

Webによる日本語プレースメントテストの開発 —外国人留学生の受け入れ拡大にむけて—

藤田 恵・平山紫帆・栗田奈美・金庭久美子・数野恵理

Abstract: 本稿は、Webによる日本語プレースメントテストの開発過程を開示し、その過程で明らかとなった課題と解決策を述べるものである。本学では、「スーパーグローバル大学創成支援」採択に伴う国際戦略の中で、留学生の受け入れを大幅に拡大する方針を示している。そこで日本語科目では、留学生数増加によるプレースメントテストの実施回数増加、1回のテストの受験者数増加への対応策として、プレースメントテストをこれまでの筆記テストからWebテストへ移行することとし、開発を行った。Webテストの開発においては、「①短期間での開発」、「②開発経費の低減」、「③オリジナルコンテンツの開発」という目標を設定し、これらを達成した。しかしながら、その過程において「④判定精度」、「⑤1回のテストにかかる時間」という課題が挙げられた。これらの解決のために試行テストを3回行い、Webテストを完成させた。Webテストの採用により、テスト実施時間の短縮と、実施者と受験者の負担軽減が実現されたことから、Webテストは留学生数増加の対応策として有効であるといえる。

Keywords: Webによるプレースメントテスト、実施者と受験者の負担軽減、日本語教育

1. はじめに

本稿は、Webによる日本語プレースメントテストの開発までのプロセスを整理し、運用開始までに挙げた課題とその解決策を示すものである。

外国語教育を行う機関では、プログラムの開始時にプレースメントテストによるレベル判定を行い、クラス配置をする例が多い。日本語教育は、その対象が留学生であるという特性から、留学生の訪日からプログラムの開始までの短い時間でレベル判定をしなければならないという時間的制約がある。

近年、大学の国際化が叫ばれており、2008年の「留学生30万人計画」、2015年の「スーパーグローバル大学創成支援」といった事業からもその動きが読み取れる。本学では「スーパーグローバル大学創成支援」採択に伴う国際化戦略「Rikkyo Global 24」の中で、留学生の受け入れを、現在の約500名から、2019年に1,000名、2024年に2,000名へと大幅に拡大する方針を示している。これに伴い、本学の日本語科目では、今後、留学生の受け入れ時期の多様化、留学形態の多様化への対応が見込まれる。これにより、日本語科目のプレースメントテストには、上述の時間的制約に加え、テスト実施回数と1回のテストで処理する受験者数の増加という負荷が、より一層かかることが予想される。

以上の背景からその対応策として、日本語科目では、プレースメントテストをこれまでの筆記によるテストからWebによるテスト（以下、Webテスト）へと移行することとした。Webテストの開発において重視したのは、「①短期間での開発」、「②開発経費の低減」、「③オリジナルコンテンツの開発」である。本稿では、①から③を実現したWebテストの開発過程を開示

し、その中で明らかとなった課題である「④判定精度」、「⑤1回のテストにかかる時間」の解決策を示す。

2. 新しいプレイスメントテストの設計

日本語科目では、これまでプレイスメントテストの実施に2日間を要しており、1日目は筆記による3つのテスト（文法、読解、作文）、2日目はインタビューテストを実施していた。新しいプレイスメントテストの形態は、前述した理由から、テスト時間と採点にかかる時間の短縮を目指し、Webテスト、筆記による作文テスト、インタビューテストの組み合わせに切り替え、全てのテストを1日で完了する設計とした。これにより、時間と労力の軽減を図る一方で、これまで受験者だけでなく、海外の協定校や学内から高い評価を得てきたインタビューテストのシステムは残すデザインとなった。

新規に開発するWebテストは、「①短期間での開発」を目指し、コンテンツ開発から試行テストを経て、運用開始までの期間を約1年とした。また、「②開発経費の低減」の実現のために、業者が事前に設計したアルゴリズムを援用することとした。そして、「③オリジナルコンテンツの開発」により、本学の日本語教育プログラムに合った高精度の判定結果が得られるようにした。次章よりWebテスト開発のプロセスを示す。

3. Webテストの開発

Webテストの開発は、12名の日本語教師がチームとなり行った。その内訳は、【a】教授2名、【b】教育講師4名、【c】兼任講師3名、【d】外部の日本語教師3名である。【a】から【d】のそれぞれの担当と開発のプロセスを図1に示す。

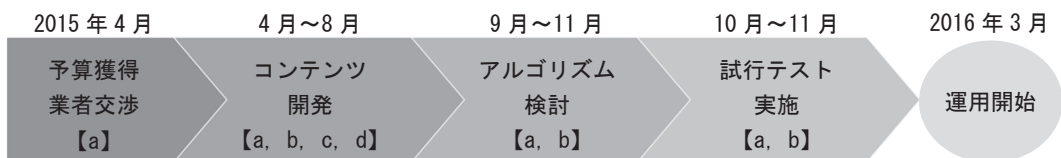


図1 Webテスト開発のプロセス（約1年）

Webテストの開発期間は、前述のとおり約1年である。予算の獲得後、業者交渉を開始し、並行して、コンテンツ開発を行った。そして、コンテンツ完成後、アルゴリズムの検討を行い、試行テストを経て、運用開始となった。本章では、コンテンツ開発から試行テスト実施までの過程を述べる。

3.1 コンテンツの開発

テストの分類には、その観点によってさまざまな名称が用いられるが、実施時期から見ると、プレイスメントテストはプログラム開始前に実施されることから「診断的評価」にあたる。これにより教育機関では、学生のレディネスを評価し、プログラムのレベルに合った判定を行い、適切なクラス配置を行う必要がある。

日本語のWebテストには、SPOT（小林他1996）やJ-CAT（今井他2012）といった外部テ

ストが公開されており、これらを用いてレベル判定を行う日本語教育機関もある。また、日本語能力試験や日本留学試験といった大規模試験の結果が用いられることもある。しかし、本学の日本語科目では、オリジナル教材を用いており、さらに技能ごとにレベルを判定し、クラス配置を行うことから、プログラムに合った高精度の判定結果を得るために、既存の外部テストは用いず、「③オリジナルコンテンツの開発」をすることとした。

Web テストの開発においては、「②開発経費の低減」の実現のため、業者が事前に設計したアルゴリズムを援用することは先に述べた。このシステムに合うように、Web テストの形式は、短答式五肢択一とすることとし、出題科目は、この形式で出題が可能な「漢字」「語彙」「文法」の3セクションを設けることとした。

本学の日本語科目では、入門から超級まで8つのレベルが設定されている¹。8つのうち、1つは技能によって能力差のある学生向けに設定された初級項目を復習するレベルである。コンテンツのレベル数を決定する際には、このレベルと、全てのセクションで高得点をとることで判定できる最上位のレベルを除外し、レベル数を6つとした。さらに、より精度の高い判定を行うために、6つのレベルをそれぞれ上位、下位に分け、合計12レベルの判定結果が出るようにコンテンツを開発することとした。

業者設計のアルゴリズムで必要とされるコンテンツ数は、各レベル75問であった。以下に、3セクション×12レベル×75問、合計2,700問の「③オリジナルコンテンツの開発」の手続きを示す。

- (1) 出題項目の選定 (1,200 問程度/1セクション) 【b: 2名/1セクション】
- (2) 作題会議 (1回目) 【a, b, c, d】
- (3) 問題文、誤答選択肢の作成 【c: 1名, d: 1名/1セクション】
- (4) 作題会議 (2回目) 【a, b, c, d】
- (5) 問題文、誤答選択肢の提出 【c: 1名, d: 1名/1セクション】
- (6) 見直し、出題問題の選定 (900 問/1セクション) 【b: 2名/1セクション】
- (7) 3セクション全問題の見直し (2,700 問) 【a: 1名】

「③オリジナルコンテンツの開発」は、【a】1名、【b】4名（うち2名は2セクションを担当）、【c】3名、【d】3名の合計11名の日本語教師で行い、その期間は4か月である。

(1) では、必要数が900問であるところを、出題項目の重複回避と、判定精度の向上のための入れ替え項目として、予備項目を含めた1,200問程度を用意した。出題項目の選定とレベル分けは、日本語科目の教材に準じて行い、外部の大規模試験の結果²も参考にした。そして、作題のための参考資料として、レベル分けされた出題項目（正解選択肢）と誤答選択肢の候補を挙げたリストを作成した。

(2) は、セクションごとに行い、作題を行う【c, d】に対して、作題のための参考資料を配付し、(3)の作成を依頼した。

(4) では、(3)の進捗状況の報告と、セクションごとに課題となった箇所を【a, b, c, d】で

1 立教大学日本語教育センター「開講レベル」<https://cjle.rikkyo.ac.jp/level> (2016.11.1 アクセス) より。実際には、J0～J8の9レベルの開講であるが、J0は「生活日本語」を学び、半期完結のクラスであるため、ここでは除外した。

2 日本語能力試験では、2009年までの出題問題と各問題の正答率、識別力を公開しており、それらを参考にした。

共有した。作題において課題となったのは、出題項目（正解選択肢）と他の問題の誤答選択肢との重なりである。これについては、本 Web テストの、問題が1問ずつランダムに画面に表示され、前に解答した問題を閲覧できないという性質上、ある程度の許容を認めることにした。(4)の後、【c, d】が作業を進め、(5)の提出となった。

(6)では、(5)で提出された各セクション1,200問程度の問題に対して、【b】2名による主観判定を行い、その結果から問題の難易度の調整および必要数900問の選定を行った。

(7)では、(6)で選定した3セクション全2,700問を【a】1名が見直し、出題問題の最終調整を行った。

以上の手続きによって、2,700問の「③オリジナルコンテンツの開発」を完了し、アルゴリズムの検討および試行テストへと移行した。

3.2 試行テストの実施

試行テストの実施によって明らかとなった課題は、「④判定精度」と「⑤1回のテストにかかる時間」である。本節では、これらの課題の検証と解決のために行った3回の試行テストの結果を示す。

3.2.1 試行テスト1回目

1回目の試行テストでは、「④判定精度」の検証を第一の目的とした。受験者は以前のプレイスメントテストのデータのある留学生を中心とし、日本語科目の8つのレベルから3～5名ずつの協力を得て行った。以下に、1回目の試行テストの概要を示す。

実施時期：2015年10月

受験者：日本語科目受講の留学生 40名

コンテンツ：「漢字」「語彙」「文法」

アルゴリズムの設定：・3問連続正解で上のレベルに移動（不正解の場合は下のレベル）

・同レベル20問連続出題時点でレベル確定

・同レベル全75問出題時点でレベル確定

1回目の試行テストでは、「④判定精度」に関して、課題が残る結果となった。その要因は、「漢字」と「文法」において、上位レベルの識別がされなかったことであり、コンテンツの見直しが必要であることが分かった。

他に、「⑤1回のテストにかかる時間」に関して検討が必要であることが明らかとなった。試行テスト前に想定していたテスト時間は、受験者の体力と心的な負担を考慮し、上限を75分としていた。しかし、1回目の試行テストでは、75分以内に3セクションを終えることができない受験者が12名おり、最長で150分となった。そのため、受験者全員が75分でテストを終えられるように、アルゴリズムの設定を見直すこととした。

3.2.2 試行テスト2回目

2回目の試行テストに向けて、「④判定精度」の向上と「⑤1回のテストにかかる時間」の短縮のために、「漢字」「文法」のコンテンツの見直しとアルゴリズムの調整を行った。

まず、「漢字」で上位レベルが識別できなかった要因は、漢字の表記を問う問題に登場する誤答選択肢が実在しない漢字語である場合、日本語能力を識別する効果が弱かったことにある。一方で、漢字の読み方を問う問題は、一定の効果が見込めるという結果であった。よって、「漢

字」では、上位レベルにおいて、漢字の表記を問う問題を避け、読み方を問う問題のみに差し換えることとした。

次に、「文法」においては、上位レベルにおいて出題問題の難易度を上げる必要があることが分かった。そのため、出題項目、問題文、誤答選択肢の見直しを行うこととし、時間的な制約から試行テストでは差し替えを見送り、運用開始までに完了させることにした。

最後に、アルゴリズムの調整では、同レベルでの連続出題数を10問までとすることで、時間の短縮を図った。

以下に、2回目の試行テストの概要を示す。

実施時期：2015年11月

受験者：日本語科目受講の留学生 37名（試行テスト1回目の受験経験者）

コンテンツ：「漢字（改定版）」「語彙」「文法」

アルゴリズムの設定：・3問連続正解で上のレベルに移動（不正解の場合は下のレベル）

- ・同レベル10問連続出題時点でレベル確定
- ・同レベル全75問出題時点でレベル確定

2回目の試行テストでは、「④判定精度」に関しては改善が見られ、「漢字」のコンテンツはこれで確定することとした。一方で、「⑤1回のテストにかかる時間」は改善が見られず、15名の受験者が75分で終わることができなかった。そのため、3回目の試行に向けて、アルゴリズムのさらなる調整を行うこととした。

3.2.3 試行テスト3回目

3回目の試行テストでは、「⑤1回のテストにかかる時間」の短縮を目指し、アルゴリズムの調整を行った。2回目の試行テストの結果から、テスト時間が長くなる受験者は「同レベル全75問出題時点でレベル確定」となっていることが分かった。よって、この出題数の上限を50問に減らし、3回目の試行テストを実施した。以下に、3回目の試行テストの概要を示す。

実施時期：2015年11月

受験者：日本語科目受講の留学生 6名（試行テスト2回目で75分以上かかった受験者）

コンテンツ：「漢字（改定版）」「語彙」「文法」

アルゴリズムの設定：・3問連続正解で上のレベルに移動（不正解の場合は下のレベル）

- ・同レベル10問連続出題時点でレベル確定
- ・同レベル50問出題時点でレベル確定

3回目の試行テストでは、アルゴリズムの調整によって、「⑤1回のテストにかかる時間」の短縮が達成された。6名の受験者とも75分より短い時間で終わることができ、最長でもテスト時間は36分であった。この試行により「⑤1回のテストにかかる時間」の課題は解決され、出題数の減少による「④判定精度」の低下も見られないことが確認できた。

3回のWebテストの試行によって、これまでに挙げた④と⑤の課題が解決され、適切なクラス配置を行うための結果が得られるようになった。また、Webテストにより、テスト実施時間と判定にかかる時間の短縮が確認でき、実施者側の負担軽減が実現可能となった。Webテストの採用は、プレイメントテストの実施回数と受験者数の増加が見込まれることへの対応策

として有効であるといえるであろう。

4. 試行テスト後のインタビュー調査

試行テスト後、受験者側の所感を明らかにするために、新旧両方のプレイスメントテストを受験した留学生2名の協力を得て、半構造化インタビューを実施した。

受験者は、まず新しいプレイスメントテストの設計について、

- ・Webテスト、作文テスト、インタビューテストと3つのテストがあるのはいい。
- ・今まで2日間で実施していたテストが1日で済むようになったのがいい。

と話しており、実施科目と日数の短縮に関する肯定的なコメントが得られた。初めの回答は、新しいプレイスメントテストの設計を行う際に、インタビューテストを残すことにした理由と合致する。そして、日数については、以前の設計では2日かけて行っていたものを、新しい設計ではWebテストの採用によって1日で実施可能となったことへの評価である。テストにかかる時間、日数の短縮は、実施者のみならず、受験者にとっても肯定的にとらえられる事項であることが分かった。

次に、Webテストについては、

- ・筆記テストより、Webテストの方がよかった。
- ・Webテストでは、問題に答えられないと下のレベルに行くなど、自分のレベルに合った問題が出されるのでいい。

といった肯定的なコメントを得た一方で、

- ・Webテストは、自国では受けたことがなかった。
- ・Webテストは（今回初めて経験したので）いいか悪いか分からない。

などのように、Webテストに不慣れであることへの指摘もあった。これらは、Webテストの判定結果が試行テストにより信頼の置けるものであることが確認できていることを伝え、さらにテスト実施前に操作説明を丁寧に行うことによって解決できるものであると考える。

5. おわりに

本稿では、Webによる日本語プレイスメントテストの開発過程を整理し、その中で明らかとなった課題と解決策を述べた。本Webテストの開発においては、「①短期間での開発」、「②開発経費の低減」、「③オリジナルコンテンツの開発」という目標を達成し、さらに課題として挙げられた「④判定精度」、「⑤1回のテストにかかる時間」の解決に至った。

本稿で示したコンテンツ開発の過程とアルゴリズムの調整は、Webテストの採用を検討している他言語や他の教育機関においても応用可能なものであろう。また、試行テスト後に実施者と受験者の双方の視点から新しいプレイスメントテストの設計を概観したことで、Webテストの採用によって得られる効果を示すことができた。今後は、試行テストよりも受験者数が増える本格的な運用でデータ収集を行い、検討を重ねることで、本Webテストの信頼性をより高めることに努めたい。

参考文献

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

小林典子・山元啓史・フォード丹羽順子 (1996) 「日本語能力の新しい測定法 [SPOT]」『世界の日本語教育』6号、pp.201-236

藤田恵・金庭久美子・栗田奈美・平山紫帆・池田伸子・丸山千歌 (2016) 「Web によるプレイ
スメントテスト開発における方略と課題」日本語教育国際研究大会 (ポスター発表)

文部科学省・外務省・法務省・厚生労働省・経済産業省・国土交通省 「『留学生 30 万人計画』
骨子」

<http://www.kantei.go.jp/jp/tyoukanpress/rireki/2008/07/29kossi.pdf> (2016.11.1 アク
セス)

文部科学省 「スーパーグローバル大学創成支援」

http://www.mext.go.jp/a_menu/koutou/kaikaku/sekaitenkai/1360288.htm
(2016.11.1 アクセス)

立教大学 「Rikkyo Global 24 立教大学国際化戦略」

<https://www.rikkyo.ac.jp/global24/> (2016.11.1 アクセス)

本稿は、日本語教育国際研究大会 (2016 年 9 月 10 日) において報告したポスター発表の内
容に、加筆、修正したものです。

本稿の執筆にあたり、立教大学異文化コミュニケーション学部の丸山千歌先生、池田伸子先
生には、多くのご示唆をいただきました。Web テストの開発は、2 名の先生と共に行い、試行
テストとインタビュー調査は、日本語教育センター事務局の皆さん、留学生の皆さんのご協
力を得て行いました。多くの皆さまにご協力いただいたことをここに記し、御礼申し上げます。

