ルーブリック作成と評価観点の「ずれ」の分析
―上級前半レベルのレポート課題―

山同 丹々子・高橋 雅子・伊藤 奈津美・藤本 朋美・安田 励子

キーワード：ルーブリック評価，レポート評価，評価の公平性，評価観点，自律学習

1. はじめに

早稲田大学日本語教育研究センターで開講されている上級前半レベルの「総合日本語6」は，2016年春学期現在，同一シラバスで4つのクラスがあり，各クラス2名の教員がチームティーチングで授業を行っている。授業において学習者は，3種類のレポートを改稿も含め6回書くことになっている。よって，本科目では授業の中でのレポート作成が大きな比重を占めている。そのため，複数の教員がレポートの添削・評価を行う際には，評価の公平性の担保が求められる。さらに，どのようなレポートが求められているのか，改稿の際はどうすればより良いレポートになるのかを学習者に明示的に提示する必要がある。

そこで，「評価の公平性」を担保するとともに，「到達目標提示」「学習者の自律的な取り組み」を目指し，教員と学習者が参照可能なルーブリック評価表を作成した。

本稿ではその作成と改善過程を報告する。

2. 先行研究

ダネル他（2014）によれば，ルーブリックはある課題について，できるようになってもらいたい特定の事柄を配置するための道具である。つまり，ある課題についての到達目標を明示的に示すことが可能である。さらに到達目標に対しての達成項目，未達成項目を示すことによりフィードバックあるいは評価表としても使用することができる。沖（2014）はルーブリック評価を導入する際の留意点として，①到達目標ごとの成績評価がどのように行われるかが学生に見えること，②成績評価が公平で客観的かつ厳格に行われること，③学習成果のフィードバックが行われることの3点を挙げている。チームティーチングで行われている科目では，評価に複数の教員が関わることが想定されるため，評価の公平性の担保は重要な課題である。複数の教員がレポート評価を行う大学の初年次教育科目におけるルーブリック評価の導入効果を検証した研究に池田・畔津（2012）がある。ルーブリック評価導入前と導入後のレポートの得点率を分析し，ルーブリック評価導入により得点分布率の著しい偏りが解消されたことを示した。ルーブリック評価導入は複数の教員が評価にかかわる際，有用であると言えよう。

3. ルーブリックの作成と使用

3-1. ルーブリックの作成

作成にあたっては先行研究の他，「総合日本語 6」で学習者に配布するレポート作成のために必要とされる要素が記載されている冊子『レポートを書くために』『レポートを書こう』に書かれているレポートの目標を参考にした。また，これまでに「総合日本語 6」でレポートの添削指導経験がある教員からの意見を基に，打ち合わせを重ね，ルーブリック評価表（資料 1 参照）を作成した。

表１　ルーブリックの構成

評価観点	下位項目（14 項目）
(1) 構成	①構成　②段落
(2) 内容	③序論　④本論　⑤結論　⑥展開・一貫性
(3) 表現の適切さ	⑦文体　⑧語彙・文型・表現　⑨文のつながり　⑩表記
(4) 形式	⑪引用の文量・情報量　⑫引用元明示　⑬参考文献欄 ⑭フォーマット

表 1 は作成したルーブリックの評価観点の構成を示している。評価観点は，「(1) 構成」「(2) 内容」「(3) 表現の適切さ」「(4) 形式」の 4 つの項目とし，それぞれに 2 つないしは 4 つの下位項目を設けた。下位項目は全部で 14 項目である。

表２　ルーブリック（一部抜粋）

	A＋(とても良い)	A（良い）	B（もう少し）	C（頑張りましょう）
(1) 構成	□序論・本論・結論の 3 部構成になっており，それぞれの分量が適切である。	□序論・本論・結論の構成になっている。	□序論・本論・結論の構成を意識しているが，整っていない。	□序論・本論・結論の構成が見られない。
	□適切な段落に分かれ，1 段落 1 トピックで書かれており，段落間のつながりが自然である。	□段落に分かれ，1 段落 1 トピックで書かれており，段落間のつながりが見られる。	□段落に分かれているが，1 段落にいくつかのトピックが入っており，まとまりがなく，段落間のつながりがおかしい。	□段落に分かれていない。

＊□には該当箇所にレ点を入れる。

表 2 は完成した基本的なルーブリック（資料 1 参照）の一部抜粋である。評価尺度は，最も高い評価を「A＋(とても良い)」として，「A（良い）」「B（もう少し）」「C（頑張りましょう）」と 4 段階にした。「総合日本語 6」として現時点の遂行目標は「A」とし，「A＋」は将来の到達目標とした。これは，現時点よりさらに上のレベルの目標を提示することにより，余裕のある学習者はそれを意識し，自律的に取り組めるよう配慮したからである。

この他にレポートの採点結果の記入欄及びコメント欄を追加した「教員評価用ルーブリック」を作成した。採点は「A＋」と「A」は「(1) 構成」から「(4) 形式」までの4つの評価観点の合計が10点満点になるようにした。内訳は「(1) 構成」の下位項目である①構成と②段落の下位2項目をそれぞれ1点の計2点とし，「(2) 内容」の③序論から⑥の展開・一貫性の下位4項目もそれぞれ1点の計4点とした。また「(3) 表現の適切さ」の下位4項目と「(4) 形式」の下位4項目は，それぞれ0.5点の計4点とした。なお，先に述べたように，現時点の遂行目標は「A」で，「A+」は将来の到達目標である。よって，「A＋」と「A」を同じ配点とした。さらに，ルーブリックの使用方法，各記述文についての詳細な説明を加えた「教員用マニュアル」を用意した。

3-2. 授業におけるルーブリックの使用

作成したルーブリックは，第1回のレポート評価に使用した。まず，各担当教員に教員用ルーブリックを配布し，使用方法等を説明した。その後，学習者が第1稿のレポートを書く前に，各担当教員は授業で基本のルーブリックを配布し，その見方や，評価項目について説明を行った。次に，学習者のレポート第1稿をフィードバックする時に，レポート本体へのコメントとともに，下位評価項目のチェック欄にチェックを記入したルーブリックを返却した。レポート第2稿提出後も同様に用いた評価を行い，「教員評価用ルーブリック」を学習者に返却した。

4. ルーブリックの評価の「ずれ」と改善

沖（2014）においても述べられているように，ルーブリックが成績評価の公平性を増大させるものであるならば，作成したルーブリック評価表もその公平性が保たれていなければならない。そこで，第1回のレポート評価に使用した後，担当教員が評価しにくかった点とフィードバックで学習者から質問が出た箇所を基に，一度改善を行った。さらに，ルーブリックを用いた評価の公平性を確認し，改善点を明らかにするため，筆者らによる評価を持ち寄り，評価観点ごとにすり合わせを行った。

4-1. 評価観点の「ずれ」

すり合わせには，筆者らにとっては初見の学習者のレポート2つ（レポートX，レポートY）を用いた。これは，初見の学習者のレポートを使用することで，公平性と客観性が保てると考えたからである。

すり合わせでは，5名が各々ルーブリックを用いて評価し，その評価尺度と評価の根拠について評価観点毎に異同を確認した。

2つのレポートに対する評価を突き合わせた結果，表3のように評価のばらつきが見られた。数字は評価者の人数，塗りつぶし部分はすり合わせの結果，決定した評価である。

表３　評価の「ずれ」－評価観点別－

		レポート X				レポート Y			
		A＋	A	B	C	A＋	A	B	C
(1) 構成	①構成	2	2	1			4	1	
	②段落		1	4			2	3	
(2) 内容	③序論		4	1			5		
	④本論		1	4			3	2	
	⑤結論		2	3	4			1	4
	⑥内容の一貫性			4	1		1	4	
(3) 表現の適切さ	⑦文体		5		1		5		
	⑧語彙・文型・表現			5				5	
	⑨文のつながり			4	1			5	
	⑩表記		5			2	3		
(4) 形式	⑪引用の分量情報量		3	1	1	1	2	2	
	⑫引用元明示		3	2			5		
	⑬参考文献		3	2	1			4	1
	⑭フォーマット	2	3			3	2		

上記表 3 を見ると，レポート X の「(1) 構成」の①構成は「A＋」と評価した者が 2 名，「A」と評価した者が 2 名，「B」と評価した者が 1 名であり，評価が「A＋」と「B」の 3 段階に評価尺度がばらついたことがわかる。一方，レポート Y の「(3) 表現の適切さ」の⑦文体は評価者全員が「A」，⑧語彙・文型・表現と⑨文のつながりは 5 名全員が「B」に評価しており，評価尺度のばらつきがなしとなる。

表 4 は表 3 の評価尺度のばらつきの幅でまとめたものである。数字は下位項目のばらつきの合計数を段階ごとに載せている。また「なし」は下位項目のうち，評価尺度のばらつきがなかった，つまり，全員の評価が一致していることを意味している。

表４　評価の「ずれ」－評価尺度のばらつき－

	評価尺度のばらつき			
	なし（評価一致）	2 段階	3 段階	4 段階
レポート X	3	9	2	0
レポート Y	5	8	1	0

全員の評価が一致した項目と 2 段階に分かれた項目の合計は，レポート X が 12 項目，レポート Y が 13 項目であった。このことから，14 の下位項目のほとんどにおいてばらつきの少ない評価ができていた。また，一番差が大きい 4 段階のばらつき「A＋～ C」の範囲が出た項目はなかった。この時点で，作成したルーブリックは一定の評価の公平性が保

たれていたといえるであろう。

一方，比較的ばらつきが大きかった3段階に評価が分かれた項目は，3項目であった。そのうち，2つ（レポートXの「(1) 構成」の①構成とレポートYの「(4) 形式」の⑪引用の分量と・情報量）は「A＋～B」の範囲でのばらつきであった。3-1. で述べたように，「A＋」は将来の到達目標であり，「A」は現時点の遂行目標である。また，「A＋」と「A」は同じ配点にしてあるため，「A＋」と「A」の評価のずれは実質的な影響はないものと考えられる。

3段階に評価が分かれた「(4) 形式」の⑪引用の分量・情報量は「A～C」の範囲でばらつきが生じており，評価に影響が出るものであった。

4-2.「ずれ」の要因と改善点

評価観点の「ずれ」が生じた要因は，評価観点項目がない場合（4-2-1.）と評価観点の記述に対する解釈の違い（4-2-2.）の二つの場合であることが判明した。以下にそれぞれの要因とその改善点を述べる。

4-2-1. 評価観点項目がない場合

ルーブリックに記載がない項目を評価し，指導する必要が生じた場合，評価者独自の解釈によって類似の評価観点が適用されたため，評価にばらつきが生じていた。

例えば，評価観点項目「(4) 形式」の第1の観点（引用の「分量」は適切か）は，レポートX，Yともに3段階に評価尺度が分かれている。ある評価者は引用の「適切さ」（引用が必要なところで行われているか）を評価する必要があると考えた。だが，この評価観点がルーブリックになかったことから，引用の「分量」の評価観点に「適切さ」を含めて評価した。そのため，「分量」のみの観点で評価した評価者との間に「ずれ」が生じたのである。

すり合わせの結果，引用の「適切さ」は，評価観点項目「(2) 内容」の第2の観点（データに基づいて意見の根拠を示しているか）で行うこととし，「教員用マニュアル」の記述にその旨を記した（表5参照）。

表5 「教員用マニュアル」（一部抜粋）

	B（もう少し）
内容	本論：データに基づいて自分の意見の論拠を示しているが，客観的で適切なデータではない。 ・データが信頼できない情報源，例えばウィキペデイアや個人のブログなどの場合。 **・データは提示されているが，意見を直接支持する論拠ではない。**

＊太字部分が記述を追加したところ

4-2-2. 評価観点の記述に対する解釈の違い

評価観点の記述，「教員用マニュアル」に対する解釈が評価者によって異なる場合，評価にばらつきが生じていた。

例えば，評価観点項目「(2) 内容」の第 1 の観点は，序論にテーマの背景と問題提起が書かれているかどうかである。レポート X には「問題提起」となる文は見られなかったが，「レポートの目的」を「問題提起」と解釈した 4 名は評価を「A」と判断した。しかし，すり合わせの結果，「問題提起」は明確に書くべきであるという見解で一致し，評価は「B」とした。ルーブリックの記述にも「はっきり」書くべきである旨を追加した（表 6 参照）。

表６ 「教員用マニュアル」（一部抜粋）

	A（良い）
内容	序論：レポートのテーマの背景と問題提起がどちらもはっきり書かれている。 ・レポートのテーマ背景と，問題提起（一つでも可）書かれている場合。 **・レポートの目的だけでは問題提起と考えない。**

＊太字部分が記述を追加したところ

他にも，「分量が適切である」「わかりやすい文章」のように，定量化されていない表現が含まれる評価観点についても，評価のばらつきや判断のしにくさが指摘された。いずれも，評価観点項目，「教員用マニュアル」の記述について，曖昧な部分を明確にした。

5. 学習者の様子

再改訂したルーブリックは，第 2 回目以降のレポート活動に使用した。学習者は徐々にその使い方に慣れてきたようで，第 1 稿を書いたあとのお互いのレポートを読むピア活動で積極的に活用していた者もいた。またレポートの目標を書く際に，「構成に気をつけて書く」「引用と自分の意見を区別する」などのルーブリックの項目を意識した目標を書いていた。

ルーブリックの使用について学習者からは，「レポートに必要な要素がはっきりわかった」「修正点がわかりやすかった」「自分のレポートを客観的に見られるようになった」という声も聞かれた。全体的な傾向として，レポート作成ごとにルーブリックの評価が上がっていく学習者が多く見られた。特に「レポート作成の際にルーブリックの項目を意識して書いた」という学習者は最後の 3 回目のレポートは将来的な目標である「A＋」の評価が多くなっていった。

6. まとめと今後の課題

本稿では，レポート評価をする際に必要なルーブリックを作成し，その作成と改善の過程について述べてきた。それは，「総合日本語 6」において，レポート作成に対する評価の比重が大きく，各教員がレポートの評価の公平性を担保することが課題であったこと，また学習者がレポート活動を通して，自律的に学習に取り組めるようになるねらいもあったからである。

作成後は，評価の「ずれ」が見られた箇所を中心に担当教員の話し合いやレポート評価のすり合わせを通し，ルーブリックと「教員用マニュアル」を再度改訂した。再改訂版を用いて第 2 回のレポートを評価したところ，評価で迷いが生じる部分が少なくなった。このことにより，ルーブリックは使用しながら記述の精度を高め，改善していく必要があるだろう。

学期末に学習者へのアンケート調査とインタビュー調査，教員へのインタビュー調査を行った。今後は，これらの結果を分析し，「評価の公平性」「到達目標の提示」「学習者の自律的な取り組み」ができていたか，ルーブリックの有効性について検証していきたい。

参考文献

池田史子・畔津忠博（2012）「複数教員によるレポート評価のためのルーブリック形式の評価表導入に関する検証」『日本教育工学会論文誌』第 36 号，53-156.

沖裕貴（2014）「大学におけるルーブリック評価導入の実際―公平で客観的かつ厳格な成績評価を目指して―」『立命館高等教育研究』第 14 号，1-90.

ダネル・スティーブンス，アントニア・レビ（2014）佐藤浩章・井上敏憲・俣野秀典（訳）『大学教員のためのルーブリック入門』玉川大学出版部

（さんどう　にこ，早稲田大学日本語教育研究センター）
（たかはし　まさこ，早稲田大学日本語教育研究センター）
（いとう　なつみ，早稲田大学日本語教育研究センター）
（ふじもと　ともみ，早稲田大学日本語教育研究センター）
（やすだ　れいこ，早稲田大学日本語教育研究センター）

資料１　総合日本語６レポートのためのルーブリック評価表

	A＋（とても良い）	A（良い）	B（もう少し）	C（頑張りましょう）
(1) 構成	□ 序論・本論・結論の３部構成になっており，それぞれの分量が適切である。 □ 適切な段落に分かれ，１段落１トピックで書かれており，段落間のつながりが自然である。	□ 序論・本論・結論の構成になっている。 □ 段落に分かれ，１段落１トピックで書かれており，段落間のつながりが見られる。	□ 序論･本論･結論の構成を意識しているが，整っていない。 □ 段落に分かれているが，１段落にいくつかのトピックが入っており，まとまりがなく，段落間のつながりがおかしい。	□ 序論・本論・結論の構成が見られない。 □ 段落に分かれていない。
(2) 内容	□ 序論：レポートのテーマの背景と問題提起がどちらもわかりやすく明確に書かれている。 □ 本論：十分な量の客観的なデータに基づいて，自分の意見の論拠が導きだされ，明確でわかりやすい説得力のある文章になっている。また，自分の意見に対する反対意見について，適切な予想ができており，その意見に明確に反論も行っている。 □ 結論：本論の内容が効果的に，かつ簡潔にまとめられ，問題提起に対する答えと，自分のレポートに関する今後の課題が明確に書かれている。 □ レポート全体の展開が，論理的思考に沿ったもので，序論から結論に至るまで，レポートの内容に一貫性がある。	□ 序論：レポートのテーマの背景と問題提起がどちらも書かれている。 □ 本論：客観的なデータに基づいて，自分の意見の論拠が導きだされている。 □ 結論：本論の内容がまとめられ，新しい内容が入っていない。問題提起に対する答えと，自分のレポートに関する今後の課題が書かれている。 □ レポート全体の展開が，理解に支障がないもので，序論から結論に至るまで，レポートの内容にほぼ一貫性がある。	□ 序論：レポートのテーマの背景と問題提起のどちらかしか書かれていない。または，はっきり書かれていない。 □ 本論：データに基づいて自分の意見の論拠を示しているが，客観的で適切なデータではない。 □ 結論：本論の内容は入っているが，あまりまとめられておらず，新しい内容が少し含まれている。問題提起に対する答えはないが，今後の課題は書かれている。 □ 段落の入れ替えや，文や段落の追加を考えた方が良い箇所がある。序論から結論に至るまで，レポートの内容に一貫性のない部分が多少ある。	□ 序論：レポートのテーマの背景と問題提起のどちらも書かれていない。 □ 本論：データのみ，または自分の意見のみしか述べられていない。 □ 結論：結論がない，または序論・本論と関連がない新しい内容がかなり含まれている。今後の課題が書かれていない。 □ レポート全体の展開や内容理解にかなり支障がある。序論から結論に至るまで，レポートの内容に一貫性がない。
(3) 表現の適切さ	□ 全体が「だ体･である体」で統一されており，「である」「のである」の使い分けも適切である。 □ レポートにふさわしいテーマに沿った語彙・文型・表現を効果的に使って論じ，全体を通じてキーワードを効果的に使っている。さらに文型･表現も適切で誤りがない。 □ 一文の長さが適切で，文にねじれがなく，文末表現・接続表現が豊富で，正しく使えている。 □ 表記に誤りがない。	□ 全体が「だ体・である体」で統一されている。 □ レポートにふさわしいテーマに沿った語彙・表現を使っており，レポートにふさわしい文型・表現にほぼ誤りがない。 □ 一文の長さが適切で，文にねじれがなく，文末表現・接続表現が正しく使えている。 □ 表記にほぼ誤りがない。	□ ほとんどが「だ体・である体」で書かれているが，「です・ます体」の箇所もある。 □ レポートにふさわしいテーマに沿った語彙・表現が十分に使われていない。または，レポートにふさわしい文型・表現に多少の改善が必要である。 □ 一文が長すぎて読みにくいが，内容は理解できる。文にねじれがあり，文末表現・接続表現に多少の改善が必要である。 □ 表記に少し誤りがある。	□ 「だ体･である体」で統一されておらず，「です･ます体」が混在している。 □ レポートにふさわしいテーマに沿った語彙・表現がほとんど使われておらず，レポートにふさわしい文型･表現にかなりの改善が必要である。 □ 一文が長すぎて読みにくく，内容も理解できない。文にねじれがあり，文末表現・接続表現にかなりの改善が必要である。 □ 表記に誤りが多い。
(4) 形式	□ レポート全体の中の引用の分量が適切で，情報量に無駄がない。 □ 引用と自分の意見を区別し，正しいやり方で引用しており，正しく引用元が示されている。 □ 引用文献（参考文献）がレポートの最後に正しい順番と形式で書かれている。 □ 指定されたフォーマット・ファイル形式で書かれている。	□ レポート全体の中の引用の分量が適切である。 □ 引用と自分の意見を区別し，ほぼ正しいやり方で引用しており，引用元が示されている。 □ 引用文献（参考文献）がレポートの最後に正しい形式で書かれている。 □ 指定されたフォーマット・ファイル形式がほぼ守られている。	□ レポート全体の中の引用の分量がやや多い。または少ない。 □ 引用元が示されているが，引用と自分の意見が区別されにくい。 □ レポートの最後にある引用文献（参考文献）に足りない情報がある。 □ 指定されたフォーマット・ファイル形式があまり守られていない。	□ レポート全体の中の引用の分量が多い。または引用がない。 □ 引用元が示されておらず，引用を自分の意見として書いている。 □ レポートの最後に引用文献（参考文献）が書かれていない。 □ 指定されたフォーマット・ファイル形式が全く守られていない。