

機関番号：11301

研究種目：基盤研究(B)

研究期間：2008～2010

課題番号：20320075

研究課題名(和文) 日本語母語話者の英語学習のための韻律評価および文法誤り検出の高度化

研究課題名(英文) Automatic prosody evaluation and grammatical mistake detection for English learning by Japanese native speakers

研究代表者 伊藤 彰則 (ITO AKINORI)

東北大学・大学院工学研究科・教授

研究者番号：70232428

## 研究成果の概要(和文)：

コンピュータによる英語学習(CALL)システムにおいて「話す」練習を可能にするために、学習者の英語発話のイントネーション・リズムを自動評価する方法、および学習者発話に含まれる文法誤りを自動的に指摘する方法を開発した。イントネーション・リズム評価においては、決定木により単語の重要度を推定する手法を開発し、ネイティブ評定者間の評定値相関に近い相関を持つスコアが計算できた。文法誤り検出では、自動生成文から n-gram を学習する手法を開発し、単語正解精度 89.2% が得られた。

## 研究成果の概要(英文)：

I have developed two methods that enable speaking exercise in a computer-assisted language learning (CALL) system: a method for evaluating prosody of an English utterance made by a learner, and a method for detecting grammatical mistakes included in the learner's utterance. As for prosody evaluation, I developed an estimation method of word importance factors using a decision tree, and obtained a high correlation to human assessment score, which is comparable to correlation between scores given by human evaluators. As for the grammatical mistake detection, I proposed a method for training an n-gram language model from artificially generated sentences with mistakes, and obtained 89.2% word accuracy.

## 交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2008年度	1,800,000	540,000	2,340,000
2009年度	1,600,000	480,000	2,080,000
2010年度	1,400,000	420,000	1,820,000
年度			
年度			
総計	4,800,000	1,440,000	6,240,000

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：e-ラーニング・コンピュータ支援教育、英語発音教育

## 1. 研究開始当初の背景

社会の国際化に伴い、外国語習得の重要性はますます高まっている。コンピュータを利用した外国語学習(Computer Assisted Language Learning: CALL)システムは手軽

で安価・かつ効率的な学習を可能にするため、現在盛んに利用されている。しかし、現在のCALLシステムは聞き取り・読み取り能力の開発が中心であり、作文については「書く」ための学習が主であった。そのため、オーラ

ルコミュニケーションに重要な発音・発話の能力開発支援は十分とは言えなかった。これは、学習者による第2言語(L2)音声における発音の良さや文の正しさをコンピュータで測ることが難しいためである。この問題を解決するため、本提案の研究代表者は、科学研究費基盤研究(B)「話者適応と文法誤りモデリングを用いた外国語教育システムの開発」(平成16～19年度)において、コンピュータでのL2発話の良さを高精度に測るための技術開発を行ってきた。この研究の成果として、L2発音の分節的な良さを計測する方法の開発(日本語話者の英語および韓国語話者の日本語)、韓国語話者の日本語発話に対する文法誤りの指摘と訂正手法の開発を行ってきた。しかし、特に日本語話者の英語学習に対して重要である韻律(ストレス、アクセント、イントネーション)の学習については、まだ十分な技術開発がなされていない。韻律評価についてはいくつかの研究があり、研究代表者らも一部手がけているが、まだ十分な性能を得るには至っていない。現在、研究代表者らは韻律評価に関する検討をさらに進めようとしているが、母語話者による韻律の評価は発音の評価よりもばらつきが大きく、コンピュータで扱うことが難しい課題である。また、日本人学習者が発話した英語からの文法誤り検出については現在検討しているところであるが、日本語-英語間には韓国語-日本語間よりもはるかに大きな文法的ギャップがあり、これまで開発したものと同様の手法によって十分な性能が得られるかどうかまだわからない状況である。これまで開発した韓国語-日本語間の文法誤り検出の枠内では、冠詞の間違い、動詞の活用、前置詞の間違い等には対処できると思われるが、文の構造自体が違う場合、使うべき動詞の間違い、単語の選択の間違い等には現在までの方法では不十分であると思われる。

## 2. 研究の目的

本研究の目的は次の2つである。

- 日本語母語話者による英語の自発的発話から、文法的な誤りを検出する手法を確立する。
- 日本語母語話者の英語発音の韻律(イントネーション、リズム)の良さを計測する手法を確立する。

このうち前者に関しては、英語学習教材の中で利用することを考慮し、比較的限定された状況(回答内容が比較的限定された対話、あるいは短い和文を英訳して発声させる課題等)における手法を開発する。このような状況では、話者が発声すべき正しい英語表現と、実際の発話中に含まれる誤りが比較的予測可能である。また、前に述べたような冠詞、前置詞、活用等の誤りに加えて、文の構

造や単語の選択誤り等にも対処できるアルゴリズムの開発を目指す。後者に関しては、当面は発声対象を単文に限り、韻律要素のうちイントネーションとリズムの評価を行う。これまで開発してきた手法と比較して、より人間の評価に近い評価値を自動的に計算するアルゴリズムの開発が目標である。また、従来の方法では、イントネーション評価を行う文について、学習者の発声だけでなく、複数のネイティブ話者による教師音声が必要である。これは、CALLシステム上での教材作成を考えると非常に大きな制約である。そこで、教師音声を必要としない韻律評価手法を確立する。このため、音声合成のためのイントネーションモデル、あるいは合成音声そのものを教師音声の代わりに用いて、ネイティブ話者による教師音声と同等の評価性能を得る方法を開発する。

## 3. 研究の方法

### (1) 文法誤り検出手法の開発

文法誤り検出に関しては、まず日本人が発声した英語の誤りの分析を行う。基礎データとしてNICT JLEコーパスを用い、日本人が犯す文法的な誤りのうち、比較的単純な処理で対処が可能である「冠詞の誤り」「動詞の変化の誤り」「前置詞の誤り」の3つのクラスでどの程度の誤りがカバーされるのかを調査する。JLEコーパスには、文法誤りとそれに対応する正解がタグとして付与されているため、その情報に基づいて誤りを調査する。また、1単語の置換・挿入・脱落だけでなく、複数の単語による誤りも調査する。

この分析に加えて、日本人が発話した(誤りを含む)英語のデータ収集を行う。文法誤りモデルの学習にはJLEコーパスを用いるが、実際にCALLシステムを使う状況においてJLEコーパスと同じような誤り傾向が得られるかどうかは検証を要する。そこで、和文英訳課題を使って日本語母語話者による英語発話を収集し、A)で作成したモデルがどの程度のカバー率を持つのか検証する。

次に、この誤り傾向を用いて、発話からの文法誤りを検出する手法を開発する。発話からの文法誤りを検出するには、まず文法誤りを含む発話を高精度に認識しなければならない。従来手法は、誤りルールによって文法誤りを含む学習者発話を予測し、それをネットワーク文法によって表現する。そして、そのネットワーク文法を用いて発話を認識する。この方法では、あらかじめ予測できた誤りにしか対応できず、実際の学習者発話に十分対応することができない。そこで提案法では、学習者が発話すべき正解となる文(正解文)に対して誤りルールを適用して文法誤りを含む文を多数生成し、それらの文からN-gram言語モデルを学習して、それを用い

て学習者発話を認識する。これを図1に示す。

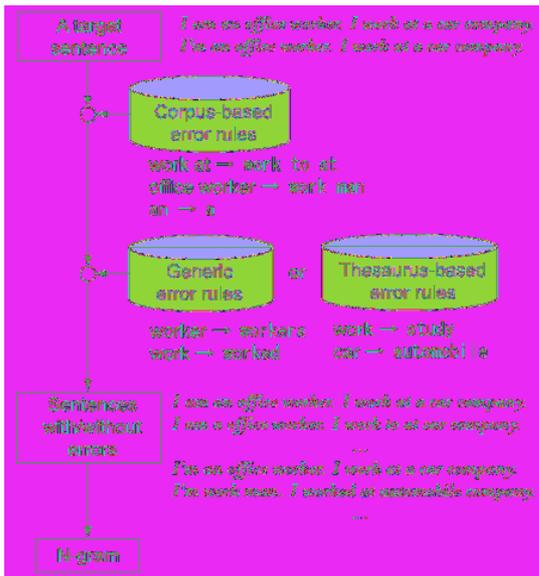


図1 文生成と N-gram を用いた学習者発話認識

これと平行して、日本人が発声した英語を高精度に認識するための音響モデルの開発を行う。一般に、日本人の英語を英語母語話者データから学習した音響モデルで認識すると、高い精度が得られない。今回の研究では、日本人が発声した英語音声を、本人が意図したとおりの単語列として認識することが求められるので、日本人英語に対応した音響モデルが必要になる。従来法としては、英語の音響モデルと日本語の音響モデルを並列に接続する方法や、話者適応手法を用いて非母語話者音声に対応した音響モデルを作成する手法がある。これに対し、本研究では日本語母語話者の英語発声の認識に特化しており、また発話者はさまざまな英語習熟度が想定される。そのため、日本語と英語の音響モデルのどちらでもない発音が含まれている可能性が高く、日本語と英語音響モデルを併用する方法は適切ではない。一方、日本人の発声した英語についてはある程度の量のデータが入手可能であるため、必ずしも話者適応手法を使わなくても音響モデルの学習が可能であると考えられる。このことから、ERJ データベースを学習データとして、さまざまな HMM 音響モデルの学習条件に対する性能評価を行い、日本人英語に対して高い認識性能を持つ音響モデルを作成する。

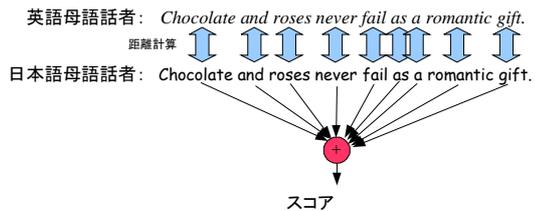
(2) 韻律評価手法の開発

韻律評価に関しては、対象として ERJ データベースを用い、従来法による評価値と英語母語話者による評定との相関を見る。次に、単語による重要度の違いにターゲットを絞

り、単語の重要度を考慮することによってシステムの評価スコアと評定者のスコアとの相関が改善できるかどうかを試す。従来法における評価では、日本語母語話者の発話における基本周波数の一次差分と、英語母語話者の同様なパラメータとのフレームごとの距離を積算することによって距離スコアを算出しており、「どの単語の韻律がどの程度良いのか」は評価していない。これに対し、評定者によるイントネーション評定においては、単語の種類や文法的役割などによって評定に与える影響が異なることが指摘されている。そこで、単語に「重要度」を設定し、単語ごとに重要度による重みをかけることでシステム評定と英語母語話者の評定との相関を高めることを目指す。これを図2に示す。そのためまず、日本語母語話者が発声する英語文中の単語を、品詞・音素数・文中の位置などの要因によってクラスに分け、それぞれのクラスについて単語に対するスコアと評定者のスコアを比較する。この作業により、どのクラスがイントネーション評定において重要なかを調べることができ、重要単語の選定についての知見を得ることができると考えられる。

さらに、発話文中の単語を自動的に最適なクラスに分け、同時にそれぞれのクラスに対して最適な重みを計算する方法を開発する。これには決定木を用い、前述の品詞や単語位置などの特徴から、最終的に計算されたスコアとイントネーションの評定値との相関が最大になるように単語クラスおよびその重みを決定する。

従来法



提案法

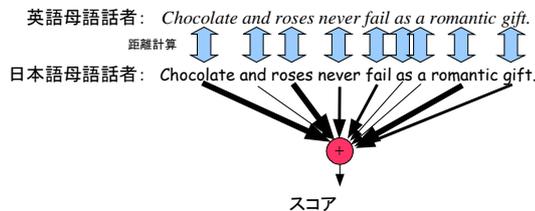


図2 単語重要度による評定スコアへの重み付けの概念図

4. 研究成果

2008年度は、合成音声を用いた日本人英語音声のイントネーション評価を中心に研究を行った。

まず、イントネーション評価のための教師音声として合成音声を用いることの有効性を検証した。英語母語話者音声を教師音声とした場合と、合成音声を教師音声とした場合について、システムによる評価がどの程度人間による評価と整合するかを相関関数により評価した。その結果、ユークリッド距離を用いた評定値計算を用いた場合には、合成音による評価の信頼性は母語話者音声をを用いる場合に比べて信頼性が低下した。しかし、距離尺度としてマハラノビス距離を用い、また評価の重要度によって距離を重みづけすることにより、合成音声を用いた評定でも母語話者音声と同等の信頼性が得られることがわかった。

次に、評定値の計算方法について検討した。従来法では、学習者の音声のイントネーションパターンと最も近い教師音声のイントネーションパターンに対する距離のみによって評定値を計算していた。これに対し、提案法では、学習者と全ての教師音声との距離を用い、これらの距離の線形結合によって評定値を計算する。線形結合の重みの最適化には重回帰分析を用いた。実験の結果、合成音を教師音声とした場合、人間による評価との相関は0.203から0.288に増加した。さらに、イントネーション評価システムに対し、イントネーションのための特徴量だけでなく、リズム評価のための特徴量を導入することを検討した。その結果、人間による評価との相関は0.36まで上昇した。

2009年度は、日本人英語イントネーション評価の高精度化と、文法誤り検出に関する基礎検討を行った。

日本人英語イントネーション評価を行うためには、学習者による英語文の発話と、複数の英語母語話者（教師話者）による同じ文の発話のイントネーションを比較する必要がある。この際に、イントネーションを直接表す特徴量と、文のリズムを表す特徴量をそれぞれ計算するが、従来はひとつの学習者発話についてそれらのスコアを独立に計算してから重み付きで足し合わせることで最終的なスコアを計算していた。これに対し本研究では、2つのスコアを教師話者ごとに最適に組み合わせ、さらにそれを重回帰分析によって統合するという方法を開発した。提案法により、人間による評価平均値とシステムの評価値の相関は最大0.54となった。英語母語話者による評価値の相関が最大0.55であることを考えると、提案法によって英語母語話者による評価と同等のイントネーション評価が実現できたと言える。

次に、日本人が話した英文に含まれる文法

誤りを自動検出する手法の改良について検討した。この方法は、以前我々が開発した方法と同じく、文法誤りルールによる誤り文生成と、生成した誤り文からのn-gram学習に基づく認識方法である。以前の研究では、一つの単語に対して文脈に関係なく誤りルールを適用していた。これに対して、我々は単語の置かれた文脈まで考慮した誤りルールを作成し、実際に適用した。また、以前の研究ではきちんと取り扱われていなかった短縮形(I' m等)をルールに組み込み、短縮形と非短縮形どちらも同様に生成できるようにルールを構成した。誤りルールを改良した結果、単語認識率は88.3%から89.2%に改善した。

2010年度は、日本人による英語発話からの文法誤り検出の高精度化について検討した。まず、日本人英語音声の認識を高精度化するため、発話者レベル別の音響モデル学習について検討した。音響モデル学習に用いる学習用音声(ERJデータベース中の95名分、約21万単語)を発話者の英語レベル別に均等に3つに分け、それぞれについてモノフォーン音響モデルを学習する。HMMの混合数を512混合まで増加させたところ、少ない混合数では分割学習の効果が高かったが、256混合以上では分割せずに学習した場合と同等以下の性能であった。これは、パラメータ数に対する学習データ数が少ないためであると考えられた。そこで、学習データの減少を補うため、MLLR話者適応を繰り返し適用する。この方法により、混合数が多い場合でも分割学習の効果により認識性能を改善することができた。

さらに、誤り指摘の精度を向上させるための基礎的な方法として、学習者の発話が想定される正解文と一致するのかどうかを識別する方法を開発した。従来は学習者発話を一旦認識し、その内容と正解文を比較して誤りを指摘していたが、提案法ではまず発話が誤りを含むのかどうかをいったん識別し、誤りがあると識別された場合のみ音声認識を行って誤りの指摘を行う。識別には、音声認識の尤度と正解文の強制アラインメントによる尤度の差を利用する。提案法により、単語正解精度は1.0ポイント上昇し、また誤り指摘の性能を示すF値は0.032(パーセント表示で3.2ポイント)改善した。

## 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7件)

1. Takuya Anzai, Seongjun Hahm, Akinori Ito, Masashi Ito and Shozo Makino,

- “Grammatical error detection from English utterances spoken by Japanese,” Proc. 2nd Asian-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA-ASC 2010), CD-ROM, 2010, 査読有
2. Yuichi Ohkawa, Motoyuki Suzuki, Hirokazu Ogasawara, Akinori Ito, Shozo Makino, “A speaker adaptation method for non-native speech using learners' native utterances for computer-assisted language learning systems,” Speech Communication, vol. 51, pp. 875-881, 2010, 査読有
  3. Akinori Ito, Tomoaki Konno, Masashi Ito, Shozo Makino, Motoyuki Suzuki, “Intonation Evaluation of English Utterances Using Synthesized Speech for Computer-Assisted Language Learning,” International Journal of Innovative Computing, Information and Control, vol. 6, pp. 1501-1514, 2010, 査読有
  4. Akinori Ito, Tomoaki Konno, Masashi Ito and Shozo Makino, “Evaluation of English Intonation based on Combination of Multiple Evaluation Scores,” Proc. Of Interspeech, pp. 596-599, 2009, 査読有
  5. Motoyuki Suzuki, Tatsuki Konno, Akinori Ito and Shozo Makino, “Automatic Evaluation System of English Prosody Based on Word Importance Factor,” Journal of Systemics, Cybernetics and Informatics, vol. 6, pp. 83-90, 2008, 査読有
  6. Tomoaki Konno, Masashi Ito, Motoyuki Suzuki, Akinori Ito, Shozo Makino, “Intonation Evaluation of English Utterances using Synthesized Speech for Computer-Assisted Language Learning,” Proceedings of International Conference on Natural Language Processing and Knowledge Engineering, pp. 202-208, 2008, 査読有
  7. Akinori Ito, Ryohei Tsutsui, Shozo Makino and Motoyuki Suzuki, “Recognition of English Utterances with Grammatical and Lexical Mistakes for Dialogue-based CALL System,” Proc. Interspeech, pp. 2819-2822, 2008, 査読有

[学会発表] (計6件)

1. 安齋 拓也, 咸 聖俊, 伊藤 彰則, “日本人英語学習者の発音レベルを考慮した音響モデルに関する検討”, 日本音響学会, 2011年3月9日, 東京
2. 安齋 拓也, 咸 聖俊, 伊藤 彰則, “日本人英語発話からの文法誤り検出”, 情報処理学会音声言語情報処理研究会, 2011年2月5日, 福山

3. 安齋 拓也, 咸 聖俊, 伊藤 彰則, 伊藤 仁, 牧野 正三, “日本人英語発話からの文法誤り検出”, 日本音響学会, 2010年9月14日, 大阪
4. 今野 智明, 伊藤仁, 伊藤彰則, 牧野正三, “合成音声を用いた英語のイントネーション評価における単語重要度適用に関する検討”, 日本音響学会, 2009年3月19日, 東京
5. 今野 智明, 伊藤仁, 伊藤彰則, 牧野正三, “合成音声を用いた英語のイントネーション評価へのリズム特徴量の導入”, 日本音響学会, 2008年9月12日, 福岡
6. Tomoaki Konno, Masashi Ito, Motoyuki Suzuki, Akinori Ito and Shozo Makino, “Evaluation of English Intonation of Japanese Learners Using Synthesized Speech,” 2<sup>nd</sup> International Workshop on Language and Speech Science, 2008年9月4日, 東京

#### 6. 研究組織

##### (1) 研究代表者

伊藤 彰則 (ITO AKINORI)  
 東北大学・大学院工学研究科・教授  
 研究者番号：70232428

##### (2) 研究分担者

なし

##### (3) 連携研究者

鈴木 基之 (SUZUKI MOTUYUKI)  
 徳島大学・大学院ソシオテクノサイエンス研究部・准教授  
 研究者番号：30282615

牧野 正三 (MAKINO SHOZO)  
 東北文化学園大学・科学技術学部・教授  
 研究者番号：00089806

大河 雄一 (OHKAWA YUICHI)  
 東北大学・教育情報学研究部・助教  
 研究者番号：60361177