| RESEARCH ARTICLE

# Co-occurrence of Epistemic Modality in Chinese

## Yiming Jia[1] ✉ and Siyang Xia[2]

**[1]**Ph.D. Candidate, Graduate School of Humanities, Nagoya University, Nagoya, Japan
**[2]**Assistant Professor, BSc in Economics, Nagoya University of Commerce & Business, Nisshin, Japan
**Corresponding Author:** Yiming Jia, **E-mail**: kaimeiask0322@gmail.com

| ABSTRACT

In Chinese, epistemic modality expressions include adverbs, auxiliary verbs, and sentence-final particles. This study focuses on the co-occurrence of various degrees of epistemic modality expressions and discusses the effect of their co-occurrence on the likelihood of sentences. First, we collected data on co-occurrence cases of various expressions from a large-scale corpus and analyzed the characteristics of the distribution of these co-occurrence patterns. Second, we examined how modal adverbs and sentence-final particles of different scales of modality affect the likelihood of sentences when they co-occur. We observed that lower certainty expressions precede higher ones, following a hierarchy of $Mod_{possibility} > Mod_{probability} > Mod_{certainty}$, regardless of their categories. Moreover, the sentence-final particle 吧*ba* diminishes the likelihood of a proposition, while the sentence-final particle （是）…的 (*shi*)…*de* strengthens it. The findings in this study contribute to a deeper understanding of epistemic modality in Chinese and offer novel resources for the acquisition of the language.

| KEYWORDS

Chinese, co-occurrence, epistemic modality, sentence-final particles

## 1. Introduction
The Chinese languages exhibit a variety of strategies for expressing epistemic modality. Although some epistemic modalities may appear semantically incompatible, they can co-occur in a single sentence, as shown in (1).

(1)      **大约**          孔乙己          **的确**      死了。

         probably      Kong Yiji      really      dead
         'Probably Kong Yiji is really dead.'

This research aims to examine the co-occurrence patterns of epistemic modal expressions in Chinese. We analyze the interaction and influence of various modal expressions on the likelihood of the interpretation of a sentence.

Previous studies have analyzed the semantic nuances and distributional patterns of epistemic modal expressions. However, most scholars focus on the usage and functions of modal auxiliary verbs and adverbs. Their co-occurrence and the effects of combining different scales of modal expressions on the likelihood of sentence interpretation have not been thoroughly examined. To address this research gap, we adopt a corpus-based approach using the Center for Chinese Linguistics (CCL) corpus (Pekin University), a large-scale Chinese corpus, to collect authentic language samples and investigate the frequency and distribution of different co-occurrence patterns.

Analyzing the correlation between modal auxiliary verbs and adverbs and their co-occurrence with sentence-final particles (SFPs) can provide valuable insights into how Chinese speakers express and interpret degrees of certainty and uncertainty in their discourse. This knowledge contributes to a deeper understanding of language variation and usage and has the potential to inform language teaching.

This paper is organized into six sections. In Section 2, we provide a comprehensive review of the relevant literature on epistemic modality. Section 3 outlines the corpus-based methodology used to collect and analyze linguistic data from the CCL corpus. We also outline the criteria for selecting modal auxiliary verbs, modal adverbs, and their combinations for our investigation. In Section 4, we present the results of different co-occurrence patterns. Furthermore, we identify the dominant co-occurrence patterns through a quantitative analysis. Section 5 discusses the results of our investigation and analyzes how different combinations of modal adverbs affect the likelihood of interpretation of a sentence. Moreover, we discuss the effect of inserting SPFs on the likelihood of interpretation of the sentence. In Section 6, we provide a conclusion that summarizes the key findings and highlights the contributions and significance of this research.

## 2. Literature Review

Modal expressions have long been discussed across languages in the literature (e.g., Von Wright 1951; Lyons 1977; Chung and Timberlake 1985; Palmer 1986). In English, modal auxiliary verbs such as can, must, and should are generally regarded as the primary resources for expressing modality, while other categories, for example, modal adverbs, are also attracting growing attention (e.g., Quirk et al. 1985; Hoye 1997; Halliday and Matthiessen 2014; Keizer 2018). In Chinese, modal expressions include modal auxiliary verbs, modal adverbs, and SFPs (e.g., Tsang 1981; Tiee 1985; Qi 2006; Xu 2008). This section provides an overview of the previous analyses of various epistemic modal expressions and their co-occurrences.

Palmer (1986) divides modality into two main categories, namely epistemic and deontic modalities. Epistemic modality, which is related to propositions, expresses a speaker's assessment of the empirical status of a proposition. Deontic modality, which is related to events, denotes speakers giving permission or laying an obligation for the event in the future. Hoye (1997) provides a detailed analysis of the co-occurrence of modal auxiliary verbs and modal adverbs. He demonstrates the use of adverbs to emphasize or change the meaning of the modal auxiliary verb in some manner. Adopting Close's (1975) classification of epistemic modal auxiliary verbs, Hoye (1997: 240) categorizes modal adverbs into three groups expressing the epistemic trichotomy of certainty, probability, and possibility, as shown in Table 1.

Table 1. Classification of Epistemic Modal Expressions in English

|  | Possibility | Probability | Certainty |
|---|---|---|---|
| Auxiliary Verbs | *might, may, could, can** | *should, ought to, would, will* | *must, can't* |
| Adverbs | *possibly, conceivably, perhaps, maybe* | *probably, quite, likely, most likely, well*** | *certainly, definitely, indeed, presumably, surely, for certain, of course, undoubtedly, necessarily* |

* Only in rhetorical questions: "Can this be love at first sight?"

** Only after could, may, and might.

Previous studies have mainly focused on three categories of modal expressions in Chinese, namely auxiliary verbs (Tsang 1981; Xu 2008); modal adverbs (see Tiee 1985; Zhang 2000; Shi 2003); and SFPs (see Zhu 1982; Wang 1987; Zhao 2019). For brief discussions on the co-occurrence of these three categories, see Qi (2006), Qi (2007), and Wan (2016).

Zhang (2000: 60–62) divides adverbs into 10 categories, and epistemic modality is classified as "evaluative adverbs," including 的确 dí què, 一定 yí dìng, 必定 bì ding.

As for the subcategories of modality, Shi (2003: 21–22) classifies modal adverbs from a semantic perspective into two categories, namely epistemic modality and deontic modality, based on Lyons' (1977) and Palmer's (1986) work. Epistemic modality is further divided into two subcategories, that is, deductive and speculative. First, epistemic deductive reasoning confirms the factuality of a proposition, and it can be further divided into three sub-categories. These sub-categories are confirmation (e.g., 想必 xiǎng bì), assertion (e.g., 确实 què shí, 的确 dí què), and indication (e.g., 才 cái). Second, epistemic speculation is the speculation of the truth of a proposition, and it can be divided into two subcategories, namely possibility (e.g., 大约 dà yuē, 大概 dà gài) and necessity (e.g., 必然 bì rán, 必定 bì dìng). Regarding the order of sequential use of modal adverbs, Shi (2003:24) proposes that several types of adverbs can be flexible. For example, adverbs indicating possibility can appear before or after adverbs indicating assertion. Shi (2003:25–26) further summarizes four principles of the constraints on the order of modal adverbs, which are the principle of scope, the principle of subjective degree, the principle of coherence, and the principle of prominence.

Xu (2008:265) highlights that seven auxiliary verbs express epistemic modality, which belongs to the three dimensions of epistemic modality. These dimensions included possibility, represented by 可能 kě néng and 能 néng; necessity, represented by 应该 yīng gāi and 该 gāi; and prediction, represented by 要 yào, 得 děi, and 会 huì.

Qi (2006:7) argues that four categories—modal adverbs, modal auxiliary verbs, SFPs, and exclamatory particles—can express mood. These four categories can co-occur in one sentence, forming 15 combinations, including the co-occurrence of modal auxiliary verbs.

Although previous analyses have established a relatively comprehensive understanding of the Chinese modality system, few studies have explored the co-occurrence patterns of various types of modal adverbs and the influence of co-occurrence on semantics. This study examines the co-occurrence patterns of epistemic modal expressions in Chinese through three research questions: (1) What are the characteristics of the distribution of co-occurrence patterns of epistemic modal expressions? (2) How do different scales of modality affect the likelihood of sentences when they co-occur? (3) How do the SFPs affect the likelihood of sentences?

## 3. Methodology

Due to the absence of uniform classification criteria in previous studies on Chinese modal auxiliary verbs and adverbs, this study adopts the classification method of modal auxiliary verbs and adverbs in English (Table I) proposed by Hoye (1997) to compare the co-occurrence of epistemic modal expressions in Chinese. We classify Chinese modal auxiliary verbs and adverbs according to the possibility/probability/certainty scale. Furthermore, we refer to the Xiandai Hanyu Yuqi Chengfen Yongfa Cidian (transl. Dictionary of the Usage of Mood Components in Modern Chinese) to identify the categories of each word selected for investigation and cross-reference the results with previous research. To identify these three epistemic scales, we consult dictionary definitions of the words and incorporate our introspection as native speakers of Chinese. The classification of Chinese epistemic auxiliary verbs and adverbs is listed in Table 2.

Table 2. Epistemic Modal Expressions in Chinese

| | | |
|---|---|---|
| Auxiliary verbs | Possibility (AUX$_{possibility}$) | 可能 *kě néng*，容易 *róng yì* |
| | Probability (AUX$_{probability}$) | 会 *huì*，应该 *yīng gāi* |
| | Certainty (AUX$_{certainty}$) | 必 *bì* |
| Adverbs | Possibility (ADV$_{possibility}$) | 似乎 *sì hū*, 好像 *hǎo xiàng*, 仿佛 *fǎng fú*, 或许 *huò xǔ*, 兴许 *xīng xǔ*, 也许 *yě xǔ* |
| | Probability (ADV$_{probability}$) | 恐怕 *kǒng pà*, 想必 *xiǎng bì*, 势必 *shì bì*, 大半 *dà bàn*, 大约 *dà yuē*, 大概 *dà gài*, 大抵 *dà dǐ*, 多半 *duō bàn* |
| | Certainty (ADV$_{certainty}$) | 绝对 *jué duì*, 确实 *què shí*, 的确 *dí què*, 一定 *yí dìng*, 必然 *bì rán*, 必定 *bì dìng*, 铁定 *tiě dìng*, 肯定 *kěn dìng*, 指定 *zhǐ dìng*, 准 *zhǔn*, 准保 *zhǔn bǎo* |

To analyze how different scales of modality affect the likelihood of sentences when they co-occur, we combine the six subcategories and obtain a total of 12 combinations, namely AUX$_{possibility}$–AUX$_{probability}$, AUX$_{possibility}$–AUX$_{certainty}$, AUX$_{probability}$–AUX$_{certainty}$, ADV$_{possibility}$–ADV$_{probability}$, ADV$_{possibility}$–ADV$_{certainty}$, ADV$_{probability}$–ADV$_{certainty}$, AUX$_{possibility}$–ADV$_{probability}$, AUX$_{possibility}$–ADV$_{certainty}$, AUX$_{probability}$–ADV$_{possibility}$, AUX$_{probability}$–ADV$_{certainty}$, AUX$_{certainty}$–ADV$_{possibility}$, and AUX$_{certainty}$–ADV$_{probability}$. We also examine the reverse versions of the 12 combinations, resulting in a total of 24 combinations. All instances of co-occurrence appear in the same main clause. To avoid any ambiguity, we exclude negations such as 不可能 *bù kě néng*, 不会 *bù huì*, and 不应该 *bù yīng gāi* and interrogative sentences. The results are presented in the next section.

## 4. Results and Discussion

English modal verbs expressing the same modality can only co-occur in a coordinated structure within the same clause (Palmer 1974: 97), as shown in (2).

(2)      John *can* and *will* help his friends.

On the other hand, Chinese modal verbs can be used more freely in co-occurrence. The results of the investigation of the co-occurrence of epistemic modal auxiliary verbs are compiled in Table 3.

| Table 3. Co-occurrence of Epistemic Modal Auxiliary Verbs | | |
|---|---|---|
| | Tokens | Reversed word order |
| $AUX_{possibility}$–$AUX_{probability}$ | 11,844 | 333 |
| $AUX_{possibility}$–$AUX_{certainty}$ | 1 | 2 |
| $AUX_{probability}$–$AUX_{certainty}$ | 0 | 539 |

Table 3 elucidates that Chinese modal auxiliary verbs that express possibility and probability co-occur most frequently (in total, 12,177 tokens). The frequency of the co-occurrence of 可能 *kě néng* and 会*huì*, as an instance of the $AUX_{possibility}$–$AUX_{probability}$ order, is significantly high (96.6%). Moreover, 9,998 out of 11,767 cases of the co-occurrence of 可能 *kě néng* and 会 *huì* are adjacent ones (i.e., 可能会 *kě néng huì*). In the 1,769 non-adjacent cases, other types of adverbs (e.g., 更 *gèng* "more," 就 *jiù* "then," 还 *hái* "also," 将 *jiāng* "will," 突然 *tū rán* "suddenly," and 最终 *zuì zhōng* "ultimately") may appear between 可能 *kě néng* and 大约 *dà yuē*. Furthmore, 可能 *kě néng* may also be used at the initial position of a clause, followed by the subject. The co-occurrence of $AUX_{possibility}$ and $AUX_{certainty}$ is notably infrequent, and the rules of Chinese prosody could be a contributing factor to this contrast. Liu (1990) has proposed that monosyllabic and disyllabic synonyms, including allogeneic synonyms, should generally be combined with words of the same number of syllables in terms of syllable collocation. As the auxiliary verb 必 *bì*, which expresses certainty, is a monosyllabic word, its conjunction with auxiliary verbs that have two syllables, such as 可能 *kě néng* and 容易 *róng yì*, happens less frequently. On the other hand, the combination of 必 *bì* and the monosyllabic auxiliary verb 会 *huì* is frequently observed.

Consequently, it is possible to infer, except in some unusual circumstances, that epistemic modal auxiliary verbs used to express a judgement of a lower degree of likelihood of the empirical status of a proposition tend to occur in positions that precede those of a higher degree of likelihood, as shown in (3).

(3)      $Mod_{possibility} > Mod_{probability}$

The examples of each co-occurrence pattern are as follows:

(4)      $AUX_{possibility}$–$AUX_{probability}$
不　　　吃　　　米饭　　**可能**　**会**　　得　　　糖尿病。
NEG　　　　　eat　　rice　　$AUX_{pos}$　$AUX_{pro}$　get　　diabetes
'You may get diabetes by not eating rice.'

(5)      $AUX_{probability}$– $AUX_{possibility}$
它　　就　　**会**　　**可能**　　危害　　整个　　社会的　稳定　　发展。
It　　then　$AUX_{pro}$　$AUX_{pos}$　endanger　whole　society's　stable　development
'It can endanger the stable development of the whole society.'

(6)      $AUX_{possibility}$–$AUX_{certainty}$
甚至　　连　　读者　　也　　**可能**　**必**　　在　　同样的　不自觉的　矛盾之中。
ADV　　ADV　reader　also　$AUX_{pos}$　$AUX_{cer}$　in　　same　unconscious　contradiction
'Even the reader is probably in the same unconscious contradiction.'

(7)      $AUX_{certainty}$– $AUX_{possibility}$
因此　　他们　**必**　更　　**容易**　误入　歧途。
CONJ　3-PL　$AUX_{cer}$　more　$AUX_{pos}$　go　astray
'So they are more likely to go astray.'

(8)      $AUX_{certainty}$– $AUX_{probability}$
若　　娶　　了　　你，　**必**　**会**　　加倍　爱惜　哩。
SBJV　marry　PRT　2-SG,　$AUX_{cer}$　$AUX_{pro}$　doubly　cherish　SFP

Table 4: Co-occurrence of Epistemic Modal Adverbs

|  | Tokens | Reversed word order |
|---|---|---|
| ADV$_{possibility}$–ADV$_{probability}$ | 13 | 9 |
| ADV$_{possibility}$–ADV$_{certainty}$ | 247 | 61 |
| ADV$_{probability}$–ADV$_{certainty}$ | 144 | 7 |

Table 4 shows the contrast between different word order patterns. Word order patterns with a lower degree of likelihood preceding a higher degree of likelihood are preferred to reversed ones, especially the co-occurrence of ADV$_{possibility}$–ADV$_{certainty}$ (80.2%) and ADV$_{probability}$–ADV$_{certainty}$ (95.4%). Despite a lower ratio, ADV$_{possibility}$–ADV$_{probability}$ is used more frequently than ADV$_{probability}$–ADV$_{possibility}$ (59.1%). The frequency of the co-occurrence of 也许 *yě xǔ* and 的确 *dí què*, as an instance of the ADV$_{possibility}$–ADV$_{certainty}$ order, is higher than other combinations (17.8%). Regarding the ADV$_{probability}$–ADV$_{certainty}$ order, 想必 *xiǎng bì* and 一定 *yí dìng* co-occur most frequently (31.3%).

Similar to the distribution of auxiliary verbs, the co-occurrence patterns of epistemic modal adverbs follow the hierarchy of likelihood in (9). Examples of each co-occurrence pattern are as follows:

(9)     Mod$_{possibility}$ > Mod$_{probability}$ > Mod$_{certainty}$.

(10)    ADV$_{possibility}$–ADV$_{probability}$
**好像**    家里人       **大抵**    都         是         这样。
ADV$_{pos}$    family       ADV$_{pro}$    all         be         same
'It seems to be the same in most families.'

(11)    ADV$_{probability}$– ADV$_{possibility}$
这       失眠        **大概**    **也许**    不是    因为        这个。
this     insomnia     ADV$_{pro}$    ADV$_{pos}$    NEG     because of     this
'This insomnia is probably not because of this.'

(12)    ADV$_{possibility}$–ADV$_{certainty}$
…等等    办法，    **也许**    **一定**    能够    推迟        审判。
other    way       ADV$_{pos}$    ADV$_{cer}$    AUX     postpone      trial
'…and other ways, it may certainly be possible to postpone the trial.'

(13)    ADV$_{certainty}$– ADV$_{possibility}$
这一      规律    **的确**    **似乎**    是       适用    于    这    一纲        的。
this      rule     ADV$_{cer}$    ADV$_{pos}$    be       apply to      this   syllabus      SFP
'This rule does seem to apply to this syllabus.'

(14)    ADV$_{probability}$–ADV$_{certainty}$
**大概**    **准**    是       个       好人。
ADV$_{pro}$    ADV$_{cer}$    be       CL      good person
'Probably he/she will be a good person.'

(15)    ADV$_{certainty}$– ADV$_{probability}$
他的      "精神"    也       **的确**    **多半**    "失落"    了。
His       spirit    also     ADV$_{cer}$    ADV$_{pro}$    lost      PRET
'His 'spirit' is indeed mostly 'lost' as well.'

Table 5. Co-occurrence of Epistemic Modal Auxiliary Verbs and Epistemic Modal Adverbs

| | Tokens | Reversed word order |
|---|---|---|
| $AUX_{possibility}$–$ADV_{possibility}$ | 11 | 310 |
| $AUX_{possibility}$–$ADV_{probability}$ | 4 | 38 |
| $AUX_{possibility}$–$ADV_{certainty}$ | 61 | 136 |
| $AUX_{probability}$–$ADV_{possibility}$ | 12 | 7891 |
| $AUX_{probability}$–$ADV_{probability}$ | 2 | 2568 |
| $AUX_{probability}$–$ADV_{certainty}$ | 38 | 19051 |
| $AUX_{certainty}$–$ADV_{possibility}$ | 0 | 14 |
| $AUX_{certainty}$–$ADV_{probability}$ | 0 | 13 |
| $AUX_{certainty}$–$ADV_{certainty}$ | 1 | 3 |

Table 5 shows that auxiliary verbs are more likely to follow the adverbs, regardless of their meaning. The word order patterns of $ADV_{certainty}$–$AUX_{probability}$ and $ADV_{possibility}$–$AUX_{probability}$ co-occur most frequently. Furthermore, 18,782 out of 19051 cases of the former pattern are co-occurrences of $ADV_{certainty}$ and 会 *huì*. Concerning the latter pattern, 7,139 out of 7,891 cases are $ADV_{possibility}$ and 会 *huì*. The reason why 会 *huì* is used so frequently may be attributed to the fact that it is one of the most frequently used typical modal verbs in modern Chinese (Wang (2002: 18)). Similarly, the combination of auxiliary verbs and adverbs expressing probability confirms this explanation (2,468 out of 2,568 cases of 会 *huì*).

The examples of each co-occurrence pattern are as follows:

(16)     $AUX_{possibility}$–$ADV_{probability}$
世间　　的　　　生意　**可能**　**大抵**　如此。
world　of　　　business　$AUX_{pos}$　$ADV_{pro}$　same
'The business of the world is probably about the same.'

(17)     $ADV_{probability}$– $AUX_{possibility}$
这　　**大概**　**容易**　产生　　腐败。
this　$ADV_{pro}$　$AUX_{pos}$　cause　corruption
'This is probably likely to cause corruption.'

(18)     $AUX_{possibility}$–$ADV_{certainty}$
他　　**可能**　**确实**　就　　是　　杀人犯。
3-SG　$AUX_{pos}$　$ADV_{cer}$　indeed　is　　murderer
'He may indeed be the murderer.'

(19)     $ADV_{certainty}$– $AUX_{possibility}$
肥胖　**确实**　**可能**　影响　心脏　的　　正常　功能。
obesity　$ADV_{cer}$　$AUX_{pos}$　affect　heart　of　normal　function
'Obesity may indeed affect the normal function of the heart.'

(20)     $AUX_{probability}$–$ADV_{possibility}$
你　　也　　**会**　**好像**　没　　看见　的。
2-SG　also　$AUX_{pro}$　$ADV_{pos}$　NEG　see　　SFP
'You will also look like you didn't see it.'

(21)     $ADV_{possibility}$– $AUX_{probability}$

咱们　　　**也许**　　**会**　　　一起　　共事　　　的。
1-PL　　ADV$_{pos}$　AUX$_{pro}$　together　work　　SFP
'We'll probably work together.'

(22)　　AUX$_{probability}$–ADV$_{certainty}$
地球　　温度　　　　　就　　**会**　　**必然**　　增加。
Earth　　temperature　　then　　AUX$_{pro}$　ADV$_{cer}$　increase
'The Earth's temperature will then definitely increase.'

(23)　　ADV$_{certainty}$– AUX$_{probability}$
他　　　**准**　　**会**　　吓跑　　了。
3-SG　　ADV$_{cer}$　AUX$_{pro}$　scare off　SFP
'He will be scared off.'

(24)　　ADV$_{possibility}$ – AUX$_{certainty}$
二者之间　　　　**似乎**　**必**　　有　　　一　　　个　　　是非。
between the two　ADV$_{pos}$　AUX$_{cer}$　have　　one　　CL　　　right and wrong
'There seems to always have a right and a wrong between the two.'

(25)　　ADV$_{probability}$– AUX$_{certainty}$
在党的　　　　　**大概**　**必**　　是　　　少数。
those in the party　ADV$_{pro}$　AUX$_{cer}$　be　　　minority
'Those in the party probably must be a minority.'

## 5. Discussion

In this section, we discuss how different modal adverbs and their combinations affect the likelihood of interpretation of sentences. We also explore the co-occurrence of modal adverbs with SFP 吧 *ba*, which expresses uncertainty, and （是）…的 (*shi*)…*de*, which emphasizes propositions. Furthermore, we discuss how the two SFPs affect the likelihood of interpretation of the sentence.

First, we assumed that the unmarked proposition 孔乙己死了 *Kong Yiji Sile* (Kong Yiji is dead) expresses the highest degree of likelihood or a real-world fact, as shown in (26a). We added epistemic modal adverbs from three categories to the proposition (26b–d) and combined them with different scales (26e–g).

(26)　　a. Unmarked: 孔乙己死了 'Kong Yiji is dead'。
　　　　b. ADV$_{possibility}$: **或许**孔乙己死了。
　　　　c. ADV$_{probability}$: **大约**孔乙己死了。
　　　　d. ADV$_{certainty}$: 孔乙己**的确**死了。
　　　　e. ADV$_{possibility}$–ADV$_{probability}$: **或许**孔乙己**大约**死了。
　　　　f. ADV$_{possibility}$–ADV$_{certainty}$: **或许**孔乙己**的确**死了。
　　　　g. ADV$_{probability}$–ADV$_{certainty}$: **大约**孔乙己**的确**死了。

To visualize the likelihood of these seven cases, we put them all on one continuum, as illustrated in Fig. 1. Instead of a numerical value, we denoted likelihood by a range. We found that, as an unmarked proposition, (26a) denotes a probability of 100%, which can be indicated by 1. Meanwhile, (26b–d) represents an incremental scale of likelihood. As for (26e), the co-occurrence of possibility and probability describes a stronger likelihood than (26b) does. Because of 的确 *dí què*, (26f) and (26g) indicate higher points on the scale of likelihood than (26b) and (26c).
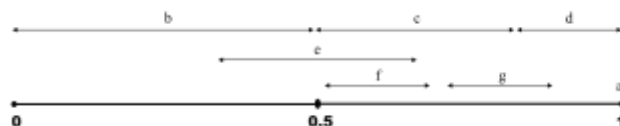


Figure 1. Likelihood scales of modal adverbs.

Second, we considered the co-occurrence patterns of modal adverbs with the SFPs 吧 *ba* and（是）…的 (*shi*)…*de*. Examples with 吧 *ba* are provided in (26a'–g').

(26)　　a'. ??　　孔乙己死了**吧**。
　　　　b'.　　**或许**孔乙己死了**吧**。
　　　　c'.　　**大约**孔乙己死了**吧**。
　　　　d'. ??　　孔乙己**的确**死了**吧**。
　　　　e'.　　**或许**孔乙己**大约**死了**吧**。
　　　　f'.　　**或许**孔乙己**的确**死了**吧**。
　　　　g'.　　**大约**孔乙己**的确**死了**吧**。

We observed that 吧 *ba* diminishes the likelihood conveyed by each modal adverb in (26b'), (26c'), and (26e'–g'). Furthermore, (26a') is more likely to be understood as a confirmation to the addressee rather than as a subjective estimation by the speaker. The combination in (26d') sounds distinctly odd, probably due to the conflict between the certainty of 的确 *dí què* and 吧 *ba*. As mentioned in Qi's (2007: 130) work, modal adverbs and SFPs are more likely to co-occur when they express the same subjective degree. On the other hand, in (26f') and (26g'), although 的确 *dí què* co-occurs with 吧 *ba*, the presence of the adverbs 或许 *huò xǔ* and 大约 *dà yuē* reduces the certainty conveyed by 的确 *dí què* and makes these sentences sound more natural. The overall tendency shifts toward the left in a continuum, as illustrated in Figure 2.
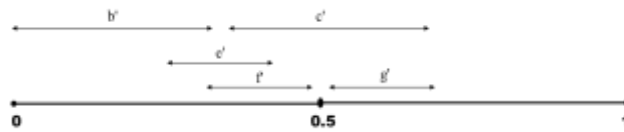


Figure 2. Likelihood scales of modal adverbs with the SFP吧*ba*.

Finally, we considered the co-occurrence patterns of modal adverbs with the SFPs (是)...的 (*shì*)...*de*, as shown in (26a''–g'').

(26)　　a''.　　孔乙己**是**死了**的**。
　　　　b''.　　**或许**孔乙己**是**死了**的**。
　　　　c''.　　**大约**孔乙己**是**死了**的**。
　　　　d''.　　孔乙己**的确是**死了**的**。
　　　　e''.　　**或许**孔乙己**大约是**死了**的**。
　　　　f''.　　**或许**孔乙己**的确是**死了**的**。
　　　　g''.　　**大约**孔乙己**的确是**死了**的**。

All seven cases are acceptable examples of co-occurrence with (是)...的 (*shì*)...*de*, and they serve to reinforce certainty. The overall tendency shifts toward the right in a continuum, as shown in Figure 3.
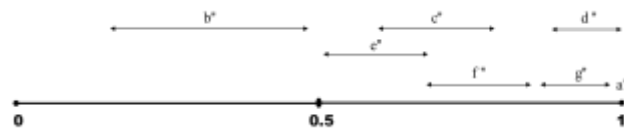


Figure 3. Likelihood scales of modal adverbs with the SFP (是)...的 (*shì*)...*de*.

## 6. Conclusion

This study investigated the patterns of co-occurrence of epistemic modality expressions in Chinese, providing insights into the intricate connection between modal adverbs and SFPs in expressing different levels of certainty and uncertainty. By conducting a thorough analysis of linguistic data from the CCL corpus, we identified prevalent co-occurrence patterns and explored the semantic factors that shape their usage.

According to our findings demonstrate that the co-occurrence patterns of auxiliary verbs and adverbs follow the hierarchy of likelihood $Mod_{possibility} > Mod_{probability} > Mod_{certainty}$. We observed that the occurrence of adverbs of different scales can affect the likelihood of a proposition. Taking SFPs into account, we also found that the co-occurrence with the SFP 吧 *ba* is primarily associated with expressions indicating a lower level of certainty, whereas the SFP (是)...的 (*shì*)...*de*, which emphasizes propositions, reinforces the certainty of the entire sentence.
The findings obtained from this study provide significant insights into the ways in which Chinese language users utilize particular lexical elements to express different levels of epistemic modality. This statement highlights the complex nature of linguistic structures and the ability of specific lexical components to alter the meaning of sentences.

However, it is important to acknowledge that this study does have several limitations. The utilization of the CCL corpus as a source of linguistic data, albeit offering a strong basis, limits our understanding to the patterns and usages that are encompassed within it. Therefore, it is possible that the results may demonstrate a certain level of specificity to the particular situations and dialogues contained within the corpus.

In order to gain a deeper understanding of the versatility and adaptability of the co-occurrence patterns identified in our study, we propose that future research should investigate their potential application in various linguistic situations, including casual spoken conversation and digital communication platforms. Furthermore, a thorough examination of the impact of cultural and social elements on the utilization of epistemic modality expressions has the potential to provide profound contextual understandings of language usage and development.

In conclusion, the findings obtained from this study have the potential to greatly contribute to the field of language education by enhancing educators' understanding of the nuanced use of epistemic modality expressions in the Chinese language. Consequently, this facilitates the creation of enhanced educational resources and instructional approaches that closely coincide with the subtle linguistic nuances and practical applications of these phrases, thus augmenting the effectiveness and pertinence of language instruction.

**Conflicts of Interest:** The authors declare no conflict of interest.
**Publisher's Note**: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers.

## References

[1] Chung, S. & Timberlake, A. (1985). Tense, Aspect, and Mood. In Timothy Shopen (eds) *Language Typology and Syntactic Description, 3: Grammatical Categories and the Lexicon*. Cambridge etc.: Cambridge University Press, 202-258.
[2] Close, R. A. (1975). *A Reference Grammar for Students of English*. London: Edward Arnold.
[3] Halliday, M. A. K. & Matthiessen, C. M. I. M. (2014) *Halliday's Introduction to Functional Grammar*, 4th ed. Abingdon: Routledge.
[4] Hoye, L. (1997). *Adverbs and Modality in English*. London & New York: Longman.
[5] Keizer, E. (2018). Modal Adverbs in FDG: Putting the Theory to the Test. *Open Linguistics*, 4(1), 356-390. https://doi.org/10.1515/opli-2018-0019.
[6] Lyons, J. (1977). *Semantics (I and II)*. Cambridge: Cambridge University Press.
[7] Palmer F. R. (1974). *The English Verb*, London: Longman.
[8] Palmer F. R. (1986). *Mood and Modality*. Cambridge: Cambridge University Press.
[9] Quirk, R., Greenbaum, S., Leech G., & Svartvik J. (1985). *A Comprehensive Grammar of the English Language*, London & New York: Longman, 1985
[10] Tiee, H. H-Y. (1985) Modality in Chinese. In N-K. Kim & H. H-Y. Tiee, (eds) *Studies in East Asian Linguistics*. Los Angeles: Department of East Asian Languages and Cultures, University of Southern California, 84-96.
[11] Tsang, C.-L. (1981) A Semantic Study of Modal Auxiliary Verbs in Chinese,". Ph.D. Dissertation. California: Stanford University.
[12] 刘叔新. (1990). *汉语描写词汇学*. 北京：商务印书馆.
[13] 齐春红. (2007). 语气副词与句末语气助词的共现规律研究. *云南师范大学学报(哲学社会科学版)*(39), 125-130.
[14] 齐沪扬. (2006). 句子中语气成分的同现和交替使用. *中文自学指导*(190), 7-11.
[15] 齐沪扬. (2011). *现代汉语语气成分用法词典*. 北京：商务印书馆.
[16] 史金生. (2003). 语气副词的范围、类别和共现顺序. *中国语文*(292), 17-31.
[17] 万光荣. (2016). *基于语料库的分句语气异类组配研究*. 北京：中国社会科学出版社.
[18] 王振来. (2002). 论能愿动词的语义类别. *辽宁工学院学报(社会科学版)*(01), 16-18.
[19] 徐晶凝. (2008). *现代汉语话语情态研究*. 北京：昆仑出版社.
[20] 张谊生. (2000). 现代汉语副词的性质、范围与分类. *语言研究*(39), 51-63.
[21] 赵春利. (2019). *现代汉语句末助词研究*. 北京：商务印书馆.
[22] 朱德熙. (1982). *语法讲义*. 北京：商务印书馆.
[23] 詹卫东、郭锐、常宝宝、谌贻荣、陈龙. (2019). 北京大学CCL语料库的研制，《语料库语言学》2019年第6卷第1期，总第11辑，71-86.