

**科学研究費助成事業 研究成果報告書**

平成 28 年 6 月 17 日現在

機関番号：12102

研究種目：基盤研究(C) (一般)

研究期間：2011～2015

課題番号：23520744

研究課題名(和文) 発話を促す実用的スピーキングテストの開発とピア評価システムの確立

研究課題名(英文) Establishment of a practical speaking test and peer assessment system to promote speech production

研究代表者

平井 明代 (HIRAI, Akiyo)

筑波大学・人文社会系・教授

研究者番号：00312786

交付決定額(研究期間全体)：(直接経費) 3,700,000円

研究成果の概要(和文)：本研究は、中高大の授業で使いやすいストーリー・リテリング・スピーキング・テスト(SRST)とその評価尺度(ルーブリック)の開発を目的とした。中学及び高等学校で学習する文法や表現が発話時に使えるかを評価できるテストにするために、オリジナルSRSTと評価尺度を改良し、ストーリーの中にどの程度の文法項目を入れ、どのような指示を出せば、ターゲットになった項目を使いながら再話を行うことができるかを調査した。そして、外部試験との高い併存的妥当性のある評価尺度を開発した。また、その尺度を使用したピア評価では、内容面よりも、言語的観点の評価が比較的、妥当性が劣り苦手とする生徒が多いことを明らかにした。

研究成果の概要(英文)：This study aims to develop a practical classroom-based Story Retelling Speaking Test (SRST) with an assessment rubric that fits the needs of junior and senior high schools as well as universities. To this end, first, we modified the original SRST in order to find a) how many target grammar items or expressions should be integrated in a story to promote the use of the target items in the subsequent retelling task, and b) what kind of instructions or hints should be given in the test. Next, we made a rubric for this revised SRST and confirmed it had high concurrent validity with an external speaking test, even under peer assessment conditions. However, we found that students scored their peer's linguistic aspects less reliably than content aspects of the rubric criteria.

研究分野：英語教育、言語評価

キーワード：スピーキング評価 ストーリーリテリング ピア評価 技能統合

## 1. 研究開始当初の背景

中学校及び高等学校学習指導要領解説(文部科学省, 2008, 2010)では、「聞く」、「読む」、「話す」、「書く」の4技能を統合的に活用できるコミュニケーション能力の育成を目指すこと、及び文法指導も「話す」、「書く」という言語活動と一体的に扱うことが強調されている。

しかし、これらの技能統合的スピーキング活動は増えてきているものの、評価に関しては、実施や採点にあたる実用性に関する問題点以外に、(a)技能統合的スピーキングテストをどのように作成し(17%)、(b)評価すればよいのか(28%)、また(c)どのようにすれば信頼性の高い評価尺度が作成できるのか(23%)という、高校教員から具体的な評価方法についての懸念が挙げられている(Murray et al, 2012)。

このような背景から、中高大の授業で使いやすい妥当性及び信頼性の高いスピーキングテストと評価尺度を開発する必要がある。

## 2. 研究の目的

本研究は、中高大の授業で手軽に活用できる実用性のあるスピーキングテスト及び評価尺度の研究・開発することを目標とする。その目標を達成するために、主に次の5点に焦点を当てた。

(1)ストーリー(テキスト)を読んで再話する組み合わせだけでなく、聞いて再話する場合に作成しやすいように合成音声ソフトを使って、SRSTに使用するテキストが作れるかを検討すること、(2)SRSTを複数の外部テストと比較し、その特徴と併存的妥当性を明確にすること、(3)SRSTの評価尺度としてEBB尺度を利用しているが、文法・語彙観点の信頼性および実用性を高めるためEBB尺度を改良すること、(4)より中高の授業のニーズに合うように、学習した文法項目や表現などをテキストに取り込んで、それらが見えるかを評価できるようなストーリー・リテリング・スピーキングテスト(SRST)とその評価尺度を開発すること、(5)教師評価だけでなく生徒同士が評価し合うピア評

価においても妥当性と信頼性を確保できる評価尺度を目指すことを目的とする。

## 3. 研究の方法

「2. 研究の目的」(1)に関して、リスニングからスピーキング技能に移るSRSTを組み合わせたスピーキングテストの開発に向けて、リスニング用テキストを読み上げソフト(Globalvoice English)で代用することができるか調べた。読み上げソフトで作られる合成音声スピーチとネイティブ・スピーカーによる読み上げスピーチを次の3つの観点で比較した。(a)理解可能度(どれくらい内容を理解できる音質を有しているか)、(b)明瞭度(どれくらい1つ1つの単語や音がはっきりと聞き取れるか)、(c)自然さ(直感的にどれくらい聞き手にとって自然に聞こえるか)である。分析方法は、合成音声と自然音声で作成した文をディクテーションさせ、その再生率を比較した。

「2. 研究の目的」(2)に関して、学生にSRSTと、電話で受験するVersant(Pearson Education, 2008)及び面接式のStandard Speaking Test(SST; ALC Press, 2010)の計3つのスピーキングテストを学生に受験してもらった。そして、それらのテストの共通点と相違点を調べるために、その言語機能と得点、発話を分析し、併存的妥当性を調査した。

「2. 研究の目的」(3)に関して、オリジナルのEBB尺度と改良したEBB尺度、及び改良したEBB尺度から分析的評価形式にした多特性評価尺度を用意した。そして、学生のSRSTパフォーマンスをそれらの3つの評価尺度で比較した。分析方法は、主に、ラッシュモデルのFACETS及び重回帰分析による比較から最も妥当な尺度を検討した。

「2. 研究の目的」(4)に関して、SRSTのストーリー内に使わせたい文法項目や表現に下線及び太字にし、かつ指示文も注意を促すように工夫した。比較のために、ストーリーは2種類用意し、異なる文法項目を使用し、受験してもらった。このような操作によって、どの程度、再話時にその項目を使用し

たかを、観点別指標の1つに加え採点した。

「2. 研究の目的」(5)に関して、SRSTが実際に高校生同士のピア評価においても、高い妥当性を示すかを複数の高校で行い、教師評価とピア評価による採点を比較し、相関等で分析を行った。

#### 4. 研究成果

「3. 研究の方法」に沿って分析を行い、次の点を明らかにした。

(1) 合成音声の場合、自然音声では一般的に弱く速く発話される機能語まで規則的に発話されるため、明瞭度が高くその部分の再生率がよかった。このことから若干自然さに欠けるものの、レベルの低い学習者にも十分利用できることがわかった。

(2) SRST、Versant、SSTの3つのテストを比較した結果、SRSTは、Versantよりは多く、SSTよりは少ない言語機能を引き出すこと、

SRSTは他の2つのテストと中程度の相関を持ち、初級者・中級者グループで弁別力を発揮すること、タスクと言語機能と、評価時に重きを置く要素の相違により、テスト得点の違いが説明され得ることがわかった。各テストの相対的な利点を示すことができたことは、評価の目的と状況に応じて適切なテストを選ぶ際に有益な情報となる。

(3) オリジナルEBB、改良版のEBB、及び多特性評価尺度で、改良版EBBが最も信頼性及び妥当性が高くなった。また、4観点から3観点にすることによって実用性も向上したが、この実用性の面だけ見ると、多特性評価尺度が最も優れていた。よって、総合的には改良版EBBをSRSTのオリジナルバージョンの評価尺度として利用できると判断した。

(4) SRSTを中高生のニーズに合うように、学習した文法や表現を2、3種類ストーリーに入れて、再話させた場合、項目の種類が多く、再生するのが難しいことがわかった。今後、ターゲットにした項目を1種類に絞った

改良版のSRSTで実施する必要があることがわかった。

(5) 上記(4)のSRST(ターゲットバージョン)をピア評価させた場合、教師評価より妥当性が下がったが、教室使用では十分な程度の妥当性を示した。ただし、ターゲット項目を評価するのに困難さを感じている生徒が多かった。今後、この点は改良の余地がある。

(6) これまで作成したSRSTを試用し、修正し計15のテストを実際に使えるように準備した。

以上のことを明らかにし、読んで話すだけでなく、聞いて話す技能統合的なSRSTとその評価尺度の開発を行うことができた。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文](計11件)

平井明代 (2015)「授業を活かすストーリーテリング・テストの活用」大塚英語教育研究会『大塚フォーラム 33』, 49-69. (査読なし)

Hirai, A. (2014) A review of four studies on measuring vocabulary knowledge. *Vocabulary Learning and Instruction* 3 (2), 85-92. December 2014. (査読あり)

doi: <http://dx.doi.org/10.7820/vli.v03.2.hirai>

Retrieved from <http://vli-journal.org/>

Hirai, A., & Koizumi, R. (2013). Validation of Empirically-Derived Rating Scales for a Story Retelling Speaking Test. *Language Assessment Quarterly*, 10: 398-422. (査読あり)

doi: 10.1080/15434303.2013.824973

Hirai, A., Fujita, R., Ito, M. & O'ki, T. (2013).

Washback of the Center Listening Test on Learners' Listening Skills and Attitudes. *ARELE (Annual Review of English Language Education in Japan)*, 24, 31-45. (査読あり)

Hirai, A., Fujita, R. & O'ki, T. (2013). 「センターリスニングがもたらすリスニング学習意欲への影響：大学種別・入試形態・専攻ごとの分析に基づく考察」(The Influence of the Center Listening Test on Listening Learning Motivation: An Analysis Focusing on University Type, Admission Type, and Major). *JACET (The Japan Association of College English Teachers) Journal*, 57, 59-81. (査読あり)

平井明代、横内 裕一郎、佐瀬 文香 (2013) 「大学入試センター試験英語の発音を測定するための考察」『平成 24 年度リスニングテストの実施結果や成果等を検証し、その改善を図るための調査研究に関する報告書』独立行政法人大学入試センターリスニングテスト検証研究会, 65-88. (査読なし)

Koizumi, R., & Hirai, A. (2012). Comparing the story retelling speaking test with other speaking tests. *JALT (Japan Language Testing Association) Journal*, 34, 35-60. (査読あり)

Hirai, A., Ito, N., & O'Ki, T. (2011). Applicability of peer assessment for classroom oral performance. *JLTA Journal*, 14, 41-59. (査読あり)

[学会発表](計 9 件)

Hirai, A. (2015.10.16). Encouraging integrated speaking tasks and assessment in class. "3rd British Council New Directions in English Language Assessment: Quality and Consequence, in association with KELTA (*Korea English Language Testing Association*). at JW Marriott

Dongdaemun Square, Seoul, Korea.

平井明代 (2015.8.23) 「技能統合的スピーキング活動と評価」全国英語教育学会第 41 回熊本研究大会。『日本の英語教育の将来：「話すこと」の評価方法』熊本学園大学(熊本県熊本市)。

平井明代 (2014.11.8) 「テストで伸ばすスピーキング指導と評価」大塚英語教育研究会 11 月例会、筑波大学大塚キャンパス校舎(東京都文京区)。

Hirai, A. (2014.9.20). Toward a practical speaking assessment to facilitate learning in the classroom. The JLTA Symposium at Ritsumeikan University. (京都府京都市)

Hirai, A. (2014.6.14). A review of the studies on the overestimation of the vocabulary tests and an effective way to measure vocabulary knowledge. Discussant of the JALT Vocabulary Sig. Symposium at Kyushu Sangyo University. (福岡県福岡市)

平井明代 (2013.12.22) 「授業を活かすストーリー・リテリング・スピーキングテスト (SRST) とその評価方法」関西英語教育学会、関西英語教育学会、龍谷大学・大阪梅田キャンパス。(大阪府大阪市)

平井明代、藤田亮子、大木俊英 (2013.8.10) 「センターリスニングがもたらすリスニング学習意欲への影響：大学種別・入試形態・専攻ごとの分析に基づく考察」、全国英語教育学会 第 39 回北海道大会、北星学園大学。(北海道札幌市)

横内裕一郎、平井明代 (2013.8.10) 「筆記による発音・アクセント問題についての研究」

全国英語教育学会 第 39 回北海道研究大会、  
北星学園大学。(北海道札幌市)

Hirai, A., Fujita, R., Matsuzaki, H. (2011. 10.29). Washback effects of the National Center Listening Test on Japanese students' listening ability and their attitudes toward studying English listening. The 15<sup>th</sup> Annual Conference of JLTA at Momoyama Gakuin University. (大阪府和泉市)

〔図書〕(計 1 件)

平井明代(編著 2012)『教育・心理系研究のための データ分析入門：理論と実践から学ぶ SPSS 活用法』東京：東京図書．全 259 頁．

〔その他〕

ホームページ等

平井HP：

<http://www.u.tsukuba.ac.jp/~hirai.akiyo.ft/index.htm>

筑波大学リポジトリ：

<https://tsukuba.repo.nii.ac.jp/>

## 6．研究組織

### (1) 研究代表者

平井 明代 (HIRAI, Akiyo)

筑波大学・人文社会系・教授

研究者番号：00312786

### (2) 研究分担者

大木 俊英 (O'KI, Toshihide)

白鷗大学・教育学部・准教授

研究者番号：90580861

### (3) 連携研究者

小泉 利恵 (KOIZUMI, Rie)

順天堂大学・医学部・准教授

研究者番号：70433571