

Deducing root phenomena from clause sizes: New evidence from question marker embeddability*

Yuyang Liu
Yale University

Yitong Luo
Yale University

1 Introduction

This talk focuses on Mandarin sentence-final question markers *ba*, *ma*, and *meiyou*:¹

- (1) *Declarative base*
Zuotian xia-le yu.
yesterday fall-PFV rain
'It rained yesterday.'
- (2) a. *Ba: Confirmation question*
Zuotian xia-le yu **ba**?
yesterday fall-PFV rain BA
'It rained yesterday, didn't it?'
- b. *Ma: Yes-no question*
Zuotian xia-le yu **ma**?
yesterday fall-PFV rain MA
'Did it rain yesterday?'
- c. *Meiyou: Yes-no question in the perfective aspect*
Zuotian xia yu **meiyou**?
yesterday fall rain MEIYOU
'Did it rain yesterday?'

There is a common practice of placing them in the same syntactic projection:

- (3) HighAttP > LowAttP > SQP > IForceP > OnlyP > SAspP (> TP)
(Pan 2019: 200, adapted)

While *meiyou* is embeddable, *ba* and *ma* are considered to be unembeddable, but recent discoveries along the lines of (4) invite us to reconsider this view.

* We thank Veneeta Dayal for advice; Shifan Chen, Yulin Du, Lingxiao Gao, Xi Huang, Chengwei Liang, Jiamu Zhu & twelve anonymous survey participants for judgments; and Miriam Butt, Richard Luo, Zoltán Szabó, Milena Šereikaitė, Jim Wood, Ka-Fai Yip & Xuetong Yuan for feedback. All errors are our own.

¹ Glossing abbreviations: 1 = first person, 2 = second person, 3 = third person, CLF = classifier, COMP = complementizer, COP = copula, PFV = perfective, PL = plural, POSS = possessive, SG = singular.

(4) *Embedded ma question*

John xiang zhidao [xia yu le **ma** zuotian]_↑.

John want know fall rain LE MA yesterday

‘John wants to know whether it started raining yesterday.’

(Bhatt & Dayal 2020: 1122)

In this talk, we reassess the embeddability of Mandarin question markers and show that the data paradigm can be best accounted for if we allow syntax and semantics to conspire in selectional constraints.

(5) *A glimpse of the proposal*

- a. We split the IForceP into three layers.
- b. We attribute the non-co-occurrence of question markers to semantic mismatches.
- c. We argue that the middle layer triggers perspectival centering.

Roadmap

- ii. Traditional views
- iii. New data
- iv. Two ingredients of the proposal
- v. Conclusion

2 Traditional views**2.1 Existing studies**(6) a. *Zhu’s generalization*

tense particles < imperative/interrogative particles < attitude/emotion particles
(Zhu 1982: 207–214, adapted)

b. *Paul’s cartography*

low complementizers < force complementizers < attitude complementizers
(Paul 2014: 83, adapted)

c. *Pan’s cartography*

SAspP < OnlyP < IForceP < SQP < LowAttP < HighAttP
(Pan 2015: 830, adapted; cf. Pan 2018, 2019, 2022)

Pan’s argument in favor of realizing *ba*, *ma*, and *meiyou* as one single head comes from the observation that they cannot co-occur as in (7–9).²

² Note that (8b) is acceptable under a different use of *ba* (which we label as *ba*₂), and (9b) is acceptable with a homophonous evidential marker that’s typically written using a different character (which we dub *ma*₂).

(7) *Co-occurrence of ba and ma: **

(Han 1995: 102)

a. *Zhangsan shi laoshi **ba ma**?

Zhangsan COP teacher BA MA

b. *Zhangsan shi laoshi **ma ba**?

Zhangsan COP teacher MA BA

(8) *Co-occurrence of ba and meiyou: **

a. *Zuotian xia yu **ba meiyou**?

yesterday fall rain BA MEIYOU

b. *Zuotian xia yu **meiyou ba**?

yesterday fall rain MEIYOU BA

(9) *Co-occurrence of ma and meiyou: **

a. *Tamen yijing chufa **ma meiyou**?

3PL already set.off MA MEIYOU

b. *Tamen yijing chufa **meiyou ma**?

3PL already set.off MEIYOU MA

(Pan 2019: 49)

2.2 Embeddability distinction

(10) *Embedded meiyou question: OK*

[Zhangsan zuotian qu kai hui **meiyou**] hen guanjian.

Zhangsan yesterday go hold meeting MEIYOU very crucial

‘Whether Zhangsan went to hold a meeting yesterday is very crucial.’

(Pan 2019: 101)

These two particles signal an urgent request from the speaker for an answer to the question (Chao 1965, a.o.). The relevant readings can be made more accessible with the aid of expressions like ‘in the end’:

(i) a. Zuotian (daodi) xia yu **meiyou ba₂**?

yesterday in.the.end fall rain MEIYOU BA₂

‘Tell me: did it rain yesterday (in the end)?’

b. Tamen (daodi) yijing chufa **meiyou ma₂**?

3PL in.the.end already set.off MEIYOU MA₂

‘Tell me: did they already set off (in the end)?’

Whether *ba* & *ba₂* and *ma* & *ma₂* should respectively be analyzed uniformly is beyond the range of the data we can adequately address here, and we refer interested readers to Li 2006, Deng 2015, Kim 2019 and so on for detailed discussion.

- (11) a. *Embedded ba question:* *
 *[Hansi shi Deguoren **ba**] de wenti hen guanjian
 Hans COP German BA COMP question very crucial
 Intended: ‘The question whether Hans is German is very crucial.’
 (Pan 2019: 101)
- b. *Embedded ma question:* *
 *[Zhangsan mingtian hui lai **ma**] hen guanjian.
 Zhangsan tomorrow will come MA very crucial
 Intended: ‘Whether Zhangsan will come tomorrow is very crucial.’
 (Pan 2019: 100)

2.3 Lexical account

- (12) a. *ba*: [+root]
 b. *ma*: [+root]
 c. *meiyou*: [−root]

3 New data

3.1 Reevaluating *ma*

Recent findings show that *ma* can be embedded under restricted conditions:

- (13) John xiang zhidao [xia yu le **ma** zuotian]_↑. = (4)
 John want know fall rain LE MA yesterday
 ‘John wants to know whether it started raining yesterday.’
 (Bhatt & Dayal 2020: 1122)

Prompted by Bhatt & Dayal, we add an example involving variable binding (14). We use this and the pronoun behavior (15) as evidence that the embedded question is not a quotation.

- (14) *Embedded ma question:*
 1SG = *speaker and* 3SG = *matrix subject*
 [Mei ge ren]_i dou xiang zhidao [ta_i de fenshu bi wo_j gao **ma**]_↑.
 every CLF person all want know 3SG POSS score than 1SG high MA
 ‘Everyone_i wants to know whether their_i score is higher than mine_j.’
- (15) *Quoted ma question:*
 1SG = *subject of the quoting predicate and* 3SG = *someone else*
 [Mei ge ren]_i dou xiang zhidao: [“ta_j de fenshu bi wo_i gao **ma**?”]_↑.
 every CLF person all want know 3SG POSS score than 1SG high MA
 ‘Everyone_i wants to know, “is their_j score higher than mine_i?”’

The embeddability of *ma* depends heavily on the matrix predicate type. While predicates like ‘want to know’ accept *ma* embedding, other predicates like ‘know’ reject it:

- (16) #Xiaoming_i zhidao [ta_i de fenshu bi wo_j gao **ma**]_↑.
 Xiaoming know 3SG POSS score than 1SG high MA
 Intended: ‘Xiaoming_i knows whether his_i score is higher than mine_j.’

This contrast between ‘want to know’ and ‘know’ is not a rogative-responsive one (Dayal 2023). ‘Forget’ is responsive but patterns like ‘want to know’ and allows *ma* embedding:

- (17) Xiaoming_i wangji-le [ta_i de fenshu bi wo_j gao **ma**]_↑.
 Xiaoming forget-PFV 3SG POSS score than 1SG high MA
 ‘Xiaoming_i forgot whether his_i score is higher than mine_j.’

We dub predicates like ‘want to know’ *wonder*-type and those like ‘know’ *know*-type.

3.2 Reevaluating *ba* and *meiyou*

The embeddability of *ba* and *meiyou* is not sensitive to the contrast between *wonder*-type and *know*-type predicates.

- *Ba* is always unembeddable (18).
- *Meiyou* is always embeddable (19).

(18) *Embedded ba question*

- Wonder-type predicate: **
 *Xiaoming_i xiang zhidao [ta_i de fenshu bi wo_j gao **ba**]_↑.
 Xiaoming want know 3SG POSS score than 1SG high BA
 Intended: ‘Xiaoming_i wants to know whether his_i score is higher than mine_j.’
- Know-type predicate: **
 *Xiaoming_i zhidao [ta_i de fenshu bi wo_j gao **ba**]_↑.
 Xiaoming know 3SG POSS score than 1SG high BA
 Intended: ‘Xiaoming_i knows whether his_i score is higher than mine_j.’

(19) *Embedded meiyou question*

- Wonder-type predicate: OK*
 Xiaoming_i xiang zhidao [ta_i jige **meiyou**].
 Xiaoming want know 3SG pass MEIYOU
 ‘Xiaoming_i wants to know whether he_i passed.’
- Know-type predicate: OK*
 Xiaoming_i zhidao [ta_i jige **meiyou**].
 Xiaoming know 3SG pass MEIYOU
 ‘Xiaoming_i knows whether he_i passed.’

3.3 Reevaluating other question markers

Other question markers considered to surface in IForceP include *bu* and the *wh*-operator (Pan 2019). *Bu* signals a yes-no question when added to the end of a declarative base:

(20) *Matrix bu question*

Ni mingnian hai hui zai lai **bu**?
 2SG next.year still will again come BU?
 ‘Will you still come again next year?’
 (Pan 2019: 43)

Here, we also add A-not-A questions:

(21) *A-not-A questions*

a. *V-not-V question*

Ta **xi-bu-xihuan** zhe ben shu?
 3SG like-not-like this CLF book
 ‘Does he like this book?’
 (Huang 1991: 306)

b. *Shi-not-shi question*

Ni **shi-bu-shi** xihuan yuyongxue.
 2SG SHI-not-SHI like pragmatics
 ‘Is it pragmatics that you like?’
 (Ye 2021: 355)

	<i>Ba</i>	<i>Bu</i>	<i>Ma</i>	<i>Meiyou</i>	Op_{WH}	A-not-A
Noun complement	✗	✗	✗	✓	✓	✓
Clausal subject	✗	✗	✗	✓	✓	✓
Clausal object of <i>know</i> -type predicates	✗	✗	✗	✓	✓	✓
Clausal object of <i>wonder</i> -type predicates	✗	✓	✓	✓	✓	✓

Table 1 Embeddability of Mandarin question markers

4 Two ingredients of the proposal

4.1 First ingredient: Split IForceP proposal

(22) *Mandarin left periphery under the split IForceP proposal*

HighAttP > LowAttP > SQP > HighIForceP > PerspP > LowIForceP > OnlyP > SAspP

a. HighIForceP: *ba* *always unembeddable*

b. PerspP: *bu* and *ma*

c. LowIForceP: *meiyou*, Op_{WH} , and A-not-A *always embeddable*

4.2 Non-co-occurrence of question markers

We build on the Hamblin-Karttunen semantics of questions and assume that a question can be a singleton or nonsingleton set as in [table 2](#).

	LowIForceP	PerspP	HighIForceP
	$p \rightarrow \text{meiyou} \rightarrow \{p, \neg p\}$	$\{p\} \rightarrow \text{bu} \rightarrow \{p, \neg p\}$	d. $p \rightarrow \text{ba} \rightarrow p$
a.	$p \rightarrow \text{A-not-A} \rightarrow \{p, \neg p\}$	c. $\{p\} \rightarrow \text{ma} \rightarrow \{p\}$	
b.	$p \rightarrow \emptyset \rightarrow \{p\}$		

Table 2 Semantic types of the input and output of question markers

(23) *Independent motivations*

- a. [Krifka 2015](#)
- b. [Karttunen 1977](#) and [Dayal 2023](#)
- c. [Dayal 2023](#)
- d. [Yuan 2021](#) and [Ye 2024](#)

Question marker	Output set	<i>Nandao</i>	<i>Ya</i>
<i>ma</i>	singleton	✓	✗
<i>bu</i>	nonsingleton	✗	✓
<i>meiyou</i>	nonsingleton	✗	✓
Op _{WH}	nonsingleton	✗	✓
A-not-A	nonsingleton	✗	✓

Table 3 Distribution of *nandao* and *ya*

Biased question word *nandao* *Nandao* triggers a bias towards the negation of the propositional content ([Xu 2017](#)):

(24) *Nandao*

(The speaker cannot understand why student A didn't go back home in the summer vacation.)

Nandao ni bu xiang jia ma?

NANDAO 2SG not miss home MA

'Don't you miss your home?'

([Xu 2017](#): 50)

(25) Ta chi pingguo bu ya?
3SG eat apple BU YA
'Does he eat apples?'

The distribution of ya is also naturally accounted for if we assume that ya selects a non-singleton set as in [table 3](#).

We interpret the PerspP (i.e., the layer that houses *bu* and *ma*) à la Dayal 2023 in the sense that its head carries a piece of not-at-issue information about the p(otential)-activeness of its complement for a phonologically null perspectival center (PRO) in its specifier position.

- In a matrix question (26a), PRO is bound by the speaker, accessible in a very high position in the left periphery.
- In an embedded question (26b), PRO is bound by an argument of the matrix predicate, such as the matrix subject.

(The question of whether it started raining is p-active from the perspective of the speaker.)

‘Did it start raining?’

(The question of whether it started raining is p-active from the perspective of the matrix subject John.)

'John wants to know whether it started raining.'

(Bhatt & Dayal 2020: 1122, adapted)

The complement question must be p-active for the perspectival center: the sentence radical should not rule out the possibility of the perspectival center's ignorance about the answer to the question as defined in (27).

8

- For *know*-type predicates (29a), the matrix content is contradictory to the not-at-issue information, leading to unacceptability.
- For *wonder*-type predicates (29b), the matrix content goes in the same direction with the not-at-issue information, indicating that the embedded question is actually p-active for the matrix subject.

(27) *Minimal requirement of p-activeness*

$$\forall Q [\forall x [\text{p-active-for}(Q, x) \rightarrow \Diamond \neg \text{know}(x, Q)]]$$

(Dayal 2023: 11)

(28) *Denotation of Persp^o*

$$\llbracket \text{Persp}^o \rrbracket = [\lambda Q. \lambda x: \text{p-active-for}(Q, x). Q]$$

(Dayal 2023: 11)

(29) a. *Know-type predicate: Incompatible with the p-activeness requirement*

$$[\text{know}(J, Q) \rightarrow \neg(\Diamond \neg \text{know}(J, Q)) \rightarrow \neg \text{p-active-for}(Q, J)]$$

#John zhidao [xia yu le **ma**]_↑.

John know fall rain LE MA

Intended: ‘John knows whether it started raining.’

b. *Wonder-type predicate: Compatible with the p-activeness requirement*

$$[\text{want-to-know}(J, Q) \rightarrow \Diamond \neg \text{know}(J, Q)]$$

John xiang zhidao [xia yu le **ma**]_↑.

John want know fall rain LE MA

‘John wants to know whether it started raining.’

(Bhatt & Dayal 2020: 1122, adapted)

Importantly, the p-activeness requirement is not a syntactic selectional restriction.

- This is supported by the contrast of *want to know* and *know*.
- Matrix illocutionary force is also predicted to have an effect on embeddability, which is borne out as in (30).

(30) *Matrix illocutionary force effect on know-type predicates*

a. *Declarative: #*

#Xiaoming_i zhidao [ta_i de fenshu bi wo_j gao ma]_↑.

Xiaoming know 3SG POSS score than 1SG high MA

‘Xiaoming_i knows whether his_i score is higher than mine_j.’

b. *Interrogative: ✓*

Xiaoming_i zhi-bu-zhidao [ta_i de fenshu bi wo_j gao ma]_↑.

Xiaoming know-not-know 3SG POSS score than 1SG high MA

‘Does Xiaoming_i know whether his_i score is higher than mine_j?’

In sum, Persp^o introduces the p-activeness requirement of the complement question for the perspectival center.

- Nothing in the syntax stops *know*-type predicates from taking PerspPs.
- However, *know*-type embedding is ruled out due to a semantic contradiction between the p-activeness requirement and the matrix propositional content.

4.4 Arguing against the lexical account

Given the novel data, a (feature-based) lexical account (as implicated in Paul 2015 or Pan 2019) is untenable:

(31) *Embedding under p-active contexts*

a. *Clausal subject*: *

*[Zuotian xia-le yu **ma**] hai bu qingchu.
 yesterday fall-PFV rain MA yet not clear
 Intended: ‘Whether it rained yesterday is not clear yet.’

b. *Prepositional object*: *

*Xiaoming dui [zuotian xia-le yu **ma**] hen haoqi.
 Xiaoming about yesterday fall-PFV rain MA very curious
 Intended: ‘Xiaoming is very curious about whether it rained yesterday.’

5 Conclusion

Takeaway messages

- Mandarin question particles show three-layer tiered embeddability which corresponds to their positions in the syntactic spine.
- Crucially, novel data regarding *bu* and *ma* suggest that they are Persp^o, which obligatorily introduces a p-activeness requirement.
- Syntax and semantics conspire in licensing question embedding.
- Root phenomena are decomposable: the embeddability of a question marker depends on both the clause size and the p-activeness requirement.

References

Bhatt, Rajesh & Veneeta Dayal. 2020. Polar question particles: Hindi-Urdu *kya*:. *Natural Language & Linguistic Theory* 38(4). 1115–1144. <https://doi.org/10.1007/s11049-020-09464-0>.

- Chao, Yuen Ren. 1965. *A grammar of spoken Chinese*. Berkeley, CA: University of California Press.
- Dayal, Veneeta. 2023. The interrogative left periphery: How a clause becomes a question. *Linguistic Inquiry*. https://doi.org/10.1162/ling_a_00507.
- Deng, Dun. 2015. The syntacticization of illocutionary forces and the root vs. non-root distinction: Evidence from the sentence-final particle *ba* in Mandarin. *Lingua* 162. 32–55. <https://doi.org/10.1016/j.lingua.2015.04.004>.
- Hamblin, Charles Leonard. 1973. Questions in Montague English. *Foundations of Language* 10(1). 41–53. <http://www.jstor.org/stable/25000703>.
- Han, Yang Saxena. 1995. A pragmatic analysis of the *ba* particle in Mandarin Chinese. *Journal of Chinese Linguistics* 23(2). 99–128. <http://www.jstor.org/stable/23756541>.
- Huang, Cheng-Teh James. 1991. Modularity and Chinese A-not-A questions. In Carol Georgopoulos & Roberta Ishihara (eds.), *Interdisciplinary approaches to language: Essays in honor of S.-Y. Kuroda*, 305–332. Dordrecht: Springer. https://doi.org/10.1007/978-94-011-3818-5_16.
- Karttunen, Lauri. 1977. Syntax and semantics of questions. *Linguistics and Philosophy* 1(1). 3–44. <https://doi.org/10.1007/BF00351935>.
- Kim, Juhya. 2019. *Speech act intensification in Mandarin*. Swarthmore, PA: Swarthmore College BA thesis.
- Krifka, Manfred. 2015. Bias in commitment space semantics: Declarative questions, negated questions, and question tags. *Semantics and Linguistic Theory (SALT)* 25. 328–345. <https://doi.org/10.3765/salt.v25io.3078>.
- Li, Boya. 2006. *Chinese final particles and the syntax of the periphery*. Leiden: Leiden University dissertation.
- Pan, Victor Junnan. 2015. Mandarin peripheral construals at the syntax-discourse interface. *The Linguistic Review* 32(4). 819–868. <https://doi.org/10.1515/tlr-2015-0005>.
- Pan, Victor Junnan. 2018. Derivation of the apparent narrow scope of sentence-final particles in Chinese: A reply to Erlewine (2017). *Studies in Chinese Linguistics* 39(2). 99–125. <https://doi.org/10.2478/scl-2018-0004>.
- Pan, Victor Junnan. 2019. *Architecture of the periphery in Chinese: Cartography and minimalism*. London: Routledge. <https://doi.org/10.4324/9781315158228>.
- Pan, Victor Junnan. 2022. Deriving head-final order in the peripheral domain of Chinese. *Linguistic Inquiry* 53(1). 121–154. https://doi.org/10.1162/ling_a_00396.
- Paul, Waltraud. 2014. Why particles are not particular: Sentence-final particles in Chinese as heads of a split CP. *Studia Linguistica* 68(1). 77–115. <https://doi.org/10.1111/stul.12020>.
- Paul, Waltraud. 2015. *New perspectives on Chinese syntax*. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110338775>.
- Xu, Beibei. 2017. *Question bias and biased question words in Mandarin, German and Bangla*. New Brunswick, NJ: Rutgers University dissertation.
- Ye, Shumian. 2021. From maximality to bias: Biased A-not-A questions in Mandarin Chinese. *Semantics and Linguistic Theory (SALT)* 30. 355–375. <https://doi.org/10.3765/salt.v30io.4826>.

- Ye, Shumian. 2024. Pianxiang ji xing wen ju de dongtai yuyi yanjiu [Dynamic semantics of biased polar questions in Mandarin]. Ms., Peking University.
- Yuan, Xuetong. 2021. Extracting commitments: Rising intonation and *ba* in Mandarin. *Sinn und Bedeutung (SuB)* 25. 908–923. <https://doi.org/10.18148/SUB/2021.V25I0.974>.
- Zhu, Dexi. 1982. *Yufa jiangyi [Lectures on grammar]*. Beijing: Comercial Press.

Yuyang Liu
Department of Linguistics
Yale University
370 Temple St
New Haven, CT 06511
yuyang.liu@yale.edu

Yitong Luo
Department of Linguistics
Yale University
370 Temple St
New Haven, CT 06511
yitong.luo@yale.edu