次世代の日本語教員のための教育評価
―日本の教育評価研究における理論的考察―
Assessment for the Next Generation of Japanese Language Teachers

安達万里江, 関西学院大学, 京都外国語大学大学院（博士後期課程）
Marie Adachi, Kwansei Gakuin University, Kyoto University of Foreign Studies

1. はじめに
1.1 本研究の背景

本研究における日本語教育現場は、日本の高等教育機関である（以下、大学）。

2018 年 11 月に発表された文部科学省の諮問機関における中央教育審議会「2040 年に向けた高等教育のグランドデザイン（答申）」（以下、答申）では、1954 年以降、42 の答申の中で、今後 20 数年先を見据え、各大学における「実現すべき方向性」が挙げられている。その中でも本研究では、次の２点に着目する。

（1）18 歳人口は、2040 年には、88 万人に減少し、現在の７割程度の規模となる推計が出されていることを前提に、各機関における教育の質の維持向上という観点からの規模の適正化を図ったうえで、社会人および留学生の受け入れ拡大が図られていくこと。
（2）高等教育機関がその多様なミッションに基づき、学修者が「何を学び、身に付けることができるのか」を明確にし、学修の成果を学修者が実感できる教育を行っていること。このための多様で柔軟な教育研究体制が各高等教育機関に準備され、このような教育が行われていることを確認できる質の保証の在り方へ転換されていくこと。

以上のように、少子化による学生数の減少傾向や、国際社会による留学生数の増加といった様々な要因が含まれる中、教育（学習／学修）の「質の保証」が求められ、「確認できる」ことが挙げられる。

そのため、大学教員に「何を教えたか」だけでなく、「何を学び、身に付けることができたか」に関して説明を求める大学が増えつつある。留学生を対象に日本語教育を行っている日本語教員も例外ではない。では、今日の日本語教育学界において、教育に関する研究の現状はどうだろうか。

本田他（2019）では、学会誌『日本語教育』101 号から 170 号に掲載された論文 280 本のうち、「教育と社会」に関する論文 48 本より分析している。

その結果、「誌名に『教育』とあるにもかかわらず、『教育と社会』に分類された論文は全体の 17%」であった（p.47）。また、「『日本語の授業』で実施する形成的評価や学習者主動型評価などをあつかった研究は現在も遅れていると言わざるを得ないだろう」と指摘している（p.52）。そこで筆者は、今日の日本語教育学界において、教育面でも特に、日本語の授業で実施する評価をあつかった研究、つまり、「評価研究」が喫緊の課題であると考えた。

1.2　本研究の目的

そこで本研究では、大学における「評価研究」に関し、まず、教育学の先行研究より文献レビューを行い、現状を概観する。特に、教育学の中でも「形成的評価」や「学習者主動型評価」などに関する理論について取り扱っている「教育方法学」の「教育評価論」に注目する。

次に、外国語・第二言語教育の流れを含む日本語教育における評価研究より、「日本の大学」「パフォーマンス評価」「アカデミック・ライティング」に焦点を絞り、「日本語ライティングの評価研究」関連の研究成果を取り上げる。そして、次世代の日本語教員にとって必要な評価研究とは何かを考察する。最後に、実践的かつ応用可能な評価活動や教育評価研究の一例を示すことを目的とする。

2.　文献レビュー

2.1　教育学における評価研究

「形成的評価」や「学習者主動型評価」などに関する理論は、教育学の中でも「教育方法学」の「教育評価論」で論じられている。そこでの形成的評価とは、指導の途中で行われる評価のことで、近年では「学習のための評価」として再定義する動向が見られる。また、学習者主導型評価とは、学習者自身が評価活動へ参加することであり、評価活動自体を学習の機会として捉える「学習としての評価」という理論的展開を踏まえつつ、教育者が行う指導と評価の充実を図っていくことが求められている（西岡他, 2015 p.7）。ここからわかることは、本研究で定義すべき教育評価とは、「授業・学習評価」であると言える。続いて、これらを含めた日本における教育評価研究にはどのような流れがあるかについて、歴史的概略について述べる。

2.1.1　日本における教育評価研究の歴史（戦後～現代）

日本における教育評価研究の理論は欧米からのものが多い。今日でも見られる標準テストは、戦前の頃より教育測定（メジャメント）の要素が強く、「評価＝ネブミ」という結果となっていた。第二次世界大戦直後の日本において、アメリカより「エバリュエーション（Evaluation）」が紹介され、「教育評価」と翻訳されることになった（田中, 2008 p.33）。その概念の創始者であるタイラー（Tyler, 1949）によれば、エバリュエーションの目的は、「ネブミ」をすることではなく、評価行為を通じて得られた情報をもとにして、教育活動を吟味・改善することにあると考えられていた。そして、問題の核心は「どのように測定・評価するのか」の前提に、「何のために何を測定・評価するのか」という教育目標論（到達度評価）へと追究がなされていく（田中, 2008 p.34）。

その後、ドメイン準拠評価とスタンダード準拠評価という２つの方法論を含み持つ「目標に準拠した評価」が 2001 年度の指導要録改訂を機に導入された（田中, 2010）。他方、アメリカでは 1990 年代よりウィギンズ（Wiggins, 1998）が評価の文脈で「オーセンティック（真正）」という概念を使用し始めた。教育評価において、「実社会」「生活」「リアルな課題」が強調され、「真正」な課題に取り組ませ、その様相を評価しなくてはならないと考えられるようになった（田

中, 2008 p.73）。そして、「真正の評価」を代表する新しい評価方法として、パフォーマンス評価とポートフォリオ評価がある。それらの理論や実践が Hart（1994）によって『パフォーマンス評価入門』としてまとめられ、21 世紀に入り、田中耕治氏の監訳（2012）によって日本にも広まっていった。

西岡他（2015）によると、パフォーマンス評価とは、知識やスキルを使いこなすことを求める問題や課題などへの取り組みを通して評価する評価方法の総称である。また、ルーブリックとは、パフォーマンス評価のツールであり、ポートフォリオ評価とは、教育・学習活動を評価するためのアプローチであると紹介されている（p.10）。日本でのパフォーマンス評価に関する研究は、初中等教育が先行している（松下, 2012）。では、日本の大学（高等教育）の現状はどうだろうか。

2.1.2　高等教育での評価研究

日本の大学におけるパフォーマンス評価に関する代表的な研究には、松下（2012, 2016, 2017）が挙げられる。そこでは、学習評価を「間接評価と直接評価」と「量的評価と質的評価」の 2 軸からなるに二次元図式によって 4 つのタイプに整理している（図 1）。

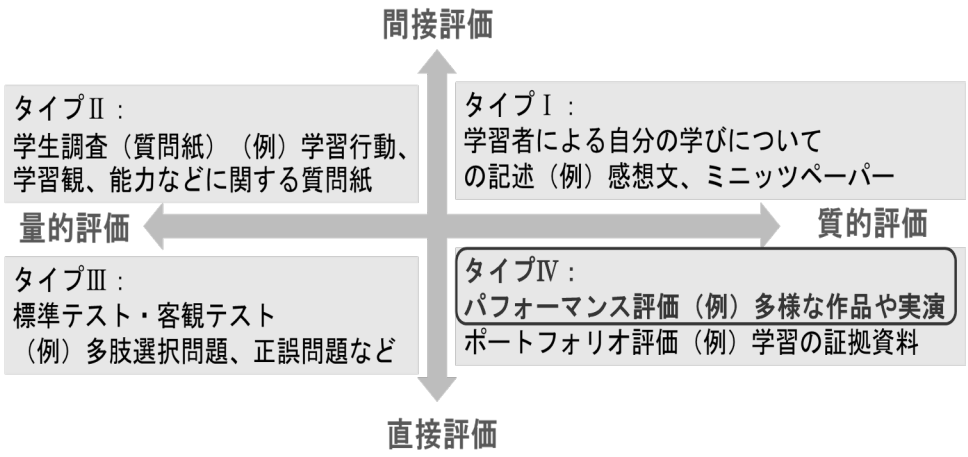

図 1　学習評価の構図（松下 2012, 2016, 2017 をもとに作成）

従来、高校までは図 1 にあるタイプⅢの標準（客観）テストが最も用いられている。他方、日本の大学における授業評価は、タイプⅡによる学生調査と、本研究で取り上げるタイプⅣによるパフォーマンス評価が一般的である。ここでは、松下（2012, 2016, 2017）以降、タイプⅣの最新の研究動向として、斎藤（2019）を挙げる。

斎藤（2019）は、日本人大学生のレポートのパフォーマンス評価にルーブリックを用い、教員と学生に評定（数値化）させ、学生たちには教員評価とのズレに対し、なぜズレが生じたのか、理由を自由記述させている。その自由記述をテキ

ストマイニングによって分析を行っていた。その結果、「質的・直接評価であるパフォーマンス評価（ルーブリックを含む）は、量的指標化しても統計分析に耐えられる信頼性を担保し、かつ、学生の学びのために活用できることを実証的に示せる」と述べている。

続いて、日本語教育の評価研究より「日本の大学」「パフォーマンス評価」「アカデミック・ライティング」に焦点を絞り、「日本語ライティングの評価研究」関連の研究を概観する。そして、斎藤（2019）の研究結果を踏まえた上で、将来的に日本語教育に還元できる研究手法について考察し、実践的かつ応用可能な評価活動や教育評価研究の一例を示したい。

2.2 日本語教育における評価研究

現在、日本語教育おける評価研究は多種多様にあるが、「日本の大学」「パフォーマンス評価」「アカデミック・ライティング」に焦点を絞り、「日本語ライティングの評価研究」関連の研究を概観する。そのため、安達（2019）では以下の条件に絞り、文献レビューを行った（表 1）。尚、本セクションは安達（2019）より加筆・修正を行った。

表 1　文献の選定基準（安達, 2019 より抜粋）

上位条件
日本語教育学界において、アカデミック・ライティングの評価研究を組織的、継続的に行い、かつ汎用性のある評価基準を作成している。
下位条件
1）評価の対象者が「日本の大学の学部に在籍する留学生」であり、評価者が「日本語教員」である。 2）評価項目ごとの評価観点について調査・分析を試みている。 3）評価者（日本語教員）だけでなく、被評価者（日本語学習者）にも直接還元できそうな研究である。

表 1 で上位条件として挙げた文献を①、下位条件として挙げた文献を②とし、下記に示す。本稿では文献レビュー①・②とし、計 19 本取り上げる。

文献レビュー①15 本：田中他（1998）～田中（2017）
文献レビュー②４本：伊集院（2017）、伊集院他（2018）、脇田（2016, 2017）

留学生を対象とした「アカデミック・ライティング」の評価研究を概観すると、大きな流れは田中他（1998）がはじまりであることがわかる。そして、日本の大学に在籍する留学生を対象に、20 数年にわたり日本語のアカデミック・ライティングの評価研究を行ってきた田中真理氏の研究論文（共著も含む）の研究動向を年表にまとめ、図 2 に示す。

年	研究動向
1998～	**日本語ライティング教育における評価研究のはじまり** ・４因子の抽出：「正確さ」「構成・形式」「内容」「豊かさ」（田中他，1998a） ・「客観的に図るための基準となるガイドラインの作成」が課題（田中他，1998b）
2004～ 2006	**評価基準・評価法の策定** ・公開されている日本語・英語の評価基準のレビュー（田中・長阪，2004） ・「マルチプル・トレイト評価法」の採用を提案（田中，2005） ・評価基準作成過程の公開とガイドラインの作成・改良（田中・長阪，2006）
2009～ 2011	**評価基準・評価法の検証** ・ワークショップによる評価基準の検討⇒課題「評価のずれ」（田中他，2009） ・評価者の意識を変える⇒課題は評価の不一致の究明（田中・長阪，2009） ・プロトコル分析⇒評価のプロセスは複雑であることがわかる（田中・坪根，2011）
2014～ 2016	**テキスト完成と「評価一致」に向けた研究** ・『Good Writingへのパスポート　読み手と構成を意識した日本語ライティング』発表（田中・阿部，2014）対象は留学生＋日本人大学生・大学院生 ・「評価一致」に向けた研究①「言語」に焦点化した訂正（田中，2015） ・クラスター分析等⇒評価者「内容面」「構成面」異なる基準（坪根・田中，2015） ・評価の不一致をもたらす要因（田中，2016a）とその対応策の考察（阿部・田中，2016）
2017～ 2018 （現在）	**近年の研究動向** ・「評価一致」に向けた研究②「構成」に焦点（田中，2016ｂ　田中他，2017） ・「日本語ライティング評価支援ツール」（自動採点）の開発（2018年HP上で公開）

図２　文献レビュー①の研究動向に関する年表（安達, 2019 より作成）

文献レビュー①のいくつかの論文は、多くの日本語教育研究者によって取り上げられおり、文献レビュー②伊集院（2017）、伊集院他（2018）、脇田（2016, 2017）も同様である。

上記①②の文献レビューによって明らかになった共通点として、評価が高くなる要因も、評価がずれる要因も「内容面」ということであった。しかし、内容面に焦点をしぼった先行研究は見当たらなかった。そのため、今後は内容面での量的・質的な調査・分析が日本語教育の評価研究として必要であると考える。そこで、高等教育・日本語教育から見る実践的かつ応用可能な評価活動や教育評価研究の手法をまとめる。

2.3　評価研究の手法

まず、評価研究においても、評価を数値化し、高・中・低位群にわけ特徴を見ることや、評価（評定）を決定づけた要因が何かを探るため、自由記述（アンケート）調査やインタビュー調査を行っている。

脇田（2016）は、ルーブリックを数値化し検証している。そこでは、学習者の自己評価が、教員による評価と一致するところが多かったとも述べている（脇田, 2017）。他方、伊集院（2017）は、上述にもある田中他（1998）等を先行研究として取り上げ、踏襲し、因子分析を行っている。また、評価基準のずれに関しては、近年急速に整備が進んでいる「作文コーパス」や「テキストマイニング」等を用い、分析が行われている。

さらに、斎藤（2019）の分析手法と比較すると、脇田（2016, 2017）同様、松下（2012, 2016, 2017）の流れを取り入れ、ルーブリックを数値化し、検証していることがわかる。加えて、脇田（2016, 2017）とは異なり、量的にデータを分析

していることで、量的指標化しても統計分析に耐えられる信頼性を担保していた。また、伊集院（2017）との共通点として「KH Corder」という計量テキスト分析ソフトを用い、「テキストマイニング」分析を行っていた。両者の異なる点は、伊集院（2017）の場合、教員の評価コメントのみを対象に分析を行っているが、斎藤(2019)では、教員評価と大学生の自己評価のずれについて、なぜずれが生じたかを考えさせ、自由記述させたコメントを対象に分析を行っていた。

3. まとめと考察

3.1 まとめ

これまでをまとめると、日本の大学における「評価」に関し、まず、教育学の中でも「教育方法学」の「教育評価論」で論じられている評価研究に注目し、文献レビューを行った。本研究であつかう教育評価研究とは授業・学習評価であり、「形成的評価」や「学習者主動型評価」などに関する評価には、パフォーマンス評価という新しい評価方法が日本の大学でも取り入れられ、そこでは松下（2012, 2016, 2017）以降、斎藤（2019）によってパフォーマンス評価による質的・量的研究が進められていることがわかった。

次に、外国語・第二言語教育の流れを含む日本語教育における評価研究より、「日本の大学」「アカデミック・ジャパニーズ」「パフォーマンス評価」に焦点を絞り、「日本語ライティングの評価研究」関連の研究成果を取り上げた。共著も含め、20 年以上にわたる田中氏の研究成果（1998～2018）、それらを踏襲した伊集院（2017）、伊集院他（2018）、脇田（2016, 2017）は、ルーブリックによる数値化や、テキストマイニング等、新たな分析手法を取り入れ、評価のずれを見ていた。それらの研究手法は斎藤（2019）と共通していたが、斎藤はさらに、教員評価と大学生の自己評価のずれについて、なぜずれが生じたかを考えさせ、自由記述させたコメントを対象に分析を行っていた。

3.2 考察

文献レビューを通し、次世代の日本語教員に必要な評価研究とは、「日本語（言語面）だけでなく内容面も評価できるようになるための授業・学習評価の研究」であると考える。そして、引き続き評価研究を進めていくにあたって、新たな観点を 4 つ、これまでの研究を踏襲する点を 3 つにまとめる。最後に、本調査を踏まえた予備調査について、今後の研究デザインを示す。

3.2.1 新たな観点（4 点）

まず、これまでの研究とは異なる、新たな観点 4 つを以下に示す。

(1) 「日本の大学」「留学生」「中級」にしぼる。
(2) 中級の作文＝「意見文」とし、テーマをしぼる。
(3) 評価のずれを「内容面」にしぼって分析する。
(4) これまで以上に学習者も評価活動に参加する。

上記（1）は、「日本の大学の留学生」が現状において中級が多いという背景があるからである。「中級が多い」とした根拠には、日本語能力試験 N2・N3 の受験者数が最多であることが挙げられる（日本語能力試験，2019）。しかし、日本語能力試験は、「書く」能力を判定はしていない。そこで新たな方法として、日本語学習者作文評価システム「jWriter」を用い、作文データの日本語レベルをそろえることを考えた。

上記（2）は、先行研究では被験者となる日本語学習者や日本人大学生など日本語レベルが異なることもあり、作文課題も異なっていた。しかし、今後おこなう研究ではテーマをしぼっておこなうことを検討する。そこには、上記（3）の「内容面」にしぼって分析したいという意図があるからである。

これまでルーブリックは教員（同士）の評価のためのツールや、学習者の自己評価のためのものが主であった。しかし、今後おこなう研究では、上記（4）のように、その両方を含め「評価活動」とし、同じルーブリックを用い評価研究の実施を考えた。

3.2.2　これまでの研究を踏襲する点（3 点）

続いて、これまでの研究**を踏襲する点** 3 つを以下に示す。

（ⅰ）作文データの収集法
（ⅱ）評価者である日本語教員
（ⅲ）分析方法

（ⅰ）作文データの収集法として近年の田中氏の論文より、作文データコーパス「I-JAS」（エッセイ）の使用を選択した。現在、国内・教室環境（JJC・JJE）116 編中 10 編を抽出済みである。テーマ（プロンプト）は「ファーストフードと家庭料理」(600～800 字)で、同条件によるデータを今後は独自に収集する。

（ⅱ）評価者である日本語教員は、例えば坪根・田中(2015)を参考に、主に 10 年以上大学においてライティング教育に関わる日本語教員を 15 名集めた。

（ⅲ）分析方法は、近年の伊集院（2017）、伊集院他（2018）、脇田（2016, 2017）、斎藤（2019）より、評価（評定）のずれを見ることを念頭に置き、まずは、ルーブリックによって数値化する。そこでは、評定平均、標準偏差、問題パラメータ平均等より分析する。

そして、アンケートまたはインタビューで得た「評価の要因」（自由記述）に対し、計量テキスト分析ソフト「KH Corder」を使用し、テキストマイニングによって分析する。

3.2.3　今後の研究デザイン

現在、来年の本調査に向け、予備調査を計画・実施を進めている。その研究デザインについて、以下表 2 にまとめる。

表 2　今後の研究デザイン（予備調査）

①作文データの収集
・作文データコーパス「I-JAS」（エッセイ）より 10 編+独自に収集 7 編，計 17 編
・日本語学習者作文評価システム「jWriter」でデータの日本語レベルをそろえる。
②調査方法
（Ⅰ）評価者 15 名に上記①をルーブリックで評定してもらう。
※評価者＝10 年以上大学においてライティング教育に関わる日本語教員
（Ⅱ）評定理由について，評価者にアンケートまたはインタビュー調査
③分析方法
（Ⅰ）の分析方法：評定平均／標準偏差／問題パラメータ平均
（Ⅱ）の分析方法：テキストマイニング（KH Corder を使用）

上記①～③を予備調査として一巡したのち、調査協力者である評価者 15 名とふり返り・改善をおこない、本調査では日本語学習者も含めた評価活動を実施する計画である。

4.　おわりに

本研究において、次世代の日本語教員に必要な評価研究を（ルーブリックを含む）パフォーマンス評価に関する研究とし、日本語（言語面）だけでなく内容面も評価できる日本語教員を目指すことを次の課題としたい。

そのため、具体的には、「評価のずれを「内容面」にしぼって分析する」ことや、「これまで以上に学習者も評価活動に参加する」ことを念頭において研究を進めることが必要であると考える。その成果が、これからの授業・学習の質の保証を担保し得るものとなるよう、今後も尽力していきたい。

謝辞

本研究は、科学研究費助成事業（学術研究助成基金助成金）若手研究 19K13248「中級日本語学習者の文章における『内容面』の評価統一に向けた評価研究と評価活動」（代表：安達万里江）の研究の一部である。

参考文献

安達万里江（2019）「日本語ライティングの評価研究―文献レビューより「評価の統一」に向けた課題を探る―」『日本語教育連絡会議論文集 Vol.31』オンライン公開 http://renrakukaigi.kenkenpa.net/ronbun/2018021.pdf（2019 年 7 月 7 日）128-137.

阿部新・田中真理（2016）「グループによるライティング評価における個人評価点の統一パターン」宇佐美洋編『評価を持って町に出よう「教えたこと・学んだことの評価」という発想を超えて』54-69. くろしお出版

伊集院郁子（2017）「作文と評価 日本語教育的観点から見たよい文章」李在稿編『文章を科学する』38-57. ひつじ書房

伊集院郁子・小森和子・奥切恵（2018）「大学教員によるライティング評価の観点を探る」Ishikawa, S. (Ed.). Learner Corpus Studies in Asia and the World Vol.3. Papers from LCSAW2017. Kobe, Japan: Kobe University. 159-176.
斎藤有吾（2019）『大学教育における高次の統合的な能力の評価―量的 vs.質的、直接 vs.間接の二項対立を超えて』東信堂
田中耕治（2008）『教育評価』岩波書店
田中耕治編（2010）『よくわかる教育評価　第 2 版』ミネルヴァ書房
田中真理・阿部新（2014）『Good Writing へのパスポート―読み手と構成を意識した日本語ライティング―』くろしお出版
田中真理、阿部新、影山陽子、佐々木藍子、坪根由香里（2017）「ヨーロッパ日本語学習者のライティング（エッセイ）分析：総合的評価とマルチプルトレイト評価結果を参照して」『ヨーロッパ日本語教育論集』22,75-93.
田中真理（2005）「日本語教育におけるライティング評価」『日本の教育年間 2005 年版』くろしお出版
田中真理（2015）「ライティング研究とフィードバック」大関浩美編『フィードバック研究への招待』107-138. くろしお出版
田中真理（2016a）「パフォーマンス評価はなぜばらつくのか？アカデミック・ライティング評価における評価者の『型』」宇佐美洋編『評価を持って町に出よう「教えたこと・学んだことの評価」という発想を超えて』くろしお出版
田中真理（2016b）「日本語教育の立場から　L1, L2 双方向から考える日本語アカデミック・ライティング：「構成」面について」『ESL Writing in East Asia: Practice, Perception and Perspectives』134-153. Shobi Printing Co. Ltd
田中真理・坪根由香里・初鹿野阿れ（1998a）「第二言語としての日本語における作文評価基準―日本語教師と一般人の比較―」『日本語教育』96, 1-12.
田中真理・坪根由香里・初鹿野阿れ（1998b）「第二言語としての日本語における作文評価―「いい」作文の決定要因―」『日本語教育』99, 60-71.
田中真理・坪根由香里（2011）「第二言語としての日本語小論文における good writing 評価―そのプロセスと決定要因」『社会言語科学』第 14 巻第 1 号, 210-222.
田中真理・長阪朱美（2004）「日本語と英語を目標言語とするライティング評価基準の展望：第二言語としての日本語のライティング評価基準作成に向けて」『第二言語としての日本語の習得研究』7 号 214-253.
田中真理・長阪朱美（2006）「第 2 言語としての日本語ライティング評価基準とその作成過程」『世界の言語テスト』.253-276, くろしお出版.
田中真理・長阪朱美（2009）「ライティング評価の一致はなぜ難しいか―人間の介在するアセスメント―」『社会言語科学』第 12 巻第 1 号, 108-121.
田中真理・長阪朱美・成田高宏・菅井英明（2009）「第二言語としての日本語ライティング評価ワークショップ―評価基準の検討―」『世界の日本語教育』19, 157-176.

中央教育審議会（2018）『2040 年に向けた高等教育のグランドデザイン（答申）』http://www.mext.go.jp/b_menu/houdou/30/11/1411368.htm（2019 年 6 月 25 日）
坪根由香里・田中真理（2015）「第二言語としての日本語小論文評価における「いい内容」「いい構成」を探る―評価観の共通点・相違点から―」『社会言語科学』第 18 巻第 1 号, 111-127.
西岡加名恵、石井英真、田中耕治編（2015）『新しい教育評価入門―人を育てる評価のために』有斐閣
本田弘之、岩田一成、義永美央子、渡部倫子（2019）『［改訂版］日本語教育学の歩き方　初学者のための研究ガイド』大阪大学出版会 47, 52.
松下佳代（2012）「パフォーマンス評価による学習の質の評価―学習評価の構図の分析にもとづいて―」『京都大学高等教育研究』18 号, 75-114.
松下佳代（2016）「第 1 章　アクティブラーニングをどう評価するか」松下佳代、石井英真編『アクティブラーニングの評価』東信堂
松下佳代（2017）「学習成果とその可視化（特集　高等教育研究のニューフロンティア）」『高等教育研究』20. 75-114.
脇田里子（2016）「ライティング・ルーブリックの実践」『コミュカーレ』第 5 号, 21-50
脇田里子（2017）『思考ツールを利用した日本語ライティング』大阪大学出版会
Hart, D.（1994）『Authentic Assessment: A Handbook for Educators』Dale Seymour Pubn、田中耕治監訳（2012）『パフォーマンス評価入門「真正の評価」論からの提案』ミネルヴァ書房
Tyler, R.W（金子孫市監訳）（1978）『現代カリキュラム研究の基礎－教育課程編成のための』日本語教育経営協会（原著 1949 年）
Wiggins, G.（1998）Educative Assessment: Designing Assessments to Inform and Improve Student Performance. Jossey-Bass.

ウェブサイト

多言語母語の日本語学習者横断コーパス（International Corpus of Japanese as a Second Language：I-JAS）https://chunagon.ninjal.ac.jp/static/ijas/about.html（2019 年 7 月 7 日）
日本語能力試験（2019）「過去の試験データ」https://www.jlpt.jp/statistics/archive.html（2019 年 7 月 20 日）
GoodWriting Rater https://goodwriting.jp/rater（2019 年 7 月 7 日）
jWriter https://jreadability.net/jwriter/（2019 年 7 月 7 日）
KH Coder https://khcoder.net/（2019 年 7 月 7 日）