

発声者が知覚する音声の音色 — その客観的評価 —

中山 一郎 岡田 稔枝 中川 みかほ

大阪芸術大学音楽学科
〒585-8555 大阪府南河内郡河南町東山469
Tel. (0721) 93-3781 (内線3235) E-mail: nkym1@gold.ocn.ne.jp

あらまし： これまで筆者らは、発声時に発声者自身が聴取する音声（自己聴取音）の音色を、自己聴取音に相当する音声を発声者がイコライザとアッテネータを操作して、自己聴取音と出来るだけ似かよった音声（シミュレート音）を作成する方法（遅延音帰還法）を用いて、周波数軸上で定量化してきた。しかしながら、自己聴取音は発声者本人しか聴取できないが故に、シミュレート音との類似性を客観的には評価できなかった。そこで本研究では、男性の話し声の母音/a/を対象に、1)自己聴取音とシミュレート音、及び、2)シミュレート音とその音色を実験者が意図的に変化させて作成した音声（加工音）、との2種類の類似度評価実験を行うことによって、自己聴取音とシミュレート音の類似性が原理的には客観的に評価できること、及びその類似性が十分に高いことを明らかにした。

キーワード： 自声聴取、音声の音色、遅延音帰還法、シミュレート音、類似度評価、客観的評価

Objective evaluation of voice timbre in autophonic production

Ichiro NAKAYAMA, Toshie OKADA, and Mikaho NAKAGAWA

Department of Musicology, Osaka University of Arts
Higashiyama, Kanan-cho, Minamikawachi-gun, Osaka, 585-8555 Japan
Tel. (0721) 93-3781 (Ext. 3235) E-mail: nkym1@gold.ocn.ne.jp

Abstract: What sort of voice timbre do we perceive in autophonic production, hearing our own voice during phonation, and can we measure the timbre itself quantitatively? The authors have showed that the timbre can be examined quantitatively as a function of frequency using a delayed feedback method created by the author, in which a sound was generated by the subjects themselves so as to simulate the sound through a loudspeaker as accurately as possible in its timbre and loudness to the sound perceived during phonation. However, the judgement of similarity of voice timbre between the perceived and the simulated sounds was left to the subjects themselves, not objectively, because the voice timbre in autophonic production cannot be examined by anyone other than the speaker oneself.

The purpose of the present study is to show a method to evaluate the similarity between the simulated and perceived sounds objectively, when male subjects phonated a vowel/a/ in the spoken mode. Two similarity judgement tests were conducted, one (TestA) between the perceived and the simulated ones, and the other (TestB) between the simulated and the processing sounds, with an amplitude of a frequency band of the simulated sound altered, using the same criterion of the judgement of the similarity as in TestA. Through these tests, it was found objectively that the similarity between the simulated and perceived sounds was fairly well.

key words: Autophonic production, Voice timbre, Delayed feedback method, Simulated sound, Evaluation of similarity, Objective evaluation

1. はじめに

我々は、自分自身の声をどのようなものとして聴いているのであろうか？ この、自声聴取 (autophonic production) という聴覚の根本問題の一つに関しては、自声聴取が日常経験そのものであるにも拘わらず、発声中に聴く音声（以後、自己聴取音と記す）のラウドネスや音色などの基本的な音響諸量が、定量化されてこなかった。ラウドネスについては、自分の声は過小に知覚されていることが定性的に述べられてきたに止まっていたし[1]、また、音色については、発声音を録音／再生して聴く音声（録音／再生音）とは差異がある、とこれまた定性的に語られるに止まっていた。

先に筆者の一人は、自己聴取音の特性は定量的に取り扱うことが可能であり、ラウドネスに関しては、同じレベルの外部入射音の聴取時に比べて、相対音圧レベルで約10dBに相当するだけ過小知覚していること[2]、また、音色に関しては、録音／再生音と比較して低周波数域を相対的に大きく（約5 dB）、高周波数域を小さく（-5 dB）聴取していること[3]を明らかにしてきた。これらの結果は、独自に開発した遅延音帰還法[4]によって初めて定量化された。この方法は、自己聴取音に相当する音声を、被験者がイコライザとアッテネータを操作して、自己聴取音と出来るだけ似かよった音声（シミュレート音）をスピーカから発生させることによって作成する方法である。

しかしながら、自己聴取音は発声者本人しか

聴取できないが故に、遅延音帰還法によって作成されたシミュレート音と自己聴取音の類似性を客観的には評価できなかった。

そこで本研究では、男性が話し声の母音/a/を発声する場合を取り上げて、2種類の類似度評価実験を行うことによって、これら2音の類似性が原理的には客観的に評価できることを明らかにする。

2. 方法

自己聴取音とシミュレート音の類似度の客観的な評価は、以下に述べるように、自己聴取音とシミュレート音、及び、シミュレート音とそれを実験者が意図的に音色を変化させて作成した音声（加工音）の類似度評価を行うことによって可能となる。なお、本研究では、シミュレート音は1/oct.band単位でイコライザを操作して、また、加工音は研究の第一段階として、1/1 oct.bandのある一帯域のみの音圧レベルを変化させて作成することを前提にする。即ち、音色を決定するための多次元的な周波数のパラメータを一次元と仮定する。

2-1 シミュレート音の作成

筆者らの従来の研究と同じ遅延音帰還法を用いた。実験のブロック・ダイアグラムをFig.1に示す。男性の発声者（類似度評価実験の被験者を兼ねる）が通常の話し声で発声した母音/a/（約2秒間。音高:125Hz(nominal)）を、口唇前方30cmの位置で収音し、2.5秒の遅延をかけて前方のスピーカから発生させる。発声者は手

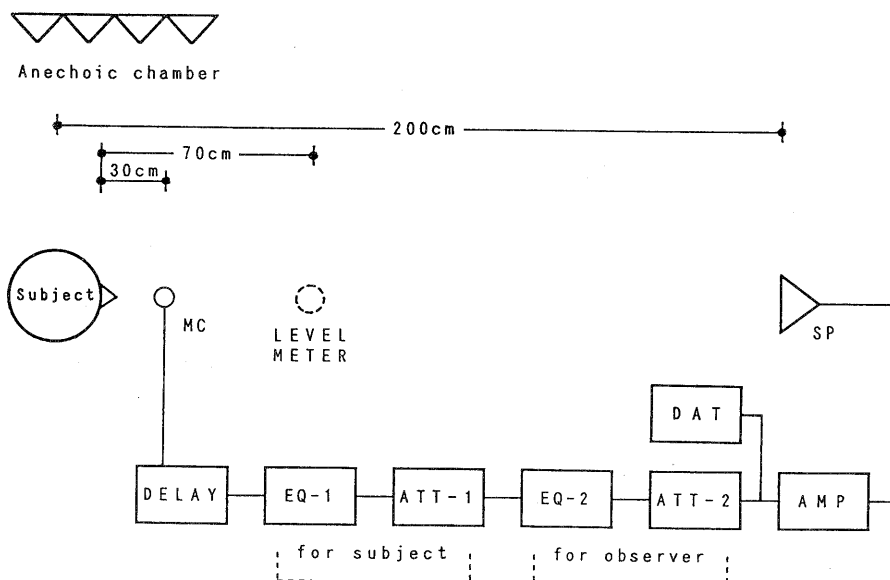


Fig.1 Block diagram of an experimental set-up in an anechoic chamber.

元のイコライザ (1/oct.band. 125Hz~8kHzの7バンド) とアッテネータを用いて、遅延音の音色とラウドネスを自己聴取音と出来るだけ似かよったものにする。この操作は、自己聴取音をスピーカによって空気音に変換して実現することを意味する。作成後、スピーカへの入力信号をDATに録音する。

実験はすべて、吸音処理した半無響状態の防音室内で行った。発声者は正常な聴力を有する本学音楽学科音楽工学コースの学生3名である。

2-2 加工音の作成

シミュレート音の音色を意図的に変化させて、類似度評価実験に用いる加工音を作成する。収録したシミュレート音の一つの周波数バンドごとに、音圧レベルを ± 3 dB、及び、 ± 6 dBの4通りに変化させて (即ち、一帯域のみの音圧レベルを変化させて、他の帯域は固定)、28通りの加工音を作成した。

2-3 自己聴取音とシミュレート音の類似度評価の解釈法

いま、ある発声者の、ある周波数バンドに着目し、シミュレート音 (Sと略記) の音圧レベルを (例えば) $+3$ dB、及び、 -3 dB変化させた加工音を、それぞれK1、K2とする。①自己聴取音 (Jと略記) とSの対 (評価対Aと記す) の類似度評価を行い、お互いが (例えば) 「少し違う」という評価であったとする。②次に、①で行ったと同じ評価の基準を用いて、SとK1の対、及び、SとK2の対 (評価対Bと記す) について評価を行い、共に「少し違う」という評価であったとする。③するとこの場合、評価対Aと評価対Bの類似度の評価が等しいので、Sのそのバンドの振幅を $+3$ dB、または -3 dB変化させた値が、Jの真の値であると解釈することができる。④もし、評価対Aの評価が評価対Bのそれより類似度が高いならば、Jの真の値は少なくとも、評価対Bの評価の領域内に存在することになる。他の各周波数バンドについても、同様の論理で自己聴取音の真の値を求めることができる。

2-4 類似度評価実験

2-3で述べたごとく、1) 先ず、評価対Aについての類似度評価を行い、2) 続いて、64対の評価対B (シミュレート音と加工音の対 (28対) と、呈示の順序効果を見るために、その順序を入れ替えた対の合計56対、及び被験者の信頼性評価のための8対。対中の2音でラウドネス平衡を行っている。) についての類似度評価をスピーカ受聴で行った。評価は64対を1試行として、被験者一人につき3試行行った。被験者には、対の初めに呈示された音声の音色に、続いて呈示された音声の音色がどの程度類似しているかを、「同じ」「少し違う」「かなり違う」「非常に違う」の4つのカテゴリで評価させた。

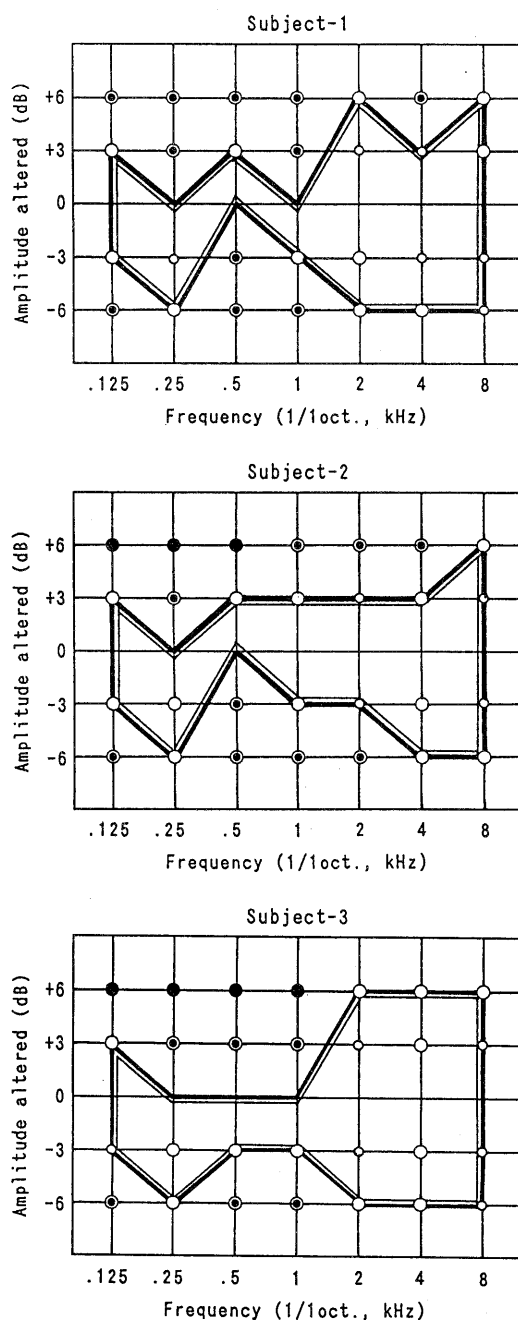


Fig.2 The judgements of similarity of voice timbre between the simulated and the processing sound, as a function of frequency (1/oct.band), with the amplitude of a frequency band altered.
○: same, ○: a little different, ◎: fairly different, ●: very different.

なお、対中の各音声の間隔は0.5秒とし、同じ対を続けて2回呈示するために、対どおしの間隔を1秒、また、異なる対の間隔を4秒とした。

3. 結果と考察

自己聴取音とシミュレート音の評価対Aに関しては、発声者1と2は「少し違う」、発声者3は「同じ」と評価した。また、評価対Bに関しては、全発声者で順序効果は認められなかった。シミュレート音を基準として、類似度評価の4つのカテゴリーにそれぞれ0、1、2、3点の評点を与え、発声者ごとに評価値を平均（3試行、及び対中の正・逆に亘る平均）して、結果を、シミュレート音の周波数特性を0 dBとしてFig.2示す（○:「同じ」、○:「少し違う」、◎:「かなり違う」、●:「非常に違う」）。また、評価対Aで発声者1と2が評価した「少し違う」を、類似度評価の信頼性の許容限界と考えて、それに該当する領域を同図に二重線で囲って示す。2-3で述べたように、この領域内に自己聴取音の真の値が存在することになる。この領域は十分に納得できる狭さであり、従って、各発声者の作成したシミュレート音は、十分に信頼できるものであると考えることができる。なお、全発声者に共通して、高周波数域では比較的「同じ」が多く、この高域成分の音色評価への影響力は小さいことを示している。これは、発声される音声中に、もともと広域成分が含まれていないことが原因と推察される。一方、低

周波数域では図中の領域は狭く、音色の変化への影響力が大きいものと考えられる。

4. まとめ

男性の話し声の/a/を対象に、自己聴取音とシミュレート音、及びシミュレート音と加工音との2種類の類似度評価実験を行い、自己聴取音とシミュレート音の類似性が原理的には客観的に評価できること、及びその類似性が十分に高いことが明らかになった。最後に、発声者諸氏、並びに本研究に有効な助言をいただいた本学 山田真司博士に深謝する。

参考文献

- [1] E. Borg and S.A. Counter, "The Middle-Ear Muscles," *SCIENTIFIC AMERICAN* August 62-68(1989).
- [2] 中山一郎ほか, "発声者が知覚するラウドネス 一過小知覚量の定量化一," 日本音響学会講演論文集 1-3-21, 431-432(1997.9).
- [3] I. Nakayama, "Voice timbre in autophonic production compared with that in extraphonic production," *J. Acoust. Soc. Jpn.(E)* 18(2), 67-71(1997).
- [4] I. Nakayama *et al.*, "Perception of loudness in autophonic production by singers," *Proc. of 1st. Int. Conf. on Music Percep. & Cog. (Kyoto, Japan)* KD3-1, 359-364(1989).