

歌唱のうまさに対する印象評価構造

高田兼輔^{†1} 金礪愛^{†1} 菊池英明^{†1}

概要: 近年、カラオケの普及により他人に歌唱を聴かせる機会が増えている。それに伴い、歌唱の評価に対するニーズが高まっているが、人の感性を考慮した歌唱の評価構造は十分に明らかになっていない。本研究では、人が歌唱のうまさの評価する際の印象評価構造を共分散構造分析を用いて明らかにする。歌唱収録、印象評定実験を経て印象評定結果を分析したところ、評価者の評価傾向が大きく2つのグループに分かれた。それぞれのグループの共分散構造分析の結果としてモデルの適合度はそれぞれ 0.445, 0.444 となった。モデルの構造について考察したところ、歌唱のうまさは「技術力」を重視して評価する傾向の人と「影響力」を重視して評価する傾向の人に大別できることが分かった。

キーワード: 歌唱, 印象, 共分散構造分析

Impression Evaluation Model for Goodness of Singings

KENSUKE TAKADA^{†1} AI KANATO^{†1}
HIDEAKI KIKUCHI^{†1}

Abstract: Recently, spread of karaoke has increased the chances for others to listen to their singing. Although the need for evaluation of singing is increasing, the evaluation model of singing taking into consideration human sensibility is not sufficiently clarified. This research aimed to clarify impression evaluation model for goodness of singings by using Covariance Structure Analysis (SEM). After dividing evaluators into 2 groups, we analyzed by using SEM. In result, GFI of group1 was 0.445, GFI of group2 was 0.444.

Keywords: Singing, Impression, Covariance Structure Analysis

1. はじめに

近年、通信カラオケがエンターテインメントの一つとして普及している。比較的低価格で長時間楽しめることから、様々な場面で老若男女問わず利用されている。このような背景により、人に歌唱を聴かせる機会が大幅に増えたことに伴い、「うまく」歌いたいと考える歌唱者にとって歌唱評価のニーズが高まっている。

他者からうまいと評価された自身の歌唱や、自身がうまいと評価した他者の歌唱でもカラオケの採点機能では良い評価を得られないことがある。例えば、プロがカラオケで自分の曲を歌っても満点が取れないというのは有名な話である。しかし、例えカラオケの採点での評価が低くとも、プロの歌唱を聴いて下手だと思える人は少ないであろう。このことから、現行のカラオケ採点システムでは人の感性を十分に反映できていないと言える。

近年歌唱に関しての研究が増えており、中野ら[1]は歌唱の評価には相対音高やビブラートに関わることを明らかにした。竹内ら[2]は人の主観評価に近いカラオケ採点法として、歌唱者の歌声から抽出したピッチ周波数と平均律音階

の距離を利用して歌唱評価を行うことを提案したが、人が歌唱に持つ印象との関係性までは明らかにしていない。金礪ら[3]は、ポピュラー音楽における歌声を聴いた際に感じる多様な印象を音響信号から自動推定することを試みた。この研究では「うまさ」の観点を含めているものの、その評価構造までは明らかにしていない。

そこで本研究では、人が歌唱のうまさの評価する際の印象評価構造を明らかにすることを目指す。歌唱のうまさに対する印象評価構造を明らかにするにあたり、今後の研究の足がかりとして、今回はポピュラー音楽に焦点を当て、ワンフレーズの短い単位での歌唱で形成される印象を扱う。

本研究は、まず歌唱収録を行い、収録した歌唱音声を用いて印象評定実験を行う。共分散構造分析(SEM)を用いて印象評定結果を分析することによって、歌唱のうまさに対する印象評価構造を明らかにする。

2. 歌唱音声収録

本章では印象評定実験に用いる歌唱音声収録について説明する。2.1に収録の条件を、2.2にその条件を満たす収録

^{†1} 早稲田大学
Waseda University



図 1 歌唱曲の楽譜
Figure 1 Score for singers

方法を述べる。

2.1 収録の条件

印象評定における歌唱音声以外の要素の影響の排除、歌唱者属性の均衡などを考慮して以下の条件で収録する。

- メロディ・歌詞が統一されていること
- 評価者にとって、未知のメロディ・歌詞であること
- 誰にでも歌唱可能な曲であること
- 男女複数人ずつのデータであること

2.2 収録の詳細

(1) 歌唱曲:

2.1 の条件を満たす歌唱曲として、[2]で用いた楽曲を使用する。(BPM:144, 9秒程度のワンフレーズの楽曲)(図1参照)

(2) 歌唱者:

男女各 20 名

(3) 歌唱者の条件:

アマチュアの歌唱者に限る

(4) 場所:

早稲田所沢キャンパス フロンティアリサーチセンター
菊池英明研究室防音室

(5) 機材:

コンデンサーマイク (RODE NT2-A)
オーディオインターフェース (Roland Rubix 330)

(6) 収録環境:

歌唱者には防音室内でヘッドフォンを装着してもらい、歌唱曲の伴奏のみを聴いてもらいながら歌唱を行ってもらい。歌唱中は自身の歌唱もマイクを通してヘッドフォンから聴こえるようにし、歌唱者が歌いやすいように配慮した。また、本研究では歌唱のうまさについての分析を行うため、歌唱者には歌唱者自身が一番うまいと思える歌い方で歌唱を行ってもらった。

3. 印象評定実験

歌唱のうまさに対する印象の評価構造を明確にすることを本研究の目的とするため、歌唱に対して多様な評価語を用いて印象評定実験を行う。以下には、印象評価語の収集方法、了解性調査と同義性調査による評価語の絞り込みについて述べたうえで、印象評定実験の方法を述べる。

3.1 印象評価語の収集

評価語を選定するために、元となる語句を集める必要がある。そこで、評価語の一部として適切なプロセスを経て収集した 223 語の評価語[2]を使用する。それらは以下の条件に該当する評価語である。

- 研究上重要であると考えられる語
- 専門的に使用されている語
- 日常的に使用されている語

ただし、[2]ではこれらの評価語のうち以下の評価語を除いている。

- 歌唱技術を直接的に評価している語
- 1語の形容詞でない語
- 聴取者に生じる反応を表した語

歌唱のうまさを表す語句には、歌唱技術を表す語句や聴取者に生じる反応を表す語句も含まれるため、本研究ではそれらの語句を以下の方法で再収集する。

- アマチュアがうまさの判断基準としている語句
- プロがアマチュアの歌唱を評価する際に用いる語句

以上の過程を経て 27 個の評価語を加え、合計 250 語の評価語を収集した。

3.2 了解性調査

歌唱のうまさを表現する際に用いる評価語は一般人でも理解できる言葉を用いる必要がある。そのため、上記の評価語収集にて集めた 250 語に対して、各評価語の了解性を測るためのアンケート調査を行った。調査の詳細は以下のとおりである。

(1) 回答者:

33 名 (20 代男性 18 名, 20 代女性 15 名)

(2) 対象:

250 語の印象評価語

(3) 教示文:

あなたが歌のうまさを表現する際に「○○な歌だからうまい」と言ったときに、そのような歌が想像できる語句に対して「使う」を、想像できない語句に対して「使わない」を選択してください

(4) 回答方法:

「使う」「使わない」の 2 択

アンケートへの回答を元に各評価語の了解性の度合いを求める。アンケートにてその評価語を「使う」と回答した人数の割合を求めることで度合いを算出した。以下の表 1 に了解性調査の一例を示す。

本研究では了解性の度合いが 0.7 以上のものを採用し、その結果 250 語の評価語の中から 93 語を選定した。

表 1 評価語の了解性度合いの例

Table 1 Examples of intelligibility

心のこもった	1.00	すごい	0.94	突き抜けるような	0.88	悲鳴のような	0.21
感情のこもった	0.97	かっこいい	0.94	素敵な	0.88	気の抜けた	0.21
綺麗な	0.94	力強い	0.91	声量のある	0.88	ロリっぽい	0.21
優しい	0.94	独特な	0.91	真っすぐな	0.88	雄々しい	0.18
明るい	0.94	心地よい	0.91	耳に残る	0.88	未熟な	0.18
透き通った	0.94	惹き付けられる	0.91	元気な	0.88	上振った	0.18
聴きやすい	0.94	柔らかい	0.91	気持ちのよい	0.88	熱い	0.15
単調な	0.94	聞き惚れる	0.91	感情表現豊かな	0.88	舌つたらずな	0.15
素晴らしい	0.94	美しい	0.88	抑揚のある	0.88	毒が抜けるような	0.12
表現力のある	0.94	迫力のある	0.88	プロのような	0.88	締めぎみな	0.09

表 2 評価語の組み合わせと類似性の例

Table 2 Examples of similarity

澄み切った	透き通った	1.00	素晴らしい	魅力的な	0.45
エネルギッシュな	パワフルな	1.00	いきいきとした	元気な	0.45
優しい	温かい	0.91	いきいきとした	テンションの高い	0.45
力強い	パワフルな	0.91	迫力のある	存在感のある	0.45
単調な	棒読みのような	0.91	元気な	力強い	0.36
迫力のある	圧倒させるような	0.91	勢いがある	声量のある	0.36
爽やかな	澄み切った	0.91	迫力のある	テンションの高い	0.36
個性的な	個性的な	0.91	勢いがある	ロックな	0.36
心のこもった	感情のこもった	0.91	鳥肌の立つ	存在感のある	0.27
美しい	綺麗な	0.91	盛り上がる	ロックな	0.27
感情表現豊かな	表現力のある	0.91	勢いがある	メリハリがある	0.27
ぬくもりのある	温かい	0.82	躍動感のある	メリハリのある	0.18
優しい	柔らかい	0.82	声量のある	安定している	0.18
安定している	ブレていない	0.82	澄み切った	真っすぐな	0.09
力強い	声量のある	0.82	爽やかな	ロックな	0.00

3.3 同義性調査

印象評定実験を行う上で、似たような意味の評価語は必要ないため、評価語を選定する必要がある。そこで本調査では了解性調査によって選定された 93 語を対象に同義性調査を行い、同義性の高い評価語の統合を行った。調査は以下の 2 段階を経た。

(i) 類似性調査アンケート

まず、93 語全ての評価語の組み合わせについて、「似たような意味」と回答した人数の割合を類似性とする。類似性調査の結果の一例を表 2 に示す。

本調査では類似性の値が 0.6 以上であった評価語の組み合わせ (計 89 対) を次の同義性調査アンケートの対象とした。

(ii) 同義性調査アンケート

89 対について、2つの評価語が表す歌声がどの程度似ているのかを 7 段階で評価してもらった。評価が高い組み合わせについて評価語の統合を行った。統合した結果を表 3 に、了解性調査・同義性調査を経て選定した印象評定実験に用いる評価語を表 4 に示す。なお、表では右側の語句を左側に統一したことを示しており、評価語の意味が近いものごとに行を区切っている。

なお、印象評定実験の前に以下の評価語を除いた。

表 3 統合を行った評価語

Table 3 Merged evaluation words

統合後の評価語	統合した評価語
聴き惚れる	惹きつけられる, つい聴いてしまう
気持ちのよい	心地よい
心のこもった	感情のこもった
感情表現豊かな	表現力のある
音程が取れている	ブレていない, 正確な
個性的な	個性のある, 独特な, 癖のある
力強い	パワフルな
圧倒させるような	迫力のある, インパクトのある, 存在感のある
元気な	盛り上がる, エネルギッシュな
勢いがある	テンションの高い
温かい	ぬくもりのある
柔らかい	優しい
穏やかな	落ち着いたきのある, ゆったりとした

表 4 印象評定実験に採用した評価語

Table 4 Evaluation words adopted for experiment

メリハリがある	力強い	<u>元気な</u>	声量のある
<u>真っ直ぐな</u>	<u>爽やかな</u>	<u>勢いがある</u>	<u>躍動感のある</u>
<u>いきいきとした</u>	<u>澄み切った</u>	情熱的な	聴き惚れる
<u>柔らかい</u>	気持ちのよい	ふわっとした	個性的な
繊細な	<u>丁寧な</u>	<u>楽しそうな</u>	<u>圧倒させるような</u>
温かい	<u>心のこもった</u>	<u>抑揚のある</u>	穏やかな
<u>和む</u>	<u>堂々とした</u>	音程が取れている	<u>感情表現豊かな</u>
<u>味がある</u>	ロックな	<u>存在感のある</u>	<u>響きのある</u>
鳥肌の立つ	ポップな	<u>リズム感がある</u>	<u>余裕のある</u>
<u>深みのある</u>	可愛い	自然な	<u>かっこいい</u>
<u>甘い</u>	<u>キレイがいい</u>	プロのような	<u>ビブラートがいい</u>

(※下線の語は 4.3.4 の分析に用いた語句を示す)

- メタ的な評価語
「プロのような」
- うまさ以上に歌唱のジャンルに関わりの強い評価語
「ロックな」「ポップな」

3.4 印象評定実験

収録した歌唱データに対して、前述の評価語によりそれぞれ 7 段階で評価してもらった。詳細は以下の通りである。

- (1) **評価者:**
24 名 (20 代男性 11 名, 20 代女性 13 名)
- (2) **刺激:**
40 種類 (男女 20 名ずつ) の歌唱データ
- (3) **回答方法:**
44 語の評価語
0 (印象なし) ~6 (印象大) の 7 段階

評価者ごとに点数を標準化し、その点数の合計を算出するが、最初に評価者間の相関を分析した結果、他の評価者と負の相関を持った評価者（1名）がいたため、そのデータを除き、残りのデータで分析を行った。

4. 共分散構造分析(SEM)を用いた分析

4.1 共分散構造分析について

共分散構造分析は構造方程式モデリング、通称 SEM と呼ばれ、統計資料の背後にある抽象的な事象を議論できることが特徴である。また、涌井ら[6]は共分散構造分析について以下のように述べている。

「共分散構造分析とは直接観測することができない潜在変数を導入し、その潜在変数と観測変数との因果関係を同定することにより社会現象や自然現象を理解するための統計的アプローチである。」

本研究では、「歌唱のうまさ」という潜在変数を、印象評定実験で得られた評価語の得点という観測変数を用いて関係性を明らかにする。

4.2 モデル構築

印象評定実験の結果を標準化し、得られた点数を本実験における印象評価得点とした。その得点をもとに分析を行う。

「うまさ」と印象評価語の間である第二層の潜在変数を同義性調査の結果を参考に「技術力」「表現力」「影響力」「迫力」としてパス図を出力した。しかし、モデルの適合度が 0.324 であり、十分な説明力とは言えないため、追加検討を行う。

4.3 追加検討

4.3.1 評価者のクラスタリング

評価語の解釈が多様になったため評価者による回答のばらつきが大きくなってしまっている可能性がある。そこで、クラスタ分析を行うことで、評価者の評価傾向を分ける。Ward法を用いて階層的クラスタリングを行った結果を図3に示す。この結果をもとに、評価者を二つのグループに分けて分析を行う。

4.3.2 グループごとの分析

被験者番号 7,8,11,15,23 を第1グループ、それ以外を第2グループとして共分散構造分析を行う。

実験結果を分析したところ、第1グループの評価者によるSEMモデルの適合度は0.362、第2グループの評価者によるSEMモデルの適合度は0.385となった。

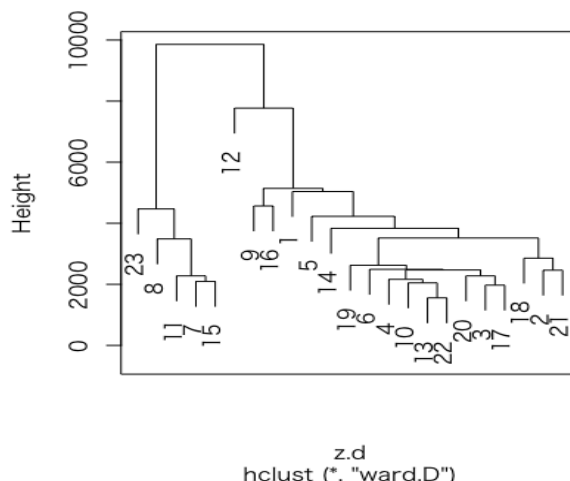


図3 評価者を対象としたクラスタ分析の結果

Figure 3 Result of cluster analysis focused on evaluator

どちらのグループも十分な説明力とは言えないため、更なる追加検討として評価語ごとの級内相関係数を算出し、歌唱のうまさに関わっていないと考えられる評価語を除外する。

4.3.3 評定者間の一致率

級内相関係数 (Interclass Correlation Coefficient, ICC) を求めることにより、評定者の回答間における回答のばらつきを求めた結果、評定項目によって 0.136-0.281 の値をとることがわかった。Landis ら[7]によると、0.00-0.20 は“poor”、0.21-0.40 は“fair”であり、“poor”と判断された信頼性の低い評価語を除くことで、モデルの適合度向上を目指す。

以上の手順により、「メリハリがある」「力強い」「声量のある」「情熱的な」「聴き惚れる」「気持ちの良い」「繊細な」「温かい」「穏やかな」「音程が取れている」「味がある」「鳥肌の立つ」「可愛い」「自然な」「個性的な」を除いた。

4.3.4 再分析

残った評価語（表4の下線部を引いているもの）26語を用いてグループごとに再分析を行った。

(1) 第1グループ

クラスタリング後の分析と同じ条件で再分析したところ、適合度は 0.446 まで向上した。モデルのパス図は図4の通りである。パス係数は「影響力」が最も高く 0.43 であった。「技術力」は 0.33、「表現力」は 0.32、「迫力」は 0.35 となっており、表現力や迫力も技術力と同等に評価に影響を与えていることが分かる。

(2) 第2グループ

第2グループも同じように再分析を行った。その結果、適合度は 0.444 まで向上した。モデルのパス図は図5の通り

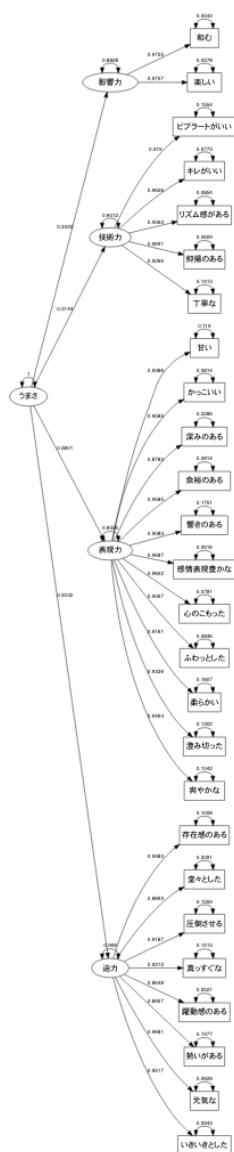


図 3 第1グループのSEMモデル
 Figure 3 Model for group1

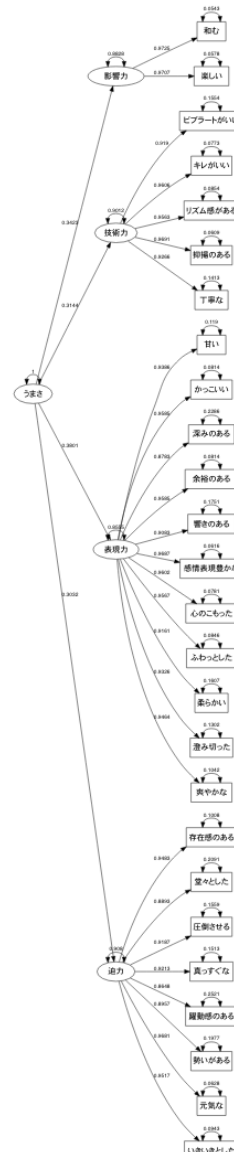


図 4 第2グループのSEMモデル
 Figure 4 Model for group2

である。パス係数は「技術力」が最も高く、1.23であった。「迫り」は0.68, 「表現力」は0.64, 「影響力」は0.49であり、第2グループではこれまで通り技術力が評価に影響していることが分かる。

この分析結果より、歌唱のうまさは「技術力」が重要であると評価する傾向の人と「影響力」が重要であると評価する傾向の人に大別することができる。

5. 課題と展望

評価者のクラスタリングや評価語の削減などの追加検討を行ったが、モデルの適合度を十分に向上させることはできなかった。今後の課題として、印象評定実験での評価者を増やすために、印象評定実験をもう一度行う必要があ

ると考える。

今後の展望として、モデルの適合度を上げた後、評価語と音響特徴量の関係もモデル化し、人の感性を考慮した自動歌唱評価システムの構築を目指す。

参考文献

- [1] 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. Vol. 48, No. 1, pp. 227-236, 2007.
- [2] 竹内ら. 人の主観に近いカラオケ採点法. 電気学会論文誌 C. Vol.130, No.6, pp.1042-1053
- [3] 金礪ら. 歌声の印象評価尺度の構築に基づく多様な印象の自動推定手法. 情報処理学会論文誌. Vol. 57, No. 5, pp. 1375-1388, 2016.
- [4] 竹内ら. 人の主観に近いカラオケ採点法. 電気学会論文誌 C. Vol.130, No.6, pp.1042-1053.
- [5] 涌井ら. 図解でわかる 共分散構造分析. 日本実業出版社. 2003.
- [6] Landis, et al. Biometrics. 33, pp.159-174, 1977.