

同時通訳データの分析に基づく
漸進的な英日翻訳に関する研究

笠 浩一朗

概要

近年、グローバル化の進展により、国際会議や国際シンポジウムが盛んに開催されている。また、海外に拠点を置く会社も増えており、国際的なビジネスも盛んに行われている。このような異言語間コミュニケーションを必要とする場面を支援する情報技術として、音声翻訳システムが注目されており、その開発が進められている。

欧米では、TC-STAR や GALE と呼ばれる音声翻訳システムの開発を目的とした大規模なプロジェクトが進められている。国内でも、成田空港や関西国際空港で音声翻訳システムの被験者実験が実施されている。さらに、音声翻訳に関する国際ワークショップが毎年開催されており、音声翻訳システムの翻訳性能を競うコンテストが実施されるなど、システムの実用化に向けた動きが加速しつつある。

実際に市販された音声翻訳システムとしては、音声翻訳機能を搭載したモバイルマルチメディアプレーヤ「VoToL」、中国語と英語との会話を自動翻訳できる携帯電話サービス「JAJAH Babel」、英語とスペイン語の双方向の音声翻訳に対応した iPhone アプリ「Jibbiggo」などが挙げられる。

しかし、現在の音声翻訳システムの多くは、話者の発話が終了してから訳出を開始する逐次通訳である。逐次通訳は、講演など一人の話者が話し続ける独話を対象とした場合、話者が発話を休止することがないため利用することができない。また、二人以上の話者が交代で話す対話を対象とした場合、話し手が話してから聞き手に伝わる間に時間差が生まれ、コミュニケーションの円滑さを妨げることになる。これらの問題を解消するために、話者の発話に追従して訳出する同時通訳システムの開発が進められている。

同時通訳は人間の言語活動において最も高度であるとされており、システムを実現するための有効な方法については、これまでほとんど明らかにされていない。このような現状を打破するために、プロの同時通訳者の振る舞いを観察し、得られた知見を活用することは一つの方法である。

そこで本研究では、大規模な同時通訳データを用いて実際の同時通訳者の訳出方略に関する知見を獲得し、得られた知見をもとに対話を対象とする英日同時通訳シ

システムを開発した。さらに、独話を対象とした同時通訳システムの開発を目的に、同時通訳データベースの講演データを用いて、同時通訳者の話速に関する定量的な分析を与えた。

本論文は、全5章から構成される。第1章は本論文の序論であり、同時通訳システムの開発に関する研究の動向を示すとともに、本論文の位置づけとアプローチを述べたものである。

第2章では、同時通訳者の訳出方略を獲得するための基盤データとなる同時通訳データベースの設計、構築、及び、利用について詳述する。本データベースの構築では、多様な側面から同時通訳者の発話を分析するために、対話と独話の英日双方向の通訳音声収録する。また、研究資源としての利便性を考慮し、音声データを文字化し、フィラーや言い直しなどの非文法的な言語現象に談話タグを付与する。さらに、発話単位での時間情報や対訳対応関係などの言語情報を人手によりアノテーションする。

第3章では、英日同時翻訳のための訳文生成手法を提案する。英語と日本語のように、大きく異なる性質を持つ言語間の同時翻訳では、対応する語の生起順序が異なるため、標準的な訳文を生成したのでは同時性の高い翻訳を実現することは困難である。例えば、英語はSVO構文であるのに対し、日本語はSOV構文であるため、英語話者発話のOの部分の発話が完了するまでVの訳出を待たなければならない。他にも、英語では時間を表す副詞が一般に文末に生起するのに対して、日本語では文頭に生起するため、発話が完了するまで訳出を開始できない。一方、同時通訳データベースを観察すると、英日同時通訳者は、標準的な訳文ではなく、原文の語順に近い語順の日本語文を訳文とすることにより、早い段階での訳出を可能にしていることがわかる。提案手法では、そのような訳出方法を参考に、日本語として容認可能な範囲で、原文の語順に近い語順の日本語訳を入力と同時的に生成する。本論文では、日本語として容認可能な文が満たすべき性質を、依存構造の観点から検討する。また、そのような性質を満たす訳文を生成する手法について述べる。提案手法では、日本語の語順に関する制約は比較的緩やかであるという特徴を利用し、日本語文として容認可能な範囲内で、原文の語順に準拠した訳文を生成する。訳文としての容認可能性は、日本語依存文法に基づいて検査する。すなわち、依存文法に基づく制約を満たし、かつ、英文の語順に最も近い日本語文を同時翻訳における訳文とする。ただし、日本語文として不自然でない場合に限り、倒置、すなわち、後方修飾性を満たさない依存関係を含む訳文についても生成を許容する。提案手法を実装したシステムを用いて翻訳実験を行い、翻訳品質、及び、同時性の観点から本手

法を評価した。その結果、容認可能な翻訳品質をもち、かつ、入力に対して同時性が高い英日同時翻訳を実現するために、提案手法が有効であることを確認している。

第4章では、英日同時通訳における通訳者の話速に関する定量的分析について述べる。本分析では、特に、同時通訳者の話速に影響を与えられいくつかの要因に焦点をあて、その影響の程度について分析を与える。具体的には、大局的な要因として「講演の話速」及び「通訳者の個性」、また、局所的な要因として、講演中の「話し手の発話状態」及び「訳出の進行状況」に着目し、話速データを比較することによりそれらの要因が及ぼす影響の程度を明らかにする。定量的分析のために、本研究では、同時通訳データベースの英日通訳データを用いた。分析では、このデータベースの中から22の英語講演データに対する複数の通訳者による88例の英日同時通訳データを抽出し、使用した。また、音声認識ツールを用いて発声時刻情報を自動推定することにより、単語レベルでの話速データを作成し、分析に使用した。分析の結果、講演者の平均話速と通訳者の平均話速の間には相関関係はなく、また、講演の違いに比べ通訳者による話速の違いの方が大きいことを確認した。さらに、原発話が確定前より確定後の方が通訳者の話速が速く、講演者の発話が休止中の方が発話中に比べて通訳者の話速が速いことを確認した。

第5章では、本研究についてまとめる。最後に今後の課題として、同時通訳システムの開発のための研究課題と将来の展望について述べる。

目次

第1章	まえがき	15
1.1	音声翻訳システム	15
1.2	音声翻訳に関する研究動向	17
1.2.1	話し言葉処理の頑健性を備えた音声翻訳	17
1.2.2	非言語情報を用いて曖昧性を解消する音声翻訳	18
1.2.3	多言語を対象とした音声翻訳	18
1.2.4	音声翻訳における翻訳対象の拡張	19
1.2.5	訳出の同時性を備えた音声翻訳	19
1.3	本論文の目的	20
1.4	本論文の内容と構成	21
第2章	同時通訳データベースの構築	25
2.1	はじめに	25
2.2	データベースの設計	26
2.3	音声データの収録	26
2.3.1	独話音声の収録	26
2.3.2	対話音声の収録	27
2.4	データベースの構築	29
2.4.1	音声データの文字化	29
2.4.2	音声データの視覚化	31
2.4.3	対訳対応データの作成	32
2.4.4	収録環境情報の作成	34
2.5	データベースの構成と規模	34
2.6	データベースの有用性	37
2.7	まとめ	39
第3章	英日同時翻訳のための依存構造に基づく訳文生成手法	41
3.1	はじめに	41

3.2	異言語間対話における同時通訳の効果	42
3.3	英日同時通訳者の訳出	44
3.3.1	英日間の語順の違いによる問題	45
3.3.2	同時通訳者の訳出方法	45
3.4	訳文生成手法	46
3.4.1	容認可能な日本語文とその依存構造	47
3.4.2	アルゴリズム	51
3.5	システム構成	55
3.5.1	解析部	55
3.5.2	変換部	58
3.5.3	生成部	58
3.5.4	同時翻訳処理の流れ	58
3.6	同時翻訳システムによる翻訳実験と評価	58
3.6.1	翻訳品質と同時性の評価指標	60
3.6.2	実験環境と実験データ	61
3.6.3	実験結果	62
3.7	まとめ	68
第4章	同時翻訳システムにおける話速制御のための通訳者の発話速度の分析	69
4.1	はじめに	69
4.2	同時通訳者の話速に影響を与える要因	70
4.3	分析データ	71
4.3.1	単語発声時刻の推定	72
4.3.2	話速の計算	76
4.3.3	話速変動の視覚化	76
4.4	講演の設定に関わる要因が同時通訳者の話速に及ぼす影響	77
4.4.1	講演者の話速と通訳者の話速との関係	77
4.4.2	通訳者の個性と通訳者の話速との関係	78
4.5	講演の進展により変化する要因が通訳者の話速に及ぼす影響	81
4.5.1	原発話の進行状況と話速変動の関係	81
4.5.2	講演者の発話状態と話速変動の関係	82
4.6	まとめ	83

	9
第5章 あとがき	85
5.1 本論文のまとめ	85
5.2 今後の課題と将来への展望	86

図一覽

2.1	収録の様子	27
2.2	独話の収録の様子	28
2.3	対話の収録の様子	29
2.4	対話設定シート	30
2.5	文字化データ(独話)	31
2.6	文字化データ(対話)	32
2.7	視覚化ツール(独話)	34
2.8	視覚化ツール(対話)	35
2.9	対訳対応ツール	36
2.10	収録環境情報ファイル	37
3.1	通訳者を介した異言語間対話の模式図	44
3.2	英語文(E1)に対する同時通訳者の訳出例	46
3.3	英語文(E2)に対する同時通訳者の訳出例	47
3.4	標準的な日本語訳(J1)と同時通訳者による日本語訳(S1)の依存構造	48
3.5	標準的な日本語訳(J2)と同時通訳者による日本語訳(S2)の依存構造	50
3.6	本手法のアルゴリズム	52
3.7	出力処理の例	53
3.8	システムの構成図	56
3.9	英文(E3)の“Denver”が入力された時点の処理($L=2$)	57
3.10	英文(E3)の”I want to fly from San Francisco”が入力されるまでの 処理の流れ	59
3.11	英文(E3)の”to Denver”から文末まで入力されたときの処理の流れ	60
3.12	文の長さの分布	62
3.13	文の長さとの翻訳精度の関係	64
3.14	文の複雑さと翻訳精度の関係	65
3.15	話者発話時間と話者発話終了から訳出終了までの時間の関係	67

4.1	単語発声時刻の推定（日本語）	73
4.2	単語発声時刻の推定（英語）	74
4.3	単語発声時刻データのサンプル	75
4.4	話速の視覚化ツール	77
4.5	通訳者の話速の分布	78
4.6	講演者の話速と通訳者の話速の関係	79
4.7	通訳者と通訳者の話速の分布	80
4.8	原発話の進行状況（原発話の確定前の通訳者発話と確定後の通訳者発話）	82
4.9	講演者の発話状態（発話中と休止中）	83

表一覽

2.1	同時通訳データベースの収録様式	26
2.2	文字範囲の特徴を表現する談話タグ	33
2.3	音や非言語的なイベントを表現する談話タグ	33
2.4	データベースの規模 (独話)	38
2.5	データベースの規模 (対話)	38
3.1	逐次通訳データの基礎統計	45
3.2	係り文節の数と倒置の起こる割合	49
3.3	評価基準	61
3.4	翻訳品質の実験結果	63
3.5	翻訳誤りの原因	66
3.6	同時性の評価	66
3.7	述部の訳出制御の設定を $L = 1, 2, 3$ にした場合の実験結果	67
4.1	通訳者の属性	71
4.2	分析データの規模	71
4.3	単語発声時刻の推定精度	75
4.4	英語講演者の話速の基礎統計	76
4.5	英日同時通訳者の話速の基礎統計	76
4.6	原発話の進行状況と通訳者の話速	82
4.7	講演者の発話状態と通訳者の話速	83
4.8	講演者の発話状態 (速い場合と遅い場合) と通訳者の話速	83

第1章 まえがき

1.1 音声翻訳システム

近年，交通や通信手段の進歩とそれに伴うグローバル化の進展により，技術や文化の国際的な交流を目的とした国際会議やシンポジウムが盛んに開催されている．また，海外に拠点を置く会社も増えており，国際的なビジネスも盛んに行われている．このような異言語間コミュニケーションを必要とする場面を支援する情報技術として，音声翻訳システムが注目されており，その開発が進められている．

音声翻訳システムは，音声認識，言語翻訳，音声合成という3つのモジュールにより構成されており，音声認識により話者の発話を認識し，言語翻訳により認識した原言語を目的言語に翻訳し，音声合成により翻訳結果を音声に変換するものである．

音声・言語処理技術の進展に伴い音声翻訳システムの研究は盛んに行われている．欧州と米国では，それぞれ TC-STAR[18, 32] と GALE[56] と呼ばれる音声翻訳システムの開発を目的とした大規模なプロジェクトが進められている．国内でも，日本電気 (NEC) が，2004 年に成田空港で英語から日本語への翻訳が可能な携帯端末の評価実験を，国際電気通信基礎技術研究所 (ATR) が，2005 年に関西国際空港で日英・日中双方向の音声翻訳機能を備えた携帯端末の被験者実験を [25]，情報通信研究機構 (NICT) が，2008 年に北京五輪の旅行者を対象に携帯電話を使った音声翻訳システムのモニター実験をそれぞれ実施している [13]．さらに，音声翻訳に関する国際ワークショップ (IWSLT: International Workshop on Spoken Language Translation) が毎年開催されており，音声翻訳システムの翻訳性能を競うコンテストが実施されるなど，システムの実用化に向けた動きが加速しつつある．

実際に市販された音声翻訳システムとして，NEC が，2006 年に音声翻訳機能を搭載したモバイルマルチメディアプレーヤ「VoToL」を販売した [3]．また，米国の JAJAH が，2008 年に IBM と共同で開発した中国語と英語との会話を自動翻訳できる携帯電話サービス「JAJAH Babel」を開始した [2]．さらに，InterACT (International Center for Advanced Communication Technologies)，カーネギーメロン大学，カールスルーエ技術研究所が共同で，2009 年 10 月に英語とスペイン語の双方向の音声

翻訳に対応した iPhone アプリ「Jibbiggo」を公開している [1] .

音声翻訳システムの実用化に至るまでには、その要素技術である音声認識、言語翻訳、音声合成の性能をそれぞれ向上させる必要がある。音声認識、及び、音声合成は、主に音声情報処理技術による解決が求められる。一方、音声翻訳システムにおける言語翻訳処理は、話し言葉の原言語テキストを入力とし、話し言葉の目的言語テキストを出力とするため、主に自然言語処理技術による解決が求められる。ここでは、自然言語処理の観点に立ち、言語処理に関する研究動向に焦点を当てる。

音声翻訳システムにおける言語翻訳処理に関する研究には、以下に挙げるように多くの課題が残されている。

(1) 話し言葉処理の頑健性に関する課題

音声翻訳では、入力話し言葉であるため、フィラー、言い直し、言い淀み、言い誤りなどの非文法的な現象を含んでいる。そのため、新聞や書物などの書き言葉を対象とした翻訳技術では、これらの言語現象を含む文を上手く翻訳することができない。また、自動音声認識によって文字化された原文を翻訳対象とするため、音声認識誤りへの対処も求められる。例えば、認識結果の一部に誤りが含まれていても、正しく認識された部分だけから、訳文を生成するなどの処理が求められる。

(2) 非言語情報を用いた曖昧性解消に関する課題

実際の通訳者は、話者の発話内容のみの情報を用いて翻訳するわけではなく、音響的な情報や話者の動作や表情などの非言語情報を翻訳結果に反映している。これらの情報は、書き言葉には存在しない情報であるため、一般的な機械翻訳技術では扱うことができない。

(3) 多言語化に関する課題

近年の翻訳技術の主流は、大規模な対訳データを用いたコーパスベースの手法である。そのため、既存の音声翻訳システムの多くは、大規模な対訳データが存在する特定の言語間を翻訳対象としている。ユーザの利便性を考慮すると、あらゆる言語間に対応した音声翻訳システムの開発が求められる。

(4) 翻訳対象の拡張性に関する課題

これまでの音声翻訳システムの多くは、比較的短文が短く、文の構造が単純な旅行対話を翻訳対象としている。翻訳対象を旅行対話に限らずに、対象ドメインを広げることができれば、音声翻訳システムが活用される場面が増えると考えられる。

(5) 訳出の同時性に関する課題

音声翻訳システムの多くは、話者の発話が終了してから訳出を開始する逐次通訳である。一人の話者が話し続ける講演などの独話を対象とした場合、話者が発話途中で話を止めることはないため、逐次通訳は適さない。また、二人以上の話者が交代で話す対話を対象とした場合、話し手が話してから聞き手に伝わる間に時間差が生まれ、コミュニケーションの円滑さを妨げることになる。これらの問題を解消するために、話者の発話に追従して訳出する、すなわち、同時性を備えた音声翻訳システムの開発が求められる。

これらの課題を解決するため、これまでにいくつかの研究が行われており、次の1.2節では、これらの課題の観点から音声翻訳に関する研究動向を概観する。1.3節では、本論文の目的について述べ、1.4節では、本論文の内容と構成について述べる。

1.2 音声翻訳に関する研究動向

1.2.1 話し言葉処理の頑健性を備えた音声翻訳

90年代、音声翻訳システムの開発が急速に進展したものの、音声翻訳システムを実環境下で使用する場合、話者発話に含まれるフィラーや言い直しなどの非文法的な表現や雑音などによる音声認識誤りなどの原因により、翻訳性能が著しく低下するという問題があった。

非文法的な表現や音声認識誤りに頑健な翻訳システムとして、古瀬ら [15] は、非文法的な表現などで構文解析できない場合に、構文構造が解析スコアが全体として最適になるように入力文を分割し、分割した単位ごとに部分翻訳し、結合する手法を提案している。また、脇田ら [72] は、誤った音声認識結果のうち信頼性が高い部分のみを翻訳する手法を提案している。さらに、Leeら [33] は、文の重要な構成要素である名詞や動詞のみを翻訳することにより、頑健性を向上させる手法を提案している。

最近の音声翻訳システムは、話し言葉処理の頑健性を向上させるために、携帯電話やPDAなどの携帯端末に音声認識結果を表示し、ユーザが認識結果を確認できるようになっており、認識誤りがある場合には、再度、発話することによって音声認識誤りに頑健なシステムを実現している [20, 47]。

1.2.2 非言語情報を用いて曖昧性を解消する音声翻訳

実際の通訳者は、話者の発話内容だけでなく、発声の韻律情報、話者の身振りや表情などの非言語現象を用いて、発話内容の構文的な曖昧性を解消している。

非言語情報を用いた音声翻訳の研究としては、発声の韻律情報を用いた研究がある。例えば、Takezawaら [64] は、日本語の肯定文か疑問文かを文末の抑揚の違いにより判定する方法を音声翻訳システムに組み込んでいる。また、Wahlsterら [71] は、ドイツ語の疑問形と命令形の構文が同じであるため、抑揚の違いにより疑問形か命令形かを判定するとともに、語彙の曖昧性解消に抑揚の違いを用いている。これらの研究のように、韻律情報を用いることによる構文的、語彙的な曖昧性解消はいくつか行われており、音声翻訳システムに実際に組み込まれている。

1.2.3 多言語を対象とした音声翻訳

これまでの音声翻訳システムの多くは、特定の言語間の翻訳を実現するものであり、多言語の音声翻訳を実現するものはほとんどない。多言語翻訳の試みとしては、音声翻訳に関する国際コンソーシアム C-STAR に所属する4ヶ国の研究機関（日本のATR、ドイツのカールスルーエ大学、米国のカーネギーメロン大学、韓国の韓国電子通信研究所）が、1999年に合同で音声翻訳実験を実施した [73]。この実験では、各研究機関の音声翻訳システムが自国語の音声認識と音声合成を担当することで、多言語翻訳を実現した。近年では、アジア・太平洋電気通信標準化機関 (ASTAP) において音声翻訳を構成する機能モジュールをインターネットで相互接続するための標準化が進められており、NICT がアジア音声翻訳先端研究コンソーシアム (A-STAR) と共同で、2009年に日本語、中国語、韓国語、タイ語、インドネシア語、マレーシア語、ベトナム語、ヒンディー語、英語の8ヶ国語の多言語音声翻訳を実現したネットワーク型音声翻訳システムの開発を行っている [19]。

多言語への適用可能な翻訳手法に関する研究では、古瀬ら [14] が、構成素境界解析と用例利用型処理を組合せた言語に依存しない手法を提案している。また、現在の翻訳手法の主流は、大規模な対訳コーパスを用いたコーパスベースの翻訳手法 [6, 7] であり、対訳コーパスを構築できれば、言語に合わせて翻訳モデルの学習方法を調整する必要はあるものの、ほとんどの言語に適用可能である。

1.2.4 音声翻訳における翻訳対象の拡張

これまでの音声翻訳システム多くは、翻訳対象としてホテルや空港などの旅行対話を想定していた [20, 31, 34, 47, 49]。これは、利用者数や利用頻度などの観点から実用性が高いこと、旅行対話で話される発話文は比較的短く、簡単であり、使用される語彙も少ないため、高い翻訳精度が期待できることが理由として挙げられる。旅行対話以外を翻訳対象とした音声翻訳システムでは、病院の診察を対象としたシステムの開発が行われているものの [48]、旅行対話と同様に、使用される語彙が限定される環境での利用を想定したものである。

近年、旅行対話に比べ、文が長く、複雑で、語彙が豊富な講演を対象とした音声翻訳への試みがいくつかみられるものの [21, 29]、対話のように実環境で動作する技術レベルにない。

1.2.5 訳出の同時性を備えた音声翻訳

現存するほとんどの音声翻訳システムでは、逐次通訳に基づいているものの、逐次通訳は、一人の話者が話し続ける講演などの独話に対しては利用できない。また、対話に対しても、話し手が話してから聞き手に伝わる間に時間差が生まれ、対話の効率や円滑さが低下する。このため、同時性を備えた翻訳手法が必要である。

同時通訳を実現するために、解析や生成、分割などの要素技術に関する研究が行われている。以下にそれぞれの研究事例を概観する。

これまでに漸進的な構文解析手法が提案されており、英語では、文脈自由文法に基づくもの [24, 41, 54]、範疇文法に基づくもの [16, 44]、木接合文法に基づくもの [22] などがある。また、日本語では、文脈自由文法に基づくもの [4]、依存文法に基づくもの [51] などが提案されている。これらの漸進的構文解析の研究では、大規模な音声コーパスの構築とともに解析精度が向上しており、同時通訳の分野だけでなく、同時通訳と同じく漸進的な構文解析が求められる字幕生成などの分野でも利用が検討されている。

漸進的な生成に関する研究としては、語順が比較的近い言語間に対する同時通訳手法として提案されているものの [5, 8]、日本語と英語のように語順が大きく異なる言語間にも応用できるかは明らかでない。日本語・英語間に対する同時通訳手法として、日英翻訳においては、節ごとに変換するための規則及びその適用法が [45]、英日翻訳では、人手で作成した変換規則で語順の問題を解決する手法が提案されているものの [40]、英日間の語順の問題を解決するための変換規則を人手で網羅的に作

成することは容易ではないという問題がある。

同時通訳の分割に関する研究では，丸山ら [37] は，独話を対象とした日英同時通訳の処理単位として述部を中心とするまとまりである節に着目し，節境界単位を提案しており，節境界単位を自動的に検出するツールを開発している．西光ら [59] は，話し言葉音声を対象として，係り受けとポーズ・フィラーの情報に着目して，処理単位を検出する手法を提案している．清水ら [60, 61, 62] は，日本語の話し言葉の処理単位として，プロの同時通訳者が原言語から目標言語に変換する自然なタイミングである音声翻訳単位を提案しており，音韻情報や句読点情報を用いて自動的に検出している．加藤ら [22] は，日英同時通訳者の翻訳単位を同時通訳者のポーズをもとに検出している．これらの研究は独話を対象としている．独話を対象とした同時通訳の研究では，翻訳単位についての検討が進んでいるものの，訳出タイミングなど他の要素技術について検討したものはほとんどない。

1.3 本論文の目的

1.2.5 節で述べたように，同時通訳手法に関する研究は存在するものの，その数は少なく，比較的語順が近い言語間を対象としたものや人手で変換規則を構築するものに限られている．特に，独話を対象としたものは，漸進的な構文解析，及び，翻訳単位の検出といった同時通訳の要素技術に関する研究に限られており，実験的なシステムの開発に取り組んだ研究はこれまでほとんど存在していないのが現状である．その要因の一つとして，大規模な同時通訳者の発話データを収録したコーパスがなく，実際の同時通訳者の訳出メカニズムを解明することが十分に行われていないことが挙げられる．

そこで本研究では，同時通訳者の発話データを大規模に収集し，実際の同時通訳者の訳出方略に関する知見を獲得し，獲得した知見をもとに対話を対象とした同時通訳手法を提案し，その手法を実装した同時通訳システムを開発することを目的とする．また，独話に関しては，同時通訳者の訳出方略に関する知見を獲得することを目指す．同時通訳者の訳出方略を獲得するための重要な分析対象としては，通訳パターン，遅延時間，訳出タイミング，話速などがあり，これらを体系的に分析する必要がある．本研究ではその中の一つである同時通訳者の話速に着目して分析する．同時通訳は，話し手の発話の進行に同期しながら訳出を進める通訳スタイルである．このため，同時通訳システムの話速制御は，話し手の発話の進行状況に強く依存する．実際の同時通訳者も，話し手の発話に追従するという制約のもとで，聴

衆にとって聞きやすい訳出を実現するために、自らの話速を柔軟に制御しているものと思われる。その意味で、同時通訳者の話速の変化の様相を明らかにし、同時通訳システムの話速制御を開発する上で参考になる知見を獲得することは極めて重要である。

具体的には、以下の研究に取り組む。

- 同時通訳データベースの構築
プロの同時通訳者の発話データが収録された大規模な音声コーパスを構築する。
- 英日同時翻訳のための訳文生成手法の開発
同時通訳データベースの英日同時通訳者の発話にみられる語順の入れ替え、倒置、言い直しに着目し、依存文法に基づいて訳出タイミングを決定する訳文生成手法を開発する。
- 同時通訳者の話速に影響を及ぼす要因の定量的分析
同時通訳データベースの独話データを用いて、同時通訳者の話速に影響を及ぼす要因を定量的に分析する。

1.4 本論文の内容と構成

本論文の構成は以下の通りである。

第2章では、同時通訳者の訳出方略を獲得するための基盤データとなる同時通訳データベースの設計、構築、及び、利用について詳述する。本データベースの構築では、多様な側面から同時通訳者の発話を分析するために、対話と独話の英日双方向の同時通訳者の音声データを収録する。また、研究資源としての利便性を考慮し、音声データを文字化し、フィラーや言い直しなどの非文法的な言語現象に談話タグを付与する。さらに、発話単位での時間情報や対訳対応関係などの言語情報を人手によりアノテーションする。

第3章では、英日同時翻訳のための訳文生成手法を提案する。英語と日本語のように、大きく異なる性質を持つ言語間の同時翻訳では、対応する語の生起順序が異なるため、標準的な訳文を生成したのでは同時性の高い翻訳を実現することは困難である。例えば、英語はSVO構文であるのに対し、日本語はSOV構文であるため、英語話者発話のOの部分の発話が完了するまでVの訳出を待たなければならない。他にも、英語では時間を表す副詞が一般に文末に生起するのに対して、日本語では文頭に生起するため、発話が完了するまで訳出を開始できない。一方、同時通訳デー

データベースを観察すると、英日同時通訳者は、標準的な訳文ではなく、原文の語順に近い語順の日本語文を訳文とすることにより、早い段階での訳出を可能にしていることがわかる。提案手法では、そのような訳出方法を参考に、日本語として容認可能な範囲で、原文の語順に近い語順の日本語訳を入力と同時的に生成する。本論文では、日本語として容認可能な文が満たすべき性質を、依存構造の観点から検討する。また、そのような性質を満たす訳文を生成する手法について述べる。提案手法では、日本語の語順に関する制約は比較的緩やかであるという特徴を利用し、日本語文として容認可能な範囲内で、原文の語順に準拠した訳文を生成する。訳文としての容認可能性は、日本語依存文法に基づいて検査する。すなわち、依存文法に基づく制約（特に、後方修飾性制約）を満たし、かつ、英文の語順に最も近い日本語文を同時翻訳における訳文とする。ただし、日本語文として不自然でない場合に限り、倒置、すなわち、後方修飾性を満たさない依存関係を含む訳文についても生成を許容する。提案手法を実装したシステムを用いて翻訳実験を行い、翻訳品質、及び、同時性の観点から本手法を評価した。その結果、容認可能な翻訳品質をもち、かつ、入力に対して同時性が高い英日同時翻訳を実現するために、提案手法が有効であることを確認した。

第4章では、英日同時通訳における通訳者の話速に関する定量的分析について述べる。本分析では、特に、同時通訳者の話速に影響を与えると考えられるいくつかの要因に焦点をあて、その影響の程度について分析を与える。具体的には、大局的な要因として「講演の話速」及び「通訳者の個性」、また、局所的な要因として、講演中の「話し手の発話状態」及び「訳出の進行状況」に着目し、話速データを比較することによりそれらの要因が及ぼす影響の程度を明らかにする。定量的分析のために、本研究では、同時通訳データベースの英日通訳データを用いた。分析では、このデータベースの中から22の英語講演データに対する複数の通訳者による88例の英日同時通訳データを抽出し、使用した。また、音声認識ツールを用いて発声時刻情報を自動推定することにより、単語レベルでの話速データを作成し、分析に使用した。分析の結果、講演者の平均話速と通訳者の平均話速の間には相関関係はなく、また、講演の違いに比べ通訳者による話速の違いの方が大きいことを確認した。さらに、原発話が確定前より確定後の方が通訳者の話速が速く、講演者の発話が休止中の方が発話中に比べて通訳者の話速が速いことを確認した。本論文の目的である同時通訳システムの開発という観点から、本分析の結果を捉えると、同時通訳システムの訳出速度の制御方法として、講演全体の平均話速を講演に関係なく一定に保ちつつ、原発話の確定後に話速を上げ、休止中に話速を上げることが一つの方法で

あることが示された。

第2章 同時通訳データベースの構築

2.1 はじめに

これまでに、音声・言語に関する研究資料として活用することを目的に、国内外で大規模な音声・言語コーパスが構築されている。例えば、アメリカ英語の書物のコーパスである Brown Corpus、イギリス英語の書物のコーパスである LOB Corpus、及び、日本の新聞記事に各種の言語情報が付与された京都テキストコーパスが構築されている。また、話し言葉処理の研究を目的に、アメリカ英語の話し言葉コーパスである CSPAE (Corpus of Spoken, Professional American-English)、イギリス英語の話し言葉コーパスである LLC (London-Lund Corpus of Spoken English)、日本語音声収録された日本語話し言葉コーパス (CSJ)[35] が構築されている。さらに、機械翻訳のためのコーパスとして、EU の議会議事録が収録された欧州の複数の言語の対訳コーパスである Europarl Parallel Corpus、英語と南アジアの言語の対訳コーパスが収録された EMILLE Corpus、ATR の対訳コーパスである SLDB (Spoken Language DataBase)[65] や BTEC (Basic Travel Expression Corpus)[26] がある。しかしながら、これまで同時通訳の研究のための言語資源はほとんどなく、大規模な同時通訳データの構築が望まれてきた。

そこで本研究では、同時通訳者の訳出方略を獲得するための基盤データを構築することを目的に、大規模な同時通訳者の発話データが収録された同時通訳データベースを構築する。本データベースが多様な研究目的に利用されることを考慮し、言語情報や時間情報をアノテーションしている。本データベースは、全体で約 182 時間の音声を収録し、音声の文字化、視覚化、及び、言語分析を実施した。収録規模は単語数にして約 100 万語に達しており、同時通訳音声としては世界最大のコーパスである。

本章の構成は、まず 2.2 節では、同時通訳データベースの構築目的と設計について述べる。2.3 節では、データベースの収録方法について述べる。2.4 節では、データベースの構築について述べ、2.5 節では、データベースの有用性について言及する。

表 2.1: 同時通訳データベースの収録様式

項目	詳細
談話形態	対話, 独話
対象言語	英語, 日本語
通訳スタイル	同時通訳
メディア	音声, テキスト

2.2 データベースの設計

同時通訳データベースでは、英日間の同時通訳に関する研究資源として広く活用されることを考慮して設計しており、一人が話し続ける独話（講演）と二人が交互に話す対話の二つの談話形態の同時通訳データを収録した。特に、対話は、音声翻訳システムが求められる状況として空港やホテルでの旅行対話を中心に収集した。対象言語は英語と日本語であり、英日・日英の双方向の音声を収録した。本データベースの収録内容を表 2.1 に示す。

2.3 音声データの収録

同時通訳データベースは、実環境下での同時通訳者の音声を収集することを重視し、教室レベルの収録環境を採用し、収録した。実際の収録状況を図 2.1 に示す。同時通訳者は、話者の発話内容だけで通訳しているのではなく、話者の表情や振舞いも通訳する上で重要な情報となる。そのため、通常の同時通訳者の環境とほぼ同じ環境になるように、通訳者は話者をガラス越しに観察できる通訳用ブースで通訳した。スタンドマイクを使用して収集した音声データは、サンプリング周波数 16kHz、16 ビットでデジタル化し、デジタルオーディオテープ (DAT) に収録している。また、同時通訳者には、プロの通訳者を起用し、高い通訳レベルを保証している。

2.3.1 独話音声の収録

独話データの収録では、講演者が通訳者の存在を意識することなく、自分のペースで発話できるようにするために、講演者が通訳者の音声を聞こえないようにした。同時通訳者は、講演者の発話の様子が見えるように、通訳ブースに入り、ヘッドホン越しに講演者の発話を聞きながら、同時通訳するようにした。本収録では、実際



図 2.1: 収録の様子

の講演ではなく、模擬講演の形式を採用し、講演者には講演するテーマを指示できるようにした。また、同一講演に対し、最大4人の同時通訳者を配置している。図 2.2 に収録の様子を示す。すなわち、データベースには、1つの講演者発話データに対応する複数の通訳データが存在することになる。これにより、収録した通訳データは多様性を備えており、同一の講演に対する複数の通訳事例を比較分析することが可能である。

2.3.2 対話音声の収録

対話の収録では、2名の話者(英語話者と日本語話者)と2名の通訳者(英日通訳者と日英通訳者)を用意した。図 2.3 に収録の様子を示す。実際の通訳者を介した異言語対話では、一人の通訳者が双方向の通訳を行うことが普通であるが、通訳の品質を確保するため英日と日英で別の通訳者を設置した。また、話者としては、相手話者の発話が通訳された音声のみを聞ける環境を設定した。一方、通訳者は、通訳の品質を確保するために、双方の音声を聞ける環境を設定した。さらに、可能な限り自然な対話を収集するため、話者役割と対話タスクを対話設定シートを用いて設定し、基本的には自由発話という様式で収録した。図 2.4 に対話設定シートを示す。収録した対話のドメインは、旅行対話として以下のものを設定した。



図 2.2: 独話の収録の様子

- ホテル

- チャックイン・チェックアウト (対面)
- 観光案内 (対面)
- レストラン紹介 (対面)
- 電話予約 (非対面)

- 空港

- チェックイン (対面)
- 観光案内 (対面)
- 航空券の予約 (非対面)

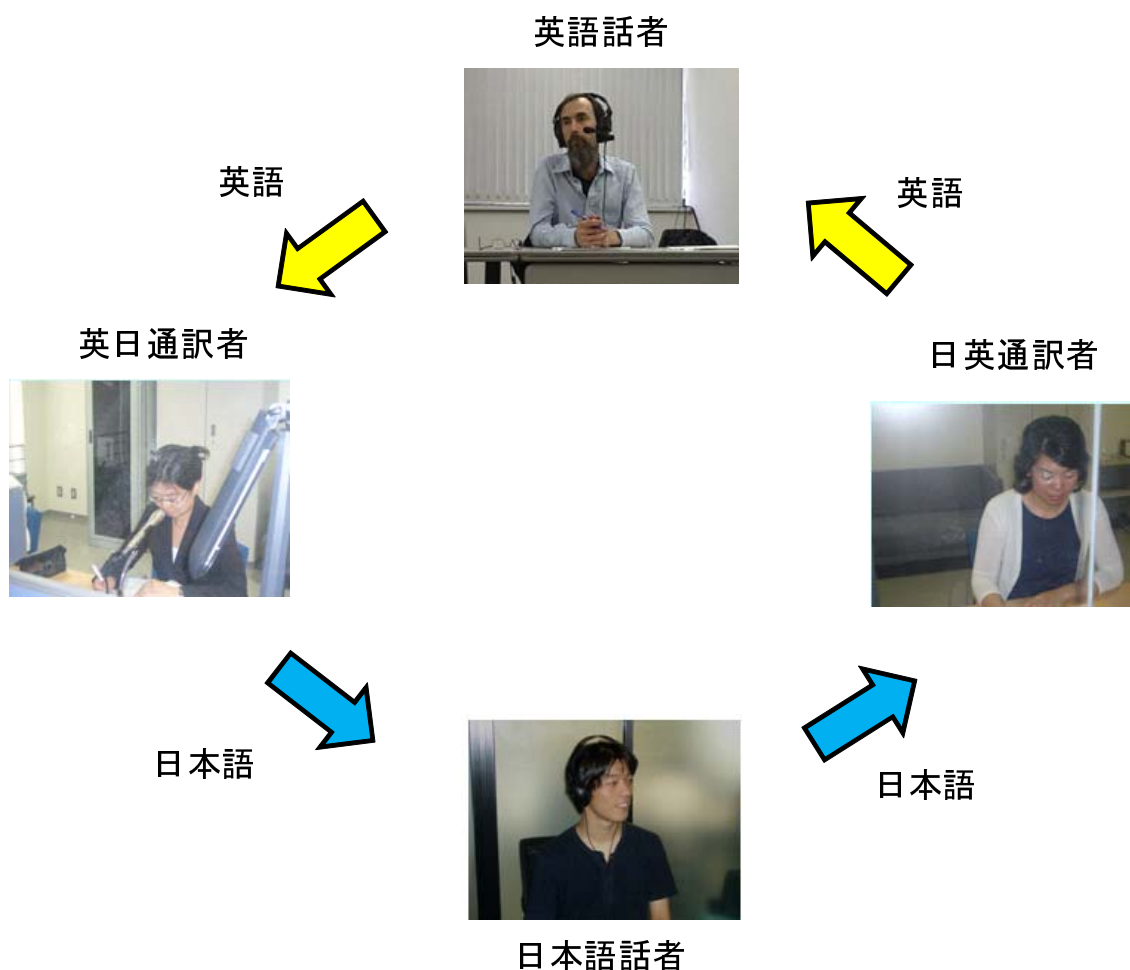


図 2.3: 対話の収録の様子

2.4 データベースの構築

音声・言語処理の研究資源として収録した同時通訳者の音声データを利用するときの利便性を考慮し、収録した音声データを文字化し、談話情報、時間情報、対訳対応情報などを付与した。以下では、各情報の付与方法の詳細を記述する。

2.4.1 音声データの文字化

音声データの文字化は、日本語話し言葉コーパス (CSJ) の書き起こし基準 [28] に準拠した。独話データの文字化データのサンプルを図 2.5 に示す。また、対話データの文字化データのサンプルを図 2.6 に示す。文字化作業は収録した全音声データに対して実施した。以下に主な基準を示す。

英語話者用設定シート		日本語話者用設定シート																	
Dialogue Sheet for Simultaneous Interpreting Corpus		同時通訳コーパスの会話設定シート																	
Sheet No.	CI001-A-EG (Check-in 001 English guest)	シート番号	CI001-A-JS (Check-in 001 Japanese staff)																
Title	Hotel Check-in	タイトル	ホテルのチェックイン																
<table border="1"> <tr><td>Place & Time</td><td></td></tr> <tr><td>Place</td><td>Front Desk</td></tr> <tr><td>Hotel name</td><td>Miyako Hotel</td></tr> <tr><td>Time</td><td>10th December 5:00 p.m.</td></tr> </table>		Place & Time		Place	Front Desk	Hotel name	Miyako Hotel	Time	10th December 5:00 p.m.	<table border="1"> <tr><td>場所・時刻</td><td></td></tr> <tr><td>場所</td><td>ホテルのフロント</td></tr> <tr><td>ホテル名</td><td>都ホテル</td></tr> <tr><td>時刻</td><td>12月10日 夕方5時</td></tr> </table>		場所・時刻		場所	ホテルのフロント	ホテル名	都ホテル	時刻	12月10日 夕方5時
Place & Time																			
Place	Front Desk																		
Hotel name	Miyako Hotel																		
Time	10th December 5:00 p.m.																		
場所・時刻																			
場所	ホテルのフロント																		
ホテル名	都ホテル																		
時刻	12月10日 夕方5時																		
<table border="1"> <tr><td>Role</td><td></td></tr> <tr><td>Your Role</td><td>Name: Jason Terry Job: Guest Age: 20</td></tr> <tr><td>Partner Role</td><td>Name: Hotel Receptionist Job: Hotel Receptionist Age: About 30</td></tr> </table>		Role		Your Role	Name: Jason Terry Job: Guest Age: 20	Partner Role	Name: Hotel Receptionist Job: Hotel Receptionist Age: About 30	<table border="1"> <tr><td>配役</td><td></td></tr> <tr><td>自分</td><td>名前: 内藤 浩二 (ないとう こうじ) 職業: ホテルのフロント係 年齢: 30歳</td></tr> <tr><td>相手</td><td>名前: お客様 職業: お客様 年齢: 20歳前後</td></tr> </table>		配役		自分	名前: 内藤 浩二 (ないとう こうじ) 職業: ホテルのフロント係 年齢: 30歳	相手	名前: お客様 職業: お客様 年齢: 20歳前後				
Role																			
Your Role	Name: Jason Terry Job: Guest Age: 20																		
Partner Role	Name: Hotel Receptionist Job: Hotel Receptionist Age: About 30																		
配役																			
自分	名前: 内藤 浩二 (ないとう こうじ) 職業: ホテルのフロント係 年齢: 30歳																		
相手	名前: お客様 職業: お客様 年齢: 20歳前後																		
<table border="1"> <tr><td>Purpose</td><td>Hotel Check-in</td></tr> <tr><td>Your Reservation</td><td>Name: Jason Terry Number of Guest: 1 (adult) Period of Stay: 2night (10th October) Room: Suite</td></tr> </table>		Purpose	Hotel Check-in	Your Reservation	Name: Jason Terry Number of Guest: 1 (adult) Period of Stay: 2night (10th October) Room: Suite	<table border="1"> <tr><td>目的</td><td>フロント係としてチェックインの応対 予約の有無を確認 予約の確認</td></tr> <tr><td>手順</td><td>部屋の案内 及び キーを渡す 朝食の時刻の確認 (午前7時~9時) チェックアウトの時刻 (午前10時)</td></tr> </table>		目的	フロント係としてチェックインの応対 予約の有無を確認 予約の確認	手順	部屋の案内 及び キーを渡す 朝食の時刻の確認 (午前7時~9時) チェックアウトの時刻 (午前10時)								
Purpose	Hotel Check-in																		
Your Reservation	Name: Jason Terry Number of Guest: 1 (adult) Period of Stay: 2night (10th October) Room: Suite																		
目的	フロント係としてチェックインの応対 予約の有無を確認 予約の確認																		
手順	部屋の案内 及び キーを渡す 朝食の時刻の確認 (午前7時~9時) チェックアウトの時刻 (午前10時)																		
<table border="1"> <tr><td>Question</td><td>Inquire checkout time Inquire breakfast time Inquire room charge Inquire payment way (Is Credit card acceptable?)</td></tr> </table>		Question	Inquire checkout time Inquire breakfast time Inquire room charge Inquire payment way (Is Credit card acceptable?)	<table border="1"> <tr><td>予約</td><td></td></tr> <tr><td>予約の名前</td><td>Jason Terry</td></tr> <tr><td>宿泊期間</td><td>2泊</td></tr> <tr><td>宿泊人数</td><td>1人</td></tr> <tr><td>部屋の種類</td><td>スイート(1001号室)</td></tr> </table>		予約		予約の名前	Jason Terry	宿泊期間	2泊	宿泊人数	1人	部屋の種類	スイート(1001号室)				
Question	Inquire checkout time Inquire breakfast time Inquire room charge Inquire payment way (Is Credit card acceptable?)																		
予約																			
予約の名前	Jason Terry																		
宿泊期間	2泊																		
宿泊人数	1人																		
部屋の種類	スイート(1001号室)																		
<table border="1"> <tr><td>部屋の情報 (1001号室)</td><td>場所: 10階 行き方: フロントの横のエレベータを使用</td></tr> </table>		部屋の情報 (1001号室)	場所: 10階 行き方: フロントの横のエレベータを使用	<table border="1"> <tr><td>その他の情報</td><td>宿泊代: スイート (1泊) 5万円 支払い方法: クレジットカード (JCBなど)、現金</td></tr> </table>		その他の情報	宿泊代: スイート (1泊) 5万円 支払い方法: クレジットカード (JCBなど)、現金												
部屋の情報 (1001号室)	場所: 10階 行き方: フロントの横のエレベータを使用																		
その他の情報	宿泊代: スイート (1泊) 5万円 支払い方法: クレジットカード (JCBなど)、現金																		

図 2.4: 対話設定シート

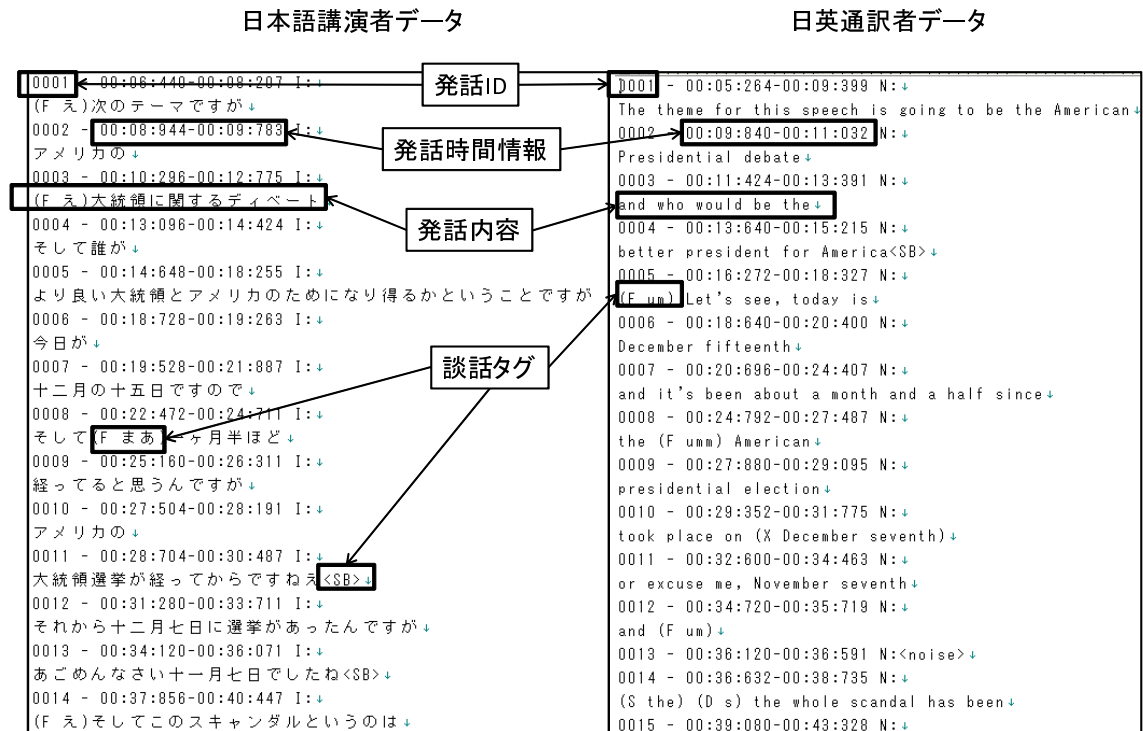
- 発話単位

話者, 及び, 通訳者の音声を 200msec 以上のポーズ (無音声区間) で分割し, 無音区間で挟まれた発話区間を発話単位とした.
- 表記方法

日本語音声については, 片仮名で表記する「発音形」と漢字仮名まじりで表記される「基本形」の2種類で表記した.
- タグ情報の付与
 - 発話 ID

全発話に通し番号を付与した
 - 時間情報タグ

発話単位ごとに発声の開始時刻と終了時刻を付与した



ー 談話タグ

話し言葉に特有の言語的現象であるフィラー（「えー」、「あー」など）や言い淀みに談話タグを付与した。付与した談話タグは2種類あり、一つは文字範囲を指定して、その範囲の特徴を表現するもの、もう一つが、音や非言語的なイベントを表現するものであり、それぞれの種類の談話タグを表 2.2 と表 2.3 に示す。

2.4.2 音声データの視覚化

音声データを聴くだけで通訳者の発話を分析することは簡単ではないため、話者や通訳者の発話の様子を視覚的に観察できることが望ましい。そこで、文字化データの時間情報をもとに話者と通訳者の発声タイミングを視覚的に表示するツールをスクリプト言語 Perl で作成した。独話音声、対話音声を視覚化したデータをそれぞれ図 2.7、図 2.8 に示す。

<p>0001 - 00:08:768-00:13:608 N: Yeah, hello, good morning<SB> I have a ticket reserved for December fifteenth</p> <p>0002 - 00:14:096-00:16:712 N: from Narita to Los Angeles,</p> <p>0003 - 00:16:952-00:19:304 N: but I would like to change, please<SB></p> <p>0004 - 00:27:640-00:28:552 N: That's correct<SB></p> <p>0005 - 00:32:928-00:34:895 N: It's (A J L six two; JL62)<SB></p> <p style="text-align: center;">英語話者データ</p>	<p>0001 - 00:09:640-00:11:239 I: はい<SB>おはようございます<SB> ハイ<SB>オハヨーゴザイマス<SB></p> <p>0002 - 00:12:056-00:12:455 I: (F うー) (F ウー)</p> <p>0003 - 00:13:791-00:15:600 I: 十二月十五日に ジューニガツジューゴニチニ</p> <p>0004 - 00:15:928-00:20:287 I: 成田からロサンジェルスまでチケットを予約してるんですけども変更したいんですが<SB></p> <p style="text-align: center;">英日通訳者データ</p>
<p>0001 - 00:04:136-00:04:528 I: Hello<SB></p> <p>0002 - 00:05:232-00:06:648 I: This is Japan Airlines<SB></p> <p>0003 - 00:07:552-00:08:112 I: May I help you?<SB></p> <p>0004 - 00:21:640-00:24:088 I: So, you got a ticket (F a) (W a; for a) flight</p> <p>0005 - 00:24:604-00:26:416 I: from Narita to Los Angeles leaving on</p> <p>0006 - 00:26:632-00:27:532 I: the fifteenth of December?<SB></p> <p style="text-align: center;">日英通訳者データ</p>	<p>0001 - 00:02:928-00:03:735 N: もしもし<SB> モシモシ<SB></p> <p>0002 - 00:04:280-00:05:783 N: 日本航空でございます<SB> ニホンコーケーデゴザイマス<SB></p> <p>0003 - 00:20:584-00:20:848 N: (D なり) (D ナリ)</p> <p>0004 - 00:21:048-00:24:719 N: 十二月十五日出発の成田からロサンジェルスですな<SB></p> <p style="text-align: center;">日本語話者データ</p>

図 2.6: 文字化データ (対話)

2.4.3 対訳対応データの作成

同時通訳に関する訳出方略を明らかにするために、話者発話とそれに対応する通訳者発話の対訳データを大規模に分析することは重要である。対訳対応付けを自動化する試みもあるが [63]、同時通訳発話にはフィラーや言い淀み、意訳や誤訳、訳出の省略などが頻出するため、高い精度で対応付けることは困難であり、現実的には作業の多くを人手に頼らざるを得ない。人手による対訳対応付け作業のコストを軽減するために、それを支援するツールを作成した。図 2.9 に対訳対応付け作業支援ツールを示す。作業者は、対訳対応付けのために画面上の操作ボタンをクリックすればよい。対応付け作業に対しては、アノテーションマニュアルを作成し、デー

表 2.2: 文字範囲の特徴を表現する談話タグ

タグ	タグの意味	例
(A)	基本形に漢字，仮名以外の文字の使用	(A イーユー; EU)
(D)	断片化した語	あと (D よす)
(F)	フィラー	(F えー)
(L)	ささやき声や独り言などの小さな声	(L アノコレナンダツケ)
(O)	外国語や古語，方言など	(O ザッツファイン)
(Q)	固有名詞 (店名)	(Q タコヘー)
(R)	人名	(R アンディプティット)
(S)	辞書に未登録の口語表現の出現	(S ほりゃ)
(W)	言い間違い，転訛，発音の弱化	(W ミダリ; ヒダリ)
(X)	言い直し	午後 (X ニジ) 十二時
(?)	聞き取り，語彙同定，漢字表記に自信無し	(? タオンゲー)
(笑)	笑いを伴う発話	(笑 なにそれ)
(雑音)	雑音を伴う発話	(雑音 以上です)

表 2.3: 音や非言語的なイベントを表現する談話タグ

タグ	タグの意味	例
< FV >	母音不確定音	それで < FV >
< H >	母音の非語彙的延長	どのフライト < H >
< I >	相手話者の発話による発話の打ち切り	違い < I >
< C >	ポーズにより長単位が分割されている	バスターミ < C >
< 笑 >	笑い (同時に喋ってはいない)	< 笑 >
< 息 >	呼吸音	あるわけで < 息 >
< SB >	文末	名古屋行きです < SB >

タグの均一性を保っている。対訳対応付けの基準として、

- 発話単位を最小単位とする
- 可能な限り細かく対応を付ける
- 時系列上で対応を交差させない

などを設けている。ただし、フィラーや雑音のみの発話単位や、通訳者が省略したことにより話者発話に対応先がない発話に対しては、対応先なしとすることを許した。

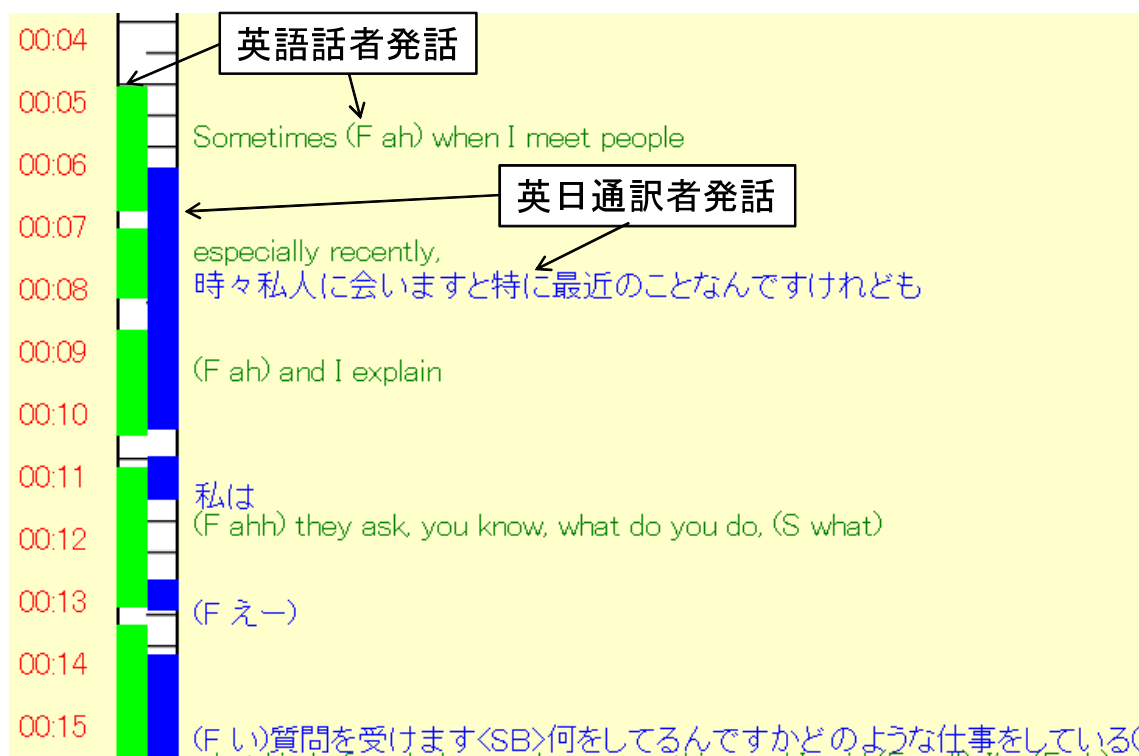


図 2.7: 視覚化ツール (独話)

2.4.4 収録環境情報の作成

本データベースは、大規模であり、多様な目的で利用されることも考えられる。研究目的によっては、音声データや文字化データには現れない情報が重要となる。このため、本データベースの汎用性を考慮し、各講演、対話ごとに、収録に関する情報を記述した収録環境情報ファイルを作成した。図 2.10 に対話の収録環境情報ファイルを示す。収録環境情報としては、収録日時、場所、時間、機器、トピック、話者役割（対話のみ）、話者情報、通訳者情報（経験年数）を記載している。なお、話者、及び、通訳者の名前は、仮名である。

2.5 データベースの構成と規模

同時通訳データベースは、現在までに、収録時間にして約 182 時間、単語数（日本語は形態素数）にして約 100 万単語を収録している。独話データの基礎統計を表 2.4 に示す。また、対話データの基礎統計を表 2.5 に示す。同時通訳データベースは、音声データファイル、文字化データファイル、環境データファイルの 3 つから構成

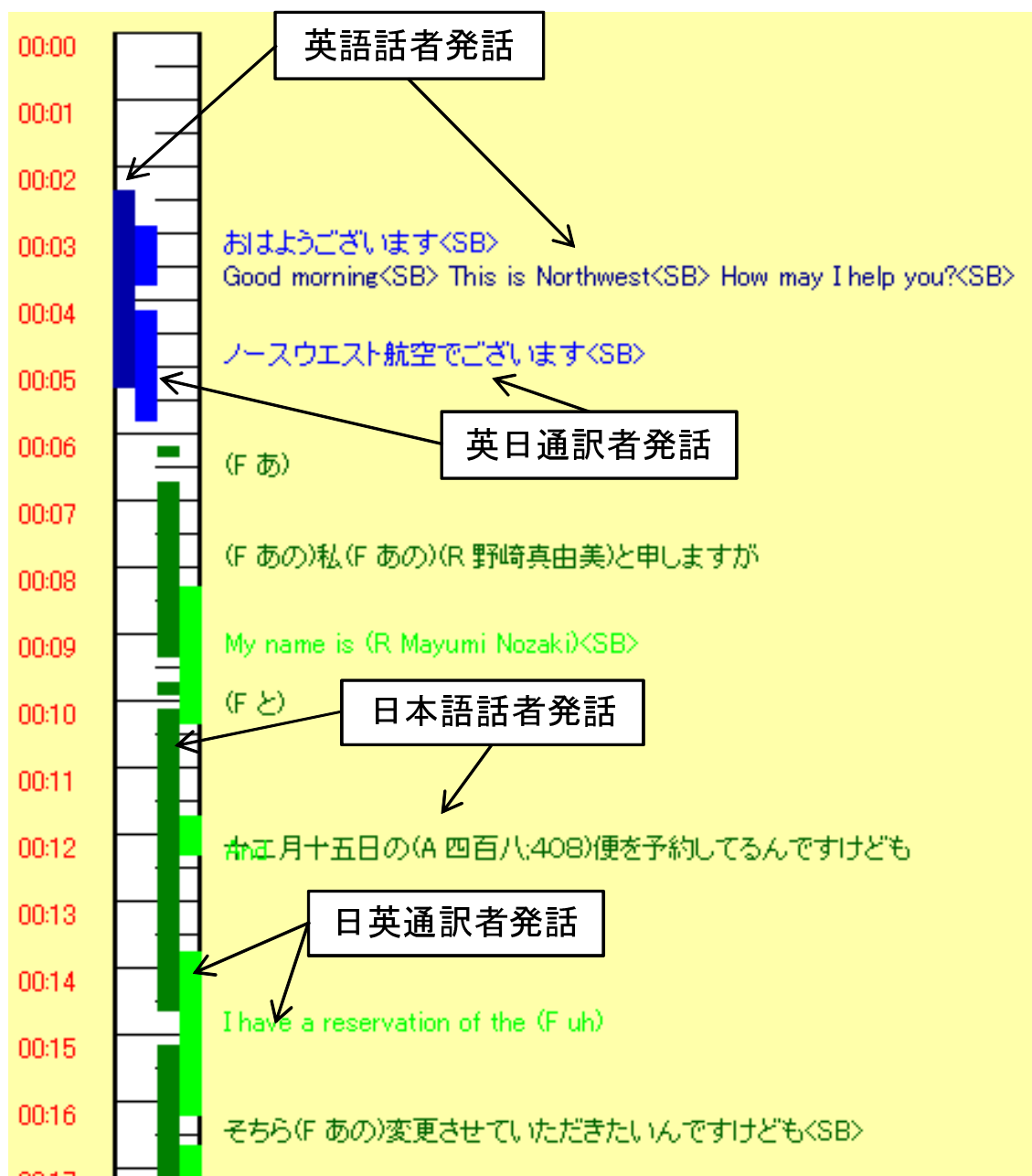


図 2.8: 視覚化ツール (対話)

されている。

- 独話音声データファイル
1つの講演につき英語と日本語からなる1つの多重音声ファイルを格納した。
- 対話音声データファイル
話者1人と通訳者1人の多重音声ファイルを1ファイルとし、2つのステレオ

#	講演者発話	通訳者発話
0	0001 - 00:05:128-00:06:599 N: The topic of this speech 0002 - 00:07:168-00:08:551 N: is Christmas<SB>	0001 - 00:06:880-00:09:535 I: (F えー)スピーチのトピック(雑音 (F うー)は)クリスマスです<SB>
1	0003 - 00:08:176-00:08:287 I: <雑音>	0002 - 00:08:176-00:08:287 I: <雑音>
2	0004 - 00:09:624-00:10:519 N: (noise This) is a little bit 0005 - 00:10:936-00:12:111 N: more of an interesting 0006 - 00:12:399-00:12:983 N: topic<SB>	0003 - 00:10:872-00:11:983 I: (F えー)これは(F まあ) 0004 - 00:12:704-00:15:943 I: ちょっと皆さんにとっては関心深い(F いー)トピックなんではないでしょうか<SB>
3	0007 - 00:13:520-00:15:448 N: It is one that 0008 - 00:15:711-00:16:664 N: interested 0009 - 00:16:672-00:17:328 N: in 0010 - 00:17:328-00:18:096 N: interested in<SB>	0005 - 00:18:064-00:21:264 I: (F うー)我々はこのトピックには関心があると思
4	0011 - 00:18:696-00:19:520 N: But I 0012 - 00:19:872-00:20:576 N: grew up 0013 - 00:20:871-00:21:656 N: in Europe<SB>	0006 - 00:21:344-00:23:375 I: (F んー)で私自身はヨーロッパで育ちました<SB>
5	0014 - 00:22:128-00:24:360 N: My parents were missionaries there	0007 - 00:23:880-00:26:343 I: 私の(F おー)両親は(X 伝道者)布教者でした<SB>
6	0015 - 00:24:928-00:27:120 N: and we traveled constantly<SB>	0008 - 00:26:552-00:27:855 I: (Q で)したがって常に 0009 - 00:28:320-00:29:639 I: (F えー)移動してましたわ(サです)<SB>
7	0016 - 00:27:856-00:29:815 N: We traveled all over Europe	0010 - 00:29:912-00:31:639 I: ヨーロッパ中を移動しましたし

図 2.9: 対訳対応ツール

音声ファイルを作成した。

- 文字化データファイル

英語話者発話テキスト，英日通訳者発話テキスト，日本語話者発話テキスト，日英通訳者発話テキストの4つに分け，データの種類，性質ごとにディレクトリで区分して格納した。

- 環境データファイル

独話，及び，対話の通訳データの収録環境に関する情報を記述したテキストファイルを，1通訳事例につき1ファイルとして作成した。ファイルには，収録時間，通訳者の年齢，経歴年数，講演トピック，対話のタスク名，話者役割などが記載されている。図 2.10 に環境データファイルを示す。

会話ID: DTEPT033
収録日: 2002年1月11日
収録場所: インターグループ大阪本社内インタースクール
収録時間: 7分44秒
防音環境: 教室
マイクロホン: SENNHEIZERヘッドセットHMD280
DAT: SONY デジタルオーディオテープデッキDTC-ZE700
サンプリング周波数: 16kHz
発話形態: 演技
ドメイン: 旅行
トピック: 旅行対話
タイトル: 電話での航空機予約変更・キャンセル
言語パターン: 英語－英日通訳者－日本語－日英通訳者
対面・非対面: 非対面
通訳者への説明: 事前に設定シートを見せるが通訳時には見せない
英語話者役割: 客
日本語話者役割: 航空会社のチケット予約係
英語話者情報: Walker__Steven__男__21__アメリカ__英会話講師
日本語話者情報: 藤沢 茜__女__32__和歌山__会社員
英日通訳者情報: 本田__礼子__女__53__大阪府__26年
日英通訳者情報: 林__喜子__女__38__奈良県__10年
コメント: 特に無し

図 2.10: 収録環境情報ファイル

2.6 データベースの有用性

同時通訳データベースは、同時通訳を多様な観点から分析することを考慮して設計されており、これまでに、通訳者の訳出方法を理論的・実証的に研究する通訳理論学、プロ通訳者を要請するための教育方法を研究する通訳教育学、コミュニケーションに関する事柄を多様な観点から研究するコミュニケーション学の観点から同時通訳を分析するために利用されている。具体的には、通訳理論学の観点から、講

表 2.4: データベースの規模 (独話)

項目		単語数	発話数	収録時間 (分)
話者	英語	90,249	8,422	695
	日本語	84,278	6,529	597
	合計	174,527	14,951	1,292
通訳者	英日	266,050	25,507	1,639
	日英	127,991	16,083	1,265
	合計	394,041	41,590	2,904
合計		568,568	56,541	4,196

表 2.5: データベースの規模 (対話)

項目		単語数	発話数	収録時間 (分)
話者	英語	107,850	14,223	1,678
	日本語	106,258	16,485	1,678
	合計	214,108	30,708	3,356
通訳者	英日	116,776	15,286	1,678
	日英	91,743	13,719	1,678
	合計	208,519	29,005	3,356
合計		422,627	59,713	6,712

演の通訳者の遅延時間の分析したもの [52], 訳出遅延時間と巻返しの関係についての分析したもの [67], 及び, 同時通訳者の訳出タイミングを分析したもの [63] が, 通訳教育学の観点から, フィラー, 及び, ポーズと聴きやすさの関係を分析したもの [66, 68] が, コミュニケーション学の観点から, 同時通訳と逐次通訳におけるコミュニケーションの円滑さ, 及び, 効率を分析したもの [50] がある. 同時通訳データベースを利用して多様な観点から同時通訳を分析した研究成果が多数存在することは, 同時通訳データベースの有用性の一端を示している.

2.7 まとめ

本章では、同時通訳データベースの設計、収集、開発、及び、利用について述べた。本データベースは、182時間、100万単語を収録しており、世界最大規模の対訳音声コーパスである。構築の目的は、異言語間コミュニケーション支援環境の実現を目指した、音声翻訳技術の向上、及び、同時通訳理論の構築であり、現在も、音声・言語処理技術を活用して、有用な言語情報のアノテーションを行っている。これまでに、同時通訳データベースを利用して多様な観点から同時通訳を分析した研究成果が多数あり、これは同時通訳データベースの有用性を示している。

第3章 英日同時翻訳のための依存構造に基づく訳文生成手法

3.1 はじめに

近年の音声処理技術の発展にともない、音声翻訳は機械翻訳における重要な研究テーマになりつつある。従来の音声翻訳システムの多くは、ターンや文を単位とした翻訳処理が実現されており、話者の発話の終了を待ってから翻訳を開始するものとして実現されている [12, 20, 34, 47]。このような音声翻訳システムを介して対話を遂行すると、話者発話時間とは別に翻訳処理時間や訳出時間がかかるため、全体の対話時間が増大する。また、聞き手にとっては、話者が発話を開始してから翻訳が開始されるまでが待ち時間となるため、それが対話の円滑さを妨げる要因となる。より自然な対話を可能にするために、同時通訳者のように話者の発話の終了を待たずに翻訳を開始することは一つの方法である [39]。

同時通訳では、英語と日本語のように、互いに違った性質をもつ言語間の場合には、対応する語の生起順序が異なり、訳出タイミングにおいて困難な問題を引き起こす。例えば、英語はSVO構文であるのに対し、日本語はSOV構文であるため、英語話者発話のOの部分の発話が完了するまでVの訳出を待たなければならない。他にも、英語では時間を表す副詞が一般に文末に生起するのに対して、日本語では文頭に生起する傾向があるなどの違いが原因となる [17]。

しかしながら、実際の同時通訳者は、様々な訳出方略を駆使することにより上述の問題を克服していると考えられる。英日同時通訳における方略の一つは、英語文の語の生起順序に近い語順で日本語訳を生成することである。同時的な翻訳処理に関する研究はそれほどなく、語順が近い言語間における同時的な変換方法や [5, 8]、日英翻訳において節ごとに変換するための規則及びその適用法 [45] が提案されている段階にあり、上述のような訳出方略を実現するための方法論についてはこれまで提示されていない。

そこで本章では、英日同時通訳者による訳出例を参考に、英語対話文の入力に対してできる限り同時進行的に日本語訳文を生成するための手法を提案する。本手法

では、日本語の語順に関する制約は比較的緩やかであるという特徴を利用し、日本語文として容認可能な範囲内で、原文の語順に準拠した訳文を生成する。訳文としての容認可能性は、日本語依存文法に基づいて検査する。すなわち、依存文法に基づく制約（特に、後方修飾性制約）を満たし、かつ、英文の語順に最も近い日本語文を同時翻訳における訳文とする。ただし、日本語文として不自然でない場合に限る。倒置、すなわち、後方修飾性を満たさない依存関係を含む訳文についても生成を許容する。

本手法に基づき、英日同時翻訳の実験システムを作成した。システムは、解析、変換、生成から構成され、英語対話文を形成するチャンクが先頭から入力されるたびに各処理を実行し、入力と同時進行的な訳文生成を実現する。実験では、ATIS コーパス [36] を使用した。また、人手により訳文の翻訳品質を4段階で評価し、さらに、訳出遅延時間を計測し、文単位での訳出と比較して評価することにより、容認可能な程度の翻訳品質を保ち、かつ、原文の入力に対して高い同時性を備えた英日同時翻訳における本手法の効果を確認した。

本章の構成は以下の通りである。3.2節では、異言語対話における同時通訳の効果を対話の効率と円滑さの観点から論じる。3.3節では、英日同時通訳者の訳出方略について論じる。次に、3.4節で、提案する訳文生成手法について述べ、3.5節で本手法を導入した同時翻訳システムについて述べる。3.6節では、本システムによる英日同時翻訳実験の結果について述べる。

3.2 異言語間対話における同時通訳の効果

これまで、異言語間対話を支援する音声翻訳システムの通訳スタイルの多くは逐次通訳であった。これは、これまでの言語・音声技術では同時通訳を実現することが難しかったこと、対話は独話に比べ文の長さが短く同時通訳の必要性はないと考えられていたことが原因として挙げられる。

大原らは、異言語間対話に同時通訳を導入する必然性を明らかにするために、同時通訳データベースを用いて対話の効率と円滑さの観点から同時通訳と逐次通訳を比較し、対話を対象とした同時通訳システムの有用性を評価している [50]。本節では、大原らの論文の評価結果をもとに、対話を対象とした同時通訳の必要性について述べる。

- 対話の効率

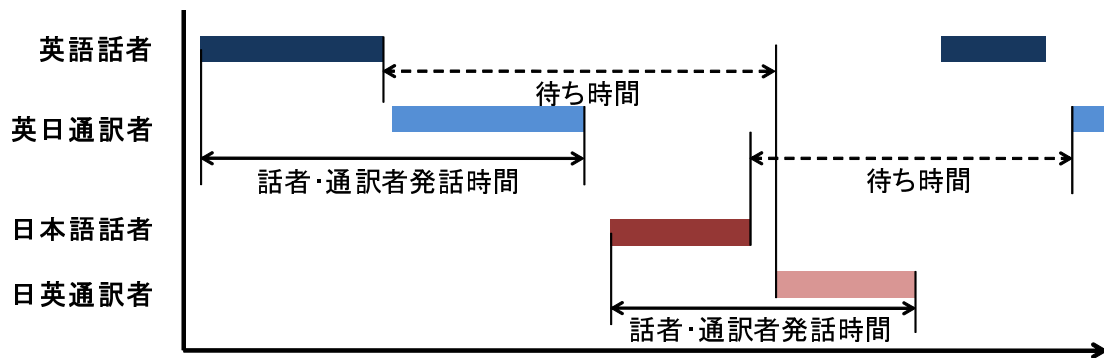
対話は、話者間で発話を繰り返しながらあるタスクを達成することを目的とし

ている。そこで、対話にかかる時間は、対話の効率さを示すものと考えることができる。そのため、本評価では、対話の効率を示す指標として、話者が発話を開始してから通訳者が訳出し終わるまでの時間（以下、話者・通訳者発話時間）を用いた。同時通訳と逐次通訳を介した英日間対話の模式図を図 3.1 に示す。逐次通訳と同時通訳の話者・通訳者発話時間は、図 3.1 の (a)、及び、(b) 内で矢印（実線）で示されている。逐次通訳では、話者の発話と通訳者の発話が重ならないため、話者・通訳者発話時間は話者の発話時間の 2 倍程度になる。一方、同時通訳では、話者の発話が終了した段階で、通訳者が訳出を開始することがあるため、話者・通訳者発話時間が逐次通訳に比べ短くなることを期待できる。

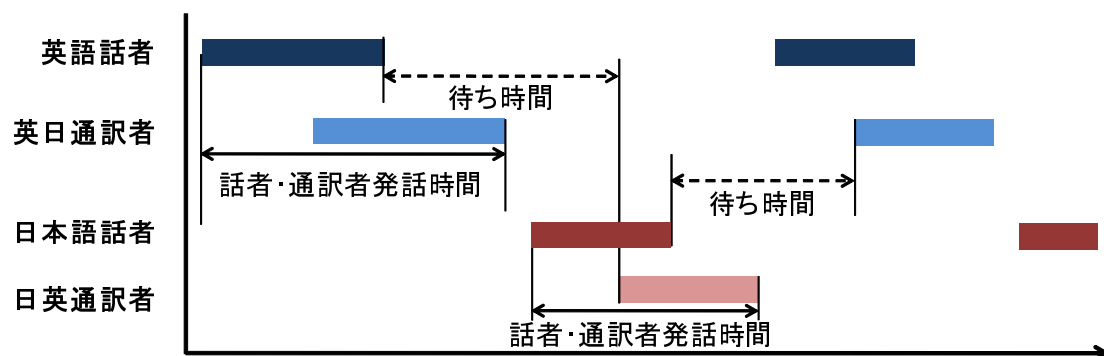
- 対話の円滑さ

円滑な対話を実現するには、話者交替が自然な間隔で行われることが望ましい。話者交替の間隔が必要以上に長くなると、話者は違和感を感じ、円滑な対話を損なわれる恐れがある。しかし、通訳者を介した異言語間対話では、一方の話者は片方の言語のみしか理解できないことが一般的であり、相手側の通訳結果を聞くまで相手の発話を理解できない。すなわち、話者は発話を終了してから、相手側の通訳が始まるまでの間、待ち時間が発生する。そのため、対話の円滑さを示す指標として、片方の話者が発話を終了してから、相手側の通訳が開始されるまでの時間（以下、待ち時間）を用いた。待ち時間を図 3.1 の (a)、及び、(b) 内に矢印（破線）で示す。逐次通訳では、話者の発話が終了してから通訳者の発話が開始されるため、待ち時間は話者の発話時間と通訳者発話時間を足した時間より長くなる。一方、同時通訳では、話者の発話と通訳者の発話が重なるため、待ち時間は話者の発話時間と通訳者発話時間の合計よりも短くなることを期待できる。

対話の効率と円滑さを同時通訳と逐次通訳で比較するための分析データについて述べる。同時通訳のデータは、同時通訳データベースに収録された 216 対話を用いた。逐次通訳のデータには、同時通訳データベースを収録した環境と同じ環境で、同じトピックの対話を実施して収集した 12 対話を用いた。逐次通訳データの基礎統計を表 3.1 に示す。同時通訳のデータに比べ、逐次通訳のデータは少ないが、逐次通訳の通訳形態を考慮すると、時間的な特徴が収録データ間で大きく異なることは少ないこと、及び、同時通訳と比較するための参考データとして用いることを考慮し、比較データとして用いた。



(a) 逐次通訳を介した英日異言語間対話の模式図



(b) 同時通訳を介した英日異言語間対話の模式図

図 3.1: 通訳者を介した異言語間対話の模式図

話者発話時間の合計と話者・通訳者時間合計の比は、同時通訳の場合は 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、大幅に短縮した。この結果より、逐次通訳に比べ同時通訳の方が対話の効率が高い。

通訳を介した異言語対話におけるターンごとの待ち時間の平均は、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。

本評価より、対話の効率と円滑さの観点からみると、対話を対象とした機械翻訳の通訳スタイルとして同時通訳が有効であるといえる。

3.3 英日同時通訳者の訳出

本節では、英日同時翻訳において英語と日本語の語順の違いにより生じる問題について述べる。

表 3.1: 逐次通訳データの基礎統計

項目		単語数	発話数	収録時間 (分)
話者	英語	3,592	405	72
	日本語	3,123	515	72
	合計	6,715	920	144
通訳者	英日	4,367	428	72
	日英	3,623	336	72
	合計	7,990	764	144
合計		12,357	1,684	288

3.3.1 英日間の語順の違いによる問題

同時通訳データベース [69] は、同時通訳者による発話を大規模に収集した音声コーパスである。このデータベースに収録された旅行対話音声のうち、英語話者によって発話された対話文

(E1) You can pick your ticket up at the counter today.

(E2) There are three sushi restaurants near this hotel.

の翻訳について考える。英語文 (E1), (E2) の標準的な日本語訳は、例えば

(J1) 今日、カウンターでチケットをお取りいただけます。

(J2) ホテルの近くにお寿司屋が三軒あります。

である。日本語訳 (J1), (J2) の訳出をできる限り早い段階で行ったとしても、文頭の「今日」「ホテル」に対応する英語 “today”, “hotel” が原文の文末に生起するため、話者の発話が終了する前に通訳者が発話を開始することは不可能である。

3.3.2 同時通訳者の訳出方法

一方、同時通訳データベースに収録されている、英語文 (E1) と (E2) に対する訳出の実例をそれぞれ図 3.2 と図 3.3 に示す。これらの図では、英語話者と英日同時通訳者の発話のタイミングを視覚化し表示している。図 3.2 と図 3.3 は、同時通訳者が話者の発話途中の段階で訳出を開始していることを示している。

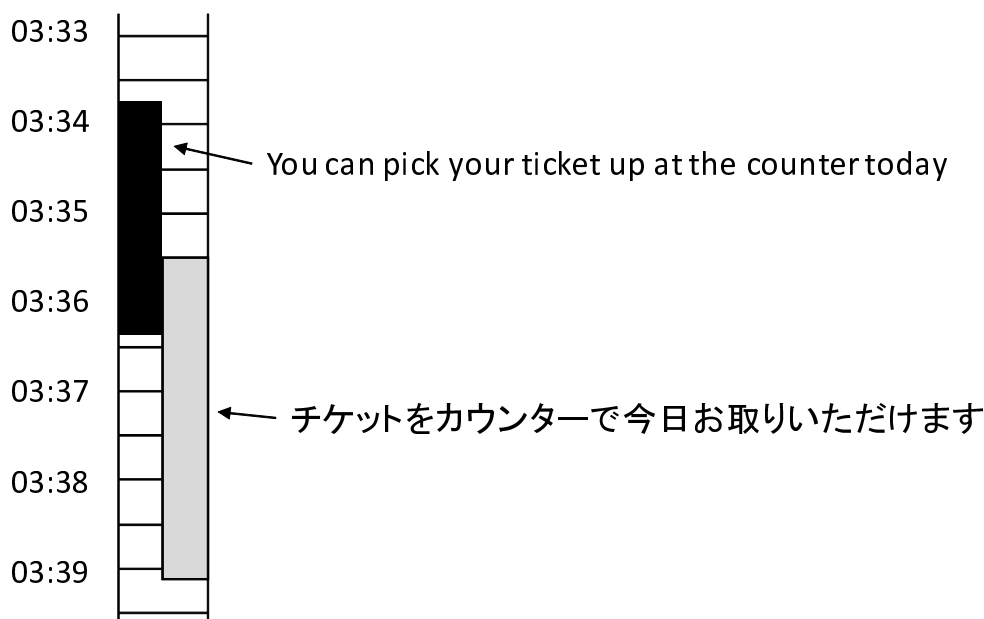


図 3.2: 英語文 (E1) に対する同時通訳者の訳出例

ここで同時通訳者が訳出した日本語文をみると、同時通訳者は英語文 (E1) と (E2) を

(S1) チケットをカウンターで今日お取りいただけます。

(S2) お寿司屋さんが三軒あります、ホテルの近くに。

と訳出しており、これらは、標準的な日本語訳 (J1), (J2) とは語順が異なる。具体的には、日本語訳「今日」と「ホテルの近くに」が、標準的な日本語訳では文頭に生起しているのに対して、同時通訳者の日本語訳では、文のより後方で生起している。これは、同時通訳者の日本語訳が、標準的な日本語訳に比べて、より英語文の語順に近くなる場合があることを意味している。このように、英日同時通訳者は、語順の入れ換えを巧みに活用し、日本語訳を英語文の語順に近い語順で訳出することにより、話者発話との同時進行的な訳出を可能にしている。

本章では、このような同時通訳者の訳出方略を、計算機による英日同時通訳に応用することを試みる。

3.4 訳文生成手法

本節では、3.3節で示した訳出方略を実現するための日本語訳文生成手法について述べる。本手法では、英語入力文を構成するチャンクの対訳表現（以下、対訳チャ

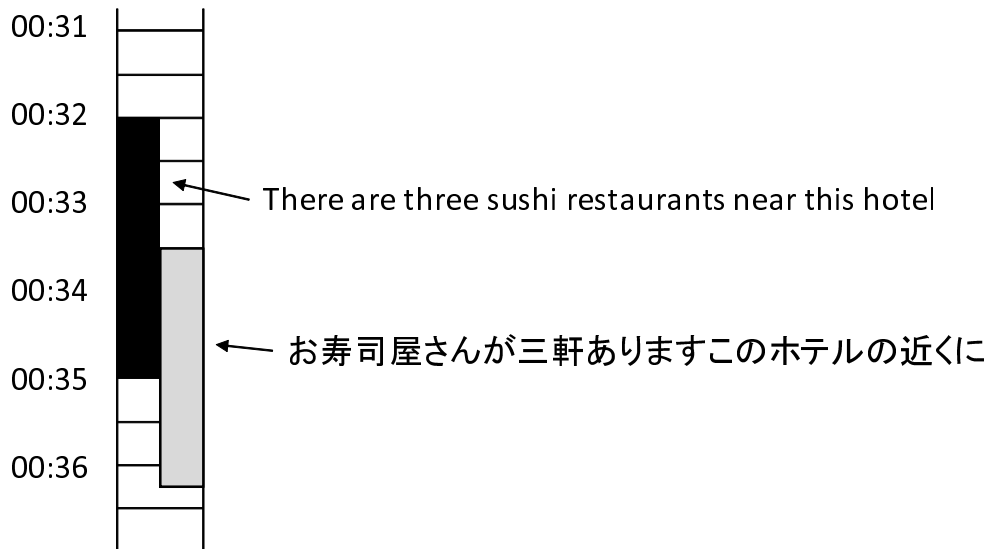


図 3.3: 英語文 (E2) に対する同時通訳者の訳出例

ンク) が, その生起順序に従って順次, 生成部に入力されることを前提とする. また, 対訳チャンク間に成立する係り受け関係についても同様に順次入力されることとする. 例えば, 英語文 (E1) において “You can pick your ticket up” までが入力された段階では, 「チケットを」「お取りいただけます」という対訳チャンクとその間の係り受け関係が生成部に入力されている.

本手法における中心的なアイデアは, 訳文として容認可能な品質を保つという制約のもとで, できる限り早く対訳チャンクを出力するための制御方法にあり, 容認可能性は依存文法に基づいて判定する. なお, 本手法に基づく英日同時翻訳システムについては 3.5 節で詳述する.

3.4.1 容認可能な日本語文とその依存構造

3.3 節で述べた同時通訳者の日本語訳は, 標準的な語順とは異なる語順で構成されているものの, 対話翻訳システムが出力する日本語としては容認できる. 本節では, そのような語順の入れ換えができる条件を依存文法に基づいて検討する.

後方修飾性

日本語には, 文節間に係り受けと呼ばれる依存関係が存在し, 係り文節が受け文節に依存するという関係として定めることができる. 英語文 (E1) の標準的な日本語

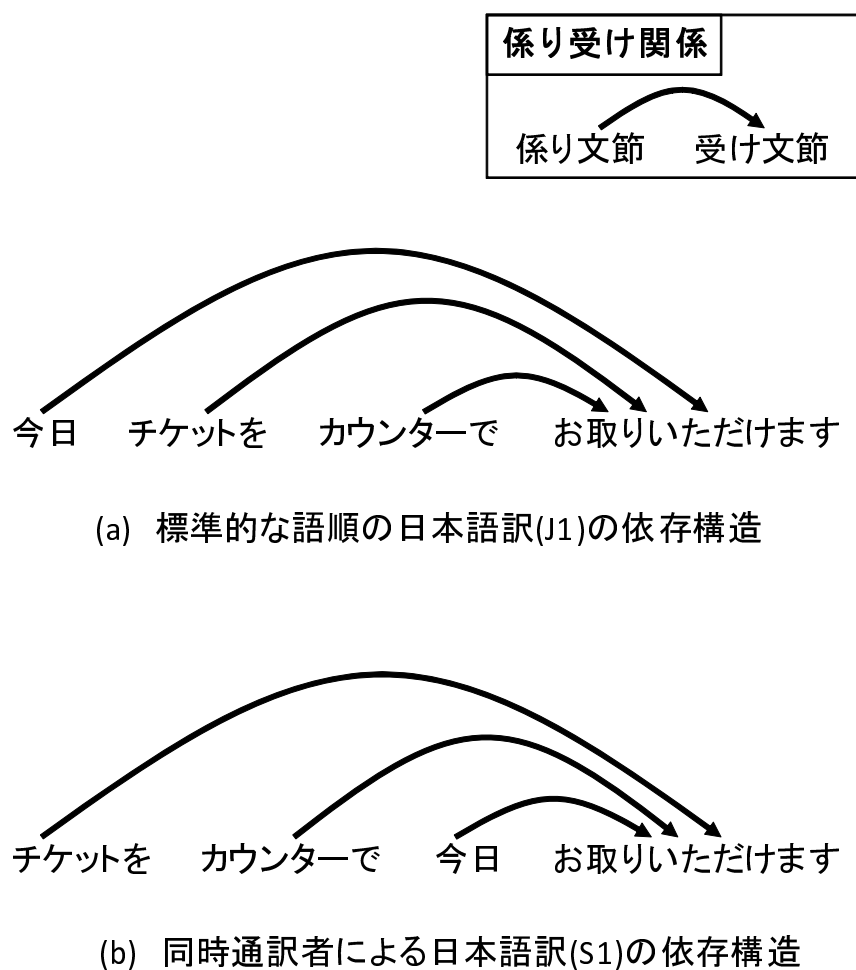


図 3.4: 標準的な日本語訳 (J1) と同時通訳者による日本語訳 (S1) の依存構造

訳 (J1) と同時通訳者の日本語訳 (S1) の依存構造をそれぞれ図 3.4 の (a) と (b) に示す。図 3.4 の各矢印が係り受け関係を示しており、矢印の始点と終点がそれぞれ係り文節と受け文節を示している。標準的な日本語訳 (J1) の係り受け関係は、すべて前方の文節から後方の文節に係っている。このような性質は一般に後方修飾性と呼ばれる。一方、同時通訳者の日本語訳 (S1) も、標準的な日本語訳 (J1) に対して語順が入れ換わっているものの、すべての係り受け関係が後方修飾性を満たしている。係り受け関係における後方修飾性は、日本語における基本的な性質であり、係り受け関係が後方修飾性を満たしていることが、日本語として容認されるための重要な要因になると考えられる。

そこで本手法では、システムが生成する日本語訳文がたとえ標準的な語順でなくても、すべての係り受け関係が後方修飾性を満たしていれば訳文として容認するというように、訳文の語順に関する制約を緩和し、より原文の語順に近い語順の訳文

表 3.2: 係り文節の数と倒置の起こる割合

文節数	全体	倒置	割合 (%)
1	2,023	32	1.58
2	1,376	103	7.49
3以上	600	95	15.83

を生成する。

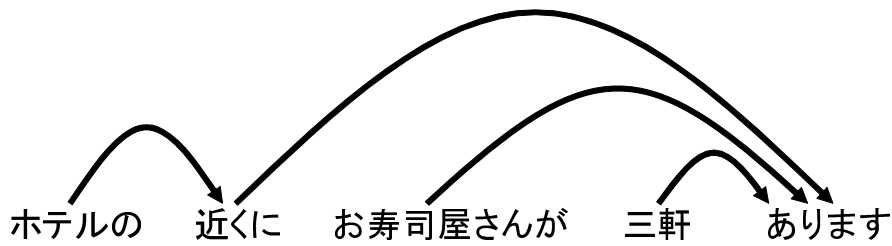
倒置

英語と日本語とでは、述部の生起する位置が大きく異なる。一般に、英語では述部が文頭に近い位置に生起するのに対して、日本語では文末に生起する。そのため、述部は早い段階で入力されてもその訳出は最後となり、述部の訳出遅延が増大するという問題が発生する。この問題は、前節で述べた語順の制約の緩和だけでは解決できない。例えば、図 3.5(a) の日本語訳 (J2) の依存構造では、述部「あります」は、他のすべての対訳チャンク「近くに」、「お寿司屋さんが」、「三軒」の係り先であるため、後方修飾性を満たすには文末で述部を生成する以外にない。

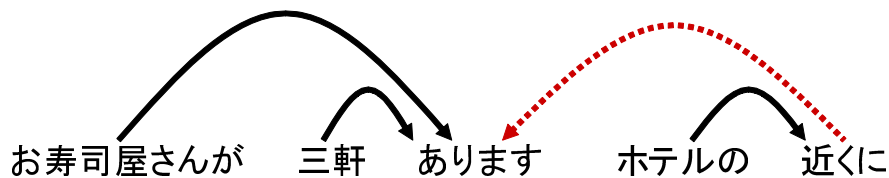
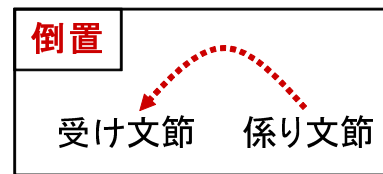
一方、同時通訳者の訳出例 (S2) の依存構造 (図 3.5(b)) では、述部「あります」を文末ではなく、「ホテルの近くに」より前方で訳出している。そして、「近くに」から「あります」に係る係り受け関係が後方修飾性を逸脱している。本論文の以下では、後方修飾性を逸脱する係り受け関係を倒置と呼ぶ。倒置は話し言葉においてたびたび起こるため、倒置を含む日本語文であっても容認できる可能性がある。倒置を許容することにより、より原文の語順に近い訳文を生成できることが期待される。

しかし、その一方で、倒置がむやみに現れると聞き手の理解に支障をきたす可能性がある。倒置が許容される条件を明らかにするために、日本語話し言葉における倒置の特徴を調査した。調査には、名古屋大学 CIAIR 音声対話コーパス [38] における係り受け関係が付与された日本語発話 7781 ターンを利用した。その中には、倒置が 230 個存在した。調査結果より、倒置には、

1. 受け文節の 92.2% が述部である
2. 受け文節に係る係り受けの数が多いほど倒置が起こりやすい (表 3.2 参照)
3. 1 文内に生起する倒置は 2 個以内である



(a) 標準的な語順の日本語訳(J2)の依存構造



(b) 同時通訳者による日本語訳(S2)の依存構造

図 3.5: 標準的な日本語訳 (J2) と同時通訳者による日本語訳 (S2) の依存構造

という特徴があることがわかった。

本研究では、これらの特徴を満たす倒置の生成を許容する。すなわち、生成する日本語訳がすべての係り受け関係が後方修飾性を満たすか、あるいは、倒置となる係り受けが、

1. 倒置の受け文節が述部である
2. 倒置の受け文節に係る文節の数が L 以上である (L の値は実験により決定)
3. 1 文内に生起する倒置は 2 個以内である

を満たすこととする。これらの条件を設けることにより、訳質の低下を抑えつつも語順の制約を緩和することができる。

3.4.2 アルゴリズム

本節では、前節で示した語順制約の緩和を活用して、より原文の語順に近い訳文を入力と同時的に生成する方法について述べる。本手法では、入力との同時的な生成のために、新しい対訳チャンクが生成部へ入力されるたびに、図 3.6 に示すアルゴリズムを実行する。アルゴリズムは、後方修飾性を満たすための出力制御と、倒置を活用するための出力制御の 2 段階で構成される。

英文 “I’ll go to the airport with my friends by taxi next Monday.” に対応する対訳チャンク「行きます」、「空港へ」、「友達と」、「タクシーで」、「来週の月曜日に」が生成部にこの順番で入力された場合の処理の例を図 3.7 に示す。図 3.7 の各列は、左から順に、「生成部に入力された対訳チャンク」、「依存構造」、「各処理時点を表す記号」、「実行される出力判定（図 3.6 の (1) (2), または (3) のいずれか)」、及び、「判定される対訳チャンク」、「出力」に対応している。これ以降では、図 3.7 の訳出例を用いて、処理の詳細について説明する。

後方修飾性の出力制御

もし、すべての対訳チャンクが生成部に入力されてから訳出するのであれば、すべての係り受け関係が後方修飾性を満たす訳文を生成するのは難しいことではない。すなわち、

出力条件 1： 自らを係り先とするすべての対訳チャンクが出力済みである

を満たした対訳チャンクを出現順に出力すればよい。しかし、入力と同時的に訳出する場合には、たとえ入力途中で出力条件 1 を満たした対訳チャンクを出力しても、その後に新たに入力された対訳チャンクがその出力済みの対訳チャンクに係る場合には、後方修飾性を逸脱することになる。例えば、図 3.7(a1) の時点では、「行きます」だけが入力されているので、この時点では出力条件 1 を満たすためそのまま出力できる。しかし、次に入力される「空港へ」が「行きます」に係るため後方修飾性を逸脱する。

そこで、後から入力される対訳チャンクの係り先となる可能性を考慮するために、出力条件 1 に加えて、

出力条件 2： 新たに生成部に入力された対訳チャンクではない

を満たすこととする。例えば、図 3.7(a1) の時点で、「行きます」は自らに係る対訳チャンクがないため出力条件 1 を満たすものの、入力されたばかりの対訳チャンク

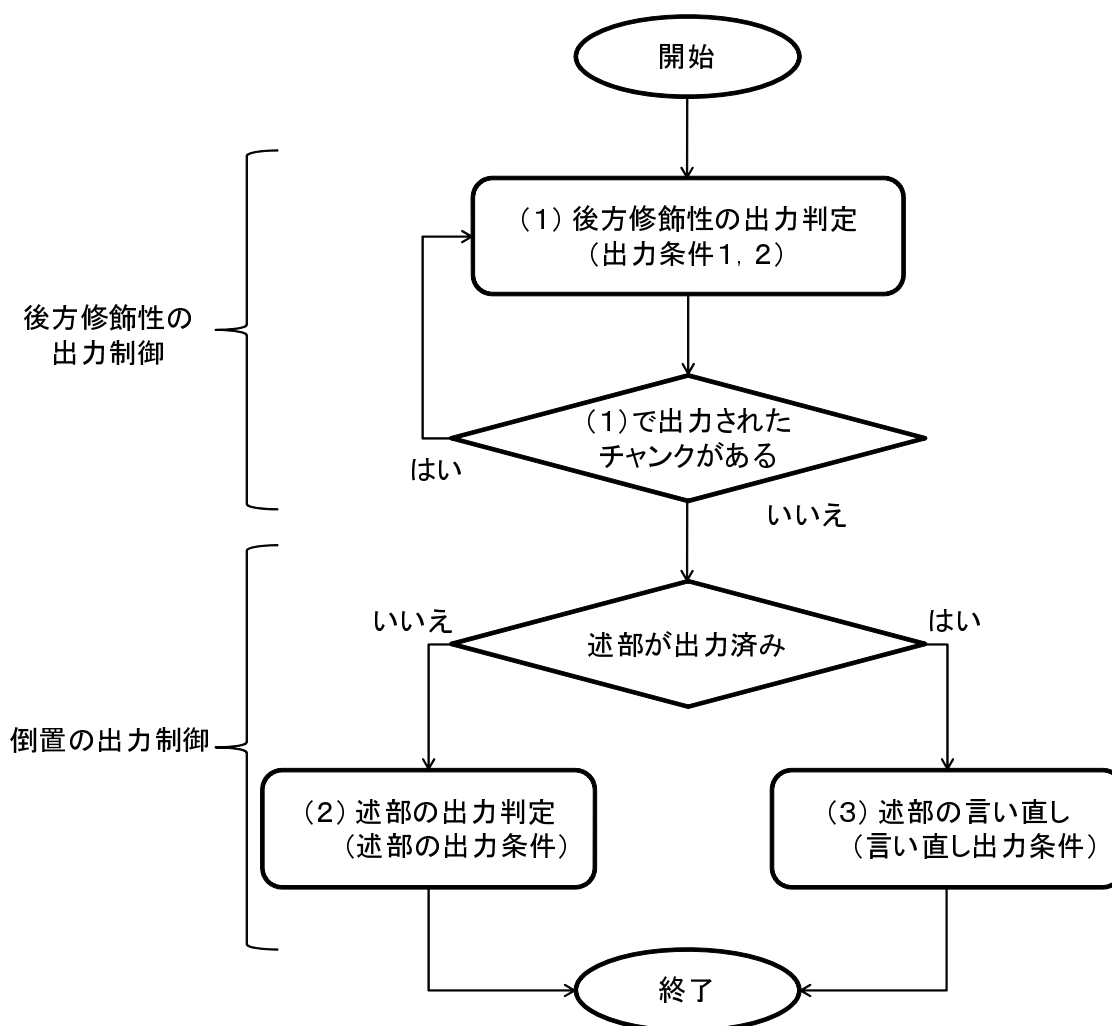


図 3.6: 本手法のアルゴリズム

であり，出力条件2を満たさないため，出力しない．もちろん，出力条件2を用いても後方修飾性を逸脱することはあるが，新しく入力された対訳チャンクがその直前の対訳チャンクに係るかどうかを検出できるので，後方修飾性を逸脱する可能性は低くなる．

出力制御のアルゴリズムでは，出力条件1, 2を用いた出力判定(1)の後に，後方修飾性の判定処理で出力されたチャンクがある場合には，再度，後方修飾性の出力判定(1)を実行している(図3.6参照)．これは，対訳チャンクが出力されることにより，他の対訳チャンクが新たに出力条件1を満たす場合があるためである．





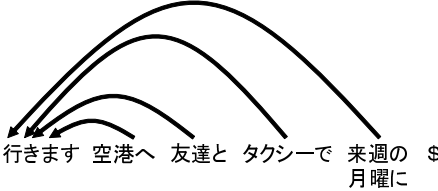
入力	依存構造	記号	出力判定	出力
行きます	行きます	(a1)	(1)「行きます」	
		(a2)	(2)「行きます」	
空港へ	 行きます 空港へ	(b1)	(1)「行きます」	
		(b2)	(1)「空港へ」	
		(b3)	(2)「行きます」	
友達と	 行きます 空港へ 友達と	(c1)	(1)「行きます」	
		(c2)	(1)「空港へ」	空港へ
		(c3)	(1)「友達と」	
		(c4)	(1)「行きます」	
		(c5)	(1)「友達と」	
		(c6)	(2)「行きます」	行きます
タクシーで	 行きます 空港へ 友達と タクシーで	(d1)	(1)「友達と」	友達と
		(d2)	(1)「タクシーで」	
		(d3)	(1)「タクシーで」	
		(d4)	(3)「行きます」	
来週の月曜日に	 行きます 空港へ 友達と タクシーで 来週の月曜に	(e1)	(1)「タクシーで」	タクシーで
		(e2)	(1)「来週の月曜日に」	
		(e3)	(1)「来週の月曜日に」	
		(e4)	(3)「行きます」	
\$	 行きます 空港へ 友達と タクシーで 来週の月曜に \$	(f1)	(1)「来週の月曜日に」	来週の月曜日に
		(f2)	(3)「行きます」	行きます

図 3.7: 出力処理の例

倒置の出力制御

倒置の出力制御では，3.4.1 節で述べた倒置が許容される条件の 1. ~ 3. を満たす訳文を生成する．倒置の条件 1. より，倒置の受け文節は述部に限られるため，述部に相当する対訳チャンクに対する出力条件を緩和する．また，倒置の条件 2 より，述部を出力していない場合には，

述部の出力条件： 自らを係り先とする対訳チャンクのうち
 L 個以上が出力済みである

を述部が満たせば、それを出力する。図 3.7 は、述部の出力条件が $L = 1$ であった場合の処理の例である。「友達と」まで入力された時点（図 3.7(c1) から (c3) を参照）では、出力条件 1 と 2 の双方を満たすので、「空港へ」を出力する。次に、「友達と」は出力条件 2 を満たさないため、出力しない。そして、述部「行きます」は、出力条件 1 を満たさないため後方修飾性の出力判定 (1) では、出力しない。しかし、述部の出力判定 (2) では、出力済みである「空港に」が係っているので、述部に係る文節が 1 個以上出力済みとなり述部の出力条件を満たすため、「行きます」を出力する（図 3.7(c6) 参照）。

しかし、述部の出力条件のみによる制御では、述部を出力した後にその述部に係る対訳チャンクが 3 つ以上出力される可能性があり、その場合に倒置の条件 3. を逸脱する。例えば、図 3.7(f1) では、「友達と」、「タクシーで」、「来週の月曜日」から述部「行きます」への係り受け関係が倒置になっている。そこで、述部が出力済みであるときに、

言い直し条件： 倒置が 3 個以上発生している

を満たす場合には、述部をもう一度出力することにより倒置の生成を回避する。図 3.7(f2) では、倒置となる係り受けが 3 個に達するため言い直し条件を満たす。述部「行きます」を言い直した結果、倒置となる係り受けの生成が解消されている。

複雑な文の出力処理の例

上述の例では、すべての対訳チャンクが主節の述部に係る単純な文を対象に説明を与えたが、より複雑な文でも同様に処理することができる。

例えば、前置詞句をともなう長い名詞句を含む文の場合でも、対訳チャンク間の修飾関係に基づき処理を実行する。“I need a flight from Minneapolis to Long Beach on June twenty six.” では、3 つの前置詞句がいずれも “a flight” に係るため、名詞句の訳出は、対訳チャンク「ミネアポリスから」、「ロングビーチへ」、「6月26日の」をこの順に出力してから「フライトを」を出力する。

また、連体修飾節や副詞節などの従属節を含む複文においても、同様に、出力条件に従って出力を制御する。例えば、“Show me the flights that go from Baltimore to Seattle.” では、“that” 以下は “the flights” に依存する連体修飾節であり、“that” 以降の出力後に “the flights” の対訳チャンクを出力することになる。すなわち「私にボルティモアからシアトルへ行く飛行機を示して下さい」と訳出する。また、“You must contact us if you want to cancel your reservation.” においては、まず、“us” の

対訳チャンク「私達に」を訳出する．ここで，述部の出力条件 ($L=1$ の場合) が満たされるため，この時点で，“must contact”の対訳チャンクを訳出する．最後に，“if”以降の副詞節の対訳チャンクを訳出する．「私達に連絡しなければなりません，予約をキャンセルしたいならば」のように，倒置を含む文として出力される．

3.5 システム構成

3.4 節で示した訳文生成手法を実装した英日同時通訳システムの構成を図 3.8 に示す．

本システムは，構文トランスファー方式に基づいており，英語の依存構造から日本語の依存構造への変換により訳文を漸進的に作り上げる．依存構造の変換として実現することにより依存文法に基づく検査を容易に実行することができ，3.4.2 節で示した出力制御アルゴリズムを自然に導入することができる．なお，同時翻訳に適した対訳が与えられたデータを大量に利用することができれば，本手法をコーパスを用いた方式によって実現できる可能性がある．

本システムは，解析部，変換部，生成部から構成されている．解析部では，入力された英単語列の依存構造を漸進的に解析する．変換部では，英単語列の依存構造に対して構造変換及び語彙変換を実行し，日本語の依存構造に変換する．生成部は，日本語の依存構造をもとに 3.4.2 節で示した出力制御アルゴリズムを用いて日本語訳を生成する．述部の出力条件が $L = 2$ の場合に，英文

(E3) I'll want to fly from San Francisco to Denver next Monday.

の“Denver”が入力されたときの処理の流れを図 3.9 に示す．以下では図 3.9 を用いて各部の処理を説明する．

3.5.1 解析部

解析部では，入力された単語列を辞書を用いて品詞タグ付けし，単語とその品詞情報を用いて baseNP[9] をまとめあげる（図 3.9 の (1) 参照）．次に，文脈自由文法に基づく漸進的依存構造解析手法 [23] に従って入力された英文の依存構造を漸進的に解析する．この解析手法では，主辞付き構文規則を用いて漸進的構文解析法に従い入力文の断片に対する句構造を生成する（図 3.9 の (2) 参照）．また，図 3.9 の (2) の NP(?) のようにまだ入力されていない語（未入力語）についても予測し解析す

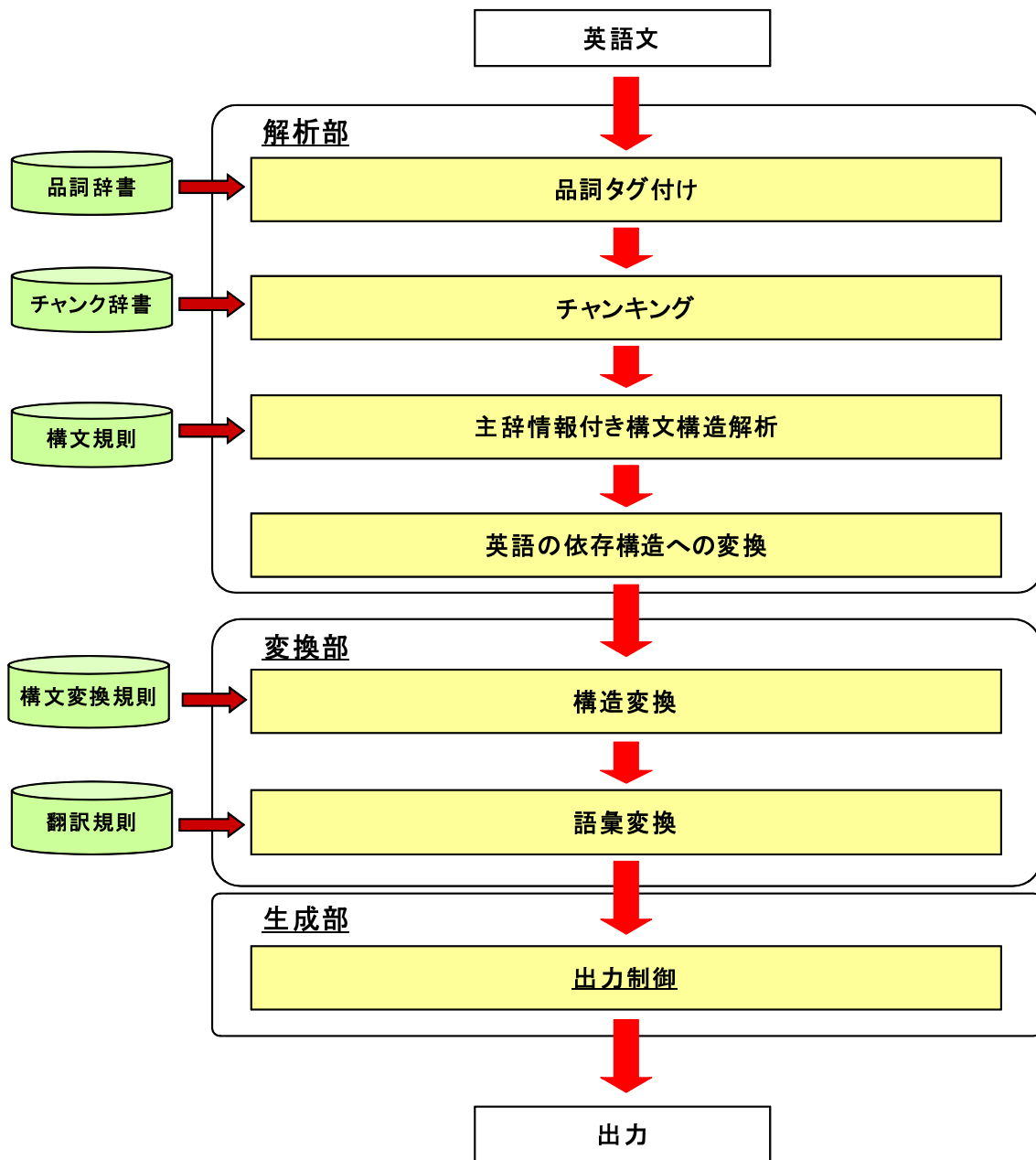


図 3.8: システムの構成図

る。続いて、句構造の主辞情報の品詞をもとに、Collins の規則 [10] に従って句構造をチャンク間の依存構造に変換する (図 3.9 の (3) 参照)。

一般に、漸進的に構文解析すると、入力文の断片に対して膨大な数の構文構造が生成されるという曖昧性の問題が発生する。本システムでは、構文解析に用いる文法規則に確率を付与し、構文構造の生成に利用した規則の確率からその構文構造の確率をもとめ、確率が最大となる構文構造を選択する。

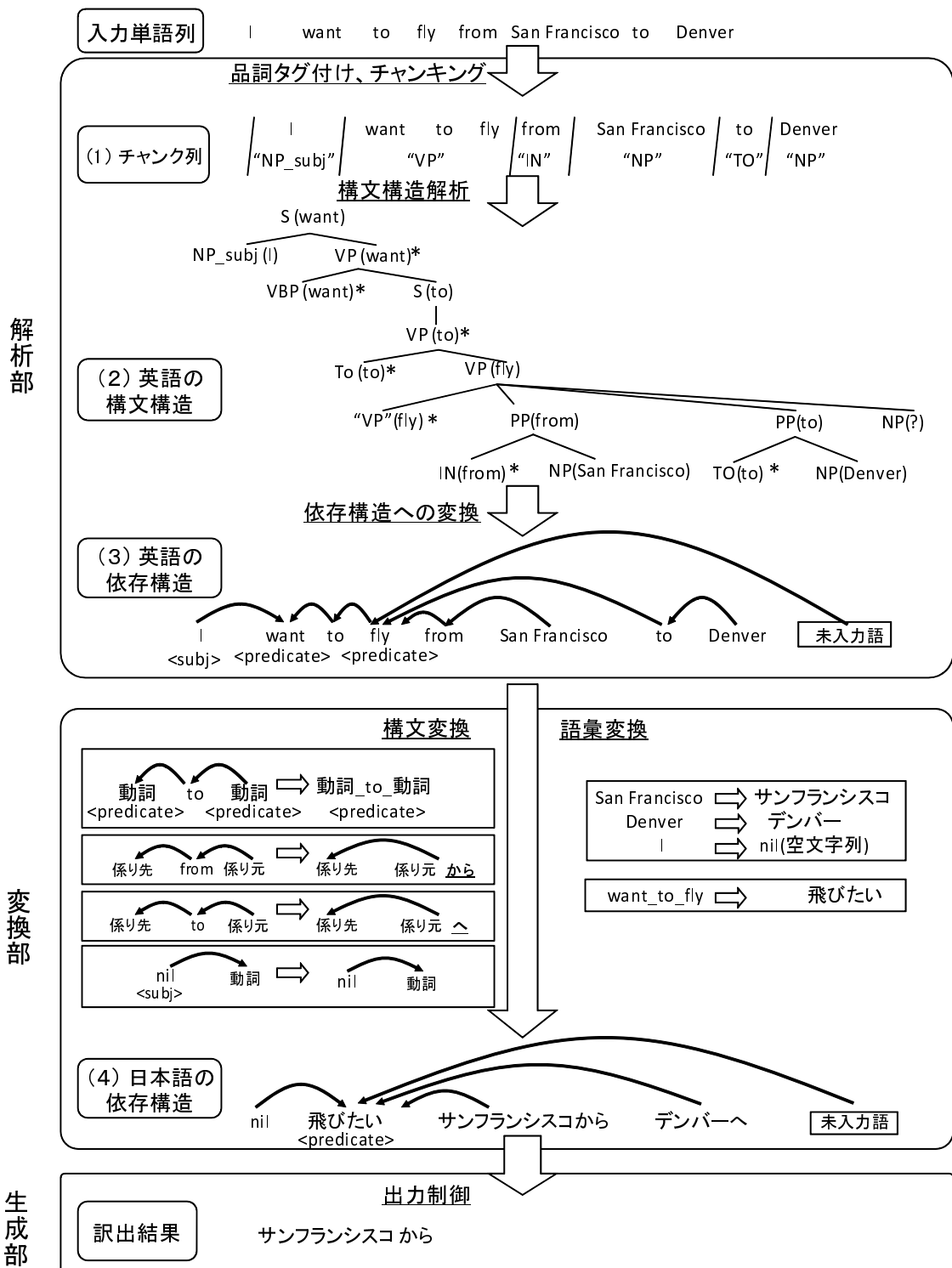


図 3.9: 英文 (E3) の “Denver” が入力された時点の処理 (L = 2)

文法規則には、主語 (subject) と目的語 (object) の情報が付与されており、構文解析することにより主語と目的語になるチャンクを解析する。さらに、品詞が動詞で

ある単語を述部 (predicate) とみなす。図 3.9 (3) の例では、主語は “I” であり、述部は “want” と “fly” の 2 つである。

3.5.2 変換部

変換部では、構造変換と語彙変換を実行することにより、英語の依存構造を日本語の依存構造に変換する (図 3.9 の (4) 参照)。構造変換では、構文変換規則により、助動詞や to 不定詞を本動詞と結びつけ、述部をまとめたり、前置詞をそれに係るチャンクの助詞に変換するなどの処理を行う。例えば、“want to fly” は述部として一つのチャンクにまとめる。前置詞 “from” は、それに係る “San Francisco” に付属する助詞「から」に変換する。語彙変換では、語彙変換規則を用いて対訳チャンクに変換する。

3.5.3 生成部

生成部では、3.4.2 節で説明した出力制御アルゴリズムに従い訳出する。図 3.9 の例では、「サンフランシスコから」は、それに係る対訳チャンクがないため、出力する。一方、「デンバーへ」は、新たに入力された対訳チャンクであるため、この段階では出力しない。また、述部「飛びたい」には、未出力の「デンバー」が係り、かつ、「飛びたい」に係る訳出済みの対訳チャンクは「サンフランシスコから」ただ 1 つであるため、述部の出力条件 ($L=2$) を満たさないことから、この段階では出力しない。

3.5.4 同時翻訳処理の流れ

本システムにおいて、英文 (E3) が入力されたときの入力途中の構文構造、英語の依存構造、日本語の依存構造、及び、出力を図 3.10 と 3.11 に示す。図 3.10 に文頭から “San Francisco” までの処理の流れを示す。また、図 3.11 に “to Denver” から文末記号 \$ が入力されるまでの処理を示す。

3.6 同時翻訳システムによる翻訳実験と評価

本手法により、日本語として容認可能で、かつ、話者の発話に対して同時に訳文を生成できることを示すために、本手法を実装した同時翻訳システムを用いて翻

入力	構文木	英語の依存関係	出力
		日本語の依存関係	
I	<p>S (want) NP_subj (I) VP (?)*</p>	<p>I VP(?)</p>	
want to fly	<p>S (want) NP_subj (I) VP (want) * VBP (want) * S (to) VP (to)* To (to)* VP (fly) "VP" (fly)* PP(from) PP(?) NP(?)</p>	<p>I want to fly PP(?) PP(?) NP(?)</p>	
from	<p>S (want) NP_subj (I) VP (want) * VBP (want) * S (to) VP (to)* To (to)* VP (fly) "VP" (fly)* PP(from) PP(?) NP(?) IN(from) NP(?)</p>	<p>I want to fly from NP(?) PP(?) NP(?)</p>	
San Francisco	<p>S (want) NP_subj (I) VP (want) * VBP (want) * S (to) VP (to)* To (to)* VP (fly) "VP" (fly)* PP(from) PP(?) NP(?) IN(from) NP(San Francisco)</p>	<p>I want to fly from San Francisco PP(?) NP(?)</p>	

図 3.10: 英文 (E3) の”I want to fly from San Francisco”が入力されるまでの処理の流れ

訳実験を行った。実験では、翻訳品質と同時性の観点から本手法を評価した。

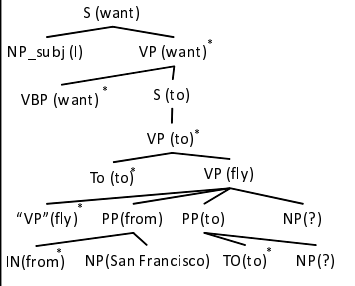
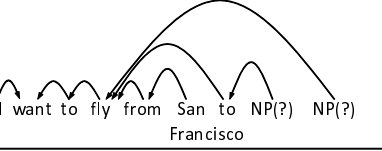
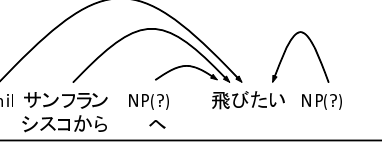
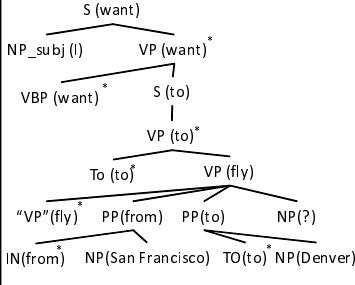
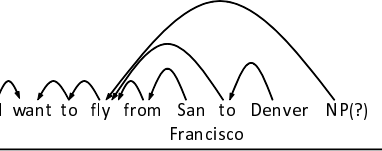
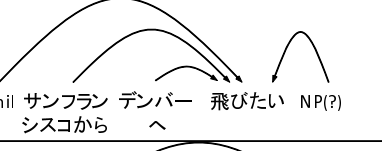
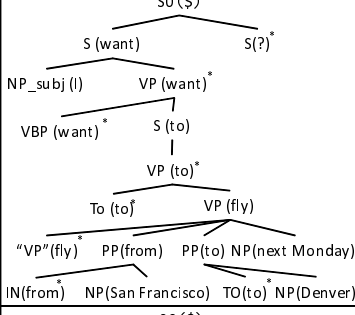


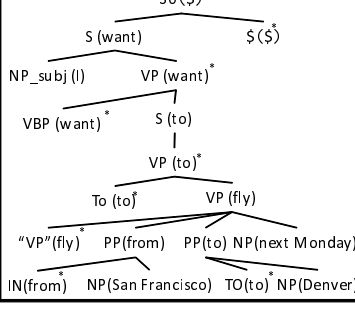
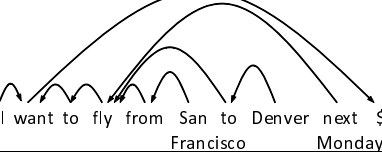

入力	構文木	英語の依存関係		出力
		日本語の依存関係		
to				サンフラン シスコから
Denver				
next Monday				デンバーへ 飛びたい
\$				来週の 月曜日に

図 3.11: 英文 (E3) の”to Denver”から文末まで入力されたときの処理の流れ

3.6.1 翻訳品質と同時性の評価指標

翻訳品質と同時性を評価する指標について説明する。翻訳品質の評価には、訳文を4段階で評価した結果を用いた。評価基準を表3.3に示す。

一方、同時性の評価尺度としては、訳出遅延チャンク数を用いる。訳出遅延チャ

表 3.3: 評価基準

評価ランク	説明
A (Perfect)	会話文として問題がない
B (Fair)	会話文としてぎこちない, 助詞や言葉遣いに間違いがあるものの, 言いたいことはほとんど分かる
C (Acceptable)	訳文は情報が欠けていたり, 断片的な情報しか分からないが, 原文の主要な情報は正しく理解できる
D (Nonsense)	会話文として重要な情報が抜けているか, 原文の意味が正しく理解できない

ンク数とは, あるチャンクが入力されてからチャンクに対応する訳が出力されるまで(ただし, 言い直しされた場合は, 言い直し表現が出力される時点まで)に入力されたチャンクの数である. 以下の式で, 平均訳出遅延チャンク数 D を求める.

$$D = \frac{\sum_{k \in K} d_k}{n} \quad (3.1)$$

ここで d_k は, あるチャンク k の訳出遅延チャンク数を, n は全入力文の総チャンク数を, K は全入力文の全チャンクの集合を表す.

訳文の同時性を評価する別の指標として話者・通訳者発話時間を計測する. 話者・通訳者発話時間とは, 原文の入力を開始してから訳出が終わるまでの時間である. 本研究では, 入力と出力は, とともに一定の話速(7.43 モーラ/秒)で行われると仮定して¹, 算出した. モーラ (mora) とは, 音節より小さい単位である [30]. 日本語の場合, 長母音や二重母音は短母音の2倍の長さとし, また, 母音の後ろの子音は短母音と同じ長さを持つと定めるのが一般的であり, 本研究もこの基準に準拠している.

3.6.2 実験環境と実験データ

本システムを Windows PC (Intel Core2 CPU 1.86Ghz, 2GB RAM) 上に JRE1.6.0 を用いて実装した. 翻訳対象としては, 構文木が付与された英語対話文である必要があり, 本実験では, Penn Treebank[36] に収録されている ATIS コーパスの全 578 文を用いた. 578 文の長さの分布を図 3.12 に示す. 1 文あたりの平均単語数 (チャンク数) は, 7.53 単語 (5.66 チャンク) であり, 最大単語数 (チャンク数) は 34 単語 (29 チャンク) である. 比較的長い文 (11 単語以上の文) も 103 文含まれており, 翻

¹同時通訳データベースの対話データを用いて話速を調査し, その結果に基づいて設定した.

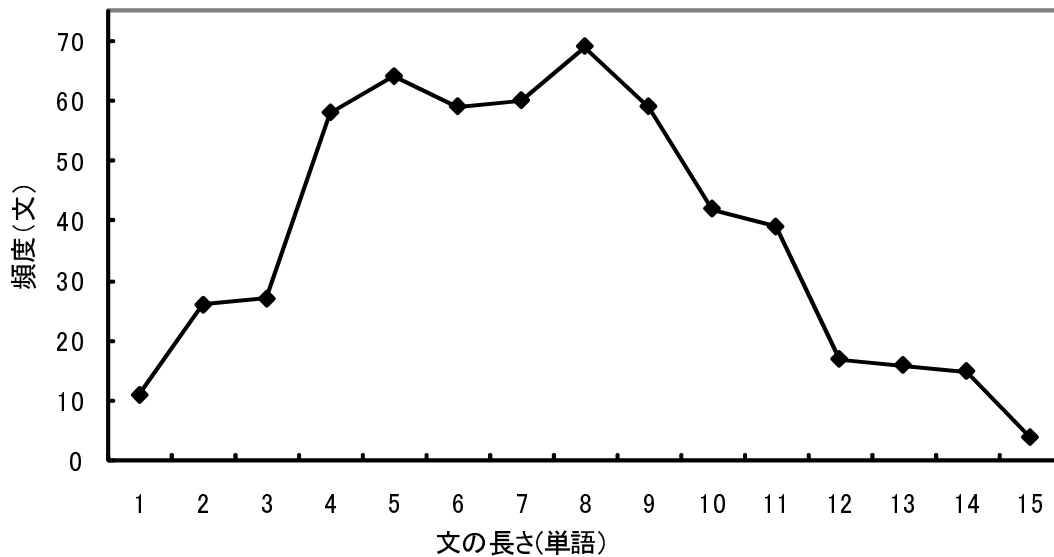


図 3.12: 文の長さの分布

訳処理における同時性の効果の検証に利用できる。578 文の種類の内訳は、単文が 297 文、述部を含まない名詞句や前置詞句が 220 文、関係代名詞節や連体修飾節を含む文が 39 文、その他に重文などが 21 文であった。

構文解析に用いた文法規則として、ATIS コーパスに付与された構文木データから獲得した 327 規則を用いた²。翻訳対象の文に対して翻訳家が参照訳を作成し、その参照訳をもとにチャンク単位で対訳辞書を作成した。原文を漸進的に構文解析する処理では、解析途中の構文候補が膨大な数となるため、解析効率の観点から、解析途中の構文候補に付与された確率が 1.0×10^{-15} 以下になる場合には、その構文候補を枝刈りした。

3.6.3 実験結果

翻訳品質

本手法の有効性を示すために、実験システムの翻訳品質を評価した。倒置の活用が品質に与える影響を調べるために、後方修飾性の出力制御のみを利用して訳出する場合（すなわち、倒置を活用しない場合）と、後方修飾性に加えて倒置を活用（ $L = 2$ ）して訳出する場合の 2 種類で実験した。また、本手法の比較対象として、原文の語

²文献 [55] の手法に従って文法規則を変換して、獲得した。

表 3.4: 翻訳品質の実験結果

手法	翻訳品質			
	A	A+B	A+B+C	D
本手法 (後方修飾性のみ)	116 (20.1%)	351 (60.7%)	414 (71.6%)	164 (28.4%)
本手法 (後方修飾性 + 倒置 L=2)	111 (19.2%)	312 (54.0%)	411 (71.1%)	167 (28.9%)
原文の語順通り	15 (2.6%)	61 (10.6%)	267 (46.2%)	311 (53.8%)
文単位の手法	121 (20.9%)	372 (64.4%)	439 (76.0%)	139 (24.0%)

順通りに訳出する方法（訳出の同時性を重視した方式），ならびに，原文がすべて入力されてから翻訳を開始する方法（以下，文単位の手法）（翻訳品質を重視した方式）を実装し，それぞれ評価した．なお，文単位の手法では，対訳辞書と文法規則は本手法と同じものを使用し，また，標準的な語順で生成するための生成規則を新たに作成し使用した．

実験結果を表 3.4 に示す．容認可能性の観点から，翻訳品質が A, B, C のいずれかである場合を正解とした．本手法（後方修飾性の出力制御のみを利用した場合）の翻訳精度（すなわち，A+B+C の割合）は 71.6% であった．原文の語順通りに訳出した場合と比べて 25.4% 高く，また，文単位の手法に対しては 4.4% の低下に留まっており，容認可能な程度の翻訳品質が保たれていることがわかった．

一方，本手法のうち，倒置導入による翻訳品質への影響を調べるため，「倒置なし」と「倒置あり」とで結果を比較すると，倒置が発生することによって新たに翻訳誤りとなった文は 3 文であった．翻訳精度の低下は 0.5% であり，倒置利用の許容可能性が示された．ただし，高い翻訳品質を備える A+B に着目すると，その割合は，倒置を許容することによって，6.7% 低下することとなった．これは，「倒置なし」において A または B と評価された文のうち，倒置が発生したものが 40 文存在し，そのうち 39 文において翻訳品質が C に低下したためである．例えば，英語文

(E3) On Tuesday I'd like to fly from Detroit to Saint Petersburg.

は，「倒置なし」では

(J3-1) 火曜日にデトロイトからサンクトペテルスブルクへ乗りたい

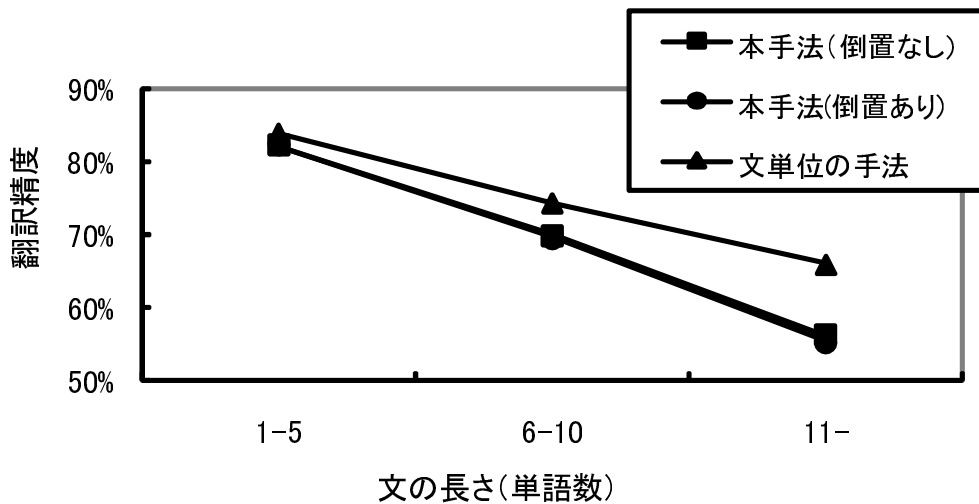


図 3.13: 文の長さとの翻訳精度の関係

と訳出され翻訳品質が A と判定されたが、「倒置あり」では

(J3-2) 火曜日にデトロイトから乗りたいサンクトペテルスブルクへ

という訳出になり翻訳品質が C と判定された。このように、訳文に倒置が含まれると、きれいな訳文 (fair) であるとは評価されず、倒置を許容する限り品質の低下を避けることは難しい。なお、品質が低下しなかった 1 文は、構造上は後方修飾に逸脱する依存関係が存在するものの、表層上では倒置が生じているようには見えない訳文が偶然生成されたことによる。

本手法は、文が長くなるほど同時性においてその効果が大きくなると考えられるが、その一方で、文が長くなり複雑さが増すと、翻訳品質が低下する可能性がある。そこで、入力文の長さ、ならびに、複雑さと翻訳品質との関係を調べた。文の長さとの関係を図 3.13 に、文の複雑さとの関係を図 3.14 に、それぞれ翻訳方式 (倒置あり、倒置なし、文単位) ごとに示す。なお、文の複雑さとしては、文に含まれる動詞の数を用いた。図 3.13 において、本手法は文が長くなると翻訳精度が低下するものの、文単位の手法との間で大きな差はない。また、図 3.14 では、文の複雑さに対する低下傾向は、本手法と文単位の手法とでほぼ同じであり、同時性の効果が示される文においても品質はある程度保たれることが示された。

本手法 (L=2) における翻訳誤りの原因を表 3.5 に示す。なお、誤りの原因が複数ある場合には重複を許している。語順誤りのうち、等位接続詞に関するものが 33 文、Wh-疑問詞に関するものが 36 文存在した。

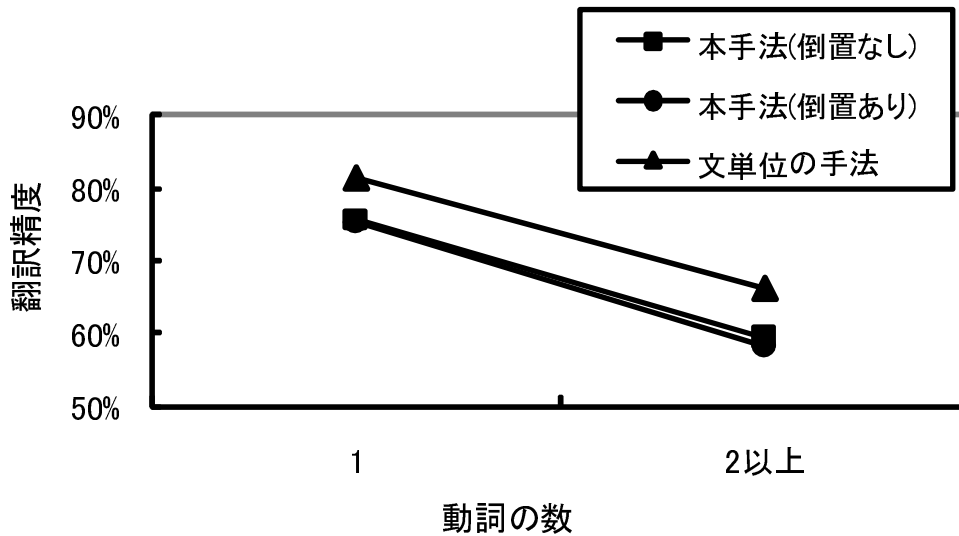


図 3.14: 文の複雑さと翻訳精度の関係

等位接続詞を含む文では，英語と日本語とでその依存構造が対応しない場合があることが主な要因である．本手法で用いた主辞付与規則 [10] に従えば，例えば “X and Y” の場合，先頭の “X” が主辞となり，“X” に “and” と “Y” が依存することになる．それに対して，日本語「XとY」では「Xと」が「Y」に依存する．この問題を解決する方法として，等位接続詞を含む構文規則の主辞を後方の “Y” とするように主辞付与規則を修正することが考えられる．これを採用すると，33 文のうちの 20 文 (60.6%) について問題の改善が見込まれる．特に，このような等位接続詞は長い文に含まれることが多く，その対処により長い文の翻訳精度が向上することが期待できる．

一方，Wh-疑問詞を含む文における語順誤りは，“Which is the latest flight?”（標準訳：「最終便のフライトはどちらですか」）のように，補語が主語よりも先に出現する文において生じるものであり，依存構造に従って訳出すると be 動詞と補語が分離することになるためである．この問題を解消するには，Wh-疑問詞と be 動詞をまとめる変換規則を追加すればよい．

なお，語彙的曖昧性や不適格性の問題は，同時翻訳に特有のものではなく，すべて一般的な対話翻訳においても同様に生じる原因である．

表 3.5: 翻訳誤りの原因

原因	文数
語順の問題	74
語彙の曖昧性	61
不適格性	57
その他	8

表 3.6: 同時性の評価

手法	平均訳出遅延 チャンク数	話者・通訳者 発話時間 (秒)
本手法 (後方修飾性のみ)	2.13	5.78
本手法 (後方修飾性 + 倒置 $L=2$)	2.08	5.75
文単位の手法	3.61	6.88

同時性

入力との同時性を評価するために、平均訳出遅延チャンク数と平均話者・通訳者発話時間を計測した。本手法の後方修飾性の出力制御のみを利用して訳出した場合、本手法の後方修飾性の出力制御と倒置を活用 ($L = 2$) して訳出した場合、文単位の手法で訳出した場合の平均遅延チャンク数と平均話者・通訳者発話時間を表 3.6 に示す。文単位の手法に比べて、本手法では平均訳出遅延チャンク数、話者・通訳者発話時間がともに減少しており、入力との同時性が向上したことが確認できた。また、倒置を活用する場合の方が倒置を活用しない場合に比べて、平均訳出遅延チャンク数が僅かであるが減少しており、倒置の活用が同時性の向上に有効であることが確認できた。

本手法と文単位の手法の場合の話者発話時間と話者発話終了から訳出終了までの時間の関係を図 3.15 に示す。図 3.15 では、本手法と文単位の手法の回帰方程式を与えた³。図 3.15 より、文単位の手法に比べ本手法を用いることにより、話者発話時間が長くなるほど訳出終了までの時間が減少しており、本手法は長い発話であるほど同時性の向上が期待できることを確認した。

³本手法の回帰方程式は対数回帰で、文単位の手法は線形回帰で回帰法的式をもとめている。これは、線形回帰と対数回帰で回帰方程式を計算したときに、決定係数が大きくなる回帰方程式をそれぞれ採用した。決定係数が大きいほど、実際のデータが回帰直線の周辺に分布していることを意味している。

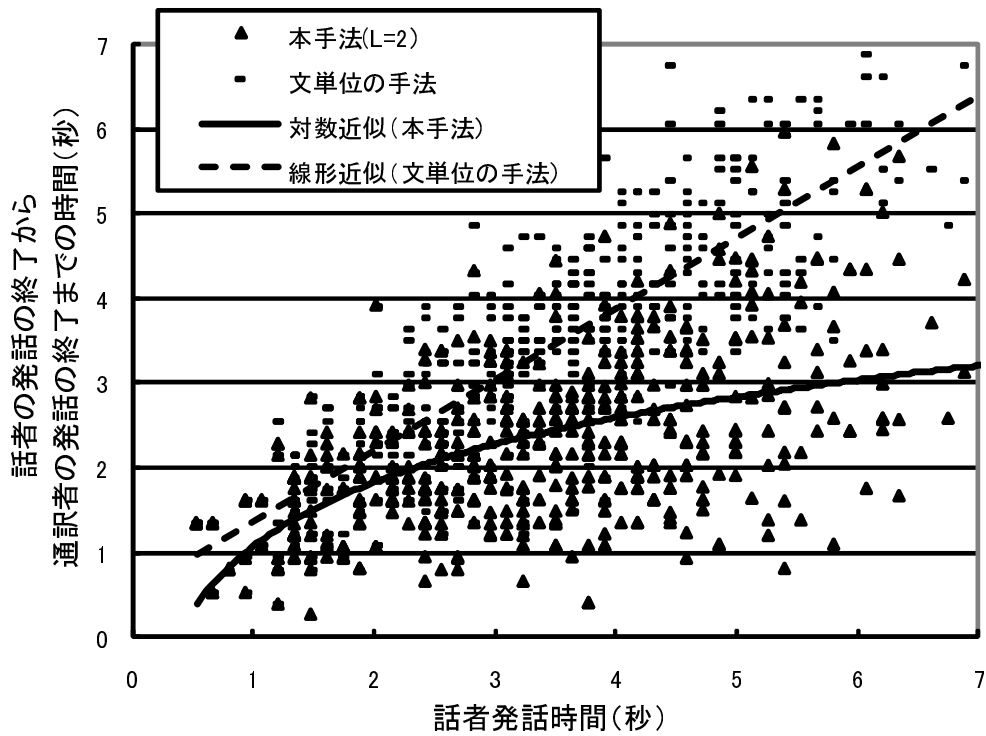


図 3.15: 話者発話時間と話者発話終了から訳出終了までの時間の関係

表 3.7: 述部の訳出制御の設定を $L = 1, 2, 3$ にした場合の実験結果

L の値	翻訳精度			平均訳出遅延 チャンク数	倒置を 含む文数	言い直しを 含む文数
	A	A+B	A+B+C			
1	58 (10.0%)	200 (34.6%)	347 (60.0%)	1.87	270	15
2	111 (19.2%)	312 (54.0%)	411 (71.1%)	2.08	62	1
3	116 (20.1%)	345 (59.7%)	414 (71.6%)	2.12	18	0

倒置の活用

述部の出力条件の L の値を 1 から 3 まで変更した場合の実験結果を図 3.7 に示す。 $L=1$ と $L=2$ の場合を比較すると、 $L=1$ では 270 文で倒置が活用されており、同時性は向上するものの、翻訳精度は 11.1% 低下した。 $L=2$ と $L=3$ の場合を比較すると、 $L=2$ の方が僅かに同時性が向上する一方で、翻訳精度の低下はほとんど見られ

なかった。

3.7 まとめ

本章では、日本語訳の構成要素間の依存構造に基づいて訳文を生成する英日同時翻訳手法を提案した。本手法を導入した英日同時翻訳システムを用いて、ATIS コーパス 578 文の翻訳実験を行った。その結果、本手法による翻訳精度は 71.6% であり、翻訳品質において、本手法の利用可能性を確認した。また、文単位の手法と比較して、本手法の方が遅延チャンク数と話者・通訳発話時間が減少しており、同時性が向上することを示した。

本章では、日本語訳文の生成において、語順の入れ換えを積極的に行うことにより翻訳処理の高い同時性を実現する手法を提案した。しかし、すべての対話文に対して本手法を適用すると、結果的に倒置や言い直しが多頻度で現れることになり、翻訳品質の観点からは必ずしも好ましくない。したがって、本手法を状況に応じて選択的に利用することが考えられる。例えば、長い文を翻訳する場合、そのまま逐次的に訳出すると遅延が大きくなる傾向にあるため、本手法を適用することが有効であると予想される。対話文の性質と手法の効果の程度との関係について、今後、考察していく予定である。

第4章 同時翻訳システムにおける話速制御のための通訳者の発話速度の分析

4.1 はじめに

同時通訳では、話し手の発話の進行に同期しながら訳出を進める通訳スタイルである。話し手の発話とそれに対応する訳との時間差が短いほど、より自然なコミュニケーションを実現できるが、一方で、時間差が短過ぎると、適切な訳を生成するのに必要な原発話の内容が得られないため、自然な訳文を生成できない場合がある。そのため、適切なタイミングで訳出することが求められる。講演などの独話では、話者が発話し続けるため、適切なタイミングで訳出するには、訳出する速度を適切に調整することが求められる。そのため、同時通訳システムを構築する上で、同時通訳者の話速の様相を分析し、分析から得られた知見をシステムの話速制御に活用することが考えられる。

しかしながら、これまで同時通訳者の話速の変化について分析した研究はほとんどない。例えば、木佐は、放送通訳における通訳の聞きやすさと話速との関係について [27]、また、松山は、逐次通訳における訳出時間との関連において通訳者の発話速度について [43] それぞれ分析している。しかし、これらはいずれも、同時通訳を対象としておらず、また、講演あるいは放送全体の話速のみを扱っており、話速の様相については言及していない。このような現状の理由として、同時通訳者の話速の変化を定量的に分析するには、通訳者による発声タイミングが細かい粒度で付与されたデータが大量に必要であり、これまでそのような大規模データを利用することは難しかったことが挙げられる。

本章では、英日同時通訳における通訳者の話速に関する定量的分析について述べる。本研究では、特に、同時通訳者の話速に影響を与えると考えられるいくつかの要因に焦点をあて、その影響の程度について分析を与える。具体的には、大局的な要因として、「講演の話速」及び「通訳者の個性」、また、局所的な要因として、講

演中の「話し手の発話状態」及び「訳出の進行状況」に着目し、話速データを比較することによりそれらの要因が及ぼす影響の程度を明らかにする。

定量的分析のために、本研究では、同時通訳データベースの英日通訳データを用いた。分析では、このデータベースの中から22の英語講演データに対する複数の通訳者による88例の英日同時通訳データを抽出し、使用した。また、音声認識ツールを用いて発声時刻情報を自動推定することにより、単語レベルでの話速データを作成し、分析に使用した。

本章の構成は以下の通りである。次の4.2節で、同時通訳の話速に影響を及ぼす要因について論じる。4.3節では、分析に使用したデータについて概説し、4.4節で大局的要因に着目した分析の結果について、4.5節で局所的要因に着目した分析の結果についてそれぞれ述べる。

4.2 同時通訳者の話速に影響を与える要因

同時通訳は、話し手の発話を聞きながらその翻訳を発声するという、極めて特殊な行為であり、通訳者の話速は様々な要因の影響を受けることになる。本研究では、同時通訳が行われる講演の設定に関わる要因（以下、大局的要因）と同時通訳の進行に応じて変化していく状況・状態に関わる要因（以下、局所的要因）に着目し、それらが話速に及ぼす影響の程度について、定量的観点から分析を与える。具体的には以下の要因に関して分析する。

- 大局的要因

同時通訳者の話速に対して、まず、話し手の話速が大いに影響することが予想される。一般には、同一規模の発話をより短い時間で遂行するためには、発話の速度を高めざるを得ないためである。しかし、その一方で、通訳者の個性として、早口で話す通訳者もいれば、ゆっくりとした速度で話す通訳者も存在する。大局的要因として、話し手の話速と通訳者の個性に注目し、その影響の程度を調査する。

- 局所的要因

講演の同時通訳が進んでいく中で、同時通訳者が置かれる状況や状態は刻一刻と変化する。同時通訳とはいっても、実際には、原発話に数秒程度遅れながら訳出が遂行されるため [52]、通訳者にとっては、原発話が完了していない状況（内容が確定していない状況）と完了している状況（内容が確定している状況）

表 4.1: 通訳者の属性

属性	属性値
性別	男 2 人, 女 15 人
年齢	29 ~ 56 歳 (平均 42.8 歳)
通訳歴	5 ~ 30 年 (平均 14.5 年)

表 4.2: 分析データの規模

項目	英語講演	英日通訳
講演数	22	88
発話単位数	5,053	18,414
単語数 (形態素数)	28,065	141,179

が存在する。このうち、確定していない状況での通訳は、同時通訳に特有な現象である。また、一方で、話し手が講演している中でも、間をあけるときの息継ぎするとき、話す内容を考えるとき、など、話し手の発話が停止している時間が存在する。同時通訳者は通常「聴く」行為と「話す」行為を同時に遂行することが必須であるが、上記のような時間は、通訳者は「話す」行為に徹することができる状態にある。そこで局所的要因として、原発話の進行状況（発話内容の確定前 / 確定後）、及び、話し手の状態（話し中 / 休止中）に着目し、その影響を調査する。

4.3 分析データ

同時通訳者の話速変動を定量的に分析するために、同時通訳データベース [38] を使用した。本研究では、独話データのうち、英語講演（22 講演）とその英日同時通訳を対象とした。各講演の長さは、8 分 ~ 11 分の間である。英日同時通訳には、17 名の通訳者を起用しており、いずれの講演もそれぞれ 4 人の異なる同時通訳者によって通訳が遂行されている。ただし、17 名の通訳者のうち 2 人の通訳者は 1 講演だけ、それ以外の 15 人の通訳者は複数の講演（4 ~ 12 講演）に対して通訳している。通訳者の年齢、性別、通訳歴を表 4.1 に示す。また、分析に使用したデータの規模を表 4.2 に示す。

同時通訳データベースにおける文字化は、日本語話し言葉コーパス (CSJ)[35] の

書き起こし基準に準拠しており，講演者ならびに通訳者の音声を 200ms 以上のポーズ(無音)で分割して得られる音声区間を発話単位と定めている．すべての発話単位に対して，開始時刻と終了時刻が人手で付与されている．また，講演者発話と通訳者発話との対訳対応について人手で分析し，発話単位を最小単位とした対応付けが与えられている [63] ．

4.3.1 単語発声時刻の推定

話速の変動をきめ細かく分析するためには，発話単位に与えられた時刻情報では十分でなく，単語などの細かい単位での情報が不可欠である．しかし，大量のデータに対する単語レベルでの時刻情報の付与を人手で行うことは現実的でなく，自動化することが望ましい．

本研究では，音声認識エンジンを用いた発声時刻推定を実行し，単語発声時刻を決定した．本分析では，英語話者と英日通訳者の発話，すなわち，英語発話と日本語発話の単語発声時刻を推定した．以下では，日本語と英語の場合の推定方法について説明する．

日本語音声における発声時刻の推定

書き起こしデータの発話の開始時刻と終了時刻の情報を用いて，発話単位の文字データと音声データに分割し，以下の手順で発話ごとに単語発声時刻を推定する．
図 4.1 に日本語音声における単語発声時刻を推定の処理の流れを示す．

1. 形態素分割と読みの付与

書き起こしデータの発話を形態素解析器茶筌 [42] を用いて，形態素に分割し，カタカタの読みを付与する．

2. 音素変換

音声認識エンジン Julius に付属の読み付与ツールキットの規則に基づいて，茶筌が出力するカタカナの読みを音素列に変換する．

3. 形態素単位で時間情報を付与

音声認識エンジン Julius を用いて音素列と音声を形態素単位で対応付ける．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．

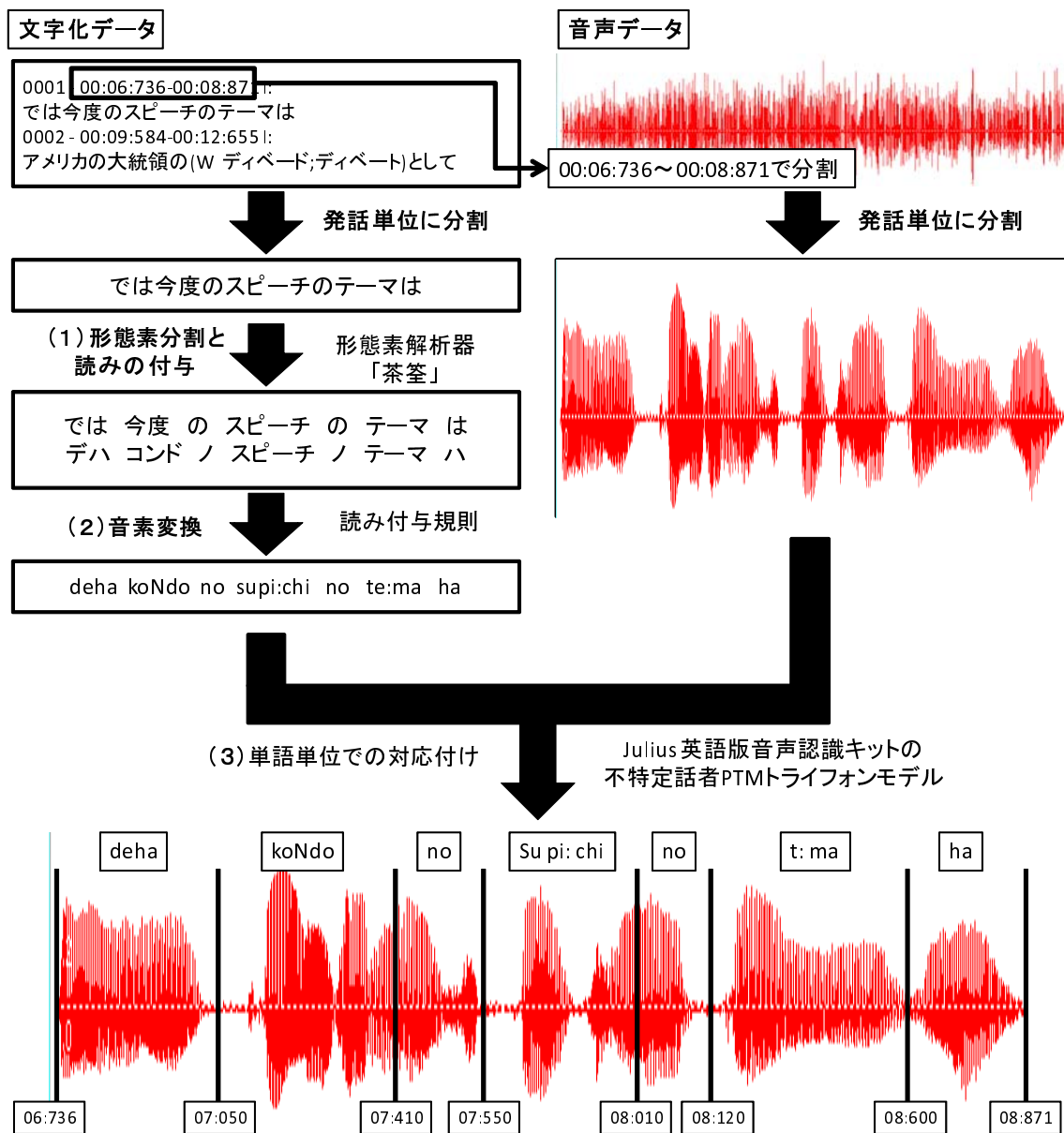


図 4.1: 単語発声時刻の推定 (日本語)

英語音声における単語発声時刻の推定

書き起こしデータの発話の開始時刻と終了時刻の情報を用いて、発話単位の文字データと音声データに分割し、以下の手順で発話ごとに単語発声時刻を推定する。図 4.1 に英語音声における単語発声時刻を推定の処理の流れを示す。

1. 単語分割

書き起こしデータの発話を空白文字により単語に分割する。

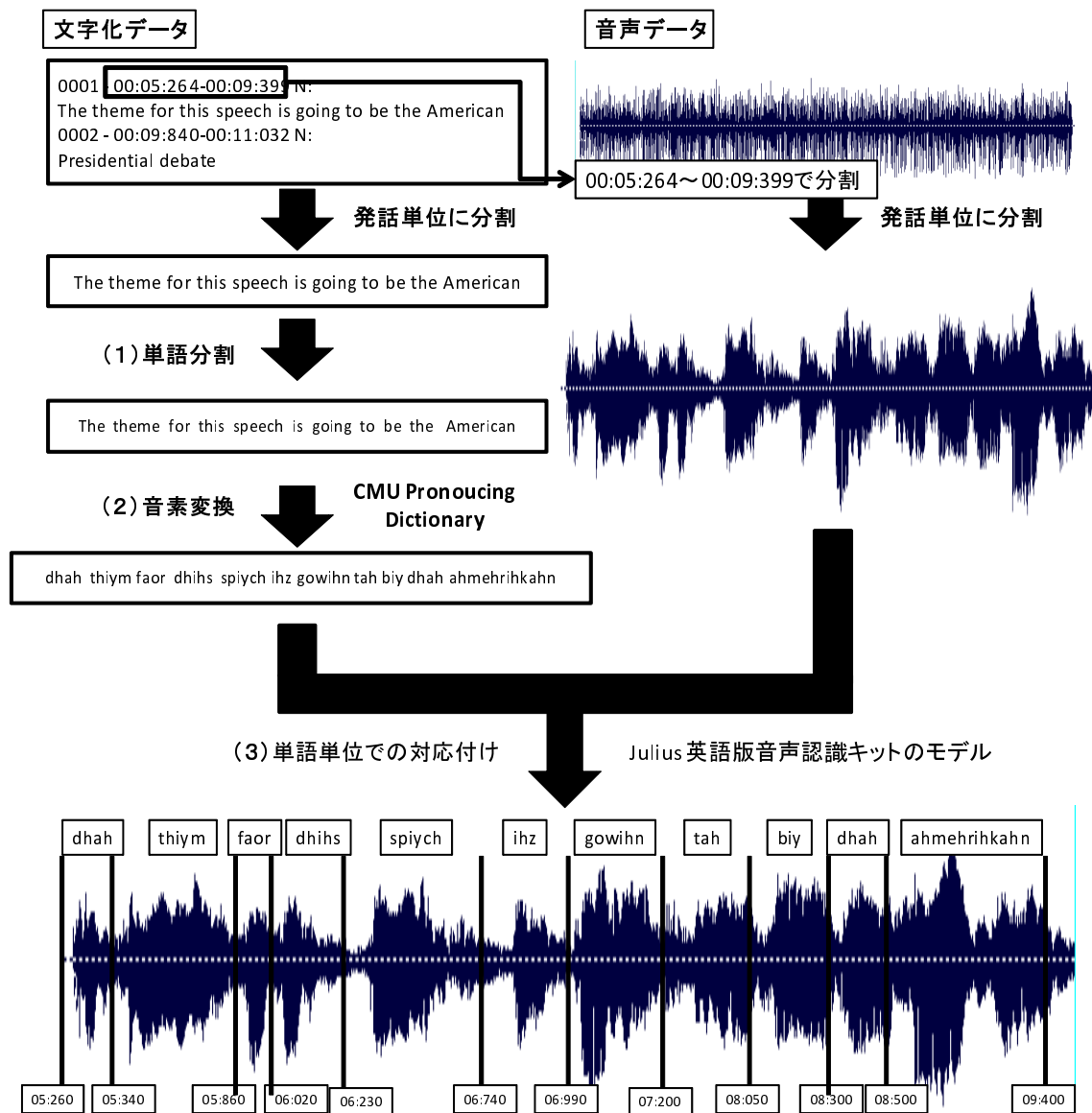


図 4.2: 単語発声時刻の推定 (英語)

2. 音素変換

The CMU Pronouncing Dictionary version 0.6[19] を用いて，単語を音素列に変換する．

3. 単語単位で時間情報を付与

音声認識エンジン Julius を用いて音素列と音声形態素単位で対応付ける．音響モデルには，Julius 英語版音声認識キットのモデルを用いた．

上記の方法で，英語講演者による単語，ならびに，英日通訳者による形態素に時刻情報を付与した．付与した単語発声時刻データのサンプルを図 4.3 に示す．付与

発話ID-単語ID	出現形	読み	音素記号	開始時刻(秒)	終了時刻(秒)
0001-0001	では	デハ	de ha	6.74	7.05
0001-0002	今度	コンド	ko Ndo	7.05	7.41
0001-0003	の	ノ	no	7.41	7.55
0001-0004	スピーチ	スピーチ	su pi: chi	7.55	8.01
0001-0005	の	ノ	no	8.01	8.12
0001-0006	テーマ	テーマ	te: ma	8.12	8.60
0001-0007	は	ハ	ha	8.60	8.87
0002-0001	アメリカ	アメリカ	amerika	9.58	10.17
0002-0002	の	ノ	no	10.17	10.32
0002-0003	大統領	ダイトウリョウ	daitou ryou	10.32	11.15
0002-0004	の	ノ	no	11.15	11.46
0002-0005	ディベード	ディベード	dibe: do	11.46	12.04
0002-0006	として	トシテ	to shite	12.04	12.65
0003-0001	どっち	ドッチ	do qchi	13.66	14.00
0003-0002	が	ガ	ga	14.00	14.30
0004-0001	アメリカ	アメリカ	amerika	15.18	15.56
0004-0002	にとって	ニトツテ	nito qte	15.56	15.93
0004-0003	いい	イイ	ii	15.93	16.09
0004-0004	大統領	ダイトウリョウ	daitou ryou	16.09	16.66
0004-0005	か	カ	ka	16.66	16.82
0004-0006	と	ト	to	16.82	17.09
0005-0001	いう	イウ	iu	17.30	17.41

図 4.3: 単語発声時刻データのサンプル

表 4.3: 単語発声時刻の推定精度

	日本語	英語
平均誤差	0.028	0.033

率は、英語で 95.0% (21,863/23,012)、日本語で 98.7%(121,219/122,765) であった。この方式は、音響データと読みデータの対応付けにより時刻を推定する方法であり、不明瞭に発声された単語や読み方が間違っている単語に対しては、時刻情報を付与することができない。付与された時刻情報の精度を、無作為に選んだ英語及び日本語の 20 発話単位に対して調べた。表 4.3 に単語発声時刻の推定精度を示す。両言語とも人手で付与したデータとの差は平均で 0.03 秒程度と僅かであり、自動付与した時刻情報の利用可能性を確認した。

表 4.4: 英語講演者の話速の基礎統計

項目	数値
シラブル数	37,360
発話時間(秒)	9,428
平均話速(シラブル/秒)	3.96

表 4.5: 英日同時通訳者の話速の基礎統計

項目	数値
モーラ数	291,277
発話時間(秒)	36,163
平均話速(モーラ/秒)	8.05

4.3.2 話速の計算

4.3.1 節で作成した単語発声時刻の情報を用いて、講演者と通訳者の平均話速を算出した。話速の計算には、日本語はモーラ、英語はシラブルを用い、それぞれの単位を「モーラ/秒」と「シラブル/秒」で表す。シラブル(syllable)は、母音を中心にして、母音の前後に1個または複数個の子音を伴って構成するものである。シラブルの計算には Syllable Counter というツールを利用した。

分析に用いたデータの講演者と通訳者の話速の基礎統計をそれぞれ表 4.4、表 4.5 に示す。

4.3.3 話速変動の視覚化

話速データの数値を観るだけででは、同時通訳者の話速の変化の様子を捉えることは困難である。また、通訳者の話速と講演者の話速の変化の様子を視覚化して、同時に表示して観察すると、講演者の話速と通訳者の話速の関係を観察しやすい。そこで、話速を視覚的に捉えることができるツールを用いて作成した。話速の視覚化ツールを図 4.4 に示す。本ツールは、講演者の話速とその講演を通訳した4人の通訳者の話速の様子を同時に表示することができる。また、観察目的によって、講演者と1人の通訳者だけ表示させたり、2人の通訳者だけを表示したりすることが可能である。また、英語と日本語の話速の単位が異なることを考慮して、平均話速

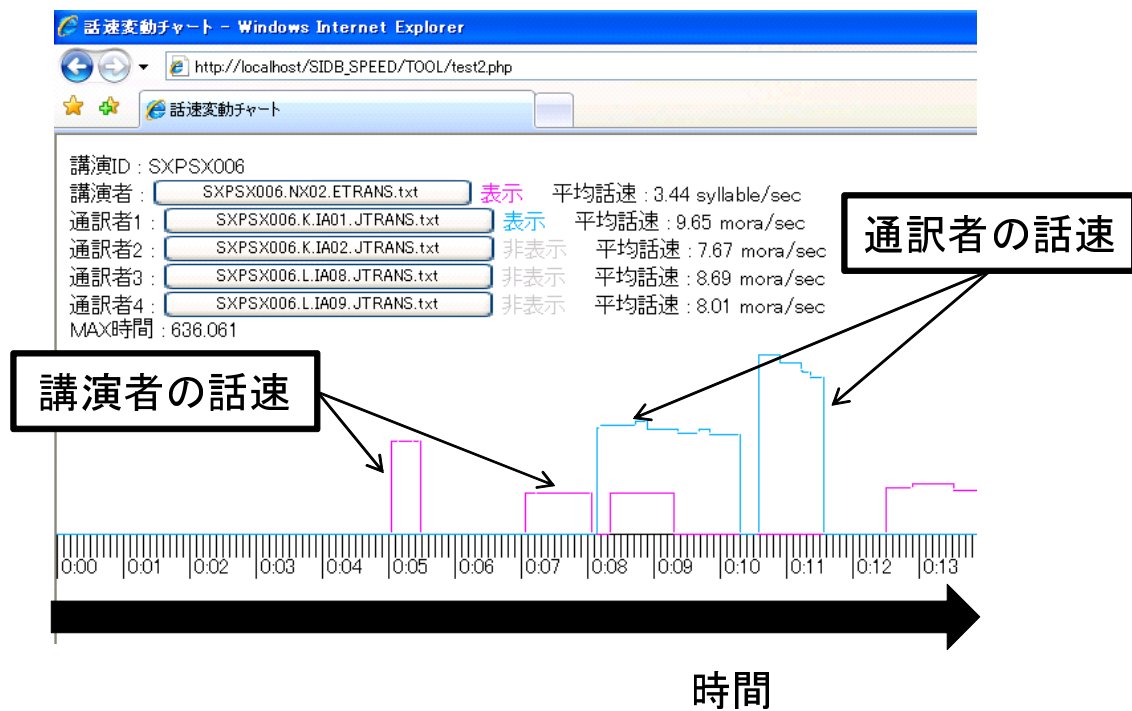


図 4.4: 話速の視覚化ツール

に対する比で表示させることも可能である。

4.4 講演の設定に関わる要因が同時通訳者の話速に及ぼす影響

本節では、大局的要因が同時通訳者の話速に及ぼす影響の分析として、「講演者の話速」による影響、及び「通訳者の個性」による影響を調査する。図 4.5 に 88 講演分の通訳データにおける通訳者の平均話速の分布を示す。各講演における通訳者の話速の平均値は 8.10 (モーラ/秒) であった。また、各講演における通訳者の話速は 6.93 ~ 9.28 (モーラ/秒) の間に分布しており、通訳者の話速は通訳者や講演ごとに大きく異なる。

4.4.1 講演者の話速と通訳者の話速との関係

講演ごとの講演者の平均話速と通訳者の平均話速の関係を図 4.6 に示す。図 4.6 の横軸は講演者の話速、縦軸は通訳者の話速を示している。図 4.6 より、講演者と通訳者の話速に相関関係がほとんどないことがわかる。実際、講演者の話速と通訳者

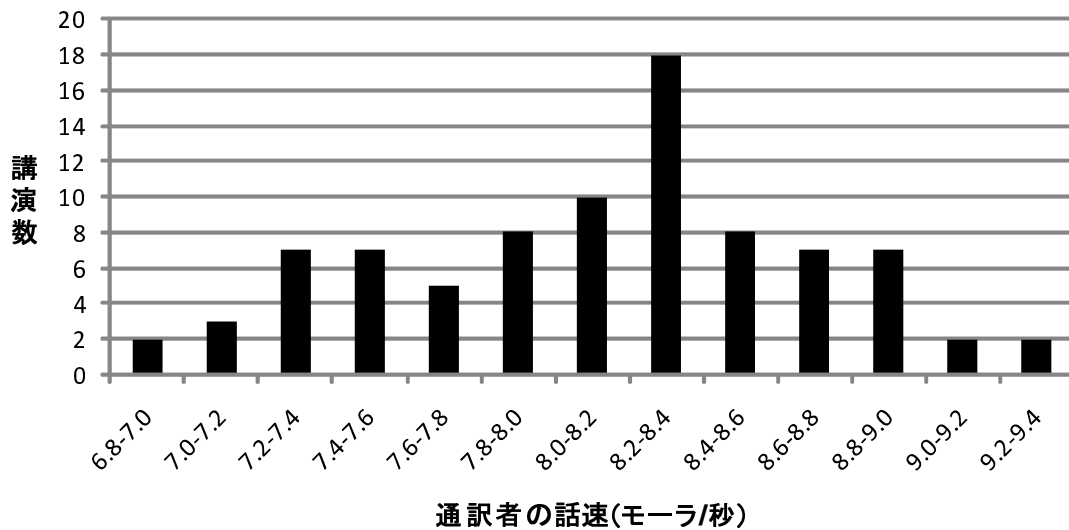


図 4.5: 通訳者の話速の分布

の話速の相関係数¹は0.18であり、講演者の話速と通訳者の話速の間に相関は認められなかった。

一方で、個々の通訳者ごとにみれば、講演者の話速と通訳者の話速との間に相関関係があることも考えられる。しかし、分析データでは、一人の通訳者が通訳した講演数は平均5.2講演であり、統計的な分析を与えるには十分でない。ただし、12講演を通訳した通訳者が一人存在しており、その通訳者に対して講演者の話速と通訳者の話速の相関関係を調査したところ、相関係数は0.05であった。一人の通訳者のみのデータであり、参考に過ぎないものの、個々の通訳者で観察した場合にも講演者の話速と通訳者の話速の間に相関関係は認められなかった。以上より、同時通訳者の話速に対する講演者の話速の影響は大きくないことが明らかになった。

4.4.2 通訳者の個性と通訳者の話速との関係

通訳者の話速は、通訳者によって大きく異なると考えられる。通訳者の個性と通訳者の話速との関係を分析するために、通訳者ごとの話速の分布を調査した。複数

¹相関係数：本分析ではピアソンの積率相関係数を用いている。相関係数の絶対値が示す相関の強さの目安を以下に示す。

0.2 未満	：ほとんど無相関
0.2 以上～0.4 未満	：弱い相関がある
0.4 以上～0.7 未満	：中程度の相関がある
0.7 以上	：強い相関がある

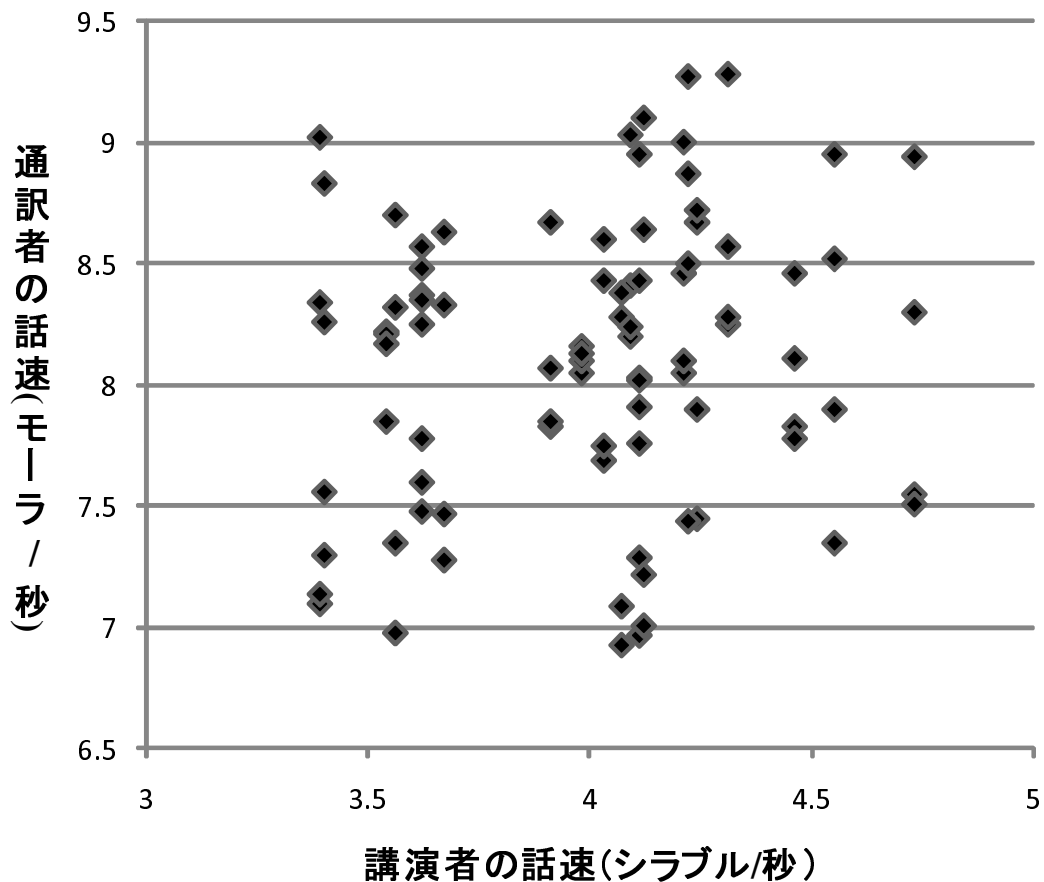


図 4.6: 講演者の話速と通訳者の話速の関係

の講演を通訳した15人の通訳者の講演ごとの平均話速の分布を図4.7に示す。図4.7の横軸は通訳者の平均話速で昇順にソートしたものであり、奥行き軸が通訳者の話速を、縦軸が頻度を表している。図4.7より、ある通訳者における講演の違いによる話速の偏りに比べ、通訳者の違いによる話速の偏りの方が大きいことがわかる。

各通訳者の話速の間に有意な差があることを確認するため、一元配置分散分析を用いて検証した。一元配置分散分析とは、3つ以上の平均値を比較するための統計的手法であり、標本全体のばらつきを群間の違いによって説明できる部分と説明できない部分に分解し、群間の違いによって説明できるばらつきが、それ以外によるばらつきに比べて十分に大きければ、標本のばらつきが生じた原因を群間の違いによるものだとするという考え方に基づいている。本分析では、群間の違いによって説

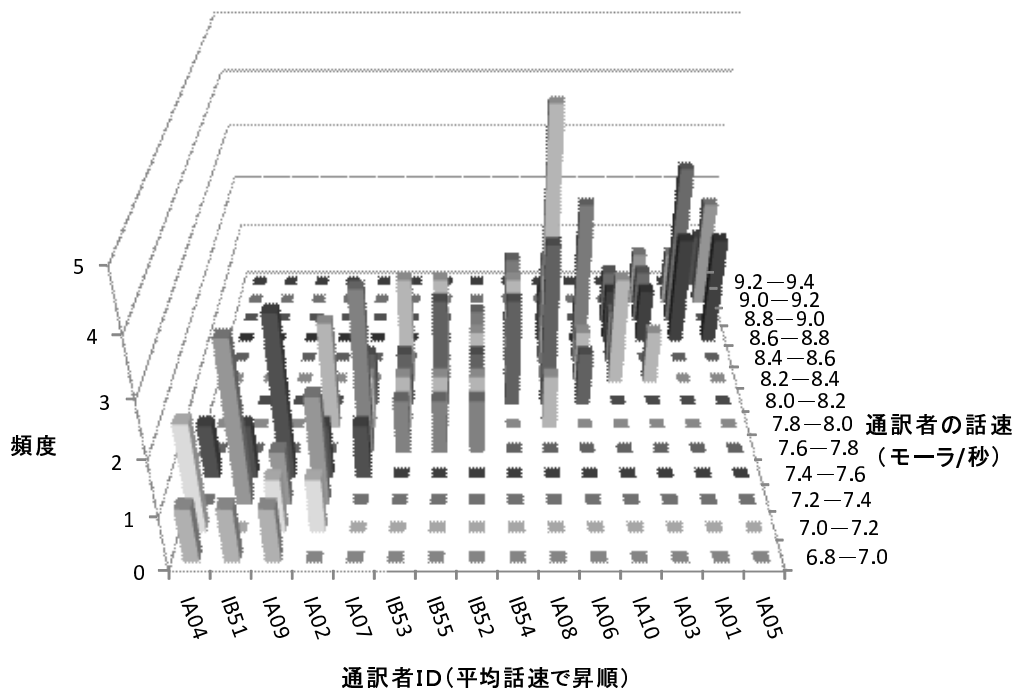


図 4.7: 通訳者と通訳者の話速の分布

明できるばらつきが，通訳者間の違いによって説明できるばらつきに対応する．全体の平均値を A ，通訳者の集合を T ，通訳者の話速の平均値を a_t ，通訳者の人数を n とすると，通訳者間の違いによるばらつきは以下の式で計算される．

$$\frac{\sum_{t \in T} (A - a_t)^2}{n - 1} \quad (4.1)$$

上記の式で計算される値は群間平均平方と呼ばれる．一方，群間の違い以外によるばらつきは，本分析では通訳者内の講演の違いによるばらつきである．通訳者 t が通訳した講演の集合を S_t ，通訳者 t が講演 s を通訳したときの通訳者の話速を a_{ts} ，通訳者 t が通訳した講演数を m_t とすると，通訳者内の講演の違いによるばらつきは，以下の式で計算される．

$$\frac{\sum_{t \in T} \sum_{s \in S} (a_t - a_{ts})^2}{\sum_{t \in T} (m_t - 1)} \quad (4.2)$$

上記の式で計算される値は群内平均平方と呼ばれる．本データの群間平均平方は 1.596，群内平均平方は 0.058 であった．群間のばらつきが群間以外のばらつきに比べて十分大きいことを示すには，群間平均平方を群内平均平方で割った値 (F 値と呼ばれる) が群間の自由度，群内の自由度，有意水準により決まる F 分布の値に比べて

大きければよい。本データの F 値は 27.5 であった。本データの群間と群内の自由度がそれぞれ 16 と 71 であり、 $F(16, 71, 0.01) \leq F(16, 60, 0.01) = 2.84$ なので、有意水準 1% で各通訳者の話速の平均値の間に有意な差が認められた。このことから、通訳者によって話速が異なることが統計的にも示された。

4.5 講演の進展により変化する要因が通訳者の話速に及ぼす影響

本節では、同時通訳者の話速に影響を及ぼす講演内の要因として、「原発話の進行状況」と「講演者の発話状態」に着目し、同時通訳者の話速に及ぼす影響を調査した。

4.5.1 原発話の進行状況と話速変動の関係

同時通訳では、原発話が終わらない段階で、その原発話に対する訳出を開始することになる。原発話が終了する前の通訳者発話では、訳出内容が確定していないため、訳出内容の確定状況に応じて話速を制御し、話速が遅くなると考えられる。一方、原発話が終了した後は、訳出内容が確定しているため、話速が速くなると考えられる。このように、原発話の進行状況に応じて、同時通訳者は発話速度を柔軟に制御していると思われる。そこで、通訳者の発話を、その原発話が終了する前（確定前）と終了した後（確定後）という二つの状況に分割し、話速を比較分析した。本分析では、通訳者発話と対訳関係にある原発話との対訳対応情報が必要となる。対訳対応情報には、発話単位をもとに人手により付与したものを利用した。図 4.8 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關係を示す。この例では、講演者の発話“ and that’s where the melody comes from ”が通訳者発話「ですのでそこからメロディーが出てきます」と対訳対応関係にあり、講演者の発話が 01:30:99 に終了しているため、通訳者発話の内、01:30:30 から 01:30:99 までに発話を開始した単語列「ですのでそこ」は確定前、それ以降の単語列「からメロディーが出てきます」は確定後となる。

分析結果を表 4.6 に示す。原発話が確定している場合の平均話速は 8.40(モーラ/秒)であり、原発話が未確定である場合の平均話速 7.59(モーラ/秒) に比べ 11% 高かった。この結果は、訳出内容が確定した後の方が確定する前に比べ話速が速くなるという直観と一致している。

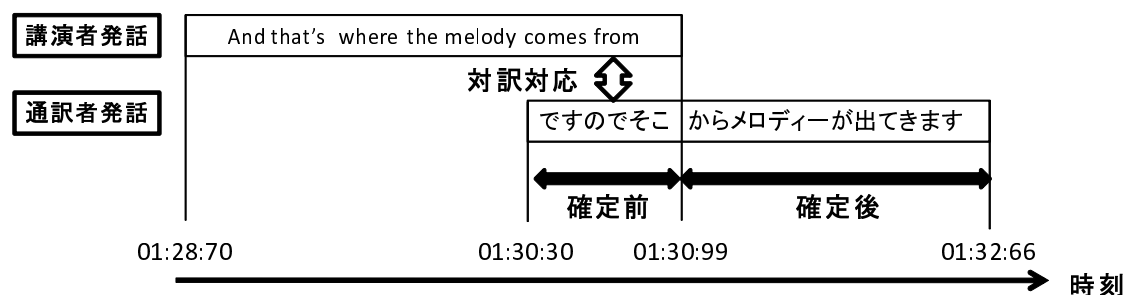


図 4.8: 原発話の進行状況 (原発話の確定前の通訳者発話と確定後の通訳者発話)

表 4.6: 原発話の進行状況と通訳者の話速

	確定前	確定後
モーラ数	103,002	171,113
発話時間 (秒)	13,574	20,380
平均話速 (モーラ/秒)	7.59	8.40

4.5.2 講演者の発話状態と話速変動の関係

同時通訳者は、講演者の発話を聴き、その発話内容を翻訳し、翻訳結果を発話している。すなわち、通訳者発話は、講演者が発話中の場合と休止中の場合とでは、休止中の方が翻訳や翻訳結果の発話に集中できるため、話速が速くなると予測できる。また、講演者の発話が速くなると、同時通訳者の話速も速くなると予想できる。そこで、まず、同時通訳者の発話を講演者が発話している状態（発話中）と発話していない状態（休止中）に分類し、通訳者の話速を分析した。図 4.9 に講演者と通訳者の発話の時間の流れと講演者の発話状態を示す。分析結果を表 4.7 に示す。講演者の発話状態が休止中の場合の通訳者の平均話速は 8.34(モーラ/秒) であり、講演者が発話中の通訳者の平均話速 7.94(モーラ/秒) に比べて 5%向上した。これは、講演者の発話が休止中の場合の方が通訳に専念でき話速が速くなるという直観と一致する結果であった。次に、講演者が発話中の同時通訳者の発話を、講演者の話速がその講演者の平均話速よりも速い場合と遅い場合に分類し、通訳者の話速を分析した。ただし、講演者の話速が速い場合と遅い場合の分類は単語単位で行い、各単語の話速の計算には、前後 2 単語の加重平均を用いた。分析結果を表 4.8 に示す。講演者の話速が速い場合の通訳者の平均話速は 7.99(モーラ/秒) であり、講演者の話速が遅い場合の通訳者の平均話速 7.90(モーラ/秒) とほとんど変わらなかった。す

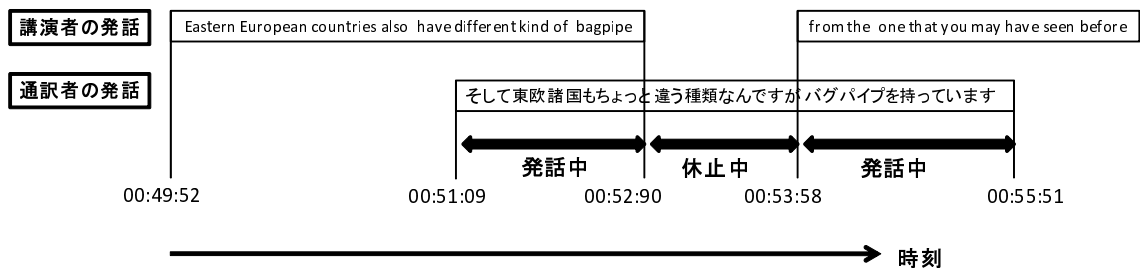


図 4.9: 講演者の発話状態 (発話中と休止中)

表 4.7: 講演者の発話状態と通訳者の話速

	発話中	休止中
モーラ数	207,155	84,122
発話時間 (秒)	26,075	10,089
平均話速 (モーラ/秒)	7.94	8.34

表 4.8: 講演者の発話状態 (速い場合と遅い場合) と通訳者の話速

	速い	遅い
モーラ数	103,463	103,692
発話時間 (秒)	12,951	13,124
平均話速 (モーラ/秒)	7.99	7.90

なわち，講演者の話速が速い場合と遅い場合では，通訳者の平均話速はほとんど変わらないことが明らかになった．

4.6 まとめ

本章では，同時通訳における通訳者の話速に影響を与える要因を分析した．分析では，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用いた．講演全体の通訳者の平均話速に与える要因として，講演者の話速，通訳者の個性に着目して分析した．その結果，以下のことを確認した．

- 講演者の話速と通訳者の話速の間の相関係数は 0.18 であり，相関は認められない．

- 各通訳者の話速の間に統計的に有意な差があり，講演の違いによる違いに比べ通訳者による話速の違いが大きい．

さらに，講演内で通訳者の話速に影響を与える要因として，原発話の進行状況と講演者の発話状態に着目して，定量的に分析した．通訳者発話を原発話の進行状況や講演者の発話状態に基づいて分類するために，通訳者発話，及び，講演者発話に対して単語単位の発声時刻を自動的に付与した．本分析により，同時通訳者の話速には以下の特徴があることを定量的に確認した．

- 原発話が確定後の平均話速 8.40(モーラ/秒) は，原発話が確定前の平均話速 7.59(モーラ/秒) に比べて 11% 高く，確定後の方が確定前より通訳者の話速が速い．
- 講演者の発話状態が休止中の場合の通訳者の平均話速は 8.34(モーラ/秒) であり，講演者が発話中の通訳者の平均話速 7.94(モーラ/秒) に比べて 5% 高く，講演者の発話状態が休止中の方が発話中より話速が高い．
- 講演者の話速が速い場合の平均話速は 7.99(モーラ/秒) であり，講演者の話速が遅い場合の平均話速 7.90(モーラ/秒) とほとんど差がない．

分析結果から，同時通訳者の話速は，講演者の話速にほとんど影響を受けておらず，各自の話速を維持するように制御していることが推測される．また，同時通訳者は，原発話の内容が確定しているか否か，あるいは，講演者が発話しているか否かとの関連で，話速の制御に関する方略を行使していることが結論づけられる．

以上の成果から，同時通訳者システムの訳出速度の制御方法として，講演全体の平均話速を講演に関係なく一定に保ちつつ，原発話の確定後に話速を上げ，休止中に話速を上げることが有効な方法であることが明らかになった．

なお，本研究の成果は，通訳学の観点からも大いに意義がある．というのも，同時通訳に関する方法論が十分に体系化されていない現状においては，同時通訳者の振る舞いに関する科学的知見を積み上げていくことが極めて重要であるためである．本研究は通訳者の話速に関する現象の一部を捉えるものであり，今後の体系化に役立つものであると考えられる．

第5章 あとがき

5.1 本論文のまとめ

本論文では、同時通訳システムの開発を目的に、同時通訳データベースを構築し、実際の同時通訳者の訳出方略を用いた訳文生成手法を提案した、また、英日同時通訳者の話速に影響を与える要因を大域的な要因と局所的な要因に分けて定量的に分析した。

まず、第1章では、同時通訳システムに関する研究を概観するとともに、現状において、同時通訳システムの開発を目指した研究はあるものの、その数は少なく、実用的な同時通訳システムを開発する段階にはないことを述べた。また、同時通訳システムの開発に向けて、大規模な同時通訳者の発話データを収録したコーパスを構築し、実際の同時通訳者の訳出メカニズムを分析し、システムへの応用を試みるアプローチの必要性について論じた。

第2章では、同時通訳データベースの設計、収集、開発、及び、利用について述べた。本データベースは、182時間、100万単語を収録しており、世界最大規模の対訳音声コーパスである。構築の目的は、異言語間コミュニケーション支援環境の実現を目指した、音声翻訳技術の向上、及び、同時通訳理論の構築であり、現在も、音声・言語処理技術を活用して、有用な言語情報のアノテーションを行っている。

第3章では、日本語訳の構成要素間の依存構造に基づいて訳文を生成する英日同時翻訳手法を開発した。本手法では、日本語の語順に関する制約は比較的緩やかであるという特徴を利用し、日本語文として容認可能な範囲内で、原文の語順に準拠した訳文を生成する。訳文としての容認可能性は、日本語依存文法に基づいて検査する。すなわち、依存文法に基づく制約（特に、後方修飾性制約）を満たし、かつ、英文の語順に最も近い日本語文を同時翻訳における訳文とする。ただし、日本語文として不自然でない場合に限り、倒置、すなわち、後方修飾性を満たさない依存関係を含む訳文についても生成を許容する。本手法を導入した英日同時翻訳システムを用いて、ATIS コーパス 578 文の翻訳実験を行った。その結果、本手法の利用可能性を確認した。また、文単位の手法と比較して、本手法の方が遅延チャンク数と話

者・通訳発話時間が減少しており，同時性が向上することを示した．

第4章では，同時通訳における通訳者の話速に影響を与える要因を分析した．分析では，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用いた．講演全体の通訳者の平均話速に与える要因として，講演者の話速，通訳者の個性に着目して分析した．その結果，講演者の話速と通訳者の話速の間に相関は認められないこと，各通訳者の話速の間に統計的に有意な差があり，講演の違いによる違いに比べ通訳者による話速の違いが大きいことを報告した．さらに，講演内で通訳者の話速に影響を与える要因として，原発話の進行状況と講演者の発話状態に着目して，定量的に分析した．通訳者発話を原発話の進行状況や講演者の発話状態に基づいて分類するために，通訳者発話，及び，講演者発話に対して単語単位の発声時刻を自動的に付与した．本分析により，同時通訳者の話速は，原発話が確定した方が話速が速いこと，講演者の発話状態が休止中の方が話速が速いことを述べた．分析結果から，同時通訳者の話速は，講演者の話速にほとんど影響を受けておらず，各自の話速を維持するように制御していることが推測される．また，同時通訳者は，原発話の内容が確定しているか否か，あるいは，講演者が発話しているか否かとの関連で，話速の制御に関する方略を行使していることが結論づけられる．以上の成果から，同時通訳者システムの訳出速度の制御方法として，講演全体の平均話速を講演に関係なく一定に保ちつつ，原発話の確定後に話速を上げ，休止中に話速を上げることが有効な方法であることが明らかになった．

5.2 今後の課題と将来への展望

本研究が残した課題と将来への展望について述べる．

- 同時翻訳に適した翻訳単位の研究

本論文の3章で提案した英日同時翻訳システムでは，処理単位として句単位を用いた．それに対して，柏岡ら [21] は，日英翻訳の処理単位として，述部を中心としたまとまりである節に着目し，ルールベースで節の境界を高精度に検出している．しかし，同時翻訳に適した処理単位は翻訳する言語によって異なるため，必ずしも句や節といった言語的な単位が適してるとは限らない．著者らは，日英対話翻訳の同時翻訳の処理単位として，対訳対応関係に着目して定義した翻訳単位について検討しており [57]，統計的な検出法により高精度に検出している [58]．さらに，村田ら [46] は，講演の日英同時翻訳のための処理単位を獲得するため，遅延時間を考慮して講演者発話を分割し，同時通訳者が分割

した単位に対して対訳データを付与する作業を進めている。

- 日英同時翻訳に関する研究

本論文の3章で提案した英日同時翻訳手法では、日本語の依存文法を活用した訳文生成手法を用いており、日本語を訳文とする同時翻訳への応用は可能であるものの、日英翻訳のように日本語以外の言語へ翻訳する場合には適用できない。日本語と英語間の対話を実現するためには、日本語から英語への翻訳手法が必要不可欠であり、現在、日英翻訳の研究に取り組んでいる。

- 統計的な同時翻訳手法に関する研究

現在、大規模なコーパスの構築や Web 上の言語データの大規模化により、対訳対応付けと自動評価法 [11, 53, 70] を用いた統計的な翻訳手法 [6, 7] が注目されている。2章で提案した訳文生成手法は、日本語の依存文法を用いたルールベースの翻訳手法であり、適切な翻訳単位や訳出タイミングを統計的に学習する同時翻訳手法を開発することが今後の課題の一つである。

- 同時通訳者の話速に関する分析結果の同時通訳システムへの応用

本論文の4章で実施した同時通訳者の話速の分析では、原発話の進行状況や講演者の発話状態により同時通訳者が話速を制御していることが分かっており、同時通訳システムの話速制御にそれらの知見を組み込むことが今後の課題の一つである。同時通訳システムの話速制御に組み込むには、本研究で明らかになった同時通訳者の話速の変化が聴き手の聞きやすさにどのような影響を与えるかを分析し、話速に関する知見が聴き手の聞きやすさを向上させるものであるかを検証する必要があると考えている。

- 同時通訳者の訳出メカニズムの分析に関する研究

本論文では、同時通訳データベースを構築し、同時通訳者の訳出方略を解明するために、同時通訳者の話速に影響を与える要因を分析した。しかし、同時通訳者の訳出方略を解明するには、訳出遅延時間、訳出タイミング、訳出パターンなど、多様な観点から分析を進める必要がある。また、本分析では、英語から日本語への同時通訳者を対象に分析を行ったが、日本語から英語への同時通訳では、また別の特徴が得られる可能性がある。さらに、講演を対象とした同時通訳の分析を行ったが、独話と対話とでは通訳者の訳出方法も異なることが考えるため、分析の余地がある。

今後の展望としては、同時通訳データベースを用いて対話、及び、独話における同時通訳者の訳出方略を体系的に分析し、対話、及び、独話を対象とした音声認識や音声合成も備えた実用可能な同時通訳システムを開発することが挙げられる。同時通訳システムが実現できれば、近年、急速に進展している音声翻訳システムの実用化への流れが、より一層進むことが期待できる。特に、講演を対象とした音声翻訳システムでは、同時翻訳技術は必須であり、実用化に向けた大きな壁を超えることになる。

謝辞

本論文をまとめるにあたり，多大な御教示と御尽力をいただきました，名古屋大学教授の坂部俊樹先生に厚く謝意を申し上げます。また，本論文の詳細について，貴重な御示唆と御指導をいただきました，名古屋大学教授の阿草清滋先生に深く感謝いたします。

本研究を進め，まとめるにあたり，日頃より懇切丁寧な御指導と御鞭撻をいただきました，名古屋大学准教授の松原茂樹先生に厚く感謝いたします。また，松原茂樹先生には，公私共に様々な御相談にのっていただきました。心より感謝の気持ちを申し上げます。

本研究の初期の段階より幅広い角度から御指導と御討論をいただきました，名古屋大学名誉教授で，現在，豊橋科学技術大学理事・副学長の稲垣康善先生に深く感謝いたします。

本研究を進めるにあたり，また，研究室生活を送るにあたり，御指導と御支援をいただきました，名古屋大学教授の河口信夫先生，助教の山口由紀子先生，加藤芳秀先生，大野誠寛先生に深く感謝いたします。

本研究の初期の段階より有意義な御議論をいただきました，名古屋大学准教授の外山勝彦先生，助教の小川泰弘先生，元助手の杉野花津江先生，名古屋産業大学准教授の福田ムフタル先生に感謝いたします。

多言語翻訳に関する共同研究のため，タイの Kasetsart 大学に滞在したときに，研究，及び，生活の両面で大変お世話になった，Kasetsart 大学准教授の Asanee Kawtrakul 先生とその研究室の皆さまに感謝いたします。

研究に関する討論をはじめ，研究以外の面でもいろいろお世話になりました，松原研究室，河口研究室，旧稲垣研究室の皆様に心から感謝いたします。特に，同時通訳データベースの構築にご協力頂いた，相澤靖之先輩，高木亮先輩，大原誠君，遠山仁美さん，小野貴博君，于海貝さん，同時通訳システムの開発にご尽力頂いた渡邊善之先輩，水野敦君に深く感謝致します。

翻訳結果の評価にご協力頂いた，名古屋大学大学院国際言語文化研究科の修了生で，翻訳家の和氣祥子さんに感謝致します。

研究活動を行うにあたりいろいろとお世話になりました，名古屋大学松原研究室秘書の土井ひとみさん，上松可奈さん，名古屋大学旧稲垣研究室秘書の伊藤千佳子さんに深く感謝いたします．

最後に，改めまして，本論文をまとめるにあたりご支援をいただいたすべての皆様に心より御礼申し上げます．

発表文献リスト

種別	論文名	関連する章
国際会議	Koichiro Ryu, Shigeki Matsubara, Nobuo Kawaguchi, Yasuyoshi Inagaki. Bilingual Speech Dialogue Corpus for Machine Interpretation Research. <i>Proceedings of the Oriental Chapter of International Coordinating Committee on Speech Databases and Speech I/O (O-COCOSDA-2003)</i> , pp. 217-224 (2003).	2章
学術雑誌	笠 浩一朗, 松原 茂樹, 稲垣 康善. 英日同時翻訳のための依存構造に基づく訳文生成手法. 電子情報通信学会論文誌, Vol. J92-D, No. 6, pp. 921-933 (2009).	3章
国際会議	Koichiro Ryu, Shigeki Matsubara, Yasuyoshi Inagaki. Simultaneous English-Japanese Spoken Language Translation Based on Incremental Dependency Parsing and Transfer. <i>Proceedings of the 21st International Conference on Computational Linguistics and the 44th Annual Meeting of the Association for Computational Linguistics (COLING/ACL-2006)</i> , pp. 683-690 (2006).	3章
国際会議	Koichiro Ryu, Atsushi Mizuno, Shigeki Matsubara, Yasuyoshi Inagaki. Incremental Japanese Spoken Language Generation in Simultaneous Machine Interpretation. <i>Proceedings of the Asian Symposium on Natural Language Processing to Overcome Language Barriers</i> , pp. 91-95 (2004).	3章
学術雑誌	笠 浩一朗, 松原 茂樹. 同時通訳者の話速に影響を及ぼす要因の定量的分析. 通訳翻訳研究, No. 9, pp. 21-32 (2009).	4章

参考文献

- [1] iPhone アプリ Jibbiggo の公式 HP. <http://www.jajahbabel.com/>.
- [2] JaJah Babel の公式 HP. <http://www.jajahbabel.com/>.
- [3] モバイルマルチメディアプレーヤ Votol の公式 HP. <http://121ware.com/votol/>.
- [4] 秋葉友良, 田中穂積. 拡張部分木を用いた漸進的構文解析. 情報処理学会第 45 回全国大会講演論文集, Vol. 3, pp. 175–176, 1992.
- [5] Jan W. Amtrup. Chart-based incremental transfer in machine translation. In *Proceedings of the 6th International Conference of Theoretical and Methodological Issues in Machine Translation*, pp. 188–195, 1995.
- [6] 荒牧英治, 黒橋禎夫, 柏岡秀紀, 加藤直人. 用例ベース翻訳の確率的モデル化. 自然言語処理, Vol. 13, No. 3, pp. 3–19, 2006.
- [7] Peter F. Brown, Vincent J. Della Pietra, Stephen A. Della Pietra, and Robert. L. Mercer. The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, Vol. 19, pp. 263–311, 1993.
- [8] Francisco Casacuberta, Enrique Vidal, and Juan Miguel Vilar. Architectures for speech-to-speech. In *Proceedings of the Workshop on Speech-to-Speech Translation: Algorithms and System*, pp. 39–44, 2002.
- [9] Kenneth Ward Church. A stochastic parts program and noun phrase parser for unrestricted text. In *Proceedings of the 1st Conference on Applied Natural Language Processing*, pp. 136–143, 1988.
- [10] Michael John Collins. A new statistical parser based on bigram lexical dependencies. In *Proceedings of the 34th Annual Meeting of the Association for Computational Linguistics*, pp. 184–191, 1996.

- [11] George Doddington. Automatic evaluation of machine translation quality using n-gram co-occurrence statistics. In *Proceedings of the 2nd International Conference on Human Language Technology Research*, pp. 257–258, 2002.
- [12] Robert E. Frederking, Alan W. Black, Ralf D. Brown, John Moody, and Eric Steinbrecher. Field testing the tongues speech-to-speech machine translation system. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, pp. 160–164, 2002.
- [13] 藤井哲也. 北京オリンピック期間における音声翻訳モニタ実証実験：携帯電話音声翻訳サービスの有用性能を検証. *電子情報通信学会誌*, Vol. 91, No. 12, pp. 1079–1080, 2008.
- [14] 古瀬蔵, 山本和英, 山田節夫. 構成素境界解析を用いた多言語話し言葉翻訳. *自然言語処理*, Vol. 6, No. 5, pp. 63–91, 1999.
- [15] 古瀬蔵, 山田節夫, 山本和英. 頑健な多言語音声翻訳のための不適格入力の分割処理. *情報処理学会論文誌*, Vol. 42, No. 5, pp. 1223–1231, 2001.
- [16] Nicholas J. Haddock. Incremental interpretation and combinatorial categorial grammar. In *Proceedings of the 10th International Joint Conference on Artificial Intelligence*, pp. 661–663, 1989.
- [17] 萩原直之. 英日同時通訳作業に必須とされる time lag の翻訳の合理化による短縮とその裨益. *英米評論*, 桃山学院大学紀要, No. 5, pp. 25–74, 1991.
- [18] Harald Hoge. Project proposal TC-STAR: Make speech to speech translation real. In *Proceedings of the 3rd International Conference on Language Resources and Evaluation*, pp. 136–141, 2002.
- [19] Chiori Hori, Sakriani Sakti, Michael Paul, Noriyuki Kimura, Yutaka Ashikari, Ryosuke Isotani, Eiichiro Sumita, and Satoshi Nakamura. Network-based speech-to-speech translation. In *Proceedings of the 6rd International Workshop on Spoken Language Translation*,, p. 168, 2009.
- [20] Ryosuke Isotani, Kiyoshi Yamada, Shinichi Ando, Ken Hanazawa, Shinya Ishikawa, and Kenichi Iso. Speech-to-speech translation software on PDAs for

- travel conversation. *NEC Research and Development*, Vol. 42, No. 2, pp. 197–202, 2003.
- [21] Hideki Kashioka and Takehiko Maruyama. Segmentation of semantic unit in Japanese monologue. In *Proceedings of International Conference on Speech Database and Assessments*, pp. 87–92, 2004.
- [22] 加藤芳秀, 松原茂樹, 稲垣康善. 確率木接合文法に基づく漸進的構文解析. 情報処理学会研究報告, NL-166, pp. 15–22, 2005.
- [23] 加藤芳秀, 松原茂樹, 外山勝彦, 稲垣康善. 主辞情報付き文脈自由文法に基づく漸進的な依存構造解析. 電子情報通信学会論文誌, Vol. J86-D-II, No. 1, pp. 86–97, 2003.
- [24] 加藤芳秀, 松原茂樹, 外山勝彦, 稲垣康善. 確率文脈自由文法に基づく漸進的構文解析. 電気学会論文誌, Vol. 122-C, No. 1, pp. 2109–2119, 2002.
- [25] 菊井玄一郎, 竹澤寿幸, 水島昌英, 葦苅豊, 中村哲, 佐々木裕, 河井恒, 清水徹, 山本誠一. 携帯型端末を用いた音声自動翻訳システムとフィールド実験. 電子情報通信学会技術研究報告, MoMuC105-264, pp. 11–16, 2005.
- [26] Genichiro Kikui, Eiichiro Sumita, Toshiyuki Takezawa, and Seiichi Yamamoto. Creating corpora for speech-to-speech translation. In *Proceedings of the 8th European Conference on Speech Communication and Technology*, pp. 281–284, 2003.
- [27] 木佐敬久. 放送通訳の聞きやすい速度とは? - ビデオ調査による視聴者の反応 - . 放送研究と通訳, Vol. 48, No. 3, pp. 40–57, 1998.
- [28] 小磯花絵, 間淵洋子, 西川賢哉, 斉藤美紀, 前川喜久雄. 『日本語話し言葉コーパス』の書き起こしの仕様について. 平成 15 年度国立国語研究所公開研究発表会論文集, pp. 27–28, 2003.
- [29] Muntzin Kolss, Matthias Wolfel¹, Florian Kraft¹, Jan Niehues¹, Matthias Paulik, and Alex Waibel. Simultaneous German-English lecture translation. In *Proceedings of the 5rd International Workshop on Spoken Language Translation*, pp. 175–181, 2008.
- [30] 窪園晴夫, 本間猛. 音節とモーラ, 英語学モノグラムシリーズ 15. 研究社, 2002.

- [31] Alon Lavie, Chad Langley, Alex Waibel, Fabio Pianesi, Gianni Lazzari, Paolo Coletti, Loredana Taddei, and Franco Balducci. Architecture and design considerations in NESPOLE!: a speech translation system for ecommerce applications. In *Proceedings of the 1st International Conference on Human Language Technology Research*, pp. 31–34, 2001.
- [32] Gianni Lazzari. TC-STAR: a speech to speech translation project. In *Proceedings of the 3rd International Workshop on Spoken Language Translation*, pp. 14–15, 2006.
- [33] Youngjik Lee, Seung-Shin Oh, and Jun Park. Usability considerations of speech-to-speech translation system. In *Proceedings of the International Conference on Spoken Language Processing*, pp. 369–372, 2004.
- [34] Fuhua Liu, Yuqing Gao, Liang Gu, and Michael Picheny. Noise robustness in speech to speech translation. *IBM Tech Report RC22874*, 2003.
- [35] 前川喜久雄, 籠宮隆之, 小磯花絵, 小椋秀樹, 菊池英明. 日本語話し言葉コーパスの設計. *音声研究*, Vol. 4, No. 2, pp. 51–61, 2000.
- [36] Mitchell P. Marcus, Beatrice Santorini, and Mary Ann Marcinkiewicz. Building a large annotated corpus of English: the Penn Treebank. *Computational Linguistics*, Vol. 19, No. 2, pp. 310–330, 1993.
- [37] 丸山岳彦, 柏岡秀紀, 熊野正, 田中英輝. 日本語節境界検出プログラムCBAPの開発と評価. *自然言語処理*, Vol. 11, No. 3, pp. 39–66, 2004.
- [38] Shigeki Matsubara, Takahisa Murase, Nobuo Kawaguchi, and Yasuyoshi Inagaki. Stochastic dependency parsing of spontaneous Japanese spoken language. In *Proceedings of the 19th International Conference on Computational Linguistics*, Vol. 1, pp. 640–645, 2002.
- [39] 松原茂樹. 同時通訳の工学と科学 –次世代自動通訳技術の実現に向けて–. *情報処理*, Vol. 49, No. 6, pp. 617–623, 2008.
- [40] 松原茂樹, 浅井悟, 外山勝彦, 稲垣康善. 不適格表現を活用する漸進的な英日話し言葉翻訳手法. *電気学会論文誌*, Vol. 118-C, No. 1, pp. 71–78, 1998.

- [41] Shigeki Matsubara, Satoru Asai, Katsuhiko Toyama, and Yasuyoshi Inagaki. Chart-based parsing and transfer in incremental spoken language translation. In *Proceedings of the 4th Natural Language Processing Pacific Rim Symposium*, pp. 521–524, 1997.
- [42] 松本裕治, 山内啓, 山下達雄, 平野善隆, 松田寛, 高岡一馬, 浅原正幸. 形態素解析システム『茶釜』version 2.3.3 使用説明書. 奈良先端科学技術大学院大学松本研究室, 2003.
- [43] 松山晶子. 英日逐次通訳とノートテーキング 訳出時間に着目した考察 . 通訳翻訳研究, No. 8, pp. 1–15, 2008.
- [44] David Milward. Incremental interpretation of categorial grammar. In *Proceedings of the 7th Conference of European Chapter of the Association for Computational Linguistics*, pp. 119–126, 1995.
- [45] Hideki Mima and Osamu Furuse. Simultaneous interpretation utilizing example-based incremental transfer. In *Proceedings of the 17th International Conference on Computational Linguistics and 36th Conference of the 36th Annual Meeting of the Association for Computational Linguistics*, pp. 855–861, 1998.
- [46] 村田匡輝, 大野誠寛, 松原茂樹, 稲垣康善. 講演の同時通訳のための対訳データの作成と分析. 第8回情報科学技術フォーラム講演論文集, E-018, pp. 293–294, 2009.
- [47] Satoshi Nakamura, Konstantin Markov, Hiromi Nakaiwa, Genichiro Kikui, Hisashi Kawai, Takatoshi Jitsuhiro, Jin-Song Zhang, Hirofumi Yamamoto, Ei-ichiro Sumita, and Seiichi Yamamoto. The ATR multilingual speech-to-speech translation. *IEEE Transactions on Audio, Speech and Language Processing*, Vol. 14, No. 2, pp. 365–376, 2006.
- [48] 中尾雪絵, Manny Rayner, Beth Ann Hockey, Pierrette Bouillon, Nikos Chatzichrisafis, 神崎享子, 井佐原均. 医療診断用音声通訳システム MedSLT における日本語規則の作成. 言語処理学会第11回年次大会発表論文集, pp. 546–549, 2005.

- [49] Hermann Ney, Franz Josef Och, and Stephan Vogel. The RWTH system for statistical translation of spoken dialogues. In *Proceedings of the 1st International Conference on Human Language Technology Research*, pp. 1–7, 2001.
- [50] 大原誠, 松原茂樹, 笠浩一朗, 河口信夫, 稲垣康善. 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—. *通訳研究*, No. 3, pp. 32–52, 2003.
- [51] 大野誠寛, 松原茂樹, 柏岡秀紀, 加藤直人, 稲垣康善. 節境界に基づく独話の漸進的係り受け解析. *電子情報通信学会論文誌*, Vol. J90-D, No. 2, pp. 556–566, 2007.
- [52] 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*, No. 7, pp. 51–64, 2007.
- [53] Kishore Papineni, Salim Roukos, Todd Ward, and Wei jing Zhu. BLEU: a method for automatic evaluation of machine translation. In *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics*, pp. 311–318, 2002.
- [54] Brian Roark. Probabilistic top-down parsing and language modeling. *Computational Linguistics*, Vol. 27, No. 2, pp. 249–276, 2001.
- [55] Brian Roark and Mark Johnson. Efficient probabilistic top-down and left-corner parsing. In *Proceedings of the 37th Annual Meeting of the Association for Computational Linguistics*, pp. 421–428, 1999.
- [56] Salim Roukos. Recent results on MT evaluation in the gale program. In *Proceedings of the 3rd International Workshop on Spoken Language Translation*, pp. 16–17, 2006.
- [57] 笠浩一朗, 松原茂樹, 稲垣康善. 日英同時翻訳のための日本語対話文の分割. *言語処理学会第13回年次大会発表論文集*, pp. 935–938, 2007.
- [58] 笠浩一朗, 松原茂樹, 稲垣康善. 対訳コーパスを用いた同時翻訳単位の検出. *言語処理学会第15回年次大会発表論文集*, pp. 881–884, 2009.

- [59] 西光雅弘, 高梨克也, 河原達也. 係り受けとポーズ・フィラーの情報をを用いた話し言葉の段階的チャンキング. 電子情報通信学会技術研究報告, NLC2005-104, pp. 247–252, 2005.
- [60] 清水徹, 中村哲, 河原達也. 音声翻訳単位の推定における句読点情報の効果. 電子情報通信学会技術研究報告, SP2008-100, pp. 127–131, 2008.
- [61] 清水徹, 中村哲, 河原達也. 同時通訳者の知識と韻律情報を用いた講演文章のチャンキング. 情報処理学会研究報告, SLP-68, pp. 81–86, 2008.
- [62] 清水徹. 同時通訳者の知識を用いた講演文章のチャンキング. 情報処理学会第70回全国大会講演論文集, Vol. 2, pp. 69–70, 2008.
- [63] 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回全国大会講演論文集, 第2巻, pp. 87–88, 2002.
- [64] Toshiyuki Takezawa, Tsuyoshi Morimoto, Yonishori Sagisaka, Nick Campbell, Hitoshi Iida, Fumiaki Sugaya, Akio Yokoo, and Seiichi Yamamoto. A Japanese-to-English speech translation system: ATR-MATRIX. In *Proceedings of the 5th International Conference on Spoken Language Processing*, pp. 957–960, 1998.
- [65] Toshiyuki Takezawa, Tsuyoshi Morimoto, and Yoshinori Sagisaka. Speech and language databases for speech translation research in ATR. In *Proceedings of the Oriental COCOSDA Workshop*, pp. 148–155, 1998.
- [66] 遠山仁美, 松原茂樹. 同時通訳における聴きやすさとポーズの関係～同時通訳コーパスを用いた被験者実験による分析. 通訳研究, No. 5, pp. 113–128, 2005.
- [67] 遠山仁美, 松原茂樹. 訳出遅延時間と訳出開始タイミングに着目した同時通訳者の原発話追従ストラテジーに関する分析. 通訳研究, No. 6, pp. 137–156, 2006.
- [68] 遠山仁美, 松原茂樹. 英日同時通訳者発話におけるフィラーの出現と聴きやすさとの関係. 通訳研究, No. 7, pp. 39–49, 2007.
- [69] Hitomi Tohyama, Shigeki Matsubara, Nobuo Kawaguchi, and Yasuyoshi Inagaki. Construction and utilization of bilingual speech corpus for simultaneous machine interpretation research. In *Proceedings of the 9th European Conference on Speech Communication and Technology*, pp. 1585–1588, 2005.

- [70] Joseph Turian, Luke Shen, and I. Dan Melamed. Evaluation of machine translation and its evaluation. In *Proceedings of the Machine Translation Summit IX*, pp. 386–393, 2003.
- [71] Wolfgang Wahlster. Robust translation of spontaneous speech: a multi-engine approach. In *Proceedings of the 17th International Joint Conference on Artificial Intelligence*, pp. 1484–1493, 2001.
- [72] 脇田由実, 河井淳, 飯田仁. 意味的類似性を用いた音声認識正解部分の特定法と正解部分のみ翻訳する音声翻訳手法. *自然言語処理*, Vol. 5, No. 4, pp. 111–125, 1998.
- [73] 横尾昭男, 竹沢寿幸, 山本誠一. C-STAR 音声翻訳国際共同実験. *情報処理*, Vol. 40, No. 9, pp. 908–909, 1999.