

# 総合報告

## Web時代の音声・言語技術

Speech and Natural Language Processing in the Web Information Era

中村 哲 磯谷亮輔 乾 健太郎 柏岡秀紀 河井 恒  
河原達也 木俣 豊 黒橋禎夫 隅田英一郎 関根 聡  
鳥澤健太郎 堀 智織 松田繁樹

### abstract

本稿では、現在、独立行政法人情報通信研究機構（NICT）において産学官で進められている MASTAR プロジェクトにおける音声・言語処理について紹介するとともに、日本及び海外、特に米国の動向について紹介する。音声・言語処理は近年飛躍的な進歩を遂げているが、この理由の一つとして挙げられるのが、大量のデータ（コーパス）を収集し、用例や統計モデルと機械学習により、自動的に処理系を構築するコーパスベース技術が確立されたことである。Web上の情報を利用することで、Web以前には不可能であった様々な音声言語処理が可能になりつつある。

キーワード：自然言語処理，音声処理，Web，機械翻訳，情報アクセス

### 1. はじめに

1970年代の後半に、雑音のある通信路モデルに基づき、音声の時間的、特徴量的な揺らぎを統計的にモデル化し、モデルパラメータの推定と、認識を統一的な枠組みで計算できる隠れマルコフモデル（HMM: Hidden Markov Model）の原型が提案された。対象言語の音声コーパス（本稿ではメタ情報付き音声・テキストデータをコーパスと呼ぶ）とテキストコーパスを大量に収集す

ることでモデルパラメータの推定を行い入力音声の認識を行う技術である。その後、多くの改良を経て最近では不特定話者の連続音声認識が、かなり高いレベルまで到達している。この統計的なモデリングは、言語翻訳でも1990年代に入り本格化し、統計翻訳という名前で一つの大きな流れを構成するに至っている。統計翻訳では、同じ意味を持つ異なる言語の文対（対訳コーパスと呼ぶ）と、対象言語のテキストコーパスを大規模に収集し、これらから統計翻訳のモデルパラメータの推定を行

中村 哲 正員 独立行政法人情報通信研究機構知識創成コミュニケーション研究センター  
E-mail satoshi.nakamura@nict.go.jp  
磯谷亮輔 正員 独立行政法人情報通信研究機構知識創成コミュニケーション研究センター  
E-mail ryosuke.isotani@nict.go.jp  
乾 健太郎 東北大学大学院情報科学研究科システム情報学専攻  
E-mail inui@ecei.tohoku.ac.jp  
柏岡秀紀 独立行政法人情報通信研究機構総合企画部企画戦略室  
E-mail hideki.kashioka@nict.go.jp  
河井 恒 正員 シニア会員 独立行政法人情報通信研究機構知識創成コミュニケーション研究センター  
河原達也 正員 京都大学学術情報メディアセンター  
木俣 豊 正員 独立行政法人情報通信研究機構知識創成コミュニケーション研究センター  
E-mail kidawara@nict.go.jp  
黒橋禎夫 正員 京都大学大学院情報科学研究科知能情報学専攻  
E-mail kuro@i.kyoto-u.ac.jp  
隅田英一郎 正員 独立行政法人情報通信研究機構知識創成コミュニケーション研究センター  
E-mail eiichiro.sumita@nict.go.jp  
関根 聡 ニューヨーク大学コンピューター・サイエンス学科  
E-mail sekine@cs.nyu.edu

鳥澤健太郎 独立行政法人情報通信研究機構知識創成コミュニケーション研究センター  
堀 智織 正員 独立行政法人情報通信研究機構知識創成コミュニケーション研究センター  
E-mail chiori.hori@nict.go.jp  
松田繁樹 正員 独立行政法人情報通信研究機構知識創成コミュニケーション研究センター  
Satoshi NAKAMURA, Ryosuke ISOTANI, Yutaka KIDAWARA, Eiichiro SUMITA, Chiori HORI, Shigeki MATSUDA, Members, Hisashi KAWAI, Senior Member, Kentaro TORISAWA, Nonmember (Knowledge Creating Communication Research Center, National Institute of Information and Communications Technology, Kyoto, 619-0289 Japan), Kentaro INUI, Nonmember (Graduate School of Information Sciences, Tohoku University, Sendai-shi, 980-8579 Japan), Hideki KASHIOKA, Nonmember (Strategic Planning Department, National Institute of Information and Communications Technology, Koganei-shi, 184-8795 Japan), Tatsuya KAWAHARA, Member (Academic Center for Computing and Media Studies, Kyoto University, Kyoto-shi, 606-8501 Japan), Sadao KUROHASHI, Member (Graduate School of Informatics, Kyoto University, Kyoto-shi, 606-8501 Japan), and Satoshi SEKINE, Nonmember (Computer Science Department, New York University, New York, 10003 U.S.A.).  
電子情報通信学会誌 Vol.94 No.6 pp.502-517 2011年6月  
©電子情報通信学会 2011

う。これまで必要だったルールベース翻訳のルールの作成が不要になり、ドメイン（翻訳の対象範囲）の拡大、翻訳言語対の増加が容易になったことが大きな利点となっている。これらの二つの技術は、いずれもコーパスを収集し、そのコーパスからモデルパラメータを推定する点の特徴である。この際に問題となる三つの課題がある。一つ目は、統計モデルの推定の性質からパラメータ推定のためのコーパスは多いほどよいから、質の良い大規模なコーパスをいかに多く効率的に収集するかである。二つ目は対象となるアプリケーションのドメインに合致したモデルを構築するために、対象ドメインから学習コーパス、辞書をいかに効率的に多く収集するかである。最後に、三つ目として、対象ドメインだけでなく言語現象として日々変貌する言語現象をコーパス、辞書として、いかに捉えるかが課題である。

1991年に欧州原子核研究機構（CERN）のティム・バーナーズ＝リーがWWW（World Wide Web、以後Web）に関する情報、ブラウザ等を公開して以来、Webには多様で膨大な情報が集まるようになった。現在、Webには、全世界の全言語の情報（画像を含む）を加えるとゼタバイト（10の21乗）オーダーの量があるといわれている。この中には、多言語の広告のページ、情報案内のページ、情報提供のページ、E-Commerceのページ、更には、一般の利用者のブログ、最近ではTwitterなどの情報が含まれている。多様な利用者によって生成されるテキストには生きた言語が含まれ、日々変化していく実社会を射影した世界を構成している。このWebに蓄積、射影される言語現象は、課題1と3への対抗策となるといえる。

本稿ではWeb時代の音声・言語処理について、特に、現在、独立行政法人情報通信研究機構（NICT）において産学官で進められているMASTARプロジェクト（<http://www.mastar.jp>）における音声・言語処理について紹介するとともに、日本及び海外、特に米国の動向について紹介する。

MASTARプロジェクトは、統計的な音声・言語処理を行うための音声・言語コーパス、及び、それらを取り扱うツール群を、産業のインフラとして捉え、共通基盤として研究開発し、提供していくプロジェクトである。これまでの研究開発フェーズでデータを収集、プロトタイプシステムを構築し、その後、実用化に結び付けるといった従来の開発プロセスでなく、実際に使用される場面でのデータを直接収集し、それをパラメータ推定に用いることで、研究開発フェーズから実際の場面での性能向上を直接行う、課題2を解決する新しい研究開発プロセスであると考えられる。更に、Web上の情報を利用することで、世の中にある固有名詞の取り込み、多言語辞書の構築やコーパス収集、単語間の関係抽出、信頼性などの解析を行うことも可能になり、課題1

と3への対応も実現できる。

本プロジェクトの目標は、Web、ネットワークを利用し、産業界、社会とリンクした形で持続的に、音声・言語資源を蓄積、成長していく研究開発の仕組みを作ることである。具体的には、以下の四つの研究開発を行う。

- (1) 総合科学技術会議の社会還元加速プロジェクトの一つに選定されたネットワーク音声翻訳に関する技術開発、社会実験、社会還元を行う。
- (2) 産業界とマニュアルなどを対象としたWeb2.0型の機械翻訳サービスを行い、共通辞書、共通コーパスの蓄積、翻訳技術研究のポジティブ成長サイクルを確立する。
- (3) ユニバーサルコミュニケーションの一環として、あらゆる利用者へ情報を届けるための音声対話インタフェース技術の技術開発、社会実験、社会還元を行う。
- (4) 産業インフラとして世界的言語資源を構築し、社会に還元する。

上記の研究開発を進めるため産学官からも新しくメンバーを迎えて、2008年に5年間のプロジェクトを発足させた。

また、構築したコーパス、辞書、ツール等を、産業界、アカデミアに配信する仕組みとして、「高度言語情報融合フォーラム」(ALAGIN Forum: Advanced Language Information Forum. <http://www.alagin.jp> 会長: 辻井潤一東大教授, 副会長: 喜連川優東大教授及び榎並和雅 NICT 理事) を2009年3月に産学官体制で発足させた。本フォーラムでは、これまで述べたコーパス、辞書、ツールなどの言語資源を広く会員に配信、共有し、産学官の共同研究の場を提供することを目指している(正会員78社, 特別会員127名, 2011年1月現在)。

本稿では、以下、2., 3. にMASTARプロジェクトにおける言語処理研究及び音声処理研究の狙いとこれまでの成果について述べ、4. に日本におけるWebを用いた音声・言語処理の動向、5. に米国における言語処理研究の動向について紹介する。(中村 哲)

## 2. NICT MASTAR プロジェクトにおける言語処理研究

### 2.1 Web時代における言語処理研究の新展開

インターネットの普及が言語処理技術に新しい潮流をもたらしていることはいうまでもない。例えば筆者の一

人が学生として研究を始めた20年前であれば言語処理技術が華々しく実用化されるというのは絵空事ではなかった。しかしながら、現在では言語処理技術は様々な場所で使われており、今後更に実用化の舞台が増えることは想像に難くない。また、産業界の若い世代の熱気には隔世の感を抱かざるを得ない。

本章ではそのような状況を背景に発足したNICT MASTARプロジェクト、並びに、その兄弟プロジェクトともいえる情報信頼性プロジェクトの現況について述べる。より具体的には、自動翻訳技術、Web情報分析システムWISDOM、概念辞書の三つのプロジェクトに焦点を当てて解説する。

## 2.2 Web時代の自動翻訳技術

インターネットでの使用言語を調べると、上位10位までで84%のシェア（日本語は第4位でシェア7%）になる。日本語以外の9言語から日本語への自動翻訳システムが作れば、インターネット上の情報の84%が読めるようになり、日本人の情報の受信能力を10倍以上高くできる。逆もまたしかり。これらの10の言語は、文字、単語、文法など大きく異なるので、自動翻訳技術には個別言語に強く依存しないことが求められる。

### 2.2.1 コーパスベース翻訳技術

計算機の基本性能が格段に進歩したこと、言語データ（大量の文章の蓄積や辞書）が計算機上に集積されるようになったこと、音声認識で統計的モデル化が大成功したこと、などを受けて、自動翻訳の研究において、対訳コーパス（同じ意味の原文と訳文の対を集めたもの）から翻訳知識を自動構築するコーパスベース翻訳技術が興り、現在、主流となっている。コーパスベースの翻訳技術は言語固有知識の構築に関わる人手が不要なので、前記の多言語翻訳という課題を解く有力な手段となる。

### 2.2.2 対訳コーパス収集技術

コーパスベース翻訳のインパクトのある定説の一つが「量が質を決める」である。様々な実験から経験的に、「対訳コーパス量を増やせば翻訳品質が改善する」ことが分かっているので、対訳コーパスを効率的に収集することが重要になる。NICTは、日本語と外国語の対訳コーパス構築に注力しており、現時点で、日本語文とその対訳の対を単位として数えて総数2,700万という世界最大規模に達している。また、対訳コーパスは順次公開をしている（例えば、Wikipedia日英京都関連文書対訳コーパス、<http://alaginrc.nict.go.jp/WikiCorpus/>）。

対訳コーパスの構築には、Webクローリング（Webデータの自動収集）のようにコンピュータ中心のアプローチのほかに、外部機関との提携など、人中心のアプローチがある。前者の技術では、文書単位で対応する日

本語と外国語のデータから文対応を自動的に抽出する技術が非常に有用であり、新聞、科学技術論文、特許など様々な分野の対訳コーパスの構築に活用してきた。また、後者については、翻訳のホスティングサービス「みんなの翻訳」というユニークな試みを行っており以下で説明する。

「みんなの翻訳」サイト（<http://trans-aid.jp/>）は、2009年4月に公開された。2010年11月の時点で、1,445人の利用者があり、5,503件の翻訳文書が登録されている<sup>(1)</sup>。また、アムネスティインターナショナル日本などのNGO/NPO、英国リーズ大学などの教育機関が積極的に活用している。

「みんなの翻訳」の特徴は、①公開されている翻訳には、「一定の条件下で、二次的著作物を作成し、それを公開してもよい」というライセンスが付与されるため、適切な使用であれば、翻訳を第三者が利用できるということ、②東京大学で開発された高機能な翻訳支援エディタQReditを誰もが利用できること、③三省堂の協力による「グランドコンサイス英和辞典」をはじめとする多様なリソースが翻訳に利用できることである。

「みんなの翻訳」はボランティアの翻訳者を主たる利用者と想定している。ボランティアの翻訳者は、様々な文書を翻訳している。例えば、ソフトウェアのマニュアルやブログなどがある。マニュアルの日本語訳はユーザーにとっては大変有り難いものであるし、ブログの翻訳はマスメディアが注目しない場所や人々について光を当てるものといえる。このように、ボランティアの翻訳者は、世の中に貢献しているし、そのことが翻訳することの強い動機付けになっている。また、日本のボランティアの翻訳者は、現在、数千人程度であるが、高学歴の主婦や企業OBなど潜在的なボランティア翻訳者の数は、数十万人程度と推定される。これは非常に大きな翻訳のパワーになると期待できる。翻訳をしたい人が、簡単に翻訳ができる環境を提供すれば、現状よりも、もっと多くの人が翻訳をするようになり、同時に対訳コーパスも規模を拡大していくと考えている。

### 2.2.3 更に一つ上の技術

言語処理研究が進んでいない言語では高精度の形態素解析プログラムがない場合も多々あり、翻訳品質を目的関数にして文字から始めて翻訳に適した単位を決定するという新しい手法を開発した。

また、対訳コーパスが幾ら巨大でも全ての固有名詞を含むことは不可能である。固有名詞の翻字（奈良をNaraに翻訳するように音で翻訳する）を自動化することも重要であり、NICTは現在、世界最高レベルの性能を実現している。

NICTは、翻訳技術の多言語性だけを追求しているのではなく、日本語を中心にして高精度翻訳の実現を目的

としている。コーパスベース翻訳は対象言語によらず一定の性能を達成できるが、高精度を目的とする場合には、構文の相違に正面から取り組む必要があり、そのためのアルゴリズムも開発している。

### 2.3 Web 情報分析システム WISDOM

NICT 知識処理グループでは、膨大な数の Web ページの中から、ユーザが信頼性の高い情報を発見する支援を行う Web 情報分析システム WISDOM を開発している。WISDOM は、ユーザが入力した分析対象文に対して関連する Web ページを検索した上で、Web ページの外観の分析、発信者情報の抽出、分析対象文に関連する主要文と対立文の抽出、意見表現の抽出などを行い、それらの分析結果を集約する情報分析システム<sup>(2)</sup>である。以下にその概要について述べる。

#### (1) 情報分析基盤の構築

WISDOM の開発において、最初に着手したのは分析対象となる Web ページの収集と集積であった。分析対象とする Web ページを偏りのない情報源とするために、約 500 万/日の日本語 Web ページのクロールを毎日行い、6 億ページ規模の Web アーカイブを構築している。更に、その収集した Web ページに対してアダルトページや SPAM ページなどを排除した約 1.2 億ページを対象に標準フォーマットと呼ぶ XML 文書アーカイブを作成して、検索用のための Web ページのインデックスを生成する。このような処理を常時行い、常に最新の Web アーカイブになるように自動的に更新しつつ、更に標準フォーマットのアーカイブも更新する情報基盤を確立している。

#### (2) 主要対立表現の抽出

WISDOM では、京都大学黒橋研究室で開発された TSUBAKI によって検索された Web ページの中から上位 1,000 件を対象として次のように主要・対立表現を抽出する。各ページから分析課題に関する 15 文程度の重要文を選択し、複合名詞、括弧で囲まれた方言、述語項構造などを主要表現の候補として集約する。その後、表記揺れの吸収や同意語表現のマージ、部分全体関係にある表現のマージなどを段階的に行い、純度の高い主要表現を集約する。更に、対立表現の抽出においては述語項構造について述語が否定になっているものや、反義語に置き換わっているものがあれば抽出する。

#### (3) 評価表現の抽出

評価表現の抽出においては、まず最初に評価表現のタイプを評価・感情・メリット・出来事・採否・当為の六つに分類分けして、評価表現を付与したコーパスを生成した。そして、そのコーパス上で、文中の単語の形態素

情報や評価表現辞書に登録されている単語の極性情報を組成として SVM による機械学習で評価極性の判断を行い、その結果を評価表現として自動的に抽出している。

#### (4) 情報発信者の解析

情報の発信者については、政府関係や企業、非営利団体、個人、匿名などのサイト運営者の分類を、HTML の構造を利用して、発信者情報が記述されていると想定できる領域等から抽出している。

以上のような処理によって、WISDOM は入力された分析対象文に対して自動的に分析を行う。例として、ホメオパシー治療の例を図 1 に示す。ホメオパシー治療とは「治療の対象となる症状を起こす物質を極度に希釈して投与することによって、身体其自然治癒力を引き出すもの」とされている。WISDOM に分析対象文として「ホメオパシーは効果がある」と入力すると、「任意団体」からの情報は肯定意見が多く、「新聞社」からの情報は否定的な意見が多いことが分かる。また、学術会議などでは、「根拠なく、荒唐無稽である」との談話を発表していること等も分かり、ホメオパシーの効果の程は確かでない。分析結果は、発信者ごとに肯定・否定の割合が異なっている。その結果、利用者が効果を信じる場合には否定意見として述べられている内容をリスクとして理解しながら行動することが容易になる。このように WISDOM は、情報検索だけでは見付けられない情報の概要を、簡単に手に取るように明らかにできる。

### 2.4 概念辞書

近年、社会はますますグローバル化、複雑化し、様々な事象の間に普通の人々の想像もつかない重大な連関をもたらしている。この結果、例えば、小さな離島に関する領土問題が日本を代表する企業の主力製品の製造へと飛び火するなどという、いわば「風が吹けば桶屋がもうかる」風でありながら重要な出来事が頻繁に起きつつある。

こうしたかけ離れた事象間の連関を認識し適切に意思決定しようとしたとき、Web は、その意思決定のベースとなる情報源の第 1 候補であろう。ところが、インターネットへの入り口として広く使われている検索エンジンが提示するのは、ユーザが高々数十ページしか読めない一群の Web ページにすぎない。そのような Web ページに、上述したような重大でありながら意外な連関が現れると期待するのには無理がある。

NICT の言語基盤グループで開発している概念辞書は、言語処理の基盤である辞書の巨大なものを言語処理技術の開発者、研究者に提供する、という研究開発上の必然性と、上述したような情報アクセスへの問題意識があって開発を開始した。基本的には億単位の Web ページから自動的に抽出された、単語間に成立する多種多様



図1 WISDOMの分析画面例(レポート画面)



図2 概念辞書のブラウジング 右下に洗濯機のトラブルとしてアトピーが表示されている。

な意味的関係の集積であり、現在 250 万語をカバーしている。抽出された関係は、意味ネットワークとして統合され、Web ブラウザ上で、自動抽出手法が認識できた限りにおいて、単語間の意味的关系を列挙することが可能である。これにより、例えば「洗濯機」が引き起こす意外なトラブルとして「アトピー」があり、その意外な解決法として 10 円玉を洗濯機に入れておく、という主婦の知恵があるといった情報が容易に取得できるようになる(図2)<sup>(3)</sup>。

こうした概念辞書関連技術の成果の一部は既に ALAGIN フォーラム (<http://www.alagin.jp>) において、データ若しくはツール、Web サービスの形で商用利用

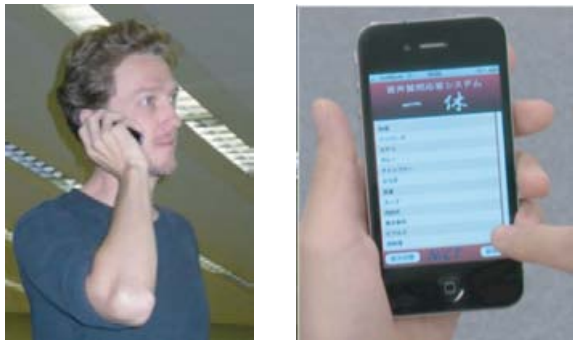
可、無償で公開されている。

また、概念辞書構築技術を拡張した音声質問応答システム「一休」も開発されている<sup>(4)</sup>。これは、スマートフォンに音声で入力された質問文、例えば「デフレを引き起こすのは何ですか?」を「X が Y を引き起こす」のような語彙統語パターンに変換し、その言い換え、例えば「X が Y を悪化させる」「X が Y の原因である」といったものを自動認識して、大量の Web 文書からパターンマッチングによって回答を抽出するものである。ここでの重要なポイントとして、パターンの多義性がある。一休では、「X の Y」のような多義なパターンも適切に取り扱える。こうした多義なパターンは往々にして高頻度であり、回答の抽出では大変重要である。例えば、「カキの食中毒」は「カキが食中毒を引き起こす」と同義であるが、この同義性は X, Y に相当する単語、カキ、食中毒が特定の意味クラス、例えば、食品、疾病に所属する限りにおいて妥当である。一休では Web 文書を用いた単語クラスタリングの結果を用いて、こうした同義性も利用して回答を抽出している。

また、いわゆる情報検索的な手法、足切りは一切行っておらず、例えば、デフレの原因として、「ネット」、「某有名大企業」といった意外な回答もパターンにマッチする限り提示される。「ネット」はネットオークションがデフレを悪化させる要因であるという主張であり、「某有名大企業」はその巨額の内部留保がデフレを悪化させているという主張である。企業についてはその後マスコミでも同主旨の記事が発表されている。こうした処理により、一休は「デフレの原因」から例えば「特定の特急内で食べるべき駅弁」「特定の病気の治療法」「特定の川で釣れる魚」「シュールレアリズムの提唱者」まで多種多様な質問に対する回答を列挙することができる。こうした回答は単語で表され、スマートフォンやブラウザの画面上で数百個の回答を検討することも十分に可能であり、通常の検索エンジンでは発見が不可能な、意外でありながら有用な情報を容易に発見することができる。また、これはスマートフォンをサーバに接続することで動作し、Web 6 億ページからの回答抽出をほぼリアルタイムで行うことができる。つまり、スマートフォンへの音声入力によって、通常の検索エンジンでは見つけられない有用な回答をいつでもどこでも入手できる。(図3。なお、デモビデオが <http://www2.nict.go.jp/x/x161/> から御覧頂ける。) なお、音声認識器は 50 万語をカバーするもので、NICT 音声コミュニケーショングループが開発したものを利用している。

## 2.5 まとめ

本章では、NICT で実施されている言語処理研究を解説した。これらの成果の一部は既に、ALAGIN や一般公開を介して、社会へと還元されており、また、今後更



(a) スマートフォンへの音声入力 (b) 回答のブラウジング

図3 音声質問応答システム「一休」の動作状況

に公開範囲を増やして更なる社会貢献を進めていく予定である。

(乾 健太郎, 木俣 豊, 黒橋禎夫,  
隅田英一郎, 鳥澤健太郎)

### 3. NICT MASTAR プロジェクトにおける 音声処理研究

NICTでは、MASTAR プロジェクトを通して、多言語音声翻訳システム及びその国際標準化、多言語対話応答システムの研究開発を行っている。本章では、各々の研究開発における現在の取組み、及び、今後の展開について述べる。

#### 3.1 多言語音声翻訳への取組み

音声翻訳は、入力音声をもとに言語へ翻訳し、その翻訳結果を音声合成で出力することにより異なる言語を話す利用者間の音声対話を実現する技術である。本節では、多言語音声翻訳の実現に重要な三つの要素技術（音声認識、翻訳、音声合成）について、NICTにおける取組み及び、本システムを用いた大規模実証実験について述べる。また最後に、音声翻訳システムの国際標準化への取組みについて述べる。

##### 3.1.1 多言語音声認識

現在主流の音声認識システムは、統計的音声認識手法を基礎としている。音声の中の個々の音素の振舞いや、単語の並びなどを統計モデルで表現し、様々な仮説の中から最も高い確率が与えられる単語列を認識結果として出力する。したがって、これらの統計モデルの精度が音声認識性能に大きく影響する。NICTでは、(株)国際電気通信基礎技術研究所（ATR）において長年にわたって収集してきた大規模音声コーパスに加え、インドネシア語やベトナム語などの多言語音声コーパスの収集を行い、多言語音声認識システムの研究開発を行っている。

3.1.4で述べる大規模実証実験の一環として実サービ

表1 音声コーパスサイズ

言語	話者数	文章数	発話時間	タスク
日本語	4,500	226,673	387.6	旅行会話
英語	930	236,737	257.6	旅行会話
中国語	540	213,352	251.0	旅行会話
インドネシア語	400	84,000	79.5	ニュース音声
ベトナム語	30	23,424	40.5	ラジオ音声

ス中の音声翻訳アプリケーション VoiceTra<sup>(5)</sup>は、日本語、英語、中国語、インドネシア語、ベトナム語の5言語について音声入力をサポートしている。これら五つの言語の音声認識に用いた音響モデルは、表1に示す音声コーパスにより学習した。言語モデルの学習には、日本語、英語は100万文、中国語は50万文、インドネシア語、ベトナム語は16万文の旅行会話文を用いた。

更に、全国5地方での実証実験期間中に収集された日本語約6万文、英語約1万7,000文、中国語約1万5,000文について人手による書き起こしを行い、音響、言語モデルの両方について学習を行った。また、VoiceTra等で収集された書き起こしのない数百万文規模の音声データに対して、音声認識結果中の単語や文の信頼度が高い音声区間を用いて音響、言語モデルを再推定する、教師なし学習を行うことにより、音声認識性能の改善を図っている。

#### 3.1.2 多言語翻訳

多言語音声翻訳システムの翻訳部では、全言語対に共通のシステム<sup>(6)</sup>を用いている。標準的なフレーズベース統計翻訳を基本として、幾つかの実用レベルの機能追加（固有名詞対訳の登録機能、翻字等）がなされており、旅行用の多言語対訳コーパスをモデル学習に用いている。図4は、旅行会話のテストデータ500文を用い、20の外国語から日本語への翻訳の評価実験を行った結果である。図に示すように、NICTの翻訳システム（黒）は、広く利用されている多言語ソフトウェア（白）よりも高い翻訳率が得られることが分かる。

#### 3.1.3 多言語音声合成

##### (1) 韓国語音声合成

音声合成を多言語化する活動の一環として、従来の日英中3言語に加えて、韓国語の音声合成システムを開発した。音声合成方式はHMM音声合成であり、音声コーパスは女性話者1名が発声した約10時間の朗読音声である。

##### (2) ボイスフォントの自動生成

ユーザが自分に似た声で翻訳結果を出力したいというニーズに対応するため、Web上に大量に存在する音声

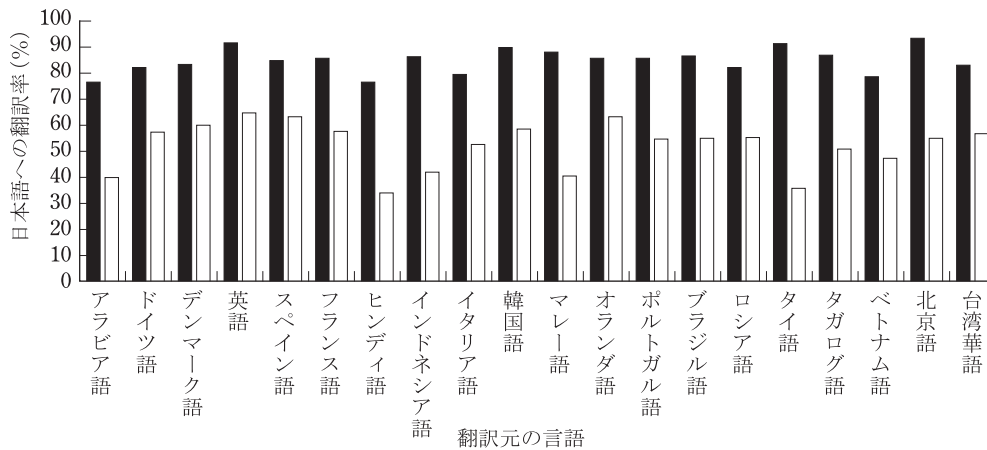


図4 翻訳率の比較 広く利用されているソフト（白）とNICTのソフトウェア（黒）と比較。縦軸が日本語への翻訳率、横軸が翻訳元の言語。

コンテンツをクロールし、音質の高い声質変換が可能な平均声HMMを自動的に学習する手法の研究を進めた。テキストが未知の音声に対して大語彙音声認識を適用し、誤りを含む音素列を参照してHMMの教師なし学習を行う場合、音素正解率が80%程度以上であれば、良好な音質の合成音を得られることが明らかになった<sup>(7)</sup>。

### (3) 音源モデルの高精度化

HMM音声合成の音質向上を目指して、音源モデル高精度化の研究を進めた。有声及び無声音源をそれぞれIIR及びFIRフィルタでモデル化し、音源及び声道モデルのパラメータを同時に学習する手法を開発し、ボコーダ型合成音特有のブザー感(buzziness)を軽減した<sup>(8)</sup>。

#### 3.1.4 大規模実証実験

NICTでは、本節で述べた多言語音声翻訳システムを用いた大規模実証実験を行っている。

##### (1) 全国5地方での実証実験

2009年12月から2010年3月にかけて、北海道から九州まで全国5地方の観光施設等約370か所において、音声翻訳の大規模な実証実験を行った<sup>(9)</sup>。本実験は総務省の委託事業「地域の観光振興に貢献する自動音声翻訳技術の実証実験」として実施されたもので、NICTは、受託した全ての事業者に対して音声翻訳技術を提供するとともに、実験システム構築、運用、データ分析等の面で全面的にサポートした。対象言語は日⇄英中韓の4言語(3言語対)で、実験期間中に計約20万件のアクセスが記録された。

地元特有の会話の翻訳精度を高めるため、各地方の固有名詞、固有表現(地元特有の言い回し)を事前に収集し、認識・翻訳・合成の辞書とモデルを各地方向けにカ

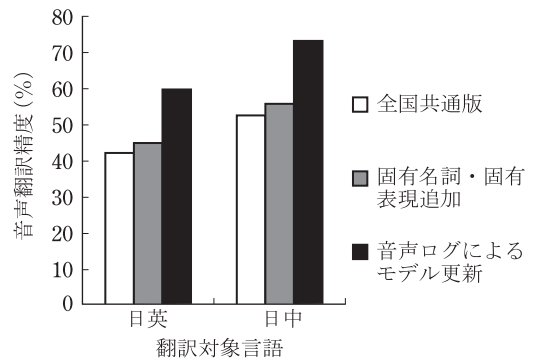


図5 全国5地方での音声翻訳実証実験におけるモデル更新の効果 縦軸は音声翻訳結果が許容範囲である発話の割合。事前の固有名詞・固有表現追加及び実験中の音声ログを利用した更新により性能改善が見られる。

スタマイズした。更に、実験期間中に蓄積された実利用音声データ(音声ログ)を書き起こし、対訳を作成して、それらを用いてモデルを更新し、システムにフィードバックした。音声ログから別途抽出した500発話をテストセットとしてモデル更新の効果を評価した結果を図5に示す。図の縦軸は、主観評価で音声翻訳結果が許容範囲内の発話の割合である。

利用者アンケートでは、伝達・理解度合いに関する質問に対し、半分程度以上伝わったあるいは理解できた割合は、日本語側で5割、外国語側で7割程度であった。北海道における事業者への利用意向調査では、本システムが商用化された場合の利用料金として、半数を超える回答者が月額1,000円以上と答えるなど、観光現場において自動音声翻訳へのニーズが確実に存在することを示唆する結果が得られた。

##### (2) 音声翻訳アプリ VoiceTra

音声翻訳技術の性能改善及び周知を目的として、ス

スマートフォン用の音声翻訳アプリケーション VoiceTra を開発し、2010年7月29日から同年末までの予定で無料公開した。VoiceTra は、21言語の双方向翻訳に対応しており、うち5言語については、音声による入出力が可能である。10月末時点での VoiceTra のダウンロード数及びアクセス数は、29万及び290万件となっている。

VoiceTra は、ネットワーク型システムを採用しており、ユーザが発話した音声と翻訳結果はログとして音声翻訳サーバに蓄積される。表2に、音声ログ100件を無作為抽出し、聴取により内容を分類した結果を示す。約半数が旅行会話的な発話の翻訳に利用されていることが分かる。

### 3.1.5 ネットワーク型音声翻訳技術の国際標準化への取組み

音声翻訳の実現には、翻訳対象となる言語の知識及び大規模な音声言語資源が必要であることから、個別の組織が、全言語対、全ドメインの音声翻訳を実現することは困難である。そこで NICT では、ネットワークを介して世界中に分散している音声認識、音声合成、翻訳モジュールを接続し、全ての言語対を音声翻訳できるネットワーク型音声翻訳の実現を目指している。本項ではネットワーク型音声翻訳の実現に必要なモジュール間通信プロトコルとデータフォーマットの国際標準化活動について述べる。

#### (1) アジアにおける国際標準化活動

これまで NICT は、アジアにおけるネットワーク型音声翻訳の先端研究を目的として、2006年に ATR (日本)、ETRI (韓国)、NECTEC (タイ)、BPPT (インドネシア)、CASIA (中国)、CDAC (インド) と共同でアジア音声翻訳先端研究コンソーシアム (A-STAR) を発足させ、2008年には IOIT (ベトナム)、I2R (シンガポール) が加盟して8か国の研究機関と共同研究を行ってきた。2007年にはアジア・太平洋電気通信標準化機関 (ASTAP) にて標準化活動を開始し、2009年7月、世界で初めてインターネットを介して、異なるアジア言語を話す複数話者間で、旅行対話を対象とした音声翻訳システムを用いて実時間音声対話に成功した。このネットワーク型音声翻訳技術をアジアにとどまらず世界

で用いられる標準化技術にすべく、標準化活動を AS-TAP から ITU-T に移行した。

#### (2) ITU-T における国際標準化活動

2009年10月ITU-T (International Telecommunication Union Telecommunication Standardization Sector) の SG16 (Multimedia coding, systems and applications), WP2 (Applications and systems), Q21 (Multimedia architecture)/Q22 (Multimedia applications and services) において、ネットワーク型音声翻訳技術の標準化を始動した。NICT がエディタとなり、①ネットワーク型音声翻訳のサービス要求条件と機能、及び、②アーキテクチャにおける要求条件の2本の勧告草案を作成し、2010年10月14日に勧告 F. 745 及び勧告 H. 625 として承認された。

### 3.2 多言語対話応答システムへの取組み

多言語対話応答システムは、利用者の発話した音声を認識する、①多言語音声認識、認識されたテキストを理解/応答文を生成する、②多言語対話制御、生成された応答文を音声として出力する、③多言語音声合成から構成されている。①多言語音声認識、及び③多言語音声合成については、3.1 に述べた多言語音声翻訳のシステムを応用し、多言語対話応答向けの音声対話コーパスを用いたシステムの研究開発を進めている。②多言語対話制御では、一問一答式の応答生成ではなく、対話の状況に応じた応答生成を目指し、制御機構の研究開発を進めている。そのため、人の対話行動を音声対話コーパスから学習する機能と事前に準備されている対話シナリオによる応答生成機能を統合可能な枠組みとして、統計的対話制御モデル<sup>(10)</sup>を利用したシステムの研究を進めている。

現在、多言語対話応答システムの研究開発を進めるにあたり、タスクは観光案内に限定し、音声対話コーパスの収集、プロトタイプシステムの構築、そのプロトタイプシステムを利用した実証実験を実施してきている。京都観光案内のエキスパートガイドと(模擬)観光旅行者との一日観光の計画立案を行う対話(1対話約30分)を複数の状況(対面、非対面、Wizard of Oz)で収録しコーパスとして整備している(対面対話100対話は ALAGIN フォーラムを通じて公開)。プロトタイプシステムとしては、観光旅行中に持ち歩いて利用できる携帯形システムと、観光案内所などで大画面での表示や旅行者の動作も考慮したマルチモーダル形システムの開発を行っている。プロトタイプシステムの対話制御の機構としては、WFST を利用しており、対話の状態に応じて応答文の生成が行われる。これらのプロトタイプシステムを利用し、京都において10日間の実証実験を実施した。実験後のアンケート調査では、多言語対話応答システムへの期待は高く、現状のプロトタイプシステムへの

表2 VoiceTra の発話の分類

分類	比率 (%)
無音	11
無効発話 (非音声など)	11
明確な旅行会話	9
旅行会話と解釈可能	42
旅行会話以外の内容	27



機能強化など課題を明確にすることができた。

### 3.3 まとめ

本章では、NICTのMASTARプロジェクトにおいて研究開発を行っている多言語音声翻訳システム、多言語対話応答システムについて述べた。今後は、多言語音声翻訳システムで現在対応している音声入力可能な言語数を5言語から21言語まで拡張する予定である。また、これまでの大規模実証実験で得られた大規模音声コーパスや知見を基に、よりドメインの広くかつ発話様式の異なる日常会話や、ビジネス会話での音声翻訳性能の改善を目指す予定である。音声翻訳の国際標準化では、ITU-Tで標準化されたネットワーク型音声翻訳技術の展開を目指し、アジア言語の拡大にとどまらず、欧米の研究機関とも連携して多言語音声ネットワークの輪を広げていく予定である。多言語対話応答システムでは、多言語への展開、タスク・ドメインの広がりに対する適応性を高め、多言語対話応答システムの性能改善を目指す予定である。

(中村 哲, 磯谷亮輔, 柏岡秀紀,  
河井 恒, 隅田英一郎, 堀 智織, 松田繁樹)

## 4. 日本における Web を用いた 音声・言語処理の動向

Webの出現は、自然言語処理や音声処理の研究の方法論に大きな変革をもたらした。Webは情報の膨大な集積であるが、音声・言語処理研究から見た場合に、巨大なテキストデータベース(コーパス)とみなすことも、巨大な知識ベースとみなすこともできる。1990年代に音声・言語処理において、コーパスに基づく方法論が主流になったが、新聞記事に代表されるように、表記

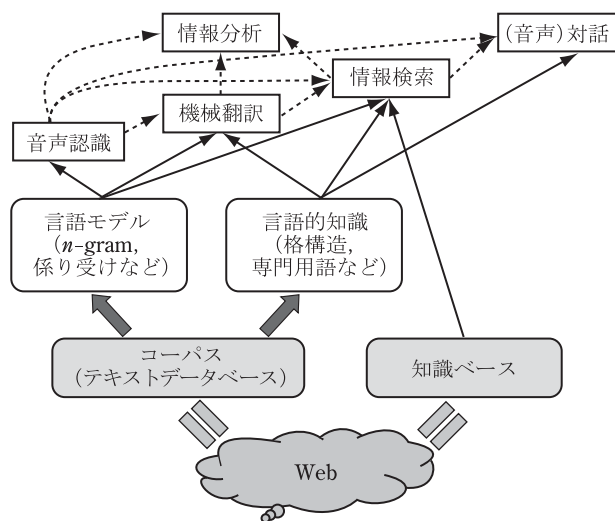


図6 Webを用いた音声・言語処理の概要

の上でも文法的にも整ったものをベースとしていた。更に、対話や検索に用いる知識源も、関係データベースやマニュアルなどの整ったものが主に用いられていた。また、格構造や翻訳のための知識なども、基本的には人手で記述されていた。

ところがこの10年ほどの間に、上記全てにおいてWebに基づく方式が研究開発され、確立されつつある。図6にその概要を示す。本稿では、我が国における研究を中心に、音声処理と言語処理のそれぞれについて紹介する。

### 4.1 Webを用いた言語処理研究

#### 4.1.1 Webの言語コーパスとしての整備

Web上の膨大なテキストが言語コーパスとして極めて高価値であることは誰もが認めるものであったが、日本においては著作権法の問題があり、これをコーパスとして整備することは難しい状況であった。しかし、その一部改正が2009年6月に可決・公布、2010年1月1日に施行され、情報検索サービス及び情報解析のためのWebページの複製が適法化されたことにより、今後、Webコーパス整備は大きく加速されていくものと考えられる<sup>(11)</sup>。

日本においてWebを大規模に収集する試みの先駆的なものは文科省e-Societyプロジェクト(2003~2007年)の一部として行われた東京大学喜連川研究室における活動で、その後も文科省プロジェクト「多メディアWeb解析基盤の構築及び社会分析ソフトウェアの開発」(2009~2012年)に引き継がれ、既に過去10年以上約160億ページのアーカイブが構築されており、今後このデータが研究コミュニティに広く利用可能となることが期待されている。また、国立国会図書館においても2010年から「インターネット資料収集保存事業」として日本国内の公的機関のWebページを中心とした収集、保存が行われており、今後収集対象の拡充が予定されている。

Webページ集合を言語コーパスとして利用するためには、単に収集するだけでなく、手間のかかる整備が必要となり、これを個々の研究者が行うとそれだけで疲弊してしまう。そこで、科研特定領域「情報爆発時代に向けた新しいIT基盤技術の研究」においては、支援班という枠組みを活用して、言語処理研究の基盤としての大規模コーパス整備を行い、開放型検索基盤TSUBAKIを構築した。ここでは、当初は東京大学で開発されたクローラで2007年に収集した約1億日本語Webページを対象として、文区切り判定、形態素・構文解析、単語及び係り受けインデクシングなどの処理を行い、API等によって検索結果(ヒット数、スニペット等)及び元のページを無制限に取得可能とした<sup>(12)</sup>。

TSUBAKIは、4.2.1で述べる音声認識用言語モデル

の構築を含め、様々な用途で言語コーパス、検索基盤として活用されている。

一方、Web上のバーティカルなコンテンツを学術利用のために提供する試みも始まっている。ヤフー株式会社では国立情報学研究所（NII）と協力して質問応答テキスト「Yahoo!知恵袋」コーパスの公開を進めている。最近では楽天株式会社から旅行施設のレビューデータ等がNII, ALAGIN フォーラムを通して公開されている。

#### 4.1.2 Webからの知識獲得とそれに基づく言語処理の高度化

Web言語コーパスの魅力は、何といてもその大きさにある。図7はこのインパクトを端的に示したものである。横軸は用いるコーパスサイズで、100万文からスタートして、およそ4倍ずつ、一番右は16億文（Web約1億ページに相当）である。縦軸はそのコーパスサイズで獲得できる知識の量及び解析精度であり、いずれもコーパス量の増加に応じて増加・改善されることが見て取れる。この図中で格フレームとは、「誰が何をどうした（例えば、学生が論文を書く）」のような述語項構造を、述語ごとに、どのような名詞が項となり得るかを記述したものであり、これはテキストの解析結果から確からしいものを抽出し、クラスタリングすることで全自動で獲得可能である<sup>(13)</sup>。

このような格フレームのカバレッジ、すなわちテスト文の述語項構造が格フレームとして既に学習されているかどうかは、当然のことながらコーパス量の増大とともに大きく改善する。そして、これに伴って、「彼はドイツ語も話す」が「彼がドイツ語を話す」と解釈する格解析や、文章中の省略を補う省略解析など、格フレームを参照して行う言語解析の精度も大幅に向上する<sup>(14)</sup>。更に、「景気が後退する」と「景気が冷え込む」のように項に依存する同義述語を出現文脈の類似度によって認識

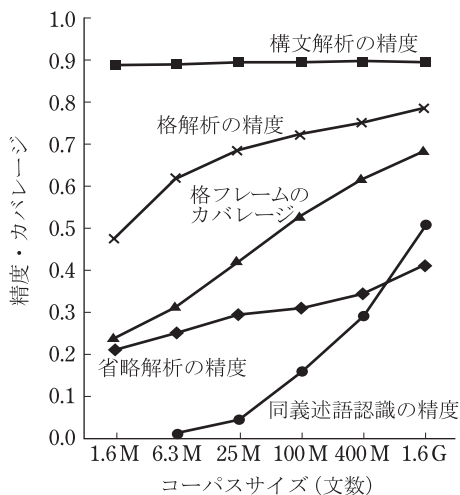


図7 コーパスサイズ増加に伴う言語処理の高度化

する精度も、同様に学習コーパスの増大に応じて大きく改善されている<sup>(15)</sup>。

更に、未知語やそのクラスの獲得<sup>(16)</sup>、肯定・否定表現の獲得<sup>(17)</sup>、Webに特徴的な企業名などの隠語の獲得<sup>(18)</sup>、Wikipediaからの様々な知識獲得<sup>(19)</sup>など、Webの大規模コーパスを用いて様々な知識を獲得する研究が活発に行われている。

#### 4.1.3 Webの情報検索・情報分析

Webは、巨大な知識ベースとして、既に我々の生活を支える社会基盤の一つであるといっている。しかしながら、情報洪水・情報爆発といわれるようにその情報量は膨大であり、更に違法・有害情報や情報信頼性の問題もある。言語処理が少しでも高度化すればその分だけ人々のWeb利活用が効率化・健全化されることから、言語処理にとってはまさに「相手にとって不足はない」という状況である。

Webを対象とする言語処理では、テキスト量の膨大さ、ドメインの広さ（ロングテイル）、口語調・新語などを含む表現の多様性など、チャレンジ的な要素が多い。しかし、クラスタ計算機やクラウドなどによって大規模並列計算も可能となり、前節で述べたように人手ではなくWebから自動獲得した知識がWebの解析を高度化するという自己循環的な言語処理の深化が始まりつつある。

Webの情報検索においては、企業のトップページを探すようなNavigational Queryはほとんど解決したといつてよいが、欲しい情報を探すInformational Queryでは現状のサーチエンジンで満足できることは少ない。調べたいこと、疑問に思うことは自然文で表現することが自然であるが、システム側では自然文クエリ及びWebテキストの言語的構造や同義表現を正確かつ柔軟に扱う必要がある。先に述べた検索基盤TSUBAKIでは、格フレームに基づく言語解析や、自動獲得した同義表現を用いてこの問題の解決を試みている。

更に、従来のサーチエンジンのようにクエリにマッチするページのリストを返すのではなく、様々な観点から情報を分析・集約するシステムの研究開発も活発に行われており、NICTの情報分析システムWISDOMはその一例である（2.参照）。また、4.1.1で述べた「多メディアWeb解析基盤」プロジェクトでは、時系列Webアーカイブを用いて、係り受け関係の時系列変化を分析することで話題追跡を行い、これを三次元可視化するシステムの開発を行っている<sup>(20)</sup>。

Webテキストにおける特徴的な言語解析としては、ブログなどのいわゆるCGMコンテンツ（CGM: Consumer Generated Media）における評判分析、すなわち肯定・否定意見の解析があり、企業のマーケティング戦略や政党・政策に対する民意調査などでも活用されつつ

ある。日本におけるブログ評判分析の先駆的研究としては BlogWatcher の開発がある<sup>(21)</sup>。ここでも、肯定・否定表現の語彙獲得と、個別のテキストの肯定・否定解析とが相補的に進展している。更に、ブログの解析において、テキストのテンス（時制）やモダリティ（様相：話者の判断や心的態度）の解析を正確に行うことにより、例えば「買った」のか「買うつもりだがまだ買っていない」かなどを区別し、個人の経験情報の意味的分類を行う「経験マイニング」などの試みもある<sup>(22)</sup>。

## 4.2 Web を用いた音声処理研究

### 4.2.1 Web を用いた音声認識用言語モデルの構築

1990 年代にディクテーションシステムに代表される大語彙連続音声認識システムの研究開発が世界的に行われた際には、主に新聞記事コーパスが利用された。これは、大規模に整った文章が容易に入手できることから自然の流れといえる。2000 年代になって、Web 上のテキストを利用するようになった。フリーの音声認識ソフトウェア Julius のディクテーションキットに含まれている言語モデルは、2003 年頃に Web テキスト（20 億形態素）から構築されたものである<sup>(23)</sup>。文の長さや文字種、文字  $n$ -gram などを用いて、日本語の文章らしくないもののみを除外していたが、これはいわば静的で汎用的なモデルである。

これに対して、検索エンジンを用いて特定のドメインに特化した言語モデルを構築する研究が行われるようになった。ドメインの代表的なキーワードを元に検索クエリを生成して、収集されたテキストを選別することにより言語モデルを学習する。特に、日本語では書き言葉と話し言葉で大きな違いがあるので、話し言葉らしい文をフィルタリングすることが重要であることが指摘されている<sup>(24)</sup>。

更に、個々の音声（発話や講演）に対して、一旦音声認識を行った上で、その認識仮説からキーワードを抽出し、関連する Web ページを検索して、動的に言語モデルを再構成し、再度音声認識をやり直す方法も提案されている<sup>(25)</sup>。これは、（音声認識システムにおける）未登録語への対応や、音声認識に基づいて情報検索を行う上で特に効果的と考えられる。未登録語や新しい単語の読み（発音）を自動推定するために、Web 上の知識源（Wikipedia や「はてなキーワード」）を利用することも試みられている。

### 4.2.2 Web を用いた音声対話システム

従来の音声対話システムは、バスの検索やフライトの検索のように、バックエンドとして関係データベースの検索を前提としたものが主流であった。この場合は、検索に必要な項目（フィールド）やその単語が限定できるので、その抽出を目標とすることで言語理解や対話制御

を容易に設計することができた（図 8）。

これに対して、一般的な文書を検索対象とする場合、すなわち情報検索をバックエンドとする場合、そのような方法論は採用できない。ただし、単純に音声入力 Web 検索するアプリケーション（音声検索：Voice Search）では、音声認識と検索エンジンを組み合わせるだけで事足りる。ドメインや対象文書を限定することで、Web を知識ベースとした対話システム、例えば観光ガイドや料理支援等を行うシステムが構成できる。

情報検索と質問応答を組み合わせた上で、プロアクティブに（ユーザから明示的な要求がなくても気を利かせて）情報提示を行うことを目指したのが、「情報コンシェルジェ」<sup>(26)</sup>である。京都に関する Wikipedia の文書に基づいて、京都案内を行うシステムを構築した。更に、日々動的に更新される Web 上のニュースを情報源として対話を行うシステムを研究開発している。効果的な対話を行うために、（スポーツや経済などの）ドメイン固有の情報抽出を導入している<sup>(27)</sup>。

### 4.2.3 Web サービスを用いた音声処理研究

以上は、図 6 に示したような、Web をテキストコーパスあるいは知識ベースとみなして行う研究の紹介であった。Web 上には種々の音声メディアがあり、その活用も考えられるが、現時点では一部の音声コンテンツに対して、字幕付与や検索インデックス作成のために、音声認識が試みられている程度である。

これに対して、Web をサービス提供の場とみなして、音声言語処理研究に活用する試みが行われている。その先駆的な例が Podcastle であり、Podcast の音声認識結果を不特定多数のユーザに修正してもらう枠組みを提供し、その結果に基づいて音声認識の言語モデルの改善を図っている<sup>(28)</sup>。

また 5. でも言及するが、近年米国では、Amazon 社の Mechanical Turk を用いて、多数の人に効率良く（有償だが安価で）音声・言語データの収集やアノテーションを行ってもらおう試みが多数行われているが、日本

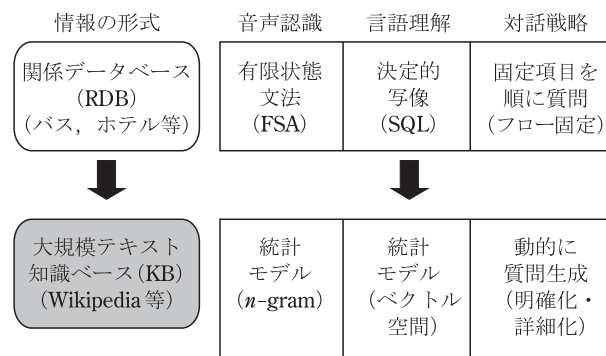


図 8 音声対話システムにおけるデータベース検索から汎用的情報検索へのシフト

においても楽天リサーチ社のモニタ誘因サービスと Web ベースの音声入力サーバを組み合わせて、大規模な音声データ収集を行った事例が報告されている<sup>(29)</sup>。

(河原達也, 黒橋禎夫)

## 5. 米国における言語処理研究の動向

### 5.1 米国における言語処理研究の背景

米国の大学の言語処理研究は、主に、米国国防省 (DARPA), 米国国立科学財団 (NSF), 企業との共同研究からの資金によって運営されている。特に、ペンシルバニア大学, 南カリフォルニア大学, カーネギーメロン大学のような、言語処理の研究を大規模に行っている大学は、国防省からの予算に大きく依存しており、活発に研究される分野は国防省の意向に大きく左右されることが多い。例えば、最近の機械翻訳の研究は、アラビア語, 中国語から英語への機械翻訳が中心であったり、情報抽出の対象となる情報はテロリストの活動などの分野に偏っていたりする。ただ、そのような大枠の制限はある中でも、具体的なテーマの設定には研究者が積極的に関わっており、国防省だけで研究内容が決まるというものではないことにも注意が必要である。国立科学財団が提供する研究資金は、日本の文科省の科研費に当たり、基礎研究に対して研究費を交付する機関である。日本と同じように主に大学の研究者から提案されたプロポーザルを、パネルと呼ばれる審査員が審査する形式であり、受理される割合は 7~25% くらいといわれている。技術的な先端性が評価されるが、それが世の中でどのように役に立つものになるかというのも重要視される指標である。また、大学と企業との共同研究も盛んであり、言語処理の分野では IBM, マイクロソフトやグーグルが競争資金等の形で研究資金を提供している。また、スタンフォード大学などでは、大学に在籍しながら自分で会社を立ち上げる形のベンチャー活動が活発であり、元気のいい大学のステータスとなっている場合が多い。このような様々な環境のもと、大学での研究が進められるわけであるが、Web 時代の言語処理研究の波はあらゆる場所に流れ込んできている。

以下では、特に米国の自然言語処理における組織横断的なプロジェクトである TAC, KDD のプロジェクトと、自然言語処理技術の共通基盤技術の研究動向、機械翻訳、情報アクセスに関わる研究動向を紹介する。

### 5.2 TAC

The Text Analysis Conference (TAC)<sup>(30)</sup> は米国国防省がスポンサーをし、米国国立標準技術研究所 (NIST: U.S. National Institute of Standards and Technology) が運営をしているプロジェクトである。プロジェクトは評価を中心とし、タスクを設計し決定する運営者とその評

価にシステムを参加させる参加者という構図になっている。プロジェクトの目的は、データと評価を提供し自然言語処理の研究を推進することであり、3 年目となる 2010 年は以下の三つのタスクが行われた。

- Knowledge Base Population (KBP)

本タスクでは、人名と組織名に関する属性情報を、テキストから抽出し、それを既存の不完全なデータベースを補うことを実質的なタスクとしている。既存のデータベースには Wikipedia の infobox を利用している。2010 年の KBP のタスクは、文章中の固有名をデータベースのエントリとリンクさせる Slot Linking のタスクと、データベースにある空欄の属性を埋めるタスクである Slot Filling のタスクがある。

- Recognizing Textual Entailment

あるテキストが他に与えられた言明の意味を成立させるかどうかという「含意」についての研究を行っているプロジェクトである。この評価型プロジェクトはヨーロッパの Pascal というプロジェクトにおいて 2004 年に開始されたが、2008 年からは TAC の枠組みの中で運営されている。2010 年は二つのタスクが行われた。Main Task では、一つの言明に対して、10 の文が与えられ、それぞれの文が言明を含意しているかどうかを判断するタスクである。これは、2009 年の TAC summarization のデータを用いており、複数のタスク間の連携が行われている。KBP Validation Pilot では、上記で説明した KBP の Slot Filling のタスクにおいて、属性の値になった要素と、それを抽出した文を与え、その文が、属性と要素間の関係を含意しているかを判断するタスクである。

- Summarization

要約に関する技術に取り組んでいるプロジェクトであり、2010 年には二つのタスクが行われた。Guided Summarization では、単なる単語頻度などの単純な技術ではなく、より深い自然言語処理の利用を促すことを目的に、あらかじめ決められたトピックに関する 10 個のドキュメントから、そのトピックに関する側面 (Aspect) を含む 100 単語の要約を生成するタスクである。Automatically Evaluating Summaries Of Peers (AESOP) は要約の自動評価の手法に関するタスクである。

### 5.3 KDD

Knowledge Discovery and Dissemination (KDD)<sup>(31)</sup> は The Intelligence Advanced Research Projects Activity がスポンサーになり、2011 年から始まる自然言語やデータベースの研究を含めたプロジェクトである。目的は、米国国防省系の情報アナリストが、新聞、内部及び

外部のレポート、人名地名他に関する様々なデータベースを駆使して情報収集を行っている作業の助けになるようなシステムを構築することになる。ここでの自然言語処理の一つの役割は、テキストからの固有表現抽出やその曖昧性解消等も含めた情報抽出にある。スポンサーの意向としては、なるべくすぐに使える技術を開発することにあり、他の国防省系（DARPA など）の大規模プロジェクトと同じく企業が中心となった大きなグループで参加する形になっている。

#### 5.4 共通基盤的技術

共通基盤技術としては、形態素論、構文解析、係り受け解析、意味論、言い換え表現、含意、推論など幅広い自然言語処理の分野の研究が行われている。一時期に比べて、与えられた何らかのタスクを最新の機械学習のアルゴリズムを使って精度の高いシステムを作るという研究は影を薄めてきているが、逆に言い換えや含意など意味論の方に移っても、シンプルな統計手法を使った近似的な研究もまま見られる。大規模コーパスを使って、頻度や共起を数えることにより、何らかの意味を抽出する研究も行われているが、国際会議 NAACL2010 で行われた、Web as Corpus Workshop が低調であったのを見ると、単なる大規模なコーパスを使っただけの研究には限界があることが認識されてきており、そこで抽出した知識とは何なのか、その精度を応用として使えるものにするためにはどうしたらいいかという研究に重心が移ってきている。

#### 5.5 Mechanical Turk

Web の到来や計算機能力の飛躍的な向上とともに、大規模なコーパスが集まるようになり、その中から言語的な知識、意味的な知識を抽出して、言語処理の役に立たせようという研究が盛んになっている。しかしながら、コーパスは文字の羅列でしかなく、抽出するための手掛かりは、単語の周りのコンテキストを使うか、単語自身の表記情報を使うしかない。そのような方法で知識を抽出する有名な方法には、文脈類似度（Distributional Similarity）や語彙構文規則（Lexico-syntactic Pattern）といったものがあるが、最近では、このような方法によって抽出できる情報の種類と質には限界があることが研究者の中で認識されるようになってきている。これは、究極的には、コーパスには何らの正解やグラウンディングがされていないことによるものであり、集めた知識は、「人間が見て初めて何か意味のあるように解釈できるようなもの」でしかなく、コンピュータがそのまま推論や演えき等の意味処理に使えるような情報ではない。そこで、人間による知識作成、または、自動的に抽出した知識の確認作業が必要となるわけである。4. で言及されたが、そのため、最近、非常に注目され利用

されているのが Amazon 社が提供している Mechanical Turk (MTurk)<sup>(32)</sup> というサービスである。

MTurk は、Web 上で実行できるシンプルな仕事のマーケットプレイスであり、何らかの仕事 (HITs: Human Intelligence Tasks) を提供する仕事提供者 (Requester) と、その仕事を遂行する作業員 (Workers) から成る。HITs は非常に簡単な html のような形式で書くことができ、CSV 形式の入出力によって仕事内容の提供と保存がされる。ここで驚異的なのは、その HITs の単価である。一般的には 3~5 分かかかるような HITs に対しての単価は 5~20 セント程度である。本章の筆者である関根が関与した WePS (Web People Search)<sup>(33)</sup> の正解作成は全部 MTurk で作成したが、数十万円程度のお金で、6 万 Web ページに対する正解データ作成を行った。1 判断当たり 5 人の作業員を当てているので、1 判断当たり 1~数円程度で作業を進められた計算になる。

この MTurk は、人手による知識作成に対して大きなインパクトを持った革命的な出来事であると考えている。当然、1 作業 1 円程度では、作業の質が問題になる。そのために 1 作業に対して 5 名の作業員を当てることや「荒らし (spammer)」への対処が必要となる。「荒らし」とは、作業の内容をきちんと行わずに高速に答えを返すような作業員を指すが、これに対しては、作業時間や作業量を監視する、中に正解がきちんと分かっている問題を埋め込んでおく、正解率が高い作業員にはボーナス (Reward) を出すといった方法で、かなりの「荒らし」の撲滅ができることが経験的に分かってきている。

WePS のみならず、様々な分野で正解作成や知識作成の作業を MTurk で行う言語処理研究が進められている。また、MTurk に関するワークショップが NAACL 2010 で行われており、100 ドルで何ができるか、というコンテストも行われた。

#### 5.6 ソーシャルメディア

Twitter や facebook など、いわゆるソーシャルメディアが台頭し、そこに書かれた文章やつぶやきに対する自然言語処理の研究が行われている。NAACL2010<sup>(34)</sup> では、#SocialMedia というワークショップが行われ、短い文章内での表現や議論の分析や、その中での質問応答、情報検索、情報抽出、リンク解析への言語利用等の研究が発表されている。また、2010 年に行われた Web People Search では、新たに Twitter 中にある企業名の曖昧性解消タスク（企業名としての apple と果物を区別するようなタスク）が取り上げられた<sup>(35)</sup>。

#### 5.7 機械翻訳技術

米国国防省 (DARPA) 予算で、音声翻訳を中心とし

た大規模プロジェクトである GALE<sup>(35)</sup>が走っている。2006～2010年に約300億円を支出して、英語、中国語、アラビア語について、非常に大規模なデータ（1万時間以上の音声コーパスや数百万文の対訳コーパス）を収集し、BBN社、IBM社、SRI（Stanford Research Institute）がリーダーとなって多数の機関で編成された三つのチームで、新聞、Web、ニュース放送等を対象として、高い性能目標のもとに競争的に実施されてきた。結果によって選択と集中が行われ、現在はBBNチームが中英翻訳、ア英翻訳を実施、IBMチームがア英翻訳を実施している。

毎年、性能は向上しているが、ア英翻訳が中英翻訳より大幅に性能が良いこと、ジャンルでは新聞が最も性能が良いことは変わらない。中英翻訳が難しいことから構文の利用が研究のホットな課題となっており、構文の利用がうまくいかずフレーズベースの統計翻訳だけで十分だとの主張がなされていた2005年前後とは様変わりしてきている。近年、言語学的構文、自動的に取得される形式的構文、句構造、依存構造、原言語側での利用、目的言語側での利用など、様々な変種が提案されている。しかも、実際にフレーズベースの統計翻訳を性能面で上回るようになってきた。

GALE以外にも、DARPA予算で、戦地で利用可能な携帯形の音声翻訳システムを目的としたTRANS-TAC、戦闘員が外国語資料を発見した現場で翻訳できるようにするため光文字認識と翻訳を組み合わせたMADCAT、戦場の特徴である騒々しい環境での音声認識と音声ベースの音源定位を補完する視線追跡技術を中心に置くRATSなど応用研究が盛んに行われた。NSF（National Science Foundation）助成の機械翻訳研究のトピックは、上記研究が掲げる目標を直接サポートするプロジェクトから、アラスカのネイティブ言語自動翻訳といったトピックまで多岐にわたる。

米国の機械翻訳のもう一つの特徴は、マイクロソフトやグーグルなど、企業による統計翻訳の研究開発が強い点である。マイクロソフトは自らの製品マニュアルや技術情報の英文を多言語化するコストを削減する強い社内ニーズがあり、グーグルは「世界中の情報を整理し、世界中の人々がアクセスできて使えるようにする」という使命のもと、多言語情報の英語化は必須であり、両者とも多言語化しやすい統計翻訳を中心に研究を進めている。両者は、Web上の無料翻訳、クラウド翻訳など様々な局面で競合している。ごく最近の新しいニュースとして、翻訳会社と統計翻訳の研究をしている企業の連携がある。ライオンブリッジ社とIBMの提携、SDL社によるLanguage Weaver社の合併は、翻訳会社の持つ対訳コーパスを最新の統計翻訳技術で有効活用しようという企てといえる。

## 5.8 情報アクセス技術

一般に以下のような四つの技術をまとめて「情報アクセス技術」と呼ぶ。

- ・ 情報検索：大規模な文書集合からユーザーが欲しい情報を含む文書を検索する
- ・ 情報抽出：世の中にある固有名、関係、イベントなどの情報を文書の中から構造化された形で抽出する
- ・ 要約：長い文章から重要な部分だけを抜き出す
- ・ 質問応答：自然言語で書かれた質問文に対し、その答えを的確に返す

このほかに、評判抽出や文書分類などの技術もあるが、それらは情報抽出や情報検索の部分的な技術、類似技術と捉えることができる。

このような情報アクセス技術は、米国では1980年代から国防省やNISTの大規模プロジェクトとして遂行されてきた。Tipster, TREC, TIDES, GALE, AQUA-INT, TAC, KDDなどがそれに当たる。固有表現や質問応答といったタスクや、情報抽出の関係抽出、イベント抽出といった概念はこの流れの中で生まれてきたものであり、情報アクセスだけに限らず、翻訳や自然言語の基礎技術に対しても様々な影響を与えてきた。

米国での情報アクセス技術に関する技術動向としては、固有表現などの限られた種類の分類問題に帰着される技術を機械学習手法で解くような研究はピークを過ぎ、対象が幅広いために、教師なしで知識を学習したり、クラウドを使って知識を作成し、Knowledge Bottleneckといわれる、知識作成の困難さを解決しようという方向の研究が盛んになっている。その際に重要なのは、どの部分の問題を手で解き、どの部分の問題を大規模コーパスやクラウドを使うかといったタスクと方法論の関係にある。最終的に知識を作成するということは、世界をどのような側面で切り取るかという問題に帰着する。固有表現の種類が3種類なのか9種類なのか200種類なのかは、その技術の対象とその応用を考えながら人間が考えていくことによって効率的な作成が行われる。それに対して、ある職業についている有名人のリストを作成する作業は人手では不可能であり、大規模なコーパスなどを使った方法論が適切である。このような知識獲得の方法自身の研究<sup>(36)</sup>や、それを使った応用研究が盛んになってきている。（隅田英一郎、関根 聡）

## 6. 今後の展望

以上ではNICT MASTARプロジェクトで開発されてきた音声言語処理技術、並びに国内外における研究動向を報告した。報告中で何度も言及されたように、イン

ターネット, Web の存在は, それ以前であれば絵空事でしかなかった音声・言語処理アプリケーションを可能にしつつある. また, これまでに紹介してきた Amazon Mechanical Turk や, 「みんなの翻訳」は, やはり, これまでにない音声・言語アプリケーションの開発体制のあり様を示唆している. 今後, このように予想もつかなかったようなアイデアも数多く出てくることであろう.

研究という観点から昨今の状況を見渡しても, これまで広く研究上の困難として認識されてきた課題が Web 等の存在により, 意外と簡単に回避されてしまうという事例も多くなってきたように思える. 例えば, ルールベースの機械翻訳における膨大なルールの整備という困難な問題が, 統計ベースの翻訳手法によってあたかも存在しなかったような位置付けになったり, 言語処理の永遠の課題とも思えた単語の意味分類や, 複雑な言語表現間の同義性の認識が, 大量の Web 文書を用いた単語クラスタリングによって, オープンドメインの音声質問応答システムを一応開発できる程度までは, あっさりと実現されてしまったといった事実がその例である.

もちろん, こうした問題の回避は, 本稿中でも指摘があったように限界があり, いずれは同じ困難な問題に立ち返る必要があるかもしれない. しかしながら, 過去においては, 困難な課題の手前で長期間止まらざるを得なかったものが, 一通り, アプリケーションの開発まで進むことができるようになったという事実は, アプリケーションのコンセプトを磨き, 更に社会的に有用なアプリケーションに思いを巡らすことが可能になったという意味で非常に大きなステップであろう. また, このステップは, 同時に, 音声・言語アプリケーションの究極の課題である, 言語表現の理解に関する深い洞察を与えてくれるものと考えべきであろう. 別のいい方をすれば, 従来重要であると認識されてきた音声翻訳, 対話システム, 情報アクセスが具体的に稼働し始めた経験を踏まえ, それらの更に一步先を考える舞台が整いつつあるのではなかろうか. 更にいえば, 今後は基盤的技術開発を着実に進めることはもちろんであるが, 音声・言語の意味, 理解, 社会的な意義付けに深く思いを巡らせ, 非常に速い速度で進化する研究の方向性を読み, 大胆な一手を打ち出すチャンスではないかと考えている.

(鳥澤健太郎)

## 文 献

- (1) 内山将夫, 阿辺川 武, 隅田英一郎, 影浦 峯, 「みんなの翻訳」言語処理学会第 15 回年次大会論文集, pp. 184-187, March 2009.
- (2) S. Akamine, D. Kawahara, Y. Kato, T. Nakagawa, Y. I. L.-Suematsu, T. Kawada, K. Inui, S. Kurohashi, and Y. Kidawara, "Organizing information on the web to support user judgments on information credibility," The 4th International Universal Communication Symposium (IUCS2010), pp. 122-129, Beijing, China, Oct. 2010.
- (3) K. Torisawa, S.D. Saeger, J. Kazama, A. Sumida, D. Noguchi, Y. Kakizawa, M. Murata, K. Kuroda, and I. Yamada, "Organizing the web's information explosion to discover unknown unknowns," New Gener. Comput., vol. 28, no. 3, pp. 217-236, Aug. 2010.
- (4) 鳥澤健太郎, 「情報爆発と音声アプリケーションの可能性」第 84 回音声言語情報処理学会研究発表会研究報告, vol. 2010-SLP-84, no. 17, pp. 1-6, Dec. 2010.
- (5) VoiceTra アプリサポートページ, <http://mastar.jp/translation/voicetra.html>
- (6) C.-L. Goh, T. Watanabe, M. Paul, A. Finch, and E. Sumita, "The NICT translation system for IWSLT 2010," IWSLT 2010, pp. 139-146, Paris, France, Dec. 2010.
- (7) J. Ni and H. Kawai, "An unsupervised approach to creating web audio contents-based HMM voices," Interspeech 2010, pp. 849-852, Chiba, Japan, Sept. 2010.
- (8) Y. Shiga, T. Toda, S. Sakai, and H. Kawai, "Improved training of excitation for HMM-based parametric speech synthesis," Interspeech 2010, pp. 809-892, Chiba, Japan, Sept. 2010.
- (9) 河井 恒, 磯谷亮輔, 安田志志, 隅田英一郎, 内山将夫, 松田繁樹, 葦苺 豊, 中村 哲, "H21 年度全国音声翻訳実証実験の概要," 2010 秋季音響論集, no. 3-9-6, pp. 99-102, Sept. 2010.
- (10) C. Hori, K. Ohtake, T. Misu, H. Kashioka, and S. Nakamura, "Weighted finite state transducer based statistical dialog management," ASRU 2009, pp. 490-495, Merano, Italy, Dec. 2009.
- (11) 喜連川 優, 牧野二郎, "IT の進展と法制度の非想定領域: 著作権法一部改正, 47-6, 47-7 による検索エンジンの合法化について," 人工知能誌, vol. 25, no. 5, pp. 621-627, 2010.
- (12) 新里圭司, 柴田知秀, 河原大輔, 黒橋禎夫, "大規模日本語ウェブ文書を対象とした開放型検索エンジン基盤の構築," 言語処理学会第 13 回年次大会, pp. 1117-1120, 2007.
- (13) 河原大輔, 黒橋禎夫, "格フレーム辞書の漸次的自動構築," 自然言語処理, vol. 12, no. 2, pp. 109-131, 2005.
- (14) R. Sasano, D. Kawahara, and S. Kurohashi, "The effect of corpus size on case frame acquisition for predicate-argument structure analysis," IEICE Trans. Inf. & Syst., vol. E93-D, no. 6, pp. 1361-1368, June 2010.
- (15) 柴田知秀, 黒橋禎夫, "文脈に依存した述語の同義関係獲得," 情処学自然言語処理研報, 2010-NL-199, 2010.
- (16) 村脇有吾, 黒橋禎夫, "形態論的制約を用いたオンライン未知語獲得," 自然言語処理, vol. 17, no. 1, pp. 55-75, 2010.
- (17) 鍛冶伸裕, 喜連川 優, "自動構築した評価コーパスからの評価表現辞書の構築," 日本データベース学会論文誌, vol. 6, no. 1, pp. 41-44, 2007.
- (18) A. Fujii, "Modeling slang-style word formation for retrieving evaluative information," Pacling2009, pp. 290-295, Sapporo, Japan, Sept. 2009.
- (19) 中山浩太郎, 伊藤雅弘, M. Erdmann, 白川真澄, 道下智之, 原隆浩, 西尾章治郎, "Wikipedia マイニング 近未来チャレンジキックオフ編," 人工知能学会論文誌, vol. 24, no. 6, pp. 549-557, 2009.
- (20) M. Itoh, M. Toyoda, and M. Kitsuregawa, "An interactive visualization framework for time-series of web graphs in a 3D environment," 14th International Conference on Information Visualisation IV, 2010, pp. 54-60, London, U.K., July 2010.
- (21) 奥村 学, "blog マイニング——インターネット上のトレンド, 意見分析を目指して——," 人工知能誌, vol. 21, no. 4, pp. 424-429, 2006.
- (22) K. Inui, S. Abe, K. Hara, H. Morita, C. Sao, M. Eguchi, A. Sumida, K. Murakami, and S. Matsuyoshi, "Experience mining: building a large-scale database of personal experiences and opinions from web documents," 2008 IEEE/WIC/ACM International Conference on Web Intelligence, pp. 314-321, Sydney, Australia, Dec. 2008.
- (23) 河原達也, 武田一哉, 伊藤克亘, 李 晃伸, 鹿野清宏, 山田篤, "連続音声認識コンソーシアムの活動報告及び最終版ソフトウェアの概要," 信学技報, SP2003-169, NLC2003-106, pp. 79-84, Dec. 2003.
- (24) 翠 輝久, 河原達也, "ドメインとスタイルを考慮した web テキストの選択による音声対話システム用言語モデルの構築," 信学論(D), vol. J90-D, no. 11, pp. 3024-3032, Nov. 2007.
- (25) 伊藤友裕, 西崎博光, 関口芳廣, "WEB 上の類似記事を利用し

た音声文書の認識性能の改善,” 情処学音声言語情報処理研報, SLP-59-9, 2005.

- (26) 河原達也, 川嶋宏彰, 平山高嗣, 松山隆司, “対話を通じてユーザの意図・興味を探り情報検索・提示する情報コンシェルジュ,” 情報処理, vol. 49, no. 8, pp. 912-918, 2008.
- (27) 吉野幸一郎, 河原達也, “Web からの情報抽出を用いた対話システムの評価,” 人工知能学会研究会資料, SLUD-B002-04, 2010.
- (28) 緒方 淳, 後藤真孝, “PodCastle: ポッドキャスト音声認識のための集合知を活用した言語モデル学習,” 情処学音声言語情報処理研報, 2010-SLP-80-10, 2010.
- (29) 西村竜一, 宮森翔子, 鈴木健太郎, 河原英紀, 入野俊夫, “安心ウェブの実現に向けた大人・子ども発話のネット収集実験,” 情処学音声言語情報処理研報, SLP-77-19, 2009.
- (30) TAC ホームページ, <http://www.nist.gov/tac/>
- (31) KDD ホームページ, [http://www.iarpa.gov/solicitations\\_kdd.html](http://www.iarpa.gov/solicitations_kdd.html)
- (32) MTurk ホームページ, <https://www.mturk.com/mturk/welcome>
- (33) WePS ホームページ, <http://nlp.uned.es/weps>
- (34) NAACL 2010 workshop ホームページ, <http://naaclhlt2010.isi.edu/workshops.html>
- (35) GALE, <http://www.darpa.mil/ipto/programs/gale/gale.asp>
- (36) Semantic Knowledge Discovery, Organization and Use ホームページ, <http://nlp.cs.nyu.edu/sk-symposium>

(平成 22 年 12 月 3 日受付 平成 23 年 2 月 16 日最終受付)



なかむら せつし  
中村 哲 (正員)

1981 京都工繊大卒。1992 京大博士 (工学) 了。シャープ中央研究所, 奈良先端大, ATR, 独立行政法人情報通信研究機構けいはんな研究所長を経て, 現在, 奈良先端大教授。ATR フェロー。音声言語情報処理の研究に従事。2010 文部科学大臣賞受賞。



いわたに りょうご  
磯谷 亮輔 (正員)

1985 東大・工・計数卒。1987 同大学院修士課程了。同年日本電気株式会社入社。2009 から独立行政法人情報通信研究機構に転向。主として音声認識, 自動通訳の研究開発に従事。日本音響学会会員。



いぬい けんたろう  
乾 健太郎

東北大大学院情報科学研究科教授。言語情報の自動編集など自然言語処理研究に従事。Computational Linguistics, 自然言語処理, 各編集委員, COLING/ACL Best Asian NLP Paper Award, 言語処理学会年次大会最優秀発表賞等各受賞。



かしわが ひでき  
柏岡 秀紀

1993 阪大大学院基礎工学研究科博士後期課程了。博士 (工学)。1993~2009 ATR 音声コミュニケーション研究所。2006 から独立行政法人情報通信研究機構。現在, 同機構音声コミュニケーション研究室長。1999 から奈良先端科学技術大学院大客員准教授兼任。



かわい ひろし  
河井 恒 (正員: シニア会員)

1984 東大・工・電気卒。1989 東大大学院工学研究科電子工学専攻了。博士 (工学)。1989~2000 (株) KDD 研究所, 2000~2004 ATR, 2004~2009 (株) KDDI 研究所, 2009 から独立行政法人情報通信研究機構に勤務。音声の認識・合成, 音響信号処理, 音声翻訳の研究開発に従事。



かわはら たつや  
河原 達也 (正員)

京大大学術情報メディアセンター教授。京大博士 (工学)。音声認識・理解及び音声対話システムに関する研究に従事。情報処理学会音声言語情報処理 (SLP) 研究会主査。



きくち たか  
木俣 豊 (正員)

1988 神戸大卒。1990 同大学院了。1999 同大学博士 (工学)。(株)神戸製鋼所, 内閣府を経て現在, 独立行政法人情報通信研究機構ユニバーサルコミュニケーション研究所長。けいはんな連携大学院教授。ユビキタスコンピューティング, コンテンツ管理, 情報分析技術の研究に従事。



くろはし ただふみ  
黒橋 禎夫 (正員)

1994 京大大学院工学研究科電気工学第二専攻博士課程了。博士 (工学)。2006 から京大大学院情報学専攻教授。自然言語処理, 知識情報処理の研究に従事。言語処理学会 10 周年記念論文賞受賞。



すみ た えいいちろう  
隅田 英一郎 (正員)

1982 電通大学院修士課程了。1999 京大博士 (工学)。現在, 独立行政法人情報通信研究機構多言語翻訳研究室長, 神戸大大学院工学研究科客員教授。機械翻訳, e ラーニングを研究。IPSJ, NLP, ASJ, ACL, IEEE 各会員。



せきね さとし  
関根 聡

1992 マンチェスター工科大計算言語学修士号。1998 ニューヨーク大 (NYU) コンピュータサイエンス学科博士号, 同年 NYU の研究助教授, 2005 から研究准教授。2010 から楽天技術研究所ニューヨークの所長を兼任。自然言語処理の研究に従事。



とりさわ けんたろう  
鳥澤 健太郎

1992 東大・理・情報科学卒。1995 同大学院情報科学専攻博士課程中退。同年同専攻助手。博士 (理学)。2001 北陸先端大助教授。同大学准教授を経て, 現在, 独立行政法人情報通信研究機構情報分析研究室長。けいはんな連携大学院教授。言語処理研究に従事。



ほり ともこ  
堀 智織 (正員)

1994 山形大・工・電子情報卒。1997 同大学院博士前期課程了。1997~1999 山形大・人文・助手。2002 東工大大学院博士 (学術)。同年 NTT CS 研。2004~2006 CMU。2007 ATR。2009 独立行政法人情報通信研究機構。2011 日本 ITU 協会賞国際協力賞受賞。



まつだ しげあき  
松田 繁樹 (正員)

2003 北陸先端大博士後期課程了。同年, ATR 音声コミュニケーション研究所研究員。現在, 独立行政法人情報通信研究機構専攻研究員。博士 (情報)。2010 文部科学大臣賞受賞。