Information-seeking Questions and Rhetorical Questions in Emotion Expressions

Helena Yan Ping Lau
Department of Chinese and Bilingual Studies
The Hong Kong Polytechnic University
Hong Kong

helena.lau@connect.polyu.hk

Sophia Yat Mei Lee
Department of Chinese and Bilingual Studies
The Hong Kong Polytechnic University
Hong Kong

ym.lee@polyu.edu.hk

Abstract

This paper explores the interaction between emotions and two types of questions, namely information-seeking questions and rhetorical questions. Corpus data shows that rhetorical questions (60.3%) are more frequently used in social media than information-seeking questions (39.7%). Of the two types of questions, approximately 94% of rhetorical questions are used to express emotions, while only 23% of information-seeking questions contain emotions. Given that rhetorical questions do play an important role in emotion expressions, we examine the interaction between of rhetorical questions and emotions in terms of question type. Various syntactic structures are proposed for the identification of different emotions. We believe that the linguistic account of different types of rhetorical questions in emotion expressions will paint a fuller picture of the nature of emotion.

1   Introduction

Information-seeking questions (IQs) generally aim to elicit an answer, while rhetorical questions (RQs), expecting no answer, aim to achieve a pragmatic goal, such as to emphasize, to persuade, to show emotions etc. (Frank, 1990; Roberts & Kreuz, 1994). As a form of figurative language, rhetorical questions usually imply meaning that go beyond the literal. It is generally believed that rhetorical questions are a rather productive means of expressing or evoking emotions, in particular the negative ones (Roberts & Kreuz 1994; Gibbs et al. 2002; Lee 2017). Yet, it has been a challenging task to distinguish rhetorical questions from information-seeking questions, as both of them have the structure of a question. In addition, given that figurative language is frequently used for emotion expressions (Kövecses, 1990, 2003; Lakoff & Johnson, 1980; Fussell & Moss, 1998; Gibbs et al., 2002), the disregard for the interaction between rhetorical questions and emotions has greatly restricted the classification and detection of emotions.

This paper aims to study the use of information-seeking questions and rhetorical questions in social media, and explore the interaction between rhetorical questions and emotions in terms of question type. We propose various syntactic structures that can be used to identify different emotions.

2   Related Work

Rhetorical questions are generally regarded as an effective persuasive device (Petty, 1981; Frank, 1989). As a form of figurative language, rhetorical questions are sometimes studied in a more general way. A great deal of work indicated that figurative language is commonly used to express emotions (Kövecses, 1990, 2003; Lakoff & Johnson, 1980; Fussell & Moss, 1998; Gibbs et al., 2002), especially the intense ones (Fainsilber & Ortony, 1987; Fussell, 1992). Roberts & Kreuz (1994) examined the discourse goals of 8 types of figurative devices, namely hyperbole, idiom, indirect request, irony, understatement, metaphor, rhetorical questions, and simile. They found that rhetorical questions are used to express both positive and negative emotions, with the latter being more frequent. Leggitt & Gibbs (2000) investigated people's emotion reactions to different figurative devices. They showed that rhetorical questions are used to alert or challenge addressee's problem or behavior. There-

fore, rhetorical questions are prone to evoke negative emotions, such as *anger*, *disgust*, and *contempt*. In addition, speakers of rhetorical questions appear to feel more negative emotions than that of other figurative devices. Rhetorical questions are also perceived as having very negative intent. Lee (2017) suggested that there is a close interaction between figurative language and emotion. She found that about one-third of the social media posts contain figurative devices, among which rhetorical questions are the most frequently used one (37%). She illustrated that rhetorical questions are particularly productive in evoking negative emotions, i.e. *sadness* and *anger*.

Despite the important role rhetorical questions play in emotion expressions, existing classification models have mainly developed for other forms of figurative language, such as irony and sarcasm (Davidov et al., 2010; González-Ibáñez et al., 2011; Reyes et al., 2013). The automatic identification of rhetorical questions has received little attention (Bhattasali et al., 2015; Ranganath et al., 2016), let alone the distinctive structures of rhetorical questions for the identification and detection of different emotions.

3   Corpus Data and Annotation

The corpus was made up of 8,529 posts randomly extracted from *Sina Weibo*, one of the most popular social media sites in China (Lee, 2015). Each post contains no more than 140 characters, and emoticons are taken into account for the annotation.

The corpus was annotated by two annotators. Five basic emotions were annotated in each post, namely *happiness*, *sadness*, *anger*, *fear*, and *surprise*. Some posts contain more than one emotion, and all of them were labelled. For the identification of the use of information-seeking questions and rhetorical questions, all the 8,529 posts were read through. For each post that contains both emotion and question, annotators would be asked to determine whether or not the tagged emotion is concerned with the question identified. If not, the question would be regarded as "no emotion". According to Lee (2017), questions can roughly be categorized into open question and closed question. She further classified the questions into 10 subtypes, including A-not-A, alternative, echo, particles, *wh*-questions and so on. Although both information-seeking questions and rhetorical ques-

tions contain open and closed questions, only rhetorical questions were further classified as our preliminary observation suggests that most information-seeking questions do not express emotions. In addition to the 10 subtypes proposed in Lee (2017), 4 were added, given the existence of other types of rhetorical questions, including *which*, *where*, *when* and *others*.

In Chinese, open questions refer to questions with *wh*-words such as *why*, *what*, *how*, etc., and they aim to elicit an open-ended answer. Closed questions refer to questions represented in the form of A-not-A structure, alternative, echo, particle or other question words that require a pre-determined answer. A-not-A questions are formed with an affirmative and its negative counterpart juxtaposed, and the respondent can choose either the affirmative or its negative counterpart as the answer (Li and Thompson, 1981). Consider (1).

(1) 你喜不喜歡台灣菜？
    (Do you like Taiwanese food?)

Alternative questions explicitly provide two or more possible options which are mostly connected by the morpheme 還是 (or), as in (2).

(2) 你喜歡台灣菜還是日本菜？
    (Do you like Taiwanese food or Japanese food?)

Echo questions have the form of a declarative sentence but end with a question mark in the written form. Particle questions refer to questions that end with a sentence-final particle, such as 嗎, 呢. Rhetorical interrogation markers such as 難道, 何必 etc. are grouped into others.

4   Data Analysis

4.1   Inter-annotator Agreement

In order to evaluate the annotation tasks, two annotators were asked to annotate the same set of data which comprises 500 posts. Table 1 shows the inter-annotator agreement calculated using Cohen's Kappa coefficient. The Kappa scores for both emotion annotation and question type annotation are high, indicating that the quality of annotation is satisfactory. Apart from that, the agreement of question type annotation is higher than emotion annotation. This may be attributed to the fact that

emotions are subjective in nature, but question types are not.

	Kappa Score
Emotion	0.663
Question Type	0.918

Table 1 - Inter-annotator Agreement

4.2 Corpus Analysis

Of the 8,529 posts, 3,671 posts (43%) do not contain any emotions, while 5,137 emotions are identified in 4,858 posts (57%). That means, some posts express more than one emotion. Among the five emotions, *happiness* has the highest frequency (49.7), followed by *sadness* (20.2%), *anger* (12.4%), *surprise* (7.4%), and *fear* (4.9%).

The total number of questions identified is 900, among which 357 (39.7%) are information-seeking questions and 543 (60.3%) are rhetorical questions. Figure 1 shows the distribution of emotions per question type. This is calculated relative to the total number of each question type.
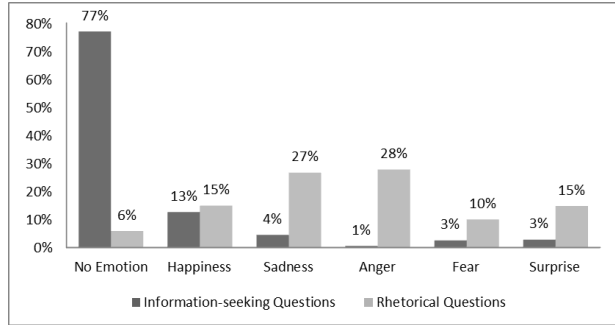

Figure 1 - Emotion Using IQs and RQs

From Figure 1, we can see that a lot more than half (77%) of information-seeking questions are not associated with any emotions, and less than a quarter of them express *happiness*. The remaining 11% are used to express *sadness* (4%), *fear* (3%), *surprise* (3%), and *anger* (1%). As for rhetorical questions, the vast majority (94%) of them do evoke emotions. Unlike information-seeking questions, rhetorical questions have a tendency to negative emotions, especially *anger* (28%) and *sadness* (27%). The *happiness* and *surprise* emotion account for 15% of rhetorical questions respectively. *Fear* has the weakest connection (10%) with rhetorical questions, which may be due to the small number of posts containing *fear*. In order to compare the role information-seeking questions and rhetorical questions play in emotion expressions, Figure 2 shows the distribution of question type per emotion in all posts.
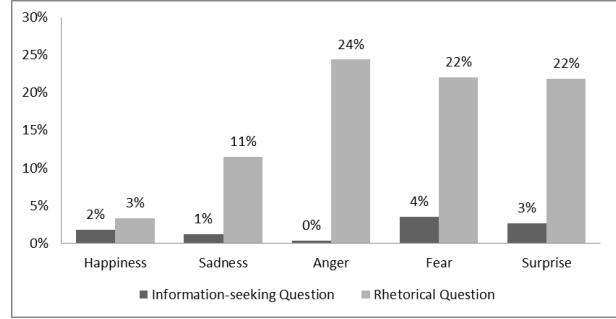

Figure 2 - Distribution of Question Type per Emotion in All Posts

Figure 2 is calculated relative to the total number of posts of a given emotion type. On the one hand, Figure 2 shows that information-seeking questions are not often used to express emotions. Among the emotions, *fear* is most frequently expressed via information-seeking questions, accounting for only 4%. On the other hand, Figure 2 shows that rhetorical questions are rather productive in expressing emotions as more than one fifth of posts containing *anger*, *fear*, and *surprise* are expressed by means of rhetorical questions. The *anger* emotion has the greatest tendency (24%) to be expressed via rhetorical questions, followed by *surprise* (22%), *fear* (22%), *sadness* (11%), and *happiness* (3%).

In sum, Figure 1 and 2 not only demonstrate the significant role rhetorical questions play in expressing emotions as compared to information-seeking questions, they also illustrate that rhetorical questions are productive in expressing negative emotions which is in line with Roberts & Kreuz (1994). In the next section, we will further discuss the interaction between rhetorical questions and emotions in terms of question type.

5 Interaction between RQs and Emotions

Lee (2017) indicated that rhetorical questions are the most frequently used figurative devices for evoking emotions. Although the distribution of rhetorical questions was shown, the study focused more on the general picture of the use of various figurative devices, such as metaphor etc. Hence, only a small number of rhetorical questions (i.e. 38 tokens) were identified and the scattered distribution could not clearly show the correlation between each type of rhetorical questions and emotions. Therefore, we further classify the 543 rhetorical questions into 14 subtypes, as shown in Table 2.

		Close Class Question					Open Class Question								
	Series of Q	A-not-A	Alternative	Echo	Particle	Others	How	How many/ much	What	Which	Who	Why	Where	When	**Total**
No Emotion	1	7	0	2	10	1	2	0	5	0	0	3	2	0	**33**
Happiness	5	27	0	2	14	4	7	0	7	0	2	15	1	0	**84**
Sadness	11	17	4	0	27	5	15	5	25	0	5	34	1	1	**150**
Anger	39	6	1	20	37	4	5	0	12	1	7	21	3	0	**156**
Fear	11	6	2	1	5	1	8	0	13	0	0	5	4	0	**56**
Surprise	9	0	2	15	22	9	3	2	10	0	0	11	0	0	**83**
Total	**76**	**63**	**9**	**40**	**115**	**24**	**40**	**7**	**72**	**1**	**14**	**89**	**11**	**1**	**562**

Table 2 - Distribution of Types of Rhetorical Questions Used

Table 2 shows that different types of rhetorical questions may show different preferences for a particular emotion. While more than a half of the posts containing a series of questions are used to express *anger*, only a couple of them are used to express *happiness*, and the rest are almost evenly distributed for the expressions of *sadness*, *fear*, and *surprise*. Frank (1989: 734) mentioned the use of clustering of rhetorical questions as "…in each case the question is re-stated for emphasis, in slightly different form. This makes for a stronger impact on the hearer; a strategy that most likely would be unnecessary if these were simply informational questions, but is a highly effective device for persuasion…". Lee (2017) regarded it as the strategy writers use to draw readers' attention to their strong emotions. Consider (3).

(3) 學校你是要把我趕盡殺絕嗎？停電？你居然停電？你居然停電了？？？涼快，涼快，涼快 [微風]
(Does this school trying to kill me or something? It went out of electricity? Seriously? It's out of electricity??? Cool down, cool down, cool down [breeze])

In (3), the rhetorical questions are restated to emphasize the fact that the electricity has been turned off. The writer intentionally used a series of questions to increase the emotion intensity which cannot be reached by stating only once. Ekman & Cordaro (2011: 365) defined *anger* as "the response to interference with our pursuit of a goal we care about…". Thus, the purpose of the use of a series of questions is to vent one's anger to someone who evokes the emotion.

As for closed questions, particle questions are most frequently used to form rhetorical questions, one-third of which are used to express *anger*. Among the five emotions, *anger* is often expressed in the pattern of "還…….嗎?", as in (4).

(4) 如果每件事我們都能自己完成那還用你教嗎，那你早就下崗了
(If we can accomplish everything on our own, why would we even need you? You'd have been sacked long ago)

Although this pattern may also appear in posts of other emotions, the use of 還 in rhetorical questions shows a tendency toward *anger*. Liu (2000) indicated that the use of 還 in rhetorical question is not to make objective statements but to express one's attitude and make the statement even stronger. It reflects how well a fact is established to the writer. To the writer in (4), students being incapable of accomplishing everything on their own is understandable as it is 你 (your) responsibility to teach them. This kind of rhetorical questions is uttered when what other people do does not reach the writer's standard or live up to his expectation. This explains the strong correlation between the pattern and *anger*.

"Declarative + 好嗎/麼/嘛" is another frequent pattern used to express *anger*, as in (5).

(5) 我真是討厭死夏天多蟲的季節了好麼！！
(I really hate summer and all the bugs so damn much, okay??)

In (5), 好麼 is obviously not a question used to seek information, but to highlight how much the writer hates summer. This pattern is occasionally

found to express *happiness* as in (6), and rarely, if not never, found to express other emotions. It is observed that the semantic polarity of the verb(s) or adjective(s) in the declarative sentence may give some hints about the emotion expressed. That is, rhetorical questions are likely used to express *happiness* (/*anger*) when there is one or more than one positive (/negative) verb or adjective found in this kind of sentences.

(6) [愛你]今晚還是很炫酷的好嗎
　　([Love you] We actually showed off a lot tonight, okay?)

"Declarative + question word" can also be found in A-not-A questions. Among those A-not-A forms, 有沒有[1] and 好不好 are the most common ones. While the latter does not have a clear semantic orientation pointing to a particular emotion, the former tends to associate with *happiness*, as in (7).

(7) 我瘦了，有木有。[哈哈][哈哈][哈哈][哈哈]
　　(I lost weight, yea? [haha][haha][haha][haha])

In (7), the writer used 有木有 to "re-confirm" the statement he made. In fact, the writer holds firm to his belief that he did lose some weight as hinted by the period and emoticons he used. Although an answer may be given in spoken context, most readers would not bother to comment on this kind of questions in social media platforms. Hence, the purpose of the question is to restate the statement that the writer is confident of or happy with.

"Rhetorical interrogation" markers including 難道, 豈, 何必, 何苦 are labelled as *others*. Among these markers, 難道 tends to correlate with *surprise*. All of the tokens of 豈 appear in the structure of "豈 + 不", expressing either positive or neutral emotion (i.e. *happiness* and *surprise*). 何必 and 何苦 are typically used for evoking negative emotions, with the former indicating both *anger* and *sadness*, and the latter *sadness*.

Among the six subtypes of open questions, four of them including *how*, *how many/much*, *what*, and *why* have a strong connection with *sadness*. *Who* question is an indicator of *anger*, *sadness*, and occasionally *happiness* whereas *where* is more evenly used in evoking various emotions (except for *surprise*). *Which* and *when* are rarely used to form rhetorical questions.

---

[1] 有木有 is often found in Weibo as a netizen transformation of 有沒有.

As for *how* questions expressing *sadness*, about one-fifth of the posts are formed with "……有多……", as in (8).

(8) 又是同一家麵館，同一張桌子，同一碗麵。我是有多沒創意。。。多愛念舊。
　　(It's the same noodle restaurant, same table and same bowl of noodles. How dull am I…how nostalgic.)

In (8), the writer grumbles about himself being so dull and nostalgic. Although this pattern can also be used to express other emotions, it is observed that if the subject is the first-person pronoun 我 (I) or in its possessive form 我的 (my), the sentence is likely expressing *sadness*. In (8), the writer is discontented with what he did but is not being able to take control of or willing to change.

W*hat* questions are quite often used for expressing *sadness* and *fear*. We observe that 怎麼辦 demonstrates a close relationship with both *sadness* and *fear*, comparing to the other three emotions. Yet, 怎麼辦 used for the expression of *sadness* and *fear* cannot be distinguished without understanding the situation provided in the context. Examples are exemplified as in (9) and (10).

(9) 吃了八個鍋貼感覺剛開胃怎麼辦！想要更多！
　　(Feels like I just started after eight dumplings, what should I do! Craving for more!)

(10) 睡不著睡不著睡不著怎麼辦 [衰]
　　(Can't fall asleep, can't fall asleep, can't fall asleep, what should I do [sad])

Although both (9) and (10) pose a rhetorical question with 怎麼辦, the expressed emotions are different. (9) is tagged as *fear* because the writer can have some more just if he wants to, and 怎麼辦 only implies that he fears that he may have to take the consequence of having more. (10) is labelled as *sadness* because the writer is suffering from insomnia helplessly. Thus, the use of 怎麼辦 can help identify and distinguish *sadness* and *fear* from the other emotions, but more tokens are needed to investigate what kind of contextual information we need to discriminate *sadness* and *fear*.

Why questions are the most frequently used open questions and approximately 40% of *why* questions are in the form of "why + (…) + 這/那麼". The tendencies of this pattern being used for each emotion in descending order are as follows: *sadness* (38%), *anger* (26%), *happiness* (24%),

surprise (9%), and *fear* (3%). The statistics are of concern to automatic emotion classification. We found that the adjective(s) following 這/那麼 may help determine what emotion the rhetorical question expresses. Consider (11).

(11) 為什麼感覺這麼鬧心這麼煩
　　 (Why am I feeling so hectic and annoyed)

(11) is annotated as *anger* as suggested by the adjectives 鬧心 (hectic) and 煩 (annoyed) following 這麼. It is also observed that if the subject refers to the writer himself, the emotion expressed is likely a negative one, such as *sadness* or *anger*.

6　Conclusion

In this paper, we identify the use of information-seeking questions and rhetorical questions in emotion expression. Results show that 94% rhetorical questions are used to express emotions, while only 23% of information-seeking questions are associated with emotions. We demonstrate that different types of rhetorical questions may have different preferences for a particular emotion. In addition, various syntactic structures of rhetorical questions are proposed for emotion identification. We believe the linguistic account of rhetorical questions in emotion expressions will provide a clearer picture of the nature of emotion.

Acknowledgements

References

Bhattasali, S., Cytryn, J., Feldman, E., & Park, J. 2015. Automatic Identification of Rhetorical Questions. In *ACL (2)* (pp. 743-749).

Davidov, D., Tsur,O., & Ari Rappoport. 2010. Semi-supervised Recognition of Sarcastic Sentences in Twitter and Amazon. In *Proceedings of the 14th Conference on Computational Natural Language Learning*. 107–116.

Ekman, P., & Cordaro, D. 2011. What is Meant by Calling Emotions Basic. *Emotion Review*, *3*(4), 364-370.

Fainsilber, L. and Ortony, A. 1987. Metaphorical Uses of Language in the Expression of Emotions. *Metaphor and Symbol*, *2*(4), 239-250.

Frank, J. 1990. You Call that a Rhetorical Question?: Forms and Functions of Rhetorical Questions in Conversation. *Journal of Pragmatics*, *14*(5), 723-738.

Fussell, S. R. 1992. *The Use of Metaphor in Writ-ten Descriptions of Emotional States*. Unpublished manuscript, Carnegie Mellon University.

Fussell, S. R. & Moss, M. M. 1998. Figurative Language in Emotional Communication. *Social and Cognitive Approaches to Interpersonal Communication*, 113-141.

Gibbs, R. W., Leggitt, J. S., & Turner, E. A. 2002. What's Special about Figurative Language in Emotional Communication. In Fussell, S. R. (Ed.), *The Verbal Communication of Emotions: Interdisciplinary Perspectives*. Mahwah, New Jersey: Lawrence Erlbaum Associates, 125-149.

González-Ibánez, R., Muresan, S., & Wacholder, N. 2011. Identifying Sarcasm in Twitter: A Closer Look. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Vol 2*, 581-586.

Kövecses, Z. 1990. *Emotion Concepts*. New York: Springer-Verlag.

Kövecses, Z. 2003. *Metaphor and Emotion: Language, Culture, and Body in Human Feeling*. Cambridge University Press.

Lakoff, G. & Johnson, M. 1980. *Metaphors We Live by*. Chicago: University of Chicago Press.

Lee, S. Y. M. 2015. A Linguistic Analysis of Implicit Emotions. In Lu, Q. & Gao, H. H. (Eds.) *Chinese Lexical Semantics: 16th Workshop, CLSW 2015, Beijing, China. May 9-11, 2015. Revised Selected Papers (Vol. 9332)*, 185-194. Cham: Springer.

Lee, S. Y. M. 2017. Figurative Language in Emotion Expressions. In Wu Y., Hong JF., Su Q. (Eds.) *Chinese Lexical Semantics: 18th Workshop, CLSW 2017, Leshan, China. May 18-20, 2017. Revised Selected Papers (Vol. 10709)*, 408-419. Cham: Springer.

Leggitt, J. S., & Gibbs, R. W. 2000. Emotional Reactions to Verbal Irony. *Discourse Processes*, *29*(1), 1-24.

Li, Charles N. & Thompson, Sandra A. 1981. *Mandarin Chinese: A Functional Reference Grammar*. Berkeley: University of California Press.

Liu, Feng-hsi. 2000. The Scalar Particle hai in Chinese. *Cahiers de linguistique-Asie orientale*, *29*(1), 41-84.

Petty, R. E., Cacioppo, J. T., & Heesacker, M. 1981. Effects of Rhetorical Questions on Persuasion: A Cognitive Response Analysis. *Journal of Personality and Social Psychology*, *40*(3), 432.

Ranganath, S., Hu, X., Tang, J., Wang, S., & Liu, H. 2016. Identifying Rhetorical Questions in Social Media. In *ICWSM* (pp. 667-670).

Reyes, A., Rosso, P., & Veale, T. 2013. A Multidimensional Approach for Detecting Irony in Twitter. *Language resources and evaluation*, *47*(1), 239-268.

Roberts, R. M., & Kreuz, R. J. 1994. Why do People Use Figurative Language? *Psychological Science*, *5*(3), 159-163.