

平成 30 年 6 月 15 日現在

機関番号：32689

研究種目：基盤研究(C) (一般)

研究期間：2014～2017

課題番号：26370746

研究課題名(和文) 発話自動採点システムの開発と英語教育プログラムへの導入

研究課題名(英文) Development of automated scoring system for L2 speech and examination of its practicality

研究代表者

近藤 悠介 (Yusuke, Kondo)

早稲田大学・グローバルエデュケーションセンター・准教授

研究者番号：80409739

交付決定額(研究期間全体)：(直接経費) 3,300,000円

研究成果の概要(和文)：本研究では発話能力の育成に焦点を当てた大規模な英語教育プログラムにおける自動採点システム導入の可能性を検討した。導入を検討したプログラムは、発話能力の育成に焦点を当てたプログラムであり、学習の対象となる表現が適切に使用できるかどうかを判定するタスクを作成し、このタスクにおける発話を英語教員が採点し、この点数を予測する発話自動採点システムを開発した。システムの予測精度を検証したところ、教師による点数との一致度は74%であった。本研究で提案した枠組みを用いてクラス分け試験および到達度試験を自動採点システムによって行うことができる可能性は高いと判断した。

研究成果の概要(英文)：We developed an automated scoring system for speech by Japanese learners of English and examined its practicality in an English language program. The system assesses the students' responses to tasks that were created to elicit the expressions learned in the program. The exact agreement is 74% between our system and the human scores.

研究分野：言語テスト

キーワード：言語テスト 英語の発話能力 自動採点

1. 研究開始当初の背景

(1) 2006年にはTOEIC SWの運用が開始され、同時にTOEFLではSpeaking Sectionが導入された。また、国内の大学入試において、英語の試験は読解、聴解だけでなく、「話す」、「書く」技能の試験が取り入れられつつある。このような動向に鑑みて、国内の英語教育の多くの場面で「話す」、「書く」能力の評価・測定を導入しようという気運が高まっているが、これらの能力の評価が取り入れられている試験の例は少ない。理由として、評定者の訓練を行い、「話す」、「書く」能力を直接評価する試験を実施する人的、時間的コストが膨大であることが挙げられる (Clouser, Margolis, Clyman, & Ross, 1997; Zechner, Higgins, Xi, & Williamson, 2009)。さらに、訓練を行っても評定者による評価のぶれがなくなることはないこと、信頼性の高い評定者を継続的に確保する難しさなども「話す」、「書く」能力の評価が導入されていない理由のひとつである。また、「話す」、「書く」能力の試験では、試験が大規模なものであれば、受験者が即時的なフィードバックを得ることができないということもこの種の試験の問題点として挙げられる。これらの問題点を解決する方法のひとつとして、自動採点システムの導入が考えられる。自動採点システムを導入することによって、この種の試験に関わる人的、時間的コストが削減でき、ぶれのない採点が可能となり、即時的なフィードバックを受検者に与えられる。

(2) 発話能力を直接評価しようとした場合、いくつかの問題が挙げられる。まず、評定者の訓練、試験の実施、評価値の分析などに膨大な人的・時間的コストがかかる。訓練を施した評定者を継続的に確保することも難しい。さらに、評定者の導入は評定者間の評価の不一致という新たな問題を引き起こす。発話自動採点システムの開発は、この問題を解決するひとつの手段として提案されている。

(3) 発話自動採点の研究は1990年ごろから発表され (Bernstein, Cohen, Murveit, Rtschev, & Weintraub, 1990)、現在ではVersant (Pearson Knowledge Technologies, 2011) や TOEFL Practice Online で利用されているSpeechRater (Higgins, Xi, Zechner, & Williamson, 2011; Zechner, Higgins, Xi, & Williamson, 2009) など実用化されているものもある。また、音声認識技術を利用した英語学習用ソフトウェアもいくつか実用化されている (Akahane-Yamada et al., 2004; Chiu, Liou, & Yeh, 2007; Tsai, 2003)。

(4) 研究代表者は、平成21年度～22年度に若手研究(B)「英語学習者のための韻律自動評価システムの構築とその評価」(課題番号21720209)の助成を受け、英語学習者のための音読自動採点システムを開発した。イソッ

ブ童話の『北風と太陽』をアジア人英語学習者101人が読み上げたものをヨーロッパ言語参照枠に基づき訓練された英語教師が評価した。このデータの発話の特徴量と評価値の関係を検証し、発話の特徴量から評価値を予測するシステムを構築した。人間の評定者が付与した評価値とシステムが算出した評価値を比較したところかなりの精度の一致が見られ、音読における自動採点の実用性が示唆された。

(5) この研究成果をもとに、本研究代表者は、平成23年度～25年度に基盤研究(C)「英語学習者の総合的な発話能力を自動評価するシステムの構築とその評価」(課題番号23520721)の助成を受け、既存のシステムを発展させ、英語学習者のさまざまな発話を誘出するタスクを作成し、それらの発話を自動評価できるシステムの構築を試みた。音読の自動評価では、主に韻律に焦点を当てた採点を行ったが、この研究では、発話の適切さや文法的な誤りも採点の対象とした自動採点システムの開発を行った。発話誘出タスクの作成は、ヨーロッパ言語参照枠(Common European Framework of Reference: CEFR)を参照し、談話完成タスクを基本に翻訳タスクや読み上げタスクを含めた。このタスクを使用して、日本人、台湾人、韓国人英語学習者から発話データを収集し、これらの発話に評定者が評価値を付与し、評価値と発話の特徴量の関係を検証した。既存データにおける評価値と発話の特徴量をもとに機械学習で使用されている分類法を採用し、新たな受検者の評価値を予測するシステムを構築した。

2. 研究の目的

(1) これらの成果を踏まえ、本研究は、さまざまな発話誘出タスクを使用し、発話評価を予測するシステムの開発を目的とした。

3. 研究の方法

(1) 本研究で対象とする英語教育プログラムは、授業1回ごとに受講者が習得すべき表現が設定されている。そのため、自動採点システムで使用されたタスクは、これら習得すべき表現が適切に使用することができるかを問うタスクである。以下に例を示す。

You (A) want to end your conversation. What would you say in the conversation below?

A: () .

B: See you.

このようなタスクを40問作成し、63人の英語学習者から発話を収集した。未回答の121を除き、2399発話を評価の対象とした。例に示したようなタスクへの応答を表1に示す評価基準を参照し、評定者が評価した。評価基準は本研究で独自に作成したものである。

表 1
評価基準

点数	記述
3	状況に合った表現が使用されている。
2	意図は理解できるが、状況に合った適切な表現ではない。
1	状況に合っていない(状況を理解していない)

この評価基準を用いてすべての発話を 3 人の教員が評価した。評定者による評価のばらつきを考慮し、すべての評定者が 3 と評価したものを 1、それ以外を 0 とした。14 人の評定者が発話を評価したが、図 1 に示すように発話によって異なる。

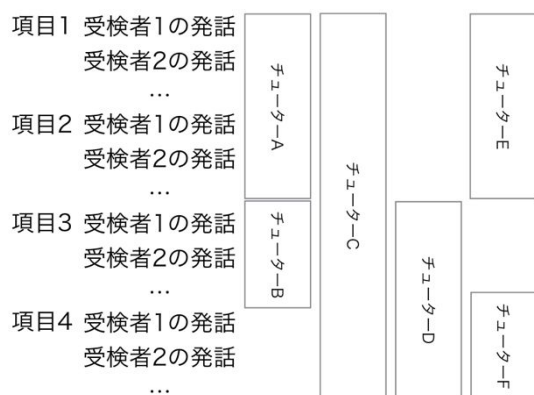


図 1 評定者の分担

また、評定者が評価した発話の 20% は重複した発話であり、このデータを利用し評定者内信頼性を算出し、信頼性の低い評定者の評価は利用しなかった。その場合も、2 人の評定者が 3 と評価した発話は 1、それ以外を 0 とした。

4 . 研究成果

- (1) これらの項目に関して個別にサポートベクターマシンおよびナイーブベイズ分類器を用いて評価を予測した。5 分割交差検証法を用いて個々の項目での予測精度を推定した。項目により差はあるが、どちらの予測方法においても平均値は.74 であった。
- (2) 本研究では英語教育プログラムの内容に沿った発話自動採点システムを開発し、システムの導入可能性を検討した。予測精度をもとに考えると、本システムは英語教育プログラムにおけるクラス分け試験および到達度確認試験として導入が可能であると考えた。

< 引用文献 >

Akahane-Yamada, R., Kato, H., Adachi, T., Watanabe, H., Komaki, R., Kubo, R., ... Kawahara, H. (2004). ATR CALL: A speech perception/production training system utilizing speech technology. Proceedings of

The 18th International Congress on Acoustics III, 2319-2320.

Bernstein, J., Cohen, M., Murveit, H., Rtschev, D., & Weintraub, M. (1990). Automatic evaluation and training in English pronunciation. Proceedings of the International Conference on Spoken Language Processing. 1185-1188.

Chiu, T. L., Liou, H. C., & Yeh, Y. (2007). A study of web-based oral activities enhanced by automated speech recognition for EFL college learning. Computer Assisted Language Learning, 20(3), 209-233. doi:10.1080/09588220701489374

Clauser, B. E., Margolis, M. J., Clyman, S. G., & Ross, L. P. (1997). Development of automated scoring algorithms for complex performance assessments: a comparison of two approaches. Journal of Educational Measurement, 34, 141-161. doi:10.1111/j.1745-3984.1997.tb00511.x

Higgins, D., Xi, X., Zechner, K., & Williamson, D. (2011). A three-stage approach to the automated scoring of spontaneous spoken responses. Computer Speech & Language, 25(2), 282-306. doi:10.1016/j.csl.2010.06.001

Tsai, P. H. (2003). A duet of pedagogy and technology—an evaluation of My ET, a computer assisted pronunciation training system made in Taiwan. Proceedings of the 21st International Conference on English Teaching and Learning in the R.O.C. 439-452.

Zechner, K., Higgins, D., Xi, X., & Williamson, D. M. (2009). Automatic scoring of non-native spontaneous speech in tests of spoken English. Speech Communication, 51, 883-895. doi:10.1016/j.specom.2009.04.009

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 4 件)

近藤悠介・石井雄隆. (2017). 「英語学習者の発話自動採点システムの開発と英語教育プログラムへの導入可能性の検討」『Language Education & Technology』54. 23-40.

近藤悠介・石井雄隆. (2015). 「機能シラバスに基づいた英語教育プログラムにおける発話自動採点システム導入可能性の検討」『外国語教育メディア学会第 55 回全国研究大会予稿集』. 90-91.

Kondo, Y., & Ishii, Y. (2015). Demonstrating practicality of automated scoring system for L2 speech in English language program.

Proceedings of the 20th International Conference of Pan-Pacific Association of Applied Linguistics. 110-111.

近藤悠介. (2015). 日本人英語学習者の短い発話を自動採点するシステムの実現可能性の検討. 情報処理学会第 77 回講演論文集. 第 4 巻. 497-498.

〔学会発表〕(計 10 件)

石井雄隆・金丸敏幸・小野雄一・近藤悠介 (2017 年 11 月 18 日). 「テクノロジーと教育・学習・評価-自然言語処理・学習支援システム・自動採点の観点から-」外国語教育メディア学会 (LET) 関東支部第 139 回(2017 年度秋季)研究大会, 東洋大学, 埼玉.

中野美知子・大和田和治・荊紅涛・後藤裕介・近藤悠介・吉田諭史. (2015 年 7 月 25 日). E-learning の現状とこれから. 言語文化教育学会 2015 年度公開シンポジウム, 早稲田大学, 東京.

小泉利恵・近藤悠介・斉藤一弥. (2015 年 5 月 23 日). 英語教育における音声指導と評価:スピーキングと発音. 外国語教育メディア学会(LET)関西支部 2015 年度春季研究大会シンポジウム, 大阪電気通信大学, 大阪.

Kondo, Y., & Ishii, Y. (2016). What can programming language do for our analysis? : The usefulness of jupyter notebook in the analysis of language testing. Proceedings of the 21st International Conference of Pan-Pacific Association of Applied Linguistics. 55-56.

近藤悠介・石井雄隆. (2016,5 月 21 日). 「コンピュータによる発話自動採点の導入可能性を考える: MyET の利用調査」. 外国語教育メディア学会(LET)関西支部関西支部春季大会. 神戸学院大学, 兵庫.

中野美知子・近藤悠介・大和田和治・吉田諭史・荊紅涛. (2015 年 8 月 29 日). JACET-ICT 調査研究特別委員会シンポジウム ICT を活用した異文化間コミュニケーション能力の養成. 大学英語教育学会第 54 回 (2015 年度) 国際大会, 鹿児島大学, 鹿児島.

石井雄隆・近藤悠介. (2015 年 8 月 5 日). 機能シラバスに基づいた英語教育プログラムにおける発話自動採点システム導入可能性の検討. 外国語教育メディア学会 (LET) 第 55 回全国研究大会, 千里ライフサイエンスセンター, 大阪.

Kondo, Y. & Ishii, Y. (16, May, 2015). Evaluation of the reliability for L2 speech rating in discourse completion test. 2015 Joint International Methodology Research Colloquium. National Institute for International Education (NIIED), Seoul, South Korea.

近藤悠介. (2015 年 5 月 11 日). 英語教育プロ

グラムにおける発話自動採点. 早稲田大学理工学研究所 特別研究「次世代 e-learning に関する研究」部会「次世代 e-learning」公開フォーラム(第 6 回), 青山学院大学, 東京.

中野美知子・大和田和治・近藤悠介・吉田諭史・上田倫史. (2014 年 8 月). ICT 調査研究特別委員会報告 ICT を活用した英語教育: Wiki、電子教材、発話自動採点. 大学英語教育学会第 53 回 (2014 年度) 国際大会, 広島.

〔図書〕(計 1 件)

近藤悠介. (2015). 「第 13 章 英語学習者を対象とした発話自動採点システム」. 中野美知子 編著. 『英語教育の実践的探究』. 溪水社.

6. 研究組織

(1) 研究代表者

近藤 悠介 (KONDO, Yusuke)
早稲田大学・グローバルエデュケーションセンター・准教授
研究者番号: 80409739

(2) 研究分担者

石井 雄隆 (ISHII, Yutaka)
早稲田大学・大学総合研究センター・助手
研究者番号: 90756545

中野 美知子 (NAKANO, Michiko)
早稲田大学・名誉教授
研究者番号: 70148229

吉田 諭史 (YOSHIDA, Satoshi)
早稲田大学・グローバルエデュケーションセンター・講師
研究者番号: 00608838