

# 平成13年度 研究開発成果報告書

## 「大規模コーパスベース音声対話翻訳技術の研究開発」

### 目 次

- 1 研究開発課題の背景
- 2 研究開発分野の現状
- 3 研究開発の全体計画
  - 3-1 研究開発課題の概要
    - 3-1-1 サブテーマ
  - 3-2 研究開発目標
    - 3-2-1 最終目標
    - 3-2-2 中間目標
  - 3-3 研究開発の年度別計画
  - 3-4 研究開発体制
- 4 研究開発の概要（平成13年度）
  - 4-1 研究開発実施計画
    - 4-1-1 研究開発の計画内容
    - 4-1-2 研究開発課題実施計画
  - 4-2 研究開発の実施内容
- 5 研究開発実施状況（平成13年度）
  - 5-1 実音響環境での音声認識技術の研究開発
    - 5-1-1 序論
    - 5-1-2 委託業務の内容
    - 5-1-3 委託業務の効果
    - 5-1-4 他の研究機関における類似研究開発及び協力関係状況
    - 5-1-5 まとめ、今後の課題等
  - 5-2 音声言語統合技術の研究開発
    - 5-2-1 序論
    - 5-2-2 委託業務の内容
    - 5-2-3 委託業務の効果
    - 5-2-4 他の研究機関における類似研究開発及び協力関係状況
    - 5-2-5 まとめ、今後の課題等
  - 5-3 コーパスベース対話翻訳技術の研究開発
    - 5-3-1 序論
    - 5-3-2 委託業務の内容
    - 5-3-3 委託業務の効果
    - 5-3-4 他の研究機関における類似研究開発及び協力関係状況
    - 5-3-5 まとめ、今後の課題等
  - 5-4 コーパスベース音声合成技術の研究開発

- 5-4-1 序論
- 5-4-2 委託業務の内容
- 5-4-3 委託業務の効果
- 5-4-4 他の研究機関における類似研究開発及び協力関係状況
- 5-4-5 まとめ、今後の課題等

## 5-5 総括

### 参考資料、参考文献

(添付資料)

#### 1 研究発表、講演、文献等一覧

## 1 研究開発課題の背景

母語以外の言語の習得には長年にわたる学習過程を必要とすることから、学習を必要とせずに外国語のコミュニケーションを可能とする自動翻訳技術は人類の共通の夢となっている。自動翻訳技術は、まず書物を翻訳するための技術である機械翻訳技術として、1950年代に研究開始された。1970年代から世界の通商や人の移動が大きく増加したことを受け、1980年代に入って、音声言語、すなわち話し言葉によるコミュニケーションを可能とする音声翻訳技術の研究も開始された。1980年代からの日本での基礎研究の結果、音声認識技術、言語翻訳技術、音声合成技術などの、音声翻訳技術を構成する各要素技術は著しく進歩した。音声翻訳技術自体も一部限定された用途における日本語英語間の対話で、一文毎の翻訳がある程度実現できる段階に達している。

しかし、音声認識技術は、限られた環境下では利用可能であるが、様々な実環境下で種々の使用者の利用を可能にするという意味では、音声翻訳技術の要素技術として満足できる性能には達していない。言語翻訳技術については、個別の言語対毎の翻訳規則を人の内観に頼って開発しており、新たな言語対やドメイン（話題）での利用を可能にするためには、新たに多くの開発作業を要する。更に、言語翻訳技術が有する最大の問題点の一つは、音声認識誤り、発話の不完全性、翻訳時に使用する知識や用例の不足等の可能性を考慮して決定されるべき翻訳結果の信頼度を示す指標がない点である。このため、相手言語の知識のない使用者が信頼して実使用環境下で利用できる段階には到っていない。

21世紀に入って、国境を越える人や物さらには情報の交流はますます増加しており、グローバル化された多言語社会において、異なる言語を話す人の間で互いの母語によるコミュニケーションを可能とする多言語音声対話翻訳技術への期待は、一層高まっている。この要望に応え得る多言語音声翻訳技術を研究開発し、実使用環境での利用可能性を実証することが、本研究開発の目的である。

## 2 研究開発分野の現状

CPUやメモリ等の半導体技術の発展によるコンピュータの高性能化に伴い、特定の話題や表現に対する音声翻訳という制限付きではあるが、音声翻訳装置の開発・実用化がなされつつある。しかし、これらの装置では音声認識機能と言語翻訳機能が機能的には切り離され、音声認識結果が単なるテキスト形式で情報が言語翻訳部に授受される構成が採用されている。更に言語翻訳機能としては従来の文書の翻訳のために開発された規則ベース翻訳技術が採用されている。このため、研究開発課題の背景に記述された問題点が未解決のまま残されており、実際の環境下での使用に際してはさまざまな制限や困難が生じる。

一方、機械翻訳技術の研究の中心は、情報化社会の進展に伴うコーパスの急激な増加を受けて、機械学習に基づいてシステムを構築するコーパスベース翻訳技術に移りつつある。コーパスベース翻訳技術では、整備された大規模な対訳コーパスを準備できれば、規則ベ

ース翻訳方式の最大の問題点の一つである移植性の困難さを解消できるメリットを有している。しかし、音声翻訳技術の研究については、テキストの翻訳の場合と異なり、音声言語の対訳コーパスとしてコーパスベース翻訳に利用可能なほどに十分な量を有するコーパスが存在しないこと等から、研究の進展が進んでいない。また、音声認識技術についても、発話全体にわたる尤度を最小化するという基準を採用している。これはディクテーションのように、読み上げられた文章を音声認識技術を用いてコンピュータに入力し、後に人手で認識誤りを訂正することを想定した応用には適しているが、音声翻訳技術のように認識結果を翻訳するような処理には必ずしも適した基準ではない。このため、本研究開発課題の解決はますます重要な課題となりつつある。

### 3 研究開発の全体計画

#### 3-1 研究開発課題の概要

異言語間のスムーズなコミュニケーションを可能とするためには、話し手同士の関係、発話者の意図、文化的背景、場面、文脈といった発話外の状況を理解した上で、発話された内容を翻訳することが必要であり、このような機能を備えた音声翻訳システムの実現が究極のゴールということになる。

しかしながら、話し手同士の関係、発話者の意図、文化的背景、場面、文脈などの情報を適切に利用する音声翻訳技術を実現することは、現状の技術では不可能であり、長期的な基礎研究が必要である。一方、発話の中には前述の発話外の状況を利用せず、一文毎の表層情報のみを使用した翻訳であっても、相互に理解可能な場合も多く存在する。現在、実環境下で使用可能な多言語音声翻訳技術の実現は極めて要望の大きい急務の課題であることを考えると、前述の究極のゴールに向けて、一定期間毎に逐次適切な目標を設定し、それを達成する具体的な方策の立案と実施が不可欠である。

当面（今後4年間程度）達成すべき目標は、様々な実環境で話された音声言語を一文毎の表層情報のみを使用して翻訳する技術を確立し、異なる言語を話す人と人との実際のコミュニケーションの場面で、どの程度的確に情報を伝え得るのかを実データに則して検証することである。そのために、様々な実環境で種々の利用者の使用を可能とする音声認識技術、実環境下の多様な発話に対応できる言語翻訳技術の研究開発が必要である。

特に、言語翻訳技術については、従来、高度な知識をもった専門家の内観に基づき規則を構築していく構文トランスファー方式及びそれに一部用例翻訳を利用した方式が主に使用されてきた。構文トランスファー方式は、十分な量の対訳コーパスがなくても開発が可能であるという利点を有しているが、ドメインのカバレッジを拡大するために高度な知識をもった専門家の内観に基づき規則を再構築する必要があるという大きな欠点を有している。またカバレッジを客観的に知る方法がない。

一方、当研究所ではこれまでコーパスベース翻訳手法の研究開発を行ってきた。音声言語、特に対話は文字言語に比較して一発話の平均単語数が少ないことから、稠密なコーパ

スが収集可能であり同手法を効果的に適用できる。さらにこの手法を使うと分担してコーパスの収集ができるという利点があるため新しいドメインへの適用が容易となる。このため、大規模なコーパスを利用して言語翻訳を行うコーパスベース翻訳手法を中核的な技術と位置づけ、本技術の研究開発と共にコーパスの開発手法についても研究開発を進める。これらの要素技術を密結合して、信頼度指標を伴った翻訳結果を出力できるコーパスベース音声翻訳技術の研究開発を実施する。

具体的には、音声対話翻訳技術として最も広範囲な利用が想定される、海外旅行中の会話を対象に、多言語音声翻訳技術の研究開発を行う。言語対としては、利用可能な地域や話者数などの相手言語の持つ種々の影響力や、言語としての構造の疎遠なども考慮し、ほとんどの日本人がある程度の会話運用能力を有する英語を対象とした日英音声翻訳技術と、逆にほとんどの日本人が知識を持たない中国語を対象とした日中音声翻訳技術及びその他特定の言語と日本語との音声翻訳技術とする。

なお、音声翻訳技術という研究テーマの性質上、各国の研究機関との研究協力が重要と考えられる。このため、各国の研究機関と研究協力体制を確立し、当研究機関で中心的に研究開発を進める研究テーマと、相手研究機関との密接な研究協力の下で行う研究テーマ、相手研究機関の研究成果を研究開発に活かす研究テーマの選択を明確化し、並行的に研究を進めることとする。例えば、日本語の音声認識、音声合成、日英および日中の言語翻訳は当研究機関が中心的に研究を進める。対訳コーパスの開発については、言語対に応じて相手研究機関との密接な研究協力の下で進める。更に、英語、中国語の音声データベースの収集などについては、相手研究機関の研究成果を活かすなどの選択を行う。

実環境下で使用可能な多言語音声翻訳技術とそのための要素技術の研究開発を行い、音声翻訳技術の実使用環境での利用可能性を実証することが、本研究開発の目的であることから、研究開発の進め方としては研究期間中に定期的にフィールドでの評価試験を含む各種の評価試験を実施し、次期定期評価試験までの具体的な目標値を設定することにより、総合的な研究の進捗を加速する。更に、実用に繋がるテストベッドを構築し、実環境での評価・データ収集を実施する。以下、実環境での音声翻訳技術を使用可能とすることを主目的とする実音響環境での音声認識技術、音声認識結果と言語翻訳結果の信頼度指標を考慮して音声処理と言語処理を統合する音声言語統合技術、様々な言語対やドメイン(話題)での適用を効率的に可能とするコーパスベース対話翻訳技術、更にコーパスベース音声合成技術の各サブテーマについての研究内容、方針、研究手法等について述べる。

### 3-1-1 サブテーマ

#### (1) 実音響環境での音声認識技術

音声認識は、近年、長足の進歩を遂げている。この理由は、確率モデルと音声コーパスの整備が当研究所を含む研究機関により組織的になされたことによる。現在用いられている隠れマルコフモデルは、1970年代後半に提案された確率モデルに基づく手法であり、発話に伴う音声の特徴空間における時間的、空間的揺らぎを適切に表す特長を有している。

しかしながら、音声翻訳を目指した場合、現在の技術の性能は実際の利用環境では、未だ不十分と言わざるを得ない。実際に利用される環境では、種々の発話様式（発話スタイル）の発話が生じ、環境には、環境雑音、残響が存在するためである。本サブテーマでは、より実環境に近い環境での頑健な音声認識技術の確立を目指す。このような実環境における変動の要因は、一般に明示的に規則で表現できる種類のものでなく、これまで音声認識で一定の成功を収めたように、ある程度以上のコーパスと、構造・規則を反映した確率的モデルを用いる手法を適用するアプローチが最も有望である。そのためには、実際の状況で大量のコーパスを収集する必要がある、音声翻訳システムを利用しながら、コーパスを収集し、研究を進めるプロセスが必要となる。それには、実際の音響環境に頑健な音声認識が第一に重要な機能となる。本サブテーマでは、本プロジェクトで対象とする音声翻訳の課題に対し、実音響環境で頑健な音声認識を実現するための「音環境適応型音声認識技術」、実環境での音声翻訳性能を向上するための発話スタイル変形への頑健性を実現する「発話スタイル適応型音声認識技術」、音声翻訳が対象にする言語対を容易に増やすための「多言語音声認識技術」、実環境における使用において高い認識精度を確保するための「適応的入力発話リジェクション技術」の4つの研究開発を目標とする。

#### ア. 音環境適応型音声認識技術

音声認識性能は、昨今かなり進歩したが、実際に音声翻訳が利用されるような音響環境を考えた場合、音響雑音や部屋の残響、マイクロフォンの特性、伝送系の特性や雑音などの影響が大きく性能を劣化させる。そこで、実環境における音声翻訳システムの性能を高めるため、音響環境に頑健な音声認識を行うための、カルマンフィルタに基づいた定常、非定常雑音推定とフィルタリング手法を確立する。さらに、実環境における使用状況をより広くするため、遠隔発話の音声認識技術の確立を試みる。そのためには、複数のマイクロフォン素子により構成されるマイクロフォンアレーで指向性を制御する方法、音響雑音に影響を受けない発話顔画像を利用する方法を検討する。具体的には、4素子から8素子程度のアレーの利用による遠隔発話音声認識、ある程度の状況を仮定した中での非定常的雑音混入音声の認識、画像情報を統合した音声区間判定、音声認識手法の確立を目指す。

#### イ. 発話スタイル適応型音声認識技術

コーパスと確率モデルに基づく現在の手法は、学習データに含まれない入力音声に対して極めて脆弱である。従って、大量のコーパスから学習された音声認識システムでも、実環境において学習データにない発話スタイルの発話が入力されるとたちまち認識できなくなる。そのため、発話スタイル変形を分析し、種々の発話様式の音声を予測しながら認識する方法を確立し、実環境における音声認識性能の向上を達成する。たとえば、読みあげ音声と対話音声、さらには講演音声では発話スタイルが大きく異なるが、すべてのスタイルの音声を収集することは不可能に近く、全てのスタイルの音声を用いて学習したとしても性能劣化が避けられない。本研究では、このような発話スタイル変形に頑健な音声認識手法の確立を目指す。

#### ウ. 多言語音声認識技術

より多数の言語に対して音声翻訳を行うことは極めて重要であるが、対象とする言語毎に新たにコーパスを集め音声認識装置を学習するのは効率が悪い。言語間には、音響的、言語的に類似性が存在し、これらの類似性を利用すれば新しい言語に対し、少量のデータで音声認識システムを構築できる可能性がある。まず、英語と中国語に関して、大語彙音声認識システムを構築し、類似性の考察を行った後、国際音素記号体系を基本に、ユニバーサルな音響モデルを構築する手法の確立を目指す。

#### エ. 適応的入力発話リジェクション技術

音声翻訳システムは、人間-機械-人間の系である。これまでの音声翻訳は、音声認識誤りの有無にかかわらず正解入力として翻訳する構成であった。しかし、高い精度を保証するには、入力発話の信頼度を検証して、発話者に知らせ、必要により再発声を要求するリジェクション技術が必要となる。リジェクションするためには、本研究では、入力発話の音響的信頼性、言語的信頼性、ドメインとの整合性を検証し、入力発話に適応的にリジェクションを行う方法を確立する。

### (2) 音声言語統合技術

音声言語は音声としての音響的（物理的）特徴、言語としての統語的意味的（言語的）特徴を持ち、いずれも情報伝達に対して重要な役割を持つ。従って音声言語を正しく効率的に処理するためにはこれらを統合的に扱う必要がある。例えば、音声認識を行うためには音響的知識に加えて正確な言語的知識（言語モデル）が不可欠である。また、言語的単位である文の区切りを認識するためには、ポーズ長等の韻律情報、用言が終止形であるといった統語的情報、主語と述語の間の意味的關係などを総合的に利用する必要がある。さらに、処理結果の信頼度の自動評価を行うためには、音声認識誤り、発話の不完全さ、言語処理に使用する知識の不足等の可能性を総合的に考慮する必要がある。ところが、現状の音声言語処理システムは音声認識と言語処理とを直列に接続して、前者の出力が後者の入力となるようにした疎結合形式であるため、先に上げたような音声から言語にまたがる処理は限定的にしか実現されておらず、実世界で利用するシステムとしては不十分である。

そこで、本サブテーマでは音声認識と言語処理をシステムとして統合するために、「適応型音声言語モデル」「発話構造解析技術」「音声言語評価・最適化技術」という3つの研究項目を設定する。最初の2つは、音声情報と言語情報を考慮することによって、音声認識と言語処理との間の mismatches や gaps を解消する技術の確立を目指す。また、残りの1つは音声翻訳結果の自動評価法の開発とこの評価法を用いた最適解の探索やシステム制御に関する検討を行う。なお、翻訳結果や処理系の評価については試験データ（コーパス）に対する精度等のデータに基づく定量的なアプローチを取る。この手法が実世界に対して有効であるためには前提となるコーパスが対象世界の妥当なサンプルでなければならない。本テーマではこれを達成するために必要なコーパスの設計法、整備法についても検討を行

う。

#### ア. 音声言語モデリング技術

現状の音声認識用の言語モデルは単語の隣接関係を大量のコーパスから学習したNグラムに基づいている。しかし、このようなモデルにはいくつかの問題点がある。まず、学習コーパスへの依存性が高いため、新たな分野に適用すると性能が劣化する（すなわち当該分野の大量のコーパスを必要とする）。また、個々の単語に対して発音が固定的に対応付けられているため、構文に依存した発音の変形といった現象が捉えられていない。さらに、日本語や中国語に見られるように言語によっては単語境界や表記法が自明でないため、多言語化にあたっては学習コーパス上の表記のゆれや未知語境界の推定誤り等の問題が避けられない。本小項目では、まず、音素レベル、句レベル等多様なレベルで言語現象の一般性を捉えることによってドメインへの依存性の低い言語モデルを検討するとともに、構文に依存した発音の変形等のモデル化を目指す。また、この検討と並行して対象ドメインへの言語モデル適応手法も検討する。言語モデルの多言語化については「隠れマルコフモデル」等の統計的手法による単語境界の自動推定などをもとに検討を進め、中国語等への適用を目指す。

#### イ. 発話構造解析技術

話し言葉では意味的な切れ目を表す句読点等が明示されないため、韻律情報と発話内容の双方から処理単位を推定する必要がある。しかし、現状の音声認識系は発話内容と関係なくポーズによって区切られる単位で処理を行っているため、認識結果が言語処理にとって適切な意味的単位にはなっていない。また、現状の音声認識では誤りの混入が避けられず後段の言語処理に対して悪影響を及ぼす。本小項目ではポーズ長などの発話の物理的特徴に加えて、言語モデルやコーパスから抽出される統語的、意味的知識を用いることによって、現在の音声認識では困難な意味的まとまりの検出を行う。また、同様の知識と音声認識段階で得られる信頼度の情報を用いることによって認識結果の修正、および、部分的な情報抽出を試み、言語処理に対して最大限有用な情報を出力する手法を検討する。

#### ウ. 音声言語評価・最適化技術

音声認識処理と翻訳等の言語処理とを統合して全体として最適化する手法を確立する。まず、最適化の大前提として、処理対象を客観的に表現するデータの構築、すなわち、コーパスの構築が必要である。

コーパス構築にあたっては、既存のコーパスをパラフレーズ実験等によって拡張するとともに、音声認識システムを利用して、実環境下の音声対話を自動的に書き起こし、コーパスを拡張していく。これらを体系的に行うため、従来あまり検討されて来なかった、コーパスの網羅性やサンプルとしての妥当性に関する検討を実証的に行う。また、これと並行して、システムの性能に対する適切な評価関数の決定、すなわち与えられたコーパスに対してシステムの性能を定量的に評価する手法について検討する。具体的な手法としては、



正解文例とシステム出力文の間の類似度を適当な照合アルゴリズムによって計算し評価結果に対応づける手法などが考えられる。以上の検討をもとに、システム出力に対する信頼性の計算手法、および、システムパラメータの最適化手法の確立を目指す。

### (3) コーパスベース対話翻訳技術

従来の機械翻訳システムは規則によって動作を制御する形式のものを中心に研究開発されてきた。規則が中心的に用いられてきた主な理由としては、多様な言語現象に関するデータを網羅的に集めるのは容易でないこと、特に十分な量の対訳データを確保するのは困難であることが挙げられる。即ち、人間の類推能力を活用して言語現象を抽象化して言語データの不足を補完することにより、翻訳システムのカバレッジを拡大しようとしてきたと考えられる。しかし、このような実現形態では、他のドメインにシステムを移植したり、新たなデータに合うようシステムを改良したりするのが容易でない。

即ち、用意されたデータに素早く適用できるようにシステムを構成するコーパスベースの手法の実現が急務である。また、コーパスベースの手法であれば、多言語への展開も容易であると考えられる。しかし、現時点ではコーパスベースの手法は狭いドメインを対象として実現されているに過ぎず、翻訳精度も構文トランスファー方式を上回っているとは言い難い。そこで、本サブテーマでは、対話に関する大量のデータを収集するとともに広いドメインに適用可能な「コーパスベース言語変換技術」の実現を目指す。

#### ア. コーパスベース言語変換技術

音声翻訳に関する潜在的な要請を踏まえ、日本人が海外旅行する際の会話支援、日本国内で外国人旅行者に対する会話支援を対象として、実際に行われる会話の対訳データを収集する。そして、この対訳データを直接的に利用して翻訳する用例ベースの翻訳手法と、対訳データを統計的に処理して統計モデルを作成しそれを利用して翻訳する統計的翻訳手法を検討する。いずれのアプローチにおいても、検討に使用するドメインや言語対への依存性を排除するように務め、新たな言語対や異なるドメインに容易に適用可能なコーパスベースの手法として確立する。具体的には、用例ベース翻訳手法では、事前に準備するデータへの依存性が高いことから、短文への適用性が高いのに対し、長文への適用性が低いことが予想されるので、表現単位毎に分割して適用する等の頑健性の向上を目指す。また、統計的翻訳手法では、データ量に依存して翻訳モデルが大きくなることが予想されるので、翻訳モデルの効率的な作成方法の確立とともに、訳文生成のための計算時間の削減を進める。また、本課題のベースとなる言語データの収集は、アメリカやヨーロッパに比べ、アジア地域では立ち遅れていたが、最近では中国や韓国等でも国家的プロジェクトとして進められている。特に韓国では言語データの分析も精力的に行われている。このような情勢を踏まえ、言語データの効率的な収集や分析手法の確立にも留意する。

### (4) コーパスベース音声合成技術

コーパスベース音声合成においては、音声コーパスの規模が大きいほど音韻的・韻律的

多様性が広がるため音質的に有利である。このため、近年、音声コーパスを大規模化する傾向が強まっている。しかしながら、コーパスの大規模化には、(1)音声合成システムの開発コストの増大、(2)このため、多様な話者を用意することが困難、(3)所要記憶容量が大きいため携帯情報機器への搭載が困難、という負の側面がある。また、コーパス規模を拡大するにつれて音質改善量は次第に飽和するため、むやみにコーパスを拡大しても意味がない。そこで、100時間程度の音声コーパスを作成し、その範囲内でコーパス規模と合成音の音質との関係を定量的に解明する。また、インターネット技術を活用した評価実験の導入を通して実験参加者層の拡充を図り、主観評価データの信頼性・普遍性を高めることにより、単位選択基準の精度を向上する。

#### ア. コーパス設計

コーパスベース音声合成では、合成単位と呼ばれる、数音素程度の長さの音声波形を接続して合成音声を生成する。合成単位の出現頻度には偏りがあり、その分布はドメインによって異なる。出現頻度の高い合成単位ほど合成音の音質への寄与が大きいが、聴覚的に弁別可能なしきい値以下の差分しか持たない合成単位を複数用意する必要はない。しかし、無計画にコーパス規模を大きくするとそのような合成単位が数多く含まれるようになり、そのほとんどは音質の改善につながらない無駄な部分となる。そこで、音声コーパスの発声者が読み上げるテキストを最適に設計することにより、無駄の少ない音声コーパスを作成する手法を開発する。

#### イ. 知覚実験による単位選択の高品質化

合成単位を音声コーパスから抽出する際の判定基準は、聴覚上の自然性とよく対応のとれたものでなければならない。しかし、人間の感じる自然性の背景には複数の要因が存在しているため、単純に自然性に関する全体的な印象を調べる形式の知覚実験を積み重ねるだけでは明快な結果を得ることが困難であった。そのため知覚的印象との対応関係の精度が曖昧なままの物理尺度を単位抽出の判定基準として用いることが多く、合成音の品質向上が阻害されてきた。コーパスの設計、接続の方法などの合成手法の諸要素別に知覚的感度を精密に測定し、その結果に基づいた知覚的自然性に対する予測モデルを構築することによって音質の向上を図る。

#### 研究開発課題の概要に現れるキーワードリストと説明

・コーパス：形態素などのタグが付され、コンピュータで処理可能な言語資料（音声言語を含む）をコーパスと呼ぶ。コーパスベースとは、コーパスを直接的若しくはその統計的な性質を利用して音声言語処理を行う技術を総称として、コーパスベースと呼ぶ。

・隠れマルコフモデル (HMM)：シンボル出力に対して状態遷移が確定できないマルコフモデルで、音声信号のような非定常信号源の特徴を近似する手法として、広範囲に使用されて

いる。

- ・音響モデル：音声認識の際に使用する入力波形の音響的特徴を表すモデルで、近年はほとんど隠れマルコフモデルで表現される。

- ・言語モデル：音声認識の際に使用する単語のつながり方を示すモデルで、近年はほとんどNグラムで表現される。

- ・カルマンフィルタ：状態空間モデルと観測モデルで時系列信号の性質を表現する手法で、予測値を求めるのに使用する際にはプレディクション、雑音に埋もれた過去の値を検出する際にはスムージング、現時点の値を求める際にはフィルタリングと呼ばれる。

- ・国際音素記号体系：現存する人類の諸言語において語の意味の対立に貢献している言語音に、アルファベットを基盤とした記号を組織的に割り振った記号体系で、国際音声学協会（IPA）が提案している体系。

- ・Nグラム：単語間の接続を遷移確率で表現した言語モデルで、N個の接続を表現する場合をNグラムと呼ぶ。2個及び3個の接続を表現する場合は、特にバイグラム、トライグラムと呼ぶ。

- ・用例ベース：予め人手により求められた規則に基づくのではなく、収集された用例を直接使用して行う自然言語処理技術を用例ベースと呼ぶ。例えば、用例ベース翻訳手法などである。実例型と呼ばれることもある。

- ・換言処理：ほぼ同一の意味や類似の意味を有する表現に変換することを換言処理と呼ぶ。自然言語処理の新しい手法として、最近注目を浴びている。

## 3-2 研究開発目標

### 3-2-1 最終目標（平成18年3月末）

「大規模コーパスベース音声対話翻訳技術の研究開発」

- (1) コーパスベース翻訳技術に基づく実環境で使用可能な音声翻訳技術を実現すること  
その具体的な実現例として、通常の短期滞在の海外旅行での会話で一般的に現れる表現に対して、「日常生活のニーズを充足し、限定された範囲内では業務上のコミュニケーションができる」とTOEIC協会で評価されているクラス（470点から730点）の平均以上の日本人による翻訳と同等の質の翻訳を実現可能な日英音声翻訳技術を実現すること。なお、英日音声翻訳の性能については、日英音声翻訳の性能と同レベルであること（注）

- (2) 大規模で稠密な言語コーパスの開発とコーパスの網羅性などの特性を評価する手法の確立

注：日本語や中国語等については、外国人の英語のコミュニケーション能力を数量化する TOEIC や TOEFL などの数量化された指標がないため、翻訳の質は主観評価による。

#### ア. 実音響環境での音声認識技術

- (1) 小規模マイクロフォンアレーと発話顔画像を用いて雑音源のある実音響環境での遠隔発話の日本語、英語、中国語の音声認識が可能なこと
- (2) 不適切な入力、ドメイン外の入力発話をリジェクトする能力を有すること

#### イ. 音声言語統合技術

- (1) 未登録語が現れるなどの実環境において頑健な音声認識が可能な日本語言語モデルを構築すること。また、これと同等のタスクに対する中国語、英語の言語モデルを構築すること
- (2) 発話の境界、発話内の基本要素が正しく解析できること
- (3) 音声翻訳システムの自動評価と最適化が可能であること。そのために必要なコーパスの収集に関するガイドラインが存在すること

#### ウ. コーパスベース対話翻訳技術

- (1) 対訳データ量が十分に与えられる場合、極端に長くない旅行会話のテキスト入力に対し、TOEIC800 点の日本人による翻訳と同程度の日英翻訳性能を実現すること
- (2) 対訳データがあまり多くない場合、極端に長くない旅行会話のテキスト入力に対し、中国語会話能力中級以上の日本人による翻訳と同程度の日中翻訳性能を実現すること
- (3) 上記二項目の実現に十分な日英、日中の対訳データ及び対訳辞書などの収集を行うこと

#### エ. コーパスベース音声合成技術

- (1) コーパス規模と音質の関係を明らかにし、効率的なコーパス設計手法を開発すること
- (2) 同手法の有効性を示すため、コーパスベース手法で到達可能な最高品質のテキスト音声合成システムの開発を目指し、また、同手法が日本語以外の言語にも適用可能であることを示すこと

### 3-2-2 中間目標（平成16年3月末）

「大規模コーパスベース音声対話翻訳技術の研究開発」

- (1) 通常の短期滞在の海外旅行での会話で現れる表現の多く網羅できるコーパスを開発すること
- (2) 研究室環境で、最終目標と同等レベルの技術を達成すること
- (3) 実環境での試験・データ収集・評価を実施できるテストベッドを開発すること

ア. 実音響環境での音声認識

- (1) 中規模マイクロフォンアレーを用いて雑音源のある実音響環境での遠隔発話の日本語英語の音声認識が可能なこと
- (2) 発話様式として不適切な入力発話をリジェクトする能力を有すること

イ. 音声言語統合技術

- (1) 発話の境界が正しく解析できること。誤りの一部が修復できること
- (2) コーパス収集に関するガイドラインを作成し、システムの自動評価が可能であること

ウ. コーパスベース対話翻訳技術に関する研究開発

- (1) 対訳データ量が十分に与えられる場合、旅行会話短文のテキスト入力に対し、TOEIC750点の日本人による翻訳と同程度の日英翻訳性能を実現すること
- (2) 対訳データがあまり多くない場合、旅行会話短文のテキスト入力に対し、中国語会話能力中級程度の日本人による翻訳と同程度の日中翻訳性能を実現すること
- (3) 上記(1)(2)の実施に必要な日英、日中の対訳データ及び対訳辞書などの収集を行うこと

エ. コーパスベース音声合成技術

- (1) 100時間程度のコーパスを用いた日本語 TTS システム、および 20 時間程度のコーパスを用いた中国語 TTS システムを実現すること

3-3 研究開発の年度別計画

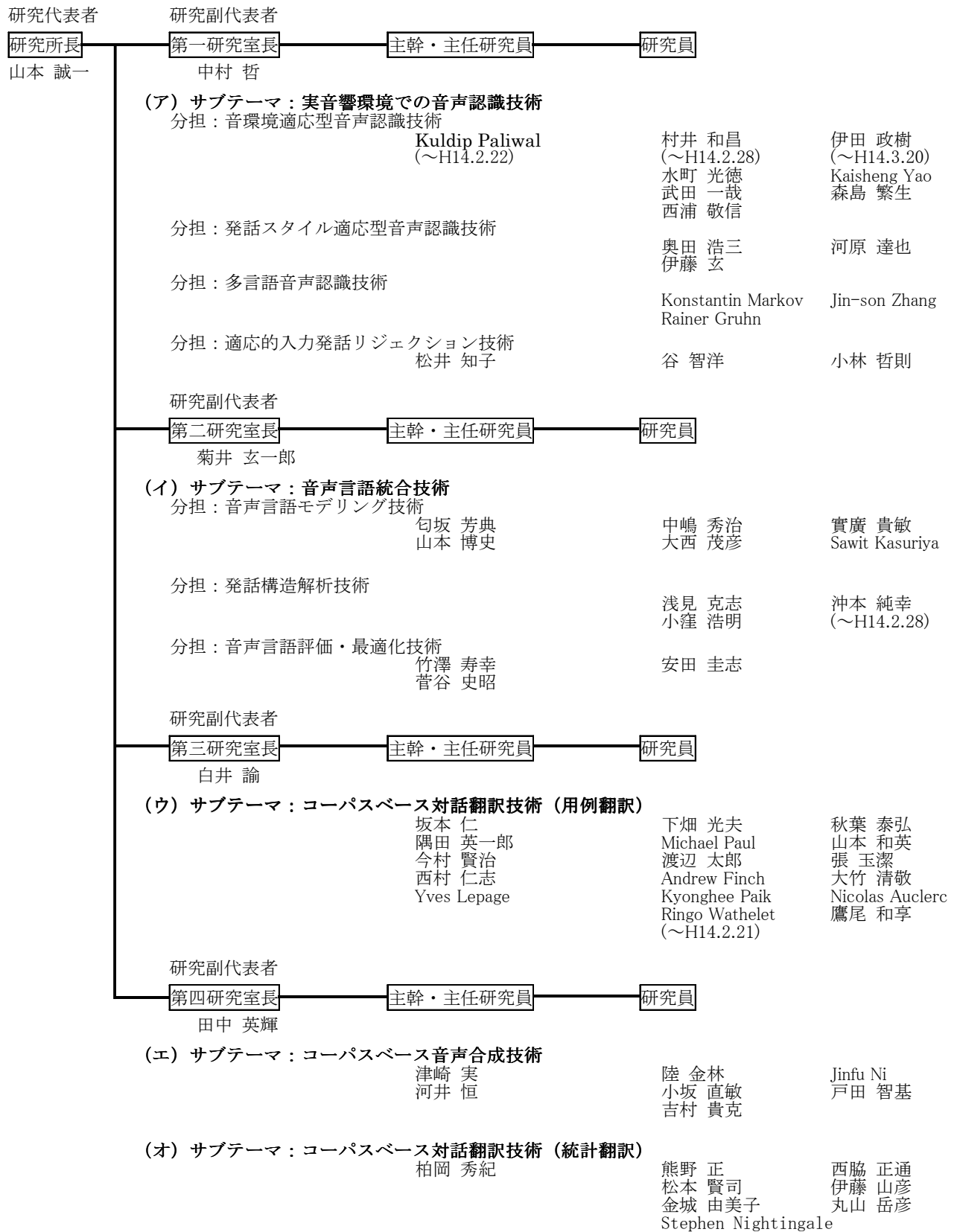
(金額は非公表)

研究開発項目	13年度	14年度	15年度	16年度	17年度	計	備考
「大規模コーパスベース音声対話翻訳技術の研究開発」							
ア 実音響環境での音声認識					→		
イ 音声言語統合技術					→		
ウ コーパスベース対話翻訳技術（用例翻訳）					→		
エ コーパスベース音声合成技術					→		
オ コーパスベース対話翻訳技術（統計翻訳）					→		
間接経費							
合計							

注) 1 経費は研究開発項目毎に消費税を含めた額で計上。また、間接経費は直接経費の30%で計上（消費税を含む。）。

2 備考欄に再委託先機関名を記載

### 3-4 研究開発体制



## 4 研究開発の概要（平成 13 年度）

### 4-1 研究開発実施計画

#### 4-1-1 研究開発の計画内容

##### ア. 実音響環境での音声認識

- ①規模マイクロフォンアレーを用いて雑音源のある実音響環境での遠隔発話の日本語、英語の音声認識するための実環境データ収集と基本方式検討
- ②発話様式データベースの収集と適応入力リジェクションの基本方式検討

##### イ. 音声言語統合技術

- ①（言語モデル作成用）多言語対応コーパス解析法の検討
- ②コーパスとのマッチングによる誤り訂正法の検討
- ③パラフレーズによるコーパス拡張法の検討

##### ウ. コーパスベース対話翻訳技術（用例翻訳）

- ①類似用例検索による翻訳可能性の検討と統計翻訳基本モデルの基本設計
- ②中日翻訳用基本パターン生成の検討と中国語および日本語の換言処理の基本設計
- ③日英中の旅行会話 20 万文の整備

##### エ. コーパスベース音声合成技術

- ①聴覚実験による合成音声評価

##### オ. コーパスベース対話翻訳技術（統計翻訳）

- ①新聞コーパス収集と日本語記事対応手法検討



4-1-2 研究開発課題実施計画

(金額は非公表)

研究開発項目	第1四半期	第2四半期	第3四半期	第4四半期	計	備考
「大規模コーパスベース音声対話翻訳技術の研究開発」						
ア 実音響環境での音声認識				→		
イ 音声言語統合技術				→		
ウ コーパスベース対話翻訳技術 (用例翻訳)				→		
エ コーパスベース音声合成技術				→		
オ コーパスベース対話翻訳技術 (統計翻訳)				→		
小計						
間接経費						
合計						

- 注) 1 経費は研究開発項目毎に消費税を含めた額で計上。また、間接経費は直接経費の30%で計上(消費税を含む)。  
 (合計の計は、「3-1の研究開発課題必要概算経費」の総額と一致)
- 2 備考欄に再委託先機関名を記載。

## 4-2 研究開発の実施内容

### ア. 実音響環境での音声認識

- ①中規模15chマイクロホンアレーの試作と発話者方向検出、遅延和ビームフォーマの実装を行った。また、2素子ペア型減算型マイクロホンアレーを発展させ音声認識に利用するための基礎検討を行った。今後、このシステムを利用し実環境データ収録を進める。
- ②発話速度の速い発話、誤り訂正発話などのデータ収集を行い分析を行った。発話速度などの発話様式に応じた可変フレームレート、マルチパスモデルによる方式の検討を進めた。

### イ. 音声言語統合技術

- ①多言語対応コーパス解析法については1) 複数言語のコーパスを使った言語モデルの適応処理、および、2) 対訳情報を使ってアジア系言語のテキストを単語に分割する処理の研究に着手した。
- ②誤り訂正法については、大量のコーパスを使って誤り箇所を訂正する手法に関する調査と研究環境の整備を行った。
- ③パラフレーズによるコーパス拡張法の検討については、既存のパラフレーズ手法によるデータ収集とその結果の検討を行った。

### ウ. コーパスベース対話翻訳技術（用例翻訳）

- ①用例翻訳エンジン、パターン翻訳エンジン、統計翻訳エンジンの基本設計と試作を行なった。いずれも、後述の③で整備された18万文を学習データとして使用し、それぞれの翻訳エンジンで使用する翻訳知識をそれぞれ異なる機械学習方式により学習する。用例翻訳エンジンは、単語の一致またはシソーラス（類語集）上の距離により、入力文と最も類似するデータベース中の文を抽出し、対訳辞書を使用して訳文を生成する。パターン翻訳エンジンは、日英それぞれの構文解析結果のうち部分木の対応関係をあらかじめ取り出しておき、翻訳知識として利用する。統計翻訳エンジンは、原言語と目的言語の単語の対応関係を確率値としてあらかじめ計算しておき、その確率モデルを用いて翻訳を行なう。また、これら3種類の翻訳エンジンから得られる訳文を最適に組合せるための訳文品質自動評価法の検討を開始した。
- ②原言語と目的言語間の言語変換処理と、原言語または目的言語の単一言語内の換言処理に分けて検討し、それぞれ基本方式と試作を行なった。技術の汎用性を目指す観点から、相互依存を極力排除するよう努めた。言語変換処理では、目的言語の単一言語コーパスを利用して、変換処理により得られた断片的な訳をこなれた表現に再構成する方法を検討した。換言処理では、中国語と日本語のそれぞれについて、語順の入れ換えや同義語への置換等を言語知識として定義することにより、異なる表現を代表的な表現に対応付けたり、1つの表現を複数の異なる表現に展開する基本動作を確認した。
- ③市販の例文集などを参考にして、旅行会話に関する基本表現を、日英対訳で20万文収集

した。日本語および英語表現の妥当性を検証するとともに、口語の中国語訳を付与することにより、日英中の3カ国語対照の対訳データベースを構築した。日本語と英語については形態素解析を行ない、単語単位の基本情報付与を行なった。それに合わせて、例文中に新たに出現した単語などに対し、対訳辞書整備を行なった。

#### エ. コーパスベース音声合成技術

- ①音素片接続時の音素環境が入れ替わることによる自然性の低下について、日本語の音素の組み合わせについて網羅的に知覚評価実験を実施した。これによって得られた評価値は、素片接続選択実施のためのコスト表として活用した場合の合成音声について予備的な試聴を実施した。

#### オ. コーパスベース対話翻訳技術（統計翻訳）

- ①日経新聞社から貸与された日本語記事と英語の翻訳には対応関係がないため、これを回復するための手法を検討した。考案した手法は1) あらかじめ日本語記事中にある会社名や出版日によって対応する英語記事候補を選択し、2) 共有する名詞の度合いで記事間の近さを計測する。英語の記事の名詞を日本語にするために市販の翻訳システムを使用した。

## 5 研究開発実施状況（平成13年度）

### 5-1 実音響環境での音声認識技術の研究開発

#### 5-1-1 序論

現在までの音声認識の研究では、大量の音声コーパスに基づく隠れマルコフモデルやN-gram言語モデルなどの確率モデルの研究が中心となってきた。確率モデルは発話に伴う音声の特徴空間における時間的、空間的揺らぎを適切に表す特長を有している。しかしながら、音声翻訳を目指した場合、現在の技術の性能は実際の利用環境では、未だ不十分と言わざるを得ない。実際に利用される環境では、種々の発話様式（発話スタイル）の発話が生じ、環境には、環境雑音、残響が存在するためである。本サブテーマでは、より実環境に近い環境での頑健な音声認識技術の確立を目指す。具体的には、本プロジェクトで対象とする音声翻訳の課題に対し、実音響環境で頑健な音声認識を実現するための「音環境適応型音声認識技術」、実環境での音声翻訳性能を向上するための発話スタイル変形への頑健性を実現する「発話スタイル適応型音声認識技術」、音声翻訳が対象にする言語対を容易に増やすための「多言語音声認識技術」、実環境における使用において高い認識精度を確保するための「適応的入力発話リジェクション技術」の4つの研究開発を目標とする。

特に、平成13年度は、マイクロホンアレーを用いた音環境適応型音声認識技術として、中規模マイクロホンアレーによる遠隔発話受音システム、小規模マイクロホンアレーによ

る遠隔発話受音の検討を行った。また、種々の環境に於ける発話スタイルの変化への適応として、誤り訂正発話、発話速度の速い発話の認識性能改善について検討を行った。

### 5-1-2 委託業務の内容

平成13年度は5-1-1の方針に基づき、下記の研究を遂行した。

- ①中規模15chマイクロホンアレーの試作と発話者方向検出、遅延和ビームフォーマの実装を行った。また、2素子ペア型減算型マイクロホンアレーを発展させ音声認識に利用するための基礎検討を行った。今後、このシステムを利用し実環境データ収録を進める。
- ②発話速度の速い発話、誤り訂正発話などのデータ収集を行い分析を行った。発話速度などの発話様式に応じた可変フレームレート、マルチパスモデルによる方式の検討を進めた。

### 5-1-3 委託業務の効果

平成13年度の研究の結果、下記の研究成果を得た。

- ①試作した中規模マイクロホンアレーによる遠隔発話受音システムがほぼリアルタイムに音源同定、遅延和ビームフォーミングを達成した。小規模マイクロホンアレーでは、3chによる減算型処理で音声認識を対象としたSNR改善が行える見通しがあった。
- ②誤認識時の誤り訂正発話に対しマルチパスモデルと話者適応化の利用で誤り訂正時の誤認識率が約60%回復した。

### 5-1-4 他の研究機関における類似研究及び協力関係状況

実環境における音声認識の研究は、非定常雑音に対処するアルゴリズムの研究と、マイクロホンアレーを利用した研究との2つが現在のトレンドであるが、いずれの研究でも当研究所の研究が分野をリードする位置づけとなっている。

### 5-1-5 まとめ、今後の課題等

上記のように平成13年度は、中規模マイクロホンアレーを用いた遠隔発話音声受音システムのプロトタイプを作成し、音声認識との結合を検討した。また、小規模マイクロホンアレーの基礎検討を進めた。一方、音声認識の性能を高めるために不可欠な発話スタイルへの適応化について、誤り訂正時の発話スタイル変形、発話速度の速い発話の取り扱い方法の検討を行い、基本的な方式の確立を行った。これらの進捗を鑑みて、平成14年度は下記の課題を中心に研究を進める。

- ①小規模マイクロホンアレーを用いて雑音源のある実音響環境での遠隔発話の日本語、英語の音声認識するための実環境データ収集と基本方式検討

## ②発話様式データベースの収集と適応入力リジェクションの基本方式検討

### 5-2 音声言語統合技術の研究開発

#### 5-2-1 序論

本サブテーマの目標は音声認識（サブテーマ1）と翻訳処理（サブテーマ3）を統合して最適な音声翻訳処理を実現するための手法を開発することである。具体的な研究項目として、音声認識処理に対して言語的な手がかりを与える「適応型音声言語モデル」、音声認識結果を解析して自動修正等を行う「発話構造解析技術」という2つの要素技術と、これらの指標となる「音声言語処理系の自動評価」のあわせて3つを設定している。なお、処理系の評価については試験データ（コーパス）に対する精度等のデータに基づく定量的なアプローチを取るため、実世界を反映したコーパスの設計法、整備法についても検討を行う。

#### 5-2-2 委託業務の内容

先にあげた3つの項目に対応して今年度は以下の研究活動を行った。

- ①適応型言語モデルについては、1) 複数言語のコーパスを使った言語モデルの適応処理、および、2) 対訳情報を使ってアジア系言語のテキストを単語に分割する処理の研究に着手した。
- ②発話構造解析については、研究環境の整備を行うとともに、誤り箇所を既知とした上で、大量のコーパスとのマッチングでどの程度訂正可能かを検討した。
- ③音声言語処理系の自動評価に関しては、コーパス収集を中心に検討を行い、木構造表現を用いたパラフレーズ手法によるデータ収集を行った。

#### 5-2-3 委託業務の効果

先にあげた①から③の活動の結果、以下のような効果があった。

- ①1) および 2) に関してそれぞれ基本的な手法を考案し、小規模な評価実験を行った結果、これらの手法が有望であることを確認した（国際会議に投稿）。
- ②DP を使った誤り訂正手法を実装し研究用のツールとして整備した。
- ③「チャートを用いたパラフレーズ法」を用いてコーパスの拡張と整備を行った。その結果、1文に対して平均で100文のパラフレーズの生成を行うことができた。

#### 5-2-4 他の研究機関における類似研究および協力関係状況

- ①の1) については類似の研究は見当たらない。但し、要素技術として使用した統計的翻訳についてはアーヘン工科大学（独）、南カリフォルニア大学（米）などで活発な研究が行われており、これらの機関の研究者とはATRへの招待や学会活動等で意見交換を行っている。2) については人手で作成された学習データを用いた分割手法はNTT等で研

究されているが、本研究ではそのようなデータを必要としない点が異なる。最近、西欧系言語以外に対する翻訳研究の必要性が高まりつつあるが、本テーマはその要素技術に位置付けられる。

- ②誤り訂正法に関しては松下電器の研究所で目的は同じであるが別のアプローチによる研究が行われており、今後、交流を進める予定である。
- ③パラフレーズの研究は日本国内で盛んであるが、主にいかにして自動的にパラフレーズを行うかに主眼が置かれており、本研究のように人間からいかに大量の同義文を引き出すかということはあまり試みられていない。

#### 5-2-5 まとめ、今後の課題等

以上のように当年度予定していた研究項目は概ね予定とおりに進捗した。今後の課題として、①の言語モデルに関してはNグラムを超えたモデルの導入、②の音声言語解析に関しては翻訳出力の改善という観点からの認識結果修正処理の検討、③の評価に関しては翻訳性能と認識性能との関係のより詳細な分析とこれを行うための実世界に対応したコーパスの作成が挙げられる。

### 5-3 コーパスベース対話翻訳技術の研究開発

#### 5-3-1 序論

従来の機械翻訳システムは規則によって動作を制御する形式のものを中心に研究開発されてきた。規則が中心的に用いられてきた主な理由としては、多様な言語現象に関するデータを網羅的に集めるのは容易でないこと、特に十分な量の対訳データを確保するのは困難であることが挙げられる。しかし、このような実現形態では、他のドメインにシステムを移植したり、新たなデータに合うようシステムを改良したりするのが容易でない。このため、用意されたデータに素早く適用できるようにシステムを構成するコーパスベースの手法の実現が急務である。また、コーパスベースの手法であれば、多言語への展開も容易であると考えられる。

音声翻訳に関する潜在的な要請を踏まえ、日本人が海外旅行する際の会話支援、日本国内で外国人旅行者に対する会話支援を対象として、実際に行われる会話の対訳データを収集する。そして、この対訳データを直接的に利用して翻訳する用例ベースの翻訳手法と、対訳データを統計的に処理して統計モデルを作成しそれを利用して翻訳する統計的翻訳手法を検討する。いずれのアプローチにおいても、検討に使用するドメインや言語対への依存性を排除するように務め、新たな言語対や異なるドメインに容易に適用可能なコーパスベースの手法として確立する。

具体的には、用例ベース翻訳手法では、事前に準備するデータへの依存性が高いことから、短文への適用性が高いのに対し、長文への適用性が低いことが予想されるので、表現単位毎に分割して適用する等の頑健性の向上を目指す。また、統計的翻訳手法では、データ量に依存して翻訳モデルが大きくなることが予想されるので、翻訳モデルの効率的な作

成方法の確立とともに、訳文生成のための計算時間の削減を進める。

また、本課題のベースとなる言語データの収集については、先にあげた項目に対して今年度は以下の研究を実施した。

### 5-3-2 委託業務の内容

①用例翻訳エンジン、パターン翻訳エンジン、統計翻訳エンジンの基本設計と試作を行なった。いずれも整備された18万文を学習データとして使用し、それぞれの翻訳エンジンで使用する翻訳知識をそれぞれ異なる機械学習方式により学習する。

用例翻訳エンジンは、単語の一致またはシソーラス（類語集）上の距離により、入力文と最も類似するデータベース中の文を抽出し、対訳辞書を使用して訳文を生成する。パターン翻訳エンジンは、日英それぞれの構文解析結果のうち部分木の対応関係をあらかじめ取り出しておき、翻訳知識として利用する。統計翻訳エンジンは、原言語と目的言語の単語の対応関係を確率値としてあらかじめ計算しておき、その確率モデルを用いて翻訳を行なう。また、これら3種類の翻訳エンジンから得られる訳文を最適に組み合わせるための訳文品質自動評価法の検討を開始した。

②中日翻訳手法については、原言語と目的言語間の言語変換処理と、原言語または目的言語の単一言語内の換言処理に分けて検討し、それぞれ基本方式と試作を行なった。技術の汎用性を目指す観点から、相互依存を極力排除するよう努めた。言語変換処理では、目的言語の単一言語コーパスを利用して、変換処理により得られた断片的な訳をこなれた表現に再構成する方法を検討した。換言処理では、中国語と日本語のそれぞれについて、語順の入れ換えや同義語への置換等を言語知識として定義することにより、異なる表現を代表的な表現に対応付けたり、1つの表現を複数の異なる表現に展開したする基本動作を確認した。

③対訳コーパスの整備については、市販の例文集などを参考にして、旅行会話に関する基本表現を、日英対訳で20万文収集した。日本語および英語表現の妥当性を検証するとともに、口語の中国語訳を付与することにより、日英中の3カ国語対照の対訳データベースを構築した。

更に、日本語と英語については形態素解析を行ない、単語単位の基本情報付与を行なった。それに合わせて、例文中に新たに出現した単語などに対し、対訳辞書整備を行なった。

### 5-3-3 委託業務の効果

上にあげた研究活動の結果、以下のような効果があった。

①類似用例検索による翻訳可能性の検討と統計翻訳基本モデルの基本設計については、コ

コーパスベースの考え方による3種類翻訳エンジンを試作し基本動作を確認した。

- ②中日翻訳用基本パターン生成の検討と中国語および日本語の換言処理の基本設計については、換言処理と言語変換による翻訳システムの基本部品技術の実現可能性を示した。
- ③20万文の旅行会話に関する日英中の3カ国語対照の対訳データベースを構築し、学習用データと評価用データを定義した。なお、20万文のうち、2万文を評価試験文として分離し、音声翻訳システムの性能を定点観測するために使用することとした。すなわち、本プロジェクト期間を通じて、方式検討やシステムが参照するデータベースは残り18万文に限定される。

#### 5-3-4 他の研究機関における類似研究及び協力関係状況

旅行会話翻訳としては、イタリアのITC-irstがイタリア語と英語間の統計翻訳を検討しており、データ等の共有と手法について意見交換を行なった。

#### 5-3-5 まとめ、今後の課題等

コーパスベース対話翻訳を実現するため、特徴の異なる3つのエンジン、用例翻訳エンジン、パターン翻訳エンジン、統計翻訳エンジンの基本動作を確認した。平成14年度には引き続きこの3つのエンジンの改良を進める。また、これら3種類の翻訳エンジンから得られる訳文を最適に組合せるための訳文品質自動評価法を検討し、最尤出力を選択する仕組みを実現する。中日翻訳に関しては、要素技術を結合することによりプロトタイプの構築を行なう。

中間評価時の性能目標として、日英翻訳では、旅行会話に現れる短文のテキスト入力に対しTOEIC750点の日本人による翻訳と同程度の翻訳性能を目指す。中日翻訳では、日英の性能に準ずる翻訳性能を目指す。

また、それに必要となる日英、日中の対訳データ及び対訳辞書などの収集を進める。

### 5-4 コーパスベース音声合成技術の研究開発

#### 5-4-1 序論

音声合成技術は、音声翻訳システムの出力機能、すなわち翻訳結果を音声として利用者に提示する機能を実現するものである。

音声翻訳システムから利用者に対して必要な情報が正確かつ円滑に伝わり、しかも違和感なく受け入れられるためには、合成音声の明瞭性・自然性が最も重要である。ATRの発明であるコーパスベース音声合成技術は、自然性の点で最も有望な技術であるが、明瞭性は実用レベルに達しているものの自然性はなお不十分である。

自然性の改善に寄与する要素技術は、読み・アクセント生成、韻律生成、音声素片選択、音声コーパス、韻律補正のための信号処理など多岐にわたるが、本委託研究においては、



過去に技術的解明が十分に行われていない素片選択と音声コーパス拡充を中心に研究開発を行う。素片選択においては、知覚特性とよく対応の取れたコスト関数の研究、およびそれにもとづく効率的な探索アルゴリズムの研究が主な課題である。一方音声コーパスに関しては、コーパス規模とそれにより実現される音質との関係の解明、さらに規模対音質の点で効率の高いコーパスを設計する手法の研究が主な課題である。

#### 5-4-2 委託業務の内容

本年度は、以下に示すように、主に素片選択の改良に取り組んだ。また、多言語音声翻訳の実現のために、日本語以外の言語、特に中国語に関しても日本語と技術の共通化を図りつつ高品質なテキスト音声合成システムの開発を行った。

##### ①素片選択用コスト関数の高精度化

音素片接続時の音素環境が入れ替わることによる自然性の低下に関して、日本語の音素の組み合わせについて網羅的に知覚評価実験を実施した。この結果を素片接続選択におけるコスト関数に組み込み、合成音声について予備的な評価試験を実施したところ、従来の単純な物理指標に基づいた選択に比して良好な音質をもたらすことが確認できた。

##### ②素片選択アルゴリズムの改良

母音間接続に起因する不連続感の発生を回避するため、音素単位とダイフオン単位を組み合わせたアルゴリズムを開発した。同アルゴリズムと従来法とを聴取実験により比較したところ、70%の選好率が得られ、有効性が実証された。

##### ③中国語音声合成システムの開発

20 時間程度の音声コーパスを収集した。このうち約 1 時間分の音声データについて音素・韻律ラベリングおよび言語情報付与を行い、音声合成に使用できる状態にした。

また、日本語用に開発した音声素片選択アルゴリズムのコスト関数を中国語向けに簡略化して中国語音声合成システムのプロトタイプを作成した。

#### 5-4-3 委託業務の効果

素片選択技術に関して、コスト関数および選択アルゴリズムの両面で進展があり、自然性の高い合成音の実現という目標に一步近づいた。コスト関数に関する研究の一環として得られた心理量の測定データは、今後心理量と対応性の高い客観的評価尺度を設計する上での基本データとなる。

中国語音声合成の研究は、初期段階ながら他社・他機関の発表しているデモ音声と比較して遜色のない音声を得られ、日本語版と要素技術の共通化を図りつつ研究を進める方針が妥当であるとの感触が得られた。

#### 5-4-4 他の研究機関における類似研究及び協力関係状況

コーパスベース音声合成技術は、研究レベルでは主流となった感があり、国外では、AT&T(ATR から供与した技術を元に研究を開始した)、Lucent, Microsoft, CMU, BT など、

国内では NTT, 東芝, IBM, 名工大, 東工大などがコーパスベースの枠組みで研究を行っている。しかし、音質面で実用レベルに近いと言える例は、世界的に見ても AT&T の商用 TTS(英語)などごくわずかであり、国内では皆無と言える。

要素技術に関しては、知覚実験を通して素片選択のコスト表を作成するというアプローチは AT&T、IPO などの研究機関も実施している。これらの対象言語である英語、オランダ語では音韻の種類が日本語に比べて多いこともあり、日本語とは異なる単位での接続を前提とした評価を、規模を縮小して行っている。今回の日本語での聴取実験は規模としてはそれらを上回るものである。

なお、研究リソースの選択的・集中的投入という観点から、本研究開発課題では素片選択および音声コーパス設計・作成を主な研究対象とする。韻律生成に関しては、現時点で最有望と思われる HMM ベースの方式を開発した名工大・東工大と教員・学生を客員として招聘するなど協力関係を保っている。一方韻律変形に関しては、高品質な信号処理技術を有する和歌山大学と随時情報交換を行っている。

#### 5-4-5 まとめ、今後の課題等

コーパスベース音声合成技術の枠組みの下で自然性の高い合成音声の実現を目指して研究開発を進め、一定の成果が得られた。次年度は、素片選択において残された課題すなわち韻律誤差・声質等の評価尺度の研究、コスト関数における重みの最適化、探索の高速化等に取り組むとともに、コーパス規模と音質の関係の検討を行う。また中国語に関しては、音声コーパスを整備するとともに、素片選択のコスト関数を中国語向けに最適化する研究を行う。

#### 5-5 総括

平成13年度の研究開発期間はわずか3ヶ月弱であったが、4つのサブテーマ、すなわち①実音響環境での音声認識技術の研究開発、②音声言語統合技術の研究開発、③コーパスベース対話翻訳技術の研究開発、④コーパスベース音声合成技術の研究開発に於て、当初研究計画通りの成果が得られた。

すなわち、

①「実音響環境での音声認識技術の研究開発」については、中規模マイクロホンアレーを用いた遠隔発話音声受音システムのプロトタイプの開発を行い、音声認識技術との結合を検討することにより、遠隔発話音声の認識に見通しをつけた。

②「音声言語統合技術の研究開発」については、パラフレーズ手法の検討により大規模コーパス開発に道筋をつけると共に、認識誤り箇所が既知とした場合に、言語情報の使用による訂正の可能性を検討した。

③「コーパスベース対話翻訳技術の研究開発」については、コーパスベース翻訳技術として、文単位の用例ベース翻訳、句単位の用例ベース翻訳、単語単位の統計的翻訳手法を実装し、基本動作を確認した。更に、日英中の旅行会話20万文の対訳データベースを整備した。

④「コーパスベース音声合成技術の研究開発」については、音声合成の際の素片選択アルゴリズム及びコスト関数の検討で進展が得られ、自然性の高い合成音の実現に一步近づいた。

特筆すべきものとしては、20 万文の旅行会話コーパスに関して任意に選択されたテスト文に対しては、翻訳性能としては既に中間目標を達成している。一方、任意に選択されたテスト文が実環境での発話の特徴を十分に表わしているかは未だ不明である。このため、現在別途収集の実環境でのデータを用いてテスト文の妥当性に関して検証を行う予定である。

以上のように、本研究開発は、全体計画に沿って順調に進んでおり、中間目標達成の見通しは明るい。

(添付資料)

1 研究発表、講演、文献等一覧

決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発059	国内大会	言語処理学会第8回年次大会	丸山 岳彦、柏岡 秀紀	「文節訳の出現順序」を考慮した翻訳単位の決定	2002.03.16 ~2002.03.20	発表済み
2001音-発060	国内大会	言語処理学会第8回年次大会	柏岡 秀紀	講演同時通訳データのアライメント	2002.03.16 ~2002.03.20	発表済み
2001音-発061	誌上	自然言語処理	今村 賢治	構文解析と融合した階層的句アライメント		投稿中
2001音-発063	国内研究会	電子情報通信学会音声研究会	戸田 智基、河井 恒、津崎 実、鹿野 清宏 (奈良先端大)	日本語テキスト合成における音素単位とダイフォン単位に基づいた単位選択	2002.01.24	発表済み
2001音-発064	国内大会	日本音響学会 2002 年春季研究発表会	戸田 智基、河井 恒、津崎 実、鹿野 清宏 (奈良先端大)	日本語テキスト音声合成における母音中心での接続を考慮に入れた単位選択	2002.03.18 ~2002.03.20	発表済み
2001音-発065	国内大会	日本音響学会 2002 年春季研究発表会	Jinfu NI, Hisashi KAWAI	Automatic Detection of Mandarin Prosodic Boundary Using Syntactic Information	2002.03.18 ~2002.03.20	発表済み
2001音-発066	国内大会	日本音響学会 2002 年春季研究発表会	河井 恒、津崎 実	波形接続型音声合成における音素環境代替コストの予測	2002.03.18 ~2002.03.20	発表済み
2001音-発067	国内大会	日本音響学会 2002 年春季研究発表会	実廣 貴敏、山本 博史、菊井 玄一郎	用例型構文解析による句構造情報を利用した言語モデル	2002.03.18 ~2002.03.20	発表済み
2001音-発068	国内大会	日本音響学会 2002 年春季研究発表会	Kaisheng YAO, Satoshi NAKAMURA	Mixture of Factor Analyzed HMM	2002.03.18 ~2002.03.20	発表済み
2001音-発069	国内大会	日本音響学会 2002 年春季研究発表会	奥田 浩三、南角 吉彦 (名古屋工業大学)、中村 哲	音響特徴パラメータの相関を利用した音素継続時間長の正規化	2002.03.18 ~2002.03.20	発表済み

決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発070	国内大会	日本音響学会 2002年春季研究発表会	Jin-song ZHANG, Satoshi NAKAMURA	A Bootstrap Approach to Develop Statistical Word Segmentation System for Chinese	2002.03.18 ~2002.03.20	発表済み
2001音-発071	国内大会	日本音響学会 2002年春季研究発表会	Jin-song ZHANG, Konstantin MARKOV, Tomoko MATSUI, Rainer GRUHN, Satoshi NAKAMURA	Developing Robust Baseline Acoustic Models for Noisy Speech Recognition in SPINE2 Project	2002.03.18 ~2002.03.20	発表済み
2001音-発072	国内大会	日本音響学会 2002年春季研究発表会	Tomoko MATSUI, F.K. SOONG (Bell Labs), B.H. JUANG (AVAYA Labs)	Classification Design for Verification of Multi-class Recognition Design	2002.03.18 ~2002.03.20	発表済み
2001音-発073	国内大会	日本音響学会 2002年春季研究発表会	Konstantin MARKOV, Tomoko MATSUI, Rainer GRUHN, Satoshi NAKAMURA	Robust Speech Recognition in Diverse Noisy Environments	2002.03.18 ~2002.03.20	発表済み
2001音-発074	国内大会	日本音響学会 2002年春季研究発表会	西浦 敬信、Rainer GRUHN、中村 哲	マイクロホンアレーとビデオカメラを用いた遠隔発話音声翻訳会議システムの試作	2002.03.18 ~2002.03.20	発表済み
2001音-発075	国内大会	日本音響学会 2002年春季研究発表会	水町 光徳、中村 哲	2ch 減算型ビームフォーマの線音源への適用	2002.03.18 ~2002.03.20	発表済み
2001音-発076	国内大会	日本音響学会 2002年春季研究発表会	小窪 浩明、林 輝昭(メディアックス)、山本 博史、菊井 玄一郎	後続仮説数の管理機能を用いた単語グラフの仮説数削減	2002.03.18 ~2002.03.20	発表済み
2001音-発077	国内大会	日本音響学会 2002年春季研究発表会	津崎 実、河井 恒	STRAIGHT と PSOLA による単語音声分析合成音の比較評価	2002.03.18 ~2002.03.20	発表済み
2001音-発078	国内大会	日本音響学会 2002年春季研究発表会	津崎 実、河井 恒	素片接続型音声合成器の単位選択特徴量についての検討:聴覚モデルと MFCC による選択の聴取評価	2002.03.18 ~2002.03.20	発表済み

決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発080	国内大会	日本音響学会 2002 年春季 研究発表会	Norbert BINDER, Rainer GRUHN, Satoshi NAKAMURA	Recognition of Non-Native Speech Using Dynamic Phoneme Lattice Processing	2002. 03. 18 ~2002. 03. 20	発表済み
2001音-発082	国内研究会	電子情報通信学会音声言 語情報処理研究会	村井 和昌、中村 哲	マルチモーダル音声認識のため の顔検出	2002. 02. 01 ~2002. 02. 02	発表済み
2001音-発083	国際会議	LREC 2002 (Third International Conference on Language Resources and Evaluation)	Taro WATANABE, Mitsuo SHIMOHATA, Eiichiro SUMITA	Statistical Machine Translation on Paraphrased Corpora	2002. 05. 29 ~2002. 05. 31	採録決定
2001音-発084	国際会議	ACL-02 (Association for Computational Linguistics 40th Anniversary Meeting)	Taro WATANABE, Eiichiro SUMITA	Phrase Decoding Algorithm for Statistical Machine Translation	2002. 07. 07 ~2002. 07. 12	査読不可
2001音-発085	国際会議	ACL-02 (Association for Computational Linguistics 40th Anniversary Meeting)	Mitsuo SHIMOHATA, Eiichiro SUMITA	Paraphrasing for Normalization of Expressions Based on a Parallel Corpus	2002. 07. 07 ~2002. 07. 12	査読不可
2001音-発086	誌上	情報処理学会論文誌	下畑 光夫, 隅田 英一郎	形態素情報の体系間変換		投稿中
2001音-発087	国内大会	言語処理学会第8回年次大 会	中嶋 秀治、山本 博史、渡辺 太郎	機械翻訳によって生成された追 加テキストを使った言語モデル の適応	2002. 03. 16 ~2002. 03. 20	発表済み
2001音-発088	国内大会	言語処理学会第8回年次大 会	山本 博史、菊井 玄一郎	教師なし学習による文の分割	2002. 03. 16 ~2002. 03. 20	発表済み
2001音-発089	誌上	電子情報通信学会論文誌	伊田 政樹、中村 哲	雑音 GMM の適応化と SN 比別マル チパスモデルを用いた HMM 合成 による高速な雑音環境適応化		投稿中

決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発090	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Michael PAUL, Eiichiro SUMITA	Corpus-based Generation of Numeral Classified Using Phrase Alignment	2002.08.24 ~2002.09.01	投稿中
2001音-発091	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Yves LEPAGE	Easy-to-Build and Easy-to-Use Light Portable Machine Translation Engines	2002.08.24 ~2002.09.01	投稿中
2001音-発092	国内研究会	情報処理学会第148回自然言語処理研究会	西村 仁志、坂本 仁	コーパスから自動抽出した表現パターンを用いる日本語文生成	2002.03.04 ~2002.03.05	発表済み
2001音-発093	国内大会	日本認知科学会第19回大会	金城 由美子、森本 郁代 (大阪大学大学院)	非言語的情報にもとづく同義性について	2002.06.14 ~2002.06.16	投稿中
2001音-発094	国内大会	言語処理学会第8回年次大会	宮木 衛 (豊橋技術科学大学)、増山 繁 (豊橋技術科学大学)、山本 和英	2名詞による連体修飾語の換言可能性に関する考察	2002.03.16 ~2002.03.20	発表済み
2001音-発095	国内大会	言語処理学会第8回年次大会	山本 和英	テキストからの語彙的換言知識の獲得	2002.03.16 ~2002.03.20	発表済み
2001音-発096	国内大会	言語処理学会第8回年次大会	山本 和英	換言と言語変換の協調による機械翻訳モデル	2002.03.16 ~2002.03.20	発表済み
2001音-発097	国内大会	言語処理学会第8回年次大会	張 玉潔、山本 和英、坂本 仁	換言コーパスを利用した中国語換言処理	2002.03.16 ~2002.03.20	発表済み
2001音-発098	国内大会	言語処理学会第8回年次大会	Yves LEPAGE, Eddy TAILLEFER, 白井 諭	類推に基づく構文解析の変定数の影響評価	2002.03.16 ~2002.03.20	発表済み
2001音-発099	国内大会	言語処理学会第8回年次大会	大竹 清敬、山本 和英	形容詞述語文の換言事例の分析	2002.03.16 ~2002.03.20	発表済み
2001音-発100	国内大会	言語処理学会第8回年次大会	Nicolas AUCLERC, Nick CABELL, Yves LEPAGE	Open Distributed Architecture for NLP Systems Using Services Described in XML	2002.03.16 ~2002.03.20	発表済み

決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発101	国内大会	言語処理学会第8回年次大会	鷹尾 和享、今村 賢治、柏岡 秀紀	英和/和英辞典を用いた英語換言語の抽出	2002.03.16 ~2002.03.20	発表済み
2001音-発102	国内大会	言語処理学会第8回年次大会	伊藤 山彦、谷田 泰郎 (TIS)、柏岡 秀紀、田中 英輝	決定木学習を利用した講演文からのトピックフレーズの抽出	2002.03.16 ~2002.03.20	発表済み
2001音-発103	国内研究会	第2回 話し言葉の科学と工学ワークショップ	奥田 浩三、河原 達也、中村 哲	発話速度の補正を用いた音響モデルの構築	2002.02.28 ~2002.03.01	発表済み
2001音-発104	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Yujie ZHANG, Kazuhide YAMAMOTO	Paraphrasing of Chinese Utterances	2002.08.24 ~2002.09.01	投稿中
2001音-発105	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Mitsuo SHIMOHATA, Eiichiro SUMITA	Identifying Synonymous Expressions from a Bilingual Corpus for Example-based Machine Translation	2002.08.24 ~2002.09.01	投稿中
2001音-発106	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Yasuhiro AKIBA, Taro WATANABE, Eiichiro SUMITA	Using Language and Translation Models to Select the Best among Outputs from Multiple MT Systems	2002.08.24 ~2002.09.01	投稿中
2001音-発107	その他	ATR Journal	山本 誠一	実世界で利用可能な音声翻訳技術の実現に向けて	2002.02.08	発表済み
2001音-発108	国内大会	言語処理学会第8回年次大会	金城 由美子、柏岡 秀紀、田中 英輝	同格的な節を補部とする名詞について	2002.03.16 ~2002.03.20	発表済み
2001音-発109	国内大会	日本音響学会 2002 年春季研究発表会	正木 芽衣子 (奈良先端大)、柏岡 秀紀、Nick CAMPBELL (HIP/CREST)	物語朗読における異なる話速と発話スタイル間の発話時間長制御について	2002.03.18 ~2002.03.20	発表済み
2001音-発110	国内大会	日本音響学会 2002 年春季研究発表会	秦 早穂子 (奈良先端大)、柏岡 秀紀、Nick CAMPBELL (HIP/CREST)	同一表記である疑問文・平叙文の韻律パターンによる判別	2002.03.18 ~2002.03.20	発表済み



決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発111	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Hideharu NAKAJIMA, Hirofumi YAMAMOTO, Taro WATANABE	Language Model Adaptation with Additional Text Generated by Machine Translation	2002.08.24 ~2002.09.01	投稿中
2001音-発112	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Taro WATANABE, Eiichiro SUMITA	Bidirectional Decoding for Statistical Machine Translation	2002.08.24 ~2002.09.01	投稿中
2001音-発113	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Kiyohiro OHTAKE, Kazuhide YAMAMOTO	Applicability Analysis of Corpus-derived paraphrases toward Example-based Paraphrasing	2002.08.24 ~2002.09.01	投稿中
2001音-発114	国内大会	2002年度第16回人工知能学会全国大会	速見 悟 (産業技術総合研究所)、板橋 秀一 (筑波大学)、小林 哲則 (早稲田大学)、竹澤 寿幸	日本音響学会研究用連続音声データベース	2002.05.29 ~2002.05.31	採録決定
2001音-発115	国際会議	21st International Unicode Conference	Yves LEPAGE	Contribution and Limits of the Use of Unicode in Approximate Pattern-matching	2002.05.14 ~2002.05.17	採録決定
2001音-発116	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Kazuhide YAMAMOTO	Acquisition of Lexical Paraphrases from Texts	2002.08.24 ~2002.09.01	投稿中
2001音-発117	国際会議	COLING2002 (The 19th International Conference on Computational Linguistics)	Kazuhide YAMAMOTO	Machine Translation by Interaction between Paraphraser and Transfer	2002.08.24 ~2002.09.01	投稿中
2001音-発118	国際会議	ICME2002	Kazumasa MURAI, Satoshi NAKAMURA	Real Time Face Detection for Multimodal Speech Recognition	2002.08.26 ~2002.08.29	投稿中

決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発121	国際会議	ICME2002	Satoshi NAKAMURA, Kazuo HIYANE (Mitsubishi Res. Institute), Futoshi ASANO (National Ins. of Advanced Industrial Science and Tech.), Yutaka KANEDA (Tokyo Denki Univ.), Takeshi YAMADA (Tsukuba Univ.), Takanobu NISHIURA, Tetsunori KOBAYASHI (Waseda Univ.), Shiro ISE (Kyoto Univ.)	Design and Collection of Acoustic Sound Data for Hands-Free Speech Recognition and Sound Scene Understanding	2002. 08. 26 ~2002. 08. 29	投稿中
2001音-発122	誌上	Speech Communication	Kaisheng YAO, Kuldip K. PALIWAL, Satoshi NAKAMURA	Noise Adaptive Speech Recognition Based on Sequential Noise Parameter Estimation		投稿中
2001音-発123	国際会議	TMI2002 (The 9th Conference on Theoretical and Methodological Issues in Machine Translation)	Eiichiro SUMITA, Kenji IMAMURA	Tutorial on Example-Based Machine Translation	2002. 03. 13 ~2002. 03. 17	発表済み
2001音-発124	国内大会	国語学会 2002 年度春季大会	丸山 岳彦	文の階層構造とテモ節の分布	2002. 5. 19	投稿中
2001音-発125	誌上	自然言語処理	吉田 辰巳 (豊橋技科大学)、大竹 清敬、山本 和英	サポートベクトルマシンを用いた中国語解析実験		投稿中
2001音-発126	誌上	電子情報通信学会論文誌	戸田 智基、河井 恒、津崎 実、鹿野 清宏	素片接続型日本語テキスト音声合成における音素単位とダイフォニ単位に基づく素片選択		投稿中

決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発127	国内大会	情報処理学会第64回全国大会	安田 圭志、菅谷 史昭、竹澤 寿幸、山本 誠一、柳田 益造 (同志社大学)	決定木学習を用いた翻訳システム自動選択手法の検討	2002.03.12 ~2002.03.14	発表済み
2001音-発128	国際会議	LREC 2002 (Third International Conference on Language Resources and Evaluation)	Andrew FINCH, Ezra BLACK	Beyond Tag Trigrams : New Local Features for Tagging	2002.05.29 ~2002.05.31	採録決定
2001音-発129	誌上	電子情報通信学会論文誌	中嶋 秀治、山本 博史、渡辺 太郎	機械翻訳によって生成された追加テキストを使った統計的言語モデルの適応		投稿中
2001音-発130	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Hisashi KAWAI, Minoru TSUZAKI	Acoustic Measures vs. Phonetic Features as Predictors of Audible Discontinuity in Concatenative Speech Synthesis	2002.09.16 ~2002.09.20	投稿中
2001音-発131	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Tomoki TODA, Hisashi KAWAI, Minoru TSUZAKI, Kiyohiro SHIKANO (NAIST)	A Segment Selection Algorithm for Japanese Concatenative Speech Synthesis Allowing Concatenation at Vowel Center	2002.09.16 ~2002.09.20	投稿中
2001音-発132	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Mitsunori MIZUMACHI, Satoshi NAKAMURA	The 2ch Hybrid Subtractive Beamformer Applied to Line Sound Sources	2002.09.16 ~2002.09.20	投稿中
2001音-発133	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Masaki IDA, Satoshi NAKAMURA	AURORA2 Evaluation of HMM Composition-based Rapid Model Adaptation Using a Priori Noise GMM	2002.09.16 ~2002.09.20	投稿中

決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発134	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Hiroaki KOKUBO, Teruaki HAYASHI (Mediax), Hirofumi YAMAMOTO, Genichiro KIKUI	Curtialment of Word Hypotheses Using Attributes of Successive Hypotheses	2002. 09. 16 ~2002. 09. 20	投稿中
2001音-発135	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Hirofumi YAMAMOTO, Konstantin MARKOV, Kozo OKUDA	Combination of Multiple Recognizer Outputs Based on Maximum Word Acceprance Rate Path Selection	2002. 09. 16 ~2002. 09. 20	投稿中
2001音-発136	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Jin-song ZHANG, Satoshi NAKAMURA	Modeling Varying Pauses to Develop Robust Acoustic Models for Recognizing Noisy Conversational Speech	2002. 09. 16 ~2002. 09. 20	投稿中
2001音-発137	誌上	電子情報通信学会論文誌	奥田 浩三、松井 知子、中村 哲	誤認識時の言い直し発話における発話スタイルの変動に頑健な音響モデル構築法		投稿中
2001音-発138	誌上	電子情報通信学会論文誌	奥田 浩三、河原 達也、中村 哲	尤度基準による分析周期・窓長の自動選択手法を用いた発話速度の補正と音響モデルの構築		投稿中
2001音-発139	国際会議	Joint International Conference of SNLP-Oriental COCOSDA 2002	Shigehiko ONISHI, Hirofumi YAMAMOTO, Genichiro KIKUI, Yoshinori SAGISAKA	A Statistical Word Model Using Word-class Specific Constraints for Handling Out-of-vocabulary Words in Speech Recognition	2002. 05. 09 ~2002. 05. 11	採録決定
2001音-発140	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Kozo OKUDA, Tatsuya KAWAHARA, Satoshi NAKAMURA	Speaking Rate Compensation Based on Likelihood Criterion in Acoustic Model Training and Decoding	2002. 09. 16 ~2002. 09. 20	投稿中

決裁番号	発表方法	雑誌名・国際会議名	発表者	タイトル	発表(予定)日	状況
2001音-発141	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Katsushi ASAMI, Toshiyuki TAKEZAWA, Genichiro KIKUI	Topic Detection of an Utterance for Speech Dialogue Processing	2002.09.16 ~2002.09.20	投稿中
2001音-発142	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Kaisheng YAO, Kuldip K. PALIWAL, Satoshi NAKAMURA	Noise Adaptive Speech Recognition with Acoustic Models Trained from Noisy Speech Evaluated on AURORA-2 Database	2002.09.16 ~2002.09.20	投稿中
2001音-発143	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Jinfu NI, Hisashi KAWAI	Design of a Mandarin Sentence Set for Corpus-based Speech Synthesis by Use of a Multi-tier Algorithm Taking Account of the Varied Prosodic and Spectral Characteristics	2002.09.16 ~2002.09.20	投稿中
2001音-発144	国際会議	ICSLP 2002 (7th International Conference on Spoken Language Processing)	Norbert BINDER, Rainer GRUHN, Satoshi NAKAMURA	Phoneme Lattice Processing for Robust Recognition of Non-Native Speech	2002.09.16 ~2002.09.20	投稿中
2001音-発145	国内大会	日本音響学会 2002年春季研究発表会	沖本 純幸、山本 博史、隅田 英一郎、菊井 玄一郎	大規模用例を用いた音声認識誤り訂正の検討	2002.03.18 ~2002.03.20	発表済み
2002音-発020	国内大会	電子情報通信学会音声研究会	陸 金林、河井 恒	中国語音節の音素と声調の代替接続による自然性劣化の知覚的評価	2002.03.28 ~2002.03.29	発表済み
2002音-発021	国内大会	日本音響学会 2002年春季研究発表会	陸 金林、河井 恒	中国語音声合成のための音声セグメンテーションについて	2002.03.18 ~2002.03.20	発表済み
2002音-発022	国内大会	日本音響学会 2002年春季研究発表会	陸 金林、河井 恒	波形素片接続時の中国語音節環境代替による自然性劣化の知覚的評価	2002.03.18 ~2002.03.20	発表済み

