

本論文は「Japanese Prosody Analysis and its Application to Computer-Aided Language Learning Systems (日本語の韻律解析と発音教育システムへの応用)」と題し、韻律を主体とする日本語発音教育システムへの応用を念頭に置き、特殊拍音素識別、アクセント型・イントネーションタイプ識別手法の開発とその基となる人間の基本周波数(F0)知覚特性を調べたもので、全7章からなり、英文で記述されている。また、付録で、実際に構築した単語アクセント型発音教育システムが述べられている。

第1章は「Background: Current speech processing techniques applied to CALL systems」であって、発音教育に主眼を置いた計算機援用言語学習システム(Computer Aided Language Learning system, CALL system)の内容と一般的な構成を概観し、対象となる音声の特徴と必要となる音声処理技術について述べている。特に、韻律に関連して重要となる特徴量の正規化について整理している。さらに、CALLシステムにおいて学習者に提示する発音修正の媒体と指針について、過去の研究例を参照した上で、その方向性を示している。最後に第2章以降の論文の構成を述べている。

第2章は「Formulations of phone durations considering speaking rate variations and its application to Japanese double-mora identification」と題し、外国人にとって学習が難しいとされる日本語特殊拍音素の識別手法を開発している。長母音等の特殊拍音素は持続時間が知覚の重要な特徴であるが、単純に長短だけで対応する普通音素との識別を行うと発話速度が変化した場合、精度の高い識別ができない。先行音と後続音の音素長への影響が線形和で表されるとした上で発話速度による各音の持続時間の変化を線形回帰により数式化し、それを基に、発話速度を推定して特殊拍音素の識別を行う手法を開発し、実験によりその有効性を示している。

第3章は「Japanese rhythm: Consideration on isochronal mora timing」と題し、日本語のリズムの基本とされるモーラ等時性について、音響的な面から考察している。等時性を表すメトロノームのビートにあわせて発声した音声を対象として分析を行い、ビート位置が音節開始時点に対応する子音開始時点よりも母音あるいは破裂開始時点により良く対応することを示した上で、モーラ等時性が母音-子音(VC)を単位として実現される傾向にあるとしている。

第4章は「Japanese pitch accent in sentence level based on F0mora representation」と題し、文音声中心アクセント句のアクセント型識別手法を開発している。これは、フレーム毎のF0等の特徴量とする従来の方式と異なり、モーラ毎の代表F0値に着目するもので、対数F0パターンの線形回帰から得られるF0のモーラ内目標値と、隣接モーラF0値との対数軸での差分を考慮した上で、発話による変動をガウス分布でモデル化することにより、高精度の識別結果を得ている。モーラ単位として、従来のCVと前章のVCを考慮し、モーラ代表F0値として、平均値と目標値を挙げて、それらの組み合わせについて詳細な実験的検討を行っている。

第5章は「Acoustic parameters related to Japanese sentence final intonation」と題し、外国人のための日本語教材に挙げられた6種の文末イントネーションについて、母語話者の知覚能力を調べた上で、文末の平均モーラ長と文全体に対する割合、文末の対数F0パターンの傾きとF0変化値、文末の対数パワーパターンの傾きを特徴量とし、前章と同様のガウス分布のモデル化により80%程度の識別率を達成している。さらに、非母語話者を対象とした知覚実験により、上記のパラメータのイントネーションタイプ知覚における効果を調べ、F0変化値、文末の文全体に対する割合が有効であることを示した。

第6章は「F0 parameters related to human pitch perception for Japanese accent and intonation」と題し、人間のピッチ知覚を良く表現するモーラ代表F0値について、知覚実験により詳細に調べている。これは、楽器(MIDI)音のピッチを半音より細かい単位で調整可能なツールを用意し、モーラ単位の音声のピッチと同一になるように被験者に調整させるもので、得られるMIDI音のピッチを音声ピッチの知覚された値として、第4章の種々のモーラ代表値との比較検討を行っている。その結果、VC単位の有用性を示し、F0平均値とF0目標値の間に適切なモーラ代表F0値が定義されるとしている。

第7章は「Conclusion and future works」であって、本研究で得られた成果を要約し、将来の課題としてCALLシステムへの応用を述べている。

以上を要するに、本論文は、日本語音声のリズムとアクセント・イントネーションについて、知覚面を中心に分析し、その結果を基に、発話速度を考慮した特殊拍音素の識別、F0の変化を考慮したアクセント型とイントネーションタイプの識別手法を開発した上で、それを利用した日本語発音教育システムの可能性を指摘している。人間の韻律知覚過程を意識した従来にない発音教育システム構築の指針を与えるとともに、韻律生成・知覚に対する新しい知見を与えるもので、音声基礎、音声処理技術分野の発展に大きく寄与するものであり、電子情報工学に貢献するところが少なくない。

よって、本論文は博士(工学)の学位請求論文として合格と認められる。