別紙1－1

論文審査の結果の要旨および担当者

報告番号	※ 甲　第 12872 号

氏　　名　宮崎 千明

論 文 題 目

Analysis and Automatic Generation of Japanese Characters' Utterances

（日本語におけるキャラクタ発話の分析とその自動生成）

論文審査担当者

主査	名古屋大学	教授	佐藤 理史
委員	名古屋大学	教授	古橋 武
委員	名古屋大学	教授	河口 信夫
委員	名古屋大学	准教授	笹野 遼平

別紙１－２

論文審査の結果の要旨

宮崎千明さん提出の論文「Analysis and Automatic Generation of Japanese Characters' Utterances (日本語におけるキャラクタ発話の分析とその自動生成)」は、特定のキャラクタらしい発話の言語的特徴の分析と自動生成手法についてまとめたもので、全６章から構成されている。

第１章では、まず、本論文における「キャラクタ」および「発話」を定義し、研究の目的を提示する。近年、対話エージェントの実現において、エージェントにキャラクタ性を付加することが重要視されている。現在は、それぞれのキャラクタ毎に発話を人手で用意しているが、このやり方では多くの異なるキャラクタを実現する場合には多大なコストが発生する。発話へのキャラクタ性付与に関連するこれまでの研究をレビューしたのち、キャラクタらしい発話の自動生成を、文のスタイル変換の一種として捉えるという本論文の立場を明らかにし、「特定のキャラクタに適した発話を自動生成するにはどうしたら良いか」という問いを提示する。さらに、この問いを、「話者のキャラクタらしさを表す言語的特徴にはどのようなものがあるか」および「どのようなアルゴリズムを用いれば、話者のキャラクタらしさを表す言語的特徴を持った発話を生成できるか」の２つに分解し、本論文の第２章から第５章の位置づけを示す。

第２章では、日本のキャラクタの発話に現れる言語的特徴を分析し、それらの分類体系を示している。具体的には、語彙選択、モダリティ、文構造、音声、表記など13種類の分類カテゴリを提示し、架空のキャラクタ17種類の発話に含まれるキャラクタに特徴的な言語表現の約90％をカバーすることを示した。さらに、13種類のそれぞれの言語的特徴の有無とキャラクタらしさの関係を調べ、相関関係が観察されることを示した。

第３章では、特定のキャラクタらしさを持つ発話を自動生成する一つの方法として、文末表現を自動変換する手法を示している。日本語では、文末表現がキャラクタらしさを表す重要な言語的特徴の一つである。キャラクタ属性（性別・年代・居住地）が付与されたTwitterデータからそれぞれの属性に特徴的な文末表現を収集し、それぞれのキャラクタに応じて文末表現を置換することよって、発話の自動生成を実現する。変換時に、文末表現の直前の述語の品詞や文末表現が内在する意味（たとえば、否定や疑問）を考慮することにより、約95%の発話に対して、文法的・意味的に違和感のない変換を実現するとともに、約90%の発話を特定のキャラクタに適したものに変換できることを実験的に示した。

第４章では、第３章で示した方法を拡張し、発話に含まれるすべての文節において、そこに含まれる機能語を確率的に変換する方法を示している。この方法により、第３章の方法で問題となった２つの問題、すなわち、文中・文末でのスタイルの不統一、および、生成される言語表現の画一性、を解消する。ここでは、特定のキャラクタらしさを強調するように人間の手で書き換えられた発話コーパスを利用し、人間による書き換え操作を文節単位で抽出し、変換規則として用いる。さらに、その書き換え操作の出現確率に従って、変換規則を選択する。この変換方法により、特定のキャラクタの発話として認知される発話が、変換前と比較して、最大約30％増加すること、さらに、発話に含まれる表現の多様性が向上することを実験的に確認した。

第５章では、キャラクタらしさの表現力を高める手段として音変化表現に着目し、その分類体系を示すとともに、音変化表現を人為的に生成するための知識を整理した形で示している。分類体系は11大分類、34中分類に区分され、全部で137種類のパターンをカバーする。そしてそれぞれのパターンに対し、音変化現象の実際とその生起環境（対象なる語、および、周辺の語の文法的・表記的特徴）をまとめている。音変化表現を半自動的に検出するアルゴリズを考案し、小説やコミックのキャラクタの発話に現れる音変化表現の約80%が、137種類のパターンでカバーされること、および、音変化パターンの情報を利用することで、推定性能が向上するキャラクタが存在することを、実験的に示した。

最後の第６章では、本論文をまとめるとともに、今後の方向性について筆者の考えを示している。

このように、本論文は、発話のキャラクタらしさに着目し、日本語においてどのような言語表現・形式がキャラクタらしさと結びついているのかという言語学的分析、および、発話を特定のキャラクタらしい発話に自動変換する具体的なアルゴリズムを示したものである。この分析と技術は、言語の個人性に迫るという科学的な側面と、これからの対話エージェントの実現に有用な技術となりうるという工学的・応用的側面を併せ持ち、学術的にも工学的にも、その価値は高い。よって、博士論文として十分な内容であると判断する。