

同一タイトル楽曲の歌詞による可視化

小松明日香

現在発表されている J-POP において、異なる歌詞の楽曲に同一のタイトルがつけられていることがある。これを、本研究では“同一タイトル楽曲”と呼ぶ。このような楽曲をタイトルで検索するとき、目的でない多くの楽曲もヒットしてしまう。本研究では、歌詞を用いて楽曲の特徴を反映したタグクラウドを作成し、そのタグクラウドによって同一タイトル楽曲内で特定の楽曲を選択するのに適した特徴とその表示方法を明らかにすることを目的とした。

同一タイトル 5 タイトルから 6 曲ずつ、合計 30 曲の歌詞を用いて、大学生の男女 10 人を被験者として評価実験を行った。楽曲の特徴は歌詞に含まれる単語の重み、単語の品詞、単語の極性の 3 つとした。単語の重みは楽曲中の出現頻度に基づく TF-IDF 値である。単語の品詞は形態素解析器 ChaSen で判定された単語の品詞で、先行研究の結果を参考に名詞、動詞、感動詞、連体詞のいずれかとした。単語の極性は日本語評価極性辞書と単語感情極性対応表を用いて、ポジティブかネガティブかを判定した。タグクラウド作成の際、単語の重みは単語のフォントサイズで、単語の品詞と極性はそれぞれ単語の色で表現した。そして、「フォントサイズを TF-IDF 値の小さい方をより細かく分割するか、均等分割するか」、「サビのみを用いるか、歌詞全体を用いるか」、「色を付けない(ノーマル)か、品詞によって色を使い分けるか、極性によって色を使い分けるか」という条件の組み合わせ(2×2×3)で 12 種類作成した。種類ごとに同一タイトルの 6 曲のタグクラウドを A3 用紙 1 枚に印刷し(アーティストなど歌詞以外の情報は記載しない)、楽曲の特徴の違いがわかりやすいかどうかなどを被験者に判定してもらった。

「フォントサイズの分割」では、TF-IDF 値の小さい方をより細かく分割したものが 64%、均等に分割したものが 34% 選択され、TF-IDF 値の小さい方をより細かく分割した方が楽曲どうしの違いがわかりやすいという結果を得た。「サビのみと歌詞全体」では、サビのみが 54%、歌詞全体が 46% で、楽曲の比較のしやすさに違いはほとんど見られなかった。「ノーマルと品詞と極性」では、ノーマルが 8%、品詞が 54%、極性が 38% となり、品詞によって色を使い分けたものが一番多く選択された。また、同一タイトル楽曲のタグクラウドの中から自由に 1 曲を選んでもらった結果、アーティスト名など歌詞以外の情報がなくても、今回作成したタグクラウドによって楽曲を区別し、聞きたいと思った曲、興味を持った曲を選択できること、自分が聞いたことのある曲であると識別できるということがわかった。

本研究では、単語の大きさは TF-IDF 値の小さい方をより細かく分割し、品詞によって色を使い分けたタグクラウドが、より楽曲どうしの違いがわかりやすく、楽曲を選択しやすいということが明らかにされた。

(指導教員 緑川信之)