

言葉とコミュニケーションの科学技術を ささえる

—音声言語情報処理研究会—

武田一哉 名古屋大学情報科学研究科

音声言語情報処理分野の概要

「音声言語」という言葉は英語で Spoken Language と表現される情報のありかたを表していて、単に「『音声』と『言語』」を表す言葉ではありません。音声言語情報には、テキストでは伝えられないニュアンスや、話し言葉に特有の言語現象、対話、他の情報伝達手段との相互作用、といったさまざまな現象がかかわっています。当研究会が冠する「音声言語情報処理」は、これらの諸現象の解明とモデル化を通して、人間にとって自然で親しみやすい情報処理系を構築するという、長期的な研究目標を表す言葉でもあるのです。

音声言語情報処理に関する研究対象は、人間の認知活動からインタフェースシステムなどのアプリケーションに至るまで幅広く、音声言語諸現象は、音声科学、自然言語処理、計算言語学、インタフェース、認知科学、人工知能などさまざまなコミュニティで取り上げられてきました。音声言語情報処理研究会(SIG-SLP と略記)は、このような背景の中、さまざまな視点から音声言語情報を横断的に議論する場を設けることを目的として 1994 年に設立されました。以来 SIG-SLP では、学際的な雰囲気の中でさまざまな問題が議論されてきました。これまで研究会で行われたパネル討論の一例をご覧いただければ、その一端がご理解いただけると思います。「なぜ音声認識は使われないか・どうすれば使われるか？」(第 1 回)、「話し言葉の文法構築は可能か？」(第 6 回)、「インタラクティブなシステムの評価をどのように考えていくか」(第 7 回)、「統計的言語処理／音声言語処理における大規模言語データベースの利用」(第 11 回)、「音

声言語対話処理の将来像を探る」(第 22 回)、「モバイル・インターネット時代の対話記述」(第 38 回)、「擬人化／顔インタフェース再考」(第 52 回)

SIG-SLP では研究の方法論もさまざまですが、音声言語情報処理諸過程への数理モデルの適用と、音声言語現象の注意深い観察に基づく知見の獲得が大きな 2 つの方法論といえます。前者は劣化音声の回復、音響モデルや言語モデルの構築といった研究課題の中心的な研究手段です。一方後者は、対話制御の戦略、複数インタフェースの統合といった研究において主として用いられる方策です。いずれも人間を含む複雑で揺らぎを持つ系の研究ですので、大規模な研究用コーパスの存在が前提であり、SIG-SLP では常に研究用コーパスに大きな関心が寄せられてきました。

音声言語情報処理技術を応用する産業製品市場は、音声認識・合成・対話システムを中心に拡大を続けています。製品の性能も多様な研究努力の成果を受け、着実に向上しており、今後さらに多くの応用が開拓されることが期待されます。より多くの利用者・利用形態を想定した製品開発のために、要素技術の頑健性や利便性の向上が、ますます重要な研究課題となりつつあります。

研究会の活動

SIG-SLP では、年 2 回「音泉研究会」と称する合宿形式の研究会を開催しており、参加者全員の自己紹介やナイトセッションでの打ち解けた討論が好評で、毎回多くの参加者がいます。毎年 12 月には電子情報通信学会の「音声」、「言語理解とコミュニケーション」の両研究会と共同で「音声言語シンポジウム」を開催しており、延べ 200 名を超える参加者がいます。さらにここ数年「音声言語情報処理技術デベロッパーズフォーラム」と題して、音声認識を中心に、応用技術に的を絞った研究会を開催し、企業研究者の注目を集めています。今後、音声言語処理技術のベンダとサービスプロバイダを中心に、企業技術者の情報交換の場に発展していくことが期待されています。

SIG-SLP では、複数の研究機関の研究者が協力して、研究基盤や評価基盤の整備を行うワーキンググループ活動を推進しています。これまでに、連続音声認識技術、擬人化エージェント技術、雑音下音声認識技術に関する 3 つのワーキンググループが活動を行い、日本語大語彙連続音声認識システム Julius、擬人化エージェントツールキット Galatea、雑音下音声認識評価基盤 CENSREC などの整備を行ってきました。今年度には、後述する新

しい研究分野である「音声文書処理」に関するワーキンググループが立ち上がりました。これらの研究基盤は広く関係研究者に共用されて、我が国の音声言語情報処理研究に大きく貢献しています。

さらに、日本語大語彙連続音声認識システム、擬人化エージェント技術については、ワーキンググループ活動終了後にコンソーシアムを立ち上げ、自主財源によりソフトウェアの改良・配布ならびに講習活動を行っています。

研究動向

限られた紙面ですが、ここで SIG-SLP を中心とした音声言語情報処理分野の最新研究動向のごく一部を紹介したいと思います。

(1) 音声認識アルゴリズムの発展

SIG-SLP ワーキンググループ活動を起源とした連続音声認識フリーソフトウェア Julius の開発が、京大、奈良先端大、名工大等において継続的に進められてきました。最新の認識技術を取り入れつつ、多様な研究目的やアプリケーション開発に利用可能となっており、我が国の音声言語情報処理研究の進展に大きく貢献しています。一方企業では、携帯電話や組み込みシステムなどのプラットフォーム上への音声言語処理系の実装、カーナビを中心に多言語化が開発課題となっています。要素アルゴリズムの移植性、音響・言語モデルの汎用性とモデル適応方法に関する研究の重要性が増しています。

(2) 音声対話システムの運用とデータ収集

奈良先端大のグループは、生駒市のコミュニティセンターに「たけまる」音声対話システムを設置し、長期間にわたる運用実験の結果を報告しています。このシステムは大語彙連続音声認識結果に対して、一種のテーブル参照を行うことで発話内容の解析と応答文の生成を行う非常にシンプルな設計となっています。このため「たけまる」は移植性や頑健性に優れたシステムとなっており、そのアーキテクチャを受け継いだ対話システムが各地で運用されるに至っています。このような音声対話システムの実運用はさまざまな機関で進められており、収集されたデータや知見に基づき、さらなるシステムの改良が進められています。

(3) 音声ドキュメント処理へのチャレンジ

国立国語研究所・情報通信研究機構（旧通信総合研究所）・東工大が中心となって「日本語話し言葉コーパス (CSJ)」が収集・公開され、これまで書き言葉を主たる対象としてきた統計的言語モデルが、日本語話し言葉につ

いても研究され始めました。話し言葉特有の多様な現象を統計的言語モデルの枠組みで処理することで、CSJ に収録されている講演音声の認識精度が向上しています。さらにこの技術を、講義音声、録音議事記録や放送音声等に应用することで音声をセマンティックコンテンツ化できれば、膨大な量の音や映像を情報資源としてより高度に活用することができます。このような「音声ドキュメント」の本格的研究を進めるために、研究会では新たなワーキンググループを立ち上げ、その研究資源の整備を開始しました。

将来展望

先ごろ首相官邸で取りまとめられた長期戦略指針「イノベーション 25」において、音声言語情報処理技術は「多様な人生を送ることができる社会」、「世界に開かれた社会」を実現するための重要な技術と位置づけられました。人間にとって最も身近な情報伝達手段である音声言語情報処理研究の重要性は、今後も変わることはないでしょう。SIG-SLP では、このような社会の要請に応えるために、これからも着実に活動を続け、幅広い問題を議論する場を提供していきたいと考えています。

(平成 19 年 7 月 10 日受付)



武田一哉 (正会員)

kazuya.takeda@nagoya-u.jp

1985 年名大修士課程修了。ATR、KDD 研究所などを経て、現在名古屋大学情報科学研究科教授。音響信号処理、音声情報処理、行動信号処理などの研究に従事。SLP 主査。

SLP