

「高度メディア社会の生活情報技術」
平成11年度採択研究代表者

ニック キャンベル

(株)国際電気通信基礎技術研究所 プロジェクトリーダー

「表現豊かな発話音声のコンピュータ処理システム」

1. 研究実施の概要

話し手の発話には、言語情報の他に感情、意図、態度などが含まれており、それらは文字上の意味を越えた心理的な意味合いを伝達する。これらの情報処理を自動化するため、コンピュータによる統計モデルを研究する。本研究では、研究成果物の一つとして「感情」、「意図」、「態度」などを含む大規模音声データベースの構築(自然対話約70%、スタジオ収録音声10%、特定用途用20%)を目標としてあげている。従って初年度は、研究素材データの収集を進めた。自然対話のデータ収集を開始すると同時に、特定用途音声データ収集を行い、収録手法の検討を行った。音声波形接続型音声合成システムのための単位データベース作成を始め、応用システムの試作まで行った。また、音声合成アルゴリズムや分析ツールの開発を行い、来年度の音声データ分析に活用する予定である。

2. 研究実施内容

(1) 音声言語グループ

本研究チームは音声合成技術、音声認識技術、韻律情報処理技術の検討を行った。

- ・ 混合正規分布モデルに基づく声質変換アルゴリズムの改良を、音声分析合成システムSTRAIGHTと周波数軸伸縮により行なった。音質、話者性の評価を行い、有効性が確認できた。
- ・ 高度な不特定話者音韻モデルとMLLRによる話者適応アルゴリズムにより、音声データベース自動音韻ラベル付与アルゴリズムを開発した。
- ・ 自然発話音声データベースおよび書き起こしテキストからフィラーの統計的生成モデルを作成して、大語彙連続音声認識システムに組み込み、有効性を確認した。
- ・ 2話者のバイリンガル音声データの収集とデータベース化を行なった。また、顔画像も含めた1話者分の音声データの収録とデータベース化を行なった。

(2) 発話様式グループ

本研究チームは音声合成のための韻律分析とモデル化、そのために必要な、多

様な発話形式による自然音声データベースの構築を進める。

- ・ 自然対話データ、特に話者態度を示す音声データの収集方法について検討し、所望するデータの収集方法に関して、具体的なデータベース設計手法を提案した。
- ・ 発話態度、意図、感情によるカテゴリーごとに、典型的ピッチパターンの抽出、分析を行い、そのモデル化を行った。
- ・ テキスト、映像、音声の3メディアの全組合せを提示素材として、聴覚実験を行い、人の感情、態度の認知機能における音声メディアの役割について検討した。

(3) 意味構造グループ

本グループでは、音声に関わる韻律や音色により、聞き手に伝わる印象、意味の認識にどのような影響が現れるか考察した。具体的には、「りきみ」と呼ばれる現象や感動詞「ああ」の韻律情報の考察を行なった。この対照研究として、中国語の応答表現、感動詞に関して検討も行なった。音韻レベルの考察としては、長音、または二重母音が、何らかの話者の意図や感情と結びついていることに関して準備的考察を行なった。また、反復などの発話スタイル、イントネーションが音声要素に与える影響、聴覚印象に与える影響に関して検討した。さらに、これまで多く論じられている旧情報・新情報の理論を音声の枠組みから捉えた実験を行なった。この他に、ヨーロッパ諸語とアジア諸語の「呼びかけ」表現の音声的観察、日本語と中国語の卓立（プロミネンス、フォーカス）は音響的観察から、「卓立強調」の表現意図に寄与している要因を検討した。日本語の習得程度と韻律情報の理解には関係があるのか、という応用的な観点から教育への実験的アプローチも行なった。

(4) 比較言語文化グループ

本グループでは、従来研究の文献検討および実験方法の検討を行った。

- ・ 感情音声や態度音声に関する従来の研究について文献の調査を行い、国際的な関連研究のまとめを作成した。（別紙文献リスト120P）
- ・ 日本で行っているデータ分析手法との対比検討を目的として、実験手法による検討を行う。そのため具体的には、意味的に中立な文を素材として、発話形式別に現在収録中である。
- ・ ストレス下にある感情音声の収集を目的として、空軍操縦士のフライトシミュレーター操縦時の音声収録の具体的計画に入った。

(5) コミュニケーション支援グループ

本グループでは、アプリケーションとしてコミュニケーション支援システムの提案、実装、評価を行なった。

- ・ 音声コーパスの作成について、日常対話に利用できる音声合成用として設計されたテキストコーパスを、病状の進行に伴い、声を失うことが予期される筋萎縮性側索硬化症（ALS）患者1名に読んでもらい、音声波形接続型音声合成システム用の音声コーパスを作成した。
- ・ システム実装について、任意の文章を表情豊かな合成音声に変換するコミュニケーション支援システムを作成した。システムには患者自身の利用に向けて患者が日常よく用いる文章を別ウィンドウでリスト表示する機能を装備した。出力される合成音声とGUIについて評価を行い、良好な結果を得た。

(6) 対話音声グループ

本グループでは、人間の日常的な言語使用において、意図・態度・感情がどのようにしてパラ言語によって表出されているかの解明を目指し、会話データの収録、音響的な分析、モデル化を行っている。また、態度・感情の表出に異常のある自閉症患者に対しても同様な分析を行い、通常人と比較することで、パラ言語による伝達の本質を探ることも1つのテーマとしている。本年度は、会話データ収録の計画について話し合い、他グループと共有可能なコーパスの設計について検討した。自閉症患者のデータについては、試験的に収録し、簡単な分析を行った結果、態度・感情の表出に異常をきたすことにより、音調が正しく使えないだけでなく、話速や句切りにも影響があることが示唆された。

(7) システム統合グループ

本グループでは、発話形式研究のためのコーディネートと応用システムの研究を行なった。

- ・ 音韻、韻律データのインタラクティブ・ラベリング・システムの構築。
- ・ 波形接続型音声合成用データベース作成システムの構築。具体的には、PDSソフトウェアを用いたラベリングツールを開発。本システムは聴取者の聴覚的印象を音声波形に直接マッピングし、ラベリングを行った。
- ・ 音声信号の韻律変換ソフトウェアの開発。
- ・ 波形素片接続型（音素単位）単位選択ソフトウェアの開発。
- ・ キー入力に依存しない入力方式を用いる音声合成用GUIの開発。
- ・ 波形編集方式のための新たな単位選択アルゴリズムの検討。
- ・ 自動ラベリングを行いながら韻律バランスを計算し、不足する韻律パターンを含む文章を自動作成の上、提示して、データベースを作成するシステムの構築。
- ・ 音声とテキストを入力とするHTKを用いた音韻セグメンテーションソフトウェアの開発。

3. 主な研究成果の発表（論文発表）

なし