

日本語教育における Can-do statements 研究の現状と課題

島田 めぐみ

日本大学大学院総合社会情報研究科

The Review of Can-do Statements Study in Japanese Language Education

SHIMADA Megumi

Nihon University, Graduate School of Social and Cultural Studies

In the Common European Framework of Reference for Languages (CEFR), the different levels of each skill are indicated by the 'Can do' descriptors. The use of can-do statements increased as awareness of the CEFR broadened, and studies considering various aspects of such can-do statements are now being conducted. In this paper, I review the research on can-do statements in the field of Japanese language education from the viewpoints of scale development, use in large-scale assessments, development for Japanese for specific purposes, and use in educational institutions. While can-do statements scales are expected to be developed in various fields in the future, the need to proceed with various validations and develop and utilize environmentally-adapted can-do statements, as well as the demand for the promotion of continuous validation, have been raised as issues.

1.はじめに

近年、日本語教育分野では、Can-do statements に関し、様々な角度から研究が行われている。このように Can-do statements の認知度が上がり、多くの場面で取り上げられるようになったのは、CEFR (Common European Framework of Reference for Languages, ヨーロッパ言語共通参照枠) の影響が大きいと考えられる。CEFR では、各レベルの各技能が言語能力記述文 (descriptors) によって記述されており、この記述文が「～ができる」という Can-do の形式で示されているためである。Council of Europe (2018: 32) によると、レベル尺度に Can-do が最初に用いられたのは看護師のトレーニングの分野であり、欧州評議会において Can-do が言語教育分野で用いられるようになったのは 1970 年代だと言う。日本語教育の分野における Can-do statements に関する研究は、約 20 年後の 1990 年代に始まる。本稿では、日本語教育分野における Can-do statements に関わる研究を、大きく、尺度開発、大規模テストにおける

利用、目的別日本語のための開発、教育機関での利用の観点から概観した上で、今後の展望を述べる。ただし、必要に応じ、他言語の事例も取り上げていく。

なお、本稿では、Can-do statements 項目を難易度順にリスト化した Can-do statements 尺度を示す場合は「Cds」、Can-do 項目の文を示す場合は「記述文」の用語を用いることとする。

2.Cds の開発に関わる研究

2.1 尺度の開発

記述文のリストを尺度として利用するためには、適切な分析を行い開発する必要がある。初期の研究に、ALTE (Association of Language Testers in Europe) の Cds 開発がある。ALTE (2002) によると、ALTE では、「Social and Tourist」「Work」「Study」の領域を設定し、400 項目ほどからなる「ALTE Can do」を作成した。そして、1998 年、UCLES (University of Cambridge Local Examinations Syndicate) の英語テス

トの成績と Cds 自己評価の結果を関連づけるために、データを収集し分析を行っている。両者の間には明確な関連性があり、テスト得点の解釈を Can-do で示すことが可能となった。ALTE のように、Cds 開発に自己評価を用いているものに、TOEIC Can-Do や英検 Can-do がある。日本語教育分野においては、三枝 (2004) の日本語 Cds をはじめ、多くの Cds が主に日本語学習者の自己評価の結果に基づいて作成されている。

以下では、自己評価に基づいて開発された三枝らの日本語 Cds と、教師評価に基づいて開発された CEFR の例を見ていく。

日本語教育分野において、本格的な Cds 尺度の開発が最初に行われたのは、三枝令子氏を中心に行われたものである (三枝, 2004)。Cds 調査 (自己評定方式) が日本語能力の評定尺度として有効に機能するかを明らかにするために行われたものであるが、後述する日本語教育学会・国際交流基金 (1999) から継続して行われている。対象は、当時の日本語能力試験 1 級と 2 級レベル¹と設定された。記述文の作成にあたっては、TOEIC Can-do, ALTE Can Do, 大学の日本語教育のシラバス, 日本語学校などの授業活動内容や使用しているテキストの内容から行動記述を抽出し、その他先行研究や日本語教科書を参考にしている。最終的に 60 項目が確定された。日本語学習者に対し、自己評価調査を行い、信頼性と妥当性を検証している。信頼性は、いずれの調査でも .9 台であった (最終調査 (60 項目, 794 名) の場合: $\alpha = .98$)。因子分析の結果の検討、日本語能力試験得点との関係、教師評定の結果との関係、大学のプレースメントテストの結果との関係など様々な観点から妥当性検証が行われ、改定が繰り返された。

一方、CEFR は教師評価に基づいて開発されている。CEFR の Cds は、1993 年から 1997 年に実施されたスイス国立研究プロジェクト (Swiss National Research Project) の結果に基づいている (Council of Europe, 2018)。このプロジェクトでは、次の 3 段階

で開発・検証が行われた。Council of Europe (2018 : 40) によると、第 1 段階として、既存の言語能力尺度を分類し、新しい記述文を作成した。第 2 段階では、4 名から 12 名の言語教師を対象としたワークショップが 32 回開催され、言語教師が記述文について、各カテゴリーへの分類、明確性、正確性、適切性などについての評価、各能力段階への並べ替えを行った。第 3 段階では、250 名の言語教師がそれぞれ 10 名の学習者について記述文を用いて評定し、その記述文の難易度判定のデータをラッシュモデルにより分析した。ここでは要点をまとめるのみにとどめるが、各段階における具体的な方法は、野口・大隅 (2014)、投野 (2013) に詳細がまとめられている。

以上、尺度開発の手段として、自己評価を用いるものと教師評価を用いるものがあることを述べた。日本語教育においては、自己評価に基づいた研究が多いが、自己評価の問題点として、記述文に示される行動に関する経験の有無が影響すること (三枝, 2004 ; 島田, 2006 ; 根岸, 2006)、個人の社会・文化的背景、言語不安や性格要因が影響すること (玉岡ほか, 2005) などが指摘されている。一方で、教師評価の場合、教師が教室外の学生のパフォーマンスについて評価できるのか、という問題もある。根岸 (2008) では、日本で外国語として英語などを教授する場合は、「教師は現実の生活場面に基づく can-do アンケートには回答できないことが多い」(p.52) ことがわかったと言う。日本語教育においても、日本国外の場合は同様のことが言えるであろう。どちらの方法にもメリット、デメリットがあるが、いずれの方法でも、最後の段階では、何らかの統計手法により難易度が特定され、尺度化されるという点は共通している。

2.2 CEFR と日本語 Cds

Cds 開発に CEFR を利用する例は多い。そこで、日本語 Cds 尺度に関する研究を、CEFR との関連から概観する。

2010 年に国際交流基金により開発された JF 日本語教育スタンダードは、CEFR の考え方を基礎に、「相互理解のための日本語」を理念として開発された (国際交流基金, 2017 : 5)。JF 日本語教育スタン

¹日本語能力試験は、2010 年の改定前、現在の「N1, N2」という名称ではなく、「1 級, 2 級」という名称を用いていた。「1 級」がもっとも難しいレベルである。

ダードは、CEFR の Cds と国際交流基金が作成した JF Can-do の 2 種類を提供している。JF Can-do は、「CEFR Can-do を日本語の使用場面に合わせて具体化したもの」²と説明されているが、JF Can-do の構築には、CEFR の記述文を詳細に分析し、各レベル記述文の特徴を明らかにすることから始めている（塩澤ほか、2010）。また、「コミュニケーション言語能力の育成を目指した教育実践の支援ツール」（塩澤ほか、2010：24）として、記述文をデータベース化した「みんなの Can-do サイト」³が公開されている。

JF 日本語教育スタンダードは、CEFR の考え方に基づいて開発されているが、欧州で開発された CEFR を日本語に適用する際、尺度をそのまま用いることができるかという疑問から行われた研究もある。例えば、伊藤（2017）では、CEFR や JF 日本語教育スタンダードでは読字能力について具体的に示されていないことを指摘し、CEFR と JF 日本語教育スタンダードの言語構造的能力を参考に読字能力の Cds を作成している。記述文の例をあげると、A1 レベルでは「すでに知っている文字であれば、ほとんど間違えることなく、文字と音の対応で読むことができる」、C2 レベルは「一貫して正しい読みをし、自身の専門分野外の難読漢字などにも精通している」などである（p.61）。

野口ほか（2019、2020）は、CEFR の Cds の自己評価調査を日本語学習者に実施し、CEFR の記述文の順序性が日本語でも同様であるかを検討している。野口ほか（2019）は受容的能力、野口ほか（2020）は産出的能力とやりとりを取り上げている。その結果、「場面によって話し方が変わることを理解し、広範囲の慣用表現や口語表現がわかる」など、順序性が大きく異なる記述文があるものの、「聞く」「読む」に関しては、目標言語が日本語でも概ね適用可能であること（野口ほか、2019）、産出的言語能力尺度についても小さい修正で概ね適用できるが、「やりとり」尺度に関しては、文言の修正や項目の差し替えなど

が必要であることが指摘されている（野口ほか、2020）。

記述文の修正、差し替えなどが必要だとする研究とは異なる見解も存在する。2019 年度日本語学校研究大会（日本語教育振興協会主催）の「今、話題の CEFR A2 って何？」の分科会において講師をつとめた奥村三菜子氏は、「CEFR を利用するには、日本語や中国語の場合は文字の問題があるのではないか」という趣旨の質問に対し、興味深い回答をしている。日本語教育振興協会のウェブページに掲載されている該当部分を引用する。

例えば、「看板や掲示が理解できる」というのは、A1・A2 の言語活動として挙げられています。これを、「日本語や中国語では文字が難しいから B レベル・C レベルだ」とするのは、（略）CEFR の能力記述の趣旨とは異なります。「看板を見て理解する」という行動そのものは、B レベル（自立した言語使用者）や C レベル（熟達した言語使用者）の活動とは考えづらいからです。よって、「手書きパッドに文字を書き写して意味を調べれば、看板や掲示が理解できる」や「そばにいる人に読み方を教えてもらえるなら、看板の意味がわかる」といった条件付きの記述にすれば、A レベルの言語活動とすることが可能であろうと、私は考えています。⁴

奥村の主張は、CEFR を日本語教育で活用する際の解決策につながる可能性がある。今後は、記述文にこのような修正を施すことが有効なのか、実証的な研究が求められるだろう。

2.3 Cds の妥当性に関する研究

Cds 尺度の妥当性に関する研究を概観する。いずれの Cds 研究も、信頼性（ α 係数）は .9 以上の高い値が示されている。その一方、自己評価という形の Cds 尺度が真に学習者の日本語能力を反映するものになっているかという観点、すなわち妥当性については Cds 研究における解決すべき課題の 1 つとなっており、妥当性を検証する研究が多くなされてい

² <https://jfstandard.jp/publicdata/ja/render.do>（2020. 5. 16 閲覧）

³ <https://jfstandard.jp/cando/top/ja/render.do>（2020. 5. 22 閲覧）

⁴ <https://www.nisshinkyō.org/news/pdf/20190821y.pdf>（2020.5.18 閲覧）

る。その中でもっとも多いのが、外的側面からの証拠を求めるものである。外的側面からの証拠とは、他の変数との間に理論上想定される相関パターンが実際にも示され得るという証拠であり（平井，2006），ある評定尺度で高い得点を得た者は，同じ特性を測定している別の評定尺度でも高い得点を得ることになり，両者の相関は高くなるはずである。三枝（2004）では，日本語能力試験と Cds の相関を調べている。1 級では $r = .203$ （272 名），2 級では $r = .312$ （89 名）と，低い値である。一方，島田ほか（2006）では，初級から上級までの学習者を対象としたプレースメントテストと Cds（三枝（2004）の Cds）の間で計算し， $r = .804$ （91 名）という結果を報告し，三枝（2004）での値の低さは級別に計算したため，切断効果によるものであると分析されている。

玉岡ほか（2005）は，20 項目からなる Cds 自己調査と日本語能力テスト得点を 34 名の日本語学習者に実施したところ，両者の間には有意な相関は見られず，重回帰分析の結果，Cds の 4 技能のうちいずれの技能からも日本語能力を予測することができなかったと言う。この要因として，個人の社会・文化的背景，言語不安や性格要因などから，自己の日本語能力を過大あるいは過小評価することが考えられるとしている。このように相反する結果が報告されているが，記述文の特徴，対象者の特性，回答方法，データ数からも少なからず影響を受けることが考えられるため，これらの結果を単純に比較することはできない。

以上は客観テストとの関連を見たものである。Cds の多くは言語行動を表しているものであるため，上記のテストで扱われている「読み」「聞き」の能力評価のほか，「書き」や「読み」，また「やりとり」のパフォーマンスの結果と比較することが望まれる。比較するテストが知識を測定するものであったら，Cds とは測定対象にずれが生じていると言えるからである。

パフォーマンス評価の結果との関連を検討した研究は多くないが，口頭能力を扱ったものに，教育機関の日本語学習者を対象に実施した青木（2006）がある。三枝（2004）の Cds 調査票（60 項目）を用いて，76 名からデータを収集している。青木（2006）

では，Cds 調査とオーラルインタビューとの相関は $r = .63$ ，読解テストとは $r = .55$ ，聴解テストとは $r = .56$ であった。そのほか，村上（2009）もパフォーマンステストを取り上げている。村上（2009）は地域の生活者を対象としたもので，読み書き技能判定課題と口頭能力測定課題（インタビュータスク，ロールプレイタスク，絵を見て話すタスク）の結果と Cds 自己評価調査（30 項目）の結果の間の相関を 129 名のデータから分析している。「自分の名前をカタカナで書く」という Cds 項目と，実際に名前を書くという課題との間の相関は， $r = .69$ ，「日常見かける禁煙などの漢字の意味がわかる」という Cds 項目と「危険」の意味を母語で書くテスト項目の結果との間には， $r = .73$ と，Cds が示す行動とテストのパフォーマンスが一致する場合は，比較的高い相関係数が得られている。これらの結果を受け，「Cds が実際の運用を反映している可能性が高い」（p.31）と述べている。単なるパフォーマンステストとの比較ではなく，Cds 項目で提示される言語行動をテスト課題として課して検証している点で貴重な研究と言える。今後は，このような妥当性検証により Cds の記述内容の検討が求められる。

Cds 自己評価の結果に影響を与える要素を検討したものに，入江（2012），鹿島ほか（2011）などがある。入江（2012）は，記述文の要素が自己評価に影響を与える要素について「聞く」の Cds に注目して，学習者 185 名のデータから分析している。例えばドラマや映画などに関する Cds では「字幕がなくても」という記述を入れることにより，聴解テスト問題との相関が高くなると報告している。鹿島ほか（2011）では，学習者が Cds 自己評価調査に回答する際，どのようなことを想定したかをインタビュー調査により検討している。その結果，学習者によって想定する内容が異なる項目があった。例えば，「自分の意見や考えを日本人に十分説明することができますか」という項目について「具体的にどのような内容を想定したか」尋ねたところ，話題を「日常的なレベル」と考える者と，「宗教」や「社会問題」など重い問題と考える者がいた。後者には，なかなか自分の意見を伝えられなかったという経験を持つ者もあり，自己評価も低かったと言う。中津原（2013）は認知的

妥当性 (cognitive validity) を検証する手法に、どのように回答したかを問う口頭レポートをあげているが、鹿島ほか (2011) のインタビューはこれにあたる。認知的妥当性の観点から述べると、「自分の国と日本の違いを話すことができる」という項目は、項目作成者の意図した認知的プロセスを経ているとは言えない結果であったと言える。

また、日本語教育分野では珍しい DIF 分析も行われている (野口ほか, 2007)。DIF とは、テスト得点 (この場合、自己評価調査の得点) の能力水準が同じであっても、ある項目に「できる」と回答しやすい受験者集団 (例えば、ある母語集団) が存在する場合、DIF (differential item functioning, 特異項目機能) があると言う。野口ほか (2007) では、三枝 (2004) のデータから、中国語母語話者 (269 名) と韓国語母語話者 (383 名) を取り上げて分析し、中国語母語話者が「できる」と回答しやすい項目、韓国語母語話者が「できる」と回答しやすい項目がそれぞれ複数あったと言う。前者については、漢字と文化的な共通性が影響しており、後者は、言語構造、学校生活、相手との距離の取り方に関する共通性が影響している可能性があることを指摘している。

様々な角度から妥当性の検証が行われているが、全体的には外的側面からの証拠を求めている研究が多い。妥当性はあるひとつの検証で確認できるものではなく、様々な観点からの検証が必要である。2.1 で取り上げた CEFR の検証作業も参考となるであろう。

3. 大規模テストにおける利用

3.1 妥当性検証のための利用

日本語教育分野において、Cds の研究としてもっとも初期のものは、日本語教育学会・国際交流基金 (1999) の「Can-do-statements 調査報告—日本語能力試験の妥当性を検証する試み—」だと思われる。これは、日本語教育学会の試験分析委員会によって、日本語能力試験の妥当性を検証する目的で行われた研究である。この報告書によると、ETS (Educational Testing Service) がアメリカの大学生を対象に行った研究を参考にしたと言う。それは「College Students' Knowledge and Beliefs: A Survey of Global

Understanding」というプロジェクトで、Cds 自己評価の結果をテスト (FSI (Foreign Service Institute) のインタビューテスト, The MLA-Cooperative Foreign Language Test の聴解テストと読解テスト) の結果と比較したものである (Barrows, et al., 1981)。

日本語教育学会・国際交流基金 (1999) では、林四郎の分類による言語行動から、日本語非母語話者が遭遇すると考えられる行動を選んで Cds 項目 (書く, 読む, 聞く, 話す, 複合のスキル, 合計 42 項目) としている。Cds 自己評価調査は、各項目について「できる」から「できない」までの 5 段階評価で回答を求めるものである。日本語能力試験 1 級のモニター試験受験者 90 名を対象に実施し、受験者による自己評価得点のほか、日本語能力試験モニター版の結果と 1997 年度日本語能力試験の結果が得られている。これら 3 種類の得点の間の相関は、Cds 総合得点とモニター版の間は $r = .35$, Cds 総合得点と 1997 年度日本語能力試験の間は $r = .42$ であったと報告されている。この結果から、Cds 自己評価調査と日本語能力試験 1 級は中程度の相関を示し、Cds 自己評価調査の信頼性が高いこと (総合得点の α 係数が .964) から、Cds 自己評価調査と日本語能力試験とは共通の能力成分を共有していると分析されている。そのほか、Cds 自己評価の因子分析、項目ごとの分析などを行っている。肝心の日本語能力試験の妥当性検証については、「まとめと今後の課題」では触れられておらず、「はしがき」に「この調査の手法そのものが試験の妥当性を検証する手段になり得ることが示唆された」と述べるにとどまっている。

日本語教育学会の試験分析委員会が上記の研究を行ったきっかけは、アメリカの ETS の元テスト開発部長である Protase E. Woodford 氏との懇談会 (1997 年) にて、日本語能力試験の妥当性検証の方法として、Cds 調査を紹介されたことであると、同書の「はしがき」で述べられている。Woodford 氏は 1997 年の来日時に、当時ジェトロビジネス日本語能力テスト (現在の BJT ビジネス日本語能力テスト) を実施していた日本貿易振興会 (JETRO, 現在の日本貿易振興機構) のテスト開発部門をも訪れ、同テストの妥当性検証について同様に Cds 調査を紹介している。このことは記録には残っていないが、日本貿易振興

会にて同テストの開発に携わっていた筆者はその場において、ジェトロビジネス日本語能力テストでも受験者を対象に、テストの妥当性検証を目的に Cds 自己評価調査を行うきっかけとなったことを記憶している。2.3 で Cds 尺度の妥当性検証に、既存のテストを外的基準として相関を検討する研究を見たが、興味深いことに、初期の日本語教育における Cds 研究は、既存のテストの妥当性を検証するために Cds 自己調査を用いていたのである。

このように、1997 年より、テストの妥当性検証の方法の 1 手段として、2 つの大規模テストに関し Cds 調査が行われた。日本語教育学会・国際交流基金（1999）は、日本語教育分野において Cds 自己評価調査を利用した研究の嚆矢であり、Cds 尺度の開発を行った、先述の三枝（2004）へと引き継がれた。

3.2 測定結果の解釈のための利用

現在は、大規模テストの多くが得点解釈に Cds を利用するようになった。あるレベルや得点を得た受験者が実際にどのようなことができるのかということが示されるようになり、利用者からも理解しやすいものとなっている。外国語テストの中で TOEIC は早い段階から Cds リストを公表している。ETS は、1990 年代後半、TOEIC の得点から受験者が英語を使ってどのような活動ができるかを推測することの重要性、すなわち、得点を解釈する重要性を認識し、受験者の得点と Cds 調査の結果をもとに TOEIC の得点と受験者が遂行できる言語行動との対応づけを行い、2000 年に「TOEIC Can-Do Guide」を公開している（The Chauncey Group International, 2000）。その後も調査を継続し、TOEIC の実施団体一般財団法人国際ビジネスコミュニケーション協会の公式ページでは、2011 年のデータ（有効回答数 10,149 名）の分析結果を利用者向けに公開している⁵。そのほか、英語テストでは、TOEFL は「Performance Descriptors for the TOEFL iBT test」⁶、実用英語技能検定は「英検

⁵ https://www.iibc-global.org/toeic/special/target/list_readin.html (2020.3.15 閲覧)

⁶ <https://www.ets.org/s/toefl/pdf/pd-toefl-ibt.pdf> (2020.3.15 閲覧)

Can-do リスト」⁷をそれぞれ公開している。

日本語のテストでもっともはやく Cds リストを公開したのは、BJT ビジネス日本語能力テストの前身のジェトロビジネス日本語能力テストだと思われる。3.1 で述べた通り、最初はテストの妥当性検証の一環として Cds 自己評価調査を実施しているが、その後、2003 年と 2004 年の受験者を対象としたアンケート調査をもとに「JLRT Can do レポート」を作成し、公開している。このレポートの目的は、このテストの評価が実際にどのような日本語運用能力を反映したのかを明らかにするためであると当時の「JLRT Can do レポート」に記載されている⁸。現在は 2013 年から 2015 年の受験者のデータに基づいた CAN DO レポートが公開されている⁹。また、2018 年に「BJT ビジネス日本語能力テストに関する Can-do statements 調査研究事業報告書」¹⁰を公開している。この報告書によると、BJT 受験者の得点解釈の指標を示すために、葦原・小野塚(2014)および葦原(2014)が開発したビジネス日本語 Cds を用いていると言う。

日本語能力試験に関する Cds は、試験実施主体者が行ったものではないが、三枝（2004）の研究の一環として、日本語能力試験 1 級および 2 級合格者の遂行可能な言語行動を明らかにするという研究がなされている（島田ほか, 2006）。また、長沼ほか(2007)によると、2010 年の日本語能力試験の改定にあたって、CEFR を元に Cds の開発を行っている。長沼ほか（2007）ではその試行版を公表し、受験者に対する調査結果を報告している。その後、2010 年に日本語能力試験が改定されたのち、2010 年と 2011 年に調査が行われ、「日本語能力試験 Can-do 自己評価調

⁷ <https://www.eiken.or.jp/eiken/exam/cando/> (2020.3.15 閲覧)

⁸現在は、JETRO 公式ウェブページから該当ウェブページが削除されているため、島田（2010：268）の記載内容とプリントアウト資料を参考にした。「JLRT」は、多枝選択形式の「聴読解テスト」を指す。

⁹ <https://www.kanken.or.jp/bjt/about/evaluation.html> (2020.3.15 閲覧)

¹⁰ https://www.kanken.or.jp/bjt/survey_reports/data/survey_reports_2018.pdf (2020.5.15 閲覧)

査レポート」の最終報告を2012年に公開し¹¹、テスト利用者向けに「日本語能力試験 Can-do 自己評価リスト」を公開している¹²。

これらの大規模テストはいずれも、受験者のテストの得点と自己評価調査の結果の関連を分析し、各レベルの受験者が遂行できると考えている Cds を明らかにしている。

根岸 (2008) は、今後は「英語のテスト結果のフィードバックにより具体的な can-do statements が含まれるようになっていくだろう」(p.57) と述べている。日本語のテストでそのようなフィードバックを実現している大規模テストは管見の限りない。実現させるためには、単に受験者の自己評価調査の結果を分析するだけではなく、ある能力の受験者が実際どのような項目を遂行できるのか詳細な検討が必要となるであろう。

4. 目的別日本語のための Cds

東京外国語大学は、「JLPTUFS アカデミック日本語 Can-do」¹³を開発している。東京外国語大学の日本語教育プログラムでは8つのレベルが設定されているため、Cds は8つのレベルで構成されている。開発にあたっては、学習者による自己評価データ(学期開始期と終了期)と Cds との付き合い合わせにより、各記述文がレベルにあった段階的な難易度を示しているか検証している(鈴木・藤森, 2014)。公式ウェブページによると¹⁴、このCds活用の可能性として、(1)教育プログラムの策定・改善、(2)学習者の日本語学習の接続(アーティキュレーション)の参加情報の提供、(3)単位互換を含む大学間連携が挙げられている。ここからもわかるように、このCdsは東京外国語大学以外の教育機関においても使用できるもの

¹¹中 https://www.jlpt.jp/about/pdf/cds_final_report.pdf (2020.3.15 閲覧)

¹² <https://www.jlpt.jp/about/candolist.html> (2020.3.15 閲覧)

¹³ <http://www.tufs.ac.jp/common/jlc/kyoten/development/ajcan-do/gaiyo/gaiyo1-3.html> (2020.5.15 閲覧)

¹⁴ <http://www.tufs.ac.jp/common/jlc/kyoten/development/ajcan-do/gaiyo/gaiyo1-3.html> (2020.5.15 閲覧)

となっている。

外国人就労者のために開発されたものに2.3で取り上げた村上(2009)がある。これは、A市で就労する外国人を対象として開発されたものである。A市の日本語能力評価基準では、CEFRに基づいてなされているが、6レベルに加え0レベルを加え、7レベルとしている¹⁵。そして、判定対象は、0から2レベルとしている。項目の作成にあたっては、まず、生活の中で想定される日本語使用場面を抽出し、そのうち、調査の結果実際に遭遇する機会が多い項目を選定し、CEFRの項目と合わせて120項目のリストを作成している。その後、領域、技能別に整理し、代表性のある項目を選別して最終的に30項目からなるCdsを開発している。先述の通り、Cdsの妥当性検証のために、パフォーマンステストと各項目の相関が検討されている。

ビジネス日本語に関しては、葦原・小野塚(2014)がBJTビジネス日本語能力テストの測定対象能力を分析し項目を作成し、ビジネス日本語Cdsの開発を行っている。予備調査では、BJT模擬テストの結果とCds自己調査の結果の間は、 $r = .151$ と相関が低かったため、記述文の内容や表現について対象者にインタビューを行い、その結果に基づき、記述文の修正を行っている。最終的に、BJTで測定されているビジネス日本語能力を反映した尺度になったと言う。葦原・小野塚(2014)はBJTビジネス日本語能力テストが測定している能力が対象となっているが、葦原ほか(2017)では、さらに汎用性のある高度外国人材のための日本語フレームワーク構築のために、Cds尺度の開発を試みている。Cds尺度の開発プロセスは、「直観的手法」「質的調査法」「量的調査法」があるが(投野2013)、葦原ほか(2017)では「直観的手法」の結果、すなわち、先行研究からのCds収集、Cds執筆、Cdsの整備などが報告されている。なお、「直観的手法」とは、2.1で紹介したCEFRの開発・検証の「第1段階」に相当する。

介護分野では、「介護の日本語 Can-do ステートメント(KCDS)」がある(西郡, 2019)。JF日本語教

¹⁵ 現在のCEFR補遺版では大きく7レベルが設定されているが、当時のCEFRは6レベルであった。

育スタンダードの A2 と B1 のリストを参照し、介護福祉施設における職員と候補者へのヒヤリング、介護の日本語教育経験、既存の資料、日本語教育法の記述や理念を参考に 97 項目からなる Cds を作成している。74 名を対象に調査を実施し、これらの項目で示された言語行動がどの程度できるのかを調査している。この KCDS には、例えば「困っていたり、体調が悪そうな利用者に、「大丈夫ですか」「一緒に〇〇しましょうか」など、短い簡単な言葉で声かけができる」「利用者のお話の中に「厠（かわや）」「えもんかけ」などの高齢者特有の古い言葉や、「あずましい（東北地方の方言で「気持ちがいい）」など方言を使った短い表現などが入っていても聞いて理解できる」など、介護分野ならではの特徴的な記述文が含まれている。

2019 年より「特定技能」での新たな外国人材の受け入れが開始したが、今後、各分野における Cds の開発も求められていくであろう。

5. 教育機関における利用

5.1 日本国内の教育機関

現在は、独自の Cds を開発する教育機関も多いが、初期に開発されたものに島田ほか（2007）がある。この Cds は、東京学芸大学の日本語学習者を対象とし、三枝（2004）で開発された日本語 Cds を元にして作成している。三枝（2004）の Cds は、対象者の目標使用言語の領域を「大学生の生活」としており、日本語能力試験 1 級と 2 級レベル用として開発されているが、島田ほか（2007）の対象者は日本語能力の低いレベルも含まれているため、23 項目を追加し、最終的に 81 項目からなる Cds を開発している。東京学芸大学の受講生を対象に調査を実施し、その結果から、日本語科目群の各レベルにおける言語行動目標を設定している。これらの試みにより、海外協定校の日本語教員や留学希望者に対して具体的な情報を提供できる、東京学芸大学の日本語教員に対し有益な情報を提供できる、受講者に対し明確な目標レベルを提示できるなどの利点をあげている。さらに、77 項目に改定された Cds の自己評価調査を海外協定校 5 校でも実施し、東京学芸大学の各レベルと各協定校各学年の対応表を作成した（島田ほか、

2009）。その結果、単位互換およびその他の成績決定に貴重な情報を提供できるとしている。この Cds の妥当性については、客観テストとの関連性を見るなどが行われている（島田、2010）。

島田ほか（2007）は、日本語科目群の各レベルの目標基準作りに Cds を利用したが、同様の試みは、日本大学の短期交換プログラムにおいても行われ、日本語クラスのレベル設定を Cds にて明示化している（保坂、2009）。保坂（2009）では、123 項目の Cds 調査票を作成、5 レベルにおける Cds 一覧を作成している。

言語能力の評価を目的とした場合、一度使用したテストを同じ受験者に再度使用することは適切ではないが、Cds 自己評価は、同じ受験者に対し異なる時期に同一の調査票を用いることができる。この利点をいかして、コース前後の結果を比較する研究もある。先述の青木（2006）は、三枝（2004）の Cds を利用し、入学時と 10 ヶ月後の卒業時に自己評価調査を行い、平均値を比較し、卒業時の方が高いことを確認している。岡山大学でも、日本語コースの受講者が、学期開始後から終了前の約 3 ヶ月の期間で、自己の日本語能力の変化をどのように感じているか CEFR のスイス版自己評価チェックリストを用いて検証している（坂野・大久保、2012）。その結果、各レベルの全ての領域で、自己評価の平均値が有意に高くなり、全項目において平均値が高くなっていたことを報告している。坂野・大久保（2012）は、Cds を利用することにより、試験では測定できない言語能力が見えてくるという可能性を示唆している。また、日本語コースの検証に Cds が利用される場合もある。明治大学では、日本語コースの受講生が自身の日本語能力の伸びをどのように認識しているかを探るために、コースの前後に Cds 自己評価調査を実施し、両者を比較している（柳澤、2019）。その結果、日本語教育の質の保証に向けた検証の方法の一つに Cds 調査が活用できると述べられている。このように、コースの前後に同一の Cds 自己評価調査を実施し、その結果を比較することにより、学習者の能力の変化を検討すること、また、コースを検証することが可能となる。

5.2 日本国外の教育機関

言語教育分野において開発されている Cds の多くは教室外での言語行動が中心となっている。しかし、国外の日本語教育機関の場合、学習者は教室外で日本語に接する機会は多くなく、そのような記述文が示す行動に関し経験がない学習者も多い。この状況は、日本における英語教育と共通する。日本における英語教育分野では、長沼・宮嶋 (2006) が、外部指標としての一般的な Cds (CEFR や TOEIC Can-Do など) だけではなく、個別の学習状況に合わせた内部指標としての Cds を開発することが重要だとし、清泉女子大学の英語の授業の学習活動を項目化し、学習状況に合わせた大学内部指標としての「清泉アカデミック Can-Do 尺度」を開発している。記述文は、生活場面や職場場面での言語行動ではなく、「授業で練習した表現を使って、日常的な話題に関してペアで会話 (ロールプレイ) ができる」(p.58)、「英語で書いた原稿を見て、自分で文法的な誤りを直すことができる」(p.61) など、大学の英語の授業で目的としていきたい学習行動を取り上げている。

長沼・宮嶋 (2006) の記述文は、島田ほか (2007) の東京学芸大学の日本語の Cds とはまったく異なる視点から作られている。また、汎用性のある「JLPTUFS アカデミック日本語 Can-do リスト」とも異なり、清泉女子大学の具体的な授業での活動に特化した内容となっている。日本の大学生の中には、教室外ではまったく英語を使用しないという学生が存在することは十分考えられ、長沼・宮嶋 (2006) の取り組みは実態に即したものであると言える。また、このような Cds 尺度により、具体的な目標を教師も学生も認識することができる。

日本語教育分野の Cds 研究では、長沼・宮嶋 (2006) のような具体的な授業活動に即した Cds 開発研究は、今まで存在しないのではないかと考えられるが、その中、島田・董 (2019) では、学習者が教室内外で実際に行う言語行動をもとに、中国 A 大学 Cds フレームワークを開発することとし、まず、「聞く」「読む」に関する A 大学 Cds を作成している。さらに、島田ほか (2020) では、内部指標である A 大学 Cds で示された各言語行動が外部指標である CEFR のいずれのレベルに相当するか検討している。具体的に

は、A 大学 Cds 自己評価のデータと、同時に実施した CEFR の Cds の自己評価データの結果をあわせて IRT の段階応答モデルを適用し、分析している。

今後さらに教育現場において Cds を利用する機会が増えていくと考えられるが、各教育現場の状況に応じた Cds 研究も必要となるだろう。

6.おわりに

以上、日本語に関する Cds 研究に関し、尺度開発、大規模テストにおける利用、目的別日本語 Cds の開発例、教育機関における事例を概観した。本稿では比較的大規模に開発された Cds を中心に取り上げたため、教室場面や教材などでの Cds 利用などの研究は取り上げることができなかったが、実際は、様々な場面で活用が進められている。

今後、Cds の利用はさらに進むと考えられ、様々な分野における Cds 尺度も開発されていくことが予想できる。しかし、項目の特徴、回答者の特性などが結果とどう関わるかなど、まだ明らかになっていないことも多く、様々な妥当性検証を進める必要があるだろう。

教育現場においても、今後各教育機関で Cds が開発され、レベルの記述、単位互換の根拠、日本語コースの検証など多くの場面で使用が進むと考えられる。すでに日本国内の機関では、多くの研究が進み、知見が蓄積されてきている。一方、国外に目を向けると、学生が Can-do の自己評価に真面目に取り組まないという声を聞くことがある。ひとつの要因として、記述文の内容が日本国内での言語使用を想定しているということが考えられる。今後は、環境に合わせた Cds 開発、利用を進めていく必要もあるだろう。

長い期間かけて開発された CEFR も継続的に検証研究がなされて、2018 年に新たに、CEFR 補遺版 (CEFR Companion Volume with New Descriptors)¹⁶が公開された。日本語教育分野においても、継続的な検証を進めていくことが求められる。

¹⁶

<https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989> (2020.5.15 閲覧)

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—プロフィシェシー・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, 18-44
- 葦原恭子(2014)「外国人のビジネス日本語能力の評価に関する基礎研究—評価システムの構築をめざして—」(平成 26 年科学研究費助成事業 研究成果報告書)
<https://kaken.nii.ac.jp/ja/file/KAKENHI-PROJECT-2520535/22520535seika.pdf> (2020.5.16 閲覧)
- 葦原恭子・奥山貴之・塩谷由美子・島田めぐみ (2017) 「高度外国人材に求められるビジネス日本語フレームワークの構築—直観的手法を中心に—」『琉球大学国際教育センター紀要』1, 1-14
- 葦原恭子・小野塚若菜 (2014) 「高度外国人材のビジネス日本語能力を評価するシステムとしてのビジネス日本語 Can-do statements の開発—BJT ビジネス日本語能力テストの測定対象能力に基づいて—」『日本語教育』157, 1-16
https://doi.org/10.20721/nihongokyoiku.157.0_1
- 伊藤秀明 (2017) 「拡張・精緻化のための読字能力の能力記述文試案作成—CEFR/JFS の言語構造的能力を参考に—」『日本語教育』168, 55-62
https://doi.org/10.20721/nihongokyoiku.168.0_55
- 入江友理 (2012) 「Can-do statements の記述に含まれる要素が自己評価に及ぼす影響—「聞く」技能に焦点を当てて—」『言語教育評価研究』2, 15-25
- 鹿嶋彰・保坂敏子・島田めぐみ (2011) 「Can-do statements 項目から回答者は実際何を想定するか」『日本語教育方法研究会誌』18(2), 32-33
https://doi.org/10.19022/jlem.18.2_32
- 国際交流基金 (2017) 『JF 日本語教育スタンダード【新版】利用者のためのガイドブック』
https://jfstandard.jp/pdf/web_whole.pdf (2020.5.8 閲覧)
- 三枝令子 (2004) 『日本語 Can-do-statements 尺度の開発 平成 13～15 年度科学研究費補助金 (基盤研究 B1) 研究成果報告書』
- 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, 23-39
- 島田めぐみ (2010) 「自己評価 Can-do statements に関する—考察—客観テストとの比較を通して—」『東京学芸大学紀要総合教育科学系』61, 267-277
- 島田めぐみ・大隅敦子・熊谷龍一・董博・野口裕之 (2020) 「国外日本語教育機関における Can-do statements と CEFR 能力記述文の間の項目困難度比較—受容技能を例に—」『2020 年度日本語教育学会春季大会予稿集』, 294-299
- 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, 75-88
[info:doi/10.20649/00000370](https://doi.org/10.20649/00000370)
- 島田めぐみ・董博 (2019) 「中国の日本語教育機関における Can-do statements フレームワーク (読む, 聞く) の構築」『東アジア日本語教育日本文化研究学会 2019 年度国際学術発表大会発表要旨文』, p.41
- 島田めぐみ・野口裕之・谷部弘子・斎藤純男 (2009) 「Can-do statements を利用した教育機関相互の日本語科目の対応づけ」『日本語教育』141, 90-100
https://doi.org/10.20721/nihongokyoiku.141.0_90
- 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—」『東京学芸大学紀要人文社会科学系 I』58, 495-504
- 鈴木美加・藤森弘子 (2014) 「A Can-do リスト開発プロセスにおける学習者の自己評価とその分析」『東京外国語大学留学生日本語教育センター論集』40, 53-68
- 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」『留学生教育』9, 65-78
- 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
- 中津原文代 (2013) 「能力記述としての Can-do statements とテストの妥当性を検証する「社会

- 的・認知的枠組み」(Socio-cognitive Framework) について」『言語教育評価研究』3, 44-53
- 長沼君主・大隅敦子・和田晃子・伊東祐郎・熊谷龍一・野口裕之 (2007) 「JLPT 日本語能力記述文作成の試み—日本語能力試験 (JLPT) Can-Do Statements 試行版の分析から—」『2007 年度日本語教育学会秋季大会予稿集』, 215-218
- 長沼君主・宮嶋万里子 (2006) 「清泉アカデミック Can-Do フレームワーク構築の試みとその課題と展望」『清泉女子大学紀要』54, 43-61
- 西郡仁郎 (2019) 「介護福祉の日本語教育の現状と支援者の育成—介護の日本語 Can-do ステートメントを中心に—」『日本語教育』172, 18-32
- 日本語教育学会・国際交流基金 (1999) 『Can-do-statements 調査報告—日本語能力試験の妥当性を検証する試み—』
- 根岸雅史 (2006) 「GTEC for STUDENTS Can-do Statements の妥当性検証研究概観」『ARCLE REVIEW』1, 96-103
- 根岸雅史 (2008) 「英語教育における最近の評価動向」『日本語教育』136, 49-58
- 野口裕之・大隅敦子 (2014) 『テストニングの基礎理論』研究社
- 野口裕之・大隅敦子・熊谷龍一・島田めぐみ (2019) 「CEFR の日本語への適用可能性—受容的能力の場合—」『日本テスト学会第 17 回大会発表論文抄録集』, 116-117
- 野口裕之・大隅敦子・熊谷龍一・島田めぐみ (2020) 「CEFR の日本語への適用可能性—産出的言語活動および (言葉の) やりとりの場合—」『2020 年度日本語教育学会春季大会予稿集』, 173-178
- 野口裕之・熊谷龍一・脇田貴文・和田晶子 (2007) 「日本語 Can-do-statements における DIF 項目の検出」『日本言語テスト学会研究紀要』10, 106-118 https://doi.org/10.20622/jltaj.10.0_106
- 坂野永理・大久保理恵 (2012) 「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」『大学教育研究紀要』8, 179-190
- 平井洋子 (2006) 「測定の妥当性からみた尺度構成—得点の解釈を保證できますか—」吉田寿夫 (編) 『心理学研究法の新しいかたち』, 21-49, 誠信書
- 房
- 保坂敏子 (2009) 「短期交換プログラムにおける日本語クラスのレベル設定の試み—Can-do-statements を利用して—」『日本語教育方法研究会誌』16(1), 2-3 https://doi.org/10.19022/jlem.16.1_2
- 村上京子 (2009) 「外国人就労者のための日本語 ‘Can Do’ statements の開発—パフォーマンス・テストによる妥当性の検討—」『言語教育評価研究』1, 21-33
- 柳澤絵美 (2019) 「Can-do statements を用いた学期開始時と終了時における日本語能力の自己評価の比較—ET 日本語における教育の質の保証に向けて—」『明治大学国際日本学研究』11(1), 123-150
- Association of Language Testers in Europe (2002) *The ALTE can do project*. <https://www.cambridgeenglish.org/images/28906-alte-can-do-document.pdf> (2020.3.15 閲覧)
- Barrows, T. S. and others (1981) *College Students' Knowledge and Beliefs: A Survey of Global Understanding. The Final Report of the Global Understanding Project*. Educational Testing Service.
- Council of Europe (2018) *Common European framework of reference for languages: Learning, teaching assessment. Companion Volume with new descriptors*. <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989> (2020.5.16 閲覧)
- The Chauncey Group International (2000) *TOEIC Can-Do Guide*. https://www.ets.org/Media/Research/pdf/TOEIC_CAN_DO.pdf (2020.5.11 閲覧)

付記

本研究は JSPS 科学研究費補助金 17H02351 (代表: 野口裕之) の助成を受けて行われたものである。

(Received: June 18, 2020)

(Issued in internet Edition: July 1, 2020)