

## 科学研究費助成事業 研究成果報告書

平成 26 年 6 月 27 日現在

機関番号：82404

研究種目：若手研究(B)

研究期間：2012～2013

課題番号：24700582

研究課題名(和文) 吃音の言語治療のための訓練音声合成および自動評価手法の開発

研究課題名(英文) Development of method for speech synthesis and automatic evaluation for stuttering treatment

研究代表者

越智 景子(OCHI, Keiko)

国立障害者リハビリテーションセンター(研究所)・研究所 感覚機能系障害研究部・流動研究員

研究者番号：20623713

交付決定額(研究期間全体)：(直接経費) 2,100,000円、(間接経費) 630,000円

研究成果の概要(和文)：本研究では、吃音の言語訓練の補助を目的とし、発話速度を遅くする訓練のために発話速度を表示するシステムを作成した。非吃音被験者に対してモデル音声の発話速度を保持する実験を行った結果、とくに遅い発話において目標に合わせることが困難であった。そこで、発話速度の視覚的なフィードバックの短期的な効果を調べる実験を行った結果、フィードバックがある場合は発話速度を低減させることができ、フィードバックを外した後もその効果が保持されることがわかった。さらに、発話速度の表示システムを貸し出して連日練習を行った結果、4人中3人は7日以内で目標とした遅い発話速度をフィードバックがない場合でも再現できるようになった。

研究成果の概要(英文)：In this study, a new home-based computerized program that consists of visual feedback of articulation rate was developed to assist adults who stutter (AWS) to slow down their speaking rate. Three experiments were setup to examine the efficacy of this new program: First, when recording the normal subjects' speaking rate during passage reading with no visual feedback, participants were not able to slow down their speech to the slow target speech rate level. In second experiment where visual feedback was presented to seven AWS, participants were able to slow down their speech to the target level. This effect was maintained even after removing the visual feedback. In third experiment, four AWS utilized this new program to train their speech rate at their home (5 minutes per day). The results showed that three AWS were able to slow down their speech to the target level after 1-week home-based treatment.

研究分野：リハビリテーション科学・福祉工学

科研費の分科・細目：言語聴覚療法学

キーワード：吃音 発話速度 韻律

## 1. 研究開始当初の背景

吃音は有症率が1%と言われ、繰り返し・引き延ばし・阻止といった中核症状を特徴とする。それ以外の非流暢と、中核症状が原因で生じた二次的行動の進展による重症化・症状の固定化により患者の社会参加の側面に影響が及ぶことがある(Healey, et al., 2004)。

根本的な治療法はないものの、言語療法によって症状を軽快させることができる(Guitar, 2005)。言語療法は副作用の危険性がある薬物療法や、常時携帯が必要な支援機器を用いる方法とは異なり、最終的には独力で流暢な発話を可能にすることを旨とするものである。成人吃音の言語療法では、流暢性形成法(Fluency shaping approach (Mallard and J.S. Kelley, (1982))が最も一般的に行われている。

この訓練方法は、発話全体を訓練によって変更することで流暢性を向上させる治療法であり、低年齢児を除いた幅広い年代の吃音児・者に用いることができる。訓練では、治療者と一緒に文章を読み上げる斉読、後に続いて読み上げる追唱、斉唱後の単独音読といった練習や、中核症状を抑えるために発話速度を適度に遅く調整した発声の習得などを行う。

効率的に新たな発話方法を習得するためには日常的な練習が必要である。訓練者を伴わずに練習を行う場合、発話速度等が適切であるか患者自身が客観的に判断することは困難である。そこで、音声信号処理技術を用いて発話が適切であるか否かを測定し、練習効果が向上すると考えられる。

コンピュータを使った訓練の有効性は海外においては示されている(Euler, et al. 2009)。しかし、日本語母語話者である吃音者の訓練を目的としたものの開発および治療効果の評価は少ない。

また、発話の速さを低減させる言語訓練は、吃音のみではなく構音障害の言語訓練で広く用いられているが、遅い発話は休止区間が発話区間より大きく伸長することが報告されているため(比企他, 1968)、休止を除いた発話区間の調音の速さを遅くすることは困難であると考えられる。また、調音の速さの調整についての研究は行われているが(Tsuao, 2006)、日本語において遅い速度に関して複数段階の速さを調査した研究やその発話速度を保持することに関する研究は少ない。

また、斉唱・復唱の練習において多様な語句、文に対応するためには、定型句の録音では不十分である。そのため、音声合成技術によって任意の文の日本語音声を適切な速度で作成し、教示音声とすることが効果的であると考えられる。合成音声に対して斉唱・復唱が行えるようにするには、アクセント・リズム・イントネーションといった音声の韻律的特徴に発話速度が与える影響を考慮することによって、より自然な発話に近づけることが必要であると考えられる。

日本語の任意速度での音声合成では音素の継続長などを柔軟に制御できる隠れマルコフモデルを用いた手法(Iwano, et al. 2002)において、基本となる「速い」「普通」「遅い」の3種類の速さの録音音声をもとに補間により所望の発話速度を得る手法が開発されているが、発話速度の違いによる韻律の構造の変化は考慮していない。

音声の基本周波数パターン(以下、F0パターン)の生成過程モデル(Fujisaki and Hirose, 1984)(以下、F0モデル)は、日本語の韻律的特徴のうち特に重要なF0パターンを定量的に記述することができる。そのため、発話速度によるポーズ、また基本周波数生成過程も出るで記述できる発話単位の変化を分析することによって自然な教示音声を合成する上で効果的であると考えられる。

## 2. 研究の目的

本研究では、流暢性形成法の言語訓練を補助することを目指し、流暢性形成訓練の中でも特に発話速度のゆっくりとしたものへの調整に焦点を当てて、練習者が単独で練習が可能となる練習システムの開発を行う。

そのため、まず、さまざまな発話速度を保持できるか否かを調査する。さらに、発話速度の自動評価システムを実装し、マイクより録音した練習音声の発話速度を視覚的にフィードバックできるようにする。さらに、その発話速度のフィードバックにより発話速度の低減がしやすくなるかどうかを検証する。また、音声合成への応用を目指し、発話速度の低減による韻律的特徴の変化を分析する。

## 3. 研究の方法

### (1) モデル音声の追唱後の発話速度の調整

モデルとして示された教示音声と同じ速度に自分自身の発話の速さを合わせて発話する実験を行った。ゆっくりした速度になるように時間伸長した教示音声を被験者に提示し、モデル音声に続いて同じ速さで2回追唱させた。その後追唱の時と同じ速さを保つように指示して、追唱の文とは異なる100文節程度の文章を音読させた。モデル音声は同一の文を時間伸長させた音声1文である。モデル音声の速さは3段階で、ランダムな順序で各速度1回ずつ計3回提示した。また、実験に先立ち、普段の速度での音読を収録した。

録音した音声資料から単位時間当たりのモーラ数を発話速度として算出した。ただし、本研究はリハビリテーションでの応用のため、休止を除いた音声区間から発話速度を算出した。すなわち音声区間において単位時間あたり発話されたモーラ数を発話速度と定義する。

### (2) 発話速度の視覚的フィードバックの短期的な効果の検証

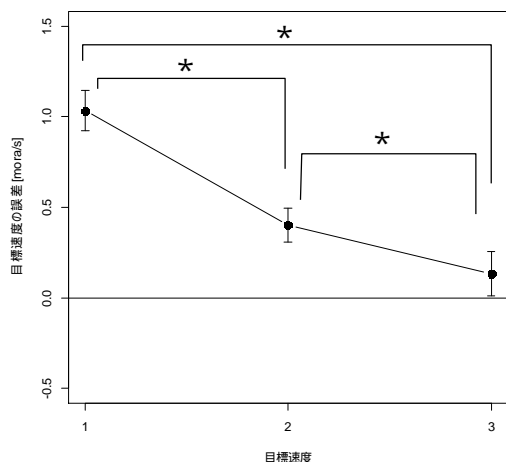


図1 各目標速度における発話速度の誤差

目標速度と音読速度との差。誤差棒は標準誤差を表す。

発話者の音声をマイクにより録音し、音声認識技術(A Lee, T Kawahara, 2009)により発話速度を計測し、視覚的に提示するシステムを実装した。そのフィードバックシステムを用いて発話者に発話速度を提示する実験を行った。すなわち、画面にそれまでの発話の発話速度と目標速度が表示される。実験課題は(1)と同様に教示音声によって与えられた発話速度を保持して文章を音読するものであり、目標速度は時間伸長したモデル音声の1文の2度の追唱によって与えた。実験は3セッションから構成され、初めのセッション A1 では発話速度のフィードバックを与えず、次のセッション B ではフィードバックを与え、最後のセッション A2 では再びフィードバックを外した状態で音読課題を実施した。各セッションでは異なる6段階の目標速度をランダムな順序で与えた。音読は各試行同じ15文を使用した。

### (3) 発話速度と韻律的特徴

発話速度が変化することによる音声の韻律的特徴の変化を調べるため、(2)の実験における音声の分析を行った。対象は非吃音者の発話で、速度の視覚的なフィードバックを伴わないセッション A1 の音声とした。

### (4) 発話速度のフィードバックを用いた速度調整の訓練効果の検証

発話速度を視覚的にフィードバックするシステムの訓練効果を検証するため、被験者にフィードバックシステムを搭載したパーソナルコンピュータを貸与し、それを用いて連日発話速度の調整訓練を行った。目標速度は被験者ごとに1つに設定し、練習するよう指示した。1回につき10文を発話速度のフィードバック行いながら行う。1日目以降の練習では、フィードバックを伴う音読に先立って、フィードバックがない状態で覚えた速さ

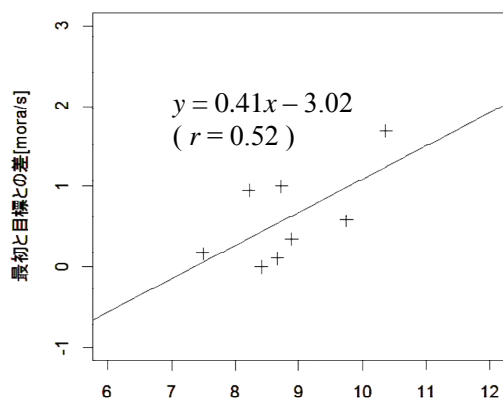


図2 普段の発話速度と、目標速度1のときの初めの50モーラにおける平均速度の目標速度からの誤差

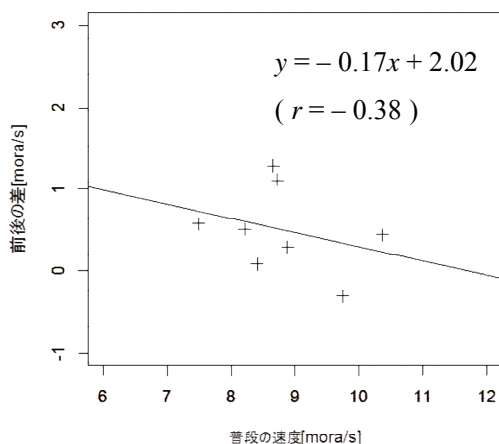


図3 普段の発話速度と目標速度1のときの初めのブロックと最後のブロックでの速度の差

での音読を5文について行わせた。

## 4. 研究成果

### (1) モデル音声の追唱後の発話速度の調整

図1に非吃音者8名(男性5名,女性3名)を対象とした、目標とするモデル音声の速度ごとの音読速度の誤差を示す。発話速度は音声認識ソフトウェア Julius(Lee and Kawahara 2009)を用いて測定した。モデル音声の発話速度の測定結果は、低い順に目標速度1, 2, 3はそれぞれ3.34 mora/s, 4.46 mora/s, 5.59 mora/sであった。各被験者の発話速度は50モーラごとの平均値を1ブロックとして50モーラごとの平均値をとった。その結果、目標速度1および目標速度2のときの音読速度は、目標速度3ときの音読速度より有意に大きかった。

また、図2に各被験者における普段の発話

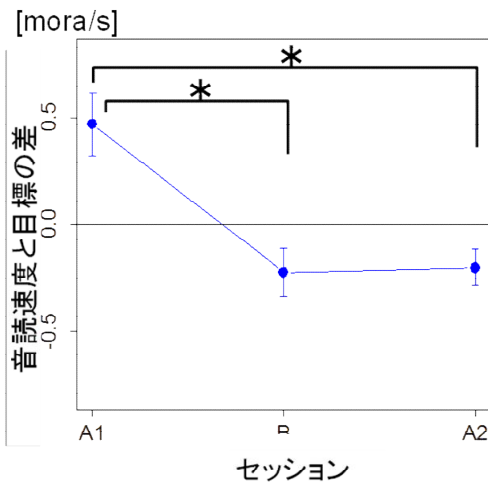


図 4 非吃音群の各セッションでの発話速度 誤差棒は標準誤差を表す。

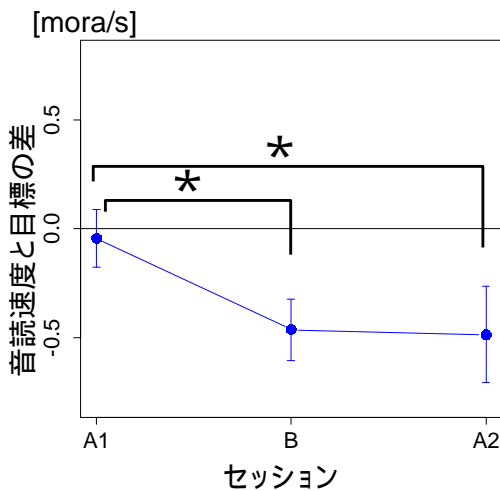


図 5 吃音者群の各セッションでの発話速度 誤差棒は標準誤差を表す。

速度と、目標速度 1 のときの第 1 ブロックの平均速度の目標速度からの誤差を示す。普段の速度と発話調整課題の初めの速度の間に正の相関がみられた。それに対し、図 3 に示す、各被験者の普段の発話速度、および初めと最後のブロックでの速度の差の間には相関がみられなかった。

以上より、遅い話速の中でもとくに普段の発話速度よりも離れて低い発話速度は、モデル音声の追唱のみでは再現できないことがわかった。また、普段の音読速度が速い者ほど目標速度より高い速度で音読することがわかった。

(2) 発話速度の視覚的フィードバックの短期的な効果の検証

被験者として、吃音者(軽度~中等度 7 名)と非吃音者(7 名)を対象とした各セッションの発話速度の目標からの誤差をそれぞれ図 4、図 5 に示す。セッション間で音読速度と目標速度との差を比較した結果、非吃音群および非吃音群は共に、セッション B とセッション A2 がセッション A1 より有意に遅かった。

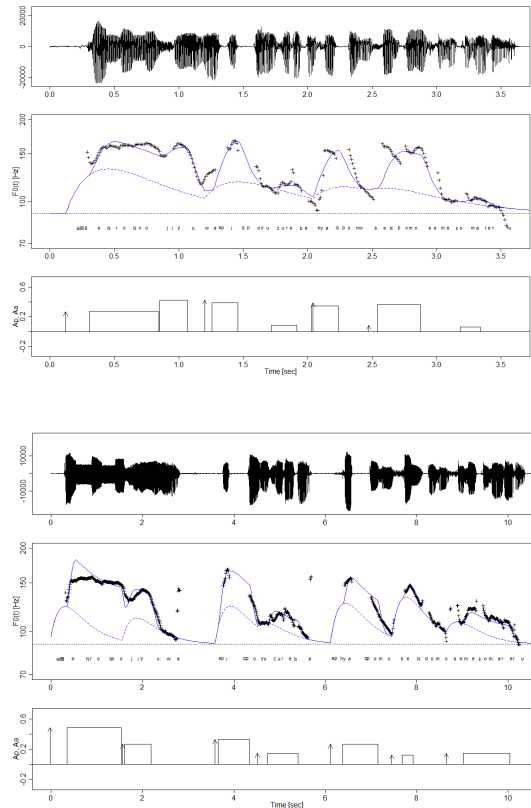


図 6 普段の速さ(上段)、およびゆっくりの速さ(下段)で発話した文の F0 パターンの F0 モデルによる分析例

したがって、視覚フィードバックによって速度低減することができ、かつフィードバックをなくした後も短期的に話速低減の効果が持続したといえる。

(3) 発話速度と韻律的特徴

図 6 に、同一発話者が同一文を異なる発話速度で音読したさいの F0 パターンと、その F0 モデルによる分析例を示す。各発話中の上段が音声波形、中段が F0 パターン、下段が F0 モデルの指令を表す。指令はアクセント句に対応しステップ状のアクセント指令と、フレーズに対応するインパルス状のフレーズ指令からなる。発話速度が普段の速さの場合に比べ、遅い場合はアクセント指令の数の増加はないものの、フレーズ指令の数は遅い話速において増加しており、各文節頭に生起している。

次に、韻律的特徴のうち文内のポーズの発話速度による影響を示す。各文内に含まれるポーズの平均モーラ長と文のモーラ数で正規化したポーズ数を図 7、各文の平均モーラ長と各文の平均ポーズ長とを図 8 に示す。発話者ごとにばらつきがみられるものの、平均ポーズ長と発話速度の間には正の相関が見られる。したがって発話速度を遅くするさいにポーズも伸長していることがわかる。以上の結果を用いることで、遅い発話速度の音声

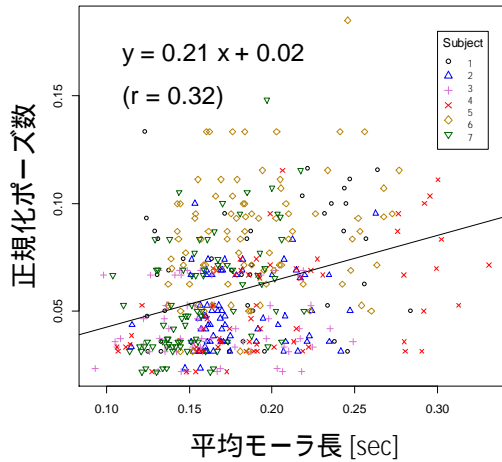


図7 文の平均モーラ長と文の長さで正規化したポーズ数

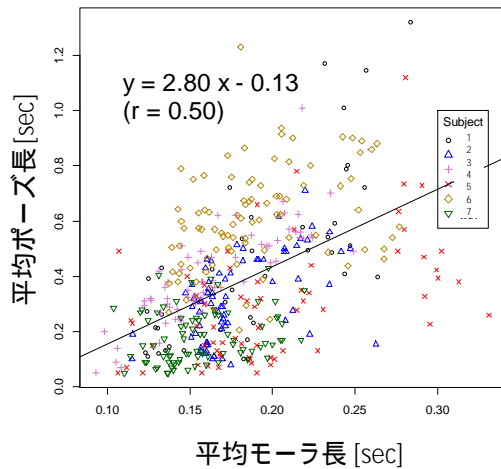


図8 文の平均モーラ長と平均ポーズ長

を合成するうえで、単に時間伸縮を行った音声に比べより自然に近い音声を得ることが可能になると考えられる。

(4) 発話速度のフィードバックを用いた速度調整の訓練効果の検証

図9に各被験者(吃音者4名)の練習開始から7日目までの音速速度の変化を示す。Dの被験者1名を除いて、7日以内に速度が目標より遅くなっている。したがって4名中3名は繰り返しの音読訓練により短期間に目標速度を下回るようになったものの1名についてはより長期の訓練が必要であると考えられる。

5. 主な発表論文等  
〔雑誌論文〕(計1件)

越智景子, 森浩一, 岡田美苗, 錦戸信和,  
音声の追唱後の発話速度制御, 日本音響

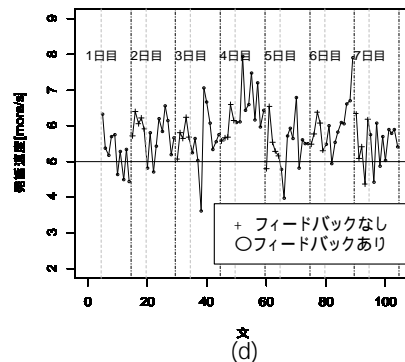
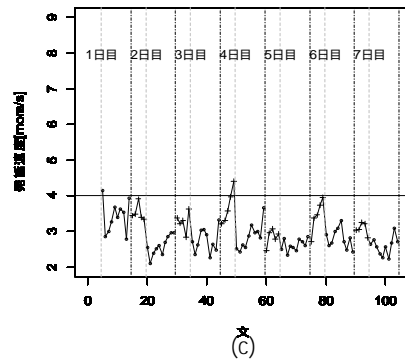
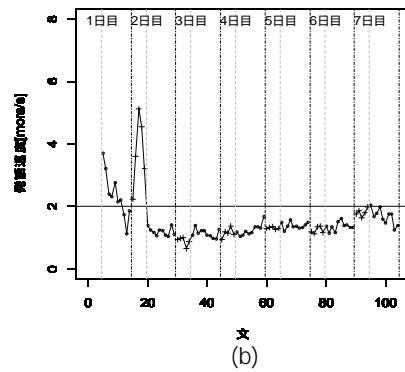
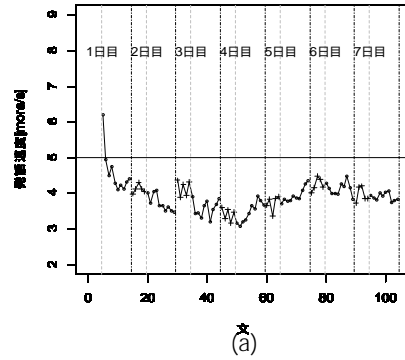


図9 各被験者の発話速度の推移。グラフの横線は目標速度を表す。

会春季講演論文集, pp.567-568, 2013(査読なし)

〔学会発表〕(計2件)

越智景子, 森浩一, 酒井奈緒美, 青木淳,  
“視覚的にフィードバックを与えた際の吃音者・非吃音者の発話速度の調整”, 第

5 8 回日本音声言語医学会，高知市，  
2013-10-17

Keiko Ochi, Koichi Mori, Naomi Sakai,  
and Jun Aoki, “Accuracy of articulation  
rate control with visual feedback in  
persons who do and do not stutter”,  
10<sup>th</sup> Oxford Dysfludncy Conference,  
Oxford, 2014-07-18

## 6 . 研究組織

### (1)研究代表者

越智 景子 (OCHI, Keiko)

国立障害者リハビリテーションセンター  
(研究所)・研究所 感覚機能系障害研究  
部・流動研究員

研究者番号：20623713

### (2)連携研究者/研究協力者

森 浩一 (MORI, Koichi)

国立障害者リハビリテーションセンター  
(研究所)・研究所 感覚機能系障害研究  
部・部長

研究者番号：60157857

酒井 奈緒美 (SAKAI Naomi): 国立障害者リ  
ハビリテーションセンター(研究所)・研究  
所感 覚機能系障害研究部聴覚言語機能障  
害研究部・研究員

研究者番号：60415362

### (参考文献)

H. A. Euler, et, al. (2009) Sprache, Stimme,  
Gehör, vol. 33, no.4, pp. 193-202

H. Fujisaki and Hirose (1984) Journal of  
Acoustic Society Japan (E), Vol.5, no.4,  
pp.322-327

B. Guitar (2005) Stuttering: An Integrated  
Approach to Its Nature and Treatment (3rd  
Edition), Lippincott Williams & Wilkins

E. C. Healey, et, al. (2004) Contemporary  
Issues in Communication Science  
and Disorders, 31, 40-48B.

K. Iwano, et, al. (2002) In Proc. IEEE  
Workshop on Speech Synthesis, pp. 219-222

A.R. Mallard and J.S. Kelley (1982)  
Journal of Fluency Disorders Vol. 7, no.2,  
pp. 287-294

A Lee and T Kawahara (2009) Asia-Pacific  
Signal and Information Processing  
Association Annual Summit and Conference  
(APSIPA ASC)

比企他(1967)音響学会論文誌 23(5), pp. 4-317

Y. C. Tsao, *et al.* (2006) J. Speech, Lang., Hear.

Res., vol. 49, pp. 1156-1164, 2006