

研究題目	音声処理技術を用いたインテリジェント発音教育システムの開発	報告書作成者	広瀬啓吉
研究従事者	峯松信明（共同研究者），河合剛（東京大学工学系研究科博士課程，現在、電気工学科助手）		
研究目的	<p>国際化の進展によって、多くの日本人が海外で生活し、また逆に、留学生など多くの外国人が日本で生活するようになった。このため、非母国語言語の教育が重要な課題となっているが、これを従来のような教師の指導によっては、需要に答えることができないことは明らかである。この観点から、既に計算機を利用した言語の自己学習の試みが各所で行われているが、これらは、教師の発音をまねることを基本としたものであり、学習者が、自分の発音の問題点を自身で判断することが求められていた。従って、学習者にとってはどこに着目して学習するかが必ずしも明確ではなく、ややもすれば声帯模写訓練になってしまい、効果的な学習が行えないという欠点があった。ここでは、最近、急速に進歩している音声分析・認識・合成等の音声情報処理技術を応用して学習者の発音の問題点を自動的に抽出し、それに従って、どのように発音を矯正するかを具体的に学習者に示すインテリジェントな発音教育システムを日本語と英語について開発する。これによって、実際の教師なしに効果的な発音教育を行うことが可能となり、今後のさらなる国際化の進展に十分対応可能な体制が構築される。</p> <p>発音教育に対する要請は国によって異なると考えられる。まず、米国においては、移民に対する米語教育が主体であり、発音教育システムに対するニーズはそれほど高くなく、楽しみながら会話能力を養うことを目的とした対話システムなど、発音能力の向上は副次的な目的とした研究が多い。発音に関しては就業希望者の公平な選別のための発音技量判別システムといった研究があるが、音声認識技術の単なる応用分野の開拓といった色彩が強い。これは、米語が世界的に力を持った言語であるため、米国人に他言語を学習しようとする意欲が乏しいことに起因する。次に欧州においては、多言語を有するため事情が大きく異なり、発音学習に対するニーズが大きい。しかしながら、欧米語を対象とするため、既に学習者は相当の発音能力を有しており、高度な発音技量の獲得が期待されている。これは、現在の音声情報処理技術で自動化することは困難である。一方、我が国では、日本人の英語学習、外国人の日本語学習といった高いニーズがあり、しかも、初歩の発音教育からスタートさせる必要がある。この様な場合、初歩の発音教育を確実に行うことが、その後の正しい発音獲得のために不可欠であるが、そのための教師の人材確保は非常に困難な状況である。今回、開発する発音教育システムは、特にこの様な、初歩の学習者に正しい発音を教育することを目的としており、音声認識をはじめとする現在の音声情報処理技術の枠内で有効なシステムを構築することが可能である。</p>		

研究内容

外国語学習者の発音した音声の特徴を教師のそれと比較し、音声認識技術を利用して発音の誤りを自動的に検出し、発音矯正のための具体的なインストラクションを学習者に与えるとともに、発音修正した合成音声を提示する説明書の図 1 のような発音教育システムを計算機上で開発する。具体的に対象とする発音の誤りの主なものは、日本語特殊拍における不適切な拍長制御、日本語発音における英語音素置換と誤アクセント型付与、英語発音における母音の挿入と不適切な強勢、である。これらについて、適切な学習が行いうる対象を選定した上で、発音誤り自動検出、矯正指示提示、修正音声合成等のアルゴリズムを開発して、システムを作成する。さらに、発音教育の実験を行い、効果的な発音教育方策を開発してシステムに実装することを研究目的とする。ここで、発音誤りの検出については、従来のシステムでは、単に学習者音声の音響的特徴と教師のそれとの距離の大小比較、あるいは音声認識器による認識結果を利用していたのに対し、本研究のシステムでは、学習対象言語の母国語話者による知覚実験で発音の許容範囲を設定して行う。また、多くの話者の発音で作成した学習対象言語の音素モデルと学習者言語の音素モデルとを用意し、学習者の発音がどちらに近いかで発音の良し悪しを判定する。また、合成音声の提示については、主に、抑揚の誤り矯正を目的とし、波形編集合成の手法により、発声者の抑揚を矯正した音声を合成するとともに子音についても矯正対象とする。従って、本研究のシステムの主要機能は、隠れマルコフモデル (HMM) と音声認識技術を利用した教師音声と学習者音声の自動セグメンテーションとアラインメント、音声認識による発音の自動評価、基本周波数パターンの抽出と教師音声と学習者音声間での比較、波形編集合成による矯正音声の生成、教師音声と学習者音声の特徴のディスプレイ上での比較表示、等である。教育方策に関しては、①知覚実験による人間による発音評価の定式化と、その音声認識技術による自動化、②学習者の言語の特徴と学習対象言語の特徴との比較とそれを考慮した発音教育手法の開発、③教師音声と学習者音声の韻律的特徴（基本周波数パターン等）の比較と発声誤りの自動検出手法の開発、④韻律的特徴のモデルによる定式化と発声誤り矯正の自動指摘手法の開発、⑤学習者自身の音声による矯正された発音の音声の生成と提示、⑥効果的な教育のための画面表示手法開発等を行う。これらの成果を統合して、最終的にインタラクティブな発音教育システムを開発する。なお、③については、対象学習者の制約から日本語教育を例にとり、欧米人とアジア人の差を研究対象とする。

<p>研究のポイント</p>	<p>計算機援用発音教育システムは、既に国内外でいくつか発表され、市販されているが、教師の発声を繰り返しまねることを基本としたものである。しかしながら、発音習得の目的は、教師の発音と同一の発音をすることではなく、学習対象言語らしい発音を行うことであって、従って、学習者にとって、自分の発声のどこが問題なのか把握しづらく、必ずしも効果があがっていない。また、これらのシステムは一方的に教師音声を提供するもので、インタラクティブな学習という観点から不足していた。音声認識技術を用いて、評価を自動的に行う試みもいくつか報告されているが、単に音声認識率で発声の善し悪しを判断するもので、それが、人間による評価とは対応せず、教育玩具の域を出ないものであった。これに対し、本研究は、母国語話者を対象とした知覚実験によって発音の許容範囲を定量化した上で、学習者の発音の善し悪しを判断するという顕著な特徴を有するとともに、学習者にとって分かりやすい矯正指示を合成音声とともに提示するという従来ないシステムとなっている。また、従来の様に、単に学習対象音声のHMMを利用した音声認識によって学習者の発音誤りを検出するのではなく、外国語を発音する際、学習者は自身の言語の発音によって置き換えを行うという仮定のもとに、学習者の言語についてもHMMを用意し、発音の誤りパターン（発音誤りの書き換え規則）を記述することによって、より精度の高い発音誤り検出を可能としている。さらに、従来、積極的に行われていなかったアクセント型等の韻律の学習も可能としている。</p>
<p>研究結果</p>	<p>日本語特殊拍における不適切な拍長制御に関しては、まず特殊拍と非特殊拍で意味の対立する単語（minimal pair）を選定し、拍長を制御して合成した音声を日本語話者に聴取させ、特殊拍音素と非特殊拍音素（通常音素）の音素長の判断曲線を求めた。次に、音素HMMによる音声認識を利用した強制アライメントにより学習者音声の特殊拍音素と通常音素の長さを自動測定する手法を開発した。これによって求めた音素長を上記の判断曲線と比較し、学習者の発音が日本語話者にどの程度正しく聴取される可能性があるかの割合を示すとともに、矯正の指示を行うシステムを開発した。欧米人、中国人の初級学習者を対象とした発音学習実験を行い、システムの有効性を確認した。日本語発音における英語音素置換に関しては、日本語と英語の音素HMMをそれぞれ用意し、可能な発音誤りを記述した書き換え規則により検出するシステムを構築した。さらに、拍長教育システムと組み合わせ、日本語学習者を対象としたシステムを開発した。図2に学習者とシステムの関係を示す。英語発音における母音の挿入に関しては、同様に書き換え規則によって母音挿入を記述し、図3の様に検出するシステムを構築した。日本語発音における誤アクセント型付与に関しては、まず、アクセント型により意味対立をおこす単語のminimal pairを選定し、合成音声を用いた知覚実験により、アクセント型判断曲線を求めた。次に、音素HMMによる強制アライメントにより音声をモーラ単位で区分し、ピッチ抽出を行ってアクセント型判断曲線と比較し、学習者音声のアクセント型が母国語話者に正しく知覚される割合を求める図4の様なシステムを開発した。英語発音における不適切な強勢に関しては、スペクトル・パワー・ピッチ・音節長をパラメータとする音節HMMにより、強勢音節を自動抽出するシステムを開発し、その際の重みの異なりに着目して、検出する手法を構築した。発音矯正された音声の提示に関しては、学習者の音声をもとにTD-PSOLAにより適切な韻律で修正して合成する手法を開発した。その際、母音挿入、子音発音誤りも修正することも可能とした。以上の成果を統合し、トータルな発音教育システムを開発した。</p>
<p>今後の課題</p>	<p>学習者にとって自身の発音習熟度レベルが可能な発音教育システムを構築した。実際に、発音教育に有効であることが明らかとなったが、より対象者を拡大した発音教育実験によってその有効性を確認する必要がある。また、本研究では、学習対象を主として単語に限定しているが、これは、初級者の発音学習という目的に沿ったものではあるものの、さらに、今後、文に対象を拡大することの必要性があることも事実である。対象を文に拡大した場合、HMMによる強勢アライメント等の要素技術の性能が低下することが予想され、これを解決する研究が今後必要となる。また、文の場合、発音誤りが複雑に重畳して発生することになるので、矯正の指示について検討することが求められる。</p> <p>今回、開発したアクセント型の学習システムでは韻律のモデルを特に仮定していないが、現在、基本周波数パターン生成過程モデルを仮定したより精密な韻律制御が可能なシステムを開発中である。現在のシステムでは、declination現象が対象外となっており、対象を文へ拡大する際に問題が発生する可能性があったが、生成過程モデルを仮定することによりこれが解決される。</p>

図1. 発音教育システムの概略。このシステムは、①音声認識技術として開発された音素 HMM による音素の自動切り出しとそれによる教師音声と学習者音声の自動アラインメント、②音素 HMM の認識結果に基づく発音の自動評価、③基本周波数パターンの抽出と教師音声と学習者音声の時間同期した比較、④波形編集合成による学習者音声の矯正、⑤教師音声と学習者音声のディスプレイ上での比較表示、等の機能を有する。

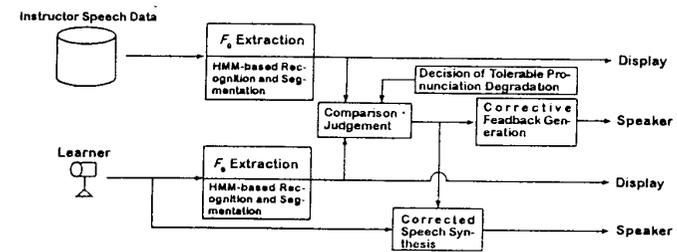


図2. 日本語発音教育システムの画面表示の例。学習者に音素の異なりによって意味が対立する単語対を発声させ、どの程度正しい発音として許容できるかを百分率表示した上で発音矯正指示を表示する。これは「恋」と「好意」を発声した場合であり、/o/の長さを知覚実験結果と照らし合わせて発音の良さを定量的に判定している。また、音素 HMM による発音評価により、「好意」の長母音が日本語の/o/になっていないことを検出している。

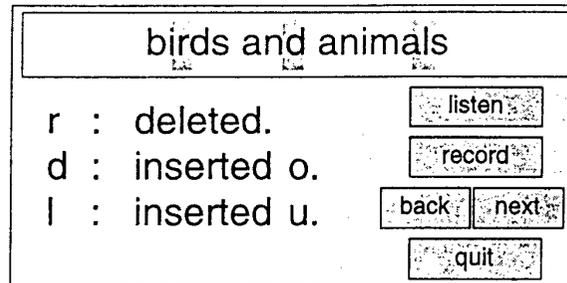
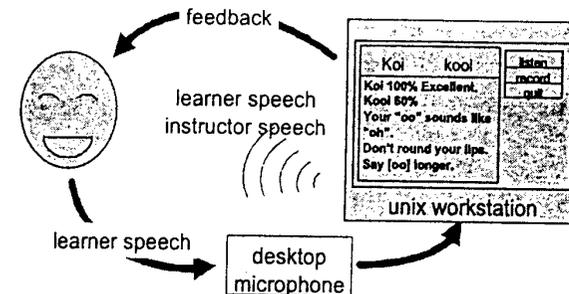


図3. 英語発音教育システムの画面表示の例。日本人話者の英語発音における母音挿入誤りと子音脱落誤りが表示されている例。

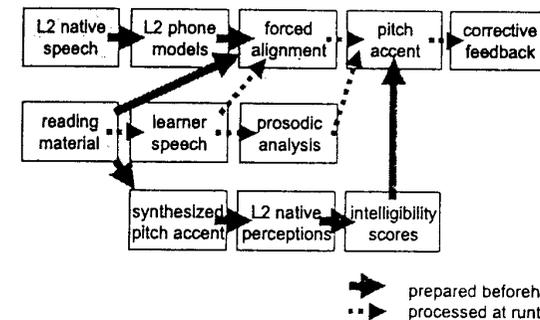


図4. 日本語アクセント型発音教育の処理の流れ。

(注:フローチャート図,ブロック図,構成図,写真,データ表,グラフ等 研究内容の補足説明にご使用下さい。)