

談話完成タスクにおける第二言語の音声認識率向上の検討

近藤 悠介[†] 中野 美知子^{††}

早稲田大学 オープン教育センター[†]

早稲田大学 教育・総合科学学術院^{††}

1. はじめに

早稲田大学オープン教育センターが設置する Tutorial English という科目は、スピーキングを中心とした英語の運用能力の向上を目指した科目で、受講者 4 人に対しチューター1 人の少人数制授業である。この授業での受講者の学習をサポートするために、さまざまな仕組みが用意されている。受講者は受講前にインターネットを介して受検する WeTEC と呼ばれるレベル分けテストを受検し、解答終了後即座に受講するレベルを決定される。授業中のフィードバックやチューターからのコメントは Tutorial Site と呼ばれる専用のサイトでやり取りされる。さらに、この授業は 20 課で構成されており、7 課、14 課、20 課の後に Lesson Review Test と呼ばれる復習問題が用意されており、このテストもネットワークを介して受検することができ、解答終了後に正解率などのフィードバックが与えられる。

このようにネットを介した学習をサポートするためのさまざまなサービスが提供されており、受講生 4 人に対しチューター1 人という少人数制授業で受講者のスピーキング能力の向上に重点を置いているが、受講者のスピーキング能力を直接評価・測定するは用意されていない。

発話能力の評価・測定が実施されない理由としては、評定者の育成に時間がかかる、評価が評定者間で一致しない、発話能力を評価・測定するテストの実施に人的、時間的コストがかかるなどが挙げられる。これらの問題を解決するために発話能力の自動採点システムの構築が望まれる。

現在、運用されている発話自動採点システムはいくつかあるが、これらのシステムはそれぞれの目的で構築されているため、これらのシステムを Tutorial English における発話能力の評価・測定のためのテストとして採用できない。そこで本研究では Tutorial English の授業内容を反映し

た談話完成タスク (Discourse Completion Task: DCT) を用いて英語の発話能力を自動的に採点するシステムの構築を計画している。本発表ではこのシステムの構築の基礎となる音声認識の精度に関して報告する。

2. DCT の項目

Tutorial English は Beginner、Basic、Pre-Intermediate、Intermediate、Pre-Advanced、Advanced の 6 つのレベルで構成され、それぞれのレベルには 20 のレッスンが含まれている。本研究ではこの中から下から 2 番目のレベルである Basic を対象とした。

それぞれのレッスンにはテーマが設定されており、例えば、Asking about personal plan というテーマのレッスンでは以下のような表現が学習対象となっている。各課ではこのような表現が使用できる状況が与えられロール・プレイをしたり、自分自身について話す実際的な練習を行っている。

- Where are you up to this weekend?
- Do you have any plan for the week end?
- Are you busy this week end?
- I have to work all weekend.
- I'm visiting my grandparents in Osaka.
- I have a biology test on Friday.

本研究では、このような表現を受験者から誘出できるような Basic レベルの DCT を 132 項目作成した。以下はその一例である。

If you (A) would like to ask your friend what he/she plans to do over the weekend, how would you say in the conversation below?

A: ()

B: I have to work all weekend!

3. データ

本研究で使用したデータは日本人大学生約 80 人より提供された発話データである。発話データ提供者は録音専用のウェブサイトにアクセスし

A study on the enhancement of word recognition rate of second language learners' speech in discourse completion task
Yusuke Kondo, Open Education Center, Waseda University
Michiko Nakano, Faculty of Education and Arts and Integrated Sciences, Waseda University

個別にデータを提供した。ひとつの項目に対し78人から86人分の発話が提供された。発話の総数は10617である。

4. 音声認識機

Hidden Markov Model Toolkit を用いて隠れマルコフモデルに基づいた音声認識機を作成した。音響モデルの作成には6300の発話を含む TIMIT Acoustic-Phonetic Continuous Speech Corpus [1]および日本人英語学習者による約80000の発話が収録された ERJ データベース[2]を利用した。言語モデルは132項目それぞれの受検者の回答から2-gramモデルを作成した。

5. 認識結果

5.1 男女混合モデル (項目別)

Unit 3 の項目 1 から 6 の認識結果を表 1 に示す。

表 1: Unit 3 の認識結果

Item	N	H	D	S	I	C	A
3-1	688	602	44	42	20	.87	.84
3-2	656	656	28	40	43	.89	.83
3-3	740	603	45	92	36	.81	.76
3-4	604	468	61	75	48	.77	.70
3-5	600	514	46	40	37	.86	.80

注: N: 総語数 H: 正しく認識できた語の数 D: 脱落の数 S: 置き換えの数 I: 挿入の数 C: H/N A: (H-I)/N

5.2 話者適応 (個人別)

男女混合、男女別に音響モデルを作成し、Unit 17 の項目 1 から 7 の発話を認識した結果をそれぞれ表 2、3 に示す。

表 2: 男女混合モデルの認識率

	平均	最少	最大
男性	70.0	42.4	97.8
女性	64.0	34.7	90.7

注: 認識率は(正しく認識できた語数-挿入語数)/総語数

表 3: 男女別モデルの認識率

	平均	最少	最大
男性	61.0	35.4	88.5
女性	76.7	56.5	93.9

注: 認識率は(正しく認識できた語数-挿入語数)/総語数

男女別の音響モデルを用い、各個人の Unit 1 から 3 での発話で話者適応 (Maximum Likelihood Linear Regression) を行った結果を表 4 に示す。

表 4: 話者適応後の認識率

	平均	最少	最大
男性	64.3	38.8	88.5
女性	76.5	57.1	91.8

注: 認識率は(正しく認識できた語数-挿入語数)/総語数

図 1 に男女混合モデル、男女別モデル、話者適応後での平均値の変化を示す。

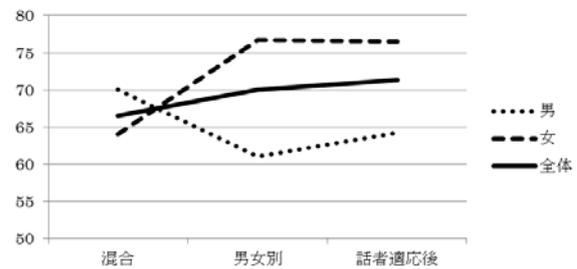


図 1: モデル別の平均値の変化

6. 考察

Unit 3 の結果を見るとある程度許容できる認識率を得た項目もあるが、項目によるばらつきは大きく、認識率を下げている要因を探る必要がある。また、男女混合モデルと男女別モデルとの認識率を比較すると、男性は男女別モデルにしたことにより認識率は下がり、女性では認識率が上がることが分かった。この原因として、男女による英語の熟達度の差が考えられる。一般的に女性は男性に比べ熟達度が高く、このことが認識率の差として現れた可能性がある。また、話者適応の効果は男女ともに小さいものであった。認識率を向上させるため、男女の熟達度の差、話者適応の方法を検討する必要がある。

DCT ではかなり高い認識率を得ることできる項目もあるため、これを利用し、文法的な正確さ、発話の的確さ、発話の丁寧さを評価できる可能性がある。

参考文献

- Garofolo, J. S., Lamel, L. F., Fisher, W. M., Fiscus, J. G., Pallett, D. D., Dahlgren, N. L., & Zue, V. (1993). TIMIT Acoustic-Phonetic Continuous Speech Corpus. Philadelphia: Linguistic Data Consortium.
- 峯松信明, 富山義弘, 吉本啓, 清水克正, 中川聖一, 壇辻正剛, 牧野正三. (2003). "英語 CALL 構築を目的とした日本人及び米国人による読み上げ英語音声データベースの構築", 日本教育工学会論文誌 (「第二言語学習とその支援に関する教育工学研究」特集号) Vol.27, No.3, 259-272.

謝辞

本研究は、独立行政法人日本学術振興会学術研究助成基金助成金(基盤研究(C) 課題番号: 23520721)の助成による。