

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 10 日現在

機関番号：32402

研究種目：基盤研究（B）

研究期間：2009～2012

課題番号：21320105

研究課題名（和文）音読とシャドーイングに音声情報処理技術を活用した外国語評定システムの開発

研究課題名（英文）Development of an Automatic Evaluation System of Reading-Aloud and Shadowing Performances by Japanese EFL Learners Using the Latest Speech Technology

研究代表者 山内 豊（YAMAUCHI YUTAKA）

東京国際大学・商学部・教授

研究者番号：30306245

研究成果の概要（和文）：

パソコン上で最新の音声情報技術の一つである Goodness of pronunciation (GOP) を使って英語学習者のシャドーイングを自動評価できるシステムを開発した。英語熟達度の異なる学習者が初見でシャドーイングした音声データについて、自動評価得点とベテラン英語教員による手動採点の相関係数は 0.85 となり、本システムの有効性が明らかになった。さらに、自動評価得点と TOEIC 得点の相関も 0.82 となり、シャドーイングと TOEIC の得点を回帰分析することによって、シャドーイングの出来具合から学習者の総合的英語力を TOEIC 得点として予測できることもわかった。

研究成果の概要（英文）：

Although shadowing is considered to be effective in improving EFL learners' listening and speaking skills, it is both difficult and time consuming to assess shadowing performance. To solve this problem, an automatic assessing system has been newly developed using the latest speech information technology: GOP (Goodness of Pronunciation). The computer automatically compares the learner's shadowed speech with an acoustic model based on the original text on a phoneme basis and computes numerical scores which indicate how close the learner's shadowed speech is to the original. To confirm the validity of this automatic assessment, an experiment was conducted. Japanese EFL learners with proficiency levels ranging from 200s to 900s measured by TOEIC scores were requested to record their shadowing into the computer while listening to passages, which had been never presented before, read by a native speaker of English at a speed of approximately 140 words per minute. Then the scores obtained by the automatic system were compared to those rated manually by veteran EFL instructors. A high correlation ($r=0.85$) was observed between the two; thus this system is significantly valid and also saves much time and energy in assessing shadowing performance. Since the automatic scores of the learners were also highly correlated with their TOEIC scores ($r=0.82$), the shadowing performance can predict their overall proficiency scores. Based on the regression model, this system can predict and present TOEIC score levels. Therefore this automatic assessing system enables EFL learners to understand objectively how well they have conducted shadowing and also grasp their own proficiency levels.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2009 年度	6,600,000	1,980,000	8,580,000
2010 年度	4,600,000	1,380,000	5,980,000
2011 年度	2,700,000	810,000	3,510,000

年度			
年度			
総計	13,900,000	4,170,000	18,070,000

研究分野：人文学

科研費の分科・細目：言語学・外国語教育

キーワード：教育評価・測定

1. 研究開始当初の背景

シャドーイングは、外国語の発音やリズムやイントネーションを習得でき、言語処理の高速化・自動化が促進され、聞く・話す技能も伸ばせる効果的な練習法として注目されている。このため、シャドーイングは、近年外国語教育の現場で注目され、中学・高校から大学の授業まで広く実践されるようになってきている。

一方、シャドーイングを、妥当性・信頼性・実行可能性をともなって評価するのは容易ではない。教員が全体的印象で評価する方法は、客観性に問題が残る。客観的な評価法として、シャドーイングに使用した原文中の何割の音節や単語を再生できたかを計算する再生率法が提唱されてきた（門田 2007, 山内・峯松・羅・川村 2009）。

本研究プロジェクトでは、シャドーイングを、音響モデルに基づいた GOP (Goodness of Pronunciation) という客観的な指標を使って、自動的に分析・評価するシステムを開発し、自動評価得点は、ベテラン英語教師による手動評価得点と高い相関をもつことが確認された。ただし、これらの客観的な評価方法にも課題がある。音節や単語の再生率に基づく方法は、韻律面を直接測定していない。GOP による自動評価法も、音素単位の音響モデルとの類似度に基づく評価であり、韻律面は直接測定していない。そこで、再生率に基づく評価や音響モデルによる自動評価が、韻律面をどのくらい反映できるかについて明らかにするための実験を実施した。TOEIC 得点に基づく異なる総合的熟達度の学習者に、英語特有のリズムやイントネーションを要求する強調構文や付加疑問文が含まれる英語パッセージをシャドーイング録音してもらった。この録音音声について、ベテラン英語教員による単語の再生率に基づく得点、と韻律に基づく得点、GOP に基づいた自動評価得点を算出した。この3得点を分析した結果、再生率と韻律評価との相関係数は 0.578 にとどまったのに対し、GOP による自動評価法は、韻律評価と 0.772 という高い相関係数を示した。この結果から、GOP に基づく自動評価は、韻律面もある程度カバーできており、韻律面まで含めた評価指標としての妥当性が確認された（山内・峯松・羅・川村 2010）

シャドーイングの出来を評価する方法として、ネイティブ・スピーカーやベテラン教員が全体的印象で評価する方法は、客観性に問題がある。提示文に含まれる単語の何割を正確に再生できたかに基づいて計算する方法は、客観性はあるが作業に多大な時間とエネルギーを要するという問題がある。

2. 研究の目的

- (1) 英語教育と音声情報処理技術の専門家が共同で、パソコン上で音声情報処理技術 (GOP: Goodness of Pronunciation) を駆使して自動的に測定・評価するシステムを開発する。自動評価結果をベテラン英語教員の手動評価結果と比較することで、自動評価システムの妥当性と信頼性を検証した。
- (2) 自動評価に強い影響を与えている要因は何かを、6つの評価観点を設定して調査した。

3. 研究の方法

- (1) 英語熟達度の異なる日本人英語学習者 (TOEIC 得点で 200 点台から 900 点台) が、発話速度 140wpm で提示された 331 語のままとまりのある英文を初見でシャドーイングした音声データを、コンピュータとベテラン日本人英語教員が評価を行った。両者の得点を比較し、自動評価点の有効性を検証した。
 - (2) 音響モデルを基にした GOP (Goodness of Pronunciation) による自動評価に強い影響を与えている要因を調べるため、以下の6つの観点を設定した。
 - ① 5段階による全体的印象評価
 - ② 英文パッセージに含まれる総単語数に対して、シャドーイングで正しく再生された単語は何語あったかを数える全単語再生率に基づく評価
 - ③ 英文パッセージ全体に対する5段階の発音評価
 - ④ 英文パッセージの中でポイントになる単語や語句をいくつか事前に選定し、それらができているかどうかに基づく発音評価
 - ⑤ 英文パッセージ全体に対するリズム・イントネーションの5段階評価
 - ⑥ 英文パッセージの中でポイントになる単語や語句のリズム・イントネーション評価。
- 以上の6つの観点を設定したあと、英語熟達度の異なる日本人英語学習者 36 名 (TOEIC

得点で 200 点台から 900 点台) が同一の英文パッケージをシャドーイングした音声を実動評価し、続いて、同じ音声データを英語教育のベテラン教員が、6 つの観点ごとに採点した。教員による手動評価得点を観点ごとに平均し、GOP を基にした自動評価に対する各観点との Pearson の相関係数を算出した。

4. 研究成果

(1) 日本人英語学習者のシャドーイングについて、GOP による自動採点と手動採点の相関係数は 0.85 となり、コンピュータによる自動採点がシャドーイングの測定・評価システムとして有効であることが明らかになった。また、自動採点による得点と TOEIC による総合的英語熟達力の相関も 0.82 となり、自動評価システムが外的妥当性も十分備えていることがわかった。さらに、シャドーイングと TOEIC の得点を回帰分析することにより、シャドーイングの出来具合から学習者の総合的英語力を TOEIC 得点として予測・算出することができる。そこで、このシャドーイング自動評価システムに、TOEIC 得点自動予測機能も搭載することもできた。

(2) 観点ごとの手動評価得点と GOP を基にした自動評価の Pearson の相関係数については、全体的印象とは 0.843、全単語再生率とは 0.733、ポイントになる単語の発音とは 0.750、全体的発音とは 0.781、ポイントになる語句のリズム・イントネーションとは 0.435、全体的リズム・イントネーションとは 0.391、TOEIC スコアに基づく総合的熟達度とは 0.895 という結果が得られた ($p < 0.01$)。以上の結果から、自動評価との相関が高く、自動評価に与える影響が強いと考えられる要因は、全体的印象評価、発音評価、全単語再生率、総合的熟達度であり、韻律面は比較的弱く、今後は韻律面をより精巧に評価する機能を本システムに搭載していくことが重要であることが明らかになった。

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 8 件)

- ① Y. Yamauchi, K. Husky and A. Kawamura, Development of a Network-based Learning System to Enhance EFL Vocabulary Knowledge 『東京国際大学論叢商学部編』 2009, 査読有, 第 79 号 105-124.
- ② 山内 豊, 峯松信明, 羅徳安 TOEIC スコアが予測可能? 「シャドーイング自動評定システム」 『多聴多読マガジン』 査読無, 2009, 17 号, 30-32.
- ③ Dean Luo, Nobuaki Minematsu, Yutaka Yamauchi, Keikichi Hirose, “Analysis

and Comparison of Automatic Language Proficiency Assessment between Shadowed Sentences and Read Sentences,” *Proc. ISCA Workshop on Speech and Language Technology in Education*, 査読有, 2009, 13-13

- ④ 山内 豊, 峯松信明, 羅徳安, 川村明美, 「音声情報処理技術を活用したシャドーイング自動評価システムの開発」, 『外国語教育メディア学会 49 回全国研究大会講演集』, 査読有, 2009, 178-179
- ⑤ 山内 豊, 峯松信明, 羅徳安, 川村明美, 「シャドーイングの評価方法: 再生率と音響モデルによる評価法は韻律面をどれだけ反映できるのか?」 『外国語教育メディア学会 50 回全国研究大会講演集』, 査読有, 2010, 178-179.
- ⑥ D. Luo, Y. Qiao, N. Minematsu, Y. Yamauchi, and K. Hirose, “Regularized-MLLR speaker adaptation for computer-aided language learning system”, *Proceedings of INTERSPEECH*, 査読有, 2010, 594-597
- ⑦ Yutaka Yamauchi, “Investigation on the possibility of shadowing as an overall proficiency test”, *Proceedings of the 2nd Combined Conference of the Applied Linguistics Association of Australia and Applied Linguistics Association of New Zealand*, 査読有, 2011, 134-134
- ⑧ 山内豊, 峯松信明, 羅徳安, 川村明美, 西川恵, 「シャドーイングの自動評価の妥当性に関する観点別評価面からの考察」, 『外国語教育メディア学会第 51 回全国研究大会発表論文集』, 査読有, 2011, 62-63.

[学会発表] (計 13 件)

- ① D. Luo, N. Minematsu, Y. Yamauchi, & K. Hirose, “Analysis and comparison of automatic language proficiency Assessment between shadowed sentences and read sentences”, ISCA Workshop on Speech and Language, Technology in Education, 2009 年 9 月 3 日 Wroxall Abbey Estate, U.K.
- ② D. Luo, N. Minematsu, Y. Yamauchi, “Analysis and Utilization of MLLR Speaker Adaptation, Technique for Learners’ Pronunciation Evaluation”, INTERSPEECH 2009, 2009 年 9 月 7 日 Brighton Centre, Brighton, U.K.
- ③ 山内 豊, 峯松信明, 羅徳安, 川村明美 「コンテンツ・シャドーイングと英語熟達度の関係」, 外国語教育メディア学会関東支部大会 2009 年 6 月 13 日 拓殖大学

- ④ 山内 豊, 峯松信明, 羅徳安, 川村明美
「音声情報処理技術を活用したシャド
ーイング自動評価システムの開発」, 外
国語教育メディア学会全国研究大会
2009年8月6日, 流通経済大学(神
戸市)
- ⑤ 山内 豊, 峯松信明, 羅徳安, 川村明美,
「総合的英語熟達度がシャドーイング・
パフォーマンスに与える影響の分析」,
外国語教育メディア学会関東支部大会
2010年6月5日, 大東文化大学
- ⑥ 山内 豊, 峯松信明, 羅徳安, 川村明美
「シャドーイングの評価方法: 再生率と
音響モデルによる評価法は韻律面をどれ
だけ反映できるのか?」外国語教育メデ
ィア学会全国研究大会 2010年8月6日
横浜サイエンスフロンティア高校
- ⑦ D. Luo, Y. Qiao, N. Minematsu, Y.
Yamauchi, and K.
Hirose, Regularized-MLLR speaker
adaptation for computer-aided
language learning system, INTERSPEECH
2010, 2010年9月27日, 幕張メッセ国
際会議場
- ⑧ D. Luo, Y. Yamauchi, and N.
Minematsu “Speech analysis for
automatic evaluation of shadowing”,
Second Language Studies: Acquisition,
Learning, Education and Technology,
2010年9月23日 早稲田大学・国際会
議場
- ⑨ Yutaka Yamauchi, ”How to assess
shadowing performance: Manual or
automatic?” 16th World Congress of
Applied Linguistics, 2011年08月25
日, Beijing Foreign Studies University,
China
- ⑩ Yutaka Yamauchi, ”Investigation on
the possibility of shadowing as an
overall proficiency test”, 2nd
Combined Conference of the Applied
Linguistics Association of Australia
and Applied Linguistics Association of
New Zealand, 2011年12月02日
Australian National University,
Australia
- ⑪ 山内豊, 峯松信明, 羅徳安, 川村明美,
「シャドーイングの自動評価の妥当性に
関する観点別評価面からの考察」 外国
語教育メディア学会第51回全国研究大
会, 2011年08月07日, 名古屋学院大
学
- ⑫ 山内豊, 峯松信明, 羅徳安, 川村明美,
西川恵「英語シャドーイング評価の試み:
全体的印象評価と観点別評価の関係」
外国語教育メディア学会関東支部 126
回研究大会 2011年06月11日

- 学習院女子大学
- ⑬ 山内豊, 峯松信明, 川村明美, 西川恵
英語シャドーイング評価における評価
者間のずれと信頼性に関する分析 外
国語教育メディア学会関東支部第127
回研究大会, 2011年11月12日
拓殖大学
〔図書〕(計0件)

〔産業財産権〕
○出願状況(計0件)
名称:
発明者:
権利者:
種類:
番号:
出願年月日:
国内外の別:

○取得状況(計0件)
名称:
発明者:
権利者:
種類:
番号:
取得年月日:
国内外の別:

〔その他〕
ホームページ等

6. 研究組織

(1) 研究代表者

山内 豊 (YAMAUCHI YUTAKA)
東京国際大学・教授
研究者番号: 30306245

(2) 研究分担者

峯松 信明 (MINEMATSU NOBUAKI)
東京大学・大学院工学系研究科・教授
研究者番号: 90273333
川村 明美 (KAWAMURA AKEMI) 東京
国際大学・言語コミュニケーション学部・
准教授
研究者番号: 30326996
ハスキー・ケイ (HUSKY KAY) 東京国際
大学・商学部・准教授
研究者番号: 50237955
染谷 泰正 (SOMEYA YASUMASA) 関
西大学・外国語学部・教授
研究者番号: 40348454
西川 恵 (NISHIKAWA MEGUMI) 東海
大学・外国語教育センター・准教授
研究者番号: 10453705

(3) 連携研究者