資 料 4

令和４年度「生活 Can do（案）の量的検証の分析（概要）」

１．分析の目的

本調査では、一般社団法人 日本語教育支援協会（2021）2021.7.15 日本語教育小委員会「令和２年度文化庁委嘱「日本語教育の参照枠」Can do の量的検証に関する調査報告書」（以下、【文献１】とする）の手順に基づき、「生活 Can do」（案）のレベル尺度の妥当性について量的検証を行った。

２．評定尺度の妥当性

今回の分析は、【文献１】と同様、ラッシュ系モデルの１つである評定尺度モデル（rating scale model）を用いて行い、それぞれの Can do の困難度をロジット（log odds units：logit）尺度の値で表わしている。各 Can do のロジット値はプラスに大きくなるほど困難度が高くなることを表す。評定尺度モデルを適用するにあたり、モデルの適合度（モデル選定の妥当性を示す指標）、一次元性（Can do 項目が全体として等質性をもつこと）などを確認しておく必要がある。今回の分析データについて、これらの指標が妥当であることを事前確認した上で分析を行っている。また、今回の分析は【文献１】と同様な手続きを踏んでいるものの、「【文献１】における各 Can do のロジット値と今回の分析における各 Can do のロジット値は、等化（同じ尺度に調整すること）を行っておらず、直接比較することはできない」という点については注意が必要である。

レベルごとの Can do 項目のロジット値

Can do 項目のレベルごとに、Can do 項目のロジット値の記述統計量を表１に示す。

表１ レベル別 Can do 項目のロジット値の記述統計量

レベル	平均値	標準偏差	中央値	最小値	最大値	レンジ	標準誤差
A1	-1.39	0.65	-1.30	-2.96	-0.15	2.82	0.11
A2	-0.38	0.65	-0.42	-1.79	0.67	2.46	0.08
B1	0.32	0.48	0.39	-0.79	1.18	1.97	0.08
B2	1.04	0.45	1.08	0.18	1.80	1.62	0.10

さらに言語活動ごと、レベルごとの中央値は以下の表２のようになっている。なお、平均値ではなく、中央値を用いたのは、【文献１】と同様、ロジット値が大きく外れた項目による影響を少なくするためである。

表２　言語活動別・レベル別中央値

レベル	聞くこと	読むこと	やり取り	発表	書くこと
A1	-0.88	-1.30	-1.73	-1.24	-1.25
A2	-0.45	-0.64	-0.21	-0.29	-0.02
B1	0.28	-0.20	0.48	0.55	0.56
B2	0.87	0.58	0.94	1.64	1.31

図１は、レベルごとのロジット値の平均値をグラフで示したものである。図の横軸は意図した Can do 項目のレベル（A1～B2）を示し、各点がそれぞれの Can do 項目に対応しており、図の縦軸の値がロジット値の平均値を示している。それぞれの言語活動について、A1 から B2 へとレベルが上がるごとに Can do 項目のロジット値が高くなっている（難しくなっている）ことが確認できた。また、A1 レベルの Can do 項目以外については、【文献１】と同様に、「聞くこと」、「読むこと」という受容系言語活動の方が「発表」、「書く」の産出系言語活動よりもロジット値が低い（易しい）という傾向がみられた。一方、【文献１】でみられた「やり取り」で A1 レベルの Can do 項目のロジット値が高い（難しい）という状況はみられず、むしろ「やり取り」の A1 レベルの Can do 項目では、他の言語活動の A1 レベルの Can do 項目よりもロジット値が低い（易しい）という結果になった。

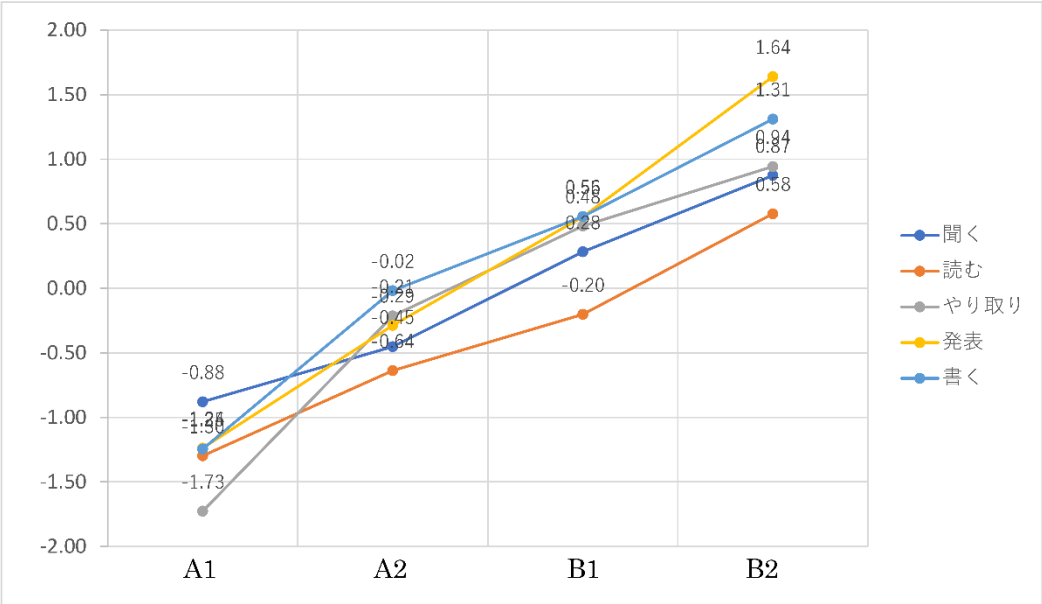

図１　レベル別・言語活動別のロジット値の中央値

３．レベルの逆転項目

（１）レベルの逆転項目の認定と概略

「隣接するレベルのロジット値の中央値を上回る（または下回る）ロジット値をもつ項目」をレベルの逆転項目（以下、逆転項目）と定義し、そのような項目を洗い出し、その原因を考察する。なお、一般に、評定尺度の質問紙調査で質問項目について「ポジティブな表現に対するネガティブな表現」を「逆転項目」と呼ぶが、本報告における「逆転項目」はそれとは異なり、【文献１】と同様に独自の選定方法と定義による名称である。

逆転項目の選定方法は以下の通りである。まず、各レベルのロジット値の中央値を求めた（表２）。前述のように、隣接するレベルのロジット値の中央値を上回る（または下回る）ロジット値をもつ項目を逆転項目として認定した。たとえば、「聞くこと」の A1 の中央値は－0.88 であるから、A2 項目でありながら－0.88 よりも低いロジット値をもつ項目は逆転項目として認定する。同様に A1 の項目でありながら、A2 の中央値である－0.45 よりも高いロジット値をもつ項目も逆転項目となる。以上の方法で逆転項目を認定したところ、全体で 13 項目が該当した。そのうち、隣接するレベルよりもさらに離れたレベルでの（つまり２レベルにわたる）逆転項目はなかった。レベル別、言語活動別の内訳は以下の表３の通りである。逆転項目の項目数の割合で見ると、「発表」、「読むこと」が多く、「聞くこと」、「書くこと」が少ない。

表３　言語活動別・レベル逆転項目数

	A1		A2		B1		B2				
言語活動	調査項目数	逆転項目数	調査項目数	逆転項目数	調査項目数	逆転項目数	調査項目数	逆転項目数	調査項目合計	逆転項目合計	逆転項目割合%
聞くこと	3	0	3	1	6	0	3	0	15	1	6.7
読むこと	4	0	5	2	6	2	4	0	19	4	21.1
やり取り	10	1	10	0	12	2	7	1	39	4	10.3
発表	3	0	3	2	4	1	4	0	14	3	21.4
書くこと	4	0	4	1	4	0	3	0	15	1	6.7
計	24	1	25	6	32	5	21	1	102	13	12.7

（２）逆転項目と経験率

今回の調査では、それぞれの Can do について経験の有無に関する調査も行った。表４は、今回抽出した逆転項目について、Can do ごとに全回答者数に対する「経験がある」と回答した学習者

の人数の割合（%）（以下、経験率とする）をロジット値と合わせて示したものである。

逆転項目全体をみると、経験率が低い項目は想定したレベルより上位のレベルの中央値よりもロジット値が高く、難しいとされる傾向にあり、経験率が高い Can do 項目は想定したレベルより下位のレベルの中央値よりもロジット値が低く、易しいとされる傾向にある。たとえば、Q210142 と Q210182 はともに B1 レベルと想定されているが、Q210142 は経験率が低く、ロジット値は B2 レベルの中央値（0.58）よりも高くなっており、Q210182 は経験率が高く、ロジット値は A2 レベルの中央値（−0.64）よりも低くなっている。

このことから、逆転項目について、想定したレベルよりも難しいと回答している Can do については、経験がないがゆえに、実際よりも難しいと自己判断している可能性が考えられる。

表 4　逆転項目のロジット値と経験率

Can do ID	Logit	経験率	言語活動	レベル	中央値基準のレベル
Q110142	-1.05	65.1%	聞くこと	A2	↓A1
Q210082	-1.39	69.1%	読むこと	A2	↓A1
Q210102	0.12	32.6%	読むこと	A2	↑B1
Q210142	0.85	17.2%	読むこと	B1	↑B2
Q210182	-0.79	56.9%	読むこと	B1	↓A2
Q310122	1.05	17.9%	やり取り	B1	↑B2
Q310262	0.36	38.7%	やり取り	B2	↓B1
Q310322	-0.15	18.7%	やり取り	A1	↑A2
Q310342	-0.29	55.9%	やり取り	B1	↓A2
Q410022	0.67	14.3%	発表	A2	↑B1
Q410082	-1.79	74.3%	発表	A2	↓A1
Q410092	-0.51	48.7%	発表	B1	↓A2
Q510052	0.59	35.5%	書くこと	A2	↑B1

※「中央値基準のレベル」は上記の逆転項目の選定方法により推定したレベルである。

４．レベルの代表項目

当該レベルの項目としてロジット値で示される困難度が適正であると判断される Can do 項目を当該レベルについて代表する項目として代表項目と呼ぶことにする。代表項目の選定手順は、【文献１】と同様な手順で行った。たとえば、「聞く」の A1 の非逆転項目の平均値は－0.90、A2 の非逆転項目の平均値は－0.70 である。ここから、A1 と A2 の閾値として、－0.90 と－0.70 の平均値である－0.80 が求められる。同様にして、A2 と B1 の閾値として－0.23 が求められる。これにより A1 と A2 の閾値－0.80 と A2 と B1 の閾値－0.23 の間にある、－0.80 以上で－0.23 以下の項目が A2 の代表項目として認定される。なお、A1 と A2 の閾値－0.80 未満の A1 項目が A1 の代表項目として認定される。

今回の調査では、各言語活動のレベル別分類のうち、調査項目数が 3 項目しかないものや 4 項目しかないものがあるため、代表項目数が 2 項目以下となっているレベル別分類も多い。特に「発表」の A2 レベルについては代表項目が 1 項目しかなかったため、A1 レベルとして想定されている Can do 項目 Q410052（言語活動：発表、CEFR レベル：A1、ロジット値－0.539）を「発表」の A2 レベルの準代表項目として追加した。この Can do 項目は逆転項目には選別されていないが、代表項目の選別の基準では A1 レベルの代表項目の範囲に収まらず、ロジット値が A2 レベルとなった項目である。

代表項目数（準代表項目を含む）の合計は 73 項目で、その内訳を表 5 に示した。

表５　言語活動別・レベル別代表項目数

	A1		A2		B1		B2			
言語活動	調査項目数	代表項目数	調査項目数	代表項目数	調査項目数	代表項目数	調査項目数	代表項目数	調査項目合計	代表項目合計
聞くこと	3	2	3	2	6	6	3	3	15	13
読むこと	4	4	5	2	6	4	4	4	19	14
やりとり	10	9	10	6	12	7	7	5	39	27
発表	3	2	3	2	4	2	4	3	14	9
書くこと	4	3	4	2	4	2	3	3	15	10
計	24	20	25	14	32	21	21	18	102	73

※「調査項目数」は逆転項目を除く前の数であり、「代表項目数」には準代表項目を含む。

本節では、便宜的に閾値を求め、それに基づいて代表項目を提示した。【文献１】と同様、ここではあくまでも機械的な振り分けを行ったに過ぎないが、代表項目として認定されたものは生活Can doの記述文として妥当であるといえる。なお、代表項目として認定されなかったCan do項目の多くは、逆転項目の選別の基準と代表項目の選別の基準が異なるために、ロジット値が代表項目の範囲に収まらなかったものである。

代表項目に選別されなかったCan do項目は、日本語のCan doとして不適切であるということを意味しているわけではない。本調査の代表項目は今後も引き続き行っていく質的な検討のための基礎資料に過ぎない。

５．まとめ

（１）「生活Can do」（案）のレベル尺度の妥当性について

全体的にみて、Can doレベルが上がるとともに学習者の自己評価による困難度も上がっており、学習者の自己評価を用いたレベル尺度について十分な妥当性を確認できたといえる。また、代表項目の割合は【文献１】における割合（55.9%）よりも今回の割合（71.6%）の方が高く、この点においても比較的よい結果が得られたものと考える。

個別の項目をみると、【文献１】と同様、一部の項目についてレベル間で逆転を起こしているものが存在している。ただし、逆転項目の割合は【文献１】における割合（17.1%）よりも今回の割合（12.7%）の方が低かった。これらの逆転項目については、Can do記述文の質的な側面から再度検討していくことが必要であろう。

（２）今後行うべき調査について

まず、本調査において代表項目として選定した項目数が少ないレベル別言語活動別の生活Can doの項目をさらに増やすために、今後も同様の調査を継続することを提案する。

さらに、上述の逆転項目が生じる原因について、それぞれのCan do記述文の場面設定の妥当性、日本独自の生活習慣や制度の影響、各言語への翻訳の影響、Can do記述文の作成時に想定した難易度が適切に表現されているかどうかなどについて改めて検討し、レベル内の項目数のバランスも考慮しながら、Can do記述文の調整、代表項目の調整、レベルの調整などを行うべきである。

本調査では、日本語教育機関などでの学習者を対象とした自己評価に基づく分析を行ったが、日本語教師を対象とした学習者評価に基づくデータを収集して分析することも検討すべきであると考える。

以上