

動画スクリプトを利用した多肢選択問題の自動生成

吉田 葵, 青柳 龍也, 来住 伸子

津田塾大学

現在、津田塾大学では、専門課程の内容重視の英語教育を目的とした英語教材協調開発システム COOLL を構築している。COOLL 上にマルチメディアコンテンツの情報を集約し、そのコンテンツに対して問題作成などの教材開発を行い英語学習者に提供する。学習教材は人手で開発されるため、開発コストがかかる。そこで、本論文では動画のスクリプト中に含まれる単語を利用した多肢選択問題を自動的に生成する手法を提案する。また、自動生成した問題の検証及び実証実験を行った。本提案手法により、スクリプトが用意された動画であれば、動画に関連した問題を自動生成できるため、効率的に教材を開発することが可能になる。

Automatic Generation of Language Multiple-choice Quizzes Based on Video Transcripts

Aoi YOSHIDA, Tatsuya AOYAGI, Nobuko KISHI

Tsuda College

The rise of the Internet has enabled ESL learners to watch authentic videos in English more easily than before. ESL learners often need added learning materials to enhance their learning utilizing authentic videos. Nonetheless, learning materials are not enough since the development of them takes a great deal of cost and time. We have designed the algorithm that automatically generates multiple-choices based on video transcripts. By embedding this algorithm in COOLL, Tsuda College's online learning system, learning materials can be efficiently provided for ESL learners. We also report the result of experiments where 37 students answered the automatically generated quizzes.

1. はじめに

1. 1 研究背景・目的

内容重視の英語教育の目標は、学習者の専門分野の内容を英語で学ぶことにより、英語スキルと専門分野の知識の両方を同時に獲得することにある[1]。内容重視の英語教育は、学習者の英語レベルや専門性のレベルを問わず効果的であり、専門課程の大学生を対象とした、学究領域中心の指導にも用いられる。内容重視の英語教育で利用する教材は、英語をコミュニケーション・ツールとして用いる人に向けて作成されたものである必要がある。さらに、学習者の学習意欲を損なわないために、学習者の英語スキル、専門性のレベルにあった教材である必要があり、書籍だけに限ら

ず、動画などのマルチメディアコンテンツを利用し、学習者の学習意欲を引き出すことも重要である。

近年、YouTube をはじめとしたオンラインビデオ共有サービスの普及により、英語による様々なマルチメディアコンテンツがインターネット経由で入手可能になった。これらのコンテンツの中には、国際機関や非営利団体が、教育利用を認めたものも多く、これらを効率よく教材として利用できれば、教材開発のコストを低く抑えることができる。

現在、津田塾大学では、マルチメディアコンテンツを集約し、内容重視の英語教育を実施するために必要な教材の開発を支援する英語教材協調

開発システム COOLL (Collaborative Open Language Learning) を運用している。2010 年 10 月からは学外公開も始まった。

COOLL は、Web 上に点在するマルチメディアコンテンツの情報やそのコンテンツに関するレビューを集約し、ユーザ間で情報共有できる機能を持つ。そして、COOLL では、集約されたコンテンツに対して、英語学習者の動画に対する理解を助けるために、英語字幕や日本語字幕を作成する機能を持ち、さらに、動画の内容に基づいた問題といった動画と関連のある学習教材を、様々な分野の専門家が協調して開発することを支援する。また、開発された教材は、COOLL と補完的に用意された Moodle を用いて、英語学習者に提供される (図 1)。

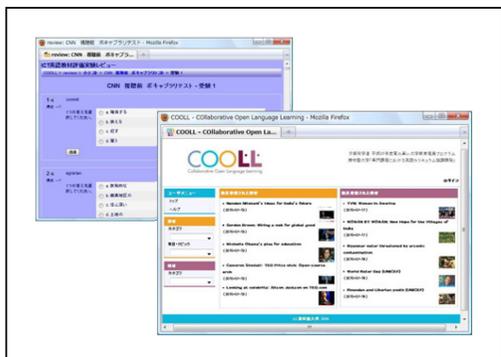


図 1 COOLL と Moodle 上の学習教材
(COOLL : <http://cooll.tsuda.ac.jp>)

現在、COOLL 上に登録されているコンテンツ数が 550 であるのに対して、提供されている学習教材は 90 であり、コンテンツ数に対して、学習教材は十分に提供されているとはいえない。これは、学習教材の開発がすべて人手で行われていて、時間や手間がかかるからである。また、自習のために使われる学習教材は個人の学習履歴や英語能力に適したパーソナルなものを提供した方が高い学習効果を得られるが、人手で行うとなると、さらに多くの時間や手間がかかってしまう。

そこで、本論文では、まず、教材開発の効率化の第一歩として、動画スクリプトを利用し、動画に関連した多肢選択問題を自動生成する手法を提案する。自動生成する多肢選択問題は、英単語の適切な日本語訳を選ぶ問題である。

1. 2 関連研究

問題の自動生成に関する研究は、プログラミング分野など様々な分野で行われている [3][4]。英語学習における多肢選択問題の自動生成 [5] では、問題文や正答誤答等のタグ付けして分類したデータをデータベースに蓄積し、そのデータを再構築して出題しているが、オンライン動画の数は日々増えるので、問題データの蓄積が間に合わなくなってしまうと考える。また、穴埋め問題の自動生成の研究 [6] では、Web をコーパスとして利用し、出題文の決定や不正解選択肢の検証を行っているが、英単語のみの日本語訳を Web から誤用かどうかを判断するのは難しい。

2. 多肢選択問題自動生成手法

自動生成する多肢選択問題は、動画のスクリプト中に含まれる英単語の適切な日本語訳を 4 つの選択肢から選ぶ問題である。自動生成アルゴリズムは、出題単語の選定、不正解選択肢の選定、日本語訳の選定、問題の生成の、4 工程からなる。

教材開発の効率化を目的とするため、動画のスクリプトさえあれば問題を素早く大量生成できるのを第一に考え、英文構造の解析などは行わない、非常にシンプルな手法である。

2. 1 出題単語の選定

出題単語の選定は、単語が習熟度や難易度に基づいて順位づけされていて、さらに、各単語の品詞情報を持つリストを用いる。

まず、動画のスクリプトを単語ごとに分け、各単語がリストに含まれるかどうかを調べる。含まれる場合はその品詞と順位を調べる。その単語の品詞がリストに複数あった場合には、先頭の品詞を調べる。次に、含まれなかった単語に対して、語幹処理を行う。そして、語幹処理した各単語が、リストに含まれるかどうかを調べ、含まれる場合は同様の処理をする。そして、リストに含まれる単語を順位でソートし、先頭から問題数分の単語を出題単語として選定する。

今回は、JACET8000 を利用し、出題単語を選定する。JACET8000 は、大学英語教育学会が作成した、日本人英語学習者が学ぶべき単語のリストである [7]。British National Corpus と、日本の英語教材や試験に使われる表現のサブコーパス等をもとに、1 位から 8000 位まで順位付けさ

れている。1000 語単位で 8 段階にレベル分けされていて、各レベルに英語学習度の目安が設けられている。7001~8000 位がレベル 8 となり、「日本人英語学習者の一般的な単語学習の最終到達目標」とされている (図 2)。また、JACET8000 には、各単語の順位に加え、品詞情報も含まれている。

JACET8000 では、8000 位に近い方が「難しい」単語であるので、動画スクリプト中のリストに含まれる単語を順位で降順にソートして、選定することにする。

Level 1 (順位1~1,000位)	中学校の英語教科書に頻出する基本的な単語
Level 2 (順位1,001~2,000位)	高等学校初級レベルの単語
Level 3 (順位2,001~3,000位)	大学入試センター試験に出てくる単語。英検2級。
Level 4 (順位3,001~4,000位)	大学受験、大学一般教養の初級レベル
Level 5 (順位4,001~5,000位)	難関大学受験、大学一般教養レベル。 英検準1級。TOEIC400~500点。
Level 6 (順位5,001~6,000位)	英語を専門としない大学生やビジネスマンが目指すべきレベル。 英検準1級。TOEIC600点。
Level 7 (順位6,001~7,000位)	英語専攻の大学生や英語を仕事で使うビジネスマンの到達目標。 英検1級。TOEIC95%以上。
Level 8 (順位7,001~8,000位)	日本人英語学習者の一般的な単語学習の最終到達目標。

図 2 JACET8000 レベル一覧

2. 2 不正解選択肢の選定

不正解選択肢は正解選択肢と文法的特徴や何らかの観点で共通点を持っている必要がある。文法的特徴として品詞情報を用い、さらに、同程度の順位の単語は同程度に難しいので、これらを不正解選択肢の選定に利用する。

出題単語と同じ品詞の単語を、JACET8000 の出題単語の順位から小さい方へ向って探索し、不正解選択肢となる 3 単語を選定する。異なる出題単語で、不正解選択肢として同じ単語が選定された場合、あるいは、出題単語が他の問題で不正解選択肢として選定された場合でも、特に考慮しない。

2. 3 日本語訳の選定

出題単語と不正解選択肢として選定された 3 単語の日本語訳を、英和辞書データから選定する。多義語である場合には、英和辞書データの先頭の訳を利用する。各単語の日本語訳に重複が起きた場合には、不正解選択肢の単語を選定し直す。

今回使用した英和辞書データ GENE95 には、

品詞情報が含まれていないため、日本語訳を選定する際に品詞は考慮していない。

2. 4 問題の生成

出題単語と選定した日本語訳を、1 つの問題として出題可能な形に整形し出力する。

今回は、Moodle にインポート可能な Moodle XML フォーマットで出力し、Moodle を利用して出題する (図 3)。

```
<?xml version='1.0' encoding='Shift-JIS' ?>
<quiz>
  <question type='multichoice'>
    <name>
      <text>動画 1-1</text>
    </name>
    <questiontext format='html'>
      <text>insert</text>
    </questiontext>
    <answernumbering>123</answernumbering>
    <shuffleanswers>1</shuffleanswers>
    <single>true</single>
    <answer fraction='100'>
      <text>挿入する</text>
      <feedback><text></text></feedback>
    </answer>
    <answer fraction='0'>
      <text><つついて離れない</text>
      <feedback><text></text></feedback>
    </answer>
    <answer fraction='0'>
      <text>知らせる</text>
      <feedback><text></text></feedback>
    </answer>
    <answer fraction='0'>
      <text>もてなす</text>
      <feedback><text></text></feedback>
    </answer>
  </question>
</quiz>
```

図 3 Moodle XML フォーマット出力例

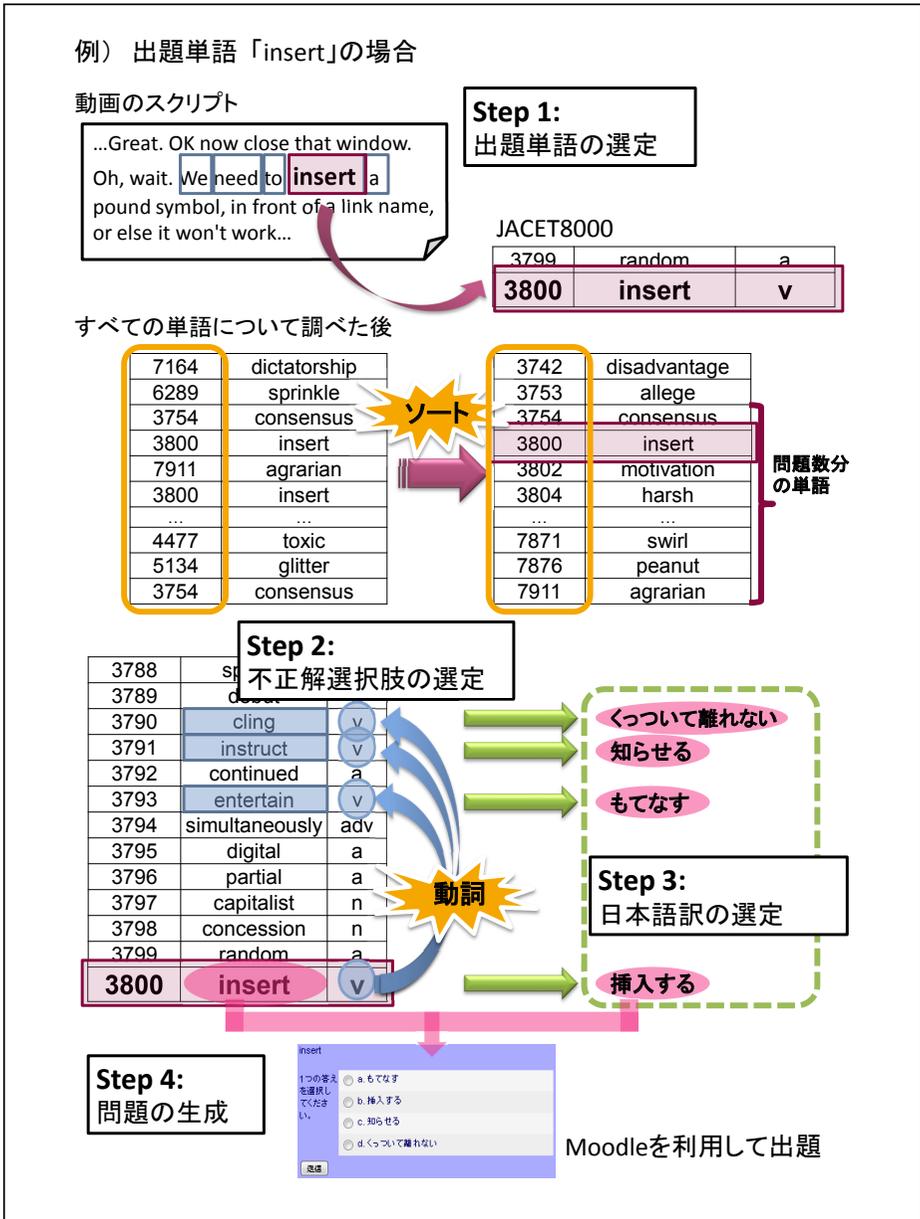


図 4 多肢選択問題自動生成手法の流れ

3. 自動生成した問題の検証

2節で述べた手法を用いて、8つの動画に対し、各25問、計200問の4択問題を自動生成し、問題の質を検証した。また、大学生37名に出題単語200個を提示し、各単語に対して「意味がわかる」「たぶんわかる」「意味はわからない」「初めてみた」の4段階で評価する認識度調査を行った。学生の自己申告によるTOEFLスコア(TOEIC、TOEFL IBT、TOEFL ITPの各スコアをTOEFL ITPに換算したスコア[8])は、420～620点である。

検証に用いた8つの動画は、あらかじめスクリプトが用意されている、2つのWebサイト(CNN Student News[9], Howcast.com[10])から選定した。各動画の長さは3分程度、スクリプトの語数は450～700語である。

3. 1 出題単語の検証

3. 1. 1 出題単語のJACET8000でのレベル

出題単語として選定された単語は、JACET8000のレベル4(大学一般教養レベル)以上が78%を占め(図5)、大学生向けの問題としては妥当なレベルだと考えられる。

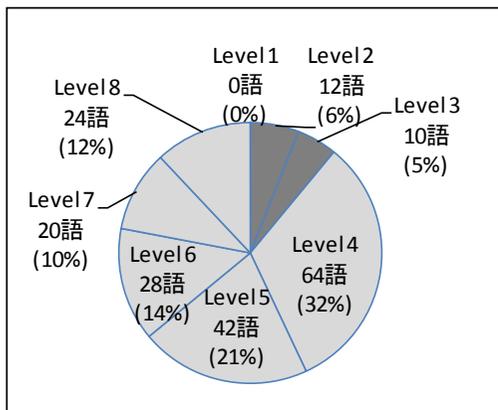


図5 出題単語のJACET8000でのレベル

また、認識度調査で、各単語に対し、「初めて見た」と答えた学生の割合とJACET8000での順位との相関係数は0.455で、相関があった(図6)。さらに、TOEFLスコアを基準にして学生を3グループに分けたとき、JACET8000のレベルにおける未知の単語が含まれる割合がグループごとに異なっていた(図7)。

これらの結果から、出題単語の選定の基準として、JACET8000を用いると、学習の目的や学習者の英語能力に応じて、出題単語の選定条件を調整することができると言える。

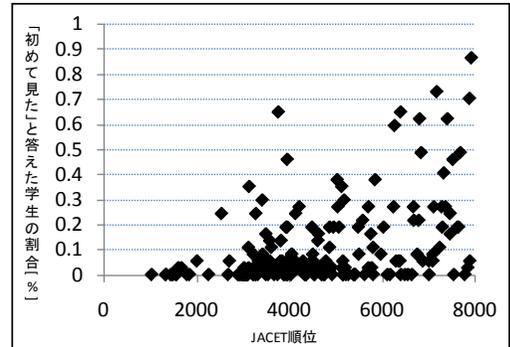


図6 認識度とJACET8000での順位の関係

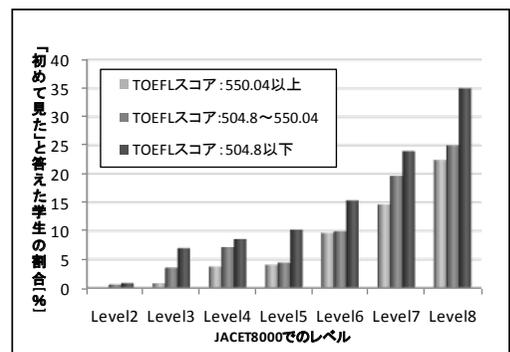


図7 TOEFLスコアによる認識度の違い

3. 1. 2 出題単語候補にはならなかった単語

動画スクリプト中で、JACET8000に含まれない単語は、重複を除いた全単語の19.2%であった。固有名詞や数詞(例 America, two)、ハイフンでつながれた単語(例 cancer-fighter, self-esteem)、ある分野に固有な単語(例 blueberry, hacker)、難しい単語(例 pungent, edible)などがリストに含まれない単語である。

これは、JACET8000が一般的な英語学習者を対象としたリストであるためだと考えられる。しかし、動画の内容を理解するためには必要な単語も多く含まれているので、今後、動画の内容理解まで踏み込んだ問題を自動生成するためには、JACET8000に加え、専門用語が含まれる単語集を利用するなど、このような単語も出題単語とし

て選定できるような手法を検討する必要がある。

3. 2 正解選択肢の検証

正解選択肢 200 個のうち、動画スクリプト中の品詞と正解選択肢の品詞が異なるものは、22 個だった。これらは、語幹処理を行う前の時点でリストに含まれてしまっていた単語（例 -ing, -ed, -s）、複数の品詞情報を持つ単語である。

また、正解選択肢の日本語訳で、文脈に完全に合っていなかったものは、27 個で、品詞情報の間違いから日本語訳の品詞は合っていないが意味的には合っているものは、6 個だった。前者の単語は、まったく違う意味を複数持つ単語（例 straw, medium）である。

正解選択肢の日本語訳が文脈に合ったものではないということは、動画に関連した問題であることに沿わないので、改善する必要がある。英文構造の解析を行わないシンプルな手法を目指しているので、集合知を取り入れ、文脈に合った日本語訳を正解選択肢とする手法を検討したい。

3. 3 不正解選択肢の検証

不正解選択肢 600 個のうち、不正解選択肢を選定する際に指標にした品詞と選択肢の日本語訳の品詞が違う選択肢は 78 個であった。

今回利用した辞書データに品詞情報が含まれていないにも関わらず、全選択肢の 87%で、品

詞情報を考慮した日本語訳を選択できていた。それほど悪くはないが、これは、品詞情報を持つ辞書データを利用することに変更することで解決できる。

4. 自動生成した問題を用いた実証実験

4. 1 実験内容

認識度調査と同じ大学生（37 名）に対し、3 節で検証した 200 問の多肢選択問題を動画とともに提供し、問題を解かせるという実証実験を実施した。実際に自動生成した問題の効果があるか、実際に運用する上で問題点がないかなどを検証した。学生には動画視聴前後で、同じ多肢選択問題を解かせ、動画視聴前後での正解数の増減を調べた。動画視聴前の解答終了時に正答の提示はしない。

さらに、動画視聴後には、英語教育の専門家によって作られた、動画の内容に関する正誤問題 10 問を解かせ、動画の内容を理解しているかどうかを調べた。

4. 2 実験環境

多肢選択問題の出題、解答には、Moodle を利用した。解答時、学生には辞書やすべての資料等を参考にすることを禁止した。解答のための制限時間は設けず、学生は満足がいくまで解答することができた。

	3.1.2 出題単語候補にはならなかった単語			3.2 正解選択肢の検証			3.3 不正解選択肢 の検証
	総単語数	リストに 含まれる単語	リスト外の単語	品詞が異なる	文脈に 合っていない	品詞×文脈○	不正解選択肢
動画 1	271	225	46	2	6	0	11
動画 2	229	196	33	2	4	1	13
動画 3	234	205	29	4	3	1	9
動画 4	297	247	50	3	4	1	11
動画 5	208	155	53	2	2	0	12
動画 6	218	150	68	4	1	0	8
動画 7	195	166	29	2	7	0	5
動画 8	233	180	53	3	0	2	9
合計	1885	1524	361	22	27	5	78

図 8 自動生成した問題の検証

動画の視聴にはブラウザ上の Flash Player を利用した。視聴時間は、最長で 20 分とし、その時間内であれば、何度繰り返しても、一時停止、巻き戻しをして見てもよいとする。さらに、電子辞書や Web 上の検索サイト等を参考にすることに制限は設けず、学生が動画を用いて自習する場合と同様の環境を提供した。

また、Moodle に不慣れな学生に対しては、必要に応じて操作補助を行った。

4. 3 実験結果

正解数の動画視聴前後の平均はそれぞれ 150.86 個、172.05 個であり、t 検定により有意な差が認められた ($p < 0.05$) (図 9)。さらに、学生の TOEFL スコアと視聴前後の正解数との相

関係数はそれぞれ 0.584、0.509 であり、共に高い相関があった (図 10)。また、TOEFL スコアと視聴前後での正解数の伸びとの相関係数は -0.500 と、高い負の相関があり、英語能力が低い学生の方が高い学生よりも点数の伸びが高く、効果的であったといえる (図 11)。

また、内容理解問題の正解率は 80%以上であり、ほとんどの学生が動画を理解できていた。

5. まとめと考察

本論文で提案した手法を用いて自動生成した問題の品質は、検証により概ね妥当なものであると言える。今後、問題の品質を保証するために、英語教育の専門家にも検証してもらい、それを元に評価する。

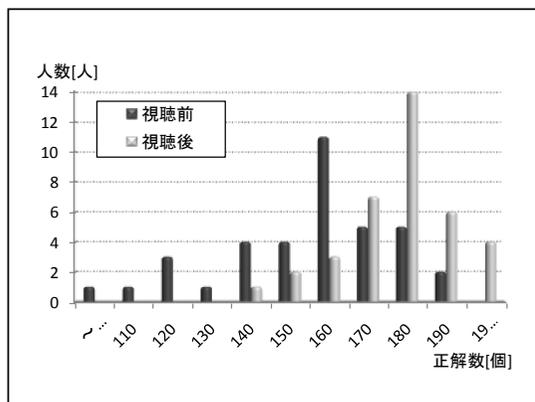


図 9 視聴前後の正解数

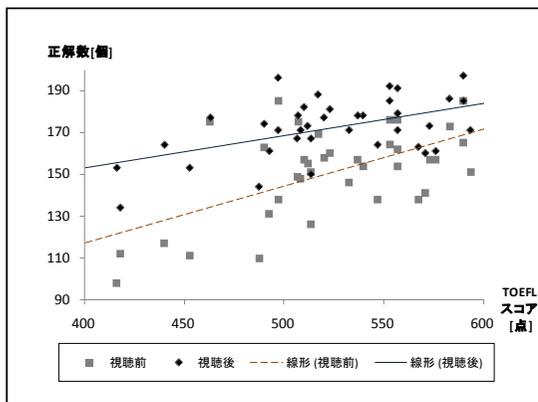


図 10 TOEFL スコアと視聴前後の正解数

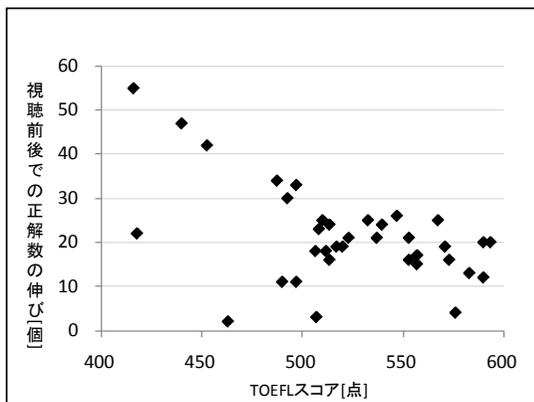


図 11 TOEFL スコアと正解数の伸びの関係

また、自動生成した問題に対し、動画視聴前後で正解数の伸びを観察することができた。伸びの要因としては、同じ問題を2回解いたこと、問題と関連する動画を視聴したことなど、複数の要因が考えられる。今後、対照実験を行って明らかにしていきたい。さらに、学生の TOEFL スコアによって効果の違いが見られたことから、この自動生成手法が、学習者に対してパーソナルな問題を提供できる可能性を示している。

COOLL に登録されているコンテンツの中で、スクリプトが用意されているものすべてに、本論文で提案した手法を用いて問題を自動生成すると、220 の学習教材を提供できるようになる。これは、現在、提供されている学習教材の数の約 2.4 倍であり、学習教材の増加に貢献できる。

謝辞

この研究は、平成 20 年度 文部科学省「質の高い大学教育推進プログラム」採択、津田塾大学「専門課程における英語カリキュラム協調開発」により実施している。

参考文献

- 1) Brinton, D.M., Snow, M.A., & Wesche, M.B., Content-Based Second Language Instruction, 2003
- 2) 生田和重, David Brooks, 動画を活用した英語学習教材と学習事例, 日本教育工学雑誌, vol.26(4), pp.371-376, 2003
- 3) 柏原昭博, 久米井邦貴, 梅野浩司, 豊田順一, プログラム空欄補充問題の作成とその評価, 人工知能学会論文誌, vol.16, No.4, pp.384-391, 2001
- 4) 長谷川信, 磯本征雄, 久村優, 教科書に基づく穴埋め問題の活用と効果, 電子情報通信学会技術研究報告.ET, 教育工学, vol.109(193), pp.13-18, 2009
- 5) 神谷健一, 田中省作, 北尾謙治, データベース・ソフトウェアを用いた英語学習教材の自動作成, 自然言語処理, vol.16(2), pp.45-58, 2009
- 6) 隅田英一郎, 菅谷史昭, 山元誠一, 英語能力測定のための空所補充問題の自動生成手法, 電子情報通信学会技術研究報告.TL, 思考と言語, vol.104(503), pp.17-22, 2004
- 7) 大学英語教育学会基本語改訂委員会, 大学英語教育学会基本語リスト JACET List of 8000 Basic Words, 2003
- 8) <http://www.ets.org>
- 9) <http://edition.cnn.com/studentnews/>
- 10) <http://www.howcast.com/>