

口頭能力と自己評価の関連性について

李 在鎬 (早稲田大学)、伊東 祐郎 (国際教養大学)、鎌田 修 (南山大学)
嶋田 和子 (アクラス日本語教育研究所)、坂本 正 (名古屋外国語大学)
六川 雅彦 (南山大学)、由井 紀久子 (京都外国語大学)

要旨

本研究グループでは、JOPT (Japanese Oral Proficiency Test: <http://jopt.jp>) という、口頭能力を評価するためのテストを開発している。本研究では、学習者の自己評価と JOPT と OPI の関連性を統計的に分析する。具体的な研究課題は、「学習者の自己評価と口頭能力にはどの程度、どのような関連があるのか」である。次の3点の調査結果を報告する。1) 「読むこと」と口頭能力の相関は強い ($r=.677$) が、「書くこと」と口頭能力の相関は弱い ($r=.286$) こと。2) 「書くこと」と「話すこと」は相関が弱く ($r=.171$)、「聞くこと」と「話すこと」は相関が強い ($r=.560$) こと。3) 口頭能力と自己評価の関連の度合いは「読む>話す>聞く>書く」の順であること。

【キーワード】 言語テスト、口頭能力、統計分析、自己評価

Keywords: Language Test, Oral Proficiency, Statistical Analysis, Self-Assessment

1 背景と目的

外国語教育における口頭能力の重要性は言うまでもないことである。しかし、その測定・評価の方法となると、信頼性と妥当性の確保、時間的制約等の理由で一般への共有化は非常に遅れている。この課題に対して、日本語教育分野において利用可能な口頭能力の測定ツールとして ACTFL (The American Council on the Teaching of Foreign Languages : 米国外国語教育協会) が開発した OPI (Oral Proficiency Interview, 鎌田ほか (編著) 2020) があり、評価の妥当性、教育実践への応用性の高さから、日本語教育分野においても広く知られている。ただ、その実施にはテストの資格取得・更新などが求められることにより、国内外の日本語教育現場において十分に実用化されているとは言えない。これらの課題を受け、本研究グループは 2013 年より科学研究費補助金基盤研究 (A) による支援を受け、JOPT (Japanese Oral Proficiency Test: <http://jopt.jp>, 李ほか 2019) と称する口頭能力を評価する言語テストを開発している。

本研究では、2018 年に収集したデータを使い、調査協力者の自己評価とアカデミック領域の JOPT のレベル判定と OPI のレベル判定の関連性を統計的に分析した結果を報告する。この試みは、日本語学習者の自己評価の傾向を複数の口頭能力評価と関連づけながら捉えることを目指すものであり、最終的には JOPT の妥当性を検討するものになると考えている。

2 先行研究

本研究は、口頭能力評価と自己評価の関連性を論じるものであり、本研究の調査結果を報告する前に、自己評価に関する先行研究をレビューする。とりわけ自己評価の有用性や妥当性に関する論考を紹介する。

まず、自己評価の有用性を示す研究として、McNamara & Deane (1995)では、自己評価は評価のタイプとしては学習者主導型評価であること、言語学習に対する責任感や自律心の形成において有効であることを指摘している。また、トムソン木下 (2008) では、自己評価が形成的評価のツールとして有用であることも論じている。これらの論考から、自己評価には、一定の有用性があることが確認できるであろう。

次に、自己評価の妥当性をめぐる研究では、学習者は自身の言語力をある程度、正確に自己評価できるとする分析 (Blanche & Merino 1989, Ross 1998) と、自己評価では言語能力を予測することができないとする分析 (玉岡ほか 2005) がある。どちらの主張もデータに基づくものであるため、一方が正しく、もう一方が正しくない結論づけることは難しい。こうした両面性の背景には、自己評価というものが持つ本質的な複雑性が関係している。というのは、自己評価は様々な要因に影響を受けるもので、複雑な心理現象だからである。例えば、伊東ほか (2008) では、タスクの経験有無によって自己評価が変わることを指摘している。また、近藤 (2012) では、年齢や学習動機や学習の不安感や文化的背景によっても自己評価が変わると指摘している。こうした多様な要素に影響を受けることから、質問の仕方を工夫する必要があること (Ross 1998 参照)、客観テストと組み合わせること (島田 2010 参照) の重要性も指摘されている。

以上で示した先行研究の指摘を踏まえた上で、以下では口頭能力評価における自己評価の有用性を検討する。先行研究に対して、本研究のオリジナルな観点として、主観評価である口頭能力評価と自己評価の関連性を検討するものであることや複数の口頭能力評価から自己評価を捉える研究であることが挙げられる。

3 データと分析方法

本研究で使用したデータは、李ほか (2018) の調査で収集したものである。具体的には2018年に浜松と東京で行ったものであり、50名の調査協力者が参加してくれた。本研究では、欠損値を除く43名の調査協力者の分析結果を報告する。

表1 調査協力者の属性

	男性	女性	合計
漢字圏	2	10	12
非漢字圏	16	15	31
合計	18	25	43

表1は、調査協力者の属性として「性別*漢字圏・非漢字圏」でクロス集計したものである。なお、調査協力者の国籍としては、「アメリカ、インド、インドネシア、タイ、フィリ

ピン、ブラジル、ベトナム、ペルー、ミャンマー、レバノン、韓国、香港、台湾、中国」の14カ国である。これらの調査協力者に対して、JOPTとOPIを実施した。そして、日本語を使った4技能として、「読むこと」、「書くこと」、「聞くこと」、「話すこと」に対して、「全くできない」、「できない」、「できる」、「よくできる」の4スケールで自己評価してもらった。

こうした手続きで収集したデータに対して、相関分析とクロス集計を行った。

4 結果と考察

自己評価に関する具体的な結果を報告する前に、JOPTとOPIの対応を示す(表2)。なお、JOPTの特徴については、李ほか(2019)を、JOPTのレベル判定については、李ほか(2018)を参照してほしい。

表2 JOPTとOPIのクロス集計

		OPI			合計
		初級	中級	上級	
JOPT	低	5	6	0	11
	中	0	9	1	10
	上	0	2	6	8
	優	0	1	13	14
合計		5	5	18	43

表2では、JOPTの4スケール(低、中、上、優)とOPIの3スケール(初級、中級、上級)の対応をクロス表で示した。「JOPT中」と「OPI中級」、「JOPT優」と「OPI上級」のように顕著な対応があるところを灰色で示す。

次に、自己評価データの信頼性を確認するために、Cronbachのアルファ係数を計算した。結果は $\alpha=0.704$ となり、一般的に「信頼性がある」とされる0.8を下回っているものの、項目数が少なかったことなどを考慮すれば、まずまずの結果であると言える。これを踏まえて、さらに属性間で極端な差がないか確認した。

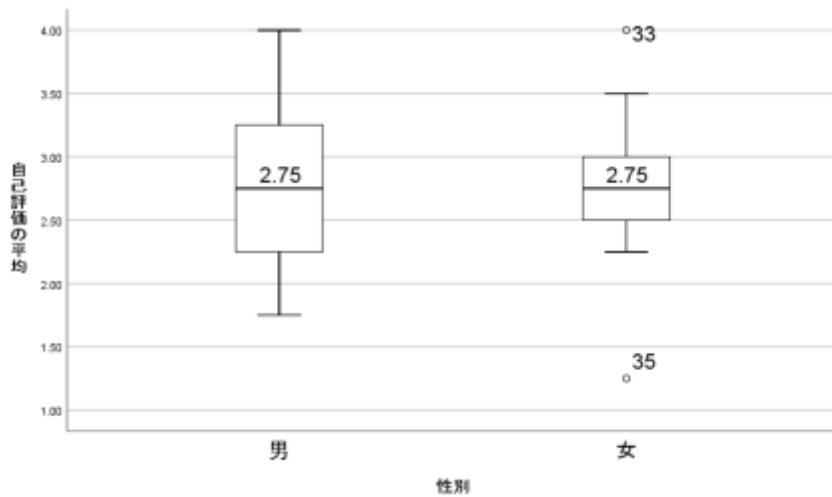


図1 性別による自己評価の箱ひげ図

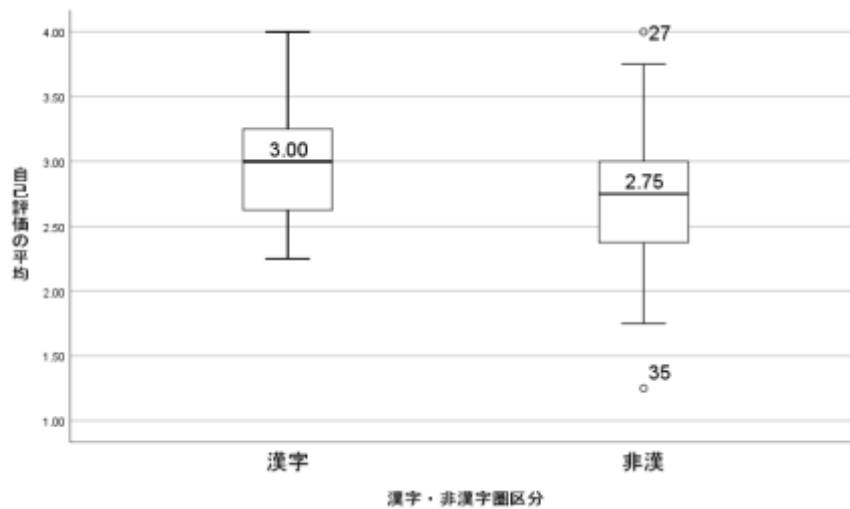


図2 漢字圏と非漢字圏の区分と自己評価の箱ひげ図

図1と図2は調査協力者の属性による自己評価の分布を示したものであるが、「全くできない=1、できない=2、できる=3、よくできる=4」として、箱ひげ図を作成した。図1の性別を見た場合、中央値としては男性も女性も2.75となり、性別による評価結果には大差はない。図2の漢字圏と非漢字圏に関しては漢字圏のほうが3.0と評価し、非漢字圏のほうは2.75と評価しており、漢字圏がやや高めに評価していることが明らかになった。図1と図2によって属性による極端な差は見られないことを確認し、自己評価の平均と標準偏差を求めた。

表3 4技能の自己評価の平均値と標準偏差

	平均値	標準偏差
自己評価（読む）	2.74	0.848
自己評価（聞く）	3.05	0.688
自己評価（書く）	2.49	0.798
自己評価（話す）	2.93	0.737

表3の平均値を見ると、「聞くこと」に対する評価値がもっとも高く、「書くこと」に対する評価値がもっとも低い。この結果は、本研究の調査協力者は「聞くこと」に対して、相対的に自信を持っているのに対して、「書くこと」に対しては、苦手意識を持っている可能性を示唆する。

次に自己評価と口頭評価の対応を見るためにスピーアマンの順位相関係数による相関係数を計算した。

表4 口頭評価と自己評価の相関

	OPI	JOPT	自己評価 （読む）	自己評価 （聞く）	自己評価 （書く）
JOPT	.827**				
自己評価（読む）	.584**	.677**			
自己評価（聞く）	.278	.302*	.492**		
自己評価（書く）	.147	.286	.424**	.312*	
自己評価（話す）	.101	.348*	.313*	.560**	.171

#**：相関係数は 1% 水準で有意（両側） *：相関係数は 5% 水準で有意（両側）

表4では、OPI（3スケール：初級=1、中級=2、上級=3）とJOPT（4スケール：低=1、中=2、上=3、優=4）と自己評価（4スケール：全くできない=1、できない=2、できる=3、よくできる=4）の相関を示しているが、口頭能力評価と自己評価の相関に注目した場合、以下の3点が言える。

- 1) OPIに関しては「読むこと」と相関しているのに対して、JOPTに関しては、「読むこと」「聞くこと」「話すこと」と相関している。
- 2) JOPT が評価する口頭能力と、自己評価の「書くこと」はどちらも産出活動であることを考えると、口頭能力が高ければ高いほど、「書くこと」に対しても、自己評価が高いことが予想されるが、表4の結果は、それを支持しない。
- 3) 自己評価同士の相関を見ると、「書くこと」と「話すこと」は、どちらも言語産出に関わる活動であるにも関わらず、相関が低い ($r=.171$)。一方、「話すこと」と「聞くこと」は産出と理解に関わる別の活動であるにも関わらず、相関が高い ($r=.560$)。

以上の観察から得られる帰結として、「話すことができる=書くことができる」という単純

化は成立しないことが明らかになった。

次に、自己評価と口頭評価の具体的な関連を調査した。

表5 自己評価（話す）と JOPT のクロス表

		自己評価（話す）				合計
		全くできない	できない	できる	よくできる	
JOPT	低	0	5	6	0	11
	中	0	2	5	3	10
	上	1	2	4	1	8
	優	0	1	8	5	14
合計		1	10	23	9	43

表6 自己評価（話す）と OPI のクロス表

		自己評価（話す）				合計
		全くできない	できない	できる	よくできる	
OPI	初級	0	2	3	0	5
	中級	0	4	10	4	18
	上級	1	4	10	5	20
合計		1	10	23	9	43

表5と表6では、口頭能力評価と特に関連が深い「話すこと」に対する自己評価と JOPT と OPI のクロス表を示した。特に顕著な対応がある部分を灰色で示す。このデータから、JOPTの低レベルと OPIの初級は、「できない」と「できる」が同比率で回答しているのに対して、レベルがあがるにつれ、「できる」「よくできる」の比率が上がっていることが確認できる。

次に、「読むこと」「書くこと」「聞くこと」と JOPT の評価の関連を調査した。

表7 自己評価（読む）と JOPT のクロス表

		自己評価（読む）				合計
		全くできない	できない	できる	よくできる	
JOPT	低	0	9	1	1	11
	中	3	4	3	0	10
	上	0	0	7	1	8
	優	0	0	8	6	14
合計		3	13	19	8	43

表8 自己評価（書く）と JOPT のクロス表

		自己評価（書く）				合計
		全くできない	できない	できる	よくできる	
JOPT	低	2	4	4	1	11
	中	1	6	3	0	10
	上	1	4	3	0	8
	優	0	4	7	3	14
合計		4	18	17	4	43

表9 自己評価（聞く）と JOPT のクロス表

		自己評価（聞く）				合計
		全くできない	できない	できる	よくできる	
JOPT	低	1	0	9	1	11
	中	1	3	4	2	10
	上	0	0	7	1	8
	優	0	0	9	5	14
合計		2	3	29	9	43

表7の「読むこと」に関しては、JOPTのレベルとほぼ対応しているのに対して、表8の「書くこと」に関してはJOPTのレベルが低い人（口頭能力が低い人）でも、「できる」という回答が多く、JOPTのレベルが高い人（口頭能力が高い人）でも、「できない」という回答が多いことから、明確な対応はないと解釈できる。これは、表4の「書くこと」とJOPTの相関が低い事実とも合致する結果である。最後に、聞くことに関しては、口頭能力に比べ、自己評価が高めであることが確認できる。

5 まとめ

本研究は、日本語学習者の自己評価と口頭能力にはどのような関連があるのかを調査したものであり、以下の3点が明らかになった。

1. 「読むこと」と口頭能力の相関は強い ($r=.677$) が、「書くこと」と口頭能力の相関は弱い ($r=.286$)。
2. 「書くこと」と「話すこと」は相関が弱く ($r=.171$)、「聞くこと」と「話すこと」は相関が強い ($r=.560$)。
3. 口頭能力と自己評価の関連の度合いは「読む>話す>聞く>書く」の順である。

*謝辞

本研究は科研費（基盤研究（A）25244023、17H00919）の研究成果である。

<参考文献>

- 伊東田恵・川口恵子・太田理津子 (2008) 「外国語能力の自己評価における言語タスク経験の有無」『日本言語テスト学会研究紀要』 11, pp.156-172.
- 鎌田修・嶋田和子・三浦謙一 (編著) (2020) 『OPIによる会話力の評価』 凡人社.
- 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.
- 島田めぐみ (2010) 「自己評価 Can-do statements に関する一考察：客観テストとの比較を通して」『東京学芸大学紀要 総合教育科学系』 61(2), pp.267-277.
- 玉岡賀津雄・松下達彦・元田静 (2005) 「日本語版 Can-do Scale はどれくらい正確に日本語能力を測定しうるか」『広島大学留学生教育』 9, pp.65-78.
- 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, pp. 27-37.
- 李在鎬・嶋田和子・伊東祐郎・鎌田修・坂本正・由井紀久子・六川雅彦 (2018) 「口頭能力テスト「JOPT」と「OPI」の対応に関する調査報告」(日本語教育学会 2018 年度秋季大会) .
- 李在鎬・伊東祐郎・鎌田修・坂本正・嶋田和子・西川寛之・野山広・六川雅彦・由井紀久子 (2019) 「日本語口頭能力テスト「JOPT」開発と予備調査」『日本語プロフィシエンシー研究』 (7), pp.28-49.
- Blanche, Patrick & Merino, B. J. (1989) Self-Assessment of Foreign Language Skills: Implications for Teachers and Researchers. *Language Learning* 39(3): 313-340.
- McNamara, M., & Deane, D. (1995). Self-assessment activities toward autonomy in language learning. *TESOL Journal* 5: 18-23.
- Ross, S. (1998). Self-assessment in second language testing: A meta-analysis and analysis of experiential factors. *Language Testing* 15: 1-20.