

探討華文流行情歌歌詞與產生情感的關聯性

Power of love songs: that trigger emotions.

方瑜瑄 / Hung Yee Shuen

摘要

本研究為了探討華文流行情歌歌詞與產生情感的關聯性，利用了 X-**village** 所學習的 python 套件，進行了論壇感情版的文字分析，熱門歌曲歌詞分析，及熱門歌手歌曲分析。從中我們得到了熱門文字後，進行文字雲。研究中我們選擇了兩個台灣熱門的論壇，PTT 和 DCARD 的男女感情版進行分析，在歌曲方面，我們則分析了 3 個熱門情歌庫，1) KKBOX 2016-2018 熱門情歌歌曲，2) 2014-2016 金曲獎男歌手得主- 林俊傑的熱門情歌，3) 家喻戶曉的情歌王- 周杰倫。經過分析，（一）我們發現在 3 個情歌中分別出現比較多的字眼為‘愛，幸福，分手，男人，女生，離開，重要，朋友，自己，只是，如果’。（二）而從論壇感情版中，我們搜尋到的熱門字為，‘男友，男生，女友，感情，分手，前任，負責，安全，自己，愛，放手，很重要’。（三）在論壇和歌詞對比下，我們發現了有至少 12 個常見的共同熱門字。研究依據結果進行討論後，發現這樣的現象可能以心理學之原理‘核心衝突關係主題’（Core Conflict Relationship Theme, CCRT），從中理解隱含於情歌歌詞中的有接近的愛情觀內涵。當看到相近於自己心裡感受的字眼，聆聽者會認為歌詞表達了自己的內心感受間接愛上這首歌。如果這個假設成立，那我相信在未來，流行音樂也可以加入音樂治療和輔導，提高情緒的治愈的效果等。

關鍵字：愛情，歌曲，文字雲

一， 前言

‘音樂在我們生活中無處不在，也與我們的日常息息相關’這句話就是我想做這個研究的開始。音樂一直帶有治愈人心的效果，也許你會推薦失戀的他聽一首‘分手快樂’，也許你會獻給你的伴侶一首‘甜蜜蜜’，但是我們是否有真的思考過自己喜歡一首歌時的原因，牽動你的心的是歌曲的歌詞嗎還是旋律呢？在國立新竹教育大學的一項研究中，‘華語流行情歌之核心衝突關係主題分析’，他們的研究結果發現，熱門歌曲的意境和歌詞有高比例會偏向一些情感主題如失戀，思念，孤獨等等。他們也發現在這些情感表達的歌曲也會有很高的聽者偏好。除此，在 Steinberg Triangular Theory 中，也有提到當我們聽到某些關鍵字時，可能會牽動我們的情感，回憶等等。所以這此，本研究有了幾個值得探討的研究問題：

- 1) 探討情歌歌詞與聽者的選擇是否有關聯。
- 2) 探討熱門歌曲歌詞與聽者的情感是否有關聯。
(是不是歌詞裡有失戀者會想到的關鍵字會更觸動他們的心呢)

二， 研究方法

本研究利用的研究方式為找到論壇中的情感熱門字，探討在抒發情緒時通常會有什麼共同的字，或是普遍大家在抒發情感時會寫什麼關鍵字。除此，在探討歌詞的部分，我們會搜索熱門情歌的歌詞會出現的共同文字。

在此我們選擇了兩個台灣熱門的論壇，PTT 和 DCARD 的男女感情版進行分析，而在歌曲方面，我們則分析了 3 個熱門情歌庫，1) KKBOX 2016-2018 熱門情歌歌曲
2) 2014-2016 金曲獎男歌手得主- 林俊傑的熱門情歌
3) 家喻戶曉的情歌王- 周杰倫

三， 研究工具

這次的研究使用的是從 X-VILLAGE 課程中學習到的 python 套件。（Requests, Beautifulsoup, Numpy, Wordcloud, jieba）

四， 研究分析

探討情感關鍵字

1) PTT BOY-GIRL 板

在探討情感關鍵字，首先利用的 **python requests**，爬蟲，取得 PTT 8 月 3 號與 4 號的男女版中的熱門文章，再讀入 txt 檔案，進而進行文字雲來分析。

Step1

用以下的code來執行分析PTT感情版上的文章

```
#PTT 心情版爬蟲
import requests
from bs4 import BeautifulSoup

from urllib.request import urlopen

url = 'https://www.ptt.cc/bbs/Boy-Girl/index.html'

for num in range(0,5):
    response = requests.get(url)
    soup = BeautifulSoup(response.text, 'html.parser')
    tag_name = 'div.title a'
    articles = soup.select(tag_name)

    #其他page的文章
    paging = soup.select('div.btn-group-paging a')

    #看有多少Page
    print(paging)

    otherpage_url = paging[1]['href']
    next_url = 'https://www.ptt.cc'+otherpage_url
    url = next_url

    f = open('love1.txt', 'w',encoding='utf-8')
    for article in articles:
        print(article.text)
        f.write(article.text)
```

Step2

以下code是分析PTT文章後，做出文字雲。

```
#PTT 文字雲
import matplotlib.pyplot as plt
from wordcloud import WordCloud
import jieba

text_from_file_with_apath = open('ptt_04.txt' , "r" , encoding="utf8").read()

wordlist_after_jieba = jieba.cut(text_from_file_with_apath, cut_all = True)
wl_space_split = " ".join(wordlist_after_jieba)

my_wordcloud = WordCloud(font_path="C:\Windows\Fonts\setofont.ttf").generate(wl_space_split)
plt.imshow(my_wordcloud)
plt.axis("off")
plt.show()
```



	A	B
1	word	count
2		66
3	女生	15
4	重要	8
5	維繫	8
6	再連絡	8
7	男友	7
8	長相	6
9	果然	6
10	自己	5
11	女友	4
12	朋友	4
13	避嫌	3
14	個性	3
15	態度	3
16	突然	3
17	改變	3
18	親密關	3
19	配不上	3
20	男生	3
21	爭執	3
22	心裡	3
23	累	3
24	外表	2
25	價值觀	2
26	啪啪	2
27	女性朋友	2
28	半夜	2
29	起床	2
30	發現	2

2) DCARD

在探討情感關鍵字，我們利用python爬蟲抓出DCARD版中的感情版前30篇熱門文章，再進行文字雲探討在文章中出現最多的關鍵字。

```
import requests
from bs4 import BeautifulSoup
import re
url = 'https://www.dcard.tw/f/relationship'
resp = requests.get(url)
soup = BeautifulSoup(resp.text, 'html.parser')
dcard_title = soup.find_all('h3', re.compile('PostEntry_title_'))
print('Dcard 感情版熱門前三十文章標題：')
for index, item in enumerate(dcard_title[:30]):
    print("{0:2d}. {1}".format(index + 1, item.text.strip()))
```

Dcard 感情板前30篇熱門文章

Dcard 感情版熱門前三十文章標題：

1. 因為要聽你說話
2. 分手後怎麼處理「情傷」
3. 15歲與18歲
4. 有個不懂3C的女友真好
5. 前男友有新對象，該怎麼調適
6. 男生的聖人模式
7. 妳知不知道我要負責妳的安全？
8. 提醒自己
9. 從前車馬很慢，書信很遠，一生只夠愛一個人。

```
import matplotlib.pyplot as plt
from wordcloud import WordCloud
import jieba

text_from_file_with_apath = open('dcard_30.txt', "r", encoding="utf8").read()

wordlist_after_jieba = jieba.cut(text_from_file_with_apath, cut_all = True)
wl_space_split = " ".join(wordlist_after_jieba)

my_wordcloud = WordCloud(font_path="C:\\Windows\\Fonts\\setofont.ttf").generate(wl_space_split)
plt.imshow(my_wordcloud)
plt.axis("off")
plt.show()
```



	A	B	
1	word	count	
2		30	
3	男友	4	
4	男生	3	
5	感情	2	
6	女友	2	
7	文長	2	
8	分手	1	
9	處理	1	
10	情傷	1	
11	前男友	1	
12	新對象	1	
13	調適	1	
14	聖人模式	1	
15	負責	1	
16	安全	1	
17	提醒	1	
18	自己	1	
19	從前	1	
20	一生	1	
21	愛	1	
22	放下	1	
23	前女友	1	
24	變胖	1	
25	爽	1	
26	陪	1	
27	成長	1	
28	曖昧	1	
29	停	1	
30	損點	1	
31	網路	1	
32	交友	1	
33	現實	1	
34	慎入	1	
35	報備	1	

探討情歌歌詞與情感是否有相關

1) 情歌歌詞

在探討情歌歌詞的部分，我們利用 python 爬蟲抓出 KKBOX 2016-2018 年的熱門歌曲歌單，讀入 txt 檔案，再進行文字雲，探討在歌詞中出現最多的關鍵字是否和情感 (PTT,DCARD) 中出現的字有相同。

• KKBOX 2016-2018 年 熱門歌曲歌詞

[kkbox_top50.txt]



```
1 #KKBOX TOP 50 2016-2018 songs
2 別堆砌懷念讓劇情 變得狗血
3 深愛了多年又何必 毀了經典
4 都已成年不拖不欠
5 浪費時間是我情願
6 像謝幕的演員 眼看著燈光熄滅
7
8 來不及 再轟轟烈烈
9 就保留 告別的尊嚴
10 我愛你不後悔 也尊重故事結尾
11
12 分手應該體面 誰都不要說抱歉
```

```
import matplotlib.pyplot as plt
from wordcloud import WordCloud
import jieba

text_from_file_with_apath = open('kkbox_top50.txt' , "r" , encoding="utf8").read()

wordlist_after_jieba = jieba.cut(text_from_file_with_apath, cut_all = True)
wl_space_split = " ".join(wordlist_after_jieba)

my_wordcloud = WordCloud(font_path="C:\\Windows\\Fonts\\setofont.ttf").generate(wl_space_split)
plt.imshow(my_wordcloud)
plt.axis("off")
plt.show()
```




word	count
	1713
我們	84
走	55
愛	37
如果	30
因為	30
一起	29
不是	24
幸福	23
慢慢	23
只是	23
時間	22
成全	22
記得	21
回憶	20
自己	20
陪	20
離開	19
哭	19
心	19
愛情	19
天空	18
世界	18
聽	18
忘	16
別說	16
不要	15
遇見	15
相遇	15
懷念	14
留下	14

• 周杰倫 30 首熱門情歌歌詞



	A	B	
1	word	count	
2		1436	
3	愛	57	
4	走	42	
5	離開	36	
6	什麼	32	
7	不能	32	
8	等	31	
9	男人	30	
10	怎麼	25	
11	知道	22	
12	愛情	21	
13	回憶	17	
14	笑	16	
15	一起	15	
16	時間	14	
17	不是	14	
18	失去	14	
19	微笑	13	
20	從	13	
21	太快	13	
22	出來	13	
23	夢	13	
24	不該	13	
25	挽回	13	
26	一天	13	
27	錯	13	
28	故事	12	
29	難過	12	
30	發現	12	
31	陪	12	

（五），結果與結論

在以上的分析中，我們可以看到在論壇 DCARD 和 PTT 中，爬蟲到的文字有 300 多字（兩天的熱門文章），而情感抒發的熱門文字包括，【女生，男友，我們，自己，聯繫，重要，長相，陪，累，改變，朋友，感情，分手，情傷，前任，愛，寂寞】等等。

而在 KKBOX 熱門歌曲，林俊傑與周杰倫熱門情歌歌詞中的常見文字包括，

【我們，自己，只是，如果，愛情，懂了，沉默，害怕，擁抱，寂寞，離開，男人，回憶，陪，分手】。

在結合兩個論壇和熱門歌曲歌詞的熱門文字中，我們發現其實有很多重疊的字包括，【自己，分手，愛情，離開，累，沉默，寂寞】等等，出現的次數【例：愛，在 DCARD,PTT,三種歌曲 list 中都會出現，出現次數超過 200 次】，而【幸福，分手，寂寞，懂了】，也出現了超過 80 次。

其實從以上的共同熱門文字上，不難發現其實蠻多都和情感（戀愛階段，熱戀到分手裡會出現的情緒有關），例：愛情，男友，女友，自己，累，分手，回憶，擁抱。例如剛剛提到的愛，幸福，分手，寂寞，這些都是在感情中會遇到的心情，但也有發現負面字眼為居多。

這個研究結論，越多和論壇中有重疊字的歌曲，越有高的聆聽點擊率，也是一些獲獎的歌曲。【例：周杰倫 – 算什麼男人，林俊傑 – 記得】。

所以在這個研究中，我們的結果符合了前面的研究假設，也回應了前面的研究問題。所以我們可以假設，當人們聽到一首歌的歌詞中有符合自己的情感關鍵字，他們會更偏好愛上這首歌。

（六），研究建議

本研究限於台灣兩年內 2016-2018，基於取得的樣本和樣本年齡，所以我們假設這個研究假設只會在年齡介於 15-30 的台灣人民。除此，目前有 越來越多的閱聽人所聽的歌曲不再限於 華語，在英文、日韓語的部分的閱聽人口越來越多，尤其是時下的青年所佔的 比率最高，因此建議未來研究可將外語 的部分納入。

另外，由於在進行這項研究時，電腦常常出現編碼的問題，在產出 csv 檔和文字雲時會出現亂碼的情況，而在進行很多次的補救後還是無解，所以最後用在台灣購入的電腦才能執行，這也是這研究中最大的遺憾。下一次也可以試著以外語（韓文或英文）進行研究，也許會有不一樣的發現，不過這可能還是要顧及到文化的差異，或是喜歡偶像心態等等。

在探討這個研究問題時，假設如果歌詞真的可以牽動人的情感，情緒，記憶，那我相信未來在音樂治療，輔導的領域，可以再深入的研究是否可以達到治愈人心的效果。

此外，以後的研究也可以延伸文創歌曲，探討熱門歌曲的共同特點，再建議創作者加入這些元素，讓歌曲更完美，更接近人心。

參考文獻

王柏蓁（2007）。從流行音樂看台灣的愛情價值觀：1998-2007。《網路社會學通訊》，
許育光（2015）。華語流行情歌之核心衝突關係主題分析。

Sternberg, R. J. (1986). A triangular theory of love. *Psychological Review*, 2, 119-135
