

同時通訳コーパスの設計と構築

松原 茂樹^{1,2} 相澤 靖之³ 河口 信夫^{1,4} 外山 勝彦^{1,3} 稲垣 康善³

名古屋大学統合音響情報研究拠点¹

名古屋大学言語文化部²

名古屋大学大学院工学研究科³

名古屋大学大型計算機センター⁴

This paper describes a large-scale spoken language corpus of simultaneous interpreting, which has been constructed at the Center for Integrated Acoustic Information Research (CIAIR), Nagoya University. The corpus, among other things, has the following characteristics: (1) English and Japanese speeches are recorded in parallel, (2) the data contain monologue and dialogue speeches, and (3) the exact beginning and ending times are provided for each utterance. We have collected a total of about 65 hours of speech data and transcribed them into ASCII text files (about 367,000 morphemes in 22,000 utterance units). This paper also outlines the software tools which we have developed for the investigation of the corpus. The corpus will be made publicly available in the near future.

1. はじめに

通訳研究では、人間の通訳プロセスを解明することを目標として、心理学や言語学、認知科学などのさまざまな分野を背景に研究が進められている。なかでも、実証論的アプローチに基づく研究では、収録した音声进行调查し分析することにより、通訳モデルの構築、およびその検証を実施する。実際、そのような研究がいくつか行われており（たとえば、亀井 1998; 永田 1997）、事例の詳細な観察を通して通訳プロセスに関する定性的な説明が提示されてきた。しかしながら、それに対して定量的分析を与えるには、これまでの研究で使用されたデータのサイズは必ずしも十分ではなかった。

MATSUBARA Shigeki, AIZAWA Yasuyuki, KAWAGUCHI Nobuo, TOYAMA Katsuhiko, and INAGAKI Yasuyoshi, "Design and Construction of Simultaneous Interpreting Corpus"

Interpretation Studies, No. 1, December 2001, pages 85-102.

(c) 2001 by the Japan Association for Interpretation Studies

最近では、計算機性能の向上や記憶媒体の大容量化を背景に、大規模な音声・言語データを処理することが可能となっており、今後、通訳研究の発展のために多くの研究者が共有可能な通訳データを大規模に収集し、蓄積することが望まれる。

一方、情報工学の分野では、近年、音声・言語処理技術の発展にともない、自動音声翻訳がますます重要な研究テーマとなっている。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 (Takezawa, 1998; Watanabe, 2000)、特定タスクドメインでの異言語間対話の実現可能性が次第に明らかになりつつある。ただし、対象となる対話は、発話者だけでなく通訳者の発話権も保障することが前提となっており、通訳の形態はいわゆる逐次通訳に留まっている。今後は、より自然なクロスリンガルコミュニケーション支援装置の実現に向けて、高い同時性を備えた翻訳技術、すなわち、同時通訳技術の開発が望まれる。

同時通訳システムを実現する上で、従来の機械翻訳研究では検討を要しなかった新たな研究課題に取り組む必要がある。というのも、機械翻訳システムでは、いわゆる *how-to-say* が重要視され (辻井 1989)、翻訳の精度や品質の向上を目指して研究が進められてきたが、同時通訳システムでは、それに加えて訳出タイミング、すなわち *when-to-say* も重要となるためである。遅れの程度が低く、話者発話にできる限り追従した訳文生成処理が望ましいが、一方で、翻訳の正確さと即時性との間にはトレードオフ関係が存在するため、同時通訳に適した訳文生成技術を備えることも重要である。

自動同時通訳の実現に向けた試みは始まったばかりであり (松原 1998; Matsubara, 1999)、まずは同時通訳の対訳コーパスの作成を通して同時通訳者の振舞いを分析することが効果的な方法であると考えられる。実際、音声言語の対訳コーパスは自動通訳技術の向上に重要な役割を果たしてきた (江原 1992; Morimoto, 1994)。これは、話者発話とその通訳発話が収録された大規模音声・言語データであり、通訳者のテクニックを獲得するための基礎データ、統計的手法に基づく翻訳処理の対訳用例、さらには、翻訳システムを評価するためのテストセットとしての役割を担うことが期待できる。

本論文では、著者らが進めている同時通訳コーパスの設計と構築について述べる。このコーパスは、名古屋大学統合音響情報研究拠点(CIAIR)(武田 2000)の言語論グループにおける研究プロジェクトの一環として作成している。本コーパスの特徴として、

- 英日・日英通訳の音声およびその文字化データを収録したパラレルコーパスである。文字化データには事前分析が与えられており、話し言葉に特有な言語現象に関するタグなどが付与されている。
- 講演を対象とした独話データ、ならびに旅行ドメインでの対話データから構成される。いずれにおいても、使用頻度が比較的高いトピックを採用しており、汎用性を備えた音声コーパスである。
- 同一の独話に対して、経験度が異なる複数の通訳者の通訳音声を収集している。音声通訳システムの現状レベルでは、経験度の高い通訳者音声のほうが有用で

あるとは必ずしも言えず、そのような観点に基づく比較分析に利用できる。

- 話者、および通訳者の発話をポーズで分割し、各発話に、開始・終了時刻を付与している。発話を対応付けることにより、通訳発話の遅れなどの時間情報を獲得することができる。

などがあり、自動通訳技術の向上はもちろん、通訳会話の分析や通訳プロセスのモデル化のための基礎データとしても活用できる。2000年までに全体で約65時間の音声を収録し、その文字化および言語分析を完了した。文字化データのサイズは形態素数にして36万7千程度であり、音声対訳コーパスとしては大規模なクラスに属する。

また本研究では、同時通訳コーパスの活用を支援することを目的に、データ分析用ソフトウェアツールを開発した。これらは、ファイルアクセス、音声データの閲覧、言語データの整備作業などにおける支援ツールとして利用できる。いずれもWWWサーバ上で実行可能なソフトウェアとして実現しており、ユーザはブラウザを使用することにより、インターネット上でのデータ参照を容易に行うことができる。

本論文の構成は以下のとおりである。次章では、同時通訳コーパスの設計について検討する。3章で音声の収録について、また、4章でそのデータベース化について述べる。5章ではコーパス分析用ツールについて説明する。

2. 同時通訳コーパスの設計

音声言語対訳コーパスを設計するにあたり、収録する音声データのタイプをどのように選択するのかはきわめて重要となる。同時通訳を対象とした、科学的小および工学的研究で利用可能な言語資料を作成することがわれわれの収録目的であり、収録対象はもちろん同時通訳音声となる。しかしながら、大規模コーパスの収集は容易な作業ではないことを勘案すると、将来における利用可能性も考慮し、汎用性を備えた設計とする必要がある。また、収録できるデータの量には限度があるため、適切な状況下での同時通訳音声を収録することが必要となる。

ところで、これまでに作成された音声言語対訳コーパスは、いずれも対話を収録の中心と位置付けている(江原1992; Morimoto, 1994)。これは、自動翻訳電話などの音声対話翻訳システムを開発する上での基礎資料として活用することが収録の目的であったことに起因している。しかし、対話のような双方向コミュニケーションだけでなく、講演や講義などの独話音声、すなわち片方向のものもまた通訳研究の対象であり、とくに同時通訳においては重要となる。実際、社会における現在の通訳需要からみても、将来の自動独話音声通訳に対する期待は大きいといえよう。

そこで本研究では、多様なデータを収集するために、独話(monologue)および対話(dialogue)の同時通訳者音声をいくつかのトピックを設定して収録することとした。まずは、英語と日本語を対象言語として選定し、その双方向音声を収録した。本研究で構築した同時通訳コーパスの収録内容を表1に示す。

表1 同時通訳コーパスの内容

発声様式	独話、対話
通訳様式	同時通訳
使用言語	英語、日本語
データ様式	音声、テキスト

独話音声としては、英語話者の講演発声と通訳者による日本語発声、ならびに日本語話者による講演発声と通訳者による英語発声、をそれぞれ収録している。さらに、社会的に使用頻度の高いテーマおよび内容を採用した。具体的には、政治、経済、歴史、文学、言語、技術、都市、環境などをテーマとして選定し、それに関する講演を音声形式で収録した。なお、同時通訳は豊富な訓練を必要とする高等技術であり、通訳者の熟練度によりその結果に大きな違いが生じる。また、音声通訳システムの現状レベルでは、経験度の高い通訳者音声のほうが有用であるとは必ずしも言えない。熟練度の影響を調査するために、同一の講演に対して経験年数が異なる複数の同時通訳者を用意し、その通訳結果を収録している。

一方、対話ドメインとして、これまでに構築されているいくつかの対話データベース (Morimoto, 1994) と同様、「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以下に示すトピックを定め、模擬会話を収録した。

- 航空会社、ホテル、レストラン、レンタカーの受付係と客との予約、およびその変更や取消に関する電話会話
- 空港の入国審査や税関での係員と旅行者との対面会話
- 空港のサービスカウンターなどの受付係と旅行者との、アクセス情報やホテル紹介に関する対面会話
- ホテルのフロント係などと客とのチェックインやツアー相談に関する対面会話

3. データの収集

3.1 収録の環境

名古屋大学統合音響情報研究拠点では、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ひとつとしており (河口 2001)、本データの収録で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。すべて同一のスタンドマイクを使用し、話者とその通訳者の音声は、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) に複数チャンネル環境で収録している。収録時間の総数を表 2 に示す。

表 2 音声の収録時間

項目		話者 (分)	通訳者 (分)
独話	英日	456	684
	日英	451	680
	合計	907	1,364
対話	英日	404	404
	日英	404	404
	合計	808	808
合計		1,715	2,172

3.2 独話データの収集

独話の収録では、英語あるいは日本語の講演に対して、1 人もしくは 2 人の同時通訳者が通訳を行った。講演者は聴衆に向かって発声をするが、講演者には通訳者の音声伝わらないようにし、聴衆の様子をうかがいながら自らのペースで発声できるようにしている。独話における話者音声の収録状況を図 1 に示す。一方、同時通訳者は通訳用ブースに入り、ヘッドホンから流れる話者音声に対して、講演者の振舞いを見ながら通訳を進める。独話における通訳者音声の収録状況を図 2 に示す。講演の聴衆は、ヘッドホンを通して通訳音声を聴くことができる。



図 1 収録の状況（話者）



図 2 収録の状況（通訳者）

なお、われわれは著作権の観点から模擬講演を収集することとしたが、ある程度の品質を備えた講演を行うことを被験者に要求することは難しい。そこで発声内容をあらかじめこちらで制御できるといった利点も考慮し、あらかじめ講演原稿をわれわれが用意して、話者はそれを参照しながら発声するという形態を採用した。話者には、事前に内容を十分把握し、講演者としてできる限り自然に発話することを依頼したが、

場合によっては読み上げに近い発声となった場合もあった。また、通訳者が講演を実際に同時通訳する場合には、原稿やレジュメなどをもとに、あらかじめ講演内容を把握しておくことが一般的である。そこで、実際に近い状況を設定するために、さらにはある程度の通訳の品質を確保するために、収録の直前に通訳者に講演原稿を渡し、内容を把握してもらうようにした。

1 講演あたりおよそ 10～30 分程度の音声を収録している。実際の同時通訳の現場での 1 人あたりの連続通訳時間に比べるとやや長い場合もあった。収録作業手順の関係上、今回はそのような収録時間を設定したが、これが通訳の品質の低下を招いた可能性もある。これについては今後の収録で改善を図る予定である。

3.2 対話データの収集

対話の収録では、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。遂行される対話のドメインによって電話による会話や対面による会話が想定されるため、話者間の仕切りの設置または撤去により、非対面・対面それぞれの会話環境を作り上げた。会話における話者の発話権を確保するために、話者は相手の話者の発声を通訳した結果のみを聞くことができることにした。通訳者は、対話の全体的な把握をするために、担当する話者の音声だけでなく、もう一方の話者音声も聞き取ることができる環境を作成した。これにより、通訳データの品質を高めることが可能となる。なお、対話様式は模擬的であるものの、できる限り自由な発話を収集するため、話者役割と対話タスクのみを事前に設定している。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を取りたい航空券の種類と枚数を、また、チケット受付を担当する話者には発券可能なフライトの状況を提示する。なお、1 対話あたりの収録時間は 1 分から 16 分である。

4. データベースの構築

4.1 音声データの文字化

収集した音声データの文字化作業は人手で行った。文字化データの作成にあたり、以下の項目について決定する必要がある。

- **表記方法:** 音声の文字化では、作業者の言語能力によってその表記の仕方が揺れることが多い。同一の単語であっても、その表記が複数存在すれば、言語処理に利用する場合にもその扱いが難しくなる。できる限り統一化できるような基準を準備することが望ましい。
- **発話の単位:** 話し言葉では、書き言葉と異なり、文を定義することは容易ではない。作業者にとって比較的揺れの少ない客観的な基準を設定することが望ましい。

- **談話タグの種類:** データに付与するタグの選択が重要である。一般的には、タグの情報が豊富であるほど多くの利用目的に合致したデータの提供が可能となる。しかしその一方で、タグを多く保持するほど収録時間あたりのコーパス作成コストが増大するため、結果的に収録できるデータ量が少なくなる。

以上の項目の中でも、談話タグの種類については、同時通訳研究のための言語資料として本コーパスを活用することを目的としたものであることを考慮する必要がある。本研究では、データの品質と数量の双方を考慮し、2種類のデータを作成することを試みた。すなわち、対話データでは品質の高いデータを目指し、高度なタグ付けを行うこととし、一方、独話データについては量を重視し、比較的シンプルなタグのみを付与した。現在までに文字化が完了している独話データの量を表3に、対話データの量を表4に示す。

表3 現在までに文字化された独話データの量

項目		形態素数	発話数
話者	英語	54,875	2,849
	日本語	62,570	2,024
	合計	117,445	4,873
通訳者	英日	100,923	5,011
	日英	64,119	3,615
	合計	165,042	8,626
合計		282,487	13,499

表4 現在までに文字化された対話データの量

項目		形態素数	発話数
話者	英語	22,528	2,231
	日本語	21,026	1,847
	合計	43,554	4,078
通訳者	英日	25,066	2,609
	日英	16,355	1,953
	合計	41,421	4,562
合計		84,975	8,640

以下では、対話音声データの文字化について説明する。同時通訳研究で文字化データを利用する上で、以下に述べる点が重要である。

- 話者と通訳者の双方の発話の特徴を明らかにするために、話し言葉に特有の言語学的現象、たとえばフィラー（「えーと」や「あの一」など）や言い淀みなどのタグを付与する必要がある。
- 話者発話とそれに対応する通訳者発話との発声タイミングに関する関係を明らかにするために、発声時刻情報が有用である。
- 前の項目に加え、より細かな言語単位に対する発声タイミングを推測するために、音声の一般的な文字表記だけでなく、その読みに関する情報（発音形データ）も加えることが望ましい。

これらの項目について検討した結果、本コーパスでは、国立国語研究所が中心となって構築を進め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することとした。すなわち、話者および通訳者の発話音声を 200 msec 以上のポーズ（無音声区間）で分割することにより発話を定義し、その開始および終了に関する時間タグを付与する。また、フィラーや言い淀み、言い直しなどの言語タグを与えるとともに、日本語音声に限って発音形データを併記した。ただし、文節切り作業は現在のところ行っていない。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 3 (p. 98) に、英日通訳者発話のサンプルを図 4 (p. 98) に、それぞれ示す。

4.2 収録環境情報の作成

一般に、コーパスは利用者によってその使用目的は異なるが、目的によっては音声データや文字データには現れない各種情報が重要となる場合がある。このため、収録セッションごとに収録に関する情報を含んだテキストファイルを作成した。収録日時・場所・時間・機器や発話トピック、話者役割および話者・通訳者情報などの項目を掲載している。利用者は、収録環境ファイルの内容を参照することにより、目的のデータを取り出すことができる。

4.3 データベースの構成

本データベースは、音声データファイル、文字化データファイル、および環境データファイルから構成される。これらのファイルを UNIX ファイルシステムに従って階層化することにより整備した。基本構成を図 5 (p. 99) に示す。

前章で説明したとおり、話者の音声とその通訳者の音声を同時に録音し、多重音声ファイルとして収納している。すなわち、独話音声の収録では、1 つのセッション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、また、対話音声の収録では 2 つの多重音声ファイルを作成した。ここで、対話音声の文字化では、同一の言語を発声する 2 人の音声（英語話者と日英通訳者、あるいは日本語話者と英日通訳者）を同一のテキストファイルとして作成することも考えられる。しかし、本データベースでは、文字化ファイルも音声ファイルに従い、別々のファイルとして実現し、データの

種類、性質ごとにディレクトリで区分して収納することとした。このようにすることにより、片方向の通訳、すなわち話者音声とそれに対応する通訳者音声の双方を分析対象とする場合にも利用しやすくなる。ただし、本データの時間情報を利用すれば、2人の発話データをひとつの対話データとして合成することは機械的に行える。

5. コーパス分析用ソフトウェアの開発

本研究で構築した同時通訳コーパスは、通訳理論研究のための言語資料として活用可能であるものの、効率のよい談話分析を進めるために、データ利用環境を整備することが望ましい。そのような環境を構築する上で、複数の研究者間で互いに容易にアクセスできること（データの共有化）、およびデータを閲覧しやすいこと（データの視覚化）が重要となる。そこで本研究では、インターネット上で閲覧可能なコーパス利用支援ツールを作成した。コーパスを WWW サーバ上に載せ、CGI (Common Gateway Interface) 技術を使ってユーザの目的に合致した形式でデータをブラウザ上に表示する。なお、本ツールの CGI スクリプトの記述には、テキスト処理言語 Perl を用いた。

5.1 コーパス分析支援ツール

同時通訳データを使用した談話分析のために、データベース参照用ソフトウェアを作成した。これは、WWW サーバに置かれたデータベースをインターネット上で参照するためのソフトウェアである。ユーザはブラウザ上で GUI (Graphical User Interface) 操作のみによって、参照したいデータの閲覧が可能となる。図 6 (p. 99) に選択用ウィンドウを示す。閲覧したいファイル名をマウスで指示することにより、そのファイル内容が図 7 (p. 100) に示すようにテキスト形式で表示される。

本コーパスは、前節で説明したようにいくつかの事前分析が行われており、談話タグや発話時刻データなどが付与されている。しかし、これらのデータは分析の内容によっては不要な情報となる場合もあり、ユーザの使用目的に応じて表示形式を変更可能となることが望ましい。そこでわれわれは、データ表示形式の選択用ツールを作成した。形式の選択もテキストファイルの内容を表示したブラウザ上でマウスのみで操作可能である。データの表示方法として以下のものがある。

- **原データ**: タグ付きのテキストデータである。言語分析用タグや時間分析用タグ、および読み分析用タグ（発音形データ）が付与されている。
- **タグ除去データ**: 原データに付与されているタグを取り除いたデータを表示する。取り除きたいタグの種類（言語分析・時間分析・読み分析）を選択することができる。時間タグのみ付与されたデータの表示例を図 8 (p. 100) に示す。
- **日本語形態素解析データ**: 日本語形態素解析システム「茶筌」(松本 1999) による日本語データの形態素分析結果を表示する。データの表示例を図 9 (p. 100) に示す。ただし、現在のところ英語データの解析については未対応である。

- **タグを数える:** 言語分析用タグの種類ごとにその発生頻度を計算し、表示する。現在、本ツールを使用することにより、同時通訳コーパスに収録したデータの定性的分析を進めている。

5.2 発話文対応付け作業支援ツール

対訳データの分析を詳細に行うときには、対応する話者発話と通訳者発話を比較参照することになる。とくに、同時通訳の自動化に利用する場合には発話文対応が定められていることがその前提となる。なぜなら、文より大きな言語単位間での自動翻訳は技術的にまだ容易ではなく、今のところ文を単位とした翻訳技術の向上が検討される必要があるためである。

このような対訳アライメント作業を自動化するための研究が行われているが、そのほとんどが新聞や技術文書などの書き言葉データを対象としている。ところが、通訳発話には意識や誤訳も頻繁に見られるとともに、翻訳が省略されることもしばしばあることから、精度の高いアライメントは困難であり、その作業は人手に頼らざるを得ない。われわれは、人手による対訳アライメント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。ツールの表示画面を図 10 (p. 101) に示す。本ツールは Perl による CGI プログラムと JavaScript を併用して作成している。画面では、話者発話文と通訳者発話文が表形式で平行で表示されており、表の各行がひとつの発話対応を示す。ユーザは、表の左右両端のボタンをマウスクリックするだけでアライメント作業を行える。すなわち、ある通訳者発話が、その左隣ではなく左上の話者発話に対応する場合には、ユーザはその通訳者発話の右側の欄の “ ” をクリックすることにより、対応付けができる。人手で対応付けされたデータについてはブラウザ上で閲覧することができ、前節で説明したツールと同様、通訳データの分析に使用できる。

5.3 音声データの視覚化

同時通訳コーパスの対話データには、話者および通訳者発話の開始時刻および終了時刻が付与されている。このような時間タグは、通訳者の発声タイミングの調査および分析に有用であり、同時通訳システムの研究においても適切な通訳単位を形式化する上で利用できる。しかし、定性的な分析を与えるためには、発話間の発声タイミングを直観的に理解可能な形式で提供することが望ましい。とくに、この対話コーパスでは、ひとつの対話セッションについて、英語話者、日本語話者、英日通訳者、日英通訳者の 4 つの文字化テキストが存在するため、それら各発話間の時間的關係を把握することは難しい。

そこで、データの文字化作業で付与した発声時間情報をもとに、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作成した。対話に対する音声データの視覚

化の例を図 11 (p. 102) に示す。図中、右側の発話が左のグラフ内に記された時間に行われたことを表している。グラフの左から順に、英語話者、英日通訳者、日本語話者、日英通訳者の発声時間であり、話者と通訳者の発声の重なり具合がわかる。この視覚化データを用いて、注目すべき対訳データ（発話対）を特定し、前節で説明した分析支援ツールを使用してそのデータを参照することができる。

6. まとめ

本論文では、名古屋大学で構築を進めている CIAIR 同時通訳データベースの設計と実現について述べた。コーパスは、同時通訳者を介した独話および対話データから構成されており、同時通訳研究のための言語資料として、また、同時通訳処理に有用な知見を獲得するための基礎データとして活用できる。現在、作成したコーパスの調査・分析を行っており、その結果については稿を改めて報告したい。

今後も引続きデータの収集を進めるとともに、言語処理に有用な各種タグ付けや対訳アライメント作業を行うことにより、より品質の高いデータベースとして実現することを検討している。なお、本論文で述べたコーパスデータについては、さらに整備を進めた上で、できる限り早い段階で公開する予定である。通訳技術の向上のために、通訳科学の観点からの成果が不可欠である。通訳理論研究者の方々にも大いに利用していただき、自動音声翻訳の分野はもちろん、通訳研究の発展のために本データベースがその一端を担えることができれば幸いである。今後の収録に対して要望等があれば、忌憚なき意見を著者らまでお寄せ願いたい。

謝辞：日頃ご支援下さる名古屋大学統合音響情報研究拠点 (CIAIR) のメンバーの皆様に深く感謝いたします。データの作成にあたり有益なご助言をいただきました ATR 音声言語通信研究所の隅田英一郎、竹沢寿幸の両氏に感謝いたします。本研究で作成したコーパスは、日本語話し言葉コーパスの書き起こし基準に準拠いたしました。基準策定にあられた皆様に感謝いたします。(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。ご尽力いただきました横尾祐郁氏に感謝します。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) によります。

著者紹介：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、名古屋大学言語文化部助手。自然言語処理、音声言語処理、機械翻訳の研究に従事。連絡先：〒464-8601 名古屋市千種区不老町 (Email: matubara@lang.nagoya-u.ac.jp)

相澤靖之 (Aizawa, Yasuyuki) 名古屋大学大学院工学研究科博士前期課程修了。現在、日本電気株式会社。在学中は音声翻訳システムの研究に従事。

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、名古屋大学大型計算機センター助教授。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。

外山勝彦 (Toyama, Katsuhiko) 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、名古屋大学大学院工学研究科助教授。同大学統合音響情報研究拠点言語論グループリーダー。人工知能基礎論、自然言語理解の研究に従事。

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。現在、名古屋大学大学院工学研究科教授。同大学大型計算機センター長。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。

編集部注：本論文の内容に関連して、2001年7月21日に行なわれた日本通訳学会第4回例会において「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」と題する松原茂樹氏の講演が行なわれました。松原氏のご好意により、同講演の要旨(pdfファイル)およびオンライン・スライドショー(PowerPoint)が学会のホームページに掲載されています。URLは次のとおりです(2001年12月1日現在) = <http://ux01.so-net.ne.jp/%7Ea-mizuno/meeting1.html>

【参考文献】

江原暉将・小倉健太郎・篠崎直子・森元逞・樽松明 (1992) 「電話またはキーボードを介した対話に基づく対話データベース ADD の構築」『情報処理学会論文誌』第33巻第4号: 448-456

亀井千秋 (1998) 「日英同時通訳実証研究」『通訳理論研究』第14号: 96-103 通訳理論研究会

河口信夫・牛窪誠一・松原茂樹・岩博之・梶田将司・武田一哉・板倉文忠 (2001) 「走行車室内音声対話収録システムの開発」『電子情報通信学会論文誌』J84-D-II, 第6号: 909-917

前川喜久男・籠宮隆之・小磯花絵・小椋秀樹・菊池英明 (2000) 「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・浅井悟・外山勝彦・稲垣康善 (1998) 「不適格表現を活用した漸進的な英日話し言葉翻訳手法」『電気学会論文誌』118-C, 第1号: 71-78

松本裕治・北内啓・山下達雄・平野義隆 (1999) 「日本語形態素解析システム『茶筌』Version 2.0 使用説明書」『Information Science Technical Report』NAIST-IS TR9908 奈良先端科学技術大学院大学

- 永田小絵 (1997) 「同時通訳・訳出の比較—香港返還記念式典のスピーチから」『通訳理論研究』第 13 号: 4-23 通訳理論研究会
- 武田一哉・板倉文忠 (2000) 「文部省 COE プログラム統合音響情報研究拠点(CIAIR)—音声・音響情報処理の多角的研究」『日本音響学会誌』第 56 巻第 11 号: 748-751
- 辻井潤一 (1989) 「機械翻訳における文章の生成」『人工知能学会誌』第 4 巻第 6 号: 645-651
- Matsubara, S., Toyama, K. and Inagaki, Y. (1999). Sync/Trans: Simultaneous Machine Interpretation between English and Japanese. In N. Foo (Ed.), *Advanced Topics in Artificial Intelligence, Lecture Note in Artificial Intelligence*, Vol. 1747: 134-143.
- Morimoto, T., Uratani, N., Takezawa, T., Furuse, O., Sobashima, Y., Iida, H., Nakamura, A., Sagisaka, Y., Higuchi, N., and Yamazaki, Y. (1994). *A Speech and Language Database for Speech Translation Research*. Proceedings of the 3th International Conference of Spoken Language Processing, 1791-1794.
- Takezawa, T., Morimoto, T., Sagisaka, Y., Campbell, N., Iida, H., Sugaya, F., Yokoo, A. and Yamamoto, S. (1998). *Japanese-to-English Speech Translation System: ATR-MATRIX*. Proceedings of the 5th International Conference of Spoken Language Processing, Vol. III, pp. 957-960.
- Watanabe, T., Okumura, A., Sakai, S., Yamabana, K., Doi, S. and Takahashi, K. (2000). *An Automatic Interpretation Software for Travel Conversation*. Proceedings of Workshop on Multi-Lingual Speech Communication, 21-24.

図3 対話文字化データのサンプル (英語話者)

```
0001 - 00:01:973-00:04:837 N:  
Okay (X welcome to) welcome to the west coast <SB>  
How may I help you? <SB>  
0002 - 00:08:616-00:09:254 N:  
Okay <SB>  
0003 - 00:09:850-00:11:468 N:  
And let me see your passport please  
0004 - 00:12:776-00:13:566 N:  
and your entry card <SB>  
0005 - 00:23:573-00:23:844 N: <breath>  
0006 - 00:23:933-00:24:690 N:  
Okay  
0007 - 00:25:103-00:25:799 N:  
and (F uh)  
0008 - 00:26:769-00:28:889 N:  
may I ask you about your travel purpose? <SB>
```

図4 対話文字化データのサンプル (英日通訳者発話)

```
0001 - 00:04:993-00:07:457 I:  
(F あ)ウエストコーストによくいらっしゃいました<SB>  
じゃあいかがいたしましょう<SB>  
(F ア)ウエストコーストニヨクイラッシュイマシタ<SB>  
ジャーイカガイタシマショー<SB>  
0002 - 00:07:608-00:08:422 I:  
何かご用でしょうか<SB>  
ナニカゴヨーデショーカ<SB>  
0003 - 00:11:380-00:13:544 I:  
(F え)パスポートを(D ごら)(F あー)見せて頂きたいと思ひます<SB>  
(F エ)パスポートオ(D ゴラ)(F アー)ミセテイタダキタイトオモイマス<SB>  
0004 - 00:13:781-00:15:560 I:  
(F え)それから入国カードをお願いいたします<SB>  
(F エ)ソレカラニューコクカードオオネガイイタシマス<SB>  
0005 - 00:25:020-00:25:750 I:  
分かりました<SB>  
ワカリマシタ<SB>  
0006 - 00:29:013-00:29:417 I: <息>  
0007 - 00:29:707-00:31:673 I:  
旅行の目的<H>お願いできますか<SB>
```

図 5 同時通訳データベースの構成

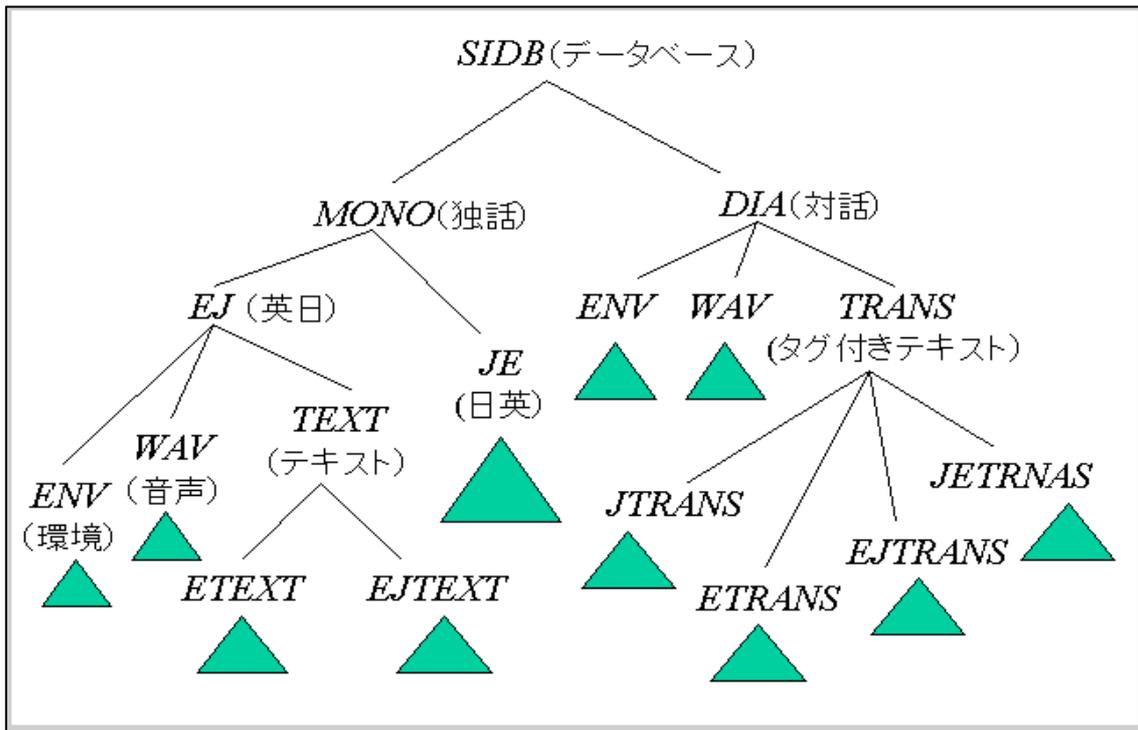


図 6 分析用データの選択

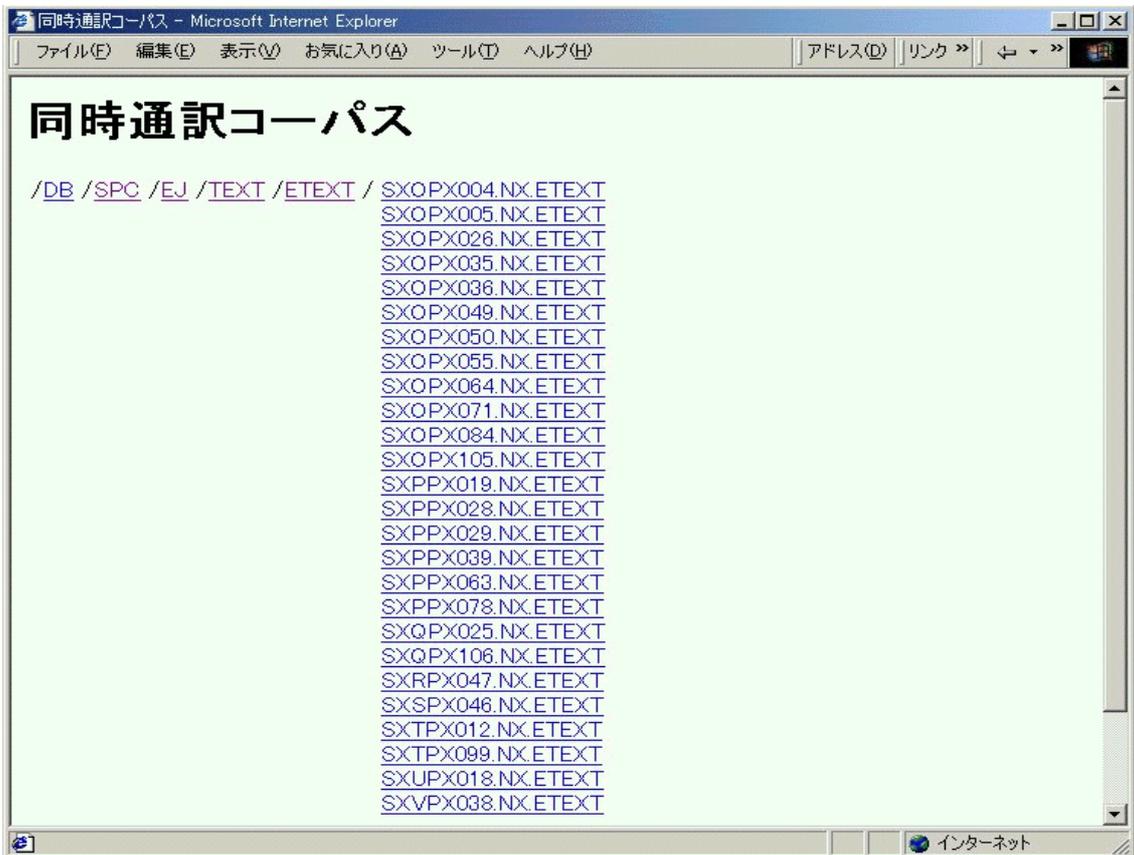


図7 ファイル内容のテキスト表示

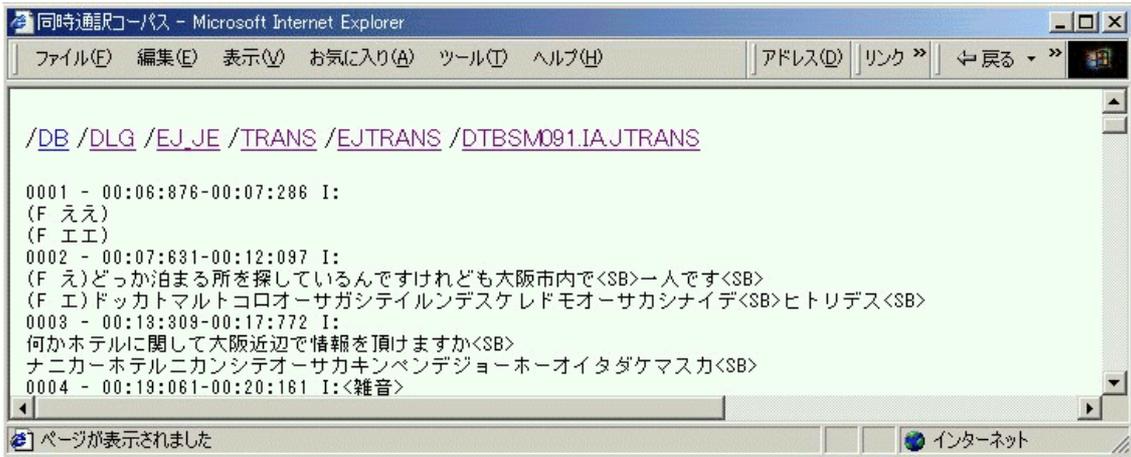


図8 タグ除去データの表示 (時間タグのみ表示した例)

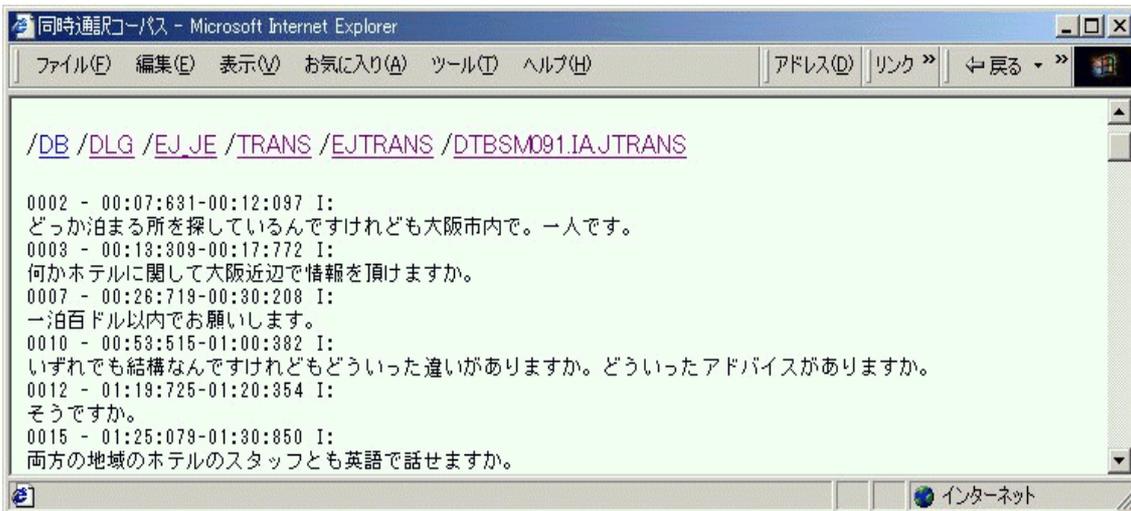


図9 形態素解析結果の表示

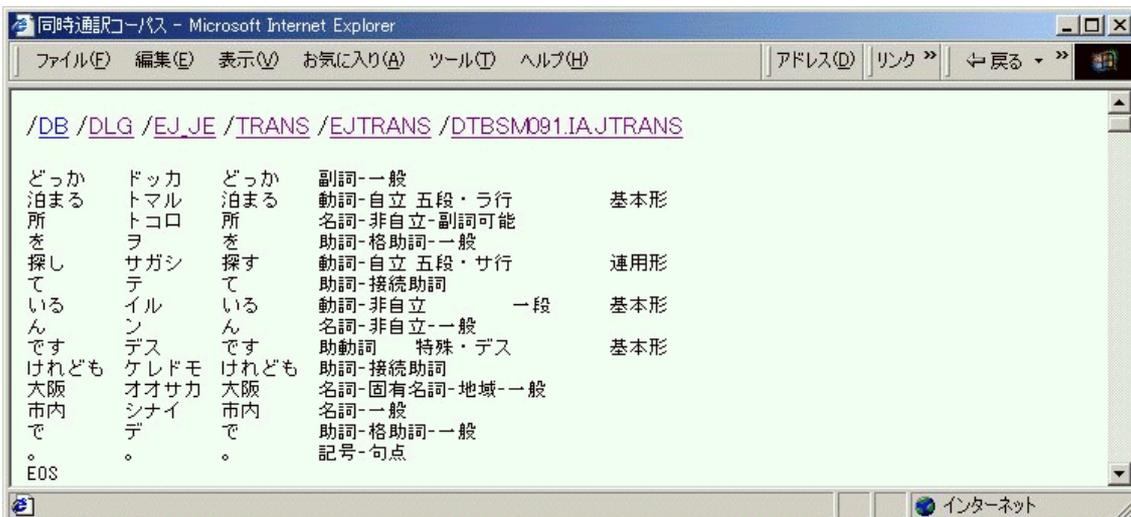


図 10 発話対応付け作業支援ツール

#	講演	通訳
0	(V_tongue) Good afternoon, ladies and gentlemen.	こんにちはみなさま。
1	Thank you for inviting me to this honorable session I am told that this session will center around a discussion.	この(F お)権威あるセッション(R にゅ;に)(F お)(R おむ <FV>)お迎えいただきましてありがとうございます。
2	(V_tongue) This discussion will focus on the role the United Nations will be expected to play within the international community in the upcoming century.	(F う<H>)今日<H>お<H>話したいのは(は<PS>)(A ユーエヌ; UN)の(F お<H>)機能のことです。
3	(V_breath) It will also deal with the role Japan will be expected to assume within the United Nations itself.	(F え<F え>)国際社会で(F え<H>)二十一世紀において(F え)期待される国連の役割です。
4	To provide the materials for this discussion, I will give a short and simple presentation which will contain many figures.	また日本(にほん)が(F は<H>)この(F え<H>)XQ (R きゅうよのなかで)国連の中でどのような機能を期待されてるかということです。
5	Since the cold war ended, the importance of the UN has substantially increased.	(F え<H>)またここで簡単に(F え<H>)XR お<H>はなしゆる; お話するために(F え<H>)数字を含めてお話ししたいと思います。
6	This is without doubt.	冷戦が終わりました(F え)国連の重要性(R ての;てのは)高まりました。
7	After the first World War, many nations felt it was necessary to create an international organization that would bring about international peace, order, and security.	(F あ<H>)これは<FV>疑う余地はございません。

図 11 音声データの視覚化

