

## 3S-4

日本語教育支援システムにおける  
教師・生徒発声音比較の 時間軸の非線形伸縮による自動化

金井 浩\* 星 晋\*\* 菊地重昭\*\* 川添良幸\* 牧野正三\*\*\* 城戸健一\*\*\*

\* 東北大学情報処理教育センター \*\* 東北工業大学通信工学科

\*\*\* 東北大学応用情報学研究センター

1. まえがき 外国人留学生に対する日本語教育を効果的に行なうためには、単に模範音声を繰返し聞かせるだけでは充分ではなく、日本語の文に含まれる種々の音声的特徴に関して、各々模範となる教師の音声と対比し、それに基づき教示することが重要となる。具体的には、日本語発声文のホルマント、パワーパターン(ストレス)、ピッチパターン(アクセント・イントネーション)等の特徴量に関して各々評価・表示する訳であるが、その際、正しい評価を行なうためには、教師発声の模範音声との発音のタイミングを時間的に正規化して、対比表示を行なう必要がある。

そこで本論文では、音声認識技術を応用して、生徒と教師発声音声間の各音素の特徴に基づいて相互の時間軸を自動的に伸縮した上で、両者の発聲音声間の特徴量の違いをホルマント、ピッチ、パワー、発声速度毎に別々に評価する方法を提案する。

2. 原理 教師による模範音声と学習者の音声を、各々10KHz、12bitでA/D変換後、フレーム長25.6ms、フレーム間移動距離10msで、各フレーム毎に15次の線形予測分析を行ない、その後、以下に示す音声的特徴を求める。

(1) 動的計画法(DP)を用いた時間軸の伸縮 線形予測係数から計算できる低次のLPCケプストラム係数(1次~15次)を成分とするベクトルを各フレーム毎に計算し、教師模範音声と生徒音声間のベクトルのユークリッド距離(平滑された対数スペクトル同志の距離)に関して、動的計画法(DP)を利用して、整合を行なう。この方法によって、両者の各フレーム毎の平滑化スペクトル同志が対応するように、教師の音声に対して生徒の音声の時間軸を非線形伸縮することができる。

(2) 音声的特徴量に関する教師・生徒の違いの評価 非線形伸縮された生徒音声と教師音声は、フレーム毎の対応がとれているから、次の正規化されたフレーム毎の特徴量の違いの評価を行なうことができる。

(a) ホルマント周波数の違いの評価 線形予測係数に基づいてスペクトル包絡を計算し、低周波から順に第3ホルマント周波数まで求め、その各々の軌跡を教師・生徒の音声に関して対比表示し、違いを示す。

(b) ピッチ周波数(基本周波数)の違いの評価 線形予測分析した後、得られた残差波形の自己相関の顕著なピークから、ピッチ周波数を決定し、その軌跡を教師・生徒の音声に関して対比表示する。特に、日本語の文の発声においては、例えば「雨」と「鉛」の区別のようにピッチアクセントが支配的となるため、ピッチの遷移状態は、日本語の教育上、特に重要なものの一つと考えられる。また、ここでは予め教師の音声の各母音の中心フレームをラベル付けしておくことにより、より明確なピッチ遷移の評価を可能としている。

(c) パワーの違いの評価 時間軸を正規化されたフレームの各フレーム毎の音声パワーを対数值(dB)で計算し、教師・生徒の対比表示を行なう。

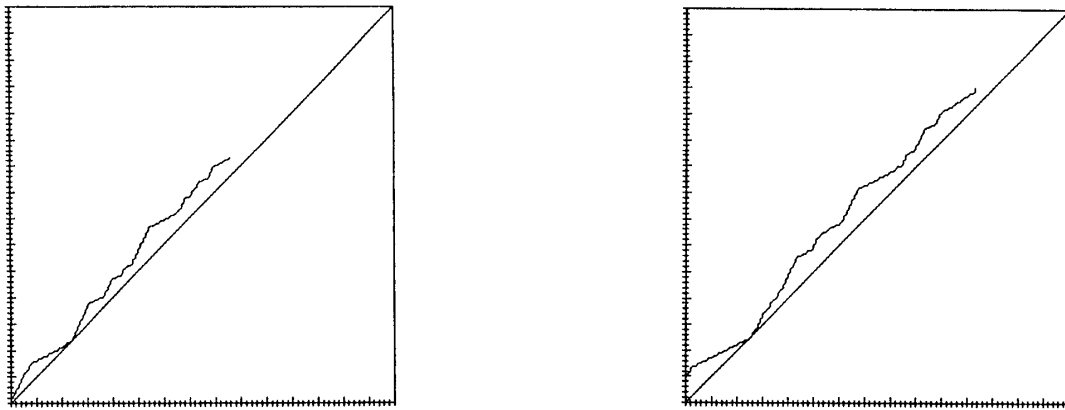
(d) 文の発声のタイミング(持続時間制御)の評価 動的計画法(DP)によって得られた整合結果における最適DPパスから、教師と生徒の局部的発声速度の変化の違いの評価を行なうことができる。この特徴量は、例えば「びったり」という促音に関する発声のタイミングの学習・評価上重要となる。

3. 実験 模範データとして、日本放送出版協会の音声ガイドテープに納められているNHK女性アナウンサーの音声を教師音声とし、生徒として本学の中国人留学生(女性)の音声を用いた。図1に前節で述べた方法によって時間軸伸縮を行なった後の各特徴量の対比表示結果を示す。生徒である留学生の音声に関しては、特に『雨』と『鉛』におけるピッチパターンの変化に乏しいことがわかる。

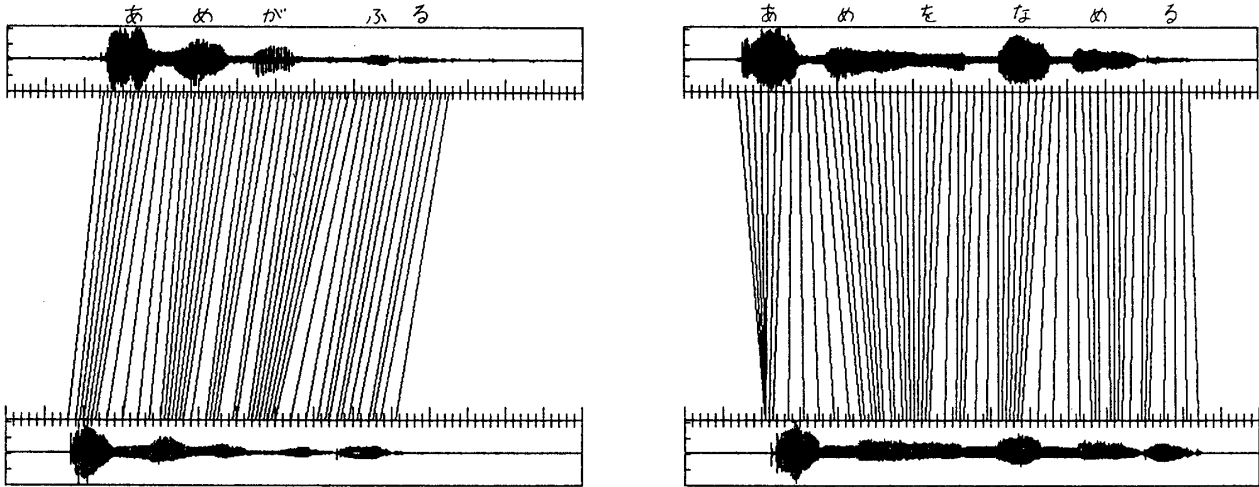
4. まとめ 動的計画法に基づいて時間軸の非線形伸縮を行なうことによって、日本語教育支援システムにおける、教師音声に対する生徒の発声音の評価の自動化のための基本となる手法を提案し、実験例を示した。今後、実際に留学正にたいする日本語教育の現場に於て、本システムの使用を行ない、各特徴量ごとの評価方法、さらには効果的な教育方法を確立して行く予定である。

参考文献 H.SAKOE & S.CHIBA, IEEE trans. Vol.ASSP-26, No.1, pp.43-49, 1978.

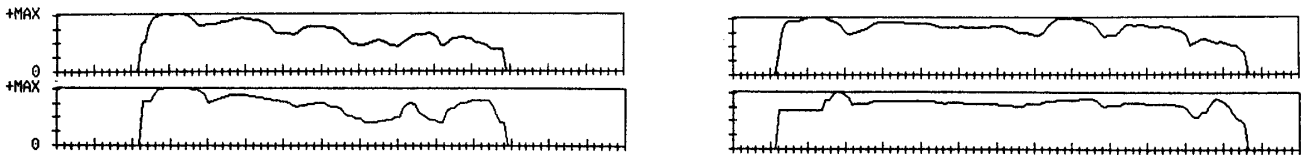
謝辞 データ整理をして頂いた東北工業大学 日比野悦子さんに謝意を表す。



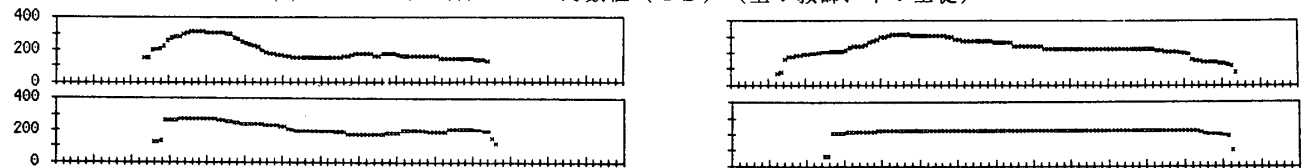
(1) 動的計画法により得られたDPパス (横軸: 生徒、縦軸: 教師発声音のフレーム)



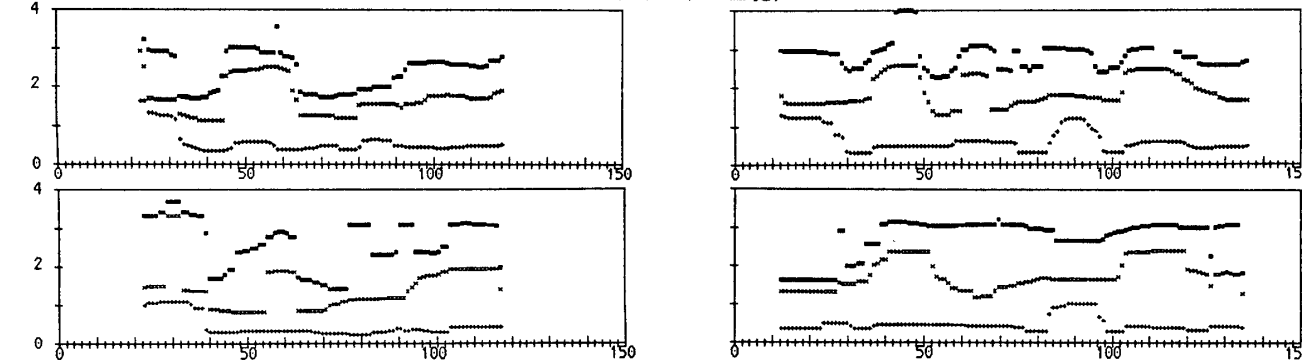
(2) 教師と生徒発声音波形と、発声のタイミング整合結果 (上: 教師、下: 生徒)



(3) フレーム毎の音声パワー対数値 (dB) (上: 教師、下: 生徒)



(4) ピッチ周波数 (Hz) (上: 教師、下: 生徒)



(5) 第1~第3ホルマント周波数 (kHz) (上: 教師、下: 生徒)

図1、時間軸伸縮後の教師・生徒発声音の音声特徴量の対比表示

(a) 『雨が降る』

(b) 『鉛を舐める』