

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 6 月 10 日現在

機関番号：12201

研究種目：基盤研究 (C)

研究期間：2010～2012

課題番号：22500147

研究課題名（和文） 音声障害の音響的評価法の標準化に向けた基盤的研究

研究課題名（英文） Fundamental study toward standardization of acoustic evaluation of pathologic voice

研究代表者

粕谷 英樹 (KASUYA HIDEKI)

宇都宮大学・大学院工学研究科・名誉教授

研究者番号：20006240

研究成果の概要（和文）：

音声治療現場における EBM の機運の広がりによって、音声障害の音響的評価法の標準化の要求が増している。しかもこの評価法は、誰でも最新の評価技術を利用できるようにするために、Web 技術で実現することが望ましい。本研究では、KayPENTAX 社の MDVP、アムステルダム大学の Praat、宇都宮大学の L-Voice の 3 つの評価法の性能を自然音声と合成音声を用いて比較検討した。基本周期のゆらぎ測定には L-Voice、振幅ゆらぎには Praat、加法的喉頭雑音の計測には L-Voice がそれぞれ適していることが分かった。また、L-Voice を Web 技術によってインターネットを介して利用できるようなパイロットシステムを製作し、その有効性を明らかにした。臨床的な評価の結果、言語聴覚士や医師の操作性が重要であることが分かった。

研究成果の概要（英文）：

Needs for establishing standard methods are increasing to evaluate pathologic voice acoustically, which is primarily motivated by a trend of evidence-based medicine at voice related clinics. Furthermore, the methods are expected to be realized based on web-based technology so that many users are able to access to the most contemporarily organized evaluation system. We made a comparative study of three well-known methods to make an acoustic evaluation of pathologic voice using both natural and synthetic speech: MDVP of KayPentax which is a rather defacto standard in this area, Praat delivered by University of Amsterdam, and L-Voice developed at Utsunomiya University. L-voice provided most reliable performance in the measurement of jitter and glottal noise, whereas Praat showed the best results for shimmer detection. A pilot system on the basis of L-Voice which is accessible over the internet has been implemented using web-technology and its efficacy has been shown in clinical assessment experiments. The clinical evaluation also indicated importance of GUI of the system.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2010年度	1,400,000	420,000	1,820,000
2011年度	1,300,000	390,000	1,690,000
2012年度	600,000	180,000	780,000
年度			
年度			
総計	3,300,000	990,000	4,290,000

研究分野：音声科学工学

科研費の分科・細目：情報学，知覚情報処理・知能ロボティクス

キーワード：音声障害，音響的評価，Web 技術

1. 研究開始当初の背景

音声を職業とする教員、アナウンサー、歌手、俳優などだけでなく、一般社会でもコミュニケーションに果たす音声の役割の重要性が認識され、音声障害に対する社会的関心が高まっている。このような社会的ニーズに応えるために、音声障害の治療や言語音声生活の衛生指導の実施を目的にした“音声外来”を設置する医療機関も増えている。そこでは EBM (Evidence Based Medicine) が必須要件であり、音声治療の効果の観察、診断の支援、音声の衛生指導などのために、客観的な評価法の一つである“音声の音響的評価”が日常的に実施されている。しかし、標準的な評価法が存在せず、評価法も限定的なために、治療機関相互あるいは研究者間で資料や知見の交換に不便を来している。

現在、臨床および研究分野で利用されている音声障害の主な音響的評価法（装置、システム）は3種類ある。米国 Kay-Elementrics（現 Kay-Pentax）で Delisky 博士を中心に開発された CSL MDVP (Computer Speech Lab, Multi-Dimensional Voice Program)、オランダのアムステルダム大学から提供されている Praat、それに我々が開発した L-Voice である。

本研究課題に密接に関連する我が国の2つの学会の学術講演会で、最近立てつづけに音声障害の音響的評価法に関する特別セッション・セミナーが企画されるとともに国際的な学術誌 (Folia Phoniatica et Logopaedica, 以下“Folia”と略す。) に、特集記事が掲載された。これらの最近の研究活動を通して言えることは、比較的広く利用されている3種類の評価技術にあっては、その性能の相互比較調査研究によって高精度で合理的な評価法として統一し、不足する部分についてはロバストな評価法を開発し、臨床で使いやすい評価システムを実現することである。

2. 研究の目的

声帯結節、ポリープ、パーキンソン病や脳血管障害などに伴う音声障害を治療する際の治療効果の観察や診断の支援、更に声（音声）の衛生指導のために、EBM (Evidence Based Medicine) に沿った音響的評価（検査）は必須である。しかし、標準的な評価法が確立されていない上に、機能が限定的な装置しか入手できないため、この分野の発展は遅れている。本研究は、“音声障害の音響的評価法”の標準化に向けた調査・研究ならびに比較実験を行って、音響的評価法の推奨基準を提言し、インターネットを介したパイロットシステムを構築して、臨床応用の基礎を確立することを目的にする。

3. 研究の方法

相当数の病的及び正常音声標本を収録して、上で述べた3つの評価法の性能を比較検討する。さらに基本周期や振幅のゆらぎを高精度に制御できる音声合成システムを開発して、合成音声による3者の評価も行う。それらの結果を総合的に判断して、標準化に向けた方向性を示す。

Web 技術を用いた音声障害の音響的評価システムのパイロットモデルを製作して、臨床場面で実際に使用しながら評価し、原理的な面や使用法の面から、音響的評価法の標準化に向けた基盤を確立する。

4. 研究成果

(1) 自然音声を用いた音響評価法の比較研究

本研究では、上述の KayPENTAX 社の MDVP (Multidimensional Voice Program)、Amsterdam 大学の Praat、宇都宮大学で開発された L-Voice の性能を比較評価した。音響特徴量として、声帯振動のゆらぎの程度を示すジッタ (PPQ)、シマ (APQ) と、声門閉鎖不全により発生する喉頭雑音の少なさの程度を示す調波対雑音比 (HNR) を対象とした。L-Voice では規格化雑音エネルギー（全帯域 NNEa）から HNR に換算した。MDVP の HNR は対象とする周波数帯域が異なるため直接的に比較することはできなかった。

使用した音声サンプルは、持続して発声した母音 [a] の定常区間であり、病的音声 186 例、正常音声 190 例であった。

ジッタ (PPQ) についての分析結果の対応を図 1、2 に示す。横軸は Praat の PPQ [%] である。図 1 の縦軸は L-Voice の PPQ [%] である。

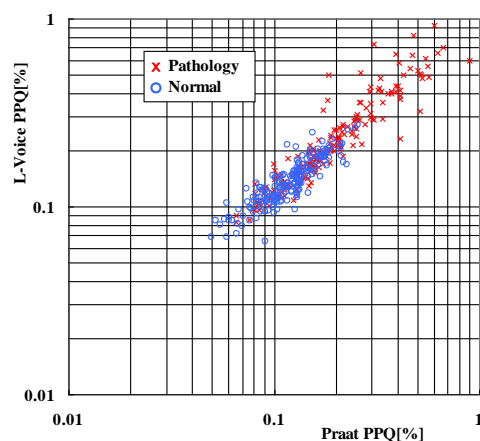


図 1 Praat と L-Voice の比較 (PPQ)

図より L-Voice の方が若干大きな値を示すが両者の値に大きな違いはない。これに対し図 2 の縦軸は MDVP の sPPQ であるが、全体的にかなり大きな値を示している。軸の値より MDVP の sPPQ が 1 を超えているデータは

表示されていないが、病的 37 例、正常 4 例が存在している。シマ(APQ)に関しても、3者の間にはPPQと同様の傾向が見られた。調波対雑音比(HNR)について比較した結果、L-Voiceの方が若干大きめの値となるが、大きな違いがないことが分かった。

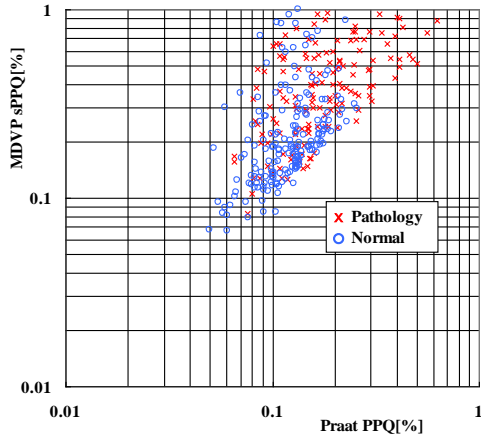


図2 PraatとMDVPの比較 (APQ)

音響パラメータをもとに病的／正常の判別の正答率を調べた。ここでは病気と正常のサンプル数がほぼ同数であることから、病気を病気と正しく判断する割合と、正常を正常と正しく判断する割合の平均値をもって正答率とした。表1に結果を示す。雑音の欄はPraatとL-VoiceはHNR[dB]であるが、MDVPではNHR[%]である。表から分かるように正答率はほぼ同程度であるが、PPQではMDVPの75.0%，APQではPraatの74.0%である。喉頭雑音に関してはPraatのHNRで78.9%が最も優れていた。表には無いが、比較したパラメータ全体ではL-voiceの高周波帯域でのNNEbが79.2%で最高値を示した。

表1 病気／正常の判別の正答率

		PPQ	APQ	雑音
Praat	正答率[%]	71.0	74.0	78.9
	閾値	0.14	1.40	22.0
L-Voice	正答率[%]	71.0	72.0	76.6
	閾値	0.20	1.70	23.2
MDVP	正答率[%]	75.0	72.0	66.6
	閾値	0.30	1.4	0.14

3種の音声評価法から得られる特徴量には違いがあるが、病気／正常の判別の正答率からは概ね同じ結果となっている。このことから、値が大きめにシフトしても、また分布に広がりを持っていても、音声の持つ特徴に符合する特徴量を求めていると考えることができる。

(2) 合成音声を用いた音響評価法の比較研究
合成音声を用いて前節で述べた3種類の

音響的評価法による測定精度の比較実験を行った。合成方式は加法的雑音モデルに基づいたARX分析・合成システムのフォルマント型合成部の改良型で、有声音源はRK波形である。音声合成器の仕様は、1)標本化周波数は自由に指定できる、2)フォルマント周波数は指定時間内で一定、3)ピッチ(周期)曲線の概形は放物型、4)ジッタはRK波形の声門閉鎖点(GCI)を制御して与える、5)シマはRK波形の振幅パラメータ(AV)に加える、6)加法的雑音量は正規乱数の σ^2 で制御する、である。なお、ジッタは周波数領域での群遅延量を操作して標本化周波数以上の精度で制御した。フォルマント周波数は成人男性の母音「あ」に固定した。また、ピッチ曲線は初期周期10msから約300msで最小周期9.2msとなり、以後緩やかに長くなる特性である。ジッタは、3段階、シマは4段階、加法的雑音量は4段階に変化させた。なお、すべての組合せの音声を作成してはいない。

比較した音響特徴量は、ジッタ(PPQ)、シマ(APQ)、調波対雑音比(HNR)とした。なお、L-Voiceでは全帯域の規格化雑音エネルギーNNEa [%]からHNR [dB]に換算した。MDVPのNHRは周波数帯域が異なるため直接、他と比較はできないので、後述する方法でHNRに変換して比較した。理論的PPQに対する分析結果を図3に示す。理論的PPQは合成時に生成した周期系列から直接求めた値である。

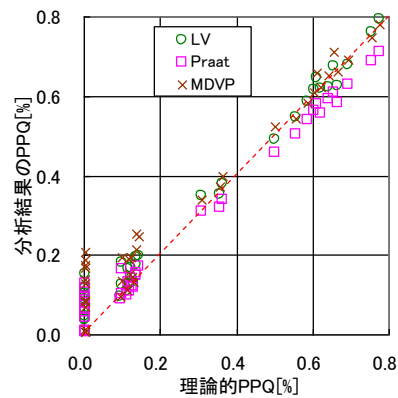


図3 3種類のPPQと理論値の関係

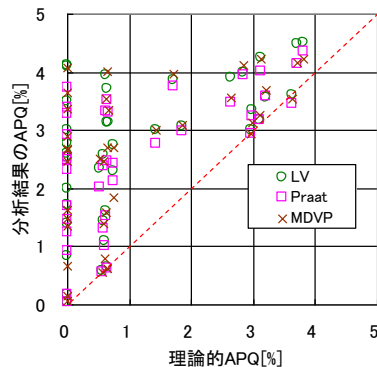


図4 3種類のAPQと理論値の関係

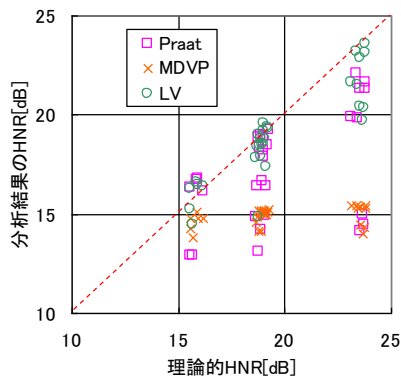


図5 3種類のHNRと理論値の関係

分析結果のPPQはシマや雑音成分の影響を受けていると考えられるが、影響は少ない。理論的APQに対する分析結果を図4に示す。理論的APQも合成時の音源振幅系列から求めたものであるが、分析結果のAPQはジッタや雑音成分の影響を大きく受けており、値が大きい方に分布している。図5に理論的HNRに対する分析結果を示す。理論的HNRも合成時の調波信号と雑音信号から求めた。L-VoiceのNNEa [%]は次式を用いてHNR [dB]に換算した。

$$HNR = 10 \log((100 - NNEa) / NNEa)$$

MDVPのNHRは1.5~4.5 kHzでの雑音成分を0~4.5 kHzの帯域での調波成分で割ったものであるため、次のようにしてHNRに変換した。1)理論的HNR(x)とMDVPで使用する帯域での理論的HNR(y)の関係式 $y = 0.9987x - 6.587$ を導き、 $x = y + 6.6$ と近似した。2)NHRの逆数の対数(dB値)を求め6.6を加えた。

図から明らかなように、PPQについては3つの方法の違いによる差異は少ないが、APQはジッタや雑音の影響を強く受け変動が大きく評価されることが、また、HNRも影響を受け雑音が多めに見積もられることがわかった。このことからAPQはジッタや雑音の影響を強く受け、理論的APQとは大きく異なる。しかしこのことはAPQの定義そのものに関係するので、誤差と見なすことの適否は検討を要する。表2に分析結果のRMS誤差を示す。

表2 3種類の測定値のRMS誤差

二乗誤差		Praat	LV	MDVP
PPQ	全体	0.053	0.055	0.079
	< 0.1	0.068	0.083	0.116
	≥ 0.1	0.041	0.028	0.046
APQ		1.742	1.947	1.868
HNR		3.724	1.514	5.518

PPQについては閾値0.1で2分したが、実際の多くの音声がこの範囲にあるからである。

表より、PPQに関しては0.1以上でL-Voiceが、APQではPraatが、HNRではL-Voiceが最もRMS誤差が小さい。HNRでは国際的に最も広く利用されているMDVPがL-Voiceより4 dBも大きいRMS誤差を示した点は注意が必要である。

(3)インターネットを介した音声障害の音響評価システムの研究

音声障害の音響的評価の標準化に向けた基盤的研究の核心の一つは、音響分析・評価システムを、インターネットを介して誰でも利用できるように製作したうえで公開し、広く利用者の意見や評価を得ながら改良していくことである。

そこで、Web技術を用いてWebアプリケーションとして実現することを試みた。音声の録音・分析・評価といった一連の作業をWebブラウザから行えるようにした。したがって、Webブラウザが使えるパーソナルコンピュータ(PC)とインターネット環境さえあれば所定の手続きを経て誰でも利用できるようになる。

図6に製作したパイロットシステムの概要を示す。システムはクライアント側とサーバ側からなる。サーバ側は、Webサーバ、音響分析サーバ、データベースサーバ(RDBMS)の3つのサーバからなる。それぞれの特徴は以下のようなものである。

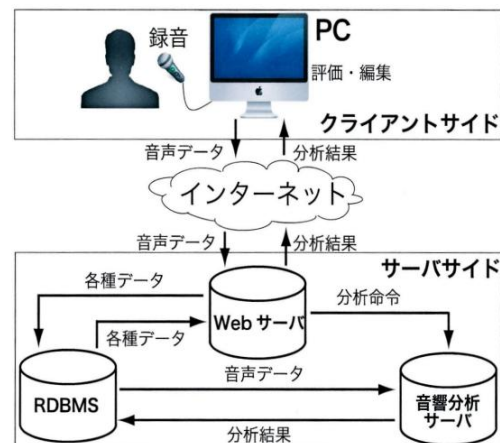


図6 音声評価システムの構成

- ① Webサーバはクライアントの要求に応じて適切なWebページを返す。Webページの更新は動的に行われる。
- ② 音響分析サーバは送られてきた音声信号を分析し、詳細な分析を行う必要があるかどうかの前処理を行い、周期性が極めて不安定な場合はそれで停止し、安定性があれば、詳細な分析をする。
- ③ RDBMSサーバは必要に応じて話者情報、音声信号、分析・評価結果、などのデータの保存・管理を行いながら、柔軟な検索

索・修正機能を提供する。

図7に分析結果の表示画面の例を示す。結果の表示だけでなく、氏名、年齢、性別などによる検索、他データとの比較、などの機能を備えている。



図7 分析・評価画面の例

本パイロットシステムのサーバは宇都宮大学に設置されている。研究分担者の所属する福島県立医科大学からサーバにインターネット経由でアクセスして、実際に音響評価を行った結果、音響の評価法については実用上問題ないが、操作性やデータセキュリティの面で今後改良が必要であることが分かった。実用化に向けた一層の改善が次の研究課題である。

5. 主な発表論文等

[雑誌論文] (計2件)

- ① 粕谷英樹, 声質の伝える情報とその関連量, 日本音響学会誌, 査読有, Vol. 68, No. 10, 2012, 520-526.
- ② Kasuya, H., Mori, H., Yoshida, H. and Ebihara, S., Longitudinal changes of selected voice source parameters, Pro. Interspeech 2010, 査読有, Vol.1, 2010, 2570-2573.

[学会発表] (計12件)

- ① 谷亜希子, 多田靖宏, 岡野渉, 菅野和広, 荒川愛子, 小針香奈, 大森孝一, 法串, 学校教員における音声障害の治療経過, 日本音声言語医学会, 2012年10月19日, 大阪国際交流センター (大阪市) .
- ② 矢口勝年, 森大毅, 粕谷英樹, 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, 日本音響学会研究発表会, 2012年9月21日, 信州大学 (長野市) .
- ③ 菊地義信, 粕谷英樹, 森大毅, 合成音声

による音声分析ソフトウェア MDVP, Praat, L-Voice の比較, 日本音響学会研究発表会, 2012年3月15日, 神奈川県 (横浜市) .

- ④ 多田靖宏, 西条聡, 谷亜希子, 鈴木輝久, 大森孝一, 当科における音声障害診療の研修プログラム, 日本音声言語医学会, 2011年10月6日, ホテルグランドヒル市ヶ谷 (東京都) .
- ⑤ 菊地義信, 粕谷英樹, 森大毅, 音声障害の音響評価のための分析ソフトウェア MDVP, Praat, L-Voice の比較, 2011年9月21日, 島根大学 (松江市)
- ⑥ 粕谷英樹, 声質の伝える情報とその関連量, 日本音響学会研究発表会, 2011年9月20日, 島根大学 (松江市) (招待講演) .
- ⑦ 粕谷英樹, 森大毅, 吉田肇, 海老原敏, 声の年齢変化に関する一考察, 日本音響学会研究発表会, 2011年3月9日, 早稲田大学 (東京都) .
- ⑧ Kasuya, H., Yoshida, H., Ebihara, S. and Mori, H., Age-related changes in selected voice source parameters of males: A longitudinal study, 4th World Voice Congress, Sept. 2010 (Seoul, Korea) (査読有) .

6. 研究組織

(1)研究代表者

粕谷 英樹 (KASUYA HIDEKI)

宇都宮大学・大学院工学研究科・名誉教授
研究者番号: 20006240

(2)研究分担者

森 大毅 (MORI HIROKI)

宇都宮大学・大学院工学研究科・准教授
研究者番号: 10302184

菊地 義信 (KIKUCHI YOSHINOBU)

国際医療福祉大学・保健医療学部・准教授
研究者番号: 20091944

多田 靖宏 (TADA YASUHIRO)

福島県立医科大学・医学部・講師
研究者番号: 70363760

(3)研究協力者

矢口 勝年 (YAGUCHI KATSUTOSHI)

宇都宮大学・大学院工学研究科・博士前期課程