

利用者の印象に基づく音楽レコメンドサービス ー音楽データベースの感情空間へのマッピングに関する考察ー

韓 語佳[†] 中野 美由紀[‡] 小口 正人[†]

[†]お茶の水女子大学 〒112-8610 東京都文京区大塚 2-1-1

[‡]津田塾大学 〒187-8577 東京都小平市津田町 2-1-1

E-mail: [†]{han.yujia, oguchi}@is.ocha.ac.jp, [‡]miyuki@tsuda.ac.jp

あらまし 本稿では、利用者の印象に基づき適切な音楽の推薦を行うことを目的とし、音楽 DB としての Spotify が提供する音楽の定量的評価値 (Arousal, Valence 等) を対象として、印象語 (楽しい, 悲しい等) と感情空間、音楽の関係について改めて検討する. Spotify は数十万曲の楽曲を提供しているが、全ての音楽を調べることには限界があるため、本研究では中国語及び日本語の曲を選び、Spotify が提供する音楽データの数値により、感情空間上にマッピングを行った. 中国人と日本人にそれぞれアンケート調査を実施し、該当する楽曲と感情空間の関係について尋ねた. また、アンケートの結果に基づき、感情空間における楽曲の位置及び印象語に関しての考察を行った.

キーワード 音楽推薦, 感情空間, 音楽 DB

1. はじめに

音楽は様々な場面で聞かれており、ウェブ上でも様々な音楽を配信するシステム (Google Play Music/YouTube Music, iTunes, Spotify, 等) が構築され、サービスが提供されている. 気分に合わせて楽曲の推薦は、今までのジャンルの楽曲の推薦と異なり、Russell らの感情空間[1,2]を利用した手法であり、近年から提案されている. 先行研究[3,4]では感情空間上に Arousal - Valence (AV) 値を使って音楽 DB をマッピングし、入力情報として好みの画像を選んでもらうことにより、その画像の感情空間上の値を利用し、音楽を推薦するサービスを構築している.

本稿では、利用者の印象に基づき適切な音楽の推薦を行うことを目的とし、音楽 DB としての Spotify が提供する音楽の定量的評価値 (Arousal, Valence 等) を対象に、印象語 (楽しい, 悲しい等) と感情空間、音楽の関係について改めて検討する. 音楽の推薦は、音楽そのものが持つ特徴量 (メロディ, 調, リズム等) を用いた評価は従来から研究されてきたが、音楽の特徴量から人手で情報を付与するためには、何を特徴量とするか等、全曲を調べる処理コストが高い. そのため、対象となる楽曲は多くても千曲程度であり、多くの音楽提供サイトが百万曲以上提供していることを考慮すると、その手法を現実の推薦に利用することは難しく、限界が感じられる [5,7,8]. 既存の音楽提供サイト Spotify は百万曲近くの楽曲を提供しており、その楽曲に関連する曲名、演奏者、作曲者と併せ、Spotify[6] が付与した数十近くの曲の特徴 (Valence, Energy, popularity, tempo 等) が数値として付与され、公開されている.

筆者らは Spotify の楽曲情報 DB を利用することで従

来にない多数の音楽に関する情報から音楽の推薦を検討している. 音楽を利用する場面は様々な事例が考えられるが、介護施設等で入所者の好みに合わせた曲を選ぶ等、状況、時間、季節などを考慮しつつ、自分とは異なる世代の人の好みに合った楽曲推薦を行うことを最終目標としている.

Spotify で紹介されている楽曲は百万曲近くあり、全ての音楽を調べることは限界がある. また、音楽はその楽曲が生み出される文化と密接に関係していることも多い. 例えば、J-POP, K-POP 等は音楽のジャンルとしてしばしば用いられている. 本研究では、中国語の歌詞及び日本語の歌詞の曲を選び、中国人と日本人にそれぞれアンケート調査を行い、該当する楽曲がマッピングされる感情空間の位置の適切さ、および楽曲に関連する印象語の適切さについて尋ねた. この報告では、感情空間における楽曲の位置及び印象語に関するアンケート結果をまとめ、その分析について報告する.

2. 感情空間

Russell らは Arousal と Valence の 2 次元によって感情を表す Arousal-Valence(AV)空間を提案した. Arousal と Valence はそれぞれ[-1, 1]の実数を取る. Arousal は -1 に近づくにつれて落ち着いた(calm)感情を表し、+1 に近づくにつれて興奮した(excite)感情を表す. Valence も同様に -1 に近づくにつれて、ネガティブな感情を表し、+1 に近づくにつれてポジティブな感情を表す.

図 1 では、印象語が AV 空間上にプロットされている例を示す. AV 空間上にプロットされる各印象語の座標はユーザスタディの結果によるものであり、原点からの角度が感情の意味合い、長さが感情の強さを意味する.

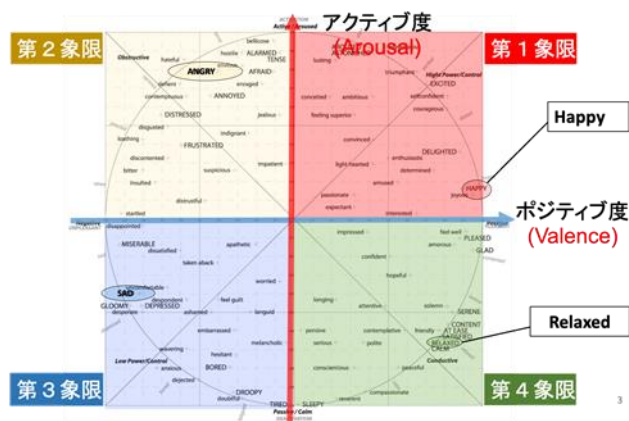


図 1 AV 空間

このことを踏まえ、音楽を Spotify 音楽情報 DB が提供する値を用いて感情空間上へ射影し、対応付け、推薦を行うことを目指す。

3. Spotify のデータ

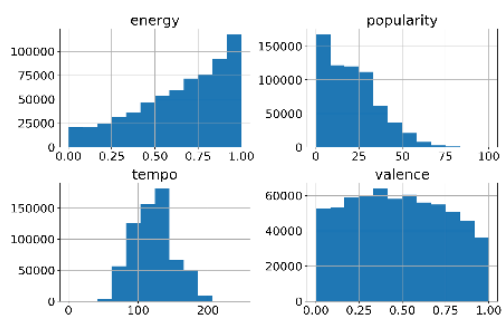


図 2 音楽データベースの楽曲分布

表 1 特徴量と曲の印象の関係

artist_name	name	popularity	energy	tempo	valence
The Beatles	Yesterday - Remastered	65	0.179	96.529	0.315
DREAMS COME TRUE	Miraiyosouzu	22	0.479	131.993	0.476
BABYMETAL	Gimme Chocolate!! - live	32	0.998	109.994	0.181
Ludwig van Beethoven	Symphony No. 7 in A Major, Op. 92: II. Allegretto	44	0.0356	65.048	0.0466
Kyary Pamyu Pamyu	ファッションモンスター	39	0.974	160.009	0.504
Michael Jackson	Beat It	54	0.799	138.825	0.906

ここでは、先行研究[3]の結果（図 2、表 1、2018 年の Spotify 楽曲情報に基づく）を引用した。

Spotify API から取り込んだ約 65 万曲の音楽情報には、数十の特徴量がある。先行研究[3]では、各特徴量のデータ分布、分散などについて全曲を調べた結果、図 2 で示したように「energy」「popularity」「tempo」「valence」は値が分散していて使える可能性が高いと考えている。また、表 1 で示したように、いくつかの良く知られた曲を取り上げ、Energy と Valence の数値が異なると曲調が大きく異なることを確認した。この結果を受け、本稿においても、レコメンドするための主要な特徴量として、valence と energy を用い、各曲

に付された値により感情空間上にマッピングする。Spotify が付与した valence はそのまま感情空間の valence とし、energy の値を arousal の値として利用する。さらに、前節の感情空間では $[-1,1]$ の範囲であるが、Spotify の valence と energy は $[0,1.0]$ の範囲に合わせ、感情空間のマッピングも $[0, 1.0]$ としている。

4. 個別の曲に関するマッピングの調査

4.1 中国語の曲の選択と Spotify 上のデータ数値

表 2 選択した中国曲と対応する印象語

曲名	印象語
● 套马杆/Horse Pole - Namika Tso	light hearted, convinced, enthusiastic
● 两只蝴蝶/Two Butterflies - Pang Long	light hearted, convinced, enthusiastic
● 小苹果/Little Apple - Chopstick Brothers	elated
● 最炫民族风/The Most Dazzling Folk Style - Luo Yan Si	high power/control, excited
● 常回家看看/Often Go Home to Have a Look - Gong Yue	a little active
● 难忘今宵/Can't forget tonight - Gong Yue	feel guilt
● 青花瓷/Blue and White Porcelain - Jay Chou	a little positive, impressed
● 羞答答的玫瑰静悄悄的开/A Shy Rose Is Silently Blooming - Hu Xia	melancholic
● 千里之外/Far Away - Jay Chou	distrustful
● 北京欢迎你/Beijing Welcomes You - Gong Yue	a little low power/control

本節の実験では表 2 に示したように、合計十曲の中国語の曲を選択し、それに沿ってアンケートの調査を行った。中国のお年寄りでも知られている有名な曲を選び、中国人と日本人にそれぞれアンケートを実施した。

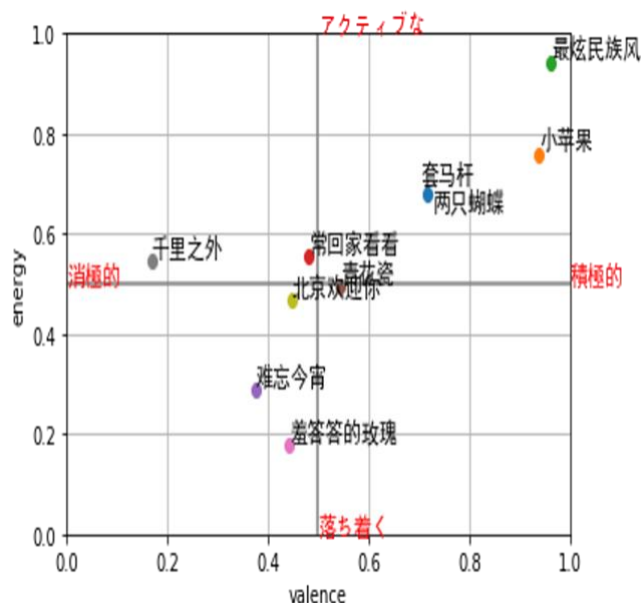


図 3 中国語の曲の感情空間上の位置

図 3 は、Spotify から取り出した中国語の曲のデータで作った散布図である。

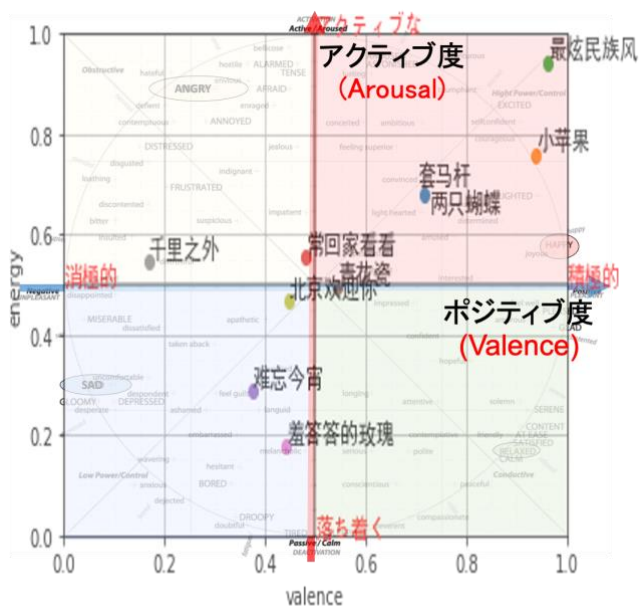


図 4 感情空間上の印象語との関係

図 4 では、その上に、AV 空間に印象語をつけた感情空間を利用した。これで、対応の位置を通じて、曲の印象語を取ることができるとともに、音楽を推薦することもできる。表 2 の右欄は、感情空間から取り出した印象語であり、今回のアンケートに使用した。

4.2 日本語の曲の選択と Spotify 上のデータ数値

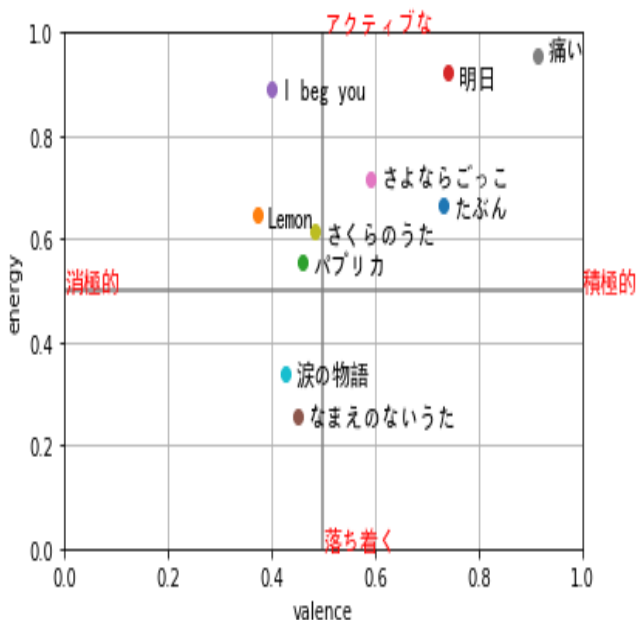


図 5 日本語の曲の感情空間上の位置

表 3 選択した日本曲と対応する印象語

曲名	印象語
● たぶん - YOASOBI	enthusiastic, light hearted
● Lemon - 米津玄師	impatient, suspicious
● パブリカ - 米津玄師	a little active and a little negative
● 明日はきっといい日になる - 高橋優	triumphant
● I beg you - Aimer	afraid
● なまえのないうた - DATEKEN	languid
● さよならごっこ - amazarashi	convinced
● とても痛い痛がりたい - EZFG, TOUYU	high power/control, excited
● さくらのうた - 高橋優	a little active
● 涙の物語 - 有里知花	worried

表 3 は、本節の実験で選択した日本語の曲である。筆者らが好きな十曲を選んだ。同じように、Spotify から取り出したデータで散布図作り、感情空間を利用して対応の位置を通じて、曲の印象語を選択した。表 3 の右欄は、感情空間から取り出した印象語であり、今後の調査に使う予定である。

合わせて見てみると、第 4 象限に散布する曲があまり見られなかった。また、筆者らが好きな曲にはアクティブな傾向があることがわかった。

5. アンケート内容

今回のアンケート調査では、前述の日本語と中国語の曲を用い、中国語のアンケートは Tencent のアンケートサービスを利用し、日本語のアンケートは Google Form を利用した。

中国語の曲については、合計 37 人の中国人 (28 人) と日本人 (9 人) からアンケートの回答が得られた。

日本語の曲については、合計 35 人の中国人 (23 人) と日本人 (12 人) からアンケートの回答が得られた。

中国語の曲と日本語の曲について、曲ごとに Spotify の URL を提示し、曲の試聴を可能とした上で、以下の二つの設問の回答を得た。

- ① 曲が感情空間に散布している位置は合っていると思うか。
- ② 曲が感情空間に対応する印象語は合っていると思うか。

各設問には以下 5 段階の選択肢を設けた。

- ① 合っている
- ② やや合っている
- ③ どちらでもない
- ④ やや合っていない
- ⑤ 合っていない

各曲ごとの回答を得た後に、以下の二つの質問を設定した。

- ① 全体的にこの感情空間に散布している曲は適切

な位置にあると思うか。

- ② 曲とそれに対応する印象語とは合っていると思うか。

6. アンケート結果

中国語と日本語のアンケート結果を用いて、感情空間上の位置の適切さと印象語の適切さについて検討する。

6.1 位置と印象語の適切さに関する検討

十曲全部に関する結果を合算したものを用いた。それぞれの図については後ほど詳しく紹介する。

中国語と日本語の各 10 曲の回答を合計した結果によると、中国語の曲に関して、76.4%以上の人が、曲が感情空間上の位置が「合っている」または「やや合っている」と答えた（図 6）。また、62.2%の人が、曲とそれに対応する印象語とは「合っている」または「やや合っている」と答えた（図 11）。

日本語の曲に関しては、69.1%以上の人が、曲が感情空間に散布している位置が「合っている」または「やや合っている」と答えた（図 14）。また、63.7%の人が、曲とそれに対応する印象語とは「合っている」または「やや合っている」と答えた（図 16）。

今回のアンケート結果について、T 検定を利用して評価した。「合っている」を+2, 「やや合っている」を+1, 「どちらでもない」を0, 「やや合っていない」を-1, 「合っていない」を-2として、T 値を求めた。その結果、T 値は表 4 が示す通りである。

表 4 選択した曲が計算した T 値

	位置の適切さ	印象語の適切さ
中国語の曲	16.348	8.275
日本語の曲	9.176	12.8501

5 % 優位水準となる T 値 2.776（1 % 有意水準の T 値は 4.604）と比較しても、中国語、日本語の位置、印象語のいずれの T 値も大きな値となっている。したがって、音楽配信サービスが提供する音楽情報を利用して楽曲を感情空間にマッピングした結果は十分に利用可能と判断できる。

6.2 中国語の曲について

中国語の曲のアンケート結果について示す。

6.2.1 感情空間上の位置に関する結果

まずは、感情空間上の位置が適切かどうかという問題の結果である。

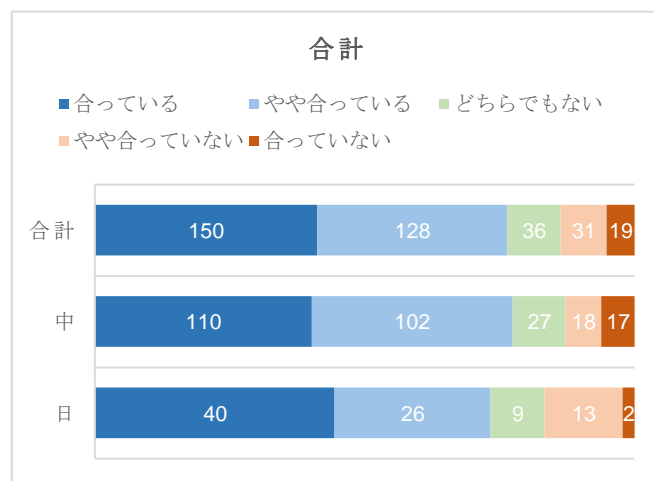


図 6 中国語の十曲合計の位置の適切さ

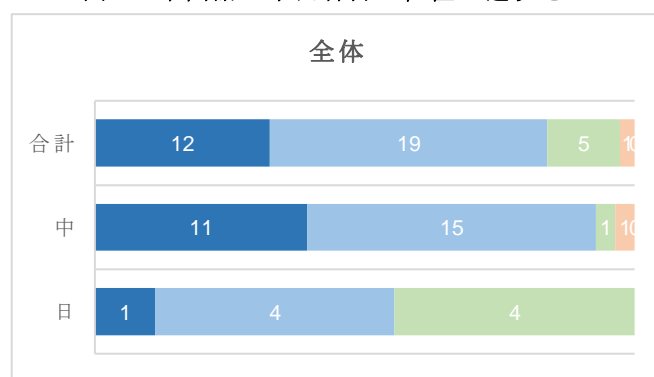


図 7 中国語の曲全体に対する位置の適切さ

図 6 は、十曲の結果を合計したものであり、図 7 は、中国人と日本人が全体的に見て、位置が適切かどうかという最後の質問の結果である。

その中で、中国人と日本人で異なる結果が出た曲がいくつかある。その例を図 8 と図 9 に示す。

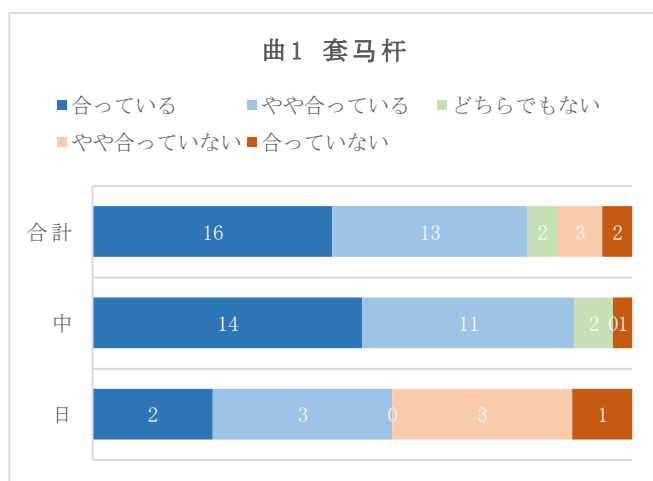


図 8 異なる結果が出た中国語の曲 1

曲 1 は、日本人のデータの数はまだ少ないため、数が増えたら結果が変動する可能性があると考えている。

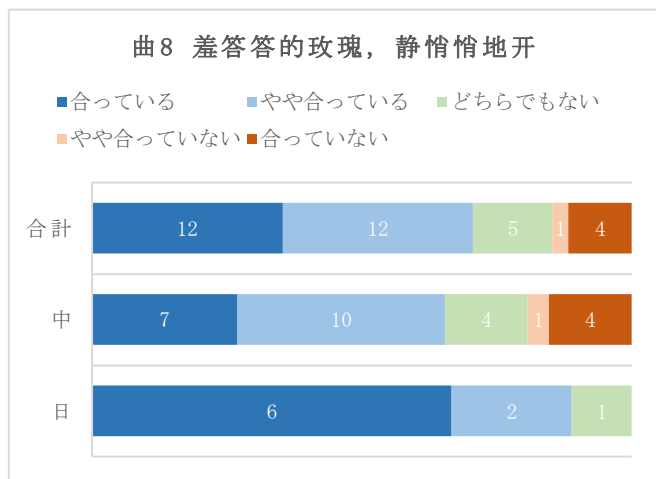


図9 異なる結果が出た中国語の曲8

曲8は、中国人は中国語の歌詞がわかっており、先入観があるため、曲を聞いていなくても答えられると思われる。

6.2.2 感情空間上の印象語に関する結果

次に、印象語が適切かどうかという問題の結果である。

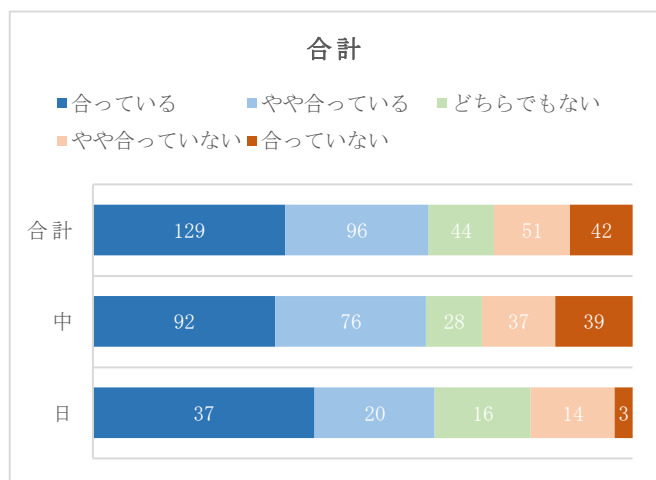


図10 中国語の十曲合計の印象語の適切さ

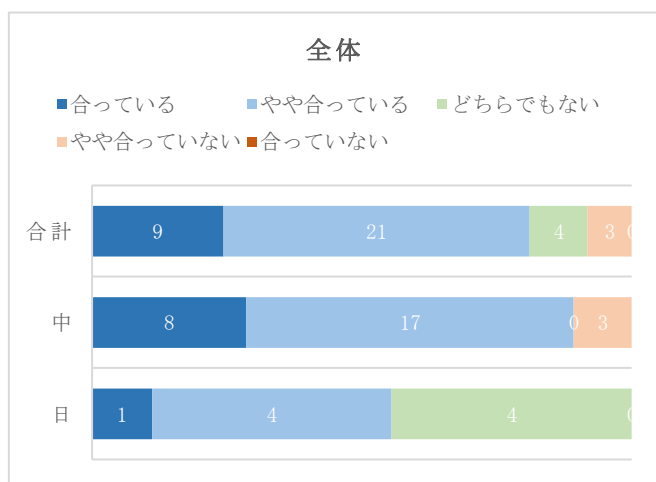


図11 中国語の曲全体に対する印象語の適切さ

図10は、全曲の結果を合計したものであり、図11は、中国人と日本人が全体的に見て、印象語が適切かどうかという最後の質問の結果である。

その中で、中国人と日本人で異なる結果が出た曲がいくつかある。その例を図12に示す。

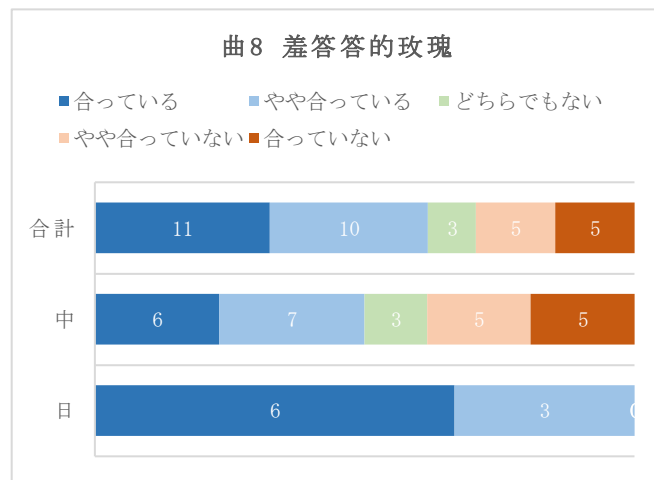


図12 異なる結果が出た中国語の曲8

曲8については、中国人は中国語の曲の歌詞がわかっており、先入観があるため、曲を聞いていなくても答えられると思われる。

そして、中国人と日本人で同様の結果が出た曲がいくつかある。その例を図13に示す。

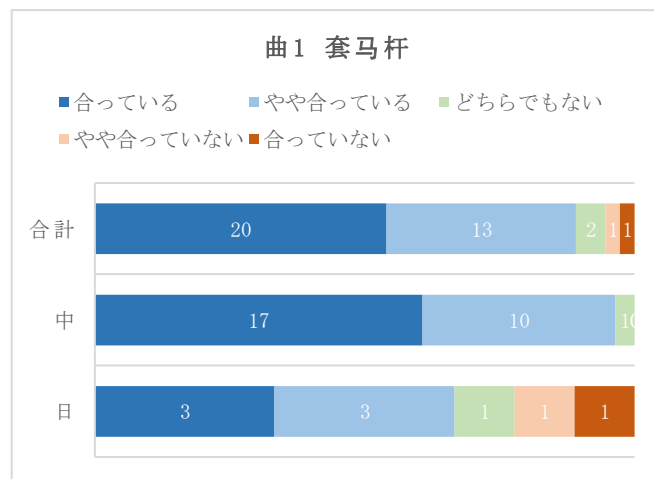


図13 同様の結果が出た中国語の曲1

曲1は第1象限にある曲である。全結果によると、第1象限にある曲は合っていると思う傾向が高いことがわかった。

また、全体的な結果によると、象限の左側にある曲は合っていないと思う傾向が高いこともわかった。先行研究の象限の左側にある曲は合っていないと思う傾向があるという指摘に合致している。

今後は、曲が感情空間にある位置をさらに正確に表示するように、特徴量の選択などを調整したいと考えている。

6.3 日本語の曲について

次に、集計した日本語の曲のデータについて示す。

6.3.1 感情空間上の位置に関する結果

まずは、感情空間上の位置が適切かどうかという問題の結果である。

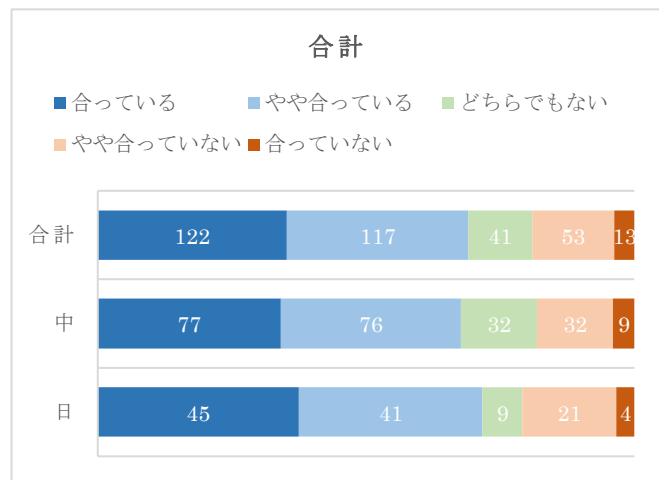


図 14 日本語の十曲合計の位置の適切さ

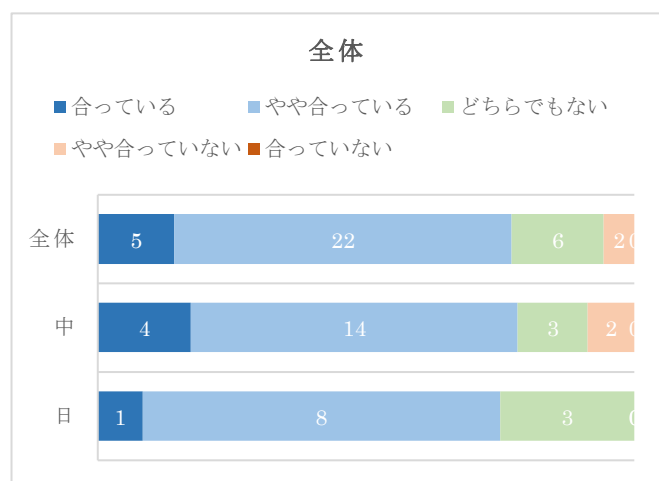


図 15 日本語の曲全体に対する位置の適切さ

図 14 は、全曲の結果を合計したものであり、図 15 は、中国人と日本人が全体的に見て、位置が適切かどうかという最後の質問の結果である。

6.3.2 感情空間上の印象語に関する結果

次は、印象語が適切かどうかという問題の結果である。

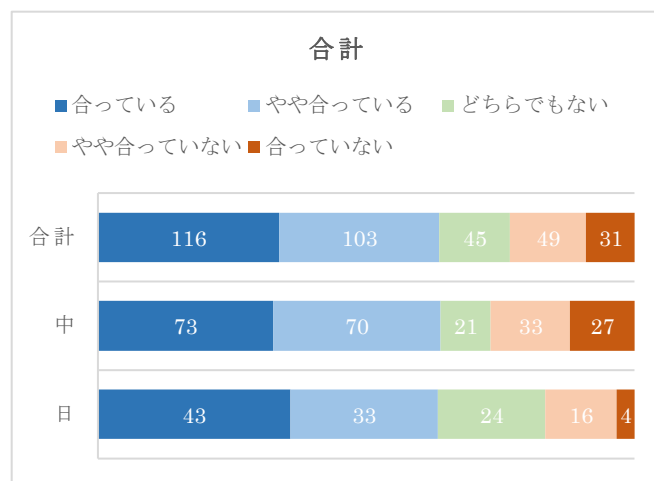


図 16 日本語の十曲合計の印象語の適切さ

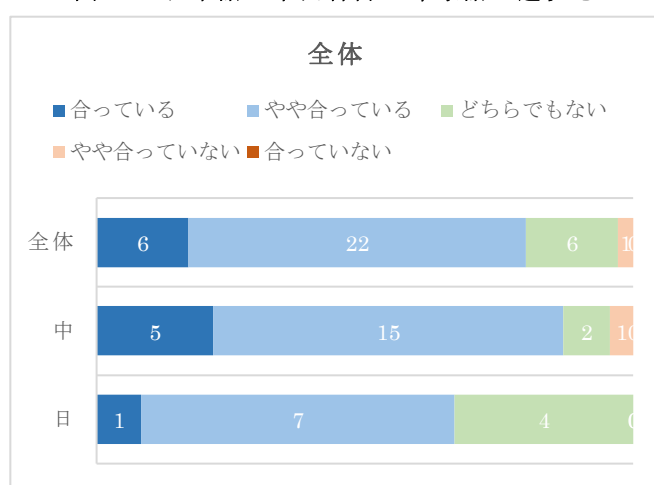


図 17 日本語の曲全体に対する印象語の適切さ

図 16 は、全曲の結果を合計したものであり、図 17 は、中国人と日本人が全体的に見て、印象語が適切かどうかという最後の質問の結果である。

それらの中で、中国人と日本人で異なる結果が出た日本語の曲と、位置より印象語が合うという結果が出た日本語の曲がいくつかある。

6.3.3 異なる結果が出た日本語の曲 1

日本語の曲も中国人と日本人で異なる結果が出た曲がいくつかある。その例を図 18 と図 19 に示す。

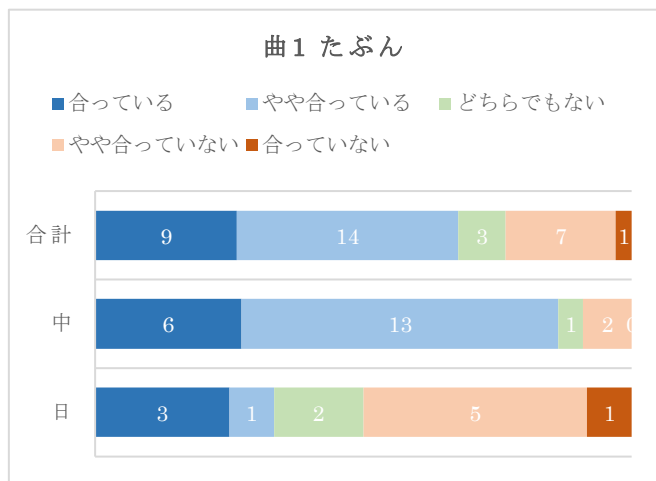


図 18 位置の適切さで異なる結果が出た日本語の曲 1

日本語の曲 1 では、中国語の曲 8 と同じように、歌詞がわかっているかわからないかにより、結果の傾向が違ふと思われる。

この曲のリズムは熱中のであり、軽快な雰囲気であるが、歌詞はネガティブに聞こえる。歌詞がわかっている日本人はより多くの人合っていないと答えた。

日本人から「negative で落ち着いた印象で切ない曲だと思います」、「落ち着いた印象(歌詞に引きずられているかも)」というコメントをもらった。

中国人からにも「そんなに積極的ではなく、少し左へ」というコメントをもらった。

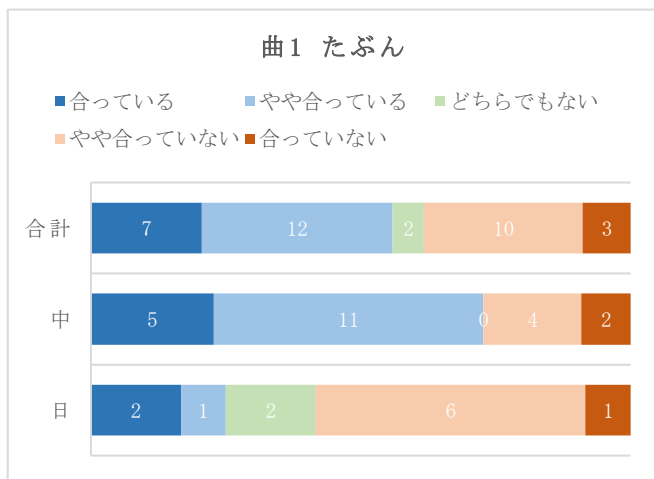


図 19 印象語の適切さで異なる結果が出た日本語の曲 1

日本語の曲 1 の印象語は「enthusiastic, light hearted」(熱中の, 気さくな)である。

日本人も中国人も、位置に関する問題の結果より、印象語が合わないという答えを選んだ人が少し増えた。この曲について、曲の位置は対応する印象語よりもっと適切だと考えられる。

日本人から「熱唱していないので、熱中のではなさそうだが、気さくそう」、「明るい曲というより重めな

曲です。」というコメントをもらった。

中国人からも「ちょっと悲しいです」、「情熱的です」というコメントをもらった。

6.3.4 印象語より位置が合うという結果が出た日本語の曲 2

日本語の曲は印象語より位置が合うという結果が出た曲がある。その例を図 20 と図 21 に示す。

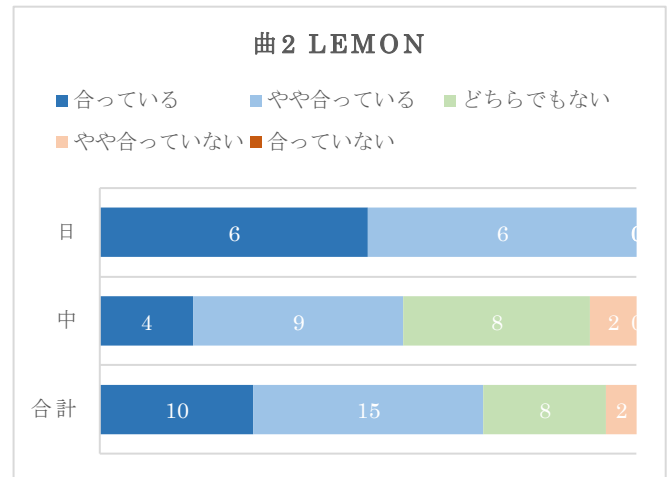


図 20 日本語の曲 2 の位置に関する結果

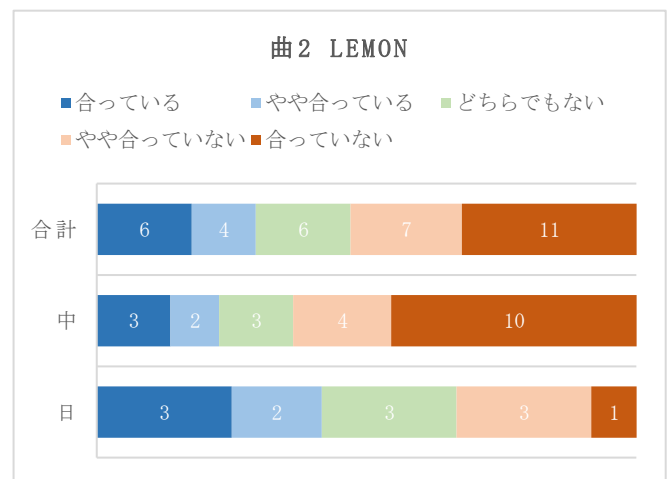


図 21 日本語の曲 2 の印象語に関する結果

日本語の曲 2 の印象語は「impatient, suspicious」(せっかち, 怪しい)である。

日本人から「悲しげな曲なのでせっかちや怪しいではないと思います」、「座標は納得いくが、感情空間の単語にやや違和感を覚えた」というコメントをもらった。

中国人からにも「悲しくて妻を思う男のように歌っています」、「この曲は落ち着いた上でやや積極的な感情を持っています。右の上で、曲 2 の近くにあると思います。」というコメントをもらった。

曲のデータで座標をつけて、周りの音楽を適切に推薦することが可能となるが、この感情空間の印象語自体は心理学を背景として提示されたものである。心理

学としての「impatient, suspicious」という感情表現は、音楽を表現する場合にはあまり用いられない、あるいは、表現されることが少ないと考えている。

6.3.5 位置より印象語が合うという結果が出た日本語の曲 8

日本語の曲は位置より印象語が合うという結果が出た曲がいくつかある。その例を図 22 と図 23 に示す。

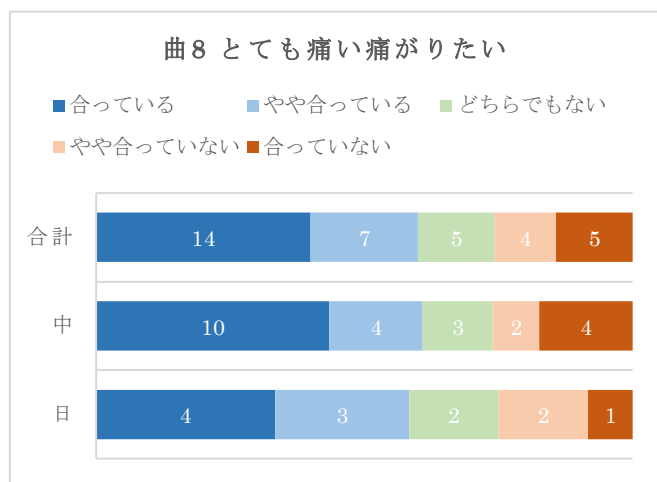


図 22 日本語の曲 8 の位置に関する結果

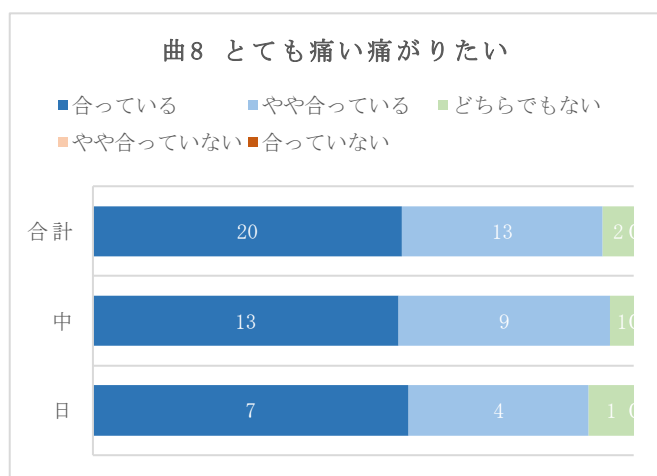


図 23 日本語の曲 8 の印象語に関する結果

日本語の曲 8 の印象語は「「high power/control, excited」(ハイパワー・コントロール, 興奮的)」である。

日本人から「明るい曲ではないので、そこまで端には位置しないと感じた」、「positive な印象は受けませんでした」、「歌詞がマイナスの感じがする」というコメントをもらった。

中国人からも「個人的にはこのマイナス感情が深く、聞くに堪えないと思います」、「メロディが強くない、感情はマイナスでもプラスでもないと思います」というコメントをもらった。

アンケート結果からは、この曲の位置に対応する印象語を認めてはいるが、歌詞および楽曲はネガティブ

に聞こえる。そのため、より多くの人がこの曲の位置は適切ではなく、もっとマイナスな感情を表す感情空間の左側に行ったほうがいいと思う傾向があると思われる。

また、日本人にこの傾向がより強いのは、歌詞がわかっていて、マイナスの感情を歌詞からもっと深く感じられるためと思われる。

7. 今後の課題

本研究では、中国語の曲を選び、中国人と日本人にそれぞれの音楽における感情空間上の位置及び印象語の適切さについてアンケートを行い、その結果を分析した。

今回は、中国語の曲と日本語の曲を選んでアンケートを用いた分析を行ったが、今後はさらに楽曲を増やし、異なるジャンルの曲に対する文化的な差異等について調査を行いたい。また、得られた結果を用いることで、介護施設等で入所者の好みに合わせた曲を選ぶ等、状況、時間、季節などを考慮しつつ、自分とは異なる世代の人の好みに合った楽曲推薦を行うシステムの構築を行いたい。大量の音楽 DB の中から、利用者が望む楽曲を適切に推薦できる手法を開発し、適切かつ従来にない多様な曲を推薦することを目指したい。

参考文献

- [1] Russell, J. A. (1980), "A circumplex model of affect. Journal of Personality and Social Psychology, 39(6), 1161-1178.
- [2] PALTOGLOU, Georgios, THELWALL, Michael, "Seeing stars of valence and arousal in blog posts", IEEE Transactions on Affective Computing, 2013, 4.1: 116-123.
- [3] 中野他「利用者の印象に基づく音楽レコメンドサービス：画像を利用した利用者指向の音楽レコメンド手法の考察」, 情報処理学会第 81 回全国大会, pp.383-384, 2019 年 2 月.
- [4] 迫木他「画像の印象に合った音楽の推薦システムの提案」, DEIM2018-P2-3
- [5] SOLEYMANI, Mohammad, et al. 1000 songs for emotional analysis of music. In: Proceedings of the 2nd ACM International workshop on Crowdsourcing for multimedia. ACM, 2013. p. 1-6.
- [6] Spotify <https://www.spotify.com/jp/>
- [7] 糸井勇貴,他." 楽曲の風景特徴化に基づく風景ウェア楽曲推薦システム", DEIM Forum, A8-3, 2017
- [8] 新穂龍太朗,他." 画像の印象に合う楽曲の自動推薦システムに関する研究", 映像情報メディア学会メディア工学研究会技術報告, ME2013-7, pp.23-26, 2013