

# 「学習者参加型評価」を導入する有効性について —学部留学生の「日本語口頭表現」授業の事例研究を通して—

小川 都

## 1. はじめに

佐々木 (2006) によれば、日本語教育では 1980 年代から、従来の文型積み上げ式の学習から学習者集団のニーズ分析を重視するコミュニカティブ・アプローチへと教育観が変化した。その後、90 年代半ばから学習者が社会との関わりの中で「自律的学習」を重視する構成主義的学習観が支持されるようになった。一方、大学における学部留学生 (以下は留学生とする) のための日本語教育においても変化があった。細川 (2002) によれば、留学生の自律的学習能力が必要とされ、他者との関わりの中で自ら大学日本語教育の内容に参加していく「学習者参加型」日本語教育が提案された。

その中で、留学生が日本語学習に積極的な参加を促す「学習者参加型評価」は従来の教師主導の評価に替わるものとして提唱された。そして、大学の日本語教育における「学習者参加型評価」の実施は、留学生が明確な学習の到達点 (評価基準) を認識し、自分自身の日本語能力をより適切に評価できる高い「自律的な評価力」を身につけ、最終的に大学での全体的な学習への積極的な参加に繋がっていくのではないかと考えられる。

本研究は大学の日本語教育における「日本語口頭表現」の 1 年間の授業を通して、「学習者参加型評価」を取り入れたクラス活動を行い、受講した留学生達の学習の到達点 (評価基準) への認識が如何に変化したのか、またその認識の変化は学習効果にどのような影響を与えたのかを量的に分析し、「学習者参加型評価」を取り入れるクラス活動の有効性を検討したい。なお、紙幅のため個別の質的分析は別稿に譲りたい。

## 2. 「学習者参加型評価」

「学習者参加型評価」とは従来の教師主導の評価に替わるものとして、学習者が評価に参加することである。その中で、細川 (2002) は、「自己責任型の自己評価」の重要性を述べたうえで、他者評価のコメントを取り入れた「相互自己評価」とでもいうべき方法が有効であると主張した。また、村田 (2004) は留学生の自己評価および他者による評価を組み合わせることは「学習者同士の観察・評価の視点を広げ」「メタ認知力を高め、自律的な評価力を高めていくために役立つ」と述べ、他者評価の重要性も示唆した。さらに、森・衣川 (2010) では、他者評価をするという活動が自身のモニタリングや評価につながる可能性について考察し、評価する側とされる側と共に行きつぎが生まれる他者評価活動を繰り返し行うことで、モニタリングの基準の意識化がより促進されるのではないかと述べている。

しかし、村田 (2004) は「学習者による他者評価は効果的と言われながらも、自己評価と組み合わせると効果的に利用し発表の改善につなげていけるような具体的な方法

はまだ確立されていない」と指摘した。また、高木（1992）は他者評価に関し、評定基準や判断基準を持たないために主観的になりやすく、相手への気遣いから実際より良い評価をするなどの問題点を指摘した。

### 3. 研究の目的

これまでの先行研究によって、学習者の自己評価と学習者同士による相互評価の重要性が明らかになった。しかし、自己評価と他者評価との相互関係によって、学習者の自律的評価力が如何に向上したか、また、その相互関係によって学習の到達点（評価基準）への認識の変化について触れていない。酒井（2003）は「学習者に対して、教える側が無意識的、意識的に学習者に期待する基準が伝わらないことがある」また、「学習者からどこまで正確さを要求されているのか理解できないと訴えられたのだが、教える側が行っている判断基準が学習者には理解できず、ズレを生じる原因となったようである」と教える側と学習者側との間で、評価基準についてズレが生じることを指摘した。

今回のクラス活動は従来の教師主導の評価に替わって、留学生同士による評価（留学生の自己評価、同じクラスの仲間による他者評価）、および授業を担当する教師による評価（以下教師評価<sup>1</sup>とする）も同時に実施した。本研究では、先行研究を踏まえ、「学習者参加型評価」というクラス活動の学習プロセスについて数量的に分析し、留学生が評価活動を共有し、内省と他者からのフィードバックを組み合わせることによって3評価間で生じるズレに気づき、その相互関係によって学習の到達点（評価基準）への認識が深まることを検証したい。また、学習の到達点（評価基準）への認識を深めることは、留学生の学習効果にどのように影響するのかについても検証したい。

## 4. 事例研究の概要

### 4.1 対象

研究対象は、私立大学に在学する留学生1年生のために設置された「日本語口頭表現」クラスである。授業は週1回90分、前期と後期各15回、通年30回実施された。対象者は該当クラスの受講者男性7名、女性5名の全12名で、国籍は中国8名、韓国4名であり、全員すでに上級レベルに十分達している学習者である。

### 4.2 クラス活動内容

#### 4.2.1 前期のクラス活動

前期はインタビューと他者紹介、スピーチの実施方法、質疑応答に応用できる討論の仕方、および自己評価と他者評価の仕方について説明し、練習を繰り返した。さらに、その都度フィードバックを行った。そして、前期の学習効果を確認するため、ス

---

<sup>1</sup> 教師評価も他者評価の一種であるが、本研究では、同じクラスの留学生の仲間による他者評価と区別し、教師評価を学習到達点（評価基準）への認識のバロメーターとして第3の評価とする。

ピーチテストを実施した。

①課題

テーマは「日本の生活の中で自分自身の関心事」について自分の周りにいる日本人にインタビューし、それを報告すること。

②発表内容

前期テストの発表内容は表1のようになっている。

表1 前期テストの発表内容

発表者 <sup>2</sup>	発表順番	インタビュー相手	インタビューのテーマ
A	1	アルバイト先の社長	子供の教育について
B	2	日本の高校生	高校生の制服について
C	3	日本人の友人	お勧めできる観光地
D	4	日本人大学生	将来の夢について
E	5	日本人の友人	日本の結婚式について
F	6	日本人の友人の父親	日本のIT関係の仕事について
G	7	日本人の友人	Lineの使用について
H	8	日本人大学生	就職活動について
I	9	アルバイト先の日本人	女性の一人暮らし
J	10	大家さん	外国人入居者について
K	11	日本人大学生	中国に短期留学した時の感想
L	12	同じゼミ日本人学生	外国人の日本語について

#### 4.2.2 後期のクラス活動

後期は社会科学系専門分野における口頭発表に必要なスキルを説明し、また、論理的な考え方や資料・データの収集法、および発表時に使用するパワーポイントの制作方法についても具体的に説明した。その後、学生同士の協同活動によって、発表テーマについて検討し、原稿を書き、発表のリハーサルと同時に自己評価と他者評価の練習、および教師評価からのフィードバックを行った。そして、最後に個人発表テストを行った。

①課題

テーマを自分の専門分野から選び、それについて専門用語の意味を理解し、実例やデータを用いて説明すること。また、それに対し自分の考えを発表すること。

<sup>2</sup> 12人の発表者をA～Lで標記したもの。前期と後期は同じ人を指す。

## ②発表内容

後期テストの発表内容は表2のようになっている。

表2 後期テストの発表内容

発表者	発表順番	専門分野	発表テーマ
A	1	経営	起業について
B	2	経営	企業の不祥事
C	3	商	コカ・コーラ社の公益広告について
D	4	経営	中小企業について
E	5	経営	ブラック企業について
F	6	経営	企業のM&Aについて
G	7	経済	所得格差とジニー係数との関係
H	8	商	マーケティング・ブランドについて
I	9	商	ソーシャルメディアについて
J	10	経営	ベンチャー企業について
K	11	商	共感覚マーケティング
L	12	経済	定価と利益の関係

### 4.3 評価の仕方

前期・後期におけるテストはそれぞれ12回実施した。前期はスピーチを録音し、後期は発表の様子を録画した。音声、映像データは評価のフィードバックに使用した。

テスト時間は1人に付き20分程度である。評価の仕方については、まず、発表者による発表を15分間行い、さらに発表内容に対し5分間の質疑応答を行った。次に、評価シートを全員に配布し、聞き手となる11名の留学生による他者評価をその場でしてもらった。記録した音声と映像を本人に渡し自己評価を授業後に提出してもらった。そして、担当教師による評価については、その評価の妥当性や適切性を確認するため、テストの評価基準、および本研究の対象者の日本語レベルを熟知している同大学の日本語教師1名と共同で評価を行い、今回の教師評価として留学生同士による他者評価と共に発表者である留学生にフィードバックした。

### 4.4 評価基準

評価基準はCouncil of Europe(2004)のCan-do一覧を参考にし、シラバスの到達目標、および留学生の日本語レベルに照合して設定した。その評価基準は表3のようになっている。

評価シートは教材として使用した『日本語口頭発表と討論の技術—コミュニケーション・スピーチ・ディバートのために—』に掲載している評価表を参考に筆者によって制作したものを使用した。スピーチ評価シートは参考資料1、個人発表の評価シートは参考資料2のようになっている。

「学習者参加型評価」を導入する有効性について

表3 テストの評価基準

評価の観点	到達目標
内容・活動	自分の専門分野からテーマを選び、専門用語などを参考文献に書かれている説明だけで発表内容として発表するのではなく、自らその意味を理解し、実例やデータを用いて説明すること。また、それに対し自分の考えも発表すること。
談話構成	・明確かつ体系的に展開できる。
	・明確に詳しく述べるができる。
	・根拠を示しながら説明できる。
	・利点と不利な点、賛成や反対の理由を挙げられながら説明できる。
流暢さ	・言いたいことを比較的困難なく表現できる。
	・間があいたり行き詰ったりすることはあるが、人の助けを借りずに話を続けられる。
語彙	・複雑な考えを述べるのは難しいが、自分の専門学習に関する専門用語や専門的な知識について述べられる。
文法	・母語の影響や誤りが見られても、それに自覚し修正しながら正確に文法を使うことができる。
発音	・母語の発音の影響があっても、間違えに気づき、修正しながらはっきりとした発音やアクセントで話すことができる。

## 5. クラス活動の結果分析

前期テストおよび後期テストにおいて、それぞれの評価シートを5件法で自己評価・他者評価、教師評価を行った。そして、それぞれの評価値を100点満点に換算し得点化した。その結果は表4の通りである。

表4 3 評価の得点の記述統計

前期テスト	度数	最小値	最大値	平均値	標準偏差
自己評価前	12	42.00	88.00	67.83	13.91
他者評価前(11×12)	132	58.00	88.00	74.67	7.32
教師評価前	12	42.00	88.00	57.33	15.43
後期テスト	度数	最小値	最大値	平均値	標準偏差
自己評価後	12	53.00	84.00	67.5	11.26
他者評価後(11×12)	132	60.00	77.00	71	5.94
教師評価後	12	56.00	80.00	68.67	7.09

以上の得点を用いて、「学習者参加型評価」を取り入れた全30回のクラス活動を通して、以下の二つの仮説を数量的に検証した。分析にあたって、統計ソフトのPASW Statistics Base 18を使用した。

仮説①自己評価・他者評価・教師評価間のズレを認識することによって、留学生は学習の到達点（評価基準）への認識が深まる。

仮説②「学習者参加型評価」の実施は留学生の学習効果の向上につながる。

### 5.1 「自律的な評価力」について

国際交流基金（2011）によれば、学習者が自分自身を評価する視点をもつことは自分自身に合った学習方法を探し自分の力で学習を進めていくことにつながる、と学習

者の自己評価を推奨している。しかし、適切な評価のためには明確な学習到達目標の設定、およびそれについて判断する「自律的な評価力」が必要とされる。自己評価だけではなく、他者評価にも同様なことが言える。評価する主体として、個人的な性格・環境・文化などの相違、また評価基準に対する認識のズレによって評価者内安定性と評価者間一致度<sup>3</sup>が低い可能性も看過できない。

しかし、継続的な訓練を通して、留学生の主観的な評価が内省と他者からのフィードバックとの相互関係によって、評価者内安定性と評価者間一致度が増し自律的な評価力が高くなるものなのであろうか。先述のように、村田（2004）は、主観的な評価活動を通して、内省とフィードバックとの相互作用の組み合わせは、留学生同士の観察・評価の視点を広げ、メタ認知力を高め、自律的な評価力を高めていくために役に立つと主張した。そこで、本研究は1年間を通して実施した「学習者参加型評価」を取り入れたクラス活動を通じて、留学生の前期と後期のテストの自己評価・他者評価の得点を用いて、その評価の差の有無、およびその変化について分析した。

まず、評価活動を実施した前期の時点では留学生による評価者内安定性と評価者間一致度について帰無仮説を立て検証し、評価者の級内相関係数<sup>4</sup>を計った。

仮説1：前期テストにおける留学生12名の自己評価と他者評価間に差がない。

仮説1を証明するには、前期テストにおける留学生12名の自己評価の得点と他者評価の得点の平均値を用いて、対応のあるt検定を行った。その結果、仮説1は棄却された ( $t(11) = -2.39, P < 0.05$ )。

表5 前期テストにおける自己評価と他者評価のt検定の結果

	N	平均	標準偏差	t	p
前期テスト自己評価	12	67.83	13.91	-2.39	0.036
前期テスト他者評価	12	74.67	7.33		

つまり、前期テストの時点では、留学生12名の自己評価と他者評価間に有意差が認められた。留学生の自己評価は他者評価より平均的に6.84点低い。さらに、前期テストの評価者一致度を見るのに級内相関係数を計ったところ、信頼係数は $r=0.663$ であり、やや低めである。自己評価と他者評価の間の差は留学生が自分自身を評価することへの不慣れ、また他者への評価に対する不安、および評価基準への認識の不明確なことから生じたと考えられる。前期テストの時点では、留学生の自己評価と他者評価における評価者内安定性と評価者間一致度が低いと言えよう。

そして、留学生の後期テストの得点を用いて、継続的な評価活動を終えた時点での自己評価と他者評価の変化についても分析した。

<sup>3</sup> 評価者間の採点結果が高い一致度を示すことと同一評価者が同一解答を採点した場合に採点結果のぶれが小さいことが望まれる。そのため、評価者内安定性と評価者間一致度指標が用いられる。

<sup>4</sup> 『研究社日本語教育辞典』により評価者内安定性と評価者間一致度を計るには級内相関係数を用いる。

仮説 2：後期テストにおける留学生 12 名の自己評価と他者評価間に差がない。

仮説 2 を証明するには、後期テストにおける留学生 12 名の自己評価の得点と他者評価の得点の平均値を用いて、対応のある t 検定を行った。その結果、2 つの評価に有意差が認められなかった ( $t(11) = -1.14, P > 0.05$ )。そして、後期テストの評価者の級内相関係数を計ったところ、信頼係数は  $r = 0.913$  となり、かなり高くなった。

表 6 後期テストにおける自己評価と他者評価の t 検定の結果

	N	平均	標準偏差	t	p
後期テスト自己評価	12	67.5	11.26	-1.14	0.281
後期テスト他者評価	12	71	5.94		

つまり、後期の個人発表テストが終わった時点では、留学生 12 名の自己評価と他者評価間に生じていた差が縮まった。継続的な自己評価と他者評価の訓練によって、留学生の主観的な評価が内省と他者からのフィードバックとの相互関係によって、評価者内安定性と評価者間一致度が高くなったと考えられる。

## 5.2 評価基準への認識

留学生による評価についてトムソン木下 (2008) は「従来教師の特権であった評価を学習者コミュニティーの中で学習者たちと共有し、学習者オートノミーを育てていく力を持っている」と述べていた。しかし、トムソン氏は「一般的な問題として、自己評価に対する教師や学習者の不信感が挙げられる」また、「(学習者) が自己評価に馴染みがなく、うまく対応できない場合などは、無理に自己評価を課しても成果が上がらないことがある」、さらに、「自己評価は本質的に絶対評価なので、大学など教育機関の評価規定と相いれない」と学習者の自己評価が直面する問題点についても指摘した。これらの問題点を俯瞰すると、自己評価だけに期待を寄せても学習者の「自律的評価力」の向上につながらないことが分かる。そこで、教師評価を学習到達点 (評価基準) への認識のバロメーターとして導入し、教師と学習者との相互作用によって、学習者の評価基準への認識が如何に変化したのかを検討したい。

仮説 3：前期テストにおける留学生 12 名の自己評価と他者評価は教師評価との間に差がない。

前期のスピーチテストにおける留学生 12 名の自己評価と他者評価の間に有意差があるため、それぞれの得点と教師評価の得点を用いて、対応のある t 検定を行った。その結果、自己評価と教師評価との間は  $t(11) = 2.72, p < 0.05$  となり、また他者評価と教師評価との間は  $t(11) = 5.43, p < 0.05$  となり、両方とも有意差が認められた。

表 7 前期テストにおける自己評価と教師評価の t 検定の結果

	N	平均	標準偏差	t	p
前期テスト自己評価	12	67.83	13.91	2.72	0.02
前期テスト教師評価	12	57.33	15.43		

表8 前期テストにおける他者評価と教師評価の t 検定の結果

	N	平均	標準偏差	t	p
前期テスト他者評価	12	74.67	7.33	5.43	0.000
前期テスト教師評価	12	57.33	15.43		

仮説4：後期テストにおける留学生12名の自己評価と他者評価、および教師評価との間に差がない。

後期テストにおける留学生12名の自己評価と他者評価の間に有意差が認められなかったため、それぞれの評価の得点と教師評価の得点を用いて、1要因で繰り返しのある分散分析を行った。その結果、3評価の差の主効果が認められなかった ( $F(2, 22) = 1.11, p > 0.05$ )。

表9 後期3評価の分散分析の結果

評価	平均値	標準偏差	N	F	p
自己評価	67.5	11.26	12	1.11	0.347
他者評価	71	5.94	12		
教師評価	68.67	7.09	12		

つまり、前期テストの時点では3評価間にズレが生じ、留学生の自己評価と他者評価の間に評価者内安定性と評価者間一致度が低く、また、評価基準に対する認識がまだ不十分であることが分かった。そして、後期テストの時点では3評価間に有意差が認められず、留学生による評価の評価者内安定性と評価者間一致度が高まり、教師評価に近づいたことは評価基準に対する認識が深まったと考えられる。

### 5.3 3評価による得点の変化

留学生12名の評価基準に対する認識の全体的な変化を確認するために、自己評価・他者評価、および教師評価を用いて、評価活動の回数が重なることで、それぞれの評価がどのように変化していたのかを分析した。前期テスト、および後期テストの3評価の評価回数ごとの得点の推移は以下のとおりである。(図3と図4に提示している評価回数は表1と表2に提示している発表順番に一致する。)

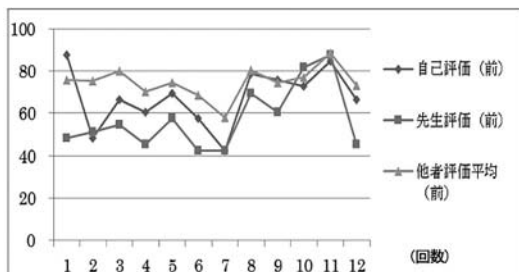


図3 前期テスト3評価の得点

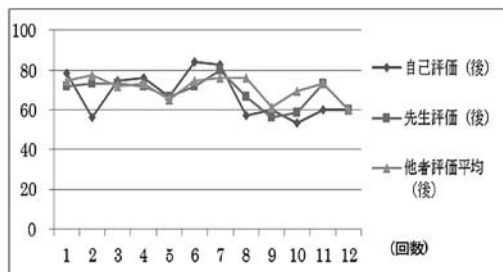


図4 後期テスト3評価の得点



図3を見ると、前期テストの時点では3評価の得点において、自己評価と他者評価は教師評価より高いことが分かる。その中、他者評価が最も教師評価から離れている。しかし、7回目付近から3評価の得点が接近しつつある。また、教師評価から依然として離れているが、自己評価と他者評価の間では距離が縮められている。図4を見ると、後期テストの時点では3評価の得点において、自己評価が教師評価からある程度上下に変動しているが、それに対し他者評価が教師評価と寄り添って変動していることが見られる。

以上、前期と後期2つのテストにおける3評価の回数ごとの得点の推移から、クラス活動を実施する回数を重ねていくうちに留学生12名の評価は教師評価に漸近する様子が見られる。つまり、留学生は自己評価と留学生同士による他者評価を通して、内省の機会と他者からのフィードバックを得られた。その結果、留学生がより多角的に自分自身の日本語能力を観察するようになり、評価の評価者内安定性と評価者間一致度も高くなった。さらに、教師評価を学習到達点（評価基準）への認識のバロメーターとして留学生に提示したことは、トムソン氏が指摘した自己評価に対する留学生の不安と不信感を軽減することにつながり、教師評価との相互作用によって留学生の評価基準への認識が深まったと考えられる。

## 5.4 学習効果と評価基準への認識との関係

### 5.4.1 全体的な学習効果

留学生12名の「日本語口頭表現」前期と後期の教師評価を用いて、学習効果を検証した。具体的には、前期テストと後期テストにおける教師評価の間に差があるかについてt検定を行った。

仮説5：前期テストと後期テストにおける教師評価の間には差がない。

その結果、有意差が認められ仮説5が棄却された ( $t(11) = -3.89, p < 0.05$ )。後期テストの平均得点が前期より11.34点有意に高かった。

表10 前期テストと後期テストにおける教師評価のt検定結果

	N	平均	標準偏差	t	p
前期テスト教師評価	12	57.33	15.43	-3.89	0.003
後期テスト教師評価	12	68.67	7.09		

### 5.4.2 学習効果について個別的分析

前期テストと後期テストにおける留学生12名のそれぞれ12回分の評価（自己評価と他者評価を含む）と教師評価の間に差があるかについてもt検定を行った。

仮説6：前期スピーチテストにおける留学生12名のそれぞれの評価と教師評価の間には差がない。

仮説7：後期個人発表テストにおける留学生12名のそれぞれの評価と教師評価の間には差がない。

分析した結果、前期テストにおいては、留学生のそれぞれの評価と教師評価との間

には全てにおいて有意差が見られた。それに対し、後期テストにおいては、留学生 12 名中 6 名の評価と教師による評価との間には有意差が見られた。(表 11、12 を参照)

表 11 前期テスト留学生のそれぞれの評価と教師評価との t 検定の結果

分析対象者 <sup>5</sup>	対応サンプルの差		t 値	N	p (両側)
	平均値	標準偏差			
ペア A	-15.17	15.58	-3.37	12.00	0.01
ペア B	-22.42	12.22	-6.35	12.00	0.00
ペア C	-20.92	13.06	-5.55	12.00	0.00
ペア D	-16.08	14.90	-3.74	12.00	0.00
ペア E	-10.00	9.14	-3.79	12.00	0.00
ペア F	-17.42	12.25	-4.93	12.00	0.00
ペア G	-15.67	14.74	-3.68	12.00	0.00
ペア H	-23.67	18.71	-4.38	12.00	0.00
ペア I	-13.08	12.22	-3.71	12.00	0.00
ペア J	-20.67	13.78	-5.20	12.00	0.00
ペア K	-21.67	12.03	-6.24	12.00	0.00
ペア L	-14.92	11.63	-4.44	12.00	0.00

表 12 後期テスト留学生のそれぞれの評価と教師評価との t 検定の結果

分析対象者	対応サンプルの差		t 値	N	p (両側)
	平均値	標準偏差			
ペア A	3.00	9.34	1.07	12.00	0.312
ペア B	-4.55	8.61	-1.75	12.00	0.110
ペア C	-6.91	7.03	-3.26	12.00	0.009
ペア D	-5.27	8.39	-2.08	12.00	0.064
ペア E	-10.91	6.77	-5.34	12.00	0.000
ペア F	-5.64	5.39	-3.47	12.00	0.006
ペア G	-7.64	4.48	-5.66	12.00	0.000
ペア H	13.00	5.59	7.72	12.00	0.000
ペア I	-3.45	5.16	-2.22	12.00	0.051
ペア J	6.27	3.90	5.33	12.00	0.000
ペア K	-3.18	11.45	-0.92	12.00	0.379
ペア L	-4.55	8.61	-1.75	12.00	0.110

先述したように、留学生のそれぞれの評価は、教師評価との差が縮まれば学習の到達点（評価基準）への認識が深まったと考えられる。そのため、仮説 7 の t 検定によって教師評価と有意差が見られなかった 6 名の留学生は、「学習者参加型評価」のクラス活動を通して、最終的に学習の到達点（評価基準）への認識が有意に深まったと考

<sup>5</sup> 分析対象者は留学生 12 名のそれぞれの評価と教師評価をペアで標記している。(例：留学生 A による 12 人分の評価と先生評価はペア A とする。) A~L は表 1 と表 2 の発表者と一致する。

えられる。そこで、学習の到達点（評価基準）への認識と学習効果とは一体どのように関係しているのかを分析した。

教師評価と有意差ありの留学生6名をグループ1とし、教師評価と有意差なしの留学生6名をグループ2とし、そして、2つのグループの学習効果について以下の3つの帰無仮説を立て検討した。

仮説8：前期テストにおける留学生グループ1とグループ2の教師評価の間には差がない。

仮説9：前期テストと後期テストにおける留学生グループ1に対する教師評価の間には差がない。

仮説10：前期テストと後期テストにおける留学生グループ2に対する教師評価の間には差がない。

学習効果の向上を促進する要因を検証するため、留学生グループ1とグループ2の前期テストと後期テストにおける教師評価を用いて、学習の到達点（評価基準）への認識が深まったこと、および前後のテストの違いという二要因の分散分析を行った。その結果、前期テストの時点では、留学生12名の学習効果間には差が見られず、また、留学生グループ1の6名の前期テストと後期テストの教師評価の間にも差が認められなかった。しかし、留学生グループ2の6名の教師評価は前期テストより後期テストの方が有意に高いことが確認できた。つまり、前後のテストの要因の主効果が認められず、学習の到達点（評価基準）への認識が深まったことの主効果が認められた ( $F(1, 5)=15.771, p<0.05$ )。

表 13 2 グループの学習効果における分散分析

	教師評価 (M+SD)		有意差の有無	F	
	前期テスト	後期テスト		テスト	交互作用
グループ1	60.67+16.94	69.67+7.09	15.771*	0.314	0.292
グループ2	54+14.49	67.67+7.61			

N=6

\* $p<0.05$

## 6. 結論

今回の事例研究は大学の日本語教育における「日本語口頭表現」の一年間のクラス活動の結果を数量的に分析したものである。大学の日本語教育に「学習者参加型評価」を導入したことで、留学生の学習の到達点（評価基準）への認識が深まったこと（仮説①自己評価・他者評価・教師評価間のズレを認識することによって、留学生は学習の到達点（評価基準）への認識が深まる。）を検証した。

まず、「学習者参加型評価」を実施するのにあたって、留学生による評価の評価者内安定性と評価者間一致度について検討した。帰無仮説1・2を検証した結果、前期テストの時点では、留学生による評価の評価者内安定性と評価者間一致度が欠けていたが、継続的な訓練によって後期テストでは、評価者内安定性と評価者間一致度が高

まったことが確認できた。

次に、「学習者参加型評価」を継続的に実施することで、学習の到達点（評価基準）への認識が深まることを検討した。帰無仮説 3・4 を検証した結果、前期テストの時点では、留学生の評価基準への認識がまだ不十分であったが、後期テストでは、評価基準に対する認識が深まったと考えられる。

さらに、前期テストおよび後期テストにおける 12 回分の 3 評価の得点推移図を用いて分析した。その結果、前期テストの得点は評価の 7 回目付近から 3 評価の得点が接近し始め、後期テストでは、3 評価の得点推移が非常に漸近する形となった。つまり、クラス活動の回数を重ねるうちに留学生の評価基準への認識が深まり、結果的に、継続的な「学習者参加型評価」の実施は学習の到達点（評価基準）への認識に繋がる傾向が見られた。

そして、「学習者参加型評価」を取り入れたクラス活動は如何に学習効果に影響を及ぼしたのか（仮説②「学習者参加型評価」の実施は留学生の学習効果の向上につながる。）について検証した。帰無仮説 5 を検証した結果、全体からみると、前期テストより後期テストの平均得点が 11.34 点も有意に高いことが確認できた。また、学習効果について個別に検討した。帰無仮説 6・7 を検証した結果、前期テストと後期テストにおける留学生による評価（自己評価と他者評価を合わせ）と教師評価のズレを検討し、評価基準に対する認識の変化が明確に見られなかった留学生 6 名をグループ 1、また認識の変化が明確に見られた留学生 6 名をグループ 2 に分別した。そして、帰無仮説 8・9・10 を検証した結果、2 グループの前期テストおよび後期テストにおける教師評価の間には差が見られなかったが、グループ 2 の 6 名の教師評価は前期テストより後期テストの平均得点が 13.67 点も有意に高いことが確認できた。つまり、全体的な後期テストの平均得点の向上に大いに貢献したのは認識の変化が明確に見られたグループ 2 の留学生であることが分かった。以上によって、学習の到達点（評価基準）への認識を深めることは、学習効果を高めることに繋がる可能性が証明された。

今回の研究は留学生の主観に基づく認識とその主体的・創造的参加を前提として、留学生が持つ思考・意思・価値観等を重視したクラス活動である。留学生が学習内容およびその評価基準を認識することで、「自律的評価力」を養成し授業への積極的な参加を経験して学習効果を高めたとと言える。

ドイツの心理学者レビンが事例研究を社会活動で生じる諸問題について小集団での基礎的研究でそのメカニズムを解明し、得られた知見を社会生活に還元して現状を改善することとして提唱している。今回の事例研究は留学生 12 名の 1 年間のデータを用いて数量的な検証を行った結果であり、普遍的な応用ができるかについては今後更なる研究が必要だと考えられる。また、30 回のクラス活動について具体的な質的分析は別稿に譲りたい。

## 参考文献

- 国際交流基金 (2011) 『学習を評価する—日本語教授法シリーズ 12—』 ひつじ書房
- 社団法人国際日本語普及協会 (2009) 「学習者参加型カリキュラムの開発—『リソース型生活日本語』の発展的活用を目指して—」平成 20 年度文化庁日本語教育研究委嘱「生活者としての外国人」のための日本語教育事業外国人に対する実践的な日本語教育の研究開発報告書
- 佐々木倫子 (2006) 「パラダイムシフト再考」国立国語研究所 (編) 『日本語教育の新たな文脈—学習環境, 接触場面, コミュニケーションの多様性』アルク, pp. 259-282
- 酒井たか子 (2003) 「日本語の学習指導を考える」『日本語教育のための心理学』海保博之・柏崎秀子編著、新曜社、初版第 2 刷、pp. 181-182
- 高木裕子 (1992) 「高次元での『自己評価表』を巡って」『関西外国語大学留学生別科日本語教育論集』3 号、pp. 77-98
- トムソン木下千尋 (2008) 「海外の日本語教育の現場における評価—自己評価の活用と学習者主導型評価の提案—」『日本語教育』136 号、pp. 27-37
- 細川英雄 (2002) 「合意形成としての評価—総合活動型日本語教育における教師論のために」『早稲田大学日本語研究教育センター紀要』15 号、pp. 105-117
- 細川英雄 (2004) 「新時代の日本語教育をめざして 第 9 回日本語教員養成における理論と実践の統合—早稲田大学大学院日本語教育研究科『実践研究』の試み」『日本語学』第 23 巻 第 12 号、pp. 68-78
- 森仁美・衣川隆生 (2010) 「自己モニタリングの基準の意識化を促進する他者評価の在り方」『日本語教育方法研究会誌』vol. 17 No. 2、pp. 26-27
- 村田晶子 (2004) 「発表訓練における上級学習者の内省とピアフィードバックの分析学習者同志のビデオ観察を通じて—」『日本語教育』120 号、pp. 63-72
- Council of Europe (2004) 『外国語の学習、教授、評価のためのヨーロッパ共通参照枠』、初版、吉島茂、大橋理枝 (訳、編)、朝日出版社

## 参考資料 1 (前期スピーチテストの評価シート)

発表者 \_\_\_\_\_ 評価者 \_\_\_\_\_  
インタビューの相手 \_\_\_\_\_

( ) の中に、評価 (5=非常によい、4=よい、3=普通、2=不足、1=非常に不足) を記入し、  
\_\_\_\_\_ に気付いたことや感想を書きなさい。

### I 準備

1. レジューメの作成について \_\_\_\_\_ ( )

### II 内容・構成

2. 背景や理由を述べたか \_\_\_\_\_ ( )

3. 内容の順序は分かりやすく構成されていたか \_\_\_\_\_ ( )

## 研究論文

4. インタビューの具体的な方法を述べたか \_\_\_\_\_ ( )
5. テーマに沿ってインタビューを実施したか \_\_\_\_\_ ( )
6. 結論や内容をうまくまとめたか \_\_\_\_\_ ( )

### Ⅲ発表技術

7. (質疑応答) 質問にうまく答えられたか \_\_\_\_\_ ( )
8. 用語の選び方は適当か \_\_\_\_\_ ( )
9. 声の大きさ・発話のスピードは適当か \_\_\_\_\_ ( )
10. 発音・アクセントについて聞き取りやすいか \_\_\_\_\_ ( )

## 参考資料 2 (後期個人発表テストの評価シート)

発表者:

発表テーマ:

評価者:

( ) の中に、評価 (5=非常によい、4=よい、3=普通、2=不足、1=非常に不足) を記入し、 \_\_\_\_\_ に気付いたことや感想を書きなさい。発表内容・構成について

### 発表内容・構成について

1. テーマは課題にあっているか \_\_\_\_\_ ( )
2. 全体の構成 (順序立てて話す) できているか \_\_\_\_\_ ( )
3. 論点・要点 (強調するところ) の提示をしたか \_\_\_\_\_ ( )
4. 説明 (事実と感想を別々に話すこと) は分かりやすいか \_\_\_\_\_ ( )
5. 発表の始め方と終わり方は適切であるか \_\_\_\_\_ ( )

### 態度と日本語運用力について

6. 発表時間は守っているか \_\_\_\_\_ ( )
7. 関連資料や図表・データなどの工夫があるか \_\_\_\_\_ ( )
8. レジューメは明確で、分かりやすいか \_\_\_\_\_ ( )
9. 参考文献は明確に提示したか \_\_\_\_\_ ( )
10. 声の大きさ・速度は適切で、はっきりと発音したか \_\_\_\_\_ ( )
11. 正し日本語で流暢に話したか \_\_\_\_\_ ( )
12. 最後まで丁寧語で話したか \_\_\_\_\_ ( )
13. 謙虚かつ誠実な態度で話したか \_\_\_\_\_ ( )
14. 間違った言い方に気づき自ら修正したか \_\_\_\_\_ ( )
15. 質疑応答はうまくできたか \_\_\_\_\_ ( )

## 付記

本稿の執筆にあたりご多忙の中ご指導くださいました西郡仁朗先生、貴重なご意見を頂きました査読委員の先生方に深く感謝いたします。

(おがわ みやこ・首都大学東京大学院博士後期課程)