

令和 6 年 6 月 27 日現在

機関番号：44101

研究種目：基盤研究(C)（一般）

研究期間：2017～2023

課題番号：17K02765

研究課題名（和文）同時通訳の訳出方略の分析のための柔軟な対訳対応付け手法の開発

研究課題名（英文）Development of a flexible alignment method for the analysis of translation strategies in simultaneous interpretation

研究代表者

笠 浩一郎（Ryu, Koichiro）

三重短期大学・その他部局等・准教授

研究者番号：40397451

交付決定額（研究期間全体）：（直接経費） 1,700,000円

研究成果の概要（和文）：同時通訳の訳出方略を分析するためには、同時通訳者の通訳データに対して様々なアノテーションを行う必要がある。特に、対訳対応付けは、必要不可欠なアノテーションの情報と言っても過言ではない。しかし、同時通訳は、通常の通訳とは大きく異なる状況で通訳されたデータであるため、意識や省略が多く、対訳対応付けは簡単ではない。本研究では、対応なしの対応を許容するなどの工夫を行った対訳対応データを作成し、同時通訳者の訳出方略の分析を行い、新たな知見が得られた。

研究成果の学術的意義や社会的意義

同時通訳者がどのようにして同時通訳という高度な通訳を実現しているかについては、解明されていないことが多い。そのため、同時通訳者の方略を解明する研究を進めていく必要があるが、その際に大きな課題になるのが、原文と訳文の対応付けである。本研究では、同時通訳者の通訳データに対して単語単位での対訳対応付けを手で行い、その対応データをもとに、同時通訳者の訳出方略の分析を行うことで、新たな知見が得られた。

研究成果の概要（英文）：In order to analyze the translation strategy of simultaneous interpretation, it is necessary to perform various annotations on the interpretation data of the simultaneous interpreter. In particular, it is no exaggeration to say that the mapping of bilingual translations is indispensable annotation information. However, since simultaneous interpretation is data interpreted in a situation that is very different from normal interpretation, there are many arbitrary translations and omissions, and it is not easy to correspond to parallel translations. In this study, we created bilingual correspondence data that allowed correspondence without correspondence, and analyzed the translation strategy of simultaneous interpreters, and obtained new knowledge.

研究分野：自然言語処理

キーワード：機械翻訳

1. 研究開始当初の背景

現在、機械翻訳への応用、及び、認知学的知見の獲得などを目的として、同時通訳の訳出方略の科学的分析が盛んに行われている。本研究者は、これまでに大規模な同時通訳データベースの開発に携わり、単語単位で時間情報を付与して通訳者の発話速度の変化を分析するなど、多くの研究成果を挙げている。

一方、同時通訳の訳出方略の分析において、同時通訳の対訳データに対して、単語レベルでの対訳対応を付与することは非常に重要であり、対訳対応をもとに得られた知見は少なくない。例えば、同時通訳ではどの程度遅れながら訳出しているかを分析する場合には、名詞、及び、動詞の対訳対応をもとに遅れ時間を計算している。また、同時通訳者の脳の中の翻訳メモリで記憶されている原言語を図る場合にも、単語の対訳対応から推定して算出試みがされている。しかし、これらの研究では、対訳対応を自動で簡単に取れるものに絞る、または、人手により分析に必要な中規模のデータに付与するなどの工夫をして実施している。

2. 研究の目的

高度な言語処理を行う同時通訳が、どのような訳出方略によって行われているかを解明する上で、単語レベルの対訳対応付けは重要であるが、簡単に対応を付けられない問題が存在する。本研究では、タグ付き対訳対応付けと専用閲覧ツールを開発することで、同時通訳の対訳対応付けで問題となる「文対応が明確でない」、「同時通訳に特有な言語現象が頻出する」、「分析目的によって必要な対訳対応が異なる」という問題を解決する。

さらに、タグ付き対訳対応付けを自動化する手法を検討する。自動化に向けて、十分な学習データが必要であるため、同時通訳データに対して人手により対訳対応付けを実施する。

また、作成した対訳対応付けデータを用いて、同時通訳の訳出方略の分析において、従来に比べより詳細な分析が行えるため、新たな知見の獲得、及び、従来研究で得られた知見の再検証を行う。

3. 研究の方法

同時通訳データの対訳対応付けには、通常対訳データの対訳対応付けにはない下記の4つの特長があるため、一般的な対訳対応付け手法を適用したのでは上手く対訳対応付けできない。

特殊性1：文対応が明確でない

新聞、及び、特許文書などの書き言葉での単語レベルの対訳対応を考える場合には、あらかじめ文対応が定まっておらず、文単位内で単語単位の対訳対応を付与すればよい。一方、同時通訳では、原発話の1文に必ずしも対応する1文で翻訳するわけでない。

特殊性2：同時通訳に特有な言語現象が頻出する

同時通訳には、意識、省略、繰り返し(同じ訳を再度訳出する)、言い誤りなどの言語現象が生起する。意識は、複数の原言語を一つの訳語に意識するため多対1対応が発生する。省略は、対応が取れない1対0対応が発生する、繰り返しは、1対多対応が発生する。言い誤りは、訳語が誤っているため、本来は対訳対応ではないため、どう扱うべきか明確にするのが困難である。

特殊性3：時間情報を手掛かりにできる

同時通訳データベースには、その後の研究で発話単位、及び、単語単位での時間情報が付与されており、一般的な対訳データのアライメントとは異なり、時間情報を手掛かりに対訳候補を絞ることができる。例えば、原単語が発話される前に訳語が発話されないなどの制約を設けるなど。

特殊性4：分析の目的によって必要な対訳対応が異なる

同時通訳の訳出方略の分析では、分析の目的の違いにより付与すべき対訳対応が異なることがある。例えば、ある単語が繰り返し訳出された場合、訳出遅延時間を測るときには最初に訳出された単語を対応として取ることが望ましいが、同時通訳者の記憶を探りたい場合には最後に訳出された単語を対応として取ることが望ましい。

上記で述べた同時通訳の対訳対応の特殊性を考慮して、新たな対訳対応付け方法の提案、及び、自動的に対訳対応を獲得する手法の開発に向けて、同時通訳データに対訳対応付けを実施する。獲得した対訳対応を用いて同時通訳の訳出方略の再検証を行い、新たな知見の獲得を試みる。新たな対訳対応付け方法では、単語レベルの対訳対応を付与し、さらに対訳対応に意識、省略、繰り返し、言い誤りなどの言語現象情報をタグとして付与する。また、同時通訳の訳出方略の分析で対訳対応を利用する場合には、専用の閲覧ツールを開発することで、同時通訳者の方略の分析に活用する。

4. 研究成果

(1) 同時通訳者が作業記憶を用いて通訳していることは広く共有された認識である。そのため、

通訳者の作業記憶について分析することは、同時通訳者の通訳プロセスを解明する上で重要なテーマであると言える。

一方、同時通訳は、原発話の内容を保持する以外にも、聴取、理解、翻訳、訳出などの複数の作業を同時に遂行しなければならないため、通訳者は各作業への注意の配分を意識的、もしくは、無意識的に行っていると考えられる。また、一般に、ある作業を重視し、それに対する注意の配分が増えた場合、他方が疎かになることは容易に想像できる。例えば、二重課題法の結果からも、情報を保持することと情報処理とが互いに干渉することを示されている。

そこで本研究では、同時通訳者の作業記憶を占有する情報の増減が、通訳者の訳出にどのような影響を与えるかについて定量的に分析した。本分析では、作業記憶を占有する情報が増えたときに影響を受けると予想される対象として、通訳者の発話速度のような通訳者自身の発話の様相から観察されるものと、原発話と通訳者の訳出との関係から観察されるものに焦点をあてて分析した。

分析では通訳者の訳出から観察できる現象として、同時通訳者の「発話率」と「発話速度」について調査し、以下のことを確認した。

- ・作業記憶の占有量が増えると、通訳者の発話速度はほとんど変化しないものの、発話する割合が増える。

さらに、話者の発話と通訳者の訳出との関係から観察できる現象として、訳出遅延時間と訳出率について調査し、以下のことを確認した。

- ・作業記憶の占有量が1以上の場合、作業記憶の占有量が増えれば、それに比例して訳出遅延時間が増える。作業記憶の占有量が5の場合は、占有量が1の場合に比べ、訳出遅延時間が1.2秒長くなる。

- ・作業記憶の占有量が増えると、それに比例して訳出率が低下する。

以上の結果から、同時通訳では、通訳者の作業記憶の占有量が増えると、同時通訳者の訳出の発話率、訳出遅延時間、及び、訳出率に影響を与えることが確認された。

(2) 同時通訳における訳出困難な場面を自動的に検出できれば、発話の内容や候補の訳語などを提示することにより、通訳者の情報処理及び記憶にかかる負担を軽減させる支援環境の実現が期待できる。

同時通訳において通訳者の訳出に困難をもたらす場面として、問題を触発するポイントについて議論されている。その中で、通訳者の情報処理に負担をかける要因として、高密度の原発話、未知の固有名詞、数字、構文構造などが挙げられ、このような場面では、訳出の失敗を招く可能性が高いと指摘されている。訳出の失敗には、誤訳、欠落などが含まれる。本研究では、訳出困難な箇所を特定する手がかりとして、訳出の失敗の現象の一つである語の欠落に注目した。

同時通訳における「原発話の特徴」あるいは「訳出状況の特徴」と「語の欠落の発生」との関係性を明らかにするために、講演者話速、訳出遅延時間、品詞、構文構造上の深さという4つの要因を取り上げ、定量的分析を行った。分析には、同時通訳データベースに収録された88講演分の英日同時通訳データを対象とした。分析の結果、以下のことを確認した。

- ・講演者話速が速いとき、語の欠落率が高くなる。
- ・訳出遅延時間が長いとき、欠落率が高くなる。
- ・副詞の欠落率は59.9%であり、平均欠落率33.1%より26.8%高く、他の品詞の欠落率との差が大きい。名詞の欠落率は25.2%、数詞の欠落率は29.1%であり、平均欠落率よりそれぞれ7.9%、4.0%低い。
- ・構文構造上の深さが深い単語は、浅い単語と比べ、欠落率が高い。

(3) 同時通訳における順送り方略の適用に影響する言語的特徴を明らかにするために、原文の書き起こしからテキストへの翻訳との訳出順の比較に基づき、同時通訳の訳出順に関する定量的な分析を行った。

分析では、依存関係にあるチャンク対を対象とした。分析の結果、同一原文に対する異なる訳者による同様の作業条件での訳(翻訳間、および、同時通訳間)は、約88%の割合で訳出順が一致していることがわかった。翻訳と同時通訳の訳出順を比較した場合、約84%の割合で訳出順が一致しており、作業条件の違いにより一致率が異なった。また、翻訳間で訳出順が一致しない場合は、一致する場合より、同時通訳で順送りを適用する割合が高いことがわかった。

同時通訳の訳出順の選択に影響する言語的特徴として、依存関係、後置修飾する句・節の長さ、および、チャンク間の距離、三つの要因を調べた。調査の結果、以下を確認した。

- チャンク間の依存関係が同時通訳の訳出順の選択に影響する。特に、関係節、名詞修飾節に対して順送りを適用する割合が高い。そのほかに、in, of, forなどの前置詞による名詞修飾、副詞節、補足節に対する順送りの適用率も高く、関係節・名詞修飾節と合わせて、同時通訳で順送り方略を必要とする構文であるといえる。
- 後置修飾する句・節が長いほど、順送りを適用する傾向がある。
- 依存関係にある2チャンク間の距離が長いほど、順送りを適用する傾向がある。

5. 主な発表論文等

〔雑誌論文〕 計5件（うち査読付論文 4件 / うち国際共著 1件 / うちオープンアクセス 0件）

1. 著者名 蔡 仲熙、笠 浩一朗、松原 茂樹	4. 巻 20
2. 論文標題 同時通訳における順送り方略の適用に影響する起点言語の特徴	5. 発行年 2020年
3. 雑誌名 通訳翻訳研究	6. 最初と最後の頁 25 ~ 40
掲載論文のDOI (デジタルオブジェクト識別子) 10.50837/its.2002	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Zhongxi Cai, Koichiro Ryu, and Shigeki Matsubara	4. 巻 1
2. 論文標題 Statistical Analysis of Missing Translation in Simultaneous Interpretation Using A Large-scale Bilingual Speech Corpus	5. 発行年 2018年
3. 雑誌名 Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018)	6. 最初と最後の頁 4282-4288
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する
1. 著者名 笠浩一朗, 松原茂樹	4. 巻 18
2. 論文標題 同時通訳における作業記憶の負荷と訳出との関係	5. 発行年 2019年
3. 雑誌名 通訳翻訳研究	6. 最初と最後の頁 147-158
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 蔡仲熙, 笠浩一朗, 松原茂樹	4. 巻 18
2. 論文標題 同時通訳における語の欠落に影響を及ぼす要因の分析	5. 発行年 2019年
3. 雑誌名 通訳翻訳研究	6. 最初と最後の頁 133-146
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 蔡仲熙, 笠浩一朗, 松原茂樹	4. 巻 118
2. 論文標題 講演の同時通訳における文の訳出難易度の推定	5. 発行年 2019年
3. 雑誌名 信学技報	6. 最初と最後の頁 37-41
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計5件 (うち招待講演 0件 / うち国際学会 1件)

1. 発表者名 Zhongxi Cai, Koichiro Ryu, Shigeki Matsubara
2. 発表標題 What Affects the Word of Target Language in Simultaneous Interpretation
3. 学会等名 International Conference on Asian Language Processing (国際学会)
4. 発表年 2020年

1. 発表者名 蔡 仲熙, 笠 浩一朗, 松原茂樹
2. 発表標題 翻訳文との比較に基づく同時通訳の訳出方略に関する分析
3. 学会等名 日本通訳翻訳学会第20回年次大会
4. 発表年 2019年

1. 発表者名 翻訳文との比較に基づく同時通訳の訳出方略に関する分析
2. 発表標題 翻訳との比較による同時通訳における順送り方略の活用の分析
3. 学会等名 電子情報通信学会思考と言語研究会
4. 発表年 2020年

1. 発表者名 蔡仲熙, 笠浩一朗, 松原茂樹
2. 発表標題 同時通訳における訳出遅延の短縮に有用な訳出方略の獲得
3. 学会等名 日本通訳翻訳学会第19回年次大会
4. 発表年 2018年

1. 発表者名 蔡仲熙, 笠浩一朗, 松原茂樹
2. 発表標題 同時通訳における語の欠落に関する定量的分析
3. 学会等名 日本通訳翻訳学会第18回年次大会
4. 発表年 2017年

〔図書〕 計0件

〔産業財産権〕

〔その他〕

-

6. 研究組織

氏名 (ローマ字氏名) (研究者番号)	所属研究機関・部局・職 (機関番号)	備考

7. 科研費を使用して開催した国際研究集会

〔国際研究集会〕 計0件

8. 本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関