

# 嚥下内視鏡検査におけるスコア評価基準によって分類された 嚥下音の特徴比較

Comparison of features of swallowing sounds  
classified by the score for videoendoscopic evaluation

杉山 翔<sup>†</sup>  
Sho Sugiyama

小澤 恵子<sup>‡</sup>  
Keiko Kozawa

森谷 季吉<sup>§</sup>  
Sueyoshi Moritani  
宮城 茂幸<sup>‡</sup>  
Shigeyuki Miyagi

坂本 眞一<sup>‡</sup>  
Shin-ichi Sakamoto

酒井 道<sup>‡</sup>  
Osamu Sakai

## 1 はじめに

肺炎は長らく日本の疾患別死亡の第 4 位であったが、厚生労働省の 2011 年の報告によると、脳血管疾患を抜き第 3 位となった [1]。肺炎による死亡者の中で高齢者が占める割合は極めて高く、また、高齢者の肺炎の約 70% 以上が誤嚥性肺炎であると言われている [2][3]。

誤嚥性肺炎のリスクを低減するためには嚥下機能評価が必要である。スクリーニングのための手法として、唾液を飲み込む動作を 30 秒間繰り返し行う反復唾液嚥下テストや、冷水 3ml を嚥下する改訂水飲みテストのプロトコルが標準化されている。改訂水飲みテストでは、むせがあるかどうかあるいは湿性嚔声の有無の確認を行う。実際の嚥下時の咽頭の状態や喉頭の動きを確認するために、内視鏡を鼻から挿入して、直視下で嚥下状態を見る嚥下内視鏡検査 (videoendoscopic examination of swallowing, VE) や、X 線透視下で造影剤を飲み込ませ、透視画像によって嚥下状態を確認する嚥下造影検査 (Videofluoroscopic examination of swallowing, VF) が知られている。これらは医療機関における限られた状況下での検査であって、日常的に利用できない。

一方、スクリーニングの一手法として知られている頸部聴診法は、ベッドサイドで簡単に実施することができるので、今後の高齢化社会において、在宅治療やグループホーム等での利用が期待できる。しかし、頸部聴診法によるスクリーニングのためには、専門医や言語聴覚士といった熟達した人が検査を実施する必要がある。

これまでに、頸部聴診法における客観的評価のための研究が行われてきた。高橋らの研究 [4] では、嚥下機能の評価の重要な手掛かりとなる嚥下音を検出するために、加速度トランスデューサを使用している。加速度トランスデューサを用いた研究では、健常者が食塊を嚥下する際の産生音を検出し、嚥下音の産生部位や、食塊が舌根から食道入口部分を通過するのに所要する時間が明らかにされている [5]。さらに、嚥下直後の呼吸音に着目することで、液体振動音、嗽音、むせを評価指標とする主観評価が、妥当かつ有効であることなども明らかにされて

いる [6]。そのほかに、嚥下音を聴診器を用いて採取した際に、嚥下性無呼吸の期間に主たる嚥下音が発生していることや、その前後の嚥下性無呼吸の期間がどの程度なのか、若者、高齢者、嚥下障がい者の 3 群に対して、それぞれ明らかにしている研究報告もある [7]。しかし、医師による所見と頸部聴診音との対照はまだ十分ではなく、客観的な評価法に発展させるためには、今後共対照研究を蓄積する必要がある [8]。

そこで、本研究では兵頭らによる嚥下内視鏡スコア評価基準 [9] と、嚥下性無呼吸時間前後における嚥下音の基本的な特徴量との対応関係を調べた。Borr らの研究 [7] では若年者、高齢者、嚥下障がい者の各群に対して、嚥下聴診音の時間波形に関連する各種パラメータの平均値が示されている。特に嚥下障がい者の嚥下性無呼吸時間は若年者と比較すると 1.74 倍、高齢者と比較すると 1.16 倍である。このことから嚥下性無呼吸時間を中心とした嚥下音の特徴に障がいの程度による変化が現れるのではないかと考えた。あらかじめ医師によりスコア付けされた嚥下障がい者を、障がいの程度が高い群、中程度の群、低い群の 3 群に分類し、それらの各群および健常者群に対し、嚥下性無呼吸時間を含む嚥下音の特徴を抽出した。特徴量としては、後述の通り嚥下音のスペクトルから得られる諸量を用いた。各特徴量について、各群間で有意差があるかどうか多重検定 [10] を行った結果を報告する。

## 2 実験方法

ここでは具体的な嚥下音の採取方法およびそれらの解析方法について述べる。

### 2.1 対象者

まず、本研究を遂行するにあたり、本学の倫理委員会の承認を得た。対象は本研究への参加に同意した健常な男性 3 人 (21 - 47 歳)、女性 6 人 (21 - 22 歳) および嚥下障がいのある男性 9 人 (60 - 98 歳)、女性 17 人 (65 - 97 歳) の計 35 名である。嚥下障がいのある男性および女性は、いずれも社会医療法人誠光会草津総合病院に入院中であって VE による嚥下機能評価を受けている。その結果から、兵頭らの「嚥下内視鏡検査のスコア評価基準」 [9] に基き、医師により嚥下機能がスコア化されて

<sup>†</sup>滋賀県立大学大学院, Graduate school of The University of Shiga Prefecture

<sup>‡</sup>滋賀県立大学, The University of Shiga Prefecture

<sup>§</sup>草津総合病院, Kusatsu General Hospital

いる。4 項目におけるスコアの合計値から表 1 に従って対象者を群分けした。なお、健常者については合計スコアを 0 とみなし A 群に分類した。

表 1 群分け基準

群	合計スコア	人数	L(ms)
A	0	9	700
B	1 ~ 4	8	800
B	5 ~ 9	13	800
C	10 ~ 12	5	800

## 2.2 嚥下音の取得方法

嚥下音の記録は市販の咽喉マイク (SH-12iK) (図 1) とノートパソコン (OS : Windows 7 Professional 32bit, CPU : Intel(R) Core(TM) i3-2370M, RAM : 4.00GB) を用いて行った。その際、サウンド編集ソフトウェアは Audacity (Version : 2.1.0) を使用し、サンプリングレート  $F_s$  を 8000Hz, 録音ゲインを 0.7 に設定した。被験者の輪状軟骨直下気管外側皮膚上に咽喉マイクを設置し (図 2), 被験食を嚥下させて産生する嚥下音を検出した。検出した嚥下音は wav ファイル形式としてコンピュータ内に保存した。また、被験食の種類は唾液、ゼリー、とろみ、水の 4 種類とし、唾液を除いた全ての被験食はスプーンを用いて口に運んだ。



図 1 咽喉マイク



図 2 咽喉マイクの位置

## 2.3 嚥下音の切り出し

中山らの研究 [5] によると、嚥下音の主たる産生は食塊が喉頭蓋を通過するときである。また、食塊が舌骨から食道通過までにかかる時間は、健常者で平均 649.4msec であることが示されている。Borr らの研究 [7] では、嚥下障がい者の嚥下性無呼吸時間は、老年者と比較すると約 1.16 倍であることが示されているので、嚥下障がい者の食塊通過時間は 753.3msec 程度と推測される。従って嚥下性無呼吸時間に産生される嚥下音を切り出すためには、健常者に対して 700msec, 嚥下障がい者に対して 800msec の区間長を設定すれば十分であると考えた。

さて、あらかじめ予備実験として、11 名の嚥下障がい者の嚥下性無呼吸時間と、嚥下音の最大値発生場所との関連性を調べた。VE 映像よりホワイトアウトが発生する時刻 ( $t_1$ ) と呼吸再開 ( $t_2$ ) の時刻を求め、 $t_1$  と  $t_2$  の間において嚥下音が最大となる時刻  $t_m$  を調べた。その結果およそ  $(t_1 - t_m) : (t_m - t_2) = 1 : 1$  であった。

嚥下障がい者に対して嚥下音を取得する際には、同時に VE 映像を記録できるので、上述の方法をそのまま適用でき、ホワイトアウトの発生 ( $t_1$ ) と呼吸再開 ( $t_2$ ) との

間に嚥下音が最大となる時刻の前後で、区間長 800msec となるように嚥下音の切り出しを行った。

健常者の嚥下音取得時は VE 映像がないが、嚥下の動作を意識的に制御できることから、嚥下を行った時刻がほぼわかる。そこで、その前後で嚥下音が最大となる位置を中心として、区間長が 700msec となるように切り出しを行った。

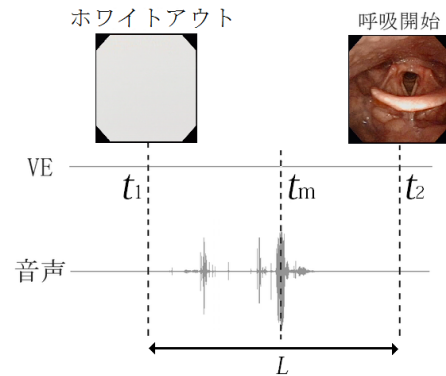


図 3 音声データの切り出し

## 2.4 特徴抽出

切り出された  $N$  個のデータを  $x_N[n]$  ( $n = 0, 1, \dots, N-1$ ) と書く。  $x_N[n]$  の離散時間フーリエ級数

$$X_N[k] = \sum_{n=0}^{N-1} x_N[n] e^{-j \frac{2\pi k}{N} n} \quad (k = 0, 1, \dots, N-1) \quad (1)$$

を求めた。そして、  $X_N[k]$  から振幅スペクトル  $|X_N[k]|$  を計算し、周波数領域における以下の 4 つの特徴量を抽出した。

振幅スペクトルの最大値

$$A_{\max} = \max |X_N[k]| \quad (k = 0, 1, \dots, N/2 - 1) \quad (2)$$

ピーク周波数

振幅スペクトルが最大値をとる時の周波数

$$f_{\max} = F_s \frac{2\pi}{N} \arg \max_k |X_N[k]| \quad (3)$$

ただし、  $F_s$ : サンプリングレート

周波数の加重平均

振幅スペクトル  $|X_N[k]|$  ( $k = 0, 1, \dots, [N/2]$ ) を度数分布とみなし、周波数の平均を以下の式 (4) で求めた。ただし、  $f_k = F_s \frac{2\pi}{N} k$

$$\bar{f} = \sum_{k=0}^{[N/2]} \frac{f_k |X_N[k]|}{\sum_{k=0}^{N-1} |X_N[k]|} \quad (4)$$

周波数の標準偏差

同様に、周波数の分散を以下の式 (5) で求め、周波数の標準偏差  $\sigma_f$  を特徴量とした。

$$\sigma_f^2 = \sum_{k=0}^{[N/2]} \frac{(f_k - \bar{f}) |X_N[k]|}{\sum_{k=0}^{N-1} |X_N[k]|} \quad (5)$$

2.5 検定手法

全ての群の中から 2 群ずつ選び、抽出した各種特徴量の平均に対して有意差検定を行った [10]。まず、ある特徴量について、全ての標本に対して正規性検定を行った。次に、選んだ 2 群の両方に正規性が認められている場合は、Welch の t 検定を行い、1 つでも認められていない場合は、Mann-Whitney の U 検定を行った。そして、4 群の多重比較を行うため、Holm 法を用いて全ての p 値を補正し、その上で有意水準を  $\alpha = 0.01$  および  $\alpha = 0.05$  として検定を行った。以上の手順で、各種特徴量において検定を行った。

3 結果

まず、各群における特徴量の平均値を表 2 に示した。そして、各種特徴量と合計スコアとの関係を図 4~ 図 7 の散布図に示し、それぞれの特徴量における有意差検定の結果を表 3~ 表 6 に示した。 $A_{max}$  については、 $\alpha = 0.05$  としても有意差が認められた組み合わせはなく、合計スコアとの相関係数は 0.31 となった。 $f_{max}$  については、 $\alpha = 0.01$  とすると A-B 間および A-C 間、さらに  $\alpha = 0.05$  とすると B-D 間においても有意差が認められた。合計スコアとの相関係数は 0.17 となった。 $\bar{f}$  については、 $\alpha = 0.01$  とすると C-D 間、さらに  $\alpha = 0.05$  とすると B-D 間においても有意差が認められた。合計スコアとの相関係数は -0.07 となった。 $\sigma_f$  については、 $\alpha = 0.01$  とすると C-D 間および B-C 間、さらに  $\alpha = 0.05$  とすると A-D 間、A-B 間、A-C 間においても有意差が認められた。合計スコアとの相関係数は -0.04 となった。

表 2 各群における各種特徴量の平均値

	$A_{max}$	$f_{max}$ (Hz)	$\bar{f}$ (Hz)	$\sigma_f$
A 群	6.50	1164	1451	800
B 群	10.19	1519	1496	706
C 群	20.32	1486	1602	1062
D 群	20.13	1253	1299	688

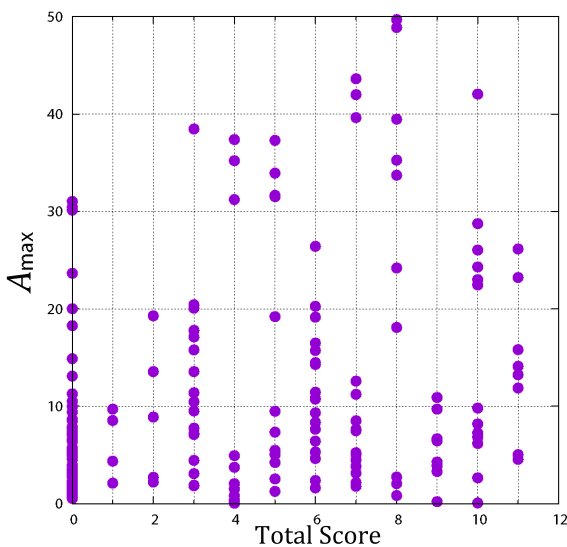


図 4  $A_{max}$  と合計スコア

表 3  $A_{max}$  における検定結果

2 群	p 値	補正 p 値
B-D	0.009572	0.057432
B-C	0.013371	0.066855
A-C	0.283037	1.132148
A-D	0.364596	1.093788
A-B	0.396598	0.793196
C-D	0.440421	0.440421

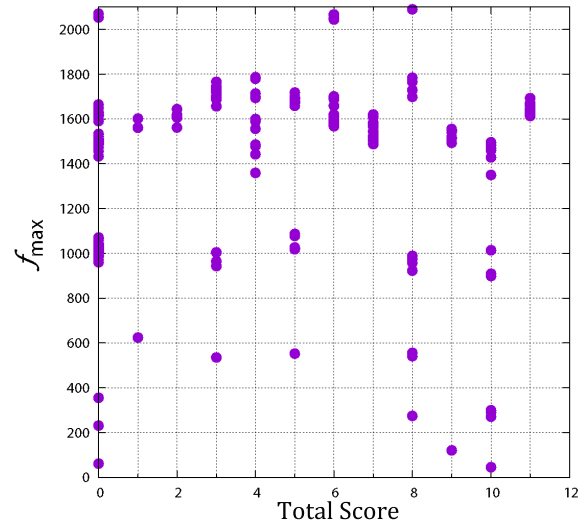


図 5  $f_{max}$  と合計スコア

表 4  $f_{max}$  における検定結果

2 群	p 値	補正 p 値
A-B	0.000014	0.000084
A-C	0.000027	0.000135
B-D	0.009902	0.039608
C-D	0.024831	0.074493
B-C	0.317228	0.634456
A-D	0.486964	0.486964

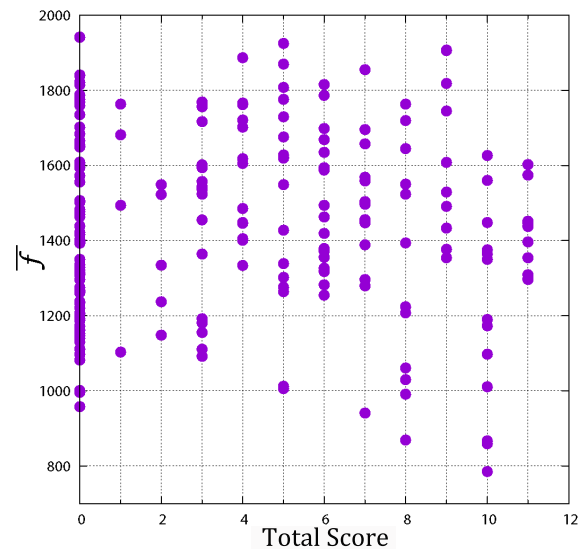
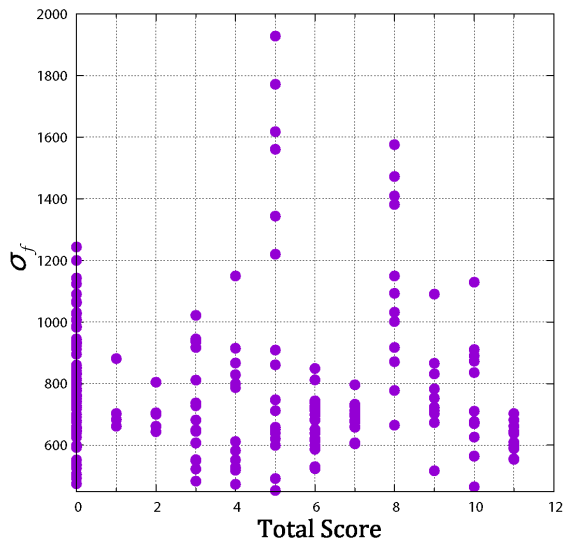


図 6  $\bar{f}$  と合計スコア

表 5  $\bar{f}$  における検定結果

2 群	p 値	補正 p 値
C-D	0.00066	0.00398
B-D	0.00242	0.01208
B-C	0.01337	0.05348
A-D	0.01394	0.04181
A-C	0.05384	0.10768
A-B	0.39660	0.39660

図 7  $\sigma_f$  と合計スコア表 6  $\sigma$  における検定結果

2 群	p 値	補正 p 値
C-D	0.00118	0.00708
B-C	0.00175	0.00875
A-D	0.00434	0.01736
A-B	0.00526	0.01577
A-C	0.01822	0.03645
B-D	0.66154	0.66154

#### 4 考察

上述したように、今回用いた特徴量は、合計スコアとの間に大きな相関が見られなかった。その一因として、嚥下によって発生した音声以外の音声が多く含まれていることが考えられる。本研究では、健常者の嚥下音を 700ms で切り出し、嚥下障がい者の嚥下音を 800ms で切り出したが、嚥下障がい者に関しては、障がいの程度毎に区間長を定めていない。区間長を障がいの程度別に定めることで、嚥下音をより正確に捉えることができると考えられる。また、検査と同時に嚥下音の録音を行っており、音声信号には多くの雑音が含まれていることが推測できる。雑音としては検査中の周りの人の声や、皮膚との摩擦によって生じる音等が考えられ、それらに対して、ノイズ除去等の前処理を行う必要があると言える。また、嚥下音の採取を行う際に用いている咽喉マイクが、検出した音声信号のある周波数帯の成分を強調している可能性があるため、今後はマイクの特性を考慮する必要があると考えられる。

#### 5 おわりに

本研究では、嚥下音の採取を行った被験者に対し、嚥下障がいの程度による群分けを行い、各群において得られた特徴量の平均を比較した。今回用いた特徴量は、いずれも合計スコアとの間に大きな相関は見られなかった。有意水準  $\alpha = 0.01$ ,  $\alpha = 0.05$  において、特徴量の平均に有意差のある組み合わせが散見されたため、嚥下障がいの程度による識別の可能性があることが示唆された。

#### 謝辞

本研究においてデータ収集に協力いただいた、社会医療法人誠光会草津総合病院の栄養サポートチームの皆様に、この場を借りて御礼申し上げます。

本研究の一部は大阪ガス福祉財団の助成により実施された。

#### 参考文献

- [1] 厚生労働省, 平成 27 年版厚生労働白書 資料編, p.10.
- [2] Teramoto S, Fukuchi Y, Sasaki H, Sato K, Sekizawa K, Matsuse T, High incidence of aspiration pneumonia in community- and hospital-acquired pneumonia in hospitalized patients: a multicenter, prospective study in Japan, *Journal of the American Geriatrics Society*, Vol.56, No.3, pp.577-579 (2008).
- [3] 大類孝, 超高齢社会における誤嚥性肺炎の現状, *日本老年医学会雑誌*, Vol.50, No.4, pp.458-460 (2013).
- [4] Takahashi K, Groher ME, Michi K, Methodology for Detecting Swallowing Sounds, *Dysphagia*, Vol.9, No.1, pp.54-62 (1994).
- [5] 中山裕司, 高橋浩二, 宇山理沙, 平野薫, 深澤美樹, 南雲正男, 嚥下音の発生部位と音響特性の検討 - 健常成人を対象として -, *昭和歯学会雑誌*, Vol.26, No.2, pp.163-174 (2006).
- [6] 平野薫, 高橋浩二, 宇山理紗, 道健一, 頸部聴診法による嚥下時発生音の評価の指標に関する検討: 嚥下後の呼気音の聴覚的評価と嚥下透視所見および音響特性との関連について, *日本口腔科学会雑誌*, Vol.50, No.2, pp.82-89 (2001).
- [7] Borr C, Hielscher-Fastabend M, Lucking A, Reliability and Validity of Cervical Auscultation, *Dysphagia*, Vol.22, No.3, pp.225-234 (2007).
- [8] 大宿 茂, "頸部聴診法", *老年歯科医学*, Vol. 28, No. 4 pp. 331-336, 2013.
- [9] 兵頭政光, 西窪加緒里, 弘瀬かほり, 嚥下内視鏡検査におけるスコア評価基準 (試案) の作成とその臨床的意義, *日本耳鼻咽喉科学会会報*, Vol.113, No.8, pp.670-678 (2010).
- [10] 高木英行, 使える!統計検定・機械学習-II-3 群以上の場合の有意差検定, *システム/制御/情報*, Vol.58, No.10, pp.432-438 (2014).