日本語オンラインプレースメントテスト（LEAP）の開発
実装・試行・評価

佐藤 礼子・榎原 実香・山元 啓史（東京工業大学）

Development of Online Placement Test of Japanese, LEAP

SATO Reiko・EBARA Mika・YAMAMOTO Hilofumi, Tokyo Institute of Technology

要旨：LEAP はレベル分けのためのオンラインテストで、短い文を聴いて一字分の空欄に回答する SPOT テストを基に開発された。ローマ字を自動的にひらがな変換する機能を有する。試行結果から、オンラインでも SPOT-用紙版と同様の回答行動である、問題文から目を離さず聞く行動が実現できたようであった。既存のプレースメントテストによるレベル分けと対照したところ、LEAP の得点は初級前半と中級学習者の間、中級前半と中上級学習者の間で異なり、日本語レベルの違いをある程度判別できることが示された。また、初級レベルで LEAP 得点の個人差が大きかったことから、初級において音声聴取能力が LEAP 得点に影響を及ぼすことが推測された。

キーワード：SPOT 形式、評価、簡易テスト、文法項目、難易度

1. はじめに

本研究の目的は簡易にできる Web 版プレースメントテストを開発することである。テストの形式として実績のある SPOT 形式（小林ら, 1996; 音声を聞きながら括弧にひらがな 1 文字を書く形式の時間制限のあるテスト）を採用した。SPOT-Web 版（小林, 2005）では Web 化に伴い、回答形式が四肢選択に変わったこと、問題項目をランダムにしたこと、音声が明瞭になったことで、SPOT-用紙版とは異なる評価法になったことが報告されている。本研究では SPOT-用紙版の、スピードテスト、ひらがな 1 文字という簡易な回答行動、クリアな単音を聞き取るのではなく多少の雑音の中でも文脈を推測しなければ回答できない音声の開発を目標とした[注1]。各問題文の構成は SPOT-用紙版の形式に準じてはいるが、単語・表現・括弧の位置のいずれにおいても SPOT と異なるため、本研究では LEAP (Language assEssment for Active Performance)と呼び、SPOT と区別する。

2. 開発・仕様

問題文は長くても 30 文字程度で、助詞、述語の活用、文法要素について 1 文あたり 1 文字分を回答させる。問題文の例は「きょうの試験、思った（よ）りむずかしかったな。」「今日中に先生に会（え）ればいいんですが。」で、漢字にはルビを付した。問題は、読み方、文体をはじめ諸要因の異なる C、D、E、F、G の 5 バージョン、各版 30 問ずつ作成した。読み上げスピードは自然な速さとし、各版 3 分以内の長さであった。録音は Zoom

を用いて行い、音声にはカセットテープのような音質になるようノイズ[注2]を加えた。

SPOT-Web 版は四肢選択形式で、音声とともに問題文 1 問とその選択肢が表示され、マウスで回答すると次の問題文の画面に遷移する方式であった。LEAP では、ローマ字入力でひらがな 1 文字を括弧に入力し、回答する形式にした（図 1)。30 問すべての問題文が 1 ページに提示され、各問題文の音声（問題間のポーズは 2 秒）が連続して流れる。音声とともに Tab キーで次の括弧に移動し、ローマ字でひらがなを入力する形式にした。IE8 を除く一般的なブラウザを使い、JavaScript を有効、日本語入力 IME オフ（自動設定）で実施できるようにした。受験方法を指示したページで、ローマ字入力方法とテストの練習ができる。練習問題は 10 問で回数に制限はない。本番は 1 回のみ受験できるように制限を施した。本番は、音声読み上げが終了するか Submit ボタンを押すかで完了する。

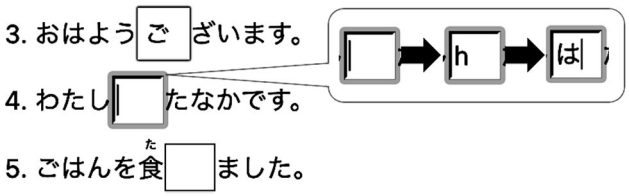

図 1：LEAP テストのローマ字入力場面（練習問題）

3. LEAP テストの試行

3.1 システムの動作

テスト受験した学生複数名に受験した際に思ったことをインタビューした。日本語入力システムを OFF にする機能も動作し、ローマ字・仮名変換に対する問題はなかった。テスト時間はほぼ音声再生時間と若干の通信ラグであり、全員の学生がテストを中断することなく受験できたことが確認できた。

3.2 日本語レベル別の LEAP 得点の分析

日本語コース開始時に、オンラインでプレースメントテスト（語彙・文法、四肢選択形式）[注3]と LEAP を実施した。自己申告により初級者は LEAP-C、F、G を、中級者は C、D、E を受験したため、本稿では全ての学習者が受験した LEAP-C の結果を分析する。

受験者は 147 名、LEAP-C の平均得点（以下、LEAP 得点）は 23.0 であった（最大値 30、最小値 2、SD = 6.1)。図 2 にプレースメントテストで判定した日本語レベル別の LEAP 得点を示した（初級前半 n=37、初級後半 n=23、初中級 n=14、中級前半 n=25、中級後半 n=35、中上級 n=13)。LEAP 得点の平均値は初級前半からレベルが高くなるにつれて高くなり、中級後半・中上級では満点に近かった。初級前半・初級後半は中級レベル（初中級～中上級）と比べて得点にばらつきが大きかった。

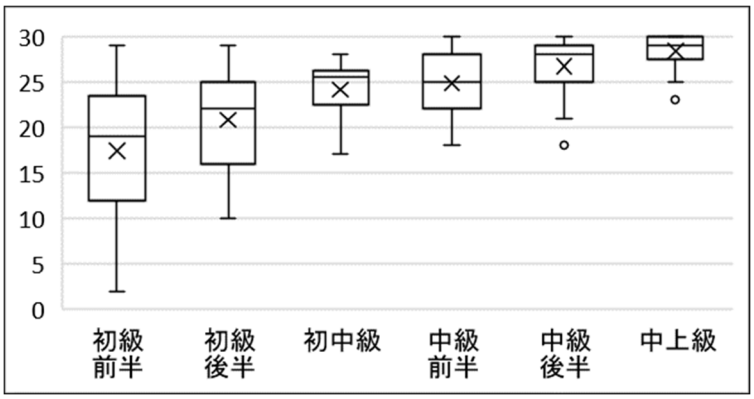

図 2：日本語レベル別の LEAP 得点

LEAP 得点について Kruskal-Wallis 法で検定を行い（$\chi^2(5) = 65.57, p < .001$）、Wilcox の順位和検定で多重比較を行った。有意水準は 5%とした。その結果、初級前半と中級（初中級・中級前半・中級後半・中上級）の間、初級後半と中級後半・中上級の間、初中級と中上級の間、中級前半と中上級の間に差が認められた。このことから、LEAP は日本語レベルによって得点が異なることが示され、初級学習者と中級学習者の間、中級の中でも中級前半学習者と中上級学習者の間で得点に差があることが確認された。しかしながら、より細かいレベル判定に必要となる、初級前半と初級後半の間、中級のレベル間については有意な差はみられなかった。

4. 考察

4.1 システム全般

SPOT-Web 版は四肢選択であったが、LEAP ではローマ字入力でひらがな 1 文字を回答する方式にしたことで、オンラインでも SPOT-用紙版と同様の回答行動、すなわち問題文から目を離さず聞く行動が実現できたようであった。

4.1 LEAP 得点と日本語レベルとの関係

LEAP 得点が初級からレベルが上がるにつれ高くなったことから、LEAP は既存のプレースメントテストと同様の傾向を示したと言える。その一方、初級前半・初級後半レベルで LEAP 得点の個人差が大きく、あまり回答できない学習者もいれば、正答率が低い（難しい）項目でも聞き取れる学習者もいた。今回、音声にノイズを入れることで、音声の聞き取りのみで回答することを回避しようとしたが、ヘッドホン等で聴取することにより、自然な環境での聴解と比べて音声が聞き取りやすくなったことも考えられる。ただし、初級と中級の間で得点の差があったことから、ノイズの効果は一定程度あったと考えられる。

5. おわりに

今回開発した LEAP では、初級と中級のレベルの違い、中級の前半と中上級のレベルの

違いを判別できた。ただし、理解を伴わなくても音声を聞き取れる能力が結果に影響すること、日本語能力が高いレベルでは天井効果がみられること、より細かいレベル判定が困難であるという課題が残された。音声の明瞭さを操作する、理解や記憶をさせながら音声を聴取させるなど、より実際の日本語運用能力を測るテストになるよう改善したい。

注

注 1 不明瞭な録音による SPOT-用紙版の得点が Web 版になると高い得点になっている（小林, 2008）。

注 2 録音後の編集作業では、Audacity のイコライザ機能を用い、コサインカーブ（200Hz-400Hz, 1000Hz-1400Hz）部を-10db に変更して加工を施した。

注 3 初級前半の学生は日本語未習者であるため、プレースメントテストは受けていない。

参考文献

小林典子・フォード丹羽順子・山元啓史. 1996.「日本語能力の新しい測定法「SPOT」」『世界の日本語教育』6 : 201-218.

小林典子. 2005.「言語テスト SPOT について - 用紙形式から WEB 形式へ -」『筑波大学留学センター日本語教育論集』20: 67-82.

小林典子. 2008.『言語テスト SPOT-WEB 版の開発と解答行動の研究　科学研究費補助金基盤研究(B) 研究成果報告書(1)(2)』 筑波大学.