

定例研究会要旨

日時：平成 29（2017）年 5 月 24 日 18:00～20:00

会場：東京外国語大学 語学研究所

「ヨーロッパにおける日本語学習者の日本語作文のテキストマイニング
—その言語的特徴に関するパイロットスタディー—」

"Analyses by Text-Mining of Essays Written by Learners of Japanese Language in Europe:
A Pilot Study on Their Linguistics Characteristics"

発表者：阿部 新（東京外国語大学大学院国際日本学研究院准教授 / 日本語教育，日本語学）
ABE Shin

本発表では、ヨーロッパにおける日本語学習者の日本語作文の言語的特徴をテキストマイニングによって探索した。

本研究では、科研費による研究プロジェクト「日本語ライティング評価の支援ツール開発：「人間」と「機械」による評価の統合的活用」（研究課題番号：26284074）によって2014年11月から2016年3月にわたって収集された、ヨーロッパ10か国（ベルギー、クロアチア、フランス、ドイツ、ハンガリー、イタリア、ロシア、セルビア、スロベニア、スペイン）とアメリカ合衆国の日本語学習者の作文データを分析した。作文のプロンプトは5種類あったが、それらのうち、本研究では比較対照と論証を行うA1、A2の2種類のプロンプトに対して書かれたもの（A1:136編，A2:153編）を分析対象とした。

分析にはKH Coderというフリーソフトを使用した。データ中から形態素解析によって語を自動的に抽出し、執筆者の個人的情報（性別、日本での学習歴、日本語能力試験受験経験、日本語学習歴、日本語能力）や、作文に関する経験（母語でのアカデミックライティングの受講経験、今回の執筆での注意点）と共起する語を分析することで、本研究で分析対象とするデータの特徴を探索した。

その結果、まず、データ全体および1作文あたりの抽出語数と異なり語数は、以下の表1のような結果となった。

	プロンプト A1	プロンプト A2
総抽出語数	50,610	56,325
異なり語数	2,777	3,013
作文数	136	153
1作文あたり抽出語数	372.13	368.13
1作文あたり異なり語数	20.42	19.69

また、共起ネットワークによる分析の結果、名詞・形容詞・動詞・形容動詞などの内容語については、日本語能力、学習歴の長さ、学習経験の多様な学習者のほうが多様な語を使用している傾向が見られ、また、性別による違いがいくつか見られた。また、助詞・助動詞などの機能語については、日本語学習歴が長い場合に多様な語の使用が見られたものの、日本語学習歴が短い学習者、日本での学習経験のない学習者、日本語能力試験の受験経験のない学習者は、「よ」「ね」といった助詞の使用が多いという結果が示された。こういった、「よ」「ね」が文章の文末で使用される場合については、文章の書き方の指導が必要となるであろうということが示唆された。