

同時通訳を介した異言語間対話の時間的特徴
～ 逐次通訳との比較に基づく対訳コーパスの分析 ～

大原 誠¹ 松原 茂樹^{2,3} 笠 浩一郎¹ 河口 信夫^{2,3} 稲垣 康善⁴

名古屋大学大学院工学研究科¹

名古屋大学情報連携基盤センター²

名古屋大学統合音響情報研究拠点³

愛知県立大学情報科学部⁴

In recent years, there have been a number of studies on simultaneous machine interpretation to develop the environments for supporting multilingual communications. The effectiveness of simultaneous interpreting in cross-language conversation, however, remains to be investigated. This paper describes the characteristic features of dialogues through simultaneous interpreting in comparison to those through consecutive interpreting. We have investigated them using the CIAIR simultaneous interpreting corpus. As a result, we have confirmed that simultaneous interpreting is effective in significantly increasing both the efficiency and smoothness of cross-language conversation.

1. はじめに

同時通訳と逐次通訳という2つの通訳スタイルの比較は、通訳論における重要な研究対象の一つであり、これまでに多くの議論が重ねられてきた(たとえば、Gile, 2001 など)。通訳スタイルの違いは、一般に、通訳者の訳出した文章の語彙的・構文的特性や通訳作業中のノートテイキングなど、さまざまな形となって現われる。これらはいずれも、通訳者の情報処理過程を解明するための重要な研究資料であり、それゆえに、2つのスタイルを比較するというテーマは通訳研究に携わる多くの研究者の関心を集め続けてきた。しかしながら、通訳研究の目的は、必ずしも通訳者の振舞いの分析だけに留まるものではなく、その周辺に現われる現象も重要な研究対象となる。通訳スタイルの違いによって生

OHARA Makoto, MATSUBARA Shigeki, RYU Koichiro, KAWAGUCHI Nobuo, and INAGAKI Yasuyoshi,
“Temporal Features of Cross-Lingual Communication Mediated by Simultaneous Interpreting:
An Analysis of Parallel Translation Corpus in Comparison to Consecutive Interpreting.”
Interpretation Studies, No. 3, December 2003, pages 35-53.

(c) 2003 by the Japan Association for Interpretation Studies

じる他の側面として、通訳を介した異言語間対話コミュニケーションの進行プロセスに関する違い、特に、話者の発話タイミングや聴者の聴取タイミングに与える影響を挙げることができる。コミュニケーション形態の観点から 2 つの通訳スタイルを比較することは、話者や聴者の立場から通訳を捉えるという意味で意義深い視点を提供する。

一方、自動通訳の分野においては、音声・言語処理技術の発展を背景に、音声通訳に関する研究が 90 年代前半から活発に進められている(松原 2001a)。すでに、対話翻訳の実験システムがいくつか提案されており(Watanabe, 2000; 菅谷 2002)、特定タスクドメインでの異言語間対話の実現可能性が明らかになりつつある。ただし、これまで開発されたシステムのほとんどは逐次通訳として実現されており、対象となる対話は、発話者だけでなく通訳システムの発話権も保証することを前提としている。実際、対話翻訳システムにおける望ましい通訳スタイルについては、これまでほとんど議論されてこなかった。その理由は、単に技術的な問題によるものであり、開発の困難さを考慮すると同時通訳を採用するという選択肢はそもそも存在しなかったためである。自動通訳システムの実用化が進み、その次世代システムについて検討されつつある現在、自動通訳システムの理想的な通訳スタイルについても議論を要するようになってきたといえよう。事実、漸進的な対話翻訳技術に関する研究がいくつか行われ、同時通訳に適した解析、変換、および生成技術の開発も進行しており(Amtrup, 1999; Matsubara, 1999; Mima, 1998)、今後の発展が期待される。

逐次通訳を介した異言語間対話を遂行する上で生じる問題として、通訳を介することによる対話の効率および円滑さの低下が挙げられる。異言語間対話の話者は、一方の言語のみに精通しているため、もう一方の言語でやり取りが進められている間は、対話における基本行為である「話す」、「聞く」といういずれの行為にも携わらない時間となる。そのような時間が長くなると、対話タスクの遂行に要する時間が長くなると同時に、話者に適した対話のテンポを維持することも容易ではなくなる。その一方で、同時通訳は、話者の発声中でも訳出を開始し、できる限り話者発話に追従しながら通訳を進めるものであり、対話の効率および円滑さの低下の軽減が期待できる。ただし、同時通訳は、逐次通訳に比べ、訳出に関する時間的制約が大きい分、訳文の品質が低下することは否めない。コミュニケーションの形態上、同時通訳の導入が不可欠である独話通訳と異なり¹⁾、異言語間対話に同時通訳を導入する必然性は必ずしも明らかではない。実際の同時通訳者を介した対話を調査することにより、同時通訳機能を備えた対話翻訳技術の有用性を、システムのユーザである話者および聴者の観点から評価することが望まれる。

そこで本論文では、自然な異言語間コミュニケーション支援環境の通訳形態として同時通訳を採用することの効果について述べる。異言語間対話の円滑さ、および効率の観点から同時通訳データを定量的に分析し、逐次通訳を介した対話との比較を通して、その効果の程度について考察する。

本研究では、データの分析対象として、名古屋大学同時通訳音声コーパス(松原 2001b)の 2001 年度版を使用する。これは、海外旅行をタスクドメインとする同時通訳・逐次

通訳を介した英語・日本語間対話の音声および文字化データであり、名古屋大学統合音響情報研究拠点 (CIAIR) における音声データベースの整備の一環として収集している (Kawaguchi, 2002)。延べ 72 時間分の音声収録、文字化、談話タグ付与、ならびに、対訳対応付け等の作業が完了しており、30 万単語規模のコーパスが構築されるに至っている。本研究では、同時通訳を介した対話の効率を評価するために、対話の遂行に要する時間を、また、円滑さについて評価を与えるために、一方の言語のみによって会話のやり取りがなされている時間を調査し、逐次通訳との間で比較した。特に、話者の発声時間が長い発話に着目し、上述の時間が逐次通訳の場合と比べてどの程度短縮されるかについて考察した。分析の結果、円滑さを備えた効率的な異言語間対話を遂行するために、通訳形態として同時通訳を採用することの効果を確認した。さらに、同時通訳と逐次通訳の大きな違いの一つである、話者発話とそれに対応する通訳者発話の時間的重なりに着目し、話者発話時間、および通訳者発話時間と重なりの有無との関係について調査した。

本論文の構成は以下の通りである。次の 2 章で、同時通訳を介した異言語間対話の効率と円滑さについて説明する。3 章では、対話同時通訳コーパスとその基礎統計情報を示し、4 章で、通訳を介した対話の時間的特徴について分析した結果を報告する。5 章では、話者、および通訳者発話時間と発話の重なりの有無の関係について述べる。

2. 同時通訳を介した異言語間対話

本章では、まず、逐次通訳と同時通訳について概説する。次に、通訳を介した異言語間対話について述べ、円滑さを備えた効率的な異言語間対話を遂行するための要因について考察する。また、逐次通訳を介した異言語間対話と、同時通訳を介したそれを模式的に比較することによって、同時通訳技術が異言語間対話に有効に機能する可能性について論じる。

2.1 逐次通訳と同時通訳

通訳には、大きく分けて逐次通訳 (consecutive interpretation) と同時通訳 (simultaneous interpretation) の 2 つのスタイルがある。その違いは、通訳者の訳出開始が、話し手が一区切り発言し終わった後か (逐次通訳)²⁾、あるいは、話者が発言を終わらないうちか (同時通訳) にあり、それは話者発話とその通訳者発話が時間的に重なるか否かといった違いになって現われる。これらのスタイルの比較については、これまで訳出結果を用いた分析が中心であり、特に、訳文のスタイルや品質等に関して詳細な議論が与えられてきた。一般的には、同時通訳では話者は自分のペースで発話することができ、通訳者の訳出状況をそれほど気にする必要がないといった特徴が、逐次通訳では訳出に関する時間的制約が緩く、訳の品質が保たれやすいといった特徴がある。

2.2 通訳を介した異言語間対話

逐次通訳、同時通訳を介した英語・日本語間対話の模式図を、それぞれ図 1(a)、(b)

に示す。この対話は、英語話者、英日通訳者、日本語話者、および日英通訳者の 4 人の対話参加者によって遂行されるが、使用言語で参加者を区分すると、2 つの同一言語間の対話が並行して進められているとみなすこともできる。すなわち、日本語話者と英日通訳者による日本語対話、ならびに、英語話者と日英通訳者による英語対話が互いに同期しながら進行している。

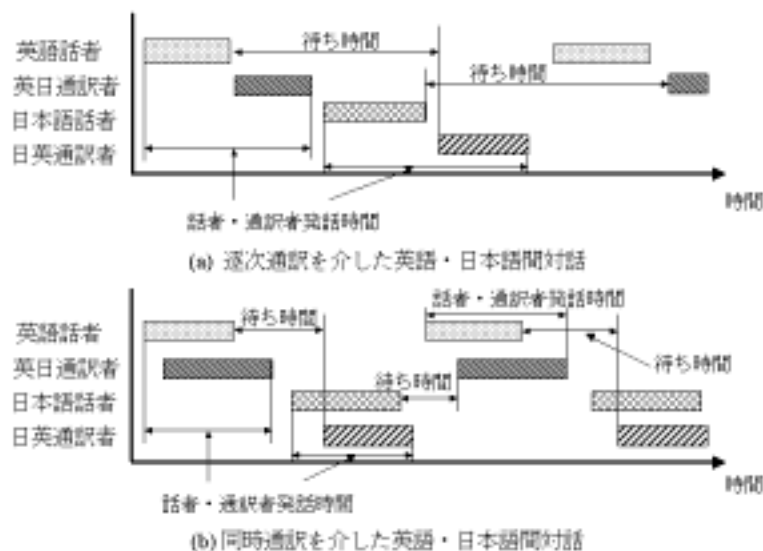


図 1 通訳を介した異言語間対話の模式図

2.2.1 対話の効率

対話とは、あるタスクを達成することを目的として、話者間で協調的に遂行される共同作業であり、それは、発話という情報伝達行為の連続によって実現される。この意味で、各発話に要する時間は、対話の効率の一端を表現しているといえる。ただし、異言語間対話における相手への情報伝達行為には、当然、通訳者による発声も含まれる。このため、話者発話の開始からその通訳者発話の終了までに要した時間(以下、話者・通訳者発話時間)が、異言語間対話の効率を計る指標となり、それが短いほど効率がよいということになる。図 1 内に話者・通訳者発話時間を矢印で示す。

逐次通訳では、通常、話者発話と通訳者発話が時間的に重ならないため、話者・通訳者発話時間は、話者の発話時間と通訳者の発話時間の合計よりも長い。たとえば、通訳に話者発話と同じだけの時間が必要だとすれば、話者・通訳者発話時間は、話者による発話時間の約 2 倍となる。一方、同時通訳を介した対話では、話者発話が終了したときすでにある程度の訳出が行われている可能性がある。話者発話終了後には、訳出されていない残りの内容のみを訳出すればよく、逐次通訳に比べ話者・通訳者発話時間が短くなる。よって、同時通訳を介することによって、より効率的な対話を遂行できる可能性がある。

2.2.2 対話の円滑さ

円滑な対話を実現するために、話者交替がスムーズに行われる必要があり、特に、一方の話者の発話終了後、もう一方の話者が発話を開始するまでに、不自然な程度の長さの間が生じることは望ましくない。ところが、異言語間対話では、話者は一方の言語のみに精通しているのが普通である。相手側の通訳者の発話を聞くことによるのみ、相手発話の内容を理解することができるため、話者は自らの発話ターンを終えた後、相手側の通訳者による発声開始を待つことになる。この意味で、本論文の以下では、話者が発話を終了してから相手通訳者の発話が始まるまでの時間を待ち時間と定義し、異言語間対話の円滑さの指標とする。図 1 内に、英語、日本語話者の待ち時間を矢印で示す。

逐次通訳の場合、話者の発話が終了してから通訳が始まるため、待ち時間は、通訳者の発話時間と相手話者の発話時間の合計よりも長くなる。一方、同時通訳を介した対話では、話者の発話に対する通訳者の発話が逐次通訳よりも早く終わることに加えて、相手話者の発話途中でもそれに対する訳出が始まることもあり、話者の待ち時間は逐次通訳の場合に比べて短くなる。このため、より円滑な対話を遂行できる可能性がある。

3. CIAIR 対話同時通訳コーパス

前章で述べたように、円滑さを備えた効率的な異言語間対話を実現するために、同時通訳技術が効果的に機能する可能性がある。しかし、対話に同時通訳を導入することによる効果の程度については必ずしも明らかでない。これについて調査するために、実際の同時通訳データを用いて定量的分析を与えた。データとしては、名古屋大学 CIAIR⁴⁾ において 2001 年度に収集した、同時通訳者を介した英語・日本語間音声対話を用いた。本章では、分析に使用した対話同時通訳コーパスの構築とその基礎統計について述べる。

3.1 収録

対話ドメインとして、「旅行会話」を採用し、空港やホテルなど、海外旅行において想定されるトピックを定め、模擬会話を収録した。トピックとして以下のものを採用した。

- 空港でのチェックイン
- 空港サービスカウンターでの観光案内
- 空港からホテルへの電話による問い合わせ
- 飛行機の予約、変更、キャンセル
- ホテルの予約、変更、キャンセル
- ホテルのチェックイン、チェックアウトなど、フロントでの会話

CIAIR では、実音響環境下での音声データを大量に収集することを目的の一つとしており、収録は教室レベルの録音環境を採用した。同時通訳では、対象となる音声だけでなく、その発声者の表情や振舞いも重要な情報となるため、通訳者は話者をガラス越しに観察可能な専用のブースに入り通訳を行う。話者、通訳者は全て同一の接話マイクを使用し、

サンプリング周波数 16kHz、16bit でデジタル化し、デジタルオーディオテープ (DAT) に複数チャンネル環境で収録している。

英語話者と日本語話者による異言語間対話に対して、通訳の品質を高めるために、英日および日英の 2 人の同時通訳者を用いている。会話における話者の発話権を確保するために、話者は相手話者の発声を通訳した結果のみを、また、通訳者は、担当する話者の音声のみを聞き取れることにした。

なお、対話様式は模擬的であり、話者の発話する内容もある程度制御したいという要求があるが、その一方で、できる限り自由な発話を収集することも考慮し、話者役割と対話タスクのみを事前に設定し、話者に提示している。たとえば、航空券の予約対話であれば、客を担当する話者には予約する航空券の種類、枚数などを、また、受付を担当する話者には発券可能なフライトをシートとして提示した。

3.2 音声データの文字化

文字化は、国立国語研究所を中心に構築が進められている日本語話し言葉コーパス (CSJ) (前川 2000) の書き起こし基準に準拠して行った。話者および通訳者の音声を 200msec 以上 (文末では 50msec 以上) のポーズで分割することにより発話単位を定め、その開始および終了の時間タグを付与した。また、フィラーや言い淀み、言い直しなどには言語タグを付与した。対話音声の文字化データのうち、英語話者発話のサンプルを図 2 に、英日通訳者発話のサンプルを図 3 に、それぞれ示す。

0021 - 01:47:544-01:52:392 N:
(F uh) Here is your claim tag<SB> You may use this to reclaim your luggage when you arrive<SB>
0022 - 01:59:808-02:02:296 N:
To the baggage claim people on the other side<SB>
0023 - 02:08:960-02:10:863 N:
Can I help you with anything further?<SB>
0024 - 02:17:208-02:18:032 N:
Oh, I see<SB>
0025 - 02:29:400-02:32:175 N:
(F well) From here can you see that escalator over there?<SB>
0026 - 02:37:224-02:42:424 N:
Okay, so just take that escalator up to the third floor and from there you should see immigration<SB>

図 2 文字化データのサンプル(英語話者)

0031 - 01:49:096-01:54:359 I: (F えー)ここに(F いー)引き取りのタグがあります<SB>荷物をお取りになるときにお出してください<SB>
0032 - 02:01:104-02:04:743 I: (F えー)そちらの(F お)荷物受け取りの所に職員がおりますので<SB>
0033 - 02:10:240-02:12:383 I: (F うー)何か(F あー)ご質問ございますか<SB>
0034 - 02:18:096-02:18:975 I: (F あー)そうなんですか<SB>
0035 - 02:30:128-02:33:432 I: (F あ)そうですね<SB>ここからエスカレーターそこにございますね<SB>
0036 - 02:37:992-02:41:736 I: (F えー)そうですねえ<SB>そのエスカレーターを乗って三階まで行ってください<SB>
0037 - 02:41:736-02:41:991 I:<息>
0038 - 02:42:040-02:44:671 I: (F えー)そこに(F い)出国審査の(F お)ところございます<SB>

図 3 文字化データのサンプル(英日通訳者)

3.3 同時通訳データの対訳対応付け

通訳開始の時間的特徴の抽出や、通訳パターンの獲得、通訳単位の認定など、通訳者音声を詳細に分析するためには、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付けられている必要がある。このため、人手による対訳対応付け作業を実施し、対訳対応データを作成した。対応付け作業に当たり、以下を仕様として設定した。

- 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付けを行う。
- 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。

3.4 コーパスの統計データ

2001年度に構築した CIAIR 同時通訳コーパスの基礎統計データとして、収録時間、発話時間、形態素(単語)数、異なり形態素(単語)数、発話単位数、発話文数、およびフィラーの数を調査した。表 1 に調査結果を示す。ここで、形態素(単語)数は、英語は単語数とし、日本語は「茶筌」(松本 2000)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(単語)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数え上げた。

表 1 対話同時通訳コーパス(2001 年度版)の基礎統計データ

項目	英語話者	英日通訳者	日本語話者	日英通訳者
収録時間(分)	1082	1082	1082	1082
発話時間(秒)	17865	17495	18534	17128
形態素(単語)数	73513	77790	75064	64306
異なり形態素(単語)数	2713	2855	2601	2615
発話単位数	10858	11354	13176	10813
発話文数	7889	9595	8725	6991
フィラー	2376	4396	4892	2712

さらに、日本語話者データと日英通訳者データに対する対訳対応付け、および英語話者データと英日通訳者データに対する対訳対応付け作業を実施した。その結果、日英で 7621 個、英日で 7691 個の発話対応を得た。これらのデータに対して、以下の項目を調査した。

- 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、および通訳者の発話時間
- 話者、および通訳者発話の形態素(単語)数

表 2 に結果を示す。なお、表中の値は全対応における平均を表す。

表 2 対応済みデータに関する統計データ

項目	日英対応	英日対応
重なり時間(秒)	0.65	0.63
開始遅れ時間(秒)	1.88	1.68
終了遅れ時間(秒)	1.78	1.81
話者発話時間(秒)	2.21	2.19
通訳者発話時間(秒)	2.15	2.20
話者発話形態素(単語)数	9.33	9.47
通訳者発話形態素(単語)数	8.49	10.29

4. 同時通訳コーパスの特徴分析

2 章では、異言語間対話において逐次通訳を介する場合と同時通訳を介する場合を比較し、同時通訳技術が効果的に機能する可能性について論じた。本章では、名古屋大学同時通訳音声コーパスを用いて、異言語間対話の円滑さと効率について定量的分析を与

える。本分析のために、逐次通訳を介した同一ドメインの対話を新たに収録し、比較対象として用いた。作成した逐次通訳対話データの規模を表 3 に示す。本分析の対象である同時通訳対話データと比べると、その規模は必ずしも大きくない。しかしながら、逐次通訳という形態を考慮すると、その時間的特徴が収録データ間で大きく異なることは少ないと考えられ、また、参考データとして利用することを勘案し、これらを比較対象として使用した。なお、本論文では、ターンを単位とした話者発話と通訳者発話に注目する。発話単位間の対応データから対話のターン間の対応データを自動的に作成し、分析ではそれを対訳対応データとして利用した。

表 3 逐次通訳対話データの基礎統計データ

項目	英語話者	英日通訳者	日本語話者	日英通訳者
収録時間(分)	72	72	72	72
発話時間(秒)	895	980	768	922
形態素(単語)数	3592	4367	3123	3623
発話単位数	405	428	515	336
発話文数	154	363	368	132

4.1 対話の効率に関する分析

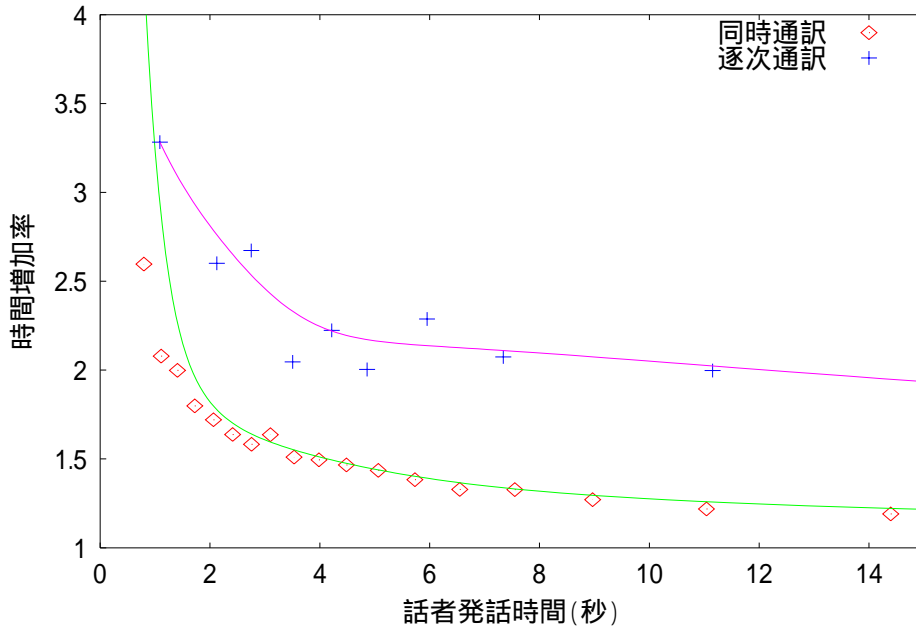
4.1.1 話者・通訳者発話時間による比較

異言語間対話の効率を評価するために、まず、同時通訳を介した場合と逐次通訳を介した場合のそれぞれについて、日英対応の話者・通訳者発話時間を調べた。同時通訳を介した対話としてコーパスの 216 対話に含まれる 4233 対応を、逐次通訳を介した対話として 12 対話に含まれる 152 対応を対象とした。同様に、英日対応の話者・通訳者発話時間の調査においては、同時通訳を介した対話としてコーパスの 216 対話に含まれる 4330 対応を、逐次通訳を介した対話として 12 対話に含まれる 153 対応を対象とした。

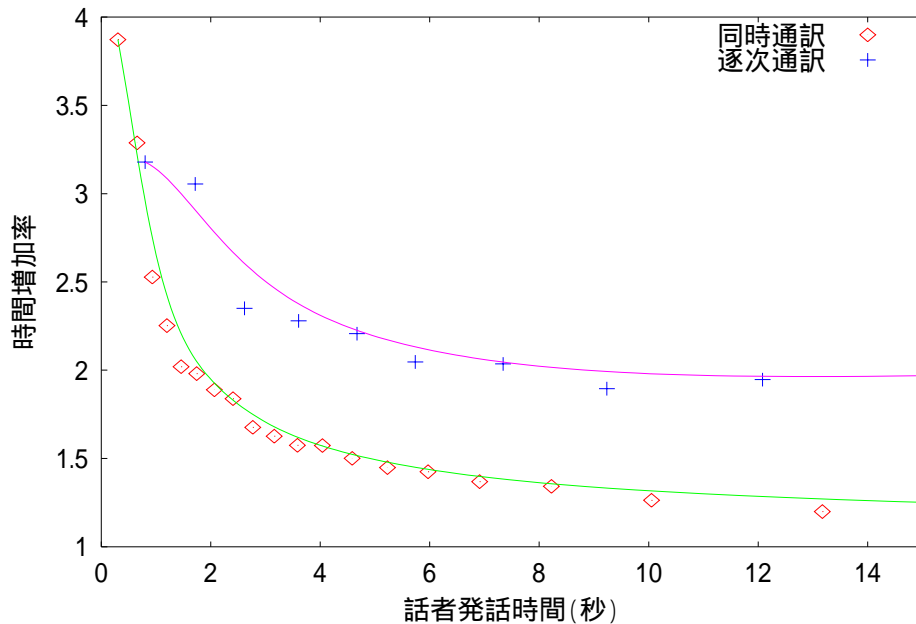
異言語間対話の話者・通訳者発話時間は、同一言語間対話と比べ、通訳が介入する分だけ増加する。この意味で、話者発話時間とそれに対する話者・通訳者発話時間の比を時間増加率と呼び、その値をターンごとに計算した。話者発話時間と時間増加率との関係を図 4(a)、(b) に示す。同時通訳を介した対話については、すべての話者・通訳者発話を話者発話時間でソートした上で 20 等分し、それぞれの時間増加率の平均値を図上に示した。また、逐次通訳の場合には、データの過疎性を考慮し、10 等分して図上に示した。³⁾

日英通訳においては、逐次通訳における時間増加率は、話者発話時間が大きくなるとともに低下しているものの、その減少の度合は緩やかであり、時間増加率はおよそ 2 に収束している(図 4(a))。それに対して同時通訳では、話者発話時間が長くなるとともに時間増加率はおよそ単調に減少している。特に、話者発話時間が 5 秒を超える発話においては、

時間増加率は平均して 1.27 となり、逐次通訳の 2.01 を大きく下回る。また、時間増加率が 1.5 以下の話者・通訳者発話は、通訳者のみが発声する時間が話者発話時間の半分以下であることを意味し、それは同時通訳の話者・通訳者発話の 58.0% を占める。英日通訳においても、多少の数値の差はあるものの、日英通訳と同様の傾向が見て取れる (図 4(b))。



(a) 話者発話時間と時間増加率の関係 (日本語話者・日英通訳者)



(b) 話者発話時間と時間増加率の関係 (英語話者・英日通訳者)

図 4 話者発話時間と時間増加率の関係

話者発話時間と時間増加率の比（逐次通訳/同時通訳）の関係を図 5 に示す。これは、同時通訳を介した場合の時間増加率における効果の程度を示す。逐次通訳を介した対話については、すべての話者・通訳者発話を話者発話時間でソートした上で 5 等分し、それぞれの時間増加率の平均値を調べた。同時通訳については、5 等分した逐次通訳とデータ範囲を合わせて 5 つに分割し、同様に増加率の平均値を調べた。そして、それらの比を図上に示した。話者発話時間が短い発話に関しては、時間増加率の比の値が小さく、同時通訳を介した効果の程度は低い。たとえば、日英通訳において、話者発話時間が 2.4 秒以下の短い発話における時間増加率の平均はほとんど変わらないのに対して、話者発話時間がより長い発話に関しては、時間増加率の平均の比がおよそ 1.5 となり、同時通訳を採用することによって話者・通訳者発話時間が大幅に短縮される（図 5）。

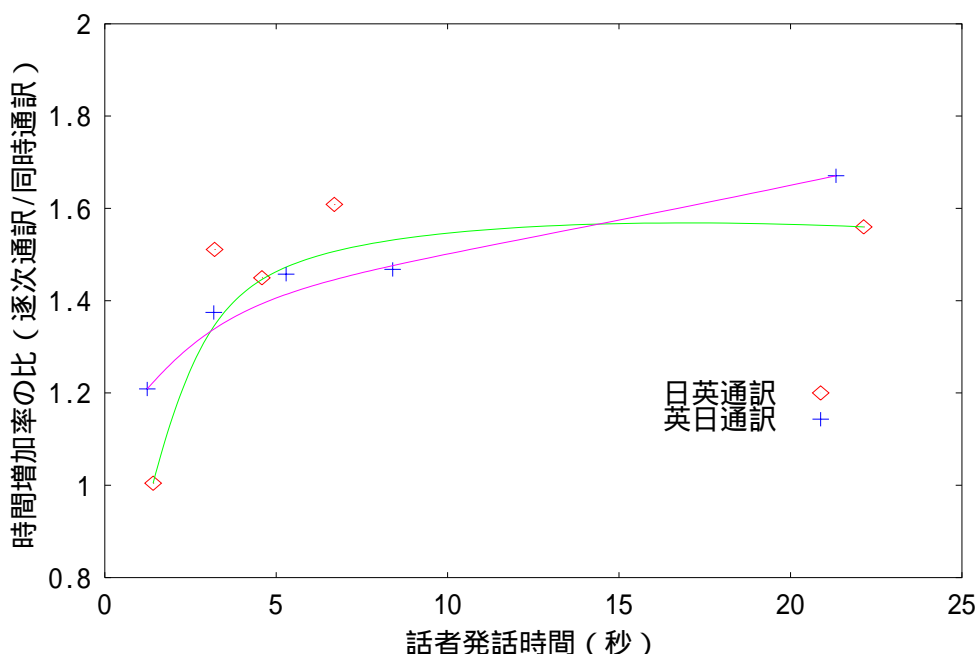


図 5 話者発話時間と時間増加率の比の関係

4.1.2 対話時間による比較

次に、同時通訳を介した対話全体に対する効果を評価するため、分析対象としたすべての対話について、話者発話時間の合計および話者・通訳者発話時間の合計を調べた。両者の関係を図 6 に示す。同時通訳を介した 216 対話に対する値、および逐次通訳を介した 12 対話に対する値をそれぞれ点で、また、通訳を介さない場合を実線で示している。同時通訳を介した対話のほとんどが逐次通訳に対する点と直線の間分布しており、これは逐次通訳を介するよりも効率よく対話を遂行できることを示している。実際、話者発話時間の合計と話者・通訳者時間合計の比は、216 対話の平均で 1.38 であり、逐次通訳の場合における 2.03 に比べ、大幅に短縮する。

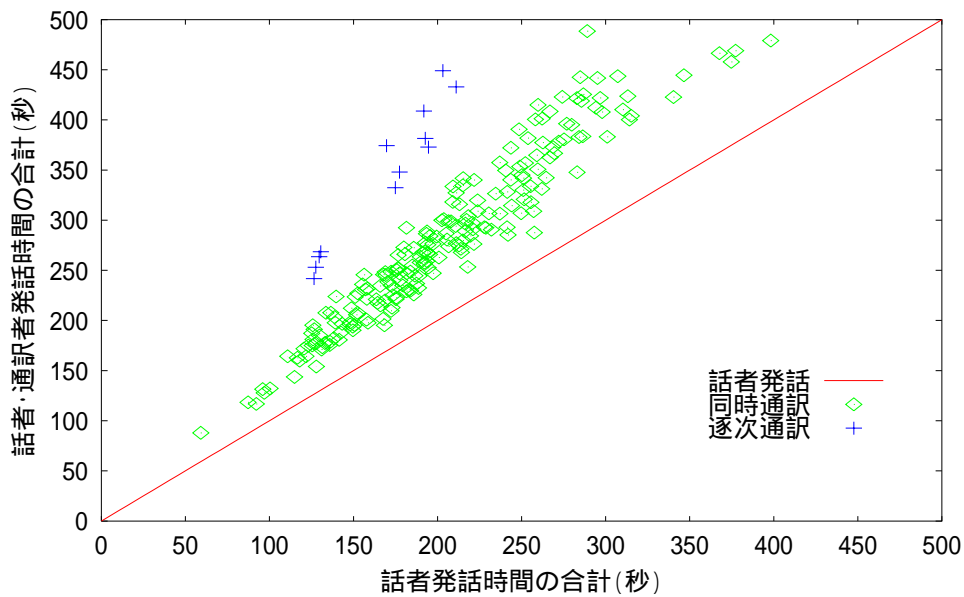


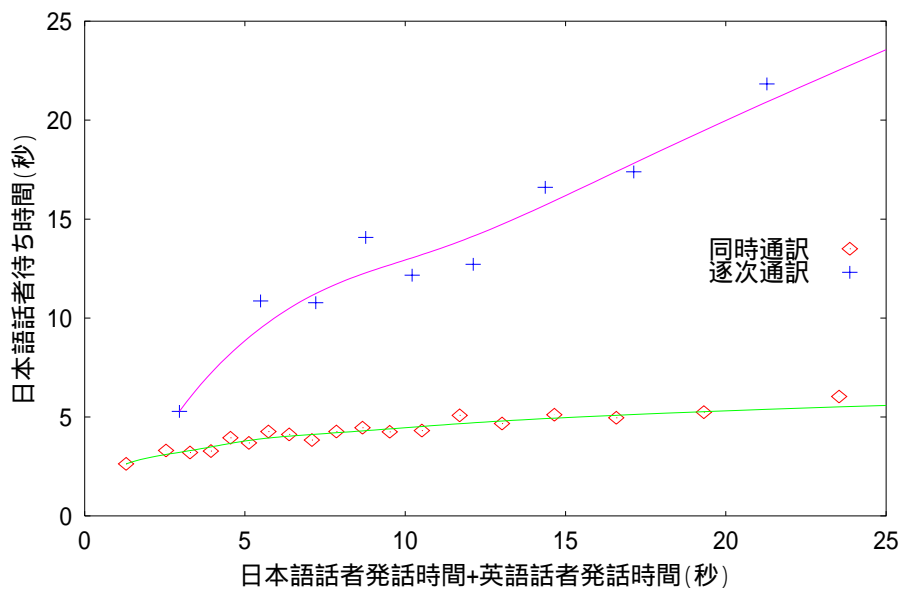
図 6 対話における話者発話時間と話者・通訳者発話時間の関係

4.2 対話の円滑さに関する分析

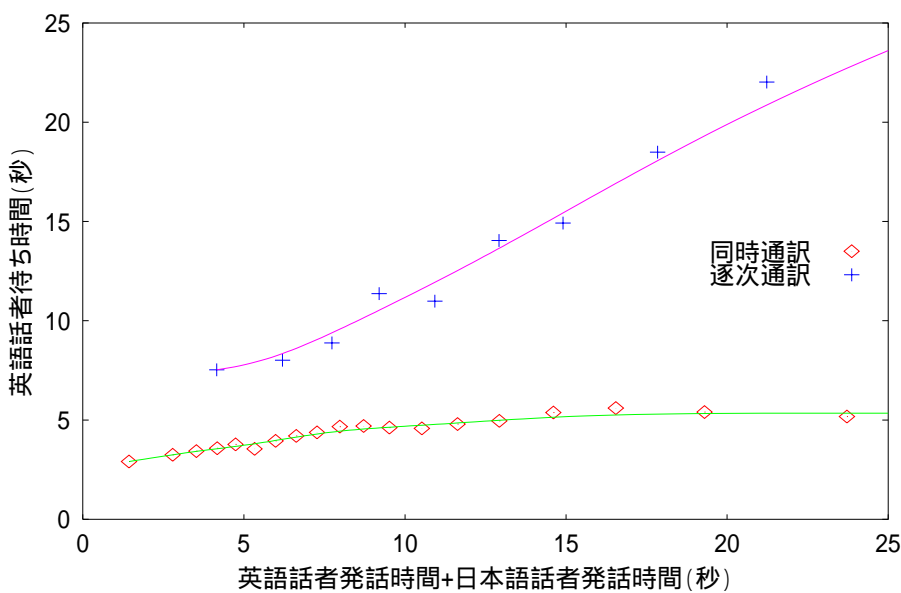
同時通訳を介した異言語間対話の円滑さを評価するために、英語話者および日本語話者の待ち時間を調査した。216 対話の英語話者ターン 4330 個のうち、英語話者に待ち時間が生じたのは、全部で 3860 ターンであり、待ち時間の平均は 4.4 秒であった。待ち時間が生じなかったターンの主な理由は、日本語話者の割り込み発話によるものであり、英語話者発話と日英通訳者発話が時間的に重なるケースが見られた。一方、逐次通訳では、英語話者の待ち時間の平均は 15.4 秒であった。同様に、日本語話者ターン 4233 個のうち、日本語話者に待ち時間が生じたのは、全部で 4064 ターンであり、待ち時間の平均は 4.3 秒であった。一方、逐次通訳では、日本語話者の待ち時間の平均は 14.5 秒であった。

異言語間対話における英語話者および日本語話者の待ち時間を、逐次通訳を介した場合と同時通訳を介した場合との間で比較した。結果を図 7(a)、(b)に示す。図では、話者による発話時間の長さが待ち時間の長さに及ぼす影響を考慮し、両者の間の関係をグラフ化した。すなわち、英語話者に対しては、英語話者から日本語話者へと発話権が移った区間における両話者の発話時間の合計をターンごとに計算し、それらを時間ごとにソートし等分割した上で、それぞれ待ち時間の平均値をもとめた。日本語話者に対しては、日本語話者から英語話者へと発話権が移った区間に対して同様に実施した。なお、分割は、同時通訳データについては 20 等分、逐次通訳データについては 10 等分とした。

同時通訳においては、話者発話時間が大きくなるにつれ増加の傾向を示すものの、増加の傾きは逐次通訳に比べてなだらかである。また、話者発話時間の長さにかかわらず、同時通訳における待ち時間の平均は逐次通訳のそれより短縮される傾向にあり、円滑さを備えた異言語間対話の遂行のために、同時通訳を採用する効果が示された。



(a) 逐次通訳/同時通訳を介した対話の日本語話者待ち時間



(b) 逐次通訳/同時通訳を介した対話の英語話者待ち時間

図 7 逐次通訳/同時通訳を介した対話の話者待ち時間

話者発話時間と待ち時間の比(逐次通訳/同時通訳)の関係を図 8 に示す。逐次通訳において待ち時間が長くなるるときとは、英日通訳者発声、ならびに、日本語話者発声の時間が長いときであるが(図 1 および図 7 参照)、その程度が大きいほど、同時通訳を介することにより待ち時間の増加が抑制される。実際、日・英語話者発話の合計が 5 秒程度の発話では、逐次通訳を介した場合の待ち時間の平均は同時通訳の場合の 2.2~2.3 倍程度であるが(図 8)、話者発話時間が長くなるにつれ倍率は増加の傾向を示す。

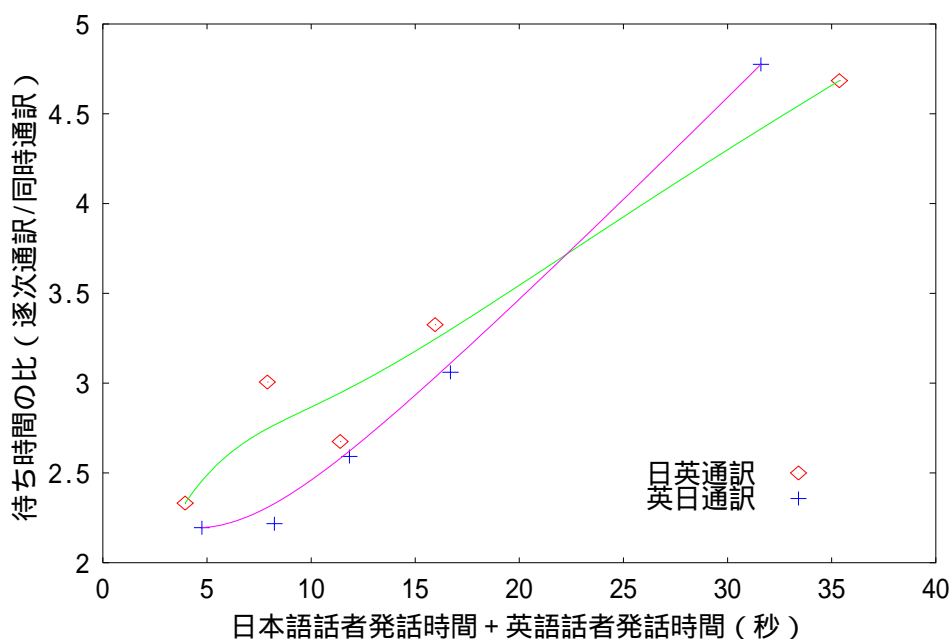


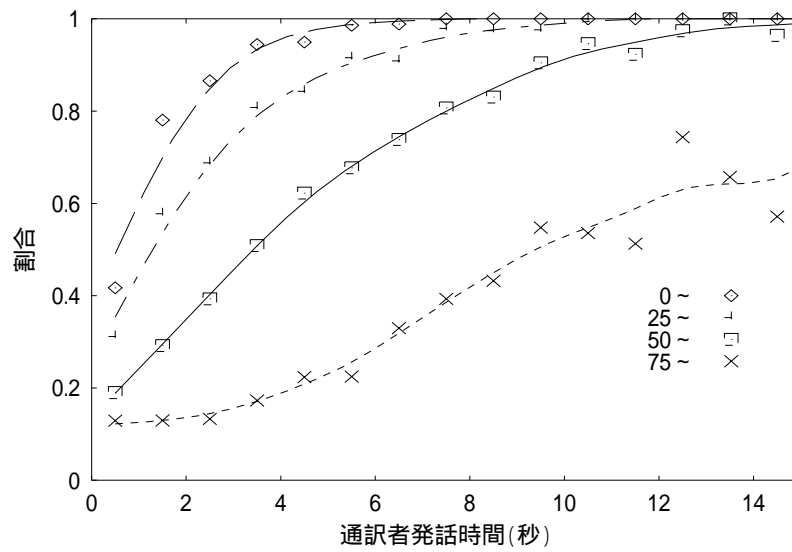
図 8 逐次通訳/同時通訳を介した対話における話者の待ち時間の比

5. 発話の重なり分析

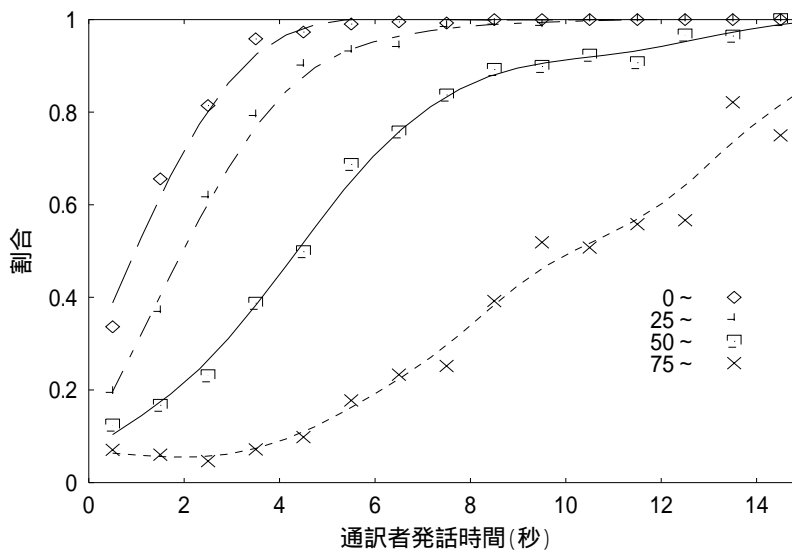
逐次通訳を介した対話と同時通訳を介した対話の最も大きな違いの一つは、話者発話とそれに対応する通訳者発話が時間的に重なるか否かであるが、同時通訳を介した場合でも重なりのあるときとないときがある。また、発話によって大きく重なったり、あまり重ならない場合が存在する。我々は、その違いには、話者発話、および通訳者発話の時間的要因が大きく寄与すると考え、同時通訳コーパスの 216 対話に含まれる 4233 の日英対訳対応発話および 4330 の英日対訳対応発話を用いて、話者、および通訳者発話時間と重なりの有無との関係について調べた。

5.1 通訳者発話時間と発話の重なり関係

通訳者が発声に長い時間を要するような場合には、自らの短期記憶容量の制限の観点からみて、通訳者は話者発話に重ねて発声せざるを得ないと予想される。そこで、通訳者の発話時間と時間的重なり度合との関係を調べた。通訳者発話時間は、発話単位間に存在するポーズの時間を省いて算出した。重なりのある話者・通訳者発話の割合を図 9 に示す。ただし、重なり程度として 4 段階の条件(重なりあり、25%以上、50%以上、75%以上)を設定し、各条件を満たす発話の割合を、通訳者発話をその時間長に従って 1 秒間隔で分類したそれぞれの平均によってグラフ上に記した。



(a) 日英通訳



(b) 英日通訳

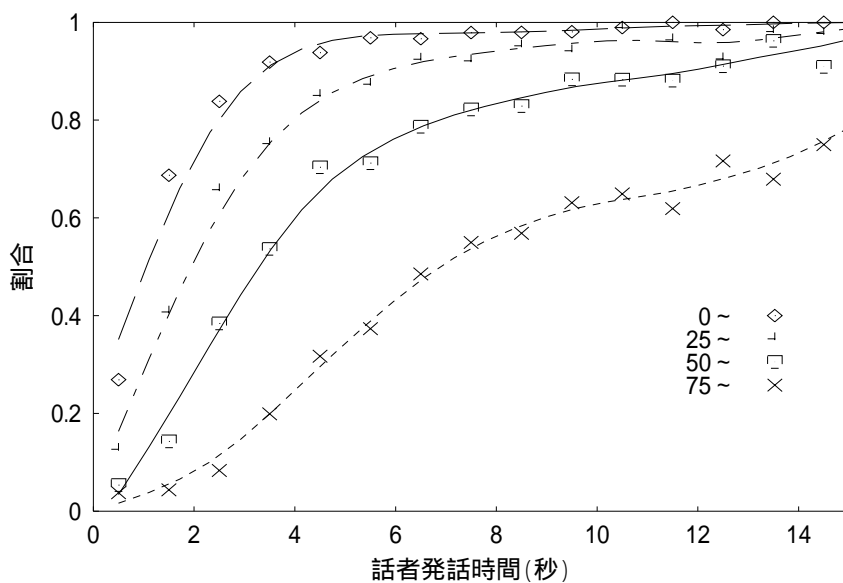
図 9 通訳者発話時間と発話の重なり関係

図 9 から、重なりが無い話者・通訳者発話の多くが、通訳者の発話時間が短い場合であることが分かる。実際、日英通訳者発話時間が 2 秒以下であるものが、重なりの無い話者・通訳者発話 731 例の約 79%を占めている。

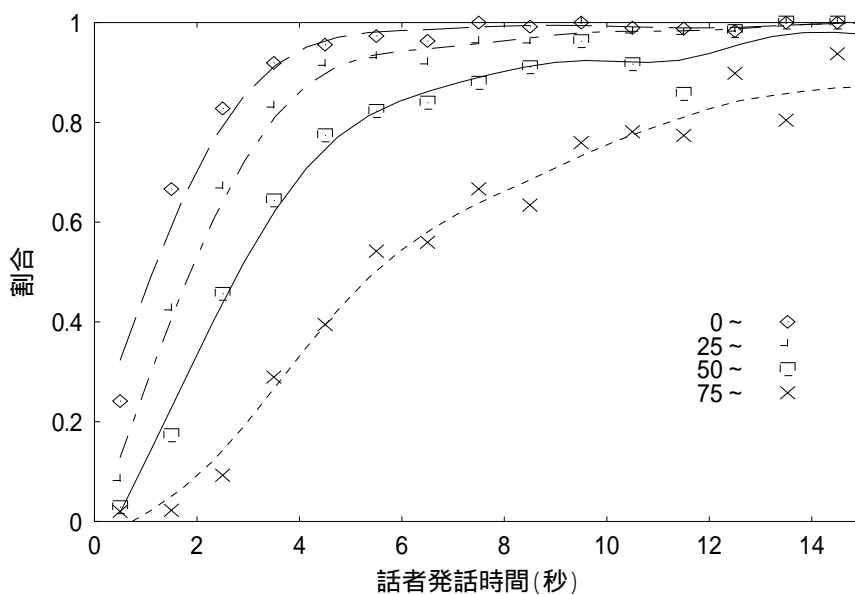
5.2 話者発話時間と発話の重なり関係

英語・日本語間対話における英語話者の待ち時間とは、英語話者の発声終了時から日英通訳者の発声開始時までを指すが(図 1 参照)、待ち時間を減少させる上で、日英通訳者の発声開始タイミングが大きなポイントとなる。逐次通訳の場合におけるそのタイミングは、

日本語話者発話の終了時以降となるが、同時通訳では、日本語話者の発声開始後、早いタイミングで訳出を開始し、日本語話者発話に重ねて発声することにより、待ち時間を減らせる可能性がある。特に、話者発話時間が長いときほど、発話の重なるの効果が大きいと予想される。そこで、同時通訳コーパスの日英、英日の双方の話者・通訳者発話を用いて話者発話時間と発話の重なるの有無との関係を調べた。重なりのある話者・通訳者発話の割合を図 10 に示す。図 9 と同様、話者発話をその時間長に従って一秒間隔で分類し、それぞれの平均をグラフ上に記した。



(a) 日英通訳



(b) 英日通訳

図 10 話者発話時間と発話の重なるの関係

図 10 から、話者の発話時間が長いほど、実際に通訳者による発声重なりが生じることが確認できた。実際、3 秒を超える日英話者発話の 98%以上に対して、通訳者発話の重なりが起こっている。逆に、重なりの無いターン 731 例の約 87%は、発話時間が 3 秒以下であるものであった。時間が短い話者発話に対しては、たとえ通訳者が重ねて発声しなかったとしても、結局は、話者発話開始後まもなく通訳者発話を開始できるため、待ち時間の増減に大きな影響はない。

6. まとめ

本論文では、通訳を介した異言語間対話の効率、ならびに円滑さに着目し、同時通訳を導入することの効果について、実際の同時通訳音声データを分析することにより考察した。分析には、英語・日本語間対話音声を収録した名古屋大学 CIAIR 同時通訳コーパスを用い、216 対話における日英および英日対訳対応付き発話データを対象とした。その結果、同時通訳を介した対話では、

- 対話時間増加率は、同時通訳で 1.38、逐次通訳で 2.03 であり、逐次通訳を介する対話に比べ、同時通訳を介する対話の効率は大幅に上がる。また、5 秒を超える日本語発話に対して通訳を行ったときの発話時間増加率は、同時通訳で平均して 1.27 であり、逐次通訳の 2.01 に比べかなり小さい。これは、英日通訳においても同様である。
- 通訳を介した対話におけるターンごとの話者待ち時間の平均は、英語話者の場合同時通訳で 4.4 秒、逐次通訳で 15.4 秒、日本語話者の場合、同時通訳で 4.3 秒、逐次通訳で 14.5 秒であり、同時通訳を介することにより対話の円滑さが大幅に向上する。
- 話者発話時間、もしくは、通訳者発話時間が 3 秒を超える対訳対応発話のほとんどが重なり発話となる。

ことが明らかになり、異言語間対話支援環境としての同時通訳技術の有用性を確認した。

本研究で実施した調査は、通訳対象となる言語(英日通訳、日英通訳など)、対話のタスクドメイン(旅行対話、会談対話、インタビュー対話など)、さらには、音声のタイプ(対話と独話)などデータの性質によって結果が異なると予想される。名古屋大学 CIAIR では、講演の同時通訳など、さまざまなタイプの音声データを収録しているが、これらのデータを用いて本論文と同様の分析を実施し、本研究の結果と比較することは、通訳環境に適応可能な同時通訳システムを実現する上で有用なデータになると予想される。

謝辞: 対訳対応データの作成にご協力いただいた名古屋大学稲垣研究室の諸氏に感謝致します。(株)インターグループには、データの収録にご協力いただきました。ご尽力下さいました横尾祐郁氏に感謝致します。本研究の一部は、文部科学省科学研究費補助金 COE 形成基礎研究費(課題番号 11CE2005「多元音響の統合的理解」代表:名古屋大学 板倉文忠 教授)、ならびに(財)中島記念国際交流財団研究助成によります。

著者紹介:大原 誠 (Ohara, Makoto)名古屋大学工学部卒業。現在、名古屋大学大学院工学研究科博士前期課程在学中。機械翻訳、文書処理の研究に従事。連絡先:〒464-8603 名古屋市千種区不老町 (Email: makoto@inagaki.nuie.nagoya-u.ac.jp)

松原 茂樹 (Matsubara, Shigeki) 名古屋大学大学院工学研究科博士後期課程修了。工学博士。現在、名古屋大学情報連携基盤センター助教授。ATR 音声言語コミュニケーション研究所客員研究員。自然言語処理、音声言語処理、機械翻訳の研究に従事。

笠 浩一郎 (Ryu, Koichiro) 名古屋大学工学部卒業。現在、名古屋大学大学院工学研究科博士前期課程在学中。機械翻訳の研究に従事。

河口 信夫 (Kawaguchi, Nobuo) 名古屋大学大学院工学研究科博士後期課程満了。工学博士。現在、名古屋大学情報連携基盤センター助教授。音声言語インタフェース、モバイルコミュニケーションの研究に従事。

稲垣 康善 (Inagaki, Yasuyoshi) 名古屋大学大学院工学研究科博士後期課程修了。工学博士。名古屋大学名誉教授。現在、愛知県立大学情報科学部教授。オートマトン・言語理論、計算論、ソフトウェア基礎論、人工知能、自然言語処理の研究に従事。

【註】

- 1) 現在、独話同時通訳の大規模データの収集が進められており (Kashioka, 2002; Matsubara, 2002)、その詳細な分析を通して、独話同時通訳システムの実現可能性が検討されている (Aizawa, 2000; 柏岡 2001; 高木 2002)。
- 2) 「一区切り」の定義は、話者および通訳者の発話スタイルや場面によって異なる。逐次通訳を介した対話では、「一区切り」の単位として対話ターンが採用される場合が多く、本論文でも原則として対話ターンを訳出単位とした逐次通訳データを使用する。
- 3) 図4における各プロットごとに時間増加率の分散をもとめ、通訳形態ごとに平均値を計算した。英日通訳の場合では同時通訳が 0.98、逐次通訳が 0.29、また、日英通訳の場合では同時通訳が 4.55、逐次通訳が 0.76 であった。いずれも逐次通訳で分散が小さく、このことから逐次通訳データのばらつきの少なさが示される。
- 4) CIAIR = the Center for Integrated Acoustic Information Research of Nagoya University (名古屋大学統合音響情報研究拠点)。プロジェクトの詳細については <http://www.ciair.coe.nagoya-u.ac.jp/jpn/> 参照。[編集部註]

【参考文献】

Aizawa, Y., Matsubara, S., Kawaguchi, N., Toyama, K., and Inagaki, Y. (2000). Spoken Language Corpus for Machine Interpretation Research. *Proc. 6th International Conference on Spoken Language Processing*, Vol. III: 398-401.

- Amtrup, J. (1999). Incremental Speech Translation. *Lecture Notes in Artificial Intelligence*, Vol. 1735
- Gile, D. (2001). Consecutive vs. Simultaneous: which is more accurate? *Interpretation Studies*, No. 1: 8-20
- Kashioka, H. (2002). Translation Unit Concerning Timing of Simultaneous Translation. *Proc. 3rd International Conference on Language Resources and Evaluation*, Vol. I: 142-146
- Kawaguchi, N., Matsubara, S., Takeda, K., and Itakura, F. (2002). Multi-Dimensional Data Acquisition for Integrated Acoustic Information Research. *Proc. 3rd International Conference on Language Resources and Evaluation Conference*, 2043-2046.
- Matsubara, S., Toyama, K., and Inagaki, Y. (1999). Sync/Trans : Simultaneous Machine Interpretation between English and Japanese. N. Foo (Ed.): *Advanced Topics in Artificial Intelligence, Lecture Notes in Artificial Intelligence*, Vol. 1747: 134-143
- Matsubara, S., Takagi, A., Kawaguchi, N., and Inagaki, Y. (2002). Bilingual Spoken Monologue Corpus for Simultaneous Machine Interpretation Research. *Proc. 3rd International Conference on Language Resources and Evaluation*, Vol. I: 153-159
- Mima, H., Iida, H., and Furuse, O. (1998). Simultaneous Interpretation Utilizing Example-based Incremental Transfer. *Proc. 17th International Conference on Computational Linguistics and 36th Annual Meeting of the Association for Computational Linguistics*. 855-861
- Watanabe, T., Okumura, A., Sakai, S., Yamabana, K., Doi, S., and Hanazawa, K. (2000). An Automatic Interpretation System for Travel Conversation. *Proc. 6th International Conference on Spoken Language Processing*, Vol. IV: 444-447
- 柏岡秀紀・田中英輝(2001)「講演の同時通訳データの分析」『言語処理学会第7回年次大会発表論文集』433-436
- 高木亮・松原茂樹・稲垣康善(2002)「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」『言語処理学会第8回年次大会発表論文集』383-386
- 菅谷史昭・竹澤寿幸・隅田英一郎・匂坂芳典・山本誠一(2002)「音声翻訳システム: ATR-MATRIXの開発と評価」『情報処理学会論文誌』第43巻第7号:2230-2241
- 前川喜久男・籠宮隆之・小磯花絵・小椋秀樹・菊池英明(2002)「日本語話し言葉コーパスの設計」『音声研究』第4巻第2号:51-61
- 松原茂樹(2001a)「同時通訳システムの実現を目指して-自動通訳技術の現状と今後の課題-」『日本通訳学会2001年度第4回特別例会資料』
<http://ux01.so-net.ne.jp/%7Ea-mizuno/meeting1.html>
- 松原茂樹・相澤靖之・河口信夫・外山勝彦・稲垣康善(2001b)「同時通訳コーパスの設計と構築」『通訳研究』第1号:85-102
- 松本裕治・北内啓・山下達雄・平野善隆・松田寛・高岡一馬・浅原正幸(2000)「日本語形態素解析システム 茶釜 version 2.2.0」『奈良先端科学技術大学院大学松本研究室』