

英語プレゼン学習支援アプリにおける採点機能の検討

窪田 知也, 堀 智子, 吉本 定伸, 小嶋 徹也*1,

*1 東京工業高等専門学校

Consideration on scoring function in English presentation learning support application

Tomoya Kubota, Tomoko Hori, Sadanobu Yoshimoto, Tetsuya Kojima *1

*1 National Institute of Technology, Tokyo College

近年グローバル化が進み、企業の会議や国際的な発表の場において、英語でのプレゼンテーションを行う機会が増えている。本研究では主に学生が効果的な英語プレゼンテーションを行えるようにするために、理想的なプレゼンテーションと学生のプレゼンテーションの相違点を客観的に評価し、採点することのできるシステムの開発を行い、実用可能なアプリケーションを開発することが目標である。

キーワード:英語プレゼンテーション, 採点機能, ピッチ

1. はじめに

近年グローバル化が進むにつれ英語プレゼン技術の向上が求められている。とりわけ英語の発音は効果的なプレゼンをするために不可欠なため、学生を対象としたプレゼン時の英語発音向上を目的としたアプリケーションの開発を行う。

従来、英語の発音練習をする場合は、お手本となる音声を学習者が聞き、アクセントや抑揚、スピードなどをお手本に合わせて発音練習を繰り返し行い、正しい発音を学ぶことが多い。しかし、母国語と異なる英語でのプレゼンテーションでは、お手本となるプレゼンテーションとの相違点を見つけることは難しく、どの程度お手本のプレゼンテーションと一致しているのかを評価するためには、ピッチとポーズと音響インテンシティが重要とされている。(1)

これまでの本研究のプロジェクト(1)(2)(3)において、英語プレゼン学習支援アプリケーションの機能は、お手本の音声(wavファイル)を再生する機能、学習者の発音した声の特徴をリアルタイムで表示する機能(図1)、音声再生中にリアルタイムで英文を表示する機能、音声のピッチを抽出する機能(4)(5)を開発してきた。また、本研究で取り扱う採点機能(スピーチスコアリングシステム(SSS))を考案した。

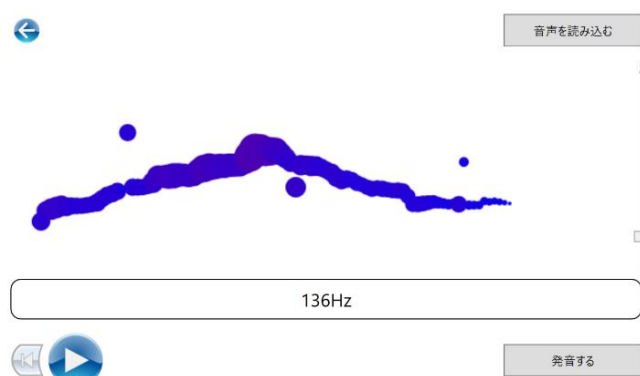


図1 リアルタイムでの音声特徴表示

本研究では英語プレゼン学習支援アプリケーションにおいて、お手本となる音声の特徴と学習者が発音した音声特徴の類似性に注目して、採点を行うためのアルゴリズムについて検討する。今回は音声特徴の中のピッチに着目してお手本の音声と学習者の音声を比較し、採点に利用することを検討する。

2. 採点機能の検討

2.1 前処理

採点を行うために音声ファイルからピッチ抽出を行うが、このデータをそのまま採点に用いてしまうと、ノイズなどの障害により正しく採点することが難しいため、採点を行う前に前処理を行う。以下に前処理の手順を示す。

2.1.1 ローパスフィルタ (LPF)

まず初めにピッチ抽出においてノイズ及び高周波成分が存在しているとピッチ抽出に問題が生ずることがあるため、音声データのノイズ及び高周波成分の除去を行うために、ローパスフィルタ (以下 LPF) を用いる。LPF 適用前後の音声データを図 2, 図 3 に示す。

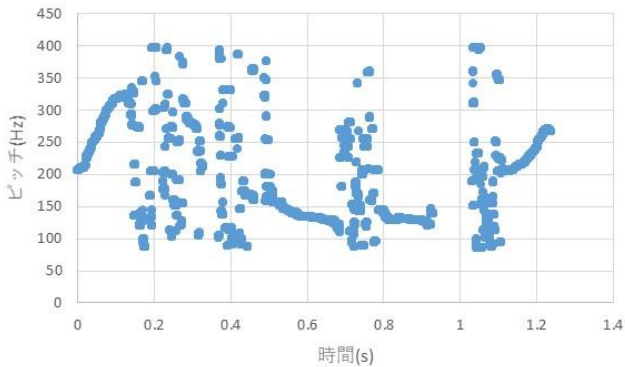


図 2 LPF 適用前

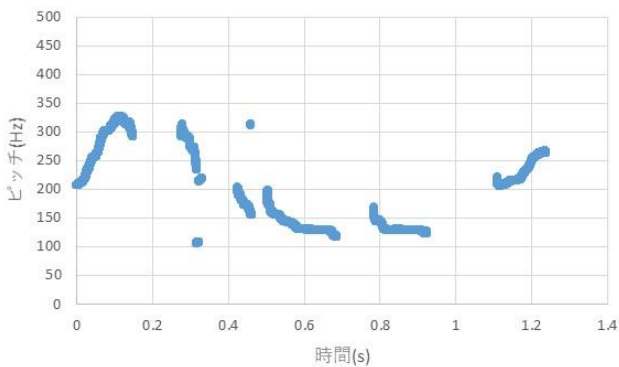


図 3 LPF 適用後

このように LPF を用いることによってノイズが大幅に改善されたのが分かる。

2.1.2 セミトーン変換

採点を行う際に性別や年齢による周波数の違いによる採点への影響を抑えるために、LPF 後のピッチ抽出したデータにセミトーン変換を適用する。セミトーン変換の式を(1)に示す。またセミトーン変換後のピッチデータを図 4 に示す。

$$\text{Semitone} = 12 \frac{\log_{10} \frac{f_1}{f_2}}{\log_{10} 2} \quad (1)$$

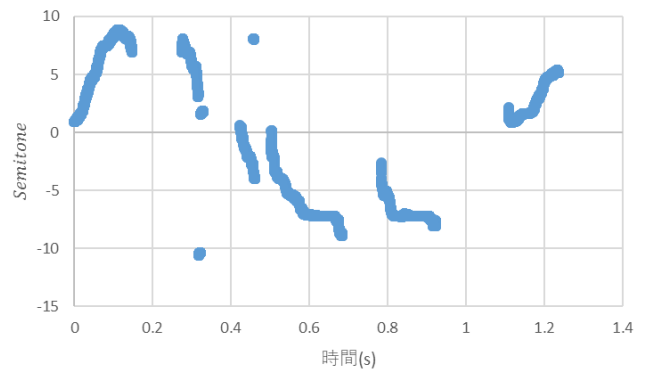


図 4 セミトーン変換後のデータ

図 3 及び図 4 より、ほぼピッチ抽出結果に影響なく変換することができていると分かる。

2.1.3 時間のずれ修正

お手本のデータと学習者のデータでは時間のずれが生じているため、ずれを修正し正しく採点が行えるようにする必要がある。時間のずれを修正するために、音声が発見される以前の無音の部分に当たる不要なデータの削除を行う。不要なデータ削除前後のデータを図 5, 図 6 に示す。

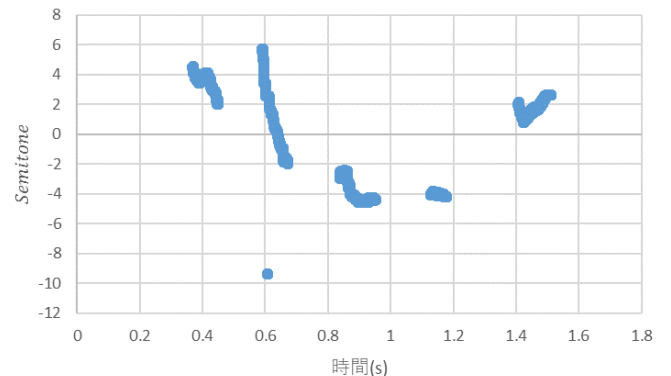


図 5 不要データ削除前のデータ

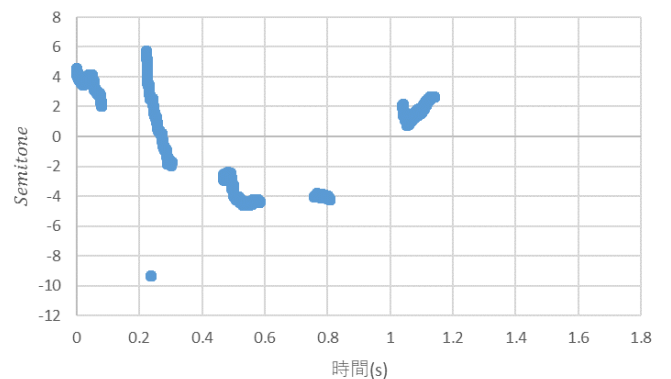


図 6 不要データ削除後のデータ

このように不要なデータが削除されているのが分かる。

2.1.4 今回使用するピッチデータ

今回採点に利用したピッチデータを図4及び図7に示す。図4に示したピッチデータがお手本のピッチデータであり、図7に示したピッチデータが学習者のピッチデータである。

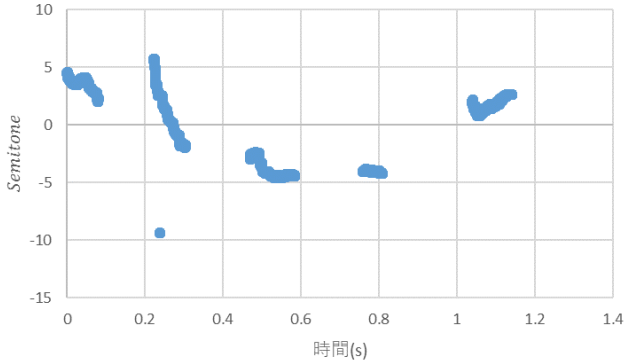


図7 学習者のピッチデータ

2.2 採点方法の検証

まず基本的な採点方法として、セミトーン変換後のピッチデータの値を直接用いて採点する方法（採点方法1）及び、セミトーン変換後のピッチデータを用いてピッチの傾きを計算しピッチの変化に着目して採点する方法（採点方法2）の検証を行う。

2.2.1 採点方法1

採点方法1ではセミトーン変換されたピッチデータの値を直接比較していく方法を用いた。計算に用いた式を(2)に示す。

$$a_n = \frac{u_n - m_n}{m_n} \quad (2)$$

a_n 計算結果

m_n お手本のピッチデータ

u_n 学習者のピッチデータ

(2)式によって時間ごとのお手本のピッチデータとの誤差率を計算し、(2)式の計算結果から誤差率の平均値を求め、以下に示す(3)式を用いて得点とする。

$$(1 - a_{ave}) \times 100 = A \quad (3)$$

a_{ave} 誤差率の平均値

A 最終的な得点

2.2.2 採点方法2

採点方法2ではセミトーン変換されたピッチデータの値から以下に示す(4)式を用いてお手本のピッチデ

ータ及び学習者のピッチデータから傾きを求める。

$$s_n = d_{n+1} - d_n \quad (4)$$

s_n 計算結果

d お手本または学習者のピッチデータ

以下に示す(5)式を用いて誤差率を計算し、(5)式の計算結果から誤差率の平均値を求め、(3)式を用いて得点とする。

$$a_n = \frac{us_n - ms_n}{ms_n} \quad (5)$$

a_n 計算結果

ms_n お手本のピッチデータから求めた傾き

us_n 学習者のピッチデータから求めた傾き

2.3 採点方法の検証

図4と図7に示すお手本と学習者のピッチデータを用いて、採点方法1と採点方法2で誤差率の計算を行った結果を、以下の図8及び図9に示す。

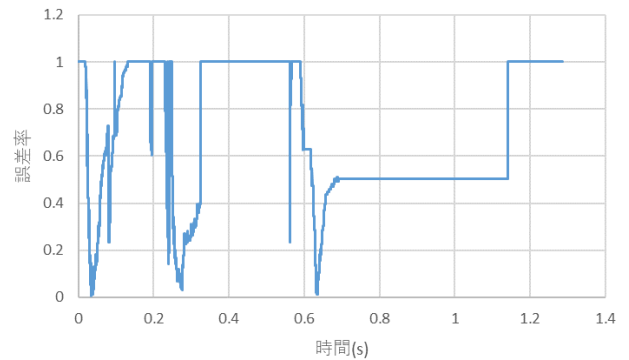


図8 採点方法1の誤差率

採点方法1の点数 30.549点

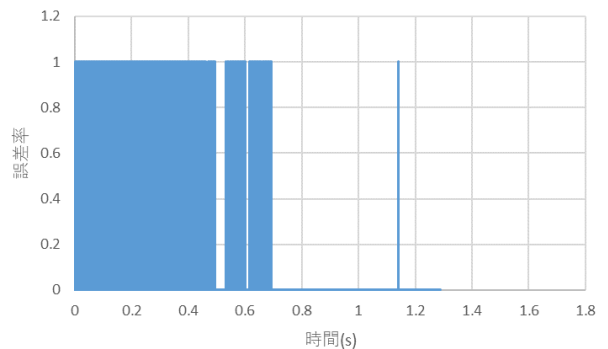


図9 採点方法2の誤差率

採点方法2の得点 96.188点

図8に示す採点方法1の採点結果から直接ピッチデータの値を比較する方法では、お手本のピッチデータの値と学習者のデータの値が大きく異なっていたため30.549点と低い点数であった。

また図 9 に示す採点方法 2 の採点結果から傾きを計算し採点する方法では、お手本または学習者のピッチデータ内に同じ値のデータが数十個続く場合もあり、このとき傾きが 0 となる時刻が非常に多くでてきてしまう。さらに傾きが変わる時刻の時だけ誤差率が 1 を示し、次の時刻では傾きが 0 になるため、得点に偏りができてしまう。これらの原因からこの方法では 96.188 点と非常に高い点数となった。

以上の結果から採点方法を見直し、採点に関して有効だと考えられる計算結果のデータのみを抽出し、その計算結果を用いて採点が行えるように改良を行った。改良を行った採点方法 1 を採点方法 3 とし、改良を行った採点方法 2 を採点方法 4 とする。

改良を行った採点方法で誤差率の計算を再度行った結果を図 10、図 11 に示す。

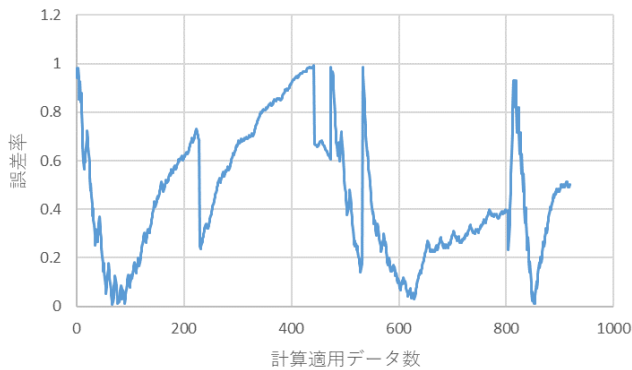


図 10 採点方法 3 の誤差率
採点方法 3 の得点 54.734 点

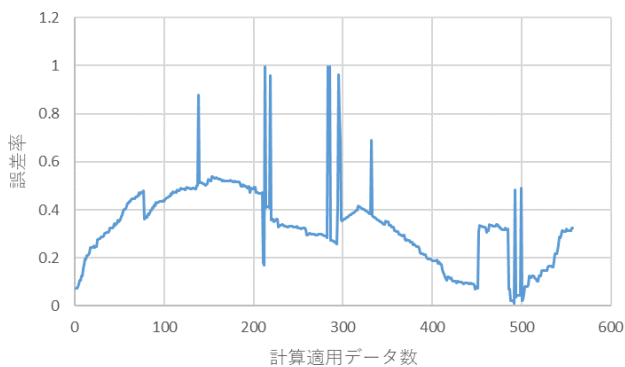


図 11 採点方法 4 の誤差率
採点方法 4 の得点 67.077 点

3. 考察

採点結果から当初想定していた改良前の採点方法

では、音声の類似性に関する採点において有効でない計算結果を多く採点結果に含んでしまったため、採点に偏りができてしまった。採点方法を見直し採点に有効と考えられる計算結果を抽出し、採点を行うように改良した結果、採点方法 3 での得点が 54.734 点、再度採点した採点方法 4 での得点が 67.077 点となった。この採点結果からピッチの特徴を捉え採点することができたと考えられる。

4. おわりに

ピッチの類似性に着目して採点を行うアルゴリズムの検討を行った。とりわけ改良を行った採点方法 3 と採点方法 4 について比較的良い結果が得られたと考えられる。今後は英語プレゼン学習支援アプリケーションに今回検討した採点機能を実装するとともに、更なる採点方法向上のために録音されたプレゼンの採点区間を分割する、より英語プレゼンにおいての特徴を捉えた採点するため、単語ごとに採点が行えるように改良する等、精度の高い採点を実現できるようにする必要がある。

謝辞

本研究は、JSPS 科研費 25370680 の助成を受けたものである。

参考文献

- (1) 松永竜太郎:英語プレゼンテーションのための音声学習支援ソフトウェアの開発, 教育システム情報学会要旨, (2014)
- (2) 橋積裕紀:英語プレゼンテーション学習支援ソフトウェアの開発, 教育システム情報学会要旨, (2015)
- (3) 西原悠貴:英語プレゼンテーション学習支援ソフトウェアの開発, 教育システム情報学会要旨, (2016)
- (4) 斉藤収三, 加藤勝洋, 寺西昇:“音声の基本周波数の特性について”, 日本音響学会誌, vol.14, no.2, pp.111-116,(1958)
- (5) 矢野博夫・日高新人・橋秀樹:音響インテンシティによる音場の可視化, https://www.jstage.jst.go.jp/article/souonseigyoy1977/15/4/15_4_170/_pdf, (2009)