

## 対象年齢が異なる百科事典の文章構造の特徴分析

橋本 舞子

本研究では、対象年齢が異なる百科事典の文章の特徴を、文章の複雑さを表す指標を用いて分析する。文章の書き手は、対象とする読み手の年齢によって書き方を変え、それは文構造に表れる。小宮は小学校から大学までの4段階の日本史の教科書を対象に文の長さ、1文中の述語数、主語と述語の距離、主語の有無、係りの次数という5つの指標について、どのように変化していくかを分析した。しかし小学校から大学まで段階的に変化するという仮説が証明できなかった指標もあった。その原因は対象文数が少ないこと、記述内容に偏りがあること、教科書検定を経ている小中高校の教科書と、大学で使用される教科書を対象にしたことが考えられる。そこで、本研究では専門家が執筆し、出版者が編集しており、記述されている内容に偏りがなく、文章量も教科書に比べて多い百科事典を対象にし、その違いを分析することとした。

分析の対象として、小学校高学年から中学生向けのエンカルタ総合大百科 2007 と、小学校低学年向けのエンカルタキッズ百科 2007 から共通する項目を無作為に17個ずつ選択した。文章を複雑にする構文を基に、指標 1文あたりの主述の組の数、主述の組に対する述語数の割合、動詞が文頭に来て、形容語になっている文節の数、1文あたりの主述の組ごとの主語有の数、1文あたりの挿入句の数、挿入句の文節数、主述間の距離、係りの次数、主語に係る次数、1文あたりの受身の文節数、1文あたりの文節数、の11個を設定した。百科事典の文章を解析ソフトJUMANとKNPを用いて形態素解析、構文解析し、指標を抽出し、2つの百科事典の指標に差があるかを検証するため、t検定を行った。

その結果、11個の指標のうち、 $t_{11}$ 、 $t_{12}$ 、 $t_{13}$  については高学年向けの方が多くことが分かった。一方、 $t_{14}$ 、 $t_{15}$ 、 $t_{16}$  に有意差が見られず、 $t_{17}$  は、低学年向けの方が多くなるという結果が得られた。主述の組に対する述語数の割合、係りの次数は、指標の値と文章の分かりやすさは必ずしも対応していないという考察が得られた。また、1文あたりの挿入句の数は、有意差は見られるが、文構造の複雑さによるものではなく、文の内容によるものであるという考察が得られた。主語に係る次数は、仮説と異なる結果になったが、主語に多くの修飾を付けた文の方が、低学年には分かりやすい文章である可能性がある、という考察が得られた。以上の考察から、対象とする読み手の年齢による文章構造の違いと、文構造の複雑さは一致しない部分もある、と言える。

今回、エンカルタ総合大百科 2007、エンカルタキッズ 2007 を使用したが、比較する年齢に差があまりなかった。そこで、今後の課題として、中学生以上、大人向けの百科事典と比較することで、より幅広い、年齢層の変化による指標の変化を分析が考えられる。

(指導教員 中山伸一)