

研究ノート

声楽的発声における母音知覚
—声楽経験および基本周波数の影響—

遠藤 希美*・川原 繁人**・皆川 泰代*

Perceiving Vowels in Singing Voice:
The Effects of Listeners' Singing Experience and Fundamental Frequencies

Nozomi ENDO*, Shigeto KAWAHARA** and Yasuyo MINAGAWA*

SUMMARY: It is more difficult to accurately perceive vowels that are produced in singing than those uttered in normal speech. Previous studies have shown that the perceptual difficulty of vowel perception increases as F0 goes higher. However, no studies have examined whether having experiences of classical singing influences the ability to perceive vowel quality produced in high-pitched singing, the gap which the current study intends to fill. The results of the current study suggest that higher F0 makes sung vowels difficult to perceive, replicating the previous findings, and that singers and non-singers show different perception patterns under some circumstances.

Key words: speech perception, vowels, singing, F0, harmonics, F1, F2

1. 背景

声楽家の歌声は通常の話声と比較して歌詞が聞き取りにくいといわれている。特に、ソプラノ歌手の歌唱のように、非常に高い基本周波数(F0)での歌唱の母音と通常の話声の母音を比較した場合、音響的に異なったものとなり、母音の明瞭性はF0が高くなればなるほど低下する(Sundberg 1987)。

Morozov (1965)によると、プロの男性歌手・女性歌手によって歌われた音節の母音の知覚では、男性の歌声も女性の歌声も低いF0では知覚が容易であったが、男性歌手ではF0が390Hz以上、女性歌手の場合はF0が523Hz以上で正答率が低くなった。非常に高い基本周波数を音源とすると、スペクトルの倍音構造が希薄になり、聴取者が母音を判断するには貧弱な信号になるということが考えられる。合成音声により統制さ

れた刺激においても、F0が高くなればなるほど母音の判別成績は低下したことから、F0による影響は非常に大きいといえる(Majewski and Hollien 1967)。

歌声の母音の知覚について、実際の歌唱の場面ではさまざまな手がかりにより母音が判別しやすくなる。Nelson and Tiffany (1968)によると、判別成績は、母音が単独で発声された場合に比べて、隣接する子音があるときの方が上昇した。つまり子音に存在する母音の同時調音(coarticulation)の情報が母音知覚の助けになっていると推察される。Smith and Scott (1980)はさまざまな方法で自由に歌われた母音を用いた場合、母音の判別成績は歌い方に左右されることを明らかにした。母音の明瞭度は発声方法にも影響を受け、カウンターテナーの歌手が同じ音域を胸声とファルセットで歌った場合、ファルセットで正答率が高く、また語頭が母音のみである場合と子音が隣接

*慶應義塾大学文学部 (Faculty of Letters, Keio University)

**慶應義塾大学言語文化研究所 (The Institute of Cultural and Linguistic Studies, Keio University)

する場合には、語頭に子音が隣接する場合で正答率が高かった (Gottfried and Chew 1986)。実際の歌唱の場面では歌手の口の動きという視覚的な手がかりもあり、Jesse and Massaro (2010) によると、ソプラノ歌手が歌唱を行っているときの顔の動画と歌われたフレーズを同時に呈示した場合、フレーズのみや動画のみの場合と比較して歌詞の聞き取りの成績が上昇した。

聴取者の歌詞に対する知識や理解も歌の知覚に影響を及ぼすことがわかっている。Fine, Ginsborg and Barlow (2009) によると、ソプラノ歌手によって歌われた、歌詞の意味などを理解した歌曲と理解していない歌曲では、理解した歌曲で歌詞の聞き取りの成績が上昇した。プロの音楽家が聴取した場合とそうでない人が聴取した場合とでは、プロの音楽家において歌詞の聴取成績が高く、歌詞に対する知識・経験・注意が歌の知覚に影響を及ぼすことが示唆されている (Fine and Ginsborg 2007, Ginsborg, Fine and Barlow 2011)。

先行研究では、隣接する子音および他の文脈を用いたさまざまなトップダウンな手がかりが歌声の母音の知覚に重要であることが示唆されているが、声楽的な発声そのものの経験が母音の知覚にどのように影響を及ぼすかは未解明である。Fine and Ginsborg (2007), Fine et al. (2009), Ginsborg et al. (2011) ではプロの音楽家とそうでない実験参加者での歌唱の知覚の相違を示しているが、これは歌詞への理解や注意の影響を示すものであり、声楽的な発声の経験が単独で発声された母音そのものの知覚特性に及ぼす影響を示したものではない。声楽的な発声は通常の母音の発声とは調音運動および音響的な特性が異なるが、声楽経験者は調音運動の経験から母音を推定しやすいという可能性や、音響的な特性を聞きなれているために母音を同定・弁別しやすいという可能性もある。しかし、単独で母音が発声された場合、音楽家と一般人では知覚が異なるか、音楽家で母音の知覚の成績が高くなるかといったことは研究されていない。声楽の経験者は声楽の発声経験によって、歌詞などのトップダウンな知識とは独立し

て、母音の知覚そのものの能力が向上することも考えられる。本研究では、F0と声楽経験が声楽的に発声された母音の知覚にどのような影響を及ぼすか、ソプラノ歌手が発声した日本語における5つの母音について検討した。

2. 方法

2.1 実験参加者

実験参加者は、日本語を母語とする声楽家群17名(平均年齢21歳)および一般群20名(平均年齢22歳)であった。音楽大学で声楽を専攻している者および専攻した者を声楽家群、声楽を専攻していない一般大学に通う学生を一般群とした。声楽家群の声楽の経験年数の平均は8.71年、標準偏差は4.00年であった。刺激は女性の発声による母音のため、参加者は全員女性とした。参加者は聴覚に異常はなかった。

2.2 刺激

刺激には音楽大学の声楽専攻を卒業した日本語を母語とするソプラノ歌手3名の音声を用いた。ソプラノ歌手3名のうち、1名はドラマティックソプラノ、2名はリリックソプラノであった。刺激に用いた音声は日本語の5母音/i/, /e/, /a/, /o/, /u/であり、各母音のF0は440Hzと880Hzであった。刺激の長さは各母音1500msであり、刺激の平均音圧レベルを75 dBに統一した¹⁾。

録音は防音室で行った。録音にはコンピュータ(ProBook4730s, HP)にアンプ(US-366, TASCAM)を接続し、アンプにマイク(ATM98, audio-technica)を接続した装置を用いた。防音室の中にはマイクのみ設置した。音声の分析および加工は音声解析ソフトであるPraat(Boersma and Weenink 2015)を用いた。F0の440Hz条件、880Hz条件の順番で各母音を1つずつ、ビブラートをかけずにできるだけ正確に、5000ms程度で約3回発声された音声録音した。音の高さに対する手がかりとして、発声の前に440Hz条件の場合は440Hzの純音を、880Hz条件では880Hzの純音を

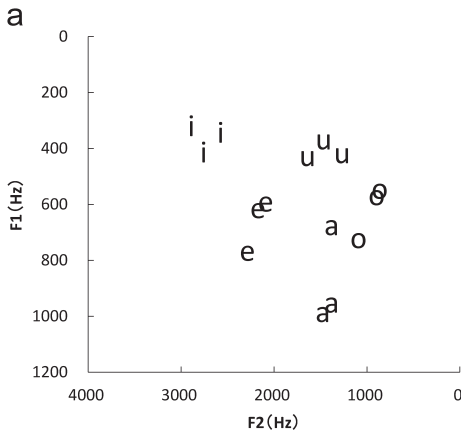


図1a 地声におけるフォルマント分布

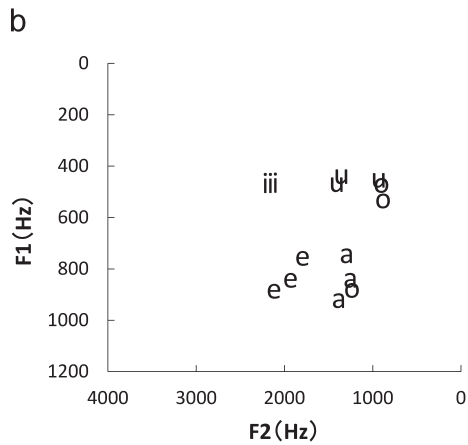


図1b 440Hz条件におけるフォルマント分布

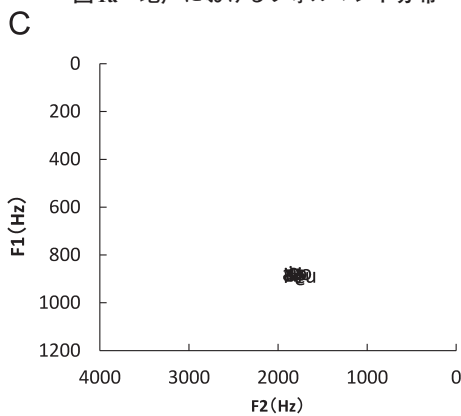


図1c 880Hz条件におけるフォルマント分布

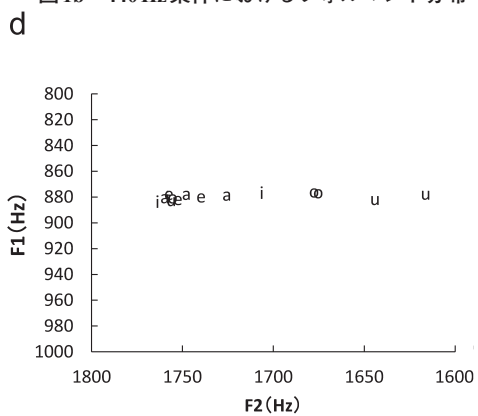


図1d 880Hz条件における拡大されたフォルマント分布

図1 地声, F0が440Hz条件, 880Hz条件におけるフォルマント平均分布

コンピュータ (LIFEBOOK SH75/M, FUJITSU) を用いて呈示した。複数の録音の中から440Hz条件の場合には誤差上下1Hz以内, 880Hz条件の場合には誤差上下2Hz以内となった定常母音部分を約1500ms切り取った。切り取った音声の平均音圧が75 dBとなるように加工した。切り取った音声の立ち上がり時間および減衰時間はそれぞれ50msとなるよう windowing加工した。

解析の参考として, 地声における /i/, /e/, /a/, /o/, /u/ を録音した。各話者における地声のF0の平均は話者1で271Hz, 話者2で273Hz, 話者3で221Hzであり, 3話者全体の地声のF0の平均は255Hzであった。スペクトルのピークとしてフォルマントを, 3話者の地声, 440Hz条件, 880Hz

条件の各条件で計測した。フォルマントの推定においてはLinear Predictive Coding (LPC) を用いてスペクトル包絡のピークを算出した。地声および440Hz条件ではPraatの公式サイトでのマニュアルで女性声のフォルマント分析に推奨されている通り (http://www.fon.hum.uva.nl/praat/manual/Sound_To_Formant_burg_.html), 5500Hz以下の範囲内で5個のフォルマントの推定を行った。ただし, 880Hz条件では倍音のみが強く現れ, 安定したフォルマント推定が不可能になったため, 5500Hz以下のすべての倍音成分が含まれるよう6個のフォルマントの推定を行った。スペクトルのピークは通常, 声道の共鳴を反映するが, 880Hz条件では, 必ずしもフォルマント

が声道の共鳴と一致していないことも考えられる。よって倍音自体の分析を考察の節で報告する。3話者の地声におけるフォルマント平均分布を図1a, 440Hz条件におけるフォルマント平均分布を図1b, 880Hz条件におけるフォルマント平均分布を図1cに示した。また, 880Hz条件における拡大されたフォルマント平均分布を図1dに示した。図1では話者ごとのフォルマント分布を区別できないため, 440Hz条件, 880Hz条件における各話者のF1, F2の具体的な数値を付録に示した。

2.3 手続き

聴覚実験は防音室にて個別に行った。刺激の呈示にはコンピュータ (LIFEBOOK SH75/M, FUJITSU) とヘッドフォン (HD 280 pro, SENNHEISER) を用いた。コンピュータに直接ヘッドフォンを接続した。刺激呈示の制御は刺激呈示ソフト (SuperLab, Cedrus) で行った。刺激の音圧を75 dBに統制するため, 疑似耳 (4153, Brüel & Kjær) と騒音計 (2250, Brüel & Kjær) を用いて音圧を計測した。

1試行で2つの刺激を順番に呈示し, 実験参加者は刺激を聞き2つの刺激の呈示順序を二肢強制選択法 (2AFC方式) で答えるよう求められた。歌手ごとに各F0で5つの母音を総当たりで組み合わせさせた (/a/-i/, /a/-u/, /a/-e/, /a/-o/, /i/-u/, /i/-e/, /i/-o/, /u/-e/, /u/-o/, /e/-o/, 計10対)。2つの刺激の間隔 (Inter-Stimulus Interval: ISI) は500msであった。組み合わせた刺激を8回繰り返し, そのうち4回は呈示順序を入れ替えた。刺激の呈示順序はランダムであった。各歌手でブロックを分け, 3ブロックとした。順序効果を統制するためにラテン方格を用いてブロックの順序を統制した。1ブロックあたりの試行数は母音の組み合わせ10対×F0要因2条件×繰り返し8回の計160試行であった。刺激呈示中はコンピュータの画面中央に注視点を呈示した。刺激呈示後すぐにコンピュータに回答画面を呈示した。回答画面では各刺激対の呈示順序について母音を示すローマ字2対をハイフンでつなぎ, 左右に2通り呈示した。回答画

面の文字の呈示位置は左右を入れ替えることでバランスをとった。実験参加者には, 刺激を聴いて画面に示されている呈示順序のどちらが正しい順序であるか画面の左の順序であればキーボードのFキー, 右の順序であればJキーで回答するよう, 文章および口頭で教示を行った。2つ目の刺激呈示終了からキーボードでの回答までの時間を計測し, 反応時間とした。

2.4 解析方法

各刺激対について知覚の成績を求めるために, 二肢強制選択法における信号検出理論で用いられる d'_{2AFC} (以下 d' と表記) の値 ($(Z(Hit) - Z(FA))/\sqrt{2}$) を実験参加者ごとに算出した。刺激対の呈示順序2通り (刺激/X/, 刺激/Y/を対とした場合, /X/-/Y/と/Y/-/X/) のうち一方を呈示順序A (/X/-/Y/), もう一方を呈示順序B (/Y/-/X/) とし, 呈示順序Aの正答率を P_A としたとき, P_A をHit (/X/-/Y/の刺激に対し, /X/-/Y/と解答), 呈示順序Bの正答率を P_B としたとき, $(1 - P_B)$ をFalse Alarm (/Y/-/X/に対し, /X/-/Y/と解答) とした。 $1/\sqrt{2}$ は, 2AFC法における信号検出理論の修正法である (Macmillan and Creelman 2005, pp. 167-170)。 d' の値が高くなればなるほど, 母音の弁別がより正確であるという指標となる。たとえば, Hitが0.5, FAが0.5のように回答がランダムで行われた場合, d' は0となる。また, 正答率が90%で, 反応バイアスがない場合, Hitが0.9, FAが0.1となる。 $Z(Hit) - Z(FA) = 2.56$ であるので, 2AFCの修正をかけて $d' = 1.81$ である。 d' の理論的な最大値は無限だが, 最大値の近似としては99%の正答率で, 反応バイアスがない場合3.29となる。 d' が負の値をとる場合は, HitがFAよりも低い, つまり, 正答率が不正答率よりも低いことを示す。

歌手3人分, すなわち3ブロック分をまとめたうえで, 母音の組み合わせを母音要因とし, 各刺激対について算出された d' を用いて声楽家要因2水準×F0要因2水準×母音要因10水準の分散分析を行った。この3名分の正答率は, 話者1で0.75

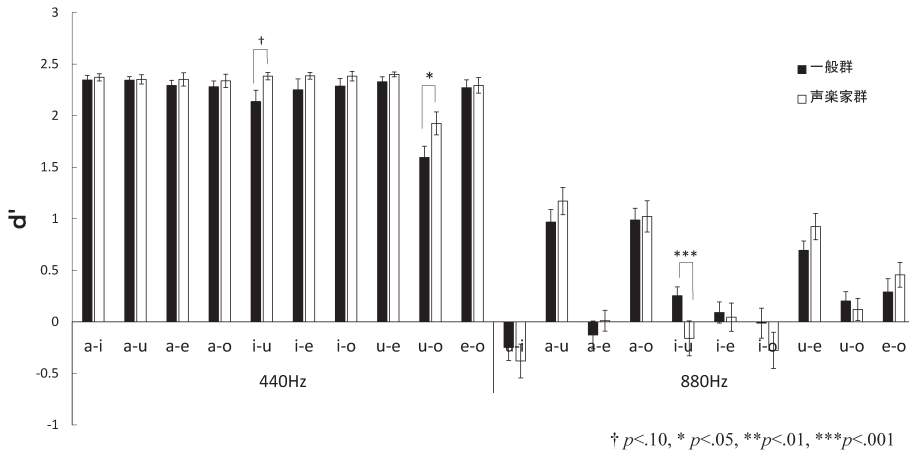


図2 異なるF0条件 (440Hz, 880Hz) における各母音条件の一般群と声楽家群のd'

(S.E.=0.01), 話者2で0.79 (S.E.=0.01), 話者3で0.76 (S.E.=0.01) となったが, 本研究では, より一般的かつ頑健に母音間の知覚的な距離を算出することを目的とするため, 話者3名分の正答率を合わせてd'を算出した。また, 各刺激対における反応時間を3ブロック分まとめたうえで実験参加者ごとに算出し, 声楽家要因2水準×F0要因2水準×母音要因10水準の分散分析を行った。多重比較を行う場合はライアン法を用いた。

3. 結果

3.1 成績

各F0における各母音について声楽家群と一般群のd'を図2に示す。図2より, 440Hz条件は880Hz条件よりd'が高く, 440Hz条件ではu-o条件が他の母音の組み合わせと比較しd'が低い傾向にあり, 声楽家群の方が一般群よりd'が高い傾向にあった。一方で, 880Hz条件では/i/の母音を含む組み合わせはd'が低く, 一般群の方が声楽家群よりd'が高い傾向にあった。

声楽家要因×F0要因×母音要因の分散分析を行った結果, 上記の傾向はほぼ支持された。F0要因に有意な主効果があり(表1), 440Hz条件の方が880Hz条件よりd'が高かった ($F_{(1,35)} = 2286.49, p < 0.001$)。声楽家要因×F0要因×母

表1 声楽家要因・基本周波数要因・母音要因におけるd'についての分散分析表

変動因	SS	df	MS	F	p
A	0.36	1	0.36	0.57	0.46
AS	22.45	35	0.64		
B	709.04	1	709.04	2286.49	0.00***
AB	0.68	1	0.68	2.20	0.15
ABS	10.85	35	0.31		
C	49.29	9	5.48	35.76	0.00***
AC	1.27	9	0.14	0.92	0.50
ACS	48.24	315	0.15		
BC	42.21	9	4.69	27.17	0.00***
ABC	3.38	9	0.38	2.17	0.02*
ABCS	54.37	315	0.17		

† $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

注: Aは声楽家要因, Bは基本周波数要因, Cは母音要因, Sは実験参加者を示す。

音要因の有意な交互作用がみられたため ($F_{(9, 315)} = 2.17, p < 0.05$), 単純交互作用検定を行い(表2), 有意な単純交互作用があった場合は単純・単純主効果検定を行った(表3)。

単純交互作用検定の結果, /u/-/o/条件における声楽家要因とF0要因の有意な単純交互作用があった ($F_{(1,350)} = 4.20, p < 0.05$)。単純・単純主効果検定を行った結果, 440Hz条件での/u/-/o/条件において声楽家要因の有意な単純・単純主効果があり, 声楽家群でd'が高かった ($F_{(1,700)} = 5.18, p < 0.05$)。

表2 声楽家要因×基本周波数要因×母音要因交互作用における単純交互作用検定

	SS	df	MS	F	p
AB (a-i)	0.11	1	0.11	0.62	0.43
AB (a-u)	0.17	1	0.17	0.94	0.33
AB (a-e)	0.03	1	0.03	0.15	0.70
AB (a-o)	0.00	1	0.00	0.01	0.91
AB (i-u)	2.00	1	2.00	10.76	0.00**
AB (i-e)	0.15	1	0.15	0.80	0.37
AB (i-o)	0.60	1	0.60	3.20	0.07 [†]
AB (u-e)	0.11	1	0.11	0.62	0.43
AB (u-o)	0.78	1	0.78	4.20	0.04*
AB (e-o)	0.09	1	0.09	0.51	0.48
誤差項		350	0.19		
AC (440Hz)	0.92	9	0.10	0.63	0.78
AC (880Hz)	3.73	9	0.41	2.55	0.01**
誤差項		630	0.16		
BC (一般群)	17.99	9	2.00	11.58	0.00***
BC (声楽家群)	27.59	9	3.07	17.76	0.00***
誤差項		315	0.17		

[†] $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

注：Aは声楽家要因，Bは基本周波数要因，Cは母音要因，Sは実験参加者を示す。

/i/-/u/条件における声楽家要因とF0要因の有意な単純交互作用があった ($F_{(1,350)}=10.76$, $p < 0.01$)。単純・単純主効果検定を行った結果，440Hz条件での/i/-/u/条件における単純・単純主効果が有意傾向にあり，声楽家群でd'が高かった ($F_{(1,700)}=2.87$, $p < 0.1$)。一方で，880Hz条件での/i/-/u/条件における声楽家要因の有意な単純・単純主効果があり，一般群でd'が高かった ($F_{(1,700)}=8.12$, $p < 0.01$)。

一般群におけるF0要因と母音要因の有意な単純交互作用があった ($F_{(9,315)}=11.58$, $p < 0.001$)。声楽家群におけるF0要因と母音要因の有意な単純交互作用があった ($F_{(9,315)}=17.76$, $p < 0.001$)。単純・単純主効果検定の結果，表3より，一般群でも声楽家群でも母音要因の各条件に周波数要因の有意な単純・単純主効果があり，440Hz条件でd'が高かった。声楽家要因の各条件と周波数要因の各条件における母音要因の単純・単純主効果があった。440Hz条件における母音要因の単純・単純主効果についてライアン法を用いて多重比較を

行った結果を表4に示す。表4より，他の母音の組み合わせと比較して，u-o条件は有意にd'が低かった。880Hz条件における母音要因の単純・単純主効果についてライアン法を用いて多重比較を行った結果を表5に示す。

3.2 反応時間

反応時間について声楽家要因×F0要因×母音要因の分散分析を行ったところ，すべての要因についての主効果，声楽家要因とF0要因の交互作用 ($F_{(1,35)}=6.71$, $p < 0.05$) が有意であり，声楽家要因×F0要因×母音要因の交互作用が有意傾向にあった ($F_{(9,315)}=1.85$, $p < 0.1$) (表6)。下位検定の結果，440Hz条件における/u/-/o/条件の弁別の困難性 (表4参照) に起因する反応時間の遅れはみられなかったため，図3には一般群と声楽家群，440Hz条件と880Hz条件ごとに反応時間をまとめた。反応時間は，440Hz条件では880Hz条件と比較して速く，声楽家群と一般群を比較すると，声楽家群でより遅い傾向にあった。

声楽的発声における母音知覚

表3 声楽家要因×基本周波数要因×母音要因交互作用における単純・単純主効果検定

	SS	df	MS	F	p
A (440Hz, a-i)	0.01	1	0.01	0.03	0.86
A (440Hz, a-u)	0.00	1	0.00	0.00	0.96
A (440Hz, a-e)	0.03	1	0.03	0.17	0.68
A (440Hz, a-o)	0.03	1	0.03	0.16	0.69
A (440Hz, i-u)	0.56	1	0.56	2.87	0.09 [†]
A (440Hz, i-e)	0.17	1	0.17	0.86	0.35
A (440Hz, i-o)	0.08	1	0.08	0.44	0.51
A (440Hz, u-e)	0.05	1	0.05	0.24	0.62
A (440Hz, u-o)	1.01	1	1.01	5.18	0.02 [*]
A (440Hz, e-o)	0.00	1	0.00	0.02	0.88
A (880Hz, a-i)	0.16	1	0.16	0.82	0.37
A (880Hz, a-u)	0.38	1	0.38	1.95	0.16
A (880Hz, a-e)	0.17	1	0.17	0.90	0.34
A (880Hz, a-o)	0.01	1	0.01	0.06	0.81
A (880Hz, i-u)	1.58	1	1.58	8.12	0.00 ^{**}
A (880Hz, i-e)	0.02	1	0.02	0.10	0.76
A (880Hz, i-o)	0.64	1	0.64	3.30	0.07 [†]
A (880Hz, u-e)	0.48	1	0.48	2.50	0.11
A (880Hz, u-o)	0.06	1	0.06	0.32	0.57
A (880Hz, e-o)	0.25	1	0.25	1.31	0.25
誤差項		700	0.19		
B (一般群, a-i)	61.82	1	61.82	331.73	0.00 ^{***}
B (一般群, a-u)	17.39	1	17.39	93.30	0.00 ^{***}
B (一般群, a-e)	53.77	1	53.77	288.54	0.00 ^{***}
B (一般群, a-o)	15.34	1	15.34	82.33	0.00 ^{***}
B (一般群, i-u)	32.59	1	32.59	174.90	0.00 ^{***}
B (一般群, i-e)	42.90	1	42.90	230.20	0.00 ^{***}
B (一般群, i-o)	48.63	1	48.63	260.96	0.00 ^{***}
B (一般群, u-e)	24.55	1	24.55	131.76	0.00 ^{***}
B (一般群, u-o)	17.82	1	17.82	95.64	0.00 ^{***}
B (一般群, e-o)	36.04	1	36.04	193.41	0.00 ^{***}
B (声楽家群, a-i)	69.58	1	69.58	373.38	0.00 ^{***}
B (声楽家群, a-u)	12.81	1	12.81	68.73	0.00 ^{***}
B (声楽家群, a-e)	50.34	1	50.34	270.15	0.00 ^{***}
B (声楽家群, a-o)	15.87	1	15.87	85.18	0.00 ^{***}
B (声楽家群, i-u)	59.47	1	59.47	319.11	0.00 ^{***}
B (声楽家群, i-e)	50.35	1	50.35	270.19	0.00 ^{***}
B (声楽家群, i-o)	65.04	1	65.04	349.03	0.00 ^{***}
B (声楽家群, u-e)	20.04	1	20.04	107.53	0.00 ^{***}
B (声楽家群, u-o)	29.95	1	29.95	160.70	0.00 ^{***}
B (声楽家群, e-o)	31.01	1	31.01	166.42	0.00 ^{***}
誤差項		350	0.19		
C (一般群, 440Hz)	8.41	9	0.93	5.74	0.00 ^{***}
C (一般群, 880Hz)	31.44	9	3.49	21.45	0.00 ^{***}
C (声楽家群, 440Hz)	3.32	9	0.37	2.26	0.02 [*]
C (声楽家群, 880Hz)	52.98	9	5.89	36.14	0.00 ^{***}
誤差項		630	0.16		

[†] $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

注：Aは声楽家要因，Bは基本周波数要因，Cは母音要因，Sは実験参加者を示す。

表4 440 Hz条件における母音要因の多重比較表

	a-i	a-u	a-e	a-o	i-u	i-e	i-o	u-e	u-o	e-o
a-i		2.66	2.63	2.62	2.57	2.63	2.65	2.68	2.07*	2.59
a-u			1.25	1.24	1.19	1.25	1.27	1.30	0.69*	1.21
a-e				2.37	2.32	2.38	2.40	2.43	1.82*	2.34
a-o					1.26	1.32	1.34	1.37	0.76*	1.28
i-u						2.27	2.29	2.32	1.71*	2.23
i-e							2.27	2.30	1.69*	2.21
i-o								2.52	1.91*	2.43
u-e									0.95*	1.47
u-o										2.12*
e-o										

* $p < 0.05$

注：表の値は平均値の差であり，MSe=0.16, df=630であった。

表5 880 Hz条件における母音要因の多重比較表

	a-i	a-u	a-e	a-o	i-u	i-e	i-o	u-e	u-o	e-o
a-i		1.38*	0.25*	1.31*	0.36*	0.38*	0.16	1.12*	0.47*	0.68*
a-u			1.13*	0.07	1.02*	1.00*	1.22*	0.26*	0.91*	0.70*
a-e				1.06*	0.11	0.13	0.09	0.87*	0.22	0.43*
a-o					0.95*	0.93	1.15*	0.19	0.84*	0.63*
i-u						0.02*	0.20	0.76*	0.11	0.32*
i-e							0.22	0.74*	0.09	0.30*
i-o								0.96*	0.31*	0.52*
u-e									0.65*	0.44*
u-o										0.21
e-o										

* $p < 0.05$

注：表の値は平均値の差であり，MSe=0.16, df=630であった。

表6 声楽家要因・基本周波数要因・母音要因における反応時間についての分散分析表

変動因	SS	df	MS	F	p
A	54655993.43	1	54655993.43	11.40	0.00**
AS	167832235.05	35	4795206.72		
B	136566452.83	1	136566452.83	53.21	0.00***
AB	17219335.17	1	17219335.17	6.71	0.01*
ABS	89830530.05	35	2566586.57		
C	9373669.20	9	1041518.80	7.35	0.00***
AC	1068567.74	9	118729.75	0.84	0.58
ACS	44666210.23	315	141797.49		
BC	6035319.27	9	670591.03	7.18	0.00***
ABC	1552921.87	9	172546.87	1.85	0.06 [†]
ABCS	29404651.00	315	93348.10		

[†] $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

注：Aは声楽家要因，Bは基本周波数要因，Cは母音要因，Sは実験参加者を示す。

声楽的発声における母音知覚

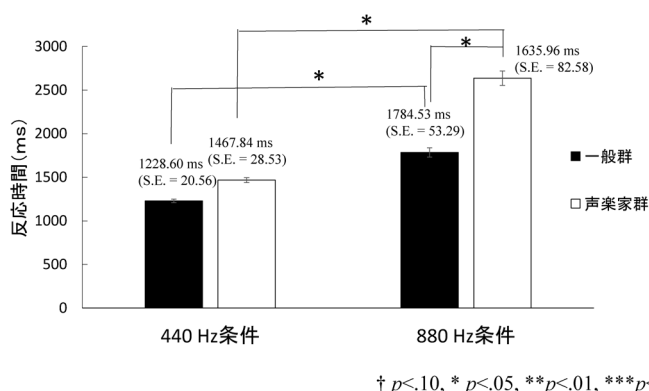


図3 440 Hz条件と880 Hz条件ごとの声楽家群と一般群における反応時間の比較

上記の傾向は統計的にも支持され、声楽家要因とF0要因の交互作用について、単純主効果検定を行ったところ、880 Hz条件における声楽家要因の有意な単純主効果があり、声楽家群で反応時間が一般群より長かった ($F_{(1,70)} = 18.10$, $p < 0.05$)。一般群におけるF0要因の有意な単純主効果があり、880 Hz条件において反応時間が440 Hz条件より長かった ($F_{(1,35)} = 11.07$, $p < 0.05$)。声楽家群におけるF0要因においても有意な単純主効果があり、880 Hz条件において反応時間が440 Hz条件より長かった ($F_{(1,35)} = 48.85$, $p < 0.05$)²⁾。

4. 考察

本研究は声楽的に発声された母音の知覚についてF0と声楽経験の影響を明らかにすることを目的とした。実験の結果から、880 Hzという高いF0では母音の弁別が困難となることが示された(図2)。声楽経験については母音の知覚の成績全体の向上に影響を及ぼすものではなく、440 Hz条件における/u/と/o/の母音のペアについての知覚成績のみ声楽経験者で高くなったこと、880 Hz条件の/i/と/u/のペアでは声楽経験者で弁別が困難となったことから、ある特定の条件で声楽経験が母音の知覚に影響を及ぼすことが明らかになった。

440 Hz条件において/u/と/o/の母音のペアは、他と比較して全体的にd'が有意に低く、弁別が困難であったが、このペアでのみ、声楽家群の方で

有意にd'が高く弁別の成績が良かった。刺激発話者のフォルマントを地声と刺激の440 Hz条件で比較すると440 Hz条件でのみ/u/と/o/の母音はF1, F2空間で近接していた。このために全体的にこのペアでd'が低くなったと考えられる。また、声楽家群の方が一般群より成績が良かったのは声楽経験が関係しており、声楽家は歌声における/u/と/o/の音響特徴を聞きなれているためであると考えられる。

高いF0で母音の知覚が困難になったことについては、Morozov (1965) が示す通り、880 Hz条件では倍音構造が希薄になり各母音のフォルマント構造が類似するためであると考えられる。すなわち、図1dが示すように、880 Hz条件ではF1がF0とほぼ同じ値を示しており、通常の話声と異なるフォルマント分布であったためとも考えられる。

880 Hz条件では、全体的に/i/に対する他の母音との弁別成績が低いが、声楽家群において特に成績が低くなる傾向があった。特に/i/と/u/のペアは声楽家群で有意にd'が小さく、一般群より弁別が困難となっていた。一般群ではd'が正の値であったにも関わらず、声楽家群ではd'が負の値であり、声楽家群におけるd'がより0に近い値を示していることから、声楽家群では母音の弁別が困難かつ知覚が逆転していることが示唆された。図1dより、880 Hz条件において/i/と/u/ではF1はほぼ同じ値であるがF2は/i/の方がわずかに高かった。このことから一般群ではF2を手がか

りに/i/と/u/の母音を弁別できたが、声楽家群ではF2を手がかりにせず、別の方略によって母音の弁別を試みたためd'が低くなったと推測することもできる。880Hz条件において声楽家群は一般群よりも有意に反応時間が長かったが、この反応時間の結果も声楽家群の弁別の方略が異なるという解釈を支持する。

それでは声楽経験者はどのような方略を用いて880Hz条件で母音の知覚を行ったのであろうか。声楽経験者は880Hz条件のような高いピッチの発声では、F1とF0の周波数を接近させることによりF1のパワーを上げる(Sundberg 1977)ため、通常の話声で使用されるようなフォルマントの手がかりを無視していた可能性が考えられる。通常の話声におけるF1の周波数より発声されたF0が低い場合、通常の話声のフォルマント周波数が維持されるが、通常の話声におけるF1の周波数より発声されたF0が高い場合、そのF1の周波数は通常の話声における高さではなく、発声されたF0と同じ高さまで引き上げられる。このメカニズムは高いF0において振幅を増幅する歌手の方略に由来する。ソプラノ歌手が通常の話声と同じ状態で調音を行った場合、声道で振幅を増幅可能な周波数に倍音をあてはめることができない。ソプラノ歌手は顎の開きによりF1をF0まで引き上げることで倍音を共鳴させ、高いF0においても大きな振幅を得ることが可能となる(Lindblom and Sundberg 1971)。発声において音圧レベルを上げるには呼気流のエネルギーを強める方法が一般的だが、声楽家はこの呼気流のエネルギーにより声帯を痛めてしまう可能性がある。F1をF0に近づけることで得られる振幅の増幅により、高いF0での発声において効率的に音圧レベルを上げることが可能になる(Sundberg 1987)。上記の顎の開きによる発声法により、調音が困難になる狭母音/i/の母音で通常の話声と異なる調音となる³⁾。声楽経験者はこの発声における調音の経験から、高周波数の声楽的発声の母音知覚では、特に/i/の母音の場合にF1およびF2に注意を向けず、一般群のようなF2を手がかりとした弁

別ができなかったため/i/-/u/や/i/-/o/の弁別が困難となったと可能性がある。また、/i/以外の母音の組み合わせにおいては、声楽家群でもF2を手がかりとした母音の弁別を行っていた可能性も考えられる。

本研究では3名の歌手の歌声を用いたため、話者ごとに正答率が異なる可能性が考えられる。付録2より、440Hz条件では/a/と/o/について話者2で近接していたが、正答率は話者間で有意差がみられなかった。440Hz条件の/u/と/o/では話者3で特に近接しており、正答率も話者3で有意に低かった(話者2>話者3: $t_{(72)}=6.48, p<0.01$, 話者1>話者3: $t_{(72)}=5.05, p<0.01$)。話者の歌唱の音響特性により、正答率にも違いがみられた。

上述したF1やF2を弁別の手がかりに用いない声楽家群の知覚方略について更なる検討を行うため、880Hz条件の音響特性について、更なる解析を試みた。まずはF3を母音弁別の手がかりとした可能性を検討したが、d'とF3の母音間の周波数の差にはどの群にも有意な相関はなかった。

880Hz条件についてはフォルマントの分析の限界も考えられる。LPCを用いたフォルマント分析は母音の音響的な情報を示す一般的な手法として用いられているために本研究でも報告したが、F0が880Hzの場合、各倍音をスペクトル包絡のピークとして検出してしまうため、結果的に図1cと図1dは第1倍音と第2倍音の分布を示していることになっているかもしれない。また、本研究の880Hz条件のように、スペクトルが疎な場合には、スペクトル包絡のピークが知覚的な手掛かりになっていない可能性もある。よって、倍音構造そのものに注目すると、倍音の強さが母音間で異なる(図4)。そのため、第1倍音から第4倍音までの平均の強さ(intensity)を求め(表7)、各倍音の振幅および倍音間の振幅の差分を算出し、倍音構造の指標とした。これらの指標について、各母音の組み合わせごとにさらに差分を求め、その値とd'との相関を求めた。その結果、どの群でも有意な相関はみられなかった。今回の解析からは倍音構造が声楽経験者の方略と関係していること

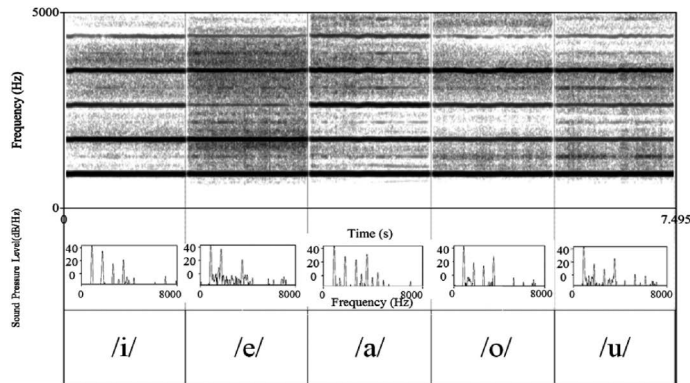


図4 880 Hz条件における各母音のスペクトログラム(上)とスペクトル(下)

表7 880 Hzにおける第1倍音から第4倍音までの平均の強さ (dB)

	/i/	/e/	/a/	/o/	/u/
H1	69.27	69.39	69.20	69.67	69.70
H2	57.37	56.86	55.95	48.41	48.60
H3	47.26	42.03	52.44	48.97	47.12
H4	49.68	52.26	54.18	53.06	47.01

は示されなかった。今後、声楽経験者における高周波数の母音知覚の方略について合成音声などにより細かく刺激を設定した実験を行うことで考察する必要がある。

本研究ではF0が2倍も異なる条件で比較したため、F0の影響が強く出たが、440Hzから880Hzの間にある高さではどのように母音の知覚が変化し、その変化には声楽経験の影響が現れるかを検討することができなかつた。今後の課題として、440Hzから880Hzの間にある周波数や440Hzより低い周波数、880Hzより高い周波数など、他のさまざまなF0で発声された母音の知覚における声楽経験の影響を今後の実験で検討する必要がある。

本研究では、声楽的に発声されたF0の高い母音は弁別が困難になることが示され、先行研究と一貫した結果が得られた。さらに本研究の結果、これまでにほとんど検討されていなかった声楽的発声の母音知覚における声楽経験の影響を示すことができた。すなわち、声楽的な発声の経験が母音の弁別に影響することが示された。F0が

440Hzの場合、声楽的な発声の経験により弁別が容易になったが、880Hzでは声楽的な発声の経験がバイアスとなり知覚が困難になることがd'の結果および反応時間から示された。これらの結果をさらに発展させることで発声経験という運動能力と知覚能力との関係にもさまざまな示唆を与えるものと思われる。

謝 辞

本研究は科学研究費基盤研究(A) 15H01691、戦略的研究基盤形成支援事業の補助を受けて行われた。本研究は第30回日本音声学全国大会で発表をした内容を元に新たな解析、考察を加えたものである。本研究にご協力下さったDonna Erickson先生、刺激作成にあたり音声データを提供して下さいましたソプラノ歌手の皆様、聴取実験の参加者の皆様、学会にて貴重な助言を下された皆様に感謝申し上げます。

〔注〕

- 1) 音圧レベルは母音によって異なり、その違いが母音の知覚に影響している可能性もある。特に開母音では口が大きく開くために音圧が高くなるといった指摘もあるが、開口度と音圧レベルの相関は必ずしも安定しているとはいえない(Kawahara, Erickson and Suemitsu 2017, Paker 2008)。音圧レベルが歌声における母音知覚の弁別に関わっている可能性については、今後の研究の課題とする。

- 2) 反応時間のデータはシャピロ-ウィルク検定により正規分布に従わなかったため ($p < 0.05$), ノンパラメトリック検定で有意差が出た比較を再検討したところ, 同様の結果が得られた。880Hz条件における一般群と声楽家群の比較では, ウィルコクソンの順位和検定の結果, 声楽家群で有意に反応時間が長くなった ($Z=9.26, p < 0.001$)。ウィルコクソン符号付き順位検定の結果, 一般群では440Hz条件と比較し880Hz条件で有意に反応時間が長くなり ($Z=11.71, p < 0.001$), 声楽家群でも440Hz条件と比較し880Hz条件で有意に反応時間が長くなった ($Z=11.17, p < 0.001$)。
- 3) 通常の話声での/i/と880Hzにおける/i/について, 開口度の観点だけでなく, 舌の高さの観点からも調音が特殊であった可能性がある。ソプラノ歌手の母音の調音について, X線写真を用い舌の状態を調べた研究によると, 通常の話声では異なる状態にある母音の調音が, F0が高い場合 (940Hz), 比較的后寄りの状態に収束する (Johansson, Sundberg and Wilbrand 1985)。このことから, 通常の話声では前舌母音である/i/の調音と比較して, 880Hz条件の/i/の調音は後寄りの特殊な状態であったと考えられる。

参考文献

- Boersma, P. and D. Weenink (2015) "Praat: Doing phonetics by computer [Computer program]." Version 5.4.15, <http://www.praat.org/> (accessed August 1, 2015)
- Fine, P. and J. Ginsborg (2007) "How singers influence the understanding of sung text." *Proceedings of the International Symposium on Performance Science, European Association of Conservatoires*, 978–990.
- Fine, P., J. Ginsborg and C. Barlow (2009) "The influence of listeners' singing experience and the number of singers on the understanding of sung text." *Proceedings of the International Symposium on Performance Science*, 51–56.
- Ginsborg, J., P. Fine and C. Barlow (2011) "Have we made ourselves clear? Singers' and non-singers' perceptions of the intelligibility of sung text." *Proceedings of the International Symposium on Performance Science, European Association of Conservatoires*, 111–116.
- Gottfried, T. L. and S. L. Chew (1986) "Intelligibility of vowels sung by a countertenor." *The Journal of the Acoustical Society of America* 79 (1), 124–130.
- Jesse, A. and D. W. Massaro (2010) "Seeing a singer helps comprehension of the song's lyrics." *Psychonomic bulletin and Review* 17(3), 323–328.
- Johansson, C., J. Sundberg and H. Wilbrand (1985) "X-ray study of articulation and formant frequencies in two female singers." *Proceedings of SMAC* 83(1), 203–218.
- Kawahara, S., D. Erickson and A. Suemitsu (2017) "The phonetics of jaw displacement in Japanese vowels." *Acoustic Science and Technology* 38(2), 99–107.
- Lindblom, B. E. and J. Sundberg (1971) "Acoustical consequences of lip, tongue, jaw, and larynx movement." *The Journal of the Acoustical Society of America* 50(4B), 1166–1179.
- Majewski, W. and H. Hollien (1967) "Formant frequency regions of Polish vowels." *The Journal of the Acoustical Society of America* 42 (5), 1031–1037.
- Macmillan, N. and D. Creelman (2005) *Detection theory: A user's guide*. 2nd edition. Mahwah: Lawrence Erlbaum Associates Publishers.
- Morozov, V. P. (1965) "Intelligibility in singing as a function of fundamental voice pitch." *Soviet Physics-Acoustics* 10, 279–283.
- Nelson, H. D. and W. R. Tiffany (1968) "The intelligibility of song." *The NATS Bulletin* 25, 22–28.
- Parker, S. (2008) "Sound level protrusions as physical correlates of sonority." *Journal of Phonetics* 36, 55–90.
- Smith, L. A. and B. L. Scott (1980) "Increasing the intelligibility of sung vowels." *The Journal of the Acoustical Society of America* 67 (5), 1795–1797.
- Sundberg, J. (1977) "Singing and timbre." *Music room acoustics* 17, 57–81.
- Sundberg, J. (1987) *The science of the singing voice*. North-east Illinois: Northern Illinois University Press.

(Received Dec. 24, 2016, Accepted Jul. 20, 2017)

声楽的発声における母音知覚

付録1 440HzにおけるF1, F2 (Hz)

	/i/		/e/		/a/		/o/		/u/	
	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
S1	448.95	2153.40	821.49	1903.43	753.05	1314.92	527.93	897.04	445.70	1354.52
S2	449.81	2246.31	893.60	2083.58	900.59	1387.54	876.32	1253.02	379.91	1441.37
S3	455.99	2211.52	757.31	1778.05	842.92	1316.31	469.47	918.73	446.71	951.20

付録2 880HzにおけるF1, F2 (Hz)

	/i/		/e/		/a/		/o/		/u/	
	F1	F2	F1	F2	F1	F2	F1	F2	F1	F2
S1	878.16	1756.77	878.73	1757.16	879.54	1747.50	878.69	1678.44	878.94	1615.16
S2	882.32	1763.12	882.40	1750.78	880.55	1761.47	881.03	1751.79	881.53	1753.95
S3	878.40	1705.43	881.02	1739.18	879.12	1724.56	881.22	1676.84	880.57	1644.97