

青山学院大学 社会情報学部 2021 年度 卒業論文

PLSA と MAP 推定を用いた日本語歌詞の印象推定手法の評価

学籍番号 18118047

氏名 黒川 皇輝

指導教員 宮治 裕 教授

2022 年 1 月

論文要旨

近年スマートフォンとオンライン音楽配信サービスの発達により、人々は膨大な数の楽曲を手軽に聴けるようになった。人々がその膨大な楽曲の中から自分の嗜好にあった楽曲を見つけ出すことは困難であるため、楽曲推薦サービスの需要が高まっている。

楽曲推薦サービスの中でも人間の気分や状況に基づいて楽曲を推薦するサービスをコンテキストウェアシステムと呼ぶ。そのシステムの中でも日本語歌詞から人間が感じる印象を推定し、ユーザの気分に合わせて楽曲を推薦する研究は未だ発展途上であった。

本研究ではPLSAによって歌詞中に出現する単語を分析することで、4クラスの潜在的トピックにクラスタリングする。そして得られたモデルパラメータをMAP推定することで、潜在的トピックを喜怒哀楽の4つの印象に制限し、歌詞中に出現する単語から感じとれる印象を推定する。印象は感性平面上にプロットすることで表現する。歌詞をフレーズごとに分割して、フレーズに登場する単語の印象を合計して正規化することでフレーズの印象とする。フレーズの印象を合計して正規化することで歌詞全体の印象とする。

歌詞とフレーズに対して推定された印象が妥当であるか有効性を示すために実験をおこなった。その結果喜の印象推定は妥当であることが確認できた。しかし、その他の印象における推定結果は妥当であるとは言えなかった。喜のクラスだけが印象推定で有効性を示せた理由はMAP推定に用いた事前知識の数が十分にあったことが一番大きな要因として考えられる。事前知識の数が少ないとMAP推定する際に無情報事前分布を与えるため、MAP推定前の潜在トピックに属するモデルパラメータと変わらない。そのためうまく印象を推定できなかったと考慮する。したがって事前知識を十分確保できた場合、本論で設計し

た印象推定手法が有効であると期待される。

謝辞

本研究を進めるにあたり，多くの方々にご指導ご鞭撻を賜りました．指導教員の宮治裕教授からは多大なご指導を賜り，時には過去文献をひも解くヒントなどご教示いただき感謝の念に堪えません．ありがとうございました．調査の実施にあたり，宮治研究室の皆さまと同大学の同期には実験参加者を務めて下さり，貴重なデータ収集にご協力いただいたことを感謝いたします．最後に，本研究ならびに学業全般にわたって経済的・心身的に支援して下さる家族に深く感謝し，お礼を申し上げます．

目次

論文要旨	i
謝辞	iii
第1章 はじめに	1
1.1 背景	1
1.2 研究目的	2
1.3 論文構成	2
第2章 関連研究	3
2.1 対話よるムード推定に基づく楽曲推薦エージェントの開発	3
2.2 物体の色や表情情報を利用した画像の印象にあった音楽推薦手法の提案	3
2.3 歌詞解析と心拍変動分析を用いた楽曲によふ感情への影響の予備的調査	4
2.4 本研究の位置付け	4
第3章 歌詞の印象推定	5
3.1 分析手法	5
3.1.1 MAP推定	5
3.1.2 PLSA	7
3.1.3 日本語版ANEW拡張データセット	8
3.2 歌詞の印象推定手法	9
3.2.1 印象の定義	9
3.2.2 歌詞データの収集	9

3.2.3	歌詞データの整形	9
3.2.4	歌詞印象推定	9
3.3	AV 平面にプロット	12
3.3.1	楽曲	12
3.3.2	フレーズ	17
第 4 章	実験	22
4.1	目的	22
4.2	方法	22
4.2.1	歌詞の印象評価	23
4.2.2	フレーズの印象評価	23
4.3	結果	24
4.3.1	歌詞の印象評価	24
4.3.2	フレーズの印象評価	37
4.4	考察	49
4.4.1	歌詞の印象評価実験の考察	49
4.4.2	フレーズの印象評価実験の考察	49
第 5 章	おわりに	51
5.1	まとめ	51
5.2	今後の課題	51
参考文献		53

第1章

はじめに

本論文では日本語歌詞から得られる印象を推定する手法の設計と実装をし，手法の効果について記述する．本章では本研究をおこなう背景となった事柄として研究目的を記述した後，次章以降の本論文の構成についてその概略を述べる．

1.1 背景

近年 Amazon Music, Spotify といった音楽配信サービスとスマートフォンなどのモバイル端末の普及により，人々は膨大な楽曲を手軽に聴くことができるようになった．それに伴い音楽配信サービスのユーザは増加傾向にある [1]．膨大な楽曲の中からユーザが自身の嗜好に合う楽曲を見つけるのは困難であるため，楽曲を推薦するシステムが研究トピックとして着目されている．

音楽推薦システムの中でもユーザの気分や状況に合う楽曲を推薦するシステムをコンテキストウェア楽曲推薦システムと呼ぶ．コンテキストとは次のように定義されている．「エンティティの状況の特徴化するのに用いられるあらゆる情報．エンティティとは，ユーザとアプリケーションとのインタラクションに関連する人や場所，オブジェクトを指し，それにはユーザ自身とアプリケーション自体も含まれる」(奥，2019, pp.300~308)．コンテキストは大まかにユーザコンテキスト，環境コンテキスト，マルチメディアコンテキストの3つに分類される．ユーザコンテキストはユーザの気分や生体情報である．環境コンテキストは位置情報や時間，天気などのユーザを取り巻く環境情報である．マルチメディアコンテキストはテキストや映像，画像などの音楽以外のメディア

からユーザが得る感性情報である。マルチメディアコンテキストの分野において、日本語歌詞からユーザが受ける印象を推定する手法は研究されているが未だ発展途上である。

1.2 研究目的

本研究の目的は日本語歌詞のテキスト情報を russell の Arousal-Valence 平面 [2] (以下 AV 平面) 上で表現することで、ユーザが感じる歌詞の印象を推定する。印象推定手法は西川ら (2011) [3] が設計した楽曲印象推定手法を参考に、日本語歌詞から読み取れる印象を推定する手法の設計と評価をおこなう。

1.3 論文構成

2章では関連研究について述べる。3章では日本語の歌詞の印象推定手法について述べる。4章では印象推定手法の有効性を確かめるために行なった実験について述べる。5章では本研究についてのまとめと考察、今後の課題について述べる。

第2章

関連研究

本章では日本語の歌詞情報から印象を推定する手法についていくつか紹介したのちに，本研究の立ち位置を説明する．

2.1 対話よるムード推定に基づく楽曲推薦エージェントの開発

角田ら (2018) [4] は印象を Plutchik の感情モデル [5] の基本8感情を用いて推定した．基本8感情は喜び，信頼，心配，驚き，悲しみ，嫌悪，怒り，予測である．歌詞情報を MeCab による形態素解析をして単語を抽出する．その後 Word2Vec を利用して，抽出した単語のベクトル変換を行ない，基本8感情との類似度を算出して，一番類似度が高い感情をテキストから得られる印象とした．この研究では喜び・悲しみなどの推定できた印象と嫌悪・怒りなどの推定できなかった印象が存在した．

2.2 物体の色や表情情報を利用した画像の印象にあった音楽推薦手法の提案

追木ら (2018) [6] は印象を AV 平面上で表現した．英語の歌詞を Word2Vec を用いて Arousal-Valence 値が既知である印象語と類似度を計算した．Arousal-Valence 値が既知である印象語は Georgios(2013)[7] らが公開されたものを使用している．歌詞ともっとも類似度が高い印象語の重みを1として正規化し，各印象の重みつけを行う．各印象語間の角度によるクラスタリング手法である spherical-kmeans[8] を用いてクラスタリングを行った．クラスタ数は4つであり，AV 平面

の各象限である．しかしこの実験では人が認識した印象とシステムが推定した印象の一致率は低かった．

2.3 歌詞解析と心拍変動分析を用いた楽曲による感情への影響の予備的調査

大木ら (2018) [9] は日本語評価極性辞書にのっとり，歌詞をポジティブ・ネガティブ・ニュートラルの3つの印象に分類した．3つの評価極性を歌詞内の各単語に付与し，単語の評価極性値を平均化することで歌詞自体への印象の評価値を算出する．歌詞をポジティブ，ネガティブ，ニュートラルの3印象にうまく推定できたが，より詳細な印象（喜び，悲しみなど）は正確に推定できない．

2.4 本研究の位置付け

上記で述べたとおり，歌詞情報から印象を推定する手法は未だ確立されていない．ユーザの気分に合う楽曲を探すために日本語の歌詞を持つ楽曲がユーザに与える印象を推定する手法を模索する．

西川らは英語の感情価単語セット ANEW[10] を用いて，歌詞を AV 平面上で表現する研究を行なった．ANEW とは英単語の 1,034 語について感情価を調査したデータセットである．この研究では歌詞の各フレーズが持つ単語の AV 平面座標を推定して，フレーズごとの平面座標を求めた．この印象推定手法は英語の歌詞での効果測定が行われたが，日本語の歌詞での効果測定は行われていないので，日本語の歌詞の場合でもこの手法が効果を発揮するのか調査する必要があると考慮する．本間ら [11] は ANEW の単語を参考にして，日本語訳した単語セットを使用し，単語の感情価と覚醒度の評定をする．その結果，AV 平面上での単語の分布は Bradley らと同様の分布を得られることが確認された．

よって，本間らが作成した日本語の単語セットを利用すれば，西川らが考案した手法を日本語の歌詞の印象推定でも一定の効果を挙げられることが予想される．

第3章

歌詞の印象推定

本研究で行った歌詞の印象推定法について説明する．最初に研究で用いた分析手法について述べたのちに，分析手法を活用した歌詞の印象推定手法について述べる．その後，分析結果を記述する．

3.1 分析手法

この節では歌詞印象推定手法で用いた分析手法について紹介する．

3.1.1 MAP 推定

MAP推定とは事前知識に基づいて未知のデータを点推定する手法である． θ を尤度， D を事前知識とすると，ベイズの定理に基づき次の式(3.1)で表せる．

$$\begin{aligned}\operatorname{argmax}_{\theta} P(\theta | D) &= \operatorname{argmax}_{\theta} \frac{P(D | \theta) P(\theta)}{P(D)} \\ &= \operatorname{argmax}_{\theta} P(D | \theta) P(\theta)\end{aligned}\tag{3.1}$$

ここで分母の $P(D)$ は尤度 θ と関係がないので無視する． $P(D | \theta)$ は尤度関数， $P(\theta)$ は事前分布， $P(\theta | D)$ は事後分布と呼ぶ．事前分布 $P(\theta)$ は複雑な計算を要する．そのため，それを回避するために事前分布の代わりに共役事前分布を用いて推定をする．共役事前分布とは尤度をかけて事後分布を求めるとその関数の形が同じになる事前分布のことである．事後分布が最大となる尤度 θ を算出するのがMAP推定である．

本研究ではカテゴリカル分布に基づくMAP推定をする．したがって事前知識 D はカテゴリカル分布に従う N 個の離散値データ $D = \{d_1, d_2, \dots, d_N\}$ と定義する．

尤度関数はカテゴリカル分布の確立質量関数で表せるので式 (3.2) となる.

$$P(D | \theta) = \prod_{n=1}^N \text{Cat}(d_n | \theta) \quad (3.2)$$

カテゴリカル分布の共役事前分布はディリクレ分布であるため事前分布を式 (3.3) で表す. ディリクレ分布はハイパーパラメータ α が大きくなるにつれて分散が小さくなる特性を持つ. ハイパーパラメータ α は対応する事象 d の観測数の初期値を意味する. α を全て 1 とした場合, ディリクレ分布は一様な無情報事前分布となり, 事後分布は尤度関数であるカテゴリカル分布と同じ分布になる.

$$P(\theta) \propto \text{Dir}(\theta; \alpha) \quad (3.3)$$

式 (3.2) と式 (3.3) を式 (3.1) に当てはめると事後分布は式 (3.4) のように表せる.

$$P(\theta | D) \propto \left\{ \prod_{n=1}^N \text{Cat}(d_n | \theta) \right\} \text{Dir}(\theta; \alpha) \quad (3.4)$$

式 (3.4) の両辺に対数をとった式 (3.5) を示す. ここで $d_{n,k}$ は K 次元ベクトル d_n の k 番目の要素を指し示す. \mathbb{C} は不定定数である.

$$\begin{aligned} \log P(\theta | D) &\propto \sum_{n=1}^N \log \text{Cat}(d_n | \theta) + \log \text{Dir}(\theta; \alpha) + \mathbb{C} \\ &= \sum_{k=1}^K \left(\sum_{n=1}^N d_{n,k} + \alpha_{n,k} - 1 \right) \log \theta_k + \mathbb{C} \end{aligned} \quad (3.5)$$

このまま偏微分しても極値を得ることはできないのでラグランジュの未定乗数法を使用する. ラグランジュ関数は式 (3.6) である.

$$L = \log P(\theta | D) + \lambda \left(\sum_{k=1}^K \theta_k - 1 \right) \quad (3.6)$$

ラグランジュ関数 L を θ_k で偏微分して 0 となる極値を求めると MAP 推定値 $\hat{\mu}_{\text{map},k}$ が得られる.

$$\hat{\mu}_{\text{map},k} = \frac{\alpha_k - 1 + \theta_k}{1 - K + \sum_{j=1}^K \alpha_j} \quad (3.7)$$

3.1.2 PLSA

PLSA[12]とは次元圧縮手法の一種であり，情報検索の分野で膨大な文書データを分類するために研究された手法である．PLSAは文書 d とその文書中に出現する単語 w の間に共通のトピックと呼ぶ潜在意味クラス z が存在することを想定として，潜在意味クラス z を確率的に推定する手法である．そして，文書 d と単語 w の共起確率 $P(d,w)$ を潜在意味クラス z を用いてモデル化する． z でモデル化した $P(d,w)$ の式を式(3.8)にグラフィカルモデルを図1に示す．

$$P(d, w) = \sum_z P(z)P(d|z)P(w|z) \quad (3.8)$$

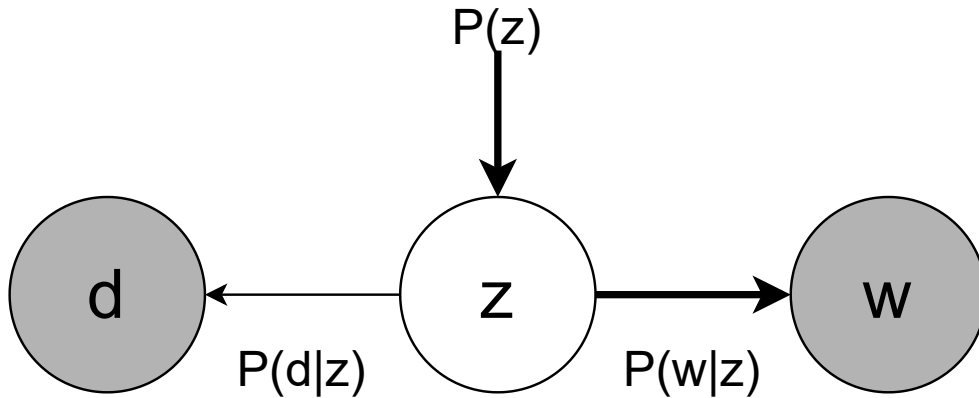


図 3.1: PLSA グラフィカルモデル

文書 d と単語 w の同時出現頻度を $N(d,w)$ とすると，対数尤度関数 $\log L$ は式(3.9)として示せる．

$$\log L = \sum_d \sum_w N(d, w) \log(P(d, w)) \quad (3.9)$$

この対数尤度関数 $\log L$ を最大にする確率変数 $P(z), P(d|z), P(w|z)$ をEMアルゴリズムによって推定する．なお推定する確率変数をモデルパラメータと呼ぶ．EMアルゴリズムとは混合分布モデルのパラメータ推定に利用できる学習アルゴリズムである． z の確率分布 $P(z|d,w)$ を予測するE(expectation)ステップと対数尤度関数の最大化をするM(maximization)ステップに計算ステップが分かれている．

それぞれのステップで行われる計算の説明を順番に述べる．式 (3.10) は E ステップの式である．初期値である $P(z), P(d|z), P(w|z)$ は絶対値が 1 以下のランダムな自然数に決定され， z の確率分布 $P(z|d, w)$ を予測する．

$$P(z|d, w) = \frac{P(d|z)P(w|z)P(z)}{\sum_z P(d|z)P(w|z)P(z)} \quad (3.10)$$

式 (3.11) (3.12) (3.13) は M ステップの式であり，それぞれの式でモデルパラメータを求めている．

$$P(d|z) = \frac{\sum_w N(d, w)P(z|d, w)}{\sum_d \sum_w N(d, w)P(z|d, w)} \quad (3.11)$$

$$P(w|z) = \frac{\sum_d N(d, w)P(z|d, w)}{\sum_d \sum_w N(d, w)P(z|d, w)} \quad (3.12)$$

$$P(z) = \frac{\sum_d \sum_w N(d, w)P(z|d, w)}{\sum_d \sum_w \sum_z N(d, w)P(z|d, w)} \quad (3.13)$$

求めたモデルパラメータを式 (3.2) に当てはめて式 (3.3) の対数尤度関数の値を求める．対数尤度関数の値を最大化するまで E ステップと M ステップを交互に繰り返し，最適なモデルパラメータを求める．

3.1.3 日本語版 ANEW 拡張データセット

本間らが開発した日本語版 ANEW の単語の類義語と同義語を WordNet[14] を用いて探索する．発見した単語に類義語元の単語が持つ Arousal と Valence の値を与えることで，日本語版 ANEW を 14,232 語に拡張した単語データセットを日本語版 ANEW 拡張データセットとする．日本語版 ANEW 拡張データセットの構成は後述する A+V+ 平面に 5,581 語，A+V-平面に 2,681 語，A-V-平面に 3,456 語，A-V+平面に 2,493 語である．

3.2 歌詞の印象推定手法

この節では歌詞の印象推定手法について紹介する。

3.2.1 印象の定義

本稿で推定する印象とは Russell の AV 平面上で表現する。AV 平面は縦軸に Arousal を、横軸に Valence を取る 2 次元平面である。Arousal 軸は正の方向に興奮を、負の方向に弛緩を表す。そして、Valence 軸は正の方向に快を、負の方向に不快を表す。AV 平面の 4 象限は大まかに喜怒哀楽の印象を表す。第 1 象限は喜びや幸福、興奮、驚きといった喜の感情グループを表す。第 2 象限は怒りや恐れ、嫌悪といった怒の感情グループを表す。第 3 象限は退屈や悲しみ、憂鬱といった哀の感情グループを表す。第 4 象限は満足や穏やか、くつろぎといった楽の感情グループを表す。AV 平面に歌詞データをプロットし、データが存在する象限から歌詞データを 4 つのグループの印象にクラスタリングする。

3.2.2 歌詞データの収集

歌詞データの収集は Uta-Net^{*1}から人気のアーティストの曲を 6,813 曲収集した。本研究では 6,813 曲から 3,000 曲を無作為に選んだ。

3.2.3 歌詞データの整形

収集した歌詞データの中に一部英語が入っていたので、英語の歌詞を削除した。そして歌詞データを歌詞のフレーズごとに分解した。形態素解析器 MeCab を用いて各フレーズから名詞・形容詞・動詞を抜き出した。歌詞から抜き出した単語の総数は 109450 語である。

3.2.4 歌詞印象推定

確率的潜在意味解析 (PLSA) を利用する。歌詞のフレーズを文書 d とし、フレーズ中に出現する単語を単語 w と定義してモデルパラメータを推定した。

^{*1} <https://www.uta-net.com/>

$P(w|z)$ はトピックから単語が観測される確率であるため，この確率の高い単語がトピックを表現する．歌詞の印象を推定するためには潜在的なトピック z を印象に制限する必要がある．通常の PLSA は文書と単語の共起確率に着目して，潜在的なトピックを推定する手法であるため，必ず潜在的なトピックが印象を表現するとは限らない．よって，日本語版 ANEW 拡張データセットを事前知識として使用し，モデルパラメータを MAP 推定することで潜在的なトピックに AV 平面の各象限を表現する．具体的に潜在的なトピック z を次の式 (3.14) のように AV 平面の各印象として定義する．

$$z \in \{A + V +, A + V-, -A - V+, -A - V-\} \quad (3.14)$$

モデルパラメータ $P(w|z)$ の対数事前分布を共役事前分布を用いて式 (3.15) に定義する． k は出現する単語の集合を表す．

$$\log P(\theta) \propto \sum_k \sum_z (\alpha_{w_k, z} - 1) \log P(w_k|z) \quad (3.15)$$

共役事前分布のハイパーパラメータ α は日本語版 ANEW 拡張データセットに含まれる単語で該当の象限に位置する場合のみ原点からの距離を入力する．それ以外の場合無情報事前分布を与える．ラグランジュの未定乗数法を使用して対数尤度関数と対数事前分布より MAP 推定値を求める MAP 推定値は式 (3.16) で表す． N は MAP 推定前の $P(w|z)$ の和である．

$$P(w_k|z)_{\text{map}} = \frac{\alpha_{w_k, z} - 1 + P(w_k|z)}{1 - K + \sum_k \sum_z \alpha_{w_k, z}} \quad (3.16)$$

$P(z)$ と $P(d|z)$ の推定には無情報事前分布を与えた．各印象ごとに単語 w が出現する確率 $P(w|z)$ を MAP 推定した．ベイズの定理より $P(z|w)$ は式 (3.17) で求められる．

$$P(z | w_k) = \frac{P(w_k | z)_{\text{map}} P(z)}{P(w_k)} \quad (3.17)$$

$P(w)$ は単語が出現する確率である．ある単語の出現数を全ての単語の出現数で割ると求まる．各フレーズ d の Arousal と Valence の値は求めた $P(z|w)$ を用いて求めた各単語 w の Arousal と Valence の値を合計．正規化して求める．式は (3.18, 3.19) である． K はあるフレーズに出現する単語数である．

$$V = \frac{1}{K} \sum_{k=1}^K ((P(A + V+ | w_k) + P(A - V+ | w_k)) - (P(A + V- | w_k) + P(A - V- | w_k))) \quad (3.18)$$

$$A = \frac{1}{K} \sum_{k=1}^K ((P(A + V+ | w_k) + P(A + V- | w_k)) - (P(A - V+ | w_k) + P(A - V- | w_k))) \quad (3.19)$$

以上で歌詞中のフレーズごとに感情価を求める。

3.3 AV 平面にプロット

この節では歌詞を AV 平面上にプロットした結果について記述する。

3.3.1 楽曲

楽曲に登場する歌詞のフレーズの AV 値の平均を歌詞の AV 値とした。図 3.2 は AV 平面に楽曲をプロットした図である。

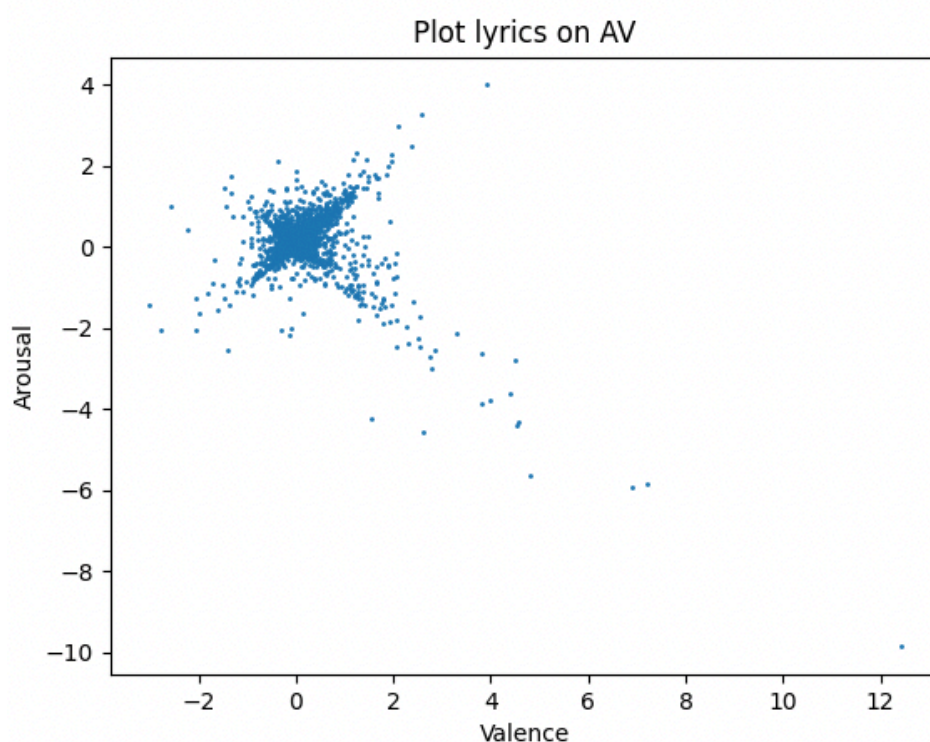


図 3.2: 歌詞 AV 平面

各印象ごとに歌詞を分類して，ward 法で原点からの距離に基づいて 3 グループにクラスタリングした。薄い肌色でプロットされている原点から一番近いグループを第 1 グループ，赤紫色でプロットされている原点から 2 番目に近いグループを第 2 グループ，黒色でプロットされている原点からもっとも遠いグループを第 3 グループとする。

第1象限のプロット図は図3.3である。

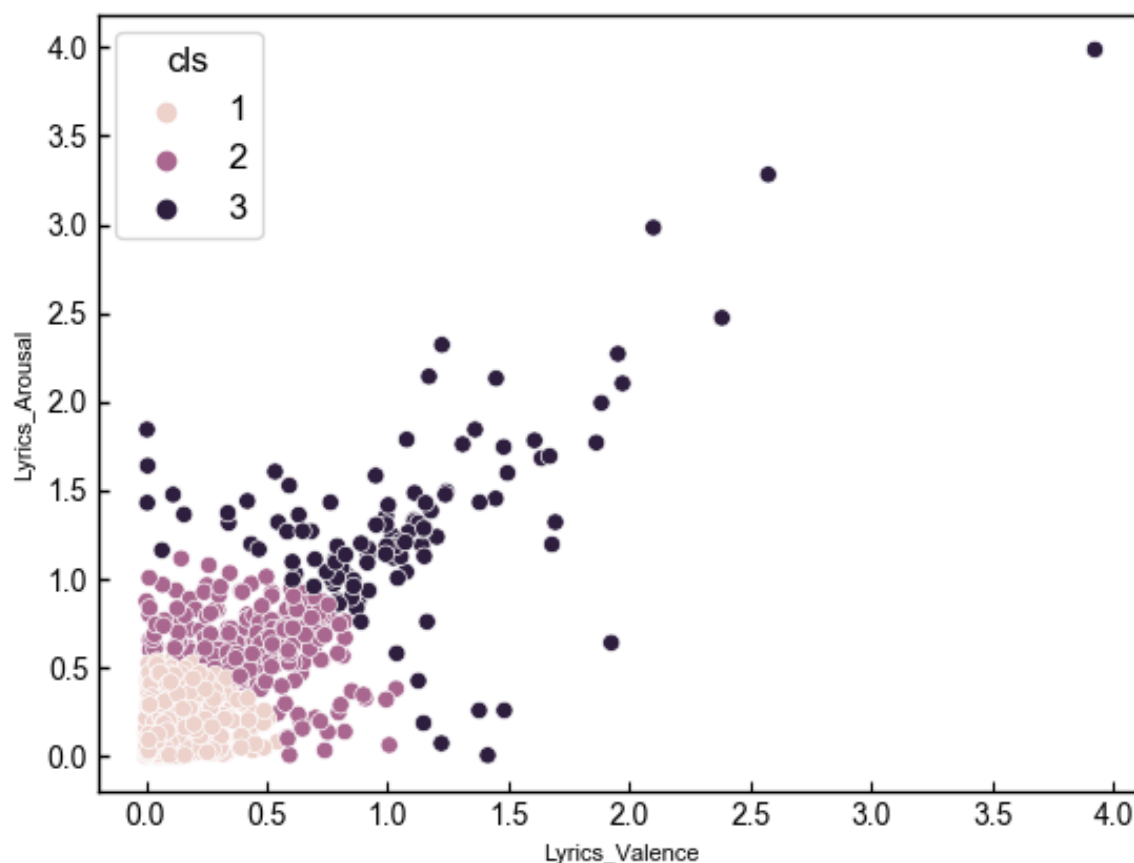


図 3.3: 歌詞 A+V+ 平面

第1象限にプロットされた楽曲は全部で1,424曲である。そのうちワード法によって分割された曲数は第1グループが1015曲，第2グループが296曲，第3グループが112曲であった。後述する実験に用いる楽曲の歌詞データはそれぞれのグループからランダムで1曲分選出する。第1グループからはコブクロの「光の誓いが聴こえた日」，第2グループからは椎名林檎の「積み木遊び」，第3グループからはDREAM COME TRUEの「冬三昧にはまだ遠い」の3曲を選出した。

第2象限のプロット図は図3.4である。

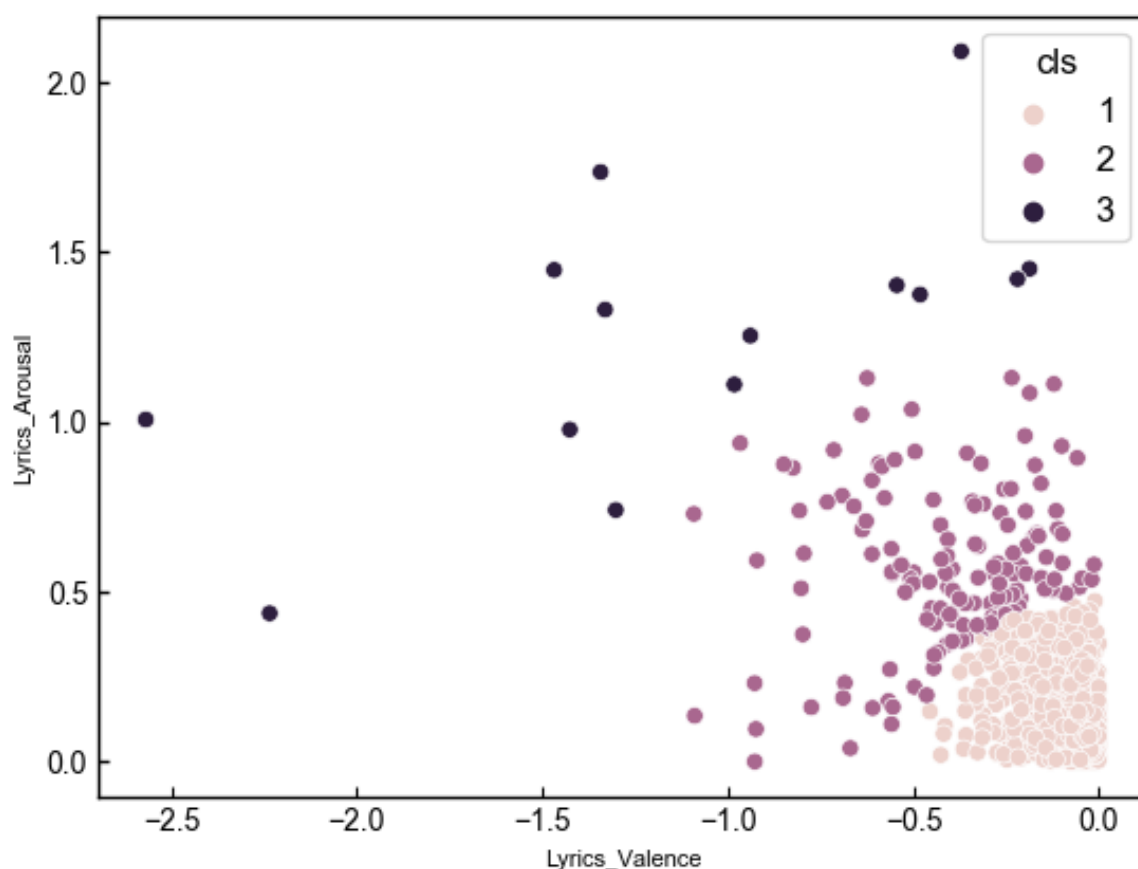


図 3.4: 歌詞 A+V-平面

第2象限にプロットされた楽曲は全部で887曲である．そのうちワード法によって分割された曲数は第1グループが728曲，第2グループが145曲，第3グループが14曲であった．後述する実験に用いる楽曲の歌詞データはそれぞれのグループからランダムで1曲分選出する．第1グループからは，第2グループからはRADWIMPSの「透明人間18号」，第3グループからはB'zの「Da La Da Da」の3曲を選出した．

第3象限のプロット図は図3.5である。

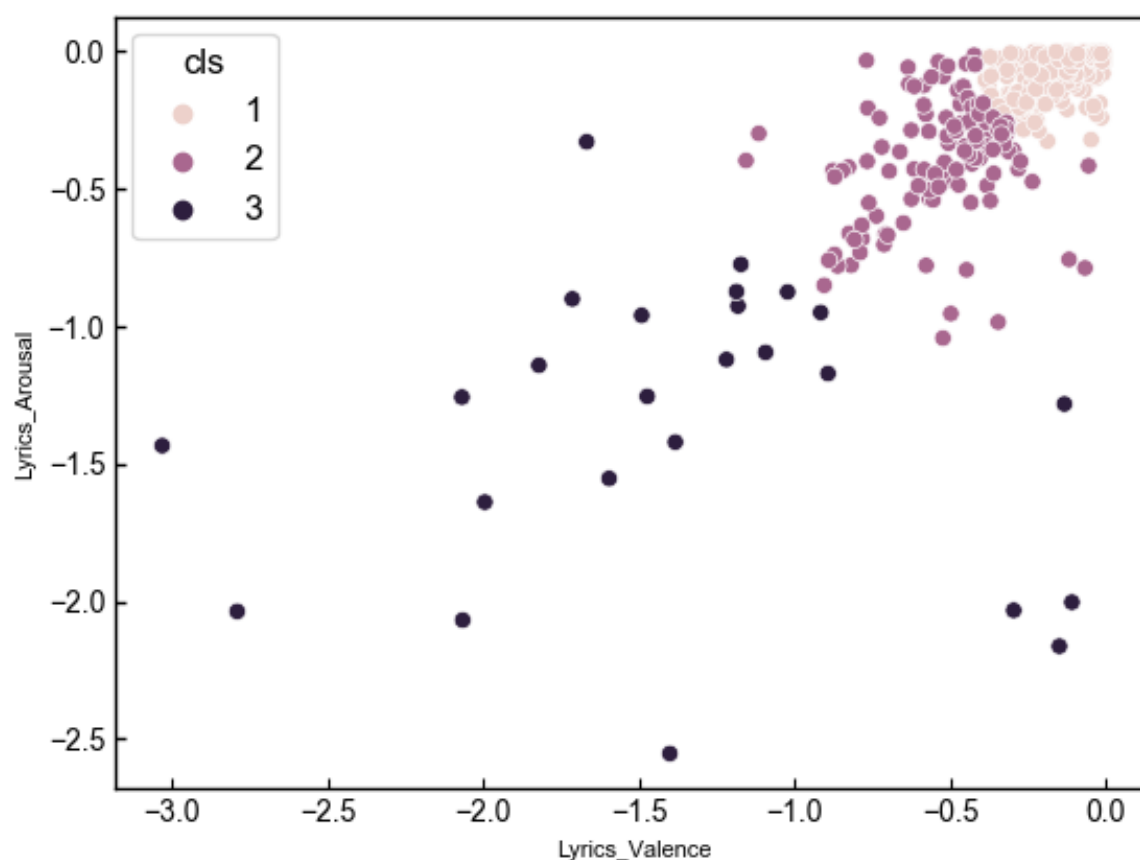


図 3.5: 歌詞 A-V-平面

第3象限にプロットされた楽曲は全部で388曲である。そのうちワード法によって分割された曲数は第1グループが236曲，第2グループが127曲，第3グループが25曲であった。後述する実験に用いる楽曲の歌詞データはそれぞれのグループからランダムで1曲分選出する。第1グループからはEvery Little Thingの「Just be you」，第2グループからはコブクロの「コイン」，第3グループからはBUMP OF CHICKENの「分別奮闘記」の3曲を選出した。

第4象限のプロット図は図3.6である。

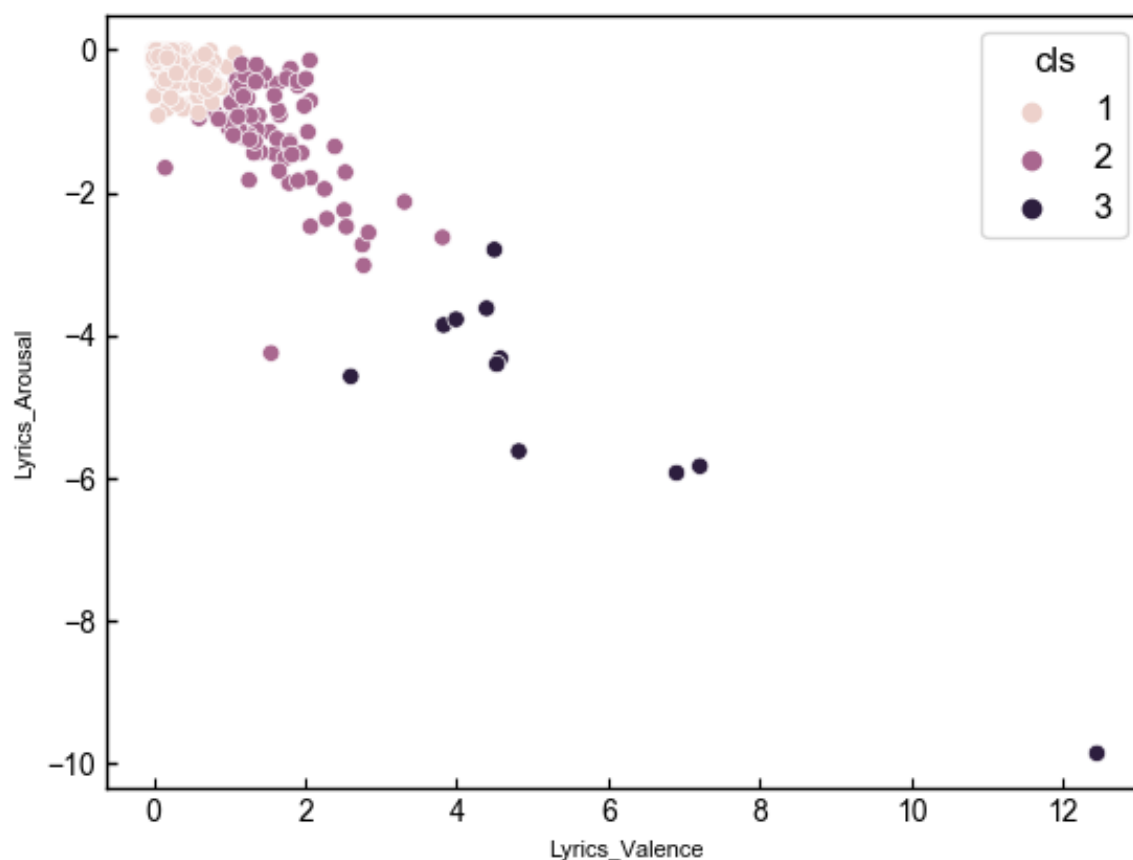


図 3.6: 歌詞 A-V+ 平面

第4象限にプロットされた楽曲は全部で302曲である。そのうちワード法によって分割された曲数は第1グループが210曲、第2グループが81曲、第3グループが11曲であった。後述する実験に用いる楽曲の歌詞データはそれぞれのグループからランダムで1曲分選出する。第1グループからはUVERworldの「若さ故エンテレケイア」、第2グループからはあいみょんの「お互い様やん」、第3グループからはスピッツの「シロクマ」の3曲を選出した。

3.3.2 フレーズ

歌詞に登場するフレーズを AV 平面にプロットした。図 3.7 は AV 平面に楽曲をプロットした図である。

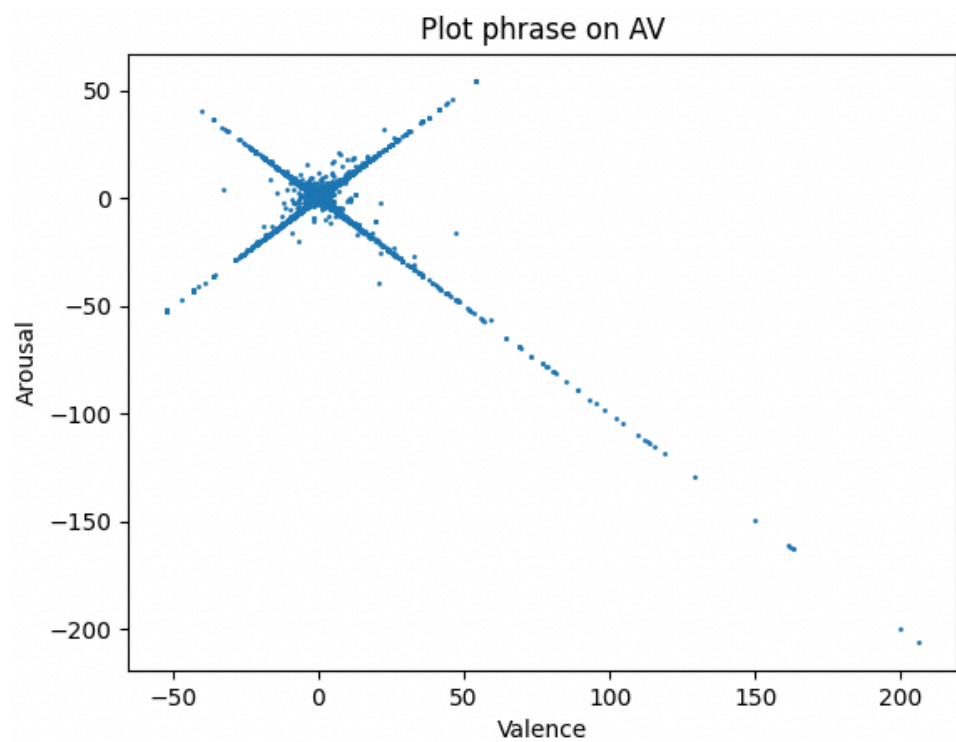


図 3.7: フレーズ AV 平面

歌詞と同じく各印象ごとにフレーズを分類して、原点からの距離に基づいて ward 法でクラスタリングした。

第1象限のプロット図は図3.8である

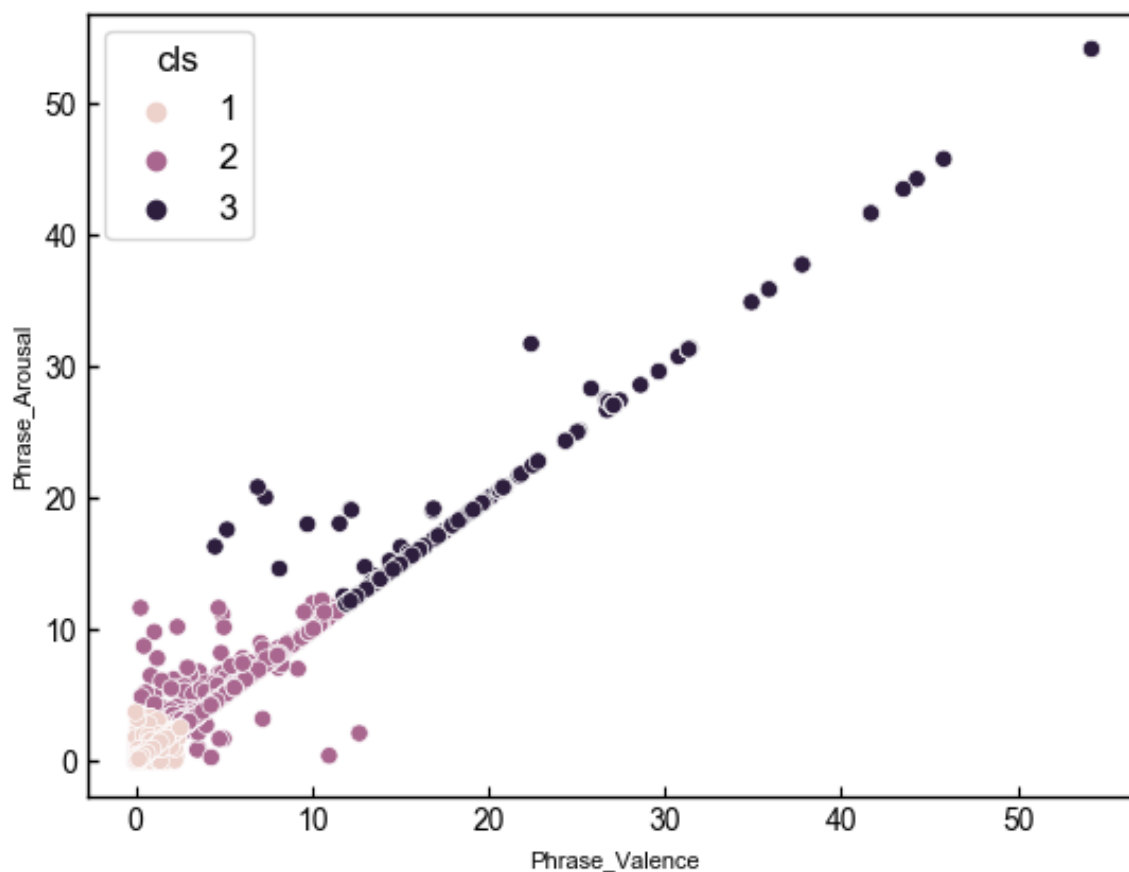


図 3.8: フレーズ A+V+ 平面

第1象限にプロットされたフレーズは全部で35935句である．そのうちワード法によって分割されたフレーズは第1グループが33,357句，第2グループが2,180句，第3グループが398句であった．後述する実験に用いるフレーズのデータはそれぞれのグループからランダムで1句選出する．第1グループからは「ひとりひとり違う世界みていくんだ」，第2グループからは「一石投じる生きる力」，第3グループからは「バタフライして」の3句を選出した．

第2象限のプロット図は図3.9である。

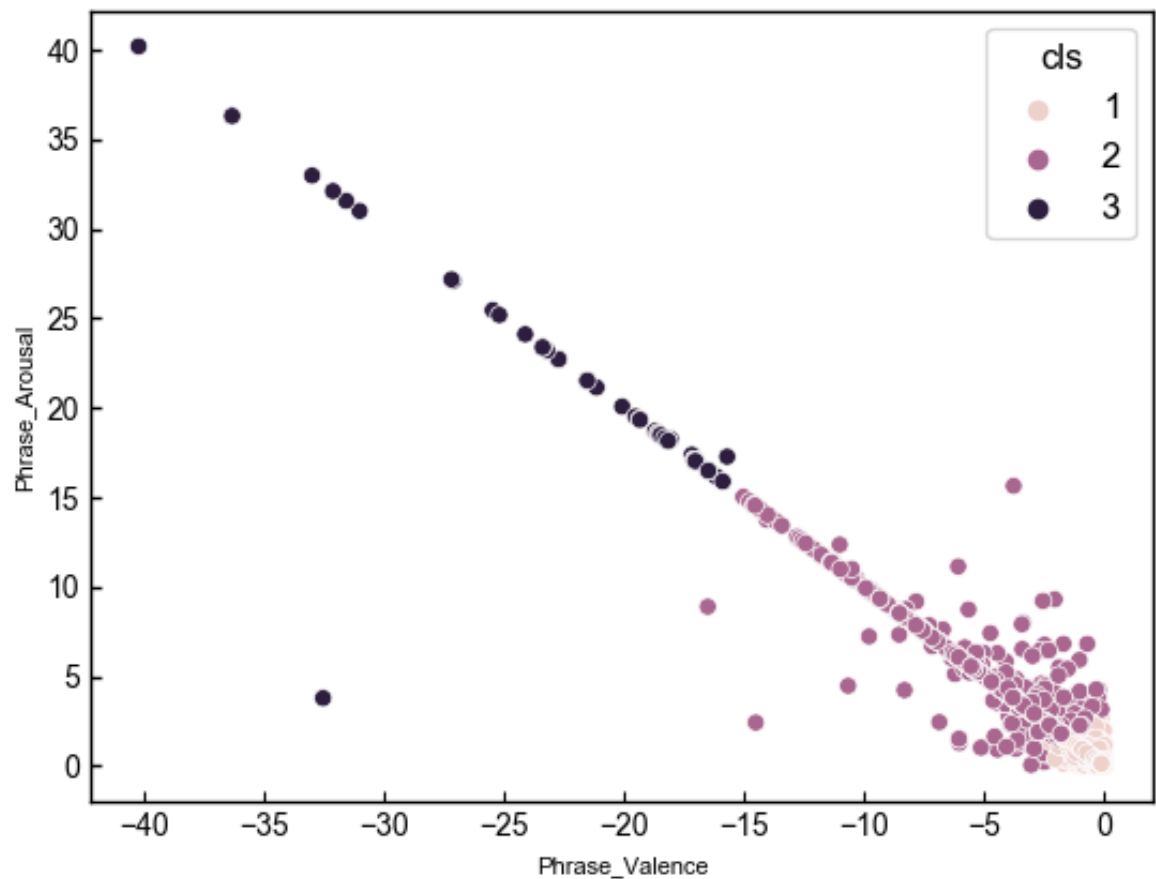


図 3.9: フレーズ A+V-平面

第2象限にプロットされたフレーズは全部で27658句である。そのうちワード法によって分割されたフレーズは第1グループが25,797句、第2グループが1,779句、第3グループが82句であった。後述する実験に用いるフレーズのデータはそれぞれのグループからランダムで1句選出する。第1グループからは「演じる意味はどこもブレたまま」、第2グループからは椎名林檎の「二つ並んだ影法師の手」、第3グループからはDREAM COME TRUEの「マイクホン争奪戦」の3句を選出した。

第3象限のプロット図は図 3.10 である。

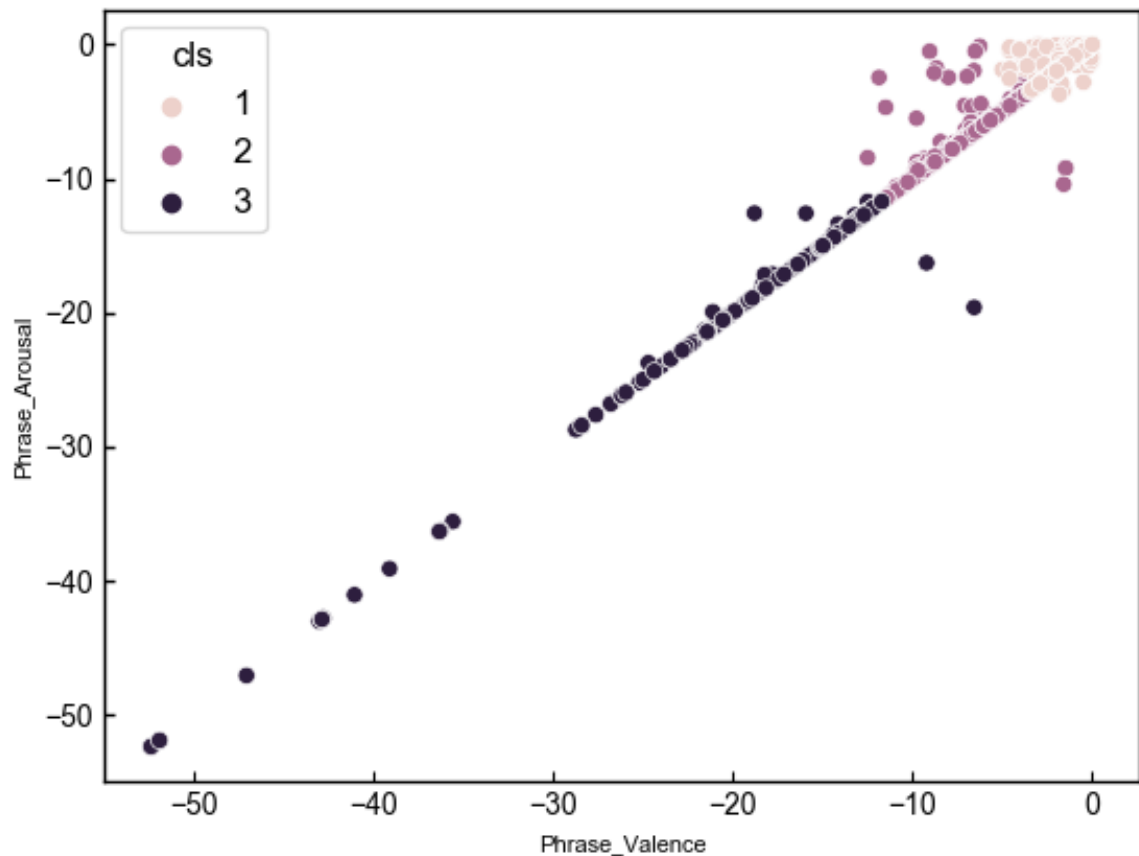


図 3.10: フレーズ A-V-平面

第3象限にプロットされたフレーズは全部で14737句である．そのうちワード法によって分割されたフレーズは第1グループが13,821句，第2グループが694句，第3グループが222句であった．後述する実験に用いるフレーズのデータはそれぞれのグループからランダムで1句選出する．第1グループからは「あたしを捨てたあなたは馬鹿で」，第2グループからは「あなたを苦しめたことでしょう」，第3グループからは「昨今のあなたは鼻につくわ」の3句を選出した．

第4象限のプロット図は図3.11である。

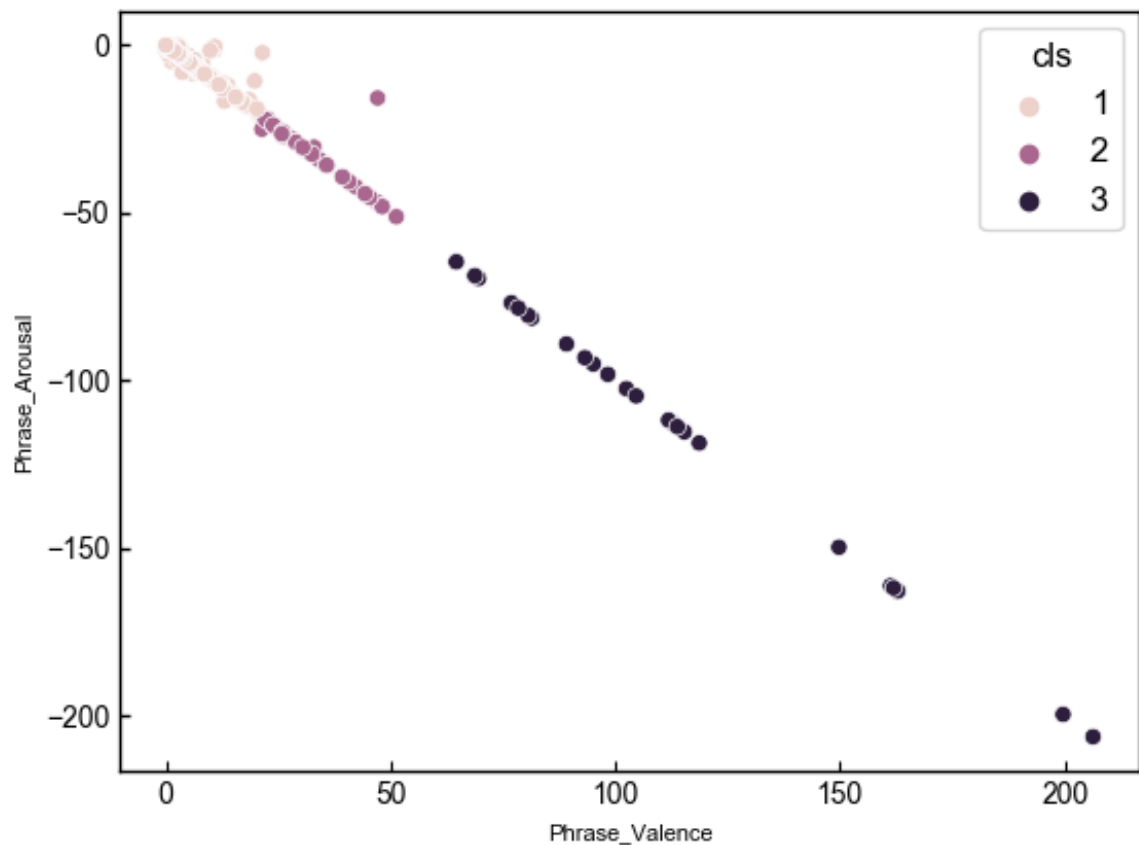


図 3.11: フレーズ A-V+ 平面

第4象限にプロットされたフレーズは全部で60,000句以上である。第4象限にプロットされたフレーズ数は多かったのを、重複削除をした後にランダムで30,000句を取り出した。そのうちワード法によって分割されたフレーズは第1グループが29,907句、第2グループが64句、第3グループが29句であった。後述する実験に用いるフレーズのデータはそれぞれのグループからランダムで1句選出する。第1グループからは「浮かぶ月にやさしく響いては消えてった」、第2グループからは「染まる真っ赤なギロチン」、第3グループからは「しぶといトラウマも全部払おう」の3句を選出した。

第4章

実験

本章では3章で述べた印象推定方法の妥当性について評価するためにおこなった実験について述べる。

4.1 目的

本節では本研究で行った実験の目的について述べる。この実験の目的は3章で述べた印象推定方法によって推定された歌詞と歌詞中に出現するフレーズの印象が妥当であるか効果を評価する。本研究で推定した歌詞と歌詞中に出現するフレーズの印象の妥当性の妥当性の検証をするために2つの実験をした。1つ目は歌詞に対して推定された印象の妥当性について評価する。2つ目は歌詞中に出現するフレーズに対して推定された印象の妥当性について評価する。

4.2 方法

本節では本研究で行った実験の方法について述べる。Google Formでアンケートを作成しデータを収集した。実験参加者は10代後半から20代前半の大学生男女11名である。男女の内訳は男子7名、女子4名である。11名は全員2つの実験に参加した。実験参加者には実験前に3章で述べた印象の定義について十分に説明をした。

- ポジティブに感じるかネガティブに感じるか。
- 興奮を感じるか弛緩を感じるか。
- 喜怒哀楽いずれの感性を感じたか。

4.3 結果

本節では本研究で行った2つの実験の結果について述べる。

4.3.1 歌詞の印象評価

歌詞の印象評価結果について述べる。喜怒哀楽の4クラスごとに詳説する。

喜クラス

喜の印象クラスにクラスタリングされた歌詞の評価結果について述べる。第1グループから選出した「光の誓いが聴こえた日」、第2グループから選出した「冬三味にはまだ遠い」、第3グループから選出した「積木遊び」の印象評価結果は以下の図4.1, 4.2, 4.3である。

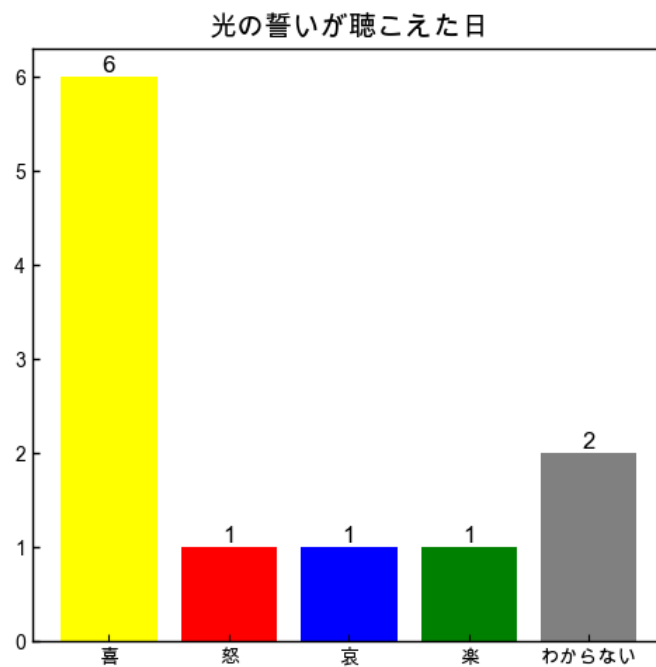


図 4.1: 歌詞 A+V+ 平面 第 1 グループ

第1グループから選出した「光の誓いが聴こえた日」の歌詞から実験参加者6名が喜のクラスに感じたと回答した。怒，哀，楽のクラスを感じたと回答した人数はそれぞれ1名ずつであり，わからないと回答した人数は2名であった。実験参加者の半分以上が推定した印象を感じたことから推定結果は妥当であるといえる。

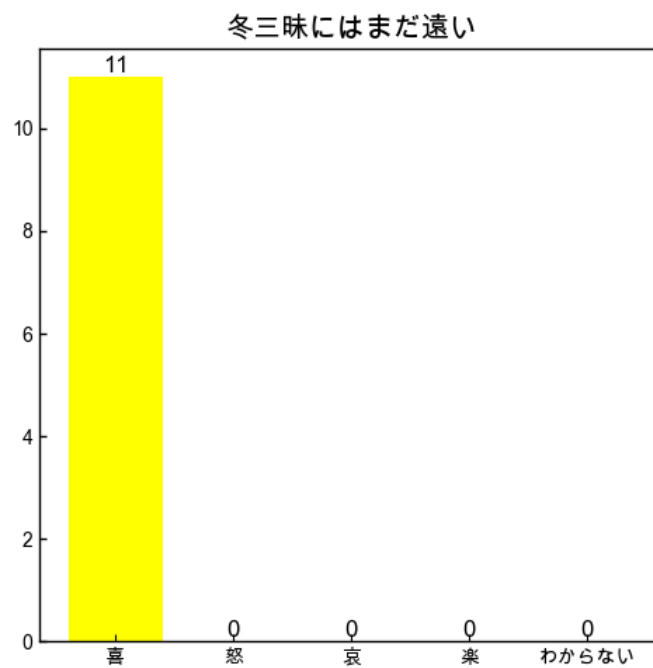


図 4.2: 歌詞 A+V+ 平面 第 2 グループ

第2グループから選出した「冬三昧にはまだ遠い」の歌詞から実験参加者11名全員が喜のクラスに感じたと回答した。したがって推定結果は妥当であるといえる。

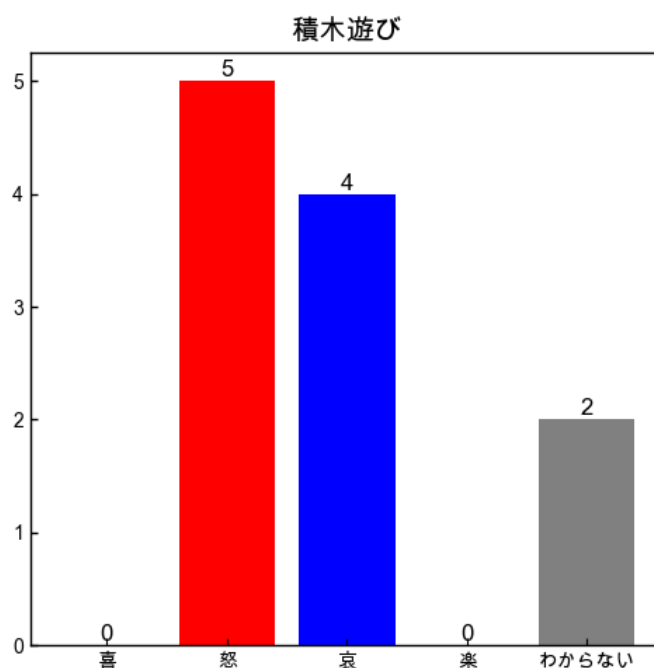


図 4.3: 歌詞 A+V+ 平面 第 3 グループ

第 3 グループから選出した「積み木遊び」の歌詞から実験参加者 0 名が喜のクラスに感じたとは回答した。怒のクラスを感じたとは回答した人数は 5 名、哀のクラスを感じたとは回答した人数は 4 名、楽のクラスと感じたとは回答した人はおらず、わからないとは回答した人数は 2 名であった。実験参加者が推定した印象を感じ取れなかったことから推定結果は不当であるといえる。

怒クラス

怒の印象クラスにクラスタリングされた歌詞の評価結果について述べる．第1グループから選出した「TEENAGE RIOT」，第2グループから選出した「透明人間 18号」，第3グループから選出した「Da La Da Da」の印象評価結果は以下の図 4.4, 4.5, 4.6 である．

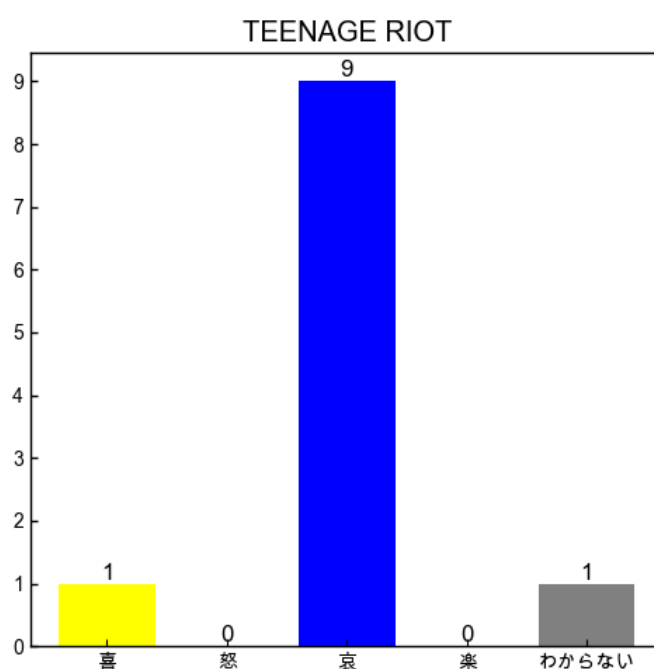


図 4.4: 歌詞 A+V-平面 第1グループ

第1グループから選出した「TEENAGE RIOT」の歌詞から実験参加者0名が怒のクラスに感じたとは回答した．喜のクラスを感じたとは回答した人数は1名，哀のクラスを感じたとは回答した人数は9名，楽のクラスと感じたとは回答した人はおらず，わからないと回答した人数は1名であった．実験参加者が推定した印象を感じ取れなかったことから推定結果は不当であるといえる．

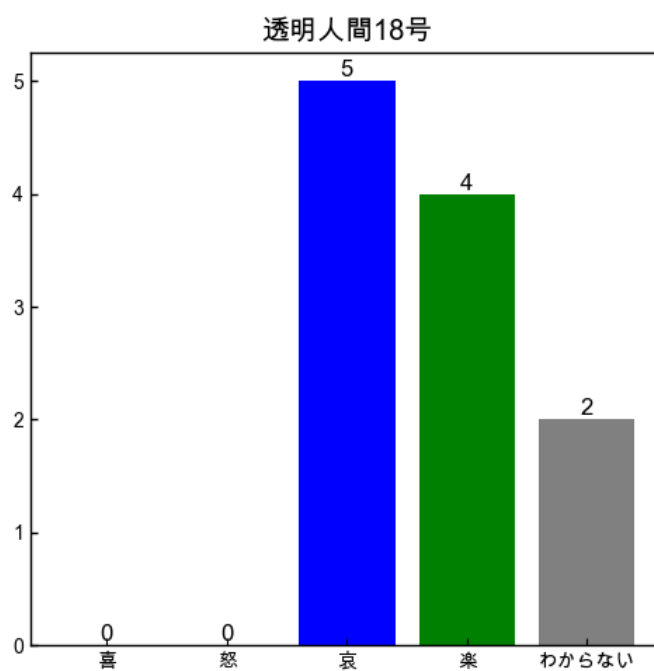


図 4.5: 歌詞 A+V-平面 第 2 グループ

第2グループから選出した「透明人間18号」の歌詞から実験参加者0名が怒のクラスに感じたとは回答した。喜のクラスに感じたとは回答した人はおらず、哀のクラスに感じたとは回答した人数は5名、楽のクラスに感じたとは回答した人数は4名、わからないとは回答した人数は2名であった。実験参加者が推定した印象を感じ取れなかったことから推定結果は不当であるといえる。

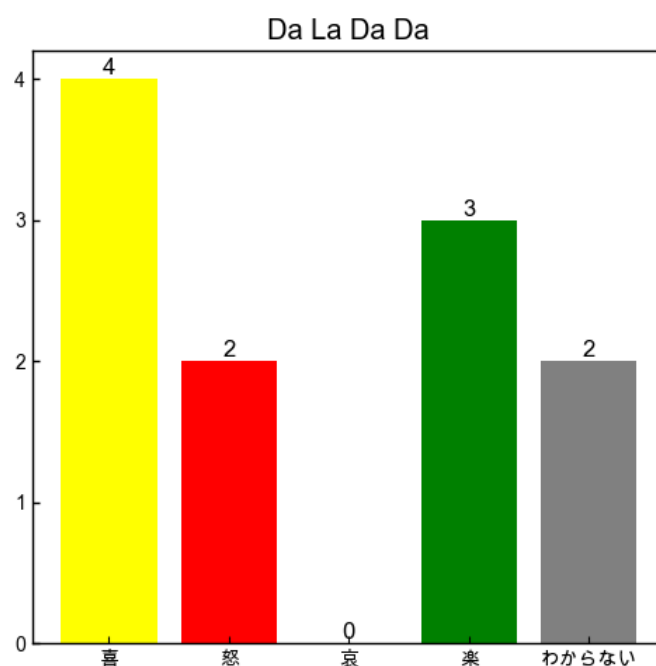


図 4.6: 歌詞 A+V-平面 第 3 グループ

第 3 グループから選出した「Da La Da Da」の歌詞から実験参加者 2 名が怒のクラスに感じたとは回答した。喜のクラスに感じたとは回答した人数は 4 名、哀のクラスに感じたとは回答した人はおらず、楽のクラスに感じたとは回答した人数は 3 名、わからないとは回答した人数は 2 名であった。実験参加者の多くが推定した印象を感じ取れなかったことから推定結果は不当であるといえる。

哀クラス

哀の印象クラスにクラスタリングされた歌詞の評価結果について述べる．第1グループから選出した「Just be you」，第2グループから選出した「コイン」，第3グループから選出した「分別奮闘記」の印象評価結果は以下の図4.7, 4.8, 4.9である．

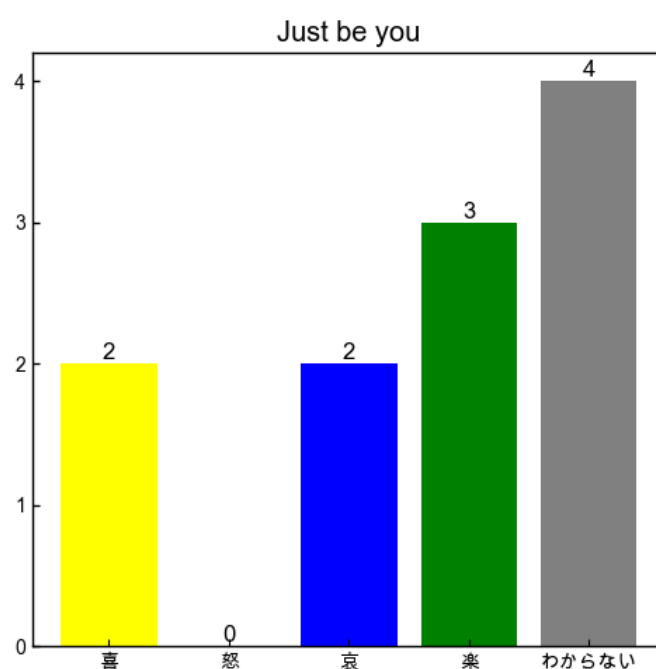


図 4.7: 歌詞 A-V-平面 第1グループ

第1グループから選出した「Just be you」の歌詞から実験参加者2名が哀のクラスに感じたとは回答した．喜のクラスを感じたとは回答した人数は2名，怒のクラスを感じたとは回答した人はおらず，楽のクラスと感じたを回答した人数は3名，わからないと回答した人数は4名であった．実験参加者の多くが推定した印象を感じ取れなかったことから推定結果は不当であるといえる．

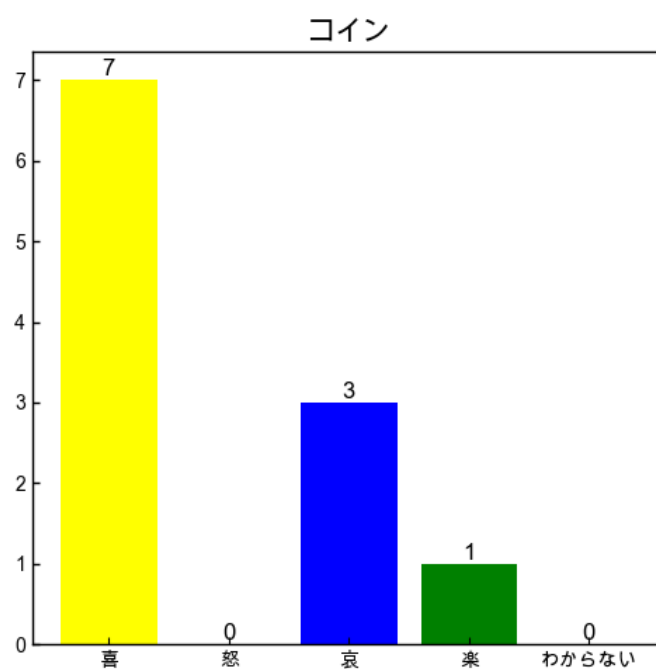


図 4.8: 歌詞 A-V-平面 第 2 グループ

第 2 グループから選出した「コイン」の歌詞から実験参加者 3 名が哀のクラスに感じたとは回答した。喜のクラスを感じたとは回答した人数は 7 名，怒のクラスを感じたとは回答した人はおらず，楽のクラスを感じたとは回答した人数は 1 名，わからないとは回答した人はいなかった。実験参加者の多くが推定した印象を感じ取れなかったことから推定結果は不当であるといえる。

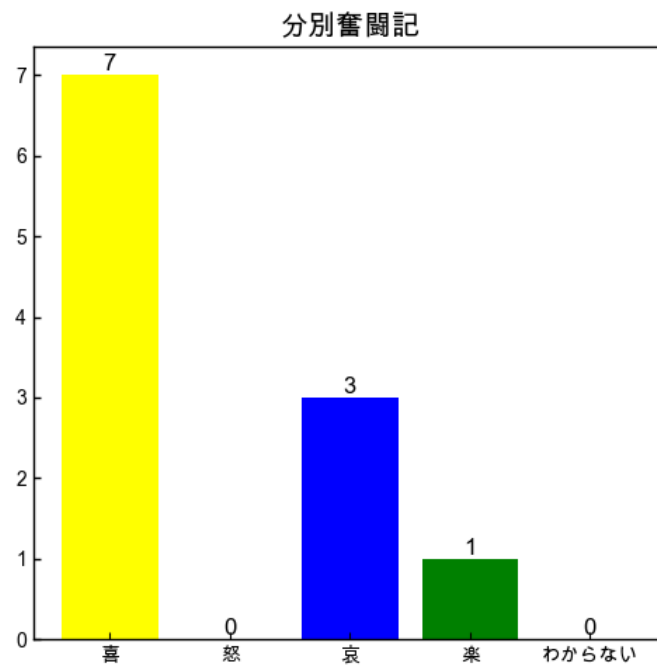


図 4.9: 歌詞 A-V-平面 第 3 グループ

第 3 グループから選出した「分別奮闘記」の実験結果は「コイン」の結果と同じであった。したがって、実験参加者の多くが推定した印象を感じ取れなかったことから推定結果は不当であるといえる。

楽クラス

楽の印象クラスにクラスタリングされた歌詞の評価結果について述べる．第1グループから選出した「若さ故エンテレケイア」，第2グループから選出した「お互い様やん」，第3グループから選出した「シロクマ」の印象評価結果は以下の図 4.10, 4.11, 4.12 である．

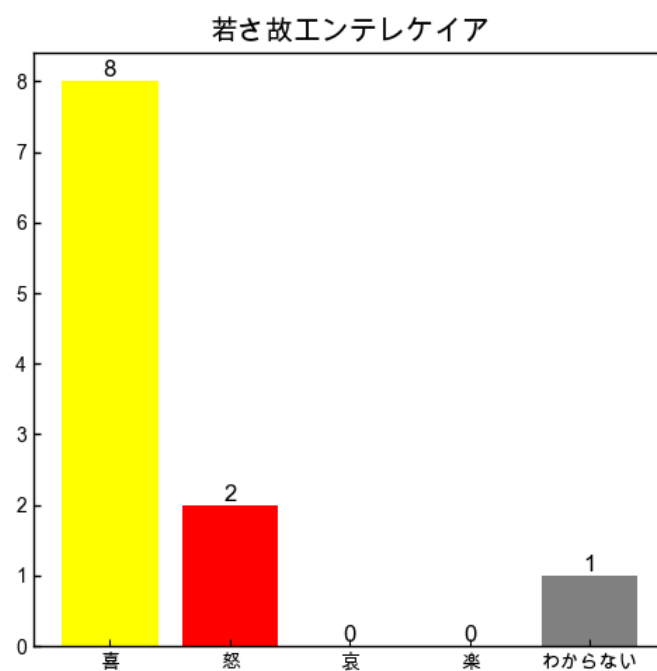


図 4.10: 歌詞 A-V+ 平面 第1グループ

第1グループから選出した「若さ故エンテレケイア」の歌詞から実験参加者0名が楽のクラスに感じたとは回答した．喜のクラスを感じたとは回答した人数は8名，怒のクラスを感じたとは回答した人数は2名，哀のクラスを感じたとは回答した人はおらず，わからないとは回答した人数は1名であった．実験参加者が推定した印象を感じ取れなかったことから推定結果は不当であるといえる．

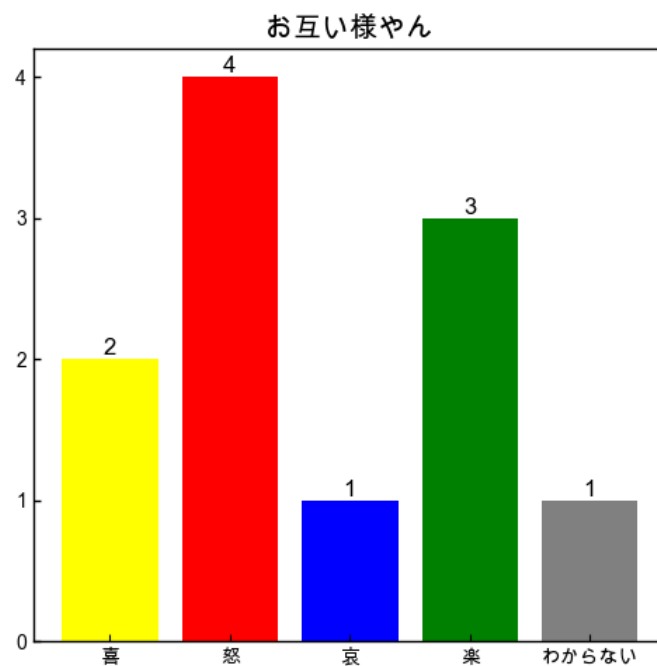


図 4.11: 歌詞 A-V+ 平面 第 2 グループ

第2グループから選出した「お互い様やん」の歌詞から実験参加者3名が楽のクラスに感じたとは回答した。喜のクラスに感じたとは回答した人数は2名，怒のクラスに感じたとは回答した人数は4名，哀のクラスに感じたとは回答した人数は1名，わからないと回答した人数は1名であった。実験参加者の多くが推定した印象を感じ取れなかったことから推定結果は不当であるといえる。

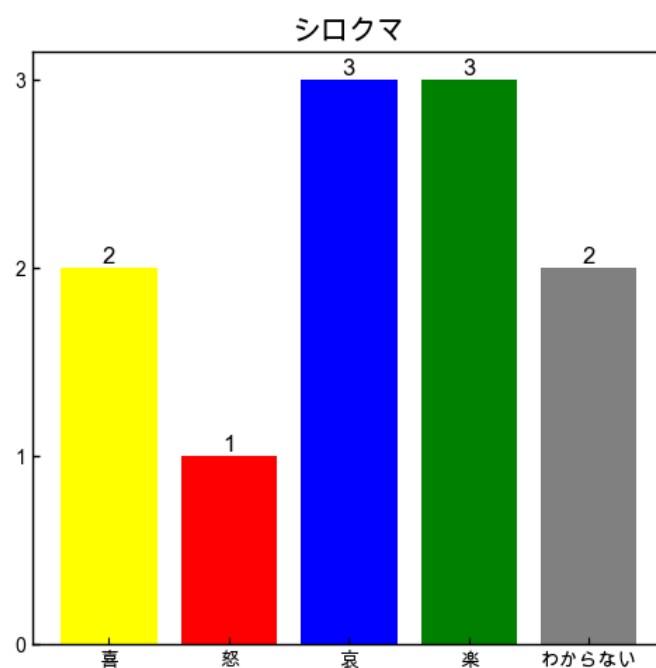


図 4.12: 歌詞 A-V+ 平面 第 3 グループ

第3グループから選出した「シロクマ」の歌詞から実験参加者3名が楽のクラスに感じたとは回答した。喜のクラスを感じたとは回答した人数は2名，怒のクラスを感じたとは回答した人数は1名，哀のクラスを感じたとは回答した人数は3名，わからないとは回答した人数は2名であった。実験参加者の多くが推定した印象を感じ取れなかったことから推定結果は不当であるといえる。

4.3.2 フレーズの印象評価

フレーズの印象評価結果について述べる．喜怒哀楽の4クラスごとに詳説する．

喜クラス

喜の印象クラスにクラスタリングされた歌詞の評価結果について述べる．第1グループから選出した「ひとりひとり違う世界をみていくんだ」，第2グループから選出した「一石投じる生きる力」，第3グループから選出した「バタフライして」の印象評価結果は以下の図4.13, 4.14, 4.15である．

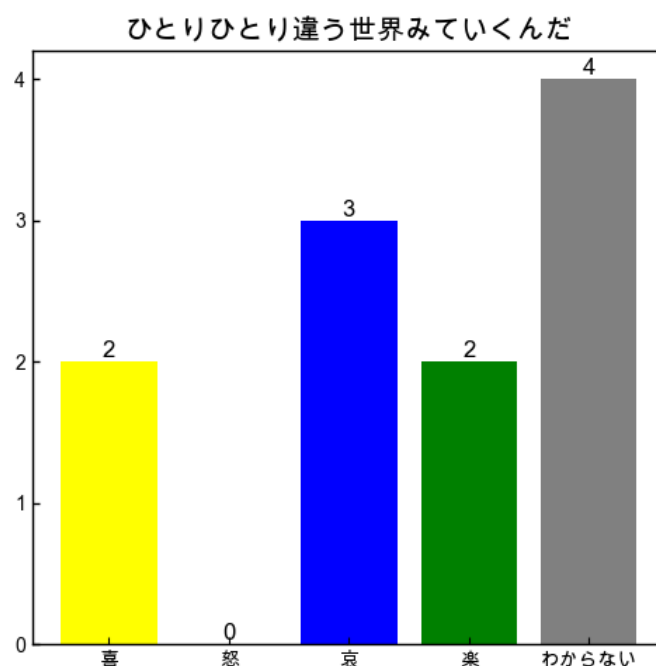


図 4.13: フレーズ A+V+ 平面 第1グループ

第1グループから選出した「ひとりひとり違う世界をみていくんだ」のフレーズから実験参加者2名が喜のクラスに感じたとは回答した．怒のクラスを感じたとは回答した人はおらず，哀のクラスを感じたとは回答した人数は3名，楽のクラスを感じたとは回答した人数は2名，わからないとは回答した人数は4名であった．実験参加者の多くが推定した印象を感じなかったことから推定結果は不当である

といえる。

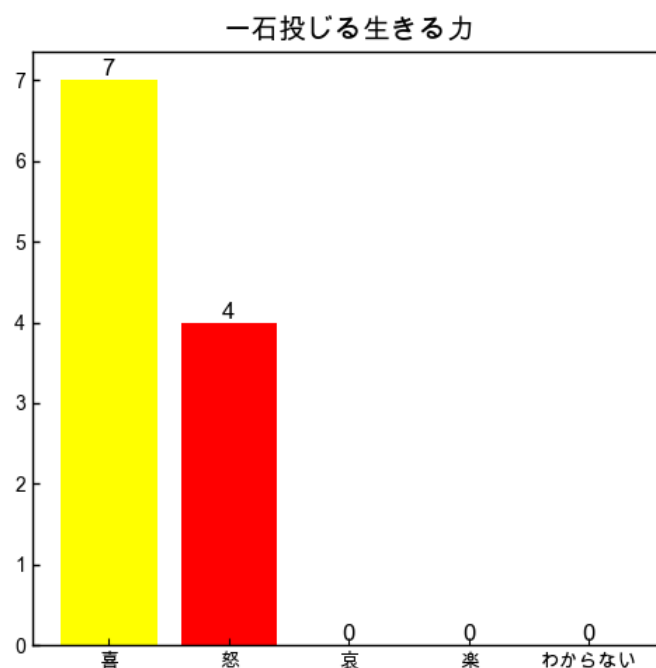


図 4.14: フレーズ A+V+ 平面 第 2 グループ

第 2 グループから選出した「一石投じる生きる力」のフレーズから実験参加者 7 名が喜のクラスに感じたと回答した。怒のクラスを感じたと回答した人数は 4 名、哀と楽のクラスを感じたまたはわからないと回答した人はいなかった。実験参加者の多くが推定した印象を感じたことから推定結果は妥当であるといえる。

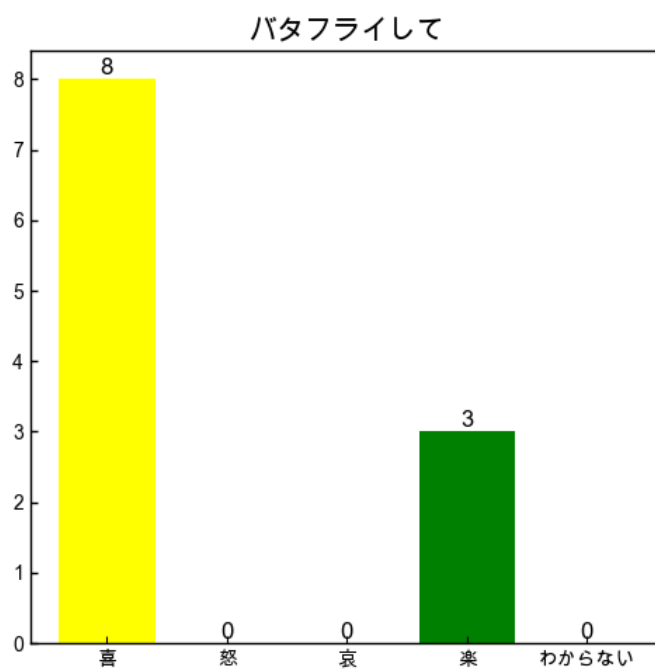


図 4.15: フレーズ A+V+ 平面 第 3 グループ

第 3 グループから選出した「バタフライして」のフレーズから実験参加者 8 名が喜のクラスに感じたと回答した。楽のクラスを感じたと回答した人数は 3 名，哀と怒のクラスを感じたまたはわからないと回答した人はいなかった。実験参加者の多くが推定した印象を感じたことから推定結果は妥当であるといえる。

怒クラス

怒の印象クラスにクラスタリングされた歌詞の評価結果について述べる．第1グループから選出した「演じる意味はどこもブレたまま」，第2グループから選出した「二つ並んだ影法師の手」，第3グループから選出した「マイクホン争奪戦」の印象評価結果は以下の図4.16, 4.17, 4.18である．

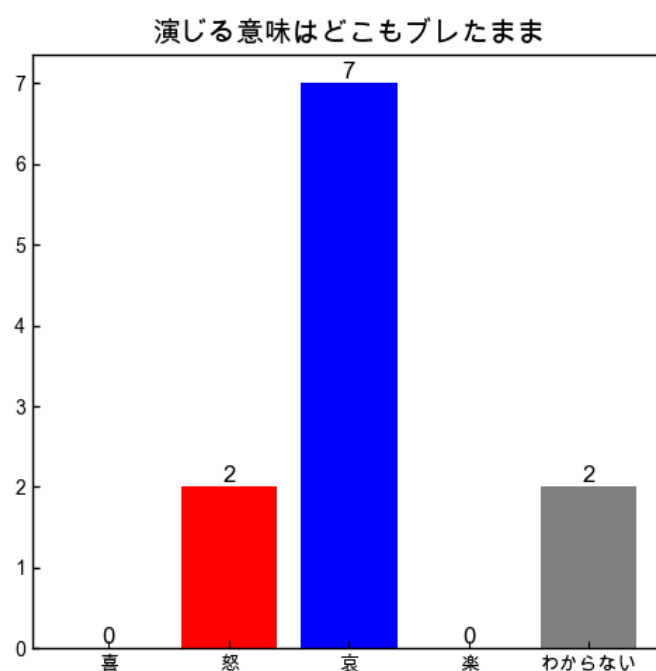


図 4.16: フレーズ A+V-平面 第1グループ

第1グループから選出した「演じる意味はどこもブレたまま」のフレーズから実験参加者2名が怒のクラスに感じたとは回答した．哀のクラスを感じたとは回答した人数は7名，喜と楽のクラスを感じたおらず，わからないと回答した人数は2名であった．実験参加者の多くが推定した印象を感じていないことから推定結果は不当であるといえる．

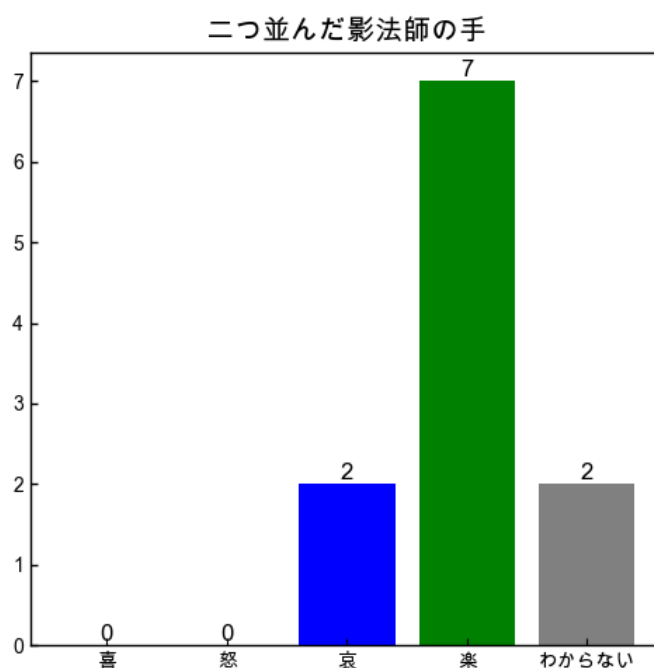


図 4.17: フレーズ A+V-平面 第 2 グループ

第 2 グループから選出した「二つ並んだ影法師の手」のフレーズから実験参加者 0 名が怒のクラスに感じたとは回答した。喜のクラスを感じた人はおらず、哀のクラスを感じたと回答した人数は 2 名、楽のクラスを感じたと回答した人数は 7 名、わからないと回答した人数は 2 名であった。実験参加者の多くが推定した印象を感じていないことから推定結果は不当であるといえる。

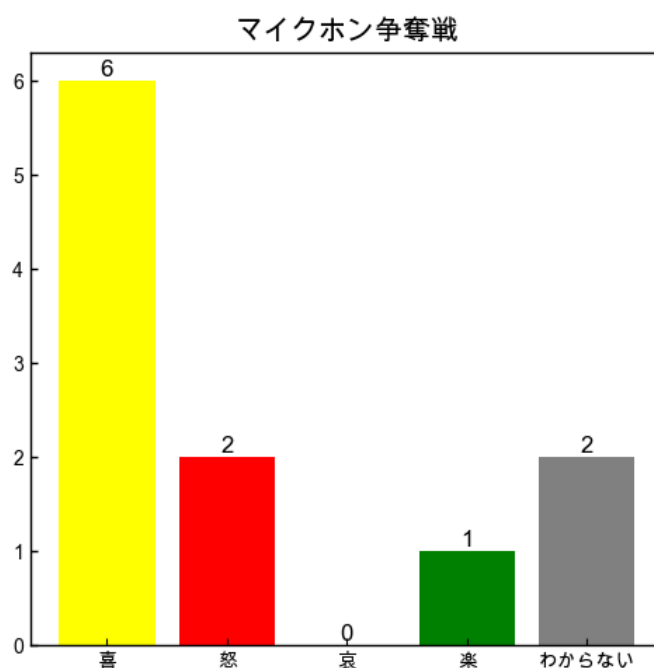


図 4.18: フレーズ A+V-平面 第3グループ

第3グループから選出した「マイクホン争奪戦」のフレーズから実験参加者2名が怒のクラスに感じたとは回答した。喜のクラスを感じた人数は6名、哀のクラスを感じたとは回答した人はおらず、楽のクラスを感じたとは回答した人数は1名、わからないと回答した人数は2名であった。実験参加者の多くが推定した印象を感じていないことから推定結果は不当であるといえる。

哀クラス

哀の印象クラスにクラスタリングされた歌詞の評価結果について述べる．第1グループから選出した「私を捨てたあなたは馬鹿で」，第2グループから選出した「あなたを苦しめたことでしょう」，第3グループから選出した「昨今のあなたは鼻につくわ」の印象評価結果は以下の図4.19, 4.20, 4.21である．

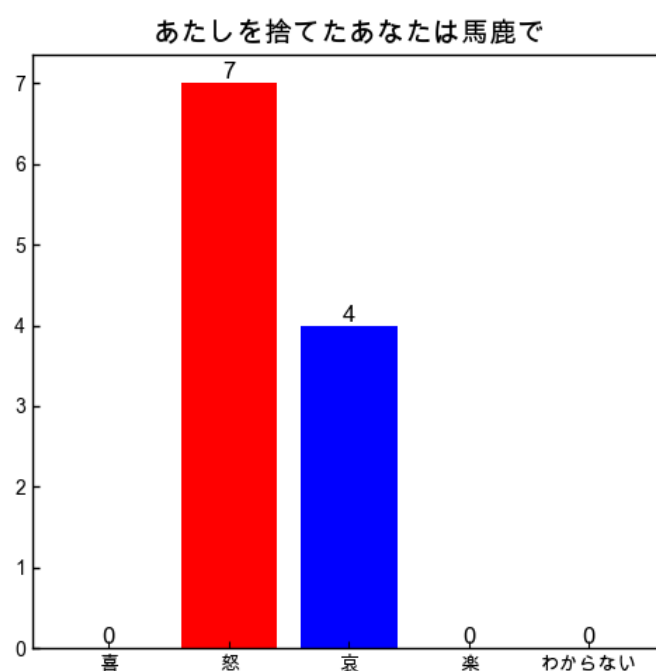


図 4.19: フレーズ A-V-平面 第1グループ

第1グループから選出した「私を捨てたあなたは馬鹿で」のフレーズから実験参加者4名が哀のクラスに感じたとは回答した．怒のクラスを感じたとは回答した人数は7名，喜と楽のクラスを感じた，またはわからないと回答した人はいなかった．実験参加者の多くが推定した印象を感じていないことから推定結果は不当であるといえる．

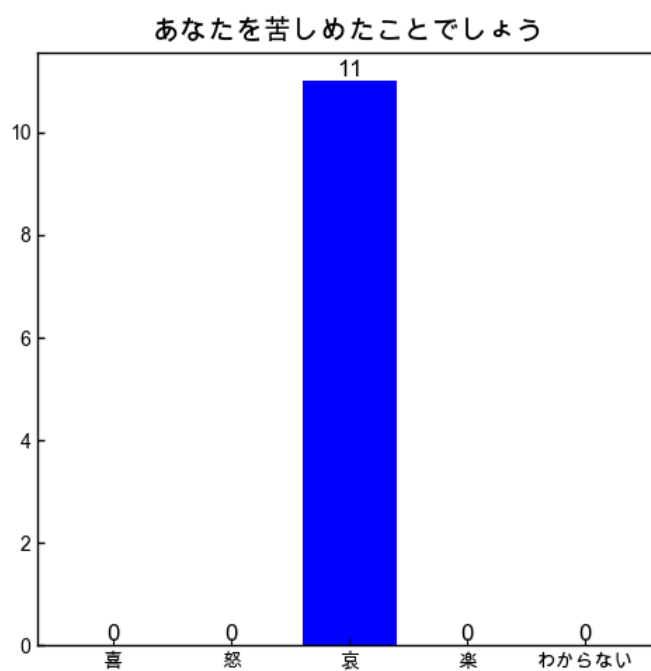


図 4.20: フレーズ A-V-平面 第 2 グループ

第 2 グループから選出した「あなたを苦しめたことでしょう」のフレーズから実験参加者 11 名全員が哀のクラスに感じたこと回答した．実験参加者全員が推定した印象を感じたことから推定結果は妥当であるといえる．

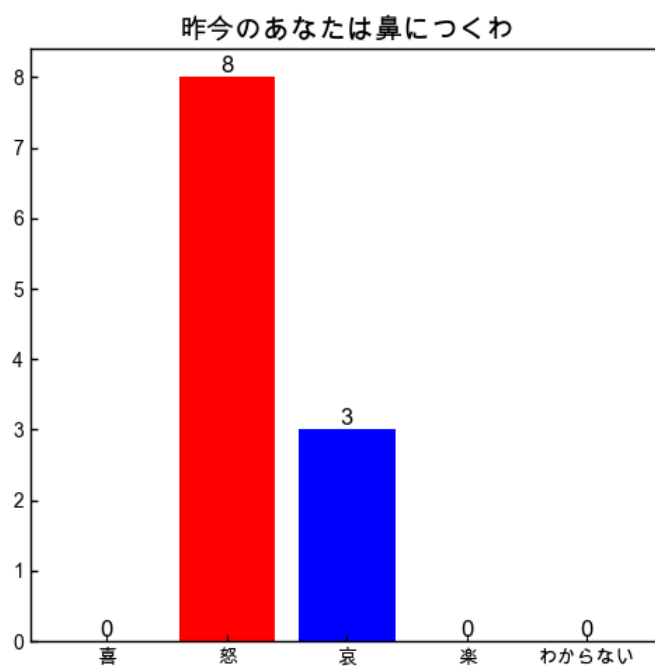


図 4.21: フレーズ A-V-平面 第 3 グループ

第3グループから選出した「昨今のあなたは鼻につくわ」のフレーズから実験参加者3名が哀のクラスに感じたとは回答した。怒のクラスを感じたとは回答した人数は8名、喜と楽のクラスを感じた、またはわからないと回答した人はいなかった。実験参加者の多くが推定した印象を感じていないことから推定結果は不当であるといえる。

楽クラス

楽の印象クラスにクラスタリングされた歌詞の評価結果について述べる。第1グループから選出した「浮かぶ月にやさしく響いては消えていった」、第2グループから選出した「染まる真っ赤なギロチン」、第3グループから選出した「しぶといトラウマも全部払おう」の印象評価結果は以下の図4.22, 4.23, 4.24である。

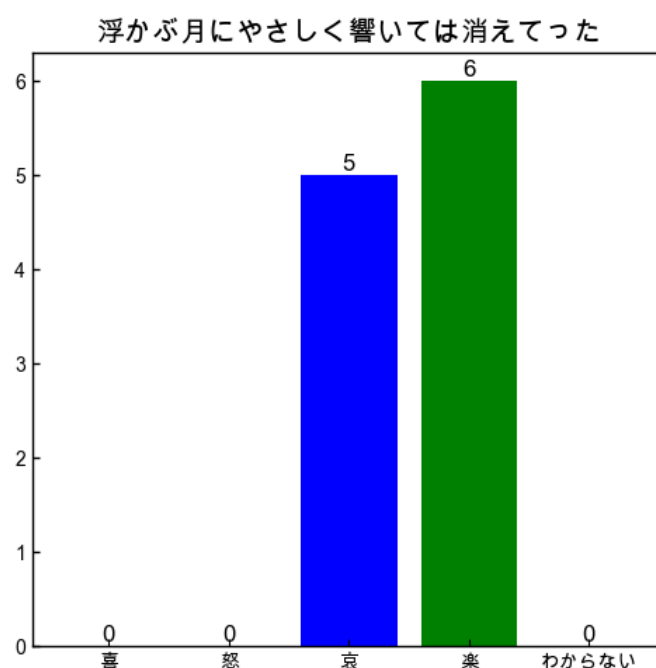


図 4.22: フレーズ A-V+ 平面 第1グループ

第1グループから選出した「浮かぶ月にやさしく響いては消えていった」のフレーズから実験参加者6名が楽のクラスに感じたと回答した。哀のクラスを感じたと回答した人数は5名、喜と怒のクラスを感じた、またはわからないと回答した人はいなかった。実験参加者の多くが推定した印象を感じたことから推定結果は妥当であるといえる。

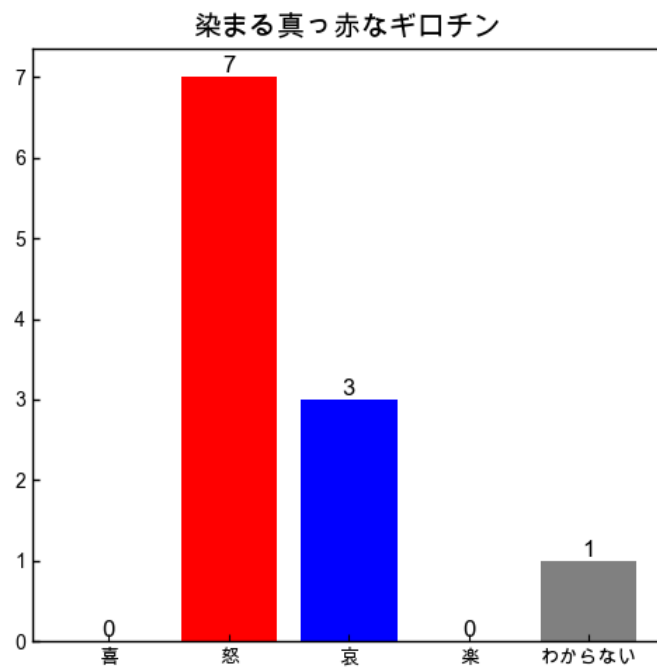


図 4.23: フレーズ A-V+ 平面 第 2 グループ

第 2 グループから選出した「染まる真っ赤なギロチン」のフレーズから実験参加者 0 名が楽のクラスに感じたとは回答した。怒のクラスを感じたとは回答した人数は 5 名、哀のクラスを感じたとは回答した人数は 3 名、喜のクラスを感じた人はおらず、わからないと回答した人は 1 名であった。実験参加者が推定した印象を感じていないことから推定結果は不当であるといえる。

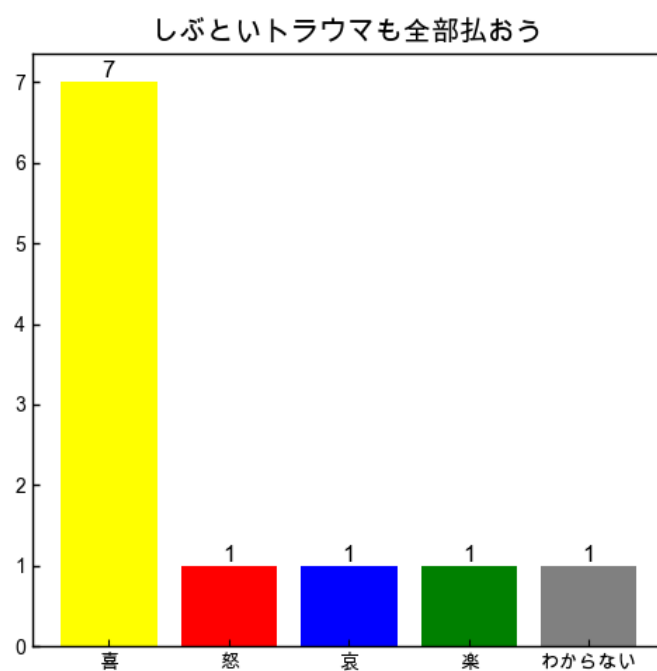


図 4.24: フレーズ A-V+ 平面 第 3 グループ

第 3 グループから選出した「しぶといトラウマも全部払おう」のフレーズから実験参加者 1 名が楽のクラスに感じたとは回答した。喜のクラスを感じたとは回答した人数は 7 名、怒と哀のクラスを感じた、またはわからないと回答した人はそれぞれ 1 名であった。実験参加者が推定した印象を感じていないことから推定結果は不当であるといえる。

4.4 考察

本節では本研究で行った2つの実験の考察について述べる。

4.4.1 歌詞の印象評価実験の考察

前節で述べた結果から喜のクラスに対しては第3グループを除き印象推定は妥当であると言える。しかし、それ以外のクラスに対してはうまく印象推定できたとは言えない。喜のクラスだけ推定がうまくいった理由は日本語版 ANEW 拡張データセットに理由がある。単語データの総数 14231 語の内、5581 語が AV 平面の第1象限に属していた。これは他の印象より 2000 語以上高い数値である。つまり象限によって事前知識が偏っていたため、印象推定の結果にも差が生まれたと考える。また、今回の実験では J-POP の楽曲を多く採用している。POPS として分類される楽曲には喜のクラスの印象を感じさせる楽曲が多い傾向にある。そのため潜在的に喜のクラスに属する楽曲をただしく印象推定する確率は高くなる。

喜のクラスの第3グループにおける印象推定が妥当でなかった理由は、MAP 推定前の $P(w|z)$ の高い単語が歌詞中に存在しているからであると考ええる。MAP 推定前の $P(w|z)$ が高ければ MAP 推定後の $P(w|z)$ も高くなる。そのため MAP 推定後のフレーズ含まれる喜のクラスを感じさせる単語の総数がその他のクラスを感じさせる単語の総数より少なくとも、喜のクラスに分類した。実際に選出された「積木遊び」の曲には内包する 19 フレーズのうち 4 句しか喜のクラスに分類されたフレーズはなく、怒のクラスに分類されたフレーズが 8 句、哀のクラスに分類されたフレーズが 5 句、楽のクラスに分類されたフレーズが 2 句という構成であった。したがって推定手法のアルゴリズムの特性上第3グループにおいて推定結果は妥当でなかったと考える。

4.4.2 フレーズの印象評価実験の考察

前節で述べた結果より第1グループを除く喜のクラスのフレーズと哀のクラスの第2グループ、楽のクラスの第3グループにおいて印象を推定できたとい

える．喜のクラスの印象推定が妥当であった理由は歌詞の印象評価実験の考察と同様に事前知識が多かったからであると考ええる．第1グループだけ妥当でなかった理由は感じとれる印象の強さが弱かったためであるためである．その証拠に多くの実験参加者がわからないわからないと回答していることが挙げられる．

哀のクラスと楽のクラスにおいても推定結果が妥当である結果を確認できたが，全体的に妥当であると評価できたのは12フレーズ中4フレーズでありフレーズにおける印象推定の精度が悪いことを示す．やはりその理由は事前知識が十分でないことからPLSAによって算出したMAP推定前の $P(w|z)$ の多くが無事前分布によるMAP推定するため，MAP推定後の $P(w|z)$ の分布に影響を残してしまうためであると考えられる．

第5章

おわりに

本章では本研究のまとめと今後の課題について述べる。

5.1 まとめ

本研究ではコンテキストウェア楽曲推薦システムの今後の発展のために日本語歌詞の印象推定手法の設計と効果測定をおこなった。従来の歌詞の印象推定手法は未だ発展途上にあり、詳細に印象推定できる手法は未だ存在しない。そのため西川らの設計した印象推定手法に基づき日本語歌詞の印象推定方法を設計した。4章で述べた実験結果より設計した印象推定手法は一部の印象において妥当であることが判明した。しかし、全体的には印象を推定することは叶わなかったと言える。一部の印象を推定できた理由としては事前知識が十分に存在したことが考えられる。

5.2 今後の課題

今後の課題は2点ある。1つ目は事前知識を増やす必要があることである。日本語版 ANEW 拡張データセットは印象においてデータ数に偏りが認められ、印象推定の結果が妥当であると判断された印象の事前知識は不当であると判断された印象よりも2,000語多く存在した。したがって推定結果が不当であった印象における事前知識を増やす必要があると考える。2つ目は歌詞から得られる印象が必ずしも1つではないということを考慮する必要があることである。歌詞全体の印象はフレーズから得られる印象の合計値であると本研究では推測して

実験をおこなった。しかし、歌詞によっては複数の印象を与える歌詞も存在するため、必ずしも1つの印象に推定することは正しいと言えない。そのため、複数の印象を推定するという観点で推定手法を設計し直す必要がある。

参考文献