

論文 / 著書情報
Article / Book Information

論題(和文)	話し言葉の音声認識と理解を目指して
Title	
著者(和文)	古井貞熙
Author	SADAOKI FURUI
出典(和文)	シンポジウム 2 0 0 2 大阪「明日を目指す科学技術」科学技術振興調整費の成果を中心として, Vol. , No. , pp. 29
Journal/Book name	, Vol. , No. , pp. 29
発行日 / Issue date	2002,

話し言葉の音声認識と理解を目指して

…プロジェクト総論…

東京工業大学大学院情報理工学研究科 古井 貞熙

書き言葉に近い音声なら高い精度で認識できるようになってきたが、言い直し、言い淀み、繰り返し、間投詞、不正確な発音などの現象を含んだ自然な話し言葉を認識しようとする、大幅に認識性能が低下してしまうのが現状である。講演や放送などの話し言葉を文字化し、その内容を把握し、再利用や情報検索に使えるようにしたいという要望や、コンピュータシステムとの音声対話への期待が高まっているが、その実現には多くの課題がある。本プロジェクトでは、話し言葉処理に関わる種々の重要な問題を取り上げ、話し言葉からその意味・内容・話し手の意図などを自動的に抽出する技術の基盤を確立することを目的としている。このために、種々の音声現象を含む大量のコーパス（データベース）を構築し、その統計的解析に基づいて、新しい音声情報処理技術を創造することを目指している。

書き言葉に関しては、新聞記事などの大量の原稿が比較的容易に入手できるが、話し言葉は、音声を録音して書き起こし、それを解析して種々の情報を付与する必要があるため、コーパスを構築するのは容易ではない。本プロジェクトでは、タスクを限定しないモノログとして講演を主体とし、対談なども含み、のべ約 700 時間、形態素にして約 700 万語規模のコーパスを作成している。その全体について、書き起こし、読みの付与、形態素解析などを行い、全体の約 10%のコアについては、韻律情報などパラ言語情報まで含めたコーパスとする。

話し言葉に適応した形態素解析ツールを作成することも、プロジェクトの重要な目標の一つである。このために、コアの部分については、人手による処理を含めた精密な解析を行い、形態素解析ツールの学習に用いる。これを用いたコーパスの解析により、話し言葉の統計的言語モデルを作成し、さらに、種々の世界知識、言語的／パラ言語的情報を抽出・体系化する。同時に話し言葉に適応した音素モデルを構築し、これらを用いて、種々の話し言葉に対応できる音声認識技術の実現をはかる。さらに、認識結果を用いた音声の自動要約技術を構築し、要約された内容をキーワード、要約文、およびその合成音声などで提示する。この目標達成のために、情報科学・工学系だけでなく、音声学、言語学などに代表される人文系言語諸分野の研究者を含めた協力関係により、多面的かつ総合的なアプローチを展開する。

これまでの研究の結果、話し言葉コーパスに基づく音素および言語モデルの構築により、講演に対する音声認識誤りを、プロジェクト開始時の 55 %から、約 1/2 に減らすことに成功した。発声速度、言い直し頻度、および未知語率が、認識誤りの変動の主要な要因であることが明らかになった。音声認識結果から、各単語の重要度（情報量）、各単語の音響的および言語的信頼度、単語連鎖の文法的正しさ、および各単語の係り受け構造に基づく単語連鎖の意味的正しさに基づいて、重要文を抽出し、さらに各文を圧縮することによって、自動的に要約を作成する方法を提案した。人が作成した要約を基準とした評価により、提案法の有効性が確認されている。