

小特集—音声研究の新たな方向を探る—

「音声研究の新たな方向を探る」小特集にあたって*

板橋 秀一 (筑波大学電子・情報工学系)*

43.72.-p

音声・言語情報処理技術は、マルチモーダルインタフェースの中核をなす技術として注目を集めており、より良いヒューマンインタフェース実現のためには、より自然な音声対話・理解技術が求められている。最近では伝統的な音声認識や音声合成の研究の他に、音声コーパス (データベース)、音声入出力システム評価、言語・話者認識、対話処理、音声検索・要約、話題抽出、話速変換、福祉応用、インターネット、マルチメディア処理、マルチモーダル入出力に関連したテーマが増えてきている。

実際の場面でも音声認識や音声合成がある程度利用されるようになってきたが、音声認識では不特定多数の話者や雑音環境、方言や多言語への対応等が求められている。音声合成では、音質の向上はもちろんであるが、任意の声質での合成等が求められている。応用面からは福祉への対応が求められよう。このような背景のもとに、最近、音声・言語関連の大型プロジェクトが幾つか開始されたので、本号ではそれらに焦点を合わせている。

本小特集は、20世紀最後の年に21世紀につながる音声研究のあり方を、幾つかの視点から語っていただくことを意図しており、その内容は大きく4部に分けられる。第1部は本稿で、本小特集の紹介と位置づけを述べる。第2部では「大型研究が目指す新たな方向」と題して、最近発足した5件の音声関係大型プロジェクトの代表者にその目指すところや抱負を述べていただく。第3部ではCREST, PRESTO, IPA等による「プロジェクト研究所感」を5人の代表者の方に、プロジ

ェクト立ち上げのころの意気込みや研究を遂行していく上での苦労話、抱負等をエッセイ風に語っていただく。第4部では「新たな方向を探る研究者座談会」を行い、第一線で活躍されている方々に音声研究の今後の方向等について語っていただく。

音声認識・言語処理に関する研究は、大規模音声・言語コーパスと確率モデルに基づく統計的認識手法の発展により、近年大幅な進歩を遂げた。しかし、より精密な音響モデルの構成とその環境・話者への適応、より精密な言語モデルの構成とそのタスクへの適応、音響モデルと言語モデルの統合等数多くの課題が残されている。音声合成の研究では、PSOLA, CHATR等コーパスベースの研究が広く進められている。一方、脳機能の解明に関する研究が盛んに行われており、それに関連して音声生成・知覚や聴覚機能に関する研究テーマも取り上げられている。

第2部及び第3部のプロジェクト研究では、音声の基礎から応用にわたる広い範囲をカバーしている。すなわち、脳活動の計測、神経回路モデル、コミュニケーション理論、音声知覚等の基礎的研究をはじめとして、音声分析・合成・認識・聴覚等の音声・音響処理、それに実世界で利用可能な音声翻訳技術の実現や音声言語処理技術の言語教育への応用等広範囲にわたっている。これらの研究の多くが、実証的な研究を進めるために観測データの収集や大規模音声・言語コーパス構築等を重要なテーマとして取り上げていることが最近のプロジェクト研究の特徴として上げられる。

音声研究のための資源である音声・言語コーパスを扱う組織として、アメリカではLDCが1992年に、ヨーロッパではELRAが1995年に設立されて、音声・言語研究の進展に大きな貢献をしている。それに比べて、日本を含むアジアではその取り組みの遅れが目立つ。日本にもこのような組

* Foreword to the special issue on Searching for New Directions of Speech Research.

** Shuichi Itahashi (Institute of Information Sciences and Electronics, University of Tsukuba, Tsukuba, 305-8573)

組織が必要であるということから、自然言語処理関係の研究者が中心となって音声研究者も加わり、「言語資源共有機構 (GSK)」が設立された。その設立シンポジウムを昨年度開催し、現在ホームページの整備を進めている段階であるが、予算の裏付けのあるしっかりとした組織に発展させる必要がある。

ELRA の主催により第 1 回 LREC (言語資源と評価国際会議) が 1998 年にスペインのグラナダで¹⁾、第 2 回が 2000 年 6 月にギリシャのアテネで開催された²⁾。1 回目に 500 名、2 回目は 600 名の参加者を得て、論文数はそれぞれ 200 件と 280 件である。言語資源と評価に関する国際会議にこれだけの論文と参加者が集まることは、それだけこの分野に対する関心が高いことを意味する。最近のアテネの会議では、従来から構築が進んでいる西欧諸言語のコーパスのほかに、ポルトガル語、ロシア語、スロベニア語、ラテンアメリカ諸語の音声コーパスや車内音声、家電製品用音声コーパス構築の報告等が見られた。LDC からは、話題抽出・追跡を目的とした多言語放送ニュースコーパスの発表があった。このように、諸言語の音声コーパス構築プロジェクトが各国で進められており、それを基盤にして、より実用面を意識した音声研究が行われていることが伺える。

その一例としてオランダ語に関するデジタル言語基盤構築プロジェクトが挙げられる。1998 年から 2002 年までの 5 年計画で、1 千万語規模の現代オランダ語コーパスを構築する計画である。この計画では、HLT (人間言語工学) における学界と産業界の間の共同研究の促進を目的の一つに挙げている。

もう一つの例は、産業界独自の言語資源センターの設立である。LDC, ELRA ができたとはいえ、実際に商業ベースで使用できるコーパスの量は限られていることから、産業界では商用に利用できるコーパスの構築が望まれている。モトローラ社から自社グループで使用できる言語資源構築のために HLT データ資源センター (HDLRC) を設立したことが報告された。このような組織の必要性は理解できるが、商用への利用がある程度済んだ段階では、一般への公開を考えてほしいものである。

今後は音声コーパスやソフトウェアの積極的な公開が求められるよう。自然言語処理の世界では、構文解析・形態素解析ソフトウェアが公開・利用されていることは良く知られている。音声処理の世界でも、各種の音声コーパスはもちろんのこと、分析・変換・合成システムや音声認識用基本ソフトウェアが公開・利用されていることは周知のとおりである。情報処理の世界では、デファクトスタンダードになるようなデータやソフトウェアを世界に発信していくことが研究の活性化につながっていくと考えられる。

音声に関連した大規模なプロジェクトがこれだけ多数同時に進行する状況はこれまでになかったことである。それだけ音声研究に対する期待が多きいと考えられる。この機会に大規模音声・言語コーパスを構築することはもちろんであるが、編集・ラベル付け等のコーパス構築用ソフトウェアを整備して公開することが望まれる。できれば話者認識や多言語処理も視野に入れたいところである。音声研究の新たな方向を探るために、本特集号が参考になることを期待したい。

文 献

- 1) A. Rubio, N. Gallardo, R. Castro and A. Tejada, Eds., *Proceedings of the First International Conference on Language Resources and Evaluation (LREC)*, Granada, Spain, 28-30, May (1998).
- 2) M. Gavrilidou, G. Carayannis, S. Markantonatou, S. Piperidis and G. Stainhaouer, Eds., *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC)*, Athens, Greece, 31 May-2 June (2000).

関連ホームページ

LDC: <http://www ldc.upenn.edu/>

ELRA: <http://www.icp.grenet.fr/ELRA/>

COCOSDA: <http://www.itl.atr.co.jp/cocosda/>

Oriental COCOSDA: <http://www.milab.is.tsukuba.ac.jp/o-cocosda/>

NIST: <http://www.nist.gov/>

板橋 秀一

昭和 39 年東北大・工・通信卒。昭 45 同大学院(博)修了。同年東北大通研助手。昭 47 電子技術総合研究所入所。昭 52 ストックホルム王立工科大学客員研究員。昭 57 筑波大電子・情報工学系助教授。現在同教授。工博。音声・画像・自然言語処理の研究に従事。電子情報通信学会、情報処理学会、人工知能学会、言語処理学会、認知科学会、IEEE, ASA, ESCA 各会員。