

4J 実験実習

「RADWIMPS の変遷に見る歌詞語彙の「標準化」に関して」

学籍番号: 22059

氏名: 来間 空

1. 背景

ロックバンド RADWIMPS は、デビュー当時はマニアックで個性的な歌詞を歌うバンドとして、一部の若者から熱狂的に支持されていた。筆者自身、彼らの紛ぐ膨大な言葉数や、時に哲学的できえある独特な恋愛観・死生観に救われたファンの一人である。

しかし、2016年の映画「君の名は。」の世界的ヒットをきっかけに、彼らは子供から大人まで誰もが知る国民的アーティストへと変貌を遂げた。一つとしてその活躍を誇らしく思う反面、かつての彼らが持っていた理解されない独白や刺すような言葉の鋭さが、洗練された綺麗な言葉へと置き換わってしまったのではないかという、寂しさを感じていたことも事実である。

一般的に、限られたファンのために歌うときは自分らしい独特的な言葉が許容されるが、大衆に向けて歌うときは誰にでも伝わるわかりやすい言葉を使う必要が出てくる。多くの人に曲を届けるためには、難しい表現を避け、共感しやすい言葉を選ぶようになるからである。

本研究は、単なる一バンドの歌詞分析にとどまらず、サブカルチャーを出自とする表現者がメインストリームへと移行する過程で、いかにして個人の作家性と大衆への適応という要素を言語レベルで折り合わせているのかを解明するモデルケースとなると考える。現代の音楽市場において、個性はいかにして国民的な共感へと翻訳されるのか。そのメカニズムを明らかにすることは、文化社会学的に重要な意味を持つ。

そこで本研究では、筆者が肌感覚として抱いていた歌詞の変化を、データを用いて客観的に検証することを目的とする。

本研究の仮説は以下の通りである。

「RADWIMPS の歌詞は、国民的アーティストへの変化に伴い、難解で独特的な言葉が減り、誰もが使う標準的でわかりやすい日本語へと変化している。」

この検証によって、アーティストが多くの人々に愛され続けるために、どのように自分らしさとわかりやすさのバランスを取っているのか、そして我々ファンは何を得て、何を失ったのかを明らかにすることができる。

2. 方法

① 本研究では、RADWIMPS の歌詞が国民的アーティストへの移行期にいかなる変容を遂げたかを検証するため、以下のデータセットおよび分析手法を用いた。

分析対象として、RADWIMPS の人気曲ランキング[1]から昇順に計 40 曲を選定し、2016 年を境界として対象楽曲を以下の 2 群に分類した。

- group old : 2016 年以前に発表されたアルバムおよびシングルから選定した 20 曲
- group new : 2016 年以降に発表されたアルバムおよびシングルから選定した 20 曲

歌詞データは、各楽曲の正規歌詞カードを参照し、テキストファイルとして電子化したものを用いた。

- ② 定量的分析には統計解析環境「R」を用い、形態素解析エンジンには「MeCab」およびR用パッケージ「RMecab」を採用した。

歌詞テキストに対して形態素解析を行い、品詞分解を実施した。分析の精度を高めるため、助詞や助動詞などの機能語を除外し、内容を表す主要品詞（名詞、動詞、形容詞）のみを抽出対象とした。また、活用のある語については原形（基本形）に統一して集計を行った。

- ③ 仮説を多角的に検証するため、以下の3つの指標を設定した。

(1) 語彙の一般性

歌詞に使用されている語彙が、一般的な日本語と比較してどの程度標準的であるかを測定する指標である。

比較対象として、現代日本語の書き言葉を均衡に収集したコーパス「BCCWJ（現代日本語書き言葉均衡コーパス）」[2]の語彙頻度データを用いた。歌詞に含まれる各単語をBCCWJの出現頻度ランクと突合し、上位ランク（高頻度語）に含まれる語彙の割合（基本語彙含有率）や、平均的なランク値を算出することで、歌詞の「平易さ」や「一般性」を定量化する。

(2) 語彙の多様性

歌詞の中でどれだけ多様な言葉が使われているかを測定する指標である。

単純な異なり語数（Type）を総語数（Token）で割る指標（TTR）は、歌詞の長さに依存しやすい欠点がある。そのため、総語数の平方根で異なり語数を割る「Guiraud Index（ギローの指標）」を採用した。

$$Guiraud\ Index = \frac{Type}{\sqrt{Token}}$$

この値が高いほど語彙が豊かで表現が多彩であり、低いほど同じ言葉が繰り返される平易な構成であることを示す。

(3) 感情極性

歌詞に含まれる感情の方向性と揺らぎを測定する指標である。

日本語極性辞書[3]を用い、各単語に付与された感情スコア（-1.0～+1.0）を歌詞に適用した。楽曲ごとのスコア平均値によって「ポジティブ／ネガティブ」の傾向を、分散（標準偏差）によって「感情の振れ幅（情緒の安定性）」を分析する。

3. 結果

3.1 分析データの基本統計量

表 1 曲の平均総語数

Group	Songs	Mean_Token
New	20	414.5
Old	20	477.4

表1より、歌詞の長さを表す1曲あたりの平均総語数（Token）について確認する。Group Old は平均477語、Group New は平均415語であった。つまり、両グループ間に大きな差は見られず、分析対象として十分な分量が確保されていることがわかる。

3.2 語彙の一般性

グループ別基本語彙率を図 1 に示す。

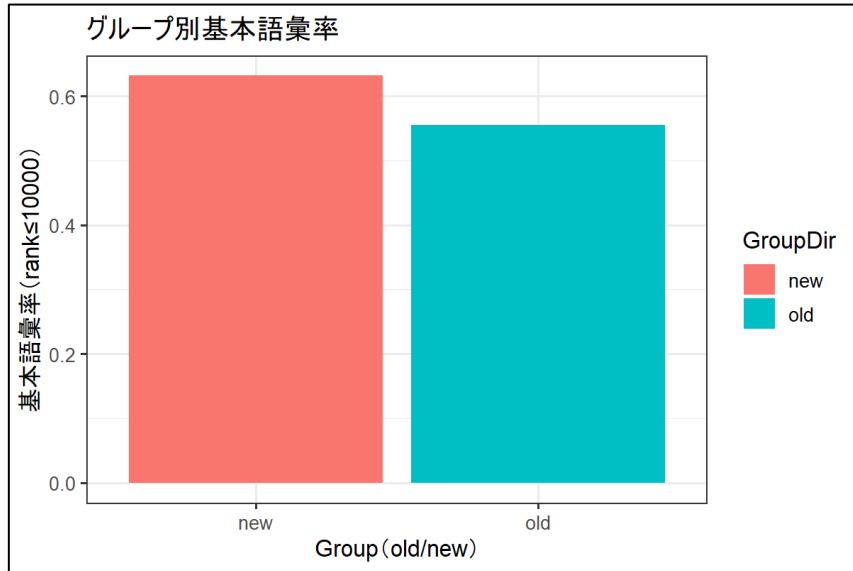


図 1 グループ別基本語彙率

図 1 は BCCWJ の頻度ランク上位 1 万語（日常的に極めて頻繁に使用される語彙）が歌詞全体に占める割合を示したものである。Group Old の平均含有率が 55.5% であったのに対し、Group New では 63.2% へと上昇した。t 検定の結果、両群には 5% 水準で有意な差が認められた ($p = 0.031$)。

3.3 語彙の多様性

語彙の多様性指標である Guiraud Index を表す図を図 2 に示す。

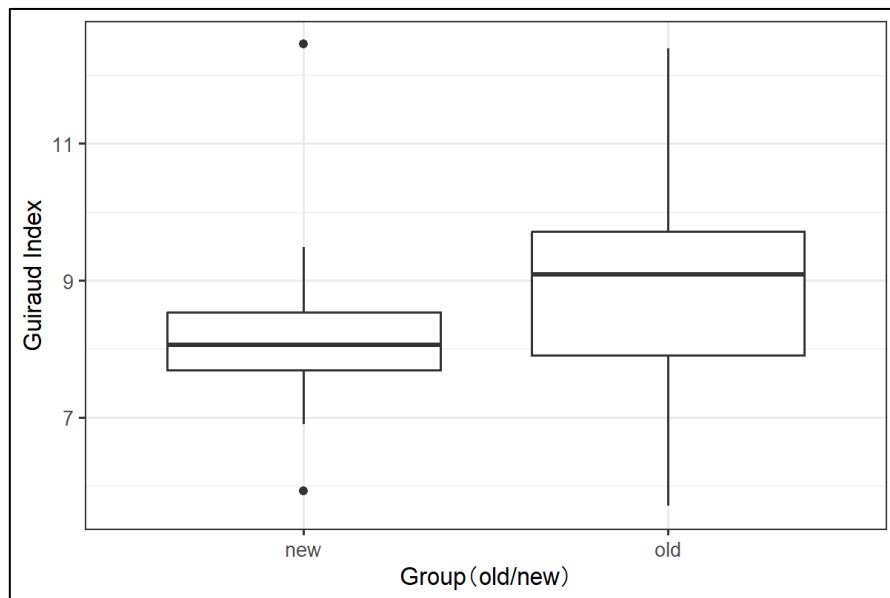


図 2 グループ別の Guiraud Index

図 2 より、Group Old の Guiraud Index 平均値は 8.75、Group New の平均値は 8.25 であった。数値上では Group New において多様性が低下しているが、t 検定による平均値の差の検定結果は $p=0.283$ であり、統計的な有意差は認められなかった。

3.4 感情極性

本分析では、歌詞に含まれる語彙のポジティブ・ネガティブ度を数値化し、その平均値と標準偏差を比較した。日本語極性辞書を用いた感情スコアの算出結果を図3に示す。

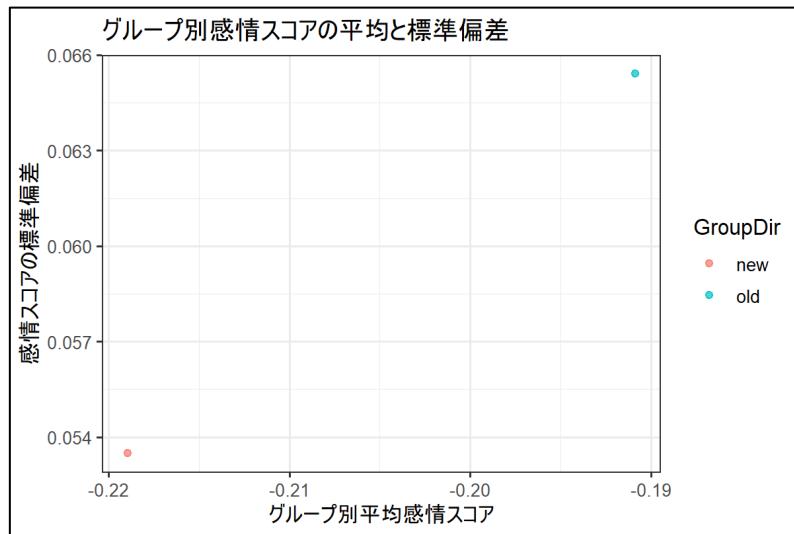


図3 グループ別感情スコアの平均と標準偏差

図3より、まず、感情スコアの平均値について確認する。分析の結果、Group Oldは-0.191、Group Newは-0.219となり、両群ともに負の値（ネガティブ傾向）を示した。 t 検定の結果、両群に統計的な有意差は認められなかった ($p = 0.146$)。

一方で、感情のばらつきを示す標準偏差に着目すると、Group Oldの0.065に対し、Group Newでは0.053となり、Newにおいて値が縮小している。

4. 考察

4.1 語彙の一般性

BCCWJを用いた語彙の一般性分析において、最も重要な知見は、Group Newにおける「基本語彙含有率（頻度ランク上位1万語）」の有意な上昇 ($p < 0.05$)である。結果3.2より、この統計的に有意な変化は、RADWIMPSがサブカルチャーの旗手から国民的ポップ・アイコンへと変貌する過程で、歌詞を構成する語彙を、特殊なものから標準的なものへと意図的に置換したことを示している。一般に、映画主題歌やCMソングといったタイプアップ楽曲は不特定多数の聴衆に向けて発信されるため、アーティストとファンが共有する文脈に依存した言葉選びはリスクとなる。したがって、Newにおける基本語彙の増加は、彼らが自身の個人的な想いを、誰にでも伝わる公共の言語へと翻訳する能力を獲得したことの証明である。

4.2 語彙の多様性

Guiraud Indexを用いた語彙の多様性分析において ($p = 0.283$)。結果3.3一見すると変化なしと捉えられるこの結果は、考察4.2と照らし合わせることで、RADWIMPSという作家の本質的な変化を浮き彫りにする。

通常、大衆化を志向するポップ・ミュージックにおいては、サビの単純な反復や、覚えやすいフレーズのリフレインを極端に多用するため、語彙の多様性は著しく低下する傾向にある。しかし、今回の分析結果は、彼らが国民的作家となった後も、歌詞における情報の密度を維持し続けていることを示唆している。つまり、彼らの変容は「言葉を削ぎ落として単純にする」ことではなく、「言葉の量は保ったまま、その質を難解なものから平易なものへ置換する」ことにあったと言える。

かつての彼らは、溢れ出る感情を制御不能なまま言葉に変換し、多様な語彙を乱れ打ちにすることで若者の共感を得ていた。対して現在は、その高い言語能力を維持しつつ、より多くの人に届くよう計算された構成の中で多様な言葉を配置している。

4.3 感情極性

結果 3.4 より、感情分析の結果を見ると、old と new で数値に大きな変化はなく、統計的な差も見られなかった。当初は「国民的な人気者になれば、歌詞も明るくポジティブになるだろう」と予想していたが、結果はそれを裏切るものだった。むしろ、数字の上では今の方が少しネガティブになっている。しかし、これは彼らが暗くなつたことを意味するのではない。RADWIMPS が多くの人々に支持される理由は、映画「君の名は。」で見られるような切なさや胸が締め付けられるような想いを歌っているからである。つまり、大衆が彼らに求めているのは底抜けの明るさではなく、誰もが共感できる哀愁なのだ。今回の結果は、彼らが国民的作家になつても、その一番の武器であるセンチメンタルズムを手放さなかつたことを証明している。

ここで最も重要な発見は、感情のばらつきが New になって小さくなつてきている点である。昔の彼らは、社会への怒りを爆発させた曲や極端に暗い曲など、感情の起伏が激しく、不安定であった。しかし今は、そうした極端な感情の波が抑えられ、一定の範囲内でコントロールされている。結論として、彼らの感情表現の変化は、ネガティブからポジティブへという単純なものではなく、自分勝手で不安定な感情のぶつけ合いから、誰もが安心して感動できるような整えられた切なさへと成熟したと言える。

4.4 データ量に関して

テキストデータの総量である。分析 1 で示した通り、RADWIMPS の楽曲は 1 曲あたり平均 400 語以上の語彙を含んでいる。これを 20 曲積算すると、各群約 8,000~10,000 語規模のコーパスとなる。個人の文体や語彙傾向を分析する計量言語学 (Stylometrics) の分野において、数千語~1 万語程度のデータ量は、著者の特徴を安定して抽出するのに必要十分な水準とされている。

5. 結論

本研究の出発点である「RADWIMPS の歌詞は、国民的アーティストへの変化に伴い、難解で独特的な言葉が減り、誰もが使う標準的でわかりやすい日本語へと変化している」という仮説は、本分析のデータによって強く支持されたと結論づけられる。その最大の根拠は、語彙の一般性分析において、誰もが日常的に使用する「基本語彙」の含有率が Group New で統計的に有意に上昇した点にある。かつての彼らが多用していた辞書にない造語や難解な表現（独特的な言葉）は減少し、代わりに多くの聴衆が共有する「標準的な日本語」への明確なシフトが確認された。

ただし、この変化は単なる語彙力の低下や単純化ではない。語彙の多様性 (Guiraud Index) が大きく低下しなかつた事実は、彼らが言葉の豊かさを保ち続けていることを意味する。つまり、彼らは言葉を減らしたのではなく、言葉の選び方をマニアックな独白から標準的な対話へと高度に最適化したのである。また、作家性の核である感情についても、ネガティブな要素（切なさ）を維持しつつ、その振れ幅を安定させることで、万人が安心して共感できるエンターテインメントへと昇華させている。

以上のことから、RADWIMPS が国民的ポップ・アイコンへと変貌を遂げられた要因は、作家としての個性を殺して大衆に迎合したからではない。むしろ、センチメンタルな作家性を維持しつつ、それを標準的な日本語というパッケージで包み込む高度な言語的適応を成し遂げた点にある。本研究の結果は、現代の音楽市場においてアーティストが広く大衆に受容されるためには、作品のクオリティだけでなく、自身の作家性を社会の共通言語へと翻訳する標準化のプロセスが不可欠であることを示唆している。

参考文献

- [1] 「RADWIMPS の人気曲ランキング」, JOYSOUND, <https://www.joysound.com/web/search/artist/38875/ranking>, 2026 年 1 月 22 参照.
- [2] 「『現代日本語書き言葉均衡コーパス』短単位語彙表(Version 1.0)」, 国立国語研究所, <https://repository.ninjal.ac.jp/records/3234>, 2021 年 4 月 23 日公開.
- [3] 高村大也, 「単語感情極性対応表」, http://www.lr.pi.titech.ac.jp/~takamura/pndic_ja.html, 2026 年 1 月 22 日参照.

感想

今回の R を用いたテキストマイニングは非常に新鮮なものだった。統計的な処理は python や MATLAB でおこなったことはあったが、テキストを形態素解析し、それを様々な手法で分析したのは初めての経験でとても有意義な実験だと感じた。また、本授業で扱ったツールの中で、特に自身の研究に有用であると感じたのは、RMeCab を用いた形態素解析と、極性辞書による感情分析の手法である。まだ自分の卒業研究のテーマは決まってないが、これまで定性的な印象論で語られがちだったこの分野に対し、今回の授業で実践したように、テキストデータを数値化し、客観的な指標として可視化するアプローチを取り入れたいと考えている。

授業に関して、RMeCab の最適なインストール方法が人によって違ったりしたため、もう少し汎用的なインストール方法を確立するべきだと感じた。また、授業 1 回目では環境構築がメインで、実質的な実験は 1 回しかなかったため、R を十分に習得できていないと感じた。よって、もう少し環境構築の時間を減らして（事前にやってもらうとか）、実験にかける時間を増やすべきだと感じた。

全体的に、この実験はほかの実験と違って、少し文系的な要素があり斬新であったため、有意義な実験であった。