「声年齢」の評価に関する音声学的研究（１）

筑波大学名誉教授
博士（学術）城生佰太郎

１．目的

世の中では、いろいろな分野で実年齢とは異なる「〇〇年齢」というものが考えられている。具体的には、脳年齢、体年齢、肌年齢、筋年齢…などがあるが、たとえば筋年齢を評価する方法は、次に示す表 1 のようになっており、単純に数値化されているのでたいへんにわかりやすい。

表１　筋年齢

筋年齢	30 歳	35 歳	40 歳	45 歳	50 歳	55 歳	60 歳
腕立て	40 回	35 回	30 回	25 回	20 回	10 回	5 回
腹筋	40 回	35 回	30 回	25 回	20 回	10 回	5 回
懸垂	10 回	8 回	5 回	3 回	2 回	1回以下	―
握力（男）	35kg	33 kg	30kg	28 kg	26kg	24 kg	22 kg
握力（女）	28 kg	26kg	24 kg	22 kg	20kg	19 kg	18 kg

声についても、すでに何人かの人たちがさまざまな立場から「声年齢」にチャレンジしているが、いずれも医学的な色彩の濃い視点からの指摘であり、言語音に特化した音声学的視点からの試みはまだない。そこで、朗読をメインのテーマとして集まっている「朗読アンサンブルれもんの会」としては、言語音だけを徹底的に追究する音声学の視点から実験音声学の方法を用いることによって「声年齢」を評価する方法を科学的に模索することを目的とする。

なお、この研究を推進してゆくメンバーは、次のとおりである（敬称略）

代表者：平尾登紀子（有限会社げんごろう代表）
実働研究メンバー：城生佰太郎（筑波大学名誉教授）、高村めぐみ（愛知大学準教授）、桐越舞（大東文化大学非常勤講師）、鈴木修子（早稲田大学日本語教育非常勤講師）、平尾麻衣子（有限会社げんごろう）

２．方法

2.1.先行研究

ネット上には、「声タイプ診断」とか「声が老化する原因と対策」「声のアンチエージング」などの見出し語が躍っている。これだけを見ていると、さぞかし声年齢に関する研究は盛んに行われているのではないかと思われがちである。しかし、内容をみてみると、これらの先行研究のほとんどが医学系の視点に立った見解であり、必ずしも私たちの求めているものとは一致しない。

たとえば、一例を引けば、

> 声の老化の原因はさまざまですが、どれも声帯が関係していると言われています。
> 声帯は声をつくる部分の筋肉なので、何もしなければ足腰と同じように衰えていきます。
> 足腰が衰えると身体の柔軟性がなくなったり、激しい運動が大変になったりするものですが、声帯の場合は上記のような症状として現れます。
> また、声帯が老化すると声だけでなく体にも影響を及ぼします。
>
> ●声帯の老化が体に及ぼす影響
> 体に力が入らなくなる
> 食べ物が飲み込みにくくなる
> （http://josei-bigaku.jp/koenoanchieizingu3611/「女性の美学」より引用）

となっており、生理的な側面からとらえた老化に関する一般的所見に終始している。

しかし、私たちが求めているのはこのような実年齢に即した「老化」の見極めではなく、もっともっと言語音に特化した、声そのものに関する踏み込んだ観察であろう。たとえば、実年齢が 20 歳であっても、声だけを聞くと中年の人ではないかと思われるような事例とか、その逆に、いかにも若い人の声だと思われる人が、実年齢では 50 代であったなどという事例は、いったいなぜ起こるのか。そのなぞに迫ることができるような科学的な方法論こそが、本研究の取るべき道であろう。

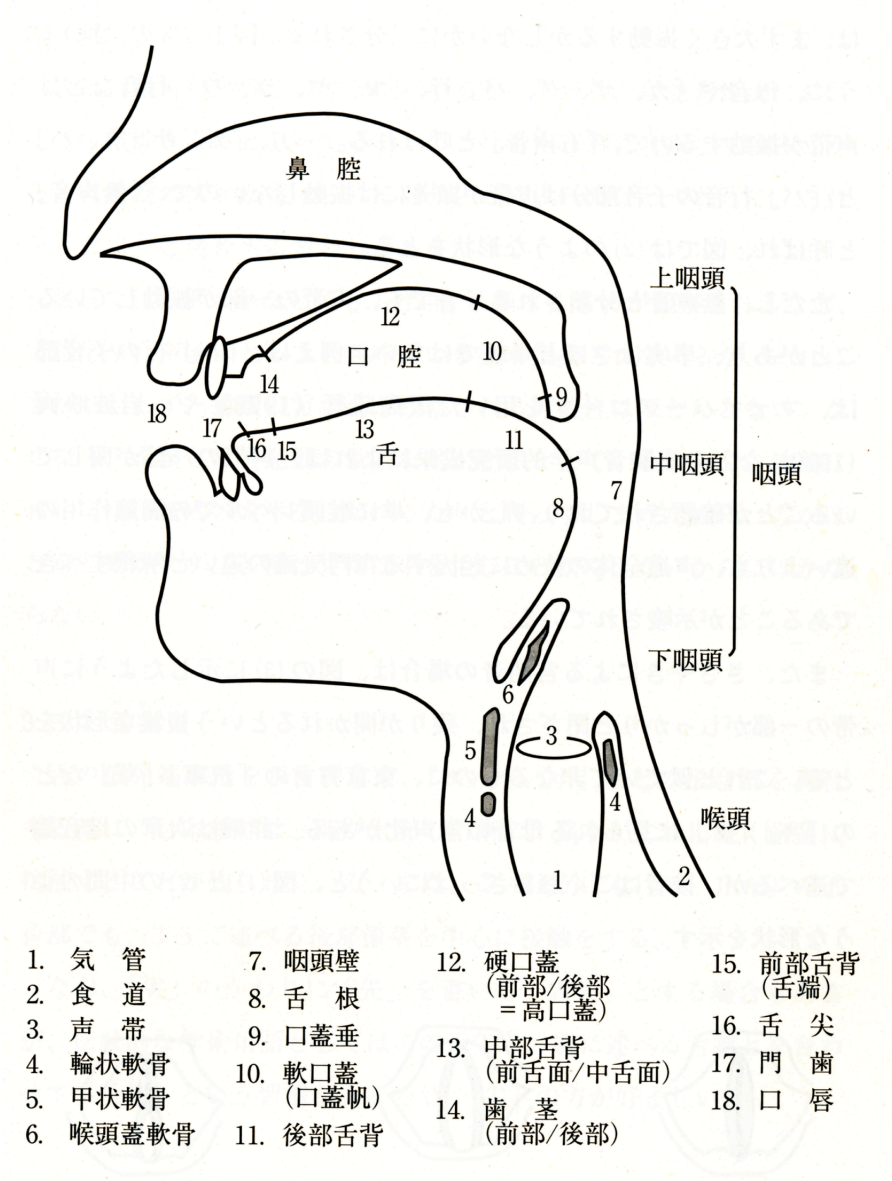
2.2.実験音声学的方法

本稿で用いられる方法は、フランス人のカトリック神父 Rousselot(ルスロ)によって 1889 年にパリで創始された La phonétique expérimentale(実験音声学)である。カトリック神父が実験音声学を創始したという社会的背景には、中世のヨーロッパで多発していた構音障害の診断と治療に際し、カトリック教会も献身的な活動をしていたという事情がある。ただし、医学的な方法ではなく、こちらはまさに文化系的方法とセンスが中核となって発展してきた学問である。

今日では、実験音声学の中身が細分化されているので、ここにそのすべてを述べることは本稿の目的に照らし合理的ではないので、本稿ではそれらのうちからサウンド・スペクトログラム(以下 SPG と略)を用いた音響音声学的研究に限定するということだけを銘記しておく。

2.3.発声と調音

ヒトの声は、肺臓から流れ出る「呼気」によって声帯を振動させ、これが口腔の中で舌・口蓋・口唇・鼻腔などの諸器官によって共鳴することで言語音となる（図１）



（城生佰太郎 2008:20 より転載）

図１　音声器官図

ここで重要なのは、(1)発声と(2)調音をしっかりと区別することである。

(1)の発声というのは、文字通り声を出すことである。普通は、母音の「ア」を出しながら声の調子をつかむことが多い。ちなみに、「ス」「ツ」「ヒ」などの子音部分を使って声の調子を調べることはできない。なぜなら、これらは音声記号では[s][ts][ç]などと表記される音で、声帯振動を伴わない無声音だからである。

(2)の調音というのは、声帯を振動させて声となった音波を、口腔の中で舌・口蓋・口唇・鼻腔などの諸器官によって共鳴させ、独特の響きを持った言語音を作り出すことである。

人には、個人差というものがある。外見で考えるのが最もわかりやすいが、人はみな顔かたちが違う。そのおかげで、A さんと B さんや C さんを区別することができる。言語音もこれと似ていて、明確に個人差がある。親しい人からの電話なら、はじめの「モシモシ」だけでも話し手がだれなのか、見当がつくのはこのためである。この「個人差」は、上に述べた(1)のレベルにも(2)のレベルにもある。すなわち、声そのものの違いと、調音の仕方の違いである。

ここで確認をしておくと、本研究で私たちが追究するターゲットは上に述べた(1)だけではなく、これに(2)を加えた総合判断を行おうということである。

2.4.SPG(サウンド・スペクトログラム)とは

こんにち、言語音の音響音声学的研究にもっとも多く誓われているのが SPG である。実際の SPG のサンプルを、図 2 に示す。

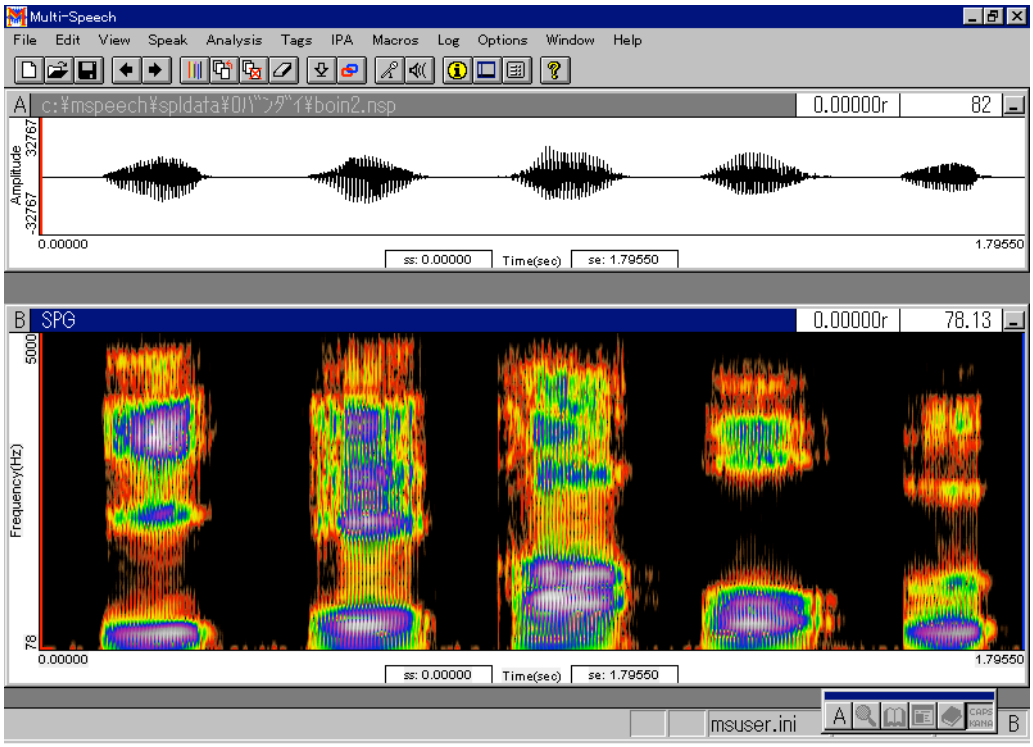


図 2　SPG による日本語 5 母音

下段にカラー表示されている図は、日本語の「イエアオウ」5 母音を解析した SPG だが、縦軸が共鳴周波数(Hz 表示)、横軸が時間長(ms 表示)、色の濃淡が共鳴の強さをそれぞれ表している。なお、上段に黒い波形で描かれているものはアナローグ表示による原波形である。また、いずれも左から右へと時系列上に要素が配列されている。言語学の音韻論や音声学の音響音声学で一般的に利用されているのは、これらの音響情報のうちの一部であり、結論的にいえば、F1 と F2 の相対的な位置関係だけが重視されている。

なお、F1 というのは「第 1 フォルマント」の略であり、同様にして F2 というのは「第 2 フォルマント」の略である。SPG 上では、これらのフォルマントは横軸方向に太く濃く安定して一定時間伸びている帯状のパタンによって表されている。また、この図では「濃く」というのは色分けされたもののうち、白っぽくなっている部分が最も「濃く」、次いで紫、青、緑、黄、橙、赤、の順に薄くなって行くものと約束されている。

そこで、図 2 に示された左端の「イ」に関するフォルマント情報を読み取ってみると、もっとも低い位置に白から紫を含んだカタマリが確認できる。これが、F1 である。この図だけからでは細かい数値は読み取れないが、ほかの表現方法を用いるとほぼ 320Hz であることがわかる。次に、その上を見ると先ほどのよりはいくぶん見劣りがするが、やはり紫を中核とする帯状のカタマリが見える。これが F2 である。F1 と同様に他の方法で定量化するとおよそ 2200Hz であることがわかる。

このような操作を、順次「エ」「ア」「オ」「ウ」に行ってゆくと、結果として次のような F1 と F2 の相対的な位置関係を簡素化した図が得られる(図 3)。これを、F1・F2 模式図という。なお、図 4 に参考までにフォルマント数値も示しておく。

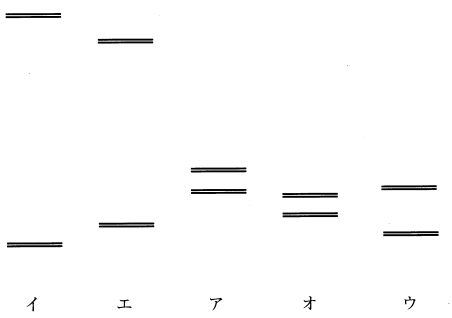


図 3　F1・F2 模式図

	F_1	F_2
イ	320	2200
エ	440	2100
ア	800	1200
オ	440	800
ウ	330	1100

図 4　フォルマント数値(Hz)

以上で明らかなように、F1 と F2 の情報さえ押さえておけば、5 母音の音響的なクラス分けが可能になる。すなわち、「イ」「エ」「ウ」は比較的 F1 と F2 が離散しているタイプ（散音型）であり、逆に「ア」「オ」は互いに近づくタイプ（緊音型）であるとまとめることができる。また、F1 の動きに注目すると、「イ」「ウ」で最小、「ア」で最大になっており、「エ」「オ」は両者の中間にあるところから、生理面との対応では開口度（下顎の開き）と F1 が

相関を持っていることがわかる。すなわち、下顎が開けば F1 は上昇し、反対に下顎が閉じれば F1 は下降するということである。

いっぽう、F2 のほうは、「イ」で最高値を示し、その後は「エ」→「ア」→「オ」の順に下降し、最後に「ウ」でほんの少し上昇している。このことは、舌位置と関係がある。すなわち、舌が前寄りに構える「前舌母音群」では F2 は高めに出るが、舌が後ろ寄りに構える「後舌母音群」では、F2 は下降するということである。

2.5.由紀さおりの声を科学する

声の問題と言うと、すぐに思いつくのが歌声である。筆者も、かつて何度もテレビ局から歌声の分析と評価にかかわる番組の出演依頼を受けたことがあるが、ここでは比較的新しく出演した「クローズアップ現代」（通称クロゲン）でおこなった由紀さおりの声の分析について簡単に述べることとする。

この番組は、「世界を魅了する日本の歌謡曲：由紀さおり　ヒットの秘密」と題して、2012 年 2 月 21 日に NHK 総合テレビで放送されたもので、ビデオ・リサーチの調査では視聴率 15.3%という、この手の番組としては異常な高さであった。それだけ、多くの人々が声の問題に関心を持っているという証拠でもあろうかと思われる。

さて、声の良し悪しなどと言うと、必ずからんでくるのが好みの問題である。たとえば、「あの人の声は確かに良く響くが、かえってうるさい」とか、「透明感は感じられるが、冷たい」というような評価がつきまとうのである。歌を趣味として鑑賞する際にはこのような視点こそが重要である。しかし、私たち音声学者はこういった観点から音声の問題を取り上げない。なぜなら、それらは個人個人に備わっている感性の問題だからであり、科学とは一線を画すべき性質のものであると考えるからである。すなわち、科学とは一定の手順さえ踏めば、いつ、だれが、どこで行ってもほぼ同じ結果が保証されるもの（これを「再現性の保証」という）でなければならないからである。

この点で、私がクロゲンの番組で採用した方法は、音響データ上に現れた「共鳴の良さ」であった。では、その「共鳴の良さ」をどのようにして見極めるのか？　ここで活躍するのが、SPG である。

具体例として、由紀さおりのヒット曲のうちから選んだ「パフ」の 1 節を取り上げる。

＜音楽入る＞

ところで、いま聞いていただいた部分をすべて SPG で音響解析してしまうとやや長くなるので、ここでは前半の部分だけをカットしてもう一度お聞かせする。

＜短尺版音楽入る＞

図 5 は、以上の音源を SPG によって解析したものである。

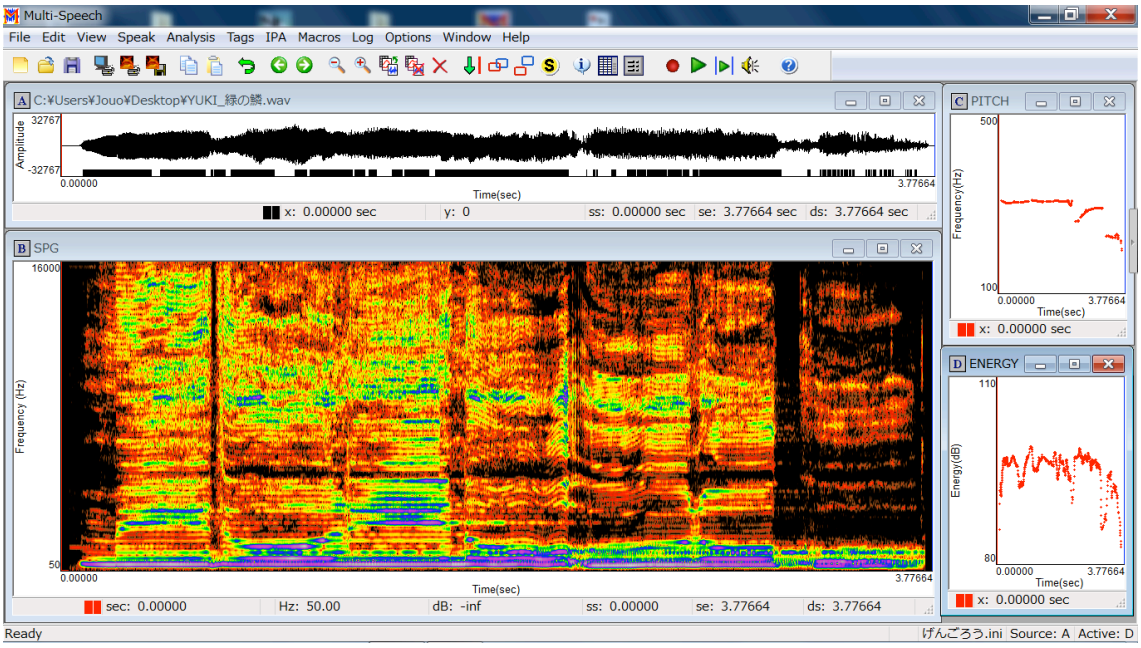


図 5　由紀さおり「パフ」より(1)

「♪みどりのうろこ　鳴らして泣いた♪」という部分だが、まず中央にカラー版で出ている SPG を見て驚異的なのは、上限を 16KHz に取ってあるにもかかわらず、ほぼ上のほうまでフルにフォルマントが濃い色で「ピカピカと光っている」ことである。先ほど説明したように、フォルマントの強い部分は濃い色になるのだが、常人ではこんなに高い周波数域までしっかりと濃く出るなどということはまれであり、さすがは由紀さおりの声は一流であることを立証している。しかも、実年齢を考え合わせるとこのデータはまさに驚異的と言わざるを得ない。

次に、右端の上には 1 秒あたりの声帯振動回数を表すピッチ（基本周波数）が取ってあるが、「みどりの」に該当する部分がほぼ平らで安定しているという点も、さすがである。なぜなら、言語音と歌声との決定的な差異は、歌声が、

(1)個人の有する声域のほぼ全域を使う
(2)ピッチが一定区間で安定している

というのに対し、言語音は

(3)個人の有する声域のうちほぼ低いほうだけを使う
(2)ピッチはたえず微量の変動を示し、一瞬たりとも同一のピッチを持続しえない

という点に、それぞれの特徴があるからにほかならない。なお、ついでのことに、右端の下に示してあるのは物理的強度を示す intensity だが、これも「みどりのうろ」までに該当する部分ではほぼ一定の強度が保たれており、さすがはプロの発声であると感心させられる。なお、中核的な SPG による表示を、図 6 に示した拡大図によってもう一度確認してみる。

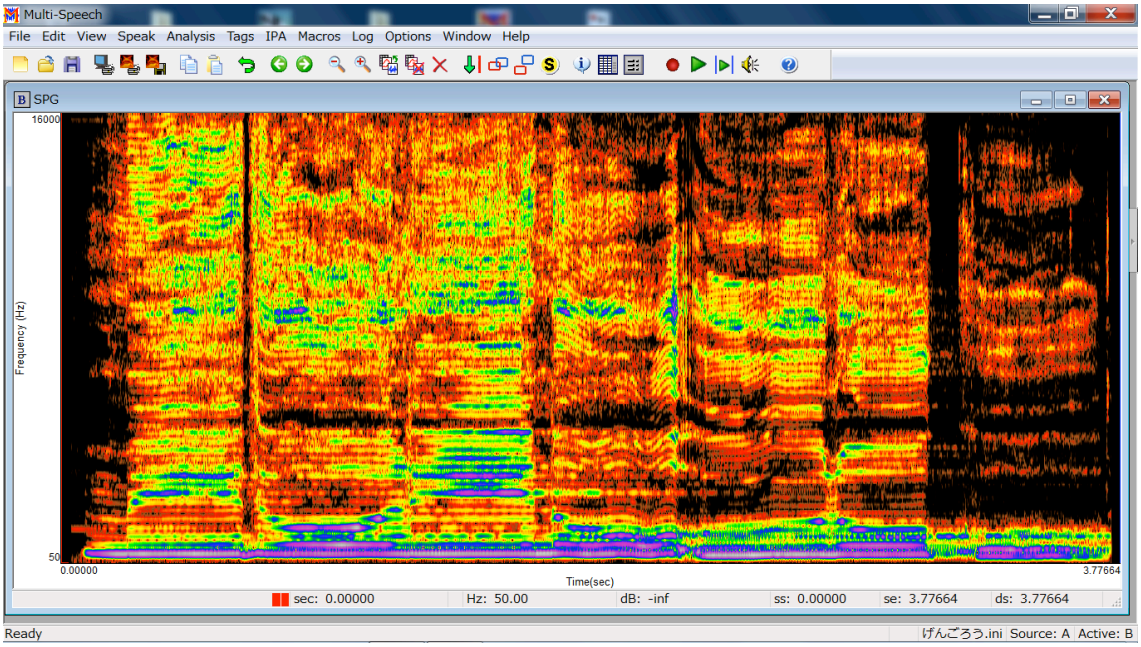


図 6　由紀さおり「パフ」より(2)　拡大図

今度は、左端に出ている周波数の目盛の上限が 16KHz になっているのが読める。また、一番左に出ているカタマリは「み」に当たる部分なので、母音フォルマントは[i]であることがわかる。さらに、このデータから、日本語の母音の中では [i] の共鳴がもっとも良いことが窺知される。以上の結果から、「共鳴の良さ」に注目することによっても声の良し悪しを科学的に評価することが可能ではないかと考えられる。

2.6.言語音を対象とする評価方法

前振りがかなり長くなったが、ここからが本研究に直接関係する部分となる。今年の初夏に、有限会社げんごろうの代表平尾登紀子が、自らが担当する朗読教室において「朗読アンサンブルれもんの会」の皆さんのご協力を得てボウリング調査用の録音をとり、これを城生佰太郎が Multi speech 3700 を用いて試験用の音響解析を行った。本日ご報告するのはこの時に収録した音声資料による。

2.6.1.録音の手順と被験者

録音は、Olympus ボイストレック V-803 により、2017 年 6 月 24 日に NHK 教室にご出

席の皆さまの中から以下の８名の方々にご協力をいただき、平尾登紀子が音声を収録した。被験者となっていただいた諸氏のフェイス・シートは、次のとおりである。

表２　被験者

資料番号	イニシャル	年齢	言語形成地
①	NF	64	秋田県小坂町
②	MS	70	神奈川県横浜市
③	NM	70	東京都渋谷区
④	MC	70	埼玉県北埼玉
⑤	YK	44	埼玉県浦和市
⑥	MN	64	神奈川県横浜市
⑦	AK	73	東京都杉並区
⑧	MK	83	静岡県磐田市

なお、これらのうち⑦の AK さんだけが男性であり、あとはすべて女性である。

2.6.2.調査票

今回のボウリング調査用の調査票は、以前に城生佰太郎が作成した全６項目から成る「実証実験用音読データ」のうちから、3 項目に絞ったものを使用した。実際に用いた調査項目は、以下のものである。

１．発声（母音を長く引く）

イーーーーー、イーーーーー、イーーーーー。
エーーーーー、エーーーーー、エーーーーー。
アーーーーー、アーーーーー、アーーーーー。
オーーーーー、オーーーーー、オーーーーー。
ウーーーーー、ウーーーーー、ウーーーーー。

２．発話速度（できるだけ早口で言う）

発泡スチロール、発泡スチロール、発泡スチロール
赤巻紙、青巻紙、黄巻紙。　赤巻紙、青巻紙、黄巻紙。
こじゅうと（小姑）の灯油事故。こじゅうとの灯油事故

３．無理のない範囲で大声を出す

そんなことって、ないでしょうーーーー？？　↑　（思いっきり低→高へ）
そんなことって、ないでしょうーーーー！！　↓　（思いっきり高→低へ）

2.6.3.音響解析

2.6.1 で述べた器材によって収録された音響データは、すべて音声ファイル形式変換ソフト Digion Audio 2 によって wave 形式のファイルに変換し、これを米国 KAY 社製音声解析ソフト Multi speech 3700 を用いて解析した。サンプリング・レートは 44.1KHz、量子化 16 ビットのモノーラルである。なお、Multi speech 3700 の入力はモノーラル仕様となっている。しかし、wave 化された音源はステレオになっていたことと、収録時の音圧に多少のばらつきがあったことの 2 点を解消するため、音声編集用ソフト Cool edit 2000 を併用して音源をモノーラル化するとともに、正規化をかけて入力レベルを一定にした[1]。

今回は、先にも述べたように全くの初めての試みなので、デモ用として上記表 2 の被験者が調音した「イ」だけを取り上げて、目の付けどころについて論じることとする。なお、本稿で用いた記録図は SPG のみであり、物理的な高さを表すピッチや強さを表すインテンシティーについては、次回以降の検討事項とする。また、評価に際しては主として目視によるアナローグ評価を採択している。ただし、繰り返しになるが今回が全くの初めての試みになるので、今後とも同一の評価方法を続けて採択して行くかどうかは、現段階では不明である。

2.6.4.有識者の目視による SPG の評価

SPG から読み取るべきポイントは、先に述べた由紀さおりの声の分析で明らかなように、暫定的に高次フォルマントの「光り具合」に求めることが考えられる。ただし、この「光り具合」というのを目視のみによって求める限りはある意味で主観的な評価基準であるとも言える。

科学論文では一般に、「主観的な観察」は悪であり、「客観的な観察」こそが善であるという発想が特に理科系を中心とした常識としてまかり通っている。これは、あたかもマクドナルドの接客マナーのように、画一的な視点で徹頭徹尾押し通してしまおうとする発想であり、一面で合理的であるかもしれないが、問題がないわけではない。なぜならば、「客観的な観察」を認知させるために取られる方法が、多くの場合一定のまとまった数量を背後に抱えた統計処理に代表される「ディジタル至上主義」に支えられているために、多くの時間と労力を必要とするからにほかならない。

いっぽうで、いわゆる文化系的発想では、物事をできるだけ多面的にとらえようとする。その際に、ヒトが生得的に持っている感覚尺度がものをいうことになる。つまり、これがいわゆる「主観的観察」ということで、先に挙げた「ディジタル至上主義」に対して「アナローグ擁護主義」とでも名付けておけばわかりやすい。この、文化系的発想による「主観的観察」の正当性に関しては、たとえば中国における前漢の時代から知られている、

---

[1] 正規化の値は、全データ一律で 85%にした。

十目の視る所十手の指す所それ厳なる哉　　―礼記―

などの文言が良い例であろう。すなわち、多くの人が認めることにはそれなりの真理が宿っているという至極妥当な指摘であり、幾世代にもわたって耳を傾けるべき人生哲学的発想であるとして重視されている。

ここで、もうひとつディジタル至上主義に対する反例を挙げておこう。たとえば、幼稚園の運動会では定番となっている「玉入れ」だが、紅組と白組とでどちらのほうにより多くの玉が入っているかで勝敗が競われることは周知のとおりである。さて、この競技では必ず最後に一つずつ玉を勘定しなければおしまいにはならない。この発想は、いかにもディジタル至上主義である。なぜならば、見ただけで赤のほうが圧倒的に多いとか、逆に白のほうが多いなどというケースが決して珍しくもないにもかかわらず、多くの手間暇をかけて一つずつ数え上げてゆくというのは、いかにも不合理な行動だからである。つまり、目視による主観的な観察であっても、「十目の視る所十手の指す所それ厳なる哉」が成立するような状況ならば堂々とアナローグ擁護主義を高く掲げて、短時間で手間のかからぬ合理的な処理をするという知恵もまた、時には有効であるということである。ただし、赤白拮抗していて、見た目ではどちらか勝敗を決しがたいような場合もある。ディジタルが必要とされるのはまさにこういう特殊な場合に限る。要は、状況の見定めなのであって、考えることをせずに機械的に何でもかんでもディジタル主義で押し通すという現代の科学論文のありかたには苦言を呈しておきたい。

というわけで、今回はSPGに慣れている有識者に限定して、目視による「光り具合」に関する意見が一致するものなのかどうか、この点についての予備的な調査を行った。方法としては、2.6.1に述べた被験者のうちから男性話者⑦を除外した7名分のSPGを同一の条件でビットマップ・ファイルにして5人の有識者に配信し、それぞれ目視による「光り具合」に序列をつけてもらった。その結果は、後述する4で述べるが、結論的にいえばそれほどのばらつきもなく、ほどほどにまとまっていたといえる。

３．結果

以下に、SPGを中心にして実験結果を述べる。1件ずつ音声も流すが、まずは目視によって次の2点を確認していただきたい。なお、資料番号⑦は黒一点の男声であったので、等質性の確保から今回は省いてある。

1)SPGの目盛の上限は15KHzに取ってある。先に見た由紀さおりとは比べ物にならないが、高周波数域がこの数値を基準にしてどこまで伸びているのかに注目することによって、「声年齢」の判断材料とする。

2)SPGの「光り具合」に注目する。2.4で述べたように、フォルマント成分の強さは、白→紫→青→緑→黄→橙→赤の順になっているので、青系に寄ったあらわれをし

ているのかどうかで、ある程度「声年齢」の判断材料とすることができる。

3.1.資料番号①の[i]

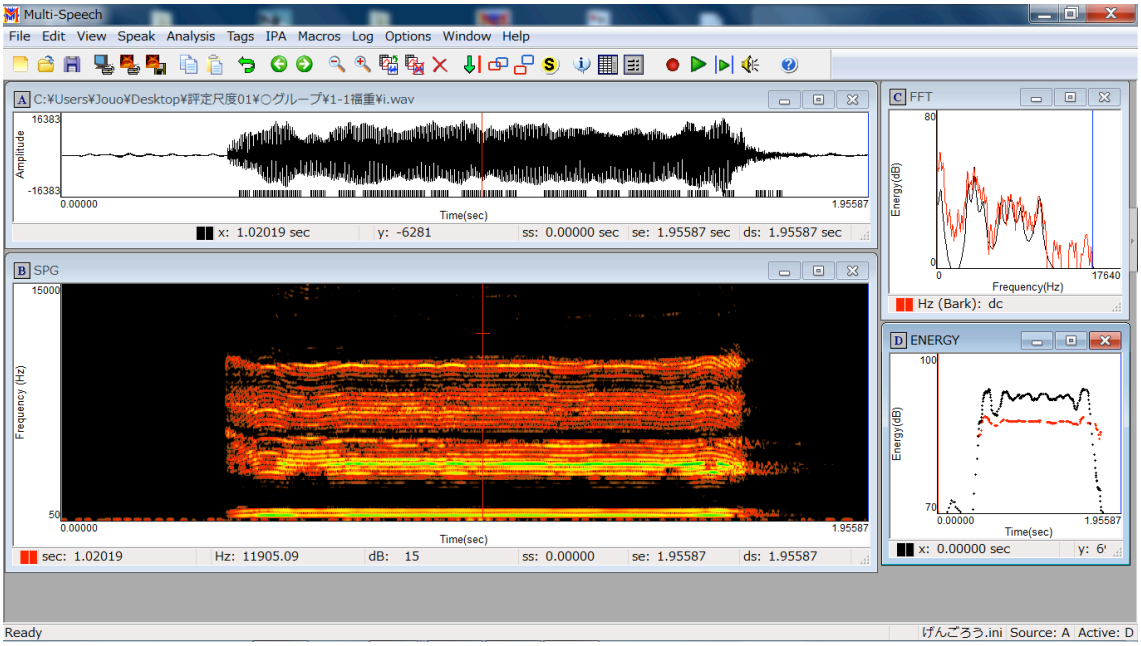


図 7　資料番号①の[i]

3.2. 資料番号②の[i]

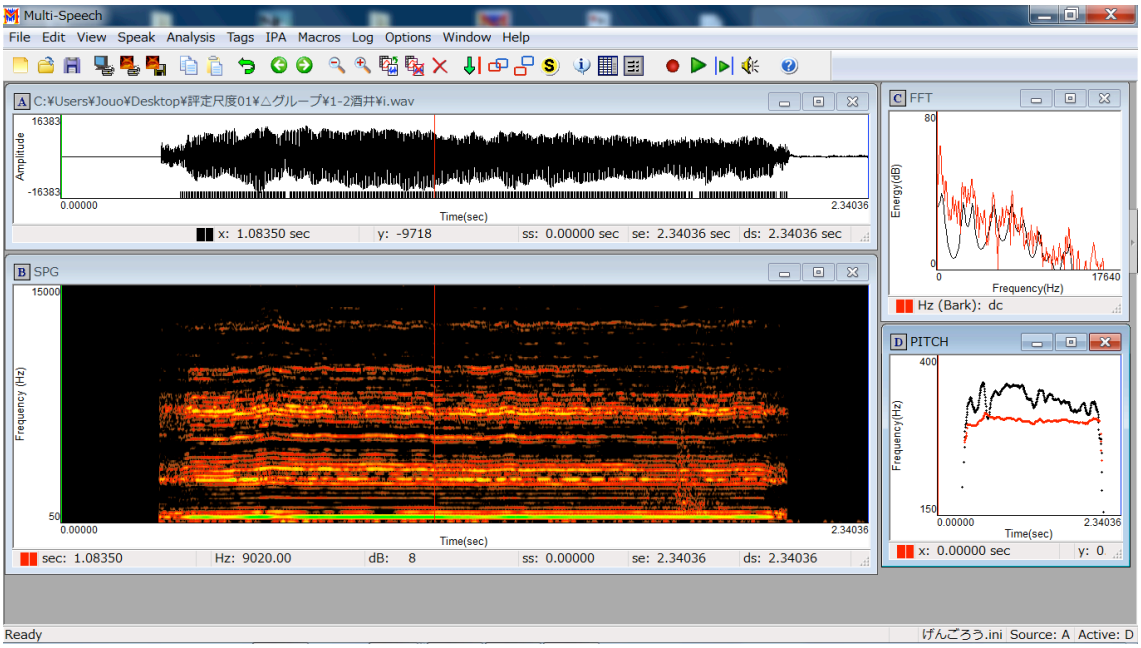


図 8　資料番号②の[i]

3.3. 資料番号③の[i]

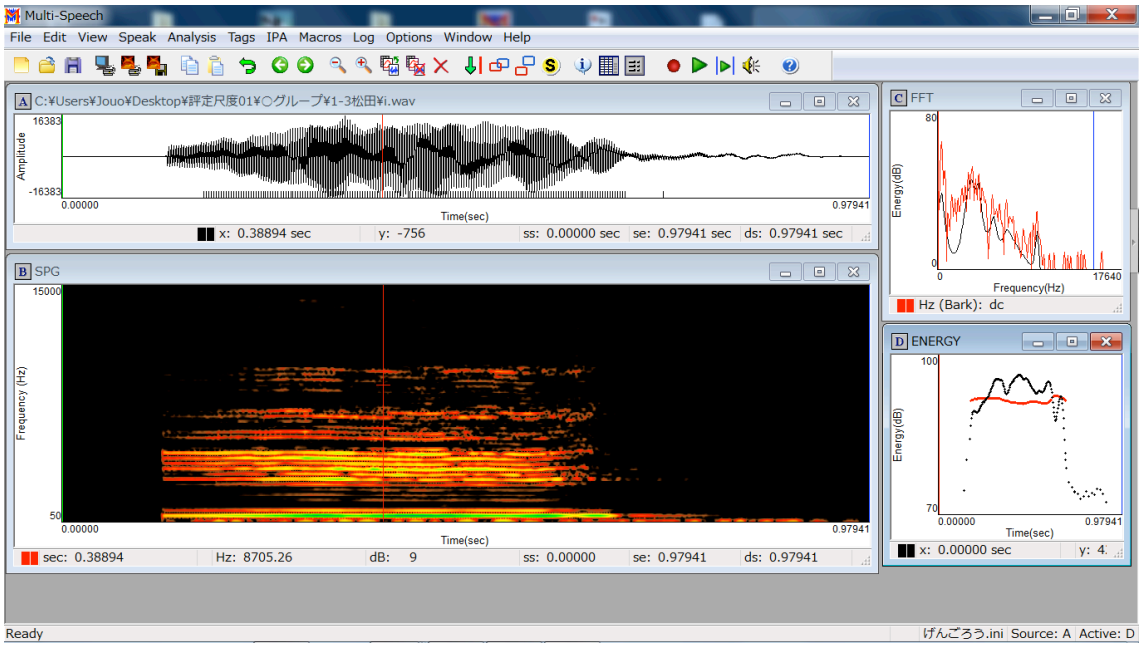


図 9　資料番号③の[i]

3.4. 資料番号④の[i]

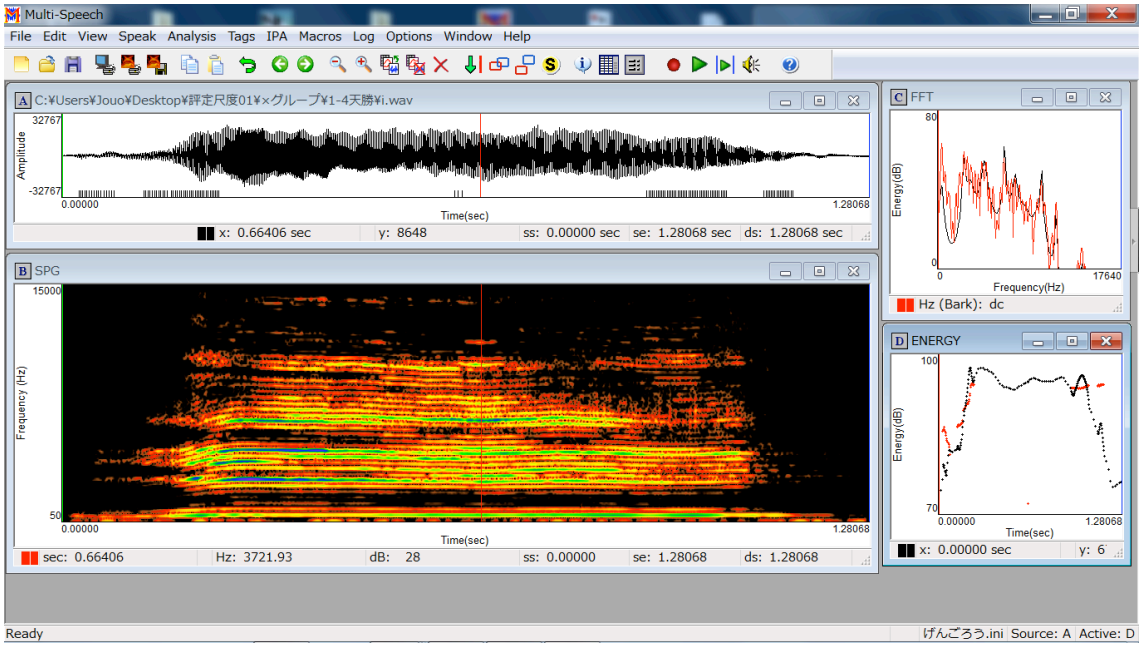


図 10　資料番号④の[i]

3.5. 資料番号⑤の[i]

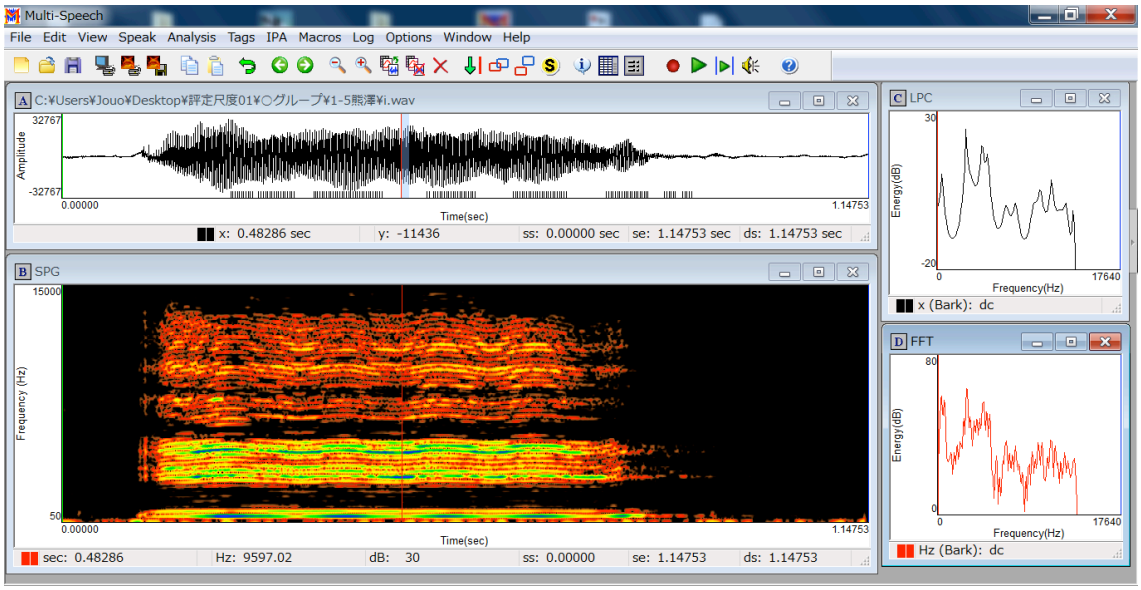


図 11　資料番号⑤の[i]

3.6. 資料番号⑥の[i]

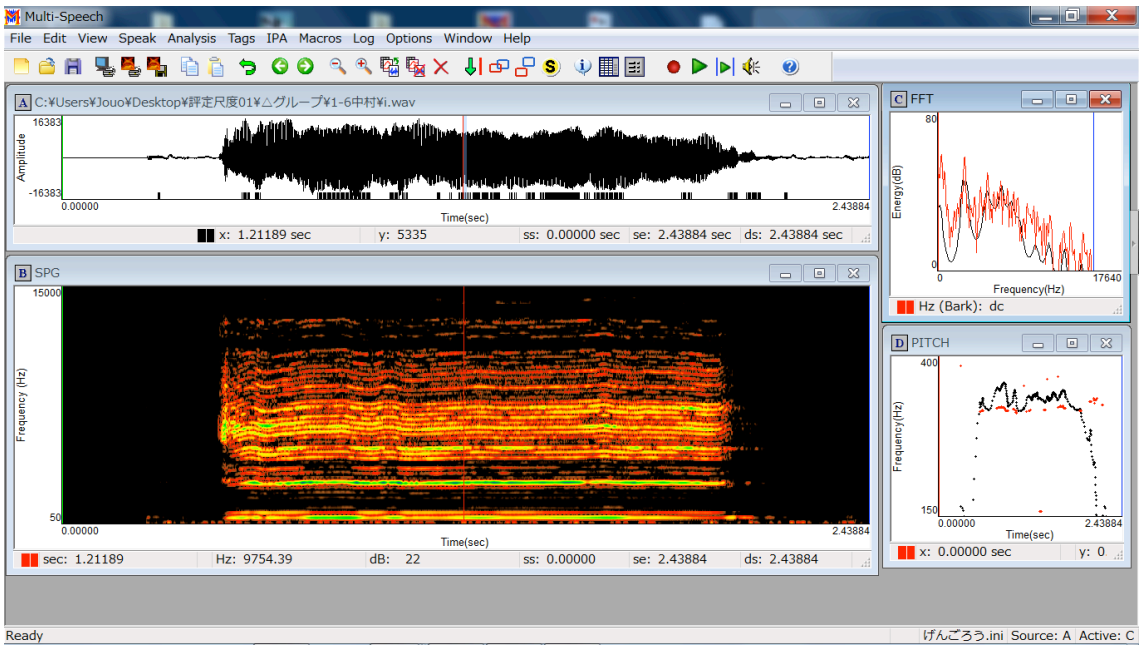


図 12　資料番号⑥の[i]

3.7. 資料番号⑧の[i]

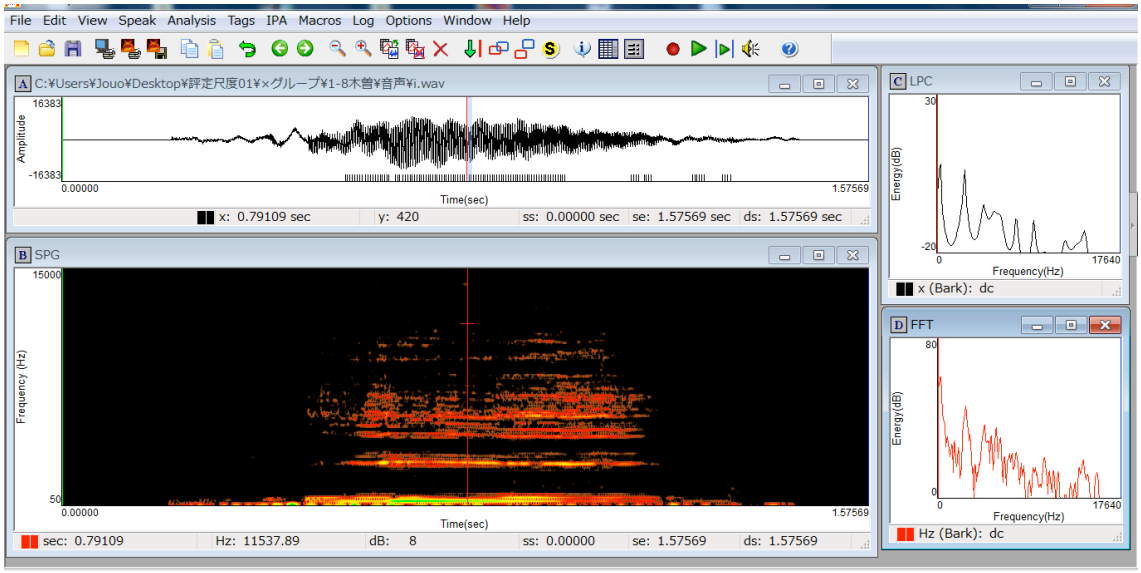
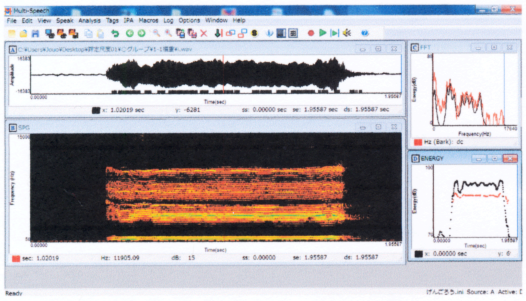
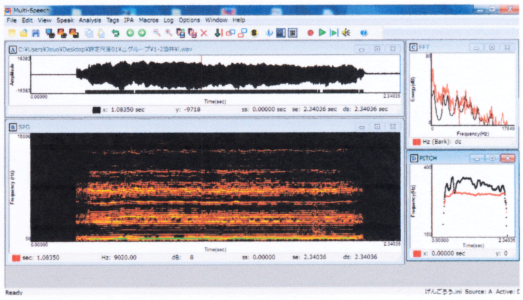


図 13　資料番号⑧の[i]

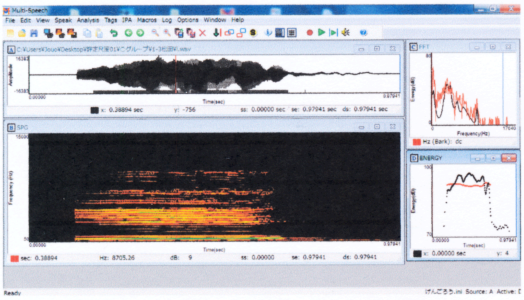
これらを 1 件ずつ見ていると、細部に関しては気づきやすいが、大まかに見て「良い声グループ」に属すのかどうかということは、必ずしも分かりやすいというわけではない。そこで、次に上に示した 7 枚の図を図 14 で一望のもとに並べてみる。



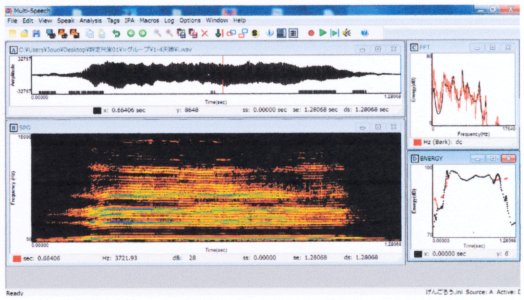
1-1_i



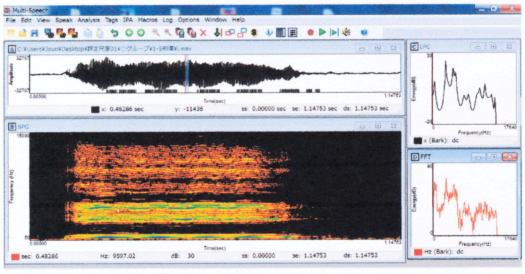
1-2_i



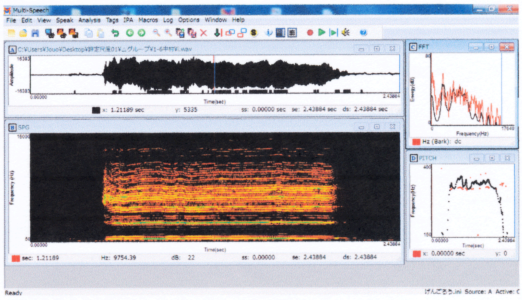
1-3_i



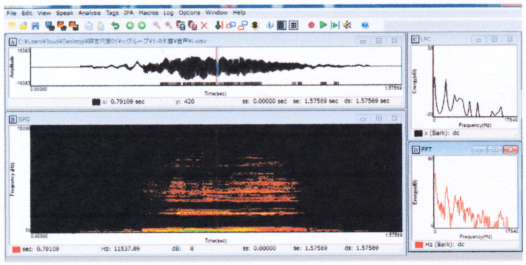
1-4_i



1-5_i



1-6_i



1-8_i

図 14　比較図

４．考察

4.1. 資料番号①の[i]

まずは、本稿における声年齢を評価する際の目の付けどころだが、SPG 上で

(1)フォルマントが「ピカピカと輝いている」か？

(2) 高次フォルマントに至るまで総じてフォルマントが良く出ているか？

という 2 点に求めている。このことを音響的な側面から表現すれば、(1)は音響的なパワーの強さを含意し、(2)は声道内における共鳴の良さを含意している。

したがって、これら(a)(b)の 2 点をクリアしている声には「○」を、これらのうちどちらか一方をクリアしている声には「△」を、そしてこれらのうちどちらもクリアできなかった声には「×」を与えることとする。

以上の基準に照らすと、図 7 に示した資料番号①は上記の(1)に関しては問題ない。しかし、(2)に関しては高次フォルマントに着目すると、残念ながら上方 1/4 程度が真っ黒になっていて反応していない。このことは、音響的な観点からは共鳴がそれほど優れてはいないということを含意する。したがって、総合評価としては「△」とする。

なお、右端にも上段と下段に図が示してあるので、それらのうちから右端上段の図について触れておく。ここに示されたものは、LPC（黒のライン）と FFT（赤のライン）である。このうち黒の LPC とは、linear predictive coding(線形予測符号化)の略で、入力された音声からフォルマントを予測することができる便利なアルゴリズムである。この図では、スペクトル包絡（黒のライン）を山のように束ねている大きな塊が 4 個確認できる。もちろん、SPG を目視しても結果は同じで 4 本となるが、慣れないうちは幅の広いフォルマントを全体で 1 本と数えるのか、それとも半分で切って 2 本と数えるのか迷うこともある。そのような際に、このアルゴリズムはとても有益である[2]。

なお、参考までに、この資料番号①による調音を聞いていただきたい。

＜音声入る＞

4.2. 資料番号②の[i]

図 8 に示した資料番号②は、

(1)フォルマントの「輝き」が見られない

(2)ただし、フォルマントの出方と本数は SPG の目視からだけでは 5 本と判断すべきか 6 本とすべきか、迷うところである。また、高次フォルマントも出てはいるが

[2]重ね描きしてある赤のラインは FFT(fast Fourier transform=高速フーリエ変換）と呼ばれるアルゴリズムだが、ここでは省略する。

薄い。そこで、上に述べた 4.1.と同様に LPC を援用する。そうすると、山となってあらわれているピークの数は 6 本であることが確認できる。

以上の理由で、評価としては「△」とする。
なお、参考までに、この資料番号②による調音を聞いていただきたい。

＜音声入る＞

4.3. 資料番号③の[i]
図 9 に示した資料番号③は、

(1)フォルマントの「輝き」が足りない
(2)特に高次フォルマントを中心とした出方が良くない

とことになる。したがって評価は(1)(2)ともに満足していないので「×」となる。
なお、参考までに、この資料番号②による調音を聞いていただきたい。上に指摘した以外にも、発声にところどころ不安定さがみられ、いわゆる「うねり」が観察されるが、この点も音響データには出ている。右端の下側に示された赤で表示されているドットはピッチを表しているのだが、この出方が上下に微動しており、資料番号③のピッチが不安定であることを示している。
なお、参考までに、この資料番号③による調音を聞いていただきたい。

＜音声入る＞

4.4. 資料番号④の[i]
図 10 に示した資料番号④は、

(1)フォルマントの「輝き」が見られる
(2)高次フォルマントまで良く出ているが、さすがに 15KHz 近傍では、かなり薄くなっている。また、フォルマントの本数も、LPC とのすり合わせによって 6 本と判定できる

という結果となるので、(1)(2)ともに満足しているので評価は「○」となる。
なお、参考までに、この資料番号④による調音を聞いていただきたい。

＜音声入る＞

4.5. 資料番号⑤の[i]

図 11 に示した資料番号⑤は、

(1)フォルマントの「輝き」が見られる

(2)高次フォルマントまで良く出ており、15KHz 近傍まで迫っている。また、フォルマントの本数も、LPC とのすり合わせによって 7 本と判定できる

という結果となるので、(1)(2)ともに満足しているので評価は「○」となる。

なお、参考までに、この資料番号⑤による調音を聞いていただきたい。

＜音声入る＞

4.6. 資料番号⑥の[i]

図 12 に示した資料番号⑥は、

(1)フォルマントの「輝き」が見られる

(2)しかし、高次フォルマントの出が良くなく、フォルマントも 5 本と判定できる

という結果となるので、(1)は満足しているが(2)がクリアされないため、評価は「△」となる。

なお、参考までに、この資料番号⑥による調音を聞いていただきたい。

＜音声入る＞

4.7. 資料番号⑧の[i]

図 13 に示した資料番号⑧は、

(1)フォルマントの「輝き」が見られない

(2)高次フォルマントが全く出ておらず、総じてフォルマントの出が悪い

という結果となるので、(1)(2)ともに満足されていないので、評価は「×」となる。

なお、参考までに、この資料番号⑧による調音を聞いていただきたい。

＜音声入る＞

4.8.有識者による判定

先に、2.6.4.で述べたように、これらの SPG データをあらかじめ 5 人の有識者に等条件で画像ファイル化し、判定に関するアンケートを取った。結果は次の通りとなった。

A 氏　⑤＞⑥＞①＞④＞②＞③＞⑧
B 氏　①＞⑤＞⑥＞④＞②＞③＞⑧
C 氏　④＞⑤＞①＞⑥＞③＞②＞⑧
D 氏　④⑤＞①⑥＞②③＞⑧
E 氏　④⑤＞①②⑥＞③＞⑧

A~C 氏は、逐一微細な点などにも注視しながら一つずつ丁寧に順番をつけたが、D,E 氏はもう少し大雑把に見てカテゴリー化した評価をしている。そこで、以上のアンケート結果を整理するにあたって、A~C 氏の評価には 1 番に 7 点を与え、以下順次最下位の 7 番まで 1 点ずつ減じた点数を与える。次に、D 氏の場合は、A グループの④⑤に 7 点、B グループの①⑥に 5 点、C グループの②③に 3 点、D グループの⑧には 1 点を与える。最後に E 氏の場合は、A グループの④⑤に 7 点、B グループの①②⑥に 5 点、C グループの③に 2 点、D グループの⑧に 1 点を与えて集計した。結果を表 3 に示す。

表 3　集計結果

	①	②	③	④	⑤	⑥	⑧
A	5	3	2	4	7	6	1
B	7	3	2	4	6	5	1
C	5	2	3	7	6	4	1
D	5	3	3	7	7	5	1
E	5	5	2	7	7	5	1
合計	27	19	14	29	33	25	5

表 3 に示したように、結果は

⑤＞④＞①＞⑥＞②＞③＞⑧

となった。ちなみに、この結果と私がこれまで 4.1.~4.7.で述べてきたことをすり合わせると、私は⑤④を「○」、①⑥②を「△」、③＞⑧を「×」と評価してきたのできわめて良く一致しているということになる。ただし、カテゴリー化する際に上の結果では「○」と「△」の分水嶺がわずか 2 点差というのは、いささか気になるところではある。

５．結語

以上で見たように、本稿では SPG を目視することによって声年齢を判定する評価基準が設けられるのではないかという可能性について論じた。しかしながら、なにぶんにもはじめての試みであり、今後の検討に期するところも大きい。そこで、本稿のまとめとして、以下に現在までのこの研究におけるメリットとデメリットについて述べておく。

利点：なによりも、低額予算で手間暇をかけずに声年齢が評価できそうであるという点は、大きな利点である。また、被験者の負担も、ごくわずかな調音をすれば良く、負担が少なく便利である。

問題点：まだ研究が始まったばかりなので、ここに述べたことがらの信頼性を、今後多くの事例に基づき、検証して行かなければならない。また、「声年齢」と謳っておりながら、現状では声の良し悪しを判定するにとどまっている。冒頭の表 1 に示したような細分化は、少なくとも私たちの分野では不可能でありまた不毛でもあるが、今後は何らかの基準を持ちこんで、多少なりとも目標に近づける試みを行わなければならない。その最短にある発想は、おそらく現在設けている「○△×」によるカテゴリー化であろうと考えている。

【参考文献および資料】

城生佰太郎(2008[1],2010[2])『一般音声学講義』勉誠出版

NHK 総合テレビ制作(2012) 「世界を魅了する日本の歌謡曲：由紀さおり　ヒットの秘密」（クローズアップ現代）2012 年 2 月 21 日放送

http://josei-bigaku.jp/koenoanchieizingu3611/「女性の美学」