文章中で用いられる同一語句のくり返しの定量的研究

東北大学大学院文学研究科言語科学専攻国語学専攻分野
鯨井綾希

論文要旨

本論文は、日本語の文章中で用いられる同一語句のくり返しの諸相を、主に定量的な観点から分析した研究である。

本論文の第 1 部・序論では、文章中で用いられる同一語句のくり返しを研究する意義と研究する上での背景を述べるとともに、目下の研究課題を以下のように示した。

（1）a. 同一語句のくり返しという現象は文章中にどんな形でどのくらい現れるものなのか。

b. 同一語句のくり返しという現象は文章のどのような特徴に連動して増減するのか。

また、この研究課題に対し、以下のような解決方法を示した。

（2）（1a）を解決する研究方法：

『現代日本語書き言葉均衡コーパス』を利用して、多くの文章を分析対象に取り上げ、それぞれで同一語句のくり返しの量を種々の観点から計量し、平均的な出現量や出現量の振れ幅、形態ごとの出現の量的様相などを分析する。

（3）（1b）を解決する研究方法：

『現代日本語書き言葉均衡コーパス』に含まれる様々なレジスターや文章内における構造・展開の様相を把握できる言語要素（接続表現など）を利用して、それらと同一語句のくり返し量との連動関係を調査し、変動の様相を分析する。

第 2 部・本論では、（1）の研究課題を（2）および（3）の視座から考察し、研究課題の解決を試みた。第 3 部・補論では、本論の補足ではあるが、同一語句のくり返しと他の言語要素とを関連づけ、その相互関係について考察した。第 4 部・結論では、第 2 部と第 3 部の分析結果をまとめ、同一語句のくり返しがどのような文章上の特徴の中で使用されるのかを総合的な立場から明らかにした。特に第 2 部を総合的に捉え直した結果、文章中で用いられる同一語句のくり返しの特徴的な側面として、以下の点を指摘できた。

（4）名詞による同一語句のくり返しの出現量

a. 文章のまとまりごとにおおよそ 15%から 30%の割合で名詞のくり返しを行う。

b. ある程度の長さを持った文章では、同一語句のくり返しは必ず用いられる。

（4a）から、文章中においてはあくまでも新出の語句がその大勢を占め、それを補足する形で、くり返しが行われていることを指摘できる。また、（4b）から、同一語句のくり返しは必ずしも量的に多いものではないが、その使用は文章中で必ず見られるという特徴を持つことも分かる。

（5）同一語句のくり返しが用いられる特定の場面

a. 論の帰結（結論やまとめ）の文脈でくり返しが集中的に用いられる。

b. 対比の文脈でくり返しが集中的に用いられる。

（5）から、同一語句のくり返しが用いられる代表的な場面を明らかにできた。

（6）文章展開に伴う同一語句のくり返しと新たな見出し語の増減（4 段階）

a. 話題の開始時（第 1 段階）では、新たな見出し語は増加し、同一語句のくり返しは減少する。

b. 話題の前半部（第 2 段階）では、新たな見出し語は増加を続け、同一語句のくり返しは増加に転じる。

c. 話題の後半部（第 3 段階）では、新たな見出し語は増加を続け、同一語句のくり返しは減少に転じる。

d. 話題の終了時（第 4 段階）では、新たな見出し語は減少に転じ、同一語句のくり返しは増加に転じる。

（6）から、（5）のような特定の場面のみならず、一つの話題の展開に伴う同一語句のくり返しの用法の変化が明らかになった。

（7）同一語句のくり返しの割合が大きくなる場合の書き手の意識

読み手を意識して対他性を帯びる文脈：

a. 読み手に何らかの行動を促し影響を及ぼうそうとする意識を持つ場合。

b. 自らの意図を読み手に明確に伝えようとする意識を持つ場合。

（4）から（6）が、専らくり返しの文章中での現れ方を分析したものであるのに対し、（7）は、そのようなくり返しがどのような書き手の心理的変化に基づいて出現するのかを明らかにしている。

第 2 部が本論文における考察の軸となる研究であるのに対し、第 3 部は同一語句のくり返しと関わる周辺領域を研究対象として取り込み、それらと同一語句のくり返しとの関係からどのようなことが言えるかを検討する部分である。

まず、第 3 部の第 1 章では同一語句のくり返しに付く連体詞としての「この」「その」と、そうした形式が付かない「ゼロ」形式のくり返しとを比較し、それぞれの形式がどのような文構造・情報構造の下で用いられるのかを分析し、それぞれの差異を記述した。それにより、「この／その／ゼロ」の各形式が付加されて行われる名詞のくり返しが、様々な使用上の偏りを持っていることを明らかにした。

第 3 部の第 2 章では、名詞句のくり返しと互換性を持ち得る「それ」という指示代名詞の使用規則を検討することを通して、くり返しを使う場面と「それ」を使う場面の差異化を試みた。ここで明らかにした「それ」の使用規則は、「それ」に先行詞となる名詞句が正しく格納され得る条件であるため、この規則から逸脱したり、先行詞が確定不能な状況や先行詞に曖昧性の生じる状況が出たりした場合には、文章中では必然的に同一語句のくり返しが用いられると考えられる。

第 3 部の第 3 章では、同一語句のくり返しの使用状況から見て、接続表現の特性の違いを論じた。同一語句のくり返しに注目して接続表現を分析することで、接続表現単独の分析では浮かび上がらせることのできない側面を見出し、くり返しと接続表現とを結びつけて論じることが可能であることを示すことができた。

論文審査結果の要旨および担当者

提　　出　　者	鯨井　綾希
論文審査担当者	（主査）教授　　齋藤　倫明 教授　　小林　隆 教授　　後藤　斉 准教授　大木　一夫 准教授　甲田　直美
論　　文　　名	文章中で用いられる同一語句のくり返しの定量的研究

本論文は、電子化コーパスと、コーパスを処理するコンピュータを用いて、日本語の文章における同一語句のくり返しの定量的分析を行ない、その実態と機能を明らかにすることを目指したものである。

具体的には、『現代日本語書き言葉均衡コーパス』に見られる大量の文章を分析対象として取り上げ、それらにおける同一語句のくり返しの量を計量し、平均的な出現量や出現量の振れ幅、形態ごとの出現の量的様相などを分析した。また、同コーパスに含まれる様々なレジスターや文章における構造・展開の様相を把握できる言語要素（接続表現など）を利用し、それらと同一語句のくり返し量との連動関係を調査・分析した。

その結果、くり返しの割合の多寡に対応する文章上の特徴として、論の帰結（結論やまとめ）・対比の文脈でくり返しが集中的に用いられることが明らかになった。また、文章展開におけるくり返しに関し、一つの話題については、話題の開始時・話題の前半部・話題の後半部・話題の終了時において、新たな見出し語の増減とくり返しの増減に一定の傾向が見られることが判明した。さらに、ひとつひとつの文章内での語の追加に伴う内容展開に注目し、名詞が一語追加される度に Type-Token Ratio を計測することで、その変動の在り方と文章展開の対応関係を明らかにした。その他、対他性の高い文脈においては同一語句のくり返しの割合が大きくなるといった、書き手の意識の側面についても言及した。

文章のジャンルとくり返しの関係については、くり返しが非常に多い文章として法律文書やその手引書、非常に少ない文章として一人称を用いて書かれた文章が挙げられることが指摘されている。また、書き言葉と話し言葉におけるくり返しの量的側面を比較することで、媒体の性質がどのようにくり返しの使用に作用するのかが詳細に分析されている。

本論文は、統計的手法を駆使することによって、従来、あまり明らかにされてこなかった同一語句のくり返しという言語的事象の全体像を定量的に明らかにし、日本語の文章構造、文脈展開における同一語句のくり返しの機能の解明に貢献している。その成果は、日本語文章論・計量国語学の今後の発展に大きく寄与するものとして高く評価できる。

よって、本論文の提出者は、博士（文学）の学位を授与されるに十分な資格を有するものと認められる。