審 査 結 果 の 要 旨

（１）研究の目的に意義や独創性があるか。

本研究文は、語彙的結束性に関し、テキストに現れる語の出現量からテキスト一般およびに個別テキストの特徴を探索することをねらいとしている。結束性は文章を一つの統一体としてまとめ上げるために必要な性質の一つとされるが、関連領域である日本語学・国語学、言語学・言語教育研究、自然言語処理分野においては、指示・代用などに代表される文法的結束性に関する研究、個別あるいは小規模データの分析、語彙的結束性の情報収集の方法の研究が中心であった。それに対し、本研究は、大規模コーパスを利用し、語彙の計量的特性、テキスト中の語彙の分布を調べ、これらがテキストの構造にどのように関係しているかを、これまでに開発されてきた方法に加えオリジナルの分析方法を用いて探究している。さらに、従来、量的研究について指摘されてきた「計量結果のみ」という問題に対し、「量的現象から言語の性質を導く」という立場でアプローチを試みている。以上の点で、独自性の高い、意義のある研究であると思われる。

（２）研究の方法は当該学問分野において妥当なものか。

本研究は計量言語学、コーパス言語学分野に該当する研究であり、その方法論である大規模コーパスを利用した計量的方法を採っている。本研究で扱う情報は出現する語の数、語と語の関係を距離などで量的に計測した結果である。これらの情報を利用して、テキストの持つ普遍的な性質や個別的な性質の解明を試みており、目的にも合致する妥当な方法である。

（３）研究資料やデータの収集と分析が適切になされているか。

本研究が利用したコーパスは、2006 年から国立国語研究所を中心として開発された『現代日本語書き言葉均衡コーパス』（Balanced Corpus of Cotemporary Written Japanese, 略称 BCCWJ）である。データ規模が大きい（約１億語）ために事例の収集が容易であること、レジスター（言語変種）に多様性があるためにレジスター間の比較ができること、また公開されたコーパスであるために第三者による検証が可能であることを利用理由としている。事例収集では、BCCWJ の特性を十分に把握したうえで、各章の調査目的に応じて、使用するデータや、語の長短の単位選択を適切に行い、同一語彙の距離の数量化など独自の計量化の方法を用いている。それらの情報は、統計的手法による分析に加え、数量情報をグラフに表わすことにより、たとえば延べ語数と TTR(Type/ Token Ratio)の相関、品詞別の同一語彙の出現間隔、多義語の出現状況等と語彙的結束性の関係の分析が行われている。また、多義語の語義の出現状況や、共起語率の分布と語彙的結束性の関係をレジスター間で比較し考察するなど、現代日本語テキストの語彙的結束性を多角的に分析している点は高く評価できる。

（４）研究の考察と結論が妥当であり，学術的な水準に達しているか

計量言語学やコーパス言語学においても、量的把握と質的把握のバランスが重要だとされるが、本研究ではその点での配慮も見られ、語彙的結束性に関する総合的な考察が試みられてい

る。第２章「テキスト全体の計量的特性と語彙的結束性との関係」では、量的分析結果から一般的テキスト、およびレジスター別の語彙的結束性の傾向を考察している。その結果、平均使用度数の高いテキストと低いテキストには文体的な特徴があること、テキスト全体の TTR と品詞別の TTR の関係として、名詞類の相関が最も高く、普通名詞の使用に TTR が大きく左右されることを明らかにしている。第３章「テキスト中の語の分布と語彙的結束性」では量的分析に加え質的なアプローチを組み合わせて考察している。高頻度語の出現間隔はレジスターによって違いがあることを示し、多義語については１件ずつその語義を分類して計量化し、語義は一つの語義に偏りやすく（名詞の意味の場合は約 70-80%）、その偏りは出現間隔が短いほどおこりやすいと指摘している。第４章「テキストの構造と語彙的結束性」では個別のテキストを事例として取り上げ、量と質の両側面から考察を加えている。共起語率によってレジスター別の結束性を観察し、法律、白書、国会議事録等が高く、新聞、ベストセラー、雑誌等が低いことを示している。さらに、段落間の非対称的類似度を利用してテキストの結束性を調べ、隣接段落以外にも結束性の高い段落があることを明らかにしている。

計量的に得られた情報を中心としつつも、テキストに個別に分析した結果を含めて総合的に考察した結論は、学術的にも高い水準に達している。

（５）取得学位にふさわしい意義や成果が認められるか

本論文は、計量言語学・コーパス言語学の分野においても先端の研究であり、学位の取得に十分な水準に達しており、研究的にも大きな意義を有していると評価できる。さらに、本研究の結果は、計量言語学の学術的な発展に対し多様な可能性を拓くものになると考えられる。同時に、言語コーパスの開発にとっても、データ化における情報タグの項目の決定などに関して、多くの示唆があると思われる。

今後への期待としては、執筆者自身も課題として挙げていることであるが、質的研究の側面の充実であり、そのための方法論についての検討が求められる。また、他領域との共同等によって、研究成果の体系化や実践への応用等に向けた研究を展開することが期待される。

以上の点から、審査委員会は全員一致して、本研究が東京学芸大学大学院連合学校教育学研究科の博士（学術）学位授与に十分に相応しい研究であると評価した。