

声楽的発声における母音知覚—声楽経験および F0 の影響—

遠藤希美(東京大学大学院総合文化研究科), 川原繁人(慶應義塾大学言語文化研究所),
皆川泰代 (慶應義塾大学文学部)

1. 目的

声楽家の歌声は通常の話声と比較して聴取が困難であるといわれている。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、音響的に全く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シラブルの母音の知覚について、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=390 Hz 以上、女性歌手の場合は F0=523 Hz 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、スペクトル上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。人工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 & Hollien, 1967)。歌声の母音の知覚の成績は隣接する子音や歌声の音色、発声法、口の動き、歌詞に対する知識など、様々な手がかりによって向上する (Gottfried & Chew, 1986 ; Jesse & Massaro, 2010 ; Ginsborg, Fine, & Barlow, 2011)。しかし、実験対象者の声楽経験の有無によって母音の知覚が向上するかを検討した先行研究はない。声楽の経験者は声楽的な発声による音声の知覚が容易になるか。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、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。参加者は全員女性であり、聴覚に異常はなかった。

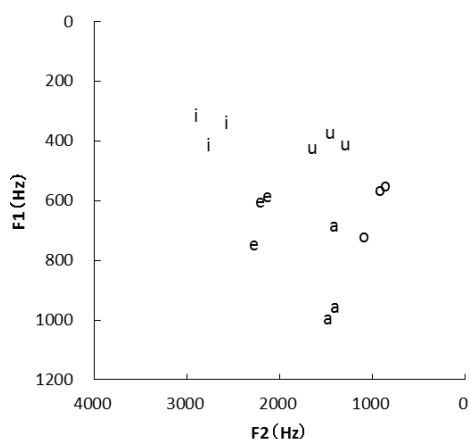
2.2. 刺激

刺激には音楽大学の声楽専攻を卒業したソプラノ歌手 3 名の音声を用いた。刺激に用いた音声は/a/, /i/, /u/, /e/, /o/の 5 母音であり、各母音につき F0=440 Hz と F0=880 Hz であった。刺激の長さは 1500 ms であり、刺激の音圧レベルを 75 dB に統一した。

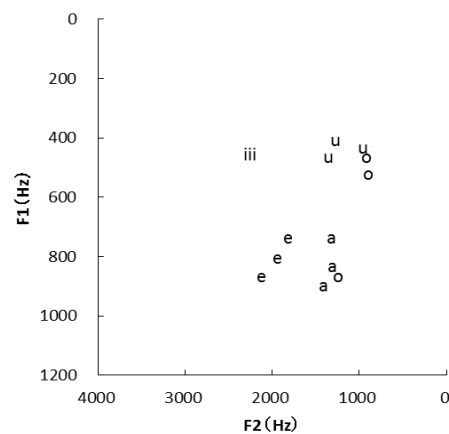
録音は防音室で行った。録音にはコンピュータ (Probook4730s, HP) にアンプ (US-366, TASCAM) を接続し、アンプにマイク (ATM98, audio-technica) を接続した装置を用いた。

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（Praat, University of Amsterdam）を用いた。各母音について F0 ごとに 5000 ms 程度発生してもらった。F0=440 Hz の場合には誤差上下 1 Hz 以内、F0=880 Hz の場合には誤差上下 2 Hz 以内となった部分を 1500 ms 切り取った。切り取った音声を音圧が 75 dB となるように加工した。

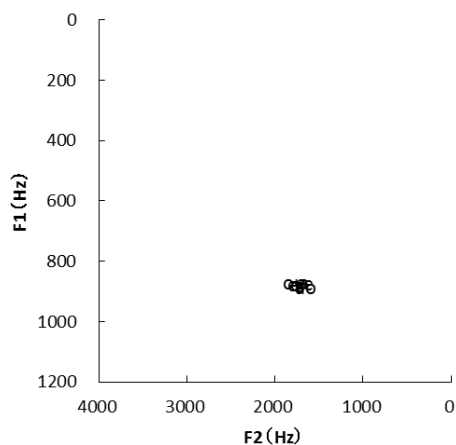
解析の参考として、地声における /a/, /i/, /u/, /e/, /o/ を録音した。地声の F0 の平均は 225 Hz であった。3 話者の地声におけるフォルマント平均分布を図 1a, F0=440 Hz におけるフォルマント平均分布を図 1b, F0=880 Hz におけるフォルマント平均分布を図 1c に示した。また、F0=880 Hz における拡大されたフォルマント平均分布を図 1d に示した。



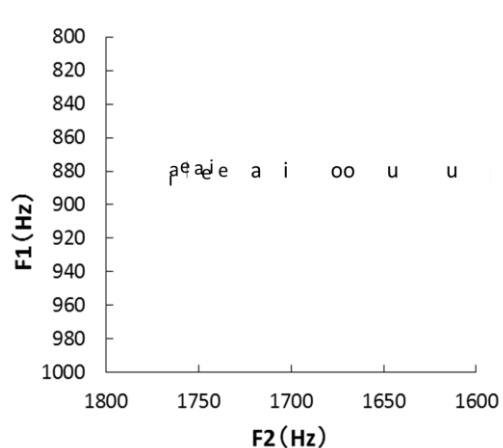
(a) 地声におけるフォルマント分布



(b) F0=440 Hz におけるフォルマント分布



(c) F0=880 Hz におけるフォルマント分布



(d) F0=880 Hz における拡大されたフォルマント分布

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

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,

FUJITSU) とヘッドフォン (HD 280 pro 64 Ω, 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, Brügel & Kjaer) と騒音計 (2250, Brügel & Kjaer) 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た (二肢強制選択法: 2AFC 方式)。歌手ごとに各 F0 で 5 つの母音を総当たりで組み合わせた。2 つの刺激の間隔は 500 ms であった。組み合わせた刺激を 8 回繰り返す、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。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。刺激呈示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. 結果の処理法

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} の値 $(Z(\text{Hit}) - Z(\text{FA})) / \sqrt{2}$ を参加者ごとに算出した。刺激対の呈示順序 2 通りのうち一方を呈示順序 A、もう一方を呈示順序 B とし、呈示順序 A の正答率を P_A (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 P_A (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 & Creelman 2005: 167-170)。歌手 3 人分、すなわち 3 ブロック分をまとめた上で、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。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上で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

歌手 3 人分、すなわち 3 ブロック分をまとめた上で各刺激対について、母音の組み合わせを母音要因とし、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。紙面の都合上一部省略し主要な結果のみ述べる。

各 F0 における各母音について声楽家群と一般群の d' を示したグラフが図 2 である。分散分析の結果、声楽家要因と F0 要因と母音要因に有意な交互作用があった ($F_{(9, 315)}=2.17$,

$p < 0.05$). F0 要因に有意な主効果があり、440 Hz 条件の方が 880 Hz 条件より d' が高かった ($F_{(1,35)}=2286.49, p < 0.05$). 声楽家要因と F0 要因と母音要因に有意な交互作用があったため、単純交互作用検定を行った。/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り ($F_{(1,350)}=10.76, p < 0.05$), 単純・単純主効果検定を行った結果、880 Hz 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 ($F_{(1,700)}=8.12, p < 0.05$). /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り ($F_{(1,350)}=4.20, p < 0.05$), 単純・単純主効果検定を行った結果、440 Hz 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 ($F_{(1,700)}=5.18, p < 0.05$). 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 ($F_{(9,315)}=11.58, p < 0.05$). 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 ($F_{(9,315)}=17.76, p < 0.05$). 詳細な統計値は省略するが、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、440 Hz 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。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は省略した。

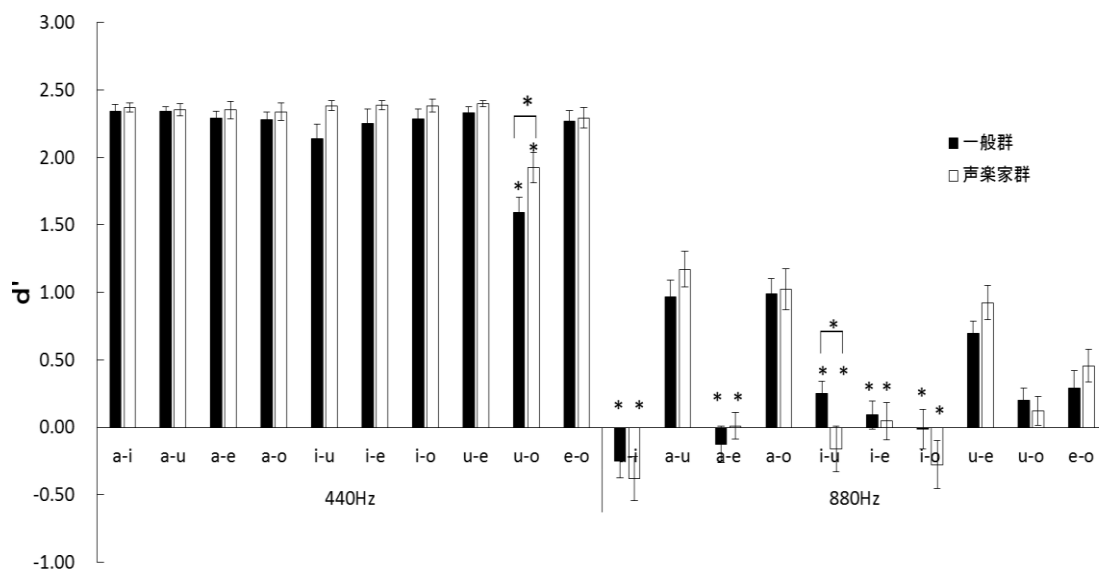


図 2: F0 要因の各条件における各母音条件の声楽家群と一般群の d'

3.2. 反応時間

各刺激対について 3 ブロック分をまとめた上で反応時間を参加者ごとに算出した。反応時間について声楽家要因 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。声楽家要因と F0 要因の有意な交互作用があった ($F_{(1,35)}=6.71, p < 0.05$)。声楽家要因と F0 要因

の有意な交互作用があったため、単純主効果検定を行った。図3は一般群と声楽家群、440 Hz条件と880 Hz条件ごとに反応時間をまとめたグラフである。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 ($F_{(1,35)}=48.85$, $p<0.05$)。

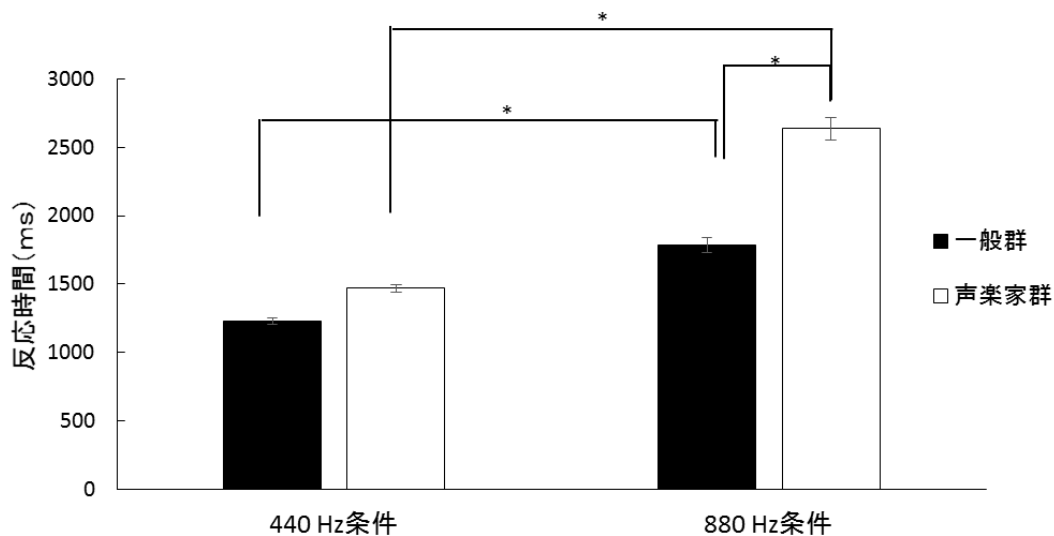


図3: 440 Hz条件と880 Hz条件ごとの声楽家群と一般群における反応時間の比較

4. 考察

本研究の目的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調べることであった。実験の結果から、F0が高い方が母音の知覚が困難であることが示された。Morozov (1965)が示す通り、F0=880 Hz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。声楽経験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条件における/u/と/o/の母音のペアは、他の母音のペアと比較して全体的にd'が有意に低かった。/u/と/o/の母音のペアのみ、声楽家群の方で有意にd'が高く知覚の成績が良かった。地声のフォルマントと刺激のフォルマントの特徴を440 Hz条件で比較すると440 Hzの/u/と/o/の母音は非常に近い特徴を持っていたが、声楽家群の方が一般群より成績が良かったのは、声楽家はその特徴を聞きなれているためであると考えられる。

880 Hz 条件において、/i/と/u/のペアは声楽家群で有意に d' が小さく、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' が負の値であった。図 1d より、880 Hz 条件において/u/と/i/では F1 に変化はないが F2 は/i/の方がわずかに高かった。反応時間の結果より、880 Hz 条件において声楽家群は一般群よりも反応時間が長かったことから、一般群では F2 を手がかりにし/u/と/i/の母音を弁別したが、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であることが示唆された。また、880 Hz 条件では、声楽家群・一般群ともに/i/に対する他の母音との弁別成績が低く、声楽家群において特に成績が低くなる傾向があった。図 1d によると、/i/の母音は他の母音と比較して F2 にばらつきがあり弁別が困難な音響特性を持っている可能性がある。このことから、未経験者は F2 を手がかりにして弁別を行うが、声楽経験者は他の方略を用いたと考えられる。声楽経験者はどのような方略を用いて $F_0=880$ Hz で母音の知覚を行ったか、人工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 F_0 が 2 倍も異なる条件で比較したため、 F_0 の影響が強く出たが、440 Hz から 880 Hz 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った。今後の課題として、440 Hz から 880 Hz の間にある周波数や 440 Hz より低い周波数、880 Hz より高い周波数など、他のさまざまな F_0 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したい。

参考文献

- Ginsborg, J., Fine, P., & Barlow, C. (2011). Have we made ourselves clear? Singers' and non-singers' perceptions of the intelligibility of sung text. In A. Williamon, D. Edwards, & L. Bartel (Eds.), *Proceedings of the International Symposium on Performance Science* (pp. 111-116). European Association of Conservatoires (AEC).
- Gottfried, T. L., & Chew, S. L. (1986). Intelligibility of vowels sung by a countertenor. *The Journal of the Acoustical Society of America*, 79(1), 124-130.
- Jesse, A., & Massaro, D. W. (2010). Seeing a singer helps comprehension of the song's lyrics. *Psychonomic bulletin & review*, 17(3), 323-328.
- Majewski, W., & Hollien, H. (1967). Formant frequency regions of Polish vowels. *The Journal of the Acoustical Society of America*, 42(5), 1031-1037.
- Macmillan, N., & Creelman, D. (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.
- Sundberg, J. (1987). *The Science of the Singing Voice*. Northern Illinois: Northern Illinois University Press.