいかどうかを判別する確率であり、日本語と英語のように語順の異なる言語には極めて効果的である。この手法を、英語講演TEDに対して適用し評価を行ったところ、同時通訳経験1年のプロの通訳とほぼ同等の性能が得られることが明らかになった。

平成26年度は、五月雨に翻訳結果を出力する方法をさらに改良した。方法として、フレーズベース統計翻訳における文分割を、形態素情報を使って分割する方法、さらにTree-to-string 翻訳での部分構文構造を考慮して分割する方法を提案し、高精度化を実現した。この手法を、英語講演TEDに対して適用し評価を行ない、右確率を利用した方法に比べて大幅な改善を確認した。

平成27年度は、品詞情報をもとに、訳文に単語順序の入れ替えが発生するかを予測するモデルを構築し、同時通訳精度の改善を試みた。また、ニューラル翻訳を実装し、統計翻訳のリランキングに用いることで性能改善できることを示した。

平成28年度は、音声認識にボトムアップパーズを適用し、次に来る構文要素を機械学習により予測し、それにより訳出を行うかどうかを判定する方式の改良を行った。また、注意型ニューラル翻訳の研究を進めた。

#### (2) コミュニケーション評価：

初年度(H24)は、提案法における翻訳の精度と同時性への影響を調査するために、旅行対話のデータを利用し翻訳単位の長さと同訳精度の関係を検証し、フレーズテーブルとright 確率を用いた翻訳単位の決定方法により、翻訳開始および翻訳処理の時間を減少できることを確認した。

平成25年度には、音声同時通訳システムの評価法として、発話チャンクを検出し訳出を行う時間と、翻訳性能のカーブから性能評価を行う手法を提案した。

平成26年度は、発話チャンクを検出し訳出を行う時間と翻訳性能から、同時通訳の優劣を人手による評価をもとに評価する手法を提案し、TED講演に対し有効性を確認した。

平成27年度は、音声認識における置換、挿入、脱落誤りに対する発話者の反応測定を行った。誤り単語の品詞や、役割によって認知負荷が異なることが明らかとなった。

平成28年度は、音声認識における置換、挿入、脱落誤りに対する発話者の反応測定を行

ない、認知負荷の測定、改良を行った。

(3) 同時通訳コーパス構築、プロトタイプ構築：

初年度（H24）は、日本語と英語を原言語とした同時通訳コーパスの収集とアノテーションを行った。原言語側のコーパスとして、日本語にはNHKニュースとCSJ（Continuous Speech of Japanese: 日本語の学術講演音声）を対象に、英語にはTED講演を対象に同時通訳を行った。収集時間は日本語が合計16.5時間、英語が23時間である。同時通訳者として、熟練度合いから3レベル（上級(S)、中級(A)、初級(B)）の通訳者に同一の講演の同時通訳を依頼した。この同時通訳音声および原言語音声に対し、発話内容だけでなく、通訳開始時刻などの時間情報もアノテーションを行い、日本語1.6時間、英語16.5時間のアノテーションが完了した。この同時通訳コーパスを用いて、機械翻訳システムのチューニングを行い、同時通訳システムの性能向上を行うことができた。現在の同時通訳システムが、経験年数1年のプロの通訳者と同等の性能を達成した。

平成25年度は、英語のTED講演のうち、34ないし46講演に対し、経験年数15年、4年、1年のプロの同時通訳者に通訳をしてもらい、同時通訳コーパス20時間の収録を終了すると共に、これらのコーパスの書き起こし、時間的なアライメントを行った。

平成26年度は、約15分の講演・ニュースを英日24本、日英60本、合計21時間の同時通訳収録を行った。通訳の経験年数は平成25年度の収録と同じである。そのうち、英日の日本語24本の6時間分の書き起こしを行った。また、講義アーカイブ90分×35本（計3,150分=52.5時間）の日本語書き起こしを行った。

平成27年度は、同時通訳（日英）9時間分（Aクラス18データ、Bクラス24データ）同時通訳（英日）7時間分（30データ）の書き起こし、講義10コマ分の日英翻訳を行った。

平成28年度は、講義5コマ分の日英翻訳を行った。

#### 5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計3件）

Quoc Truong Do, Tomoki Toda, Graham Neubig, Sakriani Sakti and Satoshi Nakamura, "Preserving Word-level Emphasis in Speech-to-speech Translation," *Transactions on Audio, Speech and Language Processing*, vol.25, no.3, 2017, 査読有  
DOI: 10.1109/TASLP.2016.2643280  
三浦明波, Graham Neubig, Sakriani

Sakti, 戸田智基, 中村哲, "中間言語情報を記憶するピボット翻訳手法," *自然言語処理*, vol.23, no.5, pp499-528, 2016.12, 査読有

DOI: 10.5715/jnlp.23.499

Keigo Kubo, Sakti Sakriani Watiasri, Graham Neubig, Tomoki Toda and Satoshi Nakamura, "Structured Adaptive Regularization of Weight Vectors for a Robust Grapheme-to-Phoneme Conversion Model," *IEICE Transactions on Information and Systems*, vol.E97-D, no.6, pp1468-1476, Jun. 2014, 査読有  
DOI: 10.1587/transinf.E97.D.1468

〔学会発表〕（計6件）

中村哲, "音声研究と自然言語研究の融合に向けて-音声翻訳研究の過去と未来-" *信学技報*, Vol. 116, No. 378, pp. 121-122, Dec. 2016, NTT 武蔵野研究開発センタ（東京都・武蔵野市）（第18回音声言語シンポジウム）（招待講演）

Quoc Truong Do, Sakriani Sakti, Graham Neubig, Satoshi Nakamura. "Transferring Emphasis in Speech Translation Using Hard-Attentional Neural Network Models." *InterSpeech*, pp. 2533 - 2537, Sep. 2016, San Diego (米国)

Yusuke Oda, Graham Neubig, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura. "Syntax-based Simultaneous Translation through Prediction of Unseen Syntactic Constituents" *The 53rd Annual Meeting of the Association for Computational Linguistics (ACL) Long Paper Track*. pp. 198-207. July 2015, 北京（中国）

Yusuke Oda, Graham Neubig, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura. "Optimizing Segmentation Strategies for Simultaneous Speech Translation" *The 52nd Annual Meeting of the Association for Computational Linguistics (ACL) Short Paper Track*. pp. 551-556. June 2014, Baltimore (米国)

Shinnosuke Takamichi, Tomoki Toda, Yoshinori Shiga, Sakriani Sakti, Graham Neubig and Satoshi Nakamura. "Improvements to HMM-Based Speech Synthesis Based on Parameter Generation with Rich Context Models" *Proc. of INTERSPEECH*, pp.364-368, Aug. 2013, Lyon (フランス)

Tomoki Fujita, Graham Neubig, Sakriani Sakti, Tomoki Toda, Satoshi Nakamura. "Simple, Lexicalized Choice of Translation Timing for Simultaneous Speech Translation" *Proc. of*

INTERSPEECH, pp.3487-3491, Aug. 2013,  
Lyon (フランス)

〔図書〕(計0件)

〔産業財産権〕

出願状況(計1件)

名称: 構文評価装置、翻訳装置、構文評価方法及び構文評価プログラム

発明者: 小田悠介、Graham Neubig、中村哲他

権利者: 同上

種類: 特許

番号: PCT/JP2016/055186

出願年月日: 2016年9月9日

国内外の別: 国内

取得状況(計1件)

名称: テキストデータ分割装置、テキストデータ分割方法、テキストデータ分割プログラム及び翻訳装置

発明者: 藤田朋希、Graham Neubig、中村哲他

権利者: 同上

種類: 特許

番号: P150011337

取得年月日: 2017年2月17日

国内外の別: 国内

〔その他〕

ホームページ等

## 6. 研究組織

### (1) 研究代表者

中村 哲 (NAKAMURA, Satoshi)

奈良先端科学技術大学院大学・情報科学研究科・教授

研究者番号: 30263429

### (2) 研究分担者

松本 裕治 (MATSUMOTO, Yuji)

奈良先端科学技術大学院大学・情報科学研究科・教授

研究者番号: 10211575

サクリアニ サクティ (SAKTI, Sakriani)

奈良先端科学技術大学院大学・情報科学研究科・助教

研究者番号: 30625083

高道 慎之介 (TAKAMICHI, Shinnosuke)

東京大学・大学院情報理工学系研究科・特任助教

研究者番号: 90784330

Neubig Graham (NEUBIG, Graham)

奈良先端科学技術大学院大学・情報科学研究科・助教 (H24~H28)

研究者番号: 70633428

戸田 智基 (TODA, Tomoki)

奈良先端科学技術大学院大学・情報科学研究科・准教授 (H24-H27)

研究者番号: 90403328

Duh Kevin (DUH, Kevin)

奈良先端科学技術大学院大学・情報科学研究科・助教 (H24~H27)

研究者番号: 80637322

小町 守 (KOMACHI, Mamoru)

奈良先端科学技術大学院大学・情報科学研究科・助教 (H24)

研究者番号: 60581329

### (3) 連携研究者

( )

研究者番号:

### (4) 研究協力者

( )