

## CEFR 読解指標 (CDS) に基づく作題例の 妥当性に関わる特性抽出の試み 2

谷 誠司, 宮崎 佳典, 安 志英, 元 裕璟

An Attempt to Extract Characteristics regarding the Validity of Example Problems Based on the CEFR Reading Comprehension Index (CDS) (2)

Seiji TANI, Yoshinori MIYAZAKI,  
Jiyoung AN, Yukyoung WON

### 抄 録

本稿は、谷・宮崎・安・元 (2022) で調査できなかった (DIALANG の) CEFR 読解 CDS (Can-Do Statements) の下位尺度 (「情報や要点を読み取る」、「説明を読む」、「読むこと全般」) を対象にして行った調査の結果を報告する。調査内容は、本研究チームで開発した「日本語 CEFR 読解指標 (CDS) 推定支援 Web アプリケーション」が提案する CDS との適合度上位 5 例文が、① CDS とどのくらい一致しているかと② CDS で記述されている「(読解) 行動」を問える問題にどのくらいできるか、の 2 点から日本語教育経験者 4 名による 5 段階評定と評定理由のコメントを分析した。

キーワード: CEFR (Common European Framework of Reference for Languages)、  
読解、CDS (Can-Do Statements)、例文適合度調査

### 1. はじめに

本研究チームでは (DIALANG の) CEFR 読解 CDS (Can-Do Statements) に対応した日本語例文 (約 1000 例文) を収集し、機械学習の手法を用いて入力例文に対応する CDS を推測・分類する研究ならびに実験を行ってきた。(宮崎・平川・高田・谷、2019) また、同分類手法を実装した「日本語 CEFR 読解指標 (CDS) 推定支援 Web アプリケーション」(以下、Web アプリケーション) の開発にも取り組んでいる。(宮崎・Vuong・谷・安・元、2020) (図 1)

「推定結果」のページは、推定結果の表示と、推定結果の解説とを併記する形式で表示されています。推定結果の表示は、推定結果の表示と、推定結果の解説とを併記する形式で表示されています。推定結果の表示は、推定結果の表示と、推定結果の解説とを併記する形式で表示されています。

| 予測CDS (推定アルゴリズム:SVM を使用) |         |  |       |         |        |
|--------------------------|---------|--|-------|---------|--------|
| CDS番号                    | CEFRレベル | CDS  | CDS番号 | CEFRレベル | CDS    |
| ①                        | B1      | 例文 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、レポートの目的と関連性をすばやく確認することができる。 | ⑩     | A1      |        |
| ②                        | B1      | 例文 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採集することができる。                           | ⑪     | B1      |        |
| ③                        | B1      | 例文 手紙、パンフレット、短文的な文書といった、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。                     | ⑫     | A2      | 例文 公共電 |
| ④                        | A2      | 例文 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指示を理解することができる。                                      | ⑬     | A1      |        |
| ⑤                        | A2      | 例文 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。                                | ⑭     | B2      | 例文 読む  |
| ⑥                        | A1      | 例文 もっとも一般的で日常的な状況でよく出くわす、簡単な指示に示しているような、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。         | ⑮     | B1      | 例文 自分  |
| ⑦                        | B1      | 例文 筆者が特別な立場や視点から取り上げた、現代の話題に関する記事やレポートを理解できる。                                      | ⑯     | A2      | 例文 自分  |
| ⑧                        | B1      | 例文 専門用語を確認するために辞書が使えらるのであれば、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。                             | ⑰     | A2      |        |
| ⑨                        | B1      | 例文 文章における議論の大き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。                                  | ⑱     | A2      | 例文 もっと |
| ⑩                        | B1      | 例文 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。   | ⑲     | A1      | 例文 身近な |
| ⑪                        | B1      | 例文 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。   | ⑳     | B2      | 例文 「推  |
|                          |         | 例文 手紙、パンフレット、新聞の短い事件記事のような簡易に書かれたテキストの文から特定の情報を                                    |       |         |        |

図1 Webアプリケーションの「推定結果」のページ (部分的に拡大表示)

さらに、谷・宮崎・安・元 (2020) では (DIALANG の) CEFR 読解 CDS からの日本語読解テストの開発を行った。具体的には (DIALANG の) CEFR 読解 CDS から 2 つの下位尺度 (「必要な情報を探し出す」と「手紙やメールを読む」) に属し、A1～B2 レベルの CDS (それぞれ 5 個と 6 個、計 11 個) を対象に問題を作成し、韓国語日本語学習者に受検してもらった。作問には Web アプリケーションから CDS との適合度が最も高い例文を使用し、韓国での日本語教育経験者に作問依頼をした。出題した問題分析の過程で作問者から、CDS と例文レベルの不一致や CDS で記述されている読み方ができない例文があったなどの事後コメントが得られた。

これを受け、谷・宮崎・安・元 (2022) では谷他 (2020) で使用した CDS を対象に、Web アプリケーションが提案する CDS との適合度 (以下「Web アプリケーションの適合度」) の上位 5 例文が、① CDS とどのくらい一致しているか (以下「一致度」) と② CDS で記述されている「(読解) 行動」を問える問題にどのくらいできるか (以下「作問可能度」) の 2 点から調査を行った。その結果、次の 4 点が明らかになった: 1) 「Web アプリケーションの適合度」と評定者の評定 (「一致度」と「作問可能度」) の相関関係は、「必要な情報を探し出す (30 例文)」では「一致度」とは  $r=.55$  ( $p<.01$ )、「作問可能度」とは  $r=.57$  ( $p<.05$ ) となり、やや相関があった。「手紙やメールを読む (25 例文)」では「一致度」とは弱い相関あり ( $r=.23$  (n.s.))、「作問可能度」とほとんど相関がなかった ( $r=.09$  (n.s.))。2) B2 レベルの CDS の例文は評定者による評定点が低い傾向にあった。3) 「Web アプリケーションの適合度」が同程度に高い例文であっても、評定者の評価が低かったり評定者間の評定が割れたりする例文があった。4) 同じ CDS の例文であっても例文ごとの文字数や例文の専門性に違いがあった。

本稿では、谷他 (2022) で報告できなかった (DIALANG の) CEFR 読解 CDS の下位尺度 (「情報や要点を読み取る」、「説明を読む」、「読むこと全般」) を対象にした調査結果を報告する。

## 2. 調査の概要

調査協力者は韓国における日本語教育経験者 4 名であり、調査は 2020 年 10 月～12 月と 2022 年 10 月～11 月に行った。調査方法は次のとおりである：1) 「情報や要点を読み取る」、「説明を読む」、「読むこと全般」に該当する A 1～B 2 レベルの各 CDS (それぞれ 7 個、3 個、6 個、計 16 個) を対象にした<sup>1</sup>。2) Web アプリケーションから CDS との適合度の上位 5 例文を選択した。対象の CDS が 16 個なので、16 個の CDS × 5 例文 = 80 例文を得た。3) 調査協力者にこの 80 例文を見せ、①「一致度」と②「作問可能度」の 2 点から 5 段階評定と評定理由のコメントを依頼した。

## 3. 調査の結果

### 3.1 全体的な調査結果

本節では、調査結果の全体的な傾向を報告する。まず本調査では 4 名の評定者がいるので、評定者間信頼性係数 (クロンバック  $\alpha$ ) を調べた<sup>2</sup>。「一致度」においては「情報や要点を読み取る (35 例文)」では  $\alpha = .67$ 、「説明を読む (15 例文)」では  $\alpha = .65$ 、「読むこと全般 (30 例文)」では  $\alpha = .58$  となった。「作問可能度」では「情報や要点を読み取る (35 例文)」では  $\alpha = .79$ 、「説明を読む (15 例文)」では  $\alpha = .15$ 、「読むこと全般 (30 例文)」では  $\alpha = .42$  となった。「説明を読む (15 例文)」と「読むこと全般 (35 例文)」の「作問可能度」は  $\alpha$  係数の数値が低い、クロンバック  $\alpha$  は「1- 各項目の分散の和 / 合計点の分散」の計算式であり、分母に来る合計点の分散の数値が小さいことが原因と考えられる。「説明を読む (15 例文)」と「読むこと全般 (35 例文)」の「作問可能度」については、評定者間信頼性係数が低いことを考慮しながら、分析を進めることにする。

次に調査対象の 80 例文は CDS の適合度が高い上位 5 つの例文であり、適合度を推定する値 (0%～100% の値を取り、Web アプリケーション上では「信頼度」と表示。) が算出されている。この「Web アプリケーションの適合度」と評定者 4 名の「一致度」の評定点の合計点 (5 点満点 × 4 名 = 20 点満点) との相関係数 (スピアマンの順位相関係数) を確認した。その結果、「情報や要点を読み取る (35 例文)」では  $r = .38$  ( $p < .05$ )、「説明を読む (15 例文)」では  $r = .13$  (n.s.)、「読むこと全般 (30 例文)」では  $r = -.10$  (n.s.) となった。「Web アプリケーションの適合度」と評定者 4 名の「作問可能度」の評定点の合計点 (5 点満点 × 4 名 = 20 点満点) との相関係数 (スピアマンの順位相関係数) は、「情報や要点を読み取る (35 例文)」では  $r = .18$  (n.s.)、「説明を読む (15 例文)」では  $r = .47$  ( $p < .05$ )、「読むこと全般 (30 例文)」では  $r = .25$  (n.s.) となった。「情報や要点を読み取る (35 例文)」の「一致度」と「説明を読む (15 例

<sup>1</sup> 2020 年 10 月～12 月には「情報や要点を読み取る」を、2022 年 10 月～11 月には「説明を読む」と「読むこと全般」を調査した。調査協力者の 4 名のうち 3 名は 2 回の調査で共通している。

<sup>2</sup> 平井 (2012) には「評定者が 2 名の場合は、相関係数やカッパ係数 (第 7 章) を用いる。評定者が 3 名以上の場合は、アルファ係数を使用する」とあるため、本研究ではアルファ係数 (クロンバック・アルファ) を用いた。

文)」の「作問可能度」だけが「Web アプリケーションの適合度」との間に中程度あるいは弱い相関があったが、それ以外は両者には相関がみられなかった。

また、CDS ごとの特徴を見るために、1 つの CDS に対応している 5 つの例文に関する 3 つの数値（「Web アプリケーションの適合度」、「一致度」の評定得点、「作問可能度」の評定得点）の平均値を出した（表 1 ～表 3）。

表 1 は「情報や要点を読み取る」にある 7 個の CDS についての結果である。平均値を基準にすると、9（B 1）と 13（A 1）はすべてにおいて平均値より数値が低い。7（B 2）と 12（A 2）は「Web アプリケーションの適合度」が低く、8（B 2）が評定者の「作問可能度」が低い。

表 1：「情報や要点を読み取る」にある CDS（7 個）の結果

| CDS 番号<br>(CEFR<br>レベル) | CDS  | Web アプリ<br>ケーション | 評定者（4 名）      |                 |
|-------------------------|--|------------------|---------------|-----------------|
|                         |  | 適合度の<br>平 均 値    | <一致度><br>の平均値 | <作問可能度><br>の平均値 |
| 7（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を理解できる。   | 40.62            | 16.20         | 16.40           |
| 8（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。                                     | 58.28            | 16.60         | 15.80           |
| 9（B 1）                  | 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。  | 35.36            | 14.40         | 15.40           |
| 10（B 1）                 | 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。  | 54.98            | 17.60         | 18.20           |
| 11（B 1）                 | 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。  | 71.48            | 18.00         | 18.20           |
| 12（A 2）                 | 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。                                | 39.39            | 16.60         | 17.80           |
| 13（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。 | 39.14            | 12.60         | 12.40           |
| 平均                      |  | 48.46            | 16.00         | 16.31           |

表 2 は「説明を読む」にある 3 個の CDS に対しての結果である。平均値を基準にすると、21（A 2）がすべてにおいて平均値より数値が低い。

表 2：「説明を読む」にある CDS（3 個）の結果

| CDS 番号<br>(CEFR<br>レベル) | CDS  | Web アプリ<br>ケーション | 評定者（4 名）           |                      |
|-------------------------|--|------------------|--------------------|----------------------|
|                         |  | 適合度の<br>平 均 値    | < 一致度 ><br>の 平 均 値 | < 作問可能度 ><br>の 平 均 値 |
| 19 (B 1)                | 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。              | 74.36            | 14.00              | 16.60                |
| 20 (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。 | 64.64            | 14.80              | 17.00                |
| 21 (A 2)                | 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。                  | 53.93            | 13.20              | 14.60                |
| 平均                      |  | 64.31            | 14.00              | 16.07                |

表 3 は「読むこと全般」にある 6 個の CDS についての結果である。平均値を基準にすると、26 (A 2) と 27 (A 1) がすべてにおいて平均値より数値が低い。22 (B 2) は「Web アプリケーションの適合度」が低く、24 (A 2) と 25 (A 2) は評定者の「一致度」が低い。

表 3：「読むこと全般」にある CDS（6 個）の結果

| CDS 番号<br>(CEFR<br>レベル) | CDS  | Web アプリ<br>ケーション | 評定者（4 名）           |                      |
|-------------------------|--|------------------|--------------------|----------------------|
|                         |  | 適合度の<br>平 均 値    | < 一致度 ><br>の 平 均 値 | < 作問可能度 ><br>の 平 均 値 |
| 22 (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。           | 30.59            | 15.00              | 15.60                |
| 23 (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を理解することができる。                         | 48.93            | 14.40              | 16.20                |
| 24 (A 2)                | 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。                                   | 42.38            | 11.60              | 16.00                |
| 25 (A 2)                | 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。                                  | 56.71            | 13.00              | 16.00                |
| 26 (A 2)                | 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。     | 39.50            | 11.80              | 14.60                |
| 27 (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。 | 33.44            | 12.80              | 13.60                |
| 平均                      |  | 41.93            | 13.10              | 15.33                |

### 3.2 個別的な調査結果

本節では、前節で評定者による「一致度」と「作問可能度」の数値が低い CDS を中心に個々の例文を見ていく。数値が低かった CDS は「情報や要点を読み取る」では 9（B 1）と 13（A 1）、「説明を読む」では 21（A 2）、「読むこと全般」では 26（A 2）と 27（A 1）であった。

図 2 から図 4 は、「情報や要点を読み取る（35 例文）」、「説明を読む（15 例文）」、「読むこと全般（30 例文）」の「一致度」と「作問可能度」の各評定得点（5 点満点）を平均値にし、散布図にしたものである。（紙面の関係上、縦軸及び横軸ともに 0 から始まっていない。）1 つの CDS に 5 例文あるので、例文は CDS 番号（CEFR レベル）例文識別記号（a～e）で表示されている。

図 2 を見ると、まず 9（B 1）の例文（実線の枠）において、例文 c と e が「一致度」と「作問可能度」の数値が低い。

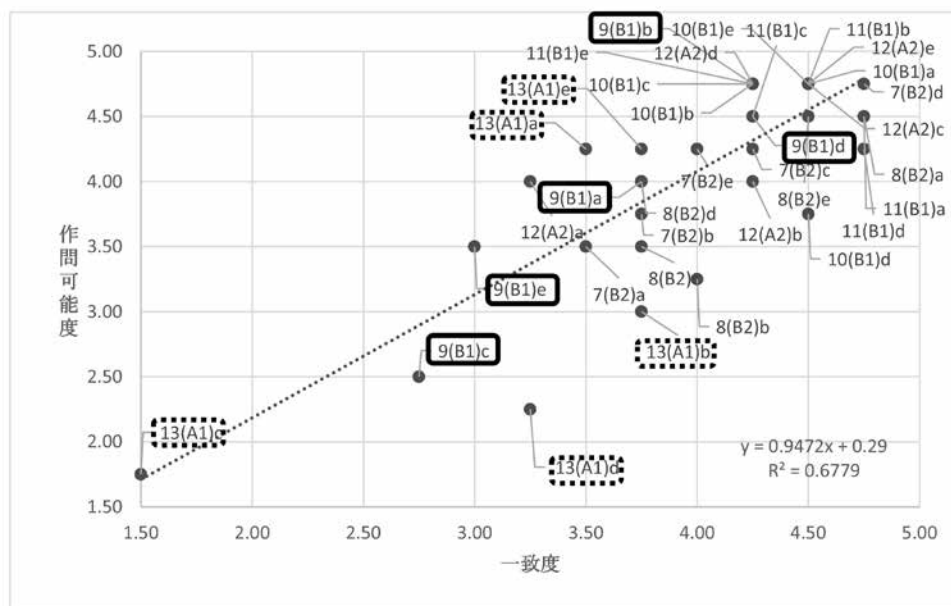


図 2：「情報や要点を読み取る」にある 35 例文の一致度と作問可能度の散布図

c の例文は「村上春樹論—コミュニケーションの物語」<sup>3</sup>という専門図書から取られている。そのため、判定者のコメントには、文章が文学的である点や内容の抽象度

<sup>3</sup> 林 正著、2002 年、専修大学出版局

が高い点、複文や名詞修飾が多い点、そして背景知識なしには B 1 レベルの日本語能力を問うには向かないことが指摘されている。例文 e は経済・金融情報の配信事業を行うブルームバークからのインターネット記事であり、アベノミクス下の経済政策についてエコノミストが評価をする記事である。例文の第 1 文から「量的質的緩和」「円相場」「下落」といった専門用語が出てきているため、評定者のコメントからも政治や経済に関する知識が不可欠であり B 1 レベルには難しい、情報や要点を読み取る問題や文中の表現を選ぶ問題であれば使えるが議論の大まかな流れや内容を問うような問題は難しい<sup>4</sup>との指摘があった。ただし、政治や経済の例文でも B 1 レベルにふさわしいと判断される例もある。例文 d は日本経済新聞からの記事であったが、外国人日本旅行者を地方にいかにつ誘致するかについての記事であり、評定者の「一致度」も「作問可能度」も 4 以上の平均点で高い評定を得ている。

同じく図 2 の 13(A 1) の例文(点線の枠)においては、例文 c と d が「一致度」と「作問可能度」の数値が低い。例文 c は鍋用の固形スープの素の広告で「うまさ溶けだす、〇〇〇〇(商品名)」というキャッチコピーだけが例文として提示されている。評定者のコメントでは、例文としての分量が少なすぎる点、「鍋」という漢字や「溶けだす」という複合動詞の難しさ、鍋や固形出汁についての背景知識の有無を問う問題になる点、情報量が少ないので問題が作りにくい点などが指摘された。例文 d はお茶販売会社のニュースレターからの例文であり、抹茶を使ったレシピ(抹茶アフォガード、抹茶スムージー、抹茶エスプレッソ)が写真と共に紹介されている。評定者のコメントでは、レシピで使われている写真が中心であり文章は写真の理解を助けるためのものである点、「抹茶」自体を既有知識であることを前提にしている点、読解というより絵を見て推測する問題になりそうな点などの指摘があった。

図 3 を見ると、21(A 2) では「作問可能度」の平均値は高いが、「一致度」は例文 a・b・e で低い(実線の枠)<sup>5</sup>。

<sup>4</sup> CDS9(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。」(下線部は筆者)とある。

<sup>5</sup> 「説明を読む(15 例文)」の「作問可能度」については、評定者間信頼性係数が低いこともあるので、分析においては「一致度」の部分を中心にみていく。

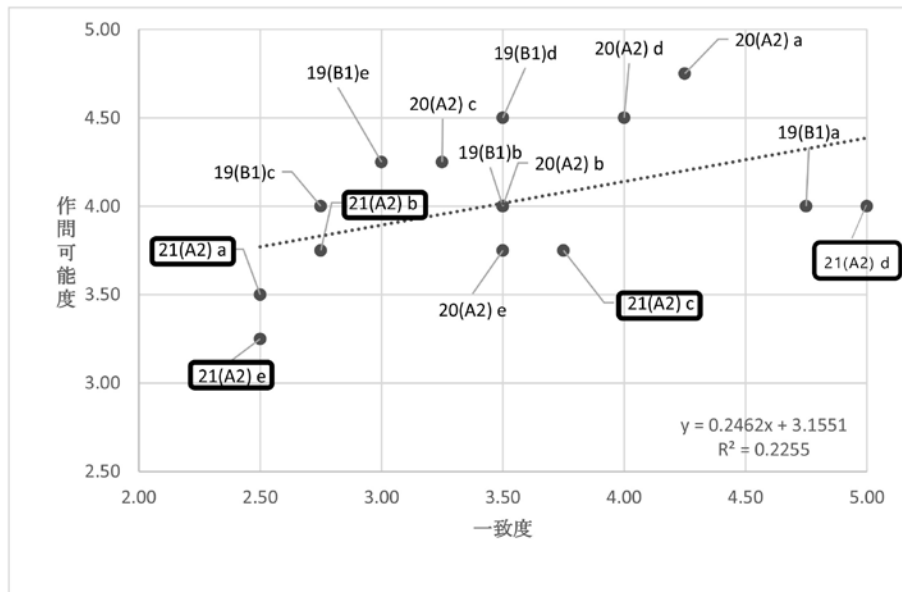


図 3：「説明を読む」にある 15 例文の一致度と作問可能性の散布図

表 4 が 21 (A 2) の a ～ e の例文である。「一致度」の低い例文 a ・ b ・ e に対する評定者のコメントには、語彙のレベルが A 2 にしては高い点、CDS 21 では「短く簡単」とあるが<sup>6</sup>文が短く簡単ではない点、画像があるほうがいい点が指摘された。

表 4 21 (A 2) の 5 つの例文 (a ～ e) <sup>7</sup>

| 例文識別記号<br>(a ～ e) | 例文  |
|-------------------|---|
| a                 | 1. 引き抜く 2. 向きを反対にさせる 3. カチッとなるまで強く押しこむ自動的にガスが噴出されます |
| b                 | 電源コードのプラグが異常に発熱する場合はプラグをぬき、電気店にご相談ください。             |
| c                 | 強く押す、まわす  |
| d                 | あたためスタート  |
| e                 | ご注意 カップを持ってお湯を注ぐことは、ドリッパーが倒れ火傷する恐れがありますのでおやめください。   |

<sup>6</sup> CDS 21 (A 2) は「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。」（下線部は筆者）とある。

<sup>7</sup> 例文の出典情報は次のとおりである。

a : Ban デオドラントパウダースプレーの使用説明文

b : フィリップス コーヒーメーカー取り扱い説明書 [http://www.philips.co.jp/shared/assets/jp/manuals\\_local/Discontinued\\_products/coffee\\_maker/HD7634.pdf](http://www.philips.co.jp/shared/assets/jp/manuals_local/Discontinued_products/coffee_maker/HD7634.pdf) (URL は 2023 年 1 月 11 日現在でもあるが、例文収集日である 2016 年 3 月 2 日当時と、掲載内容が異なる。)

c : <http://medical.mt-pharma.co.jp/di/file/dc/vlx.htm> (URL は 2023 年 1 月 11 日現在でもあるが、例文収集日である 2016 年 3 月 2 日当時と、掲載内容が異なる。)

d : 電子レンジ (引用情報不明)

e : Blendy ドリッパーコーヒーの注意事項



図 4 を見ると、まず 26 (A 2) の例文（実線の枠）において、例文 a と e が「一致度」の数値が低い。<sup>8</sup>

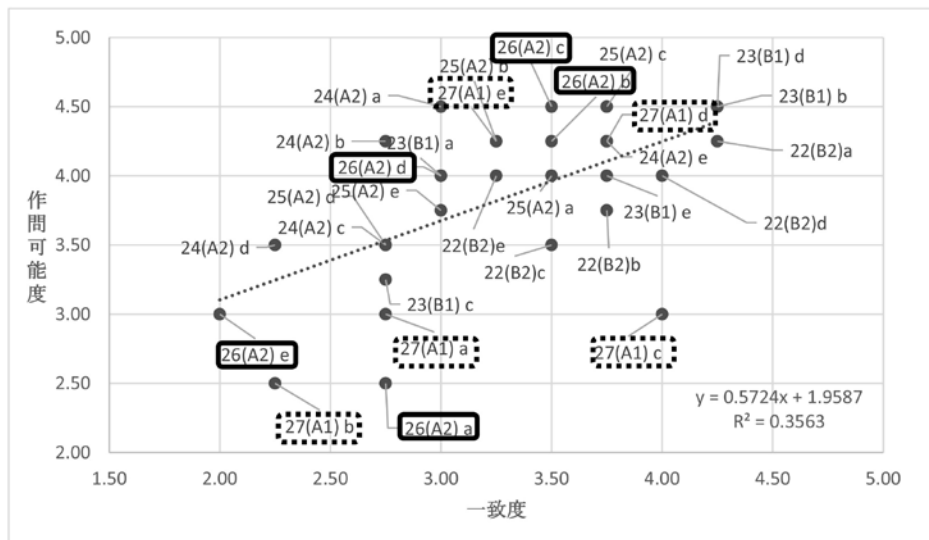


図 4：「読むこと全般」にある 30 例文の一致度と作問可能性の散布図

表 5 は 26 (A 2) a～e の 5 つの例文である。a の例文は、男性向けファッション雑誌・情報誌の記事の一部である。評定者のコメントには、例文中にある「タートルズ」がキーワードであるがこの単語自体が「世界的に共通して使われる単語」<sup>9</sup>でなく意味も不明である点、「タートルズ」がよくわからないため問題作成も難しい点の指摘があった。e の例文は言語教育学の書籍からの引用である。評定者のコメントには文章が短いわりには「従事」と「頻繁」といった A 2 レベルには難しい語が使われている点、文章が短すぎて作問が難しい点、テキストを要約した文章をいくつか提示しどれが合っているかを選択する問題が可能な点の指摘があった。

<sup>8</sup> 「読むこと全般（35 例文）」の「作問可能性」については、評定者間信頼性係数が低いこともあるので、分析においては「一致度」の部分を中心にいく。

<sup>9</sup> CDS 26 (A 2) は「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。」（下線部は筆者）とある。

表 5 26 (A 2) の 5 つの例文 (a ~ e) <sup>10</sup>

| 例文識別記号<br>(a ~ e) | 例文  |
|-------------------|---|
| a                 | ピザもスパゲッティも、家で作ればもっと楽くなる。まずは自由は発想で“タートルズ”からスタートだ！  |
| b                 | 聖バレンタインデーは毎年 2 月 14 日の祝日です。世界各国でお祝いされるものの、ほとんどの国が休日とはしていません。この日は家族も含め、自分の大切な人にカードや花やチョココレートなどを贈りますが、国によってさまざまな形式が見られます。 |
| c                 | 日本のクリスマスに欠かせないものといえば、ケーキとチキン。母親が作るのではなく、父親が仕事帰りに買ってくるのが一般的ですが、ホテルに豪華なケーキを注文することもあれば、手軽にコンビニで買ってくる姿も見られます。               |
| d                 | ウェルカムボードウェディングアイテムの中でも特にオリジナリティを出したいのがウェルカムボード。ナチュラル系やスイート系など、コンセプトに合わせたウェルカムボードでゲストをお迎えしましょう♪                          |
| e                 | 我々は、毎日のようにコミュニケーションに従事しているが、コミュニケーションとはどういうものかについて、そう頻繁に考えることはないようだ。  |

27 (A 1) の例文 (点線の枠) において、例文 a と b は「一致度」の数値が低い。a と b の例文はともにインターネット上からの引用で、それぞれ「あなたは知ってる? ○○○ (SNS のツール名) をにぎわす話題」と「今月もよく頑張りました! 自分へのご褒美 応援キャンペーン」<sup>11</sup> という例文である。例文 a に対する評価者のコメントには、「知ってる」や「にぎわす」「話題」といった語が A 1 レベルには難しい点、文章が短すぎて作問が難しい点が挙げられた。例文 b に対する評価者のコメントには、「頑張る」「ご褒美」「応援」といった語が A 1 レベルには難しく特に非漢字圏の学習者に難易度が高すぎる点、日本社会に関する背景知識がなければ「応援キャンペーン」が何であるか理解することが難しい点、キーワードとなる語彙が難しく絵などの補助があっても難しい点、文章が短すぎて作問が難しい点が挙げられた。

#### 4. 結語

本調査では、次の点が明らかになった。1) 「Web アプリケーションの適合度」と評定者 4 名による評定点の相関関係は、「情報や要点を読み取る (30 例文)」では「一致度」とやや相関あり、「説明を読む (15 例文)」では「作問可能度」と弱い相関があっ

<sup>10</sup> 例文の出典情報は次のとおりである。

a : POPEYE, 「LET'S DIY!」、マガジンハウス、2015 年 9 月号第 40 巻、p.92

b : アルク、<http://www.alc.co.jp/speaking/article/kihon/83.html> (リンク切れ)

c : アルク、<http://www.alc.co.jp/speaking/article/kihon/31.html> (リンク切れ)

d : インターネット記事、「100 均で賢く可愛く? ウェディングアイテムを DIY」<https://locari.jp/posts/18418> 電子レンジ (URL は 2023 年 1 月 11 日現在でもあるが、例文収集日である 2016 年 3 月 2 日当時と、掲載内容が異なる。)

e : 山内 進編著、2003 年、『言語教育学入門』、p.23 大修館書店

<sup>11</sup> 例文の出典情報は <http://event.rakuten.co.jp/campaign/point-up/reward/> (リンク切れ) である。

た。2) 谷他 (2022) では B 2 レベルの CDS の例文については評定者による評定点が低い傾向があったが、今回の調査では「情報や要点を読み取る」の 9 (B 1) と 13 (A 2)、「説明を読む」の 21 (A 2)、「読むこと全般」の 26 (A 2) と 27 (A 1) において評定者による評定点が低く、異なる結果が出た。3) B 1 レベル以下の例文の場合、背景知識や前提知識を要求する例文や専門用語が多い例文、想定レベルに合わない難しい単語が使われている例文に対して評定者による評定点が低い傾向があった。4) A 2 や A 1 レベルの場合、文章量が少ない例文もあるが、文章量が少ないと作問が難しいと判定されるケースもあった。

今後は、本調査結果を踏まえ、CEFR 読解 CDS からの問題作成支援の 1 つとして、CDS ごとのサンプル問題の提供へとつなげていきたい。

### 謝辞

調査に協力してくださった先生方に感謝申し上げます。本研究は JSPS 科研費基盤研究 (C) 18K00722 「CEFR 読解 CDS に基づいた問題仕様書の開発と妥当性の検証」の助成を受けたものです。

### 参考文献

- 谷誠司, 宮崎佳典, 安志英, 元裕璟. 2020. 「CEFR 読解指標に基づく日本語能力テスト開発の試み」, 『常葉大学大学院国際言語文化研究科研究紀要』創刊号 43—54.
- 谷誠司, 宮崎佳典, 安志英, 元裕璟. 2022. 「CEFR 読解指標 (CDS) に基づく作題例の妥当性に関わる特性抽出の試み」, 『常葉大学大学院国際言語文化研究科研究紀要』第 3 号 33—43.
- 宮崎佳典, 平川遼汰, 高田宏輝, 谷誠司. 2019. 「CEFR 読解指標に基づく日本語例文分類手法」, 當作靖彦 (監修), 李在鎬 (編) 『ICT × 日本語教育 — 情報通信技術を利用した日本語教育の理論と実践』54—67. 東京: ひつじ書房.
- 宮崎佳典, Vuong Hong Duc, 谷誠司, 安志英, 元裕璟. 2020. 「CEFR Companion Volume に対応した日本語例文自動分類手法」, 『日本學報』第 125 輯 153—175. 韓国日本学会.
- 平井明代 (編著). 2012. 『教育・心理系研究のためのデータ分析入門』東京: 東京書籍.

(2023 年 1 月 12 日受理)

