

韓国人日本語学習者による自己評価と韓国人日本語教師による難易度評価 の調査結果比較：CEFR 読解Can-Do Statementsを対象に

谷 誠司, 宮崎 佳典

Assessing CEFR Reading Can-Do Statements Difficulty for Korean
Learners of Japanese : Differences in Korean Learners and Teachers
Perspectives

Seiji TANI, Yoshinori MIYAZAKI

2018 年 11 月 9 日受理

抄 録

Common European Framework of Reference for Languages (以下 CEFR) の能力記述文 (Can-Do Statements, 以下 CDS) を基にして開発された DIALANG Self-Assessment Statements (以下 CEFR-DIALANG CDS) にある読解CDSを使用して、韓国人日本語学習者の自己評価データ (361名) と韓国人日本語教師の難易度評価データ (30名) を主にラッシュモデルで比較分析した。その結果、次の4点が明らかになった。1) 学習者評価と教師評価において全体の傾向としてCEFRの順序性が維持されていたが、どちらの評価でもB2とC1のCDSの間に差が見られなかった。2) A1とA2のCDSは教師が学習者より易しいと評価するが、B2のCDSは教師が学習者よりも難しいと評価した。3) 文章の特徴に関して具体的な記述がされていないCDSや語彙力について記述しているCDSは、学習者評価と教師評価においてモデルから外れたCDSと判断された。4) 学習者評価と教師評価の間に差があるCDSがあることが確認された。ただし、学習者自己評価と教師難易度評価は評価をする際に念頭に置く対象が同一ではないので、結果を一般化するには更なる調査が必要である。

キーワード : CEFR・DIALANG Self-Assessment Statements・読解Can-Do Statements・学習者評価・教師評価

1. はじめに

CEFRは2001年に欧州評議会 (Council of Europe) が発表して以来、世界の外国語教育に影響を与えている。特に技能別に6つのレベルで具体的にどのようなことができるかを記述したCDSを利用し、コース・教材・テストが様々な言語で開発されている。

日本語教育においてもCEFRの影響を受けた「JF日本語教育スタンダード」が発表され、「JF日本語教育スタンダード」準拠の教材（「まるごと 日本のことばと文化」）も発行されている。またCEFR 準拠の日本語テスト（OJAE¹やフランスのSELF²等）も開発されている。

CEFRなどの能力尺度を利用して教材やテストを開発する際は、CDSの記述内容に沿って例文を選択したりタスクを作成したりすることになり、それだけCDSの役割は大きい。同一CDSに対して学習者評価と教師評価を比較することは、CDSの改善につながることも期待できる。

そこで本稿ではCEFR-DIALANG読解CDS³を使用した韓国人日本語学習者による学習者評価データと韓国人日本語教師による教師評価データを比較分析し、その結果を報告する。

2. 先行研究

CEFRに関連するCDSの尺度化とレベル分割点の決定の過程に関する先行研究をいくつか紹介する。

CEFRのCDSは、まず既存の言語能力尺度からCDSを収集し、外国語教員グループによるCDSの質的な調査を経て、教師によるCDSの量的調査（CDSを5段階で難易度評価をするなど）からCEFRのCDSの尺度化とレベル分割点を決定している。（Council of Europe、2001；North、2005；野口・大隅、2014）

CEFR-DIALANG CDSはまずCEFRにあるすべてのCDSを読解、聴解、ライティング別に集め、最も簡潔かつ具体的なCDSを選択した。選択したCDSをI can（「○○ができる」）の形式に直し、またより簡潔な表現に修正した。修正されたCDSはすべて言語教育や言語テストの専門家から確認を受け、ヨーロッパ内の様々な言語に翻

¹ CEFR 準拠 日本語 口頭 産出 能力 評価 法 （Oral Japanese Assessment Europe）
<https://ejale.wordpress.com/ojae/ojae-intro/>

² 「SELF (Système d'évaluation en langues à visée formative) は、フランス国立研究機構 (ANR) の公募プロジェクト「革新的教育を目指した先駆的研究」に採択されたInnovalangues プロジェクトの一環として開発中のCEFR準拠の診断つきオンライン評価システムである。」（東・代田・永田、2017）

³ DIALANGとは、欧州委員会 (European Commission) の資金援助を受け、欧州の14言語の言語運用力をCEFRのレベルに基づいて診断する無料オンライン・システムである。（<https://dialangweb.lancaster.ac.uk/>）CEFR-DIALANG CDSは受験者のレベルを事前に把握するため自己の言語能力を自己評価するために開発された。（Alderson,2005）CEFRのCDSとCEFR-DIALANG CDSの違いは、前者は概括的な記述であるのに対し、後者は個別・具体的な記述になっている点である。

例：CEFR「聞く」B2レベル：「長い話や講義を理解することができ、適度になじみのある話題であれば、複雑な議論でもついていくことができる。ほとんどのTVニュースや時事問題を扱った番組を理解できる。標準語の映画なら大半理解できる」

CEFR-DIALANG「聞く」B2レベル：「話題に比較的なじみがあり、話者が話の筋を明確に示してくれば、長い話や複雑な議論の流れを理解することができる」「標準語のテレビニュース、ドキュメンタリーのような時事問題を扱った番組、生放送のインタビュー、トークショー、演劇、映画のほとんどを理解することができる」

このようにCEFR-DIALANG CDSはCEFRでは1つのレベルに含まれる項目を切り離し、それぞれ独立させている。（根岸、2006）また、CEFR-DIALANG CDSには数は少ないが、CEFRのCDSにはない独自のCDSもある。（Alderson、2005）本研究では中島・永田(2006)や根岸(2006)に従い、CEFR-DIALANG CDSをCEFRのCDSと考え、研究を進めた。

訳され、それぞれの言語学習者による学習者評価調査が実施され、CDSの尺度化とレベル分割点を決定している。またCEFRのCDSとの比較もしており、CEFRの設定したレベルとは異なっているCEFR-DIALANG CDSもあった⁴。(Council of Europe, 2001 ; Alderson, 2005)

CEFR-J⁵はCEFRに準拠しながら日本の英語教育での利用を目的に構築された英語能力尺度である。まず教育経験や直感からCDSの試作品(α版)を作り、α版を専門家による質的分析や中高教員によるCDSの並べ替え調査を実施、中高大の英語学習者への学習者評価を経て、CDSのレベルの変更や表現の改訂を行って、CDSの尺度化とレベル分割点を決定している。(投野、2013)

日本語教育の分野ではCEFRのCDSを使った研究は、大隈・野口・熊谷・石毛・長沼・和田・伊東(2006)や坂野(2015)、宮崎・谷(2015)、谷・宮崎・高田(2017)などがあるが、どれも学習者評価による研究であり、学習者評価と教師評価を比較した研究は著者らの調べる限り多くは行われていないようである。

3. 調査方法

3. 1. 協力者

1) 韓国人日本語学習者

韓国の3つ大学で日本語を学習している韓国人大学生および大学院生361名である。協力者の特徴は表1のようにまとめられる。

表1：調査協力者(韓国人日本語学習者)の特徴⁶

大学	ソウル：K大学(110名)、D大学(46名)、プサン：S大学(205名)
年齢	19歳～28歳(一部、大学院生が30代)
学年	1年生：96名(26.6%)、2年生：81名(22.4%)、3年生：81名(22.4%)、4年生：88名(24.4%)、大学院生：12名(3.3%)
専攻	日本語関連専攻・複数専攻者：138名(38.2%)、他専攻者：215名(59.6%)
日本語学習期間	6か月未満：75名(20.8%)、1年～2年：93名(25.8%)、2年～3年：44名(12.2%)、3年～4年：55名(15.2%)、4年以上：85(23.5%)
新JLPT合格者	N1合格：77名(21.3%)、N2合格：26名(7.2%)、N3合格：11名(3.0%)、N4合格：9名(2.5%)、N5合格：1名(0.3%)

⁴ Alderson(2005)はCEFR-DIALANG 読解CDSの中でCEFRのレベル設定と異なっていたCDSが4つあると報告しているが、Council of Europe(2001)で公開されているCEFR-DIALANG 読解CDSでは以下の2つのCDSしか該当せず、残り2つのCDSは含まれていない。CEFRのレベルと異なっているCEFR-DIALANG 読解CDSは以下の通り。CDS23「自分の専門分野に関連する通信文(手紙・メールなど)を読んで、楽に必要な意味が把握できる。(CEFRではB2、DIALANGではB1) CDS29「ときどき辞書を使用すれば、どんな文書でも理解することができる。」(CEFRではC1、DIALANGではB2)

⁵ <http://www.cefr-j.org/>

⁶ 無回答者がいる質問項目については、各割合を合計しても100%にならない。

2) 韓国人日本語教師

韓国で日本語を教えている30名の韓国人日本語教師である。協力者の特徴は表2のようにまとめられる。30名中、22名が大学に勤務されており、勤務先の大学は学習者評価を実施した大学ではあるが、高校や中学等に勤務されている先生もいる。

表2：調査協力者（韓国人日本語教師）の特徴

日本語教師歴	5年未満：7名(23.3%)、5年～10年未満：6名(20.0%)、10年～15年未満：7名(23.3%)、15年以上：8名(26.7%)、不明：2名(6.7%)
勤務校種	大学：22名(73.4%)、高校：4名(13.3%)、中学：1名(3.3%)、不明：3名(10.0%)
学位	博士：22名(73.3%)、修士：6名(20.0%)、不明：2名(6.7%)
専門	日本語教育：7名(23.2%)、日本文学：6名(20.0%)、通訳・翻訳：5名(16.7%)、日本語学：5名(16.7%)、社会言語学：2名(6.7%)、その他：3名(10.0%)、不明：2名(6.7%)

3. 2. 調査時期と手順

1) 韓国人日本語学習者

調査は3回に分けて行った。(①S大学：2011年11月～12月の間に2回、②K大学：2012年5月に1回、③K大学とD大学：2012年12月にそれぞれ1回) 授業時間内に簡単な説明(調査の目的、調査依頼者、調査方法、回答方法、個人情報への扱いなど)とともに調査票を配布し授業時間外に回答してもらい、その後回収した。

2) 韓国人日本語教師

2013年6月～8月の期間、調査票のファイルを協力者にメールで送り、回答した調査票ファイルを送り返してもらった形で実施した。

3. 3. 調査票

1) 韓国人日本語学習者

CEFR-DIALANG読解CDS(31項目)を使用した。原文は英語で書かれているので、根岸(2006)の参考資料を参考に日本語訳をし、さらに日本の大学院(社会学系の修士課程)を修了した韓国人に韓国語に翻訳してもらったものを使った。各項目への回答は、根岸(2006)の指摘や大隈他(2006)の例に従い、まずその項目内容の経験有無をチェックし、経験があれば「1.全然できなかった」から「5.問題なくできた」、経験がなければ「1.全然できないと思う」から「5.問題なくできると思う」の5件法で該当するところにチェックしてもらった。

2) 韓国語教師

韓国語学習者と同様、CEFR-DIALANG読解CDS (31項目) を使用し、各CDSの難しさを6件法 (1.初級 (下)、2.初級 (上)、3.中級 (下)、4.中級 (上)、5.上級 (下)、6.上級 (上)) で判断してもらった。判断する際には今教えている学習者を念頭に置いて判断をするようにとは指示をしていない。

4. 結果と考察

4. 1. 素点による分析

CEFR-DIALANG読解CDS (31項目) における学習者評価と教師評価を素点から分析した結果が表3である。なお、学習者評価と比較するために教師評価は点数を逆転させた。(「1.初級 (下)」→6、「6.上級 (上)」→1) 人数が項目ごとに異なっているのは、学習者評価の場合は「経験あり」にマークした者のみを分析対象としたためであり、教師評価の場合は無回答のケースがあったためである。どちらの評価においても、項目番号が進む(CEFRのレベルが上がる)につれ、平均点が下がっている。またCDS 1、2、3、5、7、13、14は学習者評価と教師評価の平均値の差が1.0以上で大きい。学習者評価は標準偏差が全体として1.00前後であった。一方、教師評価は6件法であることもあり、標準偏差では1.50以上のCDSが5つあり、特にB2レベルのCDSの評価にばらつきが見られた。信頼性係数の推定値 (α 係数) は学習者評価が.97であり、教師評価も.97であった。

表3：素点による分析結果

CDS No.(CEFR レベル)	記述内容 (要点のみ抜粋)	学習者評価			教師評価		
		平均点	標準偏差	人数	平均点	標準偏差	人数
1 (A 1)	簡単な情報が含まれたテキストや簡潔な描写のテキスト	3.91	0.98	334	4.93	0.68	30
2 (A 1)	身近な名前、単語、基本的な表現からできている非常に短い簡単なテキスト、テキストのある部分を読み返す	4.11	1.00	345	5.27	0.81	30
3 (A 1)	短く簡単に書かれた指示 (特に絵を含む) に従う	3.96	1.04	335	5.60	0.76	30
4 (A 1)	もっとも一般的で日常的な状況でよく出くわす、簡単な掲示、なじみのある名前、単語、または非常に簡単な句を認識する	4.05	0.97	335	4.87	0.96	30
5 (A 1)	葉書、短く簡単なメッセージ	3.75	1.10	276	4.97	0.84	30
6 (A 2)	もっとも頻度の高い単語を含んだりする短くて簡単なテキスト	3.81	1.06	311	4.57	1.17	30
7 (A 2)	日常的な言葉で書かれた短くて簡単なテキスト	3.88	1.11	337	4.93	0.81	30

8(A 2)	自分の仕事に関連した短くて簡単なテキスト	3.72	1.09	236	4.50	1.06	30
9(A 2)	簡単な言語資料(例: 広告、パンフレット、メニュー、時刻表など)、特定の情報を見つける	3.57	1.08	267	4.47	1.09	30
10(A 2)	簡潔に書かれたテキスト(例: 手紙、パンフレット、新聞の短い事件記事)、特定の情報を見つける	3.50	1.09	243	4.20	1.17	30
11(A 2)	短く、簡単な個人的な手紙	3.76	1.17	269	4.71	0.99	28
12(A 2)	身近な話題について日常の定型の手紙やファックス	3.67	1.08	234	4.07	1.31	30
13(A 2)	日常生活で出くわす機器の簡単な使用方法(例: 公衆電話)	3.59	1.08	236	4.83	0.93	30
14(A 2)	通り、レストラン、駅のような公共の場所や職場にある標識や掲示	3.64	1.09	272	5.20	0.91	30
15(B 1)	自分の専門分野や関心のある話題に関して簡潔に書かれたテキスト	3.67	1.09	233	4.27	1.21	30
16(B 1)	手紙、パンフレット、短い公的な文書といった、日常的な文章、一般的な情報を見つける	3.59	1.09	227	4.40	1.08	30
17(B 1)	長いテキストや複数の短いテキストをざっと目を通す、課題を遂行するために必要な情報を探す	3.49	1.05	239	3.23	1.52	30
18(B 1)	なじみのある話題に関する簡単な新聞記事、重要な点を認識する	3.53	1.07	218	4.10	1.04	30
19(B 1)	はっきりと主張が書かれたテキスト、主要な結論を把握する	3.62	1.12	229	3.90	1.25	30
20(B 1)	文章の議論の大まかな流れは把握できる、詳細に認識できるわけではない	3.12	0.91	256	3.57	1.26	30
21(B 1)	個人の手紙、出来事・感情・希望の表現を理解する。文通できる	3.69	1.08	219	3.77	1.17	30
22(B 1)	機器に関する簡単な使用説明	3.5	1.13	223	4.24	1.04	29
23(B 2)	自分の専門分野に関連する通信文(手紙・メールなど)	3.46	1.11	202	3.64	1.39	28
24(B 2)	辞書が使えれば、自分の専門以外の専門的な記事	3.48	1.11	218	2.87	1.65	30
25(B 2)	読む目的やテキストの種類に応じて読む速度や読み方を変える、様々な種類のテキスト、かなり楽に読む	3.22	1.08	179	2.70	1.64	30
26(B 2)	広汎な語彙力、頻度の低い語彙や句にはいくらかてこずる	3.06	1.10	197	3.03	1.14	30
27(B 2)	広範囲にわたる専門的な話題についてのニュース、記事、レポートの内容と関連性をすばやく確認	2.96	1.10	153	2.60	1.74	30

28(B 2)	筆者が特別な立場や視点から取り上げた、現代の問題に関する記事やレポート	3.19	1.12	176	2.93	1.56	28
29(C 1)	ときどき辞書を使用すれば、どんな文書でも	3.24	1.15	238	3.57	1.45	30
30(C 1)	難しい箇所を読み返す、自分の専門分野以外、新しい機械や手順についての長い複雑な説明、細かいところまで理解	2.91	1.06	156	2.73	1.48	30
31(C 2)	実質的にあらゆる形式の書きことば	2.61	1.11	150	1.77	1.38	30

* 標準偏差が1.5以上のものは太字にした。

4. 2. ラッシュモデルによる分析

ラッシュモデルの分析ソフト (WINSTEPS ver3.73) で分析した。本研究では質問紙を使用しているので、Bond and Fox (2007) に従い、ミスフィット (不適合: misfit) の基準を0.6~1.4の範囲にした。

学習者評価データを分析した結果、アンダーフィット (underfit)⁷ に該当した回答者を除き、168名のデータを使って、再分析した。Suzuki (2015)⁸ に従い、外れ値の影響を受けるが、回答者よりも項目により反応するアウトフィット平方平均値 (outfit mean square) を見ると、5つのCDS (20・26・29・30・31) が上限の1.4を超え、アンダーフィットと判断された (表4)。

CDS20 (B 1) は「文章における議論の大まかな流れを認識することができるが、必ずしも詳細に認識できるわけではない。」とあり、文章に関する具体的な特徴が記述されていない。また、CDS29 (C 1) 「ときどき辞書を使用すれば、どんな文書でも理解することができる。」やCDS31 (C 2) 「抽象的であったり、構造的に複雑であったり、高度に口語的であるような、文学的な文章や非文学的な文章を含む、実質的にあらゆる形式の書きことばを理解し、解釈することができる。」では、「どんな文書」や「あらゆる形式の書きことば」とあり、これらも具体的な文書・文章をイメージしにくいことがアンダーフィットにつながったと考えられる。CDS26 (B 2) は「広汎な語彙力を持っているが、頻度の低い語彙や句にはいくらかてこずるかもしれない。」

⁷ ラッシュモデルにおけるミスフィットには2種類あり、本研究の場合、0.6未満の場合をオーバーフィット (overfit)、1.4より高い場合をアンダーフィット (underfit) になる。オーバーフィット (overfit) とはモデルに適合しすぎているケースを指し、あまり問題とはされないが、アンダーフィット (underfit) とはモデルに適合していないケースを指す。(小泉・飯田、2010) そのためアンダーフィット (underfit) は「有害な存在 (あってはならない)」であり「削除を検討する対象になる場合がある」(静、2005) と言われている。本研究ではオーバーフィットの回答者はそのまま残し、アンダーフィットの回答者データを除いて再分析をした。

⁸ ラッシュモデルではデータがどの程度モデルにフィット (fit: 適合) しているかをインフィット値 (infit) とアウトフィット値 (outfit) という2つの指標で提示する。また、インフィット値は回答者により反応し、アウトフィット値は項目により反応すると言われている。Suzuki (2015) では日本に主に留学中の韓国人日本語学習者を対象に読解CDSを使った調査を行い、項目数の少ないCDS調査であるため回答者よりも項目 (CDS) により反応するアウトフィット値を使用している。

とあり、読解ではなく、語彙力に関する記述がされていることがアンダーフィットになった原因と考えられる。CDS30 (C1) は「もし難しい箇所を読み返すことができれば、それが自分の専門分野に関連していなくても、新しい機械や手順についての長い複雑な説明を細かいところまで理解できる。」とあり、どのような文章であるかについての記述はされているものの、専門外の文章である点や新しい機械や手順についての長い複雑な説明といった記述が具体的にイメージしにくいことがアンダーフィットの原因になったと推測される。

教師評価の回答者数は30名である。ラッシュモデルの分析に必要とされる最小サンプルサイズは100以上とされるので (大友、1996)、今回の分析結果はあくまでも参考値である。教師評価データを分析した結果、アンダーフィットに該当した回答者が4名いた。個別の回答データを確認し、アウトフィット平方平均値が最も高く、実際の回答データもほぼすべて4で回答をしていた1名とすべてを6と回答していた回答者1名の合わせて2名を除き、28名のデータを使って、再分析した。アウトフィット平方平均値を見ると、3つのCDS (20・26・27) がアンダーフィットであった。

CDS20と26は学習者評価と教師評価のどちらにもアンダーフィットになっている。学習者評価と同様、CDS20は文章に関する具体的な特徴が記述されていないこと、CDS26は読解ではなく、語彙力に関する記述がされていることが原因と考えられる。CDS27 (B2) 「さらに詳細に読む必要があるかどうかを決定するために、広範囲にわたる専門的な話題についてのニュース、記事、レポートの内容と関連性をすばやく確認することができる。」とあり、文書のタイプ (ニュース、記事、レポート) についての記述はされているが、「広範囲にわたる専門的な話題」は具体性に欠けるためアンダーフィットになったと考えられる。

表4：ラッシュモデルによる分析結果

CDS No. (CEFRレベル)	学習者評価			教師評価			学習者評価の推定値と教師評価の推定値の差
	難易度推定値	標準誤差	アウトフィット平方平均	難易度推定値	標準誤差	アウトフィット平方平均	
1 (A1)	-1.47	0.13	0.94	-2.45	0.43	0.59	0.98
2 (A1)	-2.34	0.13	0.87	-2.24	0.48	1.11	0.10
3 (A1)	-1.93	0.13	0.91	-2.37	0.67	1.01	0.44
4 (A1)	-2.08	0.13	0.81	-1.54	0.35	1.16	0.54
5 (A1)	-0.71	0.13	0.73	-1.88	0.40	0.53	1.17
6 (A2)	-1.26	0.13	1.04	-0.88	0.30	1.39	0.38
7 (A2)	-1.62	0.13	0.87	-2.03	0.39	1.01	0.41
8 (A2)	-0.60	0.14	1.14	-0.75	0.29	0.71	0.15
9 (A2)	-0.05	0.13	0.84	-0.72	0.29	1.06	0.67
10 (A2)	0.28	0.14	0.60	-0.22	0.26	1.00	0.50
11 (A2)	-0.80	0.14	0.83	-1.27	0.34	0.86	0.47
12 (A2)	-0.27	0.14	0.78	0.04	0.26	0.76	0.31

13 (A 2)	-0.10	0.14	0.86	-1.39	0.35	0.59	1.29
14 (A 2)	-0.44	0.13	1.06	-1.85	0.43	0.82	1.41
15 (B 1)	-0.34	0.14	0.71	-0.25	0.27	1.05	0.09
16 (B 1)	-0.10	0.15	0.65	-0.35	0.28	0.34	0.25
17 (B 1)	0.28	0.14	0.82	1.39	0.24	0.70	1.11
18 (B 1)	0.33	0.14	0.65	-0.09	0.26	0.58	0.42
19 (B 1)	0.14	0.14	1.03	0.16	0.25	1.23	0.02
20 (B 1)	0.88	0.14	1.44	0.56	0.24	1.40	0.32
21 (B 1)	-0.26	0.15	1.03	0.40	0.24	0.97	0.66
22 (B 1)	0.31	0.14	0.97	-0.33	0.27	1.04	0.64
23 (B 2)	0.33	0.15	0.66	0.54	0.24	1.18	0.21
24 (B 2)	0.10	0.15	1.25	1.99	0.25	1.31	1.89
25 (B 2)	1.14	0.16	0.93	2.23	0.25	0.75	1.09
26 (B 2)	1.54	0.16	2.10	1.37	0.23	1.91	0.17
27 (B 2)	2.22	0.18	1.10	2.57	0.27	1.50	0.35
28 (B 2)	1.37	0.16	0.91	1.79	0.25	1.22	0.42
29 (C 1)	0.36	0.14	1.50	0.87	0.24	1.04	0.50
30 (C 1)	1.98	0.18	1.46	2.05	0.24	0.95	0.07
31 (C 2)	3.12	0.19	1.50	4.66	0.41	0.89	1.54

*アウトフィット平方平均値が0.6未満（オーバーフィット）ものは太字を、1.4以上（アンダーフィット）ものは太字に下線を引いている。また、学習者評価の推定値と教師評価の推定値の差が1.00以上のCDSには網掛けをしている。

学習者評価と教師評価を比較するために学習者評価の推定値と教師評価の推定値の差を算出し、その差が1.00以上であるCDSを探したところ、7つのCDSが該当した。（5・13・14・17・24・25・31）（表4 & 図1）

CDS5 (A 1)「葉書などに書かれた、短く簡単なメッセージを理解することができる。」やCDS13 (A 2)「公衆電話のような、日常生活で出会う機器についての簡単な使用方法の記述を理解することができる。」、CDS14 (A 2)「通り、レストラン、駅のような公共の場所や職場にある標識や掲示を理解することができる。」については、教師のほうが学習者よりもこれらのCDSを易しいと評価している。

CDS17(B 1)「長いテキストや複数の短いテキストをざっと目を通して、課題を遂行するために必要な情報を探することができる。」やCDS24(B 2)「専門用語を確認するために辞書が使えるのであれば、自分の専門以外の専門的な記事を理解することができる。」、CDS25(B 2)「読む目的やテキストの種類に応じて読む速度や読み方を変えながら、様々な種類のテキストをかなり楽に読むことができる。」、CDS31(C 2)「抽象的であったり、構造的に複雑であったり、高度に口語的であるような、文学的な文章や非文学的な文章を含む、実質的にあらゆる形式の書きことばを理解し、解釈することができる。」については、教師のほうが学習者よりもこれらのCDSを難しい

と評価している。

レベルで見ると、A1とA2レベルのCDSは教師が学習者よりも易しいと評価し、B2レベルのCDSでは教師が学習者より難しいと評価する傾向があるようではあるが、同一レベルのCDSと比較した場合、なぜこれらの7つのCDSが学習者と教師の評価において差が出るのかについては更なる調査が必要である。



図1：学習者評価と教師評価の比較（ロジット値）

CEFR-DIALANG読解CDS（31項目）の難易度推定値をCEFRの6レベルごとにまとめ、箱ひげ図にしたものが図2である。学習者評価と教師評価はほぼ同じような動きをしており、CEFRのレベルが上がるにつれて困難度も上がり全体の傾向としてCEFRの順序性が維持されている。しかし、B2とC1の間では学習者評価では差がなく、教師評価ではC1の方が易しいと判定されている。また、A1とA2のCDSは教師が学習者よりも易しいと判定し、B1は両者とも同じ、B2以上になると、教師のほうが学習者よりも難しいと判定していることが分かる。

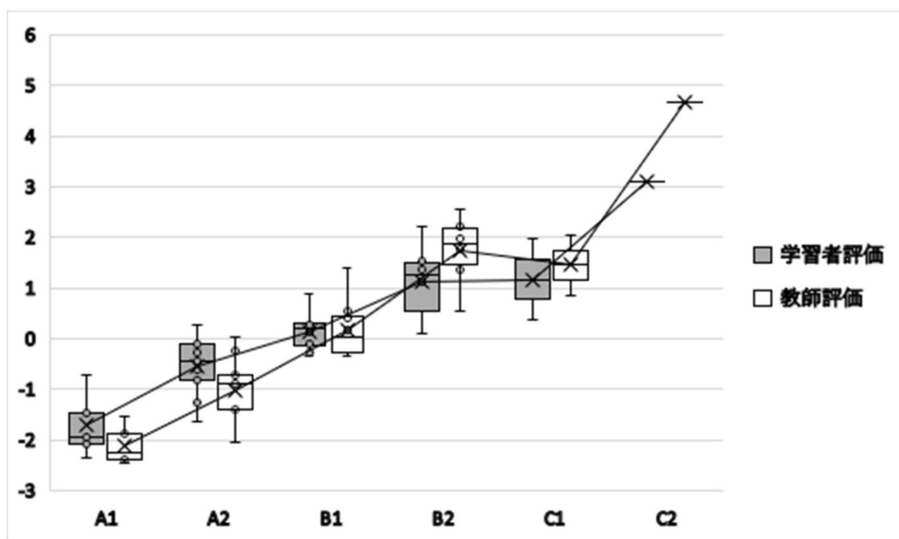


図 2：学習者評価と教師評価の箱ひげ図（ロジット値）

5. まとめ

本稿ではCEFRのCDSを基にして開発されたCEFR-DIALANG読解CDSを使用して、韓国人日本語学習者の自己評価データ（361名）と韓国人日本語教師の難易度評価データ（30名）を主にラッシュモデルで比較分析した。その結果、次の4点が明らかになった。1）学習者評価と教師評価において全体の傾向としてCEFRの順序性が維持されていたが、どちらの評価でもB2とC1の間に差が見られなかった。2）A1とA2のCDSは教師が学習者より易しいと評価するが、B2のCDSは教師が学習者よりも難しいと評価した。3）文章の特徴に関して具体的な記述がされていないCDSや語彙力について記述しているCDSは、学習者評価と教師評価においてモデルから外れたCDSと判断された。4）学習者評価と教師評価の間に差があるCDSがあることが確認された。ただし、学習者自己評価と教師難易度評価は評価をする際に念頭に置く対象が同一ではないので、結果を一般化するには更なる調査が必要である。

謝辞

調査に協力してくださった先生方及び学生さんに感謝申し上げます。

本研究はJSPS 科研費基盤研究（C）18K00722の助成を受けたものです。

参考文献

Alderson, J. C. (2005) Diagnosing Foreign Language Proficiency. New York: Continuum.

坂野永理（2015）「日本語学習者を対象にしたCEFRスイス版自己評価チェックリス

- トの検証」『第二言語としての日本語の習得研究』18,70-85.
- Bond, T., & Fox, C. (2007) *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (2nd ed). New York: Routledge.
- Council of Europe. (2001) *Common European Framework of Reference for Languages : Learning, Teaching, Assessment*. Cambridge : CUP. (吉島茂・大橋理枝・奥総一郎・松山明子訳編 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)
- 東伴子・代田智恵子・永田道子 (2017) 「行動主義にもとづいたヨーロッパにおける日本語オンラインテストの開発－新しい評価基準をめざして－」『Proceedings of The 7th International Conference on Computer Assisted Systems for Teaching & Learning Japanese (CASTEL/J)』, 162-167.
- 宮崎佳典・谷誠司 (2015) 「韓国人日本語学習者に対するCEFR読解尺度の妥当性調査」『Proceedings of The 6th International Conference on Computer Assisted Systems for Teaching & Learning Japanese (CASTEL/J)』, 127-130
- 野口裕之・大隅敦子 (2014) 『テストングの基礎理論』研究社
- North, B. (2005) *The Development of a Common Framework Scale of Descriptors of Language Proficiency Base on a Theory of Measurement*. System, 23, 445-465.
- 中島正剛・永田真代 (2006) 「CEFRの日本人外国語学習者への適用可能性」『外国語教育研究』 9, 5-24.
- 根岸雅史 (2006) 「CEFRの日本人外国語学習者への適用可能性の向上について」『言語情報学研究報告』14, 79-101. Retrieved from http://www.coelang.tufts.ac.jp/common/pdf/research_paper14/079.pdf
- 大隅敦子・野口裕之・熊谷龍一・石毛順子・長沼君主・和田敦子・伊東祐朗 (2006) 「日本語能力試験can-do-statements(試行版)とCEFR-DIALAGとの対応付けの試み」『第5回国際日本語OPIシンポジウム』 発表資料 欧州日本語OPI研究サークル
- 大友賢二 (1996) 『項目応答理論入門』大修館書店
- 静哲人 (2005) 『基礎から深く理解するラッシュモデリング』 関西大学出版部
- Suzuki, Y. (2015) *Self-assessment of Japanese as a second language: The role of experiences in the naturalistic acquisition*. Language Testing, 32, 63-81.
- 谷誠司・宮崎佳典・高田宏輝 (2017) 「ラッシュモデルによるCEFR(Common European Framework of Reference for Languages) 読解Can-do statementsの分析：韓国人日本語学習者を対象にした自己評価調査を基に」『常葉大学外国語学部紀要』33, 61-75.
- 投野由紀夫 (2013) 『CAN - DOリスト作成・活用 英語到達度指標CEFR - Jガイドブック』大修館書店