

# 合成音声による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

音声障害の程度を客観的に評価するために音響分析を行うことが広く行われ, その目的の音声分析ソフトウェアが開発され利用されている. しかし分析ソフトウェアの違いにより得られた音響特徴量には差異が生ずる. これまで筆者らは発声した音声を用いて比較を行ってきたが<sup>[1]</sup>, 基準値がないため分析ソフトウェアの精度を評価することが難しかった. 本研究では, 合成音声を用いて比較実験を行った.

## 2 実験方法

### 2.1 音声合成

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<sup>[2]</sup>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. 仕様は 1) 標準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<sup>2</sup>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準化周波数以上の精度で制御した.

合成条件を示す. 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.

### 2.2 音響分析ソフトウェア

比較評価した音声分析ソフトウェアは, MDVP (Kay PENTAX社), Praat<sup>[3]</sup>(Amsterdam大学), L-Voice(以後LVと略記)<sup>[4]</sup>(宇都宮大学)である. 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, LV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HNR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.

## 3 結果

PPQの理論値に対する分析結果を図1に示す. PPQの理論値とは合成時に生成した周期系列から直接求めた値である.

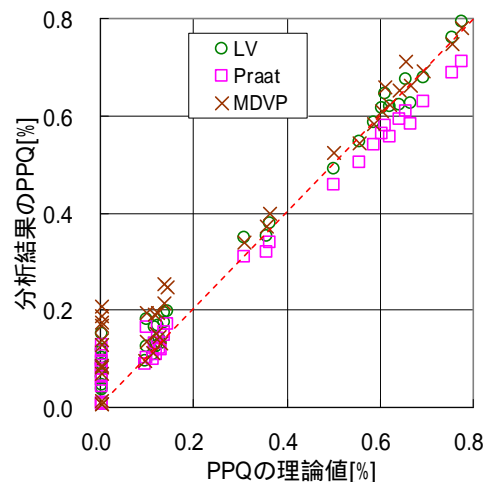


図1 PPQの理論値と分析結果

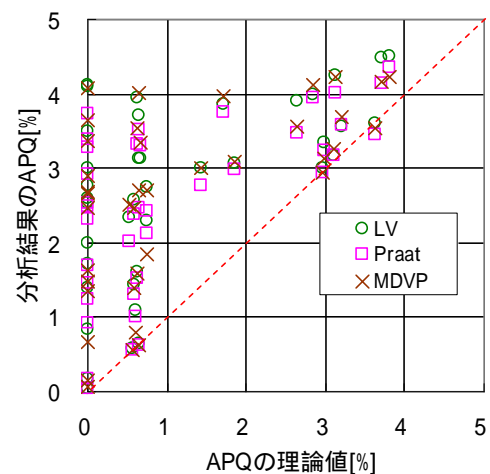


図2 APQの理論値と分析結果

\* Comparison of voice analysis software MDVP, Praat and L-Voice by using synthesized speech signal, by KIKUCHI, Yoshinobu (International Univ. of Health and Welfare), KASUYA, Hideki (Utsunomiya University), MORI, Hiroki (Utsunomiya University).

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、図からわかるように影響は少ない。APQ の理論値に対する分析結果を図 2 に示す。APQ の理論値も合成時の AV 系列から求めたものであるが、図からわかるように分析結果の 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 3 に HNR の理論値に対する分析結果を示す。HNR の理論値も合成時の調波信号と雑音信号から求めた。LV の  $NNEa[\%]$  は  $10 \times \log_{10}((100 - NNEa) / NNEa)$  により HNR [dB] に換算した。MDVP の HNR 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) HNR の理論値( $x$ )と MDVP で使用する帯域での HNR の理論値( $y$ )の関係式  $y = 0.9987x - 6.587$  を導き、 $x = y + 6.6$  と近似した。2) HNR の逆数の対数(dB 値)を求め 6.6 を加えた。

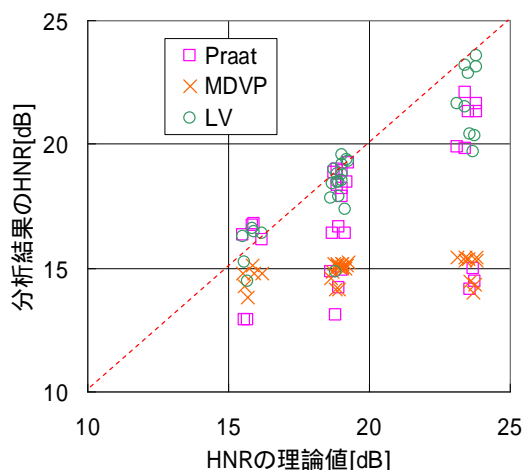


図 3 HNR の理論値と分析結果

#### 4 考察

図 1 ~ 図 3 に示したように、PPQ についてはソフトウェア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。図 4 に Praat の APQ に対する LV と MDVP のそれらの関係を示す。図 2 では広く分布していたが、図 4 では分析ソフトウェアによる差異が少ないことがわかる。このことから APQ はジッタや雑音の影響を強く受け、APQ の理論値とは大きく異なる。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 1 に分析結果の RMS 誤差を示す。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<sup>[1]</sup>。表より、PPQ に関しては 0.1 以上で LV が、APQ では Praat が、HNR では LV が最も RMS 誤差が小さい。

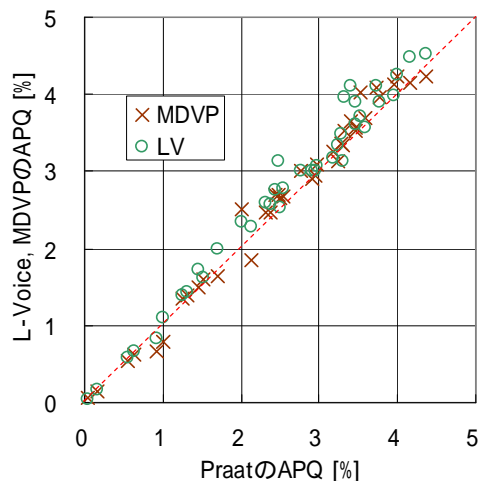


図 4 Praat の APQ との比較

表 1 各パラメータの RMS 誤差

RMS誤差		Praat	LV	MDVP
PPQ [%]	Whole	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 [dB]		3.724	1.514	5.518

#### 5 おわりに

ジッタ、シマ、及び加法的雑音量を変えて作成した 39 例の合成母音を用いて、3 種類のソフトウェアの PPQ、APQ、HNR の測定精度を比較した。PPQ と APQ では、RMS 誤差に違いが生じたが、実用上この違いが大きな問題になるとは思わない。しかし HNR では国際的に最も広く利用されている MDVP が LV 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。

謝辞

本研究の一部は科研費(22500147)によった。  
参考文献

- [1] 菊地, 粕谷, 森, 音講論, 2-9-9, 2011.
- [2] 大塚, 粕谷, 音響誌, (58)7, 386-397, 2002.
- [3] <http://www.fon.hum.uva.nl/praat/>
- [4] 朱達中他, 音講論, 1-8-7, 1995.