

## Web アプリケーションによる音声障害評価システムの構築\*

矢口勝年, 森大毅, 粕谷英樹 (宇都宮大)

## 1 はじめに

声帯結節, 喉頭がん, ポリープといった喉頭疾患では音声障害を伴うことが多くある。この喉頭疾患の診断や治療の支援, 経過観察や治療効果の確認には音声の評価が必須であるといえる。しかし, 主観的な評価では評価の曖昧性や不安定性を避けることはできない。そこで, 客観的な評価法として PC を利用した音声の音響的評価が有効である。

従来このような目的ではスタンドアローンの音響分析ソフトウェアが利用されていたが, 音声障害の音響的評価のための標準的なシステムを Web 技術の上に構築し, インターネットを通して広く利用できるようにすることにより, 世界中の言語聴覚臨床・研究従事者が共通の基盤上で評価できるようになるばかりでなく, 評価結果の統計データを共通資源として利用できる, いわばソーシャルな音声障害評価システムの実現にも道が開かれる。

この目的を実現するために, 本研究では音声障害の音響的評価を行うシステムを Web 技術に基づいて構築した。本稿ではその概要を述べる。

## 2 音声障害の評価

## 2.1 音声障害の音響的特徴

喉頭に何らかの障害が発生すると, 音声に影響が出る。疾患によって正常な音声とは異なる音声の特徴が見られ, これを音声障害という。持続母音に対する音声障害の音響的特徴として基本周期のゆらぎ, 振幅のゆらぎ, 声門体積流波形のスペクトルゆらぎ, 喉頭雑音の周波数特性と大きさなどがあげられる。経験を積んだ耳鼻科医は, これらの変化を聞き分け患者の声から疾患をある程度推定できる。

これまでも客観的な音声の評価法として, 音声の音響分析に関する研究が行われてきた。言語聴覚臨床・研究 [1] の現場で利用されてきた代表的な音声分析ソフトウェアとして, MDVP [2] が挙げられる。これは, ピッチや振幅のゆらぎなど最大 22 種の音響パラメータによる分析を行える。音声障害だけでなく, 言語聴覚士による発声訓練などにも利用されているが, 高価なソフトウェアであり, しかも使いこなすためには専門的知識を必要とするため, 医療・言語聴覚訓練従事者の誰にでも使えるソフトウェアではない。他にも, 喉頭がんのスクリーニングを目的とした嘔声評価システムがある [3]。

## 2.2 音響パラメータの算出方法

音声評価における音響パラメータには多くの種類があり, 目的によって使用するパラメータも変化する [1]。

本研究で使用する音響パラメータは, 周期変動指数 (Period Perturbation Quotient; PPQ) と振幅変動指数 (Amplitude Perturbation Quotient; APQ), 規格化雑音エネルギー (Normalized Noise Energy; NNE) [3, 4, 5] である。基本周期のゆらぎを定量化する音響パラメータとして PPQ, 振幅のゆらぎのパラメータとして APQ, 喉頭雑音のパラメータとして NNE を使用している。声帯に何かしらの障害が存在すると, これらのパラメータが変化することがわかっている [5]。

変動指数 (Perturbation Quotient; PQ) は周期や振幅のゆらぎを定量化するための音響パラメータであり, 次の式によって与えられる。

$$PQ = \frac{1}{N-2} \sum_{n=2}^{N-1} \left| \frac{u(n) - u_L(n)}{u_L(n)} \right| \times 100 \quad (\%) \quad (1)$$

ここで,  $u(n)$  は周期あるいは振幅系列,  $N$  は系列の長さ,  $u_L(n)$  は系列の三点移動平均である。 $u(n)$  が基本周期系列のものを PPQ,  $u(n)$  が最大振幅系列のものを APQ と呼ぶ。

NNE は喉頭雑音を定量化するための音響パラメータであり, 推定した雑音の平均パワーを音声信号の平均パワーで規格化したものである。これは音声信号に含まれる喉頭雑音の量をパーセントで表す。持続母音音声の周期信号と喉頭雑音の分離には楕円フィルタを用いる [3, 4, 5]。NNE には全帯域を用いたもの (NNEa) と 1.5 kHz 以上の帯域を用いたもの (NNEb) があり, これらは次式で定義される。NNEb は, 低周波帯域をフィルタによって除外し, 高周波帯域での雑音成分を評価している。

$$NNEa = \frac{\sum_{n=0}^{N-1} w^2(n)}{d_L \sum_{n=0}^{N-1} s^2(n)} \times 100 \quad (\%) \quad (2)$$

$$NNEb = \frac{\sum_{n=0}^{N-1} w_{HPF}^2(n)}{d_L \sum_{n=0}^{N-1} s_{HPF}^2(n)} \times 100 \quad (\%) \quad (3)$$

ここで,  $s(n)$  は音声信号であり,  $w(n)$  は楕円フィルタによって得られた音声信号の喉頭雑音成分,  $s_{HPF}$  と  $w_{HPF}$  はそれぞれ  $s(n)$  と  $w(n)$  を遮断周波数が 1.5 kHz の高域通過フィルタに通したもので,  $N$  は信号の長さ,  $d_L$  は楕円フィルタの次数に依存した補正項である。

\* Construction of an evaluation system for voice disorder by web application. by YAGUCHI, Katsutoshi, MORI, Hiroki, KASUYA, Hideki, (Utsunomiya University)

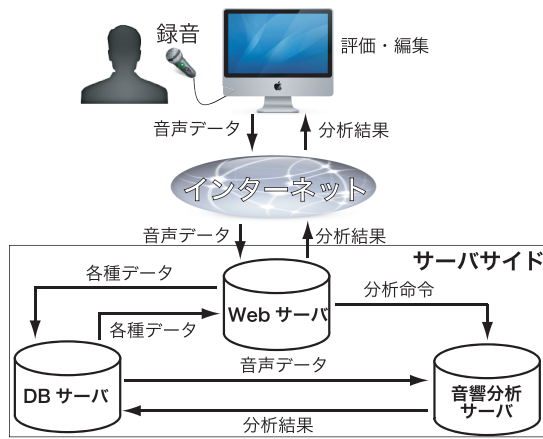


Fig. 1 音声障害評価システムの概要

### 3 システム概要

本システムはWebアプリケーションである。音声の録音、分析、評価といった音声障害の評価において必要である作業の全てをWebブラウザ上から行える。Webアプリケーションは、Webブラウザが入ったPCとインターネット環境さえあれば利用でき、ソフト管理の手間が減ることによって利用者の利便性が向上する。

Fig. 1 に本システムの概略図を示す。本システムは大きく2つに分けることができる。一つ目はクライアントサイドであり、もう一つはサーバサイドである。サーバサイドは、Webサーバ、音響分析サーバ、データベース(DB)サーバからなる。ここでは、順に説明を行う。

1. Webサーバはクライアントからの要求に応じて適切なWebページを返答する。Webアプリケーションであるため、要求に応じてWebページを変更する必要があり、PHPによって動的にWebページを生成している。
2. 音響分析サーバは、送られてきた音声データの分析を行う。分析可能であれば、上述の4種のパラメータについて分析を行う。
3. DBサーバは必要に応じてデータの保存や閲覧を行う。RDBMSによる素早いデータの検索や保存を可能にしており、同時アクセスによるデータの矛盾を防止する。

Fig. 2 に分析結果の表示画面を示す。ただ単に分析結果を表示するだけでなく、検索機能も実装している。名前や年齢、性別による検索機能があり、他データとの比較が可能である。ここでは、F0の経年変化を表すグラフや、過去のデータを重ねたレーダーチャートを表示する。

加齢によって音声に変化することが知られている[6, 7]。過去データとの比較は年齢による音声変化の観測であり、音声障害と年齢の関係を観察することで

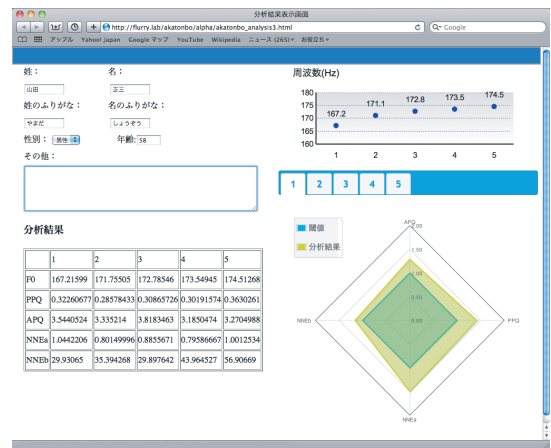


Fig. 2 分析結果表示画面

ある。このシステムでは視覚的に過去のデータとの比較ができ、音声障害と年齢による音声変化の観測を同時に行うことができる。

### 4 おわりに

客観的な音声障害の評価として、音響評価を行うシステムの構築をした。音声の録音・分析・評価といった一連の流れがWebブラウザ上から利用でき、加齢による音声変化を考慮した音声障害評価が行える。今後は、使いやすさに関する主観評価実験を行う。本システムは、言語聴覚臨床の現場で広く利用可能な音声の音響評価システムとして、近日中の公開を目指している。

謝辞 本研究の一部は、科学研究費補助金(課題番号22500147)の助成を得た。

### 参考文献

- [1] 吐師道子, “音響分析の言語聴覚臨床及び研究への応用に伴う諸問題,” 音講論(秋), 325-326, 2009.
- [2] KayPENTAX, Multi-Dimensional Voice Program.
- [3] 森大毅, 小田和浩明, 小野辰矩, 伊藤友一, 粕谷英樹, “喉頭がんスクリーニングのためのインターネットを利用した音声の音響評価システム,” 日本音響学会誌, 62-3, pp.193-198, 2006.
- [4] 長谷川清雅, 粕谷英樹, “病的音声検出のための規格化雑音エネルギー(NNE)の性質,” 日本音響学会誌, 43-3, pp.182-183, 1987.
- [5] 粕谷英樹, “声の検査法 基礎編 第2版,” 日本音声言語医学会編, 1994.
- [6] 西尾正輝, 田中康博, 新美成二, “加齢に伴う音声の変化,” 音声言語医学, 50, pp.6-13, 2009.
- [7] H. Kasuya *et al.*, “Longitudinal Changes of Selected Voice Source Parameters,” Proc. Interspeech 2010, 2570-2573, 2010.