声年齢の評価に関する音声学的研究(2)1

筑波大学名誉教授 博士(学術) 城生佰太郎

1 前回発表2の要旨

声年齢の評価は、筋年齢や皮膚年齢の評価のように、単純には行かない。その理由のひとつに個人ごとに異なる感性の問題があるからである。そこで、本研究ではその主観的な感性を除いた部分で、音声学の視点から科学的な側面に限定した場合に、声年齢の評価に関してどのような提案ができるかを探ることを当面の研究目的とした。

前回までの暫定的な結果は、SPG(サウンド・スペクトログラム)を用いて、(1)データの「光り具合」、(2)高周波数域までフォルマントが伸びているか、の2点に注目することによって、○×△による簡便な判定方法が可能ではないかという見通しが得られたということであった3。そこで、今回は若干のデータを追加することによって、前回得られた結果の追験と新たな提案を行うことを目的とする。

2 今回の分析資料について

平尾登紀子代表が中心となって、2017年8月3日と5日に女性6名分の音声を録音した。 ただし、今回の被験者は朗読をある程度の期間体験した人たちのグループ3名と、そうでない人たちのグループ3名とに分けた人数構成とした。それぞれのフェイスシートは、次のとおりである。

a)長期経験者

MM 氏 78 歳 福岡県出身

FT 氏 69 歳 愛知県出身

ME 氏 70 歳 兵庫県出身

b)短期経験者

KT氏 56 歳 埼玉県出身

YS 氏 68 歳 埼玉県出身

AI 氏 74 歳 千葉県出身

¹ 本稿は、2017年12月28日に「げんごろう」本社において行われた口頭発表を文字化したものである。

 $^{^2}$ 2017 年 9 月 19 日に当研究会にて行った発表をさす。内容は、脚注 3 に記したジャーナルに公刊されている。

³ 詳細は、有限会社げんごろう発行の電子ジャーナル『声と健康に関する研究成果』第1号、(http://www.gengoro.net/) に掲載されている城生佰太郎(2017)を参照。

器材は、前回と同様 Olympus ボイストレック V-801 によって行い、分析資料もfilの母音に限定した。なお、サンプリング・レート 44.1KHz、16 ビットのステレオで録音された元資料を、鈴木修子が編集して音種ごとに切り分けてファイル化するとともにモノーラルに変換した。

3 分析方法

分析方法も、前回と同様に城生佰太郎が Multi speech 3700 を用いて行った。ただし、今回は SPG 表示を得る以外に「声のゆらぎ」と声量についても追加調査をすることとしたので、エナジー(音圧)とピッチ(高さ)に関しても、あわせて解析を行った。

なお、分析のもととなった音源はいろいろな状況下で収録されているので、音圧がバラバラである。このことは、後の段階で目視による判定を行う際の重要な根拠となる SPG の「光り具合」に甚大な影響を与えることが予測される。したがって、解析にかかる前にすべての音声資料を Cool edit 2000 にかけて音圧に対して 55%の正規化を施した。

4 分析結果

まずは、前回と同様に SPG の解析結果を示す。

4.1 MM 氏 (図1)

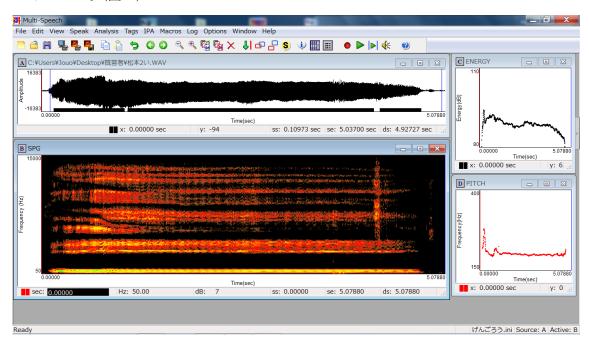


図 1

4.2 FT 氏

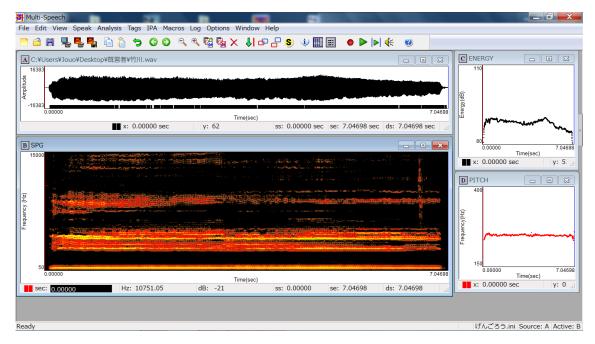


図 2

4.3 ME 氏

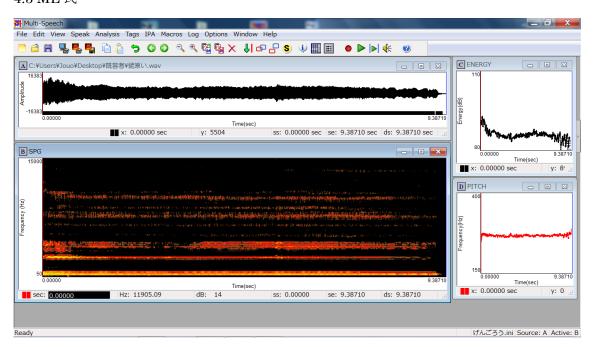


図3

4.4 AI 氏

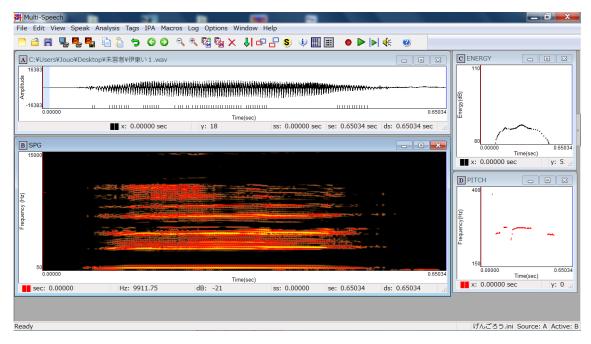


図 4

4.5 KT 氏

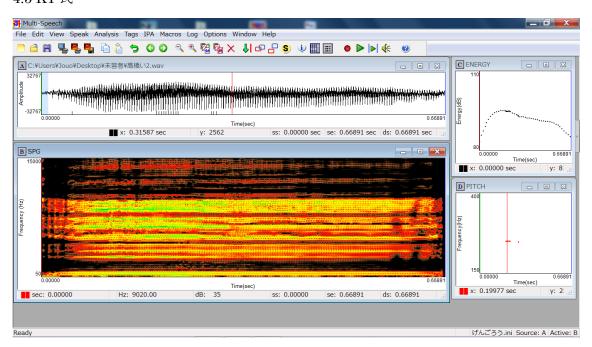


図 5

4.6 YS 氏

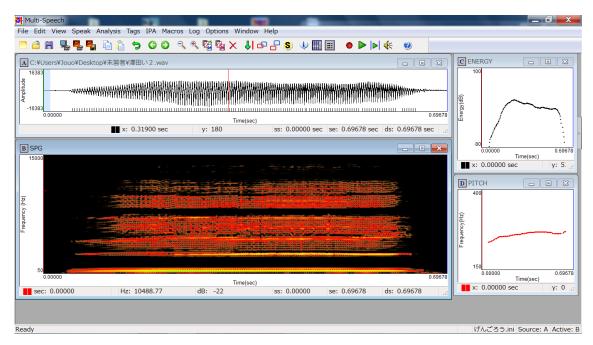


図6

4.7 統計データ(1)音圧

表 1	
dB	SD
87.60	4.58
87.27	3.29
85.31	2.75
81.08	7.65
91.54	2.97
91.62	6.14
	dB 87.60 87.27 85.31 81.08 91.54

左欄は平均値を、右欄は標準偏差値をそれぞれ示す。左欄の値が大きいほど音圧が大きいとみられる。一方、右欄の値は小さいほど安定した発声で、いわゆる声の「ゆらぎ」が少ないものとみられる。

4.8 統計データ(2)ピッチ

	表 2	
長期	Hz	SD
MM	203.21	12.4
FT	250.06	2.58
ME	269.49	3.33
短期		
ΑI	267.55	17.41
KT	270.51	19.84
YS	263.79	4.63

左欄は平均値を、右欄は標準偏差値をそれぞれ示す。左欄の値が大きいほど声が高いと みられる。一方、右欄の値は小さいほど安定した発声で、いわゆる声の「ゆらぎ」が少な いものとみられる。

5 考察

5.1 SPG を根拠とする解釈

前回の発表で暫定的に得られた目視による等級化という方法を用いると、今回の 6 名分のデータは次のようになる (図 7)。

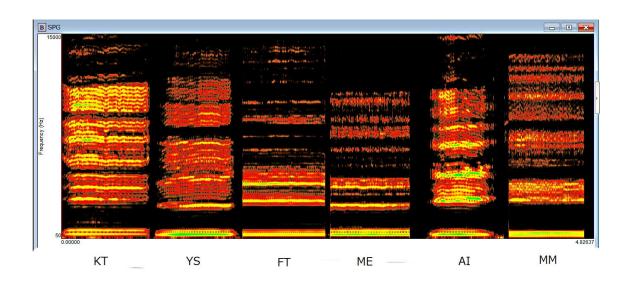


図 7

すなわち、(1)フォルマントが良く光っているか、(2)フォルマントが高周波数域まで伸びているか、の2点に注目した評価である。したがって、評価結果は

KT>AI>YS>FT>MM>ME

となる。ところで、この結果は先の第2節で述べた被験者情報とすり合わせると、意外とも思われそうな結果となっている。すなわち、朗読の長期経験者である MM,FT,ME の諸氏が、いずれも経験の浅い KT,AI,YS の諸氏よりも見劣りするということになるからにほかならない。

しかし、このことは実は驚くには値しない。すなわち、ヒトは生まれつき顔かたちが違うように、そして得意分野も違うように、それぞれ個性を持っているからにほかならない。つまり、端的にいえばもともと良い声を持って生まれた人は大した訓練をしなくてもそこそこの声を出すことができるということである。たしかに、訓練をすれば元の状態よりは良くなるであろう。しかし、そこには自ずと限界があり、その限界を超えてまでも訓練が有効であるとは言えない。

日本人は努力による美談を好むという精神主義者が多いので、特にだれにでも分かりやすいスポーツの世界ではこうしたことが好んで語り伝えられている。しかし、何事もがむしゃらに努力しさえすれば良いというものではない。己の限界を見極めるということもまた、科学による社会貢献の一面であることを強く主張しておく。

5.2 音圧とピッチ

前項の結論だけをみると、あたかも訓練はほとんど意味を持たないのではないかという早とちりな誤解を生じるかもしれない。音圧とピッチの統計データが示すものは、こうした疑問を払拭するものである。

まず音圧だが、表 1 に示したように標準偏差値を比較してみると訓練を積んだ人たちのグループでは $2.75\sim4.58$ (平均で 3.54)と値が小さいほうに収束しているが、訓練をそれほど積んでいないグループでは $2.97\sim7.65$ (平均で 5.58)と比較的拡散している。

同様にしてピッチのほうも、表 2 に示したように訓練を積んだ人たちのグループでは $2.58\sim12.40$ (平均で 6.10)と値が比較的小さいほうにまとまっているが、訓練をそれほど積んでいないグループでは $4.63\sim19.84$ (平均で 13.96)となっていて、ほぼ 2 倍の広がりを見せている。このことは、訓練によって声のゆらぎなどを技術的に逓減することができるという証拠にほかならない。

以上で明らかにしたように、本稿では、

- (1)声の良し悪しについては生得的なものがあることを率直に認める必要があること。
- (2)その上でしかるべき訓練を積めば、その人なりの進歩が期待できること

の2点を主張したことになる。

6 結語

前回と同様に、今回もケース・スタディを行ったに過ぎず、まだまだ「声年齢」の評価 に迫るためには日時を要する。しかしながら、今後とも基礎研究を積むことによって将来 的には光明が見えてくることは確かなことと信じている。

なお、今回の発表は、あらかじめ予定されていた発表者が個人的な都合によって十分な 発表準備ができなかったため、私がいわばその穴埋めとして、発表の1週間前に急きょ準 備したものである。このため、十分な吟味ができていない点をお詫び申し上げる。

【文献】

城生佰太郎(2017)「声年齢の評価に関する音声学的研究(1)」、『声と健康に関する研究成果』第1 号、有限会社げんごろう http://www.gengoro.net/「朗読ビジネス」コーナー