

## 特定のファン層が好む歌詞の特徴分析

大谷紀子研究室

1372133 渡部誠也

### 1. 研究の背景

音楽の好みは個人によって異なり、同じアーティストのファンでも好みの曲はそれぞれである。同じセットリストでコンサートを行っても、地域や客層によって盛り上がる曲が異なるという感想も聞かれる。コンサートで観客をより満足させたり、特定の人たちの心に響くような曲を提供したりするには、聴者の属性と音楽に関する嗜好の関係を明確にする必要がある。現在、大谷研究室では個人の感性に即した楽曲の自動生成システムを開発しているが、本システムで使用されている感性モデル生成技術により、特定のファンに好まれる歌のメロディと和音進行に関するルールの獲得が試みられている。しかし、歌の印象はメロディだけでなく、歌詞にも左右されると考えられるため、歌詞の分析も必要である。

本研究では、年齢や出身地などの属性が共通している聴者のグループごとに、好みの歌の歌詞に関する特徴を明確にすることを目的とする。フォークを歌う男女のデュオ「ワライナキ」を対象として、コンサートを訪れた観客に関するデータと歌詞を分析する。

### 2. 研究の内容

聴者の属性データは、ワライナキの大阪、奈良、東京、名古屋、京都におけるコンサートツアーの観客 209 名を対象として、アンケートにより取得した。アンケート項目は、お気に入りの曲、お気に入りでない曲、コンサートで良かったと思う曲、年齢層、性別、出身地、居住地、好きな色、好きな味付け、好きな料理、コンサートを見た回数、ワライナキを知ったきっかけ、チケットの購入場所、好きなお酒、CD にして欲しい曲、作って欲しいグッズである。お気に入りの曲、お気に入りでない曲のアンケート項目では、ベストアルバム・ニューアルバムに収録されている 25 曲は選択肢から選び、他の曲は自由記述で回答する。

歌詞の分析に先立ち、形態素解析により、歌詞に含まれる単語の品詞と語幹を抽出し、名詞、動詞、形容詞の語幹を有効語とする。また、特定の属性のファン層が好む曲が複数あった場合、該当する曲をすべてまとめてグループとする。お気に入り曲と答えた人数の分布で $+0.5\sigma$ 以上の範囲に入る 17 曲をワライナキの「人気曲」とする。

形態素解析結果と作成したグループを元に、下記の 3 つを調べることで、共通するルールを獲得する。

- ① 重要度に基づく歌詞の特徴語
- ② 歌詞に含まれる単語の品詞の頻度
- ③ 歌詞の特徴

本研究では、グループでより多く使われる有効語をグループの特徴語とみなす。グループに属する曲の有効語の総数を  $A$ 、有効語  $T$  の出現個数を  $tfreq(T)$ 、グループに該当する曲数を  $G$ 、グループに属する曲の中で有効語  $T$  を含む曲数を  $dfreq(T)$  としたとき、有効語  $T$  のグループにおける重要度  $w(T)$  は式

(1)で求められる.

$$w(T) = \frac{tfreq(T)}{A} \frac{dfreq(T)}{G} \quad (1)$$

算出した重要度  $w(T)$  の値が高いものをグループにおける特徴語とする. グループの特徴語と人気曲の特徴語を比較し, 算出した重要度  $w(T)$  が人気曲よりグループの方が高い特徴語をグループの特徴語とする. 歌詞に含まれる名詞, 動詞, 助動詞, 形容詞, 助詞の割合のうち, グループに属する曲すべてに共通し, いずれの人気曲とも一致しない値をグループの歌詞に含まれる単語の品詞の頻度における特徴とする. サビを3回以上リピートする, 4文字熟語が多く使われているなど, 人気曲に該当するものがなく, グループに該当する曲にのみ共通する歌詞からとれる特徴をグループの歌詞の特徴とする.

### 3. 結果と考察

上記の分析の結果, 以下のルールが得られた.

- 年齢層が低いファンは歌詞にスキャットが入っている曲, または名詞が約 42%, 動詞が 20%, 助動詞が約 10%, 助詞が 26%の曲を好む.
- 好きな色が紫のファンは名詞が 29%以下, 助詞が 36 以上の曲を好む.
- 血液型が AB 型のファンは特徴語「時」「場所」が含まれる曲を好まない.
- 名詞が 34%以上, 動詞が 18%以上, 助詞が 30%以上の曲は CD 音源, ライブ共に好まれる.

作成された 13 グループのうち, ルールを獲得できたのは上記 4 グループだけだった. お気に入りの曲の数でグループを作成していることから, 人気上位の回答数の多い曲でなければ回答数に大きな差が出ず, 多くのグループを作成できなかった. また, アンケート用紙の「お気に入りの曲」の回答欄では, 曲によって選択肢と自由記述で回答方法を分けたことで, 選択肢の存在する人気の曲は回答数が多く, 自由記述に曲名を書く必要がある曲は回答数が極端に少ないという結果になった. アンケートに回答するときに, すべての曲で回答する労力を同じにする改善が必要だと考える.

歌詞の特徴語によってルールを獲得できたのは 1 つだけだった. 歌詞の特徴語では, 特定のファン層が好む曲のグループの特徴語として得られたものはすべて人気曲の特徴語に該当し, グループ固有の特徴語を見つけることはできなかった. 歌詞の特徴語をルールとして獲得できたグループは, ルールを獲得できたグループの中で, 唯一ファンの好む曲のグループではなく, ファンの好まない曲のグループであったことが, 歌詞の特徴語をルールとして得ることができた要因と考えられる. 4 つのルールのうち 3 つのルールでは品詞の割合がルールとして獲得できた. 獲得できた 3 つのルールに共通することは獲得できた 3 つすべてのルールに名詞と助詞が含まれていることである. 今回の分析では, 助詞を含む場合と助詞を含まない場合の 2 パターンで品詞の割合を抽出した. 結果, 助詞を含まない場合ではルールを獲得できず, 助詞を含んだ場合に 3 つのルールを獲得することができた. 以上のことから, 単語の品詞の頻度におけるルール分析において, 助詞が最も重要な品詞であったと考えられる.

人気曲の特徴語に「ている」の話し言葉である「てる」という単語があった. 同様に「ん」という単語も特徴語も抽出できた. 形態素解析のときに, 「出来るんだろう」の「出来る」と「だろう」間の「ん」を名詞として認識したことが要因と考えられる. 正しい文法で作られているとは限らない歌の歌詞だからこその特徴語であるといえる. 本研究では形態素解析に「MeCab」を使用した, 別の形態素解析ソフトウェアを使用することでまったく違う分析結果から新しいルールを獲得できると考えられる.