

声楽家の演奏スタイルと音場との時間的融合に関する研究

加藤 浩介 安藤 四一

神戸大学大学院自然科学研究科

Email: kkato@kobe-u.com

建築音響と音楽音響の橋渡しの研究として、音源の時間的なファクターを定量的に評価し、音楽とホール音響の融合を試みる。聴者と演奏者に最適な音場条件が、音源の自己相関関数(ACF)の有効継続時間の最小値(τ_e)_{min}から求まるという Ando の理論 [安藤四一著・酒井博之・佐藤伸一共訳, 建築音響学—音楽演奏・音響空間と聴衆との融合, シュプリンガーフェアラーク東京, 2000] に基づき、声楽の楽譜と演奏スタイルから歌声の(τ_e)_{min}を予測することを本研究の目的とした。特に本稿では、歌詞の違いと演奏テンポの違いが τ_e (遅れ時間ごとの信号の繰り返し成分の減衰) にどの程度影響を及ぼすかに着目し、4人の被験者の歌声を無響室録音してACFの解析を行った実験結果を報告する。最も重要な発見は、1) 歌詞の種類によって humming > melisma singing > lyric singing > non-voiced singing の順に (τ_e)_{min} は短くなる; 2) 演奏のテンポが3倍になっても(τ_e)_{min} はほとんど変化しない; 3) 音の高さが変動しないモチーフは、音の高さが変動するモチーフと比べて(τ_e)_{min} が有意に長い(p < 0.05)という点である。

On the Blending of Vocal Music with the Sound Field by Different Interpretation Styles

Kosuke Kato and Yoichi Ando

Graduate School of Science and Technology, Kobe University

Concert halls do not have ideal conditions for all music programs. It is well known that the most preferred conditions for both listener and performer are determined by the minimum value of the effective duration of the running autocorrelation function (ACF) of sound signals, (τ_e)_{min} [Y. Ando 1998 Architectural Acoustics-Blending Sound Sources, Sound Fields, and Listeners. AIP/Springer-Verlag, New York]. An attempt is made here to estimate (τ_e)_{min} of vocal music by rating various kinds of interpretation styles of singers. The present results showed that (τ_e)_{min} of the ACF of a voice source varies with lyrics and fluctuation of pitch but not music tempo. Significant findings are 1) Values of (τ_e)_{min} are relatively longer in order of: humming > melisma singing (singing with "la" syllables) > singing with lyrics > non-voiced (breath noise) singing; 2) Values of (τ_e)_{min} of fast vocal music may not be shorter than those of slow tempo music; 3) Values of (τ_e)_{min} of vocal music with pitch fluctuation is shorter than those of music with constant pitch (p < 0.05).

1. Introduction

1.1 音楽とホール音響の融合

音響条件はそれぞれのホールによって異なる。残響時間ひとつをとって見ても、長いホール短いホール様々である。例えば、一般に教会の残響時間はオペラハウスに比べて長い。教会音楽をオペラハウスで演奏すると響きが足りないであろうし、逆にオペラ音楽の多くは教会で演奏すると響きすぎであろう。このように全ての音楽演奏に適した音響条件は存在しないため、双方を考慮すること、すなわち音楽演奏とホール音響の融合を目指すことが望ましい。

Ando は、人間の最もプリミティブな反応である主観的プリファレンスに対応する音源の時間的ファクターに関する研究を行った。オーケストラ演奏、ピアノ演奏、スピーチなどを用いてプリファレンス実験を行い、音源のランニング自己相関関数 (ACF) の有効継続時間の最小値(τ_c)_{min} が聴者の初期反射音の遅れ時間、残響時間などの音場の時間的ファクターに比例関連を持つことなどを明らかにした [1]。また、ステージ上の演奏者の演奏のしやすさがこの (τ_c)_{min} で評価できることも明らかになっている [2-4] (付録参照)。しかし、音楽の τ_c を求めるには通常は無響室録音された演奏ソースの解析が必要であるため、データの数が限られていた。そこで、数多く存在するすべての音楽に対応するため、楽譜と演奏スタイルから τ_c を予測する方法の研究が望まれた。

Taguti は自動演奏ピアノを用いてピアノ音源の τ_c を解析し、1) 演奏の速さまたはテンポ 2) ダイナミクス 3) 音の歯切れ 4) 同周期性 5) ダンパーペダルの踏み方が音源の τ_c を決定する主要な要素であることを明らかにした。例えば、速いテンポの演奏では τ_c が短くなる。ダンパーペダルを踏むと τ_c が長くなる。 τ_c の最小値は音符の継続時間 (NOD) におおよそ一致す

る。スタッカートの鋭さが増すと τ_c が短くなるが、60ms より短くなることはない。これはピアノ音の再生のメカニズムと関連がある。ということが明らかになっている [5]。

別の音源に歌声がある。声楽演奏にはダンパーペダルはない。しかし、i) 発声方法 ii) 歌詞 iii) ビブラート iv) 音の立ち上がりという、ピアノ音源にはない要素をもちユニークである。そこで、ピアノ音源の τ_c の決定要素から「ダンパーペダルの踏み方」を除いた 4 要素に新しくこれら 4 要素を加える。つまり合計 8 要素を定量的に評価し、声楽演奏の τ_c の回帰的な予測を行うことを本研究の目的とする。

1.2 声楽演奏の τ_c 予測に関する研究実績

発声方法の違いが τ_c に与える効果についてはこれまでに検証を終えた。E4 “oh” 母音およびアクセント・スタッカートがなく 4 音符と 8 音符のみで構成される単純な譜面を falsetto, medium falsetto, operatic singing という 3 種類の発声で 5 人のテノール歌手に演奏してもらい、無響室録音した歌声の ACF を解析した。その結果、falsetto または medium falsetto の (τ_c)_{min} は operatic singing の (τ_c)_{min} と比べて有意に長いことが明らかになった [6, Journal of Sound and Vibration に投稿中<accepted>]。

これに続く研究として、本稿は歌詞の違いおよび演奏テンポの違いが声楽演奏の τ_c に与える効果をそれぞれ検討することにする。

1.3 声楽演奏に関する他の研究

声楽演奏に関する従来の研究は、発声器官のメカニズムの解析 [7,8] や歌声のスペクトル解析 [8-11] が中心である。声楽演奏に好ましい音場についての研究 [12-14] もあるが、本研究のように時間的ファクターに着目して歌声と音場との定量的な関係を考察した研究はない。

2. 実験

2.1 実験1：歌詞を変化させた場合

方法 歌は他の楽器と異なり、歌詞が存在する。そこで、humming, *melisma* singing, lyric singing, non-voiced singing という4種類の歌詞の効果を検討した。humming は完全に口を閉じて“m”の歌詞で歌い、*melisma* singing は全ての音符に“la”の歌詞をつけて歌い、lyric singing は実際の曲の歌詞をつけて歌い、non-voiced singing は声帯を振動させず口を動かして息音のみで曲の歌詞を歌うとする。

被験者は Soprano・Alto・Tenor・Bass 各1人ずつの合計4人とした。楽譜は Fig. 1 に示す2モチーフのうち Motif I (FP)を用いた。移調を行って最初の音程の高さを、Soprano=D5、Alto=G4、Tenor=D4、Bass=G3 とした。演奏のテンポは歌い手の目の高さで1.0m 前方に配置したビジュアルメトロノームによって4分音符=120bpm (beat per minute) に保たれた。試行回数は10回とした。歌声は、歌い手の口の前方25±1cm、側方5±0.5cm に無指向性のマイクロフォンを配置して無響室録音を行った。

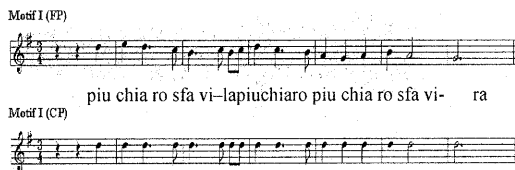


Fig. 1 Motif I (FP)と Motif (CP)の楽譜

結果 ランニングACFの積分区間2Tは、Mouriにより推奨区間 $[2T]_r = 30(\tau_e)_{\min}$ が提案されている[15]ため $2T = 1.0s$ とし、ランニングステップは0.1sで解析を行った。 τ_e の変動の例を Fig. 2 示す。また、歌声の $(\tau_e)_{\min}$ を解析した結果を Table 1 および Fig. 3 に示す。4種類の歌詞に対して個人ごとに一元配置の分散分析を行った結果、全ての場合で有意差 ($p < 0.01$)があった。

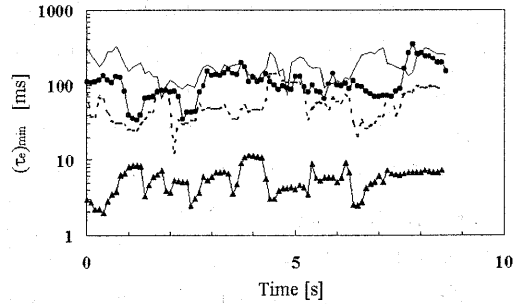


Fig. 2 τ_e の変動の例 (被験者KK, Tenor)

— : humming, —●— : *melisma* singing
 ... : lyric singing, —▲— : non-voiced singing

Table 1 被験者ごとの $(\tau_e)_{\min}$ (平均値±標準偏差)

	$(\tau_e)_{\min}$ [ms]				
	CT (Sep)	YA (Alt)	KK (Ten)	DT (Bas)	Global
Humming	73 ± 14	51 ± 13	64 ± 18	70 ± 20	65 ± 18
Melisma	49 ± 6	36 ± 5	34 ± 11	26 ± 3	36 ± 11
Lyric	33 ± 8	28 ± 6	14 ± 4	18 ± 2	23 ± 10
Non-voiced	-	2.1 ± 0.6	2.0 ± 0.2	-	2.0 ± 0.4

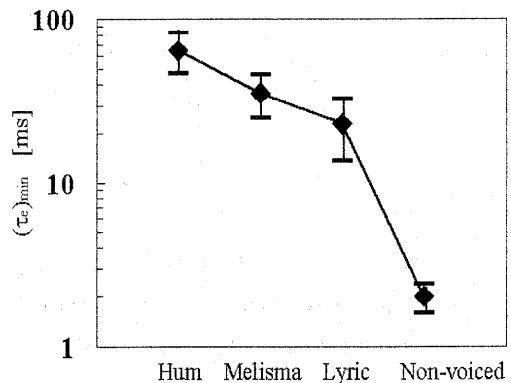


Fig. 3 被験者全体の $(\tau_e)_{\min}$ (平均値±標準偏差)

2.2 実験2：演奏テンポを変化させた場合

方法 声楽演奏もピアノ演奏の場合と同様、演奏の速さにより τ_e が変化すると考え、検証を試みた。被験者は実験1と同じとした。楽譜は Motif I (FP) および Motif I から音程の変化をなくした Motif I (CP)を用いた。移調を行って出

だしの音程の高さを、Soprano=D5、Alto=G4、Tenor=D4、Bass=G3 とした。歌詞は全て”la” (*melisma singing*) とした。演奏のテンポは4分音符=60, 120, 180bpm の3種類とした。試行回数および録音方法は実験1と同様にした。結果 積分区間 $2T=1.0$ [s]、ランニングステップ $R_s=0.1$ [s] で計算した τ_e の変動の例を Fig. 4 に示す。また、歌声の $(\tau_e)_{\min}$ を解析した結果を Table 2 および Fig. 5 に示す。

被験者 YA (Alto) における Motif (FP) の 60bpm の場合を除き、演奏テンポが変わっても $(\tau_e)_{\min}$ は有意に変化しなかった。しかし、各テンポ、被験者ごとに一元配置の分散分析を行ったところ、被験者 CT, 180bpm の場合を除き Motif I (FP) の $(\tau_e)_{\min}$ は Motif I (CP) に対して有意に短い ($p < 0.05$) という結果を得た。

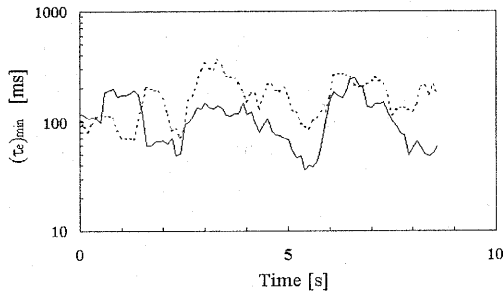


Fig. 4 τ_e の変動の例 (被験者 YA; Alt)

—: Motif I (FP), 120 bpm, $(\tau_e)_{\min}=36$ [ms].

---: Motif I (CP), 120 bpm, $(\tau_e)_{\min}=67$ [ms].

Table 2 被験者ごとの $(\tau_e)_{\min}$ (平均値±標準偏差)

Motif	bp m	$(\tau_e)_{\min}$ [ms]					Global
		CT (Sop)	YA (Alt)	KK (Ten)	DT (Bas)		
Motif I (FP)	60	54 ± 11	25 ± 4	42 ± 6	27 ± 5	36 ± 15	
	120	49 ± 6	36 ± 4	34 ± 11	26 ± 3	36 ± 11	
	180	59 ± 9	32 ± 9	39 ± 13	31 ± 9	40 ± 15	
Motif I (CP)	60	65 ± 11	43 ± 7	51 ± 5	35 ± 3	49 ± 14	
	120	58 ± 9	48 ± 12	49 ± 9	40 ± 9	49 ± 11	
	180	58 ± 13	51 ± 14	56 ± 12	42 ± 5	52 ± 13	

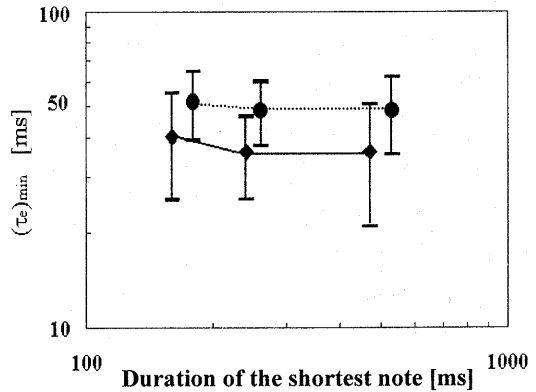


Fig. 5 被験者全体の $(\tau_e)_{\min}$ (平均値±標準偏差)

—◆—: Motif I (FP), ...●...: Motif I (CP)

横軸は Motif I の中で最も短い、8分音符の長さを表す。60, 120, 180 [bpm] に対応する 8分音符の長さはそれぞれ 500, 250, 167 [ms] である。

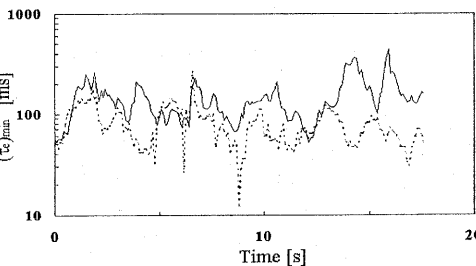
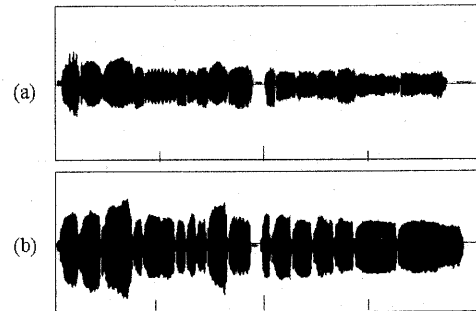


Fig. 6 被験者 YA の実波形(a) と被験者 KK の実波形 (b) ならびに τ_e 変動の比較 (Motif I, 60 bpm)

---: 被験者 YA, $(\tau_e)_{\min}=30$ [ms].

—: 被験者 KK, $(\tau_e)_{\min}=55$ [ms].

3. 考察

3.1 実験1に対する考察

歌詞に含まれるノイズ成分が多いほど τ_e が短くなる傾向があるため、humming > melisma singing > lyric singing > non-voiced singing の順に τ_e は短くなる。という仮説を立てて実験を行ったが、その通りの結果が得られた。

個人ごとに見れば τ_e に有意差があるが、個人差も大きい。例えば、被験者 CT (Soprano) の lyric singing が $(\tau_e)_{\min} = 33 \pm 8$ [ms]であるのに対して、被験者 KK (Tenor) の lyric singing は $(\tau_e)_{\min} = 14 \pm 4$ [ms]である。この差が、パートの違いによるものなのか、個人の歌い方の違いによるものなのかは今回の結果からは判断できないので、今後被験者数を増やして検討する。

また、声楽演奏が他の楽器の演奏と大きく異なる点に、骨伝導音の効果がある。特に今回扱った humming のような歌詞は低い周波数を多く含み、骨伝導の効果が大いことが予想される。我々は、声楽演奏に最適な音場条件が骨伝導音の大きさと τ_e によって算出できると仮説している。このうちの τ_e が歌詞によって大きく異なるという今回の発見は意義深い。

3.2 実験2に対する考察

声楽演奏の場合はピアノの場合[5]と異なり、 τ_e の最小値が音符の継続時間 (NOD) に一致しなかった。それどころか、演奏のテンポが3倍になっても $(\tau_e)_{\min}$ はほとんど変わらなかった。速いテンポで歌うときにはエンベロープの変動が小さくなり、音高がはげしく変動することによる効果を打ち消していると推測する。

例外的に、被験者 YA の Motif I (FP), 60bpm の $(\tau_e)_{\min}$ が他の被験者と比べて短かった理由を検討してみたい。Fig. 6 の(a) はこのセッションにおける被験者 YA の歌声の典型的な波形である。注目すべきは 5s と 18s 付近に存在す

るビブラートである。Fig. 6 の τ_e の変動を見ると、その付近で小さな値になっていることがわかる。定量的な考察には至らないにしても、このことはビブラートが声楽演奏の τ_e を短くする要因であることを示唆している。なお 8-9 [s] 付近で τ_e の値が急激に小さくなっているのは、息継ぎの時間が 0.8 秒間に及んだことによるノイズの効果であると考えたため、最小値の候補からは除外した。

3.3 総合的な考察

今回は、合唱経験が 10 年以内の choir singer の被験者で実験したからか、被験者 KK, Motif I (FP), 180 bpm の $(\tau_e)_{\min} = 39 \pm 13$ [ms]に見られるように、再現性が乏しかった。トレーニングによって、あるいは豊富な経験をもつ被験者を採用することによって再現性を確保できるかについては今後検討する。

Mouri が提案する、推奨積分区間 $[2T]_r = 30(\tau_e)_{\min}$ は騒音のラウドネスのためのものであるため、声楽演奏に相当であるかどうかは未検証である。プリファレンス予測やピッチ同定に適切な積分区間も検討する必要があるだろう。

4. 結論

今回の実験結果から次のことがいえる。

1) 歌詞の種類によって声楽演奏の τ_e は大きく変化する。 τ_e の長さは humming > melisma singing > lyric singing > non-voiced singing の順に短くなる。

2) ピアノの場合と異なり、声楽演奏では演奏のテンポは τ_e を決定する主要な要素でない。

3) 音の高さが変動しないモチーフは音の高さが変動するモチーフと比べて τ_e が長い。

未検討の 5 要因すなわち、ダイナミクス、音の歯切れ、同期性、ビブラート、音の立ち上がりの効果について今後検証を行う。

参考文献

1. Y. Ando 1998 Architectural Acoustics-Blending Sound Sources, Sound Fields, and Listeners. AIP/Springer-Verlag, New York.
2. I. Nakayama 1984 Acustica **54**, 217-221. Preferred time delay of a single reflection for performers.
3. S. Sato, Y. Ando, and S. Ota 2000 Journal of Sound and Vibration **232**, 27-37. Subjective preference of cellists for the delay time of a single reflection in a performance.
4. D. Noson, S. Sato, H. Sakai and Y. Ando 2000 Journal of Sound and Vibration **232**(1), 39-51. Single Responses to Sound Fields with a Simulated Refraction.
5. T. Taguti and Y. Ando 1997 in Music and Concert Hall Acoustics. (Y. Ando and D. Noson, editors) Chapter 23. Academic Press, London. Characteristics of the short-term autocorrelation function of sound signals in piano performances.
6. K. Kato and Y. Ando 2001 17th ICA Proceedings Volume III. Blending voice source and temporal factor of sound field by selecting different singing types.
7. K. Tom and I. Titze 2001 J. Acoust. Soc. Am., **110**, 1667-1676. Vocal intensity in falsetto phonation of a countertenor: An analysis by synthesis approach.
8. J. Sundberg 1987 Northern Illinois University Press, DeKalb, IL. The science of the singing voice.
9. S. Ternstöm and J. Sundberg 1989 J. Acoust. Soc. Am., **86**, 517-522. Formant frequencies of choir singers.
10. G. Bloothoof, E. Bringmann, M. Cappellen, J. Luipen, and K. Thomassen 1992 J. Acoust. Soc. Am., **92**, 1827-1836. Acoustics and perception of overtone singing.
11. H. Smith, K. Stevens, and R. Tomlinson 1967 J. Acoust. Soc. Am., **41**, 1262-1264. On an unusual mode of chanting by certain Tibetan lamas.
12. A. Gade 1989 Acustica **54**, 217-221. Investigations of musician's room acoustics conditions in concert halls, Part II.
13. A. H. Marshall and J. Meyer 1985 Acustica **58**, 130-140. The directivity and auditory impression of singers.
14. S. Ternstöm 1999 Journal of the Acoustical Society of America **105**, 3563-3574. Preferred self-to-other ratios in choir singing.

15. K. Mouri and K. Akiyama 2001 Journal of Sound and Vibration **241**, 87-95. Preliminary study on recommended time duration of source signals to be analyzed, in relation to its effective duration of the auto-correlation function.

付録 ランニング ACF の有効継続時間の最小値 $(\tau_e)_{\min}$ について

音源の ACF から抽出される時間的なパラメータである有効継続時間 τ_e (Fig. A1 参照) は、遅れ時間ごとの音源信号の繰り返し成分がどの程度継続的であるかを表す。とりわけランニング ACF の最小値 $(\tau_e)_{\min}$ は曲の中で最も活動的な箇所の情報を含み、聴者と演奏者にとって好ましい音場条件を決定する重要なパラメータである。以下にその計算方法を示す。

聴者に対する最適初期反射音の遅れ時間 $[\Delta t_l]_p$ および最適残響時間 $[T_{sub}]_p$ は、全反射音振幅を A とすると、式 (1) および式 (2) で表せることがわかっている[1]。

$$[\Delta t_l]_p \approx (1 - \log_{10} A)(\tau_e)_{\min} \quad (1)$$

$$[T_{sub}]_p \approx 23 (\tau_e)_{\min} \quad (2)$$

演奏者に対する最適初期反射音の遅れ時間 τ_p は、k, c を演奏者の主観的属性により決まるパラメータとすると、式 (3) で表せることがわかっている[2-4]。

$$\tau_p = \left(\log_{10} \frac{1}{k} - c \log_{10} A \right) (\tau_e)_{\min} \quad (3)$$

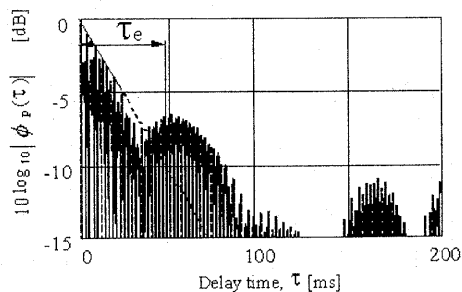


Fig. A1 音源の正規化 ACF の典型例