

音素長と音高が歌声の印象評価に与える影響

深澤 実紅^{†,*1} 北原 玲奈[†] 久保田 新菜[†] 志々目 樹[†] 戸田 菜月^{†,*2} 竹本 浩典^{†,*3}
高橋 純^{††,*4}

[†]千葉工業大学先進工学部 〒275-0016 千葉県習志野市津田沼2丁目17-1

^{††}大阪芸術大学 〒585-0001 大阪府南河内郡河南町東山469

E-mail: ^{*1}s20c3095uz@s.chibakoudai.jp, ^{*2}s17c3087mt@s.chibakoudai.jp, ^{*3}hironori.takemoto@p.chibakoudai.jp,
^{*4}johnandiamo@gmail.com

あらまし オペラ歌唱は豊かな声量、特有の響き、高音での発声の特徴で、これまで主に母音の発声技術が研究されてきた。しかし、歌詞を明瞭に伝えるには子音にも高度な発声技術が必要と考えられるが、その詳細は明らかでなかった。そこで本研究では、プロのオペラ歌手4名と声楽の経験のない学生22名の歌唱音声进行分析し、音素長や発声タイミングの違いを分析した。その結果、プロでは語頭の子音が学生より有意に長いこと、発声タイミングが有意に早いこと、破裂音の閉鎖区間が短い傾向があることなどが明らかになった。さらに、これらの違いが歌唱の印象に及ぼす影響を調べるため、録音した音声の音高や音素長を操作し印象評価実験を行った。その結果、音高のみを正確にした歌声よりも、音高を正確にし、かつ音素長をプロの平均に合わせた歌声の方が、有意に評価が向上する可能性が示唆された。

キーワード オペラ歌唱、発声技術、子音、音素長、発声タイミング、歌のうまさ

1. はじめに

歌唱力、すなわち歌のうまさに関する研究は主に母音に着目して行われてきた。歌手のフォルマントの有無やヴィブラートの程度、音程、すなわち基本周波数とその動的成分など、母音の音響特性が歌のうまさに大きく影響することが知られている[1-3]。また、相対音高やヴィブラートに関する音響特徴量を用いて歌唱力の評価を自動で行う研究もなされている[4]。一方、子音の歌唱技術や、子音が歌のうまさに及ぼす影響についての研究は少数である。例えば、子音は母音の知覚を助けるという研究[5]や、男性の訓練された歌手は歌手でない者より無声破裂子音の有声開始時間が長い傾向があるという研究[6,7]、特定の子音の長さが歌唱のグルーブ感に影響を与えるという研究がある程度である[8]。

そこでわれわれは、特に子音に着目して歌のうまさを研究している。本研究では、プロの声楽家（プロ）4名と声楽指導を受けたことのない学生（学生）22名が歌唱した武島羽衣作詞・瀧廉太郎作曲の「花」を録音し、その歌声に音素ラベリングを行って子音および母音の長さ（音素長）や発声タイミングの違いについて比較検討した。そして、これらのプロと学生の違いが歌のうまさを印象に与える影響を検討するため、学生の歌声の音素長や発声タイミングをプロに近づけた場合やプロの歌声を学生の音素長や発声タイミングに近づけた場合の印象を評価したので報告する。



Fig. 1 アンケートで示した楽譜

2. 材料と方法

2.1. アンケート調査

Fig. 1は「花」の冒頭である。言葉を明瞭に伝えるために、①～⑫のどの部分を特に注意して歌うのか、複数選択可能として11名にアンケートした。回答者は、歌唱指導者4名、歌唱指導者兼プロの歌手4名、プロの歌手1名、音楽学校に所属している歌手1名、プロとしては活動していないアマチュア歌手1名であり、全て声楽の指導を受けた経験を有していた。

2.2. 歌唱音声の録音と音素長の分析

プロのオペラ歌手4名（ソプラノ1名:Sop, テノール2名:Ten1, Ten2, バリトン1名:Bar, 以下プロ）と学生22名（男性12名:M01～M12, 女性10名:F01～F10, 以下学生）が武島羽衣作詞・瀧廉太郎作曲の「花」を歌唱した歌声を録音した。実験参加者は、MIDIで作成した「花」の伴奏をイヤホンで聴きながら1番を歌唱した。録音はプロのオペラ歌手は静音な部屋、学生は千葉工業大学の無響室で実験を行った。録音はiPhone（サンプリング周波数:44.1 kHz, 量子化ビット数:16 bit）を用いた。

音声波形に各音素のラベルをPraat [9]を用いて手動で付与し、音素長を計算した。また、各音素に対して楽譜で指示された発声開始時刻と実際の発声開始時刻の

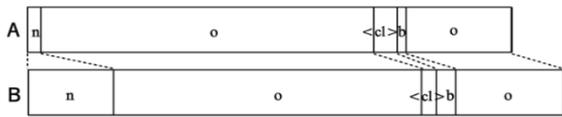


Fig. 2 学生の音声の音素長と発声タイミングの変更 (A: 原音声, B: 音素長伸縮)

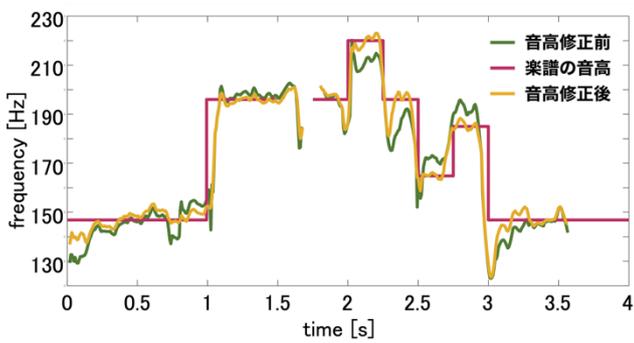


Fig. 3 学生の音声の音高修正

差（発声タイミング）を求めた。そして、プロと一般学生で音素長や発声タイミングに有意な差があるかを検討するために、プロと学生の音素長に対して分散分析を行い、有意水準 5%で Bonferroni 法を用いて多重比較を行った。

2.3. 学生の歌声の編集

22名の学生の歌声から「花」の第2小楽節「上り下りの船人が」の部分を選択して、(1) 原音声、(2) Fig. 2で示すように原音声の発声タイミングと音素長をプロの平均とした歌声（音素長伸縮）、(3) Fig. 3で示すように音高を楽譜通りに修正した歌声（音高修正）、(4) その両方を操作した歌声（音素長伸縮+音高修正）の4種類（合計88種類）を準備した。音高修正が必要であったのは、学生は必ずしも楽譜で指示された音高で歌唱できていなかったためである。そこで、基本周波数の平均値が楽譜で指示された音高に一致するように修正した。なお、全ての編集操作は Praat [9]で行った。

2.4. プロの歌声の編集

4名のプロの歌声から「花」の第2小楽節の部分を選択して、(1) 原音声、(2) 原音声の発声タイミングと音素長を学生の平均とした歌声（音素長伸縮）、(3) 音高を学生の平均に修正した歌声（音高修正）、(4) その両方を操作した歌声（音素長伸縮+音高修正）の4種類（合計16種類）を準備した。なお、全ての編集操作は Praat [9]で行った。

2.5. 歌のうまさの印評価実験

編集した歌声に対して、歌のうまさの印象を評価した。実験参加者は健常な聴力を持つ20名の学生であった。実験参加者は、同一学生の4種類の歌声のいずれか2種類の音声をヘッドホンまたはイヤホンで連続して試聴した後、「どちらの音声の方がどの程度上手に聞こえるのか」を5段階（+2: 1つ目の方がとてもう

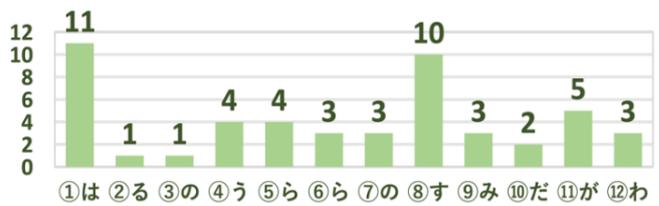


Fig. 4 アンケートの結果

まい, +1: どちらかといえば1つ目の方がうまい, 0: どちらも言えない, -1: どちらかといえば2つ目の方がうまい, -2: 2つ目の方がとてもうまい) で評価した。プロの音声についても同様に評価実験を行なった。評価総数は学生の音声は132回でプロの音声では24回であった。実験回答時間に制限はなく、各音声を一回以上聞いた後に評価した。評価値の集計後、シェフェの一対比較法で分析して尺度図を算出し、有意水準5%の Tukey の HSD 検定で評価に差があるか検討した。

2.6. 自然性の評価実験

原音声も含めて印象評価に用いた全ての歌声の自然性を評価した。実験参加者は健常な聴力を持つ20名の学生であった。実験参加者は、使用したすべての音声の自然性を評価した。ヘッドホンまたはイヤホンで音声をひとつずつ試聴した後、「音声がどの程度自然に聞こえるか」を5段階（+2: 自然である, +1: どちらかといえば自然である, 0: どちらも言えない, -1: どちらかといえば不自然である, -2: 不自然である）で評価した。評価総数は学生の音声88回、プロの音声16回で実験回答時間に制限はなく、各音声を一回以上聞いた後に評価し、20名の平均値を求めた。

3. 結果と考察

3.1. アンケート結果

Fig. 4 はアンケートの結果である。単語の語頭かつ音楽的なアクセントになる「①は」と「⑧す」で頻度が高かった。この結果から、声楽の指導を受けた経験のある歌手はこれらの「語頭」のモーラを言葉を明瞭にするために重要と考えていることが明らかになった。

3.2. 音素長の比較

Table 1 は子音、母音、子音+母音の音素長の3つに対して多重比較を行った結果である。プロと一般学生で音素長に有意な差がある音素をピンク（長い）と水色（短い）で示している。

子音については、プロでは一般学生より語頭の子音（Table 1 の★と▲）の全てが有意に長かった。語頭以外の子音では、破裂音の破裂部で有意に長い傾向が見られた。無声破裂音では気音部、有声破裂音では母音への過渡部が長かった。しかし、破裂音の直前の閉鎖区間はプロの方が有意に短い傾向も見られた。これらのことから、プロは破裂音では、明瞭度を上げるために音素長を長くする一方、音声が途切れないようにに閉

Table. 1 多重比較の結果

| 分類 | 歌詞 | 子音 | 母音 | 子音+母音 | 分類 | 歌詞 | 子音 | 母音 | 子音+母音 | 分類 | 歌詞 | 子音 | 母音 | 子音+母音 | 分類 | 歌詞 | 子音 | 母音 | 子音+母音 |
|----|----|--------|----|--------|----|----|--------|----|--------|----|----|--------|----|--------|----|----|--------|----|--------|
| ★ | は | h | a | ha | ★ | の | n | o | no | ★ | か | <cl> k | a | <cl>ka | ★ | な | n | a | na |
| | る | r | u | ru | | ぼ | <cl> b | o | <cl>bo | | い | | i | i | | が | <cl> g | a | <cl>ga |
| | の | n | o | no | | り | r | i | ri | | の | n | o | no | | め | m | e | me |
| | う | | u | u | | く | <cl> k | u | <cl>ku | | を | w | o | wo | | な | n | a | na |
| | ら | r | a | ra | | だ | <cl> d | a | <cl>da | | ず | <cl> z | u | <cl>zu | | に | n | i | ni |
| | ら | r | a | ra | | り | r | i | ri | | く | <cl> k | u | <cl>ku | | に | n | i | ni |
| | の | n | o | no | ▲ | ふ | f | u | fu | ▲ | は | h | a | ha | ▲ | た | <cl> t | a | <cl>ta |
| ▲ | す | s | u | su | | な | n | a | na | | な | n | a | na | | と | <cl> t | o | <cl>to |
| | み | m | i | mi | | び | <cl> b | i | <cl>bi | | と | <cl> t | o | <cl>to | | べ | <cl> b | e | <cl>be |
| | だ | <cl> d | a | <cl>da | | と | <cl> t | o | <cl>to | | ち | <cl> c | i | <cl>ci | | き | <cl> k | i | <cl>ki |
| | が | <cl> g | a | <cl>ga | | が | <cl> g | a | <cl>ga | | る | r | u | ru | | | | | |
| | わ | w | a | wa | | | | | | | | | | | | | | | |

★: 小学節の頭かつ語頭 <cl>: 破裂音の閉鎖区間
 ▲: 語頭
 色分け: オレンジ: プロの方が発声タイミングが早い, 黄緑: プロの方が発声タイミングが遅い, 紫: プロの方が音素長が長い, 青: プロの方が音素長が短い

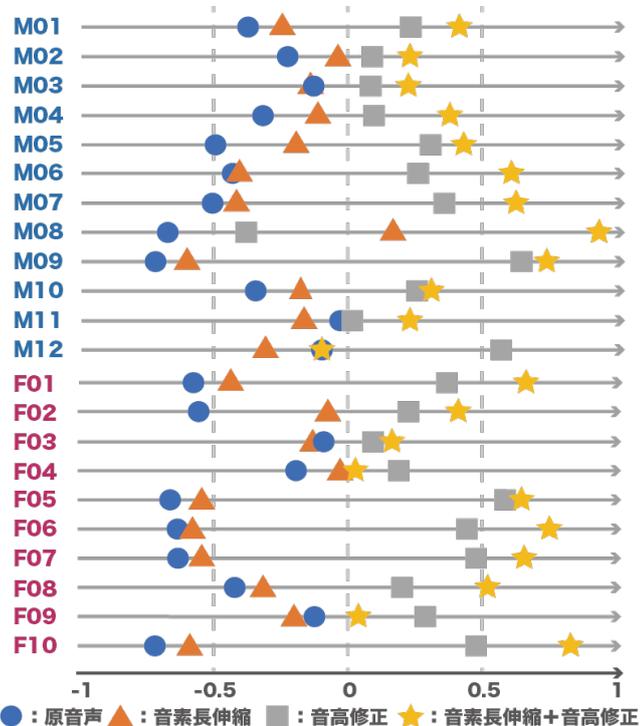


Fig. 5 学生の歌声の印象評価の結果

鎖区間は短くしていると考えられる。鼻音の/n/は、語頭ではプロで有意に長かった。しかし、語中では同じ鼻音の/m/と同様に有意に短い有意差が見られなかった。これは、鼻音はパワーが小さいので、語頭では音素長を大きくして明瞭性を増すが、語中では特に音素長を操作していないためと思われる。

母音については、語頭でも語中でもプロの方が長い場合も短い場合もあり、全体として一貫した傾向はみられなかった。

子音+母音については、全ての語頭の部分で有意に長い傾向が見られた。これはアンケート結果とも一致していた (Fig.4)。このことから、プロは語頭では母音

Table 2 M08, M12, F09 の自然性の評価

| 実験参加者 | 原音声 | 音素長伸縮 | 音高修正 | 音素長伸縮+音高修正 |
|-------|-------|-------|------|------------|
| M08 | -0.35 | 0 | 0.05 | 0.15 |
| M12 | 1.15 | -0.55 | 1.25 | -0.1 |
| F09 | 1.7 | 0.4 | 1.45 | 0.5 |

←不自然 -2 -1 0 1 2 自然→

ではなく子音を長くして歌唱しているといえる。

3.3. 発声タイミングの比較

Table 1 は発声タイミングの多重比較の結果も示している。プロと一般学生で発声タイミングに差がある音素の単純主効果を分析し、有意水準 5%以下の部分をオレンジ (早い) と黄緑 (遅い) で示している。

プロで有意に発声タイミングが早かったのは語頭の中でも小楽節の頭 (★) であった。この部分は直前に休符があるため発声タイミングを早めやすく、子音の音素長も長くなっていると考えられる。

3.4. 学生の音声の評価実験の結果

Fig.5 は学生の歌声に対する印象評価実験の結果である。基本的に、原音声、音素長伸縮、音高修正、音素長伸縮+音高修正の順で評価が高くなる傾向が見られた。原音声より音素長伸縮、音高修正、音素長伸縮+音高修正の評価が有意に高い学生はそれぞれ 4 名 (M05, M08, M10, F02), 20 名 (M11, F03 以外), 19 名 (M12, F04, F09 以外) であった。また、音高修正より音素長伸縮+音高修正の評価が有意に高い学生は 8 名 (M04, M06, M07, M08, F01, F06, F08, F10) であった。これらの結果は、楽譜通りの音高で歌うことが歌のうまさの評価において重要であること [1-3] を裏付けるとともに、少なくとも多くの学生では音高が不正確であり、音高を楽譜通りに修正することで、音素長伸縮が歌のうまさをさらに向上させる効果があることを示す。なお、カラオケの自動採点機能 [10] でも

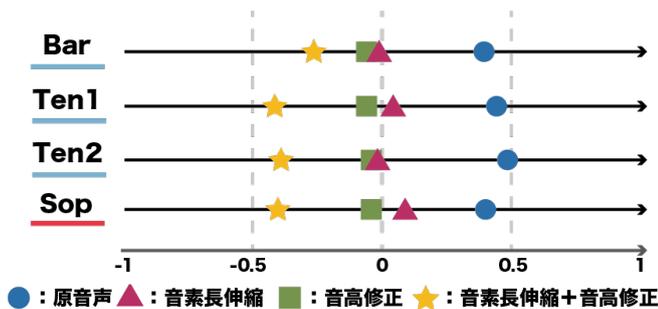


Fig. 6 プロの歌声の印象評価の結果

Table. 3 プロの音声の自然性の印象評価の結果

| 実験参加者 | 原音声 | 音素長伸縮 | 音高修正 | 音素長伸縮+音高修正 |
|-------|------|-------|------|------------|
| Bar | 1.65 | 1.3 | 1.75 | 0.9 |
| Sop | 2 | 0.1 | 1.85 | 0.3 |
| Ten1 | 1.95 | -0.15 | 1.95 | -0.2 |
| Ten2 | 1.9 | 0.25 | 1.7 | -0.1 |

不自然 ← -2 -1 0 1 2 → 自然

Table. 4 楽譜の音高と実際の音高の音符間の平均との差(cent)

| 歌詞 | | の | ぼ | り | く | だ | り | の | ふ | な | び | と | が | | |
|----|-------|-------|-------|------|------|-------|-------|------|-------|-------|------|-------|-------|-------|-------|
| 男性 | 学生の平均 | -42.9 | 56.7 | 1.3 | 7.1 | -50.7 | -54.2 | 76.0 | -24.7 | -51.8 | -0.7 | -32.9 | 0.0 | -51.5 | -18.1 |
| | Bar | 17.4 | 41.6 | 13.6 | -1.1 | -18.5 | -0.3 | -7.3 | -1.7 | -12.8 | 10.0 | -10.1 | -15.7 | -5.8 | -11.6 |
| | Ten1 | -23.2 | -41.2 | 4.1 | -0.9 | -22.1 | -59.1 | 16.6 | 20.9 | -23.1 | 5.3 | -2.7 | 8.2 | -10.4 | -6.9 |
| | Ten2 | -10.9 | -3.1 | 25.9 | 50.6 | -2.0 | -10.3 | 15.8 | -28.0 | -4.7 | 16.3 | -4.6 | 30.5 | 37.8 | 13.0 |
| 女性 | 学生の平均 | -8.3 | 40.8 | -9.0 | 9.9 | -26.3 | -36.4 | 43.8 | -16.4 | -23.6 | -7.1 | -8.7 | -1.8 | -26.9 | -23.9 |
| | Sop | 7.1 | 25.1 | 46.4 | 53.8 | -19.0 | -39.8 | 22.0 | -0.3 | 27.6 | 12.7 | 4.1 | 29.3 | 18.9 | 31.6 |

：楽譜の音程と50cent以上の差があった部分

楽譜通りの音高で歌唱しているかを評価している。

また、全体として原音声、音素長伸縮、音高修正、音素長伸縮+音高修正の順で評価が高くなったが、M03, M08, M11, M12, F03, F04, F09の7名ではこの順ではなかった。しかし、評価値間の有意差を考慮すると、この順序から逸脱していたのはM08, M12, F09の3名であった。M08では音素長伸縮と音高修正、M12とF09では原音声と音素長修正、音高修正と音素長修正+音高修正の間に順序が入れ替わる方向に有意差があった。M08は原音声の自然性が低かった(Table 2)。これは、楽譜の示す音高と発声タイミングから乖離していたためと思われる。特に発声タイミングの乖離が大きかったため、音素長伸縮により発声タイミングが修正されて評価が向上したと考えられる。M12とF09では音素長伸縮を行うと自然性が原音声より低下した(Table 2)。これは、プロと比べて一部の音素が極端に短く(約1/10~1/7)、音素長伸縮により不自然性が増したためと考えられる。

3.5. プロの音声の評価実験の結果

Fig.6はプロの歌声に対する印象評価実験の結果を示しており、Table 4は楽譜上の音高と実際の音高の音符間の平均との差を示している。なお、±50 cent以内を「正しい音高」と定義している[11]。

すべてのプロにおいて、原音声、音素長伸縮、音高修正、音素長伸縮+音高修正の順で、歌のうまさの評価が低下した。これは、オペラ歌唱の声であっても、音高が不正確な場合や不適切な音素長の場合には歌のうまさの印象を低下させる結果となった。

また、Bar, Ten1, Ten2の3名では、音素長伸縮、音高修正、音素長伸縮+音高修正のいずれにおいても、原音声に比べて評価が有意に低下した特に男性学生では、音高が不正確な箇所が多く、音高修正によって音高が不正確になり、歌のうまさの印象を低下させたと考えられる。これらの結果は、学生の音声の評価実験の結果を裏付けた。

しかし、Ten1の音素長伸縮と音素長伸縮+音高修正の自然性が低い結果となっていた(Table 3)。Ten1の原音声は録音環境における残響が大きかったため、音素長修正によって歌声が不自然になったと考えられる。Ten1の結果はBarやTen2と同様の傾向を示したが、残響による不自然さがこの結果影響している可能性がある。このため、今後は残響の除去や再録音も検討する必要がある。

4. まとめ

本研究では、プロのオペラ歌手4名と一般学生22名の音素長や発声タイミングの比較を行った。音素長については、プロは語頭の子音を学生よりも有意に長く歌唱していた。また、プロは破裂音の閉鎖区間が有意に短いが無声破裂音では気音部、有声破裂音では母音への過渡部を長く発声していた。これは、歌声が途切れないように閉鎖区間を短くしつつ、破裂音を印象付けるようにそれ以外の部分を長く発声していると考えられる。発声タイミングについては小楽節の頭の部分で早かった。これにより、小楽節の頭を明瞭に発声し、印象付けていると考えられる。

そして、これらの違いが歌のうまさの印象に与える

影響を調べるため、歌声に操作を加えて印象を評価した。学生 22 名の歌声の音素長や発声タイミングをプロに近づけた場合は、22 名中 19 名で音素長伸縮+音高修正の歌声が原音声より有意に評価が高くなった。しかし、ある程度音高が正確でなければ、音素長伸縮により評価が向上しないことも示唆された。一方プロの歌声の音素長と発声タイミング、音高を学生の平均に近づけた場合は、原音声、音素長伸縮、音高修正、音素長伸縮+音高修正の順に、評価が低下する傾向が見られた。これらの結果から、プロのオペラ歌手のように豊かな響きのある歌声であっても、音高や音素長が適切でないと歌のうまさの評価は低下すると考えられる。

しかし、母音と子音の適切な音素長の比率については、検討した母音と子音の組み合わせが少なかったため、何らかの法則があるかどうか明らかにすることはできなかった。また、歌手のフォルマント[3]の有無が適切な音素長の比率にどのような影響を与えるについても明らかにすることはできなかった。今後はこれらの点について追及する必要がある。

謝 辞

本研究では JSPS 科研費 23K11172 の支援を受けた。また実験に参加していただいた全ての方々から感謝する。

文 献

- [1] 齋藤ら, “歌声らしさの知覚モデルに基づいた歌声特有の音響特徴量の分析,” 音響学会誌, 64 (5), pp.267-277, 2008.
- [2] 齋藤, 後藤, “歌唱指導による歌声中の音響特徴の変化:歌唱力評価に寄与する音響特徴の検討,” 音講論 (春), pp.583-586, 2009.
- [3] Sundberg, “THE SCIENCE OF THE SINGING VOICE,” Northern Illinois University Press, 1987.
- [4] 中野ら, “楽譜情報を用いない歌唱力自動評価手法,” 情報処理学会論文誌, 48, pp.227-236, 2004.
- [5] Smith & Scott, “Increasing the intelligibility of sung vowels,” J Acoust Soc Am, 67(5), pp.1795-1797, 1980.
- [6] McCrea & Morris, “Comparisons of Voice Onset Time for Trained Male Singers and Male Nonsingers During Speaking and Singing,” Journal of Voice, Vol. 19, No. 3, pp.420-430, 2005
- [7] Lisker & Abramson, “A Cross-Language Study of Voicing in Initial Stops: Acoustical Measurements,” Word, 1964,20, pp.384-422, 2015.
- [8] 的場ら, “歌唱のグルーブ感の構成要因の分析,” 情報処理学会研究報告, Vol.2014-MUS-102, No.12, pp.1-5,2014.
- [9] Boersma & Heuven, “Praat, a system for doing phonetics by computer,” Glot International, 5 (9/10), pp.341-345, 2001.
- [10] ヤマハ株式会社, 株式会社第一興商, 安間幸成, 橘聡, “歌唱音声評価装置, カラオケ採点装置及びそのプログラム,” 特許広報 P4091892, 2008-5-28
- [11] Demorest & Pfordresher, “Singing Accuracy