

## 科学研究費助成事業 研究成果報告書

平成 29 年 8 月 18 日現在

機関番号：62615

研究種目：若手研究(B)

研究期間：2014～2016

課題番号：26750217

研究課題名(和文) 吃音の言語訓練の自然性向上のためのコンピュータを用いた支援

研究課題名(英文) A study on computer-based speech therapy system to improve speech naturalness

研究代表者

越智 景子(OCHI, Keiko)

国立情報学研究所・情報学プリンシプル研究系・特任研究員

研究者番号：20623713

交付決定額(研究期間全体)：(直接経費) 2,500,000円

研究成果の概要(和文)：発話の障害の一種である吃音の言語訓練では流暢性形成訓練が広く用いられている。流暢形成訓練に含まれる柔らかい声立て(軟起声)の自動判定には従来は立ち上がりにかかる時間が用いられるため、時間をかけた発声が練習される可能性があった。本研究では、立ち上がりにかかる時間によらない音響特徴量から自動判別する手法を開発し、高い判別率を得た。さらに、軟起声の評価を視覚的にフィードバックする言語訓練システムを作成し、それを使って長期的な訓練を行うことによって、軟起声の指示をしない普段の発声でも軟起声を達成することができた。さらに、その発声のさいの構音速度もゆっくりになることを回避できた。

研究成果の概要(英文)：Fluency shaping is widely used in therapy for stuttering, one of speech disorders. Soft voice onset, which is included in the fluency shaping, was conventionally measured by the rise time of the volume and that might cause the unnatural prolongation of the first syllable. In this study, we proposed the acoustic features which were related to the periodicity at the beginning of the phonation to identify the soft onset. We obtained high discrimination rate of 99% by the experimental evaluation. We also developed a home-based training system which visually fed back the acoustic features and showed that one participant could phonate soft onsets without the instruction for soft onset after a long-term training.

研究分野：音声信号処理

キーワード：言語訓練 吃音 音声

## 1. 研究開始当初の背景

言語障害の一種である吃音は、音や語の一部の繰り返し・引き伸ばしと、発話の開始や継続ができなくなる阻止（難発）とを中核症状とし、その他の非流暢も伴う。進展（悪化）に伴い発話困難のため社会参加に影響が及ぶ場合もある。言語訓練として一般的に用いられている流暢性形成法<sup>s</sup>では、発話速度をゆっくりとしたものへの調整や、軟起声とよばれる立ち上がりが柔らかな楽な発声が行われる(Guitar, 1998)。

効果的な訓練を行うためには、訓練施設以外に自宅でも日常的に練習を行い、最終的には習得した発話パターンを自然に聞こえるものへと近づけていくことが重要である。軟起声の自己判定は困難であるため、訓練室外での練習をしやすいするためには、コンピュータ等によって練習の発声を自動的に評価し、提示することが有用であると考えられる。コンピュータは、各種の言語障害の訓練や外国語学習で、幅広く行われており、訓練を円滑にすることができる。吃音については、英語においてWebster(1980)らがコンピュータを利用した流暢形成性訓練を提案し、そのドイツ語版の治療研究(Euler et al., 2009)では、長期的な効果が報告されている。日本語においては、モデル音声と一緒にゆっくりした音読を行う斉唱訓練をコンピュータの画面読み上げソフトを使って実施し、吃音症状が軽快した症例が報告されている(小内山ら, 2013)。また、流暢性形成訓練を受けた後の発話の自然性は、治療前に比べると良好であるものの、非吃音者の発話と比べると少ない(Franken et al., 1992)ことが問題であり、流暢性形成訓練で習得する発話パターンは可能な限り自然に聞こえるものを目指すことが望まれる。

これまで、軟起声の客観的指標として音声波形の振幅の包絡線の傾きが提案され、

臨床家の聴覚印象との一致が示されている(Peters, 1986)。しかし、こうした波形をもとにした手法は、傾斜の測定の基準となる点が一意に定められないため(Orlikoff, et al., 2009)自動処理の上では問題となる。さらに、振幅をゆっくりと増加させることを中心とした教示方法では、最初の音が十分な音量になるまでに時間がかかり、1音目が必然的に引き伸ばされる。

## 2. 研究の目的

軟起声の指導過程で音量を徐々に上げて声立てを行うことが多いが、1音目を不自然に伸ばした発話パターンが練習され、日常で使いにくい印象を与える恐れがある。

そこで、時間によらない音響的特徴の違いから軟起声の判別を行えば、音圧の包絡線のみを指標に使った場合のように、緩やかな傾きでの出だしを作るために1音目の発声をゆっくりと不自然になるほど伸ばしてしまうことが避けられる。それによって、吃症状を低減しつつできる限り自然性が高い発話が練習・習得できると考えられる。本研究では、立ち上がりの時間によらない軟起声の自動判定手法の開発とそれによる訓練効果の検証を目的とする。

## 3. 研究の方法

本研究ではまず、軟起声の音響特徴を利用した自動判別手法を開発し、吃音者および非吃音者の軟起声の指示のもとでの発声を比較した。その後、軟起声の自動判別結果を視覚的にフィードバックする訓練システムを開発し、訓練効果を検証した。具体的には以下の方法で研究を行った。

### (1) 軟起声の自動評価

Hillenbrand(1994)は氣息性の発声は信号の周期性と高い相関があることを報告している。息を出しながらの起声では始めに声帯は接触せずに振動し始め(Orlikoff et

al., 2009)、徐々に接触するようになる。声門の閉鎖時に一部に空いている場所がある状態で声帯が振動する場合、発声される音声の振幅スペクトルが通常発声に比べて高周波帯域で低い値をとることがシミュレーションで示されている(Stevens, 2000)。

本研究では、声帯が開いた状態からの起声を軟起声と定義し、閉じた状態からの起声を硬起声と定義した。軟起声の判別のために、音声の周期性に関わる2つの特徴量を提案した。一つは時間差(dT)で、

$$dT = T_l - T_p \quad (1)$$

により定義される。 $T_l$ はRoot-mean-square (RMS)レベル、 $T_p$ は周期性がそれぞれ閾値を超えた時刻である。軟起声では声門が完全に閉じる前に声帯に小さい振動が生じていることから、2つの時間に時間差ができると考えられるためである。もう一つは氣息性発声の定量化に用いられている(Bickley, 1982)倍音振幅比(H1-H2)で、第1倍音と第2倍音の振幅の比の対数である。

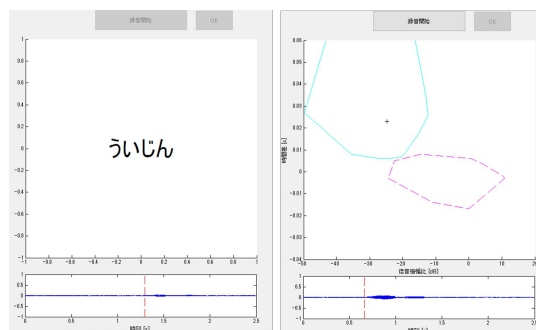
自動判別の学習データとして、吃音の臨床経験がある言語聴覚士(Speech Therapist, ST)4名の単独母音および母音から始まる単語発声を使用した。発声には硬起声、軟起声および長い軟起声が含まれる。長い軟起声は、訓練の初期段階にわかりやすいために使われる1モーラ目を伸ばした発声であり、本研究では長い軟起声を含む軟起声と硬起声の2クラスの判別を行った。判別にはSupport Vector Machine (SVM)を用い、one-speaker-leave-outにより判別率を評価した。

## (2) 吃音者・非吃音者の軟起声の比較

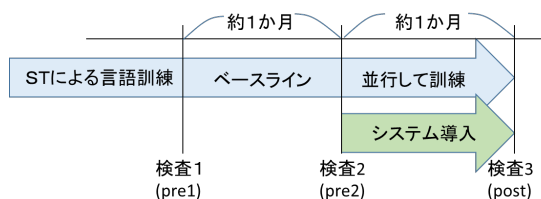
言語訓練の経験のない吃音者(PWS)8名、非吃音者(PWNS)9名が5母音と単語発声(母音から始まる10-40単語)を1-3回軟起声の指示の元で発声した。

## (3) 視覚的なフィードバックの短期訓練

上記の軟起声の判別手法をもとに、軟



**図1 軟起声訓練システムの画面表示**、軟起声が発声できたかどうかを視覚的にフィードバックするシステムを作成した。図1に訓練システムのスクリーンショットを示す。練習者の発声を録音してその特徴量を算出し、プロットする。比較のため言語聴覚士による軟起声および硬起声の発声の範囲を多角形で図示し、練習者の発声がどの位置にあるかがわかるようにしている。作成した視覚提示により、短期的な軟起声の訓練の実験を行った。協力者は吃音者10名、非吃音者8名である。実験では、12-13単語または単独母音を8回ずつ練習させた。(4) 提案言語訓練システムによる長期訓練吃音者1名について長期的な訓練効果を調べた。吃音は中等度で本訓練システム使用前に3か月言語訓練を受け軟起声の練習に取り組んでいる。研究協力期間はベースラインとして訓練システムの導入前の1か月、システム導入後の1か月である。システム導入後は訓練システムを1日10分弱使用した。図2に示す時期に検査を3回実施した。



**図2 言語訓練のスケジュール**

## 4. 研究成果

### (1) 軟起声の自動評価

図3、図4にそれぞれ母音/a/の硬起声、軟起声、長い軟起声の波形および起声部分の

周波数スペクトルの例を示す。軟起声では発声開始直後に振幅が大きくなる前に

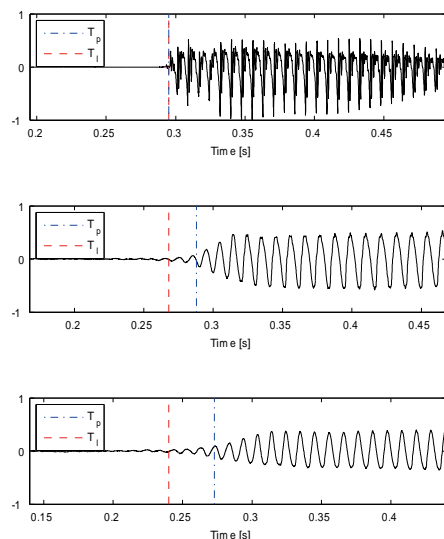


図 3 硬起声(上)、軟起声(中)、長い軟起声(下)の音声波形

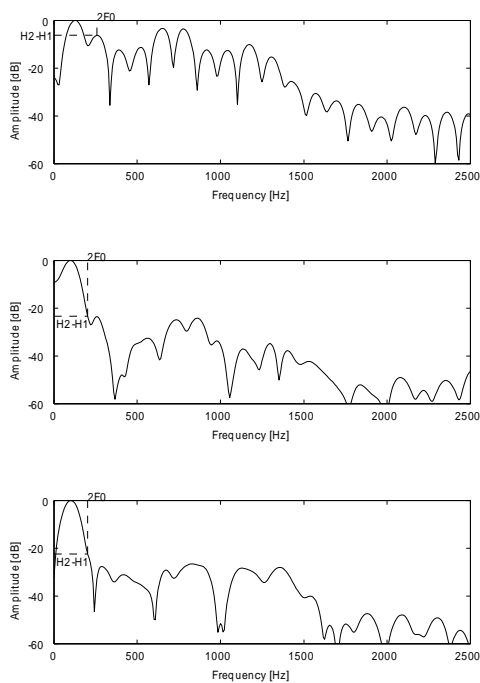


図 4 硬起声(上)、軟起声(中)、長い軟起声(下)のスペクトル

sin 波に近い波形が見られる。また、図 4 では硬起声に比べ軟起声で H1-H2 の値が大きい。

図 5 に硬起声、軟起声、長い軟起声の提案する特徴量の分布を示す。軟起声と硬起声が分かれて分布していることがわかる。

図 6 に従来法および提案手法による誤判別

率を示す。RT1、RT2 はそれぞれ Peters ら (1986) および Koike (1967) による立ち上がり

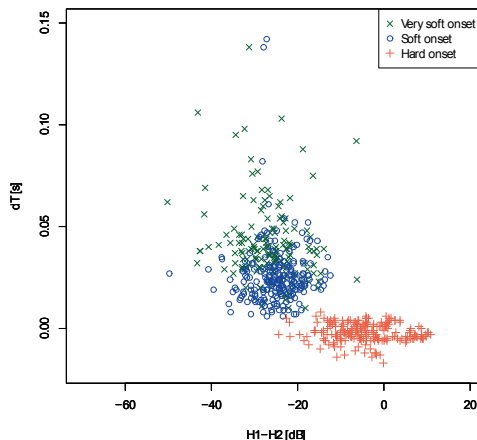


図 5 特徴量の分布

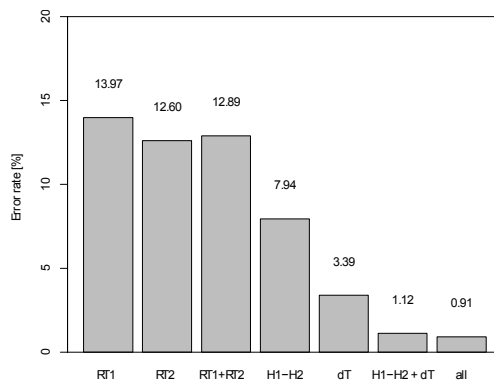


図 6 判別結果

りにかかる時間を自動抽出できるよう改変したものである。提案する特徴量(H1-H2 + dT)のみで従来法に比べ低く 1%程度の誤判別率を得た。また、この特徴量は話者の性別によらず判別可能である。

(2) 吃音者・非吃音者の軟起声の比較

提案する軟起声の判別特徴量を算出したところ、軟起声の指示のもとでの発声の半数以上が言語聴覚士による硬起声の範囲に入った人数は、非吃音者では 9 名中 0 名であるのに対し、吃音者では 8 名中 4 名であり吃音者のほうが有意に硬起声で発声した者が多かった(χ<sup>2</sup>検定,  $p < 0.05$ )。吃音者については短期的な軟起声の導入では硬起声での発声が残ることがあるため、その場合はより継続的な訓練が必要であることが示唆される。

(3) 視覚的フィードバックの短期訓練

1 セッション目ではフィードバックなしで軟起声の文章での説明を読んで発声し、2

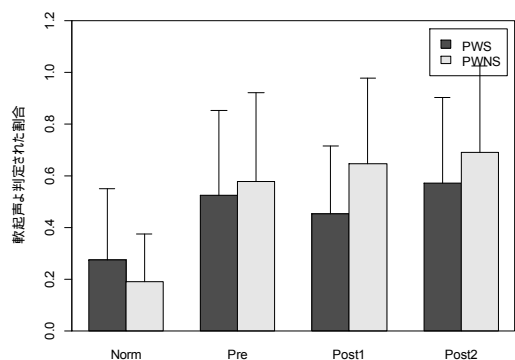


図 7 短期訓練の結果

セッション目はフィードバックありで発声した。最初に普段の発声を録音し、その後セッション1の前後、セッション2の後にそれぞれ単語または単独母音の発声を録音し、それぞれ、Norm、Pre、Post1、Post2と呼ぶことにする。図7に各テストでの発声が軟起声と判定された割合を示す。ここで軟起声の判定は、(1)で用いた言語聴覚士の発声を学習データとした線形判別によりおこなった。テストと吃音・非吃音の2群の2要因分散分析を行った結果、テストのみに有意な主効果を認め ( $p < 0.05$ )、下位検定の結果 Norm とその他のテストの間に有意差を認めた ( $p < 0.05$ )。被験者別に見ると Post1 で Pre より軟起声の割合が増えていたのは吃音者では 10 名 5 名、非吃音者で 8 名中 5 名であった。短期的なフィードバックでは不十分な被検者については継続的な訓練が必要となると考えられる。

#### (4) 長期訓練の効果の検証

図8にベースラインおよびシステム導入前後の検査で発声した普段の発声、すなわち軟起声の指示がない発声が軟起声に判別された割合を示す。また、図9に訓練中の軟起声および普段の発声の構音速度を示す。意識的に軟起声を発声する場合はゆっくり発声しているが、長期訓練後は比較的早い

普段の発声でも軟起声ができる割合が増えていることを表している。したがって立ち

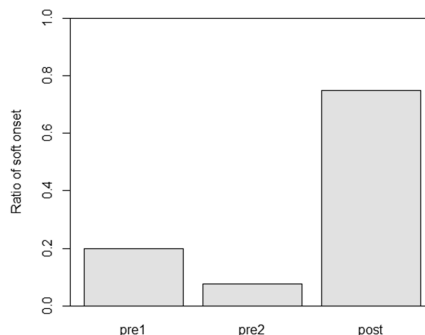


図 8 普段の発声の軟起声の割合

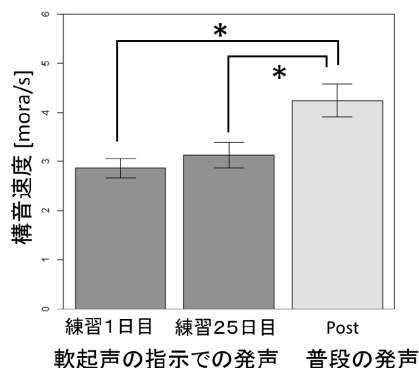


図 9 構音速度

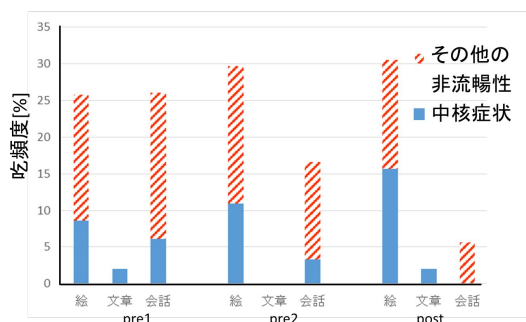


図 10 吃頻度

上がりの時間を基準にしたさいのような引き延ばした発声ではなく、短い時間で軟起声を発声できるようになったといえる。図10に3回の成人吃音検査時の吃頻度を示す。会話については吃頻度の低下が見られた。絵の説明についても検査1 (pre1) では阻止は8s以上や10s以上のものがみられたが、pre2では6s以下、postでは4s以下のみ見られ、症状が緩和された。

5. 主な発表論文等

〔雑誌論文〕

1. Keiko Ochi, Koichi Mori, Naomi Sakai, Jun Ogura, “Accuracy of Articulation Rate Control with Visual Feedback in Persons who do and do not Stutter,” *Procedia - Social and Behavioral Sciences*, Vol. 193, no. 30 pp. 217-222, 2015.

〔学会発表〕(計5件)

1. 越智景子, 酒井奈緒美, 森浩一, 他1名, “柔らかい起声の音響的特徴,” 日本音響学会春季講演論文集, 2014.
2. 越智景子, 森浩一, 酒井奈緒美, “軟起声の言語訓練システムの開発,” 第60回日本音声言語医学会総会・学術講演会, 名古屋, 2015-10-15/10-16, p.86.
3. Keiko Ochi, Koichi Mori, Naomi Sakai and Nobutaka Ono, “Automatic Discrimination of Soft Voice Onset Using Acoustic Features of Breathly Voicing,” *Proc. Interspeech*, pp. 267-271, 2016.
4. 吉澤健太郎, 越智景子, “やってみよう軟起声,” 日本吃音・流暢性障害学会第4回大会, 所沢, 2016-9-2/9-3.
5. 越智景子, 森浩一, 酒井奈緒美, 北條具仁, 小野順貴, “音響特徴を用いた軟起声自動評価システムの提案とこれに基づく長期訓練の一例,” 日本吃音・流暢性障害学会第4回大会, 所沢, 2016-9-2/9-3.

〔図書〕

なし

〔産業財産権〕

なし

〔その他〕

なし

6. 研究組織

(1)研究代表者

越智 景子 (OCHI, Keiko)

国立情報学研究所・情報学プリンシプル研究系・特任研究員

研究者番号：20623713

(2)研究分担者

なし

(3)連携研究者

なし

(4)研究協力者

森 浩一 (MORI, Kochi)

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：6015857

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・研究員

研究者番号：60415362