EXPLORING THE RELATIONSHIP BETWEEN MOOD AND CREATIVITY IN ROCK LYRICS

Xiao Hu
Library and Information Science Program
Morgridge College of Education
University of Denver
xiao.hu@du.edu

Bei Yu
School of Information Studies
Syracuse University
byu@syr.edu

ABSTRACT

The relationship between mood and creativity has been widely studied in psychology, however, no conclusion is reached in terms of which mood triggers high creativity, positive or negative. This paper provides new insights to this on-going argument by examining the relationship between lyrics creativity and music mood. We use three computational measures to gauge lyrics creativity: Type-to-Token Ratio, word norms fraction, and WordNet similarity. We then test three hypotheses regarding differences in lyrics creativity between music with different moods on 2715 U.S. rock songs. The three measures led to consistent findings that lyrics of negative and sad songs demonstrate higher linguistic creativity than those of positive and happy songs. Our findings support previous studies in psycholinguistics that people write more creatively when the text conveys sad or negative sentiment, and contradict previous research that positive mood triggers more unusual word associations. The result also indicates that different measures capture different aspects of lyrics creativity.

1. INTRODUCTION

Music is a product of human's creativity, and yet few studies have been done to analyze musical creativity using computational methods [2]. In the meantime, progress has been made in the area of literature and language creativity (i.e., linguistic creativity). In this study, we borrow the measures of linguistic creativity to examine lyrics, the textual part inherently integrated in many music pieces, aiming to provide an alternative approach to music creativity research that is complementary to modeling music composition and music audio. This study is expected to provide new insights to the relationship between mood and creativity in general in that the argument on whether positive or negative mood trigger higher creativity remains inconclusive in psychology research [4].

In Western English dictionaries, creativity is defined as "…the ability to transcend traditional ideas, rules, patterns, relationships, or the like, and to create meaningful new ideas, forms, methods, interpretations" [18]. Based on this definition, when measuring creativity, a central task is to identify new or unusual patterns. In this study, we apply three linguistic measures to gauging lyrics creativity signified by vocabulary richness and unusual word associations.

This research is expected to contribute to research on creativity in music, psychology and linguistics. Besides, mood has been identified as a new metadata type or facet of music in recent years. Findings in this study will help analyzing music mood from a new angle, lyrics creativity.

2. RELATED WORK

To date creativity in lyrics has rarely been studied. However, research in the following related areas has inspired and informed this research.

2.1 Lyrics and Music Mood Classification

Lyrics have been used in predicting music mood, either standalone (e.g., [6] [8]) or being combined with music audio (e.g., [9], [10], [20]). These studies identified lyric features that were effective in mood classification such as higher-order bag-of-words features (e.g., trigrams and bigrams) [6], psycholinguistic and stylistic features [8] [9]. In terms of the relationship between lyrics and music mood, Hu and Downie [10] found lyrics were less effective for classifying negative and passive categories, while Schuller et al. [20] revealed lyrics were more helpful on the classification of valence (positive and negative feelings). These studies are insightful but none of them examined the aspect of creativity. Although mood classification is not the focus of this study, findings on the relationship between mood and lyrics creativity suggest adding lyrics creativity features may help improve mood prediction accuracy.

2.2 Linguistic Creativity

In the text domain, creativity is a new topic compared to other aspects of text analytics such as sentiment analysis and topic detection. There have been several workshops on linguistic creativity which focused on “unusual” language phenomena, such as metaphor and plot. As one of the few studies on metrics and tools for measuring linguistic creativity, [21] proposed a set of creativity measures which inspired this study. The measures will be described in details in Section 4.

2.3 Mood and Creativity

Mood and creativity, as two inherent traits of human nature, have long been studied in psychology, sociology and cultural studies (e.g., [1]), however, research on the relationship between mood and creativity has been inconclusive. Some studies support that positive mood increases cognitive flexibility and thus enhances creativity [12]. For example, Isen et al. [13] found that human subjects gave more unusual word associations under positive-affect conditions, which suggested positive relationship between positive mood and creativity. At the same time, other studies support that negative moods may promote artistic creativity and that positive moods restrain it [4][15]. Consequently, a context-dependent view is increasingly accepted recently [5].

3. HYPOTHESES

In this research we focus on the relationship between mood and linguistic creativity in lyrics. We test three hypotheses to investigate the relationship between mood and creativity. The mood categories studied in this paper are listed in Section 5.1, Table 1.

Hypothesis 1 (H1): Lyrics in sad mood category are more creative than those in happy mood categories.

Hypothesis 2 (H2): Lyrics in negative mood category are more creative than those in sad mood categories.

Hypothesis 3 (H3): Lyrics in active mood category are more creative than those in passive mood categories.

H3 is based on the connection between creativity and active brain activities. High creativity is more likely to be observed in lyrics with intense emotion than in calm ones.

4. LYRICS CREATIVITY MEASURES

Measuring creativity is difficult because the evaluation could be subjective to some extent. To obtain robust result we adopt multiple measures to gauge lyrics creativity. We can draw strong conclusion if all measures led to consistent results. The first measure that we adopt is Type-to-Token Ratio, which has long been used to measure vocabulary richness in creative writing [14]. However, we have also been cautioned that Type-to-Token Ratio may not be reliable for texts shorter than 350 words [7]. This is particularly relevant to this study because the average length of lyrics is about 200 words.

The other measures used in this study are inspired by the work of Zhu and colleagues [21]. As one of the few papers applying computational measures to predicting linguistic creativity, [21] proposed 13 linguistic measures and built a linear regression model to detect measures with more prediction power. Two of the most salient measures in [21] are adopted in this study to measure lyrics creativity. Both of them are psycho-linguistic measures.

4.1 Type-to-Token Ratio

Type-to-Token Ratio is defined as the number of unique terms in a piece of text divided by the number of total terms. It is often used to measure the vocabulary richness of text. Specifically, this measure (denoted as *r* thereafter) is defined as:

$$r = C_{uniq}(x)/n \tag{1}$$

where C_{uniq} denotes number of unique words in a piece of lyrics and *n* is the total number of words in it. In calculating this measure, we removed function words (also called stopwords) in the lyrics because they do not carry independent meanings and thus do not add to vocabulary richness. As vocabulary richness is related to creativity, a lyric with higher value of *r* is regarded as more creative.

4.2 Word Norms Fraction

This measure is to capture how “usual” a text is. In cognitive psychology experiments, *word norms*, which represent associations between words, were collected by asking human subjects to freely recall associative words (responses) when they see the cue words (stimuli). Therefore, lyrics with high occurrences of word norms should indicate high “usualness” and thus low creativity since creativity often corresponds to unusual patterns.

Several existing word norms thesauri have been built by different researchers in different countries. Because we are going to analyze U.S. rock lyrics, we choose a thesaurus developed in U.S. to prevent cultural impact on word associations. Specifically, we use the Free Association Norms built by researchers in University of South Florida [16]. This word norms dataset contains 72,176 pairs of associated words.

Using the Free Association Norms, word norms fraction (denoted as *f* thereafter) is calculated as:

$$f = C_{norm}(x, y)/n \tag{2}$$

Where $C_{norm}(x, y)$ is the count of word pairs that appear in the Free Association Norms, and n represents number of words in the text.

As lyrics are written in lines and, like a sentence, a lyric line can be regarded as a relatively independent unit, we calculate the f measure for each lyric line and use the average across all lines as a song's creativity score. A song with lower value of f measure is regarded as more creative than a song with higher value of f measure.

4.3 WordNet Similarity

WordNet [3] is an English lexicon with marked linguistic relationships among word senses including hyponyms, hypernyms, holonym, entailment, etc. With hyponyms and hypernyms, WordNet can be seen as a hierarchy of word concepts and from which similarities between concepts can be calculated. There are quite a few similarity measures defined to leverage WordNet. Following [21], we also used *path similarity* in this study but we adopted a different software implementation, namely the Word::Similarity module in CPAN [17]. Path similarity between two concepts is calculated by counting the nodes between them in the WordNet concept hierarchy. The similarity score belongs to the interval of (0, 1]. If two word senses are in the same *synset*, meaning the two are synonyms in WordNet, the similarity score would be 1. The more nodes on the path connecting two word senses, the smaller the similarity would be. If two word senses do not have a path in between, then the similarity score is -1.

Just as word norms fraction, we take one lyric line as the unit to calculate the WordNet similarity (denoted as s thereafter). Stopwords are removed and all pairs of remaining words in a lyric line are considered:

$$s = \frac{\sum S_{path}(x, y)}{n} \tag{3}$$

where $S_{path}(x, y)$ denote one word pair in a lyric line and n represents number of words in the line. The score of a piece of lyrics is the average of scores across all lines.

One noteworthy issue is that WordNet is organized by word senses instead of words. Because of the doubts surrounding the effectiveness of Word Sense Disambiguation (WSD), we did not conduct WSD and instead adopted a simplified approach that uses the highest similarity between all senses of two words.

Since s measures the similarity among words, a lyric with lower s value would be regarded as more creative than one with higher s value.

5. EXPERIMENT AND RESULTS

To test our hypotheses, we conducted an experiment to calculate and compare the aforementioned measures on a dataset of lyrics with mood labels.

5.1 The Dataset

This study adopts the Mood Tag Dataset[1] (MTD) used in the Audio Tag Classification task in the Music Information Retrieval Evaluation eXchange (MIREX) 2010. This dataset include 3,469 songs in 18 mood categories that cover positive, negative, active and passive moods. In the MTD, each song is labeled one or more mood categories according to the social tags applied to it [11]. It is noteworthy that more than 90% of the songs in the MTD are in the genre of rock and more than 95% of them are U. S. songs. Therefore, findings of this study are limited to U. S. rock lyrics that are written in English.

For the experiment, we constructed our dataset by combining mood categories in the MTD into the categories required by the hypotheses. Combinations of categories are shown in Table 1.

Lyrics in this study were downloaded from an online lyric database, LyricWiki.com. One unique nature of lyrics is that repetitions are very common (e.g., chorus is usually repeated multiple times). However, the creative measures (e.g., Type-to-Token Ratio, r) punish repetitions, which is likely to be unfair in the case of lyrics. Unlike repetitions in other genres of text, lyrics usually repeat a whole line or paragraph. If a line is creative, then repeating it is still, if not more, creative. Therefore, to alleviate such bias, repetitions of entire lines and paragraphs as well as notations in the lyrics were removed.

In order to avoid any bias caused by lyric length, we excluded songs with lyrics that are too long (> 500 words) and too short (< 100 words). As the experiment results suggested (see Section 6), the measures are indeed more or less sensitive to lyric length. We also balanced the datasets by setting the same number of songs in each comparable categories (e.g., positive vs. negative). In the cases where one category had more songs than the other as provided by the MTD, a random selection was conducted in the larger category. Table 1 shows the combination of categories and lyrics statistics. There are in total 2715 unique songs in this experiment.

5.2 The Results

5.2.1 Happy vs. Sad

The creativity measures on happy and sad songs are presented in Table 2. A t-test was conducted to examine the

[1] http://www.music-ir.org/mirex/wiki/2010:Audio_Tag_Classification

significance of the differences. The results indicate higher creativity level (higher *r*, lower *f* and *s*) in sad lyrics based on all the three creativity measures, and the difference is consistently significant across all measures. Our hypothesis *H1* is then not rejected.

	Category	Categories in MTD	#. of songs	avg. lyric length (st.dev.) in words
H1	Happy	glad, cheerful, gleeful	842	218.80 (77.19)
	Sad	sad, mournful, gloomy, brooding	842	201.61 (73.36)
H2	Positive	glad, cheerful, gleeful, confident hopeful, exiting	1470	220.45 (78.64)
	Negative	sad, mournful, gloomy, brooding angry, aggressive	1470	203.29 (75.62)
H3	Active	aggressive, angry, exciting, gleeful	861	222.44 (83.05)
	Passive	calm, dreamy	861	206.07 (76.90)

Table 1. Lyrics categories and statistics.

Category	*r*	*f*	*s*
Happy	0.5543	0.0563	-0.6344
Sad	0.6042	0.0502	-0.6430
p-value	**<0.0001**	**0.0030**	**0.0398**

Table 2. Results of Hypothesis 1

5.2.2 Positive vs. Negative

Measures on positive and negative songs are presented in Table 3. A *t*-test was conducted to examine the significance of the differences. We observed higher creativity level (higher *r*, lower *f* and *s*) in negative lyrics based on all the three measures. The difference is significant according to each of the measures. Our hypothesis *H2* is not rejected.

Category	*r*	*F*	*s*
Positive	0.5953	0.0557	-0.6366
Negative	0.6523	0.0490	-0.6431
p-value	**<0.0001**	**<0.0001**	**0.0399**

Table 3. Results of Hypothesis 2

5.2.3 Active vs. Passive

Measures on active and passive songs are presented in Table 4. A t-test was conducted to examine the significance of the differences. We have observed lower WordNet similarity (higher creativity) in passive songs, but higher Type-to-Token Ratio (higher creativity) in active songs (although the difference was not significant for *r*). In addition, there was no significant difference in terms of unusual word associations (as indicated by the *f* measure). Hence our hypothesis H3 is not consistently supported by all measures.

Category	*r*	*f*	*s*
Active	0.5881	0.0525	-0.6322
Passive	0.5775	0.0526	-0.6419
p-value	0.0648	0.4804	**0.0149**

Table 4. Results of Hypothesis 3

6. DISCUSSION

To further understand the relationship between lyric creativity and mood categories, we manually examined the 10 most creative and 10 least creative songs for each measure (except for *f* where 306 songs had the smallest value, 0). The category distributions of these songs are listed in Tables 5 to 7. A general trend across these tables is that the most creative songs include more sad and negative songs while the least creative songs consist of mostly happy and positive songs. This observation is consistent with the results of statistical tests on Hypotheses 1 and 2. The trend regarding active and passive songs differs across measures, and thus once again we cannot draw consistent conclusion on hypothesis 3.

Besides statistics, it is helpful to look at the lyrics themselves. The most creative song as measured by Type-to-Token Ratio (*r*) is Elton John's "Tiny Dancer". Part of its lyrics is presented below. It is indeed more creative than most songs, and it happens to be the only happy song among the 10 most creative ones. This discrepancy with the general trend is worth further study in the future.

```
   ...
Ballerina you must've seen her
Dancing in the sand
And now she's in me always with me
Tiny dancer in my hand

Jesus freaks out in the street
Handing tickets out for God
Turning back she just laughs
The boulevard is not that bad
   ...
```

As the second most creative song selected by WordNet Similarity (*s*), "Something in the Way" by Nirvana (Sad, Negative and Passive) contains the following lyrics:

```
And the animals I trapped have all become my pets
And I'm living off of grass and the drippings from
my ceiling
It's okay to eat fish 'cause they don't have any
feelings
Something in the way mmm
...
```

As it can be seen, this piece of lyrics indeed has some unusual combination of words or concepts.

10 Most Creative Songs (0.931 ≤ *r* ≤ 0.971)	Happy	1	Sad	5
	Positive	2	Negative	4
	Active	3	Passive	2
10 Least Creative Songs (0.105 ≤ *r* ≤ 0.203)	Happy	9	Sad	1
	Positive	9	Negative	1
	Active	1	Passive	4

Table 5. Categories of most and least creative songs measured by Type-to-Token Ratio (*r*)

306 Most Creative Songs (*f* =0)	Happy	76	Sad	98
	Positive	106	Negative	142
	Active	84	Passive	113
10 Least Creative Songs (0.276 ≤ *f* ≤ 0.369)	Happy	4	Sad	3
	Positive	6	Negative	2
	Active	6	Passive	3

Table 6. Most and least creative songs as measured by Word Norm Fraction (*f*)

10 Most Creative Songs (-0.895 ≤ *s* ≤ -0.844)	Happy	2	Sad	6
	Positive	2	Negative	6
	Active	0	Passive	3
10 Least Creative Songs (-0.244 ≤ *s* ≤ 0.088)	Happy	5	Sad	0
	Positive	7	Negative	3
	Active	4	Passive	2

Table 7. Most and least creative songs as measured by WordNet Similarity (*s*)

The fact that the word norm fraction measure (*f*) was 0 for 306 songs is interesting. A close inspection of the lyrics reveals these 306 songs have short lyric lines. This helps attribute the low *f* value to the way the measure was calculated. The word pairs were formed with each line of lyrics, and when the lines were short, there were few word pairs which resulted in fewer pairs matching the norms.

Another observation is that there are controversial songs. “Gangster Tripping” by Fatboy Slim was listed as the No.1 least creative song by Type-to-Token Ratio (*r*) but was the No.1 most creative song selected by WordNet similarity (*s*) and was among the most creative songs measured by word norm fraction (*f*). A close examination of the lyrics uncovers the reason: there are many word repetitions and thus its *r* value is very low. However, words in each line are neither similar nor with usual associations. A typical snippet of the lyrics is shown below:

```
...
We gotta kick that gangster shit
C'mon we gotta kick that gangster shit
C'mon we gotta get that
get that get that get that get that get that get
that get that get that get that

It's what we're doin' when a
What we're doin' when a
What we're doin' when a fatboy's slippin'
...
```

Such discrepancy on a single song discloses the limitations of the creativity measures. Type-to-Token Ratio just captures one kind of creativity, but it biases against creativity of repetitive patterns while repetition is a common feature of lyrics and does not necessarily indicate less creativity. Word norm fraction heavily relies on the given association norms. Furthermore, it favors lyrics with shorter lines since there are fewer word pairs in shorter lines, which possibly leads to lower scores (higher creativity). On the contrary, WordNet similarity favors longer lyric lines as those lines potentially contain more word pairs with similarity value of -1 (no WordNet path between the words) which contributed to a lower *s* value (higher creativity). In addition, WordNet similarity is limited by the hypernym and hyponym hierarchy which is only available for nouns and verbs. This analysis suggests that combination use of multiple measures gives the most comprehensive estimation of creativity as a multi-faceted linguistic phenomenon.

7. CONCLUSIONS AND FUTURE WORK

In this study we examined the relationship between mood and creativity in U.S. rock lyrics. We used three computational measures, Type-to-Token Ratio, Word Norm Fraction, and WordNet Similarity, to gauge lyrics creativity, and then compared the difference in creativity between lyrics in various mood categories. Because the three measures capture different aspects of linguistic creativity, our result suggests combination use of multiple measures to gauge lyric creativity.

We have also found that sad and negative lyrics correspond to higher linguistic creativity based on all three measures. This result supports previous studies on psycholinguistics that people write more creatively when the text conveys sad or negative sentiment, but contradict previous research that positive mood triggers more unusual word associations. One interpretation is that the impact of mood on the task of writing a piece of text with certain theme (like lyrics) is different from that on recalling free association between words. The former one involves more specific description. Furthermore, the correlation between creative writing and negative emotion is actually reflected in vocabulary of human languages. Sentiment lexicons in different languages share a common feature that negative words outnumber positive words. Schrauf and Sanchez [19] also found that people recall more negative words than positive words. These phenomena suggest that humans are actually better equipped with richer word choices when it comes to describe negative emotions.

This study focuses on lexical creativity. As future work it is worth investigating other dimensions of linguistic creativity: syntactic, morphological, and semantic creativity.

8. REFERENCES

[1] J. R., Averill, K. K., Chon, and D. W. Haan: "Emotions and Creativity, East and West," *Asian Journal of Social Psychology*, 4, pp.165-183. 2001.

[2] I. Deliège and G.A. Wiggins: *Musical creativity: multidisciplinary research in theory and practice*, Psychology Press, New York, 2006.

[3] C. Fellbaum (*eds.*): *WordNet: An Electronic Lexical Database*. MIT Press. 1998.

[4] K. Gasper: "When Necessity Is The Mother of Invention: Mood and Problem Solving," *Journal of Experimental Social Psychology*, 39, 248 –262, 2003.

[5] J. M. George and J. Zhou: "Understanding When Bad Moods Foster Creativity and Good Ones Don't: The Role of Context and Clarity of Feelings," *Journal of Applied Psychology*, 87(4), 687-697, 2002

[6] H. He, J. Jin, Y. Xiong, B. Chen, W. Sun, and L. Zhao: "Language Feature Mining for Music Emotion Classification via Supervised Learning From Lyrics," In *Proceedings of Advances in the 3rd International Symposium on Computation and Intelligence*, 2008.

[7] C. W. Hess and K. M. Sefton: "Sample Size and Type-Token Ratio for Oral Language of Preschool Children," *Journal of Speech and Hearing Research*, vol. 29, pp. 129-134. 1986.

[8] Y. Hu, X. Chen, and D. Yang: "Lyric-Based Song Emotion Detection with Affective Lexicon and Fuzzy Clustering Method," In *Proceedings of the 10th International Conference on Music Information Retrieval*, 2009.

[9] X. Hu and J. S. Downie: "Improving Mood Classification in Music Digital Libraries by Combining Lyrics and Audio", In *Proceedings of the 10th annual joint conference on Digital libraries*,2010.

[10] X. Hu and J.S. Downie: "When Lyrics Outperform Audio for Music Mood Classification: a Feature Analysis", In *Proceedings of the 11th International Symposium on Music Information Retrieval*, 2010.

[11] X. Hu, J. S. Downie, A. Ehmann: "Lyric Text Mining in Music Mood Classification", In *Proceedings of the 10th International Symposium on Music Information Retrieval*, 2009.

[12] E. R. Hirt, E. E. Devers, and S. M. McCrea: "I Want to Be Creative: Exploring The Role of Hedonic Contingency Theory in The Positive Mood-Creativity Flexibility Link," *Journal of Personality and Social Psychology*, 94(2), pp. 214-230. 2008

[13] A. M. Isen, M. M. Johnson, E. Mertz and G. F. Robinson: "The Influence of Positive Affect on The Unusualness of Word Associations," *Journal of Personality and Social Psychology*, 48(6), pp. 1413-1426, 1985

[14] D.A. Majid, A-G. Tan and K-C. Soh: "Enhancing Children's Creativity: An Exploratory Study on Using the Internet and SCAMPER As Creative Writing Tools", *Korean Journal of Thinking and Problem Solving*, 13(2), pp. 67-81, 2003.

[15] A. M, Mendes: "The Dark Side of Creativity: Biological Vulnerability and Negative Emotions Lead to Greater Artistic Creativity," *Personality and Social Psychology Bulletin* 34(12), pp.1677-86, 2008

[16] D. L., Nelson, C. L., McEvoy, and T. A. Schreiber: *The University of South Florida word association, rhyme, and word fragment norms*. 1998.

[17] S. Patwardhan and T. Pedersen: "Using WordNet-based Context Vectors to Estimate the Semantic Relatedness of Concepts," *In the Proceedings of the EACL 2006 Workshop Making Sense of Sense - Bringing Computational Linguistics and Psycholinguistics Together*. Trento, Italy. 2006.

[18] J. R. Robinson: "Webster's Dictionary Definition of Creativity," *Online Journal for Workforce Education and Development*, 3(2), 2007, Available at: http://opensiuc.lib.siu.edu/ojwed/vol3/iss2/2

[19] R. W. Schrauf and J. Sanchez: "The Preponderance of Negative Emotion Words across Generations nnd Across Cultures," *Journal of Multilingual and Multicultural Development*, 25(2-3), pp. 266-284, 2004.

[20] B. Schuller, C. Hage, D. Schuller, and G. Rigoll: "Mister D.J., Cheer Me Up!: Musical and Textual Features for Automatic Mood Classification", *Journal of New Music Research*, 39(1), pp. 13-34, 2010.

[21] X. Zhu, Z. Xu, and T. Khot: "How Creative Is Your Writing? A Linguistic Creativity Measure from Computer Science and Cognitive Psychology Perspectives," *In NAACL 2009 Workshop on Computational Approaches to Linguistic Creativity*, 2009.