

博士論文の要旨及び審査結果の要旨

氏名 細井 尚子
学位 博士 (学術)
学位記番号 新大院博 (学) 第 210 号
学位授与の日付 平成 27 年 9 月 24 日
学位授与の要件 学位規則第 4 条第 1 項該当
博士論文名 源氏物語のテキストマイニングと特徴抽出

論文審査委員 主査 教授・藤澤 延行
副査 教授・松原 幸治
副査 教授・赤林 伸一
副査 助教・山縣 貴幸

博士論文の要旨

本論文は全四章と付録から構成されている。第一章は本研究の背景と本論文の構成、第二章は源氏物語の時代背景、作者について、第三章はテキストマイニングと数量化Ⅲ類、クラスタリングの導入による特徴抽出法についてとその適用結果と考察、第四章は結論、そして最後に付録として源氏物語の執筆順序に関する文学的考察がされている。

第一章では、源氏物語が執筆された時代についてと文学的課題としての源氏物語の複数作者説について述べた。すなわち、源氏物語の作者は一般的には紫式部といわれているが、時代的背景から作者を特定しがたく、文学的には彼女以外にも作者が存在するともいわれている。本研究では、コンピュータの導入による文学作品の解析法を新たに提案し、その解析法を源氏物語に導入することで、源氏物語の複数作者説に対して、科学的な立場から評価し、源氏物語の作者について考察することを目的とする。

第二章では、源氏物語の時代背景と作者について考察するにあたり、助動詞の用法に着目する方法について述べ、それにより作者の違いを明らかにする手段とすることの意義を考察した。特に、過去を表す助動詞であるき、けりにおいてその用法に作者の特徴があることが示された。

第三章では、コンピュータを用いた助動詞の抽出法として、テキストマイニング法、さらに源氏物語各章ごとの助動詞の類似性を見るための数量化Ⅲ類による解析とグループ化の指標としてのクラスタリングを導入することで、各章ごとの関連性について考察する方法を示した。その結果に男よると、源氏物語は紫の上系と玉蔓系に大きく分かれ両者には関連性が少ないことが示された。このことから、両系の作者は異なる可能性が高いことが示された。

第四章は、以上の源氏物語の作者について考察した結果をまとめた結論である。

最後に、源氏物語の執筆順序に関する文学的考察を各帖ごとにまとめた結果を付録として示した。

審査結果の要旨

本研究の内容を総括すると、源氏物語の作者について、その時代背景と文学的立場からの作者についての考察とコンピュータを用いた科学的立場からの統計処理を組み合わせた研究内容であるといえる。すなわち、源氏物語の作者を助動詞の用法に着目して考察するという文学的立場からの研究を行い、さらに、テキストマイニング、数量化Ⅲ類とクラスターリングといった科学的手法を導入することで、源氏物語の複数作者説を支持する新たな成果を得たことは文学上の大きな成果である。一方、コンピュータを用いて、文学作品の特徴抽出を短時間で可能にする技術の導入は、文学の世界ではこれまでにない画期的な成果でもある。これらの研究成果は、源氏物語の作者についての新たな考察を加えただけでなく、より広く古典と呼ばれる文学作品の特徴抽出法として広く応用可能であり、サイエンティフィックアートと呼ばれる広い意味での自然科学研究のさらなる発展に寄与する大きな成果にも繋がる。

以上のことから、本研究は博士(学術)の学位論文として十分であると認定した。