

CEFRチェックリストを使った日本語能力の自己評価の変化

坂野 永理・大久保 理恵

Changes in Self-Assessment of Japanese Language Ability
Using the CEFR Self-Assessment Checklists
Eri BANNO, Rie OKUBO

要旨

本稿ではCEFRのスイス版自己評価チェックリストを使い、岡山大学の日本語コース受講者86名を対象に、学期開始後から終了前の約3か月の期間で自己の日本語力の変化を感じているかどうか、またその変化はどの言語領域において現れるのかを調査した。調査の結果、初級、中級、中上級の全てのレベルにおいて、学生の自己評価の平均値が学期開始後より終了前の方が有意に高いという結果が現れた。また、各レベルの全ての領域で終了前の方が開始時より自己評価の平均値が有意に高く、全ての項目で終了前のほうが開始時より平均値が高いという結果となった。本調査により、受講生は学期終了時には自己の日本語力をより肯定的に評価するようになったことが明らかになった。

キーワード：自己評価、日本語、CEFR、Can-do statements

1. はじめに

岡山大学の日本語コースは、平成24年現在、初級1、初級2、中級入門、中級1、中級2、上級1、上級2まで7レベルのクラスからなる。各クラスは一学期15週間で、一週間に90分授業4コマからなり、合格すれば次学期に一つ上のレベルのクラスが履修できることとなっている。ただし、最上級の上級2に関しては、開講クラスは週4コマであるが、週1コマでも履修可能である⁽¹⁾。学習者が一学期の間に次のレベルに行く日本語力を得たかどうかは、クラス内での試験等により判断されるが、学習者自身が自己の日本語力についての伸びをどう感じているのかは定かではない。

近年Can-do statementsによる学習者の自己評価が、日本語教育においても注目を浴びている。外国語教育におけるCan-do statementsは、学習者が目標言語でできることを記述したもので、「かんたんな自己紹介ができる」という易しいものから、「ネイティブスピーカー同士の活発な会話についていくことができる」という難易度の高いものまで様々な難易度の項目があり、「話す」「読む」など領域も多岐にわたる。

日本語教育においては、JF日本語教育スタンダード(国際交流基金, 2010)やアカデミック・ジャパニーズに焦点を置いたJLC日本語スタンダード(東京外国語大学留学生日本語教育センター, 2011)など、機関が独自で作成したものも公開されているが、外国語学習のCan-do statementsで最も知られているのは、CEFR(ヨーロッパ言語共通参照枠)によるものである(吉島・大橋他, 2004)。CEFRは、2001年に欧州評議会の言語教育政策を基に作成されたものであり、

ヨーロッパの外国語教育現場で広く普及している（細川・西山，2010）。

本調査は、CEFRのスイス版自己評価チェックリストを使い、岡山大学の日本語コース受講者を対象に、学期開始時から終了前の約3か月の期間で、自己の日本語力の変化を感じているかどうか、またその変化はどの言語領域において現れるのかを調査するものである。本稿では、まず、授業開始時と終了時にCan-do statementsを使用した先行研究について述べ、今回使用したCEFRの自己評価チェックリストについて説明する。次に、今回の調査と調査結果について報告し、最後に考察を述べる。

2. 先行研究

日本語教育におけるCan-do statementsに関する研究は、その開発に関するもの（森本他，2011；塩澤他，2010など）、学習者の特性による差を論じたもの（島田，2010）などもあるが、本節では試験との関連性を見た研究と授業開始時と終了時を比較した研究について述べる。

試験との関連性の研究については、玉岡他（2005）が自作のCan-do statements 20問と文字・語彙・文法22問、読解14問の試験をとの相関を見ている。結果は、相関係数はどれも有意ではなく、Can-do Scaleは日本語能力を測定する意味での妥当性が低いと論じている。島田他（2006）は、三枝（2004）の作成したCan-do statementsを使い、日本語能力試験の1級と2級、プレースメントテスト各々との相関を測定した。日本語能力試験の総点とCan-do statementsとの相関は1級で0.20、2級で0.31と低かったが、プレースメントテストとの相関は0.80と高かった。島田他は日本語能力試験との相関が低かった理由として、級別に測定したことによる分布の切断効果によるものとしている。Can-do statementsについては、その信頼性も高かったことから、日本語能力試験の結果を反映する尺度として有効であると論じている。これに追随するものとして、島田他（2007）は、三枝の開発したCan-do statementsを改訂したものを使い、プレースメントテストとの相関を調べ、0.73という高い相関関係を報告している。一方、今井（2010）は、島田他（2006）が扱ったデータは正規分布ではないため、相関が高くなりやすいことを指摘し、結果の解釈は慎重にすべきと警告している。今井は、島田他（2007）に修正を加えたCan-do statementsとJ-CATのセクションごとの相関を調べ、0.41から0.59とある程度の相関が見られたことを報告している。

授業開始時と終了時を比較した研究には、青木（2006）、原（2010）、鈴木他（2012）がある。青木（2006）は、三枝（2004）で作成されたCan-do statementsを用い、入学時とその約10ヶ月後の卒業時に76名の学生に調査を行い、その平均値を比較した。その結果、卒業時の方が、「読む」「書く」「話す」「聞く」の4技能全てで平均値が高くなっていた。中でも伸びの大きい項目は、コースの中で実際に行ったことを含む項目、インターアクションが必要ではないモノログや準備の可能な項目、入学時の平均値が高くない項目であったことを報告している。また、テストとの相関においては、入学時は0.64、卒業時は0.40であり、入学時の方が相関が高いことを報告している。さらに、得点上位のグループのほうがテストとの相関が高いという結果から、能力の高い学習者のほうが自己評価を正しく行えることも指摘している。

原（2010）は、学習者が言語活動を経験することにより、自分がどんな言語活動ができるかを正しく自己評価できるようになると考え、学習者17名に学期開始時と終了時の二回、授業で行う内容に合ったCan-do statementsに回答させた。使用したCan-do statementsはCEFRの能力記述文から選定したものである。調査では、全ての項目において、学期終了時のほうが一様に自己評価が上昇したという結果が得られた。原は、この結果を学習者が自分の能力が伸びたことを無

意識に感じ取った結果であるとしていると共に、学習者が言語活動を経験することの重要性を指摘している。

鈴木他（2012）は、JLC 日本語スタンダードの中から194項目のCan-doの調査を2010年10月（学期開始時）と2011年7月（学期終了時）に行った。2度の調査の調査対象者は異なっているが、二つの調査を比較すると、おおむねどのレベルでも学期終了時のほうが「できる」という回答が増えており、レベルが上がるにつれ、「できる」という回答が増えていたことを報告している。

青木（2006）及び鈴木他（2012）の研究から、授業開始時と終了時の比較では、授業開始時よりも約10ヶ月後の終了時ほうが評価が高くなることが明らかになっている。しかしながら、授業開始時と終了時の違いを比較する場合、日本の大学では一学期約15週間の授業が一般的であることを考えると、一学期での評価の違いをみることも必要ではないかと考えられる。また、授業で行った項目はその評価が学期開始時より終了時に高くなったことは、原（2010）で確認されたが、授業で行った項目のみではなく、より広い範囲の自己評価を行い、総合的な語学力の変化について知る重要性もあると思われる。

3. CEFRとスイス版自己評価チェックリスト

CEFR（ヨーロッパ共通言語参照枠）は、ヨーロッパの言語教育の向上のために一般的基盤を与えることを目的としたものである（吉島・大橋，2004）。CEFRでは、学習者のレベルをA1からC2の六段階に分け、A1、A2を「基礎段階の言語使用者」、B1、B2を「自立した言語使用者」、C1、C2を「熟達した言語使用者」としている。それぞれのレベルには、学習者の熟達度レベルを明示化した能力記述文が例示されており、これを参考にして各々の学習者がどのレベルであるかを把握できるようになっている。

ヨーロッパ言語ポートフォリオ（ELP）は、CEFRの目的を実行する教育的ツールである（ヨーロッパ日本語教師会・独立行政法人国際交流基金，2005）。スイス版ELPにはSchneider & North（2000）の自己評価チェックリストが掲載されているが、これは前述のA1からC2までの六つのレベルがあり、レベルごとに学習者ができることをチェックするシートである。チェックリストは、Listening, Reading, Spoken Interaction, Spoken Production, Strategies, Language Quality, Writing の7つの領域から成る。各々のレベルの各領域の項目数は表1の通りである。

表1 スイス版自己評価チェックリスト項目数

領域	A1	A2	B1	B2	C1	C2
Listening	4	6	6	6	6	1
Reading	8	8	8	8	8	6
Spoken Interaction	8	12	7	7	4	1
Spoken Production	2	6	6	6	4	2
Strategies	3	3	3	3	3	1
Language Quality	0	4	4	4	4	4
Writing	5	8	8	8	8	8
全体	30	47	42	42	37	23

4. 調査方法及び対象者

調査には前述の自己評価チェックリスト (Schneider & North, 2000) を使用した。このチェックリストは、学習者自身ができることをチェックし、チェックした項目が80%以上であれば、そのレベルに達しているとみなされるものである。今回の調査では、できるものにチェックするという方式ではなく、各項目について自分が日本語でどの程度できるかを5「常にできる」から1「全くできない」までの5段階で評価させた。

調査票は英語版、中国語版、韓国語版を用意した。英語版はSchneider & North (2000) を使用し、それを基に英語の堪能な中国人話者、韓国人話者により、中国語版、韓国語版の各々を作成した。作成された中国語版、韓国語版は、それぞれ別の中国語、韓国語話者により校正、確認が行われた。

調査は平成23年度後期と平成24年度前期に行い、各々の学期で、学期開始から約1ヶ月後とその約3か月後の学期終了前の2回、同じ調査票を配布し、調査対象者に回答してもらった。

調査対象者は岡山大学日本語コースを受講している留学生の内、学期開始後と学期終了前の両方の調査に参加した留学生86名である⁽²⁾。日本語コースは初級1、初級2、中級入門、中級1、中級2、上級1、上級2と7レベルに分かれているが、それを3グループに分け、初級1と初級2の受講生20名（初級レベル）にはCEFRチェックシートのA1とA2、中級入門と中級1の受講生37名（中級レベル）にはA2とB1、中級2以上の受講生29名（中上級レベル）にはB1とB2の調査票を与えた。各グループの受講生の出身国・地域は表2の通りである。中国語話者には中国語版、韓国語話者には韓国語版、その他の受講生には英語版が配布された。

表2 調査対象者の国・地域

初級レベル		中級レベル		中上級レベル	
国・地域	人数	国・地域	人数	国・地域	人数
中国	7	アメリカ合衆国	14	中国	13
アメリカ合衆国	3	フランス	7	韓国	7
ベトナム	3	中国	6	アメリカ合衆国	3
バングラデシュ	2	韓国	4	トルコ	2
インドネシア	1	ドイツ	2	フランス	2
オーストラリア	1	スペイン	1	台湾	2
韓国	1	ベトナム	1	計	29
タイ	1	ミャンマー	1		
ドイツ	1	ロシア	1		
計	20	計	37		

5. 結果

5.1 記述統計量

初級、中級、中上級それぞれのレベルの学期開始後と学期終了前の記述統計量を表3から表5に示す。表からは、全てにおいて学期終了前のほうが学期開始後より値が上がっているのがわかる。また、各項目について学期開始後と学期終了前の平均値の差を調べたところ、全ての項目で学期終了前のほうが学期開始後より値が高かった。

表3 初級レベルの記述統計量

領域	学期開始後		学期終了前	
	平均	標準偏差	平均	標準偏差
Listening	3.08	0.65	3.74	0.72
Reading	2.90	0.82	3.45	0.83
Spoken Interaction	3.20	0.68	3.67	0.60
Spoken Production	3.06	0.90	3.68	0.57
Strategies	3.29	0.88	3.69	0.86
Language Quality	3.04	0.80	3.44	0.67
Writing	2.74	0.81	3.35	0.68
全体	3.03	0.66	3.57	0.56

表4 中級レベルの記述統計量

領域	学期開始後		学期終了前	
	平均	標準偏差	平均	標準偏差
Listening	3.43	0.62	3.99	0.60
Reading	3.34	0.64	3.86	0.62
Spoken Interaction	3.67	0.74	4.18	0.67
Spoken Production	3.36	0.68	3.90	0.67
Strategies	3.46	0.77	4.11	0.77
Language Quality	3.39	0.71	3.84	0.75
Writing	3.21	0.59	3.74	0.66
全体	3.42	0.59	3.94	0.59

表5 中上級レベルの記述統計量

領域	学期開始後		学期終了前	
	平均	標準偏差	平均	標準偏差
Listening	3.50	0.62	4.00	0.55
Reading	3.53	0.57	3.93	0.53
Spoken Interaction	3.60	0.67	3.80	0.62
Spoken Production	3.32	0.65	3.70	0.76
Strategies	3.61	0.64	3.96	0.58
Language Quality	3.46	0.67	3.73	0.70
Writing	3.24	0.71	3.66	0.63
全体	3.45	0.59	3.82	0.55

5.2 信頼性と相関

初級, 中級, 中上級それぞれのレベルの学期開始後と学期終了前の回答についての信頼性 (α 係数) は表6の通りである。全てにおいて α 係数が非常に高く, この調査票の信頼性は高いことが明らかになった。

表6 レベルごとの信頼性

レベル	学期開始後	学期終了前
初級	0.98	0.98
中級	0.99	0.99
中上級	0.99	0.99

各レベルにおいて, それぞれの領域が互いに関連しているかどうか, 相関係数を求めた結果は以下の通りである (表7-12)。全体と各領域との相関は非常に高い。また, 領域間の相関も高いものが多いが, 初級レベル学期終了前のReadingに関しては, Spoken InteractionやStrategiesとの相関が低いという結果となった。

表7 初級レベルの領域ごとの相関 (学期開始後)

	Listening	Reading	Spoken Interaction	Spoken Production	Strategies	Language Quality	Writing	全体
Listening	---	0.634	0.835	0.656	0.704	0.626	0.586	0.844
Reading		---	0.644	0.366	0.569	0.453	0.677	0.794
Spoken Interaction			---	0.833	0.884	0.625	0.759	0.949
Spoken Production				---	0.794	0.610	0.628	0.796
Strategies					---	0.547	0.679	0.867
Language Quality						---	0.826	0.744
Writing							---	0.874
全体								---

表8 初級レベルの領域ごとの相関 (学期終了前)

	Listening	Reading	Spoken Interaction	Spoken Production	Strategies	Language Quality	Writing	全体
Listening	---	0.656	0.690	0.738	0.548	0.572	0.718	0.895
Reading		---	0.192	0.389	0.239	0.407	0.767	0.729
Spoken Interaction			---	0.723	0.883	0.634	0.482	0.780
Spoken Production				---	0.534	0.684	0.656	0.797
Strategies					---	0.552	0.522	0.735
Language Quality						---	0.790	0.766
Writing							---	0.884
全体								---

表9 中級レベルの領域ごとの相関（学期開始後）

	Listening	Reading	Spoken Interaction	Spoken Production	Strategies	Language Quality	Writing	全体
Listening	---	0.744	0.762	0.744	0.616	0.692	0.745	0.870
Reading		---	0.613	0.703	0.476	0.563	0.791	0.819
Spoken Interaction			---	0.887	0.834	0.776	0.701	0.917
Spoken Production				---	0.797	0.826	0.783	0.936
Strategies					---	0.841	0.668	0.827
Language Quality						---	0.736	0.859
Writing							---	0.888
全体								---

表10 中級レベルの領域ごとの相関（学期終了前）

	Listening	Reading	Spoken Interaction	Spoken Production	Strategies	Language Quality	Writing	全体
Listening	---	0.830	0.744	0.726	0.740	0.760	0.784	0.887
Reading		---	0.681	0.643	0.691	0.672	0.790	0.856
Spoken Interaction			---	0.865	0.870	0.842	0.761	0.923
Spoken Production				---	0.889	0.814	0.744	0.897
Strategies					---	0.867	0.781	0.915
Language Quality						---	0.753	0.894
Writing							---	0.903
全体								---

表11 中上級レベルの領域ごとの相関（学期開始後）

	Listening	Reading	Spoken Interaction	Spoken Production	Strategies	Language Quality	Writing	全体
Listening	---	0.805	0.848	0.816	0.647	0.767	0.856	0.922
Reading		---	0.700	0.672	0.660	0.687	0.660	0.825
Spoken Interaction			---	0.905	0.784	0.855	0.894	0.951
Spoken Production				---	0.784	0.846	0.907	0.940
Strategies					---	0.737	0.753	0.824
Language Quality						---	0.807	0.892
Writing							---	0.942
全体								---

表12 中上級レベルの領域ごとの相関（学期終了前）

	Listening	Reading	Spoken Interaction	Spoken Production	Strategies	Language Quality	Writing	全体
Listening	---	0.844	0.803	0.753	0.718	0.622	0.666	0.872
Reading		---	0.700	0.720	0.618	0.538	0.705	0.843
Spoken Interaction			---	0.837	0.882	0.756	0.791	0.926
Spoken Production				---	0.732	0.797	0.816	0.923
Strategies					---	0.815	0.753	0.863
Language Quality						---	0.811	0.846
Writing							---	0.907
全体								---

5.3 学期開始後と終了前の変化

初級、中級、中上級それぞれのレベルにおいて、学期開始後と終了前の変化があったかを対応のある t 検定を使い、分析した。その結果、初級レベルで $t(19) = 4.50$, $p = 0.0002$, $r = 0.72$, 中級レベルで $t(36) = 6.99$, $p < 0.0001$, $r = 0.76$, 中上級レベルで $t(28) = 4.32$, $p = 0.0002$, $r = 0.63$ となり、全てのレベルにおいて有意差が認められた。 r は効果量であり、.30 で中程度の効果量、.50 で大きな効果量があると解釈されるが（水本・竹内, 2008）、全てのレベルで大きな効果量が認められ、学期開始後と終了前には学習者の自己評価に変化が生じたことが明らかとなった。

それぞれのレベルにおいて、領域別に変化があったかどうかを対応のある t 検定を使って分析したものが表13から表15である。初級、中級レベルにおいては、全ての領域で有意差が現れ、効果量も大きいという結果となった。中上級レベルにおいては、全ての項目で5%水準で有意差が現れたが、Spoken InteractionとLanguage Qualityで効果量が大きいと判断される目安の.50を下回った。

表13 初級レベルの学期開始後と終了前の平均値の差の検定

領域	t	df	p	r
Listening	4.88	19	0.0001	0.75
Reading	3.15	19	0.0053	0.59
Spoken Interaction	4.53	19	0.0002	0.72
Spoken Production	3.49	19	0.0025	0.63
Strategies	2.59	19	0.0180	0.51
Language Quality	2.52	19	0.0207	0.50
Writing	3.54	19	0.0022	0.63

表14 中級レベルの学期開始後と終了前の平均値の差の検定

領域	<i>t</i>	<i>df</i>	<i>p</i>	<i>r</i>
Listening	6.73	36	0.0000	0.75
Reading	5.66	36	0.0000	0.69
Spoken Interaction	5.24	36	0.0000	0.66
Spoken Production	4.97	36	0.0000	0.64
Strategies	6.41	36	0.0000	0.73
Language Quality	5.52	36	0.0000	0.68
Writing	5.83	36	0.0000	0.70

表15 中上級レベルの学期開始後と終了前の平均値の差の検定

領域	<i>t</i>	<i>df</i>	<i>p</i>	<i>r</i>
Listening	5.30	28	0.0000	0.71
Reading	4.51	28	0.0001	0.65
Spoken Interaction	2.14	28	0.0414	0.37
Spoken Production	3.35	28	0.0024	0.53
Strategies	3.02	28	0.0054	0.50
Language Quality	2.22	28	0.0347	0.39
Writing	3.47	28	0.0017	0.55

効果量が大きいと判断されなかった中上級レベルのSpoken InteractionとLanguage Qualityについて、中級2と上級レベルの学生を分けて平均値の差を調べたのが表16である。Spoken Interactionに関しては、上級レベルの学生が学期開始後のほうが値が高くなっている。また、Language Qualityに関しても上級レベルの学生の値の変化は小さい。中級2に比べて上級の学生の平均値の差が小さいことが結果に影響を与えたのではないかと推測される。

表16 中級2と上級レベルの2領域における平均値

レベル	Spoken Interaction		Language Quality	
	学期開始後	学期終了前	学期開始後	学期終了前
中級2	3.46	3.80	3.33	3.68
上級	3.82	3.80	3.67	3.81

6. 考察

本稿ではCEFRのスイス版自己評価チェックリストを使い、日本語コース受講者を対象に、学期開始後から終了前の約3か月の期間で、自己の日本語力の評価に変化がみられるかを調査した。調査の結果、初級、中級、中上級の全てのレベルにおいて、学生の自己評価の平均値が学期開始後より終了前の方が有意に高いという結果が現れた。また、各レベルの全ての領域において終了前の方が開始時より自己評価の平均値が有意に高く、全ての項目で終了前のほうが開始時より平均値が高いという結果となった。

これは学期開始後から終了までの期間で、学習者が自己の日本語力について、日本語でできることが増えたと感じ、自己の日本語力をより肯定的に評価するようになったということである。3ヶ月という短い期間でも、日本で生活し、日本語の授業を履修することにより、日本語力の伸

びが感じられるようになったということであろう。教師は、授業中の学生の観察、課題や試験により学生の日本語力を測るが、一学期15週間という期間では、その伸びがはっきりと見えない場合もある。特に中級以上のクラスになると、すでにある程度の日本語力があるため、初級に比べて伸びが見えにくいと感じる教師も多い。今回の結果は、外からは明白には観察できなくても、学習者自身は伸びを感じていることを意味する。学習者のこのような実態を教師が知っておくことは、授業を行う上で有益であろうと考える。また、学習者の自己評価値の上昇は、学習者が現在のクラスから次学期一つ上のレベルのクラスに上がることに對する肯定的な判断材料ともなる。今回の結果は、岡山大学日本語コースにおいて、授業内での評価のみではなく、学習者からの評価においても学習者が次のレベルに行く準備ができていることを支持するものである。これは、現在の岡山大学日本語コースの授業がそれぞれのレベルで適切に行われ、全体のコースが問題なく運営されていることを示す一つのデータとなりえるであろう。

また、初級、中級、中上級の各々のレベルで、全ての項目の自己評価の平均が終了前のほうが開始時より高く、全ての領域で自己評価の変化に有意差が現れたという結果も興味深い。原(2010)は、授業で行った項目は、学期終了前の自己評価が高くなると論じたが、今回使用したチェックリストの項目数は77項目（初級レベル）あるいは84項目（中級、中上級レベル）あり、授業でこれを全て行うのは不可能である。学習者が授業以外の生活の場で、これらの項目を実際に行ったかどうかは確認できないが、授業で全てを行わなくても、学習者ができることは増えていくということと言えるであろう。ただし、これが第二言語学習環境に特有のものなのか、外国語学習環境においても言えるのかは今後の研究の必要があると思われる。

ただし、領域別の差については、今回、中上級レベルのSpoken InteractionとLanguage Qualityで、有意差は5%水準で有意ではあったが、効果量は大きくはないと判断された。この一つの理由としては、中級2に比べ、上級クラスの学生の自己評価の差が学期開始後と終了前で小さかったことが挙げられる。上級2（23年度までは「上級」という名称）は、週4コマ開講されているが、選択制で週1コマでも受講可能である。今回の調査対象者を調べたところ、平均受講コマ数は週1.8コマであった。また、上級レベルになると、話すスキルよりも読む、書くに重点が置かれることも多い。これらの要因が学生の日本語力の伸びの実感に影響したのではないと思われるが、これについては今後調べて行く必要のある課題である。

7. 今後の課題とまとめ

本稿ではCEFRのスイス版自己評価チェックリストを使い、岡山大学の日本語コース受講者を対象に、学期開始後と終了前の自己の日本語力の変化について調査した。

今後の課題として、まず各項目の精査を挙げたい。調査に使用されたCEFRのスイス版自己評価チェックリストは元々ヨーロッパ言語を対象として作られたもので、日本語には当てはまらないと思われる項目もある。今後は各項目を分析し、項目の適切性について検討する必要がある。第二に、受講者のデータを増やし、レベル別ではなく、クラス別の分析を行いたい。今回の調査は86名の参加者を3グループに分けて行ったが、今後はデータ数を増やし、「初級1」と「初級2」などクラスレベルでの違いを見ていきたい。特に今回、授業前後の自己評価の差が小さいと指摘した上級クラスについては、データを増やして見て行く必要があると考える。最後に、今回はB2レベルまでを調査項目としたが、上級の学生にとっては易しい項目も多いため、Cレベルの項目を入れて調査を行いたいと考えている。

Can-do statementsは、先行研究においては、テスト得点との相関を分析し、言語能力を測る

手段としての妥当性に疑問を呈す研究もあるが、試験は学習者の持つ言語能力全てを測れるものではない。試験によって測れない言語能力が自己評価によって見えてくる可能性もある。今回の調査で、学習者は確実に自己の日本語力の伸びをとらえていることが明らかになったことは、Can-do statementsが学習者自身による自己の言語能力の意識化の指標として、今後大いに活用できる可能性を示したのではないかと考える。

注

- (1) 平成23年度までは上級は一レベルしかなく、全6レベルのコースであった。また、上級クラスは週1コマでも履修可能であった。
- (2) 一つのレベル（A1 など）内の一つの領域（Listening など）の項目の中で複数の未回答の項目があった場合、その回答者はデータから削除した。

引用文献

- 青木惣一（2006）「学習開始時と学習終了時の Can-do statements 比較－プロフィシエンシー・テスト、オーラル・インタビューを外部基準として－」『アメリカ・カナダ大学連合日本研究センター紀要』29, pp. 18－44.
- 原やす江（2010）「能力記述文（CDS）を利用した日本語授業のシラバス・デザインの試み－ヨーロッパ共通参照枠（CEFR）を外部指標とする日本語能力の可視化に向けて－」『城西国際大学紀要』18（2）, pp. 53－81.
- 細川英雄・西山教行（2010）『複言語・複文化主義とは何か－ヨーロッパの理念・状況から日本における受容・文脈化へ』くろしお出版
- 国際交流基金（2010）『JF 日本語教育スタンダード』国際交流基金
- 水本篤・竹内理（2008）「研究論文における効果量の報告のために」『英語教育研究』31, pp. 57－66.
- 森本由佳子・塩澤真季・小松知子・石司えり・島田徳子（2011）「コミュニケーション言語活動の熟達度を表す JF Can-do の作成と評価」『国際交流基金日本語教育紀要』7号, pp. 25－42.
- 三枝令子（2004）『日本語 Can-do-statements 尺度の開発 平成13年度～平成15年度科学研究費補助金研究成果報告書』
- Schneider, G. & North, B. (2000) *Fremdsprachen können – was heisst das?* Zürich: Rüeegger.
- 島田めぐみ・三枝令子・野口裕之（2006）「日本語 Can-do statements を利用した言語行動記述の試み－日本語能力試験受験者を対象として－」『世界の日本語教育』16, pp. 75－88.
- 島田めぐみ・谷部弘子・斎藤純男（2007）「日本語科目における言語行動目標の設定－Can-do-statements を利用して－」『東京学芸大学紀要 総合教育科学系』58, pp. 495－505.
- 島田めぐみ（2010）「自己評価 Can-do statements の母語別分析」『東アジア日本語教育・日本文化研究』13, pp. 79－93.
- 塩澤真季・石司えり・島田徳子（2010）「言語能力の熟達度を表す Can-do 記述の分析－JF Can-do 作成のためのガイドライン策定に向けて－」『国際交流基金日本語教育紀要』6号, pp. 23－39.
- 鈴木美加・藤森弘子・藤村知子・鈴木智美・中村彰・坂本恵・花蘭悟・伊集院郁子（2012）「日本語学習における目標記述をめぐって－全学日本語プログラムの Can-do リスト作成に向けて－」『東京外国語大学留学生日本語教育センター論集』38, pp.155－166.

- 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号, pp.65-78
- 東京外国語大学留学生日本語教育センター（2011）『JLC日本語スタンダード2011改訂版』東京外国語大学大学留学生日本語教育センター
- ヨーロッパ日本語教師会・独立行政法人国際交流基金（2005）『ヨーロッパにおける日本語教育事情とCommon European Framework of Reference for Languages』独立行政法人国際交流基金
- 吉島茂・大橋理枝他（訳・編）（2004）『外国語教育II 外国語の学習, 教授, 評価のためのヨーロッパ共通参照枠』朝日出版社