

## 学位論文審査結果の要旨

専攻名	システム工学 専攻	氏名	松下 真也
学位論文題目	少数言語における保存活動の支援に関する研究		
主査・副査	主査	高瀬 治彦	㊟
	副査	若林 哲史	㊟
	副査	川中 普晴	㊟
	副査		㊟
<p>審査結果の要旨</p> <p>本学位論文は、少数言語の保存活動を支援するため、収集した音声データを文字に書き起こし、その解析を行う作業の一部を計算機により自動化するものである。</p> <p>話者が少ない少数言語は、少なからず絶滅の危機に瀕しており、その保護・保存が求められている。これらの作業を計算機で自動化しようとする、対象言語固有の事前知識や、機械学習をするための多量のデータが必要である。しかし少数言語に対して、これらはいずれも乏しい状況であり、既存の手法の適用は困難である。本学位論文では、そのことをふまえて、これらの作業のなかでも初期の作業にあたる音声の書き起こし・書き起こしテキストの解析に着目し、これらの計算機による支援を試みた。</p> <p>音声の書き起こしでは、音声データを、その発音を表すテキストに変換する。この際に、音声データ上に、認識したい音のどれがいつ現れたのかを特定する必要がある。本学位論文では、保存対象の言語を構成する音が不明瞭な段階でこれを行うため、画像処理技術を応用した手法を提案した。この手法は、音声の周波数に関する特徴量の時間変化を表した図の上で、認識したい音と音声データの類似点を検出し対応づけることで、音の種類と位置を検出する。この成果は、音声の書き起こし作業の進捗に大きく貢献する。</p> <p>書き起こしテキストの解析では、与えられたテキストから最終的にはその意味をとらえる。一般的にこの解析は、分かち書きや形態素解析、構文解析、意味解析などを組み合わせて行う。本学位論文では、これらの解析のうち、最も基本的な、与えられたテキストを意味のある塊に分割する「分かち書き」に焦点を当て支援した。特に、事前知識・データ量が不十分である状況をふまえて、データ量が不足した状況下での教師なし分かち書き手法について検討した。そのために、1段階目で不可分な塊を見つけ、その結果を利用して2段階目で正確な分かち書きを行う段階的な手法を提案した。この成果により、事前知識・データ量が不十分な状況下でも、その後の解析に役立つ分析結果を得ることができるようになる。</p> <p>以上の研究成果は、新規性と有用性を兼ね備えており、工学的に有用である。そのため、博士(工学)の学位論文審査として合格と判定する。</p>			