

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 12 日現在

機関番号：14501

研究種目：基盤研究(C)

研究期間：2011～2014

課題番号：23501108

研究課題名(和文) 学習履歴データを活用した外国語リスニング学習支援システムの開発及び実践的評価

研究課題名(英文) Development and evaluation of learning support system based on the analysis of historical data

研究代表者

康 敏 (Kang, Min)

神戸大学・国際文化学研究科・教授

研究者番号：60290425

交付決定額(研究期間全体)：(直接経費) 3,100,000円

研究成果の概要(和文)：本研究は、英単語のリスニング学習支援を行うCALLシステムの構築について研究し、システムの開発、実装及び有効性の検証を行ったものである。システムは、個々の学習者または学習グループの英単語学習履歴データから聞き取れないまたは間違いやすい音声を検出したうえ、検出した音声をターゲットとした訓練教材を自動生成する。これによって、授業時間外でも学習者が自身の音声聞き取りの弱점에気づき、その弱点を克服するための訓練を行うことができ、自律学習を促すことは可能となる。システムを実際に学習者に利用させた結果、システム利用の有効性を確認すると同時に、実践的運用においては必要な改善点も明らかにした。

研究成果の概要(英文)：The purpose of this research was to study how to develop learning systems to support English vocabulary listening. A language learning support system has been implemented and evaluated. The system can detect errors of phoneme perception by analyzing personalized historical data in word recognition and generate training materials concerned the errors. The system can help learners recognize their weak points easily in phoneme perception and take personalized training to overcome the weak points by themselves outside the classroom. Although there are some points that should be improved, the learning effectiveness by the system has also been clarified in evaluation experiments.

研究分野：科学教育・教育学・教育工学

キーワード：教授学習支援システム 学習履歴データの分析 データマニング 統計的アプローチ 誤りの検出 外国語学習支援 リスニング学習 自律学習

1. 研究開始当初の背景

外国語学習においては、教育の現場で「読めるが聞けない」ことが問題視され、音声言語によるコミュニケーション能力の育成が語学教育の重要な課題の一つとされている。その原因の一つは、授業時間外に母語のような音声インプット環境の欠如と考えられ、いつでもどこでも Web 経由で利用できる CALL システムを用いた学習スタイルが広がりつつある。CALL システムを用いた学習は時間と場所を問わない利点があるだけでなく、開発済みの教材を利用する以外に、学習者の進捗状況に応じて教材を作成することができ、自律学習を向上させる効果も期待できる。さらに、個々の学習者の学習履歴データが容易に蓄積できるのも CALL システムによる外国語学習の重要な利点の一つである。学習履歴データは個々の学習者のペースに合わせた学習教材を提供し、自律学習を実現するための判断尺度になるほか、その解析によって、学習プロセスやつまづきやすい箇所などを明らかにして学習者と教育者にフィードバックすることにも用いられている。

しかし、教育の現場においては、CALL 授業がまだ積極的に展開されていると言えない状況にある。開発済み教材を用いる CALL システムの多くは、教材の内容が固定されがちであり、教師側による手軽な変更が難しく、個々の学習者の進捗状況に応じた教材作成には対応できない。また、教師による教材作成に対応できるシステムであっても、音声の録音と編集、画像の録画と編集などの作業が必要であるため、授業の準備時間が長くなる問題が依然として残っている。さらに、教師にはシステムへの理解やコンピュータに関する一定の専門知識などが要求され、CALL の学習促進効果を引き出すには教師が重要な要因の一つと指摘されている。これらのことから多くの CALL システムにはユーザビリティの向上が求められていると考えられる。近年、教材作成に柔軟に対応できる Learning Management System (LMS) は CALL 授業を行うためのツールとして注目され、実践的利用についての報告も多く見受けられる。Moodle のような LMS は汎用性が高く、OSS でもあり、導入しやすいメリットもあるが、その高い汎用性は逆に利用しやすさを損ねて CALL 授業展開の困難さの一因と考えられる。多くの LMS においても、外国語学習に欠かせないリスニング学習教材の作成に音声の録音と編集、画像の録画と編集などの作業が必要である。また、編集した音声・画像ファイルの教材への取り込みには HTML に関して一定の知識が作成者としての教師に求められる。

このように、CALL 授業を展開していくためには、ユーザビリティを向上し、柔軟かつ手軽に教材作成できるシステムの開発が望まれる。また、CALL の学習促進効果を引き出し、自律学習につなげるために、学習履歴

データに対するより高度な分析結果をフィードバックできる機能も必要である。手軽に音声つき教材を作成できる機能をシステムに備えるには、合成音声の利用が考えられる。音声合成技術の急速な進歩に伴って、現段階の合成音声は明瞭さ (Intelligibility) だけでなく、自然さ (Naturalness) においてもかなり自然音声に近づいてきている。市販の電子辞書の多くはすでに音声合成機能が実装されている。しかし、教育への合成音声利用の有効性と実用化についての研究は非常に少ないため、合成音声を利用した外国語学習システムに関する実践的報告がほとんどない。筆者らはこの点に着目し、これまでに英語と中国語を対象にして合成音声の学習者・教師評価を調査し、合成音声の外国語教育への利用について研究を行ってきた。その成果の一部として、学習者の合成音声に対する評価を参照して選択的に利用することを提案している。

一方、学習履歴データの活用に関しては、多くの e-Learning ツールと同様に、既存の CALL システムには、初歩的な統計解析手法による学習履歴データの集計等の機能が実装されているが、学習内容に踏み込んだ解析などより高度な解析までまだ至っていない。近年、情報科学分野で盛んに使用されているデータマイニング手法を e-Learning に応用し、種々のツールに取り入れる研究が数多く見られる。学習プロセスの詳細な再現、異常学習プロセスの抽出、また外国語学習の分野については誤答に関するフィードバックなどが学習履歴データのデータマイニングによって実現されている。しかし、CALL におけるより高度な学習履歴データ分析の研究報告はまだ極めて少なく、特に、学習内容に踏み込んだデータ解析に関するシステムの開発研究はまだ見当たらない。

2. 研究の目的

本研究の目的は、学習履歴データを活用し、ユーザビリティを向上した外国語学習支援システムの開発と実践的評価を行うことである。ユーザビリティの向上は、利用目的をリスニング学習に限定し、システムに備える機能を特定することによって実現する。リスニング学習に特有の教材テンプレートを開発し、合成音声を取り入れることによって教材作成機能の使い勝手の向上を図る。また、データマイニング法などを利用してより高度な学習履歴データの分析機能をシステムに実装することも目標としている。リスニングの際の音声知覚についての誤認などを学習履歴データから分析することによって、個々の学習者またはグループの音声聞き取りにおける特徴を抽出して学習者と教師にフィードバックし、トレーニングによる改善を図り、自律学習を可能にすることを旨とする。

3. 研究の方法

システムは主に教材作成・利用部分と学習履歴データ分析部分からなる。

リスニング学習用のテンプレートをデザインし、音声合成エンジンを実装した教材作成システムの開発から始める。学習用テンプレートは主に練習問題やテスト問題を作成するテンプレートから構成する。教材作成インターフェースでは、教材に必要な音声を事前に用意した音声ファイルのアップロードと音声合成エンジンによる音声ファイル作成という二つの機能を備える。また、単語学習を中心にしてテンプレート、教材作成及び教材利用という三つの視点からユーザビリティテストを行い、システムの使い勝手を評価する。

上述のシステムを用いて、学習履歴データを収集し、学習内容に踏み込んだ高度な学習履歴データ解析手法を開発して、学習履歴データの分析機能を実装していく。学習者は、音声聞き取りの際に、どの音声を聞き取れないか、どの音声を聞き間違いやすいか、どの音声は単語認知の妨げになるウィークポイントであるかなどのことを学習した内容から割り出し、それらのウィークポイントに合わせたトレーニングによって改善を図る。最後には、システムを完成して実践的評価を行い、システム利用の有効性とシステムの改善すべき点を明らかにする。

4. 研究成果

研究成果はシステムの使い勝手の向上に関する部分と高度な学習履歴データの分析機能に関する部分からなる。

(1) 使い勝手の向上を図るためには、まず英語のリスニング学習用のテンプレートをデザインし、音声合成エンジンを実装した教材作成システムの開発を行った。また、学習者へのフィードバックに焦点を当てた履歴データ分析機能を開発し、実験的利用によって利用者によるユーザビリティ評価を実施した。

リスニング学習用テンプレートは単語学習に焦点を当てた。実装した単語ディクテーションテンプレートは単語の意味理解に着目したトレーニングに対応し、二者択一問題作成テンプレートは、音声への知覚のトレーニングを行う際、ミニマルペアを使った発音の聞き分けの練習問題作成に対応することにした。教材作成インターフェースには、教材に必要な音声を事前に用意した音声ファイルのアップロードと音声合成エンジンによる音声ファイル作成と二つの機能が備わっている。合成音声を利用する場合、これまでの研究で得た学習者による合成音声評価データを用いて推測した学習者評価結果を教師に提示して教材に取り入れるか否かの判断材料を示すことができる。

フィードバックに焦点を当てた履歴データ分析機能としては、クラス毎または学習者毎の回答状況の提示機能以外に、回答状況

の並べ替え表示機能、回答状況の音素別表示機能及び他者の誤答を含んだ間違いやすい誤答パターン一覧機能を開発している。実施したユーザビリティテストは教材利用後のフィードバックに関する視点から行ったものであり、学習者からポジティブな評価を得たほか、これらのフィードバック機能により、確実に学習者に刺激を与えたことも確認できている。

(2) 学習内容に踏み込んだ学習履歴データの分析に関しては、リスニング理解に支障を来す可能性のある音素への誤認の検出に焦点を当て、シーケンシャルデータマイニング法 (GSP 法) とベイズの定理による解析手法を提案した。

着目した履歴データは単語ディクテーション問題への回答である。

単語認知はリスニング理解の基本的なステップであり、教育現場では、単語のディクテーションがしばしば授業に取り入られている。CALL 授業の場合、容易に蓄積できる履歴データの一つである。これまでの多くの音声知覚モデルによると、単語音声の知覚は単語を構成する各音素の知覚と密接な関係がある。解析手法の提案はこれに基づくものである。単語聞き取りの際の誤認から、音素のような下位レベルの音韻的要素に対する誤認を検出することによって、音素への誤認をウェークポイントとして繰り返し学習し、単語に対する認知能力を高めることにつながる。

① GSP 法による誤認パターン検出手法

この手法の流れとしては、まず個々の学習者または学習グループの誤答したすべての単語を履歴データから抽出し、音素に分解する。次に正解単語を構成する音素と比較し、正答音素と誤答音素をペアにした誤認パターンを抽出する。これはペアリングのステップである。それから、誤認パターンをシーケンシャルデータに生成し、シーケンシャルデータマイニングでよく用いる GSP 法を使って誤認パターンの特徴を検出する。

正答音素と誤答音素のペアリングにおいては、音素の一対一の位置関係から検出を行う逐次比較検出手法、英語のリスニングにおいて一般的に指摘されている日本人学習者が区別しにくい英語子音ペアを取り入れた枠付き検出手法及びこの二つの手法の組み合わせという三種類のアルゴリズムを提案した。手法の評価では、英語オーラルの授業を受講する二つのクラスで単語ディクテーションテストを実施し、収集した履歴データに対して提案手法を用いて分析した結果、それぞれのクラスの音素における誤認パターンの特徴を検出したことが確認できた。また個々の学習者による評価も実施した。さらに、正答音素と誤答音素のペアリングにおける三種類のアルゴリズムの精度について検討を行い、組み合わせの手法が最もよい精度を

もたらずことも明らかにした。

GSP 法による誤認パターン検出の手法を利用して、学習者はどの発音を聞き取れず、ほかのどの発音に間違っただけでなく、少ない履歴データからの検出も可能ということも特徴である。一方、一回のみの単語ディクテーションテストから検出した誤認パターンは学習者の聞き取りにおける本当の弱点かどうか、この手法では判断しにくい。そこで提案したのは、ベイズの定理による解析手法であった。

② ベイズの定理による解析手法

ベイズの定理の応用として、単語認知における聞き取れない音素の推測に関しては、二つの標本空間は $\Omega_1 = \{A_1, A_2\}$ 、と $\Omega_2 = \{B_1, B_2, \dots, B_j\}$ になる。 A_1 と A_2 はそれぞれ「単語が認知できる」と「単語が認知できない」という二つの事象と定義し、 Ω_2 内の任意の事象 B_j はある音素 Ph_j を含む単語とすると、ベイズの公式にあった事後確率は事象 A_2 において次のように計算することができる。

$$P(A_2|B = Ph_j) = \frac{P(B = Ph_j|A_2)P(A_2)}{P(B = Ph_j|A_1)P(A_1) + P(B = Ph_j|A_2)P(A_2)}$$

各事象の定義からは、各確率は以下の意味を持つようになる。

- 事前確率 $P(A_1)$: 単語が認知できる確率である。
- 事前確率 $P(A_2)$: 単語が認知できない確率である。
- 尤度 $P(B=Ph_j|A_1)$: 認知できる単語の中で音素 Ph_j を含む単語の割合である。
- 尤度 $P(B=Ph_j|A_2)$: 認知できない単語の中で音素 Ph_j を含む単語の割合である。
- 事後確率 $P(A_2|B=Ph_j)$: 音素 Ph_j を含む単語なら聞き取れない確率である。

従って、事後確率を計算することによって、ある音素を含む単語を認知する確率が低ければ、その音素は単語認知に妨げる音声知覚のウィークポイントである可能性高いことが考えられる。

手法の評価では、JACET 語彙リストの英単語を二つのレベルに分け、システム経由で単語ディクテーションテストを実施した。テスト毎の履歴データとテストの累積履歴データ別にベイズの定理による推測結果を比較した結果、テスト毎の結果が激しく変動することに対して、2、3回のテストの累積履歴データからは安定した推測結果が得られることが判明した。即ち、英単語のリスニング学習においては、40個~60個の学習履歴データがあれば、ベイズの定理によって正しく学習者の音素知覚のウィークポイントを見つけ出すことは可能である。

ベイズの定理によって累積学習履歴デー

タから正しく音声知覚のウィークポイントを見つけ出した上に GSP 法によって誤認パターンを検出するという組み合わせの手法はそれぞれの手法を単独に用いるより、学習者の実際の弱点を見つけ出すことに有効であるとの結論に至っている。

③ 訓練教材の自動作成

自律学習を可能にするシステムの開発を目指して、システムにおいては、音声聞き取りのウィークポイントの検出機能の開発だけでなく、聞き取りのスキルを向上するための教材自動生成機能の開発も試みた。

教材自動生成機能によって、2種類のテンプレートに対応した訓練教材が生成される。学習者の誤認パターンに基づいて、予め用意したミニマルペアの単語一覧からランダムに誤認の音素を含むペアを複数抽出し、音声合成エンジンによって単語の音声を生成し、二者択一問題形式の教材を作成している。聞き取れない音素についてのトレーニングのあと、単語認知の向上につながるかどうか、学習者が確認できるように、誤認の音素を含む単語を中心にして単語ディクテーション問題形式の教材を自動生成している。自動生成に当たっては、単語の選択に工夫を施していた。評価実験では、被験者に対するインタビューの結果に照らして、被験者によってシステムを利用した効果が見られたケースと見られないケースという分かれた結果になったが、同時にシステムの改善すべき点も明らかにした。

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 9 件)

- ① 康 敏、鄒 亜亨、柏木 治美、大月 一弘、鏑木 誠、単語認知における音声知覚の弱点を分析する手法についての評価、日本教育工学研究報告集、査読無、JSET15-1、2015、27-34
- ② Yaheng ZOU、Harumi KASHIWAGI、Yi SUN、Kazuhiro OHTSUKI and Min KANG、Developing and Evaluating a Test Generation Module to Support Personalized Phoneme-based Training、Proceedings of the 22nd International Conference on Computers in Education、査読有、2014、809-811
- ③ Min Kang、Yaheng Zou、Harumi Kashiwagi、Kazuhiro Ohtsuki and Makoto Kaburagi、Analysis of E-Learning Logs to Estimate Students' Phonemic Perception Confusion in English Word Recognition、Advances in Web-Based Learning、Wang J.、Lau R.、Eds、Springer、査読有、2013、320-329
- ④ Yaheng Zou、Harumi Kashiawgi、Kazuhiro Ohtsuki and Min Kang、Development of

Material Automatic Generation System Based on the Analysis of Phonemic Errors in English Vocabulary Listening, Advances in Web-Based Learning, Wang J., Lau R., Eds, Springer, 査読有、2013、340-349

- ⑤ Yaheng ZOU, Kashiwagi HARUMI, Kazuhiro OHTSUKI, Min KANG, Developing a Learning Support System to Detect Phonemic Errors in Data of Listening Learning Log, Proceedings of the International Conference on e-Commerce, e-Administration, e-Society, e-Education and e-Technology, 査読有、2013、1145-1154
- ⑥ 蔣 傑翔, 鄒 亜亨, 柏木 治美, 康 敏, 英単語学習支援システムにおける記憶メカニズムの活用についての検討、日本教育工学研究報告集、査読無、2013、JSET13-1、117-124
- ⑦ 鄒 亜亨, 柏木 治美, 大月 一弘, 康 敏, リスニング学習履歴データに基づく誤りパターンの検出システム、日本教育工学会論文誌、査読有、36 卷(Suppl.)、2012、49-52
- ⑧ 鄒 亜亨, 吉田 均, 柏木 治美, 大月 一弘, 康 敏, フィードバック機能を重視したリスニング学習支援システムの開発、日本教育工学研究報告集、査読無、JSET12-1、2012、179-186
- ⑨ Min Kang, Yaheng Zou, Harumi Kashiwagi, Kazuhiro Ohtsuki and Makoto Kaburagi, An analysis of learners' response to diagnostic rhyme test, Proceedings of Annual international conference on education & E-Learning, 査読有、2011、110-114

[学会発表] (計 7 件)

- ① 康 敏, 鄒 亜亨, 柏木 治美, 大月 一弘, 鏑木 誠, 音声知覚弱点の推測についての検討、日本教育工学会第 29 回全国大会、2013. 9. 21、秋田大学 (秋田県)
- ② 鄒 亜亨, 柏木 治美, 大月 一弘, 康 敏, 英単語学習におけるリスニング教材自動生成システムの試作、日本教育工学会第 29 回全国大会、2013. 9. 21、秋田大学 (秋田県)
- ③ 康 敏, 鄒 亜亨, 柏木 治美, 大月 一弘, 鏑木 誠, DRT 単語を利用した合成音声リスニングテストの結果、日本教育工学会第 28 回全国大会、2012. 9. 15、長崎大学 (長崎県)
- ④ 鄒 亜亨, 柏木 治美, 大月 一弘, 康 敏, 英単語学習における誤りパターンの検出、日本教育工学会第 28 回全国大会、2012. 9. 15、長崎大学 (長崎県)
- ⑤ 蔣 傑翔, 鄒 亜亨, 柏木 治美, 康 敏, 記憶の特徴を考慮した英単語学習支援システム開発の試み、日本教育工学会第 28

回全国大会、2012. 9. 15、長崎大学 (長崎県)

- ⑥ 康 敏, 鄒 亜亨, 柏木 治美, 大月 一弘, 鏑木 誠, 英単語学習における聞き取り困難な音声要素の自動検出の試み、日本教育工学会第 27 回全国大会、2011. 9. 19、首都大学 (東京都)
- ⑦ 鄒 亜亨, 柏木 治美, 康 敏, リスニング学習支援システムにおけるフィードバック機能、日本教育工学会第 27 回全国大会、2011. 9. 18、首都大学 (東京都)

6. 研究組織

(1) 研究代表者

康 敏 (KANG, Min)

神戸大学・大学院国際文化学研究所・教授
研究者番号：60290425

(2) 研究分担者

柏木 治美 (KASHIWAGI Harumi)

神戸大学・国際コミュニケーションセンター・教授

研究者番号：60343349

大月 一弘 (OHTSUKI, Kazuhiro)

神戸大学・大学院国際文化学研究所・教授
研究者番号：10185324

(3) 連携研究者

鏑木 誠 (KABURAGI, Makoto)

放送大学・兵庫学習センター・客員教授
研究者番号：40093504