

## まる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

武田素子・千葉朋美

〔キーワード〕 『まると 日本のことばと文化』、テスト開発、JF 日本語教育スタンダード、eラーニング

### 〔要 旨〕

国際交流基金が運営している日本語学習のためのプラットフォーム「JF にほんご eラーニング みなど」では、日本のことばと文化を総合的に学ぶ「まると日本語オンラインコース」を開講している。本稿では、初年度は A1のみであった「まると日本語オンラインコース」に A2のコースを追加するにあたり、ユーザーがいつでもオンライン上で気軽に受けることができ、自身の日本語能力にあったコースを選ぶための目安となるよう開発した「おすすめコース診断テスト」の開発過程について述べる。本テストは「もじとことば」「かいわとぶんぼう」「ちょうかい」「どっかい」の4セクションで構成され、問題は CEFR を参照し、各コースで扱われている Can-do から出題するのに相応しいものを選定しながら作成した。本テストの受験者7307名のデータを分析した結果、高い信頼性が得られ、ユーザーにとってコース選びの一助となり得ることがうかがえた。

### 1. はじめに

国際交流基金は2016年7月より日本語学習のためのプラットフォーム「JF にほんご eラーニング みなど」(以下、「みなど」)において、日本のことばと文化を総合的に学ぶ「まると日本語オンラインコース」(以下、まるとコース)を開講している(武田ほか 2017)。まるとコースは、国際交流基金が開発したコースブック『まると 日本のことばと文化 かつどう』及び『まると 日本のことばと文化 りかい』(以下、『まると』)のカリキュラムとシラバス、素材を基に開発されたオンラインコースで、2018年9月現在、『まると』の「入門 (A1)」にあたる A1-1、A1-2、「初級1 (A2)」にあたる A2-1、A2-2、「初級2 (A2)」の1～10課にあたる A2-3の計5つのコースを開講しており、11課～18課にあたる A2-4コースは、2018年10月に開講予定である(表1)。

表1 「みなと」で開講しているまるごとコース (2018年10月時点)

レベル	解説言語	オンラインコース	コースブック	
A1	英語、インドネシア語、タイ語、 スペイン語、ベトナム語	まるごと A1-1 (かつどう)	入門 (A1)	1課-10課
		まるごと A1-1 (かつどう・りかい)		11課-18課
		まるごと A1-2 (かつどう)		
		まるごと A1-2 (かつどう・りかい)		
A2	英語、スペイン語	まるごと A2-1 (かつどう・りかい)	初級1 (A2)	1課-10課
		まるごと A2-2 (かつどう・りかい)		11課-18課
	英語	まるごと A2-3 (かつどう・りかい)	初級2 (A2)	1課-10課
		まるごと A2-4 (かつどう・りかい)		11課-18課

「みなと」では、各種日本語オンラインコースが開講されており、各コースのコース概要ページに記載されたコース目標 Can-do や JF 日本語教育スタンダード (以下、JFS)<sup>(1)</sup>に基づいて表示された A1~C2の日本語レベルなどの情報から、ユーザーは自分にあったコースを選び、受講することができるようになっている。初年度は A1レベルのコースのみであったまるごとコースに A2レベルのコースを追加するにあたり、オンラインでの自学自習という状況において、ユーザーがどのコースを受けたらいいのか迷うことが予想された。そこで、ユーザーがまるごとコースを選択する際に自分の日本語能力にあった適切なコースを判断する手段の一つとして「おすすめコース診断テスト」(以下、本テスト)を開発し、「まるごとコース紹介ページ」(<https://www.marugoto-online.jp/info/>)に追加することとした。本稿では、本テストの開発の過程を示し、テスト公開後の利用状況や分析結果について報告する。

## 2. 調査と開発の方針

### 2.1 各国語のおすすめコース診断テスト調査

本テストの開発方針を検討するため、まず各国語の類似テストを調査し、どのような構成・設問で作成されているか分析した。調査は、自身の語学力を測り、それに基づいて受講すべき自学自習のオンラインコースやラジオ講座などを判断する目安とするためのテストで、インターネット上でいつでも無料で受けられるものを対象とした。調査を行った主なテストは表2に示した通りである。調査の結果、ユーザーがあるレベルに到達しているかどうかの判断としては、ユーザーがレベルを選んでレベルごとに違う問題を解くものと、全レベル共通の問題を解くものの2タイプあることがわかった。多くのテストが文法、聴解、読解など複数のセクションで構成されており、問題の形式は多肢選択式が採用されていることが多く、自由作文などの産出の課題は少なかった。受験時間については、アンスティチュ・フランセ日本のフランス語テストの読解と文作成に制限時間が設定されていた以外は、時間制限は設けられていなかった。また、Instituto Cervantes のスペイン語テストはアダプティブテストであり、British Council

の英語テストでは解答する際に確信度を「Certain」「Fairly Sure」「Not Sure」の3レベルで尋ねていたが、多くのテストでは解答の正誤によって判定を出しているようであった。

筆者らは各国語のテストを実際に受験してみることで、テストの目的は受験後のオンラインコースなどの受講であるため、気軽さが重要で、ユーザーにとって負担にならない問題数で判定できるものがよいこと、受験方法や問題の解き方が複雑でない方がよいなどの気付きを得た。

表2 各国語のテスト調査対象一覧

実施機関 実施機関の URL	言語	テスト名	テストの問題	テストの構成
NHK 出版 <a href="https://eigoryoku.nhk-book.co.jp/">https://eigoryoku.nhk-book.co.jp/</a>	英語	英語力測定テスト 2016 <sup>(2)</sup>	レベル別 (基礎編、応用編)	文法 5問 会話・表現 5問 リスニング 5問
British Council <a href="http://learnenglish.britishcouncil.org">http://learnenglish.britishcouncil.org</a>	英語	free online English test	全レベル共通 (A1～C2)	文法・語彙 25問
アンスティチュ・フラン セ日本 <a href="http://e-francais.institutfrancais.jp">http://e-francais.institutfrancais.jp</a>	フランス語	レベルチェック・ テスト	全レベル共通 <sup>(3)</sup> (A1～B2)	聴解 大問3問 (小問10問) 読解 大問4問 (小問14問) 文法・作文 大問2問 (小問5問)
Deutsche Welle <a href="https://learngerman.dw.com">https://learngerman.dw.com</a>	ドイツ語	Placement Test	レベル別 (A1、A2、B1)	聴解 大問8問 (小問22問) 読解 大問6問 (小問13問) 文法・表現 大問16問 (37問)
Goethe Institut <a href="https://www.goethe.de">https://www.goethe.de</a>	ドイツ語	あなたのドイツ語 力を試してみよう	全レベル共通 (A1～C2)	聴解 大問4問 (小問10問) 読解 大問6問 (小問13問) 文法・語彙 大問2問 (小問7問)
Instituto Cervantes <a href="https://ave.cervantes.es">https://ave.cervantes.es</a>	スペイン語	PROFICIENCY TEST	セクションごとの アダプティブ	語彙・構造 30問 読解 大問1問 (小問5問) 聴解 大問1問 (小問5問)

## 2.2 開発の方針

本テストの目的は、ユーザーの受験時点での日本語能力を測り、レベル認定を行うことではなく、ユーザーの日本語能力にあったまるとコースをおすすめすることである。

まるとコースは『まると』のカリキュラム・シラバスに沿って開発されている。『まると』は JFS に準拠して制作されたコースブックであり、レベルは JFS の6段階 (A1～C2) の言語熟達度によって示されている。また、日本語を使って何がどのようにできるかという能力 (課題遂行能力) に重点をおき、各課の学習目標は Can-do の形で表されている (来嶋ほか 2012)。本テストも、このような『まると』の枠組みに沿い、開発の方針を以下のように定めることとした。

ア. JFS の言語熟達度の尺度に沿って、まるとコースで開発が予定されている『まると』入門 (A1)、初級1 (A2)、初級2 (A2) の3段階の判定を出す。

イ. 設問は個々の言語知識を問うものではなく、まるとで学習目標として設定されている

課題遂行能力に重点をおいた内容とする。

ウ. オンラインで自学自習をしているユーザーを対象としているため、インターネット上でいつでも受験でき、気軽に受けられる構成とする。

### 3. テストの開発

テストの開発にあたっては、磯村ほか (2011)、熊野ほか (2013) の試験開発のプロセスを参考に、テスト全体の構成と設問形式の検討をし、CEFR (Common European Framework of Reference for Languages: Learning, teaching, assessment) の共通参照レベルを参照しながら、各レベルでの出題に相応しい『まるごと』の Can-do の選定を行った。その後、選定した Can-do に基づいて問題を作成し、全セクションの問題がそろった段階で試行を行い、試行結果の分析と修正の工程を経て完成させた。以下、各工程で具体的にを行った作業について報告する。

#### 3.1 テスト開発過程

##### ①テストの構成の検討

まるごとコースでは、各コースの最後に到達度を測るために終了テストを設けている。『まるごと』の活動編はコミュニケーション言語活動を目標に、理解編はコミュニケーション言語能力を養うもので、文法・文型・語彙などを文脈の中で理解し、使えるようになることを目標にしている (来嶋ほか 2014)。まるごとコースの終了テストは、活動編と理解編をあわせて学んだ後に受験するものであるため、該当課の言語活動 Can-do と、それを支える言語構造能力の確認ができるよう、両編で扱われている場面・文脈の中で学習項目を問うという原則で作成している。具体的には、『まるごと』の理解編の「テストとふりかえり」部分を参考に、「もじとことば」「かいわとぶんぼう」「ちょうかい」「どっかい」の4セクションで構成している。

本テストも、まるごとコースの終了テストと同様に「もじとことば」「かいわとぶんぼう」「ちょうかい」「どっかい」の4セクションで構成し、気軽に受けることができるよう、最大でも30問程度に収められるよう構成を考えていくこととした。

##### ② Can-do の選定(「ちょうかい」「どっかい」)

気軽に受験することができることを考慮した場合、各段階 (入門 (A1)、初級1 (A2)、初級2 (A2)) から出題できる問題数は、セクションごとに2問程度となる。『まるごと』で扱われているどの Can-do から出題するのが適当かを検討するため、まず、JFS が熟達度の尺度として採用している CEFR の「共通参照レベル：全体的な尺度」と、「共通参照レベル：自己評価表」の能力記述文を確認した。

次に、『まるごと』の入門 (A1)、初級1 (A2)、初級2 (A2) で扱われている全152の Can-

doを「話しことば（受容・産出・やりとり）」と「書きことば（受容・産出・やりとり）」のカテゴリにあわせて分類を行った<sup>(4)</sup>。

Can-doを一覧表にし、各レベルでの言語活動がどのように変わるのかを概観した上で、「話しことば（受容・産出・やりとり）」に関するCan-doから「ちょうかい」を、「書きことば（受容・産出・やりとり）」に関するCan-doから「どっかい」の出題に適したCan-doを、言語活動のカテゴリとトピックに偏りが出ないように選んだ。表3は、各コースで扱われているCan-doの中から「ちょうかい」と「どっかい」で出題するCan-doとして選定したものである。

表3 「ちょうかい」「どっかい」で選定したCan-do

コース	ちょうかい	どっかい
A1-1	Can-do5（ <u>社交的なやりとりをする</u> ） パーティーやイベントで初めて会った人に、名前、出身、仕事などをたずねたり、答えたりすることができる。	
	Can-do25（ <u>情報交換する</u> ） 友人と会う日を決めるときに、次週のスケジュールを短い簡単な言葉で教え合うことができる。	
A1-2		Can-do30（ <u>必要な情報を探し出す</u> ） 地域で有名な祭りなど、催し物のポスターを見て、開催日、場所など、ごく基本的な情報を探し出すことができる。
		Can-do44（ <u>情報や要点を読み取る</u> ） ごく短い簡単な文で書かれたブログを見て、写真などを手がかりに、ブログを書いた人が何をしたら、どこへ行ったかなどを理解することができる。
A2-1	Can-do12、13、14（ <u>情報交換する</u> ） 道に迷ったとき、目的地への行き方について、短い簡単な言葉で人に質問したり、説明したりすることができる。	Can-do16（ <u>手紙やメールを読む</u> ） 待ち合わせをしている友人からの遅刻を知らせる短い簡単なメールを読んで、内容をだまかに理解することができる。
A2-2	Can-do27（ <u>共同作業中にやりとりをする</u> ） 友人とピクニックなどの準備をするために、だれが何を持っていくかなど、短い簡単な言葉で確認や指示をしたり、受けたりすることができる。	Can-do36（ <u>必要な情報を探し出す</u> ） 出張者のスケジュール表などの短い簡単なテキストを見て、会議や視察など、同行に必要な情報を探し出すことができる。
A2-3	Can-do22（ <u>店や公共機関でやりとりをする</u> ） テーマパークの案内所で、係員にイベントの時間や場所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。	Can-do27（ <u>手紙やメールを読む</u> ） <sup>(5)</sup> 友人からの年賀状や誕生日カードに書いてある「あけましておめでとうございます」「お誕生日おめでとうございます」など、定型の簡単なメッセージを読んで、理解することができる。
A2-4	Can-do41（ <u>共同作業中にやりとりをする</u> ） 職場などで、電気の消し忘れや紙の無駄遣いなど、同僚の不適切な行動を短い簡単な言葉で注意したり、注意されたときに対応したりすることができる。	Can-do37（ <u>情報や要点を読み取る</u> ） 短い簡単な文で書かれていれば、観光地においてあるノートに書かれた旅行者のコメントを読んで、内容をだいたい理解することができる。

東ほか（2017）では、自動採点のコンピューターベースのテストでは、選択式の問題による受容能力測定が中心となり、一見コミュニケーション／行動中心アプローチ的な行動と相反するよう見えるが、学習者が日本語を使う実際のコミュニケーション場面を想定し、その状況やインターアクションの流れを考慮することにより、オーセンティックな活動場面に近づけることはできるとしている。「ちょうかい」の作成にあたり、「道に迷ったとき、目的地への行き

方について、短い簡単な言葉で人に質問したり、説明したりすることができる。」(図1)のように、ユーザーが聴き手の一人となり課題を達成する内容となるよう留意して作成した。また、熊野ほか(2013)で、聴解で文字を選択肢にした場合、A1レベルの学習者にとって聴解能力とは別の文字を読む能力が必要になると指摘されていることから、「ちょうかい」の選択肢は全てイラストや写真とすることにした。

また、「どっかい」では、CEFRの「共通参照レベル：自己評価表」の「理解すること(読むこと)」のA1~A2レベルで示されているように、A1では「掲示やポスター、カタログ」、A2では「ごく短い簡単なテキスト、広告や内容紹介のパンフレット、メニュー、予定表、簡単に短い個人的な手紙」が読む対象となるよう選び、問題を作成した(図2)。



図1 「ちょうかい」の問題例 (A2-1)



図2 「どっかい」の問題例 (A2-2)

### ③「基本文」の選定(「かいわとぶんぼう」)

「ちょうかい」と「どっかい」で出題するCan-doを選定した後に、「かいわとぶんぼう」で出題する項目を選んだ。『まるごと』の「かいわとぶんぼう」は、モデル会話で場面と文脈を理解した上で、コミュニケーションのために使われている言語構造の学習へと誘導するよう作成されている。理解編で学習項目として示されている基本文のうち、各コースの「話すこと(やり取り・表現)」のCan-doを達成するために必要な新出表現で、ルールを知っていなければ答えられない可能性が高い項目を選び、出題することとした。

表4は、出題項目として選んだ『まるごと』の基本文と、その基本文が学習項目として挙げられているCan-doを一覧にまとめたものである。出題する項目の選定にあたっては、「ちょうかい」と「どっかい」のトピックやCan-doとの重なりは避けるようにした。また、まるごとコース内の終了テストでは「かいわとぶんぼう」部分は、学習したことの確認として、空欄補充とタイピングの両形式で出題しているが、本テストでは空欄補充の形式での出題とした(図



図3 「かいわとぶんぼう」の問題例 (A1-2)

3)。空欄補充のみにした理由は、「かいわとぶんぼう」セクションは、日本語でのタイピング力を測るものではなく、該当する会話にあった表現（基本文）が使えるかどうかを測るための問題であり、タイピングミスによる失点を防ぐためである。

表4 「かいわとぶんぼう」で選定した基本文と Can-do

コース	『まると』の基本文	Can-do
A1-1	いえに エアコンが あります。 いえに ねこが います。	17 友人に自分の家を説明するとき、「部屋が一つあります」「テレビはありません」など、簡単な言葉で言うことができる。
A1-2	きのう デパートに 行きました。	5 友人や近所の人に、休みの日にどこへ行ったかたずねたり、「楽しかったです」などの簡単な感想を交えながら、答えたりすることができる。
A2-1	私は 東京に すんでいます。 私は でんしゃの かいしゃで はたらいて います。	1 初めて会った人の前で自己紹介するとき、自分や家族がどこに住んでいるか、何をしているかなど、短い簡単な言葉で話すことができる。
A2-2	たなかさんは 楽しかったと 言っています。	50 友人と、共通の友人の誕生日プレゼントを買うために、買う物や予算などについて、短い簡単な言葉で話し合うことができる。
A2-3	おしろの 中が 見られます。	16 ホテルのフロントなどで、ツアーを紹介してもらうために、興味がある活動や、行きたい観光地名を言うことができる。
A2-4	私は 母に しまわれました。 私は 先生に 絵を ほめられました。	48 子供時代の習い事や学校生活などについて、短い簡単な言葉で友人に話すことができる。

④ことば、漢字の選定(「もじとことば」)

「もじとことば」では、各レベルの言語活動に特徴的なトピックを選び、そのトピックに必要な語彙を出題することとした(表5)。CEFRの使用語彙の領域のCan-doを参考に、A1では「人物や場所」、A2では「仕事や自由時間に関わる身近な日々の事柄」などのトピックを中心に、他のセクションで出題するトピックと重ならないように出題項目の選定を行った。しかしながら、「もじとことば」は他セクションと比較して、その語彙を知っているかどうかという言語知識に大きく左右されるため、「ことば」「漢字」を6問ずつにして出題項目を増やし、1問当たりの配点を他のセクションより低くすることとした。

表5 「もじとことば」で選定したトピックと語彙

コース	CEFR Can-do	ことば(出題項目)	漢字(出題項目)
A1-1	・特定の具体的な状況に関して、基本的な単語や言い回しのレパートリーを持っている。ただしそれらの間の繋がりはない。(A1)	T1 わたし(ちち)	T2 たべもの(肉)
A1-2		T7 まち(バス)	T7 まち(南口)
A2-1	・生活上の単純な要求に対応できるだけの語彙を持っている。(A2.1)	T2 きせつと天気(さむい)	T5 外国語と外国文化(大学)
A2-2		T3 私の町(あかい)	
A2-2	・基本的なコミュニケーションの要求を満たすことができるだけの語彙を持っている。(A2.1)		T8 けんこう(頭)
A2-3		T1 新しい友だち(なきます)	T2 店で食べる(注文)
A2-4	・馴染みのある状況や話題に関して、日常的な生活上の交渉・取引を行うのに十分な語彙を持っている。(A2.2)	T6 ネットショッピング(れいぞうこ)	T6 ネットショッピング(機能)

出題形式は、「ことば」は意味に関わる知識（形態と意味）を問う問題とし、画像を見て、その名称を答える四択の多肢選択問題とした（図4）。「漢字」はまるごとコースでは、「漢字の意味」と「漢字の読み」を扱っているが、本テストの他セクションでは該当コースで習う漢字は漢字で表記し、意味の確認も兼ねた問題としたので、本セクションでは「漢字の読み」のみを出題することとした。また、まるごとコースでは、『まるごと』でのなぞ

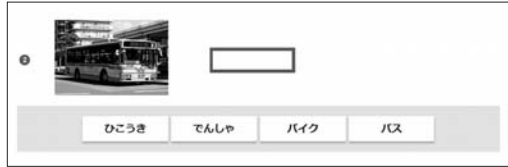


図4 「ことば」の出題例 (A1-2)

り書きに代わり、A1レベルよりタイピングの練習を取り入れているため、この部分をタイピング形式での出題とし、日本語でのタイピング力も測る形式で出題することとした（図5）。

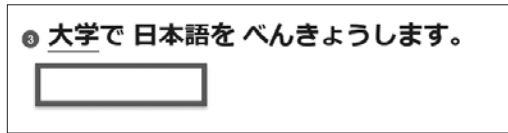


図5 「漢字」の出題例 (A2-1)

②～④の順で、出題項目の選定と問題の作成を行った後、全体での出題トピックのバランスの確認を行った。なお、問題の指示文は各国語とし、日本語表記は『まるごと』に則り分かち書き、該当コース既出のものは漢字で表記することとした。

### 3.2 テストの配点、判定基準

3.1で述べた手順で出題する問題を決めていき、問題数は「もじとことば」「かいわとぶんぼう」「ちょうかい」「どっかい」の4セクションで、合計33問となった。問題構成と配点は表6に示す通りである。

表6 本テストの構成

セクション コース		もじとことば		かいわと ぶんぼう (6問)	ちょうかい (6問)	どっかい (9問)	合計	
		ことば (6問)	漢字 (6問)					
入門 (A1)	A1-1	2点	2点	4点	4点	4点(2点×2)	16点	32点
	A1-2	2点	2点	4点	4点	4点	16点	
初級1 (A2)	A2-1	2点	2点	4点	5点	4点	17点	33点
	A2-2	2点	2点	4点	4点	4点(2点×2)	16点	
初級2 (A2)	A2-3	2点	3点	4点	4点	4点(2点×2)	17点	35点
	A2-4	2点	2点	5点	4点	5点	18点	
合計		25点		25点	25点	25点	100点	



表6に基づき、判定基準については、下記のようにした。

- ・ 0点以上32点未満 入門 (A1) (A1-1、A1-2コース)
- ・ 32点以上65点未満 初級1 (A2) (A2-1、A2-2コース)
- ・ 65点以上100点 初級2 (A2) (A2-3、A2-4コース)

なお、まるごとコースは総合的な日本語能力を身につけることを目標としたコースであるため、いずれかのセクションの点数が極端に低い場合は、より上の段階のコースを受講するのは難しいこと、本テストは多肢選択式のテストであるため実際の日本語能力に関わらず偶然正解してしまう場合もあるということを考慮し、判定基準に下記の条件をつけることとした。

- ア. 32点以上でも、1つでも0点のセクションがあった場合には、入門 (A1)
- イ. 32点以上でも、入門 (A1) レベルの問題の正答率が60%未満の場合は、入門 (A1)
- ウ. 65点以上でも、初級1 (A2) から出題した A2レベルの問題の正答率が60%未満の場合は、初級1 (A2)

上記の判定基準に基づき、実際のテストの判定画面ではおすすめのコースと、セクションごとの点数、受験にかかった時間を表示することにした (図6)。開始画面で名前を入力した場合 (任意) には判定画面にも名前が表示され、PDF 化や印刷をして、「まるごと教師サポート付きコース<sup>6)</sup>」などの受講生がその結果を教師に提出したりするといった利用もできるようにした。

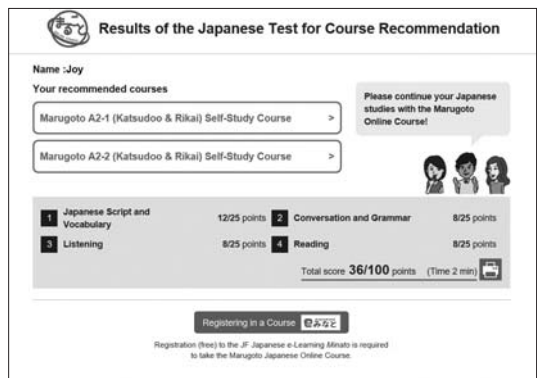


図6 テストの判定画面

### 3.3 テストの試行

国際交流基金関西国際センターで実施している研修の研修生に協力を依頼し、テストの試行を行った。対象は、国際交流基金関西国際センターで『まるごと』を使ってゼロから約5ヶ月間日本語を学んできた学習者18名で、試行を行った時点で『まるごと』の入門 (A1)、初級1 (A2) の全課 (1課~18課) と、初級2 (A2) の3課~10課までの学習を終えていた。

本テストはオンラインで行うものではあるが、実際のサイト上での画面作成は問題完成後に予定されており、受験環境は異なるものの問題自体の検証には問題ないと判断し、試行は紙媒体で行った。そのため、漢字の読みは手書きで、「ちょうかい」の音声は2回ずつ流す形で実施した。制限時間は設けず任意の時間としたが、最も短い人で10分、最も長い人で37分で提出

した。試行の結果は表7の通りである。

表7 テストの試行の結果概要

受験者数	最高点	最低点	平均点	標準偏差
18名	98点	47点	77.44点	13.84

テストの信頼性を調べるために、クロンバックの $\alpha$ 係数をもとめた。クロンバックの $\alpha$ 係数は、0.00から+1.00の数値となり、1.00に近ければ近いほど信頼性の高いテストだと解釈でき、一般的には0.80が一つの目安になるとされている(中村 2002)。本テストの試行の結果の $\alpha$ 係数は0.81であったため、信頼性において問題はないと判断した。

次に、項目分析のため、各問題の正答率と識別力を算出した(図7)。正答率は各問題の難易度を表す指標であり、「プレテのように、難易度が異なる項目で構成されたテストであれば、むずかしいと想定された問題群ほど数値は低く、やさしいと想定された問題群ほど高い数値で示されるのが理想」(伊東 2005:164)だとされている。また、識別力は一般的には識別力を表す点双列相関係数が0.30以上であれば、受験者を識別できている良問であるとされている(中村 2002)。つまり、識別力の高い問題というのは能力の高い受験者群と低い受験者群とをよく弁別することができる問題であるということである。

試行では識別力が0.3に満たない問題が15問(「もじとことば」7問、「かいわとぶんぼう」1問、「ちょうかい」3問、「どっかい」4問)見られた。そのうち、8問(「もじとことば」3問、「かいわとぶんぼう」1問、「ちょうかい」1問、「どっかい」3問)は正答率が100%で全員正解であったために識別力が0であった。また、1問(「もじとことば」)は正答率が0%で全員不正解であったために識別力が0であった。今回の試行では、同じ進度で学習している学習者に対して実施したため、このような結果となった可能性が高い。識別力が0.3に満たない問題のうち、イラストや指示文の曖昧さ、設問の順番が原因で受験者を惑わせていたと思われる2問については、問題の検討を行い、イラストと指示文の修正をして、テストを完成させた。

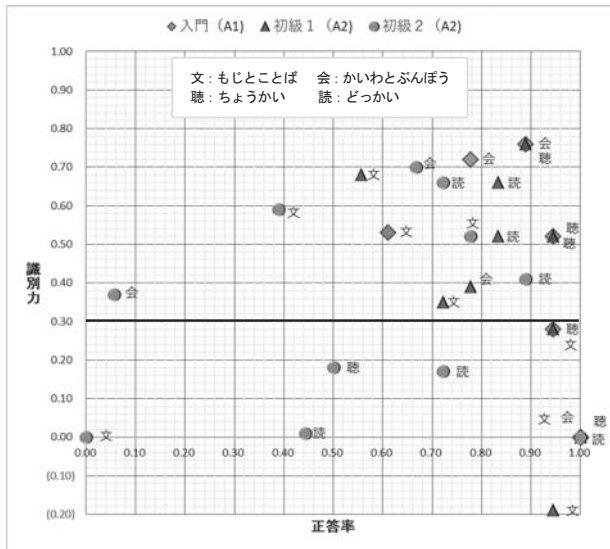


図7 試行の結果 各問題の正答率と識別力

## 4. テストの利用状況と分析

### 4.1 テストの利用状況

本テストは、2017年8月1日のA2-1コース公開と同時に、まるとコース紹介ページに追加し、どのコースから始めたらよいかわからないユーザーには、本テストの受験を勧めるようにしている。公開から2018年7月31日までの1年間で、延べ8564名の受験があった。受験時の解説言語は、公開時には日本語、英語のみであったが、2018年

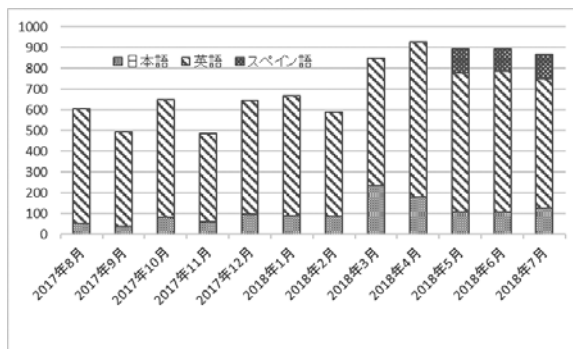


図8 テストの受験者数の推移

5月16日にスペイン語が追加され、現在解説言語は3言語備えている。受験時の解説言語と受験者数の推移は図8に示す通りで、公開以降受験者数は増加傾向にあることから、多くのユーザーに活用されていることがうかがえる。

また、本テストはパソコン、スマートフォンいずれの端末からも受験可能であるが、受験時の端末の比率は、パソコンとスマートフォンが2:1であり、スマートフォンからの受験も多く見られた。受験者全体の平均受験時間は16分37秒（標準偏差12分34秒）であり、本テストが気軽に受けられるテストとして、ユーザーに受け入れられていることが推察された。

### 4.2 テストの検証

本テスト1年間の延べ8564名の受験者のデータのうち、受験時の解説言語が日本語であった1257名は、日本語母語話者など本来対象としない人が受験している可能性があるため分析対象から除き、7307名を分析の対象とした。本稿で対象とした7307名の平均点は100点満点中51.78点で、標準偏差は28.00であった。また、各問題の出典となるコースの段階（入門（A1）、初級1（A2）、初級2（A2））別の結果概要は、表8に示す通りであり、問題の出典となるコースの段階が上がるにつれて、平均点が下がることが確認された。

表8 問題の出典となるコースの段階別に見たテストの結果概要

問題の出典となるコースの段階	項目数	配点	平均点 (標準偏差)	最小値	最大値
入門 (A1)	11	32	20.97点 (14.14)	0点	32点
初級1 (A2)	11	33	17.13点 (18.38)	0点	33点
初級2 (A2)	11	35	13.53点 (23.33)	0点	35点

これらのデータをもとに古典的テスト理論 (classical test theory: CTT) を用いて、テストの受験結果の分析を行った。まず、本テストの信頼性を調べるために、試行の時と同様にクロンバックの $\alpha$ 係数をもとめたところ0.95であり、本テストの信頼性は高いことがわかった。

次に、各問題項目を検証するために、識別力と正答率を算出した (図9)。識別力については、全ての問題が、一般的に良問と言われている0.3以上であり、受験者の日本語能力を識別するのに適した問題であったことがうかがえた。

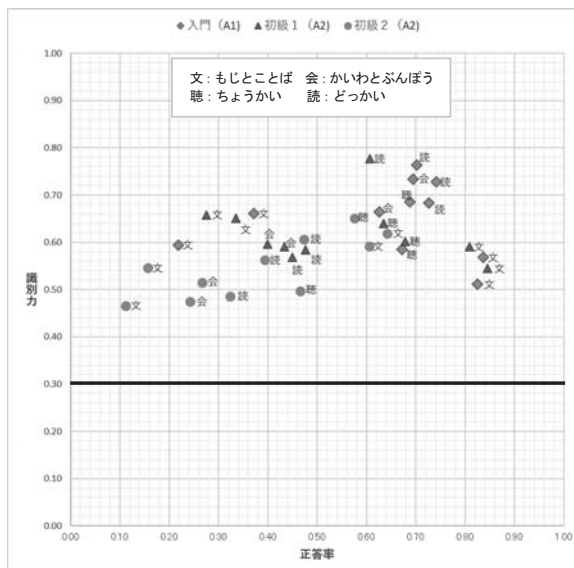


図9 テストの結果 各問題の正答率と識別力

問題の出典となるコースの段階別に、各セクションの平均正答率を確認すると、本テストの「かいわとぶんぼう」「ちょうかい」「どっかい」では、問題の出典となるコースの段階が入門 (A1)、初級1 (A2)、初級2 (A2) と上がるにつれて、平均正答率が下がっていたことから、各コースを受けるための能力を問うのに適当な問題であったことが推察された (表9)。一方、「もじとことば」では、入門 (A1) の平均正答率と初級1 (A2) の平均正答率にほぼ差がなかった。これは、入門 (A1) の問題として出題した「漢字の読み」の問題 (えきの 南口 にいます。) の正答率が22.00%で初級1 (A2) として出題した問題の正答率 (33.70%、27.80%) より低かったことと、初級1 (A2) の問題として出題した「ことばの意味」の問題 (あかい) の正答率が84.60%で入門 (A1) として出題した問題の正答率 (82.50%、83.70%) より高かったためである。この2間に関しては、問題の出典となるコースの段階にあうよう今後修正を検討する必要がある。

表9 問題の出典となるコースの段階別に見た平均正答率

問題の出典となるコースの段階	もじとことば	かいわとぶんぼう	ちょうかい	どっかい	全体
入門 (A1)	56.38%	66.30%	68.35%	72.73%	64.82%
初級1 (A2)	56.75%	41.95%	66.00%	51.5%	54.31%
初級2 (A2)	38.15%	25.85%	52.55%	40.20%	39.09%

### 4.3 おすすめコースの判定

本テストでは、3.2で説明した通り受験者の得点と設定した条件に基づいて、入門（A1）、初級1（A2）、初級2（A2）の3つの段階で判定を行い、おすすめのコースを診断している。本テストの問題と構成が、おすすめコースの診断に適切であったかを見てみたい。

まず、判定結果（おすすめコース）ごとに、テストの結果を確認したところ、7307名の受験者のうち、入門（A1）は2610名、初級1（A2）は2017名、初級2（A2）は2680名であった。判定結果別の平均点、標準偏差、平均受験時間は表10の通りである。

表10 判定結果ごとの該当者数と各セクションの平均点・平均受験時間（標準偏差）

判定結果 (おすすめコース)	該当者数	各セクションの平均点（各25点満点）				100点満点	平均 受験時間
		もじと ことば	かいわと ぶんぼう	ちょうかい	どっかい	合計	
入門（A1）	2610名	6.19点 (4.30)	3.52点 (4.90)	6.39点 (6.69)	4.78点 (5.61)	20.88点 (14.57)	10分10秒 (11分7秒)
初級1（A2）	2017名	11.72点 (3.34)	10.07点 (3.98)	17.40点 (5.21)	14.12点 (4.23)	53.31点 (8.24)	21分43秒 (12分27秒)
初級2（A2）	2680名	18.47点 (5.01)	18.76点 (5.05)	22.46点 (3.48)	21.03点 (3.64)	80.71点 (10.33)	19分4秒 (11分19秒)

いずれのセクションでも判定されたコースの段階が上がると点数が高くなっており、判定段階と平均点が比例していることがわかった。セクションごとの傾向を見てみると、「もじとことば」は他セクションと比べ入門（A1）と判定された人の平均点は高い方だが、初級2（A2）と判定された人の平均点は18.47点と最も低く、標準偏差も大きい。「もじとことば」は、作題上ある特定の語彙を知っているかという該当レベルの課題を遂行するために必要な言語知識の一部を問うことになる。課題達成に必要な語彙は、より上のレベルになれば個人によって異なり、個人が持っている知識の有無に影響を受けやすいためであると考えられる。「かいわとぶんぼう」は全体的に他セクションより平均点が低い傾向にあり、その傾向は入門（A1）と判定された人に特に顕著であった。そこで、入門（A1）と判定された人のデータを確認したところ、入門（A1）と判定された人の54.18%（1414人）が0点だった。CEFRの能力 Can-do 「文法的正確さ」では、A1レベルは「学習済みのレパートリーの中から、限られた、いくつかの単純な文法構造や構文を使うことはできる」、A2レベルは「いくつかの単純な文法構造を正しく使うことができるが、依然として決まって犯す基本的な間違いがある—例えば、時制を混同したり、性・数・格などの一致を忘れていたりする傾向がある。しかし、本人が何を言おうとしているのかはたいていの場合明らかである。」であり、Aレベルでは文法的な正確さは完全ではない。ただし、多肢選択式で正誤を出すという本テストの性質上、正確な文法の使用が求められることから、他セクションと比べて低い平均点となっていると考えられる。また、「ちょうかい」と「どっかい」は、入門（A1）から初級1（A2）までの平均点の開きが、初級1（A2）

から初級2 (A2) までよりも比較的大きい。この2つのセクションでは、個々の言語知識を問うものではなく、課題遂行能力により重点をおいた設問となっていることから、A1とA2の各レベルで可能な言語活動の違いが結果に反映されていると言えるだろう。

次に、3.2で設定した基準に基づき、本テストが適切におすすめコースを診断できているかを確認するため、判定結果 (おすすめのコース) 別に問題の出典となるコースの段階別問題群の平均点と標準偏差をもとめた。表11に示す通り、問題の出典となるコースの段階別に判定結果 (入門 (A1)、初級1 (A2)、初級2 (A2)) を見ると、判定結果が上がるにつれて、平均点も上がっていることがわかる。また、判定結果別に問題の出典となるコースの段階をみると、難しい問題になるほど平均点が下がっていることがわかった。このことから、本テストの構成と判定基準、各段階だと判定するのに採用した問題は、おすすめコースを診断するのに適切であったと考えられる。

表11 判定結果と問題の出典となるコースの段階別問題群の平均点 (標準偏差)

		問題の出典となるコースの段階別問題群の平均点 (標準偏差)			
		入門 (A1) 32点満点	初級1 (A2) 33点満点	初級2 (A2) 35点満点	
判定結果	入門 (A1)	8.65点 (6.38)	7.05点 (5.81)	5.17点 (5.27)	
	初級1 (A2)	25.41点 (3.20)	16.97点 (4.50)	10.90点 (5.49)	
	初級2 (A2)	29.62点 (2.53)	27.06点 (3.66)	23.35点 (6.80)	

また、本テストで、判定の基準として設けた3つの条件が機能していたかを確認するため、この条件に当てはまった受験者の内訳を確認した。条件ア「32点以上でも、1つでも0点のセクションがあった場合には、入門 (A1)」に該当したのは2610名中264名 (入門 (A1) 判定者の10.1%)、条件イ「32点以上でも、入門 (A1) レベルの問題の正答率が60%未満の場合は、入門 (A1)」に該当したのは2610名中394名 (入門 (A1) 判定者の15.1%)、条件ウ「65点以上でも、初級1 (A2) から出題したA2レベルの問題の正答率が60%未満の場合は、初級1 (A2)」に該当したのは2017名中92名 (初級1 (A2) 判定者の4.6%) であった。

条件ア「32点以上でも、1つでも0点のセクションがあった場合には、入門 (A1)」に該当した人が0点だったセクションの内訳は表12の通りである。「ちょうかい」に関しては、テスト受験時に音声を聞く環境になかったため0点となった受験者がいたことも予測されるが、この結果を見ると、4セクションのうち「かいわとぶんぼう」と「ちょうかい」が判定に影響していたことが推察された。まるごとコースは、聞く・話す・読む・書くの4技能全てに関わる活動を通して、総合的な日本語能力を身につけることを目標としたコースであるため、いずれかのセクションの点数が極端に低い状況でより上の段階のコースを受講するのは難しいと思われる。上記のような結果から、限られた問題数で構成した本テストにおいて、ア～ウの条件付けは、ユーザーにあったコースを判定する際に一定の効果があったのではないかと考えられる。

表12 入門（A1）と判定された受験者の合計点と0点であったセクション

0点であったセクション	合計点32点以上 65点未満の受験者	合計点65点以上 の受験者	条件A該当者全体 (条件A該当者の中での割合)
「もじとことば」	11名	0名	11名 ( 4.17%)
「かいわとぶんぼう」	140名	2名	142名 (53.79%)
「ちょうかい」	74名	7名	81名 (30.68%)
「どっかい」	15名	0名	15名 ( 5.68%)
「もじとことば」と「かいわとぶんぼう」	2名	0名	2名 ( 0.76%)
「かいわとぶんぼう」と「どっかい」	1名	0名	1名 ( 0.38%)
「ちょうかい」と「どっかい」	12名	0名	12名 ( 4.55%)

## 5. おわりに

オンラインコースでの学習は、多くのユーザーが自学自習という環境にあり、自身の現在の日本語能力にあったコースを選ぶのが難しいと考えられる。本稿では、JFSに準拠して開発した本テストが、分析の結果、信頼性と妥当性がある程度確保できていることがわかり、ユーザーにとってコース選びの一助となり得ることがうかがえた。今後もオンラインでの学びを助ける手だてを模索していきたい。

### 〔注〕

- <sup>①</sup>JF 日本語教育スタンダードは「相互理解のための日本語」を理念として開発された日本語の教え方、学び方、学習成果の評価のし方を考えるためのツールである（国際交流基金 2017）。
- <sup>②</sup>開発時に調査した「NHK 英語テキストフル活用 BOOK 15分でわかる！英語力測定テスト2016」は、現在は非公開となっている。
- <sup>③</sup>アンスティチュ・フランセ日本のフランス語のテストでは、テストの冒頭で自身のフランス語力を答える質問があり、「自分や他人を紹介することができる。家族や趣味について簡単なフランス語で話せる。」「自分の学歴、身の回りの状況について話せる。さしあたって必要な事柄を説明できる。」「フランス語が話されている地域を旅行しているときに起こりそうな、たいいてい事態に対処することができる。」「さまざまな話題について、明確かつ詳細に表現できる。」のうち、後者の2つを選んだ場合には、聴解が大問2問（小問9問）追加される構成となっている。
- <sup>④</sup>開発の過程でCEFR Can-doと『まるごと』のCan-doを整理する際には、国際交流基金が運営する「みんなのCan-do サイト」（<https://jfstandard.jp/cando/top/ja/render.do>）を用いて行った。そのため、本稿の表3、表4内のCan-do番号は『まるごと』のものだが、Can-doの記述はJF Can-doの文言を示す。
- <sup>⑤</sup>Can-do27は『まるごと』初級2（A2）活動編9課のCan-doだが、このCan-doのもととなったJF Can-doはA1レベルのものである。そのため、作題にあたっては初級2（A2）の理解編9課の読解や、JF Can-do 124「旅行中の出来事について書かれた家族や友人からの短い簡単なはがきやメールなどを読んで、内容を大まかに理解することができる。」を参考にした。
- <sup>⑥</sup>「まるごと教師サポート付きコース」は、「まるごと日本語オンラインコースサイト」での自学自習に、課題の添削やライブレッスンなど教師によるサポートが付いたコースである（千葉ほか 2018）。

〔参考文献〕

- 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
- 伊東祐郎 (2005) 「プレースメント・テストの妥当性確認の試み」『東京外国語大学留学生日本語教育センター論集』31、161-174
- 来嶋洋美・柴原智代・八田直美 (2012) 「JF 日本語教育スタンダード準拠コースブックの開発」『国際交流基金日本語教育紀要』8、103-117
- 来嶋洋美・柴原智代・八田直美 (2014) 「『まるごと日本のことばと文化』における海外の日本語教育のための試み」『国際交流基金日本語教育紀要』10、115-129
- 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 (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』、国際交流基金
- 国際交流基金 (2013) 『まるごと 日本のことばと文化 かつどう』(入門 A1)、三修社
- 国際交流基金 (2013) 『まるごと 日本のことばと文化 りかい』(入門 A1)、三修社
- 国際交流基金 (2014) 『まるごと 日本のことばと文化 かつどう』(初級1 A2)、三修社
- 国際交流基金 (2014) 『まるごと 日本のことばと文化 りかい』(初級1 A2)、三修社
- 国際交流基金 (2014) 『まるごと 日本のことばと文化 かつどう』(初級2 A2)、三修社
- 国際交流基金 (2014) 『まるごと 日本のことばと文化 りかい』(初級2 A2)、三修社
- 国際交流基金「みんなの Can-do サイト」<<https://jfstandard.jp/cando/top/ja/render.do>> (2018年9月4日)
- 武田素子・熊野七絵・千葉朋美・笠井陽介・石井容子・前田純子・北口信幸 (2017) 「『まるごと (A1) 日本語オンラインコース』サイトの開発」『国際交流基金日本語教育紀要』13、133-140
- 千葉朋美・武田素子・廣利正代・笠井陽介 (2018) 「『まるごと (A1) 教師サポート付きコース』の運用と成果—オンラインコースにおける学習者支援—」『国際交流基金日本語教育紀要』14、51-66
- 中村洋一 (2002) 『テストで言語能力は測れるか—言語テストデータ分析入門—』、桐原書店
- 東伴子・代田智恵子・永田道子 (2017) 「行動主義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—新しい評価基準をめざして—」『CASTEL/J 2017 予稿集』162-167