

同時通訳者の話速変動の定量的分析

于 海貝†

笠 浩一朗‡

松原 茂樹‡

† 名古屋大学大学院情報科学研究科

‡ 名古屋大学情報連携基盤センター

1 はじめに

同時通訳は、人間の極めて高度な言語処理活動であり、その訳出プロセスを分析することは、認知科学、言語学、通訳学など、幅広い分野において、重要な意義がある。小野らは、同時通訳者の訳出遅延について詳細な分析を与えているが [1]、発話速度に関する知見についてはこれまでほとんど分析されていない。そこで本稿では、名古屋大学同時通訳データベース [2] を用いて、同時通訳者の話速変動を定量的に分析する。原発話の進行状況や講演者の発話状態が同時通訳者の話速に与える影響を分析するために、同時通訳データに対して単語単位での時間情報を付与した。

2 同時通訳者の話速変動

同時通訳者の発話速度は、通訳対象となる原発話が発話中と発話後とでは、違いがあると考えられる。例えば、原発話が確定した発話後の方が発話速度が速くなると予測できる。

また、講演者の発話状態が、発話速度が速い場合と遅い場合、あるいは、発話していない場合とでは、違いがあると考えられる。例えば、発話していない状態では、通訳者は訳出に専念できるので発話速度が速くなると予測できる。

そこで本研究では、原発話の進行状況、及び、講演者の発話状態と通訳者の話速変動との関係性を分析する。

3 分析に用いたデータ

同時通訳者の話速変動を大規模なデータを用いて分析するために、名古屋大学同時通訳データベースを使用した [2]。本分析では、原発話の進行状況や講演者の発話状態による話速変動を分析するために、講演者発話と通訳者の発話に対して単語単位で発話の開始と終了時刻を付与した。

3.1 名古屋大学同時通訳データベース

名古屋大学同時通訳データベースでは、講演の同時通訳音声データが収録されている。音声は、200 ミリ秒以上のポーズで単位分割され、各発話単位には、発話 ID、開始・終了時間、対訳対応情報が付与されている。

Corpus-based Analysis of Simultaneous Interpreter's Utterance Speed

† Haibei Yu (Nagoya University)

‡ Koichiro Ryu (Nagoya University)

‡ Shigeki Matsubara (Nagoya University)

{yu,ryu,matubara}@el.itc.nagoya-u.ac.jp

	講演者発話	対訳対応	通訳者発話
0	001-00:05:120-00:08:204 N:For my final topic I will speak about		001-00:06:328-00:09:304 I:私の最後のお話として
1	002-00:08:616-00:12:592 N:the little differences between Tokyo and Kansai		002-00:12:056-00:16:255 I:東京と関西の違いについてお話をしたいと思います

図 1: 対訳対応済みの同時通訳データの例

表 1: 分析データ基礎統計

	英語講演	英日通訳
講演数	23	92
発話時間 (s)	9,912	35,133
文数	1,207	9,293
発話単位数	5,809	20,078
単語数	32,886	149,989

る。講演者発話と通訳者発話の対訳対応が与えられたデータの例を図 1 に示す。本分析には、英語講演 23 講演と各講演に 4 人分の英日通訳の合計 92 講演分を使用した。英語講演と英日通訳の基礎統計を表 1 に示す。

3.2 単語発声時刻の推定

発話単位の音声データと書き起こし文に対して、以下の手順で単語単位の時間情報を付与する。

1. 書き起こし文を単語に分割し、各単語に読みを付与する

- 日本語：形態素解析器「茶筌」[3] が出力するカタカナの読みを音素列に変換し、単語の読みとした。音声認識エンジン Julius[4] 付属の読み付与ツールキットの規則を変換規則として用いた。
- 英語：The CMU Pronouncing Dictionary version 0.6[5] を用いて読みを付与した。

2. 単語とその音声信号区間とを対応付け

音響モデルを用いて、各単語の読みに含まれる音素を音声と対応付ける。対応付けには、音声認識エンジン Julius[4] を用いた。日本語の音響モデルには、Julius 付属の不特定話者 PTM トライフォンモデルを用いた。英語の音響モデルは、Julius 英語版音声認識キットのモデルを用いた。

3. 単語の発話開始・終了時刻を、対応する音声信号の時間情報から推定する。

本手法が付与した時刻と人手で付与した時刻の誤差が 100 ミリ以下を正解とすると、日本語の正解率は 99.2%、英語の正解率は 95.8%であった。

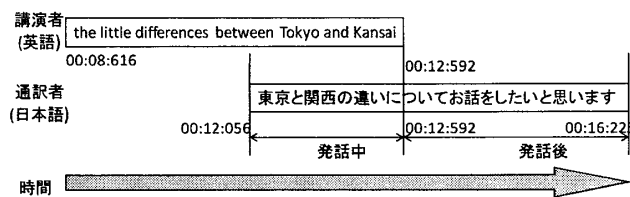


図 2: 原発話の進行状況

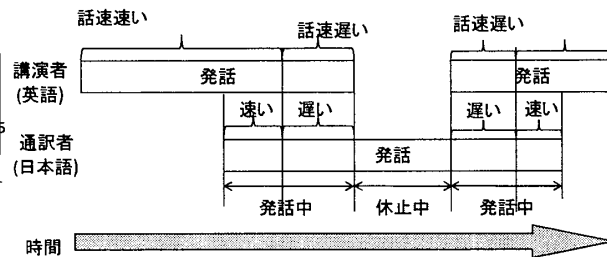


図 3: 講演者の発話状態

表 2: 原発話の進行状況による話速の比較

	発話中	発話後
モーラ数	91,179	216,039
発話時間 (s)	11,522	23,611
話速平均 (mora/s)	7.91	9.15

表 3: 講演者の発話状態による話速の比較

	発話中			休止中
	速い	遅い	全体	
モーラ数	153,792	125,830	279,622	27,596
発話時間 (s)	17,028	14,719	31,747	3,386
話速平均 (mora/s)	9.03	8.55	8.80	8.15

4 通訳者の話速変動の分析

作成した単語発声時刻の情報を用いて同時通訳者の話速変動を分析した。分析では、原発話の進行状況と講演者の発話状態に着目した。話速の計算には、日本語はモーラ数、英語はシラブル数を用いた。日本語通訳者と英語講演者の平均話速は、それぞれ (8.73mora/s), (4.12syllable/s) であった。

4.1 原発話の進行状況による話速の分析

同時通訳者は原発話の進行状況に応じて、発話速度を柔軟に制御していると思われる。特に、講演者の発話が完了する前と後とでは、訳出内容が定まっていなかったり、あるいは、定まっているという違いがあり、それは発話速度に影響することが予想される。そこで、通訳者の発話中と発話後という2つの状態に着目し、同時通訳コーパスを用いて、その発話速度を比較した。図2に、対訳対応関係にある講演者発話と通訳者発話の時間的関係を示す。この例では、通訳者発話のうち、00:12:056から00:12:592までは講演者が発話中、それ以降は発話後であり、この場合、両者の話速が比較対象となる。

分析結果を表2に示す。発話中の話速は7.91(mora/s)であり、発話後の話速9.15(mora/s)に比べて16%低下している。発話中は、通訳者は対応する訳文内容が定まっていなかったため、話速が遅い。しかし、講演者の発話が終わると、通訳者が講演者の内容を十分に把握できているので、話速が速くなると考えられる。

4.2 講演者の発話状態による話速変動の分析

同時通訳者の話速は、講演者の発話状態に応じて変動すると考えられる。例えば、発話中と休止中では、休止中の方が訳出に専念できるため、通訳者の話速が速くなることが予想される。そこで、講演者の発話状態を、発話中と休止中に分類し、さらに発話中を、各単語の話速が各講演者の平均話速よりも速い場合と遅い場合に分類した(図3参照)。講演者の各単語の話速の計算には、前後2単語加重平均を用いた。

分析結果を表3に示す。講演者が休止中の通訳者の話

速は8.15(mora/s)であり、講演者が発話中の通訳者の話速8.80(mora/s)に比べて7%低下した。また、講演者の話速が速い場合の通訳者の話速は9.03(mora/s)であり、講演者の話速が遅い場合の通訳者の話速8.55(mora/s)に比べて6%速かった。

5 おわりに

本稿では、名古屋大学同時通訳データベースを用いて、同時通訳者の話速変動を分析した。原発話の進行状況の発話中における通訳者の話速は、発話後の話速と比べ16%低下した。講演者の発話状態によって、同時通訳者の話速に違いが見られた。

謝辞: 本研究を進めるにあたり音声認識エンジン Julius についてご教示頂いた京大の河原達也先生、及び、名工大の李晃伸先生に感謝致します。

参考文献

- [1] 小野 貴博, 遠山 仁美, 松原 茂樹: 大規模音声コーパスを用いた日英・英日同時通訳における訳出遅延の比較分析, 通訳研究 第7号, pp.49-64 2007.
- [2] 名古屋大学同時通訳データベース, <http://sidb.el.itc.nagoya-u.ac.jp/>
- [3] 日本語形態素解析システム 茶筌 <http://chasen.naist.jp/hiki/ChaSen/>
- [4] 大語彙連続音声認識システム Julius, <http://julius.sourceforge.jp/>
- [5] CMU Pronouncing Dictionary <http://www.speech.cs.cmu.edu/>