

音声障害の音響評価のための音声分析ソフトウェア MDVP, Praat, L-Voiceの比較*

○菊地義信 (国際医療福祉大・保), 粕谷英樹 (宇都宮大・名誉教授),
森大毅 (宇都宮大・工)

1 はじめに

音声の音響分析を行い音声障害の評価を行うことが広く行われ, この目的でいくつかの音声分析ソフトが開発され利用されている。しかし得られた音響特徴量は数値的に異なる場合が多く比較する上で問題となる。本研究では, 代表的な音響特徴量について音響評価ソフトウェア間でどの程度異なるかを調べた。また, 病気の音声を含む400名弱の音声を用い病的/正常の判別を行い, その能力の違いについても調べた。

2 音響分析

2.1 音声特徴量

本研究では音声評価ソフトウェアとしてMDVP (KayPENTAX 社), Praat[1](Amsterdam 大学), L-Voice[2](宇都宮大学)を使用した。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。

2.2 音声サンプル

使用した音声サンプルは持続して発声した母音[a]であり, 定常区間から1秒以内の長さで切り出した過渡区間を含まない音声である。音声評価ソフトウェアの分析アルゴリズムの関係で, 音声の開始部分と終了部分の数ピッチの音声区間の扱われ方が異なるがこれについては無視した。

音声サンプルの内訳は病気186名, 正常190名である。

3 結果

3.1 音響特徴量

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。

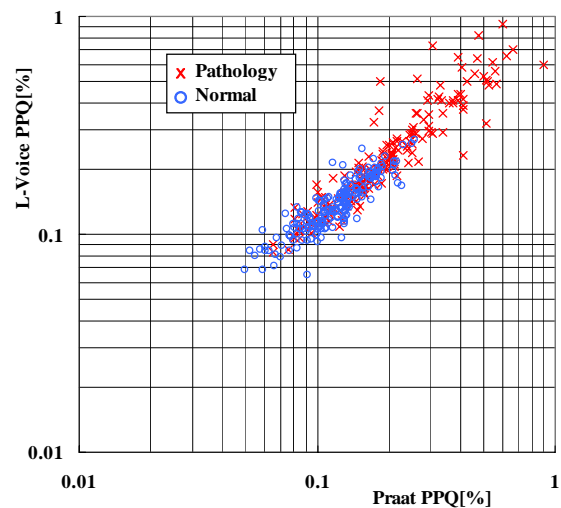


図1 PraatとL-VoiceのPPQの対応

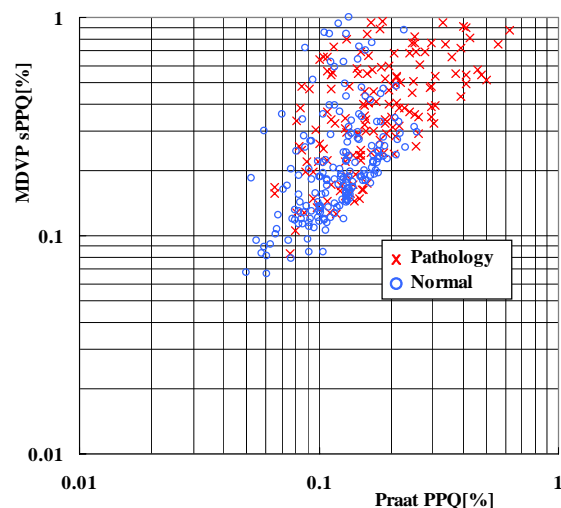


図2 PraatとMDVPのPPQの対応

これに対し図2の縦軸はMDVPのsPPQで

* Comparison of voice analysis software MDVP, Praat and L-Voice for acoustic evaluation of dysphonic voice, by KIKUCHI, Yoshinobu (International Univ. of Health and Welfare), KASUYA, Hideki (Utsunomiya University), MORI, Hiroki (Utsunomiya University).

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)についての分析結果の対応を図 3 に示す。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る。

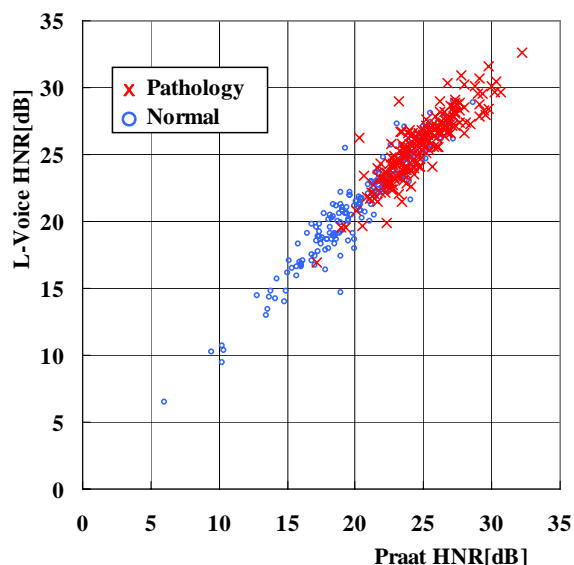


図 3 Praat と L-Voice の HNR の対応

表 1 病気／正常の判別の正答率
雑音の欄:Praat, L-Voice は HNR[dB],
MDVP は NHR[%]

		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

3.2 病気／正常の判別

音響パラメータをもとに病気／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。

4 考察

4.1 音響パラメータ

Praat と L-Voice が出力する音響特徴量には強い相関があるが、3つのパラメータともに、L-Voice の方が僅かではあるが大きめの値となった。PPQ, APQ に関しては、サンプリング周期以上の精度で周期を推定するために内挿を行っているが、その方法の違いによるものと考えられるが、具体的には今後の検討課題である。一方 HNR では、L-Voice の場合、ピッチ同期の楕円フィルタが関与するので、PPQ の違いの根拠と同じ可能性があるが、雑音量を推定する際に構成した楕円フィルタに混入する雑音の白色性を仮定したことも一因かもしれない。今後詳細に検討したい。

MDVP が出力する特徴量は他の 2 種に比べ、平均的に大きめかつ広い範囲に分散する傾向にある。

4.2 病気／正常の判別

3種の音声評価ソフトウェア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。

5 おわりに

ここでは 376 例の発声した音声を用い音声評価ソフトウェアの比較を行ったが、求めた特徴量の真の値は不明である。より比較を確かなものにするために、今後は合成音声を用いるとともに、3者の計測アルゴリズムの特徴と分析結果の関係についても検討したい。

謝辞

本研究の一部は科研費(22500147)によった。

参考文献

- [1] <http://www.fon.hum.uva.nl/praat/>
- [2] 朱達中他, 音講論, 1-8-7, 1995.