

博士論文要旨

論文題名：外国ルーツ高校生を含む『高校生日本語作文コーパス』の構築と計量的研究

立命館大学大学院文学研究科
人文学専攻博士課程後期課程

マツモト サトミ

松本 理美

本論文は、2部7章で構成される。まず、序章で研究の背景、目的、意義について述べる。I部は『高校生日本語作文コーパス』の設計と構築に関する2章から成る。第1章では『高校生日本語作文コーパス』の設計と構築の基本方針について説明し、第2章ではコーパス構築の過程で必要となる接続節のアノテーション基準の策定について詳述する。II部は『高校生日本語作文コーパス』の利用に関する3章から成る。第3章では本コーパスを利用し、高校生作文と国語教科書の複文構造について詳細な計量分析を行う。第4章では国語教科書の連体節構造についての分析を行い、第5章では高校生作文と国語教科書における連用節のテ形節に着目した分析を行い、高校生作文と国語教科書における接続節の使用実態と複文構造の特徴を明らかにする。最後に終章で本研究をまとめる。

本研究の目的は、i) 日本の高校に在籍し、多様な日本語習得環境にある高校生の日本語作文を収集し、外国ルーツ高校生を含む『高校生日本語作文コーパス』を構築すること、ii) 構築したコーパスを用いて、高校生作文と国語教科書における日本語（特に複文）に関する計量分析を行い、多様な高校生の日本語作文と国語教科書における接続節の使用実態および日本語の特徴を明らかにすることである。

日本の高校に通う高校生は日本語の習得環境により、大きく3つのグループに分けることが出来る。1 つめは、日本で生まれ育ち、親、本人ともに日本語を母語として習得した日本人高校生のグループである。2 つめは、海外で生まれ育ち、親、本人ともに共通の外国語を母語とし、自分の意志で日本に留学してきた留学生のグループであり、来日前後に体系的、段階的な日本語教育を受けている。そして3つめは、親、本人の国籍・母語、生まれ育った環境などは様々で、親の仕事や結婚などにより、親に伴って来日した高校生（以下、外国ルーツ高校生）のグループである。外国ルーツ高校生の日本語教育環境には、大きな個人差・地域差があり、その4割程度は「日本語指導が必要な生徒」である。

このような3つのグループの高校生を対象とした日本語作文データの構築や研究は希少で、高校生作文を対象とした日本語の使用実態調査はほとんど行われていないのが現状である。

本研究では、これら3つのグループの高校生作文と国語教科書を対象に、文法的視点からも

分析が可能となる『高校生日本語作文コーパス』を構築する。コーパス構築の過程で接続節情報を付与するにあたり、接続節の用法分類基準を策定する。また、このコーパスを利用して高校生作文と国語教科書の複文に着目した計量分析を行う。

本論文では、I部で『高校生日本語作文コーパス』の構築と設計について詳述する。本コーパスは、日本の高校に通う外国ルーツ高校生（主に日本語指導が必要な高校生）、留学生、日本人高校生という、日本語習得環境の異なる高校生が書いた日本語作文（全282編）を収録した高校生作文コーパス（メインコーパス）と、小学校・中学校・高校の国語教科書の著作物（全22編）を収録した国語教科書コーパス（サブコーパス）により構成される。

高校生作文コーパスは、調査協力校より収集した、テーマ、作成の条件や環境などをすべて統一した作文に、形態論情報や接続節情報を付与したデータを主たるデータとする多様なデータにより構成される。また、国語教科書コーパスは、説明文・論説文を主とした著作物に高校生作文と同様のアノテーションを施したデータなどにより構成される。

コーパス構築に際し、すべてのデータに、自動解析が可能な形態論情報のみならず、接続節情報を手作業により付与した。接続節情報を付与する用法分類基準は、英語―日本語自動翻訳のために構築された「鳥バンク」節間意味分類基準を、日本語の接続節の分類のみに特化して修正することにより策定した。この基準策定により、接続節アノテーションの精度と、本コーパスを用いた接続節の研究の信頼性を保つことができる。

II部で、本コーパスを利用した高校生作文と国語教科書の日本語に関する調査、分析について、詳しく述べる。

高校生の日本語作文における複文構造の分析により、多様な高校生の日本語リテラシー（主に作文力）の実態について文法的な視点で論じた。1文あたりの文字数、接続節数は日本人高校生が最も多いが、語彙の豊富さに関しては必ずしも日本人高校生が高いとは言えず、作文の内容に依存する可能性が示唆される。また、高校生作文には連体節の出現割合が低く、連用節が多用されていること、連用節においてテ形節の連鎖が多い外国ルーツ高校生は話しことばのような文体で作文を書いていることなどが明らかになった。

国語教科書における複文率はどの学校種においても総じて高いが、複文構造には学校種による特徴が見られた。連体節では、学校種の上昇とともに補足語修飾節の割合が減少し、内容補充節、相対名詞節の割合が増加することがわかった。また、学校種の上昇に伴い減少する傾向にある連用節では、条件・譲歩節の割合が、高校生作文より高く、学校種の上昇に伴い出現割合も上昇するという調査結果を得た。

本研究で構築した『高校生日本語作文コーパス』を用いた計量分析では、多くの調査項目で、高校生作文のグループ間の差、国語教科書の学校種間の差、高校生作文と国語教科書の差が確認された。本研究により、高校生作文や国語教科書における日本語の使用実態の調査結果が作文教育、国語教育、日本語教育で有効に利用できること、複文が文章の文体的特徴や文章レベルを捉える指標のひとつとなり得ることが示された。