

中国における日本語学習者の研究論文の評価に関する研究

—Analytic scoring 評価に基づいて—

(要約)

広島大学大学院文学研究科

博士課程後期人文学専攻

学生番号：D193511

氏名：高帥

本研究は、中国語を母語とする日本語学習者の研究論文に対する従来の評価基準を検討し、複数の調査によって新たな評価基準（「新評価基準」）を提示した。さらに、従来の評価基準と比較し、信頼性、妥当性と評価者間の評価一致度の観点から新評価基準の優位性を検証した。

国際交流基金の2018年度日本語教育機関調査では、世界の日本語学習者数は3,851,774人で、学習者数が最も多いのは26.1%を占める中国であると報告されている。1998年度の同じ調査と比べ、中国の日本語学習者数は20年間で4倍に増加し、2012年度調査から中国は世界で最も日本語学習者の多い国となっている。

日本語教育の急速な普及と発展の中で、中国においても日本語・日本文化を専門とする学科を有する高等教育機関が急増し、学習者が日本語でレポートや論文を書く機会が多くなり、その評価研究も盛んになっていった。このような背景の下、中国の「四年制大学各専攻教育の国家スタンダード」（普通高等学校本科專業類教学質量国家標準）（以下、「新国標」）と「四年制大学外国言語文学類専攻教育指南」（普通高等学校本科外語語言文学類專業教学指南）（以下、「指南」）に沿った、日本語で書かれた研究論文の評価基準の開発が期待された。そこで、本研究は、第二言語として日本語を学んでいる中国語を母語とする日本語学習者の研究論文を評価するための、信頼性と妥当性の高い評価基準の作成を目指した。100名以上の日本語教師の協力を得て、5件の調査を行い、その結果に基づいて、新たな Analytic scoring 評価基準「日本語研究論文評価基準」（「新評価基準」）を開発した。

本研究は、以下の6章から構成されている。

第1章「序論」では、本研究の背景、目的と構成を述べた。

第2章「先行研究と本研究の課題」では、日中における日本語学習者のライティング評価研究に関する先行研究を概観し、先行研究が残した課題を問題点として3点にまとめて論じた。

1つ目は、評価の研究では、評価のプロセスや留意点に関する研究が多く、具体的な評価基準が示されている研究はほとんどないこと。

2つ目は、試験や課題における作文を対象とする評価の研究はあるが、対象を日本語の研究論文に特化した評価の研究は極めて少ないこと。

3つ目は、論文の評価には、評価者のライティング教育経験、背景知識などによる個人差が影響する可能性があるため、評価者間の評価一致度を高めることが求められるが、一致度に関する言及や調査が極めて少ないこと。

以上の先行研究の課題を踏まえ、本研究では、以下のように研究課題を設定した。

課題I. 研究論文を対象に、Holistic scoring (以下、HS) 評価法と Analytic scoring (以下、AS) 評価法を調査によって比較し、以下の観点からどちらが優位なのかを明らかにした。

I-1 HS 評価と AS 評価、どちらが信頼性が高いか。

I-2 HS 評価と AS 評価、どちらが評価者間の評価一致度が高いか。

課題II. 第二言語としての日本語研究論文評価における優位性の高い評価方法に対応する評価基準はどのように構築するのか。

II-1 中国教育部により公表された「新国標」と「指南」に適応する、日本語による研究論文の評価基準はどのように構築するのか。

II-2 その評価基準の信頼性と妥当性はどのように検証するのか。

課題III. 課題IIで構築した新評価基準の実用性はどのように検証するのか。

III-1 大学で現行の評価基準と新基準とではどちらが信頼性が高いか。

III-2 大学で現行の評価基準と新基準とではどちらが評価者間の評価一致度が高いか。

第3章「Holistic scoring 評価と Analytic scoring 評価の比較研究」では、課題Iの達成を目的とした調査1を行い、その結果をまとめた。まず、研究論文のためのHS評価基準とAS評価基準を設定した。次に、中国人日本語教師(9名)が論文(2編)を上記2種類の評価基準を用いて別々に評価した(調査1-1)。その結果により、AS評価基準の評価項目全体は内的一貫性があり、AS評価基準の信頼性が高いことを検証した。(課題I-1)

また、評価者間の評価一致度を比較するために、中国人日本語教師(8名)が論文(5編)を上記2種類の評価基準を用いて別々に評価した(調査1-2)。得られた結果により、変動係数とケンドールの一致係数を計算し、AS評価はHS評価より評価者間の評価一致度が高いことを検証した。(課題I-2)

課題Iの調査結果から、第二言語としての日本語研究論文を評価するには、全体的印象を基に単一の得点でライティングの評価をするHS評価より、評価

基準や観点を詳細に示す AS 評価の方の優位性が高いことが明らかになった。  
(課題I)

第4章「日本語学習者の研究論文の Analytic scoring 評価基準の開発」では、課題IIの達成を目的とした調査2を実施した結果をまとめた。まず、「新国標」と「指南」に示されている人材育成に求められた能力を基に、AS 評価基準の構成概念を得た。(課題II-1)

10年以上の日本語教育歴があり、かつ5年以上の日本語論文指導歴を持つ中国人日本語教師(18名)を選出し、デルファイ法を用いて「新評価基準」原案を作成した(調査2-1)。高い信頼性と妥当性を備える評価基準を作成するために、中国人日本語教師(128名)に「新評価基準」原案の各評価小項目の必要度を5段階評価するアンケート調査を行った(調査2-2)。調査結果を用いてクロンバック $\alpha$ 係数を計算し、「新評価基準」原案の評価項目の信頼性が高いことが明らかになった。また、探索的因子分析(主因子分析法、プロマックス回転)を行い、5因子を持っている因子構造が確認できた。さらに、評価項目の信頼性を保証した上で、探索的因子分析の結果に基づく仮説モデルを示した。最後に、検証的因子分析を行い、仮説モデルを最も適合するモデルに修正し、修正した「新評価基準」原案の構成概念妥当性が検証できた。修正した「新評価基準」原案に「評価大項目」「段位」「得点」と「成績」の欄を追加改良し、最終的に「新評価基準」とする「日本語研究論文評価基準」を作成した。(課題II-2)

課題IIの調査結果から、この評価基準が第二言語としての日本語研究論文評価における優位性の高い評価方法であり、かつ信頼性と妥当性が兼ね備え、「新国標」と「指南」に適応する評価基準であることを検証した。(課題II)

第5章「『新評価基準』と『現行基準の比較研究』」では、課題IIIの達成を目的として実施した調査3の結果をまとめた。日本語を第二言語とする日本語学習者の研究論文を対象に、ある大学が採用している「現行論文評価基準」と「新評価基準」を比較する次の調査を行った。中国人日本語教師(16名)が評価対象の論文(4編)を「現行論文評価基準」と「新評価基準」を用いて別々に評価し、結果により、両者の評価項目全体のクロンバック $\alpha$ 係数とケンドールの一致係数を計算した。(課題III-1、課題III-2)

課題IIIの調査結果から、課題IIで構築した「新評価基準」は本研究に用いた

「現行論文評価基準」より、信頼性も評価者間の評価一致度も高いこと、さらに、実用性があることが検証された。（課題Ⅲ）

第6章「結論」では、本論内容のまとめと今後の課題と展望について述べた。本研究で言及しきれなかった、①変動係数の測定により評価者間の評価一致度を分析すること、②「現行論文評価基準」と「新評価基準」に対する評価者の内省に関する調査、③「新評価基準」の日本語教育への応用および「新評価基準」の自動評価システムの開発についても今後の課題として明記した。

本研究の意義は、以下の4点である。

1. 中国高等教育における日本語研究論文を評価の要点を明らかにした上で、高い信頼性・妥当性がある、日本語研究論文の評価基準を作成した。

この評価基準には、中国の「新国標」と「指南」に求められた人材育成の能力を踏まえ、日本語研究論文に特化した、具体的な評価項目が示されている。本研究の成果は、中国における日本語研究論文に関する評価基準をはじめて提示した。

2. 「新評価基準」と「現行論文評価基準」を用いて、日本語教師による評価を実施し、評価者の評価一致度の観点から「新評価基準」の優位性が検証されたので、今後、現行論文評価基準に代わって利用可能である。。

複数の評価者がいた場合でも、「新評価基準」は評価者間の評価一致度がより高く、客観性があることが明らかになったことから、広い範囲で利用可能である。

3. 本研究で示した「新評価基準」は、中国における現場の日本語教師にとって、日本語による研究論文や作文の指導の際の指標として、教育にも応用できる。

「新評価基準」は評価基準や観点を詳細に示すAS評価の具体的な評価項目が示されているので、評価者間でこれらの項目を用い「いい日本語研究論文はどのようなものか」の共通認識を持てば、評価者の個人的要因による評価の不一致の程度が減少し、評価者間の評価の一致度を向上させることができる。また、評価者が論文を評価することだけでなく、教師が論文作成を指導することにも、論文作成者自身が論文を修正することにも広く利用できる。「新評価基準」はその評価項目が日本語研究論文や作文の作成指導の指標となるので、ライティング教育において、教師が学習者全員に対して授業内容を明確にでき、

評価得点によって個別指導のポイントも明確にできる。この具体的な指標によりにより、日本語学習者の研究論文や作文の作成能力を向上させることが期待できる。

4. 第二言語による研究論文や作文の評価基準の作成過程のモデルとして、日本語以外でも利用でき、汎用性が高い。

本研究では、中国国内に限定することを明記したとはいえ、中国以外のところで求められる能力が中国と異なる場合でも、評価項目の選出方法は汎用的に利用でき、基本的な評価項目導出法の研究として有効と考える。世界各国の外国語研究論文評価基準の作成および活用のために必要な資料を提供できたといえよう。