

In-situ or Ex-situ: the Wh-Adjunct *Hé* ‘how, why’ in Old Chinese

Yunhui Gao

A thesis

submitted in partial fulfillment of the
requirements for the degree of

Master of Arts

University of Washington

2026

Committee:

Zev Handel

Andrew Hedding

Gian Duri Rominger

Program Authorized to Offer Degree:

Asian Languages and Literature

©Copyright 2026

Yunhui Gao

University of Washington

Abstract

In-situ or Ex-situ: the Wh-Adjunct *Hé* ‘how, why’ in Old Chinese

Yunhui Gao

Chair of the Supervisory Committee:

Zev Handel

Department of Asian Languages and Literature

This thesis explores the syntactic behavior of the wh-adverbial *hé* ‘how, why’ in Old Chinese by focusing on the constructions “VP-*hé*” and “*hé*-VP”. Unlike the wh-nominal *hé* ‘what’, whose wh-fronting has been well discussed in the literature, the wh-adverbial *hé* receives little attention, especially in the construction “VP-*hé*”. The main innovation of this thesis is to correlate the syntax of *hé* with its reconstructed pronunciation *gʰaj, which motivates a morphological decomposition of *hé* within the framework of feature movement and the Copy Theory of Movement. A concise answer to the question posed in the title is that the *how*-wh-*hé* is an in-situ wh-word but the *why*-wh-*hé* can be analyzed either as an in-situ or an ex-situ wh-word.

Chapter I introduces the motivation for this study, which comes primarily from the fact that wh-adverbials cannot surface sentence-finally in Mandarin, whereas they can do so in Old Chinese.

After a brief literature review, a methodological issue of doing historical generative syntax, namely, how to compensate for the lack of native intuition is discussed.

Chapter II provides descriptive syntax and semantics of the wh-word *hé*, revealing that both types of the wh-adverbial *hé* can surface at either the left or the right periphery. Then, the etymology of *hé* is clarified in a comparative Sino-Tibetan perspective. Some OC-internal evidence supporting the decomposition of *hé* is provided in this chapter.

Chapter III compares four solutions to the structure of wh-adverbial *hé* in “VP-*hé*”: object analysis (Aldridge 2010, Shi 2015), adjunct analysis, predicate analysis, and resultative complement analysis. This study chooses to analyze this kind of *hé* as a wh-adjunct for both empirical and theoretical reasons.

Chapter IV subcategorizes the wh-adjunct *hé* into two types: *why*-wh-*hé* (i.e. high wh-adjunct) and *how*-wh-*hé* (i.e. low wh-adjunct) based on OC-internal evidence (i.e. Intervention Effect) and cross-linguistic evidence from Mandarin and Japanese. This chapter reaches a preliminary conclusion that both types of the wh-adjunct *hé* base-generate at the right periphery but the *why*-wh-*hé* moves in LF to [Spec, CP].

Chapter V starts with a clarification of the necessity of a movement analysis for the *why*-wh-*hé*. To derive a feature movement analysis of *hé*, the remaining part of this chapter focuses on two kinds of question particles: $Q_{[y/n]}$ and $Q_{[wh]}$ in Old Chinese and other languages.

Chapter VI temporarily turns its attention to modern Chinese dialects in order to compensate for the lack of certain syntactic diagnostics in OC. It is assumed that the preverbal Q-particle $kəʔ^5$ in Jiānghuái Mandarin is etymologically related to the OC wh-word *hé*. Since $kəʔ^5$ is demonstrated

to undergo the rightward $\text{Neg}^0\text{-to-C}^0$ movement, *hé* can also reasonably move rightwards. This chapter provides a different solution to the analysis of the *why-wh-hé*, in contrast to the preceding two chapters.

Chapter VII summarizes main findings and conclusions of this thesis and gives further clarifications on some contentious issues.

TABLE OF CONTENTS

Abbreviations	iv
Chapter I Introduction	1
1.1 Old Chinese	1
1.2 Origin of the study	2
1.2.1 a syntactic difference between OC and Mandarin wh-adverbials	3
1.2.2 the <i>how</i> -wh-word in Japhug	6
1.3 A brief literature review	11
1.4 Methodology	15
1.4.1 lack of native intuition	16
1.4.2 near-native intuition	18
 Chapter II Descriptive Syntax and Semantics of the Wh-word <i>Hé</i>	 18
2.1 Interpretational restrictions of <i>hé</i> at periphery positions	19
2.2 Etymology of <i>hé</i>	21
2.3 Distinctions between <i>hé</i> and <i>hú</i>	22
2.4 The polysemy of <i>hé</i> : a modern view	26
 Chapter III Clause-final <i>Hé</i>: Complement, Adjunct or Predicate	 28
3.1 The object analysis	29
3.2 The adjunct analysis	35
3.3 The predicate analysis and multiple wh-questions in OC	39
3.4 The resultative complement analysis	48
 Chapter IV Two Types of Wh-adjunct <i>Hé</i>	 51

4.1 Terminological issues	51
4.2 Reason-wh <i>hé</i> ‘why’ vs manner-wh <i>hé</i> ‘how’	55
4.3 Some inspirations from Japanese	60
4.4 Categorization of wh-words	65
 Chapter V Deriving a Feature Movement Analysis of <i>Hé</i>	 71
5.1 Necessity of a Movement Analysis	72
5.2 Feature Movement	76
5.3 Q _[wh] and Q _[y/n]	78
5.3.1 question particles in OC	79
5.3.2 Q _[wh] and Q _[y/n] in other languages	86
 Chapter VI Relating <i>Hé</i> to the Q-particle <i>Kǎ?</i> in Jiānghuái Mandarin	 89
6.1 Preverbal Q-particles in Chinese dialects	90
6.2 Etymology and a brief history of K-VP questions	92
6.3 Basic data of K-VP questions in Jiānghuái Mandarin	93
6.4 Comparisons among four kinds of yes-no questions	95
6.4.1 matrix and embedded readings	96
6.4.2 compatibility with <i>-ne</i> (or equivalents)	99
6.4.3 compatibility with modals	100
6.4.4 compatibility with preverbal adjuncts	101
6.4.5 compatibility with negative contexts	101
6.4.6 compatibility with the adverb <i>dàodī</i> ‘actually’	102
6.4.7 Intervention Effect	103
6.4.8 embedded null subject question	104

6.5 The Neg ⁰ -to-C ⁰ movement of <i>kəʔ⁵</i> ‘Q’	106
6.5.1 theoretical hypothesis	106
6.5.2 rightward movement of <i>kəʔ⁵</i>	108
6.5.3 correlating negation and interrogative	111
Chapter VII Conclusion and Further Remarks	114
7.1 Conclusion	114
7.2 Clarification on some contentious issues	118
7.2.1 decomposition of <i>hé</i> (*g ⁵ aj)	118
7.2.2 rightward movement	119
References	122

Abbreviations

1	first person
2	second person
3	third person
ACC	accusative
BA	the grammatical word <i>bǎ</i> in Mandarin
CAUS	causative
COMPL	complementizer
CONJ	conjunction
CTH	Clausal Typing Hypothesis
CTM	Copy Theory of Movement
DE	the grammatical word <i>de</i> in Mandarin
DECL	declarative
DM	Distributed Morphology
EMC	Early Middle Chinese
EMPH	emphasize
FACT	factual non-past
FOC	focus
FUT	future
GEN	genitive
HON	honorific

IE	Intervention Effect
INDEF	indefinite
IPFV	imperfective
IRR	irrealis
JM	Jianghuái Mandarin
LCA	Linear Correspondence Axiom
LOC	Late Old Chinese
MC	Middle Chinese
MP	modal particle
NEG	negation
NOM	nominative
NOML	nominalization
OC	Old Chinese
PAST	past
PERF	perfective
POSS	possessive
PROG	progressive
Q _[wh]	wh-question particle
Q _[y/n]	yes-no question particle
RC	resultative complement
REL	relative clause

RM	rightward movement
SENS	sensory
SFP	sentence-final particle
SG	singular
TOP	topic

Chapter I Introduction

Chapter I first discusses a few definitions of Old Chinese (OC) in order to clarify the object of this research. Then, I present some basic data to show the origin of this study, followed by the main hypothesis concerning the morphological decomposition of the OC wh-word *hé*. Subsequently, a brief literature review is provided, and I conclude with a discussion of some relevant methodological issues.

1.1 Old Chinese

OC is a classical language in the Sino-Tibetan family. Chinese, along with the Tibeto-Burman branch, constitutes a top-level branch in the family (Matisoff 2003:1). OC is canonically defined as “varieties of Chinese before the unification of Qín dynasty in 221 BCE” (Baxter and Sagart 2014:1), and more precisely, 11-3 centuries BCE. The lower bound of OC can be extended to the end of the Eastern Hàn dynasty (220 AD), which is also referred to as the Late Old Chinese (LOC) period. Unlike their 2014 definition, Baxter and Sagart (1998) treat OC as “the language of the Chinese classical texts of the Zhōu dynasty [i.e. 11 ca. BCE-256 BCE]”. Also, Handel (2003:544) defines OC as “the Chinese language as spoken in the first half of the first millennium BC”.

It may not be what the authors originally intended to emphasize, but as can be seen from different definitions, some scholars emphasize the nature of OC as once having been a colloquial language, while others focus on its textual attributes. A question that can fundamentally affect the plausibility of this research is to what extent OC texts can reflect the spoken language at that time. Wang (1962 [2017]:1) points out that Classical Chinese (古代漢語) includes two systems 文言

wényán (literary language) and 古白話 *gǔbáihuà* (early vernacular Chinese). Furthermore, *wényán* contains Old Chinese (上古漢語) spoken in the pre-Qín (-221BCE) period as well as the written language after the pre-Qín period that imitates pre-Qín classical texts. Wang (idem.) clearly points out that OC texts are based on the colloquial language of the pre-Qín period. Under Wáng Lì's definition, the current study has as its object the Old Chinese that has its colloquial basis from the pre-Qín period. Therefore, most of the primary sources cited in this study come from the pre-Qín period, but some important Eastern Hàn texts (e.g. *Lùnhéng* 論衡) will also be employed. Ideally, OC texts that do not reflect colloquial language should be excluded from the data. Therefore, making the distinction between colloquial and written languages is of special importance in doing historical generative syntax since it makes no sense to propose syntactic movements for classical texts that had already lost their colloquial basis.

As I have alluded to above, there are apparent genre distinctions between EOC (ca. 11- ca. 6 BCE) and LOC (ca. 5 BCE -ca. 3 AD). EOC texts are represented by Oracle Bone Inscriptions, *Shījīng* (The Book of Odes) and *Shàngshū* (The Book of Documents) that are written in rhyming or succinct styles. LOC texts such as *Zuǒzhuàn* (The Zuo Tradition), *Lúnyǔ* (Analects), *Mèngzǐ* (Mencius), *Guóyǔ* (Discourses of the States) are longer prose narratives or sayings literature, thus naturally contain more dialogue. These textual characteristics make LOC texts a more suitable corpus for doing historical syntax. And without exception, it is in these texts that the syntactic phenomena we will be discussing are most frequently attested.

1.2 Origin of the study

The inspiration of this thesis comes from two sources: one is the conspicuously different syntactic distribution between OC and Mandarin *wh*-adverbials¹, and the other is the *how-wh*-word in Japhug.

1.2.1 a syntactic distinction between OC and Mandarin wh-adverbials

In Mandarin, *wh*-adverbials cannot surface in sentence-final positions, unless there is a pause between them and preceding clauses:

$$(1)^2$$

a. wèishenme	nǐ	zuótiān	méi	lái	xuéxiào?
why	you	yesterday	NEG	come	school
‘Why didn’t you come to school yesterday’					

b. [TopP [TP nǐ zuótiān méi lái xuéxiào]], wèishenme t_{TP} ?
 you yesterday NEG come school why
‘You didn’t come to the school yesterday, why?’

(2) a. **zěnmē** zuò yáng ròu?
 how make lamb
 ‘How do (you) make lamb?’

b. * [TP zuò yáng ròu], zěnmē t_{TP}?
make lamb how
Intended: ‘As for making the lamb, how (to do it)?’

c. wǒ bù lái, zěnmē?

¹ The “wh-adverbial” here is purely a descriptive term referring to the OC wh-word *hé* when it means ‘how’ and ‘why’. Starting from 3.3, I will call them using a more analytical term “wh-adjunct”.

² I use *pinyin* (with characters omitted) to transcribe standard Mandarin, but IPA to transcribe other languages and dialects. In transcribing OC texts, I use *pinyin* plus characters.

I NEG come how

‘What if I just do not come?’

#‘What is the manner x , such that I do not come in x ’

(1a) and (2a) exhibit that *how*-and-*why* wh-words in Mandarin normally appear at left peripheries.

In (1b), TP can be analyzed as scrambled to the domain of the TopP. But topicalization of the complement TP of *zěnmē* ‘how’ is not permitted (2b). Among wh-adverbials, as indicated by the comparison between (1b) and (2b), *zěnmē* ‘how’ is more incompatible with sentence-final positions than *wèishēnmē* ‘why’ because even if there is a pause, (2b) is ungrammatical. Although (2c) is grammatical, *zěnmē* ‘how’ can only trigger rhetorical interpretations in negative contexts.

However, OC wh-adverbials do not have such restrictions:

(3) 其為舟車何？以為車以行陵陸，舟以行川谷。 (Mòzǐ, Jiéyòng)

qí wéi zhōu chē hé?

it make boat carriage why

‘Why did (they) make boats and carriages?’

yǐ wéi chē yǐ xíng líng lù,
because make carriage for travel hill land

zhōu yǐ xíng chuān gǔ.
boat for travel river valley

‘Because (they) made carriages for traveling over hills and land, and boats for traveling through rivers and valleys.’

(4) 且許子何不為陶冶...何許子之不憚煩？ (Mencius, Téngwéngōng Shàng)

qiě Xǔ zǐ hé bù wéi táo yě?

moreover Xu master why NEG do potter founder

‘Moreover, why does Master Xu not work as a potter or a founder?’

hé Xǔ zǐ zhī bù dān fán
 why Xu master GEN NEG afraid.of trouble
 ‘Why isn’t master Xu afraid of troubles?’

(5) 君子于役，如之何勿思？ (Shījīng, Wángfēng)

jūnzǐ yú yì, rú zhī hé wù sī?
 gentleman FUT military.service do it how NEG long.for
 ‘(My) gentleman is going to perform the military service and how could (I) not long for him?’

(6) 子欲居九夷。或曰：陋，如之何！子曰：君子居之，何陋之有？ (Analects, Zihǎn)

zǐ yù jū jiǔyí.
 master wish reside barbarian.tribe
 ‘The Master wished to reside in barbarian tribes.’

huò yuē lòu, rú zhī hé?
 someone say shabby do it how
 ‘Someone said, (that is) shabby, how can (you) deal with it?’

zǐ yuē, jūnzǐ jū zhī,
 master say gentleman reside it
 hé lòu zhī yǒu
 where shabby GEN have
 ‘The Master said, (if) a gentleman resides there, where in can it be shabby?’

(7) 公曰，何如之謂仁？子曰，不淫於色。 (Dàdài Lǐjì, Qiānshèng)

gōng yuē, hé rú zhī wèi rén
 lord say how do it be.called benevolent
 ‘The lord said, what manner of doing it (may) be called benevolent?’

zǐ	yuē,	bù	yín	yú	sè
Master	say	NEG	indulge	in	sensual.desire

‘The Master said, not indulging in sensual desire.’

I provided rather detailed contexts through (3)-(7) to assist readers in examining the interpretations of *hé*. For example, the context in (3), namely *yǐ* ‘because, for’ clearly indicates that *hé* asks about reason. (3) and (4) show that when *hé* means ‘why’, it can either precede or follow a TP. On the one hand, (6) tells us that when it means ‘how’, *hé* can stay at the right periphery; on the other hand, it shows that *hé* can also mean ‘where’. (7) further shows that when *hé* means ‘how’, it can also stay at the left periphery.

In a word, both adverbial uses of *hé* can stay at either left or right peripheries. This descriptive phenomenon raises a few questions: at which periphery does the wh-word *hé* base-generate; are movements needed to explain the syntactic distribution of *hé*; if movements exist, what moves? Essentially, we will be focusing on the structure “VP-*hé*” in OC. The rest of this thesis seeks answers for the questions raised here.

1.2.2 the *how*-wh-word in Japhug

In this subsection, I discuss the *how*-wh-word in Japhug since it suggests the possibility of doing morphological decomposition of the OC *hé* (*g^ʃaj) and further informs the question of what potentially moves in “VP-*hé*”. Intuitively, if a wh-word moves, all of it moves. However, if a wh-word can be morphologically decomposed, a logical possibility is that not every component needs to move.

Japhug is a conservative language in the Gyalrongic branch of Sino-Tibetan with complex morphology. An interesting fact about Japhug is that it does not have an independent word for

‘how’, and naturally, there are apparent morphological correlations between wh-nominals and wh-adverbials. According to Jacques (2021:198-208), some of the wh-words in Japhug are: *tɛ^{hi}* (<*t^{hi}) ‘what’, *ɛu* ‘who’, *t^hɣstuy* (<*t^histuy) ‘how many’, *t^hɣjtɛu* (<*t^hijtɛu) ‘when’, *ɣotɛu³/ɣoj* ‘where’, *tɛ^hindza* ‘why’. As can be seen, the wh-nominal *t(ɛ)^{hi}* ‘what’ serves as the morphological base of ‘how many’, ‘when’ and ‘why’. In Japhug, ‘how’ is expressed with the combination of *tɛ^{hi}* ‘what’ plus verbs *fse* ‘be like’ or *stu* ‘do like’. For example, consider (8), adapted from (36) in Jacques (2021:200):

- | | | | | |
|-----|------------------------------------|------|--------------|-------------------|
| (8) | tchi | zo | tú-wy-stu | p ^h ɣn |
| | what | EMPH | IPFV-do.like | be.efficient:FACT |
| | ‘How should we do for it to work?’ | | | |

It can be understood that ‘how’ in Japhug is formed by the construction ‘what...do’. Interestingly, the OC construction *rú...hé* ‘do...how’ (6) exhibits a similar pattern, namely, a wh-word co-occurs with a light verb ‘do’ and the whole construction asks about manner. The equivalent of the OC light verb *rú* ‘do’ in Mandarin is *zuò*, but Mandarin does not have such a ‘do...wh’ construction. Therefore, the fact that OC wh-adverbials can stay at the right periphery but Mandarin cannot might have something to do with the light verb ‘do’. For example, the OC *rú* ‘do’ possibly functions as an operator that licenses the sentence-final *hé*. However, this possibility is ruled out empirically because verbs that have no semantic correlation with *rú* ‘do’ also appear in the construction “VP-*hé*”, such as *wèi* ‘call’ (III.1).

One thing that Japhug differs from OC is that *tɛ^{hi}* ‘what’ can form rhetorical questions on its own:

³ *Tɛu* is the morpheme for locative. (Jacques 2021:206)

- (9) **azo** **te^{hi}** a-pu-u-ŋu-a
 1SG what IRR-IPFV-be-1SG
 ‘How can it be me?’

Hé can pose rhetorical questions without light verbs when it stays at the left periphery (10), but it needs a light verb *rú* ‘do’ to license rhetorical readings if staying at the right periphery (11):

- (10) 以至仁伐至不仁，而何其血之流杵也？ (Mencius, Jinxīn Xià)

yǐ	zhì	rén	fǎ	zhì	bù	rén,
by	most	benevolent	fight.against	most	NEG	benevolent

ér	hé	qí	xuè	zhī	liú	chǔ	yě?
CONJ	how	3.GEN	blood	GEN	CAUS.flow	pestle	SFP

‘When the most benevolent lord wages a punitive campaign against the least benevolent one, how could (he) have caused the blood of the people to flow and even float pestles of the mortars?’

- (11) 不能正其身，如正人何？ (*Analects*, *Zilù*)

bù	néng	zhèng	qí	shēn,	rú	zhèng	rén	hé?
NEG	be.able	rectify	3.GEN	body	do	correct	others	how

'If one cannot rectify his own person, how can (he) rectify others?'

Jacques (2021:203) points out, in a more complex example, that *te^{hi}* ‘what’ can mean ‘as for how X it becomes’ which further shows the semantic correlation between the *what*-wh-word and the *how*-wh-word in Japhug. But essentially, Japhug does not have an equivalent wh-word to *how* or Mandarin *zěnmě* ‘how’. Unlike the *how*-wh-adverbial, Japhug has an independent *why*-wh-adverbial. See the pair in (12) for instance (idem. 201):

- (12) a. a-tɾɛime, tɕʰi ɲu-tu-nɾɛ ɲu?

1SG.POSS-lady	what	SENS-2-laugh	be:FACT
‘My lady, why are you laughing?’			

b. **te^hindza** ɲuɪ tuɪ ɣɹwu ɲu
 why SENS-2-cry be:FACT
 ‘Why are you crying?’

Both (12a) and (12b) ask about reason, but (12a) uses the *what*-wh-word, while (12b) uses an independent reason-wh *te^hindza*. Etymologically, *te^hindza* ‘why’ is a compound of *te^hi-undza* ‘what-cause’, but its internal structure is synchronically opaque. As in Japhug, OC can also use the *what*-wh-word when asking about reason:

(13) 占之為何？ (Shuōyuàn, Zhèngjiàn)

Zhàn	zhī	wèi	hé?
divine	it	for	what
‘For what purpose do (they) divine it?’			

Cross-linguistically, it is common to use *what*-wh-words to paraphrase wh-adverbials. For example, *why* can be paraphrased by *for what reason*, the Mandarin *wèishénme* ‘why’ can be paraphrased by *wèi-le-shénme* ‘for-PERF-what’, and the Japanese *naze* ‘why’ can be paraphrased as *nani-o* ‘what-ACC’⁴. But on the other hand, languages differ in whether morphological details of the paraphrasing are opaque from a synchronic perspective. For example, the internal structure of Japhug *te^hindza* ‘why’ (< *te^hi-undza* ‘what-cause’) is not clear for ordinary native speakers, but the internal structure of Japanese *nani-o* is uncontroversially clear to ordinary native speakers.

⁴ The paraphrased wh-adverbials can show different syntactic behavior from independent wh-adverbials. 4.3 will address the problem of Mandarin *wèi-le-shénme* ‘for-PERF-what’ and Japanese *nani-o* ‘what-ACC’ in more detail.

In the field of OC phonology, the wh-word 何 *hé* is only reconstructed with one pronunciation *g^ʰaj (Baxter and Sagart 2014). But thinking from the perspective of writing systems, a single character does not need to record only one sound, especially when there are multiple sounds that are semantically correlated. We can do a thought experiment for the Japhug case: how would Japhug speakers use Chinese characters to record their *what/how*-wh-word *tɛ^{hi}* and *why*-wh-word *tɛ^{hi}indza*? A possible way of doing this is to use the character 何 *hé* to record all of these wh-words which is similar to how the character 何 functions in the Japanese writing system. In Japanese, 何 can be read as *naze* ‘why’, *nani* ‘what’ or *nan* ‘what (N)’. Suppose that ancient Japhug speakers used 何 to record *tɛ^{hi}indza* ‘why’ and this lexical item was lost over time while only *tɛ^{hi}* ‘what’ continues to be used in reason questions today. Without horizontal comparisons, historical linguists today would have no way to know that *why*-wh-words in Japhug once had finer morphological structures and would conclude that ‘what’ and ‘why’ are expressed by the same wh-word throughout the history of Japhug, and this is exactly what we normally think of OC wh-words.

My point here is not to doubt the validity of the phonological reconstruction of 何 *hé*, especially because there is no explicit evidence for different pronunciations for *hé* according to its meanings. What I want to address is that based on the empirical facts of Japhug—there isn’t an independent *how*-wh-word and the *why*-wh-word was once morphologically decomposable, also considering the fact that Mandarin wh-adverbials *zěnmě* ‘how’ and *wèishénme* ‘why’ can be traced back no earlier than the Buddhist text 景德傳燈錄 *Jǐngdé Chuándēnglù* (1004 AD-1007AD) composed in the Northern Sòng dynasty (Lü 2017:146, 320), OC probably didn’t have separate wh-words for ‘how’ and ‘why’, thus the multifunctionality of *hé* is reflected in syntax and it might have finer morphological structures. But the decomposition of *hé* does not necessarily mean separating it into two morphemes like *nani-o* ‘what-ACC’ in Japanese. Rather, this decomposition can be realized at

a more abstract level, such as separating two features [wh] and [Q_{y/n}]. I will elaborate on the details of the morphological/featural decomposition of *hé* in Chapter V.

1.3 A Brief Literature Review

The wh-word *hé* has been widely discussed in the literature, especially on the object fronting phenomenon. OC follows a basic word order of SVO, but it exhibits some conditional SOV orders. For instance, interrogative pronouns canonically precede, rather than follow, prepositions or verbs:

(14) 王誰與爲不善 (Mencius, *Téngwéngōng Xià*)

wáng	shuí _i	[PP yǔ t _i]	wéi	bú	shàn?
king	who	with	do	NEG	goodness

‘With whom does the king do what is not good?’

(15) 王者孰謂? (Gōngyáng Zhuàn, *Yǐngōng Yuánnián*)

wàng	zhě	shú _i	[VP wèi t _i]
claim.throne	NOML	who	be.called

‘Who is called one who claims the throne?’

In (14), *shuí* ‘who’ receives its case from a PP-internal position and gets fronted above the preposition. Similarly, *shú* ‘who’ in (15) base-generates in a VP-internal position but surfaces higher. Another noticeable point about (14) and (15) is that wh-words appear after the subject. Aldridge (2010) refers to this phenomenon as “clausal-internal wh-movement”. Meisterernst (2010) also touches upon the wh-fronting of *hé*, which is one of the three types of object fronting phenomena she summarizes. The object-fronting phenomenon in OC was lost in colloquial languages no later than the EMC period (ca. 1 BCE), due to the loss of the genitive particle 之 *zhī* (Aldridge 2023).

Unlike the well-known object fronting phenomenon of *hé*, the core concern of this study—“VP-*hé*” has been rarely mentioned in the literature, regardless of theoretical framework. Aldridge (2010:37-42) touches upon the analysis of “VP-*hé*” when discussing the lack of wh-fronting in OC. In other words, Aldridge (2010) treats *hé* in “VP-*hé*” as an object. Chapter II will discuss Aldridge’s applied object analysis in more detail.

Dong (2009) is one of the few studies that directly concern the “VP-*hé*”. She mainly analyzes the construction “VP, *hé*-yě”, in which *yě* is a sentence-final “focus marker” (Shi 2023:14). Dong’s idea is that “VP, *hé*-yě” is a bi-clausal structure, and she believes that *hé* in “VP-*hé*” is predicative. Dong (2009) does not mention the historical relationship between “VP, *hé*-yě” and “VP-*hé*” in OC. According to my preliminary observation, “VP, *hé*-yě” appeared before “VP-*hé*” and this can be shown by a comparison between (16) and (17):

(16) “古之所以貴此者，何也” (Lǎozǐ, excavated version, Mǎwángduī Bóshū Jiǎběn)

gǔ	zhī	suǒ	yǐ	guì	cǐ	zhě,	hé	yě?
ancient	GEN	REL	by	prize	this	NOML	why	SFP

‘As for why the ancients valued this—what was the reason?’

(17) 古之所以貴此道者何? (Lǎozǐ, transmitted version)

gǔ	zhī	suǒ	yǐ	guì	cǐ	dào	zhě	hé
ancient	GEN	REL	by	prize	this	Dao	NOML	why

‘Why was it that the ancients prized this Dao?’

Both (16) and (17) come from *Lǎozǐ*, but (16) is an excavated version taken from the Mǎwángduī Silk Manuscripts which may be closer to the original text than transmitted versions of *Lǎozǐ*. It can be noticed that the original “VP, *hé*-yě” later changed to “VP-*hé*”. In the EOC period, although

“VP-*hé*” is attested, as shown in (18), *hé* in this case is a wh-nominal and thus does not function the same as that in (17):

- (18) 謂之何哉？ (Shījīng, Bèifēng)
 wèi zhī **hé** zāi
 call it what SFP
 ‘What (then) can (I) call it?’

Dong (2009:213) clearly points out that *hé* in 其稱人何 *qí-chēng-rén-hé* ‘it-call-human-why’ (‘Why does it call someone human?’)⁵ is in the same clause as things it questions⁶. The “VP-*hé*” construction can still be found in Middle Chinese texts, according to Zhao (2023). Zhao (2023:43) points out 所以者何 *suǒ-yǐ-zhě-hé* ‘REL-by-NOML-what’ (‘What is the thing that someone relies on?’) probably comes from the “merging” between two clauses *suǒ-yǐ-zhě* and *hé*. Although Dong (2009) and Zhao (2023) both clarify that “VP-*hé*” is a monoclausal structure, it is worth explaining further why “VP-*hé*” does not contain two clauses, especially given that the related construction “VP, *hé-yě*” undoubtedly contains two clauses. This question is important because it bears on the fundamental issue of this research—if “VP-*hé*” really contains two clauses, the subsequent movement analysis would become meaningless. Next, I use concrete examples to show that “VP-*hé*” as a whole is an independent clause.

In the first place, *hé* can stand on its own to create a clause:

- (19) 對曰：請令魏王可。君曰：何？ (MǎwángduīBóshū, Zhànguó Zònghéngjiāshū)
 duì yuē, qǐng lìng Wèi wáng kě. Jūn yuē: **hé?**
 response say please let Wei king assent Lord say why

⁵ Taken from 公羊傳 *Gōngyáng Zhuàn*, *Yǐngōng Sinián*.

⁶ The original text is “這樣的‘何’也可以說與疑問對象處於同一個小句中了”.

‘(He) replied, please have the King of Wei assent. The lord said, why?’

However, whether *hé* can stand on its own or not is insufficient to prove whether “VP-*hé*” constitutes one or two clauses. At the same time, there is clear evidence showing that *hé* is in the same clause as its preceding elements:

(20) 王者所以有社稷何? (Báihǔtōng Délùn)

[CP [VP [DP wàng zhě suǒ yǐ] yǒu shèjì] hé]
 claim. kingship NOML REL by have altars.of soil.and.grain how

‘How did those who claim kingship come to possess altars of soil and grain?’

a. ‘What is the manner *x*, such that by *x*, those who claimed kingship come to possess altars of soil and grain?’

#b. ‘What is the thing *x*, such that those who claimed kingship and possessed altars of soil and grain have *x*?’

In traditional grammars, *wàng-zhě-suǒ-yǐ-yǒu-shèjì* (‘the fact that people who claimed throne own their countries’) is said to be an incomplete sentence, and this proves that (20) must contain only one clause. Comparing (20) to (17) and (18), we notice that if there is no nominalization marker *zhě* preceding *hé*, a sentence-final *hé* cannot be interpreted as wh-nominal, as shown by the interpretation (b) in (20).

“VP-*hé*” is included in more complex structures in (21)-(23):

(21) 說而不繹，從而不改，吾未如之何也已矣。 (Analects, Zǐhǎn)

yuè	ér	bù	yì,	cóng	éi	bù	gǎi,
pleased	but	NEG	examine	follow	but	NEG	amend
[CP1 wú	[CP2 wèi	rú	zhī	hé	yě]	yǐ	yǐ].
I	NEG	do	it	how	SFP	stop	SFP

‘If (they) are pleased yet do not examine (it), and follow (others) yet do not amend (themselves),
I can do nothing with such people.’

(22)子曰：不曰「如之何如之何」者，吾未「如之何」也已矣。(Analects, Wèilínggōng)

zǐ	yuē,	bù	yuē	rú	zhī	hé	rú	zhī	hé
master	say	NEG	say	do	it	how	do	it	how
zhě,	wú	wèi	rú	zhī	hé	yě	yǐ	yǐ.	
NOML	I	NEG	do	it	how	SFP	stop	SFP	

‘The Master said, as for those who don’t say how to do it, how to do it, I can do nothing
with such people.’

(23) 夫不深料秦之不奈我何也。(ZhànguóCè, Qícè)

fú	bù	shēn	liào	[_{DP} Qín	zhī	bú	nài	wǒ	hé]	yě.
COMPL	NEG	deep	consider	Qín	GEN	NEG	do	us	what	SFP

‘(I) did not deeply consider that Qin could do nothing against us.’

In (21), *rú-zhī-hé* ‘do-it-how’ as a whole serves as the complement of the negation *wèi*, and there are two CPs marked by two SFPs *yě* and *yǐ* respectively. In (22), the first two *rú-zhī-hé* ‘do-it-how’ are used as direct quotations and thus exclude the possibility that the clause-final *hé* constitutes an independent clause from *rú-zhī* ‘do-it’. In (23), *nài-wǒ-hé* ‘do-I-how’ is embedded in the object DP thus *hé* certainly cannot be separated from *nài-wǒ* ‘do-I’.

We’ve seen in section 1.3 that although “VP-*hé*” has more or less been mentioned in the literature, its syntactic structure is not clear. And by examining Dong (2009) and Zhao (2023), I excluded the possibility that “VP-*hé*” contains two clauses.

1.4 Methodology

This section deals with some methodological issues in doing historical/diachronic generative syntax in which I focus on two correlative questions: one is the feasibility of doing such research

without native intuition in the same sense as researching on modern languages, and the other is how to compensate for the lack of such intuition.

The first thing that needs to be clarified is the distinction between “historical syntax” and “diachronic syntax”. The word *diachronic* was first used in linguistics by F. de Saussure in the 1910s, referring to historical “development” of a language⁷. Lightfoot (1979:5) also states that one of the fundamental prerequisites of doing diachronic syntax is “to compare grammars of at least two stages of a language”. The word *historical*, on the other hand, does not necessarily indicate the “change of state”. My focus in this thesis lies not in the change of wh-movement in the history of Chinese. Rather, I focus on wh-movements in a specific period of Chinese, namely Late Old Chinese. The general purpose of this research is to explore “human mental processes and intellectual capacity” (Chomsky 1965: 46) in LOC, which is not so different from research on modern languages. The only difference is that the corpus I employ is restricted and does not have ungrammatical sentences. Therefore, this is a study of historical syntax rather than diachronic syntax. Therefore, my explanations will focus on synchronic expositions instead of on motivations of diachronic changes.

1.4.1 lack of native intuition

The thorniest question in doing historical generative syntax might be the lack of native intuition and we are unable to test native speakers’ “acceptability”, which “belongs to the study of performance” (Chomsky 1965:11). However, it was not long after Chomsky’s *Syntactic Structures* (1957), which revolutionized the way syntax can be studied, that linguists started to reflect on the question of whether a linguistic theory relies heavily on native intuition and grammaticality tests

⁷ https://www.oed.com/dictionary/diachronic_adj?tab=meaning_and_use#6909640.

can be applied to the study of classical languages. In her article applying the generative rules to diachronic changes of English auxiliaries, Closs (1965: 403, footnote 6.) discusses the feasibility of “using this notion (i.e. Chomsky 1957) of grammar in historical analysis”. Closs points out that a historical study of grammar needs to be based not only on observable data but also on unobservable data. This means we can and indeed need to predict the existence of unobservable (both grammatical and ungrammatical) data in doing historical syntax. Therefore, she concludes that “deviance as well as grammaticality can tentatively be made explicit”. The importance of predicting unobservable data is also mentioned in the field of OC phonology, as Baxter and Sagart (2014:4-5) state that in linguistic reconstructions, “prediction” is “the deductive part of the whole hypothetico-deductive approach”.

Closs (1965:403, footnote 6.) mentions a notion she called “secondary intuition” which refers to linguists’ intuition of dead languages “after several years of close association with the language”. In other words, linguists should first make use of all the observable data and when necessary, use their secondary intuition to predict which sentences should be ungrammatical or less acceptable to native speakers of a particular period. In introducing early generative approaches to syntactic changes, Lightfoot (1979:30) expressed a similar view in claiming that “philological judgment is required to ascertain whether a non-occurring sentence-type would in any case be rare in the kinds of texts available”. But Aldridge’s series of generative studies of OC syntax almost never use the term “ungrammatical”. Instead, she rigorously keeps using the terms “not attested” (Aldridge 2010:21), “unattested” (Aldridge 2023:9) or “have not found” (Aldridge 2023:22). Based on Closs (1965) and Lightfoot (1979) mentioned above, I assume it to be reasonable to predict ungrammatical or less acceptable sentences in the study of historical syntax, though they are secondary evidence bearing less importance to observable data. However, this assumption is not

only based on the fact that linguists have secondary intuition but also rests upon what I call “near-native intuition” in the next subsection.

1.4.2 near-native intuition

The generative approach holds an “abrupt” and “contingent” (rather than “gradual”) view towards syntactic change, and E-language is where syntactic change originates (Madariaga 2017:2). During the process of language acquisition, I-languages are also subject to change diachronically (Chomsky 2005). As Aldridge (2010:46) points out, “the only diachronic difference between Archaic and modern Chinese is the loss of the strong focus feature driving movement of VP-internal wh-phrases”. It means that even if it’s the same language, native speakers’ mental grammars change over time⁸. Then it should be reasonable to assume that a native speaker’s intuition is more likely to approximate that of historical speakers when they lived closer in time to the period in which the texts were composed. Luckily, there are many LOC texts that were annotated and commented on not too long after the text was composed, some in the LOC period itself, while some in the EMC period. And these ancient philologists’ paraphrases of the original texts can be regarded as “near-native intuition”. The importance of “near-native intuition” will be discussed in 3.3 when dealing with multiple wh-questions in OC.

Chapter II Descriptive Syntax and Semantics of the Wh-word *Hé*

⁸ According to the “uniformitarian hypothesis” of UG, Universal Grammar is invariant across time, while diachronic change consists in the resetting of parameters in the mental grammars of successive generations of native speakers (Roberts 2007:456). Although this hypothesis holds that native speakers of the same language have the same UG across time, it’s still safe to say that ancient people who lived closer to the LOC period generally have better “intuition” to OC than we native speakers of modern Chinese.

We have glimpsed some basic usages of the multifunctional wh-word *hé* in Chapter I (I. 3-I. 7). The most important descriptive fact is that when *hé* functions as wh-adverbial ‘how, why’, both usages can surface either at the left or the right periphery. In this chapter, I discuss usages of *hé* in more detail, revealing that although the wh-adverbial *hé* can occur relatively freely at both peripheries, each periphery poses some interpretational restrictions on *hé*. Then, I clarify the etymology of *hé* in a comparative Sino-Tibetan context, after which another OC wh-word 胡 *hú* will be compared to *hé* to provide an OC-internal argument to decompose *hé*. Lastly, I reflect on the polysemy of *hé* from modern languages and propose that “*hé*-VP” base-generates at the left periphery only when the modal *qí* follows it, yielding a reading similar to English *how come*. Generally speaking, Chapter II lays a philological foundation for subsequent theoretical analyses and it bridges chapters I and III.

2.1 interpretational restrictions of *hé* at periphery positions

A broad tendency of the wh-adverbial *hé* is that when it surfaces at the left periphery, it is more natural to mean ‘why’ and when at the right periphery, more natural to mean ‘how’. (I3) is the only example I found in which *hé* means ‘why’ but occurs at the right periphery, and (I7) is one of the few examples in which *hé* occurs at the left periphery and clearly asks about manner (i.e. means ‘how’). This means that left peripheries tend to interpret the wh-adverbial *hé* as ‘why’, while right peripheries tend to interpret it as ‘how’. When it occurs at the left periphery, it is common to find *hé* co-occurring with a demonstrative *qí* ‘that’:

(1) 夫子聖者與？何其多能也？ (Analects, Zihǎn)

fūzǐ	shèng	zhě	yú?	hé	qí	duō	néng	yě?
Confucius	sage	NOML	Q	how	that	many	ability	SFP

‘Is Confucius a sage? How can (he) be that versatile?’

(2) 何其所臨之民有饑寒不得衣食者?

(*Shuōyuàn, Jūndào*)

hé	qí	suǒ	lín	zhī	mín	yǒu
how	that	REL	govern	GEN	people	have
jī	hán	bù	dé	yī	shí	zhě?
starvation	cold	NEG	obtain	clothing	food	NOML

‘How is it that, among the people (he) governs, there are those who suffer starvation and cold and cannot obtain clothing or food?’

(3) 何許子之不憚煩?

(*Mencius, Téngwéngōng Shàng*)

hé	Xǔ	zǐ	zhī	bù	dān	fán
why	Xu	master	GEN	NEG	troubled	labor

‘Why is Master Xu not troubled by such labor?’

曰，百工之事，固不可耕且為也。

bǎi	gōng	zhī	shì,	gù	bù	kě	gēng	qiě
hundred	craftsman	GEN	affair	originally	NEG	be.able	plough	while
wéi	yě							
do	SFP							

‘It says, the affairs of a hundred craftsmen, originally cannot be done while one is ploughing.’

In (1) and (2), when there is *qí* ‘that’, “*hé*-VP” reads more like a rhetorical question (i.e. *how come* VP), especially in negative contexts. (1) probably can still be understood as asking the manner of how Confucius became that versatile, while (2) definitely is not asking about the manner in which the people came to suffer but expresses the speaker’s blame on the ruler. It becomes clearer if (1) and (2) are compared to (3). Notice that there is no *qí* ‘that’ in (3), and the responsive clause explains the reason why the Master Xu is not afraid of troubles which clearly indicates that

the question asks about reason. So, *hé* in “*hé-qí*-VP” is different from both the pure *how*-wh-adverbial and the *why*-wh-adverbial; rather, it should be interpreted as *how come* which forms rhetorical questions.

Recall from the Mandarin data in (I.1) and (I.2) that *zěnmé* ‘how’ is more incompatible with sentence-final positions than *wèishénme* ‘why’. In contrast, we have seen above that in OC, it’s ‘why’ instead of ‘how’ that is more incompatible with sentence-final positions. On the one hand, Mandarin and OC are similar in that both of their wh-adverbials show some periphery sensitivities. On the other hand, a fundamental difference between them is that Mandarin wh-adverbials strongly disfavor sentence-final positions while OC wh-adverbials can occur relatively freely in sentence-final positions. Therefore, we not only need to explain the generative process of wh-adverbials in OC but also need to parameterize the distinction between OC and Mandarin.

2.2 etymology of *hé*

According to Chen (2006), the interrogative pronoun *hé*, reconstructed as *g^faj, may have already existed in the Oracle Bone Inscriptions (ca. 13 BCE)⁹, the earliest Chinese texts. However, in OBI, the character 何, which later recorded the wh-word *hé*, was mostly used to record proper nouns (Zhao (1999 [2009]:76, Chen 2006). Chen (2006) finds only two examples of *hé* which he identifies as interrogative pronouns, one of which is:

- (4) 王夢兄戊，何從？ (Jiǎgǔwén Héjī, 17378)
- | | | | | | | |
|------|-------|---------|-----|------------------------|----------------------|-------------------------|
| wáng | mèng | xiōng | Wù, | hé _i | [_{VP} cóng | <i>t_i</i>]? |
| king | dream | brother | Wu | what | follow | |
- ‘The king dreamed of Brother Wu; what (should he) follow?’

⁹ Whether the interrogative pronoun *hé* exists in Oracle Bone Inscriptions is controversial, see the introduction of Chen (2006) for details.

(4) shows that in the earliest attested usages of *hé*, it also gets fronted to preverbal positions, just as in canonical LOC texts. About 500 years later, the interrogative pronoun *hé* became more widely attested in texts from the Warring States period (475 BCE-221 BCE) such as *Shījīng* and *Analects* :

(5) 天實為之，謂之何哉？ (Shījīng, Běimén)

tiān	shí	wéi	zhī,	wèi	zhī	hé	zāi?
heaven	truly	do	it	call	it	what	SFP

‘It was Heaven that truly did it; what can (I) say of it?’

(6) 君使臣，臣事君，如之何？ (Analects, Bāyì)

jūn	shǐ	chén,	chén	shì	jūn,	rú	zhī	hé?
ruler	command	minister	minister	serve	ruler	do	it	how

‘(When) a ruler commands his ministers and ministers serve (their) ruler, how to do it?’

Since *hé* may already be attested in the earliest Chinese texts, which represent a period approximately 2500 years after Sinitic and Tibetan languages diverged (Sagart et al. 2019, Fig 2.), it is not too surprising that *hé* is a cognate to Tibetan. According to Yu (1989:61, 63), OC 何 *hé* *ʔa is a cognate to Tibetan *gal* while 胡 *hú* *ʔa ‘what, why’¹⁰ is a cognate to Tibetan *ga*.

2.3 distinctions between 何 *hé* and 胡 *hú*

The reconstruction that Yu (1989) employs assumes that the two wh-words 何 *hé* and 胡 *hú* are homophones in OC, but Baxter and Sagart (2014)’s system reconstructs them differently: 何 *gʰaj vs 胡 *gʰa. B&S’s reconstruction correctly captures some important syntactic differences between *hé* and *hú*.

¹⁰ Yu (1989) employs a different reconstruction system than this study. I here keep the reconstructions in Yu’s paper unchanged.

Like *hé*, *hú* can form reason questions (7) and rhetorical questions (8). *Hú* can also be a VP-internal wh-argument and get fronted (9). But we never find *hú* occurring sentence-finally:

(7) 胡不歸？ (Shījīng, Bèifēng)

hú bù guī?
 why NEG return
 ‘Why not return?’

(8) 同始異終，胡可常也？ (Zuǒzhuàn, Shàogōng Qīnián)

tóng shǐ yì zhōng, **hú** kě cháng yě?
 alike begin different end how can constant SFP
 ‘(If they) begin alike but end differently, how can (this) be constant?’

(9) 胡以備之？ (Guǎnzǐ, Shěmí)

hú yǐ bèi zhī?
 what use guard.against it
 ‘With what shall (one) guard against it?’

Phonologically, *hé* (何 *g^haj) and *hú* (胡 *g^ha) differ minimally in codas (i.e. *-j vs *-∅). From a pure phonological view, such a distinction may reflect ancient dialectal variations, but if we correlate this distinction with different syntactic distributions (i.e. *-j can appear sentence-finally vs *-∅ never appears sentence-finally), this minimal pair can be explained in a homogeneous phonological system. I have alluded to the idea in Chapter I that *hé* (*g^haj) should be decomposed based on features, now the distinction between *hé* and *hú* provides morphological details of this decomposition. Since the coda *-j poses syntactic restrictions on wh-words and it is part of the final *-aj, it’s natural to separate the initial *g^h- from the final *-aj and assign them with different features to explain the syntactic behavior of *hé* (*g^haj). I leave the question of what features *g^h- and *-aj bear to Chapter V, when proposing the feature movement analysis of *hé*.

Although they follow different syntactic restrictions, it is not controversial that *hé* and *hú* are historically related. An ancient scholar 鄭玄 *Zhèng Xuán* (127 AD-200 AD), who lived in the transitional period of LOC and EMC, gave a paraphrase of (7) using *hé* as 君何不歸乎 *jūn-hé-bù-guī-hū* ‘lord-why-NEG-return-Q’ which indicates that *hú* may have disappeared in colloquial languages no later than *Zhèng Xuán*’s period. Both Matisoff (1995:74)¹¹ and Schuessler (2007:35, 273) correlate the pronunciations *hé* and *hú*, and they call *hé* an independent form while *hú* a dependent form. According to Schuessler (2007:35), the dependent form is phonologically realized as an open syllable (i.e. *g^ʰa) while in semantics and syntax, it “indicates incompleteness of the utterance and the expectation of a following element or answer” (idem). Schuessler’s statement can be naturally adapted to my analysis above in the sense that since *hú* (*g^ʰa) is a dependent form, it cannot stay in sentence-final positions.

Besides *hé* (*g^ʰaj), there is another wh-word 奚 *xī* (*g^ʰɛ) that shows different syntactic behavior from *hé*:

(10) 子奚不為政？ (Analects, *Wéizhèng*)

zǐ	xī	bù	wéi	zhèng
you.HON	why	NEG	take.part	politics

‘Why do you, sir, not take part in politics?’

(11) 衛君待子而為政，子將奚先？ (Analects, *Zǐlù*)

wèi	jūn	dài	zǐ	ér	wéi	zhèng,
Wei	lord	await	you.HON	CONJ	take.part	politics

¹¹ Matisoff (1995:74) glosses 胡 *hú* as ‘interrogative particle’, but I have found no cases where *hú* can be used sentence-finally as a yes-no-question particle like 乎 *hū*. Since Matisoff also refers to 歟 *yú* as “interrogative particle”, which is truly a yes-no-question particle, the term “interrogative particle” couldn’t have been misunderstood when Matisoff uses it to refer to 胡 *hú*. Interestingly, if *hú* in OC can really be used as both a wh-word and a yes-no-question particle, it will be beneficial to the derivation in Chapter VI. And there are indeed languages in which a wh-word and the Q are homophones. For example, ‘what’ in Nepali is *ke*, and *ke* is also used as a yes-no question particle.

‘The ruler of Wei awaits you, sir, to take part in politics; what will you put first?’

To better understand the syntax-phonology correlations revealed by *hé* and *hú/xī*, we can take a look at a similar phenomenon in Túnxī dialect (Lu 2018:123), a variety of the Huīzhōu dialect (Sinitic). In Túnxī Huī, singular personal pronouns make a distinction between weak and strong forms¹², as they are shown in table 1:

	weak form	strong form
1.SG	ɑ ²⁴	ɑ ²⁴ -le
2.SG	ŋ ⁴⁴	ŋ ⁴⁴ -le
3.SG	k ^h ə ⁴⁴	k ^h ə ⁴⁴ -le

(12) laʔ⁵-ka⁴² α?
which-CL Q

25

‘Who is it?’

α²⁴-le/*α²⁴/*α²³(cliticized)

1.SG

‘Me.’

- (13) α²⁴ tɕ^hio¹¹ xα²⁵ k^hə⁴⁴-le/*k^hə⁴⁴/* k^hə⁰
 1.SG fear 3.SG
 ‘I fear him/her.’

The different syntactic restrictions of weak and strong personal pronouns in Túnxī dialect parallel the contrast between *hé* and *hú* such that the dependent form *hú* (*g^ɛa) never appears sentence-finally and since it is prosodically less prominent, it should always be followed by other elements. The empirical facts of *hé/hú* and Túnxī Huī personal pronouns, as well as the Japhug wh-adverbials all point to the hypothesis that *hé* has a finer morphological structure that results in its different syntactic distributions, and this raises an interesting theoretical question of what moves and what is left in-situ.

2.4 the polysemy of *hé*: a modern view

We have seen in Chapter I the multifunctionality of *hé* such that it can be interpreted as ‘what’, ‘where’, ‘why’ or ‘how’. Wh-words in modern languages are also commonly multifunctional. Cheng (2021) points out that both Mandarin and Cantonese use ‘what’ and ‘how’ to express causal/reason questions. (14) shows a similarity between Mandarin and Cantonese that their *what*-wh-words can ask about reason:

- (14) a. Cantonese lei⁵ haam³ **mat**¹ aa³
 2.SG cry what SFP

‘Why are you crying?’

(15) shows that in Mandarin, *zěnmē* ‘how’ can be interpreted as ‘why’ when preceding a modal (Tsai 2008):

‘Why/how come that you will know?’

qí	lǐ	nì	yǐ,	hé	kě	wèi	yì?
3.GEN	principle	reverse	SFP	how	can	call	moral

‘Its principle is reversed, how come/why it can be called moral?’

Although both Mandarin and Cantonese use sentence-initial *wh*-words to form causal/reason questions, they differ in an important fact that Cantonese uses its *what*-*wh* word *mat*¹ while Mandarin uses its *how*-*wh* word *zěnmé* (Cheng 2021:220) to do so:

- (17) a. **mat**¹/*dim² lei⁵ hai² dou⁶ haam³ ge²?
 what/*how 2.SG PROG cry SFP
 ‘Why/how come you are crying?’
- b. **zěnme**/*shénme nǐ zài kū?
 how/*what 2.SG PROG cry
 ‘Why/how come you are crying?’

On the theoretical side, Cheng (2021:244) assumes that like *wat* ‘what’ in Dutch, the Cantonese *mat*¹ ‘what’ base-generates at the left periphery, and *mat*¹ heads an exclamative sentence. Adopting the idea of Cheng (2021) and regarding wh-adverbials in OC, Mandarin and Cantonese, the commonality is that their *how*-and-*why* wh-words are correlated; also, *hé* base-generates at the left periphery when there is the modal *qí*, and in this scenario, *hé* asks about neither pure reason nor manner, which is just like the English *how come*.

Chapter III Clause-final *Hé*: Complement, Adjunct or Predicate?

Beginning in Chapter III, we turn to the core discussion of this study, namely the syntactic structure of “VP-*hé*”. I will first discuss Aldridge (2010) and Shi (2015)’s object analysis of the clause-final *hé*, pointing out some empirical and theory-internal problems with this analysis. Then, I put forward my adjunct analysis, after which two alternative treatments for the clause-final *hé*, the predicate and resultative complement analyses, will be considered. This study chooses the adjunct analysis as the preferred solution.

3.1 the object analysis

Aldridge (2010) and Shi (2015:43) both treat *hé* in *rú-zhī-hé* ‘do-it-wh’ and (1) as an object¹³:

(1) 秦伯曰， 國謂君何？對曰，小人感，謂之不免，君子恕，以為必歸。(Zuǒzhuàn, Xīgōng Shíwǔnián)

qínbo	yuē,	guó	wèi	jūn	hé?	duì	yuē,
Qinbo	say	state	call	lord	what	reply	say
xiǎorén	qì,	wèi	zhī	bù	miǎn,	jūnzǐ	shù,
common.people	distressed	call	it	NEG	escape	gentleman	charitable
yǐ	wéi	bì	guī				
think	be	definitely	return				

‘Qinbo said, what does the state call their ruler? The reply said, common people feel distressed and think that the ruler cannot escape (from blame), while gentlemen are charitable and think that the ruler definitely will return.’

In (1), the reply indicates that the question asks about neither reason nor manner. So, *hé* in (1) is uncontrovertibly a wh-argument. I glossed *rú-zhī-hé* as ‘do-it-wh’ because in some cases, for example (1), *hé* serves as an argument, whereas in others, it serves as a wh-adverbial. (2) and (3) are also examples taken from Shi (2015:43):

(2) 吾甚恐，如之何則可？ (Mencius, Liánghuìwáng Xià)

wú	shèn	kǒng,	rú	zhī	hé	zé	kě?
I	greatly	scared	do	it	how	then	fine

‘I am greatly scared. How shall I do, then, to make it fine.’

(3) 一薛居州，獨如宋王何？ (Mencius, Téngwéngōng Xià)

yī	Xuē Jūzhōu,	dú	rú	sòng	wáng	hé
----	-------------	----	----	------	------	----

¹³ Aldridge (2010:37) refers to this *hé* as a “lower object”.

one Xue Juzhou alone do Song king how
‘How could one Xue Juzhou alone do with the King of Song’

The interpretation of *hé* in (2) and (3) is less clear than that in (1). On the one hand, the context makes it possible to interpret *hé* as asking about the manner of dealing with the speaker’s fear (2) and the king of Song (3). On the other hand, the main verb in (2) and (3) is *rú* ‘do’, but in (1) it is *wèi* ‘call’. Shi (2015:43) analyzes (2) and (3) as double object constructions, but the problem is, *rú* typically acts as a two-place predicate. Aldridge (2010:37) also notices that when *rú* takes two internal arguments (i.e. as a three-place predicate), the “second internal argument is always a wh-word, and this wh-word never undergoes fronting”. In contrast, *wèi* ‘call’ can take two objects even when there is no wh-argument, for example *wèi-zhī-xué* ‘call-it-learned’, *wèi-zhī-cán* ‘call-it-cruel’ (Aldridge 2010:35). When the verb *rú* acts as a two-place predicate, it means ‘resemble’ (4) or ‘go to’ (5) rather than ‘do’:

(4) 執轡如組，兩驂如舞。 (Shījīng, Zhèngfēng)

zhí pèi rú zǔ, liǎng cān rú wǔ
hold rein resemble weaving.thread two side.horse resemble dancer
‘Holding the reins as if weaving thread, the two side horses move as dancers.’

(5) 夏，公如齊逆女。 (Shuōyuàn, Xiūwén)

xià gōng rú qí nì nǚ
summer duke go.to Qi welcome bride
‘In summer, the Duke went to Qi to welcome (his) bride.’

This implies that if “*rú*-N-*hé*” (‘do-N-wh’) really is a double object construction, *rú* ‘do’ only acts as a three-place predicate when the lower argument is a wh-word. Shi (2015:43) explains this phenomenon from a functional perspective and believes that this has something to do with the

position of focus, but it does not answer why a transitive verb *rú* becomes ditransitive when it takes a wh-argument. Although a single verb can change its valency and transitivity is a continuum (Hopper and Thompson 1980), for example, *drink* can either be a one-place predicate in *I do not drink* or a two-place predicate in *I like drinking coffee*, it's hard to explain why a verb changes its valency only when there is a wh-argument.

Aldridge (2010:41-42) clearly specifies that “V-N-*hé*” is a low applicative structure in which *hé* is a THEME. Although Shi (2015) and Aldridge (2010) both treat this kind of clause-final *hé* as an object, their solutions differ fundamentally—Shi's solution implies that “V-N-*hé*” is generated in a one-step way, but Aldridge believes that the object is “applied” and the main verb takes an ApplP as complement. Paul and Whitman (2009) distinguish “thematic applicative” and “raising applicative” and they point out that the former is like high applicative and the latter like low applicative, as they are defined by Pylkkänen (2008). According to Paul and Whitman (2009), in thematic applicative, there is no GOAL argument, but only BENEFICIARY; in raising applicative, there is a GOAL and a transfer of the THEME. However, it's apparent that Paul and Whitman (2009)'s analyses do not apply to at least (2) and (3).

Double object constructions (e.g. *I gave him a book*) are typical raising/low applicatives and activity verbs such as *give* clearly specify a transfer of the THEME *a book*. In (2) and (3), since the verb *rú* ‘do’ is not a typical activity verb, the so-called THEME object *hé* does not undergo any transfer. Therefore, even if “*rú*-N-*hé*” is really a low applicative, it is different from the canonical double object construction type. A semantic problem of the object analysis is that, since THEMES are entities that “undergo actions or are moved, experienced, or perceived” (Carnie 2021:246), it's hard to say that *hé* in (2) and (3) receives its theta-role as THEME. *Hé* in (1) functions as a wh-argument thus can still be considered as a THEME, albeit marginally. While in (2) and (3), since it

asks about manner, it does not fall within the traditional definition of THEME. Also, *hé* in (2) and (3) cannot be regarded as either GOAL or BENEFICIARY, ruling out the possibility that “*rú*-N-*hé*” is a high applicative. Pylkkänen (2008:18) subcategorizes low applicative into two types: recipient applicative and source applicative, and *hé* in (1)-(3) apparently cannot be regarded as either RECIPIENT or SOURCE.

Even if we leave aside the problem of *hé*, (2)-(3) may not be considered as applicatives. Pylkkänen (2008:18) points out that low applicatives “relate a recipient or a source to an individual that is the internal argument of a verb”. However, *zhī* ‘it’ in (2) and *Sòng wáng* ‘the king of Song’ in (3) may not be internal arguments of the verb *rú* ‘do’ because we cannot find examples in which the verb *rú* means ‘do’ and it takes a THEME object. The point here is that even if “*rú*-N-*hé*” (‘do-N-wh’) is an applicative, it is N, rather than the wh-word *hé*, that should be regarded as the applied object because *zhī* ‘it’ and *Sòngwáng* ‘the king of Song’ are more like recipients of the action ‘do’.

Next, I use a pair of sentences (6) and (7) to clarify why the NP directly following *rú* ‘do’ is more qualified to be an applied object than the wh-word *hé*:

(6) 為天下得人者謂之仁。 (Mencius, *Téngwéngōng Shàng*)

[_{DP} wèi tiānxià dé rén zhě]_i wèi zhī_{i/*j} rén
for all.under.heaven obtain worthy.men NOML call it benevolent

‘Those who obtain worthy men for All-under-Heaven are called benevolent.’

#‘Those who obtain worthy men for All-under-Heaven call someone else benevolent.’

(7) 可謂仁乎? (Analects, *Yōngyě*)

pro kě [_{VP} wèi *t*_{pro} rén] hū?
can call benevolent Q

‘Can it be called benevolent?’

In (6), *zhī* ‘it’ must co-index with the subject DP, as indicated by the two potential interpretations, and this seems to imply that there is a *pro* in (7). When the *pro* is assigned a case by *wèi* ‘call’, it moves to the left periphery to be topicalized. The structure DP-*wèi-zhī-rén* in (6) is parallel to the *pro-wèi-t_{pro}-rén* in (7). In overt syntax, the adjective *rén* ‘benevolent’ can serve as the complement of *wèi* ‘call’ (i.e. *wèi-rén*), but we never find the verb *wèi* ‘call’ taking a single recipient as its complement, namely, there are no cases of *wèi*-DP. When *wèi* ‘call’ is followed by a recipient DP, it appears in two structures, as shown in (8) and (9):

(8) 司馬子魚謂君曰…

(*Shuōyuàn, Zhǐwǔ*)

sīmǎ zǐyú	wèi	jūn	yuē...
Sima Ziyu	call	lord	say

‘Sima Ziyu Sima Ziyu call the lord, saying...’

(9) 孔文子何以謂之文也？

(*Analects, Gōngyěcháng*)

kǒngwénzǐ _i	hé	yǐ	wèi	zhī _i	wén	yě
Kongwenzi	what	by	call	it	Wen	SFP

‘On what grounds was Kong Wenzi called Wen?’

In (8), *wèi* ‘call’ is followed by a recipient *jūn* ‘lord’ and another verb *yuē* ‘say’ introducing direct quotations. In (9), when *wèi* ‘call’ takes two DP complements, the pronominal complement *zhī* ‘it’ needs to co-index with the subject DP. (6)-(9) aim to demonstrate that if “*wèi*-N-*hé*” is a low applicative, N should be the applied object. This analysis also applies to the verb *rú* since when it means ‘do’, there are only examples like *rú-hé* ‘do-what’, but no examples like *rú*-N.

The applied object analysis is also problematic from a morphological point of view. Based on the tree in Aldridge (2010:42), her analysis should analyze *wèi-jūn-hé* ‘call-lord-what’ as Figure 1:

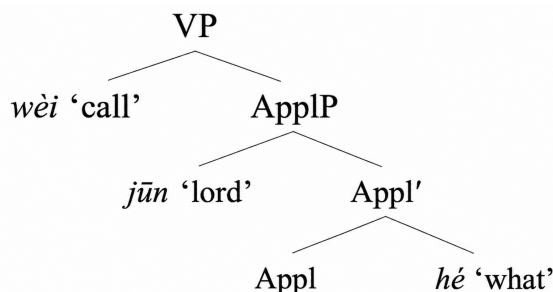


Figure 1.

Aldridge (2012) believes that *yǐ* ‘use, with’, which is traditionally treated as a verb and a preposition in OC functions as an applicative head but the Appl⁰ in Figure 1 is empty¹⁴. Aldridge (2010) does not clarify whether the applicative structure *rú-zhī-hé* has its Appl⁰ filled. Paul and Whitman (2009) mention the Raising Applicative Hypothesis (Georgala et al. 2008) which claims that “overt applicative affixes are realized uniformly as verbal suffixes”. If this hypothesis is correct, even if the Appl⁰ in Figure 1 is filled by *yǐ*, the applied object analysis needs to explain why OC violates the Raising Applicative Hypothesis because *yǐ* is not a verbal suffix, rather, it is a preposition prefixed to DPs.

Based on the above discussion, an empirical drawback of the applied object analysis is that it overlooks the fact that the construction V-N-*hé* is heterogeneous: in some cases, *hé* is a wh-argument (1), but in some cases, it is a wh-adverbial (2, 3) and thus cannot be an object. I have tried to demonstrate drawbacks of this analysis from the perspective of syntax, semantics and

¹⁴ It needs to be pointed out that Aldridge’s applicative head analysis of *yǐ* ‘use, with’ was published two years after her applied object analysis of *hé*.

morphology and it is the semantic side, namely, the fact that some clause-final instances of *hé* do not receive the theta roles that an applied object is canonically assumed to receive, is the most serious problem.

3.2 the adjunct analysis

The adjunct analysis treats the clause-final *hé* as a PP. Before I present evidence for this solution, it is necessary to clarify the definition of adjunct and the distinction between adjuncts and arguments.

In X-bar theory, adjuncts are sisters of phrasal nodes and can never be sisters to a head but (internal) arguments can be. Arguments are assigned theta-roles by heads, but adjuncts are not assigned theta-roles. In the Minimalist Program (Adger 2003:82-90), arguments are computed in syntax by the operation Merge. Take a simple transitive sentence *Paul burn letters*, for instance: the complement (i.e. internal argument) *letters* checks its [N] feature with *burn* and then they merge (first merge) to project a V at the bar-level; the specifier of the V, namely, *Paul* (i.e. external argument) checks another selectional [N] feature of V, after which the derivation projects a maximal VP. On the other hand, adjuncts are “somehow incorporated into a sentence, but not via the checking of selectional features” (idem: 88). In a word, arguments are “merged” into syntax due to selectional features but adjuncts are “adjoined” into syntax without having selectional features. While in Chomsky’s (2004) term, what distinguishes adjuncts from complements is that they undergo pair-Merge rather than set-Merge. It can be seen that adjuncts differ from arguments in many important respects, so the adjunct analysis and the applied object analysis are essentially different.

The empirical data that supports the adjunct analysis comes from the parallelism between “VP-*hé*” and “VP-PP”. In OC, PPs canonically follow the main verb (10-12), but exceptions also exist (13)¹⁵:

(10) 紀之以義，終之以仁。 (Lǐjì, Wénwángshìzì)

[_{VP} v jì	zhī	[_{PP} yǐ	yì]]
regulate	it	with	righteousness
zhōng	zhī	yǐ	rén
complete	it	with	benevolence

‘Regulate it with righteousness and bring it to completion with benevolence.’

(11) 孟孫問孝於我。 (Analects, Wéizhèng)

Mèngsūn	[_{VP} v wèn	xiào	[_{PP} yú	wǒ]]
Mengsun	ask	filial.pietty	from	I

‘Mengsun asks for filial piety from me.’

(12) 禮起於何也？ (Xúnzǐ, Lǐlùn)

lǐ	[_{VP} v qǐ	[_{PP} yú	hé]]	yě?
ritual	originate	from	what	SFP

‘From what does ritual originate?’

(13) 於何本之？ (Mòzǐ, Fēimìngshàng)

[_{VP} [_{PP} yú	hé]	v běn	zhī]
from	what	originate	it

‘From what does it originate?’

¹⁵ This restriction started to break in the transitional period between LOC and EMC, while in Mandarin, most of PPs cannot follow the main verb.

In (12) and (13), the wh-word *hé* is introduced by the preposition *yú* ‘from’, but when the preposition changes to *yǐ* ‘by’, things are different:

(14) a. 王曰，何以利吾國？ (Mencius, *LiánghuìwángShàng*)

wáng	yuē,	[_{VP PP} hé	yǐ]	v lì	wú	guó]
king	say	what	by	benefit	my	state

‘The king said, by what (can I) benefit my state?’

b. 以何驗之？ (Lùnhéng, *Wúxíng*)

[_{VP} [_{PP} yǐ	hé]	v yàn	zhī]
by	what	verify	it

‘By what (shall we) verify it?’

c. 俟孔子以何禮？ (Lùnhéng, *Xièduǎn*)

[_{VP} v sì	Kǒngzǐ	[_{PP} yǐ	[_{DP} hé	lǐ]]]
serve	Confucius	with	what	rite

‘With what rites (shall one) serve Confucius?’

When *hé* is introduced by *yǐ* ‘by’, regardless of whether it is fronted or not (14a vs 14b), we never find examples in which a PP follows a verb, namely, there is no “VP-*yǐ-hé*” (VP-by-wh), let alone “VP-*hé-yǐ*”. Aldridge (2012) analyzes *yǐ-yáng-yì-zhī* ‘APPL-sheep-change-it’ as an applicative, and the ApplP *yǐ-yáng* can either precede or follow the main verb. The different syntactic behavior between *yǐ-hé* ‘by-wh’ and *yǐ-yáng* ‘APPL-sheep’ demonstrates that *yǐ* functions differently in these two contexts: it is simply a preposition in *yǐ-hé* but is an applicative head in *yǐ-yáng-yì-zhī*. Examples (12)-(14), as well as the comparison between *yǐ-hé* and *yǐ-yáng*, show that even if there is an overt applicative morpheme *yǐ*, it may not be appropriate to analyze (1)-(3) as applicatives because *yǐ-hé* ‘by-wh’, which I treat as a PP, behaves differently from the ApplP in Aldridge (2012).

The comparison between (11)-(13) and (14a,b) also concerns the interpretation of *hé*. When the preposition is *yú* ‘from’, *hé* is a wh-nominal meaning ‘what’, as indicated by the complement of *yú* in (11), which is a personal pronoun *wǒ* ‘I’. While when the preposition is *yǐ* ‘by’, *hé* is more like a wh-adverbial ‘how’ that asks about manner, and this explains why *yú-hé* ‘from-what’ and *yǐ-hé* ‘by-how’ behave differently in syntax. In addition to the parallelism between “VP-*hé*” and “VP-PP”, the fact that adjuncts are believed to occupy the most embedded positions (Larson 1988; Cheng 1991:99) also supports the adjunct analysis since no other content constituents can be added after “VP-*hé*”¹⁶, which means *hé* is in the most embedded position.

(14c) looks like a counterexample of the above analyses because the complement DP of *yǐ* ‘by’ contains a wh-word *hé* but the PP still follows the verb *sì* ‘serve’. In fact, (14c) is structurally different from (14a,b). *Hé* in (14c) serves as the specifier of the DP *hé-lǐ* ‘what-ritual’, but in (14a,b), *hé* itself is the complement of the preposition *yǐ* ‘by’. One important insight I draw from (14c) is that under the adjunct analysis, “VP-*hé*” includes a covert preposition. This idea will be elaborated in more detail in Chapter IV when I identify two types of wh-adjuncts *hé*. I have mentioned above that even if we assume there to be a covert applicative head in “V-N-AppI⁰-*hé*”, it is not appropriate to analyze this construction as an applicative. But it needs to be clarified that this does not mean that I deny the covert existence of *yǐ* in “VP-*hé*”. What I deny is the existence of the AppI⁰ *yǐ*, but the existence of the preposition *yǐ* is compatible with my adjunct analysis. To conclude this section, I provide a pair of sentences to support the existence of covert prepositions in “VP-*hé*”:

(15) 需役聞之於謳。

(*Zhuāngzǐ, Nèipiān, Dàzōngshī*)

Xū Yì wén zhī yú Ōu

¹⁶ SFPs can be further added after *hé*, for instance (II.5).

Xu Yi hear it from Ou
 ‘Xu Yi heard it from Ou.’

(16) 吾聞之吾師。 (Zhuāngzǐ, Wàipiān Qiūshuì)

wú wén zhī [PP Ø wú shī]
 I hear it my teacher
 ‘I heard it from my teacher.’

(15) and (16) are excerpted from the same text. One of them has an overt preposition *yú* (15) but the other one does not (16). Also, (15) and (16) are parallel in meaning in that the external arguments ‘Ou’ and ‘my teacher’ both are SOURCES. Following my analysis, I provide the tree of *wèi-jūn-hé* (1) in Figure 2:

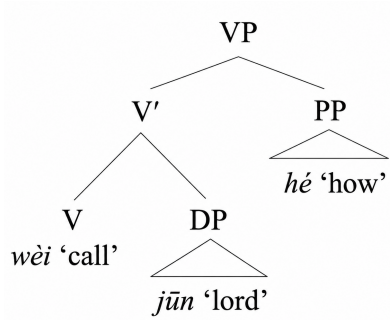


Figure 2.

Another advantage of the adjunct analysis is that it tries to reach a unified explanation for the clause-final wh-nominal *hé* ‘what’ (1) and the clause-final wh-adverbial *hé* ‘how’. Although this study focuses on examples like (2) and (3), I expect the clause-final wh-nominal *hé* ‘what’ also to be compatible with the adjunct analysis, supposing the existence of covert prepositions.

3.3 the predicate analysis and OC multiple wh-questions

Multiple wh-questions are discussed here for three purposes. One is to provide a theory-internal motivation to call into question the applied object analysis, another is to provide a predicate analysis of the clause-final *hé*, and the third is to exemplify the application of near-native intuition mentioned in I.4.2.

Multiple wh-questions are rarely mentioned in the literature, and their absolute frequency in the corpus is also very low. The interpretation of multiple wh-questions is tricky even in modern languages, let alone in a classical language. Therefore, it's necessary to resort to near-native intuition, namely, ancient annotations, because our secondary intuition is generally less reliable than ancient people's near-native intuition. In (17)-(20), I provide corresponding ancient annotations, if any, in their original texts accompanied with English translations in the footnotes, together with approximate living periods of the annotators, approximate time intervals between the original texts and ancient annotations, and lastly, reference books:

(17) 信臣精卒陳利兵而誰何¹⁷?

(*Shǐjì, Qínshǐhuáng Běnjì*)

xìn chén jīng zú chén lì bīng ér **shuí hé?**
 trusted minister elite soldier lay.out advanced weapon CONJ who what
 'Trusted ministers and elite soldiers lay out advanced weapons and (ask) who (the passersby are and) what (they carry).'

(18) 不知其誰何¹⁸?

(*Zhuāngzǐ, Yìngdìwáng*)

bù zhī qí **shuí hé**
 NEG know it who how
 'Do not know who and how it is.'

¹⁷ 如淳曰，“何”猶“問”也。 [*Ru Chun* (ca. late 2 AD-ca. early 3 AD) said, *hé* is the same as *wèn* 'ask'.] (Annotator: *Péi Yīn* 裴駰, ca. late 5 AD, after 580 years, *Shǐjì* (Vol.1) 史記 (一), Beijing: Zhonghua Bookstore, 2014, p354.)

¹⁸ 既不可名目，故不知的是何誰也。 [Now that it cannot be named, therefore do not know how it is and who it is.] (Annotator: *Guō Xiàng* 郭象, 252 AD-312 AD, after 500 years, *Zhuāngzǐ Zhùshū* 莊子注疏, Beijing: Zhonghua Bookstore, 2011, p166.)

(19) 不知為之誰何¹⁹?

(*Huáinánzǐ, Běnjīngxùn*)

bù	zhī	wéi	zhī	shuí	hé
NEG	know	do	it	who	how

‘Do not know who did it and how it was done.’

(20) 若所追者誰何？

(*Shǐjì, Huáiyīnhóu Lièzhuàn*)

ruò	suǒ	zhuī	zhě	shuí	hé
you	REL	chase	NOML	who	what

‘What you are chasing is who and what?’

The meaning of (17) is controversial. My secondary intuition leads me to gloss *shuí-hé* as two wh-words ‘who-what’, but an ancient annotator 如淳 *Rú Chún*, who lived around 300 years after the composition of the original text interpreted *hé* not as a wh-word but as a main verb meaning ‘ask’ that takes *shuí* ‘who’ as its fronted object. *Rú Chún*’s near-native intuition indeed makes more sense from a view of canonical OC syntax. What is at issue is whether wh-words themselves can function as predicates in OC. This brings us back to the core concern of this study—the syntactic structure of “VP-*hé*”. If the clause-final *hé* really is a predicate, *Rú Chún*’s annotation would be easier to understand. The predicative nature of the sentence-final *hé* can also be detected from examples in which *hé* is followed by a particle *yě*:

(21) 君子之不耕而食,何也?

(*Mencius, Jīnxīn Shàng*)

jūnzǐ	zhī	bù	gēng	ér	shí,	hé	yě?
gentleman	GEN	NEG	plough	CONJ	eat	why	SFP

¹⁹ 道無姓名, 自當然也, 故曰不知誰何也. [The Tao does not have a name, and naturally it should be like this, therefore saying that do not know who and how Tao is.] (Annotator: *Gāo Yòu* 高誘, 250 AD-284 AD, after 400 years, *Huáinánzǐ* 淮南子, Shanghai: Shànghǎi Gǔjí Chūbǎnshè)

‘As for gentlemen not ploughing yet eating—why is this so?’

(22) 追我者誰也?

(*Mencius, Lílóu Xià*)

zhuī wǒ zhě, shuí yě?

pursue I NOML who SFP

‘Who is it that is pursuing me?’

Shi (2015:4) hypothesizes that *yě* in pre-OC period was a copula, and in OC it had already grammaticalized into a modal particle but still preserved some verbal features²⁰. Hong (2010:223-226) holds a similar but more radical point of view, arguing that the sentence-final particle *yě* was a copula which is equivalent to the copula *red* in Tibetan²¹. Following Shi (2015) and Hong (2010)’s ideas, *hé* in (21) and (22) would not be a predicate because the particle *yě* functions as the predicate. Recall that there are textual variants (I.16, 17) indicating that “VP-*hé-yě*” changed to “VP-*hé*” in later periods. This historical change implies that the original copula *yě* incorporated into the wh-word *hé* and made the clause-final *hé* in later periods behave predicatively.

(17) is a good example to demonstrate the predicative nature of the sentence-final *hé* since this sentence contains a conjunction word *ér* ‘and’. According to Mei (2018:175-220)’s analysis of *ér*, (17) is a “coordinate construction” (idem:173), and *ér* connects two predicates. Therefore, the structure of (17) can be schematized in Figure 3:

²⁰ The original text is “‘也’是一個仍然保留著動詞性的語氣詞”.

²¹ Whether LOC has a copula is controversial. For example, Mei (2018:167-170) points out that EOC has a copula *wéi* 惟 but later disappeared, while LOC does not use overt copulas in copular sentences.

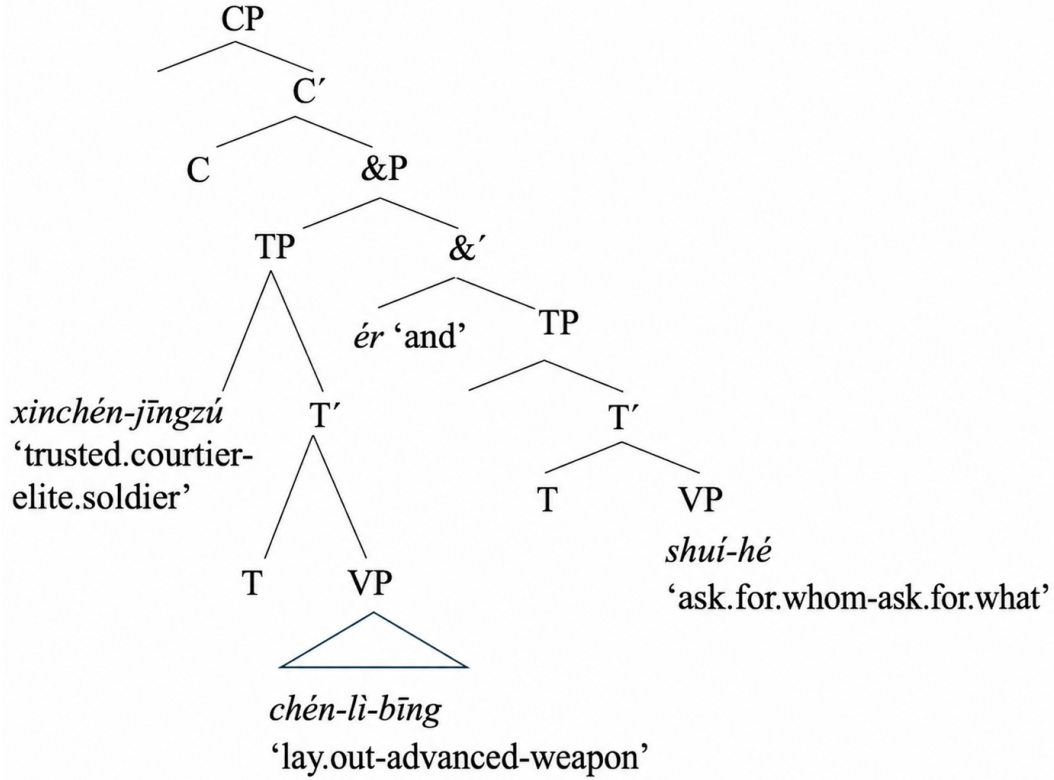


Figure 3.

Following the predicate analysis, *shuí* and *hé* can be glossed as ‘ask.for.whom’ and ‘ask.for.what’ respectively.

An ancient scholar 郭象 *Guō Xiàng* had an annotation for (18)²²: 故不知的是何誰也 *gù-bù-zhī-dí-shì-hé-shuí-yě* ‘therefore-NEG-know-this-be-what-who-SFP’ (‘Therefore do not know what and who this is’). *Guō Xiàng*’s near-native intuition reflects some interesting changes from LOC to EMC. First, the proximate demonstrative *dí* ‘this’ he used is commonly believed to be a dialectal word, and more importantly, he told us that *shuí-hé* ‘who-what’ in the original text indeed contains two wh-words, as well as the fact that wh-words themselves were not able to function as predicates at least in the EMC period, so he added a copula *shì* ‘be’. It should also be noticed that *Guō Xiàng*

²² See footnote 18.

flipped the word order of the original text such that his paraphrase used *hé-shuí* ‘what-who’ rather than the original *shuí-hé* ‘who-what’. We cannot tell for sure whether he chose the flipped word order randomly or for any syntactic reasons.

As for the word order between *shuí* ‘who’ and *hé* ‘what’ in multiple wh-questions, we can take a further look at (19) and use it as an argument to support the adjunct analysis. In (19), if *hé* ‘what’ can be analyzed as an applied object, *shuí* ‘who’ can also be treated so, and we would be concerned with explaining how a second applied object can be added. On the other hand, following the adjunct analysis, both *shuí* ‘who’ and *hé* ‘how’ can be naturally analyzed as adjuncts because there can be an infinite number of bar levels where adjuncts can be adjoined to syntax. (20) also deserves our attention because there are textual variants of this sentence, one of which²³ does not have the sentence-final *hé* ‘what’ while another one does²⁴. Here I follow the second version in believing the lack of *hé* is due to later revisions by ancient or pre-modern compilers because of their unfamiliarity of OC grammar which is a common phenomenon in the long circulation history of ancient Chinese texts. This textual variance implies that wh-words lost the ability to function as predicates thus the clause-final wh-adverbial *hé* disappeared in the colloquial language no later than the Western Hàn dynasty (202 BCE-9 AD).

In a more abstract sense, analyzing *hé* ‘how’ and *shuí* ‘who’ as applied objects or adjuncts concerns how restrictive our theory is. Kayne (1994: XIII) clarifies his goal of proposing “further severe limitations on the range of syntactic representations” in order to achieve a “restrictive theory of syntax”. An ideal syntactic theory should be able not only to generate an indefinite number of grammatical sentences but also to be restrictive enough to not over-generate ungrammatical sentences. For our concerns here, treating *hé* ‘how’ and *shuí* ‘who’ as applied objects makes the

²³ *Shiji* (Vol.8) 史記(八): Beijing: Zhonghua Bookstore, 2014, p3167.

²⁴ *Shiji* 史記: Beijing: Zhonghua Bookstore, 1959, p2611.

addition of objects less restrictive, but the adjunct analysis is more restrictive in the sense that it won't over-generate ungrammatical sentences—if we happen to find a sentence with three wh-words following a VP, it will still fall within the prediction of the adjunct analysis because it does not need to worry about the number of wh-words adjoined. Still, the adjunct analysis correctly predicts the existence of the word order *hé-shuí* ‘what-who’ (23-25):

(23) 公為何誰矣？ (WúyuèChūnqiū, Hèlǚ Jiǔnián)

gōng	wèi	hé	shuí	yǐ
duke	act.for	what	who	SFP

‘For what, and for whom, does the Duke act?’

(24) 我已為東帝，尚何誰拜？ (Shǐjì, WúwángBì Lièzhuàn)

wǒ	yǐ	wéi	dōng	dì,	shàng	hé	shuí	bài?
I	already	become	east	emperor	still	what	who	bow

‘I have already become the Emperor of the East; for what, and to whom, (should I) still bow?’

(25) 此何誰知不為福？ (Shuōyuàn, Yìwén)

cǐ	hé	shuí	zhī	bù	wéi	fú?
this	why	who	know	NEG	be	fortune

‘Who knows this is not a fortune, and why?’

Comparing (23)-(25) to (17)-(20), we can notice that *shuí-hé* ‘who-what/how’ never surfaces at the left periphery, while *hé-shuí* ‘what/why-who’ can either stay VP-internally (23), or be external to VPs (24, 25). In (25), since *hé* ‘why’ is a reason wh-adverbial, and *shuí* ‘who’ is a subject at [Spec, TP], *hé* ‘why’ should occupy a standard [Spec, CP] position. Therefore, as a functional projection higher than CP, *cǐ* ‘this’ can be a Topic (Rizzi 1997). The structure of (25) is given in Figure 4:

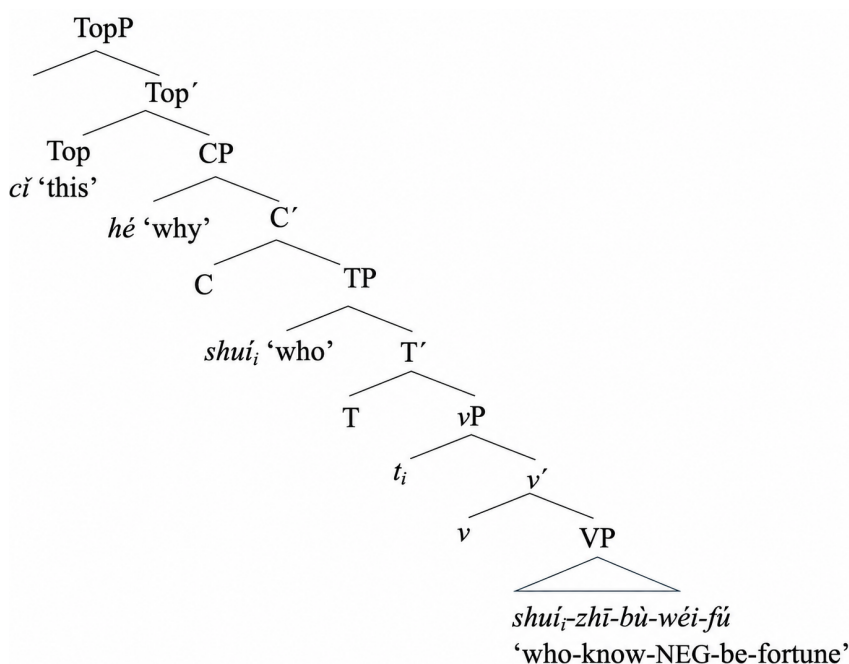


Figure 4.

In wh-questions containing only one wh-word, there are also examples to support the predicate analysis:

(26) 非諸侯而何?

(*Analects*, *Xiānjìn*)

fēi	zhūhóu	ér	hé?
not.be	feudal.lord	CONJ	what

‘If (they) are not the feudal lords, what (else could they be)?’

(27) 人事未為，天氣已見，非時而何?

(*Lùnhéng*, *Zhìqī*)

rén	shì	wèi	wéi,	tiān	qì
human	affair	not.yet	carry.out	heaven	sign
yǐ	xiàn,	fēi	shí	ér	hé?
already	appear	not.be	timing	CONJ	what

‘Human affairs have not yet been carried out, yet signs in Heaven have already appeared; if (this) is not timing, what is it?’

In (26) and (27), the conjunction word *ér* connects *hé* with another clause. Importantly, the negative marker *fěi* (*pəj) ‘not.be’ is a fusion of *bù-wéi* (*pə-g^wij) ‘NEG-be’ (Schuessler 2007:232). This indicates that *hé* has the same syntactic status as the clause containing a copula, so it is reasonable to treat *hé* as an independent predicate.

The above discussion explains the predicate analysis from a purely OC-internal perspective, while this analysis is also supported by evidence from related languages. In some modern Tibeto-Burman languages, wh-words can function as predicates on their own:

(28)	ta	pa	le	<i>Burmese</i>
	this	what	Q _{y/n}	
	‘What is this?’			

(29)	tehi ²¹³ ku ⁵⁵	iaŋ²¹³ku⁵⁵	*mo ⁵¹	<i>Lùxī Achang</i>
	this	what	Q	
	‘What is this?’			

(28) and (29) show that in Burmese and Lùxī Achang, wh-words can be predicates and there are no copulas in wh-questions. A difference between (28) and (29) is that Burmese obligatorily needs a yes-no question particle *le* but the yes-no question particle *mo*⁵¹ is not compatible with wh-questions in Achang. The predicate analysis is also supported by Tsou, an Austronesian language spoken in Taiwan. According to Tsai (2008:87-88), instrumental *how* and manner *how* in Tsou are “realized as conjuncts of the main predicate” which means that the meaning ‘How did Pasuya hit Mo’o’ is literally expressed as ‘Pasuya-hit-Mo’o-and-how’. Tsai also clearly points out that *mainenu* ‘how’ in Tsou is “syntactically realized as a matrix predicate”.

In sum, the predicate analysis of the clause-final *hé* has its merits, especially because it finds evidence from both OC and other languages, and it takes into consideration the historical correlation between “VP-*hé*” and “VP-*hé-yě*”. However, the absolute number of sentences supporting it is not sufficient.

3.4 the resultative complement analysis

In this subsection, I discuss another solution to the syntactic status of the clause-final wh-word *hé* namely, treating it as a “resultative complement”. This analysis has its commonality with the applied object analysis in that both of them treat *hé* as a complement rather than an adjunct.

In the generative tradition, *complement* is the “constituent which is the sister of X^0 ” (Sybesma 2015), while “object” is a subcategorization of *complement*. But in descriptive Chinese syntax, the term “complement” (補語 *bǔyǔ*) has been used in a different way in which it serves as a general term to refer to “a non-nominal constituent which follows the verb or adjective either directly or linked to it by *de*²⁵” (idem). The reason that “complement” has been used in such a way in Chinese syntax is because there exist some elements in Chinese that form a natural class while some of them correspond to real *complements* in English, but some do not have equivalents in English. (30)-(32) give some examples of “complement” (in brackets) in Chinese syntax:

- (30) tā shāo cài shāo de [hěn hǎo chī].
 he cook food cook DE very good eat
 ‘He cooks well.’

²⁵ *De* can be simply understood as a linking particle in-between a(n) verb/adjective and a “complement”.

(31) tā bǎ²⁶ dàngāo kǎo [hú=le]
 he BA cake bake burnt=PERF
 ‘He baked the cake **burnt**.’

(32) huǒchē wǎng běifāng kāi [qù=le]
 train towards north drive go=PERF
 ‘The train drove towards the north.’

The “complement” in Chinese can be further categorized into “degree complement” (30), “resultative complement” (RC) (31), “directional complement” (32), etc., according to their semantics. (30)-(32) show that some “complements” have English equivalents, such as *hěn-hǎo-chī* in (30) corresponds to the adverbial adjunct *well*, and *hú-le* in (31) to the adjectival complement *burnt*. But some “complements”, for example, *qù-le* in (32) do not have English equivalents.

(30) is also referred to as a verb-copying construction, and it was formed in Chinese by the 15th century (Shi 2023:256). In ancient Chinese, the equivalent of a RC in verb-copying sentences can directly follow the object, forming a “V+O+RC” construction with no copied verbs and *de*. For instance:

(33) a. 尋傷左臂復打頭破 (EMC: *ChūyàoJīng*, Vol.4.)
 xún shāng zuǒ bì fù dǎ tóu pò.
 soon hurt left arm again hit head broken
 ‘Soon he hurts his arm and then hit his head broken.’

b. tā dǎ tóu dǎ pò=le

²⁶ In descriptive and functional grammars, Mandarin *bǎ* had been regarded as an object marker which fronts the object to preverbal positions and serves as the condition to derive the SOV order. In the generative contexts, *bǎ* was analyzed as being related to aspect. For example, Zou Ke (1993) *The Syntax of the Chinese BA Construction* (*Linguistics* 31, pp715-736) regards *bǎ* as the head of a functional category and selects an aspect or particle phrase as its complement. Also see Liu Feng-Hsi (1997) *An Aspectual Analysis of BA* (*Journal of East Asian Linguistics* Vol.6, pp51-99) for an aspectual analysis of *bǎ*. I here gloss *bǎ* as BA for simplicity.

he hit head hit broken=PERF
 ‘He hits his head broken.’

c. tā bǎ tóu dǎ pò=le
 (s)he BA head hit broken=PERF
 ‘(S)he hits her/his head broken.’

(34) a. 飲酒樂。 (Zuǒzhuàn, Shàngōng Èrshínián)

yǐn jiǔ lè
 drink wine happy
 ‘Drink the wine and become happy.’

b. hē jiǔ hē [kāixīn=le] (Mandarin)
 drink wine drink happy=PERF
 ‘Drink the wine and become happy.’

(33a) and (34a) are examples of “V+O+RC” in MC and OC, and (33b, c) (34b) are their Mandarin equivalents. Based on the above introduction of RC in Chinese, we can analyze the clause-final *hé* in “V-O-*hé*” as a RC. Semantically, this analysis makes sense because answers to questions like *rú-zhī-hé* ‘do-it-how’ normally provide solutions or comments to the event *rú-zhī* ‘do-it’ which then can be regarded as a kind of “result”. But from a historical view, this analysis is not preferred to because although there are cases of “V+O+RC” in OC (34a), it is generally believed that this structure became widely attested in texts no later than EMC (Zhao 2000:59). In other words, the “V+O+RC” structure had not fully developed in OC, at least when the RC is not a wh-word. From a historical perspective, the resultative complement analysis is problematic in the sense of “complement”, but it has its merits in the sense of “resultative”. Tsai (2008:84) points out that

answers to the English *how*-questions can either be manner, instrumental or resultative, and this supports the idea that the clause-final *hé* has resultative meanings.

Chapter IV Two Types of Wh-adjunct *Hé*

We have seen four possible analyses of the clause-final *hé* in Chapter III: applied object, adjunct, predicate, and resultative complement, and I chose the second solution, so starting from this chapter, I use “wh-adjunct” to refer to the clause-final *hé*. Section 2.1 already briefly touches on the issue of subcategorizing wh-adjuncts into the *how*-type and the *why*-type according to their different syntactic distributions in overt syntax. In this chapter, some syntactic diagnostics are employed to test the hierarchical distinction between the *how*-wh-*hé* and the *why*-wh-*hé*, and the adjunct analysis proposed in 3.2 will be slightly revised based on some inspiration from Japanese. Most importantly, building upon the distinction made between the two types of *hé*, this chapter reaches a conclusion about positions in which *how*-wh-*hé* and the *why*-wh-*hé* base-generate.

4.1 terminological issue

Discussion of base-generating positions of wh-phrases correlates with the issue of “in-situ” and “ex-situ”. So, it is necessary to clarify the meaning of these two terms first. In early works of wh-in-situ, such as Huang (1982 [2014]) and Pesetsky (1987), “in-situ” refers to the phenomenon that a wh-phrase does not move in S-Structure, in other words, the movement is not realized at the PF level. Pesetsky (1987:98) clearly states that “a wh-in-situ is a wh-phrase that has not undergone wh-movement by S-Structure” and he emphasizes that an in-situ wh-phrase is “visibly” not moved

to [Spec, CP]. Tsai (1994:121) describes languages like Chinese as ones in which “wh-phrases remain in situ at S-Structure”. The early use of “in-situ” implies that even if a wh-phrase moves at LF, whether the movement is phrasal or featural, that wh-element is still in-situ.

But this term seems to have been used in another way. For instance, Aoun and Li (1993:200) argue that “in Chinese and English wh-elements stay in situ even in the LF component”, while Tan and Shen (2025) use the term “true wh-in-situ” to refer to colloquial Singapore English sentences like *John can buy the durians where*. The second use of “in-situ” extends its meaning to the LF level, and in-situness becomes no longer visibly apparent. Under the second use of “in-situ”, scholars resort to the movement of elements other than wh-phrases themselves to interpret in-situ wh-elements. Aoun and Li (1993), for example, put forward that it is a question operator that co-indexes with in-situ wh-elements that is raised to the appropriate Spec of Comp position, and that operator licenses wh-elements to be in-situ even in LF. The raising of the Q-operator in Aoun and Li’s solution is realized overtly at S-Structure.

Kotek (2014:22) believes that “no overt or covert movement is involved under the in-situ approach to questions”. This more recent view can be regarded as a radical version of extending in-situness to the level of LF because it does not even rely on the movement of other elements. By doing so, the semantics and syntax of in-situ wh-elements are made more independent from each other. Kotek (2014:23-24) introduced the Rooth-Hamblin framework, which is a semantic tool to interpret in-situ wh-words compositionally. And this semantic tool also applies to overtly moved wh-words. As Kotek concluded, “there is no reason to expect any instances of wh-movement that are caused by the semantic needs of the wh-words themselves” (idem: 24), pure syntax-internal motivations, such as an EPP feature, must be invoked to explain why in-situ wh-words must move even if they are semantically licensed. There are also other approaches such as choice functions

(Reinhart 1998) and unselective binding (Pesetsky 1987) that have been put forward to interpret in-situ wh-elements. Due to this terminological issue of “in-situ”, Li (2024:14) calls it “truly in-situ analysis” when in-situ wh-elements do not move in LF either. In this thesis, I use the term “in-situ” in its more recent sense which means if a visibly in-situ wh-element or anything related to it moves in LF, I will call this wh-element “ex-situ”. Next, I use some concrete examples to clarify the in-situ phenomenon.

Although “a fundamental property of human languages is that syntactic constituents are interpreted in positions different from the ones where they are phonetically realized” (Corver and Nunes 2007:1), there are many counterexamples to this fundamental property, and wh-in-situ is one of them. Wh-in-situ can be defined as a phenomenon in which wh-elements are phonetically realized at the same place in which they receive theta roles. From a semantic viewpoint, in-situ wh-elements are semantically interpreted in the position where they occur in overt syntax. In such cases, wh-expressions occupy the same syntactic positions as non-wh-expressions because they receive the same abstract case.

(1) and (2) further clarify the above definitions:

(1) a. [CP Whom_i C' [TP does [VP he like t_i]]]?

b. [CP Ø C' [TP He T' [VP likes them]]].

(2) a. tā xǐhuan shuí?

he like who

‘Who is the person *x*, such that he likes *x*’

b. tā xǐhuan tāmen.

he like they

‘He likes them.’

In (1), *whom* and *them* both are assigned accusative case in VP-internal positions, but *whom* is phonetically realized at a VP-external position [Spec, CP], and is defined as ex-situ. In contrast, the Mandarin wh-word *shuí* ‘who’ and the personal pronoun *tāmen* ‘they’ both remain in VP-internal positions (2), and [Spec, CP] is not filled by wh-words in overt syntax, so *shuí* ‘who’ is called an in-situ wh-word.

- (3) a. [CP How_i C' [TP does he [VP eat noodles *t_i*]]]?
 b. [CP [TP He [VP eats noodles [PP by chopsticks]]]]?

- (4) a. tā zěnmē chī miàntiáo?
 he how eat noodles
 ‘What is the manner *x*, such that he eats noodles in *x*?’

- b. tā yòng kuàizi chī miàntiáo
 he with chopsticks eat noodles
 ‘He eats noodles with chopsticks.’

(1) (2) are parallel to (3) (4) but wh-words in the former two examples are arguments, while those in the latter two examples are adjuncts. The comparison between (1) (2) and (3) (4) demonstrates that in-situness is not determined by linear word orders but by canonical positions of their non-wh counterparts. And this means that in order to determine whether a wh-element is in-situ or ex-situ, it is essential to determine the syntactic status of its non-wh counterpart. Therefore, the most natural way to derive the clause-final wh-adjunct *hé* is to regard it as being in-situ because its non-wh-counterpart PPs canonically follow the verb, as revealed by (III. 10-12). Then, we need to explain why and how the clause-final *hé* ‘how, why’ also surfaces at left peripheries.

4.2 reason-wh *hé* ‘why’ vs manner-wh *hé* ‘how’

Lin (1992) deals with the asymmetry between Mandarin *zěnmeyàng* ‘how’ and *wèishénme* ‘why’, and in Lin (2014:185), he points out that *zěnmeyàng* ‘how’ is a VP adjunct. Lin’s adjunct analysis of *zěnmeyàng* fits with our impression of the manner *hé* ‘how’ in OC in that in contrast with the reason *hé* ‘why’, it seldom moves to [Spec, CP] and at the same time, asks about manner, while it also conforms with Nakao and Obata (2009)’s dealing with Japanese accusative reason-wh *nani-o* ‘what-ACC’. The analyses of reason-wh in Mandarin and Japanese make it reasonable to also treat OC reason-wh *hé* as an adjunct but it does not base generate on the left at [Spec, CP] when surfacing at the right periphery—it needs to move to [Spec, CP] from the right periphery, a position where it adjoins to VPs. This analysis is in line with the spirit of Tsai’s (1994) analysis which assumes that wh-adverbs must undergo covert XP movement. Following this line, Tsai (2008:97) clarifies that the instrumental and resultative *how* is in-situ. The cross-linguistic investigation of *how-why* alternations in Tsai (1994, 2008) drives me to hypothesize that clause-final positions are places where the wh-adjunct *hé* ‘how, why’ base-generates. The difference between these two wh-adjuncts is that the *how*-wh-word *hé* stays in situ at the right periphery, but the *why*-wh-word *hé* moves leftwards to [Spec, CP].

(5) shows both differences and commonalities of *why*-wh-words in Mandarin and OC:

- (5) a. **shuí** **wèishénme** zhīdào zhè=ge bù shì hǎo shì?
 who why know this=CL NEG be good thing
 ‘Who knows that this is not a good thing and why (s)he knows?’
- b.* **wèishénme** **shuí** zhīdào zhè=ge bú shì hǎo shì?
 why who know this=CL NEG be good thing
 Intended: ‘Who knows that this is not a good thing and why (s)he knows?’

c. wèishénme yǒu-rén zhīdào zhè=ge bú shì hǎo shì?
 why have-people know this=CL NEG be good thing
 ‘Why there is someone who knows that this is not a good thing?’

d. shuí zhīdào zhè=ge wèishénme bú shì hǎo shì
 who know this=CL why NEG be good affair
 ‘Who knows why this is not a good affair?’

(5b) is a rough Mandarin equivalent of (III.25) 此何誰知不為福 *cǐ-hé-shuí-zhī-bù-wéi-fú* ‘this-why-who-know-NEG-be-fortune’, but it is ungrammatical to have ‘why’ preceding ‘who’. The same reading of (III.25) in Mandarin is shown in (5a). (5b) shows that in Mandarin, ‘why’ cannot scope above a subject wh-word to take matrix scope. In other words, when there is a subject wh-word in the matrix CP, Mandarin ‘why’ needs to adjoin lower than the wh-subject (5a) or takes the narrow scope (5d). But OC ‘why’ does not have this constraint. This contrast leads me to suspect that OC reason-wh *hé* and Mandarin *wèishénme* ‘why’ base-generate at different positions. The fact that ‘why’ in Mandarin only occurs at the left periphery and in OC occurs at both peripheries suggests that it base-generates at a lower position in OC.

There is empirical evidence to support that the reason-wh *hé* base-generates at different positions from the Mandarin *wèishénme* ‘why’:

(6) 其為舟車何？以為車以行陵陸，舟以行川谷。 (Mòzǐ, Jiéyòng)

qí	wéi	zhōu	chē	hé?			
it	make	boat	carriage	why			
yǐ	wéi	chē	yǐ	xíng	lù	líng,	zhōu
because	make	carriage	for	walk	land	hill	boat
yǐ	xíng	chuān	gǔ.				

for walk river valley

In (6), I gave the context to show that the reason-wh *hé* ‘why’ here must be interpreted as ‘why’ rather than ‘how’. This means that the reason-wh *hé* can linearly appear either at the left (II. 3) or the right periphery. Therefore, I argue that OC reason wh-adjunct *hé* ‘why’ needs to move to [Spec, CP] to surface at the left periphery. An important piece of evidence supporting the landing site of *hé* comes from *qí* (*gə), which is a grammatical word that frequently follows *hé*. In OC, *qí* literally means ‘that’ but it is also used as a personal pronoun. (7) and (8) show different functions of *qí*:

(7) 如之何其使斯民飢而死也? (Mencius, *Liánghuìwáng Shàng*)

pro_i rú zhī hé **qí_i** shǐ sī mín jī ér sǐ yě?
do it how he CAUS these people starving CONJ die SFP

‘How is it that he caused these people to starve and die?’

(8) 何其久也? (Shuōyuàn, *Fèngshǐ*)

[_{CP} hé _C **qí** [_{TP} jiǔ] yě]?
why that long SFP

‘Why it was such a long time?’

In (7), *qí* ‘he’ receives an agent theta role since it appears in a causative sentence, and it co-indexes with the matrix *pro*. But in (8), there is no semantic agent, nor any other theta roles, that *qí* can bear. For cases like (8), I argue that *qí* is a complementizer rather than, say, a pronoun. This analysis is also reasonable from a diachronic view because demonstratives are one of the major grammaticalization sources of complementizers (Heine and Kuteva 2002:329).

Cross-linguistic comparisons also support the idea that *qí* in (8) is a complementizer. Although wh-phrases in standard English cannot co-occur with a *that*-complementizer in a minimal clause,

in Bavarian German (Bayer 1984), Irish (Carnie 2021:368), and some older varieties of English (idem:378), a *wh*-phrase is compatible with an overt complementizer:

- (9) der Hund **der** wo gestern d' Katz bissn hod *B. Ger*
 the dog who that yesterday the cat bitten has
 'The dog that bit the cat yesterday.'

- (10) **Cad** a^L tá sa seomra? *Irish*
 what C-*wh* is in.the room
 'What is in the room?'

- (11) thy freend **which that** thou has lorn *Variety of English*
 'Your friend that you have lost.'

(9)-(11) are examples supporting the idea that the landing site of *hé* is above the complementizer, namely, [Spec, CP]. In OC, *qí* can also be used in contexts in which it acts as a complementizer more clearly:

- (12) 其或繼周者 *(Analects, Wèizhèng)*
qí huò jì Zhōu zhě
 that someone succeed Zhou NOML
 'That someone succeed (throne) of the Zhou dynasty...'

In (12), *qí* is similar to English *that* in *that*-clauses, so there is reason to believe that in (8), the reason-*wh* *hé* 'why' lands at a higher position than C. Murphy (2017:196) points out that in Mandarin, the causal-*wh* *wèishénme* 'why' cannot occur below a modal, and this restriction also

holds in OC (13) because I have found no examples like “modal-*hé*” which should be ungrammatical in OC, but only “*hé*-modal”:

(13) 仍舊貫，如之何？何必改作？ (*Analects, Xiānjìn*)

réng	jiù	guàn,	rú	zhī	hé?	hé	bì	gǎi	zuò?
follow	old	warehouse	do	it	how	why	must	change	construct

‘How about following the old warehouse? Why must (we) change its construction?’

The linear order of *hé* and the modal *bì* shown in (13), together with the same restriction in Mandarin, serves as evidence that unlike the manner-wh *hé* ‘how’, the high adjunct causal *hé* ‘why’ moves to a position higher than T.

So far, we have seen that OC shares most of the important properties of wh-adverbials in Mandarin and other languages, such as a clear distinction between wh-nominal and wh-adverbial and in more detail, between low wh-adjunct and high wh-adjunct. The special property of OC wh-words is that all the above distinctions are not phonologically realized, namely, the same root *hé* (*g^haj)²⁷ simultaneously acts as low wh-adjunct ‘how’ and high wh-adjunct ‘why’. This peculiarity of *hé* leads us to resort to covert syntax in giving it reasonable explanations. Tsai (1999) and Murphy (2017) propose that only causal-whs undergo movement and they must have moved, otherwise their crossing of interveners will violate the principle of Relativized Minimality. This claim theoretically persuades us to believe that *hé* ‘why’ base-generates at the right periphery and it moves to the left.

The distinctions regarding island sensitivity between causal-whs and other wh-adjuncts inevitably lead to two different mechanisms (i.e. unselective binding for wh-nominals and ‘when’

²⁷ At least in current reconstruction systems, there is no evidence to link different syntactic positions/functions of *hé* with different pronunciations.

‘how’ ‘where’ vs LF movement for causal wh-adverbials) for the derivation of wide-scope of in-situ wh-expressions, which is not preferred within the *Minimalism*. And based on such reasons, Murphy (2017) advocates for a “unified theory” to deal with wh-in-situ. In order to resolve the mystery of causal-whs, Murphy (2017) first agrees with Aoun and Li (1993)’s operator binding approach which says the wh-operator must merge locally to wh-adjunct. To be specific, Aoun and Li (1993) reduce the distinction between in-situ causal-whs and other whs to whether they need to have a local antecedent. Murphy (2017:205) further points out the drawback of Aoun and Li’s solution such that their theory “overgenerates and predicts that all adjuncts should be subject to island constraints”, whose prediction is, apparently not correct. Ultimately, Murphy (2017) follows Rubin’s proposal of adding a functional structure within adjuncts and relies more on semantics contributed by the operator in [Spec, ModP] to explain how different in-situ wh-adverbials can be licensed in a uniform way while at the same time, have different island constraints.

Moreover, the phonological uniformity of different types of wh-adjuncts in OC has some implications for the architecture of grammar. According to the Late Insertion principle in Distributed Morphology (Bobaljik 2017), the pieces manipulated by syntax are abstract and lack phonological content. Applying this principle for the OC multifunctional wh-adjunct *hé*, we can say that the syntactic terminal node *hé* (*g^haj) is phonologically underspecified for wh-nominal or wh-adverbial when it enters the syntax. While this is also similar to Cheng (2000:11)’s dealing with the intonation Q-morpheme in French as being underspecified for [yes/no] or [wh] features. The case of OC wh-adjuncts exhibits “mismatches between the minimal units of grammatical combination and the minimal units of sound” (idem.) and is in line with the spirit of DM.

4.3 some inspiration from Japanese

Fujii et al. (2014) centers on distinctions between reason wh-adverbials and purpose/instrumental wh-adverbials, using data from Mandarin *wèishénme* ‘why’/*wéi-le-shénme* ‘for-PERF-what’ and Japanese *nande* ‘why, what.with’/*naze* ‘why’. The Japanese *nande* is similar to OC *hé* in that it has a “dual nature”: it can either get a reason interpretation or an instrumental interpretation. They discovered that when a wh-adverbial has reason interpretation, it obeys islands; while when having purpose or instrumental interpretations, it does not obey islands. And this is why *wèishénme* ‘why’ cannot be moved out of islands to take wide scope, but *wéi-le-shénme* ‘for-PERF-what’ can. Since *nande* has both interpretations, it can either obey or disobey island constraints. The Mandarin and Japanese data they analyzed show “systematic mapping relation between the meaning of a wh-adverbial and its syntactic category” (idem: 201), and this mapping can be summarized as: reason wh-adverbial-island sensitive-analyzed as AdvPs vs purpose/instrumental wh-adverbial-island insensitive-analyzed as PPs.

Fujii et al. (2014)’s generalization conforms with my analysis of treating *hé* as a PP when it means ‘how’. Although I did not find OC sentences involving islands that would allow a direct comparison between OC and Japanese, but a diagnostic they use that resorts to linear word order also applies to OC. Inspired by Murphy (2017), I have alluded to this diagnostic in (13), and here we can take a closer look at OC data. Fujii et al. (2014) find that the relative height of reason wh-adverbial (*nande^R*) and purpose/instrumental wh-adverbial (*nande^I*) can be disambiguated by their linear orders with regard to modals. The observation is, reason wh-adjuncts cannot follow the modal, while non-reason ones can. Take their Mandarin examples (14) and (15), for instance:

- (14) Aiku **huì** wèi shénme líkāi?
 Aiku will for what leave
 ‘For what purpose will Aiku leave?’

- (15) a.* Aiku **huì** wèishénme líkāi?
 Aiku will why leave
- b. Aiku wèishénme **huì** líkāi?
 Aiku why will leave
 ‘Why will Aiku leave?’

The reason wh-adjunct *wèishénme* cannot stay lower than the future modal, while the purpose one can. This is because *wèishénme* ‘why’ is located above T’. In a broader sense, the authors follow Rizzi’s idea and claim that *why*-type adverbs are special in that they are generated very high in clause structure.

We can use the same diagnostic to test the relative height between the reason-wh *hé* ‘why’ and the manner-wh *hé* ‘how’ in OC:

- (16) 說之將何如? (Mencius, GàoziXià)
- | | | | | |
|-------|-----|--------------|--------------------|-----|
| shuì | zhī | jiāng | [PP [NP hé] [P Ø]] | rú? |
| lobby | it | will | how | do |
- ‘How will (you) lobby him?’

- (17) 其生我也，將以何益？其殺我也，將以何損？ (Huáinánzǐ, Jīngshénxún)
- | | | | | | | | |
|----|------------|----|-----|--------------|-----|------|--------------|
| qí | shēng | wǒ | yě, | jiāng | [yǐ | hé] | yì? |
| it | give.birth | I | SFP | will | by | what | gain.benefit |
| qí | shā | wǒ | yě, | jiāng | [yǐ | hé] | sǔn? |
| it | kill | I | SFP | will | by | what | suffer.harm |
- ‘(If) it gives me life, what benefit will (I) gain? (If) it kills me, what harm will [I] suffer?’

In (16), the instrumental *hé* ‘how’ follows the modal ‘will’. Unlike Japanese *nande* or Mandarin *wèileshénme* where *de* ‘with’ and *wèi* ‘for’ are phonologically the same as a preposition, *hé* does

not contain such a prepositional morpheme in its overt morphology. There is an overt preposition *yǐ* ‘by’ in (17), which is semantically parallel to (16), and this serves as evidence to assume that there is a null preposition associated with *hé* in (16). Null prepositions are also proposed in Huang (1982:535) when he deals with noun-adverb distinctions in English. In contrast, I didn’t find any examples where the reason *hé* ‘why’ precedes *jiāng* ‘will’, but there are many examples where *hé* ‘why’ precedes another modal *qí*, for example (18):

(18) 何其聲之似我君也？ (Mencius, *JinxīnShàng*)

hé	qí	shēng	zhī	sì	wǒ	jūn	yě?
why	MP	voice	GEN	resemble	I	lord	SFP

‘How come is it that (his) voice resembles that of my lord?’

(19) 人雖欲自絕，其何傷於日月乎？ (Analects, *Zīzhāng*)

rén _i	suī	yù	zì	jué,	qí	hé
person	even.if	wish	self	cut.off	it	why
shāng	yú	rì	yuè	hū?		
injure	at	sun	moon	Q		

‘Even if a person wishes to cut himself off, how could (he) injure the sun and moon?’

It is worth mentioning that although both (18) and (19) contain the word *qí*, *qí* functions differently in these two examples: it is a modal in (18) but a demonstrative in (19). So, although the reason wh-word *hé* ‘why’ precedes *qí* in (18), (19) does not constitute an exception to the summary that reason-whs cannot scope lower than modals. Drawing an analogy from the Fujii et al. (2014)’s example (24) and based on the null preposition hypothesis, the internal structure of reason *hé* ‘why’ and manner/instrumental *hé* ‘how’ can be expressed in Figure 5:

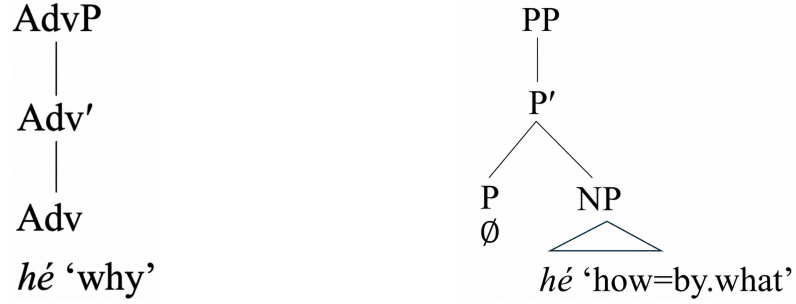


Figure 5.

Based on the discussion in this section, the adjunct analysis should be revised because only the *how-wh*-adjunct *hé* shall be treated as a PP, while the *why-wh*-adjunct *hé* is an AdvP. This conclusion explains why the *how-wh*-adjunct *hé* has a more restricted syntactic distribution—only when it occurs at the right periphery, it asks about pure manner. AdvPs in Mandarin generally have more flexible word order than PPs, as revealed by the comparison between (20) and (21):

- (20) a. tā hē kāfēi [AdvP mànmande]
 he drink coffee slowly
- b. tā [AdvP mànmande] hē kāfēi
 he slowly drink coffee
- c. hē kāfēi tā [AdvP mànmande]
 drink coffee he slowly
 ‘He drinks coffee slowly.’
- (21) a. tā [PP zài jiā=li] chī fàn
 he at home=inside eat rice
- b.?[?] [PP zài jiā=li] tā chī fàn
 at home=inside he eat rice

c.*	tā	chī	fàn	[_{PP} zài	jiā=li]
	he	eat	rice	at	home=inside
	‘He eats at home.’				

All of the word orders in (20) are grammatical, but only (21a) sounds totally natural. Therefore, it is natural to find that the *why*-wh-*hé* can freely occur at both peripheries in OC, but the *how*-wh-*hé*, which essentially is a PP, has a more restricted syntactic distribution.

4.4 categorizations of wh-words

In 4.2 and 4.3, I classified the OC wh-adjunct *hé* into two types: low wh-adjunct ‘how’ and high wh-adjunct ‘why’. By comparing OC to Mandarin and Japanese and providing some OC-internal evidence, I argued that the reason-wh *hé* undergoes leftward movement to land at [Spec, CP]. In this section, I discuss the categorization of wh-words in a broader perspective and further reflect on the distinction between reason-wh *hé* ‘why’ and manner-wh *hé* ‘how’.

Wh-words are cross-linguistically divided into different types due to the asymmetries among them, such as their sensitivity to island constraints, syntactic positions, theta roles, etc. The two basic types of wh-words are wh-arguments and wh-adjuncts and this distinction is further based on the broader distinction between arguments and adjuncts. According to Ackema (2015), some differences between arguments and adjuncts are: adjuncts are never obligatory; a prototypical argument in VPs is an NP, but a prototypical adjunct is a PP or an adverbial phrase, and in NPs, typical arguments are NPs or PPs, but typical adjuncts are PPs or APs; though both arguments and adjuncts can be wh-moved from an embedded clause to a matrix clause, it’s more difficult to extract an adjunct out of an island.

According to the summaries in Lin (2014) and Murphy (2017), at the highest level, wh-expressions can be divided into wh-nominals (arguments) and wh-adverbials (adjuncts), and among the wh-adverbials, there are low adjuncts like *how*, *when*, *where*, as opposed to high (causal) adjuncts like *why*. Low wh-adjuncts are believed to be adjoined to vP, but high wh-adjuncts are adjoined to TP. As for their sensitivity to island effects, high wh-adjuncts are island-sensitive, so they cannot be moved out of islands to take wide scope (22b), but low wh-adjuncts can move out of islands (22a):

- (22) a. nǐ xǐhuan chī [island tā **zěnmē** shāo de yáng ròu]?
 you like eat he how cook GEN lamb
 ‘What is the manner x, such that you like to eat things he cooked in manner x?’

- b.* nǐ xǐhuan chī [island tā **wèishénmē** shāo de yáng ròu]?
 you like eat he why cook GEN lamb
 Intended: ‘What is the reason x, such that you like to eat things he cooked for x?’

In English, the *why*-wh-adverbial is also different from other wh-adverbials (Fujii et al. 2014:183), that is, *where* and *when* can stay in-situ (e.g. *Who bought a book {where, when}*) but *why* cannot:
 * *Who bought a book why*. Tsai (2008:100), citing Collins’s work, points out that in the sentence *Why does everyone hate John*, *why* is raised to [Spec, CP] from a sentential adjunct position c-commanded by *everyone*, whereas *how come* is base-generated at C. The fact that OC reason-wh-*hé* can surface at the right periphery conforms with Tsai’s (2008) idea that high wh-adjuncts undergo movement from a lower position.

To continue our discussion of distinctions between *how*-wh-*hé* and *why*-wh-*hé*, let’s see another syntactic diagnostic. Tsai (2008:96) uses negation to test IEs and finds that negation blocks instrumental and manner questions but not reason questions. The same result holds in OC:

(23) 則凡可以辟患者，何不為也？

(*Mencius, Gàozǐ Shàng*)

zé fán kě yǐ bì huàn zhě, hé **bù** wéi yě?

then whatever be.able with avert calamity NOML why NEG do SFP

‘Then, as for whatever be able to avert calamity, why (would one) not do (it)?’

(24) 曾子怒曰，商，女何無罪也？

(*Lǐjì, Tán’gōng Shàng*)

Zēng zǐ nù yuē, Shāng, rǔ hé **wú** zuì yě?

Zeng master angry say Shang you how.come do.not.have guilt SFP

‘Master Zeng said angrily, Shang, how come you do not have guilt?’

(23) and (24) contain different negative markers, but the same thing is, when *hé* scopes above negations, the question cannot ask about manner or instrument. (23) and (24) differ trivially in contexts. (23) can still be understood to be a content question asking about reason, but (24) essentially is a rhetorical question. There are examples of “NEG-*hé*” in OC, but unsurprisingly, *hé* in these examples does not mean ‘why’:

(25) 星隕木鳴，國人皆恐。曰：是何也？曰：無何也

(*Xúnzǐ, Tiānlùn*)

xīng zhuì mù míng, guó rén jiē kǒng.

star fall tree groan state person all afraid

yuē, shì hé yě. yuē, **wú** **hé** yě.

say this why SFP say NEG whatever SFP

‘Stars fell and trees groaned. The people of the state were all afraid. (Someone) said, why is this (the case)? (Someone) said, it is nothing.’

(26) tā **bù** zěnmē yùndòng

he NEG however exercise

‘He does not exercise much.’

In (25), when *hé* scopes lower than negation, it is interpreted as a wh-indefinite, which is the same as the Mandarin example in (26).

Next, I show more examples from Japanese and French to illustrate the *how-why* alternations that OC and Mandarin exhibit. Japanese and Vietnamese, both of which are wh-in-situ languages, are also reported to have the same ‘how’-‘why’ alternation observed in Mandarin (Murphy 2017:197-200). In Japanese (Nakao and Obata 2009), the accusative wh-adjunct *nani-o* ‘what-ACC’ can be interpreted the same as the authentic wh-adjunct *naze* ‘why’:

- (27) a. kare-wa nani-o sawai-dei-ru no?
 he-TOP what-ACC make.noise-PROG-PRES Q
 ‘Why is he making a noise?’
- b. kare-wa naze sawai-dei-ru no?
 he-TOP why make.noise-PROG-PRES Q
 ‘Why is he making a noise?’

The authors also point out that *nani-o* ‘what-ACC’ and *naze* ‘why’ differ in animacy of the subject, speaker inference, and types of matrix predicate (bridge vs non-bridge verbs), sluicing, etc. What is the same for *nani-o* and *naze* is that they cannot be extracted from an adjunct island, which separates them from the argument *nani-o*, and this island constraint serves as an important diagnostic for the adjunct status of the accusative *nani-o*. Based on these differences and commonalities, the authors claim that *nani-o* is base-generated in a higher position in Functional Phrase but the authentic reason adjunct *naze* is base-generated in a VP-adjoined position. This view is particularly instructive for our discussion of the base-generating position of the reason-wh-*hé*. In OC, *hé-yǐ* ‘what-with’ is similar to the accusative *nani-o* ‘what-ACC’ in Japanese:

(28) 何以報德？以直報怨，以德報德。

(*Analects*, *Xiànwèn*)

hé	yǐ	bào	dé?	yǐ	zhí
what	with	repay	kindness	with	uprightness
bào	yuàn,	yǐ	dé	bào	dé
repay	resentment	with	kindness	repay	kindness

‘With what (shall one) repay kindness? Repay resentment with uprightness, and repay kindness with kindness.’

(29) 君子質而已矣，何以文為？

(*Analects*, *Yányuān*)

jūnzǐ	zhì	ér	yǐ	yǐ,	hé	yǐ	wén	wéi
gentleman	substance	CONJ	enough	SFP	what	with	refinement	SFP

‘For a gentleman, substance is enough; what need has (he) of refinement?’

Hé-yǐ ‘what-with’ occurs at left peripheries, and it can ask about either manner as a *how*-wh-phrase (28) or about reason as a *why*-wh-phrase (29). Following Nakao and Obata (2009)’s conclusion on different base-generating positions of *nani-no* and *naze*, it is reasonable to say that *hé-yǐ* ‘what-with, why/how come’ base-generates high at the left periphery but the bare *why*-wh-adjunct *hé*, like *naze*, base-generates low at a VP-adjoined position. In his discussion of the *how-why* alternation in Mandarin, Tsai (2007:204-205) mentions the difference between OC *hé-yǐ* ‘what-with’ and *yǐ-hé* ‘with-what’. Tsai refers to *yǐ-hé* as an “inner adverbial” like the English *how*, and refers to *hé-yǐ* as an “outer adverbial” like the English *why* or *how come*, which is exactly the pattern observed in (28) and (29).

In the optional wh-movement language French, *comment* ‘how’ remains in-situ (30) only when interpreted with an instrumental meaning (Aoun 1986; Tsai 1994:129), and this is exactly the same scenario as the *how*-wh-adjunct *hé* which stays in VP-internal positions:

(30) tu as ouvert la porte **comment?**
 you have opened the door how
 ‘How did you open the door?’

(31) a. tu es venu **pour** **quoi?**
 you are come for what
 ‘Why did you come?’

b.* tu es venu **pourquoi?**
 you are come why
 Intended. ‘Why did you come?’

In Tsai (1994), the contrast in (31) is used to provide a parallel example to discuss the asymmetry between Mandarin reason-*why* *wèishénme* ‘why’ and purpose-*why* *wèi-le-shénme* ‘for-PERF-what’ which was noticed earlier by Lin (1992). In Mandarin, *wèishénme* ‘why’ is like the French *pourquoi* (31a) in allowing wide-scope interpretation out of an island, while *wèi-le-shénme* ‘for-PERF-what’ is like the French *pour-quoi* ‘for-what’ (31b) in being sensitive to island effects. Aldridge (2010:34-35) points out that the wh-argument *hé* ‘what’ can take matrix scope in complex NP (32) and adjunct (33) islands:

(32) 天何慾何惡者也？ (Mòzǐ, 4)

tiān [DP CP [hé_i yù t_i hé_j wù t_j] zhě] yě?
 heaven what desire what despise NOML SFP
 ‘Heaven is one who desires what and despise what?’

(33) 何恃而不恐？ (Guóyǔ, Lǚ 1)

[hé_i shì t_i] ér bù kǒng?
 what depend CONJ NEG fear

‘Based on what are (you) not afraid?’

Unfortunately, I didn’t find any examples where the wh-adjunct *hé* is involved in islands. Due to limitations of the corpus data, we cannot distinguish between the *how-wh-hé* and the *why-wh-hé* based on their sensitivity to island effects. But based on their different linear word orders with regard to modals, it is clear that the wh-adjunct *hé* ‘how, why’ has two subtypes and the two types of *hé* follow different syntactic derivations, which is a conclusion predicted by Kayne (1994)’s Linear Correspondence Axiom (LCA). The LCA claims that if two elements differ in linear word order, they must also differ hierarchically. This means the fact that the *how-wh-adjunct hé* tends to stay VP-internally, while *why-wh-adjunct hé* tends to stay at the left periphery, naturally predicts that the two types of *hé* have a hierarchical distinction in LF.

Chapter V Deriving a Feature Movement Analysis of *Hé*

In Chapter IV, we have reached the conclusion that the *how-wh-hé* base-generates low in VP-adjoined positions and stays in-situ; the *why-wh-hé* base-generates low at the right periphery though, moves leftwards to check the [wh] feature at [Spec, CP], thus can be analyzed as an ex-situ wh-word. Following this analysis, Chapter V aims to figure out details of the movements involved with the wh-adjunct *hé*. A crucial question that needs to be answered is what element moves. I first clarify the necessity of a movement analysis, then briefly introduce the notion of feature movement under the Minimalist Program. In what follows, the attention of this chapter turns to two kinds of question particles: $Q_{[wh]}$ and $Q_{[y/n]}$. The elaboration of these two question particles centers on two aspects from a cross-linguistic perspective—whether the two particles are

morphologically made distinct, and whether a wh-word can co-occur with a question particle in a minimal clause.

5.1 necessity of a movement analysis

One of the necessities of a movement analysis to deal with the wh-adjunct *hé* comes from Intervention Effects (IEs). IEs “refer to the phenomenon where an interrogative wh-expression cannot follow a quantifier or a focus expression” (Li 2024). Tsai (2008:97)’s formalization of IEs is shown in (1):

- (1) * wh_i...X...wh_i..., where X is a scope-bearing element, and the wh-dependency is created by LF movement.

The existence of IEs implies that there are LF movements happening but are blocked by interveners. Therefore, if IEs are operative in the wh-adjunct *hé*, we can further verify the existence of wh-movement of *hé* in LF. Next, I use Mandarin examples to explain the IEs, then we take a look at OC.

Li (2024:12), based on Tsai (2008)’s differentiation of high and low wh-adjuncts, points out that the high-adjunct *wèishénme* ‘why’ cannot follow the strong quantifier *dōu* ‘all’ (2a), and the low-adjunct *zěnmeyàng* ‘how’ cannot precede this intervener *dōu* ‘all’ (2b):

- (2) a.* měi=gè xuéshēng **dōu** wèishénme bù lái? (Li 2024:12)
 every=CL student all why NEG come
 Intended: ‘Why do not each of the students all come?’
- b.* měi=gè xuéshēng zěnmeyàng **dōu** huì mǎi fáng
 every=CL student how all will buy apartment

Intended: ‘How will each student all buy apartments?’

In OC, we can actually find examples where *hé* follows a quantifier, such as (3):

(3) 皆何以然?

(*Shuōyuàn, Xiūwén*)

jiē *hé* *yǐ* *rán*

all what by so

‘By what are all (these things) so?’

But in (3), *hé* is not a wh-adjunct. Rather, it is a fronted object. There are many cases of *jiē-hé* ‘all-what’ in the corpus, but in none of them does *hé* act as a wh-adjunct. The absence of such examples indirectly suggests that as with Mandarin, OC does not allow the *why-wh-hé* to follow, at least the quantifier *jiē* ‘all’. There are also no examples of *hé-jiē*, even when *hé* is a wh-nominal, let alone when *hé* is a *how-wh-adjunct*. Viewing from the quantifier *jiē* ‘all’, it seems that IEs also hold in OC because the *why-wh-hé* cannot follow *jiē* ‘all’ and the *how-wh-hé* cannot precede *jiē* ‘all’. But there are potential counterexamples to this generalization:

(4) 孔子曰，回！來！若獨何不願乎？

(*Shuōyuàn, Zhǐwǔ*)

Kǒngzǐ *yuē*, Huí, lái, ruò **dú** **hé** bù yuàn hū

Confucius say Hui come you alone why NEG desire Q

‘Confucius said, Hui, come here, why do you alone desire nothing?’

(5) 是其始死也，我獨何能無慨然？

(*Zhuāngzǐ, Wàipiān, Zhìlè*)

shì *qí* *shǐ* *sǐ* *yě*, *wǒ*

this she begin die SFP I

dú **hé** *néng* *wú* *kǎirán*

alone how.come be.able without sorrow

‘When she first died, how come I alone was able to be without sorrow?’

(6) a.*	nǐ	dāndú	wèishénme	méiyǒu	yuànwàng
	you	alone	why	not.have	aspiration
b.?	nǐ	wèishénme	dāndú	méiyǒu	yuànwàng
	you	why	alone	not.have	aspiration
c.	wèishénme	dāndú	nǐ	méiyǒu	yuànwàng
	why	alone	you	not.have	aspiration
‘Why you alone have no aspirations?’					

In (4) and (5), the *why-wh-hé* follows *dú* ‘alone’. The adverb root *dú* ‘alone’ is still used in Mandarin, but disyllabifies as *dān-dú* ‘single-alone’. However, the Mandarin equivalent of (4) is ungrammatical, as indicated by (6a). (6b) and (6c) further tell us that when *wèishénme* ‘why’ is raised a little higher to scope above ‘alone’ but lower than the subject, the sentence becomes more acceptable but still not natural. It’s only when *wèishénme* ‘why’ is raised to scope above both ‘alone’ and the subject, that the sentences become fully natural. Whether *dú* ‘alone’ in OC can be regarded as an intervener like *jiē* ‘all’ is not certain, but the comparison between (4) and (6a) reveals that *why-wh*-words in OC and Mandarin are subject to different syntactic restrictions. Although both *dú* ‘alone’ and *jiē* ‘all’ look alike in semantics as quantifiers, *dú* almost always occurs in negative sentences when it co-occurs with *hé*, while *jiē* does not have such dependency on negation. If *dú* ‘alone’ really is an intervener, the grammaticality of (4) and (5) implies that LF movement of the *why-wh-hé* does not happen, otherwise the movement should have been blocked by *dú* ‘alone’ and renders ungrammaticality.

Another diagnostic Li (2024) employs to distinguish between the two types of *hé* is their linear orders relative to modals, which is something we have tested on the modal ‘will’ in (IV.14-17).

Here we compare the modal ‘be.able’ in OC and Mandarin:

- (7) a. Lǐ Bái **wèishénme** **néng** xiū chē?
 Li Bai why be.able fix car
 ‘Why is Li Bai able to fix cars?’

- b.* Lǐ Bái **néng** xiū **wèishénme** chē?
 Li Bai be.able fix why car
 Intended: ‘Why is Li Bai able to fix cars?’

- (8) a.* Lǐ Bái **zěnmeyàng** **néng** xiū chē?
 Li Bai how be.able fix car
 Intended: ‘What is the manner *x*, such that Li Bai is able to fix cars in manner *x*?’

- b. Lǐ Bái **néng** **zěnmeyàng** xiū chē?
 Li Bai be.able how fix car
 ‘How can Li Bai fix cars?’ (adapted from Li 2024)

- (9) 不罪其吏，而徒攻社，能何復塞？ (Lùnhéng, Shùngǔ)

bù	zuì	qí	lì,	ér	tú
NEG	blame	those	official	but	merely
gōng	shè,	néng	hé	fù	sè
attack	altar	be.able	how	further	remedy

‘(If one) does not blame those officials but merely attacks the altars, how could (this) be able to further remedy (the disaster)?’

- (10) 不畏于天，將何能保？ (Zuǒzhuàn, Wéngōng Shíwǔnián)

bú	wèi	yú	tiān,	jiāng	hé	néng	bǎo
----	-----	----	-------	-------	-----------	-------------	-----

NEG awe towards heaven will why be.able preserve
'(If one) does not stand in awe of Heaven, why can (one) preserve (the state)?'

(7)-(9) show that OC and Mandarin follow the same rule here: 'why' cannot follow 'be.able' and 'how' cannot precede 'be.able'. This means that IEs also exist in OC, which is also the conclusion Wang (2021) reaches. As Li (2024:14) points out that "IEs are usually analyzed as a failure of covert movement or failure in forming a dependency", IEs serve as an important diagnostic for covert movement and feature movement. However, OC data exhibits some heterogeneities: (4) indicates the absence of LF movement but (9) and (10) prove the existence of LF movement. In order to reach a unified analysis of IEs in OC, it is natural to resort to a solution of feature movement. In theory, *wh*-adverbials are generally believed to move. For example, Reinhart (1998:30) believes that *wh*-adverbials are "never very good in situ" but surfacing positions of *why*-*wh-hé* (i.e. both left and right peripheries) clearly indicate that it can remain in-situ. This apparent contradiction also drives us to the feature movement analysis because the core idea of feature movement is that in-situ *wh*-elements do not need to move in LF, rather, only features move. It is also suggested in Cheng and Rooryck (2000) that *wh*-adverbials involve feature movement in LF. In the next section, I elaborate on feature movement in a more general context.

5.2 feature movement

Before we get into details of feature movement, it's necessary to review the core argument that supports the LF movement analysis of *wh*-in-situ. An important reason to propose LF movement for in-situ *wh*-elements is that although they do not move in overt syntax, they can take wide-scopes. (11) is an example adapted from Cheng (2009:770) to illustrate the embedded and wide scopes of in-situ *wh*-phrases:

(11) [_{CP1} {Which student} _{C1'} knows [_{CP2} where _{C2'} [Mary bought {which book}]]]]?

- a. {Bill, Mike, Hery} knows where Mary bought which book.
- b. Bill knows where Mary bought {book A, book B, book C} and Sue knows where Mary bought {book D, book E, book F}.

In (11), *which student* and *where* move to the specifier positions of each CP respectively while *which book* remains in-situ. In the answer (11a), each individual in the set provides a potential answer to the matrix wh-phrase *which student* in (11), therefore is not answering the in-situ wh-phrase and in this case *which book* has an embedded scope. In contrast, each individual in the set of (11b) answers the in-situ wh-phrase, and *which book* has a wide scope in this interpretation.

Starting from the 1990s, other approaches were put forward to deal with wh-in-situ. Lin (2014:193-202) summarized different approaches to wh-in-situ into four major theories: the quantificational theory, the operator movement theory, the unselective binding theory, and the alternative semantics theory. Among these four theories, the quantificational theory is what the standard covert movement approach adopts. In analyzing wh-phrases as quantificational operators, they move at LF to produce a quantifier-variable representation. In Huang (1982)'s covert movement analysis, Mandarin wh-words move to the same position in LF just as where wh-words overtly move in English, namely, [Spec, CP]. Aldridge (2010) analyzed OC wh-argument *hé* 'what' as wh-indefinites rather than quantificational operator and they move to a clause-internal position—the edge of *vP*. The quantificational analysis of wh-in-situ seems to presuppose that moved (overtly or covertly) wh-elements land at a position higher than TP that can scope over the whole TP, but the non-quantificational analysis does not require them to move that high. In OC, an overtly moved wh-word can land as high as [Spec, CP] and at the same time be wh-indefinite.

Under the minimalist framework, we should expect “movement to operate just with formal features, rather than categories” (Hornstein et al. 2005:324, cf. Cheng (2000)). More specifically, “the movement parameter are to be accounted for not in terms of a given feature (strong or weak), but in terms of presence or absence of strong features” (Hornstein et al. 2005:303). So, it is reasonable to hypothesize that if *why-wh-hé* stays in-situ, what actually moves is only the initial **g^ɕ* of *hé* (**g^ɕaj*) which serves as morphological representation of the strong [wh] feature, and what has been left in-situ is the final **-aj* which bears the weak [*Q_{y/n}*] feature.

Some preliminary evidence for this morphological decomposition has been provided in 1.2.2 and 2.3. Here I provide four more pieces of OC-internal evidence to support this analysis. Firstly, the sentence-final yes-no-question particle 乎 *hū* (**g^ɕa*) shares a similar low vowel as *hé* and it strictly stays at the right periphery. Secondly, there exists a **k-* prefix in pre-OC (see footnote 2. in Matisoff 2003:87). Thirdly, early in the 1950s, Wáng Lì classified OC pronouns into three series according to their initial types: the **z-* series refer to people, the **ɣ-* series refer to things and the zero-initial series refer to locatives (Wang 1980 [2015]: 279-284). This means that initials affect semantics of wh-words in OC, and this is the philological foundation to propose that the initial of *hé* bears a [wh] feature. Lastly, in his generative study of OC grammar, Mei (2018:5-16, 183) also decomposes grammatical words in OC to explain their syntactic behavior. For instance, in 言 *yán* (**ŋan*), 焉 *yān* (**ʔan*) and 爰 *yuán* (**g^wan*), the final **-an* is a separate morpheme that has connecting and anaphoric functions. All of these pieces of evidence demonstrate that an element smaller than a syllable in OC, for example a velar stop sound can be a meaning/feature-bearing unit.

5.3 *Q_[wh]* and *Q_[y/n]*

Based on the discussion of feature movement in 5.2, this section focuses on two major question particles: $Q_{[wh]}$ and $Q_{[y/n]}$. Question features can be instantiated in different ways. For example, sentence-final particles in Chinese and Japanese bear question features while in French, an intonation morpheme can bear the question feature (Cheng and Rooryck 2000). Since I mainly use Chinese examples in this study, it is worth mentioning the issue of question particles under the context of feature movement.

The Clausal Typing Hypothesis (CTH) (Cheng 1991) predicts that no languages employ question particles and syntactic wh-movement at the same time, and no languages have “optional movement” of wh-words. Aldridge (2010) analyzes the fronting of the wh-argument *hé* ‘what’ in OC as syntactic wh-movement and since OC uses question particles as well, she claims OC “seems to be a blatant violation of the Clausal Typing Hypothesis”. According to this study, the wh-adverbial *hé* in OC also violates the prediction of CTH. Cheng (1991:52-108) also realizes that there are counterexamples of CTH, like optional wh-fronting languages (e.g. French) and multiple wh-fronting languages (e.g. Polish). Besides these typical exceptions, Tan and Shen (2025) argue that colloquial Singapore English is a language with both true wh-in-situ and true wh-movement, which also challenges the CTH.

5.3.1 question particles in OC

In this subsection, I introduce basic facts about Q-particles in OC and clarify some controversial points. There are several sentence-final particles that can be used in questions in OC, including 乎 *hū* (* g^{fa}), 歟 *yú* (* la^{28}), 也 *yě* (* grA), etc. Widely attested though, *hū* is not purely a

²⁸ This character is not included in Baxter and Sagart (2014)’s reconstruction system. So the reconstruction is taken from Schuessler (2008:55).

question particle. Rather, it's a complex SFP related to interrogativity and modality. Based on Yao (2015:370-372), I give prototypical usages of *hū* in (12)-(15):

(12) 秦伯曰，晉國和乎？對曰，不和。 (Zuǒzhuàn, Xīgōng Shíwǔnián)

Qínbo yuē Jìn-guó hé **hū?** duì yuē, bù hé.
 Qinbo say Jin-state peace Q reply say NEG reconcile
 ‘Qinbo said, will the state of Jin make peace? The reply said, it will not.’

(13) 吾何執？執御乎？執射乎？ (Analects, Zǐhǎn)

wú hé zhí? zhí yù **hū?** zhí shè **hū?**
 I what expertise expertise charioteering Q expertise archery Q
 ‘What is my expertise? Is my expertise charioteering? Is my expertise archery?’

(14) 夫今之歌者其誰乎？ (Zhuāngzǐ, Shānmù)

fú jīn zhī gē zhě qí shuí **hū?**
 COMPL nowadays GEN sing NOML MP who Q
 ‘Nowadays, who indeed is the one who sings?’

(15) 學而時習之，不亦說乎？ (Analects, Xuéér)

xué ér shí xí zhī, bú yì yuè **hū?**
 study CONJ frequently review it NEG also happy Q
 ‘Study and frequently review it, isn’t it also a pleasure?’

(12) clearly shows that *hū* can be used as a yes-no question particle in OC because the second clause uses the negative marker *bù* to answer the question. *Hū* in (13) has a subtle difference from that in (12), although both of them look like yes-no question particles. But if we take into consideration the context of (13), and refer to the definition that “a yes-no question asks for a confirmation or denial of a single proposition” (Huang et al. 2009:236), it can be detected that *hū* here actually is asking the addressee to choose between two propositions, and thus forms a

disjunctive question rather than asking for the kind of answer in (12), which is an answer to a single proposition. (14) is tricky in that it seems that a wh-word *shuí* ‘who’ can co-occur with a $Q_{[y/n]}$ in OC²⁹, which is strictly forbidden in modern Chinese. However, *hū* in (14) is a SFP equivalent to Mandarin *ne*, rather than a $Q_{[y/n]}$. When *hū* is equivalent to *ne*, it normally co-occurs with a modal *qí*, while the Mandarin equivalent of *qí-shuí-hū* ‘MP-who-SFP’ is *shì-shuí-ne* ‘FOC-who-SFP’, with the wh-word being marked by the focus marker *shì*. *Hū* in (15) has the same nature as that in (14), namely as an SFP, and it is c-commanded by negation in (15). We can also find the equivalency of *hū* in (14) and (15) to Mandarin *ne* from the fact that the interpretations of (14) and (15) have presuppositions of the addressee: (14) actually implies ‘there is nobody who can still sing’, while (15) implies ‘it is happy’. Now, let’s see some Mandarin examples:

(16) a. wǒ=de tècháng shì qí mǎ háishì shèjiàn ***ma/(ne)?**
 I=GEN expertise be ride horse or archery $Q_{[y/n]}/SFP$
 ‘Is my expertise riding horses or archery?’

b. wǒ=de tècháng shì qí mǎ **ma?**
 I=GEN expertise be ride horse $Q_{[y/n]}$
 ‘Is my expertise riding horses?’

c. wǒ=de tècháng shì qí mǎ **ne?**
 I=GEN expertise be ride horse SFP
 ‘How come that my expertise is riding horses?’=
 ‘My expertise is not expected to be riding horses.’

d. wǒ=de tècháng shì bú shì qí mǎ **ne/*ma**
 I=GEN expertise be NEG be ride horse SFP/ $Q_{[y/n]}$

²⁹ They can co-occur, but not in cases like (14). See discussion of (20)-(23) below.

‘Is my expertise riding a horse’

(16a) is a Mandarin equivalent of (13) and is referred to as a disjunctive question. (16d) is an A-not-A question, which “request[s] the addressee to choose between a positive and negative alternative provided in the question” (Huang et al. 2009: 236). A-not-A questions are sometimes regarded as a subtype of disjunctive questions (idem), because both types (16a, d) are compatible with the SFP *ne*, but not the yes-no question particle *ma*. The real yes-no question, on the other hand, shows the opposite pattern in its compatibilities of *ne* and *ma* (16b, c). (16b) is a real yes-no question, while if *ma* is changed to *ne*, (16c) can only be understood as the speaker having a negative presupposition about the proposition being asked.

The above descriptive facts raised the necessity to clarify the function of the particle *ne*. In generative literature, *ne* is sometimes treated as a yes-no Q-particle like *ma* (e.g. Aoun and Li 1993; Lin 2014) or as an “optional wh-question particle” (Cheng et al. 1996:44). However, it may not be precise to describe *ne* as a wh-question particle, no matter if it is optional or not. Shi (2023:30) already pointed out that *ne* in Mandarin is not a genuine Q-particle. Rather, he points out that *ne* “confirms the actuality of an event”, and “when the particle *ne* is added to a wh-question, the force of the question is weakened; in most cases, wh-questions with this particle are actually rhetorical questions” (idem). An example of wh-question ending with *ne* is given in (17):

- (17) *nǐ* *xǐhuan* *shuí* ***ne***
 you like who SFP
 ‘Who do you like?’ (asking in a polite way)

(17) is of no difference with *nǐ-xǐhuan-shuí* ‘you-like-who’ regarding their semantics, while they only differ in pragmatics.

Though he correctly points out a problem in Aoun and Li (1993), Shi himself also made a descriptive mistake concerning the OC yes-no question Q-particle, which is very important to our theoretical concerns. Shi (2023:31) believes that OC Q-particle *hū* is the same as Mandarin *ma* in that they both cannot co-occur with *wh*-expressions within a minimal clause. As for Mandarin, this is exactly what (18) shows:

- (18) a.* nǐ xǐhuan **shuí** ma?
 you like who Q
 Intended: ‘Who do you like?’

- b. ni xǐhuan **shui** ma?
 you like whoever Q
 ‘Do you like anyone?’

- c. [CP1 tā zhīdào [CP2 nǐ xǐhuan **shuí^A/shuí^B** ma]?
 he know you like who/whoever Q
 A: ‘Who is the person *x*, such that he knows that you like *x*?’
 B: ‘Does he know that there exists a person *x*, such that you like *x*?’

- (19) Zhāng Sān zhīdào shuí mǎi=le shū?
 Zhang San know who buy=PERF book
 A: ‘Who does Zhang San know bought books?’
 B: ‘Zhang San knows who bought books.’

Both (18a) and (18b) contain only one clause while the former one is ungrammatical, but the latter one is grammatical. The difference is that iff *shuí* ‘who’ is realized in a neutral tone, it is compatible with *ma* in a minimal clause (18b). Otherwise, it needs to be not in the same minimal clause with *ma* (18c, interpretation A). This means that when *shuí* ‘who’ (in its citation tone) wants to take

matrix scope, it must not co-occur with *ma* in a minimal clause. (19) is a classic example from Huang (1982 [2014]:151) which is used to support the LF movement of Mandarin in-situ wh-elements. In other words, the reason (19) is ambiguous is that there is no *ma*. The effects of phonological realization of a wh-word on sentential interpretation as shown in (18) also apply at (19). Traditionally, the two interpretations in (19) are not explained by taking into consideration the phonology of *shuí* ‘who’, which can be seen from the fact that the Mandarin transcriptions were originally not marked with tones.

Cheng and Rooryck (2000:14) pointed out that the French intonation Q-morpheme is a root morpheme, and this explains why French in-situ wh-words cannot stay in embedded clauses. Similarly, Mandarin *ma* and OC *hū* are also root morphemes, thus both of them can potentially check a feature in the matrix clause. They also point out that in French, “either the intonation morpheme or wh-movement checks the Q feature represented by *est-ce que*, but not both” in which *est-ce que* is what they call “overt Q morpheme” (idem:11). And this also explains the ungrammaticality of (18a) in that both *shuí* ‘who’ and *ma* can check the Q feature on C⁰ ([yes/no] or [wh]), but there is only one specifier position in matrix that bears the Q feature. In summary, I totally agree with Shi’s observation of the restrictions exemplified by Mandarin data, but this is not the case in OC:

(20) 白其蹄者以何事乎？ (Lùnhéng, Shízhī)

bái	qí	tí	zhě,	yǐ	hé	shì	hū?
CAUS.white	it	hoof	NOML	for	what	do	Q

‘One who makes its (black horse’s) hoof white, for what does he do that?’

(21) 眾多欲以何趨乎？ (Lùnhéng, Zìjì)

zhòng	duō	yù	yǐ	hé	qū	hū?
-------	-----	----	----	----	----	-----

the.many generally wish for what strive Q
 ‘For what do the many generally wish to strive?’

(22) 何為泣乎? (Lùnhéng, Féngyù)

hé wéi qì **hū**
 why do cry Q
 ‘Why do you cry?’

(23) 治天下將奈何乎? (Hánfēizǐ, WàichǔshuōShàng)

zhì tiānxià **jiāng** nài hé **hū?**
 govern All-under-Heaven will do what MP
 ‘How will (he) govern All-under-Heaven?’

In most of the cases where a wh-word co-occurs with *hū* in a minimal clause, there is a modal in the matrix clause, like *jiāng* in (23), and in this case, *hū* is a modal particle rather than a real Q_[y/n]. But I do find three examples in which *hū* is a real Q_[y/n] and it co-occurs with a wh-word in a minimal clause, as can be seen in (20)-(22) where there are no modals in the matrix clauses. These kinds of sentences are very rare even in OC, while (20)-(22) come from the same text. This inevitably raises questions about their reliability. But scholars generally believe that *Lùnhéng* is a reliable LOC text and it in some sense reflects the colloquial language at that time (ca. 80 CE). I will leave the question of whether (20)-(22) is a dialectal phenomenon or reflects some historical changes within the OC period but simply based on a synchronic view and assumes that a wh-word was permitted to co-occur with a real Q_[y/n] in the colloquial language of OC. It is then worth considering how native speakers of OC and modern Chinese have different mental grammars regarding wh-elements.

5.3.2 Q_[y/n] and Q_[wh] in other languages

In 5.3.1, I pointed out that unlike modern Chinese, OC wh-words can co-occur with a yes-no-question particle in the same minimal clause. To better understand this important difference between OC and Mandarin, I present some Japanese and Korean data in this subsection. Discussion of such co-occurrence patterns aims to reflect on the relationship between wh-word and question particle when they co-occur. Furthermore, this relationship has implications on our movement analysis because if one of them is a copy of another, it clearly indicates that one of them moves. I will address this issue in more detail when the Copy Theory of Movement is mentioned in Chapter VI. In this subsection, I focus on empirical data.

In Japanese and Korean, both of which are also wh-in-situ languages (24, 27a), wh-words can freely co-occur with a Q_[y/n] in a minimal clause (25, 28):

- (24) Mary-ga nani-**ka**-o yonda
 Mary-NOM what-Q-ACC read.PAST
 ‘Mary read something.’

- (25) nani-**ka** kaimashita **ka**
 what-Q bought.polite Q
 ‘Did you buy anything?’

- (26) [CP₁ John-ga [CP₂ Mary-ga nani-o katta **ka**] sitteiru] (Hagstrom1998:16)
 John-NOM Mary-NOM what-ACC bought Q know
 ‘John knows what Mary bought.’

The Japanese Q-particle *ka*³⁰ is referred to as a “clause-typing particle” because it always occurs sentence-finally (McCready and Davis 2020:657) while it can also mark the indefiniteness of a noun (*nani-ka* in 24), and it is commonly believed to be the same morpheme as the clause-final *ka*, though some scholars believe that the indefinite *ka* and clause-final *ka* are accidental homophones (idem:659). Although it is “clause-typing”, *ka* is not a root morpheme like the French intonation morpheme and Mandarin *ma*, because *ka* can also appear in embedded environments (26). And what makes the embedded *ka* different is that it cannot be omitted when marking an embedded interrogative sentence but can be omitted when marking a matrix interrogative (idem:657). This is because in Japanese the Q feature on C⁰ can either be checked by *ka* or by an intonation morpheme which is also a root morpheme. And due to this, the intonation morpheme cannot check the Q feature on the embedded C, thus the embedded *ka* must be retained.

Also see some Korean examples:

- (27) a. Mary-ka mwues-inka s-ass-ta
 Mary-NOM what-INDEF buy-PAST-DECL
 ‘Mary bought something.’
- b. mues-inka s-ass-ni
 what-INDEF buy-PAST-Q
 ‘Did you buy anything?’

Korean is similar to Japanese in that it also adds a suffix to wh-words to form wh-indefinites, but unlike Japanese suffixal *ka* is phonologically the same as the Q-particle *ka*, Korean suffixal *inka*, also sometimes transcribed as *ina* or *nka*, according to their phonological environments, is opaque

³⁰ Japanese has another Q-particle *no* which comes from *no da* or *no desu*. *No* is more often used by females or children in informal situations (Makino and Tsutsui 1986 [2022]: 322-324).

in its phonological relation to *ni* ‘Q’. On the other hand, Korean also has other forms of Q-particles, such as *ka*. The Q-particle *ka* is rarely used in colloquial Korean, rather, it’s historically used in colloquial Korean while nowadays it’s only preserved in formal written language. But whether it be *inka/ina/nka* or *ni/ka*, there more or less are some phonological correlations between them, namely, same main vowels (i.e. [i] or [a]). Cheng and Rooryck (2000:12) mention that the Q-particle *ni* in standard Korean is underspecified for [wh] or [y/n] features, as indicated by the ambiguity of (28):

- (28) *nwu(kwu)-ka* *pakkey* *w-ass-ni* (Choe 1994:276)
 who/someone-NOM outside come-PAST-Q
 ‘Who is at the door?’ or ‘Is there someone at the door?’

- (29) a. *Swuni-ka* *NWU³¹-lul/*nwu-lul* *cohala-no* (idem:278-279)
 Swuni-NOM *who-ACC/someone-ACC* *like-Q_[wh]*
 ‘Who does Swuni like?’; * ‘Does Swuni like someone?’

- b. *ni-ka* *nwu(-lul)/*NWU-lul* *cohaha-na*
 you-NOM *someone-ACC/*who-ACC* *like-Q_[y/n]*
 ‘Do you like someone?’; * ‘Who do you like?’

Though the Q-particle is ambiguous, *nwu* employs phonological apparatuses to disambiguate a wh-word (with emphatic stress) from an indefinite pronoun (without stress), as instantiated by the comparison between (29a) and (29b). While this is exactly the same strategy as how Mandarin disambiguates a wh-word (citation tone) from an indefinite pronoun (neutral tone).

³¹ Since in Korean, *-kwu* can be omitted when *nwu* means ‘someone’, Choe uses capital letters NWU to refer to the wh-word *nwu*, while small letters *nwu* exclusively refers to ‘someone’.

Cross-linguistically, there seems to be a tendency for one of the wh-words in a language and its indefinite pronoun to share the same morphological base while at the same time, wh-words are inclined to choose the heavy phonological form (i.e. citation tone; stressed) and indefinite pronouns to choose the light phonological form (i.e. neutral tone; unstressed). However, in some Korean dialects like Kyengsang dialect (Choe 1994:278-279), the wh-question particle is morphologically different from the yes-no question particle, as (29) shows. So, we can say that in Kyengsang dialect the Q-particle is underlyingly specified for either a [wh] feature or a [y/n] feature, and we noticed that *na* ‘Q_[y/n]’ also shares the main vowel [a] with the suffixal indefinite morphemes.

The important takeaways of this chapter are: the existence of IEs in OC supports the LF movement of *hé*, and the movement is realized by means of moving features; wh-words and Q-particles show some morphological correlations cross-linguistically

Chapter VI Relating *Hé* to the Q-Particle *kəʔ* in Jiānghuái Mandarin

In the preceding two chapters, the sentence-final *why-wh-hé* is analyzed as an in-situ wh-word which moves leftward to the left periphery, and what moves in LF is its initial **g^ℓ*- which bears the [wh] feature. There is another logical possibility though, which is to treat the left-periphery *hé* as being in-situ and it moves rightward to land in sentence-final positions. To demonstrate this possibility, the attention of Chapter VI turns to modern languages. The reasoning here proceeds as follows: *hé* is etymologically related to the preverbal question particle *kəʔ* in Jiānghuái

Mandarin; $kəʔ^5$ undergoes the rightward Neg^0 -to- C^0 movement; Universal Grammar of native speakers of the same language do not change over time³².

Chapter VI is structured as follows: 6.1 introduces the synchronic distribution of preverbal Q-particles in modern Chinese dialects; 6.2 discusses the etymology of these preverbal Q-particles and sketches their history; 6.3 and 6.4 show empirical data of four kinds of yes-no questions in Chinese, revealing that our main concerns—K-VP questions exhibit the most similarities to A-not-A questions and Negative Particle Questions; 6.5 derives the rightward Neg^0 -to- C^0 movement for K-VP questions and evidence supporting this derivation comes from both OC and dialectal data.

6.1 preverbal Q-particles in Chinese dialects

Chinese languages canonically use sentence-final (i.e. postverbal) question particles in yes-no questions, but some Chinese dialects and Tibeto-Burman languages (e.g. Naxi Yi, Rouruo, Jingpo, Dulong, Muya Qiang, see Yue-Hashimoto 1991) also employ preverbal question particles. These dialects mainly include Southwest Mandarin, Jiānghuái Mandarin, and Min dialects, as indicated by the read symbols in circles 1, 2 and 3, respectively, in Figure 6 (Cao *eds.* 2008, picture 101):

³² See the “uniformitarian hypothesis” of UG in footnote 8.

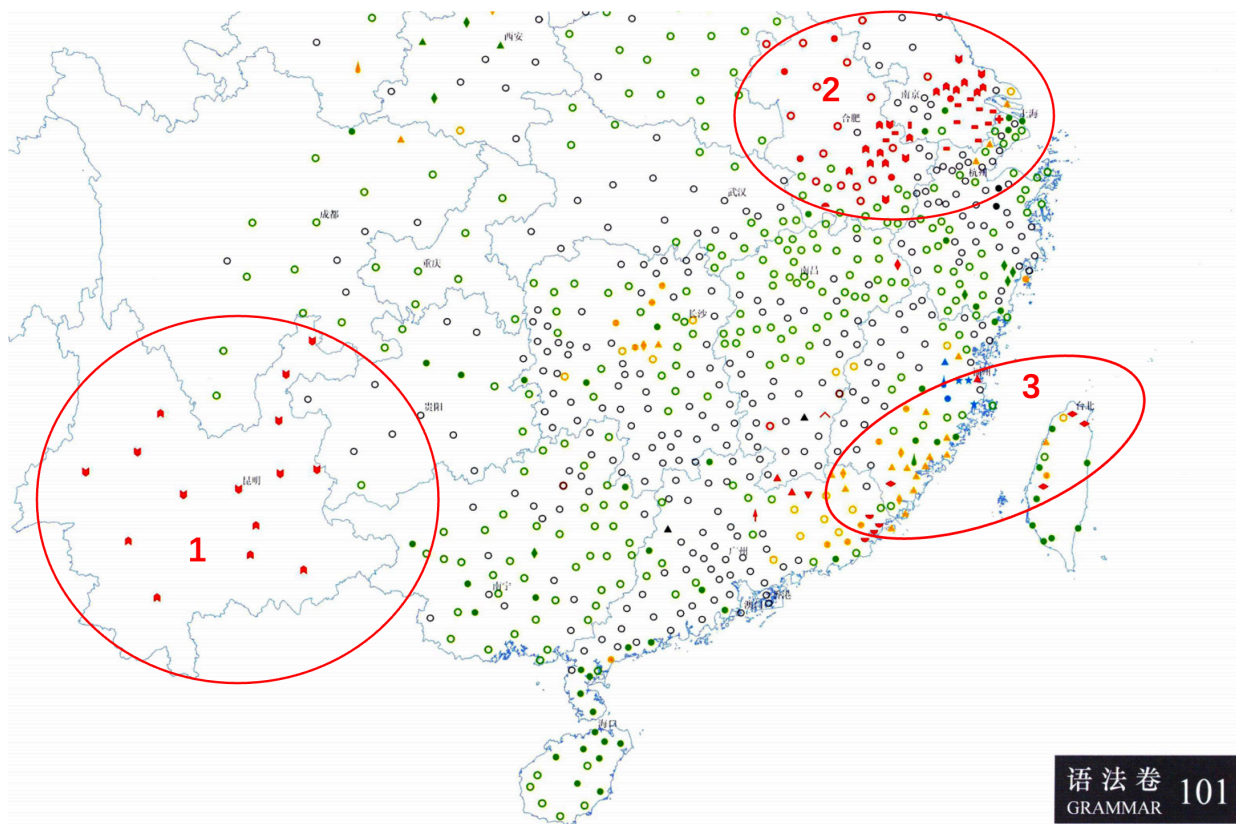


Figure 6.

The particle is pronounced as $kəʔ^5$, $k^həʔ^5$, $xəʔ^5$ in some Jiānghuái Mandarin dialects³³, and as $k^h a ʔ^2$ (Huang 2025), *kam* (Yue-Hashimoto 1991) in some Min dialects. Although their etymologies are not necessarily the same, initials of the particle are either velar stops or velar fricatives. Based on phonological similarity, yes-no questions formed by preverbal Q-particles are normally referred to as “K-VP” questions in the literature. In what follows, I mainly examine data from Jiānghuái Mandarin, but K-VP questions in other dialects will be compared to Jiānghuái Mandarin where appropriate.

³³ The Jiānghuái Mandarin data comes from my native dialect spoken in the Tónglíng 銅陵 city located in the central-south Ānhuī province.

6.2 etymology and a brief history of K-VP questions

According to Zhang (1990:134-150), preverbal Q-particles in modern Chinese dialects can be traced back to the interrogative adverb 曷 *qǐ* (*C.q^hajʔ>khj+jX³⁴) used in pre-Qín periods. In its early uses, *qǐ*-VP forms both neutral and rhetorical questions. Starting from the Eastern Han dynasty (25AD-220AD), the neutral-type K-VP questions were written as 可 *kě*-VP (also see Jiang 1990 for similar conclusions). Zhang (1990) also points out that in the Five Dynasties and Ten Kingdoms period (907AD-979AD), *qǐ* was replaced by 可 *kě* (*kh^hajʔ>khaX), and approximately at the same time, *kě* (*khaX*) developed a consonantal coda which is the origin of glottal stops in its modern reflexes, such as *kəʔ⁵*, *khəʔ⁵*, *xəʔ⁵*, etc. Zhang (1990), however, does not trace preverbal Q-particles in modern dialects to the wh-word *hé* in OC.

Based on the phonological similarity between *hé* (*g^haj) and *kě* (*k^hajʔ) in OC, as well as evidence from the writing system (i.e. 可 *kě* serves as the phonetic element of 何 *hé*), I assume that in OC, the preverbal Q-particle *kě* has the same origin as the wh-word *hé*. Furthermore, the preverbal Q-particle 可 *kəʔ⁵* in Jiānghuái Mandarin should also have the same origin as the OC *hé* ‘wh’ and *kě* ‘Q’. Another reason to trace *kəʔ⁵* to *hé* instead of *qǐ* in OC is that *kəʔ⁵* never scopes higher than subjects (5f), but *qǐ* can scope either lower (1) or higher (2) than subjects (examples taken from Zhang 1990:135):

(1) 子豈識之？

(*Zuǒzhuàn, Chénggōng Èrnián*)

zǐ qǐ shí zhī

you.HON Q know it

‘Do you, sir, know it?’

³⁴ This indicates Middle Chinese transcriptions.

(2) 豈女為之與?

(Mòzǐ, MíngguǐXià)

qǐ rǔ wèi zhī yú
Q you do it SFP
'Was it you who did it?'

This analysis can also be supported by cross-linguistic evidence: in Nepali, the yes-no question particle *ke* is morphologically the same as its *what*-wh-word *ke*:

(3) **ke** tyo tapāĩ-ko ghar ho *Nepali*
Q that 2.SG-GEN house be.SG
'Is that your house?'

(4) (***ke**) yo ke ho?
Q this what be.SG
'What is this?'

(3) and (4) show that the question particle *ke* must occur at sentence-initial positions and it cannot be used in wh-questions. Also in Japanese, *ka* can either mark the indefiniteness of wh-words (e.g. *nani-ka*, 'what-INDEF, whatever') or mark yes-no questions, which provides further evidence that wh-words and Q-particles can be morphologically related.

One of the reasons that drove me to connect *hé* to Q-particles in modern dialects is that due to the fact that we do not have native speakers of OC, some important syntactic diagnostics cannot be applied, but they can be tested in modern dialects.

6.3 basic data of K-VP questions in Jiānghuái Mandarin

This subsection introduces basic descriptive facts of K-VP questions in Tónglíng Jiānghuái Mandarin (JM). Consider first the set of sentences in (5):

- (5) a. η^{213} $kəʔ^5$ $\epsilon i^{213}xu\tilde{\epsilon}^0$ $tɕ^hiəʔ^5$ $jaŋ^{35}zəu^{44}$ $(a^5)?$
 you Q like eat lamb (MP)
 ‘Do you like eating lamb?’
- b. η^{213} $\epsilon i^{213}xu\tilde{\epsilon}^0$ $tɕ^hiəʔ^5$ $jaŋ^{35}zəu^{44}$ $maŋ^5/*maŋ^{51}$
 you like eat lamb Q/*MP
 ‘Do you like eating lamb?’
- c. η^{213} $kəʔ^5$ $\epsilon i^{213}xu\tilde{\epsilon}^0$ $tɕ^hiəʔ^5$ $jaŋ^{35}zəu^{44}$ $Amaŋ^{51}/*Bmaŋ^5?$
 you Q like eat lamb MP/*Q
 A: ‘Do you actually like eating lamb or not?’
 *B, intended: ‘Do you like eating lamb?’
- d.* η^{213} $tsə^{213}məʔ^5$ $tsə^{44}$ $jaŋ^{35}zəu^{44}$ $maŋ^5?$
 you how cook lamb Q
 Intended: ‘How do you cook lamb?’
- e.* η^{213} $kəʔ^5$ $tsə^{213}məʔ^5$ $tsə^{44}$ $jaŋ^{35}zəu^{44}?$
 you Q how cook lamb
 Intended: ‘How do you cook lamb?’
- f.* $kəʔ^5$ η^{213} $\epsilon i^{213}xu\tilde{\epsilon}^0$ $tɕ^hiəʔ^5$ $jaŋ^{35}zəu^{44}$
 Q you like eat lamb
 Intended: ‘Do you like eating lamb?’

This Mandarin variety has two approaches to expressing yes-no questions, one of which is using a sentence-final Q-particle $maŋ^5$ (5b) and the other one is to use a preverbal Q-particle $kəʔ^5$ (5a), but the two approaches cannot co-occur (5c). (5b) and (5c) aim to show the difference between the Q-particle $maŋ^5$ and the modal particle $maŋ^{51}$: $maŋ^{51}$ must be used in interrogative (5c) or

imperative sentences; thus, it results in ungrammaticality in the declarative sentence (5b). (5d) and (5e) exhibit that yes-no question particles cannot be used in wh-questions, whether the particle is preverbal or postverbal, which is different from OC because in OC, the question particle *hū* can occur in wh-questions (V. 22).

An important fact revealed by (5f) is that when there is a subject, for instance η^{213} ‘you’, $kəʔ^5$ cannot precede the subject. At the same time, subjects base-generate in VP-internal positions and move to [Spec, TP] (Koopman and Sportiche 1991). This means that $kəʔ^5$ in JM surfaces within the domain of TP. However, the Clausal Typing Hypothesis (Cheng 1991) claims that the interrogative force of a sentence is necessarily realized within the domain of CP, either via a question particle base-generated in C^0 or via movement to that position. Nepali puts its Q-particle *ke* at the left periphery (3) and Japanese puts its Q-particles *ka* and *no* at the right periphery (6)

(6) anata=wa	nani=o	shiteiru	no	<i>Japanese</i>
you=TOP	what=ACC	do.PROG	Q	
‘What are you doing?’				

Even in English, which is a language without overt question particles, it is proposed that there is a covert sentence-initial question morpheme in its direct and indirect questions (Baker 1970). Taken together, both empirical data and theoretical prediction show that $kəʔ^5$ ‘Q’ in JM surfaces at an exceptional syntactic position compared with its counterparts in other languages. Therefore, we are led to a movement analysis to derive K-VP questions in JM.

6.4 comparisons among four kinds of yes-no questions

In (V.16), we have glimpsed some attributes of A-not-A questions, here we examine them in more details. Enlightened by some syntactic diagnostics in Cheng et al. (1996), I apply 10

diagnostics in this section as to compare K-VP questions in JM to three other kinds of yes-no questions in Mandarin—Negative Particle Questions (NPQs, 7), A-not-A questions³⁵ (8), and *ma*-headed yes-no questions (9):

(7) nǐ xǐhuan chī yáng ròu bu³⁶
 you like eat lamb NEG.Q
 ‘Do you like eating lamb?’

(8) nǐ xǐ-bù³⁷-xǐhuan chī yáng ròu
 you li-NEG-like eat lamb
 ‘Do you like eating lamb?’

(9) nǐ xǐhuan chī yáng ròu ma
 you like eat lamb Q
 ‘Do you like eating lamb?’

As instantiated by (7), the morpheme *bu* in Mandarin NPQs has dual status—it is used both as a negative marker and a yes-no question particle synchronically. But originally, *bu* was merely a negative marker in OC, and it developed into a Q- particle in MC (Aldridge 2011).

6.4.1 matrix and embedded readings

(10) a. [CP₁ ta³¹ ɕio²¹³ təʔ⁵ [CP₂ nǐ kəʔ⁵ lɛ³⁵=ləʔ⁰]] JM
 he know you Q come=PERF
 ‘He knows whether you have already come.’
 b. [CP₁ ta⁵³ kəʔ⁵ ɕio²¹³ təʔ⁵ [CP₂ nǐ lɛ³⁵=ləʔ⁰]]?

³⁵ A-not-A questions have several other variants, see Huang (1991) for details.

³⁶ This negative marker is pronounced as *bù* [pu⁵¹] in its citation tone, but in NPQs, it must be realized in neutral tone.

³⁷ This negative marker can either be pronounced as *bù* (i.e. citation tone) or *bu* (i.e. neutral tone).

‘Does he know that you have already come?’

‘Does he know whether you have already come?’

NPQs can have embedded and matrix readings at the same time:

‘He already knows whether you have come.’

‘Has he already known that you came?’

Intended: 'He already knows whether you have come.'

‘Has he already known that you came?’

scope, while *méiyǒu* can have either matrix or embedded scopes. Therefore, NPQs are the same as

K-VPs in that they both can have either embedded or matrix readings, but they differ in whether both readings can be realized at the same time. This difference concerns whether a Q-particle is root morpheme. Root morphemes are morphemes those can only be used in matrix clauses. In this sense, *bu* is a root morpheme thus (12a) is ungrammatical, but *kəʔ⁵* is not. When both embedded and matrix clauses are interrogatives, Mandarin needs to use an A-not-A question in the matrix to express the meaning of (10c), as shown in (13):

- (13) [CP₁ tā zhīdào=bù=zhīdào [CP₂ nǐ lái=lái **méiyǒu**/***bu**]]?
 he know=NEG=know you come=PERF NEG.Q
 ‘Does he know whether you have already come?’

The non-root morpheme *kəʔ⁵* is similar to the Japanese *-ka* because *-ka* can also be used in embedded and matrix clauses simultaneously:

- (14) [kare=wa [anata=ga mō kita **ka** CP₂] shitteimasu **ka** CP₁] *Japanese*
 he=TOP you=NOM already come Q know.polite Q
 ‘Does he know whether you have already come?’

(15) exhibits that *ma*-headed yes-no questions cannot have embedded readings, namely, *-ma* is a root morpheme. And naturally, *-ma* cannot occur in embedded and matrix clauses simultaneously:

- (15) a. [CP₁ tā zhīdào [CP₂ nǐ lái=lái] ma]
 he know you come=PERF Q
 ‘Does he know that you have already come?’

- b. * [CP₁ tā zhīdào [CP₂ nǐ lái=lái ma]]

he know you come=PERF Q
 Intended: ‘He knows whether you have come.’

And (16) tells us that A-not-A questions are the same as K-VPs because they are compatible with all of the three syntactic environments:

(16) a. [CP₁ tā zhī=bù=zhīdào [CP₂ nǐ lái=lái]]
 he kn=NEG=know you come=PERF
 ‘Does he know that you have already come?’

b. [CP₁ tā zhīdào [CP₂ nǐ lái=bù=lái]]
 he knows you come=NEG=come
 ‘He knows whether you will come.’

c. [CP₁ tā zhī=bù=zhīdào [CP₂ nǐ lái=bù=lái]]
 he kn=NEG=know you come=NEG=come
 ‘Does he know whether you will come or not?’

6.4.2 compatibility with *-ne* (or equivalents)

(17)-(20) exhibit the compatibility between four types of yes-no questions and modal particle:

(17) ta³¹ kəʔ⁵ tɕhi⁵⁵ nẽ⁰ JM
 he Q go MP
 ‘Will he go?’

(18) a.* tā qù bu ne
 he go NEG MP
 Intended: ‘Will he go?’

- (23) *tā* **huì** *qù* *bu*
 he will go NEG
 ‘Will he go?’

- | | | | |
|---------------|------------|----|----|
| (24) tā | huì | qù | ma |
| he | will | go | Q |
| ‘Will he go?’ | | | |

6.4.4 compatibility with preverbal adjuncts

Among the four types of yes-no questions, only A-not-As are incompatible with preverbal adjuncts:

- (25) ta⁵³ kəʔ⁵ **tein**³¹ **tsʰaŋ**³⁵ tɕʰi⁴⁴ JM
 he Q often go
 ‘Does he go often?’

- (26)* tā cháng qù=bú=qù?
 he often go=NEG=go
 ‘Does he often go?’

- (27) **tā** **cháng** **qù** **bu?**
 he often go NEG
 ‘Does he often go?’

- (28) **tā** **cháng** **qù** **ma?**
 he often go NEG
 ‘Does he often go?’

6.4.5 compatibility with negative contexts

(29)-(33) show grammaticalities of the four types of questions when they occur in negative sentences:

(29)*	ŋ ²¹³	kəʔ ⁵	pəʔ ⁵	tɕ ^h i ⁵⁵	ɕi ⁵³ ia ²¹³ t ^h u ³⁵	<i>JM</i>
	you	Q	NEG	go	Seattle	

Intended: ‘Will you not go to Seattle?’

(30)?	nǐ	bù	qù	Xīyǎtú	bu
	you	NEG	go	Seattle	NEG.Q

Intended: ‘Will you not go to Seattle?’

(31)*	nǐ	méiyǒu	qù=guo	Xīyǎtú	méiyóu
	you	NEG	go=EXPE	Seattle	NEG.Q

Intended: ‘Haven’t you been to Seattle?’

(32)*	nǐ	bù	qù=bú=qù	Xīyǎtú
	you	NEG	go=NEG=go	Seattle

Intended: ‘Will you not go to Seattle?’

(33)	nǐ	bú	qù	Xīyǎtú	ma
	you	NEG	go	Seattle	Q

‘Will you not go to Seattle?’

Both K-VPs (29) and A-not-As (32) cannot occur in negative contexts, but *ma*-headed questions (33) can. As indicated by the comparison between different types of NPQs (30 vs 31), *bù* is more acceptable than *méiyǒu* in negative contexts.

6.4.6 compatibility with the adverb *dàodǐ* ‘actually’

Adverbs also help in differentiating between different types of yes-no questions. For example, the adverb ‘actually’:

- (34) ɲ²¹³ tɔ⁵⁵ ti²¹³ kəʔ⁵ tɕ^hi⁵⁵ ɕi⁵³ ia²¹³ t^hu³⁵ JM
 you actually Q go Seattle
 ‘Are you actually going to Seattle or not?’

- (35)* nǐ **dàodǐ** qù Xīyǎtú bu
 you actually go Seattle NEG.Q
 Intended: ‘Are you actually going to Seattle or not?’

- (36)* nǐ **dàodǐ** qù Xīyǎtú ma
 you actually go Seattle Q
 Intended: ‘Are you actually going to Seattle or not?’

- (37) nǐ **dàodǐ** qù=bú=qù Xīyǎtú
 you actually go=NEG=go Seattle
 ‘Are you actually going to Seattle or not?’

(34)-(37) exhibit that K-VPs and A-not-As can cooccur with the adverb ‘actually’ while the other two types of questions cannot do so. However, this pattern is expected because in its earliest uses, K-VPs were accompanied with rhetorical meanings, which is semantically compatible with the adverb ‘actually’.

6.4.7 Intervention Effects

The intervener ‘only’ can be used to diagnose whether LF movements happen:

- (38)* ɲ²¹³ kəʔ⁵ **kuaŋ**⁵³ ɕi²¹³xuẽ tɕ^hiəʔ⁵ jaŋ³⁵ zəu⁵⁵ JM

you Q only like eat lamb

Intended: ‘Do you only like eating lamb?’

(39)* nǐ **zhǐ** xǐhuan chī yáng ròu bu
 you only like eat lamb NEG.Q

Intended: ‘Do you only like eating lamb?’

(40) nǐ **zhǐ** xǐhuan chī yáng ròu ma
 you only like eat lamb Q

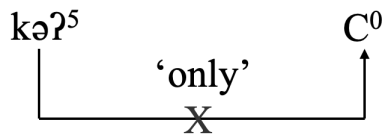
Intended: ‘Do you only like eating lamb?’

(41)* nǐ zhǐ xǐ=bu=xǐhuan chī yáng ròu
 you only li=NEG=like eat lamb

Intended: ‘Do you only like eating lamb?’

(38)-(41) show that only *ma*-headed yes-no questions are insensitive to IEs. The existence of IE in (38) provides strong evidence to show that the Q-particle $kəʔ^5$ moves rightward in LF but it is blocked by the intervener ‘only’. The IE in (38) is shown in (42):

(42)



6.4.8 embedded null subject question

This subsection tests whether embedded interrogatives allow null subjects:

(43) o²¹³ pəʔ⁵ ɛio²¹³ təʔ⁵ [CP *pro* kəʔ⁵ io⁵⁵ tɕ^hi⁵⁵ ɛi⁵³ ia²¹³ t^hu³⁵] JM
 I NEG know Q need go Seattle

‘I do not know whether (I) need to go to Seattle.’

(44) wǒ bù zhīdào [_{CP} *pro* yào=bú=yào qù Xīyǎtú]
 I NEG know need=NEG=need go Seattle

‘I do not know whether (I) need to go to Seattle.’

(45)* wǒ bù zhīdào [_{CP} *pro* yào qù Xīyǎtú bu]
 I NEG know need go Seattle NEG.Q

Intended: ‘I do not know whether (I) need to go to Seattle.’

(46) wǒ bù zhīdào [_{CP} *pro* yào qù Xīyǎtú ma]
 I NEG know need go Seattle Q

‘Do not I know that (I) need to go to Seattle?’

#‘I do not know whether (I) need to go to Seattle.’

(43)-(46) again show that K-VPs and A-not-As pattern the same as both of them allow embedded null subjects, as indicated by the *pro*. On the other hand, NPQs and *ma*-headed questions do not allow embedded null subjects.

The diagnostics employed through 6.4.1 to 6.4.8 are summarized in the following table:

Table2. Comparison between four types of yes-no questions

Diagnostics	K-VP	NPQ	A-not-A	<i>ma</i> -yes-no
matrix reading	✓	✓	✓	✓
embedded reading	✓	✓	✓	✗
cooccur with <i>-ma</i> (or equivalent)	✗	✗	✗	✗
cooccur with <i>-ne</i> (or equivalent)	✓	✗	✓	✗
cooccur with modals	✓	✓	✗	✓
cooccur with preverbal adjuncts	✓	✓	✗	✓
occur in negative sentences	✗	✗	✗	✓

cooccur with ‘actually’	✓	✗	✓	✗
Intervention Effect	✓	✓	✓	✗
allow embedded null subjects	✓	✗	✓	✗
occur in matrix and embedded clauses simultaneously	✓	✓	✓	✗

Table 2 indicates that among the 11 diagnostics, A-not-As share 9 of the syntactic restrictions with K-VPs, followed by NPQs which share 8 of them with K-VPs, while *ma*-headed yes-no questions only share 4 of them with K-VPs.

6.5 the Neg⁰-to-C⁰ movement of *kəʔ⁵* ‘Q’

6.5.1 theoretical hypothesis

I have shown in 6.4 that K-VPs are syntactically the most closely related to A-not-As, while “strongly speaking, a-VP questions (i.e. K-VPs) are A-not-A questions”(Gasde 2004:310). Huang (1991:324-325) also points out that K-VPs in Taiwanese are “counterparts of the A-not-A questions”. More specifically, he treats the [A not AB] type as “a variant of the *kam*-question”. Huang (1991) and Huang et al. (2009:254) propose an intermediate Q to derive A-not-A questions, and this Q occupies the same position as negation, namely, after the subject and before the verb, as indicated by Figure 6 (Huang et al. 2009:254). Interestingly, Figure 7 shows that this position also holds *kəʔ⁵* ‘Q’:

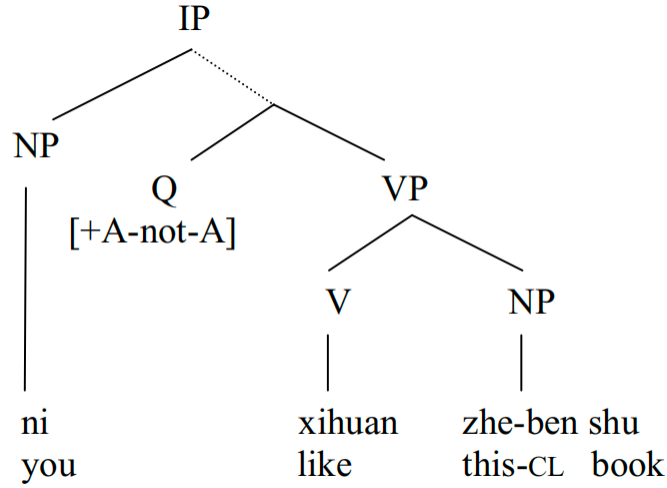


Figure 7.

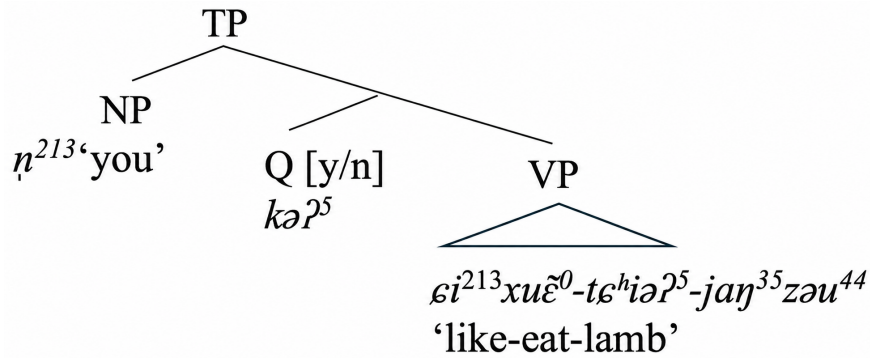


Figure 8.

The authors claim that A-not-A questions need to undergo LF movement to the domain of C “for the assignment of its scope in Logical Form (LF)” (idem) which is exactly the same proposal as the one I derived in 6.3. Based on similar syntactic restrictions between K-VPs and NPQs, and the Neg⁰-to-C⁰ movement in Mandarin NPQs (Cheng et al. 1996), it’s reasonable to hypothesize that the Q-particle *ka*²⁵ also undergoes Neg⁰-to-C⁰ movement. To make this hypothesis more convincing, 6.5.2 clarifies why *ka*²⁵ moves rightwards; then, 6.5.3 proves the correlation between negation and interrogative, namely, why *ka*²⁵ occupies a negation head. Data from both OC and modern Chinese dialects will be employed in the next two subsections.

6.5.2 rightward movement of *kəʔ*⁵

The OC Q-particle *hū* canonically occurs in sentence-final positions, but there are some exceptions:

(47) 禮之本末，將於此乎在？ (Zuǒzhuàn, Shàngōng Wǔnián)

lǐ	zhī	běn	mò,	jiāng	<i>i</i> yú	cǐ	hū	zài <i>i</i> ?
ritual	GEN	root	branch	MOD	at	here	Q	be

‘Will the root and branches of ritual lie herein?’

(48) 吾乃今於是乎見龍？ (Zhuāngzǐ, Tiānyùn)

wú	nǎi	jīn	yú	shì	<i>i</i>	hū	[jiàn	lóng] <i>i</i> ?
I	then	nowadays	at	here		Q	behold	dragon

‘Is it nowadays that I behold the dragon here?’

(49) 生乎由是？死乎由是？ (Xúnzǐ, Quànxué)

<i>i</i>	sheng	hū	[yóu	shì] <i>i</i> ?	<i>i</i>	sǐ	hū	[yóu	shì] <i>i</i> ?
	be.born	Q	from	here?		die	Q	from	here

‘(Am I) born from here? (Am I) die from here?’

In the above three examples, what Q separates apart forms a constituent, namely, *zài-yú-cǐ* ‘be-at-here’, *yú-shì-jiàn-lóng* ‘at-here-see-dragon’ and *yóu-shì-shēng* ‘from-here-be.born’ are constituents, respectively. Logically, there are two ways to derive the linear orders in (47)-(49). One is to say that the elements to the right of Q undergo right-dislocation (RD), while the other is to say that it is *hū* ‘Q’ that moves. Within the two approaches, I choose the latter one because what needs to be right-dislocated in (47) is a single copular *zài*, while in (48) and (49) are VP and PP, respectively. Therefore, the Q-movement analysis can be more unified in theory. Within the Q-

movement analysis, there also exist two possibilities. One solution is that Q scrambles leftwards, and the other one is that Q moves rightward in LF.

Scrambling has been well discussed in Japanese, and this phenomenon is more often reported in SOV languages. Tsujimura (2014:242) defines this phenomenon as “the type of sentence in which some constituents are not in the basic order of SOV is called a scrambled sentence”. The fact that OC is not an SOV language makes the scrambling analysis suspicious from a theoretical perspective. At the same time, the author defines RD as “a movement of some element(s) to the right of the verb...notably observed in colloquial speech” (idem:242). Without considering word order typology, the scrambling analysis is reasonable to some extent in that (47)-(49) do not violate any principles that are strictly prohibited in scrambled sentences. Tsujimura (2014:245) gives examples of some constituents that cannot be separated. For instance, in *takai-kutu* ‘expensive-shoe’, *kodomo-no-sensei* ‘child-GEN-teacher’, the adjective and genitive noun cannot be separated from their head nouns. Also, coordinated phrases and a constituent that belongs to an embedded relative clause cannot be separated. Due to these characteristics, scrambling is said to be “clause bound”(idem:247), but RD does not have this clausal restriction. Agbayani and Ishii (2024:221) mention that “single Right Dislocation of a ‘true adjunct’ is disallowed”, in which what they call “true adjunct” is something like *riyû-mo-naku* ‘without any reason’, thus right-dislocation of *yôu-shi* ‘from-here’ in (49) should be prohibited. Still, they point out that RD only applies to a non-predicate XP (idem:223), therefore the predicate *zài* ‘be’ in (47) is not eligible for RD. Based on these arguments, it is more reasonable to analyze (47)-(49) as scrambling rather than RD if Q does not move.

Now we turn to the solution of Q-movement. Unlike RD, which can happen due to pragmatic or prosodic reasons (Agbayani and Ishii 2024), scrambling needs syntactic motivations, as Tanaka

(2001) claims that “there is no direct syntactic movement of the right-dislocated phrase”. In (47) and (48), *hū* ‘Q’ linearly follows PPs and in (49) it follows a verb which indicates that there are no functional projections that Q can be attracted to. Therefore, *hū* checks no features in (47)-(49) thus it’s not possible for it to undergo leftward scrambling that requires a syntactic motivation. Then, there remains the last logical possibility, namely, *hū* undergoes rightward movement in LF to check the [Q] feature in sentence-final positions.

Besides evidence from OC, some Chinese dialects also support the rightward movement of *kəʔ⁵*. K-VP questions in JM are incompatible with sentence-final Q-particles, as shown in (50):

- (50)* *ta⁵³* *kəʔ⁵* *tɕhi⁵⁵* *maŋ⁰*
 he Q go Q
 Intended: ‘Will he go?’

But in some dialects (e.g. Yáchéng Jūnhuà 崖城軍話, Liu 2018) and languages (e.g. Lùxī Achang), preverbal Q-particles can co-occur, or even must co-occur with sentence-final particles:

- (51) *ni²¹* *ʔa⁵⁵* *kʰu⁵⁵* *pɕi³²⁴ kein³³* **(miu⁵⁵)* *Yáchéng Jūnhuà*
 you Q go Beijing NEG.Q
 ‘Will you go to Beijing?’

- (52) *a²⁵* *tshu⁵⁵* *kəʔ⁵* *pi⁵¹=tau³³* *mɔ⁵¹* *Lùxī Achang*
 two year Q give=PERF Q
 ‘Has (someone) given (you) two years’ (rent)?’

In Yáchéng Jūnhuà, *miu*⁵⁵ ‘NEG.Q’ is obligatorily used in K-VP questions, and in Lùxī Achang, *mɔ*⁵¹ ‘Q’ can still be used when there is a preverbal particle³⁸. Similarly, K-VP questions in OC also frequently contain sentence-final particles (see examples in Zhang 1990:136 and Jiang 1990).

The cooccurrence of two Q-particles in JM, OC, Yáchéng Jūnhuà and Lùxī Achang can be explained from the perspective of Copy Theory of Movement (CTM). CTM aims to explain the “displacement property” of human languages under the Minimalist Program. This theory states that “a trace is copy of the moved element that is deleted in the phonological component” (Corver and Nunes 2007:2). According to the standard definition of CTM, a lower copy gets erased in PF and only the highest copy shall be pronounced. If sentence-final Q-particles in these languages project CPs, namely, being higher copies, preverbal particles can be regarded as lower copies of moved Q-particles.

6.5.3 correlating negation and interrogative

Although K-VP questions in JM, including most of the other dialects, cannot be used in negative sentences, a few dialects do not have this restriction. Cao *eds.* (2008, picture 101) reports the existence of K-Neg-VP construction in Fúzhōu 福州 dialect, Zànhuáng 贊皇 dialect (Southwest of Héběi province), Dìngnán 定南 dialect (far South of Jiāngxī province), and Hépíng 和平 dialect (Northeast of Guǎngdōng province). These dialects imply the existence of an uninterpretable [NEG] feature within the Q-particle *kəʔ*⁵.

³⁸ (52) was collected in a natural conversation. However, in elicitation sessions, my consultants judged sentences containing two Q-particles to be unacceptable. The preverbal Q-particle *kəʔ*⁵ in Lùxī Achang and JM are accidentally homophonous, and they shall not be cognates.

Based on phonological evidence, Yue-Hashimoto (1991:187) believes that preverbal Q-particles *kam* and *kʰaʔ* in Southern Min dialects “suggest a fusion of 𠂔³⁹+NEG”, and NEG is attracted to preverbal Q-particles⁴⁰. Yue-Hashimoto (1991) uses the term “attract” in a non-generative sense, but it gives us an inspiration that there is some syntactic movement happening in K-VP questions.

The correlation between negation and interrogative also comes from diachronic changes of K-VP questions in the history of Chinese. I have alluded to this in 6.2 that according to Zhang (1990), the neutral 𠂔 *kě*-VP questions (i.e. non-rhetorical uses) appeared after the formation of rhetorical K-VP questions. Jiang (1990) also finds that the earliest uses of 𠂔 *kě*-VP questions in the Eastern Hàn dynasty trigger rhetorical rather than neutral questions. Huang (2025) finds that during the late Míng dynasty (ca. 1581CE), K-VP questions normally ask rhetorical questions, while they rarely ask neutral questions.

The above facts show that K-VP questions gradually lost rhetorical uses during their development. Cross-linguistically, rhetorical questions are closely related to negation, as Han (2005) clarifies that “a rhetorical positive question has the illocutionary force of a negative assertion”, and this explains why *kəʔʂ* ‘Q’ occupies the same position as a negation head and bears [*u*NEG] feature. From a diachronic view, the reason [NEG] feature on *kəʔʂ* is uninterpretable is that it has not fully lost the semantic restriction inherited from its early uses. Cheng et al. (1996) assume that negative markers in Cantonese and Taiwanese are in “dual status”, which means they are simultaneously negation markers and Q-particles, and thus can be formalized as bearing [Q, NEG]

³⁹ 𠂔 *kě* is the character used to write *kəʔʂ* ‘Q’ in JM.

⁴⁰ Yue-Hashimoto’s (1991) idea of phonological fusion is not uncontroversial. For instance, Huang (2025) believes that the *-m* ending comes not from the negative marker, rather, *kam* has its etymology as 敢 *gǎn* (*kʰamʔ>*kamX*)

features. Adopting their way of feature representation, the Q-particle $kəʔ^5$ should bear the features $[iQ, uNEG]$, where $[Q]$ is an interpretable feature.

In sum, Chapter VI demonstrates the plausibility of a rightward movement analysis of the OC wh-word *hé*, providing an alternative solution to the canonical leftward movement analysis derived in chapters IV and V. This demonstration relies heavily on modern Chinese dialects, and it rests on the premise that OC native speakers and modern Chinese native speakers share the same UG, which is referred to as the Uniformitarian Hypothesis of UG (see footnote 8 in 1.4.2). Otherwise, the JM Q-particle $kəʔ^5$ would be unable to support the validity of the rightward movement of the OC *hé*.

Now I have demonstrated that the OC wh-word *hé* ($*g^saj$) can reasonably move rightwards, but unlike the leftward movement, in which the question feature associated with the initial $*g^s$ moves to check the Q feature, it is the feature associated with the final $*-aj$ that moves to check the Q feature when *hé* moves rightwards. This means that depending on the directionality of movement, features associated with different feature-bearing units move in LF. In light of the distinction between the two wh-words *hé* and *hú* mentioned in 2.3, we can better understand why it is the feature associated with $*-aj$ that moves when *hé* undergoes rightward movement in LF—*hé* (g^saj) has the $*-j$ ending so it can occur sentence-finally while *hú* (g^sa) does not have the $*-j$ ending so it never occurs sentence-finally. A theoretical problem of the Neg^0 -to- C^0 movement needs to be clarified: unlike the view of Cheng et al. (1996), Cheng and Sybesma (2026:400) believe that “ C^0 may be occupied by a sentence final particle, but no movement takes place into”. They also point out that “when movement is not possible, the relevant position is targeted by movement later (at LF) or by insertion” (idem), while the movement into C^0 proposed for *hé* is indeed LF movement.

Chapter VII Conclusion and Further Remarks

This chapter first recapitulates main findings of the present study, then provides further clarification on some contentious issues.

7.1 Conclusion

This study is motivated by differences in the linear order of wh-adverbials between OC and Mandarin. It focuses on a rarely discussed OC construction “VP-*hé*” in which the wh-word *hé* can be interpreted either as a wh-nominal ‘what’ or as a wh-adverbial ‘how’ or ‘why’. Based on evidence from OC, Japanese and English, I propose the existence of a covert preposition before *hé* in order to reach a unified analysis of the syntactic structure of “VP-*hé*”. However, the central concern of this thesis is the wh-adverbial *hé*. The wh-adverbial *hé* can be categorized into two types: *how*-wh-*hé* and *why*-wh-*hé* and they differ hierarchically in LF. Although both types of *hé* can surface at either periphery, the low wh-adjunct *hé* ‘how’ tends to surface at the right periphery while the high wh-adjunct *hé* ‘why’ tends to surface at the left periphery.

Adopting the adjunct analysis, I summarize conclusions to the question posed in the title of this thesis, namely, whether the OC wh-adjunct *hé* is in-situ or ex-situ: when *hé* ‘how come’ co-occurs with the modal *qí* at the left periphery (i.e. *hé-qí*-VP), it base-generates there and does not move thus is an in-situ wh-word; when it means ‘how’, *hé* base-generates at VP-internal positions and remains in-situ (i.e. VP-*hé*_{how}); when it means ‘why’, both in-situ and ex-situ solutions are possible, depending on which periphery *hé* surfaces at. When surfacing at the right periphery (i.e. VP-*hé*_{why}), *hé* can be analyzed as base-generated there, namely, being in-situ, and the feature associated with the initial *g^s- moves leftwards to [Spec, CP] to check the [wh] feature; but the

why-wh-hé can also be analyzed as an ex-situ wh-word with the feature associated with the final *-aj moving rightwards if it base-generates at the left periphery (i.e. *i*-VP-*hé_i*). When surfacing at the left periphery (i.e. *hé_{why}*-VP), *hé* does not undergo movement thus is an in-situ wh-word. The above conclusions are further summarized in the following table:

Table 3. Summaries of the major conclusions

Interpretation of <i>hé</i>	Surfacing position	In-situ or Ex-situ	Notes
‘how.come’	left periphery	In-situ	co-occurs with the modal <i>qí</i>
‘how’	right periphery	In-situ	canonical position
‘why’	right periphery	In-situ	rarely attested
		or Ex-situ	rightward movement
	left periphery	In-situ	canonical position

From my point of view, the in-situ analysis of the *why-wh-hé* is preferred to the ex-situ analysis. First of all, the in-situ solution provides a unified analysis of the *why-wh-hé*, regardless of which periphery it surfaces at. More importantly, evidence for the ex-situ solution comes primarily from yes-no question particles in modern dialects instead of OC wh-words.

Next, I provide tree diagrams for sentences extracted from (II.3), (II.4) and (III.3), repeated here as (1), (2) and (3) respectively.

(1) 其為舟車何？ (Mòzǐ, Jiéyòng)

qí wéi zhōu chē hé?
 it make boat carriage why
 ‘Why did (they) make boats and carriages?’

(2) 何許子之不憚煩？ (Mencius, Téngwéngōng Shàng)

hé Xǔ zǐ zhī bù dān fán

why Xu master GEN NEG troubled labor
 ‘Why is Master Xu not troubled by such labor?’

(3) 如宋王何？

(*Mencius, Téngwéngōng Xià*)

rú sòng wáng **hé**
 do Song king how
 ‘How to deal with the King of Song?’

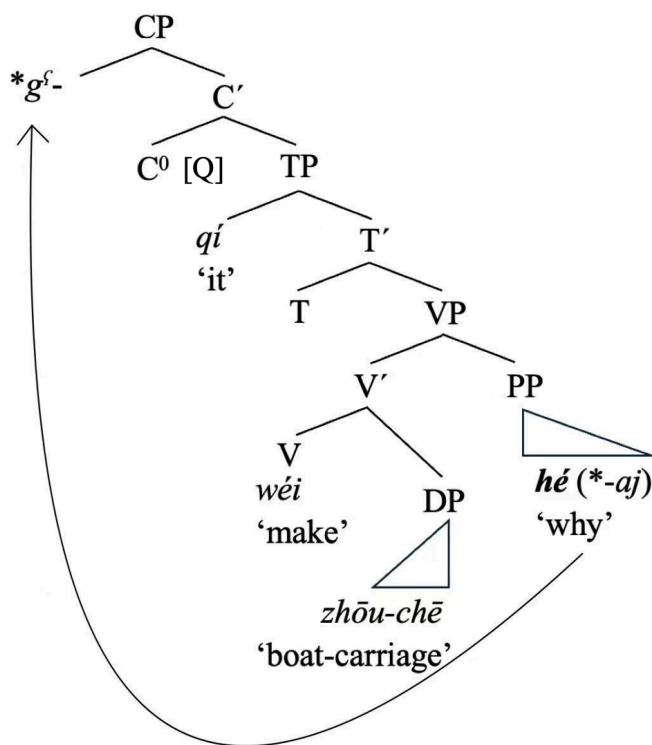


Figure 9. (tree for 1, leftward movement, in-situ)

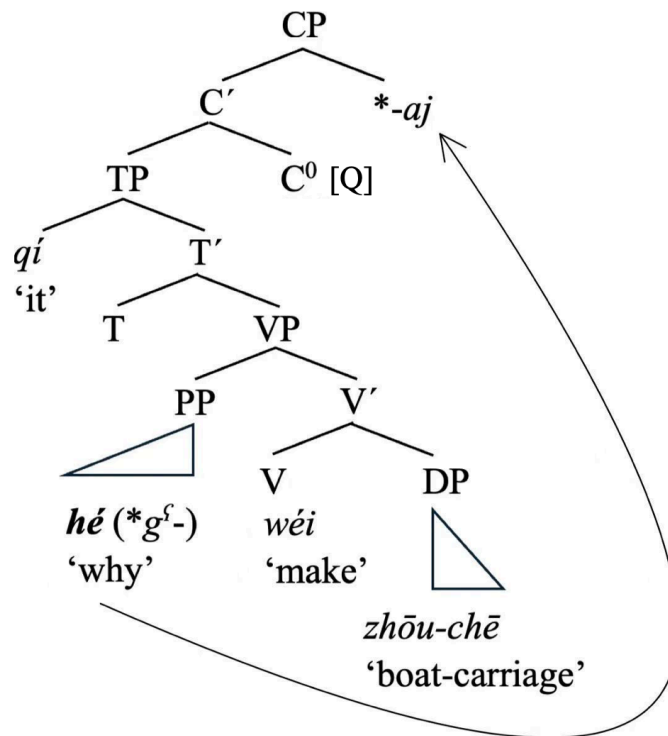


Figure 10. (tree for 1, rightward movement, ex-situ)

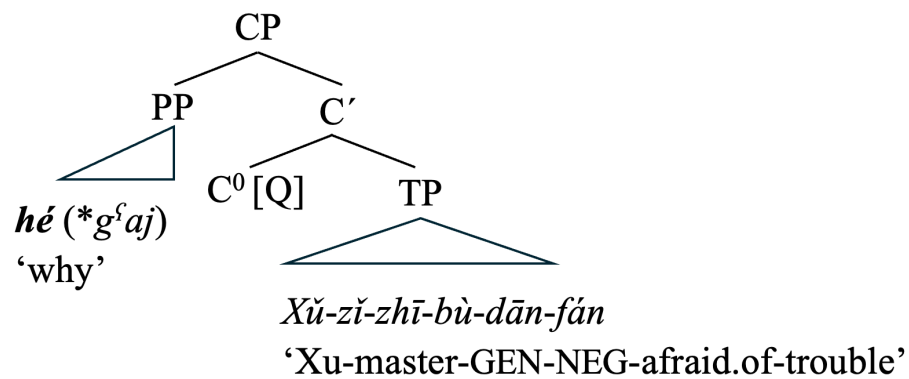


Figure 11. (tree for 2, no movement)

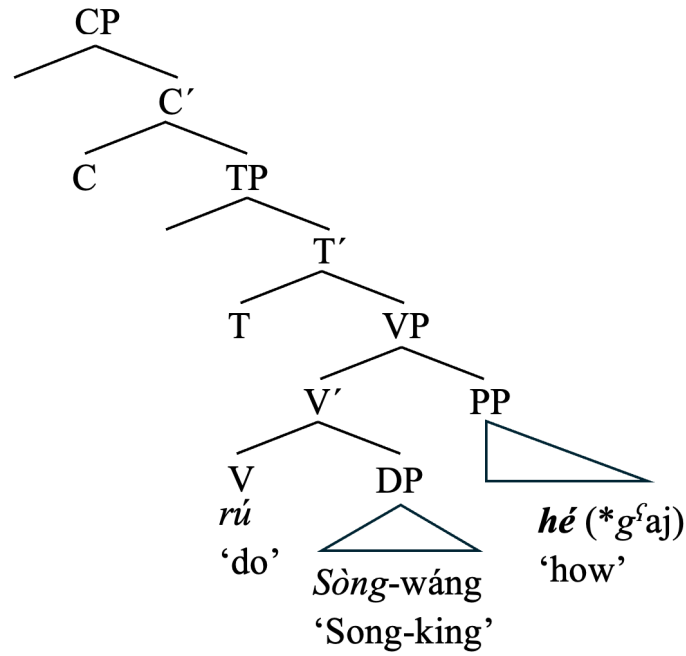


Figure 12. (tree for 3, no movement)

When the *why-wh-hé* surfaces at the right periphery, both in-situ and ex-situ approaches are possible, as indicated by the parenthetical descriptions in the figure captions. In contrast, when the *why-wh-hé* surfaces at the left periphery, no movement is involved since *hé* base-generates here. As in Figure 11, there is no movement in Figure 12 either when *hé* base-generates at the right periphery as a low wh-adjunct ‘how’.

7.2 clarification on some contentious issues

This section briefly clarifies three contentious issues that can fundamentally affect the plausibility of this study from a more general perspective.

7.2.1 decomposition of *hé* (*g'aj)

The decomposition of *hé* is also evidenced by syntactic theories in general. Pesetsky (2000) points out that wh-expressions can be decomposed into a question feature plus a nominal expression. Li (2024:14) cites Pesetsky (2000) to support the idea that the Mandarin *shénme* ‘what’ contains a question feature Q and a nominal expression. The decomposition of *shénme* can be made clearer from a diachronic view. In Early Modern Chinese (ca. 8 AD), *shénme* ‘what’ was written as 是物 (*dzyeX mjut*) or 底物 (*tejX mjut*), and the latter one literally translates as ‘what-thing’. And the morpheme *hé* was also once combined with *mjut* ‘thing’ in MC, forming a disyllabic word 何物 (*ha mjut*) ‘what-thing’. Therefore, we can say that the difference between ‘what’ in OC and that in (E)MC lies in that in OC, ‘what’ inherently contains the sememe ‘thing’, but in later periods, this sememe became overt.

7.2.2 rightward movement

The RM analysis derived in Chapter VI may be the most vulnerable point of the present study because syntactic movements in natural languages, especially A' movements that target specifier positions, canonically move leftwards⁴¹. Rightward operations were put forward early in Chomsky (1957 [2002]: X), referred to as “affix lowering” (e.g. English *-ed* and *-s*). Such movements were abandoned in later literature for empirical and theoretical reasons. To name one of them, the lowering analysis of the plural suffix *-s* cannot be applied to nouns that have irregular plural forms (e.g. *sheep*, *deer*). Huang (2015) believes that aspect markers *-zhe*, *-le*, *-guo* in Mandarin also undergo affix lowering. It needs to be clarified that although affix lowering is a kind of rightward operation, it’s more a morphological operation than a phrasal syntactic movement. Importantly, affix lowering is an operation towards hierarchically lower positions, but the rightward movement

⁴¹ Rightward wh-movements are reported to exist in ASL (Neidle et al. 1997).

proposed in Chapter VI targets at hierarchically higher positions, following the analysis of Tang (1989)⁴² which claims that sentence-final question particles in Chinese languages are in the C⁰ positions. Typical rightward movements include Heavy NP Shift and Right Node Raising (e.g. *John likes and Bill dislikes TV shows about vampires*, Citko 2011:68-70). Since this study only concerns rightward-moved elements that are phonologically weak, those moved rightwards but are phonologically heavy will not be mentioned hereafter.

Worse still, the RM analysis seems to constitute a violation of the Antisymmetric Theory because the Linear Correspondence Axiom (LCA) states that “adjunctions must always be to the left, never to the right” (Kayne 1994: XIII), while under my analysis, the adjunction of *hé* can be to the right. However, Takano (2003) proposes that LCA is not absolute—although languages overwhelmingly exhibit antisymmetric attributes, they still allow the existence of symmetric features. Following the distinction made by Chomsky (2004) on Pair-Merge (forming adjunction structures) and Set-Merge (forming complements and specifiers), Takano (2003:524) claims that “since Pair-Merge is exempt from antisymmetry, adjuncts can surface in either direction”. Therefore, even adopting the Antisymmetric Theory, the RM analysis of the wh-adjunct *hé* has its plausibility.

Although abandoned, the earliest view of rightward movement echoes some more recent studies because they reveal that RM has phonological requirements—rightward-moving elements are either phonologically light or heavy. For instance, the affixes mentioned in Chomsky (1957 [2002]: X) are phonologically light; Bennett et al. (2013) summarize the syntactic behavior of the clitic pronoun in Irish as “lightest to the right”. Still, Feng (1996) motivates the fronting of wh-nominal *hé* ‘what’ to prosody, namely, when it cliticizes to verbs, it becomes phonologically weak.

⁴² Tang, T.-C. 1989. *Studies on Chinese Morphology and Syntax 2*. Taipei: Student Book Corporation, cited in Cheng et al. (1996).

The proposals made in these different languages conform to my analysis on the rightward-moving element within *hé*, because **-aj* is also phonologically weak as compared to the whole syllable **g^haj*.

The phonological weakness of **-aj* also conforms with the in-situ and rightward movement analysis of *why-wh-hé* when it surfaces at left peripheries. Hornstein et al. (2005:305) define the morphological property of strong features as being “can only be checked by lexical items or projections of lexical items, but not by sets of formal features that by themselves do not constitute a lexical item”. Since a single final **-aj* does not constitute a lexical item, *hé* does not bear strong features thus can be licensed in-situ. Similar ideas are also found in Huang (2015) which claims that if features are strong, movements happen in overt syntax; and if features are weak, movements happen in LF.

Several issues remain unresolved. In some cases, arguments about OC have to rely primarily on external evidence from other languages, since relevant OC-internal data are limited. For instance, one of the most important diagnostics to distinguish *why-wh*-words and *how-wh*-words is based on their different sensitivity to island effects, but due to the lack of OC sentences where the *wh*-adverbial *hé* is involved in islands, I have to rely on intervention effects (i.e. linear word orders) to tell the difference between *why-wh-hé* and *how-wh-hé*. The argument developed through chapters IV to VI focuses on the syntax of *why-wh-hé* while the question of how the in-situ *how-wh-hé* is licensed is left untouched. Since this study mainly focuses on explaining the phenomena in OC, it does not answer the question of how to parameterize differences between OC and Mandarin *wh*-adverbials. Adopting the Overt Parametrization Hypothesis (Tang 2003:37) which claims that only semantic features are universal, while phonetic features are allowed to vary across languages, phonology parameterizes these differences and this is in line with the decomposition of

hé. Similar to Aldridge (2010)’s judgement of the OC wh-nominal *hé* ‘what’ which is claimed to violate the CTH, the wh-adverbial *hé* may also violate the CTH because it moves due to syntactic motivations and at the same time, OC employs question particles.

Generally speaking, the problems raised in 7.2.1 and 7.2.2 do not affect the overall validity of this thesis, especially because RM is not only reported in non-Chinese languages but also in Chinese. For instance, MC question particles undergo the Neg-to-Q movement (Aldridge 2011). The morphological decomposition of *hé* has its philological basis, but the proposed correlation between syntax and this decomposition requires further empirical support.

REFERENCES

- Ackema, Peter. 2015. Arguments and Adjuncts. In T. Kiss, & A. Alexiadou eds., *Syntax-Theory and Analysis: An International Handbook* (Vol. 1). Berlin: De Gruyter Mouton, pp. 246-273.
- Adger, David. 2003. *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.
- Agbayani, Brian and Ishii, Toru. 2024. A prosodic approach to multiple right dislocation in Japanese. *Journal of Japanese and Korean Linguistics*, Vol.31, pp. 219-229.
- Aldridge, Edith. 2010. Clause-internal wh-movement in Archaic Chinese. *Journal of East Asian Linguistics*, Vol.19, No.1, pp. 1-36.
- Aldridge, Edith. 2011. Neg-to-Q: The historical origin and development of question particles in Chinese, *The Linguistic Review*, Vol.28, pp. 411-447.
- Aldridge, Edith. 2012. PPs and Applicatives in Late Archaic Chinese. *Studies in Chinese Linguistics*, Vol.33, No.3, pp. 139-164.
- Aldridge, Edith. 2023. The Loss of Object Focus and Wh-movement in Early Middle Chinese. *Journal of Historical Syntax*, Vol.7, pp. 1-40.
- Aoun, Joseph. 1986. *Generalized Binding*. Foris Publications, Dordrecht.
- Aoun Joseph, and Li Audrey Yen-hui. 1993. Wh-Elements in Situ: Syntax or LF. *Linguistic Inquiry*, Vol. 24, No. 2, pp. 199-238.
- Baker, C.L. 1970. Notes on the description of English questions: the role of an abstract question. *Foundations of Language*, Vol 6, No 2, pp. 197-219.
- Baxter, William and Sagart, Laurent. 1998. Word Formation in Old Chinese. In Jerome L. Packard ed. *New Approaches to Chinese Word Formation*. Berlin: Mouton de Gruyter, pp. 35-77.
- Baxter, William H. and Sagart, Laurent. 2014. *Old Chinese: A New Reconstruction*. New York: Oxford University Press.
- Bayer, J. 1984. COMP in Bavarian syntax. *The Linguistic Review*, Vol.3, pp. 209-274.

- Bennett, Ryan, Elfner, Emily, and McCloskey James. 2013. Lightest to the right: An anomalous displacement in Irish. *Linguistic Inquiry*, Vol.47, No.2, pp. 169-234.
- Bobaljik, Jonathan David. 2017. Distributed Morphology. In *Oxford Research Encyclopedia of Linguistics*. (<https://doi.org/10.1093/acrefore/9780199384655.013.131>)
- Cao, Zhiyun ed. 2008. *Hànyǔ Fāngyán Dìtǔjī (Yǔfǎ Juàn)* [Linguistic Atlas of Chinese Dialects (Grammar)], Beijing: The Commercial Press.
- Carnie, Andrew. 2021. *Syntax: A Generative Introduction* (Fourth Edition). Wiley-Blackwell.
- Chen, Nianfu. 2006. Jiǎgǔwén 何 hé, 此 cǐ Yòngwéi Dàicí Kǎolùn [A Study of the Pronominal Uses of hé and cǐ in Oracle Bone Inscriptions]. *Zhōngguó Yǔwén* [The Chinese Language], Vol.5.
- Cheng, Lai-Shen Lisa. 1991. On the Typology of Wh-Questions. PhD Dissertation, MIT.
- Cheng, Lai-Shen Lisa. 2000. Moving just the feature. In U. Lutz, G. Müller, & A. von Stechow eds., *Linguistik Aktuell/Linguistics Today*. Amsterdam: John Benjamins. pp. 77-99.
- Cheng, Lai-Shen Lisa. 2009. Wh-in-situ, from the 1980s to Now. *Language and Linguistics Compass* 3/3, pp. 767-791.
- Cheng, Lai-Shen Lisa. 2021. What-as-why sentences in Cantonese. In G. Soare ed., *Studies in Generative Grammar*. Berlin: De Gruyter Mouton, pp. 219-246.
- Cheng Lai-Shen Lisa, Huang James C.-T. and Tang Jane C.-C. 1996. Chinese negative particle questions: a dialectal comparison. In James R. Black and Virginia Motapanyane eds., *Microparametric Syntax and Dialectal Variation*, John Benjamins Publishing Company, pp. 41-78.
- Cheng Lai-Shen Lisa and Rooryck Johan 2000. Licensing Wh-in-situ. *Syntax*, Vol.3, No.1, pp. 1-19.
- Cheng Lai-Shen Lisa and Sybesma Rint. 2026. Analyticity as an epiphenomenon. In Andrew Simpson eds. *New Insights into Theoretical Syntax from Asian Languages: Studies in honor of C.-T. James Huang*. John Benjamins Publishing Company, pp. 380-404.
- Nakao Chizuru and Obata Miki. 2009. When ‘What’ Means ‘Why’: On Accusative wh-adjuncts in Japanese. *Proceedings of the 32nd Annual Penn Linguistics Colloquium*, Volume 15.1, pp. 153-161.
- Choe, Sook Hyon. 1994. Syntactic wh-movement in Korean and licensing. In Young-Key Kim-Renaud ed., *Theoretical Issues in Korean Linguistics*. Stanford Linguistics Society, pp. 275-302.
- Chomsky, Noam. 1957 [2002]. *Syntactic Structures*. Berlin & New York: Mouton de Gruyter.
- Chomsky, Noam. 1965. *Aspects of the Theory of Syntax*. MA: MIT Press.
- Chomsky, Noam. 2004. Beyond explanatory adequacy. In *Structures and beyond: The cartography of syntactic structures*, Vol. 3, A. Belletti ed. Oxford: Oxford University Press, pp. 104-131.
- Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry*, Vol. 36, pp. 1-22.
- Citko, Barbara. 2011. *Symmetry in Syntax: Merge, Move and Labels*. Cambridge: Cambridge University Press.
- Closs, Elizabeth. 1965. Diachronic Syntax and Generative Grammar. *Language*, Vol. 41, No. 3, pp. 402-415.
- Corver, Norbert and Nunes Jairo. 2007. From trace theory to copy theory. In Corver, Norbert and Nunes Jairo eds. *The Copy Theory of Movement*. Amsterdam/Philadelphia: John Benjamins Publishing Company, pp. 1-9.

- Dong, Xiufang. 2009. Hànyǔ Zhōng Wèn Yuányīn de Yíwèncí Jùfǎ Wèizhi de Lìshǐ Biànhuà [Diachronic Change of the Position of Interrogative Words Asking for Causes]. *Lìshǐ Yǔyánxué Yánjiū* [Studies in Historical Linguistics], Vol.2. Beijing: The Commercial Press, pp. 209-220.
- Feng, Shengli. 1996. Prosodically constrained syntactic changes in early archaic Chinese. *Journal of East Asian Linguistics*, Vol.5, pp. 323-371.
- Fujii Tomohiro, Takita Kensuke, Yang Chung-Yu Barry, and Tsai Dylan Wei-Tien. 2014. Comparative remarks on wh-adverbials in situ in Japanese and Chinese. In Mamoru Saito *ed.*, *Japanese Syntax in Comparative Perspective*. Oxford: Oxford University Press.
- Gasde, Horst-Dieter. 2004. Yes/no questions and A-not-A questions in Chinese revisited. *Linguistics*, Vol.42, No.2, pp. 293-326.
- Georgala Effi, Waltraud Paul and John Whitman. 2008. Expletive and Thematic Applicatives. *Proceedings of the 26th West Coast Conference on Formal Linguistics*. C. B. Chang and H. J. Haynie *eds.*, pp. 181-189.
- Hagstrom, Paul. 1998. Decomposing Questions. PhD Dissertation. MIT.
- Han, Chung-Hye. 2005. Deriving the interpretation of rhetorical questions. *The Proceedings of WCCFL 16*, pp. 1-17.
- Handel, Zev. 2003. A Concise Introduction to Old Chinese Phonology. In Matisoff (2003), pp. 543-574.
- Heine, Bernd and Kuteva, Tania. 2002. *World Lexicon of Grammaticalization*. Cambridge: Cambridge University Press.
- Hong, Bo. 2010. *Hànyǔ Lìshǐ Yǔfǎ Yánjiū* [Studies in Chinese Historical Grammar]. Beijing: The Commercial Press.
- Hopper, Paul and Thompson, Sandra. 1980. Transitivity in Grammar and Discourse. *Language*, Vol. 56, No.2, pp. 251-299.
- Hornstein Norbert, Nunes Jairo, Grohmann Kleanthes K. 2005. *Understanding Minimalism*. Cambridge: Cambridge University Press.
- Huang, C.T. James. 1982. Logical Relations in Chinese and the Theory of Grammar. PhD Dissertation, MIT.
- Huang, C.T. James. 1982 [2014]. Move Wh in a Language without Wh-movement. In Richard Kayne *et al. eds.* *An Annotated Syntax Reader: Lasting Insights and Questions*. Wiley-Blackwell.
- Huang, C.T. James. 1991. Modularity and Chinese A-not-A questions. In *Interdisciplinary Approaches to Language, Essays in Honor of S.-Y. Kuroda*, Carol Georgopoulos and Roberta Ishihara *eds.*, Dordrecht: Kluwer Academic Publishers, pp. 305-332.
- Huang, C.T. James. 2015. On syntactic analyticity and parametric theory. In Audrey Li, Andrew Simpson and Wei-Tien Dylan Tsai *eds.*, *Chinese Syntax in a Cross-linguistic Perspective*, Oxford: Oxford University Press, pp. 1-48.
- Huang, C.T. James, Li Audrey Y.-H. and Li Yafei. 2009. *The Syntax of Chinese*. Cambridge: Cambridge University Press.
- Huang, Yanxuan. 2025. Cóng Fǎnjiéwèn dào Zhōngxìngwèn—Shìlùn Cháozhōu Fāngyán “K-VP” Wènjù de Láiyuán [From Rhetorical Question to Neutral Question: The Origin of the “K-VP” Structure in the Chaozhou Dialect]. *Yǔyánxué Lùncóng* [Essays on Linguistics], Vol.4, pp. 120-127.
- Jacques, Guillaume. 2021. *A Grammar of Japhug*. Berlin: Language Science Press.
- Jiang, Lansheng. 1990. Yíwèn Fùcí Kě Tànyuán [On the Origin of the Interrogative Adverb Kě]. *Gúhànyǔ Yánjiū* [Studies in Archaic Chinese], Vol.2.
- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

- Koopman Hilda and Sportiche Dominique. 1991. The position of subjects. *Lingua*, Vol.85, pp. 211-258
- Kotek, Hadas. 2014. *Composing Questions*. PhD Dissertation. MIT.
- Larson, Richard. 1988. On the Double Object Construction. *Linguistic Inquiry*, Vol.19, pp. 335-392.
- Li, Haoze. 2024. Intervention effect in Mandarin Chinese. In *Oxford Research Encyclopedia of Linguistics* (<https://oxfordre.com/linguistics/view/10.1093/acrefore/9780199384655.001.0001/acrefore9780199384655-e-1079>)
- Lightfoot, David W. 1979. *Principles of Diachronic Syntax*. Cambridge: Cambridge University Press.
- Lin, Jo-Wang. 1992. The Syntax of *Zenmeyang* ‘how’ and *weishenme* ‘why’ in Mandarin Chinese. *Journal of East Asian Linguistics*, Vol.1, pp. 293-331.
- Lin, Jo-Wang. 2014. Wh-Expressions in Mandarin Chinese. In James Huang, Y.-H. Audrey Li, and Andrew Simpson eds., *The Handbook of Chinese Linguistics*. Wiley-Blackwell, pp. 180-207.
- Liu, Chuntao. 2018. Yáchéng Jūnhuà Yíwènjù “Ā+VP+Miù” de Bǐjiào Fēnxī [Analyses of Interrogative Sentence “A+VP+Miu” in Yacheng Junhua]. *Yǔyán Kēxué* [Linguistic Science], Vol.17, No.2, pp. 173-184.
- Lu, Wen. 2018. Aspects of the Grammar of Tunxi Hui: A Transitional Sinitic Language. PhD Dissertation, Hong Kong University.
- Lü, Shuxiang. 2017. *Jīndài Hànyǔ Zhǐdàicí* [Demonstrative Pronouns in Early Modern Chinese], Jiang Lansheng ed., Beijing: The Commercial Press.
- Madariaga, Nerea. 2017. Understanding Grammars through Diachronic Change. *Frontiers in Psychology*, Vol.8, pp. 1-16.
- Makino, Seiichi and Tsutsui, Michio. 1986 [2022]. *A Dictionary of Basic Japanese Grammar*. The Japan Times Publishing.
- Matisoff, James A. 1995. Sino-Tibetan Palatal Suffixes Revisited. In Nishi Yoshio, James A. Matisoff, Nagano Yasuhiko. eds. *New Horizons in Tibeto-Burman Morphosyntax*. Senri Ethnological Studies, No. 41. Osaka, pp. 35-91.
- Matisoff, James A. 2003. *Handbook of Proto-Tibeto-Burman: System and Philosophy of Sino-Tibetan Reconstruction*. Berkeley: University of California Press.
- McCready, Elin and Davis, Christopher. 2020. Sentence-final particles in Japanese. In Wesley M. Jacobsen and Yukinori Takubo eds., *Handbook of Japanese Semantics and Pragmatics*. De Gruyter Mouton, pp. 655-683.
- Mei, Guang. 2018. *Shànggǔ Hànyǔ Yǔfǎ Gāngyào* [An Outline of Classical Chinese Grammar]. Shanghai: Shanghai Education Press.
- Meisterernst, Barbara. 2010. Object Preposing in Classical and Pre-Medieval Chinese. *Journal of East Asian Linguistics*, Vol.19, No.1, pp. 75-102.
- Murphy, Andrew. 2017. Toward a unified theory of wh-in-situ and islands. *Journal of East Asian Linguistics*, Vol.26, No.2, pp. 189-231.
- Neidle Carol, Kegl Judy, Bahan Benjamin, Aarons Debra, and Dawn MacLaughlin. 1997. Rightward wh-movement in American Sign Language. In Dorothee Beerman, David Leblanc and Henk Van Riemsdijk eds., *Rightward Movement*. John Benjamins Publishing Company, pp. 247-278.

- Paul, Waltraud and Whitman, John. 2009. Applicative structure and Mandarin ditransitives. In Duguine, M., S. Huidobro and N. Madariaga eds., *Argument Structure and Syntactic Relations from a Cross-Linguistic Perspective*. Amsterdam-Philadelphia: John Benjamins, pp. 261-282.
- Pesetsky, David. 1987. Wh-in situ: Movement and unselective binding. In E. J. Reuland and A. G. B. ter Meulen eds., *The Representation of (In)definiteness*. Cambridge, MA: MIT Press, pp. 98-129.
- Pesetsky, David. 2000. *Phrasal movement and its kin*. Cambridge, MA: MIT Press.
- Pylkkänen, Liina. 2008. *Introducing Arguments*. Cambridge, MA: MIT Press.
- Reinhart, Tanya. 1998. Wh-in-situ in the Framework of the Minimalist Program. *Natural Language Semantics* Vol.6, No.1, pp. 29-56.
- Rizzi, Luigi. 1997. The Fine Structure of the Left Periphery. In Kayne Richard, Zanuttini Raffaella, and Leu Thomas eds., *An Annotated Syntax Reader*. Wiley-Blackwell, 2014, pp. 379-399.
- Roberts, Ian. 2007. *Diachronic Syntax*. Oxford: Oxford University Press.
- Sagart, Laurent et al. 2019. Dated language phylogenies shed light on the ancestry of Sino-Tibetan, *Proceedings of the National Academy of Sciences of the United States of America* 116 (21), pp. 10317-10322 (<https://doi.org/10.1073/pnas.1817972116>).
- Schuessler, Alex. 2007. *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.
- Schuessler, Alex. 2008. *Minimal Old Chinese and Later Han Chinese*. Honolulu: University of Hawai'i Press.
- Shi, Yuzhi. 2015. *Hànyǔ Yǔfǎ Yǎnhuàshǐ* [The Evolution of Chinese Grammar]. Nanchang: Jiangxi Education Press.
- Shi, Yuzhi. 2023. *The Evolution of Chinese Grammar*. Cambridge: Cambridge University Press.
- Sybesma, Rint. 2015. Complement (and Object), In Rint Sybesma et al. eds. *Encyclopedia of Chinese Language and Linguistics* (Online). Brill. (https://doi.org/10.1163/2210-7363_ecll_COM_00000097)
- Takano, Yuji. 2003. How antisymmetric is syntax. *Linguistic Inquiry*, Vol.34, No.3, pp. 516-526.
- Tan Rachel, and Shen Zheng. 2025. True wh-movement and wh-in-situ in one language: Evidence from Colloquial Singapore English. *Journal of Linguistics*, pp. 1-24.
- Tanaka, Hidekazu. 2001. Right-dislocation as scrambling. *Journal of Linguistics*, Vol. 37, No. 3, pp. 551-579.
- Tang, Sze-Wing. 2003. *Hànyǔ Fāngyán Yǔfǎ de Cānshù Lǐlùn* [A Parametric Theory of Chinese Dialectal Grammar]. Beijing: Peking University Press.
- Tang, Sze-Wing and Cheng, Siu-Pong. 2014. Aspects of Cantonese Grammar. In James Huang, Y.-H. Audrey Li, and Andrew Simpson eds., *The Handbook of Chinese Linguistics*. Wiley-Blackwell, pp. 601-628.
- Tsai, Wei-Tien Dylan. 1994. On nominal islands and LF extraction in Chinese. *Natural Language and Linguistic Theory*, Vol.12, No.1, pp. 121-175.
- Tsai, Wei-Tien Dylan. 1999. The Hows of *Why* and the Whys of *How*. *UCI Working papers in Linguistics*, pp. 155-184.
- Tsai, Wei-Tien Dylan. 2007. Chóngwēn Wèishénme Wèn Zěnmeyàng, Zěnmeyàng Wèn Wèishénme: Tán Hànyǔ Yíwèn jù hé Fǎnshēn jù zhōng de Nèi Wài Zhuàngyǔ [The hows of *why*

and the whys of *how* revisited: On inner and outer adverbials in Chinese interrogatives and reflexive sentences]. *Zhōngguó Yǔwén* [The Chinese Language], Vol.3, No.318, pp. 195-207.

Tsai, Wei-Tien Dylan. 2008. Left periphery and how-why alternations. *Journal of East Asian Linguistics*, Vol.17, No.2, pp. 83-115.

Tsujimura, Natsuko. 2014. *An Introduction to Japanese Linguistics* (Third Edition). Wiley-Blackwell.

Wang, Aiqing. 2021. The Intervention Effect in Late Archaic Chinese and Modern Mandarin. *Studies in Chinese Linguistics*, Vol.42, No.1, pp. 29-96

Wang, Li, 1962 [2017]. *Gudài Hànyǔ* (Dìyī Cè) [Classical Chinese (the First Volume)]. Beijing: Zhonghua Bookstore.

Wang, Li. 1980 [2015]. *Hànyǔ Shǐgǎo* [A Draft of the History of Chinese Language]. Beijing: Zhonghua Bookstore.

Yao, Zhenwu. 2015. *Shànggǔ Hànyǔ Yǔfǎshǐ* [A History of Old Chinese Grammar]. Shanghai: Shanghai Classics Publishing House.

Yu, Min. 1989. Hànzàng Tóngyuán Zipǔgǎo [A list of cognates between Chinese and Tibetan (Draft)]. *Mínzú Yǔwén* [Minority Languages], Vol.1.

Yue-Hashimoto, Anne O. 1991. Stratification in Comparative Dialectal Grammar: A Case in Southern Min. *Journal of Chinese Linguistics*, Vol.19, No.2, pp. 172-201.

Zhang, Min. 1990. *Hànyǔfāngyán Fǎnfǔwènjù de Lèixíngxué Yánjiū: Gòngshífēnbù jīqí Lìshíyùnhán* [A Typological Study of Yes-no Questions in Chinese Dialects: in Diachronic Perspective]. PhD Dissertation, Peking University.

Zhao, Changcai. 2000. Hànyǔ Shùbǔjiégòu de Lìshíyánjiū [A Diachronic Study of the Predicate-Complement Structure in Chinese]. PhD Dissertation, Graduate School of Chinese Academy of Social Sciences.

Zhao, Changcai. 2023. Zhōnggǔ Hànyǔ Fójīng de Shèwènjù [A Study of Rhetorical Questions in the Medieval Chinese Translation of Buddhist Sutras]. *Lìshǐ Yǔyánxué Yánjiū* [Studies in Historical Linguistics], Vol.20, pp. 35-50.

Zhao, Cheng. 1999 [2009]. *Jiǎgǔwén Jiǎnmíng Cídiǎn* [A Concise Dictionary of Oracle Bone Inscriptions]. Beijing: Zhonghua Bookstore.