

# Information-seeking Questions and Rhetorical Questions in Social Media

Helena Yan Ping Lau and Sophia Yat Mei Lee

*The Hong Kong Polytechnic University*

This study examines the distinctive features of information-seeking questions and rhetorical questions from the semantic and syntactic perspectives. 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, 94% of rhetorical questions are found to evoke emotions, while only 23% of information-seeking questions are associated with emotions. It indicates that rhetorical questions are a rather productive means of expressing emotions. In view of this, we explore the characteristics of the two types of questions in terms of their lexical choices and syntactic behaviors. Various linguistic cues and syntactic structures are proposed to distinguish the differences between information-seeking questions and rhetorical questions. It is believed that the study will be directly beneficial to the identification of rhetorical questions and help enhance the existing automatic emotion classification system.

## 1. Introduction

Information-seeking questions (IQs), as suggested by its name, generally aim to make a request for information or for 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 and Kreuz, 1994). As a form of figurative language, rhetorical questions usually convey a more complicated meaning that goes beyond the literal. Although a rhetorical question may lead to a discussion or elicit a response from the interlocutor(s), the speaker who poses the question does not mean to, or not even interested in getting an answer to the question. For example, the rhetorical question *aren't you ashamed of yourself?* conveys the intended meaning of *you ought to be ashamed of yourself* instead of asking whether or not the hearer is ashamed of himself. Given that rhetorical questions are a rather productive means of expressing or evoking emotions, in particular the negative ones (Roberts and Kreuz 1994; Gibbs et al. 2002; Lee 2017), it is believed that being able to distinguish between information-seeking questions and rhetorical questions may be directly beneficial to the identification of emotions. Yet, it has been a challenging task to distinguish the two question types from each other, as both of them have the structure of a question.

This paper attempts to explore how information-seeking questions and rhetorical questions are used in social media in terms of emotion expressions. We examine the distinctive features of information-seeking questions and rhetorical

questions from the semantic and syntactic perspectives, i.e. the lexical choices and syntactic behaviors of the two types of questions. Various linguistic cues and syntactic patterns are proposed which can be used to differentiate the two question types from each other.

## 2. Related Work

Information-seeking questions are typically used to elicit an answer, while rhetorical questions are used to make a statement without expecting a direct answer. With regard to the relation between questions and emotions, previous research has extensively focused on rhetorical questions and only a few studies have been done on investigating information-seeking questions. Quan et al. (2010) analyzed emotion expressions in Chinese at sentence level. They suggested that sentences without the presence of negation marker, conjunction or question mark do not convey any emotions if they do not contain any emotional words, while sentences with the presence of the three items express emotions even if they do not contain any emotional words. They indicated that interrogatives (including both information-seeking questions and rhetorical questions) can be used to express any emotions, in particular the *anxiety* emotion. Lau and Lee (2018) explored the interaction between emotions and both types of questions in social media. They illustrated that approximately 23% of information-seeking questions are associated with emotions, whereas 94% of rhetorical questions are used to express emotions. It reflects the important role rhetorical questions play in emotion expressions in social media.

Generally speaking, rhetorical questions are regarded as an effective persuasive device (Petty, 1981; Frank, 1990). As a form of figurative language, rhetorical questions are often studied in a more general way. Previous studies indicated that figurative language is commonly used to express emotions (Kövecses, 1990, 2003; Lakoff and Johnson, 1980; Fussell and Moss, 1998; Gibbs et al., 2002), especially the intense ones (Fainsilber and Ortony, 1987; Fussell, 1992). The frequent use of figurative language for emotion expressions can partly be due to “the subjective nature of emotional experiences appears to lend itself to figurative expression” (Fussell and Moss, 1998: 113). Roberts and Kreuz (1994) examined the discourse goals of eight types of figurative devices, namely hyperbole, idiom, indirect request, irony, understatement, metaphor, rhetorical question, and simile. They suggested that rhetorical questions are used to express both positive and negative emotions, with the latter being more frequent. Leggitt and 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. Therefore, 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 also illustrated that rhetorical questions are particularly productive in evoking negative emotions, i.e. *sadness* and *anger*. Drawing

from the insight of Lee (2017), Lau and Lee (2018) further explored the use of rhetorical questions in emotion expressions. Various linguistic cues and syntactic structures are proposed for the identification of five different emotions, namely *happiness*, *sadness*, *anger*, *fear*, and *surprise*.

Although researchers have attempted to investigate the syntactic structures of rhetorical questions in Mandarin Chinese (Alleton, 1988; Hsiao, 2017), the proposed patterns may not account for the majority of posts in social media due to the register of text. In view of this, we explore the structures of rhetorical questions, and investigate how they can be distinguished from information-seeking questions by their actual use in social media.

### 3. Corpus Data and Annotation

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

The corpus was annotated with emotions and questions by two annotators (Lee, 2015; Lau and Lee, 2018). Following Lee et al. (2013), 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. Questions were annotated at clause level. In other words, both main interrogatives and embedded interrogatives were identified and tagged. For each post that contains both an emotion and a question, annotators would be asked to determine whether or not the tagged emotion is concerned with the question identified. If so, that means the identified question expresses such an emotion. Otherwise, the question would be regarded as “no emotion”. Consider examples (1) and (2).

- (1) 簡直就是我的寫照，為神馬我那麼苦逼[淚][淚][淚]  
(That is just the portrayal of my life, why am I feeling so miserable [tears][tears][tears])
- (2) 你們懷孕的時候都玩些啥呀。。我好無聊啊。。虛度光陰浪費生命，我心痛哇。。。[淚][傷心]  
(What did you folks play with during your pregnancy... I'm so bored... It's such a waste of time and life, I'm broken-hearted...[tears][sad])

In example (1), the *sadness* emotion is expressed by means of the rhetorical question. Not aiming to elicit an answer, the rhetorical question in (1) is used to make the statement that “I (the writer) am feeling so miserable”. Since the tagged emotion corresponds to the identified question, the emotion tag (i.e. *sadness*) was not removed. As for example (2), the *sadness* emotion is expressed not by the information-seeking question but the remaining parts of the post. In that case, the emotion tag was

removed so that the statistics can truly reflect the interaction between questions and emotions.

#### 4. 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, more than one emotion may be involved in a single post. As for emotion annotation, Figure 1 demonstrates the distribution of each type of emotion. Among the five emotions, *happiness* has the highest frequency (49.74%), followed by *sadness* (25.48%), *anger* (12.44%), *surprise* (7.40%), and *fear* (4.94%).

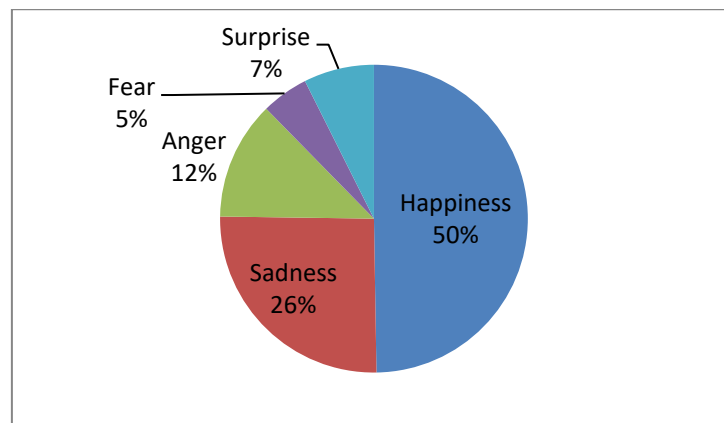


Figure 1 - Distribution of Emotion Types

As for question type annotation, 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 2 shows the distribution of emotions per question type. This is calculated relative to the total number of each question type.

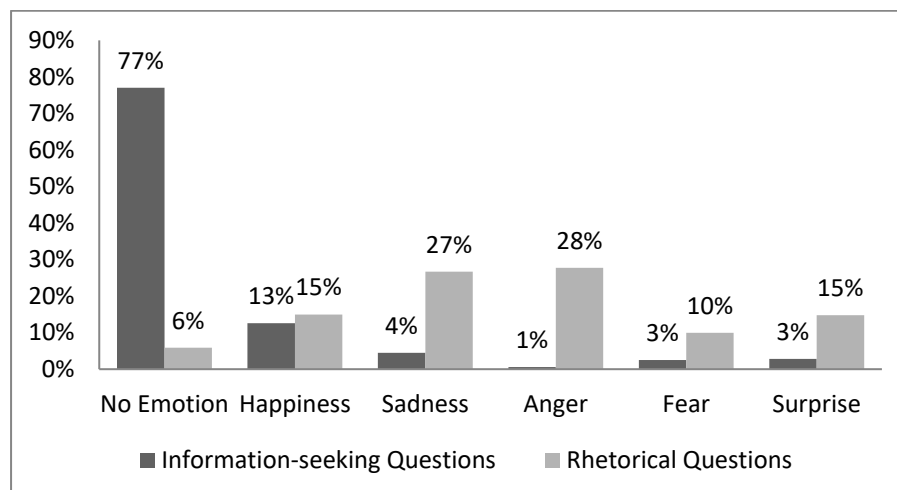


Figure 2 - Emotions Expressed Using IQs and RQs

Figure 2 shows that the vast majority (94%) of rhetorical questions do evoke emotions, whereas only 23% of information-seeking questions are associated with emotions. Of all information-seeking questions, 13% of them express the *happiness* emotions, and the remaining 11% are used to express *sadness* (4%), *fear* (4%), *surprise* (3%), and *anger* (1%). Unlike information-seeking questions, rhetorical questions show a tendency towards negative emotions, in particular *anger* (28%) and *sadness* (27%). In order to support the claim that the differentiation between information-seeking questions and rhetorical questions is important for emotion studies, Figure 3 illustrates the distribution of question type per emotion in all posts.

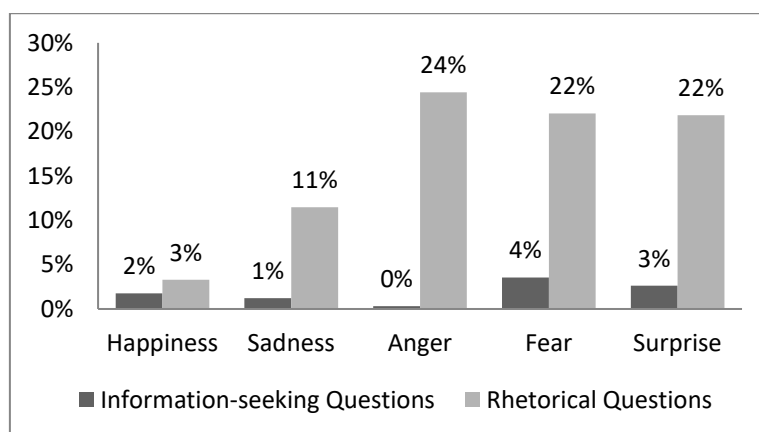


Figure 3 – Distribution of Question Type per Emotion in All Posts

Figure 3 is calculated relative to the total number of posts of a given emotion type. Figure 3 further illustrates that information-seeking questions are not often used to express emotions, while 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. As compared to information-seeking questions, not only are rhetorical questions more often found in social media, they are even more productive in emotion expressions. In the following sections, various distinctive features of the two question types at both lexical and syntactic levels will be proposed.

## 5. The Distinctive Features of IQs and RQs at the Lexical Level

Although both information-seeking questions and rhetorical questions have a structure of a question, they can be distinguished from each other from the semantic perspective. That is, they have different lexical choices in forming a question. In the following subsections, we will explore the linguistic cues for the differentiation in terms of the use of question words, the verb 知道, and the adverb 又.

### 5.1 The Use of Question Words

At the lexical level, we investigate the use of question words in information-seeking questions and rhetorical questions, such as the *wh*-words and sentence-final particles (i.e. 嗎, 呢, and 吧) in terms of their collocational behaviors.

### 5.1.1 The Use of “Why Questions”

Among all the *wh*-words such as *what*, *when*, *how* etc., *why* is the most frequently used form. We compare the occurrence of *why* question in information-seeking questions and rhetorical questions as shown in Table 1.

	怎麼	腫麼	咋	炸	雜	為什麼	為神馬	為嘛	為毛	為何	為啥	幹嘛	Total
IQs	4	1	0	0	0	7	0	0	0	0	4	0	16 (13.22%)
RQs	41	3	10	1	1	31	2	1	1	4	4	6	105 (86.78%)
Total	45	4	10	1	1	38	2	1	1	4	8	6	121 (100%)

Table 1 – The Occurrence of *Why* in Information-seeking Questions (IQs) and Rhetorical Questions (RQs)

Some lexical items in Table 1 are found in the dataset as the internet slangs which are derived from 怎麼 and 為什麼 on account of their phonetic similarity. Given that 怎麼 and its derived forms can also refer to *wh*-words such as *how* etc., we manually read through each token and remove those that are not referring to the meaning of *why* from Table 1. Table 1 shows that netizens prefer the standard forms of question words to the internet slangs in forming an information-seeking question, whereas both forms may be used to form a rhetorical question. Therefore, the use of internet slangs may serve as an indicator for distinguishing the two question types from each other.

Table 1 also illustrates that rhetorical questions are more frequently formed with the word *why*. While 86.78% *why* questions are rhetorical questions, only 13.22% are information-seeking questions. In order to prove that the high occurrence of *why* in rhetorical questions is not because of the large number of rhetorical questions identified, Table 2 demonstrates the distribution of *why* question per question type.

	Occurrence of <i>why</i>	Total no. of questions	Percentage of <i>why</i> question per question type
IQs	16	357	4.48%
RQs	105	543	19.34%

Table 2 – The Occurrence of *Why* Questions

Table 2 shows that about one-fifth of the rhetorical questions are formed with *why*. The frequency of the use of *why* in rhetorical questions is four times higher than that in information-seeking questions. Therefore, we compare how the two question types are formed with the use of *why* in terms of their collocational behaviors. It is observed that “這/那麼” often collocates with the *wh*-word *why* in rhetorical question, but rarely does “這/那麼” co-occur with *why* in information-seeking questions. Table 3 illustrates the frequencies of “*why* + 這/那麼” which are calculated relative to the

total number of *why* questions in each question type and the total number of each question type, respectively.

	Occurrence of “ <i>why</i> + 這/那麼”	Percentage of “ <i>why</i> + 這/那麼” per <i>why</i> question in each question type	Percentage of “ <i>why</i> + 這/那麼” per question type
IQs	1	6.25%	0.28%
RQs	42	40.00%	7.73%

Table 3 – The Occurrence of “*Why* + 這/那麼”

Table 3 shows that the lexical item 這/那麼 is often found to collocate with *why* in rhetorical questions, accounting for 40% of *why* questions in rhetorical questions. As for the co-occurrence of “*why* + 這/那麼” in information-seeking questions, only one token is found, and it comprises 6.25% of *why* questions in information-seeking questions. To show how frequently the pattern is used in each question type, we also estimate the occurrence of the pattern per each question type on a percentage basis. The occurrence of the pattern accounts for approximately 7.7% of the rhetorical questions, and less than 1% of the information-seeking questions. Examples of a rhetorical question and an information-seeking question formed in the pattern of “*why* + 這/那麼” are given in (3) and (4), respectively.

(3) 為什麼感覺這麼開心這麼煩  
(Why am I feeling so hectic and annoyed)

(4) 你最近怎麼這麼多感慨?  
(Why have you been so emotional lately?)

(3) is a rhetorical question that the writer uses to express his/her *anger* emotion. It is observed that people often use this pattern to refer to the state of themselves or other animate or inanimate objects instead of the direct addressee. In most cases, an adjective is usually placed after 這/那麼 to describe or evaluate the situation or emotion state of the writer himself or the object to which he is referring. (4) is the only information-seeking question found in the dataset that are formed with this pattern. Unlike those rhetorical questions formed with this pattern, the second person pronoun 你 is the addressee from whom the writer wants to elicit an answer. Therefore, if the writer is not expecting an answer from the addressee using the second person pronoun, “*why* + 這/那麼” can serve as a distinctive structure which can be used to distinguish rhetorical questions from information-seeking questions.

### 5.1.2 The Use of “怎麼 + Modal Verb”

The *wh*-word 怎麼 conveys the meaning of *why*, *how* and *what* in different contexts. When 怎麼 is used to denote *why* and *how*, it sometimes collocates with modal verbs, such as 可以, 可能, 能, and 會, as indicated in Table 4.

	怎麼 + 可以	怎麼 + 能	怎麼 + 可能	怎麼 + 會	Total	Percentage of “怎麼 + modal verb” per question type
IQs	0	0	0	1	1	0.28%
RQs	7	7	1	2	17	3.13%

Table 4 – The Occurrence of “怎麼 + Modal Verb”

Table 4 shows that 怎麼 is often collocated with certain modal verbs such as 可以 and 能 to form a rhetorical question. In our dataset, only a few tokens containing 怎麼 are found to co-occur with the modal verbs 可能 and 會. Examples of a rhetorical question and information-seeking question formed with this pattern are exemplified as in (5) and (6).

- (5) 這樣吃下去，怎麼能不長胖。  
(If I continue eating like that, how can I not gain weight.)
- (6) 你怎麼會知道！  
(How do you know!)

(5) is a rhetorical question used to state the fact that the writer will definitely gain some weight if he/she continues to eat in that way. (6) is the only information-seeking question that is formed with this pattern. Similar to example (4), the second person pronoun 你 is used in (6) as the direct addressee with whom the writer would like to raise the question. Therefore, if the writer does not use the second person pronoun to seek an answer, the pattern “怎麼 + modal verb” is another unique structure that is mostly used to form a rhetorical question.

### 5.1.3 The Use of “不/沒 + Particle”

In this subsection, we mainly explore the use of three common question particles, namely 嗎, 呢, and 吧 when they collocate with a negation marker. Although both a rhetorical question and an information-seeking question can appear in the structure of “不/沒 + particle”, the former is more frequently formed with this pattern, as demonstrated in Table 5.

	不/沒 + 吧	不/沒 + 嗎	不/沒 + 呢	Total	Percentage of “不/沒 + particle” per question type
IQs	4	0	0	4	1.12%
RQs	5	14	3	22	4.05%

Table 5 – The Occurrence of “不/沒 + Question Particle”

From Table 5, we can see that the structure “不/沒 + particle” is more often used to form a rhetorical question, especially those questions that end with 嗎 and 呢. The



frequent use of negation markers in rhetorical questions is because the markers highlight the statement the writer makes. An example is exemplified as in (7).

- (7) 中國的航班晚點可以不這麼頻繁嗎???!!!  
(Can the flight delays in China not being that often???)

In (7), the writer attempts to express his *anger* emotion by making a point that the flight delays in China are so often with a rhetorical question. This may be the strategy writers use to make for a strong impact on the hearers which cannot be done without the presence of the negation marker.

## 5.2 The Use of the Verb 知道

In this section, we mainly examine the use of the epistemic verb 知道. It is observed in our dataset that 知道 is often collocated with the negation marker 不. The occurrence of the use of 知道 and 不知道 in both question types is illustrated as in Table 6.

	知道	Percentage of 知道 per question type	不知道	Percentage of 不知道 per question type
IQs	12	3.36%	0	0%
RQs	4	0.74%	9	1.66%

Table 6 – The Occurrence of 知道 and 不知道

It is suggested in Table 6 that 知道 can be used to form both information-seeking questions and rhetorical questions, whereas 不知道 can only be used to form rhetorical questions. Hence, 不知道 is a distinctive phrase that can be used to differentiate between the two types of questions. As for the use of 知道, rhetorical questions containing 知道 are either formed with A-not-A structure, as in (8), or the rhetorical interrogation marker 難道, as in (9).

- (8) 媽媽 20 歲多織給她自己的衣服，到現在也不過時哦!。。。是不是她早就知道以後會是個女兒呢?  
(This is the clothes that mum knitted for herself at the age of 20, and it does not look out of date!... Did she know that she would have a daughter in the future?)
- (9) 難道連我媽都知道我是個吃貨!  
(Don't tell me that even my mum knows that I am a foodie!)

As for information-seeking questions containing 知道, they are mainly formed with *wh*-words or the negation marker 不 that serves as a question particle. Examples are exemplified in (10) and (11), respectively.

- (10) 我想知道這個怎麼判  
(I want to know how this is to be judge)
- (11) 這地方知道不?  
(Do you know this place?)

### 5.3 The Use of the Adverb 又

Previous studies indicated that the adverb 又 is used in rhetorical questions to strength the tone of the statement (Zhang 2004; Yin 2008). It is observed in our dataset that 又 appears only in rhetorical questions, as in Table 7.

	又	Percentage of 又 per question type
IQs	0	0%
RQs	11	2.03%

Table 7 – The Occurrence of the Adverb 又

As shown in Table 7, people tend not to use the adverb 又 to form an information-seeking question, but it may be used to form a rhetorical question. Hence, it can serve as an indicator of the rhetorical questions. Consider (12).

- (12) 十個人中便有八個人想自主創業，可是又有誰呢?  
(Eight out of ten people want to run their own business, but who really starts it up?)

In (12), 誰 ‘who’ is a *wh*-definite which is not used to ask who really start up a new business. Instead, the question is used to make a point that most people want to run their own business, but no one ends up doing that. In that case, the adverb 又 does not convey any concrete meaning but strengthens the statement made by the writer.

## 6. Distinctive Features of IQs and RQs at the Syntactic Level

In addition to the distinctive features of information-seeking questions and rhetorical questions at the lexical level, we also explore the structures that can be used to discriminate between the two types of questions. In the following subsections, we discuss three distinctive patterns formed with tag questions.

### 6.1 The Use of “Declarative + 好嗎/麼/嘛”

A tag question in Chinese comprises a tag verb in three interrogative forms, namely V-not-V<sup>1</sup>, V-particle, and Neg-V-particle (Hsin, 2016). A tag question is a quick question with the tag being attached to end of a sentence to ask for a response or confirmation (Chao, 1968; Wang, 1965; Li and Thompson, 1981). In our corpus

<sup>1</sup> V-not-V is named as A-not-A in this study as A-not-A is more commonly used in previous studies.

data, we found that all the questions appear in the structure of “declarative + 好嗎/麼/嘛” are rhetorical questions, as shown in Table 8.

	D + 好嗎	D + 好麼	D + 好嘛	Total	Percentage of “declarative + 好嗎/麼/嘛” per question type
IQs	0	0	0	0	0%
RQs	5	7	4	16	2.95%

Table 8 – The Occurrence of “Declarative + 好嗎/麼/嘛”

Table 8 shows that questions formed in this pattern are not used to seek information. Instead, those questions are used to strengthen the statement the writer asserted. Lau and Lee (2018) suggested that the pattern is frequently found to express the *anger* or *happiness* emotion. Consider (13).

- (13) 我媽說我皮膚像外地人一樣粗…外地人皮膚超好的好嗎！！  
(My mum said my skin is as rough as foreigners...Foreigners have soft skins, okay!!)

In (13), the first sentence indicates that the writer’s mother presupposes that foreigners have rough skins and makes comparison between the foreigners and the writer. However, the writer uses the rhetorical question formed with 好嗎 to refute the connotation assumed by his/her mother. Hence, it is not surprising that the pattern is only found in rhetorical questions.

## 6.2 The Use of “Declarative + A-not-A”

In addition to V-particles, a tag question can also be formed with A-not-A. Hsin (2016) proposed that some verbs in the form of V-not-V (i.e. A-not-A) can be used in a rhetorical way, such as 好不好, 行不行, and 可以不可以. Hsin (2016: 90-91) indicated that these tags display “a rhetoric function of refutation” and “the proposition of the host sentence is expressed with strongest assertion and is regarded as the absolute truth by the speaker, so unquestionable”. Apart from 好不好, other forms such as 有沒有 and 有木有 (the netizen transformation of 有沒有) are also commonly found in rhetorical questions (Lau and Lee, 2018). The occurrence of the pattern of “declarative + A-not-A” is shown as in Table 9.

	D + 好不好	D + 有沒有	D + 有木有	Total	Percentage of “declarative+ A-not-A” per question type
IQs	1	1	1	3	0.84%
RQs	5	1	19	25	4.60%

Table 9 – The Occurrence of “Declarative + A-not-A”

Table 9 illustrates that “declarative+ A-not-A” is more often used to form a rhetorical question than an information-seeking question, with 有木有 being even more

frequently used form. In rhetorical questions containing “declarative + 好不好”, the purpose of the questions is not to seek any information, but to rebut the statement made by the addressee, as in (14).

- (14) 這裡湘菜為主，連油麥菜都是辣的。友友大呼太好吃，以後不要去好親婆那裡吃飯了！多多聽完，白她一眼曰：“呸，什麼話？親婆做菜很拿手的好不好？”  
(Hanun cuisines are mainly served here; even the Indian lettuce is spicy. *Youyou* exclaimed that the food is absolutely delicious, and she is not going to Grandma’s for meals! *Duoduo* glanced at her right after, and said: “Hey, what? Grandma cooks really well, okay?)

Different from 好不好, “declarative + 有沒/木有” is used to restate the statement that the writer strongly believes. Consider (15).

- (15) 自己先贊個字體很漂亮有木有?  
(Let me give myself the thumbs up first! The writing is beautiful, yea?)

In (15), the writer uses 有木有 to “re-confirm” the proposition he/she made. Although an answer may be given in that case, the only answer the writer expects is the positive one, which contrasts with the purpose of an information-seeking question. Besides, the writer indeed hold firm to his/her belief that his/her writing is nice as he gave his/her writing the thumbs up. Therefore, we can make a conclusion that “declarative + A-not-A” are more likely used to form rhetorical questions.

### 6.3 The Use of “A-not-A + Sentence/ Phrase(s)”

Apart from forming a tag question, 有沒/木有 can also precede a sentence or phrase(s) to form a A-not-A question as in (16).

- (16) 有没有什么。。医治疼痛的良方呢？  
(Is there any cure for pain?)

Example (16) is an information-seeking question asking if someone have any recommendations. In an attempt to examine whether or not the position of A-not-A affects the question type, we count the occurrence of the pattern of “A-not-A + sentence/phrase(s)” is shown as in Table 10.

	有沒+ sentence/ phrase(s)	有木有+ sentence/ phrase(s)	Total	Percentage of “有沒/木有 + sentence/ phrase(s)” per question type
IQs	11	6	17	5.04%
RQs	3	0	3	0.55%

Table 10 – The Occurrence of “有沒/木有 + Sentence/ Phrase(s)”

Table 10 demonstrates that if A-not-A comes before the sentence/phrase(s), the question is more likely an information-seeking question. Therefore, we assume that the position of the A-not-A form matters. If A-not-A precedes the sentence/phrase(s), it is probably an information-seeking question; if A-not-A is placed after the sentence, it is presumably a rhetorical question.

## 7. Conclusion

This paper explores the characteristics of information-seeking questions and rhetorical questions from the semantic and syntactic perspectives. The corpus data shows that rhetorical questions (60.3%) are more frequently used in social media as compared to information-seeking questions (39.7%). Moreover, rhetorical questions are more productive in evoking or expressing emotions than information-seeking questions. We propose various linguistic cues and syntactic patterns for the purpose of distinguishing between rhetorical questions and information-seeking questions. At the semantic level, we find that the structures of “why + 這/那麼”, “怎麼 + modal verb”, and “negation + question particle” are more often used to form rhetorical questions than information-seeking questions. We also suggest that the verb 知道 is more commonly found in information-seeking questions than rhetorical questions, and “negation + 知道” is only found in rhetorical questions. It is also observed that the adverb 又 is only used to form rhetorical questions. At the syntactic level, it is suggested that “declarative + tag question (i.e. V-particle/ A-not-A)” are more often found in rhetorical questions. However, if A-not-A comes before a sentence or phrase(s) as in “A-not-A + sentence/ phrase(s)”, the question is likely an information-seeking one.

In sum, we believe the linguistic account of information-seeking questions and rhetorical questions will provide a clearer picture of the features of the two types of questions, and be directly beneficial to the identification of rhetorical questions as well as to emotion studies.

## ACKNOWLEDGEMENTS

This work is supported by a General Research Fund (GRF) project sponsored by the Research Grants Council (Project No. B-Q50Z) and a PolyU Faculty Research Grant (Project no. 1-ZVEK).

## REFERENCES

- ALLETON, VIVIANE. 1988. The so-called “rhetorical interrogation” in Mandarin Chinese. *Journal of Chinese Linguistics*, 278-297.
- CHAO, YUEN REN. 1968. *A grammar of spoken Chinese*. Berkley: University of California Press.
- FAINSILBER, LYNN and ORTONY, ANDREW. 1987. Metaphorical uses of language in the expression of emotions. *Metaphor and Symbol*, 2(4), 239-250.
- FRANK, JANE. 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 written descriptions of emotional states*. Unpublished manuscript, Carnegie Mellon University.
- FUSSELL, SUSAN R. and MOSS, MALLIE M. 1998. Figurative language in emotional communication. *Social and Cognitive Approaches to Interpersonal Communication*, 113-141.
- GIBBS, RAYMOND W., LEGGITT, JOHN S., and TURNER, ELIZABETH A. 2002. What's special about figurative language in emotional communication. *The Verbal Communication of Emotions: Interdisciplinary Perspectives*, edited by Fussell, S. R., 125-149. Mahwah, New Jersey: Lawrence Erlbaum Associates.
- HSIAO, KATHERINE PEI-YI. 2017. Rhetorical wh-questions in Chinese and feature movement. *UST Working Papers in Linguistics*, 9, 1-19.
- HSIN, AI-LI C. 2016. An analysis of Chinese tag questions with a cross-linguistic comparison to English tags. *Taiwan Journal of Linguistics*, 14(1), 69–119.
- KÖVECSES, ZOLTAN. 1990. *Emotion concepts*. New York: Springer-Verlag.
- KÖVECSES, ZOLTAN. 2003. *Metaphor and emotion: language, culture, and body in human feeling*. Cambridge: Cambridge University Press.
- LAKOFF, G. and JOHNSON, M. 1980. *Metaphors We Live by*. Chicago: University of Chicago Press.
- LAU, HELENA YAN PING and LEE, SOPHIA YAT MEI. 2018. Information-seeking questions and rhetorical questions in emotion expressions. *Proceedings of Chinese Lexical Semantic Workshop (CLSW2018)*. May 26-28, 2018. Chiayi County, Taiwan.
- LEGGITT, JOHN S. and GIBBS, RAYMOND W. 2000. Emotional reactions to verbal irony. *Discourse Processes*, 29(1), 1-24.
- LEE, SOPHIA YAT MEI, CHEN, YING, CHU-REN. HUANG, and SHOUSHAN LI. 2013. Detecting emotion causes with a linguistic rule-based approach. *Computational Intelligence, Special Issues on Computational Approaches to Analysis of Emotion in Text*. WileyBlackwell.
- LEE, SOPHIA YAT MEI. 2015 A linguistic analysis of implicit emotions. *Chinese Lexical Semantics: 16<sup>th</sup> Workshop, CLSW 2015, Beijing, China. May 9-11, 2015, (Vol. 9332)*, edited by Qin Lu and Helena Hong Gao, 185-194. Heidelberg: Springer.
- LEE, SOPHIA YAT MEI. 2017. Figurative language in emotion expressions. *Chinese Lexical Semantics: 18th Workshop, CLSW 2017, Leshan, China. May 18-20, 2017, (Vol. 10709)*, edited by Yunfang Wu, Jia-Fei Hong and Qi Su, 408-419. Heidelberg: Springer.
- LI, CHARLES N. and THOMPSON, SANDRA A. 1981. *Mandarin Chinese: A functional reference grammar*. Berkeley: University of California Press.
- PETTY, RICHARD E., CACIOPPO, JOHN T., and HEESACKER, MARTIN. 1981. Effects of rhetorical questions on persuasion: A cognitive response analysis. *Journal of Personality and Social Psychology*, 40(3), 432.
- QUAN, CHANGQIN, HE, TINGTING., and REN, FUJI. 2010. Emotion analysis in blogs at sentence level using a Chinese emotion corpus. *Proceedings of the 6th International Conference on Natural Language Processing and Knowledge Engineering (NLPKE-2010)*, Beijing, 2010, 1-8.

- ROBERTS, RICHARD M., and KREUZ, ROGER J. 1994. Why do people use figurative language?. *Psychological Science*, 5(3), 159-163.
- WANG, WILLIAM S-Y. 1965. Two aspect markers in Mandarin. *Language*, 41:457-70.
- YIN, SHU LIN. 2008. Hai and You in rhetorical questions [試談“還”和“又”的反問用法]. *Nankai Linguistics [南開語言學刊]*, 1, 013. [In Chinese]
- ZHANG PING. 2004. Hai which expresses the mood of rhetorical question and You which emphasizes the mood of rhetorical question [表反問語氣的“還”與加強反問語氣的“又”]. *Journal of Social Science of Hunan Normal University [湖南師範大學社會科學學報]*, 02. [In Chinese]