

# ライティング・ループリックの実践

脇 田 里 子

## 1 はじめに

学部留学生に対する日本語ライティングの授業において、数年来、筆者はレポートに関するアンケート調査を実施している。そのアンケートの中に、それまでに学習者がレポートを提出した授業名、レポートの題名、レポートの字数、レポートに対するフィードバックの有無に関する質問項目を設定している。レポートに対するフィードバックに関して、語学のライティングの授業を除いたほとんど全ての授業では、教員から学習者に対して、レポートに関するフィードバックがなされていないことが明らかになった。レポートに対するフィードバックがなければ、自分の書いたレポートがどのように評価されているのかわからない、といった学習者の不満の声は、以前からよく聞かれる。

また、筆者は、ある専門科目の授業で、レポートをフィードバックする際、「内容・構成・表現など」の3項目に関して、A～Dの4段階評価と全体コメントを添えて返却したことがある。授業後、ある学習者が自分のレポートがなぜその評価になったのか、理由を説明してほしいという申し出を受けた経験がある。教員が大まかな評価レベルを示してフィードバックをした場合であっても、学習者からその評価基準について説明してほしいという内容であった。

このように、レポートのフィードバックをしていない場合、フィードバックが無いことに対する不満の声が上がり、また、フィードバックしている場合も、評価判定基準について、更に、詳細な説明を求める声がかかることから、フィードバックの方法は大きな課題であることがわかる。

---

『コミュニカール』5 (2016) 21-50

©2012 同志社大学グローバル・コミュニケーション学会

なお、教員がレポートを評価する際には、教員個人が有する内的な評価項目や評価基準に鑑み、評定することが多い。そのため、同一科目を複数の教員で担当し、同一基準で成績を評価し、複数の教員で成績を管理する体制でなければ、教員同士においても、評価方法の共通認識を持つことが少ないと言える。よって、レポートを評価する場合、評定者によって、評価項目や評価基準は微妙に異なるため、同じレポートを評価しても、評価結果が異なることが十分にあり得る状況が生じる。

そこで、本研究では、学習者のフィードバックの不十分さに対する不満や教員による評価判定の差異を小さくするために、ライティングの成果物に対する評価として、評価観点と段階別の評価基準を明示したルーブリック評価を検証する。本研究の目的は、ルーブリックによる評価を導入する際の留意点や成果を明らかにすることである。

本研究の構成は、以下の通りである。第2章にて、日本語教育におけるライティング評価に関する先行研究を述べる。第3章にて、アメリカで発祥したルーブリック評価の概要について、また、第4章にて、ライティング・ルーブリックを使った実践方法について、それぞれ説明する。第5章にて、ルーブリックを使った実践の結果と考察を述べ、第6章にて、本研究を総括し、今後の課題について言及する。

## 2 日本語ライティングの評価に関する先行研究

日本語教育において、初級から中級レベルの「書き」は「作文」と呼ぶことが多い。本研究において、「作文」は、「自分が思ったことや感じたことを書くことで、その思ったことや感じたことの裏付けがなくても問題がない」(井下, 2014: 8) とし、本研究では「作文」という用語は用いない。本研究で対象とする日本語ライティングは、主として上級レベルの外国人学習者が日本語を駆使して、自分の考えとその裏付けをある一定以上の字数で述べた文章とし、初級～中級レベルで行われる文型学習定着のための日本語作文は対象としない。なお、先行研究において、研究者が「作文」という用語を使用している場合は、そのまま「作文」と表記する。

本章では、日本語ライティングの成果物に対する評価について、主だった

4つの先行研究として、村上（2007）の評価基準、日本留学試験（2010）「記述」の評価基準、石橋（2012）の評価基準、田中他（2014）の評価基準を示す。そして、これらの評価方法よりも、次章で詳述するルーブリック評価の方がフィードバックとしてより望ましいことを提案する。

## 2.1 村上（2007）の評価基準

村上（2007: 133）によれば、作文の評価は主観に左右されるため信頼性に欠けると言われ、従来から点数化することへの抵抗が大きかったとし、作文の授業では添削が中心で、点数で採点されることがあまりなかったと述べている。しかし、2002年に大学入学試験として、大規模試験の日本留学試験（主催は日本学生支援機構、以下JASSO）が始まり、その中に「記述」問題（試験時間20分間、400字程度）が出題されたことから、作文の評価に関する関心が高まってきたという。2002年から2009年までの「記述」問題の採点は、「文法能力」と「論理能力」の2項目につき各3点、計6点満点で採点されていた。

JASSOによる日本語試験（聴解、聴読解、読解）と「記述」問題の得点の相関を調べるため、村上（2007: 135）が用いた「記述」問題の評価方法は、5つの観点に基づく、分析的（analytic）評価を採用した。分析的評価とは、複数の評価の観点を設け、各観点別に評価を出す方法を指す。その5つは、①正確さ、②文体の統一、③表現の多様性・書きことば表現、④文のわかりやすさ、⑤内容である。「記述」を評価するために、5つの観点とそれぞれの観点の満点の数値が示されたが、それぞれの観点をどの程度厳しく採点するかは評定者に委ねられる。

### 村上（2007）による「記述」評価の観点（84点満点）

#### ①正確さ（20点満点）

文法、語彙、表記、原稿用紙の使い方などの正確さを見る。

#### ②文体の統一（4点満点）

文末が「です・ます」体か、「だ・である」体で統一されているか。

#### ③表現の多様性・書きことば表現（20点満点）

表現の多様性、レポートの表現としての適切性を見る。

ひらがなの多用は評価を下げる。

④文のわかりやすさ (20点満点)

接続助詞・接続詞、指示語の用法、文の首尾一貫性などを見る。

⑤内容 (20点満点)

首尾一貫した主張、根拠の妥当性を見る。

## 2.2 日本留学試験 (2010) の「記述」の評価基準

その後、2010年の日本留学試験の「記述」問題から、試験時間が30分、字数が400~500字に拡大され、表1に示す50点満点の採点基準が示された。この採点基準は、全体的な印象をもとに単一の得点評価を行う全体 (holistic) 評価を採用している。2002年の2項目からなる6点満点の採点基準と比較

表1 日本留学試験「記述」の採点基準 (2010年以降)

2009年12月

得点	基準
50点	(レベルS) 課題にそって、書き手の主張が、説得力のある根拠とともに明確に述べられている。かつ、効果的な構成と洗練された表現が認められる。
45点	(レベルA) 課題にそって、書き手の主張が、妥当な根拠とともに明確に述べられている。かつ、効果的な構成と適切な表現が認められる。
40点	
35点	(レベルB) 課題にほぼそって、書き手の主張が、おおむね妥当な根拠とともに述べられている。かつ、妥当な構成を持ち、表現に情報伝達上の支障が認められない。
30点	
25点	(レベルC) 課題を無視せず、書き手の主張が、根拠とともに述べられている。しかし、その根拠の妥当性、構成、表現などに不適切な点が認められる。
20点	
10点	(レベルD) 書き手の主張や構成が認められない。あるいは、主張や構成が認められても、課題との関連性が薄い。また、表現にかなり不適切な点が認められる。
0点	(NA)* 採点がなされるための条件を満たさない。

レベル A, B, C については、同一水準内で上位の者と下位の者を区別して得点を表示する。

すると、「レベル S、レベル A、レベル B、レベル C、レベル D、NA」の 6 段階を示し、それぞれの段階の具体的な基準内容が示されている。この採点基準においても、評定者が基準内容をどの程度厳しく採点するかは評定者に委ねられる。

### 2.3 石橋（2012）の評価基準

石橋（2012）は、中国人学習者を対象に、第 1 言語使用が第 2 言語による日本語作文の質に与える影響を研究している。その中で、作文の質的分析として、分析的评价を採用している。表 2 に示すように、「内容」「構成」「言語形式」の 3 つのカテゴリーと、それらの下位分類として 11 項目、および、それらの評価基準を設定した。なお、この評価基準は、石橋（2002）の研究で使用した日本語作文の評価基準と同じである。石橋（2012）が提案してい

表 2 石橋（2012: 46, 表 3.2）による作文の質的分析の下位項目と評価基準

カテゴリー		基準
内容	1 具体的記述	生き生きとした例や考えをサポートする表現がある。
	2 アイディアの発展	主要アイディアの説明および精緻化がよくできている。
	3 全体的明確さ	理解しやすく順序だてられたアイディアが提示されている。
	4 興味	読み手の注意を引きつける。想像をかきたてられる。洞察力、味方が独特。
	5 主題	書き手の明確で適切な主要アイディアがある。テキストへの表出は非明示的な場合もある。
構成	6 書き出し	何について書こうとしているのかの焦点、あるいはポイントがはっきりしており、読み手を引きつけ、次にくる内容への準備ができています。
	7 論理的つながり	アイディアが段落内において論理的につながっている。
	8 結論	結論が全文を通してのまとめ、述べられていることに基づいている。
	9 まとめ	アイディアが作文を通して主要ポイントに関連している。
言語形式	10 語彙	知的な語彙を使い、変化がある。語彙のレベルが適切。
	11 言語形式の多様性	文の始まり、句、従属節、ディスコースマーカーの多様性

る「内容」「構成」「言語形式」の3つの観点による評価項目は、管見の限り、日本語教員の中でよく用いられている項目であるように思われる。ただし、3つの観点の下位項目の内容は研究者によって多少異なる場合が多い。

また、11項目の評定法として、5段階評価（5非常に優れている、4やや優れている、3普通、2やや劣る、1非常に劣る）を採用している。なお、作文の質の評定は主観が伴うため、日本語教育経験の長い2名の教員が評価に当たるよう配慮している。

## 2.4 田中他（2014）の評価基準

田中他（2014）は、それまでの日本語ライティングの評価研究の集大成として、分析的評価による Good Writing のための評価基準を提案した（表3）。田中他（2014: 6）によれば、Good Writing とは、言いたいこと（メイン・アイデア）が明確に分かり、それが文章全体を通して一貫している文章で、内容のみでなく、形式（構成）がしっかりし、読み手に対する配慮のあるものとしている。この Good Writing は、英語の Essay Writing を参考にしているが、従来の論証型の文章に限らず、活動報告書などの多様な文章を書くことも念頭におき、文章の種類（ジャンルとモード）、文章目的の適合性、文章構成（パラグラフや文章全体のマクロ構成）、「読み手」に対する意識を高めることを提案している。

田中他（2014）では、評価基準は大きく5つのトレイト、A「読み手」、B「内容」、C「モード（ナラティブ、描写、説明、論証など）」、D「構成・結束性」、E「日本語（言語面）」と、それらの下位項目として12項目（または、25項目）を設定している。この評価の観点の新しい点は、A「読み手」とC「モード」を設定している点である。また、評価項目の数としても、ライティング評価としては、最大の項目数であると思われる。なお、書き手、および、読み手に対するチェックリストとして、表3の評価基準を活用する際には、各評価項目に対して、「○・△・×」の3段階でチェックするように指示している。

以上、4つのライティング評価に対する具体的な評価基準を提示した。全てのライティングの評価基準を検討したわけではないが、ライティングの評

表3 田中他(2014: 巻末資料)による Good Writing のための評価基準(一部抜粋)

トレイト		基準説明	
A 読み手	配慮	・読み手を意識しているか。 ・読み手に対する配慮があるか。	
	面白さ	・読み手にとって興味深いか。	
B 内容	メイン・アイデア	・メイン・アイデア(一番言いたいこと)が明確か。	
	トピック・センテンス	・トピック・センテンスの内容が適切で明確か。	
	サポート	・メイン・アイデアやトピック・センテンスが適切にサポートされているか。	
C モード	モード・タイプの選択	・適切なモードが選ばれているか。	
	モード・タイプの展開	・選んだモードに合った展開がなされているか。(D 構成とも関係)	
D 構成・ 結束性	文章全体の 構成(マクロ 構成)	構成	・マクロ構成が明確か。(序論、本論、結論に相当する部分があるかなど。)
		一貫性	・序論と結論が呼応しているか。
		メタ言語	・予告のメタ言語(「まず」「次に」)などが効果的に使われているか。
	パラグラフ間	結束性	・パラグラフとパラグラフのつながりがスムーズか。
		メタ言語	・パラグラフとパラグラフをつなぐ「メタ言語」(「まず」「次に」)などが効果的に使われているか。
		指示語	・指示語(「この」「その」)などが効果的に使われているか。
	パラグラフ内 (ミクロ構 成)	構成	・パラグラフが1文で終わることなく、2文以上で構成されているか。
		結束性	・パラグラフ内の文と文のつながりがスムーズか。
		接続表現	・文と文をつなぐ接続表現(「たとえば」「しかし」)などが効果的に使われているか。
指示語		・指示語(「この」「その」)などが効果的に使われているか。	
E 日本語 (言語面)	a 正確さ	文法	・文法が正確か。
		文型・構文	・文が正確に組み立てられているか。
		語彙・表現	・語彙・表現が正確か。
		表記	・表記(漢字・ひらがな・カタカナ)が正確か。 ・誤字・脱字・漢字の変換ミスがないか。
	b 適切さ (レジス ター)	文型・構文	・単文の羅列や同じ文系の繰り返しではなく、バラエティー豊かな文型が使われているか。 ・日本語らしい構文が使われているか。 ・文の長さが適切で、読みやすいか。 ・意味が複数に解釈できる多義文になっていないか。
		語彙・表現	・語彙・表現が適切に使われているか。
		文末スタイル	・文末スタイル(「普通体」あるいは「丁寧体」)が適切に使われ、基本的に統一されているか。
		表記	・漢字やカナ。 ・句読点・符号が適切に、かつ統一して使われているか。 ・数字(一、二、三と1、2、3)や西暦と元号などが適切に、かつ統一して使われているか。

価方法としては、全体的評価（日本留学試験の「記述」の評価）よりも、各観点別に評価する分析的評価（村上 2007, 石橋 2012, 田中他 2014）の方が多いいと言えよう。その際に、評価する観点項目やその数、観点の段階とその記述などが問題になる。村上（2007）では、Hamp-Lyons & Henning（1991）の Multi-Trait-Multi-Method Matrix 法を使った研究結果に着目し、1つの評価方法を別のテストに転用することは難しく、作文ごとに評価の観点や段階を変えた方がよいという主張に着目している。筆者もライティングごとに、その評価の観点や評価項目を設定することに同意する。

また、上記に示した4つのライティング評価基準を比較することによって、次のことが明らかになった。相違点は、評価項目の観点、評価項目の数、評価全体に占める項目ごとの評価比重や採点方法が挙げられる。一方で、共通点は、評定者が基準内容をどの程度厳しく採点するかが評定者に委ねられることである。確かに、ライティングというパフォーマンスを評価する上で、評定者による個人差が生じることはある程度止むを得ない面もあるが、この差異を最小限にできる評価方法が望ましい。本研究では、評価項目ごとに、その項目のレベルとその評価基準を記述するルーブリックが、教員間の個人差を小さくする評価として適しているのではないかと考える。次章では、そのルーブリック評価について説明する。

### 3 ルーブリック評価とは

ルーブリック（Rubric）評価は、1980年代から、アメリカで、ポートフォリオ（Portfolio）評価<sup>1</sup>とともに、絶対評価の判断基準表を意味する用語として用いられ、現在、アメリカの数多くの高等教育機関が導入・活用している。ルーブリック評価は、客観的な知識や理解を判断するための評価ではなく、思考・判断などのパフォーマンス系の評価において、予め、設定した評価項目とそのレベル設定による達成度を判断するものである。松下他(2013)によれば、ルーブリックは、複数の規準（criterion）とレベル、それを説明する記述語からなる評価基準表のことであり、評価基準であるだけでなく、学習目標を具体的に示すために用いられているという。アメリカにおいて、ルーブリックは、個別の授業、もしくは、個別機関の活用のみならず、機関



を越えて活用可能なルーブリックの開発・運用が、全米カレッジ・大学協会（Association of American Colleges & Universities, 以下、AAC & U と略記）の下で進められている（吉田, 2011: 2）。

AAC & U は、2005 年から 10 年計画で「教養教育とアメリカの約束 (Liberal Education & America's Promise, LEAP)」という取り組みを進めている。その一部として、2007 年から 2010 年にかけて、バリュー (Valid Assessment of Learning in Undergraduate Education, VALUE, 「学士課程教育における適切な学習評価」、以下、バリューと略記) プロジェクトを立ち上げた。バリュープロジェクトでは、学士課程教育段階における「本質的学習成果 (Essential Learning Outcomes)」をめぐる学生の達成度を定義づけ、記述し、評価すること、すなわち、バリュールーブリックの開発が進められている。なお、バリュープロジェクトでは、e ポートフォリオとルーブリックを組み合わせて評価する方法が取られている。

このプロジェクトの下、大学教員やその他の学術関係者などの開発チームがバリュールーブリックの作成に携った。そのルーブリックには、数多くの高等教育機関のほとんどのルーブリックにおいて見出される本質的な規準が含まれている。つまり、各成果領域において学生の学習の質を判断する際に重要と考えられた規準の要約であるという意味で、「メタ」ルーブリックである（吉田, 2011: 6）。

2015 年現在、バリュールーブリックは、16 種類作成されており、本質的学習成果としての 16 領域に対応している<sup>2</sup>。その 16 領域の 1 つに、本研究と関連する「文章コミュニケーション (Writing Communication)」が見られる。次ページの表 4 に、松下 (2012) がこれを日本語に訳したものを示す。

表 4 のルーブリックは、5 つの評価観点、および、各観点に関する 4 つのレベルの評価基準からなる。松下 (2014: 242) によれば、1 から 4 の数字は学年や成績を表示するものではなく、レベル 1 は入学してくる学生にみられるレベル、レベル 4 は学士号を授与される学生に期待されるレベルを表し、レベル 2・3 は文字どおりレベル 1 からレベル 4 に向かう道のりのマイルストーンを表しているという。その意味では、大学 4 年間という長期にわたる学生の学習成果の評価を示した長期的ルーブリックと言える。

よって、バリュールーブリックをそのまま、個々の授業のライティング評価に利用することは困難である。バリュールーブリックは、複数の大学で利用可能な「メタ」ルーブリックとして設計されているため、個々の大学・学科・科目の文脈に合わせるためには、ルーブリックをローカライズすることになる。

松下（2014: 243）は、バリュールーブリックを用いた評価の手続きにおいて、ルーブリックの修正（**modification**）と調整（**calibration**）が重要であるとしている。また、ルーブリックの修正は個々の大学・学科・科目に合わせたローカライズに欠かせないとし、AAC & Uのバリュールーブリックのケーススタディから、下記のようなローカライズが可能であると指摘している。

- ・バリュールーブリックのうち、必要なルーブリックだけを選択する。
- ・規準やレベルはそのまま、記述語の表現を変える。
- ・記述語に加えて、規準やレベルの表現も変える。
- ・規準やレベルを削ったり、加えたりする。
- ・複数のバリュールーブリックを組み合わせ、新たに1つのルーブリックを作る。

一方、ルーブリックの調整とは、「ルーブリックを使う人々が一堂に会して、ルーブリックがどうデザインされておりどう適用されるべきなのかについて共通理解を築くプロセス」としている（松下, 2014: 244）。ルーブリックの調整は、ファシリテータのもと、以下のような手続きで行われるという。

- ・ルーブリックを読み合わせ、解釈を共有する。
- ・学生の作品事例を用いながら、1つの規準（または1つのレベル）について各自が採点する。
- ・採点終了後、参加者は、自分の採点の根拠について作品の該当箇所を引きながら説明する。同じ得点を与えた参加者ごとに説明していく。
- ・以上のことを少なくとも2つの作品事例で行う。

このように、バリュールーブリックを修正や調整し、個々の大学・学科・科目の文脈に合わせたルーブリックを作成することになる。

表4 パリユーループリックにおける「文章コミュニケーション」  
定義

文章コミュニケーションとは、文章作成（ライティング）によって自分の考えを展開・表明することである。文章コミュニケーションには、多くのジャンルと文体を使って書くようになることが含まれる。また、多様な文章作成テクノロジーを使って書くことや、テキスト・デス・イメーჯをうまくミックスさせることなども含まれる。文章コミュニケーション能力はカリキュラム全体を通じて、繰り返し書く経験を重ねることによって発達する。

	4 キャプストーン	3 マイルストーン	2 マイルストーン	1 ペンチマーク
文章作成の 文脈と目的	文脈・読者・目的について完璧な理解を示し、それによって、与えられた課題に対応し、作品のあらゆる要素に焦点をあてている。	文脈・読者・目的、与えられた課題（例えば、読者・目的、文脈を結びつけること）に明確に焦点をあてている。	文脈・読者・目的や与えられた課題（例えば、読者の認知や了解事項への気づき）を見せ始めること）への自覚を示している。	文脈・読者・目的や与えられた課題（例えば、読者としての読者や自己の期待）に対し最低限の注意を示している。
内容の展開	適切で関連性があり説得力に富む内容を用いることによって、科目の習得ぶりを示すとともに、書き手の理解したこととを伝え、作品全体を形づくっている。	適切で関連性があり説得力に富む内容を用いることによって、学問分野の文脈の中でアイデアを探究し、作品全体を形づくっている。	適切で関連性のある内容を用いることによって、作品の大半を通じて、アイデアを展開・探究している。	適切で関連性のある内容を用いることによって、作品の何かで、シンプルなアイデアを展開している。
ジャンルと 学問分野の 約束事	特定の学問分野や文章作成課題に関連する広範な約束事（構成、内容、提示、書式、文体選択を含む）に対し、細かい注意を向けうまく遂行している。	特定の学問分野や文章作成課題に関連する重要な約束事（構成、内容、提示、文体選択を含む）を一貫性をもって使用している。	特定の学問分野や文章作成課題にふさわしいものとして、期待されること（基本的構成、内容、提示など）に従っている。	基本的構成や提示のしかたについて一貫した体系をとおしている。
資料（ソース）と根拠（エビデンス）	当該の学問分野やジャンルにふさわしいアイデアを展開するため、質が高く、信頼でき、関連性のある資料をうまく使いこなしている。	当該の学問分野やジャンルの中に位置づくアイデアを裏づけるために、信頼でき、関連性のある資料を一貫して使っている。	当該の学問分野やジャンルにふさわしいアイデアを裏づけるために、信頼できる（もしくは関連性がある）資料をとおしている。	アイデアを裏づけるために、資料をとおしている。
構文と技法を操ること	読み手に明確かつ流暢に意味を伝えることができる格調ある言葉遣いをしている。ほとんど全く誤りがない。	読み手に意味を伝える直截的な言葉遣いをしている。滅多に誤りがない。	文章に教か所誤りを含むが、明確に意味を伝える言葉遣いをしてしている。	用語法に誤りがあるために、意味の伝達が妨げられるような言葉遣いをしてしている。

(松下2012, p.103, 資料3-4より引用)

注：ペンチマークレベルのパフォーマンス（1のセル）を満たさない作品事例にはゼロを割り当てること。

## 4 ライティング・ルーブリックを使った授業実践

本章では、筆者が担当する学部留学生を対象にしたライティングの授業で使用するライティング・ルーブリックについて説明し、その授業実践の概要を述べる。さらに、ルーブリック評価の特徴と比較するために、従来の教員個人の基準による評価も行う。

### 4.1 実践で利用するライティング・ルーブリック

前章で述べたように、個々の科目に適したルーブリックを作成するには、バリュールーブリックの修正や調整が不可欠で、その作業に膨大な時間と労力がかかる。そのような過程を経て、松下他（2013）は、日本人学生を対象にしたライティング・ルーブリックを作成している<sup>3</sup>。このルーブリックには3つの特徴がみられる。1つ目は、問題解決、論理的思考、文章コミュニケーションの観点から評価する点である。2つ目は、妥当性と信頼性を備えている点である。3つ目は、1回限りの利用に限らず、長期的ルーブリックとしても使える点である。このルーブリックの評価基準は、筆者が実践している日本語ライティングの授業における問題解決する論証型レポートの評価にも利用できる評価観点であることを確認し、これに多少の修正と調整を加えることにした。

なお、第2章の日本語ライティングの評価に関する先行研究で示したように、管見の限り、第二言語の日本語ライティングにおけるルーブリックの研究はほとんど見られない。本来ならバリュールーブリックを元に、第二言語の日本語ライティングのルーブリックをゼロから作成することが望ましい。しかし、日本人学生対象の日本語ライティング・ルーブリックの研究が進められているため、その研究成果を日本語の第二言語教育に応用して作成することとした。

表5に、本研究で利用するライティング・ルーブリックを提示する。これは松下他（2013）のルーブリックを元に、筆者が日本語学習者によるライティング評価に利用するため、次の2点について、修正、調整したものである。

1つ目は、評価観点に新たに「文法・語彙」の観点を加えた点である。松下他（2013）のルーブリックの評価観点は6つで、「背景と問題」「主題と結

表5 日本語ライティング・ルーブリック

観点	問題解決		論理的思考			文章表現	
	背景と問題	主題と結論	根拠と事実・データ	対立意見の検討	全体構成	表現ルール	文法・語彙
観 点 の 説 明	・与えられたテーマから自分で問題を設定する。	・設定した問題に対し、展開してきた自分の主張を関連付けながら、結論を導く。	・主張を支える根拠を述べている。 ・根拠の真実性を立証する信頼性の高い事実・データを示す。*1	・自分の主張と対立する意見を取り上げる。 ・それに対して論駁(問題点の指摘)を行う。	・問題の設定から結論に至る過程を論理的に組み立て、表現する。	・研究レポートとしてのルールを守る。	・文法上の誤りがなく、意図したことを適切な語彙を使用し、日本語で意思伝達できる。
レ ベ ル 3	・与えられたテーマから自分で問題を設定している。  ・論ずる意見も含め、その問題を取り上げた理由や背景について述べている。	・問題に対し、主張を関連付けながら、結論を導いている。  ・結論は一般論にとどまらず、独自性がある。	・根拠を述べている。  ・根拠となる事実・データが2つ以上示されている。	・自分の主張と対立する意見を2つ以上取り上げている。  ・それら全てに対して、論駁(問題点の指摘)を行っている。	・問題設定から結論までが論理的に組み立てられている。  ・記述の順序、段落と段落のつながりがよい。	・指定されたフォーマットを使い、レポートの文字数が守られている。  ・引用部分と自分の文章の区別を明示している。レポートの最後に引用文献を書いている。	・文法上の誤りがほとんどない(助詞、動詞の活用など)。  ・レポートにふさわしい文体(だ、である体)・語彙・表現が使われる。自分の言いたいことが意思伝達できる。  ・1つの長さが長すぎない(60字以内)。修飾・被修飾の関係が明白である。主語と述語が対応している。
レ ベ ル 2	・与えられたテーマから自分で問題を設定している。  ・その問題を取り上げた理由や背景について述べている。	・問題に対し、主張を関連付けながら、結論を導く。	・根拠を述べている。  ・根拠となる事実・データが1つ示されている。	・自分の主張と対立する1つの意見を取り上げている。  ・それに対して論駁(問題点の指摘)を行っている。	・問題設定から結論までが論理的に組み立てられている。  ・記述の順序、段落と段落のつながりが大体よい。	・1つの段落は1つの内容でまとめる。段落に中心文と支持文がある。	
レ ベ ル 1	・与えられたテーマから自分で問題を設定している。  ・その問題を取り上げた理由や背景の内容が不十分である。	・結論を述べているが、主張との関連が不十分である。	・根拠を述べている。  ・根拠となる事実・データが明らかではない。	・自分の主張と対立する意見を取り上げている。  ・それに対して論駁(問題点の指摘)を行っていない。	・問題設定から結論までのアウトラインがたどれる。  ・記述の順序や段落と段落のつながりがよくない。	*2	*2
レ ベ ル 0	・与えられたテーマではない問題設定になっている。 または、問題設定が曖昧である。	・結論が述べられていない。あるいは、結論が曖昧である。	・根拠を述べていない。	・自分の主張と対立する意見を取り上げていない。	・問題設定から結論までのアウトラインをたどることが難しい。	・上記の3つのどれにも該当しない。	・上記の3つのどれにも該当しない。

松下佳代他 (2013)「VALUE ルーブリックの意義と課題」第17回大学教育研究フォーラム、発表資料を参考に加筆・修正

\*1 信頼できる事実・データとは、大学、公的機関、学会、新聞など公共性・信頼性の高い情報を指す。一方、匿名など作成者名がない情報、個人のブログなどは、信頼性が低いと判断される。

\*2 3つの条件をすべて満たす場合は「レベル3」、2つの場合は「レベル2」、1つの場合は「レベル1」と、満たさない場合は「レベル0」とみなす。

<春学期レポートのルーブリック評価>

レベル	背景と問題	主題と結論	根拠と事実・データ	対立意見の検討	全体構成	表現ルール	文法・語彙
3							
2							
1							
0							

<秋学期レポートのルーブリック評価>

レベル	背景と問題	主題と結論	根拠と事実・データ	対立意見の検討	全体構成	表現ルール	文法・語彙
3							
2							
1							
0							

論」「根拠と事実・データ」「対立意見の検討」「全体構成」「表現ルール」であった。それに、日本語母語話者ではない、外国人学習者特有のライティングの問題である日本語の「文法・語彙」の評価観点を加え、7つの評価観点にした。日本語教育の中級レベルまでのライティングは、学習者の作成した日本語文章の語彙・表現・文法などの観点から添削することが多い。そして、上級学習者になっても、日本語の文法や語彙面の誤りは依然として見られる。そのため、「文法上の誤りがなく、意図したことを適切な語彙を使用し、日本語で意思伝達できる」という文法・語彙の観点を加えた。

2つ目は、各評価観点に関する4つのレベルの評価基準の記述を短く、平易簡明な表現に言い換えた点である。また、隣り合うレベルとの違いに関する記述には、その違いに注目しやすいように下線を引いた。これは、この評価表を主として利用する者が、日本語母語話者ではない外国人学習者であることに配慮したためである。

また、本研究のルーブリックでは、評価基準表の下に、レポートに対する評価基準のチェック表を付けた。学習者は自分のレポートが各評価観点のどのレベルに一致しているかについて、シート下のチェック表に印を入れ、学習者が自己評価しやすい環境を提供した。

なお、本研究のルーブリックの作成過程において、4.3で詳述する5名の評価者（日本語教員）が一堂に会して、共通理解を築くプロセスは経ていない。その主な理由は、筆者以外の4名の評価者は、今回の実践対象の授業担当教員ではないためである。よって、授業の目的やレポート課題などをどのようにルーブリックへ反映するかは筆者が検討し、作成したルーブリックについて、後日、他の評価者に説明した。

## 4.2 授業実践の概要

はじめに、実践の対象とした授業の基本情報を表6に示す。実践対象の科目は、日本の某大学の文系学部留学生2年生を対象にした「日本語ライティング2」である。授業の目的は、卒業論文につながる論理的なレポートが書けるようになることである。授業期間は2013年度秋学期(週2回×15週)で、この実践の研究協りに同意した学部留学生は6名(入学時の日本語能力は、日本語能力試験N2レベルが1名、N1レベルが4名、N1レベル以上が1名。以下、「学習者」と称す。)である。学習者6名は、2013年度春学期に、レポート・ライティングの基本的な書き方の授業を受講し、ライティングに関する基礎知識を得ている。

次に、評価対象のレポートについて述べる(表7)。6名の学習者が作成した2種類のレポートを対象とする。1つは春学期の「日本語ライティング1」

表6 授業の基本情報

1	科目	日本の某大学、学部留学生2年生対象「日本語ライティング2」
2	授業の目的	卒業論文につながる、論理的なレポートが書けるようになること
3	期間	2013年度秋学期(90分×週2回×15週)
4	研究協力者	6名(東アジア出身の留学生)
5	4の日本語能力(1年半前、入学時)	J-CAT*において、日本語能力試験(JLPT)のN2レベル 1名、N1レベル 4名、N1以上レベル 1名
6	4のライティング経験	2013年度春学期に、「日本語ライティング1」(週2回)において、レポート・ライティングの基礎を学ぶ。

\* J-CAT (Japanese Computerized Adaptive Test) は、日本語学習者を対象にしたインターネット上の日本語能力自動判定テストである。

表7 評価対象のレポート

		春学期期末レポート	秋学期期末レポート
1	レポート完成時期	2013年7月	2014年1月
2	レポートの型	論証(問題解決型)レポート	
3	テーマ	母国の社会問題	日本の社会問題
4	字数	4,000字	4,000~6,000字
5	レポートの条件	特になし	対立意見の提示 それに対する反論

における期末レポートである。学習者は、レポートの書き方に関する教科書の内容に沿って、レポートを書くための基礎的な表現、構成などを学んでいる。学習者は期末レポートとして、4,000字程度で、母国の社会問題をテーマに問題解決型の論証レポートを作成している。もう1つは秋学期の「日本語ライティング2」における期末レポートである。授業では、アンケート調査型レポートと賛否型レポートを作成後、期末レポートとして、4,000～6,000字で、日本の社会問題をテーマに、問題解決型レポートを作成している。なお、秋学期のレポートにおいては、自分の意見に対する反対側の意見提示とそれに対する反論を書くように指示した。

### 4.3 評価の方法

本実践では、ルーブリック評価と従来の評価を比較することによって、ルーブリック評価導入の成果や課題を探ることが目的である。ルーブリックは学習者の利用を前提にしているため、日本語教員による評価だけでなく、学習者にも自己評価させ、それらの結果を総合的に分析、考察する。

教員の評価にあたり、授業担当者1名を含む5名の日本語教員がレポートを評定する。そして、5名の教員の評価の平均値を教員側の評価として扱う。なお、日本語教員5名は、全員が日本語教員歴10年以上で、レポート・ライティング指導の経験を有しているため、教員による評価の信頼性は高いと思われる。

教員は6名の学習者が作成した2種類のレポートの全て12編を評価する。教員による評価は3種類あり、表8の①③は「教員個人の内的基準による「総合」、「内容」、「構成」、「表現・文法・語彙」の観点におけるA～D判定の

表8 レポート評価の方法

(1) 5名の日本語教員による評価		(2) 6名の学習者による自己評価	
春学期レポート	秋学期レポート	春学期レポート	秋学期レポート
① 個人基準による評価*	③ 個人基準による評価*	⑥ 個人基準による評価	⑧ 個人基準による評価
② ルーブリック評価	④ ルーブリック評価	⑦ ルーブリック評価	⑨ ルーブリック評価
⑤ 2種類の評価の相関性に関するコメント		⑩ ルーブリック評価に対するコメント	

\* 個人基準の「総合」評価に占める「内容」「構成」「文法等」の比重に関しても回答する。



4段階評価」、②④は「ルーブリック評価（表5）」、⑤は「①と②の2つの評価に関するコメント」である。

表8の①③「個人基準による評価」に関して説明する。ライティング評価では、教員個人の内的基準や観点によって評価する場合がある。本実践では、レポート全体に対する「総合」評価、そして、レポートの「内容」、「構成」、「表現・文法・語彙」の3つの観点に関して、A～Dの4段階で評価する。こうした「個人基準による評価」は、多くの教員がレポート評価において採用している評価方法であるため、本研究でも採用した。なお、個人の基準による「総合」評価は、まず先に、全体評価をしてから、「内容」、「構成」、「表現・文法・語彙」の分析的評価をするように依頼した。教員は評価方法についての説明を受けた後、①③「個人基準による評価」、②④「ルーブリック評価」、⑤「コメント」の順に、2つのレポートを同時に評価した。また、①③「個人基準による評価」に関して、「総合」評価に占める「内容」「構成」「文法等」の比重に関する問いにも回答する様求めた。

次に、学習者による自己評価について述べる。学習者は自分が作成した2つのレポートに対して、教員による評価と同様に、⑥⑧「学習者個人の内的基準による「総合」、「内容」、「構成」、「表現・文法・語彙」の観点におけるA～D判定の4段階評価、⑦⑨「ルーブリック評価（表5）」、⑩「ルーブリック評価に対するコメント」を述べる。学習者の自己評価は、単に、学習者自身のレポートに対する内省を知るためだけでなく、教員による評価と比較し、その差異について考察するために、使用する。

学習者による自己評価の時期について補足する。秋学期の授業初めに、学習者は春学期の期末レポートを評価した。その際に、教員は秋学期の期末レポートも同様に、個人基準による評価とルーブリックを用いて、評価することを事前に伝えた。秋学期の最後に、学習者は秋学期の期末レポートを評価した。

なお、日本語教員も学習者も、ルーブリック評価を使って評価することは初めてであった。そのため、評価の前に、ルーブリックの評価方法についての説明を丁寧に行った。

ルーブリックはパフォーマンスを評価するためのツールであり、数値評価

に馴染まない。しかし、教員個人の基準による評価とルーブリック評価を比較するため、ルーブリックの全体評価の数値化を試みた。その方法は、ルーブリック表の各評価観点のレベルをそのまま点数化し、その合計の点数によって、全体評価を4段階に分けるものである。すなわち、表5のルーブリック表の「背景と問題」「主題と結論」「根拠と事実・データ」「対立意見の検討」「全体構成」「表現ルール」「文法・語彙」の7つの観点について、3レベルなら3点、2レベルなら2点、1レベルなら1点、0レベルなら0点と点数化し、合計21点満点とした。そして、全体評価は、16～21点がA評価、11～15点がB評価、6～10点がC評価、0～5点がD評価の4段階に区分する。また、表8①②「教員基準による評価（総合・内容・構成・表現）」の全体評価も数値化を試みた。「教員基準による評価」はA～Dの4段階で評定しているが、教員の平均値を算出するため、これらをA=4.0, B=3.0, C=2.0, D=1.0とした。教員平均値の評定はA= 3.5～4.0, B= 2.6～3.4, C= 1.5～2.4, D= 1.0～1.4に置き換え、算出する。

## 5 ルーブリック評価を使った実践結果と考察

### 5.1 教員の個人基準による評価とルーブリック評価

学習者6名が作成した春学期レポートと秋学期レポートの計12編に対する教員5名による評価結果を表9と表10に示す。表中、6名の学習者は学習者A～学習者F、また、レポートを評価した5名の教員は教員G～教員Kと示す。表中の「学習者A春」とは、学習者Aが春学期に作成したレポートを指す。5名の教員評価の平均値を教員評価とし、これをレポートの評価値として扱う。2つの評価の全体評価を比較するために、教員個人の基準による「総合」評価と、ルーブリック評価の総合点をともに、A～Dの4段階で表した。

表9と表10において、教員平均の「総合」評価とルーブリックの全体評価の4段階評価を見ると、全12編のレポート中、同じ評価レベルになったものが9編(75%)、残り3編(25%)は評価レベルが1段階の差(例、学習者Bの春学期レポートは「総合」がC評価、ルーブリックがD評価)になっている。また、教員5名の「総合」評価とルーブリック評価が同じ評価レベ

表9 教員の基準による評価とルーブリック評価 (学習者A～C)

	学習者A春	学習者A秋	学習者B春	学習者B秋	学習者C春	学習者C秋
教員G：総合評価	C	A	C	B	C	C
教員H：総合評価	C	B	D	D	B	B
教員I：総合評価	A	A	C	B	B	B
教員J：総合評価	B	A	C	B	B	B
教員K：総合評価	A	A	D	A	C	B
<b>教員平均：総合評価 *1</b>	<b>B (2.6)</b>	<b>A (3.8)</b>	<b>C (1.6)</b>	<b>B (2.8)</b>	<b>B (2.6)</b>	<b>B (2.8)</b>
教員G：内容、構成、表現	C C B A A A	C C B B A A	C C B B A A	C B B C C C	C B B C C C	C C C
教員H：内容、構成、表現	C C A B B B	D D D D D A	A A C B C C	A A C B C C	A A C B C C	A A C B C C
教員I：内容、構成、表現	B A A A A A	C C B A B A	B B A B B A	B B A B B B	B B A B B B	B B B
教員J：内容、構成、表現	B C B A A A	C C B B A B	A B A B B A	B B A B B B	B B A B B B	B B B
教員K：内容、構成、表現	B B A B B A	C C C A B A	C C C B B A	C C C B B C	C C C B B C	C C C B B C
<b>教員平均：内容、構成、表現*1</b>	<b>B B A A A A</b>	<b>C C C C B A</b>	<b>B B B B B A</b>	<b>B B B B B C</b>	<b>B B B B B C</b>	<b>B B C</b>
教員G：ルーブリック 21点満点	C(9)	A(18)	C(8)	B(11)	B(12)	C(9)
教員H：ルーブリック	C(6)	B(14)	D(1)	D(2)	B(13)	B(13)
教員I：ルーブリック	A(16)	A(21)	C(8)	A(19)	B(13)	A(19)
教員J：ルーブリック	B(11)	A(21)	D(4)	B(13)	B(12)	C(8)
教員K：ルーブリック	B(15)	A(18)	D(5)	A(20)	C(6)	C(10)
<b>教員平均：ルーブリック *2</b>	<b>B(11.4)</b>	<b>A(18.4)</b>	<b>D(5.2)</b>	<b>B(13)</b>	<b>B(11.2)</b>	<b>B(11.8)</b>

\*1 各教員による評価（総合・内容・構成・表現）は A=4.0, B=3.0, C=2.0, D=1.0 とした。教員平均値は A=3.5～4.0, B=2.6～3.4, C=1.5～2.4, D=1.0～1.4 に置き換え、算出した。

\*2 ルーブリック評価の平均は、A=16～21, B=11～15, C=6～10, D=0～5 に置き換え、算出した。

\* 色付けしたセルは、同一レポートの評価において、教員間で2段階以上のレベル差がある評価を示す。

表10 教員の基準による評価とルーブリック評価 (学習者D～F)

	学習者D春	学習者D秋	学習者E春	学習者E秋	学習者F春	学習者F秋
教員G：総合評価	A	A	C	B	A	A
教員H：総合評価	A	A	B	B	A	A
教員I：総合評価	B	A	B	A	A	A
教員J：総合評価	B	A	B	B	B	A
教員K：総合評価	C	B	B	B	A	A
<b>教員平均：総合評価 *1</b>	<b>B(3.2)</b>	<b>A(3.8)</b>	<b>B(2.8)</b>	<b>B(3.2)</b>	<b>A(3.8)</b>	<b>A(4.0)</b>
教員G：内容、構成、表現	B A B A A B	C D C B A B	B A A A A A	B A A A A A	B A A A A A	A A A A A A
教員H：内容、構成、表現	A A A A A A	B B C A B B	A A A A A A	B B A A A A	A A A A A A	A A A A A A
教員I：内容、構成、表現	B B A A A A	B A B A A B	B A A A A A	B B A A A A	A A A A A A	A A A A A A
教員J：内容、構成、表現	B B A A A B	B C C B B C	B B B B C B	B B A A A A	A A A A A A	A A A A A A
教員K：内容、構成、表現	C B B A B B	C C B B B B	B B A A A A	B B A A A A	A A A A A A	A A A A A A
<b>教員平均：内容、構成、表現*1</b>	<b>B B B A A B</b>	<b>B B C B B B</b>	<b>B B A A A A</b>	<b>B B A A A A</b>	<b>A A A A A A</b>	<b>A A A A A A</b>
教員G：ルーブリック 21点満点	B(15)	A(20)	C(8)	A(19)	B(14)	A(21)
教員H：ルーブリック	C(10)	A(18)	B(14)	A(19)	B(14)	A(20)
教員I：ルーブリック	A(16)	A(21)	B(14)	A(19)	A(18)	A(21)
教員J：ルーブリック	C(10)	A(18)	C(10)	B(15)	B(13)	A(20)
教員K：ルーブリック	C(10)	B(14)	C(10)	B(14)	B(14)	A(19)
<b>教員平均：ルーブリック *2</b>	<b>B(12.2)</b>	<b>A(18.2)</b>	<b>C(11.2)</b>	<b>A(17.2)</b>	<b>B(14.6)</b>	<b>A(20.2)</b>

\*1 各教員による評価（総合・内容・構成・表現）は A=4.0, B=3.0, C=2.0, D=1.0 とした。教員平均値は A=3.5～4.0, B=2.6～3.4, C=1.5～2.4, D=1.0～1.4 に置き換え、算出した。

\*2 ルーブリック評価の平均は、A=16～21, B=11～15, C=6～10, D=0～5 に置き換え、算出した。

\* 色付けしたセルは、同一レポートの評価において、教員間で2段階以上のレベル差がある評価を示す。

ルで評定したものは、教員5名が評定した全レポート60編のうち42編(70%)である。よって、教員個人の基準による評価とルーブリック評価は、概ね相関している。つまり、本実践の教員個人の評価基準には、今回のルーブリックの評価観点と共通する部分が多いと言えよう。

次に、1つのレポートに対する教員5名の評定差に関して述べる。教員個人の評価基準による評定に差異があることは当然予想される。12編のレポートの教員基準による評価において、2段階以上の差異が認められたのは、「学習者A春」「学習者B秋」「学習者D春」の3編のみであった。興味深いことに、その3編は、ルーブリック評価においても、教員間において2段階以上の評定差が見られる。このことから、教員個人の基準による評価とルーブリック評価がある程度相関があることを示すと同時に、ルーブリック評価において、評価規準を記述しても、その規準の適合性に関する判断は、評価する教員次第であることも示している。

例えば、学習者Bの秋学期レポートの評価に関して、教員HはD評価、教員KはA評価と大きな差異が生じた。教員Hと教員Kに確認したところ、これは評定のミスではない。学習者Bは多くの文献から自分の意見を支持する根拠を示し、学習者Bの主張を展開した。その根拠の書き方に関して、教員Hは、先行研究を長々と引用しただけで、自分の考えを述べたものでないと評定した。一方、教員Kは、先行研究を引用した上で、自分の意見を上手にまとめていると評定した。このように、同じレポートであっても、教員間の評定が大きく分かれることは、まれに見られる。そして、このことは、ルーブリック評定であっても、各教員がもつ内的基準が最優先され、ルーブリック評価規準の適合性に対する判断が、必ずしも、同じではないことを示している。つまり、ルーブリックを利用した評定においても、評価者によって評価の差異が生じる場合がある。これはルーブリックを利用する際に留意すべき点である。

表11は、教員の内的基準による「総合」評価における「内容」:「構成」:「文法等」評価の比重を表している。「内容」:「構成」:「文法等」の比重は、「3:2:1」が2名、「2:2:1」が2名、「1:2:1」が1名であった。共通して言えることは、「文法等」の比重が最も軽く、「構成」と「内容」は同じ比重か、

表 11 教員の基準による「内容」:「構成」:「文法等」の比重

教員	「総合」評価における「内容」:「構成」:「文法・表現・語彙」の評価の比重	左の比重にした理由	「内容」とは何か
G	3:2:1	上級レベルのライティングでは、レポートの内容、オリジナリティが最も比重が高く、次に、論理的に理解しやすい文章構成、最後に日本語の文法が間違いがなく、適切な語彙や表現を用いていることが大切だと考える。	「他の人と違う内容の面白さ」「一貫した論旨」
H	2:2:1	文法・表現にはあまり重きを置いていません。その理由は、科目の目的がレポート・論文の書き方を学ぶことにありますし、書き手は(超)上級レベルの学習者なのでそれほど問題がないという前提に立っているからです。	「議論の構築性(説得的に議論を構築できるか)」
I	1:2:1	多くない時間数の授業を通して、「内容」を深める思考的な訓練を母語以外で身につけるのは難しいと思われる。授業で学んだことを評価したいので、構成は、型なので、授業にまじめにとり組み、努力すれば、身につけられると思うから。	「主張の独自性・確実性」だと考えるが、それが見られない場合も、「主張の一貫性」があり、「飛躍がないこと」。
J	2:2:1	今回のレポート執筆の目的は、社会問題を取りあげることとその解決策の提示なので、その中身と解決策が読み手に伝わるかどうかのほうが、日本語の能力より重要だと思われたので、上記のような割合にした。	「主張の伝達」「意見提示」
K	3:2:1	私自身がレポートでもっとも重視する「説得力」に対して、3:2:1の順で影響力があると考えているため。	問題提起から結論までの「議論の一貫性」。加えて、「オリジナルな視点を含んだ発展性」。

表 12 教員の基準による評価とルーブリック評価の相関に関するコメント

教員	教員の基準による評価とルーブリックによる評価の相関はどの程度あるか。
G	教員のレポートの評価基準とルーブリックの評価基準は相関性が高いと思う。「対立意見」の提示は短いレポートでは必ずしも必要としていない。そして、「文法・語彙」観点の全体に占める割合は、個人の基準ではルーブリックより大きいと思う。また、ルーブリックの観点ごとに、各レベルを明文化し、認定をしている点が素晴らしい。今まで「内容」や「構成」と、漠然と捉えていた観点が、ルーブリックの具体的な記述によって、レベル別によく認識できるようになった。
H	数値で計算していないのでははっきりとは言えませんが、全体の相関は比較的高いように思われます。細かく見ると、春は「内容」と「文法・語彙」で一致度にずれがあります。「内容」の評価に関して、教員の評価はオリジナリティや問題提起の際の意見の表明がない場合、減点していないからだと思います。文法・表現に関しては、ルーブリックの方が厳しい評価になっています。また、学習者Eについて教員とルーブリック間のずれが目立ちます。
I	漠然ととらえている「内容」「構成」の印象に関係があると思う。特に、「対立意見の検討」を項目として立てることで、漠然と「内容の深まり」と言っていた部分を言語化して評価できると思う。また、「表現のルール」に示された項目は、漠然と「読みやすさ・こなれている」などの印象であらわしていた評価のポイントになり、学生に評価の説明をする際にわかりやすく提示できる。
J	評価項目の指定があったためか、ルーブリックによる評価の方がばらつきがあるように感じた。特に、春学期のレポートで「対立意見」に触れるものが少なかったため、その違いが顕著だった。
K	「観点」項目では8割程度相関している。「レベル」に関しても一定の相関性はあるが、例えば「根拠」や「対立意見」についての「提示数」や、「文法・語彙」の「文字数」などは非常に疑問なので、自身の評価基準とはしていない。

「構成」よりも「内容」の方が大きい比重にしていることである。各教員がイメージしている「内容」とは、「論理の一貫性」や「独自性」が共通した要素と言える。

最後に、表12に、教員の基準による評価とルーブリック評価の相関に関するコメントを示した。表12のコメントからも、教員の基準による評価とルーブリック評価は相関が高いことがわかるが、各教員がルーブリックの全ての評価項目を重視しているわけでも、各評価項目を同じ比重で評価しているわけでもないことがわかる。しかし、ルーブリックの規準を読んだことで、教員が「内容」や「構成」の評価に対する漠然ともっていたイメージを具体化し、確認できたことがわかる。

以上、個人基準による評価とルーブリック評価を比較してきた。教員が一人で評定する際、評価観点や観点ごとの評価比重などを学習者に明らかにせず、評定することが多い。今回の実践において、教員個人評価の「内容」「構成」「文法等」の評価比重や、教員が「内容」としているものは別の言葉で表現すると、「論理の一貫性」や「主張の独自性」であることがわかった。そして、ルーブリック評価の評価観点や評価規準に触れることにより、教員がそれまで内的な評価項目として漠然とイメージしていたものを、はっきりと認識できるようになったと言える。このように、ルーブリックの評価規準を明示することで、それまでの漠然としていた評価観点や評価基準を明確に示したことに、大きな意味があると言えよう。

なお、今回のルーブリック評価では評価観点間の重みづけは行っておらず、全て同じ比重で評価した。教員が授業で使用する際には、その授業目的を反映し、特定の評価観点到重みづけをすることや評価観点を変更するなどして、各自のライティングクラスに合わせたルーブリック評価表を作成することが望ましいであろう。

## 5.2 学習者の自己評価（個人基準による評価とルーブリック評価）

今回の実践では、学習者にも自分のレポートに対して、教員と同様に、表8に示した⑥⑧「学習者個人の内的基準による「総合」、「内容」、「構成」、「表現・文法・語彙」の観点におけるA～D判定の4段階評価」、⑦⑨「ループ

リック評価（表5）」、⑩「ループリック評価に対するコメント」を実施した。全員の学習者は、これまでのレポート作成において、自分のレポートに対して「レポートでよくできたところ」、「うまくできなかったところ、改善が必要なところ」といった全体的な振り返りの経験をしていた。しかし、今回、示したような評価項目ごとの4段階評価やループリック評価の経験はなく、学習者はより厳格な自己評価に慣れていないと言える。

表13は、学習者個人の内的基準による「総合」、「内容」、「構成」、「表現・文法・語彙」の自己評価とループリックの自己評価、並びに、それらの教員平均の評価を示したものである。ここでは、教員平均の評価を基準に学習者の自己評価を比較することによって、学習者がどの程度評定できているかについて述べる。まず、「総合」評価に関して、教員評価と学習者の評定が一致しているのは、12編中5編のみであった。教員評価よりも自己評価の方が高く評定しているのは2編（学習者B春、学習者B秋）で、教員評価よりも自己評価の方が低く評価しているものは5編（学習者D秋、学習者E春、

表13 教員平均による評価と学習者による自己評価の比較

教員平均による評価	学習者A春	学習者A秋	学習者B春	学習者B秋	学習者C春	学習者C秋
総合評価	B (2.6)	A (3.8)	C (1.6)	B (2.8)	B (2.6)	B (2.8)
内容, 構成, 表現	B   B   A	A   A   A	C   C   C	C   B   A	B   B   B	B   B   C
ループリック 21点満点*	B(11.4)	A(18.4)	D(5.2)	B(13)	B(11.2)	B(11.8)
学習者による自己評価	学習者A春	学習者A秋	学習者B春	学習者B秋	学習者C春	学習者C秋
総合評価	B	A	B	A	B	B
内容, 構成, 表現	B   B   A	B   A   A	B   C   B	A   B   A	A   A   C	A   C   C
ループリック 21点満点	A(17)	A(18)	A(18)	A(18)	A(16)	B(12)

教員平均による評価	学習者D春	学習者D秋	学習者E春	学習者E秋	学習者F春	学習者F秋
総合評価	B(3.2)	A(3.8)	B(2.8)	B(3.2)	A(3.8)	A(4.0)
内容, 構成, 表現	B   B   B	A   A   B	B   B   C	B   B   B	B   A   A	A   A   A
ループリック 21点満点*	B(12.2)	A(18.2)	C(11.2)	A(17.2)	B(14.6)	A(20.2)
学習者による自己評価	学習者D春	学習者D秋	学習者E春	学習者E秋	学習者F春	学習者F秋
総合評価	B	C	C	C	B	B
内容, 構成, 表現	A   A   B	C   B   C	C   B   C	B   C   C	B   B   C	B   B   C
ループリック 21点満点	B(14)	A(17)	C(10)	A(16)	B(14)	A(19)

\* 各教員／学習者による評価（総合・内容・構成・表現）はA=4.0, B=3.0, C=2.0, D=1.0とした。教員平均値はA=3.5～4.0, B=2.6～3.4, C=1.5～2.4, D=1.0～1.4に置き換え、算出した。

\*\* 色付けしたセルは、教員平均と学習者の評価を比較し、評価が高かった方を示す。



学習者 E 秋、学習者 F 春、学習者 F 秋) である。一方、ルーブリック評価においては、教員平均と学習者の評価が一致しているのは、12 編中 8 編であった。教員平均よりも自己評価の方が高く評定しているのは 4 編(学習者 A 春、学習者 B 春、学習者 B 秋、学習者 C 春) で、教員評価よりも自己評価の方を低く評定しているものはなかった。

以上に示したように、学習者の自己評価を教員評価と比較した場合、個人の内的基準による「総合」評価よりも、ルーブリック評価の方が教員評価と一致するところが多い。これは、ルーブリック評価が 7 つの評価観点とレベルごとの評価規準を明示しているため、「内容」、「構成」、「表現・文法・語彙」といった評価者個人の評価観点よりも、ルーブリックの具体的な規準を示した方が評価しやすかったためであろう。なお、学習者による「総合」評価においては、春と秋のレポートの評価において教員評価よりも低かった学習者が 2 名(学習者 E、学習者 F) いた。この 2 名は自分に厳しい評価を課していると言える。その一方で、学習者 B に関しては、春と秋のレポートに対して、総合評価もルーブリック評価も教員評価より高く評定していることから、自己評価が甘い傾向にあると言える。

表 14 は学習者のルーブリック評価に対するコメントを示している。ルーブリックを使って自己評価することによって、評価の観点を知り、学習者のレポートに足りない点を理解しようとする態度が伺えたと言えよう。本実践においては、レポートを作成した後にルーブリックを示すのではなく、レポートを作成する前にルーブリックを提示し、評価観点や評価基準を明示した。このように、事前の評価ポイントの提示が学習者のライティング作成における指標になっている可能性が高い。よって、教員がレポートの課題提示と同時に、評価観点や評価基準などを示すことが望ましいと考える。

## 6 おわりに

本研究では、レポートのフィードバックの不十分さに対する学習者の不満や、教員間による評定の差異を小さくするために、日本語ライティングでは、まだ、あまり導入されていないルーブリック評価について検討してきた。本研究では、アメリカのバリュールーブリックを元に開発された日本人学生用



表 14 学習者のルーブリック評価に対するコメント

学習者	ルーブリック評価を使ったコメント
A	細かい項目に分けて評価をすることで、自分のレポートの問題点がわかった。
B	文章表現、文法・語彙はもちろん大事だが、やはり、根拠と主題を支える事実もしっかり書かないとダメだと気づいた。
C	春学期のレポートは、自分が感じたものを主題として作成したレポートであり、さまざまな情報を手に入れるのは簡単でした。それをまとめるのはなかなか難しかったです。
D	春学期のレポートは結構うまくできていたと思いましたが、春学期のレポートのよくない部分が目に入りました。特に、秋学期のレポートと比べると、根拠とデータの部分が足りないと思います。 ルーブリック評価をして、自分のレポートをもっと細かく評価することができました。今までは、漠然と構成、内容、方法だけで評価しましたが、この評価によって、自分のレポートの改善点がわかるようになりました。
E	ルーブリック評価によって、評価の方法を考えなおすようになった。実際に、どうやって評価が行われるのか、どのような観点からチェックしていけばいいのかについて、知りました。
F	自分では大体よくできていたと思ったが、ルーブリックを使って、対立意見の検討に関しては全然できなかったことがわかりました。そして、1文の長さが長いところが結構あったのではないかと思います。

のライティング・ルーブリック（松下他 2013）に修正を加えたライティング・ルーブリックを作成し、それをライティングの授業実践にて利用した。研究の目的は、ルーブリックによる評価を導入する際の留意点や成果を明らかにすることであった。

実践において、日本語教員 5 名が、12 編のレポートに対して、教員個人の内的基準による評価（「総合」、「内容」、「構成」、「表現・文法・語彙」の観点における A～D 判定の 4 段階評価）と、今回作成したルーブリックを使った評価を行った。その結果、教員個人の基準による評価とルーブリック評価は、全体評価の結果において 70% 前後が一致したことから、概ね相関していると言えよう。

しかし、あるレポートに対する教員の評定において、1 名の教員は最上位の評定を、1 名の教員は最下位の評定を下す事例も生じた。このことから、ルーブリック評価において、評価規準を明示しても、その規準の適合性に対する判断は、評価者次第であるとも言えよう。これはルーブリックを利用する際に留意すべき点であり、可能であれば、複数の教員で評定することが望ましいであろう。

また、教員個人の内的基準による評価においては、評価観点や観点ごとの評価比重などを明らかにせず評定することが多い。今回の実践において、教員個人評価の「内容」「構成」「文法等」の評価比重は、「3:2:1」と「2:2:1」が半々であった。教員が「内容」としているものを明言化すると、「論理の一貫性」や「主張の独自性」に集約できることを示した。さらに、ルーブリック評価の評価観点や評価規準に触れることにより、教員がそれまで内的な評価項目として漠然とイメージしていたものが何であったか、はっきりと認識できるようになったことは、大きな意味があると言えよう。

この実践では、6名の学習者においても、各学習者の内的基準による自己評価とルーブリックによる自己評価を実施した。その結果、学習者によっては、教員評価よりも低く自己評価する者もいれば、高く自己評価する者もいた。そして、学習者のルーブリック評価は、内的基準による自己評価よりも、教員評価と一致する傾向が高かった。次に、学習者はルーブリックを使って自己評価することによって、評価の観点を知り、自分のレポートに欠けているところを理解しようとする態度がうかがえた。

以上のことから、ルーブリック評価を利用すれば、レポートのフィードバックが不十分だとする学習者の不満を小さくすることができるであろう。また、ライティングの評価規準を明示したルーブリック評価を利用することによって、各教員の内的基準による評価と比べ、教員間の評定の差異は縮小していると思われる。しかし、今回の実践において、教員間で大きく評定が分かれた事例も生じたことから、複数の教員が完全に一致した評定を下すことは簡単ではないと言えよう。

最後に、ルーブリック評価の導入に関する今後の課題について述べる。本実践の評価対象にしたレポート数は12編であり、分析したポート数は十分とはいえ、今後、更に検証を進める必要がある。また、今回作成したルーブリックは、評価した5名の教員の合意のもと作成したものではなかったため、評価観点、評価規準、評価比重の点において、各教員が下しているライティング評定と多少の相違が生じている。そのため、授業で実践する際には、担当教員が授業の目的や学習者のレベルなどを把握した上で、より科目に適応したルーブリックを作成することが望ましい。そして、学習者にルーブリッ

クを提示する時期は、今回と同様に、レポートの課題提示と同時に提示することが望まれる。そうすれば、学習者がレポートを作成上、評価の指標を参照し、自己学習を促すことも可能となるであろう。

## 注

- 1 ポートフォリオとは、近藤（2012: 146）によれば、学習者が何かの目的のために、自己の学びの過程と結果、そして、それに対する内省を長期的に記録したファイルを指すという。また、ポートフォリオによる評価は、通常の試験では評価しにくい学習意欲や文章表現力、ボランティアの精神や行動力、作品制作の力量と実績などを評価するために開発された評価方法であるという。
- 2 AAC & U のバリエーションルーブリック（VALUE Rubric）の 16 領域は次の通りである。
  - 知的・実践スキル（Intellectual and Practical Skills）
    - ・探求と分析力（Inquiry and analysis）
    - ・批判的思考力（Critical thinking）
    - ・創造的思考力（Creative thinking）
    - ・文章コミュニケーション（Written communication）
    - ・口頭コミュニケーション（Oral communication）
    - ・読解力（Reading）
    - ・量的分析リテラシー（Quantitative literacy）
    - ・情報リテラシー（Information literacy）
    - ・チームワーク（Teamwork）
    - ・問題解決力（Problem solving）
  - 個人的社会的責任（Personal and Social Responsibility）
    - ・市民参加…地域とグローバル（Civic engagement—local and global）
    - ・異文化間の知識と能力（Intercultural knowledge and competence）
    - ・倫理的思考力（Ethical thinking）
    - ・生涯学習に対する基盤と能力（Foundations and skills for lifelong learning）
    - ・グローバルな学習（Global learning）…2013 年追加
  - 統合的・応用的学習（Integrative and Applied Learning）
    - ・統合的学習（Integrative learning）

- 3 松下他（2013）が作成したルーブリックは、新潟大学歯学部「大学学習法」（1年次前期2コマ）の授業で使用されたもので、レポート評価についての共同研究の成果によるものである。なお、この授業の教育目標は「自ら問題を見つけ、必要な情報を収集、分析、統合し、論理的に思考して問題を解決する基本的な能力を身につける」である。

## 参考文献

- 石橋玲子（2002）『第2言語習得における第1言語の関与』風間書房
- 石橋玲子（2012）『第2言語による作文産出の認知心理学的研究』風間書房
- 市嶋典子（2014）『日本語教育における評価と「実践研究」』ココ出版
- 井下千以子（2014）『思考を鍛えるレポート・論文作成法 第2版』慶應義塾大学出版会
- 宇佐美洋（2014）『「非母語話者の日本語」はどのように評価されているか』ココ出版
- 近藤ブラウン妃美（2012）『日本語教師のための評価入門』くろしお出版
- ステーブンス, D. 他, 佐藤浩章監訳（2014）『大学教員のためのルーブリック評価入門』玉川大学出版部
- 田中真理・阿部新（2014）『Good Writing へのパスポート』くろしお出版
- 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』第18号, pp.75-114.
- 松下佳代・高橋雄介他（2013）「VALUE ルーブリックの意義と課題－規準とレベルの分析を通して－」, 第17回大学教育研究フォーラム発表資料
- 松下佳代（2014）「学習成果としての能力とその評価－ルーブリックを用いた評価の可能性と課題－」『名古屋高等教育研究』第14号, pp.235-255.
- 村上京子（2007）「作文の評価とその関連要因」藤原雅憲他編『大学における日本語教育の構築と展開』ひつじ書房, pp.133-147.
- 吉田武大（2011）「アメリカにおけるバリエーションルーブリックの動向」『関西国際大学教育総合研究所紀要叢書』第4号, pp.1-12.
- Hamp-Lyons, L., and Henning, G. (1991) Communicative Writing profiles: An Investigation of the Transferability of a Multiple-Trait Scoring Instrument Across EFL Writing Assessment Contexts. *Language Learning* 41 (3): pp.337-373.

参考 Web (2015 年 10 月 25 日アクセス)

(独) 日本学生支援機構、日本留学試験、「記述」採点基準

<http://www.jasso.go.jp/eju/saitenkijun.html>

全米カレッジ・大学協会、バリュールーブリック、文章コミュニケーション

<http://www.aacu.org/value/rubrics/written-communication>

## Educational Practice of the Writing Rubric

Riko WAKITA

**Keywords:** Japanese Writing, Evaluation, Assessment, Rubric, Educational Practice

The purpose of this study was to investigate the use of rubrics as a means of improving the feedback given on academic essays written in Japanese by undergraduate students from overseas studying in Japan. I introduced a Japanese writing rubric that was based on the writing rubrics discussed in Matsushita et al. in 2013 to an academic writing class. The rubric included the seven aspects (1 background and problem, 2 theme and conclusion, 3 argument(s) and data, 4 opposing opinion(s), 5 constitution, 6 rule of expression, and 7 grammar and lexicon(s)) and each aspect of the rubric included four levels of quality.

The results of this study suggest that using a rubric for assessment can help students to improve their Japanese academic writing. One possible reason for this is that the aspects and levels of the rubric helped to make the expectations and criteria of the written assignment explicit for both the students and the teacher. The results of the study also show the potential of rubrics to help students further improve their writing if they are given to the students before they are asked to complete the written assignment. Furthermore, it was determined that the scores on the essays given using the rubric as a means of assessment closely corresponded to the teachers' written evaluations.