

# 歌声改善システムに関する研究

鬼頭 良太

指導教員：辻 孝吉

## 1 はじめに

発声練習には長い期間が必要である上、効率的に進めるには自分の状態を判断できる人のもとで行うことが重要となる。

そこで、本研究では人の発声の悪癖を判定し、発声者の課題を見つけることを目的とする。

本研究で考察する悪癖は団子声の判定となる。

これは、判定の目的として、団子声は声道を塞ぐ発声であり響きが低減される発声のため、訓練の妨げになることが考えられるためである。

そのため本研究では人の音声情報のみで悪癖の判定を行うシステムの構築の足がかりとして団子声の判定指標の特定を行う。

詳細な流れとしては、発声された音声の定常区間を切り出し、その音声に対し OpenSMILE を用いた音声分析を行う。この分析結果から判別に有効な特徴量を検討する。

## 2 先行研究について

歌声に関する研究については、歌声を自動評価する研究[1]、歌声の声区を自動判定する研究[2]、第二フォルマントの数値によって喉が締まった発声を判定するもの[2]、歌声に存在する響きに相当する指標を求めたもの[3]、YUBA メソッドに利用する表声と裏声の判定を助ける研究[4]、舌骨筋の稼働に着目した音声分析を行ったもの[5]などが挙げられる。

しかし、悪癖を自動判定する研究は存在しない。

このため、発声訓練は問題のある発声に対して対症療法的に行うために、発声を判別することが必要となる。したがって、今回の研究の目的は特定の発声訓練の前段階として、訓練が必要な声であるかを判定するという目的となる。

## 3 発声について

人間の発声とは声帯の緊張度や声門下圧、呼気流量などによって形成される。これらの要素に関わるものが喉頭周りの筋肉であるが、これらの筋肉は意識的に動かすことが困難である。

そのために歌唱音声は特に訓練していなければその人の通常発声に使用している筋肉に大きく依存することとなり発声に難がある人が多い。

口頭を支える筋肉の偏りからなる喉を閉めたような発声や、口頭を押し下げるために舌を下方に硬直させた団子声発声など、発声する悪癖は数多い。

他にも、高音域の発声に難がある、つまりは内喉頭筋の動きに問題のある人や、体づくり、呼吸法など様々な改善対象が存在しているが、内喉頭筋については先述した YUBA メソッドなどのように訓練法が確立している上に関連研究も数多い。

体づくりや呼吸法に関しては喉とは異なり練習を間違えることが少ないため、本稿では外喉頭筋に関連した問題を取り扱う。

## 4 団子声について

この発声方法は喉頭を舌で無理やり押し下げて発声を続けると生じるもので、力んだ舌が喉の空間を押しえ込み、歌声が響かず内側にこもったような響きとなる。

喉頭をさげる意識があると陥りやすい。この発声方法の定義はフースラーの書籍から抜粋した[6]。

人間の発声は喉頭を吊っている筋肉のバランスによって声道の状態を決定する。喉頭を挙上する筋肉、さげる筋肉が二種類ずつ存在していて、高音発声時に挙上筋肉のみが動くとき喉が締まったような響きとなるため喉頭を下げる必要がある。この下げる運動に舌が付随して動く場合に団子声という状態に陥る。

## 5 openSMILE について

openSMILE (open-source Speech and Music Interpretation by Large-space Extraction) とは音声の特徴抽出や分類を行うことのできるオープンソースのツールである。取得できる特徴量には MFCC (メル周波数ケプストラム係数)、各種フォルマント周波数とそのパワー、ジッター、シマー、などがあるこの openSMILE では 88 種類ほどの特徴量を取得することができる。本研究ではこちらの openSMILE を用いて団子声判別に寄与する特徴量を見つけることを目的とする。

## 6 研究手法

無音の室内において母音の発声をしてもらい、その特徴を得る。音声の種類は通常の発声と団子声の二種類取得する。

取得する音声は「お」母音を選択する。これは舌を押し下げた団子声発声が行いやすいためである。

これらの音声に対し分析し、特徴量の検討を行う。

発声者は 20 代男性 5 人、女性 2 人で行った。

具体的な流れとしては以下のフローの通りを行う

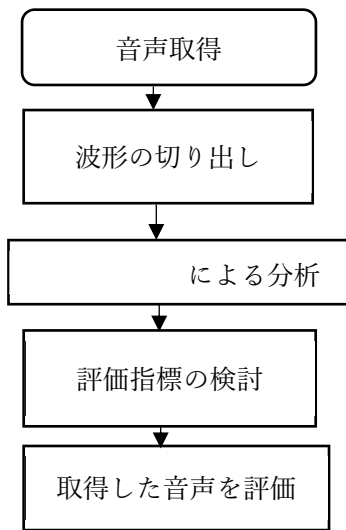


図:実験のフローチャート

## 実験結果

実験の結果を以下の表にまとめる。

表:各特微量

人	性別	MFCC1	MFCC2	MFCC3	MFCC4	ジッター	シマー	ジッター標準偏差	HNR
通常A	男	45.499	24.338	14.728	-43.632	0.00319	0.177	0.5950179	8.81
通常B	男	38.831	13.510	13.408	-44.953	0.00198	0.199	0.56510633	9.4
通常C	男	41.899	18.777	15.788	-36.892	0.00389	0.0039	0.6886965	9.4
通常D	男	34.581	24.554	11.250	-55.818	0.00728	0.197	0.4079375	9.89965
通常E	女	43.493	29.267	-33.87	-44.589	0.00546	0.178	0.33736655	14.7720
通常F	女	40.406	1.7627	-14.13	-51.247	0.00419	0.354	0.4483961	12.5932
団子H	男	34.100	3.1886	11.004	-13.099	0.01309	0.487	2.4180775	10.2262
団子K	男	36.836	20.877	14.405	-35.308	0.00613	0.2434	5.1724086	12.5918

この結果からジッターの標準偏差が団子声判定に利用できると考えられる。

この結果以外にも特徴を取得したが、ジッターの標準偏差ほどに明確な差が生じた特徴量は存在しなかった。

## 考察

実験結果からジッターの標準偏差に大きな違いが生じることがわかった。これは団子声の外喉頭筋ではなく舌で無理やりに喉頭を下方へ押し下げて喉頭位置を安定させた発声のため、音声発声に伴い発声が不安定になるためと考える。

## 結び

団子声など悪癖を判定する指標を特定した。今後はこの特徴量を用いたシステムの製作が求められる。しかし、本研究では団子声発声を行えるものが一人しかいなかったため更なる研究として団子声の音声を取得することが求められる。

団子声は響きを低減した発声であったため音声内の響きと雑音の比率を表す HNR も判別に利用できると考えていたが、通常の音声や歌唱音声と比較して大した違いが見られなかった。

また、今回は定義をもとに発声を取得したので信頼性は少々高まったとは考えられるが、歌唱音声との明確な差異を考察するために歌唱経験や訓練を受けたことのある者を被験者に選ぶ必要がある。

団子声は特別優れた発声ではないために特別訓練を行う必要がないと考えていたが、歌唱音声との判別や、訓練を受けた者の団子声発声との比較を行うことで更に詳細に判定ができると考えられるためである。

そのため今後は訓練を受けた者の音声取得や異なる被験者による団子声発声取得が課題となる。

## 参考文献

- [1] 中野他”楽譜情報を用いない歌唱力自動評価手法”, 情報処理学論, 48 巻 1 号, pp. 227-36, 2007-01-15
- [2] 平山他”ポピュラー歌唱における高音域の声区と発声状態の判別手法”, 情報処理学会第 74 回全国大会, 2S-10, 2012
- [3] Sundberg, J.”The Science of the Singing Voice”, p.226, 1987
- [4] 浅野, ”裏声判別指標を用いたボイストレーニングソフトウェア”, 三重大学修士論文, 2013
- [5] 平山他”筋電センサーを用いた歌声分析のための喉頭音源分析”, 第 11 回情報科学技術フォーラム, E-010, 2012
- [6] Frederick Husler. Yvonne Rodd=Marling(1965). Singing The Physical Nature of the Vocal Organ. (フレデリック・フースラー. イヴォンヌ・ロッド=マーリング. 大熊文子(訳) (1987). ”うたうこと”. 音楽之友社), 1987, 182p, ISBN4-276-14252-0