Causing Emotion in Collocation: An Exploratory Data Analysis

Pei-Yu Lu, Yu-Yun Chang and Shu-Kai Hsieh
Graduate Institute of Linguistics, National Taiwan University
No. 1, Sec. 4, Roosevelt Road, Taipei, Taiwan, 10617
{emily.lulala, yuyun.unita, shukai} @gmail.com

Abstract

This paper aims to seek approaches in investigating the relationships within emotion words under linguistic aspect, rather than figuring out new algorithms or so in processing emotion detection. It is noted that emotion words could be categorized into two groups: *emotion-inducing* words and *emotion-describing* words, and *emotion-inducing* words would be able to trigger emotions expressed via *emotion-describing* words. Hence, this paper takes the social network Plurk, the emotion words are from the study on Standard Stimuli and Normative Responses of Emotions (SSNRE) in Taiwan and the National Taiwan University Sentiment Dictionary (NTUSD) as corpus, combining with Principle Component Analysis (PCA) and followed collocation approach, in order to make a preliminary exploration in observing the interactions between *emotion-inducing* and *emotion-describing* words. From the results, it is found that though the retrieved Plurk posts containing *emotion-inducing* words, polarities of the induced *emotion-describing* words contained within the posts are not consistent. In addition, the polarities of posts would not only be influenced by emotion words, but negation words, modal words and certain content words within context.

Keywords: sentiment analysis, emotion word, collocation.

1. Introduction

Sentiment analysis has recently become a prevalent trend in the field of natural language processing, and has wide applications for industry, policy making, sociology, psychology and so on. Various approaches have been proposed with impressive experimental or computational evidence, from document-level analysis to sentence-level or even phrasal-level analysis [1]. Among most studies, the Sentiment/Emotion-labeled Lexicon is taken as an indispensible lexical resource for the improvement of emotion classification accuracy. However, by assuming the static correspondence of word-emotion, most studies have neglected the fact that emotion words are not fixed with specific valence but are influenced under diverse contexts.

On account of contextual effects of emotion, [2-4] have firstly introduced a notion inspired by cognitive linguistics - emotion cause event - that refers to “the explicitly expressed arguments or events that trigger the presence of the corresponding emotions.” A set of linguistic cues is proposed to detect the cause events, resulting a valuable corpus resource for the task of emotion classification.

Despite that there are some explicit causers that might trigger emotions via context, a recent large-scaled interdisciplinary emotion research project [5, 6] has focused on the **emotion words** and found that they do help capture the emotion perceptions [7], and can thus be employed in emotion-related processing tasks. As designed in [5, 6] emotion words can be further grouped into *emotion-inducing* (情緒誘發詞) *emotion-describing* words (情緒描述詞). Emotions are mainly divided into two polarities: positive and negative. *Emotion-inducing* words encode the underlying repository knowledge to be able to elicit *emotion-describing* words. Therefore, in this study, we assume that the *emotion-inducing* word can be treated as the pivot in emotion detection of the sentences, and the way the *emotion-inducing* word interacts with its collocational context would be the key to a deeper understanding of emotional processing in texts.

Instead of seeking new approaches and algorithms in emotion detection, this paper aims to emphasize on seeking other possibilities in context-based emotion detection through investigating the relationships of emotion polarity between *emotion-inducing* and collocated content words. We carry out an exploratory data analysis with the assistances of programming technique and linguistic resolution on data inspection, in order to make prediction on the potential underlying linguistic cues within emotions embedded in context. Since taking *web as corpus* is convenient for its easy access and availability of voluminous data, one of the popular social network in Taiwan, Plurk, is considered in our study.

2. Literature Review

2.1 Emotion Classes

Constructing a gold standard emotion classification has long been an unsolved issue among various research fields, such as philosophy [8, 9] biology [10], linguistics [11, 12], neuropsychology [13] and computer science [14, 15]. Regardless of the disagreement and not having consensus on one emotion class, some parts of emotions are widely shared amid diverse emotion classes proposed by previous studies [16-18], which are happiness, sadness, fear, and anger. However, since our study is based on the approach of [2], we simply follow the five emotion classes adopted in the paper, which the emotion classification is firstly presented by [18] happiness, sadness, fear, anger, and surprise.

2.2 Emotion Words in Context

In sentiment analysis, the fact that emotion of a word changes based on contexts has been mostly neglected, which might lead to diverse polarities. Thus, in recent studies, researchers started to take this issue into consideration while exploring word sentiments.

In addition, since words may contain various senses and further evoke diverse emotions based on contexts, the need of a list of emotion lexicon would be practical and could be applied to a number of purposes. [19] introduced the approach of using Mechanical Turk provided by Amazon's online service of crowdsourcing platform for a large amount of human annotation on numerous linguistic tasks [20, 21].

To be more specific, emotion lexicon or also known as emotion words that are covered in sentiment detection and classification (for example, happy, sad, angry and so on) are mostly *emotion-describing* words, which are words that directly express and describe emotions. On the other hand, for words that have the potential to evoke or arouse emotions under context, are grouped as *emotion-inducing* words, such as holiday, homework, weekend, Monday and so on.

Since *emotion-inducing* words contain certain underlying implicit linguistic cues to evoke emotions, many studies work on different approaches to inspect the context-based emotion words. For example, [22] uses the technique of crowdsourcing and Mechanical Turk method to help annotate the lexicon that have the possibility to evoke emotions, and evaluate the results with inter-annotator agreement.

Other studies take the emotion cause event to help figure out the causers of emotions within context. As mentioned by [23], a cause of an emotion is suggested to be one event. Therefore, a cause event could be referred to a cause that could immediately trigger an event, as stated by [4]. [3] expresses in the paper that emotion cause detection is one of cause event detection, therefore some typical patterns that are used in cause event detection, such as because and thus, could be applied to emotion cause detection. Additionally, they have included some manually and automatically generalized linguistic cues to further explore emotion cause detection.

In this study, the experimental results of Chinese emotion word list in [5] are included, which obtain the valences (from 1 to 9) of word polarities in both *emotion-describing* and *emotion-inducing* words, in order to investigate whether given an *emotion-inducing* word along with context could the sentiment prediction model envision its possible evoked emotions presented via *emotion-describing* words.

3. Methodology

In order to investigate the implicit linguistic cues that might shift the polarities of emotion words, the analysis by applying Principle Component Analysis (PCA) and collocation of *emotion-inducing* words are considered. Through PCA, the distribution and relationships between *emotion-inducing* and *emotion-describing* words could be revealed and presented visually via the powerful plots in R. In addition, since PCA tends to exhibit the groups of emotion words that might have strong interactions between *emotion-inducing* and *emotion-describing* words, the approach by inspecting the collocations of *emotion-inducing* words would help figure out the linguistic cues that might lead to the interactions in context. Three materials taken in this study include Plurk corpus, emotion words from the study on Standard Stimuli and Normative Responses of Emotions (SSNRE) in Taiwan, and National Taiwan University Sentiment Dictionary (NTUSD, [24]).

3.1 Material

Like Twitter, Plurk is one of most popular social networks and micro-blogging service in Taiwan. Since Plurk can be easily and freely accessed through Plurk API 2.0[1] and along with its enriched emoticon information, a total of 43959 posts has been retrieved and used in this study.

Regarding the emotion words adapted from the project SSNRE, these words are categorized into two groups: *emotion-inducing* words and *emotion-describing* words. While the emotion of *emotion-inducing* words is recessive and needs to be triggered by the context, the emotion of *emotion-describing* words is dominant and exists in its semantic sense. That is, although *emotion-inducing* words have explicit polarities in experimental results, its polarities will be affected by the context, such as *emotion-describing* words in the same sentences. Based on the changeable polarities of *emotion-inducing* words, the paper treats *emotion-inducing* words as the target of observation. In the study of SSNRE, 395 *emotion-inducing* words and 218 *emotion-describing* words has been underwent three psychological experiments with a 9-point likert scale, which includes four to six perception parameters. In the 9-point likert scale, the number 9 refers to the greatest positive emotion; whereas, the number 1 indicated the most awful negative emotion. That is, emotion words that are more than five points would belong to positive emotion and those lower than five points would be assigned as negative emotion. Within the 395 *emotion-inducing* words, 140 words are with positive emotion and 255 words are with negative emotion; as to the 218 *emotion-describing* words, there are 58 words tagged as positive emotion and 160 words tagged as negative emotion. Since *emotion-inducing* words are to induce and trigger emotions, we assume that if a sentence

---

1 http://www.plurk.com/API
2 http://ckipsvr.iis.sinica.edu.tw/

contains an *emotion-inducing* word, the induced emotions will be revealed via *emotion-describing* words with the same polarity.

NTUSD is a list of positive and negative *emotion-describing* words that is constructed by [24], containing 9,365 positive and 11,230 negative *emotion-describing* words.

In this paper, *emotion-describing* words from SSNRE will be combined with NTUSD to enlarge the *emotion-describing* word list (which would be called as *mixed emotion-describing word list* in this paper).

3.2 Preparation for Processing Principal Component Analysis (PCA)

Reducing dimensions for preserving the most representative variables, PCA is a multivariate analysis that reveals the internal structure of the data in a way that best explains the variance in the data with a smaller number of variables. Words distributed based on independent variables, *emotion-inducing* words. Since there are many unknown factors that might influence the interactions between *emotion-inducing* and *emotion-describing* words, applying PCA would be a choice to provide a quick glance of the interaction strength in between, and helps fast investigation in figuring out sets of emotion words with strong relationships. (relationships between *emotion-inducing* words and *emotion-describing* words) Therefore, when having large amount of data, PCA would be suitable for a preliminary data exploration.

Through the analysis of PCA, the distribution of relationships between *emotion-inducing* words and *emotion-describing* words would be presented from R plots for further exploration. For running PCA in R, some variables related to *emotion-inducing* words and *emotion-describing* words need to be prepared which are stated as below.

Every post in retrieved Plurk data containing any one of 395 *emotion-inducing* words will be collected into our *ad hoc* database. After the collection of 20461 posts, the sentences are word-segmented into 710,908 tokens and tagged by Chinese Knowledge Information Processing (CKIP) tool[2]. Then, the sentiment score for each sentence would be calculated by the *mixed emotion-describing word list*, which includes 9,423 positive and 11,390 negative *emotion-describing* words. The calculation treats each positive *emotion-describing* word as one point, and each negative *emotion-describing* word as a minus one point. The final sentiment score for each sentence would be the sum of the occurrences of positive and negative *emotion-describing* words within each sentence. The final sentiment score for each sentence could then be grouped into three types of emotion polarities: positive, negative and

[2] http://ckipsvr.iis.sinica.edu.tw/

neutral.

From the PCA results, it is found out due to simple evaluation in calculating the final sentiment score, although posts that are identified as positive / negative emotion, there are some posts that might actually possess opposite emotion. Therefore, since the polarity of emoticons could imply the real emotion of a post [25], the Plurk emoticons are then included in order to get a more accurate result before processing collocation. As done in previous study [26], the polarities of Plurk posts are not only automatically classified but also manually evaluated using emoticons; thus in this paper, only posts that the polarities from final sentiment score meet with the assessed polarities in [26], would be preserved.

Two types of data are prepared for running PCA: one is the posts with positive *emotion-inducing* words, but calculated with negative emotion from final sentiment score; and the other is the posts with negative *emotion-inducing* words, but calculated with positive emotion from final sentiment score.

All the *emotion-inducing* and *emotion-describing* words from the prepared dataset are calculated with ratio of frequency. Additionally, since the distribution of emotion words' frequency probabilities presents a long tail in plot, which such long tail in statistics would be hard for processing a significant result, this study only preserves the emotion words that the probabilities are over the third quantile into PCA.

3.3 The Analysis Approach with Collocation

The results of PCA show sets for *emotion-inducing* and *emotion-describing* words with strong interactions, the reason for an explanation is not revealed which will be discussed in section 4. However, there might be some linguistic cues that could be observed for expressing the differences and further identifying the polarity changing of *emotion-inducing* words via context. Since the events within posts also possess underlying emotions and might affect *emotion-inducing* words in triggering the polarities of *emotion-describing* words, the approach by studying frequently collocated events with *emotion-inducing* words, is applied to help investigate the implicit polarities of events that might have an influence to the emotions triggered by *emotion-inducing* words. According to [27], the purpose of collocation is to explain the way in which meaning arises from language text. [27] indicates words that occur physically together have a stronger chance of being mention together and words do not occur at random in a text.

We propose, via investigating the collocation of *emotion-inducing* words which is widely used in corpus analysis, the causes for illustrating the relationships could be unveiled. Using the result observed from PCA, the span of *emotion-inducing* words' collocation is set to three (the preceding three words and succeeding three words of the *emotion-inducing* words) and calculated into frequency. Only the top three collocation words for each *emotion-inducing* word are selected for examining the emotion polarities.

4. Results and Discussion

In this section, the results from PCA listed below would be discussed with illustrative examples revealed from R plots, and further applied with collocational approach for exploration. In PCA, two various types of results evaluated from final sentiment score are discussed via R plots, including posts with positive *emotion-inducing* words but with an overall negative sentiment, and posts with negative *emotion-inducing* words but with an overall positive sentiment. Additionally, the illustrative examples taken for discussion from PCA results, would all be circled with dotted lines in the following plots. Furthermore, the collocations of positive and negative *emotion-inducing* words would be investigated, in order to find out linguistic cues to help illustrate the interactions between *emotion-inducing* and *emotion-describing* words.

4.1 Analysis in PCA

Figure 1 presents posts with positive *emotion-inducing* words but with an overall negative sentiment via PCA analysis. For the illustrative examples in the plot, two *emotion-describing* words *ke3 shi4* 可是 'however' and *bu4 neng4* 不能 'can not' (in black color) and three *emotion-inducing* words *yun4 dong4* 運動 'exercise', *shui4 jue4* 睡覺 'sleep' and *wan2* 玩 'play' (in grey color) imply that there are strong interactions within them. Therefore, it is roughly observed that *emotion-inducing* words such as *yun4 dong4* 運動 'exercise', *shui4 jue4* 睡覺 'sleep' and *wan2* 玩 'play', might be affected by the *emotion-describing* words *ke3 shi4* 可是 'however' and *bu4 neng4* 不能 'can not', and lead to an overall negative emotion in posts.

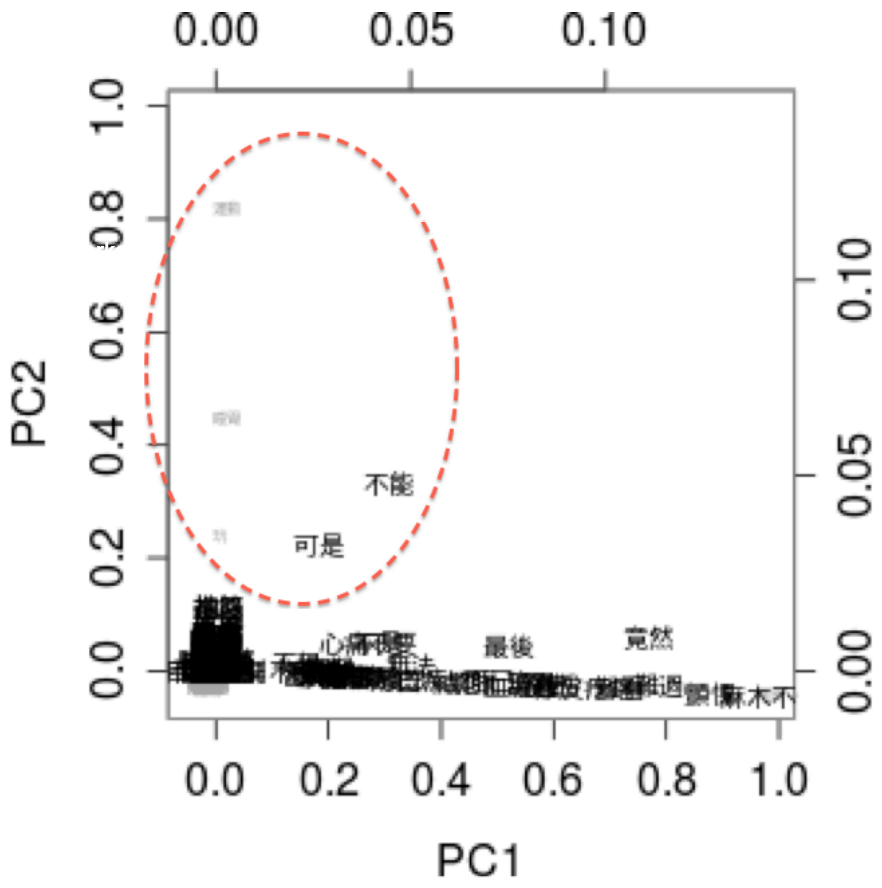

Figure 1. Negative Emotion but with Positive *Emotion-Inducing* Words

Results with negative *emotion-inducing* words but with an overall positive sentiment are expressed in Figure 2. There are two groups of illustrative examples in Figure 2.

For the first group (the top circle), there are strong interactions between four *emotion-describing* words *hen3 duo1* 很多 'many', *gan3 jue2* 感覺 'feel', *shi2 jian1* 時間 'time', and *xi1 wang4* 希望 'hope' (in black color) and one *emotion-inducing* word *kao3 shi4* 考試 'test' (in grey color). Therefore, it could be firstly imply that *emotion-inducing* word *kao3 shi4* 考試 'test' might be influenced by *emotion-describing* words, such as *hen3 duo1* 很多 'many', *gan3 jue2* 感覺 'feel', *shi2 jian1* 時間 'time', and *xi1 wang4* 希望 'hope', and cause polarity shifting from negative to positive in context.

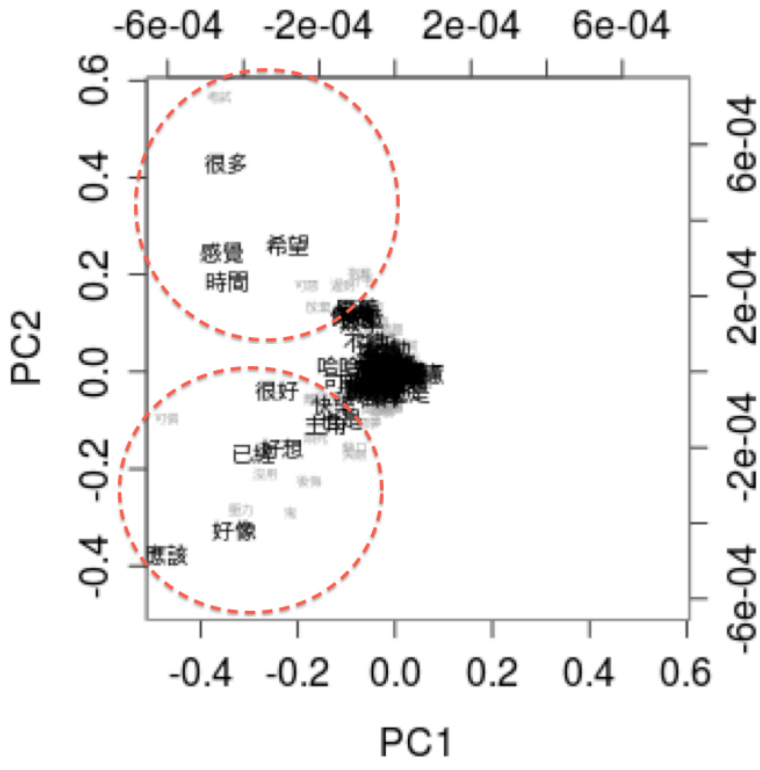

Figure 2. Positive Emotion but with Negative *Emotion-Inducing* Words

For the second group of illustrative examples in Figure 2 (the bottom circle), it is approximately find that four *emotion-describing* words *ying1 gai1* 應該 ‘should’, *hao3 xiang4* 好像 ‘seem’, *yi3 jing1* 已經 ‘already’, and *hao3 xiang3* 好想 ‘really want to’ (in black color) and four *emotion-inducing* words *ya1 li4* 壓力 ‘pressure’, *gui3* 鬼 ‘ghost’, *hou4 hui3* 後悔 ‘regret’, and *mei2 yong4* 沒用 ‘useless’ (in grey color) might have stronger interactions in context, in order to change the overall polarity from negative to positive than the other emotion words.

4.2 Collocations of *Emotion-inducing* Words

Though our previous assumption in the relationships between *emotion-inducing* and *emotion-describing* words is ‘positive *emotion-inducing* words would trigger positive *emotion-describing* words; and negative *emotion-inducing* words would trigger negative *emotion-describing* words’, the results discovered by PCA are apart from the assumption: [1] there are some positive *emotion-inducing* words that might arouse negative emotion-describing words and cause an overall negative emotion in posts; while, [2] there are some negative *emotion-inducing* words that might trigger positive *emotion-describing* words and lead to an overall positive emotion in posts.

Since nouns and verbs could be taken as linguistic cues in expressing events, only the top three frequently collocated nouns or verb within the collocations of *emotion-inducing* words (event collocations, for short) are considered in this paper.

4.2.1 Collocations of Positive *Emotion-Inducing* Words

The event collocation results of the three positive *emotion-inducing* words presented in Figure 1 (*yun4 dong4* 運動 'exercise', *shui4 jue4* 睡覺 'sleep', and *wan2* 玩 'play'), are listed in Table 1.

Therefore, as presented in Table 1, situations that posts containing positive *emotion-inducing* words, which might lead to an overall negative emotion are as below: 1) *emotion-inducing* word *yun4 dong4* 運動 'exercise' with event collocations such as *tou1 lan3* 偷懶 'lazy' and *chou1 jin1* 抽筋 'cramps'; 2) *emotion-inducing* word *shui4 jue4* 睡覺 'sleep' with an event collocation such as*han2 leng3* 寒冷 'cold'; 3) *emotion-inducing* word *wan2* 玩 'play' with event collocations such as *jia4 ri4* 假日 'holidays' and *ke3 xi1* 可惜 'unfortunately'. In above cases, the co-occurrences might shift the emotion polarity into negative ones

Table 1. The Emotion Polarities of Collocation of Positive *Emotion-Inducing* Words

Positive *emotion-inducing* words	*yun4 dong4* 運動 'exercise'	*shui4 jue4* 睡覺 'sleep'	*wan2* 玩 'play'
First Collocation	*shui4 jue4* 睡覺 'sleep'	*shi2 er4 dian3* 十二點 'twelve o'clock'	*da3 nao4* 打鬧 'roughhouse'
Polarity	+	0	+
Second Collocation	*tou1 lan3* 偷懶 'lazy'	*han2 leng3* 寒冷 'cold'	*jia4 ri4* 假日 'holidays'
Polarity	−	−	+
Third Collocation	*chou1 jin1* 抽筋 'cramps'	*xia4 ke4* 下課 'class dismissed'	*ke3 xi1* 可惜 'unfortunately'
Polarity	−	+	−

4.2.2 Collocations of Negative *Emotion-Inducing* Words

The event collocation results of the six negative *emotion-inducing* words presented in Figure 2 (*kao3 shi4* 考試 'test', *chi2 dao4* 遲到 'being late', *ke3 lian2* 可憐 'poor', *ya1 li4* 壓力 'pressure', *gui3* 鬼 'ghost' and *li2 kai1* 離開 'leave'), are listed in Table 2.

Furthermore, as shown in Table 2, posts containing negative *emotion-inducing* words might tend to an overall positive emotion in the circumstances as below: 1) *emotion-inducing* word *kao3 shi4* 考試 'test' with event collocations such as *xi1 wang4* 希望 'hope', *ma1 ma1* 媽媽 'mom', and *xiao4 lu4* 效率 'efficiency'; 2) *emotion-inducing* word *chi2 dao4* 遲到 'being late' with event collocations such as *se1 che1* 塞車 'traffic jam', *shang4 ban1* 上班 'work', and *tong2 shi4* 同事 'colleague'; 3) *emotion-inducing* word *ke3 lian2* 可憐 'poor' with event collocations such as *ba4 ba4* 爸爸 'dad' and *nan2 ren2* 男人 'man'; 4) *emotion-inducing* word *ya1 li4* 壓力 'pressure' with event collocations such as *jin4 du4* 進度 'schedule'; 5) *emotion-inducing* word *gui3* 鬼 'ghost' with event collocations such as *tai2 wan1* 台灣 'Taiwan'; 6) *emotion-inducing* word *li2 kai1* 離開 'leave' with event collocations such as *ren2 sheng1* 人生 'life', *kao3 shi4* 考試 'test' and *wan3 an1* 晚安 'good night'. Due to the positive emotion polarity of the events, the polarities of posts with negative *emotion-inducing* words turn into positive ones

Table 2. The Emotion Polarities of Collocation of Negative *Emotion-Inducing* Words

Negative emotion-inducing words	***kao3 shi4* 考試 'test'**	***chi2 dao4* 遲到 'being late'**	***ke3 lian2* 可憐 'poor'**	***ya1 li4* 壓力 'pressure'**	***gui3* 鬼 'ghost'**	***li2 kai1* 離開 'leave'**
First collocation	*xi1 wang4* 希望 'hope'	*se1 che1* 塞車 'traffic jam'	*ba4 ba4* 爸爸 'dad'	*jin4 du4* 進度 'Schedule'	*zuo4 meng4* 作夢 'dreaming'	*ren2 sheng1* 人生 'life'
Polarity	+	+	+	+	−	+
Second collocation	*ma1 ma1* 媽媽 'mom'	*shang4 ban1* 上班 'work'	*nan2 ren2* 男人 'man'	*ji1 yin1* 基因 'gene'	*tai2 wan1* 台灣 'Taiwan'	*kao3 shi4* 考試 'test'
Polarity	+	+	+	−	+	+
Third collocation	*xiao4 lu4* 效率 'efficiency'	*tong2 shi4* 同事 'colleague'	*ya2 tong4* 牙痛 'toothache'	*fan2 nao3* 煩惱 'trouble'	*chong3 wu4* 寵物 'pets'	*fan2 nao3* 煩惱 'trouble'
Polarity	+	+	−	−	−	+

5. Conclusion

Sentiment/Emotion analysis has been one of the most important fields in NLP and computational intelligence. Different machine learning algorithms coupled with different feature combinations are proposed and have gained great achievement. Nonetheless, it is still a formidable task due to the permanent-in-context properties and the covert way we process emotions. In this paper, we argue that a static word list of emotion-labeled information would not suffice. As a preliminary step, we conduct an exploratory multivariate analysis (PCA) based on the Plurk corpus, NTUSD and SSNRE, and find out that *emotion-describing* words such as some negation words, modal words and certain content words would affect the polarities of posts, regardless of the *emotion-inducing* words' polarities. That is, the polarities of posts are beyond expectation. Nevertheless, as an exploratory analysis, in the limited amount of data, the findings need deeper development and further research for more complete evidence.

The collocation information has been a widely used contextual cue in corpus-based syntactical-semantic analysis. However, in computational sentiment analysis, the use of collocation does not focus on investigating the implicit linguistic cues but on its explicit frequency values. Since this kind of underlying embedded linguistic features has been long neglected, these would only improve the accuracy of the sentiment detection, but also leverages a Chinese Emotion Lexicon that will be created in the future.

References

[1] T. Wilson, J. Wiebe, and P. Hoffmann, "Recognizing contextual polarity in phrase-level sentiment analysis," in *Proceedings of the conference on Human Language Technology and Empirical Methods in Natural Language Processing*, 2005, pp. 347-354.

[2] Y. Chen, S. Y. M. Lee, S. Li, and C.-R. Huang, "Emotion cause detection with linguistic constructions," in *Proceedings of the 23rd International Conference on Computational Linguistics*, 2010, pp. 179-187.

[3] S. Y. M. Lee, Y. Chen, S. Li, and C.-R. Huang, "Emotion Cause Events: Corpus Construction and Analysis," in *LREC*, 2010.

[4] S. Y. M. Lee, Y. Chen, and C.-R. Huang, "A text-driven rule-based system for emotion cause detection," in *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, 2010, pp. 45-53.

[5] C. M. Cheng, H. C. Chen, and S.-L. Cho, " Affective words," *In A Study on Standard Stimuli and Normative Responses of Emotion in Taiwan,* 2012.

[6] "Standard Stimuli and Normative Responses of Emotions (SSNRE) in Taiwan," *Project website at http://ssnre.psy.ntu.edu.tw/,* 2012.

[7] L. F. Barrett, K. A. Lindquist, and M. Gendron, "Language as context for the perception of emotion," *Trends in cognitive sciences,* vol. 11, pp. 327-332, 2007.

[8] W. James, "II.—WHAT IS AN EMOTION?," *Mind,* pp. 188-205, 1884.

[9] B. de Spinoza, *The collected works of Spinoza* vol. 2: Princeton University Press, 1985.

[10] C. Darwin, "On the origins of species by means of natural selection," *London: Murray,* 1859.

[11] Z. Kövecses, *Metaphor and emotion: Language, culture, and body in human feeling*: Cambridge University Press, 2003.

[12] A. Wierzbicka, *Emotions across languages and cultures: Diversity and universals*: Cambridge University Press, 1999.

[13] J. H. Turner, "The Evolution of Emotions in Humans: A Darwinian–Durkheimian Analysis," *Journal for the theory of social behaviour,* vol. 26, pp. 1-33, 1996.

[14] A. Ortony, *The cognitive structure of emotions*: Cambridge university press, 1990.

[15] R. W. Picard, *Affective computing*: MIT press, 2000.

[16] P. Ekman, "Expression and the nature of emotion," *Approaches to emotion,* vol. 3, pp. 19-344, 1984.

[17] R. Plutchik, *Emotion: A psychoevolutionary synthesis*: Harper & Row New York, 1980.

[18] J. H. Turner, *On the origins of human emotions: A sociological inquiry into the evolution of human affect*: Stanford University Press Stanford, CA, 2000.

[19] S. M. Mohammad and P. D. Turney, "Emotions evoked by common words and phrases: Using Mechanical Turk to create an emotion lexicon," in *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, 2010, pp. 26-34.

[20] C. Callison-Burch, "Fast, cheap, and creative: evaluating translation quality using Amazon's Mechanical Turk," in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1*, 2009, pp. 286-295.

[21] R. Snow, B. O'Connor, D. Jurafsky, and A. Y. Ng, "Cheap and fast---but is it good?: evaluating non-expert annotations for natural language tasks," in *Proceedings of the conference on empirical methods in natural language processing*, 2008, pp. 254-263.

[22] S. M. Mohammad and P. D. Turney, "Crowdsourcing a word–emotion association lexicon," *Computational Intelligence,* 2012.

[23] L. Talmy, *Toward a cognitive semantics, Vol. 1: Concept structuring systems*: The MIT Press, 2000.

[24] L. W. Ku and H. H. Chen, "Mining opinions from the Web: Beyond relevance retrieval," *Journal of the American Society for Information Science and Technology,* vol. 58, pp. 1838-1850, 2007.

[25] 孫瑛澤, 陳建良, 劉峻杰, 劉昭麟, and 蘇豐文, "中文短句之情緒分類."

[26] M.-Y. Chen, H.-N. Lin, C.-A. Shih, Y.-C. Hsu, P.-Y. Hsu, and S.-K. Hsieh, "Classifying mood in plurks," in *ROCLING*, 2010.

[27] J. Sinclair, *Corpus, concordance, collocation*: Oxford University Press, 1991.