

「日本語教育の参照枠」二次報告（案）
－日本語能力評価の考え方について－

文化審議会国語分科会日本語教育小委員会

令和3年1月28日

「日本語教育の参照枠」二次報告（案）
 ー日本語能力評価の考え方についてー
 目 次

はじめに

1. 日本語能力評価の現状と課題	… 1
2. 「日本語教育の参照枠」における言語教育観に基づく評価の三つの理念	… 3
3. 「日本語教育の参照枠」における日本語能力観と評価の考え方	… 5
(1) 日本語能力観について	
(2) 言語使用者・学習者の熟達度を構成する能力及び言語能力	
(3) 客観的に日本語能力を測定する上での評価の基礎的な概念について	
4. 「日本語教育の参照枠」における多様な評価の在り方と事例	… 8
(1) 主な熟達度評価の在り方	
(2) 筆記試験によらない評価の事例	
5. 日本語能力の判定試験と「日本語教育の参照枠」の対応関係を示す方法	… 16
(1) 日本語能力の判定試験と「日本語教育の参照枠」の対応関係を示すことの意味	
(2) Council of Europe (2009) Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)における対応付けの手続き	
(3) 国内の外国語試験と CEFR の尺度への対応付けの事例	
6. 社会で活用される日本語能力の判定試験に求められる要素	… 20
(1) 試験開発に関する基本的な考え方	
(2) 社会的ニーズに応える日本語能力判定の在り方について	
<参考>	
参考文献	… 25
参考資料 1 評価の種類	… 26
参考資料 2 日本語の能力判定に係る試験等一覧	… 33

○ はじめに

本報告は、令和 2 年 1 1 月に文化審議会国語分科会日本語教育小委員会が取りまとめた「日本語教育の参照枠」一次報告に基づき、国内外の日本語学習者が日本語の習得段階に応じて、求められる日本語教育の内容及び方法を明らかにし、外国人等が適切な日本語教育を受けられ、評価できるようにするため、外国人の日本語能力の判定基準及び評価の在り方について示し提言するものです。

外国人の日本語能力を判定する方法として国内外で様々な試験が実施され、個々の指標に基づき、レベルや判定基準等が設定されていますが、学習・教育内容の多様化が進む中、各試験が判定する日本語能力についての共通の指標を整備し、利用できるようにすることが必要となってきました。

現在、地域の日本語教室では、外部の試験団体が実施する試験の測定結果に基づく評価を必要としない者も少なくないという状況がある一方で、日本語の熟達度を判定する共通かつ簡易な評価指標がなく、評価ツールなども整備されていないことが課題となっています。そこで、いわゆる試験による評価ではなく、日本語教育の個々の現場で実施できるパフォーマンス評価の方法と事例及びポートフォリオによる評価や自己評価などの代替的評価の方法と事例についても幅広く示していくことが必要になってきています。

令和元年 6 月 28 日に公布・施行された「日本語教育の推進に関する法律」（令和元年法律第 48 号）第 22 条においては、日本語教育を受ける者の日本語能力に応じた効果的かつ適切な教育が行われるよう、教育課程の編成に係る指針の策定、指導方法及び教材の開発・普及、並びにその支援、その他の必要な施策を講ずる旨の規定が盛り込まれました。

政府は、この法律に基づき、「日本語教育の推進に関する施策を総合的かつ効果的に推進するための基本的な方針」を令和 2 年 6 月 23 日に閣議決定し、日本語教育の推進の基本的な方向や具体的施策例などの内容等を定めました。この中で、「ヨーロッパ言語共通参照枠」（以下、CEFR という。）を参考に、日本語の習得段階に応じて求められる日本語教育の内容・方法を明らかにし、外国人等が適切な評価を受けられるようにするため、「日本語教育の参照枠」を作成することとしています。

一方、平成25年、文化審議会国語分科会日本語教育小委員会の下に設置された「課題整理に関するワーキンググループ」において取りまとめられた「日本語教育の推進に向けた基本的な考え方と論点の整理について」（報告）に、「論点3. 日本語教育の標準と日本語能力の判定基準について」が盛り込まれました。これらを踏まえ、令和2年度から日本語教育小委員会に「日本語能力の判定基準に関するワーキンググループ」を設置し、審議を進め、今般「日本語教育の参照枠」二次報告の取りまとめに至ったものです。

本報告は6章で構成されています。

第1章では、日本語能力評価の現状と課題について整理しました。

第2章では、「日本語教育の参照枠」における言語教育観に基づく評価の三つの理念を示し、第3章で、「日本語教育の参照枠」における日本語能力観と評価の考え方を示しました。その上で、第4章で、日本語教育における多様な評価の在り方を事例とともに分かりやすく紹介しています。

第5章では、「日本語教育の参照枠」と日本語能力の判定試験や評価の対応関係を示す方法について、手続きや事例を示しています。試験によらない多様な評価の方法を示すことで、今後、日本語能力の判定における選択肢が広がっていくことが期待されます。

第6章では、社会で活用される日本語能力の判定試験に求められる要素として、試験開発に関する基本的な考え方や社会的ニーズに応える適切な日本語能力判定の在り方について提言しています。

さらに、参考資料として、日本語の能力判定に係る試験等一覧を収録しました。

今後、本報告が日本語能力の判定を行う際に参照され、多様な日本語教育の現場で用いられ、関係者の日本語能力評価に関する理解が深まるとともに、国内外で多様な日本語教育の目的に応じた質の高い日本語の試験が開発され、適切に実施され、より良い日本語の能力判定の方法が普及することにより、日本語教育の一層の推進に寄与することを望むものです。

1. 日本語能力評価の現状と課題

(1) 日本語能力の評価について

日本語能力の評価について、現在、以下のような現状と課題がある。

- 世界中で国境を越えた人の移動が進み、複数の言語を使用し、複数の社会に生きる人々が増えている。このような状況において、国内外を移動する日本語学習者が継続的に自らの日本語の熟達度を客観的に把握したり、具体的な学習目標を立て自律学習を進めたりするために、有効に活用できる評価の仕組みが求められている。
- 現在、国内外の多様な日本語学習者に対する日本語能力の評価に対応できる、国としての共通の指標等が示されていない。そのため、国内及び海外の日本語教育の現場では、個々に独自に定めた指標による評価が行われている。しかし、国内外の人の往来や、生活、留学、就労等に応じた日本語学習の目的の多様化を受けて、国内外で共通して参照できる包括的な評価の枠組みが必要となっている。
- 日本語に関する知識だけではなく、生活、留学、就労等に必要な日本語のコミュニケーション能力が求められるようになってきたことから、そのようなコミュニケーション能力を測定し判定することが求められている。しかし、日本語教育機関や日本語教師等に対して、日本語学習の目的や言語活動に応じた多様な評価の方法についての選択肢が十分に示されていない。
- 地域の日本語教室では、様々な日本語のレベルの外国人を随時受け入れている状況にあるが、日本語の熟達度を判定する指標がない中で、レベルに基づくクラス分けが困難なために教室運営に支障をきたすなどの問題が生じている。生活者が集まる地域の日本語教室において、共通的に用いることができ、かつ簡易な評価指標がなく、評価ツール等も整備されていない。
- 地域の日本語教室等で日本語を学ぶ学習者の中には、試験団体が実施する試験の測定結果に基づく評価を必要としない者も少なくない。また、地域の日本語教育に評価はなじまないとの意見もあるが、評価とは正確な測定を第一義とする試験と同義ではなく、本来多様で豊かなものである。そのため、日本語教育の個々の現場で実施できるパフォーマンス評価の方法と事例及びポートフォリオによる評価や自己評価などの代替的評価の方法と事例についても幅広く示していくことが必要である。

(2) 日本語能力を判定する試験について

日本語能力を判定する試験については、現在、以下のような現状と課題がある。

- 外国人等の日本語能力を判定する方法として国内外で様々な試験（約 20 の機関・団体）が実施され、各試験の目的に応じて、得点の解釈基準やレベル設定、レベル判定基準等が設定されているが、学習・教育内容の多様化が進む中、各試験が測定する日本語能力についての測定結果を相互に参照できる枠組みを整備し、利用できるようにすることが必要となっている。
- 例えば、留学生にとって必要な日本語の知識や能力を測る試験で示された日本語能力のレベル判定基準が、そのまま「生活者としての外国人」や就労者に対する日本語能力評価や学習目標の指標として用いられてしまうことがあり、留学生とは異なる目的、場面で言語活動を行う人の日本語能力について、適切な判定がなされていない。
- 現行の日本語能力を判定する試験においては、日本語による言語活動のうち、「読むこと」、「聞くこと」を評価するものは多く存在するが、「話すこと」、「書くこと」、いわゆる産出に関する言語能力を評価するものが少なく、その評価のための基準も明確に示されているとは言えない状況にある。
- 日本語能力を判定する試験においては、これまでカンニング等の不正行為や合格証の偽造等の問題が度々指摘されている。特に、社会的ニーズに応える日本語能力を判定する試験実施機関に関しては、試験の実施に際し、最大限対策が求められる。
- また、日本語能力を判定する試験については、その出題基準や過去問題、採点方法等の情報が十分に公開されているとは言えないものもあり、日本語教育機関及び日本語学習者が必要な試験を選びにくい状況にある。また、試験及び実施団体の適格性が担保されているかどうか分かりにくいという指摘がある。
- 国内外で日本語能力の証明のための試験実施の需要は拡大している。多様なニーズに応えられるよう、AI 技術の活用やオンラインによる受験が可能となるような試験の開発を促進し、日本語能力の判定機会の充実を図る必要がある。
- 特定技能等の在留資格では、入国要件に一定の日本語能力が課せられている以外に、業種や職能に応じた日本語能力を判定するための試験が求められるようになってきている。例えば、介護に関する職業に従事する上で必要な日本語能力を測定する試験など、個別の分野に特化した日本語能力の判定試験が開発されるようになってきている。しかし、職業分野ごとに求められる日本語能力の判定基準については策定が進んでいない。

以上のような現状と課題に鑑み、国内外を移動する日本語学習者に対する日本語能力の評価に対応できる、国としての共通の指標等を示し、かつ多様な評価の方法を幅広く示すことによる包括的な評価の仕組みを整備することが必要である。

2. 「日本語教育の参照枠」における言語教育観に基づく評価の三つの理念

- 国内外における日本語学習者の日本語の習得段階に応じて求められる日本語教育の内容及び方法を明らかにし、外国人等が適切な日本語教育を受けられ、評価され、また自己評価できるようにするため、「日本語教育の参照枠」の考え方にに基づき、日本語能力を判定する参照枠の在り方について策定する。
- その際、「ヨーロッパ言語共通参照枠（CEFR：Common European Framework of Reference for Languages: Learning, teaching, assessment）」（以下、CEFR（2001）という。）を参考に、言語知識を測定する筆記試験等による評価だけでなく、パフォーマンス評価及びポートフォリオによる評価などの代替的評価を含めた評価の在り方を示す。
- 国内外で様々な日本語能力を測定し判定する試験が実施され、個々の指標に基づき、レベルや判定基準等が設定されているが、学習・教育内容の多様化が進む中、各試験はそれぞれ異なる目的をもって開発されるため、レベルや判定基準は各試験で異なるものが設定されるものであるが、一方でそれらの背後に汎用性の高い共通の参照枠を整備し、利用できるようにする必要もある。
- 「日本語教育の参照枠」では、「日本語教育の推進に関する法律」第一条に掲げる「多様な文化を尊重した活力ある共生社会の実現に資するとともに、諸外国との交流の促進並びに友好関係の維持及び発展に寄与する」ことを理念として示し、言語教育観の柱として以下の三つを示した。全ての指標はこの考えに基づいて示されている。

① 日本語学習者を社会的存在として捉える

学習者は、単に「言語を学ぶ者」ではなく、「新たに学んだ言語を用いて社会に参加し、より良い人生を歩もうとする社会的存在」である。言語の習得は、それ自体が目的ではなく、より深く社会に参加し、より多くの場面で自分らしさを発揮できるようになるための手段である。

② 言語を使って「できること」に注目する

社会の中で日本語学習者が自身の言語能力をより生かしていくために、言語知識を持っていることよりも、その知識を使って何ができるかに注目する。

③ 多様な日本語使用を尊重する¹

各人にとって必要な言語活動が何か、その活動をどの程度遂行できることが必要か等、目標設定を個別に行うことを重視する。母語話者が使用する日本語の在り方を必ずしも学ぶべき規範、最終的なゴールとはしない。

¹ 本項目は、CEFR第1章3節「複言語主義(plurilingualism)とは何か」(吉島・大橋訳編(2014))で示している「究極目標としては「理想的母語話者」を考えるとといったようなことはなくなる」という部分を参照しつつ、日本語教育の現状を踏まえて定めたものである。ここで示す「多様な日本語使用」とは、地域の多様な言語使用を尊重する意味から、必ずしも共通語を規範とするものではないということの意味する。

- この三つの言語教育観の柱に基づき、「日本語教育の参照枠」における三つの評価の理念を示す。下記の三つの評価の理念のうち、①と②は評価の目的、つまり何のために評価を行うかについての理念であり、③は、評価を行う上での手法に関する理念である。

① 生涯にわたる自律的な学習の促進

「日本語教育の参照枠」における評価は、生涯にわたる自律的な学習の促進²を目的とする。

② 学習の目的に応じた多様な評価手法の提示と活用推進

「日本語教育の参照枠」では、日本語を使用して、何が、どのように、どれくらいできるのかを言語能力記述文等を用いて具体的に示すとともに、それがどの程度達成できたかを把握するために、多様な評価手法を提示し、その活用を後押ししていくための考え方や事例を示す。

③ 評価基準と評価手法の透明性の確保

日本語学習者、教師ばかりでなく、一般の日本人等にとっても参照しやすい、日本語で「できること」に注目した評価基準を示し、その評価手法の透明性を確保することを通して、日本語教育に関わる全ての人の中で評価に関する共通認識を醸成する。これにより、日本語学習者がいつ、どこにいても、一貫した学びを継続できる環境の整備を目指す。

² CEFR (2001) p.1 にも「CEFR はさらに学習者の熟達度のレベルを明示的に記述し、それぞれの学習段階で、生涯を通して学習進度が測れるように考えてある(吉島・大橋訳・編 2014,p.1)」とある。

3. 「日本語教育の参照枠」における日本語能力観と評価の考え方

(1) 日本語能力観について

- 評価を行うにあたっては、その対象となる日本語能力をどのように捉えるかという能力観を明確にする必要がある。「日本語教育の参照枠」一次報告では、日本語能力観について、行動中心アプローチに基づき日本語の熟達度を五つの言語活動ごとに示し、必要なことから学んでいくことを重視するとした。
- 行動中心アプローチとは、多様な背景を持つ言語の使用者及び学習者を、生活、就労、教育等の場面において、様々な言語的/非言語的な課題 (tasks) を遂行する社会的存在として捉える考え方のことである。
- 課題 (tasks) には、社会の中で目的を持って行う言語的/非言語的行動の全てが含まれる。買い物をする、交通機関を利用すること、娯楽として映画や読書を楽しむこと、地域社会・学校・職場などでよりよい人間関係を構築すること、職を得て働き収入・やりがいを得ていくこと、教育を受け教養を身に付けることなど、多岐にわたる。
- 行動中心アプローチにおける言語教育の目標とは、言語の使用者及び学習者がそれぞれの社会で求められる課題を遂行できるようになることである。したがって、学習者は、文法や語彙の難易度、言語活動間のバランスにかかわらず、課題を遂行するために必要な事柄から学ぶことができる。

(2) 言語使用者・学習者の言語能力熟達度を構成する能力

- CEFR (2001) では、「人間の全ての能力は、言語使用者がコミュニケーションを行う力に何らかの形で寄与することから、あらゆる能力はコミュニケーション能力の一部と考えてよい。それでも、言語とはそれほど緊密に関わらないものを、狭義の言語能力の範疇に含まれるものから区別することは意義あることだろう。」として、言語使用者 / 学習者の言語熟達度を構成する能力及び言語能力を、次の四つに整理して示している。

① 一般的能力

一般的能力とは、叙述的知識（世界・社会文化・異文化などについての知識）、技能とノウ・ハウ（生活や余暇・社会的・異文化間・職業的な技能）、実存的能力（態度・動機・価値観・信条・認知的スタイル・性格）、学習能力（言語とコミュニケーションに関する意識・音声意識と技能・学習技能・発見技能）から構成される。

② コミュニケーション言語能力

コミュニケーション言語能力は、語彙、音韻、統語論に関する知識や技能である言語構造的な能力、言語の社会文化的な条件下での言語使用と関連する社会言語能力、談話の構成能力のような言語素材を使うときの機能面に関する能力である言語運用能力から構成される。「日本語教育の参照枠」では、これらの能力に基づき「能力 Can do」を示している。

③ コミュニケーション言語活動

言語活動は、受容、産出、やり取り、仲介³の四つのモードから構成される。「日本語教育の参照枠」では、コミュニケーション言語活動として五つの言語活動（受容的：「聞くこと」及び「読むこと」、産出：「話すこと（発表）」及び「書くこと」⁴、やり取り：「話すこと（やり取り）」）についての言語能力記述文を「活動 Can do」として示している。

④ コミュニケーション言語方略

方略とは、言語活動を行う上で駆使する、分からない言葉などに対する推測や質問、聞き取りにくい言葉について聞き返したりする行動を指す。また、コミュニケーション言語方略は四つのモード（受容、産出、やり取り、仲介）ごとに整理することができる。「日本語教育の参照枠」では、受容、産出、やり取りについて「方略 Can do」を示している。

- ①から④のうち、日本語能力として評価の対象となるのは②、③、④であり、「日本語教育の参照枠」一次報告では②、③、④の能力を表す言語能力記述文（Can do）を示している。

(3) 客観的に日本語能力を測定する上での評価の基礎的な概念

- 日本語能力を測定する際には、その原則となる基礎的な概念について整理しておく必要がある。CEFR（2001）では、評価の議論には、伝統的に基本となる以下の三つの概念があるとしている。この三つの概念は主に試験などで客観的に日本語能力を測る際の原則となるものである。さらに細分化したものについては「6. 社会で活用される日本語能力を測定する試験に求められる要素」で詳述する。

³ 本報告では、「日本語教育の参照枠」一次報告でもおいて示した通り、CEFR(2001)を主に参考として検討を行うこととするため、仲介活動等についての検討は今後の検討課題とする。

⁴ CEFR(2001)「共通参照レベル：自己評価表」に従って整理した。産出：「書くこと」については、「書く産出活動（書くこと）」、「書かれた言葉のやり取り」に分けて言語能力記述文が示されている。

- ・妥当性 (validity)

あるテストや評価の手法が、当該の状況で、測定目的としたものと、実際に測定しているものとが一致しており、またそこで集められた情報が当該の学習者の熟達度を正しく示している場合に、そのテストや評価は妥当性が高いといえることができる。

- ・信頼性 (reliability)

その評価の手法によって、ぶれのない安定的な評価結果を出すことができるか。古典的なテスト理論では、テストの測定精度を信頼性係数 (reliability coefficient) で表す。

- ・実行可能性 (feasibility)

その評価の手法が現実的に実行可能であるかどうか。

※参考：CEFR における評価

【第九章 評価(Assessment)】(CEFR 2001, 吉島・大橋訳 2014 p.199)

- ・ CEFR では、評価を言語学習のプログラムの広義の評価 (evaluation)⁵ という広い問題ではなく、限定的な意味でのアセスメント (assessment) として扱っていることから、本報告でも、評価については熟達度評価を中心に示すこととする。

広義の評価
(evaluation)

- ・ 学習者の熟達度についての評価 (assessment)
- ・ ある方法や言語教材の効率性
- ・ 言語学習のプログラムで実際に産出されたディスコースの種類や質
- ・ 学習者/教師の満足度
- ・ 教育の効率性

⁵ evaluation については吉島・大橋訳 (2014) では「総括」と訳出している。しかし、evaluation を総括と訳すことについては一般的であるとは言えず、また、10 ページに出てくる「総括的評価 (Summative assessment)」との混乱を避けるため、本報告では evaluation を「広義の評価」とした。

4. 「日本語教育の参照枠」における多様な評価の在り方と事例

- CEFR (2001) では評価を「学習者の熟達度についての評価 (assessment)」として扱っていることから、本報告でも、評価については熟達度評価を中心に示すこととする。
- CEFR (2001) では、評価についての論点として、「評価の方法や伝統はさまざまであるが、あるアプローチ (例：教師による評価) より、別のアプローチ (例：公的な試験) の方が、教育上の効果において絶対に優れていると考えるのは間違いである。本書にある共通参照レベルのような、一連の共通基準の主要な利点は、まさにお互いに異なる評価の形式でも対応付けが可能になることである。」ことを挙げ、教育の目的に応じて様々な手法を組み合わせたり、取捨選択したりして評価を行うことを推奨している。
- 評価の在り方については、試験によるものとそうでないものがある。試験によらない評価は、代替的評価とも呼ばれ、試験では測ることが難しいとされる、言語を用いた課題遂行能力や学習過程における様々な気付きや学びを把握するための評価の方法のことを指す。言語能力の熟達度の評価は、コースで設定した学習目標や学習者の特性に応じて、試験と試験によらない評価を組み合わせることで総合的に実施していくことが望ましい。

(1) 試験

ある教育プログラムにおける試験では、ある期間内に扱った学習目標の到達度を測る試験と、学習した教育プログラムの構成やカリキュラム、教材に関わらず、その時点で何ができるかという熟達度を測る試験がある。熟達度を測る試験については試験団体が実施する試験を受ける場合もある⁶。試験による評価の利点は以下の通りである。

- ・学習対象の学習成果を表わす熟達度を、より一般的な言語能力尺度上に位置づけて表わすことができる。
- ・異なるテスト間の測定道具としての共通性および違いを明確にできる。
- ・国や機関を超えて共通に参照できる、日本語能力を評価する枠組みや構成概念の設定及び、測定道具 (テスト) の仕様を検討する基本設計図として活用できる。
- ・学習者が自身で言語能力の目標設定や評価 (到達点の確認, 調整) についての見通しを持つことができる。

⁶ 試験については、従来型の紙ベースの試験のほかに、近年では CBT (Computer-Based Test) が導入されている。

(2) パフォーマンス評価

パフォーマンス評価とは、学習者に例えばロールプレイやエッセイなどの言語的な課題を与え、その遂行の度合いを評価することをいう。パフォーマンス評価は到達度、あるいは熟達度を測る試験として実施する場合と、試験によらない評価として実施する場合がある。

日本語教育においては産出的言語活動（「話すこと（やり取り）」、「話すこと（発表）」、「書くこと」）についての到達度や熟達度を測ることが多く、特に就労場面における産出的活動についての熟達度を測ることについては、社会的に高いニーズがある。パフォーマンス評価の利点は以下の通りである。

- ・教師と学習者の双方がパフォーマンスに関する評価基準を共有することで、評価の透明性を高めることができる。
- ・学習者は与えられたパフォーマンス課題に対して、評価基準を基にした明確なフィードバックを得ることができる。
- ・単にできた / できない、だけの評価だけでなく、何が、どのくらいできたのか、について、多様な観点から評価を行うことができる。

パフォーマンス評価を行う際には、ルーブリックによる評価を行う場合がある。ルーブリックとは、例えば言語的課題（例えば「家族を紹介する」などの言語的タスク）の達成度と文法的正確さや使用語彙の範囲、発音などの質的側面等の観点を組み合わせた評価基準表のことをいう。ルーブリックによる評価の利点は以下の通りである。

- ・評価の観点を明示することで、個人的な価値判断による影響を避け、達成すべき目標を学習者と共有することができる。

パフォーマンス評価の例

- ・ACTFL-OPI(The American Council on the Teaching of Foreign Languages-oral proficiency interview)は、汎言語的に使える会話能力テストであり、ACTFL 言語運用能力基準に基づいて、初級から超級までの 10 レベルで判定される。
- ・国際交流基金(2016)では、JF 日本語教育スタンダードに準拠したロールプレイテスト(A1～C1)を公開している。
- ・豊田市が運営する「とよた日本語学習支援システム」では、「とよた日本語能力判定」を開発し、独自の指標によるパフォーマンス評価を行っている（豊田市 2011）。

(3) 自己評価

自己評価とは、言語能力記述文のリストで構成された自己評価表などを用いて、自身の言語熟達度を把握することの他、学習に対する振り返りを記述し、学習の過程で読み返したりすることを通して、自律的な学習能力を育成することを目的とした評価活動のことを指す。自己評価の利点は以下の通りである。

- ・試験による評価と同様に、学習者が自身で言語熟達度の目標設定や評価(到達点の確認、調整)についての見通しを持つことができ、自身の学習に対する意識を高めることができる。
- ・学習についての振り返りの記述を継続的に記し、ポートフォリオに記録として残すことにより、学習者及び、その周りの人々に学びのプロセスを示すことができる。

自己評価の例

- ・ **CEFR** スイスプロジェクトの自己評価チェック表 (Self-assessment checklist)⁷
 では、各レベルの言語能力記述文についての評価を、自分で行うばかりでなく、教師による評価欄が設けられている。また、個々の言語能力記述文が自分の目標かどうかのチェック欄があり、学習者にその項目の重要性も考えさせることができる。この項目によって何をいつ学ぶかについて教師が一方的に決めるのではなく、学習者が自ら判断することによって自律的な学習を促す効果がある。
- ・ 文化審議会国語分科会 (2012) 『「生活者としての外国人」に対する日本語教育における日本語能力評価について』では、「生活者としての外国人」向けの「日本語学習ポートフォリオ」を公開しており、その中で学習者による自己評価と教師によるチェックリストの例を「生活上の行為達成の記録 (p.54)」として示している。

⁷ <https://www.laits.utexas.edu/fi/sites/laits.utexas.edu.fi/files/Self%20Assessment%20Checklist%20European.pdf> (令和 2 年 10 月 30 日閲覧)

(4) 相互（ピア）評価

相互(ピア)評価とは、学習者とその周りの人が相互に評価を行うことである。クラスメートや家族、就労場面であれば職場の人などから、日本語の使用についてのコメントを得ることで、自分の熟達度を多面的に把握することができる。相互(ピア)評価の利点は以下の通りである。

- ・クラスメートや周りの人々から学習についてフィードバックを得ることで、学習に対する動機を高めることができる。また、他の学習者の評価に関わることを通して自己評価に対する内省を深めることができる。

相互（ピア）評価の例（掲載予定）

(5) ポートフォリオによる評価

ポートフォリオによる評価とは、多様な広がりを見せる学習者の達成成果を学習者の様々な必要性、性質や資質に応じて記述し、評価することである。CEFR(2001)では、複文化・複言語能力の育成を推進するため、ヨーロッパ言語ポートフォリオ(European Language Portfolio: ELP)⁸を用いて学習者一人一人がさまざまな面から自分の言語発達を記録できるようになっている。

ポートフォリオは、ファイル形式で作成することにより、筆記試験の結果、パフォーマンス評価で使ったルーブリック、自己評価チェックリスト、相互(ピア)評価で行った、他の学習者からのコメントシートを格納することができ、学習者や教師をはじめとする学習者の周りの人々は適宜これらの評価結果を参照することで、総合的な評価を行うことができる。

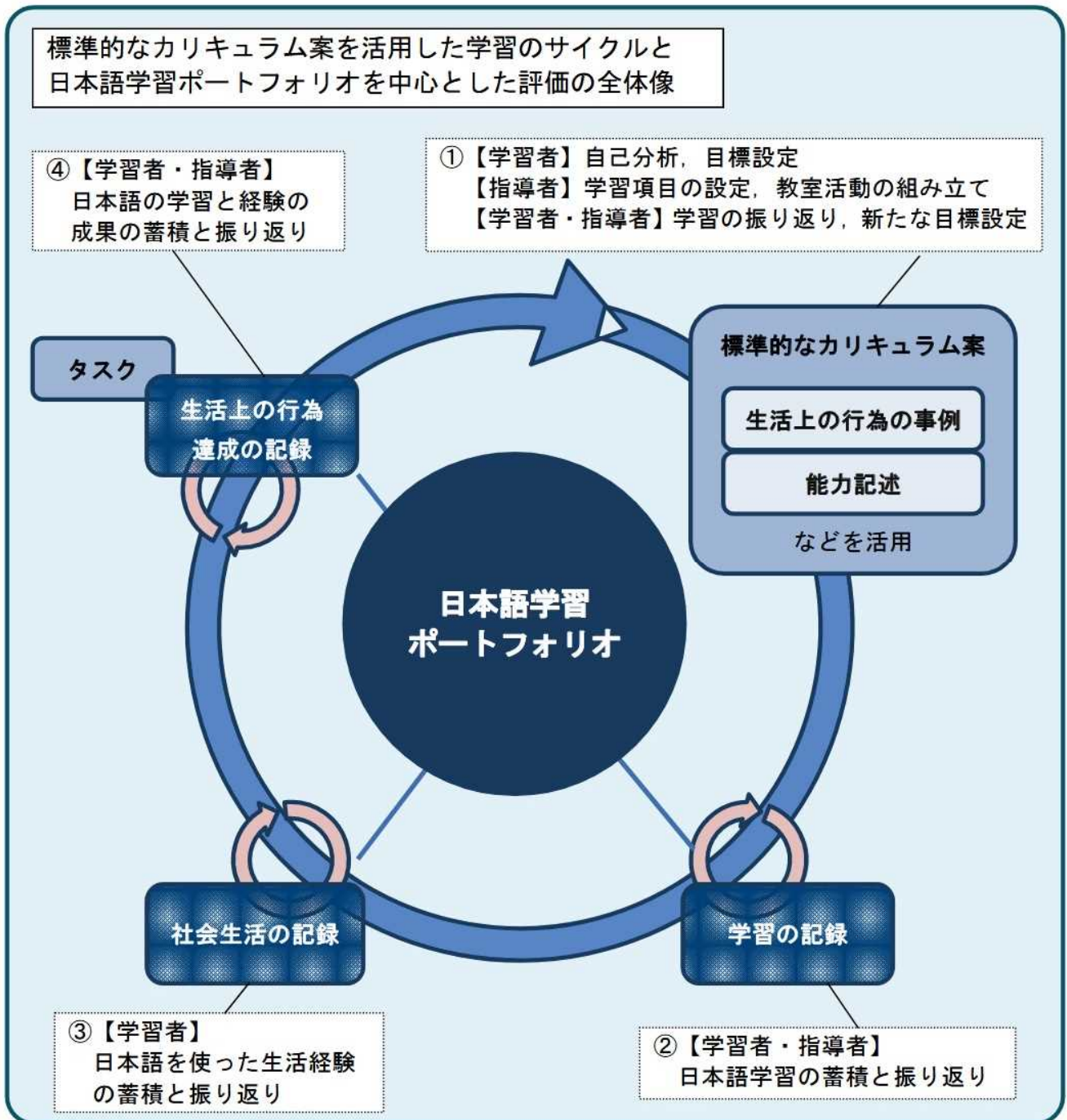
- ・ 学習者自身、教師、学習者の周りの人が、学習者の言語の熟達度の成長の過程を通時的に把握することができる。
- ・ 学習者が学習機関を移動し、新たに学習を始める際にポートフォリオを示すことで、教師はその学習者がこれまで学んできた内容と熟達度が把握でき、適切な教育内容を準備することができる。
- ・ ポートフォリオにおける振り返り記述などについて、記述内容についての評価の観点を盛り込むことで、質的な評価が可能となる。また、評価の各段階に得点を与えることによって、数値化が難しい評価対象に総合点を与えることができる。

ポートフォリオ評価の例

- ・ 欧州評議会が公開しているヨーロッパ言語ポートフォリオ(European Language Portfolio: ELP)は、一つまたは複数の言語を学ぶ学習者が言語の学びと文化の学びについて記録し、振り返ることができるものであり、言語パスポート、学習の記録、学習に関する資料の三つから構成される。ヨーロッパ言語ポートフォリオは、学習者の年齢や目的によって多種多様なものが公開されており、学習者の自律的な学習、複言語・複文化の学びを促進するために利用されている。

⁸ <https://www.coe.int/en/web/portfolio> (令和2年8月20日閲覧)

- ・文化審議会国語分科会（2012）は、「生活者としての外国人」向けの「日本語学習ポートフォリオ」を公開し、ポートフォリオによる評価の方法を示している。また、文化庁の補助事業等を利用し、各地域の機関や団体が地域の実情や教育目的に応じたポートフォリオを作成している。



「生活者としての外国人」に対する日本語教育における日本語能力評価について (p.19)

- ・国際交流基金の **JF** 日本語教育スタンダードでは、ポートフォリオを評価の中心の一つとして位置づけ、世界各地でそれぞれの現場独自のポートフォリオを作成し、評価活動を行っている（国際交流基金 2017）。

5. 日本語能力の判定試験と「日本語教育の参照枠」の対応関係を示す方法

(1) 日本語能力の判定試験と「日本語教育の参照枠」の対応関係を示すことの意味

現在、国内外で実施されている日本語能力の判定試験（約 20 の機関・団体）は、各試験の目的に応じて、得点の解釈基準やレベル設定、レベル判定基準等が設定されている。これらの試験が「日本語教育の参照枠」との対応付けを行うことによって、各試験が測定する日本語能力についての測定結果を相互に参照できる枠組みを整備し、利用できるようにすることが可能となり、異なる試験間の通用性が高まる。また、共通の指標での評価を得ることができることで、受験者はどの試験を受験しても、熟達度のレベルについて、個別の試験の独自性や特質を勘案した上で、測定結果を相互に参照できる枠組みに基づいた教育的なフィードバックを得ることができる。

(2) Council of Europe (2009) Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR)における対応付けの手続き

Council of Europe (2009)では、CEFR の尺度への対応付けのために次の 5 つの手続きを示している。

① CEFR への理解を深める (Familiarisation)

対象となる資格・検定試験の対応付けを行う専門家集団（パネル・メンバー）に対し、CEFR、そのレベル区分、言語能力記述文への理解を深める研修を行うこと。対象となる試験の対応付けを行う委員に対し、CEFR レベルを理解するためのトレーニングを実施する。トレーニングは、事前課題とワークショップに分けられる。トレーニングには以下の a) から i) がある。

事前課題：

- a) CEFR (2001) の第 3 章第 6 節を読んで各レベルの弁別的特徴を理解する。
この際、各レベルの言語活動だけでなく、機能、概念、文法、語彙などの例示尺度についても十分理解する。
- b) コーディネーターによって作成された、対象となる試験を CEFR と対応付けする上で必要となる観点（第 3 章～第 5 章の各節末の問いより抜粋）をまとめたチェックリストを確認する。
- c) CEFTrain (<http://www.helsinki.fi/project/ceftrain/>) にアクセスし、各レベルの弁別的特徴を示したパフォーマンスに実際に触れ、言語能力記述文の分析を通して CEFR のレベルをさらに深く理解する。

ワークショップ（約 3 時間）：

＜導入活動＞

- d) Council of Europe (2009) 付録 A 1 の表 (CEFR 第 3 章第 6 節の短縮版) を用いて、レベルの異なる言語能力記述文をレベル順に並べ替える活動を行う。
- e) CEFR (2001) に収録されている「自己評価表」に基づいて自分ができる外国語について自己評価を行う。さらに、外国語能力の質的側面に関して、「話し言葉の評価表」または「話し言葉の流ちょうさ」や「文法的正確さ」についての言語能力記述文を用いて自己評価を行うこともできる。その後、他の参加者との共有・議論を行う。

＜言語能力記述文の質的分析＞

- f) CEFR (2001) に収録されている言語能力記述文を、レベルを伏せてバラバラにした上で並び替えてレベル付けを行い、なぜそのレベルをつけたのかについてグループで検討を行う。(いくつかのカテゴリーにまたがる場合は、言語能力記述文の合計が 40 を超えない程度とする)
- g) CEFR (2001) に収録されている「自己評価表」に用いられている個々の言語能力記述文をバラバラに切り離しレベルを伏せた切片を準備し、それらを正しい位置に配置しなおす活動を行う。

＜産出技能評価の準備＞

- h) CEFR の言語能力評価基準表の穴埋めまたは並べ替えタスクを行う。
CEFR の「話すこと」から始めるのであれば、Council of Europe (2009) の付録 C 2 の表、「書くこと」から始めるのであれば、C 4 の表を用いる。(対応付けの対象となる試験に産出技能の評価がない場合でも必ず行うこと)
- i) ビデオに撮られた学習者のパフォーマンスを用いて、CEFR レベルを説明する。

② 対象となる資格・検定試験を自己点検し、明確化する (Specification)

資格・検定試験の問題内容や問題タイプについての自己点検を行い、当該試験の出題範囲およびレベルが CEFR と対応付けられること。また、CEFR と対応付かない領域について記述をすること。さらには、内容分析に基づき、CEFR の言語能力記述尺度を用いた当該の試験のプロフィールを描くこと。

自己点検にあたっては、Council of Europe (2009) 付録のセクション A 2 の書式のチェックリストを利用して、対応付けの対象となる試験の内容分析を行う。セクショ

ン A2 の書式は全部で 24 あり、内容は以下の通りである。

A1-7：対応付けの対象となる試験の概要に関する書式

A8：対応付けの対象となる試験の CEFR レベルの最初の推定に関する書式

A9-22：対応付けの対象となる試験問題内容に関する書式（A9-18：コミュニケーション言語活動および A19-22：コミュニケーション言語能力）

A23：対応付けの対象となる試験の CEFR と対応付けられた出題範囲とレベルの主張のためのプロフィールの図示に関する書式（必ずしも試験内容の低位分類名と一致している必要はない）

A23：対応付けの対象となる試験の CEFR レベルの最終的な推定に関する書式（A8 の書式と異なる推定になった場合はその理由についても明記する）

③ 標準化トレーニングを行い、レベルを設定する (Standardisation training and benchmarking)

パネル・メンバーが基準設定（資格・検定試験のスコアを CEFR に対応付けること）を行うため、試験課題と実際のパフォーマンス例に基づいて、パネル・メンバーの間で CEFR レベルに関する一貫した共通認識を得ること。

話すこと、書くことの実際のパフォーマンスについて、Council of Europe (2009) 付録のセクション C1～C4 の評価表を用い、次の三つの段階に分けてレベル判定のトレーニングを行う。その後、対応付けの対象となる試験に関するレベル付けがされていないパフォーマンスの判定を行い、合意形成を行う。また、評価者間や評価者内の信頼性の分析も行う。

第一段階：レベルが確定しているパフォーマンスについての解説を C2, C3（話すこと）および C4（書くこと）の評価表を用いてコーディネーターが行う。C1 の全体尺度を最初に用いてもよい。

第二段階：レベルが確定しているパフォーマンスの判定をコーディネーターのアドバイスを受けながらグループで行う。

第三段階：レベルが確定しているパフォーマンスの判定を個々に行う。

リスニング、リーディング、言語能力についても同様の段階を踏んでトレーニングを行う。出題されるテキストのレベルそのものではなく、問題の難易度との組み合わせにより、受験者の能力を位置づけることに注意する。

④ 基準を設定し、CEFR の段階別表示に位置付ける (Standard setting Procedures)

パネル・メンバーがグループでの数次の審議を経て資格・検定試験のスコアを CEFR の段階別表示に位置付けること。

いくつかの統計的な手法を用いて、受験者のデータを CEFR のレベルに分割し、対応付けの対象となる試験におけるスコアのそれぞれのレベルの境界を明らかにする。この作業に使用する統計の手法は、3つのグループに分類される。

- 1) 実証的なデータを使用せずにテスト項目を専門家が評定して判定する手法
- 2) 実証的なデータを使用せずに受験者集団を焦点にした手法
- 3) テストの項目困難度と受験者集団についての実証的なデータを IRT (Item Response Theory) により分析する手法

各グループについての具体的な手法については、Council of Europe (2009) に示されている。

⑤ 妥当性を検証する (Validation)

上記①～④の手続きが適切に行われているか、質的、量的な方法に則り継続的に検証する。

(3) 国内の外国語試験と CEFR の尺度との対応付けの事例

① 公益財団法人日本英語検定協会

Dunlea, J. (2009,2010), 公益財団法人日本英語検定協会 (2018) では、欧州評議会が示している CEFR の尺度への対応付けの手法を用いて英検の各級及びライティングスコアと CEFR レベルとの対応付けを示している。

② 一般財団法人進学基準研究機構

(Center for Entrance Examination Standardization(CEES))

英語コミュニケーションテスト GTEC においては、2016 年度から 2017 年度にかけて CEFR の尺度との対応付けを行い、Pre A1/A1, A1/A2, A2/B1, B1/B2, B2/C1 の各閾値を設定した。2019 年度以降、GTEC 受験者や教師へのフィードバックとして CEFR を用いることを決定している。

6. 社会で活用される日本語能力の判定試験に求められる要素⁹

(1) 試験開発に関する基本的な考え方

試験は目的に応じて、開発すべきものであり、本報告は大規模試験において最低限必要となる要素を示すものである。それぞれの概念の説明の下の項目はチェックリストである。ただし、これらの項目はあくまでも例示として挙げているものであり、作成する試験の用途や目的に応じて項目の検討が行われるべきである。

① 有用性 (usefulness)

テストは妥当性、真正性、信頼性、波及効果等のさまざまな観点から評価されるが、そのテストの総合的な価値を個々の観点から見た価値の総和として捉える概念である。

② 妥当性 (validity)

テストが測定目的とした能力や特性（一般化して「構成概念」と呼ばれる）を確かに測定しているか否かを表わす概念で、「構成概念妥当性」を中心に据えて、それを確認する方法により妥当性の異なる側面が強調される。

<試験作成・評価に関する妥当性>

- 測ろうとしている言語能力を測っているか。
- 測定する言語能力の理論的基礎（根拠）となる文献を明示しているか
- 試験作成の過程で測るべき知識や能力の一覧を作成しているか。
- 測ろうとしている言語能力を適切に評価できるような採点を行っているか。
- 採点基準に基づいた適切な証拠が得られるような問題となっているか。
- 測ろうとしている言語能力に対して、設問数のバランスの偏りがないか。
また、試験全体の構成概念のバランスと合致しているか。
- 測ろうとしている言語能力と直接関連しない設問や内容が含まれていないか。

<試験実施に関する妥当性>

- 受験者の身体的、心理的、経験的な特徴に配慮した試験になっているか。
- 試験内容が受験者間の公平性を保つものになっているか。
- 試験内容や能力基準に関する十分な情報が受験者に与えられているか。
- 試験の指示や説明が問題の意図を十分に伝えるものとなっているか。
- 試験解答のプロセスが測定意図と照らし合わせて適切なものとなっているか。
- 特別な支援を必要とする受験者に対する配慮がなされているか。
- 試験内容についての情報、受験者情報の管理が適切に行われているか。

⁹ 各要素の説明については野口・大隅（2014）pp.11-23 の説明をもとに一部修正

③ 真正性 (authenticity)

試験の問題項目がその試験で測定しようとしている目標言語使用領域における現実の課題をどの程度反映しているかの度合いをいう。例えば、読解力を測定する場合には、現実社会で実際に使用された文書を問題文とすることが望ましい。ただし、外国語学習の初級レベルでは既習の語彙・文法・漢字・言いまわしなどに配慮した文章を新たに書き起こすこともある。また、試験の問題項目は、現実の言語使用を反映していることが望ましい。

- 設問はどれくらい現実の言語使用場面を反映しているか。
- 受験者の言語能力が一定レベル以上であることを想定している試験においては、現実社会で実際に使用されている文書を活用できているか。
- 言語活動の場面やタスクにおいては、現実の言語使用を反映しているか。

④ 信頼性 (reliability)

テストの測定精度を表わす概念で、そのテストの測定結果、すなわち、得点に含まれる測定誤差が小さいほど、そのテストの信頼性が高いという。一般に、測定の標準誤差や信頼性係数で表わされる。

- 熟達度が変わらないと考えられる期間内に何度実施しても同じ結果を得ることができるか。
- 適切な統計手法を用いて「内的一貫性」、「項目弁別力」等についての検証を行っているか。
- 対面式の会話試験などで受験者のパフォーマンス能力を測る場合、試験官の間での評価のばらつきがないか、また、試験官のパーソナリティや印象が受験者のパフォーマンスに影響を与えていないか。
- パフォーマンス能力の測定において、試験官内での評価のばらつきがないかの検証を行っているか。
- 試験問題の項目困難度のバランスが想定される受験者の能力帯と合っているか。

⑤ 実行可能性 (feasibility)

テスト開発及びその後の安定的かつ継続的運営を可能にするのに関わる人的・経済的資源に関連し、テストを物理的・経済的に成り立たせるための前提条件のことをいう。具体的には、問題開発、テスト実施（採点、結果の通知）、データの分析などテストの実施及び結果の検証などに関する一連の流れを継続的に実行可能かということの問題にする。

- 設問作成に時間がかかりすぎていないか。必要以上の時間がかけられていないか。
- どれくらいの作業、労力で実施可能か。十分な人員を配置できているか。
- 採点のための適切な時間を確保できるか。
- 上記を勘案して、年間を通して安定的に実施できる回数で実施しているか。
- 受験者の受験環境が過度に負担をかけるものとなっていないか。

⑥ 波及効果 (washback effect)

試験の内容が受験者や教師、教育機関、企業、それら関係者を含む社会に与える影響のことを言う。例えば、外国語教育機関が外国語試験の出題傾向に合わせて学習内容やカリキュラムを決めるなど。

- 教育機関のカリキュラム改善に役に立っているか。
- 受験者に学習方法の改善を促すフィードバックを与えることができているか。
- 共通参照枠に照らし合わせるなど、透明性の高いフィードバックを与えているか。
- 試験結果を解釈する十分な情報が与えられ、現実のコミュニケーション場面での言語使用との対応を示しているか。
- 全体的なレベルだけでなく、個々の言語コミュニケーション活動や言語コミュニケーション能力に関する学習者のプロフィールを示しているか。
- 初級レベル等での学習上の配慮など、真正性から逸脱する点について、受験者に情報が与えられているか。

(2) 社会的ニーズに応える日本語能力判定の在り方について

① 日本語能力の判定のための試験の開発促進

- 日本語による言語活動のうち、「読むこと」、「聞くこと」のテストは多く存在するが、「話すこと（やり取り）」、「話すこと（発表）」、「書くこと」の言語能力を測定するテストの開発・普及が求められる。
- 幅広い受験ニーズ及び能力判定の需要に応えるため、従来型の紙ベースによる試験のほかに、オンラインによる受験が可能となる CBT（Computer-Based Test）による試験の実施による受験機会の拡大が求められる。
- 専門的な内容を測定する目的の試験開発のためには、当該分野の言語使用状況の調査・分析が必要となる。例えば介護などで必要とされる職務の分析と並行して、その職務を遂行するために必要な日本語の知識・内容に関する実証的な検討を行い、言語能力記述文を策定していくという手順を踏むことが適当である。
- 社会的なニーズに応えられる日本語能力を測定し、それに基づいて能力レベルを判定する試験を開発し安定的に実施していくためには、試験開発に関する専門人材の育成が不可欠である。

② 試験実施機関に求められる要素

- 日本語の大規模テストにおいては、試験実施機関は試験の実施に向けて一層の対策が求められる。不正防止のための対応策や適切な検証を行った上で例えば信頼性係数の推定値などテストの性能に関する検証結果の公表を義務付けたり、検証結果を第三者が確認したりするなどの対応が必要である。
- 試験実施団体は、受験希望者が試験を選ぶ上で必要となる情報を公表するよう努める必要がある。また、適格性を有する試験を選定する際には、試験の信頼性及び妥当性に関する根拠資料の提出を求め、試験・評価の専門家の意見を踏まえる必要がある。
- 「日本語教育の参照枠」に基づき、今後、日本語能力の評価が必要となる外国人材の活動や業種等による分野別の日本語能力の評価が行われるようになると想定されるが、試験実施機関側には、試験により測定・評価・判定する日本語の分野やレベルを社会に広く明示することが必要である。
- 「話すこと」、「聞くこと」などの言語活動別に求められる能力レベルが示されることにより、「話すこと」、「聞くこと」に関する評価手法や試験が日本語学習及び日本語の能力レベル判定に有効に活用されることが望まれる。

- 新規に開発された試験では、試行試験の結果を公表し、専門家からの評価を仰ぐことが必要である。ただし、試験の仕様や機密に触れる事項で継続的に試験を実施することの障りとなることはこの限りではない。この場合でも、継続的に実施された試験に関して試験団体の中に言語テストの専門家がスタッフとして存在して、言語テスト理論に基づく分析結果を試験の企画・作題・実施等の部門と共有していることが要請される。

③日本語能力判定の有効な活用に向けて

- C E F R レベルと試験との対応付けの方法を示すだけでなく、レベル尺度を体験したり、スタンダードセッティングのワークショップを開催したりするなど、活用に向けた研修機会の確保や評価担当者の育成に対する支援も必要である。
- 試験による日本語能力判定だけでなく、日頃の教室活動の中で学習者の日本語能力の進歩やレベルの推定ができるような手法とその事例を示していくことで、多様な評価の在り方を周知していくことが重要である。

参考文献

- 国際交流基金 (2016) 「JF 日本語教育スタンダードに準拠ロールプレイテストテスター用マニュアル」 <https://jfstandard.jp/roleplay/ja/render.do> (令和 2 年 8 月 25 日閲覧)
- 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 (令和 3 年 1 月 25 日閲覧)
- 豊田市 (2011) 『とよた日本語学習支援ガイドライン』
- 日本英語検定協会 (2018) 「英検ライティングスコアと CEFR レベル対応付け調査報告書」 <https://www.eiken.or.jp/eiken/group/result/pdf/eiken-score-cefr.pdf> (令和 2 年 8 月 25 日閲覧)
- 野口裕之・大隅敦子 (2014) 『テストニングの基礎理論』 研究社
- 文化審議会国語分科会 (2012) 「「生活者としての外国人」に対する日本語教育における日本語能力評価について」 https://www.bunka.go.jp/seisaku/kokugo_nihongo/kyoiku/nihongo_curriculum/index_4.html (令和 2 年 8 月 25 日閲覧)
- Center for Entrance Examination Standardization(2018) 「GTEC スコアと CEFR J レベル関連付け調査報告」 <https://www.benesse.co.jp/gtec/schoolofficials/research/pdf/doc-2018-01.pdf> (令和 2 年 8 月 25 日閲覧)
- Council of Europe (2001) Common European Framework of Reference for Languages: Learning, teaching, assessment. Cambridge University Press. (吉島茂・大橋理枝 訳・編 (2014) 「外国語の教育Ⅱ 外国語の学習, 教授, 評価のためのヨーロッパ共通参照枠 (追補版)」 朝日出版社) <https://www.goethe.de/ins/jp/ja/spr/unt/kum/ger.html> (令和 2 年 8 月 25 日閲覧)
- Council of Europe (2009) Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR) <http://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=0900001680667a2d> (令和 2 年 8 月 25 日閲覧)
- Council of Europe (2011) Relating language examinations to the Common European Framework of Reference for Languages: Learning, teaching, assessment (CEFR) Highlights from the Manual. https://www.ecml.at/Portals/1/documents/ECMLresources/2011_10_10_relex_E_web.pdf?ver=2018-03-21-100940-823 (令和 2 年 8 月 25 日閲覧)
- Council of Europe (2018) CEFR Companion Volume with New Descriptors <https://rm.coe.int/cefr-companion-volume-with-new-descriptors-2018/1680787989> (令和 2 年 8 月 25 日閲覧)
- Dunlea, J. (2009,2010) 「英検と CEFR との関連性について研究プロジェクト報告」 https://www.eiken.or.jp/center_for_research/pdf/market/report_02.pdf (令和 2 年 8 月 25 日閲覧)

参考資料 1 : 評価の種類

CEFR (2001) では、評価の種類について以下の項目を挙げている。表 1 は網羅的なものではなく、ある用語が右に置かれるか、左に置かれるのかも重要ではないとしている。

表 1 評価の種類¹⁰

1	到達度評価 Achievement assessment	熟達度評価 Proficiency assessment
2	基準準拠型評価 (NR) Norm-referencing	規準準拠型評価 (CR) Criterion referencing
3	合否型基準準拠型評価 Mastery leaning CR	連続型基準準拠型評価 Continuum CR
4	継続的評価 Continuous assessment	定点評価 Fixed point assessment
5	形成的評価 Formative assessment	総括的評価 Summative assessment
6	直接評価 Direct assessment	間接評価 Indirect assessment
7	運用評価 Performance assessment	知識評価 Knowledge assessment
8	主観的評価 Subjective assessment	客観的評価 Objective assessment
9	チェックリスト評定 Checklist rating	パフォーマンス評定 Performance rating
10	印象評価 Impression judgment	指針に基づいた判断 Guided judgment
11	全体的評価 Holistic assessment	分析的評価 Analytic assessment
12	シリーズ評価 Series assessment	分野別評価 Category assessment
13	他者評価 Assessment by others	自己評価 Self assessment

1 到達度評価 / 熟達度評価

到達度評価は、特定の目的の達成の度合いを評価し、学習したことを評価する。それゆえ、当該の週や学期に行った 勉強、教科書、シラバスに関連する。到達度評価は各々の授業に基づいた、内部の見方を反映している。つまり、限られた学習範囲の目標（試験範囲）にどの程度到達したかを見るということである。

熟達度評価は、実世界の問題に対して、学習者が、何ができるか何を知っているかの評価である。これは外部からの見方を反映している。

教師は教育へのフィードバックを得ようとして到達度評価の方に自然に関心を向けがちである。雇用者、教育行政の管理者、大人の学習者は、熟達度評価、つまり、成果や何ができるようになったかの方に、より関心があるだろう。到達度評価の利点は学習者の経験との差が少ないことである。熟達度評価の利点は誰でもその学習者のいる位置が分かることであり、結果が明確なことである。

¹⁰ 表の翻訳と各項目の説明は、吉島茂・大橋理枝 訳・編 (2014) pp.205-214 を抜粋し、一部修正した。

2 基準準拠型評価 / 規準準拠型評価

基準準拠とは、学習者を例えば得点順位に並べ、一緒に学習している他の学習者との相対的な位置を明らかにする評価である。規準準拠は、周りの学習者の力量とは無関係に、その教科の学習者本人の力量だけを純粹に評価するもので、基準準拠に対置されるものでもある。基準準拠はクラスという範囲の中で行うこと（あなたはクラスで 18 番目です）も、人口統計的な同類集団（あなたは 21,567 番目です、上位 14%にいます）など、あるテストを受けた学習者グループの範囲の中で行われることもある。

規準準拠は、個人個人のテスト結果が規準表全体のどこに位置しているか分かるようにするもので、熟達度を垂直軸に、関連領域を水平軸にとって図示するものである。これには（a）それぞれのテスト / モジュールがカバーしている関連領域の定義、および（b）「区切り点」、つまり特定の熟達度の水準に達していると認定できるテストの点の特定が必要となる。

3 合否型基準準拠型評価 / 連続型基準準拠型評価

合否型基準準拠は、単一の「最低限の能力の標準」や「区切り点」を決めて、学習者を「合格者」「不合格者」に分ける仕組みである。このやり方では、学習目標達成の度合いは問わない。

連続型基準準拠は、当該の分野において、個々の力量が予め決められた連続体の中のどの位置にあるかを示すものである。

CEFR は合否型でも連続型でも利用できる。連続型で用いられたレベルの尺度は共通参照レベルに照合することができる。合否型で設定される目標は CEFR の提案したカテゴリーとレベルの概念表の中に位置づけることができる。

4 継続的評価 / 定点評価

継続的評価は教師、またある場合には学習者から見た授業コース全体を通しての授業中の言語運用、課題やプロジェクトの評価である。最終的な成績は授業コース / 学年 / 学期全体を反映する。

定点評価とは、ある特定の日、すなわち、普通は授業コースの最後か開始以前に行われる試験やその他の評価に基づいて、成績が与えられ判断が下されることである。以前にあったことは問題ではなく、その人が今できることが決定的に重要なのである。

評価は、何らかの判断を下すために一定の点でとり行われる授業コースの外のものとみなされることが多い。継続的評価は、授業コースの中に組み込まれており、その授業の終了時に何らかの総合的なやり方で評価する時に使われる。宿題や、学習の強化のための定期的達成度テストとは別に、継続的評価は教師や学習者によるチェックリストや表の形をとることもある。

5 形成的評価 / 総括的評価

形成的評価は、学習の進み具合や学習者の強み、弱点に関する情報を集める継続的な評価である。教師はこれらの情報を授業コースの計画や学習者へのフィードバックに役立てることができる。形成的評価という言葉は、広い意味で用いられることが多く、質問紙や話し合いから得られた数量化できない情報も含まれる。

総括的評価は授業コースの終わりにこれまでの成果を成績としてまとめるものである。それは必ずしも熟達度の評価ではない。事実、総括的評価の多くは標準準拠型の定点評価であり、達成度評価である。

6 直接評価 / 間接評価

直接評価は、学習者が実際にしていることを評価することである。例えば、小グループで何かを論じているところで、評価者がそれを観察して、基準となる表と比較し、言語運用を表の中の最も適当なカテゴリと一致させ評価を下す。間接評価は通常紙面テストを用い、実行可能だと考えられる技能を評価する。

直接評価は事実上、話すこと、書くこと、やり取りでの聞くことに限られる。というのも、受容的活動は直接見ることができないからである。例えば、読むことについては、学習者に適当な解答欄をチェックさせる、文を完成させる、質問に答えさせるなどして、理解の証を出させることで、間接的に評価するしかない。言語の使用の幅とその把握の程度は、基準との一致度を判断して直接的にも評価できるし、あるいはテスト問題の答えを解釈し、一般化することで間接的にも評価できる。古典的な直接テストは面接であり、古典的な間接テストはクローズ・テスト (cloze test) である。

7 運用評価 / 知識評価

運用評価を行うためには、学習者が実際の発話か、書いた文書の実例を提示しなければならないが、それらは直接テストによって得られる。

知識評価では、学習者がさまざまな種類の質問に答えることになるが、その質問は、学習者がどの程度言語的な知識を持っており、その使い方をどの程度把握しているかを証明するものでなければならない。

能力を直接に測ることはできない。それを測ろうとする場合、運用の幅から熟達度についての一般化を行うしかない。熟達度というのは、実際に使用された能力と考えてよいだろう。この意味であらゆるテストは、運用例を証拠としてその根底にある能力を推定しようとするのだが、全てのテストが測っているのは実際には運用のみである。

8 主観的評価 / 客観的評価

主観的評価は、評価者の判断によって決められる。普通、これは運用の質に対する判断のことである。

客観的評価は、主観性を排した評価である。普通、これは例えば選択肢式のテストのように、各項目に該当する正解が一つしかないような間接テストを意味する。

しかし、主観性 / 客観性の問題は、これよりはるかに複雑である。間接テストは、しばしば「客観式テスト」とされているが、これは採点者が確定的な正解をもとに、受験者の回答を正しいとするか誤りとするかを決め、正しいとした回答の数を数えて最終的な結果を出す類のものを指している。この過程をもう一段階先に進め、それぞれの質問に正解が一つだけしかないようにし、採点者の誤りを防ぐために機械で採点することもしばしばである。実際、この意味での「客観式テスト」の客観性は、多少強調され過ぎているくらいがある。というのは、誰かが、評価という行為をより制御しやすいテストの実施技術に還元するということを決定したからである。そして、誰かがテストの細目を書き、他の誰かが、その細目の中の特定の項目を実際に測定可能な形に書いたのかもしれない。結局は、誰かが、出題される可能性のある他の項目ではなく、その特定の項目をこのテストのために選んだのである。これらの決定は全て何らかの主観的な判断を伴うので、このような形式のテストは客観採点式テストと呼んだ方がよいのではないだろうか。

直接運用評価では、たいてい評価者の判断を基に成績がつけられる。すなわち、学習者がどの程度上手に言語を運用したかという判断は主観的に行われるのであり、関連する諸事情を考慮に入れたり、ガイドラインや基準や経験に照らしたりして決められる。言語やコミュニケーションは非常に複雑であり、自動化に馴染まず全体が個々の部分の総和よりも大きいことから、主観的な方法には利点がある。特定のテスト項目が実際には何をテストしているのかを明確にすることは困難な場合が多い。従って、能力や運用力の特定の面に焦点を当てたテスト項目というのは、表面的にはともかく、実際はそううまくは機能しない。

しかし、公平を期すためには、すべての評価はできる限り客観的でなければならない。内容の選択や言語運用の質に関する主観的な決定に、個人的な価値判断が影響することは可能な限り減らさなければならないし、特に総括的評価が行われる場合には尚更である。というのは、テストの結果は、その評価を受けた人の将来を決めるために第三者によって使われることが多いからである。

9 チェックリスト査定 / 尺度評定

尺度査定：いくつかのレベル、あるいはレベル帯から構成される尺度に基づいて、学習者が特定のレベルにある、または特定のレベル帯の範囲内にいることを判断する。

チェックリスト査定：特定のレベルやモジュールに関連があるとされる事項のリストに照らして、学習者に対する判断を下す。

「尺度査定」では、学習者をいくつかあるレベル帯のどれかに当てはめることに主眼が

置かれる。強調されるのは垂直性であり、学習者が尺度内のどの程度上の位置まできたのか、という点である。それぞれのレベルやレベル帯が意味するものは、その尺度の言語能力記述文によって明らかにされていなければならない。さまざまな分野に対して、それぞれ複数の尺度があるかもしれないし、それらは表として同じページに記載されていてもよいし、別のページに記載されていてもよい。それぞれのレベルやレベル帯について定義があるかもしれないし、一つおきにしか定義されていなくてもよい。もしくは、上、中、下のレベルに当たるものにしか定義がないこともありうる。

これに代わるものとして、チェックリストがある。これはそのリストに出ている項目に関連する分野を達成できたかどうかをチェックすることが主眼となる。つまり、水平性が強調されるのであり、そのモジュールの内容をどのくらい達成することができたかが重要になる。チェックリストは、質問紙のように、要点を列挙したような形で示すこともできる。一方、それは、車輪のような形で示すこともできるし、また別の形で示すことも可能である。答え方は、「はい / いいえ」だけかもしれないし、もっと細かい形（例えば「0から4」の段階付けで示されるなど）かもしれないが、その場合それぞれに表示がきちんと付いた目盛りがあり、その表示も定義されていることが望ましい。

言語能力記述文は、独立した、当該のレベルと対応した基準項目になっているので、これを基にして特定のレベルのチェックリストを作ることも、全てのレベルに関する査定尺度や表を作ることも両方可能である。

10 印象評価 / 指針に基づいた判断

印象：学習者の授業中の言語運用の経験に基づいて行われる完全に主観的な判断で、特定の評価に関して何の基準にも基づかない評価。

指針に基づいた判断：何らかの基準に基づいて、意図的に評価しようという意識を持って判断を行うことによって、印象のみによる判断を補い評価者の主観性が減じられる判断。

「印象」とは、教師や学習者が、授業中や宿題などの出来具合のみを基にすることによって、判断を下す場合のことを指している。主観的な査定、特に、継続的な評価で使われる査定は、反省や記憶を基にしている場合が多いが、その際に使われている反省や記憶の焦点は、対象者を一定の期間意識的に観察することによって定まってくる可能性が高い。非常に多くの学校でこの方法が実施されている。

「指針に基づいた判断」とは、上に述べたような印象判断が、一連の評価方法を通じて、判断が熟慮に基づいたものになった場合を指している。そのような方法は、(a)何らかの手順に従った評価が行われていること、及び/または(b)各評点または成績の間を区別できる明確な基準を設けていること、そして(c)標準化のために評価者が何らかの訓練を受けることを意味している。

指針に基づいて判断を行う利点は、このようにして評価する人たちの中で共通基準が確立できれば、下される判断の一貫性が劇的に増すことである。もし、言語運用の実例や他の評価方法との関連が固定的「水準点」として示されているならば、一貫性はさらに向上

する。この点の重要性が強調されるのは、さまざまな学問分野で、次のような研究結果が重ねて確認されていることによる。すなわち、評価者が充分訓練されていないまま判断を下した場合、評価者の厳しさの差が、学習者の実際の能力差と同じくらい大きくなり、学習者に対する評価結果がほとんど偶然で決まったも同然になりかねないということである。

共通参照レベルの尺度の言語能力記述文は、上の(b)のような明確化された基準を示すのに利用できるし、既存の基準によって表示されている標準値が、共通レベルのどの辺りに相当するのかを位置づけるのに使うこともできる。

11 全体的評価 / 分析的評価

全体的評価というのは、包括的で統合的な判断を下すことである。さまざまな評価側面の比重は評価者の直感によって定められる。分析的評価は評価側面の一つ一つを別個に見る。この区別には二通りの仕方がある。すなわち、(a) 何を評価するか、(b) どのようにしてレベル帯や級や得点が与えられるかである。ある部分では分析的な評価を行い、別の部分では全体的な評価を行うというように、組み合わせて評価が行われるような方法が採られることもある。

(a) 評価対象：

「話すこと」や「言葉のやり取り」のように、包括的な分野に対して何か一つだけ得点や級を与えるように評価しようとする場合もある。他の、もっと分析的なやり方の中には、学習者の言語運用の中で、相互に独立したいくつかの側面ごとに別々に結果を出すことを評価者に要求するやり方もある。さらに、評価者が学習者に対する包括的な印象を記し、それぞれの分野別に分析的な評価を行い、その後熟慮して全体的な判断を下すようなやり方もある。

(b) 結果の算定：

学習者の言語運用を観察して、全体的な観点から尺度の言語能力記述文に当てはめるといいうやり方があるが、この場合尺度が全体的なもの（包括的な尺度を一つだけ用いるとき）である場合も、分析的なもの（3～6分野に分かれて表になったもの）である場合もある。このようなやり方は、結果を算定するのに計算を用いない。結果は一つの数値で表されるか、複数の分野にわたって「電話番号」型に羅列したものとなる。他の、さらに分析的なやり方としては、分野別に何らかの評点を与え、それらを合計してその学習者の評価値とし、さらに場合によってはその評価値を成績に変換するというやり方がある。この方法の場合、分野別に配点比重を変えて計算するというのが典型的なやり方である。つまり、さまざまな分野がそれぞれ同等の価値を持っているとは見なされないということである。

12 シリーズ評価 / 分野別評価

分野別評価は、単独の評価課題で学習者の言語運用を評価基準表に照らして判断するやり方である。

シリーズ評価は、相互に関連性のない複数の評価課題を行い、それを一つの全体的な査定結果として、それぞれの段階の内容が示された尺度、例えば0～3や1～4などで、表現するものである（この場合の評価課題は、他の学習者や教師とのロールプレイという形を取ることが多い）。

シリーズ評価は、分野別評価では一つの分野の評価が他の分野の評価に影響するという傾向に対処する方法の一つである。初級レベルでは課題の達成に重点が置かれ、その学習者は何ができるのかということ、単なる印象からではなく実際の言語運用から教師・学習者が評価したものを基にして、チェックリストを埋めていくことが意図される。上級レベルでは、言語運用の中の特定の側面の熟達度を示すような課題が与えられるであろう。結果はその学習者の輪郭像として報告される。

13 他者評価 / 自己評価

他人による評価：教師または評価者による判断

自己評価：自分自身の熟達度の判断

上で述べた評価技術の多くには、学習者自身に関与することも可能である。今までになされた研究から、（例えば、ある授業コースに入れてもらえるかどうかなどの）「高い賭け金」がかかっていない限り、テストや教師による評価は自己評価によって効果的に補完できるといわれている。自己評価の正確さが増すのは以下の場合である。

- (a) 明確な基準をもった熟達度の言語能力記述文に基づいて評価が行われる。
- 及び / または、
- (b) 評価が具体的経験と関連する。

この経験そのものがテスト課題であってもよい。また、学習者が評価を行うために訓練を受ければ、自己評価はさらに正確なものとなるだろう。このように系統立てて行われる自己評価と教師による評価やテストとの相関係数（一致の妥当性のレベルの指標）は、教師が行った評価と評価の間や、テストとテストの間や、教師による評価とテストとの間で通常みられる相関係数と同じくらいに高いこともある。

しかし、自己評価の最大の可能性は、それを学習者の動機付けや意識を高めることに使うことにある。学習者が自分の長所に気づき、弱点を認識し、学習の方向付けをさらに効果的なものにする手助けをすることである。

参考資料2：日本語の能力判定に係る試験等一覧

1. 試験の概要

(1)能力評価の名称	(2)実施機関・団体	(3)目的	(4)開始年	(5)能力評価の主な対象	(6)年間受験者等数	(7)合格者/レベル認定者数	(8)年間実施回数・実施時期(国内)	(9)実施地(国内・海外)	(10)テスト方式・PC利用等	(11)評価を行う言語能力・活動等	(12)評価を行うレベル	(13)試験科目と時間	(14)解答方法・方式	(15)採点・判定方法	(16)結果通知方法及び通知に要する時間	(17)受験費用	(18)試験問題の公表	(19)URL
日本語能力試験(JLPT)	独立行政法人国際交流基金と公益財団法人日本国際教育支援協会の共催 ※海外では国際交流基金が各地機関の協力を得て実施。(台湾では公益財団法人日本台湾交流協会と共催で実施。) ※国内では日本国際教育支援協会が実施	日本語能力を測定し、認定することを目的とする。	1984年	原則として日本語を母語としない人	【国内】439,085人 【海外】729,450人(2019年)	全レベル合計：408,400人	2回(7月,12月)	【国内】47都道府県 【海外】87カ国・地域	☑紙媒体	☑読むこと(読解) ☑聞くこと(聴解) ☐書くこと(記述) ☐話すこと(発表) ☐話すこと(やりとり) ☑その他(言語知識(文字・語彙、文法))	N1 N2 N3 N4 N5	<N1> 言語知識(文字・語彙・文法)・読解：110分 聴解：60分 <N2> 言語知識(文字・語彙・文法)・読解：105分 聴解：50分 <N3> 言語知識(文字・語彙)：30分 言語知識(文法)・読解：70分 聴解：40分 <N4> 言語知識(文字・語彙)：25分 言語知識(文法)・読解：55分 聴解：35分 <N5> 言語知識(文字・語彙)：20分 言語知識(文法)・読解：40分 聴解：30分	☑多肢選択形式	・マークシート式回答を機械で読み込み採点 ・日本語教育、心理統計分野の専門家が得点および試験全体の動向を確認 ・IRTに基づいてスコアを算出	☑本人 ・日本国内受験者：受験者全員に「合否結果通知書」、さらに合格者には「日本語能力認定書」を送付 ・海外受験者：受験者全員に「認定結果及び成績に関する証明書」、さらに合格者には「日本語能力認定書」を交付	【国内】5,500円 【海外】実施都市によって異なる	☑無	https://www.jlpt.jp/
JPT日本語能力試験	一般社団法人日本語能力試験実施委員会	ビジネスシーンや日常的な場面、状況において、高度で機能的なコミュニケーション能力を客観的に測定、評価することを目的としている。	1985年	日本語を母語としないビジネス関係者と日本語学習者	【国内】1,102人 【海外】32,481人(2020年)	テスト結果は合格・不合格ではなく、聴解5点～495点、読解5点～495点、合計10点～990点のスコアで5点刻みで表示される。	・実施回数：年間12回 ・実施時期：毎月一回	【国内】9都市 【海外】6か国	☑紙媒体 ☑PC利用(会場受験)	☑読むこと(読解) ☑聞くこと(聴解) ☐書くこと(記述) ☐話すこと(発表) ☐話すこと(やりとり) ☐その他	・レベル評価でなく点数制(10点～990点)方式 ・点数区間別(880/740/610/460/220点以上)聴解・読解スキルの目安の案内	聴解 45分 読解 50分	☑多肢選択形式	・解答の読み込み * CBT:受験者の解答をデータ保存 * PBT:マークシート式答案用紙を機械でスキャンして保存 ・「得点等化方式(百分位 Conversion table)」により点数を算出	☑本人 ・結果発表は受験8日～10日後、JPT APP(アプリ)にて通知 ・JPT APPにID(メールアドレス)とパスワード入力ログイン後、成績結果を確認	【国内】 ・費用：6,000円 ・試験方式：紙媒体、PC利用(CBT) 【海外】 ・費用：実施都市によって異なる ・試験方式：PC利用(CBT)	☑有(不定期)	https://www.jptest.jp/
ACTFL-OPI	ACTFL(全米外国語教育協会)／Language Testing International ※非公式なテスト(Advisory OPI)については所属機関にテスト資格保持者がいれば可能 ※テストの団体：日本OPI研究会	ACTFLによって開発された汎言語的に使える会話能力テスト。「OPI」とは、oral proficiency interview(オーラル・プロフィシエンシー・インタビュー)の頭文字で、外国語の口頭運用能力を測定するためのインタビューテスト。	1986年	言語学習者	非公開	非公開	随時	国内・海外問わず	☑電話 ☑対面 ※正式なテストは電話で行われる。Advisory OPIは対面等の方法が用いられる。	☐読むこと(読解) ☐聞くこと(聴解) ☐書くこと(記述) ☐話すこと(発表) ☐話すこと(やりとり) ☐その他	初級-下～超級	インタビュー 約20～30分	☑口述式	・終了後にOPIを行ったテストターが録音を聞き直し、ガイドラインに照らしながら、被験者の口頭運用能力がどのレベルにあるかを判定 ・正式なACTFLとしての判定結果を出すためには、さらに第2テストターが同じテープを聞き、その結果が第1テストターと一致したとき、初めて判定を決定 ・第1テストターと第2テストターとの判定が異なったときは、第3テストターが判定に加わる	☑本人	134ドル	☑無	日本語OPI研究会(テストターの団体) http://opi.jp/

(1)能力評価の名称	(2)実施機関・団体	(3)目的	(4)開始年	(5)能力評価の主な対象	(6)年間受験者等数	(7)合格者/レベル認定者数	(8)年間実施回数・実施時期(国内)	(9)実施地(国内・海外)	(10)テスト方式・PC利用等	(11)評価を行う言語能力・活動等	(12)評価を行うレベル	(13)試験科目と時間	(14)解答方法・方式	(15)採点・判定方法	(16)結果通知方法及び通知に要する時間	(17)受験費用	(18)試験問題の公表	(19)URL
日本語NAT-TEST	・ 専門教育出版「日本語NAT-TEST運営委員会」 ・ 海外での試験は、各国の運営本部が同委員会の指導・監督の下、厳正に実施	日本語能力を判定することを目的とする。	1989年(2007年に現在の名称に変更)	日本語を母語としない日本語学習者	69,947人(2020年) 164,179人(2019年) 114,168人(2018年)	30,392人(2020年)	6回以上(2月、4月、6月、8月、10月、12月)	【国内】2都市 【海外】16か国	☑紙媒体	☑読むこと(読解) ☑聞くこと(聴解) ☐書くこと(記述) ☐話すこと(発表) ☐話すこと(やりとり) ☑その他(言語知識(文字・語彙・文法))	初級から上級までの5レベル(1Q、2Q、3Q、4Q、5Q)	●1Q 言語知識(文字・語彙・文法)・読解(110分) 聴解(約60分) ●2Q 言語知識(文字・語彙・文法)・読解(105分) 聴解(約50分) ●3Q 言語知識(文字・語彙)(30分) 言語知識(文法)・読解(70分) 聴解(約40分) ●4Q 言語知識(文字・語彙)(30分) 言語知識(文法)・読解(60分) 聴解(約35分) ●5Q 言語知識(文字・語彙)(25分) 言語知識(文法)・読解(50分) 聴解(約30分)	☑多肢選択形式	・ マークシート式回答を機械で読み込み採点 ・ 合格点は全体の解答を統計解析して算出(回ごとに変動) ・ 合格の条件は「総得点が合格点を上回っていること」と「各分野の得点が配点の25%以上であること」を同時に満たしていること	☑本人 試験日から3週間以内にサイトで合格発表、同時に成績表・合格証を郵送 ☑機関 合格発表日以降、成績証明を照会可能、公知の住所に郵送	【国内】5,500円 【海外】実施都市によって異なる	☑無 ※過去問題を収録した問題集を刊行	http://www.nat-test.com/
J.TEST 実用日本語検定	株式会社語文研究社「日本語検定協会・J.TEST事務局」	日本語能力を測定することを目的とする。	1991年	日本語を母語としない日本語学習者	73,157人【国内】29,050人【海外】44,107人 (2020年5月中止分を除く過去6回分)	●A-CLレベル試験 21,121人中 「特A級」83人「A級」109人「準A級」362人 「B級」613人「準B級」2,263人 「C級」3,802人 認定なし13,889人 ●D-Eレベル試験 36,224人中 「D」級6,075人 「E」級13,369人 認定なし16,780人 ●F-Gレベル試験 15,812人中 「F」級14,708人 「G」級14,890人 認定なし6,214人	6回(1月、3月、5月、7月、9月、11月)	【国内】15都市 2021年3月より24都市 【海外】12か国・地域 2021年3月より13か国・地域	☑紙媒体	☑読むこと(読解) ☑聞くこと(聴解) ☑書くこと(記述) ☐話すこと(発表) ☐話すこと(やりとり) ☑その他(言語知識:文法・語彙・漢字力)	●A-CLレベル試験 1000点満点 特A級 930点以上 A級 900点以上(CEFR C2) 準A級 850点以上 B級 800点以上 準B級 700点以上(CEFR C1) C級 600点以上(CEFR B2) ●D-Eレベル試験 700点満点 D級 500点以上(CEFR B1) E級 350点以上(CEFR A2) ●F-Gレベル試験 350点満点 F級 250点以上(CEFR A1) G級 175点以上	☑多肢選択形式 ☑記述式(A-C,D-Eのみ)	読解・聴解試験 ・ マークシート式回答を機械で読み込み採点 記述試験 ・ 記述式回答を複数の採点担当者が採点	☑本人 ☑機関 ・ 試験日から約4週間後に郵送 ・ レベル認定者には認定証も同封 (試験実施後約3週間でウェブサイト検索可能)	【国内】4,800円 ※日本の学校に在籍する中学生及び高校生は受験料免除 * 2021年5月より5,200円 ※日本の学校に在籍する中学生及び高校生は受験料50%免除 【海外】国・地域によって異なる。	☑有(年1回以上) ☑有(持ち帰り可)	https://j-test.jp/n-ewjtest https://j-test.jp/cefr-jtest	
ACTFL及びLTIの習熟度試験(日本語版)	ACTFL(全米外国語教育協会) & Language Testing International (LTI) & iJapan株式会社(日本におけるLTIの正式な代理店)	各テストは、個人の特定の習熟度レベルを適切に判断し、最終的には有効かつ正当な言語運用能力を測定するように設計されている。	1992年	日本語を話す者を対象に日本語の口頭運用能力を測定します。	オンライン受験の場合、受験者の場所は特定不可のため、全世界の年間受験者数を提供: 約90万人(OPI&OPIc & LPT&WPT & RPT)	合格や不合格等の判定を行っていない	随時	OPIテスト(電話)利用可能時間は、月曜から金曜の午前8時から午後10時まで(米国東部標準時)。 OPIc / WPT / LPT / RPT: コンピュータベースのテスト(CBT)のため、いつでもどこでもネット環境があれば受験可; また、認定されたテストセンターでの受験も可能	☑PC利用(在宅受験)(OPIc / WPT / LPT / RPT) ☑電話(OPI)	☑読むこと(読解: RPT) ☑聞くこと(聴解: LPT) ☑書くこと(記述: WPT) ☑話すこと(やりとり: OPI & OPIc)	For Speaking Tests (OPIc or OPI) ACTFL Rating Scale: CEFR Rating Scale: Superior ... C2 Advanced High ... C1 Advanced Mid ... B2.2 Advanced Low ... B2.1 Intermediate High ... B1.2 Intermediate Mid ... B1.1 Intermediate Low ... A2 Novice High ... A1 Novice Mid ... 0 Novice Low ... 0	☑多肢選択形式 (LPT, RPT) ☑口述式(OPI, OPIc) ☑記述式(WPT)	受験者の言語運用能力を基準に基づいて評価。 スピーキングテスト(OPIおよびOPIc)は、ACTFL認定のテスターによって評価。 テストが完了すると、テスト内容はLTIのACTFLテストボックスに送信され、レビュー、スコアリングされる。 (WPT、LPT及びRPTの情報について、版元と確認中)	☑本人 ☑機関 PDF証明書は3~5日(営業日)でメールで送付される。 24~36営業時間以内にクライアントのパスワードで保護されたサイトに投稿される。会社、機関、または機関は、受験者に結果を通知する方法がある。バイリンガルの資格があるかどうかを判断する。	OPI: 都度要相談 OPIc(Eng): JPY 7500 OPIc(Other): JPY 9500 WPT(Eng): JPY 7500 WPT(Other): JPY 9500 LPT: JPY 5800 RPT: JPY 5800	☑無	https://www.languagestesting.com/lti-information/gen-eral-test-descriptions	

BJTビジネス日本語能力テスト	公益財団法人日本漢字能力検定協会	ビジネス場面で必要とされる日本語コミュニケーション能力を測定することを目的とする。	1996年	日本語を母語としないビジネス関係者、日本のビジネス社会で働くことを目指している学生	志願者数： 【国内】3,326人 【海外】1,716人 (2019年)	J1+ : 249人 J1 : 568人 J2 : 2,150人 J3 : 1,317人 J4 : 367人 J5 : 88人	随時	【国内】33都市 【海外】17か国と1地域 (2020年11月27日現在)	☑ P C 利用 (会場受験)	☑読むこと (読解) ☑聞くこと (聴解) ☑書くこと (記述) ☑話すこと (発表) ☑話すこと (やりとり) ☑その他 (言語の基礎力：語彙・文法、敬語・待遇表現)	J1+ : どのようなビジネス場面でも日本語による十分なコミュニケーション能力がある (600~800点) J1 : 幅広いビジネス場面で日本語による適切なコミュニケーション能力がある (530~599点) J2 : 限られたビジネス場面で日本語による適切なコミュニケーション能力がある (420~529点) J3 : 限られたビジネス場面で日本語によるある程度のコミュニケーション能力がある (320~419点) J4 : 限られたビジネス場面で日本語による最低限のコミュニケーション能力がある (200~319点) J5 : 日本語によるビジネスコミュニケーション能力はほとんどない (0~199点)	聴解：約45分 聴読解：約30分 読解：30分 計：約105分	☑多肢選択形式	・受験者がP C画面上で入力した解答を即時自動採点 ・I R Tに基づいた統計処理によりスコアを算出	☑本人 ・受験後すぐにスコアレポートを交付 ・ピアソンVUEのウェブサイトを受験結果を確認でき、翌日以降に成績認定書を印刷可能	【国内】7,000円 (税込) ・公的支援： ①留学生就職支援事業における受験料の全額・一部補助 ②文部科学省「留学生就職促進事業」における学習支援 【海外】 実施都市によって異なる	☑無	https://www.kanken.or.jp/bjt/
日本留学試験(EJU)	独立行政法人日本学生支援機構	日本の大学等で必要とする日本語力 (アカデミック・ジャパニーズ) 及び基礎学力の評価を行うことを目的とする。	2002年	外国人留学生として、日本の大学 (学部) 等に入学を希望する者	【国内】46,744人 【海外】12,587人 (2019年度) 第1回 【国内】25,237人 【海外】6,682人 第2回 【国内】21,507人 【海外】5,905人	合格・不合格はなくスコアのみ	2回 (6月・11月)	【国内】16都道府県 【海外】14か国(18都市)	☑紙媒体	☑読むこと (読解) ☑聞くこと (聴解) ☑書くこと (記述) ☑話すこと (発表) ☑話すこと (やりとり) ☑その他	レベルは設けていない	記述30分 読解40分 聴解・聴読解55分	☑多肢選択形式 ☑記述式	・読解、聴解・聴読解 マークシート式解答用紙を機械で読み込み採点し、I R Tに基づいてスコアを算出 ・記述 記述式解答用紙を採点基準に基づき採点	☑本人 試験実施後約6週間後の指定日に郵送にて通知 (オンライン出願者は同指定日よりオンライン上で確認が可能) ☑機関 成績利用校として登録している大学等は、所定の手続きにより同指定日よりオンラインで受験者の成績照会が可能	【国内】7,560円 (1科目のみの受験者) 14,040円 (2科目以上の受験者) 【海外】 実施都市によって異なる	☑有 (年1回以上)	https://www.jaso.go.jp/ryugaku/study/eju/index.html
標準ビジネス日本語テスト (STBJ)	一般社団法人応用日本語教育協会	一般的な日本語の語彙や表現だけではなく、ビジネスに求められる敬意表現やビジネスマナーに関する知識、実践的なコミュニケーション能力の測定を目的とする。	2007年	日本語を母語としない者 (日本企業や日系企業へ就職を希望する者、既に就職した者、日本への留学を希望する者)	【国内】151人 【海外】4,135人 (2019年)	BJ1~BJ5までのレベル判定のため、合否の判定はなし。	随時	【国内】随時 【海外】3か国	☑紙媒体	☑読むこと (読解) ☑聞くこと (聴解) ☑書くこと (記述) ☑話すこと (発表) ☑話すこと (やりとり) ☑その他 (言語知識：語彙)	・BJ1 (800-1000点) CEFR C1相当 ・BJ2 (650-799点) CEFR B2相当 ・BJ3 (450-649点) CEFR B1相当 ・BJ4 (250-449点) CEFR A2相当 ・BJ5 (0-249点)	テスト形式： テスト1：聴解能力テスト (1) テスト2：聴解能力テスト (2) テスト3：発話能力テスト テスト4：聴読解能力テスト (2) テスト5：聴読解能力テスト (2) テスト6：読解能力テスト テスト7：語彙能力テスト 試験時間 (全体)：約85分 ※聴解問題があるため、問題により若干の誤差がある。	☑多肢選択形式	マークシート式回答を機械で読み込み採点	☑本人 受験後約10日後に、各国外務省ウェブサイト上で、受験番号・パスワードにより結果を照会、または郵送にて通知。 ☑機関 「法務省が告示した日本語教育機関」「日本語教育振興協会認定校 (維持会員でなくても可)」及び「大学、専門学校等の学校法人」に書類を発行。試験日の原則3週間後から発行可能。	【国内】インハウス試験 4,500円 準会場試験 現在検討中 【海外】実施都市によって異なる	☑無	https://www.ajiea.net/stbj/
J-CAT 日本語テスト	一般社団法人日本語教育支援協会	日本語のeラーニング教材開発、ICT技術に基づくシステム開発を行い、インターネットを通して日本語教育機関に幅広く提供することにより、日本語・日本事情教育を発展させることを目的とする。	2008年	日本語を母語としない者	【国内】約20,000人 (2019年度) 2017年度：23,830人 2016年度：17,814人	該当しない	随時	実施地を限定していない。	☑ P C 利用 (会場受験) ☑ P C 利用 (在宅受験) ☑ C A T	☑読むこと (読解) ☑聞くこと (聴解) ☑書くこと (記述) ☑話すこと (発表) ☑話すこと (やりとり) ☑その他 (文法、語彙)	CEFR A 1 からC1までをワンスケールで評価	試験科目：語彙、文法、読解、聴解の4セクション 時間：全セクション通して実施。約45分~90分	☑多肢選択形式	・受験者がP C画面上で入力した解答を即時自動採点 ・I R Tに基づいてスコアを算出	☑機関 ・テスト終了と同時に得点が表示	3プラン別・ Challenge:1000円 Performance:2,000円 Certificate:4,000円	☑無	https://j-cat.jalesa.org/

アルクの電話による日本語会話テスト JSST	株式会社アルク	日本語会話力テスト。日本語に関する知識を「その時、その場」で組み合わせ、話を創造する力を測る。	2008年	日本語を母語としないビジネスパーソン	1,500人 (2019年)	10段階評価のレベル判定のため、合否無し	随時	実施地を限定していない。	☑電話	☐読むこと（読解） ☑聞くこと（聴解） ☐書くこと（記述） ☑話すこと（発表） ☑話すこと（やりとり） ☑その他（言語知識：文法、語彙、発音、流暢さ、単語）	レベル1～10の10段階評価 レベル1～3：初級 レベル4～8：中級 レベル9～10：上級	10問の質問をランダムに出題。 所要時間約15分	☑口述式	録音された会話を複数の評価担当者が確認して評価	☑機関 ・申込者にメール通知 受験期間終了後10営業日後にウェブサイトにて公開	・コメントシート付 税抜5,500円 ・コメントシートなし(スコアレポートのみ) 税抜4,500円	☑無	https://www.alc.co.jp/jsst/
TOPJ実用日本語運用能力試験	一般財団法人アジア国際交流奨学財団	日本語を外国語として勉強している学習者には、語彙や文法、文型などの基礎能力の到達度を測定することはもとより、日本に滞在している外国人の生活や仕事などにおけるコミュニケーションの中での疑問点を通して、日本社会の中での習慣と文化に対する理解能力を測定することも目的としている。	2008年	・日本語を母語としない外国人対象 ・日本語を外国語として勉強している学習者 ・日本に滞在している外国人	25,145人 (2019年) 内訳 【国内】 299人 【海外】 24,846人	初級：23,632人 A-4：2,227人 A-5：10,935人 その他：10,470人 中級：1,239人 A：28人 B：81人 C：262人 その他：868人 上級：274人 A：13人 B：35人 C：63人 その他：153人	6回 奇数月	【国内】 5都市 【海外】 6か国	☑紙媒体	☑読むこと（読解） ☑聞くこと（聴解） ☐書くこと（記述） ☐話すこと（発表） ☐話すこと（やりとり） ☐その他	上級 聴解 35分 筆記 85分 初級・中級 聴解 35分 筆記 55分	☑多肢選択形式	マークシート式回答を機械で読み込み採点	☑本人 ☑機関 ・試験実施後、2週間以内に試験結果検索ページで、試験の決定級・得点の結果等を調べることができる。 ・試験実施後、1か月以内に個人に試験結果通知並びに認定書を送付。	【国内】 初級・中級・上級 5,000円 【海外】 実施都市によって異なる	☑無	https://www.topj-test.org	
とよた日本語能力判定 (対象者判定テスト)	豊田市国際まちづくり推進課 ※平成20年度から令和元年度までは名古屋大学に委託し、令和2年度からは（公財）豊田市国際交流協会に委託して実施	日本語の会話力、読み書き能力を測り、「とよた日本語学習支援システム」が運営する日本語教室での支援対象者を認定するため。	2008年	豊田市内に在住・在勤で、「とよた日本語学習支援システム」が運営する日本語教室の受講を希望する外国人住民	139人	139人	9回	【国内】 愛知県豊田市 ・企業内での日本語教室 ・地域での日本語教室	☑紙媒体 ☑電話 ☑対面	☑読むこと（読解） ☑聞くこと（聴解） ☑書くこと（記述） ☐話すこと（発表） ☑話すこと（やりとり） ☐その他	(レベル) 6 熟達段階 5 深化段階 4 拡大段階 3 自立段階 2 要支援段階 1 基礎段階 0 未学習段階 ※「とよた日本語学習支援システム」では0,1レベルの人を対象に支援している。	・「読む・書く」の判定 ・「聞く・話す」の判定 各判定約15分	☑多肢選択形式 ☑記述式 ☑口述式	「対象者判定の手引き」に沿って採点。採点後は、同手引きのフローチャートに基づきレベルを判定	☑本人 判定実施後、1週間以内に本人に通知	無料	☑無	http://www.city.toyota.aichi.jp/index.html
J-cert生活・職能日本語検定 (J-cert)	公益財団法人 国際人財開発機構	日本語習熟度の検定試験を行い、日本への留学・研修ないしは日本企業への就職に役立たせることを目的とする。	2009年 2017年 GNK→J-cert 名称変更	日本語を母語としない内外の外国人	【国内】 0人 【海外】 7,729人 (2019年)	Aコース受検 189人 C1(上級) 51人 B2(準上級) 45人 B1(中級) 24人 不合格69人 Bコース受検 7,339人 A2.2(準中級) 3,711人 A2.1(初級) 1,612人 不合格 2,016人	【国内】 0回 【海外】 4回 (2019年) 3・5・9・11月の 第一日曜日 (基本スケジュール)	【国内】 1都市 【海外】 9か国	☑紙媒体 ☑対面 ☑その他 (録音音源の再生)	☑読むこと（読解） ☑聞くこと（聴解） ☑書くこと（記述） ☑話すこと（発表） ☑話すこと（やりとり） ☑その他（言語知識：文字・語彙・文法） ☑その他（社会文化知識） ※話すこと書くことはマスター級にて実施	C2（マスター級） C1（上級） B2（準上級） B1（中級） A2.2（準中級） A2.1（初級） A1（基礎級） 2021年3月より基礎級認定予定	マスターコース 総合100分 会話30分 Aコース 文法・読解60分 聴解30分 社会文化40分 Bコース 社会・文化(読解)60分 聴解30分 文字・語彙・文法45分	☑多肢選択形式 ☑記述式 ☑口述式	・社会文化、聴解、文字・語彙・文法試験、筆記総合 ☑マークシート式採点・判定 ・口述試験 ☑録音された会話を複数の採点担当者が確認して採点 ※口述試験はマスター級にて実施	☑本人 ☑機関 ☑受験後、4週間以内に合否結果通知書を自宅、または所属団体に郵送。 ☑合格者の合否結果通知書には認定書を記載	【国内】 マスターコース 15,000円 Aコース 5,000円 Bコース 4,000円	☑無	https://www.j-cert.org/
実践日本語コミュニケーション検定 (PJC)	株式会社サーティファイ「コミュニケーション能力認定委員会」	日本で就労する際に必要とされる日本語能力試験N1～N2相当の「ビジネス会話レベル」の日本語コミュニケーション能力を測定することを目的とする。	2013年	日本語を母語としない者	約2,000人 国内：海外=9:1	受験者全員をレベル認定	随時	【国内】 5都市 【海外】 4か国	☑紙媒体 ☑P C利用 (会場受験)	☑読むこと（読解） ☑聞くこと（聴解） ☐書くこと（記述） ☐話すこと（発表） ☐話すこと（やりとり） ☑その他 (言語知識：語彙、表現)	A(B2) B(B1) C D E	80分 (読解30分、聴解・聴読解50分)	☑多肢選択形式	マークシート式回答を機械で読み込み採点	☑本人 ☑機関 ・ペーパー試験 マークシート返送から2週間程度でレベル証明書を送付。 ・Web試験 結果は試験後即時で画面表示。レベル証明書は試験後1週間以内に後追いで送付。	・評価試験(ペーパー試験) 5,000円 ・評価試験 (Web試験) 4,400円 ・アセスメント試験 2,200円	☑無	https://www.sikaku.gr.jp/c/pjc/

浜松版日本語コミュニケーション能力評価システム (HAJAC)	公益財団法人浜松国際交流協会	浜松市の多文化共生を目指す取り組みの一環として、外国人住民の自立と生活ステージの向上を図るために、個人が持つコミュニケーション能力を浜松版日本語コミュニケーション能力評価システム (HAJAC)で視覚化 (見える化) し、その評価を活用することで、外国人住民、日本人支援者、日本社会のそれぞれに具体的な波及効果をもたらすことを目的とする。	2013年	浜松市内に在住・在勤の日本語を母語としていない外国人	約120人 (外国人学習支援センター日本語教室やHICE主催の日本語教室申込者のプレテスト及び修了時テストで実施)	約120人	随時 (申込時)、年2回 (修了時テスト)	静岡県浜松市外国人学習支援センター	☑対面	☐読むこと (読解) ☐聞くこと (聴解) ☐書くこと (記述) ☐話すこと (発表) ☐話すこと (やりとり) ☐その他	C7~C0	半構造化インタビュー (レベルによってはロールプレイ/逆質問も行う) 15分間	☑口述式	評価者は音声聞き返ししながら、評価基準表を元に学習者のレベルを判定する。判定は2名の評価者で実施する。	☑本人 インタビュー評価シートへ記入し、学習者へ対面でフィードバックする。	無料	☑無	http://hice.jp/publish/tools.html
実践日本語コミュニケーション検定・ブリッジ (PJC Bridge)	株式会社サーティファイ「コミュニケーション能力認定委員会」	日本での就学に必要な日本語能力試験 N5~N3レベルの基礎的な日本語能力を測定することを目的とする。	2015年	日本語を母語としない者	約4,000人 国内：海外 = 4:6	受験者全員をレベル認定	随時	【国内】10都市 【海外】11か国	☑紙媒体 ☑P C利用 (会場受験)	☑読むこと (読解) ☑聞くこと (聴解) ☐書くこと (記述) ☐話すこと (発表) ☐話すこと (やりとり) ☐その他 (言語知識：語彙、表現)	A(B1) B(A2) C(A1) D E	60分 (読解40分、聴解20分)	☑多肢選択形式	マークシート式回答を機械で読み込み採点	☑本人 ☑機関 ・ペーパー試験 マークシート返送から2週間程度でレベル証明書を送付。 ・Web試験 結果は試験後即時で画面表示。レベル証明書は試験後1週間以内に後追いで送付。	・評価試験(ペーパー試験) 5,000円 ・評価試験 (Web試験) 4,400円 ・アセスメント試験 2,200円	☑無	https://www.sikaku.gr.jp/c/pjc/bridge/
OPIc (日本語版)	一社) Global 8 がLTI(Language Testing International)許諾のもと日本における契約・実施を展開。LTIは1992年設立され、世界最大の外国語教育に関する学会ACTFL(全米外国語教育協会)の独占的語学テスト実施機関として、40ヶ国以上で14ヶ国語のOPIc語学テストを提供。	OPIの受験機会を拡大すべく i B T (internet based test) 化したもの	2016年 (OPIc日本語版のサービス)	日本語を母語としない者を対象に日本語の口頭運用能力を測定する。	非公開 (参考値としてサービス初年度 2016年実績 国内：8人 海外：642人)	当テストは合格/不合格の概念が無く、レベル認定が行われる。認定レベルは非公開	随時	要求により調整可能 海外は他の認定機関にて実施対応	☑P C利用 (会場受験) ☑その他 (在宅は監督行為や認定利用目的で対応可能性あり)	☐読むこと (読解) ☐聞くこと (聴解) ☐書くこと (記述) ☐話すこと (発表) ☐話すこと (やりとり) ☐その他	Superior~NoviceLowで評価。尚、評価期待レベルにより申込レベル(FORM)を5段階で設定。それぞれの評価レベルの意味やCEFRとの関係は以下URL参照下さい。 http://www.global8.or.jp/Evaluation%20Indicators_Global8.pdf	説明・準備に20分 話すこと (やりとり)に概ね30分 (レベルや回答時間による)。	☑口述式	・口述試験 ・終了後にクラウドに録音された (やりとり) データをACTFL公認のRaterがガイドラインに照らしながら、被験者の口頭運用能力がどのレベルにあるかを評価	☑本人 或いは ☑機関 (法人会員) に1週間以内に通知	実施並びに結果通知方法により異なるが、10,000円+消費税を基準に設定	☑無	http://global8.or.jp/opic.html
JLCT (外国人日本語能力検定)	一般社団法人外国人日本語能力検定機構 (J L C T)	言語コミュニケーション能力を測ることを目的とする。	2018年	原則として日本語を母国語としない人	【国内】13人 【海外】24,231人 (2019年)	(JCT1) 合格者：76人 (JCT2) 合格者：361人 (JCT3) 合格者：863人 (JCT4) 合格者：3,885人 (JCT5) 合格者：10,052人	年間5回 (3月、5月、7月、9月、11月) (2019年)	【国内】1都市 【海外】23か国	☑紙媒体 ☑P C利用 (会場受験) (中国のみ)	☑読むこと (読解) ☑聞くこと (聴解) ☐書くこと (記述) ☐話すこと (発表) ☐話すこと (やりとり) ☐その他 (言語知識：文字・語彙・文法)	JCT1→CEFR C1 JCT2→CEFR B2 JCT3→CEFR B1 JCT4→CEFR A2 JCT5→CEFR A1	(JCT1) 語彙文法読解 110分 聴解 60分 (JCT2) 語彙文法読解 105分 聴解 50分 (JCT3) 語彙 30分 文法読解 70分 聴解 40分 (JCT4) 語彙 30分 文法読解 60分 聴解 35分 (JCT5) 語彙 25分 文法読解 50分 聴解 30分	☑多肢選択形式	・マークシート式回答を機械で読み込み採点 ・中国WEB試験は正答をプログラミングして即時採点	☑本人 ☑機関 ・国内は社団本部から、海外は海外の各エージェントから通知 ・合格者全員に合格証書と成績通知書を各国エージェント宛て送付。所定手続を行った者には日本国内の学校宛に成績通知書を発行 採点結果は試験後3週間前後。証明書発送は3~4週間前後 中国WEB試験の採点結果は即日。証明書発送は1週間前後	【国内】- 【海外】実施都市によって異なる	☑無 ※過去問題集を発行	www.jlct.jp

ONi T 口頭ビジネス日本語試験	有限会社アイシーアイ	第一に、受験者がビジネス場面における日本語を「話す」力を証明したり、自身の弱点を客観的に理解することで、より成長し、充実した社会人生活を送るための指標となることである。 第二に、企業側が受験者の得手不得手を客観的に把握することで、配属の際のミスマッチ防止や適切な研修の実施が可能になるため、企業の人材戦略の強化の一助になることである。	2018年	主に日本語を使用して仕事に従事する必要がある企業への就職を希望する学生、内定者、入社して日の浅い社員などを想定している。	非公表	非公表	年間を通じていつでも受験可能	国内会場は、現在アイシーアイ会場（東京）のみ （一定の条件を満たした場合にオンライン受験も可能）	☑ P C 利用（会場受験） ☑ P C 利用（一定の条件を満たした場合在宅受験）	☐読むこと（読解） ☐聞くこと（聴解） ☐書くこと（記述） ☐話すこと（発表） ☐話すこと（やりとり） ☐その他（6つの評価観点：発音、文法、語彙、流暢さ、構成、聞き手への配慮/場面に応じた言葉遣い）	●レベル0～7の8段階 ・レベル7（スコア255～300点） ・レベル6（スコア230～254点） *レベル6,7「高度な業務対応可能」 ・レベル5（スコア205～229点） ・レベル4（スコア180～204点） *レベル4,5「社外業務対応可能」 ・レベル3（スコア130～179点） ・レベル2（スコア80～129点） *レベル2,3「社内業務対応可能」 ・レベル1（スコア1～79点） *レベル1「日常会話レベル」 ・レベル0（スコア0点） *レベル0「評価不能」 ※CEFRとのレベル対照はHPを参照。 尚、主にCEFR B1-C2レベルの受験者を対象にしたテストであるため、A1,A2レベルの受験者の受験は推奨しない。	●話す力：質問と回答時間合わせて約1時間。 ・ウォームアップ：40秒 ・問題1 状況説明能力：40秒 ・問題2 指示確認能力：40秒 ・問題3 伝言能力：60秒 ・問題4 状況連絡能力：40秒 ・問題5 確認能力：60秒 ・問題6 報告相談能力：60秒 ・問題7 意見説明能力：60秒 ・問題8 報告能力：60秒 ・問題9 敬語運用能力：40秒 ・問題10 反対意見説明能力：90秒 ・問題11 解決方法提案能力：120秒 ・問題12 意見根拠説明能力：120秒	☐口述式 問題は音声や画像で提示し、音声を録音して解答。	・コンピュータを介して録音された解答を第三者機関で認定された複数の評定者が評価。 ・評価の信頼性の確保は評定者間信頼性に拠る。 ・すべての評価について第三者機関（一般社団法人日本語コミュニケーション協会が設置する評価、日本語教育の専門家からなる試験評価委員会「TEC」）が信頼性をチェック。	☑本人 ☑機関 約1ヶ月後、受験者もしくは企業ご担当者へ評価結果を記載した「公式評価レポート」を送付。	15,000円+税/回 ※追加でコンサルティンクを希望の場合別途5,000円+税/回	☐無	https://www.orainihongotest.org/
日本語能力評価試験（J P E T）	学校法人駿河台学園日本語能力評価試験協会	学習・ビジネスを含む日常生活シーンでの日本語運用能力を測定することを目的とする。（熟達度テスト）	2019年	日本語を母語としない外国人留学生・ビジネスパーソン	5,679人（直近2020年1月～2020年12月）	1000点満点の無段階評価のため合否無し	11回（4月を除く毎月1回）	【国内】6都市 【海外】3か国 ※2021年度予定	☐紙媒体	☑読むこと（読解） ☑聞くこと（聴解） ☐書くこと（記述） ☐話すこと（発表） ☐話すこと（やりとり） ☐その他（言語知識(文字・語彙、文法)）	1000点満点の無段階評価 下記はCEFRとの対応表 得点：CEFR 1000～940：C2 939～890：C1 889～780：B2 779～430：B1 429～350：A2 349～300：A1	聴解・読解（合計90分）	☐多肢選択形式	・マークシート式回答を機械で読み込み採点 ・令和3年度より I R T を活用してスコアを算出	☑本人：受験後10日間で郵送通知 ☑機関：依頼後2日間で機関に直接郵送通知	【国内】5,000円 【海外】実施都市によって異なる（3,000円前後）	☑有（年3回）	https://www.jp-et-sundai.jp/
国際交流基金日本語基礎テスト（JFT-Basic）	独立行政法人国際交流基金	主として就労のために来日する外国人が遭遇する生活場面でのコミュニケーションに必要な日本語能力を測定し、「ある程度日常会話ができ、生活に支障がない程度の能力」があるかどうかを判定する	2019年	日本語を母語としない外国人、その中でも、主として就労のために来日する外国人を対象とする	【海外】7,971人（2019年度）	基準点到達者数（2019年度） 【海外】2,737人	2021年3月から、年6回（奇数月）実施予定	【海外】7か国（実績） 【国内】47都道府県（2021年3月開始予定）	☑ P C 利用（会場受験）	☑読むこと（読解） ☑聞くこと（聴解） ☐書くこと（記述） ☐話すこと（発表） ☐話すこと（やりとり） ☐その他（文字と語彙、会話と表現）	就労のために必要な、「ある程度日常会話ができ、生活に支障がない程度の能力」の目安として、A2レベルの一定程度の日本語力を持っているかどうかを判定	文字と語彙会話と表現 聴解 読解 計60分（セクションごとの解答制限時間はない）	☐多肢選択形式	受験者がPC画面で入力した解答を即時自動採点、IRTに基づいて総合得点（尺度得点）を算出し、判定基準点に到達したかどうかを判定	☑本人 テスト終了時のPC画面に総合得点と判定結果を表示。また、受験後5営業日以内に判定結果通知書をウェブサイト上で発行。	【国内】7,000円 【海外】国によって異なる	☐無	https://www.jpft-go.jp/jft-basic/index.html

2. 社会で活用される試験に望まれる要素

能力評価の名称	(1)テストスペック (試験の細目表)	(2)サンプル問題の 公表	(3)得点配分と合否 の判定方法の公開	※パフォーマンス評価 等の場合、評価者 に対する研修の実施	(4)第三者評価等の 体制	(5)IRTに基づく 得点等化	(6)結果分析への IRT活用	(7)特に配慮を要する 受験者への対応	(8)情報セキュリティ 管理体制	(9)不正・偽造等 防止対策	(10)安全確保等 対策整備
日本語能力試験	有	有	有	—	有	有	有	有	有	有	有
JPT日本語能力試験	有	有	無	—	有	有	有	有	有	有	有
ACTFL-OPI	無	無	無	有	有	無	無	—	有	有	—
日本語NAT-TEST	有	有	有	—	無	無	無	有	有	有	有
J.TEST 実用日本語検定	有	有	有	—	有	無	有	有	有	有	無：検討中
ACTFUL及びLTIの習熟度試験 (日本語版)	有	有	有	有	有	有	有	有	有	有	有
BJTビジネス日本語能力テスト	有	有	無 (合否判定なし)	—	有	有	有	有	有	有	有
日本留学試験(EJU)	有	有	有	有	有	有	有	有	有	有	有
標準ビジネス日本語テスト (STBJ)	有	有	有	—	無：検討中	無：検討中	無：検討中	有	有	有	有
J-CAT 日本語テスト	有	有	無	—	無	有	有	無	有	有 (Certificateプラン のみ)	無
アルクの電話による日本語会 話テスト JSST	有	有	無	有	無：検討中	無	無	無	有	有	有
TOPJ実用日本語運用能力 試験	有	無	無	—	有	無：検討中	無：検討中	有	有	有	有
とよた日本語能力判定 (対象者判定テスト)	有	無	無	有	無	無	無	有	有	無	有
J-cert生活・職能日本語検定 (J-cert)	有	有	有	有	有	無	無	有	有	有	有
実践日本語コミュニケーション 検定 (PJC)	有	有	有	—	有	無：検討中	無：検討中	有	有	有	有
浜松版日本語コミュニケーシ ョン能力評価システム (HAJAC)	無	無	無	有 ※不定期	無	無	無	無	有	無	有
実践日本語コミュニケーション 検定・ブリッジ (PJC Bridge)	有	有	有	—	有	無：検討中	無：検討中	有	有	有	有
OPIc (日本語版)	有	有 ※但し、英語	有 ※但し、日本での一 般概念とは異なる	有	有	有	有	有	有	有	有
JLCT (外国人日本語能力検 定)	有	有	無	無：検討中	有	無：検討中	無：検討中	有	有	有	有
ONIT口頭ビジネス日本語 試験	有	有	無	有	有	無	無	無	有	有	無
日本語能力評価試験 (JPET)	有	有	無	有	無：検討中	無：検討中	無：検討中	有	有	有	有
国際交流基金日本語基礎テ スト (JFT-Basic)	有	有	有 (合否の判定方 法)	—	有	有	有	無	有	有	有

3. 「日本語教育の参照枠」レベル尺度との対応付けに向けて

能力評価の名称	(1)当該テスト独自のCan do	(2)当該テスト独自のCan doの検証	(3)CEFRのレベル尺度との対応付け	(4)CEFRのレベル尺度との対応付けの検証	(5)検証結果の公表	(6)検証方法の公表
日本語能力試験	有	有	無	有	有	有
JPT日本語能力試験	有	有	有	有	有	有
ACTFL-OPI	無	無	有	有	有	有
日本語NAT-TEST	無	無	有	有	無	無
J.TEST 実用日本語検定	有	有	有	有	無	無
ACTFUL及びLTIの習熟度試験（日本語版）	有	有	有	有	有	有
BJTビジネス日本語能力テスト	有	有	有	無	無	無
日本留学試験(EJU)	無	無	無	無	無	無
標準ビジネス日本語テスト(STBJ)	有	無	有	有	有	有
J-CAT 日本語テスト	有	無	有	有	有	無
アルクの電話による日本語会話テスト JSST	有	有	無	無	無	無
TOPJ実用日本語運用能力試験	有	有	有	有	無	無
とよた日本語能力判定（対象者判定テスト）	有	無	無	無	無	無
J-cert生活・職能日本語検定(J-cert)	有	有	有	有	無	無
実践日本語コミュニケーション検定(PJC)	有	有	有	有	無：検討中	無：検討中
浜松版日本語コミュニケーション能力評価システム(HAJAC)	有	無	無	無	無	無
実践日本語コミュニケーション検定・ブリッジ(PJC Bridge)	有	有	有	有	無：検討中	無：検討中
OPIc（日本語版）	有	有	有	有	有	有
JLCT（外国人日本語能力検定）	無	無	有	有	有	有
ONIT口頭ビジネス日本語試験	有	有	有	有	有	有
日本語能力評価試験(JPET)	有	有	有	有	無	無
国際交流基金日本語基礎テスト(JFT-Basic)	有	有	有	有	無：検討中	無：検討中