

英語学習者音声のリズムの評価に用いる テキストの必要条件

Requirement of Texts for Evaluation of Rhythm in Learners' English Speech

中村 静, 匂坂芳典

Shizuka NAKAMURA and Yoshinori SAGISAKA

要旨

英語音声のリズムの評価では、学習者の習熟度の表れ方や評価者の評価の正確さは、そこで用いるテキストの特性に依存する。評価者によって付与された主観評価値を予測する客観評価モデルをより効率的に構築することを目指して、英語音声のリズムの評価に用いるテキストの必要条件について検討した。客観評価モデルは、学習者と母語話者の対応する音節持続時間の差異に基づいて構築した。英語のリズムは、強音節と弱音節とが交互に繰り返されることで知覚される。必要条件の単位には、その繰り返しの途中でよく知覚される強音節に着目して、1テキストに含まれる強音節の個数を用いた。予測実験の結果、音節持続時間の差異を用いて線形回帰モデルによって客観評価を行う場合には、1テキストに強音節が3個含まれることが、英語のリズムの評価に有効なテキストの必要条件であることがわかった。

1 緒論

国際化社会が進行するにつれて、英語運用能力の習得の需要が増大してきた。それに伴い、学習者の英語音声の評価に関して、人間による主観的な評価手法に代えてコンピュータを用いた客観的な評価手法の確立が望まれてきた。この客観評価手法では、英語教師の労力の軽減、評価基準の共通化、多量なデータの迅速な処理、CALL (Computer-assisted language learning) システムへの応用等を目的としている。

音声には、音韻性と韻律性の2面がある。日本語を母語とする学習者の英語音声の客観評価に関する先行研究では、主に自動音声認識技術が適用されて、音韻性の評価を目的とするものが多くなされてきた [例えば、1-2]。一方、韻律性の評価を目的とした先行研究では、韻律性を反映する音響特徴量である、強度、基本周波数、持続時間等を学習者と母語話者とで比較して、その類似度を用いた評価手法が検討されてきた [例えば、3-5]。

英語音声の韻律性は、強音節と弱音節との対比に基

づくリズムによって特徴付けられる [6-7]。このリズムは日本語音声には存在しない。したがって、日本語を母語とする学習者が発話した英語の韻律性は、リズムによく表れることが期待される。また、高等学校学習指導要領解説 [8] にも記載されているように、リズムは意図や感情を伝える上で重要な役割を担っており、学習者はそれを的確に習得することが望まれている。そこで、本研究では、韻律性の中で、英語音声の学習において重要な学習課題であるリズムの習熟度を評価の対象とする。

英語音声のリズムの知覚には、音節の心理的な大きさや高さや長さを伝える強度や基本周波数や持続時間といった音響特徴量が関与するが、本研究ではまず持続時間に焦点を絞る。その理由は、音節持続時間は、それ自体で長さを伝えるのに加えて、その持続時間の中で変化する音節の強度や基本周波数とも強い相関を持っていて、英語のリズムを形成するのにとくに重要な貢献をすることが実験的にも確かめられている [9] ためである。

本研究の主観評価では、英語音声のリズムの習熟度について7段階に振り分ける評価尺度を使う。客観評価では、英語教師が文毎に付与した主観評価値を予測する。この予測には、英語のリズムの基盤であり、学習者の特徴がよく表れることが期待される、音節の持続時間 (以降、音節持続時間と呼ぶ) の、学習者と母語話者との差異を利用する。

英語音声のリズムの評価では、学習者の習熟度の表れ方や評価者の評価の正確さは、そこで用いるテキストの特性に依存するために、そのテキストの特性が、予測精度に大きく影響することになる。先行研究では、テキストの長さによって異なる予測精度が得られている [10-11]。本研究では、客観評価モデルをより効率的に構築することを目指して、英語音声のリズムの評価に用いるテキストの必要条件について検討する。

本論文の構成は、次の通りである。2章では、まず、学習者による発話や評価者による評価が緻密になされるように考慮したうえで、本研究の実験に用いるテキストの基本的性質を揃えるための、1テキストの長さ

と単語の難易度に関する基準を設ける。その後、英語音声のリズムの構造と知覚特性の観点から、テキストの必要条件を設定する。3章では、2章で設けたテキストの基本的性質についての基準を満たす音声試料と、主観評価値を準備し、それらを用いて構築する客観評価モデルについて述べる。4章では、2章で設定した必要条件の妥当性について、音節持続時間の差異と主観評価値との相関に基づいて実験的に検証するために、構築した客観評価モデルを使って予測実験を行う。

2 リズムの評価に用いるテキストの必要条件の設定

テキストの必要条件に関してより明確な結果を得るためには、本研究の実験に用いる音声試料の各テキストの基本的性質を揃えておく必要がある。本章では、まず、学習者による発話や評価者による評価が緻密になされるように考慮したうえで、テキストの基本的性質を揃えるための、1テキストの長さや単語の難易度に関する基準を設ける。その後、この緻密な発話や評価の範囲内で、英語音声のリズムの構造と知覚特性の観点から、テキストの必要条件を設定する。

2.1. テキストの基本的性質の統一

2.1.1. 1テキストの長さ

学習者による発話や評価者による評価が緻密になされるようにするためには、1テキストの長さの限界を把握する必要がある。そのために行った予備実験では、1テキストの長さや、学習者音声の音響的特徴や、複数の評価者による評価傾向の類似の度合との関係を調べた。1テキストの長さ、学習者音声の音響的特徴、評価傾向の類似の度合を表す客観的数値には、各々、音節数、学習者と母語話者との対応する音節持続時間の二乗誤差を1テキストにわたって平均した値、主観評価値の評価者間での相関係数を用いた。実験の結果、緻密な発話や評価に用いる音節数の限界は、11から14程度であることがわかった。この音節数を、1テキストの長さに関する基準とする。

2.1.2. 単語の難易度

学習者とその発音を知らないような常用でない単語がテキストに含まれていることは、学習者が意図して発話した英語のリズムを評価することを困難にする。本研究での評価の対象は、特定の単語の音韻性ではなくて、文全体にわたる韻律性のうちのリズムの習熟度である。そのため、話者によってテキストに含まれる単語に関する知識の差がないことが望ましい。そこで、ほとんどが中学必修レベルの簡単な単語で構成され、固有名詞が含まれないことを、単語の難易度に関する基準とする。

2.2. リズムの知覚特性に基づく必要条件の設定

強勢は英語音声のリズムの構造を特徴づける。強音節が強勢を受けていると認知されるのは、それが周りの強勢のない弱音節より、持続時間が長いことや、声が高いことや高いことによって、際立っているためである [7]。これらの強音節と弱音節とが交互に繰り返されることで、リズムが感知される [9]。本研究では、この繰り返しの中でよく知覚される強音節の特徴に着目して、1テキストに含まれる強音節の個数をテキストの必要条件の単位とする。

英語では、1テキストに1個以上の強音節が含まれる。1テキストに強音節が1個含まれる場合、強音節同士による間隔は形成されない。このようなテキストの例を、表1の一番目に示している。この表では、強音節を“@”の記号で示している。1テキストに強音節が2個含まれる場合、それらの強音節の間に1個の間隔が形成されて、その絶対的な長さが知覚される。この例を、表1の二番目に示している。1テキストに強音節が少なくとも3個含まれる場合、それらの隣接する強音節が形成する間隔が2個以上存在して、その周期的な繰り返しによるリズムも感知される。この例を、表1の三番目に示している。これらの理由から、評価者がリズムを感知するためには、1テキストに強音節が少なくとも3個含まれる必要がある。

前節では、学習者による発話や評価者による評価が緻密になされるように、テキストの基本的性質について基準を設けた。その基準を満たすように次章で選出したテキストには、表2に示すように、各々強音節が3個含まれていた。そこで、1テキストに強音節が3個含まれることを、テキストの必要条件に定める。この必要条件の妥当性については、音節持続時間の差異と主観評価値との相関に基づいて、4章で実験的に確かめる。

母語話者が英語音声を発話する場合に、強勢の位置や程度は、強勢同士が近づきすぎるのを避けて規則的な間隔を保とうとするという英語の一般の規則によって、変化する場合がある [12]。しかし、全ての母語話者がこの変化を同様に起こすとは限らない [7]。また、普段音声を聴取する場面では、母語の聴取者であっても、強勢の程度にまでは注意が払われない [6]。これらの英語音声学で指摘されているリズムの特性が母語話者の音声サンプルで表れる実態を考察した結果、評価者間で強勢の認知の違いによる影響を小さくして、より明解な結果を得るために、本研究では、弱音節とは対照的に最も際立つ第1強勢のある音節のみを強音節として扱うことにした。以降では、第1強勢のある音節のみを強音節と呼び、その他を弱音節と呼ぶ。

表1 1テキストに強音節が1, 2, 3個含まれるテキストの例. 音節境界は“・”の記号で示している. 強音節 (@) と弱音節 (-) は2.2で述べた定義に基づく.

| 1テキストに含まれる強音節の個数 | テキスト |
|------------------|---|
| 1 | I'm a·mused. - - @ |
| 2 | I'm a·mused by the man. - - @ - - @ |
| 3 | I'm a·mused by the man and his ver·y fun·ny jokes. - - @ - - @ - - - - - @ |

表2 実験に用いるテキスト. 各テキストには, 強音節が3個含まれている. 音節境界は“・”の記号で示している. 強音節 (@) と弱音節 (-) は2.2で述べた定義に基づく.

| | テキスト |
|---|---|
| A | I'm a·mused by the man and his ver·y fun·ny jokes. - - @ - - @ - - - - - @ |
| B | Why won't you wait un·til Fri·day when he's back? - - - @ - - @ - - - @ |
| C | Thank you ver·y much for eve·ry·thing that you did for us. - - - @ - @ - - - @ - - |
| D | The boys have sold some of the flow·ers. - @ - @ - - - @ - |
| E | I was ter·ri·bly an·noyed with the man for beat·ing the dog. - - - - - @ - - @ - - - - @ |

3 主観評価値を予測する客観評価モデルの構築

本章では, 前章で設けたテキストの基本的性質についての基準を満たす音声試料と, 主観評価値を準備し, それらを用いて構築する客観評価モデルについて述べる.

3.1. 実験に用いる試料

3.1.1. 学習者と母語話者の音声サンプル

実験に用いる音声試料は, 前章で設けたテキストの基本的性質についての基準を満たすテキストが含まれている「日本人学生による読み上げ英語音声データベース (English speech database read by Japanese students (以降, ERJデータベースと略す)) [13]」の中から選択した. このデータベースには, 統一された環境で録音された, 幅広い習熟度の学習者によって発話された英語音声格納されている. 実験には, 表2に示す, 前章で設けたテキストの基本的性質についての基準を満たすテキストを5個選んだ. 表2の各テキストの2行目に“@”の記号で示すように, 各テキストには強音節が3個含まれていた.

学習者音声には, 合計106サンプルを選定した. 話者は日本語を母語とする大学生106名であった. 1テキスト当たりのサンプル数は約21であった. ERJデータベースが作成された過程では, 練習時や録音時に母語話者による音声サンプルが参考として提示されるこ

とはなかった. また, 話者にはあらかじめ強勢の位置と程度が提示され, 各自の読み上げ練習の後に録音が行われた.

ERJデータベースには, 学習者と同じテキストが読み上げられた母語話者音声も格納されている. 母語話者音声には, 前述の学習者音声に対応する20名の母語話者が発話した58サンプルを選んだ. 1テキスト当たりのサンプル数は約11であった.

3.1.2. 主観評価値

学習者音声の全サンプルに対して, 英語教師が主観評価を行った. 評価者は, 英語音声学の知識を有し, 日本語を母語とする話者に対する英語教育経験を持つ5名であった. 評価者に3.1.1で選んだ母語話者音声の話者は含まれていない.

評価には, 文全体にわたる英語のリズムの習熟度についての, -3 (非常に習熟度が低い) から+3 (非常に習熟度が高い) までの7段階に振り分ける評価尺度を用いた. 主観評価値は, 文毎に付与された. 評価者には, 個々の音声サンプルを複数回聴くことを許した.

3.2. 線形回帰モデルを用いた客観評価モデル

学習者は, 母語話者の英語のようにリズムを制御することを目指している. そこで, 学習者音声を同一テキストが発話された母語話者音声と比較して, 計測された対応する音声区分の持続時間の差異が小さいこと

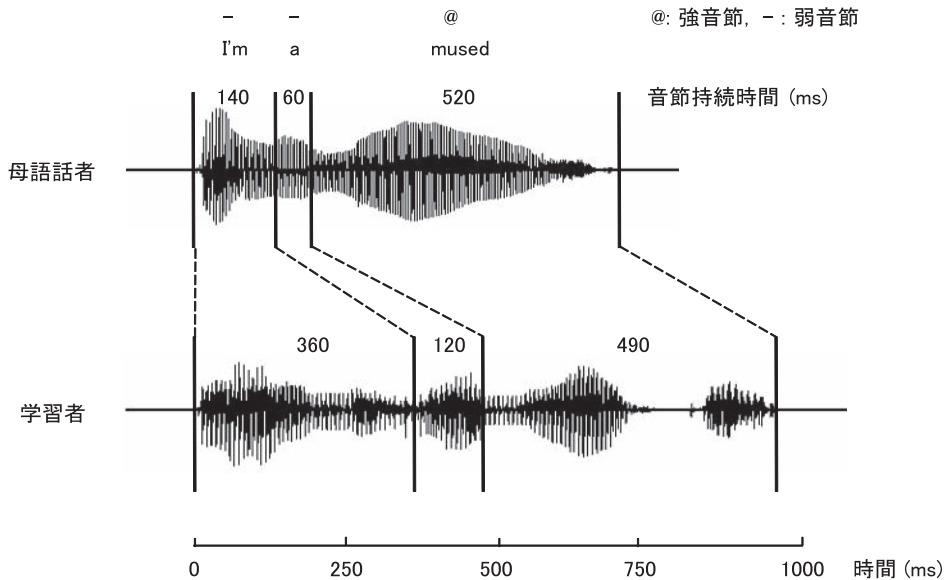


図1 音声波形を用いて示した、音節持続時間の母語話者と学習者との比較の例。

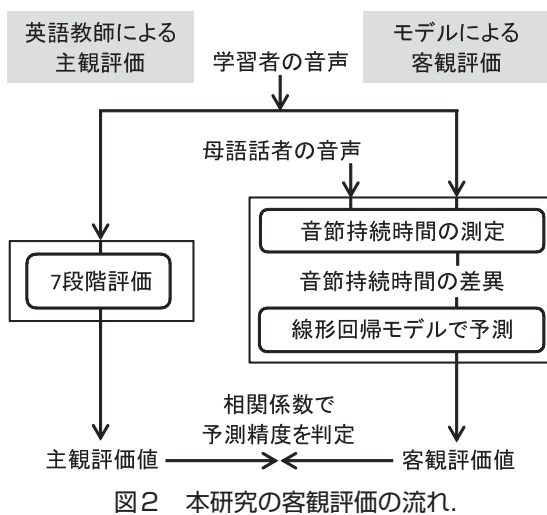


図2 本研究の客観評価の流れ。

を、英語のリズムの習熟度が高いことを表す特徴量として扱った。持続時間の差異は、図1に示すように、英語のリズムに深くかわり、学習者の特徴がよく表れることが期待される各音節で計測した。この音節持続時間について、学習者に系統的な特徴があることを、母語話者との比較によって次項で確かめている。これらの音節持続時間は、音声認識器 (HTK [14]) により10ms単位で行った自動音素アラインメントの結果を用いて計算したものに、英語音声学の知識を有する専門家による手作業の修正を加えたものである。

本研究の客観評価の流れを図2に示す。学習者音声と母語話者音声と比較して、特徴量である音節持続時間の差異を計算する。それを用いて主観評価値を予測する客観評価モデルを構築する。客観評価モデルが算出する客観評価値が主観評価値に近いほど、客観評価モデルの予測精度が高いことになる。この精度は、主観評価値と客観評価値との相関の強さで判断する。

客観評価モデルは、変数の寄与を単純にして理解し

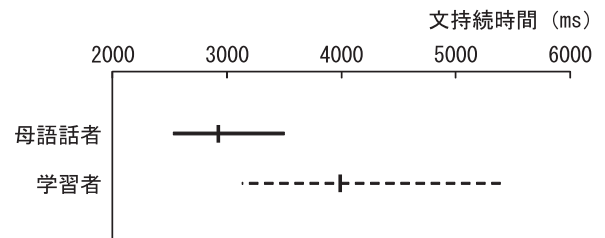


図3 同一テキストを読み上げた際の、母語話者 (実線) と学習者 (点線) の、文持続時間の揺らぎの例。母語話者は学習者ほど範囲が広がってはいないが、母語話者でもかなりの揺らぎがある。

やすくするために、線形回帰モデルを用いて構築した。学習者 i のテキスト j の主観評価値である $S_{(i,j)}$ の予測値 $\tilde{S}_{(i,j)}$ を求めるには、以下の線形回帰モデルの式を使った。

$$\tilde{S}_{(i,j)} = \alpha \sqrt{\frac{1}{n} \sum_{k=1}^n (D_{(i,j,k)} - \hat{D}_{(j,k)})^2} + \beta$$

説明変数には、学習者 i のテキスト j の k 番目の音節持続時間 $D_{(i,j,k)}$ と、母語話者のテキスト j の k 番目の音節持続時間 $\hat{D}_{(j,k)}$ の二乗誤差を用いた。 α と β は、主観評価値とその予測値との誤差が最小になるように決定した。これらのデータの予備的処理については、3.2.2と3.2.3で述べる。

3.2.1. 学習者の音節持続時間の特徴

音節持続時間の学習者と母語話者との特徴の差異を詳細に分析するために、まずは母語話者と学習者の文の持続時間 (以降、文持続時間と呼ぶ) の特徴を把握する。図3は、表2のテキストAを例として、母語話

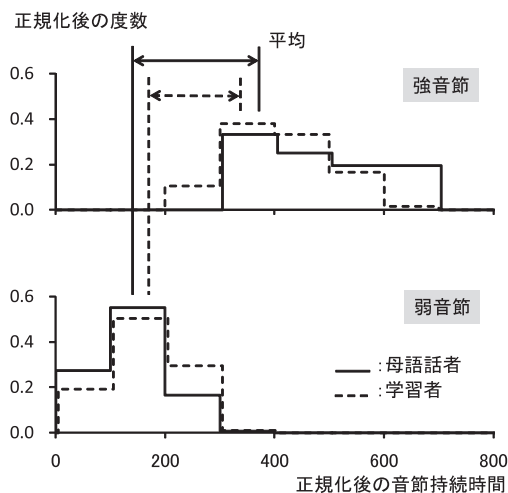


図4 学習者と母語話者の、強音節（図上）と弱音節（図下）の持続時間の度数分布。学習者音声では、母語話者音声と比較して、強音節の伸長が少なく、弱音節の短縮が少ない傾向にある。

者と学習者の文持続時間の範囲を横線で表したものである。短い縦線は、話者間の平均値を示す。上が母語話者、下が学習者の結果である。

このように、学習者でも母語話者でも話者間で揺らぎがある。この揺らぎは、どのテキストでも確認された。揺らぎが最も大きいテキストでは、最も長い母語話者は最も短い母語話者の約1.5倍であった。学習者の場合では、約2.6倍であった。実測の音節持続時間には、このような文持続時間の違い、すなわち、全体的な発話の速さの違いによる影響が含まれている。この影響を除いて、リズムの特徴の差異を的確に計測するためには、学習者音声の文持続時間を規範の母語話者音声の文持続時間に揃えた後に音節持続時間を比較する必要がある。

音節持続時間の学習者と母語話者との特徴の差異を調べるために、学習者と母語話者の音節持続時間の度数分布を比較した。図4では、テキストAを例にとり、図上に強音節、図下に弱音節の結果を示している。

前述のように、文持続時間の違いによる影響を除いて比較するために、図4では、全ての音声サンプルの文持続時間を母語話者間平均文持続時間で正規化して、この伸縮後の音節持続時間を横軸にとっている。度数については、話者数と音節数とで正規化して、それらの影響を除いた後の数値を縦軸にとっている。話者は学習者23名と母語話者11名である。

音節の構成等によって変動的ではあるが、一般に、強音節は強くゆっくりと発話されるため、その持続時間は弱音節より長い。実際に、図4の例では、強音節は長く、母語話者と学習者の平均持続時間は、各々372と332である。一方、弱音節は短く、母語話者と学習者の平均持続時間は、各々129と160である。学習者の母語話者との差異に着目すると、強音節は伸長が少なく、弱音節は短縮が少ない傾向にある。この傾向は、

他のテキストでも確認された。これらの強音節と弱音節の持続時間の学習者の特徴は、先行研究 [15] でも確認されている。本研究では、このような学習者のリズムの特徴を反映する音節持続時間の母語話者との差異を利用して客観評価を行う。

3.2.2. 目的変数としての主観評価値

3.1.2で述べたように、主観評価では、各々の1文全体の音声サンプルに対して、評価者毎に1個の主観評価値が与えられた。その結果、各音声サンプルには合計5個の評価値が付与された。

本研究の客観評価モデルで用いる線形モデルでは、1個の主観評価値が1個の音声サンプルに対応していることが望ましい。そのため、音声サンプル毎に主観評価値の代表値を算出した。この代表値は、著者らの先行研究 [15] の手法に従って、5名の評価者の評価特性を揃えるように、7段階に振り分ける評価尺度で付与された各評価者の主観評価値の度数分布が、上限・下限まで一様に近づくように補正したうえで、評価者間平均を求めた値とした。このようにして求めた主観評価値の代表値を、以降では単に主観評価値と呼ぶ。

3.2.3. 説明変数としての音節持続時間の差異

図3に示した通り、学習者は母語話者よりゆっくり発話するため、文持続時間は母語話者より学習者のほうが長い傾向にある [5, 10, 15]。学習者の母語話者とのリズムの特徴の差異を的確に計測するために、先行研究 [10] の手法に基づいて、学習者の文持続時間を母語話者の文持続時間の代表値で正規化した後に、音節持続時間の差異を計測した。

母語話者音声の規範には、音節毎に母語話者間の平均持続時間を使った [16-17]。これにより、母語話者間の持続時間の揺らぎの影響を統計的に最小限に抑えることができるため、母語話者に共通した英語の傾向を取り出せる可能性が高いと考えられる。これは、学習者との差異の計測に用いる規範に適している。

4

予測実験による必要条件の妥当性の実験的検証

2章では、英語音声のリズムの構造と知覚特性の観点から、テキストの必要条件について検討した。その結果、緻密な発話や評価がなされる範囲内で、強音節と弱音節とが交互に繰り返されることで形成される英語音声のリズムを感知できるように、1テキストに強音節が3個含まれることを必要条件に設定した。英語音声のリズムの構造から考察すると、強音節が3個の場合と、2個、1個の場合とでは、評価者によるリズムの感知のされ方が大きく異なると思われる。本章では、必要条件の妥当性について、音節持続時間の差異と主観評価値との相関に基づいて実験的に検証するた

表3 1テキストに含まれる強音節の個数の違いによる、主観評価値と客観評価値との相関係数の比較。

| 1テキストに含まれる強音節の個数 | トレーニング | テスト |
|------------------|--------|------|
| 3 | 0.70 | 0.68 |
| 2 | 0.52 | 0.49 |
| 1 | 0.38 | 0.34 |

めに、構築した客観評価モデルを使って予測実験を行う。

まず、必要条件を満たすテキスト、つまり、1テキストに強音節が3個含まれるテキストの音声試料を用いて、予測実験を行った。客観評価モデルは、前章までに述べたようにデータを準備して構築した。データセットは、106のサンプルで構成された。3分割交差検定法によってモデルのトレーニングとテストを行った。予測精度を示す主観評価値と客観評価値とのテストデータに対する相関係数は、表3に示す通り、0.68であった。

次に、必要条件を満たさないテキスト、つまり、1テキストに強音節が2個、または、1個含まれるテキストの音声試料を用いて、比較実験を行った。ERJデータベースには、表2の5種類の各テキストから順次句や単語を省いて短くしたテキストも含まれている。1テキストに含まれる強音節が2個と1個のテキストをこの中から選び出して、参照群に利用した。表1では、強音節が3個含まれる表2のテキストAと、それに対応する強音節が2個、または、1個含まれるテキストを示している。各々のデータセットは、上述の実験と同数の話者による同数のサンプルで構成された。主観評価値は、3.2.2で述べたように用意した。

表3に示すように、テストデータに対する相関係数は、強音節が2個と1個の場合に、各々0.49と0.34であった。強音節が3個の場合より、各々0.19と0.34弱まった。強音節が3個と1個、3個と2個の場合には、各々有意水準1%で有意差がみられた。学習者と母語話者との対応する音節持続時間差異を用いて線形回帰モデルによって客観評価を行う場合には、1テキストに強音節が3個含まれることが、2個や1個より英語のリズムの評価に有効なテキストの必要条件であることがわかった。

5 結論

英語音声のリズムの評価では、学習者の習熟度の表れ方や評価者の評価の正確さは、そこで用いるテキストの特性に依存する。評価者によって付与された主観評価値を予測する客観評価モデルをより効率的に構築することを目指して、英語音声のリズムの評価に用いるテキストの必要条件について検討した。客観評価モデルは、学習者と母語話者の対応する音節持続時間の

差異に基づいて構築した。

学習者による発話や評価者による評価が緻密になされる範囲内で、英語音声のリズムの構造と知覚特性の観点から、テキストの必要条件を検討した。英語音声のリズムの構造から考察すると、強音節が3個の場合と、2個、1個の場合とでは、評価者によるリズムの感知のされ方が大きく異なると思われる。強音節と弱音節とが交互に繰り返されることで形成される英語音声のリズムを感知できるように、1テキストに強音節が3個含まれることを必要条件に設定した。

この必要条件の妥当性について、音節持続時間の差異と主観評価値との相関に基づいて実験的に検証するために、構築した客観評価モデルを使った予測実験を行った。その結果、1テキストに強音節が3個含まれる場合には、2個や1個の場合より、予測精度が有意に向上することが実験的に確かめられた。したがって、音節持続時間の差異を用いて線形回帰モデルによって客観評価を行う場合には、1テキストに強音節が3個含まれることが、英語のリズムの評価に有効なテキストの必要条件であることがわかった。

本研究では、客観評価モデルに用いる音響特徴量として、それ自体で長さを伝えるのに加えて、その持続時間の中で変化する音節の強度や基本周波数とも強い相関を持つ、音節持続時間を利用した。客観評価の予測精度の向上を目的とした著者らの先行研究 [11] では、客観評価モデルに用いる音響特徴量についてすでに検討され、持続時間に強度で重み付けした値の利用による精度の向上が確認されている。今後は、このような学習者の特徴量の抽出に関する結果と、本研究で得られたような評価の対象に関する結果とを融合させて、精度の高い客観評価モデルを効率的に構築することを目指す。

謝辞

この研究の一部には、日本学術振興会特別研究員奨励費No. 21・06162と早稲田大学特定課題研究助成費2010B-367による援助を受けた。

参考文献

- [1] 中川聖一, A. A. Reyes, 鈴木英之, 谷口泰広. “音声認識技術を利用した英会話CAIシステム.” 情報処理学会論文誌, 38, 517-526 (1997).
- [2] 河合 剛, 石田 朗, 広瀬啓吉. “2言語の音響モデルを用いた音声認識による非母語発音誤りの検出と発音評価.” 日本音響学会誌, 57, 569-580 (2001).
- [3] Y. Yamashita, K. Kato, and K. Nozawa. “Automatic scoring for prosodic proficiency of English sentences spoken by Japanese based on utterance comparison.” *IEICE Trans. Inf. Syst.*, E88-D, 496-501 (2005).
- [4] A. Ito, T. Konno, M. Suzuki, and S. Makino. “Improvement of automatic English prosody evaluation based on word clustering using a decision tree.” *IEICE Trans. Inf. Syst.*, J91-D, 358-366 (2008).
- [5] 中村 静. 音韻持続時間の差異に基づく英語学習者音声

のリズム制御に対する客観評価. 2008年度早稲田大学大学院国際情報通信研究科修士論文 (2009).

- [6] D. Jones. *An Outline of English Phonetics* (W. Heffer & Sons LTD., Cambridge and Maruzen Company LTD., Tokyo, 1960), 245-247.
- [7] P. Roach. *English Phonetics and Phonology* (Cambridge University Press, Cambridge, 2009), 64-88.
- [8] 文部科学省. 高等学校学習指導要領解説 外国語編・英語編 (開隆堂, 東京, 2010), 9-32.
- [9] I. Lehiste. *Suprasegmentals* (MIT Press, Cambridge, 1970), 106-153.
- [10] S. Nakamura, H. Tsubaki, Y. Kondo, M. Nakano, and Y. Sagisaka. "Tempo-normalized measurement and test set dependency in objective evaluation of English learners' timing characteristics." Proc. ICPhS (Saarbrücken, Germany), 1733-1736 (2007).
- [11] S. Nakamura, S. Matsuda, H. Kato, M. Tsuzaki, and Y. Sagisaka. "Objective evaluation of English learners' timing control based on a measure reflecting perceptual characteristics." Proc. IEEE ICASSP (Taipei, Taiwan), 4834-4840 (2009).
- [12] P. Ladefoged. *A Course in Phonetics* (Harcourt Brace Jovanovich, Texas, 1975), 102-103.
- [13] N. Minematsu, K. Nishina, and S. Nakagawa. "Read speech database for foreign language learning." *J. Acoust. Soc. Jpn.*, **59**, 345-350 (2003).
- [14] S. Young, G. Evermann, T. Hain, D. Kershaw, G. Moore, J. Odell, D. Ollason, D. amnd Povey, V. Valtchev, and P. Woodland. *The HTK Book (for HTK Version 3.4)* (Cambridge University Engineering Department, Cambridge, 2006).
- [15] S. Nakamura. "Analysis of relationship between duration characteristics and subjective evaluation of English speech by Japanese learners with regard to contrast of the stressed to the unstressed." *J. Pan-Pacific Association of Applied Linguistics*, **14**(1), 1 -14 (2010).
- [16] S. Nakamura, H. Kato, and Y. Sagisaka. "Effects of mora-timing in English rhythm control by Japanese learners." Proc. INTERSPEECH (Brighton, England), 1539-1542 (2009).
- [17] S. Nakamura, and Y. Sagisaka. "A requirement of texts for evaluation of rhythm in English speech by learners." Proc. ICPhS (Hong Kong, China), 1438-1441 (2011).