

原著

音声波形の振幅の包絡線を用いて音声の起声を評価する試み

青木 俊仁^{1),2)*}, 松坂 碧唯¹⁾, 宮崎 早梨¹⁾, 山本 祐未¹⁾, 渡邊 いつか¹⁾, 坂本 和也²⁾,
浅岡 拓希²⁾, 池 聡¹⁾, 吉村 知佐子¹⁾, 光内 梨佐¹⁾, 上松 智幸¹⁾, 稲田 勤¹⁾,
石川 裕治¹⁾, 宇高 二良²⁾, 世木 秀明³⁾

Pilot studies on evaluating vocal attack by the amplitude envelope of speech waveforms.

Toshihito Aoki, SLHT, PhD^{1),2)*}, Aoi Matsusaka¹⁾, Sari Miyazaki¹⁾, Yumi Yamamoto¹⁾, Itsuka Watanabe¹⁾,
Kazuya Sakamoto, SLHT²⁾, Hiroki Asaoka, SLHT²⁾, Satoshi Ike, SLHT, MS¹⁾, Chisako Yoshimura, SLHT, MS¹⁾,
Risa Mitsuuchi, SLHT, MS¹⁾, Tomoyuki Uematsu, SLHT, MS¹⁾, Tsutomu Inada, SLHT, MA¹⁾,
Yuji Ishikawa, SLHT¹⁾, Jiro Udaka, MD, PhD²⁾, Hideaki Seki, PhD³⁾

要 旨

小児の音声の起声の客観的かつ簡便な評価に資する目的で、音声波形の振幅の包絡線の立ち上がりの速度、オーバーシュート率をパラメータとして音声の起声の種類を区別できるかを検討した。対象は起声を意図的にコントロールできる健常成人とした。その結果、硬起声の立ち上がり速度とオーバーシュート率は、軟起声および気息性起声の立ち上がり速度とオーバーシュート率より有意に高いことが明らかとなった。これらのことから、音声波形の振幅の包絡線の立ち上がりの測定は、音声の起声の客観的かつ簡便な評価として活用できると考えられ、小児の音声の起声の評価にも活用できる可能性が示唆された。小児の音声の起声について、音声波形の振幅の包絡線の立ち上がり速度とオーバーシュート率の2つのパラメータで評価できるかを検討することが今後の課題である。

キーワード：音声波形の振幅の包絡線、音声の起声、音声の起声の評価

Abstract

We conducted experiments to determine whether a measure of the rise time of the amplitude envelope of speech waveforms could differentiate between vocal attacks. A hard attack had a higher rate of rise and overshoot compared to soft and breathy attacks. The findings confirm that measuring the rise time of the amplitude envelope of speech waveforms is an appropriate method for evaluating vocal attack. Evaluating vocal attack in children is needed to be addressed in future studies.

KeyWords: amplitude envelope of speech waveforms, vocal attack, evaluating vocal attack

1) 高知リハビリテーション専門職大学 リハビリテーション学部 リハビリテーション学科 言語聴覚学専攻
Division of Speech-Language-Hearing Therapy, Department of Rehabilitation, Faculty of Rehabilitation, Kochi Professional
University of Rehabilitation

2) 医療法人 真樹会 宇高耳鼻咽喉科医院
Udaka Otolaryngology Clinic

3) 千葉工業大学 情報工学科
Department of Information Engineering, Chiba Institute of Technology

*Corresponding author : aoki@kochireha.ac.jp

はじめに

起声には、発声時に声門閉鎖と呼気の流出がほぼ同時に生じる軟起声と発声前に声門を固く閉じてから呼気が流出し発声する硬起声、発声の開始時に摩擦音/h/を伴う気息性起声がある¹⁾。母音「あ」の発声でイメージすると、普通の「あ」が軟起声で、「っあ」のように発声前に促音を伴うものが硬起声、「はあ」と発声前に囁き声を伴うものが気息性起声である。起声は発声において最も重要な段階の一つに位置付けられており²⁾、Voice Therapyの一環として起声方法を硬起声から軟起声あるいは気息性起声に変容させる方法^{1~6)}や、反対に硬起声を促す方法⁶⁾などを用いることがある。

小児に発生する音声障害を小児嚁声という。小児嚁声の多くが硬起声の習慣をもっているとされており⁷⁾、小児嚁声を助長する要因の一つと考えられている。そのため、小児の音声障害の臨床において起声の評価は重要性が高い。起声の評価の方法には、聴覚心理的評価、発声時の声帯振動を電気的インピーダンスの変化として測定するエレクトログロトグラフ（以下、EGG）や高速度ビデオ内視鏡による声帯運動の撮影⁸⁾、呼気流量や呼気流量などの空気力学的検査²⁾などがあげられる。聴覚心理的評価は、主観的であり普遍性や妥当性に乏しく、再現性は低い。EGGや高速度ビデオ内視鏡検査、空気力学的検査はいずれも高額な機器を必要とする欠点がある。また、経鼻的にファイバーを挿入する内視鏡検査は、協調性や忍耐力に欠ける小児では実施困難なことが多い。そのため、ヨーロッパ喉頭学会のガイドラインでは、小児の音声障害の客観的評価として音響分析が推奨されている⁹⁾。

Koike²⁾は、音声の起声の立ち上がり時間を「発声開始から音声波形の振幅の包絡線が定常振幅に達するまでの時間」と定義し、音声の起声を音響分析的に評価する指標として提唱している。Petersら¹⁰⁾は、音声の起声の立ち上がり時間から算出した音声波形の振幅の包絡線の立ち上がりの傾斜を指標として音声の起声の評価が可能であることを明らかにしている。しかし、Petersらの方法は音声波形の振幅

の包絡線の立ち上がりの傾斜を算出するための基準点が容易に定められず⁸⁾、臨床への応用は困難である。

一般に、音声の起声における音声波形の振幅の包絡線は、発声開始時に最低点から立ち上がり、振幅が最大点にオーバーシュートした後に定常振幅に至る。音声波形の振幅の包絡線の最低点と最大点は音響分析ソフトにより自動的に求めることが可能である。この振幅の包絡線の最低点と最大点を基準点とすれば、音声波形の振幅の包絡線の立ち上がりの傾斜が容易に算出できる。また、音声波形の振幅の包絡線のオーバーシュートは硬起声ほど大きくなると予測されるが、音声波形の振幅の包絡線のオーバーシュートと音声の起声の種類との関連を検討した研究はない。このオーバーシュートも音響分析ソフトを利用すれば容易に算出可能である。音声の起声の種類と音声波形の振幅の包絡線の立ち上がりの傾斜およびオーバーシュートとの関連が明らかになれば、特に小児の音声障害の臨床への応用が期待できる。

本研究では、小児の音声の起声の客観的かつ簡便な評価に資する目的で、音声の起声における音声波形の振幅の包絡線の立ち上がりの傾斜とオーバーシュートと音声の起声の種類との関連を検討した。

方法

1. 対象

対象は、医療機関での音声障害の診断や既往のない健常成人で、A大学に通う学生52名（男性27名、女性25名、平均年齢 20.7 ± 2.4 歳）であった。本研究で対象を健常成人としたのは、起声の意図的なコントロールが小児では容易ではないためである。

本研究は、高知リハビリテーション専門職大学研究倫理審査委員会の承認を得て実施した（承認番号KPUR2024E0012）。参加者には、本研究の方法およびデータや個人情報の取り扱い等について書面および口頭で説明し、書面にて同意を得た。

2. 音声の録音およびEGGデータの記録方法

音声の録音およびEGGデータの記録はいずれも暗騒音15dB(A)の防音室内で行った。

音声の録音にはコンデンサーマイクロフォン NEEWER® NW-700を用いた。発声時に呼気がマイクロフォンに当たって生じるポップ音防止のためマイクロフォンから10cmの位置にポップガードを設置して音声を録音した。被験者に楽な高さで大きさでの軟起声「あ」および硬起声「っあ」、氣息性起声「はあ」の3種類の起声方法で約3秒間発声させ、それぞれ3回録音した。いずれの発声方法も事前に練習を行い、それぞれの発声が可能となったことを確認して録音を開始した。

EGGデータの記録はGlottal Enterprises社製のEG2-PCX2を用いて記録した。被験者の頸部の甲状軟骨上部の左右の位置に電極を設置し、音声の録音と同時にEGGデータを記録した。

音声およびEGGデータは、マイクロフォンおよびEG2-PCX2をRoland製のオーディオインターフェースRubix24でNEC製パーソナルコンピュータN1475 CALと接続し、サンプリング周波数44.1kHz、量子化ビット数16bitで同時に記録した。音声およびEGGデータの同時記録には、Glottal Enterprises社製の声門電位計波形補正および声門開閉度測定用信号処理プログラムソフトPhaseComp1.3.9を用いた。

3. 音声およびEGGデータの解析による起声の種類判定方法

音声およびEGGデータの解析には、音響分析用フリーソフトPraat (version 6.4.22) (Boersma & Weenink, University of Amsterdam) を用いた。EGGデータは声帯振動に関わる信号のみを抽出するため、100Hzのハイパスフィルターおよびスムージング処理を行った。音声波形およびEGGの声門閉鎖状況の解析結果から、検査者に指示された軟起声、硬起声、氣息性起声、それぞれの起声方法で被験者が発声できているかどうかを判定した。起声の種類のそれぞれの判定方法を図1に示した。発声前の呼気漏出および声門閉鎖が認められない場合は軟起声であり（図1-a.）、発声前の呼気漏出がなく、声門閉鎖が認められる場合は硬起声（図1-b.）、発声前の呼気漏出はあるが声門閉鎖は認められない場合は氣息性起声（図1-c.）である。検

査者に指示された正しい起声方法で発声できていないと判定された音声は集計から除外した。

4. 音声波形の振幅の包絡線を用いた音声の起声の解析方法と音声の起声を評価するパラメータ

音声波形の振幅の包絡線の描出および解析は、Praatを用いて行った。解析には各音声の起声の種類の3回のデータのうち2回の平均値を用いた。なお、EGG波形のアーチファクトなど不備の認められたデータは集計から除外した。

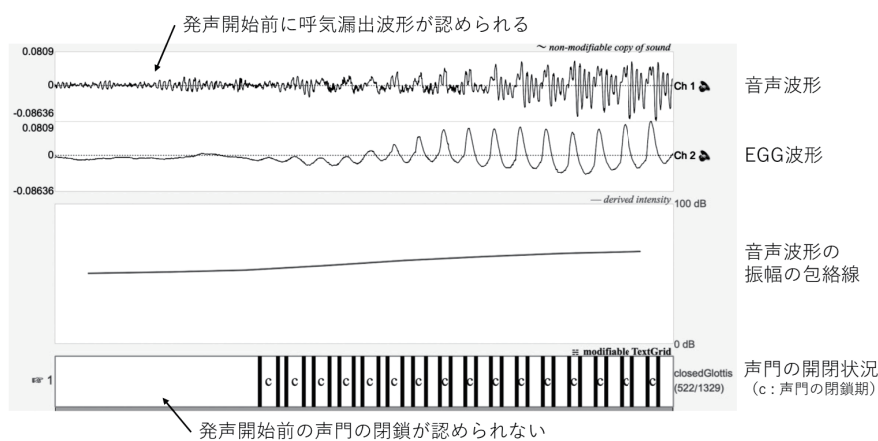
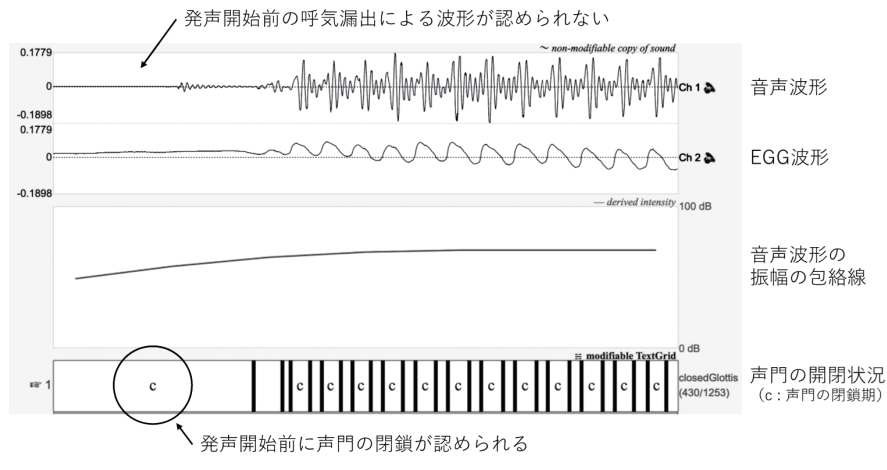
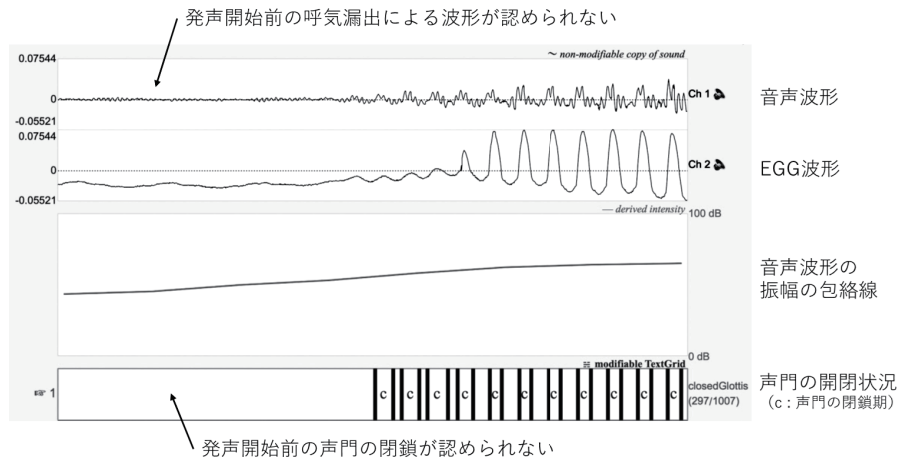
Praatで描出された振幅の包絡線（図2）の情報から、振幅の最小値（Pmin）とその時間、振幅の最大値（Pmax）とその時間、振幅の定常状態の平均値（Pstd）を解析し、算出した。また、Pminの時間およびPmaxの時間の差、すなわちPminからPmaxに至るまでの時間（Ptime）を算出した。さらに、PmaxとPstdの振幅差（Pmax-Pstd）を求めた（図3）。これらの解析結果から、以下の数式を用いて、①どれだけ素早く起声されたかの指標となる立ち上がり速度、②発声が定常状態となる前の余分な音圧上昇の指標となるオーバーシュート率を算出し、音声の起声のパラメータとした。いずれのパラメータも大きくなるほど硬起声に近づくことと予測される。

$$\frac{P_{\max}-P_{\min}}{P_{\text{time}}} \text{ (dB/msec.)} \cdots \cdots \text{① 立ち上がり速度}$$

$$\frac{(P_{\max}-P_{\text{std}})}{P_{\text{std}}} \times 100 (\%) \cdots \cdots \text{② オーバーシュート率}$$

5. 統計学的分析方法

正しい起声方法で発声できていなかったり、不備の認められたデータを除外し、20名（男性6名、女性14名、平均年齢20.9±1.9歳）が分析対象となった。独立変数を音声の起声の種類、すなわち軟起声、硬起声、氣息性起声とし、従属変数を各音声の起声のパラメータ、すなわち立ち上がり速度およびオーバーシュート率としてFriedman検定を用いて比較した。多重比較にはSteel-Dwass法を用いた。統計解析はStatcel 4¹¹⁾を用いて行い、危険率5%を有意水準とした。



a. 軟起声：発声開始前の呼気漏出波形，声門閉鎖ともに認められない． b. 硬起声：発声開始前の呼気漏出波形は認められないが，声門閉鎖が認められる． c. 氣息性起声：発声開始前の呼気漏出波形は認められるが，声門閉鎖は認められない．

図1 起声の種類別のPraatによる音声波形とEGG波形

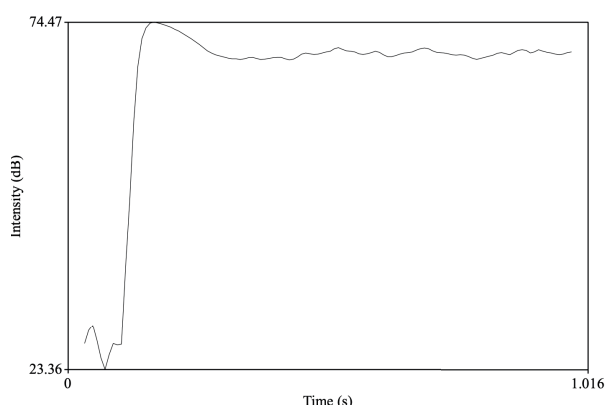


図2 Praatで描出した音声の振幅の包絡線の例（硬起声）

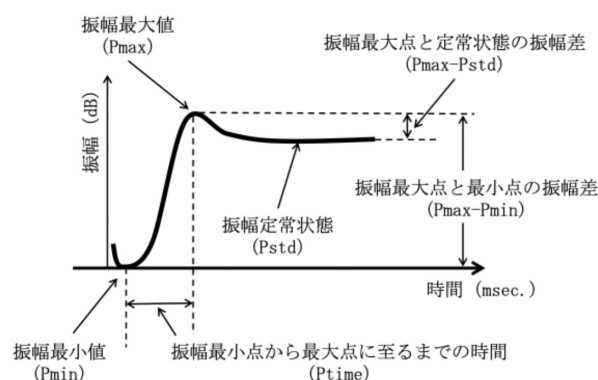


図3 音声の起声部分の振幅の包絡線の変化モデル

結果

音声の起声の種類別の記述統計量を表1に示した。Pmin(dB)およびPtime(msec.)は硬起声が最も低値であり、Pmax(dB)およびPstd(dB)は硬起声が最も高値であった。

立ち上がり速度の中央値[四分位範囲]は、軟起声が106.14[84.79-136.31](dB/msec.)、硬起声が326.51[288.17-509.58](dB/msec.)、気息性起声が86.60[61.28-113.99](dB/msec.)であり、硬起声が軟起声および気息性起声に比し有意に高かった($p < 0.01$) (図4)。オーバーシュート率の中央値[四分位範囲]は、軟起声が2.27[0.64-2.91](%)、硬起声が7.48[4.28-9.77](%)、気息性起声が3.88[2.13-7.26](%)であり、硬起声が軟起声に比し有意に高く($p < 0.01$)、気息性起声も軟起声に比し有意に高かった($p < 0.05$) (図5)。

考察

1. 音声波形の振幅の包絡線を用いた音声の起声の評価の是非

先に述べたように、音声の起声の評価を音声波形の振幅の包絡線の立ち上がりの傾斜を用いて行う試みは、Petersら¹⁰⁾によって既に行われている。Petersら¹⁰⁾の方法は、音声の起声部分には立ち上がりの初期・中期・後期に3つの傾斜があり、3つの傾斜の交点を基準点として中期の傾斜を求めるものである (図6-a.)。しかし、この方法における音声の起声の立ち上がりの傾斜の基準点を求めるのは容易ではな

く、複雑な操作が必要となることが指摘されている⁸⁾。そのため、Orlikoffら⁸⁾は、高速度ビデオ内視鏡検査とEGGを組み合わせでデジタルキモグラムを作成する方法を提案している。しかしながら、この方法では高額な機器を要すること、一定の苦痛を伴う経鼻的な内視鏡検査は小児の音声障害の臨床への応用には適さないことなどのデメリットがある。本研究の方法では、傾斜を算出するための基準点PminおよびPmaxは一意に定めることが可能 (図6-b.)であり、Praatなどの音響分析ソフトを用いれば自動的に測定することができる。音声を録音することができれば容易に測定が可能であるため、小児の音声障害の臨床への応用に適していると考えられる。ただし、音声波形の振幅の包絡線の立ち上がりの傾斜の算出における、本研究の方法とPetersら¹⁰⁾の方法の違いが音声の起声の評価の精度にどのような影響を与えるかについて、今後検討する必要がある。

2. 音声の起声の評価する指標としての音声の起声の立ち上がり速度の有用性

本研究において、音声の起声の種類と音声の振幅の包絡線を用いた音声の起声の立ち上がり速度との関連を検討した結果、硬起声の立ち上がり速度は、軟起声および気息性起声の立ち上がり速度に比し高く、硬起声は軟起声や気息性起声よりも音声の起声の立ち上がりの傾斜が急峻であることが明らかとなった。Koike²⁾は、音声の起声は立ち上がり時間の

表1 音声の起声の種類別の記述統計量

起声の種類	Pmin(dB)	Pmax(dB)	Pstd(dB)	Ptime(msec.)
軟起声	27.68 (6.05)	67.14 (5.30)	66.12 (5.25)	0.45 (0.24)
硬起声	22.42 (2.86)	76.33 (3.87)	70.77 (4.43)	0.17 (0.07)
気息性起声	27.51 (5.29)	72.45 (5.06)	69.28 (5.11)	0.69 (0.38)

数値の下の () 内の数値は標準偏差を表す

Pmin : 振幅最小値, Pmax : 振幅最大値, Pstd : 振幅定常状態の平均値,

Ptime : 振幅最小点から最大点に至るまでの時間

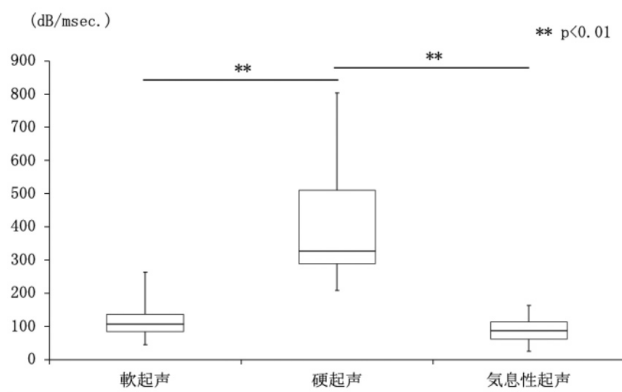


図4 起声の仕方が音声の起声の立ち上がり速度に与える影響

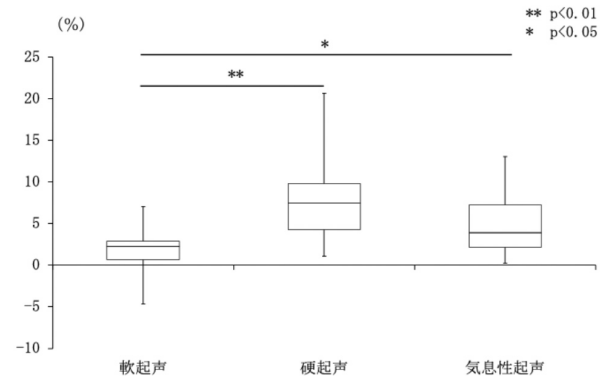
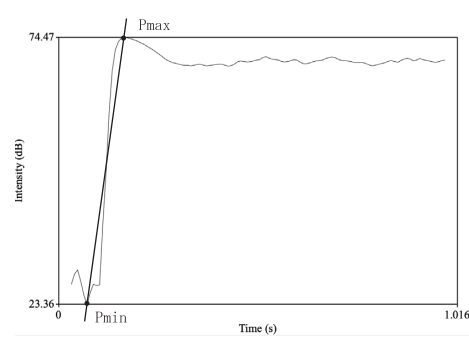
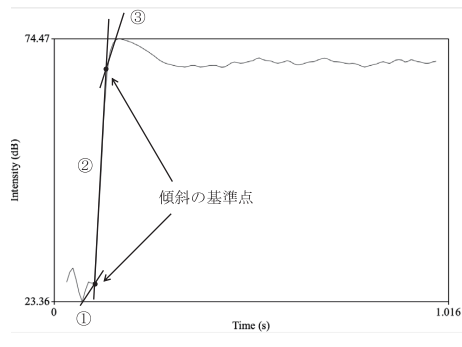


図5 起声の仕方が音声の起声のオーバーシュート率に与える影響



a. Petersら（1986）の方法：音声の起声の包絡線の3つの傾斜の交点を基準点とし、②の傾斜を求める。b. 本研究の方法：PmaxとPminを基準点として音声の起声の包絡線の傾斜を求める。

図6 Petersら（1986）と本研究の音声の起声の振幅の包絡線の傾斜算出の基準点

長さによって特徴づけられ、軟起声は音声の起声の立ち上がり時間が長く、硬起声は音声の起声の立ち上がり時間が短くなることを明らかにしている。本研究の結果は、硬起声が軟起声や気息性起声に比べて速く立ち上がることを示しており、Koike²⁾の結果と一致する。これらのことから、本研究の音声の起

声の立ち上がり速度は硬起声を軟起声および気息性起声と区別する指標として有用であると考えられた。

3. 音声の起声を評価する指標としての音声の起声のオーバーシュート率の有用性

本研究の結果、硬起声のオーバーシュート率は、

軟起声のオーバーシュート率に比し高く、硬起声は軟起声よりも定常状態になる前の余分な音圧上昇が大きいことが明らかとなった。硬起声は発声前に声門が固く閉じてから呼気の流出が生じ¹⁾、爆発的に呼気流量を増加させた発声となる²⁾ため、音声の過度な音圧上昇を伴うものと考えられる。また、気息性起声のオーバーシュート率は、軟起声のオーバーシュート率に比し高く、気息性起声も軟起声に比べ定常状態になる前の余分な音圧上昇が認められた。気息性起声と軟起声の音声の立ち上がりの傾斜は、起声直後はほぼ同じであるが、途中から気息性起声が軟起声よりも増加することが報告されている²⁾。そのため、気息性起声のオーバーシュート率が軟起声のオーバーシュート率よりも増加した可能性が考えられる。

これらのことから、本研究の音声の起声のオーバーシュート率は起声の種類を区別する指標として有用と考えられた。

まとめ

小児の音声の起声の客観的かつ簡便な評価に資する目的で、音声波形およびEGGデータから軟起声、硬起声、気息性起声に分類し、それぞれの音声波形の振幅の包絡線から、音声の起声の立ち上がり速度およびオーバーシュート率を算出し、両者の関連を比較した。対象は起声を意図的にコントロールできる健常成人とした。その結果、音声の起声のパラメータである立ち上がり速度とオーバーシュート率は、いずれも音声の起声の種類を区別する指標として有用であることが明らかとなった。これらのことから、音声波形の振幅の包絡線の立ち上がりの測定は、音声の起声の客観的かつ簡便な評価として活用可能と考えられ、小児の音声の起声の評価にも活用できる可能性が示唆された。小児の音声の起声について、音声波形の振幅の包絡線の立ち上がり速度とオーバーシュート率の2つのパラメータで評価できるかを検討することが今後の課題である。

文献

- 1) 小池靖夫：音声言語医学領域の用語とその解説 (Ⅱ). 音声言語医 31(4) : 427-430, 1990.
- 2) Koike Y: Experimental Studies on Vocal Attack. 耳鼻咽喉科臨床 60(8) : 663-688, 1967.
- 3) 新美成二：小児嗄声の取り扱い. 廣瀬肇(編) : 耳鼻咽喉科・頭頸部外科 MOOK 4. 金原出版, 東京, 1987, pp10-18.
- 4) 渡辺陽子, ダイアン BM, 三浦真弓 : 声帯結節に対するVoice Therapy. 耳鼻と臨 29(6) : 949-952, 1983.
- 5) 丘村熙：小児結節の治療方針. 耳鼻と臨 29(6) : 956-957, 1983.
- 6) 三瀬和代：音声治療の基本的考えと実際. 喉頭 30(2) : 67-71, 2018.
- 7) 前川彦右エ門, 長谷川寛治, 伊藤督夫・他 : 学童嗄声の経時的観察-小児嗄声の予後とその取り扱い-. 耳鼻臨床 68(12) : 1487-1511, 1975.
- 8) Orlikoff RF, Deliyski DD, Baken RJ, et al: Validation of a Glottographic Measure of Vocal Attack. J Voice 23(2) : 164-168, 2009.
- 9) Dejonckere PH, Bradley P, Clemente P, et al: A basic protocol for functional assessment of voice pathology, especially for investigating the efficacy of (phonosurgical) treatments and evaluating new assessment techniques. Guideline elaborated by the Committee on Phoniatrics of the European Laryngological society (ELS). Eur Arch Otorhinolaryngol 258(2) : 77-82, 2001.
- 10) Peters HFM, Boves L, van Dielen ICH: Perceptual judgment of the abruptness of voice onset in vowels as a function of the amplitude envelope. J Speech Hear Disord 51(4) : 299-308, 1986.
- 11) 柳井久江 : 4 Steps エクセル統計 第4版, オーエムエス出版, 埼玉, 2015.