

潜在ランク理論による CEFR(Common European Framework 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

Analysis of reading Can-do statements in CEFR(Common European Framework for Languages) using Laten Rank Theory

Seiji TANI

2015年11月20日受理

抄 録

潜在ランク理論は Shojima(2007) によって提案された新しいテスト理論である。潜在ランク理論は素点ベースの古典的テスト理論と異なり、潜在的な変数（能力値）を対象とする点では項目応答理論と同じであるが、受験者の能力を連続尺度上で評価する項目応答理論とは違い、段階評価をする。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潜在ランク理論で分析をし、その結果の一部を報告する。

キーワード：潜在ランク理論, CEFR(Common European Framework for Languages), 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

潜在ランク理論は Shojima(2007) によって提案された新しいテスト理論である。潜在ランク理論は素点ベースの古典的テスト理論と異なり、潜在的な変数（能力値）を対象とする点では項目応答理論と同じであるが、受験者の能力を連続尺度上で評価する項目応答理論とは違い、段階評価をする点異なる。

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

2. 方法

2.1. 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 名が本調査に参加してくれた。361 名の協力者の特徴は表 1 のようにまとめられる。361 名のうち、すべて無回答やすべて同じ回答など適切な回答が得られていないデータを除き、最終的に 350 名分のデータが分析の対象となった。

表 1：調査協力者の特徴

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%)

注：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。

2.2. 調査時期と手順

調査は 3 回に分けて行った (1 回目：S 大学で 2011 年 11 月～12 月、2 回目：K 大学で 2012 年 5 月、3 回目：K 大学と D 大学で 2012 年 12 月)。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。

2.3. 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 (社会学系の修士課程) を修了した韓国語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。

2.4. 分析方法

潜在ランク理論による分析には Exametrika5.3(shojima,2008) を使用した。分析は

データの数が 10,000 以下で小さいため、自己組織化マップ (self-organizing map, SOM) のメカニズムを利用した設定で行った。また、自己評価は 5 件法のリッカート尺度を用いているため拡張モデルの 1 つである段階的ニューラルテスト (Graded neural test: GNT) を採用し、CEFR は言語能力レベルを A1 レベルから C2 レベルまでの 6 段階で設定しているため、潜在ランク数を 6 とした。各項目に対する単調増加制約や事前分布、目標潜在ランク分布などの指定はしなかった。なお、実際の分析にあたっては清水 (n.d.) を参考にさせてもらった。

3. 結果

3.1. 基本統計量

根岸 (2006) の指摘に従い、「経験あり」にマークした回答のみを分析対象とした。CEFR-DIALANG にある読解 CDS を使用した自己評価の記述統計は表 2 のとおりである。また信頼性係数の推定値 (α 係数) は .97 であった。

表 2：CEFR-DIALANG の読解 CDS を使用した自己評価の記述統計

	基本統計量
受検者数	350
項目数	31
最小値	5
最大値	153
中央値	73.5
平均値	74.28
分散	1706.70
標準偏差	41.31

表 3 に CEFR-DIALANG にある読解 CDS を使用した自己評価の TRP,LRD と RMD を示した。テスト参照プロファイル (TRP) を見ると、およそ 10 点刻みでランクが上がっており、後述する項目参照プロファイル (item reference profile, IRP) においても、傾きには違いはあるがすべての項目で単調増加であったことから、潜在ランク理論の前提である、潜在ランクの順序配置条件は満たされていると考えた。潜在ランク分布 (LRD) では、ランクが最も低い Rank1 に該当する学習者が最も多く、Rank 5 が最も少ないことが分かる

表 3：CEFR-DIALANG にある読解 CDS を使用した自己評価の TRP,LRD と RMD

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5	Rank 6
テスト参照プロファイル (TRP) ⁱ	81.99	90.07	102.06	113.65	125.12	132.55

潜在ランク分布 (LRD) ⁱⁱ	124	39	52	51	24	60
ランク・メンバーシップ 分布 (RMD) ⁱⁱⁱ	109.19	51.20	52.01	49.83	26.78	60.99
相対 TRP	0.53	0.58	0.66	0.73	0.81	0.86
相対 LRD	0.35	0.11	0.15	0.15	0.07	0.17
相対 RMD	0.31	0.15	0.15	0.14	0.08	0.17

「テストが受検者たちをいったい何段階で評価できるのかという分解能を知る手がかり」(荘島、2010) であるテスト適応度では、小泉・飯村 (2010) に従い、CFI と RMSEA を確認した。その結果、CFI は 0.997^{iv}、RMSEA は 0.004^v となり、今回設定したランク数が 6 つのモデルの適合度は十分高いと判断した。

3.2. 潜在ランク理論によるランクと CEFR の設定したレベルとの関係

潜在ランク理論による CEFR-DIALANG の読解 CDS の項目識別度と項目難易度は表 4 のようになった。

表 4：潜在ランク理論による CEFR-DIALANG の読解 CDS の項目識別度と項目難易度

CDS 番号 (CEFR レベル)	CDS 記述内容	項目識別度	項目難易度
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	0.50	1
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	0.44	2
3(A1)	短く簡単に書かれた指示 (特に絵を含む)	0.52	2
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	0.44	2
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	0.46	2
7(A2)	日常的な言葉で書かれた短くて簡単なテキスト	0.57	2
5(A1)	葉書などに書かれた、短く簡単なメッセージ	0.46	3
8(A2)	自分の仕事に関連した短くて簡単なテキスト	0.44	3

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、特定の情報を見つける	0.43	3
11(A2)	短く、簡単な個人的な手紙	0.55	3
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	0.46	3
13(A2)	公衆電話のような、日常生活で出くわす機器についての簡単な使用方法	0.42	3
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	0.44	3
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	0.44	3
21(B1)	個人の手紙を読んで、友達や知り合いと文通し	0.43	3
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、特定の情報	0.47	4
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	0.46	4
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必要な情報を探す	0.46	4
18(B1)	なじみのある話題に関する簡単な新聞記事	0.46	4
19(B1)	はっきりと主張が書かれたテキストの主要な結論	0.45	4
22(B1)	機器に関する、明瞭に書かれた簡単な使用説明	0.46	4
23(B2)	自分の専門分野に関連する通信文（手紙・メールなど）	0.51	4
24(B2)	辞書が使えるば、自分の専門以外の専門的な記事	0.31	4
29(C1)	ときどき辞書を使用すれば、どんな文書でも	0.39	4
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。	0.29	5
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	0.34	5
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	0.45	5

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	0.25	6
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。	0.37	6
30(C1)	難しい箇所を読み返すことができれば、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	0.37	6
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	0.29	6

潜在ランク理論による分析で CEFR-DIALANG の読解 CDS の順位性が維持されているかを明らかにするために、CEFR が設定している CDS のレベルと潜在ランク理論によるランクの関係を暫定的に表 5 のような対応関係があると設定した。

表 5：CEFR のレベルと潜在ランク理論の暫定的な関係

CEFR のレベル	潜在ランク
C2	6
C1	5
B2	4
B1	3
A2	2
A1	1

次に縦軸に CEFR のレベル、横軸に潜在ランク理論によるランクを置いて、それぞれの CDS を並べると、表 6 のようになった。CEFR が設定しているレベルと潜在ランク理論によるランクが一致している CDS は網がかかっている部分にある CDS で、全 31 の CDS 中、8 つの CDS であり、一致率は 25.8% である。また、隣接レベルに入っている CDS も一致しているとした場合、CEFR と潜在ランク理論のランクが一致している CDS は 26 になり、一致率も 83.8% になる。また、スピアマンの順位相関係数を算出したところ、 $\rho = .84$ であった。

表 6：CEFR のレベルと潜在ランク理論によるランクの対応関係

		潜在ランク理論によるランク					
		1(A1)	2(A2)	3(B1)	4(B2)	5(C1)	6(C2)
CEFR	C2						31
	C1						30
	B2				23,24	25,28	26,27
	B1			15,21	16,17, 18,19, 22	20	
	A2		6, 7	8, 9, 11,12, 13,14	10		
	A1	2	1, 3, 4	5			

3.3. 日本語熟達度テスト (新 JLPT) の取得級と潜在ランク理論によるランクの関係

調査票で調査協力者の全般的な日本語熟達度を知るために新 JLPT の所得級に関する質問項目を入れておいた。潜在ランク理論によって調査協力者は 6 ランクに分けられるが、それぞれのランクにいる調査協力者の新 JLPT の取得級は表 7 のようになった。表 7 を見ると、上位ランクにいる調査協力者のほうが新 JLPT の上位級に合格している割合が多く、ランクが下がるごとに新 JLPT の取得級が下がっていくことが分かる。

表 7：日本語熟達度テスト (新 JLPT) の取得級と潜在ランク理論によるランクの関係

		Rank1	Rank2	Rank3	Rank4	Rank5	Rank6
人数		124(35.4%)	39(11.1%)	52(14.9%)	51(14.6%)	24(6.9%)	60(17.1%)
新 JLPT	N1	0(0%)	2(5.1%)	9(17.3%)	13(25.5%)	11(45.8%)	40(66.7%)
	N2	0(0%)	3(7.7%)	5(9.6%)	5(9.8%)	5(20.8%)	8(13.3%)
	N3	3(2.4%)	2(5.1%)	0(0%)	2(3.9%)	2(8.3%)	1(1.67%)
	N4	4(3.2%)	1(2.6%)	2(3.8%)	2(3.9%)	0(0%)	0(0%)
	なし	117(94.3%)	31(79.5%)	36(69.2%)	29(56.9%)	6(25%)	11(18.3%)

3.3. 自己評価の総得点と潜在ランク理論によるランクの関係

潜在ランク理論によるランクごとの自己評価の総得点の違いをみるためにランクごとの自己評価の総得点のヒストグラムを描いた (図 1)。平均値ではランク間でお

よそ 10 点の差があったが、ランクごとの得点分布をみると、ランク間で違いがあることが分かる。

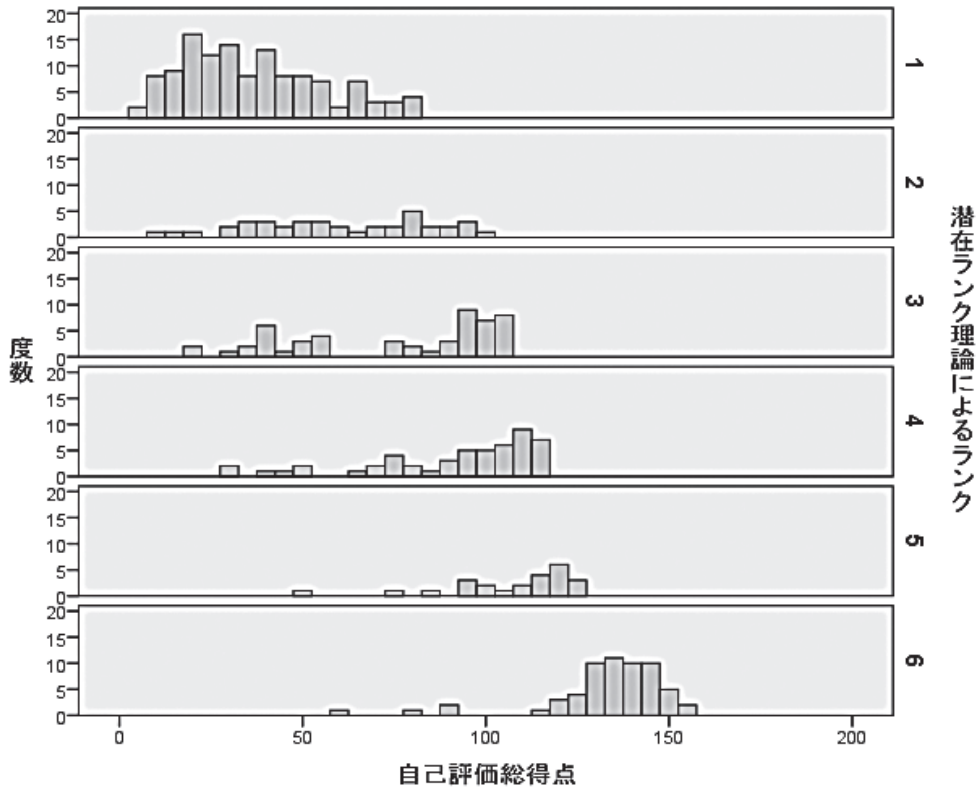


図 1：潜在ランク理論によるランクごとの CEFR-DIALANG の読解 CDS 自己評価の総得点のヒストグラム

4. まとめと今後の課題

CEFR-DIALANG の読解尺度にある CDS を使い、韓国語日本語学習者に対して 5 件法で自己評価をしてもらった結果を潜在ランク理論から分析を行った。その結果、CEFR が設定しているレベルと潜在ランク理論によるランクとが一致している CDS は 31 の CDS のうち、8 つの CDS であり、一致率は 25.8%であったが、隣接レベルの CDS も一致していたとした場合、一致している CDS は 26 になり、一致率も 83.8% になった。また、スピアマンの順位相関係数を算出したところ、 $\rho = .84$ であったことから、大局的な観点からは CEFR が設定したレベル設定が韓国語日本語学習者にも適用できると言える。

潜在ランク理論によるランク分けと全般的な日本語熟達度の関係を調査協力者の新 JLPT の取得級から分析したが、上位ランクにいる調査協力者のほうが新 JLPT の

上位級に合格している割合が多く、ランクが下がるごとに新 JLPT の取得級が下がっていくことが明らかになった。また、自己評価の総得点とランクの関係においても、ランクが上がるごとに自己評価の総得点が高いことが分かった。

今後は項目参照プロファイルの結果やランクメンバーシッププロファイル等を詳細に分析することで、CEFR が設定したレベルから外れた CDS について分析を加えていきたい。

謝辞

本研究は JSPS 科研費基盤研究 (C) 26370619 「テキスト分析に基づく韓国人日本語学習者用読解尺度開発とその検証」の助成を受けたものです。

参考文献

- 印南洋 (2012). 「SEM 入門」竹内理・水本篤 (編著), 『外国語教育研究ハンドブック』. 東京: 松柏社
- 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」JLTA(Japan Language Testing Association)Journal, 13, 91-109.
- 小山由紀江・木村哲夫 (2011). 「Neural Test Theory を使った Can-do Statements の分析」統計数理研究所共同研究レポート, 254, 59-77.
- 清水裕士 (n.d.). 「Exametrika を用いたを用いた GHQ60 の潜在ランクの推定」 Retrieved from <http://norimune.net/material>
- 清水裕士 (2014). 「心理尺度への潜在ランク理論の応用 精神的健康調査票 (GHQ) を用いた順序的評価」岡山地域部会第 53 回研究会講演資料 2014 年 10 月 31 日 岡山理科大学 Retrieved from <http://norimune.net/2199>
- 清水裕士・大坊郁夫 (2015). 「潜在ランク理論による精神的健康調査票 (GHQ) の順序的評価」『心理学研究』85(5) Retrieved from https://www.jstage.jst.go.jp/article/jjpsy/advpub/0/advpub_85.13225/_pdf
- Shojima, K. (2007). Neural test theory. DNC Research Note, 07-02. Retrieved from <http://www.rd.dnc.ac.jp/~shojima/ntt/jpaper.htm>
- 荘島宏二郎 (2010). 「ニューラルテスト理論—学力を段階評価するための潜在ランク理論」植野真臣・荘島宏二郎 (編著) 『学習評価の新潮流』東京: 朝倉書店.
- 荘島宏二郎 (2014). Exametrika (Version 5.3) [Computer software]. Retrieved from <http://www.rd.dnc.ac.jp/~shojima/exmk/jindex.htm>
- 荘島宏二郎 (n.d.). 「潜在ランク理論 (ニューラルテスト理論)」 Retrieved from <http://www.rd.dnc.ac.jp/~shojima/ntt/jindex.htm>
- 横内裕一郎 (2014). 「能力記述文による自己評価と実際のスピーキング能力の関係—英検 Can-do リストと CEFR-J を使って—」『EIKEN BULLETIN』26, 218-230

-
- ⁱ 「特定の潜在ランクに属する受験者がテストで取ることが予想される期待得点」(横内、2014)
- ⁱⁱ 「特定の潜在ランクに何名の受験者が該当するかを示す潜在ランクの分布」(横内、2014)
- ⁱⁱⁱ 「受験者が各潜在ランクに当てはまる確率を示したもの。木村(2013)によると、『潜在ランクの推定値が同じだったとしても、RMPの違いによって、受験者に異なるフィードバックを返すことができる』ということで、同じ潜在ランクに属する複数の受験者のパフォーマンスを比較することも可能。」(横内、2014)
- ^{iv} 「CFIは.00から1.00の範囲で変動し、.95以上が好ましいと解釈されます」(印南、2012)
- ^v 「RMSEAは0.05以下で0.00に近いほどよい指標です」(印南、2012)