

同時通訳コーパスを用いた通訳者発声タイミングの分析

高木 亮† 松原 茂樹† 稲垣 康善†

†名古屋大学大学院工学研究科 †名古屋大学言語文化部 / CIAIR
 atakagi@inagaki.nuie.nagoya-u.ac.jp

1 はじめに

自然な異言語間対話支援環境の実現を目指し、著者らは、同時通訳システムに関する研究を進めている [3]。同時通訳システムの実現においては、生成する訳文の品質だけでなく、その出力タイミングもまた重要であり、実際の同時通訳者の振舞いを調査し、分析することは効果的である。名古屋大学統合音響情報研究拠点 (CIAIR) では、音情報処理技術の高度化を目指し、多様な音声・言語データベースの整備を進めており、その一環として、同時通訳音声データベースを作成している [4]。

そこで本稿では、同時通訳コーパスを用いて、通訳者の訳文発声タイミングを分析する。このコーパスでは、発話単位ごとに発声開始及び終了の時刻が付与されている。話者発話と通訳者発話との間の対訳対応データが与えられており、両発話の時間的關係を観察することができる。これらの時間情報データを用いて、英日及び日英同時通訳における通訳単位、訳出開始の遅れ時間、及び、通訳発声速度を調査し、英日及び日英同時通訳の訳文生成タイミングに関する特徴を明らかにした。

2 同時通訳コーパス

名古屋大学 CIAIR では、これまでに、日英双方向の同時通訳音声収録し、同時通訳コーパスを構築している [4]。2001 年の収録では、英語と日本語の独話音声とその同時通訳音声を収集している。英語音声として、政治や経済、歴史の講演を、日本語音声として、計算機科学に関連したテーマの講演を収集している。また、1人の講演者に対して、経験度の異なる4人の通訳者による同時通訳音声を収録している。日本語講演では、講演者はスライドを用いており、通訳者はそのスライドを見ながら通訳を行う。書き起こしは、日本語話し言葉コーパス (CSJ)[2] の書き起こし基準に準拠しており、話し言葉に見られるフィラーや言い直し、言い誤りなどの言語現象に対してタグを付与している。また、200ms 以上のポーズ、または、文末の 50ms 以上のポーズで、発話を単位分割し、すべての発話単位に、開始時刻、終了時刻を付与している。また、人手によって対訳対応づけ作業を行っており、現在までに、日英、英日ともに、16講演分の対訳対応データを作成している [8]。

3 コーパスの基礎統計データ

同時通訳コーパス (2001 年版) の基礎的な統計データとして、収録時間、発話単位数、形態素 (単語) 数、異なり

表 1: 同時通訳コーパスの統計データ

項目	英語	英日	日本語	日英
収録時間 (分)	239	955	146	585
発話時間 (秒)	11458	40473	5690	22289
形態素数	35474	165127	21708	63872
異なり形態素数	3506	6115	1457	3130
発話単位数	5573	20496	4505	12468
発話文数	1263	7423	651	3315
フィラー	1251	8067	1482	3261

表 2: 対応済みデータに関する統計データ

項目	英日対応数:1717		日英対応数:1441	
	合計	平均	合計	平均
重なり時間 (秒)	1845	1.07	1232	0.86
開始遅れ時間 (秒)	5482	3.19	5955	4.13
終了遅れ時間 (秒)	5383	3.13	5829	4.04
話者発話時間 (秒)	6593	3.84	5077	3.52
通訳発話時間 (秒)	6214	3.62	5388	3.74
話者発話形態素数	22360	13.02	18572	12.89
通訳発話形態素数	25926	15.10	15554	10.79

形態素 (単語) 数、発話時間、及び、談話タグを調査した。表 1 に調査結果を示す。表中の英語、英日、日本語、日英は、それぞれ英語講演、英日同時通訳、日本語講演、日英同時通訳のデータを示す。

形態素 (単語) 数は、英語は単語数とし、日本語は茶筌 [5] の解析結果に基づいて形態素数を計算した。異なり形態素数の調査では、英語は表記が異なる単語を、日本語は基本形が異なる単語を調べた。

1人の講演者音声に対して、異なる4人の通訳者による同時通訳を収録したため、通訳の収録時間は講演のほぼ4倍となる。一方、通訳者音声の発話文数は、英日通訳で話者音声の約6倍、日英通訳で話者音声の約5倍となっており、これは、通訳者は話者の文を複数の文に分割して訳すことがあるという通訳理論における経験的知見 [6] を裏付けている。

日本語話者発話	日英通訳者発話
0001 - 00:05:120-00:05:879 N: 本日は	0001 - 00:07:096-00:07:596 I: Today,
0002 - 00:06:352-00:08:672 N: 派生文法に基づく日本語処理 0003 - 00:08:911-00:10:848 N: とその応用というタイトルで	0002 - 00:08:004-00:09:199 I: (F ah) Japanese 0003 - 00:10:672-00:15:167 I: processing based on (W deliberation;derivation) of grammar and (F uh) (X its pa) its application 0004 - 00:15:632-00:18:039 I: is a theme of (W a) by (F ah) lecture,
0004 - 00:11:168-00:13:240 N: (F えー) 我々の研究を紹介します (SB)	0005 - 00:18:279-00:19:655 I: (W I've;i'd) like to report (F ah) 0006 - 00:19:848-00:20:879 I: research in this area(SB)
0005 - 00:15:872-00:16:912 N: (F えー) まず (F えー)	0007 - 00:21:184-00:21:615 I: First,
0006 - 00:17:200-00:19:919 N: コンピューターによる日本語処理についてですが	0008 - 00:22:567-00:26:472 I: (F ah) Japanese processing by computer is the first topic(SB)
0007 - 00:20:376-00:20:719 N: (F えー)	

図 1: 対訳対応済みの同時通訳コーパス

4 対訳対応データ作成と分析

4.1 対訳対応データの作成

通訳者音声の詳細に分析するには、話者の発話とその通訳者の発話が比較的小さな単位で対応づけられていることが望ましい。そこで、人手によって対訳対応づけ作業を行い、対訳対応データを作成した [8]。その際、

- 発話単位を対応の最小単位とする。
- できる限り細かい対応づけを行う。
- フィラーなどの非言語現象のみからなる発話単位や、通訳者の省略や補足などで話者発話に対応先がない発話に対して「対応先なし」を許容する。

という条件を設定して実施した。

4.2 対応済みデータに関する統計データ

対訳対応が付いている講演データとその通訳データにおいて、対応先が存在する話者発話、及び、その通訳者発話に対して、以下の項目を調査した。表 2 に結果を示す。表中の「合計」は 16 講演分の合計であり、「平均」は 1 対応あたりの値を表す。

- 通訳者の発声において話者音声と重なっている時間
- 通訳者の発話開始・終了遅れ時間
- 話者・通訳者の発話時間
- 話者・通訳者発話の形態素数

4.3 通訳者の訳文発声タイミングの分析

同時通訳では、通訳者は、話者の発話途中で訳出を開始する。これは話者の発話の一部を通訳単位として捉え、その訳を早い段階で訳出しているためであると予想される [1, 7]。そこで、対訳対応が付いているコーパスを分析し、通訳者が用いている通訳単位とその発声タイミングを調査した。

4.3.1 日英同時通訳の分析データ

対訳対応が付いている日本語講演 16 組について調査を行った。コーパスの例を図 1 に示す。左側が日本語話者発話、右側が日英通訳者発話を表す。

通訳者の通訳単位は、訳出可能なタイミングを決定するための手がかりになる可能性がある。そのような単位を明らかにするために、対応先が存在する話者発話、通訳者発話の組に関する特徴を調べた。対応先が存在する話者発話の形態素数の分布を図 2 の実線で示す。平均形態素数は 12.89 であり、20 形態素までで全体の 83.1% を占める。

このうち、形態素数の少ない発話は、通訳単位を構成している可能性があり、このような発話を検出したときには即座に対訳を生成できるということになる。形態素数が 4 以下の発話 (全体の 16.0% にあたる 231 組) の特徴を調べた。主な特徴を以下に挙げる。括弧内にその頻度を示す。

- 話者の発話が接続詞である (84 組)
話者 そして
通訳者 Then
- 話者の発話の主語である (42 組)
話者 ユーザーが
通訳者 the user,
- 前置詞句として訳出が可能である (65 組)
話者 構文解析では
通訳者 in syntactic parsing,

次に、通訳者の発話開始時間の遅れ (話者の発話開始時間と通訳者の発話開始時間の差) を調査した。小さな遅れで通訳を行えるのには、それなりの理由が存在すると予想され、それは早いタイミングで訳出を開始するための

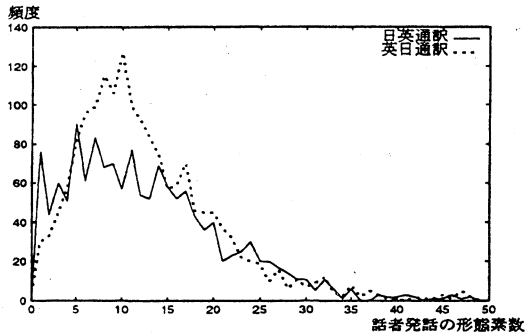


図 2: 1 対応あたりの話者発話の長さの分布

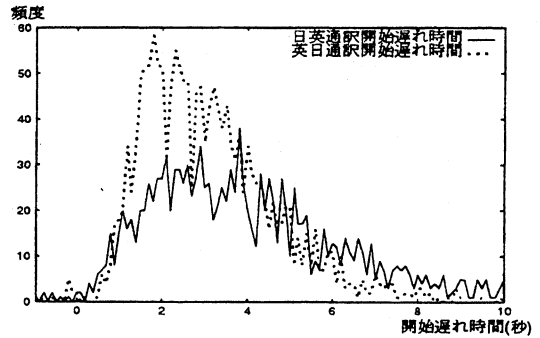


図 3: 開始遅れ時間の分布

テクニックとして利用できる可能性がある。分布を図 3 の実線で示す。平均開始遅れ時間は 4.1 秒である。

通訳者の開始遅れ時間が小さい対応組の上位 30 個 (0.79 秒以下) を対象として、追従が可能になる要因を調査した。これらの対応に関して見られた主な特徴として以下のようなのものが挙げられる。

- 講演に用いたスライドの情報による予測 (6 組)

話者はスライドを用いて講演を行っている。通訳者もこのスライドを見ることができるため、スライドに記された情報を参照することにより、話者が次に話す内容を予測し、追従することができる。

話者 背景ですが

通訳者 This slide shows the background of us.

- 接続詞 (11 組)

話者が接続詞を用いた場合に通訳者はその後ろの発話を待たずに訳出できる。

話者 したがって

通訳者 Therefore

- 主語の補間 (4 組)

日本語話者は主語を省略する場合がある。このときに、通訳者は文脈から一般的な主語 “we”, “I” などを用いて対応するため、早い段階での訳出が可能となる。

話者 分類したところ

通訳者 we analyzed such patterns

- 主語の訳出 (3 組)

主語にあたる部分は、日本語、英語ともに文頭にくるため即座に訳出できる。

話者 ユーザーが

通訳者 the user,

- フィラーの挿入 (1 組)

話者の次の発話を待つ間に、フィラーを発声し間を埋める。

話者 以上で終わります。

通訳者 ah This completes my lecture.

4.3.2 英日同時通訳データの分析

対訳対応が付いている英語講演 16 組について調査を行った。対応先が存在する話者発話の形態素の分布を図 2 の点線で示す。日英通訳の調査と同様、形態素数が 4 以下の発話 (全体の 9.90% にあたる 170 組) を調べた結果、以下のように日英通訳と同様の特徴が見られた。

- 話者の発話が接続詞である (36 組)

話者 For example,

通訳者 例えば

- 話者の発話が主語である (32 組)

話者 some of the machines

通訳者 そのうちの機械のいくつか

- 節・句を構成 (48 組)

話者 Someone once said that

通訳者 ある人が言いました。

次に、日英通訳データと同様に通訳者の発話開始時間の遅れを調査した。分布を図 3 の点線で示す。平均で 3.2 秒の遅れとなった。

通訳者の開始遅れ時間が小さい対応組の上位 30 個 (0.96 秒以下) を調べたところ、以下のような要因が見られた。

- 主語の部分の翻訳 (7 組)

英語でも日本語でも主語は文頭、もしくは、それに近い位置に現れる。このため、主語の部分に関しては後ろに続く語に関係なく翻訳が可能となる。

話者 I arrived in Japan

通訳者 え 私 日本に到着しましたのが

- フィラーの挿入 (4 組)
次に続く話者発話を待つためフィラーを挿入して間を繋ぐ。

話者 these are supporting countries.

通訳者 あー というような国が支援をしているわけ
であります。

- 接続詞 (4 組)
話者の発話が接続詞の場合は、すぐに翻訳が可能となる。

話者 but after walking up and down

通訳者 え しかし え一道を歩いているうちに

- 副詞句・前置詞句 (2 組)
日本語の語順は英語に比べ柔軟であり、文頭の副詞句や前置詞句は先に翻訳することが可能となる。

話者 Here in Kansai

通訳者 ん 関西では

4.3.3 日英通訳と英日通訳の比較

通訳の開始遅れは、英日通訳の方が日英通訳よりも、1.0 秒ほど小さい。その要因として、(1) 日本語は英語に比べ語順が柔軟であるため、英語単語の出現順に従って翻訳することにより同時性の高い通訳が可能になるといこと、(2) 英語講演のトピックが一般的なものであったので、通訳者の背景知識によって予測が可能であったこと、(3) 日本語では動詞が文末に現れるため、日英通訳における訳文の文構造が早い段階では定まりにくいこと、などが考えられる。

4.3.4 通訳者の発話速度の分析

同時通訳における訳出が遅れる原因として、ある発話に対する翻訳に時間がかかることによるものと、前の発話に対する訳出が遅れており、現在の発話の訳出を開始できないことによるものがある。通訳者は、話者発話に追従するために、自らの発声速度を柔軟に調節していると考えられる。そこで、対訳対応が付いたコーパスにおいて、対訳対応箇所の通訳者の発話速度を調査した。発話速度は、発話単位あたりのモーラ数で計算した。英語のモーラ数は単母音で1モーラ、長母音で2モーラ、母音の後ろの子音で1モーラとした [9]。日本語のモーラは茶筌 [5] の解析結果を利用して定めた。結果を図4に示す。

通訳が進むにつれて、発話速度が増加している。通訳者は同時性を高めるために早い段階で訳出を始めるが、開始当初は、入力された発話の、対応する訳文においての役割が定まらないために速度は遅い。しかし、話者の発話が進み、通訳者が次第に話者の内容を把握するのに伴って、出来るだけ早い段階で訳出できるよう、通訳者の短期記憶に訳出する情報が留まらないように徐々に速度を上げていることが考えられる。

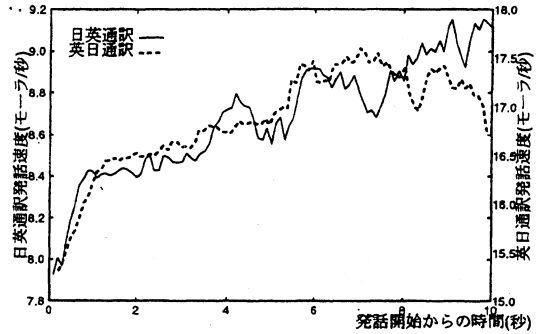


図 4: 通訳者発話速度

5 まとめ

本稿では、対訳対応付き同時通訳コーパスを用いて、英日及び日英通訳における同時通訳単位、訳出開始の遅れ時間、及び、通訳者発話速度を調査し、通訳者の訳文生成のタイミングを分析した。同時通訳において、話者が接続詞を用いたときや主語にあたる部分が特定できたときには即座に訳出できる可能性があること、また、訳出可能な情報の量に応じて発話速度を制御することにより、訳出遅れの少ない訳文生成が可能になる、など同時通訳の訳文生成タイミングに関する特徴を明らかにした。

参考文献

- [1] 船山中他, “同時通訳における処理単位について”, 通訳理論研究, Vol.6, No.1, pp.4-13 (1996).
- [2] 前川喜久男, 籠宮隆之, 小磯花絵, 小塚秀樹, 菊池英明, “日本語話し言葉コーパスの設計”, 音声研究, 第4巻第2号, pp.51-61 (2000).
- [3] S. Matsubara, K. Toyama, Y. Inagaki, “Sync/Trans: Simultaneous Machine Interpretation between English and Japanese.”, In N. Foo (Ed) *Advanced Topics in Artificial Intelligence, Lecture Note in Artificial Intelligence*, Vol.1747, pp.134-143 (1999).
- [4] 松原茂樹, 相澤靖之, 河口信夫, 外山勝彦, 稲垣康善, “同時通訳コーパスの設計と構築”, 通訳研究, No.1, pp.85-102 (2001).
- [5] 松本裕治ら, 形態素解析システム『茶筌』version 2.2.1 使用説明書, <http://chasen.aist-nara.ac.jp/>
- [6] 水野的, “日英同時通訳研究ノート”, 通訳理論研究, Vol.5, No.2, pp.4-21 (1995).
- [7] 新崎隆子, “同時通訳と逐次通訳における情報処理”, 通訳理論研究, Vol.4, No.2, pp.40-46 (1994).
- [8] 高木亮, 松原茂樹, 稲垣康善, “同時通訳コーパスの対訳アライメント手法とその評価”, 情報処理学会第64回全国大会 (2002).
- [9] 田窪行則ら, “岩波講座 言語の科学 2 音声”, 岩波書店, (1998).