

「国際交流基金日本語基礎テスト」の開発

—生活場面でのコミュニケーションに必要な言語能力 (A2レベル) を判定する CBT—

熊野七絵・戸田淑子・安達祥子

〔キーワード〕 外国人材受入れ(特定技能)、生活場面、A2レベル、CEFR/JFS の Can-do、CBT

〔要 旨〕

国際交流基金日本語基礎テストは、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言語能力を測定し、「ある程度日常会話ができ、生活に支障がない程度の能力」として A2レベルの一定程度の能力水準に達しているかどうかを判定することを目的とし、開発した。CEFR および JF 日本語教育スタンダードの考えに基づき、日本での生活場面に絞った Can-do を元に真正性のある問題を作成することで、高い信頼性と妥当性を備えたテストを実現した。また、CBT 方式により、頻回受験、即時結果通知、漏洩防止などを備えるとともに、馴染みのない受験者への対応を工夫することで、テストの円滑実施を実現した。

1. 国際交流基金日本語基礎テストの概要

「国際交流基金日本語基礎テスト」(Japan Foundation Test for Basic Japanese、以下、JFT-Basic) は、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言語能力」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ことを目的としたテストである⁽¹⁾。対象者は日本語を母語としない外国人で、主として就労のために来日する外国人を対象としている。在留資格「特定技能1号」を得るために必要な日本語能力水準を測るテストとしても活用されている。

ヨーロッパの言語共通参照枠である“Common European Framework of Reference for Languages: Learning, teaching, assessment”(以下、CEFR、Council of Europe 2001) および、CEFR に沿って国際交流基金が「相互理解のための日本語」を理念とする日本語教育のための枠組みとして開発した「JF 日本語教育スタンダード」(以下、JFS、国際交流基金 2017a)の考えに基づき、課題遂行において「何がどれだけできるか」を示した能力記述文(以下、Can-do)を参照して開発されている。

本テストは「文字と語彙」「会話と表現」「聴解」「読解」の4セクション構成で、CEFR、JFS の Can-do に基づいた問題が出題され、4セクションの解答結果から総合的に「生活場面でのコミュニケーションに必要な言語能力」を測定し、既述の「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能力水準に達しているかを判定するものである。

本テストはCBT（Computer Based Testing、以下、CBT）方式⁽²⁾により実施され、海外各国のテスト会場でコンピュータを使用してテスト問題を出題し、判定結果はテスト終了後に即時通知される。2019年4月から実施を開始し、2020年8月現在まで隔月で累計6か国において実施されている。

2. 開発の経緯

2018年6月に閣議決定された「骨太の方針2018」において、「新たな外国人材の受入れ」が盛り込まれ、同12月には出入国管理及び難民認定法改正法案が可決された⁽³⁾。同改正において、新たな在留資格「特定技能」が創設され、2019年4月から施行されている。特定技能の在留資格に係る制度の運用に関する基本方針では、「特定技能1号」の資格要件について、法務省が定める基準に即し、14の特定産業分野の技能試験と日本語試験を実施することが定められている。「特定技能1号」による外国人材受入れは5年間で34万人が想定され、外国人材受入れニーズが高い9か国（ベトナム、フィリピン、カンボジア、中国、インドネシア、タイ、ミャンマー、ネパール、モンゴル）を主な対象国とし、試験の開発や実施に向けた検討が進められることとなった。

国際交流基金は、①大規模試験である日本語能力試験（以下、JLPT）の開発、海外での実施経験⁽⁴⁾、②CEFRを参照して開発したJFSおよび各種テストの開発経験⁽⁵⁾、③eラーニングやCBTの開発経験などがあった⁽⁶⁾。そこで、これらの知見、実績をもとに、新たな日本語能力判定試験の開発に着手することとした。

新たな日本語試験の開発にあたり、(1) 外国人が日本社会で生活、就労する上で必要な日本語コミュニケーション力を備えているかを、国際的標準⁽⁷⁾を踏まえ確認することができる能力判定テストであること、(2) 多くの外国人材の継続的な入国に資する海外での頻回受験、即時結果通知、在留資格要件ともなるハイスタークスのテストに必要な本人確認やカンニング、漏洩防止などの条件を備えていること、の2点を要件とした。

3. 先行事例調査

3.1 他国の在留資格要件となる言語テスト調査

開発にあたり、他国の移民や就労者の在留資格要件として利用されている言語テストについて、調査を行った。韓国では、就労の在留資格用にEPS-TOPIK（A2）を開発し、協定締結国において実施している。欧州各国では、移民の在留査証再発給の条件として、社会統合のための知識と言語能力を確認するため、言語についてはCEFRに基づく大規模総合テストの受

験結果を活用するものが多い。例えば、フランスは1年目に DELF (A1)、5年目までに DELF (A2)、イタリアは CILS (A2) の証明が必要である。移民向けの試験として開発されたものとしては、言語コース等を受講後に受験するドイツの DTZ (B1)、オランダの StaatexamenNT2 (B1) があり、各国内で実施されている。英国、オーストラリア、カナダなど英語圏では、IELTS が利用されている。上記のうち、CBT での実施はオランダの StaatexamenNT2 のみで、他は全て PBT (Paper Based Testing) で実施されている。

入国前後の在留資格としては、A2レベルを求めるテストが多く、JFT-Basic の判定レベル (A2) と共通していた。一方、就労者向けに入国前に CBT で実施されるテストの前例はなかった。

3.2 CBT の日本語能力判定テスト

日本語能力レベルを総合的に判定することを目的とし、CBT で実施されているテストには、以下のものがある。簡易に日本語運用力を測定するテストとして、プレースメントテストなどに広く使われ、現在は「TTBJ」上で Web 版を利用できる「SPOT」(酒井ほか 2015)、項目応答理論に基づくコンピュータアダプティブテストとして開発され、レベル分け等で利用されている「J-CAT」(今井ほか 2012)、ビジネス日本語能力を測るために開発され、現在は CBT 方式でも実施され、高度外国人材の日本語能力の参考指標の一つとしても活用されている「BJT」(葦原・小野塚 2014)、「生活者としての外国人」のための日本語能力判定のために開発された「とよた日本語能力判定」(豊田市 2015)、CEFR 準拠の多言語オンラインテストとしてフランスの高等教育機関で開発された「SELF」の日本語テスト(東・代田 2019) などである。

いずれも、各対象や目的に合わせて開発され、入門から上級 (A1-B2) などの幅広いレベルの問題を出題し、解答結果から、受験者がどのレベルに属しているかを判定するものである。一方、JFT-Basic は、日本国内で外国人が遭遇する「生活場面でのコミュニケーションに必要な言語能力」を測定し、基礎段階である A2レベルの一定の能力水準に達しているか判定することが目的であり、海外での CBT 方式での大規模実施を前提とし、在留資格要件ともなるハイステークスなテストであることから、同目的に合わせたテストを新規で開発、実施することとした。

4. テスト開発サイクル

JFT-Basic の開発にあたっては、CEFR とその補遺版、テスト開発マニュアル “Manual for Language Test Development and Examining” (ALTE 2011)、CBT に関する国際テスト委員会のガイドライン “International Guidelines on Computer-Based and Internet Delivered Testing”

(ITC 2005) を参考にした。開発は、Planning (計画) → Design (作題) → Try-out (試行) → Informing stakeholders (関係者への通知) の段階を踏んだ (ALTE 2011: 60)⁽⁸⁾。図1の開発サイクルに示したように、「Design (作題)」は、Test specifications (テストの構成や方針の設定)、Assembling test (問題作成、質の調整)、Live test material (CBT 用素材作成、CBT 化) の順に行い、「Try out (試行)」としてプレテストを行い、分析、検証した。

4.1 テストの構成や方針の設定

まず、テストの全体構成や構成概念を検討した。CBT の配信仕様や会場、即時自動採点の制約から、「作文」「会話」セクションの大規模実施は困難であることも勘案し⁽⁹⁾、「文字と語彙」「会話と表現」「聴解」「読解」の4セクション構成とした。CEFR や JFS を参照しながら、各セクションとカテゴリーのねらいを決定し、カテゴリー、レベルごとの該当 CEFR Can-do、問題文や選択枝に使用する語彙・表現・文法・イラスト、その表記・表示方法、設問のパターンなどの出題方針を定めた。また、CBT 方式で効率的に能力を測定、判定するため、想定受験時間を60分とし、受験時間内に、高い信頼性で測定できるテストとして望ましい問題数、ラッシュモデルによる項目応答理論 (IRT) に基づく分析、等化において、統計上必要な問題数などの条件を踏まえて、各セクションの大問、小問などの問題数を設定した。

4.2 問題作成、CBT 化

問題作成と CBT 化の流れは、図1の①-⑧に示したとおりである。

- ① セクションごとに参照する CEFR のコミュニケーション言語活動・言語能力のカテゴリーに該当する JFS の A1-A2 の Can-do 一覧から、具体的な Can-do 例を選定する。
- ② セクションごとに JFS の Can-do に基づき、CEFR の A1、A2. 1、A2. 2 の各レベルの Can-do などの出題方針を参照しながら、各問題を作成する。
- ③ 全セクションの問題の複数回の検討、修正を経て、質の調整を行う。
- ④ Can-do やトピックのバランス、設問や問題文、選択枝など、全体の調整を行う。
- ⑤ 外部委員の確認、コメントを受け、問題の修正の検討を行い、再修正した上で、完成する。
- ⑥ CBT の配信仕様に合わせて、画像、音声、原稿、多言語翻訳などの素材作成を行い、CBT 用の入稿フォーマットに記入し、入稿する。
- ⑦ CBT 業者により、CBT 化する。

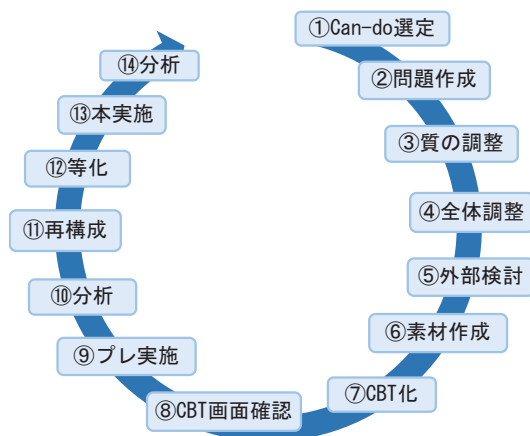


図1 JFT-Basic 開発サイクル

- ⑧ CBT化されたテスト配信用画面で複数回の確認、修正を繰り返し、CBT業者がテスト会場で動作確認を行う。

4.3 プレテスト実施と分析、検証

本実施に先立ち、プレテストを実施し、分析、検証を行った(図1、⑨~⑫)。プレテストは複数の実施予定国において、対象レベルの受験協力者221名に対して、123問、約2セット分のテスト問題を出題した。

結果は統計的に分析し、テストの信頼性、妥当性の検証や、各問題の品質評価を行った。テストの信頼性係数であるクローンバックの α は各セットで0.90、0.89であり、高い信頼性が確認できた。セクション間の平均通過率等の結果にも大きな偏りはなかった。各問題の通過率、識別力や選択枝の解答傾向などの品質評価の結果、統計上注意を要する問題は2問のみであり、問題の品質は良好であった。また、A1、A2.1、A2.2の出題方針に基づいて作題した問題が想定通りに機能しているのか、追加分析を行い、通過率はレベルが上がるにつれて低くなっており、想定通り本テストの測定に寄与していることを確認した。また、解答時間も計測し、各問題の平均解答時間などから、本実施における解答時間は60分に決定した。

次に、プレテストの結果を元に、本テストの判定基準点を設定した。基準点の設定にあたっては、Modified Angoff法を用いた⁽¹⁰⁾。専門家の予測値と受験結果との統計的な検証を行った結果、予測値の信頼性は0.92と高く、推奨される基準点が算出された。この基準点に基づき、尺度得点の得点幅(10-250点)と判定基準点(200点)を設定し、テスト問題の等化を行った。

さらに、等化後の数値を元に、受験結果のテスト情報量(TIF)分析を行い、受験者の能力値について確かな情報を得られるテスト問題構成となっていることも確認している。

また、基準関連妥当性分析として、JLPTを外的基準とし、JFT-Basicが日本語能力を計測するテストとして有効かを検証するため、JFT-BasicのプレテストとJLPTの双方を受験した受験者(計253名、N5:67名、N4:117名、N3以上:69名)の受験結果を分析した。相関分析では、JLPTの全レベルのパーセンタイルランクとJFT-Basicの尺度得点間に相関があり、JLPTのN4とN5の尺度得点とJFT-Basicの尺度得点間には、強い相関があることを確認した。また、回帰分析によって、JLPTの全レベルのパーセンタイルランク、N4とN5の尺度得点は、ともにJFT-Basicの尺度得点を予測する統計的に有意な変数となっていることを確認した。以上から、JFT-Basicが日本語能力を計測するテストとして有効であるという妥当性を検証した。

なお、2019年4月以降の本実施において能力判定に使われているテスト問題は、統計的な品質評価において、一定の基準値以上であることを確認し、等化済みの問題であり、再受験者にも対応できる十分な問題プールを備えている。各実施回のテスト結果は統計的に分析、検証し、分析結果に基づき、出題方針や基準を精緻化し、次の作題に活かしている(図1、⑬~⑭)。

5. テストの構成と問題作成

5.1 テストの構成とねらい

本テストの構成は既述のとおり、「文字と語彙」「会話と表現」「聴解」「読解」の4セクションである。テストの構成と構成概念を検討するにあたり、CEFR、JFSの考えに基づき、課題遂行のための言語コミュニケーションにおいては、「コミュニケーション言語活動」だけでなく、それを支える「コミュニケーション言語能力」も必要であると考えた。「コミュニケーション言語能力」として、基礎段階の言語使用者にとって日本での生活で「文字」が読める、書ける、基本的な語彙を持ち、使えるということは重要な要素であるため、「文字と語彙」セクションを設け、また、日常会話において必要な文法や表現を使えるかを測るため、「会話と表現」セクションを設けることで、「作文」「会話」などの言語活動を下支えする能力を測ることとした。

図2は、本テストで測定する言語能力と、その下位能力、および、それらを測るセクションを整理したものである。テストの構成概念として「生活

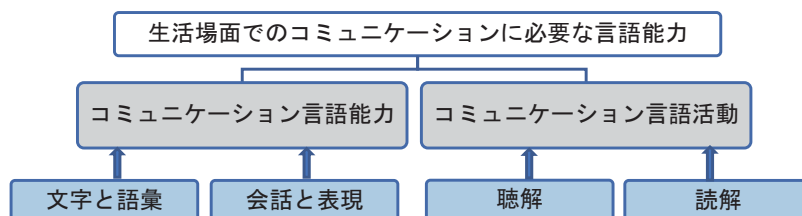


図2 本テストで測定する言語能力

場面でのコミュニケーションに必要な言語能力」を測定するために、「コミュニケーション言語活動」を「聴解」「読解」セクションで、「コミュニケーション言語能力」を「文字と語彙」「会話と表現」セクションで測定することとした。

各セクションは、さらにいくつかのカテゴリーに分かれており、それぞれのセクション、カテゴリーのねらいは表1に示した通りである。

表1 テストの構成とねらい

セクション	セクションのねらい	カテゴリー	カテゴリーのねらい
文字と語彙 約15問	生活場面で使用される日本語の文字が読めるか、基本的な語彙を持ち、使えるかを測る。	語の意味	語の意味を問う。
		語の用法	語の用法を問う。
		漢字の読み	漢字で書かれた語の読み方をひらがなで問う。
		漢字の意味と用法	漢字で書かれた語の意味と用法を問う。
会話と表現 約15問	生活場面の会話に必要な文法や表現を使えるかを測る。	文法	会話の文脈に合った適切な文法が使えるかを問う。
		表現	会話の文脈に合った適切な表現が使えるかを問う。
聴解 約15問	生活場面において、会話や指示などを聞いて、理解できるかを測る。	内容理解(社会的なやりとり)	情報交換や社会的なやりとりを聞いて、内容を理解できるかを問う。
		内容理解(店や公共機関でのやりとり)	店や公共機関でのやりとりを聞いて、内容を理解できるかを問う。
		内容理解(指示、アナウンス)	指示やアナウンス、音声メディア等を聞いて、内容を理解できるかを問う。
読解 約15問	生活場面において、手紙や掲示、説明などを読んで、理解できるかを測る。	内容理解	手紙やメッセージなどの短い簡単なテキストを読んで、内容を理解できるかを問う。
		情報検索	日常の看板や掲示、資料などから必要な情報を探し出せるかを問う。

5.2 Can-do に基づく問題作成

問題はセクション、カテゴリーごとに、A1、A2.1、A2.2の3つのレベルのCEFR Can-do を作題のレベル指標とし、各レベルに合った問題を作成している。A1レベルのCan-doに基づいた問題も出題している理由は、その中でも生活に必須のCan-doがあり、A2レベルの能力水準に達しているかを測定するのに必要であると考えたためである。A2レベルを下位レベル、A2.1とA2.2に分けているのは、A2レベルはA1レベルに比べ幅が広く、CEFRでは必要性に応じて実際のレベルに分けることが可能だとされていること（Council of Europe 2004：32）による。

さらに、問題ごとにJF Can-do 一覧や、同対象者のために生活場面に絞り同時開発中だった「JF生活日本語Can-do」一覧からCan-do例を選定することで、場面・トピック、言語活動を具体化している。選定対象となったのは、JF Can-doのうちのA1、A2のCan-do全326個と、「JF生活日本語Can-do」のA1、A2の全381個のCan-doで、JFSの15トピックを参照しトピックに偏りのないように選定している。また、Can-do例は「文字と語彙」セクションでは主に活動Can-do全般から⁽⁴⁾、「会話と表現」セクションでは主にやりとり（口頭）Can-doから、「聴解」セクションでは主にやりとり（口頭）、受容（聞く）Can-doから、「読解」セクションでは主に受容（読む）Can-doから選定している。

表2では、JFT-Basic ウェブサイトで公開している全カテゴリーのA2.2の「サンプル問題」作成にあたり、参照したCEFR Can-do、Can-do例を一覧で示す。

表2 サンプル問題で参照したCan-do

セクション	カテゴリー	CEFR Can-do	Can-do 例
文字と語彙	語の意味	【CEFR 能力 Can-do 使用語彙領域 A2.2】 馴染みのある状況や話題に関して、日常生活上の交渉・取引を行うのに十分な語彙を持っている。	友人と、映画鑑賞やスポーツ観戦などのあとで、印象に残ったことやわからなかったことなどについて、短い簡単な言葉でコメントや質問をしたり、質問に答えたりすることができる。
	語の用法		友人と、お互いの異文化体験などについて、短い簡単な言葉でコメントや質問をしたり、答えたりすることができる。
	漢字の読み	【CEFR 能力 Can-do 正書法の把握 A2】 当人が話す語彙に含まれる短い単語の音声を、（完全に標準的な綴りではない場合もあるが）割合に正確に文字化することができる。	短い簡単な文で書かれていれば、地域の掲示板、回覧板、アパートのお知らせなどを読んで、何が、いつ、どこであるか（例えば、アパートの水道管の工事）など、いくつかの情報を理解することができる。
	漢字の意味と用法		日常の出来事などについて書かれた友人の短い簡単なSNSの書き込みを読んで、内容をだまかに理解することができる。

「国際交流基金日本語基礎テスト」の開発

会話と表現	文法	【CEFR 能力 Can-do 使える言語の範囲 A2.2】 たいていの場合、言いたいことを内容的に妥協・制限したり、言葉を探したりする必要があるが、予測可能な日常的な状況に本人が対応するために必要な、基本的な言語のレパートリーを持っている。	友人に、知り合いの結婚や出産などのお祝いに、どんなもの、いくらぐらいのものを贈ったらいいか質問し、簡単な情報を得ることができる。
	表現	【CEFR 能力 Can-do 社会言語的な適切さ A2.2】 ・例えば、簡単な形で情報を交換、請求したり、意見や態度を表明したりするなどの、基本的な言語機能を実行でき、また応じることができる。 ・最も簡単な、一般的な表現や、基本的な慣習に従って、単純な形ではあるが、効果的に交際を維持することができる。	職場などで、同僚をお昼ごはんなどに誘うために、一緒に行くかどうか、短い簡単な言葉でたずねたり、誘いに答えたりすることができる。
聴解	内容理解 (社交的なやりとり)	【CEFR Can-do 社交的なやりとりをする A2.2】 時々繰り返しや言い換えを求められることが許されるなら、自分に向けられた、身近な事柄について、はっきりとした、標準語での話はたいてい理解できる。	同僚と、夏休みや年末年始などの休暇明けに、休み中何をしたか、どうだったかなど、短い簡単な言葉でやりとりするとき、相手の話をだいたい理解することができる。
	内容理解 (店や公共機関でのやりとり)	【CEFR Can-do 店や公共機関でやりとりをする A2.2】 旅行、宿泊、食事、買い物のような毎日の生活での普通の状況に対処することができる。	スーパーやコンビニなどの店員に、原材料や産地、賞味期限、ハラルかどうかなどについて質問し、いくつかの簡単な答えを理解することができる。
	内容理解 (指示・アナウンス)	【CEFR Can-do 指示やアナウンスを聞く A2】 短い、はっきりとした、簡単なメッセージやアナウンスの要点は聞き取れる。	ゆっくりとはっきりと話されれば、地域の防災に関する簡単な説明を聞いて、パンフレットなどにあるイラストや写真を手がかりに、避難場所や避難時の注意事項などを理解することができる。
読解	内容理解	【CEFR Can-do 手紙やメールを読む A2.2】 身近な話題についての日常の手紙やファックス（照会、注文、確認など）の基本的なタイプのもので理解できる。	地域の広報誌などにある、外国人向けの短い簡単なお知らせを読んで、日本語教室や交流イベントなどの場所や日時、内容など、いくつかの情報を理解することができる。
	情報検索	【CEFR Can-do 必要な情報を探し出す A2】 広告、趣意書、メニュー、参考書目録、時刻表のような、簡単な日常の資料の中から予測可能な特定の情報を見つけることができる。	薬箱に書かれている注意書きなどの短い簡単なテキストを見て、用法・用量など、必要な情報を探し出すことができる。

本テストの構成概念として、「生活場面でのコミュニケーションに必要な言語能力」を測ることを重視し、4セクション全てにおいて、主として就労のために来日する外国人が遭遇する特定の職種によらない生活場面に絞って問題を作成している。日本の生活における必要性という観点から、買い物や交通だけでなく、病院、防災などの場面・トピックも扱っている。また、素材についても、例えば「読解」セクションにおいて、読む必然性のない単なる読み物ではなく、手紙やメール、SNS などのメッセージ、日常の看板、掲示、薬の説明など、生活上のテキストを読む対象としている。

ALTE (2011: 12) には、言語テストにおいて重要な真正性 (authenticity) として、タスクや問題が実際の生活の言語活動を正確に表す situational authenticity と、受験者がタスクにかかわる自然さを表す interactional authenticity の二つの真正性が言及されている。本テストにおいても、受験者が実際に日本語を使う生活場面での言語活動を想定することによって前者を、各 Can-do で示された課題を遂行する立場となるよう設定することで後者を実現している。

5.3 セクションごとの問題例と問題作成上の工夫

5.3.1 文字と語彙

「文字と語彙」セクションでは、生活場面で使用される日本語の文字が読めるか、基本的な語彙を持ち、使えるかを測る問題を出題する。図3はカテゴリー「漢字の読み」の問題例で、下線が引かれた漢字語彙の読み方をひらがなでどう書くか、正しいものを選ぶ問題である。CEFRの能力Can-do「使用語彙領域 A2.2」「正書法の把握 A2」をレベル指標として参照した。

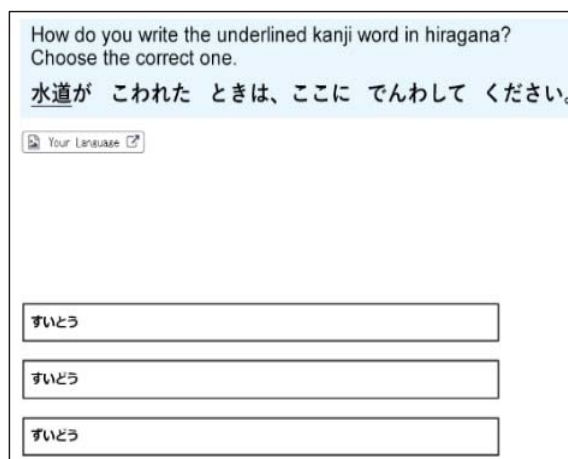


図3 「漢字の読み」の問題例

Can-do 例のように、短い簡単な文で書かれている水道についてのアパートの掲示やお知らせを見て、いくつかの情報を理解できることを達成するために必要な、漢字語彙「水道」の読み方を知っているかを測っている。

「漢字の読み」では、漢字語彙の読みを問うことで、語として、意味だけでなく正しく音がわかるか、間接的に、タイピングで文字化できるかどうかを測っている⁽¹²⁾。「漢字の読み」で測るのは、使用語彙領域、読字・書字能力であるが、課題遂行の文脈から独立したものにならないよう、Can-do 例に沿って具体的な場面、トピック、言語活動を想定して作題している。

5.3.2 会話と表現

「会話と表現」セクションでは、生活場面の会話に必要な文法や表現が使えるかを測るため、会話文を読み、文脈を理解した上で、適切な文法や表現を選ぶ問題を出題する。

図4はカテゴリー「表現」の問題例で、会話文を読み、文脈を理解した上で、適切な表現を選ぶ問題である。CEFRの能力Can-do「社会言語的な適切さ A2.2」をレベル指標として参照した。Can-do 例にあるように、職場などで同僚からお昼ごはんを誘われ、その誘いを断るという場面で、

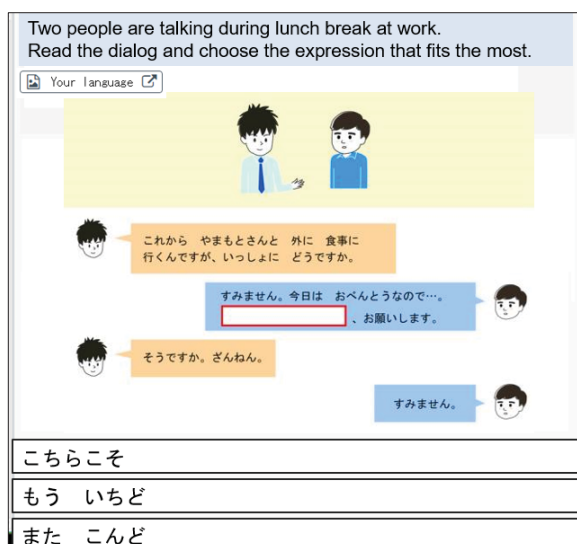


図4 「会話と表現」問題例

失礼な断り方にならずに、その後の交際も維持できるよう、「また今度、お願いします」といっ

た表現を使えるかどうかを測っている。このように、会話の文脈の中で問うことによって、間接的ではあるが、日常会話に必要な表現が使えるかを測っている。

A2の下位レベルである A2.1と A2.2の Can-do においては、レベル別の特徴が見られる。表2で示した CEFR の能力 Can-do 「社会言語的な適切さ A2.2」は「情報交換・請求、意見や態度の表明」「効果的な交際の維持」となっており、A2.1「日常的に使われる挨拶や呼びかけ」「招待や提案、謝罪」より、「社会言語的な適切さ」の広がりが見られる。「表現」カテゴリーにおいては、このようなレベル別の特徴にも留意して作題している。

5.3.3 聴解

「聴解」セクションでは、生活場面において、会話や指示などを聞いて、理解できるかを測るため、音声を聞いて、設問の答えとして正しいものを選ぶ問題を出題する。選択枝はイラストで、音声は2回まで聞ける。

図5は、カテゴリー「内容理解 指示・アナウンス」の問題例で、レベル指標として、CEFR Can-do 「指示やアナウンスを聞く A2」を参照した。Can-do 例にあるように、職場の防災訓練で担当者が、火事があったときどうするかについて話しているのを聞き、「最初に何をするか」「どこに避難するか」を聞き取れるかを測っている。A レベルでは、実際の防災訓練場面の素材をそのまま使用して、レベル判定に資する問題を作成することは難しいが、テスト作成マニュアルにもあるように、学習者の言語能力に素材や活動を合わせて (ALTE 2011 : 12)、受験者が日本で必要となる防災訓練などの場面での課題を「聞き手」として疑似的に遂行し、来日後の生活に備えられるよう作題している。

図5 「聴解」問題例

5.3.4 読解

「読解」セクションでは、生活場面において、手紙や掲示、説明などを読んで、理解できるかを測るため、テキストを読んで、設問の答えとして正しいものを選ぶ問題を出題する。選択枝は、イラストまたは文字で、設問が2問あり、設問も日本語で読む。

図6はカテゴリー「情報検索」の問題例で、CEFR の「必要な情報を探し出す A2」をレベル

指標として参照した。Can-do 例にあるように、薬局でもらった薬の説明書を見て、「薬を1度にいくつ飲むか」「薬の副作用があるかもしれないのはどれか」を探し出せるかを測る問題である。「聴解」セクション同様、受験者が来日後の生活で必要性の高い、薬の説明を読むという言語活動に「読み手」として参加することによって、言語レベルを調整した素材を使ってではあるものの、疑似的に課題遂行し、来日後の読む活動のイメージが持てるよう作題している。

このように、全てのセクションにおいて、生活場面での言語活動を想定して作題することで situational authenticity を、また、特に「聴解」

「読解」セクションにおいては、各 Can-do で示された課題を「聞き手」「読み手」として疑似的に遂行することで interactional authenticity を実現している。

Kumaru-san received instructions on medicines at the pharmacy. Read the instructions and answer questions (1) and (2).

Your Language

クマル・シュレスタ 様		2020年11月9日		
No	色・形	名前・はたらき	起 期 昼 夕 夜	注意事項
1		【ネツナレ錠】 熱を下げます。 傷水をおさえます。	1 1 1	ねむくなりやすいので、車の運転はさけてください。
			1日3回 毎食後 1回1錠【5日分】	
2		【セキドミカブセル】 せきを鎮めます。	2 2 2	薬の作用が強くなる ことがありますので、 飲酒しないでください。
			1日3回 毎食後 1回2錠【5日分】	
3		【タントレン錠】 たんの切れをよく します。	1 1	他の薬を飲んでる人 は医師に相談してくだ さい。
			1日1回 夕食後 1回1錠【3日分】	

(1)セキドミカブセルは 1回に いくつ 飲みますか。

2つ

3つ

5つ

(2)飲んだ 後、ねむくなるかもしれないのは どれですか。

ネツナレ錠

セキドミカブセル

タントレン錠

図6 「情報検索」問題例

6. CBT 方式のテストの実施と工夫

6.1 CBT 方式のテストの実施

JFT-Basic はPBT方式のテストのように、受験者がテスト会場に集合し、同一日に同一のテストを一齐に受験するのではなく、国ごとに定められたテスト会場、実施期間において、複数の実施日、時間帯に実施され、受験者は自身が予約した日時に受験する。予約は実施3日前まで可能である。

テスト会場では、受付時に公的なID等で本人確認が行われ、受験者はボディーチェックを受け、荷物や所持物をすべてロッカーに預けてから、CBT受験についての説明を受けて、会場内の指定されたブースに案内され、受験を開始する。試験時間中は監督員が常駐している。

受験者は、ブース内でコンピュータの画面に表示される問題やヘッドフォンから流れる音声をもとに、画面上で解答する。画面内でテストに関する守秘義務への同意を行った上で、任意のタイミングで受験を開始し、「文字と語彙」「会話と表現」「聴解」「読解」のセクション順に受験するが、セクションごとの解答時間はなく、60分の制限時間内に全セクションの問題を自分のペースで解答する。出題される問題は、受験者によって出し分けられている。

図7は問題画面イメージを示したものである。コンピュータ画面上にテストの問題および選択肢が表示される。設問は英語で表示されており、「Your Language」ボタンをクリックすると、

9言語の現地語の設問が表示される。受験者は、自分がわかる言語で設問を読み、場面、状況を理解した上で、解答することができる。セクション内であれば、解答し直すことも可能であるが、聴解セクションでは、前後の問題に移動して再度音声を聞くことはできない。

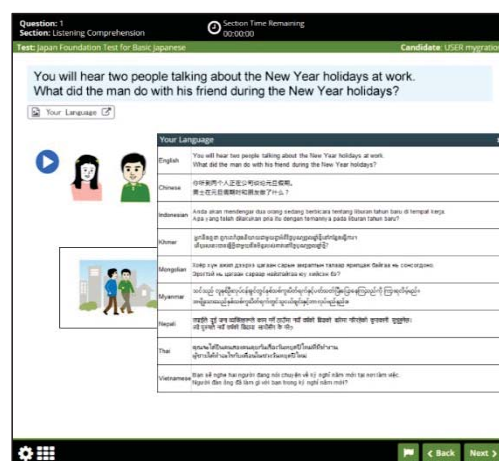


図7 問題画面イメージ

6.2 テスト結果の通知

本テストでは、テスト終了時に画面内に総合得点と判定結果が表示されるため、受験者は受験後即時に判定結果を確認できる。また、正式な判定結果通知書は受験後予約ウェブサイト上で受験者本人が確認、印刷できるようになる。

判定結果通知書には、受験者情報、受験情報、総合得点とそれに基づく判定結果、セクションごとの正答率が記載される⁽¹³⁾。総合得点は正答数そのままの素点ではなく、等化によって算出された尺度得点で表示される。尺度得点は、出題されるテスト問題の難易度の差等に関わらず、一定の能力値を示すものである。総合得点の得点範囲は10～250点で、総合得点が判定基準点（200点）以上のとき、「ある程度日常会話ができ、生活に支障がない程度の日本語能力水準に達している」と判定される。

6.3 CBT 方式のテストにおける工夫

CBT 方式のテストにおいて工夫した点は、次の2点である。

まず、設問の場面や状況設定を、英語や現地語で理解した上で解答できるようにした点である。設問は翻訳された際に場面や状況がわかりやすいように基本構成を定め、意図に沿った翻訳となるよう、翻訳の質にも留意している。

次に、CBT 方式に慣れていない受験者への対応である。特定技能の特定業種は、介護、建設、農業など、コンピュータを使う事務仕事を中心ではない業種が多いため、受験者が CBT の受験や操作に不慣れであることが懸念された。そのため、テスト開始画面内に「Your Language」の表示の仕方、音声再生ボタンやボリューム調整など操作に関する説明を加え、英語または現地語で確認してから受験できるようにした。また、画面上でエラーメッセージなど英語のみで表示される文言に関しては、ブース内に現地語の補助資料を設置するなどの配慮をした。問題画面も CBT 仕様で可能な範囲で、画像の大きさや解像度、文字の大きさやフォント、最小限のスクロールで解答できる問題文や選択枝の配置など、わかりやすく表示する工夫も行った。

また、受験者や関係者向けに、JFT-Basic ウェブサイトの「JFT-Basic とは」にテストの説

明を記載するとともに、CBT 画面や問題画面の説明を動画にした「操作動画」、実際のテスト画面に近い形式で全カテゴリーの問題を体験できる「サンプル問題」を掲載し、JFT-Basic がどんなテストなのか、事前に確認、イメージしてから、受験できるようにした。

7. 結果と今後の取組み

JFT-Basic の開発では、先に述べた要件 (1) について、国際的標準である CEFR に基づき、日本での生活場面に絞り、A1-A2の JFS の具体的な Can-do を元に、真正性のある問題を作成することで、「生活場面のコミュニケーションに必要な言語能力」を測定し、A2レベルの一定程度の水準に達しているかを判定するテストとして、高い信頼性、妥当性を備えたテストを実現した。要件 (2) について、CBT 方式での実施により、頻回受験、即時結果通知、本人確認やカンニング、漏洩防止などを備えるとともに、CBT 方式に馴染みのない受験者に対する受験時の対応、事前対応も行うことで、円滑な実施を実現した。

JFT-Basic は2019年4月から2020年8月現在まで、累計6か国（フィリピン、カンボジア、インドネシア、ネパール、モンゴル、ミャンマー）において、計10回実施され、10,346人が受験し、基準点到達率平均は35.3%である⁽¹⁴⁾。本実施からも安定実施を続けており、各回の結果の統計分析においても、品質評価、3レベル分析、テスト情報量分析を継続し、テストが高い信頼性、品質であることを検証している⁽¹⁵⁾。また、継続的な分析によって、数値の悪い問題を修正するなど、テストの品質を維持しているほか、傾向を確認し、出題方針や基準を精緻化することで、よりよい新規問題の作成、実施に向けて継続的に取り組んでいる。

実施時には、受験者に JFT-Basic に関するアンケートや聞き取り調査も行ってきたが、上記の工夫の結果、問題の難易度、受験時間、画像や翻訳などのわかりやすさについて概ね好意的な評価が得られ、CBT 方式の受験についても、「操作がわかりやすい」「設問が自分の言語なので、よく理解して解答できる」など、受験の大きな妨げにはなっていない様子も確認できた⁽¹⁶⁾。また、実施国の現地教師の訪日研修において、JFT-Basic を紹介した際には、「テストのための準備が即日本での生活に役立ちそう、テストを受ける意味がある」「生活の Can-do ベースなので、必要な日本語や社会や文化も含め、背景知識として何を教えればいいのかイメージできる」などの声があった。

主として就労のために来日を希望する受験者にとって、JFT-Basic が日本の生活場面でのコミュニケーションをイメージし、それに向けて準備することが来日後の生活に必要なことを学ぶ近道になるというのは、JFT-Basic の開発にあたり、当初から目指していたことであった。ハイスタークスのテストは、テスト自体が新しい学習の在り方を促すインセンティブとなるといふ社会的な影響力があることを考慮して、開発、運用する必要がある。

国際交流基金では、新たな外国人材の受入れ関連事業として、①JFT-Basic の開発・実施、

②Can-do や教材の整備、③現地の教師・機関支援、④現地教師訪日研修を行っている。②については、2019年9月に「JF 生活日本語 Can-do」を、2020年3月に『いろどり 生活の日本語』をウェブ上で公開し、③については、生活日本語コーディネーターや日本語専門家の派遣による現地支援活動、④については、現地教師の訪日研修も開始している。今後、JFT-Basic のさらなる説明や広報、Can-do と教材や e ラーニングの開発、現地の日本語教育支援や現地教師訪日研修など、新たな外国人材に対する日本語教育支援の充実により、学んだことがテストで測られ、日本での生活に役立つ、という一貫した取り組みが、外国人材受入れの一助となることを願い、今後もよりよい作題、安定実施に向けて JFT-Basic の開発、運用を続けていきたい。

〔注〕

- (1) 2019年9月に公開された JFT-Basic ウェブサイト (国際交流基金 2019a) の「JFT-Basic とは」にテストの説明が掲載されている。
- (2) CBT は一般的にコンピュータを使用して出題されるテスト形式を示すが、ITC (2005) では、テスト自体とテストの開発から実施、採点、評価等までを含む過程全体をテストング (testing) と呼び区別している。本稿でも、「CBT」はコンピュータを使用して出題されるテスト形式を、「CBT 方式」は受験予約、テスト会場での問題配信、自動採点、結果通知までの一連のテスト実施方式のことを示す。
- (3) 法務省のホームページ上に「新たな外国人材の受入れおよび共生社会実現に向けた取組 (在留資格「特定技能の創設等」)」についての情報が掲載、更新されている。
- (4) 「特定技能1号」の申請には、JLPT の N4 レベル以上の証明も利用できる。JLPT と JFT-Basic の違いは、JLPT は多様な背景の学習者を対象とし、幅広い場面を想定したテストであり、N5-N1 の各レベルのテストが国内外で年2回の特定の試験日に PBT (Paper Based Testing) 方式で実施され、合否結果が確認できるのは2か月後であるのに対し、JFT-Basic は主として就労のために来日する外国人を対象とし、日本国内での生活場面を想定した基礎段階 (A2) の能力を判定するテストであり、海外で年6回 (日本国内でも2021年3月より開始)、設定期間中の複数日程に CBT 方式で実施され、受験機会が多いこと、受験当日に結果が通知されることなどである。
- (5) 磯村・三矢 (2011) は JF の日本語講座のクラス分けに利用する会話、読解、作文のレベルチェックテスト、熊野ほか (2013) は JF の日本語講座で共同利用できる聴解、読解、作文、会話の A1 レベル認定試験の開発、国際交流基金 (2017b) はロールプレイトスクによる口頭テストの開発について報告している。
- (6) 「JF にほんご e ラーニング みなと」 (<https://minato-jf.jp>) をはじめとした e ラーニング教材の開発、運用において、オンラインコース制作の一環として CBT の制作、運用経験がある。武田・千葉 (2019) では、まるごと日本語オンラインコースの受講前にオンラインで受験し、自分のレベルにあったコースがどれか確認できる「おすすめコース診断テスト」の開発、運用結果を報告している。
- (7) 「国際的標準」は言語試験等の開発、検証において国際的に参照される CEFR を想定した。
- (8) 本テストは、テストングや統計の専門家、日本語教育の有識者からなる事業委員会を定期的に開催し、そこでの議論、助言を踏まえて、開発、実施している。
- (9) CBT の配信仕様やテスト会場コンピュータの言語設定等により、「作文」は文字のタイピングが困難であること、「会話」は他試験と同時にされる会場で発話が困難であること、双方とも採点が人によるため時間を要し、即時自動採点や結果通知が困難であることなど。
- (10) Modified Angoff 法は TOEFL と CEFR のレベルマッピングにも使われた統計的ワークショップ手法である。CEFR のレベルに詳しい専門家12名が、各問題について最低合格水準者が100名いたとして何名が正答するかの予測値を記入し、予測値と受験結果を統計的に分析し、推測される推奨基準点が導き出された。
- (11) 「文字と語彙」セクションの Can-do 例は、「語の意味」「語の用法」カテゴリーでは主に活動 Can-do 全

般から、「漢字の読み」「漢字の意味と用法」カテゴリーでは特に受容（読む）、産出（書く）、やりとり（文書）から選定しているが、場面・トピックを示すためのもので、あくまで参考情報という位置づけである。

- ⁽¹²⁾伊藤（2017）はCEFR、JFSの読字能力以外の言語構造能力Can-doを参考に読字能力のCan-doを試案した。「漢字の読み」で測る読字能力は伊藤が提示した「よく知っているものの範囲であれば目にしたものを単語として読むことができる」というA2レベルのCan-doとも合致する。
- ⁽¹³⁾JFT-Basicでは、セクションごとの下限点はなく、4セクション全体で総合的に測定、判定される。
- ⁽¹⁴⁾各試験の日程、実施国・都市などの実施概要、申込者数、受験者数、基準点到達率などの実施結果については、JFT-Basicウェブサイトの「実施結果」を参照のこと。
- ⁽¹⁵⁾本実施の計10回のテスト結果分析においても、テストの高い信頼性が確認されている。また、2019年に実施したプレテスト、4-9月（フィリピン1カ国で実施）、10-11月テスト（複数国にて多言語ポップアップを導入して実施）の品質評価の結果を比較し、3つの試験回の受験者層の違い等により、テスト結果や各問題の解答傾向に大きな差はないことを確認した。
- ⁽¹⁶⁾アンケート、聞き取り調査のうち、CBT形式での受験を受験者がどう捉えたかについては、八嶋ほか（2020）で具体的な調査結果を報告している。

〔参考文献〕

- 葦原恭子・小野塚若菜（2014）「高度外国人材の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
- 磯村一弘・三矢真由美（2011）「JF日本語教育スタンダード「みんなのCan-doサイト」を用いたレベルチェックテストの作成」『ヨーロッパ日本語教育』16、171-175
- 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、日本語教育学会、55-62
- 今井新悟・赤木彌生・中園博美（2012）『J-CAT オフィシャルガイド—コンピュータによる自動採点日本語テスト—』、ココ出版
- 外務省（2020）「入管法改正による新しい在留資格 特定技能の創設」
<<https://www.mofa.go.jp/mofaj/ca/fna/ssw/jp/index.html>>（2021年2月25日）
- 熊野七絵・伊藤秀明・蜂須賀真希子（2013）「JFS/CEFRに基づくJFS日本語講座レベル認定試験（A1）の開発」『国際交流基金日本語教育紀要』9、73-88
- 国際交流基金（2017a）『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』
- 国際交流基金（2017b）『JF日本語教育スタンダード準拠ロールプレイテストテスター用マニュアル』
- 国際交流基金（2019a）「JFT-Basic 国際交流基金日本語基礎テスト」ウェブサイト
<<https://www.jpff.go.jp/jft-basic/index.html>>（2020年8月30日）
- 国際交流基金（2019b）「JF生活日本語Can-do」
<https://www.jpff.go.jp/j/urawa/j_rsorcs/seikatsu.html>（2020年8月30日）
- 酒井たか子・加納千恵子・小林典子（2015）「第5章 TTB」『日本語教育のための言語テストガイドブック』、くろしお出版、86-109
- 出入国在留管理庁（2020）「特定技能総合支援サイト」<<https://www.ssw.go.jp/>>（2021年2月25日）
- 武田素子・千葉朋美（2019）「まるごと日本語オンラインコース『おすすめコース診断テスト』の開発」『国際交流基金日本語教育紀要』15、39-54
- 豊田市（2015）「第14章 とよた日本語能力判定」『とよた日本語学習支援システム ガイドライン』
<https://www.city.toyota.aichi.jp/_res/projects/default_project/_page_/001/003/583/13_14guideline.pdf>（2021年2月25日）
- 東伴子・代田智恵子（2019）「行動中心アプローチにもとづいたヨーロッパにおける日本語オンラインテストの開発—」『ICT×日本語教育』、ひつじ書房、150-165
- 法務省（2018）「特定技能制度」<http://www.moj.go.jp/isa/policies/ssw/nyuukokukanri01_00127.html>（2021年2月25日）

「国際交流基金日本語基礎テスト」の開発

- 八嶋康裕・熊野七絵・千葉朋美・廣利正代 (2020) 「JFT-Basic 受験者は CBT 形式での受験をどう捉えたか」『2020年度日本語教育学会秋季大会予稿集』、日本語教育学会、132-137
- ALTE (2011). *Manual for Language Test Development and Examining*. Council of Europe.
<https://www.alte.org/resources/Documents/ManualLanguageTest-Alte2011_EN.pdf> (2021年2月25日)
- Council of Europe (2001). *Common European Framework of Reference for Languages : Learning, teaching, assessment*. Cambridge. Cambridge University Press.、吉島茂・大橋理枝 (訳、編) (2004) 『外国語の学習、教授、評価のためのヨーロッパ共通参照枠』初版第一版、朝日出版社
- International Test Commission (2005). *International Guidelines on Computer-Based and Internet Delivered Testing*. <<http://www.intestcom.org>> (2020年8月30日)、日本テスト学会 (訳) (2017) 『CBT とインターネット配信テストニングに関する ITC ガイドライン日本語版』