

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

谷 誠司, 宮崎 佳典, 高田 宏輝

Analysis of Reading Can-do Statements in CEFR (Common European Framework of Reference for Languages) Using Rasch Model Analysis.

Seiji TANI, Yoshinori MIYAZAKI, Hiroki TAKADA

2016年11月18日受理

抄 録

本研究では CEFR-DIALANG の読解尺度にある能力記述文 (Can-do statements) を使い、韓国人日本語学習者を対象に 5 件法の自己評価をしてもらった結果を Rasch model で分析をし、その結果の一部を報告する。

キーワード：ラッシュモデル、CEFR、読解 Can-do statements、韓国人日本語学習者、自己評価調査

1. はじめに

近年、外国語教育において外国語で何ができるのか (Can-do) という考えのもとに記述された言語能力尺度に関心が集まっている。その中で欧州評議会が開発した CEFR は、ヨーロッパの枠を超え、世界の外国語教育に導入されているが、その際に CEFR の言語能力尺度が欧州語以外の言語教育にも適用可能であるのかを検討する必要がある。

本稿では韓国人日本語学習者を対象に CEFR-DIALANG¹ の読解能力記述文(以下、CDS)の自己評価調査を実施し、1) 読解 CDS の項目困難度と CEFR のレベル分けとの対応関係、2) 自己評価による読解 CDS 平均値と日本語熟達度テスト(新 JLPT)の成績との関係について分析した結果を報告する。

2. 先行研究

CEFR の CDS を使った研究は数多くされているが、CDS の順位性を扱っている先行研究をいくつか紹介する。

中島・永田(2006)では東京外国語大学の学生約 360 名(1 年～4 年生)を対象に CEFR-DIALANG の CDS を使用して主専攻語と副専攻語の能力(聞く・読む・書く)について「できる・できない」で自己評価調査を行った。項目応答理論による分析の結果、おおむね CEFR の適用可能性を肯定する結果が出たが、「読む」においては「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(A1)」と「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(A2)」が CEFR の想定レベルより難しい結果が出た。

根岸(2006)では中島・永田(2006)の研究で CEFR の想定順序と異なった CDS を取り上げ、詳細な具体例を共に提示することで自己評価(4 件法)の精度が改善されるかを東京外国語大学の学生と高校生約 700 名を対象にして調査した。その結果、4 技能すべてにおいて具体例の提示が精度向上につながったと報告している。

大隈・野口・熊谷・石毛・長沼・和田・伊東(2006)では、日本語学習者 1068 名(主な母語：中国語・韓国語・インドネシア語)を対象に旧 JLPT 試行版 CDS を使用して 1) 経験の有無と 2) 5 件法でどのくらいできるか(と思うか)を尋ねた。旧 JLPT 試行版 CDS の中から CEFR-DIALANG の CDS を比較し総合的に同一レベルにあると見られる CDS を選択し、経験ありにマークした者のみを対象に IRT 分析を行った。困難度推定値から、「聞く」と「書く」の自己評価結果は CEFR の想定している困難度とおおむね一致していたが、「読む」は一致していなかった。「読む」においては、「新

¹ 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・システムであり、受検者の自己の言語能力を自己評価するために DIALANG self-assessment statements が CEFR をもとにして作成されている。中島・永田(2006)、根岸(2006)、大隈・野口・熊谷・石毛・長沼・和田・伊東(2006)、宮崎・谷(2015)、そして本研究では DIALANG self-assessment statements、略して CEFR-DIALANG の CDS を使用している。CEFR の CDS と CEFR-DIALANG の 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 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)

聞には含まれている広告やチラシなどを見て、必要な情報がとれる。(A2 相当)」と「短いメモや年賀状などに書かれた決まりきった文を読んで、書かれている内容が理解できる(A1 相当)」がCEFRの想定レベルより難しく、「ヘッドライナーや掃除機など、操作が簡単な家電製品についての取り扱い説明書を読んで理解できる(B1 相当)」が易しいと捉えられていた。

坂野(2015)ではCEFR スイス版自己評価チェックリストを使用して、日本国内の大学で日本語を履修している日本語学習者313名(主な国籍は中国・アメリカ・韓国・フランスなど)を対象に、聞く・読む・口頭でのやり取り・口頭での産出活動・方略・言語の質・書くについてのCDSを5件法でどのくらいできるかを評価させた。領域ごとにラッシュモデルで項目難易度を算出し、各領域においてCEFRのレベルごとに平均値を算出した。その結果、すべての領域においてA1からB2まではレベルごと平均値で0.3以上の差を持って難易度が上がっていた。しかし、「聞く」・「口頭でのやり取り」・「口頭での産出活動」・「方略」・「言語の質」ではC1とC2の間で、「書く」ではB2からC2までの間で十分な難易度の差がなかったが、「読む」はA1からC2のすべてのレベルで十分な難易度の差があった。

宮崎・谷(2015)では韓国人日本語学習者(大学生)205名²を対象にCEFR-DIALANGの読解CDS(31項目)を使用して、日本語の読解能力を5件法で自己評価(「1. 全然できない」～「5. 問題なくできる」)した。素点ベースでの分析の結果、大局的にはCEFRの順位性は維持されているが、CEFRの順位性が瓦解しているCDSがあることが明らかになった。具体的には「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を例えばテキストのある部分を読み返すことができれば理解できる」(CDS2、A1)、「もっとも一般的で日常的な状況でよく出くわす、簡単な掲示にでているような、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ことができる」(CDS4、A1)、「日常的な言葉で書かれた短くて簡単なテキストを理解することができる」(CDS7、A2)、「短く、簡単な個人的な手紙を理解することができる」(CDS11、A2)、「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」(CDS14、A2)、「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」(CDS29、C1)は同等レベルのCDSと比較して、韓国人日本語学習者は易しいと捉えられていた。

以上のようにこれまでに研究ではCEFRの読解尺度に関しては全体的にCEFRが想定する順位性が維持されているが、個々のCDSに関しては順位性から逸脱しているものも見られていると報告されている。また、日本語教育の分野においてラッシュモデルを使って、CEFRの読解CDSの順位性を検証している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。

² この205名は本研究での調査に協力してくれた361名の中に含まれる。

3. 研究課題

本稿では CEFR-DIALANG の読解 CDS が韓国人日本語学習者に適用可能であるかを検証するために、ラッシュモデルを使って、CEFR-DIALANG の読解 CDS 自己評価データを分析し、また日本語熟達度テスト(新 JLPT)の取得と CEFR-DIALANG の読解 CDS 自己評価間の関係を分析する。具体的には次の 2 点を研究課題とする。

- 1) CEFR-DIALANG の読解 CDS の順位性が韓国人日本語学習者にもみられるか。また CEFR の順位性が瓦解している CDS があればどのような原因があるか。
- 2) 日本語熟達度テスト(新 JLPT)の取得級によって、CEFR-DIALANG の読解 CDS の自己評価がどのように変化するか。

4. 調査方法

4.1. 協力者

韓国の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 名が調査に協力してくれた。協力者の特徴は表 1 のようにまとめられる。

表 1：調査協力者の特徴³

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(26.6%)、2 年生：81(22.4%)、3 年生：81(22.4%)、4 年生：88(24.4%)、大学院生：12(3.3%)
専攻	日本語関連専攻・複数専攻者：138(38.2%)、他専攻者：215(59.6%)
日本語学習期間	6 か月未満：75(20.8%)、1 年～ 2 年：93(25.8%)、2 年～ 3 年：44(12.2%)、3 年～ 4 年：55(15.2%)、4 年以上：85(23.5%)
新 JLPT 合格者	N1 合格：77(21.3%)、N2 合格：26(7.2%)、N3 合格：11(3.0%)、N4 合格：9(2.5%)、N5 合格：1(0.3%)

4.2. 調査時期と手順

調査は 3 回に分けて行った(S 大学：2011 年 11 月～ 12 月の間に 2 回、K 大学：2012 年 5 月に 1 回、K 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。

4.3. 調査票

- 1) CEFR-DIALANG にある読解 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

³ 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。

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することによって得た。

- 2) 学年、専攻、日本語学習期間、日本語関連能力試験の成績などについての情報を得るためにフェイスシートを作成した。こちらも質問および回答はすべて韓国語にした。

5. 結果と考察

5.1. 基礎統計量と信頼性

CDS 調査票の基礎統計量は表 2 のようになり、信頼性係数の推定値 (α 係数) は .97 であった。なお、人数が項目ごとに異なっているのは、「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る。

表 2：基礎統計量

CDS No. (CEFR レベル)	CDS 記述内容	人数	平均値	標準 偏差
1 (A1)	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に関して概要の把握ができる。特にテキストの内容を理解するのに助けとなる絵が含まれていれば、さらに安易に概要の把握ができる。	334	3.91	0.98
2 (A1)	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を、例えばテキストのある部分を読み返すことができれば理解できる。	345	4.11	1.00
3 (A1)	短く簡単に書かれた指示 (特に絵を含む) に従うことができる。	335	3.96	1.04
4 (A1)	もっとも一般的で日常的な状況でよく出くわす、簡単な掲示にでているような、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ことができる。	335	4.05	0.97
5 (A1)	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。	276	3.75	1.10
6 (A2)	もっとも頻度の高い単語で書かれていたり世界的に共通して使われる単語を含んだりする短くて簡単なテキストを理解することができる。	311	3.81	1.06
7 (A2)	日常的な言葉で書かれた短くて簡単なテキストを理解することができる。	337	3.88	1.11
8 (A2)	自分の仕事に関連した短くて簡単なテキストを理解することができる。	236	3.72	1.09
9 (A2)	広告、パンフレット、メニュー、時刻表などの簡単な言語資料の中の特定の情報を見つけることができる。	267	3.57	1.08

10 (A2)	手紙、パンフレット、新聞の短い事件記事のような簡潔に書かれたテキストの中から特定の情報を取り出すことができる。	243	3.50	1.09
11 (A2)	短く、簡単な個人的な手紙を理解することができる。	269	3.76	1.17
12 (A2)	身近な話題について日常の定型の手紙やファックスを理解することができる。	234	3.67	1.08
13 (A2)	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。	236	3.59	1.08
14 (A2)	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。	272	3.64	1.09
15 (B1)	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を理解することができる。	233	3.67	1.09
16 (B1)	手紙、パンフレット、短い公的な文書といった、日常的な文章において必要とする一般的な情報を見つけて、理解することができる。	227	3.59	1.09
17 (B1)	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。	239	3.49	1.05
18 (B1)	なじみのある話題に関する簡単な新聞記事において重要な点を認識することができる。	218	3.53	1.07
19 (B1)	はっきりと主張が書かれたテキストの主要な結論を把握できる。	229	3.62	1.12
20 (B1)	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。	256	3.12	0.91
21 (B1)	個人の手紙を読んで、出来事、感情、希望の表現を理解することができ、友達や知り合いと文通できる。	219	3.69	1.08
22 (B1)	機器に関する、明瞭に書かれた簡単な使用説明を理解することができる。	223	3.5	1.13
23 (B2)	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。	202	3.46	1.11
24 (B2)	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。	218	3.48	1.11
25 (B2)	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。	179	3.22	1.08
26 (B2)	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。	197	3.06	1.10
27 (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。	153	2.96	1.10
28 (B2)	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を理解できる。	176	3.19	1.12
29 (C1)	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。	238	3.24	1.15

30 (C1)	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。	156	2.91	1.06
31 (C2)	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。	150	2.61	1.11

5.2. CEFR-DIALANG の読解 CDS の順位性の成否

CEFR-DIALANG の読解 CDS の順位性が韓国人日本語学習者にも当てはまるかを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ため、Bond and Fox(2007) に従い、ミスフィット⁴の基準を0.6～1.4の範囲にした。全データを分析した結果、ミスフィット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。その結果、外れ値の影響を抑えたインフィット値を見ると、3つのCDS (26・29・31) がアンダーフィット (underfit) と判断された (表3)。

アンダーフィット (underfit) のCDSを見ると、CDS26(B2)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。」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原因と考えられる。また、CDS29(C1)「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(C2)「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、逆に具体的な文書をイメージしにくいことが原因と思われる。

表3：ラッシュモデルによる分析結果

CDS No. (CEFR レベル)	推定値	標準誤差	インフィット 平方平均	アウトフィット 平方平均
2(A1)	-2.34	0.13	1.01	0.87
4(A1)	-2.08	0.13	0.86	0.81
3(A1)	-1.93	0.13	0.97	0.91
7(A2)	-1.62	0.13	0.93	0.87

⁴ 小泉・飯村 (2010) によると、「ラッシュモデリングにおいては、一般的にはインフィットの平方平均が0.70～1.30の範囲に入らない場合をミスフィット (不適合; misfit) と解釈し、0.7未満の場合をオーバーフィット (overfit)、1.30より高い場合をアンダーフィット (underfit) と呼ぶ (静, 2005)」とある。オーバーフィット (overfit) はモデルに適合しすぎている場合を、アンダーフィット (underfit) はモデルに適合していない場合を指す。オーバーフィット (overfit) はあまり問題とはしないが、アンダーフィット (underfit) は「有害な存在 (あってはならない)」ので「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。

1(A1)	-1.47	0.13	0.92	0.94
6(A2)	-1.26	0.13	1.04	1.04
11(A2)	-0.80	0.14	0.81	0.83
5(A1)	-0.71	0.13	0.77	0.73
8(A2)	-0.6	0.14	1.16	1.14
14(A2)	-0.44	0.13	1.07	1.06
15(B1)	-0.34	0.14	0.69	0.71
12(A2)	-0.27	0.14	0.76	0.78
21(B1)	-0.26	0.15	0.98	1.03
13(A2)	-0.10	0.14	0.86	0.86
16(B1)	-0.10	0.15	0.68	0.65
9(A2)	-0.05	0.13	0.85	0.84
24(B2)	0.10	0.15	1.27	1.25
19(B1)	0.14	0.14	1.06	1.03
10(A2)	0.28	0.14	0.61	0.60
17(B1)	0.28	0.14	0.81	0.82
22(B1)	0.31	0.14	1.01	0.97
23(B2)	0.33	0.15	0.67	0.66
18(B1)	0.33	0.14	0.66	0.65
29(C1)	0.36	0.14	1.50	1.50
20(B1)	0.88	0.14	1.33	1.44
25(B2)	1.14	0.16	0.93	0.93
28(B2)	1.37	0.16	0.90	0.91
26(B2)	1.54	0.16	2.15	2.10
30(C1)	1.98	0.18	1.39	1.46
27(B2)	2.22	0.18	1.06	1.10
31(C2)	3.12	0.19	1.44	1.50
平均	0.00	0.15	1.01	1.00
標準偏差	1.24	0.01	0.31	0.32

また、算出された困難度値(通常、-3が最も易しく+3が最も難しい)順にCDSを並べ(図1)、隣接するCDS間の困難度値の差が0.3以上あるところを目安にして、暫定的な分割点を設定した。例えば、A1レベルに入るCDSはCDS2、3、4であり、A2レベルに入るCDSはCDS1、7、6であることが分かる。なお、図にある「#」は回答者2名、「.」は回答者1名を表す。Mは平均、Sは1標準偏差、Tは2標準偏差を示す。

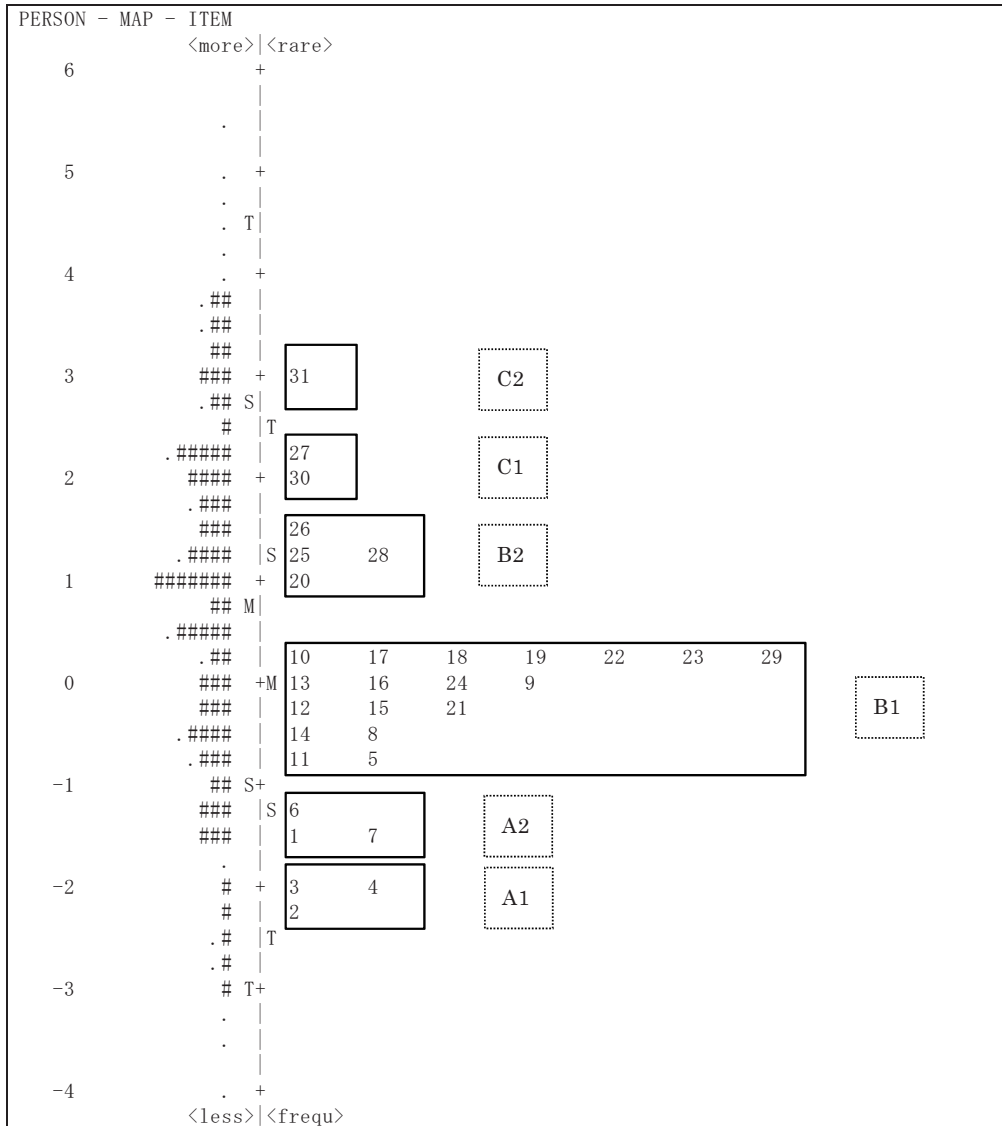
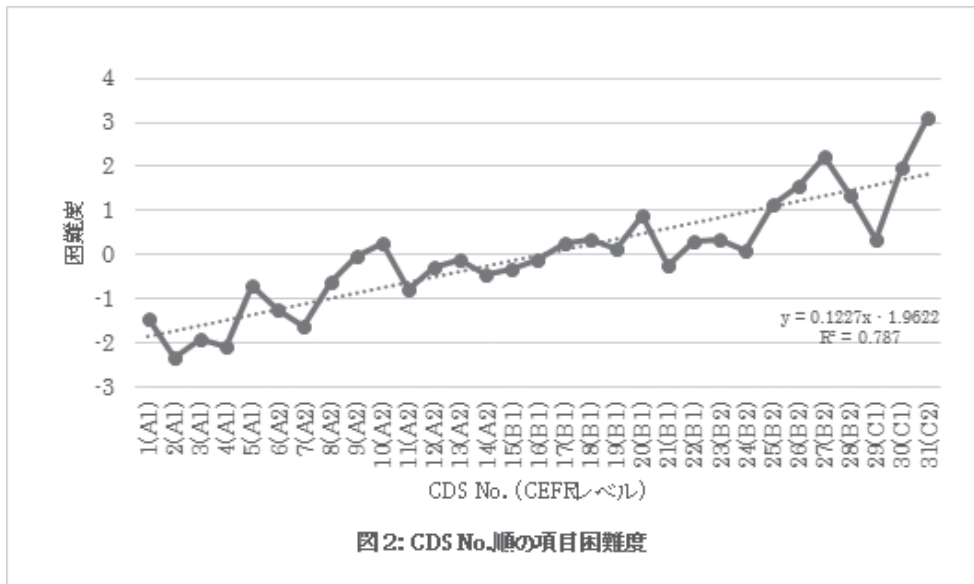


図1 Person Item Map

図2はCDSの番号順に項目困難度をグラフ化したものである。全体的な傾向としては番号の若いCDSは項目難易度が低く、番号が上がるごとに項目難易度も上がる傾向がみられ、全体的にはCEFRの順序性と一致しているといえる。

一方で、CEFRの想定しているレベルと今回の分析による項目困難度にずれがみられるCDSもある。例えば、CDS1やCDS5はCEFRレベルで最も易しいA1レベルと想定されているが、今回の分析では項目困難度が他のA1レベルのCDSと比較して高く出ている。



このように図1で設定した暫定的な分割点から外れている CDS(1・5・8・9・10・11・12・13・14・20・23・24・27・29)について、CEFR の順位性が瓦解した原因を表4にまとめた。

表4：順位性が異なった CDS とその考えられる原因のまとめ

CDS No. (CEFR レベル)	記述内容 (要点のみ抜粋)	レベル差	考えられる原因
1(A1)	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	A1 → A2 (易→難)	今回の分析で A1 レベルと判定された CDS(2・3・4)を見ると、単語や句レベルのテキストが想定されているようであるので、単語や句レベル以上のテキストと考えられたため困難度が上がったか。
5(A1)	テキスト葉書などに書かれた、短く簡単なメッセージ	A1 → B1 (易→難)	中島・永田(2006)、大隅他(2006)と同じ結果。日本のはがきは欧州のポストカードより文章量が多いためか。
8(A2)	自分の仕事に関連した短くて簡単なテキスト	A2 → B1 (易→難)	CDS 7(A2)は「 <u>日常的な言葉</u> で書かれた短くて簡単なテキストを理解することができる」とあり、CDS8は「 <u>自分の仕事に関連した</u> 短くて簡単なテキストを理解することができる」とあり、2つのCDSの違いは下線部だけである。今回の分析で A2 レベルと判定された CDS(1・6・7)を見ると、日常的で短く簡単なテキストが想定されているようであるので、「仕事に関連した」という部分が難しいと考えられたために困難度が上がったか。

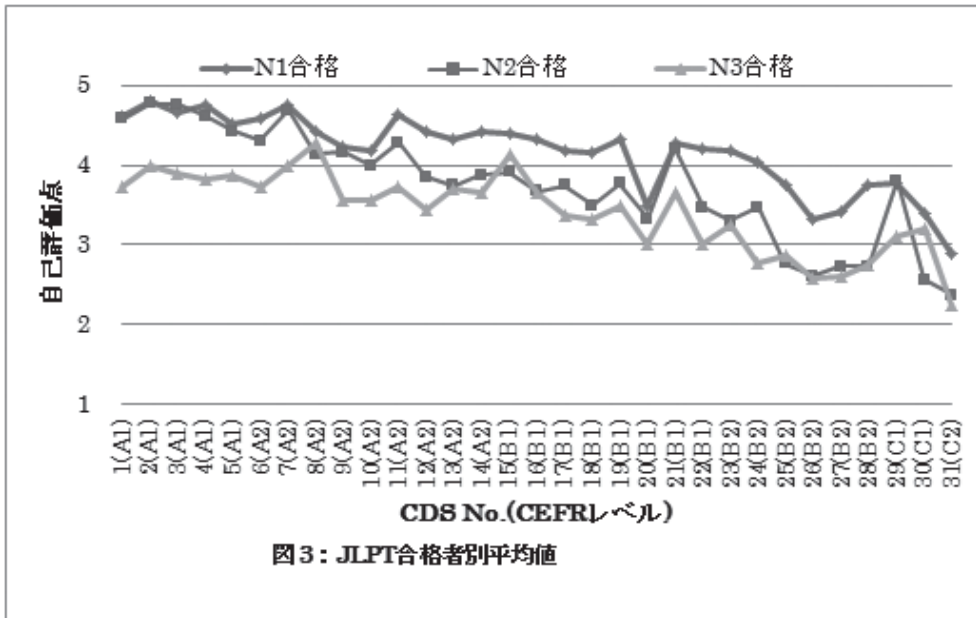
9(A2)	簡単な言語資料 (例：広告、パンフレット、メニュー、時刻表など)、特定の情報を見つける	A2 → B1 (易→難)	大隅他(2006)でも困難度が上がった。テキストの種類が難しいと考えられ、B1 レベルになったか。
10(A2)	簡潔に書かれたテキスト(例：手紙、パンフレット、新聞の短い事件記事)、特定の情報を見つける	A2 → B1 (易→難)	9 と同じ原因か。
11(A2)	短く、簡単な個人的な手紙	A2 → B1 (易→難)	CDS16 は「 <u>手紙</u> 、パンフレット、短い公的な文書といった、日常的な文章において必要とする一般的な情報を見つけて、理解することができる」、CDS 21 は「 <u>個人の手紙</u> を読んで、出来事、感情、希望の表現を理解することができ、 <u>友達や知り合いと文通</u> できる」とあり、両 CDS との CEFR の設定レベルも今回の分析結果も B1 である。CDS11 が A2 から B2 と判断された原因として親しい人との手紙のやり取りだと B1 レベルと判断されたためか。ただ、CDS11 の場合、「短く、簡単な」というテキスト上の特徴からは A2 レベルが妥当とも思われる。
12(A2)	身近な話題について日常の定型の手紙やファックス	A2 → B1 (易→難)	今回の分析で A2 レベルと判定された CDS(1・6・7)を見ると、日常的で短く簡単なテキストが想定されているようであるので、手紙やファックスは文書量が多く、やや複雑と考えられたためか。
13(A2)	日常生活で出くわす機器の簡単な使用方法(例：公衆電話)	A2 → B1 (易→難)	中島・永田(2006)と同じ結果。「機器についての簡単な使用方法」が難しいとイメージされたか。
14(A2)	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	A2 → B1 (易→難)	公共の場所や職場での掲示文はテキストが長めと複雑と判断されたためか。ただ、標識に関してはテキストの長さはかなり短いと思われるので、A2 レベルでも妥当とも思われる。
20(B1)	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	B1 → B2 (易→難)	中島・永田(2006)では弁別力が低い項目。文末がほかの項目と異なり、「できるわけではない」で終わっているため、回答者を混乱させたか。今回の分析で B1 レベルと判定された CDS と比較すると、記述の抽象度が高く、テキストが具体的にイメージしにくい。

23(B2)	自分の専門分野に関連する通信文(手紙・メールなど)	B2 → B1 (難→易)	CDS15は「 <u>自分の専門分野</u> や <u>関心のある話題</u> に関して簡潔に書かれたテキストを理解することができる」とあり、CEFRの設定レベルも今回も分析でもB1であった。CDS23がB2設定レベルにかかわらず、今回の分析ではB1と判定されたのは「 <u>自分の専門分野</u> 」だと易しく感じるためか。
24(B2)	辞書が使えるれば、自分の専門以外の専門的な記事	B2 → B1 (難→易)	中島・永田(2006)では弁別力が低い項目。辞書使用が認められることで能力が低くてもできるといった回答者が多かったためか。
27(B2)	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	B2 → C1 (易→難)	「 <u>広範囲にわたる専門的な話題</u> 」や「 <u>内容と関連性をすばやく確認</u> 」の部分が難しく感じさせたか。
29(C1)	ときどき辞書を使用すれば、どんな文書でも	C1 → B1 (難→易)	24と同じ原因か。

5.3. 日本語熟達度テスト(新 JLPT)の取得級と自己評価による CDS 平均値の関係

日本語熟達度テスト(新 JLPT)の取得級によって、各 CDS への自己評価がどのように変化するかを見るために、新 JLPT の N1 合格者～ N3 合格者の各 CDS に対する自己評価の平均値をグラフ化した(図3参照)。

CDS1～10まではN1とN2の間に平均値の差がない。A1とA2のCDSなので、差が出ないと思われる。また、CDS21(B1)とCDS29(C1)はN1とN2の間で差がない。CDSの記述内容(CDS21「個人の手紙を読んで、友達や知り合いと文通」、CDS29「ときどき辞書を使用すれば、どんな文書でも」)から「友人間での手紙」や「辞書使用」という部分が原因と考えられる。CDS8(A2)・CDS13(A2)・CDS15(B1)・CDS16(B1)・CDS23(B2)・CDS25～28(すべてB2)・CDS30(C1)はN2とN3の間では差がなかったり逆転していたりする。下位レベルのCDSの記述内容(CDS8「自分の仕事に関連した短くて簡単なテキストを理解することができる。」、CDS13「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。」、CDS15「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を理解することができる。」、CDS16「手紙、パンフレット、短い公的な文書といった、日常的な文章において必要とする一般的な情報を見つけて、理解することができる。」)から、「自分の仕事」「短くて簡単」「日常生活」「自分の専門分野」「日常的」という表現が原因と考えられる。また、N3程度の学習者はB2以上のCDSを正確に自己評価できない可能性がある。



6. まとめと今後の課題

3. の研究課題に従って、結果をまとめる。

1) CEFR-DIALANG の読解 CDS の順位性が韓国人日本語学習者にもみられるか。また CEFR の順位性が瓦解している CDS があればどのような原因があるか。

ラッシュモデルでの分析の結果、CDS26・29・31 がアンダーフィット (underfit) と判断された。CDS26(B2) 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。」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ミスフィットになった原因と考えられる。また、CDS29(C1)「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(C2)「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、逆に具体的な文書をイメージしにくいことがミスフィットになった原因と思われる。

また、CEFR の順位性が瓦解している CDS もあった。CDS の記述内容や他の CDS との比較によって、①葉書、②自分の専門分野、③親しい人との手紙のやり取り、④辞書使用、といった内容が含まれる CDS は想定しているレベルより易しくなること、一方で ①日常生活で会う機器についての簡単な使用方法、②広告、パンフレット、メニューなどは難しくなることが分かった。

2) 日本語熟達度テスト (新 JLPT) の取得級によって、CEFR-DIALANG の読解 CDS の自己評価がどのように変化するか。

A1 と A2 といった下のレベルの CDS では N1 取得者と N2 取得者には差がない。「親しい人と手紙のやりとり (CDS21、B1 レベル)」と「辞書使用 (CDS 29、C1 レベル)」の CDS では N1 取得者と N2 取得者には差がない。「自分の仕事」「短くて簡単」「日常生活」「自分の専門分野」「日常的」という表現がある A2 や B1 レベルの CDS や B2 レベル以上の CDS において N2 取得者と N3 取得者の間には差がなく、非熟達者は過大評価をしやすいという研究結果 (Kruger & Dunning, 1999) から、N3 レベルの学習者はこれらの CDS を正確に自己評価できていない可能性も考えられる。

今後の課題としては、まず、ミスフィットの回答者が多かったので、ミスフィットの回答者の回答を 1 つ 1 つ見ていき、なぜミスフィットが起こったかの原因を明らかにしていく必要がある。また、ミスフィットであった CDS や CEFR が設定した順位性が瓦解した CDS の扱いについても教師による CEFR-DIALANG の読解 CDS の難易度評価調査等の結果も踏まえながら、慎重に決めていくべきであろう。

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) 26370619 の助成を受けたものです。

参考文献

- Bond, T., & Fox, C. (2007) *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (2nd ed.). New York, NY: Routledge.
- Kruger, J., & Dunning, D. (1999) Unskilled and Unaware of it : How difficulties in recognizing one's own incompetence lead to inflated self-assessment. *Journal of Personality and Social Psychology*, 77, 1121-1134.
- 小泉利恵・飯村英樹 (2010) 「ニューラルテスト理論の特徴：古典的テスト理論・ラッシュモデリングとの比較から」『日本言語テスト学会研究紀要』13, 91-109.
- 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 於 University of Hawaii, Kapiolani Community College (米国, ハワイ州)
- 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.
- 大隈敦子・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
- 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

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.

静哲人 (2005) 『基礎から深く理解するラッシュモデリング 項目応答理論とは似て異なる測定のパラダイム』大阪：関西大学出版部

