

学習者主導型評価をもとに言語能力を探る—Can do テストの可能性  
Possibilities of learner-directed assessments: A case for a can-do test

三ツ木紗奈子、カンザス大学  
Sanako Mitsugi, University of Kansas

## 1. はじめに

言語教育は学習者に言語自体を教えるだけでなく、自ら言語学習を続けていける技術と能力をつけていくことを目標とすべきだと言われている (Nunan, 1988)。そのためには、学習者が自分自身の「学び」に自覚的になり、それを把握、内省することが求められている。第二言語教育において、このような学習者主導型評価の必要性が強調されて久しい (e.g., McNamara & Deane, 1995)。日本語教育では、近年、ポートフォリオやジャーナル、ブログなどを用いた学習者参加型評価の実践が多く取り上げられている。

自己評価は学習者の自立心の形成、及び、学習意欲の向上につながるとされ、その質的側面の可能性に関心が高まっている (青木 2008)。しかしながら、自己評価の意義は認めつつも、信頼性や妥当性に疑いを持つ教師や研究者は少なくない (トムソン木下 2008)。自己評価の信頼性については、多くの研究がなされてきた。学習者による自己評価の結果と教師による評価との相関を調べた Oskarsson (1981) の研究では、両者に有意な相関があったとしているが、そうではないとする研究もあり (Blanche & Merino, 1989)、一貫した答えは出ていない。日本語教育では、自己評価の結果と実際の言語能力が必ずしも一致するとは限らないと指摘する研究や (玉岡・松下・元田 2005)、客観的評価や教師評価との相関はなかったとする研究がある (Thomson, 1996)。研究デザインや評価のツールの違いにより、一元的な比較が難しいことに加え (Bachman & Palmer, 1989)、自己評価には、学習者の心理状態、過去に経験したタスクや文化的価値が付随する可能性があることがこの研究分野を複雑にしているだろう (Thomson, 1996; 玉岡・松下・元田 2005)。

## 2. 研究の目的と方法

### 2.1. 調査の目的

本研究は、Can-do テスト形式の自己評価ツールの妥当性、有用性を考察する。主に、以下の2点を検証した。(1) Can-do テスト形式の自己評価と客観的評価(読解テスト、口頭テスト)との間に関係性があるか。(2) Can-do テストによって、学習者の日本語能力を推計することができるか。

### 2.2. 被験者

カンザス大学で日本語を学んでいる学部生 36 名を対象に調査を行った。被験者の性質理解のため、簡単な言語学習背景アンケートを行った。その結果は表 1 に報告した通りである。被験者は日本語 2 年、3 年後期修了間近の学生で、日本語専攻の学生 18 名を含む。第一言語は、1 名を除き、全員が英語であった。調

査に必要なタスクを完了しなかった被験者 2 名と第一言語が非英語の被験者を除き、最終的に 33 名を分析対象にした。

表 1. 被験者の背景情報

被験者係数	人数
	総数 33
性別	女性 19 (57.5%)
	男性 14 (42.4%)
専攻	日本語専攻 18 (54.5%)
	その他 15 (45.4%)
履修授業	2年後期 25 (75.7%)
	3年後期 8 (24.2%)
入学前の日本語学習	なし 20 (66.6%)
	あり 13 (39.3%)

### 2.3. Can-do テスト

項目の選択には、Kondo-Brown (2005) の自己評価項目を元に作成した (付録 1)。口頭能力について評価する 12 個の質問を用意した。被験者はタスクがどの程度出来るかを「全く出来ない (Not at all)」、「非常に困難である (With great difficulty)」、「難しい (With some difficulty)」、「比較的良くできる (Relatively well)」、「簡単にできる (Quite easily)」の五段階で評価した。Can-do テストは言語背景アンケートと共に、口頭テスト実施の直前に行った。

### 2.3. 読解テスト

読解テストは近藤ブラウン (2012, pp. 194) で報告されているハワイ大学日本語プログラム到達度評価用のテストを参考に作成した。本校のカリキュラムを考慮し、テスト項目が本校の学習目標基準、学生のレベルに合うと思われるか、日本語科の学科長、及び、日本語プログラムの教員からフィードバックを得ながら推敲した (付録 2)。このテストは 6 題の四肢選択問題から成り、学期末最終週に授業時間の 15 分を使って、クラスごとに一斉実施した。

### 2.4. 口頭テスト

本校では、アメリカ教育省の Title VI の東アジア研究の一貫として、2 年生の日本語からそれより上の日本語を履修する学生を対象に ACTFL OPI の言語能力基準に則った口頭面接を行っている<sup>1</sup>。本研究では、この試験の判定を口頭能力のインデックスとして使用した。口頭テストは、日本語プログラムの院生アシスタントが、授業外で個別に行った。

<sup>1</sup> ACTFL OPI では、トレーニングを受けた認定試験官によって、定められた基準に沿って、面接の実施と判定を行わなければならない。本校の口頭面接は、それらの必要条件を全て満たしているわけではない。正規 OPI と区別するため、「口頭テスト」と呼ぶことにする。

### 3. 結果

#### 3.1. 評価ツールの結果

はじめに、Can-do テスト結果を報告する。質問項目のクロンバックアルファ係数は非常に高いものだった ( $\alpha = .937$ )。これより、項目間には内的一貫性があり、信頼性が高いテストだったと言える。次に Can-do テストの項目をタスクの難易度の観点から、4つに分類した (レベル1: 項目 1-3、レベル2: 項目 4-6、レベル3: 項目 7-9、レベル4: 項目 10-12)。この分類で計算した平均値、及び、標準偏差は、表2の通りである。平均値を目視すると、レベルが上がるにつれて、Can-do テストの得点は低くなっているのが分かる。観察された差が統計的に有意であるか、一元配置の分散分析を使って比較したところ、平均値には有意な差があった ( $F(3, 96) = 167.61, p < .001$ )。ボンフェローニ補正法を用いて、多重比較検定を行った結果、全てのレベル間で平均値に有意差があることが検証できた ( $ps < .001$ )。つまり、自己評価の結果は項目の難易度を具現していると言える。

表2. 自己評価のタスク平均得点と標準偏差

レベル	口頭能力	平均値	標準偏差
1	基礎的な語彙、語句の産出ができる。	13.3	1.44
2	初歩的な語句や定型表現を使って、話すことができる	11.06	1.69
3	生産的に言語を使って会話できる。自分の身の回りのトピックについて、質問にその場で対応できる。	8.69	2.03
4	客観的、抽象的なトピックについてまとまった会話ができる。レジスターを効果的に使い、説明ができる。	5.51	1.9

次に、読解テストの結果を報告する。読解テストは、6点満点に対して、平均が4.21点、標準偏差が1.51であった。平均から±標準偏差の範囲をばらつきの指標とみなして、算出すると3点から5点となった (正確には2.71点から5.71点)。その範囲内には、63.6%にあたる21名が入っていた。このことから、テストの判別力は妥当だったと言えるだろう。

最後に口頭テストの結果を報告する。全体の78.7%の被験者がOPIの初級にあたる1から3点をとっていた。そのうち、一名を除く全員が2年修了直前の被験者だった。日本語教育学会が期待する「基礎レベルを超えた日本語能力」は中級一下であるとされているが、3年修了時になると、87.5%の学生が中級のレベルに達していた。口頭テストの結果と対応するOPI基準、被験者の分布数は表3に示した通りである。

表 3. 口頭テスト得点分布

対応OPI判定	得点	クラス配置			
		2年生	(%)	3年生	(%)
初級一下	1	14	(56)	0	-
初級一中	2	8	(32)	0	-
初級一上	3	2	(8)	1	(12.5)
中級一下	4	1	(4)	4	(50)
中級一中	5	0	-	2	(25)
中級一上	6	0	-	1	(12.5)

## 3.2. テスト間の関係性

次に、本研究で用いた3つの評価ツール間の関係性を検証した。読解テストの得点、口頭テストの得点、Can-do テストの総得点、Can-do テストの4レベルの得点を係数にピアソン積分相関係数を算出した。口頭テストと読解テストの二組の得点には、強い正の相関関係があった ( $r < .494, p = .003$ )。このことから、読解テストの得点が高くなればなるほど、口頭テストの得点も高くなる傾向にあると言える。これらのテストは、産出に関する「話す」技能と知覚に関する「読む」技能という、異なる言語能力を評価するものなので、相関関係があることは必ずしも直感的ではないかもしれない。

一方で、Can-do テストと口頭テストが共に「話す」技能を評価するツールであるので、2組の得点の間には相関があると推測したが、Can-do 総得点と口頭テストの得点間には、弱い正の相関の傾向が見られたのみで、この傾向は統計的に有意ではなかった ( $r < .304, p = .086$ )。更に、Can-do テストの総得点は、読解テストの得点とも、有意な相関がなかった ( $r < .050, p = .784$ )。つまり、Can-do テストを行っただけでは、学習者の知覚技能も産出技能も分かりかねるという結果となった。相関関係のデータは表4に示した。

表 4. 評価ツール間の相関

	口頭テスト	読解テスト	Can-do合計	レベル1	レベル2	レベル3	レベル4
口頭テスト		.494 **	.304	.255	.459 **	.133	.049
読解テスト	.494 **		.050	-.016	.068	.082	-.007
Can-do合計	.304	.050		.500 **	.730 **	.776 **	.752 **
レベル1	.255	-.016	.500 **		.489 **	.043	.066
レベル2	.459 **	.068	.730 **	.489 **		.340	.280
レベル3	.133	.082	.776 **	.043	.340		0.621 **
レベル4	.049	-.007	.752 **	.066	.280	.621 **	

\*\* Correlation is significant at the 0.01 level (2-tailed).

Can-do テストのレベル別の得点を見ると、レベル2の得点と口頭テストの得点には正の相関があることが見てとれる ( $r = .459, p = .007$ )。Can-do テスト総得点との間にはなかった、何らかの関係が見受けられる。相関関数からは、二つ

の変数間の関係以上のことは分かりかねるので、Can-do テストレベル 2 の得点を説明係数、口頭テストの得点を非説明係数とし、単回帰分析を行った ( $R^2 = .211$ ,  $F(1,32) = 8.291$ ,  $p = .007$ )。検定によると、重決定係数(寄与率)は低く、Can-do テストレベル 2 の得点は、口頭テストの結果の推計に役に立つ変数ではないという結果となった ( $\beta = .459$ ,  $t = 2.879$ ,  $p = .007$ )。

次に Can-do テストレベル 2 の得点に加え、先の相関分析で口頭テストと有意な相関があった読解テストの得点を加えて、重回帰分析(強制投入法)を行った<sup>2</sup>。分析の結果、読解テストの得点 ( $\beta = .465$ ,  $t = 3.358$ ,  $p = .002$ ) と Can-do テストレベル 2 の得点 ( $\beta = .428$ ,  $t = 3.087$ ,  $p = .004$ ) が口頭能力を説明する効率的な説明変数で、その重回帰式は当てはまりがいいということが分かった ( $R^2 = .427$ ,  $F(2,32) = 11.156$ ,  $p < .001$ )。Can-do テストレベル 2 の得点を元にした単回帰式モデルは口頭テストの得点の分散の 21.1%しか推計できなかったのに対して、読解テストの得点を変数として加えた重回帰式は 42.7%の分散を予測できることが分かった。

#### 4. 考察と結論

本研究では、学習者主導型の評価ツールとである Can-do テストの妥当性、有用性を検証した。使用した Can-do テストは、その内部一貫性と適切な難易度推移にも関わらず、学習者の口頭能力を十分に予測できるツールとは言えないことが明らかになった。Can-do テストの難易度別 4 レベルの下位群を分析した結果、レベル 2 の「初歩的な語句、定型表現を使って話すことができる」という項目群の得点が口頭テストに多少関わっていることが分かった。しかしながら、レベル 2 の得点が口頭テストに及ぼす純粋な影響力を検証したところ、十分に推計できるものとは言えなかった。しかし、Can-do テストレベル 2 の得点を読解テストの得点と組み合わせて検定した際には、口頭テストの結果を推定できるということが明らかになった。

Johnstone (2000) は Can-do テストなどの「代替アセスメント」は客観的評価に単一で替わることは出来ないが、他の評価ツールと組み合わせることで、効果的なツールとなると議論している。また、近藤ブラウン (2012) は、学習者の言語能力の判定は単一のテストに頼らず、複数の情報を元に判断することで、妥当性や信頼性の高い評価が行えるとしている。本研究の結果は、自己評価と客観的評価を組み合わせることで、学習者の口頭能力をより効率的に推定できることを示し、これらの主張を支持する結果だと言える。

Can-do テストの項目は、日本語の達成度を中心においた教育指標と考えることもできる。学習者にとって、このような汎言語的な基準に照らして自分の実力がどこにあるのか評価する経験は有益だろう。しかしながら、アメリカの日本語教育の現場では、日本語話者との接触を通じて可能となるインフォーマルな自己評価の機会が、日本の日本語学習者と比べて、少ないのが必然であろう。本研

<sup>2</sup>表 4 に示す全ての係数を説明変数として、ステップワイズ法で重回帰分析も行った場合、Can-do テストレベル 2 と読解テストの得点のみが回帰式に投入された。その他の変数は有意な変数とならなかった。

究の被験者の中には、評価すべき「タスクを経験したことがないので、わからない」という意見もあった。自己評価の機会が少ないことは、過大評価や過小評価を招き、ツールの信頼性の低下につながりかねない。普段の教室活動から、ロールプレーなどで表現学習の練習を行ったり、成果、内省、学習過程のサイクルの機会を増やしたりすることで、自己評価をより効率的に活用できるようになるだろう。

## 参考文献

- Bachman, L., & Palmer, A. (n.d.). The construct validation of self-ratings of communicative language ability. *Language Testing*, 6(1), 14–25.
- Johnstone, R. (2000). Context-sensitive assessment of modern languages in primary (elementary) and early secondary education: Scotland and the European experience. *Language Testing*, 17(2), 123–146.
- Kondo-Brown, K. (2005). Differences in language skills: Heritage language learner subgroups and foreign language learners. *The Modern Language Journal*, 89(4), 563–581.
- McNamara, T., & Deane, D. (1995). Self-assessment activities: Toward autonomy in language learning. *TESOL Journal*, 5(1), 17–21.
- Nunan, D. (1988). *The Learner-Centered Curriculum*. Cambridge: Cambridge University Press.
- Oskarsson, M. (1981). Subjective and objective assessment of foreign language performance. In R. J. A. (Ed.), *Directions in Language Testing* (pp. 225–239). Singapore: Singapore University Press.
- Thomson, C. K. (1996). Self-assessment in self-directed learning: Issues of learner diversity. In R. Pemberton (Ed.), *Taking control: autonomy in language learning* (pp. 71–91). Hong Kong: Hong Kong University Press.
- 近藤ブラウン妃美 (2012). 『日本語教師のための評価入門』 東京: くろしお出版.
- 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」 『日本語教育』, 136号, 27–37.
- 小山悟 (1996). 「自立学習促進の一助としての自己評価」 『日本語教育』, 88号, 91–103.
- 青木直子 (2008). 「日本語を学ぶ人たちのオートノミーを守るために」 『日本語教育』 138号, 33–42.
- 玉岡賀津雄・松下達彦・元田静香 (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」 『留学生教育: 広島大学留学生教育センター』 9号, 65–78.

付録 1

Self-ratings of oral tasks

Please indicate how well you can carry out the following tasks in Japanese using the 1-5 scale below.

- 1- Not at all
- 2- With great difficulty
- 3- With some difficulty
- 4- Relatively well
- 5- Quite easily

Item

- 1) [ ] Count to 10 in the language.
- 2) [ ] Say the days of the week.
- 3) [ ] Give the current date (month, day, year).
- 4) [ ] Introduce myself in social situations, and use appropriate greetings and leave-taking expressions.
- 5) [ ] Order a meal in a restaurant.
- 6) [ ] Give biographical information about myself (place of birth, composition of family, early schooling etc.).
- 7) [ ] Give directions on the street.
- 8) [ ] Describe my present job, studies, or other major life activities accurately and in detail.
- 9) [ ] Sustain everyday conversation in casual style Japanese with my Japanese friend.
- 10) [ ] Sustain everyday conversation in honorific/humble polite style Japanese with a person older than I am.
- 11) [ ] Describe the U.S. educational system in some detail.
- 12) [ ] State and support my position on a controversial topic with examples and reasons (e.g., cigarette smoking, birth control, environmental pollution).



付録2

Class: ( )  
 Name: ( )

Read the following email message from Yoko (KU student) to her friend, Haruko. Choose the best answer according to the content of the message and place a checkmark (✓). Yoko wrote this message to Haruko in early April. If you do not know the answer, make your best guess. You will have 10 minutes.

春子さん  
 お元気ですか。アルバイトは忙しいですか。  
 こちらは春休みも終わり、もうすぐ期末試験が始まります。今学期は遊んでばかりいたの、試験には受からないかもしれません。スプリングブレイクには、友達とシカゴに行ってきました。レンタカーをかりて、ドライブに出かけたり、買い物をして楽しんでました。でも、最後の日はホテルで寝ていました。ミシガン湖へ行くつもりだったんですが、前の晩ビールをたくさん飲まされて、具合がわるくなってしまったんです。  
 ところで、カンザス大学の日本文化クラブで、おり紙を教えることになりました。楽しみです、うまくいかどうかちょっと心配です。日本人だからといって、じょうずにおり紙が教えられるわけじゃないです。おり紙の本を買おうと思って、こちらの本屋でさがしてみたんですが、いいものはありません。悪いです、夏休みはこちらへ来る時におり紙の本を買って来てもらえませんか。無理言って、すみません。では、夏休みに、春子さんに会えるのを楽しみにしています。

よう子

**1. This semester, Yoko:**

- a. ( ) Seems to have studied hard.
- b. ( ) Does not seem to have studied hard.
- c. ( ) Seems to have been busy with her study.
- d. ( ) Does not seem to have been busy with her part-time job.

**2. What does Yoko say about her final exams?**

- a. ( ) She might fail the exams.
- b. ( ) She will definitely fail the exams.
- c. ( ) She might pass the exams.
- d. ( ) She will definitely pass the exams.

**3. On the last day of her Spring break Yoko:**

- a. ( ) Went shopping.
- b. ( ) Slept in the hotel room.
- c. ( ) Rented a car and went driving.
- d. ( ) Reluctantly drank a lot of beer.

**4. In what context, did Yoko Mention “Michigan Lake”?**

- a. ( ) She went there.
- b. ( ) She did not plan to go there.
- c. ( ) She ended up not going there.
- d. ( ) She got sick while driving there.

**5. Yoko is worried about her origami lessons because:**

- a. ( ) She has never taught origami before.
- b. ( ) Even though she is Japanese, she is not good at origami.
- c. ( ) Some Japanese students might find her origami lessons boring.
- d. ( ) Being Japanese does not mean that she can teach origami well.

**6. Which of the following statement is true about Yoko’s summer break?**

- a. ( ) Yoko will visit Haruko in Japan.
- b. ( ) Haruko will meet Yoko in Kansas.
- c. ( ) Haruko will teach origami at the Japanese Culture Club.
- d. ( ) Haruko has offered to give Yoko an origami book as a souvenir.